QUANTWAY COLLEGE WITH COREQUISITE



Carnegie Math Pathways
WestEd

LibreTexts™

Quantway College with Corequisite

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide and accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

This text was compiled on 06/12/2025

TABLE OF CONTENTS

Licensing

1: College Module N

- 1.1: Introduction to Quantitative Reasoning
  - 1.1.1: Exercise N.1
- 1.2: Millions, Billions, or Trillions?
  - 1.2.1: Preparation N.2
  - 1.2.2: Growth Mindset Activity
  - 1.2.3: Exercise N.2
- 1.3: A Trillion Dollars of Student Loans
  - 1.3.1: Preparation N.3
  - 1.3.2: Exercise N.3
- 1.4: Computing and Interpreting Percentages
  - 1.4.1: Preparation N.4
  - 1.4.2: Exercise N.4
- 1.5: The National Debt…How Big Is It?
  - 1.5.1: Preparation N.5
  - 1.5.2: Exercise N.5
- 1.6: The Great Pacific Garbage Patch
  - 1.6.1: Preparation N.6
  - 1.6.2: Exercise N.6
- 1.7: An Index for Your Salary
  - 1.7.1: Preparation N.7
  - 1.7.2: Exercise N.7
- 1.8: Screening Tools–How Effective Are they?
  - 1.8.1: Preparation N.8
  - 1.8.2: Exercise N.8
- 1.9: A Closer Look at Screening Tools
  - 1.9.1: Preparation N.9
  - 1.9.2: Exercise N.9
- 1.10: Stress Reappraisal

2: College Module M

- 2.1: Mathematical Models–What's Best?
  - 2.1.1: Preparation M.1
  - 2.1.2: Exercise M.1
- 2.2: Linear Models
  - 2.2.1: Preparation M.2
  - 2.2.2: Exercise M.2
- 2.3: Data Is Trendy
  - 2.3.1: Preparation M.3
  - 2.3.2: Exercise M.3
- 2.4: Four Equations…Two Models

- 2.4.1: Preparation M.4
- 2.4.2: Exercise M.4
- 2.5: Modeling a New Type of Growth
  - 2.5.1: Preparation M.5
  - 2.5.2: Exercise M.5
- 2.6: How long Does That Take?
  - 2.6.1: Preparation M.6
  - 2.6.2: Exercise M.6
- 2.7: A Trust in Social Security
  - 2.7.1: Preparation M.7
  - 2.7.2: Exercise M.7
- 2.8: Multiple Variables Weighing on Your mind?
  - 2.8.1: Preparation M.8
  - 2.8.2: Exercise M.8

3: College Module S

- 3.1: Interpreting Charts and Graphs
  - 3.1.1: Preparation S.1
  - 3.1.2: Exercise S.1
- 3.2: Measuring Success
  - 3.2.1: Preparation S.2
  - 3.2.2: Exercise S.2
- 3.3: Statistical Analysis Process
  - 3.3.1: Preparation S.3
  - 3.3.2: Exercise S.3
- 3.4: Sampling and Error
  - 3.4.1: Preparation S.4
  - 3.4.2: Exercise S.4
- 3.5: Exploring Climate Data
  - 3.5.1: Preparation S.5
  - 3.5.2: Exercise S.5
- 3.6: Exploring Climate Variability
  - 3.6.1: Preparation S.6
  - 3.6.2: Exercise S.6
- 3.7: Modeling with Normal Distributions
  - 3.7.1: Preparation S.7
  - 3.7.2: Exercise S.7
- 3.8: Significance in Observational Studies
  - 3.8.1: Preparation S.8
  - 3.8.2: Exercise S.8
- 3.9: Significance in Experiments
  - 3.9.1: Preparation S.9
  - 3.9.2: Exercise S.9

4: Corequisite Module N

- 4.1: Corequisite- Estimation Vs Calculation
- 4.2: Corequisite- Conversion, Estimation, Percentages and Ratios, Scientific Notation
  - 4.2.1: Practice Problems Corequisite N.2
- 4.3: Corequisite- Interest, Doubling Time
  - 4.3.1: Practice Problems Corequisite N.3
- 4.4: Corequisite- Percentages and Ratios, Absolute and Relative Change, Rates
  - 4.4.1: Practice Problems Corequisite N.4
- 4.5: Corequisite- Percentages and Taxes, Ratios, Proportions and Dimensional Analysis
  - 4.5.1: Practice Problems Corequisite N.5
- 4.6: Corequisite- Dimensional Analysis, Unit Conversion
  - 4.6.1: Practice Problems Corequisite N.6
- 4.7: Corequisite- Buying Power, Consumer Price Index, Absolute and Relative Change, Reading and Interpreting Graphs
  - 4.7.1: Practice Problems Corequisite N.7
- 4.8: Corequisite- Percentages, Division, Extracting Relevant Information
  - 4.8.1: Practice Problems Corequisite N.8 and N.9

5: Corequisite Module M

- 5.1: Corequisite- Graphing Points, Linear Equations and Models
  - 5.1.1: Practice Problems Corequisite M.1
- 5.2: Corequisite- Solving Linear Equations, Slope
  - 5.2.1: Practice Problems Corequisite M.2
- 5.3: Corequisite- Trendlines
  - 5.3.1: Practice Problems Corequisite M.3
- 5.4: Corequisite- Linear Models Practice
  - 5.4.1: Practice Problems Corequisite M.4
- 5.5: Corequisite- Linear vs Exponential Models
  - 5.5.1: Practice Problems Corequisite M.5
- 5.6: Corequisite- Exponential Models (Practice and Application)
  - 5.6.1: Practice Problems Corequisite M.6
- 5.7: Corequisite- Multivariable Models
  - 5.7.1: Practice Problems Corequisite M.7

6: Corequisite Module S

- 6.1: Corequisite- Graphical Literacy
  - 6.1.1: Practice Problems Corequisite S.1
- 6.2: Corequisite- Graphical Literacy
  - 6.2.1: Practice Problems Corequisite S.2
- 6.3: Corequisite- Four-Step Process, Interpreting Statistical Studies, Understanding Experimental Design
  - 6.3.1: Practice Problems Corequisite S.3
- 6.4: Corequisite- Characteristics of Samples and Sample Data
  - 6.4.1: Practice Problems Corequisite S.4
- 6.5: Corequisite- Distribution, Shape, and Terminology

- 6.5.1: Practice Problems Corequisite S.5
- 6.6: Corequisite- Computing Measures of Center; Computing and Interpreting Z-scores
  - 6.6.1: Practice Problems Corequisite S.6

Index

Glossary

Detailed Licensing

Detailed Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

CHAPTER OVERVIEW

1: College Module N

1.1: Introduction to Quantitative Reasoning
1.1.1: Exercise N.1
1.2: Millions, Billions, or Trillions?
1.2.1: Preparation N.2
1.2.2: Growth Mindset Activity
1.2.3: Exercise N.2
1.3: A Trillion Dollars of Student Loans
1.3.1: Preparation N.3
1.3.2: Exercise N.3
1.4: Computing and Interpreting Percentages
1.4.1: Preparation N.4
1.4.2: Exercise N.4
1.5: The National Debt…How Big Is It?
1.5.1: Preparation N.5
1.5.2: Exercise N.5
1.6: The Great Pacific Garbage Patch
1.6.1: Preparation N.6
1.6.2: Exercise N.6
1.7: An Index for Your Salary
1.7.1: Preparation N.7
1.7.2: Exercise N.7
1.8: Screening Tools–How Effective Are they?
1.8.1: Preparation N.8
1.8.2: Exercise N.8
1.9: A Closer Look at Screening Tools
1.9.1: Preparation N.9
1.9.2: Exercise N.9
1.10: Stress Reappraisal

This page titled 1: College Module N is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.1: Introduction to Quantitative Reasoning

Quantway Cohort Contract

By signing below, I agree to fulfill the following requirements for participation in this Quantway course, and acknowledge that I understand the requirements for continued enrollment.

Commitments to the class:

- I commit to successfully completing this Quantway course with the members of my cohort.
- I commit to supporting all of my cohort members to understand mathematics.
- I commit to supporting all of my cohort members to complete the Quantway course.
- I will come to class every day prepared to participate in all classroom activities.
- I will contribute to creating a productive classroom atmosphere that supports everyone learning.
- I will keep an open mind and a positive attitude, and will be willing to try out new learning strategies and study skills.
- I will _____

_____

Commitments to my group:

- We commit to coming to class and working with our group everyday
- We commit to_____

_____

- We commit to_____

_____

Printed Name:_____Signature:_____Date:_____

Printed Name:_____Witness: _____Date: _____

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- you are a necessary and integral part of the classroom community.
- you should feel comfortable asking questions and participating in classroom discourse, no matter your language background.
- quantitative literacy empowers people by giving them tools to think for themselves, to ask intelligent questions of experts, and to respectfully question authority confidently.
- quantitative reasoning is important for decision-making in our democratic society.

By the end of this collaboration, you should be able to

- work effectively in groups.
- understand the course expectations and what is necessary to succeed in the course.

PROBLEM SITUATION: THE VACATION PROPERTY PUZZLE

A woman writes the following question to Ask Marilyn, a newspaper help columnist:

Nine members of our family will be renting a vacation property. The total rental fee is $3,600 for 10 days. People will be staying for a varying number of days. I say the first step in figuring how much each person will owe for his or her share is to divide the fee by nine [—the number of people that will be going on the vacation]. My husband says we should start by dividing the fee by 10—the number of days we have the rental. Who is right?[1]

(1) Discuss the problem situation with your group. Make sure everyone understands the problem. Investigate the two methods proposed by the husband and wife to determine the monetary implications.

(2) Explore different solutions to the Vacation Property Puzzle. As a group, pick a solution to share.

_____

[1] http://parade.com/473876/marilynvossavant/sharing-the-cost-of-a-vacation-rental/

This page titled 1.1: Introduction to Quantitative Reasoning is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.1.1: Exercise N.1

MAKING CONNECTIONS TO THE COLLABORATION

(1) What are your career goals? What do you see yourself doing in the next five to 10 years?

(2) Why did you enroll in this class? Are you required to take this class or did you choose it as an elective? Please explain.

(3) What grade do you realistically expect of yourself in this class? What do you have to do in this class to earn your expected grade or higher? Please explain.

(4) What was your most recent mathematics course (college or high school)? How well did you do? Could you have done better and if so, how could you have done better? Did you do better than you expected, and if so, why?

(5) The meme below was shared on various social media outlets. Are you persuaded by the quantitative information in this meme? Explain why or why not?



Exercises are also available online. The online versions of Exercises contain one or more **Learning** sections and a **Questions** section. The **Learning** sections in online Exercises match what is in this workbook and no score is given for these sections. The **Questions** section in online Exercises contains similar questions, but the numbers and figures vary from your workbook, so they are good for extra practice. Completing the **Questions** section will generate a score.
Tip: You can use the print version or the **Learning** sections online for practice. Then you can do the **Questions** section for additional practice and a score.

This page titled 1.1.1: Exercise N.1 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.2: Millions, Billions, or Trillions?

INTRODUCTION

Take a moment to discuss the following questions in your group:

- How big is a million? A trillion? A billion? Can you think of any good ways to describe these amounts?
- How can you compare them?
- Can you think of an example of using each in real life?

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- keeping track of the number of decimal places in large numbers is useful for rough estimations.
- there are huge differences between measures of a million, a billion, and a trillion.

By the end of this collaboration, you should be able to

- estimate the approximate size of the product of several large quantities.

PROBLEM SITUATION: HOW MUCH IS A TRILLION?

In the following questions, you will be presented with various situations. After reading the situation, consider your own estimation strategies before sharing your ideas in your group.

(1) Fill in the table below by describing your estimation strategy for each situation. Include your final estimated quantity. Write if your estimated quantity is in the millions, billions, trillions or some other size, such as tens of millions.

Quantity	**Describe your strategy and explain your estimate**
(a) In 2020, there were about 1.1 million working people in the U.S. who made at or below the federal minimum wage of $7.25 per hour. Estimate the amount of money needed per year to raise wages to $15 per hour. Hint: Assume a person works 40 hours per week for 52 weeks a year.	
(b) Estimate the amount of money needed to buy everyone in the United States a box of Girl Scout cookies. Hint: Assume a box of cookies cost $5 and the population of the U.S. is about 334 million.	
(c) Estimate the amount of money needed to pay for public college tuition and fees for four years. In 2021, there were approximately 14.63 million college students in the U.S. The average cost of college per year was about $25 thousand.	
(d) If the U.S. national debt was $30 trillion in March of 2022 and was evenly divided among people living in the United States, estimate how much each person's share would be. Hint: The population of the U.S. was about 334 million in 2022.	

(2) Now, use a calculator to refine your estimates.

Quantity	**Calculated Amount**

(a) In 2020, there were about 1.1 million working people in the U.S. who made at or below the federal minimum wage of $7.25 per hour. Calculate the amount of money needed per year to raise wages to $15 per hour. Hint: Assume a person works 40 hours per week for 52 weeks a year.	
(b) Calculate the amount of money needed to buy everyone in the United States a box of Girl Scout cookies. Hint: Assume a box of cookies cost $5 and the population of the U.S. is about 334 million.	
(c) Calculate the amount of money needed to pay for students to attend the University System of Ohio for four years. In 2020, the total students enrolled in this system was about 526 thousand, and tuition and fees were about $13,000 per year.	

(3) In Question 2 above, we tried to find more precise quantities. Why are the results still referred to as **estimates**?

(4) (a) If the U.S. budget was about $6.011 trillion in 2022, estimate the percentage of the budget which would be needed to pay for your estimate in Question 1a, rounded to $18 billion, for raising the minimum wage to $15 per hour. Round to the nearest tenth of a percent.

(b) If the U.S. budget is about $3.8 trillion, calculate the percentage of the budget which would be needed to pay or your calculation in Question 2a above, for raising the minimum wage. Round to the nearest hundredth of a percent.

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

This page titled 1.2: Millions, Billions, or Trillions? is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.2.1: Preparation N.2

Below is a table that describes some large numbers in words, standard decimal notation, and scientific notation. Notice how scientific notation keeps track of the number of decimal places to the right of the first digit but before the decimal point. Keeping track of the number of decimal places in quantities can be useful for some estimation problems. Recall that if a number does not have a decimal point displayed, then the decimal point is assumed to be at the right end of the number.

Quantity in Words	Decimal Notation	Scientific Notation	Number of places the decimal point was moved
One hundred twenty-two and one tenth	122.1	1.221×10^2	2
One thousand forty-five	1,045	1.045×10^3	3
Three million, four hundred seventy-five thousand, three hundred twenty and seventy-one hundredths	3,475,320.71	3.47532071×10^6	6
One billion, four hundred seventy-two million, three hundred forty-one thousand, nine hundred eighty-two and three tenths	1,472,341,982.3	1.4723419823×10^9	9
One Trillion	1,000,000,000,000	1×10^{12}	12

How to Round Numbers

To estimate with large numbers, we often round to a convenient place and then use that value in subsequent calculations. Rounding to a certain place is done in two steps:

- First, we need to identify the place value to round;
- Then, If the number to the immediate right of the place value you wish to round to is 5 or higher, you increase the digit you are rounding to by one and change all of the numbers to the right of it to zeroes. If the number to the immediate right of the place value you wish to round to is 4 or lower, you simply change to zeroes all of the numbers to the right of the place value to which you are rounding.

For example,

- 12,650 rounded to the nearest thousand is 13,000
- 12,499 rounded to the nearest thousand is 12,000
- 65,452,120 rounded to the nearest million is 65,000,000
- 12,811,120 rounded to the nearest million is 13,000,000

Tens and Hundreds of Thousands/Millions/Billions/Trillions

When writing numbers using words it is often helpful to look at numbers as products of powers of 10 listed above: thousands, millions, billions, trillions.

For example,

- 65,000 can be written as 65 * 1,000 or simply 65 thousand
- 95,000 can be written as 95 * 1,000 or simply 95 thousand
- 12,000 can be written as 12 * 1,000 or simply 12 thousand

The three numbers listed above are in tens of thousands.

Similarly,

- 850,000,000 can be written as 850 * 1,000,000 or simply 850 million

- 122,000,000 can be written as 122 * 1,000,000 or simply 122 million
- 650,000,000 can be written as 650 * 1,000,000 or simply 650 million

All three numbers listed above are in hundreds of millions.

For example, suppose we wanted a quick estimate of the yearly total expenses for a project in China that is expected to cost approximately $10 for every person in the country per day. By using estimates, we may be able to perform the calculations by hand reasonably quickly, as follows, to obtain a rough answer. We could estimate the population of China as 1.5 billion, use the ten dollars per person, and estimate the number of days in a year as 350. The product of these three numbers is:

$350 \times 10 \times 1.5 \times 10^9$

If we leave off the 10^9 we get:

$= 350 \times 15$

which is 10 times 350 plus another half of that: 3500 + 1750 = 5250, so now we have:

$= 5250 \times 10^9$

which is 5,250 billion or 5.25 trillion. So the cost would be about $5.25 trillion.

Recall that, when we say something like "5250×10^9 is 5,250 billion" we are using these facts:

Number	Power of Ten	Words
1,000	10^3	Thousand
1,000,000	10^6	Million
1,000,000,000	10^9	Billion
1,000,000,000,000	10^{12}	Trillion

Now, use the information above to match the given numbers to tens or hundreds of thousands/millions/billions/trillions.

• 15,000,000 • 250,000,000,000 • 76,000 • 879,000,000 • 876,000,000,000,000 • 58,000,000,000,000	• tens of thousands • tens of trillions • hundreds of millions • hundreds of trillions • hundreds of billions • tens of millions

(1) Use the above estimation strategy, together with any others you are familiar with, to decide if the following quantities are in the millions, billions, trillions, or some other size, such as tens of millions. **Note**: For some of these questions you could find the exact values, but for others, such as populations, any number would be an estimate.

(a) The number of seconds you have been alive if you are 23 years old.

(b) The amount of money needed to send all adults in the United States to a public college for four years. Assume there are about 258 million adults in the United States and that the cost of tuition and fees to a public college is about $25,000.

(c) The number of words in the novel War and Peace.

(2) Refine your above estimates by using a calculator and obtaining any missing information you need.

(a) The number of seconds you have been alive.
For example, for a 10-year-old: 60 * 60 * 24 * 365 * 10 = 315,360,000 or about 315 million seconds.

(b) The amount of money needed to send all adults in the United States to a public college for four years. Assume there are about 258 million adults in the United States and that the cost of tuition and fees to a public college is about $25,000.

(c) The number of words in the novel *War and Peace*. Using an internet search, we found that sample pages of an e-book version of *War and Peace* have about **300** words on each page. This e-book version is **1307** pages long.

Preparations are also available online. The online versions of Preparations contain one or more **Learning** sections and a **Questions** section. The **Learning** sections in online Preparations match what is in this workbook and no score is given for these sections. The **Questions** section in online Preparations contains similar questions, but the numbers and figures vary from your workbook, so they are good for extra practice. Completing the **Questions** section will generate a score.
Tip: You can use the print version or the **Learning** sections online for practice. Then you can do the **Questions** section for additional practice and a score.

After Preparation N.2 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration N.2, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
convert between millions, billions, and trillions.	
compute basic percentages and ratios.	
express one number as a percent of another.	

This page titled 1.2.1: Preparation N.2 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

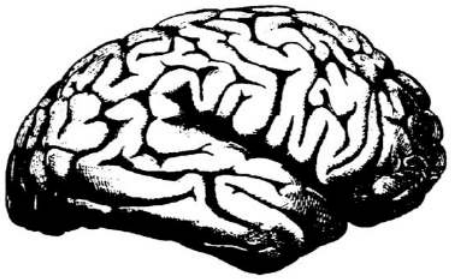
1.2.2: Growth Mindset Activity

YOU CAN GROW YOUR BRAIN

New Research Shows the Brain Can Be Developed Like a Muscle

Many people think of the brain as a mystery. We don't often think about what intelligence is or how it works. And when you do think about what intelligence is, you might think that a person is born either smart, average, or dumb—either a "math person" or not—and stays that way for life.

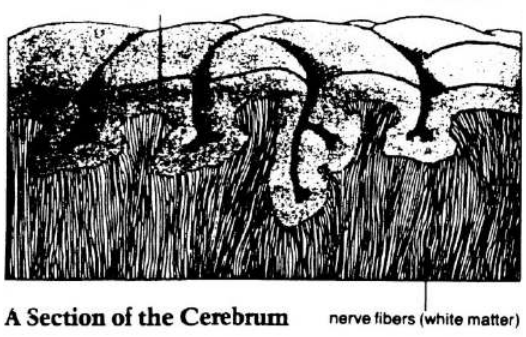
But new research shows that the brain is more like a muscle—it changes and gets stronger when you use it. Scientists have been able to show just how the brain grows and gets stronger when you learn.

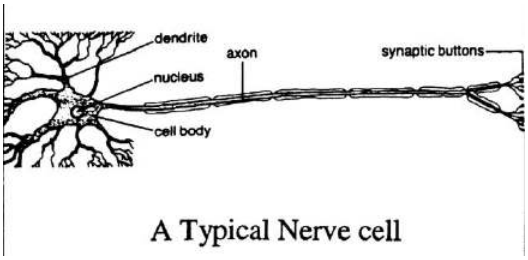


Everyone knows that when you lift weights, your muscles get bigger and you get stronger. A person who can't lift 20 pounds when they start exercising can get strong enough to lift 100 pounds after working out for a long time. That's because muscles become larger and stronger with exercise. And when you stop exercising, the muscles shrink and you get weaker. That's why people say "Use it or lose it!"

But most people don't know that when they practice and learn new things, parts of their brain change and get larger, a lot like the muscles do. This is true even for adults or older teenagers. So it's not true that some people are stuck being "not smart" or "not math people." You can improve your abilities a lot if you practice and use good strategies.

Inside the outside layer of the brain—called the cortex—are billions of tiny nerve cells, called neurons. The nerve cells have branches connecting them to other cells in a complicated network. Communication between these brain cells is what allows us to think and solve problems.





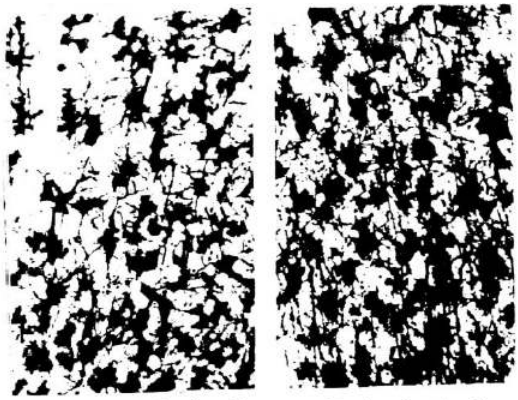
A Typical Nerve cell

When you learn new things, you actually grow more of these tiny connections and they get stronger. The more you challenge your mind to learn, the more your brain cells grow. Then, things that you once found very hard or even impossible to do—like understanding history, doing algebra, or completing complicated paperwork—become easier. The result is a stronger, smarter brain.

How Do We Know That The Brain Can Grow Stronger?

Scientists started thinking the human brain could develop and change when they studied adult animals' brains. They found that animals who lived in a challenging environment, with other animals and toys to play with, were different from animals who lived alone in bare cages.

The animals who lived alone just ate and slept all the time, but the ones who lived with different toys and other animals were always active. They spent a lot of time figuring out how to use the toys and how to get along with other animals.



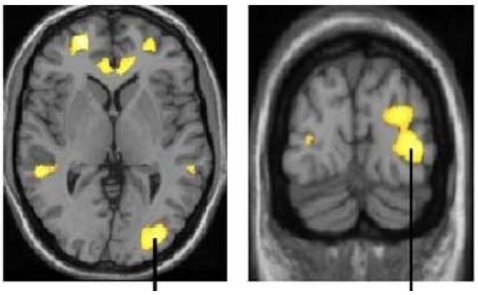
Nerves in brain of animal living in bare cage.

Brain of animal living with other animals and toys.

These animals had more connections between the nerve cells in their brains. The connections were bigger and stronger, too. In fact, their whole brains were about 10% heavier than the brains of the animals who lived alone without toys.

The adult animals who were exercising their brains by playing with toys and each other were also "smarter" –they were better at solving problems and learning new things.

Can Adults Grow Their Brains?

Scientists have recently shown that adults can grow the parts of their brains that control their abilities—like the ability to do math or even to juggle. 

In Yellow: Parts of the brain that grew when adults learned to juggle

doi:10.1371/journal.pone.0002669.g001

In one study, scientists found a group of adults who were not jugglers. They taught half how to practice juggling in the right way. These people practiced for a long time and got much better at juggling. The other half didn't practice, and didn't get better.

Next, the scientists used a brain scanner to compare the brains of the two groups of people. They found that the people who learned how to juggle actually grew the parts of their brains that control juggling skills. Their brains had changed, so they actually had more ability.

This was surprising because these people said before the study that they couldn't juggle—just like some people say they're "not good at math" or "not smart at school." But when they kept trying and learned good strategies for practicing, they actually learned and grew their brains.

Two Parts of Your Brain That Need to Grow: The "Knowing" Part and the "Know How" Part

The brain doesn't just learn things; it also learns how to do things. Scientists have shown that people can grow two types of intelligence: the "knowing" type and the "know how" type. The "knowing" type of intelligence is what most people think about when they are learning new things in science or English or history. But you also need to learn skills that let you use your brain in a smarter way.

How do scientists know that you can grow the "know how" part of your brain? One example comes from the brains of taxi drivers in London, England. The streets of London are like a terrible maze. They are very hard to get around in. Brain scientists found that the longer the taxi drivers spent practicing driving in the streets of London, the more their brains grew. The "hippocampus" area of the brain—which controls memory for how to do everyday things—was actually larger! This means the more they practiced driving the confusing streets of London, the more the "know how" part of their brain grew. Although you might not realize it, we all grow the "know how" part of our brains whenever we learn to do new, hard things.

People often forget that we need to grow the "know how" part of our brains. If you study a long time but use a bad strategy, you may not learn—even if you try very hard. A few students study by doing the same set of easy problems and skipping the hard ones. Or they just re-read the textbook but don't try to challenge the ideas. After all, it's easier that way. But when it comes time to do the test, they don't do well because they didn't do things that grew their brains and taught them new things. When this happens, they may even say "I must not be smart at this."

But brain scientists have shown that a stronger brain doesn't just happen to you because you put in the time. You have to practice in the right way. If a weightlifter watched other people exercise all day long, he wouldn't get any stronger. You actually have to practice the right way—and usually that means the hard way—to get better at something. In fact, scientists have found that the brain grows more when you use new and different strategies. The brain grows less when you stick to things you already know.

How do you grow the "know how" part of your brain? You do it by seeking out help from others, like teachers or older students who do well. Or by going to the academic support center or getting tutoring. And then you practice successful strategies. At first it can be hard—sometimes people feel dumb for not knowing simple things like how to take notes, how to study, or even how to sign up for their classes. Or they feel not smart if they go to a tutor. But it's actually the opposite. Practicing better strategies makes your brain smarter.

The Truth About “Smart” and “Dumb”

People aren’t “smart” or “dumb” in school. At first, no one can read or solve equations. But with practice, they can learn to do it. And the more a person learns, the easier it gets to learn new things. This is because the “knowing” and “know how” brain muscles have gotten stronger.

This is true for everyone, even adults or college students. Dr. Wittenberg, the scientist from Wake Forest University who did the research on juggling, said “We used to think adults can’t form new brain connections, but now we know that isn’t true… The adult brain is like a muscle, and we need to exercise it.”

People who don’t know this can miss out on the chance to grow a stronger brain. It does take work to learn, just like lifting more weights or becoming a better juggler does. Sometimes it even hurts! But when you feel yourself get better and stronger, you realize that all the work is worth it.

References:

Blackwell, L. A., Trzesniewski, K. H., & Dweck, C. S. (2007). Theories of intelligence and achievement across the junior high school transition: A longitudinal study and an intervention. Child Development, 78, 246-263.

Driemeyer, J., Boyke, J., Gaser, C., Buchel, C., May, A. (2008). Changes in Gray Matter Induced by Learning—Revisited. PLoS One, 3, e2669. doi:10.1371/journal.pone.0002669.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. Proceedings of the National Academy of Sciences, 97(8), 4398-4403.

Nordqvist, C. (2004, Feb 1). “Juggling makes your brain bigger – New Study.” Retrieved from http://www.medicalnewstoday.com/releases/5615.php

Questions: You Can Grow Your Brain

Questions

(1) Think about the article that you just read. What are all the reasons why scientists say that people's math ability can grow and get better with effort and practice?

Please summarize them briefly below.

(2) In the article, you learned three things:

- When you work hard and learn new things, your brain grows new connections and you get smarter.
- The more you challenge yourself, the smarter you will become.
- Smart kids are the kids who have practiced more—they have built up their reading and math "muscles."

Think about an example from your own life. What is something you weren’t good at first but then practiced using a good strategy and became really good at it? Write about it and explain how you became good at it.

(3) Not all math students know that the brain can get smarter, even though it may help them have success. And we want to get your help so that we can learn more about how to explain it to them. We're hoping you can explain—in your own words—that the brain gets smarter when people use good strategies and try hard.

Imagine a friend who is struggling in school. This friend used to do pretty well in school but now is having a hard time and is starting to feel dumb. Write a letter to your friend to encourage him or her—tell them about what you just learned about the brain and why they shouldn’t be discouraged. For example, you can tell them:

“Dear friend,

I realize that you feel discouraged in your math class because it is becoming a bit challenging for you, but after learning that our brain grows after practicing something hard I strongly encourage you to continue. I have learned that changing our studying strategies, asking for help, using resources, and practicing will help us improve what we do not understand. Life is a challenge and if everything was easy then the whole world would be boring. We have to grow our "know how" parts of our brain which requires a lot of effort. Our brain can become stronger and smarter by practicing what we do not know. So don't give up on your math class; challenge yourself.”

Here are some other general ideas you can share with them:

1. How they can get smarter if they work hard and use a good strategy.
2. How they should work hard to build their reading and math muscles.
3. How they are not dumb, they just need to practice using a good strategy.
4. How they can ask the teacher or other students to help them learn better ways to study.
5. Any other tips you have for learning in school and getting smarter.

(Don't worry about writing a perfect final draft. We just want to know how you would say this to another student in your own words.)

This page titled 1.2.2: Growth Mindset Activity is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.2.3: Exercise N.2

(1) A website recommends that you should have 3 gallons of water for every inch of goldfish length in your tank.[2]

(a) If you have one billion goldfish that are each 3 inches long, how many gallons of water would you need?

(b) With its seven salt and freshwater aquariums, totaling 12.87 million gallons, the Chimelong Ocean Kingdom theme park is considered the world's largest aquarium. How many of these record-breaking aquariums would it take to house the same one billion goldfish that are each 3 inches long. Round your answer to the nearest aquarium.

(2) The Jurassic period began approximately 200 million years ago. **Approximately** how many minutes have passed since the beginning of this period to the present day?

(3) Janetta wishes to buy a commercial building to expand her business and needs to save for the down payment.

(a) If she can save $4,000 a month, **estimate** how long it will take to save $2,000,000 if she does not earn interest on her savings. Assume (unrealistically) that the price of the building won't change while she saves.

(b) **Calculate** how long it will take to save $2,000,000 if she can save $4,000 a month.

(c) **Estimate** how much she would need to save each month to realize her goal in 15 years.

(d) Now **calculate** how much she would need to save each month to realize her goal in 15 years. Round to the nearest cent.

(4) The main hall of Grand Central Station in New York City is approximately 210,000 square feet.

(a) **Estimate** how many people could stand with enough room to turn around in the main hall of the station if there were an emergency which required that the station be used to shelter people.

(b) If there are 4 million people in Manhattan, NY, on a typical workday, **estimate** the percentage of them that could be sheltered in the station.

---

[2] http://www.allabout-aquariumfish.com/2009/02/tank-size-for-goldfish.html

This page titled 1.2.3: Exercise N.2 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.3: A Trillion Dollars of Student Loans

INTRODUCTION

Taking out a loan is usually a necessary step to prepare for a large purchase, such as a car, a house, or a college education. While there are online tools and specific equations used to calculate the amount of a monthly payment for these loans, you can usually use a much simpler set of steps to find an accurate estimate for a monthly payment. Other variables that contribute to the affordability of a loan may be harder (or impossible) to calculate. These might be things like estimating your monthly paycheck or the average starting salary for a person with a college degree in your field.

- What variables contribute to determining the amount of a monthly payment on a loan?
- What variables outside of the loan contribute to determining whether or not one can afford a loan?

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- estimation is useful in many situations and can be done with simple as well as more complex calculations.
- the interest rate of a loan and the length of a loan repayment have a significant impact on the size of each loan payment.
- the interest rate and the length of a loan repayment impact the total amount repaid.

By the end of this collaboration, you should be able to

- compute estimations related to personal budget/finance issues.
- estimate the total interest charged on a loan.
- estimate monthly payments for a variety of loans.
- determine if an estimate is an underestimate or overestimate.

PROBLEM SITUATION 1: ESTIMATING COSTS

Melinda is purchasing a car and needs to borrow $12,000 to make the purchase. She needs to decide between two loans. Both loans have an Annual Percentage Rate (APR) of 5%, but one has a term of 3 years while the other has a term of 6 years.

(1) Compare the two loans by estimating the monthly payment for each loan. To figure this out, work through the following questions in your group.

(a) **Estimate** how much the monthly payments would be if there were no interest?

3-year loan =

6-year loan =

(b) How much would the payments be if the interest was charged on the entire original principal for the entire term of the loan? Round to the nearest dollar.

3-year loan =

6-year loan =

(c) Use your answers to Questions 1(a) and 1(b) to create a **better estimate** of the two payment amounts.

3-year loan =

6-year loan =

(d) Use your estimates in Question 1(c) to **estimate** the total amount Melinda would pay for each loan.

3-year loan =

6-year loan =

(e) Use your estimates in Questions 1(c) and 1(d) to **estimate** the amount of interest Melinda would pay for each loan.

3-year loan =

6-year loan =

(f) Write a few sentences explaining to Melinda the advantages and disadvantages of each loan. One thing you might consider is the value of the car over time.

PROBLEM SITUATION 2: PAYING FOR COLLEGE

Sandra is eligible for a subsidized Stafford loan to help her pay for college. Subsidized Stafford loans are loans available to low- and medium-income families to help pay for college. While a student is in college, the U.S. government pays the interest on the loan. So, if a student borrows $5,000, then he/she still owes only $5,000 when he/she graduates from college (or leaves early).

Note that loan repayment for a Stafford loan begins six months after graduation (or after the student has left college). That is the point when interest begins to accrue (be charged). Once repayment begins, payments are typically made monthly.

(2) Sandra is not sure whether she should take the subsidized loan. She is concerned about being able to pay it back (her other option is to work part-time and attend school part-time). Suppose her financial aid award letter says she is qualified to borrow up to $4,000 per year.

How much (in total) will Sandra owe if she borrows $4,000 per year and it takes her five years to graduate from college? **Note**: This total is called her "loan balance."

(3) Typically, individuals are given a 10-year repayment period. Payments are made monthly.

(a) How many payments (total) will Sandra make during the repayment period?

(b) Pretend that no interest is charged on Sandra's loan. If she only has to repay the principal, she takes 10 years to repay, and she repays the same amount each month, what would be the amount of each monthly payment? Round to the nearest cent.

(c) Any estimate will not be exact, but it is important to have an understanding as to why the "actual" answer is more or less than an estimate and why. Is the estimate from Question 3(b) more or less than the "actual" answer?

(4) (a) If Sandra is charged an interest rate of 6.8% per year and she takes the full 10 years to repay, how much would her monthly payments be to pay off the loan? To figure this out, use an estimation strategy. You don't need exact calculations or formulas.

(b) How much total interest will Sandra pay over the 10 years?

(c) Do you think your group's estimate in Question 4(b) is an overestimate or an underestimate of the actual interest Sandra will pay? Why?

(5) Use the Average Balance Method estimate from Question 4(a). Assume that Sandra has the following minimum living expenses:

- Rent and utilities cost $1000/month
- Food and eating out cost $300/month
- Car payment and insurance cost $300/month
- Entertainment, gas, and cell phone cost $200/month
- Clothing and miscellaneous household expenses cost $100/month
- Health Insurance costs $100/month

You should also expect that about 30% of Sandra's "gross income" (her actual salary) will be deducted from her paycheck to cover various taxes, Medicare, and Social Security. Estimate how much Sandra would have to earn annually to be able to afford the loan payment.

(6) Do you think this is a reasonable salary for Sandra to expect to earn if she has a bachelor's degree? Why or why not?

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

This page titled 1.3: A Trillion Dollars of Student Loans is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.3.1: Preparation N.3

Simple Interest can be defined as interest paid or computed on only the original principal of a loan or on the amount of an account. You may have seen the formula for Simple Interest before:

$$I = Prt$$

where I is the interest, P is the principal (or amount) of the loan, r is the annual interest rate in decimal form, and t is the number of years.

(1) If the interest rate on a loan is **doubled**, what will be the effect on the total amount of interest paid on the loan? Assume simple interest is used. Choose all that apply.

(i) The total amount of interest paid on the loan will be twice as much.

(ii) The total amount of interest paid on the loan will remain the same.

(iii) The total amount of interest paid on the loan will go up by the same percentage as the percentage point increase in the interest rate.

(iv) The total amount of interest paid on the loan will be four times as much as the original amount.

(v) The total amount of interest paid on the loan will double.

PayPal is a company that operates an online payments system that is used by many vendors to receive payments from customers. Once users link their credit cards or bank accounts to PayPal, they can pay for things online without having to re-enter their sensitive card or bank details.

PayPal has a service called “PayPal Credit.” Some vendors provide a special offer of “No Payments + No Interest” if the original charges are paid in full within six months (sometimes the offer is only available for a minimum purchase of $250).

Marco is looking to purchase a new laptop computer which costs $1,200. He is considering a payment plan that would give him six months to pay for the computer, with no interest charged. For the purposes of Questions 2–4, we will assume there are no taxes or other fees that would increase the total cost of the laptop.

(2) The federal government sets a minimum wage that is currently $7.25 an hour. Most states have a minimum wage that is equal to or higher than $7.25 an hour. Wisconsin is a state with a minimum wage that is the same as the federal minimum wage. If Marco is working at a minimum wage job full time (40 hours per week) in Wisconsin, how much is his weekly salary?

(3) Using the weekly salary you calculated above, how much would Marco earn per month (assume that there are four weeks in every month)?

(4) What percentage of his monthly salary should Marco set aside to pay for the laptop within six months? As mentioned, the new laptop computer costs $1,200. Marco is considering a payment plan that would give him six months to pay for the computer, with no interest charged. Assume there are no taxes or other fees that would increase the total cost of the laptop. Round to the nearest hundredth of a percent (two decimal places). **Hint**: Use Marco’s income as calculated in Question 3.

(5) In 2023, the minimum wage in Michigan was raised to $10.10 an hour. For parts (a)-(c), assume a working week of 40 hours.

(a) How much is his weekly salary?

(b) How much would he earn per month? Assume there are four weeks in every month.

(c) What percentage of his monthly salary should Marco set aside to pay for the laptop within six months? Recall that the laptop costs $1200. Round to two decimal places.

(6) When you make payments on a long-term loan with interest, a portion of each monthly payment goes to pay the current interest due on the loan, and the remaining amount of the monthly payment is used to pay down the amount still owed on the loan (i.e., the remaining principal). Ding-Huan is purchasing a car and is paying an Annual Percentage Rate (APR) of 6%. After making a down payment, the amount Ding-Huan financed was $15,000 and his monthly payment for the 5-year term (length) of the loan is $290.00. You can think of the calculations for the portion going to interest and principal each month as simple interest problems, calculated month by month. Recall that I = Prt for simple interest.

(a) Calculate the total amount that Ding-Huan will pay over the 5-year term. How much will be principal? How much will be interest?

Total =

Principal =

Interest =

(b) How much will the monthly interest rate be, given as a percent?

(c) How much of the first monthly payment goes toward paying off the interest and how much goes toward paying off the principal? Round to the nearest cent.

(d) How much will Ding-Huan still owe on the principal after he makes the first payment (at the end of the first month of the loan)? Recall that he borrowed $15,000.

(e) How much will he still owe on the principal after he makes the second payment at the end of the second month of the loan? Round to the nearest cent.

(7) In the next collaboration, we will be discussing college student loan debt. If about 44 million Americans had student loan debt in 2021, with an average debt of around $37,000, estimate how much total student loan debt was outstanding at that time.

After Preparation N.3 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration N.3, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
convert percentages to decimals.	
decide when simple interest is appropriate and how to calculate it.	
use and interpret the terms “millions,” “billions,” and “trillions.”	

This page titled 1.3.1: Preparation N.3 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.3.2: Exercise N.3

(1) It was estimated that, in the United States, the total student loan debt surpassed total credit card debt and auto loans for the first time in 2011, with total student loan debt reaching the $1 trillion mark.

(a) Write this total loan amount in standard form (written out with all the zeros).

(b) If the average interest rate on these loans is 6%, estimate how much interest is being paid on the $1 trillion dollars of student loan debt each year. Write your answer in standard form.

(2) By federal law, the interest rate on subsidized Stafford loans changes each year, based on certain financial market rates. The interest used on a Stafford loan is determined by the first disbursement date of the loan. For Stafford loans taken out for the 2022-2023 school year, the interest rate is 4.99%.[3] Think about the situation of Sandra and her student loan ($20,000) from the collaboration.

(a) At the 4.99% rate, will Sandra pay less, pay more, or pay the same each month than she would with an interest rate of 6.8%?

(b) Estimate again how much her monthly payments would be if she is charged an interest rate of 4.99% and she takes the full 10 years to repay the loan. **Note**: It is assumed that you will use the "average balance method" which was discussed during the collaboration. Round to the nearest dollar.

(3) (a) You want to purchase a new car. You are offered a loan for $16,000 with an annual interest rate of 6% and a 36-month repayment period. Estimate how much your monthly payment would be for this loan. Use the "average balance method" to estimate the total amount you will pay over the 36 months for this $16,000 car loan. Round to the nearest dollar.

(b) Calculate an estimate how much your monthly payment would be for this loan. Round your answer to the nearest dollar.

(4) Suppose you decide before buying the car that you can't afford the monthly payment you calculated in Question 3(b), so you take out the same $16,000 loan with an interest rate of 6%, but now with a 60-month repayment period.

(a) Calculate an estimate of the total amount you would pay over the 60 months for the car loan in this case.

(b) Calculate an estimate for your monthly payment with the 60-month option.

(c) Calculate an estimate how much more you will pay for the car if you choose the 60-month loan instead of the 36-month loan.

(5) It is estimated that the average principal owed for student loans in 2021 was $37,338 per student.[4] If market rates go up and the interest rate for student loans increases from 4.99% to 6.8%, estimate how much more interest students will pay over a 10-year repayment period for this average amount owed, at the 6.8% rate as compared with a rate of 4.99%. Round all figures to the nearest dollar.

---

[3] https://www.edvisors.com/student-loans/federal-student-loans/interest-rates/

[4] https://educationdata.org/average-student-loan-debt

This page titled 1.3.2: Exercise N.3 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.4: Computing and Interpreting Percentages

INTRODUCTION

In this section, you will learn about adding a tip to a meal bill and determining sale prices. Most students find this process straightforward when they need to calculate the total sale price given the original price and the rate to add. However, many students find it difficult to calculate the original price when they are given the total sale price and the rate added.

The formula you will use in this section is:

Total price = original price + original price * tax rate,

Or, alternatively: Total price = original price * (1 + tax rate)

You should use the *distributive property* to convince yourself that the two formulae are the same. The second formula will be much more useful when you are trying to determine the original price of an item, given the total price.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- a percentage could represent a ratio which is made up of both a numerator and denominator.
- a percentage is used to measure relative change in a quantity.
- a percentage could represent a chance or probability.

By the end of this collaboration, you should be able to

- compute and understand the difference between the absolute and relative change between two quantities (including rates).
- correctly compare quantities using both absolute and relative change (including rates).
- represent chances or probabilities by both percentages and ratios.

PROBLEM SITUATION 1: COMPUTING WITH PERCENTAGES

(1) The state of Illinois has a 6.25% sales tax.[5]

(a) Calculate the amount of tax on a $120 purchase. Round to the nearest cent.

(b) What is the total sale price of a $120 purchase, including tax? Round to the nearest cent.

(2) Illinois has a 6.25% sales tax, and textbooks are taxable in Illinois.[6] The amount that a student paid at her school bookstore was $185. What was the price of the books and school supplies before tax? Round to the nearest cent.

(3) In 2022, Jacob worked as a fast food manager and:

- had a taxable income of $30,000
- was required to pay $3,398 in federal taxes.[7]

What is the federal tax rate as a percentage? Round to the nearest percent.

(4) According to Wikipedia, the information in the table below represents a breakdown of how the U.S. federal government spent money during the 2022 fiscal year:[8] Compute the percentage of the total amount of money that the U.S. federal government spent in each category and complete the table above. Round to the nearest percentage point.

Category	Billions of Dollars Spent in Fiscal Year 2022	Percentage
Defense Department	751	(i)
Social Security	1,200	(ii)
Healthcare	1,339	(iii)
Net Interest	475	(iv)
Other Mandatory	1,583	(v)
Other Discretionary	910	(vi)

PROBLEM SITUATION 2: MEASURING CHANGE WITH PERCENTAGES

There are several different ways that one can measure the change between an old value and a new value. These values could be, for example, the weight of the same person at two different points in time, or the age of the previous president and the age of the current president.

The absolute change tells us by how much the two values differ. The relative change is the absolute change compared to the reference value (earlier quantity). Relative change is often given as a percentage.

You may recall the following two formulas, where "Old" is the starting value, and "New" is the ending value:

$$Absolute\ Change = New - Old$$

$$Relative\ Change = \frac{New - Old}{Old} * 100\%$$

When working on Questions 5 and 6, think carefully about what change you are trying to measure. To measure a relative change, one must understand which value constitutes the "starting, reference value or old value" and which constitutes the "ending, final or new value." The context and the surrounding language in a question help to clarify which is which.

(5) An article about the 2016 presidential election stated: "Trump lost the popular vote by about 2.1 percentage points. Meanwhile, Trump racked up 306 electoral voters—14 percent more than Clinton."[9] Answer the following questions about this statement.

1. Does "2.1 percentage points" refer to an absolute or a relative change?
2. Does "14 percent more" refer to an absolute or a relative change?

(6) The average annual snowfall for Boise, ID, is 19.8 inches. Moline, IL, receives an average of 33.3 inches. Fill in the blanks:

(a) Moline receives, on average, _____ more inches of snow than Boise.

(b) Moline receives, on average, _____ % more snow than Boise. Round to the nearest tenth of a percent.

(c) Boise receives, on average, _____ % less snow than Moline. Round to the nearest tenth of a percent.

(d) In Question 6(a), were you computing absolute or relative change?

(e) In Questions 6(b) and 6(c), were you computing absolute or relative change?

PROBLEM SITUATION 3: PERCENTAGES AS RATES

(7) According to the CDC, 11.5% of Americans smoked tobacco in 2021.[10]

(a) What is meant by this statement? What was the numerator? What was the denominator? What assumptions could you make about the population of Americans included in this statement?

(b) Discuss in your group how this percentage (11.5%) was probably computed. Record your group's thinking below.

(8) The Insurance Institute for Highway Safety (IIHS) is an independent, nonprofit scientific and educational organization dedicated to reducing deaths, injuries and property damage from motor vehicle crashes. In 2021, the IIHS published statistics showing, in part, that almost three out of four motor vehicle crash deaths were males.[11]

(a) Use the figures from the last sentence ("almost three out of four motor vehicle deaths"). What percentage of motor vehicle crash deaths in the U.S. were males? Round to the nearest percentage point.

(b) What percentage of motor vehicle crash deaths in the U.S. were female? Round to the nearest percentage point.

(c) Do your answers to Questions 8(a) and 8(b) support the statement, "In general, males are more likely to be killed in traffic accidents than females"? Explain your reasoning. Support your answer with some possible examples.

PROBLEM SITUATION 4: PERCENTAGES, PROPORTIONS, AND RISK

(9) A flyer that went to households in the mail, from a company called Life Line Screening,[12] included the following claim:

- "Your risk of suffering a stroke is high—it affects 1 in 5 women, and 1 in 6 men."

Answer the questions below by converting the risk of suffering a stroke from a proportion to a percentage.

(a) What is the likelihood of a woman suffering a stroke? Calculate the percentage.

(b) What is the likelihood of a man suffering a stroke? Calculate the percentage. Round to the nearest percentage point.

(c) There are many different ways to compare the differences in these two risks. Find two or three ways, and write a complete sentence for the ways you chose.

(10) Some research suggests that the risk of developing type 2 diabetes by age 70 is 11%. How could you express this with a meaningful rate or ratio?

PROBLEM SITUATION 5: MEASURING CHANGES IN PERCENTAGES

(11) The unemployment rate is the percentage of people in the labor force who are unemployed and actively seeking work. Going into the Covid pandemic in January 2020, the unemployment rate in the United States was 3.5%. In May of that year, unemployment was 13.2%.[13] There are many ways that this change in unemployment rate could be described. Fill in the blanks in the following statements so they correctly describe this change.

(a) From January 2020 to May 2020 the unemployment rate increased by _____ percentage points. (Hint: You are computing an absolute change here.)

(b) From January 2020 to May 2020 the unemployment rate increased _____ percent. (Hint: You are computing a relative change here.) Round to the nearest percent.

(c) From January 2020 to May 2020 the unemployment rate _____ (fill in with “nearly” or “more than”) _____ (fill in with “tripled” or “quadrupled”).

(12) Question 7 provided information from the CDC about smoking habits in the U.S. The table below presents data from other countries.

Country	**Percentage of Smokers**
Greece	39.1
Germany	28.0
France	34.6
Sweden	8.0
Canada	17.5

(a) Write a statement comparing the smoking rates of people in Greece and Sweden, using a measure of absolute change.

(b) Write a statement comparing the smoking rates of Canada and Greece, using a measure of relative change.

(c) Describe what you learned by writing a statement comparing the smoking rates of Canada and Greece. Use phrases similar to “almost double,” “more than triple,” and “almost a fourth of.”

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

---

[5] http://www.salestaxinstitute.com/resources/rates

[6] http://blog.taxjar.com/textbooks-tax-exempt-sales-tax/

[7] https://www.irs.gov/pub/irs-pdf/i1040tt.pdf

[8] https://en.Wikipedia.org/wiki/United_States_federal_budget#Major_expenditure_categories

[9] http://www.vox.com/policy-and-politics/2016/11/9/13572112/trump-popular-vote-loss

[10] https://www.cdc.gov/tobacco/data_statistics/fact_sheets/fast_facts/index.htm

[11] https://www.iihs.org/topics/fatality-statistics/detail/males-and-females

[12] Life Line Screening, unsolicited mail, received April 14, 201

[13] https://data.bls.gov/timeseries/LNS1400000

This page titled 1.4: Computing and Interpreting Percentages is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.4.1: Preparation N.4

(1) The U.S. Census Bureau conducts a census of the U.S. population every ten years. The 2020 census found that approximately 330,000,000 people lived in the U.S. The Census Bureau also reported that 5.9% of the U.S. population was under five years of age. How many children under the age of five were living in the U.S. in 2020?

(2) Audryn added a $12 tip to a $68.73 restaurant bill. What percent tip did she leave? Round to the nearest tenth of a percent.

(3) Celeste has $40 that she plans on using to pay for lunch with a friend, including the tip. She always tips 16%. What is the maximum amount Celeste can spend on the food, so she can be sure to have enough money remaining to cover the tip? Round to the nearest cent.

After Preparation N.4 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration N.4, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
increase or decrease a quantity by a given percentage.	
represent chances or probabilities by both percentages and ratios.	
interpret a rate by stating its meaning verbally.	

This page titled 1.4.1: Preparation N.4 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.4.2: Exercise N.4

(1) The average annual snowfall for Cincinnati, OH is 23.3 inches. Buffalo, NY, receives an average of 95.4 inches. Fill in the blanks:

(a) Buffalo receives, on average, _____ % more snow than Cincinnati. Round to the nearest percent.

(b) Cincinnati receives, on average, _____ % less snow than Buffalo. Round to the nearest percent.

(2) The U.S. Census Bureau tracks changes in total population as well as populations within age groups. Between 2010 and 2020, the population 65 years and over increased at a rate of 38.6% . This increase is significantly greater than the total growth rate for the U.S. population of 7.4% during the same period.[14]

(a) Explain carefully what 38.6 percent refers to in the above statement.

(b) Explain carefully what 7.4 percent refers to in the above statement.

(c) Come up with a specific example (with numbers!) that illustrates the statement from the census. Your numbers do not have to agree with the actual U.S. population, they just need to describe a situation that agrees with the U.S. Census statement at the start of this problem.

(3) The U.S. Census Bureau also reported that in 1920 approximately 4.7% of the U.S. population was made up of individuals who were 65 years or older, while in 2020 this same age group made up about 16.8% of the population.

(a) Write a complete sentence that describes the absolute change in the number of people 65 years or older.

(b) Write a complete sentence that describes the relative change in the number of people 65 years or older.

(4) The U.S. population increased by 22.7 million people from 2010 to 2020. This was a 7.4% increase.[15] According to these figures, what was the U.S. population in the years 2010 and 2020? Round to the nearest million.

(5) The Centers for Disease Control and Prevention (CDC) frequently report statistics related to motor vehicle safety. The proportion of drivers reportedly distracted at the time of a fatal crash has increased from 7 percent in 2005 to 13 percent in 2020."[16]

(a) What does the 7 percent refer to in the above statement? What does the 13 percent refer to?

(b) If someone claimed that this represents a 6% increase in the proportion of drivers that were distracted at the time of a fatal crash, would that be true or false? Explain.

(6) In 2009, the federal minimum wage in the United States increased from $6.55 per hour (in 2008) to $7.25 per hour (in 2009). By what percentage did the minimum wage increase? Round to the nearest tenth of a percent.

(7) A Pew Research Center survey found that in 2004, 31% of adults in the U.S. supported gay marriage, and in 2021, 61% supported gay marriage.[17] Which of the three statements below are correct? Select all that apply. All results are rounded to the nearest percent or percentage point.

(i) Between 2004 and 2021, the number of people who support gay marriage increased by 97%.

(ii) Between 2004 and 2021, the number of people who support gay marriage increased by 30%.

(iii) Between 2004 and 2021, the number of people who support gay marriage increased by 30 percentage points.

(8) According to the CDC, in 2002, 4.3% of youths in the United States aged 12–13 years used alcohol. In 2019, this percentage dropped to 1.7%.[18]

(a) Write a complete sentence that describes the decline in illicit drug use among youth aged 12–13 years by measuring absolute change.

(b) Write a complete sentence that describes the decline in illicit drug use among youth aged 12–13 years by measuring relative change.

(c) Could someone mistakenly perceive one of these changes as seeming larger? Why?

(9) The following table is an excerpt of some statistics from the CDC concerning the smoking behavior of mothers in the United States in 2016:[19]

Education of Mother	Percent of Mothers 25 Years of Age and over who Smoked Cigarettes
Less than 12 years	11.7
High school diploma or GED	12.2
Some college or associate's degree	7.9
Bachelor's degree	1.0
Master's degree or higher	0.4

(a) Describe carefully the numerator and denominator that make up the 11.7% in the above table.

(b) Notice that in the first two rows, the percent changes from 11.7% to 12.2%. Find the absolute and relative change in this data. Round to nearest tenth of a percentage point and tenth of a percent.

(c) Using complete sentences, compare the two percentages discussed in Question 9(b). Your comparison should include statements about absolute change, relative change, and the relationship between the educational backgrounds of the individuals involved.

---

[14] https://www.census.gov/library/stories/2023/05/2020-census-united-states-older-population-grew.html

[15] https://www.census.gov/library/stories/2021/08/more-than-half-of-united-states-counties-were-smaller-in-2020-than-in-2010.html.

[16] https://www.forbes.com/sites/larrymagid/2012/06/10/cdc-study-too-many-teens-text-while-driving/#4397519b7302 and https://www.cdc.gov/niosh/motorvehicle/topics/distracteddriving/default.html

[17] https://www.pewresearch.org/short-reads/2022/11/15/about-six-in-ten-americans-say-legalization-of-same-sex-marriage-is-good-for-society

[18] https://www.cdc.gov/nchs/data/hus/2020-2021/SubUse.pdf

[19] https://www.cdc.gov/nchs/data/databriefs/db305.pdf

This page titled 1.4.2: Exercise N.4 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.5: The National Debt…How Big Is It?

INTRODUCTION

In this Collaboration, we will be discussing the concept of inflation. Inflation is the rate at which the general level of prices for goods and services is rising and, consequently, the buying power of currency is falling. In other words, if everything else (rent, groceries, utilities, etc.) gets more expensive, but your salary stays the same, your buying power decreases. To account for that, some state and private sector employees get annual salary adjustments based on the Consumer Price Index (CPI). Those adjustments are often referred to as Cost of Living Adjustment (COLA).

The following example illustrates the concept of adjusting for inflation. Consider a person who made $65,800 in 2022. How much would they need to make in 2023 to maintain the same buying power?

Year	Annual Expenses	Annual Salary
2022	$20,000	$65,800
2023	$22,000	

Another approach is to model this problem as a proportion and solve it for the missing quantity; Annual Salary in 2022 (denoted as x).

The reasoning behind this proportion is that in order to maintain the same buying power, the fraction of Annual Salary corresponding to Annual Expenses should be the same in 2022 and 2023.

$$\frac{Annual\ Expenses\ in\ 2022}{Annual\ Salary\ in\ 2022} = \frac{Annual\ Expenses\ in\ 2023}{Annual\ Salary\ in\ 2023}$$

$$\frac{20,000}{65,800} = \frac{22,000}{x}$$

$$20,000x = 22,000 \times 65,800$$

$$x = 72,380$$

Check:

$$\frac{20,000}{65,800} = \frac{22,000}{72,380}$$

$$30.4\% = 30.4\%$$

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- proportions are one way to compare numbers of varying magnitudes.
- percentages are useful when making comparisons across time.

By the end of this collaboration, you should be able to

- use the Consumer Price Index and proportional reasoning to adjust dollar amounts for inflation.
- compare the U.S. national debt at various times in history by both 1) adjusting for inflation, and 2) expressing the debt as a percentage of the Gross Domestic Product.

PROBLEM SITUATION 1: HOW MUCH IS THAT CAR?

(1) The average price of a new car in 1945 was $1,020. The Consumer Price Index was 18.0 in 1945 and 271.0 in 2021. Use this information to answer the following question: $1,020 in 1945 has the same buying power as how many dollars in 2021?

PROBLEM SITUATION 2: IS THE U.S. NATIONAL DEBT OUT OF CONTROL?

(2) The table below lists the U.S. national debt at various times in history. We need to adjust these figures for inflation. Adjusting them will allow us to more accurately compare the sizes of these debts to one another.

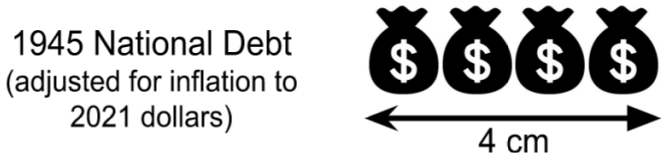
(a) Use the Consumer Price Index[21] to adjust the national debt for each year so that the debt is expressed in 2021 dollars. Enter the adjusted amounts in the table below. Round to two decimal places.

Year	National Debt (in billions of dollars)	Consumer Price Index (average annual) as reported by BLS	National Debt Adjusted for Inflation (expressed in billions of 2021 dollars)
1945	259	18.0	
1980	930	82.4	
1995	4,974	152.4	
2010	14,025	218.1	
2021	28,429	271.0	

(b) When adjusted for inflation, in which year was the national debt largest?

(c) When adjusted for inflation, in which year was the national debt smallest?

(3) When communicating quantitative information, an infographic can be a useful tool. In order to help people understand the national debt adjusted to 2021 dollar amounts, we can draw bags of money to illustrate the change. If we use a 4 cm segment (4 bags of money) to represent the 1945 national debt adjusted for inflation (see figure below), how long should the segments be for each of the other years' national debt after adjusting to 2021 amounts? Round your answers to the nearest tenth of a centimeter.



(a) How long should the segment be for **1980**? Round to one decimal place.

(b) How long should the segment be for **1995**? Round to one decimal place.

(c) How long should the segment be for **2010**? Round to one decimal place.

(d) How long should the segment be for **2021**? Round to one decimal place.

(4) According to the U.S. Treasury Department, a one-hundred-dollar bill is 0.0043 inches thick. Therefore, a stack of 100 one-hundred-dollar bills would be 0.43 inches thick. Assume that we could create stacks of one-hundred-dollar bills that would never topple over, no matter the height.

(a) How tall is the stack of money that represents the 1945 national debt, adjusted for inflation (expressed in billions of 2021 dollars)? Express your answer in miles and round to the nearest mile.

(b) How tall is the stack of money that represents the 2021 national debt? Express your answer in miles and round to the nearest mile.

PROBLEM SITUATION 3: ANOTHER VIEW OF THE U.S. NATIONAL DEBT?

(5) Now let us consider a different approach to analyzing the U.S. national debt over time. Just as the cost of dinner and a movie for Carlotta before and after college can be compared as a percentage of her income (as we thought about in the Preparation N.5), economists frequently compare the national debt at different times in history by measuring it as a percentage of Gross Domestic Product (GDP). The GDP is the estimated value of all the goods and services that are produced in a country in a given year.

(a) Fill in the table below with the national debt as a percentage of each year's GDP. Round to the nearest percentage point.

Year	National Debt (in billions of dollars)	GDP (in billions of current dollars)[22]	National Debt as a Percentage of GDP
1945	259	228	
1980	930	2,857	
1995	4,974	7,640	
2010	14,025	15,049	
2021	28,429	23,315	

(b) After adjusting the debt relative to GDP, in which year was the national debt the largest? In which year was the debt smallest? (Note: In this case, we say the national debts are being compared in relative terms.)

(c) Without adjustment, in which year was the national debt the largest? In which year was it the smallest? (Note: In this case, we say the national debts are being compared in absolute terms.)

(d) Is it better to measure the national debt in absolute terms or relative terms? Write a brief explanation.

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

---

[21] https://www.minneapolisfed.org/about-us/monetary-policy/inflation-calculator/consumer-price-index-1913-

[22] https://www.bea.gov/data/gdp/gross-domestic-product

This page titled 1.5: The National Debt…How Big Is It? is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.5.1: Preparation N.5

(1) Determine which of the following fractions are equivalent to $\frac{3}{4}$. Select all that apply.

(i) $\frac{15}{20}$

(ii) $\frac{6}{9}$

(iii) $\frac{24}{32}$

(iv) $\frac{75}{100}$

(v) $\frac{9}{16}$

(2) Mary is preparing for a party. She wants to create a snack mix that consists of 4 parts of M&M's to 7 parts of peanuts. List three possible combinations that would be an acceptable mixture of M&M's and peanuts.

(3) Solve the following equations for n:

(a) $\frac{80}{125} = \frac{n}{250}$

(b) $\frac{31.5}{n} = \frac{3.5}{540}$

(4) In May of 2023, the exchange rate was reported to be 1 U.S. Dollar ($) = 0.93 Euros (€). How many Euros are equivalent to $450? The Euro is divided into 100 cents. Round to the nearest cent.

(5) The population density of Martin County, Minnesota is reported to be 28 people/mi2. The county has a total area of 712 square miles. How many people would you expect to be living in this county?

One usually needs to take special care when comparing dollar amounts over time, as the following scenarios illustrate.

Scenario I

According to the U.S. Bureau of Labor Statistics (BLS), the average cost of white bread in the United States in 1980 was $.50 per pound while in 2022 it was $1.99 per pound.

(6) Does the rise in the cost of bread from 1980 to 2022 mean that bread is no longer affordable? Explain.

Scenario II

When Carlotta was a college student, she decided that after paying tuition and living expenses she could usually afford to go out to dinner and a movie once every other week. Now that she has graduated, she feels that she can afford dinner and a movie at least once a week.

(7) Does this mean that the cost of dinner and a movie has decreased over time? Explain.

(8) An online article[20], posted in 2020 looked at the prices of twenty-four popular video game consoles. It shows both the original price of the console, as well as the price adjusted for inflation in 2020.

Use data below for consoles released since 2000 to answer the following questions.

Console	Year	Original Price	2020 Price
PlayStation 2	2000	$299	$446.64
GameCube	2001	$199	$291.57
Xbox	2001	$299	$438.08
Xbox 360	2005	$299	$393.30

Wii	2006	$249	$321.19
PlayStation 3	2006	$499	$643.67
Wii U	2012	$299	$337.57
PlayStation 4	2013	$399	$444.96
Xbox One	2013	$499	$556.48
PlayStation 4 Pro	2016	$399	$429.69
Nintendo Switch	2017	$299	$318.77
Xbox One X	2017	$499	$525.80
Nintendo Switch Lite	2019	$199	$201.45
Xbox Series S/X	2020	$299	$299
PlayStation 5 Digital/Disc	2020	$399	$399

(a) Identify the two cheapest consoles by original price.

(b) Identify the three most expensive consoles by original price.

(c) Identify the two cheapest consoles by adjusted price.

(d) Identify the two most expensive consoles by adjusted price.

In the next collaboration, we will be comparing the U.S. national debt at different periods in time. The first step is to make sure that all the dollar amounts are expressed in comparable terms. The following table gives estimates of the U.S. national debt at various times in history.

Year	**U.S. National Debt**
1860	$64.8 million
1945	$259 billion
1980	$930 billion
1995	$4.974 trillion
2010	$14.025 trillion
2021	$28.429 trillion

(9) In order to better compare these quantities, convert all dollar amounts to billions of dollars.

Year	**U.S. National Debt**
1860	
1945	
1980	
1995	
2010	
2021	

(10) Based on this information, does it appear that the U.S. national debt continues to grow year after year? Write one reason why a person might say that the debt is not really growing during this period?

(11) In the next collaboration, we will be using the Consumer Price Index (CPI). Use the Investopedia site to learn about the CPI: https://www.investopedia.com/terms/c/consumerpriceindex.asp

How would you define what the CPI is?

After Preparation N.5 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration N.5, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
convert between millions, billions, and trillions.	
solve a proportion.	
express one number as a percent of another.	
Identify equivalent fractions.	

[20] https://www.ign.com/articles/comparing-the-price-of-every-game-console-with-inflation

This page titled 1.5.1: Preparation N.5 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.5.2: Exercise N.5

(1) The average price of a movie ticket in 1980 was $2.70. Calculate the cost of a movie ticket in 2020, using the Consumer Price Index of 82.4 in 1980 and 258.8 in 2020. Round to the nearest cent.

(2) In 2020, SONY released its newest version of the PlayStation, the PlayStation 5. The retail price was $399.99. Many consumers felt that this gaming console was overpriced. They pointed to the original PlayStation that was released in 1995 with a list price of $300. Using the fact that the CPI was 152.4 in 1995 and 258.8 in 2020, adjust the price of the original PlayStation to 2020 dollars. Which gaming console would you conclude was more expensive?

(3) Mary earns a yearly salary of $50,000 and will be receiving a $1,200 raise. Mike earns $9.25 per hour and will be receiving a 25-cent per hour raise. Who will be receiving the better raise in absolute terms? Who will be receiving the better raise in relative terms? Use the fact that Mike and Mary will both be working 40 hours per week and that there are typically 52 weeks in a year.

(4) Conduct an Internet search to find reports of the U.S. national debt for the last fifty years. How is the data reported or compared? What organization is reporting the data? Identify any possible bias in the way this organization reports their data.

(5) The table below provides national debt and population data for five different years.

(a) How much money would each person owe if the national debt were evenly distributed among all living people in the United States? Round answers to the nearest dollar and enter them in the table below.

Year	National Debt (in billions of dollars)	U.S. Population (in millions)	Individual Share
1945	259	140	
1980	930	227	
1995	4,974	266	
2010	14,025	309	
2021	28,429	332	

(b) In absolute terms, in which year do individuals owe the most?

(c) In absolute terms, in which year do individuals owe the least?

(d) Is this a reasonable way to try to compare the national debt from these years? Explain.

(6) (a) Adjust the dollar amounts you computed in Question 5(a) for inflation by expressing all the quantities in terms of the value of a dollar in 2021. Enter the individual share of the national debt from Question 5(a) then calculate the individual share in 2021 dollars. Round answers to the nearest dollar and enter them in the table below.

Year	Individual Share (from Question 5a)	Consumer Price Index (average annual) as reported by BLS	Individual Share (expressed in 2021 dollars)
1945		18.0	
1980		82.4	
1995		152.4	
2010		218.1	
2021		271.0	

(b) When adjusted for inflation, in which year do individuals owe the most?

(c) When adjusted for inflation, in which year do individuals owe the least?

(d) Is this a reasonable way to try to compare the national debt from these years? Explain.

This page titled 1.5.2: Exercise N.5 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.6: The Great Pacific Garbage Patch

INTRODUCTION

In this collaboration, you will set up and solve several proportions. Proportions are very forgiving on how you set them up. As long as you match "units," your proportion will be set up correctly.

For example, imagine a scale on a map indicates that 2 inches represents 5 miles. If two items are 4.5 inches apart on the map, how far apart are they on the earth? Take a minute to think about this in your group before moving on.

$$\textbf{Setup 1:} \quad \frac{2\ inches}{5\ miles} = \frac{4.5\ inches}{?\ miles}$$

For this setup, we have inches in the numerator on both fractions, and miles in the denominator, so our units match and can be used to give us 11.25 miles for the unknown value. The reason behind setting it up this way is that the ratio of the distance on a map to the actual distance is constant.

$$\textbf{Setup 2:} \quad \frac{2\ inches}{4.5\ inches} = \frac{5\ miles}{?\ miles}$$

For this setup, we have both original given values in the numerators. We put the second value and the unknown second value as denominators. So again, our units match and this setup will give the correct answer, 11.25 miles. The reason behind setting it up this way is that the ratio of two distances on a map is the same as the ratio of the actual distances.

Both setups are examples of proportions. Proportions are equations in the form of $\frac{a}{b} = \frac{c}{d}$. One approach to solving proportions is to write them as to $ad = bc$. This form is often easier to work with than the original setup.

For example: $\frac{2}{4.5} = \frac{5}{x}$

Is equivalent to: $2x = 4.5 \times 5$

$$2x = 22.5$$

$$x = 11.25$$

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- scaling linear distances by a factor will result in scaling areas by the square of that factor.

By the end of this collaboration, you should be able to

- compute a scale factor between two quantities having the same dimension.
- use a scale factor to convert units in a variety of contexts.

PROBLEM SITUATION: A LOT OF TRASH

The Great Pacific Garbage Patch has been receiving media attention since its discovery in 1997. The size of the Garbage Patch is not exact, because it has no specific border, and identifying a rough border depends upon the cutoff chosen for particle density (pollution level). The size of the Garbage Patch is usually measured as somewhere between 700,000 km^2 (about the size of Texas) and 15 million km^2 (about twice the size of the contiguous United States). To illustrate how large that is, the Great Pacific Garbage Patch is often described as "twice the size of Texas."

(1) Is the description, "twice the size of Texas," a reasonable estimate of the size of the Great Pacific Garbage Patch? Explain.

(2) Statistics books often use pie charts and bar graphs to portray quantities and sizes. In contrast, the image below, called an "infographic" (or "pictograph"), uses images rather than a chart or graph to represent the size of the objects in the Problem Situation.



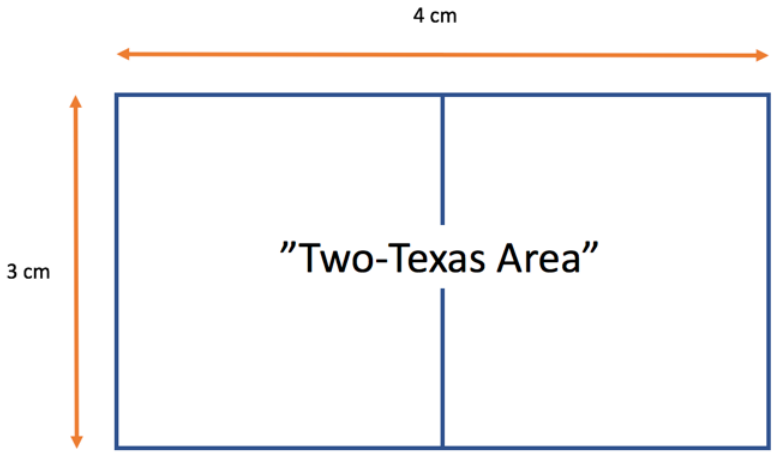
(a) Think about why the shape of Texas is a good image to help represent the Garbage Patch. Would a simple square or circle be more helpful or less helpful? Why?

(b) Does the size of the Garbage Patch in the above infographic seem reasonable? Explain why or why not.

The infographic above claims to show the size of the Garbage Patch in the Pacific Ocean. To better understand how large the infographic is claiming the Garbage Patch to be, we will create a **scale factor**. A *scale factor* is a number that when multiplied by a given quantity results in a proportional second quantity. The *scale factor* for this infographic will tell us the proportion of real-world kilometers to centimeters on the map, or how many kilometers are represented by one centimeter on the map.

(3) Let's say the distance (in centimeters) from the point shown in Australia to the point shown in Alaska on the infographic is 10.5 cm. If the actual distance from Australia to Alaska is 12,000 km, how many km does each cm on the infographic represent? Round to the nearest kilometer.

(4) Now you are going to estimate the total area of both Texas shapes from the given map. Let's imagine someone quickly sketched a rectangle to roughly represent the "two-Texas area."



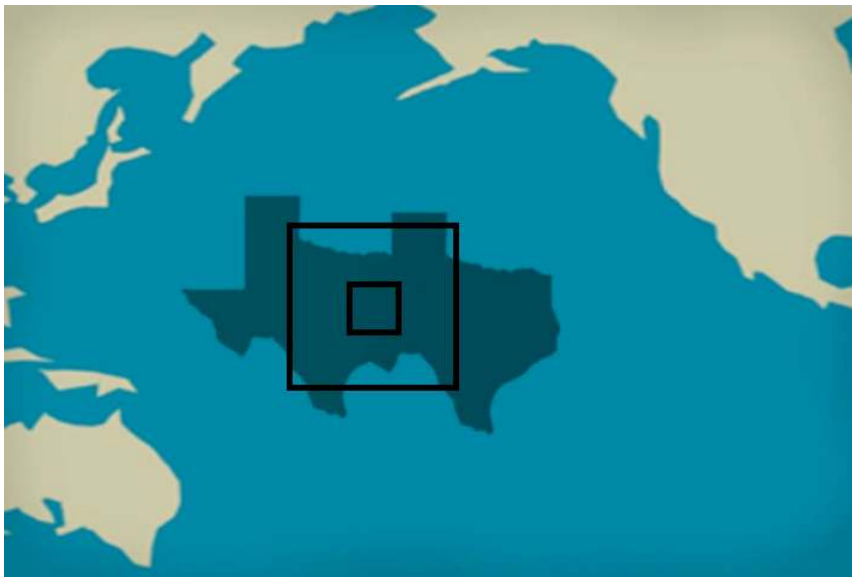
(a) Using the measurements in the sketch above, and the scale factor (km/cm) from Question 3, what would be the actual height and width (in kilometers) of the sketched rectangle on the infographic map?

(b) What actual area (in km^2) is represented by the rectangle of the "two-Texas area"? Write down your steps for doing this calculation.

(c) Is the graphical representation of the size of the Garbage Patch reasonable? Explain why or why not.

As discussed in Question 4, using a 3 cm by 4 cm rectangle somewhat accurately represents the upper estimate of area of the Great Garbage Patch of 15,000,000 km^2.

Now, let's visualize it in a slightly different way, using two squares:



- One square that represents the lower estimate of 700,000 km^2
- Another square one that represents the upper estimate of 15,000,000 km^2

(5) (a) What should be the dimensions of the smaller square that represents the area of 700,000 km^2?
Round your answer to two decimal places.

(b) What should be the dimensions of the larger square that represents the area of 15,000,000 km^2?
Round your answer to two decimal places.

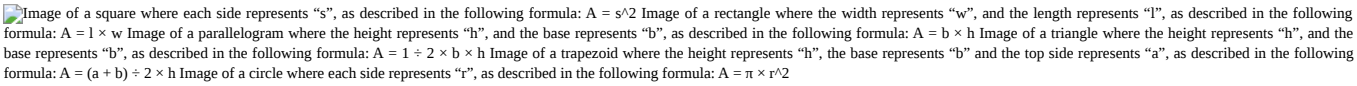
MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

This page titled 1.6: The Great Pacific Garbage Patch is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.6.1: Preparation N.6

Here are some formulas for areas of some basic geometric shapes:


Image of a square where each side represents “s”, as described in the following formula: A = s^2 Image of a rectangle where the width represents “w”, and the length represents “l”, as described in the following formula: A = l × w Image of a parallelogram where the height represents “h”, and the base represents “b”, as described in the following formula: A = b × h Image of a triangle where the height represents “h”, and the base represents “b”, as described in the following formula: A = 1 ÷ 2 × b × h Image of a trapezoid where the height represents “h”, the base represents “b” and the top side represents “a”, as described in the following formula: A = (a + b) ÷ 2 × h Image of a circle where each side represents “r”, as described in the following formula: A = π × r^2

(1) (a) Calculate the area of a circular tabletop whose diameter is 1.2 meters. Round to the nearest hundredth square meter.

(b) What would be the area of a square tabletop with a side length of 1.2 meters? Round to the nearest hundredth square meter.

(c) Which tabletop would require more paint to cover it?

A little research is required to prepare for the next collaboration. Collect the following information in your notebook and have it at hand for Collaboration N.6.

(2) What does the “Great Pacific Garbage Patch” refer to?

(3) What is the distance (in kilometers) between Alaska and Australia?

(4) What is the area of Texas in square kilometers (km^2)?

(5) What is the area of the Pacific Ocean (in km^2)?

After Preparation N.6 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration N.6, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
measure lengths.	
compute areas of circles and rectangles.	
perform computations with large numbers.	

This page titled 1.6.1: Preparation N.6 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.6.2: Exercise N.6

(1) Imagine you are planning to draw an infographic that models the Pacific Ocean as a square with 3-centimeter sides, and the Great Pacific Garbage Patch as a square within that square.

(a) The area of the Pacific Ocean is 165 million km^2. Use 15 million km^2 as the area of the Garbage Patch. What is the ratio of the area of the Garbage Patch to the area of the Pacific Ocean?

(b) If the Pacific Ocean is represented by the square with 3-centimeter sides, what is the area of this square?

(c) If the ratio between the squares is the same as the ratio in Question 1(a), what should be the area of the square representing the Garbage Patch? Round to the nearest hundredth of a cm^2.

(d) What should be the length of each side of the square representing the Garbage Patch? Round to the nearest tenth of a cm.

(e) What is the ratio of the sides of these two squares?

(f) What is the ratio of the areas of these two squares?

(2) Imagine you are planning to use circles to visually compare the populations of California (38,800,000), Texas (28,240,245), and New York (19,807,500). The size of each population will be represented by the area of each circle, not the radius. Round each answer to the nearest tenth.

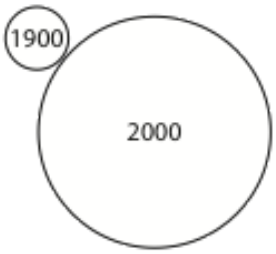
(a) If the circle for New York will have a radius of 5 cm, what would be the size (area) of that circle (which represents the population of New York)?

(b) What should the area and radius be for the circle representing the population of Texas? Round answers to one decimal place.

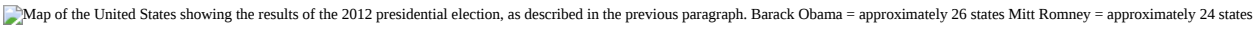
(c) What should the area and radius be for the circle representing the population of California? Round answers to one decimal place.

(3) A diagram has circles such that the area represents the populations of various countries in the world. The circle for India has a radius of 3.2 cm and the circle for China has a radius of 3.4 cm. If 18% of the world's population is in India, what percentage of the world's population is in China? Round to the nearest percent.

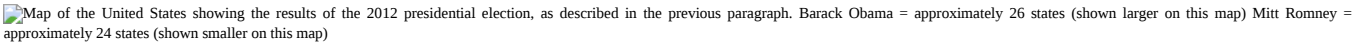
(4) Your team has been tasked with drawing an infographic to show the increase in population of the United States between 1900 and 2000. During this time, the U.S. population increased by 269% (almost quadrupling in size). A member of your team created the infographic below to depict this increase. Write a few sentences which explain to this teammate why the infographic is inaccurate.



(5) Below is a map of the continental United States, showing the results of the 2012 presidential election.[23] The darker shaded states are states in which the majority of voters voted for Barack Obama. In the lighter shaded states, a majority of voters voted for Mitt Romney. Looking only at the information provided on this map, which candidate would you say appeared to receive the majority of votes in the 2012 election? Fill in the blank: _____.


Map of the United States showing the results of the 2012 presidential election, as described in the previous paragraph. Barack Obama = approximately 26 states Mitt Romney = approximately 24 states

(6) In the next map, the size of each state has been adjusted to reflect the number of people living in the state. The larger the state's population, the larger the state appears to be on the map. Use this map to determine which candidate appeared to receive the majority of the votes in 2012. Fill in the blank: _____.


Map of the United States showing the results of the 2012 presidential election, as described in the previous paragraph. Barack Obama = approximately 26 states (shown larger on this map) Mitt Romney = approximately 24 states (shown smaller on this map)

_____

[23] http://www-personal.umich.edu/~mejn/election/2012

This page titled 1.6.2: Exercise N.6 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.7: An Index for Your Salary

INTRODUCTION

In addition to reviewing the concepts of Absolute Change and Relative Change, this collaboration introduces the concept of an Index. Similar to Relative Change, indexes are used to compare two quantities in relative terms. Later in this collaboration we will learn how to use it to keep track of changes over time. Here's an example that shows how to calculate Absolute Change, Relative Change, salary Ratio, and salary Index:

	2022 salary	**2023 salary**	**Absolute Change**	**Relative Change**
Tyler	$78,000	$88,000	$88,000 - $78,000 = $10,000	$10,000 / $78,000 * 100% = 12.82%
Amy	$87,000	$91,000	$91,000 - $87,000 = $4,000	$4,000 / $87,000 * 100% = 4.6%
Danela	$81,000	$85,000	$85,000 - $81,000 = $4,000	$4,000 / $81,000 * 100% = 4.94%

	2022 salary	**2023 salary**	**Salary Ratio**	**Salary Index**
Tyler	$78,000	$88,000	$88,000 / $78,000 = 1.1282	$88,000 / $78,000 * 100 = 112.82
Amy	$87,000	$91,000	$91,000 / $87,000 = 1.0460	$91,000 / $87,000 * 100 = 104.60
Danela	$81,000	$85,000	$85,000 / $81,000 = 1.0494	$85,000 / $81,000 * 100 = 104.94

Based on the calculations shown above, please generalize by matching the concept to the formula that can be used to calculate Absolute Change, Relative Change, Salary Ratio, and Salary Index for any two numbers.

• Salary Index • Relative Change • Salary Ratio • Absolute Change	• $\frac{New}{Old}$ • $\frac{New}{Old} \times 100$ • $\frac{New - Old}{Old} \times 100\%$ • $New - Old$

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- understanding relative changes in quantities over time can illustrate patterns of growth.
- indexes can be useful ways to compare various changing quantities to each other.

By the end of this collaboration, you should be able to

- create an index of relative changes given raw data.

PROBLEM SITUATION: CAREER PLANNING

As you read in the preparation, Gary and Anne are employees at a retail store for a cell phone provider. Gary is a Sales Associate and Anne is his Sales Supervisor. Their company offers many opportunities for growth in the organization. For example, Anne was promoted to Supervisor from Associate about three years ago. Gary is starting some long-term planning and wants to decide if he should stay with this company and seek a promotion, or if he wants to explore other options with another company, or even try a new career.

Gary has access to the following data from the past seven years about typical salaries for Sales Associates, Sales Supervisors, and Regional Sales Managers. Notice that salaries do not always increase from one year to the next. Gary wants to get a better idea about how the salaries change over time for each of these positions, to gain insight into his earning potential if he stays with his current company.

Salaries of Sales Personnel (2016–2023)

	2016	2017	2018	2019	2020	2021	2022	2023
Sales Associate	$41,500	$42,600	$42,800	$42,000	$43,100	$43,600	$44,700	$46,300
Sales Supervisor	$56,500	$57,100	$58,700	$57,300	$60,500	$59,600	$60,700	$62,600
Regional Sales Manager	$85,000	$89,500	$87,400	$82,700	$87,100	$89,000	$86,100	$91,000

(1) What do you notice about the figures in this table? What are some advantages or disadvantages of only using this table to compare salaries to make a decision? Be sure to describe the advantages and disadvantages quantitatively.

(2) (a) Make a quick sketch of a graph or visual display of these data. To make your graph easier to read, consider using different color ink to represent each of the positions.



(b) What are some advantages or disadvantages of using a graph like the one you made above to visualize data over time?

(3) Which position seems to show the most growth in salary over the past seven years? What particular quantities would be valuable for Gary to notice to help make his decision?

(4) The table below shows absolute and relative changes in salaries for each position from 2016 to 2023.

Absolute and Relative Changes from 2016 to 2023

Position	Absolute Change	Relative Change (rounded to nearest tenth of a percent)
Sales Associate	$4,800	11.6%

Sales Supervisor	$6,100	10.8%
Regional Sales Manager	$6,000	7.1%

(a) Explain what you observe in the numbers and in the absolute changes and relative changes. For example, you could explain how the Sales Associate can have the lowest absolute increase in salary but the highest relative increase.

(b) Now calculate the absolute and relative changes in salary from 2021 to 2022. Fill in the table below with your answers. Round your answers to the nearest tenth of a percent.

Absolute and Relative Changes from 2021 to 2022

Position	Absolute Change	Relative Change
Sales Associate		
Sales Supervisor		
Regional Sales Manager		

(c) Which position had the greatest absolute increase in salary? Which position had the greatest relative increase in salary? Write an explanation for what you observe in the absolute changes and relative changes.

Salary Index

One way to systematically look at a large set of relative changes over time is to create an **index**. Making an index requires you to make a lot of calculations similar to what you did in the preparation for this collaboration. A salary index will look like a table of numbers that answer the question, *"The salary for this year is what percent of the base year?"*

For example, in the preparation, you were given the following information:

In 2022, Gary earned $48,300, but he received a raise in 2023 and earned $53,400.

If Gary meets his supervisor's benchmark goals in 2023, then he will earn $59,150 in 2024.

Year	Gary's Salary
2022	$48,300
2023	$53,400
2024	$59,150

You then used the given information to answer the following questions:

Gary's 2023 salary is _____% of his 2022 salary. Round your answer to the nearest percent.

Gary's 2024 salary is _____% of his 2022 salary. Round your answer to the nearest percent.

In this example, 2022 is the base year or the reference value. Note that both the 2023 and 2024 indices are calculated by dividing the corresponding salaries by the base year salary.

Please review these questions with your group and verify that the correct percentages are rounded to the nearest percent:

- Gary's 2023 salary is 111% of his 2022 salary
- Gary's 2024 salary is 122% of his 2022 salary

Gary's Salary Index

Year	Salary Index (Base Year = 2022)
2022	100
2023	111

2024	122

(5) Now, use the salary information from the problem situation to create a salary index for each of the three positions. You will be creating three separate indices, one for Sales Associate, one for Sales Supervisor, and one for Regional Sales Manager. Use 2016 as the base year, and round each answer to the nearest tenth.

Salaries of Sales Personnel (2016–2023)

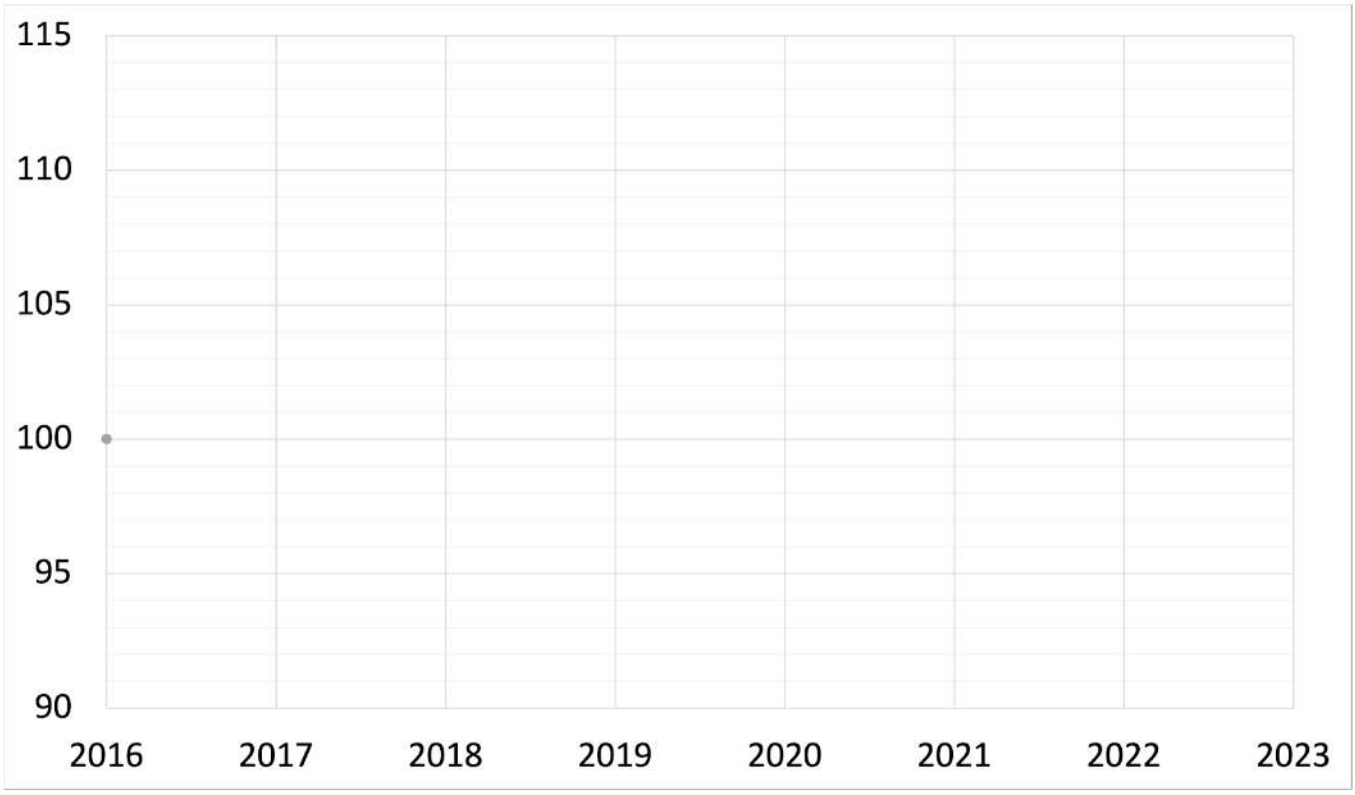
	2016	2017	2018	2019	2020	2021	2022	2023
Sales Associate	$41,500	$42,600	$42,800	$42,000	$43,100	$43,600	$44,700	$46,300
Sales Supervisor	$56,500	$57,100	$58,700	$57,300	$60,500	$59,600	$60,700	$62,600
Regional Sales Manager	$85,000	$89,500	$87,400	$82,700	$87,100	$89,000	$86,100	$91,000

Salary Index (Base Year = 2016)

Year	Sales Associate	Sales Supervisor	Regional Sales Manager
2016	100.0	100.0	100.0
2017	102.7	101.1	105.3
2018			
2019			
2020			
2021			
2022			
2023			

(6) Use quantities in your index table to explain your observations and ideas about what Gary should know about how the salaries change over time.

(7) Make a graph of the indices that you created. What are the advantages and disadvantages of using this graph to interpret how the salaries changed over time?



MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

This page titled 1.7: An Index for Your Salary is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.7.1: Preparation N.7

Gary is a Sales Associate in a retail store for a cell phone provider. Anne is the Sales Supervisor for Gary's store. In the next collaboration, you will explore ways to measure and compare how Gary's and Anne's salaries change over time.

(1) In 2022, Gary earned $48,300. In 2023, he received a raise of $5,100. Comparing Gary's increased salary to his previous salary, what would be a good estimate for the relative increase? Explain your strategy.

(2) Calculate the missing amount in each statement below and fill in the blanks:

(a) Gary's new salary is _____ times his previous salary. Round your answer to the nearest hundredth.

(b) Gary's new salary is _____% of his previous salary. Round your answer to the nearest percent.

(c) Gary's new salary is _____% more than his previous salary. Round your answer to the nearest percent.

(3) If Gary meets his supervisor's benchmark goals in 2023, then he will earn $59,150 in 2024. Calculate the missing amount in each statement below and fill in the blanks, assuming Gary meets the benchmark goals.

(a) Gary's 2024 salary is _____ times his 2022 salary. Round your answer to the nearest hundredth.

(b) Gary's 2024 salary is _____% of his 2022 salary. Round your answer to the nearest percent.

(4) If Gary meets his goals, his supervisor Anne will also receive a pay raise. Anne's salary in 2023 is $67,680, and she would receive a pay raise of $6,100 for 2024. What would be a good estimate for the percent increase in Anne's pay? Explain your strategy.

(5) Calculate the missing amount in each statement below and fill in the blanks.

(a) Anne's 2024 salary is _____ times her 2023 salary. Round your answer to the nearest hundredth.

(b) Anne's 2024 salary is _____% of her 2023 salary. Round your answer to the nearest percent.

(6) Write an explanation of how Gary's raise from 2023 to 2024 compares to Anne's raise from 22023 to 2024. Be sure to include specific quantities and mention both the absolute change and relative change in each of their salaries.

After Preparation N.7 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration N.7, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
compare two quantities using a ratio.	
describe a number as a percent of another number.	
calculate and interpret absolute and relative changes.	
create a line graph.	

This page titled 1.7.1: Preparation N.7 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.7.2: Exercise N.7

The following table shows the average price per gallon of regular gasoline in U.S. cities,[24] and the price index for petroleum imports to the United States.[25] The Petroleum Import Price Index measures the change in international value of the crude petroleum that is imported to the United States. Petroleum can be refined and distilled to create a variety of fuels, including gasoline, or other industrial chemicals.

	Average Gas Price (Annual)	Petroleum Import Price Index (June)
2012	$3.680	357.2
2013	$3.575	364.9
2014	$3.437	387.9
2015	$2.520	226.1
2016	$2.250	173.1
2017	$2.528	178.7
2018	$2.813	255.4
2019	$2.691	236.6
2020	$2.258	142.6
2021	$3.100	267.6
2022	$4.059	456.5

(1) (a) Using the data from the two consecutive years 2019-2020, calculate the relative changes in gas price and petroleum import price. Round to the nearest hundredth of a percent.

(b) Using the data from the two consecutive years 2020-2021, calculate the relative changes in gas price and petroleum import price. Round to the nearest hundredth of a percent.

The term **volatility** refers to the likelihood that something will change rapidly and unpredictably. Something with a price or value that varies greatly over time may be called **volatile**. Investors consider volatility when making investment decisions. An investor who takes risks, for example, might be interested in volatile stocks because the risk of a big decrease in value is outweighed by a chance for a big increase in value. A more conservative investor would avoid volatile stocks and look instead for something with a steadier value over time. Investors can also invest in commodities, such as crude oil, gold, and corn.

(2) Which of the two commodities in the table above, gas or import petroleum, could be described as more volatile? Please explain your reasoning using the Writing Principle

(3) With reference to the table above, calculate the overall relative changes from 2012 to 2022. Round your answers to the nearest hundredth of a percent.

(i) Overall relative change of the average gas price:

(ii) Overall relative change of the petroleum import price index:

In this collaboration, you used indices to compare how different salaries changed over time. Other kinds of indices can be used to make comparisons between other kinds of varying quantities. The Uniform Crime Reporting Program (https://ucr.fbi.gov) and the Bureau of Justice Statistics (www.bjs.gov) provide information and quantitative data about crime. Crime statistics are often reported as a “rate per 100,000 people” to help compare small percents. For example, in the table below, a robbery rate can be calculated by comparing the number of robberies in a year to the total population in that area, but these percentages can be very small and hard to compare. So the rate is then used to calculate how many people out of 100,000 would be a victim of robbery.

Estimated Crime in 2020[26]

State	Population	Number of Robberies	Robbery Rate (as a percentage, rounded to the nearest ten thousandth)	Robbery Rate (per 100,000 people, rounded to the nearest tenth)
Montana	1,084,225	280	0.0258%	25.8
North Dakota	779,094	153	0.0197%	19.7
Vermont	643,077	66	0.0103%	10.3

Notice that we can use this table to conclude that these three states have nearly the same low robbery rate.

Now answer **Questions 4-13** to complete the table below and write two or three sentences to summarize the data using the robbery rate per 100,000 people.

Estimated Crime in 2020[27]

State	Population	Number of Robberies	Robbery Rate (as a percentage, rounded to the nearest ten thousandth of a percent)	Robbery Rate (per 100,000 people, rounded to the nearest tenth)
California	39,538,223	44,915	(4)	(5)
District of Columbia	689,545	2,295	(6)	(7)
Minnesota	5,706,494	3,909	(8)	(9)
New York	20,201,249	18,302	(10)	(11)
Ohio	11,799,448	7,894	(12)	(13)

[24] https://www.eia.gov/dnav/pet/hist_xls/EMM_EPM0_PTE_NUS_DPGa.xls

[25] www.bls.gov/data

[26] https://cde.ucr.cjis.gov

[27] https://cde.ucr.cjis.gov

This page titled 1.7.2: Exercise N.7 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.8: Screening Tools–How Effective Are they?

INTRODUCTION

No medical test is perfect. Some are very good at determining when a patient has a disease, and some are very good at determining when a patient does not have a disease.

For example, you may have seen advertisements for *Cologuard*, a test for colon cancer, which claims to have a "specificity of 87%." But what does it mean for a test to have a specificity of 87%? You should be able to answer that question after completing this collaboration.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- screening tools for medical conditions or diseases are not 100% effective.
- there are at least two ways to quantify the effectiveness of screening tools.
- when a doctor decides to use a particular screening tool, the doctor has to decide whether it is more important to have a **high specificity** or a **high sensitivity**.

By the end of this collaboration, you should be able to

- interpret the meaning of a screening tool's sensitivity and specificity.

PROBLEM SITUATION: PSA

Prostate cancer is cancer of the prostate gland, which surrounds the neck of the bladder in males. Prostate cancer is a common type of cancer in men. Doctors have developed an initial screening test for detecting prostate cancer which measures Prostate Specific Antigen (PSA). In this initial screening test, the doctor takes a sample of the man's blood and measures the level of PSA. PSA is a substance which is only produced by the prostate gland. Men who have prostate cancer frequently have high levels of PSA in their blood.

However, PSA is also present in men who do not have prostate cancer. So, it is possible that a man with a high PSA level might not have prostate cancer. It is also possible that a man with a low PSA level might still have prostate cancer. This is why the PSA blood test is only used as the initial screening test. A more effective way to know if prostate cancer is present is to conduct a second screening test using a biopsy. In a biopsy, doctors cut small tissue samples from a man's prostate gland. The samples can then be tested to see if cancer is present.

Why don't all men start with the biopsy? Doctors try to avoid unnecessary biopsies because the procedure is invasive, uncomfortable, and expensive. Even though the PSA blood test is less effective than a biopsy, it is simpler to use as an initial screening tool for prostate cancer because the only thing required is a blood sample. Doctors use the results of the PSA blood test to help them decide whether to recommend a second screening (the biopsy). In this collaboration, we will be focusing on the initial screening tool: the PSA blood test.

The table below shows the risk of prostate cancer for men with certain PSA levels in their blood (measured in nanograms per milliliter, or ng/mL):[28]

PSA Level (nanograms per milliliter, ng/mL)	**Risk of Prostate Cancer**
< 0.5	7%
0.6-1.0	10%
1.1-2.0	17%
2.1-3.0	24%
3.1-4.0	27%

(1) Review the table. Based on this information, do PSA levels seem to be associated with risk of prostate cancer? Record any patterns that you notice.

Imagine you are a doctor. You decide to use the PSA blood test to screen for prostate cancer in your patients. But, at what PSA level will you decide to do a biopsy on a patient? In other words, what is your "cut-off" value?

A PSA "cut-off" value is the PSA level at which you start recommending biopsies for your patients. That is, you will recommend a biopsy for any man whose PSA level is *at or above* the "cut-off" value. Any man whose PSA level is *below* the "cut-off" value will not be given a biopsy, because prostate cancer is less likely, though still possible.

For example, suppose the PSA cut-off value is set at 2.0. A man with a PSA level of 1.3 will be told that it is unlikely he has prostate cancer, and a biopsy will not be recommended. A man with a PSA level of 2.5 will be told that, because it is likely enough that he has prostate cancer, a biopsy is recommended. In both cases, it is uncertain whether the men have prostate cancer, but the higher PSA level of 2.5 indicates that it is more likely in that man.

(2) Suppose your goal in choosing a PSA cut-off value is to be certain that you detect almost all individuals with cancer. With that goal, should the cut-off value be high or low? Explain your reasoning.

(3) Suppose your goal is different now. Imagine that your goal in choosing a PSA cut-off value is to limit the number of unnecessary biopsies. Should the cut-off value be high or low? Explain your reasoning.

When determining a cut-off value, doctors may consider how effective the screening tool will be at different values. Doctors will most likely discuss the effectiveness with their patients when administering a screening tool. When discussing how effective a screening tool is, two technical terms are frequently used: ***sensitivity*** and ***specificity***. These terms quantify probabilities or likelihoods. They are expressed as percentages.

The **sensitivity** is the likelihood that a test correctly produces a positive result for those who should test positive. For example, with the PSA blood test, sensitivity is the likelihood that the test will indicate the presence of cancer when cancer is actually present.

The **specificity** is the likelihood that a test correctly produces a negative result for those who should test negative. For example, with the PSA blood test, specificity is the likelihood that the test will indicate the absence of cancer when cancer is actually absent.

The terms *sensitivity* and *specificity* are not just used when screening for prostate cancer, but can be calculated for any test that screens for the presence or absence of a disease or medical condition.

Sometimes students have trouble remembering which of the two terms refers to patients with a disease, and which refers to patients without the disease. It may be helpful to use the mnemonic that doctors should be especially *sensitive* in how they interact with *sick* patients, those who actually have the disease.

The following table, from *JAMA: The Journal of the American Medical Association*[29], shows the specificity and sensitivity associated with using three different possible cut-off values when screening for prostate cancer:

PSA Level (ng/mL)	Sensitivity	Specificity
1.1	83%	39%
2.6	40%	81%
4.1	20%	94%

(4) Suppose you choose 1.1 as the PSA cut-off value.

(a) Write a complete sentence that explains the sensitivity associated with this screening tool and cut-off value. Imagine that you are writing this to a friend who does not understand sensitivity.

(b) Write a complete sentence that would explain to a friend the specificity associated with this screening tool and cut-off value. Again, imagine your friend does not understand specificity.

(c) With this screening tool and cut-off value, what percentage of men without cancer would get unnecessary biopsies?

(d) With this screening tool and cut-off value, what percentage of men with cancer would not get necessary biopsies?

(5) Suppose you were to choose to use 4.1 as the cut-off value.

(a) Write a complete sentence that would explain to a friend the sensitivity associated with this screening tool.

(b) Write a complete sentence that would explain to a friend the specificity associated with this screening tool.

(c) With this screening tool and cut-off value, what percentage of men without cancer would get unnecessary biopsies?

(d) With this screening tool and cut-off value, what percentage of men with cancer would not get necessary biopsies?

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

---

[28] Ian M. Thompson, et al, "Prevalence of Prostate Cancer among Men with a Prostate-Specific Antigen Level ≤4.0 ng per Milliliter," New England Journal of Medicine 350, no. 22(2004): 2239-46.

[29] Ian M. Thompson et al, "Operating Characteristics of Prostate-Specific Antigen in Men with an Initial PSA Level of 3.0 ng/mL or Lower," JAMA 294, no. 1(2005): 66-70.

This page titled 1.8: Screening Tools–How Effective Are they? is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.8.1: Preparation N.8

(1) Suppose a local food bank is selling raffle tickets as a fundraiser and claims that each raffle ticket has a 3% chance of winning one of several prizes. Additionally, suppose an animal shelter is also selling raffle tickets for a different fundraiser, and reveals that there are 1200 total tickets, and exactly 48 will win a prize.

(a) Which fundraiser offers a greater chance of winning a prize? Explain your answer using the likelihood of winning each raffle (in percentages).

(b) If you wanted to increase your chances of winning one of the raffles would it be better to buy five tickets or ten tickets? Explain your reasoning.

(c) If you bought 150 raffle tickets for the animal shelter's fundraiser, how many tickets would you expect to be winners?

(d) If you bought 150 raffle tickets for the food bank's fundraiser, how many tickets would you expect to be winners?

(e) If you bought 150 tickets for each raffle, is it possible that none of the tickets will win? Explain your answer.

(2) If there is a 61% chance that Carol gets accepted to Andover University, what is the chance, or probability, that Carol does not get accepted?

The **probability** of an event is defined as the likelihood that it will happen, and gives the chance that it *will* occur. Probability is often given as a percentage, which must always be between 0% (impossible) and 100% (certain). For instance, a team may be told they have a "30% chance of winning the game," which would mean they are more likely to lose than win. Probability is sometimes described using words such as "unlikely," "possible," "even chance," and "likely." For example, if the probability of an event is 4%, we would say it is unlikely.

It is also common to represent probabilities as ratios. As an example, if you flip a coin, what is the likelihood that the coin will come up "heads"? The chances are 1 out of 2, or ½, since there is a head on only one of the two sides of a coin. You can think of the ratio as

$$probability = \frac{number\ of\ successful\ outcomes}{number\ of\ possible\ outcomes}$$

This formula can be used to find any probability when all possible outcomes are equally likely. In the case of the coin, there was only 1 successful outcome (heads turning up) out of 2 possible outcomes (heads or tails).

(3) Suppose you roll a six-sided die.

(a) What is the probability that a 3 turns up?

(b) What is the probability that an even number turns up?

(c) What is the probability that a number greater than 2 turns up?

