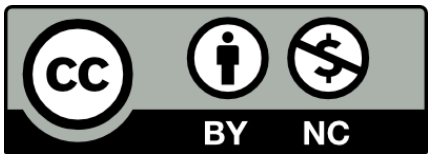
CARNEGIE MATH
PATHWAYS
WestEd



STATWAY® PATHWAY MODULE 1 v4.1

TABLE OF CONTENTS:

1.1: Setting Course Expectations and Creating Productive Classroom Norms

Statway® Cohort Contract

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

Specifically:

- I commit to successfully completing Statway with the members of my cohort.
- I commit to helping all of my cohort members understand mathematics and complete Statway.
- 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.

Printed Name:_____ Signature:_____ Date:_____

Printed Name:_____ Witness: _____ Date:_____

1.1-S: Understanding Percentages

LEARNING GOALS

By the end of this lesson, you should understand that:

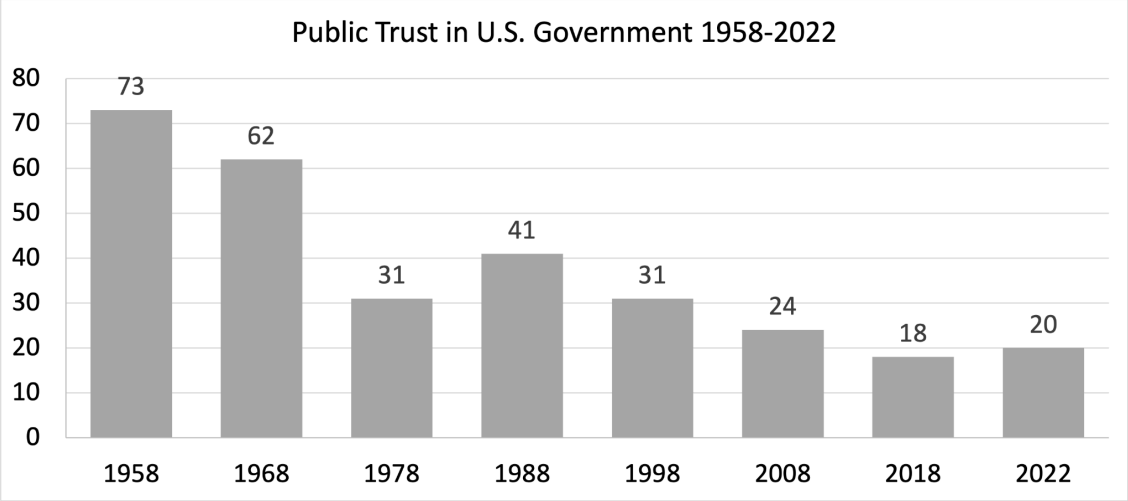
- Benchmark percentages (1%, 10%, 50%) can be used to obtain other percentages.
- Equivalent numbers can be represented as decimals, percents, and fractions.
- Converting numbers from one form to another can help us understand real-world situations.

By the end of this lesson, you should be able to:

- Use benchmark percentages (1%, 10%, 50%) to find other percentages of a number.
- Convert fractions to decimals and decimals to percentages.
- Find a percent of a number.
- Use a bar model to find a percentage of a number.
- Use fractions, decimals and percentages to solve real-world problems.

INTRODUCTION

Statistics is the science of collecting, describing, summarizing, and analyzing data to learn about our world. *Data* is information we collect from measurements and observations. Much of the data we collect and use we summarize with percentages. For example, the graph below shows how U.S. adults' trust in the federal government has changed over time[1]. The graph describes the percentage of U.S. adults who say that they trust the federal government in Washington, DC always or most of the time.



[1] https://www.pewresearch.org/politics/2022/06/06/public-trust-in-government-1958-2022/

1 Examine the graph and write a statement about the data. Use specific information from the graph.

2 What does the graph tell us about the year 2008?

3 Suppose we randomly surveyed 120 U.S. adults in the year 2008. How many of these adults do we expect would trust the federal government in Washington, DC always or most of the time? Try to answer this question without using a calculator.

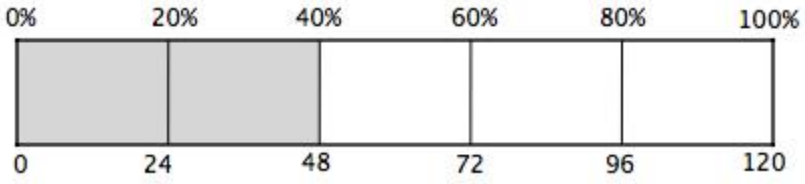
One strategy for finding a percentage of a number is to use a ***benchmark percentage***. Benchmark percentages are percentages of a number that can be determined by simple (often mental) calculations. Benchmark percentages include 100%, 50%, 25%, 20%, 10%, and 1%. Computing benchmark percentages enables us to find other percentages.

4 A Complete the following table. Find each benchmark percentage of 120. Round to the nearest person (whole number).

Benchmark Percentages	Value
100%	
50%	
10%	
1%	

B Explain how you could use the table to estimate 24% of 120.

We can use a ***bar model*** to visualize percentages. In a bar model we break up an amount into equal parts using benchmark percentages and shade in the portion of the bar that represents a specific percentage. The bar model below displays how 120 can be broken up into 20% regions. 40% of 120 is then shaded in.



5 Use a bar model to find and express 60% of 55.

6 Use a bar model to find and express 80% of 180.

YOU NEED TO KNOW

To convert a fraction to a decimal: Divide the numerator by the denominator in your calculator. For example, for the fraction 4/13, divide 4 by 13 in your calculator, and round the decimal to two decimal places. 4/13 = 0.31.

Convert a decimal to a percent: Once you have the decimal, move the decimal point two places to the right and write as a percentage. For example, 0.31 = 31%.

Finding the percent of a number: Change the percent to a decimal, and multiply the decimal by the number. For example, 31% of 80 is 0.31 x 80, which is 24.8.

TRY THESE

The following questions will help you practice your conversion skills. **Equivalent** means that numerical expressions are equal to each other. Complete the following questions about equivalent percentages, fractions, and decimals.

7 The previous study found that in the year 2018, 18% of U.S. adults trust the federal government always or most of the time.

A Write 18% as a fraction.

B Write 18% as a decimal.

For the next question, you will use a **mathematical expression**. A mathematical expression is when numbers, symbols, and operators are grouped together to show the value of something. An example of a mathematical expression is 2 + 4.

Refer to the graph from earlier showing the percentage of U.S. adults in different years who indicated that they trust the federal government in Washington, DC always or most of the time.

8 Imagine 1,500 U.S. adults were surveyed in 2018. You want to calculate how many would have said that they trust the federal government. Which mathematical expression could you use to find the answer? Select all that apply.

- ❑ $0.18 * 1500$
- ❑ $\frac{18}{100} * 1500$
- ❑ $(18\%)(1500)$
- ❑ $18\% \times 1500$

NEXT STEPS

We often use percentages to compare two numbers. When we express one number as a percentage of another number, we describe how one number is a portion of the other number.

Calculation of a percentage requires that one quantity be deemed the part and one quantity be deemed the whole. The quantity deemed the whole, which will be in the denominator of the fraction, is referred to as the reference value.

One issue that is important to many U.S. adults is taxation - the amount of federal and state taxes that Americans are required to pay each year. The table below displays the 2022 annual salaries and 2022 federal tax payments for three individuals.

Name	2022 Annual Salary	2022 Federal Tax Payment
John	36,200	2,800
Lisa	65,050	7,820
Erica	53,700	7,040

Use the data in the above table to answer the following questions.

9 Let's first look at John's 2022 annual salary and his 2022 federal tax payment. What percent of John's 2022 annual salary is his 2022 federal tax payment?

A First express your answer as a fraction

B Now express your answer as a decimal rounded to three decimal places.

C Now express your answer as a percent, rounded to one decimal place. For example, 0.123456 as a percent rounded to one decimal place would be 12.3%.

D Interpret the percentage. Explain this result in your own words.

10 Let's now look at Lisa's 2022 annual salary and her 2022 federal tax payment. What percent of Lisa's 2022 annual salary is her 2022 federal tax payment?

A First express your answer as a fraction.

B Now express your answer as a decimal rounded to three decimal places. This will give you a proportion.

C Now express your answer as a percent, rounded to one decimal place. For example, 0.123456 as a percent rounded to one decimal place would be 12.3%.

D Who paid a higher percentage of their annual salary towards federal taxes? Explain.

E Why might Lisa pay a higher percentage of her salary towards federal taxes?

11 Marcus' 2022 federal tax payment was $5,350. This payment was $\frac{1}{5}$ of his 2022 annual salary. What is his 2022 annual salary?

12 Kim's 2022 federal tax payment was $12,350. This payment was 15.6% of her 2022 annual salary. What is her 2022 annual salary?

LET'S SUMMARIZE

Remember that:

- Benchmark percentages (1%, 10%, 50%) can be used to obtain other percentages.
- Equivalent numbers can be represented as decimals, percents, and fractions.
- Converting numbers from one form to another can help us understand real-world situations.

Practice Problems 1.1-S

A study[2] in 2021 reported the following information:

- 59% of U.S. adults say that the feeling that some U.S. corporations do not pay their fair share of federal taxes bothers them a lot.
- 47% of U.S. adults say that the feeling that the U.S. tax system is too complex bothers them a lot.

1 If 4,000 U.S. adults were surveyed at the time of this study, how many of them would you expect to say that the feeling that some U.S. corporations do not pay their fair share of federal taxes bothers them a lot.

2 If 6,000 U.S. adults were surveyed at the time of this study, how many of them would you expect to say that the feeling that the U.S. tax system is too complex bothers them a lot.

3 The first row shows an example of conversion from a fraction (1/100) to a decimal (0.01) to a percent (1%). Please complete the other conversions accordingly.

Simplified Fraction	Decimal	Percent
$\frac{1}{100}$	0.01	1%
$\frac{1}{10}$		

[2] https://www.pewresearch.org/fact-tank/2021/04/30/top-tax-frustrations-for-americans-the-feeling-that-some-corporations-wealthy-people-dont-pay-fair-share/

4

Simplified Fraction	Decimal	Percent
	0.2	

5

Simplified Fraction	Decimal	Percent
		25%

6 Round the decimal to two decimal places. Enter the percent rounded to the nearest whole percent.

Simplified Fraction	Decimal	Percent
$\frac{1}{3}$		

7

Simplified Fraction	Decimal	Percent
	0.5	

8

Simplified Fraction	Decimal	Percent
$\frac{2}{3}$		

9

Simplified Fraction	Decimal	Percent
	0.75	

Now see if you can convert the fractions to decimals and then match them to the correct percentages.

10 Match the fractions to their equivalent percent form (rounded to the nearest percent).

(a) $\frac{1}{5}$	(i) 10%
(b) $\frac{2}{3}$	(ii) 20%
(c) $\frac{3}{4}$	(iii) 25%
(d) $\frac{1}{2}$	(iv) 33%
(e) $\frac{1}{3}$	(v) 50%
(f) $\frac{1}{4}$	(vi) 67%
(g) $\frac{1}{10}$	(vii) 75%

Answer the following:

11 **Estimate**: What percent is 37 of 105?

12 **Estimate**: What percent of 300 is 72?

13 **Calculate**: 16 is what percent of 42? Round to the nearest percent.

14 **Calculate**: What is 48% of 260? Round to the nearest whole number.

15 **Calculate**: What is 129% of 15?

Preparation 1.2

What is Statistics? Why Do We Study Statistics?

Statistics is about using **data** to answer questions. Data is information that we collect from our world. Data involves facts and observations that we make. Before scientists created statistics and before people used data, they would use opinions and hunches to explain how the world worked. A lot of times these explanations were wrong.

For example, people once believed that the sun revolved around the earth. When people started making observations and using measures, they discovered that this was incorrect. The earth revolves around the sun. Data helps us draw better conclusions.

Data can help us answer many types of questions.

- Students can use data to help pick a college that is best for them.
- Teachers use data to find the best ways to educate their students.
- Medical professionals use data to learn if new treatments actually work.
- Voters use data about their society and planet to create a better democracy.
- Politicians use data to better represent the people who elect them.

In statistics, we gather, summarize, and analyze data to search for answers to our questions. Let's begin with an exercise that involves using data to answer a question. Remember that data involves facts, observations, and measures about a particular topic or idea.

How significant is your birthday? Many people believe that habits, hobbies, likes, dislikes and many other things are affected by the time of year you were born. One such theory involves *chronotypes*. The term "chronotype" refers to a person's natural tendency to be most active and alert at certain times of the day. People who are most active early in the day are often labeled as morning people, early birds or larks. Those who are at their best in the evening are called night people or owls. Now imagine that you want to answer the following question: Can someone's birth month influence his or her chronotype?

INTRODUCTION

In Collaboration 1.2, you will be introduced to the statistical analysis process and learn how this process enables you to use data to make decisions. How significant is your birthday?

Many people believe that habits, hobbies, likes, dislikes and many other things are affected by the time of year you were born. One such theory involves chronotypes. The term "chronotype" refers to a person's natural tendency to be most active and alert at certain times of the day. People who are most active early in the day are often labeled as morning people, early birds or larks. Those who are at their best in the evening are called night people or owls. Now imagine that you want to answer the following

question: **Can someone's birth month influence his or her chronotype?** To answer this question, we could:

- Ask a group of people for their birthdate and ask about their chronotype.
- Examine whether people with certain birth months tend to have certain chronotypes. We are looking for a pattern in the data which shows a clear relationship between birth month and chronotype.

Before you perform this statistical investigation, you will complete an activity to identify your chronotype and check whether or not Chronotype Theory accurately predicts your chronotype based on your birth month.

Identify your Chronotype: Are you a morning person, a night person, or neither?

On a sheet of paper, for each characteristic in the table below, write down the number of the answer that best applies to you (1, 2, or 3). There are a total of 11 characteristics in the table, so you should write down 11 numbers.

Characteristic[3]	1	2	3
Most alert	Around noon	Around 6 PM	Equally alert at noon and 6 PM
Most productive	Morning	Evening	Equally productive in morning and evening
Alarm clock	Don't need it	Need multiple alarms	One alarm is enough
Favorite exercise time	Morning	Evening	Equally likely to exercise in morning and evening
Mood	Declines steadily all day	Rises steadily all day	Same throughout the day
Favorite meal	Breakfast	Dinner	I like breakfast and dinner equally
Coffee use	None or a small amount	Quite a bit	Somewhere in the middle
Class Time	I would prefer a class at 8 AM	I would prefer a class at 7 PM	No preference between a class at 8 AM or 7 PM
Morning behavior	Full of energy	Out of steam	Somewhere in the middle
Evening behavior	Out of steam	Full of energy	Somewhere in the middle
Travel	I get bad jet lag	Adapt quickly to time zone changes	I get a bit of jet lag

[3]Adapted from http://www.nasw.org/users/llamberg/larkowl.htm.

