

11.3 Collecting Data

Identifying Sampling Methods in Statistical Studies

The steps in a typical statistical study are shown below.



There are many different ways of sampling a population, but a *random sample* is preferred because it is most likely to be representative of a population. In a **random sample**, each member of a population has an equal chance of being selected.

The other types of samples given below are defined by the methods used to select members. Each sampling method has its advantages and disadvantages.

Core Concept

Types of Samples

For a **self-selected sample**, members of a population can volunteer to be in the sample.



For a **systematic sample**, a rule is used to select members of a population. For instance, selecting every other person.



For a **stratified sample**, a population is divided into smaller groups that share a similar characteristic. A sample is then randomly selected from each group.



For a **cluster sample**, a population is divided into groups, called *clusters*. All of the members in one or more of the clusters are selected.



For a **convenience sample**, only members of a population who are easy to reach are selected.

****Concept 1: Identifying Types of Samples****

1. You want to determine whether students in your school like new design of the school's website. Identify the type of sample described.

a) You list all of the students alphabetically and choose every sixth grader.

You are using a rule to select students. So, the sample is a systematic sample.

b) You mail questionnaires and use only the questionnaires that are returned.

The students can choose whether to respond. So, the sample is a self-selected sample.

c) You ask all of the students in your algebra class.

You are selecting students who are readily available. So, the sample is a convenience sample.

d) You randomly select two students from each classroom.

The students are divided into similar groups by their classrooms, and two students are selected at random from each group. So, the sample is a stratified sample.

2. You want to determine whether the people in your neighborhood like the new social media websites that provide neighborhood updates. Identify the type of sample described.

a) You ask all of the neighbors on the block where you reside.

convenience sample

b) You randomly select a neighbor from each block in the neighborhood.

stratified sample

c) You email a questionnaire to each neighbor and use only the questionnaires that are returned.

self-selected sample

Recognizing Bias in Sampling

A **bias** is an error that results in a misrepresentation of a population. In order to obtain reliable information and draw accurate conclusions about a population, it is important to select an *unbiased sample*. An **unbiased sample** is representative of the population that you want information about. A sample that overrepresents or under-represents part of the population is a **biased sample**. When a sample is biased, the data are invalid. A *random sample* can help reduce the possibility of a biased sample.

****Concept 2: Identifying Bias In Samples****

3. Identify the type of sample and explain why the sample is biased.

a) A news organization asks its viewers to participate in an online poll about bullying.

The viewers can choose whether to participate in the poll. So, the sample is a self-selected sample. The sample is biased because people who go online and respond to the poll most likely have a strong opinion on the subject of bullying.

b) A computer science teacher wants to know how students at a school most often access the internet. The teacher asks students in one of the computer science classes.

The teacher selects students who are readily available. So, the sample is a convenience sample. The sample is biased because other students in the school do not have an opportunity to be chosen.

4. Identify the type of sample and explain why the sample is biased.

a) The principal asks students at one lunch table about the quality of food served in the school cafeteria.

Cluster sample; Students at one table are likely to have different opinions than students at other tables.

b) A sports announcer wants to know how often people in town attend community sporting events. She asks every tenth person in attendance at a local soccer game.

Systematic sample; People who did not go to the soccer game were not surveyed.

****Concept 3: Selecting an Unbiased Sample****

5. You are a member of your schools yearbook committee. You want to poll members of a senior class to find out what the theme of the yearbook should be. There are 246 students in the senior class. Describe a method for selecting a random method sample of 50 students to poll.

Step 1: Make a list of all 246 seniors. Assign each senior a different integer from 1 to 246.
Step 2: Generate 50 unique random integers from 1 to 246 using the randInt feature of a graphing calculator.

Step 3: Choose the 50 students who correspond to the 50 integers you generated in Step 2.

6. You are in charge of the senior-class prom. You want to poll the members of the senior class to find out where the prom should be held. There are 415 students in the senior class. Describe a method for selecting a random sample of 70 students to poll.

Make a list of all 415 seniors. Assign each senior a different integer from 1 to 415. Generate 70 unique random integers from 1 to 415 using the random number function on a spreadsheet. Choose the 70 students that correspond to the 70 integers generated.

Core Concept

Methods of Collecting Data

An **experiment** imposes a treatment on individuals in order to collect data on their response to the treatment. The treatment may be a medical treatment, or it can be any action that might affect a variable in the experiment, such as adding methanol to gasoline and then measuring its effect on fuel efficiency.

An **observational study** observes individuals and measures variables without controlling the individuals or their environment. This type of study is used when it is difficult to control or isolate the variable being studied, or when it may be unethical to subject people to a certain treatment or to withhold it from them.

A **survey** is an investigation of one or more characteristics of a population. In a survey, every member of a sample is asked one or more questions.

A **simulation** uses a model to reproduce the conditions of a situation or process so that the simulated outcomes closely match the real-world outcomes. Simulations allow you to study situations that are impractical or dangerous to create in real life.

****Concept 4: Identifying Methods of Data Collection****

7. Identify the method of data collection each simulation describes.

a) A researcher records whether people at a gas station use hand sanitizer.

The researcher is gathering data without controlling the individuals or applying a treatment. So, this situation is an observational study.

b) A landscaper fertilizes 20 lawns with a regular fertilizer mix and 20 lawns with new organic fertilizer. The landscaper then compares the lawns after 10 weeks and determines which fertilizer is better.

A treatment (organic fertilizer) is being applied to some of the individuals (lawns) in the study. So, this situation is an experiment.

8. Identify the method of data collection each simulation describes.

a) A teacher records how many students enter the classroom and turn in their homework before sitting down at their desks.

The teacher is gathering information without controlling the individuals or applying a treatment. So, this situation is an observational study.

b) A manager uses a computer program to predict how many defective products can be expected on a particular assembly line.

The computer program reproduces the conditions of the assembly line. So, this situation is a simulation.

Recognizing Bias in Survey Questions

When designing a survey, it is important to word survey questions so they do not lead to biased results. Answers to poorly worded questions may not accurately reflect the opinions or actions of those being surveyed. Questions that are flawed in a way that leads to inaccurate results are called **biased questions**. Avoid questions that:

- encourage a particular response
- do not provide enough information to give an accurate opinion
- are too sensitive to answer truthfully
- address more than one issue

****Concept 5: Identify and Correct Bias in Survey Questioning ****

9. A dentist surveys his patients by asking, “Do you brush your teeth at least twice per day and floss everyday?” Explain why the questionnaire may be biased or otherwise introduce bias into the survey. Then describe a way to correct the flaw.

Patients who brush less than twice per day or do not floss daily may be afraid to admit this because the dentist is asking the question. One improvement may be to have patients answer questions about dental hygiene on paper and then put the paper anonymously into a box.

10. A survey question asks, “Should the town build a playground and a dog area in the park on main street?” Explain why this question may be biased or otherwise introduce bias into the survey. Then describe a way to correct the flaw.

The question addresses more than one issue, so a “yes” response does not indicate whether the person wants only the playground, only the dog area, or both. One improvement may be to ask two separate questions, one about the playground and one about the dog area.