1 RESEARCH PRINCIPLES

Here's the basis for statistical thinking:

- Statistical Realms
- Rationale for Statistics
- Treatment and Control Groups
- Random Assignment
- Research Questions
- Hypothesis Formulation
- Accept/Reject Hypothesis
- Levels of Measure
- Types of Variables

LEARNING OBJECTIVES

Upon completing this chapter, you will be able to:

1.1 Understand the various domains that utilize statistics

1.2 Understand the rationale for using statistics

1.3 Understand the two types of variables and four levels of measure

1.4 Determine the variable type: *categorical* or *continuous*

1.5 Understand the reasons for having *treatment* and *control* groups

1.6 Comprehend the rationale for random assignment

1.7 Construct research questions

1.8 Formulate hypotheses

1.9 Appropriately accept/reject hypotheses based on statistical outcomes

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

OVERVIEW—RESEARCH PRINCIPLES

This chapter provides an introduction to using and understanding statistics. Consider this analogy: When a chef learns to cook, it involves more than just a tabletop of fresh food. That chef also needs to learn how to use the tools of the trade (e.g., pots, pans, spatulas, thermometers, timers, knives, spoons, stove), know how and when to expertly use each object to get the job done, and understand how food preparation fits into the larger realm of the dining experience. The terms and concepts presented in this chapter will provide the basis for comprehending how statistics fits into the process of conducting scientific research.

As you proceed through this text, observe the diversity of examples and exercises as you process them, demonstrating the multiple broad applicability of the statistics that you'll be learning. A partial list of professions that utilize statistics includes actuary, agriculture, banking, biology, business, census, chemistry, clinical trials, communication, computer science, data science, defense, ecology, economics, education, engineering, epidemiology, finance, forestry, genetics, health care, insurance, law, machine learning, manufacturing, marketing, medicine, meteorology, pharmacology, physics, psychology, public health, research scientist, risk assessment, safety, science writing, social work, sociology, sports, survey science, telecommunications, transportation, urban planning, and zoology.

RATIONALE FOR STATISTICS

While there is a form of statistics known as **single-subject design (SSD)**, which tracks the progress of one person (e.g., "Prior to motivational interviewing intervention, Dusty was walking an average of 5,100 steps per day; now Dusty is walking an average of 8,300 steps per day"), this book focuses on a much more common form of statistics, which involves comprehending the overall characteristics and behaviors of multiple individuals, sometimes arranged in groups. Just as chefs reach for multiple utensils to prepare meals, statisticians have a variety of tests to select from to analyze and comprehend the phenomenon of interest, based on the design and type of variables involved, which is why this book contains more than one chapter:

Chapter 4 covers *descriptive statistics*, sometimes referred to as *summary statistics*, which are the most fundamental of all statistics, making it possible to summarize the data in a variable to a concise result. For example, if we had a list of data that was 1,000 pages, we could use descriptive statistics to answer questions like "What was the average test score?" or "How many left- and right-handed people took this test?"

While this kind of information is useful, Chapters 5 through 9 cover *inferential statistics*, which goes beyond averages and headcounts, providing results designed to answer questions that provide the basis for making more informed decisions.

For example, learning inferential statistics will enable you to ask and confidently answer statistical questions such as "Do children whose parents are teachers do better in school?" "What form of therapy is best for reducing depression: talk therapy, antidepressant medication, or talk therapy combined with antidepressant medication?" "Do elementary students perform

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

better on tests before or after lunch?" "Do people who have more siblings tend to be happier than people from smaller families?" "Do people who took a test prep course pass an exam at the same rate as those who didn't take the course?" Despite the variety of research questions and data that we may gather, statistics enables us to process the data and provide results that are universally accepted and understood, providing a common basis for communicating our findings.

These statistical processes provide the foundation for engaging in *evidence-based practice* (EBP), meaning that we don't just use our intuition to guess what might be an effective intervention—instead, we turn to the numbers to help us make more objective determinations. For example, statistics can help us see if one teaching method outperforms another or which outreach process renders the most people opting for flu shots. To continue the example, suppose a researcher discovers that an instructor has created a prep course that significantly increases the passing rate of a certification exam. If that researcher publishes these results, which would include not only the statistics but also the methods and materials that were used to achieve these positive results, others could read this, and possibly other related manuscripts, and adopt or adapt this method to achieve the same results with others.

Although it may seem paradoxical, one might say that the first step in conducting evidence-based practice research is not to launch a research project but rather to explore the scientific literature to find out if your research question has already been asked and answered by one or more other researchers. If not, then it may be time to design and implement a new research project; however, often researchers and practitioners will find peer-reviewed quality research publications that can be referenced and used as is or adapted to meet their needs. Evidence-based practice sometimes involves creating a hybrid model, selectively combining methods and metrics from more than one effective approach.

Without statistics, it can be difficult to know if well-intentioned interventions are having a positive, negative, or neutral effect. For example, suppose a committee believes that regular aerobic exercise will reduce stress, so the researchers at a workplace gather a group of 20 volunteer participants and provide 30-minute exercise sessions Monday through Friday for 10 weeks. Without some form of statistical analysis, it would be difficult, if not impossible, to be certain if this exercise program was actually effective in reducing stress.

LEVELS OF MEASURE AND TYPES OF VARIABLES

Some might say that half of the battle in statistics is knowing which statistical test to run. Part of what you'll need to know to select the proper statistic involves determining if a variable is a **continuous variable** or a **categorical variable**. There are two types of *continuous* variables (**interval variable** and **ratio variable**) and two types of *categorical* variables (**nominal variable** and **ordinal variable**). We'll proceed with definitions and examples of each of these.

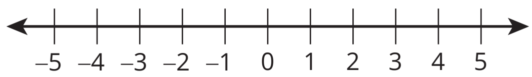
Continuous Variables

Continuous variables are the kinds of variables that you're probably most familiar with. Continuous variables are the numbers that you are accustomed to using when counting or

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

solving a math equation. Imagine a number line where 0 is in the middle; to the left of 0 are all of the negative numbers, –1, –2, –3, and *continue* forever. To the right of 0 are all of the positive numbers, 1, 2, 3, and *continue* forever like numbers on a number line:



Another way to think about this is a *continuous* number is any number that you can enter into a calculator. There are two levels of *continuous* variables: *interval* and *ratio*.

Interval

Interval variables are any number from negative infinity to positive infinity (including all fractions). The numbers are evenly spaced; for example, the distance between 3 and 4 is the same as the distance between 4 and 5. Examples of interval variables include elevation where sea level = 0; anything above sea level is a positive elevation, and anything below sea level is a negative elevation. A bank balance is an interval variable, since a bank balance could be negative (e.g., –$14.83) indicating debt, or positive (e.g., $65,780.99). Temperature is also an interval variable as it can be negative or positive (e.g., –15° . . . 93°).

Ratio

Ratio variables are just like interval variables; however, ratio variables cannot be negative. For example, the number of puppies in a box is a ratio variable, because it cannot be negative (e.g., 5 puppies, 2 puppies, 1 puppy, 0 puppies); there's no such thing as a box that contains –3 puppies. Other examples of ratio variables include the number of minutes that one is engaged in an activity, weight, height, age, IQ score, and heartbeats per minute.

> **Learning tip:** The lowest possible value for a *ratio* variable is 0; notice that the word *ratio* ends in *o*, which resembles the number *0*.

Categorical Variables

Unlike continuous variables that contain numbers, *categorical* variables contain (nonnumeric) lists. The items in these lists can be thought of as categories. For example, *Employment* is a categorical variable that contains two categories: *Employed* and *Unemployed. DishwasherStatus* is a categorical variable that also contains two categories: *Dirty* and *Clean*. A categorical variable can contain more than two categories; for example, the variable *CarColor* may contain eight categories: *Black, Blue, Brown, Gray, Green, Red, Silver, White*. There are two levels of *categorical* variables: *nominal* and *ordinal*.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

Nominal

Nominal variables contain categories that have no particular order. For example, the variable *EyeColor* could contain the following categorical values: *Amber, Blue, Brown, Gray, Green, Hazel, Violet*. Although this list is arranged alphabetically, there's really no inherent order to the variable *EyeColor*. For example, we could have presented the categories in some other order such as *Green, Hazel, Violet, Blue, Amber, Gray, Brown*, or *Brown, Gray, Hazel, Violet, Blue, Amber, Green*. Another example of a nominal variable is *VoteStatus*, where the categories would be *Voted, Did Not Vote*. Alternatively, the categories could be sequenced the other way: *Did Not Vote, Voted*.

> **Learning tip:** The word *nominal* begins with the word *no*, as in *no order*.

Ordinal

Ordinal variables are categorical variables that have an inherent ranked order among the categories. For example, the categorical variable *Meal* has the following categorical values: *Breakfast, Lunch, Dinner*. It would be irrational to arrange the categories arbitrarily, such as *Lunch, Dinner, Breakfast*. The sequence of the values in ordinal variables is typically arranged in ascending order (from first to last or from lowest to highest); for example, for the ordinal variable *Size*, it makes sense to arrange the categories in this order: *Small, Medium, Large, Extra Large*. The ordinal variable *Education* could contain categories in this order: *Did Not Complete High School, High School Diploma, Associate's Degree, Bachelor's Degree, Master's Degree, Doctorate Degree*.