1 How many 1s did you choose?

2 How many 2s did you choose?

3 How many 3s did you choose?

Chronotype Selection

Look at your answers above and use the following to determine your chronotype:

- If you picked more 1s than other choices, this indicates that your chronotype is a morning person.
- If you picked more 2s than other choices, this indicates that your chronotype is a night person.
- If you picked more 3s than other choices, this indicates that your chronotype is neither a morning person nor a night person.

*Tiebreaker. If you have a tie between your two top columns, simply select the result with which you most identify.

Write down your chronotype from this activity on your paper. You will need it for the Collaboration.

Chronotype Theory

Chronotype Theory states that there are three chronotypes and four birth months that correspond to each chronotype. The theory states that being a (a) morning person, (b) night person or (c) neither/both is related to the month when you were born, and predicts your chronotype based on your birth month. The table below shows the predictions for each birth month.

Month	Chronotype	Month	Chronotype
January	Evening	July	Morning
February	Evening	August	Morning
March	Neither/Both	September	Neither/Both
April	Neither/Both	October	Neither/Both
May	Morning	November	Evening
June	Morning	December	Evening

Use the table to identify which chronotype matches your birth month, according to Chronotype Theory. Note that on your paper.

Did the chronotype indicated by the initial activity match the chronotype predicted by Chronotype Theory using your birth month only? Make a note on your paper whether the two chronotypes matched or did not match. You will use all of this information in Collaboration 1.2.

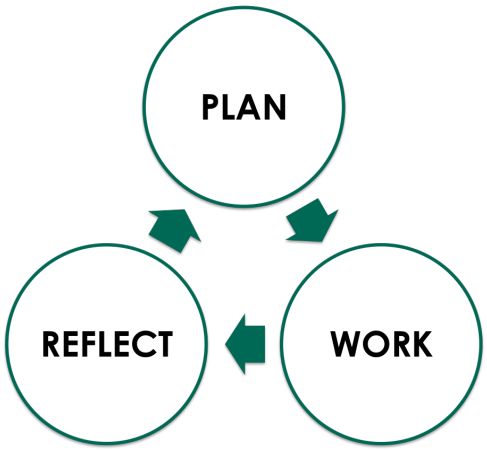
Self-Regulating Your Learning: An Introduction

One goal of this course is to increase your ability to learn efficiently and effectively. This means learning faster and learning smarter—what scientists call being a "self-regulated learner." The following section explains what this means.

Self-regulating your learning means you *plan* your work, monitor your *work* and progress, and then *reflect* on your planning and strategies and what you could do to be more effective. These are the three phases of **Self-Regulated Learning (SRL)**. They are introduced below, and will be followed up on later in the course.

> **Plan:** Before doing a problem or assignment, self-regulated learners *plan*. They think about what they already know or do not know, decide what strategies to use to finish the problem, and plan how much time it will take. Research has shown that math experts often spend much more time planning how they will do a problem than they do actually completing it. Novices, the people who are just starting out, often do the opposite.
>
> **Work:** Self-regulated learners use effective strategies as they *work* to solve problems. They actively *monitor* what study strategies are working and make changes when they are not working. When they do not know which strategy would be better, they ask for help. Self-regulated learners also keep themselves focused while they are working and pay attention to their feelings to avoid getting frustrated.
>
> **Reflect:** Usually after an assignment or problem is done, self-regulated learners take time to reflect about what worked well and what did not. Based on that reflection, they think about how to change their approach in their future. The *reflect* phase helps self-regulated learners understand more about how they learn so they can become more efficient and more effective the next time. Reflecting is important for doing a better job the next time one plans for a new problem or assignment.

You can think of these three phases as a cycle. You incorporate what you learned during the reflect stage into your next *plan* phase, making you a more effective learner as you repeat this process many times. The most effective students get in the habit of working this way:



For most people, self-regulating takes time, practice, and hard work, but it is always possible. People can improve even if, in the beginning, they did not self-regulate their learning very well. The more you practice something and the more you train your brain to think in certain ways, the easier it becomes.

Since thinking this way takes practice, you will have opportunities to practice some of these skills as you progress through this course. As you read through the course material, you will encounter activities that are designed to help you incorporate the *Plan*, *Work*, and *Reflect* phases in specific ways. Take the time to thoughtfully complete these exercises. The payoff will be worth it!

1.2: The Statistical Analysis Process

LEARNING GOALS

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.
- Chance variability can affect the statistical decision-making process.
- Unusual results can provide evidence for a claim about a population proportion.

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.
- Interpret data displayed in a dotplot.
- Evaluate the strength of evidence against a claim about a population proportion.

INTRODUCTION

In this collaboration you will be introduced to the statistical analysis process and learn how this process enables you to use data to make decisions.

How significant is your birthday? Many people believe that habits, hobbies, likes, dislikes and many other things are affected by the time of year you were born. One such theory involves *chronotypes*. The term "chronotype" refers to a person's natural tendency to be most active and alert at certain times of the day. People who are most active early in the day are often labeled as morning people, early birds or larks. Those who are at their best in the evening are called night people or owls. In this investigation you will answer the question: Can someone's birth month influence their chronotype?

CHRONOTYPE THEORY

Chronotype Theory states that there are three chronotypes and four birth months that correspond to each chronotype. The theory states that being a (a) morning person, (b) night person or (c) neither/both is related to the month when you were born, and predicts your chronotype based on your birth month. The table below shows the predictions for each birth month.

Month	**Chronotype**
November, December, January, February	Evening
March, April, September, October	Neither/Both
May, June, July, August	Morning

1A In the preparation for this collaboration, you determined your chronotype based on your personal characteristics. Does the birth-month theory correctly predict your chronotype? Explain.

1B In Preparation 1.2, you were introduced to Chronotype Theory and each student determined whether or not the chronotype they selected for themselves is the same as the chronotype predicted for them by Chronotype Theory.

What fraction of students in your group selected the chronotype that matched the chronotype predicted for them by Chronotype Theory? Express the fraction as a decimal and percentage.

1C Can you use the information collected from your group to decide if Chronotype Theory is reasonable or not reasonable for all college students?

Statistical analysis is the process of looking at data to learn about something bigger. Looking at data allows us to make **generalizations** about **populations** that are 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.

A STATISTICAL INVESTIGATION

Consider the following situation. A researcher obtains a random sample of 50 college students that is representative of all college students. Each student in the sample identified their chronotype and determined whether Chronotype Theory correctly predicted the chronotype they selected. Of the 50 students, 26 indicate that they selected the chronotype predicted for them by Chronotype Theory.

With this sample data, we can perform a statistical analysis to investigate if Chronotype Theory is reasonable.

2 Answer each of the following questions about the four steps of the statistical analysis process.

Step 1: Ask a question that can be answered by collecting data.

A What question are we trying to answer in our investigation?

Step 2: Decide what to measure and then collect data.

B What information can we get from each student to answer this question? How are the data related to the question we are trying to answer?

Step 3: Summarize and analyze the data.

C Each student identified their chronotype and compared it to the chronotype predicted from Chronotype Theory for his or her birth month. Some students found that the predicted chronotype matched the one they selected and others did not. How could we *summarize* the results for all students in the class?

D Of the 50 students in the sample, 26 indicate that they selected the chronotype predicted for them by Chronotype Theory. What percentage of students in the sample selected the chronotype that was predicted for them by Chronotype Theory?
(i) 126%
(ii) 50%
(iii) 52%
(iv) 13%

Step 4: Draw a conclusion and communicate the results.

E Given the summary of the data, how can we use the summary to answer the research question? What would we expect to learn from our data? Can we generalize, or **infer**, the results to something bigger than our class?

Language Tip
To *infer* means to use statistical evidence to make a conclusion that applies to a larger group than the sample.

NEXT STEPS

We have outlined how we might approach each of the steps in a statistical analysis in order to determine whether Chronotype Theory is reasonable. We have completed steps 1 and 2 and part of step 3. We have asked a question to be answered with data, collected data, and summarized the data. We now must think carefully about how we can analyze the data to draw a reasonable conclusion.

To analyze our data, we need to think about how the data can be used to answer our question. We will use *probability* to help in this. Talk about the following questions with your group to begin this process.

For each question below, take a minute to think about this on your own first before sharing in your group. Explain your answers.

3 Suppose Chronotype Theory is false and a person's birth month does **not** predict their chronotype. Is it possible that a student in the class would still select the chronotype that Chronotype Theory predicted for them? Explain why you think this.

4 Suppose Chronotype Theory is false and a person's birth month does not predict their chronotype. Out of all the students in the sample, about what percentage of the students should select chronotypes that matched the predictions?

(i) 100%

(ii) 50%

(iii) 33%

(iv) 3%

5 In the random sample of 50 students, 52% selected chronotypes that matched the predictions from Chronotype Theory. Does this result provide evidence in support of, or against, Chronotype Theory? Explain.

NEXT STEPS

Using Probability to Discover What May Happen by Chance

Language Tip
Chance variation **describes the type of differences we would naturally expect to see between different samples.**

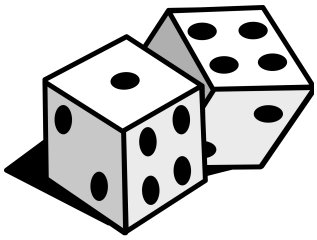
We must use additional tools to assess whether the class result provides evidence in support of, or against, Chronotype Theory. If Chronotype Theory is false and birth month does not predict chronotype, there is a ⅓ chance that a student would select the chronotype predicted for them solely due to chance. So, based on chance alone, we would expect about ⅓ of the students in the class to select the chronotype that is

predicted for their birth month. But how far above 1/3 would the fraction need to be in order to convince us that the opposite is true, that is, that Chronotype Theory is reasonable?

In the next activity, we will answer this question by learning about the values which are most probable. Probable values are those that are most likely to occur by **chance variation**.

If Chronotype Theory is false, we would expect ⅓ of the students in the class to select their predicted chronotype by chance. We will use a 6-sided die to investigate how results vary due to chance when an event has a ⅓ chance of occurring.

In a 6-sided die, each of the six outcomes {1,2,3,4,5,6} are equally likely to occur. The event of rolling a 1 or 2 should occur 33% or ⅓ of the time since this event contains two of the six outcomes that are possible.



6 We will use the following online dice rolling simulator at https://carnegiemathpathways.org/go/randomdice to roll a die 50 times and keep track of the number of rolls that produce an outcome of "1" or "2". In the simulation, select "Roll 50 virtual dice." The number of rolls is based on the number of students in the random sample. Since there were 50 students in the sample, roll the die 50 times. The outcomes of the 50 die rolls will be visible all at once. Write the fraction and proportion of die rolls that result in a "1" or "2".

Note: In statistics, a **proportion** is a number between 0 and 1. It represents a portion out of the total. We usually give proportions as decimals or percents. We can calculate a decimal proportion by dividing the **numerator** of a fraction by the **denominator**. For example, if you roll the die 50 times, and the fraction of die rolls that result in a "1" or "2" is (12/50), then you divide 12 by 50. The proportion would be 0.24. To change to a percent we multiply by 100 or move the decimal 2 places to the right. 0.24 = 24%.

7 How does the proportion of die rolls that resulted in a "1" or "2" compare to 0.33?

8 If you repeat the experiment a second time, do you think you will get exactly the same proportion of die rolls resulting in a “1” or “2”? Select the most appropriate answer below.

(i) Yes, since the outcomes of the rolls will be similar each time we roll the die 50 times.

(ii) Yes, since the outcomes of the 50 rolls are based on a random process.

(iii) No, since the outcomes will vary since this is a random process.

(iv) No, since the die has changed over time.

9 Repeat the experiment 5 times, and for each rolling of 50 die, keep track of the proportion of die rolls resulting in a “1” or “2”. Enter in all of the observed proportions from the experiment.

Roll 1
Number of die rolls resulting in "1" or “2”:
Proportion (out of 50):

Roll 2
Number of die rolls resulting in "1" or “2”:
Proportion (out of 50):

Roll 3
Number of die rolls resulting in "1" or “2”:
Proportion (out of 50):

Roll 4
Number of die rolls resulting in "1" or “2”:
Proportion (out of 50):

Roll 5
Number of die rolls resulting in "1" or “2”:
Proportion (out of 50):

10 Your instructor has constructed a **dotplot** which shows the observed proportions of die rolls that resulted in a “1” or “2” for all students in the class. **Note**: If completing this problem online, follow the instructions given online to create your own dotplot.

11 Use the dotplot to answer these questions.

A What was the smallest proportion observed?

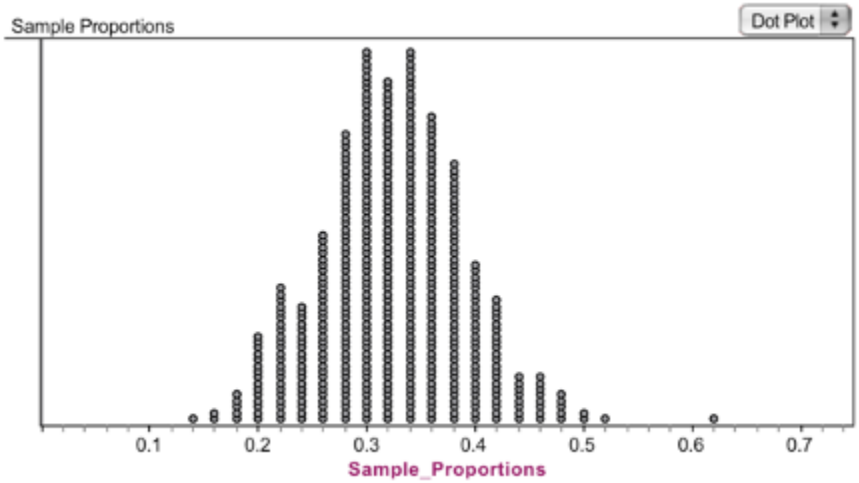
B What was the largest proportion observed?

C Did the proportions differ much from trial to trial?

D Where is the dotplot centered?

12 Summarize what the dotplot tells us. What can you learn from this graph?

A **distribution** of data (like the proportion of die rolls resulting in "1" or "2") shows the values that emerge from an experiment and how often each of the values occur. The dotplot which you created is an example of a distribution. To better understand what outcomes are most likely to occur due to chance alone, we can use a computer simulation with a larger number of trials. Below are the results of 500 simulations, where a die was rolled 50 times. Each dot represents the proportion of rolls that resulted in a 1 or 2 out of 50 in a single simulation.



