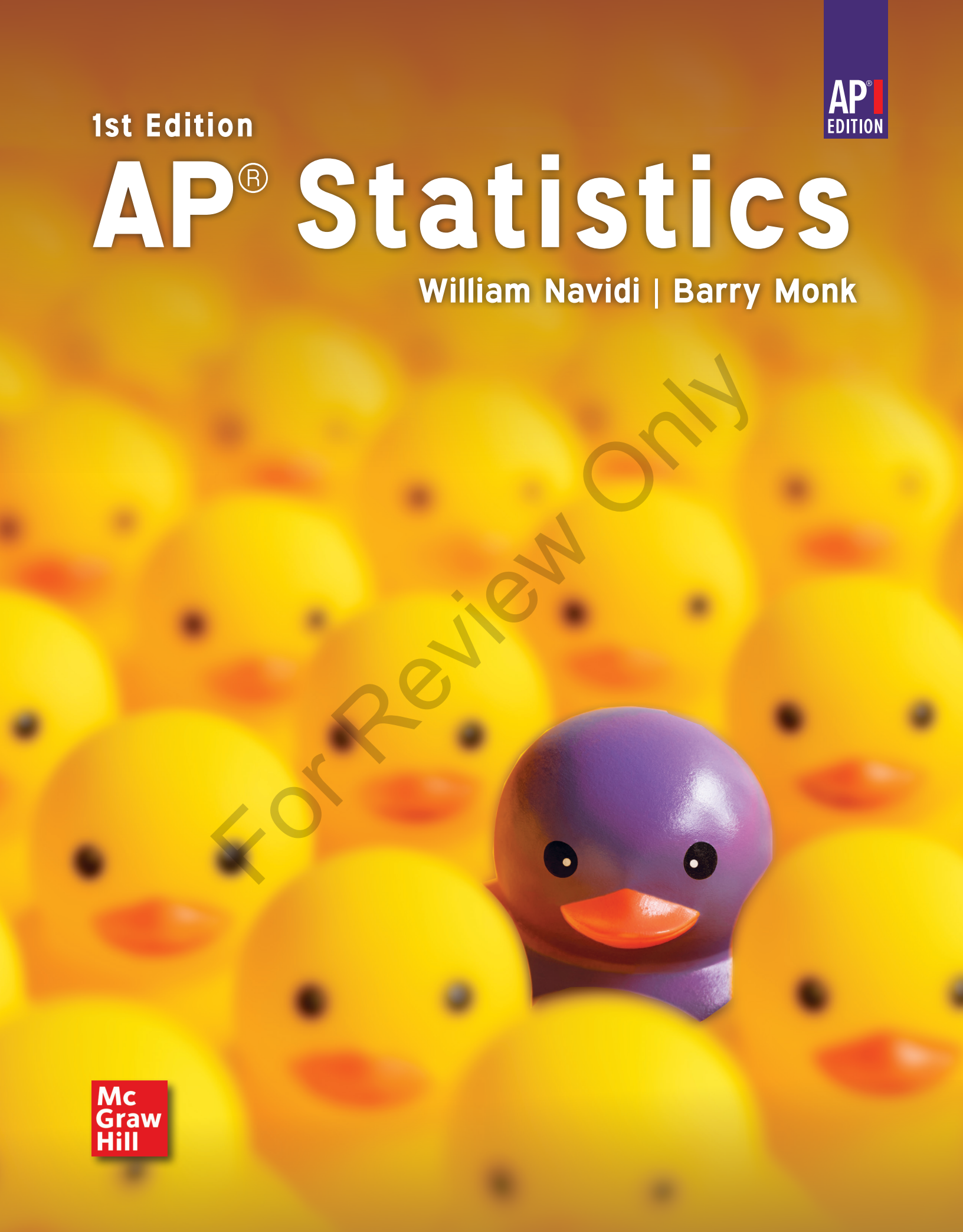
1st Edition

AP® Statistics

William Navidi | Barry Monk

AP® EDITION



Mc Graw Hill

AP® Statistics

First Edition

William Navidi
Colorado School of Mines

Barry Monk
Middle Georgia State University

AP STATISTICS

This book is printed on acid-free paper.
$PrintCode
ISBN 978-1-26-459786-4
MHID 1-26-459786-X

Front/Back Cover Image: enviromantic/iStock/Getty Images Plus/Getty Images

mheducation.com/advancedplacement

To Catherine, Sarah, and Thomas
—William Navidi

To Shaun
—Barry Monk

About the Authors

William Navidi

William Navidi is a Professor Emeritus of Applied Mathematics and Statistics at the Colorado School of Mines in Golden, Colorado. He received a Bachelor's degree in Mathematics from New College, a Master's degree in Mathematics from Michigan State University, and a Ph.D. in Statistics from the University of California at Berkeley. Bill began his teaching career at the County College of Morris, a two-year college in Randolph, New Jersey. He has taught mathematics and statistics at all levels, from developmental through the graduate level. Bill has written two Engineering Statistics textbooks for McGraw Hill.

Barry Monk

Barry Monk is a Professor of Mathematics at Middle Georgia State University in Macon, Georgia. He holds a Bachelor of Science in Mathematical Statistics, a Master of Arts in Mathematics (specializing in Optimization and Statistics), and a Ph.D. in Applied Mathematics, all from the University of Alabama. He has been teaching introductory statistics since 1992 in both classroom and online settings. He also holds a minor in Creative Writing and brings a storyteller's attention to making statistics accessible and relevant to students across disciplines.

(t)William Navidi; (b)Dawn Sherry

AP Contributors

Sarah M. Johnson is a veteran educator and AP Statistics instructor who has been teaching Mathematics and Statistics at Grand Blanc High School in Grand Blanc, Michigan, for more than 22 years. She began teaching AP Statistics early in her career and now serves as an AP Professional Learning Consultant for the College Board's Midwest Region, providing professional development and support to teachers nationwide. Sarah has contributed extensively to the AP Statistics program in multiple roles, including AP Statistics Table Leader, AP Rubric Team Member, AP Reader, AP Multiple Choice Item Author, AP Advocate, AP Pilot Reader, and National Delegate to the Academic Assembly of the College Board. Sarah lives in Michigan with her husband, a marine engineer and program manager, and together they are raising four daughters. Outside the classroom, she and her therapy dog, Ziggy, volunteer as a certified therapy dog team supporting students and community well-being.

Jared Derksen lives with his family in sunny Southern California. He has a Bachelor's and a Master's degrees in pure mathematics. He has taught mathematics students from 7th grade through college since 1991, focusing on AP Statistics beginning in 1998. Jared became an AP Reader in 2004. As a content creator, he has written material for The College Board and most recently he joined DeAnna McDonald as the co-author of *5 Steps to a 5 for AP Statistics.* In his free time, he enjoys his family, the Dodgers, good food, and great books.

David G. Pepino serves as the Mathematics Department Chair at Bishop Eustace Preparatory School and has been teaching since 2000. He holds a background in mathematics and economics, a master's degree from Villanova University, and has completed graduate studies in education at the University of Pennsylvania. David founded the school's chapter of the National Mathematics Honor Society, Mu Alpha Theta, and has chaired the Middle States Evaluation committees for Assessment and Evidence of Student Learning as well as Curriculum, Instruction, and Assessment. His extracurricular contributions include developing and implementing a European travel-abroad program, leading community outreach initiatives such as a homeless support project, and serving as a Kairos retreat group leader.

Jeff Eicher, Jr. is a Data Science professor at Eastern University. Before transitioning to teaching undergraduate and graduate students, he taught statistics and AP Statistics for 11 years to high school students, along with Math 2, Geometry, Math 3, Pre-Calculus, and AP Calculus. He was part of the College Board's Instructional Design Team in 2018-19 and has been an AP Statistics exam reader for 8 years. He served as a National Math and Science Initiative mentor, tutor, and content creator for 3 years. He is currently pursuing his PhD in Applied Data Science.

Nicholas Reeves grew up in Albuquerque and Santa Fe, NM and earned Education and Theology degrees in 2000 and 2004 from Hardin-Simmons University in Abilene, TX. He received his Master of Science degree in Statistics from Texas A&M University in 2016. A certified high school mathematics teacher, Nick has been teaching at Lubbock High School in Lubbock, TX, since 2007, teaching AP Statistics since 2010. Nick has been an Exam Reader for AP Statistics for 7 years and was named Lubbock High Teacher of the Year in 2019. He enjoys taking the experiences and interests of his students and turning them into statistical situations for his class to study.

Brief Table of Contents

Contents

Special Features

Case Study

AP Statistics Skills Practice

Writing in AP Statistics

To the AP Statistics Student

Welcome to AP Statistics — a course about risk, randomness, and reward, and about learning to think clearly in a world overflowing with information.

Every day you are bombarded with data. News headlines report percentages, companies advertise studies, social media shares graphs, and research claims compete for attention. In this environment, it is not enough to see numbers — you must learn to question them. To succeed in modern life, you must become a thoughtful and responsible consumer of data, able to identify or classify information, interpret results, and justify conclusions.

AP Statistics is not simply a math class. While you will calculate and estimate, the heart of this course is communication and statistical thinking. You will read technical passages carefully, describe patterns in context, compare results, and construct clearly written explanations. You will complete investigations that require you not only to determine answers, but to explain what those answers mean.

Throughout the course, you will explore how randomness creates variability and how uncertainty shapes real-world decisions. You will design studies, analyze data, and evaluate claims by asking: What can we conclude? What are the risks? What evidence supports this conclusion? Just as importantly, you will learn when conclusions cannot be made.

This class emphasizes careful reasoning. You will interpret graphs, justify statistical arguments, and communicate your findings with precision. The goal is not just to arrive at a number, but to understand and explain the story behind the data.

By the end of this course, you will think more critically, read more carefully, and write more clearly about evidence. You will be prepared not only for the AP exam, but for a world where informed decisions depend on your ability to analyze, question, and communicate with confidence.

Welcome to AP Statistics.

Sarah Johnson
AP Contributor

About *AP Statistics*

This thoughtfully crafted program is designed to inspire and empower students on their journey to mastering the course content and concepts. Built on the trusted foundation of *Elementary Statistics*—the best-selling college-level textbook by William Navidi and Barry Monk—this program transforms the latest AP Statistics Framework into an engaging, accessible, and impactful learning experience.

At the heart of *AP Statistics* are three guiding principles: **Clarity, Quality,** and **Accuracy**. Every concept is presented with precision and care, ensuring students can easily grasp even the most complex ideas. From the first introduction of a topic to its application on the AP Exam, the program provides a seamless learning path supported by high-quality exercises, real-world examples, and AP-style practice questions.

What sets this program apart is its connection to the diverse applications in the real world. Students explore data sets rooted in health, social sciences, and popular culture—making statistics relevant and relatable across many fields. Special features like **AP Skills Practice** and **Writing in AP Statistics** empower students to master the skills and task verbs essential for success on the AP Exam, addressing common challenges with confidence-building strategies.

With a program this engaging, the journey to AP success becomes not just achievable—but inspiring.

- Unit and chapter content are in lock-step alignment to the organization of the course and the AP Framework's topics, practices, and skills.
- The high school student-friendly design is welcoming and accessible.
- Multiple, engaging features offer students the chance to revisit and practice key concepts, helping them overcome common challenges with confidence.
- AP-style multiple choice and free response questions are available at the end of each chapter and unit to give students continuous practice with exam-style questions.
- SmartBook® creates adaptive reading experiences that are personal, purposeful, and productive. Mobile-friendly access across all devices allows students to begin a SmartBook® assignment on their laptop and finish it on their smartphone.
- ALEKS for AP Statistics (available as an add-on to the digital subscription) uses adaptive questioning to quickly and accurately determine exactly what statistics topics a student knows and doesn't know and instructs each student on the topics they are most ready to learn.

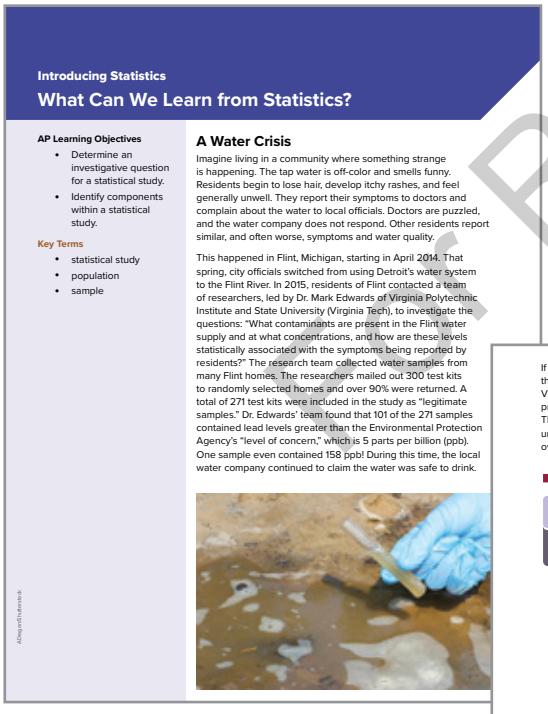
AP Success Starts Here!

AP Statistics is meticulously structured to align with the AP Framework, organizing content into five AP units and ten chapters that encompass all topics and learning objectives in the Course and Exam Description. Special attention is given to the **Statistical Practices** and **Skills**, with dedicated features guiding students through the statistical problem-solving process and helping them understand how these skills will be evaluated on the Exam. This thoughtful approach allows teachers to prioritize essential concepts and content without distraction.

High-interest features and real-world data sets immerse students in the exploration of statistical concepts, making learning both engaging and meaningful. By connecting statistics to diverse subject areas—many tied to students' everyday experiences—the program transforms abstract ideas into relevant and fascinating insights. This applied approach not only sparks curiosity but also empowers students to see the practical impact of statistics in understanding and solving real-world challenges.

Enhanced Pedagogy and Real-World Connections

Starting with the introduction, "What Can We Learn from Statistics?"—which explores the components of statistical studies and introduces the investigative question—the program presents statistical concepts in a clear and accessible manner, perfectly tailored to the AP course.



Introducing Statistics

What Can We Learn from Statistics?

AP Learning Objectives

- Determine an investigative question for a statistical study.
- Identify components within a statistical study.

Key Terms

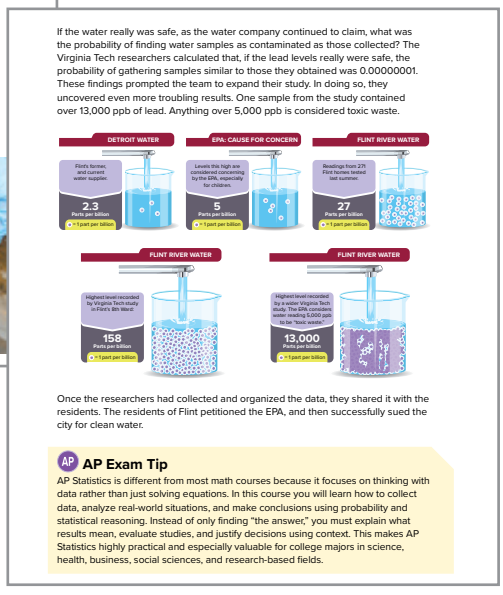
- statistical study
- population
- sample

A Water Crisis

Imagine living in a community where something strange is happening. The tap water is off-color and smells funny. Residents begin to lose hair, develop itchy rashes, and feel generally unwell. They report their symptoms to doctors and complain about the water to local officials. Doctors are puzzled, and the water company does not respond. Other residents report similar, and often worse, symptoms and water quality.

This happened in Flint, Michigan, starting in April 2014. That spring, city officials switched from using Detroit's water system to the Flint River. In 2015, residents of Flint contacted a team of researchers, led by Dr. Mark Edwards of Virginia Polytechnic Institute and State University (Virginia Tech), to investigate the questions: "What contaminants are present in the Flint water supply and at what concentrations, and how are these levels statistically associated with the symptoms being reported by residents?" The research team collected water samples from many Flint homes. The researchers mailed out 300 test kits to randomly selected homes and over 90% were returned. A total of 271 test kits were included in the study as "legitimate samples." Dr. Edwards' team found that 101 of the 271 samples contained lead levels greater than the Environmental Protection Agency's "level of concern," which is 5 parts per billion (ppb). One sample even contained 158 ppb! During this time, the local water company continued to claim the water was safe to drink.

Introducing statistics through **real-life situations** prepares students to analyze data in everyday contexts and professional settings. This approach builds confidence and equips learners with skills they can directly apply to decision-making and problem-solving in the real world.



If the water really was safe, as the water company continued to claim, what was the probability of finding water samples as contaminated as those collected? The Virginia Tech researchers calculated that, if the lead levels really were safe, the probability of gathering samples similar to those they obtained was 0.00000001. These findings prompted the team to expand their study. In doing so, they uncovered even more troubling results. One sample from the study contained over 13,000 ppb of lead. Anything over 5,000 ppb is considered toxic waste.

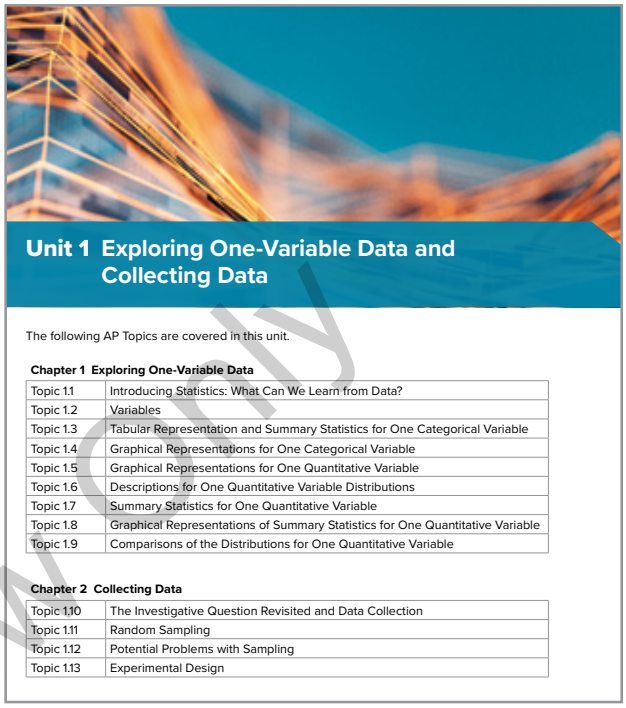
Once the researchers had collected and organized the data, they shared it with the residents. The residents of Flint petitioned the EPA, and then successfully sued the city for clean water.

AP Exam Tip

AP Statistics is different from most math courses because it focuses on thinking with data rather than just solving equations. In this course you will learn how to collect data, analyze real-world situations, and make conclusions using probability and statistical reasoning. Instead of only finding "the answer," you must explain what results mean, evaluate studies, and justify decisions using context. This makes AP Statistics highly practical and especially valuable for college majors in science, health, business, social sciences, and research-based fields.

Unit and Chapter Organization Precisely Aligns with the Framework

Each **Unit Opener** maps chapters directly to AP topics, giving students and instructors an instant view of what lies ahead and how it connects to exam preparation.



Unit 1 Exploring One-Variable Data and Collecting Data

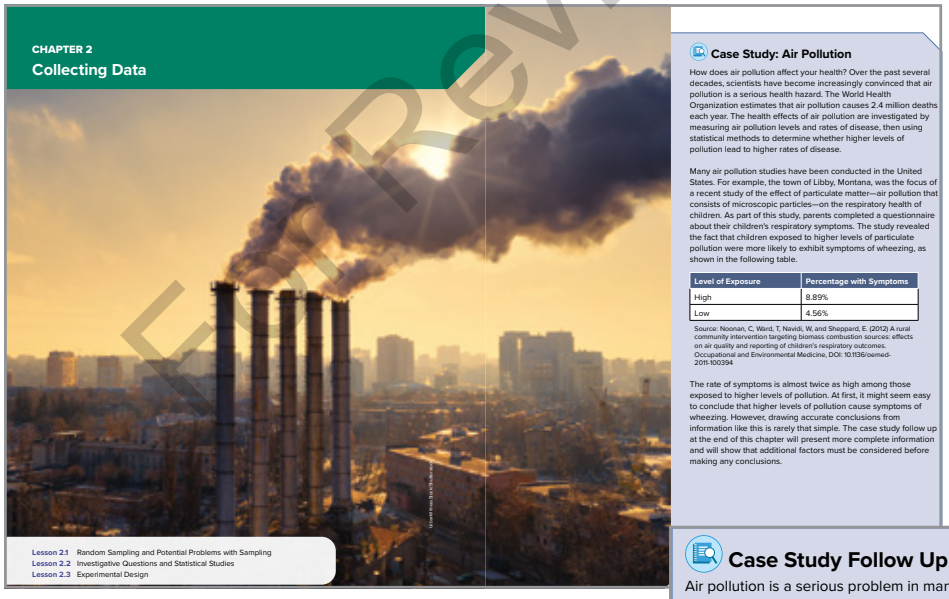
The following AP Topics are covered in this unit.

Chapter 1 Exploring One-Variable Data

Topic 1.1	Introducing Statistics: What Can We Learn from Data?
Topic 1.2	Variables
Topic 1.3	Tabular Representation and Summary Statistics for One Categorical Variable
Topic 1.4	Graphical Representations for One Categorical Variable
Topic 1.5	Graphical Representations for One Quantitative Variable
Topic 1.6	Descriptions for One Quantitative Variable Distributions
Topic 1.7	Summary Statistics for One Quantitative Variable
Topic 1.8	Graphical Representations of Summary Statistics for One Quantitative Variable
Topic 1.9	Comparisons of the Distributions for One Quantitative Variable

Chapter 2 Collecting Data

Topic 1.10	The Investigative Question Revisited and Data Collection
Topic 1.11	Random Sampling
Topic 1.12	Potential Problems with Sampling
Topic 1.13	Experimental Design



CHAPTER 2
Collecting Data

Lesson 2.1 Random Sampling and Potential Problems with Sampling
Lesson 2.2 Investigative Questions and Statistical Studies
Lesson 2.3 Experimental Design

Case Study: Air Pollution

How does air pollution affect your health? Over the past several decades, scientists have become increasingly convinced that air pollution is a serious health hazard. The World Health Organization estimates that air pollution causes 2.4 million deaths each year. The health effects of air pollution are investigated by measuring air pollution levels and rates of disease, then using statistical methods to determine whether higher levels of pollution lead to higher rates of disease.

Many air pollution studies have been conducted in the United States. For example, the town of Libby, Montana, was the focus of a recent study of the effect of particulate matter—air pollution that consists of microscopic particles—on the respiratory health of children. As part of this study, parents completed a questionnaire about their children's respiratory symptoms. The study revealed the fact that children exposed to higher levels of particulate pollution were more likely to exhibit symptoms of wheezing, as shown in the following table.

Level of Exposure	Percentage with Symptoms
High	8.89%
Low	4.56%

Source: Noonan, C, Ward, T, Navidi, W, and Sheppard, E. (2012) A rural community intervention targeting biomass combustion sources: effects on air quality and reporting of children's respiratory outcomes. Occupational and Environmental Medicine, DOI: 10.1136/oemed-2011-100394

The rate of symptoms is almost twice as high among those exposed to higher levels of pollution. At first, it might seem easy to conclude that higher levels of pollution cause symptoms of wheezing. However, drawing accurate conclusions from information like this is rarely that simple. The case study follow up at the end of this chapter will present more complete information and will show that additional factors must be considered before making any conclusions.

Each chapter begins and ends with an engaging **Case Study** designed to activate learning by asking students to apply statistical concepts to real-world situations. Through authentic data and application questions, students analyze evidence, make assumptions, and connect statistical ideas to everyday contexts.

Case Study Follow Up

Air pollution is a serious problem in many places. One form of air pollution that is suspected to cause respiratory illness is particulate matter (PM), which consists of tiny particles in the air. Particulate matter can come from many sources, most commonly ash from burning, but also from other sources such as tiny particles of rubber that wear off automobile and truck tires.

The **Case Study Follow Up** at the chapter's conclusion challenges students to apply what they've learned and use statistics to confirm or revise their initial assumptions from the opening Case Study. This approach not only deepens understanding through developing critical thinking, problem solving, and decision-making skills, it also empowers students to see the practical value of statistics in solving real-world problems.

LESSON 2.2

Investigative Questions and Statistical Studies

AP Learning Objectives

- Determine the components of an investigative question within a statistical study.
- Identify an experiment.
- Identify an observational study.
- Justify the appropriateness of generalizations for a statistical study.
- Justify the appropriateness of conclusions based on a well-designed experiment.

Key Terms

- response variable
- explanatory variable
- experiment
- experimental units

Investigative Questions

Recall that a statistical study is one in which data are collected from a sample to answer an investigative question about a larger population. In practice, a study will often have several investigative questions. For example, in a study to compare a new medical treatment to a currently used treatment, one question is likely to be "Is the new treatment more likely to cure the disease than the current treatment?" Other questions might include "Is the new treatment less likely to produce side effects?" and "Does the new treatment cure the disease more quickly?" When designing a statistical study, the most important and earliest step is to develop the investigative questions. What do we want to know? What important questions do we want to answer? The investigative questions impact every decision in the statistical process. They influence how the researcher will collect the data, what kind will be collected, how to organize and analyze the data, and what conclusions can be made from the data.