> **Learning tip:** The root of the word *ordinal* is *order*, indicating that there is an inherent *order* to the categories.

Summary of Variable Types

The levels of *categorical* (*nominal* and *ordinal*) and *continuous* (*interval* and *ratio*) variables are included for completeness, but when it comes to selecting, running, and documenting statistics, you'll need to proficiently determine if a variable is either *continuous* or *categorical*, as detailed in Table 1.1.

> **Learning tip:** Consider the acronym *NOIR*: **N**ominal, **O**rdinal, **I**nterval, **R**atio.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

TABLE 1.1 ■ Categorical and Continuous Variables

Type	Level	Example
Categorical	Nominal	Football, Hockey, Baseball
	Ordinal	Breakfast, Lunch, Dinner
Continuous	Interval	. . . −3, −2, −1, 0, 1, 2, 3 . . .
	Ratio	0, 1, 2, 3 . . .

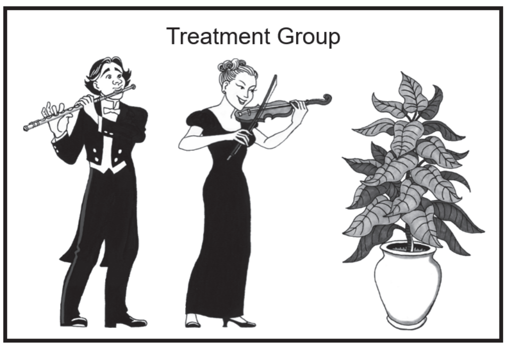
CONTROL AND TREATMENT GROUPS

You've likely heard of *control* and *treatment* groups in scientific settings. In short, the **control group** is given no intervention, and the **treatment group** is given the intervention that is expected to improve some condition. This enables us to compare the performance of the *control group* to that of the *treatment group* to determine how effective the treatment was. Depending on the study design, the *control group* isn't always given nothing; a variation of this design involves providing the members of the *control group* treatment as usual (TAU), and the *treatment group* gets a new intervention. For example, it could be considered unethical to give a placebo (e.g., a sugar pill) to a patient with a serious disease just because that person is in the control group—instead, the researcher could provide them with the drug that's traditionally used to treat their condition, while those in the *treatment group* would be issued the new drug that's under investigation.

Initially, the notion of having a *control group* may seem useless: *What's the point of having a group that we do absolutely nothing to—isn't that a waste? Why not just get one group of people and see if the treatment works on them—after all, isn't that what we want to know?*

Consider this series of five illustrated examples; the first involves only a *treatment group* (no *control group*) to determine if classical music enhances plant growth. You plant a seed in fertile soil and regularly provide the plant with water, sunlight, and 8 hours of classical music per day for a year (Figure 1.1).

FIGURE 1.1 ■ Treatment Group Only (Positive Treatment Effect)



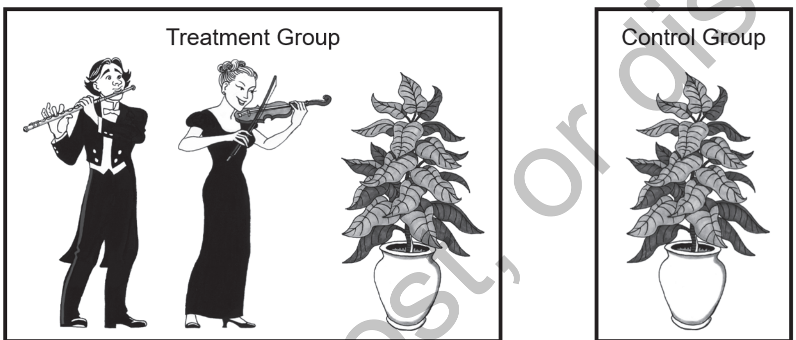
Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

In Figure 1.1, we observe that the plant that got the music produced 20 healthy leaves; hence, we could conclude that classical music has a positive effect on plant growth, but the results of this one-group (treatment-only) design may be questionable. Specifically, one may reason: *That plant got sunlight, good soil, water, and music, so how can we be sure that it was the music that actually contributed to the plant's growth? Maybe it would have grown up fine with just the sunlight, soil, water, and no music.*

To address this, we could implement a more robust model using a two-group design consisting of a *treatment group* and a *control group* (Figure 1.2).

FIGURE 1.2 ■ Treatment Group and Control Group (Neutral Treatment Effect)



In Figure 1.2, we see a two-group design involving a *treatment group* and a *control group*. The plant in the *treatment group* will be in Room 201, where the planted seed will get quality soil, sunlight, water, and daily music. The plant in the *control group* will get the same seed, soil, lighting, watering, and even the planter as the plant in the treatment group, but the plant in the control group will be placed in Room 222, which is far down the hall where the music cannot be heard. One year after planting, we observe that the plant in the *treatment group* appears to have grown identically to the plant in the *control group*—both have 20 healthy leaves; since the plant that got no music did as well as the plant that did get music, based on this comparison, it appears that the music added nothing to the plant's growth, and hence, we might conclude that the music had a *neutral effect* on plant growth.

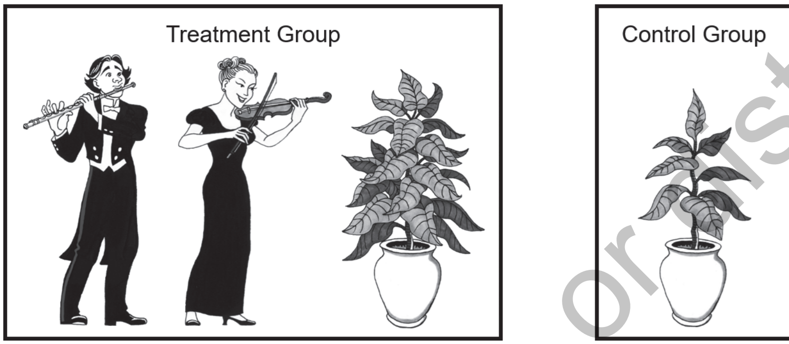
Before concluding this example, refer back to Figure 1.1, which involved only the *treatment group*; in that case, we concluded that the plant growth was positive based on the 20 healthy leaves, which may be attributable to the music. The two-group design in Figure 1.2 provides us with more information; notice that the plant in the *control group*, which got no music, also has 20 healthy leaves (just like in Figure 1.1). Further, when we compare the plant in the *treatment group* to the plant in the *control group*, we change our opinion on the meaning of those 20 healthy leaves—since the plant that got no music in the control group grew equally to the plant that got music in the treatment group, we can no longer plausibly claim that the music had a positive effect on the plant's growth; it seems that this is just normal growth for this plant.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

Figure 1.3 illustrates another two-group design, which tells a different story: The plant in the *treatment group* that was exposed to music grew considerably more compared to the plant in the *control group*. Since the only thing that's different between these two groups is that the *treatment group* got music and the *control group* did not, these results suggest that music had a *positive effect* on the plant's growth.

FIGURE 1.3 ■ Treatment Group and Control Group (Positive Treatment Effect)



In this example, Figure 1.4 illustrates that the music had a *negative effect* on plant growth; clearly, the plant in the *control group* grew better without the music compared to the plant in the *treatment group*. Had there only been the *treatment group* (and no *control group*), we might conclude that the music helped the plant produce eight healthy leaves. Statistically comparing the *treatment group* to the *control group* provides a stronger understanding of the effectiveness (or ineffectiveness) of various treatments.

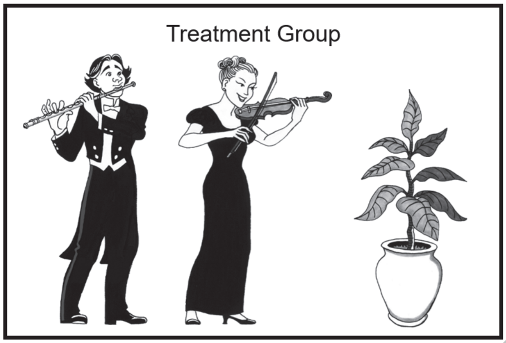
FIGURE 1.4 ■ Treatment Group and Control Group (Negative Treatment Effect)

To finalize this example, observe that Figure 1.5, depicting a plant that produced eight leaves, with no control group to compare it to, might (mis)lead us to conclude that eight leaves

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

Do not copy, post, or distribute

FIGURE 1.5 ■ Treatment Group Only (Positive Treatment Effect)



are as much as we can expect from this plant, even with the music. The consequence of having a *treatment group* only produces the same conundrum as we saw in Figure 1.1: Without the *control group* for comparison, we might mistakenly conclude that the music had a positive effect on this plant's growth when, in fact, it may have had no effect or even a negative effect.