If Chronotype Theory is false, we would expect that ⅓ of the students in the random sample would select the chronotype predicted for them solely by chance. Since a die roll also has a ⅓ chance of resulting in a "1" or "2", the distribution of proportions of die rolls displays the results we could expect to see from the class if Chronotype Theory is false.

13 In the random sample of 50 students, the proportion of students who selected a chronotype for themselves that matched the chronotype predicted for them by Chronotype Theory was 0.52. How does this proportion compare to the values in the dotplot containing 500 sample proportions?

Draw a Conclusion and Communicate the Results

Decision time! It is now time to make an inference. An inference is a decision based on the evidence we have gathered.

14 Does the sample proportion of 0.52 provide convincing evidence that Chronotype Theory is reasonable? Why or why not?

NEXT STEPS

The task you have just completed illustrates the statistical analysis process, which we have described in four steps.

Steps in a Statistical Analysis

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.

A statistical investigation is an ongoing process. Researchers analyze the results of one study and use the results to generate new research questions. Then they conduct more research. This enables them to think more carefully about how they collect data and identify ways to improve the data collection process.

The table below summarizes what we just completed for the Chronotype investigation.

Steps in Statistical Analysis	For the Chronotype Investigation
1. Ask a question that can be answered by collecting data.	Is Chronotype Theory reasonable? Does a person's birth month predict their chronotype?
2. Decide what to measure and then collect data.	For each student in the class, we collected their birth month, the chronotype they selected for themselves, and the chronotype predicted for them by Chronotype Theory.
3. Summarize and analyze the data.	We summarized the data by determining the proportion of students in the class whose selected chronotype matched their predicted chronotype. We compared the class proportion to the distribution of results that would occur solely due to chance variation.
4. Draw a conclusion and communicate the results.	If the class proportion was similar to results that would occur due to chance, we conclude that Chronotype Theory is false. If the class proportion is much greater than the results that would occur solely due to chance variation, we conclude that Chronotype Theory could be true.

TRY THESE

Consider how you would identify the steps in a statistical analysis in the following study:

A group of researchers studied women who had visited a fertility clinic. The researchers wondered if fewer than half of the women who visit the clinic would want to choose the sex of their future child.[4]

The researchers mailed a survey to women who had visited the clinic. The survey asked women if they would choose the sex of their future child, if they were able to do so. A total of 561 women responded to the survey. Of these 561 women, 229 said that they wanted to choose the sex of their future child.

Language Tip

Convincing evidence **is information that provides very strong support for a conclusion. The evidence is beyond what would be expected due to chance variation.**

The researchers did statistical analysis on the data. Based on their statistical analysis of these data, the researchers concluded that there is **convincing evidence** that *fewer than half* of the women who visit the clinic would choose the sex of a future child. This conclusion is based on the following observation:

If, in reality, at least half of women who visit a fertility clinic would like to choose the sex of a future child, it would be very unusual to observe a percentage as low as 41% in a sample of 561 women who visited the clinic (229/561 ≈ 0.41).

15 Identify the four steps of the statistical analysis process based on the study description above.

Steps in Statistical Analysis	**Study 1**
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]Tarun Jain et al., "Preimplantation Sex Selection Demand and Preferences in an Infertility Population," *Fertility and Sterility* 83, no. 3 (2005): 649-58.

4. Draw a conclusion and communicate the results.	

LET'S SUMMARIZE

Please consider the following key points:

- The statistical analysis process involves the use of data.
- The process involves using evidence from a sample to make an inference about a larger population.
- The process entails asking a question that involves the use of data, collecting data, summarizing and analyzing data, and using data to draw a conclusion.
- To use sample evidence to conclude that a theory about a population is reasonable, we must rule out chance as a plausible explanation for the sample result.

Exercise 1.2

The statistical analysis process enables us to use data to make decisions about situations when we only have a limited amount of information. The statistical investigation we conducted in this collaboration involved using data from students in class to assess whether a person's chronotype can be predicted by their birth month. Using data from the class we made an inference about the relationship between birth month and chronotype for all people. Statistical investigations allow us to use data from small samples to make generalizations about much larger populations.

Here, we will use the statistical analysis process to investigate whether a coin is fair.

A student flips a coin ten times and finds that the coin landed "Heads" on eight of the ten flips. Does this result provide convincing evidence that the coin is **not** fair? In other words, does this result provide convincing evidence that the probability that the coin lands "Heads" is **not equal** to the probability that the coin lands "Tails". Answer the questions below to investigate this research question.

1 If a coin is fair (i.e. the probability that the coin lands "Heads" is equal to the probability that the coin lands "Tails"), what is the probability that the coin lands "Heads" on a single coin flip?

(i) $\frac{1}{5}$

(ii) $\frac{1}{4}$

(iii) $\frac{1}{3}$

(iv) $\frac{1}{2}$

2 If a coin is fair and is flipped 20 times, approximately what fraction of the coin flips do you expect will land heads? Choose the best answer below.

(i) $\frac{0}{20}$

(ii) $\frac{5}{20}$

(iii) $\frac{10}{20}$

(iv) $\frac{20}{20}$

3 If a coin is fair and is flipped 20 times, approximately what proportion of the coin flips do you expect will land *heads*? Choose the best answer below.

(i) 0.40

(ii) 0. 50

(iii) 0.65

(iv) 9.75

4 Using a simulation, you will examine the proportions of *heads* that are likely if a fair coin is flipped 20 times. You will now construct a dotplot of 100 sample proportions based on random trials of 20 coin flips. Follow the instructions below to construct the dotplot in Statkey.

- Open the Statkey simulator by going to https://carnegiemathpathways.org/go/statkeysampling
- Click on Edit Proportion and set the proportion to 0.5 (the probability of getting heads).
- Click on the Sample Size (n =) and set the sample size to 20.
- Click Generate 100 samples to generate and plot 100 sample proportions in the dotplot. The number of sample proportions will be shown at the top right of the viewing window. Each dot in the dotplot represents a sample proportion from a random sample of size 20.

What sample proportions are typical when a fair coin is flipped 20 times?

(i) 0.1 to 0.3

(ii) 0.3 to 0.7

(iii) 0.7 to 0.8

(iv) 0.6 to 0.9

5 The simulation displays proportions that are likely when a fair coin is tossed 20 times. The student conducting the experiment got heads 15 times in 20 flips. The proportion of heads was 0.75.

Is this an unusual value? Statisticians often view a value as unusual if it has a 5% chance (or less) of occurring.

Select the statement below that best describes how the proportion 0.75 relates to the sample proportions in the dotplot.

(i) 0.75 is typical because the value occurs more than once in the dotplot.

(ii) 0.75 is unusual because the proportions in the dotplot are not all equal to 0.75.

(iii) 0.75 is unusual because it is greater than 0.5.

(iv) 0.75 is unusual because it happens no more than 5 times out of 100 trials.

6 Which statement below is a reasonable conclusion based on comparing the sample result of 0.80 to the simulated dotplot?

(i) The outcome of Heads" in 10 flips is very unlikely, if the coin is fair. The observed sample result is unlikely to occur by chance alone, so there is evidence to conclude that the coin is fair.

(ii) The outcome of 8 "Heads" in 10 flips is very unlikely, if the coin is fair. This observed sample result is unlikely to occur by chance alone, so there is evidence to conclude that the coin is not fair.

(iii) The outcome of 8 "Heads" in 10 flips is very unlikely, if the coin is fair. Since this outcome is possible, when a coin is fair, we cannot conclude that the coin is not fair.

(iv) Since the coin was only tossed 10 times, we cannot make a decision about whether or not the coin is fair.

Read the following study description and answer the following questions:

Researchers wanted to know if people think a task will be hard to accomplish when the instructions are difficult to read.[5] To answer this question, researchers randomly divided twenty student volunteers into two groups of 10 students each. Researchers gave instructions to each group of students using different fonts (see below). Instructions for one group were written in a large upright font. The other group was given the *same* instructions but in a font that used *hard-to-read italics*. Researchers asked students to read the directions and say how many minutes they thought the task would take. Researchers did this in order to figure out if the fonts used for the instructions made a difference.

This is the easy-to-read upright font that was used in the study.

This is the hard-to-read italic font that was used in the study.

The first group of students, those that read the instructions printed in the easy font, had an average time estimate of 8.23 minutes. The other group, the group that read the instructions in the *hard-to-read italic* font, had an average time estimate of 15.1 minutes.

Researchers concluded that such a large difference between the averages was not likely to have occurred by chance. There was evidence that people think a task will be harder when instructions are difficult to read.

The four steps of the statistical analysis process are listed below:

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.

[5]Hyunjin Song, "The Effects of Processing Fluency on Judgment and Processing Style: Three Essays on Effort Prediction, Risk Perception, and Distortion Detection" (PhD diss., The University of Michigan, 2009).

7 Which question below is a reasonable research question for this investigation?

(i) Do people like reading in different fonts?

(ii) Do people prefer reading one font to another font?

(iii) Do people think a task will be harder if the instructions for the task are harder to read?

(iv) Do people think that some instructions are easier to follow than other instructions?

8 What variables are used to answer the research question?

(i) Type of font & Amount of time a person thinks a task will take to complete.

(ii) Type of font & Amount of time a task takes to be completed.

(iii) Preferred font & Amount of time a person thinks a task will take to complete.

(iv) Preferred font & Amount of time a task takes to be completed.

9 How are the data summarized?

(i) Researchers compared the total amount of time that the two groups took to complete the tasks.

(ii) Researchers compared the average amount of time that the two groups took to complete the tasks.

(iii) Researchers compared the total amount of time that the two groups estimated that it would take to complete the tasks.

(iv) Researchers compared the average amount of time that the two groups estimated that it would take to complete the tasks.

10 What did the researchers conclude?

(i) The 10 students in the sample who read the instructions in the hard-to-read font took longer to complete the task than the 10 students who read the instructions in the easy-to-read font.

(ii) There is evidence that people will take a longer amount of time to complete a task when the instructions are harder to read.

(iii) The 10 students in the sample who read the instructions in the hard-to-read font think the task will be more difficult than the 10 students who read the instructions in the easy-to-read font.

(iv) There is evidence that people think a task will be harder when the instructions are harder to read.

Read the following study description and answer the following questions:

The United States Government recommends that to stay physically fit, middle-aged adults (ages 40 to 60) need to burn 150 to 400 calories per day doing exercise. Researchers at Minnesota State University, Mankato, wanted to learn whether middle-aged adults who used the Wii Fit video game exercised enough to meet the government's fitness recommendations.[6] The Wii Fit is a video game that includes exercises.

The researchers taught 20 middle-aged adult volunteers how to use the Wii Fit video game. On the day after they were trained, the adults exercised for 20 minutes with the Wii Fit. Researchers measured the total amount of energy each of the adults in the study used in calories. They found that the average energy used was 116 calories for the 20 minute session.

Based on the results of the study, the researchers concluded the Wii Fit video game could be a helpful form of exercise for middle aged adults. But, for exercise with Wii Fit to meet the government's recommendation, the researchers stated that the length of the exercise session should be increased from 20 minutes to 30 minutes.

The four steps of the statistical analysis process are listed below:

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.

11 Which question below is a reasonable research question for this investigation?

(i) Do people think that playing the Wii Fit video game burns calories?

(ii) Does the Wii Fit video game burn enough calories to be considered suitable exercise?

(iii) Does the Wii Fit video game burn more calories than traditional exercise?

(iv) What is the average amount of time that middle-aged adults spend playing Wii Fit video games?

12 What data did the researchers collect to answer the research question?

(i) The amount of time that the adults exercised.

(ii) The name of the adults.

(iii) The type of exercise the adults completed.

(iv) The total amount of calories the adults burned through exercising

[6] http://www.ncbi.nlm.nih.gov/pubmed/21178930

13 How are the data summarized?

(i) Researchers found the proportion of adults who exercise using the Wii Fit video game.

(ii) Researchers found the proportion of adults who prefer exercising with the Wii Fit video game over traditional exercises.

(iii) Researchers found the average amount of calories that adults consumed through exercising using the Wii Fit video game.

(iv) Researchers found the average amount of time that adults spent exercising using the Wii Fit video game.

14 What did the researchers conclude?

(i) The Wii Fit video game is a preferred exercise for some middle-aged adults.

(ii) The Wii Fit video game does not appear to burn enough calories in a 20–minute session, but a 30–minute session would possibly be enough.

(iii) The Wii Fit video game does appear to burn enough calories in a 20–minute session, but a 30–minute session would be even better.

(iv) The sample size is too small to make any reasonable conclusions about all middle-aged adults.

1.2 Monitor (survey)

Monitor your progress on learning the objectives for this unit. If you identify any objectives you need to review, go back through the unit's activities. Your responses are not graded, but will be available for your instructor to see.

Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Skill or Concept: I can …	**Rating from 1 to 5**
Identify the steps in the statistical analysis process given a description of a statistical study.	
Interpret data displayed in a dotplot.	
Evaluate the strength of evidence against a claim about a population proportion.	

1.2-S: Using Percentages to Explore Climate Data

LEARNING GOALS

By the end of this lesson, you should understand that:

- Equivalent numbers can be represented as decimals, percents, and fractions.
- Converting numbers from one form to another can help us understand real-world situations.

By the end of this lesson, you should be able to:

- Convert fractions to decimals and decimals to percentages.
- Find a percent of a number.
- Given a change in a quantity, find and interpret the percent decrease.
- Use fractions, decimals and percentages to solve real-world problems.

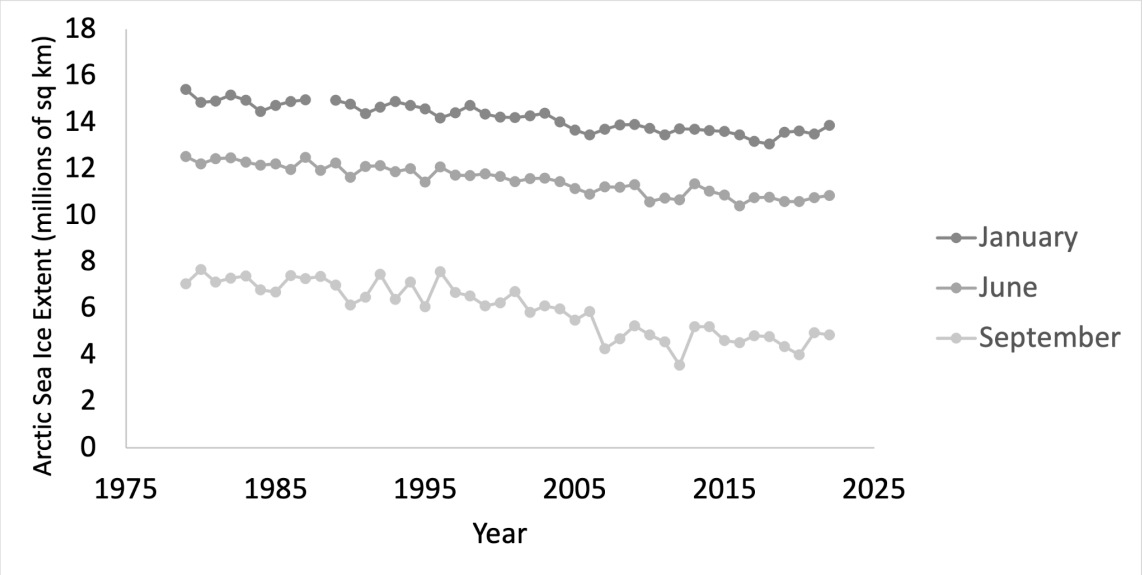
INTRODUCTION

One of the most important issues in the world today is climate change. *Climate change* is the change in the distribution of weather patterns across the globe due to increased concentrations of carbon dioxide and other heat-trapping gases in the atmosphere.