A well-developed investigative question has three components:

1. The question should clearly identify the variables we are interested in studying. There are often two variables, the explanatory variable and the response variable. The **response variable** measures the outcome of the statistical study. The **explanatory variable** helps explain or predict how the response variable changes.

Each chapter is thoughtfully divided into discrete lessons, each aligned with specific **AP Learning Objectives**, creating a clear and progressive learning path. This structured approach ensures students build a strong foundation and deepen their understanding as they advance through the material.

Key Terms are listed at the beginning of each lesson and are defined at point-of-use within the lesson to support AP Statistics vocabulary acquisition.

Support for Learning and Mastery

Within each lesson, concepts are brought to life through an **Example** and **Solution** structure, providing students with a clear, step-by-step approach to applying statistical principles and analyzing data. By breaking down complex ideas into manageable steps, the concepts become more accessible and students gain a deeper understanding and mastery of statistical processes.

Example: Identifying an Experiment

A pet food company is interested in knowing if a new formulation of goldfish food will cause goldfish to grow faster. They conduct an experiment in which they randomly select 10 goldfish of similar age, weight, and size from a local pet store. Each fish is placed in an identical fishbowl. Five will receive the new formulation of goldfish food, and the other five will receive the old formulation. They will be fed the same amounts at the same time of day for 30 days. At the end of 30 days, the researchers will measure the weights of the goldfish and compare them across the treatment groups.

a. Explain why this study is an experiment.

b. Identify the experimental units, factors, levels, and treatments.

Solution:

a. This statistical study is an experiment because treatments are being assigned to the experimental units.

b. The experimental units are the ten goldfish. The factor is food formulation. There are two levels – new formulation and old formulation. This leads to two treatments; five fish receive the new formulation of food, and five fish receive the old formulation.

tanuha2001/Shutterstock

Check Your Understanding

Determine the sampling method in each of the given studies.

1. Every 10 years, the U.S. Census Bureau attempts to count every person living in the United States. To check the accuracy of their count in a certain city, they draw a sample of census districts (roughly equivalent to a city block) and recount everyone in the sampled districts. What kind of sample is formed by the people who are recounted? Explain.
2. A public health researcher is designing a study of the effect of diet on heart disease. The researcher knows that the diets of males and females tend to differ and that males are more susceptible to heart disease. To be sure that both males and females are well represented, the study comprises a simple random sample of 100 males and another simple random sample of 100 females. What kind of sample do these 200 people represent? Explain.
3. A college basketball team held a promotion at one of its games in which every 20th person who entered the arena won a free basketball. What kind of sample do the winners represent? Explain.
4. To select people to call for jury duty, a numbered list of all resident licenses or ID cards is made. Then random numbers are generate corresponding to those numbers are selected. What kind of sampl

After each concept is explained, one or more **Check Your Understanding** exercises are strategically placed to help students assess their grasp of the material. These exercises empower students with the confidence to progress or identify areas that require further review.

Using Technology

The following steps show how to use Desmos to calculate the mean and standard deviation for the discrete probability distribution below.

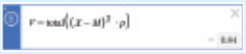
x	0	1	2	3
$P(x)$	0.1	0.4	0.3	0.2

Step 1: On the first line, enter a list of the values: $X = [0, 1, 2, 3]$.

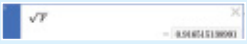
Step 2: On the second line, enter a list of the corresponding probabilities (in the same order): $p = [0.1, 0.4, 0.3, 0.2]$.

Step 3: On the third line, type M = **total**$(X \cdot p)$. This command tells Desmos to total or sum the products of each X-value with its corresponding probability. The result is the mean (expected value).

Step 4: On the fourth line, type V = **total**$((X - M)^2 \cdot p)$ This computes the variance using the formula.



Step 5: On the fifth line, type **sqrt(V)**. Desmos displays the standard deviation, which is the square root of the variance.



Using Technology features walk students step-by-step through the process of finding solutions and graphing data using the same Desmos graphing calculator that they will use on the AP Statistics Exam. Desmos is also available for students to use in the digital course.

AP Exam Tip

You will likely be asked to describe an experimental design on the AP Exam. It is important that you name each treatment group. Simply calling it "Treatment Group 1" is not sufficient. You must be specific about what treatment that group is receiving.

AP Exam Tips are integrated throughout to prepare students for what they will encounter on the exam.

Take Another Look

Another way to describe a confounding variable: A confounding variable is a variable related to the explanatory variable that can cause the response variable to have different outcomes.

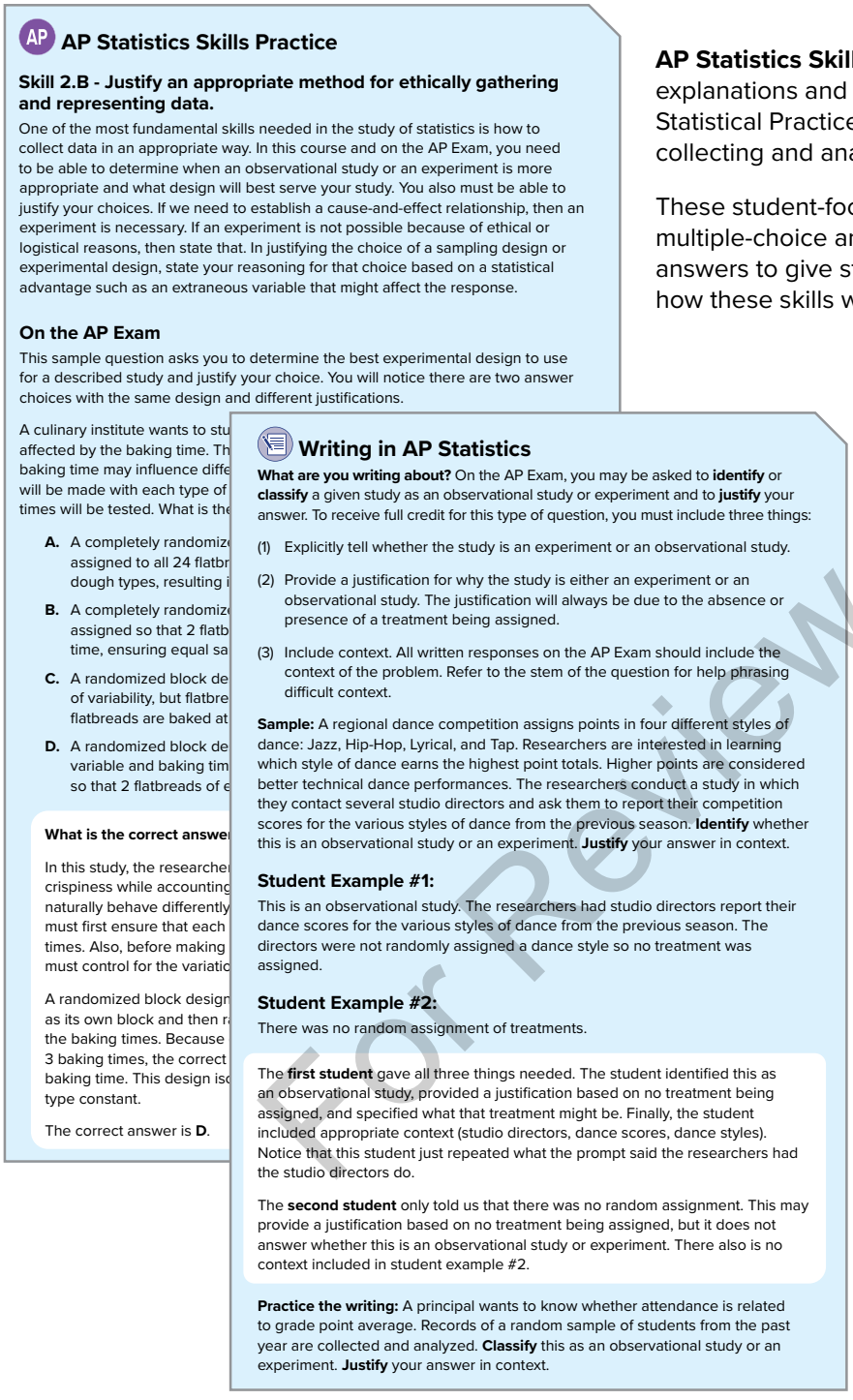
The **Take Another Look** features provide students with additional explanations that encourage reflection, break down the concept into manageable parts, and provide reinforcement to increase retention.

CAUTION

A sample is not randomly selected when observational units are deliberately chosen or volunteer themselves to be in the sample. Do not use a convenience sample when it is possible to select a random sample.

The **Caution** features highlight potential pitfalls where students might struggle, offering thoughtful tips and practical guidance to help them navigate challenges with ease.

Focus on Essential AP Skills



AP Statistics Skills Practice

Skill 2.B - Justify an appropriate method for ethically gathering and representing data.

One of the most fundamental skills needed in the study of statistics is how to collect data in an appropriate way. In this course and on the AP Exam, you need to be able to determine when an observational study or an experiment is more appropriate and what design will best serve your study. You also must be able to justify your choices. If we need to establish a cause-and-effect relationship, then an experiment is necessary. If an experiment is not possible because of ethical or logistical reasons, then state that. In justifying the choice of a sampling design or experimental design, state your reasoning for that choice based on a statistical advantage such as an extraneous variable that might affect the response.

On the AP Exam

This sample question asks you to determine the best experimental design to use for a described study and justify your choice. You will notice there are two answer choices with the same design and different justifications.

A culinary institute wants to stu
affected by the baking time. Th
baking time may influence diffe
will be made with each type of
times will be tested. What is the

A. A completely randomiz
assigned to all 24 flatbr
dough types, resulting i

B. A completely randomiz
assigned so that 2 flatb
time, ensuring equal sa

C. A randomized block de
of variability, but flatbre
flatbreads are baked at

D. A randomized block de
variable and baking tim
so that 2 flatbreads of e

What is the correct answe

In this study, the researche
crispiness while accounting
naturally behave differently
must first ensure that each
times. Also, before making
must control for the variatic

A randomized block design
as its own block and then r
the baking times. Because
3 baking times, the correct
baking time. This design isc
type constant.

The correct answer is **D**.

Writing in AP Statistics

What are you writing about? On the AP Exam, you may be asked to **identify** or **classify** a given study as an observational study or experiment and to **justify** your answer. To receive full credit for this type of question, you must include three things:

(1) Explicitly tell whether the study is an experiment or an observational study.

(2) Provide a justification for why the study is either an experiment or an observational study. The justification will always be due to the absence or presence of a treatment being assigned.

(3) Include context. All written responses on the AP Exam should include the context of the problem. Refer to the stem of the question for help phrasing difficult context.

Sample: A regional dance competition assigns points in four different styles of dance: Jazz, Hip-Hop, Lyrical, and Tap. Researchers are interested in learning which style of dance earns the highest point totals. Higher points are considered better technical dance performances. The researchers conduct a study in which they contact several studio directors and ask them to report their competition scores for the various styles of dance from the previous season. **Identify** whether this is an observational study or an experiment. **Justify** your answer in context.

Student Example #1:

This is an observational study. The researchers had studio directors report their dance scores for the various styles of dance from the previous season. The directors were not randomly assigned a dance style so no treatment was assigned.

Student Example #2:

There was no random assignment of treatments.

The **first student** gave all three things needed. The student identified this as an observational study, provided a justification based on no treatment being assigned, and specified what that treatment might be. Finally, the student included appropriate context (studio directors, dance scores, dance styles). Notice that this student just repeated what the prompt said the researchers had the studio directors do.

The **second student** only told us that there was no random assignment. This may provide a justification based on no treatment being assigned, but it does not answer whether this is an observational study or experiment. There also is no context included in student example #2.

Practice the writing: A principal wants to know whether attendance is related to grade point average. Records of a random sample of students from the past year are collected and analyzed. **Classify** this as an observational study or an experiment. **Justify** your answer in context.

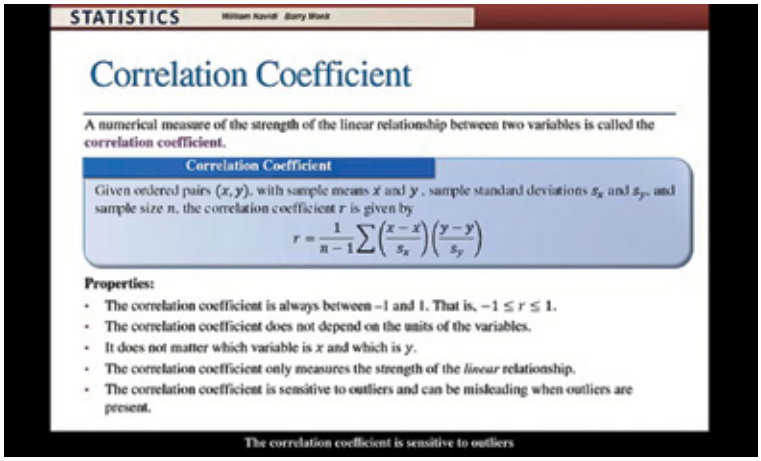
AP Statistics Skills Practice features include explanations and practical applications of the AP Statistical Practices, from formulating questions, collecting and analyzing data, to interpreting results.

These student-focused features provides sample multiple-choice and free-response questions and answers to give students a deeper understanding of how these skills will be assessed on the Exam.

Writing in AP Statistics tackles a common student challenge on the Exam. This feature provides invaluable guidance for mastering Free-Response questions and demystifies what exam readers expect when students are asked to **describe, identify, interpret,** or **justify** their answers.

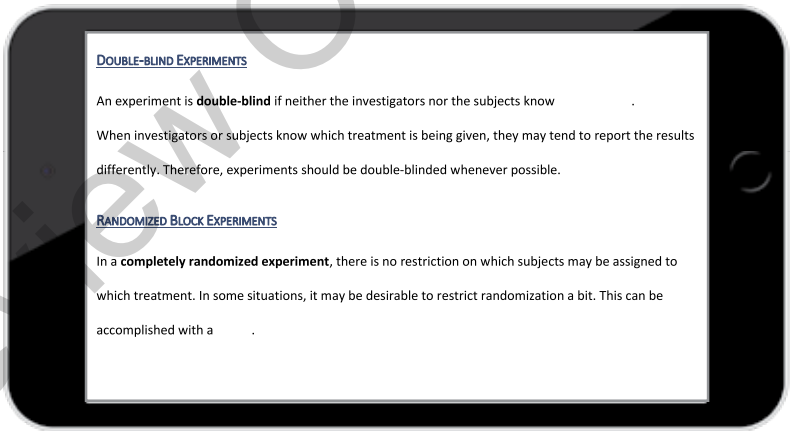
By analyzing high-scoring and low-scoring examples, students gain insight into best practices and learn how to craft responses that meet AP standards. This targeted support not only builds confidence but also equips students with the skill set to excel on their written explanations.

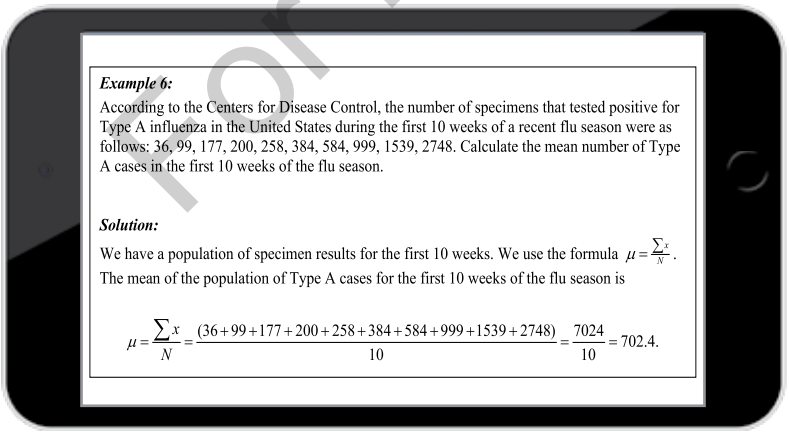
Enriched Digital Support



The **videos** in the digital course vividly illustrate essential concepts, definitions, formulas, and problem-solving methods covered in the text. They deliver clear demonstrations and diverse methods, empowering students with a deeper, more comprehensive understanding.

Guided Student Notes structure learning to promote active participation, enhancing comprehension. By emphasizing key points, they highlight essential content to support better retention and deeper mastery. Organized by AP Topic, these notes can also serve as a pre-exam review guide.





The **Corequisite Workbook** provides just-in-time support for students who need additional practice with introductory skills. Available as either an interactive eBook or a printable PDF, the workbook is organized around the key foundational skills necessary for course success. Topics are introduced with multiple examples and assessed through practice problems. Teachers can assign the workbook as needed or let students use it as a self-guided study tool.

Kahoot! gamifies the content using interactive quizzes, puzzles, true/false questions and more to help students review important material in an interactive and engaging way.

Key AP Takeaways and Practice

Lesson 2.1 Summary

Key AP Takeaways

- A census is a complete count of every unit in the population. Censuses tend to be time consuming, expensive, and often impossible to conduct.
- Most populations are too large to study each member, so samples are drawn and studied.
- Using a random mechanism to select the sample helps to ensure that the sample represents the population well.
- Simple random sampling, stratified random sampling, cluster random sampling, and systematic random sampling are all valid sampling methods.
- Avoid convenience sampling if possible.
- A study is biased if it is conducted by a method that tends to produce an incorrect result.
- Some of the most common forms of bias are voluntary response bias, undercoverage bias, nonresponse bias, and response bias.

Lesson Summaries highlight **Key AP Takeaways**, bringing focus to only the most essential concepts and skills, promoting AP-level success.

Lesson 2.2 Practice

Understanding the Concepts

For exercises 1–5, determine the appropriate word or phrase.

1. What are the three components of a good investigative question?
2. What are the two types of variables in a statistical study called?
3. What is a statistical study called in which the assignment to treatment groups is not made by the investigator?
4. What is a variable that
variables?
5. How can the results of
6. What type of study re
explanatory and respo

**For exercises 7–9, identi
and response variables f**

7. Among students in a t
instruction (graphic or
students' scores on th
8. Among students at a l
message from the school nurse, compared with no re
in the average number of hours of sleep students get

Working with the Concepts

20. You are giving me a headache: A pharmaceutical company wants to test a new drug that is designed to provide superior relief from headaches. They want to select a sample of headache sufferers to try the drug. Do you think that it is feasible to draw a simple random sample of headache sufferers, or will it be necessary to use a convenience sample? Explain your reasoning.

21. Pay more for recreati
university wants to sa
increase in their recre
you think it is feasible
necessary to use a co

Extending the Concepts

31. Draw a sample: Suppose that you are asked to determine students' opinions at your school about a potential change in library hours. Describe how you could select a sample of each of the following types: simple random sample, stratified sample, cluster sample, systematic sample.

32. *Literary Digest* poll: In the 1936 presidential election, Republican candidate Alfred Landon challenged President Franklin Roosevelt. The *Literary Digest* magazine conducted a poll in which they mailed questionnaires to more than 10 million voters. The people who received the questionnaires were drawn from lists of automobile owners and people with telephones. The magazine received 2.3 million responses and predicted that Landon would win the election in a landslide with 57% of the vote. In fact, Roosevelt won in a landslide with 62% of the vote.

a. In 1936, most people did not own automobiles and many did not have telephones. Explain how this could have caused the results of the poll to be mistaken.

b. What can be said about the response rate? Explain how this could have caused the results of the poll to be inaccurate.

c. The *Literary Digest* believed that its poll would be accurate because it received 2.3 million responses, which is a very large number. Explain how the poll could be wrong, even with such a large sample.

Lesson-level exercises offer targeted practice in Understanding, Working with, and Extending the Concepts, empowering students to broaden their understanding, strengthen problem-solving skills, and confidently apply their knowledge.

Chapter 2 Review

Reviewing the Concepts

In exercises 1–3, identify the kind of sample that is described. Justify your answer.

1. **Website ratings:** A popular website is interested in conducting a survey of 400 randomly selected visitors to the site in such a way that 200 of them will be under age 30, 150 will be aged 30–55, and 50 will be over 55.
2. **School days:** A researcher randomly selects 4 of 12 high schools in a certain region and surveys all of the administrative staff members in each school about a potential change in the ordering of supplies.
3. **Political polling:** A pollster obtains a list of registered voters and uses a computer random number generator to choose 100 of them to ask which candidate they prefer in an upcoming election.
4. **Fluoride and tooth decay:** Researchers examine the association between the fluoridation of water and the prevention of tooth decay by comparing the prevalence of tooth decay in countries that have fluoridated water with the prevalence in countries that do not.
   a. Is this a well-designed experiment or an observational study? Explain.
   b. Assume that tooth decay was less common in countries with fluoridated water. Explain why this result is due to a confounding variable.
5. **Fuel efficiency:** A study is conducted to compare the fuel efficiency of two types of car engines, A and B. Several models of car are used in the study. Within each model, three cars are randomly assigned to engine A and three to engine B.
   a. Is this a randomized block experiment or a matched pairs experiment? Explain.
   b. Assume that one type of engine had noticeably better gas mileage than the other. Could this be due to a confounding variable? Explain.

Chapter-level **Reviewing the Concepts** exercises allow students to apply their knowledge from the lessons in the chapter to bring key concepts together for a more advanced review.

Unit summaries feature **Putting It All Together**, a comprehensive overview that reinforces key concepts across chapters, and ensures a solid grasp of the unit's content. **Moving Forward with These Concepts** previews how students will expand on the AP concepts and skills gained up to this point, and prepares them for success in future units.

Unit 1 Summary

Putting It All Together

In this unit, you have learned about collecting and exploring data. At the beginning of the statistical process:

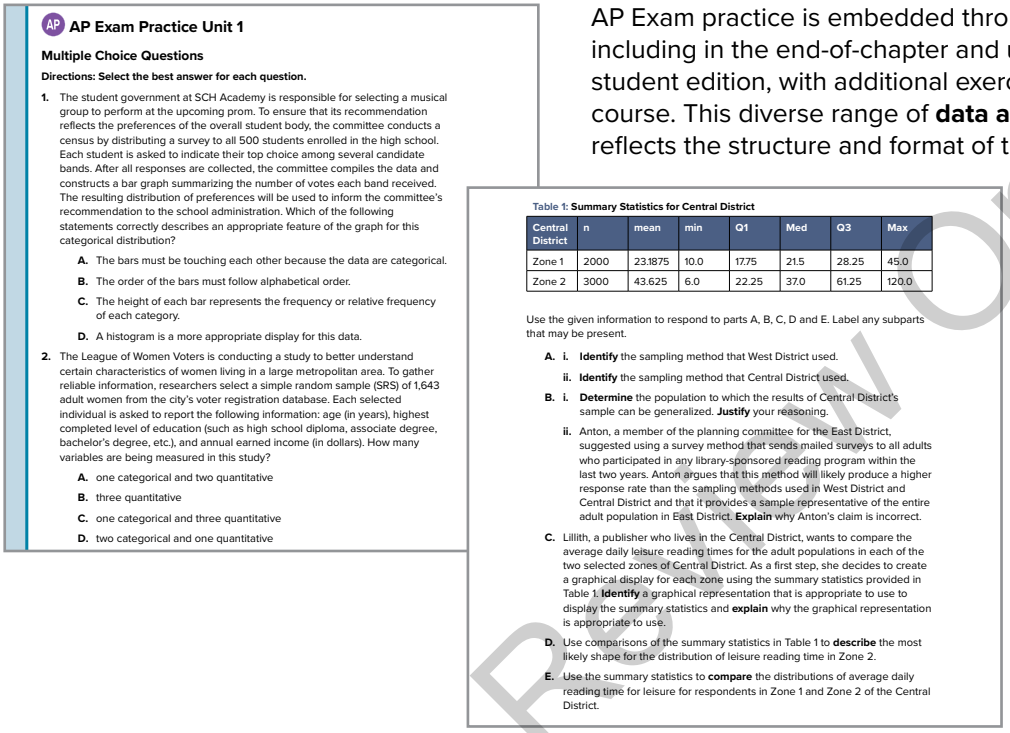
- Start with one or more investigative questions – What do we want to know about?
- Determine what kind of data to collect. What is the population that we are interested in? What variables are we interested in? Are they categorical or quantitative? If they are quantitative, are they continuous or discrete?
- What would be the best way to collect the data that we need? Do we want to do a survey? An observational study with random sampling? How are we going to reduce or prevent bias in our data? Does this statistical study call for a well-designed experiment? If so, what experimental design would be most appropriate?
- Once we get our data – how are we going to organize it? We need graphs and summary statistics, so the type of data helps inform which ones we choose. What patterns do we notice? If the data are quantitative, what can we say about its distribution?

Moving Forward with These Concepts

Now that we have our investigative questions, along with data, graphs, and summary statistics, how do we answer the questions? We have an idea of what the answer *might* be, but it is not quite enough to be certain – at least as certain as we CAN be in Statistics. Before we can answer the questions with statistical certainty, we need to learn about probability. Probability is the study of randomness. It connects the results of our data collection and exploration to answers. We begin the study of probability in the next unit.