(4) Suppose that you have a bag containing 3 green marbles, 8 blue marbles and 13 red marbles. Suppose you reach into the bag and randomly draw one marble. You will win a treat if it is a green marble. What is the probability that you win a treat?

After Preparation N.8 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration N.8, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
read information presented in a table.	
represent risks or probabilities by both percentages and ratios.	

This page titled 1.8.1: Preparation N.8 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.8.2: Exercise N.8

(1) Suppose a company is developing a new pregnancy test. The test was put through a trial in which data was collected on 500 women who took the pregnancy test. The results of the trial are summarized below, in what is called a **two-way table**. A two-way table displays the counts for certain groups. For example, the value in the first row of numbers and the first column of numbers of the table shows that 170 of the 500 women were pregnant and had a positive pregnancy test.

Now, pregnancy is not a disease, but it is a medical condition, so the terms sensitivity and specificity can still be used when describing a pregnancy test, based on whether the woman is or is not pregnant. Answer the following questions to increase your familiarity with two-way tables.

Pregnancy Test Trial Results	Woman is Pregnant	Woman Is Not Pregnant
Pregnancy Test is Positive	170	6
Pregnancy Test is Negative	34	290

(a) What is the **sensitivity** of this pregnancy test, based on the collected data? Round to the nearest tenth of a percent.

(b) What is the **specificity** of this pregnancy test, based on the collected data? Round to the nearest tenth of a percent.

(2) Suppose that 50 men with a family history of prostate cancer participate in a study. First, the men's PSA levels are tested. After that, biopsies are performed on all of them, to gauge the effectiveness of PSA screening, and determine whether biopsies should have been performed. (Outside of the study, biopsies wouldn't be performed on all of them). The following table shows the results of the study.

PSA Level (ng/mL)	**Number of Men with Prostate Cancer**	**Number of Men Without Prostate Cancer**
3 and above	4	6
2-2.99	3	9
1-1.99	2	12
0-0.99	1	13

(a) Suppose the cut-off value is set at 2. What is the sensitivity for this group? What is the specificity?

(b) Suppose the cut-off value is raised to 3. What is the sensitivity now for this group? What is the new specificity?

(c) What happened to sensitivity when the cut-off value was raised from 2 to 3? Why? Is this good or bad?

(d) What happened to specificity when the cut-off value was raised from 2 to 3? Why? Is this good or bad?

(e) Suppose the PSA cut-off value is dropped to 1. Given this PSA cut-off value, how many healthy men among the group of 50 would be told incorrectly that they may have cancer? How many men with cancer would be told incorrectly that they probably do not have cancer? How many of the men would receive a correct diagnosis?

(3) Below is a list of screening tools that are used to detect the presence or absence of certain conditions. Pick one screening tool and conduct some research to determine how effective it is. Can you find any information on the screening tool's sensitivity and specificity? Does your research indicate that there are any other quantitative measures of its effectiveness? State carefully where you obtained the information and whether or not you believe it is reliable.

(a) Mammograms are used to screen for breast cancer.

(b) There is an ELISA (Enzyme-Linked ImmunoSorbent Assay) test developed to screen for the presence of HIV, the virus that causes AIDS.

(c) The FDA recently approved a new over-the-counter AIDS test to screen for the presence of the HIV virus. "Over-the-counter" means that the test can be purchased at a pharmacy and used in the privacy of one's own home without having to send anything off to a lab for analysis.

(d) Exercise stress testing is sometimes used as a screening tool for Coronary Artery Disease (CAD).

(e) Over the counter pregnancy tests are available to quickly test for pregnancy.

(f) The Purified Protein Derivative (PPD) test is used to screen for tuberculosis. Sometimes this test is also called the Mantoux test or tuberculin sensitivity test.

(g) The Beck Depression Inventory is used to screen individuals for depression.

(4) Suppose a drug company is designing an over the counter HIV/AIDS test and they will not be able to develop a test that has both a very high sensitivity and a very high specificity. Which of these, sensitivity or specificity, is the most important to maximize? Explain your answer.

This page titled 1.8.2: Exercise N.8 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.9: A Closer Look at Screening Tools

INTRODUCTION

In this collaboration, you'll be taking a deeper dive into sensitivity, specificity and related probabilities. Let's check in on a few things to get us started. In your group, discuss the following questions:

1. Consider a deck of 52 playing cards. If you draw one card randomly, what is the probability that it is a queen?
2. Now take the complete deck again and take out all of the cards except the face-cards (just leaving the kings, queens and jacks). Now randomly draw a card. What is the probability that the card is a queen this time?
3. How do the numerators and denominators compare in each of these probabilities?

The first draw could be restated as: “Draw a card from the deck. What is the probability that it is a queen?”

The second draw would be restated as something like this: “If only the face cards are in the deck, then what is the probability of drawing a queen?”

The "if" part of this statement is adding a condition which changes the group from which you are drawing.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- conditional probability is the probability that an event will occur given that another event has already occurred.
  - it is important to be careful when interpreting phrases about conditional probabilities.
  - a two-way table is a useful tool to analyze conditional probabilities.
  - some calculations related to a tool’s effectiveness cannot be computed directly from its sensitivity or specificity.

By the end of this collaboration, you should be able to

- correctly interpret phrases or statements referring to conditional probabilities.
- calculate different types of conditional probabilities with information in a two-way table (contingency table). These conditional probabilities include specificity and sensitivity.
- create a two-way table to model the effectiveness of a screening tool.

PROBLEM SITUATION: USING SCREENING TOOLS TO MAKE DECISIONS

In the previous collaboration (N.8), you learned how difficult it is to decide on a cut-off value when trying to identify people who may have cancer. There can be benefits to certain cut-off values, but also drawbacks. A screening tool’s sensitivity and specificity can help quantify how effective the screening tool is. In this collaboration, we will examine other factors that are important when analyzing the effectiveness of screening tools.

There are a variety of probabilities, or likelihoods, associated with analyzing the effectiveness of screening tools. Specificity and sensitivity are two examples, but there are many others. Describing these probabilities in everyday language can sometimes be very difficult. Doctors and scientists who study screening tools have to be careful about how they describe and distinguish between these probabilities.

When answering questions about a test’s effectiveness, it is important to read each question carefully to determine the right proportion.

For example, consider the following two questions:

Question A: If a screening tool is given to a person without cancer, what is the probability that the screening tool will suggest that cancer might be present?

Question B: If a screening tool suggests that cancer is present in a person, what is the probability that the individual does not have cancer?

(1) Compare **Question A** and **Question B** above. Are these two questions asking the same thing, phrased in two slightly different ways? Or, are these two questions asking for different information, and their answers could be different? Explain.

(2) Suppose you were being screened for a disease, such as cancer. Unfortunately, your doctor tells you that the screening tool produced a positive result. This suggests that cancer may be present.

Looking back at **Question A** and **Question B**, which question would you like your doctor to answer? Why?

The questions **Question A** and **Question B** in the Problem Situation are examples of **conditional probabilities**, which are questions that can be put into the following form:

If [something has happened], what is the probability that [something else will happen]?

For example, Jan rolls a six-sided die and does not show the result to her friend, but tells her friend she rolled an odd number. What is the probability of her friend guessing correctly that Jan rolled a 5? You can think of this problem as: *If her friend knows an odd number was rolled, what is the probability that her friend will guess it?* This is a conditional probability. (**Note**: The answer is ⅓, since there is one correct answer, 5, and three possibilities: 1, 3, and 5.)

As we discussed in Questions 1 and 2, great care must be taken to determine exactly what calculations should be performed to find a specific conditional probability. Questions that are asking for different conditional probabilities may appear quite similar at first glance. Determining the numerator and denominator for the conditional probability may require reading the problem carefully to guarantee understanding, and may require finding the sum of a row or column in a two-way table.

Suppose we are studying the effectiveness of Screening Tool X, and we decide to investigate the research questions A and B given earlier. First, we apply Screening Tool X to 1,200 people. For each patient, Screening Tool X produces either

- a positive result if the tool suggests the presence of cancer, or
- a negative result if it suggests the absence of cancer.

Remember, as we saw in Collaboration N.8, most screening tools are not guaranteed to be correct in either case, and it is possible for either a negative result or a positive result to be wrong.

For the purposes of our study, suppose that after being tested with Screening Tool X and given an initial diagnosis, everyone undergoes a biopsy to see if cancer was or was not actually present. Table 1 displays the results of this hypothetical experiment in a two-way table:

Table 1

	Cancer is Present	**Cancer is Not Present**	**Total**
Result of Screening Tool X is Positive	53	56	109
Result of Screening Tool X is Negative	211	880	1091
Total	264	936	1200

(3) What is the *sensitivity* of this test? Remember, the sensitivity of the screening tool is the probability that the screening tool will be correct when cancer is actually present. In other words, what percentage of individuals with cancer received a positive result from the screening tool? Take a minute to try this on your own before sharing with your group. Write your answer as a percentage rounded to one decimal place.

(4) What is the *specificity* of this test? Remember, the specificity is the probability that the screening tool will be correct when cancer is not present. In other words, what percentage of individuals without cancer received a negative result from the screening tool? Write your answer as a percentage rounded to one decimal place.

(5) Recall Question A from the Problem Situation:

Question A: If a screening tool is given to a person without cancer, what is the probability that the screening tool will suggest that cancer might be present?

Use Table 1 above to answer **Question A**. Does your answer appear to be related in any way to either the specificity or sensitivity? If so, how? (**Hint**: Read through Question A carefully, thinking about a description of the numerator and denominator in that probability statement, and relating those descriptions to the problems you just completed on specificity and sensitivity.)

(6) Now think about Question B from the Problem Situation:

Question B: If a screening tool suggests that cancer is present in a person, what is the probability that the individual does not have cancer?

Use Table 1 above to answer **Question B**. Does this answer appear to be related in any way to either the specificity or sensitivity? If so, how? (**Hint**: Read through Question B carefully, thinking about a description of the numerator and denominator in that probability statement, and relating those descriptions to the problems you just completed on specificity and sensitivity.)

Now, suppose we study the effectiveness of Screening Tool X on a different population that is known to have a low cancer rate. We apply the screening tool to 1,200 people in this new population. Table 2 below displays the results of this hypothetical experiment:

Table 2

	Cancer is Present	**Cancer is Not Present**
Result of Screening Tool X is Positive	3	70
Result of Screening Tool X is Negative	10	1117

(7) Determine the specificity and sensitivity of Screening Tool X in the new population. Write your answers rounded to the nearest percent.

(8) As you did for Questions 5 and 6, answer Question A and Question B using the results of Screening Tool X in this new population.

Question A: If a screening tool is given to a person without cancer, what is the probability that the screening tool will suggest that cancer might be present?

Question B: If a screening tool suggests that cancer is present in a person, what is the probability that the individual does not have cancer?

(a) Use Table 2 above to answer **Question A**. Does your answer appear to be related in any way to either the specificity or sensitivity? If so, how?

(b) Use Table 2 above to answer **Question B**. Does this answer appear to be related in any way to either the specificity or sensitivity? If so, how?

(c) Compare your answers to Questions 8(a) and 8(b) to your answers to Questions 5 and 6 above. Why do you think the results are similar or different?

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

This page titled 1.9: A Closer Look at Screening Tools is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.9.1: Preparation N.9

(1) A high school class took a math test. To help prepare future classes for similar tests, the counselors at the school did additional research on the class's study habits. Here is the data from that research, given in a two-way table. (**Note**: You may wish to use the blank spaces provided to compute totals.)

	Studied	**Didn't Study**	Total
Passed the Test	70	4	
Failed the Test	10	32	
Total			

(a) How many students studied for the test?

(b) How many students didn't study for the test?

(c) How many students passed the test?

(d) How many students failed the test?

(e) How many students took the test?

(f) What percentage of students passed the test? Round to the nearest tenth of a percent.

(g) What percentage of students studied for the test? Round to the nearest tenth of a percent.

(h) What percentage of students who studied for the test, passed the test? Round to the nearest tenth of a percent.

(i) What percentage of students who failed the test, didn't study for the test? Round to the nearest tenth of a percent.

(2) A local businesswoman wants to measure the ability of advertisement slogans to catch people's attention. She designs two signs, each with a different slogan, places them in store windows in similar stores in a mall, and observes shoppers' reactions. She records her information in the following table:

	Slogan 1	**Slogan 2**
Number of people who stopped to take a closer look at the sign	36	27
Number of people who passed the sign without stopping	510	288

(a) Which slogan appeared to be the most effective? Explain your answer.

(b) If, on another day, only 315 people encountered the sign with Slogan 1 on it, how many people do you think might actually stop and take a closer look?

After Preparation N.9 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration N.9, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
read information presented in a two-way table.	
represent risks or probabilities by both percentages and ratios.	
interpret the meaning of a screening tool's sensitivity and specificity.	

This page titled 1.9.1: Preparation N.9 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.9.2: Exercise N.9

(1) In December 2020, the FDA published a report that included the accuracy of the Ellume COVID-19 Home Test. The article reported that the test, "correctly identified 96 percent of positive samples and 100% of negative samples in individuals with symptoms."[30]

(a) What do the percentages in the paragraph tell us about the specificity and sensitivity of the test?

(b) Suppose that this test was administered by a group of 10,000 people, of which 1% are infected with COVID-19. Fill in the following table in order to gather data on the test's effectiveness.

Hint: How many people in total have COVID-19? Where does that go in the table? How many of those would test positive?

	Number of COVID-19 Positive Individuals	**Number of COVID-19 Negative Individuals**	**Total**
Result of Ellume COVID-19 Test is Positive			
Result of Ellume COVID-19 Test is Negative			
Total			

(c) How many individuals would receive a positive test result even though they are not infected with COVID-19? Based on the above table, if someone is not infected with COVID-19, what is the probability that they will receive a positive test result?

(d) How many of the negative test results would be received by people who were infected with COVID-19? Based on the above table, if someone receives a negative result from this new test, what is the probability that the person is actually infected with COVID-19? Round to the nearest hundredth of a percentage point.

(e) Suppose that this same test was administered to 10,000 people in another country in which 10% of the population is infected with COVID-19. Fill in the following table:

	Number of COVID-19 Positive Individuals	**Number of COVID-19 Negative Individuals**
Result of Ellume COVID-19 Test is Positive		
Result of Ellume COVID-19 Test is Negative		

(f) In this population, if someone tests positive with this new test, what is the probability that the person is not infected with COVID-19? Round to the nearest hundredth of a percentage point.

(g) In this population, if someone tests negative with this new test, what is the probability that the person is infected with COVID-19? Round to the nearest hundredth of a percentage point.

(2) Suppose an even newer test for COVID-19, released in 2023, advertises that it has improved sensitivity compared to the test described in Question 1. The advertisement then states that people should use the newer test because it is more effective. The two-way table below shows the data on which the advertisement's claims were based.

	Number of COVID-19 Positive Individuals	**Number of COVID-19 Negative Individuals**
Result of COVID-19 Test is Positive	49	40
Result of COVID-19 Test is Negative	1	19910

Calculate the sensitivity and specificity of the newer test. Round to the nearest tenth of a percentage point.

(3) Consider the following situation presented in a *New York Times* blog: "Before going on vacation for a week, you ask your spacey friend to water your ailing plant. Without water, the plant has a 90 percent chance of dying. Even with proper watering, it has a 20 percent chance of dying. And the probability that your friend will forget to water it is 30 percent."[31] We can model this situation by using a hypothetical 100 ailing plants and entering what would happen into a two-way table like the one below:

	Watered	**Not Watered**	**Total**
Number of Plants that Died			
Number of Plants that Did Not Die			
Total			100

(a) Fill in the table using the percentages in the scenario. (**Hint**: Start with column totals: watered and not watered.)

(b) Now you can answer the questions the *New York Times* blogger asks. Round to the nearest percentage point.

(i) What's the chance that your plant will survive the week?

(ii) If the plant is dead when you return, what's the chance that your friend forgot to water it?

(iii) If your friend forgot to water the plant, what's the chance it'll be dead when you return?

---

[30] https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-antigen-test-first-over-counter-fully-home-diagnostic

[31] http://opinionator.blogs.nytimes.com/2010/04/25/chances-are/

This page titled 1.9.2: Exercise N.9 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

1.10: Stress Reappraisal

(to be completed before the Module 1 exam, as instructed)

As you might expect, taking a math test can be a very stressful experience. Before taking the Module 1 exam, we are going to complete a short reading exercise designed to help you perform well.

In the following pages, you will be presented with scientific research. Because scientific research articles can be long, complex, and difficult to understand we had students like you summarize the research articles discussed below.

The research presented in the following materials was conducted with people with anxiety disorders to help them cope with feelings of overwhelming anxiety. More specifically, materials provide you with information about how stress can help improve performance.

While you read each summary, please try to think about how stress can help you. Then, after each one, you will be asked to write a short summary using your own words describing how the information presented can help maximize your performance on the Module 1 exam. The information you write here will be included in messages to future Quantway and Statway students designed to provide students with the confidence to thrive in testing situations.

The following summary is of Jamieson and Mendes' 2010 study that appeared in the *Journal of Experimental Psychology*:

> In stressful situations, people experience changes in their body. They might experience these changes as "unsettled feelings" or "butterflies in their stomach," and conclude that they are nervous. However, the bodily changes that happen during stress can be good, not just bad. For instance, scientists have found that feelings of "butterflies" or nervousness indicate that the body is gathering resources to meet the demands of the situation. In other words, the body needs energy to perform and stress helps deliver this energy to where it is needed.
>
> Stress can be "good" or "bad," and the type of stress we experience depends on our perceptions and beliefs. To demonstrate how the body and mind work together, imagine you are a skier staring down a steep, icy slope with no other way off the mountain than going down this dangerous trail. Regardless of whether you like skiing, this situation would cause stress. Expert skiers may experience the stress as "excitement" if they believe they can handle the difficult trail, whereas novices are more likely to experience the stress as "fear" if the difficulty of the trail exceeds their skill level. Thus, the skier's response (excitement vs. fear) depends on how they perceive the situation and their body's responses.
>
> The negative consequences of stress are the direct result of people's beliefs that feeling anxious will make them perform poorly. However, our research from people with anxiety disorders indicates that **stress does not hurt performance, but can actually help** because our brain releases chemicals that help us think quickly. So, during the Module 1 exam, try and view stress as a coping tool.

In your own words please briefly describe how this information can help you perform well on the Module 1 exam:

_____

_____

_____

_____

The following summary is of Nock et al.'s (2011) study that appeared in the *Journal of Clinical Psychology*:

> Stress is a normal physical reaction that helps you face the challenges in your life. It is not harmful. In fact, if we did not have stress reactions we could not survive. If stress is helpful, then why do most people see it as a negative experience?
>
> Research from patients coping with anxiety disorders indicates that negative reactions to stressful situations like taking an exam in class or going on a job interview are the result of how we think about stress (also known as "cognitive appraisals"). When the "fight or flight" system activates, our brain searches for potential sources of harm. However, in modern society there is often no physical threat. When no explanation can be found, the search turns inwards and the brain invents explanations such as, "There must be something wrong with me." Nothing could be further from the truth.
>
> During stressful situations remember that your body's responses are good. Increased heart rate, sweating, and heavy breathing all help deliver oxygen where it is needed. Remember, if you find yourself experiencing stress during tests, focus on the benefits of your body's responses.

In your own words please briefly describe how this information can help you perform well on the Module 1 exam:

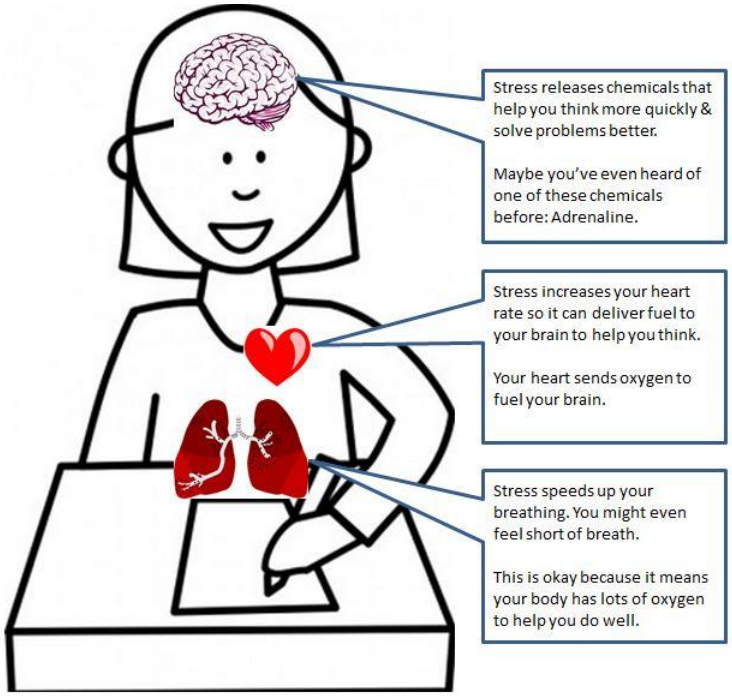
_____

_____

_____

_____

The following is an illustrative diagram that shows the biological changes that happen when we experience stress. Please take a minute to note where the changes occur and how these help us do well.



In your own words please briefly describe how this information can help you perform well on the Module 1 exam:

_____

_____

_____

_____

Great job! You have finished the reading exercise.

Remember during the Module 1 exam, we ask that you try to remind yourself that your body's responses to the stressful testing situation will help you to perform well.

Good luck!

This page titled 1.10: Stress Reappraisal is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

2: College Module M

2.1: Mathematical Models–What's Best?
2.1.1: Preparation M.1
2.1.2: Exercise M.1
2.2: Linear Models
2.2.1: Preparation M.2
2.2.2: Exercise M.2
2.3: Data Is Trendy
2.3.1: Preparation M.3
2.3.2: Exercise M.3
2.4: Four Equations…Two Models
2.4.1: Preparation M.4
2.4.2: Exercise M.4
2.5: Modeling a New Type of Growth
2.5.1: Preparation M.5
2.5.2: Exercise M.5
2.6: How long Does That Take?
2.6.1: Preparation M.6
2.6.2: Exercise M.6
2.7: A Trust in Social Security
2.7.1: Preparation M.7
2.7.2: Exercise M.7
2.8: Multiple Variables Weighing on Your mind?
2.8.1: Preparation M.8
2.8.2: Exercise M.8

This page titled 2: College Module M is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.1: Mathematical Models–What's Best?

INTRODUCTION

Media outlets regularly publish rankings of cell phone carriers, travel destinations, colleges, TVs and credit cards, just to mention a few.

Here's an example of one such ranking – the "Best Countries for Retirement," according to investopedia.com. To determine which countries are the best for retirees, we use a scoring system that measures a variety of factors, including:

- Ease of buying and owning property
- Cost of renting
- Benefits and discounts on things such as healthcare and entertainment
- Visa and residency requirements
- Healthy living
- Development and infrastructure
- Climate
- Stability of the country's political situation
- Cost of living

The top 5 countries with the highest cumulative average score across all those categories are Panama, Costa Rica, Mexico, Ecuador, Malaysia.

When creating rankings like the one above, we first need to decide what factors (variables) will contribute to the overall ranking. Then, we need to decide how those factors can be measured (quantified). Finally, we need a formula that combines all of them into one overall score.

In your group, agree on a topic that interests everyone (or most of you!), and try to come up with a strategy to create a ranking system related to that topic. For example, you could think about how you'd rank the restaurants in your area.

After agreeing on a topic, think about:

- What variables to use;
- How to measure those variables; and
- How to combine all of them into one overall score.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- quantifying variables is sometimes difficult and there may be multiple methods to do so.
- models can be used to help make real-life decisions, but it is likely that they do not take into account every possible variable.
- mathematical models provide a method to estimate values that are difficult to measure directly.
- mathematical models can be equations, graphs, tables, or described in words.

By the end of this collaboration, you should be able to

- define what is meant by the term "variable."
- describe what a mathematical model is.
- identify possible variables which may be used to model a given scenario.
- use a given mathematical model to make comparisons.
- use a multivariable mathematical model to estimate a value.

PROBLEM SITUATION 1: THE BEST COLLEGES

You read about the yearly *U.S. News* ranking of best colleges in Preparation M.1. Today, we are going to think more about this ranking system and then review another ranking system for the "best" community colleges. The media commonly uses "best" and "worst" rankings when they report on colleges. Often, these rankings are used to make decisions about what colleges are better than others.

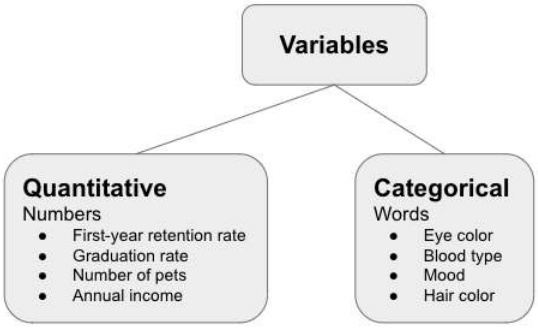
(1) Using what you know about how *U.S. News* ranks colleges, discuss what things you might include in your decision in choosing a college. That is, what factors would be important to your decision?

Important factors for a decision or situation are called **variables**. You can remember this term because variables are the things that vary, or are different, across the people or things we are measuring. For example, first-year student retention and graduation rates

are two factors that vary from school to school. These factors are also important in determining how good a school is. These factors are variables used in the *U.S. News* ranking system.

Variables, like *first-year student retention* or *graduation rates*, often have values that are numbers. These are called **quantitative variables**. Numbers can allow for more meaningful comparisons. For example, if the variable is *first-year student retention* for each college, we can use a number to give the rate at each college, such as the percentage of first-year students that go on to the second year of college. Then we can easily compare a college with an 80% retention rate to a college with a 40% retention rate. The comparison would not be as meaningful if only descriptive words had been used, such as "high" retention rate for the first college and "low" retention rate for the second.

When variables sort data into a limited number of values [such as a screening test being either positive or negative, or sizes of small, medium, and large), they are called **categorical variables**.



It can be difficult to decide what numeric value to use for certain variables. For example, *U.S. News* also uses a variable called *faculty resources*. This could be measured by average faculty salary (in dollars), money available for classroom equipment (also in dollars), or the number of hours of support faculty get from a mentor, or a combination of multiple things.

(2) Refer back to the list of variables (factors) you discussed in Question 1, and brainstorm how numeric values could be used to quantify them (express with numbers). Include the units of measurement.

A **mathematical model** is a description of a system, or situation, using mathematical concepts, language, and numbers. People often use variables to help create mathematical models; however, variables are not always needed for models. Like functions, mathematical models can be equations, tables, graphs, or descriptions in words.

A mathematical model may be created to make predictions about a situation or simply represent and explain a situation and its different components. Mathematical models are often not perfectly accurate, but they are useful ways to measure values. Many models have been shown to produce values that closely match real data. For example, meteorologists use information on wind speed, temperature, and humidity to predict the chance of rain. While meteorologists are not perfectly accurate, their models produce far more accurate predictions than guessing.

Most decisions using models are based on multiple variables. One way to compare options is to combine the models' variables in a meaningful way. The combining of variables (usually by adding, subtracting, multiplying, dividing, or averaging) into one number is called building a **mathematical model equation**.

Now, we'll study another ranking system, this time for community colleges. We will look at a model that combines the variables in a way that gives us one number for each community college. The resulting composite score will allow us to first compare and then rank a set of community colleges.

The Aspen Institute awards the Aspen Prize every other year. The award "honors those institutions that strive for and achieve exceptional levels of success for all students, while they are in college and after they graduate."[3] In addition to the prestige of winning this award, the winning college also receives a $1 million cash prize. To determine the 150 finalists for the award, the Aspen Institute collects information from 1000 community colleges related to the following variables (remember that we call them variables because their values may vary between colleges):

1. **Performance** – The **performance** score measures student success in persistence, degrees awarded, completion, and transfer;

2. **Change** – The **change** score measures consistent improvement in these areas over time; and

3. **Minority Achievement and Family Income** – The **minority achievement** score and **family income** score measure equitable outcomes for students of all racial/ethnic and socioeconomic backgrounds.

The values of these variables are given in the table below for four community colleges. Note that the table itself is a model for this situation.

College	State	Performance Score	Change Score	Minority Achievement Score	Family Income Score
Bergen Community College	NJ	299	800	178	70
Miami Dade College	FL	323	400	288	191
Northwest Iowa Community College	IA	503	200	437	99
Pearl River Community College	MS	444	600	275	118

Each of these scores is created from combining a great deal of information about many factors. For example, the **performance** score is created by using both the *first-year student retention* rate and the school's *graduation* rate. There is no easy way to guess how these scores were calculated. In order to find out, you would need to look carefully at the details of the research plan. One must be careful when combining numbers with different units.

Later in this module, you will look specifically at how to build mathematical models. For this collaboration, you should concentrate on understanding the following:

- the model that is given
- what the model is used for
- what the variables in the model measure

(3) Look at the scores in the table above. Which school do you think is the "best?" Why? (Hint: What data are you focusing on in the table?)

(4) The Aspen institute used the following mathematical model equation to calculate a score for each school. In this case, the model is represented by this equation:

R = 0.33*P + 0.33*C + 0.27*M + 0.07*F

where:

R = Overall ranking score (higher scores are best)

P = Performance score

C = Change score

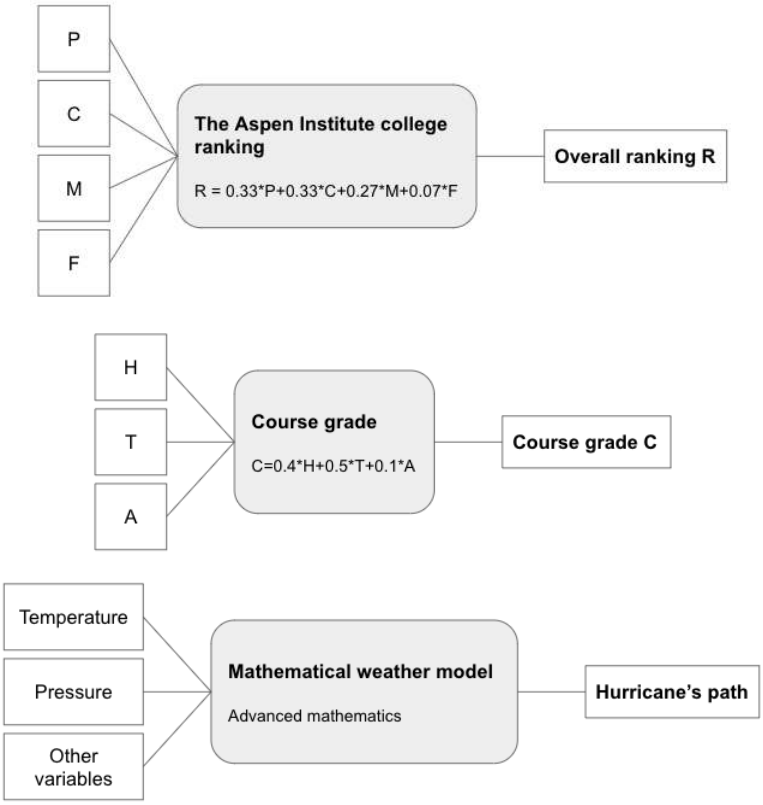
M = Minority score

F = Family income score

Use this equation to calculate the overall score for each of the schools and enter it into the table below. According to the overall score you calculated, which school is best?

College	R
Bergen Community College	(i)
Miami Dade College	(ii)
Northwest Iowa Community College	(iii)
Pearl River Community College	(iv)

Here are visual representations of the models that we've discussed so far in this collaboration:



PROBLEM SITUATION 2: WEIGHING THE MODELS

Throughout this module, we will investigate several mathematical models. Some build on the linear and exponential models that you may have seen in previous courses, while others are completely different types of models, such as polynomial models or trigonometric models. We next turn our attention to a mathematical model that estimates body size.

(5) According to the RAND study described in the Science Daily article[4] that you read in the preparation, the proportion of Americans who are severely obese continues to increase. The study used Body Mass Index, or BMI, to classify individuals as obese. We will now consider the formula for BMI, as well as its relationship to other measurements of a person's size.

(a) What variables are used to calculate BMI?

(b) How do you think the researchers in the RAND study obtained information about these variables? Imagine yourselves as the researchers. You will want a large, varied sample, and you have limited funds. What would be a good way to get people's height and weight from diverse U.S. locations?

(c) If you want to determine whether someone is obese, are there any other variables that might be important to consider?

(6) The article defines Body Mass Index as weight in kilograms divided by the square of height in meters. In terms of a mathematical model equation, we can write this as

$$BMI = \frac{weight\ (in\ kilograms)}{[height\ (in\ meters)]^2}$$

The height and weight of athletes is often publicly available. The table below lists known heights and weights of several current and former (at the time their height and weight were measured) athletes. Calculate the BMI for each of the athletes. Use the fact that one inch is 0.0254 meters and one pound is about 0.454 kilograms.

(a) Discuss in your group before completing the entries for Rory McIlroy first. Round answers to two decimal places.

(b) Complete the rest of the table for the other athletes. Round answers to two decimal places.

Athlete	Height		Weight		Body Mass Index
	Feet (') Inches (")	Meters (m)	Pounds (lbs)	Kilograms (kg)	kg/m^2
Rory McIlroy (Golf)	5' 10"		160 lbs		
Julianna Pena (UFC)	5' 7"		135 lbs		
Cam Gallagher (Baseball)	6' 3"		230 lbs		
Breanna Stewart (WNBA)	6' 4"		170 lbs		
Shaquille O'Neal (NBA)	7' 1"		325 lbs		
Maria Sharapova (Tennis)	6' 2"		130 lbs		

(7) The Centers for Disease Control and Prevention (commonly referred to as the CDC) suggests that

- BMI below 18.5 is Underweight;
- BMI between 18.5 and 24.9 is Normal;
- BMI between 25 and 29.9 is Overweight;
- and BMI of 30 or above is Obese.

What are the athletes' classifications based on their BMIs? Write "Underweight", "Normal", "Overweight", or "Obese", based on your calculated BMI.

Rory McIlroy:

Julianna Pena:

Cam Gallagher:

Breanna Stewart:

Shaquille O'Neal:

Maria Sharapova:

(8) Recall the description of the BMI calculation: "weight in kilograms divided by the square of height in meters." Let's examine this mathematical model more closely.

(a) Consider the following two scenarios. In each scenario, which person has the lower BMI?

(i) Julie and Sarah have the same weight, but Julie is taller. Which of them has a lower BMI?

(ii) Mike and Jim have the same height, but Mike weighs more. Which of them has a lower BMI?

(b) If you wanted to lower your BMI, which variable could you reasonably change?

(c) What are some limitations of the BMI model?

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

[3] http://highered.aspeninstitute.org/projects/

[4] http://www.sciencedaily.com/releases/2012/10/121001132146.htm

This page titled 2.1: Mathematical Models–What's Best? is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.1.1: Preparation M.1

Every year, U.S. News & World Report publishes a list of the best colleges in the United States. Colleges on the list are ranked based on many different factors. Below is an excerpt from an article introducing the 2022 U.S. News "Best Colleges" rankings.[1]

Use the information in the following three paragraphs to help you answer the questions below.

The host of intangibles that makes up the college experience can't be measured by a series of data points. But for families concerned with finding the best academic value for their money, the U.S. News Best Colleges rankings provide an excellent starting point for the search.

They allow you to compare at a glance the relative quality of institutions based on such widely accepted indicators of excellence as first-year student retention and graduation rates and the strength of the faculty. And as you check out the data for colleges already on your short list, you may discover unfamiliar schools with similar metrics, and thus broaden your options.

Many factors other than those spotlighted here will figure in your decision, including location and the feel of campus life; the range of academic offerings, activities and sports; and cost and the availability of financial aid. But if you combine the information on usnews.com with college visits, interviews and your own intuition, the U.S. News rankings can be a powerful tool in your quest for the right college.

(1) Why does U.S. News say that the rankings of colleges are important?

(i) The rankings allow individuals to choose the best college for themselves because everything an individual would need to know is included in the ranking.

(ii) The rankings allow the top colleges to charge more for tuition because the rankings show that they are the best.

(iii) Although the rankings cannot capture everything, the rankings allow people to compare colleges based on some common measures.

(2) Read the news article, "Severely Obese Are Fastest Growing Group of Overweight Americans, Study Finds," from Science Daily, available online at: http://www.sciencedaily.com/releases/2012/10/121001132146.htm.[2] What measure did the study use to determine whether an individual was obese?

(3) The Science Daily article defines Body Mass Index (BMI) as "weight in kilograms divided by the square of height in meters." In order to use this definition, one must be able to convert pounds to kilograms and inches to meters.

(a) If Andrew is 6 feet and 2 inches tall, what is his height in inches?

(b) If one inch is 0.0254 meters, what is Andrews's height in meters? Round to the nearest one hundredth of a meter.

(c) Andrew weighs 205 pounds. If one pound is about 0.454 kilograms, what is his weight in kilograms? Round to the nearest one hundredth of a kilogram.

(4) In dealing with geometric figures, we are often asked to compute the perimeter and the area of these figures. Recall that the perimeter of a figure is the distance around the figure and the area of the figure is the amount of surface the figure takes up.

(a) Complete the table below by writing the formula for the perimeter and area of each indicated figure. If you are not familiar with these formulas, you may need to do a quick internet search.

Figure	Perimeter	Area
Rectangle		
Triangle		
Circle		

(b) A rectangle has a length of 3.4 inches and a height of 4.5 inches. Using your formulas from Question 4(a), find the perimeter and area of the rectangle. Include units in your answer.

(c) A triangle has a base of 8 feet and an area of 51 square feet. Using your formula from Question 4(a), find the height of the triangle. Include units in your answer.

(d) The circumference of a circle is found to be 157 cm. Using your formula from Question 4(a), find the radius and the area of the circle. (Hint: Use the approximation pi = 3.14.)

After Preparation M.1 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration M.1, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	Rating from 1 to 5
substitute values into an equation or formula and simplify the result.	
perform basic arithmetic operations including understanding the order of operations in a complex equation.	
make unit conversions, specifically from inches to meters and from pounds to kilograms.	
read news articles with complex formulas.	

[1] http://www.usnews.com/education/best-colleges/articles/how-us-news-calculated-the-rankings

[2] http://www.sciencedaily.com/releases/2012/10/121001132146.htm

This page titled 2.1.1: Preparation M.1 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.1.2: Exercise M.1

(1) Jenna is using her car for work and created a model for the cost, per mile, of driving her car. It is given by the equation:

$$J = \frac{g}{22} + 0.0536$$

where J = Cost, per mile, of driving Jenna's car (in dollars)

g = Cost of gas (dollars/gallon)

(a) If the cost of gas is $3.92 per gallon, what is the cost, per mile, of driving Jenna's car? Round to the nearest cent (e.g., $0.87).

(b) Is g a variable? Explain your reasoning.

(c) What other variables (in addition to cost of gas) could have been included for this context?

(d) Let's think about one variable: the cost of insurance. Describe one way to measure that variable.

(2) In baseball statistics, slugging percentage (abbreviated SLG) is a popular measure of the power of a batter. It is calculated as total bases earned divided by at-bats:

$$SLG = \frac{S + 2 * D + 3 * T + 4 * HR}{AB}$$

where AB is the number of at-bats for a given player, and S, D, T, and HR are the number of singles, doubles, triples, and home runs, respectively. (Walks are specifically excluded from this model and do not count as at-bats.)

(a) What are the units for slugging percentage?

(b) The table below shows statistics for the New York Yankees baseball player Aaron Judge.[5] Fill in Judge's missing SLG for the years 2018, 2020, and 2022. Round your answers to three decimal places. In which year did Judge have the most power?

Year	AB	S	2B	3B	HR	SLG
2017	542	75	24	3	52	0.627
2018	413	66	22	0	27	
2019	378	57	18	1	27	0.540
2020	101	14	3	0	9	
2021	550	95	24	0	39	0.544
2022	570	87	28	0	62	

(c) List two ways a player could increase his or her power rating.

(3) In Collaboration M.1, we used BMI to determine whether athletes were obese. The BMI calculation is based on the relationship between weight and height. There are other formulas that try to account for additional variables when estimating body fat. Deurenberg and Weststrate[6] developed the following formula for estimating body fat percentages:

- Men's Adult Body Fat Percentage = 1.2*BMI + 0.23*Age – 16.2
- Women's Adult Body Fat Percentage = 1.2*BMI + 0.23*Age – 5.4

The table below provides the names and ages (at the time their height and weight were measured) of the athletes you classified in class.

Athlete	Age	BMI (from Question 6)	BMI Classification (from Question 7)	Estimated Body Fat Percentage (BF%)	BF% Classification
Rory McIlroy	27				
Julianna Pena	33				

Cam Gallagher	30				
Breanna Stewart	28				
Shaquille O'Neal	36				
Maria Sharapova	29				

Fill in the BMI and BMI classifications from Questions 6 and 7. Then complete the next two questions.

1. Use Deurenberg and Weststrate's formula to estimate each athlete's body fat percentage at the age given. Enter your answers in the "Estimated Body Fat Percentage" column.

(b) The American Council on Exercise uses the following classification system based on Body Fat Percentage:[7]

Description	**Women**	**Men**
Essential Fat	10–13.99%	2–5.99%
Athletes	14–20.99%	6–13.99%
Fitness	21–24.99%	14–17.99%
Average	25–31.99%	18–24.99%
Obese	32% +	25% +

Find the athletes' classifications based on their estimated body fat percentage and enter the information in the "BF% Classification" column of the table above. (Note: In the table above, use the middle column for Julianna Pena, Breanna Stewart, and Maria Sharapova. Use the third column for Rory McIlroy, Cam Gallagher, and Shaquille O'Neal.)

(c) Compare the classification based on BMI for Breanna Stewart to the classification based on body fat percentage. Do the classifications match? What would it mean if the classification did not match?

(4) Do you think that BMI is a reasonable estimate of body fat composition? Why or why not? Does BMI help classify people as "overweight?" Is it useful for classifying all kinds of people, at all weight levels? What good does using the BMI formula do? What harm does using the BMI formula do?

(5) Recall Deurenberg and Weststrate's formula for estimating body fat percentages:

- Men's Adult Body Fat % = 1.2*BMI + 0.23*Age – 16.2
- Women's Adult Body Fat % = 1.2*BMI + 0.23*Age – 5.4

Using these formulas, consider how the two different mathematical models work in the following scenarios.

(a) If Pedro keeps his BMI constant over time, what will happen to his estimated body fat percentage as he ages?

(i) The body fat percentage will increase.

(ii) The body fat percentage will decrease.

(iii) The body fat percentage will stay the same.

(b) Aimee and Christopher are twin sister and brother, respectively. If they have the same height and weight, who will have a greater estimated body fat percentage?

(i) Aimee.

(ii) Christopher.

(iii) They will have the same body fat percentage.

(c) This year, Nisheidi increased her BMI by 1 from last year. She also got one year older. Which increase, age or BMI, will have a greater impact on her estimated body fat percentage?

(i) Increased age by 1 year.

(ii) Increased BMI by 1 point.

(iii) They would have the same impact.

(6) Many pediatric dosages of drugs are determined by using the child's body surface area. The Mosteller formula is often used for this purpose.[8] To determine a child's body surface area (in square meters), the child's height (in inches) and weight (in pounds) are multiplied together; this number is divided by 3131 and then a square root is taken of the result. Thus

$$A = \sqrt{\frac{h * w}{3131}}$$

where:

A = the child's body surface area (in square meters)

h = the child's height (in inches)

w = the child's weight (in pounds)

(a) Use this formula to find the surface area of a child who is 4 feet tall and weighs 75 pounds. Round to the nearest hundredth of a square meter. (Note: You will need to first convert the height into inches.)

(b) To get the dosage of an allergy medicine for a child, the following formula is used:

$$D = \frac{25 * A}{1.73} mg$$

where A is the body surface area, in square meters. Calculate the dosage of this allergy medicine for this child. Round to the nearest tenth of a milligram.

(c) One practitioner forgot to convert the feet to inches and uses "4" in the equation for height. What surface area did this practitioner calculate? Round to the nearest hundredth of a square meter.

(d) What was the resulting dose from making this error? Round to the nearest tenth of a milligram.

(e) What do you think would be the result of the miscalculated prescription?

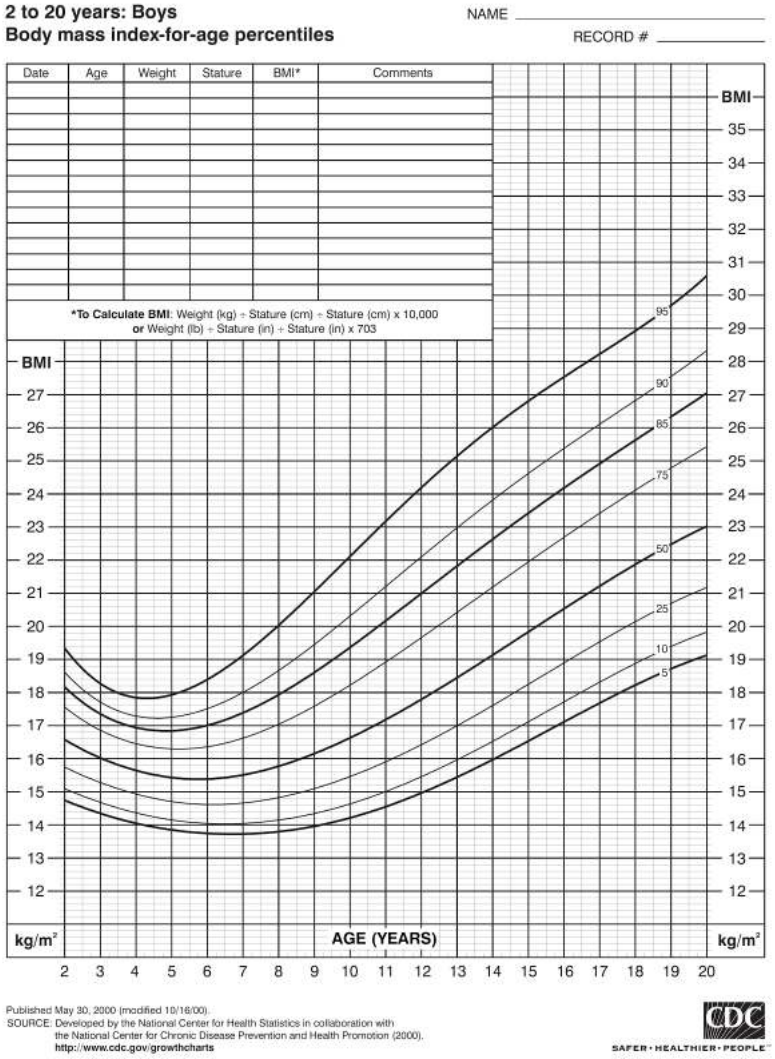
(7) In Collaboration M.1, you saw that mathematical models can be represented by equations containing variables. Equations are just one of the many types of representations that can be used to describe a mathematical model. In addition to equations, many mathematical models are represented by tables or graphs. One such graphical model is a pediatric growth chart.

Pediatric growth charts have been used by pediatricians, nurses, and parents to track the growth of infants, children, and adolescents in the United States since 1977. These growth charts consist of a series of percentile curves that illustrate the distribution of various measurements in relation to age and gender, including height, weight, head circumference, and Body Mass Index.

One of the newer charts is the Body Mass Index-For-Age (BMI-For-Age) chart for boys ages 2 to 20 years, shown in the figure below. This chart can be used to determine the percentile rank for a child, given the age and the BMI of the child.

For example, Enrique is 8 years old and he has a BMI of 17.5. According to the BMI-For-Age chart below, Enrique falls into the 75th percentile (less than the 85th percentile). This means that 75% of all 8-year-old boys have a BMI that is less than Enrique's, or only 25% of all 8-year-old boys have a BMI that is greater than Enrique's. Look at the chart, and make sure you understand how it

was used to determine that Enrique falls in the 75th percentile, given his age and BMI.



Use the chart to answer the following questions:

(a) Alex is a 12-year-old boy who weighs 120 pounds and measures 4 feet 11 inches tall. Calculate Alex's BMI using the formula from the in-class lesson

$$BMI = \frac{weight\ (in\ kilograms)}{[height\ (in\ meters)]^2}$$

Then use the chart above to determine his percentile rank.

(b) Joshua is 16 years old and falls into the 25th percentile. What is his approximate BMI?

(c) A boy is tracked from age 6 until age 15 using this BMI-For-Age growth chart. Suppose that throughout this entire time, he has always been in the 85th percentile rank. What happens to his BMI as he ages from 6 to 15?

(d) A boy is tracked from age 6 until age 15 using this BMI-For-Age growth chart. Suppose that throughout this entire time, his BMI has stayed at 17 kg/m2. What happens to his percentile rank as he ages from 6 to 15?

---

[5] https://www.baseball-reference.com/players/j/judgeaa01.shtml

[6] http://www.halls.md/bmi/fat.htm

[7] http://www.acefitness.org/blog/112/what-are-the-guidelines-for-percentage-of-body-fat/

[8] R.D. Mosteller, “Simplified Calculation of Body Surface Area,” New England Journal of Medicine 317, no. 17 (1987): 1098.

This page titled 2.1.2: Exercise M.1 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.2: Linear Models

INTRODUCTION

In Collaboration M.1, you were introduced to the idea of a mathematical model. There are many different types of mathematical models. We will first look at a type called a linear model, which is both one of the simplest and most important kinds of models. Any situation that has a "constant rate of change" is a linear model. There are many common examples of linear models, including sales tax, linear depreciation, hourly salary, and the grade (steepness) of a road.

In Preparation M.2, various definitions for linear models were presented. In this collaboration, we will use these definitions, learning about them in more detail. Check-in with your group members to make sure everyone understands the following definitions:

Definition: The **slope of a line** is the ratio of the change in the vertical value (sometimes called output, y-value, dependent variable, or response variable) to the change in the horizontal value (sometimes called input, x-value, independent variable, or explanatory variable) as we move between any two points on the line. The slope of a line is therefore a rate of change. Often it is referred to as the ratio of the rise of the graph to the run of the graph, or "rise over run." The slope can also be described as the rate at which the output values are changing (increasing or decreasing) as the input value increases by 1. It can be found by using any two points on the line. (Note: By convention, the input is on the horizontal axis and the output on the vertical axis.)

Definition: The **vertical intercept** of a line is the point at which the line intersects the vertical axis, where the input value is zero. This is also called the initial value if the input starts out at a value of zero.

Definition: The **horizontal intercept** of a line is the point at which the line intersects the horizontal axis, where the output is zero.

Definition: A **linear model** has a constant rate of change (slope). (This is not true of all models and their corresponding equations.)

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- the slope of a linear model is a rate of change.
- a linear model has a constant rate of change.
- the vertical intercept of a linear model is an initial value.
- linear models can be used to approximate real-life situations, though they are often not exact.
- data that changes in an approximately linear manner can be modeled by a linear regression model (trendline).

By the end of this collaboration, you should be able to

- create a table, graph and/or equation from a verbal description of a linear model.
- move freely between the four different modes of looking at linear models (words, tables, graphs and equations).
- find and interpret the slope of a linear model from the four different representations.
- find and interpret the horizontal and vertical intercepts of a linear model from the four different representations.
- use a spreadsheet to create a scatterplot and linear regression model (trendline) from data.

PROBLEM SITUATION 1: ENERGY BILLS

The company that provides electricity to people in and around New Orleans, Louisiana, is called Entergy. The table below shows an estimate of monthly rates charged by Entergy.[12] Use this table to respond to the questions below.

Summer Billing Cycles (May – October)	**Winter Billing Cycles (November – April)**
Energy Charges: • 6.002¢ per kWh • Monthly Customer Charge: $8.07	Energy Charges: • First 800 kWh per month at 6.002¢ per kWh • Over 800 kWh per month at 4.766¢ per kWh • Monthly Customer Charge: $8.07

For this collaboration, we will only be using the summer billing cycle. Notice that the monthly service charge of $8.07 is applied to each monthly bill, no matter how much electricity is used.

(1) Suppose a customer uses no electricity in July. What would her monthly bill be?

(2) A customer uses 900 kWh in July. What would her monthly bill be?

(3) Another customer uses 901 kWh in July. What would her monthly bill be?

(4) What is the difference between outputs in your answers to Question 2 and Question 3? What is the difference between the inputs you used to answer Questions 2 and 3? Now use these differences to calculate the slope.

(5) Explain why a linear model would represent Entergy's summer billing charges for any customer.

(6) Considering how you calculated your answers to Questions 2 and 3, find a linear equation that models any Entergy customer's bill. Use k for energy usage (in kWh), and B for Entergy's charge (in dollars).

(7) Sketch a graph of your linear model. Note: If completing this problem online, follow the instructions given online to create your graph.

(8) What is the vertical intercept of this model? How is it related to billing? Write your answer as a point (k, B), and then write a sentence for your explanation.

(9) Does this model have a horizontal intercept? What does it mean?

(10) Suppose a family makes a quarterly (3-month) budget to plan for upcoming expenses. This family estimates they will use an average of 1000 kWh per month in July, August and September. How much money must the family set aside in the budget so that they will be able to pay for the electricity they will use during the three-month period? Round to the nearest dollar, since budgets are only estimates.

PROBLEM SITUATION 2: CELL PHONE DEPRECIATION

Smartphones, like many other manufactured goods, depreciate on a daily basis. Depreciation means how much an item (e.g. a smartphone) loses in value compared to the price that was originally paid for it from the manufacturer. There tends to be two main periods in a smartphone's life when it sees the biggest drops in value: 1. the period of time after the phone is initially purchased, and 2. when the manufacturer launches a new model. The decrease in value of an item over time is called depreciation.

In September 2019, the iPhone 11 256GB was released with a price of $849.[13] Twelve months after the launch, it was found that the value of iPhone 11s depreciated about 37% (which is actually better than most smartphone depreciation!).[14]

Suppose you bought a new iPhone 11 256GB in September 2019 for $849. The phone would soon start depreciating in value. The following table shows the approximate value of your iPhone 11, by month from September 2019.

Month	**September 2019**	**October 2019**	**November 2019**	**December 2019**	**January 2020**	**February 2020**
Phone Value in Dollars	849	800	780	750	720	699

(11) It can be easier to work with models when the data are numerical. Modify the first row of the table to change the months into numbers. Compare your answers with the rest of your group and the class.

Phone Value in Dollars	849	800	780	750	720	699

(12) Now, use the table from Question 11 to create a graph of the data. Remember to label the axes. Note: If completing this problem online, follow the instructions given online to create your graph.

(13) Does the data appear to be linear? Explain your answer.

(14) What is the vertical intercept of the graph? What does this mean in relation to the situation?

(15) Answer the next three questions using the assumption that the data is approximately linear.

(a) Approximate the slope of the linear model. (Hint: When data are only approximately linear, be careful when choosing the two points you use to calculate the slope. Often a good strategy is to use the first point and last point since they represent the total change in the data.)

(b) What is the meaning of this slope in this situation?

(c) Roughly how much does the value of the phone change in a two-month period?

(16) Write an equation that gives a linear model for the value of the cell phone from Sept 2019 through February 2020. Use V for the value in dollars, and m for the number of months after Sept 2019, and use the slope from Question 15c.

(17) Suppose you want to sell your phone before it is worth less than $300. Using your model, when is the last month you could sell it and get at least $300?

(18) What is the horizontal intercept of your model? What is the meaning of this in the context of this problem? Is this reasonable?

PROBLEM SITUATION 3: CELL PHONE DEPRECIATION USING TRENDLINES

We just saw that a linear model can be used to model data that is reasonably, but not perfectly, linear. In Problem Situation 2, we used a rough linear model that may not have represented the data most accurately. There are ways to create a linear model that represents the data more carefully. Your instructor will show you how this is done now in a spreadsheet application, such as Microsoft Excel or Google Sheets. We will be using the table created in Question 11:

Months after Sept 2019	**0**	**1**	**2**	**3**	**4**	**5**
Phone Value in Dollars	849	800	780	750	720	699

(19) Create a scatterplot of the data using a spreadsheet application. Does the data look linear?

(20) We will now use the spreadsheet application to calculate the trendline. The trendline is a linear model representing the trend of the data. When using technology, the trendline is often called the line of best fit, which is the line that comes closer to all the data points, on average, than any other possible line, for data that isn't perfectly linear. If possible, display the trendline along with the scatterplot.

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

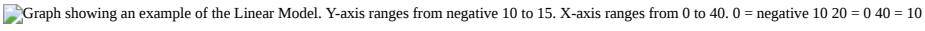
In this collaboration, we have looked at linear data in four different ways. Each way can be useful in different situations.

We have seen **VERBAL DESCRIPTIONS**, such as "Energy Charges 6.002¢ per kWh [and a] Monthly Customer Charge [of] $8.07." Many real-life problems start out as situations that we can describe in words. But we can also try to use alternate representations that make the data more clear, and easier to work with mathematically.

We have also seen **NUMERICAL REPRESENTATIONS** of data, which means the data is presented in a table that allows us to organize and carefully look over the data, such as the following:

Months after Sept 2019	**0**	**1**	**2**	**3**	**4**	**5**
Phone Value in Dollars	849	800	780	750	720	699

We have also seen **GRAPHICAL REPRESENTATIONS** of data, such as the graph given in Question 2 in Preparation M.2:



Notice that when data is presented graphically, it is very easy to see if it appears to be linear, and to read information such as intercepts from the graph.

Finally, we have created **ALGEBRAIC MODELS** of the data, which usually means to give an equation representing the data, such as V = 849 − 30m. Once we have an algebraic model, we can use it to answer questions about the data using the given information and solving for unknown quantities - for example, find the number of months it takes for a $849 cellphone to depreciate to $0 knowing that it depreciates at the rate of $30 per month.

---

[12] https://www.entergy-louisiana.com/your_home/price/

[13] https://www.macrumors.com/2021/07/16/iphone-12-depreciates-less-than-iphone-11/

[14] https://www.decluttr.com/phone-depreciation/

This page titled 2.2: Linear Models is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.2.1: Preparation M.2

In Collaboration M.2, we will be using **linear models**, which are models that can be used when the data lies along (or approximately along) a line. The following definitions are important for linear models.

Definition: The **slope** of a line is the ratio of the change in the vertical value (sometimes called output, y-value, dependent variable, or response variable) to the change in the horizontal value (sometimes called input, x-value, independent variable, or explanatory variable) as we move between any two points on the line. The slope of a line is therefore a **rate of change**. Often it is referred to as the ratio of the rise of the graph to the run of the graph, or "rise over run." The slope can also be described as the rate at which the output values are changing (increasing or decreasing) as the input value increases by 1. It can be found by using any two points on the line. (**Note**: By convention, the input is on the horizontal axis and the output on the vertical axis.)

Definition: The **vertical intercept** of a line is the point at which the line intersects the vertical axis, where the input value is zero. This is also called the initial value if the input starts out at a value of zero. A vertical intercept is always written as a point in the form (0,#).

Definition: The **horizontal intercept** of a line is the point at which the line intersects the horizontal axis, where the output is zero. A horizontal intercept is always written as a point in the form (#,0).

Definition: A linear model has a **constant rate of change** (slope). (This is not true of all models and their corresponding equations.)

Two examples of graphical linear models are shown below:[9]

Figure 1	**Figure 2**

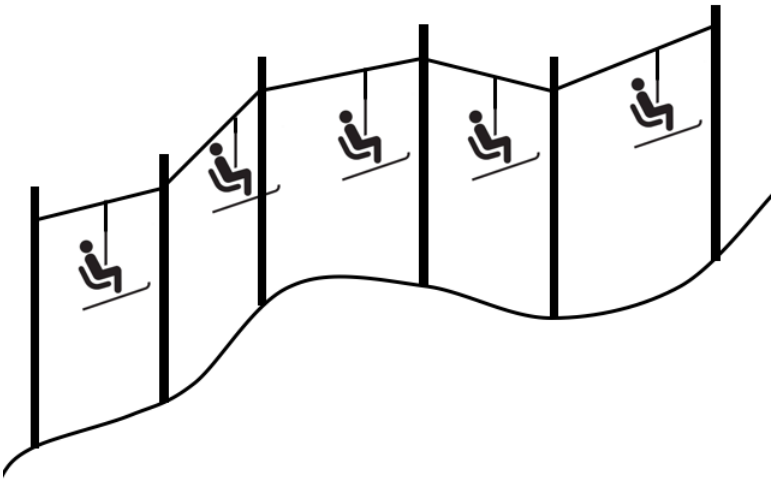
Graph showing an example of the Linear Model. Y-axis ranges from negative 2 to 10. X-axis ranges from negative 7 to 7. Slope (rate of change) = 0.5 Horizontal intercept is shown at (negative 4,0) Vertical intercept is shown at (0,2)

Graph showing an example of the Linear Model. Y-axis ranges from negative 2 to 10. X-axis ranges from negative 7 to 7. Slope (rate of change) = negative 1 Horizontal intercept is shown at (3,0) Vertical intercept is shown at (0,3)

Notice that in Figure 1, the slope is 0.5. This is because every time we move 2 units to the right on the line, the line rises 1 unit, so the rate of change is the ratio 1/2 = 0.5. Alternatively, we could say that every time we move 1 unit to the right, the line rises ½ unit, so the rate of change is again the ratio (½)/1 = 0.5. Notice that we get the same slope either way, since all linear models have a constant rate of change.

In Figure 2, the slope is −1 because every time we move one unit to the right on the line, the line falls 1 unit (a change of −1), so the rate of change is the ratio (−1)/1 = −1.

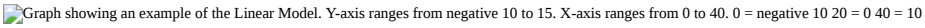
(1) The following picture shows five skiers riding up a mountain on a ski lift.



(a) Which of the skiers is on the steepest portion of the lift?

(b) If you wanted to be certain your answer to Question 1(a) is correct, what would you want to know about each of the line segments representing the lift in the picture? Consider the definitions for linear models that were just presented.

(2) Many old comic books contained ads offering money-making schemes to children reading the comics.[10] As an example, suppose a company offered to send 30 greeting cards to a child to sell to relatives, friends, and neighbors. The child was to sell them for 50 cents each, and had to send $10 back to the company. Here is a graph of a linear model representing the situation. The input (horizontal value) represents the number of cards sold, and the output (vertical value) represents the child's profit in dollars.



(a) What is the vertical intercept? What does it represent in the context of this problem?

(b) What is the horizontal intercept? What does it represent in the context of this problem?

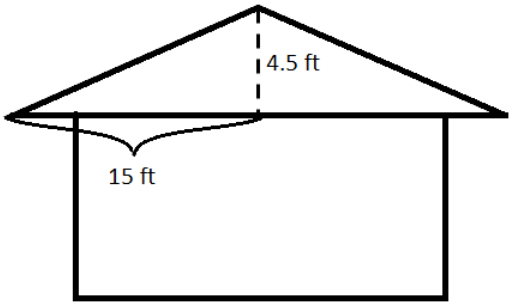
(c) What is the slope? (Hint: Since this line goes through both the horizontal and vertical axes, it would be easy to use both intercepts to find the slope.)

(d) How much money will the child earn if all 30 cards are sold and the agreed upon $10 is sent back to the company?

(3) The City of Minneapolis Construction Code contains the following requirements:

> ROOF SLOPE:
>
> The code requires a roof slope (rise divided by run) of 2:12 or greater for asphalt shingles. However, asphalt roofs sloped between 2:12 and 4:12 will require a double layer of felt paper that goes under the shingles to prevent leakage.
>
> The slope of a roof is usually written in ratio notation. For example, a roof with a slope of 2/12 would be written as 2:12. Note that roofers write slope as 2:12 rather than reducing the answer to 1:6. Why do they leave it that way? It's not because they are bad at math, rather it is because a denominator of 12 is preferred because there are 12 inches in a foot!
>
> For roof slopes of 2:12 up to 4:12, two layers of approved felt are required underneath. For roof slopes of 4:12 or greater, one layer of approved felt is required underneath.[11]

Suppose Joe lives in Minneapolis. He measures his roof, and from the gutter to the peak, it runs 15 feet horizontally and rises 4.5 feet vertically.



(a) What is the slope of Joe's roof?

(b) How many layers of felt must Joe put on his roof to meet code? Select one answer.

(i) He is not allowed to use asphalt shingles with felt.

(ii) He must use one layer of felt.

(iii) He must use two layers of felt.

(iv) His slope can't be compared to the ones in the code, so it can't be determined.

(c) Why do roofs with a lesser slope require greater protection?

After Preparation M.2 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration M.2, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	Rating from 1 to 5
read an input/output table.	
plot points from an input/output table.	
find slopes and intercepts of linear models.	

[9] Graphs created at http://www.meta-calculator.com/online/

[10] http://mentalfloss.com/article/57958/6-get-rich-quick-schemes-vintage-comic-books

[11] http://www.minneapolismn.gov/www/groups/public/@regservices/documents/webcontent/convert_272400.pdf

This page titled 2.2.1: Preparation M.2 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.2.2: Exercise M.2

(1) Just like smartphones, most car values decrease over time. For example, in 2015, a new Toyota Prius Two sold for $25,500.[15] In 2023, online car valuation sites like the Kelley Blue Book and Car Gurus list used Prius Twos for $18,500. The value of the 2015 Toyota Prius Two depreciated over time. Use this information to answer the following problems.

(a) Assume the depreciation is linear. What is the rate of change of a linear model that represents the value of the Prius over time?

(b) Create a linear model of the value of the Prius, where V is the value, and t is the number of years after 2015.

(c) What is the vertical intercept of the graph of the model? What does it represent in the context of this problem?

(d) Use the model to estimate the value of the Prius in 2026.

(2) You have just discovered that your bathtub faucet has a leak. You don't have the funds to get this fixed right away, so you decide to use water that leaks out to water your houseplants.

(a) Assume that you left a large measuring cup under the leaking faucet from 10:00 p.m. until 6:00 a.m. the next morning. In the morning, you find the measuring cup filled to the 3-cup mark. Now assume you put a large pot under the faucet. Fill in the table below to show the total amount of water you would find in the pot if you took measurements every two hours over a 24-hour period.

Hours	Amount of Water (Cups)
0	
2	
4	
6	
8	
10	
12	
14	
16	
18	
20	
22	
24	

On a separate piece of paper, or using a spreadsheet application, make a graph of the amount of water that leaked from the faucet over time by plotting these points. What do you notice about the graph? Note: If completing this problem online, follow the instructions given online to create your graph.

(b) What is the significance of the vertical and horizontal intercepts? Explain the meaning in the context of this situation.

(c) If you used a 1-gallon watering can, how long would it take to fill up? Round to the nearest hundredth of an hour. (Hint: You may need to look up how many cups are in a gallon.)

(d) What is the rate of change of the corresponding linear model? In other words, how fast is the water leaking from the faucet?

(e) Write an equation modeling the quantity of water that has leaked from the faucet over time. Be sure to define the variables you use, including units.

(f) Using your model, how much water would leak out in 11 hours and 15 minutes? Round your answer to the nearest hundredth of a cup.

(3) Obesity among adults in the United States is increasing. We want to investigate how the percentage of U.S. adults who are obese is growing. The table below shows the percentage of U.S. adults who were obese (based on BMI) for each year given.[16]

Year	Percentage of U.S. Adults Who are Considered Obese
2011	34.0
2012	34.7
2013	35.4
2014	36.0
2015	36.7
2016	37.3

When we examine the data above, we see that the percentage of adults who are obese generally increases from year to year. However, the amount of the increase is not exactly the same every year. We will treat the data as linear for this problem, even though the data is only roughly linear. Use a spreadsheet to answer the following questions.

(a) What is the trendline for the data? (Hint: First create a scatterplot.) Adjust the year to be the number of years after 2011.

(b) What is the slope of the linear model (trendline)? What does the slope mean?

(c) Using the trendline, predict the percentage of U.S. adults that will be considered obese in 2029. Round to the nearest percent.

(d) Using your model, what percentage of U.S. adults will be considered obese in the year 2150? Is this reliable?

---

[15] https://www.kbb.com/toyota/prius/2015/

[16] https://ourworldindata.org/grapher/share-of-adults-defined-as-obese?tab=chart&country=~USA

This page titled 2.2.2: Exercise M.2 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.3: Data Is Trendy

INTRODUCTION

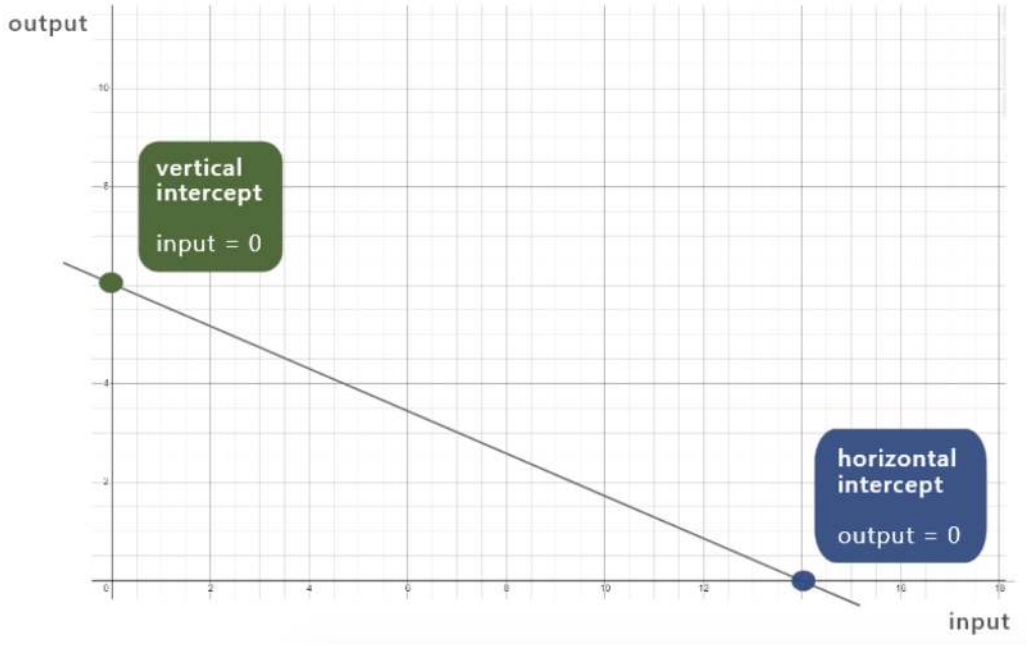
You are going to look at data that is nearly linear and approximate with a linear equation. From there, you'll be able to make predictions about what might happen in the future. This is a very common way to try to predict future events.

It is very important that you understand the concepts of slope and intercepts discussed in Preparation M.2. Please see the definitions below for a quick review:

Definition: The slope of a line is the ratio of the change in the vertical value (sometimes called output, y-value, dependent variable, or response variable) to the change in the horizontal value (sometimes called input, x-value, independent variable, or explanatory variable) as we move between any two points on the line. The slope of a line is therefore a rate of change. Often it is referred to as the ratio of the rise of the graph to the run of the graph, or "rise over run." The slope can also be described as the rate at which the output values are changing (increasing or decreasing) as the input value increases by 1. It can be found by using any two points on the line. (Note: By convention, the input is on the horizontal axis and the output on the vertical axis.)

Definition: The vertical intercept of a line is the point at which the line intersects the vertical axis, where the input value is zero. This is also called the initial value if the input starts out at a value of zero.

Definition: The horizontal intercept of a line is the point at which the line intersects the horizontal axis, where the output is zero.



Definition: A linear model has a constant rate of change (slope). (This is not true of all models and their corresponding equations.)

It is also very important that you understand how to use technology to graph data and approximate the data with a linear equation. If necessary, please review any process with your instructor.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- linear models are used to model real-world data.
- linear trends can be observed and can be created for data with fluctuations.
- data that change in an approximately linear manner can be modeled by a trendline (line of best fit or piecewise lines of best fit).
- different interpretations of trends seen in data can be made.

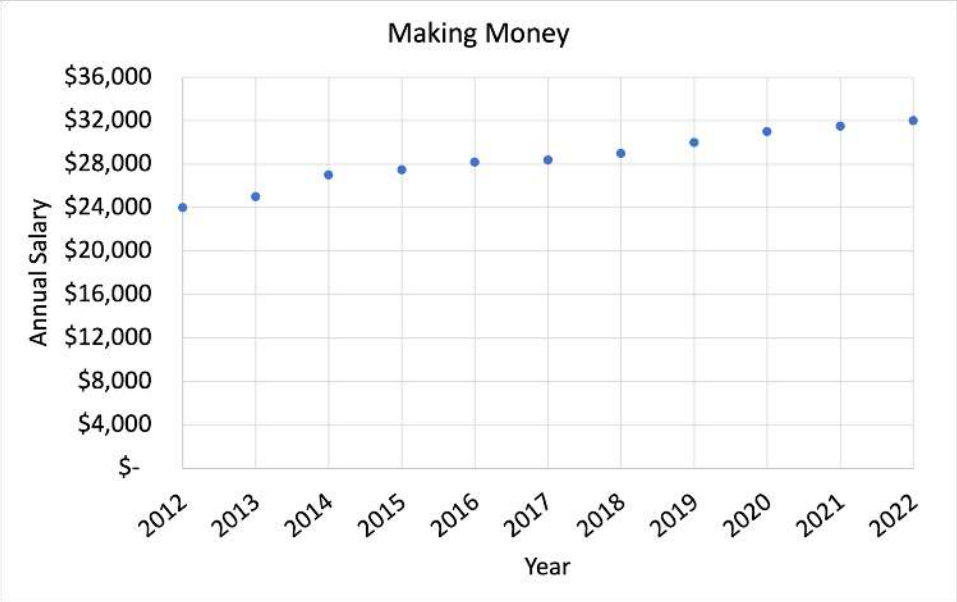
By the end of this collaboration, you should be able to

- interpret the slope and y-intercept in context from an equation or a graph.

- understand when a linear model is appropriate as a model for data.
- use a spreadsheet to create a scatterplot and linear regression model from data.

PROBLEM SITUATION 1: WAGE GROWTH

In 2022, a co-worker created the following graph to represent her annual pay over the years.



(1) Find a linear equation that models the above situation. Use P to represent your co-worker's pay in dollars and t to represent the number of years since 2012. If the trend continues, in what year would she expect to earn $36,000 a year? Take a minute to think about this on your own before sharing your ideas in your group.

It is important to note that your co-worker would be using extrapolation to predict her future salary, so her prediction is unlikely to be perfect and could even be quite a bit off. Extrapolation is the process of using a mathematical model to make predictions that are outside the range of the available data, based on the assumption that existing trends will continue. The problem with extrapolation is that the trend could change in the future, for any number of reasons, which would make the prediction unreliable. For example, your coworker's company could become very successful, resulting in 10% raises for everyone in the company in 2020; or a recession could hit, resulting in no raises for any employees from 2017-2022. Our model from Question 1 would not account for these changes. People often use extrapolation to make predictions, but there is a risk to doing this, since it assumes that current trends will not change.

PROBLEM SITUATION 2: BABY GROWTH AND DEVELOPMENT

(2) The following table shows the average weight for infant boys:[17]

Age	Average Weight for Boys (lbs)
Newborn	7.16
1 month	9.15
2 months	10.91
3 months	12.56
4 months	14
5 months	15.43
6 months	16.53
7 months	17.64

8 months	18.74
9 months	19.62
10 months	20.28

A linear model could be created which could be used to fairly accurately calculate the average weight of an infant boy, given his age in months.

(a) Create a linear model equation for the weight of infant boys.

(b) Use your model to predict the weight of a 20-month-old boy. Round to one decimal place.

(c) Is your prediction in Question 2(b) reasonable? Explain.

(d) According to your model, what would a 25-year-old weigh? Is this reasonable?

PROBLEM SITUATION 3: VIEWS ON GLOBAL WARMING

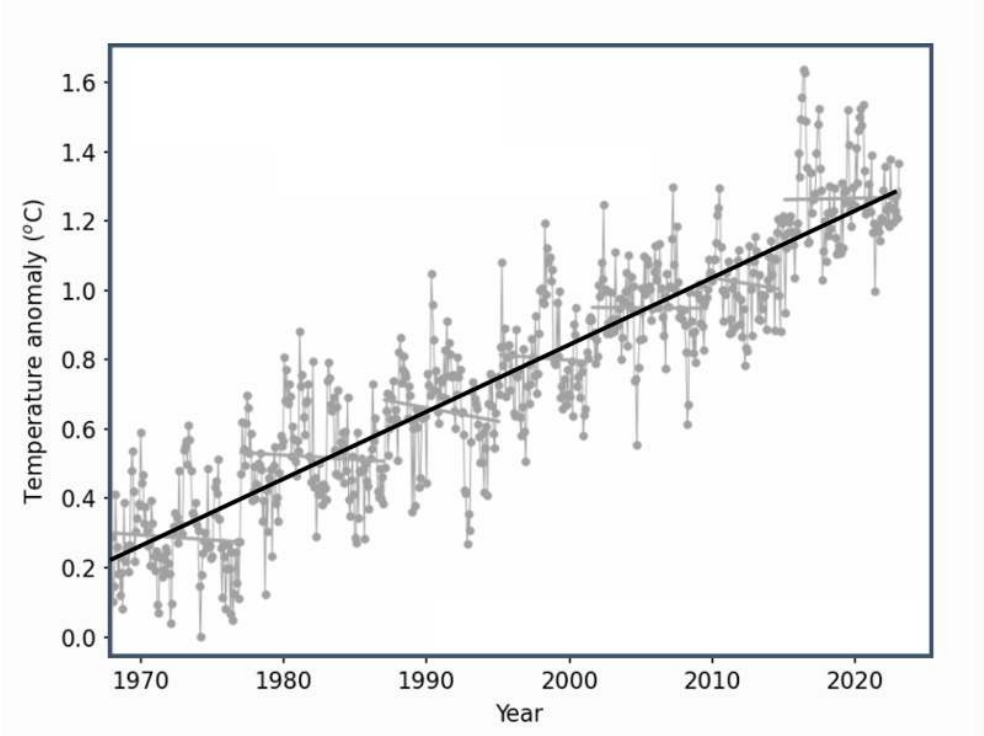
Every year, the Berkeley Earth Surface Temperature (BEST) project releases an analysis of land-surface temperature records (some going back as far as 250 years).[18] According to Berkeley Earth, "The analysis shows that the rise [or growth] in average world land temperature[is]... about 1.3 degrees Celsius." BEST states that "global warming is real," and attributes the temperature trend to a combination of volcanoes and CO2. There are others who are skeptical of the theory of "global warming" and hold a contrary view of weather data. These people believe that there are other explanations and reasons for the climate changes being experienced. Today, we are going to examine the Berkeley Earth project data and linear models for Global Surface Temperature Change to study the issue of global warming.

On the website skepticalscience.com, the data from the Berkeley Earth project has been used as the source for two different analyses of the same information. The first one is labeled "How Realists View Global Warming" and the other is "How Contrarians View Global Warming." The two infographics are reproduced in this collaboration. You will see that each includes a line or line segments along with the actual data. Use these infographics to answer the following questions.

Part I: The Realist View

Below is an infographic showing the data gathered by BEST showing the monthly "Global Temperature Anomaly" (that is, the change in average Global Temperature, in degrees Celsius) for the years 1970–2022. Shown on the graph is a trendline, also known as a line of best fit, that is a linear model representing the trend of the data. It is stated on the project's website that the baseline period is the 1850-1900 mean.[19] This means that the data from 1970–2022 is compared with the data from 1850-1900.

Use the infographic to answer **Questions 3–10**.



(3) What are the units on the horizontal axis? What is the variable?

(4) What are the units on the vertical axis? What is the variable?

(5) What does one point on the graph represent?

(6) Now look at the trendline that is drawn on the graph. What does this line represent?

(7) Do you think the trendline is reasonable? Does it represent the trends in the data?

(8) (a) Estimate the value where the trendline crosses the vertical axis (the value in 1970). Take a minute to think about this on your own before sharing your ideas in your group. Describe the meaning of the point where the trendline crosses the vertical axis.

(b) Estimate the slope of this line in degrees Celsius **per year** (although the original data is monthly, the slope will be per year). Write an interpretation of the slope in this context. Be sure to use units with your values, and use complete sentences.

Finding exact points on this line is difficult. Just do your best to estimate two points and use them to find the slope.

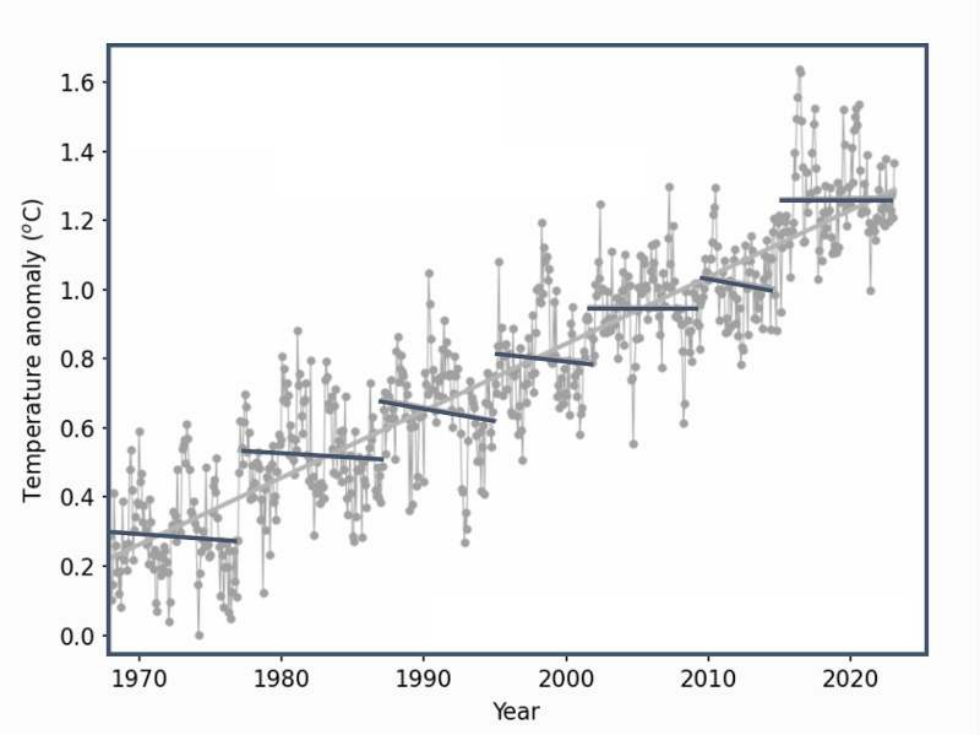
(c) Using your own estimates, estimate the equation of the line. Take a minute to try this on your own before sharing in your group.

(9) What is the model saying about the general pattern of the global temperature. Is it rising or falling?

(10) Using your linear equation from above, predict the temperature change at the end of 2030. Also predict the temperature change at the end of 2050. Do you think this is a reliable prediction for 2050? Why or why not? Explain using complete sentences.

Part II: The Contrarian View

The infographic below contains the same data points for the Global Temperature Change (the change in average Global Temperature in the years from 1970–2022) that were included in the infographic showing "How Realists View Global Warming." Also shown is a **piecewise model**, which is a model defined by multiple pieces, each of which has its own trendline that represents the trend of the data within a certain interval. This infographic shows a linear piecewise model that consists of 6 line segments—each of which can be considered a trendline for that restricted time period. This infographic is meant to depict a common view stated by contrarians that global warming is not real since there have been several periods over the last 50 years when global temperatures have declined on average. Note that the units on both axes and the data are identical to what was in the Realist View. The only difference is the trendlines that are used to analyze the data. Use this infographic to answer the questions below.



(11) Look at the line segments that are drawn on the graph. What do these lines represent as a whole?

(12) Now look at just the segment of the piecewise "best fit" model from 2010 to 2015 (the line segment **second farthest to the right**).

(a) Estimate where the line segment would cross the vertical axis (in year 1970). (**Hint**: It might be useful to use a straight-edged item to extend the line to the vertical axis. ***You can see the line is very slightly slanted***.)

(b) What is the slope value of this line segment? Give an interpretation of the slope value.

(c) Estimate the equation of the 2010 to 2015 line segment.

(13) Do you think we could use the equation you found in Question 12(c) to predict the mean Global Surface Temperature Change in 2035? Why or why not?

(14) Explain what the Contrarian Model is saying about the general pattern of global temperature. Is it rising or falling? Think about this on your own for a minute before sharing your ideas in your group.

(15) If you look at the overall trend of the data in the infographic, it is increasing. However, each line segment in the piecewise model is decreasing (ever so slightly). What does this mean? Explain what the model is saying about the general pattern of global temperature change.

(16) Come up with an example of a similar piecewise linear model. Note that models can be represented in different ways – tables, graphs, written descriptions, or equations.

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

---

[17] http://www.heightweightchart.org/babies-to-teenagers.php

[18] https://berkeleyearth.org/

[19] https://skepticalscience.com/escalator_2022.html

This page titled 2.3: Data Is Trendy is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.3.1: Preparation M.3

(1) Several years ago, Onondaga Community College, in upstate New York, introduced a meal plan that included “Flex Dollars” that were stored on students’ ID cards. The following description of flex dollars appeared on the college’s website:

To accommodate instances where you do not wish to purchase a “meal”, but would still like something to eat or drink, the College offers $150 of “Flex Dollars” included in each meal plan package. This “flexible money” may be used just like cash for any item sold in the Gordon Café, Mawhinney Café, Starbucks or Fresh Express. Do you want that Large Mocha at Starbucks or that pint of Ben and Jerry’s from Mawhinney Café? Use your flex dollars to pay for it!

Onondaga student Loralee tracked her use of flex dollars weekly during the fall semester as shown in the table below:

Week	**Balance on Loralee’s ID Card**
0	$150.00
1	$146.75
2	$142.75
3	$139.00
4	$134.00
5	$130.25
6	$126.25
7	$122.25
8	$118.50
9	$114.50
10	$110.00
11	$106.50
12	$101.50
13	$97.00
14	$93.25

(a) On a separate piece of paper, make a graph of the amount of money left on Loralee’s card over the semester. Note: If completing this problem online, follow the instructions given online to create your graph.

(b) This model is approximately linear. What are the units for the slope?

(i) $

(ii) Week/$

(iii) $/week

(iv) There are no units

(c) What is the significance of the vertical intercept?

(d) What is the significance of the horizontal intercept?

(e) Let M be the amount of money left on the card, t weeks after the beginning of the fall semester. Write a linear equation to model the data. Use Week 0 and Week 14 to calculate the slope. Round your slope to 2 decimal places.

(f) If Loralee continues to spend her flex dollars at the same rate, use the equation from part (e) to determine how much money will be on the card at the end of spring semester which also lasts 14 weeks. Use the equation from part (e) to determine the answer.

(2) A teacher surveyed her students about the amount of physical activity they do each week. She then had their body mass index (BMI) measured. The results are recorded in the table below.

Student	Active Hours per Week	BMI
A	10	16
B	3	25
C	6	24
D	8	10
E	10	16
F	8	18
G	7	21
H	2	28
I	19	9
J	14	12

(a) Using a spreadsheet application, make a scatterplot of this data.

Using the scatterplot, create a linear regression model (trendline). Write the equation of the trendline here:

(b) Using the linear model from Question 2(a), estimate the BMI of a student who is active for 12 hours a week. Round your answer to the nearest tenth.

After Preparation M.3 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration M.3, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	Rating from 1 to 5
identify the slope and the vertical intercept of a linear equation.	
identify a linear trend or data that is increasing or decreasing linearly.	
use a spreadsheet to make a scatterplot.	
create a linear regression model (trendline) using technology.	

This page titled 2.3.1: Preparation M.3 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.3.2: Exercise M.3

(1) In the chart below you will find the statistics from 2003-2023 for the number of babies born, per 10,000 live births, with Down syndrome. Over those 20 years, you can see that there has been an increase.

Year	Down Syndrome at Birth per 10,000 Live Births
2003	9.5
2008	10.3
2013	10.8
2018	11
2023	11.8

(a) Using a spreadsheet application, create a scatterplot of this data and find a trendline which approximates the data in the form y = mx + b, where x is the number of years since 2003 and y is the number of Down syndrome births per 10,000 live births. Hint: Make sure you label 2003 as year 0 and all other years as time since 2003. Write the trendline below.

(b) What does this model give as the 2036 rate of Down syndrome births per 10,000 live births? Round to 1 decimal place.

(c) Do you think your linear model accurately represents the growth in babies born with Down syndrome until 2036? Explain.

(2) The data in the chart below is for the distance (in cm) to the near point, the point nearest the eye at which the eye can accurately focus, at a person's age. Use a spreadsheet application to create a linear regression model for the data in the form y = mx + b, where x is age and y is near point. Using your model, predict the near point of someone who is 80 years old. Round your answer to 1 decimal place.

Age (years)	Near Point (cm)
10	7.5
20	9
30	11.5
40	17.2
50	52.5
60	83.3

(3) Every year, the Mauna Loa Observatory in Hawaii measures the levels of carbon dioxide (CO_2) in the atmosphere. Below are the annual mean amounts, in parts per million, for the years 2006–2022:[20]

Year	Annual mean CO_2 (in ppm)
2006	382.09
2007	384.02
2008	385.83
2009	387.64
2010	390.1
2011	391.85
2012	394.06
2013	396.74

2014	398.81
2015	401.01
2016	404.41
2017	406.76
2018	408.72
2019	411.66
2020	414.24
2021	416.45
2022	418.56

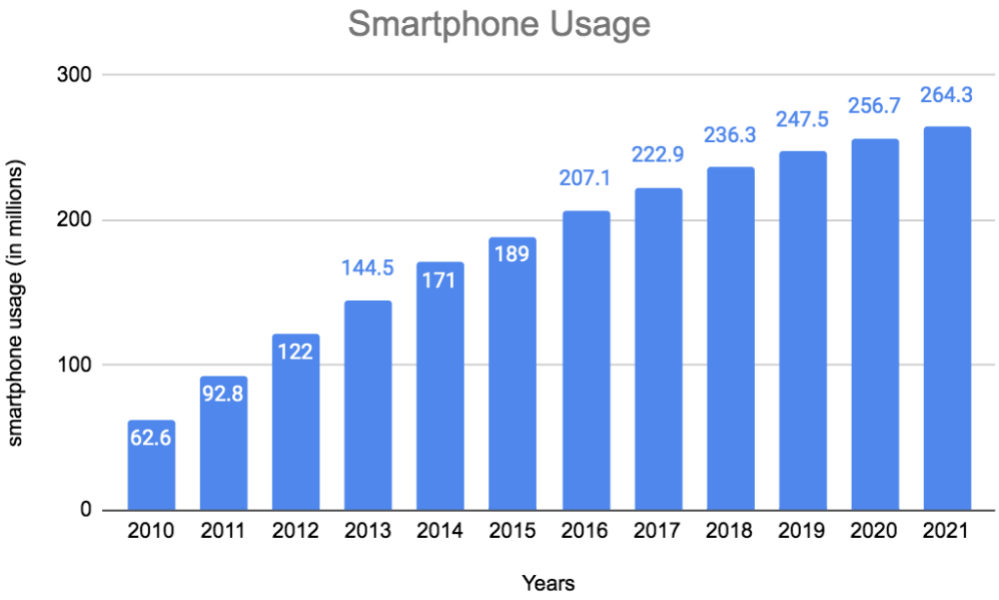
(a) Create a scatterplot of the data using a spreadsheet application, then add a trendline (linear regression model) using the spreadsheet. Write the trendline in the form y = mx + b, where x is the number of years after 2006, and y is the annual mean CO2. Round your slope to 2 decimal places and your vertical intercept to the nearest whole number. (Hint: Since x is the number of years after 2006, adjust the data when you enter it in the spreadsheet.)

(b) What are the units for the input? What are the units for the output?

(c) What is the slope? Explain what the slope tells us about the situation.

(d) Using your model, predict how much CO2 there will be in the atmosphere at Mauna Loa in 2050. Round to the nearest whole ppm.

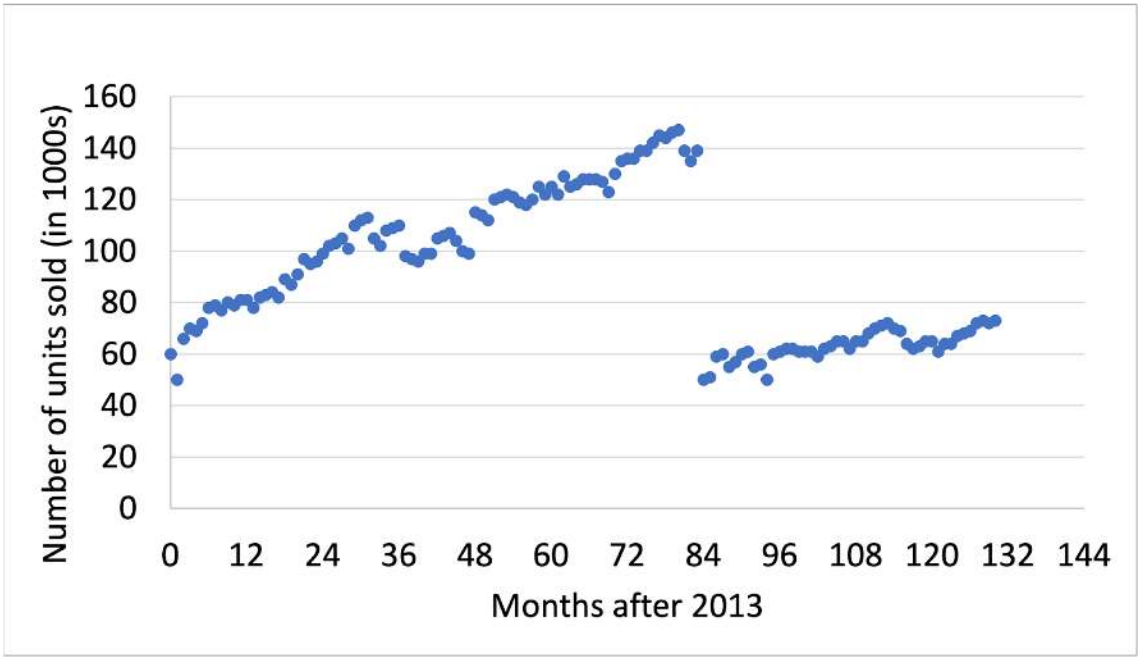
(4) The graph below shows the number of smartphone users in the U.S. (in millions) from 2010 to 2021.[21] (Note: Values for 2016–2021 were only projections.)



(a) Write the linear model for the data, in the form y = mx + b, where x is the number of years since 2010 and y is the number of smartphone users, in millions.

(b) Describe any issues (or problems) with these data that might significantly affect the model's reliability.

(5) The graph below displays the monthly sales data for a specific clothing item from 2013–2023 in the U.S. (where x is months after January 2013, and y is the number of units sold (in thousands). What trends can you identify in the sales?



[20] http://www.esrl.noaa.gov/gmd/ccgg/trends/mlo.html

[21] https://www.statista.com/statistics/201182/forecast-of-smartphone-users-in-the-us/

This page titled 2.3.2: Exercise M.3 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.4: Four Equations…Two Models

INTRODUCTION

In Unit M.2, we explored the cost of electricity in and around New Orleans, LA, during the summer. In this collaboration, we consider electrical charges during the winter.

The company that provides electricity to people in and around New Orleans, Louisiana, is called Entergy. The table below estimates the winter monthly rates charged by Entergy.

Winter Billing Cycles (November – April)
Energy Charges: • First 800 kWh per month at 6.002¢ per kWh • Over 800 kWh per month at 4.766¢ per kWh • Monthly Customer Charge: $8.07

This situation is called "Piecewise Defined." It is piecewise because the cost changes based on the amount of electricity used; cost per kWh is not constant so the graph if this situation is not a line.

For example, a customer who used 700 kWh in a month would be charged

$0.06002 × 700 + $8.07 = **$50.08**

However, a customer who used 1000 kWh in a month would be charged

$0.06002 × 800 + $0.04766 × 200 + $8.07 = **$65.62**

Notice how this customer was charged 6.002¢ for the first 800 kWh, but was then only charged 4.766¢ per kWh for the remaining 200 kWh.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- sometimes models are comprised of more than one algebraic equation.
- setting two equations equal to each other is an algebraic method for finding points of intersection.

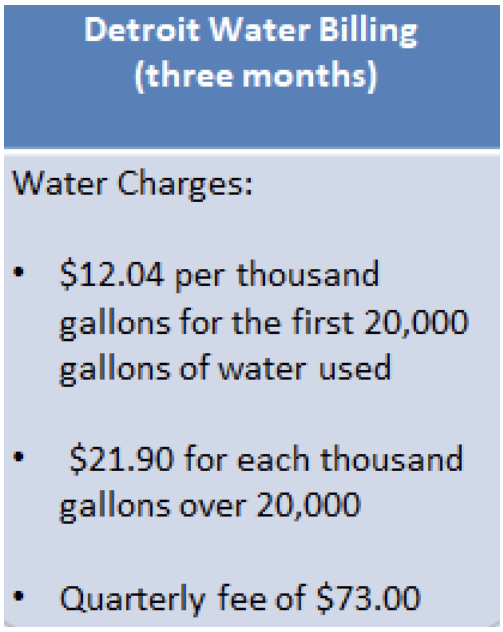
By the end of this collaboration, you should be able to

- create and use piecewise linear models.
- set two linear expressions equal to each other and solve the resulting linear equation.

PROBLEM SITUATION 1: ANOTHER UTILITY BILL

In Problem Situation 1 of Collaboration M.2, we examined electric bills in New Orleans, but did not consider the winter rates, which had tiered pricing. Let us now look at a tiered type of utility bill.

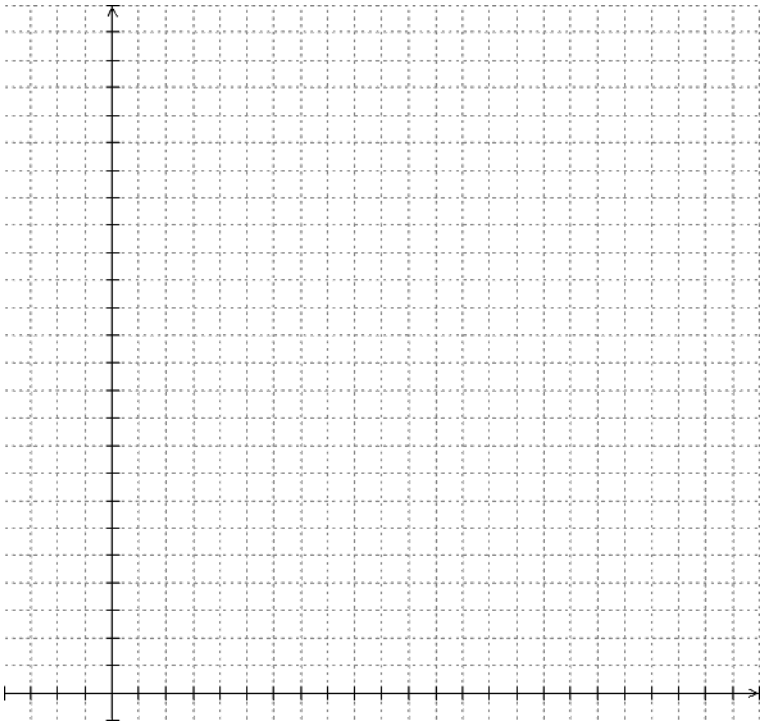
In order to conserve water, the Detroit area water system has a two-tiered system of charges. Customers are billed quarterly (every three months) for water usage. The cost to customers is $12.04 per thousand gallons for the first 20,000 gallons of water used, and $21.90 for each thousand gallons over 20,000. In addition there is always a flat quarterly fee of $73.00.



(1) In the table below, enter the amount of the quarterly bill for each of the water usage amounts listed.

Quarterly Water Usage in Thousands of Gallons	Amount of Bill in Dollars
15	
18	
20	
25	
28	
32	

(2) Plot this data as a scatterplot on the graph below. **Note**: If completing this problem online, follow the instructions given online to create your scatterplot.



(3) Construct an algebraic model that represents the amount of the quarterly water bill (in dollars) for any amount of water used. Take a minute to try this on your own before sharing your ideas in your group. Use g for water usage and W for water bill amount.

(4) Draw the graph of the model on your scatterplot diagram. Note: If completing this problem online, follow the instructions given online.

(5) How many thousands of gallons of water were used if the quarterly water bill was $229? What if the bill was $686?

PROBLEM SITUATION 2: THE COST OF MOVING

Julio is moving from one apartment to another apartment in a neighboring town. He needs to rent a truck for one day to move his belongings. Julio uses the Internet to search for companies with trucks for rent. He finds two companies that have trucks available on the day he needs one: Rent-A-Truk and Trucks-4-Less. Rent-A-Truk advertises a rate of $40 for the day, plus 60 cents a mile for every mile over 40 miles. Trucks-4-Less advertises a rate of $25 a day plus 80 cents a mile for every mile over 25 miles.

(6) What factors will Julio use to determine which company to rent from?

(7) Fill in the blank with little or no calculation: If Julio only drives _____ miles during the day, Trucks-4-Less will end up being the better deal. Explain your reasoning.

(8) Fill in the blank with little or no calculation: If Julio drives _____ miles during the day, Rent-A-Truk will end up being the better deal. Explain your reasoning.

(9) Without doing any calculations, do you know if there will ever be a time when the two rental plans would cost the same? Consider this on your own before sharing your ideas and reasoning in your group.

(10) How much total mileage used will make the cost of both rental plans about the same? Use any method you wish to make this approximation.

(11) Now we will use algebra to find the exact mileage that would make both rental plans cost the same.

(a) Set up an algebraic model that can be used to represent the cost for using Rent-A-Truk given that one plans on driving x miles:

Total cost using Rent-A-Truk plan: $x \leq 40$: R = _____.

Total cost using Rent-A-Truk plan: $x > 40$: R = _____.

(b) If a person plans on driving x miles, what would the algebraic model be to represent the cost for using Trucks-4-Less:

Total cost using Trucks-4-Less: $x \leq 25$: T = _____.

Total cost using Trucks-4-Less: $x > 25$: T = _____.

(c) Julio might want to know at what mileage the cost of the Rent-A-Truk plan equals the cost of the Trucks-4-Less plan. Use your algebraic models to write one equation, using only the variable x, that represents this situation.

Hint: You can think about this situation as R = T .

Hint: Julio estimated that this situation occurs at a mileage higher than 40.

(d) Solve the equation from (c) for x and check whether the solution agrees with your answer to Question 10.

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

This page titled 2.4: Four Equations…Two Models is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.4.1: Preparation M.4

(1) Zeke is driving from his house to the beach and wants to model how far he is from home. He is currently 50 miles from home and will drive 65 mph until he reaches the beach. Which of the following equations models this situation? Let t be time in hours and D be distance in miles from Zeke's home.

(i) D = 50 – 65t

(ii) D = 50 + 65t

(iii) D = 65 + 50t

(iv) D = 65 – 50t

(2) Ezra has a base salary of $25,000 and earns 65% commission on vacuum cleaners. This means that for each vacuum cleaner he sells, he earns an additional amount that equals 65% of the price of the vacuum cleaner. To answer the following questions, use a model where S is Ezra's total salary and v is the total dollar amount of his vacuum cleaner sales.

(a) What is the dollar amount of sales in vacuum cleaners Ezra will need to make in a year if he wants to earn $35,000? Round to the nearest cent.

(b) If vacuums are sold for $200 each, how many vacuum cleaners would Ezra need to sell in a year to make $35,000? Round to the nearest whole number of vacuum cleaners.

(3) The number of members in a professional organization increased each year for ten years, starting in 2013. The result can be modeled by the equation M = 1,200 + 250t.

(a) What do you think the variables M and t represent?

(b) What was the rate of the organization's growth? How large was the organization's membership at the start of the growth period? Explain what your answers mean in the context of the question.

After Preparation M.4 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration M.4, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	Rating from 1 to 5
recognize when a linear model is appropriate.	
write an algebraic equation for a line given the y-intercept and slope.	
solve a linear equation.	

This page titled 2.4.1: Preparation M.4 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.4.2: Exercise M.4

(1) The following table indicates the postage rates for large envelopes and packages of different weights:

Weight, Not Over (ounces)	Cost to Mail a Large Envelope ($)	Cost to Mail a Package ($)
1	0.88	1.71
2	1.08	1.71
3	1.28	1.71
4	1.48	1.88
5	1.68	2.05
6	1.88	2.22
7	2.08	2.39
8	2.28	2.56

(a) Write a sentence explaining how much it costs to send a large envelope.

(b) Write a sentence explaining how much it costs to send a package.

(c) Do you think a large envelope and a package that weigh the same will ever cost the same to mail? Explain.

(d) Create a linear model that could be used to calculate the cost to mail a large envelope, given its weight in ounces. Use E to represent the cost to mail an envelope and x to represent the weight of the envelope in ounces. Write your answer in the form E = mx + b.

(e) Create a linear model that could be used to calculate the cost to mail a package, given its weight in ounces. Let P represent the price to mail a package that weighs x ounces. Write the second linear model simplified in the form P = mx + b. (Hint: This will require two linear pieces.)

(f) Find the weight at which a large envelope and a package will cost the same amount to mail, assuming your linear models still apply to higher weights. Round to the nearest tenth of an ounce.

(2) The table below shows the salary plan for new employees at a company that sells furniture. A new employee earns $25,000 in base salary and 8% commission on furniture sold after his or her first $2,000 in sales for the year.

1. Complete the table by entering a new employee's yearly salary at each level of sales.

Sales	Yearly salary ($)
$0	
$1500	
$2000	
$5000	
$10,000	
$50,000	
$100,000	

(b) Write an algebraic model for an employee's salary if their sales total is $2000 or less. Use S to represent the annual salary.

(c) Write an algebraic model for an employee who sells furniture with a total sales price of x dollars and x is over $2,000. Let S be salary and x total sales price. Write your answer simplified in the form S = mx + b.

(d) How much in furniture sales would a new employee need in order to earn a salary of $40,000 for the year?

(3) According to an article in Forbes magazine, “While major soft drink companies continue to battle declining carbonated soft drink sales, another segment continues to grow…”[22] That segment is bottled water. In 2010, soft drink companies sold 13.75 billion gallons of carbonated soft drinks and 8.61 billion gallons of bottled water. In 2019, the amount of carbonated soft drinks sold fell to 12.01 billion gallons, while the amount of bottled water sold increased to 14.35 billion gallons.

(a) What was the rate of change for soft drink sales from 2010 to 2019? Round to two decimal places.