Let's begin with a question related to climate change: how has the amount of ice in the Arctic Ocean changed over the past 40 years?

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). Climatologists use satellite imagery to estimate these values. Sea ice is important as it serves as the habitat for animals in the Arctic region and helps regulate the climate of the Earth. It also stores water that would otherwise contribute to rising sea levels.

The graph below[7] shows the monthly Arctic Sea ice extent during the months of January, June and September, each year, from 1979 to 2022.



1 What does the graph tell us about how the Arctic Sea ice extent has changed during the time period from 1979 to 2022?

2 What could explain why the Arctic Sea ice extent is consistently smaller in September than in January?

The table below displays Arctic Sea ice extent values in specific months from 2010 to 2022. All values are in millions of square kilometers.

Table 1

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Jan	13.74	13.46	13.73	13.70	13.65	13.60	13.47	13.19	13.01	13.57	13.64	13.50	13.88
Jun	10.59	10.75	10.67	11.36	11.03	10.88	10.35	10.72	10.78	10.59	10.59	10.77	10.86
Sep	4.87	4.56	3.57	5.21	5.22	4.62	4.51	4.82	4.79	4.37	4.00	4.95	4.87

[7] http://nsidc.org/arcticseaicenews/

3 Use the data in Table 1 to answer the following questions. What percent of the January 2010 ice extent is the September 2010 ice extent?

A First express your answer as a fraction.

B Now express your answer as a decimal rounded to two decimal places.

C Now express your answer as a percent rounded to the nearest percent.

D Interpret the percentage. Explain the result in your own words.

4 Use the data in Table 1 to answer the following questions. What percent of the January 2013 ice extent is the September 2013 ice extent?

A First express your answer as a fraction.

B Now express your answer as a decimal rounded to two decimal places.

C Now express your answer as a percent rounded to the nearest percent.

D Interpret the percentage. Explain the result in your own words.

5 Use the data in Table 1 to answer the following questions. What percent of the January 2022 ice extent is the September 2022 ice extent?

A First express your answer as a fraction.

B Now express your answer as a decimal rounded to two decimal places.

C Now express your answer as a percent rounded to the nearest percent.

D Interpret the percentage. Explain the result in your own words.

6 What can we infer from the previous answers about the typical percent *decrease* in Arctic Sea ice extent from January to September each year?

7 Suppose the January 2025 Arctic Sea ice extent is 12.98 million square kilometers. Predict the September 2025 Arctic Sea ice extent. Enter the number only, in millions of square kilometers, rounded to two decimal places.

TRY THESE

The table below displays Arctic Sea ice extent values in 1979, 2000, and 2022. All values are in millions of square kilometers.

Table 2

Year	1979	2000	2022
Jan	15.41	14.22	13.88
Jun	12.53	11.67	10.86
Sep	7.05	6.25	4.87

8 Use the data in Table 2 to answer the following questions. What percent of the January 1979 ice extent is the January 2000 ice extent?

A First express your answer as a fraction.

B Now express your answer as a decimal rounded to two decimal places.

C Now express your answer as a percent rounded to the nearest percent.

D Explain in your own words what this percent indicates.

9 Use the data in Table 2 to answer the following questions. What percent of the January 2000 ice extent is the January 2022 ice extent?

A First express your answer as a fraction.

B Now express your answer as a decimal rounded to two decimal places.

C Now express your answer as a percent rounded to the nearest percent.

D Explain in your own words what this percent indicates.

10 A student examined the data in Table 2 and asserted that the ice extent is decreasing more rapidly from 2000 to 2022 than it was from 1979 to 2000. Do the data support this conclusion? Consider all of the data provided above. Explain.

11 If the trends shown continue, predict the Arctic Sea ice extent in September 2044. Explain your reasoning.

NEXT STEPS

Let's now examine a different question: do people in countries throughout the world have differing views about the seriousness of global climate change?

In 2021, adults throughout the world were asked about how concerned they were that climate change would harm them personally at some point in their lifetime. The data in the table below are similar to the survey results from a random sample of adults from five countries.[8]

12 Complete the following table by entering the percent in A-E below. Round to the nearest whole percent.

Country	Very/somewhat concerned	Not too/not at all concerned	Percent who are very/somewhat concerned
Germany	989	329	A
United Kingdom	932	381	B
Australia	1302	732	C
South Korea	887	121	D
Japan	854	300	E

13 Assuming the survey results represent the views of adults in the country, which country in the table has the least concern about global climate change?

14 Which country has the greatest concern about global climate change?

15 If a sample of 100,000 adults in Germany are chosen, how many would you expect to believe that global climate change is a very serious problem?

16 If a sample of 2,000,000 adults in South Korea are chosen, how many would you expect to believe that global climate change is a very serious problem?

8 https://www.pewresearch.org/global/2021/09/14/in-response-to-climate-change-citizens-in-advanced-economies-are-willing-to-alter-how-they-live-and-work/

LET'S SUMMARIZE

- Statistics is the science of collecting, describing, summarizing, and analyzing data to learn about our world.
- Data is information we collect from measurements and observations.
- Equivalent numbers can be represented as decimals, percents, and fractions.
- Converting numbers from one form to another can help us understand real-world situations.

Practice Problems 1.2-S

The data in the lesson showed how the Arctic Sea ice extent has declined over time. The ice extent is the lowest in the summer months as the ice melts due to warming temperatures. In September 1979 the Arctic Sea ice extent was 7.2 million square kilometers. The lowest September measurement ever recorded occurred in September 2012, where the ice extent measured 3.5 million square kilometers.

1 What percent of the September 1979 ice extent is the September 2012 ice extent? Round to the nearest percent.

2 What is the percent decrease in Arctic Sea ice extent from September 1979 to September 2012? Round to the nearest percent.

3 Suppose that the Arctic Sea ice extent from September 2012 (3.5 million square kilometers) decreased by 33%. What would the amount be, in millions of square kilometers? Round your answer to one decimal place.

Global warming is a term used to describe the Earth's long-term rise in global temperature. Gallup poll[9] conducted in March 2021 found that 65% of American adults worried personally about global warming a great deal or a fair amount. The survey result was based on an online survey of U.S. adults.

4 If the survey was based on 1250 adults, how many of the adults surveyed find global warming to be personally important?

[9] https://news.gallup.com/poll/355427/americans-concerned-global-warming.aspx

5 If the survey was based on 3500 adults, how many of the adults surveyed find global warming to be personally important?

6 Suppose that in a different survey of 2205 U.S. adults, 702 said that they felt global warming is not personally important to them. What percentage of the sample find global warming to not be personally important? Round your answer to the nearest percent.

Compare the following fractions, percentages, and decimals to identify which numbers are **larger**.

7 $\frac{1}{5}$ and 28%

8 $\frac{3}{5}$ and 0.45

9 $\frac{35}{80}$ and 33%

10 $\frac{9}{5}$ and 220%

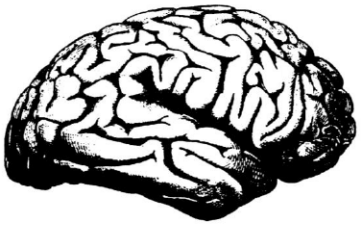
Growth Mindset Activity 1.3

YOU CAN GROW YOUR BRAIN

New Research Shows the Brain Can Be Developed Like a Muscle

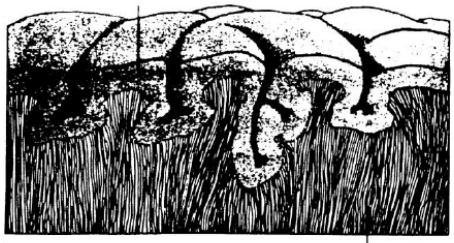
Many people think of the brain as a mystery. We don't often think about what intelligence is or how it works. And when you do think about what intelligence is, you might think that a person is born either smart, average, or dumb—either a "math person" or not—and stays that way for life.

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.

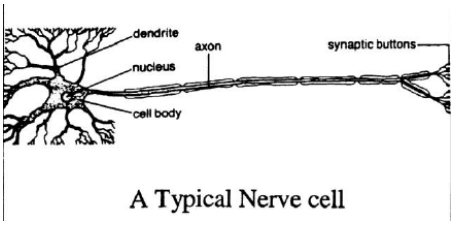


Everyone knows that when you lift weights, your muscles get bigger and you get stronger. A person who can't lift 20 pounds when they start exercising can get strong enough to lift 100 pounds after working out for a long time. That's because muscles become larger and stronger with exercise. And when you 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.

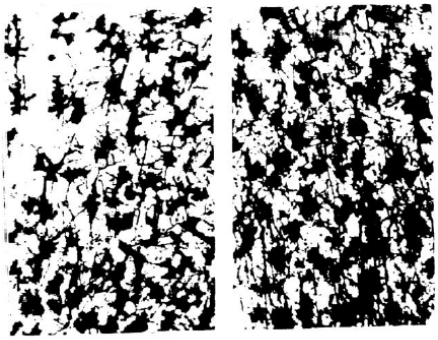


When you learn new things, you actually grow more of these tiny connections and they get stronger. The more you challenge your mind to learn, the more your brain cells grow. Then, things that you once found very hard or even impossible to do—like understanding history, doing algebra, or completing complicated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

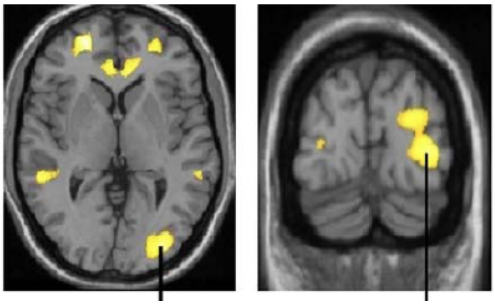
The adult animals who were exercising their brains by playing with toys and each other were also "smarter" –they were better at solving problems and learning new things.

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 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.



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

doi:10.1371/journal.pone.0002669.g001

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 https://carnegiemathpathways.org/go/juggling

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.

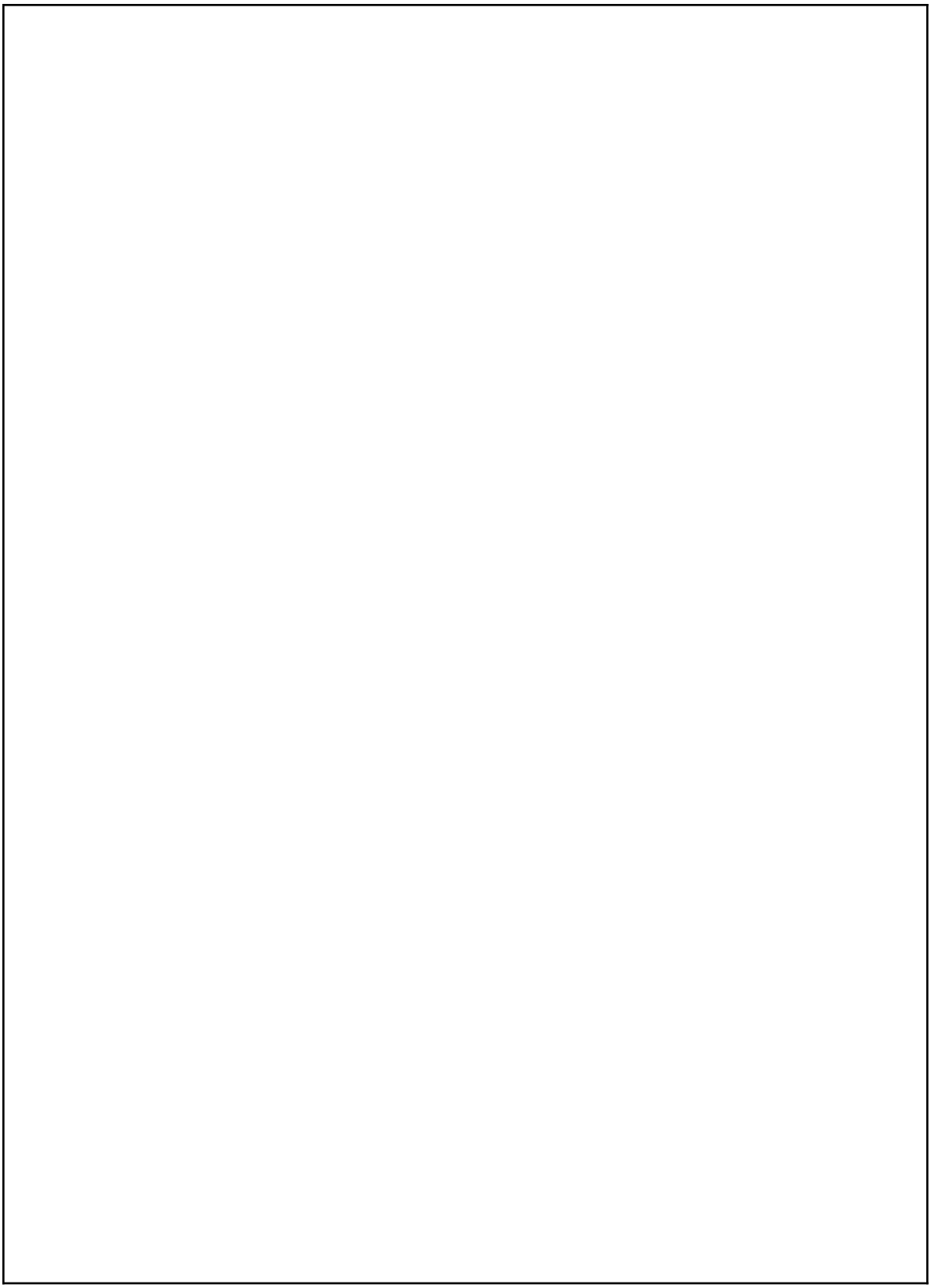
Imagine a friend who is struggling in school. This friend used to do pretty well in school but now is having a hard time and is starting to feel dumb. Write a letter to your friend to encourage him or her—tell them about what you just learned about the brain and why they shouldn't be discouraged. For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

A. How they can get smarter if they work hard and use a good strategy.
B. How they should work hard to build their reading and math muscles.
C. How they are not dumb, they just need to practice using a good strategy.
D. How they can ask the teacher or other students to help them learn better ways to study.
E. 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.)