Preparing for the AP Exam with Confidence

Students benefit from extensive AP Exam Practice opportunities in both the student edition and digital course that are designed to assess understanding and ensure thorough exam preparation. The three complete AP Practice Exams feature multiple-choice questions, including targeted question sets from Units 2 and 5, and four free-response questions. This helps students build confidence by mastering content and essential AP skills.



AP Exam Practice Unit 1

Multiple Choice Questions

Directions: Select the best answer for each question.

1. The student government at SCH Academy is responsible for selecting a musical group to perform at the upcoming prom. To ensure that its recommendation reflects the preferences of the overall student body, the committee conducts a census by distributing a survey to all 500 students enrolled in the high school. Each student is asked to indicate their top choice among several candidate bands. After all responses are collected, the committee compiles the data and constructs a bar graph summarizing the number of votes each band received. The resulting distribution of preferences will be used to inform the committee's recommendation to the school administration. Which of the following statements correctly describes an appropriate feature of the graph for this categorical distribution?
   - **A.** The bars must be touching each other because the data are categorical.
   - **B.** The order of the bars must follow alphabetical order.
   - **C.** The height of each bar represents the frequency or relative frequency of each category.
   - **D.** A histogram is a more appropriate display for this data.
2. The League of Women Voters is conducting a study to better understand certain characteristics of women living in a large metropolitan area. To gather reliable information, researchers select a simple random sample (SRS) of 1,643 adult women from the city's voter registration database. Each selected individual is asked to report the following information: age (in years), highest completed level of education (such as high school diploma, associate degree, bachelor's degree, etc.), and annual earned income (in dollars). How many variables are being measured in this study?
   - **A.** one categorical and two quantitative
   - **B.** three quantitative
   - **C.** one categorical and three quantitative
   - **D.** two categorical and one quantitative

AP Exam practice is embedded throughout the course, including in the end-of-chapter and unit questions in the student edition, with additional exercises in the digital course. This diverse range of **data and application tasks** reflects the structure and format of the actual AP Exam.

Table 1: **Summary Statistics for Central District**

Central District	n	mean	min	Q1	Med	Q3	Max
Zone 1	2000	23.1875	10.0	17.75	21.5	28.25	45.0
Zone 2	3000	43.625	6.0	22.25	37.0	61.25	120.0

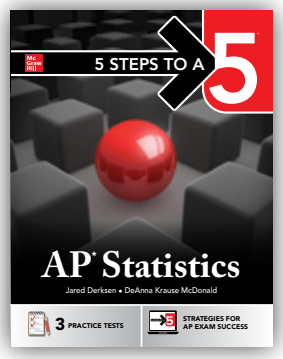
Use the given information to respond to parts A, B, C, D and E. Label any subparts that may be present.

- **A.** i. **Identify** the sampling method that West District used.
  ii. **Identify** the sampling method that Central District used.
- **B.** i. **Determine** the population to which the results of Central District's sample can be generalized. **Justify** your reasoning.
  ii. Anton, a member of the planning committee for the East District, suggested using a survey method that sends mailed surveys to all adults who participated in any library-sponsored reading program within the last two years. Anton argues that this method will likely produce a higher response rate than the sampling methods used in West District and Central District and that it provides a sample representative of the entire adult population in East District. **Explain** why Anton's claim is incorrect.
- **C.** Lillith, a publisher who lives in the Central District, wants to compare the average daily leisure reading times for the adult populations in each of the two selected zones of Central District. As a first step, she decides to create a graphical display for each zone using the summary statistics provided in Table 1. **Identify** a graphical representation that is appropriate to use to display the summary statistics and **explain** why the graphical representation is appropriate to use.
- **D.** Use comparisons of the summary statistics in Table 1 to **describe** the most likely shape for the distribution of leisure reading time in Zone 2.
- **E.** Use the summary statistics to **compare** the distributions of average daily reading time for leisure for respondents in Zone 1 and Zone 2 of the Central District.

One complete **AP Practice Exam** in the Student Edition, along with two additional, printable and online exams in the digital course, offers students abundant opportunities to understand how content and skills are assessed while building the stamina and confidence needed for AP Exam success.

Complete **Answer Keys** for the practice exams are correlated to the AP Topics and Skills and provide feedback and scoring rubrics for teachers. These can be shared with students for self-study and further guidance.

5 Steps to a 5: AP Statistics helps students build AP exam muscle memory all year long through daily engagement with the most up-to-date course content, exam format information, subject-specific strategies, 3 full practice tests, and tips for success. Print and digital bundles are available as an add-on.

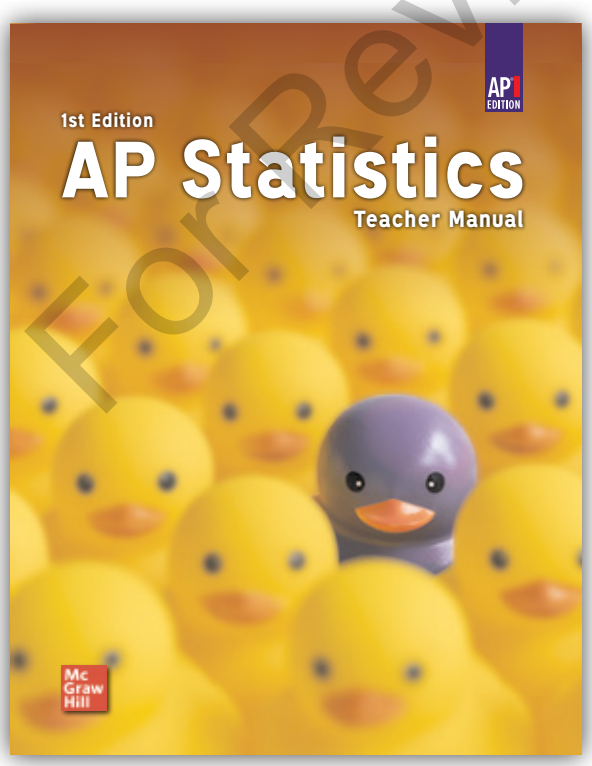


Robust AP Teacher Support

The AP Teacher Manual is designed to support new and experienced teachers with guides, activities, strategies, and answer keys. The manual is available in print, and an online version is included with the digital resources.

Key features include:

- An AP Units and Topics Pacing Guide, Using the Teacher Manual, AP Course and Exam Information, and a Sample Syllabus
- Chapter-specific pacing guides, suggested activities, teaching support/ tips, and answers to questions in the student edition
- Correlations to the AP Topics, Learning Objectives, Essential Knowledge statements, and Practices
- Content focusing on areas where students have historically struggled on the AP exam and guidance for helping them successfully overcome those roadblocks
- Extra support for the AP Skills Practice and Writing in AP Statistics features in the student edition
- Activities for individuals, pairs, and groups to enable students to apply their learning while collaborating and providing peer support
- English Learner support

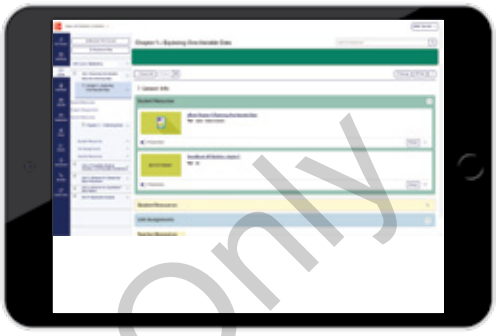


Personalized, Dynamic Digital Resources

Your Online Course on my.mheducation.com

AP Statistics is enriched with digital resources that work seamlessly with the text including videos and interactivities, AP Exam Practice, and adaptive learning tools that provide students with an opportunity to contextualize and apply their understanding.

For Students



More Practice. More Interactivity.

- Interactive eBook and adaptive SmartBook®
- *Kahoot!* games
- Project-Based Learning activities
- Lesson- and chapter-level content review quizzes and tests
- Chapter-level AP Exam practice questions
- Two complete AP practice exams
- Illustrative videos
- Guided Student Notes
- Corequisite Workbook

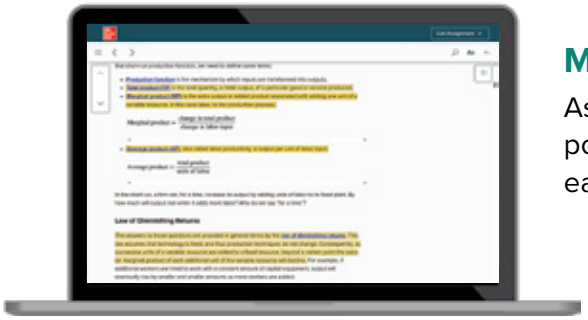
For Teachers

More Support. More Data.

- Teacher Manual
- Benchmarks and pacing
- AP correlations
- Point-of-use instructional strategies
- Accessible, customizable PowerPoint presentations
- English Learner activities
- Chapter-, and lesson-level content quizzes and assessments
- Chapter-level AP Exam-style practice questions
- Auto-graded test banks
- Answer Keys and Scoring Rubrics

Adaptive Learning with SmartBook®

SmartBook® delivers personalized, adaptive learning tailored to each student's individual needs by pinpointing knowledge gaps and focusing instruction on the concepts that require additional study. Teachers can assign a specific chapter, topic, or concept and access advanced reporting features that track individual and class progress with actionable insights to inform in-class instruction.



For Students

More Personalized. More Productive.

As students move through the material, multiple data points are captured to sequence and pace learning for each student's unique needs.

- **Focused Instruction:** Yellow highlights help students easily identify their assigned learning concepts.
- **Targeted Remediation:** Blue highlights bring focus to the content and concepts that require additional study.
- **Meaningful Practice:** Practice sets with instant feedback allow students to ask for guidance and rate their confidence level.
- **Recharged Learning:** Students can recharge their learning by accessing previously completed assignments with personalized recommendations.
- **Mobile Ready:** Assignments are accessible both online and offline with the K-12 Portal app.

For Teachers

More Control. More Actionable.

Teachers can organize assignments to suit their students' needs and align to their course outcomes while easily tracking progress at the individual and/or class level.

- **Flexible Assignments:** Assign homework down to the sub-topic level and time-on-task.
- **Manageable Content:** Assign content across multiple chapters to establish context for the learning ahead and make connections between chapters, topics, and concepts.
- **Results-based Support:** Provide personalized review assignments that target each student's areas of weakness, better preparing them for upcoming assessments.
- **Actionable Reports:** Advanced reporting features track individual and class progress with data-driven insights.

Adaptive Learning with ALEKS

Available with the digital subscription as an add-on, ALEKS AP Statistics uses adaptive questioning to quickly and accurately determine exactly what statistical topics a student knows and doesn't know and instructs each student on the topics they are most ready to learn.

Builds real mastery

ALEKS provides an individual learning pathway for each student that includes both instructional support and supplemental practice opportunities.

Closes gaps and accelerates growth

ALEKS meets students where they are and targets the exact skills they need to grow. Students can progress faster than the pace of the core classroom when they're ready, empowering ownership of learning.

Flexible and accessible

Whether used in class, in tutoring, or at home, ALEKS supports all learners through intuitive design and read-aloud options.

Actionable data and built-in tools

ALEKS includes powerful tools and insights to assess student learning and adapt instruction, including:

- Real-time reports
- Automatic assignment suggestions
- Data-driven tools group students based on learning criteria
- Support for differentiation

Adaptive AP Exam Prep ALEKS is fully aligned to both the structure of *AP Statistics* and the Course and Exam Description, helping students determine their preparedness for the AP Exam.

It creates a personalized learning experience with powerful reporting features that deliver actionable insights for every student.

Power Up with *5 Steps to a 5*

The number one choice for AP teachers, the *5 Steps to a 5* test prep guide is available as an add-on for your AP Statistics course.

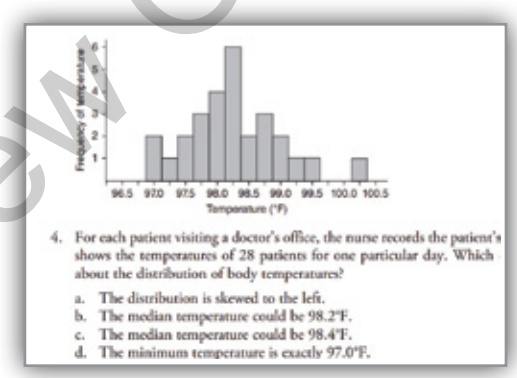
5 Steps to a 5: AP Statistics is a robust, in-class resource that reinforces critical concepts, offers extensive AP Exam practice, and helps students walk into test day feeling prepared and confident.

Each workbook includes:

- Full-length practice tests that align with the latest College Board requirements,
- Hundreds of practice exercises and answer explanations,
- An overview of all the most important AP Statistics topics and skills,
- And proven test-taking strategies from veteran AP Statistics teachers.

Review and Test Prep You Can Trust

5 Steps guides are useful tools throughout the school year, and each begins with a diagnostic test to determine student's strengths and challenges.

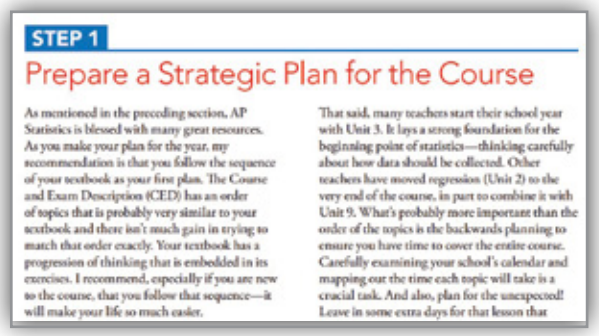


Bellringers for Every Day of the Year

5 Minutes to a 5 is a group of 180 five-minute activities that reinforce the most vital course material and give students the edge they need.

Time-saving Teacher Resources

A teacher manual is included to maximize classroom time, give tips on curriculum management, and build a strategy for the entire school year.



Visit mheducation.com/5Steps for more information.

What Can We Learn from Statistics?

AP Learning Objectives

- Determine an investigative question for a statistical study.
- Identify components within a statistical study.

Key Terms

- statistical study
- population
- sample

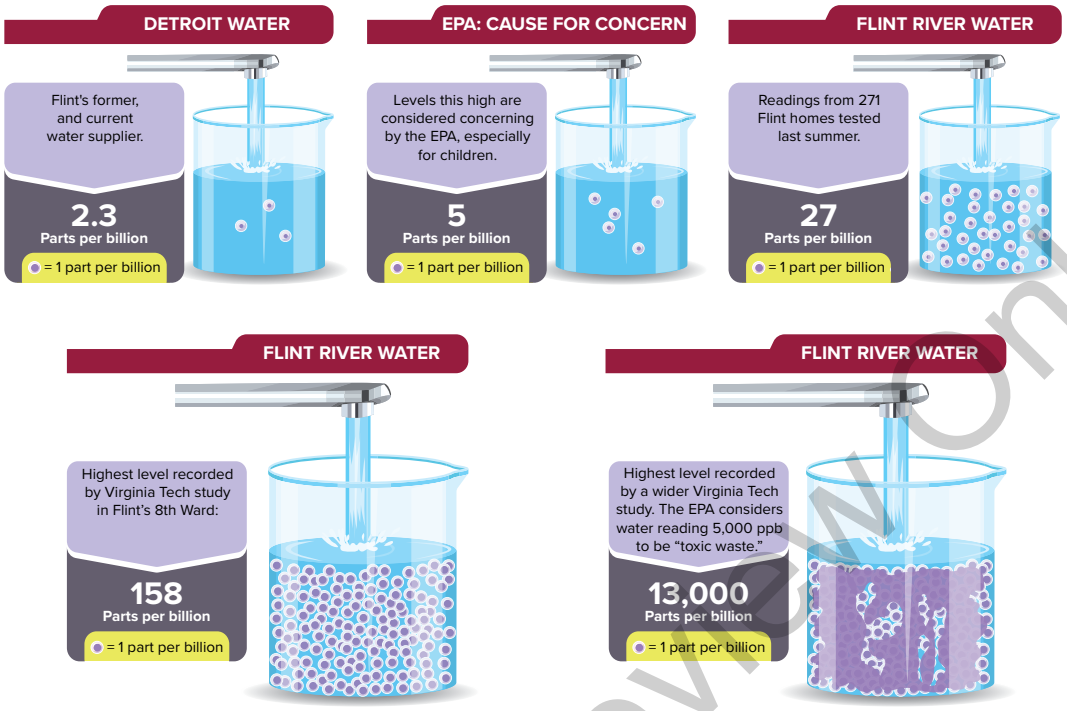
A Water Crisis

Imagine living in a community where something strange is happening. The tap water is off-color and smells funny. Residents begin to lose hair, develop itchy rashes, and feel generally unwell. They report their symptoms to doctors and complain about the water to local officials. Doctors are puzzled, and the water company does not respond. Other residents report similar, and often worse, symptoms and water quality.

This happened in Flint, Michigan, starting in April 2014. That spring, city officials switched from using Detroit's water system to the Flint River. In 2015, residents of Flint contacted a team of researchers, led by Dr. Mark Edwards of Virginia Polytechnic Institute and State University (Virginia Tech), to investigate the questions: "What contaminants are present in the Flint water supply and at what concentrations, and how are these levels statistically associated with the symptoms being reported by residents?" The research team collected water samples from many Flint homes. The researchers mailed out 300 test kits to randomly selected homes and over 90% were returned. A total of 271 test kits were included in the study as "legitimate samples." Dr. Edwards' team found that 101 of the 271 samples contained lead levels greater than the Environmental Protection Agency's "level of concern," which is 5 parts per billion (ppb). One sample even contained 158 ppb! During this time, the local water company continued to claim the water was safe to drink.

ADragan/Shutterstock

If the water really was safe, as the water company continued to claim, what was the probability of finding water samples as contaminated as those collected? The Virginia Tech researchers calculated that, if the lead levels really were safe, the probability of gathering samples similar to those they obtained was 0.00000001. These findings prompted the team to expand their study. In doing so, they uncovered even more troubling results. One sample from the study contained over 13,000 ppb of lead. Anything over 5,000 ppb is considered toxic waste.



Once the researchers had collected and organized the data, they shared it with the residents. The residents of Flint petitioned the EPA, and then successfully sued the city for clean water.

AP Exam Tip

AP Statistics is different from most math courses because it focuses on thinking with data rather than just solving equations. In this course you will learn how to collect data, analyze real-world situations, and make conclusions using probability and statistical reasoning. Instead of only finding "the answer," you must explain what results mean, evaluate studies, and justify decisions using context. This makes AP Statistics highly practical and especially valuable for college majors in science, health, business, social sciences, and research-based fields.

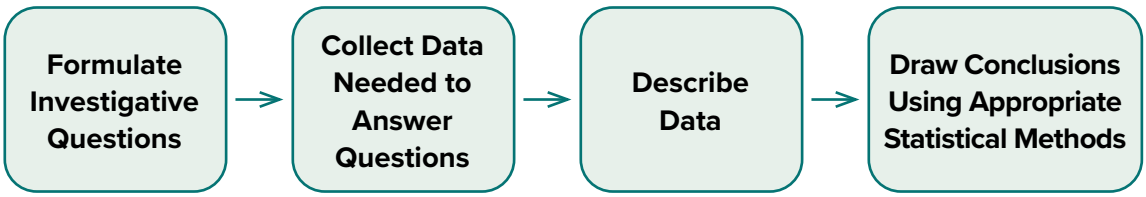
Statistical Investigative Process

The people of Flint followed the steps generally used in a statistical investigative process. These steps are:

1. Formulate investigative questions. Investigative questions for a specific study should have a defined purpose and should not be changed based on the data analysis or results. The questions should be posed so that the data required to answer them can be collected and analyzed using appropriate methods. In the case of the Flint water crisis, the initial questions were: Why are so many people falling ill, and why is there increased exposure to lead, especially among children?
2. Collect data needed to answer the questions. This is most often done through a **statistical study**. A statistical study is a study in which data are collected from a sample to answer the investigative questions about a larger population. The **population** consists of all items or individuals of interest. A **sample** is a subset of the population from which data are obtained. Statistical studies are necessary when the population is too large, or it is too difficult to collect data from every item or individual in the population. In this case, it would have been impossible to test *all* of the water coming from the Flint River for contamination, so the research team used the 300 households as their sample.
3. Describe the data. Once the data were collected, the research team used graphs and summary statistics to describe the data. Understanding how to calculate and interpret statistics armed the researchers and Flint citizens with powerful information that the city officials in court could not refute or ignore.
4. Draw conclusions, using appropriate statistical methods. If the data collected are unlikely to have occurred purely by chance, this provides convincing evidence that something else is causing the results. In the case of the Flint water crisis, the statistics the researchers calculated convinced the EPA and the judge presiding over the lawsuit that the water was unsafe to drink. They alerted the city authorities and the community to immediately stop using the water. Eventually, the water company worked with officials to provide clean water and the water crisis ended.

In the coming chapters, we will explore the processes involved in formulating meaningful questions, determining what type of data to collect and how to collect it, effectively summarizing and describing the data, and drawing sound conclusions about the topics that inspire curiosity in everyday life. We will learn how to answer the question, *"What story can we uncover from these data?"* Understanding the likelihood of events, and the uncertainty that surrounds them, provides a valuable foundation for informed decision-making and enables you to form conclusions with greater confidence and insight.

Statistical Investigative Process



Unit 1 Exploring One-Variable Data and Collecting Data

The following AP Topics are covered in this unit.

Chapter 1 Exploring One-Variable Data

Topic 1.1	Introducing Statistics: What Can We Learn from Data?
Topic 1.2	Variables
Topic 1.3	Tabular Representation and Summary Statistics for One Categorical Variable
Topic 1.4	Graphical Representations for One Categorical Variable
Topic 1.5	Graphical Representations for One Quantitative Variable
Topic 1.6	Descriptions for One Quantitative Variable Distributions
Topic 1.7	Summary Statistics for One Quantitative Variable
Topic 1.8	Graphical Representations of Summary Statistics for One Quantitative Variable
Topic 1.9	Comparisons of the Distributions for One Quantitative Variable

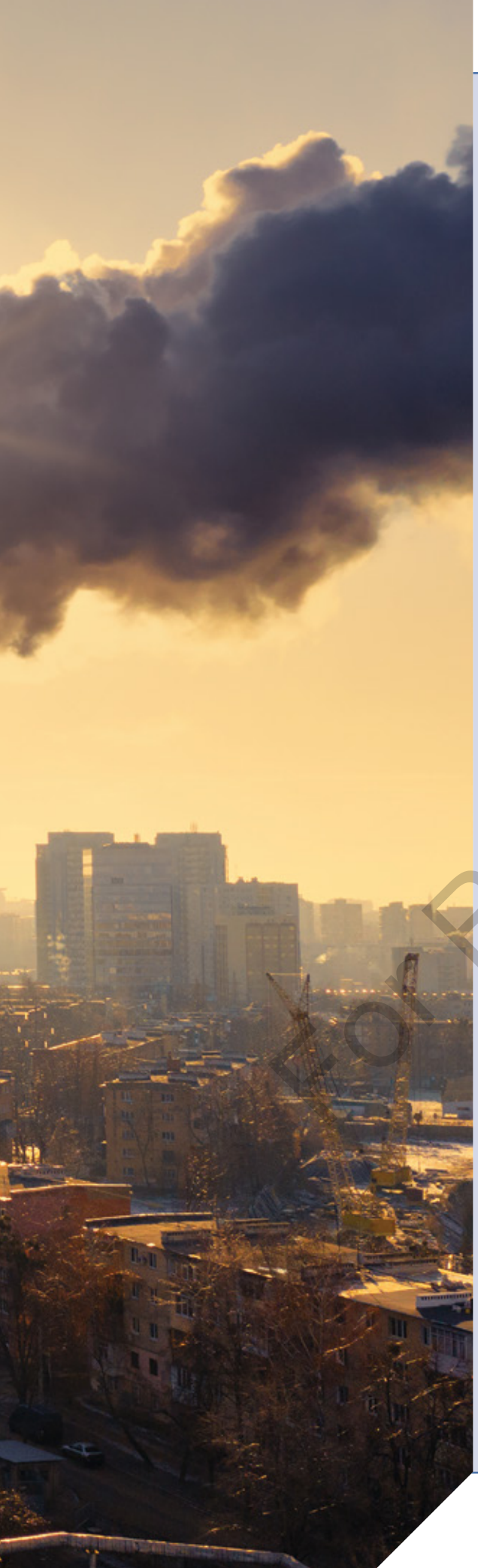
Chapter 2 Collecting Data

Topic 1.10	The Investigative Question Revisited and Data Collection
Topic 1.11	Random Sampling
Topic 1.12	Potential Problems with Sampling
Topic 1.13	Experimental Design

CHAPTER 2

Collecting Data

UrbanWitnessStock/Shutterstock



Case Study: Air Pollution

How does air pollution affect your health? Over the past several decades, scientists have become increasingly convinced that air pollution is a serious health hazard. The World Health Organization estimates that air pollution causes 2.4 million deaths each year. The health effects of air pollution are investigated by measuring air pollution levels and rates of disease, then using statistical methods to determine whether higher levels of pollution lead to higher rates of disease.