For visual clarity, this series of examples has involved one plant in each group, but to further solidify the results of a study like this, it would be good practice to have more than one plant in each group, just in case we might have unknowingly used a better seed for one plant than the other—that would be an unfair comparison. To get more robust statistical results, we could set up 30 plants in each group, and then after a year, we could calculate the average number of leaves per plant in the *treatment group* and compare that to the average number of leaves per plant in the *control group* to determine if one group significantly outperformed the other; this will be covered in Chapter 5: *t* Test and Welch Two-Sample *t* Test.

RANDOM ASSIGNMENT

Continuing with our plant example, suppose you're at the beginning of the process; you have enough planters, soil, and seeds to prepare 60 plants—30 for the control group and 30 for the treatment group. The seeds are provided in a clear container, and hence the seeds near the top and sides have been exposed to more light than the rest of the seeds. You wonder if this might have an effect on the plant's growth, so you confer with a botanist, who confirms that when this type of seed is exposed to light prior to planting, it slightly enhances growth. Since the seeds can shift around in the bag during shipping, making it impossible to know precisely how much light each seed has been exposed to, the best strategy is to randomize the seeds. This will help to reduce the bias among the groups—if we mix up the seeds prior to planting and assigning plants to (control/treatment) groups, then the light-exposed seeds have an equal chance of being assigned to the *control group* and the *treatment group*. If we have a sufficient sample size, **random assignment** should evenly distribute low- and high-exposed seeds among the groups; hence, prior to beginning any treatment(s), random assignment to the groups helps neutralize bias. The rationale and methods for randomization are covered in further detail in Chapter 2: Sampling.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

RESEARCH QUESTION AND HYPOTHESIS FORMULATION

A hypothesis is a provisional statement or statements that rationally address a research question prior to gathering and analyzing data. For example, suppose we want to research if *working with a tutor improves academic performance*. There are a variety of ways to form the research question. For clarity, we'll be using a unified approach throughout this book, wherein the research question will be phrased as a *yes/no* question: *Does working with a tutor improve academic performance?*

You can think of the hypotheses as the possible *answers* to the research question. Since our research question has two possible outcomes, *yes* or *no*, we can write the two hypotheses, anticipating the statistical results that will emerge: one for the *yes* answer and another for the *no* answer. First, we'll focus on the *no* answer:

H_0: The Null Hypothesis

Realistically, we must recognize that not all interventions are going to be effective; sometimes, an intervention may fail to have the effect that we hoped for. In this case, our research question is: *Does working with a tutor improve academic performance?* Hence, we first write the hypothesis that corresponds to the *no* answer: *Tutoring has no effect on academic performance.* This is the ***null hypothesis*** **(H_0)**, suggesting that the treatment had a *null* effect—tutoring had *no* effect on the student's academic performance. Essentially, we're saying that the intervention was ineffective.

This would be documented as follows:

H_0: Tutoring has no effect on academic performance.

H_1: The Alternate Hypothesis

Next, we consider the *yes* answer: *Tutoring has an effect on academic performance.* This is the **alternate hypothesis (H_1)**, meaning that this is the *alternative* to the null hypothesis—*tutoring had an effect on the student's academic performance.*

This would be documented as follows:

H_1: Tutoring has an effect on academic performance.

The research question and hypotheses could be presented as such:

Research question: Does working with a tutor improve academic performance?

H_0: Tutoring has no effect on academic performance.

H_1: Tutoring has an effect on academic performance.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

Notice that when we phrase the research question as a *yes/no* question, there's only a one-word difference between the null hypothesis (H_0) and the alternate hypothesis (H_1):

*H_0: Tutoring has **no** effect on academic performance.*

*H_1: Tutoring has **an** effect on academic performance.*

Hypothesis Resolution

After gathering the data and running the appropriate statistical analysis, you'll then refer back to the hypotheses (H_0 and H_1) and select the hypothesis that corresponds to the statistical results. The inferential statistics covered in Chapters 5 through 9 include guidance for resolving the hypothesis. In other words, the statistical results will give you what you'll need to know for **hypothesis resolution**: identifying which hypothesis (H_0 or H_1) corresponds to the statistical results. From there, you can concisely document your findings, citing selected numbers that R will provide in the statistical results.

ASKING AND ANSWERING RESEARCH QUESTIONS

A statistician colleague of mine once said, "I want the numbers to tell me a story." I've never heard the mission of statistics expressed so elegantly. One way to think about statistics is to conceive that the numbers in a dataset contain a story; the answers to your questions are hidden within the data, and we can use statistics to filter the data, coaxing it to reveal its secrets. Admittedly, that makes things sound a bit mystical, but in a way, that's precisely what happens when we process statistics—the formulas systematically organize and process the data to produce a concise set of results that help us understand what we could not see just by gazing at the data table.

Chapters 4 through 9 are structured to answer different types of statistical questions depending on the design and type(s) of variables involved.

Chapter 4: Descriptive Statistics reduces a long list of data into a short list of figures, enabling you to comprehend and communicate the contents of any variable. For example, you could answer questions like:

- *How many left- and right-handed people are in our group?*
- *What's the average score?*
- *What's the youngest and oldest ages in a workplace?*
- *What percentage of our staff is part-time and full-time?*
- *What's the average number of classes students are enrolled in?*
- *How many people in this group voted in the last election?*

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

Chapter 5: *t* Test and Welch Two-Sample *t* Test compares two conditions/treatments to determine if one outperformed the other or if they both performed about the same. These tests can statistically answer questions like:

- *Do tutored students have the same test scores as nontutored students?*
- *Which drug is best for reducing hypertension (high blood pressure): Drug A or Drug B?*
- *Do people take the same amount of time to eat one scoop of chocolate ice cream compared to one scoop of vanilla ice cream?*

Chapter 6: ANOVA—Tukey Test and Wilcoxon Multiple Pairwise Comparisons Test is very similar to Chapter 5: *t* Test and Welch Two-Sample *t* Test, except instead of being limited to comparing only *two groups* to each other, the statistics in this chapter can process comparisons involving *three or more groups* (notice that these examples are the same as those in Chapter 5, but with additional groups):

- *Do tutored students, nontutored students, students who use a learning app, and students who engage in study groups have the same test scores?*
- *Which drug is best for reducing hypertension (high blood pressure): Drug A, Drug B, or Drug C?*
- *Do people take the same amount of time to eat one scoop of chocolate ice cream, one scoop of vanilla ice cream, and one scoop of strawberry ice cream?*

Chapter 7: Paired *t* Test and Paired Wilcoxon Test can assess the effectiveness of a treatment wherein one group of people is pretested, then given some intervention, and then each person is given a posttest (which uses the same measurement as the pretest). If the posttest score shows a significant improvement compared to the pretest score, this suggests that the treatment was effective. This statistic can answer a variety of questions involving a single group of people:

- *To assess the effectiveness of Quick Coaching, a coach unobtrusively observes an archer shooting 10 arrows at a target and records the score (center ring = 10 . . . outermost ring = 1, misses the target = 0), and then the coach provides improvement recommendations, after which, the archer shoots 10 more arrows. The coach compares the scores before and after giving the feedback between rounds.*
- *To determine if food has an effect on mood, we ask each participant to score their mood (1 = bad mood . . . 5 = good mood), and then we provide a slice of really good chocolate cake; when they're done with the cake, we ask them to report their current mood score using the same 1 to 5 scale. Finally, we compare the mood scores before and after eating the cake.*

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

- *To find out if music helps people relax, we first take the person's pulse rate, then we play 10 minutes of soft instrumental music, and then we take the person's pulse rate again. We can then compare the first pulse rate to the second pulse rate to discover if the pulse rate changes significantly after listening to the music.*

Chapter 8: Correlation—Pearson Test and Spearman Test can be used in a single-group design to assess the relationship between two continuous variables. This statistic can answer questions like:

- *Is there a correlation between income and happiness?*
- *Is there a correlation between age and hours of sleep per night?*
- *Do students who spend more time completing a test tend to have higher scores or do students who complete the test quicker tend to have higher scores?*

Chapter 9: Chi-Square Test is used to determine if there's an association between two categorical variables. This statistic can be used to answer questions like:

- *Do all political parties have the same proportion of members in favor of voting by mail?*
- *Do Collies, Keeshonds, and Cocker Spaniels have the same preference when it comes to selecting a bone or a ball?*
- *Do the same percentage of degreed and nondegreed individuals pass a CPR certification class?*

These examples demonstrate the diverse types of research questions that can be asked and answered using a fairly concise set of statistical analyses.

GOOD COMMON SENSE

As you compute multiple statistics throughout this text, be aware that despite our diligence for precision, statistics do not *prove* or *disprove* anything; rather, statistics are generally used to help us to understand the characteristics of a group and to reduce uncertainty. Statistical results are typically written tentatively (e.g., *The results suggest that this intervention was effective in treating this problem.*) as opposed to definitively (e.g., *The results prove that this intervention was effective in treating this problem.*).

Also, recognize that statistical results reflect the characteristics or performance of the *group(s)* of data (people) in our dataset and do not provide insights regarding any particular *individual*. For example, suppose we analyze the ages of five people who are 8, 12, 22, 53, and 60; the average age is 31, but it would be wrong to think that you could arbitrarily point to any one person in that group and confidently proclaim, "You are 31 years old." Although the average is 31, it's possible that no individual in the group is 31, and in this case, none of

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

these people are in their 30s. Further, if we asked a group of people, "How many children did your parents have?" and the average turns out to be 2.3, it would be foolish (and comically gruesome) to think that all of these families have 2.3 children. In summary, statistical results pertain to the overall *group* and are not intended to plausibly characterize or make predictions about any specific *individual(s)*.