1.3: Research Questions and Types of Statistical Studies

LEARNING GOALS

By the end of this collaboration, you should understand that:

- The types of conclusions we can draw from data are based on the methods in which data are collected.
- Statistical methods can be applied to a variety of disciplines.

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

- Identify the research question in a given study.
- Distinguish between an observational study and an experiment.
- Identify the explanatory variable and response variable in an experiment.
- Identify variables of interest and population in an observational study.
- Explain why it is or is not reasonable to generalize conclusions from a particular statistical study.

INTRODUCTION

In Unit 1.2, we studied the four step process used in a statistical analysis. Step 1 in this process is "Ask a question that can be answered by collecting data." You will see in this lesson that understanding the type of research question being asked is very important. The type of research question has an impact on the method we use to collect data.

In a **statistical study**, a **population** is a set of all people or objects that share certain characteristics. A **sample** is a subset of the population used in the study. **Subjects** are the individuals or objects in the study. The subjects are often people, but can be animals, plants, or things. **Variables** are the characteristics of subjects that we study. For example, a variable might be eye color, age, educational level, salary or city. In Lesson 1.2, we asked whether a relationship existed between a person's birth month and chronotype. The population was all people. The sample was the students in our class and the variables were sets of personality traits and birth date groups.

The questions we ask in a statistical study are usually in one of two categories:

1. Research questions about a population.
2. Research questions about the **cause-and-effect relationship** between two variables. Note that a cause-and-effect relationship is a relationship between two *variables* where a known change in the first variable produces a predictable change in the second.

Research Questions about a Population

Type of Research Question	Examples
Make an estimate about the population (often about an average or proportion)	• What is the average amount of sleep community college students get at night? • What percent of community college students have jobs?
Test a claim about the population (often a claim about an average or proportion)	• Is the average amount of sleep for community college students more than seven hours? • Do more than half of community college students plan to vote in November?
Compare two populations (often comparing averages or proportions)	• Do freshman university students have higher average GPAs than freshman community college students? • Are community college students more likely to receive financial aid than university students?
Investigate a relationship between two variables	• Is there a relationship between the number of hours a full time student works at a job and his or her GPA? • Are students who drink diet soda more likely to be overweight?

Research Questions about Cause-and-Effect Relationships between Variables

Examples of this type of research question include:

- Does requiring students to do homework in a college class improve test grades?
- Does caffeine reduce the risk of dementia (memory loss associated with old age)?
- Does taking an aspirin daily reduce the risk of heart attacks in adults over the age of 50?

To answer these research questions, we investigate how one variable responds as another variable is manipulated or changed. An **explanatory variable** is the variable being modified or manipulated in the study. A **response variable** is the output variable which is used to measure the impact of changes to the explanatory variable. An **experiment** involves a change to the explanatory variable.

TRY THESE 1

1 Here is one of the studies we examined in Unit 1.2. As you read the problem, think about the following statistical ideas: (1) the population of interest, (2) the variables being studied, and (3) the type of research question that is being asked.

> A group of researchers studied women who had visited a fertility clinic. The researchers wondered if fewer than half of the women who visit the clinic would want to choose the sex of their future child.[10]
>
> The researchers mailed a survey to women who had visited the clinic. The survey asked women if they would choose the sex of their future child, if they were able to do so. A total of 561 women responded to the survey. Of these 561 women, 229 said that they wanted to choose the sex of their future child.
> The researchers did statistical analysis on the data. Based on their statistical analysis of these data, the researchers concluded that there is **convincing evidence** that *fewer than half* of the women who visit the clinic would choose the sex of a future child. This conclusion is based on the following observation:
>
> > If, in reality, at least half of women who visit a fertility clinic would like to choose the sex of a future child, it would be very unusual to observe a percentage as low as 41% in a sample of 561 women who visited the clinic ($229/561 \approx 0.41$).

A What is the research question being asked?

B Does this study ask a question about a population or about a cause-and-effect relationship between two variables?

C What is the population of interest in this study?

D What variable is being examined for each subject in the study?

[10]Tarun Jain et al., "Preimplantation Sex Selection Demand and Preferences in an Infertility Population," *Fertility and Sterility* 83, no. 3 (2005): 649-58.

2 Here is the other study we examined in Unit 1.2. As you read the problem, think about the following statistical ideas: (1) the population of interest, (2) the variables being studied, and (3) the type of research question that is being asked.

Researchers wanted to know if people think a task will be hard to accomplish when the instructions are difficult to read.[11] To answer this question, researchers randomly divided twenty student volunteers into two groups of 10 students each. Researchers gave instructions to each group of students using different fonts (see below). Instructions for one group were written in a large upright font. The other group was given the *same* instructions but in a font that used *hard-to-read italics*. Researchers asked students to read the directions and say how many minutes they thought the task would take. Researchers did this in order to figure out if the fonts used for the instructions made a difference.

This is the easy-to-read upright font that was used in the study.

This is the hard-to-read italic font that was used in the study.

The first group of students, those that read the instructions printed in the easy font, had an average time estimate of 8.23 minutes. The other group, the group that read the instructions in the *hard-to-read italic* font, had an average time estimate of 15.1 minutes.
Researchers concluded that such a large difference between the averages was not likely to have occurred by chance. There was evidence that people think a task will be harder when instructions are difficult to read.

A What is the research question being asked?

B Does this study ask a question about a population or about a cause-and-effect relationship between two variables?

C What is the explanatory variable?

D What is the response variable?

[11]Hyunjin Song, "The Effects of Processing Fluency on Judgment and Processing Style: Three Essays on Effort Prediction, Risk Perception, and Distortion Detection" (PhD diss., The University of Michigan, 2009).

NEXT STEPS

When we know what type of question a study asks and what the variable(s) are in the study, then we can move on to the second step of a statistical analysis. Step 2 in this four-step process is "decide what to measure and then collect data." There are 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.

Language Tip
A sample is *representative* of the population if it doesn't differ in any important way from 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.
The first study, about women choosing the sex of their child, is an observational study. The researchers only asked the women about choosing the sex. They were *observing* the women's responses.

Experiment

An *experiment* is used to answer questions about how one variable responds when another is manipulated. In an experiment, researchers observe how a response variable behaves as an explanatory variable is changed. Researchers actively manipulate the explanatory variable.

Because the goal of an experiment is to learn about the effect of the different treatments (changes to explanatory variable) on the response variable, it is important to apply the different treatments to similar samples. Also, the results only apply to a larger population if the sample is representative of the population.

The second study above is an experiment. The researchers changed the font of the instructions for a task to see if it would change the amount of time the students thought the task would take.

Key Difference between an Observational Study and an Experiment

An important difference between these data collection methods is that there is no attempt to influence the results in an observational study. This is different from an experiment. In an experiment, the explanatory variable is manipulated and changes in the response variable are examined.

TRY THESE 2

3 Imagine that our college is having financial problems. The college announces that it will shorten library hours to save money. The library will be closed nights and weekends. Some students think that it is okay to pay a $20 fee to the college to keep the library open nights and weekends.

We are interested in learning about the proportion of students who would pay a $20 fee to keep the library open nights and weekends. To investigate this question, we select a sample of 100 students. We ask each of the students whether he or she agrees with the $20 fee increase to keep the library open nights and weekends.

A Does this study ask a question about a population or about a cause-and-effect relationship between two variables?

B Is this an observational study or an experiment?

If it is an *observational study*, what is the population of interest? What is the question we are asking about the population?

If it is an *experiment*, what is the explanatory variable and what is the response variable?

C It is important to obtain a sample of 100 students that is *representative* of all students at the college. What is an appropriate way of selecting a sample of 100 students from the population that ensures that it is representative? As you answer this question, be sure to think about (1) the population of interest, (2) your sample of 100 students, and (3) the type of research question that is being asked.

4 We are interested in learning whether jogging for longer amounts of time decreases the resting heart rate of college students. We want to see if there is a difference between:

- The resting heart rate of college students that jog for 30 minutes three times a week for six weeks, and
- The resting heart rate of college students that jog for 15 minutes three times a week for six weeks.

To investigate this question, we will use 100 college students who do not currently jog and who have volunteered to participate as subjects in this study. Resting heart rate of each subject will be measured at the start of the study. Fifty of the students will participate in a jogging program where they get together three times a week and jog for 30 minutes. The other 50 students will get together three times a week, but will only jog for 15 minutes. At the end of six weeks, resting heart rate will be measured again.

A What is the research question?

B Does this study ask a question about a population or about a cause-and-effect relationship between two variables?

C Is this an observational study or an experiment?

If it is an *observational study*, what is the population of interest? What is the question we are asking about the population?

If it is an *experiment*, what is the explanatory variable and what is the response variable?

D We need to divide the 100 volunteers into two groups so that there is a "fair" comparison between the 30 minute and 15 minute jogging groups. What would be an appropriate way to divide the 100 volunteers into two groups so that the two groups have similar volunteers?

NEXT STEPS

Drawing Conclusions from Statistical Studies

The fourth step in the statistical process is **drawing a conclusion and communicating the results**. When making a conclusion, researchers extend beyond the data that are observed to explain what they learned from the study.

There are two types of conclusions that might be made from a study. One type of conclusion is "**generalizing from a sample to the population**." Note that researchers always seek to study a **representative sample** of a larger population. When researchers draw this type of conclusion, they are confident that what they observed in the sample is true for the larger population.

The other type of conclusion is "**cause-and-effect**." This conclusion arises from an experiment when a change in a response variable was caused by the manipulation of an explanatory variable. If a researcher manipulates a variable and this change generates an "effect" or response, the researcher can conclude that the change was due to the variable manipulation.

Language Tip

A conclusion is *reasonable* when the study design allows us to make that type of conclusion.

The table below summarizes when each of these types of conclusions is **reasonable**.

Type of Conclusion	Reasonable When
Generalize from sample to population	Observational study is conducted and the sample is representative of the population
Cause-and-effect	Experiment is conducted and groups assigned to experimental conditions are similar

The best way to choose a sample that is representative of the population is to use a **random sample** from the entire population. The best way to ensure similar groups for different experimental conditions is to use random assignment to the experimental groups. In addition, if an experiment uses both random assignment and a random sample from a population, we can draw a cause-and-effect conclusion and apply it to the larger population.

We will see more about these ideas in upcoming lessons, but without a random sample in an observational study or random assignment in an experiment, no conclusions can reliably be drawn.

TRY THESE 3

5 The SAT exam is used in admissions decisions by many four-year colleges and universities. The College Board carried out a study of 6,498 SAT essays that were selected at random from the more than 1.4 million SAT exams taken in a particular academic year.[12] For this sample of essays, 15% were written in cursive and 85% were printed in block letters. The results showed that the average score for essays written in cursive was higher than the average score for essays that were printed.

A Is this study an observational study or an experiment? Explain.

B Is it reasonable to conclude that writing the essay in cursive was the cause of the higher scores? Explain your answer.

C Is it reasonable to conclude that the average score for essays written in cursive was higher than the average score for essays that were printed for all essays written in this testing period? Explain your answer.

[12] https://archive.org/details/ERIC_ED562868/page/n5/mode/2up

6 Imagine that a psychologist was interested in finding out whether music affects the ability to remember material that has been read. The psychologist recruited volunteer students who said they liked to study while listening to music and randomly assigned them into two groups. Each group was told to read an essay about Pearl Harbor and the U.S. entry into World War II. One group read the essay in silence. The other group read the essay while music of a style of their choice played in the background. After reading the essay they took a brief test that asked the students to recall details about the essay. The psychologist concluded that students who listened to music while they were reading scored lower than students who read in silence.

A Is this study an observational study or an experiment? Explain.

B What are two possible reasons for why the students who listened to music scored lower than the students who didn't listen to music?

C The psychologist found that the difference was so large that it was unlikely due to chance variation alone. Is it reasonable to conclude that the music was the cause of the lower scores? Explain.

7 An article described a study in which researchers looked at a random sample of 500 publicly accessible web profiles posted by 18-year-olds on a social networking site.[13] The content of each profile was analyzed. Researchers drew many conclusions. One of the conclusions was that profiles that list a sport or hobby have fewer references to "risky behavior." Fewer references to "risky behaviors" means that there were fewer comments on the person's profile about sex, drugs, or violence.

A Is it reasonable to generalize the stated conclusion to all 18-year-olds with a publicly accessible web profile on the social networking site? What aspect of the study supports your answer?

B Not all social networking site users have a publicly accessible profile. Is it reasonable to generalize the conclusion "profiles that list a sport or hobby have fewer references to risky behavior" to all 18-year-old users on the site? Why or why not?

LET'S SUMMARIZE

- We can ask two types of research questions. Each type of research question is answered by a different type of study. Observational studies are used to answer research questions about characteristics of populations or comparing two populations. Experiments are used to answer research questions about cause-and-effect relationships.
- For observational studies, we should avoid making cause-and-effect conclusions, but it is possible to generalize from the sample to the population of interest if the study design incorporated random selection from the population.
- For experiments, it is possible to reach cause-and-effect conclusions if the study design uses random assignment to create the experimental groups.
- If an experiment uses both random assignment to create experimental groups and random selection from some population, it is possible to make cause-and-effect conclusions and to generalize these conclusions to the population.

[13]Megan A. Moreno et al., "Display of Health Risk Behaviors on MySpace by Adolescents: Prevalence and Associations," *Archives of Pediatrics & Adolescent Medicine* 163, no. 1 (2009): 27-34. http://www.ncbi.nlm.nih.gov/pubmed/19124700.

Exercise 1.3

"Sweet Potatoes Brighten Your Skin" is the headline of an article that appeared in the magazine *Woman's World*. The article concludes that eating sweet potatoes causes skin to be healthier because it reverses age spots, blocks harmful UV rays in sunlight, and protects against skin dryness. Consider the following four hypothetical study designs. For each study design, answer the questions that follow.

Study Design 1: Two hundred students were selected at random from those enrolled at a large college in California. Each student in the sample was asked whether he or she ate sweet potatoes more than once in a typical week. A skin specialist rated skin health for each student on a scale of 1 to 10. It was concluded that skin health was significantly better on average for the group that reported eating sweet potatoes more than once a week than it was for the group that did not.