Many air pollution studies have been conducted in the United States. For example, the town of Libby, Montana, was the focus of a recent study of the effect of particulate matter—air pollution that consists of microscopic particles—on the respiratory health of children. As part of this study, parents completed a questionnaire about their children's respiratory symptoms. The study revealed the fact that children exposed to higher levels of particulate pollution were more likely to exhibit symptoms of wheezing, as shown in the following table.

Level of Exposure	Percentage with Symptoms
High	8.89%
Low	4.56%

Source: Noonan, C, Ward, T, Navidi, W, and Sheppard, E. (2012) A rural community intervention targeting biomass combustion sources: effects on air quality and reporting of children's respiratory outcomes. Occupational and Environmental Medicine, DOI: 10.1136/oemed-2011-100394

The rate of symptoms is almost twice as high among those exposed to higher levels of pollution. At first, it might seem easy to conclude that higher levels of pollution cause symptoms of wheezing. However, drawing accurate conclusions from information like this is rarely that simple. The case study follow up at the end of this chapter will present more complete information and will show that additional factors must be considered before making any conclusions.

LESSON 2.1

Random Sampling and Potential Problems with Sampling

AP Learning Objectives

- Identify components within a statistical study.
- Identify a census.
- Identify a sampling method given a description of a study.
- Justify the appropriateness of a sampling method.
- Identify potential sources of bias in sampling methods.

Key Terms

- census
- simple random sample (SRS)
- stratified random sample
- cluster random sample
- systematic random sample
- convenience sample
- unbiased
- biased
- bias
- voluntary response bias
- undercoverage bias
- nonresponse bias
- response bias

Why We Select Samples

Why do we select samples? Recall from the Introduction and Chapter One that the entire collection of individuals about which information is sought is called the *population*, and a subset of that population is called a *sample*. A **census** is a study in which we examine every member of the population of interest. Censuses are expensive, time consuming, and often impossible to conduct. So, we select a sample, a manageable-sized group of individuals, from the population to examine.

Ideally, the sample should represent the population as closely as possible. For example, in a political poll we would like the proportions of voters preferring each of the candidates to be the same in the sample as they are in the population. Unfortunately, there are no methods that can guarantee a sample will represent the population well. The best we can do is to use a method that makes it likely that the sample will be similar to the population. The best sampling methods all rely on a random mechanism such as a coin toss, a die roll, or a random number generator. The most basic sampling method is *simple random sampling*.

Simple Random Sampling

In a **simple random sample (SRS)** of size n, each observational unit has an equal chance of being selected and every sample of size n also has the same chance of being selected.

To understand the nature of a simple random sample, think of a lottery. Suppose that 10,000 lottery tickets have been sold, and that five winners are to be chosen. What is the fairest way to choose the winners? The fairest way is to put the 10,000 tickets in a drum, mix them thoroughly, then reach in and draw five tickets out one by one. These five winning tickets are a simple random sample from the population of 10,000 lottery tickets. Each ticket is equally likely to be one of the five tickets drawn. More importantly, each collection of five tickets that can be formed from the population of 10,000 tickets is equally likely to comprise the group of five that is drawn.

Since a simple random sample is analogous to a lottery, it can often be drawn by the same method now used in many lotteries: with a computer random number generator. Suppose there are N items in the population. We number the items 1 through N. Then, we generate a list of unique random integers between 1 and N and choose the corresponding population items to comprise the simple random sample.

Sampling With and Without Replacement

When we sample two items from a population, we can proceed in one of two ways. We can replace the first item drawn before sampling the second; this is known as *sampling with replacement*. When sampling with replacement, it is possible to draw the same item more than once. The other option is to leave the first item out when sampling the second item; this is known as *sampling without replacement*. When sampling without replacement, it is impossible to sample an item more than once. Sampling without replacement is the method used most often in AP Statistics.

Example: Choosing a simple random sample

There are 300 employees in a company. The Human Resources department wants to draw a simple random sample of 20 employees to fill out a questionnaire about their attitudes toward their jobs. Describe how you would use a random number generator to select this sample.

Solution:

Step 1: List all 300 employees and number them from 1 to 300.

Step 2: Use a random number generator to generate 20 unique random numbers between 1 and 300. If a number repeats, ignore it and generate another number.

Step 3: The employees who correspond to these numbers comprise the sample.

AP Exam Tip

You may be asked on the AP Exam to describe in detail how you would choose a simple random sample. Be very clear in your answer.

1. Label each observational unit.
2. Explain in detail how you will randomize. For example, use a random number generator, ignoring any repeated numbers.
3. State that the observational units that correspond to the chosen numbers comprise the sample.

Example: Using a random number generator to select a simple random sample

To draw a simple random sample of five people from the following list of 40 people, a random number generator was used to produce five unique random numbers between 1 and 40. The numbers were: 27, 39, 20, 35, and 17. What does this mean?

1. Dan Aaron	11. Johnny Gaines	21. Jorge Ibarra	31. Edward Shingleton
2. Annie Bienh	12. Carlos Garcia	22. Maurice Jones	32. Michael Speciale
3. Oscar Bolivar	13. Julio Gonzalez	23. Jared Kerns	33. Andrew Steele
4. Dominique Bonnaud	14. Jacqueline Gordon	24. Kevin King	34. Neil Swain
5. Paul Campbell	15. James Graves	25. Frank Lipka	35. Sherry Thomas
6. Jeffrey Carnahan	16. Ronald Harrison	26. Carl Luther	36. Shequia Thompson
7. Joel Chae	17. Andrew Huang	27. Laverne Mitchell	37. Barbara Tilford
8. Dustin Chen	18. Anthony Hunter	28. Zachary Quesada	38. Jermaine Tryon
9. Steven Coleman	19. Jonathan Jackson	29. Donnell Romaine	39. Lizbet Valdez
10. Jared Dersken	20. Sarah Johnson	30. Gary Sanders	40. Katelyn Yu

Solution: Since 27, 39, 20, 35, and 17 were the numbers selected, the person who corresponds to each of those numbers is selected for the sample. Our sample includes Laverne Mitchell, Lizbet Valdez, Sarah Johnson, Sherry Thomas, and Andrew Huang.

CAUTION

Different calculators and statistical software will generate different random numbers, so your sample may not match the one shown in this example. As long as you use a correct procedure, your random sample is valid.

Example: Determining whether a sample is a simple random sample

A physical education professor wants to study the physical fitness levels of students at her university. There are 20,000 students enrolled at the university, and she wants to select a sample of size 100 to take a physical fitness test. She obtains a list of all 20,000 students, numbered from 1 to 20,000. She uses a computer to generate 100 unique random integers between 1 and 20,000 (ignoring repeats). She then invites the 100 students corresponding to those numbers to participate in the study. Is this a simple random sample? Explain why or why not.

Solution: Yes, this is a simple random sample because any group of 100 students out of the 20,000 would have been equally likely to have been chosen.

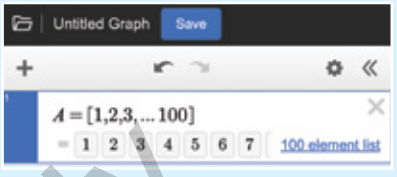
Using Technology

Using a random number generator to sample without replacement

Suppose you want to select a sample of 10 observational units from a population of 100 using a random number generator as the random mechanism. The following steps show how to use Desmos to select 10 integers between 1 and 100 without replacement.

On the Desmos graphing calculator enter the following:

Step 1: Name and enter your list of numbers. "A = [1, 2, 3, ...,100]" where 100 is the size of the population from which we are sampling. You can name your list any letter. Desmos will create a list of 100 integers that starts with 1 and ends with 100. Notice that line 1 indicates that you have created a 100-element list.



Step 2: "Random (A)" will instruct Desmos to randomly select one number from the list that you created and will show the selected number in a box below the command. In this particular case, the random number is 29.

Step 3 (and on): Repeat the command from line 2 until you have 10 unique numbers that represent your sample of 10. Ignore any repeated numbers. In this case the next two numbers are 43 and 44.

Check Your Understanding

1. A physical education professor wants to select a sample of 50 students to fill out a questionnaire about which sports they play. The professor's 10:00 a.m. class has 50 students. He uses the first 20 minutes of class to have all the students in his 10:00 a.m. class fill out the questionnaire. Is this a simple random sample? Explain why or why not.

2. To play the Colorado Lottery Lotto game, you must select six numbers from 1 to 42. Lottery officials then draw a simple random sample of six numbers from 1 to 42. If your six numbers match the ones in the simple random sample, you win the jackpot. Sally plays the lottery and chooses the numbers 1, 2, 3, 4, 5, 6. Her friend George says that this isn't a good choice since it is very unlikely that a random sample will turn out the first six numbers. Is he right?

Stratified Random Sampling

While a simple random sample is the most basic sampling method, other valid methods might be a better choice depending on the situation. In a **stratified random sample**, the population is divided into groups, called strata, where the members of each stratum are similar in some way. A simple random sample is then selected from each stratum. Stratified random sampling is useful when the strata differ from one another in some way that might affect the response. For example, if we wanted to know how voters felt about a topic, we might stratify by age group because we believe that the differing opinions of younger and older voters may affect their responses. Stratifying guarantees that our sample would include sufficiently large numbers of voters from both age groups; whereas in a simple random sample, it would be possible to generate a sample with one age group substantially underrepresented.

Example: Selecting a stratified random sample

A company has 1,000 employees, of whom 800 are full-time and 200 are part-time. The company wants to survey 50 employees about their opinions regarding benefits. Attitudes toward benefits may differ considerably between full-time and part-time employees. Why might it be a good idea to use a stratified sample? Describe how one might be selected.

Solution: If a simple random sample is selected from the entire population of 1,000 employees, it is possible that the sample will contain only a few part-time employees, and their attitudes will not be well represented. For this reason, it might be advantageous to use a stratified sample. To select a stratified sample, one would use two strata. One stratum would consist of full-time employees, and the other would consist of part-time employees. Since 80 percent of the employees are full-time and 20 percent are part-time, we might choose to select a simple random sample of 40 employees from the full-time stratum and a simple random sample of 10 from the part-time stratum. This allocation guarantees that both groups are represented in the sample at the same rates as the population. In some cases, it is desirable to sample a larger proportion from a smaller stratum, to ensure that reliable conclusions can be drawn about that stratum.

Cluster Random Sampling

A **cluster random sample** involves the division of a population into smaller groups called clusters. Ideally, each cluster mirrors the diversity of the population, with all clusters being similar to one another. A simple random sample of clusters is selected from the population. Data is collected from all observational units in each selected cluster.

Cluster sampling is useful when the population is too large and scattered for simple random sampling to be feasible. Suppose you draw a simple random sample of third period classes at your school and interviewing every member of each selected class. This would be a cluster sample, with the classes as the clusters. U.S. government agencies use cluster sampling extensively in sampling the U.S. population to measure sociological factors such as income and unemployment.

Example: Selecting a cluster sample

To estimate the unemployment rate in a county, a government agency draws a simple random sample of households in the county. Someone visits each household and asks how many adults live in the household and how many of them are unemployed. What are the clusters? Why is this a cluster sample?

Solution: The clusters are the households in the county. This is a cluster sample because a simple random sample of households is selected, and every adult in each selected cluster is part of the sample.

Take Another Look

In both cluster sampling and stratified sampling, the population is divided into groups. In stratified sampling, a simple random sample is chosen from each group. In cluster sampling, a random sample of groups is chosen and every member of the chosen groups is included in the sample.

Systematic Random Sampling

A **systematic random sample** is a method in which sample members from a population are selected according to a random starting point and a fixed, periodic interval between successive sampling units. Think of walking alongside a line of people and randomly choosing one person to start the sample, and then choosing every third person after that one. That would produce a systematic random sample.

In a systematic sample, the population items are arranged in an orderly manner. The frequency of sampling is then decided; for example, one could sample every third item, or every fifth item, or every hundredth item. Let k represent the sampling frequency. To begin the sampling, choose a starting place at random. Select the item in the starting place, along with every kth item after that.

In industry, systematic sampling is sometimes used to sample products as they come off an assembly line to check that they meet quality standards.

Example: Describe a systematic sample

Automobiles are coming off an assembly line. The safety supervisor has decided to draw a systematic sample of the cars for a detailed check of the steering system. The starting point, determined by a random number generator, will be the third car. Every fifth car after that will be sampled. Which cars will be sampled?

Solution: Start with the third car, then count by five to determine which cars will be sampled. The sample will consist of cars numbered 3, 8, 13, 18, and so on.

Figures 2.1–2.4 illustrate four valid methods of random sampling.

Figure 2.1 Simple Random Sample

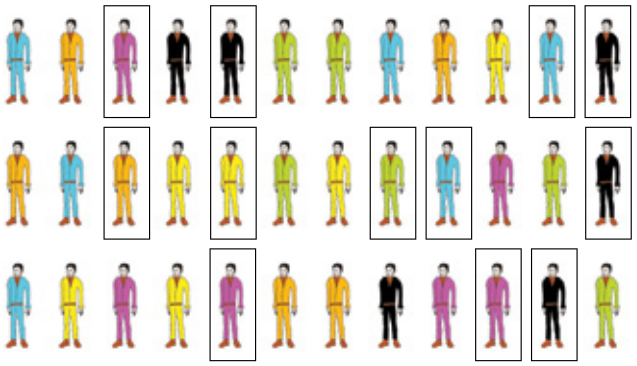


Figure 2.2 Stratified Sample

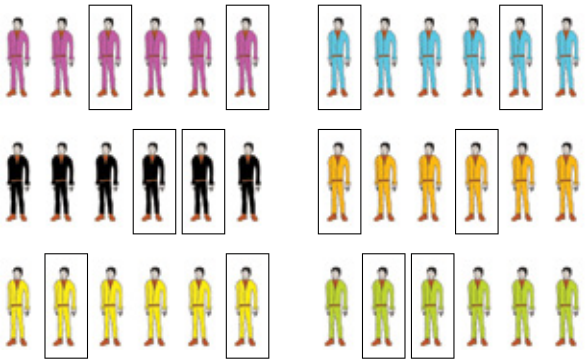


Figure 2.3 Cluster Sample

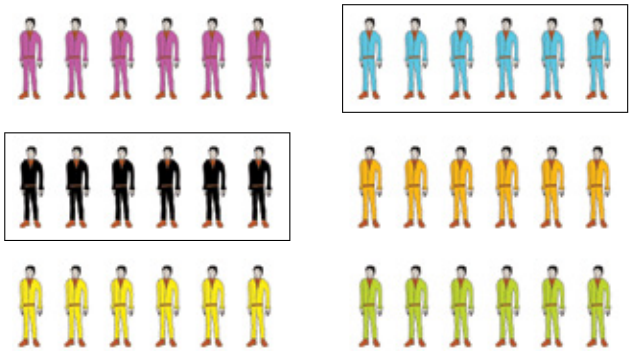
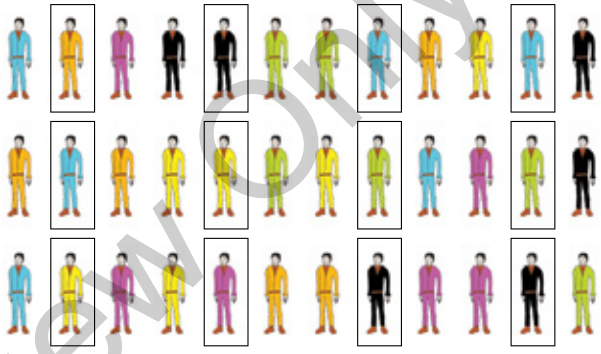


Figure 2.4 Systematic Sample



Check Your Understanding

Determine the sampling method in each of the given studies.

1. Every 10 years, the U.S. Census Bureau attempts to count every person living in the United States. To check the accuracy of their count in a certain city, they draw a sample of census districts (roughly equivalent to a city block) and recount everyone in the sampled districts. What kind of sample is formed by the people who are recounted? Explain.
2. A public health researcher is designing a study of the effect of diet on heart disease. The researcher knows that the diets of males and females tend to differ and that males are more susceptible to heart disease. To be sure that both males and females are well represented, the study comprises a simple random sample of 100 males and another simple random sample of 100 females. What kind of sample do these 200 people represent? Explain.
3. A college basketball team held a promotion at one of its games in which every 20th person who entered the arena won a free basketball. What kind of sample do the winners represent? Explain.
4. To select people to call for jury duty, a numbered list of all residents with driver's licenses or ID cards is made. Then random numbers are generated and the people corresponding to those numbers are selected. What kind of sample is this? Explain.

Convenience Sampling

In some cases, it is difficult or impossible to select a sample in a truly random way. In these cases, you can sample items by a convenient method. A sample obtained in such a way is called a **convenience sample**. Convenience samples are easier to select but may not represent the population as well as a random sample would and should be avoided if possible.

Example: Selecting a convenience sample

A construction engineer has just received a shipment of 1,000 concrete blocks, each weighing approximately 50 pounds. The blocks have been delivered on a large pallet. The engineer wants to investigate the crushing strength of the blocks by measuring the strengths in a sample of 10 blocks. Explain why it might be difficult to select a simple random sample of blocks. Describe how the engineer might select a convenience sample.

Solution: Selecting a simple random sample would require removing blocks from the center and bottom of the pallet, which would be quite difficult. One way to draw a convenience sample would be to simply take 10 blocks off the top of the pallet.

The most significant problem with convenience samples is that they may differ systematically in some way from the population. For this reason, convenience samples should not be used except in situations where it is not feasible to draw a random sample. If a convenience sample is selected, it is important to think carefully about all the ways in which the sample might differ systematically from the population. In the previous example, it is possible that blocks in different parts of the pile may have been made from different batches of mix or may have different curing times or temperatures affecting the strength of the blocks in those batches. A convenience sample could give misleading results.

CAUTION

A sample is not randomly selected when observational units are deliberately chosen or volunteer themselves to be in the sample. Do not use a convenience sample when it is possible to select a random sample.

Bias in Sampling

No study is perfect, and even a properly conducted study will generally not give results that are completely accurate. For example, suppose you want to estimate the number of students at a college who are not affiliated with a political party.

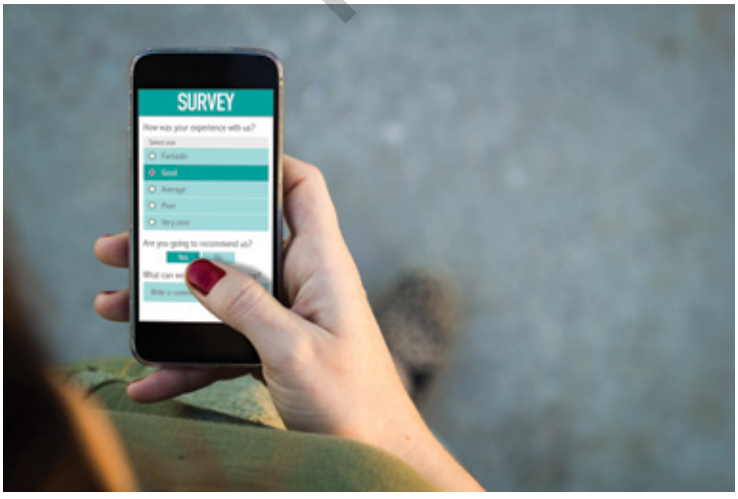
You might draw a simple random sample of students. However, just by chance, your sample probably contained a different percentage of Independent voters than the entire population of students. If you drew many simple random samples, some would have a greater percentage of Independents than in the population, and some would have a smaller percentage. But on average, the percentage of Independents in a simple random sample will be the same as the percentage in the population. A study conducted by a procedure that produces the correct result on average is said to be **unbiased**.

Now suppose that you estimated the percentage of Independents in the population by selecting a sample of students who attended a speech made by an Independent politician. On average, studies conducted in this way would overestimate the percentage of Independents in the population. Studies conducted with methods that tend to overestimate or underestimate a population value are said to be **biased**.

Bias in a sampling method is a systematic error in the sampling procedure that results in a statistic being consistently larger or consistently smaller than the parameter the statistic is used to estimate. In practice, it is important to design studies to have as little bias as possible. Unfortunately, some studies are highly biased, and the conclusions drawn from them are not reliable. Here are some common types of sampling method bias.

Voluntary Response Bias

Voluntary response bias is the result of a voluntary response survey. A voluntary response survey is one in which people are invited to visit a website, send a text message, participate in an online poll, or in some other way volunteer to express their opinions on an issue. In many cases, the opinions of the people who choose to participate in such surveys do not reflect those of the population as a whole. People with strong opinions are more likely to participate. In general, voluntary response surveys are highly biased.



Undercoverage Bias

Undercoverage bias occurs when some members of the population are more likely to be included in the sample than others, or when an entire group of the population is left out completely. For example, a telephone poll that only surveys landline phone numbers will systematically leave out those who only have mobile phones or no phones. Because older people are more likely to have landlines, they may be overrepresented in the survey, while individuals without landline telephones will be undercovered groups. It is almost impossible to avoid undercoverage bias completely, but modern survey organizations work hard to keep it at a minimum.

Example: Identifying bias

A researcher wants to estimate local residents' opinions about a proposal on an upcoming ballot. The researcher obtains a list of all landline phone numbers in the voting area and randomly selects phone numbers from that list. What is the source of bias and why may it be problematic for the researcher?

Solution: The source is undercoverage bias. Polling in this way will exclude individuals who use only mobile phones or have no phones (many younger people, or lower-income households) and overrepresent older demographics. This could lead to the opinions in the poll not being reflective of the true population.

Nonresponse Bias

People cannot be forced to answer questions or to participate in a study. In any study, a certain proportion of people who are selected to participate will refuse to do so or cannot be reached. This results in **nonresponse bias**. In many cases, the opinions of the nonresponders tend to differ from the opinions of those who do respond. As a result, surveys with many nonresponders are often biased.

Response Bias

Any pattern of inaccurate answers in a study is called **response bias**. This bias can be introduced in several ways. Here are four common types of response bias:

- *Wording the question.* Confusing, leading, or emotionally charged wording can nudge respondents toward answering in a particular way. For example, the question, "Do you agree school lunches are unhealthy and gross?" includes emotionally charged and leading words.
- *Interviewer effect.* Characteristics or behaviors of the interviewer can influence answers. For example, if a school resource officer interviews 12th grade students on their driving habits in the last 30 days, the answers to the survey may not be accurate or truthful.
- *Social desirability.* People often give answers they think the researcher wants to hear, rather than being honest. For example, when asked, people may overreport hours spent volunteering because they believe that saying they volunteer more will make others think more highly of them.
- *Inaccurate self-reporting.* People are forgetful and inaccurate guessers. If you ask a person how many hours they sleep each night on average and compare results to a sleep tracker, the results will likely vary wildly.

It is important to note that many studies exhibit more than one type of bias. For example: Parents of school-aged children were asked, “Do you support improving public schools by issuing school bonds?” Only 30 percent of those asked responded, and 95 percent of those responding supported improving public schools through school bonds. This sample exhibits nonresponse bias because only 30 percent of those sampled responded to the question. Response bias is present because parents of school children want to appear supportive of improving public schools even if they don’t want to pay taxes toward a school bond. Undercoverage bias also exists here because people who are not parents of school-aged children are unrepresented in the sample.

Example: Types of bias

Specify the types and possible direction of bias involved in each of the following situations.

a. A podcast host invites listeners to go to a website and participate in a poll about which candidate they prefer in an upcoming election. Ninety-seven percent of the listeners who participate in the poll support Candidate A.

b. A health insurance provider mailed a survey to each of its customers, asking whether they were satisfied with their coverage. Ten percent of those who received the survey responded. Eighty-four percent of the customers who responded said they were very unhappy with their coverage.

c. A student government representative polled students during the 12th grade lunch to ask how spirited they believe the student section to be at varsity football games. The results showed that most students feel that the student section at the varsity football games is extremely spirited.

d. A woman with three adorable golden retriever puppies asks people walking in an outdoor mall their opinions on animal testing in scientific studies. According to her results, 95 percent of people are against animal testing in scientific studies.

Solution:

a. Voluntary response bias. Only those who listen to the podcast could respond, and individuals are not asked directly for their opinion; instead,

they are invited to respond if they wish. Those who do respond likely have stronger opinions about the election than the general population. Thus, 97 percent is likely an overestimate of the approval rating for Candidate A.

b. Nonresponse bias. Only a small percentage of those asked for their opinion responded. Those who did respond seem to be unhappy with their coverage, and therefore the results are likely an overestimate of the actual proportion of all of the customers who are unhappy with their coverage.

c. Undercoverage bias. Those in grades 9, 10, and 11 are not represented in the sample, and the opinions of students in grade 12 are overrepresented in the sample. The results of this survey are likely to overestimate the level of spiritedness that all of the students feel is present at varsity football games.

d. Response bias. The presence of the puppies will likely sway responses to be against animal testing, thus resulting in an overestimate of people who are against it.