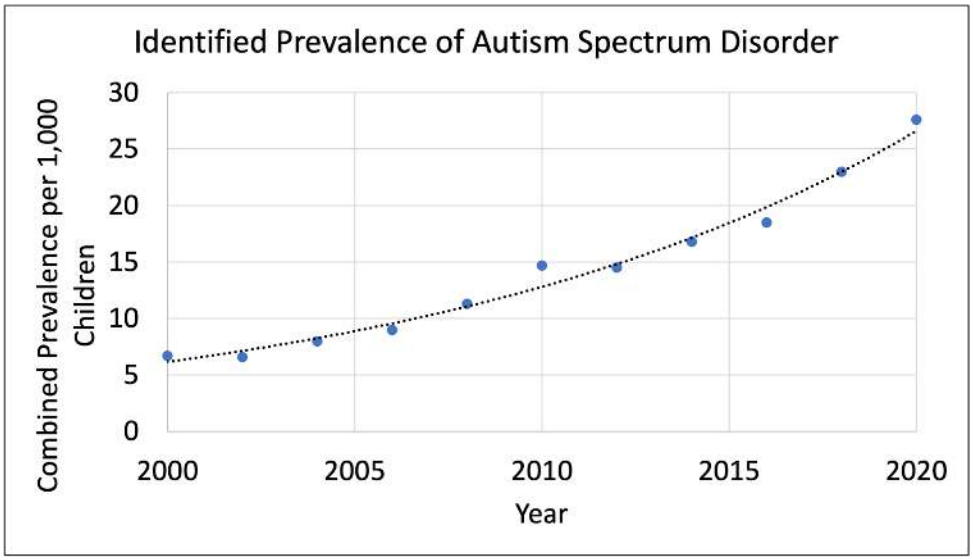
(b) What was the rate of change for bottled water sales from 2010 to 2019? Round to two decimal places.

(c) Using the information from Question 3(a), and assuming that sales are linear, write a linear model that represents soft drink sales from 2010 to 2019. Use t for the number of years after 2010.

(d) Using the information from Question 3(b), and assuming that sales are linear, write a linear model that represents bottled water sales from 2010 to 2019. Let t be the number of years after 2010.

(e) Using your equations, during what year was the amount of soft drinks and bottled water sold the same?

(4) The graph below shows the increase in the number of children with autism from 2000–2020.[23] Is the number of children with autism across these years best modeled by a linear equation, a piecewise model, or neither? Explain.



[22] https://petpla.net/wp-content/uploads/2020/05/BW-consumption-shift-2020-2.jpg

[23] https://www.cdc.gov/ncbddd/autism/data.html

This page titled 2.4.2: Exercise M.4 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.5: Modeling a New Type of Growth

INTRODUCTION

Throughout this course, we will be discussing two of the most important types of mathematical models: linear models and exponential models.

When a quantity increases or decreases by the same amount for each additional unit of the input variable (usually denoted as x), the value of the output variable (usually denoted as y) can be modeled as a linear model. Linear models have a constant additive rate of change, sometimes called a constant absolute rate of change. For example, think of getting a $500 raise every year. We will discuss this in more detail later in this collaboration.

When a quantity increases or decreases by the same percentage for each additional unit of the input variable (usually denoted as x), the value of the output variable (usually denoted as y) can be modeled as an exponential model. Exponential models have a constant multiplicative rate of change, sometimes called a constant relative rate of change. For example, think of getting a 3% raise every year. We will discuss this in more detail later in this collaboration.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- exponential models are best used to represent quantities which have a constant relative change, often expressed as a percentage or ratio.
- the doubling time or halving time associated with an exponential model does not depend on the initial value.

By the end of this collaboration, you should be able to

- set up and refine algebraic equations for exponential models in various growth and decay situations.
- discuss similarities and differences between linear and exponential models.
- use at least one strategy to “solve” an equation in an exponential model.
- create exponential models with given half-lives.

PROBLEM SITUATION: EXPONENTIAL MODELS

In Preparation M.5, you looked at some quantities which change over time. Looking back at Preparation M.5 Questions 2 and 3 (shown below for reference), think about the terminology from previous units that might apply to these questions.

Preparation M.5 Question 2: The population of a small town in Ohio was 5,500 people in the year 2010 and census data showed that the population decreased by approximately 2.5% every year until 2020. What was the approximate population in 2011? 2012? 2020? Round to the nearest person.

Preparation M.5 Question 3: The population of a small town in Michigan was also 5,500 people in the year 2010, but it decreased by an average of 150 people a year until 2020. What was the approximate population in 2011? 2012? 2020? Round to the nearest person.

Now apply this terminology to your answer for Preparation M.5 Question 4 (shown below for reference).

(1) Answer Preparation M.5 Question 4 again below by using complete sentences that include the most appropriate terminology. (**Hint**: Think about the terminology introduced in both Module M and prior Module N.)

Preparation M.5 Question 4: What is different about the population change from year to year in the small town in Michigan (in Question 3 above) and the population change from year to year in the small town in Ohio (in Question 2 above)?

Linear Models (review)

Linear models are based on constant absolute change.

We often say that linear models have a constant additive rate of change. In the context of the Michigan population example (which was increasing by 150 people every year), we would use repeated addition of 150 to calculate the population at any given year.

Exponential Models

When a quantity, such as population, is increasing with a relative multiplicative change each year, this growth can be called **constant relative change**, since the amount of change at any point in time depends upon the amount of the quantity present at that

point in time. This means the ratio of the rate of change of the quantity to its current size remains constant over time. It also means that if the quantity is growing, the more there is of the quantity, the faster it grows.

This type of growth can be modeled with an **exponential model**. Exponential models are used to represent these situations where a quantity (such as population) increases by a certain factor per unit increase in another quantity (such as time).

As with linear models, exponential models can increase, which is called **exponential growth**, or decrease, which is called **exponential decay**.

We often say that exponential models have a constant multiplicative rate of change. Let's consider the Ohio population example, which was decreasing by 2.5% every year, in another way: how to model it if the population were increasing. During Year 1, the population started at 5500 and gained 2.5%, or the population is 5500 + .025 × 5500. We could use the distributive property to write this as 5500(1 + .025), or even better, 5500(1.025).

In other words, if we have a 2.5% increase, we multiply by 1.025. Think of this as you keep 100% of the population (which is why you have the 1) and then gain 2.5% (which is why you have the .025). To find the population at the end of Year 2, you would multiply the population at the end of the 1st year by 1.025, giving 5500(1.025)(1.025), or even better, 5500(1.025)2.

(2) (a) List at least two other situations where quantities, besides population, can have constant relative change and be represented with an exponential model, and explain your reasoning.

Quantity that increases or decreases by the same percentage every day (or year, or month):

Explain your reasoning:

Quantity that increases or decreases by the same percentage every day (or year, or month):

Explain your reasoning:

(b) Refer to the examples from part (a), and create an algebraic formula that represents the changing values over time.

(c) What is “exponential” about these exponential models?

(3) Suppose that the cost of attending a certain four-year private college (tuition and fees) was $16,500 in 1977 and that the cost increased about 3.2% every year.

(a) Create an algebraic model, or formula, that could be used to estimate the cost of attending this college over time.

(b) According to your model, what would be the tuition and fees for 1987, 1997, and 2022? Round your answers to the nearest dollar.

(c) How many years from 1977 did it take for the total cost of tuition and fees to double? To triple?

(d) If the original cost of attending the college in 1977 was only $12,000, how would the doubling and tripling time change?

(e) Explain one strategy you could use to determine the year when this model would predict a cost of $42,000. Assume we are using the original cost of $16,500 for 1977.

(4) The rate of elimination of caffeine from the human body varies greatly from individual to individual. Suppose that Jacob drinks a 16-ounce coffee drink that contains about 310 mg of caffeine. His body eliminates about 13% of the caffeine every hour. Hint: Assume all caffeine is absorbed into the body as soon as the coffee is drunk.

(a) Create an algebraic model for this situation. Let t be the number of hours after drinking a 16-oz cup of coffee, and let A be the amount of caffeine in your body.

(b) How much caffeine would be in Jacob’s body right after he drank a second 16-ounce coffee drink, exactly three hours after drinking the first one?

Summary

Throughout this course, we will be discussing two of the most important types of mathematical models: linear models and exponential models.

When a quantity increases or decreases by the same amount for each additional unit of the input variable (usually denoted as x), the value of the output variable (usually denoted as y) can be modeled as a linear model. Linear models have a constant additive rate of change, sometimes called a constant absolute rate of change. Here are some additional examples of exponential models.

Context	Input Variable	Output Variable	Exponential Model
Scott's starting salary is $55,000, and he will be getting a 2% pay raise every year	Number of years x	Scott's salary y	$y=55{,}0001.02^x$
A local Payday-Loan-Cash office charges 25% interest per week on a $1,000 loan	Number of weeks w	Amount owed A	$A=1{,}0001.25^w$
A brand new Dodge Charger valued at $52,250 depreciates at the rate of 8% per year	Number of years n	Value V	$V=52{,}2500.92^n$

In general, **exponential growth models** are written as $A=P(1+r)^t$, where

- P is the starting value
- r is the **growth rate** expressed as a decimal
- A is the amount amount after time t

In general, **exponential decay models** are written as $A=P(1-r)^t$, where

- P is the starting value
- r is the **elimination rate** expressed as a decimal
- A is the amount amount after time t

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

This page titled 2.5: Modeling a New Type of Growth is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.5.1: Preparation M.5

(1) Population Growth and Investment Growth:

(a) The population of a small town was 5,500 people in the year 2010 and census data showed that the population grew by approximately 2.5% every year until 2020. What was the approximate population in 2011? 2012? 2013? 2020? Round to the nearest person.

Year	Population
2011	
2012	
2013	
2020	

(b) If Karl invested $5,500 in 2010 in an account which earned 2.5% interest every year, how much money would be in the account after one year? Two years? Three years? Ten years? Round to the nearest penny.

Year	Account Balance
2011	
2012	
2013	
2020	

(c) What are the similarities and differences between the previous two situations?

(2) The population of a small town in Ohio was 5,500 people in the year 2010 and census data showed that the population decreased by approximately 2.5% every year until 2020. What was the approximate population in 2011? 2012? 2020? Round to the nearest person.

Year	Population
2011	
2012	
2020	

(3) The population of a small town in Michigan was also 5,500 people in the year 2010, but it decreased by an average of 150 people a year until 2020. What was the approximate population in 2011? 2012? 2020?

Year	Population
2011	
2012	
2020	

(4) What is different about the population change from year to year in the small town in Michigan (in Question 3 above) and the population change from year to year in the small town in Ohio (in Question 2 above)?

(5) Jennifer owns a clothing store. She wants to increase the price of a $15 shirt at her store by 60%. She performs the calculation two ways: 15 + .6 (15) and 15 (1.6). She gets the same result with each calculation. Explain why both methods work.

(6) Jennifer has been selling a coat at her store for $125. She decides to have a “30% off everything in the store” sale. Using Question 5 above as a guide, write two math expressions which a customer could use to calculate the sale price of a coat.

After Preparation M.5 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration M.5, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	Rating from 1 to 5
recognize exponential modeling.	
evaluate expressions involving exponents.	

This page titled 2.5.1: Preparation M.5 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.5.2: Exercise M.5

The **half-life** of a drug in the blood is the amount of time it takes for half the drug to be eliminated from the body. This amount of time is fixed per individual and is independent of the initial amount of the drug. So, for example, if the half-life of a certain drug is 3 hours, and 25 mg of the drug is in the blood at 5 p.m., then 12.5 mg (half of 25 mg) of the drug is left at 8 p.m., and 6.25 mg (half of 12.5 mg) is left at 11 p.m.

1. Recall Question 4 from the collaboration: "The rate of elimination of caffeine from the human body varies greatly from individual to individual. Suppose that Jacob drinks a 16-ounce coffee drink that contains about 310 mg of caffeine. His body eliminates about 13% of the caffeine every hour." You created an algebraic model for this situation. Use your algebraic model to help you answer the questions below. (**Hint**: You may find it helpful to make a table (or use a spreadsheet) with two columns, one labeled "Time Since Drinking Coffee" and the second labeled "Amount of Caffeine in Body.")

(a) Approximately how long will it take for Jacob's body to eliminate half of the caffeine? (As mentioned above, this amount of time is called the caffeine's "half-life.") Write your answer rounded to the nearest whole hour.

(b) If Jacob drank a 12-ounce Pepsi containing 38 mg of caffeine instead of the coffee drink, how long will it take for Jacob's body to eliminate half of the caffeine? Write your answer rounded to the nearest whole hour.

(2) In Question 1 above, we assumed that the body could eliminate 13% of the caffeine present every hour. Wikipedia states that the half-life of caffeine in a healthy adult is often between 4.9 and 6 hours.[24] (Remember, the "half-life" refers to the amount of time it takes for the body to eliminate half of the caffeine currently present in the body.)

If 13% of the caffeine is eliminated every hour, does this imply a half-life between 4.9 and 6 hours? Explain. Does this answer agree with your answer to Question 1(a) above?

(3) The antibiotic *tetracycline* is used for treatment against many different bacterial infections. Suppose that an individual is given 300 mg of tetracycline every 6 hours, and the half-life of tetracycline is 8 hours. Round to the nearest mg for all answers below. (**Hint**: Assume all tetracycline is absorbed into the body immediately.)

(a) If the individual takes his first dose at noon, how much tetracycline remains in his body just before he takes his second dose?

(b) How much tetracycline is in the individual's body just right after he takes his second dose?

(c) How much tetracycline is in the individual's body just before he takes his third dose? Just after?

(4) Joan quit her job at GM when she was 35 years old. Her friend, who is a financial advisor, recommended that she leave her 401(k) retirement savings in GM's retirement plan rather than withdrawing or transferring the money to a new plan. Her friend said that, on average, Joan could expect a 6% increase per year if she left her money in the plan for many years, based on the past performance of GM's plan. Joan will not be able to add more money to the GM 401(k) account, but can open a new one at her new job.

1. Create a model for the amount of money Joan will have in her 401(k) after any number of years. Use A for the accumulated amount of money, t for the number of years, and P for the amount of money in the plan when she quit, which is called the principal.
2. What percent of the principal will Joan have if she leaves the money in the account for the 32 years until she retires? What percent increase does this represent? Round each answer to the nearest hundredth of a percent.
3. If Joan's original principal was $20,100, how much would be in the account when she retires? Round to the nearest cent.
4. Approximately how many years will it take the principal to double? How long to quadruple? Write your answers rounded to the nearest year.

[24] http://en.Wikipedia.org/wiki/Caffeine

This page titled 2.5.2: Exercise M.5 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.6: How long Does That Take?

INTRODUCTION

Let's spend a few minutes at the start of this collaboration to review some important skills and facts from Preparation M.6. Most importantly:

- Solving equations algebraically is all about “undoing” various arithmetic operations. It could be as simple as dividing/multiplying both sides of the given equation, or adding/subtracting a number from both sides of the given equation.

However, as we study more advanced mathematical models, we’ll need more advanced operations. Specifically, square roots, cube roots, and logarithms.

- If we want to undo the squaring operation, we need a special operation, the square root
- If we want to undo the cubing operation, we need a special operation, the cube root
- If we want to undo an operation when the variable is in the exponent, we need to use logarithms

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- the log operation undoes the “raise to a power” operation.
- exponential equations can be solved algebraically using logarithms.

By the end of this collaboration, you should be able to

- solve exponential equations using logarithms.
- check solutions to exponential equations obtained algebraically.

PROBLEM SITUATION: SOLVING EXPONENTIAL EQUATIONS

We use “opposite” operations to solve equations. If you can see what operations are currently applied to a variable, you can do the opposite operations to solve the equation—to find out what number the variable stands for. Some examples of operations in math are: addition, subtraction, multiplication, division, raising a number to a power, and taking a root of a number (such as the square root).

(1) Fill in the column on the right in the table below:

If you have an equation like...	**...you can solve it by using the opposite operation. Write the operation you would use to solve the equation, and then solve it.**
$x+5.4=11.6$	
$t-28=44$	
$5w=38$	
$\frac{d}{12}=3.5$	

(2) In this course, we have also learned that there is a way to “undo” exponents. Fill in the column on the right in the table below:

If you have an equation like...	**...you can solve it by using the opposite operation. Write the operation you would use to solve the equation, and then solve it.**
$x^2=36$	
$t^3=27$	

Exponential equations

Now, consider the following example. Even though both equations include the same numbers and the same variable, they are fundamentally different. It's all about the placement of a variable.

- The equation on the left is a cubic equation; the variable x is in the base of x3
- The equation on the right is an exponential equation; the variable x is in the exponent of 3x

Given equation	$x^3=25$	$3^x=25$
Given operation	The variable x is cubed	3 is raised to the power of x
"Opposite" operation	Cube root	Logarithm

An equation that has a variable in the exponent is called an **exponential equation**. We will be discussing such equations in more detail throughout this collaboration.

Sometimes, we can solve such equations by **inspection**. This means we look at an equation and see what number could go in place of the variable to **make** the equation true, such as in the equation $10^x = 100$, where we know that x = 2 because $10^2 = 100$.

(3) Solve the equations below by inspection.

Solve for x: $5^x=25$

Solve for x: $3^x=81$

In Preparation M.6, we saw that we can also undo the operation in an exponential equation using logarithms, which are often referred to as **logs**. We learned to use the log operation to solve equations when the base is the number 10 (such as $10^x=16$).

(4) Solve the equations below. Think about the operation you would use to solve the equation first and then solve it. Round to the nearest thousandth.

$10^x=135$

$10^x=12.2$

What happens if the base is something other than 10?

As we saw in Question 3 above, if you have the equation: $7^x = 49$, you might be able to use inspection to see that x = 2 because you remember that $7^2=49$. But what about an equation like $1.05^x = 1.6$?

This kind of equation cannot be solved by remembering multiplication facts that we've learned before, or by a simple use of the "Log" key on a calculator, since the calculator assumes a base of 10 (not 1.05). We need to have a way to undo the exponent in this type of equation.

To solve an exponential equation, we can use the following formula:

$$If\ a^x = b,\ then\ x = \frac{log\ b}{log\ a}$$

where a and b are numbers, a is the base of the exponential equation, and x is a variable.

If you have an equation like...	**...you can solve it using this new formula.**
$7^x = 49$	We use logarithms to solve this equation: $x = \frac{log\ 49}{log\ 7}$ Use a calculator to check that you get the solution $x = 2$
If you have an equation like...	...you can solve it using this new formula.

$1.05^t = 1.6$	The way to use logarithms to solve this equation is: $t = \frac{log\ 1.6}{log\ 1.05}$ And when you put this into a calculator, you should get a solution like t ≈ 9.6331635...= 9.633 rounded to thousandths

(5) Use the new formula to solve the equations below. Round to the nearest thousandth.

$3^x = 27$

$6^t = 200$

$1.032^t = 2$

(6) Note that the first equation in Question 5 could have been solved by inspection. Without using the new formula, how could you find an estimate for the second equation in the table?

Summary of Exponential Equations and Logarithms

Let's summarize what we know about exponential equations and logarithms so far:

- Exponential equations usually fit the format $a = b^x$ where a and b are constant.
  For example, $10^x = 24$.
- Depending on the value of b, we solve exponential equations in two ways:

Base	10	a
Exponential equation	$b = 10^x$	$b = a^x$
Solution	$x = log\ b$	$x = \frac{log\ b}{log\ a}$
Example	$Solve\ 2 = 10^x$ $x = log\ 2$ $x = 0.30$	$Solve\ 15 = 2^x$ $x = \frac{log\ 15}{log\ 2}$ $x = 3.91$

Later in this collaboration and in the corresponding exercise assignment, you will see exponential equations that don't quite the format of $a = b^x$. When that happens, try to simplify it first using addition, subtraction, multiplication or division.

Applications of Exponential Equations and Logarithms

Recall the following scenario from Collaboration M.5:

Suppose that the cost of attending a certain four-year private college (tuition and fees) was $16,500 in 1977 and that the cost increased about 3.2% every year. In Collaboration M.5, we found the model for the situation. We then thought about how we might determine the year when tuition would reach $42,000. We considered various methods, such as guess-and-check (guessing different values and checking the results), to find a solution to the equation:

$$42,000 = 16,500(1.032)^t$$

where t = 0 in 1977.

(7) In what year is tuition predicted to be $42,000?

(8) What method could you use to determine whether your solution is correct? Explain your method, then use the method to check your answer from Question 7.

Exponential equations are often seen when working with accounts or loans that use **compound interest**. Compound interest occurs when interest is calculated on both the initial principal and any accumulated interest from previous periods. (Note: For comparison, recall that simple interest does not take into account the accumulated interest, that is, the interest each year is only based on the

initial principal.) The equation for computing the amount of a loan, or in an account, that uses compound interest (assuming no change in principal) is below:

$$A = P(1 + \frac{r}{n})^{nt}$$

In this formula,

- A represents the total amount of the loan or in the account.
- P represents the principal (initial amount of the loan or money put into the account).
- r represents the interest rate, written as a decimal.
- n represents the number of compounding periods per year.
- t represents the term of the loan in years or time since opening the account in years.

(9) Suppose that you take out a $2000 loan with a compound interest rate of 4%, compounded annually (once a year). If you make no payments, and interest compounds annually, in how many years will the loan balance double (reach or surpass $4000)? Think about this on your own for a minute before sharing your ideas in your group.

Hint: when n = 1, the formula simplifies to: $A = P(1 + r)^t$.

(10) Jason has just consumed a 16-ounce cup of coffee containing 300 mg of caffeine.

(a) If caffeine is being eliminated from Jason's body at a rate of 13% every hour, how long before three quarters of the caffeine has been eliminated? Enter your answer rounded to the nearest hour. (Hint: First, set up a model for the situation and then use it to answer the question.)

(b) How much caffeine will be in Jason's blood if he drinks another 16-ounce cup of coffee 5 hours after the first?

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

This page titled 2.6: How long Does That Take? is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.6.1: Preparation M.6

(1) Fill in the blanks:

(a) If you wanted to solve the equation x + 7 = -2 you could _____ from both sides of the equation.

(b) This works because the operation of _____ "undoes" the operation of

_____.

(c) After performing this step, you arrive at the solution, x = _____.

(2) Fill in the blanks:

(a) If you wanted to solve the equation 3x = 15 you could _____ both sides of the

equation by _____.

(b) This works because the operation of _____ "undoes" the operation of

_____.

(c) After performing this step, you arrive at the solution, x = _____.

(3) Fill in the blanks:

(a) If you wanted to solve the equation $x^2 = 15$ you could _____ both sides of the equation.

(b) This works because the operation of _____ "undoes" the operation of

_____.

(c) After performing this step, you arrive at a solution, x = _____or_____. Round to the nearest hundredth.

Logarithms (base 10)

Now we would like to introduce a new operation which "undoes" exponentiation. That operation is written as log(x), and we say this as "the log of x."

The term log is an abbreviation for the mathematical term **logarithm**. A logarithm is a quantity representing the power to which a fixed number (the base, in this case 10) must be raised to produce a given number.

For example, a logarithm, or log operation, can be used to find the exponent to which one must raise the number 10 to get the number x. The expression "log(1000)" is then equal to 3, since $10^3=1000$. Another example is that "log(10)" is equal to 1, since $10^1=10$. (Note: Most calculators [or spreadsheet applications] have a key [or function] labeled "Log" which performs this operation. Note that here, the parentheses do not indicate multiplication; they indicate a process is being done.)

(4) What are the following logs?

(a) log(100) = _____

(b) log(100,000) = _____

Generalizing, since log(1000) = 3 and $10^3 = 1000$ are both true, we can replace the 1000 with 10^3 in the first equation. This gives us the equation: $\log(10^3)= 3$. Similarly, $\log(10^2)= 2$ and $\log(10^1)= 1$. This leads to the general statement that:

$$log(10^x) = x$$

Now we would like to use the log operation to solve equations. Consider the equation $10^x=16$. Notice that the exponent is not a whole number in this case. We can still solve this equation by doing something to both sides of the equation, as we did in Questions 1–3. In this case, we will "take the log" of both sides of the equation:

Initial equation	$10^x = 16$
Apply logs to both sides of the equation	$\log(10^x) = \log 16$ $x = \log 16$

The left side is explained above. The right side is rounded to the thousandth place after using the “Log” key on the calculator to find log 16.	Answer: x = 1.204

We can check this result using a calculator. If x = 1.204, we get $10^x = 10^{1.204} = 15.996$, which is about 16. (It is not exactly 16 because we rounded log 16 to three decimal places.)

(5) (a) Using your calculator (or spreadsheet application) follow the steps above to solve the equation $10^x = 25$. Round to the nearest thousandth.

(b) Using the value you found for x in Question 5(a), find 10^x. Round to the nearest whole number. Does $10^x = 25$?

After Preparation M.6 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration M.6, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	Rating from 1 to 5
solve linear equations by undoing operations.	
understand that the log operation undoes the “raise 10 to the power of…” operation.	

This page titled 2.6.1: Preparation M.6 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.6.2: Exercise M.6

(1) Jose deposited $2,500 into a CD account with an Annual Percentage Rate (APR) of 3%, compounded annually, and he must leave it in the account for at least 5 years.

(a) How much money would he have after 5 years? Round to the nearest cent.

(b) How many years would it take for the account to reach or surpass $4000?

(2) Lorrie bought a new car for $22,234 and a friend who sells used cars told her that she could expect the car to depreciate by about 10% a year for the next 5 years.

(a) How much will the car be worth after 3 years? 5 years? Round to the nearest dollar.

(b) How long will it take for the car to have a value of $10,000? Round to the nearest tenth of a year.

Recall that the *half-life* of a drug refers to how much time is needed for the amount of the drug in someone's blood to be cut in half from being broken down and/or eliminated from the body. So, if the half-life of a drug is 2 hours, and 25 mg of the drug is in the blood at 5 p.m., then 12.5 mg (1/2 of 25) of the drug is left at 7 p.m., and 6.25 mg (1/2 of 12.5) is left at 9 p.m.

(3) In some adults, the half-life for the common pain reliever acetaminophen (the active ingredient in, for example, Tylenol) is about 2 hours. If Janetta is one of those adults and she takes two 500 mg capsules of acetaminophen at 1 p.m., how many milligrams will be in her blood at 7 p.m.? (**Hint**: Think about how the half-life relates to how long it's been since she took the pain reliever.)

(4) The following exchange was found on an online forum:

User 1: I took four extra strength Tylenol yesterday. I took the last two somewhere between 3 and 5 PM. It's been over 24 hours. Would it be safe for me to drink tonight?

User 2: The half-life in a normal adult of Tylenol is 2–3 hours, so a conservative estimate would say you have about 0.39% of the original Tylenol in your body after 24 hours. I'd say you're safe.

User 1: Is that estimate taking into account that I used extra strength Tylenol?

User 2: I believe extra strength is just more of the same, so the ratio should still hold.

(a) Do you agree with User 2's original calculation? Why does User 2 call this a conservative estimate?

(b) Could you provide a more detailed explanation of User 2's second comment? What is User 2 really saying, and why doesn't the increased dosage of Tylenol affect User 2's original calculation?

(5) Luke inherited $10,000 dollars from an uncle when he was 12 years old and it was deposited in a stock market fund at that time. The account was turned over to him when he was 25 years old. Assume that the stock market fund increases an average of 7% a year.

(a) How much will be in the account when Luke turns 25? Round to the nearest cent.

(b) If, when Luke turns 25, he decides to leave the money in the account to save for a down payment on a house, how old will he be when he has $50,000 in the account? Round to the nearest whole year.

Financial institutions use the following formula to calculate the monthly payment, c, on a loan:

$$\textbf{Equation \#1: } c = \frac{P \cdot r \cdot (1+r)^N}{(1+r)^N - 1}$$

This formula can be re-written in this way:

$$\textbf{Equation \#2: } c = (c - P \cdot r) \cdot (1+r)^N$$

In this formula,

- c represents the monthly payment on a loan
- r represents the monthly interest rate. (**Note**: The monthly interest rate is simply the annual percentage rate [APR], as a decimal, divided by 12.)
- N represents the number of monthly payments, which is called the loan's term.

- *P* represents the amount borrowed, which is known as the loan's principal.

(6) Suppose Janetta would like to buy a $17,000 car. She has enough cash to pay for taxes and license fees, and she has $2,000 from the sale of her previous car. She plans to take out a loan to cover the remaining costs. She is offered an APR of 7%.

(a) If Janetta takes a 4-year loan, what will her monthly payments be? Round to the nearest cent. Use Equation #1 above.

(b) Suppose Janetta can only afford monthly payments of $150. How many months would it take her to pay off the loan? Use Equation #2 above and round to the nearest month.

(c) Does your answer seem reasonable for a car loan? Why or why not? What about the value of the car over time as the loan is paid off?

This page titled 2.6.2: Exercise M.6 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.7: A Trust in Social Security

INTRODUCTION

Social Security is an insurance program. Workers pay into the program, typically through payroll withholding where they work. They can earn up to four credits each year. In 2023, you receive 1 credit for each $1,640 of earnings, up to the maximum of 4 credits per year.[26] That money goes into two Social Security trust funds (OASI and DI) where it is used to pay benefits to people currently eligible for them. The money that is not spent remains in the trust funds.

- To qualify for Social Security retirement benefits, workers must be at least 62 and have paid into the system for 10 years or more.[27]
- Workers who wait to collect Social Security, up to age 70, will receive higher monthly benefits.
- Spouses and ex-spouses may also be eligible for benefits, based on their former partner's earnings record.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- basic exponential models can be used to investigate complex scenarios that are described in media articles.
- predictions based on models are *sensitive* to initial assumptions.

By the end of this collaboration, you should be able to

- investigate when an exponential model will reach a given value.
- quantify error in predictive values based on changes in initial assumptions.

PROBLEM SITUATION: THE FUTURE OF SOCIAL SECURITY

The Social Security Trust Funds hold the accumulated savings of the U. S. Social Security system. By law, the funds are invested in the general debt of the United States government, entirely in U.S. Treasury securities.

Projections by the Social Security Board of Trustees (the trustees) show that Social Security expenditures will exceed tax revenues each year through the end of the 75-year valuation period (ending in 2097).

One measure that analysts use to determine the strength of the Social Security Trust Fund finances is the “Trust Fund Ratio.”

$$Trust\ Fund\ Ratio = \frac{Trust\ fund\ asset\ reserves\ at\ beginning\ of\ year}{Total\ cost\ of\ running\ Social\ Security\ for\ that\ year}$$

For example, in 2010 the asset reserves of the Social Security trust fund were $2.5 trillion and the total cost of the program that year was $706.3 billion, so the Trust Fund Ratio was 354%. Another way of saying this is that the trust fund asset reserves were 3.54 times greater than the cost of the program. The Trust Fund Ratio indicates how much of recipients’ annual benefit amounts could be paid from the balance at the beginning of a given year. When the Trust Fund Ratio drops below 100%, then this means that the fund is unable to pay out the total benefits due.

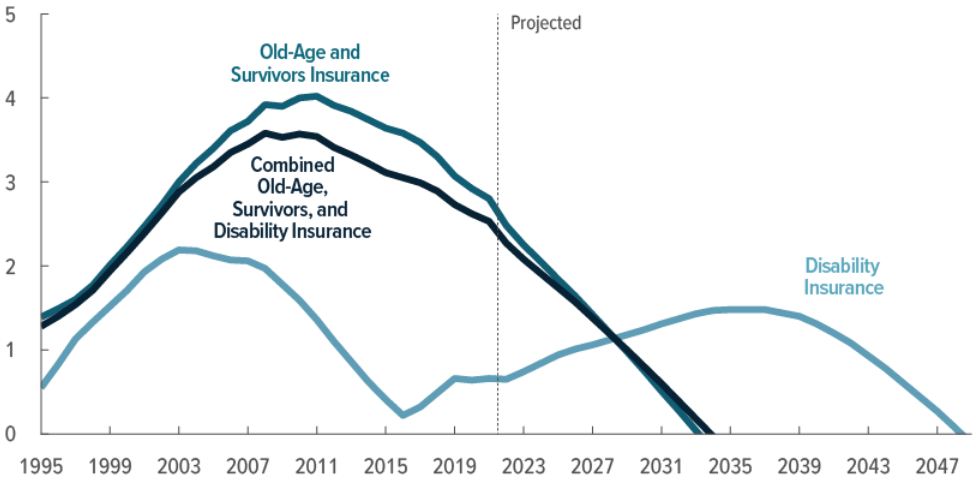
(1) (a) From 2010 to 2019, the Trust Fund Ratio declined 7.7 percentage points each year. The Trust Fund Ratio in 2010 was 354%. Using 354 as the starting amount, create a linear model and determine in what year the model predicts the Trust Fund Ratio to fall below 100. Your model can be a table, a graph, or an equation.

(b) Use the Trust Fund Ratio information for 2020 and 2022 in the table below to create a linear model and determine in what year the model predicts the Trust Fund Ratio to fall below 100. Round the slope to one decimal place.

Year	Trust Fund Ratio (%)
2020	265
2021	256
2022	234

Comparing our answers from Questions 1a and 1b, we can see that using more recent data (2020 to 2022) suggests that the Trust Fund Ratio will be depleted sooner than what the 2010 to 2019 data suggested.

It is projected that in 2023, the Trust Fund Ratio will be anywhere between 204% to 219%. Furthermore, some projections suggest that the funds will drop below 100 (96%) in 2029, which means that full payments to retirees cannot be met, and then it will drop to zero (trust fund asset reserve depletion) around 2034.[28]



Let's now explore what projected trust fund asset reserves suggest about when the fund will be unable to meet its benefit pay out obligations. The Congressional Research Service stated that at the end of 2022, the trust fund asset reserves were valued at around $2.83 trillion.[29] Using projection figures from the 2023 Annual Report of the Board of Trustees for the OASI and DI funds for 2023 and onward, we can estimate that the asset reserves of the Social Security trust fund is going to decline about 15% each year (as a result of paying out more than is being brought in).

(2) (a) Assume the 15% yearly decline estimates are correct, and create a model for the yearly value of the trust fund asset reserves after 2.83 trillion dollars was reached in 2022. Your model can be a table, a graph, or an equation.

(b) How much will the trust funds decrease in the first year after 2022?

(c) How long will it take for the value of the trust funds to decline to half of $2.83 trillion? In what year will that occur?

(3) Using projection data, when the trust fund drops below 400 billion dollars there will not be enough money to cover the benefits due to retirees. Use your Trust Funds Model from Question 2 to find when this will occur.

(4) (a) If additional funds are added to the Trust Funds as a result of policy changes, so that the value of the funds in 2022 is $5 trillion, how long would it take for the value of the funds to shrink to half?

(b) With a fund value of $5 trillion in 2022, when would these funds drop below 400 billion dollars?

(5) What amount of money is needed in the trust funds in 2022 so that the value will not drop below 400 billion dollars until 2042? You can create an equation to model this scenario with a variable for the unknown starting amount (in trillions), and using a repeated 15% decrease for 20 years. This equation is:

$$0.4 = M(0.85)^{20}$$

Solve the equation. What is the starting amount, M?

(6) What are some important factors that might contribute to the decline that is projected in the value of the Social Security Trust Funds?

(7) What could be done to address the problem of the decline in the value of the Trust Fund?

(8) Some assumptions were made in order to produce the Trust Funds Model in Question 2. These assumptions are based on predictions of a future some years off, and could be somewhat inaccurate. However, they are still the best assumptions that could be made at the time and are much better than random guesses. Make a list of these assumptions.

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

[26] https://www.ssa.gov/benefits/retirement/planner/credits.html

[27] https://www.ssa.gov/benefits/retirement/planner/agereduction.html

[28] https://www.ssa.gov/OACT/TR/2023/

[29] https://fas.org/sgp/crs/misc/R42035.pdf

This page titled 2.7: A Trust in Social Security is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.7.1: Preparation M.7

The Social Security program is a program managed by the U.S. government to make payments to individuals after they retire. Basically, current workers pay into the program (a portion of their paycheck is automatically deducted for this purpose) and people who are retired receive payments from this program, based on their income while they worked. The following text is from a March 2023 NPR story.[25] Please read the excerpt and answer the questions.

> **Social Security is now expected to run short of cash by 2033**
>
> The Social Security program is expected to run short of cash to pay promised benefits in about ten years, while a key trust fund for Medicare will run out of funds by 2031, according to new forecasts issued Friday by trustees of both programs. Social Security benefits for retirees and others are primarily paid for through payroll taxes on current workers, and are supplemented by a trust fund.
>
> Benefits paid out by the program have exceeded money coming in since 2021, and the trust fund is now expected to be depleted by 2033. That's a year earlier than forecast last year, thanks in part to slower economic growth.
>
> Unless changes are made before then to shore up the program, 66 million Social Security recipients would see their benefits cut by 23-25%.
>
> Meanwhile, the Medicare trust fund, which supplements payments to hospitals and nursing homes, is also running out of cash. That could result in an 11% pay cut to health care providers unless changes are made by 2031. That deadline is three years later than had been forecast last year.
>
> **Millions depend on Social Security, Medicare**
>
> Treasury Secretary Janet Yellen, who leads the trustees, stressed the importance of propping up both trust funds to avoid draconian cuts in benefits and provider payments.
>
> "Social Security and Medicare are two bedrock programs that older Americans rely upon for their retirement security," Yellen said in a statement. "The Biden-Harris Administration is committed to ensuring the long-term viability of these critical programs so that retirees can receive the hard-earned benefits they're owed."
>
> The primary challenge for Social Security is demographic. As aging baby boomers retire, there are fewer workers paying into the program to support the rising cost of benefits. As of last year, there were just 2.7 workers paying into the system for each person drawing Social Security benefits.
>
> Additionally, a smaller fraction of income is now subject to the payroll taxes that support Social Security.
>
> Patching the program will require higher taxes, lower benefits or some combination of the two.
>
> "The only responsible thing to do is admit that we've got to make changes and we disagree about how to do it but let's sit down and try to figure those out," said Maya Macguineas, president of the Committee for a Responsible Federal Budget. "If we wait until the last minute, they'll be much, much harder."

(1) (a) What is the main point of this article?

(b) Why has the Social Security Administration been paying out more money than expected?

(c) What does the article suggest needs to be done to address these issues?

(2) The Social Security Administration posts answers to frequently asked questions about the Social Security Trust Funds at the site https://www.ssa.gov/oact/progdata/fundFAQ.html. Use the website to help you find answers to the following questions:

(a) Which two of the following funds comprise the Social Security Trust Funds?

(i) Old-Age and Survivors Insurance (OASI).

(ii) Temporary Assistance for Needy Families (TANF).

(iii) Disability Insurance (DI).

(iv) Women, Infants, and Children (WIC).

(c) In what year in the past were the reserves in the OASI trust fund nearly depleted?

(d) How were benefits paid out when the funds were nearly depleted?

(i) Federal income tax was increased to raise the money.

(ii) Money was borrowed from other federal trust funds and repaid with interest.

(iii) Benefits were reduced for four years.

(3) According to carfax.com, the value of a new vehicle can drop by 20 percent after the first 12 months of ownership. Then, for the next four years, you can expect your car to lose roughly 10 percent of its value annually. Use the depreciation rates listed by CARFAX to find the value after 5 years of a brand new Ford F-150 purchased for $62,520. Round to the nearest dollar.

https://www.carfax.com/blog/car-depreciation

(4) Consider an item initially priced at $549.99. Find its value after three consecutive 10% markdowns. Round to the nearest cent.

After Preparation M.7 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration M.7, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	Rating from 1 to 5
set up and evaluate exponential models given starting values and rate of growth/decline.	

[25] https://www.npr.org/2023/03/31/1167378958/social-security-medicare-entitlement-programs-budget

This page titled 2.7.1: Preparation M.7 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.7.2: Exercise M.7

Some of the numbers used in the Trust Funds Model were projections. Let's investigate what happens if the initial fund value in 2022 is different and then if the rate of decline is different. To do so, answer the following questions:

(1) Using a starting value of $2.83 trillion in the trust funds in 2022, with an annual rate of decline of 15%, what would the funds be valued at in 2030? Round to the nearest billion.

(2) Let's say the funds in 2022 were valued at $4.5 trillion. The questions below consider how this new starting value affects the estimated value of the funds in 2030.

(a) What is the new estimated value of the trust funds in 2030? Round to the nearest billion.

(b) Complete the following sentence to compare this result to your calculation in Question 1. Round to the nearest whole percent:

"With a starting value of $4.5 trillion, the estimated value in 2030 is about _____ % more than the estimated value in 2030 with a starting value of $2.83 trillion."

(3) With an original starting value of $2.83 trillion in 2022, assume the actual rate of decline after 2022 was 10% greater than the 15% rate. (Notice that this refers to a 10% relative increase over the 15% rate of decline that was originally estimated in the collaboration, and not an absolute increase of 10 percentage points.) In the questions below, consider how this would affect the estimated value of the funds in 2030.

(a) What is the new estimated value of the trust funds in 2030? Round to the nearest billion.

(b) Complete this sentence to compare this result to your calculation in Question 1 above. Round to the nearest whole percent:

"The estimated value in 2030 after a yearly decline of 16.5% is about _____ % less than the estimated value in 2030 after a yearly decline of 15%."

(4) Now let us assume the starting value of the funds was $4.5 trillion *and* the rate of decline was 10% greater than was estimated in the collaboration. What is the estimated value of the trust funds in 2030? Round to the nearest billion.

(5) Read the following article. Use the quantitative information in the article to answer the questions below.

Berlin Owes Mittenwalde, Small German Town, Trillions On 450-Year-Old Loan
(Reporting by Sophie Duvernoy, editing by Paul Casciato)
BERLIN, July 18 (Reuters) - The sleepy hamlet of Mittenwalde in eastern Germany could become one of the richest towns in the world if Berlin were to repay it an outstanding debt that dates back to 1562. A certificate of debt, found in a regional archive, attests that Mittenwalde lent Berlin 400 guilders on May 28 1562, to be repaid with six percent interest per year.
According to Radio Berlin Brandenburg (RBB), the debt would amount to 11,200 guilders today, which is roughly equivalent to 112 million euros ($136.79 million). Adjusting for compound interest and inflation, the total debt now lies in the trillions, by RBB's estimates.
Town historian Vera Schmidt found the centuries-old debt slip in the archive, where it had been filed in 1963. Though the seal is missing from the document, Schmidt told Reuters that she was certain the slip was still valid. "In 1893 there was a debate in which the document was examined and the writing was determined to be authentic," Schmidt said.
Schmidt and Mittenwalde's Mayor Uwe Pfeiffer have tried to ask Berlin for their money back. Such requests have been made every 50 years or so since 1820 but always to no avail. Reclaiming the debt would bring significant riches to Mittenwalde, a seat of power in the middle ages, which now has a population of just 8,800. Red brick fragments of medieval fortifications still dot the leafy town center.
The town's Romanesque church was once the provost seat for Paul Gerhardt, one of Germany's most prolific hymn writers. Gerhardt, who lived there briefly in the 17th century, is the only noted Mittenwalde resident to date.
Schmidt and Pfeiffer met with Berlin's finance senator Ulrich Nussbaum, who ceremonially handed them a historical guilder from 1539. The guilder was put in a temporary display at the Mittenwalde museum. "This case shows that debts always catch up with you, no matter how old they are," Nussbaum told the Berliner Zeitung paper.
The debt-laden German capital would have difficulty meeting Mittenwalde's demands anyway. According to a report released by the senate finance administration in June 2012, Berlin is already close to 63 billion euros in the red.[30]
Note: $1 = 0.8188 euros when the article was written.

(a) Using information in the first paragraph of the article, create a model for the value of the loan over time, using simple interest, in guilders. Use A for the value of the loan and t for the number of years since the loan began.

Hint: The value of loan is the principal (or amount) of the loan plus any interest earned,

or A = P + I. Also recall the simple interest formula is I = Prt, where *I* is the interest earned, *P* is the principal (or amount) of the loan, *r* is the annual interest rate in decimal form, and *t* is the number of years.

(b) Create a model for the value of the loan using *compound interest*, compounded annually, in guilders. Use *A* for the value of the loan and *t* for the number of years since the loan began.

(c) Using these models, critique the accuracy of the quantitative information in the second paragraph of the article.

---

[30] http://www.huffingtonpost.com/2012/07/19/berlin-mittenwalde-debt-450-debt_n_1685134.html

This page titled 2.7.2: Exercise M.7 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.8: Multiple Variables Weighing on Your mind?

INTRODUCTION

There are many types of multivariable models, but in this collaboration, we will concentrate on models which are similar to the linear models we studied previously, except that they have multiple input variables. In these multivariable models, each input variable has a coefficient, which is a number that multiplies the variable in the model. For example, if G = 3m − 2n, then both 3 and −2 are coefficients.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- formulas using multiple variables are often used in real-world scenarios.
- multivariable models have one output variable and more than one input variable.
- a weighted average is a multivariable model which gives a better "average" for some sets of data.
- many problems are just variations of a weighted average problem, such as calculating grades and GPAs.

By the end of this collaboration, you should be able to

- interpret and use formulas with multiple variables.
- calculate the weighted average of a set of numbers in context.
- construct an expression for a weighted average problem.
- find missing information if the weighted average is known.

PROBLEM SITUATION 1: INVESTIGATING MODELS FOR RESTING ENERGY EXPENDITURE

In Preparation M.8, you read about the importance of calculating resting energy expenditure (REE). A variety of equations have been used to predict REE in critically ill patients.[33] Some of the equations used to predict REE in critically ill patients are shown below.

Many of the equations below use multiple variables to predict REE. Mathematical models that use more than one input variable to produce an output variable are called **multivariable models**.

American College of Chest Physicians equation
REE = 25 * Weight

Harris-Benedict equation	
Men: REE = 66.4730 + (13.7516 * Weight) + (5.0033 * Height) − (6.7550 * Age)	**Women:** REE = 655.0955 + (9.5634 * Weight) + (1.8496 * Height) − (4.6756 * Age)

Penn State 1998 equation
REE = (1.1 * value from Harris-Benedict) + (140 * T_{max}) + (32 * VE) − 5,340
Penn State 2003 equation
REE = (0.85 * value from Harris-Benedict) + (175 * T_{max}) + (33 * VE) − 6,433

Notes about the variables in these equations:

- REE is measured in calories per day.
- Weight is measured in kilograms.
- Height is measured in centimeters.
- Age is in years.
- T_{max} is the maximum body temperature in the past 24 hours, measured in Celsius.
- VE is a measure called *minute volume*, which is the volume of air that is inhaled and exhaled in one minute, measured in liters per minute. Normal levels are around 5–8 liters per minute.

(1) Which variables in the equations above have a positive (increasing) effect on REE? Take a minute to think about this on your own before sharing your ideas in your group. Explain how you determined your answer.

(2) Which variable in the equations above has a negative (decreasing) effect on REE? Take a minute to think about this on your own before sharing your ideas in your group. Explain how you determined your answer.

(3) Suppose Camila and Sofia, who are female twins, want to calculate their REE using the Harris-Benedict equation.

1. Using the information in the table below, calculate the predicted REE for Camila and for Sofia. Do the calculations on your own before sharing your answers in your group.

	Camila	**Sofia**
Weight (kg)	61	62
Height (cm)	165	165
Age (years)	30	30
Predicted REE		

(b) How much higher is Sofia's predicted REE than Camila's predicted REE? How does this relate to the coefficient for weight in the Harris-Benedict equation for females?

(c) Interpret the coefficient for Age in the Harris-Benedict equation for females.

(4) Imagine that you have just completed a moderate 45-minute workout and decide to measure your REE. Which of the equations might be used to support the notion that working out can raise your REE?

(5) Use the profile from Table 1 below to do the calculations (a) - (d). Round to the nearest calorie per day. Try each calculation on your own first before sharing your answers in your group. (Hint: Don't forget that, in the equations, weight is measured in kilograms, height is measured in centimeters, and temperature is measured in degrees Celsius.)

Recall also that 1 lb ≈ 0.4536 kg, 1 inch = 2.54 cm, and a formula for converting degrees Fahrenheit (F) to degrees Celsius (C) is $C = \frac{5}{9}(F - 32)$.

Table 1

	Profile #1
Sex	Male
Age	29
Weight	188 pounds
Height	5' 11"
Tmax	98.6 °F
VE	7 liters/minute

Table 2

Profile #:_____	
Formula	REE
American College of Chest Physicians	(a)
Harris-Benedict	(b)
Penn State 1998	(c)
Penn State 2003	(d)

(6) How accurate do you think these formulas are?

PROBLEM SITUATION 2: TENNIS BALL VS CAR

Multivariable models can also have inputs with non-linear relationships to the output. An example is seen in a fundamental equation in physics for kinetic energy. Kinetic energy is the amount of energy an object has due to its motion. Think about an object moving toward you. The heavier the object and the faster it is moving, the more kinetic energy it has (the harder it will hit you). The equation for calculating kinetic energy is:

$$K = 0.5 * m * v^2$$

where:

K = kinetic energy, in joules

m = mass of the object, in kilograms

v = speed of the object, in meters per second

(7) To consider the importance of kinetic energy, imagine being hit by a tennis ball that is traveling 50 miles per hour (it would probably hurt). Now imagine being hit by an average sized car moving at 65 miles per hour.

(a) To quantify the difference, calculate the kinetic energy of the tennis ball and the car. Try this on your own first before sharing your answers in your group. Round to the nearest one hundredth of a joule.

Hint: Tennis balls have a mass of 58.5 grams. An average sized car has a mass of 1500 kilograms.

Hint: Remember to convert units: 1 kilogram = 1000 grams; 1 mile per hour ≈ 0.447 meters per second.

(b) How many times greater is the car's kinetic energy compared to the tennis ball's kinetic energy? Round to the nearest whole number.

(c) If we instead compare the kinetic energy of two average sized cars, Car 1 moving at 50 miles per hour and Car 2 moving at 25 miles per hour, how many times greater is the kinetic energy of Car 1 compared to Car 2?

PROBLEM SITUATION 3: EVALUATING USED CARS

Your friend Gary is helping you search for a good used car to buy. He researched three cars yesterday and rated each car on a scale of 1 to 10 (10 is the best, 1 is the worst) in four different categories. Gary's recommendations are shown in the table below. Gary recommended that you look at Car 3, because it has the highest average rating of the three cars.

	Car 1	**Car 2**	**Car 3**
Reliability	8	6	7
Gas Mileage	6	7	5
Interior Features/Comfort	5	9	6
Cargo space	7	4	10
Standard Average (Mean)	6.5	6.5	7

(8) (a) What do you observe about how the three cars compare to each other across all categories?

(b) What is an explanation for why Car 3 has the highest overall average?

(9) What are the advantages and disadvantages of using a standard average to select the best overall car?

One way to show that some categories are more important than others would be to count that rating more times. For example, if Reliability is important, you can count it twice; if Comfort is even more important, you can count it four times. Your ratings might now look like this:

	Car 1	**Car 2**	**Car 3**
Reliability	8	6	7

Reliability	8	6	7
Gas Mileage	6	7	5
Interior Features/Comfort	5	9	6
Interior Features/Comfort	5	9	6
Interior Features/Comfort	5	9	6
Interior Features/Comfort	5	9	6
Cargo Space	7	4	10
Average (Mean)			

(10) Make a prediction for which car will have the highest overall rating and then calculate each average.

This kind of average, where some values are counted more times than others, is called a **weighted average**. The weights refer to how much each value is counted. In the above example, you could say that the categories of Gas Mileage and Cargo Space have weights equal to 1, Reliability has a weight of 2, and Interior Features/Comfort has a weight of 4.

So, the total number of weights is **eight**. (1 for Gas Mileage, 1 for Cargo Space, 2 for Reliability, 4 for Interior Features/Comfort = **8 total weights**.)

Weighted average problems are specific cases of multivariable models. You might notice that the categories for rating the cars are acting as input variables used to calculate an overall score for "Car Quality" (weighted average), which is the output variable.

(11) Which of these following represents the multivariable model used to calculate the score for Car Quality (weighted average) in the previous problem? The following variables represent the individual rating in each category: R = Reliability; G = Gas Mileage; I = Interior Features/Comfort; C = Cargo Space.

(i) $Car\ Quality = \dfrac{(R+G+I+C)}{4}$

(ii) $Car\ Quality = \dfrac{(R+G+I+C)}{8}$

(iii) $Car\ Quality = \dfrac{(2R+G+4I+C)}{4}$

(iv) $Car\ Quality = \dfrac{(2R+G+4I+C)}{8}$

(12) Gary suggests using a different weighting strategy. He agrees with all of your preferences, but wants to use bigger weights with the following formula:

$$Car\ Quality = \frac{(10R+5G+20I+5C)}{40}$$

(a) What is the weight for the Interior Features/Comfort category?

(b) How many total weights are used in this equation?

(c) Using the same ratings in each category from above, find the weighted average for each car using this new equation. How do these weighted averages compare to the earlier weighted averages? What could explain your results?

(13) Imagine using this approach in choosing a car and create your own weighted average formula based on your own personal preferences. Use the same four categories: Reliability, Gas Mileage, Interior Features/Comfort, and Cargo Space, but choose relative weights for each category based on what you think is important in a vehicle.

1. Write your weighted average formula.

(b) Explain how you chose the weights for the different categories.

(c) Use your formula to calculate the weighted average of each of the three cars.

(d) Which car has the highest weighted average according to your group's weighting? Which car would be the best out of the three for your group?

It was noted above that weighted average models are specific cases of multivariable models. There are many multivariable models that are not weighted average models, for example, the REE equations from Problem Situation 1.

(14) Consider the weighted average model from Question 12 and compare it to the Harris-Benedict equation for males' REE (both shown below).

Weighted average: $Car\ Quality = \dfrac{(10R + 5G + 20I + 5C)}{40}$

Harris-Benedict for males: REE = 66.4730 + (13.7516 * Weight) + (5.0033 * Height) − (6.7550 * Age)

(a) What differences do you notice between the equations?

(b) Why is the Harris-Benedict model for male REE not a weighted average?

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

---

[33] http://rc.rcjournal.com/content/respcare/54/4/509.full.pdf

This page titled 2.8: Multiple Variables Weighing on Your mind? is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.8.1: Preparation M.8

Read the article “Exercise and Weight Loss,” available online at http://www.health.harvard.edu/diet-and-weight-loss/exercise-and-weight-loss-the-importance-of-resting-energy-expenditure. Then answer the questions below.

(1) According to the article, why is increasing physical activity the key to losing excess weight?

(2) What is resting energy expenditure?

Calculating resting energy expenditure (REE) might be important for those seeking to lose weight—its usefulness is still being studied. But calculating REE has an important role in medical care as well. Critically ill patients under hospital care must be provided with sufficient calories to cover their REE and other energy needs.[31] Certain kinds of medical treatment cause patients to lose weight, and studying REE could help doctors and researchers understand the causes of weight loss.[32]

(3) Using an online calorie calculator, such as the one at https://www.mayoclinic.org/healthy-lifestyle/weight-loss/in-depth/calorie-calculator/itt-20402304, estimate the number of daily calories you need to maintain your current weight.

(4) REE accounts for around 60−75% of total daily calorie needs. Using your result from Question 3, estimate your resting energy expenditure.

(5) Suppose the heights for a family of four are the following:

Father: 6 feet, 3 inches.

Mother: 5 feet, 7 inches.

Daughter: 5 feet, 4 inches.

Son: 5 feet, 2 inches.

1. What is the average (mean) height of the members of this family, in feet and inches?
2. Using the formula $c = 2.54 \cdot i$, where c is height in centimeters, and i is height in inches, what is the average (mean) height of this family in centimeters? Round to the nearest cm.

After Preparation M.8 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration M.8, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	Rating from 1 to 5
calculate the mean of a set of numbers.	
evaluate formulas.	
convert units of measurement.	

[31] http://rc.rcjournal.com/content/respcare/54/4/509.full.pdf

[32] http://www.nutritionjrnl.com/article/S0899-9007(11)00307-8/abstract

This page titled 2.8.1: Preparation M.8 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

2.8.2: Exercise M.8

The *Consumer Reports* rating for cereals is created using numerous factors[34] and indicates how healthy a cereal is for you. An instructor at Yale used a "Healthy Breakfast" dataset to analyze how the *Consumer Reports* rating for various cereals is related to the nutrition information on the cereals' packages.[35] In particular, the instructor looked at the amount of fat, sugar, and fiber in the cereal. He found that the following was the best model for how these nutritional components related to the *Consumer Reports* rating, where fat, fiber, and sugar are all in grams per serving:

Rating = 53.4 − (3.48 * Fat) + (2.95 * Fiber) − (1.96 * Sugar)

(1) Maria wanted to know what the *Consumer Reports* rating for her three favorite cereals would be, but all she had was the nutrition labels from the cereal boxes. She decided to use the equation above to calculate the rating. The information from the cereal boxes is shown in the table below.

	Cereal 1	**Cereal 2**	**Cereal 3**
Fat	2	11	3
Fiber	3	12	2
Sugar	1	4	9
Rating			

1. Calculate the rating for each cereal and enter it into the table. Round to two decimal places.

(b) If Maria wanted to eat the most healthy cereal (the cereal with the highest rating), which cereal should she choose?

(c) Interpret the coefficient for Fiber in the model above.

In Collaboration M.8, you saw examples of weighted averages using relative weights.

(2) Calculate the weighted averages for two additional used cars using the weights given in the table below.

	New weights		**Car 4 Ratings**	**Car 5 Ratings**
Reliability	5		7	6
Gas Mileage	6		5	7
Interior Features/Comfort	10		8	7
Cargo Space	4		6	8

(a) Weighted average for Car 4:

(b) Weighted average for Car 5:

(3) Weighted averages can also be created using *percentages* as weights.

(a) In your weighted average equations above, how many total weights are used?

(b) For each category, what percent of the total weights are represented by that category's weight? Enter your answers in the table.

	New Weights as a Percentage of Total Weights
Reliability	
Gas Mileage	
Interior Features/Comfort	

Cargo Space

(c) In the tables below, multiply each category's percentage weight by its rating, and add up the four amounts.

	Percentage Weights (written as a decimal)		**Car 4 Ratings**		
Reliability		x	7	=	
Gas Mileage		x	5	=	
Interior Features/Comfort		x	8	=	
Cargo Space		x	6	=	
Total					

	Percentage Weights (written as a decimal)		**Car 5 Ratings**		
Reliability		x	6	=	
Gas Mileage		x	7	=	
Interior Features/Comfort		x	7	=	
Cargo Space		x	8	=	
Total					

(d) Write an equation for the weighted average using the percentage weights. Use upper case letters to represent the categories (e.g. *R*, *G*, *I*, and *C*).

A Grade Point Average (GPA) is another kind of weighted average. Letter grades are assigned a number value in order to calculate a numerical average. This table shows typical point values:

A	A−	B+	B	B−	C+	C	C−	D+	D	D−	F
4.0	3.7	3.3	3.0	2.7	2.3	2.0	1.7	1.3	1.0	0.7	0

When calculating a GPA, some grades have a higher impact than others because the weights of the grades for each course are determined by the number of credit hours in the course.

(4) Jose gets the following grades in his courses his first term in college. What would his Grade Point Average (GPA) be for that term? Enter your answer rounded to two decimal places. (Hint: You will need to change each letter grade into a point value and then use the number of credits as the weight for each grade. Remember that you need to use the total number of weights [credits] in your calculation.)

Class	**Credits**	**Grade**
Pre-calculus	5	B

Spanish I	5	A-
Chemistry I (with lab)	6	B+
Archery	1	A

To calculate a GPA after several terms, you can think of the current GPA as being weighted by the total number of credits taken in previous terms, since the credit-hour weights of the courses the student has taken previously have already been used in the calculation of the GPA from previous terms. For example, suppose another student currently has a GPA of 2.85 after 40 credits and now completes four new courses. Then that student's GPA can be calculated from this information:

Class	**Credits**	**Grade**	**Point Value**
Pre-calculus	5	B	3.0
Spanish I	5	A−	3.7
Chemistry I (with lab)	6	B+	3.3
Archery	1	A	4.0
Credits taken in previous terms	40	--	2.85
Total	57		

(5) Calculate this student's GPA. Round to the nearest hundredth.

(6) Suppose Alex wants to get a C+ or better in his math class and his instructor says a C+ is 75%–77% for that course. Below are Alex's grades and his teacher's weightings of different assessments. Calculate the minimum grade (as a percent) Alex must earn on the final to get a 75% or higher in the class. Assume that the percent on the final exam must be a whole number. (**Hint**: One approach is to set up an equation using the grades and the weighted percentages, then use a variable to represent the unknown final exam grade. The weighted average should equal the minimum score Alex would need to get on the final exam in order to earn a C+ in the class.)

Assessment	**Grade**	**Percentage Weight**
Homework	76% (average)	10%
Quizzes	82% (average)	20% (total)
Midterms **(3)**	65%, 60%, 72%	10% **each**
Final Exam	???	40%

(7) A college student, Andrea, wants to apply for a scholarship. The scholarship requires a minimum GPA of 2.5. Andrea has a GPA of 2.3 after completing 70 college credits. She is taking 16 credits this term. What is the minimum GPA Andrea will need this term in order to be eligible for the scholarship at the end of this term?

---

[34] http://www.consumerreports.org/cro/cereals.htm

[35] http://www.stat.yale.edu/Courses/1997-98/101/linmult.htm

This page titled 2.8.2: Exercise M.8 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

3: College Module S

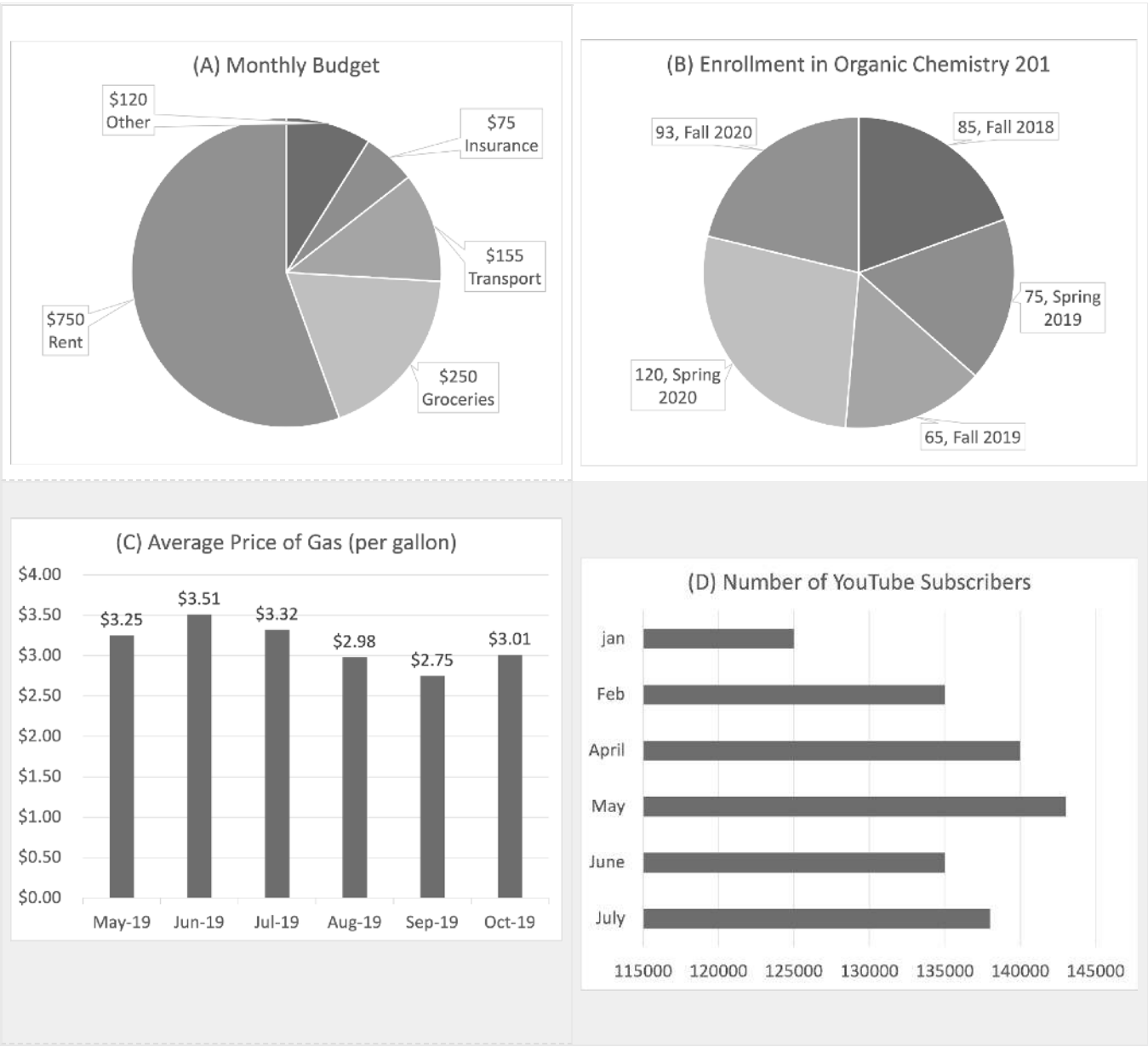
3.1: Interpreting Charts and Graphs
3.1.1: Preparation S.1
3.1.2: Exercise S.1
3.2: Measuring Success
3.2.1: Preparation S.2
3.2.2: Exercise S.2
3.3: Statistical Analysis Process
3.3.1: Preparation S.3
3.3.2: Exercise S.3
3.4: Sampling and Error
3.4.1: Preparation S.4
3.4.2: Exercise S.4
3.5: Exploring Climate Data
3.5.1: Preparation S.5
3.5.2: Exercise S.5
3.6: Exploring Climate Variability
3.6.1: Preparation S.6
3.6.2: Exercise S.6
3.7: Modeling with Normal Distributions
3.7.1: Preparation S.7
3.7.2: Exercise S.7
3.8: Significance in Observational Studies
3.8.1: Preparation S.8
3.8.2: Exercise S.8
3.9: Significance in Experiments
3.9.1: Preparation S.9
3.9.2: Exercise S.9

This page titled 3: College Module S is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.1: Interpreting Charts and Graphs

INTRODUCTION

We can help make a chart meet accessibility standards by adding a table of values or making it clear what category a value belongs to, making sure segments are clearly identifiable (or use colors that meet accessibility standards), and labeling the chart and axes. Additionally, depending on the data to be represented, we need to make good decisions about the type of chart to use. Of the four charts below, one is a poor representation of its data, and one is somewhat misleading. Discuss with your group which of the following charts is a poor representation and which one is a little misleading.



SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- bar charts and pie charts are graphical tools.
- bar charts and pie charts are used to tell a quantitative story in a visual way.

By the end of this collaboration, you should be able to

- read bar charts and pie charts.
- interpret the story behind bar charts and pie charts.
- create bar and pie charts independently, by hand and using a spreadsheet.

PROBLEM SITUATION 1: IT'S ALL IN HOW YOU LOOK AT IT

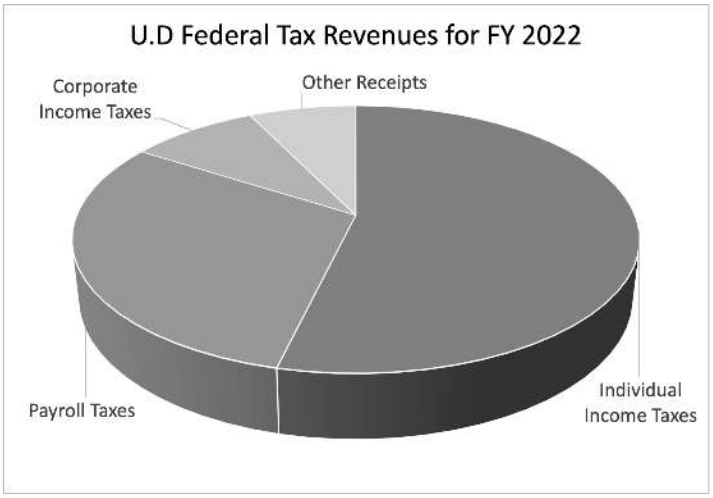
Legislators often have to make policy decisions about how to collect taxes and then spend the resulting revenue. While making these decisions, politicians are often confronted with a wide variety of data, such as from tables, charts and graphs. This information is presented to the politicians from many different sources, too. Many of those sources, such as lobbyists and unions, can have self-serving and competing agendas. In this collaboration, we will think about how the manner in which data are presented might influence politicians' decisions.

In 2022, U.S. federal revenues amounted to $4.9 trillion—$850 billion (about 21%) more than the receipts recorded in 2021. That increase was the largest in five decades. Oftentimes, the government spends more than it collects, which results in a budget deficit. In 2022, the U.S. government spent around $6.3 trillion, resulting in a $1.4 trillion deficit.[2] Such deficits occur regularly, and there are many appeals made to legislators about how money should be collected and spent.

One of the main points of contention is money collected through payroll taxes—in particular, the FICA taxes that fund Social Security and Medicare. Some lobbyists argue that payroll tax rates are too low since there is a deficit, and others argue that they are already too high, harming both companies and their employees.

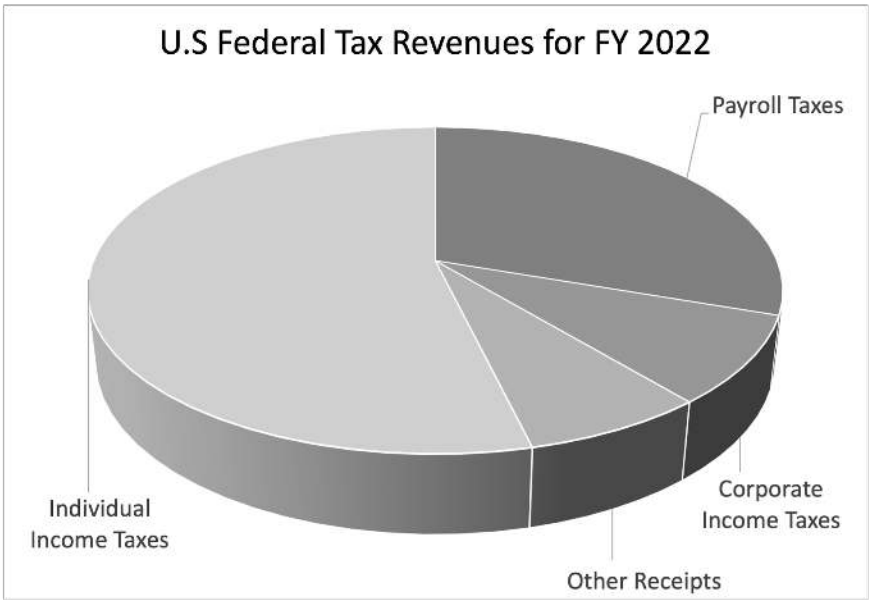
In order to present data in a quick, compelling way, people such as lobbyists often use graphs and charts, since visual data is easier to understand than raw numbers presented in a table. Figure 1 below is a 3-Dimensional (3D) pie chart presenting the tax revenue data for fiscal year 2022.[3]

Figure 1: Pie Chart Showing Tax Revenues for Fiscal Year 2022



However, this is not the only way to create a pie chart for this data. Figure 2 shows a second 3D pie chart presenting the same data.

Figure 2: Pie Chart Showing Tax Revenues for Fiscal Year 2022



1. Focus on the Payroll Taxes in the pie charts in Figures 1 and 2. Does the slice of pie representing Payroll Tax revenues seem the same size to you in both charts? Why or why not? Do you think either chart is misleading? Take a minute to think about this on your own before sharing your ideas in your group.
2. Suppose a group wishes to lobby a congresswoman to vote against an increase in payroll taxes. The group believes that payroll taxes are already high enough. Would this group show the legislator Figure 1 or Figure 2? Why? Can you think of a case in which a group would show the congresswoman the other figure instead?
3. As we learned in Preparation S.1, and can see in Questions 1 and 2, using 3D pie charts can be misleading, and may even be dishonest. The video you watched in Preparation S.1 recommended that we use 2-Dimensional (2D) pie charts to avoid these issues. Below is part of a table that gives a breakdown of how taxes were collected in fiscal year 2022. The data in the table was used to create both Figure 1 and Figure 2 above. Fill in the missing amounts and percentages. Numbers are rounded to the nearest billion dollars, and percents to the nearest tenth of a percent.

Table 1: 2022 U.S. Federal Tax Revenues

Source of Revenue	**U.S. Federal Revenue for Fiscal Year 2022 (in billions of dollars, rounded to the nearest billion)**	**Percentage (rounded to the nearest tenth of a percent)**
Individual Income Tax	$2632 billion	
Payroll Tax (FICA, etc.)		
Corporate Income Tax		8.7%
Other (Excise, Estate, etc.)		7.3%
TOTAL:	**$4896 billion**	**100%**

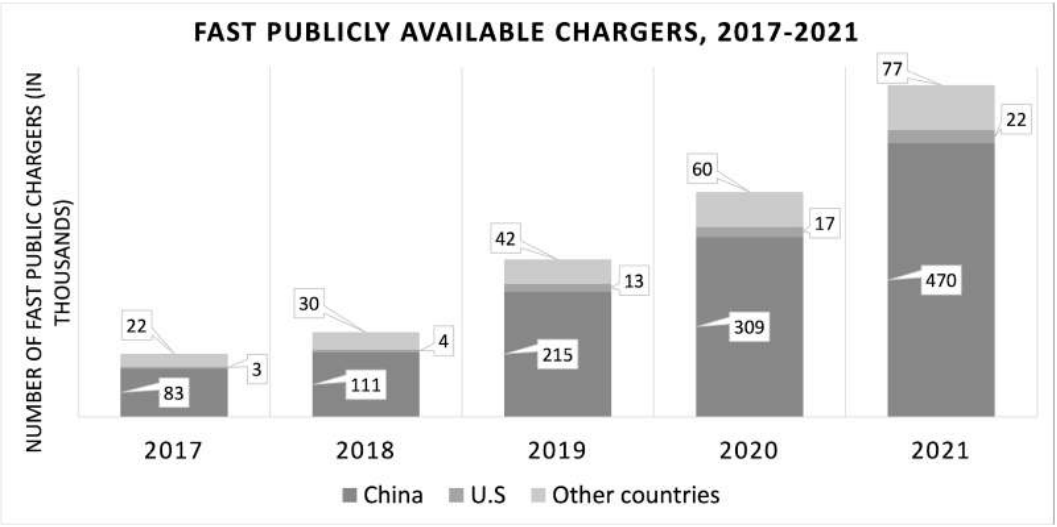
4. Were you surprised to see the actual percentage for Payroll Taxes? Do you feel either Figure 1 or Figure 2 does an adequate job representing that percentage visually?
5. The original pie charts given in Figure 1 and Figure 2 were rendered in a 3D view. Redraw this pie chart in 2D, using the guidelines given in the slideshow from Preparation S.1. Explain how the 2D chart would provide a clearer story to policy makers. Sketch the chart by hand and estimate the wedge sizes. Try to make your sketch as neat as possible.

PROBLEM SITUATION 2: HOW THE COMPETITION STACKS UP

The bar chart below (Figure 3) is called a **stacked bar chart**, as the data in each bar is made up of more than one part, where one part is stacked on top of another. The chart shows the number of fast publicly available electric car chargers for China, the U.S.,

and other countries for the years 2017 to 2021, rounded to the nearest thousand.

Figure 3: Stacked Bar Chart Showing Fast Publicly Available Chargers for the years 2017-2021



6. Using the information in the bar chart, answer the following questions.
    1. What was the total number of fast public chargers in 2020?
    2. What percentage of fast public chargers in 2020 were not in China or the U.S.? Round to one decimal place.

(7) (a) Out of the five given years, in which year was the percentage of fast public chargers outside of China and the U.S. the greatest? Discuss in your group before recording your group's answer in the given space.

2. Between 2019 and 2020, what was the **absolute change** in the number of fast public chargers in China?
3. Between 2019 and 2020, what was the **relative change** in the number of fast public chargers in China? Round to two decimal places.

(8) Now that you have examined some details about the stacked bar chart, consider the whole chart. What is the main story being told by this chart? In other words, what main trend(s) does the chart show? Are there any other possible stories being told? Discuss this with your group members.

(9) Create another bar chart. In this bar chart, use only China and the U.S. fast charger figures. Place the bars representing these two companies side-by-side for each quarter, not stacked. What story is being told by this chart? Discuss this question with the members of your group.

(10) Use the data from 2017 and 2021 to create two linear models. The first linear equation should model the number of fast public chargers available in China in a given year, and the second linear equation should model the number of fast public chargers available in the U.S. in a given year. Let the independent variable (which is the input, or x-value) be the number of years after 2017.

1. What is your linear model for fast public chargers in China? Use C for the number of fast public chargers (in thousands), and t for the number of years after 2017.
2. What is your linear model for fast public chargers in the U.S.? Use U for the number of fast public chargers (in thousands), and t for the number of years after 2017.
3. What does the slope mean for your linear model of fast chargers in China?
4. Use your models to predict the number of fast chargers in China and the U.S. in 2029. What do you think of the accuracy of these predictions? What about for the year 2061?

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

[2] https://www.cbo.gov/publication/58888

[3] https://www.cbo.gov/publication/58592

This page titled 3.1: Interpreting Charts and Graphs is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.1.1: Preparation S.1

Data are information that we collect from our world and involve facts and observations. Module S focuses on the presentation, collection, investigation, and analysis of data. These skills are critically important and necessary to understand the complex issues in our world. We begin by exploring important concepts and issues pertaining to the presentation of quantitative information.

Charts and graphs are often used to communicate quantitative information. Using a chart or graph is a visual way of telling a story that includes quantitative information, in a clear and concise manner. In the next lesson, we will focus on **pie charts** and **bar charts**. Please read the following article that gives some information about bar and pie charts:

https://chartio.com/learn/charts/how-to-choose-pie-chart-vs-bar-chart/

Use what you learned from the article to respond to the following questions:

1. When would we want to use a pie chart?
2. What should be included in a complete pie chart? Is there anything you were encouraged to avoid when using pie charts?
3. When would we want to use a bar chart?
4. What elements should be included in a complete bar chart? Is there anything you were encouraged to avoid when using bar charts? If yes, explain.

Accessibility Guidelines

The term **accessibility** refers to the design of products, devices, services, or environments so as to be usable by people with disabilities.[1]

Here are some accessibility guidelines for creating charts

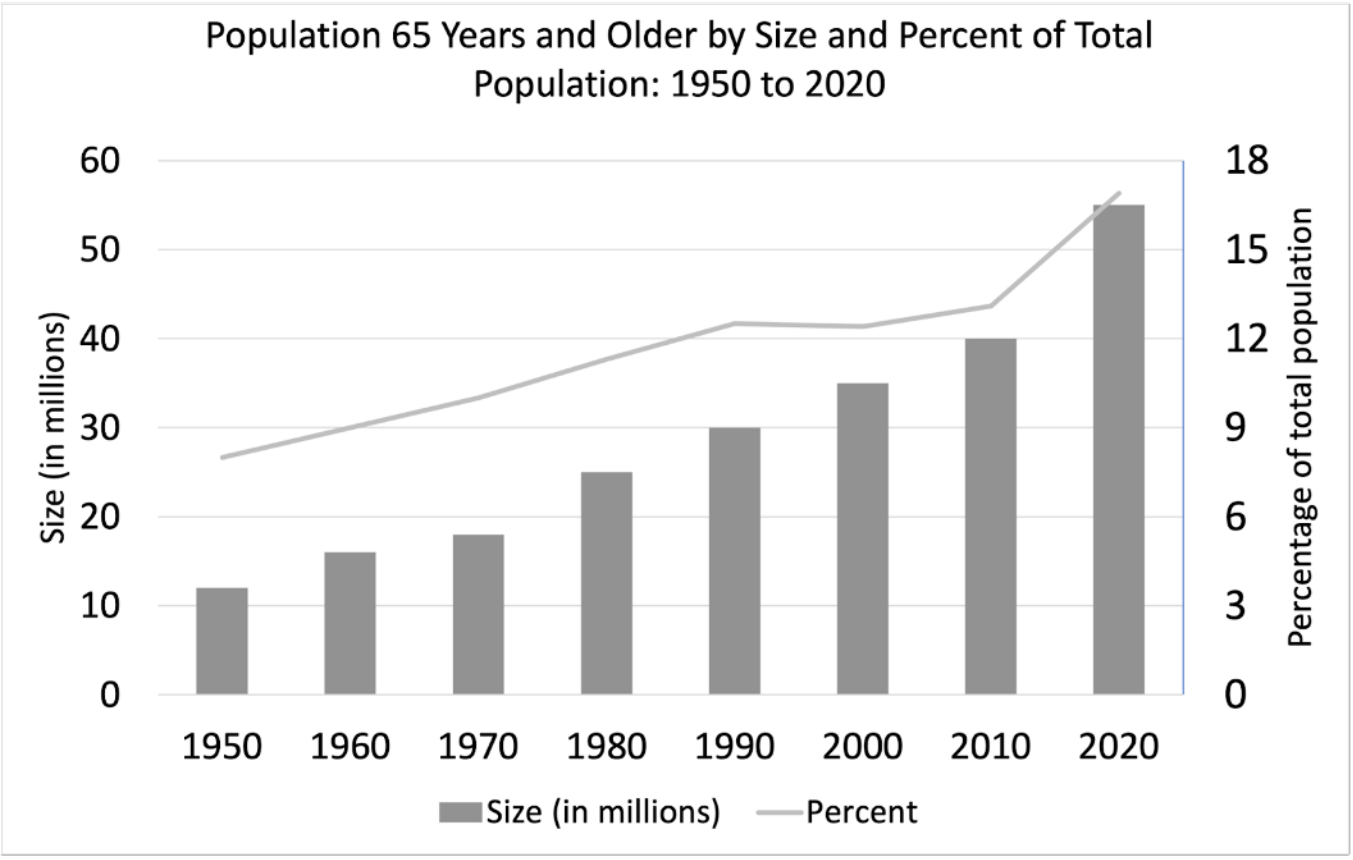
- https://accessibility.psu.edu/images/charts/

Here are some accessibility guidelines for group presentations:

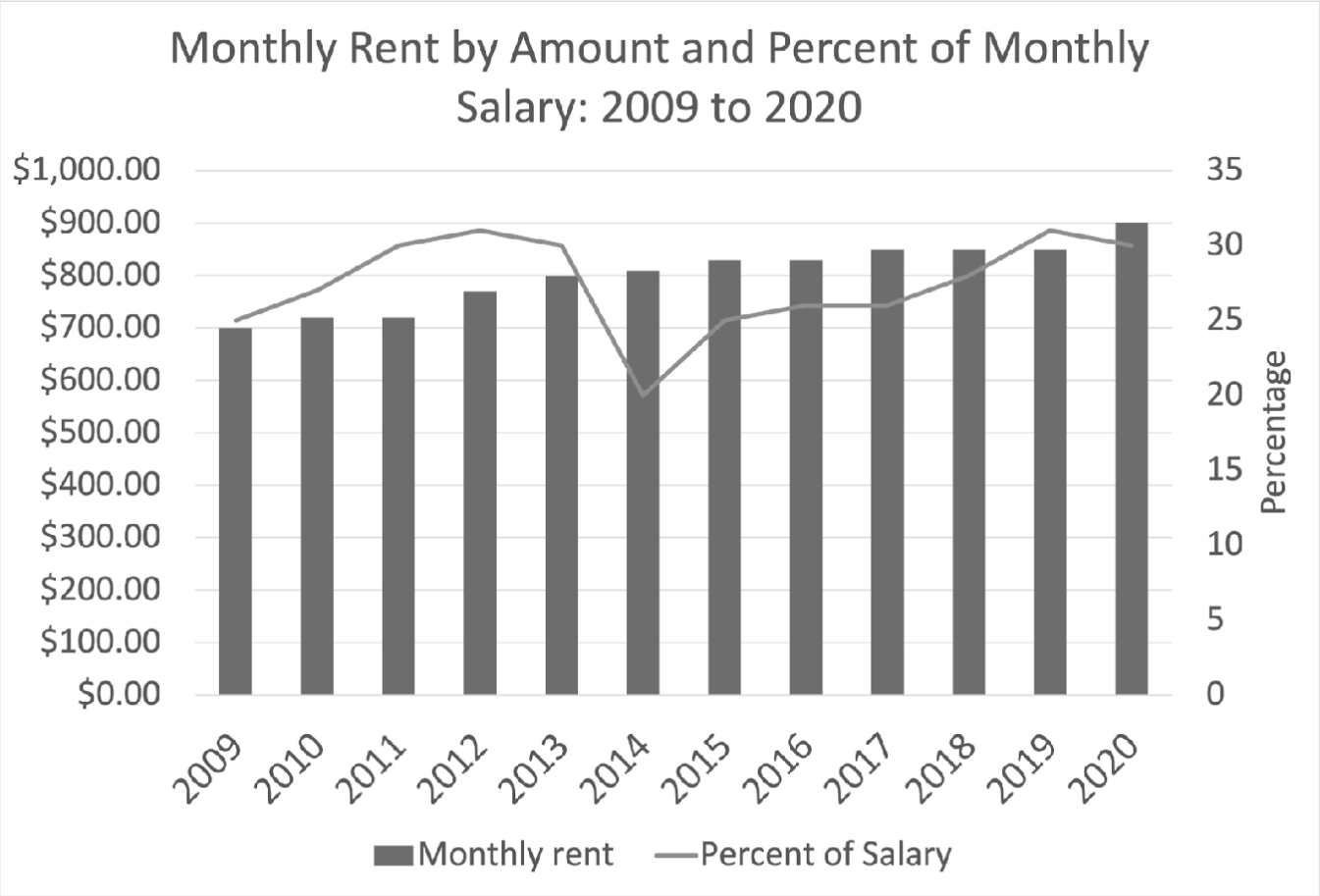
- https://www.maine.gov/rehab/accessibility_guide/presentation.shtml

You won't be tested on knowing any specific accessibility standards or best practices, but it's definitely something worth knowing as you enter a diverse workforce.

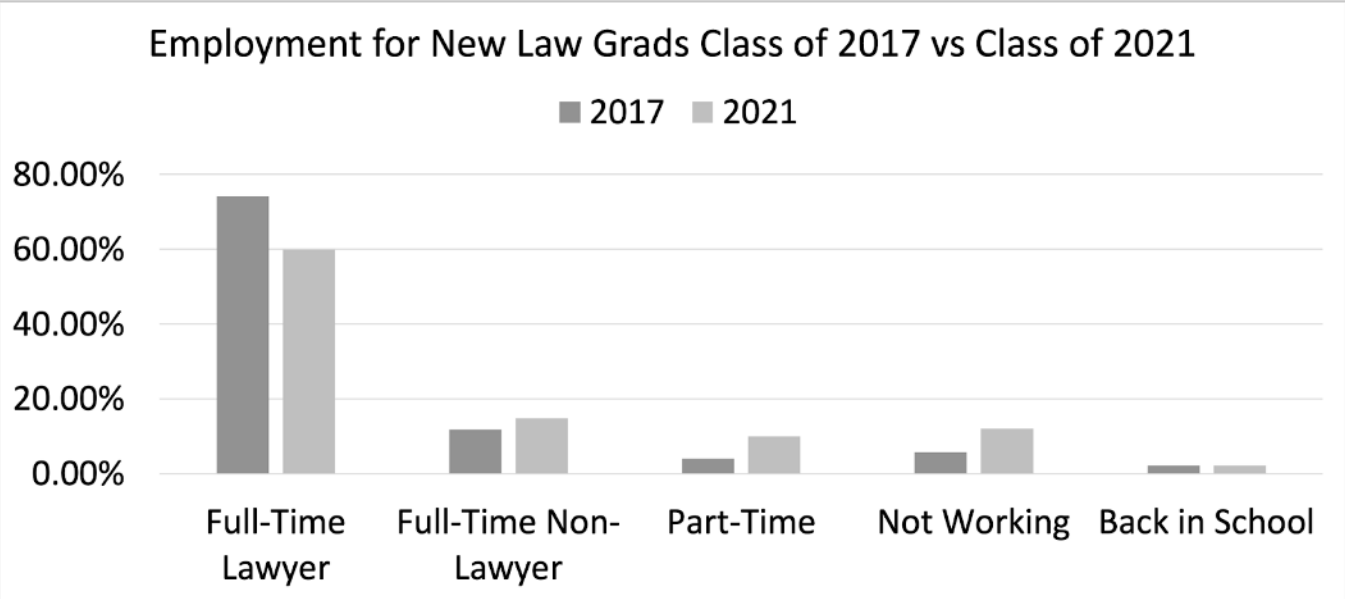
5. In the graphic below:
    1. approximately what percentage of the population was 65 years and older in 1960?
2. in which year did the population of 65 years and older exceed 20 million?



6. In the graph below, even though the rent stayed the same or increased every year in absolute terms from 2009 to 2020, we see a significant drop in the percentage of monthly salary spent on rent in 2014. Think of at least one possible explanation of this drop.



7. Below is a graph of how the 2021 graduates of law school are doing compared to the 2017 graduates. Both the 2021 and 2017 graduating class had about 44,000 students. What percentage of the full time lawyers in this study graduated in 2021?



	Full-Time Lawyer	Full-Time Non-Lawyer	Part-Time	Not Working	Back in School
2017	74.10%	11.80%	4.10%	5.80%	2.30%

2021	59.80%	14.90%	10.00%	12.10%	2.20%

After Preparation S.1 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration S.1, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
interpret the meaning of percentages of a whole.	
compute percentages of a whole.	
enter data into the cells of a spreadsheet.	

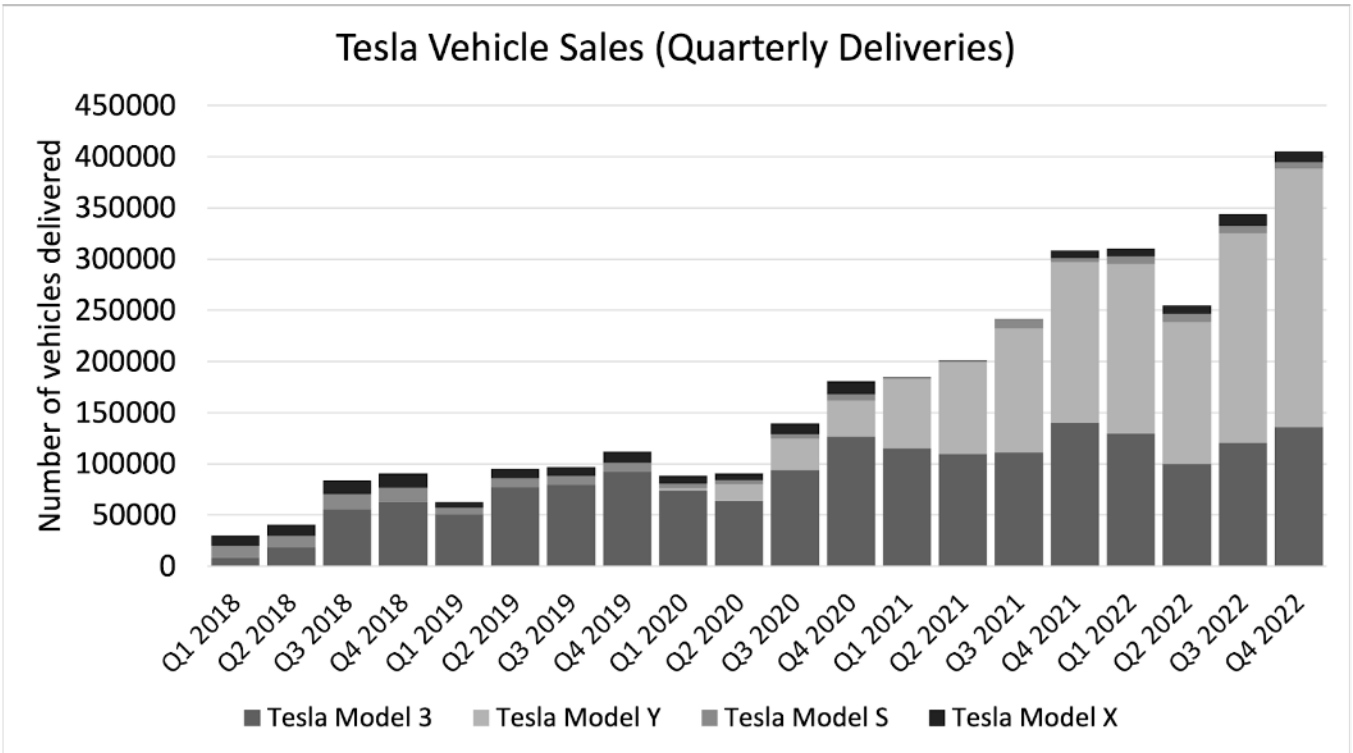
[1] https://en.Wikipedia.org/wiki/Accessibility

This page titled 3.1.1: Preparation S.1 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.1.2: Exercise S.1

The stacked bar chart below (Figure 4) shows the sales of the different models of Tesla electric vehicles by quarter (a fourth of a year, often used by businesses).[5] Use this chart to answer **Questions 1–5.**

Figure 4: Stacked Bar Chart for Tesla Vehicles Sales by Quarter



1. How many full years are covered by this stacked bar chart?
2. Is it easier to read the sales trends for a particular model over another in this chart? Why?
3. What overall story is this chart telling us about how sales of particular models compare over time?
4. Would it be easy for you to convert this to a non-stacked bar chart, using only this chart? Why or why not?
5. What would happen if a table were made using this chart, to show the trend in Model Y sales as a percentage of total sales?

1. The percentage would tend to decrease over time.
2. The percentage would tend to remain the same over time.
3. The percentage would tend to increase over time.
4. There is not enough information. We need the actual sales numbers for each quarter.

The following information was provided by the Coalition for the Homeless. Refer to these data for **Questions 6–8**. (You are encouraged to actually go to the source of the data shown in the footnote and check it out. It is an excellent example of how data can be presented in different types of charts in an interactive, clear, and informative manner.)

New York City's Homeless Shelter Population, September 2022:[6]

- Total number of homeless people in municipal shelters: 60,252
- Children: 19,310
- Adults in families: 20,178
- Single adults: 20,764

6. What percent of the people in homeless shelters in New York City in September 2022 were children? Round to the nearest whole percent.

7. What kind of chart would best compare the proportion of homeless children, adults in families, and single adults in New York City's homeless shelters in September 2022? Explain your reasoning.
8. Using Google Sheets (or another spreadsheet application), create the chart you chose in the last problem to visually compare the three populations: homeless children, adults in families, and single adults in New York City's homeless shelters in September 2022.
9. Create two bar charts in Google Sheets (or another spreadsheet application) to compare the number of kids and the number of adults who slept in New York City homeless shelters each night during September 2022 across four years. One chart should have two bars for each year (years along the horizontal axis). The other chart should have the four bars for kids grouped together on the left, then the four bars for adults grouped together on the right.

The table below summarizes additional information from the Coalition for the Homeless.

Month, Year	**Average Number of Homeless Children Each Night in Shelters in New York City**	**Average Number of Homeless Adults Each Night in Shelters in New York City**
September, 2019	22,083	40,308
September, 2020	18,745	38,507
September, 2021	14,946	32,970
September, 2022	19,310	40,942

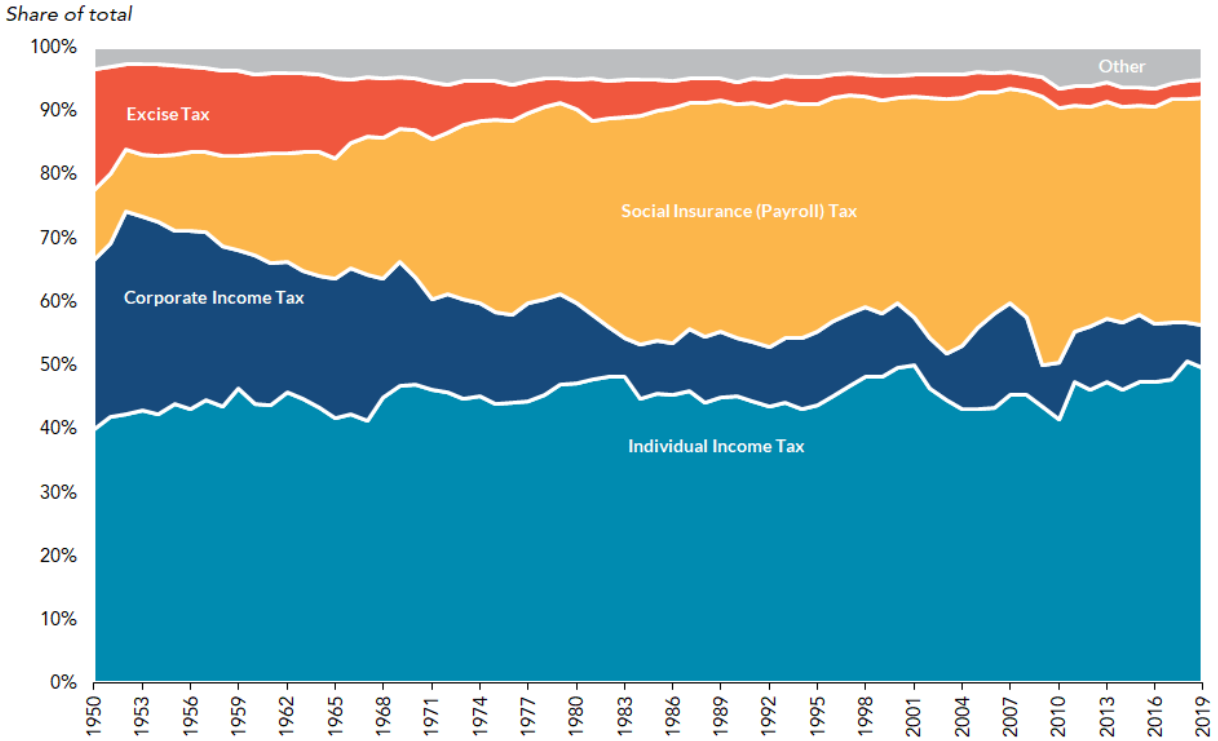
1. Are the two bar charts telling the same story? What stories are they telling? Explain your reasoning.
2. What are the advantages and disadvantages to each format? Explain your reasoning.
10. There are many types of charts and graphs that were not covered in the collaboration. One fascinating chart that combines aspects of bar charts and pie charts is called a **percentage area chart**. In this type of chart, areas are stacked (like a stacked bar chart) and scaled so that they always stack up to 100% (just as the slices in a pie chart must total 100% of the pie). Consider Figure 5 below, which shows federal tax revenues from 1950 to 2019.[7]

Figure 5: Sources of Federal Revenue from 1950–2019

FIGURE 4

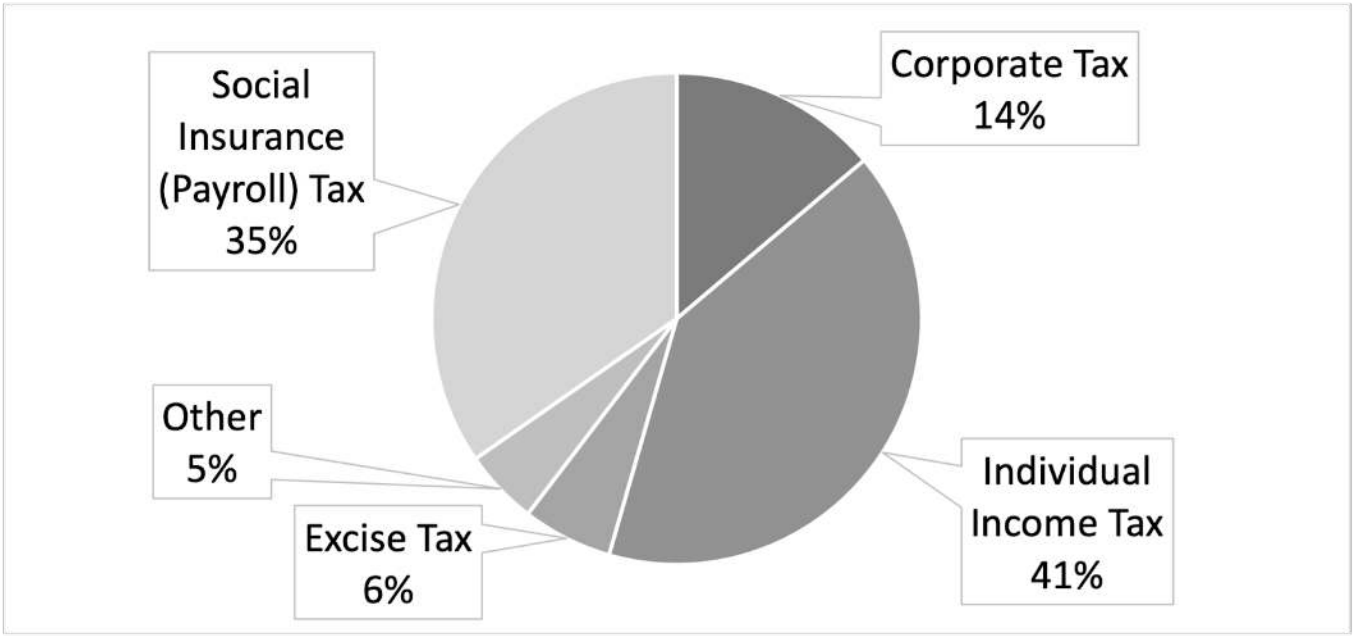
Sources of Federal Revenue

Fiscal years 1950 – 2019



Source: Office of Management and Budget. Historical Tables. Table 2.1, "Receipts by Source: 1934–2025."

To better understand it, consider the year 1975 and how the Excise tax segment above is represented in the following pie chart:



(a) What story does this chart tell about individual income tax revenue since 1950?

(b) What story does this chart tell about corporate income tax revenue since 1950?

(c) What story does this chart tell about payroll tax revenue since 1950?

Exercises are also available online. The online versions of Exercises contain one or more **Learning** sections and a **Questions** section. The **Learning** sections in online Exercises match what is in this workbook and no score is given for these sections. The **Questions** section in online Exercises contains similar questions, but the numbers and figures vary from your workbook, so they are good for extra practice. Completing the **Questions** section will generate a score.
Tip: You can use the print version or the **Learning** sections online for practice. Then you can do the **Questions** section for additional practice and a score.

---

[5] https://cleantechnica.com/2022/04/20/tesla-quarterly-sales-growth-in-7-charts/

[6] https://www.coalitionforthehomeless.org/facts-about-homelessness-2/

[7] https://www.taxpolicycenter.org/briefing-book/what-are-sources-revenue-federal-government

This page titled 3.1.2: Exercise S.1 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.2: Measuring Success

INTRODUCTION

Please take a moment to discuss the following questions in your group:

- What does it mean to be a successful student in general?
- What does it mean to be successful in this course?

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

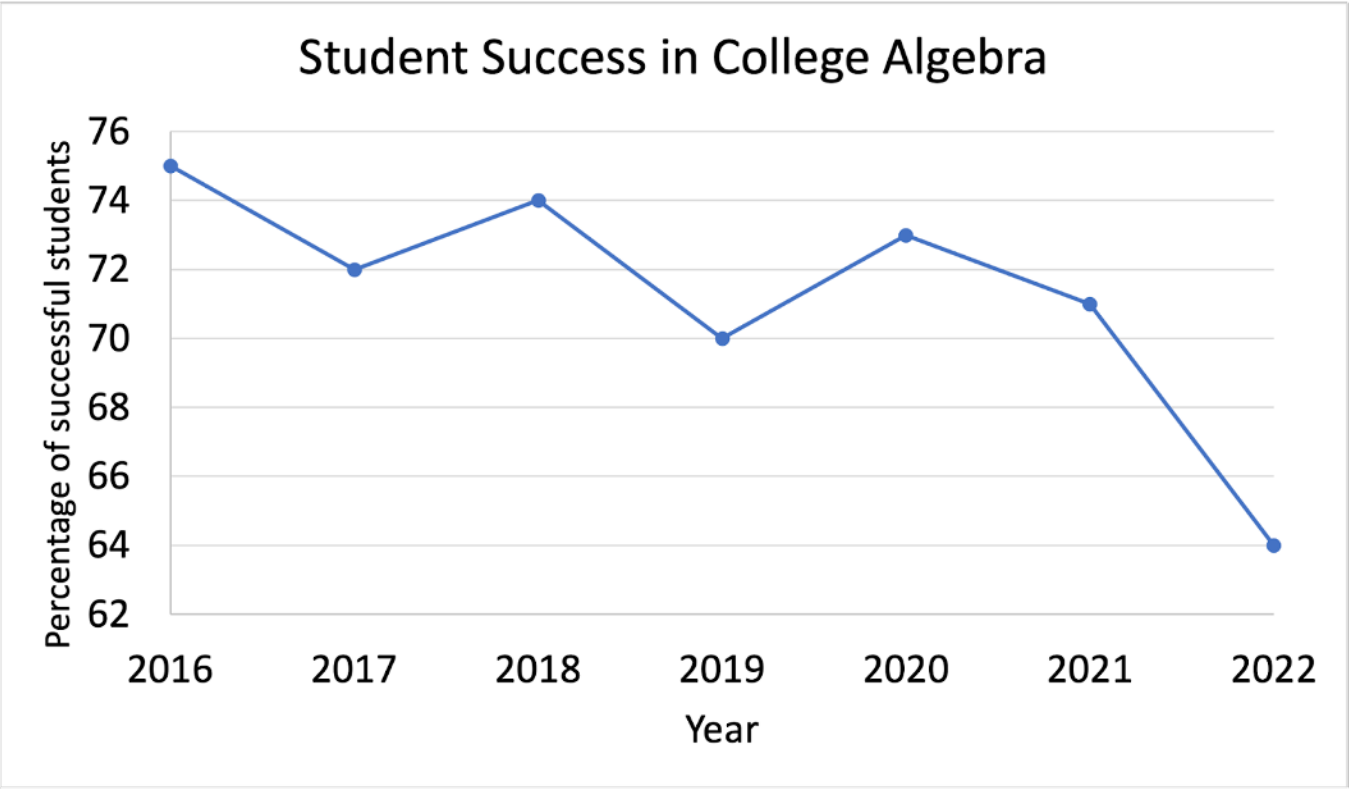
- the vertical scale of a graph can drastically affect the visual interpretation.
- reading and understanding axis labels on graphs is vital to their correct interpretation.
- graphs involving quantitative data can be biased depending on subjective categorical definitions.
- populations consist of all elements or individuals of interest.
- samples are subsets of populations.

By the end of this collaboration, you should be able to

- apply a short checklist of issues to consider when critically evaluating a graph.
- identify the research question, population, and sample in a statistical study.