1 A Is the study described above an observational study or an experiment?

B Which phrase/sentence from the study design description helped you determine this?

(i) Two hundred students were selected at random.

(ii) Each student in the sample was asked whether he or she ate sweet potatoes.

(iii) A skin specialist rated skin health for each student on a scale of 1 to 10.

(iv) There is no phrase/sentence in the description that helps determine this.

2 A Did the study use random selection for some population?

B Which phrase/sentence from the study design description helped you determine this?

(i) Two hundred students were selected at random.

(ii) Each student in the sample was asked whether he or she ate sweet potatoes.

(iii) A skin specialist rated skin health for each student on a scale of 1 to 10.

(iv) There is no phrase/sentence in the description that helps determine this.

3 A Did the study use random assignment to experimental groups?

B Which phrase/sentence from the study design description helped you determine this?

(i) Two hundred students were selected at random.

(ii) Each student in the sample was asked whether he or she ate sweet potatoes.

(iii) A skin specialist rated skin health for each student on a scale of 1 to 10.

(iv) There is no phrase/sentence in the description that helps determine this.

4 Is the conclusion "eating sweet potatoes leads to healthier skin" reasonable given the study description?

(i) Yes, this used a large sample and random selection.

(ii) Yes, I already knew eating sweet potatoes is good for your skin.

(iii) No, the sample size is insufficient.

(iv) No, observational studies cannot determine cause-and-effect relationships.

5 Is it reasonable to generalize conclusions from this study to some larger population? If so, what population?

(i) Yes, this study involves a large sample and uses random selection. The sample represents the population, which is all college students in the United States.

(ii) Yes, this study involves a large sample and uses random selection. The sample represents the population, which is all students enrolled at the college.

(iii) No, a sample of college students cannot represent any larger population.

(iv) No you cannot generalize conclusions from an observational study.

Study Design 2: One hundred people who live in Miami volunteered to participate in a statistical study. The volunteers were divided into two experimental groups based on sex, with females in group 1 and males in group 2. Those in group 1 were asked to eat six ounces of sweet potatoes daily for three months. Those in group 2 were asked not to eat any sweet potatoes for three months. At the end of the three months, a skin specialist rated skin health on a scale of 1 to 10 for each of the volunteers. It was concluded that skin health was significantly better on average for group 1 than for group 2.

6 A Is the study described above an observational study or an experiment?

B Which phrase/sentence from the study design description helped you determine this?

(i) Those in group 1 were asked to eat six ounces of sweet potatoes daily for three months. Those in group 2 were asked not to eat any sweet potatoes for three months.

(ii) One hundred people who live in Miami volunteered to participate in a statistical study.

(iii) It was concluded that skin health was significantly better on average for group 1 than for group 2.

(iv) There is no phrase/sentence in the description that helps determine this.

7 A Did the study use random selection from some population?

B Which phrase/sentence from the study design description helped you determine this?

(i) One hundred people who live in Miami volunteered to participate in a statistical study.

(ii) The volunteers were divided into two experimental groups based on sex.

(iii) Those in group 1 were asked to eat six ounces of sweet potatoes daily for three months.

(iv) There is no phrase/sentence in the description that helps determine this.

8 A Did the study use random assignment to experimental groups?

B Which phrase/sentence from the study design description helped you determine this?

(i) The volunteers were divided into two experimental groups based on sex, with females in group 1 and males in group 2.

(ii) One hundred people who live in Miami volunteered to participate in a statistical study.

(iii) It was concluded that skin health was significantly better on average for group 1 than for group 2.

(iv) A skin specialist rated skin health on a scale of 1 to 10 for each of the volunteers.

9 Is the conclusion "eating sweet potatoes leads to healthier skin" reasonable given the study description?

(i) Yes, this used a large sample and random selection.

(ii) Yes, this was an experiment and cause-and-effect relationships can be determined.

(iii) No, the experiment had a flawed design and should not be used to determine a cause-and-effect relationship.

(iv) No, observational studies cannot determine cause-and-effect relationships.

10 Is it reasonable to generalize conclusions from this study to some larger population? If so, what population?

(i) Yes, this study involves a large sample and uses random selection. The sample represents the population, which is all people who live in Miami.

(ii) Yes, this study involves a large number of subjects and uses random assignment. The subjects represent the population, which is all people in the U.S.

(iii) No, there were not enough subjects in the experiment.

(iv) No, the experiment had a flawed design and is therefore not conclusive.

Study Design 3: One hundred people volunteered to participate in a statistical study. For each volunteer, a coin was tossed in order to place them into a group. If the coin landed head up, the volunteer was assigned to group 1. If the coin landed tail up, the volunteer was assigned to group 2. Those in group 1 were asked to eat six ounces of sweet potatoes daily for three months. Those in group 2 were asked not to eat any sweet potatoes for three months. At the end of the three months, a skin specialist rated skin health on a scale of 1 to 10 for each of the volunteers. It was concluded that skin health was significantly better on average for those in group 1 than for those in group 2.

11 A Is the study described above an observational study or an experiment?

B Which phrase/sentence in the study design description helped you determine this?

(i) Those in group 1 were asked to eat six ounces of sweet potatoes daily for three months. Those in group 2 were asked not to eat any sweet potatoes for three months.

(ii) A coin was tossed in order to place them into a group.

(iii) If the coin landed tail up, the volunteer was assigned to group 2.

(iv) It was concluded that skin health was significantly better on average for those in group 1 than for those in group 2.

12 A Did the study use random selection from some population?

B Which phrase/sentence in the study design description helped you determine this?

(i) One hundred people volunteered to participate in a statistical study.

(ii) For each volunteer, a coin was tossed in order to place them into a group.

(iii) If the coin landed heads up, the volunteer was assigned to group 1.

(iv) If the coin landed tail up, the volunteer was assigned to group 2.

13 A Did the study use random assignment to experimental groups?

B Which phrase/sentence in the study design description helped you determine this?

(i) One hundred people volunteered to participate in a statistical study.

(ii) For each volunteer, a coin was tossed in order to place them into a group.

(iii) Those in group 2 were asked not to eat any sweet potatoes for three months.

(iv) It was concluded that skin health was significantly better on average for those in group 1 than for those in group 2.

14 Is the conclusion "eating sweet potatoes leads to healthier skin" reasonable given the study description?

(i) Yes. While the subjects volunteered to participate in the experiment, random assignment was used and the experiment was well-designed.

(ii) Yes. Group 1 had a higher average than Group 2.

(iii) No, the sample size is insufficient.

(iv) No, we don't know where people in the sample are from.

15 Is it reasonable to generalize conclusions from this study to some larger population? If so, what population?

(i) Yes, this used a large sample and random selection. It represents the larger population.

(ii) Yes, this used a large sample and random assignment.

(iii) No. This study involved volunteers and we have no additional information about the participants. Therefore, we can't comfortably generalize these results to any larger population.

(iv) No, you cannot generalize conclusions from an observational study.

Study Design 4: One hundred students were selected at random from those enrolled at a large college. Each of the selected students was asked to participate in a study and all agreed. For each student, a coin was tossed in order to place them into one of two groups. If the coin landed head up, the student was assigned to group 1. If the coin landed tail up, the student was assigned to group 2. Students in group 1 were asked to eat six ounces of sweet potatoes daily for three months. Students in group 2 were asked to not eat any sweet potatoes for three months. At the end of the three months, a skin specialist rated the health of each student's skin on a scale of 1 to 10. The researchers concluded that skin health was significantly better for the students in group 1 than for the students in group 2.

16 A Is the study described above an observational study or an experiment?

B Which phrase/sentence in the study design description helped you determine this?

(i) Students in group 1 were asked to eat six ounces of sweet potatoes daily for three months. Students in group 2 were asked to not eat any sweet potatoes for three months.

(ii) Each of the selected students was asked to participate in a study and all agreed.

(iii) One hundred students were selected at random from those enrolled at a large college.

(iv) The researchers concluded that skin health was significantly better for the students in group 1 than for the students in group 2.

17 A Did the study use random selection from some population?

B Which phrase/sentence in the study design description helped you determine this?

(i) Each of the selected students was asked to participate in a study and all agreed.

(ii) One hundred students were selected at random from those enrolled at a large college

(iii) If the coin landed tail up, the student was assigned to group 2.

(iv) At the end of the three months, a skin specialist rated the health of each student's skin on a scale of 1 to 10.

18 A Did the study use random assignment to experimental groups?

B Which phrase/sentence in the study design description helped you determine this?

(i) One hundred students were selected at random from those enrolled at a large college.

(ii) Each of the selected students was asked to participate in a study and all agreed.

(iii) The researchers concluded that skin health was significantly better for the students in group 1 than for the students in group 2.

(iv) For each student, a coin was tossed in order to place them into one of two groups.

19 Is the conclusion "eating sweet potatoes leads to healthier skin" reasonable given the study description?

(i) Yes, this was a well-designed experiment.

(ii) Yes, researchers concluded skin health was significantly higher for Group 1.

(iii) No, these are all college students.

(iv) No, observational studies cannot determine cause-and-effect relationships.

20 Is it reasonable to generalize conclusions from this study to some larger population? If so, what population?

(i) Yes, this study used a large sample, random selection, and random assignment. The sample represents the population, which is all college students in the U.S.

(ii) Yes, this study used a large sample, random selection, and random assignment. The sample represents the population, which is all students enrolled at the college.

(iii) No, a sample of college students from a particular college cannot represent any larger population.

(iv) No. Three months is not a long enough time for the experiment to be conducted.

1.3 Monitor (survey)

Monitor your progress on learning the objectives for this unit. If you identify any objectives you need to review, go back through the unit's activities. Your responses are not graded, but will be available for your instructor to see.

Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Skill or Concept: I can …	**Rating from 1 to 5**
Identify the research question in a given study.	
Distinguish between an observational study and an experiment.	
Identify the explanatory variable and response variable in an experiment.	
Identify variables of interest and population in an observational study.	
Explain why it is or is not reasonable to generalize conclusions from a particular statistical study.	

1.3-S: Evaluating and Simplifying Expressions

LEARNING GOALS

By the end of this lesson, you should understand that:

- Algebraic expressions consist of terms, variables, coefficients, and numbers.
- Simplifying an algebraic expression involves combining like terms.
- Evaluating an algebraic expression involves substituting in a value for one or more variables in the expression.

By the end of this lesson, you should be able to:

- Simplify an algebraic expression by combining like terms.
- Evaluate an algebraic expression.
- Use an algebraic expression to solve a contextual problem.

INTRODUCTION TO 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
$-a^3$	–1
$2b$	2
–5	Does not have a coefficient because there is no variable. This is called a **constant term** because it never changes.

1 State the number of terms in each expression:

A $3x + 4$

B $5x - 4x^2 + 2$

C 5

2 What is the coefficient of the x^2 term in the expression $5x - 4x^2 + 2$?

3 To ***evaluate*** an algebraic expression we substitute a number into the expression for one or more variables in the expression. Evaluate the following expressions for the given values of the variable(s). Show your work and simplify the expression to obtain a single number.

A $6x - 5$ when x = -3

B $2l + 2w$ when $l = 12$ and $w = 9$

C $t^2 - 5t + 4$ when $t = 4$

D $6(n + 5) + 8n$ when $n = -2$

E $-8x^3 - x + 6y + 10$ when $x = -1$ and $y = 3$

NEXT STEPS

We often evaluate algebraic expressions when we use *mathematical models* to explore relationships between real-world phenomena. Mathematical models are often formulas that describe relationships between variables in the real world.

4 A statistical study involving children from the United States and Canada analyzed children's drawings from draw-a-scientist activities over the past five decades. In this activity, children are asked to draw a scientist and the sex of the scientist is recorded. The study found that the percentage of all children who draw a man in the draw-a-scientist activity has tended to decrease over time, and that the percentage of girls who draw a man is decreasing more rapidly than the percentage of boys who draw a man.

The model $G = -0.6t + 67$ describes the percentage of **girls** who draw a man when asked to draw a scientist over time. *G* is the percentage of girls who draw a man when asked to draw a scientist; *t* is the number of years since 1985.

A Use the model to compute the percentage of girls who drew a man, *G*, in (i) 1985 (*t* = 0), (ii) 1995 (*t* = 10), (iii) 2005 (*t* = 20), and (iv) 2015 (*t* = 30).

t	***G***
0	(i)
10	(ii)
20	(iii)
30	(iv)

B Look at how the percentage of girls who draw a man when asked to draw a scientist is changing over time. Complete the following:

Since 1985, every 10 years the percentage of girls who draw a man when asked to draw a

scientist decreases by _____ percentage points.

C Use the model to predict the percentage of girls who will draw a man when asked to draw a scientist in the year 2030.

5 The weekly cost of driving a car to school for a college student is based on the number of miles she drives. The mathematical model below describes the cost, *C*, in dollars, to drive *m* miles to school each week.

$$C = 20 + 0.083m + 0.24m$$

A If the student drives 18 miles each week, what is the weekly cost? Round your answer to the nearest cent.

B If the student drives **8 miles each way, three times a week**, what is the weekly cost? Round your answer to the nearest cent.

6 Environmental clean up can be costly for businesses and governments. The mathematical model below describes the cost, *C*, in millions of dollars, to remove *x* percent of pollutants from a lake.

$$C = \frac{280x}{100-x}$$

A Use the model to find the cost of removing 20% of pollutants from the lake. Write your answer in millions of dollars.

B Use the model to find the cost of removing 50% of pollutants from the lake.

NEXT STEPS

To ***simplify*** an algebraic expression, we combine *like terms*. Like terms are terms that contain the same variable, or constants. We can reorder terms in an expression using the ***associative property of addition***. This property indicates that the order in which we add three or more numbers does not impact the sum. We can combine like terms due to the ***distributive property***. The distributive property states that for any real numbers a, b and c:

$$a(b + c) = ab + ac \qquad \text{or} \qquad a(b - c) = ab - ac$$

7 Simplify the following expressions. First reorder the terms to place the like terms next to each other, then combine like terms.

A $5x - 5 + 9x$

B $12 + 8w - 2 - 13w$

C $8t^2 - t + 4 + 7t + 15$

D $20 + 0.083m + 0.24m$

E $9(y - 3) - 40 + 2y$

F $5(x - 4) + 2.5(x - 12)$

LET'S SUMMARIZE

In this lesson, you learned that:

- Algebraic expressions consist of terms, variables, coefficients, and numbers.
- Simplifying an algebraic expression involves combining like terms.
- Evaluating an algebraic expression involves substituting in a value for one or more variables in the expression.