Check Your Understanding

1. A television newscaster asks viewers to use a mobile app to state whether they approve of the president's economic plan. Results from the survey show that the president's economic plan has an 88 percent approval rating. What type of bias is present here? Are these results likely an overestimate or underestimate?

2. A polling organization sent text messages to 1,000 people in a certain city to ask them whether they favor a tax increase to build a new school. Two hundred people replied, and 150 of them opposed the tax. What type of bias is present here.

A Big Sample Size Does Not Make Up for Bias

A sample is useful only if it is drawn by a method likely to represent the population well. If you use a biased method to draw a sample, then drawing a big sample doesn't help; a large, nonrepresentative sample does not describe a population any better than a small, nonrepresentative sample. Voluntary response surveys often draw several hundred thousand people to participate. Although the sample is large, it is unlikely to represent the population accurately, so the results are meaningless.

Check Your Understanding

Eighty thousand people attending a professional football game filled out surveys asking their opinions on using tax money to upgrade the football stadium. Seventy percent said that they supported the use of tax money. A pollster then surveyed a simple random sample of 500 voters, and only 30 percent of the voters in this sample supported the use of tax money. The owner of the football team claims that the survey conducted at the football stadium is more reliable because the sample size was much larger. Is he right? Explain.

Lesson 2.1 Summary

AP Key AP Takeaways

- A census is a complete count of every unit in the population. Censuses tend to be time consuming, expensive, and often impossible to conduct.
- Most populations are too large to study each member, so samples are drawn and studied.
- Using a random mechanism to select the sample helps to ensure that the sample represents the population well.
- Simple random sampling, stratified random sampling, cluster random sampling, and systematic random sampling are all valid sampling methods.
- Avoid convenience sampling if possible.
- A study is biased if it is conducted by a method that tends to produce an incorrect result.
- Some of the most common forms of bias are voluntary response bias, undercoverage bias, nonresponse bias, and response bias.

Lesson 2.1 Practice

Understanding the Concepts

In exercises 1–5, determine the appropriate word or phrase.

1. What is the entire collection of individuals about which information is sought called?

2. A subset of a population is called what?

3. A sample that is not drawn by a well-defined random method is called what?

4. What type of sample is one in which the population is divided into groups and a random sample of groups is drawn?

5. What type of sample is one in which the population is divided into groups and a random sample is drawn from each group?

In exercises 6–8, determine whether the statement is true or false. If the statement is false, rewrite it as a true statement.

6. In a cluster sample, the population is divided into groups, and a random sample from each group is drawn.

7. Both stratified sampling and cluster sampling divide the population into groups.

8. One reason that voluntary response sampling is unreliable is that people with stronger views tend to express them more readily.

For exercises 9–11, identify which is taken: a census or a sample.

9. A small town asks every household to report its water usage for the year to calculate total community consumption.

10. A zoo selects a subset of its visitors to give feedback on its newest exhibit.

11. A company surveys all its employees to decide whether to switch to a four-day work week.

In exercises 12–13, identify the kind of sample described.

12. Parking on campus: A college faculty consists of 400 men and 250 women. The college administration wants to draw a sample of 65 faculty members to ask their opinion about a new parking fee. They draw a simple random sample of 40 men and another simple random sample of 25 women.

13. Rate your online shopping: An online store gathers feedback on users' shopping experience by having a survey pop up on their mobile device or computer after they check out.

In exercises 14–19, specify the type of bias involved and the direction of the bias if possible.

14. A bank sent out questionnaires to a simple random sample of 500 customers asking whether they would like the bank to extend its hours. Eighty percent of those returning the questionnaire said they would like the bank to extend its hours. Of the 500 questionnaires, 20 were returned.

15. To determine his constituents' feelings about election reform, a politician sends a survey to people who have subscribed to his newsletter. More than 1,000 responses are received, and 92% of those agree strongly that election reform needs to happen.

16. A school principal wanted to know what percentage of her students have cheated on a test in the last month. She polled students and nearly all reported never having cheated on a test.

17. A media personality asks viewers to post their opinions on social media about whether the U.S. Congress is doing a good job in handling the economy. More than 100,000 people share their views online, and only 25% feel that Congress is doing a good job.

18. Students in one statistics class were asked if they preferred to take tests in class or remotely. Ten percent said they preferred to take their tests remotely.

19. One hundred people in a park on a Saturday afternoon were asked how much time they typically spent outdoors during a weekend. Of the responders, 85% reported that they spend more than 8 hours outdoors during a weekend.

Working with the Concepts

20. You are giving me a headache: A pharmaceutical company wants to test a new drug that is designed to provide superior relief from headaches. They want to select a sample of headache sufferers to try the drug. Do you think that it is feasible to draw a simple random sample of headache sufferers, or will it be necessary to use a convenience sample? Explain your reasoning.

21. Pay more for recreation? The director of the recreation center at a large university wants to sample 100 students to ask whether they would support an increase in their recreation fees to extend the hours that the center is open. Do you think it is feasible to draw a simple random sample of students, or will it be necessary to use a convenience sample? Explain your reasoning.

22. Nuclear power, anyone? In a survey conducted by representatives of the nuclear power industry, people were asked the question, "Do you favor the construction of nuclear power plants to reduce our dependence on foreign oil?'' A group opposed to the use of nuclear power conducted another survey with the question, "Do you favor the construction of nuclear power plants that can kill thousands of people in an accident?''

a. Do you think that the percentage of people favoring the construction of nuclear power plants would be about the same in both surveys? Why or why not?

b. Would either of the two surveys produce reliable results? Explain.

23. Who's calling, please? Random-digit dialing is a sampling method in which a computer generates phone numbers at random to call. Do you think that caller ID increases the bias in random digit dialing? Explain.

24. Higher taxes, anyone? A polling firm wants to estimate the proportion of people in a town who support an increase in sales tax to support public schools. They approach people in an amusement park to ask their opinion. Do you think this sampling method produces bias? Explain.

25. Military spending: A pollster wants to determine the level of support for military spending. The pollster asks 100 people watching a Memorial Day parade whether they support increased funding for the military and 80% say yes. Is this result representative of the population, or is it likely due to bias? Explain.

26. Stay on the line, please: When calling a doctor's office, callers are asked to stay on the line after the call is over to answer a brief survey. Twenty percent of callers stay on the line. Of these, 75% say they were satisfied with the call. Is this result reliable, or is it likely that it was due to bias? Explain.

Exercises 27–30 refer to the population of animals in the following list. The population is divided into four groups: mammals, birds, reptiles, and fish.

Mammals		**Birds**	
1. Aardvark	6. Lion	11. Flamingo	16. Hawk
2. Buffalo	7. Zebra	12. Swan	17. Owl
3. Elephant	8. Pig	13. Sparrow	18. Chicken
4. Squirrel	9. Dog	14. Parrot	19. Duck
5. Rabbit	10. Horse	15. Pelican	20. Turkey
Reptiles		**Fish**	
21. Gecko	26. Python	31. Catfish	36. Shark
22. Iguana	27. Turtle	32. Tuna	37. Trout
23. Chameleon	28. Tortoise	33. Cod	38. Perch
24. Rattlesnake	29. Alligator	34. Salmon	39. Guppy
25. Boa constrictor	30. Crocodile	35. Goldfish	40. Minnow

27. **Simple random sample:** Your friend generated a simple random sample of eight animals from the list of 40 animals in the list. The friend claims all eight randomly selected animals were fish. Is it possible that your friend is telling the truth?

28. **Another sample:** Draw a sample of eight animals by drawing a simple random sample of two animals from each group. What kind of sample is this? Explain.

29. **Another sample:** Draw a simple random sample of two groups of animals from the four groups and construct a sample of 20 animals by including all the animals in the sampled groups. What kind of sample is this? Explain.

30. **Another sample:** Choose a random number between 1 and 5. Include the animal with that number in your sample, along with every fifth animal thereafter, to construct a sample of eight animals. What kind of sample is this? Explain.

Extending the Concepts

31. **Draw a sample:** Suppose that you are asked to determine students' opinions at your school about a potential change in library hours. Describe how you could select a sample of each of the following types: simple random sample, stratified sample, cluster sample, systematic sample.

32. ***Literary Digest* poll:** In the 1936 presidential election, Republican candidate Alfred Landon challenged President Franklin Roosevelt. The *Literary Digest* magazine conducted a poll in which they mailed questionnaires to more than 10 million voters. The people who received the questionnaires were drawn from lists of automobile owners and people with telephones. The magazine received 2.3 million responses and predicted that Landon would win the election in a landslide with 57% of the vote. In fact, Roosevelt won in a landslide with 62% of the vote.

    a. In 1936, most people did not own automobiles and many did not have telephones. Explain how this could have caused the results of the poll to be mistaken.

    b. What can be said about the response rate? Explain how this could have caused the results of the poll to be inaccurate.

    c. The *Literary Digest* believed that its poll would be accurate because it received 2.3 million responses, which is a very large number. Explain how the poll could be wrong, even with such a large sample.

LESSON 2.2

Investigative Questions and Statistical Studies

AP Learning Objectives

- Determine the components of an investigative question within a statistical study.
- Identify an experiment.
- Identify an observational study.
- Justify the appropriateness of generalizations for a statistical study.
- Justify the appropriateness of conclusions based on a well-designed experiment.

Key Terms

- response variable
- explanatory variable
- experiment
- experimental units
- treatment
- levels
- observational study
- prospective study
- retrospective study
- sample survey
- confounding variable

Investigative Questions

Recall that a statistical study is one in which data are collected from a sample to answer an investigative question about a larger population. In practice, a study will often have several investigative questions. For example, in a study to compare a new medical treatment to a currently used treatment, one question is likely to be "Is the new treatment more likely to cure the disease than the current treatment?" Other questions might include "Is the new treatment less likely to produce side effects?" and "Does the new treatment cure the disease more quickly?" When designing a statistical study, the most important and earliest step is to develop the investigative questions. What do we want to know? What important questions do we want to answer? The investigative questions impact every decision in the statistical process. They influence how the researcher will collect the data, what kind will be collected, how to organize and analyze the data, and what conclusions can be made from the data.

A well-developed investigative question has three components:

1. The question should clearly identify the variables we are interested in studying. There are often two variables, the explanatory variable and the response variable. The **response variable** measures the outcome of the statistical study. The **explanatory variable** helps explain or predict how the response variable changes.
2. The question should define the parameter of interest and clearly identify how the data will be analyzed. Are we trying to *estimate* the value of the parameter of interest? Or are we trying to *find evidence* to conclude the parameter is greater than, less than, or just different than what has been claimed?
3. The question should specify the population to which conclusions will apply, and the types of conclusions that can be obtained.

Example: Identifying the Explanatory and Response Variables

Identify the explanatory and response variables in each of the following investigative questions.

1. For AP Statistics students at Central High School, does attending weekly AP Statistics tutoring sessions result in a higher mean unit test score?

2. Among all junior and senior students at Central High School, is the proportion of seniors who plan to take the AP Statistics exam greater than the proportion of juniors who plan to take the exam?

Solution:

1. The explanatory variable is whether a student at Central High School attends weekly AP Statistics tutoring sessions. The response variable is their mean unit test score.
2. The explanatory variable is a Central High School student's grade level. The response variable is whether the student takes the AP Statistics Exam.

Statistical Studies

Once we have established an investigative question, we conduct a *statistical study* to answer that question. Consider how scientists might address the investigative question, "Does a new organically-engineered seed produce a larger wheat yield than the two most common types of seed currently being used in the United States?" The following steps describe how a study might be conducted to determine which of three types of seed will result in the largest wheat yield.

1. Prepare three identically sized plots of land, with similar soil types.
2. Plant each type of seed on a different plot, choosing the plots at random.
3. Water and fertilize the plots in the same way.
4. Harvest the wheat, and measure the amount grown on each plot.
5. If one type of seed produces substantially more (or less) wheat than the others, then scientists will conclude that it is better (or worse) than the others.

In general, studies fall into two categories: *experiments* and *observational studies*.

Experiments

An **experiment** is a statistical study in which a treatment is assigned to each individual in the study to observe the response. The individuals (people, animals, plants, rocks, etc.) to whom treatments are assigned are called **experimental units**. For example, the wheat study is an experiment. The treatments are the three types of seeds and the experimental units are the three plots of land. The **treatments** may also be considered as values of an explanatory variable. Thus in the wheat study, the explanatory variable is the seed type, which has three possible values. Explanatory variables are sometimes called *factors*. The choices for a given factor are called the **levels** of the factor. In the wheat study, the factor is the seed, and the levels of the factor are the three types of seed.

When possible, treatments should be chosen at random. An experiment in which treatments are assigned to the experimental units at random is called a *randomized experiment*.

Example: Identifying an Experiment

A pet food company is interested in knowing if a new formulation of goldfish food will cause goldfish to grow faster. They conduct an experiment in which they randomly select 10 goldfish of similar age, weight, and size from a local pet store. Each fish is placed in an identical fishbowl. Five will receive the new formulation of goldfish food, and the other five will receive the old formulation. They will be fed the same amounts at the same time of day for 30 days. At the end of 30 days, the researchers will measure the weights of the goldfish and compare them across the treatment groups.

a. Explain why this study is an experiment.

b. Identify the experimental units, factors, levels, and treatments.

Solution:

a. This statistical study is an experiment because treatments are being assigned to the experimental units.

b. The experimental units are the ten goldfish. The factor is food formulation. There are two levels – new formulation and old formulation. This leads to two treatments; five fish receive the new formulation of food, and five fish receive the old formulation.

Check Your Understanding

1. A study involving 120 patients was done to determine whether increasing the level of C4A proteins in the blood speeds recovery from cardiopulmonary bypass surgery. Sixty patients were randomly assigned to receive plasma containing C4A proteins, while the remaining 60 were assigned to receive plasma that did not contain C4A proteins. The length of hospital stay was recorded for each patient. Those who received C4A proteins had an average stay of 9 days, while those who did not had an average stay of 12 days.

a. Why is this study considered an experiment?

b. What are the experimental units? What are the treatments?

c. What is the response variable?

2. The Aspirin Myocardial Infarction Study involved 4,524 patients who had experienced a heart attack. Of these, 2,267 were chosen at random to receive 1 g of aspirin each day, while the remaining 2,257 were given a placebo. During the next three years, 14.1% of those receiving aspirin and 14.8% of those receiving a placebo experienced another heart attack.

a. Is this study an experiment? Explain.

b. What are the experimental units? What are the treatments?

c. What is the response variable?

Observational Studies

An **observational study** is one in which the investigators do not assign treatments to the observational units. In an observational study, we simply observe what has already happened, what is currently happening, and predict what might happen in the future. There are three types of observational studies that we will look at in this course: prospective studies, retrospective studies, and sample surveys.

Prospective studies are observational studies that collect data in the present and look into the future to make predictions following trend data. **Retrospective studies** look at data that have already been collected. A **sample survey** is a particular kind of study done on human subjects (or participants) in which they are asked a set of questions. Sometimes this is referred to as polling.

Observational studies tend to be less reliable than experiments. To see why, consider a study that is intended to determine whether smoking increases the risk of heart attack. A group of smokers and a group of non-smokers volunteer to participate in a study in which they are observed for several years. During that time, a higher percentage of the smoking group experiences a heart attack. Does this *prove* that smoking increases the risk of heart attack? No. The problem is that the smoking group will differ from the nonsmoking group in many ways other than smoking, and these other differences may be responsible for differences in the rate of heart attacks.

For example, smoking is more prevalent among men than among women. Therefore, the smoking group will contain a higher percentage of men than the nonsmoking group. It is known that men have a higher risk of heart attack than women. So, the higher rate of heart attacks in the smoking group could be due to more men in the smoking group, and not to the smoking itself.

Writing in AP Statistics

What are you writing about? On the AP Exam, you may be asked to **identify** or **classify** a given study as an observational study or experiment and to **justify** your answer. To receive full credit for this type of question, you must include three things:

(1) Explicitly tell whether the study is an experiment or an observational study.

(2) Provide a justification for why the study is either an experiment or an observational study. The justification will always be due to the absence or presence of a treatment being assigned.

(3) Include context. All written responses on the AP Exam should include the context of the problem. Refer to the stem of the question for help phrasing difficult context.

Sample: A regional dance competition assigns points in four different styles of dance: Jazz, Hip-Hop, Lyrical, and Tap. Researchers are interested in learning which style of dance earns the highest point totals. Higher points are considered better technical dance performances. The researchers conduct a study in which they contact several studio directors and ask them to report their competition scores for the various styles of dance from the previous season. **Identify** whether this is an observational study or an experiment. **Justify** your answer in context.

Student Example #1:

This is an observational study. The researchers had studio directors report their dance scores for the various styles of dance from the previous season. The directors were not randomly assigned a dance style so no treatment was assigned.

Student Example #2:

There was no random assignment of treatments.

The **first student** gave all three things needed. The student identified this as an observational study, provided a justification based on no treatment being assigned, and specified what that treatment might be. Finally, the student included appropriate context (studio directors, dance scores, dance styles). Notice that this student just repeated what the prompt said the researchers had the studio directors do.

The **second student** only told us that there was no random assignment. This may provide a justification based on no treatment being assigned, but it does not answer whether this is an observational study or experiment. There also is no context included in student example #2.

Practice the writing: A principal wants to know whether attendance is related to grade point average. Records of a random sample of students from the past year are collected and analyzed. **Classify** this as an observational study or an experiment. **Justify** your answer in context.

Confounding Variables

The smoking study scenario illustrates the major problem with observational studies. It is difficult to tell whether a difference in the outcome is due to the explanatory variable (smoking) or to some other difference between the two groups of people. This is known as confounding. In the preceding example, sex was a *confounding variable*. A **confounding variable** is a variable that is related to both the explanatory and the response variables. When a confounding variable is present, it is difficult to determine whether differences in the response are due to the explanatory variable or to the confounding variable. The sex of a person is related both to smoking (men are more likely to smoke) and to heart attacks (men are more likely to have heart attacks). For this reason, it is difficult to determine whether the difference in heart attack rates is due to differences in smoking (the treatment) or differences in sex (the confounding variable).

Take Another Look

Another way to describe a confounding variable: A confounding variable is a variable related to the explanatory variable that can cause the response variable to have different outcomes.

How can we prevent confounding variables? One way is to design a study so that the confounder is not a factor. For example, to determine whether smoking increases the risk of heart attack, we could compare a group of male smokers to a group of male nonsmokers, and a group of female smokers to a group of female nonsmokers. A person's sex would not be a confounder here because there would be no differences in sex between the smoking and nonsmoking groups. Of course, there are other possible confounders. Smoking rates and rates of heart attacks vary among ethnic groups. If people in ethnic groups that are more susceptible to heart attacks are also more likely to smoke, then ethnicity becomes a confounder. This can be dealt with by comparing smokers of the same sex and ethnic group to nonsmokers of that sex and ethnic group.

Designing observational studies that are free of confounding variables is difficult. In practice, many studies must be conducted over a long period of time. In the case of smoking, this has been done, and we can be confident that smoking does indeed increase the risk of heart attack, along with other diseases. If you don't smoke, you have a much better chance of living a longer and healthier life.

Example: Identifying a confounding variable

In a study of the effects of blood pressure on health, a large group of people of all ages was given regular blood pressure checkups for a period of one year. It was found that people with high blood pressure were more likely to develop cancer than people with lower blood pressure. What is a possible confounding variable in this study?

Solution: Age is a likely confounding variable. Older people tend to have higher blood pressure than younger people, and older people are more likely to get cancer than younger people. Therefore higher cancer rates might be due to age rather than higher blood pressure.

Check Your Understanding

1. A school reports that students who consume energy drinks at school tend to get lower test scores than students who don't consume energy drinks at school. What is a potentially confounding variable? Explain.
2. The local news reports that people who carry lighters with them are more likely to die of lung cancer. Identify a potentially confounding variable and justify your answer.

Generalizations for a Statistical Study

Whether we are conducting an experiment or an observational study, the goal is to generalize the results from our sample to a larger population. There are some requirements for doing so, and not all studies can be generalized.

When individuals in a study are randomly selected from a population, it is appropriate to make generalizations about the entire population from which the sample was selected. In the goldfish example, the population of interest was all the goldfish sold at a local pet store. The 10 goldfish for the experiment were randomly selected from this local pet store, so we can generalize our conclusions to the entire population of goldfish sold there.

In some experiments, it may be unethical or difficult to randomly select experimental units to participate in an experiment. In that case, the study's experimental units are obtained from volunteers and will represent the population of experimental units similar to those who participated in the study.

Some samples chosen without randomization, such as convenience or voluntary response samples, may not represent the entire population. This can lead to bias because the people in the sample differ in some specific way from the population of interest. Such samples do not generalize well to the population we are attempting to study.

In an experiment, *randomly assigning* the experimental units to the different treatments will result in a cause-and-effect conclusion between the explanatory and the response variables. Random assignment reduces the potential for confounding variables because it spreads them across the treatment groups. Any difference observed in the response variable can then be attributed to the treatment rather than to confounding variables.

Because it is difficult and often impossible to control confounding variables in an observational study, cause-and-effect conclusions are only made from observational studies if many studies are conducted over a period of time, such as those that show a causal relationship between smoking and lung cancer. In the AP Statistics course, we will only make cause-and effect conclusions from the results of a well-designed experiment.

CAUTION

The phrase “correlation does not imply causation” is used frequently when talking about statistical studies. This is an important truth, but often students take this to mean we can *never* show a causal relationship. We *can* show a causal relationship between an explanatory and response variable, and the most straightforward way to show causation is with a well-designed experiment.

Example: Generalizing survey results to the population

Charlotte’s mother is a well-known teacher at her high school. Charlotte decides to survey students about the quality of teachers at the school. She only polls those students she knows and are friendly to her. She finds an overwhelming majority approve of the quality of teaching in her school. Can she generalize her results to the entire high school?

Solution: No. Charlotte used a convenience sample and only surveyed people likely to give a positive response. She did not use a random mechanism to select her sample. This sample likely is not representative of the entire student population, so Charlotte should not generalize her results from this sample to the entire student body of the high school.

Check Your Understanding

1. It is established that people with high blood pressure have a greater risk of developing heart disease. Assume that in an observational study of residents in a large suburban town, a group of people who exercise regularly has a lower rate of heart disease than a group that does not exercise regularly. To whom can we generalize the results of this study? Can we conclude from this study that exercising regularly lowers the risk of developing heart disease? Explain.
2. Researchers want to determine whether getting at least 8 hours of sleep per night improves student performance on a math test. A random sample of 200 high school students from a large school district is selected. The students are then randomly assigned to one of two groups. The treatment group is required to get at least 8 hours of sleep the night before the test, while the control group follows their normal sleep routine. All students take the same standardized math test the following day under identical conditions. The group that received at least 8 hours of sleep scored significantly higher on average than the control group. To whom can we generalize the results of our study? Can we conclude that getting 8 hours of sleep improves math test scores? Explain.

Lesson 2.2 Summary

AP Key AP Takeaways

- A good statistical investigative question clearly identifies the explanatory and response variables, how the data will be analyzed, and indicates what kind of conclusions can be made from the results.
- In an experiment, one or more treatments are assigned to the experimental units. In an observational study, no treatments are assigned.
- A confounding variable is a variable that is related to both the explanatory and the response variables. When a confounding variable is present, it is difficult to determine whether differences in the response are due to the explanatory variable or to the confounding variable.
- If the observational or experimental units in a statistical study are randomly selected from the population of interest, then the results of the study can be generalized to the entire population. If observational or experimental units *are not* randomly selected from the population of interest, then the results can only be generalized to individuals similar to those in the study.
- A causal relationship between the explanatory and response variables can only be shown through an experiment that uses random assignment of treatments. In an observational study, we cannot show a causal relationship–only an association between the explanatory and response variables.

Lesson 2.2 Practice

Understanding the Concepts

For exercises 1–5, determine the appropriate word or phrase.

1. What are the three components of a good investigative question?

2. What are the two types of variables in a statistical study called?

3. What is a statistical study called in which the assignment to treatment groups is not made by the investigator?

4. What is a variable that is related to both the explanatory and the response variables?

5. How can the results of a statistical study be generalized to the entire population?

6. What type of study results in a cause-and-effect conclusion about the explanatory and response variables?

For exercises 7–9, identify (a) the population of interest and (b) the explanatory and response variables from each investigative question.

7. Among students in a teacher's class, does the note-taking method used during instruction (graphic organizer or traditional notes) lead to a difference in students' scores on the unit test?

8. Among students at a large suburban high school, does receiving a reminder text message from the school nurse, compared with no reminder, cause an increase in the average number of hours of sleep students get per night?

9. Among students enrolled in a physical education class at the local community college, is there an association between the number of minutes students spend being physically active each day and their resting heart rate?

For exercises 10–12, identify whether the observational study is prospective or retrospective. Explain.

10. A public health team analyzes past school attendance records to see whether students with chronic absences later showed signs of asthma.