PROBLEM SITUATION: TRACKING STUDENT SUCCESS IN COLLEGE ALGEBRA

A student body president, Jill, was concerned about student success in her university's College Algebra class. She obtained a report from her university's institutional research office which contained a graph summarizing the success of all students in multiple sections of College Algebra at the university. The graph is shown below. Jill wanted to study the data to determine if there really was a problem with student success in College Algebra at her university. Jill knew that many colleges and universities are developing alternative course designs and mathematics pathways to increase student success in mathematics.



(1) Look carefully at the graph that Jill obtained. Discuss in your group any trends and patterns that you see in the data. Use the Writing Principle to record your group's thinking.

(2) Some people argue that the way data is presented can affect how it is interpreted. Do you agree with this statement? Could the same data be presented in a way that might drastically affect its interpretation?

(3) Let's say that you want to ask two or three questions that could help you improve your understanding of students' success in College Algebra at this university. What would your questions be? Take a minute to think of some questions on your own first before sharing in your group.

Be aware of the vertical scale!

(4) Recreate the "Student Success in College Algebra" graph using a vertical scale that starts at 0% and ends at 100%. Compare your graph with the original. How does the new vertical scale change the possible interpretations of the data? What are some advantages and disadvantages your new graph has over the original one given to Jill?

Consider Percentages, Actual Values, and Population Size!

We often gather information on an entire group of individuals or objects of interest to us. A group such as this is known as the **population**. The total number of members in the population is the **population size**. Sometimes we cannot gather information on every member of the population, so we gather information on a **sample**, which is a subset of the population. The number of members in the sample is called the **sample size**.

For example, a team of researchers is interested in studying the height of all people in the United States. The population would be all people in the U.S. and the population size would be approximately 322 million people. But it would be impractical to measure the heights of 322 million people, so the researchers could measure the heights of a sample of 1,000 people.

Now let's reconsider Jill's concern about student success in her university's College Algebra class. The graph given to Jill is missing many important details, such as the number of students enrolled in College Algebra at the university in each given year. The total number of students enrolled in College Algebra at the college each year is a population size.

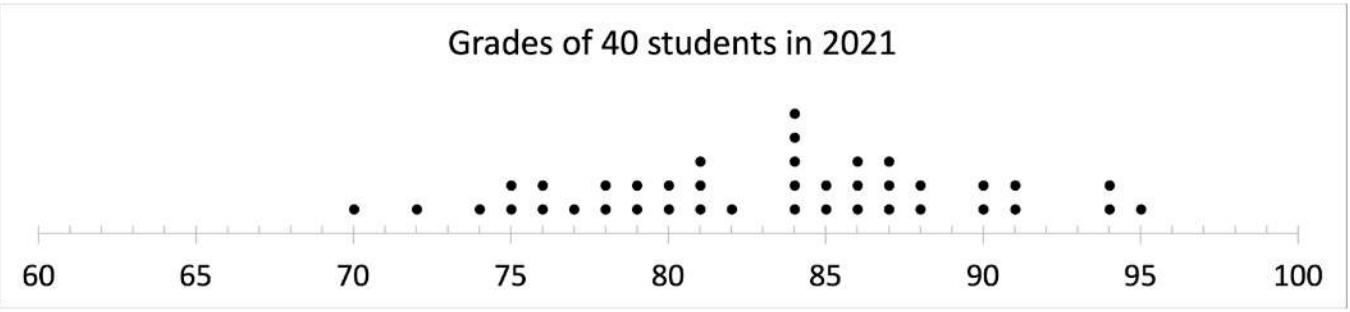
(5) Suppose that the "Student Success" graph presented data from a population of 40 students each year. In this case, would you consider the decline in student success shown in the graph between 2021 and 2022 to be significant? Explain. (**Hint**: First calculate how many students in 2021 and 2022 were successful.)

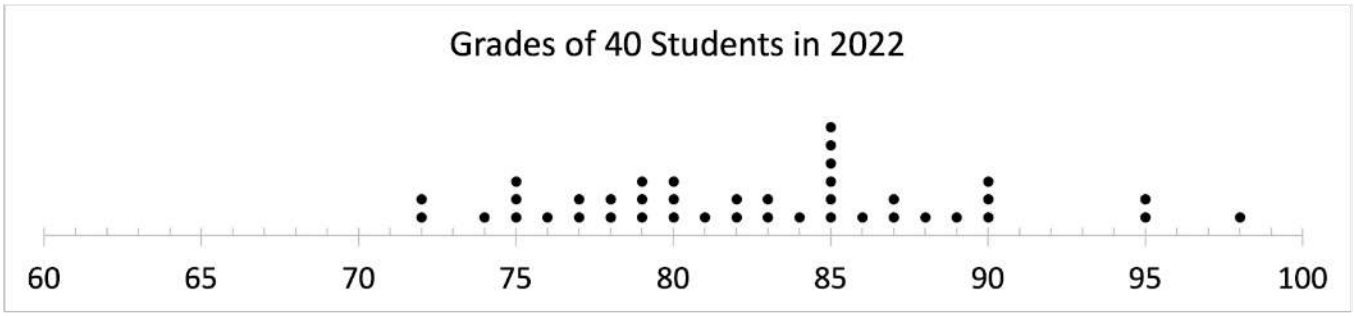
(6) Let's say that the data came from analyzing classes that typically had a combined enrollment of about 200 students (i.e., population size is 200 students each year). In this case, would you consider the decline in student success shown in the graph between 2021 and 2022 to be meaningful? Explain. (**Hint**: First calculate how many students in 2021 and 2022 were successful.

Be Wary of Inexact Terms!

Another way to display quantitative information is a **dotplot**. Dotplots are graphs in which dots represent each bit of data and are placed appropriately along a horizontal axis.

(7) The following two dotplots show hypothetical course grades (as percentages) for 40 students in 2021 and 40 students in 2022.





1. If “success” is defined as receiving 80% or higher, what percentage of the students would be considered successful each year? How does this compare with the success rate in the original graph given to Jill?
2. If “success” is redefined as receiving 74% or higher, what percentage of these students would be considered successful each year?

2021:

2022:

3. In the original graph given to Jill, “success” was not defined. Now that you have seen Questions 7(a) and 7(b), explain why she may wish to learn how “student success in College Algebra” was defined.

MAKING CONNECTIONS

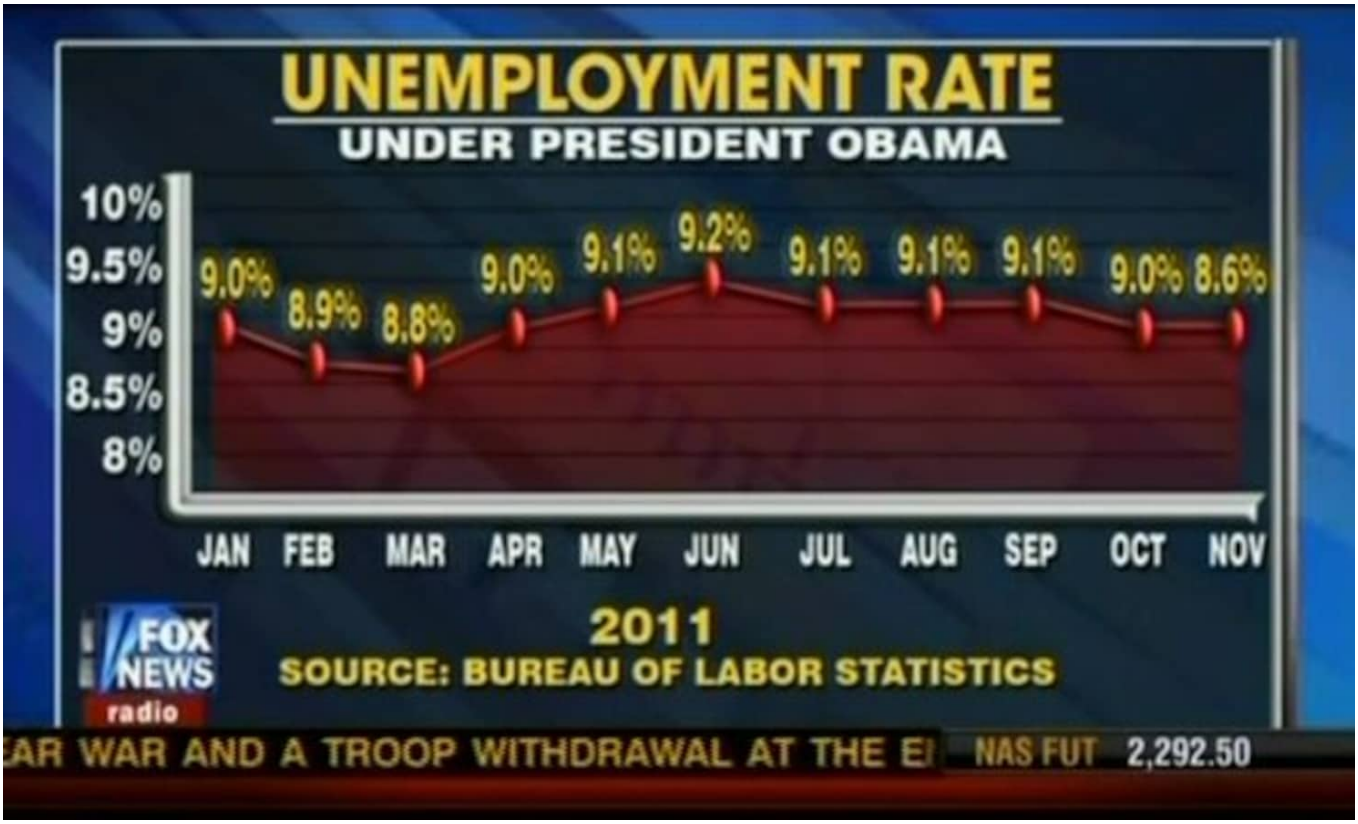
Record the important mathematical ideas from the discussion.

This page titled 3.2: Measuring Success is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.2.1: Preparation S.2

A **line graph** is another way to display quantitative information. Line graphs display data as a series of points, which are connected by line segment.

(1) In December of 2011, Fox News showed the following line graph[8] of the unemployment rate in the U.S. for 2011:



(a) According to this graph, what was the unemployment rate for June 2011?

(b) How did the unemployment rate in November 2011 compare with the unemployment rate at the beginning of the year? Compare using both absolute and relative change.

(c) Is the data accurately graphed? If not, what is wrong with the presentation of the data?

(d) Who is the source of this data?

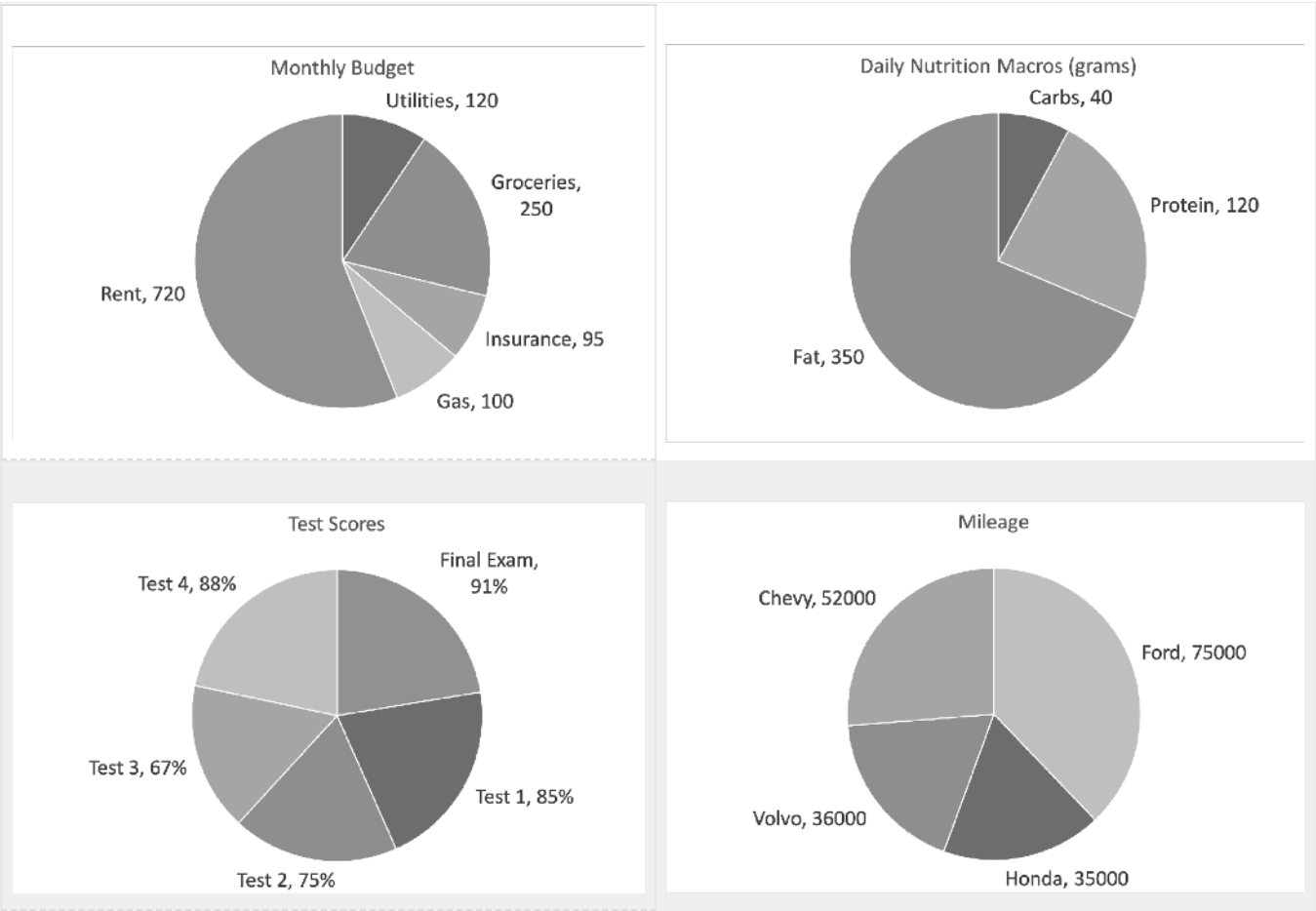
i. Fox News.
ii. Bureau of Labor Statistics.
iii. Unemployment Agency.

(2) The following table shows estimates of U.S. troop levels in Afghanistan in January of each year from 2001 through 2021.[9] What type of graph or chart do you think would best represent this data visually? Use a spreadsheet application of your choice to create that visual representation.

Year	Troop Level
2001	2,300
2005	19,500
2006	21,500
2007	25,240
2008	30,051

2009	60,065
2010	101,205
2011	102,077
2012	81,174
2013	63,673
2014	33,186
2015	12,802
2016	12,489
2017	16,500
2018	14,000
2019	14,000
2020	8,000
2021	2,500

(3) Which of the following pie charts contains data that isn't suitable to be represented in a pie chart? Check all that apply.



i. Monthly budget
ii. Daily nutrition (grams)

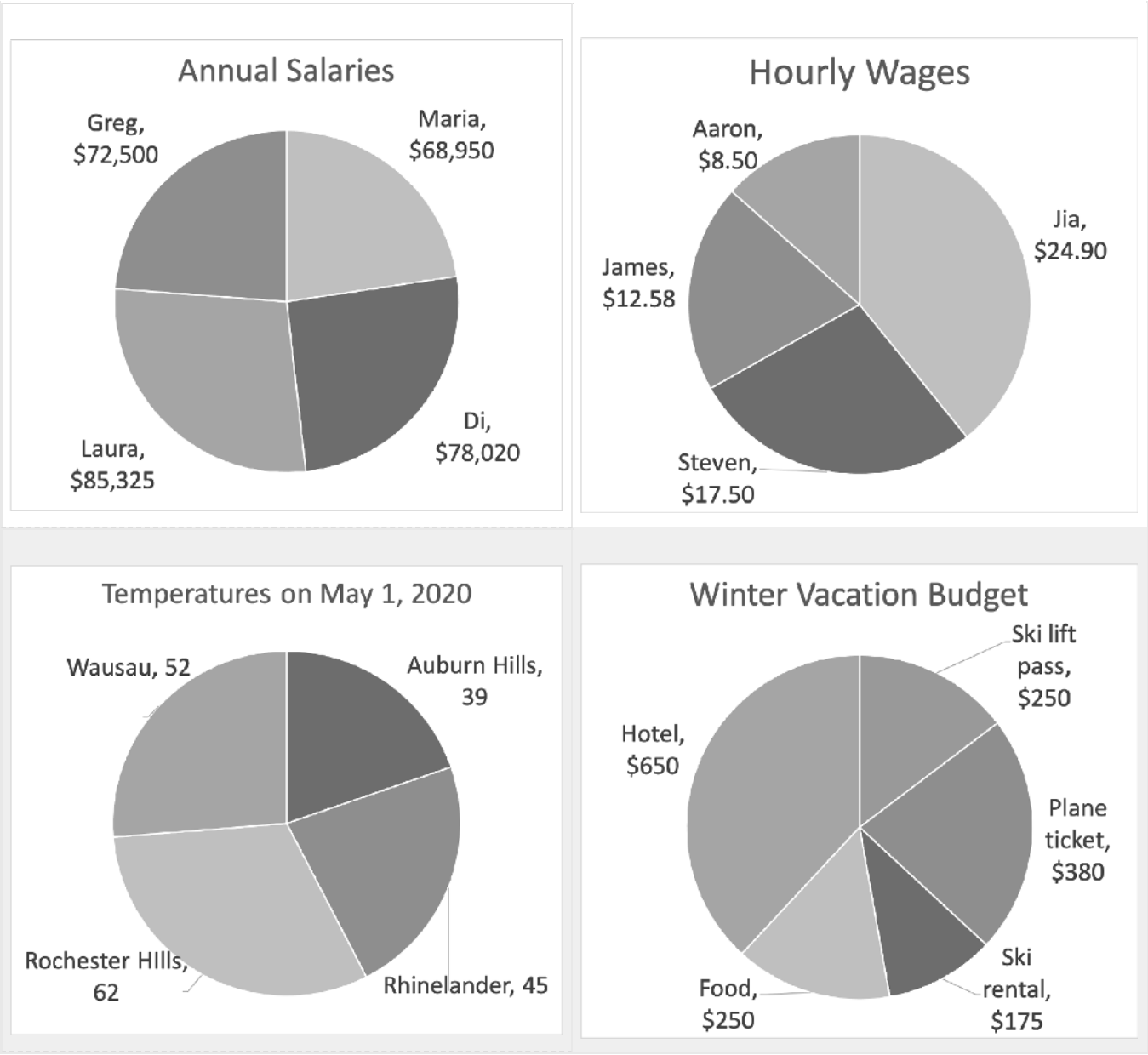
iii. Test scores
iv. Mileage

(4) Which of the following charts **incorrectly** represent the given data? Check all that apply.



i. Daily nutrition (grams)
ii. Daily sales
iii. Mileage
iv. Test scores

(5) Which of the following pie charts contains data that isn't suitable to be represented in a pie chart? Check all that apply.



i. Temperatures on May 1, 2020
ii. Winter vacation budget
iii. Hourly wages
iv. Annual salaries

After Preparation S.2 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration S.2, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
read information in a graph.	
use the correct language to describe changes in percentages.	

plot points on a graph with vertical and horizontal axes.	

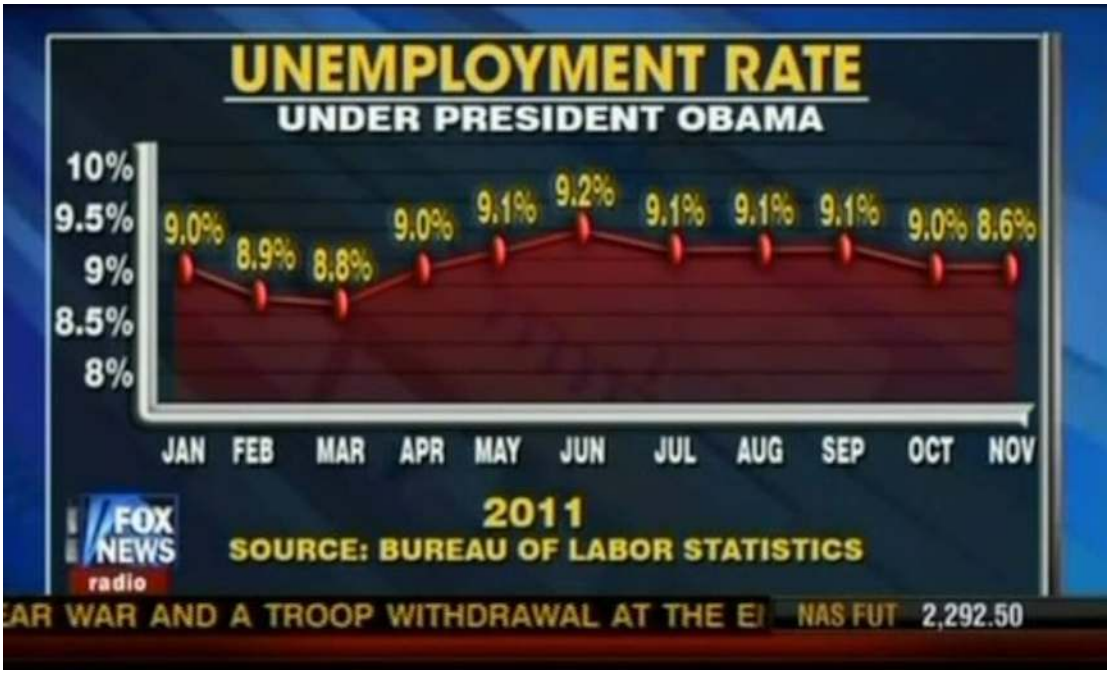
[8] http://mediamatters.org/blog/2011/12/12/today-in-dishonest-fox-news-charts/185162

[9] https://usafacts.org/articles/afghanistan-how-troop-levels-changed-over-the-course-of-americas-longest-war/

This page titled 3.2.1: Preparation S.2 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

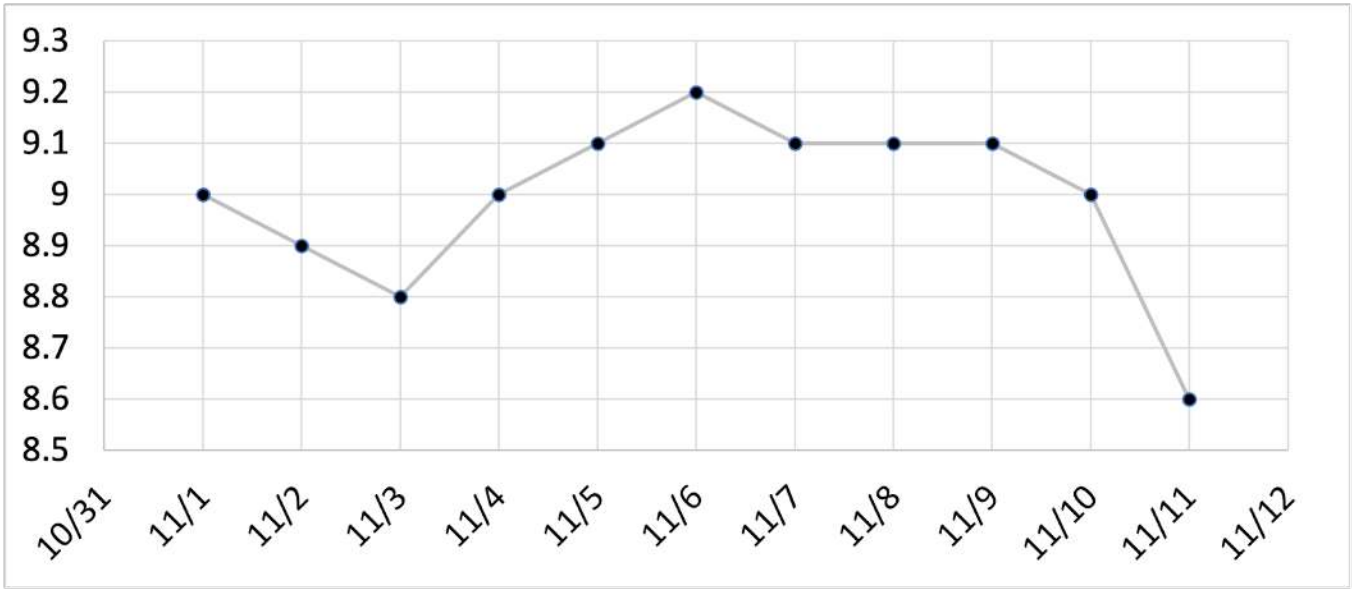
3.2.2: Exercise S.2

1. Recall from Question 1 in Preparation S.2 that, in December of 2011, Fox News showed a graph of the unemployment rate in the U.S. for 2011. This graph has been reproduced and criticized by many media outlets and blogs since then. Below is a screenshot of the graph that was posted by MediaMatters.org.



1. Why might this graph be criticized? (Hint: Review your answer from Preparation S.2 Question 1(c) and think about the concepts you learned in Collaboration S.2.)

(b) MediaMatters.org also posted a chart that looks like the one below, saying it was "… how an honest chart of the changes to the unemployment rate during 2011 looks." Compare and contrast this chart with the one attributed to Fox News.



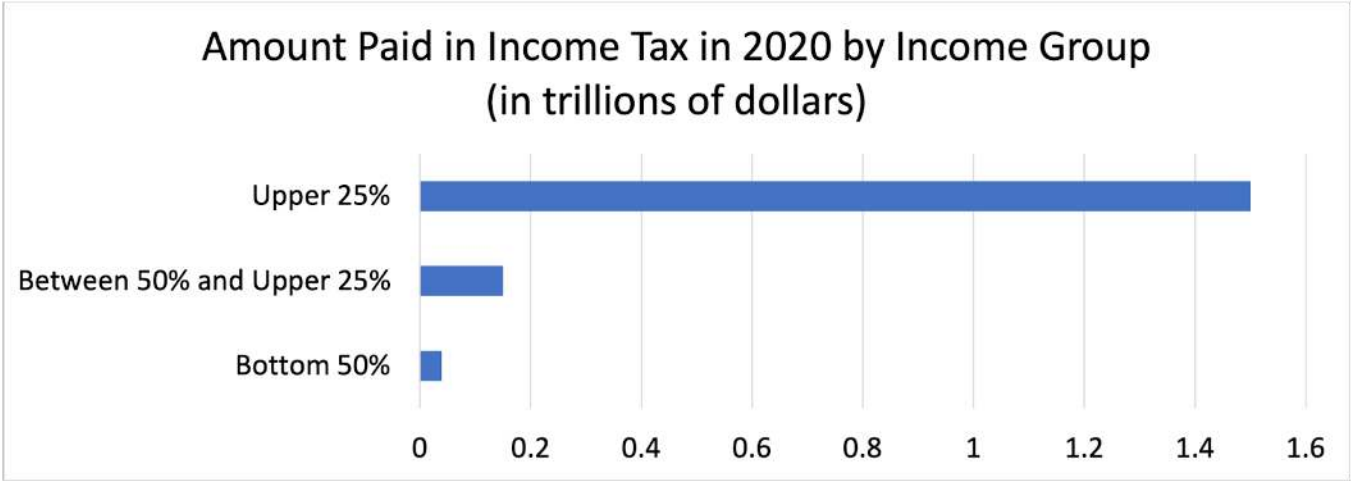
(c) If you wanted to present an honest and objective graph of the unemployment changes throughout 2011, how would you present the data? Create such a graph. Use the space below or a spreadsheet application.

(d) How would your graph change if the unemployment figures were all rounded to the nearest percent?

(e) Calculate the relative change in the U.S. unemployment rate from January 2011 to November 2011. Express the relative change as a percentage. Round to the nearest tenth.

(f) Using the U.S. unemployment rate in January 2011 and November 2011, create a linear model for the U.S. unemployment rate as a function of the number of months since January 2011. Let R be the U.S. unemployment rate (as a percent) and let t be the number of months since January 2011.

(2) The bar chart below shows the total amount of money paid in income tax in the U.S. in 2020 by taxpayers in three different income groups.[10] "Bottom 50%" refers to taxpayers with incomes that are equal to or less than the median income of all taxpayers in the U.S. "Between 50% and Upper 25%" refers to taxpayers whose incomes are greater than the median but less than incomes in the upper 25%. "Upper 25%" refers to individuals whose incomes are in the top 25% of all taxpayers' income.

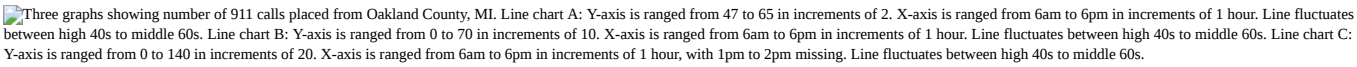


(a) Using the information in the above graph, critique the statement, "Americans in the Upper 25% income group pay more income taxes than Americans in the other income groups."

(b) Assuming that the information in the above chart is correct, does it allow one to easily compare the total amount of money paid in income tax by Americans in the different income groups? Explain carefully.

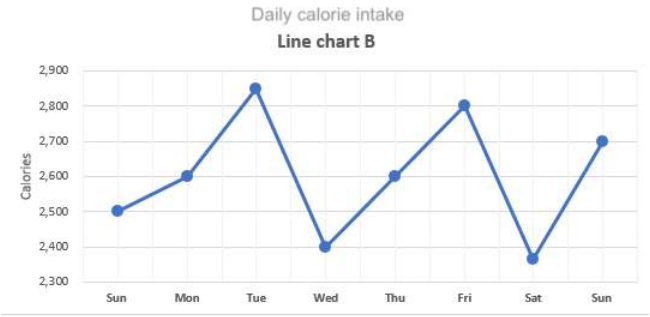
(c) What additional information, if any, could have been provided to help interpret the information being presented?

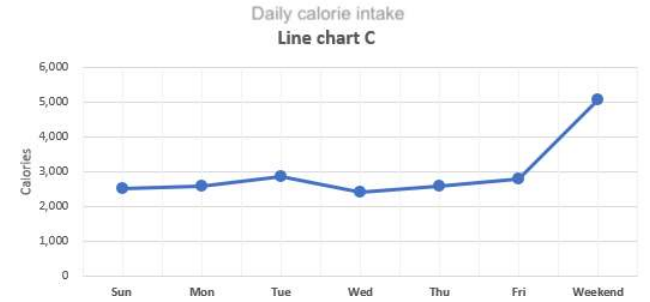
(3) Which of the following charts is misleading in how it represents the given data? Check all that apply.


Three graphs showing number of 911 calls placed from Oakland County, MI. Line chart A: Y-axis is ranged from 47 to 65 in increments of 2. X-axis is ranged from 6am to 6pm in increments of 1 hour. Line fluctuates between high 40s to middle 60s. Line chart B: Y-axis is ranged from 0 to 70 in increments of 10. X-axis is ranged from 6am to 6pm in increments of 1 hour. Line fluctuates between high 40s to middle 60s. Line chart C: Y-axis is ranged from 0 to 140 in increments of 20. X-axis is ranged from 6am to 6pm in increments of 1 hour, with 1pm to 2pm missing. Line fluctuates between high 40s to middle 60s.

(4) Which of the following charts is misleading in how it represents the given data? Check all that apply.







[10] https://taxfoundation.org/summary-latest-federal-income-tax-data-2023-update/

This page titled 3.2.2: Exercise S.2 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.3: Statistical Analysis Process

INTRODUCTION

Suppose you are asked to design a study to investigate whether students who attend colleges in rural, suburban, and urban locations in your state have differing views on the threat that climate change poses to the safety and well-being of people living in the United States. Think about how to design this particular study in a way that can be used for other similar studies. In other words, try to come up with a general step-by-step procedure that can be used to conduct similar studies.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- statistical investigations are a cycle of forming a research question, designing a study and collecting relevant data, data analysis, statistical inference, and drawing appropriate conclusions.
- appropriate data collection methods play an important part in the types of conclusions we can draw from the data.
- the sample in a study affects the population to which conclusions can be generalized.

By the end of this collaboration, you should be able to

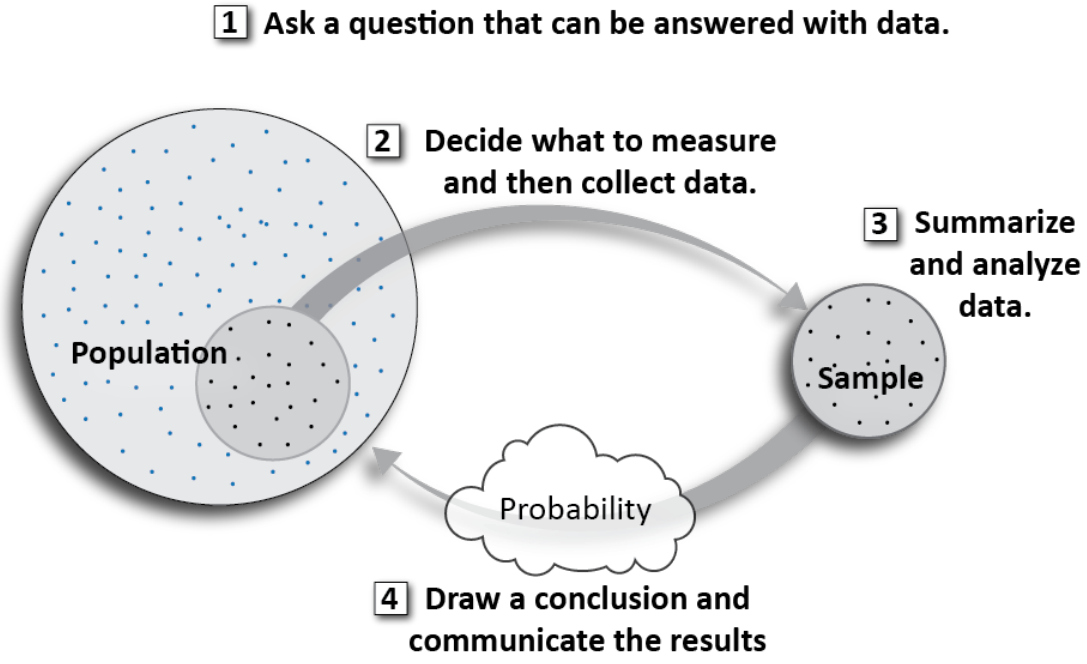
- identify the steps in the statistical analysis process given a description of a statistical study.
- distinguish between an observational study and an experiment.
- identify the explanatory variable and response variable in an experiment.
- understand the conclusions you can draw from observational studies and experiments.
- identify conditions when it is appropriate to generalize from sample to population.
- identify lurking variables and sources of bias in a particular statistical study.

PROBLEM SITUATION 1: EXAMINING SCIENTIFIC STUDIES

This collaboration introduces the statistical analysis process and enables you to examine how this process is applied in scientific studies. **Statistical analysis** is the process of using data to answer questions.

Recall that data are information that we collect from our world and consist of facts and observations. In statistical studies, often data from a sample are used to answer questions about a population. Recall that a population is the entire group of individuals or objects of interest to us and a sample is a subset of that population. Looking at sample data allows us to make **generalizations** about populations that may be large and difficult to understand.

We can think of the statistical analysis process in four steps.



Steps in a Statistical Analysis

Step 1: Ask a question that can be answered by collecting data.

Step 2: Decide what to measure and then collect data.

Step 3: Summarize and analyze the data.

Step 4: Draw a conclusion and communicate the results.

In Preparation S.3, you read two press releases from the McLean Hospital website that described recent scientific studies. The following questions ask you to determine how the statistical analysis process was applied in these scientific studies.

Consider **STUDY 1** (from Preparation S.3): "Brain Imaging Study Suggests Long-Term Steroid Use Can Lead to Significant Brain Structural and Functional Abnormalities."

1. Identify the samples and populations in this study.
2. In the table below, identify each step of the statistical analysis process for this study.

Steps in Statistical Analysis	**Brain Imaging Study**
1. Ask a question that can be answered by collecting data.	
2. Decide what to measure and then collect data.	
3. Summarize and analyze the data.	
4. Draw a conclusion and communicate the results.	

Consider STUDY 2 (from Preparation S.3): "Researchers See Promise in Potential Depression Treatment."

3. Identify the sample and population in this study.
4. In the table below, identify each step of the statistical analysis process for this.

Steps in Statistical Analysis	**Depression Treatment Study**
1. Ask a question that can be answered by collecting data.	
2. Decide what to measure and then collect data.	
3. Summarize and analyze the data.	
4. Draw a conclusion and communicate the results.	

The two studies above provide examples of the two main types of studies used to collect data: **observational studies** and **experiments**.

Observational Study

In an *observational study*, researchers observe subjects in a sample to learn about population characteristics. Researchers usually observe a sample of the population, since it is often impossible to obtain information from every member of the population. Because the goal of an observational study is typically to learn about the population, it is important that the sample be representative of the population. For example, if your population of interest has 50% females and 50% males, your sample should include roughly the same proportion of males and females. This way, if the variables you are observing turn out to be different between males and females, you will have data for both groups.

Experiment

An *experiment* is used to answer questions about how one variable responds when another variable is manipulated.

- The variable that is actively manipulated by the researcher is the **explanatory variable**.
- The variable in which changes are then measured is the **response variable**.

Since the goal of an experiment is to learn about the effect of the different treatments (changes to the explanatory variable) on the response variable, it is important to apply the different treatments to similar samples, which should be representative of the

population. The results of the experiment only apply to a larger population if the samples are representative of that population.

Often, the experiment is conducted by taking two samples (groups).

- The first group is called the **experimental group**, or **treatment group**, for which an explanatory variable is actively manipulated (treatment received).
- The second is called the **control group**, for which the same variable is not actively manipulated.

All other conditions of the experiment (except the manipulated explanatory variable) are the same for both groups.

Key Difference between an Observational Study and an Experiment

An important difference between these data collection methods is that, in an experiment, the explanatory variable is *manipulated* and changes in the response variable are examined. Often observational studies may be chosen due to ethical or logistical considerations. For example, suppose researchers wanted to determine the effects of long-term smoking on female health. It would be unethical to make subjects in an experiment start smoking to determine the effect of smoking on health, and it could take decades to get results. An observational study would avoid these issues by studying a sample of females who are already long-term smokers.

Let's again consider **STUDY 2** (from Preparation S.3): "Researchers See Promise in Potential Depression Treatment."

5. Using the information above, discuss in your group whether this study was an experiment or an observational study. If the study is an observational study, identify the variable of interest. If the study is an experiment, identify the explanatory variable and response variable.
6. This study examined results from two groups: an experimental (treatment) group and a control group. Identify these two groups in the study. What was the purpose of incorporating the control group?

Cause-and-effect and Non-causal Relationships

Another important difference between these data collection methods is the type of conclusions they generate. Statistical studies can yield two main types of conclusions about the relationship between the variables of interest: they can conclude there is a **cause-and-effect** relationship (causal) or they can conclude that there is a **non-causal** relationship.

The type of conclusion that is possible depends on the type of study that is performed. Experiments can generate causal inferences about two variables because the other variables are being controlled, so any change in the response variable can be attributed to the explanatory variable. Observational studies can only show non-causal relationships between variables, where one can conclude that two variables are related, but cannot conclude that one directly causes the other.

It is important to note that, regardless of the type of conclusion, researchers often want the conclusion to generalize from the sample to the entire population. The table below summarizes the types of conclusions from each study and when generalizing conclusions to the population is reasonable. The next unit, S.4, will discuss how to create a study with these conditions.

Type of Study	**Type of Conclusion**	**Can Generalize from Sample to Population When...**
Observation	Non-causal relationship	• The sample(s) is representative of the population.
Experiment	Cause-and-effect relationship (causal)	• Each sample is representative of the population. • The experimental and control groups are similar.

Let's again consider **STUDY 2** (from Preparation S.3): "Researchers See Promise in Potential Depression Treatment."

7. Is it appropriate to establish a cause-and-effect conclusion from this study? Describe the conditions of the study that support this type of conclusion.

PROBLEM SITUATION 2: EXAMINING NON-CAUSAL RELATIONSHIPS

Let's now discuss the final article you read in Preparation S.3, **STUDY 3**: "Abuse as a Child Linked to Longer Term Homelessness."[16]

8. Does the article suggest that a relationship exists between the variables in the study? (**Hint**: You identified the variables in Preparation S.3 Question 6.)
9. Does the design of this study enable us to conclude that violence or abuse suffered as a child causes long-term homelessness? Explain.

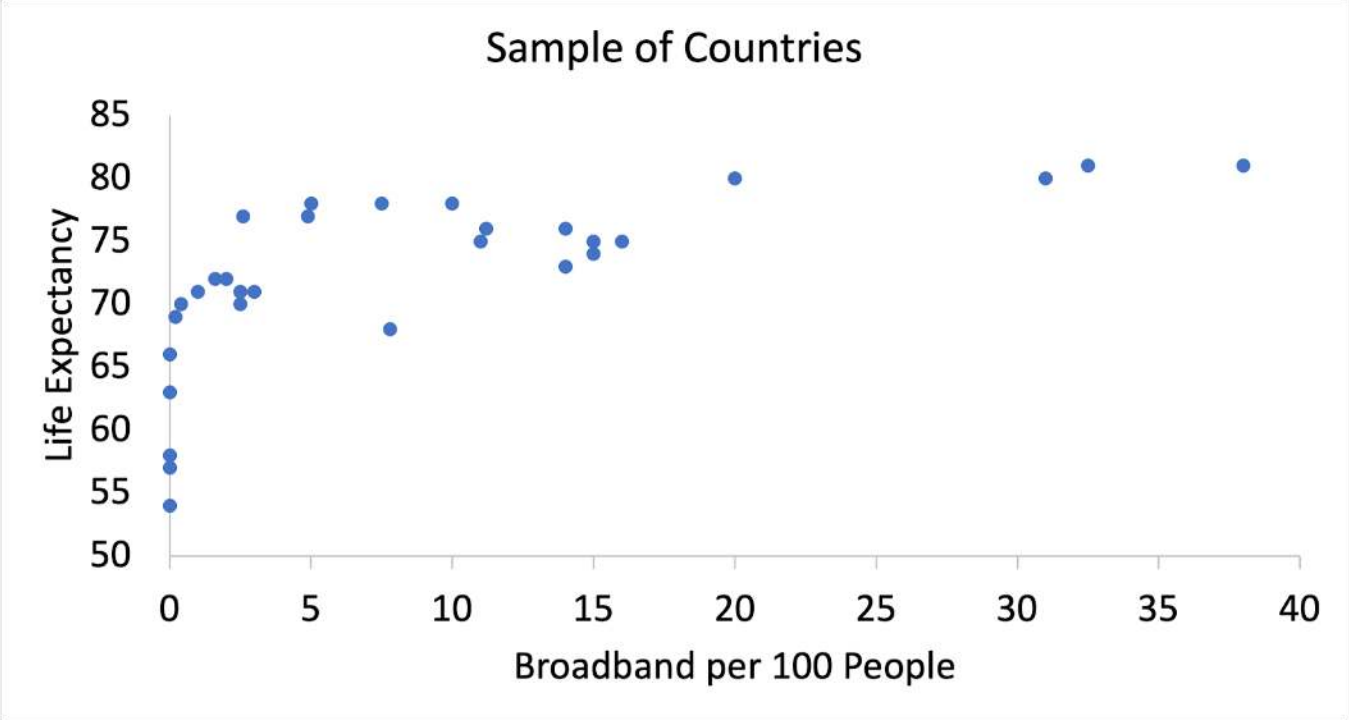
Correlation vs. Causation

A common conversation in statistics occurs around "correlation vs. causation." **Correlation** is a statistical measure of the strength of a relationship between two quantitative variables. A correlation simply implies that two variables are related, but does not imply that one causes the other. Thus, when there is a non-causal relationship between quantitative variables, we say there is a **correlation** between these variables. Note that two quantitative variables with a cause-and-effect relationship, such as exercise intensity and heart rate, can have a strong correlation; however, a correlation alone does not prove causation. Consider correlation and causation in the following question:

10. Suppose a student was interested in the relationship between two variables:

- the proportion of people who have broadband Internet service in a country, and
- life expectancy in that country (in years of age).

Broadband Internet service is generally high-speed Internet access. The scatterplot below shows the relationship between these two variables in a random sample of 30 countries. Each point represents a country. The x-coordinate of the point is the broadband Internet rate (number of people per 100 in the country who have broadband Internet service); the y-coordinate is the average life expectancy of people in the country.



(a) What type of relationship exists between the two variables? Explain.

(b) Why is it unreasonable to conclude that access to broadband Internet service *causes* citizens in the country to live longer? What other factors might explain this relationship?

Lurking variables

There is an important type of variable that influences how we must interpret non-causal relationships, such as the one in Question 10, called a **lurking variable**. A lurking variable is a variable that was not considered, but directly affects both variables being studied. A lurking variable makes it seem as though the variables are directly related. However, without the lurking variable, the variables being studied would not be related, would be related to a lesser degree, or would even be related in the opposite direction.

For example, in the study described in Question 10, the wealth of each country is a lurking variable, the idea of which you may have discussed in part (b). The wealth of a country would lead to an increase in both life expectancy and broadband rates. It may be this lurking variable that creates the appearance of a relationship between the two variables being studied.

Another commonly used example is the *positive* correlation between the number of firefighters at a fire and the amount of damage caused by the fire. One would certainly not want to interpret this correlation as, "more firefighters at a fire cause more damage." Instead, this correlation is a non-causal relationship that is attributed to a lurking variable—that is, the strength of the fire. As a fire gets stronger, more firefighters come and more damage is caused by the fire.

A positive correlation indicates that as one variable increases, the other increases; a negative correlation indicates that as one increases, the other decreases.

Bias

One other factor that should be considered when reading and interpreting statistical studies is **bias**. A study is said to suffer from **bias** if its design tends to favor certain results. For example, suppose a study found that 45% of pedestrians in Seattle drink coffee while they are out walking. If it turned out that this study observed walkers in downtown Seattle at 7:30 a.m., then we might wonder whether the results would be different if the study had been conducted at a different time of day. The time when the study was conducted could have created a bias, even if it was unintentional.

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

---

[16] http://www.smh.com.au/federal-politics/political-news/abuse-as-a-child-linked-to-longer-term-homeless-20120418-1x7hx.html

This page titled 3.3: Statistical Analysis Process is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.3.1: Preparation S.3

The next unit explores the statistical analysis process through the perspective of clinical studies and sociological studies. The three studies introduced in this preparation will be used again during the collaboration, and you will be expected to be familiar with them during Collaboration S.3.

Begin by reading the press release below about a brain imaging study:

STUDY 1:

Brain Imaging Study Suggests Long-Term Steroid Use Can Lead to Significant Brain Structural and Functional Abnormalities

Belmont, MA - Results of the first systematic brain imaging study conducted on long-term users of anabolic-androgenic steroids reveal significant brain structural and functional abnormalities, according to doctors at McLean Hospital.

An article detailing their findings, “Brain and Cognition Abnormalities in Long-Term

Anabolic-Androgenic Steroid Users,” was published last month in the journal *Drug and Alcohol Dependence*.

“Anabolic steroids are the newest of the world’s major forms of drug abuse,” said Harrison G. Pope Jr. MD[11], co-director of the Biological Psychiatry Laboratory[12] at McLean, professor of psychiatry at Harvard Medical School, and co-author of the article. “This is a public health problem that is just coming over the horizon now and we’re only beginning to understand its consequences. Widespread use of steroids did not appear until the 1980s in the United States, so this is a very young form of substance abuse in contrast to other drugs like marijuana and opiates that have been around for thousands of years.”

And while professional athletes make the headlines for steroid use, he said, that is not where the abuse is rampant.

“Professional athletes account for a tiny fraction of steroid users,” said Pope. “The great majority of steroid users are not competitive athletes at all, but are rank-and-file guys at the gym who just want to get big. This is one of the reasons that the problem has not been widely recognized, because the public has been focused on the occasional high-profile athletes who have been found to take steroids.”

There are approximately three million American men who have used steroids at some time in their lives, he said. From that number, it is estimated that one million have developed steroid dependence where they continue to take steroids for long periods of time, often in spite of the adverse effects, and this number is likely growing.

While studies have been published by Pope and others on the cardiovascular, endocrine, and cognitive effects of long-term steroid use, this is the first time that doctors have been able to systematically examine changes in the brain from long-term steroid use, said Marc J. Kaufman, PhD, director of the Translational Imaging Laboratory19 at McLean, associate professor of psychiatry at Harvard, and also one of the article’s co-authors.

“We used three types of imaging methods to examine different parts of the brain, including magnetic resonance imaging (MRI), which tells you about brain structure; functional magnetic resonance imaging (fMRI), which measures brain activity by detecting changes in blood oxygenation; and magnetic resonance spectroscopy (MRS), which looks at levels of different chemicals in the brain.”

The study ran from the fall of 2013 to the summer of 2014, using subjects drawn from a pool of 150 male weightlifters ages 35 to 55, with the average age being 42. Brain images of those reporting long-term use of steroids were compared with images from non-users.

One of the key findings with the structural imaging data, said Kaufman, centered on the amygdala, the part of the brain related to emotion regulation, aggression, anxiety, and possibly even depression, which was enlarged in those who were chronic steroid users as compared to the non-users.

“The amygdala in the steroid users wasn’t just a little bit larger than the control group, it was more than 20 percent larger,” said Kaufman. “That is a really marked difference, and it is particularly interesting because previous studies have shown that amygdala enlargement has been associated with aggression among other types of substance abuse populations.”

The study uncovered another abnormality among steroid users: a decrease in the brain levels of a sugar known as scyllo-inositol, which plays a key role in preventing neurotoxic proteins from clumping, particularly beta-amyloid protein which is known to play a role in the development of Alzheimer’s disease and other dementias.

"Clumping of beta-amyloid protein basically gums up the pipes in cells and leads to early cell death. The levels of this protective scyllo-inositol were significantly lower in steroid users," said Kaufman. "If scyllo-inositol depletion is a widespread phenomenon among steroid users," he said, "we could be seeing more cases of dementia or Alzheimer's in these men as they age."

Said Pope, "Another reason for concern is that unlike alcohol and most illicit drugs, which leave your body fairly quickly after use, steroids can be present in the system at levels 20 to 40 times greater than natural testosterone levels, 24 hours a day, seven days a week, for the entire duration of a course, or so-called 'cycle,' of steroid use."

"The body can never be rid of them because users are typically injecting long-acting steroids regularly throughout their cycle," he said. "As a result, there's never a break. The body has no moment to rest from the effects. That is quite different from most drugs that are abused. And that's why we worry that steroids might have unique forms of toxicity that we don't typically see with other drugs of abuse."

Pope also pointed out that steroids are not addictive in the sense that alcohol or drugs are addictive, but users develop a different type of dependence syndrome. He gave three reasons for steroid dependency:

- Body image: Weightlifters get big and they become fixated on wanting to stay that way. This fixation develops in much the same way that women with anorexia nervosa can weigh 87 pounds and look in the mirror and still think that they look fat. In steroid users, the pattern is simply the reverse; they may weigh 230 pounds with a 32-inch waist and still think that they are not muscular enough.
- Endocrine effects: When a male takes steroids, which are all synthetic derivatives of testosterone, the body subsequently shuts down its own production of testosterone. If that person then stops using steroids, his own testosterone production may take weeks, months, or occasionally years to return to normal. He may develop impaired erectile function, loss of sex drive, and sometimes depression when coming off steroids, which often motivates him to go back to using the drugs.
- Hedonic effects: There's a pleasure experience that is associated with taking steroids. It's not the same as getting high as with drugs or alcohol, but there is a reinforcing quality of some type. Some steroid users describe this as feeling self-confident or invincible. Scientists are convinced of this phenomenon because male hamsters in laboratory studies will self-inject steroids to the point of death if they are given the opportunity.

When trying to get off of steroids, the user has to be given medications to try and stimulate his own testosterone function while slowly tapering the external steroids, Pope said. "It's an elaborate program that requires an expert endocrinologist and sometimes also a psychiatrist, so it's a difficult process."

Kaufman pointed out that while the findings of the study are alarming, they are preliminary. "The results merit an extended, more comprehensive sample in order to build a strong scientific foundation," he said. "We are applying for funding to conduct a larger study. These results need to be followed up because they have many implications. For example, clinical trials are already underway to establish whether replacement of scyllo-inositol can be useful for preventing cognitive decline in people with early Alzheimer's, an approach that may prove useful in steroid users." In any case, said Pope, "It is urgent to pursue this area of research, given the potential implications for the millions of men who have used or who are contemplating using steroids."[13]

Now answer the questions below.

1. Which of the options below is presented in the article as one of the most serious health concerns for male steroid users?

i. Increased likelihood of dementia
ii. Breathing difficulties
iii. Extreme weight loss
iv. Heart failure

2. The following hypothetical data sets contain amygdala volumes for two samples of male weightlifters. (**Note**: Samples were used since it is not practical to obtain the amygdala volume of all male weightlifters who are steroid users or non-users.)

1. Find the mean of each sample. Round each answer to two decimal places.

Sample A (in cubic centimeters)	**Sample B** (in cubic centimeters)
1.22, 1.17, 1.08, 1.26, 1.45, 1.42	1.16, 1.64, 1.41, 1.52, 1.29, 1.61

2. Assuming that one of the samples is from the population of male weightlifters who are steroid users and the other is from the population of male weightlifters who are non-users, which sample is most likely to be from the steroid users? Explain.

Now read the press release below about a treatment for depression:

STUDY 2:

Researchers See Promise in Potential Depression Treatment

Belmont, MA - Individuals with major depressive disorder or bipolar disorder who receive low-field magnetic stimulation (LFMS) show immediate and substantial mood improvement, report McLean Hospital researchers in the August 1 issue of Biological Psychiatry.

"LFMS is unlike any current treatment. It uses magnetic fields that are a fraction of the strength but at a higher frequency than the electromagnetic fields used in TMS and ECT," explained first author Michael Rohan, PhD, a physicist at McLean Hospital's Brain Imaging Center[14] and a lecturer at Harvard Medical School.

According to Rohan, although other brain stimulation treatments like electroconvulsive therapy (ECT) and transcranial magnetic stimulation (TMS) are often effective for the treatment of depression, they typically take longer to have an impact on mood. In addition, in the case of ECT, it is associated with side effects, such as memory loss.

Likewise, antidepressant medications can be highly effective in treating depression but can take between four to six weeks before mood changes can be detected.

"Importantly, LFMS appears to have an immediate effect on mood and thus has the potential to provide relief in emergency situations," explained Rohan, who first reported on the potential use of LFMS to treat depression in a ground-breaking study in 2004. "In addition to providing quick relief from symptoms, the other exciting piece about LFMS is that no side effects have been observed."

Using a portable tabletop LFMS device that Rohan designed, the researchers conducted the study with 63 volunteers between the ages of 18 and 65. All participants were diagnosed with either major depressive disorder or bipolar disorder and were on a stable regimen of antidepressants or mood-stabilizing medications for at least six weeks. Of the study participants, 34 actively received LFMS, while the other 29 went through the process of receiving LFMS, but did not actually receive any brain stimulation. Since neither the patients nor the researchers knew which treatment each person actually received, the true effect of the LFMS could be measured.

Each participant rated their mood before and after the single 20-minute treatment of LFMS using two common self-assessment tools—a visual analog scale and the positive and negative affect schedule. The participants who actively received the LFMS treatment indicated a marked improvement in their mood, while those who did not receive LFMS reported no change in mood.

"We observed immediate improvement in mood following relatively brief exposure to LFMS," noted Rohan. "Although larger research studies are needed, we think LFMS could be a powerful tool as a rapidly acting treatment for depression either alone or in combination with medication."

Biological Psychiatry also published commentary from experts on brain stimulation who said Rohan's work on LFMS is exciting and showed great promise.

"If the results described in this study are replicated in larger studies and the effects are shown to be durable, LFMS would be a welcome addition to the clinical armamentarium in the treatment of depression, may find application in other psychiatric and neurologic diseases, and may help to inform and guide us toward future directions in neuromodulation," wrote Mouhsin Shafi, MD, Philip Stern, MD, and Alvaro Pascual-Leone, MD, PhD.

According to Rohan, additional research is already underway to find the best parameters for LFMS use in the clinical treatment of depression. He has also stated a research study to evaluate the effects of multiple compared to single treatments, and how long the antidepressant effects last following treatment.[15]

Answer the questions below.

3. Recall that important factors for a decision or situation, or the factors analyzed in a study, are called *variables*. What variables did the researchers examine in this study?

4. Doctors who are familiar with the details and results of this study commented that if the study is "replicated" and the results are shown to be "durable," this treatment approach would be a welcome addition to existing treatment options for major depressive disorder and bipolar disorder. Explain what these comments mean in your own words.

STUDY 3:

The final reading in this preparation discusses a relationship between abuse as a child and homelessness. Read the article, "Abuse as a Child Linked to Longer Term Homeless," from the *Sydney Morning Herald*, available at the link below:

http://www.smh.com.au/federal-politics/political-news/abuse-as-a-child-linked-to-longer-term-homeles s-20120418-1x7hx.html

5. The study described in the article was conducted on two groups of people. Define these two groups.
6. The researchers examined several variables that contribute to long-term homelessness. What variables did they examine?
7. (a) Which of the following statements accurately describe the New York Times article on Universal Basic Income linked below? Check all that apply.

   Finland's Basic Income Trial Boosts Happiness, but Not Employment

   https://carnegiemathpathways.org/go/nyt_finland_basic_income

   i. According to the study, proponents of the Universal Basic Income (UBI) believe that UBI will significantly reduce homelessness and substance abuse.
   ii. After the initial success of this trial, Finland is planning to implement the UBI program on a national scale in 2021. This idea is supported by all major political parties in Finland.
   iii. The study doesn't provide any detailed information about the control group. However, it's safe to assume that it consists of people who did not receive UBI.
   iv. Proponents of the Universal Basic Income (UBI) hoped to show that UBI recipients would be more likely to work.

(b) Which of the following statements accurately describe the New York Times article on the benefits of the Mediterranean Diet linked below? Check all that apply.

That Huge Mediterranean Diet Study Was Flawed. But Was It Wrong?

https://carnegiemathpathways.org/go/nyt_diet_study

i. In the original study, the trial enrolled 7,447 participants aged 55 to 80 who were assigned one of three diets: a Mediterranean diet with at least four tablespoons a day of extra virgin olive oil; the same diet with an ounce of mixed nuts; or a traditional low-fat diet.
ii. The researchers claim that the Mediterranean diet can cut the risk of type 2 diabetes and sleep apnea by about 20 percent in those at high risk.
iii. The study has come under a lot of scrutiny because of its flawed design.
iv. After re-analysing the results of the initial study, there is now consensus in the scientific community about the conclusion of the study. The Mediterranean diet does indeed cut the risk of heart attacks and strokes by about 30 percent in those at high risk.

(c) Which of the following statements accurately describe the Science Daily article on the effect of social media and well-being linked below? Check all that apply.

Social media use increases depression and loneliness:

https://www.sciencedaily.com/releases/2018/11/181108164316.htm

i. The link between the use of social media and substance abuse has been talked about for years, but a causal connection had never been proven. For the first time, University of Pennsylvania research based on experimental data connects Facebook, Snapchat, and Instagram use to increased substance abuse.
ii. The link between the use of social media and depression has been talked about for years, but a causal connection had never been proven. For the first time, University of Pennsylvania research based on experimental data connects Facebook, Snapchat, and Instagram use to decreased well-being.
iii. Each of 143 participants completed a survey to determine mood and well-being at the study's start, plus shared shots of their iPhone battery screens to offer a week's worth of baseline social-media data. Participants were then randomly

assigned to a control group, which had users maintain their typical social-media behavior, or an experimental group that limited time on Facebook, Snapchat, and Instagram to 10 minutes per platform per day.

iv. Each of 143 participants completed a survey to determine mood and well-being at the study's start, plus shared shots of their iPhone battery screens to offer a week's worth of baseline social-media data. Participants were then randomly assigned to a control group, which had users not use any social media, or an experimental group that limited time on Facebook and Twitter to 30 minutes per platform per day.

After Preparation S.3 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration S.3, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
identify the purpose of a scientific study or statistical study.	
identify variables of interest in a scientific study or statistical study.	
identify the conclusion of a scientific study or statistical study.	
identify potential limitations in a scientific study or statistical study.	

[11] http://www.mcleanhospital.org/biography/harrison-pope

[12] http://www.mcleanhospital.org/research-programs/biological-psychiatry-laboratory

[13] http://www.mcleanhospital.org/news/2015/06/15/brain-imaging-study-suggests-long-term-steroid-use-can-lead-significant-brain

[14] http://www.mcleanhospital.org/research-programs/brain-imaging-center

[15] http://www.mcleanhospital.org/news/2014/07/28/researchers-see-promise-potential-depression-treatment

This page titled 3.3.1: Preparation S.3 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

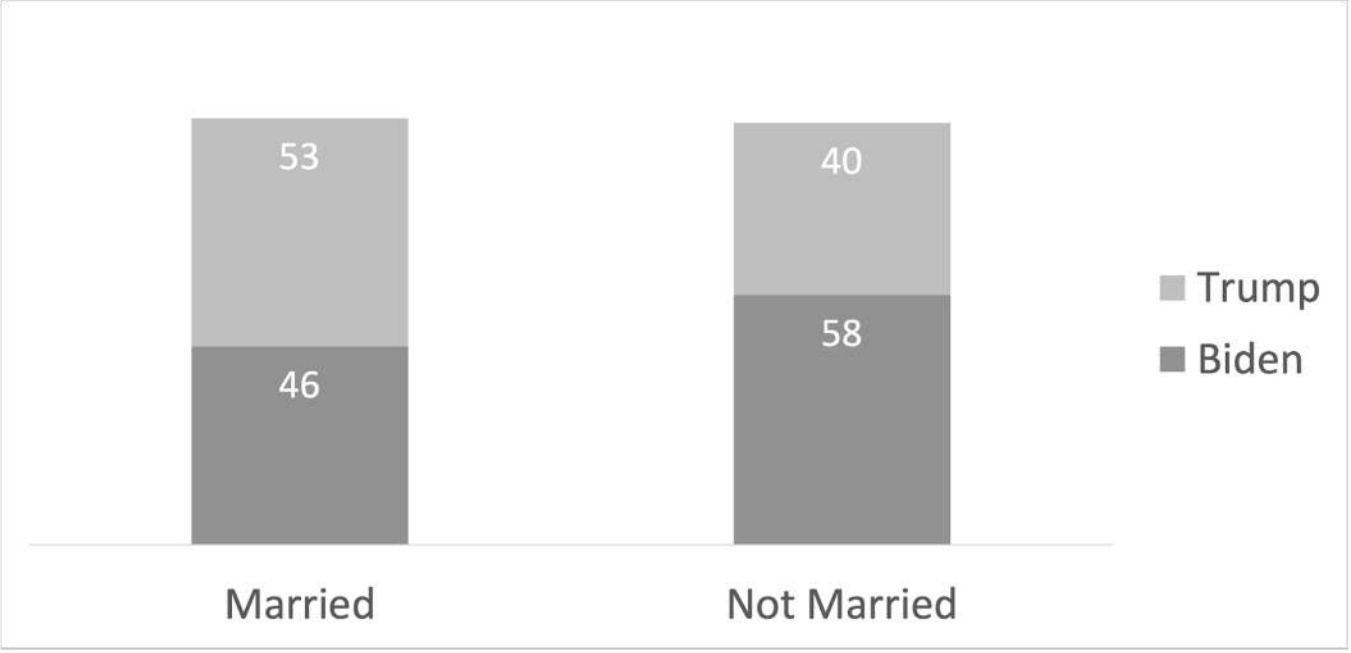
3.3.2: Exercise S.3

Now that you have some experience reading and interpreting articles with quantitative information, stretch your critical thinking muscles some more by reading part of an academic article based on quantitative research. Keep in mind that the intended audience for this article is not the general population, but, more likely, other social researchers and policy makers. Please read the introduction that begins on the first page and continues up to the "Methods" section on the second page. Also read the small section in blue on the first page. As you are reading, highlight any quantitative information the authors are using. Then answer the questions below. The PDF of the article can be found here: https://carnegiemathpathways.org/go/Caton2005.

1. Name three factors the authors identified as distinguishing homeless people from those who have never been homeless.
2. What is the sample of interest in this study?
3. Do the conclusions of this study (discussed in the section in blue) generalize to a larger population? If so, what is the largest population for whom we can generalize the conclusions? Explain.
4. The article states that this is a **longitudinal study**. In a longitudinal study, data is gathered from the same sample or population repeatedly over an extended period of time. What are some advantages to this type of study?
5. What did this study conclude were the strongest predictors of a longer duration of homelessness?
6. What kinds of biases could have occurred with this study?
7. What is a possible lurking variable that could be present in this study?
8. Does this study provide evidence that a causal relationship exists between the risk factors and homelessness? Explain.

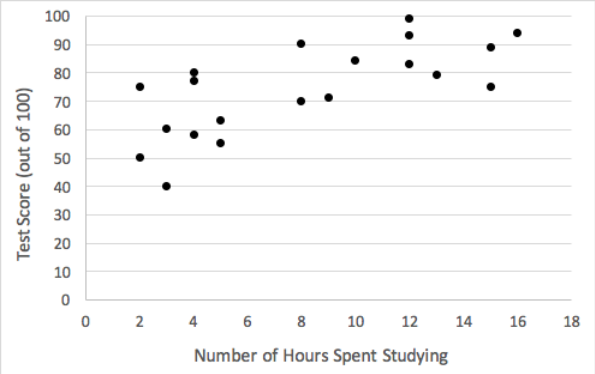
2020 United States Presidential Election: Exit Polls

The following stacked bar graph summarizes results from exit polls in the 2020 U.S. presidential election. The survey results come from questionnaires completed by approximately 15,590 U.S. voters.[17]



9. What is the sample and population of interest in this study?
10. What generalizations can be made based on these survey results? In your answer, clearly define the population to which generalizations can be made.
11. In this study, how is marital status related to preference for president?
12. What are some possible lurking variables in this study that could explain the relationship between marital status and preference for president?

13. What kinds of biases could have occurred with this study?
14. Suppose a researcher investigated the relationship between the number of hours students spent studying and their test scores. The researcher concluded that the data indicate that studying more hours causes students to have better test scores. The data are summarized in the scatterplot below.



(a) Do you agree with the researcher's conclusion? Explain. (Hint: What type of relationship exists between the two variables?)

(b) What other factors might explain this relationship?

---

[17] https://www.cnn.com/election/2020/exit-polls/president/national-results

This page titled 3.3.2: Exercise S.3 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.4: Sampling and Error

INTRODUCTION

Think about a population proportion of interest that you'd like to estimate. For example, a percentage of New Hampshire primary voters who support a particular presidential candidate or a percentage of all high school students who are concerned about climate change.

When trying to answer similar questions, it is impossible, or highly impractical, to survey every single individual in a population (that would be a census). Instead, we use representative samples to calculate sample estimates, and then generalize those results to the corresponding populations.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- statistics are numerical characteristics of samples, and parameters are numerical characteristics of populations.
- statistics vary from sample to sample and are used to estimate unknown population parameters.
- the margin of error of a sample statistic describes how close a certain percentage of all similarly constructed sample statistics are to the unknown population parameter.
- large samples provide better estimates of population parameters.

By the end of this collaboration, you should be able to

- distinguish between population parameters and sample statistics.
- create an interval estimate of a parameter based on a statistic and margin of error.
- interpret the margin of error from a statistical study.

PROBLEM SITUATION 1: POLLS IN THE 2016 PRESIDENTIAL ELECTION

The election of Donald Trump to the presidency of the United States was a shocking result to many Americans. The majority of polls and political commentators predicted that Hillary Clinton would win the election, and many predicted that Clinton's victory would be sweeping and convincing. The night before the 2016 presidential election, the *New York Times* reported, "Hillary Clinton has an 85% chance to win."

The *New York Times* and others' predictions were partially based on statewide polls from swing states such as Ohio, Pennsylvania, and Florida. Swing states in presidential elections are states in which support for the Republican candidate and the Democratic candidate is almost evenly split across voters.

Statistics and Parameters

A result of a poll or survey is an example of a **sample statistic**—a numerical characteristic of a sample. The primary role of a statistic is to estimate an unknown **population parameter**—a numerical characteristic of a population.

An example of a population parameter is a population proportion. For example, suppose that, before the 2016 presidential election, a poll in Ohio surveyed 1,000 people and concluded that 46% of people in Ohio will vote for Donald Trump and 47% will vote for Hillary Clinton. The two percentages are sample statistics that each estimate a population parameter: the proportion of all voters who will vote for Trump and Clinton, respectively.

Margin of Error and Confidence Level

Sample statistics are only estimates. Every sample statistic has an **error**—the difference between the sample statistic and the corresponding population parameter. The **margin of error** of a sample statistic specifies how close a certain percentage of similarly constructed sample statistics are to the population parameter. This percentage of similarly constructed sample statistics is called the **confidence level**.

For example, in the Ohio poll of 1,000 people, suppose that the margin of error for the sample statistic for the people who will vote for Trump (46%) is 3 percentage points. This margin of error was based on a 95% confidence level. This means that 95% of similarly obtained sample statistics will be within 3 percentage points of the actual population parameter (in this case, the population proportion of those who will vote for Trump). This also means, however, that 5% of similarly obtained sample statistics will not be within 3 percentage points of the population parameter.

Using the sample statistic and margin of error, we can create an **interval estimate** that gives us a plausible range of values for the unknown population parameter. Interval estimates have the form:

statistic ± margin of error or

(sample statistic – margin of error, sample statistic + margin of error)

The interval estimate for the example above would be 46% ± 3%. In practice, we typically don't know whether or not an interval estimate actually contains the population parameter, since the population parameter is still unknown. Actual population parameters are often unknown and impossible to determine without measuring every single member of the population. Read the following example from the same election, then answer the questions that follow.

The Muhlenberg/Morning Call poll predicted a Clinton victory. This poll was conducted in Pennsylvania a week before the election. Respondents were asked, "If the 2016 Presidential Election was being held today and the race was between Hillary Clinton, Donald Trump, Gary Johnson, and Jill Stein, who would you vote for?" The poll results for this question are shown below.

Hillary Clinton	44%
Donald Trump	**40%**
Gary Johnson	7%
Jill Stein	2%
Neither/Other	3%
Not Sure	5%

According to the "Methodological Statement" for the Muhlenberg/Morning Call poll, these results were based on "a telephone survey of 405 likely general election voters [in Pennsylvania]... between October 30 and November 4, 2016. Respondents were interviewed in English on both landlines... and cell phones." Muhlenberg/Morning Call reported that the results of the poll had a margin of error of 5.5 percentage points at the 95% confidence level.

1. Do you think the actual percentage of all likely general election voters in Pennsylvania who supported Donald Trump a week prior to the election was exactly 40%? Explain your answer.
2. What does the margin of error (5.5 percentage points) tell us about the actual percentage of all likely general election voters in Pennsylvania who supported Donald Trump a week prior to the election?

The poll result of 40% for Donald Trump is an estimate of an unknown population parameter—the proportion of all likely general election voters in Pennsylvania who supported Donald Trump in the 2016 election, at the time of the poll.

3. Use the estimate and margin of error from the Muhlenberg/Morning Call poll to find a range of values which the poll suggests might contain the proportion of the population that favors Donald Trump. Then find the range of values which the poll suggests might contain the proportion of the population that favors Hillary Clinton.

Pennsylvania ultimately voted for Donald Trump, providing Donald Trump with 20 electoral votes in his decisive electoral-map victory in the 2016 presidential election. Among all general election voters in Pennsylvania, 48.8% supported Donald Trump, and 47.6% supported Hillary Clinton.

4. Does the range of values from Question 3 capture the actual percentage of Pennsylvania general election voters who voted for Donald Trump and Hillary Clinton in the 2016 presidential election?

Interpreting Margin of Error and Confidence Level

A Pew Research Center survey, conducted June 7–July 5, 2016, questioned U.S. adults on their views about the 2016 presidential election. Among 4,602 respondents (4,172 by web and 430 by mail), the survey found that 59% of U.S. adults experienced election fatigue, i.e., they felt worn out by the amount of election coverage in the news. The survey reported a margin of error of 2.3 percentage points at a 95% confidence level.

5. Identify the *sample size* and sample statistic in this study.
6. What is the population parameter of interest in this study?

7. Use the sample statistic and margin of error to construct an interval estimate for the unknown population parameter.
8. Interpret the confidence level of 95% for this interval estimate. Can we conclude with certainty that the interval estimate contains the population parameter?

PROBLEM SITUATION 2: LEGALIZATION OF MARIJUANA

Sampling from a Finite Population

To better understand how sample statistics vary from a population parameter, we will generate random samples from a finite population with a known population parameter. We can then examine the distribution of sample statistics. From there, we can examine the errors in the sample statistics, and see how statistics can fall within and outside the margin of error.

The legalization of marijuana is an important issue across the nation. Ballot measures were introduced in several states to address this issue in the 2016 election. Suppose that in a small city (population: 100), the proportion of all adults who support marijuana legalization is 0.60. The table below displays the opinion of all 100 of these adults. Y indicates the person supports marijuana legalization, and N indicates the person does not support marijuana legalization.

Table: Support for Legalization of Marijuana

Population of the city N = 100, Population proportion = Number of adults who support legalization of Marijuana / Total number of adults = 60 / 100 = 0.60.

1	Y	26	N	51	N	76	N
2	Y	27	Y	52	Y	77	N
3	N	28	Y	53	N	78	Y
4	Y	29	N	54	Y	79	Y
5	Y	30	Y	55	Y	80	Y
6	N	31	N	56	Y	81	N
7	Y	32	Y	57	Y	82	Y
8	N	33	N	58	Y	83	Y
9	N	34	Y	59	Y	84	Y
10	N	35	Y	60	N	85	Y
11	Y	36	N	61	N	86	Y
12	Y	37	Y	62	N	87	Y
13	N	38	Y	63	Y	88	Y
14	N	39	N	64	Y	89	N
15	Y	40	N	65	Y	90	Y
16	N	41	Y	66	N	91	Y
17	Y	42	Y	67	N	92	Y
18	N	43	N	68	Y	93	N
19	N	44	N	69	Y	94	N
20	Y	4	Y	70	Y	95	Y
21	Y	46	Y	71	N	96	Y
22	N	47	N	72	Y	97	N
23	Y	48	Y	73	Y	98	Y

24	N	49	N	74	Y	99	N
25	N	50	Y	75	Y	100	Y

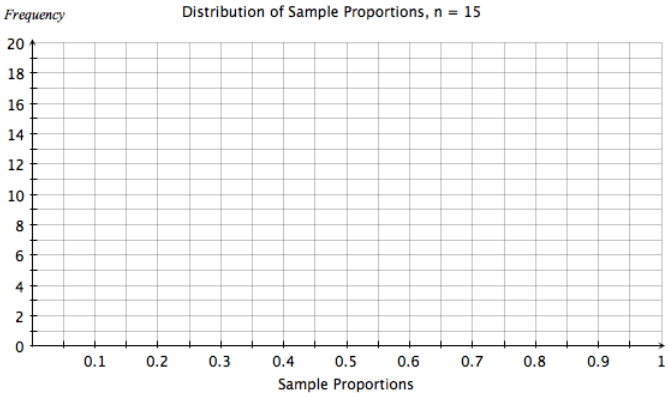
9. If you collected several random samples of size 15 from this population, and found the sample proportion for each sample (the proportion of the sample that supports marijuana legalization), how close do you think the overwhelming majority (say, 90%) of the sample proportions would be to the actual population proportion of 0.60?
10. Using a random number generator, generate a list of 15 random numbers between 1 and 100. Use the list of 15 random numbers to collect a random sample of 15 people from the population above, and label these people with Y or N. Write down the 15 numbers and their labels.

https://www.random.org/

11. Use the 15 randomly generated numbers from Question 10 to determine the proportion of your sample that supports marijuana legalization. This sample proportion is a sample statistic. In this case, the sample proportion is found by dividing the number of successes ("Y" designations) by the sample size (15). Round your answer to two decimal places.
12. What is the difference between the sample statistic that you found in Question 11 and the known population parameter, 0.60?
13. Now, imagine repeating the same sampling process to get the total of 30 different samples, each one with its own sample proportion listed in the following table. It's hard to see any patterns by just looking at this table, but it'll be much easier to do it if we represent it graphically using a dotplot.

Sample number	Sample proportion	Sample number	Sample proportion	Sample number	Sample proportion
1	0.33	11	0.6	21	0.73
2	0.47	12	0.6	22	0.73
3	0.47	13	0.6	23	0.73
4	0.47	14	0.6	24	0.73
5	0.53	15	0.6	25	0.73
6	0.53	16	0.67	26	0.8
7	0.53	17	0.67	27	0.8
8	0.53	18	0.67	28	0.8
9	0.53	19	0.73	29	0.87
10	0.53	20	0.73	30	0.93

On the graph below, create a dotplot of the sample proportions listed above. Note: If completing this problem online, follow the instructions given online to create your dotplot.



14. Estimate the center of this distribution of sample proportions. Do any of the sample proportions appear to be unusual in this distribution? Explain.

The margin of error of a sample proportion is based on the sample size and confidence level. The calculation of margin of error is beyond the scope of this class, but it is important to recognize how the margin of error changes as the confidence level changes.

Level of Confidence and Margin of Error

The table below shows the margin of error of sample proportions when the sample size is 15 (as in the questions above) for three confidence levels: 90%, 95%, and 99%. These margins of error are relatively large because the sample size (15) is so small.

The table shows how the margin of error increases as the confidence level increases. Using a wider range allows us to be more confident.

Confidence Level	Margin of Error (Percentage Points)
90%	21
95%	25
99%	33

Let's focus on the 90% confidence level and the corresponding margin of error of 21 percentage points. This means that 90% of all sample proportions will be within 21 percentage points of the population parameter.

15. Look at the distribution you created for Question 13. Are all of the sample proportions within 21 percentage points from the actual population proportion of 0.60? If not, specify the sample proportion(s) that are not within this margin of error from this population parameter.

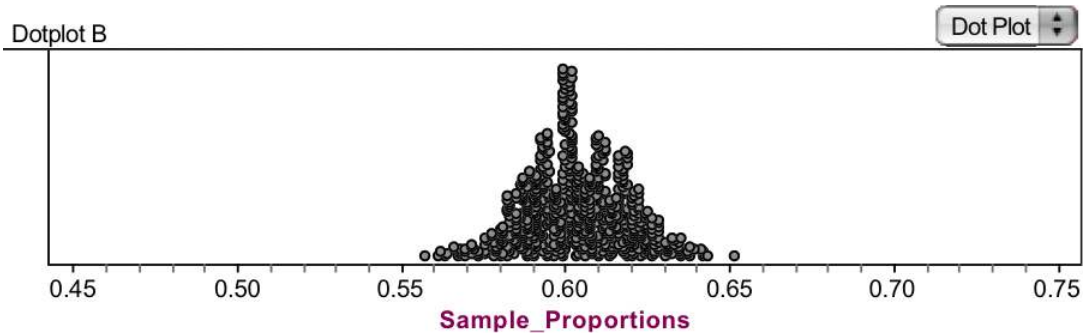
16. Using the same population of 100 people, suppose a researcher obtains one sample proportion of

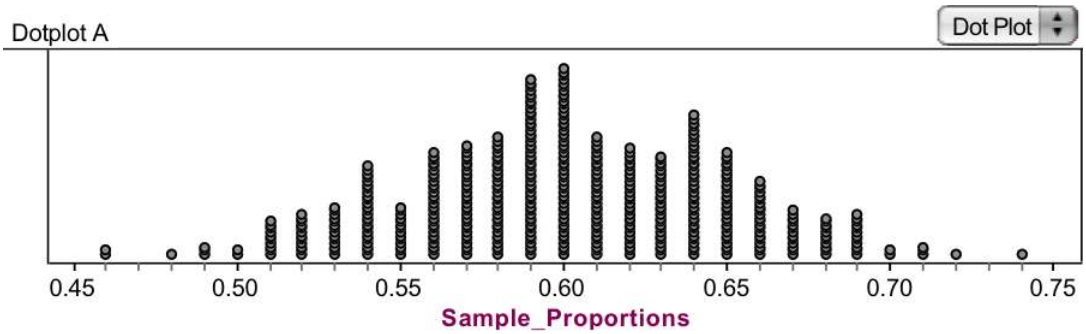
0.27 (people who support marijuana legalization) with a margin of error of 0.21 (21 percentage points). Would the researcher's confidence interval contain the population parameter? What are the implications of her stating that her confidence interval is 100% accurate when she makes conclusions about the population?

PROBLEM SITUATION 3: SAMPLE SIZE AND MARGIN OF ERROR

The sample size also impacts the margin of error of a sample statistic. To examine the role of sample size, imagine that we used computer simulations to create two distributions of sample proportions from the same population using random samples with different sample sizes.

- Dotplot A contains 500 sample proportions from random samples of size 100 taken from a population with a population proportion of 0.60.
- Dotplot B contains 500 sample proportions from random samples of size 1000 taken from a population with a population proportion of 0.60.





17. The sample proportions in the two dotplots above are estimates of the population proportion, 0.60. Which sample size, 100 or 1000, yields sample proportions that are better estimates of this population parameter? Explain.

18. Estimate the margin of error, based on a 95% confidence level, for the sample proportions in Dotplot A. Complete the following steps:

a. Estimate the interval that contains 95% of all sample proportions.
b. Calculate the width of the interval.
c. Divide the width of the interval by 2.

(19) Estimate the margin of error, based on a 95% confidence level, for the sample proportions in Dotplot B. Complete the following steps:

a. Estimate the interval that contains 95% of all sample proportions.
b. Calculate the width of the interval.
c. Divide the width of the interval by 2.

(20) What impact does sample size have on the margin of error?

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

Student Supplement - Random Number Table

21910	58787	05803	50347	46097	17957
88972	20045	94865	06784	39472	87118
12342	75597	61795	48784	79199	54950
90874	09624	84450	48548	97586	31222
39332	55983	97883	02238	90842	36771
74320	09949	82257	41008	32270	85903
89050	74066	76201	47571	17679	77816
52896	75248	35542	07290	19121	95225

49054	90435	86781	75609	70184	80534
50004	26955	67141	29751	72255	18787
50810	58229	34981	89175	08267	53219
34338	85610	15619	37936	40512	55858
54670	56390	23511	99380	66929	67740
29094	99782	26237	51792	14277	21545
14098	06961	82872	68191	02824	11308
65623	85865	58201	02901	83932	77593
14023	19760	51023	08776	16148	42084
71194	50528	30300	64057	64487	35328
62239	72499	01816	04699	30377	07289
78923	38920	68489	23086	62018	50310
03938	66370	72025	86706	78906	93835
44299	18862	28094	16628	15889	26346
28517	25172	13400	54135	38464	93086
39986	71728	32861	21363	74845	45605
01454	45263	48241	99242	70688	09326
98277	46713	84959	66456	63392	23271
70683	60259	05308	47039	09699	16387
15779	82819	30536	87665	46449	40757
11137	52296	39216	27076	10571	23703
76844	71737	26360	24480	24339	09144
40512	26132	20193	68251	13634	60177
07488	26182	95110	95156	73230	48062
98722	42051	73494	77558	12543	29759
40498	60258	62040	72315	71475	39871
30525	48274	30773	20589	21439	31808

This page titled 3.4: Sampling and Error is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.4.1: Preparation S.4

The previous unit introduced the four steps in the statistical analysis process. Steps 2–4 involve obtaining data from a sample, analyzing and summarizing the data, and drawing a conclusion. A key to these three steps is obtaining a sample that is *representative* of the population.

1. What does it mean for a sample to be representative of the population? It might be helpful for you to think of some specific populations, like students at our college or professional basketball players.

A **simple random sample** is defined as a sample formed in a manner such that each member of the population has an equal likelihood of being selected into the sample.

2. Can we always obtain a simple random sample? Consider the following example. A national political organization is interested in obtaining information about U.S. adults' views on medical marijuana. The organization plans to survey U.S. registered voters, using cell phones and landline phones. Is it possible for the organization to obtain a simple random sample of U.S. adults in this way? Explain.

A January 2017 Gallup poll reported that 47% of U.S. adults believe that President Obama will be remembered in history as an above average or outstanding president.[18] The survey was conducted via telephone interviews with 1,302 U.S. adults.

3. What are the sample and population of interest in the Gallup poll above?
4. Which statement below best describes the numerical result of 47% in the Gallup poll?

1. Approximately 47 out of every 100 adults in the U.S. believe President Obama will be remembered in history as an above average or outstanding president.
2. Approximately 47 out of every 100 adults in the U.S. in January 2017 believe President Obama

will be remembered in history as an above average or outstanding president.

3. Approximately 47 out of every 100 adults surveyed by Gallup in January 2017 believe President Obama will be remembered in history as an above average or outstanding president.

5. Let's assume that the sample of 1,302 U.S. adults is representative of all U.S. adults. Do you think it is reasonable to draw conclusions about the entire U.S. population, a population containing over 300 million people, from a representative sample of 1,302 U.S. adults? Explain.

Data Visualizations – Recall from Collaboration S.2 that a dotplot is a useful visual representation of quantitative data in which each value in the data set is displayed and groups of data values can be easily identified.

The dotplot below displays the unemployment rates (as percentages) for all 50 U.S. states and the District of Columbia in 2020.[19]



6. How many states or districts had an unemployment rate that was less than 7% in 2020?

The dotplot above displays a distribution of data—a collection of data values that satisfy a common variable. The center of a distribution is a single number that represents the typical data value.

7. Which value below best describes the center of the distribution of state unemployment rates in 2020?
   i. 10%
   ii. 5%
   iii. 6%
   iv. 7%

8. Which data value below appears to be most unusual in this distribution? A data value can be considered unusual if it is far away from the majority of values in a distribution.

i. 8.1%
ii. 5.5%
iii. 6.1%
iv. 13.5%

After Preparation S.4 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration S.4, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
identify the sample and population of interest in a statistical study.	
identify the center of a distribution from looking at a graph of the distribution.	
identify unusual values in a distribution from looking at a graph of the distribution.	

[18] https://news.gallup.com/poll/201770/americans-say-history-kind-unkind-obama.aspx

[19] https://www.bls.gov/lau/lastrk20.htm

This page titled 3.4.1: Preparation S.4 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.4.2: Exercise S.4

Prior to the 2016 election, Remington Research Group and Axiom Strategies conducted a poll in Wisconsin to find the voting preferences of likely voters. The survey was conducted on November 1 and November 2, 2016, among 2,720 likely general election voters. The survey found that 49% of Wisconsin voters supported Hillary Clinton and 41% of Wisconsin voters supported Donald Trump. The margin of error was 1.88 percentage points, based on a 95% level of confidence.

1. Identify the sample size and sample statistics in this study.
2. Identify the population parameters of interest in this study.
3. Use the information above to create an interval estimate for the proportion of all Wisconsin voters who supported Hillary Clinton.
4. Use the information above to create an interval estimate for the proportion of all Wisconsin voters who supported Donald Trump.

Let's investigate the accuracy of these interval estimates. Donald Trump ultimately won the majority of votes in Wisconsin. To identify the percentage of voters who actually voted for Donald Trump and Hillary Clinton, go to the presidential election results on the *New York Times* website: http://www.nytimes.com/elections/results/president.

5. Identify the percentage of votes which Clinton and Trump received in Wisconsin. Did the interval estimates from Questions 3 and 4 capture the actual percentages of voters who voted for these candidates in Wisconsin?

Possible Sources of Bias in the 2016 Presidential Election

Watch the video, "Why 2016 Election Polls Missed their Mark," from the Pew Research Center, available at the following link:

http://www.pewresearch.org/2016/11/15/video-why-2016-election-polls-missed-their-mark/.

6. The video describes three main sources of sampling bias which could have caused the pre-election polls to be inaccurate. Explain, in your own words, each possible source of sampling bias.

The Impact of Fake News

Read the article, "Many Americans Believe Fake News Is Sowing Confusion," from the Pew Research Center, available at the following link:

http://www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/.

7. Read the first five paragraphs of the article. What are the sample and population of interest in this study?
8. What are three sample statistics from this study?
9. What is the one population parameter that is of most interest to you in this study? What would you like to know about the population?
10. The study estimates that 64% of U.S. adults believe fake news has caused a great deal of confusion about basic facts of current events. The poll reported a margin of error of 3.6 percentage points at the 95% confidence level. Use this information to create an interval estimate for the percentage of all U.S. adults who believe fake news has caused a great deal of confusion about basic facts of current events.
11. Which statement(s) below provide(s) an accurate interpretation of the previous 95% confidence interval? Select all that apply.

    (i) If different, but similarly obtained, sample proportions were used to estimate the unknown population parameter, 95% of these sample proportions would be within 3.6 percentage points of the unknown population parameter.

    (ii) There is a 95% probability that 64% of U.S. adults believe that fake news has caused a great deal of confusion about basic facts related to current events.

    (iii) We can conclude with certainty that the proportion of U.S. adults who believe that fake news has caused a great deal of confusion about basic facts related to current events is within 3.6 percentage points of the sample statistic of 64%.

    (iv) 95% of interval estimates created from similarly obtained sample proportions would capture the actual proportion of all U.S. adults who believe fake news has caused a great deal of confusion about basic facts related to current events.

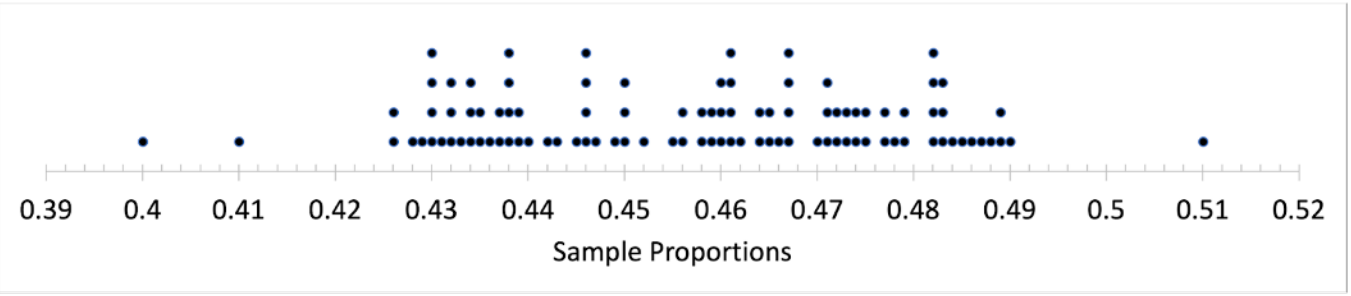
12. The study examined whether different subgroups of U.S. adults held different views on the impact of fake news. For example, the study considered whether there is a relationship between income level and views on the impact of fake news. Do the results of the study show that a correlation exists between these two variables?

Concerns about Global Warming among U.S. Adults

13. A Gallup poll[20], conducted between March 1 and March 23, 2023, surveyed U.S. adults about their concerns and beliefs regarding global warming. Among 1,009 U.S. adults, aged 18 and older, from all 50 U.S. states and the District of Columbia, the poll found that 46% believe that global warming poses a serious threat in our lifetime. The survey reported a margin of error of 4 percentage points at the 95% confidence level.

    1. Use the information about the Gallup poll to create an interval estimate for the proportion of all U.S. adults who worry a great deal about global warming.
    2. If Gallup had wanted a smaller margin of error for this population parameter, what are two ways they could have achieved it?

(14) Let's assume that 46% is the population proportion from the Gallup poll in Question 13—the actual proportion of all U.S. adults in March 2023 who believe that global warming poses a serious threat in our lifetime. A computer simulation was used to generate random samples of size 1000 from this population. The dotplot below displays 100 of these sample proportions. The margin of error is approximately 4 percentage points at the 95% confidence level.



    1. How many sample proportions are outside the margin of error from 0.46?
    2. Which sample proportions will produce interval estimates that do not capture the population parameter?

---

[20] https://news.gallup.com/poll/474542/steady-six-say-global-warming-effects-begun.aspx

This page titled 3.4.2: Exercise S.4 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.5: Exploring Climate Data

INTRODUCTION

To understand any distribution of quantitative data, we examine both its center (measured by mean, median, and mode) and variability (measured by range and standard deviation). This is extremely important in the context of manufacturing, when the goal is to produce items that not only meet the specifications, but also have as little variability as possible. For example, consider two production lines that manufacture automotive components that are supposed to weigh 25 pounds. The following table shows two samples of size 8 taken from both production lines.

	Production Line A	Production Line B
	25	25
	25.02	25.18
	25.1	25.21
	25.12	25.18
	25.17	25.25
	25.05	24.98
	24.98	24.99
	25.25	24.95
Sample Mean	**25.09**	**25.09**
Sample Standard Deviation	**0.0923**	**0.1230**

Both samples have the same sample means of 25.09. However the first production line has a lower sample standard deviation, which means that the items manufactured on the first production line have less variability.

This concept is explored in greater detail in more advanced statistics courses. At this point, we just need to remember that knowing the average or the median of data doesn't usually give us the full picture.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- distributions of data can be described based on their centers and variability.
- distributions of data can be used to explore real-world situations.

By the end of this collaboration, you should be able to

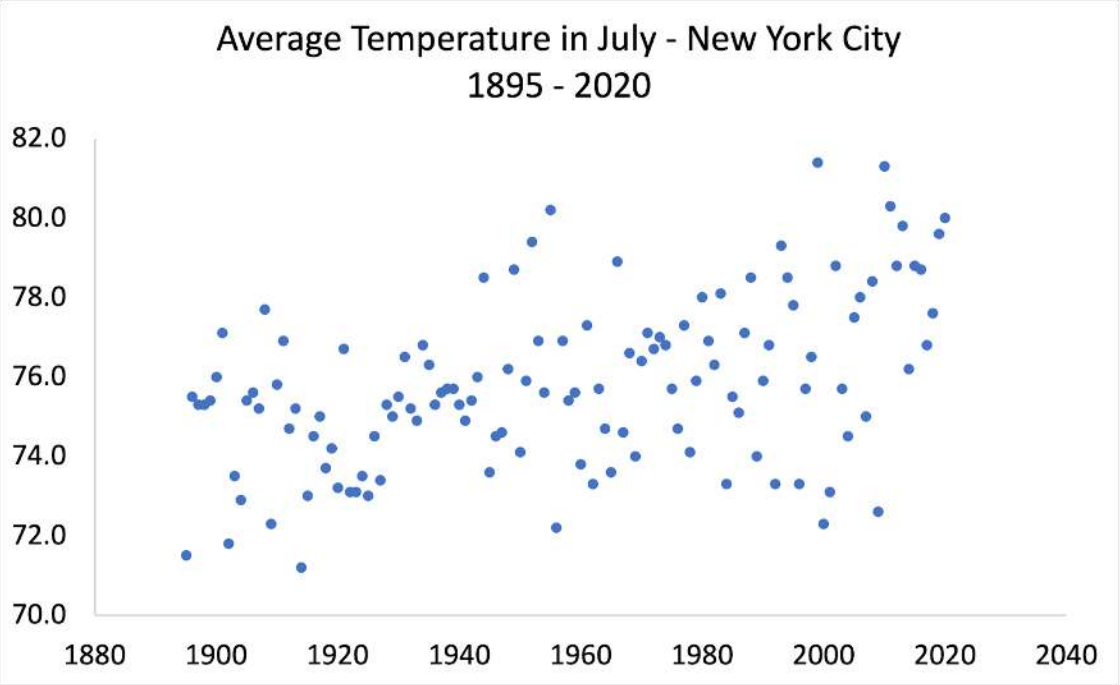
- describe the center and variability of a single-variable data distribution.
- estimate the mean of grouped data by computing a weighted mean.
- estimate the median and mode of grouped data given its frequency histogram.
- compute and interpret the sample standard deviation of a data distribution.

PROBLEM SITUATION 1: AVERAGE TEMPERATURES IN JULY IN NEW YORK CITY

Global warming is a major concern for people throughout the world. Rising temperatures, rising sea levels, drought, and extreme weather conditions threaten countries, economies, and communities in serious ways. To understand how the global climate and

local climates are changing over time, climate scientists collect and analyze a wide range of climate data to describe and predict climate change.

The scatterplot below displays the average temperatures (in degrees Fahrenheit) in July in New York City from 1895 to 2020. Each dot represents the average temperature in July for a specific year.



1. Describe any trends that you see in the graph? Does the graph support the view among scientists that temperatures across the globe are rising? Explain.
2. How could you use the data above to make an argument that average temperatures in July in New York City have increased over time? Your argument should involve quantitative information. Identify at least two different ways. Explain each approach.

Comparing Distributions

One way of examining how the average temperature in July in New York City has changed over time is to compare the average temperatures in different time periods. The tables below display the average temperature (in degrees Fahrenheit) in July in New York City in four 10-year time periods: 1901–1910, 1921–1930, 1971–1980, and 2011–2020.[22]

1901–1910

1901	1902	1903	1904	1905	1906	1907	1908	1909	1910
77.1	71.8	73.5	72.9	75.4	75.6	75.2	77.7	72.3	75.8

1921–1930

1921	1922	1923	1924	1925	1926	1927	1928	1929	1930
77.3	73.7	73.8	74.1	73.6	75.1	74	75.9	75.9	76.4

1971–1980

1971	1972	1973	1974	1975	1976	1977	1978	1979	1980
77.8	77.2	77.4	77.2	75.8	74.8	79	74.4	76.9	79.3

2011–2020

2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
80.3	78.8	79.8	76.2	78.8	78.7	76.8	77.6	79.6	80.0

3. In your group, calculate the sample standard deviations for the decade **1901-1910**. Round your answer to two decimal places.

As explained at the end of the corresponding preparation, calculating sample standard deviations is somewhat tedious, but fortunately, it only takes a few clicks to do it using Excel or Google Sheets.

- Step 1: Enter data into an Excel document or a Google Sheet
- Step 2: Click an empty cell and type in **=STDEV.S(**
- Step 3: Highlight the data you entered in Step 1 by holding the left mouse button and dragging over the data. This will populate the space after the parenthesis in the STDEV.S function.
- Step 4: Press Enter

Sample standard deviation for the decade 1901-1910 =

4. Interpret the sample standard deviation that you computed in Question 3. Carefully explain what it means.
5. Calculate the sample means and sample standard deviations for the other three decades. Complete the table below. Enter answers rounded to two decimal places. Compare the sample means and sample standard deviations. What do you notice?

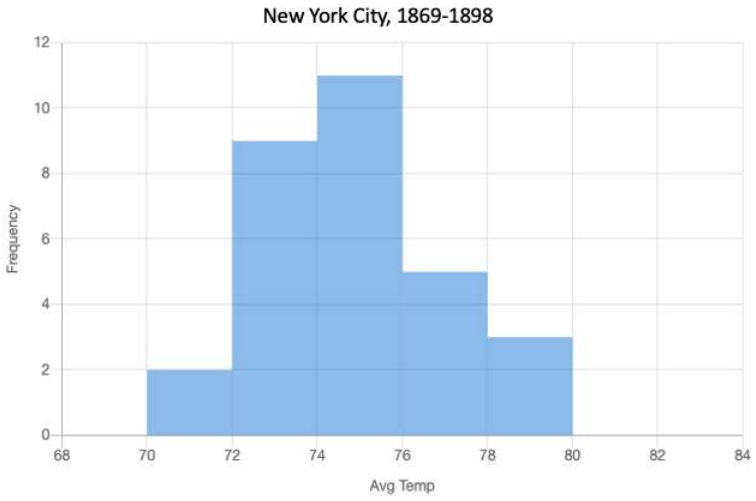
Decade	Mean	Standard Deviation
1901-1910	74.73	From Q3
1921–1930		
1971–1980		
2011-2020		

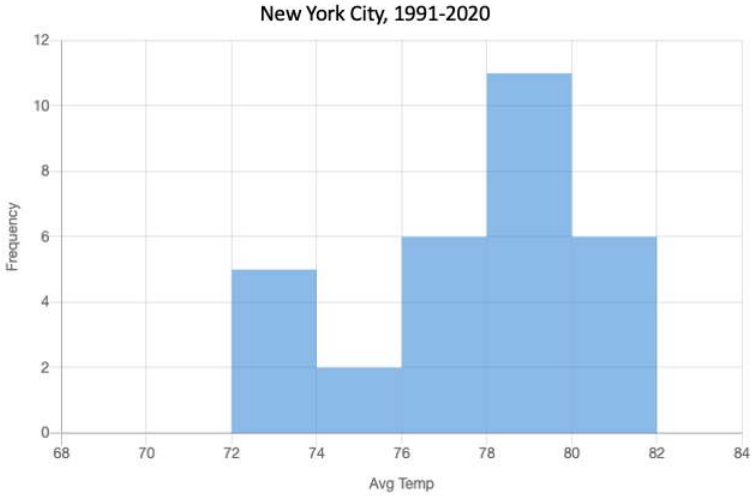
(6) In which decade: 1901-1910, 1921–1930, 1971–1980, or 2011-2020, do the average temperatures in July vary the most? How is this evident in the scatterplot of the average temperatures in July from 1895 to 2020?

Frequency Histograms

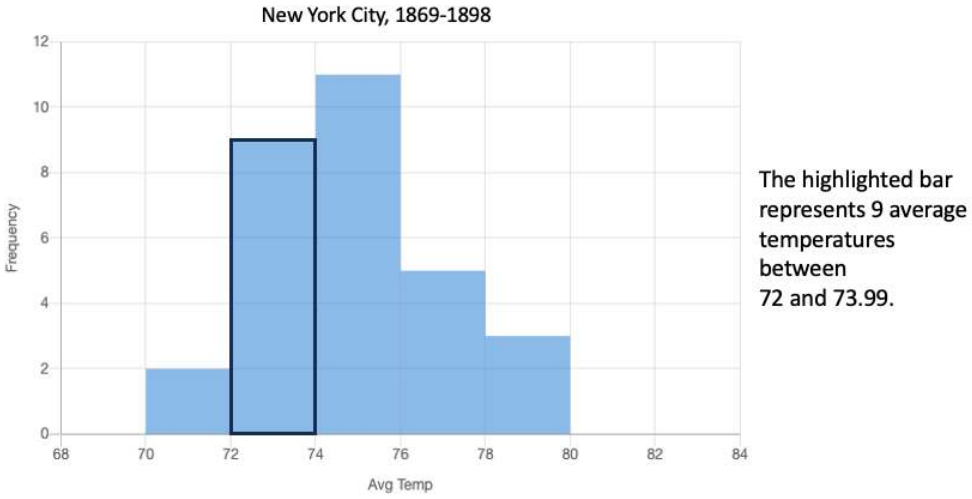
We can also compare distributions of average temperatures over 30-year time periods. The **frequency histograms** below show the average temperatures (in °F) in July in New York City from 1869 to 1898, and from 1991-2020. These time periods are the first 30 years and the last 30 years of the time frame in which data was collected. Each histogram groups the average temperatures into categories

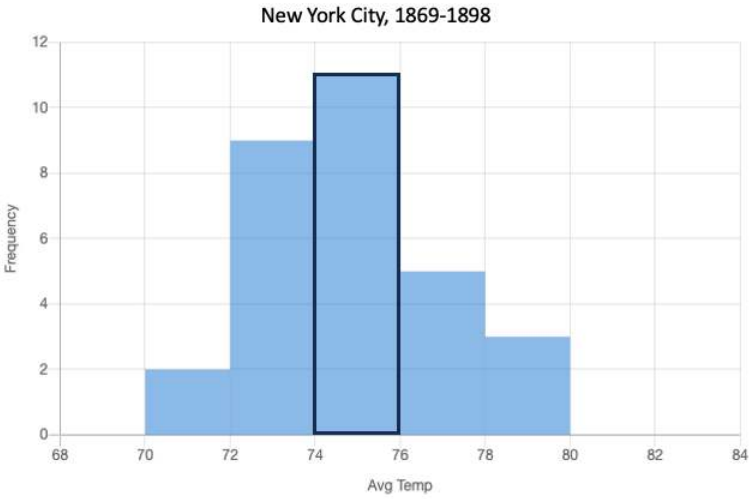
(non-overlapping groups of numbers) and shows the number of average temperatures in each category. The number of average temperatures in each category is the **frequency** of the category. The frequency of each category is how often it appears in the data, which is then the height of the corresponding bar.



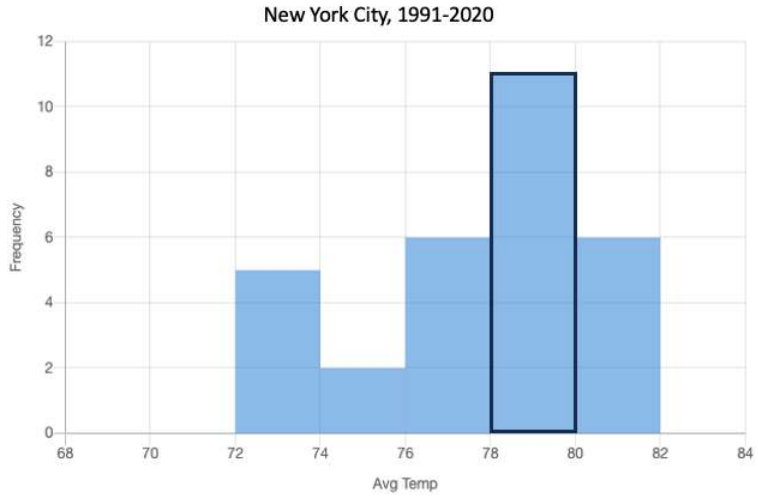


To better understand how to interpret histograms, let's examine the following…





The highlighted bar represents 11 average temperatures between 74 and 75.99.



The highlighted bar represents 11 average temperatures between 78 and 79.99.

7. (a) Estimate the mean, median, and mode of the average temperatures in the 1869 to 1898 histogram. Round to the nearest degree, if necessary.

(b) Estimate the mean, median, and mode of the average temperatures in the 1991 to 2020 histogram. Round to the nearest degree, if necessary.

8. How do these histograms support the consensus among climate scientists that climate change is leading to increased temperatures across the globe?

MAKING CONNECTIONS

Record the important mathematical discussions from the discussion.

---

[22] https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/city/time-series

This page titled 3.5: Exploring Climate Data is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.5.1: Preparation S.5

The next collaboration focuses on understanding and interpreting data distributions. In simple terms, think about a distribution as a shape of the data. In this course, we will be focusing on symmetric distributions, similar to the houseflies example shown in the following video:

https://www.khanacademy.org/math/ap-statistics/quantitative-data-ap/describing-comparing-distributions/v/classifying-distributions

To understand any distribution of quantitative data, we examine its **center and variability**.