Practice Problems 1.3-S

1 Answer the following questions about the algebraic expression below:

$$3x + 2y + 25$$

A How many terms are in the expression?

B What is the coefficient of the x term?

C Evaluate the expression when $x = 6$ and $y = -9$.

2 Evaluate the following algebraic expressions for the given values of the variable(s).

A $3v - 25$ when $v = 4$

B $2n^2 - n - 8$ when $n = -2$

C $x^2 - 9x + 1$ when $x = 5$

Using data from the draw-a-scientist study on children's drawings from 1985 to the present, we can model the percentage of **boys** who will draw a man when asked to draw a scientist over time. The equation $B = -0.23t + 92$ models this percentage. B is the percentage of boys who will draw a man when asked to draw a scientist; t is the number of years since 1985.

Use the model to compute the percentage of boys who draw a man, B, in:

3 1985 ($t = 0$).

4 1995 ($t = 10$).

5 2005 (t = 20).

6 2015 (t = 30).

7 Look at how the percentage of boys who draw a man when asked to draw a scientist is changing over time. Complete the following:

Since 1985, every 10 years the percentage of boys who draw a man when asked to draw a scientist

decreases by _____ percentage points.

8 Use the model to predict the percentage of boys who will draw a man when asked to draw a scientist in the year 2030. Write your answer rounded to two decimal places.

Simplify the following expressions.

9 $7x - 10 + x + 4$

10 $5(m - 2) - 8m - 12$

11 $5(w - 2t) + 15 + 7w + 9t$

12 $-5(n + 1) + 6(n - 8$

1.4: Random Sampling & Experiments

LEARNING GOALS

By the end of this collaboration, should understand that:

- Bias can sometimes be introduced by the way the sample is selected.
- The absence of bias is one characteristic of a good sampling plan.
- Random sampling helps protect against selection bias.
- The sample size plays an important role in determining the accuracy of a sample.
- The purpose of random assignment in the design of an experiment is to eliminate the influence of variables that we cannot control or do not know about. The purpose of a control group is for comparison with the treatment groups.
- The purpose of a placebo group is to rule out the placebo effect in an experiment.
- The purpose of blinding is to remove the influence of the participants' prior beliefs and the researchers' prior beliefs.

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

- Describe a method for selecting a simple random sample.
- Given the design of a study or experiment, identify any biases in the design.
- Identify the response variable, explanatory variable and treatments, when given the description of an experiment.
- Describe a process for randomly assigning subjects to treatment groups.
- Explain the benefits of random assignment in the design of an experiment.
- Given the description of an experiment, identify whether random assignment and direct control were part of the experimental design.
- Identify blinding, the control group, and the placebo group in an experiment.
- Explain the importance of blinding, the control group, and the placebo group in an experiment.
- Describe characteristics of a well-designed experiment.

INTRODUCTION

In Unit 1.3, we learned that in order to generalize results from a sample to the population, the sample must be **representative** of the population. A sample is representative if it does not differ in any important way from the population. This collaboration explores sampling techniques that yield representative samples.

1 Suppose our college is thinking of ways to raise money. Many students like parking spaces close to their classes. The administration is thinking of selling reserved parking spaces for $100. The college wants to know the percentage of students who would support this fee.

One way to find the percentage of students that support this fee would be to conduct a **census**. A census is a **survey** of an entire population. The college would ask every student on campus if she or he would support the fee.

Is this a reasonable plan? Why do you think so?

2 Read the following ways to sample students at our college. For each method:

- Tell whether the method would produce a sample that represents the student population.
- Explain why or why not.

A Choose four 8:00 a.m. classes at random. **Survey** all the students in each class.

B Put a **poll** on the front page of the college website. A poll is an opinion survey. Use the students who answer the question as the sample.

C Talk to students as they enter the Student Center.

3 None of the sampling methods above will produce a sample that is representative of the college's population of students. Suggest a better method to gather a representative sample.

NEXT STEPS

When we sample, our goal is for every population member to have the same chance of being selected. We want to avoid **selection bias**. This is the term used to describe a situation in which a sample differs from the population in some systematic way, so that some individuals are more likely to be selected than others. For example, imagine that we are surveying students and we choose only students entering the library. Not all students go to the library. The students in the sample may not be representative of the population of all students.

One good way to avoid *selection bias* is to select a **simple random sample.** In a simple random sample, all samples (of a given size) have the same chance of being chosen. One way to obtain a simple random sample is to put the name of every individual in the population on a piece of paper and put the names into a hat. Then we draw *n* names to include in our simple random sample of size *n*. In practice, a number is assigned to each individual and then random numbers are used to choose the sample.

Language Tip
***Bias* is a systematic (not random) deviation from the true value.**

4 **Biased samples** result when sampling methods tend to leave out certain types of population members. The three previous sampling methods all produced biased samples. There are different types of biased samples.

A One type of biased sample is a **voluntary response sample**. Good samples are chosen by researchers. In a voluntary response sample, the participants are self-selected. In other words, each participant *chooses* to participate.

Which sample below is a voluntary response sample?

(i) Choose four 8:00 a.m. classes at random. Select all the students in these four classes for the sample.

(ii) Put a poll on the front page of the college website. Use the students who answer the question as the sample.

(iii) Talk to students as they enter the Student Center.

B Another biased sample is a **convenience sample**. Convenience sampling does not use random selection. It involves using an easily available or "convenient" group to form a sample. Many samples have convenience sampling problems. Which sample below is the best example of a convenience sample?

(i) Choose four 8:00 a.m. classes at random. Select all the students in these four classes for the sample.

(ii) Put a poll on the front page of the college website. Use the students who answer the question as the sample.

(iii) Talk to students as they enter the Student Center.

5 Suppose our college has 13,000 students. The college has the names and email addresses for all students in its database.

A Suggest a way that the administration could choose a simple random sample of 150 students to survey about the parking fee proposal.

B After the administration has chosen a sample, how could they actually conduct the survey?

C The administrator is told that the college can dedicate additional time and money to conducting this study. Would it make sense for the administrator to double the sample size and survey 300 students? Explain.

NEXT STEPS

Learning from Experiments

Previously, we discussed the four step process in a statistical study. The first step in this process is to ask a question that can be answered by collecting data. In this collaboration we will explore how an experiment could be used to answer the question: "Does listening to music make a difference in how much students remember while studying?"

Experiments often begin with a researcher stating a **hypothesis**. A **hypothesis** is a **conjecture**—an informed guess based on some information you have about the topic you want to study.

6 State a hypothesis about the effect that listening to music has on the amount of material a person remembers when studying. Do you think a person remembers *more*, *less*, or about *the same* as a student who studies in silence? Explain why you think this?

7 Suppose we think that listening to music while studying decreases the amount of information memorized. Design an experiment to test this hypothesis. Explain how you will collect and analyze data to make a decision about this hypothesis. Clearly describe the explanatory variable and the response variable in the experiment.

8 A Suppose a researcher designed the following experiment: In a class of 40 students, 20 were randomly assigned to a control group, and 20 were randomly assigned to a treatment group. Subjects in the control group read a paper for 10 minutes without listening to music, while subjects in the treatment group read the same paper for 10 minutes while listening to music. After reading the paper, all students completed a quiz on the contents of the paper. The quiz scored out of 10 points. The quiz results are shown below.

Quiz Scores: Listening to Music { 5, 6, 4, 5, 4, 6, 5, 4, 3, 4 }
Quiz Scores: Not Listening to Music { 7, 8, 5, 6, 7, 6, 8, 5, 7, 6 }

Find the average score of the students who listened to music.

Find the average score of the students who did not listen to music.

B How do the averages compare?

9 Think about what you know about a well-designed study. What are two possible reasons you might observe a difference between the students who studied while listening to music and the students who studied in silence?

Significance

When a difference between groups of a variable is so large that it is *unlikely* to be caused by chance, we say it is **significant.** In this case a significant difference would mean that we believe that listening to music was the reason for the difference in quiz score averages between the two groups. We do not yet have the tools to determine if the difference between averages is due to music listening or chance variation.

For now, we need to be careful. We shouldn't draw final conclusions yet. We need to rule out other possible causes. Think about other factors, or reasons, that could explain differences in quiz scores. For example, if some students took American History recently, they might get higher scores on the quiz.

Anecdotal Data

We should not draw conclusions based solely on information from a handful of people. For example, suppose one student states that they know from personal experience that they retain more information when they study while listening to music. This evidence may be legitimate, but it is not enough for a conclusion to be made. This is an example of anecdotal evidence. An anecdote, or a personal story, may be persuasive, but it only gives information about one individual.

Another example of anecdotal evidence would be observing a person who smokes a pack of cigarettes every day and does not get cancer. One person does not provide sufficient evidence. It would not be valid to conclude that people who smoke a pack of cigarettes every day do not get cancer.

NEXT STEPS

Direct Control, Random Assignment, Blinding, and the Placebo Effect

The goal of an experiment is to determine the effect of changing a treatment on the response variable. To do this, we want to rule out other possible explanations for differences in responses to treatments. Two common strategies to help with this are **direct control** and **random assignment.**

Direct control means that if you notice that there are other variables, *besides the explanatory variable*, that might affect the response, you try to manage those variables. **Random assignment** helps us create groups that are similar.

TRY THESE

10 Consider the following description of an experiment.[14] 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 evaporation. In this experiment the participants handled uncooked chicken for 45 seconds. The participants then washed their

[14]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.

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 One variable that might affect the response is the length of time that people handled the raw chicken. The researchers controlled this by having everyone handle the chicken for the same amount of time—45 seconds. Two other variables were controlled in this experiment. What are they?

B A **control group** is a group in an experiment that does not get a treatment. Including a control group in an experiment provides a basis for making comparisons. Is there a control group in this experiment? Explain.

Blinding

Sometimes people already have ideas about whether the treatments in an experiment will be effective. These beliefs might influence the response. When participants in an experiment do not know the type of treatment they are receiving, they are said to be **blinded**. Blinding participants is a way to prevent prior beliefs about the treatment from influencing their response.

11 A shoe company wants to compare two different products for making hiking boots waterproof. Researchers create an experiment to compare two different products. In the experiment, researchers randomly give participants hiking boots. The hiking boots were waterproofed using one of the two waterproofing products. Participants then wear the boots on a hike along a mountain stream. They have to cross the stream and walk in the water in several places. After the hike, the boots are left to dry. Then the boots are sent to a lab where a technician evaluates the boots for water damage.

Do you recommend blinding only the participants, only the lab technician, or both the participants and lab technician? Explain your answer.

Placebo Effect

In experiments that use human participants, use of a control group may not be enough to establish whether a treatment really has an effect. Studies have shown that people sometimes respond in a positive way to treatments that have no active ingredients. These non-treatments that have no active ingredients may be colored water or sugar pills, and are called **placebos**. People often report that such non-treatments relieve pain or reduce other symptoms such as dizziness. Thinking that a non-treatment has helped the pain or made you feel better is an example of what is called the **placebo effect**.

Because of the placebo effect, experiments often include a control group that receives no treatment and another control group that receives a placebo. The placebo is identical in appearance (and taste, etc.) to what the people in other experimental groups receive. In such situations, participants should be blinded—we do not want people to know that they are receiving a placebo.

12 A researcher included a control group (with no treatment) and a placebo group in her experiment. How does including both groups allow her to decide when a placebo effect occurs?

13 How does including a placebo group in her experiment allow her to decide whether a particular treatment has a real effect on the response variable?

LET'S SUMMARIZE

Please consider the following key points:

- The goal of an experiment is to determine how changes to an explanatory variable affect a response variable.
- Direct control and random assignment are two strategies that allow us to rule out other possible explanations for changes to the response variable.
- Random assignment ensures that the treatment groups have similar characteristics. This helps to control for variables that we may not know about in advance and enables us to draw cause-and-effect conclusions about the effect of the explanatory variable on the response variable.
- Generally, we only draw cause-and-effect conclusions from experiments and not from observational studies.

Exercise 1.4

Collaboration 1.4 examined important characteristics of observational studies and experiments. The questions in Exercise 1.4 will assess your understanding of these concepts.

1 Imagine that you want to learn about the average number of hours, per day, that students at your college spend online. You want to select a simple random sample of 75 students from the full-time students at your college. You have a list of all full-time students, whose names are arranged in alphabetical order. Suppose that there are 5000 students at your college, of which 3500 are full-time.

Which method below obtains a **simple random sample** of 75 students from this population?

(i) Select the first 75 students who you meet on campus for the sample.

(ii) Number all the students from 1 to 5000, use a random number table or calculator to randomly select 75 numbers between 1 and 5000, determine which students correspond to those numbers on the list, and select them for the sample.

(iii) Number all the full-time students from 1 to 3500, use a random number table or calculator to randomly select 75 numbers from1 to 3500, determine which students correspond to those numbers on the list, and select them for the sample.

(iv) Send an email to all 3500 full-time students. Use the first 75 students who respond to your email as the sample.

2 Imagine you want to estimate the average amount of time per week that students at a particular college spend studying. Which method below is an example of a **convenience sample**?

(i) Select 50 students who are enrolled in English literature at the college this semester.

(ii) Select 40 students as they enter the library.

(iii) Select 50 students who participate in student clubs and student activities.

(iv) Select 40 students at random from all students at the college by randomly choosing 40 student identification numbers.

3 Imagine you want to determine the locations on campus where students most prefer to go to study. Which method below is an example of a **voluntary response sample**?

(i) Email all students a link to an online survey. The sample would include all students who respond.

(ii) Survey all first-year students at a mandatory first-year orientation.

(iii) Select 60 students who are studying in the library during one week of the semester.

(iv) Email a random sample of 200 students a link to an online survey. Require their response before they are allowed to enroll in classes for the next term.

4 A researcher at King's College in London found that infomania (information overload) has a temporary, negative effect on intelligence quotient (IQ). 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. What is the explanatory variable in this experiment?

(i) IQ test score

(ii) Number of email distractions

(iii) Whether or not the person was distracted with texts and emails

5 What is the response variable in this experiment?

(i) IQ test score

(ii) Number of email distractions

(iii) Whether or not the person was distracted with texts and emails

6 Why would it be good for the researchers to use random assignment to put each volunteer in one of the experimental groups?

(i) Random assignment places each volunteer in the treatment group that they prefer to be in. This helps to ensure that the volunteers will follow directions and participate in the entire experiment.

(ii) Random assignment helps to ensure that the treatment groups have similar characteristics.

(iii) Random assignment helps to ensure that the treatment groups will have different outcomes.

(iv) Random assignment ensures that the volunteers are unaware of which treatment group they are in so their prior beliefs and experiences don't influence the outcomes.

