Sampling Techniques

GOAL Classify samples and identify potentially biased samples.

Surveys are often conducted to gather data about a *population*. A **population** is the entire group of people or objects that you want information about. A **census** is a survey of an entire population. When it is too difficult, expensive, or time-consuming to conduct a census, a **sample**, or part of the population, is surveyed.

EXAMPLE 1 Identifying a Population and a Sample

The researchers at a hearing research center want to know if the music played during aerobics classes at health clubs is loud enough to cause hearing damage. They randomly choose 10 health clubs from the 150 health clubs in the area and measure the loudness of the music played during the aerobics classes.
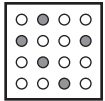
a. What is the population?   **b.** What is the sample?

SOLUTION

a. The population is the aerobics classes at the 150 health clubs.

b. The sample is the aerobics classes at the 10 randomly chosen health clubs.
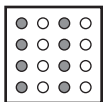
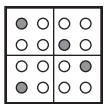Six different types of samples are listed below.

Types of Samples

Simple Random Sample

Members are chosen using a method that gives everyone an equally likely chance of being selected.
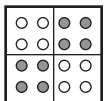
Systematic Sample

Members are chosen using a pattern, such as selecting every other person.

Stratified Sample
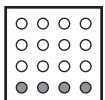
The population is first divided into groups. Then members are randomly chosen from each group.
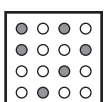
Cluster Sample

The population is first divided into groups. A sample of the groups is randomly chosen. All members of the chosen groups are surveyed.

Convenience Sample

Members are chosen because they are easily accessible.

Self-Selected Sample

Members volunteer to participate.

EXAMPLE 2 **Classifying a Sample**

The officials of the National Football League (NFL) want to know how the players feel about some proposed changes to the NFL rules. They decide to ask a sample of about 100 players. Classify the sample.

a. The officials choose the first 100 players who volunteer their opinions.

b. The officials randomly choose 3 players from each of the 32 teams in the NFL.

c. The officials have a computer generate a list of 100 players from a database that includes all of the players in the NFL.

SOLUTION

a. This is a self-selected sample because the players volunteer.

b. This is a stratified sample because the players are separated by team and randomly chosen from each team.

c. This is a simple random sample because each player has an equally likely chance of being chosen.

CHECK **Examples 1 and 2**

In Exercises 1–3, use the following information.

The editor of a snowboarding magazine wants to know the readers' favorite places to ride. The latest issue of the magazine included a survey, and 238 readers completed and returned the survey.

1. What is the population?

2. What is the sample?

3. Classify the sample.

In a random sample , each person or object has an equally likely chance of being selected. A random sample is most likely to produce a sample that is representative of a population.

A non-random sample can result in a *biased sample*. A biased sample is a sample that may not be representative of a population. In a biased sample, the population can be *underrepresented* or *overrepresented*.

Underrepresented One or more of the parts of a population are left out when choosing the sample.

Overrepresented A greater emphasis is placed on one or more of the parts of a population when choosing the sample.

Simple random, systematic, stratified, and cluster samples are preferred types of samples because the resulting samples are usually representative. The results of convenience and self-selected samples are more likely to be biased.

EXAMPLE 3 **Identifying Potentially Biased Samples**

Administrators at your school want to know if more vegetarian items should be added to the lunch menu. Decide whether the sampling method could result in a biased sample. Explain your reasoning.

a. Survey every 10th student waiting in line to purchase lunch.

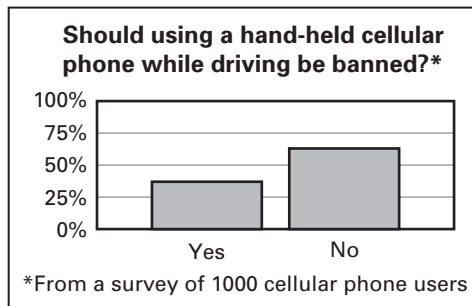**b.** Survey every 25th student who enters the cafeteria during the lunch period.

SOLUTION

a. This method could result in a biased sample because it underrepresents the students who do not purchase lunch. Some of these students may not purchase lunch because there are not enough vegetarian items on the lunch menu.

b. This method is not likely to result in a biased sample because a wide range of students will be surveyed.

EXAMPLE 4 **Analyzing the Results of a Survey**

A local newspaper conducts a survey to find out if city residents think the use of hand-held cellular phones while driving in the city should be banned. The results are shown in the bar graph.

a. Tell why the results may be biased.

b. Suggest a way to eliminate the bias.

Should using a hand-held cellular phone while driving be banned?*

*From a survey of 1000 cellular phone users

SOLUTION

a. The sample contains only cellular phone users, so cellular phone users are overrepresented.

b. The survey should be conducted using a random sample of city residents. For example, survey 100 randomly chosen people at 10 locations in the city.

CHECK **Examples 3 and 4**

4. The host of a radio show wants to know the listeners' favorite bands. He asks listeners to call the radio station and tell him their favorite bands. Could this sampling method result in a biased sample? Explain.

In Exercises 5 and 6, use the following information.

Town officials in Westford want to build a new fire station on Groton Road. A local newspaper surveys residents about the new fire station. The results are shown at the right.

5. Tell why the results may be biased.

6. Suggest a way to eliminate the bias.

Should the new fire station be on Groton Road?*	
Yes	28%
No	72%

*From a survey of 100 residents of Groton Road

Identify the population and the sample.

  1. A quality control inspector needs to estimate the number of defective computers in a group of 250 computers. He tests 25 randomly chosen computers.