KEY CONCEPTS

- Rationale and uses for statistics
- Level of data (continuous: interval, ratio; categorical: nominal, ordinal)
- Types of data (continuous, categorical)
- Treatment group
- Control group
- Random assignment
- Research question
- Hypotheses (H_0: null, H_1: alternate)
- Accepting/rejecting hypotheses

PRACTICE EXERCISES

For the following exercises, describe the basis for an experiment that would render data that could be processed statistically.

Exercise 1.1

The director of a healing center wants to determine if 30 minutes of guided meditation affects the resting pulse rate.

a. Indicate the groups (e.g., control, treatment).

b. State the research question.

c. State the hypotheses (H_0 and H_1).

Exercise 1.2

A pediatrician wants to find out if having a magician visit and perform for hospitalized children at their bedside has an effect on their anxiety.

a. Indicate the groups (e.g., control, treatment).

b. State the research question.

c. State the hypotheses (H_0 and H_1).

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

Exercise 1.3

A physical education instructor wants to know if seventh-graders and eighth-graders on a team are about the same heights.

a. Indicate the groups (e.g., control, treatment).

b. State the research question.

c. State the hypotheses (H_0 and H_1).

Exercise 1.4

A professor wants to determine which version of a course produces the best grades: in-person classroom or live online.

a. Indicate the groups (e.g., control, treatment).

b. State the research question.

c. State the hypotheses (H_0 and H_1).

Exercise 1.5

An office manager who oversees two sites wants to determine if providing a free lunch on Friday to employees who are on time for the entire week affects lateness.

a. Indicate the groups (e.g., control, treatment).

b. State the research question.

c. State the hypotheses (H_0 and H_1).

Exercise 1.6

A school principal wants to find out if providing music education has an effect on math scores.

a. Indicate the groups (e.g., control, treatment).

b. State the research question.

c. State the hypotheses (H_0 and H_1).

Exercise 1.7

A nurse wants to find out if aromatherapy helps to reduce patient stress.

a. Indicate the groups (e.g., control, treatment).

b. State the research question.

c. State the hypotheses (H_0 and H_1).

Exercise 1.8

A dairy farmer wants to find out if classical music affects how much milk cows produce.

a. Indicate the groups (e.g., control, treatment).

b. State the research question.

c. State the hypotheses (H_0 and H_1).

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

Exercise 1.9

A playground counselor wants to find out who's better at jumping rope: children enrolled in public school or children enrolled in private school.

a. Indicate the groups (e.g., control, treatment).

b. State the research question.

c. State the hypotheses (H_0 and H_1).

Exercise 1.10

A manager of a customer support call center that spans two separate rooms wants to see if running classic cartoons on a big screen with the sound muted affects employee morale.

a. Indicate the groups (e.g., control, treatment).

b. State the research question.

c. State the hypotheses (H_0 and H_1).

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

2 SAMPLING

You don't need all the data to run quality statistics; just get a sample.

- Rationale for Sampling
- Sampling Terminology
- Representative Sample
- Probability Sampling
- Nonprobability Sampling
- Sampling Bias
- Optimal Sample Size

LEARNING OBJECTIVES

Upon completing this chapter, you will be able to:

2.1 Comprehend the rationale and advantages of sampling: time, cost, feasibility, and extrapolation

2.2 Understand three-tier sampling terminology: population, sample frame, and sample

2.3 Derive a representative sample to reduce threats to external validity

2.4 Understand the utility of probability sampling options

2.5 Understand the utility and constraints of nonprobability sampling options

2.6 Identify factors that could potentially bias a sample and techniques to reduce such bias

2.7 Consider an appropriate sample size

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

OVERVIEW—SAMPLING

Think about a simple statistic that you already know, like the average, where you total the numbers and then divide by the amount of numbers that you have. For example, suppose you had the following test scores: 76, 99, and 86. The total is 261. Since there are three numbers, we'd then divide that total by 3: 261 ÷ 3 = 87; hence, we can say that the average is 87. It was fairly simple to process these data, consisting of three figures to derive the average, but what if there were no data? Clearly, this would mean that we'd have nothing to process; without data, there's nothing to average, so the question naturally emerges: *Where do all these data come from?* In cases where there's a relatively small and manageable group of data (e.g*., What's the average test score from a single class?*), it's possible to collect and process the data on everyone. Other times, the group that we're interested in is so large that it's simply not feasible to gather all of the data to answer our statistical question (e.g., *What's the average income of everyone living in a city?*). If you don't have access to the income data for every person in that city in an existing file, the **population** of that city may be too big for you or your research team to gather data from everyone so that you could calculate the average income. In this case, feasibility is an issue; instead of trying to gather data on *everyone*, we can gather a **sample** and then compute our statistics (average) based on that sample. This process will give us some insights that we didn't have before regarding the average income of people in that city. If we do things right, the data gathered from this *sample* will admittedly not include data from everyone in the city, but it can provide a viable estimate that we can work with. If we use an appropriate sampling method to collect income data from considerably less than half of the people, we'll know more about the income level of those living in that city than we did before acquiring the *sample*.

You're likely already familiar with the notion of **sampling**. Perhaps you've tasted a sample of cheese at a store, and based on that small amount, you made a decision to buy or not buy some quantity of it. At some point, you've likely had a blood sample taken consisting of a small amount of blood, which can be sent to a lab for analysis. The findings from that small sample can be used to understand the status of the rest of your blood (which was not sent to the lab); further, that small amount of blood can suggest a diagnosis and treatment strategy. You probably noticed that they didn't draw your entire blood supply and send that to the lab for analysis—that would be painful, time-consuming, expensive, and lethal. The notion is that often in statistics, acquiring *all* of the data is not always a reasonable or necessary goal. In statistics, we often make due with a sample, and in selected cases, we can use what we learn from the analysis of the *sample* to guide us in understanding the larger *population* that it was drawn from; this is known as **external validity**—taking what you've statistically learned from the *sample* and plausibly generalizing it to the larger *population* that it was drawn from. For example, if we've conducted our *sample* properly, the average income, derived from a small sample, may reasonably represent the average income of all of those employed in that community.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

SAMPLING RATIONALE

The process of relying on data from a *sample* rather than the entire *population* is rooted in the practical constraints of collecting quantities of data in the real world, specifically: time, cost, feasibility, availability, and extrapolation.

Time

Time is a potent resource. We cannot change the pace of time, and nobody can store time to use it later. Nor can we manufacture more time; all we can do is use it. Some processes are serial, such as conducting research, wherein we (1) derive a research question and corresponding hypotheses, (2) gather pertinent data, (3) compute statistical analyses on the data, (4) document the results, and (5) derive a plan of action based on the results of the data. Considering that these processes need to be carried out in order, if Step 2 involves gathering data from a sizable population (e.g., 500,000 people), this could substantially slow the process, whereas strategically gathering data from a sample would shorten the data collection process, enabling us to gather data and act upon the findings more rapidly. Further, some decisions are time dependent; you or others may need to make a decision within a specified time frame (e.g., by the end of the week). Hence, there may simply not be enough time to gather a vast amount of data, particularly if the data-gathering method involves administering lengthy interviews.

Cost

Aside from setup costs (e.g., developing a survey, acquiring computers, training staff, online survey subscription fee, participant recruitment advertising expenses), there may be costs associated with gathering data. This could involve paying trained interviewers to confer with participants, participant compensation fees, photocopying questionnaires, postage, data entry, and other per-participant expenses. Budgetary constraints often necessitate *sampling* as opposed to attempting to gather data from the larger *population*.

Feasibility

The process of gathering data has challenges. Referring back to the example of the blood sample, it'd be deadly to draw a person's entire blood volume for testing, plus it's hard to conceive of a laboratory equipped to conduct such testing. Consider another example wherein a school board wants to have the parents of each student in a school district complete an online survey. Some households may not have access to a computer with online capabilities. Also, if the survey is not available in multiple languages, this may create language barriers. Those who do have online access may opt to complete only part(s) of the survey, while others may simply choose not to respond. If the research involves in-person encounters, geographical distance may preclude remote regions.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

Extrapolation

Gathering data meaningfully can provide a sufficient representation of the population, enabling us to analyze the *sample* that we acquire and reasonably assert that we now have a better understanding of the *population* that it was drawn from—what we now know about our small *sample* reasonably pertains to the larger *population* as well. Plausibly generalizing the results of our *sample* to characterize the larger *population* is referred to as *external validity*. Other times, the goal is not to extrapolate our findings from the *sample* to comprehend the larger *population*—sometimes the goal is merely to understand those involved in our *sample* or others who are similar to the unique characteristics of those in the *sample* that we analyzed; this will be covered in the section on *nonprobability sampling*.