Center: Recall that the center of a distribution is a single number that represents the typical data value in the distribution. Three numbers are used to represent the center: mean, median, and mode. The most common of these are the mean and median.

- The **Mean** is the arithmetic average, the sum of data values divided by the number of data values.
- The **Median** is the middle of a sorted list of data values. If the sorted list of data values has an odd number of values, the median is the middle value. If the sorted list of data values has an even number of values, the median is the mean (average) of the middle two values.
- The **Mode** is the number that occurs most frequently, though there does not have to be a mode if such a number does not exist.

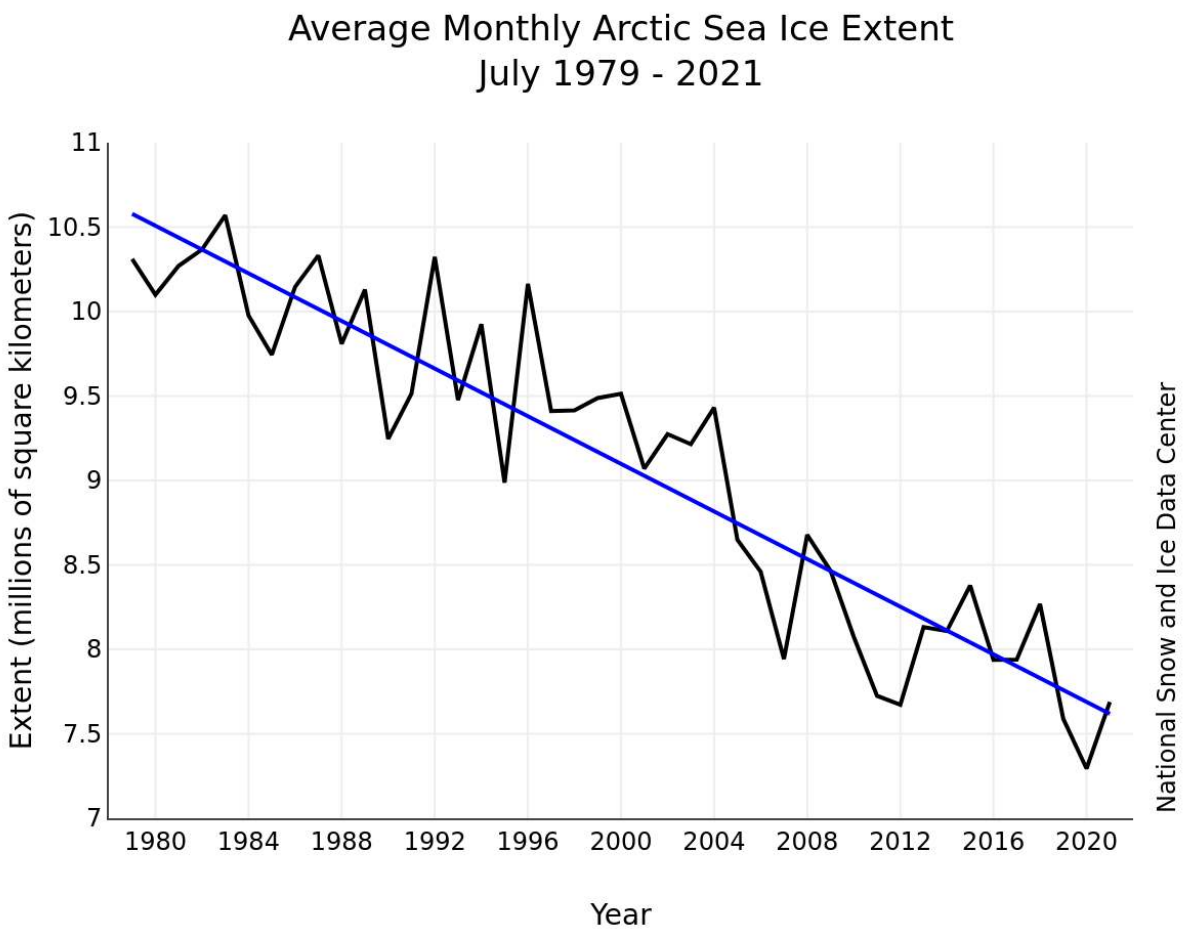
Two most commonly used measures of **variability** are range and standard deviation. Both will be discussed in more detail in this assignment.

- **Range**: The range of a data set is the difference between the maximum value and the minimum value.
- **Standard Deviation**: The standard deviation of a data set represents the typical distance of a

value from the mean.

Arctic Sea Ice Extent

Climatologists (scientists who study the Earth's climate) track a variety of climate indicators to assess the impact of climate change. One such indicator is the portion of the Arctic Sea that is covered by ice at different points throughout the year. This is called the "Arctic Sea ice extent" and is measured in millions of square kilometers (million km^2). The ice extent in the Arctic Sea is at its lowest in the late summer or early fall of each year. The graph below shows the Arctic Sea ice extent in the month of July each year from 1979 to 2021.[21]



1. In July 1979, the Arctic Sea ice extent was 10.5 million km^2. In July 2021, the Arctic Sea ice extent was 7.5 million km^2. Use this information to develop a linear model for the Arctic Sea ice extent over time. Let A represent Arctic Sea ice extent and t represent the number of years since 1979. Round the slope to three decimal places.
2. Use the linear model from Question 1 to predict the Arctic Sea ice extent in July 2025. Round your answer to two decimal places.
3. The following distributions contain the Arctic Sea ice extent in July for two 5-year periods: from 1979–1983, and from 2017–2021. The ice extents are measured in millions of km^2. Identify the mean, median, and mode for each distribution. Round to two decimal places when necessary.

Year	Ice Extent (million sq. km	Year	Ice Extent (million sq. km)
1979	10.5	2017	7.9
1980	10.4	2018	7.7
1981	10.7	2019	7.6
1982	10.8	2020	7.3
1983	11	2021	7.5
Mean		Mean	
Median		Median	

Mode		Mode	

4. What can you conclude from comparing the means of the Arctic Sea ice extent from these two periods of time?

Variability

It is often useful to examine the **variability** of a data distribution. Variability refers to how spread out the data values are. Variability is often measured in two main ways: range and standard deviation.

Range: The range of a data set is the difference between the maximum value and the minimum value.

5. Compute the range of the ice extents from 1979 to 1983, and compute the range of the ice extents from 2017 to 2021.

1979–1983: Range =

2017–2021: Range =

Standard Deviation: The standard deviation of a data set represents the typical distance of a value from the mean.

To find the standard deviation of a sample data set, we can begin by finding the difference, or deviation, of each data value from the sample mean.

6. The table below contains the Arctic Sea ice extents in July from 2017 to 2021, measured in millions of km^2. Complete the table by finding the deviation of each value in the 2017–2021 distribution. Use the mean rounded to two decimal places. (**Hint**: you calculated the mean in Question 3. Write the deviations in the third column.)

Year	**Value**	**Value − Mean (Deviation)**
2017	7.9	
2018	7.7	
2019	7.6	
2020	7.3	
2021	7.5	

7. Calculate the average of the deviations you found in Question 6 above. Does this average seem like a good estimate for the *standard deviation*, which is the *typical distance* from the mean? Why or why not?

8. Now square each of the deviations. Write the deviations in the third column and enter the squared deviations in the fourth column. Round the squared deviations to three decimal places.

Year	**Value**	**Value − Mean**	**$(\text{Value} - \text{Mean})^2$**
2017	7.9		
2018	7.7		
2019	7.6		
2020	7.3		
2021	7.5		

9. Why does squaring the deviations help correct the issue you found when you examined the average of the deviations in Question 7?

10. Now find the average of the squared deviations you found in Question 8. Round to three decimal places. Does this seem like a reasonable estimate of the standard deviation (typical distance from the mean)? Why or why not?

11. Take the square root of your answer to Question 10 (the square root of the average of the squared deviations). Round to three decimal places. Does this seem to be a better estimate of the standard deviation (typical distance from the mean)? Why or why not?

Your answer to Question 11 is very close to the actual sample standard deviation. There is only one small difference in the formal calculation to find the standard deviation of a sample. After adding up the squared deviations, instead of dividing by the number of squared deviations, n, to get the average, you divide by n − 1 (one less than the sample size). The formula is shown below. Note: The ∑ symbol represents sum, x represents a data value, and x̄ represents the sample mean.

$$Sample\ Standard\ Deviation = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$

We divide by $n - 1$ instead of n to increase the size of the sample standard deviation. We do this to ensure that the sample standard deviation does not underestimate the **population standard deviation**.

12. Calculate the *sample standard deviation* of the data set. Round the answer to three decimal places.

13. The table below contains the Arctic Sea ice extents in July over five consecutive years from 1979 to 1983. The ice extents are measured in millions of km^2. Complete the table below. Round the squared deviations (values in the rightmost column) to four decimal places.

Year	**Value for Arctic Sea Ice Extent**	**Value − Mean**	**(Value − Mean)^2**
1979	10.5		
1980	10.4		
1981	10.7		
1982	10.8		
1983	11.0		

14. Compute the sample standard deviation for the sample. Round to three decimal places.

Technology to the Rescue!

As you can tell, calculating sample standard deviation can be somewhat tedious. Fortunately, there's an easy way to do it using Excel or Google Sheets. It only takes a few clicks.

1. Enter data into Excel or Google Sheets
2. Click an empty cell and type in **=STDEV.S(**
3. Highlight the data you entered in Step 1 by holding the left mouse button and dragging across the data
4. Press Enter

15. Which statement below accurately compares the variability of the July Arctic Sea ice extents from 1979 to 1983 to the variability of the July Arctic Sea ice extents from 2017 to 2021?

i. The ice extents from 1979 to 1983 vary from their mean more than the ice extents from 2017 to 2021.
ii. The ice extents from 1979 to 1983 vary from their mean less than the ice extents from 2017 to 2021.
iii. The ice extents from 1979 to 1983 and the ice extents from 2017 to 2021 have the same variability.
iv. The ice extents from 1979 to 1983 vary from their mean half of the amount compared to the ice extents from 2017 to 2021.

After Preparation S.5 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration S.5, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
determine the mean, median, and mode of a data set.	

determine the range of a data set.
compute and interpret the sample standard deviation of a data set.

[21] http://nsidc.org/arcticseaicenews/2021/08/a-change-of-pace-2/

This page titled 3.5.1: Preparation S.5 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.5.2: Exercise S.5

The table below displays the average temperature (in degrees Fahrenheit) in July in New York City each year from 1991–2000.[24]

1991–2000

1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
77.7	74.2	80.2	79.4	79.2	73.4	75.8	76.5	81.4	72.3

1. Compute the sample mean and sample standard deviation. Round your answers to two decimal places. Feel free to use Excel or Google Sheets to calculate sample standard deviation.

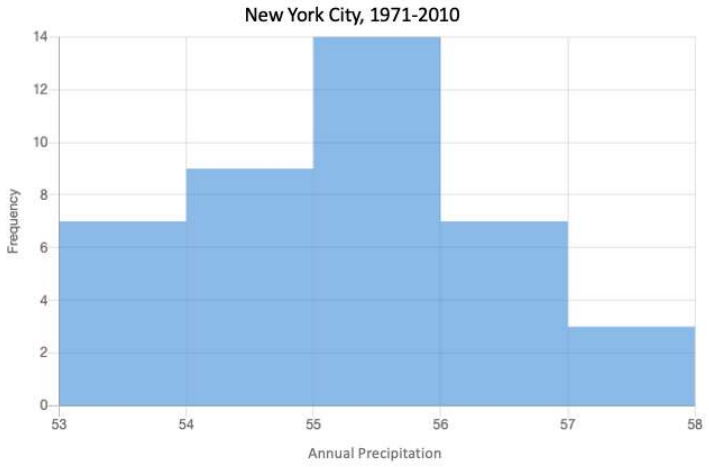
$$Sample\ Standard\ Deviation = \sqrt{\frac{\sum(x-\overline{x})^2}{n-1}}$$

The sample Mean =

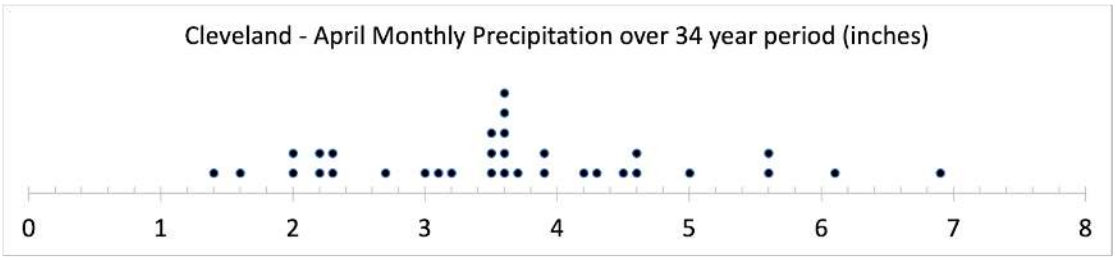
The sample standard deviation =

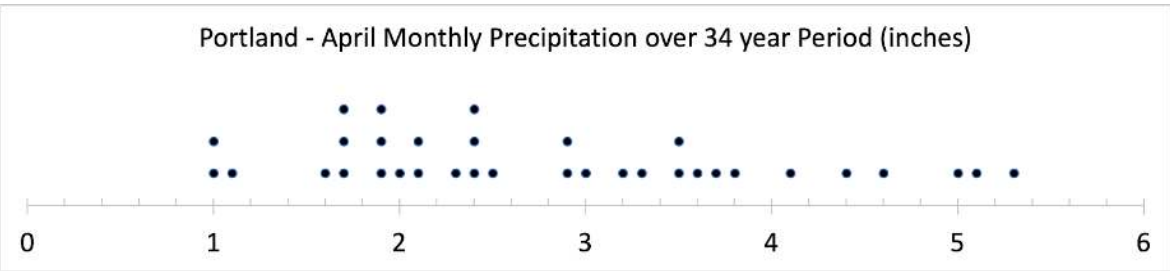
2. Interpret the standard deviation.
3. Which average temperature is the most unusual in this distribution? Explain.
4. Estimate the mean, median, and mode of the annual precipitation data in the histogram below.

   The frequency histogram below shows the annual precipitation (in inches) in New York City from 1971–2010.[25] Precipitation is rain, sleet, and snow.



5. The dotplots below display the monthly precipitation in April over a 34-year period for Cleveland, Ohio and Portland, Oregon.[26] Use the dotplots to answer the following questions.





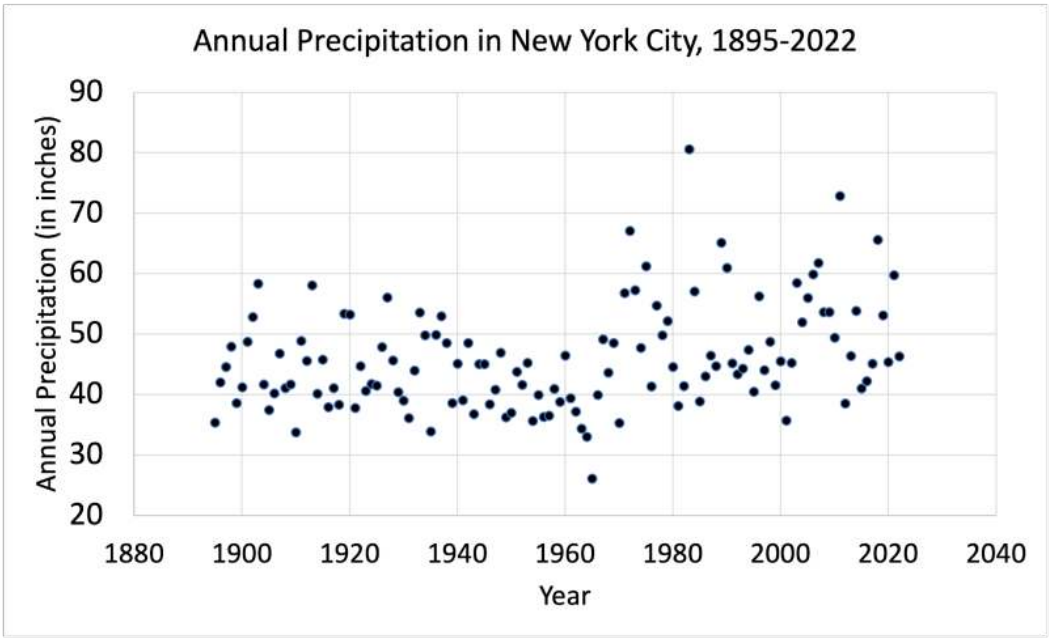
(a) Which city had the greater variability in April monthly precipitation over the 34-year period? Explain.

(b) Estimate the range of April monthly precipitation in Cleveland and Portland.

(c) Mean April monthly precipitation for Cleveland and Portland is 3.6 inches and 2.8 inches, respectively. Would it be reasonable to estimate that the standard deviations of Cleveland's April monthly precipitation is 3 inches? Would it be reasonable to estimate that the standard deviations of Portland's April monthly precipitation is 1.5 inches? Explain.

The scatterplot below displays the annual precipitation (in inches) in New York City from 1895–2022.[27]

Each dot represents the annual precipitation for a specific year.



6. What trend do you see in the graph? Do any characteristics of the way the chart was constructed affect your perception of this trend?
7. In 1895, the annual precipitation in New York City was 35.4 inches. In 2022, the annual precipitation in New York City was 46.3 inches. Use this information to develop a linear model for the annual precipitation in New York City over time. Let A represent annual precipitation and t represent the number of years since 1895. Round the slope to three decimal places.
8. Use the linear model from Question 7 to predict the annual precipitation in New York City in 2040. Round your answer to two decimal places.
9. Use the linear model from Question 7 to predict the year in which annual precipitation in New York City will reach 60 inches.

---

[24] https://www.weather.gov/okx/CentralParkHistorical

[25] https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/city/time-series

[26] https://www.ncdc.noaa.gov/cag/

[27] https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/city/time-series

This page titled 3.5.2: Exercise S.5 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.6: Exploring Climate Variability

INTRODUCTION

Consider two geographical regions: one with very little variation in July temperatures from year-to-year and another region with significantly higher variation in July temperatures. Assume that both regions have the same mean temperature of 80.

Next, suppose that this July, both regions recorded the same average temperature of 85 degrees. Subtracting the mean, and saying that both regions recorded temperatures 5 degrees higher than the mean, doesn't really give us enough information to make a meaningful comparison.

We need to somehow account for the variation between the regions. We will discuss this idea in more detail in this collaboration.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- data distributions can be characterized based on their shape.
- a Z-score describes the number of standard deviations that a data value is from the mean.
- Z-scores can be used to identify unusual values in bell-shaped distributions.
- Z-scores enable us to compare values that originate from different distributions.

By the end of this collaboration, you should be able to

- recognize the shape of a distribution of data values.
- compute and interpret Z-scores of values given the mean and standard deviation of distributions.
- interpret frequency histograms and relative frequency histograms to understand variability in a data distribution.

PROBLEM SITUATION 1: TEMPERATURES ARE RISING

The science brief you read in Preparation S.6, *The New Climate Dice: Public Perception of Climate Change*, written by scientists at the National Aeronautics and Space Administration (NASA), discusses the results of a study that compared distributions of temperature anomalies over time. This collaboration will explore the concepts presented in the brief and examine how these concepts relate to data distributions.

A *monthly temperature normal* is a 30-year average of monthly temperature data collected during a specific month at a specific geographic location.

The table below presents the monthly temperature normals (in degrees Fahrenheit, °F) in July for six climate divisions in the U.S.[28] The climate locations below represent a region in a state that has approximately uniform and consistent climate conditions. The means and standard deviations are derived from temperature data from 1991 to 2020. The means are calculated by finding the average of each July average temperature over 30 consecutive years.

Climate Location	**July Monthly Normal (°F) 1991 - 2020**	
	Mean	**Standard Deviation**
Central Massachusetts	72.6	2.4
Central Ohio	75.4	2.6
Northeast Arizona	73.8	1.7
South Central Florida	83.3	1.3
Southeast Desert Basins California	85.6	2.1
Southwest North Dakota	68.7	3.3

1. What are some reasons that climate scientists compute 30-year temperature normals? How can these statistics be used?

2. The Central Massachusetts climate division represents an area of approximately 4,200 square miles. Temperature data are collected at multiple climate stations in this region on a daily basis. Describe a specific method that you think could be used to combine daily temperature data from multiple climate stations to develop a monthly average temperature for the region. Note that the climate stations in the climate division are not uniformly distributed throughout this climate division.
3. July average temperatures in Central Massachusetts, from 1991 to 2020, have a mean of 72.6°F and a standard deviation of 2.4°F. Both of these statistics, the mean and standard deviation, are based on thirty July average temperatures from 1991 to 2020. Interpret the standard deviation. What does this tell us?
4. Compare South Central Florida's July monthly normal to Southwest North Dakota's July monthly normal. Which climate division has July temperatures that vary more from the 30-year average? Why?

A **temperature anomaly** is a measure of how much the temperature of a location varies from the location's typical or normal temperature, as determined by a reference value. Temperature anomalies can be computed in many ways. One approach is to compute a Z-score. The Z-score of a temperature is the number of standard deviations the temperature is from the typical or normal temperature. Note that to compute a Z-score, the population parameters (mean and standard deviation) must be known. When considering weather data, a recent 30-year average is acceptable for representing these parameters.

5. The tables below contain July average temperatures (°F) from 2017 to 2020 for three climate divisions in the U.S.[29] The tables also show the mean (30-year average) and standard deviation for each climate division. Your instructor will assign each group one climate division to examine. Find the Z-score of each average temperature using the formula below. Round each Z-score to one decimal place.

$$Z-score = \frac{Value - Mean}{Standard\ Deviation}$$

Central Ohio (Mean = 75.4, Std. Dev. = 2.6)

Month/Year	Average Temperature (Data Value)	Difference (Value - Mean)	Z-score
July 2017	73		
July 2018	73.8		
July 2019	75.9		
July 2020	76.9		

Southwest North Dakota (Mean = 68.7, Std. Dev. = 3.3)

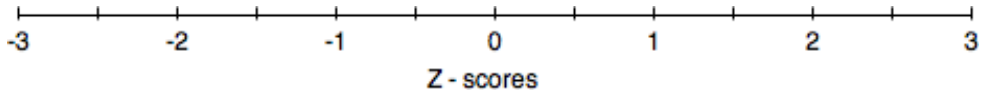
Month/Year	Average Temperature (Data Value)	Difference (Value - Mean)	Z-score
July 2017	72.2		
July 2018	69.3		
July 2019	69.3		
July 2020	70.3		

South Central Florida (Mean = 83.3, Std. Dev. = 1.3)

Month/Year	Average Temperature (Data Value)	Difference (Value - Mean)	Z-score
July 2017	82.8		
July 2018	82.4		
July 2019	82.8		

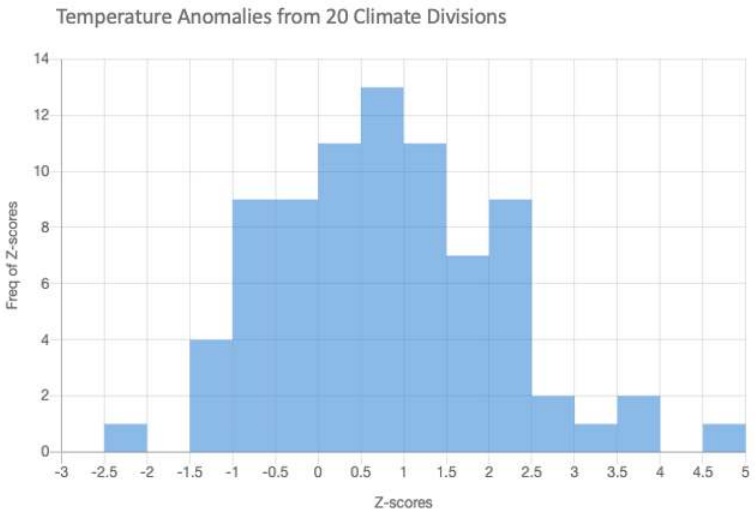
July 2020	83.0		

6. The Z-scores from Question 5 are measures of temperature anomalies. Based on the climate division that you examined, in which year was the average temperature in July the most unusual? Explain.
7. What does it mean if a temperature anomaly (Z-score) is positive? What does it mean if it is negative? What does it mean if it is 0?
8. (a) Create a dotplot by combining the twelve temperature anomalies (Z-scores) from the three climate divisions in Question 5. **Note**: If completing this problem online, follow the instructions given online to create your graph.



(b) What can we conclude about July average temperatures in these three climate divisions from 2017 to 2020? Do these temperature anomalies suggest an increase in temperature in these regions compared to the respective 30-year averages?

9. The frequency histogram below summarizes temperature anomalies (Z-scores) of July average temperatures in 2017 to 2020 for **twenty climate divisions** across the U.S. These climate divisions were selected from each geographic region in the U.S. The histogram contains eighty Z-scores. Use the histogram to answer the questions below.



(a)What is the shape of this distribution? What does this tell us about temperature anomalies in these twenty climate regions during this time period?

(b) What can we conclude about July average temperatures in these twenty climate divisions from 2017 to 2020? Do these temperature anomalies suggest an increase in temperature in these regions compared to the respective 30-year averages?

(c) Which distribution: the dotplot from Question 8, or the histogram from Question 9, provides us a better understanding of the effect of climate change on July temperatures across the U.S.? Explain.

10. In the questions above, we saw that Z-scores can be used to measure temperature anomalies and that we can combine Z-scores from multiple geographic locations and analyze the collection of Z-scores. This is what we did to create the dotplot in Question 8 and the histogram in Question 9. Suppose we instead only collected the *differences* between average temperatures and their means (30-year averages). Could we combine these temperature differences, place those differences on the dotplot (as opposed to the Z-scores we used in Questions 8 and 9) and analyze them in the same way? Explain. **Note**: A difference for one climate division is the numerator of the Z-score.

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

[28] https://www.ncei.noaa.gov/access

[29] https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/statewide/time-series

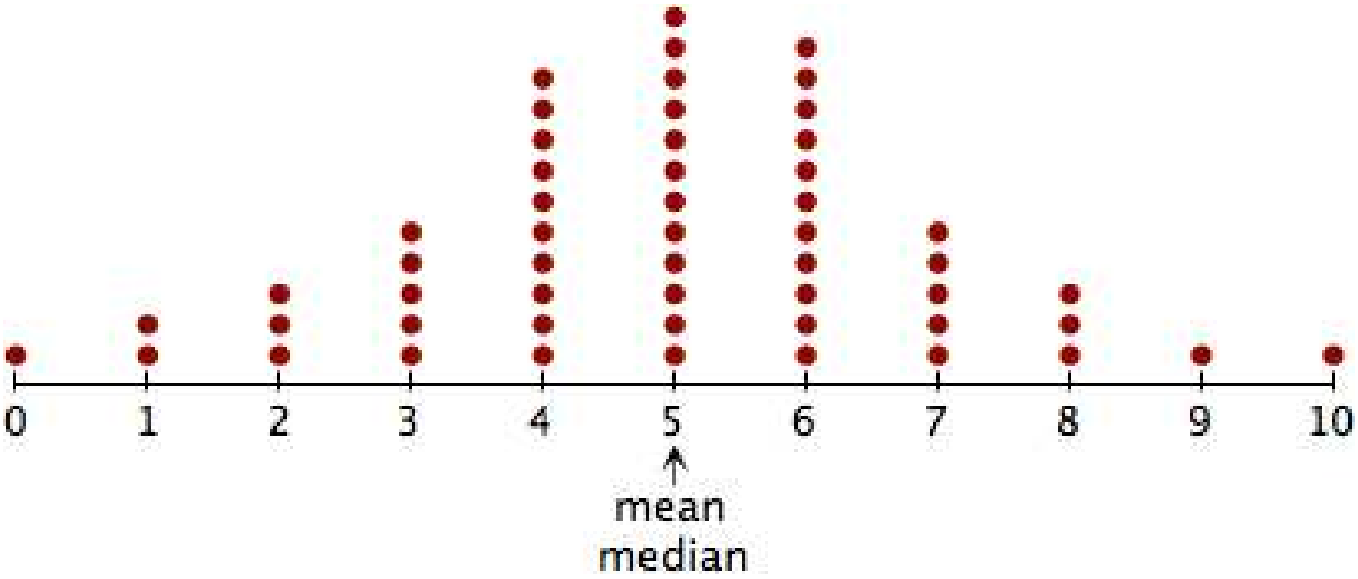
This page titled 3.6: Exploring Climate Variability is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.6.1: Preparation S.6

In the previous unit we examined how to describe the center and variability of a data distribution. Now we examine the concept of the *shape* of a distribution and ways in which individual values in a distribution can be characterized as *usual* or *unusual*.

Shape: The shape of a data distribution is a description of the overall pattern of the distribution. Below, we will discuss three basic types: bell-shaped, skewed to the right, and skewed to the left.

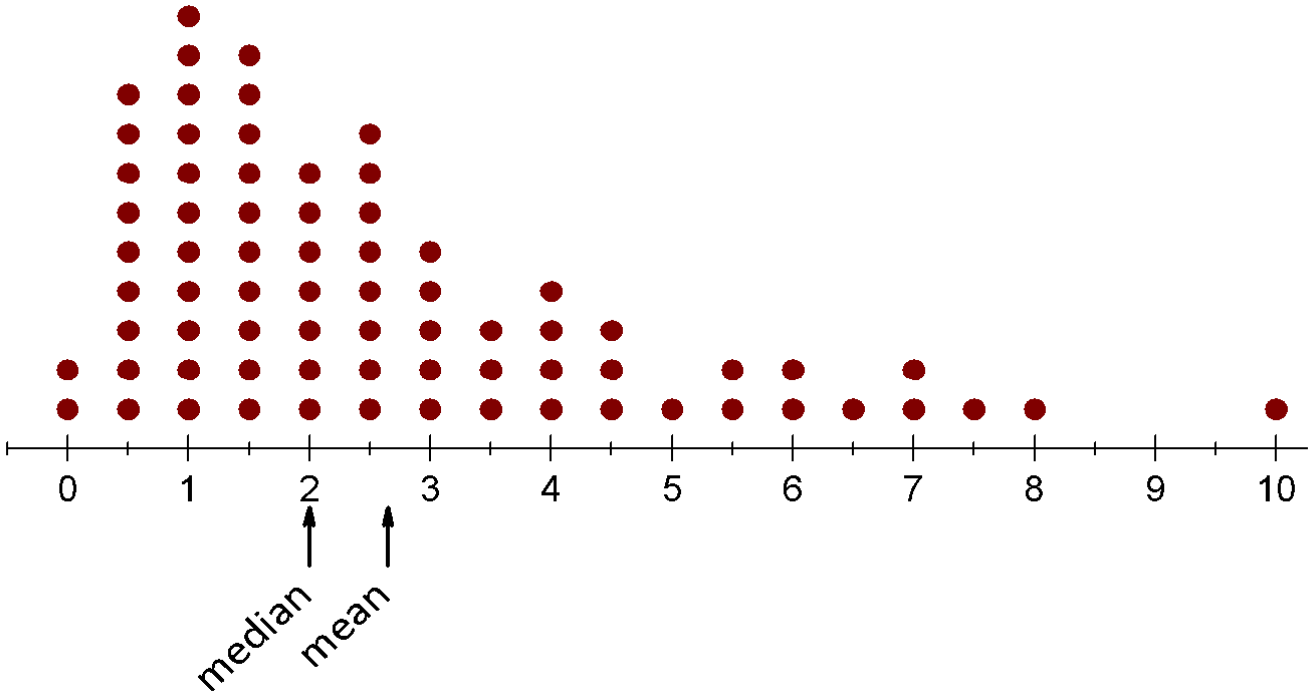
Bell-Shaped



In a **bell-shaped** distribution, the mean and median are very similar and are close to the middle of the distribution. The number of data values less than the mean is similar to the number of data values greater than the mean. The distribution is also approximately symmetric, so the left and right sides will resemble mirror images. In a perfect bell-shaped distribution, the median will also be the mode, and as you move away from the median in either direction the frequency of data points will become fewer and fewer. This is a common shape for many distributions of naturally occurring quantities, such as the distribution of the heights of a large *random sample* of men.

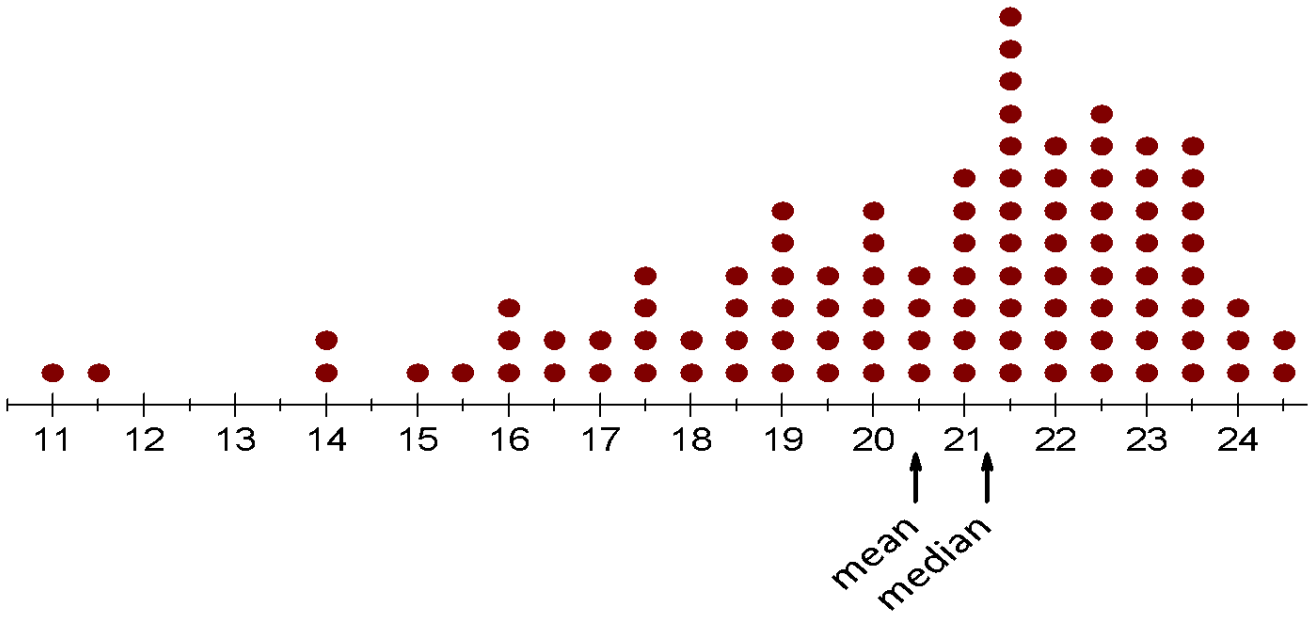
The **normal distribution** is a specific and important type of bell-shaped distribution. The normal distribution is used in probability theory and represents a *perfectly* symmetrical bell-shaped distribution, which follows an exact formula where only the mean and standard deviation are needed to draw the distribution. In the normal distribution, the mean, median, and mode are the same and represent the center of the bell. The height and width of the bell depend on the standard deviation. Normal distributions with larger standard deviations will have more spread out data, so the distribution will be shorter and wider.

Skewed to the Right Distribution



When data are **skewed to the right**, the right side trails off and is more spread out than the left side, and we say that the right "tail" is longer than the left "tail." When this occurs, the mean is pulled towards the right tail, and the mean tends to be greater than the median. The mode may be farther to the left in these distributions (e.g., in the dotplot above, the mode is 1). An example of a data set which might be skewed to the right is a random sample of salaries in a city, in which a small number of salaries are very large in comparison to the rest of the salaries.

Skewed to the Left Distribution



When data are **skewed to the left**, the left side trails off and is more spread out than the right side, and we say that the left "tail" is longer than the right "tail." When this occurs, the mean is pulled towards the left tail, and the mean tends to be less than the median. The mode may be farther to the right in these distributions (e.g., in the dotplot above, the mode is 21.5). An example of a

data set which is skewed to the left is a random sample of students' scores on an easy assignment worth 25 points, since most do very well, but a smaller number of students may still struggle or only partially complete the assignment.

Now let's consider some example data sets. Based on the description of the following data sets, determine the data set's shape and justify your answer.

1. The final grades of students in a large biology class are as follows: 15 A's, 18 B's, 10 C's, 0 D's, and 4 F's. The corresponding numerical grades are based on a scale of 100 points. The data set consists of the students' numerical grades arranged in increasing order.
2. According to the National Center for Health Statistics, in the year 2019, there were approximately 2.3 million births in the U.S. in which babies were born at full term (after a gestation period of 39–40 weeks). Typical birth weights range from 5.5 to 10 pounds. The data set consists of a random sample of 200 birth weights arranged in increasing order.
3. (a) The following frequency table shows home prices in a small community. Is the distribution skewed left, bell-shaped, or skewed right?

Price range	*Number of homes (frequency)*
$50,000 - $74,999	1
$75,000 - $99,999	4
$100,000 - $124,999	10
$125,000 - $149,999	39
$150,000 - $174,999	41
$175,000 - $199,999	12
$200,000 - $224,999	3
$225,000 - $249,999	2

(b) The following frequency table shows home prices in a small community. Is the distribution skewed left, bell-shaped, or skewed right?

Price range	*Number of homes (frequency)*
$50,000 - $74,999	32
$75,000 - $99,999	46
$100,000 - $124,999	51
$125,000 - $149,999	35
$150,000 - $174,999	9
$175,000 - $199,999	1
$200,000 - $224,999	0
$225,000 - $249,999	1

(c) The following frequency table shows home prices in a small community. Is the distribution skewed left, bell-shaped, or skewed right?

Price range	*Number of homes (frequency)*
$50,000 - $74,999	2
$75,000 - $99,999	0
$100,000 - $124,999	0
$125,000 - $149,999	1
$150,000 - $174,999	5
$175,000 - $199,999	15
$200,000 - $224,999	35
$225,000 - $249,999	27

4. The following frequency table shows the weights of nine hundred 1-yr old male Siberian Huskies. Is the distribution skewed left, bell-shaped, or skewed right?

Weight range (kilograms)	*Frequency (number of 1-yr old male Siberian Huskies)*
15 to 15.99	2
16 to 16.99	8
17 to 17.99	31
18 to 18.99	82
19 to 19.99	137
20 to 20.99	179
21 to 21.99	181
22 to 22.99	139
23 to 23.99	93
24 to 24.99	35
25 to 25.99	9
26 to 26.99	4

Z-scores (Standard Scores)

When a distribution is bell-shaped and the mean and standard deviation are known, we can find a data value's relative position by calculating its *standardized score* or **Z-score**. A Z-score describes the number of standard deviations that a data value is from the mean. The Z-score formula is shown below.

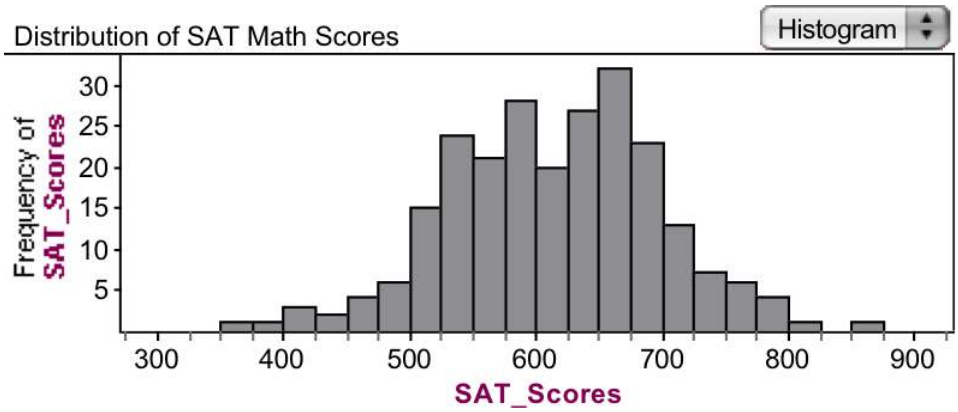
$$Z-score = \frac{Value - Mean}{Standard\ Deviation}$$

Every data value in a distribution has a Z-score. Z-scores can be negative, positive, or zero. When a

Z-score is less than or equal to −2, or greater than or equal to 2, we say that the corresponding data value is *unusual*, since the data value is relatively far from the mean. In fact, only 5% of data values in a bell-shaped distribution will have a Z-score less than or equal to −2, or greater than or equal to 2.

SAT Mathematics Test Scores

The frequency histogram below summarizes a distribution of SAT Mathematics scores from a sample of high school students in the Midwest. The distribution of SAT scores has a mean of 608 and a standard deviation of 82. Since the distribution is bell-shaped, we can use Z-scores to characterize values (test scores) as being unusual.



5. (a) SAT Mathematics scores of three students in this distribution are shown in the table below. Find the Z-score of each SAT Mathematics score. Round your answers to 2 decimal places.

Student	**SAT Math Score**	**Z-score**
Jose	720	

Destiny	575	
James	798	

(b) Do any of the three students have an unusual test score? Explain.

6. Read the following article that describes methods which NASA uses to examine and measure global warming and climate change: https://pubs.giss.nasa.gov/abs/ha02700w.html. After reading the article, explain what the researchers reported regarding temperature trends and unusual temperatures.

After Preparation S.6 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration S.6, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
identify the shape of a distribution of data values.	
interpret the standard deviation of a data set.	
compute the Z-score of a value given the mean and standard deviation of a distribution.	

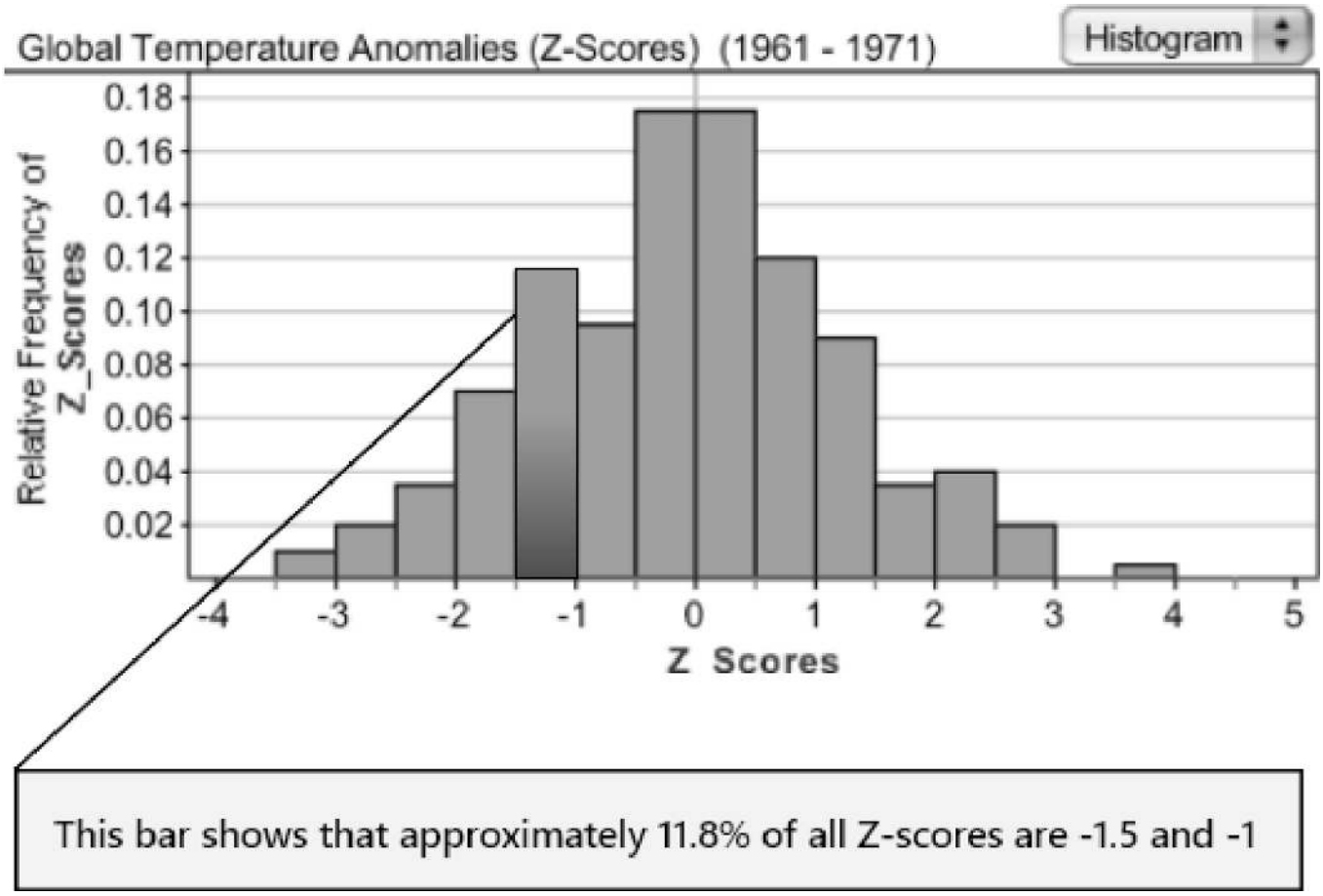
This page titled 3.6.1: Preparation S.6 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

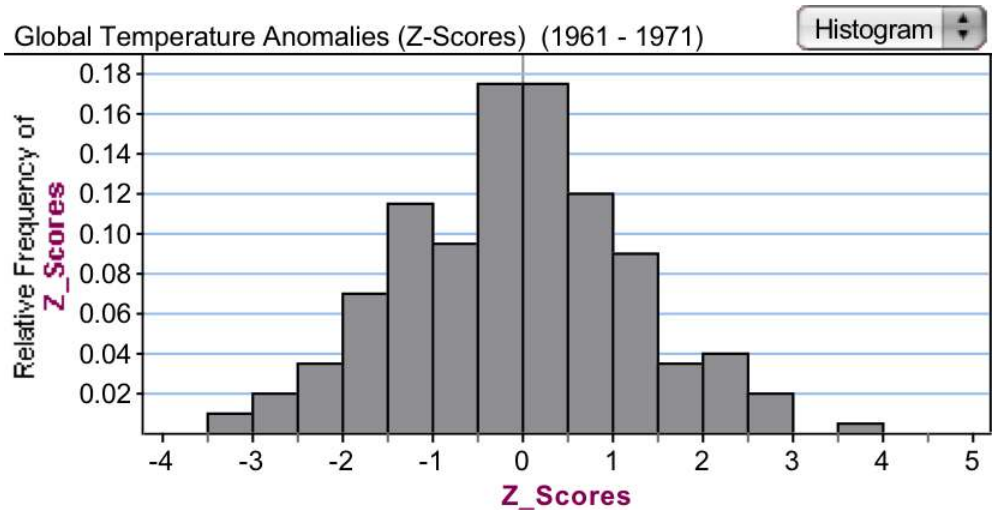
3.6.2: Exercise S.6

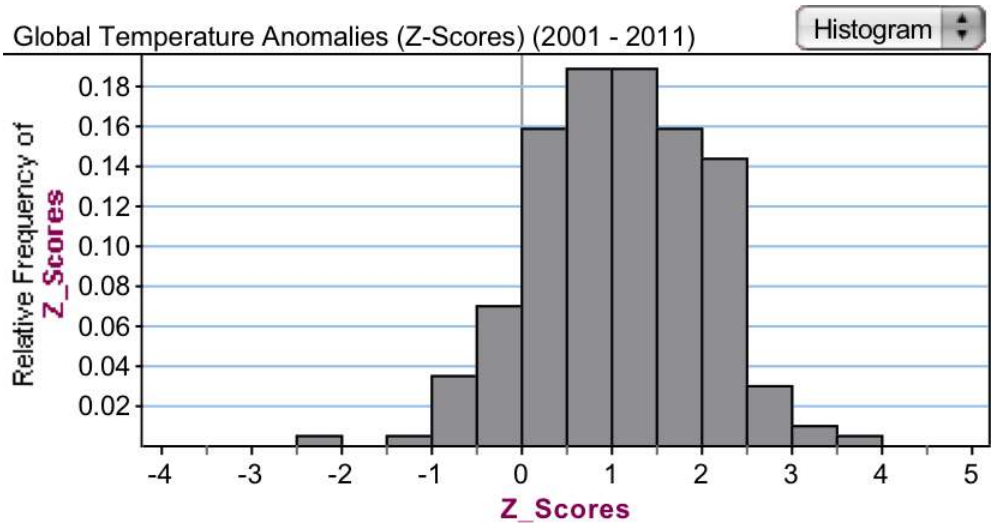
Modeling Temperature Anomalies Using Histograms

The NASA science brief you read in Preparation S.6 described how global seasonal temperature anomalies in summer months changed from 1950 to 2011.[30] The seasonal temperature anomalies are Z-scores found using 30-year averages from the base time period 1951 to 1980. The Z-scores are based on temperature data from climate stations throughout the world.

The **relative frequency histograms** below model global temperature anomalies (Z-scores) during the summer months for two time periods: 1961–1971 and 2001–2011. Each bar in the histograms represents a range of values. The height of a bar shows the relative frequency of all global seasonal temperature anomalies whose value is represented by the range of that bar. Note that we sometimes represent relative frequencies in their percentage form, as seen below. For example, 0.05 means 5%.







1. Compare the center, shape, and spread of these two distributions. How are they similar? How are they different?
2. Temperature anomalies (Z-scores) that are greater than or equal to 2.0, or less than or equal to −2.0, correspond to temperatures that are highly unusual, with respect to the base period of 1971 to 2000. How is the occurrence of highly unusual temperatures different in these two distributions?
3. Use the relative frequency histograms above to estimate the following percentages. Each percentage can be found by adding up relative frequencies.

(a) Estimate the percentage of temperature anomalies in the 1961 to 1971 time period that are less than or equal to −2.0? (**Hint**: This can be estimated by the relative frequency of the bars that are to the left of −2.0.)

(b) Estimate the percentage of temperature anomalies in the 1961 to 1971 time period that are greater than or equal to 2.0? (**Hint**: This can be estimated by the relative frequency of the bars that are to the right of 2.0.)

(c) Estimate the percentage of temperature anomalies in the 2001 to 2011 time period that are less than or equal to −2.0? (**Hint**: This can be estimated by the relative frequency of the bars that are to the left of −2.0.)

(d) Estimate the percentage of temperature anomalies in the 2001 to 2011 time period that are greater than or equal to 2.0? (**Hint**: This can be estimated by the relative frequency of the bars that are to the right of 2.0.)

4. How do these two distributions of temperature anomalies support the consensus among climate scientists that the earth is warming?

5. Z-scores can be used to assess whether data values in a certain distribution are unusual. This is appropriate when the data distribution is approximately bell-shaped. The National Oceanic and Atmospheric Administration (NOAA) reports that, from 1991 to 2020, July average temperatures in Fort Wayne, Indiana had a mean of 73.7°F with a standard deviation of 2.9°F.[31]

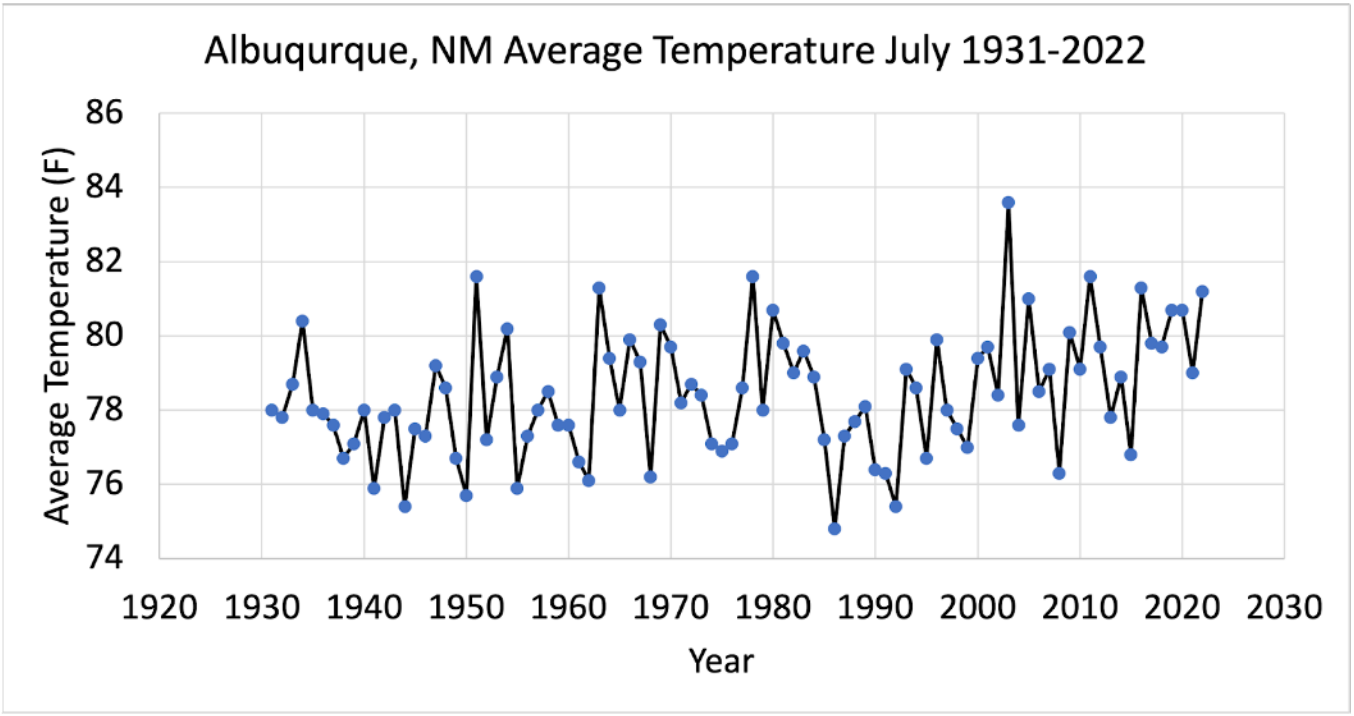
1. Is it reasonable to assume that July average temperatures in Fort Wayne, Indiana from 1991 to 2020 follow a bell-shaped distribution? Explain.

2. In 2011, the July average temperature in Fort Wayne, Indiana was 79.8°F. Compute and interpret the Z-score of this average temperature. The Z-score is found by the formula:

$$Z - score = \frac{Value - Mean}{Standard\ Deviation}$$

(c) A July average temperature in Fort Wayne, Indiana had a Z-score of −1.2. What is the July average temperature that corresponds to this Z-score?

6. The graph below displays the July average temperatures in Albuquerque, New Mexico from 1931 and 2022.[32] Historical climate data indicates that July average temperatures in this region have a mean temperature of 78.4°F and with a standard deviation of 1.67°F.



Use the mean of 78.4°F and standard deviation of 1.67°F to answer the following questions.

1. Find and interpret the Z-score of the July 2005 average temperature of 83.6°F. Round your answer to two decimal places.

2. What can you say about the Z-scores of July average temperatures in Albuquerque, NM from 2017 to 2022? **Hint**: No need to calculate them. Simply consider how each year's temperature compares to the historical average.

---

[30] https://pubs.giss.nasa.gov/abs/ha02700w.html

[31] https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/city/time-series

[32] https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/city/time-series

This page titled 3.6.2: Exercise S.6 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.7: Modeling with Normal Distributions

INTRODUCTION

Recall from Preparation S.7 that Normal distributions have three key properties:

- The mean, median, and mode are all equal.
- The normal distribution is symmetric about the mean.
- The total area below a normal curve is 1, which is 1.00 or 100%, since all (100%) of the values must fall somewhere within the distribution.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- normal distributions model bell-shaped and symmetric distributions.
- normal curves are defined based on their mean and standard deviation.
- the area below a normal curve within a specific interval represents the probability that a randomly selected value is within that interval.
- statistical inference is the process of using sample statistics to make decisions about population

parameters.

- a hypothesis test is a procedure which uses a sample statistic to make a decision about the value of an unknown population parameter.
- a P-value is the probability of obtaining a sample statistic as extreme as the one observed, assuming that the population parameter is a certain value.

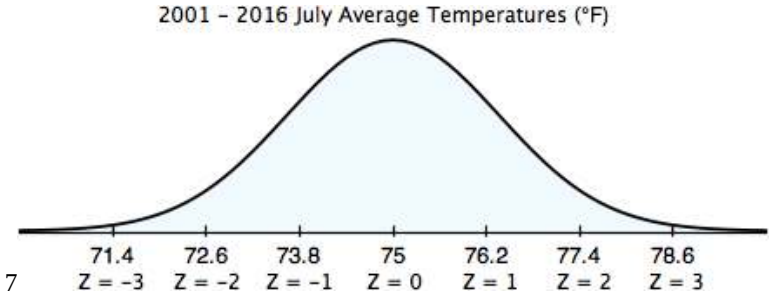
By the end of this collaboration, you should be able to

- find probabilities and percentages using a normal distribution.
- apply the Empirical Rule to determine the intervals that contain 68%, 95%, and 99.7% of data values in a normal distribution.
- perform a hypothesis test about an unknown population parameter based on sample data.
- find and interpret P-values in a hypothesis test for a population parameter.

PROBLEM SITUATION 1: SUMMER TEMPERATURES IN THE SOUTHWESTERN U.S.

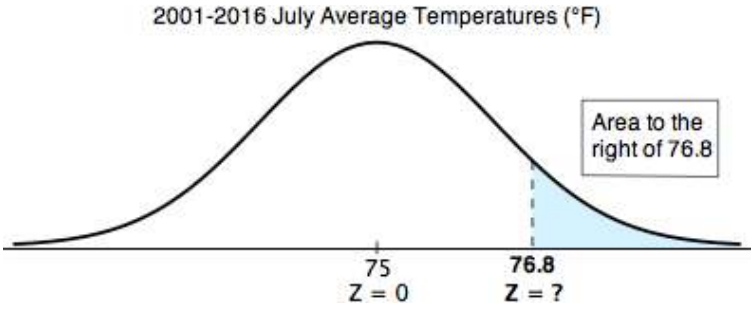
Finding Probabilities in a Normal Distribution

An important property of a normal distribution is that we can use it to find probabilities associated with any value in the distribution. From 2001–2016, July average temperatures in the Southwestern region of the U.S. were normally distributed with a mean of 75 (°F) and a standard deviation of 1.2 (°F). This normal distribution is shown below.



In the image above, note that the values below the horizontal axis correspond to the mean and the positions that are one, two and three standard deviations above and below the mean. Below these values are the corresponding Z-scores.

Suppose we want to find the probability that a July average temperature in the 2001–2016 distribution is greater than 76.8 (°F). This probability is the area to the right of 76.8 below the normal curve with mean 75 (°F) and standard deviation 1.2 (°F). The area corresponding to this probability is shown in the normal curve below.



To find the area to the right of 76.8, we can convert 76.8 to a Z-score, and then find the area to the right of the Z-score using a table created from the standard normal distribution. The standard normal distribution is the normal distribution with mean 0 and standard deviation 1. A key property of normal curves is that the area of an interval of values below any normal distribution can be found by:

1. converting the interval of values to an interval of Z-scores, and
2. finding the area of the interval of Z-scores below the standard normal distribution.

Recall that

$$Z-score = \frac{Value - Mean}{Standard\ Deviation}$$

1. Let's apply this procedure to find the probability that a randomly selected 2001–2016 July average temperature is greater than 76.8 (°F). Complete the following steps:

(a) Convert the July average temperature of 76.8 to a Z-score using the mean 75.0 and standard deviation 1.2. Round the Z-score to two decimal places.

(b) At the end of the collaboration you will find a Standard Normal Distribution table. This table provides areas below the standard normal distribution to the left of Z-scores. To use the table to find the area to the left of a Z-score: (1) go to the row that contains the Z-score's ones and tenths digits, (2) go to the column that contains the Z-score's hundredths digit, (3) find the intersection of the row and column to obtain the area.

Using the Standard Normal Distribution table, find the area to the left of the Z-score you found in Question 1(a). Write this area as a decimal and a percentage.

(c) What is the area to the right of the Z-score you found in Question 1(a)? Write this area as a decimal and a percentage. (Hint: Remember that the total area below the standard normal distribution is 1.)

(d) What is the probability that a randomly selected 2001–2016 July average temperature is greater than 76.8 (°F)? Express this answer as a percentage. (Hint: The probability that a July average temperature is greater than 76.8 (°F) is the area to the right of the Z-score for 76.8.)

2. Apply the process you used in Question 1 to find the probability that a randomly selected 2001–2016 July average temperature is less than 74.0 (°F). **Hint**: Begin by converting the average temperature of 74.0 (°F) to a Z-score. Round this Z-score to two decimal places.

PROBLEM SITUATION 2: TESTING A CLAIM ABOUT CLIMATE CHANGE

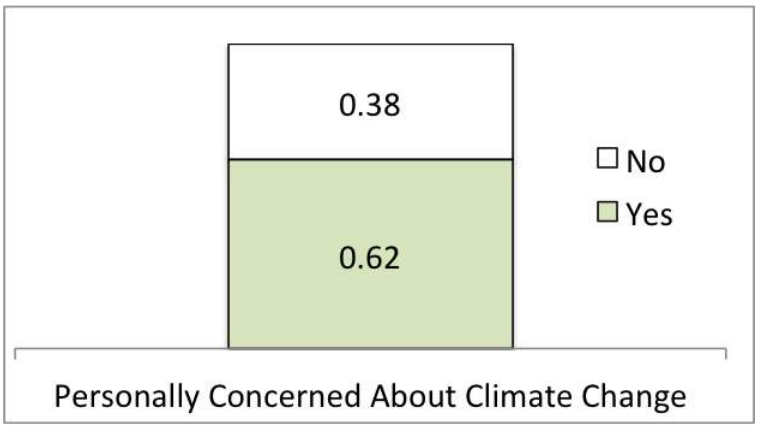
The Power of Samples

We can use normal distributions and calculating areas for certain intervals, as we did above, to help with statistical inference. **Statistical inference** is the process of using sample statistics to make decisions about population parameters. This is a powerful tool for understanding and describing large populations, as samples are often all that we can obtain. Fortunately, well-chosen samples are powerful, and when they are combined with probability tools, samples enable us to draw logical inferences about very large populations.

College Students' Views on Climate Change

A student at a Florida community college surveyed students at her college on their views about climate change. In the student's random survey of 100 other students, 62 reported that they have concerns that climate change will personally impact themselves or

a family member. The sample is representative of all students at the college. The sample results are summarized in the bar graph below.



3. 62% of the sample reported that they have concerns that climate change will personally impact themselves or a family member. 62% is a sample statistic. Identify the population parameter that corresponds to this sample statistic.
4. Another student views these sample results and claims that the sample provides evidence that more than half (i.e., more than 50%) of all students at the college have personal concerns about the effects of climate change. What do you think? Does this sample statistic enable us to conclude with certainty that more than half of all students at the college hold this view? Explain.

Using a Sample Statistic to Test a Claim About an Unknown Population Parameter

Does the observed sample proportion (0.62) provide compelling evidence that the population proportion is greater than 0.50? Sample statistics vary from sample to sample, so we must carefully consider how we use a sample to make an inference about a population. The population proportion could be greater than 0.50, but we cannot be certain. There are two possible relationships between the observed sample statistic (0.62) and the hypothesized population proportion (0.50).

Explanation 1: The actual population proportion is greater than 0.50. The observed sample proportion could reflect (be at or near) the actual population proportion. This would mean that the proportion of all students at the college who have concerns that climate change will personally impact themselves or a family member is greater than 0.50.

Explanation 2: The actual population proportion is less than or equal to 0.50. The observed sample proportion could be due solely to chance and sampling variability. The population proportion could actually be less than or equal to 0.50. This would mean that the proportion of all students at the college who have concerns that climate change will personally impact themselves or a family member is less than or equal to 0.50. In this case, the observed sample proportion differs from the actual population proportion and is a result of chance alone.

We don't know the actual population proportion, so we don't know which explanation above is correct. Even with samples that are chosen randomly and reflect the population, we can never be certain. We can, however, perform a *hypothesis test* to say with a certain degree of confidence which explanation is correct.

Hypothesis Tests

A **hypothesis test** is a procedure that uses probability to determine the likely explanation behind an observed sample statistic. To perform a hypothesis test on the sample data, we perform the following steps:

1. **Establish a null and alternative hypothesis about the unknown population parameter.**

The null and alternative hypotheses are competing hypotheses. When a hypothesis test compares a sample statistic to a hypothesized population parameter, the **null hypothesis** states that the population parameter is equal to a specific value. In this hypothesis test, the null hypothesis states that the population proportion is equal to 0.50. Thus, it states that half (50%) of all students at the college have concerns that climate change will personally impact them. The **alternative hypothesis** is a claim about a population parameter based on a claim or research question. The alternative hypothesis in this test is that the population proportion is greater than 0.50.

The hypotheses can be summarized as follows:

- Null Hypothesis: Population proportion = 0.50.
- Alternative Hypothesis: Population proportion > 0.50.

2. **Find the probability of observing a sample proportion as extreme as the one we observed, assuming the null hypothesis is true.**

The *fundamental* question is: if the null hypothesis is true and the population proportion is 0.50, how likely is it that we would obtain a random sample of 100 students whose sample proportion is 0.62 or more? We will find this probability. The procedure is explained below.

3. **Use the probability to make a decision about the hypotheses.**

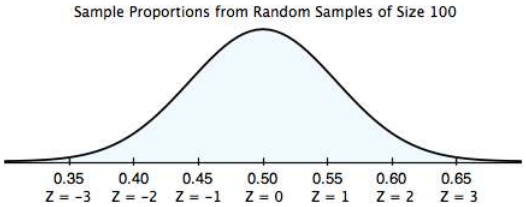
If the probability of finding such a sample proportion (0.62 or more) is very low, the observed sample proportion provides strong evidence that the population proportion is not equal to 0.50. If this occurs, we reject the null hypothesis and support the alternative hypothesis. In other words, we reject the hypothesis that the proportion of all students at the college who have personal concerns about climate change is equal to 0.50, and support the hypothesis that the population proportion is greater than 0.50. When we reject the null hypothesis, we say that the sample proportion is ***statistically significant***.

If the probability of finding such a sample proportion is not very low, we arrive at the opposite decision. In this case, we do not reject the null hypothesis and we do not support the alternative hypothesis. That is, we do not reject the hypothesis that the population proportion is 0.50, and we do not support the hypothesis that the population proportion is greater than 0.50. When we do not reject the null hypothesis, we say the sample proportion is not statistically significant.

Finding the Probability

We can find the probability of obtaining a random sample of 100 students whose sample proportion is

0.62 or more by modeling the distribution of sample proportions, from a population whose proportion is 0.50, with a normal distribution. From this population, sample proportions from random samples of size 100 are normally distributed with a mean of 0.50 and a standard deviation of 0.05. This distribution of sample proportions is displayed below.



5. Note that the mean of all sample proportions in this distribution is the assumed population proportion. Using the Empirical Rule, which sample proportions are unusual in this distribution? (**Hint**: Unusual values in a normal distribution are values that are two or more standard deviations away from the mean.)
6. The observed sample proportion is 0.62. Compute the Z-score of this sample proportion using the mean of sample proportions (0.50) and the standard deviation of sample proportions (0.05). Round the Z-score to two decimal places. Interpret this Z-score.
7. The observed sample proportion is 0.62. Compute the Z-score of this sample proportion using the mean of sample proportions (0.50) and the standard deviation of sample proportions (0.05). Round the Z-score to two decimal places.

The *percentage* of sample proportions that are 0.62 or more is equivalent to the probability of obtaining a random sample whose sample proportion is 0.62 or more. This probability is called a **P-value** (probability value). In a single sample hypothesis test, the P-value specifies the probability of randomly obtaining a sample statistic as *extreme* as the one observed if the null hypothesis is true.

8. What is the P-value for the above hypothesis test? Interpret this P-value.
9. If the population proportion is 0.50, is it likely or unlikely that a random sample of 100 students from the population would have a sample proportion of 0.62 or more?

Making a Decision about the Hypotheses

When a P-value is low, typically less than 5%, we say that the sample statistic is *statistically significant*. We then reject the null hypothesis that the population parameter is equal to a specific value, and we support the alternative hypothesis. We decide that the observed sample statistic is not due to chance alone, and instead reflects the population parameter.

Note: If a P-value other than 5% is going to be used as a cutoff, this value must be chosen before the hypothesis test is performed, to prevent confirmation bias. A P-value of 5% is typically used to ensure we are at least 95% certain before we reject the null hypothesis. For example, consider a jury trial: Would you want to convict someone, if you weren't at least 95% certain they were guilty? Consider a drug safety study. Would you want to provide a new drug to the public if you weren't at least 95% certain it was safe enough to use?

10. Is the observed sample proportion of 0.62 statistically significant? Explain.

11. (a) What should we decide about the null and alternative hypotheses?

The hypotheses can be summarized as follows:

- Null Hypothesis: Population proportion = 0.50.
- Alternative Hypothesis: Population proportion > 0.50.

(b) What can we state about the proportion of all students at the college who have concerns that climate change will personally impact themselves or a family member?

12. Can we be 100% certain about our decision from Question 11?

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

Standard Normal Distribution Table (Negative Z-Scores)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681

-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Standard Normal Distribution Table (Positive Z-Scores)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706

1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

This page titled 3.7: Modeling with Normal Distributions is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.7.1: Preparation S.7

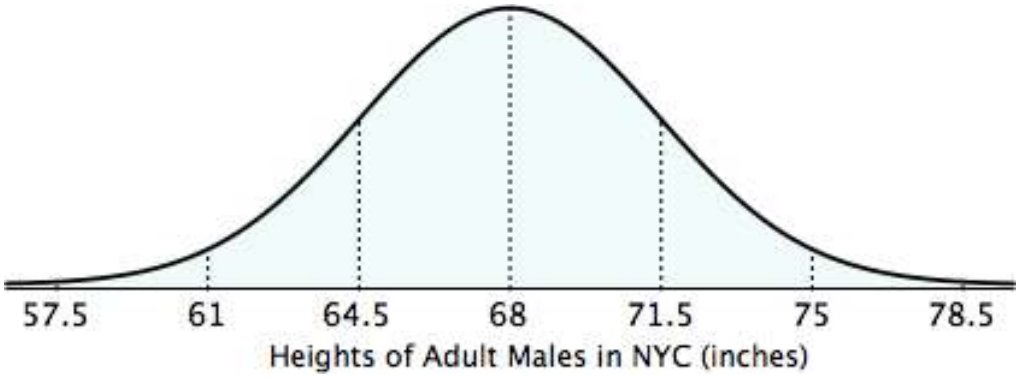
Normal Distributions

In the previous unit we saw that global temperature anomalies followed bell-shaped distributions. Temperature anomalies are just one of many real-world observations that have bell-shaped distributions. To examine properties of bell-shaped distributions and determine probabilities associated with their values, we often model bell-shaped distributions with *normal distributions*. When we model a data distribution with a normal distribution, we say that the data are **normally distributed**. Note that actual data typically will not be perfectly symmetrical and bell-shaped, so the normal distribution does not perfectly reflect the data; however, a normal distribution can still be used as a model that helps us understand and analyze the data.

Normal distributions have a wide range of applications. One application is modeling large distributions of observations in nature. For example, consider the distribution of heights of all adult men in New York City (NYC). This distribution is extremely large and contains over three million values. If we obtain a representative sample of values in this distribution, we can estimate the **population mean** and **population standard deviation**, and use these estimates of the population parameters to model the population with a normal distribution.

Suppose we obtain a representative sample of all NYC adult male heights to determine that their mean height is 68 inches with a standard deviation (SD) of 3.5 inches. We can construct a normal distribution based on these estimated parameters and draw a **normal curve** to represent the distribution. Normal curves are graphs of normal distributions. The graph of the normal distribution with a mean of 68 and SD of 3.5 is shown below.

Figure 1: Normal Curve Representing the Distribution of Heights of NYC Adult Males



Note that the normal curve is centered on the mean. This normal distribution is a theoretical model of the large distribution of heights of NYC adult males. Normal distributions are examples of *continuous probability distributions* because they enable us to find probabilities corresponding to any interval of data values.

Normal distributions have three key properties:

1. The mean, median, and mode are all equal.
2. The normal distribution is symmetric about the mean.
3. The total area below a normal curve is 1, which is 1.00 or 100%, since all (100%) of the values must fall somewhere within the distribution. (**Note**: this will be discussed later in the lesson.)

Before going further, answer the questions below about the normal distribution of heights of all adult males in NYC.

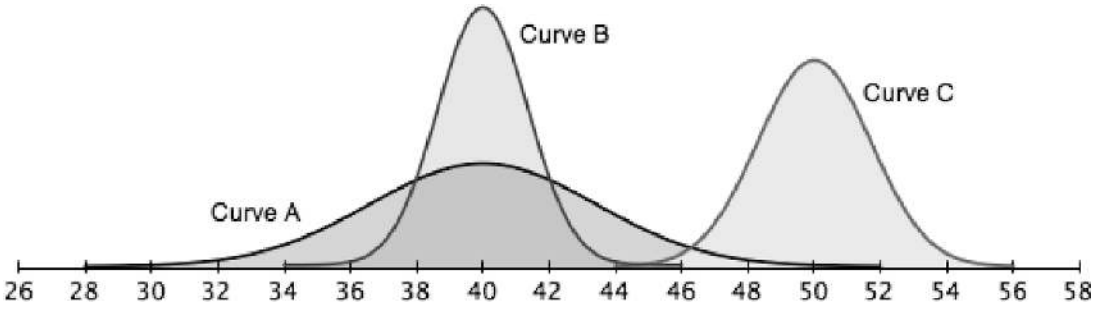
1. The mean, median, and mode of heights of adult males in NYC are equal to 68 inches. What does this tell us?
2. The distribution of heights is symmetric about the mean. What does this tell us?
3. Heights of adult males in NYC have a standard deviation of 3.5 inches. What does this tell us about the heights in this distribution?

Normal Distributions

Normal distributions have the same shape, but vary based on their center (mean) and spread (standard deviation). Figure 2 below shows the graphs of three normal distributions (Curve A, Curve B, and Curve C).

4. Match the graphs of the normal distributions with the appropriate means and standard deviations (SD) given in the questions below.

Figure 2: Curves A, B, and C



(a) Mean = 40, SD = 1.2.

(b) Mean = 50, SD = 1.5.

(c) Mean = 40, SD = 3.0.

Empirical Rule

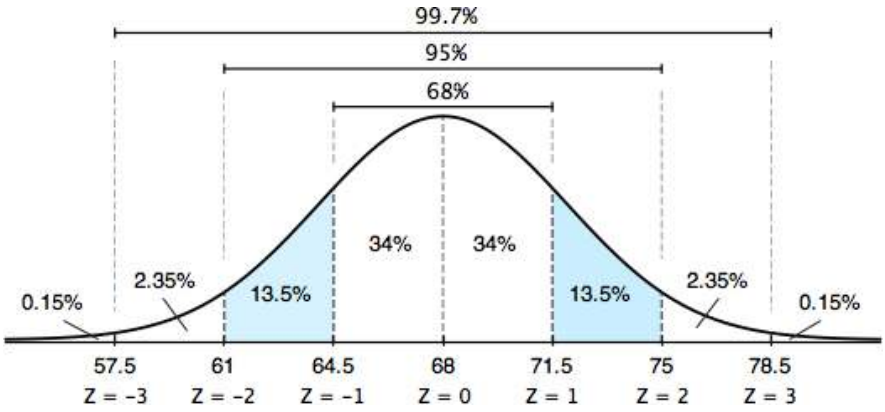
The Empirical Rule describes the variability of data values in a normal distribution. It specifies the approximate percentage of data values within any normal distribution which lie within one, two, and three standard deviations from the mean.

In any normal distribution:

- About 68% of data values lie within 1 standard deviation of the mean.
- About 95% of data values lie within 2 standard deviations of the mean.
- About 99.7% of data values lie within 3 standard deviations of the mean.

To illustrate this rule, let's apply it to the specific normal distribution with a mean of 68 and a standard deviation of 3.5. This is the distribution we saw in Figure 1 (heights of all NYC adult males). Figure 3 below applies the Empirical Rule to that distribution. Note that the values below the horizontal axis correspond to the mean and the positions that are one, two and three standard deviations above and below the mean. Below these values are the corresponding Z-scores.

Figure 3: Empirical Rule Applied to Normal Distribution with Mean = 68 and SD = 3.5



In addition to showing the intervals that contain approximately 68%, 95% and 99.7% of the data values in the distribution, Figure 3 also shows the approximate percentage of data values in specific subintervals. For example, in this normal distribution, approximately 68% of data values are between

64.5 and 71.5. This interval, 64.5–71.5, contains data values that are within one standard deviation from the mean and is said to contain the **middle** 68% of data values. The area under the distribution in this interval is then 0.68. Note that the intervals that contain 0.15% of the data include all data below 45 and all data above 75, respectively.

It is important to note that the sum of the percentages in these eight regions is 100%. The total area below a normal curve is 1, representing the fact that 100% of values in the distribution are contained within it. We can use the percentages (areas) in these subintervals to determine the percentage of data values in other intervals.

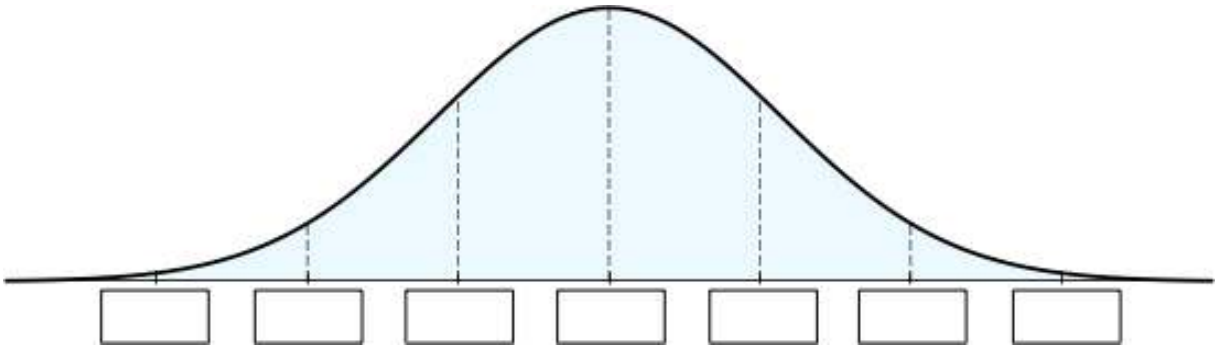
5. Use Figure 3 to answer the questions below. What percentage of data values in this normal distribution are:

(a) Between 68 and 75?

(b) Less than 64.5?

(c) Greater than 75?

6. Suppose we found that the heights of U.S. adult females are normally distributed with a mean of 63.4 inches and a standard deviation of 2.3 inches. In the normal curve below, fill in the boxes below the x-axis by specifying the mean of the normal distribution and the positions that are one, two and three standard deviations from the mean. Write each value in its corresponding box.



7. What interval contains the middle 95% of U.S. adult female heights?
8. What interval contains the middle 99.7% of U.S. adult female heights?
9. What percentage of U.S. adult females have heights less than 65.7 inches?
10. In Unit S.6, we saw that data values in a bell-shaped distribution that have Z-scores less than or equal to −2, or greater than or equal to 2, are unusual in their distribution. Using this same definition of an unusual observation, what heights would be considered unusual for U.S. adult females? (**Hint**: Identify the heights that are two or more standard deviations from the mean in both directions.)

After Preparation S.7 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

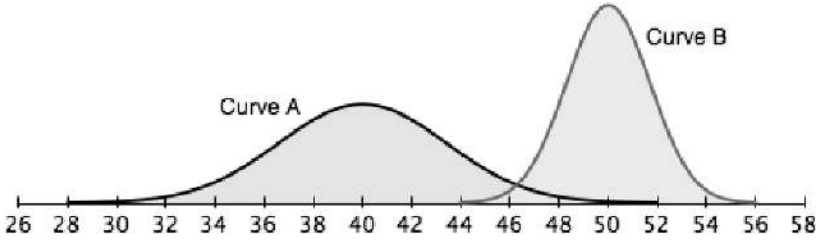
Before beginning Collaboration S.7, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
identify the shape of a data distribution.	
compute the Z-score of a data value in a data set given the mean and standard deviation of the data set.	

This page titled 3.7.1: Preparation S.7 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

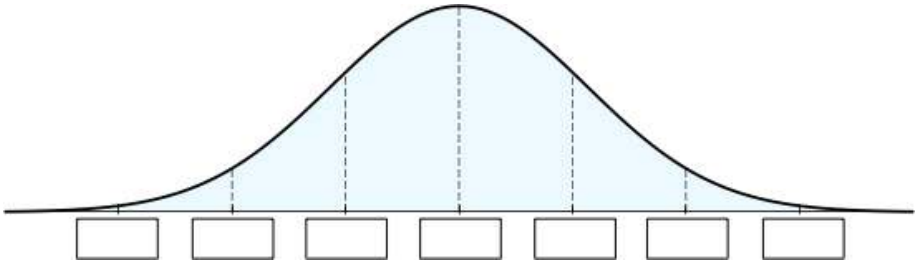
3.7.2: Exercise S.7

The image below displays two normal curves.



1. Estimate the mean of each normal curve.
2. Which normal curve has a greater standard deviation? How do you know?
3. In 2021, SAT mathematics scores of high school graduates in New Hampshire had a mean of 526 and a standard deviation of 109.[33] We will assume that this distribution of test scores is normally distributed.

(a) In the normal curve below, fill in the boxes below the x-axis by specifying the mean of the normal distribution and the positions that are one, two and three standard deviations from the mean. Write each value in its corresponding box.



(b) Using the Empirical Rule, determine the interval of SAT mathematics scores that contains the middle 95% of the distribution.

(c) Find the probability that a randomly selected SAT mathematics test score in this distribution is less than 450. Express the probability as a percentage. Round the Z-score to two decimal places, and express its P-value as a percentage rounded to two decimal places.

(d) What percentage of SAT mathematics test scores in this distribution are greater than 600? Express the probability as a percentage.

(e) A student who earned a test score of 640 claims that their score is in the top 10% of the distribution. Is the student's claim accurate? Explain.

College Students' Views on the Security of Free Speech

4. A student at a community college in Connecticut surveyed students at his college on their views about free speech in the U.S. The survey asked students if they feel that free speech is under attack in the U.S. In a random survey of 100 students, 77 stated that free speech is secure in the U.S. The sample is representative of all students at the college. Does this sample data provide convincing evidence that the proportion of all students at the college who believe that free speech is secure is greater than 75%? We can perform a hypothesis test to answer this question. Complete the steps below.

(a) What is the sample proportion in this hypothesis test?

(b) The hypotheses for this hypothesis test are:

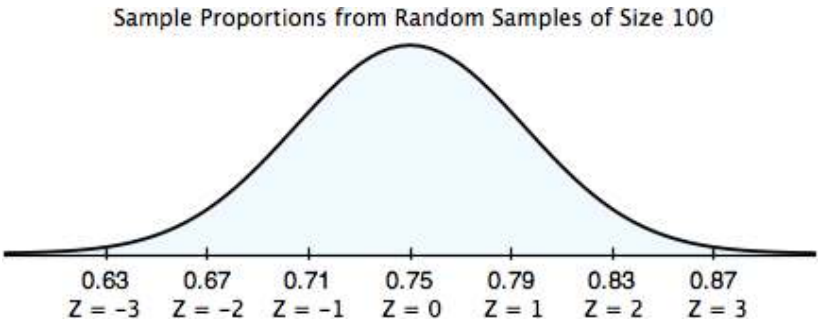
Null Hypothesis: Population proportion = 0.75.

Alternative Hypothesis: Population proportion > 0.75.

The null hypothesis states that 75% of students at the college believe that free speech is secure in the U.S. today. The alternative hypothesis is based on the research question. It states that more than 75% of all students at the college believe that free speech is secure in the U.S. today.

Assuming the null hypothesis is true (population proportion = 0.75), we will find the probability of obtaining a random sample of 100 students whose sample proportion is as extreme as the one we observed (0.77 or more).

We can find this probability by modeling the distribution of sample proportions with a normal distribution. Sample proportions from random samples of size 100, from a population whose proportion is 0.75, are normally distributed with mean of 0.75 and a standard deviation of 0.04. This distribution of sample proportions is displayed below.



The observed sample proportion is 0.77. Compute the Z-score of this sample proportion using the mean of sample proportions (0.75) and the standard deviation of sample proportions (0.04). Round the Z-score to two decimal places.

(c) The percentage of sample proportions that are 0.77 or more is equivalent to the probability of obtaining a random sample whose sample proportion is 0.77 or more. This probability is the

P-value (probability value). Find the P-value. The P-value is the area to the right of the Z-score.

(d) Let's make a decision. Is the observed sample proportion of 0.77 statistically significant? Explain.

(e) What should we decide about the null and alternative hypotheses?

(f) What should we decide about the proportion of all students at the college who believe that free speech is secure in the U.S.?

---

[33] https://nces.ed.gov/programs/digest/d21/tables/dt21_226.40.asp

This page titled 3.7.2: Exercise S.7 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.8: Significance in Observational Studies

INTRODUCTION

Please discuss the following questions with your group:

- What specific concept from earlier in Module S do you feel fairly confident about?
- Are there any concepts from Module S that you find challenging?

Please take turns, and make sure that every group member answers both questions. If possible, please help your group members understand the concepts that they find challenging.

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that

- statistical inference is the process of using sample statistics to make decisions about population parameters.
- a hypothesis test is a statistical procedure that uses probability to determine the likely explanation behind an observed difference in sample statistics.
- a difference in sample statistics is statistically significant if it leads us to reject the null hypothesis in a hypothesis test.
- sample results from an observational study can be generalized to populations when the results

are statistically significant.

By the end of this collaboration, you should be able to

- compute sample statistics and the difference in sample statistics based on an observational study.
- use probability to determine the likely explanation behind an observed difference in sample

statistics.

- use probability to assess whether a difference in sample statistics is statistically significant.

PROBLEM SITUATION: WOMEN IN LEADERSHIP

In the previous collaboration, we learned that hypothesis tests enable us to use a single sample statistic to make a decision about a single population parameter. In this collaboration, we will learn how hypothesis tests enable us to use two sample statistics to make a decision when comparing two population parameters.

The Pew Research Center poll that you read about in Preparation S.8 surveyed U.S. adults on their views regarding women in leadership positions in business and government.[34] The survey addressed a range of topics, including perceived barriers to increased leadership by women, and the impact of female leadership on society.

1. The results of the Pew survey showed that men and women have differing views on several topics. To examine this difference, we will focus on one specific question in the survey: “In general, do you think men or women in high political offices are better at handling economic conditions?” When men and women were asked this question, approximately 16% of women and 12% of men responded that women are better at handling economic conditions. The poll was conducted by asking this question of 1,835 adults (914 women and 921 men) in the U.S. What do these statistics enable us to infer about the populations?

(a) Define the samples and populations in this observational study.

(b) In this study, the sample statistics are the sample proportions 16% and 12%. Sample statistics estimate population parameters. Describe the population parameters that correspond to these two statistics.

(c) The difference in sample proportions—the sample proportion of women minus the sample proportion of men—is also a statistic, because this difference is based on sample data. Compute the difference: sample proportion of women minus sample proportion of men.

(d) The statistic you calculated in Question 1(c) shows that a difference exists between the *sample proportions* of women and men who believe that women in high political offices are better than men in high political offices at handling economic conditions. Can we decide with absolute certainty that the proportion of *all* U.S. women who think women are better than men at handling economic conditions is greater than the proportion of *all* U.S. men who feel the same way? Explain.

What Do These Sample Statistics Tell Us About the Population Parameters?

Performing a *hypothesis test* on sample statistics from two samples enables us to make a decision about the relationship between the corresponding population parameters. In the example above, there are two explanations for the observed difference of 0.04 (4 percentage points) between the sample proportions.

Explanation 1: The actual population proportions are different. The observed difference in sample proportions could reflect (be at or near) the actual difference in the population proportions. In this case, one population proportion is greater than the other.

Explanation 2: The actual population proportions are not different. The observed difference in sample proportions could be due solely to chance and sampling variability. The population proportions could actually be equal.

Hypothesis Test

Performing a *hypothesis test* on sample statistics from two samples enables us to make a decision about which explanation above is correct. To perform a hypothesis test on the sample data from the Pew Research study above, we can perform the following steps (similar to what we did in Collaboration S.7):

1. **Establish a null and alternative hypothesis about the unknown population proportions.**

Null Hypothesis: Population proportion of all U.S. women = Population proportion of all U.S. men.

Alternative Hypothesis: Population proportion of all U.S. women > Population proportion of all U.S. men.

When comparing two samples, the null hypothesis states that there is no difference between the population parameters. The alternative hypothesis is based on the research question and, in this case, states that the population proportion of women is greater than that of men.

2. **Find the probability of observing sample proportions whose difference is as extreme as the one we observed, assuming the null hypothesis is true.**

Assuming the null hypothesis is true, we will find the probability of obtaining two random samples whose difference in sample proportions (sample proportion of women minus sample proportion of men) is 4 percentage points or more. We will explain this procedure below. (**Note**: You will not have to perform the calculations or procedures described below, but it is important to understand what the

P-value means when comparing two sample statistics.)

3. **Use the probability to make a decision about the hypotheses.**

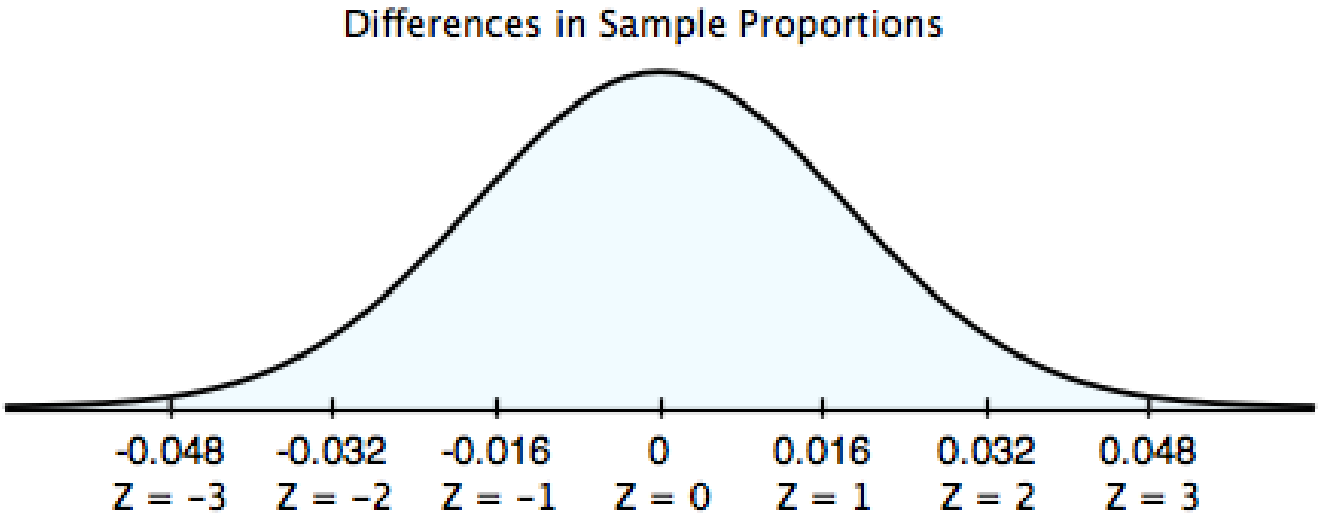
We can find the probability of obtaining a difference in sample proportions of 0.04 or more, assuming the null hypothesis is true, by modeling all possible differences in sample proportions with a normal distribution. To do this, we:

- Combine the data from both samples to establish an estimate for the common population proportion.
- Create a distribution of sample proportions for women, and create a distribution of sample

proportions for men, assuming both distributions of sample proportions come from a population with a common population proportion.

- Compute the probability of all possible differences using all possible pairs of sample proportions (sample proportion of women minus sample proportion of men) to form a distribution of differences.

The resulting distribution of differences in sample proportions is normally distributed with a mean of 0 and a standard deviation of 0.016. (Note: As noted earlier, you will not have to perform the calculations to find the mean and standard deviation of a distribution of differences. This is beyond the scope of this class.) The distribution of differences in sample proportions is displayed below.



Examining the Distribution of Differences in Sample Proportions

1. (a) Each difference in the normal distribution above is a difference between a sample proportion of women and a sample proportion of men. What does it mean for a difference to be negative? What does it mean for a difference to be 0? What does it mean for a difference to be positive? (Hint: the differences are calculated as follows: sample proportion of women minus sample proportion of men.)

(b) What do you notice about the center of this distribution of differences?

(c) What difference in sample proportions is most likely to occur? What explains why this difference occurs most frequently?

(d) What differences in sample proportions are unlikely in this distribution? (**Hint**: Unlikely observations in a normal distribution are values that are two or more standard deviations from the mean.)

Finding the P-value (Probability Value)

3. The observed difference in sample proportions is 4 percentage points (0.04) (using the original sample proportions in Question 1 above). Assuming the population proportions are equal, based on the null hypothesis, we can determine how likely it would be to randomly obtain two sample proportions, one for women and one for men, whose difference is 4 percentage points or more. To do so, we can find the probability of randomly obtaining a difference in sample proportions that is greater than or equal to 0.04. This probability is the area to the right of 0.04 under the "differences in sample proportions" normal curve shown above.

1. Convert the difference in sample proportions of 0.04 to a Z-score. Interpret the Z-score. (Hint: Use a population mean of 0 and standard deviation of 0.016, as shown in the normal distribution.)
2. Find the area to the right of the difference in sample proportions of 0.04 by finding the area to the right of the Z-score you obtained in Question 3(a). Express this area as a percentage. This is the P-value. (**Hint**: Use the Standard Normal Distribution table.)
3. If the population proportions (women and men) are equal, is it likely or unlikely that a sample proportion from a random sample of 914 women would be 4 percentage points or more greater than the sample proportion from a random sample of 921 men?

Making a Decision about the Hypotheses

Recall that when a P-value is low, typically less than 5%, we say that the difference in sample statistics is *statistically significant*. We reject the null hypothesis that the population parameters are equal, and we support the alternative hypothesis that one population parameter is greater than the other. We decide that the observed difference in sample statistics is not due to chance alone, and instead reflects the difference that exists between the population parameters.

4. Is the observed difference in sample proportions, 4 percentage points, statistically significant? Explain.
5. What should we decide about the null and alternative hypotheses?

6. What can we decide about the population proportions—the proportion of all U.S. women and the proportion of all U.S. men who believe that women in high political offices are better than men in high political offices at handling economic conditions?
7. Can we be 100% certain about our decision from Question 6?

Are Women Held to Higher Standards?

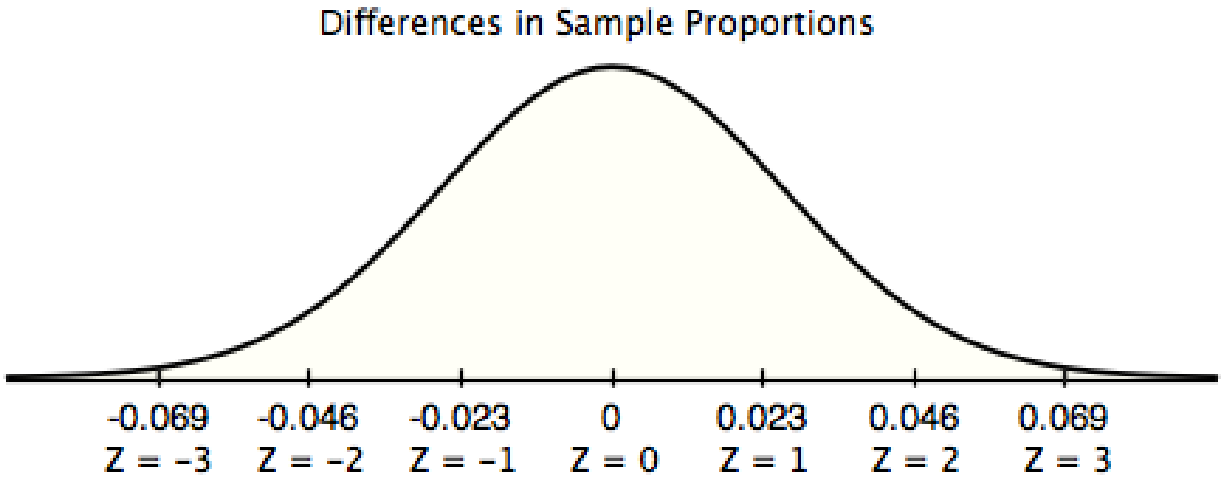
8. The Pew Research Center poll also asked survey respondents whether they agreed with the statement, "A major reason that there aren't more women in high political offices is that women are held to higher standards than men." 48% of women, and 28% of men, responded that they agreed with the statement. The poll was conducted on 1,835 adults (914 women and 921 men) in the U.S. (a) Find the difference in the sample proportions (sample proportion of women minus sample proportion of men).

Suppose you perform a hypothesis test using these sample statistics. The hypotheses for the test are:

- Null Hypothesis: Population proportion of all U.S. women = Population proportion of all U.S. men.
- Alternative Hypothesis: Population proportion of all U.S. women > Population proportion of all

U.S. men.

Assuming the null hypothesis is true, we can model all possible differences in sample proportions with a normal distribution. The distribution of differences in sample proportions has a mean of 0 and a standard deviation of 0.023 and is shown below.



(b) Convert the difference in sample proportions (from Question 8(a)) to a Z-score. Interpret the Z-score. (**Hint**: Use a population mean of 0 and standard deviation of 0.023, as shown in the normal distribution.)

(c) In this next question, first assume that the population proportions are equal, and then consider the likelihood of observing sample results that differ by the amount observed. If the likelihood is very low, they should decide that the results are statistically significant and reject the assumption that the population proportions are equal. In this example, we must ask: if the population proportions are the same, how likely would it be to obtain two large samples whose sample proportions differ by 20 percentage points or more? The likelihood seems very small, so the sample results are statistically significant and we reject the assumption that the population proportions are equal.

Based on the Z-score in Question 8(b), do you think that the observed difference in sample proportions is statistically significant? What can we decide about the population proportions? Explain.

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

---

[34] http://www.pewsocialtrends.org/2015/01/14/women-and-leadership/

This page titled 3.8: Significance in Observational Studies is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.8.1: Preparation S.8

1. The Pew Research Center conducts national polls on issues in politics, media and news, and social trends. In 2015, Pew Research polled U.S. adults on their views about women and leadership. The survey was conducted on 1,835 adults (914 women and 921 men) in the U.S. Read the first three sections (Public Says Women are Equally Qualified … , Getting the Job Done in Politics and Business, and Women in Leadership Today) of the article, "Women and Leadership," from the Pew Research Center, at http://www.pewsocialtrends.org/2015/01/14/women-and-leadership/. Then answer the questions below. (a) What is the focus of this study?

(b) When asked the question: "How much, if anything, do you think having more women in top leadership positions in business and government would do to improve the quality of life for all women?", 78% of U.S. women responded "some" or "a lot." The survey reported a margin of error of 3.4 percentage points for women at a 95% confidence level.

(i) Create an interval estimate for the proportion of all U.S. women who believe that having more women in top leadership positions in business and government would improve the quality of life for all women.

(ii) Interpret the interval estimate.

(c) When asked the question: "How much, if anything, do you think having more women in top leadership positions in business and government would do to improve the quality of life for all women?", 62% of U.S. men responded "some" or " a lot." The survey reported a margin of error of 3.5 percentage points for men at a 95% confidence level.

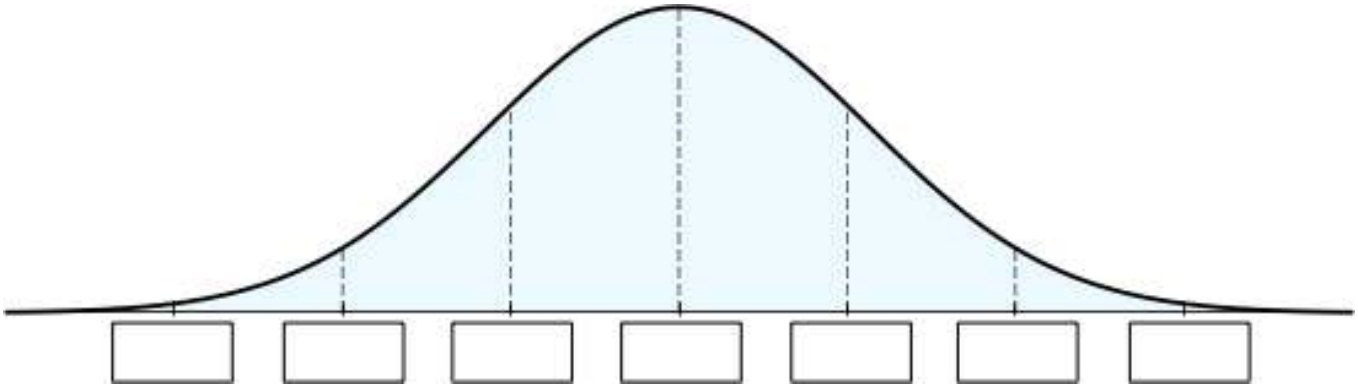
(i) Create an interval estimate for the proportion of all U.S. men who believe that having more women in top leadership positions in business and government would improve the quality of life for all women.

(ii) Interpret the interval estimate.

(d) Compare the interval estimates for U.S. women and U.S. men. Can we decide with confidence that the proportion of all U.S. women who believe that having more women in top leadership positions in business and government would improve the quality of life for all women is greater than the proportion of all men who believe the same thing? Explain.

2. A recent poll found that 20% of U.S. adults say that they trust the government in Washington, D.C., to do what is right "just about always" or "most of the time." Let's assume that 20% is the true proportion of all U.S. adults who trust the federal government to do what is right. Given this assumption about the population proportion, we can model sample proportions from this population with a normal distribution.

(a) Let's model sample proportions from random samples of size 100. Sample proportions from random samples of size 100, taken from this population, are normally distributed with a mean of 0.20 and a standard deviation of 0.04. Represent this distribution of sample proportions in the normal curve below. Fill in the boxes below the x-axis by specifying the mean of the normal distribution and the positions that are one, two and three standard deviations from the mean. Write each value in its corresponding box.



(b) In the normal distribution above, what interval contains the middle 95% of sample proportions? (Hint: Use the Empirical Rule.)

(c) A different researcher obtained a random sample of 100 U.S adults in which 25% said that they trust the federal government to do what is right. She wanted to determine how likely her sample statistic was if the population proportion is actually 20%. To answer this question, we can find the probability of randomly obtaining a sample proportion of 0.25 or more. This probability is the area to the right of 0.25 under the normal curve in Question 2(a). First, convert the sample proportion of 0.25 to a Z-score.

(d) Find the area to the right of the sample proportion 0.25 in the distribution, by finding the area to the right of the Z-score you obtained in Question 2(c). Express this area as a percentage. Interpret this percentage as a probability. (Hint: Use the Standard Normal Distribution table to find the area to the right of the Z-score.)

(e) Assuming the population proportion is 0.20, which sample statistic is more unlikely: obtaining a random sample of 100 U.S. adults in which 26% or more trust the federal government to do what is right, or obtaining a random sample of 100 U.S. adults in which 12% or less trust the federal government to do what is right? Explain.

After Preparation S.8 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration S.8, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
find the probability associated with an interval of values in a normal distribution.	
interpret the probability associated with a randomly selected sample statistic	

This page titled 3.8.1: Preparation S.8 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.8.2: Exercise S.8

1. What does it mean to say that a difference in sample statistics (e.g., a difference in sample proportions) is statistically significant?
2. A Gallup poll conducted in March 2014 reported that 54% of U.S. adults said the effects of global warming have already begun.[35] This poll was conducted on a random sample of 1,048 adults from across the U.S. The same poll was again conducted in March 2023 on 1,009 adults and found that 60% reported that the effects of global warming have already begun.[36]

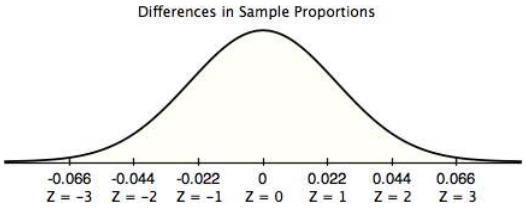
Using the sample statistics from March 2013 and March 2023, we can perform a hypothesis test to determine whether the observed difference in sample statistics indicates that the proportion of U.S. adults who worry a great deal or fair amount about global warming in March 2023 is greater than the proportion of U.S. adults in March 2014 who felt the same way. The null and alternative hypotheses are:

Null Hypothesis: Population proportion in March 2023 = Population proportion in March 2014.

Alternative Hypothesis: Population proportion in March 2023 > Population proportion in March 2014.

(a) Find the difference in the sample proportions: March 2023 sample proportion minus March 2014 sample proportion.

(b) Assuming the null hypothesis is true, we can model differences in sample proportions with a normal distribution. The distribution of all differences in sample proportions has a mean of 0 and a standard deviation of 0.022.



The observed difference in sample proportions is 0.06 (6 percentage points). Assuming the null hypothesis is true, find the P-value—the probability of randomly obtaining a difference in sample proportions of 0.06 or more. Interpret what this P-value means. (Hint: The first step is to convert the difference in sample proportions to a Z-score.)

(c) Is the difference in the sample proportions statistically significant? Explain.

(d) What can we decide about the null and alternative hypotheses?

(e) What can we decide about U.S. adults' level of worry about global warming in March 2023 compared to March 2014?

The Value of Academic Advising

4. A researcher wanted to know if college students' views on academic advising vary based on the amount of time that students are enrolled in college. She collected data from a random sample of 400 U.S. college students. The following two-way table summarizes the results of the survey. The results were analyzed based on the year of the student.

	How valuable is academic advising to your success in college?	
	Highly valuable	**Somewhat valuable**
1st year student	128	62
3rd year student	160	50

(a) Define the sample in this study and the population of interest.

(b) What proportion (percentage) of 1st year students in the sample believe that academic advising is highly valuable to success in college? Round to the nearest percent.

(c) What proportion (percentage) of 3rd year students in the sample believe that academic advising is highly valuable to their success in college? Round to the nearest percent.

(d) What is the difference in sample proportions (sample proportion of 3rd year students minus sample proportion of 1st year students)? Round to the nearest percent.

(e) A researcher used these sample statistics to perform a hypothesis test. The test examined whether the proportion of all 3rd year students who view academic advising as being highly valuable is greater than the proportion of all 1st year students who feel the same way. The null and alternative hypotheses are indicated below.