  2. The manager of the human resources department at a company wants to know if any of the company's 281 employees would take advantage of a reduced membership to a health club. The manager asks a sample of 70 randomly chosen employees.

  3. The manager of a department store wants to know how the shoppers at the store learned about the store's one-day sale. He asks 50 randomly chosen customers.

  4. A family wants to know the average number of pieces of junk mail they receive each day. They count the pieces of junk mail that they receive each day for a week and find the average.

In Exercises 5–7, classify the sample related to the following situation.

The principal of a school wants to know if the students at the school would like to have the morning announcements posted on the school's Web site.

  5. Survey the first 30 students who enter the cafeteria during the lunch period.

  6. Survey every 10th female student and every 10th male student who enters the cafeteria during the lunch period.

  7. Survey every 20th student who enters the cafeteria during the lunch period.

In Exercises 8–11, classify the sample.

  8. The programming director of a local television station decides to conduct a survey to find out if the station's viewers prefer to watch the local news at 5:30 P.M. or 6:00 P.M. He asks each viewer to call a toll-free number and state his or her preference.

  9. The manager of a movie theater wants to know how the movie viewers feel about the new stadium seating at the theater. She asks every 30th person who exits the theater each Saturday night for a month.

  10. The manager of a credit union wants to know whether its members utilize the online services offered on the credit union's Web site. He decides to call members who have been randomly chosen from a database of all of the members of the credit union.

  11. The owner of a bakery wants to know if its customers are satisfied with its selection of baked goods. She asks the first 20 people who make a purchase on a Saturday morning.

  12. City officials want to conduct a survey about an upcoming election. Describe how to generate a simple random sample, a systematic sample, and a stratified sample of the registered voters in the city.

Decide whether the sampling method could result in a biased sample. Explain your reasoning.

13. On the first day of school, all of the incoming freshmen attend an orientation program. The principal wants to learn the opinions of the freshmen regarding the orientation program. He decides to ask the first 25 freshmen that he sees.

14. The manager of an apartment building wants to know if the residents are satisfied with his service. He writes each apartment number on a piece of paper and places the pieces of paper in a hat. Then he randomly chooses 10 apartment numbers and asks the residents of the 10 apartments about his service.

15. The members of the school drama club want to know how much students are willing to pay for a ticket to one of their productions. They decide that each member of the drama club should ask 5 of his or her friends what they are each willing to pay.

16. The manager of a city bus system wants to know if the people who ride the buses are satisfied with the service. She decides to post mail-in surveys on each of the city's buses.

In Exercises 17 and 18, use the following information.

A town recently passed a leash law for dogs. A survey asks whether the town should designate an off-leash area where dogs can roam freely in the town park. The results of the survey are shown at the right.

Should Spruce Park have an off-leash area?*	
Yes	92%
No	8%

*From a survey of 100 dog owners

17. Tell why the results may be biased.

18. Suggest a way to eliminate the bias.

19. The manager of the produce department at a grocery store receives several crates of oranges. She wants to examine the shipment to determine the quality of the fruit. She decides to open each crate and examine the fruit on the top. Describe what is wrong with this sampling technique.

In Exercises 20–22, use the following information.

A high school is made up of 19% freshmen, 27% sophomores, 29% juniors, and 25% seniors. The editor of the school newspaper wants to survey a sample of 100 students.

20. Suppose the sample includes 25 students from each class. Tell why the results may be biased.

21. Suggest a way to eliminate the bias.

22. Which is more likely to give accurate results: 100 students randomly chosen from a school with 677 students, or 100 students randomly chosen from a school with 504 students? Explain.

23. A medical conference has 500 participating doctors. The table lists the doctors' specialties. A researcher wants to survey a sample of 25 of the doctors to get their opinions on a new medical device.

a. Explain why it may be better for the researcher to use a stratified sample rather than a simple random sample.

b. When a stratified sample is chosen using *proportionate allocation*, the size of each group in the sample is proportional to the size of the group in the population. Assuming the researcher uses proportionate allocation, how many dermatologists should be in the sample?

Specialty	Number of doctors
Dermatology	40
Geriatrics	120
Oncology	140
Pediatrics	100
Surgery	100

24. Measures such as mean and standard deviation, when calculated for an entire population, are called *population parameters*. When such measures are calculated for a sample, they are called *sample statistics*. In this exercise you will compare population parameters and sample statistics for some trees in an apple orchard. The grid represents the entire population of 100 trees. The values in the grid show the number of kilograms of apples produced by each tree during one year. The mean for the population is 68.9 and the standard deviation is about 25.5.

109	52	62	72	110	61	51	50	100	50
54	104	54	111	74	73	77	68	65	66
108	53	27	75	52	117	76	60	64	67
73	36	103	71	67	60	59	26	80	61
38	63	35	112	75	68	51	72	79	62
58	105	55	53	118	57	101	66	116	31
29	57	74	33	102	69	28	71	30	58
39	55	34	120	64	114	70	113	78	63
107	37	56	25	76	70	69	77	30	115
56	40	106	32	119	65	80	78	79	59

a. Use your calculator to choose a random sample of two apple trees. To do so, first enter the data from the grid in a list, L1. Now execute the command **L1(int(100*rand)+1)** where where **int(** is chosen from the **NUM** menu after pressing MATH and **rand** is chosen from the **PRB** menu after pressing MATH. Pressing ENTER twice returns two values from the grid. What is the mean and standard deviation for the sample?

b. Use a similar method to choose a random sample of 10 apple trees. What is the mean and standard deviation for this sample?

c. What do you think would happen to the mean and the standard deviation of the sample if you chose a random sample of 80 apple trees? Explain.