SAMPLING TERMINOLOGY

The process of deriving a sample involves a three-step process, wherein we move from the largest realm (the *population*) to the smallest (the *sample*).

Population

The *population* refers to every person or data item in a specific domain. In terms of people, the *population* of a country would include *all* of the people living in that country. On a smaller scale, the *population* of an organization would be *all* of the members of that organization. Regarding data items, the population would refer to *all* of the academic records for every student at a school or *all* of the medical records of patients at a hospital. You've probably surmised that when it comes to a *population*, the operative word is *all*; hence, the population is considered the largest potential pool of data in the data collection process. For example, suppose 20,000 students are enrolled in a college; the *population* is 20,000.

Sample Frame

The **sample frame** is a step down from the *population*. Whereas the *population* entails *all* of the people or data records within a specified domain, the *sample frame* is a *subset* of the *population* that you could potentially access. For example, suppose you want to email a survey to the students enrolled at a college, and you locate a website that includes the email addresses of the students, but students can freely configure their online profile: They can opt to have their email listed on this public website or not. Suppose 8,000 of the 20,000 students in the college *population* opt to have their email address posted on the website—this subset of 8,000 students that you could reach out to is the *sample frame*.

Sample

Whereas the *sample frame* constitutes the list of people (or data) that you *could* draw from, the *sample frame* may be too large. Hence, we take one final step down and select a subset from

Do not copy, post, or distribute

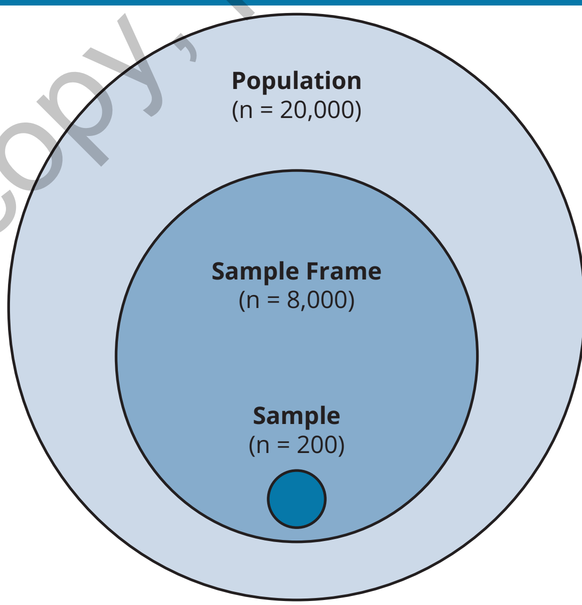
Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

the *sample frame*; this constitutes the actual *sample* that we will work with. Continuing with our university example, the *sample frame* consists of a pool of 8,000 students that we could potentially contact. Suppose this research involves conducting a 30-minute individual interview with each participant, and we provide $10 compensation for their time. If you gathered data on all 8,000 students in the *sample frame*, this would require a budget of $80,000 just for participant fees and a total of 4,000 hours of interviewing time. If you conducted back-to-back interviews for 8 hours per day, it would take 500 days (more than a year) to gather the data. Such a sizable project may not be plausible in terms of time, cost, and feasibility. Alternatively, you could select a *sample* of 200 students from the *sample frame* of 8,000. This would reduce the participant compensation fee to $2,000 and the data collection time to a total of 100 hours (12.5 days). These 200 students would constitute the *sample*.

The *population, sample frame*, and *sample* involve a system of subsets (see Figure 2.1):

- The *population* includes all people or data records in a specified domain. In this example, the *population* is the 20,000 students who are enrolled at a college.
- The *sample frame* is the portion of the *population* that you could potentially access. In this example, the *sample frame* is 8,000, which is 40% of the population (8,000 ÷ 20,000).
- The *sample* is the portion of the *sample frame* that you include in your research. In this example, the *sample* is 200, which is 2.5% of the *sample frame* (200 ÷ 8,000) and only 1% (200 ÷ 20,000) of the *population*.

FIGURE 2.1 ■ Three Sampling Tiers: Population, Sample Frame, and Sample



Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

REPRESENTATIVE SAMPLE

You've probably already encountered a **representative sample**. In an ice cream parlor, you may have requested a sample of a flavor that seemed promising. You'd then receive a very small amount of the ice cream, and after you taste it, you'd decide if you wanted more of it. You likely presumed that the small amount that you tasted was *a representative sample*, meaning that you'd reasonably expect that the rest of the ice cream in the big container was the same flavor, color, texture, and temperature as the small sample that you tasted.

Often, researchers make a deliberate effort to gather a *representative sample*, which can facilitate *external validity*, meaning that what we (statistically) learn about the *sample* is reasonably generalizable to the larger *population* that the sample was drawn from. *Probability sampling* methods have the potential for *external validity*. Alternatively, *nonprobability sampling* enables the researcher to focus only on those in the sample, with no intention to generalize those findings to the larger population, meaning that there would be no potential for *external validity*. We'll look at design options for *probability sampling* and *nonprobability sampling*.

PROBABILITY SAMPLING

Probability sampling is an equal-opportunity sampling technique, wherein every person or data record has an equal chance of being selected for the *sample*. If a *representative sample* is gathered, the results have the potential for *external validity*, meaning that what we know about those in the *sample* can plausibly be generalized to comprehend the larger population that the *sample* was drawn from. We'll explore methods for probability sampling: *simple random sampling*, *systematic sampling, proportionate and disproportionate stratified sampling*, and *area sampling*.

Simple Random Sampling

The most fundamental form of sampling is **simple random sampling**, sometimes referred to as *SRS*. This method involves randomly selecting people or data records. For a small group, we could use a low-tech random selection method (e.g., coin flip, drawing names or numbers out of a hat); for larger groups, it would likely be more efficient to use a program with a random-number generator to guide your selections. For example, suppose we had a sample frame of 60 people to select from and we wanted to randomly choose 10 for our sample. We could assign a number (1 . . . 60) to each person and then use one of these methods to gather 10 numbers that would constitute our sample: 32, 20, 55, 12, 14, 51, 19, 25, 17, and 43 (Figure 2.2). If anyone in the selected sample opts out, then we could use the same random selection process to replace them.

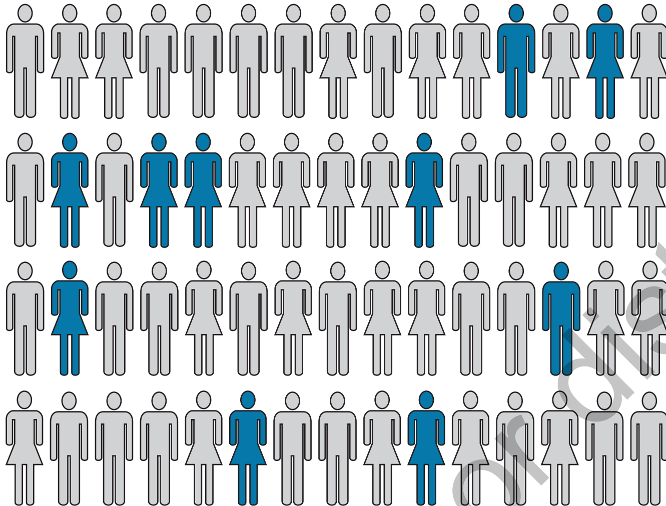
Stratified Sampling

In the prior example, the goal was simply to randomly select 10 people from a group of 60, which produced 8 women and 2 men, but suppose you needed that sample to be

Do not copy, post, or distribute

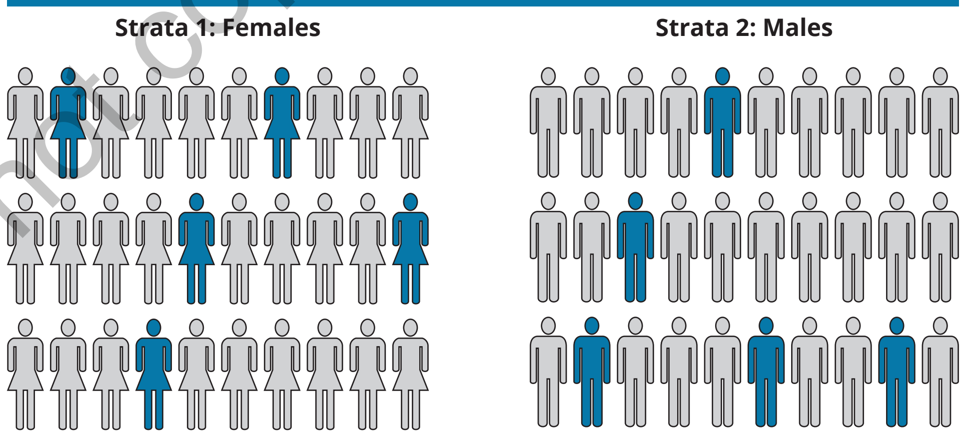
Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

FIGURE 2.2 ■ Simple Random Sampling: The Researcher Randomly Selects 10 of the 60 People



gender-balanced (5 women and 5 men); you could use a **stratified sampling** method. Instead of everyone in the *sample frame* being in one group and then making random selections, we could *stratify* this group based on *gender*, meaning that we would create two *strata* (think of a *stratum* as a *subset*): one stratum for the *women* and another for the *men*. Since our goal is to gather a sample of 10, we could then randomly select 5 from the women *stratum* and 5 from the male *stratum* (Figure 2.3).