7 Which statement below accurately describes the control group?

(i) The control group was the group of volunteers who were distracted with text and email messaging while taking the IQ test.

(ii) The control group was the group of volunteers who took the IQ test without having any distractions.

(iii) The experiment did not involve a control group.

8 Is it possible for the subjects of this study to be blinded?

(i) Yes, since the volunteers will not know which volunteers are in their group or the other group.

(ii) Yes, since the volunteers will not know whether they are being distracted during the test.

(iii) No, since the volunteers will know that they are taking an IQ test and they will know whether or not they were checking emails and messages.

(iv) No, since the volunteers will know the conditions (distracted, not distracted) under which they took the test.

9 Researchers recently designed an experiment to investigate whether dogs can be trained to recognize cancer by smell. In the experiment, dogs were trained to distinguish between people with and without cancer by sniffing exhaled breath. Dogs were trained to lie down if they detected cancer. After training, the dogs' ability to detect cancer was tested using breath samples from an unfamiliar group of people. According to the article that presented the study's results, "The researchers blinded both [the] dog handlers and [the] experimental observers to the identity of [the] breath samples."[15]

Which statement below explains the meaning of the last sentence of the study?

(i) The dog handlers and experimental observers were unaware of each other.

(ii) The dog handlers and experimental observers were unaware of which dogs received training.

(iii) The dog handlers and experimental observers were unaware of which breath samples came from people who had cancer.

[15] http://www.azcentral.com/health/news/articles/0203doctordogs.html

10 Why is blinding important for this experiment?

(i) If the dog handler or the observers knew which breath samples were from people who had cancer they could inadvertently pass some signals on to the dogs and influence the dogs' behavior.

(ii) If the dog handler or the observers knew which breath samples were from people who had cancer they could misinterpret the behavior of the dogs.

(iii) If the dog handler or the observers knew which breath samples were from people who had cancer they could tamper with the breath samples.

11 Which feature of the experiment is an example of direct control?

(i) The breath samples were administered to dogs randomly.

(ii) Dog handlers were randomly assigned to dogs to help dogs participate in the experiment.

(iii) Dog handlers were instructed to allow dogs to have the same amount of time to sniff the breath samples.

(iv) Observers were randomly assigned to observe the dogs' behaviors in the experiment.

1.4 Monitor (survey)

Monitor your progress on learning the objectives for this unit. If you identify any objectives you need to review, go back through the unit's activities. Your responses are not graded, but will be available for your instructor to see.

Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Skill or Concept: I can …	**Rating from 1 to 5**
Identify the response variable, explanatory variable and treatments, when given the description of an experiment.	
Explain the benefits of random assignment in the design of an experiment.	
Explain the importance of blinding, the control group, and the placebo group in an experiment.	
Describe characteristics of a well-designed experiment.	

1.5: Sampling Strategies, Bias, and Random Assignment in Experiments

LEARNING GOALS

By the end of this collaboration, you should understand that

- There are reasonable alternatives to simple random sampling.
- Convenience samples are poor sampling methods.
- Stratified random sampling can be preferable to simple random sampling if there is an effective way to stratify.
- Bias can also be introduced in ways other than how the sample is selected. These include response bias and nonresponse bias.
- The purpose of random assignment in the design of an experiment is to create similar groups for each experimental condition.

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

- Distinguish between simple random sampling, stratified random sampling, and systematic sampling.
- Explain why it is desirable to create similar groups when considering stratified random sampling.
- Critically evaluate a sampling method.
- Explain how the wording of survey questions has the potential to introduce bias.
- Give an example of a situation where a sensitive question or the desire to provide a socially acceptable response may introduce bias.
- Describe a process for randomly assigning experimental units to experimental conditions.
- Explain the benefits of random assignment in the design of an experiment.

INTRODUCTION

In the previous collaboration you learned that simple random sampling is a way to select a sample that avoids bias and is representative of the population. However, it is not always easy to select a simple random sample. Sometimes it may even be impossible. For example, a business may want to take a sample of its customers. If they do not have a list of their customers they can't do a simple random sample. In situations where simple random sampling is difficult or impossible, you might use other sampling strategies. In this lesson, you will learn three new sampling strategies: **systematic sampling**, **stratified random sampling,** and **convenience sampling**.

Systematic Sampling

In systematic sampling, individuals in the population are selected in a "systematic" way by choosing a number *k*, such as 100, and then choosing every *k*th (or in this case 100th) individual from the population (which should be ordered in some sequence). In general, systematic sampling works well when there are

no repeating patterns in the population list or ordered arrangement of the population that might be related to the variable you are examining.

Stratified Random Sampling

In stratified random sampling the population is divided into non-overlapping groups that have some common characteristic, like sex, ethnicity, or college major. Then subjects are randomly selected from *each group*. The subjects are then pooled together to produce one sample. Stratified random sampling is often used when the non-overlapping groups are less variable than the population as a whole. When this occurs the stratified random sample can produce more precise estimates of population characteristics than a simple random sample of the same size. This method is also used when we want to make sure that every group is included in the sample. This is important if some of the groups are large and others are small.

Convenience Sampling

One strategy that is usually a bad idea is to use a convenience sample. This method does not use random selection and involves using a readily available or "convenient" group to form a sample. Data from a convenience sample should *not* be used to generalize to *any* larger population.

1 Suppose you are interested in learning how far the people at a college football game traveled to see the game. There is a sellout crowd of 20,000 people. Assume that the seats in the stadium are numbered from 1 to 20,000.

A Describe a method for selecting a **systematic sample** of 200 people from the population of 20,000 people attending the game.

B Describe a method for selecting a **stratified random sample** of 200 people from the population of 20,000 people attending the game.

C Describe a method for selecting a **convenience sample** of 200 people from the population of 20,000 people attending the game.

D Which sampling plan above is easiest to implement?

E Which sampling plan is the most difficult to implement?

F Which plan will most likely yield a representative sample?

NEXT STEPS

In the previous lesson we discussed *selection bias*. Selection bias occurs when a sample differs from the population in some systematic way, so that some individuals from the population are more likely to be selected for the sample than others. Bias can also be introduced by the way in which data are obtained from the sample. Two ways in which this can occur are **response bias** and **nonresponse bias.**

Response Bias

Response bias occurs when the wording of the question might distort or influence the response.

2 Read the following question that was part of a Gallup survey sponsored by the American Paper Institute:

> It is estimated that disposable diapers account for less than two percent of the trash in today's landfills. In contrast, beverage containers, third-class mail, and yard waste are estimated to account for 21% of trash in landfills. Given this, in your opinion, would it be fair to tax disposable diapers?

A Do you think that the wording of this question may have influenced the way people responded to the question? If so, do you think that this was intentional? Explain your answer.

B Work with your group to write a better survey question to replace the question above.

Nonresponse Bias

Nonresponse bias occurs when it is not possible to collect data from every individual selected to be a part of the sample. For example, let's say that you carefully select a random sample of students at your college. Then as you attempt to collect data, you find that some students refuse to participate. Maybe it also turns out to be impossible to contact some of the selected students. Bias can occur if the people who respond differ in some important way from those who do not respond. Data obtained *only* from responders may not be representative of the population.

The nonresponse rate (the proportion of the originally selected sample that does not respond) for surveys and opinion polls can differ from study to study depending on the following points:

- survey topic
- the length of the survey
- the way in which the survey is conducted

Discuss the following questions in your groups.

3 Do you think the nonresponse rate would be higher for a 20-question survey or a 50-question survey? Why do you think so?

4 Do you think the nonresponse rate for a survey of college students would be higher for a survey about student loans or for a survey about saving for retirement? Why do you think so?

NEXT STEPS

In previous collaborations, we stated that random assignment helps to make experimental groups similar. In this exercise we will see how well random assignment actually works.

Random Assignment—Creating Comparable Groups

An article in the journal *Pediatrics* reported on the results of an experiment that compared recovery times for two types of hernia surgery for children.[16]

Method 1: laparoscopic repair (a surgery that uses three small incisions)

Method 2: open repair (a surgery that uses one large incision)

To compare the two treatments (hernia surgery methods), the researchers needed to create two groups of children that were similar with respect to any variables that might affect the response variable (recovery time).

Imagine that a new group of researchers thought that another variable - a child's age - might also affect his or her recovery time. The researchers wanted to *control* for age, so they wanted the two treatment groups to contain children who were similar in age. This would prevent the age variable from influencing the response variable in the experiment.

One way to do this is to randomly assign children to one of the two groups. This might be done by flipping a coin to assign each child to a group. If the coin lands *heads,* the researchers assign the child to the Method 1 group. If the coin lands *tails*, the researchers assign the child to the Method 2 group.

Let's investigate whether this method of random assignment creates similar groups. Suppose there are 30 children with hernias who volunteered to participate in the experiment. The identification numbers and ages of these 30 children are given in the following table.

[16]A. I. Koivusalo et al., "A Single-Blinded, Randomized Comparison of Laparoscopic Versus Open Hernia Repair in Children," *Pediatrics* (2009): 332–37.

Table: Identification Numbers and Ages of 30 Children

Child ID#	Age
1	12
2	11
3	9
4	8
5	11
6	10
7	11
8	10
9	7
10	6
11	12
12	10
13	10
14	9
15	10

Child ID#	Age
16	7
17	7
18	8
19	6
20	9
21	7
22	9
23	8
24	11
25	9
26	12
27	12
28	11
29	12
30	12

5 Flip a coin to determine if you will assign a child to the Method 1 group or to the Method 2 group. Start by assigning Child # 1 to a group. A toss of *heads* assigns a child to the Method 1 group, and *tails* assigns a child to the Method 2 group. Write the ages of the children assigned to each group below. This process will probably not result in exactly 15 children in each group, but that is okay.

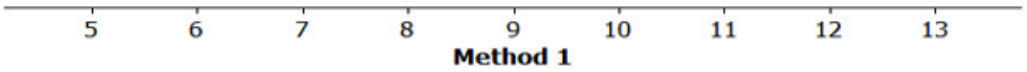
Ages of Children in Method 1 Group:

Ages of Children in Method 2 Group:

6 A Calculate the average age for each group based on Question 5.

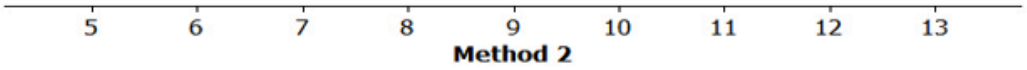
B Is one average much larger than the other, or are the average ages about the same?

7 Below are two number lines. There is one number line for each surgical method. Construct dotplots on the paper for the ages of the children in each group. **Note**: If completing this problem online, follow the instructions online to create your dotplots.



Method 1

Average Age:



Method 2

Average Age:

8 Think about the dotplots and their averages. Was the method of random assignment successful in creating groups with similar ages? Explain your answer.

9 Other variables that might affect recovery time are weight and fitness level. Do you think that our random assignment to experimental treatments (Method 1 group and Method 2 group) would create groups of similar weight and fitness level? Why do you think so?

LET'S SUMMARIZE

- Bias in surveys can be introduced by the way the sample is selected (selection bias) or by the way the data are collected once the sample is selected (response or nonresponse bias).
- Random assignment tends to create comparable groups. It does not tend to favor one condition over the other.
- Stratified random sampling is commonly used when the differences that exist among people in each subgroup are less than the differences that exist among people in the entire population.

Exercise 1.5

The article, "I'd Like to Buy a Vowel, Drivers Say", suggests that young people prefer automobile names consisting of just numbers and/or letters that do not form a word (such as Hyundai's XG300, Mazda's 626, and BMW's 325i).[17] The article says that Hyundai first planned to identify the car that was eventually marketed as the XG300 with the name *Concerto* until it realized that consumers thought XG300 sounded more "technical" and deserving of a higher price.

You are interested in finding out how much more students at your college would be willing to pay for an XG300 than a comparable car called *Concerto*.

Suppose a list of all the students at your college is available. Assume that there are 15,000 students at your college.

1 Describe a method for selecting a **simple random sample** of 300 students from this population.

2 Describe a method for selecting a **systematic sample** of about 300 students from this population.

3 Suppose the list of students also includes their ages. Describe a method for selecting a **stratified random sample** that is based on age.

[17] http://usatoday30.usatoday.com/money/autos/2001-08-09-vowel.htm

4 To create a stratified random sample for this survey, which is the best method?

(i) Divide the college into groups by major program of study and randomly pick students from each group.

(ii) Divide the college into random groups then randomly pick students from each group.

(iii) Divide the college into groups by car preference (economy car, midsize car, etc.), then randomly pick students from each group.

(iv) Divide the college into age groups (15-24, 25-34, etc.), then randomly pick students from each group.

5 Researchers were interested in learning whether people with higher levels of education use the Internet differently than people who do not have as much formal education.[18] To answer this question, the researchers used data from a national telephone survey. Approximately 1,300 households were selected for the survey, and 270 households completed the interview. What type of bias should the researchers be concerned about and why?

6 The report, "Majoring in Money: How college students and other young adults manage their money" includes the results of a survey of college students.[19] Based on the data collected, the report states that about 60% of students with credit cards pay them off each month. The report also states that the average outstanding balance on undergraduates' credit cards was about $1,183 over the 12 months prior to the survey. The survey data was from 810 college students who are currently attending a technical school, two-year undergraduate institution, or four-year undergraduate institution. What concerns do you have about generalizing these conclusions to the population of all undergraduate students?

[18] https://www.researchgate.net/publication/249683187_Digital_InequalityDifferences_in_Young_Adults'_Use_of_the_Internet

[19] https://www.salliemae.com/content/dam/slm/Media/images/Research/Majoring-In-Money-Report-2019.pdf

You want to test whether an online homework system can help students in statistics classes. Some students will do homework from a book. Other students will use the online homework system. There are 120 students signed up for Statistics. There will be four Statistics classes, and each of these classes will have 30 students. There are two professors, A and B, and each professor will teach two of the Statistics classes.

7 One way to choose the students for each treatment would be to assign students who are taking Statistics with Professor A to use book homework, and students who are taking Statistics with Professor B to use the online homework system. Explain the problem with this method of assigning students to groups.

8 Describe how you could use random assignment to decide who will be in the book homework group and who will be in the online homework group.

1.5 Monitor (survey)

Monitor your progress on learning the objectives for this unit. If you identify any objectives you need to review, go back through the unit's activities. Your responses are not graded, but will be available for your instructor to see.

Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Skill or Concept: I can …	**Rating from 1 to 5**
Distinguish between simple random sampling, stratified random sampling, and systematic sampling.	
Explain why it is desirable to create similar groups when considering stratified random sampling.	
Explain how the wording of survey questions has the potential to introduce bias.	
Describe a process for randomly assigning experimental units to experimental conditions.	