Null Hypothesis: Population proportion of all 3rd year students = Population proportion of all 1st year students.

Alternative Hypothesis: Population proportion of all 3rd year students > Population proportion of all 1st year students.

The hypothesis test returned a P-value of 0.025. Is this difference in sample proportions statistically significant?

(f) What should we decide about the null and alternative hypotheses?

(g) What can we decide about all 1st year and 3rd year students' views on academic advising? Explain.

---

[35] https://news.gallup.com/poll/167879/not-global-warming-serious-threat.aspx

[36] https://news.gallup.com/poll/474542/steady-six-say-global-warming-effects-begun.aspx

This page titled 3.8.2: Exercise S.8 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.9: Significance in Experiments

SPECIFIC OBJECTIVES

By the end of this collaboration, you should understand that:

- a difference in sample statistics is statistically significant if it leads us to reject the null hypothesis in a hypothesis test.
- a cause-and-effect relationship can be generalized to a population if the difference in sample statistics within an experiment is statistically significant.
- an experiment can establish a cause-and-effect relationship between explanatory variables and

response variables in a population.

By the end of this collaboration, you should be able to:

- compute sample statistics and the difference in sample statistics based on an experiment.
- use probability to assess whether a difference in sample statistics is statistically significant.
- interpret P-values from a hypothesis test.

PROBLEM SITUATION 1: EXAMINING THE EFFECT OF WEARABLE DEVICES ON ADULTS' LONG-TERM WEIGHT LOSS

Researchers from the University of Pittsburgh conducted a study to examine the impact of the use of fitness trackers on the weight loss of adults participating in a long-term weight loss program.[38] 470 adults aged 18–35 with a body mass index between 25–39 participated in the study. Participants were placed on a low-calorie diet, provided group counseling, and asked to increase their exercise activity over a 24-month period. Six months into the study, researchers randomly assigned participants into two groups: 233 participants were assigned to a standard intervention (SI) group, and 237 participants were assigned to an enhanced technology intervention (ETI) group. In the standard intervention group, participants engaged in self-monitoring of their diet and physical activity and entered data into a website. In the enhanced technology intervention group, participants were provided a wearable device and accompanying web interface which they used to monitor their diet and physical activity. After 24 months, researchers measured the weight loss of all participants. In the standard intervention group, the mean weight loss was 5.9 kg (13.0 pounds). In the enhanced technology intervention group, the mean weight loss was 3.5 kg (7.72 pounds).

1. What is the explanatory variable in this experiment? What values were assigned to this variable?
2. The experiment has two experimental groups. Define each of these groups. What is the purpose of randomly assigning participants into these two groups?
3. What is the response variable in this experiment?

There is an important type of variable, called a *confounding variable*, that influences how we must interpret causal relationships. A **confounding variable** is a variable that has an effect on the response variable, but was not properly controlled for in the experiment. For example, suppose a student living in a dorm wanted to do an experiment to see if offering a free t-shirt would get more people to come to his social activism club. He put up large signs on the third and fourth floors of his dorm advertising a free

t-shirt and free pizza for those who attended the next meeting. He did not put up any signs on the first and second floors. He found that more students from the third and fourth floors than the first and second floors attended the next meeting.

After talking to a statistics student, he realized he could not conclude that offering a free t-shirt increased attendance, because there were confounding variables in his experiment. Increased attendance could have been caused by the free pizza, or simply increased awareness of the meeting from the signs. The free pizza and the presence of the signs were both confounding variables in his experiment. The student could have controlled for these by only offering one incentive and by posting similar signs on the first and second floors that had the same information but did not offer a free t-shirt. These controls would have eliminated those confounding variables to isolate the effect of offering a free t-shirt.

4. Identify two controls used in this experiment. For each control, identify the corresponding confounding variable that was addressed

Exploring the Effect of Monitoring Approach on Weight Loss

The researchers performed a hypothesis test to examine the effect of integrating wearable technology into the weight loss program. The hypothesis test used the following hypotheses:

- Null Hypothesis: Mean weight loss of all adults who complete the standard intervention program = Mean weight loss of all adults who complete the enhanced technology intervention program.
- Alternative Hypothesis: Mean weight loss of all adults who complete the standard intervention program > Mean weight loss of all adults who complete the enhanced technology intervention program.

The null hypothesis states that standard intervention and enhanced technology intervention yield the same weight loss, on average. This hypothesis states that the enhanced technology intervention has *no effect* compared to the standard intervention. The alternative hypothesis states that the enhanced technology intervention has a negative impact on weight loss, compared to the standard intervention, for all participating adults.

5. Compute the difference in sample means: mean weight loss of adults in the standard intervention program minus mean weight loss of adults in the enhanced technology program.
6. The hypothesis test returned a P-value of 0.002. (**Note**: This P-value is found based on a T-distribution. Calculating this P-value is beyond the scope of this class, so it is provided.)

(a) Interpret this P-value.

(b) Recall that when a P-value is low, typically less than 5%, we say that a difference in sample statistics is statistically significant. Is the observed difference in sample means statistically significant?

(c) What can we decide about the null and alternative hypotheses?

(d) What decision can we make about these two weight loss programs?

PROBLEM SITUATION 2: DOES USING TECHNOLOGY HINDER STUDENT PERFORMANCE?

7. A mathematics professor at a large university was concerned that students' use of technology during class decreased their achievement in mathematics classes. She decided to conduct an experiment to examine whether college students' use of laptop computers with Internet access during class negatively impacted their performance on the final exam. The experiment was conducted in three College Algebra classes with a total of 86 students. The three classes were representative of all College Algebra classes at the university. The professor randomly assigned students to two groups: 44 students were provided a laptop computer with Internet access to use during class, and 42 students were not provided a laptop to use during class. The 44 students from the first group were allowed to use the laptop computers in class and were instructed to use them to support their learning of course material. The 42 students in the second group were not allowed to use computers during class. The experiment was conducted over an entire semester. At the end of the semester, all students took a cumulative final exam. The mean final exam scores for each of the two groups are displayed below.

Used Laptop Computer in Class	**Mean of Final Exam Scores**
No	82.1
Yes	79.4

(a) What is the explanatory variable and response variable in this experiment?

(b) The sample statistics in this study are sample means. What is the difference in sample means (that is, the sample mean of final exam scores of students who did not use laptops in class minus the sample mean of final exam scores of students who used laptops in class)?

(c) What does this difference in sample means suggest about the effect of laptop computer use in class on final exam grades for students who participated in the experiment?

(d) The professor performed a hypothesis test using the following hypotheses:

- Null Hypothesis: Mean final exam score of all students who don't use laptops in class = Mean final exam score of all students who do use laptops in class.
- Alternative Hypothesis: Mean final exam score of all students who don't use laptops in class > Mean final exam score of all students who do use laptops in class.

The null hypothesis states that the use of laptop computers with Internet access during class has *no effect* on final exam grades for all students. The alternative hypothesis states that the use of laptops has a negative effect, thus the mean exam score of all students

who don’t use laptops in class is greater than the mean exam score of all students who do.

The hypothesis test returned a P-value of 12%. **Interpret this P-value**.

5. Is the difference in sample means statistically significant?
6. What can we decide about the null and alternative hypotheses?
7. What decision can we make about the effect of using laptop computers with Internet access during class on final exam scores for all College Algebra students at this university?

MAKING CONNECTIONS

Record the important mathematical ideas from the discussion.

---

[38] http://jamanetwork.com/journals/jama/fullarticle/2553448

This page titled 3.9: Significance in Experiments is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

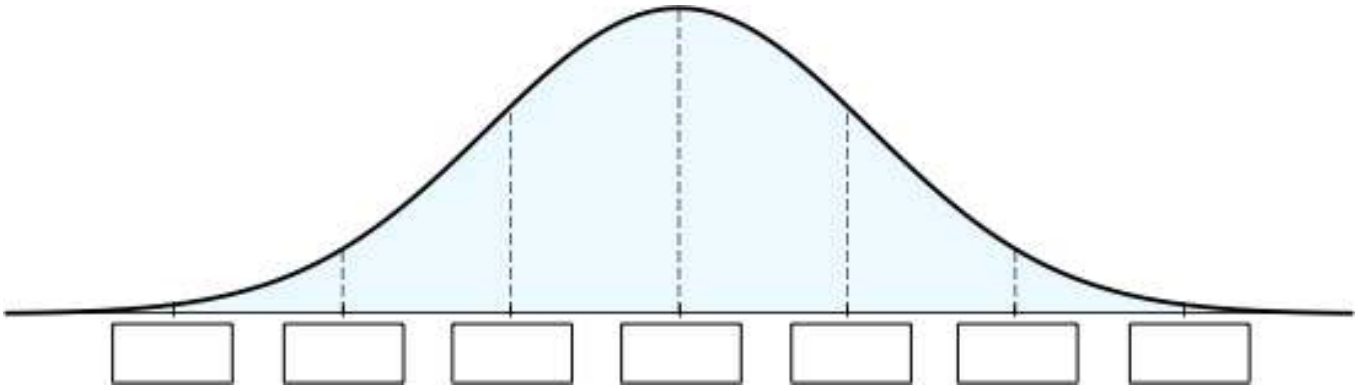
3.9.1: Preparation S.9

Statistical significance is a powerful concept in statistics. When an observed difference in sample statistics is statistically significant, the sample results enable us to reject the null hypothesis and support the alternative hypothesis. Collaboration S.8 introduced statistical significance based on sample results from observational studies. This lesson explores how sample results from experiments can be used to determine statistical significance and involves comparisons between sample means to make decisions about population means.

Exploring a Distribution of Sample Means

1. The heights of adult females in a large city are normally distributed with a mean of 65 inches and a standard deviation of 2.5 inches. Given these population parameters (population mean and population standard deviation), we can use a normal distribution to model sample means from random samples of the same size from this population. Doing so enables us to find probabilities associated with sample means, assuming the population mean is a specific value.

(a) Let's model sample means from random samples of size 100. Sample means from random samples of size 100, taken from this population, are normally distributed with a mean of 65 inches and a standard deviation of 0.25 inches. Represent this distribution of sample means in the normal curve below. Fill in the boxes below the x-axis by specifying the mean of the normal distribution and the positions that are one, two and three standard deviations from the mean. Write each value in its corresponding box.



(b) The normal curve above represents sample means from random samples of size 100. What interval contains the middle 95% of sample means? (Hint: Use the Empirical Rule.)

(c) What percentage of sample means are 64.8 inches or less? (Hint: Convert the sample mean of 64.8 to a Z-score. Use the Standard Normal Distribution table to get the area to the left of the Z-score.)

(d) Would it be likely or unlikely to obtain such a random sample of 100 adult females from this population whose mean height is 64.8 inches or less? Explain.

(e) What percentage of sample means are 65.6 inches or more? (Hint: Convert the sample mean of 65.6 to a Z-score. Use the Standard Normal Distribution table to get the area to the right of the Z-score.)

(f) Would it be likely or unlikely to obtain such a random sample of 100 adult females from this population whose mean height is 65.6 inches or more? Explain.

Hypothesis Test for a Population Mean

SAT mathematics scores of all New Hampshire high school graduates in 2021 have a mean of 526 and a standard deviation of 109.[37] An SAT preparation program in New Hampshire claims that the students who complete their program outperform this state average. Suppose we decide to perform a hypothesis test to assess whether this claim should be supported. We cannot obtain the SAT mathematics scores for all students who have completed the preparation program, but data are available on a random sample of 50 students who completed the program. We find that the sample has a mean score of 546 with a standard deviation of 78. Based on the sample results, can we decide that the mean score of all students who complete the program is greater than the state average of 526?

The mean score for the population of SAT preparation program completers is *unknown*. In order to assess whether students who complete the program outperform the state average, we will assume that the mean score for the program completers is equal to the

state average, and examine the likelihood of observing a sample mean as extreme as the one we found.

We will perform a hypothesis test using the following hypotheses:

Null Hypothesis: Mean SAT score of all students who completed the SAT preparation program = 526 (state average).

Alternative Hypothesis: Mean SAT score of all students who complete the SAT preparation program > 526 (state average).

2. Assuming the null hypothesis is true, the hypothesis test returns a P-value of 0.038. Thus, if the mean SAT score of all program completers is 526, there is a 3.8% chance of obtaining a random sample of 50 program completers whose mean score is 546 or more.

Note: since we don't know the population standard deviation of SAT scores for all students who completed the program, this P-value is found using a different type of distribution, called a T-distribution. The necessary calculations to find this P-value are beyond the scope of this class. For problem situations involving a T-distribution, the P-value will be provided.

(a) Is the observed sample mean statistically significant?

(b) What decision should we make about the null and alternative hypotheses?

(c) What decision should we make about the relationship between the mean score of those who complete the SAT preparation program and the mean SAT score for the high school graduates in the state?

After Preparation S.9 (survey)

You should be able to do the following things for the next collaboration. Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Before beginning Collaboration S.9, you should understand the concepts and demonstrate the skills listed below:

Skill or Concept: I can …	**Rating from 1 to 5**
identify the purpose of an experiment.	
identify key features of an experiment (explanatory and response variables, sample size, possible confounding variables).	
understand what it means for a sample statistic to be statistically significant.	

[37] https://nces.ed.gov/programs/digest/d21/tables/dt21_226.40.asp

This page titled 3.9.1: Preparation S.9 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

3.9.2: Exercise S.9

The Effect of Fertilizers on the Growth of Tomato Plants

1. A statistics student who enjoys gardening conducted an experiment to examine the effect of using a new eggshell fertilizer on the growth of tomato plants. The student randomly assigned sixteen Black Cherry tomato plants to two conditions: (a) eight plants were fertilized with a traditional organic fertilizer, and (b) eight plants were fertilized with an organic fertilizer based on eggshells. The student measured the growth (in millimeters) of each plant each week. The tomato plants were randomly planted throughout a large plot of land. The tomato plants were the same size when planted and were planted on the same day. The same amounts of fertilizers were used each week over a 4-week period.

The table below displays the average growth rate of each plant in the two fertilizer groups (in millimeters per week).

Traditional Fertilizer	**Eggshell Fertilizer**
4.9	5.9
5.2	5.1
4.8	5.5
4.7	5.9
4.9	4.8
4.8	4.9
4.6	5.7
5.9	5.6

(a) The experiment was conducted on two samples. Define the samples and the corresponding populations of interest.

(b) What is the explanatory variable in this experiment? What two conditions were assigned to the explanatory variable?

(c) What is the response variable in this experiment?

(d) Experimental controls are implemented to eliminate confounding variables and ensure that changes to a response variable are influenced solely by changes to the explanatory variable. Identify an experimental control and state the corresponding confounding variable that is being addressed.

(e) Compute the sample mean and sample standard deviation for each group of tomato plants. Round the values to two decimal places. Compute the difference in sample means: mean of eggshell fertilizer sample minus mean of traditional organic fertilizer sample.

(f) The student performed a hypothesis test using the following hypotheses:

Null Hypothesis: Mean growth rate of plants with eggshell fertilizer = Mean growth rate of plants with traditional organic fertilizer.

Alternative Hypothesis: Mean growth rate of plants with eggshell fertilizer > Mean growth rate of plants with traditional organic fertilizer.

The null hypothesis states that the type of organic fertilizer used has no effect on tomato plant growth. The alternative hypothesis states that the eggshell fertilizer increases tomato plant growth compared to the traditional organic fertilizer. The hypothesis test returned a P-value of

0.026. Interpret this P-value.

(g) What can we decide about the null and alternative hypotheses?

(h) What can we decide about the two fertilizers?

This page titled 3.9.2: Exercise S.9 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

4: Corequisite Module N

4.1: Corequisite- Estimation Vs Calculation

4.2: Corequisite- Conversion, Estimation, Percentages and Ratios, Scientific Notation

4.2.1: Practice Problems Corequisite N.2

4.3: Corequisite- Interest, Doubling Time

4.3.1: Practice Problems Corequisite N.3

4.4: Corequisite- Percentages and Ratios, Absolute and Relative Change, Rates

4.4.1: Practice Problems Corequisite N.4

4.5: Corequisite- Percentages and Taxes, Ratios, Proportions and Dimensional Analysis

4.5.1: Practice Problems Corequisite N.5

4.6: Corequisite- Dimensional Analysis, Unit Conversion

4.6.1: Practice Problems Corequisite N.6

4.7: Corequisite- Buying Power, Consumer Price Index, Absolute and Relative Change, Reading and Interpreting Graphs

4.7.1: Practice Problems Corequisite N.7

4.8: Corequisite- Percentages, Division, Extracting Relevant Information

4.8.1: Practice Problems Corequisite N.8 and N.9

This page titled 4: Corequisite Module N is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.1: Corequisite- Estimation Vs Calculation

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- you are a necessary and integral part of the classroom community.
- you should feel comfortable asking questions and participating in classroom discourse, no matter your language background.
- quantitative reasoning is the ability to understand and use quantitative information. It is a powerful tool in making sense of the world.
- relatively simple math can help make sense of complex situations.
- 1 billion = 1,000 × 1,000 × 1,000.
- the representations 1 billion, 1,000,000,000, and 10^9 have the same meaning.

By the end of this lesson, you should be able to

- work effectively in groups.
- understand the course expectations and what is necessary to succeed in the course.
- identify quantitative information.
- convert units from feet to miles.
- round numbers.
- name large numbers.
- work in groups and participate in discussion using the group norms for the class.

ESTIMATION

In this course, you will talk about different types of **estimation**.

- **Educated guess**: One type of estimation might be called an educated guess about something that has not been measured exactly. You could estimate the world population. This quantity cannot be measured exactly—it would be impossible to count how many people live on the earth at any given time. Scientists can use good data and mathematical techniques to estimate the population, but it will always be an estimate.
- **Convenient estimation**: Sometimes estimations are used when it is inconvenient or not worthwhile to make an exact count. Imagine that you need to know how much paint to buy to paint the baseboard trim in your house. (The baseboard trim is the piece of wood that follows along the bottom of the walls.) You need to know the length of the baseboard. You could measure the length of each wall to the nearest 1/8 inch and carefully subtract the width of halls and doors. It would be much quicker and just as effective to measure to the nearest foot or half foot. If you were cutting a piece of baseboard to go along the floor, however, you would want an exact measurement.
- **Estimated calculation**: This usually involves rounding numbers to make calculations simpler. You will find in this course that percentages are used in many contexts. One of the most important skills you will develop is understanding and being comfortable working with percentages in a variety of situations.

PROBLEM SITUATION: DOES THIS INFORMATION MAKE SENSE?

During this course, you will be presented with a number of problem situations. These problem situations will help you learn how to evaluate the types of quantitative information you may encounter in everyday life.

This problem situation asks you to use quantitative information to figure out if a statement makes sense.

Imagine that you just received a flier in the mail with the following statement:[1]

> Every year since 1980, the number of American children gunned down has doubled.

(1) What groups might have wanted to mail a flier like this? What are some social issues or political ideas that this statement might support?

Is the Information Reasonable?

The flier above includes **quantitative information**. Quantitative information uses concepts about quantity or number (this can be specific numbers or a pattern based on numerical relationships such as doubling).

You hear and see statements that include quantitative information every day. People use these statements as evidence to convince you to do things like

- vote a certain way;
- donate or give money to a cause; or
- understand a health risk.

You often do not know whether these statements are true. You may not be able to locate the information that supports these statements, but you can start by asking if the statements are **reasonable**. This means asking if the statements make sense. You will be asked if information is reasonable throughout this course. This lesson will help you to understand what is meant by this question.

(2) Do you think the statement, “Every year since 1980, the number of children gunned down has doubled” is a reasonable statement? Discuss with your group and write your reasoning below.

(3) Using only the information that was included in the flier, how can you decide if the statement was reasonable? Talk with your group about the different ways in which you might answer this question. Record some strategies below.

(4) In Question 3, you thought about how to decide if the statement below was reasonable.

"Every year since 1980, the number of American children gunned down has doubled."

For this problem, choose a starting number for gun deaths in 1980. Put this number into the table below in cell (a). Work in your group to complete the other values in the second column of the table (b) – (f). Enter the numbers in words in (g) – (l).

For example, if “1000” is in column 2, “one thousand” would be entered in the corresponding space in column 3.

Year	Number of Children Gunned Down	Rounded (using words)
1980	*(a)*	*(g)*
1990	*(b)*	*(h)*
2000	*(c)*	*(i)*
2010	*(d)*	*(j)*
2015	*(e)*	*(k)*
2020	*(f)*	*(l)*

(5) Refer to the table you created in Question 4. Does the number of children gunned down in the year 2020 seem reasonable? What kind of information might help you decide?

About This Course

This course is called a *quantitative reasoning* course. This means you will learn to use and understand quantitative information. It will be different from many other math classes you have taken. You will learn and use mathematical skills connected to situations like the one you just discussed in this lesson. You will talk, read, and write about quantitative information.

The lessons in this course will focus on three themes.

- **Social and Environmental**: You will learn how to understand information about your society, government, and world that can impact many decisions you make.
- **Financial**: You will study how to understand financial information and how to use it to make decisions in your life.
- **Health and Risk**: You will learn how to understand information about health issues and medical treatments.

This lesson is part of the “Social” theme. In this lesson, you learned about ways to decide if information is reasonable. This can help you form an opinion about an issue.

Today, the goal was to introduce you to the idea of quantitative reasoning. This will help you understand what to expect from the class. Do not worry if you did not understand all of the math concepts. You will have more time to work with these ideas throughout the course.

In this course you will learn to do the following things:

- Understand and interpret quantitative information.
- Evaluate quantitative information. Today you did this when you answered if the statement was reasonable.
- Use quantitative information to make decisions.

---

[1] Adapted from Joel Best, Damned Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists (Berkeley: University of California Press, 2001).

This page titled 4.1: Corequisite- Estimation Vs Calculation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.2: Corequisite- Conversion, Estimation, Percentages and Ratios, Scientific Notation

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- the magnitude of large numbers is seen in place value and in scientific notation.
- proportions are one way to compare numbers of varying magnitudes.
- different comparisons may be needed to accurately compare two or more quantities.
- estimation is a valuable skill.
- standard benchmarks can be used in estimation.
- there are many strategies for estimating.
- percentages are an important quantitative concept.

By the end of this lesson, you should be able to

- express numbers in scientific notation.
- use multiple computations to compare quantities.
- convert between millions, billions, and trillions.
- compute basic percentages and ratios.
- express one number as a percentage of another.

PROBLEM SITUATION: COMPARING POPULATIONS

In this problem situation, you are going to compare the populations of China and the United States, and you will compare the population of each country to the population of the world. Since the population of each of these countries is very large, you will use scientific notation to write your answers. Recall that **scientific notation** is a special way of writing very large or very small numbers. A number in scientific notation is written in the form: $a \times 10^n$ where $1 \leq a < 10$; and n is an *integer*. An **integer**, commonly known as a whole number, is a number that can be written without a fractional component. Below are two examples of very large numbers written in scientific notation.

Examples:

- In 2020, the total population of Asia was approximately 4.7 billion people.[2] You can write this as 4,700,000,000 or you can use scientific notation to write it as 4.7×109 people.
- 28,930,000 can be written in scientific notation as 2.893×107.

(1) In late-2022, the population of the United States was 333,000,000.[3] In late-2022, Earth's population was about 7.9 billion. Write these numbers in scientific notation.

(a) U.S. population in scientific notation:

(b) World population in scientific notation:

(2) What are some other ways you could compare the population of the United States to the population of the earth *besides* writing the two numbers in scientific notation? Think about forms of comparisons using both estimation and calculation.

(3) A **ratio** is an expression that compares how much there is of one thing to another. Ratios are often written as fractions. In mid-2022, the population of China was 1.412 billion.[4] Use a ratio to compare the population of China in 2022 to the population of Earth in 2022 (7.9 billion). Write your answer as a fraction.

(4) Compare China's population with the population of the United States using a ratio (with the U.S. population as the reference value). Write a sentence that interprets this ratio in the given context.

---

[2] https://www.prb.org/international/geography/asia/

[3] http://www.census.gov/popclock/

[4] https://www.prb.org/international/geography/china/

This page titled 4.2: Corequisite- Conversion, Estimation, Percentages and Ratios, Scientific Notation is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.2.1: Practice Problems Corequisite N.2

(1) Which of the following statements is true? How do you know if you are correct?

(i) A trillion is 10^{10}.

(ii) A trillion is 10 billion.

(iii) A trillion is 100 billion.

(iv) A trillion is 1,000 billion.

(2) **Equivalent** means that numerical expressions are equal to each other. Complete the following table of equivalent percentages, fractions, and decimals. The first row in the table is done for you. Recall that the first numeral to the right of the decimal point is referred to as the "tenths" place while the second numeral to the right of the decimal point is the "hundredths" place.

Simplified Fraction	**Decimal**	**Percent**
$\frac{1}{100}$	0.01	1%
$\frac{1}{10}$		
	0.2	
		25%
$\frac{1}{3}$	round to nearest hundredth	round to the nearest one percent
	0.5	
$\frac{2}{3}$	round to nearest hundredth	round to the nearest one percent
	0.75	

Estimate the following percentages without using a calculator. Explain your estimation strategy.

(3) 10.1% of 7,800

(4) 0.99% of 83,583

(5) 20% of 5,008,340

(6) Which of the following is the percent estimate of 1/3, rounded to the nearest hundredth of a percent? Explain and/or show your calculation work.

(i) 0.33%

(ii) 3.3%

(iii) 33.3%

(iv) 33.33%

(7) In industrialized countries, such as the United States, up to 30% of the population suffers from foodborne diseases each year. This means that 30 out of _____ people living in industrialized countries will likely suffer from foodborne diseases each year.

In 1994, an outbreak of illness due to ice cream contaminated with the bacteria salmonella occurred in the U.S. The outbreak affected an estimated 224,000 people.

(8) If the total population in the United States at that time was 260,000,000, which is the best estimate for the percentage of people who were affected? Show your calculation work.

(i) About 0.1%

(ii) About 1%

(iii) About 5%

(iv) About 10%

(v) About 25%

(9) Complete the following statement: _____ out of every 100,000 people in the United States were affected by the 1994 salmonella outbreak. Round to the nearest whole person.

Corequisite lessons are also available online. The **Practice Problems** section of online Corequisite lessons contains similar questions to what you see in your workbook, but the numbers and figures vary. Completing the **Practice Problems** section online will generate a score.

This page titled 4.2.1: Practice Problems Corequisite N.2 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.3: Corequisite- Interest, Doubling Time

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- population growth can be measured in terms of doubling time.
- doubling times can be used to compare population growth during different periods.
- quantitative reasoning and math skills can be applied in various contexts.
- creditworthiness affects credit card interest rates and the amount paid by the consumer.
- reading quantitative information requires filtering out unimportant information (introductory level).

By the end of this lesson, you should be able to

- calculate quantities in the billions.
- use data to estimate a doubling time.
- compare and contrast population growth via population doubling times.
- recognize common mathematical concepts used in different contexts.
- identify a complete response to a prompt that asks for connections between mathematical concepts and a context.
- write a formula in a spreadsheet.
- convert percentages to decimals
- decide when simple interest is appropriate and how to calculate it
- use and interpret the terms “millions,” “billions,” and “trillions.”

PROBLEM SITUATION 1: MEASURING POPULATION GROWTH

In this lesson, you will look at doubling times to determine how the human population of the earth has changed over time. The **doubling time** of a population is the amount of time it takes a population to double in size. Calculating doubling time helps you understand how fast a population is growing. Comparing doubling times helps you understand how growth is changing over time.

The table below gives historical estimates of the human population. The population in 8,000 BCE was estimated to be 5 million people. Two thousand years later, in 6,000 BCE, the population had doubled to 10 million people. Therefore, the population doubling time for 8,000 BCE is about 2,000 years.

Population Estimates Throughout History[5]

Year	World Population (lower bound, in millions)		Year	World Population (lower bound, in millions)
10,000 BCE	1		1850	1,262
9,000 BCE	3		1900	1,650
8,000 BCE	5		1950	2,519
7,000 BCE	7		1960	2,982
6,000 BCE	10		1970	3,692
5,000 BCE	15		1980	4,435
4,000 BCE	20		1985	4,831
3,000 BCE	25		1990	5,263
2,000 BCE	35		1995	5,674
1,000 BCE	50		2000	6,070
500 BCE	100		2005	6,454
1 CE	200		2010	6,922

1000 CE	310		2015	7,339
1800 CE	978		2020	7,753

(1) Using the information in the table, estimate the doubling times of Earth's human population. Enter your answers in the table below. Start with the year given below and estimate how long it took for the population in that year to double. The first entry is done for you. Be prepared to explain how you arrived at your answers.

Year	**Doubling Time**	**Year**	**Doubling Time**
8,000 BCE	2,000 years	1800 CE	(d)
6,000 BCE	(a)	1850 CE	(e)
3,000 BCE	(b)	1900 CE	(f)
1 CE	(c)	1970 CE	(g)

(2) What do you notice about the doubling times? What does this tell you about how the human population has changed over time?

A Writing Tip

One of the skills you will learn in this course is how to write quantitative information. A writing principle that you will use throughout the course is given below and followed by Question 3, which offers examples of how to use this principle.

Writing Principle: Use specific and complete information. Readers should understand what you are trying to say even if they have not read the question or writing prompt. This includes

- information about context, and
- quantitative information.

(3) Which of the following statements best describes the change in doubling times before 1800 CE?

(a) The doubling times decreased.

(b) Before 1800 CE, estimated population doubling times decreased from 2,000 to 1,000.

(c) The doubling times decreased from 2,000 to 1,000.

(4) Write a statement that describes the change in doubling times after 1800 CE.

PROBLEM SITUATION 2: UNDERSTANDING CREDIT CARDS

When you use a credit card, you can pay off the amount you charge each month. If you do not pay the full amount, you are borrowing money from the credit card company. This is called **credit card debt**. Many people in the United States are concerned about the amount of credit card debt both for individuals and for society in general. In this lesson, you will use skills and ideas from previous lessons to think about some issues related to credit cards.

(5) The statement below came from a website that reported statistics for credit card debt in 2022:

- "American household debt hit a record $14.6 trillion in the spring of 2021, according to the Federal Reserve."[6]

Does the statement "US households had $14,600 billion in debt in 2021," express the same amount of debt? Justify your answer with an explanation.

Understanding the Fine Print

When you apply for a new credit card, you usually receive a disclosure in the mail along with the new credit card. The information below is excerpted from the disclosure that came with a new credit card. Use this information to answer Questions 2 and 3.

Annual Percentage Rate (APR) for Purchases	**0.00%** introductory APR for 6 months from the date of account opening. After that, your APR will be **10.99%** to **23.99%** based on your creditworthiness. This APR will vary with the market based on the Prime Rate.

A **credit score** is a number assigned to a person that indicates to lenders his or her ability to repay a loan. A high credit score indicates good credit. In the following questions, you will explore how your credit score can affect how much you have to pay in interest for the purchases you make with your credit card.

(6) Juanita and Brian both have a credit card with the terms in the above disclosure form. They have both had their credit cards for more than six months.

(a) Juanita has good credit and gets the lowest interest rate possible for this card. She is not able to pay off her balance each month, so she pays interest. **Estimate** how much interest Juanita would pay in a year if she maintained an average balance of $5,000 each month on her card. Juanita is past the 6 month 0.0% introductory APR rate. Explain your estimation strategy.

(b) Brian has a very low credit score and has to pay the highest interest rate. He is not able to pay off his balance each month, so he pays interest. **Calculate** how much interest he would pay in a year if he maintained an average balance of $5,000 each month. Show your calculation. Round your answer to two decimal places.

(c) What are some things that might affect your credit score?

(7) The **Annual Percentage Rate (APR)** of a credit card is the interest rate for a full year. The **periodic rate** of a credit card is the interest rate for a month. You can find the periodic rate by dividing the APR by the number of months in a year.

(a) What is the periodic rate for Juanita's card? Round to two decimal places.

(b) Juanita has a balance of $1,082 on her January credit card statement. The statement says the minimum payment required in January is $100. Juanita makes only the minimum payment, so the rest of the balance carries over to February. Which of the following is the best estimate of how much interest she will owe in February?

(i) Less than a dollar

(ii) $5-$10

(iii) $10-$20

(iv) More than $20

Credit Cards and Cash Advances

You will use the following information from the credit card disclosure for Question 8. A cash advance is when you use your credit card to get cash instead of using it to make a purchase.

Annual Percentage Rate (APR) for Purchases	After that, your APR will be **10.99%** to **23.99%** based on your creditworthiness. This APR will vary with the market based on the Prime Rate.
APR for Cash Advances	**28.99%**. This APR will vary with the market based on the Prime Rate.

(8) Discuss each statement below and decide if it is reasonable.

(a) Jeff pays the highest interest rate for purchases he makes with his credit card. For a cash advance, he would pay $0.05 more for each dollar he receives in cash than he pays for purchases he makes with his credit card.

(b) The APR for cash advances is about two-and-a-half times as much as the lowest APR for purchases.

(9) Brian used a spreadsheet to record his credit card charges for a month.

	A	B	C
1	Credit Card Charges for the Month		
2	Gas	$56.08	
3	Groceries	$36.72	
4	Restaurant	$12.82	
5	Movie tickets	$16.00	
6			
7	Interest paid		
8			

Brian used the following expression to calculate his interest for that month's charges.

$$\frac{0.2399}{12}(B2+B3+B4+B5)$$

Which of the following statements best explains what the expression means in terms of the context?

(i) Brian added his individual charges. Then he divided 0.2399 by 12. Then he multiplied the two numbers.

(ii) Brian found the interest charge for the month by dividing 0.2399 by 12 and multiplying it by the sum of Column B.

(iii) Brian added the individual charges to get the total amount charged to the credit card. He found the periodic rate by dividing the APR by 12 months and multiplied the rate by the total charges. This gave the interest charge for the month.

---

[5] https://www.census.gov/population/international/data/worldpop/table_history.php

[6] https://www.debt.org/faqs/americans-in-debt/demographics/

This page titled 4.3: Corequisite- Interest, Doubling Time is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.3.1: Practice Problems Corequisite N.3

The purpose of this section is for you to practice with the skills and concepts from the lesson. You will see questions directly related to the lesson. You will also see questions that apply the skills and concepts to different situations. The section will sometimes have reading material that helps explain the topics from the lesson. Later in the course, you can look back at this information as you review what you have learned.

Note that the online version of these practice problems can contain variabilized figures and values.

Question 1 highlights important quantitative reasoning skills that you will learn in this course.

Quantitative Reasoning Skill: Reading and interpreting quantitative information

The two statements shown below are about children "gunned down" in America.[7] Read them carefully and decide what each means mathematically.

> Original Statement: The number of American children killed each year by guns has doubled since 1980.

> Reworded Statement (circa 2020): Every year since 1980, the number of American children gunned down has doubled.

(1) Based on the original statement and the reworded statement, which of the following comments is valid?

(i) Assume that the **original** statement is true. If approximately 100 children were killed by guns in 1980, the number of children killed by guns in 2020 was about 200.

(iii) Assume that the **reworded** statement is true. If approximately 100 children were killed by guns in 1980, the number of children killed by guns in 2020 was about 200.

This highlights the importance of reading and writing carefully about quantitative information. The original and reworded statements look very similar, but mean entirely different things:

- The original statement says that the number has doubled once from 1980 to the published date (2020).
- The reworded statement says that the number has doubled every year between 1980 and the published date (2020).

There are many types of **investments**. Putting money into a savings account or an individual retirement account are two examples of investments.

Some types of investments earn **interest** based on a percentage rate. People often estimate the doubling time of investments to predict how much money the investment will be worth in the future. An investment that earns 5% interest will double in value about every 14 years. Use this information to complete the missing values in the table below for $7,500 invested at 5% interest. Be sure to show your calculation work.

Complete the table by answering **Questions 2-4**.

Year	**Value of Investment**
2020	$7,500
	$15,000
	$30,000
2062	

(2) In what year would the value of a $7,500 investment at 5% interest be $15,000?

(3) In what year would the value of a $7,500 investment at 5% interest be $30,000?

(4) What would the value of a $7,500 investment at 5% interest be in the year 2062?

(5) Which of the following is the best estimate for the amount of time it would take the investment in the table above to reach a hundred thousand dollars? Be sure to explain your answer.

(i) Less than 40 years

(ii) Between 40 and 60 years

(iii) Between 60 and 80 years

(iv) More than 80 years

(6) Some checking accounts pay a small amount of interest on the money in the account. In this case, **interest** is money that is paid to the account holder by the financial institution issuing the checking account. The interest is a percentage of the amount of money in the account. The percentage is called the **annual interest rate**. Compare the following two offers.

- Bank of Avalon pays 0.8% with no annual fee.
- Cypress Savings pays 1.5%, but charges a $10 annual fee.

Which bank would be the better offer if you have $1,000 in an account for one year?

Use the Writing Principle with your calculations to defend your answer.

(i) Bank of Avalon

(ii) Cypress Savings

Corequisite lessons are also available online. The **Practice Problems** section of online Corequisite lessons contains similar questions to what you see in your workbook, but the numbers and figures vary, so they are good for extra practice. Completing the **Practice Problems** section online will generate a score.
Tip: You can use the print version for practice. Then you can do the **Practice Problems** section for additional practice and a score.

---

[7] Best, J. (2001). Damned lies and statistics. University of California Press: Berkeley and Los Angeles.

This page titled 4.3.1: Practice Problems Corequisite N.3 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.4: Corequisite- Percentages and Ratios, Absolute and Relative Change, Rates

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- population density is a ratio of the number of people per unit area.
- proportionality can be used to compare population densities.
- the concept of proportional representation in voting.
- how representation in the U.S. Congress is allocated.
- a relative change is different from an absolute change.
- a relative measure is always a comparison of two numbers.

By the end of this lesson, you should be able to

- calculate population densities.
- determine if two population density ratios are proportional to each other.
- compare and contrast populations via their densities.
- calculate a unit rate.
- solve a proportion by first finding a unit rate and then multiplying appropriately.
- calculate a relative change.
- explain the difference between relative change and absolute change.

PROBLEM SITUATION 1: HOW THE CENSUS AFFECTS THE HOUSE OF REPRESENTATIVES

Every 10 years, the United States conducts a census. The census tells us how many people live in each state. You can also find out how much populations have changed over time from the census data. The original purpose of the census was to decide on the number of representatives each state would have in the House of Representatives. Census data continue to be used for this purpose, but now there are many other uses for the data. For example, governments may use the data to plan for public services such as fire stations and schools.

In this problem situation, your instructor will ask your team to explore data from one of two census regions: 1) South Atlantic States or 2) Mountain States. There are tables on the next page that show the population of each state in these census regions in 2010 and 2020. You will be asked to calculate the population growth (in people) as a percentage for each state in your assigned region, and for the region as a whole. You will examine how this affects the number of representatives each state has in the House of Representatives.

The **absolute change** in a state's population tells us by how many people the population has changed. The **relative change** is the change as it compares to the earlier population. Often relative change is given as a percentage. See the table below for an example using the population change in California from 2010 and 2020.

	2020 Population (in millions)	**2010 Population (in millions)**	**Absolute Change (in millions)**	**Relative Change (in percent form)**
California	39.58	37.27	39.58-37.27=2.31	2.3137.27*100%=6.2%

Use the following data to answer Questions 1–6 for your assigned census region.

South Atlantic States

	2020 Population	**2010 Population**
Florida	21,570,527	18,801,310
Georgia	10,725,274	9,687,653
North Carolina	10,453,948	9,535,483

Mountain States

	2020 Population	2010 Population
Arizona	7,158,923	6,392,017
Colorado	5,782,171	5,029,196
Nevada	3,108,462	2,700,551

(1) Match the states below with the absolute change in population.

• Colorado • Florida • Georgia • Nevada • North Carolina • Arizona	• 2,769,217 • 766,906 • 752,975 • 1,037,621 • 407,911 • 918,465

(2) Match the states below with the relative change (in percentage form) in the population.

• Colorado • Florida • Georgia • Nevada • North Carolina • Arizona	• 9.63% • 15.10% • 14.97% • 12.00% • 10.71% • 14.73%

(3) Which of the three states from each region changed the most in absolute population?

Mountain states:

South Atlantic states:

(4) Which of the three states from each region changed the largest relative increase in population?

Mountain Region states:

South Atlantic states:

(5) Explain why the lists in Question 3 and Question 4 are not the same.

(6) Use the previous information to calculate the following absolute and relative changes in total population.

(a) The absolute change in total population for the three Mountain Region states is

_____ people.

(b) The relative change in total population for the three Mountain Region states is _____ %.

(c) The absolute change in total population for the three South Atlantic Region states is

_____ people.

(d) The relative change in total population for the three South Atlantic Region states is

_____ %.

(7) Let's now consider a state we have not explored yet: Illinois. Interestingly, the population of the state of Illinois actually decreased between 2010 and 2020. Illinois had a population of 12,864,380 in 2010. Its population was 12,822,739 in 2020.

(a) What was the absolute change in Illinois' population?

(b) What was the relative change in Illinois' population? Write a statement that describes the relative change in Illinois' population between 2010 and 2020. Think about this on your own first before sharing with your group. Round your answer to the nearest hundredth of a percent.

PROBLEM SITUATION 2: USING RATIOS TO MEASURE POPULATION DENSITY

Earth's human population has grown from about 1 billion people to about 7 billion in the last two centuries. However, populations in different regions do not always grow uniformly. For example, populations tend to increase in areas where people already live close enough to one another to find mates. On the other hand, crowded populations decrease when deadly diseases, such as smallpox or Ebola virus, sweep through them. In this lesson, you will compare geographic regions by their **population densities**. The population density of a geographic region is a ratio of the number of people living in that region to the area of the region. Population density ratios are typically reported in a way that makes it easy to compare the ratio to a standard area measurement, such as one square mile or one square kilometer.

Example: Imagine 200 people standing in an open field that measures one mile by one mile (also known as one square mile). The population density of the open field could be thought of as 200 people per one square mile. This can be expressed as a ratio. The expression below shows the population density of the open field as a ratio in fraction form.

$$\frac{200\ people}{1\ square\ mile}$$

(8) A **proportion** is the equality of two ratios. Quantities are proportional if they satisfy a proportion. Picture the open field from the example that has a population of 200 people per one square mile. Then, picture a second open field. The length of each side of the second open field is double the length of each side of the first open field (making it 2 miles by 2 miles). The amount of people standing in the second open field is double the amount of people standing in the first open field. Is the population density of the second field equal to the population density of the first field? Explain.

(9) An island in the archipelago nation known as the Maldives is threatened by global warming. The island, called the Ari Atoll, is approximately one square mile in size, and has a population of 1,260 people. What is the population density per square mile?

(10) Imagine that global warming caused flooding in Ari Atoll, so the island is now half of its former size. However, the same number of people still live in Ari Atoll.

(a) Before doing any calculations, decide if you think the population density of Ari Atoll increased, decreased, or stayed the same after the flooding. Explain your answer.

(b) Calculate the population density of Ari Atoll after the flood.

(11) Suppose that since everyone is now living closer together, the birth rate on Ari Atoll increases rapidly. Soon the population has quadrupled. What is the population density now?

(12) Picture the open field from the example that contains 200 people and measures one square mile. Now picture another open field that measures five miles by five miles and has exactly the same population density as the first open field.

(a) How many people are standing in this field?

(b) Write a mathematical expression that shows that the population density of the first open field and the population density of the second open field are equal to each other.

(13) Consider a tiny island in the Atlantic Ocean. There are 50 survivors shipwrecked there. When they first arrived on the island, they all set up camp near the waterfall in the middle of the island. About six months later, a professor who was among the survivors claims the island can only support a population density of one person per acre (an acre is 43,560 square feet, which is a little less than the size of an American football field).

If the survivors then decide to space their campsites out evenly across the island, how will this affect the overall population density of the island?

PROBLEM SITUATION 3: VOTING FOR THE HOUSE OF REPRESENTATIVES

The U.S. Congress consists of two bodies, the Senate and the House of Representatives. For legislation to be enacted, it must pass in both of these bodies. The House of Representatives consists of a fixed number of 435 voting representatives. These representatives are allocated to each state **proportionally**, based on population. (Recall that a **proportion** is the equality of two ratios. Quantities are proportional if they satisfy a proportion.) In other words, this means that the higher a state's population, the more representatives it will have. The aim of this method of allocation is for U.S. voters to have relatively equal representation in the House. (**Note**: Each state must have at least one representative.)

(14) Representatives were last allocated in 2020, when the population of the U.S. was approximately 330 million. Find the average number of people represented by a single U.S. representative in 2020. Write your answer as a ratio in fraction form.

(15) Determine how many representatives should be allocated to the following states:

Ohio (2020 population 11,799,000):

Wyoming (577,000):

California (39,538,000):

(16) The U.S. population in 2020 was approximately 330,000,000. Based on current trends, the population of the U.S. is expected to grow to 352,427,000 in 2030 when the next U.S. Census is conducted (and representatives are reallocated). However, the population of Ohio is expected to remain approximately the same. Will Ohio keep the same number of representatives? If not, how many will they now have?

(17) All members of the U.S. House of Representatives will be up for reelection in 2026. The population of the U.S. is expected to be approximately 340 million in 2027 when the congressional session begins.

Find, as a unit rate, the average number of people who will be represented by a single U.S. representative in 2027.

This page titled 4.4: Corequisite- Percentages and Ratios, Absolute and Relative Change, Rates is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.4.1: Practice Problems Corequisite N.4

The following headlines all refer to change. Identify the change as absolute or relative.

(1) “Enrollments at Northeastern University are expected to increase by 1,000!”

(i) Absolute change

(ii) Relative change

(2) “Another $3,000 tuition increase is expected.”

(i) Absolute change

(ii) Relative change

(3) “A new proposal has sales tax rates dropping from 5% to 1%, a drop of only 4%.”

(i) Absolute change

(ii) Relative change

(4) “A new proposal has sales tax rates dropping from 3% to 1%, a 67% drop!”

(i) Absolute change

(ii) Relative change

Let's say the following tax cut was applied to U.S. taxpayers who earned up to $50,000 per year. On average, these taxpayers retained 4% more of their income than before the cuts. The tax cut also gave these taxpayers a separate $900 credit on their federal tax return.

(5) How much extra money did a typical person earning $50,000 receive due to the federal tax cuts?

(6) This extra money represents what percent of the typical person’s salary? Round to the nearest tenth of a percent.

Let's say the following tax cut was applied to U.S. taxpayers earning $450,000 per year. They were taxed at a lower rate on the first $97,800 they earned, saving them 4% of this portion of their salary. The same taxpayers also received $12,750 more on their federal tax return. Fill in the blank to complete the statement below.

(7) A person earning $450,000 a year received $_____ more of their income. Round to the nearest dollar.

(8) A person earning $450,000 a year received $16,662 more or _____ % more of their income. Round to the nearest tenth of a percent.

The tables below give the dimensions of four different rectangles sets.

(9) Are the four rectangles in this set “Proportional” or “Not Proportional”?

	Width	**Length**
Rectangle 1	17	110.5
Rectangle 2	23.4	152.1
Rectangle 3	33	214.5
Rectangle 4	52.2	339.3

(10) Are the four rectangles in this set “Proportional” or “Not Proportional”?

	Width	**Length**
Rectangle 1	7.4	15.5
Rectangle 2	17	40.8
Rectangle 3	23.4	68.5

Rectangle 4	33	72.6

Let's say the United States used 22,691,000 barrels of oil a day, and worldwide, people used around 42,799,000 barrels of oil per day. This includes oil used for (among other things) fuel and manufacturing.

(11) If there were 330 million people in the United States in 2020, what was the daily consumption rate per capita (barrels per person) in the United States? Round to the nearest hundredth of a barrel.

(12) If there were 7.82 billion people in the world in 2020, which of the following statements would be correct? The U.S. rate of oil consumption per capita was about:

(i) 2 times the world rate.

(ii) 5 times the world rate.

(iii) 10 times the world rate.

(iv) 50 times the world rate.

(13) There are 42 gallons in a barrel of oil. Which of the following statements is true? The average American is responsible for about:

(i) 0.5 gallons of oil use per day.

(ii) 3 gallons of oil use per day.

(iii) 5 gallons of oil use per day.

(iv) 30 gallons of oil use per day.

This page titled 4.4.1: Practice Problems Corequisite N.4 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.5: Corequisite- Percentages and Taxes, Ratios, Proportions and Dimensional Analysis

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- units can be used in dimensional analysis to set up calculations.
- precision should be based on several factors, including the size of the numbers used and the precision of the original values. Rounding can produce large differences in results.

By the end of this lesson, you should be able to

- solve a complex problem with multiple pieces of information and steps.
- use dimensional analysis.
- investigate how changing certain values affects the result of a calculation.

PROBLEM SITUATION 1: FICA TAXES

The United States government requires that businesses pay into two national insurance programs—Social Security and Medicare—which help senior citizens pay for their personal **expenses** (things that a person or people spend money on) and health care. Businesses take money out of their employees' paychecks in order to pay the government. In other words, if you work for a business, your employer deducts Social Security and Medicare taxes from your paycheck. Also, the business pays a portion of the taxes for you. These taxes are called Federal Insurance Contributions Act (FICA) taxes.

People who own their own businesses are self-employed. They have to pay their own taxes. This
is called the self-employment tax. In this problem situation, you will calculate how much three self-employed individuals owe in self-employment tax. You will do this by using a tax worksheet called the Schedule SE. This is an **Internal Revenue Service (IRS)** tax form. The IRS is the part of the government that collects taxes. It has many different types of forms that help individuals figure out how much they owe in taxes.

(1) Below is a copy of a blank schedule SE.

Marianne Lopez has a part-time job as a math tutor. She offers tutoring services to students both in person and online. Last year, she earned $11,385 in **revenue**, or the amount of money she received from selling her service. Her expenses totaled $3,862. You are going to calculate how much self-employment tax Marianne owes. That is, you are going to determine what value should be put into the cell on line 5. Take a moment to do this on your own before sharing your calculations in your group.

Note: Marianne does not have a farm, and she does not receive social security retirement or disability benefits. She is also not a minister or a member of a religious order and does not have other income. You do not need to know which schedule or form Marianne's **net profit** (the actual amount of money made after expenses) comes from. These cells have been zeroed out to avoid confusion. You also do not need to calculate line 13.

Part I Self-Employment Tax

Note: If your only income subject to self-employment tax is **church employee income,** see instructions for how to report your income and the definition of church employee income.

A If you are a minister, member of a religious order, or Christian Science practitioner **and** you filed Form 4361, but you had $400 or more of **other** net earnings from self-employment, check here and continue with Part I . . . . . ▶ ☐

Skip lines 1a and 1b if you use the farm optional method in Part II. See instructions.

1a	Net farm profit or (loss) from Schedule F, line 34, and farm partnerships, Schedule K-1 (Form 1065), box 14, code A		**1a**	0
b	If you received social security retirement or disability benefits, enter the amount of Conservation Reserve Program payments included on Schedule F, line 4b, or listed on Schedule K-1 (Form 1065), box 20, code AH		**1b**	(0)

Skip line 2 if you use the nonfarm optional method in Part II. See instructions.

2	Net profit or (loss) from Schedule C, line 31; and Schedule K-1 (Form 1065), box 14, code A (other than farming). See instructions for other income to report or if you are a minister or member of a religious order		**2**	
3	Combine lines 1a, 1b, and 2		**3**	
4a	If line 3 is more than zero, multiply line 3 by 92.35% (0.9235). Otherwise, enter amount from line 3		**4a**	
	Note: If line 4a is less than $400 due to Conservation Reserve Program payments on line 1b, see instructions.			
b	If you elect one or both of the optional methods, enter the total of lines 15 and 17 here		**4b**	0
c	Combine lines 4a and 4b. If less than $400, **stop;** you don't owe self-employment tax. **Exception:** If less than $400 and you had **church employee income,** enter -0- and continue ▶		**4c**	
5a	Enter your **church employee income** from Form W-2. See instructions for definition of church employee income	**5a** 0		
b	Multiply line 5a by 92.35% (0.9235). If less than $100, enter -0-		**5b**	0
6	Add lines 4c and 5b		**6**	
7	Maximum amount of combined wages and self-employment earnings subject to social security tax or the 6.2% portion of the 7.65% railroad retirement (tier 1) tax for 2021		**7**	142,800
8a	Total social security wages and tips (total of boxes 3 and 7 on Form(s) W-2) and railroad retirement (tier 1) compensation. If $142,800 or more, skip lines 8b through 10, and go to line 11	**8a** 0		
b	Unreported tips subject to social security tax from Form 4137, line 10	**8b** 0		
c	Wages subject to social security tax from Form 8919, line 10	**8c** 0		
d	Add lines 8a, 8b, and 8c		**8d**	0
9	Subtract line 8d from line 7. If zero or less, enter -0- here and on line 10 and go to line 11 ▶		**9**	
10	Multiply the **smaller** of line 6 or line 9 by 12.4% (0.124)		**10**	
11	Multiply line 6 by 2.9% (0.029)		**11**	
12	**Self-employment tax.** Add lines 10 and 11. Enter here and on **Schedule 2 (Form 1040), line 4**		**12**	
13	**Deduction for one-half of self-employment tax.** Multiply line 12 by 50% (0.50). Enter here and on **Schedule 1 (Form 1040), line 15**	**13**		

(a) What value should be put into the cell on line 2?

(b) What value should be put into the cell on line 3?

(c) What value should be put into the cell on line 4a? Round your answer to two decimal places.

(d) What value should be put into the cell on line 6?

(e) What value should be put into the cell on line 9?

(f) What value should be put into the cell on line 10? Round your answer to two decimal places.

(g) What value should be put into the cell on line 11? Round your answer to two decimal places.

(h) What value should be put into the cell on line 12?

(2) Leigh Olson started a small bakery last year. She has no income to report on line 1a or line 1b. She earned $1,050 in revenue selling cookies and cupcakes. Her expenses totaled $630. In your group, work through the questions below to figure out how much self-employment tax Leigh Olson owes.

Note: Leigh does not have a farm, and she does not receive social security retirement or disability benefits. She is also not a minister or a member of a religious order. You do not need to know which schedule or form Leigh's net profit comes from. These cells have been zeroed out to avoid confusion. You also do not need to calculate line 13.

Part I Self-Employment Tax

Note: If your only income subject to self-employment tax is **church employee income,** see instructions for how to report your income and the definition of church employee income.

A If you are a minister, member of a religious order, or Christian Science practitioner **and** you filed Form 4361, but you had $400 or more of **other** net earnings from self-employment, check here and continue with Part I ▶ ☐

Skip lines 1a and 1b if you use the farm optional method in Part II. See instructions.

1a	Net farm profit or (loss) from Schedule F, line 34, and farm partnerships, Schedule K-1 (Form 1065), box 14, code A		1a	0
b	If you received social security retirement or disability benefits, enter the amount of Conservation Reserve Program payments included on Schedule F, line 4b, or listed on Schedule K-1 (Form 1065), box 20, code AH		1b	(0)

Skip line 2 if you use the nonfarm optional method in Part II. See instructions.

2	Net profit or (loss) from Schedule C, line 31; and Schedule K-1 (Form 1065), box 14, code A (other than farming). See instructions for other income to report or if you are a minister or member of a religious order		2	
3	Combine lines 1a, 1b, and 2		3	
4a	If line 3 is more than zero, multiply line 3 by 92.35% (0.9235). Otherwise, enter amount from line 3		4a	
	Note: If line 4a is less than $400 due to Conservation Reserve Program payments on line 1b, see instructions.			
b	If you elect one or both of the optional methods, enter the total of lines 15 and 17 here		4b	0
c	Combine lines 4a and 4b. If less than $400, **stop;** you don't owe self-employment tax. **Exception:** If less than $400 and you had **church employee income,** enter -0- and continue ▶		4c	
5a	Enter your **church employee income** from Form W-2. See instructions for definition of church employee income	5a 0		
b	Multiply line 5a by 92.35% (0.9235). If less than $100, enter -0-		5b	0
6	Add lines 4c and 5b		6	
7	Maximum amount of combined wages and self-employment earnings subject to social security tax or the 6.2% portion of the 7.65% railroad retirement (tier 1) tax for 2021		7	142,800
8a	Total social security wages and tips (total of boxes 3 and 7 on Form(s) W-2) and railroad retirement (tier 1) compensation. If $142,800 or more, skip lines 8b through 10, and go to line 11	8a 0		
b	Unreported tips subject to social security tax from Form 4137, line 10	8b 0		
c	Wages subject to social security tax from Form 8919, line 10	8c 0		
d	Add lines 8a, 8b, and 8c		8d	0
9	Subtract line 8d from line 7. If zero or less, enter -0- here and on line 10 and go to line 11 ▶		9	
10	Multiply the **smaller** of line 6 or line 9 by 12.4% (0.124)		10	
11	Multiply line 6 by 2.9% (0.029)		11	
12	**Self-employment tax.** Add lines 10 and 11. Enter here and on **Schedule 2 (Form 1040), line 4**		12	
13	**Deduction for one-half of self-employment tax.** Multiply line 12 by 50% (0.50). Enter here and on **Schedule 1 (Form 1040), line 15**	13		

(a) What value should be put into the cell on line 2?

(b) What value should be put into the cell on line 3?

(c) What value should be put into the cell on line 4a? Round your answer to two decimal places.

(d) How much tax does Leigh owe? Be careful here!

(3) In Question 1, you learned about Marianne Lopez. You used the form to calculate how much self-employment tax Marianne owes. Now, write the calculation you completed to answer Question 1 as a single expression. An example of a single expression is shown below.

Example: The following calculations,

19 - 4 = 15

5.75*15 = 86.25

0.3*86.25 = 25.875

Can be written as a single expression: 0.3(5.75)(19 - 4) = 25.875

(4) Look back at the single expression you wrote for Question 3. Imagine you have to explain the expression to Marianne to help her understand how you calculated the amount of self-employment tax that she owes. Answer these questions about the expression:

(a) What does the operation $11,385 − $3,862 represent for Marianne?

(b) What does the operation of multiplying by 0.9235 represent for Marianne?

(c) What does the operation of multiplying by (0.124+0.029) represent for Marianne?

Self-Employment and the Medicare Tax

The Affordable Care Act was passed by Congress and signed into law in 2010. The law reforms the health care system in the United States by protecting consumers from losing their health insurance, bringing down insurance costs, expanding access to health care services, strengthening Medicare with added benefits, and improving the quality of health care for Medicare patients.

Medicare is a federal (government-run) health insurance program for seniors, people who are 65 and older, in the United States. All taxpayers pay into the Medicare program.

In order to raise revenue to support the expansion of care under the Affordable Care Act, the U.S. government created an additional Medicare tax for individuals with higher incomes. People who are self-employed use Part II of Form 8959 to determine how much additional Medicare tax they have to pay (if any). Only individuals and couples with incomes over certain amounts have to pay this tax. These amounts appear in line 9 of Form 8959.

Part II Additional Medicare Tax on Self-Employment Income

8	Self-employment income from Schedule SE (Form 1040), Section A, line 4, or Section B, line 6. If you had a loss, enter -0- (Form 1040-PR and Form 1040-SS filers, see instructions.)	8	$215,500	
9	Enter the following amount for your filing status: Married filing jointly . . . . . $250,000 Married filing separately . . . . . $125,000 Single, Head of household, or Qualifying widow(er) $200,000	9	$200,000	
10	Enter the amount from line 4	10	$0	
11	Subtract line 10 from line 9. If zero or less, enter -0-	11	$200,000	
12	Subtract line 11 from line 8. If the result is zero or less, enter -0-			12
13	Additional Medicare Tax on self-employment income. Multiply line 12 by 0.9% (.009). Enter here and go to Part III			13

(5) Andie Henson is a single, self-employed psychiatrist. Her self-employment income from Schedule SE (Form 1040) Section A was $215,500. She did not earn wages from another job during this year. This information has already been entered in lines 8–11 of Form 8959. Use mental math to find the exact value of line 12 and add it to the form. Then, estimate the value of line 13 and record your estimate in the space below. Do not use a pencil and paper or a calculator. Briefly explain your estimation strategy.

PROBLEM SITUATION 2: PROPORTIONS IN ARTWORK

Many professionals such as graphic artists, architects, and engineers work with objects that are enlarged or shrunk. It is usually important that the objects have the same appearance despite the change in size. For example, a business logo on a billboard needs to look the same as a logo on a coffee mug.

In this collaboration, you will explore the mathematics behind these changes in size.

Imagine that this graphic is our original image. It has the following dimensions:

Width: 1.9 inches
Height: 1 inch

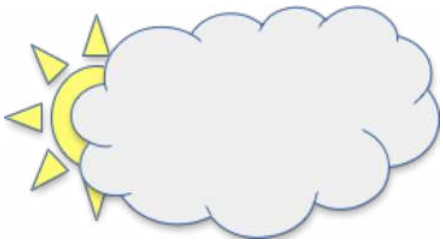


Let's say we make a change to the width only. This new graphic has the following dimensions:

Width: 2.45 inches
Height: 1 inch



Now, let's say we take the original image and we make changes to both the width and the height, but we do not maintain proportionality. This new graphic has the following dimensions:

Width: 2.2 inches Height: 1.19 inches	

And lastly, let's imagine we take the original image and make changes to both the width and the height, maintaining proportionality. This new graphic has the following dimensions:

Width: 2.55 inches Height: 1.29 inches	

Notice that it can be difficult to spot just by looking at the new image if it is proportional to the original image. In the next question, you're going to explore what can be done to check for proportionality.

(6) Suppose you were given three different kinds of dimensional changes done to the original graphic without actually seeing the new graphics. The dimensional changes are shown in the table below.

Original graphic dimensions:

Width: 1.9 inches

Height: 1 inch

Three Different Kinds of Changes					
Change One		Change Two		Change Three	
Width	Height	Width	Height	Width	Height
7.45 in	1 in	4 in	1.19 in	0.95 in	0.5 in

How could you tell which changes were proportional and which were not?

(7) You are a graphic artist hired to make a billboard for a college. The original logo is 2-1/4 inches (width) by 3-3/8 inches (length). You need to enlarge the logo so it has a length of six feet. How wide will the enlarged version be? **Hint**: Think back to the work about equal ratios in Question 1 and to the idea of using a variable to represent an unknown. What could we do if we use the variable x to stand in for the width of the enlarged version?

(8) In Question 7, you could have used the following proportion to represent the relationship between the original and enlarged objects:

$$\frac{2.25}{3.375} = \frac{4}{6}$$

Could this proportion be written in other ways? Explain.

(9) Suppose you had set up the following proportion to solve the original problem in Question 7. What steps would you use to solve the equation?

$$\frac{3.375}{2.25} = \frac{6}{x}$$

(10) Solve the equation. Round to the nearest tenth.

$$\frac{12.7}{x} = \frac{0.2}{3}$$

(11) Solve the equation. Round to the nearest tenth.

$$\frac{8,500}{4,200} = \frac{x}{5}$$

(12) Many small engines for saws, motorcycles, and utility tractors require a mixture of oil and gas. If an engine requires 20 ounces of oil for five gallons of gas, how much oil would be needed for eight gallons of gas?

PROBLEM SITUATION 3: ESTIMATING COSTS FOR CAR OR ACTIVITY

(13) If you drive a car or plan to get a car, complete (a) below. If you do not have a car, complete (b) below. Show your work.

(a) Estimate the cost per mile of driving your car based on what you actually pay for insurance and gas mileage. Don't forget about the cost of maintenance and how often you drive.

(b) If you do not have a car, estimate the cost per year of some activity or item that you pay for at least twice a week on average. For example, buying a cup of coffee or energy drink, downloading music, going to a movie, paying a babysitter, riding a bus, etc.

This page titled 4.5: Corequisite- Percentages and Taxes, Ratios, Proportions and Dimensional Analysis is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.5.1: Practice Problems Corequisite N.5

(1) Martin Binford is an author. He has no income he would report on line 1a or line 1b of his Schedule SE. He earned $163,380 last year from his books. He had $3,563 in expenses. (**Note**: He records his net profit in line 2, where it says "Net profit or (loss).") You do not need to know which schedule or form it comes from, and he is not a minister or member of a religious order. He also does not get a deduction for one-half of self employment tax. How much self-employment tax does he owe (line 12)? Round to the nearest dollar.

Part I Self-Employment Tax

Note: If your only income subject to self-employment tax is **church employee income,** see instructions for how to report your income and the definition of church employee income.

A If you are a minister, member of a religious order, or Christian Science practitioner **and** you filed Form 4361, but you had $400 or more of **other** net earnings from self-employment, check here and continue with Part I . . . . . ▶ ☐

Skip lines 1a and 1b if you use the farm optional method in Part II. See instructions.

1a	Net farm profit or (loss) from Schedule F, line 34, and farm partnerships, Schedule K-1 (Form 1065), box 14, code A		**1a**	0
b	If you received social security retirement or disability benefits, enter the amount of Conservation Reserve Program payments included on Schedule F, line 4b, or listed on Schedule K-1 (Form 1065), box 20, code AH		**1b**	(0)

Skip line 2 if you use the nonfarm optional method in Part II. See instructions.

2	Net profit or (loss) from Schedule C, line 31; and Schedule K-1 (Form 1065), box 14, code A (other than farming). See instructions for other income to report or if you are a minister or member of a religious order		**2**	
3	Combine lines 1a, 1b, and 2		**3**	
4a	If line 3 is more than zero, multiply line 3 by 92.35% (0.9235). Otherwise, enter amount from line 3		**4a**	
	Note: If line 4a is less than $400 due to Conservation Reserve Program payments on line 1b, see instructions.			
b	If you elect one or both of the optional methods, enter the total of lines 15 and 17 here		**4b**	0
c	Combine lines 4a and 4b. If less than $400, **stop;** you don't owe self-employment tax. **Exception:** If less than $400 and you had **church employee income,** enter -0- and continue ▶		**4c**	
5a	Enter your **church employee income** from Form W-2. See instructions for definition of church employee income	**5a** 0		
b	Multiply line 5a by 92.35% (0.9235). If less than $100, enter -0-		**5b**	0
6	Add lines 4c and 5b		**6**	
7	Maximum amount of combined wages and self-employment earnings subject to social security tax or the 6.2% portion of the 7.65% railroad retirement (tier 1) tax for 2021		**7**	142,800
8a	Total social security wages and tips (total of boxes 3 and 7 on Form(s) W-2) and railroad retirement (tier 1) compensation. If $142,800 or more, skip lines 8b through 10, and go to line 11	**8a** 0		
b	Unreported tips subject to social security tax from Form 4137, line 10	**8b** 0		
c	Wages subject to social security tax from Form 8919, line 10	**8c** 0		
d	Add lines 8a, 8b, and 8c		**8d**	0
9	Subtract line 8d from line 7. If zero or less, enter -0- here and on line 10 and go to line 11		**9**	
10	Multiply the **smaller** of line 6 or line 9 by 12.4% (0.124)		**10**	
11	Multiply line 6 by 2.9% (0.029)		**11**	
12	**Self-employment tax.** Add lines 10 and 11. Enter here and on **Schedule 2 (Form 1040), line 4**		**12**	
13	**Deduction for one-half of self-employment tax.** Multiply line 12 by 50% (0.50). Enter here and on **Schedule 1 (Form 1040), line 15**	**13**		

(2) Which of the following expressions can be used to compute how much self-employment tax Martin Binford owes? How do you know if your answer is correct?

(i) 0.029(0.9235)($163,380 – $3,563)

(ii) 0.124($142,800) + 0.029(0.9235)($163,380 – $3,563)

(iii) 0.124(0.9235)($163,380 – $3,563) + 0.029(0.9235)($163,380 – $3,563)

(iv) 0.029(0.9235)($163,380 – $3,563 + $17,707.20)

(3) The expression below shows another way to calculate Martin's tax.

$$0.124(142,800)+0.029(150,881.43-142,800)$$

Based on this expression, select the statement that describes how Martin's income is taxed.

(i) Martin pays 15.59% tax on his income.

(ii) Martin pays 44.3% tax on his income.

(iii) Martin pays 12.4% in tax on the first $142,800 of his income. He pays 29% on his income over $142,800.

(iv) Martin pays 12.4% in tax on the first $142,800 of his income. He pays 2.9% on his income over $142,800.

(4) Miguel is moving and wants to estimate what his electricity bill will be in his new apartment. He looks at old bills and sees that he uses around 700 kWh of electricity each month. The utility company charges $6 each month plus 6.726 cents per kWh for the first 500 kWhs and 8.136 cents for each kilowatt-hour above 500. How much will Miguel pay for 700 kWh of electricity?

Workers in the U.S. pay several types of taxes on income, including the FICA tax and the federal income tax. Your federal income tax rate is based on the amount of money you make. Income is broken into levels called tax brackets. The table below shows the tax brackets for 2021.

Taxable Income	**Taxes Owed**
$0–$9,950	10% of taxable income
$9,951–$40,525	$995.00 plus 12% of excess over $9,950
$40,526–$86,375	$4,664.00 plus 22% of excess over $40,525
$86,376–$164,925	$14,751.00 plus 24% of excess over $86,375
$164,926–$209,425	$33,603.00 plus 32% of excess over $164,925
$209,426–$523,600	$47,843.00 plus 35% of excess over $209,425
$523,601 plus	$157,804.25 plus 37% of excess over $523,600

(5) Calculate the tax for a person with taxable income of $25,000. Round to the nearest dollar.

(6) Calculate the tax for a person with taxable income of $63,500. Round to the nearest dollar.

(7) Refer to your answer for the previous question. The total tax from the previous question is what percentage of the person's taxable income? Round to the nearest one percent. Write your answer using the Writing Principle.

In planning a Thanksgiving vacation, you want to rent a car for a week and travel the Pacific Highway from San Francisco to San Diego. You want to return to San Francisco via Las Vegas, Death Valley, and Yosemite National Park. This trip covers approximately 1,500 miles. You plan to return the car with a full tank of gasoline. You are considering two options advertised by Hertz:

- Toyota Camry, 33 mpg highway, costs $465.59 plus taxes and fees, totaling $553.28.
- Toyota Prius Hybrid, 48 mpg highway, costs $634.49 plus taxes and fees, totaling $804.61.

In both cases, you must purchase gasoline, which costs $3.80 per gallon in California.

(8) Find the **total gasoline cost** of renting and driving the **Camry** for the trip:

Total gasoline cost is $_____

(9) Find the **total cost** of renting and driving the **Camry** for the trip:

Total cost is $_____

(10) Find the **total gasoline cost** of renting and driving the **Prius** for the trip:

Total gasoline cost is $_____

(11) Find the **total cost** of renting and driving the **Prius** for the trip:

Total cost is $_____

The advertised mpg for new cars is based on a speed of 55 mph. As speed increases above 55 mph, the efficiency reduces dramatically.[12]

- 3% less efficient at 60 mph
- 17% less efficient at 70 mph

Compare the efficiency of a Toyota Camry (33 mpg highway) versus a Toyota Prius (48 mpg highway). Round to the nearest tenth.

(12) Camry at 60 mph:

(13) Camry at 70 mph:

(14) Prius at 60 mph:

(15) Prius at 70 mph:

If gasoline costs $3.67 per gallon, how much money would you save by going 60 mph versus 70 mph on the trip of 1,500 miles in the

(16) Camry? $_____

(17) Prius? $_____

(18) How much longer, to the nearest minute, would it take to travel 300 miles at 60 mph versus 70 mph? Write your answer using the Writing Principle.

_____

[12] www.mpgforspeed.com

This page titled 4.5.1: Practice Problems Corequisite N.5 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.6: Corequisite- Dimensional Analysis, Unit Conversion

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- the units found in a solution may be used as a guide to the operations required in the problem—that is, factors are positioned so that the appropriate units cancel.
- units can add meaning to the numbers that result from calculations.
- formulas can be found by searching the Internet and reference books.
- a variable can be used to represent an unknown.
- using a formula requires knowing what each variable represents.

By the end of this lesson, you should be able to

- calculate quantities in the billions.
- write a rate as a fraction.
- use a unit factor to simplify a rate.
- use dimensional analysis to help determine the factors in a series of operations to obtain an equivalent measure.
- use formulas from geometry and perform calculations that involve rates and measures to support financial decisions.
- evaluate an expression.
- use the appropriate units for length, area, and volume.

PROBLEM SITUATION 1: USING DIMENSIONAL ANALYSIS

Dimensional analysis is a method of setting up problems that involves converting between different units of measurement. It is also called *unit analysis* or *unit conversion*. Many professionals—including pharmacists, dieticians, lab technicians, and nurses—use dimensional analysis. It is also useful for everyday conversions in cooking, finances, and currency exchanges. Many people can do simple conversions without dimensional analysis; however, they will likely make mistakes on more complex problems.

The advantage of using dimensional analysis is that it is a way to check your calculations. While it is always important that you develop your own methods to solve problems, this is a time when you are encouraged to learn and use a specific method. Once you have learned dimensional analysis, you can decide when to use it and when to use other methods.

(1) (a) According to Toyota's website, a 2023 Prius can get an estimated 57 miles per gallon (mpg) in the city and 56 mpg on the highway.[13] How many miles will you be able to drive in the city if you have 4.5 gallons of gas?

(b) How many gallons of gas will you need to drive 3,450 miles? Round your answer to one decimal place.

Converting Your Paycheck

This is an example of how to use dimensional analysis to solve a problem.

Sample Question: Your paycheck for two weeks came out to $1,200. You work eight hours a day, five days a week. How much are you making per minute in cents? Let's work through answering this question using dimensional analysis.

Answer: First, decide what you are looking for. Here you are looking for the rate of cents per minute, which can be written as:

$$\frac{c}{min}$$

Now, identify a rate that we know has the same numerator (the part of a fraction that represents a count of the number of parts) as the rate we are looking for. Since cents converts to dollars, we have:

$$\frac{100c}{\$1} \rightarrow \frac{c}{min}$$

Next, we want to cancel the "$" (since this unit is not in the answer we are looking for). To cancel the "$" we can multiply by a unit ratio with "$" in the numerator. Since $1,200 is the amount we make in two weeks, we can multiply by that ratio:

$$\frac{100c}{\$1} \times \frac{\$1200}{2\ weeks} \rightarrow \frac{c}{min}$$

Now, we need to cancel the unit "weeks", so our next ratio must have "weeks" in the numerator:

$$\frac{100c}{\$1} \times \frac{\$1200}{2\ \cancel{weeks}} \times \frac{\cancel{week}}{5\ days} \rightarrow \frac{c}{min}$$

Continuing this process, we finally get to the unit factor for "minute":

$$\frac{100c}{\$1} \times \frac{\$1200}{2\ \cancel{weeks}} \times \frac{\cancel{week}}{5\ \cancel{days}} \times \frac{\cancel{day}}{8\ \cancel{hours}} \times \frac{\cancel{hour}}{60\ minutes} \rightarrow \frac{c}{min}$$

(2) Is the resulting calculation reasonable? Explain.

(3) Many states have banned texting while driving because it is dangerous, but many people do not think that texting for a few seconds is that harmful.

Suppose you are driving 60 miles/hour and you take your eyes off the road for four seconds. How many feet (ft) will you travel in that time? **Hint**: Remember that there are 5,280 ft in one mile. Start with the unit you are looking for (ft). Then, create a chain of ratios, starting with one where "ft" is in the numerator, that will cancel all other units.

(4) Let's examine population densities and use these to calculate projected populations. The population density of Tokyo is 6,038 people per square kilometer (km). Use dimensional analysis to calculate how many people would live in the nation of Japan, which comprises an area of approximately 378,000 square km, if the entire nation was as dense as the city of Tokyo.

(5) Nurses are often required to calculate dosages. That is, they must check the order that a doctor has given for the administration of a drug and decide whether the dosage is correct. To calculate correctly they must convert between different metric units. For example, 1,000 milligrams (mg) = 1 gram (g); and 1,000 micrograms (mcg) = 1 mg.

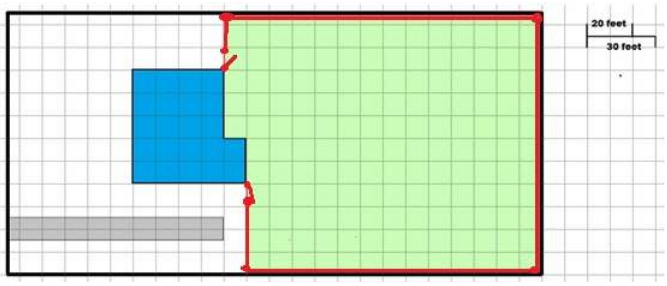
Suppose a doctor has ordered a dose of 0.1 gram of a medication. The drug comes in a solution concentration of 200 mg per milliliter. How many milliliters of this solution is required?

(6) Now, calculate how many milliliters you would need to administer 500 mg from a dosage concentration of 1 g per 3 mL. Use the same procedure as in the previous question.

PROBLEM SITUATION 2: HOME IMPROVEMENTS

Bob and Carol Mazursky have purchased a house and the small piece of land that surrounds the house. A small piece of land that can be purchased or sold is often called a lot. Bob and Carol want to make some improvements to their new property. In the following few problems, you will calculate the costs of these improvements. Scale drawings of the house and the lot are displayed in the problems. You will need these drawings to calculate improvement costs.

Figure 1

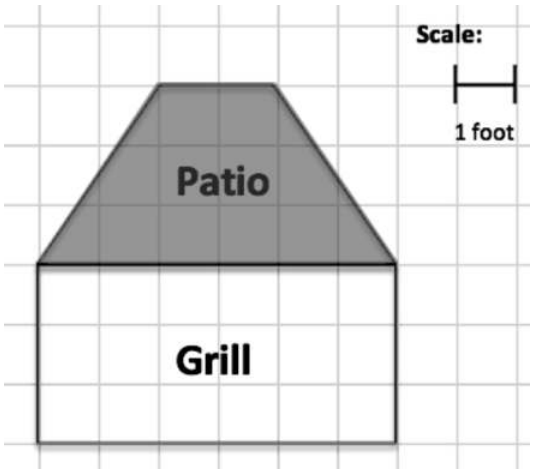


(7) Review the drawing of the house and lot (Figure 1). What does the scale mean?

(8) Bob and Carol want to get the backyard (Figure 1) fertilized and reseeded. They found an ad for Gerry's Green Team lawn service, which can be seen here. Gerry came to their house and said that the job would take about half a day and would cost about $600. Is Gerry's estimate consistent with his advertisement? Why or why not?



(9) There is a brick grill in the backyard. Bob and Carol are going to make a concrete patio in the shape of a trapezoid next to the grill (see below). The concrete slab needs to be at least 2 inches thick. They will use 40-pound bags of premixed concrete. Each 40-pound bag makes 0.30 cubic feet of concrete and costs $6.50. How much will the materials cost, including the 7.5% tax?
Outdoor Grill: Bob and Carol are going to add a trapezoidal patio adjacent to the outdoor grill in the backyard. The shaded area is made of concrete and two inches deep.



[13] https://www.toyota.com/prius/

This page titled 4.6: Corequisite- Dimensional Analysis, Unit Conversion is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.6.1: Practice Problems Corequisite N.6

(1) As of March 2023, Florence Griffith-Joyner held the women's world record for the 100-meter dash. She set the record with a time of 10.49 seconds in 1988.[14] Which of the following calculations are correctly set up to convert this speed into miles per hour?

(i) $\frac{10.49\ seconds}{100\ meters} \cdot \frac{1\ meter}{3.28\ feet} \cdot \frac{5,280\ feet}{1\ mile} \cdot \frac{1\ minute}{60\ seconds} \cdot \frac{1\ hour}{60\ minutes}$

(ii) $\frac{100\ meters}{10.49\ seconds} \cdot \frac{3.28\ feet}{1\ meter} \cdot \frac{1\ mile}{5,280\ feet} \cdot \frac{60\ seconds}{1\ minute} \cdot \frac{60\ minutes}{1\ hour}$

(iii) $\frac{10.49\ seconds}{100\ meters} \cdot \frac{1\ meter}{3.28\ feet} \cdot \frac{5,280\ feet}{1\ mile} \cdot \frac{60\ seconds}{1\ minute} \cdot \frac{60\ minutes}{1\ hour}$

(iv) $\frac{100\ meters}{10.49\ seconds} \cdot \frac{3.28\ feet}{1\ meter} \cdot \frac{1\ mile}{5,280\ feet} \cdot \frac{1\ minute}{60\ seconds} \cdot \frac{1\ hour}{60\ minutes}$

(2) Find the answer to the conversion in the previous question. Round to the nearest tenth of a mile per hour.

(3) A 2023 Toyota Prius hybrid vehicle gets 56 mpg for highway driving. The tank holds 11.3 gallons of fuel.[15] Typically the low fuel warning light comes on when approximately two gallons of fuel remain in the tank. Which of the following calculations can be used to find the distance that can be traveled after the fuel light comes on and before the car runs out of gasoline?

(i) $\frac{1}{2\ gallons} \cdot \frac{56\ miles}{1\ gallon} = 28\ miles$

(ii) $\frac{2\ gallons}{1} \cdot \frac{56\ miles}{1\ gallon} = 112\ miles$

(iii) $\frac{11.3\ gallons}{1} \cdot \frac{56\ miles}{1\ gallon} = 632.8 miles$

(iv) $\frac{2\ gallons}{1} \cdot \frac{1\ gallon}{56\ miles} = \frac{1}{28} miles$

Part of a person's water footprint is the water used for cleaning. In this question, you will calculate the cost of water for laundry and bathing. You will use the City of New York 2023 rate of $9.87/100 cubic feet of water. Calculate the cost of each of the following based on this rate. Use the conversion factor of 7.48 gallons per cubic foot.[16]

(4) A standard washing machine uses approximately 50 gallons of water per load.[17] A household washes one load of laundry per week for 52 weeks. Find the total cost per year. Round to the nearest dollar per year.

(5) According to one study, the average American shower lasts for 8.2 minutes and uses 17.2 gallons of water.[18] A person showers once a day for a year. Find the total cost per year. Round to the nearest dollar per year.

Unit rates are ratios with a denominator of "1," although they are not always written as fractions. For example, 60 mph written as a unit rate is $\frac{60\ miles}{1\ hour}$.

The language "miles per hour" implies that the operation is miles divided by hours.

Write the expression as a unit rate in fractional form.

(6) 23 mpg (miles per gallon)

(7) 5 gal/min (gallons per minute)

(8) $7.15/hr (dollars per hour)

Convert the rate into a unit rate.

(9) $\frac{250\ miles}{12.5\ gallons}$

(10) $\frac{450\ miles}{8\ hours}$

(11) $\frac{\$382}{40\ hours}$

---

[14] http://en.Wikipedia.org/wiki/100_metres#Women

[15] https://www.toyota.com/prius/

[16] https://www.nyc.gov/site/dep/water/residential-water-use.page

[17] ibid.

[18] ibid.

This page titled 4.6.1: Practice Problems Corequisite N.6 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.7: Corequisite- Buying Power, Consumer Price Index, Absolute and Relative Change, Reading and Interpreting Graphs

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- ratios provide a way of comparing the relative increase or decrease of two variables.
- index numbers are a way of comparing the relative size of a variable over time.

By the end of this lesson, you should be able to

- use proportional reasoning to find the size of a variable that remains a constant proportion of another variable.
- use index numbers to find the value of a variable relative to time (or another variable).

PROBLEM SITUATION: THE BUYING POWER OF THE MINIMUM WAGE

The United States has a federal **minimum wage**. This means that there is a law that requires employers to pay employees a minimum amount. Many states also have their own minimum wage laws that require a higher wage than the federal law. The minimum wage started in 1938 as part of a law that protected the rights of workers in many ways. The minimum wage in 1938 was set at \$0.25 per hour.[19]

Since 1938, Congress has increased the minimum wage many times to account for inflation. **Inflation** is when prices increase over time. Maybe you have heard stories from older people about how a cup of coffee used to cost a dime or a gallon of gasoline was less than a dollar. Prices for these items have increased due to inflation. **Deflation** is when prices decrease over time.

When deciding if the minimum wage should be increased, people often talk about the purchasing power of the wage. **Did the minimum wage in 1938 have more purchasing power than the minimum wage in 2021?** You will start by thinking about the specific example of buying a movie ticket for the following questions where you can compare the purchasing power of the minimum wage in 1997 to the minimum wage in 2021.

(1) The minimum wage in 1997 was \$5.15 per hour. The minimum wage in 2021 was \$7.25 per hour.

(a) What was the absolute change in the minimum wage from 1997 to 2021?

(b) What was the relative change in the minimum wage from 1997 to 2021? Round to the nearest tenth of one percent.

(2) The average price of a movie ticket in 1997 was \$4.59.[20] Which statement is correct? Be prepared to explain your answer.

(i) In 1997, a person earning minimum wage had to work less than an hour to earn enough for a movie ticket.

(ii) In 1997, a person earning minimum wage had to work more than an hour to earn enough for a movie ticket.

(3) The average price of a movie ticket in 2021 was \$9.57.[21] Which statement is correct? Be prepared to explain your answer.

(i) In 2021, a person earning minimum wage had to work less than an hour to earn enough for a movie ticket.

(ii) In 2021, a person earning minimum wage had to work more than an hour to earn enough for a movie ticket.

As we just explored, the minimum wage did not increase enough from 1997 to 2021 to keep up with the cost of a movie ticket. Remember that the **minimum wage** is the minimum amount employers must pay their employees required by federal law. A movie ticket is only a single product that someone might buy. To compare the purchasing power of the minimum wage from 1938 to 2021, you need more information about more products. You will use a tool called an index number.

Index numbers are a way to compare the relative change or difference in a data set, such as the prices of products. The change can be measured over time or over different geographic regions. There are different types of index numbers. One type is a measure of average relative change in data. In this instance, one data point is called the base and is assigned the value of 100 (meaning it represents 100%). The other figures are adjusted in proportion to the base. The plural form of the word index is indices.

The following two questions will help you understand index numbers. You will begin by looking at one product: **the Big Mac**.

(4) You are going to use 1992 as the base year for our Big Mac Price Index. Prices of the Big Mac vary with time. In 1992, a Big Mac cost about \$2.19.

(a) In 1968, a Big Mac cost about $0.49. Compare the prices by finding the ratio of the 1968 price to the 1992 price. Express the ratio as a percent. Round to the nearest percent.

(b) In 2021, a Big Mac cost about $4.93.[22] Compare the prices by finding the ratio of the 2021 price to the 1992 price. Express the ratio as a percent. Round to the nearest percent.

(c) Now let's think about how to use your answers to (a) and (b) to complete the Big Mac Price Index for the years 1968, 1992, and 2021. In parts (a) and (b), what year did you use as the base? What is the index number for 1992 and 2021?

Big Mac Price Index

Year	Index Number
1968	
1992	
2021	

The Value of the Dollar

Now you have seen how an index is created and what it means. You are now going to look at a real index that the U.S. government uses to track changes in the prices of many items over time. It is called the Consumer Price Index (CPI). You will use a version called the **Urban Consumer Price Index (CPI-U)**.

The Bureau of Labor Statistics publishes the CPI-U each month. The CPI is a measure of the weighted average of a "basket of consumer goods and services." The CPI-U basket includes transportation, food, medical care, housing, apparel, recreation, and education. The CPI represents the cost of that basket of goods and services in *base-year dollars*. This allows us to compare the buying power of a dollar in one year as compared to another. If the CPI is more than 100, then average prices are higher than in the base year. If the CPI is less than 100, then average prices are lower than in the base year. Because the CPI compares prices over time, it is one of the most frequently used statistics for identifying periods of **inflation** or **deflation** (increase or decrease of prices over time). An increase in the CPI indicates *inflation* in prices. A decrease indicates *deflation*. The following table shows the indices for selected years from 1913 to 2021.[23]

Year	Index		Year	Index		Year	Index
1913	9.9		1949	23.8		2005	195.3
1914	10.0		1950	24.1		2006	201.6
1915	10.1		1981	90.9		2007	207.3
1921	17.9		1982	96.5		2008	215.3
1922	16.8		1983	99.6		2009	214.5
1923	17.1		1984	103.9		2010	218.1
1924	17.1		1985	107.6		2012	229.6
1925	17.5		1996	156.9		2014	236.0
1936	13.9		1997	160.5		2015	237.0
1937	14.4		1998	163.0		2016	240.0
1938	14.1		1999	166.6		2017	245.1
1939	13.9		2000	172.2		2018	251.1
1940	14.0		2001	177.1		2019	255.7
1946	19.5		2002	179.9		2020	258.8
1947	22.3		2003	184.0		2021	271.0

1948	24.1		2004	188.9		

(5) What year is used as the base for the CPI-U? What does the base mean?

(6) How does the 2021 dollar compare in value to the 1983 dollar?

(7) How does the 1949 dollar compare in value to the 1983 dollar?

(8) Who had more purchasing power? Provide quantitative information to support your answer.

- The person making minimum wage ($0.25 per hour) in 1938.
- The person making minimum wage ($7.25 per hour) in 2021.

Hint: What are some ways you could compare one index number (or wage) to another? How many working hours does it take in 1938 to buy one "basket of goods"? How many working hours does it take in 2021 to buy the same "basket of goods"?

(9) Find the price of gasoline in 1981 and in 2021. Give the sources for your information. Use the CPI-U table to evaluate the statement, “Gasoline was more expensive in 2021 than in 1981.” Provide mathematical information to justify your explanation.

---

[19] https://www.dol.gov/general/aboutdol/history/flsa1938

[20] https://www.statista.com/statistics/187091/average-ticket-price-at-north-american-movie-theaters-since-2001/

[21] https://www.statista.com/statistics/187091/average-ticket-price-at-north-american-movie-theaters-since-2001/.

[22] https://www.economist.com/big-mac-index

[23] https://www.usinflationcalculator.com/inflation/consumer-price-index-and-annual-percent-changes-from-1913-to-2008/

This page titled 4.7: Corequisite- Buying Power, Consumer Price Index, Absolute and Relative Change, Reading and Interpreting Graphs is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.7.1: Practice Problems Corequisite N.7

Let's say the minimum wage in 1994 was $5.25 per hour. The minimum wage in 2021 was $7.25 per hour.

(1) What was the absolute change in the minimum wage from 1994 to 2021?

(2) What was the relative change in the minimum wage from 1994 to 2021? Round to the nearest tenth of one percent.

(3) If the average price of a movie ticket in 1994 was $6.24, which statement is correct?

(i) In 1994, a person earning minimum wage had to work less than an hour to earn enough for a movie ticket.

(ii) In 1994, a person earning minimum wage had to work more than an hour to earn enough for a movie ticket.

(4) If the average price of a movie ticket in 2021 was $6.24, which statement is correct?

(i) In 2021, a person earning minimum wage had to work more than an hour to earn enough for a movie ticket.

(ii) In 2021, a person earning minimum wage had to work less than an hour to earn enough for a movie ticket.

For the questions below refer to the CPI-U table.

This table gives the cost of a representative basket of goods and services in base-year dollars (1983). The table shows the indices for selected years from 1913 to 2021. The Consumer Price Index (CPI-U) provides an indication of how the price of consumer goods changes over time.

Let's say that in 1999, a standard cup of coffee cost $1.48. In 2021, that same cup of coffee was $2.55.

(5) What was the relative increase in the price of the cup of coffee from 1999 to 2021? Round to the nearest one percent.

(6) What was the relative increase in the CPI-U from 1999 to 2021? Round to the nearest one percent.

(7) Which of the following statements is true?

(i) From 1999 to 2021, the price of coffee increased at a higher rate than the CPI-U.

(ii) From 1999 to 2021, the price of coffee increased at a lower rate than the CPI-U.

(iii) From 1999 to 2021, the price of coffee increased at the same rate as the CPI-U.

(8) The rent on a building in 1997 was $1,300 per month. The landlord changes the rent relative to changes in the CPI-U. Use the CPI-U to calculate the rent in 2000. Round to the nearest dollar.

This page titled 4.7.1: Practice Problems Corequisite N.7 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.8: Corequisite- Percentages, Division, Extracting Relevant Information

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- percentages involve a numerator (comparison value) and a denominator (reference value).
- a percent has different uses, including being used to express the likelihood (or probability) of a certain event.
- it is important to select the correct comparison value and reference value when calculating percentages.

By the end of this lesson, you should be able to

- correctly identify the quantities involved in a verbal statement about percentages.
- convert between ratios and percentages.
- convert between the decimal representation of a number and a percent.
- read and use information presented in a two-way table.
- extract relevant information from a table.
- select the appropriate values to calculate probabilities.

PROBLEM SITUATION 1: THE LANGUAGE OF PERCENTAGES

World Health Organization (WHO) is the part of the United Nations that oversees health issues in the world. The WHO leads numerous studies on tobacco use around the world. In its study called, “Gender and Tobacco,” the organization learned that tobacco use among females is increasing.[24] For example, recent research shows that just as many young girls smoke as young boys. The report is filled with information about percentages of females who smoke, percentages of males who smoke, and the percentage of smokers who start smoking by age 10. The language used to describe this information can be difficult to understand; therefore, pay close attention to the language used to describe a percent at the beginning of this lesson. This will help you to understand the findings reported in the WHO study.

Consider the following two quantities:

- Quantity 1 (Q1): The percentage of females who smoke.
- Quantity 2 (Q2): The percentage of smokers who are female.

(1) (a) Are these two quantities equal (Q1 = Q2)?

(b) Could Q1 be greater than Q2 (Q1 > Q2)?

(c) Could Q1 be less than Q2 (Q1 < Q2)?

(d) Explain your reasoning.

(2) What information would you need to calculate these percentages?

Smoking on Two College Campuses

Questions 3 and 4 present two situations with data about smoking. You can use these situations to test your ideas from Questions 1 and 2.

(3) Suppose a study on smoking was conducted at Midland University. The following table indicates the results of the study.

	Males	**Females**
Smokers	4,572	5,362
Nonsmokers	10,284	12,736

(a) What percentage of females smoke at Midland University? Round your answer to the nearest percent. Hint: How many females are involved in the study? How many females are smokers? Think about adding rows and columns.

(b) What percentage of smokers at Midland University are female? Round your answer to the nearest percent.

(4) Suppose a study was conducted at Northwest College. The following table indicates the results of the study.

	Males	Females
Smokers	1,256	536
Nonsmokers	1,028	1,053

(a) What percentage of females smoke at Northwest College? Round your answer to the nearest percent.

(b) What percentage of smokers at Northwest College are female? Round your answer to the nearest percent.

(c) A newspaper stated that 40% of the male students at Northwest College are smokers. Is that claim reasonable? Explain why or why not.

(5) In 2021, the World Health Organization conducted a study about the prevalence of tobacco use worldwide.[25] The organization reported that 2.1% of the adult females in China smoke tobacco products. In the United States, 15.2% of adult females smoke. Using this information answer the following questions:

(a) Out of 100 adult females in China, about how many are smokers?

(b) Out of 1,000 adult females in China, about how many are smokers?

(c) Out of 100 adult females in the United States, about how many are smokers?

(d) Out of 1,000 adult females in the United States, about how many are smokers?

(e) Suppose you read that 590 out of 1,000 males in China smoke. Based on this data, what percentage of males in China smoke? Round to the nearest whole percent.

(f) Are there more female smokers in China or the United States?

(6) A teacher has collected data on the grades his students received in his morning and afternoon classes. The following tables show two different ways to represent the same data as percentages.

Table 1

	Grade A	Grade B	Grade C	Grade D	Grade F
Morning Class	12.5%	25.0%	37.5%	6.3%	18.7%
Afternoon Class	14.3%	20.0%	37.1%	8.6%	20.0%

Table 2

	Grade A	Grade B	Grade C	Grade D	Grade F
Morning Class	44.4%	53.3%	48.0%	40.0%	46.2%
Afternoon Class	55.6%	46.7%	52.0%	60.0%	53.8%

(a) Which table could be used to answer the following question: What percentage of the students who received an “A” is in the morning class? Explain your answer.

(b) Which table could be used to answer the following question: What percentage of the students in the morning class received an “A”? Explain your answer.

PROBLEM SITUATION 2: USING PERCENTAGES TO DESCRIBE THE ACCURACY OF MEDICAL TESTS

Some athletes use performance-enhancing drugs (PEDs) to improve how they do in sports. Schools, sports leagues, and other sports organizations usually do not allow the use of PEDs. These groups can administer athletes a blood or urine test to determine if the athletes are using drugs. In this problem situation, you will calculate the *probability* that such a test gives correct results and the probability that such a test gives incorrect results. **Probability** means the chance that something happens. Probabilities are often reported in percentages (%).

Five hundred athletes have undergone a test to determine if they use PEDs. A positive (+) test result indicates that the athlete is using a PED. A negative (–) test result indicates the athlete is not using these drugs. However, this test is not 100% accurate. This

means that some errors may be present in the test results. The table below shows how often the test correctly determines if athletes used PEDs.

	Athletes Using PEDs	Athletes Not Using PEDs	Total
Positive Test Result	9	5	(a)
Negative Test Result	1	485	**486**
Total	**10**	(b)	**500**

Use the numbers in the table to answer the following questions. Be careful of what numbers you use for the numerator and denominator in your calculations.

(7) The table is missing one row total (a) and one column total (b). Fill in the missing totals.

(8) You might want to know: “If an athlete is using PEDs, what is the probability that this test gives a positive result?” Use the steps below to answer this question.

(a) How many athletes are using PEDs?

(b) How many of the athletes using PEDs received a positive test result?

(c) If an athlete is using PEDs, what is the probability that this test gives a positive result?

(9) The table shows that one athlete who was using PEDs received a negative test result. This means the test incorrectly identified this single athlete. What percentage of athletes who are using PEDs incorrectly test negative? **Hint**: Think about the ratio of the number of incorrect negative results compared to the number of athletes who were using PEDs.

(10) You might want to know: “If an athlete is not using PEDs, what is the probability that this test gives a negative result?” Work through (a) - (c) below to answer this question.

(a) How many athletes are not using PEDs?

(b) How many of the athletes not using PEDs received a negative test result?

(c) If an athlete is not using PEDs, what is the probability that this test gives a negative result? Record your answer as a percentage in the space below. Round to the nearest whole percent.

(11) Think about a situation in which a school principal receives a positive result on a PED test for a student athlete. Refer back to the table from earlier. Answer these questions:

(a) What percentage of athletes not using PEDs incorrectly test positive? **Hint**: Think about the ratio of the number of incorrect positive results compared to the number of those who are not using PEDS.

(b) How should the principal think about this percentage? What recommendation would you make to the principal about what she should do when she receives a positive result on a PED test for a student athlete?

(12) Are your results from Question 8(c) and Question 9 related? Are your results from Question 10(c) and Question 11(a) related? If yes, briefly explain.

(13) For (a) - (c) below, refer back to the table from earlier.

(a) You might also want to know: “What percentage of all positive test results are from athletes who are using PEDs?” Use the table to answer that question. **Hint**: Think about the ratio of the number of correct positive results compared to the number of all positive test results.

(b) Now answer, "What percentage of all positive test results are from athletes who are NOT using PEDs?" Record your answer as a percentage in the space below. Round your answer to one decimal place.

(c) Does your answer to part (b) surprise you? Does this make you rethink your answer for Question 11(b)? Now what recommendation would you make to the principal about what she should do when she receives a positive result on a PED test for a student athlete?

(14) You might also want to know: “What percentage of all negative test results are from athletes who are not using PEDs?” Use the table to answer that question. **Hint**: Think about the ratio of the number of correct negative results compared to the number of

all negative test results.

(15) You can use a percentage to show how accurate the PED test was. A test is **accurate** when it produces very few mistakes or errors. Pick one percentage that you think best describes how accurate the test was. Explain what this percentage says about the test and why you picked this percentage.

(16) Now, think about how to use a percentage to show how inaccurate the test was. A test is **inaccurate** if it produces many errors. Pick one percentage to show how inaccurate the test was. What does this percentage say about the test and why you picked this percentage?

---

[24] https://apps.who.int/iris/handle/10665/43657

[25] https://www.who.int/publications/i/item/9789240039322

This page titled 4.8: Corequisite- Percentages, Division, Extracting Relevant Information is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

4.8.1: Practice Problems Corequisite N.8 and N.9

A teacher has collected data on the grades his students received in his morning and afternoon classes. The following tables show two different ways to represent the same data as percentages.

Table 1

	Grade A	Grade B	Grade C	Grade D	Grade F
Morning Class	12.5%	25.0%	37.5%	6.3%	18.7%
Afternoon Class	14.3%	20.0%	37.1%	8.6%	20.0%

Table 2

	Grade A	Grade B	Grade C	Grade D	Grade F
Morning Class	44.4%	53.3%	48.0%	40.0%	46.2%
Afternoon Class	55.6%	46.7%	52.0%	60.0%	53.8%

(1) What percentage of the students who received an “B” is in the morning class? Write your answer in a complete sentence.

(2) What percentage of the students in the morning class received an “B”? Write your answer in a complete sentence.

(3) What are the reference values in Table 1?

(i) The number of students in a certain class.

(ii) The number of students who earned a certain grade.

(iii) The number of students in a certain class who earned a certain grade.

(iv) The total number of students in both classes.

(4) What are the reference values in Table 2?

(i) The number of students in a certain class.

(ii) The number of students who earned a certain grade.

(iii) The number of students in a certain class who earned a certain grade.

(iv) The total number of students in both classes.

A test is administered to 500 athletes to determine if they are using performance-enhancing drugs (PEDs). It is assumed that a positive test result indicates that the athlete is using PEDs, and a negative test result indicates that the athlete is not using PEDs. However, this test is not 100% accurate, so some errors occur. The following table shows the test results for a group of athletes.

	Athletes Using PEDs	Athletes Not Using PEDs	Total
Positive Test Result	90	4	**94**
Negative Test Result	10	396	**406**
Total	**100**	**400**	**500**

Use the information in the table to answer the following questions.

(5) How many athletes are using PEDs?

(6) If an athlete is using PEDs, which of the following describes the chance that this test will give a positive result? There may be more than one correct answer.

(i) 9%

(ii) 90%

(iii) 9 out of 100

(iv) 90 out of 100

(7) What is the chance of the test giving a negative result for an athlete using PEDs?

(8) If an athlete is not using PEDs, what is the chance that this test will give a negative result?

(i) 39.6%

(ii) 99%

(iii) 4 out of 396

(iv) 4 out of 400

(9) What is the chance of the test giving a positive result for an athlete not using PEDs?

A hospital tracks the number of cases (patients) that come into its Emergency Room during each eight-hour shift. The cases are listed in categories based on the severity of the illness or injury. The categories from least severe to most severe are: stable, serious, and critical. The following table gives the data for a week.

Shift	**Stable**	**Serious**	**Critical**	**Total**
8:00 a.m. – 3:59 p.m.	250	120	45	**415**
4:00 p.m. – 11:59 p.m.	270	230	105	
12:00 a.m. – 7:59 a.m.		175	95	**460**
Total	**710**		**245**	**1480**

(10) What is the **total** number of cases for the **4:00 p.m. – 11:59 p.m.** shift?

(11) What is the number of **stable** cases for the **12:00 a.m. – 7:59 a.m.** shift?

(12) What is the total number of **serious cases** across **all shifts**?

A nursing supervisor ranks the shifts based on two different criteria. Which shift received the highest percentage of the total critical cases? For **Questions 13-15**, find the percentage of critical cases for each shift. Round to the nearest one percent.

Shift	**Percentage of Total Critical Cases**
8:00 a.m. – 3:59 p.m.	(13)
4:00 p.m. – 11:59 p.m.	(14)
12:00 a.m. – 7:59 a.m.	(15)

(16) Which shift received the highest percentage of the **total critical cases**? Rank the shifts from highest to lowest.

Now let's explore which shift has the highest percentage of critical cases compared to that shift's total cases. Round to the nearest one percent.

(17) 8:00 a.m. – 3:59 p.m.

(18) 4:00 p.m. – 11:59 p.m.

(19) 12:00 a.m. – 7:59 a.m.

(20) Rank the shifts from highest to lowest.

(21) Which of the following conclusions could be drawn?

(i) The highest number and the highest ratio of experienced nurses should be scheduled during the 12:00 a.m. to 7:59 a.m. shift.

(ii) The highest number of experienced nurses should be scheduled during the 4:00 p.m. to 11:59 p.m. shift. The highest ratio of experienced nurses should be scheduled during the 8:00 a.m. to 3:59 p.m. shift.

(iii) The highest number and the highest ratio of experienced nurses should be scheduled during the 4:00 p.m. to 11:59 p.m. shift.

(iv) The highest number of experienced nurses should be scheduled during the 12:00 a.m. to 7:59 a.m. shift. The highest ratio of experienced nurses should be scheduled during the 4:00 p.m. to 11:59 p.m. shift.

(22) Given the following statistics, rank the risk of having a child with diabetes with the highest risk on top and lowest risk on the bottom.

In general, if you are a man with type-1 diabetes, the odds of your child getting diabetes are 1 in 17. If you are a woman with type-1 diabetes and your child was born before you were 25, your child's risk is 1 in 25; if your child was born after you turned 25, your child's risk is 1 in 100.

Rank the risk of having a child with diabetes from 1 (highest risk) to 3 (lowest risk).

___ Mother with type-1 diabetes who gave birth after turning 25

___ Father with type-1 diabetes

___ Mother with type-1 diabetes who gave birth before turning 25

Mark each statement true or false.

(23) A man with type-1 diabetes has a child. The probability that the child will also have diabetes is about 6%.

(i) True (ii) False

(24) A 20 year-old woman with type-1 diabetes has a child. The probability that the child will also have diabetes is about 25%.

(i) True (ii) False

(25) A 30-year-old woman with type-1 diabetes has a child. The probability that the child will also have diabetes is about 1%.

(i) True (ii) False

This page titled 4.8.1: Practice Problems Corequisite N.8 and N.9 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

5: Corequisite Module M

5.1: Corequisite- Graphing Points, Linear Equations and Models

5.1.1: Practice Problems Corequisite M.1

5.2: Corequisite- Solving Linear Equations, Slope

5.2.1: Practice Problems Corequisite M.2

5.3: Corequisite- Trendlines

5.3.1: Practice Problems Corequisite M.3

5.4: Corequisite- Linear Models Practice

5.4.1: Practice Problems Corequisite M.4

5.5: Corequisite- Linear vs Exponential Models

5.5.1: Practice Problems Corequisite M.5

5.6: Corequisite- Exponential Models (Practice and Application)

5.6.1: Practice Problems Corequisite M.6

5.7: Corequisite- Multivariable Models

5.7.1: Practice Problems Corequisite M.7

This page titled 5: Corequisite Module M is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.1: Corequisite- Graphing Points, Linear Equations and Models

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- addition and subtraction are inverse operations.
- multiplication and division are inverse operations.
- solving for a variable includes isolating it by "undoing" the actions to it.
- multivariable models have one output variable and more than one input variable.