FIGURE 2.3 ■ Stratified Sampling: Split the Sample Frame Into Two (or More) Strata, Then Randomly Select Five From Each Strata



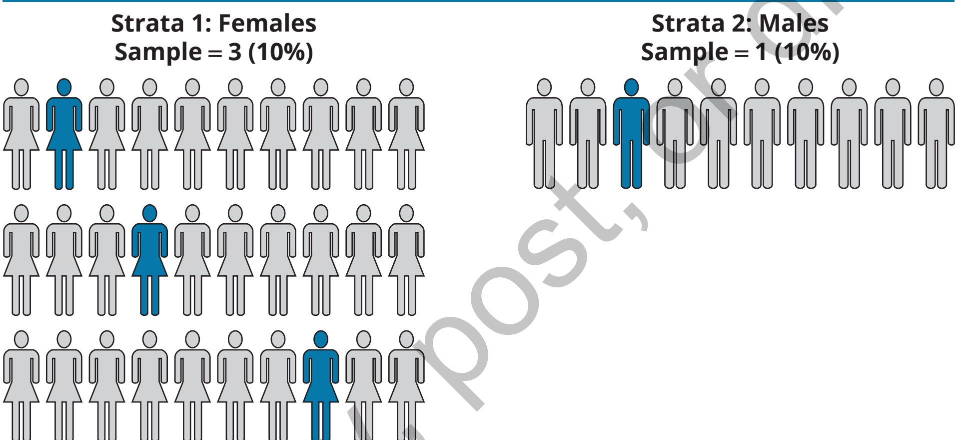
Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

Proportionate and Disproportionate Stratified Sampling

When working with *stratified sampling*, you may opt for **proportionate stratified sampling** or **disproportionate stratified sampling**. Consider a different distribution that has a total of 40 people: 30 in the women stratum and 10 in the men stratum. We could gather a *proportionate* sample, wherein we would select 10% from each stratum, meaning that we'd take the same *proportion* (10%) from each group. The women stratum has 30 people; 10% of 30 is 3, and hence, we'd randomly select 3 people from the women stratum. Next, we'd gather the same proportion from the men stratum, which contains 10 men; 10% of 10 is 1, and hence, we'd randomly select 1 person from the men stratum (Figure 2.4).

FIGURE 2.4 ■ Using Proportionate Stratified Sampling to Gather a 10% Sample (From Each Stratum) Renders Three Women and One Man



In the prior example, notice that *stratified* sampling produced only one man (10% of 10 = 1). A technique for increasing the sample size when faced with one or more small strata is to use *disproportionate* sampling, where instead of choosing a proportion (percentage) to select from each stratum, we choose a set number of individuals (or data records) to select from each stratum. For example, instead of randomly selecting 10% from each stratum, we may specify that we'll randomly select three from each stratum (Figure 2.5).

Although the number of people selected from each stratum is now the same (three women and three men), this is *disproportionate sampling* since the proportion (percentage) selected from each group is different. Randomly selecting 3 out of the 30 women constitutes a 10% sample from that stratum, whereas randomly selecting 3 out of the 10 men is a 30% sample from that stratum.

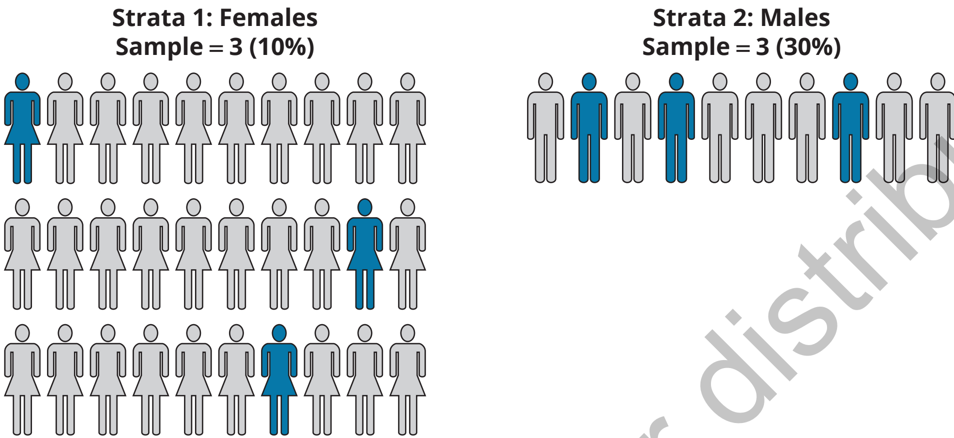
Systematic Sampling

You may have noticed that when using random sampling, it's possible that the selections may involve people (or data records) that are near each other or far apart. Sometimes you may want

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

FIGURE 2.5 ■ Using Disproportionate Stratified Sampling to Gather Three (From Each Stratum) Renders Three (10%) Women and Three (30%) Men



to gather data in a random fashion, but you'd like the selections to be evenly spaced; **systematic sampling** provides a method for doing that. For example, suppose you'd like to select 15 of the 60 people who will be attending a meeting. First, divide the number of people in the *sample frame* (60) by the *target sample size* (15); this will produce the "*k*" or *skip term* (k = 60 ÷ 15, which equals 4). Next, identify the start point, which will be a random number between 1 and k (in this case, k = 4); suppose that the random start point is 3. This means that the first selection for your sample would be the 3rd person who enters the meeting, along with every 4th person after that, so the 7th person who enters the meeting would be the second person in your sample, then the 11th person, and so on. Based on this method, you'd get an even sample of the meeting attendees in terms of their arrival times—early, on time, and late. This method can also be used with lists of data records (Figure 2.6).

Area Sampling

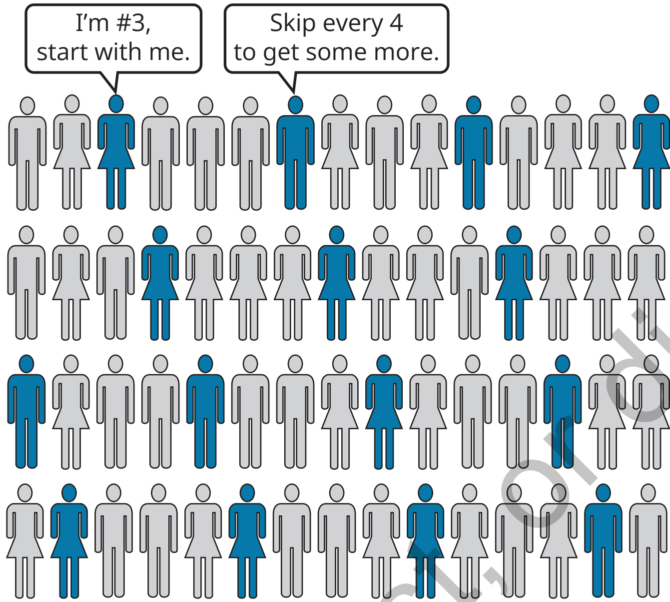
Area sampling, also known as **cluster sampling** or **multistage cluster sampling**, is useful for gathering data spanning a geographical region or when it's not possible to attain a *sample frame*. Considering that the characteristics of neighborhoods and residents can vary, *area sampling* provides a method for gathering a (more) representative sample of the area of interest. In this example, you want to gather data from 30 residents of Smalltown, consisting of 15 blocks. First, we'd number the blocks (Block 1, Block 2, Block 3 . . . Block 15). Since our goal is to survey 30 households, and there are 15 blocks, simple arithmetic tells us that we should randomly select and survey 2 households per block (30 samples ÷ 15 blocks = 2 samples per block) (Figure 2.7).

This example presumes that there are an equal number of residences and people living on each block. In reality, there may be substantial differences; for example, suppose this 15-block neighborhood consists primarily of homes that are all about the same size, except Blocks 6, 7, 11, and 12 mostly have apartments with 8 to 12 units per building, and there is a high-rise condominium on Block 9 with 100 units. In such cases, it would be appropriate to adjust the

Do not copy, post, or distribute

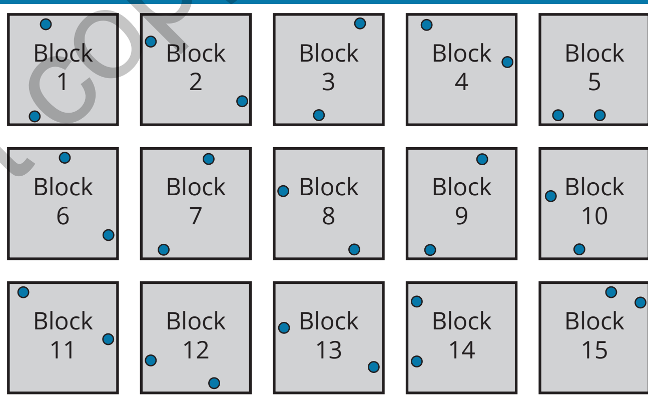
Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

FIGURE 2.6 ■ Systematic Sampling Provides Periodic Selection From the Sample Frame Beginning With a Randomly Derived Start Point



Note. The *skip term* is *k* (*k* = *sample frame* ÷ *target sample size*), and the *start point* is a random number between 1 and *k*.