11. Doctors track a group of teenagers for five years to see whether their exercise habits predict future heart health.

12. Researchers review hospital records to compare past sugar consumption habits of patients who developed diabetes with those who did not.

13. A medical researcher wants to determine whether exercising can lower blood pressure. At a health fair, he measures the blood pressure of 100 individuals and interviews them about their exercise habits. He divides the individuals into two categories: those whose typical level of exercise is low, and those whose level of exercise is high.

a. Is this an experiment or an observational study?

b. The subjects in the low-exercise group had considerably higher blood pressure, on the average, than subjects in the high-exercise group. The researcher concluded that exercise decreases blood pressure. Is this conclusion well justified? Explain.

14. A medical researcher wants to determine whether exercising can lower blood pressure. She recruits 100 people with high blood pressure to participate in the study. She assigns a random sample of 50 of them to pursue an exercise program that includes daily swimming and jogging. She assigns the other 50 to refrain from vigorous activity. She measures the blood pressure of each of the 100 individuals both before and after the study.

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

b. On the average, the subjects in the exercise group substantially reduced their blood pressure, while the subjects in the no-exercise group did not experience a reduction. The researcher concluded that exercise decreases blood pressure. Is this conclusion well justified? Explain.

15. An agricultural scientist wants to determine the effect of fertilizer type on the yield of tomatoes. There are four types of fertilizer under consideration. She plants tomatoes on four plots of land. Each plot is treated identically except for receiving a different type of fertilizer. The type of fertilizer that each plot receives is randomly assigned.

At the end of the growing season the yield from each plot is measured and compared.

a. What are the experimental units?

b. What are the treatments?

c. What is the response?

d. The yields differ substantially among the four plots. Can you conclude that the differences in yield are due to the differences in fertilizer? Explain.

16. A student journalist wants to know whether teachers in different departments at his large high school are paid differently. Five teachers are randomly sampled from each of three areas: math, English, and science, which are the three largest departments in the school. He then compares their annual salaries to see if there are differences across departments.

a. Is this a well-designed experiment or an observational study?

b. What sampling method was used? Explain.

c. What is a possible bias? Explain.

17. A social media company wants to know whether certain types of ads hold teen attention better. A researcher randomly selects 18 teens and assigns each to watch one 15-second ad from one of three categories: gaming, fashion, and food. Each teen's attention time (in seconds) is recorded to see which type keeps them watching the longest.

a. Is this a well-designed experiment or an observational study? Explain.

b. Identify the experimental units.

c. Identify the treatments.

d. What is the response variable?

18. Virtual reality (VR) technology is sometimes used in physical therapy. An article published in *Games Health Journal* reported the results of a study in which patients with shoulder injuries were randomly assigned to two groups. One group performed rehabilitation exercises in a VR environment designed to guide physical movement. The other group received the usual treatment, without VR.

a. Is this an experiment or an observational study?

b. What are the experimental units?

c. What are the treatments?

d. The individuals using VR experienced outcomes that were as good as those who had conventional therapy. The article states that VR technology is useful in physical therapy for all people with shoulder injuries. Is this conclusion justified?

19. In a study to investigate the relationship between social media usage and sleep quality among teenagers, researchers collect data from a large group of teens, measuring their average daily social media usage and their self-reported sleep quality.

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

b. The researchers observe that teenagers with higher social media usage report significantly poorer sleep quality compared to those with lower usage. The research team concludes that high social media usage causes decreased sleep quality in teenagers. Is this conclusion well justified?

20. In a certain supermarket, 50 customers who used a self-service checkout lane averaged 5.2 minutes of checkout time, while 62 customers who used a cashier averaged 6.1 minutes.

a. What is the explanatory variable?

b. What is the response variable?

c. Why is this considered an observational study and not an experiment?

Working with the Concepts

21. Air pollution and colds: A scientist wants to determine whether people who live in places with high levels of air pollution get more colds than people in areas with little air pollution. Do you think it is possible to design an experiment to study this question or will an observational study be necessary? Explain.

22. Cold medications: A scientist wants to determine whether a new cold medicine relieves symptoms more effectively than a currently used medicine. Sixty people exhibiting cold symptoms volunteer to take part in the study. Thirty of the volunteers are randomly assigned to take the new cold medication, and the other thirty are assigned to take the current cold medication. The group that took the new cold medication got relief from their symptoms faster and for a longer period than those who took the new medication. What kind of conclusion can be made from this study? To what population can we generalize the results?

23. Taxicabs and crime: A sociologist discovered that regions that have more taxicabs tend to have higher crime rates. Does having more taxicabs cause the crime rate to increase, or could the result be due to a confounding variable? Explain.

24. Remote or in class? Do students perform better on tests taken remotely? Each student in a class of 200 was allowed to choose whether to take the final exam on paper or remotely. One hundred and twenty students chose the remote test and had an average score of 85. Eighty students chose the paper test and had an average score of 79.

a. Can we conclude that all students will score higher on a remote test than on a paper test? Why or why not?

b. It turns out that the students who took the remote test had an average GPA of 3.4, while those who took the paper exam had an average GPA of 3.1. Explain how GPA is a confounding variable.

Extending the Concepts

25. The Salk Vaccine Trial: In 1954, the first vaccine against polio, known as the Salk vaccine, was tested in a large randomized study. Approximately 750,000 children were asked to enroll in the study. Of these, approximately 350,000 did not participate because their parents refused permission. The children who did participate were randomly divided into two groups of about 200,000 each. One group, the treatment group, got the vaccine, while the other group, the control group, got a placebo treatment. The rate of polio in the treatment group was less than half of that in the control group.

a. Is it reasonable to conclude that the Salk vaccine was effective in reducing the rate of polio?

b. To what population can we generalize the results of this study?

LESSON 2.3

Experimental Design

AP Learning Objectives

- Identify elements of a well-designed experiment.
- Identify experimental designs.
- Justify the appropriateness of a particular experimental design.

Key Terms

- placebo
- randomization
- extraneous variable
- replication
- control
- control group
- double-blind
- single-blind
- completely randomized experiment
- randomized block experiment
- matched pairs experiment

Well-Designed Experiments

An article in *The New England Journal* of *Medicine* (359:339–354) reported the results of a study to determine whether a drug called raltegravir is effective in reducing levels of virus in patients with human immunodeficiency virus (HIV). A total of 699 patients participated in the experiment. These patients were divided into two groups. One group was given raltegravir. The other group was given a **placebo**, which is a treatment with no known medical effect, such as a sugar pill. This resulted in two treatments in this experiment, raltegravir and placebo.

The experimenters gave raltegravir to about two-thirds of the subjects and the placebo to the others. To determine which patients would be assigned to which group, 462 of the 699 patients were randomly assigned to the raltegravir group. The remaining 237 patients were assigned to the placebo group.

The researchers examined the subjects after 16 weeks and measured the levels of virus in their blood. The outcome of this experiment was the number of copies of virus per milliliter of blood. Patients were considered to have a successful outcome if they had fewer than 50 copies of the virus per milliliter of blood. In the raltegravir group, 62 percent of the subjects had a successful outcome, but only 35 percent of the placebo group did. The conclusion was that raltegravir was effective in lowering the concentration of virus in HIV patients.

The importance of a well-designed experiment cannot be overstated. Experiments allow researchers to make verifiable discoveries and observations about the world around them. As we learned in the previous lesson, well-designed experiments are the most reliable way to prove a cause-and-effect relationship.

All well-designed experiments must include the following four components:

1. Comparison of at least two treatment groups
2. Random assignment of treatments to experimental units
3. Replication of the treatments
4. Direct control of potential extraneous sources of variation in the response variable when possible

If any of these four components is missing in the design of the experiment, a cause-and-effect conclusion cannot be determined.

Take Another Look

A well-designed experiment is the most reliable way of determining whether a given treatment *causes* a given outcome.

Random Assignment

The purpose of random assignment, also known as **randomization**, is to reduce bias by ensuring that every experimental unit, or subject, has an equal chance of being placed in any treatment group. This process spreads *extraneous variables* evenly among the treatment groups. An **extraneous variable** is anything outside the scope of the experiment that could influence the response variable. Extraneous variables can also be any trait of the experimental unit. Spreading these throughout the treatment groups ensures that the treatment groups are equivalent and that any differences between them will be due only to the treatments. Random assignment is the process that leads to drawing a cause-and-effect conclusion from the experiment.

In the raltegravir experiment, extraneous variables could possibly have affected the response of the patients to the drug–such as sex, age, general health, and how long the patient had been living with HIV. Using random assignment ensured that both treatment groups had a mixture of patients with varying levels of these factors. This helped create similar treatment groups.

AP Exam Tip

Be sure to use the correct terminology when discussing randomization. A researcher can take a *random sample* of students at a karate dojo for an observational study, or they can *randomly assign* treatments to those students in an experiment. On the AP Exam, do not use the words *"random sample"* when referring to an experiment. Using incorrect terminology will be penalized on the exam.

Replication

Replication within an experiment means more than one experimental unit is assigned to each treatment. In the raltegravir experiment, there were hundreds of experimental units in each treatment group, while the goldfish experiment from Lesson 2.2 used ten goldfish, five in each treatment group. It is not always necessary to include numerous experimental units in each treatment group, but more than one is required. The more units in each treatment group, the better.

Direct Control

Control involves intentionally keeping extraneous variables from becoming confounding variables. The idea is to keep treatment groups as similar as possible so the only difference is the active treatment. Then, any change in the response is attributed to the factor and not some potentially confounding variable.

Consider the fish in Figure 2.5. These fish are dramatically different in size and they live in very different environments. In the goldfish food formulations experiment in Lesson 2.2, the fishbowls were identical and the goldfish began as similar age, weight, and size. At the end of the experiment, any differences between the goldfish in the two treatment groups would be due to food formulation and not these other factors.

Figure 2.5

Example: Identifying a well-designed experiment

To assess the effectiveness of a new interactive digital technique for teaching arithmetic to elementary school children, a class of 30 first graders was chosen to participate in an experiment. Fifteen of the class members were randomly assigned to learn with the new method, and the remaining 15 first graders were randomly assigned the currently used method. The instructional time for both methods was the same. At the end of 8 weeks, the children were given a test to assess their knowledge. What are the treatments in this study? Explain why this is a well-designed experiment.

Solution: There are two treatment groups: 15 students using the new method of learning arithmetic, and the 15 students learning with the current method. This is a well-designed experiment because there is comparison between the 2 treatment groups. The children were randomly assigned to the treatment groups. Replication was used; there were 15 students in each treatment group. Direct control: all the children were from the same class, so they had the same teacher and arithmetic class at the same time each day. They all worked for the same amount of time. The random assignment of students to the different groups helps ensure that any other differences among the students are evenly distributed among the two treatment groups.

Example: Determining whether an experiment is well-designed

To assess the effectiveness of a new interactive digital technique for teaching arithmetic to elementary school children, parents were given the option to enroll their children in a class that would use the new method. A total of 100 children

were enrolled in this class, while 100 others were taught by the currently used method. At the end of 8 weeks, the children were given a test to assess their knowledge. Is this a well-designed experiment? Explain why or why not.

Solution: This is not a well-designed experiment because parents chose the treatments for their children rather than the assignments being made at random.

Check Your Understanding

1. A study involving 120 patients was conducted to determine whether increasing the level of C4A proteins in the blood speeds recovery from cardiopulmonary bypass surgery. Sixty patients were randomly assigned to receive plasma containing C4A proteins, while the remaining 60 were assigned to receive plasma that did not contain C4A proteins. The length of hospital stay was recorded for each patient. Those who received C4A proteins had an average stay of 9 days, while those who did not had an average stay of 12 days. Is this a well-designed experiment? Explain your reasoning.
2. The Aspirin Myocardial Infarction Study involved 4524 patients who had experienced a heart attack. Of these, 2267 were chosen at random to receive 1 g of aspirin each day, while the remaining 2257 were given a placebo. During the next three years, 14.1 percent of those receiving aspirin and 14.8 percent of those receiving a placebo experienced another heart attack. Is this a well-designed experiment? Explain your reasoning.

Control Groups and Placebo Effect

A well-designed experiment must include at least two treatment groups so that a comparison can be made. Often a **control group** is used to compare with the group receiving the active treatment being tested. The control group can receive no treatment, an old treatment, or a current treatment. In the previous goldfish example, the control group received the old food formulation. In some cases, like the raltegravir study, the control group receives a placebo.

Consider an experiment in which half of the subjects are given no treatment at all, and the other half are given a treatment which is known to have no medical effect. You might expect that the outcome would be the same for both groups. In practice, however, the subjects receiving the placebo will often experience a better outcome than the subjects receiving no treatment at all. The fact that some people feel better after taking a placebo is called the *placebo effect*. Placebos tend to provide a positive outcome because people associate receiving any treatment with improving their condition.

Blinding in Experiments

Because the placebo effect can be strong, it is beneficial in studies using a placebo to not tell the participants which treatment they are receiving. It is a further advantage if the assignment can be done so that neither the experimenters nor the subjects know which treatment has been assigned them.

Experiments like this are called **double-blind** experiments. The raltegravir experiment was a double-blind experiment because neither the patients nor the doctors treating them knew which patients were receiving the drug and which were receiving the placebo.

Experiments should be run double-blind whenever possible. Why? When investigators or subjects know which treatment is being given, they may tend to report the results differently. For example, in an experiment to test the effectiveness of a new pain reliever, patients who know they are getting the drug may report their pain levels differently than those who know they are taking a placebo. Doctors can be affected as well; a doctor's diagnosis may be influenced by knowing which treatment a patient received.

In some situations, it is not possible to run a double-blind experiment. Experiments in which either the experimental units or the researchers (but not both) know which treatment is being administered are called **single-blind**. For example, if researchers are comparing the effects of different types of exercise on sleep quality, then the subjects will know what type of exercise they are doing. This type of experiment can still be considered single-blind if the people who are interacting with the participants and recording the results are not aware of which treatment group the participants are in.

Some experiments do not lend themselves to either type of blinding. In an experiment that compares a treatment involving taking medication to a treatment involving surgery, both patients and doctors will know who received which treatment.

Example: Determining whether an experiment is double-blind

Is the following experiment a double-blind experiment? Explain.

Teachers in an elementary school want to assess the effectiveness of a new interactive technique for teaching arithmetic to their first graders. The school has four first grade classes in which arithmetic is taught by the same teacher. There are twenty students in each class. The first-grade teachers randomly assigned two of

the first grade classes to be taught with the new method, and the other two classes were assigned to be taught with the currently used method. At the end of eight weeks, the children were tested to assess their knowledge.

Solution: While it is possible that the children might not know which treatment they are receiving (which is new and which is current), the teachers will know what method was used. This experiment cannot be double-blind.

Check Your Understanding

For each of the following, determine whether it could be designed as a double-blind experiment and explain why or why not.

a. A study to determine whether remote learning is as effective as classroom learning.

b. A study to determine whether a new drug, taken as a pill, reduces the amount of time needed to fall asleep.

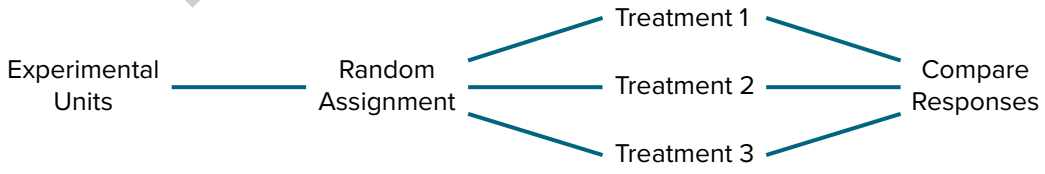
c. A study to determine whether exercising for an hour each day results in weight loss.

d. A study to determine whether a new artificial sweetener increases blood pressure.

Completely Randomized Design

The type of well-designed experiments described in the previous examples is called a **completely randomized experiment**. In this type of experimental design, there is no restriction on which subjects may be assigned which treatment; treatments are assigned to experimental units completely at random. Often the number of experimental units assigned to each treatment is the same, but it is not a requirement that treatment groups be the same size. Figure 2.6 diagrams a completely randomized design.

Figure 2.6 Completely Randomized Design



AP Exam Tip

You will likely be asked to describe an experimental design on the AP Exam. It is important that you name each treatment group. Simply calling it "Treatment Group 1" is not sufficient. You must be specific about what treatment that group is receiving.

Writing in AP Statistics

What are you writing about? On the AP Exam, you may be asked to **describe** how to randomly assign treatments to experimental units. To receive full credit, you must label all the experimental units, explain how you will randomly assign the treatments, and tie the treatment back to the experimental unit. Remember, when you are writing in AP Statistics, always include context. If the problem is about monkeys and bananas, your response must refer to monkeys and bananas, not just data points.

Sample: A new type of shin guard claims to reduce injuries for high school soccer players. Clarence is the athletic director at a district with three large high schools. Two hundred athletes play soccer across the programs Clarence oversees each season. He wants to use a completely randomized design for his study to determine whether the new type of shin guard lives up to its claim.

Describe an appropriate method Clarence could use to randomly assign the two treatments, new shin guards and standard shin guards, to 100 players each.

Student response #1

Assign each player a unique number between 1 and 200. Using a random number generator, generate 100 unique integers (ignore repeats) between 1 and 200 (inclusive). Assign those players with the 100 unique integers to wear the new shin guards for a season, and the remaining 100 players will be assigned to wear the standard shin guards.

Student response #2

Write the numbers 1 to 100 on equal sized slips of paper. Put the slips of paper into a hat. Draw out 100 slips. Those 100 get the new shin guards, and the other ones get the standard shin guards.

The **first student** included all three components necessary to completely answer the question: (1) created appropriate labels for the units/treatments, (2) described how to correctly implement the random process so that every possible assignment is equally likely, and (3) the response indicated a random process that resulted in 100 players being assigned each type of shin guard.

The **second student** used the slips of paper method, which is not inherently incorrect. They also included the required context. Unfortunately, this response (1) failed to create appropriate labels for the units/treatments, and (2) did not randomize (mix thoroughly) the slips of paper before drawing or indicate that the slips of paper were drawn without replacement.

Practice the writing: Some dance studios purchase professional choreography as opposed to using in-house choreography, where the local dance teachers create the choreography. Shatina owns a studio and wants to investigate whether there is a difference in scores between purchased choreography and in-house choreography. Shatina's studio will compete with 60 different dances in multiple categories and age groups during the year. Shatina would like to use in-house choreography for half of the dances and purchased choreography for the other half, then see whether there is a difference in the competition scores for the two types of choreography. Shatina will use a completely randomized design to conduct her investigation. **Describe** an appropriate method Shatina could use to randomly assign the two treatments, purchased or in-house choreography, to thirty dances each.

Randomized Block Design

In some situations, it is desirable to restrict the randomization a bit. For example, suppose that two reading programs are to be tested in an elementary school with children in grades one through four. If children are assigned completely at random to the programs, it is possible that one of the programs will end up with more fourth graders while the other one will contain more first graders. Since fourth graders tend to be better readers, this will give an advantage to the program assigned to that group. This possibility can be avoided by randomizing the students within each grade separately. In other words, randomly assign exactly half of the students within each grade to each reading program. This type of experiment is called a **randomized block experiment**. In the example just discussed, each grade constitutes a block.

In a randomized block experiment, the subjects are divided into blocks in such a way that the subjects in each block are the same or similar regarding a variable that is related to the outcome. Age and sex are commonly used blocking variables. Then the subjects within each block are randomly assigned a treatment.

Randomized block experiments are often used when the number of subjects is small because randomization may not prevent large differences between treatment groups. A randomized block experiment is also a good choice when there is a factor that very strongly affects the outcome. Then, even small differences between the treatment groups may influence the results. Blocking reduces variability in an experiment.

Figure 2.7 shows a diagram of a randomized block design.

Figure 2.7 Randomized Block Design

Example: Identify blocks in a randomized block experiment

In a study involving 100 males and 100 females with high blood pressure, 50 males and 50 females, chosen at random, are given a new blood pressure drug while the other subjects receive a placebo. What are the blocks in this experiment and why did researchers block on this variable?

Solution: The blocks are the males and the females. It is appropriate to block using sex because it is possible that the drugs affect the blood pressures of men and women differently.

Matched Pairs Design

A **matched pairs experiment** is a type of randomized block design in which each block consists of two subjects. There are two treatments, A and B (one of which may be a placebo). One subject in the block, chosen at random, receives treatment A, and the other receives treatment B. The two subjects in each block should be chosen to be as similar as possible regarding factors that may affect the outcome. It is common to match on age and sex.

Example: Describe the advantages of matching

In a study of the effect of a new weight-loss drug, pairs of subjects are matched on age, sex, and weight. One member of each pair receives the drug, while the other receives a placebo. What are the advantages of this experimental design?

Solution: Age, sex, and weight may all affect the outcome of a weight-loss drug. By matching on these factors, they will not affect the outcomes in this study.

A common form of matched pairs experiment is one in which individuals are matched to themselves. For example, to compare the effectiveness of two allergy medications, each subject would take one medication for a certain length of time, then the other for the same length of time. Which medication was taken first would be chosen at random. The advantage of matching subjects with themselves is that there is no difference between the two treatment groups, except for the treatment.

Matched pairs studies are sometimes conducted on identical twins. Each twin receives a different treatment. Since identical twins have identical genes, any difference in outcomes cannot be explained by genetics. For example, studies have shown that with identical twins, one of whom smokes, and the other does not, there is a higher death rate among the smokers. This indicates that the bad health effects of smoking are not explained by genetic differences between smokers and non-smokers.

Take Another Look

Matched pairs experiments can be done as follows:

a. Two different experimental units can be matched by some characteristics they share that may influence the response variable.

b. A single experimental unit can be given both treatments in randomized order.

Example: Determining whether an experiment is a matched pair experiment

A marketing firm wants to compare two website designs: Design A and Design B, to see which leads to higher customer spending. A random sample of ninety online shoppers is selected from the company's customer base. Shoppers are matched into pairs based on age, shopping history, and average past spending. Within each pair, one shopper is randomly assigned to view Design A, and the other is assigned to view Design B. The amount of money spent during the visit is recorded. Is this a matched pairs experiment? If so, what are the pairs, and to whom can the results be generalized?

Solution: Yes, this is a matched pairs experiment because similar shoppers are paired based on key characteristics and treatments are randomly assigned within each pair. The pairs are the two shoppers within each matched pair. Because the shoppers were randomly selected from the company's customer base, the results can be generalized to that customer base, but not necessarily to all online shoppers.

Check Your Understanding

1. In a study of a medication designed to lower cholesterol, subjects with high cholesterol were divided into 6 groups: Men under age 40, men aged 40–60, men over 60, women under age 40, women aged 40–60, and women over 60. Within each group, half the subjects were chosen at random to receive the new medicine, while the other half received a placebo. Is this a randomized block experiment? Explain. If so, what are the blocks?

2. To test a new ointment designed to treat acne, subjects with acne applied the ointment to one side of their face, and a placebo ointment to the other side. The side receiving the placebo was chosen at random. What type of experiment is this? Explain.

Choosing an Experimental Design

A *completely randomized design* is the simplest to implement and ensures all treatment groups are similar. However, the results from this design can be less precise if there is significant variability among the experimental units.

A *randomized block design* improves precision by grouping experimental units with similar characteristics (which may influence the response variable) into blocks before randomly assigning treatments within each block. This helps control for variability due to factors we know about. It can be more difficult and time consuming to set up a randomized block design than a completely randomized design. A *matched pairs design* is most helpful when comparing two treatments on closely related or identical subjects (as in twins), but might prove troublesome to set up with large samples. When choosing an appropriate experimental design, the statistician is tasked with balancing simplicity, control, and precision.

Example: Choosing an Experimental Design

Pool noodle javelin is a favorite pastime in the midwest region of the United States. Suppose we want to know whether throwing the pool noodle underhand or overhand results in the pool noodle traveling farther. At the last pool noodle javelin match, people from ages four through ninety-four competed. How could we design an experiment to control extraneous variables such as age and fitness level?

Solution: This experiment could be conducted in a matched pairs design, where all the participants are matched with themselves. Each person participating in the experiment will throw the pool noodles underhanded and overhanded. Which method they use on their first throw will be randomly chosen, and they will throw with the other method on their second throw. The difference in the noodles travel (underhanded and overhanded) will be recorded and compared. Using a matched pair design allows any differences in the distance to be attributed to whether the pool noodle was thrown underhanded or overhanded, and not the age of the subject, their height, overall fitness level, ability to throw, etc. If a completely randomized design were used, most of the children and elderly participants might be assigned one treatment group and more of the younger and more fit participants in the other treatment group. Our results might not accurately show which throwing method results in a longer throw.

AP Statistics Skills Practice

Skill 2.B - Justify an appropriate method for ethically gathering and representing data.

One of the most fundamental skills needed in the study of statistics is how to collect data in an appropriate way. In this course and on the AP Exam, you need to be able to determine when an observational study or an experiment is more

appropriate and what design will best serve your study. You also must be able to justify your choices. If we need to establish a cause-and-effect relationship, then an experiment is necessary. If an experiment is not possible because of ethical or logistical reasons, then state that. In justifying the choice of a sampling design or experimental design, state your reasoning for that choice based on a statistical advantage such as an extraneous variable that might affect the response.