By the end of this lesson, you should be able to

- solve for a variable in a linear equation.
- explicitly write out the order of operations to evaluate a given equation.

PROBLEM SITUATION: CALCULATING BLOOD ALCOHOL CONTENT

Blood alcohol content (BAC) is a measurement of how much alcohol is in someone's blood. It is usually measured as a percentage. So, a BAC of 0.3% is three-tenths of 1%. That is, there are 3 grams of alcohol for every 1,000 grams of blood. A BAC of 0.05% impairs reasoning and the ability to concentrate. A BAC of 0.30% can lead to shortness of breath, a blackout, and loss of bladder control. In most states, the legal limit for driving is a BAC of 0.08%.[1]

BAC is usually determined by a Breathalyzer, urinalysis, or blood test. However, the Swedish physician E.M.P. Widmark developed the following equation for estimating an individual's BAC.[2] This formula is widely used by forensic scientists:

$$B = -0.015 \cdot t + \left(\frac{2.84 \cdot N}{W \cdot r} \right)$$

B = percentage of BAC

N = number of "standard drinks" (A standard drink is one 12-ounce beer, one 5-ounce glass of wine, or one 1.5-ounce shot of liquor.) N should be at least one.

W = weight in pounds

r = the distribution rate for alcohol through the body (this value is a constant), 0.68 for males

and 0.55 for females

t = number of hours since the first drink

(1) Looking at the equation, discuss why the items (t, N, W, and r) on the right of the equation make sense in calculating BAC.

(2) Consider the case of a male student who has three beers and weighs 120 pounds. Simplify the equation as much as possible for this case. What variables are still unknown in the equation? Round values (when necessary) to the nearest thousandth.

(3) (a) Using your simplified equation from Question 2, find the estimated BAC for this student **one** hour, **three** hours, and **five** hours after his first drink.

(b) What patterns do you notice in the data from Question 3 (a)?

(4) How did you arrive at the BAC values mathematically? For example, did you multiply, add, subtract, etc., and what did you do first? Outline the steps that you took to get the answers for Question 3(a).

(5) How long will it take for this student's BAC to be 0.08, the legal limit? How long will it take for the alcohol to be completely metabolized resulting in a BAC of 0.0? Round to one decimal place.

ALGEBRAIC EXPRESSIONS

In algebra, a **term** is an individual part of an algebraic expression. Terms are separated by addition (+) or subtraction (–) signs, and can consist of numbers, variables (letters), or the product of numbers and one or more variables. In instances where a number and variables are being multiplied, the number is called a **coefficient**. For example, the expression below has three terms as shown in boxes. Notice that the last term is (–5). In the original expression, this was written as *minus* 5, but this can be rewritten as *adding negative* 5 as shown below. When breaking an expression into terms, you ask, *what is being added*?

$$-a^3 + 2b - 5$$

$$\boxed{-a^3} + \boxed{2b} + \boxed{-5}$$

Terms	Coefficients
–a3	–1
2b	2
–5	Does not have a coefficient because there is no variable. This is called a **constant term** because it never changes.

(6) State the number of terms in each expression:

(a) 3x + 4

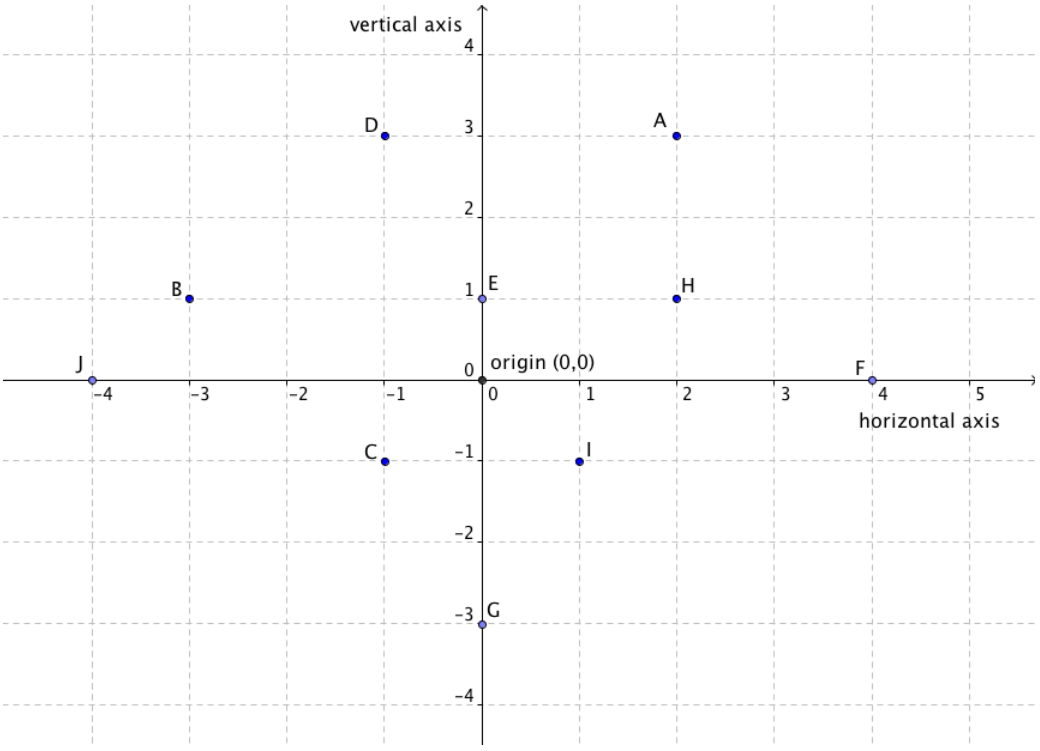
(b) 5x – 4x2 + 2

(c) 5

(7) What is the coefficient of the x^2 term in the expression $5x - 4x^2 + 2$?

GRAPHING REVIEW

This section will help you review some graphing skills that will be needed in many future lessons. In upcoming lessons, you will make graphs on a **coordinate plane** like the one shown below. You will learn about using a coordinate plane and practice how to graph points.



We will begin with some vocabulary. A coordinate plane has two axes that measure distance in two dimensions. The **horizontal axis** goes from left to right. In previous classes, you may have called this the **x-axis**. The **vertical axis** goes up and down. This is sometimes called the **y-axis**. The axes are two number lines that create a grid on the coordinate plane. (**Note**: Axis is singular and axes is plural.)

The point at which the two axes intersect or cross is called the **origin**. This point represents 0 for both axes. To the left of this point, the horizontal axis is negative; to the right it is positive. Below the origin, the vertical axis is negative; above the origin it is

positive. You can see this in the numbers along each axis above. These numbers are called the **scale**.

Each location or *point* on the coordinate plane is defined by an **ordered pair**. You can think of this as the address of a point. An ordered pair must contain two numbers. Ordered pairs are written in a set of parentheses ( ) with a comma separating the numbers, such as (2, 4). The first number is the distance and direction going left or right from the origin and the second number is the distance and direction going up or down. The ordered pair for the origin is (0, 0).

Follow these steps to find the point represented by the ordered pair (2, 3):

1. First, think about the "address" of the point. If this were a street address, the ordered pair tells you to walk 2 blocks horizontally in the positive direction (right) and then walk 3 units vertically in the positive direction (up).
2. Start at the origin. Go 2 units to the right because this is the positive side of the horizontal axis.
3. Go 3 units up.

Point A on the graph above is the point (2, 3). A few other examples from the graph are given below:

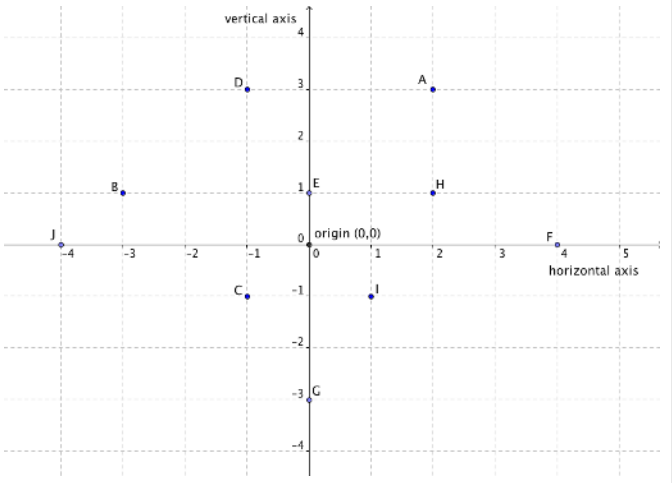
Point B: (–3, 1)

Point E: (0, 1)

Point F: (4, 0)

(8) Write the ordered pairs for the following points on the graph.

Point C:
Point D:
Point G:
Point H:
Point I:
Point J:



[1] http://en.Wikipedia.org/wiki/Blood_alcohol_content

[2] http://www.avvo.com/legal-guides/ugc/explaining-blood-alcohol-levels-bac-what-does-the-08-bac-mean-how-many-drinks-is-that-1

This page titled 5.1: Corequisite- Graphing Points, Linear Equations and Models is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.1.1: Practice Problems Corequisite M.1

(1) Use the Widmark Equation to solve this question. Recall that r = 0.68 for men and r = 0.55 for women.

$$B = -0.015 \cdot t + \left(\frac{2.84 \cdot N}{W \cdot r} \right)$$

A male student had five glasses of wine at a party. He weighs 145 pounds. How long will it take before his BAC is 0.06?

State the number of terms in each expression:

(2) $3x^2+9$

(3) 4

(4) $10x-6x^2-3$

(5) What is the coefficient of the x^2 term in the expression:

$$-x^2 - 8x - 8$$

(6) Which expressions are equivalent to the expression below? There may be more than one correct answer.

$$2(5x + 4) - 3x + 1$$

(i) 10x + 8 – 3x + 1

(ii) 18x – 3x + 1

(iii) 15x + 1

(iv) 7x + 9

Simplify the following:

(7) 3^2

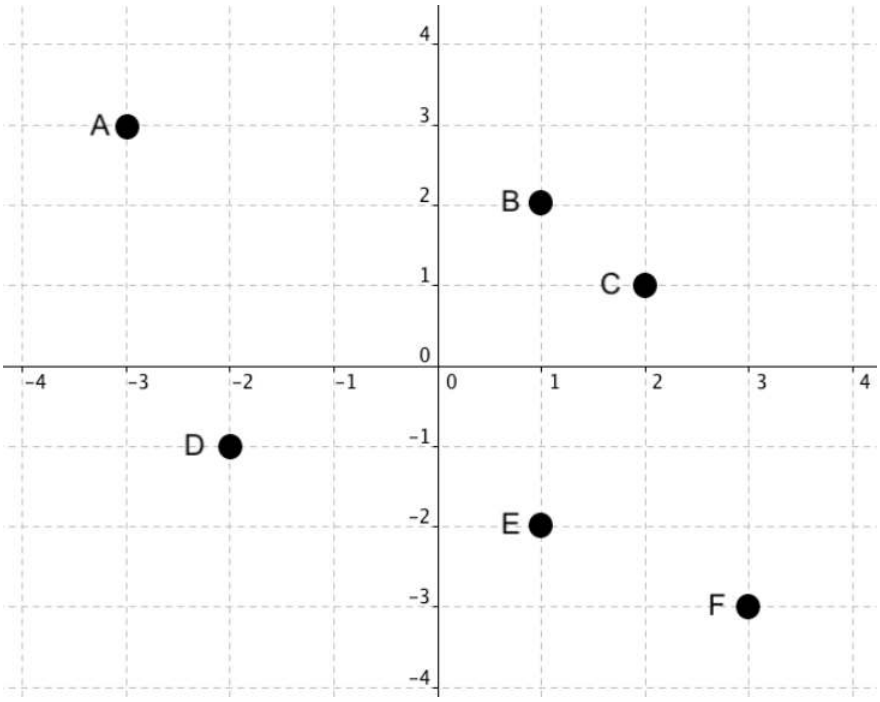
(8) -9^2

(9) $(-3)^2$

(10) -6^2

(11) $-(-10)^2$

Write the ordered pair for the following points on the graph.



(12) Point A:

(13) Point B:

(14) Point C:

(15) Point D:

(16) Point E:

(17) Point F:

This page titled 5.1.1: Practice Problems Corequisite M.1 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.2: Corequisite- Solving Linear Equations, Slope

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- linear models are appropriate when the situation has a constant increase/decrease.
- slope is the rate of change.
- the rate of change (slope) has units in context.
- different representations of a linear model can be used interchangeably.

By the end of this lesson, you should be able to

- label units on variables used in a linear model.
- make a linear model when given data or information in context.
- make a graphical representation of a linear model.
- make a table of values based on a linear relationship.
- identify and interpret the vertical intercept in context.

REPRESENTATIONS OF MODELS

In Module M, you will learn about how different models (such as equations in context) can be useful in examining some situations encountered in real life. A model is a mathematical description of an authentic situation. You can also say that the mathematical description "models" the situation. You will practice using four representations to express the models for situations. We will start the module with the more familiar linear models and then expand to other models.

Verbal Description (Words)

A verbal description explains the relationship in words. Some relationships are difficult to put into words, while in other cases a verbal description can help you make sense of what the relationship means in the context.

Numerical Representation (Table)

Another way that you could represent a relationship is in a table. Tables are helpful for recognizing patterns and general relationships or for giving information about specific values. A table should always have labels for each column. The labels should also include units when appropriate.

Graphical Representation (Graph)

A graph provides a visual representation of the situation. It helps you see how the variables are related to each other. It also helps you to make predictions about future values as well as values in between those in your table. The horizontal and vertical axis of the graph should be labeled, including units.

Algebraic Model (Equation)

You will consider how to write a mathematical equation for a relationship. In writing equations, it is always important to define what the variables represent, including units.

Summary

Throughout this course, you have learned that having the skill to move between different forms and tools is important in problem solving. Alternating among the four representations of mathematical relationships is another example of this. In some cases, you may struggle with writing an equation, but find it helpful to start with a table. You might want a graph for a visual representation, but also need to express a relationship in words. It is important that you can translate one form into another and that you can choose which form is most useful in a specific situation.

PROBLEM SITUATION: DAILY LATTE

A local coffee shop offers a coffee card that you can preload with any amount of money and use like a debit card each day to purchase coffee. At the beginning of the month (when you get your paycheck), you load it with $50. Your favorite small soy latte costs $2.63.

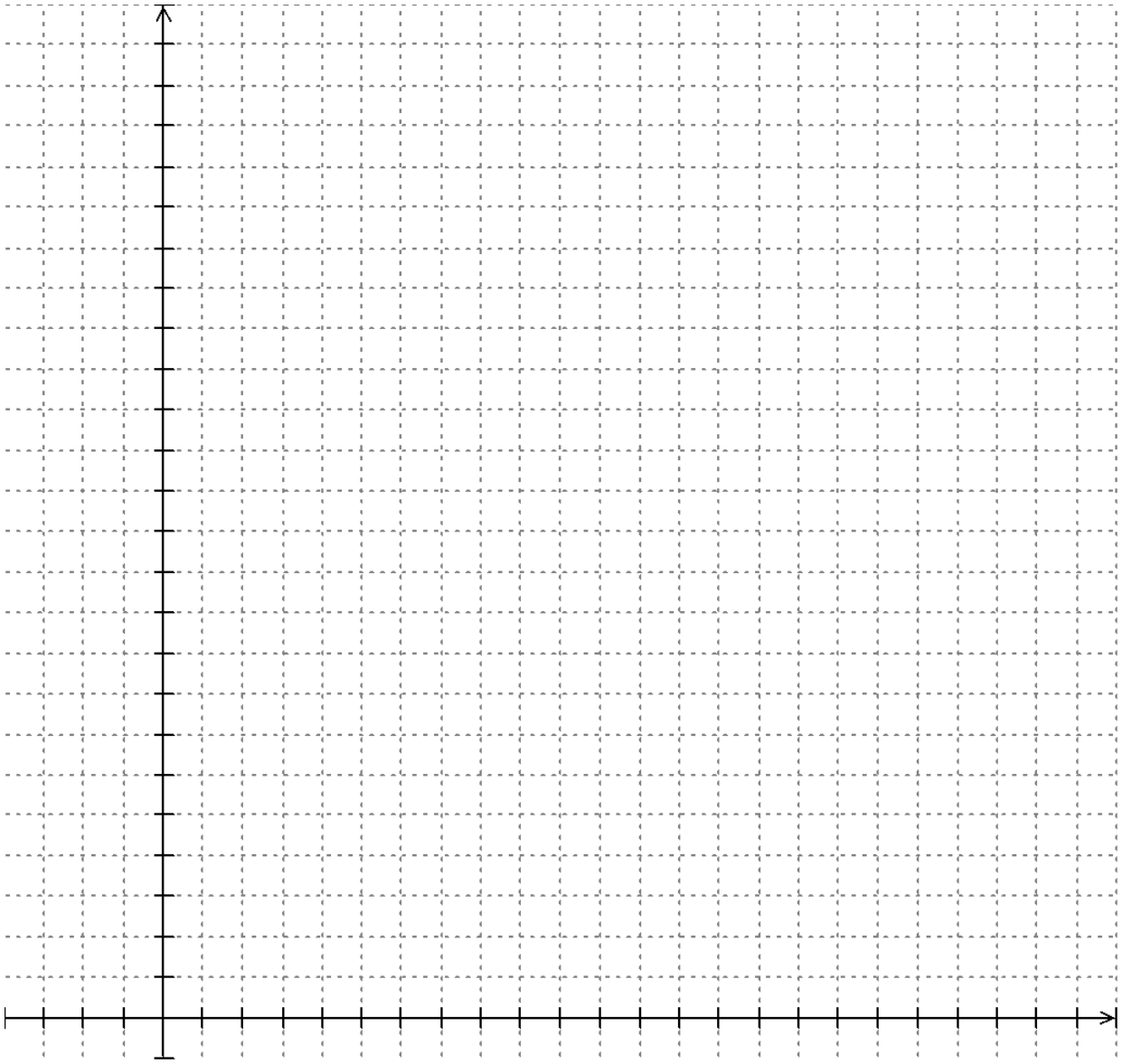
(1) Pretend you purchase a latte every weekday. *Estimate* whether the amount of money you loaded onto your coffee card will last until the end of the month.

(2) Fill in the table and graph below to help you answer the question: Will your $50 coffee card last until the end of the month? (**Note**: Any given month has about 22 weekdays; there could be as few as 20 weekdays in February, or as many as 23 in some other months.)

(a) Fill in the table.

Lattes (n)	**Amount Remaining on Card (A)**
0	50
1	50 – 2.63
2	
3	
4	
10	
15	
20	
23	

(b) Graph the data from the table in (a). Be sure to include scales and labels on the axes. Let the horizontal axis be the number of lattes purchased and the vertical axis be the amount of money remaining on the card. **Note**: If completing this problem online, follow the instructions given online to create your graph.



(c) Build a linear model. Use A to represent the remaining money on the coffee card and n to represent the number of lattes purchased.

(d) Answer the following questions:

(i) Will the card last the entire month?

(ii) If so, how much money is left at the end of the month?

(iii) If not, how many days will it last?

The two mathematical relationships in the plans are linear. They have certain important characteristics. The following terms are vocabulary words you need to know to talk about linear models and other types of mathematical relationships. You will discuss these in your group. Make sure you take good notes about what each means.

Constant rate of change: This is the characteristic that defines a linear relationship. A constant change can be seen both with the table and the graph.

Slope: This is a ratio that describes the rate of change. Be sure you understand the slope and its relation to the graph as the rise (vertical movement) of the line over the run (horizontal movement). The concept of slope is the rate of change of the quantity represented. The slope is multiplied by the (input) variable because you use it as a rate to find a new value.

Intercepts: The point at which a graph touches one of the axes. There is a vertical intercept and a horizontal intercept.

Vertical intercept: The vertical intercept is sometimes called the starting value. It is the point at which the graph touches the vertical axis. Note that this often has important meanings in context.

Horizontal intercept: The point at which the graph intersects the horizontal axis, where the value of the output is zero.

Linear Model: **A = b + mt**. This can also be written as $A = mt + b$. In either case, A is the amount of the quantity at any time t; m is the constant rate of change (the slope); and b is the initial amount (when t = 0).

(3) (a) What is the rate of change for the model above in Problem Situation: Daily Latte? What does it mean in the context of the problem situation? Remember to include units.

(b) What is the vertical intercept for the model above in Problem Situation: Daily Latte? What does it mean in the context of the problem situation?

(c) Does the model above in Problem Situation: Daily Latte have a horizontal intercept? If so, what is it and what does it mean in the context of the problem situation?

This page titled 5.2: Corequisite- Solving Linear Equations, Slope is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.2.1: Practice Problems Corequisite M.2

Terrence is very careful about tracking his gas mileage. Every time he fills his gas tank, he records how much gas he buys and the number of miles he has driven. He puts this information into a spreadsheet so he can easily calculate his gas mileage in miles per gallon.

	A	B
1	Number of Miles	Gallons of Gas
2	255	13.2
3	136	5.9
4	297	13.5
5	114	5.3

(1) Select the formula that would calculate Terrence's gas mileage.

(i) = (A2 + A3 + A4 + A5) / (B2 + B3 + B4 + B5)

(ii) = A2 + A3 + A4 + A5 / B2 + B3 + B4 + B5

(iii) = (B2 + B3 + B4 + B5) / (A2 + A3 + A4 + A5)

(iv) = B2 + B3 + B4 + B5 / A2 + A3 + A4 + A5

(2) Terrence is planning a long road trip of about 1,000 miles. The average price of gas is $3.85/gallon. Based on the data in the spreadsheet, estimate how much he should budget for gas. Round to the nearest dollar.

Lenders such as banks, credit unions, and mortgage companies make loans. The person receiving the loan usually pays the loan off in small payments over a long period of time. The lender earns money by charging **interest**, which is based on a percentage of the amount that is borrowed. There are different types of interest. Car loans are usually calculated using the formula for **simple interest**. The total amount repaid is based on the interest and the value of the original loan, called the **principal**. The formula for the total dollars needed to repay the loan, with interest, is found using the formula

$A = P + P \cdot r \cdot t$	A is the *amount* (total principal plus interest) required to repay the loan. P is the *amount* borrowed, the principal. r is the *annual interest rate*, quoted as a percent, but used as a decimal in the formula. t is the *time*, in years, taken to repay the loan (six months would be 1/2 year).

Suppose you get a loan of $5,000 at an annual interest rate of 4.25%.

(3) Use the given information to write the formula for the total amount to be repaid in t years.

Answer **Questions 4 - 14** to complete the table of values.

	t (years)	**A ($)**
0	(4)	(9)
4 months	$\frac{1}{3}$	(10)
6 months	(5)	(11)
1 year	(6)	(12)
3 years	(7)	(13)
6 years	(8)	(14)

(15) Estimate the maximum time allowed to repay the loan if you want the total payoff to be less than $7,000.

Indicate if each of the following statements is true or false.

(16) The slope of a linear model is a ratio that describes how the values represented by the model increase or decrease.

(i) True (ii) False

(17) The slope of a linear model changes depending on the values substituted into the model.

(i) True (ii) False

(18) The horizontal intercept is represented by an ordered pair in which the first value is zero: (0, __).

(i) True (ii) False

(19) The vertical intercept is sometimes called a starting or initial value.

(i) True (ii) False

This page titled 5.2.1: Practice Problems Corequisite M.2 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.3: Corequisite- Trendlines

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- linear equations can approximate nearly linear data.

By the end of this lesson, you should be able to

- find the equation of a line that estimates nearly linear data by calculating the rate of change and estimating the vertical intercept of the line.
- use approximate linear models to interpolate and extrapolate.

PROBLEM SITUATION: LIFE EXPECTANCY AND SOCIAL SECURITY

This lesson will focus on the Social Security program. Social Security provides income for people who are elderly or disabled. The following information is from the Social Security Administration website:[3]

> Social Security is financed through a dedicated payroll tax. Employers and employees each pay 6.2 percent of wages up to the taxable maximum of $160,200 (in 2023), while the self-employed pay 12.4 percent. When a person reaches “full retirement age,” he or she can claim a full retirement benefit. If someone retires before full retirement age, he or she receives a lower benefit.

(1) Use the following information from the Social Security Administration website to answer the questions below.[4]

Full retirement age is the age at which a person may first become entitled to full or unreduced retirement benefits. No matter what your full retirement age (also called "normal retirement age") is, you may start receiving benefits as early as age 62 or as late as age 70.

If You Retire Early

You can retire at any time between age 62 and full retirement age. However, if you start benefits early, your benefits are reduced a fraction of a percent for each month before your full retirement age.

The chart below lists age 62 reduction amounts and includes examples based on an estimated monthly benefit of $1000 at full retirement age.

Full Retirement and Age 62 Benefit By Year Of Birth

Year of Birth	Full (normal) Retirement Age	Months between age 62 and full retirement age		A $1000 retirement benefit would be reduced to	The retirement benefit is reduced by	A $500 spouse’s benefit would be reduced to	The spouse’s benefit is reduced by
1937 or earlier	65	36		$800	20.00%	$375	25.00%
1938	65 and 2 months	38		$791	20.83%	$370	25.83%
1939	65 and 4 months	40		$783	21.67%	$366	26.67%
1940	65 and 6 months	42		$775	22.50%	$362	27.50%
1941	65 and 8 months	44		$766	23.33%	$358	28.33%
1942	65 and 10 months	46		$758	24.17%	$354	29.17%
1943-1954	66	48		$750	25.00%	$350	30.00%

1955	66 and 2 months	50		$741	25.83%	$345	30.83%
1956	66 and 4 months	52		$733	26.67%	$341	31.67%
1957	66 and 6 months	54		$725	27.50%	$337	32.50%
1958	66 and 8 months	56		$716	28.33%	$333	33.33%
1959	66 and 10 months	58		$708	29.17%	$329	34.17%
1960 and later	67	60		$700	30.00%	$325	35.00%

(a) Who can collect full retirement benefits at age 65?

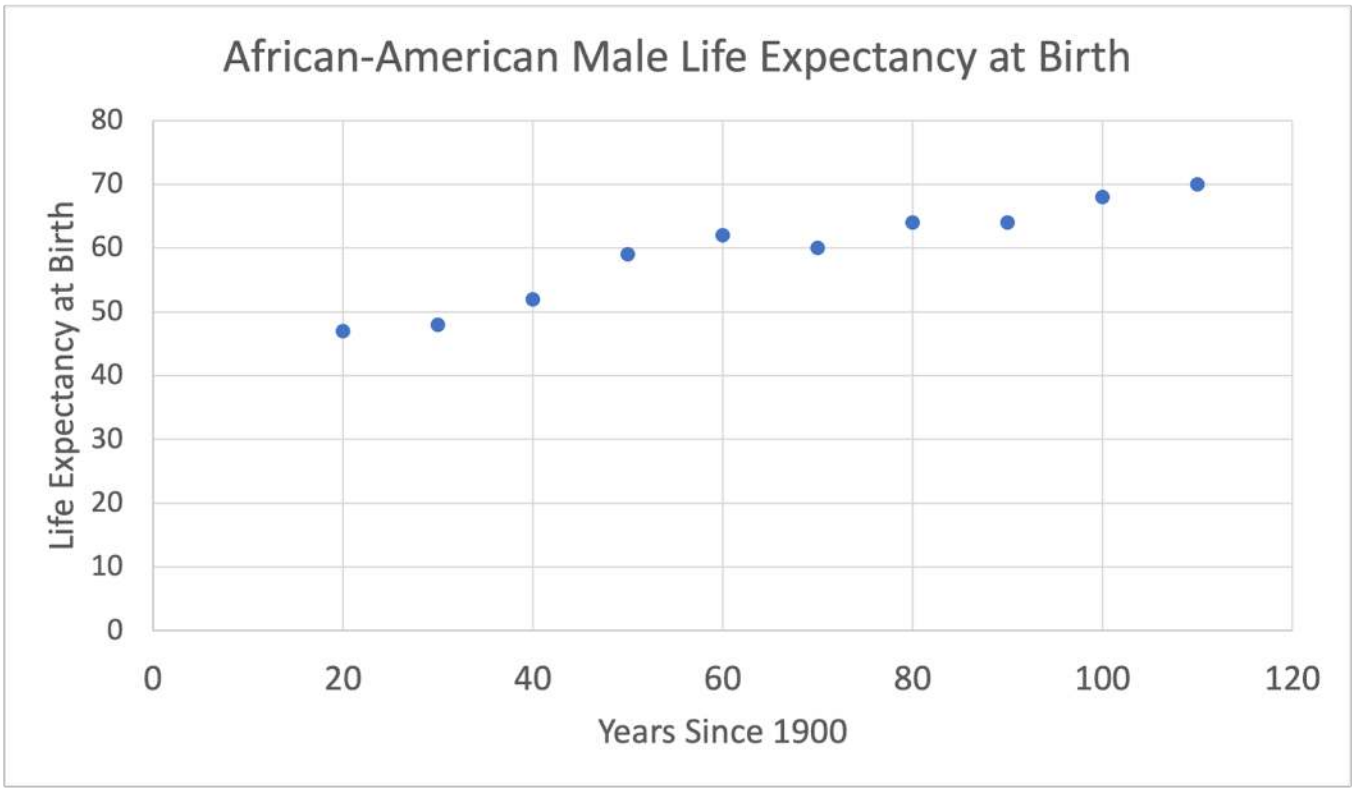
(b) At what age can someone born in 1950 collect full benefits?

(c) What is the youngest age at which you can begin collecting reduced benefits?

One proposal to reform Social Security is to increase the retirement age. This, however, raises concerns about fairness. No matter what the retirement age is, some people will pay into Social Security but die before retirement and never receive a benefit. This happens to more people when the retirement age is increased. In this lesson, you will examine the effects of raising the retirement age to 75. Specifically, you will answer the question of whether this change would have a greater impact on some groups than others.

To explore this question, you will use life expectancy data from the Centers for Disease Control. Real data rarely fall on a straight line, but sometimes data show a definite trend. If the trend is close to linear, the data can be *approximated by a linear model*. This means that a linear model gives good estimates of what the data will be if the trend continues. A model can also be used to estimate values between data points. In this lesson, you will learn to create linear models from data.

The following data show the life expectancy of African American males in the United States at birth.



(2) Find a linear model to approximate these data, letting L be the life expectancy at birth and y the year of birth measured in years after 1900.

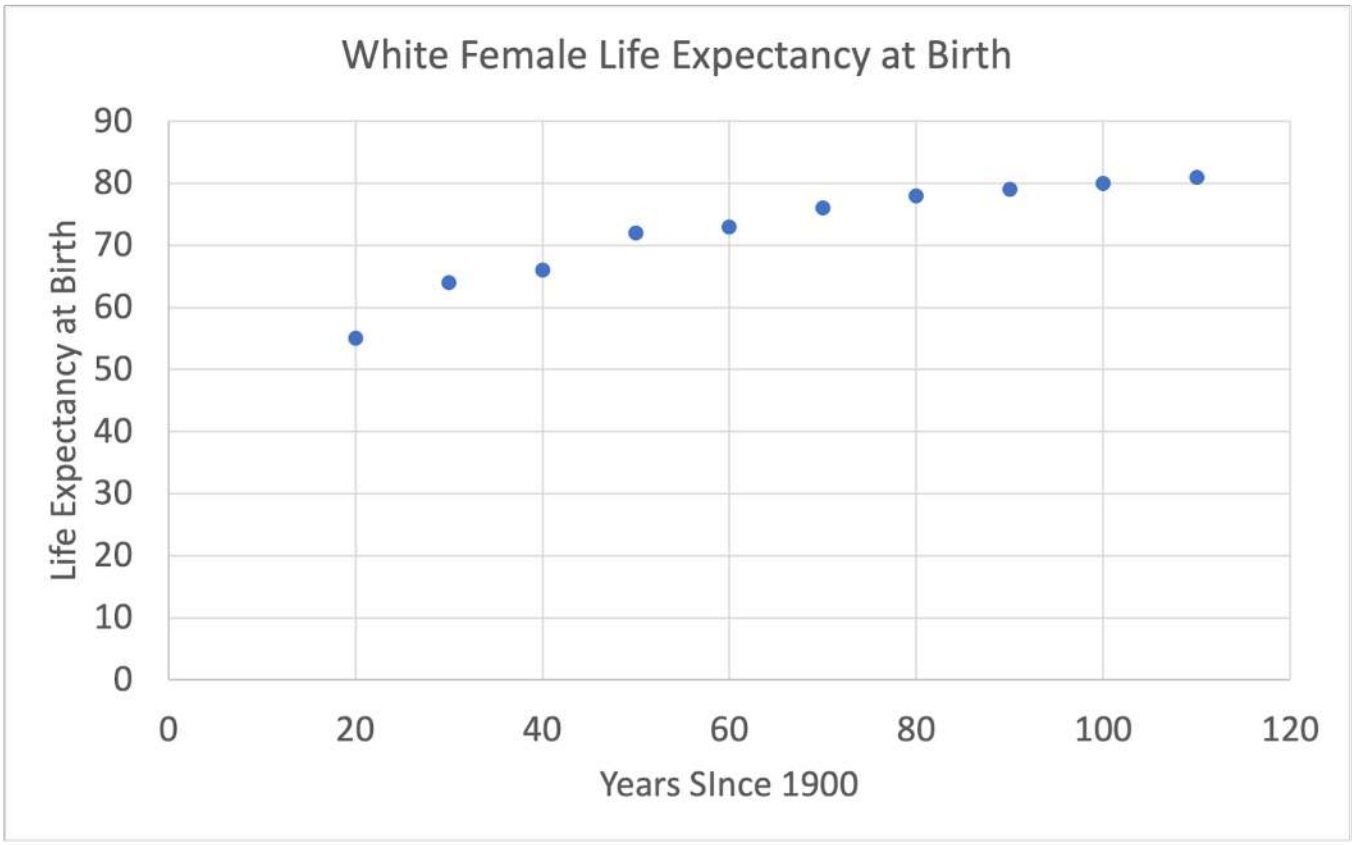
(3) In what year will African American male babies have a life expectancy of 75 years? In what year will these male babies be eligible to begin drawing Social Security if the retirement age is raised to 75?

Babies born in _____ will have an average lifespan of 75 years.

These babies will start drawing Social Security in: _____

(4) The last part of this lesson uses data about the life expectancy of a different population of Americans: white females.

Find a linear model to approximate your dataset, letting L be the life expectancy at birth and y the year of birth measured in years after 1900. Be prepared to develop the line you feel best represents the data graphically and to discuss how you found the equation of this line.



(5) When does your model predict the population will first have a life expectancy of 75 years at birth? In what year will this group begin collecting Social Security if the retirement age is raised to 75?

(6) When does your model predict the population will first have a life expectancy of 90 years at birth?

---

[3] https://www.ssa.gov/news/press/factsheets/HowAreSocialSecurity.htm

[4] http://www.socialsecurity.gov/retire2/agereduction.htm#chart

This page titled 5.3: Corequisite- Trendlines is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.3.1: Practice Problems Corequisite M.3

The table below shows data on the prevalence or rate of smoking among adults in the United States over time.[5]

Year	Prevalence of Smoking Among Adults in the United States (as a percentage)
1965	42.40
1970	37.40
1980	33.20
1990	25.50
2000	23.30
2005	20.90
2010	19.30
2013	17.80
2020	12.50

(1) Graph the data on the grid below. Add appropriate labels.



(2) Create a linear equation for the data. Let P = the percentage of adults who smoke and t = the number of years after 1965. Use the two points below to find your model.

Point 1: Year 1970, Percentage 37

Point 2: Year 2013, Percentage 18

Round your vertical intercept to the nearest whole number, and your slope to two decimal places.

(3) Describe the situation modeled by your equation. This would be a verbal description of the model (so it should match the graph and the equation). Be sure to include an interpretation of the starting value and the rate of change.

(4) Do you think your model would accurately predict the rate of smoking in 2060? Explain your answer.

(5) Based on these data, can you say that fewer adults in the United States smoked in 2010 than in 2005?

5 www.infoplease.com/ipa/A0762370.html#ixzz1bS71bHKR and https://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/index.htm

This page titled 5.3.1: Practice Problems Corequisite M.3 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.4: Corequisite- Linear Models Practice

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- linear models are appropriate when the situation has a constant increase/decrease.
- slope is the rate of change.
- the rate of change (slope) has units in context.
- different representations of a linear model can be used interchangeably.

By the end of this lesson, you should be able to

- label units on variables used in a linear model.
- make a linear model when given data or information in context.
- make a graphical representation of a linear model.
- make a table of values based on a linear relationship.
- identify and interpret the vertical intercept in context.

PROBLEM SITUATION: DATA PLANS

Data plans for cell phone plans can be purchased in two different ways. One is an unlimited data plan where the customer pays a set monthly fee for unlimited data. The other is a per-gigabyte (GB) pricing service where the customer pays a set monthly fee plus a specified amount for each GB. The GBs are not prorated (that means it doesn't matter if the customer uses 0.001 GB or 1 GB, they must still pay for the entire 1 GB). The cost of the phone itself and other fees are the same amount under each plan.

You are shopping for a data plan and need to decide which option costs less. (**Note**: The descriptions of these options are examples of verbal representations of mathematical relationships.)

- Per-GB Pricing: There is a monthly fee of $20 plus $10 per GB.
- Unlimited Data: The unlimited plan costs $60 per month.

(1) Which plan do you think is less expensive and why?

(2) (a) Use the table below to build a linear model for the **Per-GB Pricing** data plan. First, create equations for 3, 4, 5, and 10 GBs used, and then create the linear model for g GBs used.

Per-GB Pricing (P)

GBs Used (g)	Cost	Equation
0	20	P = 20
1	20 + 1(10)	P = 20 + 1(10)
2	20 + 1(10) + 1(10)	P = 20 + 2(10)
3		P =
4		P =
5		P =
10		P =
g	$20 + \underbrace{\qquad\qquad}_{\text{t times}}$	P =

(b) Use the table below to build a linear model for the **Unlimited** data plan. First, create equations for 3, 4, 5, and 10 GBs used, and then create the linear model for g GBs used.

Unlimited Data (U)

GBs Used (g)	Cost	Equation
0	60	U = 60
1	60	U =
2		U =
10		U =
g		U =

(c) Use your models to fill in the following table.

GBs Used (g)	Monthly Cost of Per-GB Plan (P)	Monthly Cost of Unlimited Data (U)
0		
1		
2		
3		
4		
5		
6		
7		
8		

(d) Using the table from Q2(c), under what conditions is P less expensive?

(e) Using the table from Q2(c), under what conditions is U less expensive?

(3) The following questions are about “Problem Situation: Data Plans.”

(a) What are the rates of change of each of the two models? What do they mean in the context of the problem situation? (Include units.)

(b) What are the vertical intercepts of each of the two models? What do they mean in the context of the problem situation?

(c) Do the two models have horizontal intercepts? If so, what are they and what do they mean in the context of the problem situation?

(4) Explain how you can tell from a graph if a linear model has a negative slope, a positive slope, or a slope of zero. Include sketches of each type of graph.

This page titled 5.4: Corequisite- Linear Models Practice is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.4.1: Practice Problems Corequisite M.4

A person's Blood Alcohol Content (BAC) after t hours can be approximated by Widmark's equation:

$$B = -0.015 \cdot t + \left(\frac{2.84 \cdot N}{W \cdot g} \right)$$

where N is the number of "standard drinks," W is his/her weight in pounds, and g = gender constant (0.68 for men and 0.55 for women). One standard drink is a 12-ounce beer, 5-ounce glass of wine, or 1.5-ounce shot of liquor.

Paula is part of a research study to verify Widmark's equation. She weighs 130 pounds. The researchers ask her to drink three beers. After one hour, her BAC is at 0.10. After three hours her BAC is at 0.08.

(1) Find an equation of the line using her BAC measurements where B is her BAC after t hours. Round decimals to the hundredths place. **Hint**: The time and BAC can be written as ordered pairs: (1, 0.10) and (3, 0.08). It might help to graph the two points and draw the line between them.

(2) What is the slope of your equation and how would you interpret this value in the context of the situation?

(3) What are the horizontal and vertical intercepts of your equation and how would you interpret these values in this scenario?

(4) Simplify Widmark's equation for Paula's case, leaving t as a variable. Round all decimals to the thousandths place.

(5) What is the slope and vertical intercept for this simplified equation?

(6) Describe the meaning of the slope and vertical intercept in context.

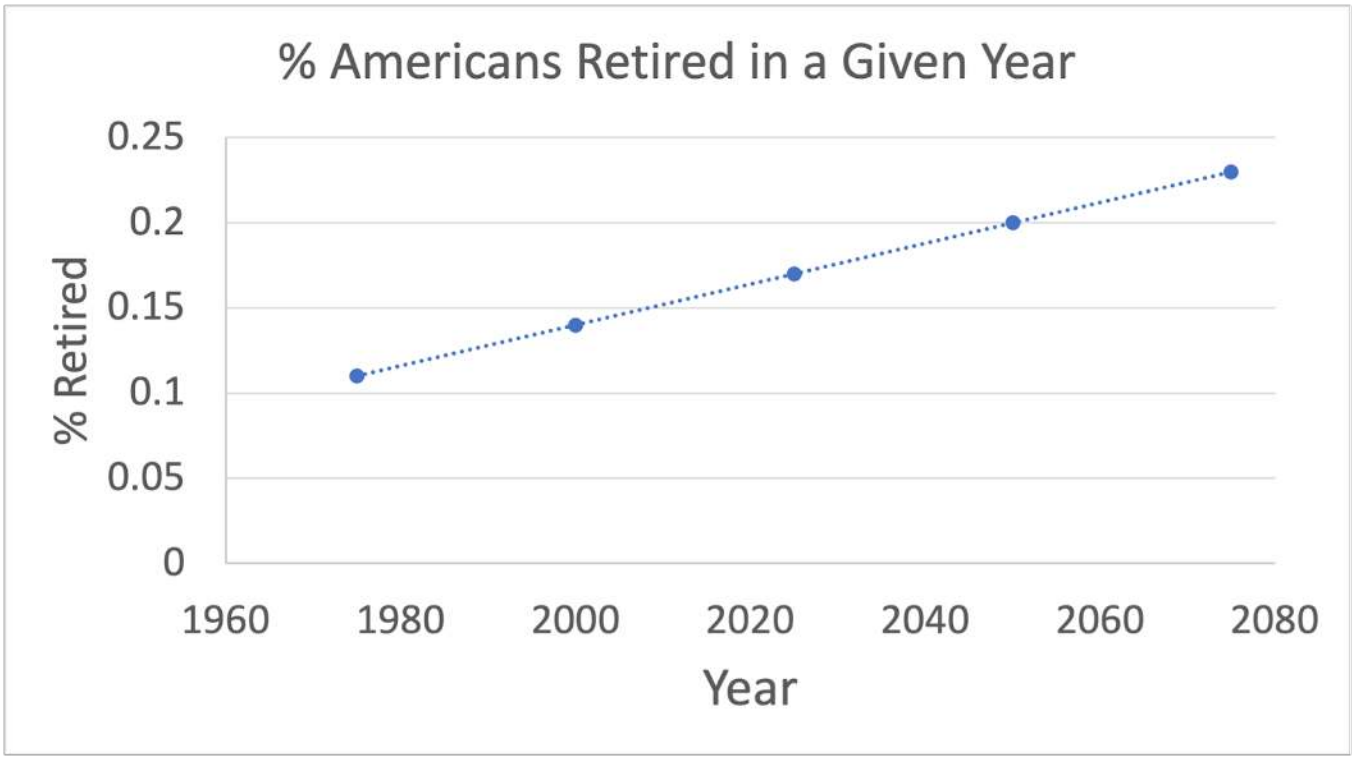
(7) Looking at the models you created in Question 1 and Question 4, compare the slopes of the two equations. Which model has the steepest slope?

(i) Widmark's equation

(ii) Research study model

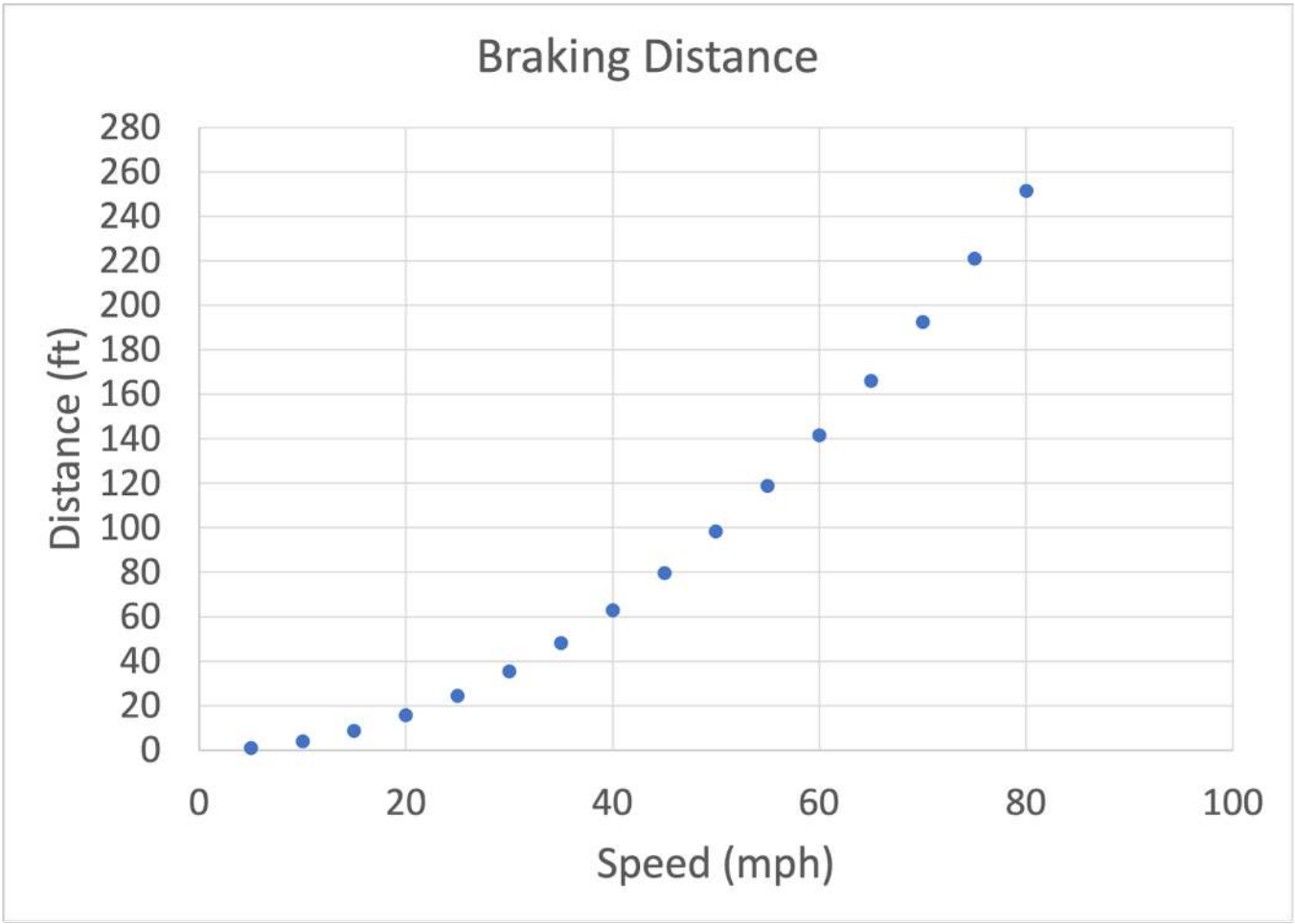
For **Questions 8-12**, do the representations depict a linear model or nonlinear model? Use the graph, equation, or table provided to determine your answer.

(8) $R = \frac{t}{873.36} - 2.15$, where R is the percent of Americans who are retired in the year t. The graph is shown below.



(9) P = 120 + 2L, where P = the perimeter of a rectangular area (in feet), and L = the length of the rectangle (in feet).

(10) The graph for braking distance versus speed for the coefficient of friction of 0.8 and a roadway grade of 5% is shown below.



(11)

Time (min)	Distance (ft)
0	3.5
2	13.5
4	23.5
6	33.5

(12) The formula for the area of a circle: $A = \pi r^2$.

This page titled 5.4.1: Practice Problems Corequisite M.4 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.5: Corequisite- Linear vs Exponential Models

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- using algebra to generalize calculations is a valuable tool.
- compounding is repeated multiplication by a compounding factor.
- compounding is best expressed in terms of exponential growth, using exponential notation.
- exponential growth models the compounding of interest on an initial investment.

By the end of this lesson, you will be able to

- write or interpret an expression that models percentage increase or decrease.
- use expressions and equations to solve contextual problems about percentage increases or decreases.
- calculate the earnings on a principal investment with annual compound interest.
- write a formula for annual compound interest.
- compare and contrast linear and exponential models.

INTRODUCTION

(1) The following question gives you practice with using percentages in expressions and equations. Match the expressions that either increase or decrease the quantity x by a given percentage with the verbal descriptions of the changes. A verbal description may be used more than once.

Answer	**Column A**		**Column B**
	x + 0.45x		a) increase x by 45%
	0.45x		b) increase x by 55%
	x – 0.55x		c) increase x by 145%
	1.45x		d) increase x by 245%
	3.45x		e) increase x by 345%
	2.45x		f) decrease x by 55%
	x + 0.55x		g) decrease x by 45%
	x – 0.45x		h) decrease x by 145%
	0.55x		

PROBLEM SITUATION: THE FIVE-YEAR CD

Suppose you invest $1,000 **principal** (the original deposit in an investment) into a certificate of deposit (CD) with a five-year **term** (the agreed upon period of time that money is in an investment) that pays a 5% annual interest rate. The **compounding period** is one year. Recall that in **compound interest**, interest is added to the starting principal and accrues after set periods of time through the duration of the whole term.

(2) How much money will you have in your account at the end of the five-year term? (Be ready to explain your calculations.) Use the table below to find a pattern and develop a formula to model the total amount accrued in a CD with annual compounding after n years, if the principal = $1,000 and the APR = 5%. Round figures in the Amount Accrued column to two decimal places.

Term	**Calculation**	**Amount Accrued**
1 year	$1,000 + $1,0000.05 = $1,000(1 + 0.05) = $1,000*1.05	$1,050.00
2 years	($1,000*1.05)(1.05) = $1,000(1.05)2	
3 years		

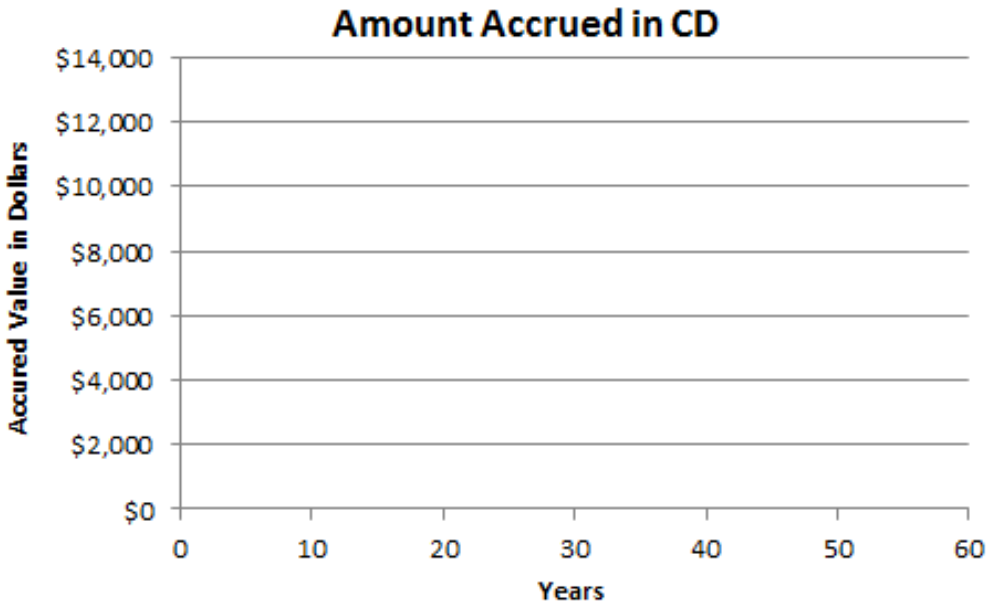
4 years		
5 years		
n years		

(3) Is your formula from the previous question linear? Explain.

(4) Using the formula and patterns you developed, fill out the following table. Round figures in the Amount Accrued column to two decimal places.

Term	Calculation	Amount Accrued
10 years		
20 years		
30 years		
40 years		
50 years		

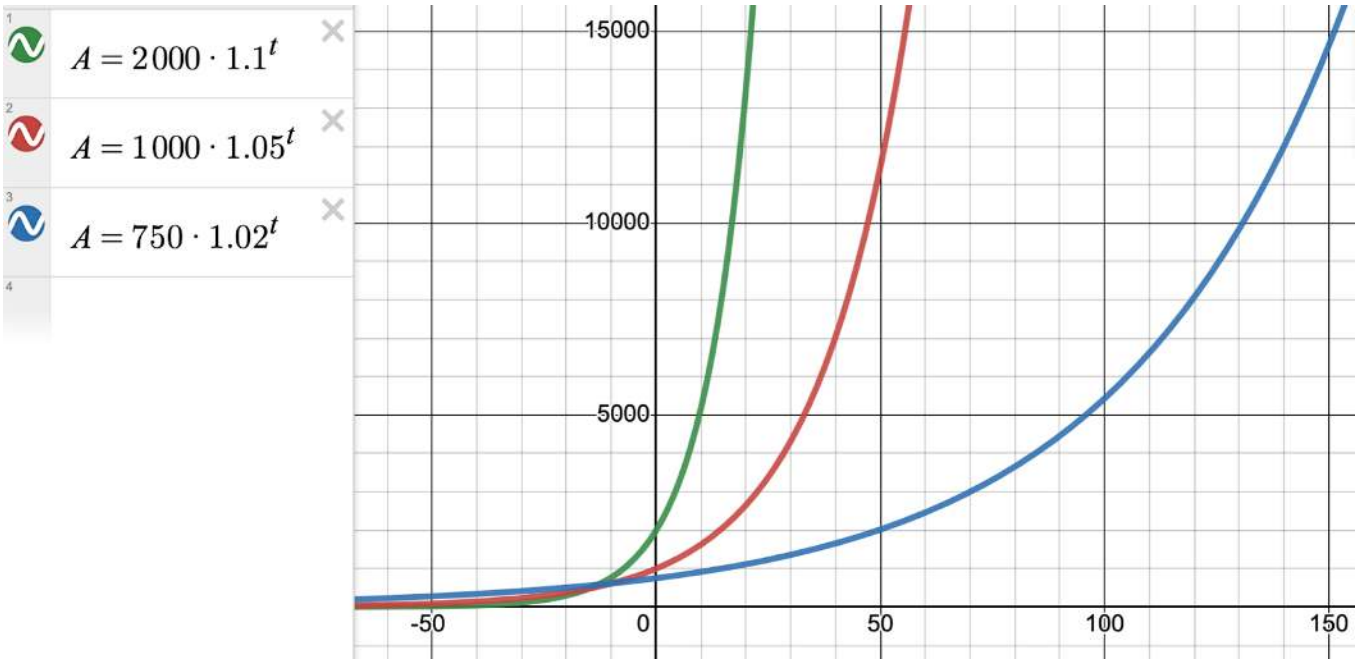
(5) Use the values from the table above to plot a graph of the model you developed. Note: If completing this problem online, follow the instructions given online to create your graph.



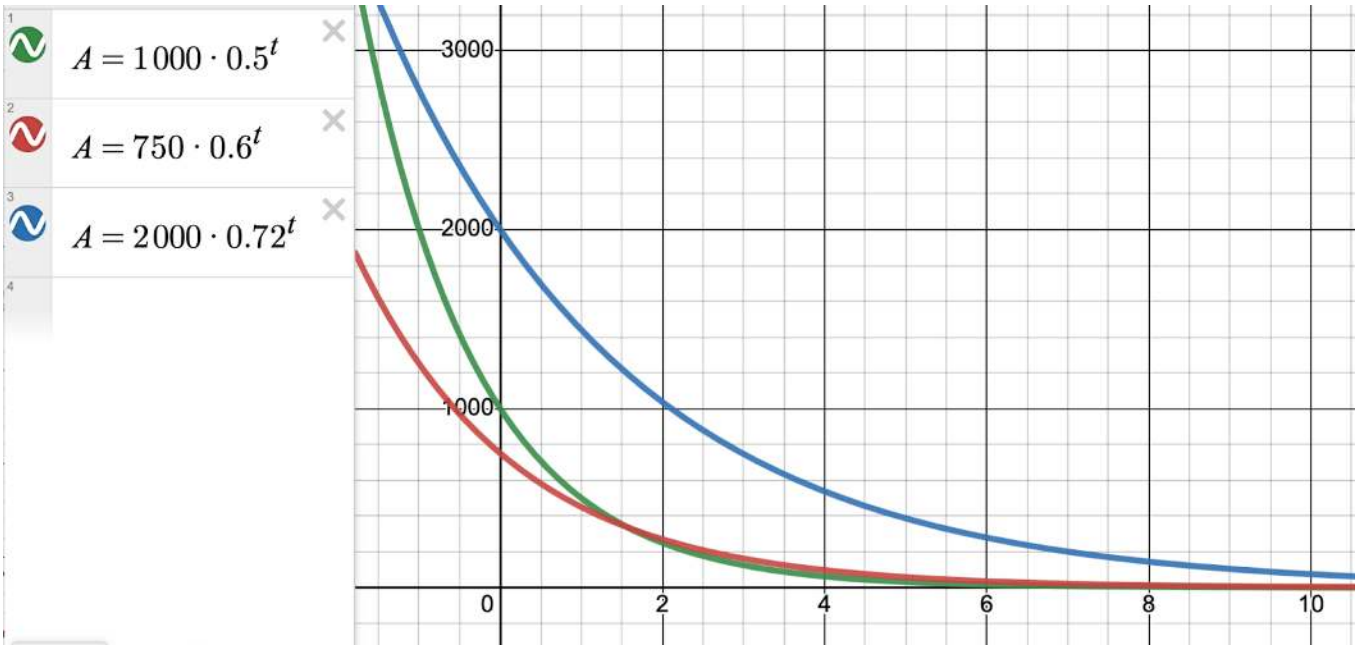
Behavior of Exponential Growth/Decay

Note from the graph, that the growth is very small at first, but then it changes dramatically. Here are some other graphs that illustrate both exponential growth and exponential decay:

Exponential growth (increasing rate of change)



Exponential decay (decreasing rate of change)



(6) Write a general formula that could be used to find the accrued amount (A) for a CD with annual compounding. Let P = the principal, r = the annual interest rate as a decimal, and t = number of years of the investment.

This page titled 5.5: Corequisite- Linear vs Exponential Models is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.5.1: Practice Problems Corequisite M.5

(1) How is a linear model different from an exponential model?

Recall the simple interest formula is:

$$A = P + P \cdot r \cdot t$$

where A is the *Amount* (total principal plus interest) required to repay the loan, P is the *Principal*, r is the *annual interest rate* (quoted as a percent, but used as a decimal), and t is the *time* in years.

Complete **Questions 2-11** in the table below with the amounts you would have if you invested $1,000 at 5% interest with simple interest and with interest compounded annually.

	Simple Interest	**Compound Interest**
1 year	(2)	(7)
2 years	(3)	(8)
3 years	(4)	(9)
4 years	(5)	(10)
5 years	(6)	(11)

(12) Is simple interest a linear model? How do you know?

For the next three questions, we will explore which investment strategy will have a higher balance after 10 years: investing $500 at 10% APR or $1,000 at 2% APR. We will assume both investments have annual compounding.

(13) After 10 years, the $500 at 10% APR would grow to $_____

(14) After 10 years, the $1,000 at 2% APR would grow to $_____

(15) Which investment will have the higher balance after 10 years? Explain your answer clearly using the Writing Principle.

This page titled 5.5.1: Practice Problems Corequisite M.5 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.6: Corequisite- Exponential Models (Practice and Application)

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- solving for a variable includes isolating it by “undoing” the actions to it.
- there are differences and similarities between exponential growth and decay.

By the end of this lesson, you should be able to

- explicitly write out the order of operations to evaluate a given equation.
- use the compound interest formula for different compounding periods.
- write an exponential decay model.

INTRODUCTION

We spent some time at the start of this module discussing how to “undo” operations to solve linear equations. We will also need to “undo” operations to solve the more complicated exponential equations. Let’s review our more familiar linear situation.

(1) Consider the following equation for the values given in (a) and (b). In each case, write out the steps (using words) that you would take to solve for the variable. Remember that we want to apply this process to exponential equations soon. So, think carefully about what you do in each step while solving.

$$y = -4x - 2$$

(a) Solve for y if x = −3. Using words, explain your steps.

(b) Solve for x if y = −3. Using words, explain your steps.

PROBLEM SITUATION 1: THE VALUE OF A CD

In the previous corequisite lesson, you used the following formula to learn how to find the accumulated amount (A) of a certificate of deposit (CD) with a compounding period of one year:

$$A = P(1 + r)^t$$

In this problem situation, you will use the same formula to explore the accumulated amount (A) of a CD with a shorter compounding period. Remember that the variable (P) represents the principal, r represents the interest rate per period, and t represents the number of periods.

(2) (a) Suppose you invest $1,000 in a two-year CD, advertised with an annual percentage rate (APR) of 12%, where compounding occurs monthly. If the APR is 12% per year, how much is it per month?

(b) Use your answer to complete the following table. Record how you found the results in the middle column.

- Principal = $1,000
- APR = 12%
- Term = two years
- Compounding period = one month

Period	Calculation	Amount Accrued
1 month		
2 months		
3 months		
6 months		
12 months		
24 months		

(3) Write a general formula that can be used to calculate the value of any CD with any number of compounding periods in a year. Define your variables. **Hint**: Use the simplified compound formula as a starting point, but recognize the key difference in this current example: instead of interest being compounded once annually, the interest in this example is compounded monthly; that is, it is compounded 12 times a year!

(4) How does the compounding period affect the accumulated amount? **Hint**: Use quarterly compounding to compute the accumulated amount after eight quarters (24 months), and compare to the results in the above table.

PROBLEM SITUATION 2: UNDERSTANDING DEPRECIATION

Depreciation is a process of losing value. For example, new automobiles lose 15% to 20% of their value each year for the first few years you own them.

(5) (a) Based on this fact, develop a depreciated-value (D) formula for a $26,000 automobile. Assume that the depreciation is 15% per year. Use the table below to guide your calculations.

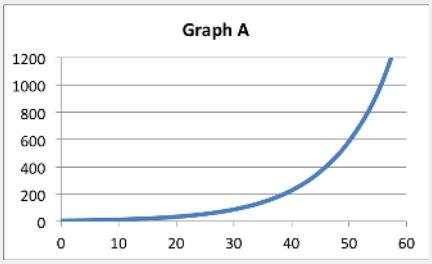
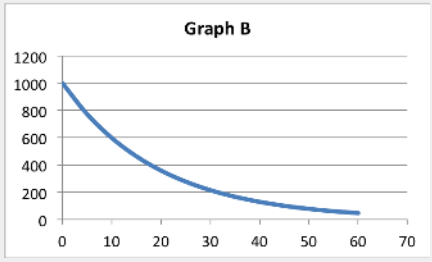
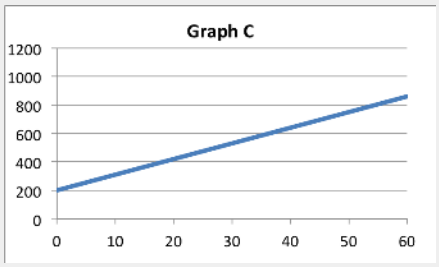
Age of the Automobile	**Calculation**	**Value**
New	N/A	$26,000
1 year old		
2 years old		
5 years old		
t years old		

(b) Use your formula from part (a) to make a graph of the value of the vehicle over time. **Note**: If completing this problem online, follow the instructions given online to create your graph.

This page titled 5.6: Corequisite- Exponential Models (Practice and Application) is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.6.1: Practice Problems Corequisite M.6

(1) Write the letter of each graph next to the equation that best matches it.

_____ $y = 1{,}000(0.95)^x$	_____ $y = 200 + 11x$	_____ $y = 5(1.1)^x$
		

In this problem, you will compare the effects of different compounding periods on the interest an investment earns. Complete the table below using the values indicated. Show the formula you used, with the correct values, in the second column.

- Principal: $1,000
- APR: 4.5%
- Time: 10 years

Complete the table below by answering Questions 2-9.

Compounding Period	**Equation Used for Calculation (with values inserted)**	**Amount Accrued After 10 Years**
Annual	(2)	(3)
Quarterly	(4)	(5)
Monthly	(6)	(7)
Daily	(8)	(9)

Certain drugs are eliminated from the bloodstream at an exponential rate. Doctors and pharmacists need to know how long it takes for a drug to reach a certain level to determine how often patients should take medications. Answer the following questions based on this information.

(10) Select the correct statement:

(i) The same amount of the drug will be eliminated in the first hour as in the second hour.

(ii) More of the drug will be eliminated in the first hour than in the second hour.

(iii) Less of the drug will be eliminated in the first hour than in the second hour.

(11) Write an exponential model for the following situation. The drug dosage is 500 mg. The drug is eliminated at a rate of 5.2% per hour. Use D = the amount of the drug in milligrams and t = time in hours.

(12) How much of the drug is left after six hours? Round to the nearest milligram.

This page titled 5.6.1: Practice Problems Corequisite M.6 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.7: Corequisite- Multivariable Models

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- multivariable models have one output variable and more than one input variable.
- formulas using multiple variables are often used in real-world scenarios.
- a weighted average is a multivariable model which gives a better "average" for some sets of data.
- many problems are just variations of a weighted average problem, such as calculating grades and GPAs.

By the end of this lesson, you should be able to

- calculate the mean of a set of numbers.
- interpret and use formulas with multiple variables.
- calculate the weighted average of a set of numbers in context.
- construct an expression for a weighted average problem.

INTRODUCTION: REVIEW OF MEAN, MEDIAN, AND MODE

The terms ***mean***, ***median***, and ***mode*** are three different kinds of averages.

(1) (a) Match the term to the correct definition.

• Mean • Median • Mode	• The average of a group of numbers. • The number that occurs most frequently. • The middle number in a list of numbers that have been arranged in order.

(b) Imagine a place that has very different daily temperatures. The table below shows the temperatures for one week. What is the mean? What is the median? What is the mode?

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
62°	90°	106°	10°	50°	0°	50°

PROBLEM SITUATION 1: REVISITING COLLEGE RANKING

The Aspen Institute awards the Aspen Prize every other year. The award "honors those institutions that strive for and achieve exceptional levels of success for all students, while they are in college and after they graduate." To determine the 150 finalists for the award, the Aspen Institute collects information from 1000 community colleges related to the following variables (remember that we call them variables because their values may vary between colleges):

1. **Performance** – The **performance** score measures student success in persistence, degrees awarded, completion, and transfer;

2. **Change** – The **change** score measures consistent improvement in these areas over time; and

3. **Minority Achievement and Family Income** – The **minority achievement** score and **family income** score measure equitable outcomes for students of all racial/ethnic and socioeconomic backgrounds.

The Aspen Institute used the following mathematical equation to calculate a score for each school. In this case, the model is represented by this equation:

$$R = 0.33 * P + 0.33 * C + 0.27 * M + 0.07 * F$$

where:

R = Overall ranking score (higher scores are best)

P = Performance score

C = Change score

M = Minority score

F = Family income score

(2) Mathematical models that use more than one input variable to produce an output variable are called **multivariable models**. There are many types of multivariable models, but first we will concentrate on models which are similar to the linear models we studied previously, except that they have multiple input variables. Could the overall ranking score formula above be considered a multivariable model? Why or why not?

(3) Weighted averages are averages that give each number in a set more (or less) influence, or **weight**, based on what the number represents. This is different from a standard averaging calculation in which you add a set of numbers and divide the total by the count of the numbers in the set. This standard type of averaging assigns the same weight to each of the numbers (each with a weight of 1, where it is counted just one time). Weighted-averaging calculations try to give a more realistic description of some sets by giving more "valuable" numbers greater weight and therefore a greater impact on the result. Does the overall ranking score formula above represent a weighted average? Why or why not?

(4) Consider again the overall ranking score formula above.

(a) Which categories are the most significant?

(b) Which category contributes the least to the overall ranking score?

(c) If this is an "average", what number are we dividing by?

PROBLEM SITUATION 2: CALCULATING GRADE POINT AVERAGE (GPA)

A Grade Point Average (GPA) is another kind of weighted average. Letter grades are assigned a number value in order to calculate a numerical average. This table shows typical point values:

A	A−	B+	B	B−	C+	C	C−	D+	D	D−	F
4.0	3.7	3.3	3.0	2.7	2.3	2.0	1.7	1.3	1.0	0.7	0

(5) When calculating GPA in high school, the **mean** is often used. With reference to both tables, calculate a student's GPA for the following four high school classes.

Class	**Grade**
English	A
Math	B+
Spanish	A-
Biology	B

(6) When calculating a college GPA, a **weighted average** is used. This means some grades have a higher impact than others because the weights of the grades for each class are determined by the number of credit hours in the class. Calculate the GPA for the following four college classes. Round to the nearest hundredth.

Class	**Credits**	**Grade**
Spanish	5	A-
Physics	4	B
History	3	C
English	3	B+

(7) Suppose Hamm was taking the four college classes above and also got an A in her zero-credit math class. How would the math class affect her GPA?

Class	**Credits**	**Grade**

Spanish	5	A-
Physics	4	B
History	3	C
English	3	B+
Math	0	A

(8) Suppose Messi was taking the same five college classes as Hamm and also got an F in his one-credit art class. How would the art class affect his GPA?

Class	**Credits**	**Grade**
Spanish	5	A-
Physics	4	B
History	3	C
English	3	B+
Math	0	A
Art	1	F

This page titled 5.7: Corequisite- Multivariable Models is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

5.7.1: Practice Problems Corequisite M.7

Use the following table of typical point values for the problems below.

A	A−	B+	B	B−	C+	C	C−	D+	D	D−	F
4.0	3.7	3.3	3.0	2.7	2.3	2.0	1.7	1.3	1.0	0.7	0

(1) Calculate the GPA for the following four college classes. Round to the nearest hundredth, if needed.

Class	**Credits**	**Grade**
English	3	A
Math	3	B+
Spanish	5	A-
Biology	4	C

(2) Calculate the GPA for the following four college classes. Round to the nearest hundredth, if needed.

Class	**Credits**	**Grade**
English	3	C
Math	3	B
Spanish	5	A-
Biology	4	C

(3) Morgan would like to apply for a college scholarship. There is a requirement of a 3.5 GPA. Suppose Morgan took the same four college classes and will also be taking a three-credit History class. If Morgan gets an A in History, what would their GPA be? Round to two decimal places, if need be.

Class	**Credits**	**Grade**
English	3	B+
Math	3	B+
Spanish	5	A-
Biology	4	C
History	3	A

(4) Evan would like to apply to a different college. There is a minimum requirement of a 3.3 GPA. Suppose Evan took the original four college classes and will also be taking a three-credit Film class. What is the **minimum** grade needed in the Film class in order to meet the 3.3 GPA requirement to apply?

Class	**Credits**	**Grade**
English	3	A
Math	3	B+
Spanish	5	A-
Biology	4	C
Film	3	???

This page titled 5.7.1: Practice Problems Corequisite M.7 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

6: Corequisite Module S

6.1: Corequisite- Graphical Literacy

6.1.1: Practice Problems Corequisite S.1

6.2: Corequisite- Graphical Literacy

6.2.1: Practice Problems Corequisite S.2

6.3: Corequisite- Four-Step Process, Interpreting Statistical Studies, Understanding Experimental Design

6.3.1: Practice Problems Corequisite S.3

6.4: Corequisite- Characteristics of Samples and Sample Data

6.4.1: Practice Problems Corequisite S.4

6.5: Corequisite- Distribution, Shape, and Terminology

6.5.1: Practice Problems Corequisite S.5

6.6: Corequisite- Computing Measures of Center; Computing and Interpreting Z-scores

6.6.1: Practice Problems Corequisite S.6

This page titled 6: Corequisite Module S is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.1: Corequisite- Graphical Literacy

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

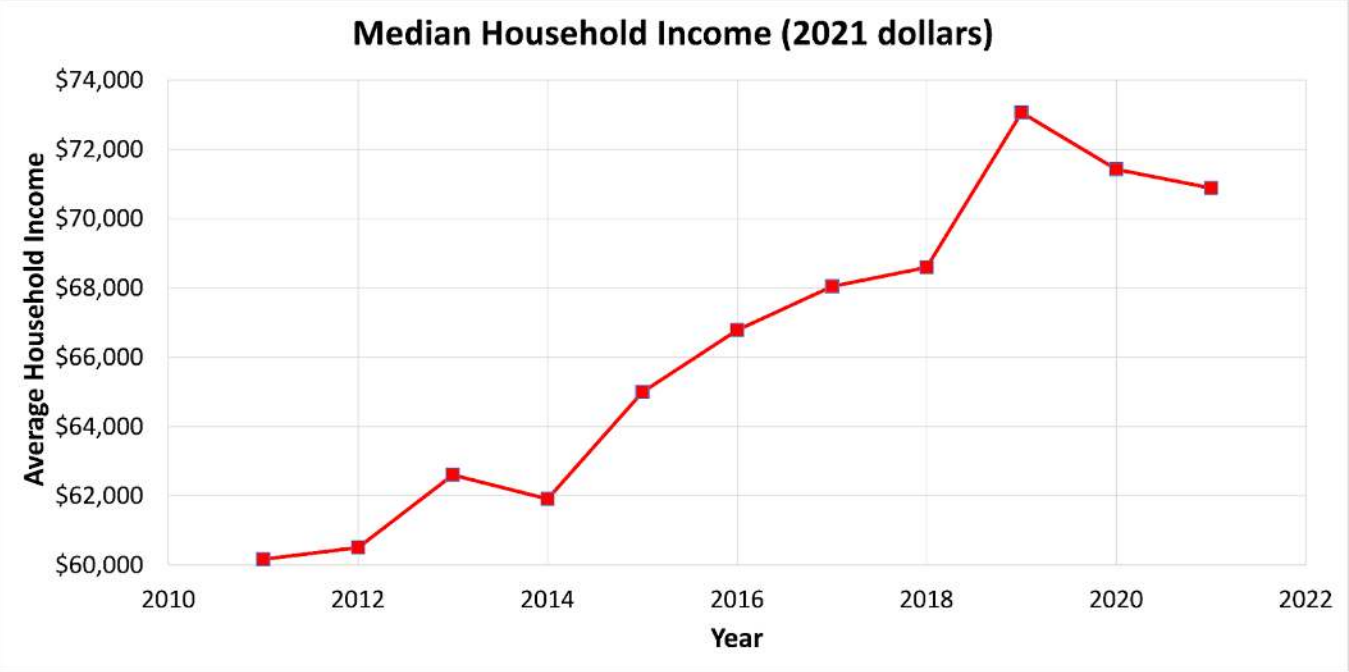
- to fully understand a pie chart, the reference value must be known.

By the end of this lesson, you should be able to

- calculate relative change and average rate of change from a line graph.
- estimate the absolute size of the portions of a pie chart given its reference value.
- use data displayed on two graphs to estimate a third quantity.

PROBLEM SITUATION 1: MEDIAN HOUSEHOLD INCOME

The graph below shows the median income of households in the United States from 2011 to 2021.[1] A household income (or family income) is the total income of all people living in the household. The median household income in the U.S. in a given year is the middle household income when all household incomes in that year are sorted from lowest to highest.



The Writing Principle

Throughout this course, you are asked to explain your thinking about things such as the meaning of data, how you calculated the data, or how you know your answer is correct, deepens your own understanding of the material. Expressing your thinking clearly is a learning tool. We can do this by following The Writing Principle.

Writing Principle: Use specific and complete information. The reader should understand what you are trying to say even if he or she has not read the question or writing prompt. This includes

- information about context, and
- quantitative information.

To help guide your writing, ask the following questions.

What is the topic of the writing task?

- What is the task telling me to do?
- Some examples are given below:
  - Describe how you found the answer.

- Explain why you think you have the right answer.
- Reflect on the process of coming up with the answer.
- Make a prediction about the next data point.
- Compare two data points or the answers to two parts of the problem.

- What information am I given to help me with the task?

1. What do you notice about this graph? Carefully describe any trends and patterns that you see in the data. Use the Writing Principle.

(2) Use Graph 1 to answer the following questions.

(a) Select the best phrase to complete this sentence: The numbers on the horizontal axis (the line across the bottom of the graph) represent

(i) income from $2,010 to $2,022.

(ii) income from $2,010,000 to $2,022,000.

(iii) years from 2010 to 2022.

(iv) years and income.

(b) Select the best phrase to complete this sentence: The numbers on the vertical axis (the line along the left of the graph) represent

(i) years from $60,000 to $74,000.

(ii) income for a household or family from $60,000 to $74,000.

(iii) income for one person from $60,000 to $74,000.

(iv) income in 2021 from $60,000 to $74,000.

(c) Estimate the median household income in:

2012:_____

2016:_____

2020:_____

(d) In what one-year period did the median household income increase the most? What is the increase?

(e) Use the estimates from part (c) to calculate the relative change in the median household income from 2012 to 2020. Round to the nearest one percent.

(f) What is the average rate of change from 2012 to 2020? What does this mean?

PROBLEM SITUATION 2: UNDERSTANDING THE NATIONAL DEBT & GDP

The increase in the U.S. national debt is a frequent topic in political and economic conversations. The *national debt* is how much the country owes. The gross domestic product (GDP) is the value of all goods that are produced in a country in a certain time period, typically one year. A country's GDP is a measure of the country's economic output.

Before examining the national debt and GDP of the U.S., let's consider the following scenario involving an individual person.

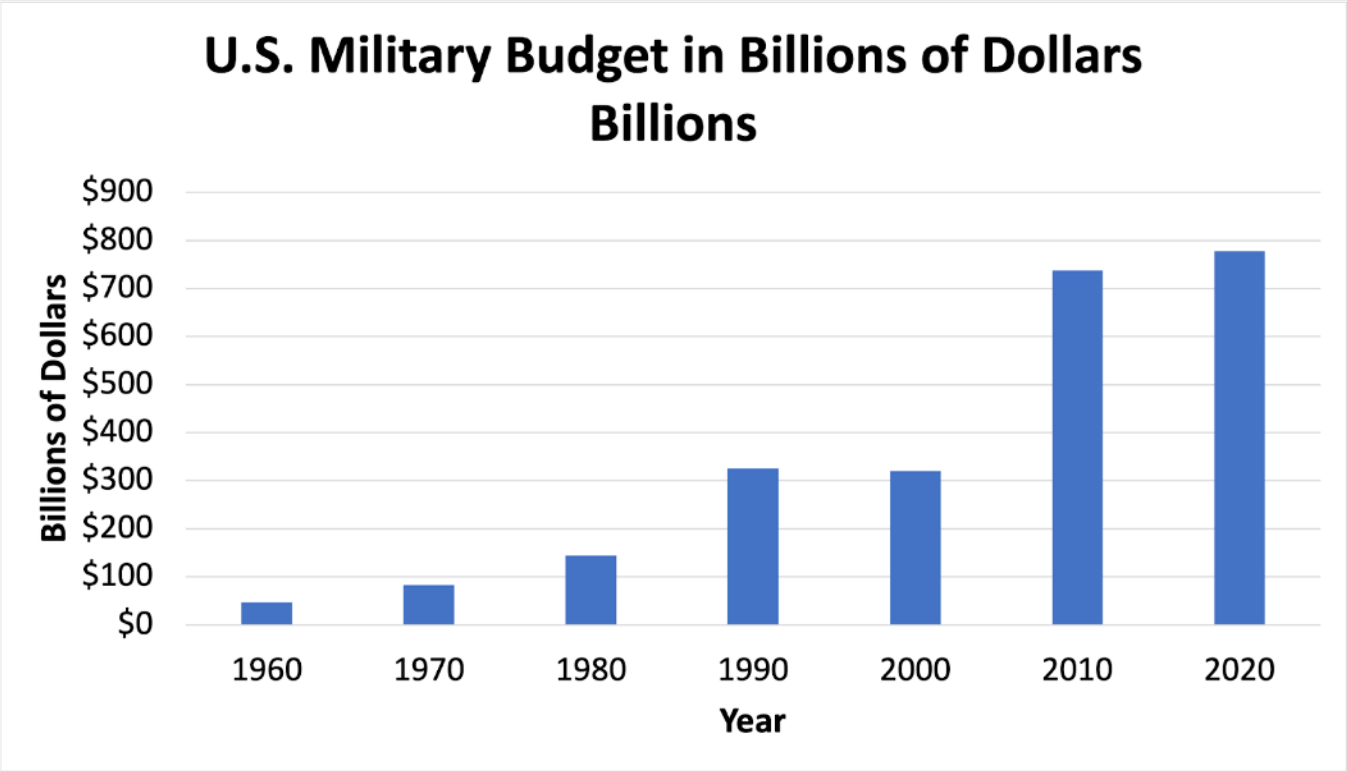
(3) Two pairs of statements about Jeff's housing expenditures are given below.

Pair 1	**Pair 2**
In 2000, Jeff spent $900 per month on housing. In 2020, Jeff spent $1,800 per month on housing.	In 2000, Jeff spent 20% of his income on housing. In 2020, Jeff spent 10% of his income on housing.

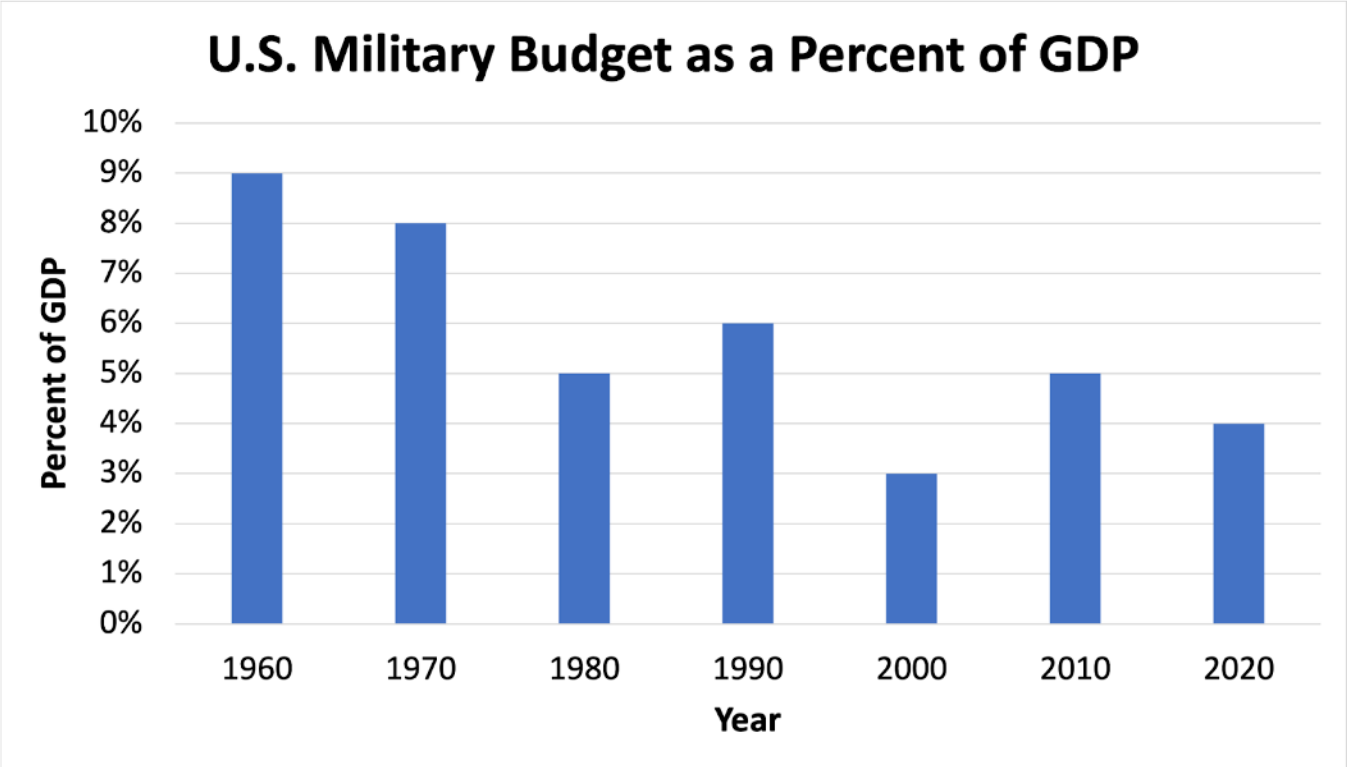
How can both pairs of statements be true?

Think about the statement, "The 2020 military budget is way out of hand and has never been higher." You will use Graphs 3 and 4 to evaluate the statement.

Graph 2[2]



Graph 3



(4) Is the statement true? Based on what information?

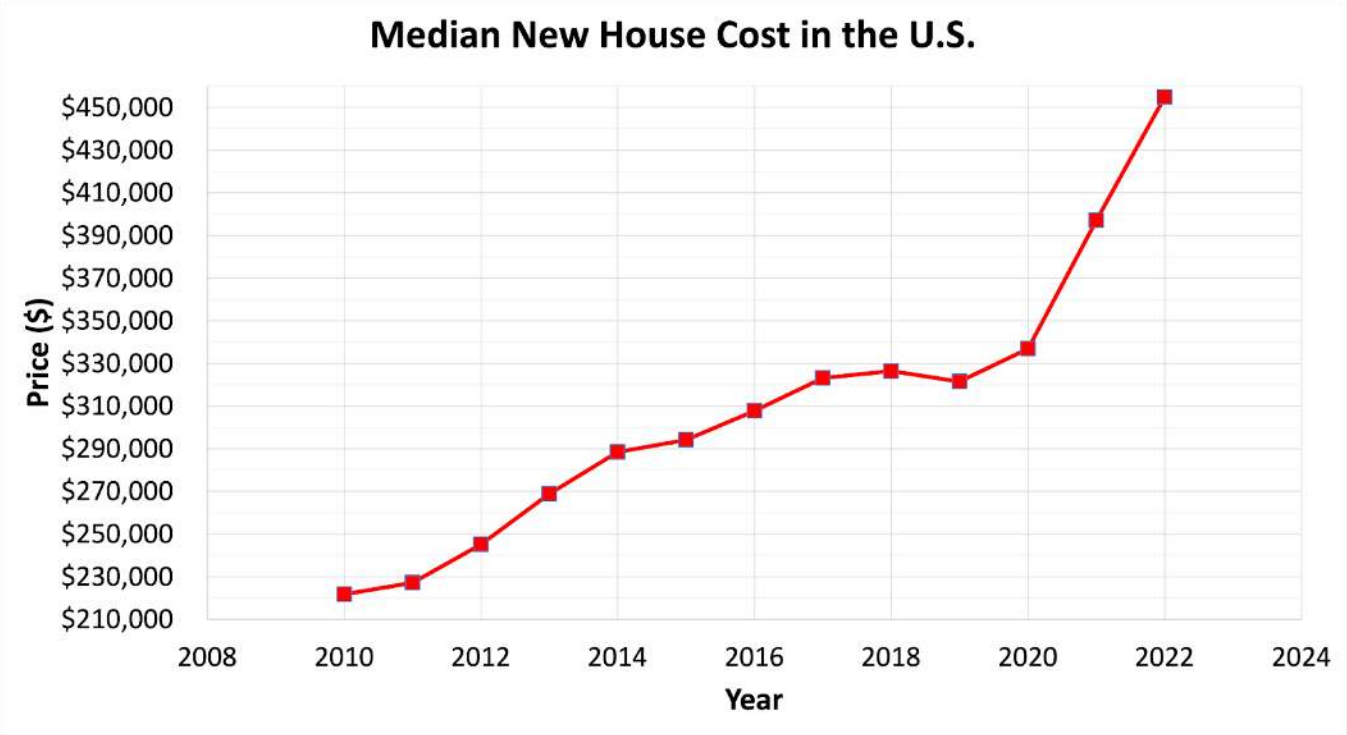
[1] https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html

[2] https://www.macrotrends.net/countries/USA/united-states/military-spending-defense-budget

This page titled 6.1: Corequisite- Graphical Literacy is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.1.1: Practice Problems Corequisite S.1

Use the graph[3] below to answer the following questions. (Note: The median is the middle number of a data set when the numeric values are arranged in order of size.)



(1) Estimate the percentage increase in new housing prices from 2010 to 2022. (Choose the best answer.)

(i) The prices increased from around $220,000 to around $450,000, about a 49% increase.

(ii) The prices increased from around $220,000 to around $230,000, about a 16% increase.

(iii) The prices increased from around $220,000 to around $450,000, about a 105% increase.

(iv) The prices increased from around $220,000 to around $230,000, about a 14% increase.

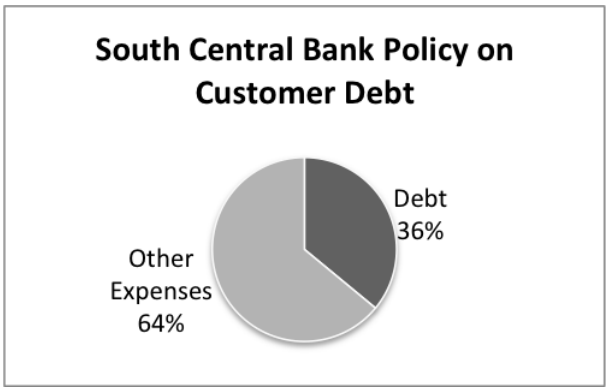
Estimate the percentage increase in new housing prices from 2011 to 2013. (Fill in the blanks.) (Fill in the blanks for **Questions 2-4**.)

The prices increased from (2) _____ to (3)_____. This is about a (4)_____% increase.

Estimate the percentage decrease from 2020 to 2022. (Fill in the blanks for **Questions 5-7**.)

The prices decreased from (5)_____ to (6)_____. This is about a (7)_____% decrease.

South Central Bank has a policy that limits the amount of debt customers may have in order to receive a loan. The following pie chart shows the *highest* percentage of debt that the bank allows.



(8) What is the reference value in this situation?

(i) Other expenses

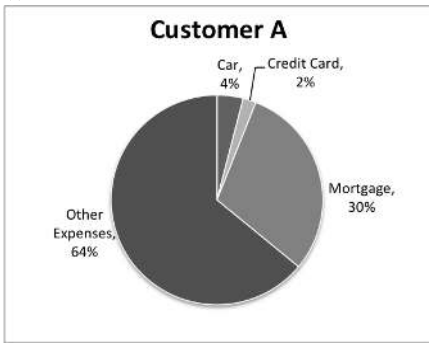
(ii) Debt

(iii) Total of other expenses and debt

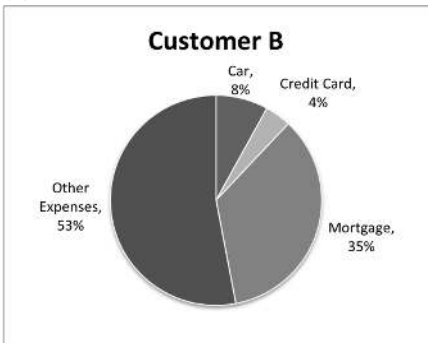
(iv) Other expenses minus debt

(9) Three graphs are given below. Each graph represents a loan customer. Each customer's debt is broken into three categories: car, credit card, and mortgage (the loan on a house). Which customer(s) meet the bank policy on the limit of the amount of debt? There may be more than one correct answer.

(i)



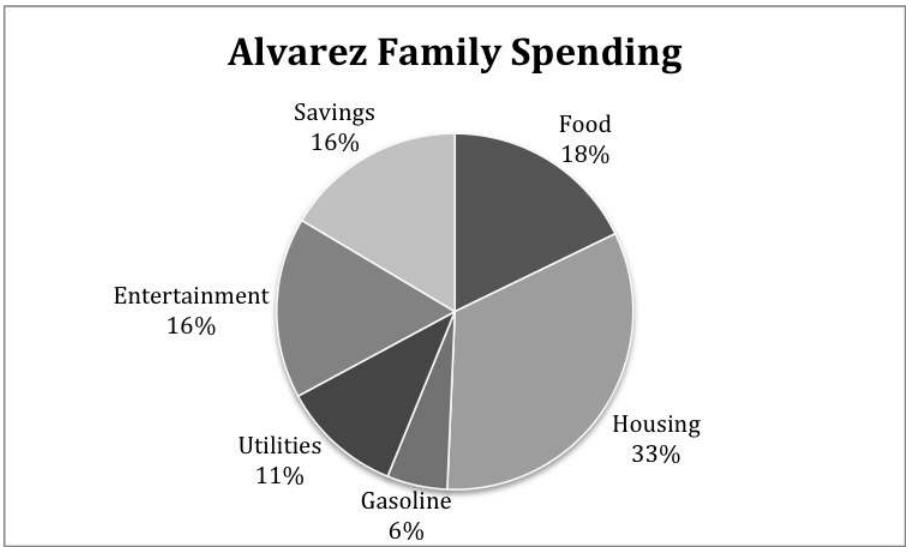
(ii)

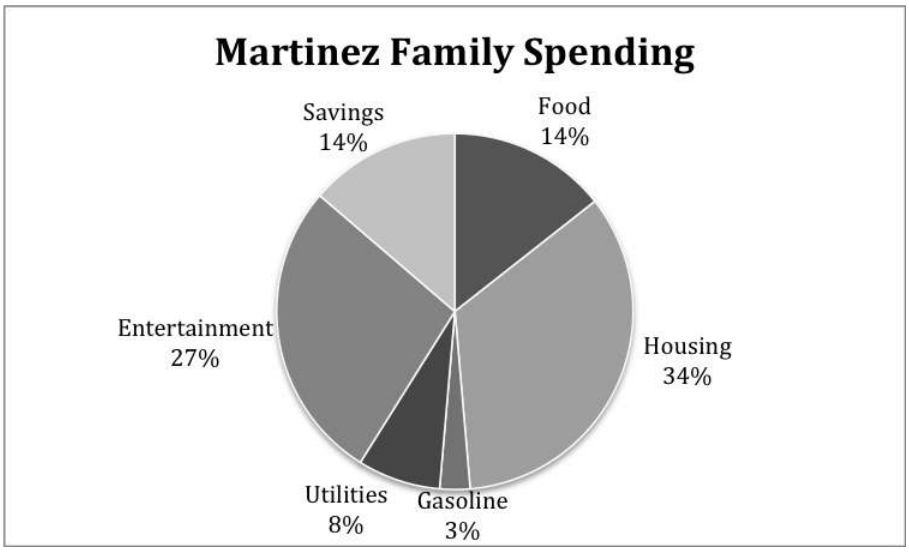


(iii)



The following two pie charts show how two families spend their money. The Alvarez family has a take-home pay of $3,650 per month and the Martinez family has a take-home pay of $7,300 per month.





(10) Select the statement that best compares how much the families spend on gasoline.

(i) The Alvarez Family spends more on gasoline than the Martinez Family.

(ii) The two families both spend $200 on gasoline.

(iii) Both families spend around $200 on gasoline. This is 6% of the Alvarez budget, but only 3% of the Martinez budget because the Martinez family starts with about twice as much money as the Alvarez family.

(iv) The Alvarez family spends more on gasoline than the Martinez family because 3% of $7,300 is more than 6% of $3,650.

Estimate the amount each family spends on food and housing. (Fill in the tables by entering answers for **Questions 11-14**.)

Martinez Family	
Food	(11)
Housing	(12)

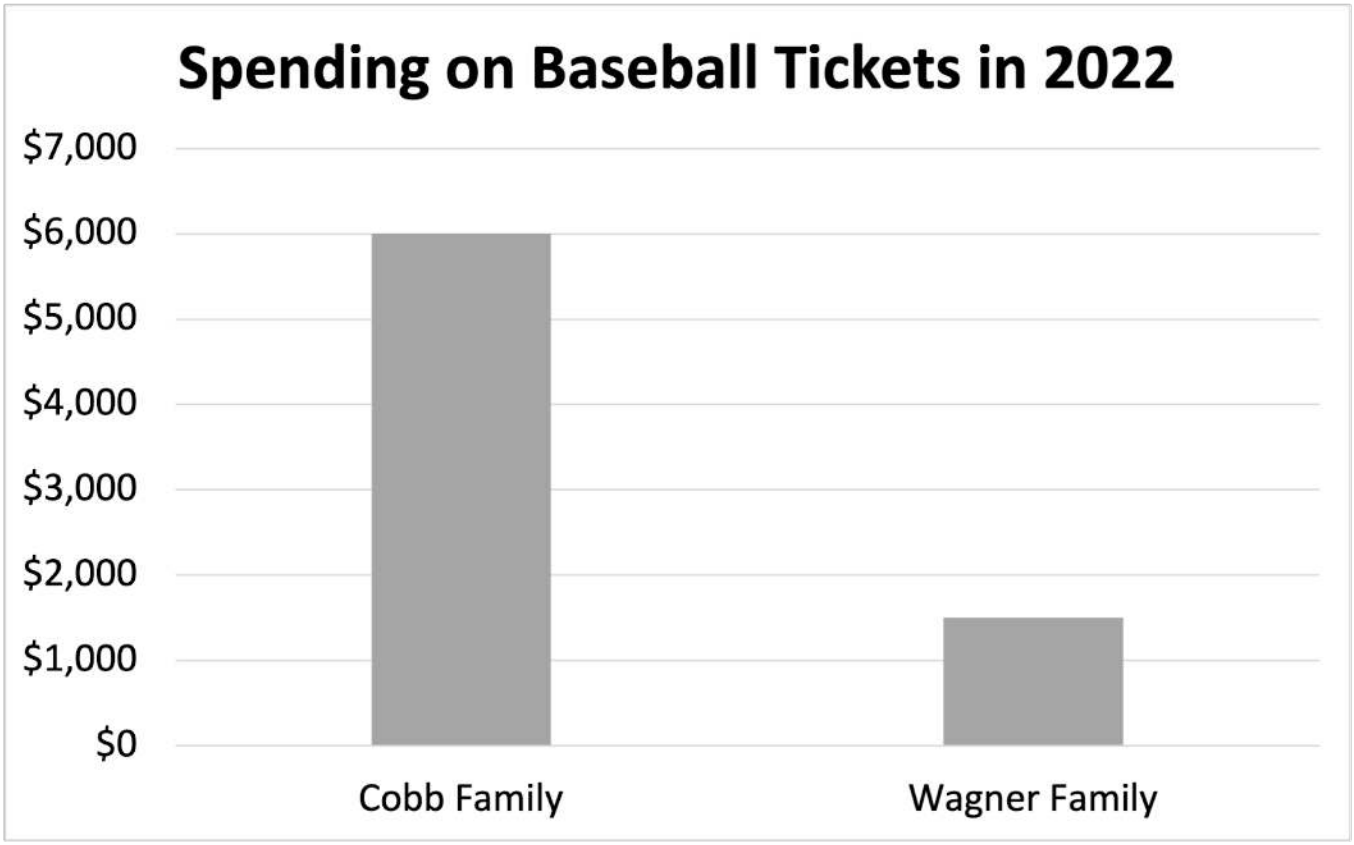
Alvarez Family	
Food	(13)
Housing	(14)

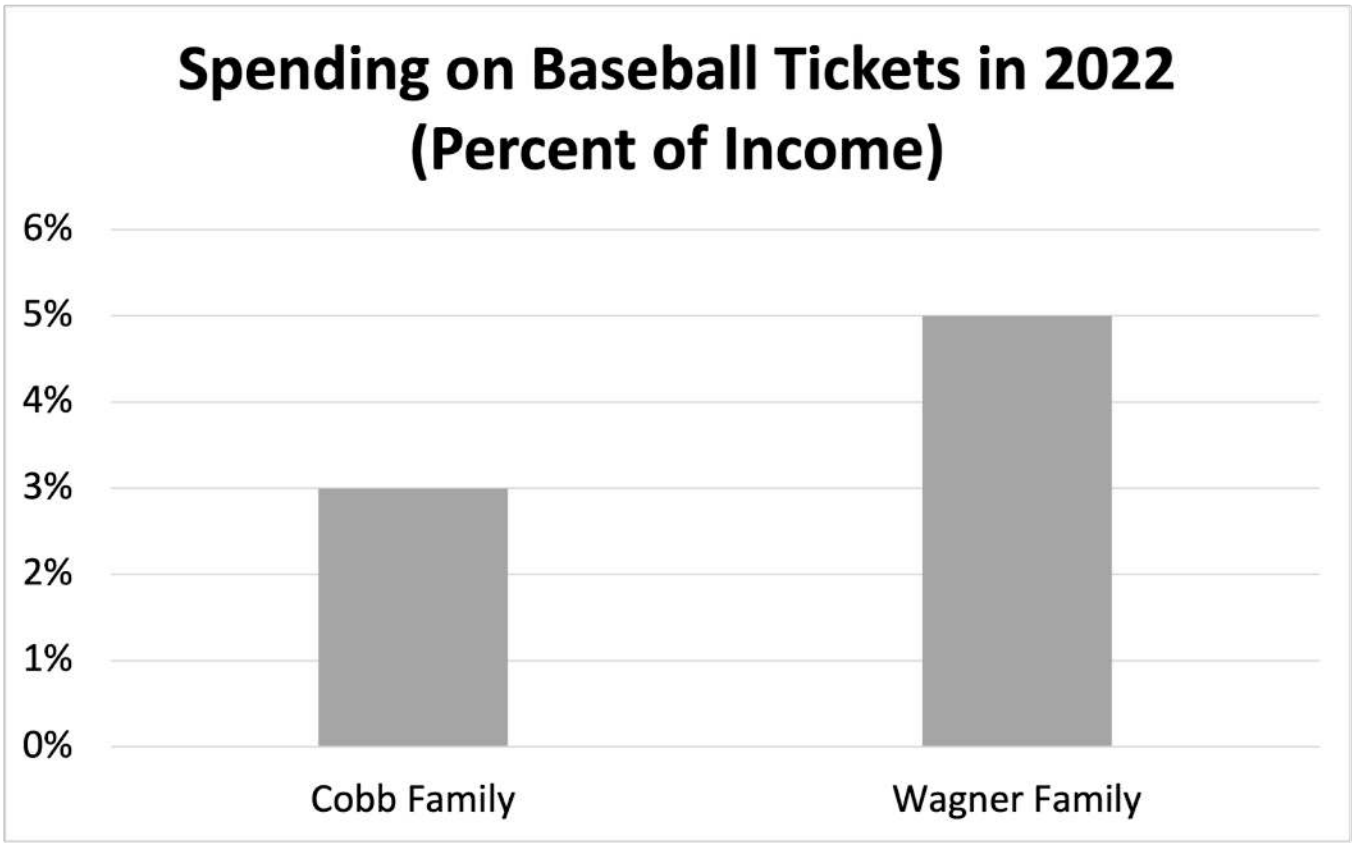
(15) Both families spend about the same percentage of their income on housing. The family with the _____ income can afford a house that has twice the payment and maintenance costs.

(i) Higher

(ii) Lower

(16) The following two graphs show how much two families spend on tickets to professional baseball games each year.





Choose the true statement based on these graphs.

(i) The Wagners must have a higher income than the Cobbs. They spend less on baseball tickets, but it is a higher portion of their total income.

(ii) The Cobbs must have a higher income than the Wagners. They spend more on baseball tickets, but it is a smaller portion of their total income.

(iii) The Wagners must have a higher income than the Cobbs. They spend more on baseball tickets, but it is a smaller portion of their total income.

(iv) The Cobbs must have a lower income than the Wagners. They spend more on baseball tickets, but it is a smaller portion of their total income.

[3] https://www.census.gov/construction/nrs/historical_data/index.html

This page titled 6.1.1: Practice Problems Corequisite S.1 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.2: Corequisite- Graphical Literacy

SPECIFIC OBJECTIVES

Students will understand that

- bar charts and pie charts are graphical tools used to tell a quantitative story in a visual way.
- the change in a quantity can be expressed as an absolute change and a relative change.

Students will be able to

- read bar charts and pie charts.
- interpret the story behind bar charts and pie charts.
- create bar charts and pie charts.
- compute absolute and relative changes.

PROBLEM SITUATION 1: THE EFFECT OF REDUCING RISK

A new drug advertises that it "reduces the risk of heart attack by 50%." In order to better understand the benefits of this drug, you will examine heart attack risk for two different groups.

Group 1 consists of individuals in Africa who are

- 40 years old,
- do not have Diabetes Mellitus,
- smoke tobacco,
- have high cholesterol, and
- have high blood pressure.

Group 2 consists of individuals in Africa who are

- 40 years old,
- do not have Diabetes Mellitus,
- do not smoke,
- have low cholesterol, and
- have high blood pressure.

The World Health Organization reports that individuals in Group 1 have a 40% chance (or risk) of suffering a heart attack within 10 years. The same report indicates that individuals in Group 2 have a 10% risk of suffering a heart attack within 10 years.

(1) What are the differences between individuals in Group 1 and Group 2?

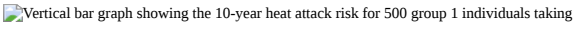
The following questions are designed to help you understand heart attack risk in Group 1.

(2) (a) If 500 people from Group 1 are observed for 10 years, how many individuals would you expect to suffer a heart attack within this time period?

(b) Create a graph that shows the risk of heart attack within 10 years for 500 Group 1 individuals.

(c) If 500 people from Group 1 are treated with the new drug that is advertised to "reduce the risk of heart attack by 50%," how many individuals would you expect to suffer a heart attack within the next 10 years?

(d) Look at the graph below. Think about how best to describe what this graph shows in one or two sentences. Record your answer in the space below.


Vertical bar graph showing the 10-year heat attack risk for 500 group 1 individuals taking new drug.

(e) Compare the graph above to your graph from part (b). What, specifically, do these graphs show that was reduced by 50%?

(f) There are two ways to report the expected drop in heart attacks due to the new drug. The *absolute change* is a subtraction of the old outcome from the new outcome. It is typically reported as a number. *Relative change* is a ratio of the amount of change in comparison to the old outcome. It is typically reported as a percentage. For each group, report both the absolute and relative changes in the number of individuals who are expected to suffer heart attacks.

Absolute Change: _____ fewer people

Relative Change: _____% fewer people

The following questions are designed to help you understand heart attack risk in Group 2.

(3) (a) If 500 people from Group 2 are observed for 10 years, how many individuals would you expect to suffer a heart attack within this time period?