FIGURE 2.7 ■ Area Sampling: Identify Target Sample Size and Divide by Number of Blocks to Derive Number of Samples to Randomly Select From Each Block



Note. Target sample size = 30 households; number of blocks = 15 blocks; number of samples to randomly select from each block: 30 samples ÷ 15 blocks = 2 samples per block.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

number of samples gathered from each block based on the population density of the block—the more people per block, the larger the sample on that block.

NONPROBABILITY SAMPLING

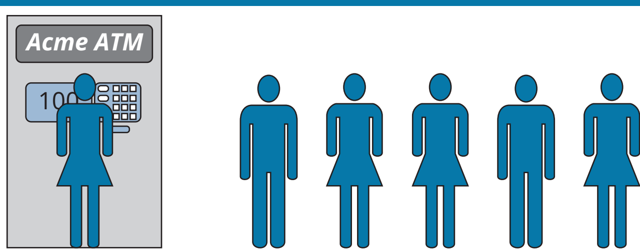
Whereas the hallmark of *probability sampling* is that each person/data item has an equal chance of being included in the *sample*, this is not the case when using **nonprobability sampling** methods. Sometimes, the *sample frame* is not readily accessible or may not exist (e.g., a list of all left-handed people, a list of all undocumented immigrants, a list of all part-time students who work full-time and provide care for a family member), which confounds the *probability sampling* process. Since not everyone has an equal opportunity to be selected for the *sample*, it follows that the data gathered via nonprobability sampling are not expected to be a *representative sample*. Further, since the *sample* does not proportionally represent the characteristics of the larger *population*, research conducted via *nonprobability sampling* does not have *external validity*; what we learn about the *sample* cannot plausibly be generalized to our understanding of the larger *population*.

Initially, it may seem as if *nonprobability sampling* would produce useless results since it lacks *external validity*, but the goal of a study may be to gain a focused understanding of a specific set of individuals, and the larger population is not the primary concern. This will become clearer as we explore examples of various forms of *nonprobability sampling* techniques: *convenience sampling, purposive sampling, quota sampling*, and *snowball sampling.*

Convenience Sampling

Convenience sampling, also known as **availability sampling**, is essentially what it sounds like: The researcher gathers the sample from people or data that's readily available (see Figure 2.8). For example, suppose you wanted to investigate how many siblings people have; you could stand at a busy ATM and survey the people who are waiting in line. Even if you were to gather a sizable sample, considering that not everyone in the community accesses this particular ATM, the results would lack *external validity*, but this sample could still have value; the bank may be planning a promotional campaign directed at families who utilize their ATMs.

FIGURE 2.8 ■ Convenience/Availability Sampling Involves Recruiting Readily Accessible Individuals or Data



Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

Purposive Sampling

Purposive sampling is used when you're interested in gathering information from a (very) specific portion of the population that is considered low prevalence in the population. In *purposive sampling*, the potential participants must meet one or several criteria. For example, suppose there's a new therapy that's been designed to ease the symptoms of patients with cancer undergoing radiation treatments. Fortunately, most of the people in the population do not meet these criteria, meaning that randomly selecting individuals from the population would be an inefficient sampling method. Potential participants must meet all the following criteria:

- 18 to 65 years old
- Cancer diagnosis
- Scheduled for 5 to 10 radiation therapy treatments
- Willing to take an experimental drug or placebo
- Not using any unprescribed substances

The researcher may need to advertise for suitable participants or coordinate with appropriate health care providers or institutions to recruit a sufficient sample. Once again, we see that those who would be in the sample for this study are unlike the people in the population; hence, *external validity* is not plausible. Given the goal of this study (to ease the symptoms of patients with cancer undergoing radiation treatments), this should serve to exemplify the value of *nonprobability sampling*—sometimes the goal of the research is to gain an understanding of *a selected portion of the population* and not the overall population. Although technically the results would not have *external validity*, one could plausibly propose that what was learned from studying this group of individuals may be applicable to others who meet these same criteria (e.g., additional patients with cancer who are undergoing radiation treatment).

Quota Sampling

Quota sampling involves determining how many people you want to sample (e.g., 100 people); the sampling process continues until the quota (of 100 people) is reached. *Quota sampling* can be useful when a sample is needed promptly. For example, suppose you want to sample 50 people as they exit a library, asking what hours they'd like the library to be open. As soon as the 50th person responds, you stop the data collection process, even if there are additional people willing to respond to your survey. From there, you can carry out statistical analysis using the data that you gathered.

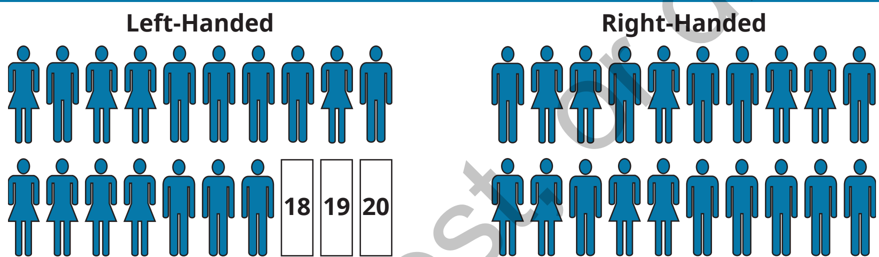
Quota sampling can also be used in a stratified data collection process. For example, suppose the library is concerned about accessibility to left- and right-handed people and wants to gather data from 20 left-handed people and 20 right-handed people; notice that we now have two strata: left-handed people and right-handed people (Figure 2.9). You could stand outside the library and ask people who are exiting, "Are you left or right handed?" and then

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

ask each person if they'd be willing to respond to a brief survey. As shown in Figure 2.9, you've gathered data from 20 right-handed people and 17 left-handed people. It makes sense that you've reached the 20-person quota for right-handed people first since most people are right-handed. At that point, if the next person indicates that they are right-handed, you'd courteously inform that person that you are currently only gathering data from left-handed people (the right-handed part of this survey is now closed). You'd continue that process until you recruit left-handed participants 18, 19, and 20. Also, the quotas don't necessarily need to be equal for each stratum; for example, our quotas might be to gather data from 15 left-handed people and 40 right-handed people. In *quota sampling*, the rule is that you stop collecting data (on a stratum) once the predetermined quota is met.

FIGURE 2.9 ■ Quota Sampling Specifies the Number of Participants Sought (in Each Stratum)



Snowball Sampling

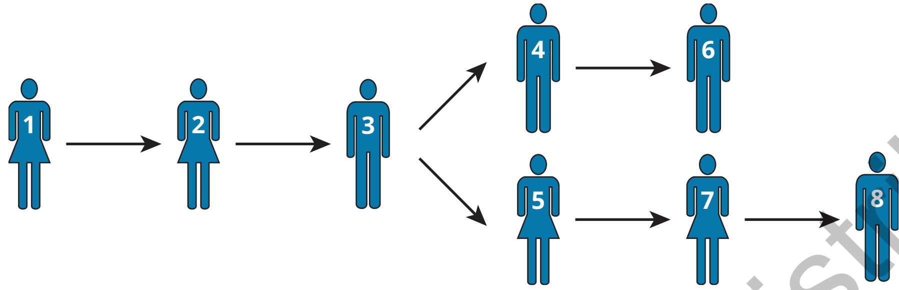
Snowball sampling is based on the process for creating an actual snowball: Start by gathering a handful of snow, then pack on more snow (that's just like the snow that you initially picked up), and continue packing on more and more snow until the snowball reaches the desired size. When it comes to people, consider the following saying: *Birds of a feather flock together*. Since people may know other similar people, *snowball sampling* can be useful when gathering a sample of people with characteristics that are uncommon or not readily identifiable; for example, a musician may know other musicians, and a single parent may know other single parents, and hence, referrals to others may be possible, which is precisely how *snowball sampling* works.

Suppose you're interested in understanding how many hours per week dyslexic students spend on their homework. Since dyslexia is a condition that doesn't present with any overt signs or symptoms, it could take some time to ask numerous passersby if they're dyslexic before you find your first potential participant for your sample. Referring to Figure 2.10, once you find someone who is dyslexic and willing to participate in your study (Person 1 in the figure), before dismissing her, you ask if she knows of anyone else who's dyslexic who might be willing to partake in your study. This first participant may contact her friend who's also dyslexic, and this could be your second participant. Upon completing your encounter with this second person, you ask if she knows of anyone else who's dyslexic, and she refers you to

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

FIGURE 2.10 ■ Snowball Sampling: Requesting Referrals From Those Who Meet the Research Criteria in Order to Recruit Others Who Meet the Research Criteria



her brother (Person 3), who also chooses to partake in the study, and he provides referrals to two more people (Persons 4 and 5), and so on.