On the AP Exam

This sample question asks you to determine the best experimental design to use for a described study and justify your choice. You will notice there are two answer choices with the same design and different justifications.

A culinary institute wants to study whether the crispiness of baked flatbread is affected by the baking time. The study will use 4 different dough recipes because baking time may influence different dough types in different ways. Six flatbreads will be made with each type of dough for a total of 24 flatbreads, and 3 baking times will be tested. What is the best design to use for the study?

A. A completely randomized design, because baking times are randomly assigned to all 24 flatbreads without accounting for differences among dough types, resulting in 8 flatbreads baked at each baking time.

B. A completely randomized design, because baking times are randomly assigned so that 2 flatbreads of each dough type are baked at each baking time, ensuring equal sample sizes for each baking time.

C. A randomized block design, because dough type is identified as a source of variability, but flatbreads are assigned to baking times only to ensure 8 flatbreads are baked at each baking time.

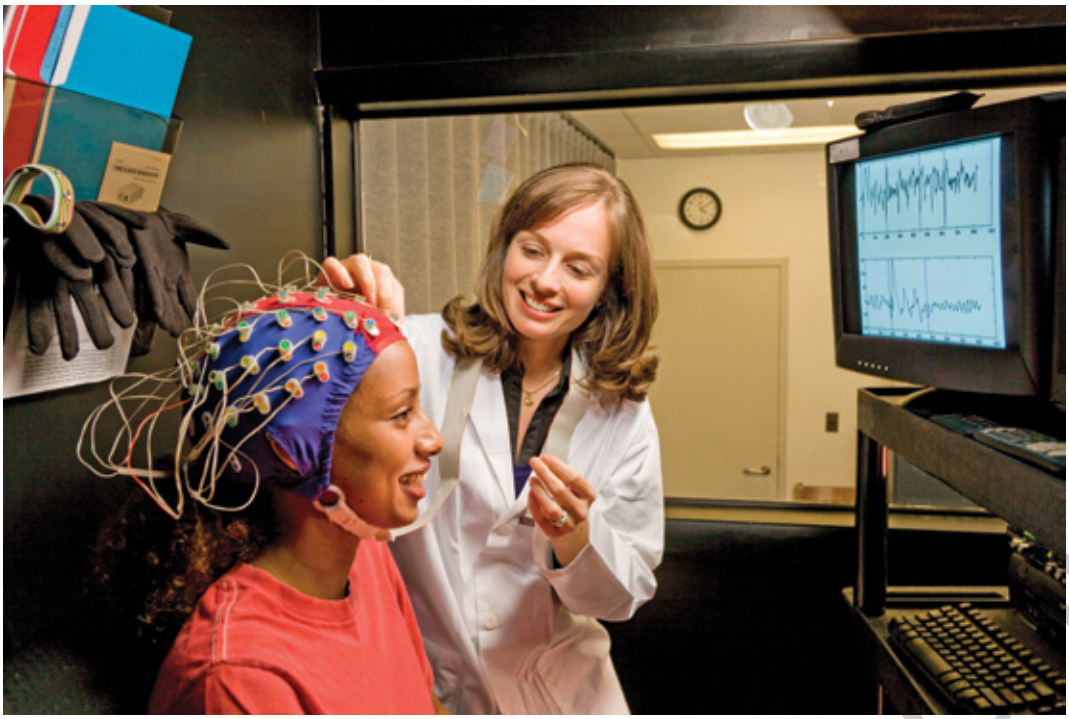
D. A randomized block design, because dough type is used as a blocking variable and baking times are randomly assigned within each dough type so that 2 flatbreads of each dough type are baked at each baking time.

What is the correct answer?

In this study, the researcher wants to compare the effect of baking time on crispiness while accounting for the fact that different dough recipes may naturally behave differently. Before interpreting differences in crispiness, you must first ensure that each dough type is treated similarly across all baking times. Also, before making conclusions about the effect of baking time, you must control for the variation created by dough type.

A randomized block design accomplishes this by treating each dough type as its own block and then randomly assigning flatbreads within each block to the baking times. Because each dough type has 6 flatbreads and there are 3 baking times, the correct allocation is 2 flatbreads of each dough type per baking time. This design isolates the effect of baking time while holding dough type constant.

The correct answer is **D**.



Ethical Considerations with Experiments

Experiments involving animals or human subjects require special ethical considerations to ensure that no harm is caused. Most academic studies using animal or human participants must be reviewed by an ethics or institutional review board, which is responsible for ensuring the study does not pose unnecessary risk. For example, when testing a new cancer treatment, it would be unethical to give one group the new treatment and another group no treatment at all.

Experiments that involve human subjects also typically require written informed consent. This consent form explains what is involved in the study, what participants will be asked to do, and any potential risks. By signing, participants indicate they understand and agree to take part. Most informed consent documents also explicitly state that participants may withdraw from the study at any time, for any reason. Researchers are also obliged to include a post-experimental debrief to ensure participants were not harmed in any way by their participation in the experiment.

Lesson 2.3 Summary

AP Key AP Takeaways

- A well-designed experiment has four components: comparison, randomization, replication, and direct control.
- Extraneous variables can introduce bias and should be minimized or avoided.
- A double-blind experiment is one in which neither the experimental units nor the researchers are aware of what treatment the participant received. A single-blind experiment is one in which either the experimental units or the researchers are not aware of what treatment was received.
- A completely randomized design randomly assigns treatments to all experimental units. It is the simplest experimental design.
- A randomized block design allows researchers to group experimental units based on a shared trait or characteristic believed to influence the response variable. Within each block, treatments are randomly assigned to the experimental units.
- A matched pairs design is a special type of block design; a block of size two or one. In a matched pairs design, experimental units are either:
    - paired together by like characteristics (e.g., twins) and one member of each pair is randomly assigned treatment A and the other receives treatment B.
    - paired with themselves and given both treatments in a randomly assigned order.
- Blocking reduces variability in an experiment.

Lesson 2.3 Practice

Understanding the Concepts

In exercises 1–5, determine the appropriate word or phrase.

1. What are the four components of a well-designed experiment?

2. What do we call a study in which neither the investigators nor the subjects know who is getting which treatment?

3. What do we call a treatment that is known to have no medical effect?

4. What do we call an experiment in which each experimental unit receives both treatments in a randomized order?

5. Why would we want to use a randomized block design for an experiment?

In exercises 6–10, determine whether the statement is true or false. If the statement is false, rewrite it as a true statement.

6. Using a placebo in an experiment helps reduce the effects of confounding variables, but it does not control for the placebo effect.

7. In a single-blind experiment, neither the subjects nor the researchers who interact with the subjects know which treatment each subject receives.

8. A matched pairs experiment is a special case of a randomized block experiment in which each block contains exactly two experimental units.

9. Replication increases the likelihood that an experiment will detect a real difference between treatments if one exists.

10. A completely randomized experiment assigns experimental units to treatments using randomization with no blocking.

11. A social media company wants to know which types of ads hold teen attention better. A researcher randomly selects 18 teens and assigns each to watch one 15-second ad from one of three categories: gaming, fashion, and food. Each teen's attention time (in seconds) is recorded to see which type keeps them watching the longest. What type of experimental design was used?

12. It is established that age is a factor in the ability to focus. How might you design the experiment from question 11 to take this into account?

For exercises 13–16, Dark or Light Mode? A researcher wants to know whether students read faster in dark mode or light mode when using their phones. Thirty teens are recruited and each read the *same* short article twice: once in dark mode and once in light mode. The order in which they read the article is randomly assigned. For each reading session, the researcher records the time (in seconds) it takes each teen to finish the article and calculates the difference in reading times between the dark mode and light mode.

13. Justify that this is a well-designed experiment.

14. What type of design was used? Explain.

15. What are the experimental units, treatments, and response variable?

16. Could this experiment be done double-blind? Explain.

Working with the Concepts

17. In a research study, 90 adults with seasonal allergies are randomly assigned to one of the two treatment groups. One group receives an herbal nasal spray, and the other group receives a saline placebo spray. Each participant uses their assigned spray every evening. The next morning, they complete an online questionnaire about symptom severity. A separate team of researchers will receive the data labeled only with participant ID numbers and dates, and will analyze the results without knowing which spray each participant received.

a. Identify the four components necessary for this to be a well-designed experiment.

b. Is this experiment single-blind or double-blind? Explain.

18. A biology instructor wonders whether background music affects how well students do on short, in-class practice quizzes. From January through April, in both Biology I and Biology II, every time a quiz is given the instructor will either play soft instrumental music or no music at all. There are 10 quizzes over the course of that time. For each quiz, the instructor records whether music was played and the average quiz score.

a. Identify the explanatory and response variables.

b. The instructor decides to block on course level – Biology I and Biology II. Sketch a diagram of what a randomized block design would look like.

c. Describe an appropriate method for randomly assigning the quizzes to the different treatments.

d. Is it possible to have blinding in this experiment?

19. A fitness researcher conducted a study to measure the effects of a new endurance supplement. All 40 volunteer runners in the study were given the supplement, and their improvement in mile-run time was measured after 3 months.

a. Identify any problems that this experimental design has.

b. Describe how the researcher could address and correct the problems identified in part (a).

Extending the Concepts

20. You are curious about whether using a new study-skills app helps students improve their quiz scores. Describe in detail an appropriate experiment that you could use to collect data to help answer your question. Be sure to name the type of design you would use and describe how you would randomize the assignment of treatments.

21. Another Salk Vaccine Trial: Another study of the Salk vaccine, conducted at the same time as the trial described in Exercise 25 from Lesson 2.2, used a different design. In this study, approximately 350,000 second graders were invited to participate. About 225,000 did so, and the other 125,000 refused. All of the participating second graders received the vaccine. The control group consisted of approximately 725,000 first and third graders. They were not given any placebo, so no consent was necessary.

a. Was this a well-designed experiment? Explain.

b. Was it double-blind? Explain.

c. Perhaps surprisingly, polio was more common among upper-income and middle-income children than among lower-income children. The reason is that lower-income children tended to live in less hygienic surroundings. They would contract mild cases of polio in infancy while still protected by their mother's antibodies, and thereby develop a resistance to the disease. The children who did not consent to participate in the study were more likely to come from lower-income families. In this study, the treatment group consisted of children who had consent to participate. The control group consisted of all first and third graders. It turned out that the results of this study seriously underestimated the effectiveness of the vaccine. Explain why.

Chapter 2 Review

Reviewing the Concepts

In exercises 1–3, identify the kind of sample that is described. Justify your answer.

1. **Website ratings:** A popular website is interested in conducting a survey of 400 randomly selected visitors to the site in such a way that 200 of them will be under age 30, 150 will be aged 30–55, and 50 will be over 55.

2. **School days:** A researcher randomly selects 4 of 12 high schools in a certain region and surveys all of the administrative staff members in each school about a potential change in the ordering of supplies.

3. **Political polling:** A pollster obtains a list of registered voters and uses a computer random number generator to choose 100 of them to ask which candidate they prefer in an upcoming election.

4. **Fluoride and tooth decay:** Researchers examine the association between the fluoridation of water and the prevention of tooth decay by comparing the prevalence of tooth decay in countries that have fluoridated water with the prevalence in countries that do not.

   a. Is this a well-designed experiment or an observational study? Explain.

   b. Assume that tooth decay was less common in countries with fluoridated water. Explain why this result is due to a confounding variable.

5. **Fuel efficiency:** A study is conducted to compare the fuel efficiency of two types of car engines, A and B. Several models of car are used in the study. Within each model, three cars are randomly assigned to engine A and three to engine B.

   a. Is this a randomized block experiment or a matched pairs experiment? Explain.

   b. Assume that one type of engine had noticeably better gas mileage than the other. Could this be due to a confounding variable? Explain.

6. **Phones and driving:** Almost all states in the United States prohibit texting while driving, and several states prohibit all phone use while driving. To determine the extent to which using a phone increases the risk of a traffic accident, a researcher examines accident reports to obtain data about the number of accidents in which a driver was using a phone.

   a. Is this a well-designed experiment or an observational study? Explain.

   b. Assume that the accident reports show that people were more likely to have an accident while using a phone. Can we conclude that using a phone while driving increases the likelihood of having an accident? Explain.

7. **Turn in your homework:** The English department at a local college is considering the use of electronic-based assignment submission in its English composition classes. They are particularly interested in seeing if submitting assignments electronically rather than in person will increase the number of on-time submissions.

To study its effects, each section of the class is divided into two groups at random. In one group, assignments are submitted by turning them in to the professor on paper. In the other group, assignments are submitted electronically.

a. What makes this a well-designed experiment?

b. What are the experimental units, treatments, and response variable?

c. Describe how you would randomly assign treatments to a class of 120 English composition students.

In exercises 8–9, explain the type of bias in the study and if it will lead to an overestimate or underestimate.

8. Longevity: A life insurance company wants to study the life expectancy of people born in 1950. The company's actuaries examine death certificates of people born in that year to determine how long they lived.

9. Political opinion: A member of the U.S. Congress sent out questionnaires to 10,000 constituents to ask their opinions on a new health-care proposal. A total of 200 questionnaires were returned, and 70% of those responding supported the proposal.

10. Investigative Question: A high school is curious about whether phone habits during the school day are related to student focus. The administration randomly selects students and assigns some to keep their phones locked in pouches during class for a week and allows others to keep their phones as usual. At the end of the week, all students take the same attention and concentration assessment. Write a clear, statistically appropriate investigative question that could be used to study this situation. Make sure you include the population of interest, explanatory variable (and any groups) and the response variable in your question.

Write About It

11. Describe the difference between a stratified sample and a cluster sample.

12. Explain why it is better, when possible, to draw a simple random sample rather than a convenience sample.

13. Describe circumstances under which each of the following samples could be used: simple random sample, stratified sample, cluster sample, systematic sample.

14. What are the primary differences between a well-designed experiment and an observational study?

15. What are the advantages of a double-blind study? Are there any disadvantages?

16. Provide an example of a study, either real or hypothetical, that is conducted by people who have an interest in the outcome. Explain how the results might possibly be misleading.

17. In a randomized study that is not double-blind, patients are aware of whether they are getting the treatment or a placebo. In this situation, discuss how the placebo effect can be influenced by the expectations of the patient or the attitude of the person administering the placebo.

AP Exam Practice

Multiple Choice Questions

Directions: Select the best answer for each question.

1. A team of environmental scientists is preparing to study how recreational use affects the ecological health of freshwater lakes within a large national park. Before beginning data collection, the researchers must formulate an investigative question in which the first component clearly identifies the variable(s) of interest:

The scientists are considering several possible investigative questions:

I. How does human activity impact the natural balance of lakes in the park?

II. Is there a relationship between weekly boating hours on a lake and the concentration of dissolved oxygen in that lake's water?

III. Why do some lakes seem more polluted than others, even when they are in the same region?

IV. Does increased tourism lead to environmental harm in lakes across the park system?

Determine the question in which the first component most appropriately guides data collection by directly specifying measurable variables.

A. Question I, because it captures the broad goal of the study and allows the researchers to explore multiple potential indicators of ecological health.

B. Question II, because it explicitly identifies two quantitative variables—weekly boating hours and dissolved oxygen concentration—that can be measured and collected systematically.

C. Question III, because comparing lakes that "seem more polluted" encourages the researchers to gather both observational and subjective data from multiple sources.

D. Question IV, because it focuses on tourism, which is the primary factor believed to cause environmental issues in lakes, allowing for causal investigation.

2. A supermarket manager wants to determine whether patrons are satisfied with the customer service they receive while shopping. The manager is considering several possible data collection methods. Which of the following scenarios best describes conducting a census?

A. The manager stations an employee near the exit and asks every 10th customer who leaves the store during a four-hour window to complete a brief customer service survey.

B. The manager uses a random number generator to select 200 customers from the store's loyalty card database and emails those selected customers with a customer-service questionnaire.

C. The manager selects three randomly chosen shopping days and surveys all customers who shop in the store only on those three days.

D. The manager collects customer service feedback from every patron who shops in the store for an entire month, ensuring that all members of the population of interest are included in the data.

3. College athletes devote a substantial amount of time to their sport, including training sessions and competition-related activities. To what extent does this commitment affect their academic performance? A team of psychology majors at a small liberal arts college conducted a study to determine whether student-athletes have higher GPAs than non-athletes. The research team contacted the transcript office and obtained the GPAs of a random sample of students at the college who were athletes and the GPAs of a separate random sample of students at the college who were not athletes. They then compared the average GPA for each group. What type of study is this?

A. Observational Study

B. Experiment

C. Sample Survey

D. Census

4. A university registrar wants to estimate the proportion of first-year students who plan to change their major before the end of the academic year. The registrar decides to select a simple random sample (SRS) of 120 students from the population of 3,200 first-year students. To obtain the sample, the registrar assigns each student an ID number from 0001 to 3200 and then uses a computer to generate 120 unique random integers in that range. The registrar explains that this procedure ensures that every possible group of 120 students has the same chance of being selected.

However, a member of the administrative staff suggests an alternative method: dividing the 3,200 students into dormitory groups, then selecting exactly 4 students at random from each dormitory to reach a total of 120. The staff member claims this method is "also an SRS since each student still has an equal chance of being chosen."

Which of the following statements correctly evaluates whether the registrar's sampling method and the staff member's suggested method both qualify as simple random samples?

A. Only the registrar's method is an SRS because every possible sample of 120 students is equally likely. The staff member's method is not an SRS because restricting the selection to exactly 4 students per dormitory prevents many possible samples of size 120 from ever being chosen.

B. Both methods are SRS procedures because each student has the same individual probability of being selected, regardless of whether the sample is drawn by dormitory or by random numbers.

C. Only the staff member's method is an SRS because dividing by dormitory ensures equal representation from each group, which is required for a simple random sample.

D. Neither method is an SRS because an SRS requires sampling with replacement, and both methods sample students without replacement.

5. A building developer is planning to construct a new complex near the coastline but is concerned about potential ground erosion, which could compromise the structural integrity of the foundation. To assess these risks, the developer decides to examine the soil conditions. The developer selects random samples of soil from each of the distinct areas including the beach, marsh, bay, and causeway. What type of sampling method is being used in this investigation?

A. A simple random sample

B. A cluster random sample

C. A convenience sample

D. A stratified random sample

6. KWPN News, a local television news station based in Phoenix, Arizona, launches an online poll to measure their viewers' opinions on the proposed water restrictions to private residences. The poll is hosted on the station's website and is promoted several times during the eleven o'clock news broadcast. Viewers are encouraged to "visit KWPN.com to let us know whether you support or oppose the water project." The poll records 18,000 responses within 24 hours and shows that 92% of KWPN viewers do not support the water restriction project. Which answer best describes the type of bias in this sample?

A. voluntary response bias

B. non-response bias

C. response bias

D. undercoverage bias

7. Which of the following demonstrates undercoverage bias?

A. During routine checkups at a community health clinic, doctors have begun collecting data on patients' medication habits. During each consultation, the doctor asks the patient whether they have been consistent with their prescribed medication routine over the past month. The patient's response is recorded in the clinic's electronic survey system and added to a database used for evaluating adherence patterns. Clinic administrators plan to analyze these responses to identify trends and develop strategies for improving medication compliance.

B. A local public health department distributes a survey to residents to gather opinions about proposed changes to the town's emergency services. Several of the survey items are written with complex wording and double negatives, such as asking whether respondents "do not oppose discontinuing certain nonessential services." Many participants report difficulty interpreting these questions, leading to inconsistent or confused responses.

C. A university research team conducts a survey to assess students' attitudes regarding online learning. To gather responses efficiently, the researchers distribute the questionnaire through university email accounts and require participants to log in with their student credentials. However, this procedure unintentionally excludes students enrolled in the university's continuing education program, who do not receive campus email addresses.

D. A city government conducts a telephone survey to measure residents' satisfaction with local public transportation. The survey team randomly selects 1,000 phone numbers using a random number generator and attempts to contact each household. However, a portion of the selected individuals do not answer the phone, have disconnected numbers, or refuse to participate. As a result, some of the randomly selected individuals are not included in the final sample.

8. Members of U.S. Department of Urban Planning conducted a study comparing a variety of demographics present in multiple cities across the country. They were particularly interested in knowing if having access to public parks can decrease the obesity rate in communities. Researchers collected data on the number of public parks in each city, the average obesity rate among residents, and the average household income. The study found that cities with more parks tend to have lower obesity rates. However, the analysis also reveals that cities with more parks tend to have higher average incomes. It is known that high-income people are less likely to be obese. Which of the following best describes the variable income?

A. Control variable

B. Confounding variable

C. Blocking variable

D. Response variable

9. A team of behavioral economists is studying how different types of digital reminders influence the likelihood that adults complete an online financial-literacy training module within a 14-day window. The population of interest is working adults employed at mid-size technology companies across the United States. The researchers plan to assign each participant to one of three reminder types: text-message reminder, email reminder, and push-notification via mobile app.

Preliminary interviews reveal substantial variation across companies in work schedules, digital-security policies, and whether employees are allowed to receive non-work notifications on employer-issued devices. Some companies even block push-notifications entirely.

Identify which experimental design would be most appropriate for achieving the researchers' goals?

A. Completely randomized experiment, because it ensures that all participants have an equal chance of receiving any reminder type, and avoids biases caused by differences across companies.

B. Randomized blocked experiment with companies as blocks, because company-level differences are likely to systematically affect how effective each reminder type is, and blocking will control these sources of variability.

C. Matched pairs design, pairing employees from different companies who have similar digital-usage patterns and randomly assigning one reminder type to each member of the pair.

D. Cluster-randomized experiment, in which each company is randomly assigned one reminder type, because this design eliminates the need to manage company-level restrictions within the randomization process.

10. An automobile manufacturer wants to compare the effectiveness of two executive training programs for recent college graduates. Each employee will attend both programs: one during their first year and the other during their second year, with the order of participation randomized for each employee to control for any order effects. A mentor will observe each employee throughout the programs and assign a score based on measurable improvements in leadership skills, decision-making, and project management. The company plans to compare the mentor-assigned improvement scores between the two programs to determine which program leads to greater skill development. Which of the following best describes the program?

A. Randomized block design

B. Completely randomized design

C. Observational study

D. Matched pairs design

Free Response Question

Directions: Show all your work. Indicate clearly which methods you use. You will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

1. A science center believes that adding a natural plant extract to the water in its butterfly habitats might increase butterfly activity. Increased activity is desirable because it makes the exhibit more engaging for visitors. The center conducts a completely randomized experiment using 60 identical butterfly enclosures. To get a baseline measurement for comparison, the activity level of the butterflies in each enclosure is measured on a scale of 0 to 10, where 0 means little movement and 10 means very high activity. Thirty enclosures are randomly assigned to receive water that contains the natural plant extract, and the other 30 enclosures receive regular water with no extract added. After one week, staff record the activity level in each enclosure on a scale from 0 to 10 and compare it to the baseline level for the enclosure.

a. Determine a valid investigative question for this study. Ensure all three components of a valid investigative question are included in your response.

b. Based on the information provided about the science center's experiment, identify each of the following.

- Experimental units
- Treatments
- Response variable

c. Describe an appropriate method the staff could use to randomly assign enclosures to receive water with the plant extract or regular water.

Case Study Follow Up

Air pollution is a serious problem in many places. One form of air pollution that is suspected to cause respiratory illness is particulate matter (PM), which consists of tiny particles in the air. Particulate matter can come from many sources, most commonly ash from burning, but also from other sources such as tiny particles of rubber that wear off automobile and truck tires.

The town of Libby, Montana, was recently the focus of a study on the effect of PM on the respiratory health of children. Many houses in Libby are heated by wood stoves, which produce a lot of particulate pollution. The level of PM is greatest in the winter when more stoves are being used and declines as the weather becomes warmer. The study attempted to determine whether higher levels of PM affect the respiratory health of children. In one part of the study, schoolchildren were given a questionnaire to bring home to their parents. Among other things, the questionnaire asked whether the child had experienced symptoms of wheezing during the past 60 days. Most parents returned the questionnaire within a couple of weeks. Parents who did not respond promptly were sent another copy of the questionnaire through the mail. Many of these parents responded to this mailed version.

The table presents, for each day, the number of questionnaires that were returned by parents of children who wheezed, the number returned by those who did not wheeze, the average concentration of particulate matter in the atmosphere during the past 60 days (in units of micrograms per cubic meter), and whether the questionnaires were delivered in school or through the mail. We will consider a PM level of 17 or more to be high exposure and a PM level of less than 17 to be low exposure.