(b) Create a graph that shows the risk of heart attack within 10 years for 500 Group 2 individuals.

(c) If 500 people from Group 2 are treated with the new drug and are observed for 10 years, how many individuals would you expect to suffer a heart attack within this time period?

(d) Look at the graph below. Think about how to describe this graph in one or two sentences.


Vertical bar graph showing the 10-year heat attack risk for 500 group 2 individuals taking new drug.

(e) Compare the graph above to your graph from Question 3 part (b). What, specifically, do these graphs show that was reduced by 50%?

(f) There are two ways to report the drop in heart attacks due to the new drug. The absolute change refers to the actual number of fewer individuals suffering heart attacks as a result of using the new drug, while the relative change (or percent change) refers to this change as a percentage. Report both the absolute and relative change in the number of individuals who suffered heart attacks. Enter values only in the spaces below.

Absolute Change: _____ fewer people

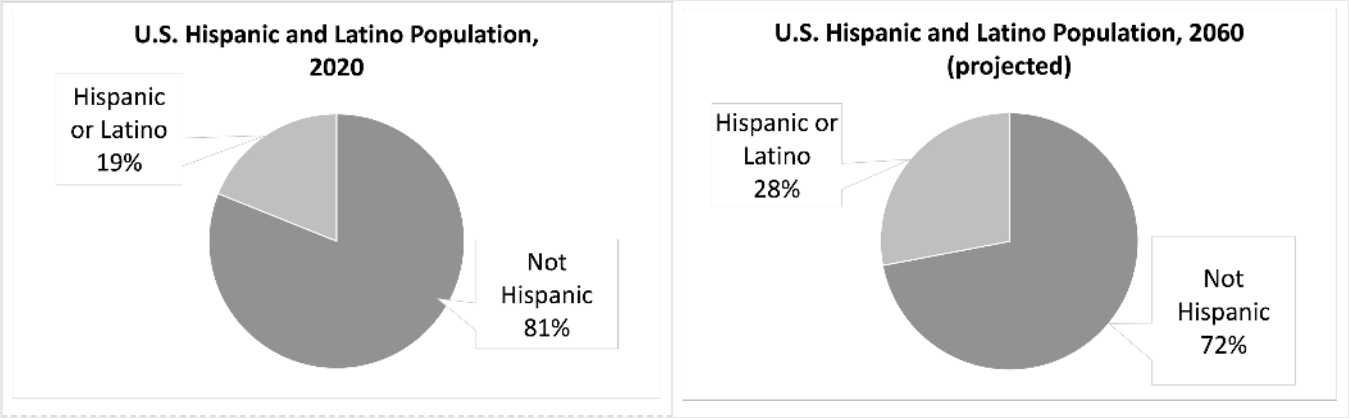
Relative Change: _____% fewer people

(4) Suppose a third group of people have only a 0.5% chance of suffering a heart attack. How do you think the absolute and relative changes in the number of individuals (out of 500) who are expected to suffer a heart attack will compare to your earlier calculations? Hint: Do the absolute and relative change calculations for this new third group. What do you notice?

(5) Based on the previous calculations, critique the statement: “A drug which reduces the risk of heart attack by 50% will most likely save many lives.”

PROBLEM SITUATION 2: READING PIE (CIRCLE) CHARTS

Refer to the graphs below to answer the following questions. This pair of graphs cannot be used to predict that the number of non-Hispanics in the United States is expected to decline between 2020 and 2060.



(6) A student claims that this pair of graphs indicates that the number of non-Hispanics in the United States is expected to decline between 2020 and 2060. Is this statement reasonable? Explain.

(7) The U.S. population in 2020 was around 330,000,000. In 2060, the U.S. population is expected to be around 404,000,000. Estimate the number of Hispanic and non-Hispanic Americans at each time.

(8) Compare the relative increases in the number of Hispanics in the U.S., and number of non-Hispanics in the U.S., from 2020 to 2060.

(9) If the percentage of people in the U.S. who are Hispanic increases at a linear rate from 2020 to 2060, predict the year in which the percentage of people in the U.S. who are Hispanic will reach 25%. Develop and use a linear model to answer this question. Let t be the years since 2020, and H be the percentage of people in the U.S. who are Hispanic.

This page titled 6.2: Corequisite- Graphical Literacy is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.2.1: Practice Problems Corequisite S.2

Think back to the work you did in Problem Situation 1 that saw you engage with the World Health Organization's report detailing the risk of individuals in each group suffering a heart attack within 10 years. Then, you used a group with 500 people. Repeat the work with a group of 1500 people. Group 1 has a 40% risk of heart attack and Group 2 has a 10% risk. The drug reduces the risk of a heart attack by 50%.

(1) How many people in **Group 1** are expected to have a heart attack **without the drug**?

(2) How many people in **Group 2** are expected to have a heart attack **without the drug**?

(3) What percentage of **Group 1** is expected to have a heart attack **with the drug**?

(4) How many people in **Group 1** are expected to have a heart attack **with the drug**?

(5) What percentage of **Group 2** is expected to have a heart attack **with the drug**?

(6) How many people in **Group 2** are expected to have a heart attack **with the drug**?

(7) Select the true statement:

(i) The size of the group does not change anything about the results for the group.

(ii) The size of the group changes the number of people who are expected to have a heart attack with the drug, but it does not change the percentage.

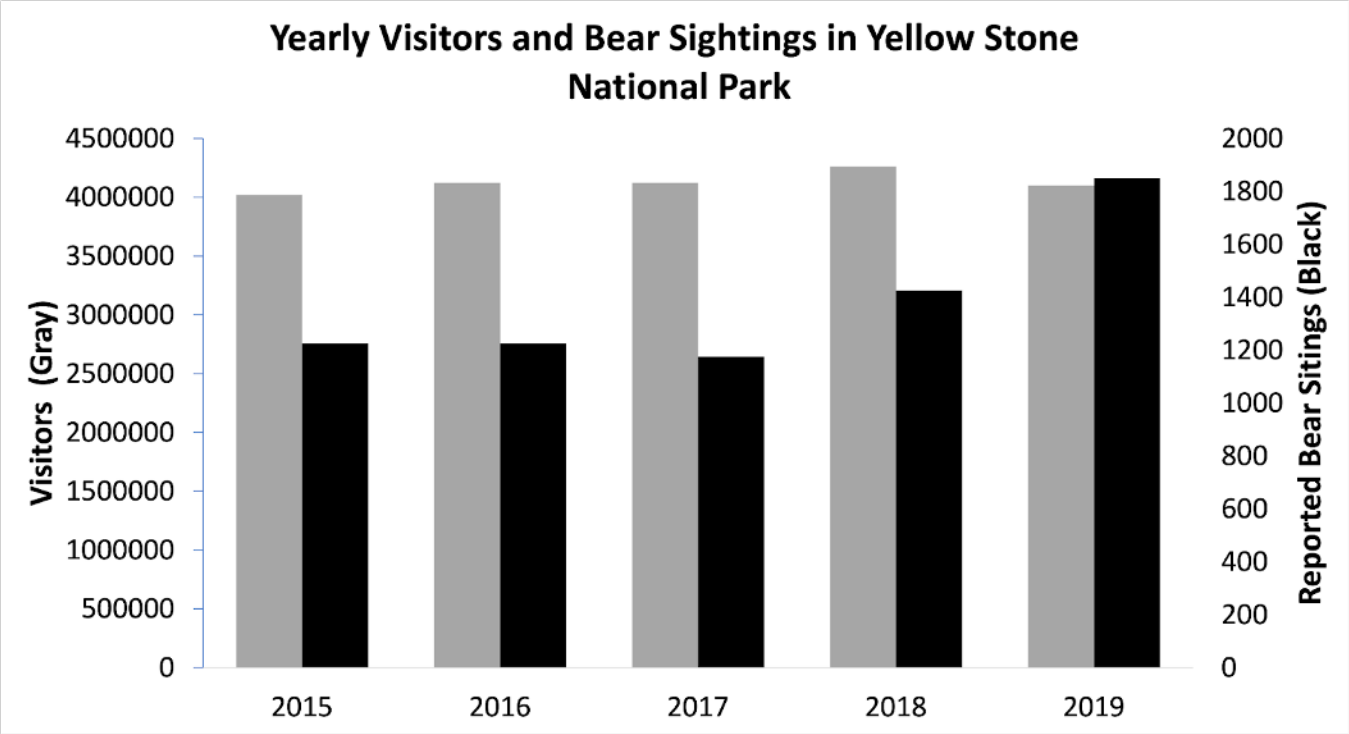
(iii) The size of the group does not change the number of people who are expected to have a heart attack with the drug, but it does change the percentage.

(iv) The size of the group changes the number and the percentage of people who are expected to have a heart attack with the drug.

BEAR SIGHTINGS

The bar graph below shows the number of bear sightings (all bears, not only grizzlies) in Yellowstone National Park from 2015 to 2019.[4] Because the number of sightings is affected by the number of visitors to the park, these data are shown as well.[5]

Note: Because the number of visitors each year is much larger than the number of bear sightings, the data are plotted on different scales. The scale on the left (ranging from zero to 4,500,000) is for the number of visitors, shown in gray. The scale on the right (ranging from zero to 2000) is for the number of bear sightings, shown in black.



(8) In what year was the percentage of visitors who saw bears the highest? _____

(9) In what two years was the percentage of visitors who saw bears the lowest? _____ & _____

(10) Estimate the average number of bear sightings over these five years.

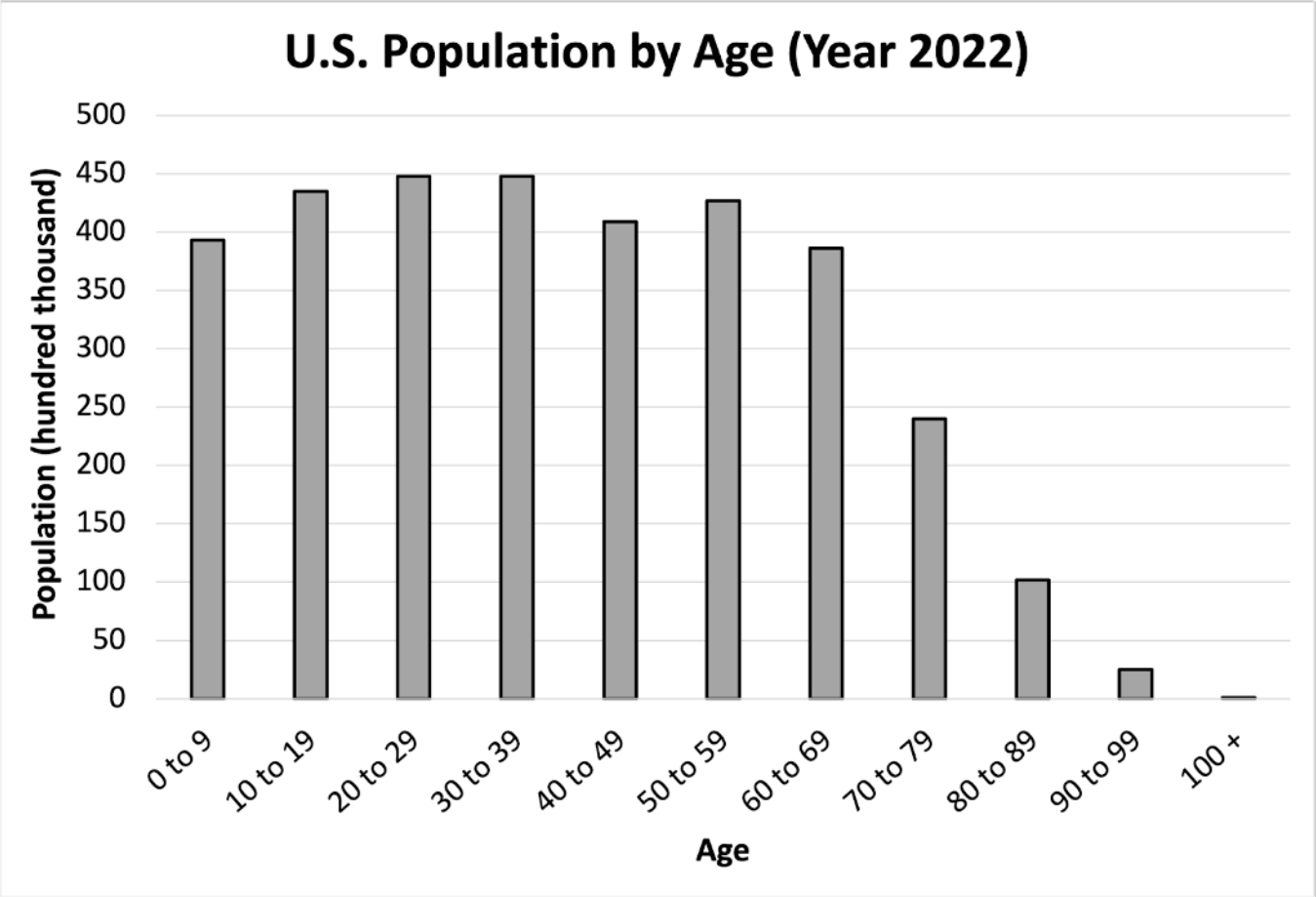
(i) 1100

(ii) 1200

(iii) 1400

(iv) 1600

The graph below is based on information from the 2020 U.S. Census.[6] Use the graph to answer the questions.



(11) Which is the best estimate of the total population for age group 0–9?

(i) 400

(ii) 40,000

(iii) 40,000,000

(iv) 400,000,000

(12) Which is the best estimate for the total population of age group 70–79?

(i) Two hundred and four

(ii) Two hundred and four thousand

(iii) Two hundred and four million

(iv) Twenty four million

(13) Which statement is the most accurate?

(i) There are fewer than 10,000,000 people age 70 and older.

(ii) There are between 10,000,000 and 20,000,000 people age 70 and older.

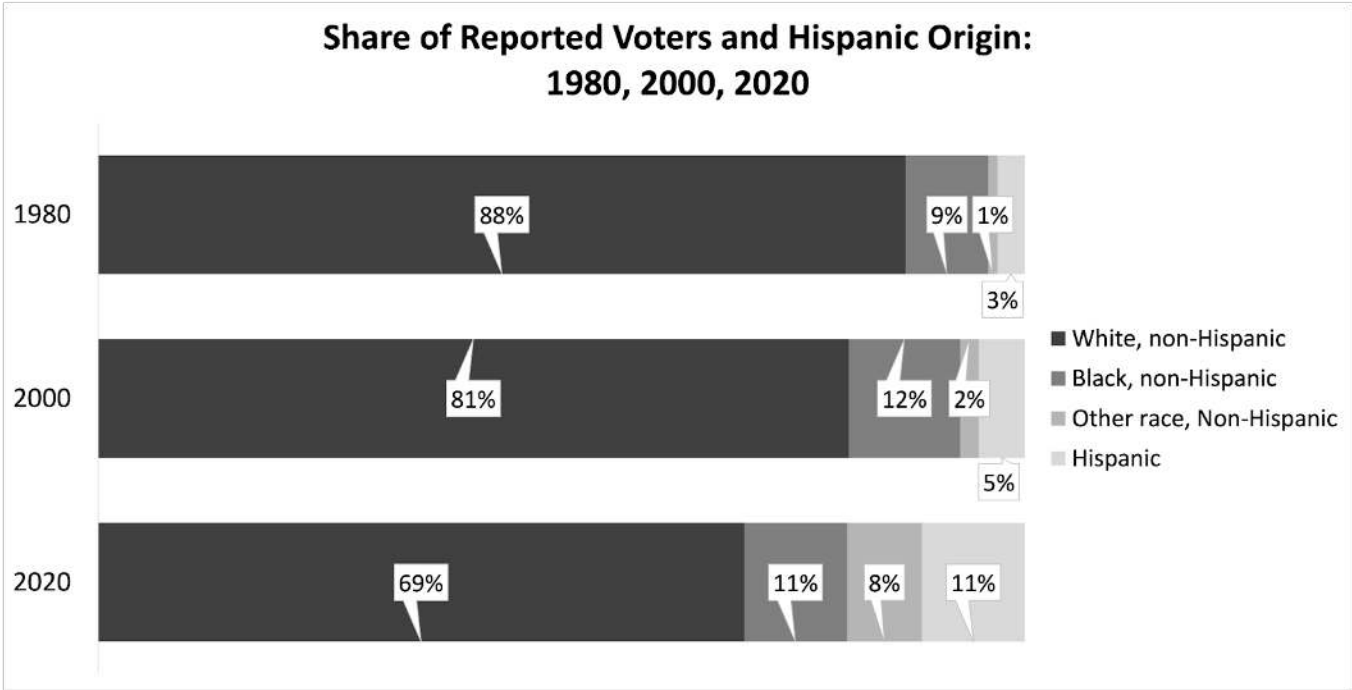
(iii) There are between 20,000,000 and 30,000,000 people age 70 and older.

(iv) There are between 30,000,000 and 40,000,000 people age 70 and older.

(14) An advertiser is considering two advertising campaigns for a product. Campaign A is most effective for ages 20–39, while Campaign B is most effective for ages 30–49. Which campaign should the advertiser choose to reach the most people?

VOTING IN PRESIDENTIAL ELECTIONS

The stacked bar graph below shows how the share of Hispanic voters and voters in racial groups has changed between the 1980, 2000, and 2020 presidential elections.[7] In 1980, 2000, and 2020, the approximate number of voters in the presidential elections were 86.5 million, 105 million, and 154.6 million, respectively.[8]



(15) Is it reasonable to conclude that the number of white, non-Hispanic voters who voted in the 2020 presidential election is less than the number of white, non-Hispanic voters who voted in the 1980 presidential election? Explain.

(16) From 1980 to 2020, determine the **absolute increase** in the number of black, non-Hispanic voters who voted in the presidential elections.

(17) From 1980 to 2020, determine the **relative increase** in the number of black, non-Hispanic voters who voted in the presidential elections.

(18) Linear Model: Using data from 1980 and 2020, develop a linear model for the data. Let S represent the share of the presidential votes cast by Hispanic voters (as a percentage) and t represent the number of years since 1980. Enter your model in format S = mt + b.

(19) Assume that the share of Hispanic voters in presidential elections changes in a linear manner over time. Use the linear model to determine in what year will the share of Hispanic voters be greater than 20%.

---

[4] https://www.nps.gov/yell/learn/management/upload/2019-BEAR-REPORT-FINAL_reduced.pdf

[5] https://www.statista.com/statistics/254231/number-of-visitors-to-the-yellowstone-national-park-in-the-us/

[6] https://www.census.gov/data/tables/time-series/demo/popest/2020s-national-detail.html

[7] https://www.census.gov/newsroom/blogs/random-samplings/2017/05/voting_in_america.html and https://www.pewresearch.org/politics/2020/06/02/the-changing-composition-of-the-electorate-and-partisan-coalitions/#electorate-has-become-more-ethnically-diverse-better-educated

[8] http://www.presidency.ucsb.edu/index.php

This page titled 6.2.1: Practice Problems Corequisite S.2 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.3: Corequisite- Four-Step Process, Interpreting Statistical Studies, Understanding Experimental Design

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

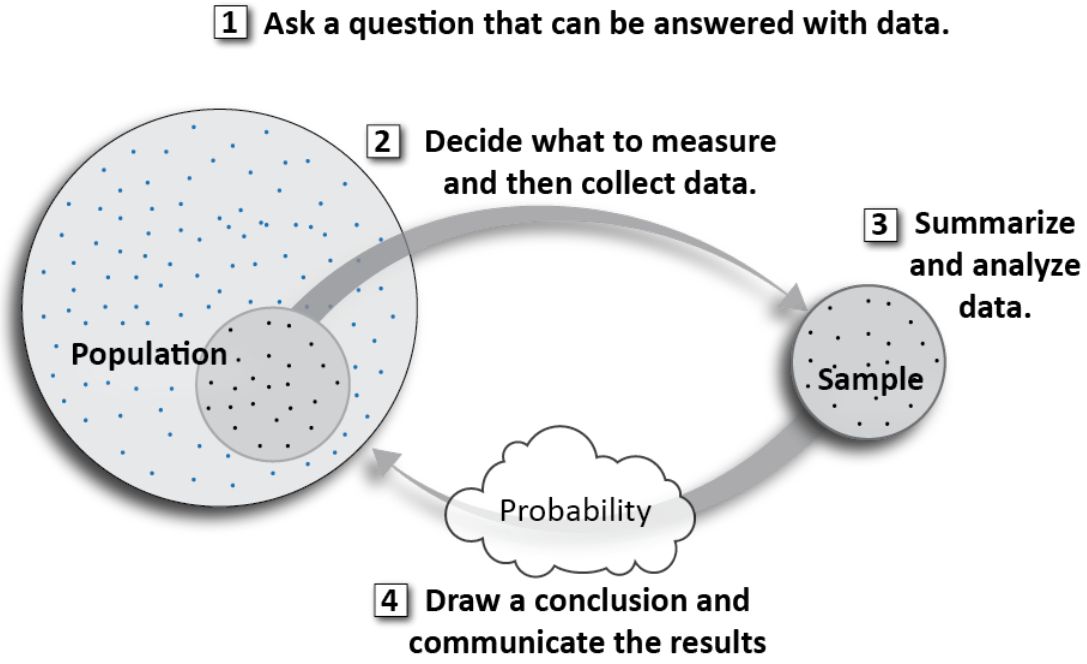
- statistical investigations are a cycle of forming a research question, designing a study and collecting relevant data, data analysis, statistical inference, and drawing appropriate conclusions.
- experiments can be used to determine whether changes to a variable lead to changes in another variable.

By the end of this lesson, you should be able to

- identify the purpose of a scientific study or statistical study.
- identify variables of interest in a scientific study or statistical study.
- identify the conclusion of a scientific study or statistical study.
- identify the explanatory variable and response variable in an experiment.

INTRODUCTION

This lesson introduces the statistical analysis process. The statistical analysis process can be summarized by the four steps below.



A statistical study is a study that involves the use and analysis of data from a sample. A sample is a collection of objects from a larger population. The examination and analysis of sample data can yield results which can be used to draw conclusions about the larger population. This process enables researchers to gain insights on very large populations.

PROBLEM SITUATION 1: WHO IS MORE PHYSICALLY ACTIVE?

Researchers from the U.S., Netherlands, and England conducted a study[9] to examine how adults in the three countries differ in their physical activity and views about physical activity. Study participants were recruited randomly from a nationally representative cohort of adults in each country. Researchers asked 540 U.S., 748 Dutch (Netherlands), and 254 English adults to wear an accelerometer on their wrist for a 7-day period to measure their physical activity. During this same 7-day period, on each day, participants self-reported on their overall level of physical activity during that day using a 5-point scale (1-Inactive, 2-Mildly Active, 3-Moderately Active, 4-Active, 5-Very Active).

The study found that adults across all three countries self-reported similar levels of physical activity. However, the data from the accelerometers showed that adults from these three countries differed in their actual levels of physical activity.

(1) What are the research questions being asked in this study?

(2) What are the populations of interest?

(3) What variables are being examined for each adult in the study?

(4) Let's examine some of the data from the study. The tables below show the number of U.S. participants and Dutch (Netherlands) participants, based on age-group, who ranked themselves as being in certain physically active categories (labeled subjective), and who were measured via an accelerometer as being in certain physically active categories (labeled objective).

Table 1: Netherlands

	Subjective (Self-ranked)				**Objective (Measured)**			
	18–39	**40–50**	**51–64**	**65+**	**18–39**	**40–50**	**51–64**	**65+**
Inactive	8	15	16	19	14	16	35	85
Mildly active	45	29	56	38	14	34	51	51
Moderately active	59	68	96	92	39	29	49	26
Active	39	25	59	55	42	34	42	34
Very active	5	9	7	9	47	34	56	17
Total	**154**	**145**	**234**	**213**	**155**	**146**	**234**	**213**

Table 2: United States

	Subjective (Self-ranked)				**Objective (Measured)**			
	18–39	**40–50**	**51–64**	**65+**	**18–39**	**40–50**	**51–64**	**65+**
Inactive	11	8	18	11	30	45	71	61
Mildly active	27	39	70	20	24	14	38	20
Moderately active	52	34	52	42	23	12	24	12
Active	34	24	32	22	31	19	25	3
Very active	17	10	7	5	34	26	24	5
Total	**141**	**116**	**180**	**99**	**141**	**116**	**182**	**101**

(a) For each age group, what is the percentage of **Dutch (Netherlands)** participants who **ranked themselves** as being "Active" or "Very Active"? Round each value to the nearest percent.

(b) For each age group, what is the percentage of **U.S.** participants who **ranked themselves** as being "Active" or "Very Active"? Round each value to the nearest percent.

(c) In which age group is the gap in percentages the greatest between the Netherlands and the U.S.? In which age group is the gap in percentages the least?

The gap in percentages is the greatest in the _____ age group, where the U.S.

percentage is _____ percentage points higher.

The gap in percentages is the least in the _____ age group, where the U.S.

percentage is _____ percentage points lower.

(d) For each age group, what is the percentage of **Dutch (Netherlands)** participants who **were measured** as being "Active" or "Very Active"? Round each value to the nearest percent.

(e) For each age group, what is the percentage of **U.S.** participants who **were measured** as being "Active" or "Very Active"? Round each value to the nearest percent.

(f) In which age group is the gap in percentages the greatest between the Netherlands and the U.S.? In which age group is the gap in percentages the least?

The gap in percentages is the greatest in the _____ age group , where the U.S.

percentage is _____ percentage points lower.

The gap in percentages is the least in the _____ age group, where the U.S. percentage is _____ percentage points lower.

(5) The study was conducted on samples of adults from each country that are *representative* of all adults in the countries. Based on this, what conclusions can be drawn from this study?

(6) Based on the results of this study, can we conclude that the Netherlands has more adults age 18 - 39 who are "Active" or "Very Active" than the U.S.? Explain.

(7) There are currently approximately 80 million U.S. adults age 18 - 39 and approximately 4.2 million Dutch (Netherlands) adults age 18 - 39. Using the results of the study, predict the number of adults age 18 - 39 in these two countries whose physical activity level would place them in the "Active" or "Very Active" categories.

Approximately _____ U.S. adults in the 18-39 age group would be labeled as active or very active.

Approximately _____ Dutch adults in the 18-39 age group would be labeled as active or very active. Round to the nearest hundred thousand.

PROBLEM SITUATION 2: UNDERSTANDING EXPERIMENTS

When a difference between groups on a variable is so large that it is *unlikely* to be caused by chance, we say the difference is **significant**. In the first part of this lesson, we examined data from a study of adults from the Netherlands, England, and the U.S. on adults' physical activity and views regarding physical activity. The variables of interest were the age of adults, country of origin, self-reported physical activity and measured physical activity. The researchers concluded that there is a significant difference in physical activity among Dutch and U.S. adults. This means that the observed difference in the physical activity levels in the samples of adults reflects the difference that exists in the actual populations of adults. Statistical studies often involve the use of an *experiment*. Researchers conduct an experiment to examine how changes in one variable impact another variable.

A Bit of Terminology

Let's take a look at some terminology used when describing experiments. When planning an experiment, it is helpful to think of the research question in the following form:

What is the effect of A on B?

In such questions, A represents an **explanatory variable** and B represents a **response variable**.

The first part of the question identifies the **explanatory variable** which we manipulate – we change it intentionally. The second part of the question identifies the **response variable** that we observe, looking for differences.

For example, we could conduct an experiment to investigate the effect of a gasoline additive on fuel efficiency. Fuel efficiency is something we all want when we drive our cars. It refers to how far we can drive on a tank of gas. In this experiment, the gasoline additive is the explanatory variable. The response variable is fuel efficiency. The additive is applied and we test for differences in fuel efficiency.

We might choose to compare fuel efficiencies for two groups of similar cars. We could create one group of cars using the gas additive and another group without it. Then we could gather data to compare the fuel efficiency of cars in the different groups.

We will conclude this lesson by considering three different experiments.

(8) Consider the following description of an experiment.[10] In this experiment, the researchers wanted to investigate whether the way people dry their hands after washing them has an effect on how clean their hands are.

An experiment was conducted to compare bacteria reduction for three different hand drying methods: using paper towels, using a hot air dryer, and allowing evaporation. In this experiment the participants handled uncooked chicken for 45 seconds. The participants then washed their hands with a single squirt of soap for 60 seconds and finally used one of the three hand drying methods. After participants dried their hands, the researchers measured the bacteria count on the participants' hands.

(a) What is the explanatory variable in this experiment? How many treatments were there in this experiment?

(b) What is the response variable in this experiment?

(c) Researchers *control* variables in an experiment when they want to prevent variables from having an impact on the response variable. Control variables are held *constant* throughout the experiment. The researchers chose to control a number of variables other than the hand-drying method that might have influenced the response variable. How did they do this?

(9) A researcher at King's College in London found that infomania (information overload) has a temporary, negative effect on intelligence quotient (IQ).[11] Imagine that a group of researchers wished to further test this conclusion. These researchers divided volunteers into two groups. Each subject took an IQ test. One group had to check email and respond to instant messages while they were taking the test. The other group took the IQ test without any distractions. Researchers found that the distracted group's average IQ test score was 10 points lower than the average IQ test score for the group that was not distracted.

(a) What is the response variable in this experiment?

(b) What is the explanatory variable, and what are its values (the individual treatments)?

(c) *Random assignment* is a feature of an experiment where the researchers randomly assign subjects to be in treatment groups. Random assignment ensures that each treatment group contains similar subjects. Explain why it would be good for the researchers to use random assignment to put each volunteer in one of the experimental groups. Why should the researchers do this rather than letting the volunteers decide which group they wanted to be in?

(d) A *control group* is often used in an experiment to establish baseline results and provide a means for examining the effect of a treatment. Identify the *control group* in this experiment.

(10) Read the following summary of an experiment published in the New England Journal of Medicine.[12] Think about the following question:

> *Do ethnic group and sex influence the type of care that a heart patient receives?*
>
> Previous research suggested that doctors are less likely to treat African-Americans and females with cardiac catheterization or coronary bypass surgery than whites and males. Cardiac catheterization and coronary bypass surgery are two treatments for chest pain or heart attacks. Scientists thought this situation was due to differences in illness severity (how serious the illness was), insurance coverage, patient preference, and health care access.
>
> To determine whether ethnic group and sex influenced the type of care that a heart patient receives, researchers visited either the meetings of the American College of Physicians or the American Academy of Family Physicians. They asked 720 primary care doctors at the meetings to watch a video of a patient and recommend care. The doctors thought the study was about how to make decisions while taking care of patients.
>
> Researchers created the eight videos by videotaping actors who described chest pain. The actors had memorized the lines from identical scripts. The actors were two black males, two black females, two white males, and two white females. The actors wore identical clothing, used identical gestures, and were taped from the same position.
>
> The researchers randomly assigned one of the videos to each of the doctors to watch. They did this by having the doctor roll a die with eight sides. For example if the doctor rolled a four, they watched video number four that featured a white female. After the video the doctor was asked what treatment they would recommend for the patient on the video.

(a) What are the explanatory and response variables in this experiment?

(b) In the experiment described, the researchers chose to control a number of variables other than ethnic group and sex that might have influenced the recommended treatment. How did they do this?

(c) Did the researchers use random assignment to create the experimental groups? If so, how did they do this?

---

[9] https://jech.bmj.com/content/72/6/471

[10] A.M. Snelling et al., “Comparative Evaluation of the Hygienic Efficacy of an Ultra-Rapid Hand Dryer vs. Conventional Warm Air Hand Dryers,” Journal of Applied Microbiology 110, No. 1 (January 2011): 19-26.

[11] http://www.drglennwilson.com/Infomania_experiment_for_HP.doc

[12] Kevin Schulman et al., “The Effect of Race and Sex on Physicians’ Recommendations for Cardiac Catheterization,” New England Journal of Medicine 340, no. 8 (1999): 618-26.

This page titled 6.3: Corequisite- Four-Step Process, Interpreting Statistical Studies, Understanding Experimental Design is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.3.1: Practice Problems Corequisite S.3

Examining Children's Stereotypes of Scientists

Children's perceptions of who a scientist is and what they look like can affect how they view science and their ability to do science. Historically, when children were asked to draw a scientist, children drew a man. In a landmark study of nearly 5000 children from the U.S. and Canada published in 1983, when asked to draw a scientist, 99.4% of children drew a man. These results, based on drawings collected between 1966 and 1977, provided researchers a clear image of children's stereotypes of scientists at that time. A *stereotype* is a thought or belief that is widely held about a particular type of person or thing, but is typically an oversimplification, and is often not true.

Since the 1970's the draw-a-scientist activity has been repeated in numerous research settings. Recently, researchers were interested in whether children's stereotypes of the sex of scientists has changed over time and whether stereotypes are different among boys and girls. They performed an analysis of data from 78 research studies that occurred from 1971 to 2016. The research studies included data on 20,860 drawings of scientists in which the sex of the scientist could be determined from children in kindergarten to grade 12. The studies were conducted in a variety of educational settings among children from a wide range of backgrounds.

The study found that: (a) the percentage of all children who draw a man in the draw-a-scientist activity has tended to decrease over time, and (b) the percentage of girls who draw a man has decreased more rapidly than the percentage of boys who draw a man.

(1) What are the research questions being asked?

(2) What is the population of interest?

(3) The *subjects* of this statistical study are children in the U.S. and Canada. What variables are being examined for each subject in the study?

The study found that girls' perceptions of scientists being male are changing more rapidly than those of boys. Using data from the study on children's drawings from 1985 to the present, we can create an equation that predicts how the percentage of girls who draw a man when asked to draw a scientist will change over time. Suppose we obtain the equation

$$G = -0.6t + 67$$

where G is the predicted percentage of all girls who draw a man when asked to draw a scientist and t is the number of years since 1985.

What percentage of girls drew a man:

(4) In the year 1985?

(5) In the year 1990?

(6) In the year 2000?

(7) In the year 2015?

(8) In what year will the predicted percentage of girls who draw a man when asked to draw a scientist be 40%?

This page titled 6.3.1: Practice Problems Corequisite S.3 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.4: Corequisite- Characteristics of Samples and Sample Data

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- a sample is a subset of a population.
- data from samples can be used to draw conclusions about populations.

By the end of this lesson, you will be able to

- compute sample means, sample medians, and sample proportions.
- construct and interpret dotplots and bar graphs.
- determine relationships between variables based on sample data in a statistical study.

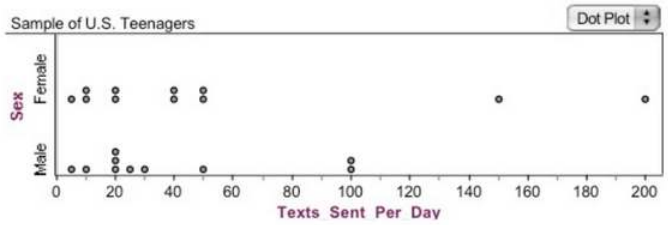
PROBLEM SITUATION 1: NUMBER OF TEXT MESSAGES SENT PER DAY

Do young females text more than young males? The Pew Research Center released a study in April 2015 titled Teens, *Social Media & Technology*[13] which presents data on U.S. teenagers that can be used to answer this question. The Pew Research Center is a non-partisan organization that conducts public opinion polling and research on U.S. politics and policy, journalism and media, the Internet, science and technology, and social and demographic trends in the U.S. The survey was administered online in late 2014 and early 2015 to a nationally representative sample of 1,060 U.S. teenagers. Each survey respondent provided information about their sex, age, and average number of text messages sent per day. The table below displays the results of a small sample of teenagers.

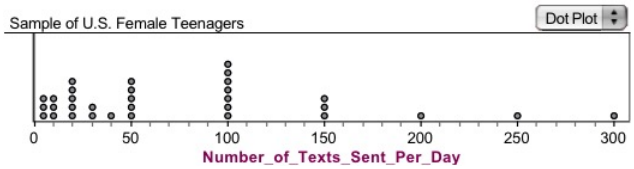
Table 1: Survey Results from a Sample of Teenagers in the U.S.

Age	Sex	Text Messages Sent Per Day		Age	Sex	Text Messages Sent Per Day
16	Male	5		17	Male	20
15	Male	100		17	Male	100
15	Male	20		15	Male	20
16	Female	20		15	Male	25
16	Female	20		15	Male	50
13	Female	40		16	Female	200
16	Female	50		14	Female	150
14	Male	10		13	Female	5
13	Male	30		16	Female	10
14	Female	50		17	Female	40
16	Female	10				

1. How could we use the sample data in the table above to determine whether young females text more than young males?
2. Below is a dotplot of the number of text messages sent per day by male and female teenagers in this sample. How do the values in the female and male samples compare?



3. The *mean* of a set of data values is the arithmetic average, the sum of data values divided by the number of data values. Determine the mean for the sample of females and for the sample of males. These values are called *sample means*. Round the sample means to the nearest whole number.
4. The *median* of a set of data values is the middle of a sorted list of data values. If the sorted list of data values has an odd number of values, the median is the middle value. If the sorted list of data values has an even number of values, the median is the mean (average) of the middle two values. Determine the median of each group. These values are called *sample medians*.
5. A data value is considered *unusual* if it is very different from the majority of the values in its distribution. We will examine statistical procedures for labeling values as unusual in later lessons. Look at the dotplot again. Do any values appear to be unusual in the dotplots? If so, explain.
6. The dotplot below contains data from a larger sample of female teenagers from the study. The dotplot contains 32 data values. Find the median of this dotplot.



7. If we want to estimate the typical number of text messages sent each day by female teenagers in the U.S., which sample result should we use: the sample median from the first sample of females, or the sample median from the second sample of females? Justify your answer.

PROBLEM SITUATION 2: ACCESSING THE INTERNET AT HOME

In addition to using sample means and medians to draw conclusions about populations, we can also use sample proportions to draw conclusions about populations. Sample proportions are found by counting the number of people in a sample that have a specific characteristic, and dividing this number by the sample size. Sample proportions are typically expressed as decimals.

The Pew Research Center conducted a survey[14] in April 2018 to gather information on U.S. adults' views about the Internet and the ways in which U.S. adults connect to the Internet. The survey was conducted via telephone interviews (landline and cellular) with U.S. adults, age 18 years or older, across all 50 states and the District of Columbia.

The survey asked respondents how they connected to the Internet at home. Respondents were asked whether they subscribed to a broadband Internet connection (a high-speed Internet connection using DSL, cable, or fiber optic service), accessed the Internet from home via their smartphone, or did neither. The results, separated according to the age group of the respondent, are displayed in the table below.

Table 2: How do you access the Internet at home?

Age Group	Broadband at home	Smartphone, No Broadband	No Smartphone, No Broadband	Total
18 - 29	236	99	17	352
30 - 49	368	126	34	528
50 - 64	369	86	89	544
65 +	263	55	211	529

8. Let's examine this sample data by computing sample proportions. Round each answer below to two decimal places.

(a) What proportion of the respondents are in the age 18 - 29 age group?

(b) What proportion of the respondents have broadband Internet access at home?

(c) What proportion of adults age 50 - 64 have a smartphone but no broadband Internet access at home?

9. For each age group, identify one sample proportion, round the sample proportion to two decimal places, and specify the population characteristic that the sample proportion can be used to estimate. Round answers to two decimal places.

10. Based on the sample data, is there a relationship between an adult's age group and the type of Internet access they have at home? Justify your answer using sample proportions.
11. Can we conclude that being an older adult in the U.S. (i.e. being age 65 or older) causes a person to be less likely to have broadband or smartphone Internet access at home? Why or why not?
12. Would it be reasonable to conclude that among all U.S. adults, those in the 18 - 29 age group have greater access to broadband Internet at home than those in the 50 - 64 age group? Explain your answer.
13. The statistical study above was conducted on 1,953 adults across the U.S. in April 2018. Imagine the same survey was administered to a different random sample of 2,000 adults across the U.S. during the same time period. Do you think that both surveys, the actual one and the hypothetical one, would yield identical results?

---

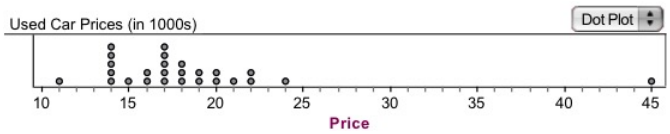
[13] http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/

[14] http://www.pewinternet.org/2018/04/30/declining-majority-of-online-adults-say-the-internet-has-been-good-for-society/

This page titled 6.4: Corequisite- Characteristics of Samples and Sample Data is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.4.1: Practice Problems Corequisite S.4

The dotplot contains the sales prices (in thousands) of 26 used cars in a used car dealership. (Hint: The minimum value in the data set is $11,000 and the maximum value is $45,000.)



(1) Find the sample mean of the data set. Round the mean to the nearest dollar.

(2) Find the sample median of the data set.

Suppose a researcher collected the following data from a random sample of students at a college. The sample is representative of all students at the college. Each survey respondent provided the following data:

- **Demographic**: Sex, Age, Major
- **Political Preferences**: Support raising the minimum wage of all workers to $15 per hour (Yes/No)

The survey data are displayed below in Table 3. Each row in the table corresponds to a survey respondent.

Table 3: Sample of Students at a College

Sex	Age	Major	Minimum Wage
M	22	Allied Health	Y
F	20	Humanities	N
F	28	Science	Y
M	19	Science	Y
M	24	Humanities	N
F	26	Allied Health	Y
F	22	Allied Health	Y
M	21	Journalism	N
F	21	Science	Y
M	23	Humanities	Y
F	25	Allied Health	Y
F	22	Journalism	Y

(3) Determine the mean of the ages of the students in the sample. Round to two decimal places.

(4) Determine the proportion (percentage) of college students in the sample who support raising the minimum wage of all workers $15 per hour. Write your answer as a percentage.

The “American dream” is an ideal that, in America, any person can succeed and prosper if they work hard, and any person can be as successful as their abilities and talents will take them.

Let's imagine a sample of U.S. adults were surveyed on their views about opportunity and wealth in the U.S. Respondents were asked, “Do you believe your family has achieved the American dream?”. The table below summarizes the results of the respondents, grouped according to their educational attainment. A person’s educational attainment refers to the highest level of education which the person has achieved. Table 4 displays percentages.

Table 4: Do you believe your family has achieved the American dream?

Educational Attainment	Has already achieved it	On way to achieving it	It is out of reach
College graduate or more	49	41	9
Some college	35	49	15
High school or less	27	48	24

(5) What does the “9” in the table represent?

(6) Based on the sample data, is there a relationship between a person’s educational attainment and their belief that their family has achieved the American Dream?

There is a _____ relationship between these two variables. The sample data indicates

that as a person’s educational attainment increases, the person is _____ likely to believe that their family has achieved the American dream.

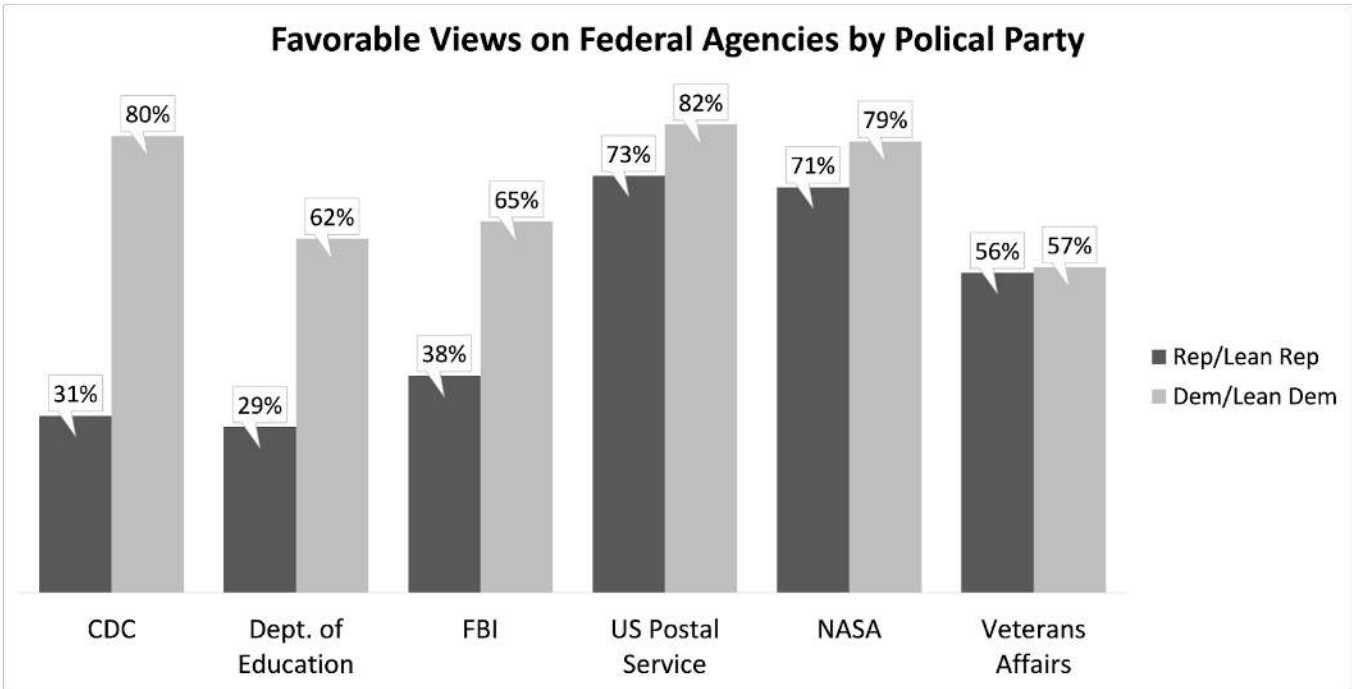
(7) What generalizations to the population(s) of interest can you make based on these survey results?

We can infer that among all U.S. adults, people who attain a higher educational level are

_____ to believe that their family has achieved the American dream.

In 2023 the Pew Research Center surveyed U.S. adults[15] to gather information on how the general public views the work of federal agencies. Pew surveyed 10,701 U.S. adults asking them to rate 16 agencies as favorable, unfavorable, or not sure. The political affiliation of respondents was also gathered and used to describe partisan gaps in American’s views of federal agencies.

Some of the survey results are displayed in the bar graph below. Use the bar graph to answer the questions that follow.



(8) Interpret the sample result of 82%.

(9) Interpret the sample result of 38%.

(10) Which pair of sample proportions provides the **greatest** evidence that a difference exists in the views of all Republicans and Democrats? Explain.

(11) Which pair of sample proportions provides the **least** evidence that a difference exists in the views of all Republicans and Democrats? Explain.

---

15 https://www.pewresearch.org/short-reads/2023/03/30/americans-feel-favorably-about-many-federal-agencies-especially-the-park-service-postal-service-and-nasa/

This page titled 6.4.1: Practice Problems Corequisite S.4 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.5: Corequisite- Distribution, Shape, and Terminology

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- a distribution of data can be described based on its center and variability.
- a distribution of data can be used to explore a real-world situation.

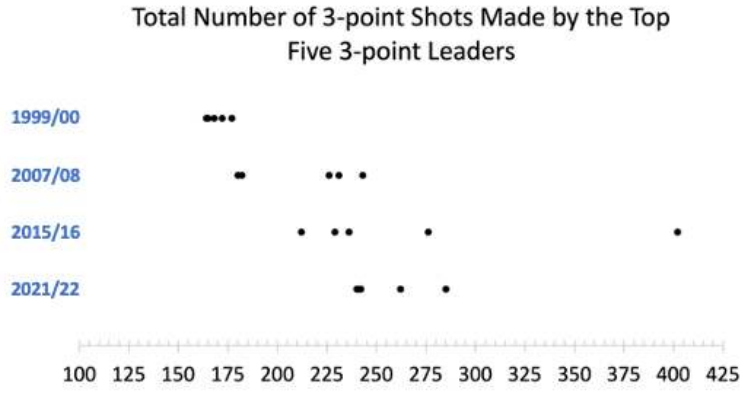
By the end of this lesson, you will be able to

- compute and interpret the sample standard deviation of a data set.
- determine relationships between variables based on sample data in a statistical study.

PROBLEM SITUATION 1: COMPARING SAMPLES

In this lesson we will examine an important concept in statistics: *variability*. The variability (or *spread*) of a data set refers to how widely data values are scattered in a data set. The greater the variability, the farther the values are from the center. The lesser the variability, the closer the values are to the center. The measure of variability we will examine is the *standard deviation*. The standard deviation describes the typical deviation from the mean.

Stephen Curry, the star point guard of the Golden State Warriors, is credited for "changing the game" in the National Basketball Association (NBA). Through his exemplary 3-point shooting, Curry has altered how NBA players, fans, team owners, and commentators view the value and importance of the 3-point shot. The number of 3-point shots attempted and made in the current NBA signifies a real shift in how the game is being played. The stacked dotplot below displays the total number of 3-point shots made by the top five 3-point leaders in the NBA during the 1999-2000, 2007-2008, 2015-2016, and 2021-2022 regular seasons.



1. Use the data from the top five 3-point leaders in these four seasons to describe how 3-point shot-making has changed over time.
2. Each sample contains five data samples. **Estimate** the sample mean x of each sample. Estimate the sample means by visually examining the stacked dotplot.

A Mean for 1999/00 sample:

B Mean for 2007/08 sample:

C Mean for 2015/16 sample:

D Mean for 2021/22 sample:

3. How do these measures of center compare?
4. Which sample appears to have greater variability? Variability refers to the spread of the data.

5. Let's examine how the values in the **2016** sample vary from the mean. Compute the standard deviation of the top eight 3-point totals in the 2016 season. The values from the 2016 sample are in the first column. You are going to work through and complete some missing values in the second and third columns. The second column contains the deviation of each value from the mean. The third column contains the squared deviation for each value in the data set.

Value	**Value − Mean**	**(Value − Mean)^2**
402	154	23716
280		
238		
233		
216		
211		
203		
200		

(6) (a) Now, we're going to compute the standard deviation of the data set. We'll do this in two steps. First, add the squared deviations in the third column together and divide the resulting sum by (the sample size minus 1). This value is called the sample variance. It is a "rough" average of the squared deviations.

The sample variance of the 2016 data set is _____.

(b) Now take the square root of your answer to part (a). This is the standard deviation of the sample. Round the value to the nearest whole number.

The standard deviation of the 2016 data set is _____.

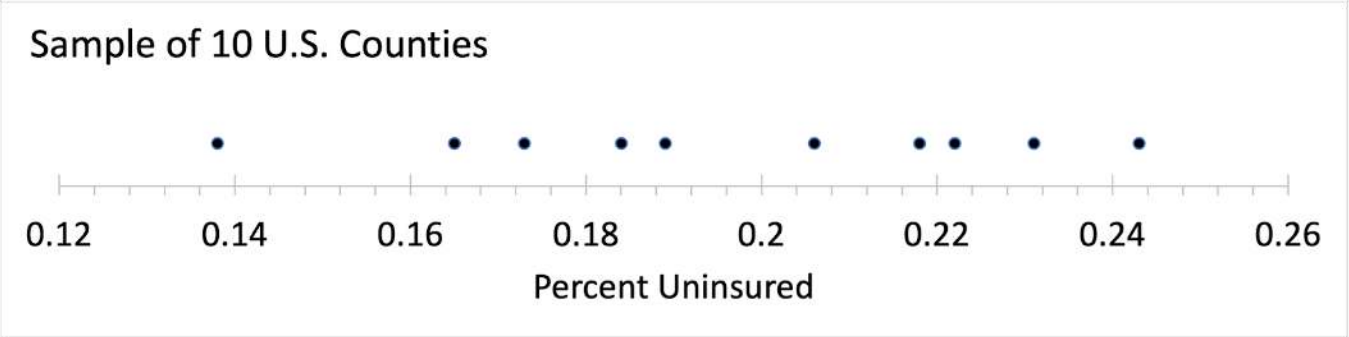
(7) Which value in the 2016 sample had the greatest impact on the size of the standard deviation? Why is this the case?

(8) Let's say the mean of a sample of data values for a particular season is 216 and the sample standard deviation is 38. Interpret this standard deviation. What does this tell us about these eight values in the sample?

Data on Counties in the U.S.

In the U.S., each state is divided into non-overlapping geographical regions called counties. Delaware, one of the smallest states in the U.S, has 3 counties. California, one of the largest states in the U.S., has 58 counties. Counties across the U.S. have different characteristics, such as their geographical size, population, political affiliation, and demographics. The U.S Census Bureau collects demographic and economic data on counties in the U.S. to understand how counties are similar and different on a variety of demographic, social, health and economic indicators. Collecting data at the county level enables counties to be the *unit of analysis* in a statistical study.

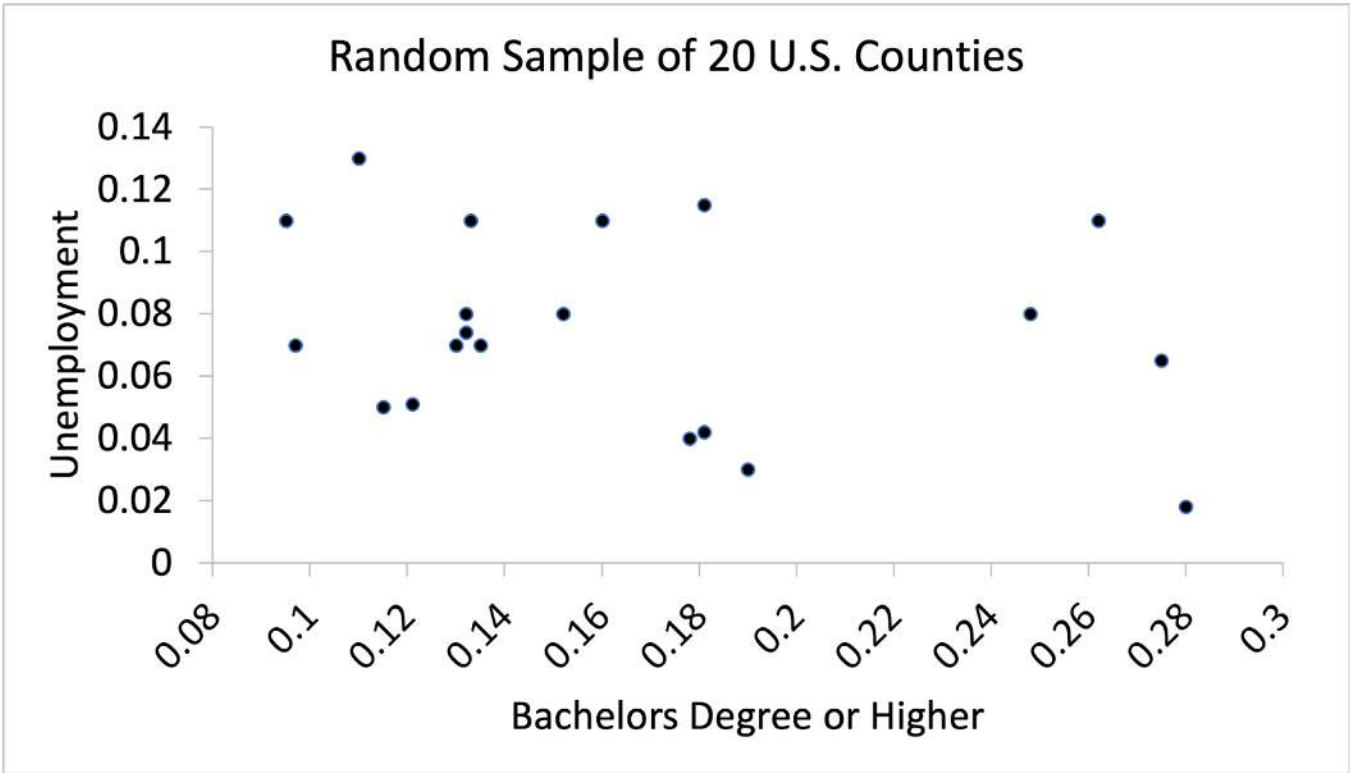
(9) One important indicator of the social and economic well-being of a county is the rate of individuals in the county who are without health insurance. The dotplot below displays the percentage of people who are uninsured in 10 counties in the U.S. in 2020 as reported in an analysis published in 2022.[16] The mean uninsured rate in this sample is approximately 0.20. Thus, on average, 20% of people living in a county in this sample lacked health insurance.



1. The lowest uninsured rate in the sample is 0.138, or 13.8%. If this county contains 22,600 people, how many people in the county are uninsured? Round to the nearest person.
2. Would it be reasonable to conclude that the standard deviation of the ten data values is 0.10? Explain.

PROBLEM SITUATION 2: EXAMINING RELATIONSHIPS IN SCATTERPLOTS

The scatterplot below displays the relationship between two variables: percentage of adults in a county who have a bachelor's degree or higher, and percentage of adults in a county who are unemployed. The data was collected from twenty counties included in the report.[17]

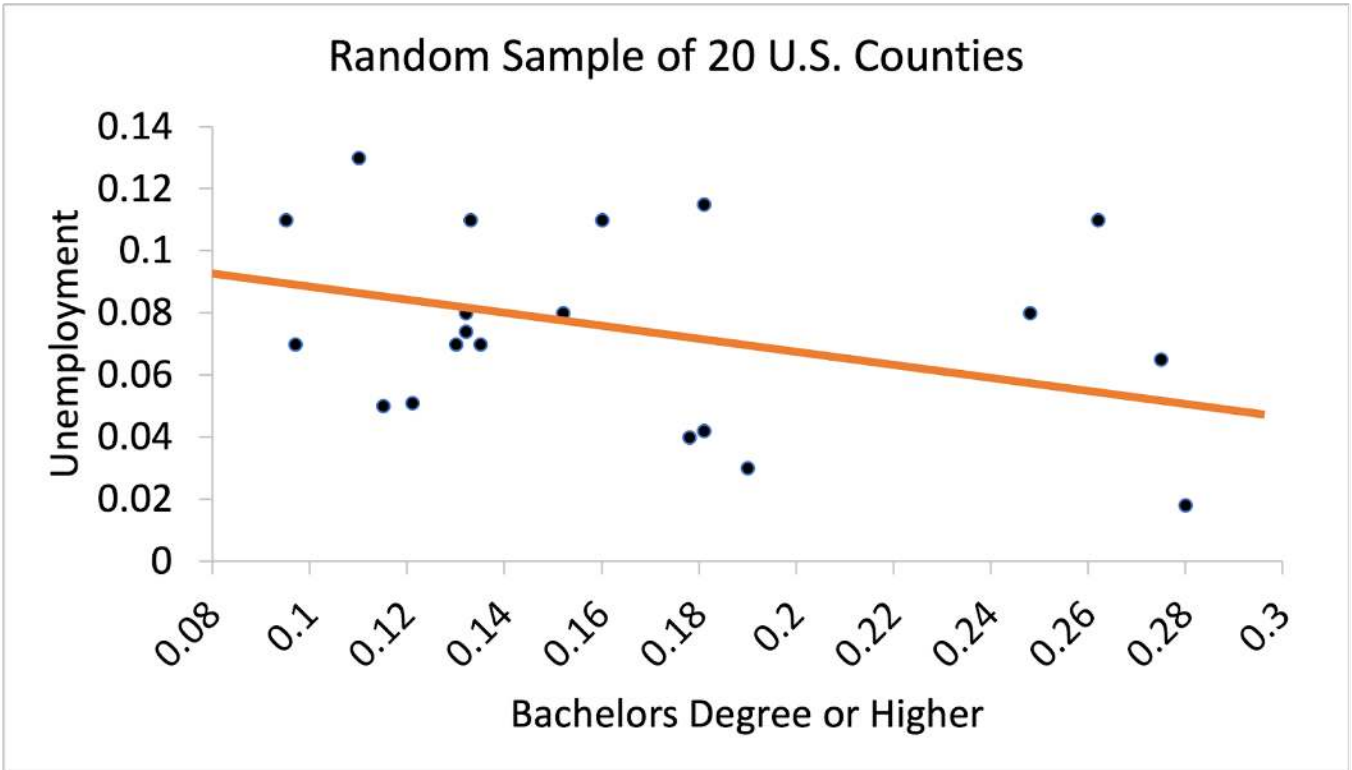


(10) What can you conclude about the relationship between these two variables based on this sample data?

(11) Without performing a calculation, determine which variable: percentage of adults in a county who have a bachelor's degree or higher, or percentage of adults in a county who are unemployed, has a **greater mean**. Explain.

(12) Without performing a calculation, determine which variable: percentage of adults in a county who have a bachelor's degree or higher, or percentage of adults in a county who are unemployed, has a **greater standard deviation**. Explain.

(13)



(a) Suppose we draw a line containing the points (0.10, 0.09) and (0.28. 0.06). Calculate the slope using these two points. Round your answer to two decimal places.

(b) Now, use the point (0.28. 0.06) to find the vertical intercept. Round your answer to the nearest thousandth (three decimal places).

(c) Now, let's use the slope and y-intercept to create our linear model. Let R be the unemployment rate of a county and B the percentage of adults in a county who have a bachelor's degree or higher.

In general, linear models are written as y = mx + b, where x is the input variable y is output variable m is the slope b is the vertical intercept In this problem, you're asked to use B for the input variable (the percentage of adults in a country who have a bachelor's degree of higher) and R for the output variable (unemployment rate). Write the linear model, and round the slope and the vertical intercept to two decimal places.

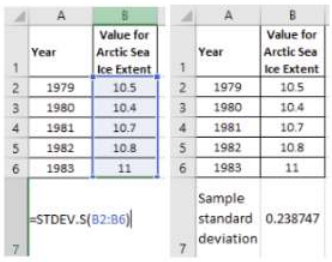
(d) Use the linear model to predict the unemployment rate of a county in which 23% of adults have a bachelor's degree or higher.

(e) Does your prediction seem reasonable? How could the prediction be improved? Explain.

Technology to the Rescue!

As you can tell, calculating sample standard deviation can be somewhat tedious. Fortunately, there's an easy way to do it using Excel or Google Sheets. It only takes a few clicks.

1. Enter data into Excel or Google Sheets
2. Click an empty cell and type in **=STDEV.S(**
3. Highlight the data you entered in Step 1 by holding the left mouse button and dragging
4. Press Enter



	A	B
1	Year	Value for Arctic Sea Ice Extent
2	1979	10.5
3	1980	10.4
4	1981	10.7
5	1982	10.8
6	1983	11
7	=STDEV.S(B2:B6)	

	A	B
1	Year	Value for Arctic Sea Ice Extent
2	1979	10.5
3	1980	10.4
4	1981	10.7
5	1982	10.8
6	1983	11
7	Sample standard deviation	0.238747

[16] https://www.census.gov/newsroom/press-releases/2022/uninsured-rates-decrease-sahie.html

[17] https://openpsych.net/paper/12

This page titled 6.5: Corequisite- Distribution, Shape, and Terminology is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.5.1: Practice Problems Corequisite S.5

The table below displays social and health data on a random sample of 11 counties in the United States. For each county, the table shows the percentage of adults who have a graduate degree, the percentage of adults who smoke cigarettes, and the percentage of children living with one parent. The data was collected from surveys in 2010 and analyzed in the 2016 report.[18]

Table 1: Random Sample of 11 Counties in the United States

County	**Percentage of Adults with a Graduate Degree**	**Percentage of Adults who Smoke Cigarettes**	**Percentage of Children Living with One Parent**
Newport News City, Virginia	8	23	43
Washington County, Florida	5	27	35
Warrick County, Indiana	9	14	22
San Mateo County, California	17	11	23
Cullman County, Alabama	5	22	27
Allegheny County, Pennsylvania	14	19	33
Bourbon County, Kansas	7	19	31
Bedford County, Virginia	9	19	23
Randolph County, Alabama	5	28	44
St. Bernard Parish, Louisiana	3	22	43
Hardin County, Ohio	6	31	29

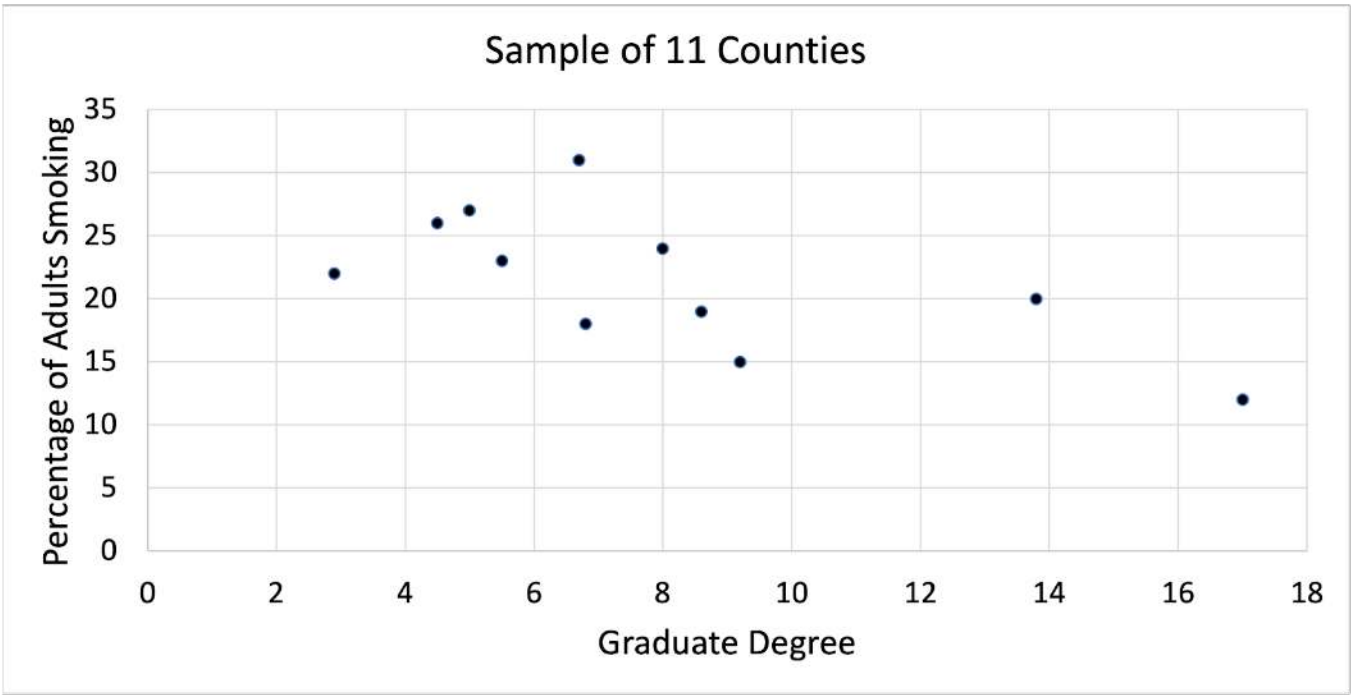
1. Which variable in the data set above has the **lowest variability**? Use the ranges or estimate the standard deviations of each variable.

(2) Which variable in the data set above has the **highest variability**? Use the ranges or estimate the standard deviations of each variable.

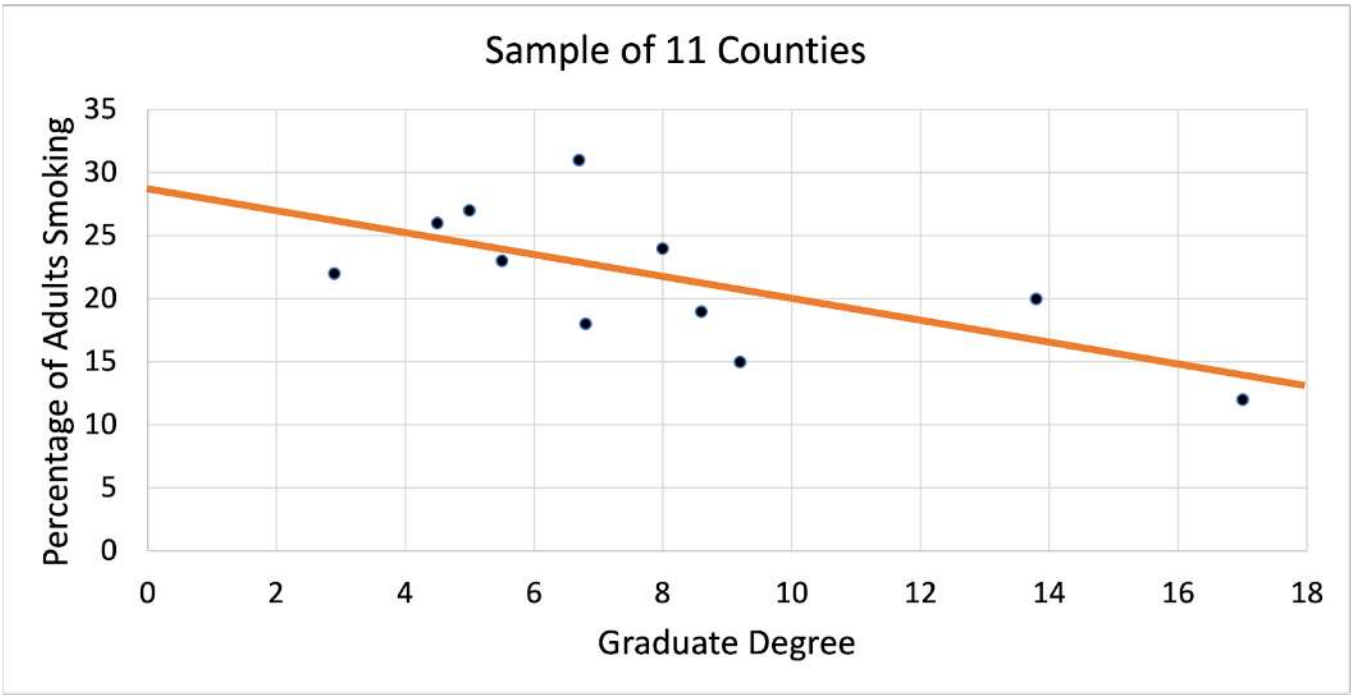
For **Questions 3-8**, compute the mean and standard deviation for each variable and complete the table. Round each answer to one decimal place.

	Percent of Adults with a Graduate Degree	Percent of Adults who Smoke Cigarettes	Percent of Children Living with One Parent
Mean	(3)	(4)	(5)
SD	(6)	(7)	(8)

(9) The scatterplot shows the percentage of adults in a county who have a graduate degree and the percentage of adults in a county who smoke cigarettes. What can you conclude about the relationship between these two variables based on the sample data?



(10) Develop a linear model to summarize the data. Use the points (0,29) and (17,12). Let S be the percentage of adults in a county who smoke and G be the percentage of adults in a county who have a graduate degree.



(11) Use the linear model to predict the percentage of adults who smoke cigarettes in a county if 12% of adults in the county have a graduate degree.

---

18 https://openpsych.net/paper/12

This page titled 6.5.1: Practice Problems Corequisite S.5 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.6: Corequisite- Computing Measures of Center; Computing and Interpreting Z-scores

SPECIFIC OBJECTIVES

By the end of this lesson, you should understand that

- distributions of data can be described based on their shape.
- distributions of data can be used to explore real-world situations.
- the Z-score of a data value can be used to assess whether a value is unusual within its distribution.

By the end of this lesson, you will be able to

- compute and interpret the mean, median and mode of a data set from its histogram.
- compute and interpret the Z-score of a data value within a bell-shaped distribution.

PROBLEM SITUATION 1: A SKEWED DISTRIBUTION

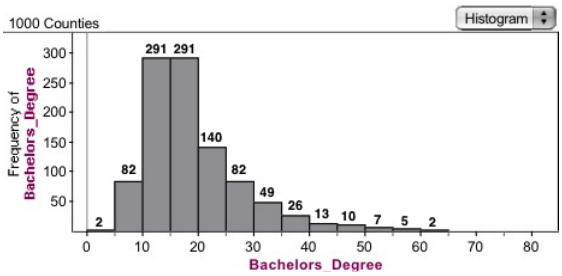
In this lesson we will examine shapes of distributions of quantitative data. The shape of a distribution describes how the values in the distribution vary. By knowing the shape of a distribution we can understand how the mean and median of the distribution compare. This is an important consideration when identifying whether the mean, median or mode best describes the center of a data set. The shape of a distribution also informs us on how we can use the distribution to determine unusual values.

County-Level Data

There are approximately 3150 counties in the U.S. The histogram below summarizes data for a random sample of 1000 counties. The sample contains about one-third of all U.S. counties. Since the random sample is sufficiently large, we can think of it as being representative of all counties in the U.S.

The following histogram is skewed. In a skewed distribution, the shape of the distribution is not symmetric so one side has a longer tail of values, and the mean and median are different. The histogram below displays the percentage of adults in a county who have a bachelor's degree or higher in 2010. This histogram is *skewed to the right.*

The number above each bar is the frequency of the group or values. The midpoint of each bar is the position on the horizontal axis at the center of the bar. For example, the midpoint of the first bar is 2.5 and the midpoint of the second bar is 7.5. The first bar can be interpreted as indicating that there are 2 counties in which 2.5% of the adults in the counties have a bachelor's degree or higher. The second bar can be interpreted as indicating that there are 82 counties in which 7.5% of adults in the counties have a bachelor's degree or higher.



(1) Describe the values in this histogram. What values are typical? What values are unusual? Explain in the context of the distribution.

(2) Since the histogram contains grouped data, we cannot use the individual data values to find the mean, median and mode. We must estimate these measures of center using the information provided in the histogram. The mean can be estimated by a weighted mean. Round the mean and median to two decimal places.

Mean =

Median =

The two modes are:

(3) Which measure: the mean, median, or mode, best represents the typical percentage of adults in a county who have a bachelor's degree or higher? Explain.

(4) What percentage of the counties in this sample have less than 20% of its adults having a bachelor's degree or higher? Write your answer as a percentage rounded to one decimal place.

(5) What percentage of the counties in this sample have between 20% and 50% of its adults having a bachelor's degree or higher?

(6) Is it unusual for a county to have a majority of adults with a bachelor's degree or higher? Explain.

(7) The mean of these data values is approximately 20 (meaning 20 percent) and the standard deviation is approximately 10 (meaning 10 percentage points). Each value in the data set has a deviation from the mean. The standard deviation tells us that the typical deviation from the mean is about 10 percentage points.

(a) Let's refer to the County-level Data on education. The mean of these data values is approximately 20% and the standard deviation is approximately 10%. How many data values are between 10% per and 30%? In other words, how many data values are within 1 standard deviation (10%) from the mean (20%)?

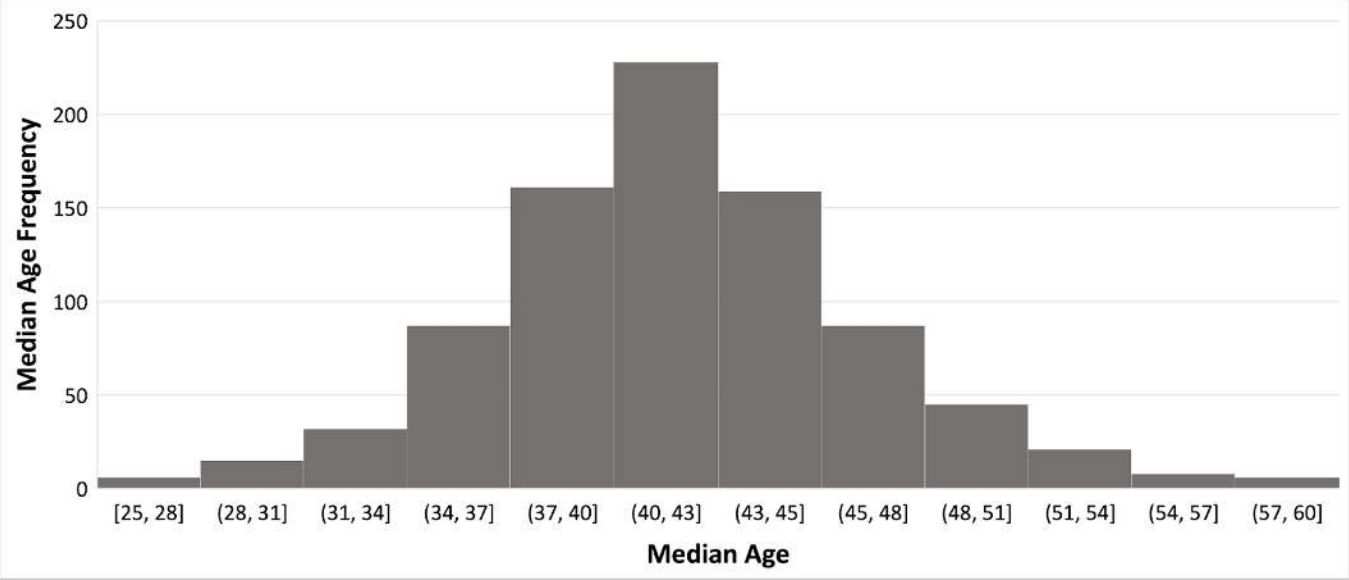
(b) How many data values are within 2 standard deviations (10%) from the mean (20%)? First, find the lower value. The mean minus 2 standard deviations is equal to _____%.

(c) How many data values are 2 standard deviations or more above the mean?

(d) What percentage of the counties in this sample are 2 standard deviations or more above the mean? Round your answer to the nearest percent.

PROBLEM SITUATION 2: EXAMINING BELL-SHAPED DISTRIBUTIONS

The highest level of education achieved is an example of a *categorical variable*. Categorical variables separate data into groups or categories, like "At Least a High School Degree" and "Bachelor's Degree or Higher."The median age of people living in a county is an example of a *quantitative variable*. Quantitative variables are variables whose values are obtained through counts or measurements. The histogram below displays the median age (in years) of people living in 1000 counties in the U.S. The data is based on the 2020 U.S. Census.[19]



This histogram is approximately *bell-shaped*. In a bell-shaped distribution, the mean and median are approximately equal, so the percentage of data values less than the mean is almost the same as the percentage of data values greater than the mean. When this occurs, the mean of the distribution is considered to be at the center of the distribution.

(8) (a) Without performing a calculation, estimate the mean of the histogram above.

(b) What values appear to be unusual? Explain.

(9) The standard deviation of the median ages in the sample above is 5 years. Interpret this standard deviation.

Z-scores

Statisticians often want to examine whether a data value is unusual within its distribution. When a distribution is bell-shaped and the mean and standard deviation are known, we can find the relative position of a data value in the distribution by calculating its *standardized score* or **Z-score**. A Z-score describes the number of standard deviations that a data value is from the mean. The Z-score formula is shown below.

$$Z - score = \frac{value - mean}{standard\ deviation}$$

Every data value in a distribution has a Z-score. Z-scores can be negative, positive, or zero. When a Z-score is less than or equal to −2, or greater than or equal to 2, we say that the corresponding data value is unusual, since the data value is relatively far from the mean. In fact, only 5% of data values in a bell-shaped distribution will have a Z-score less than or equal to -2, or greater than or equal to 2.

The table below contains the median ages (in years) of four counties in the U.S.[20]

Counties in the U.S.
County, State Median Age
Webb County, TX 29.6 Columbus County, NC 42.3 Franklin County, MA 47.2 Pickett County, TN 51.7

(10) For the 1,000 counties in the random sample, the mean of the median ages is 41.1 years and the standard deviation of the median ages is 5 years.

(a) Use these values to find the Z-score of each median age in the table above. Round Z-scores to two decimal places.

(b) Are any of these counties' median ages unusual? Explain.

(11) Another variable that has a bell-shaped distribution is the percentage of adults in a county who smoke cigarettes. This distribution has a mean of 21.5 (meaning 21.5%) and a standard deviation of 6 (meaning 6 percentage points). In one county in Alaska, 40.5% of adults smoke cigarettes.

Find the deviation from the mean and the Z-score of this data value. Round the Z-score to two decimal places.

(12) Dekalb County in Georgia had characteristics in 2020 which made it below average in comparison to other counties in the U.S. The median age of people in Dekalb County was 36.7 years. In Dekalb County, 10.8% of adults smoked cigarettes.

Characteristic	**Mean**	**Standard Deviation**	**Values in Dekalb County, GA**
Median age	41.1 years	5 years	36.7 years
Percentage of adult smokers	21.5%	6%	10.8%

Which of these characteristics is more unusual when Dekalb County is compared to other counties in the U.S.? Explain.

---

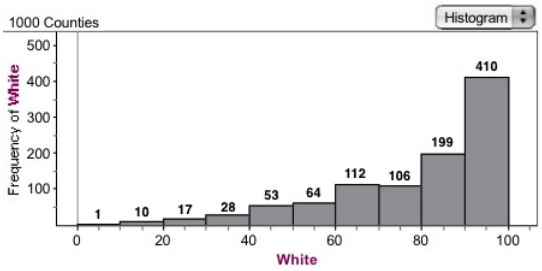
[19] https://www.census.gov/data/tables/time-series/demo/popest/2020s-counties-detail.html

[20] https://www.census.gov/data/tables/time-series/demo/popest/2020s-counties-detail.html

This page titled 6.6: Corequisite- Computing Measures of Center; Computing and Interpreting Z-scores is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

6.6.1: Practice Problems Corequisite S.6

The U.S. Census Bureau collects information on the racial composition of people in counties throughout the U.S. by asking survey respondents to identify their race. This information allows researchers to estimate the percentage of people in a county who are in various racial categories. The histogram below displays the percentage of adults who identified their race as white in 1,000 counties in the U.S. in 2020.[21] The number above each bar is the frequency of the group of values. The midpoint of each bar is the position on the horizontal axis at the center of the bar.



(1) What is the shape of the distribution: skewed to the right, skewed to the left, or bell-shaped?

(2) Use the histogram to *estimate* the mean. **Note**: The mean can be estimated by a weighted mean. Write your answer as a percentage rounded to two decimal places.

(3) Use the histogram to *estimate* the median. Write your answer as a percentage.

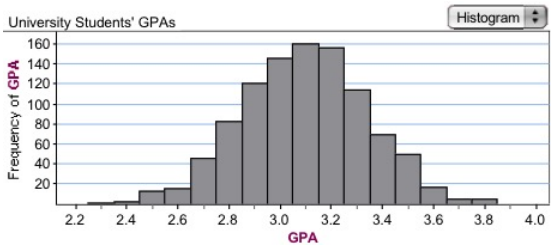
(4) Use the histogram to *estimate* the mode. Write your answer as a percentage.

(5) Which measure (mean, median or mode) best represents the typical percentage of adults who are white in this sample of counties? Explain.

(6) What percentage of the counties in this sample have 80% or more white residents? Write your answer rounded to one decimal place.

(7) What percentage of the counties in this sample have a majority of non-white residents? Write your answer rounded to one decimal place.

The GPAs of 1,000 students at a university have a bell-shaped distribution with a mean of 3.1 and standard deviation 0.25. The GPAs are summarized in the histogram below. We can use the histogram to show how the distribution is approximately symmetric to the mean.



(8) Find the approximate percentage of data values that are **0.3 or less** from the mean. Write your answer as a percentage rounded to the nearest whole percent. **Hint**: The percentage of data values that are 0.3 or less from the mean is the percentage of values that are 2.8 or less.

(9) Find the approximate percentage of data values that are **0.3 or more** from the mean. Write your answer as a percentage rounded to the nearest whole percent. **Hint**: The percentage of data values that are 0.3 or less from the mean is the percentage of values that are 3.4 or more.

Find and interpret the Z-scores of the GPAs of the following three students.

(10) Jayden: GPA = 3.4

(11) Lisa: GPA = 3.7

(12) Mark: GPA = 2.6

(13) Which of the three GPAs can be classified as unusual?

---

[21] https://mtgis-portal.geo.census.gov/arcgis/apps/MapSeries/index.html?appid=2566121a73de463995ed2b2fd7ff6eb7

This page titled 6.6.1: Practice Problems Corequisite S.6 is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Carnegie Math Pathways (WestEd) via source content that was edited to the style and standards of the LibreTexts platform.

Index

D

dire

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Quantway College with Corequisite (Carnegie Math Pathways)

Webpages: 137

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC 4.0: 98.5% (135 pages)
- Undeclared: 1.5% (2 pages)

By Page

- Quantway College with Corequisite (Carnegie Math Pathways) - *CC BY-NC 4.0*
  - Front Matter - *CC BY-NC 4.0*
    - TitlePage - *CC BY-NC 4.0*
    - InfoPage - *CC BY-NC 4.0*
    - Table of Contents - *Undeclared*
    - Licensing - *CC BY-NC 4.0*
  - 1: College Module N - *CC BY-NC 4.0*
    - 1.1: Introduction to Quantitative Reasoning - *CC BY-NC 4.0*
      - 1.1.1: Exercise N.1 - *CC BY-NC 4.0*
    - 1.2: Millions, Billions, or Trillions? - *CC BY-NC 4.0*
      - 1.2.1: Preparation N.2 - *CC BY-NC 4.0*
      - 1.2.2: Growth Mindset Activity - *CC BY-NC 4.0*
      - 1.2.3: Exercise N.2 - *CC BY-NC 4.0*
    - 1.3: A Trillion Dollars of Student Loans - *CC BY-NC 4.0*
      - 1.3.1: Preparation N.3 - *CC BY-NC 4.0*
      - 1.3.2: Exercise N.3 - *CC BY-NC 4.0*
    - 1.4: Computing and Interpreting Percentages - *CC BY-NC 4.0*
      - 1.4.1: Preparation N.4 - *CC BY-NC 4.0*
      - 1.4.2: Exercise N.4 - *CC BY-NC 4.0*
    - 1.5: The National Debt…How Big Is It? - *CC BY-NC 4.0*
      - 1.5.1: Preparation N.5 - *CC BY-NC 4.0*
      - 1.5.2: Exercise N.5 - *CC BY-NC 4.0*
    - 1.6: The Great Pacific Garbage Patch - *CC BY-NC 4.0*
      - 1.6.1: Preparation N.6 - *CC BY-NC 4.0*
      - 1.6.2: Exercise N.6 - *CC BY-NC 4.0*
    - 1.7: An Index for Your Salary - *CC BY-NC 4.0*
      - 1.7.1: Preparation N.7 - *CC BY-NC 4.0*
      - 1.7.2: Exercise N.7 - *CC BY-NC 4.0*
    - 1.8: Screening Tools–How Effective Are they? - *CC BY-NC 4.0*
      - 1.8.1: Preparation N.8 - *CC BY-NC 4.0*
      - 1.8.2: Exercise N.8 - *CC BY-NC 4.0*
    - 1.9: A Closer Look at Screening Tools - *CC BY-NC 4.0*
      - 1.9.1: Preparation N.9 - *CC BY-NC 4.0*
      - 1.9.2: Exercise N.9 - *CC BY-NC 4.0*
    - 1.10: Stress Reappraisal - *CC BY-NC 4.0*
  - 2: College Module M - *CC BY-NC 4.0*
    - 2.1: Mathematical Models–What's Best? - *CC BY-NC 4.0*
      - 2.1.1: Preparation M.1 - *CC BY-NC 4.0*
      - 2.1.2: Exercise M.1 - *CC BY-NC 4.0*
    - 2.2: Linear Models - *CC BY-NC 4.0*
      - 2.2.1: Preparation M.2 - *CC BY-NC 4.0*
      - 2.2.2: Exercise M.2 - *CC BY-NC 4.0*
    - 2.3: Data Is Trendy - *CC BY-NC 4.0*
      - 2.3.1: Preparation M.3 - *CC BY-NC 4.0*
      - 2.3.2: Exercise M.3 - *CC BY-NC 4.0*
    - 2.4: Four Equations…Two Models - *CC BY-NC 4.0*
      - 2.4.1: Preparation M.4 - *CC BY-NC 4.0*
      - 2.4.2: Exercise M.4 - *CC BY-NC 4.0*
    - 2.5: Modeling a New Type of Growth - *CC BY-NC 4.0*
      - 2.5.1: Preparation M.5 - *CC BY-NC 4.0*
      - 2.5.2: Exercise M.5 - *CC BY-NC 4.0*
    - 2.6: How long Does That Take? - *CC BY-NC 4.0*
      - 2.6.1: Preparation M.6 - *CC BY-NC 4.0*
      - 2.6.2: Exercise M.6 - *CC BY-NC 4.0*
    - 2.7: A Trust in Social Security - *CC BY-NC 4.0*
      - 2.7.1: Preparation M.7 - *CC BY-NC 4.0*
      - 2.7.2: Exercise M.7 - *CC BY-NC 4.0*
    - 2.8: Multiple Variables Weighing on Your mind? - *CC BY-NC 4.0*
      - 2.8.1: Preparation M.8 - *CC BY-NC 4.0*
      - 2.8.2: Exercise M.8 - *CC BY-NC 4.0*

- 3: College Module S - *CC BY-NC 4.0*
  - 3.1: Interpreting Charts and Graphs - *CC BY-NC 4.0*
    - 3.1.1: Preparation S.1 - *CC BY-NC 4.0*
    - 3.1.2: Exercise S.1 - *CC BY-NC 4.0*
  - 3.2: Measuring Success - *CC BY-NC 4.0*
    - 3.2.1: Preparation S.2 - *CC BY-NC 4.0*
    - 3.2.2: Exercise S.2 - *CC BY-NC 4.0*
  - 3.3: Statistical Analysis Process - *CC BY-NC 4.0*
    - 3.3.1: Preparation S.3 - *CC BY-NC 4.0*
    - 3.3.2: Exercise S.3 - *CC BY-NC 4.0*
  - 3.4: Sampling and Error - *CC BY-NC 4.0*
    - 3.4.1: Preparation S.4 - *CC BY-NC 4.0*
    - 3.4.2: Exercise S.4 - *CC BY-NC 4.0*
  - 3.5: Exploring Climate Data - *CC BY-NC 4.0*
    - 3.5.1: Preparation S.5 - *CC BY-NC 4.0*
    - 3.5.2: Exercise S.5 - *CC BY-NC 4.0*
  - 3.6: Exploring Climate Variability - *CC BY-NC 4.0*
    - 3.6.1: Preparation S.6 - *CC BY-NC 4.0*
    - 3.6.2: Exercise S.6 - *CC BY-NC 4.0*
  - 3.7: Modeling with Normal Distributions - *CC BY-NC 4.0*
    - 3.7.1: Preparation S.7 - *CC BY-NC 4.0*
    - 3.7.2: Exercise S.7 - *CC BY-NC 4.0*
  - 3.8: Significance in Observational Studies - *CC BY-NC 4.0*
    - 3.8.1: Preparation S.8 - *CC BY-NC 4.0*
    - 3.8.2: Exercise S.8 - *CC BY-NC 4.0*
  - 3.9: Significance in Experiments - *CC BY-NC 4.0*
    - 3.9.1: Preparation S.9 - *CC BY-NC 4.0*
    - 3.9.2: Exercise S.9 - *CC BY-NC 4.0*
- 4: Corequisite Module N - *CC BY-NC 4.0*
  - 4.1: Corequisite- Estimation Vs Calculation - *CC BY-NC 4.0*
  - 4.2: Corequisite- Conversion, Estimation, Percentages and Ratios, Scientific Notation - *CC BY-NC 4.0*
    - 4.2.1: Practice Problems Corequisite N.2 - *CC BY-NC 4.0*
  - 4.3: Corequisite- Interest, Doubling Time - *CC BY-NC 4.0*
    - 4.3.1: Practice Problems Corequisite N.3 - *CC BY-NC 4.0*
  - 4.4: Corequisite- Percentages and Ratios, Absolute and Relative Change, Rates - *CC BY-NC 4.0*
    - 4.4.1: Practice Problems Corequisite N.4 - *CC BY-NC 4.0*
  - 4.5: Corequisite- Percentages and Taxes, Ratios, Proportions and Dimensional Analysis - *CC BY-NC 4.0*
    - 4.5.1: Practice Problems Corequisite N.5 - *CC BY-NC 4.0*
  - 4.6: Corequisite- Dimensional Analysis, Unit Conversion - *CC BY-NC 4.0*
    - 4.6.1: Practice Problems Corequisite N.6 - *CC BY-NC 4.0*
  - 4.7: Corequisite- Buying Power, Consumer Price Index, Absolute and Relative Change, Reading and Interpreting Graphs - *CC BY-NC 4.0*
    - 4.7.1: Practice Problems Corequisite N.7 - *CC BY-NC 4.0*
  - 4.8: Corequisite- Percentages, Division, Extracting Relevant Information - *CC BY-NC 4.0*
    - 4.8.1: Practice Problems Corequisite N.8 and N.9 - *CC BY-NC 4.0*
- 5: Corequisite Module M - *CC BY-NC 4.0*
  - 5.1: Corequisite- Graphing Points, Linear Equations and Models - *CC BY-NC 4.0*
    - 5.1.1: Practice Problems Corequisite M.1 - *CC BY-NC 4.0*
  - 5.2: Corequisite- Solving Linear Equations, Slope - *CC BY-NC 4.0*
    - 5.2.1: Practice Problems Corequisite M.2 - *CC BY-NC 4.0*
  - 5.3: Corequisite- Trendlines - *CC BY-NC 4.0*
    - 5.3.1: Practice Problems Corequisite M.3 - *CC BY-NC 4.0*
  - 5.4: Corequisite- Linear Models Practice - *CC BY-NC 4.0*
    - 5.4.1: Practice Problems Corequisite M.4 - *CC BY-NC 4.0*
  - 5.5: Corequisite- Linear vs Exponential Models - *CC BY-NC 4.0*
    - 5.5.1: Practice Problems Corequisite M.5 - *CC BY-NC 4.0*
  - 5.6: Corequisite- Exponential Models (Practice and Application) - *CC BY-NC 4.0*
    - 5.6.1: Practice Problems Corequisite M.6 - *CC BY-NC 4.0*
  - 5.7: Corequisite- Multivariable Models - *CC BY-NC 4.0*
    - 5.7.1: Practice Problems Corequisite M.7 - *CC BY-NC 4.0*
- 6: Corequisite Module S - *CC BY-NC 4.0*
  - 6.1: Corequisite- Graphical Literacy - *CC BY-NC 4.0*

- 6.1.1: Practice Problems Corequisite S.1 - *CC BY-NC 4.0*
- 6.2: Corequisite- Graphical Literacy - *CC BY-NC 4.0*
  - 6.2.1: Practice Problems Corequisite S.2 - *CC BY-NC 4.0*
- 6.3: Corequisite- Four-Step Process, Interpreting Statistical Studies, Understanding Experimental Design - *CC BY-NC 4.0*
  - 6.3.1: Practice Problems Corequisite S.3 - *CC BY-NC 4.0*
- 6.4: Corequisite- Characteristics of Samples and Sample Data - *CC BY-NC 4.0*
  - 6.4.1: Practice Problems Corequisite S.4 - *CC BY-NC 4.0*
- 6.5: Corequisite- Distribution, Shape, and Terminology - *CC BY-NC 4.0*
  - 6.5.1: Practice Problems Corequisite S.5 - *CC BY-NC 4.0*
- 6.6: Corequisite- Computing Measures of Center; Computing and Interpreting Z-scores - *CC BY-NC 4.0*
  - 6.6.1: Practice Problems Corequisite S.6 - *CC BY-NC 4.0*

Back Matter - *CC BY-NC 4.0*

- Index - *CC BY-NC 4.0*
- Glossary - *CC BY-NC 4.0*
- Detailed Licensing - *CC BY-NC 4.0*
- Detailed Licensing - *Undeclared*

Detailed Licensing

Overview

Title: Quantway College with Corequisite (Carnegie Math Pathways)

Webpages: 137

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC 4.0: 98.5% (135 pages)
- Undeclared: 1.5% (2 pages)

By Page

- Quantway College with Corequisite (Carnegie Math Pathways) - *CC BY-NC 4.0*
  - Front Matter - *CC BY-NC 4.0*
    - TitlePage - *CC BY-NC 4.0*
    - InfoPage - *CC BY-NC 4.0*
    - Table of Contents - *Undeclared*
    - Licensing - *CC BY-NC 4.0*
  - 1: College Module N - *CC BY-NC 4.0*
    - 1.1: Introduction to Quantitative Reasoning - *CC BY-NC 4.0*
      - 1.1.1: Exercise N.1 - *CC BY-NC 4.0*
    - 1.2: Millions, Billions, or Trillions? - *CC BY-NC 4.0*
      - 1.2.1: Preparation N.2 - *CC BY-NC 4.0*
      - 1.2.2: Growth Mindset Activity - *CC BY-NC 4.0*
      - 1.2.3: Exercise N.2 - *CC BY-NC 4.0*
    - 1.3: A Trillion Dollars of Student Loans - *CC BY-NC 4.0*
      - 1.3.1: Preparation N.3 - *CC BY-NC 4.0*
      - 1.3.2: Exercise N.3 - *CC BY-NC 4.0*
    - 1.4: Computing and Interpreting Percentages - *CC BY-NC 4.0*
      - 1.4.1: Preparation N.4 - *CC BY-NC 4.0*
      - 1.4.2: Exercise N.4 - *CC BY-NC 4.0*
    - 1.5: The National Debt…How Big Is It? - *CC BY-NC 4.0*
      - 1.5.1: Preparation N.5 - *CC BY-NC 4.0*
      - 1.5.2: Exercise N.5 - *CC BY-NC 4.0*
    - 1.6: The Great Pacific Garbage Patch - *CC BY-NC 4.0*
      - 1.6.1: Preparation N.6 - *CC BY-NC 4.0*
      - 1.6.2: Exercise N.6 - *CC BY-NC 4.0*
    - 1.7: An Index for Your Salary - *CC BY-NC 4.0*
      - 1.7.1: Preparation N.7 - *CC BY-NC 4.0*
      - 1.7.2: Exercise N.7 - *CC BY-NC 4.0*
    - 1.8: Screening Tools–How Effective Are they? - *CC BY-NC 4.0*
      - 1.8.1: Preparation N.8 - *CC BY-NC 4.0*
      - 1.8.2: Exercise N.8 - *CC BY-NC 4.0*
    - 1.9: A Closer Look at Screening Tools - *CC BY-NC 4.0*
      - 1.9.1: Preparation N.9 - *CC BY-NC 4.0*
      - 1.9.2: Exercise N.9 - *CC BY-NC 4.0*
    - 1.10: Stress Reappraisal - *CC BY-NC 4.0*
  - 2: College Module M - *CC BY-NC 4.0*
    - 2.1: Mathematical Models–What's Best? - *CC BY-NC 4.0*
      - 2.1.1: Preparation M.1 - *CC BY-NC 4.0*
      - 2.1.2: Exercise M.1 - *CC BY-NC 4.0*
    - 2.2: Linear Models - *CC BY-NC 4.0*
      - 2.2.1: Preparation M.2 - *CC BY-NC 4.0*
      - 2.2.2: Exercise M.2 - *CC BY-NC 4.0*
    - 2.3: Data Is Trendy - *CC BY-NC 4.0*
      - 2.3.1: Preparation M.3 - *CC BY-NC 4.0*
      - 2.3.2: Exercise M.3 - *CC BY-NC 4.0*
    - 2.4: Four Equations…Two Models - *CC BY-NC 4.0*
      - 2.4.1: Preparation M.4 - *CC BY-NC 4.0*
      - 2.4.2: Exercise M.4 - *CC BY-NC 4.0*
    - 2.5: Modeling a New Type of Growth - *CC BY-NC 4.0*
      - 2.5.1: Preparation M.5 - *CC BY-NC 4.0*
      - 2.5.2: Exercise M.5 - *CC BY-NC 4.0*
    - 2.6: How long Does That Take? - *CC BY-NC 4.0*
      - 2.6.1: Preparation M.6 - *CC BY-NC 4.0*
      - 2.6.2: Exercise M.6 - *CC BY-NC 4.0*
    - 2.7: A Trust in Social Security - *CC BY-NC 4.0*
      - 2.7.1: Preparation M.7 - *CC BY-NC 4.0*
      - 2.7.2: Exercise M.7 - *CC BY-NC 4.0*
    - 2.8: Multiple Variables Weighing on Your mind? - *CC BY-NC 4.0*
      - 2.8.1: Preparation M.8 - *CC BY-NC 4.0*
      - 2.8.2: Exercise M.8 - *CC BY-NC 4.0*

- 3: College Module S - *CC BY-NC 4.0*
  - 3.1: Interpreting Charts and Graphs - *CC BY-NC 4.0*
    - 3.1.1: Preparation S.1 - *CC BY-NC 4.0*
    - 3.1.2: Exercise S.1 - *CC BY-NC 4.0*
  - 3.2: Measuring Success - *CC BY-NC 4.0*
    - 3.2.1: Preparation S.2 - *CC BY-NC 4.0*
    - 3.2.2: Exercise S.2 - *CC BY-NC 4.0*
  - 3.3: Statistical Analysis Process - *CC BY-NC 4.0*
    - 3.3.1: Preparation S.3 - *CC BY-NC 4.0*
    - 3.3.2: Exercise S.3 - *CC BY-NC 4.0*
  - 3.4: Sampling and Error - *CC BY-NC 4.0*
    - 3.4.1: Preparation S.4 - *CC BY-NC 4.0*
    - 3.4.2: Exercise S.4 - *CC BY-NC 4.0*
  - 3.5: Exploring Climate Data - *CC BY-NC 4.0*
    - 3.5.1: Preparation S.5 - *CC BY-NC 4.0*
    - 3.5.2: Exercise S.5 - *CC BY-NC 4.0*
  - 3.6: Exploring Climate Variability - *CC BY-NC 4.0*
    - 3.6.1: Preparation S.6 - *CC BY-NC 4.0*
    - 3.6.2: Exercise S.6 - *CC BY-NC 4.0*
  - 3.7: Modeling with Normal Distributions - *CC BY-NC 4.0*
    - 3.7.1: Preparation S.7 - *CC BY-NC 4.0*
    - 3.7.2: Exercise S.7 - *CC BY-NC 4.0*
  - 3.8: Significance in Observational Studies - *CC BY-NC 4.0*
    - 3.8.1: Preparation S.8 - *CC BY-NC 4.0*
    - 3.8.2: Exercise S.8 - *CC BY-NC 4.0*
  - 3.9: Significance in Experiments - *CC BY-NC 4.0*
    - 3.9.1: Preparation S.9 - *CC BY-NC 4.0*
    - 3.9.2: Exercise S.9 - *CC BY-NC 4.0*
- 4: Corequisite Module N - *CC BY-NC 4.0*
  - 4.1: Corequisite- Estimation Vs Calculation - *CC BY-NC 4.0*
  - 4.2: Corequisite- Conversion, Estimation, Percentages and Ratios, Scientific Notation - *CC BY-NC 4.0*
    - 4.2.1: Practice Problems Corequisite N.2 - *CC BY-NC 4.0*
  - 4.3: Corequisite- Interest, Doubling Time - *CC BY-NC 4.0*
    - 4.3.1: Practice Problems Corequisite N.3 - *CC BY-NC 4.0*
  - 4.4: Corequisite- Percentages and Ratios, Absolute and Relative Change, Rates - *CC BY-NC 4.0*
    - 4.4.1: Practice Problems Corequisite N.4 - *CC BY-NC 4.0*
  - 4.5: Corequisite- Percentages and Taxes, Ratios, Proportions and Dimensional Analysis - *CC BY-NC 4.0*
    - 4.5.1: Practice Problems Corequisite N.5 - *CC BY-NC 4.0*
  - 4.6: Corequisite- Dimensional Analysis, Unit Conversion - *CC BY-NC 4.0*
    - 4.6.1: Practice Problems Corequisite N.6 - *CC BY-NC 4.0*
  - 4.7: Corequisite- Buying Power, Consumer Price Index, Absolute and Relative Change, Reading and Interpreting Graphs - *CC BY-NC 4.0*
    - 4.7.1: Practice Problems Corequisite N.7 - *CC BY-NC 4.0*
  - 4.8: Corequisite- Percentages, Division, Extracting Relevant Information - *CC BY-NC 4.0*
    - 4.8.1: Practice Problems Corequisite N.8 and N.9 - *CC BY-NC 4.0*
- 5: Corequisite Module M - *CC BY-NC 4.0*
  - 5.1: Corequisite- Graphing Points, Linear Equations and Models - *CC BY-NC 4.0*
    - 5.1.1: Practice Problems Corequisite M.1 - *CC BY-NC 4.0*
  - 5.2: Corequisite- Solving Linear Equations, Slope - *CC BY-NC 4.0*
    - 5.2.1: Practice Problems Corequisite M.2 - *CC BY-NC 4.0*
  - 5.3: Corequisite- Trendlines - *CC BY-NC 4.0*
    - 5.3.1: Practice Problems Corequisite M.3 - *CC BY-NC 4.0*
  - 5.4: Corequisite- Linear Models Practice - *CC BY-NC 4.0*
    - 5.4.1: Practice Problems Corequisite M.4 - *CC BY-NC 4.0*
  - 5.5: Corequisite- Linear vs Exponential Models - *CC BY-NC 4.0*
    - 5.5.1: Practice Problems Corequisite M.5 - *CC BY-NC 4.0*
  - 5.6: Corequisite- Exponential Models (Practice and Application) - *CC BY-NC 4.0*
    - 5.6.1: Practice Problems Corequisite M.6 - *CC BY-NC 4.0*
  - 5.7: Corequisite- Multivariable Models - *CC BY-NC 4.0*
    - 5.7.1: Practice Problems Corequisite M.7 - *CC BY-NC 4.0*
- 6: Corequisite Module S - *CC BY-NC 4.0*
  - 6.1: Corequisite- Graphical Literacy - *CC BY-NC 4.0*

- 6.1.1: Practice Problems Corequisite S.1 - *CC BY-NC 4.0*
- 6.2: Corequisite- Graphical Literacy - *CC BY-NC 4.0*
  - 6.2.1: Practice Problems Corequisite S.2 - *CC BY-NC 4.0*
- 6.3: Corequisite- Four-Step Process, Interpreting Statistical Studies, Understanding Experimental Design - *CC BY-NC 4.0*
  - 6.3.1: Practice Problems Corequisite S.3 - *CC BY-NC 4.0*
- 6.4: Corequisite- Characteristics of Samples and Sample Data - *CC BY-NC 4.0*
  - 6.4.1: Practice Problems Corequisite S.4 - *CC BY-NC 4.0*
- 6.5: Corequisite- Distribution, Shape, and Terminology - *CC BY-NC 4.0*
  - 6.5.1: Practice Problems Corequisite S.5 - *CC BY-NC 4.0*
- 6.6: Corequisite- Computing Measures of Center; Computing and Interpreting Z-scores - *CC BY-NC 4.0*
  - 6.6.1: Practice Problems Corequisite S.6 - *CC BY-NC 4.0*

Back Matter - *CC BY-NC 4.0*

- Index - *CC BY-NC 4.0*
- Glossary - *CC BY-NC 4.0*
- Detailed Licensing - *CC BY-NC 4.0*
- Detailed Licensing - *Undeclared*