Even if the participant characteristic that you are seeking is readily observable (e.g., anyone who uses a wheelchair), considering that this is generally a low-prevalence condition, *snowball sampling* may still be a valuable technique; a person who uses a wheelchair may know of others who also use a wheelchair. Further, someone who uses a wheelchair may provide you with other valuable referrals that may route you to places where people who use wheelchairs can be found (e.g., rehabilitation center, wheelchair store/repair shop, hang-out areas, websites, organizations).

Another possibility is that you're interested in gathering a sample from an *invisible* or a *hidden population*. Such individuals aren't really transparent or necessarily in hiding, but they may possess a particular characteristic that you'd like to recruit in your sample, but unlike people who use a wheelchair, the characteristic that you're interested in is not readily observable, such as individuals who are a single parent, bisexual, fearful of thunderstorms (astraphobia), or chess enthusiasts.

Alternatively, some individuals may actively conceal their attributes, fearing consequences; this may include people who are involved in illegal activities or individuals with potentially embarrassing or uncommon characteristics such as a peculiar obsession, fetish, unpopular belief system, or stigmatized condition or disease. If you're fortunate enough to encounter such a person, instead of asking for names and contact information of others, you may request that participants pass your contact information along to other suitable individuals, enabling those people the option to (anonymously or confidentially) reach out to you. When people feel confident that disclosing sensitive information won't cost them consequences, they may provide honest disclosures of carefully guarded truths. To facilitate this process, keep an open mind, a nonjudgmental attitude, and a professionally positive demeanor. Genuine respect for your (potential) participants can plausibly convey that you are not a threat, which may gain you valuable cooperation, additional participants, and (more) truthful information.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

SAMPLING BIAS

Sampling bias can occur if individuals with a particular characteristic (e.g., high intelligence, low socioeconomic status, youngest, tallest) are disproportionately recruited to partake in a study. Consider the various ways that sampling bias could intentionally or unintentionally occur:

Self-Selection Bias/Voluntary Response Bias—People with a unique characteristic may be more likely to choose to partake in a study than others. For example, academically high-performing high school students may be more motivated to respond to a survey focusing on college plans.

Nonresponse Bias—People may opt not to partake in a study, choose not to respond to selected questions, or drop out of a study. For example, someone with a criminal past may resist responding to topics involving illegal activities.

Undercoverage Bias—There may be too few or no individuals in the sample with a particular characteristic. For example, in a study involving handedness (left/right-handed), it may be difficult to find (enough) individuals who are ambidextrous.

Advertising Bias—Depending on where and how participants are recruited can have an impact on the results. For example, placing recruitment bulletins on public transit (only) would exclude individuals who use other forms of transportation.

Time Bias—Responses can be affected by when a sample is attained. For example, suppose a researcher wants to gather opinions regarding the community center in a public park. Surveying people in the park between 2:00 and 4:00 PM would likely produce very different results compared to data gathered from people who are in that same park between 2:00 and 4:00 AM.

Survivorship Bias—Recruiting only people who meet a certain criterion can be problematic. For example, administering a customer satisfaction survey to people currently using a bank and excluding those who opted to close their account(s) and take their business elsewhere would likely produce incomplete findings.

Recall Bias—Considering that memory is imperfect, asking questions that involve remote events or obscure information may produce erroneous results. For example, asking participants how many times they took aspirin in the last year is likely to result in an estimate rather than an accurate number.

Exclusion Bias—Excluding selected individuals from participating can skew results. For example, administering an online survey rules out people who do not have Internet access and those who are not tech-savvy.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

Sampling bias may resonate through the research process all the way to the statistical results and possibly beyond if the findings are published. When the goal is to assess a specialized portion of the population via a form of nonprobability sampling or compare two different parts of the population, it's important to make this clear throughout your documentation. It would be appropriate to disclose such issues in terms of the limitations of such a study and explicitly discuss how one or more forms of sampling bias may have adversely impacted *external validity*.

OPTIMAL SAMPLE SIZE

You may be wondering: *What's the right sample size?* The answer to this seemingly simple question involves multiple factors, including the research design, the number of groups, the type of variables, the type of statistical analysis, and the desired robustness (power) of your findings. Researchers strive to attain an *optimal sample size*—not too few and not too many.

If a researcher gathers data on too few participants, then the results may be found to be *underpowered*, meaning that the sample was too small to produce robust/stable statistical results; this could adversely impact the solidity of the findings. For example, if you surveyed one person to determine if voting by mail is good or bad, clearly the sample ($n = 1$) is too small to have confidence in the outcome. Imagine the results: "Our research revealed that 100% of those surveyed favored (or opposed) mail-in ballots." An underpowered study could be misleading and may compromise the credibility of those involved.

Conversely, if a sample is too large (e.g., $n = 1{,}000{,}000$), the researcher has likely wasted time and money, delayed the results, and potentially delayed action(s) that would be taken based on the results.

Researchers can use **power calculations** to determine the proper sample size. Generally, statistical power between 0.7 and 0.8 indicates a viable sample size. It's possible to compute power calculations at various points in the research process:

(1) *Before* starting a research project, power calculations can provide an estimate of how many participants you'll need to sample.

(2) *During* the research project, it's appropriate to periodically compute the power to determine if you have (already) acquired a sufficient sample size. For example, suppose you are authorized to gather data on up to 300 people; after gathering data on the 220th person, you find that you have attained a power of 0.8. Having achieved sufficient power, you may opt to conclude the data collection process.

(3) *At the conclusion* of the sampling process, it is appropriate to run a final power calculation to determine the level of power achieved in this study.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

GOOD COMMON SENSE

The acronym **GIGO** (*Garbage In, Garbage Out*) is likely as old as computing. It implies that if you input erroneous data into a computer, the computer will rapidly and meticulously process the data per the instructions of the program, but you can expect the output to be erroneous too. This would be like cooking with spoiled ingredients; even if you used proper cooking techniques, you'd reasonably expect a bad meal to emerge. One can think of sampling as the starting point of the data collection process; hence, the quality of the decisions and techniques pertaining to sampling can have a substantial impact on the quality of the results.

For example, when deploying an online survey, unless comprehensive consideration is given regarding sampling prior to the launch of the survey, the data collected could be polluted with a multitude of possibly undetectable anomalies, including such incidents as responses from people who do not meet the criteria for the study, individuals who revisit the survey multiple times and provide consistent findings in an effort to skew the results, (repeated) bogus/random responses, or responses from automated software (e.g., bots). Clearly, a sample that's biased or not critically controlled can produce misleading results, prompting the researchers to take potentially inappropriate (in)actions.

KEY CONCEPTS

- Rationale for sampling
  - Time
  - Cost
  - Feasibility
  - Extrapolation
- Population
- Sample frame
- Sample
- Representative sample
- External validity
- Probability sampling
  - Simple random sampling
  - Stratified sampling
    - Proportional sampling
    - Disproportional sampling
  - Systematic sampling
  - Area sampling

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

- Nonprobability sampling
  - Availability sampling
  - Purposive sampling
  - Quota sampling
  - Snowball sampling
- Sampling bias
- Optimal sample size

PRACTICE EXERCISES

Exercise 2.1

Define the following terms and provide an example for each:

a. Population

b. Sample frame

c. Sample

d. Representative sample

e. Sampling bias

Exercise 2.2

Explain the difference(s) between a *probability sample* and a *nonprobability sample*.

Exercise 2.3

A community is considering turning a portion of a public park into a dog park. The city planning commission has selected you to survey community members to gather opinions about the project. Explain how you would use *simple random sampling*.

Exercise 2.4

An online video service has commissioned you to conduct a customer satisfaction survey; they provide you with a list of 100,000 subscribers containing their name, cell phone number, and email address. Explain how you would gather surveys from 200 subscribers using *systematic sampling*.

Exercise 2.5

At a community forum with 95 attendees, a community organizer wants to recruit 5 people who have children and 5 people who do not have children to partake in a focus group. Explain how you would recruit these 10 people using *stratified sampling*.

Exercise 2.6

Prior to building a new store in Anytown, Acme Corporation wants to conduct a survey of 240 local households. They provide you with a list of all of the addresses for each of the 80

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.

blocks of Anytown. You are also informed that the blocks are evenly populated. Explain how you would use *area sampling*.

Exercise 2.7

The board of directors of a mall wants to know how much money people spend in the food court. Explain how you would use *availability sampling*.

Exercise 2.8

A tutoring service wants you to survey students who have a learning disability. Explain how you would use *snowball sampling*.

Exercise 2.9

Acme Hospital has selected you to conduct a satisfaction survey of recently discharged patients; they provide you with a list of patients who agreed to be contacted after their hospitalization. You need to gather data from 50 minors (5 to 17 years old) and 100 adults (18 years or older). Explain how you would use *quota sampling*.

Exercise 2.10

A new afterschool program is starting up at Anytown Community Center, providing free recreation and life skills classes (e.g., cooking, first aid, music). Participants must live in Anytown, be between 7 and 17 years old, and be able to attend 2 days a week from 4:00–5:30 PM. Explain how you would use *purposive sampling*.

Do not copy, post, or distribute

Copyright ©2025 by SAGE Publications, Inc.
This work may not be reproduced or distributed in any form or by any means without express written permission of the publisher.