Date	PM Level	Number of People Returning Questionnaires	Number of People who Wheezed	School/Mail
March 5	19.815	3	0	School
March 6	19.885	72	9	School
March 7	20.006	69	5	School
March 8	19.758	30	1	School
March 9	19.827	44	7	School
March 10	19.686	31	1	School
March 11	19.823	38	3	School
March 12	19.697	66	5	School
March 13	19.505	42	4	School
March 14	19.359	31	1	School
March 15	19.348	19	4	School
March 16	19.318	3	1	School
March 17	19.124	2	0	School
April 12	14.422	10	1	Mail
April 13	14.418	9	1	Mail
April 14	14.405	8	0	Mail
April 15	14.141	3	0	Mail
April 16	13.910	4	0	Mail
April 17	13.951	2	0	Mail
April 18	13.545	2	0	Mail
April 20	13.326	3	0	Mail
April 22	13.154	2	0	Mail

1. Explain why more questionnaires given in school were returned more often than those mailed to homes.
2. As the weather gets warmer, PM level goes down because wood stoves are used less. Explain how this causes the method of response (school or mail) to be related to PM level.
3. It is generally the case in epidemiologic studies that people who have symptoms are often eager to participate, while those who are unaffected are less interested. Explain how this may cause the method of response (school or mail) to be related to the outcome.
4. Rather than send out questionnaires, the investigators could have telephoned a random sample of people over a period of days. Explain how this might have reduced any confounding.
5. What are some other potential problems with the telephone method in question 4?

Unit 1 Summary

Putting It All Together

In this unit, you have learned about collecting and exploring data. At the beginning of the statistical process:

- Start with one or more investigative questions – What do we want to know about?
- Determine what kind of data to collect. What is the population that we are interested in? What variables are we interested in? Are they categorical or quantitative? If they are quantitative, are they continuous or discrete?
- What would be the best way to collect the data that we need? Do we want to do a survey? An observational study with random sampling? How are we going to reduce or prevent bias in our data? Does this statistical study call for a well-designed experiment? If so, what experimental design would be most appropriate?
- Once we get our data – how are we going to organize it? We need graphs and summary statistics, so the type of data helps inform which ones we choose. What patterns do we notice? If the data are quantitative, what can we say about its distribution?

Moving Forward with These Concepts

Now that we have our investigative questions, along with data, graphs, and summary statistics, how do we answer the questions? We have an idea of what the answer *might* be, but it is not quite enough to be certain – at least as certain as we CAN be in Statistics. Before we can answer the questions with statistical certainty, we need to learn about probability. Probability is the study of randomness. It connects the results of our data collection and exploration to answers. We begin the study of probability in the next unit.

AP Exam Practice Unit 1

Multiple Choice Questions

Directions: Select the best answer for each question.

1. The student government at SCH Academy is responsible for selecting a musical group to perform at the upcoming prom. To ensure that its recommendation reflects the preferences of the overall student body, the committee conducts a census by distributing a survey to all 500 students enrolled in the high school. Each student is asked to indicate their top choice among several candidate bands. After all responses are collected, the committee compiles the data and constructs a bar graph summarizing the number of votes each band received. The resulting distribution of preferences will be used to inform the committee's recommendation to the school administration. Which of the following statements correctly describes an appropriate feature of the graph for this categorical distribution?

- **A.** The bars must be touching each other because the data are categorical.
- **B.** The order of the bars must follow alphabetical order.
- **C.** The height of each bar represents the frequency or relative frequency of each category.
- **D.** A histogram is a more appropriate display for this data.

2. The League of Women Voters is conducting a study to better understand certain characteristics of women living in a large metropolitan area. To gather reliable information, researchers select a simple random sample (SRS) of 1,643 adult women from the city's voter registration database. Each selected individual is asked to report the following information: age (in years), highest completed level of education (such as high school diploma, associate degree, bachelor's degree, etc.), and annual earned income (in dollars). How many variables are being measured in this study?

- **A.** one categorical and two quantitative
- **B.** three quantitative
- **C.** one categorical and three quantitative
- **D.** two categorical and one quantitative

3. A transportation researcher working for a metropolitan mobility office is studying rider behavior to inform future infrastructure planning. The city has recently expanded the number of electric scooters available for public use, and officials want to understand how riders use them during different parts of the week. The researcher is specifically interested in whether riders tend to travel longer distances on weekends than on weekdays, which may help determine where to place charging hubs and where to increase lane capacity. To investigate this, the researcher will analyze GPS-recorded trip distances automatically logged by the city's scooter-sharing system. The researcher plans to obtain a random sample of scooter trips occurring over an entire month, ensuring that both weekday and weekend rides are proportionally represented.

Which of the following best represents an appropriate investigative question for this study?

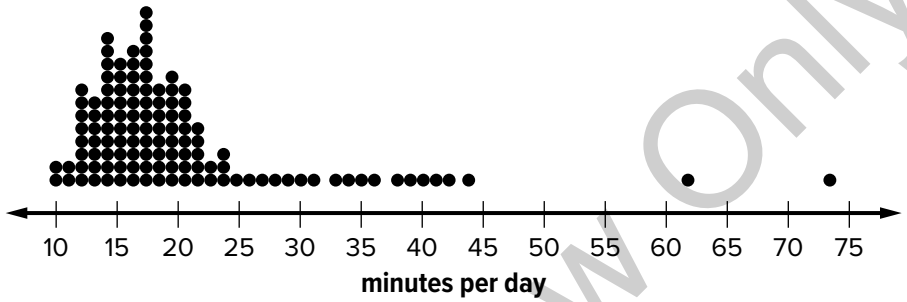
A. How many electric scooters are rented in the city during weekends compared with weekdays?

B. Do electric scooter riders in this city believe that weekend trips are more enjoyable than weekday trips?

C. Is there evidence that the mean trip distance for all electric scooter riders in the city is greater on weekends than on weekdays?

D. Do riders typically travel long distances on electric scooters, and does this vary by neighborhood?

4. A wildlife biologist records the number of minutes that 120 tagged deer spend at a particular watering site each day. The following dotplot is a summary of the data.



Which of the following best describes the distribution of the amount of time deer spend at the watering site each day?

A. The distribution is right skewed with a mean 19 minutes, a range of 70, and no apparent outliers.

B. The distribution is right skewed, centered around 19 minutes, with a small spread and no extreme values.

C. The distribution is right skewed, centered around 19 minutes, with moderate variability, and two possible high outliers at 62 and 74 minutes.

D. The distribution is uniform, with a long right tail caused by random variation, and very little spread in the central 50% of the data.

5. A botanist records the number of new leaves on forty plants after applying a growth treatment to each of the plants. The distribution of the number of leaves is right skewed with several unusually high values and a median of 14. The botanist reports that the mean number of new leaves is 17.8 with a standard deviation of 9.6. The botanist wants to describe the center of the distribution in a way that accurately reflects a "typical" plant's growth under the treatment.

Which measure of center is most appropriate, and why?

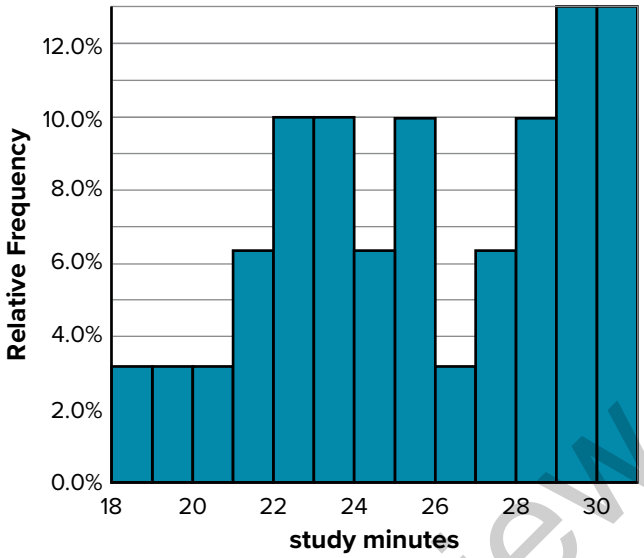
A. The mean; the mean incorporates information from all data values and is therefore always a better summary of the center.

B. The median; the median is resistant to the influence of the unusually high values that pull the distribution to the right.

C. The mean; the difference between the mean and median is small. Skewness does not affect the estimate substantially.

D. The median, because the standard deviation is large, which automatically makes the median larger than the mean.

6. A researcher records the number of minutes 30 students spent studying for a statistics quiz. She constructs both a histogram and a stem-and-leaf plot of the data. The two graphs display the same set of quantitative data. Compare the two graphs.



1	8 9
2	0 1 1 2 2 2 3 3 3 4 4 5 5 5 6 7 7 8 8 8 9 9 9 9
3	0 0 0 0

Study minutes
KEY: 3 | 0 = 30

Which answer provides the clearest explanation for why the data may appear to show different patterns or features, even though they represent the same values?

A. Histograms display the exact data values, while stem-and-leaf plots group values into intervals, which may hide or combine clusters.

B. Stem-and-leaf plots preserve the exact values and natural ordering of the data, whereas the choice of bin widths in histograms may merge nearby values and create the appearance of a single peak.

C. Histograms always show relative frequencies, but stem-and-leaf plots always show raw frequencies, so the two plots will necessarily highlight different distributional features.

D. Stem-and-leaf plots maintain the natural ordering of the values only when the dataset is small, but for larger datasets they may incorrectly reorder observations, causing the displays to differ.

7. The distribution of heights of adult male giraffes in a certain region has a mean of 17 feet and a standard deviation of 2 feet. The distribution of heights of adult female giraffes has a mean of 15 feet and a standard deviation of 1.5 feet. Approximately, what would be the height of a female giraffe with the same standardized score (*z*-score) as a male giraffe with a height of 19 feet.

A. 15.8 feet

B. 16.0 feet

C. 16.5 feet

D. 17.0 feet

8. Connect Mobile is considering a new cellular tower in the town of Springfield. To measure usage the company records the number of text messages sent per day by a sample of 20 people working in the downtown area. Following is the data set.

12	25	34	41
15	27	35	43
18	29	36	45
21	30	38	48
22	32	40	55

The analyst realizes that one of the responses, 55, was incorrectly entered and should have been recorded as 45 instead. What is the change in the interquartile range (IQR) after correcting the error?

A. The IQR decreases by 4

B. The IQR decreases by 10

C. The IQR remains unchanged

D. The IQR increases by 5

9. A nutritionist is conducting a study to understand dietary habits in a small suburban community. She recruits 12 adults of varying ages and backgrounds and asks them to record their total sodium intake from all meals and snacks over a typical day. She finds that the mean sodium intake for the 12 adults is 2,200 mg with a standard deviation of 600 mg.

The nutritionist decides to report the sodium intake in grams instead of milligrams (1000 mg = 1 g). She also needs to adjust for an error in recording by subtracting 500 mg from each person's intake. What are the new mean and standard deviation of the distribution of sodium intake?

A. Mean: 1.7 g Standard Deviation: 0.1 g

B. Mean: 2.2 g Standard Deviation: 0.1 g

C. Mean: 1.7 g Standard Deviation: 0.6 g

D. Mean: 2.2 g Standard Deviation: 0.6 g

10. Customer service agents at a large company have expressed concern regarding inequities in the distribution of call volume among team members. In response, the department manager records the number of calls handled by each of the sixty phone agents and prepares a report for the team that includes a detailed five-number summary.

Minimum	Q1	Median	Q3	Maximum
4	12	18	27	90

The manager discovers that the maximum value (90) was an error. The correct maximum value is 48

Describe how this correction to the maximum value influences the dataset.

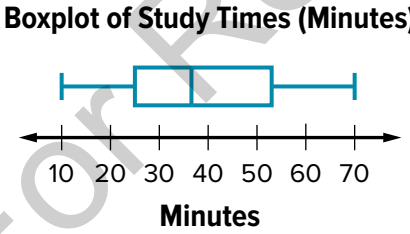
A. Only the maximum value changes; the other four summary values remain the same.

B. The maximum value and Q3 both decrease, but Q1 and the median remain the same.

C. Only the maximum and median change, because changing the largest value always shifts the center of the distribution.

D. All five values could potentially change because the entire distribution must be recomputed when an error is corrected.

11. A teacher records the number of minutes spent by 20 students studying for a statistics exam. The boxplot below summarizes the data. The teacher claims: "At least 25% of students studied for more than 60 minutes." Based on the boxplot, is the teacher's claim justified?



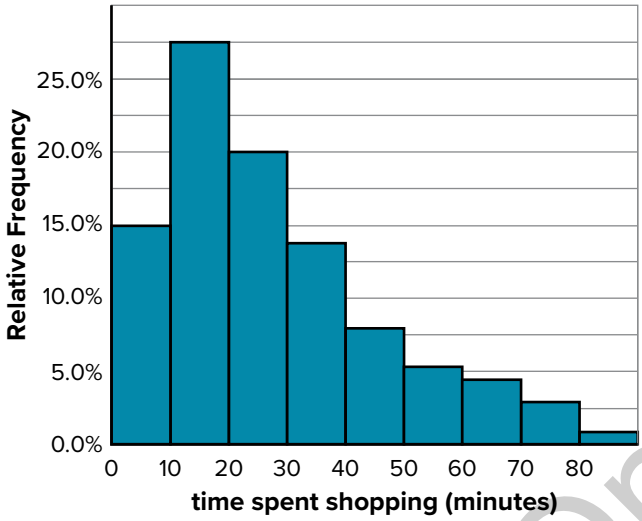
A. Yes, because the upper quartile (Q3) is less than 60 minutes.

B. Yes, because 60 minutes is below the maximum value.

C. No, because fewer than 25% of students studied more than 60 minutes.

D. No, because the median is less than 60 minutes.

12. A local store records the time (in minutes) that customers spend shopping. The histogram below displays the shopping times for a random sample of 500 customers.



Based on the distribution shown, which of the following statements about the mean and median of the distribution is most accurate?

A. The mean is approximately equal to the median because the distribution is roughly symmetric.

B. The mean is less than the median because the distribution is skewed right.

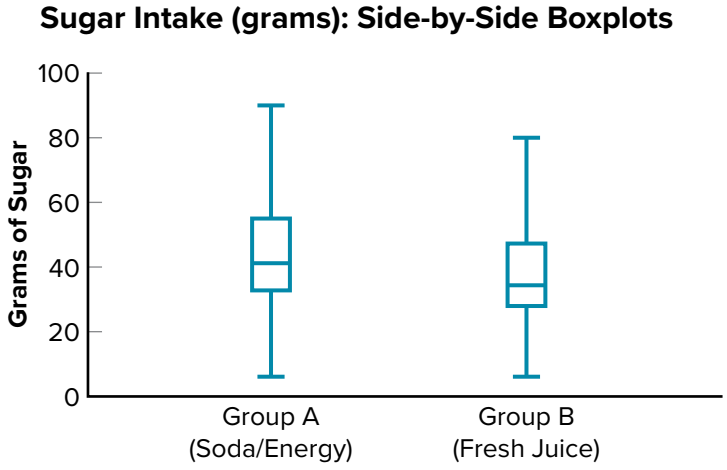
C. The mean is greater than the median because the distribution is skewed right.

D. The median is greater than the mean because the distribution contains several large outliers.

13. An endocrinologist specializing in diabetes is conducting an observational study to examine the effects of different types of dietary sugar on daily sugar intake. She is particularly interested in comparing processed sugar found in soda and energy drinks versus natural sugar found in fresh juice. The study recruits 100 adult patients from her clinic. Each patient is classified into one and only one group based on their self-reported type of drink:

- Group A: Adults who drink a soda or energy drink
- Group B: Adults who drink fresh juice

The boxplots are shown for the two distributions.



Based on the graphical summaries, which of the following is not a supported conclusion?

A. In Group A, the median is higher than Group B's median, suggesting that the typical adult in Group A consumes more sugar than the typical adult in Group B.

B. Group A has both a larger IQR and range than Group B, indicating that individual sugar intakes in Group A have more variability than in Group B.

C. Group A has a higher maximum sugar intake than Group B.

D. For Group B, the maximum value is closer to the median than the minimum is, suggesting that there are some low values pulling the mean below the median, indicating a left-skewed distribution.

14. The research and development team at Rest Easy, a company producing melatonin supplements, wants to determine whether a new sleep aid improves the quality of sleep for adults aged 30–65. The main outcome is hours of REM sleep per night, measured using a validated sleep tracker. Sleep quality can vary due to individual differences such as baseline sleep patterns, stress levels, and caffeine intake.

The researchers propose four possible experimental designs:

I. Randomly assign participants to either the supplement group or a placebo group. Compare the mean REM sleep between groups after one week.

II. Measure each participant's REM sleep for one week without the supplement, then provide the supplement for one week, and compare before-and-after mean REM sleep for each participant.

III. First, assess each participant's stress level using a standardized stress questionnaire and classify individuals into low-, medium-, or high-stress categories. Then, within each stress category, randomly assign participants to either the supplement group or the placebo group. Finally, compare the mean REM sleep between the treatment and control groups within each stress block.

IV. Allow participants to choose whether to take the supplement, then compare REM sleep between groups based on participants' self-selected choices.

Which design is the best choice for this study?

A. Design I; it minimizes confounding by comparing two randomized groups.

B. Design II; using each participant as their own control eliminates all variability from person to person.

C. Design III; stress levels are an important characteristic of the population that could confound the treatment effect.

D. Design IV; allowing participants to self-select ensures that individuals receive a treatment aligned with their preferences.

15. A biologist at Plow and Chow Nursery wants to investigate whether three different fertilizers (A, B, and C) lead to different average growth rates of tomato plants. The scientist has 45 plants, but due to varying availability of fertilizer, they assign: 15 plants to fertilizer A, 20 plants to fertilizer B, and 10 plants to fertilizer C.

The plants differ slightly in size at the start of the experiment, but they come from the same greenhouse and are randomly assigned to treatments.

Which of the following statements is most accurate regarding the design of this experiment?

A. This is not a completely randomized design because the treatment groups have unequal sample sizes.

B. This is a completely randomized design because each plant is randomly assigned to a fertilizer treatment, even though the treatment group sizes differ.

C. This is not a completely randomized design because the initial plant sizes should have been used to form blocks before randomization.

D. This is a completely randomized design only if each fertilizer is assigned to the same number of plants.

16. Professors at a college's music department want to test whether listening to classical music before an exam improves students' performance compared to listening to no music before an exam. They recruit student volunteers from two different first year English classes.

- Class 1 (Honors English): Students are assigned to listen to classical music for 20 minutes before taking a vocabulary test.
- Class 2 (Regular English): Students take the same vocabulary test without listening to any music.

The professors compare the average test scores of the two groups. Which variable most clearly acts as a confounding variable in this experiment?

A. The type of music (classical vs. none)

B. The difficulty level of the vocabulary test

C. The academic level of the class (Honors vs. Regular)

D. The students' preference for listening to music while studying

17. A school district wants to estimate the proportion of high school students who regularly eat breakfast. The district instructs each school to collect data by sending an online survey link to students. The survey is sent to students who have perfect attendance for the semester. Of those contacted, 62% report that they regularly eat breakfast.

Which of the following best explains why the sampling method is biased?

A. The sample size is too small to accurately estimate the true proportion of students who regularly eat breakfast.

B. Students with perfect attendance might be more likely to respond to online surveys, which increases sampling variability.

C. Students with perfect attendance tend to have established morning routines and therefore may be more likely than the general population to regularly eat breakfast, causing the statistic to consistently overestimate the true proportion.

D. The survey includes only high school students and does not include younger students, so the results cannot generalize to the entire district.

18. A university conducted a study to understand the time commitment students devote to independent research projects. To explore patterns in student work habits, researchers randomly selected a sample of 30 undergraduate students and recorded the number of minutes each student spent per week on independent research. Unusually low or high amounts of time spent on their projects could indicate either a lack of engagement or an exceptionally high workload. A summary of the results is shown:

Lowest 3 Observations	Q1	Median	Q3	Highest 3 Observations
0, 5, 10	15	20	35	60, 90, 120

The researcher wants to identify any unusually low or high values relative to the rest of the data. Which of the following statements about potential outliers in this dataset is most accurate?

A. 0, 90, and 120 minutes are all outliers because they are unusually small or large compared to the central bulk of the data.

B. Only 0 minutes is a lower outlier; 90 and 120 minutes are not outliers because the 1.5 IQR method does not classify them as unusually large.

C. 90 and 120 minutes are outliers because they are unusually high relative to the rest of the data, but 0 is not considered unusually low.

D. There are no outliers in this dataset because all values fall within two standard deviations of the mean.

19. A state government is planning a comprehensive study to understand the health habits of all adults age 18 and older living within the state. The goal is to use this information to design targeted public health policies, improve community health programs, and allocate resources effectively. The health department is considering two approaches for data collection:

I. Conducting a survey of a randomly selected sample of 10,000 adults from the population.

II. Conducting a census by collecting health information from every adult in the state.

The department is particularly concerned about capturing accurate and detailed data from small communities, some of which may have unique health behaviors that differ from the state average. They also want to minimize sampling error and ensure that no significant subgroups are underrepresented, while balancing logistical feasibility and cost. Which of the following statements best describes the characteristics and implications of performing a census in this context?

A. A census involves recording information from all adults in the state, providing complete data for the population, but it is time-consuming and expensive compared to a survey, and may be impossible to complete.

B. A census collects information from a randomly selected subset of adults, which reduces costs and allows for easier data collection while still giving accurate population estimates.

C. A census is unnecessary in this case because sampling error can always be eliminated by increasing the sample size in a survey.

D. A census guarantees perfect accuracy and eliminates all potential sources of error, including measurement errors and nonresponse.

20. A pharmaceutical company is testing a new medication intended to reduce blood pressure. The study recruits 120 adults diagnosed with hypertension. The participants in the study are randomly assigned to receive either the new medication or a standard treatment. The researchers want to ensure that participants' expectations about the treatment do not influence the outcomes, but it is logistically difficult and unnecessary to prevent the nurses administering the treatments and recording results from knowing which participants receive which treatment. After six weeks, participants' blood pressure readings are collected and analyzed. Which of the following statements best describes the characteristics and implications of this study design?

A. The study is single-blind because the participants do not know which treatment they receive, but the nurses administering the treatments do. This design helps reduce participant bias but does not prevent potential researcher bias.

B. The study is single-blind because both the participants and the nurses are unaware of which treatment is being administered, minimizing both participant and researcher bias.

C. The study is single-blind because the researchers analyzing the data do not know which participants received which treatment, but participants do know, reducing bias.

D. The study cannot be single-blind because the nurses know which treatment is administered; only a double-blind design is valid for reducing bias in clinical trials.

Free Response Questions

Directions: Show all your work. Indicate clearly which methods you use. You will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

The Library Council of a large city wants to study the reading habits of adults (ages 18+) across the metroplex. The survey asks, among other things, the respondent's average daily reading time for leisure (in minutes). The city is divided into several library districts, and each district may choose how it collects data from its patrons. Three of the districts use the following procedures:

West District: An online survey link is emailed to every adult with an active library card in the district's library system. The message explains that the library is gathering feedback to guide future programming and resource allocation. Participation in the survey is optional, and no incentives are offered. Two weeks after the initial message, a single follow-up reminder is sent to all nonrespondents encouraging them to complete the survey.

East District: Every adult who has participated in at least one library-sponsored reading program within the past two years is identified from the district's program-registration database. These individuals are then sent a mailed survey packet that includes questions about their average daily leisure reading time, preferred genres, and use of library services. A prepaid return envelope is included to encourage completion, and participants are informed that their responses will help the district evaluate and improve future reading programs.

Central District: The library district is organized into six library service zones that are similar in population size and demographics. To gather data on community reading habits, district officials randomly select two of the six zones using a simple random selection process. Every household within the selected zones receives a mailed survey that asks adults to report their average daily leisure reading time, along with several questions about reading preferences and access to library resources. Households are encouraged to return the survey within three weeks using the prepaid envelope provided.

Central District's two selected zones returned the data shown in the given table, representing individual answers to the question: "How many minutes per day do you typically spend reading for leisure?"

Table 1: Summary Statistics for Central District

Central District	n	mean	min	Q1	Med	Q3	Max
Zone 1	2000	23.1875	10.0	17.75	21.5	28.25	45.0
Zone 2	3000	43.625	6.0	22.25	37.0	61.25	120.0

Use the given information to respond to parts A, B, C, D and E. Label any subparts that may be present.

A. **i.** **Identify** the sampling method that West District used.

ii. **Identify** the sampling method that Central District used.

B. **i.** **Determine** the population to which the results of Central District's sample can be generalized. **Justify** your reasoning.

ii. Anton, a member of the planning committee for the East District, suggested using a survey method that sends mailed surveys to all adults who participated in any library-sponsored reading program within the last two years. Anton argues that this method will likely produce a higher response rate than the sampling methods used in West District and Central District and that it provides a sample representative of the entire adult population in East District. **Explain** why Anton's claim is incorrect.

C. Lillith, a publisher who lives in the Central District, wants to compare the average daily leisure reading times for the adult populations in each of the two selected zones of Central District. As a first step, she decides to create a graphical display for each zone using the summary statistics provided in Table 1. **Identify** a graphical representation that is appropriate to use to display the summary statistics and **explain** why the graphical representation is appropriate to use.

D. Use comparisons of the summary statistics in Table 1 to **describe** the most likely shape for the distribution of leisure reading time in Zone 2.

E. Use the summary statistics to **compare** the distributions of average daily reading time for leisure for respondents in Zone 1 and Zone 2 of the Central District.