

Lesson 13: Populations, Samples, and Generalizing from a Sample to a Population

Classwork

In this lesson, you will learn about collecting data from a sample that is selected from a population. You will also learn about summary values for both a population and a sample and think about what can be learned about the population by looking at a sample from that population.

Exercises 1–4: Collecting Data

1. Describe what you would do if you had to collect data to investigate the following statistical questions using either a sample statistic or a population characteristic. Explain your reasoning in each case.
 - a. How might you collect data to answer the question, “Does the soup taste good?”
 - b. How might you collect data to answer the question, “How many movies do students in your class see in a month?”
 - c. How might you collect data to answer the question, “What is the median price of a home in our town?”
 - d. How might you collect data to answer the question, “How many pets do people own in my neighborhood?”
 - e. How might you collect data to answer the question, “What is the typical number of absences in math classes at your school on a given day?”

- f. How might you collect data to answer the question, “What is the typical life span of a particular brand of flashlight battery?”
- g. How might you collect data to answer the question, “What percentage of girls and of boys in your school have a curfew?”
- h. How might you collect data to answer the question, “What is the most common blood type of students in my class?”

A *population* is the entire set of objects (people, animals, plants, etc.) from which data might be collected. A *sample* is a subset of the population. Numerical summary values calculated using data from an entire population are called *population characteristics*. Numerical summary values calculated using data from a sample are called *statistics*.

- 2. For which of the scenarios in Exercise 1 did you describe collecting data from a population and which from a sample?
- 3. Think about collecting data in the scenarios above. Give at least two reasons you might want to collect data from a sample rather than from the entire population.
- 4. Make up a result you might get in response to the situations in Exercise 1, and identify whether the result would be based on a population characteristic or a sample statistic.
 - a. Does the soup taste good?

- b. How many movies do your classmates see in a month?
- c. What is the median price of a home in our town?
- d. How many pets do people in my neighborhood own?
- e. What is the typical number of absences in math classes at your school on a given day?
- f. What is the typical life span of a particular brand of flashlight batteries?
- g. What percentage of girls and of boys in your school that have a curfew?
- h. What is the most common blood type of my classmates?

Exercise 5: Population or Sample?

- 5. Indicate whether the following statements are summarizing data collected to answer a statistical question from a population or from a sample. Identify references in the statement as population characteristics or sample statistics.
 - a. 54% of the responders to a poll at a university indicated that wealth needed to be distributed more evenly among people.

- b. Are students in the Bay Shore school district proficient on the state assessments in mathematics? In 2013, after all the tests taken by the students in the Bay Shore schools were evaluated, over 52% of those students were at or above proficient on the state assessment.
- c. Does talking on mobile phones while driving distract people? Researchers measured the reaction times of 38 study participants as they talked on mobile phones and found that the average level of distraction from their driving was rated 2.25 out of 5.
- d. Did most people living in New York in 2010 have at least a high school education? Based on the data collected from all New York residents in 2010 by the United States Census Bureau, 84.6% of people living in New York had at least a high school education.
- e. Were there more deaths than births in the United States between July 2011 and July 2012? Data from a health service agency indicated that there were 2% more deaths than births in the U.S. during that timeframe.
- f. What is the fifth best-selling book in the United States? Based on the sales of books in the United States, the fifth best-selling book was *Oh, the Places You'll Go!* by Dr. Seuss.

Exercises 6–8: A Census

6. When data are collected from an entire population, it is called a census. The United States takes a census of its population every ten years, with the most recent one occurring in 2010. Go to http://ri.essortment.com/unitedstatesce_rta.htm to find the history of the U.S. census.
- Identify three things that you found to be interesting.
 - Why is the census important in the United States?
7. Go to the site: www.census.gov/2010census/popmap/ipmtext.php?fl=36. Select the state of New York.
- How many people were living in New York for the 2010 census?
 - Estimate the ratio of those 65 and older to those under 18 years old. Why is this important to think about?
 - Is the ratio a population characteristic or a statistic? Explain your thinking.
8. The American Community Survey (ACS) takes samples from a small percentage of the U.S. population in years between the censuses. (www.census.gov/acs/www/about_the_survey/american_community_survey/)
- What is the difference between the way the ACS collects information about the U.S. population and the way the U.S. Census Bureau collects information?

- b. In 2011, the ACS sampled workers living in New York about commuting to work each day. Why do you think these data are important for the state to know?
- c. Suppose that from a sample of 200,000 New York workers, 32,400 reported traveling more than an hour to work each day. From this information, statisticians determined that between 16% and 16.4% of the workers in the state traveled more than an hour to work every day in 2011. If there were 8,437,512 workers in the entire population, about how many traveled more than an hour to work each day?
- d. Reasoning from a sample to the population is called making an inference about a population characteristic. Identify the statistic involved in making the inference in part (c).
- e. The data about traveling time to work suggest that across the United States typically between 79.8% and 80% of commuters travel alone, 10% to 10.2% carpool, and 4.9% to 5.1% use public transportation. Survey your classmates to find out how a worker in their family gets to work. How do the results compare to the national data? What might explain any differences?

Lesson Summary

The focus of this lesson was on collecting information either from a *population*, which is the entire set of elements in the group of interest, or a subset of the population, called a *sample*. One example of data being collected from a population is the U.S. Census, which collects data from every person in the United States every ten years. The Census Bureau also samples the population to study things that affect the economy and living conditions of people in the U.S. in more detail. When data from a population are used to calculate a numerical summary, the value is called a *population characteristic*. When data from a sample are used to calculate a numerical summary, the value is called a *sample statistic*. Sample statistics can be used to learn about population characteristics.

Problem Set

1. The lunch program at Blake Middle School is being revised to align with the new nutritional standards that reduce calories and increase servings of fruits and vegetables. The administration decided to do a census of all students at Blake Middle School by giving a survey to all students about the school lunches.

<http://frac.org/federal-foodnutrition-programs/school-breakfast-program/school-meal-nutrition-standards>

- a. Name some questions that you would include in the survey. Explain why you think those questions would be important to ask.
- b. Read through the paragraph below that describes some of the survey results. Then, identify the population characteristics and the sample statistics.

About $\frac{3}{4}$ of the students surveyed eat the school lunch regularly. The median number of days per month that students at Blake Middle School ate a school lunch was 18 days. 36% of students responded that their favorite fruit is bananas. The survey results for Tanya's seventh-grade homeroom showed that the median number of days per month that her classmates ate lunch at school was 22, and only 20% liked bananas. The fiesta salad was approved by 78% of the group of students who tried it, but when it was put on the lunch menu, only 40% of the students liked it. Of the seventh graders as a whole, 73% liked spicy jicama strips, but only 2 out of 5 of all the middle school students liked them.

2. For each of the following questions: (1) describe how you would collect data to answer the question, and (2) describe whether it would result in a sample statistic or a population characteristic.
 - a. Where should the eighth-grade class go for their class trip?
 - b. What is the average number of pets per family for families that live in your town?
 - c. If people tried a new diet, what percentage would have an improvement in cholesterol reading?
 - d. What is the average grade point of students who got accepted to a particular state university?
 - e. What is a typical number of home runs hit in a particular season for major league baseball players?

3. Identify a question that would lead to collecting data from the given set as a population, and one where the data could be a sample from a larger population.
- All students in your school
 - Your state
4. Suppose that researchers sampled attendees of a certain movie and found that the mean age was 17 years old. Based on this observation, which of the following would be most likely.
- The mean ages of all of the people who went to see the movie was 17 years old.
 - About a fourth of the people who went to see the movie were older than 51.
 - The mean age of all people who went to see the movie would probably be in an interval around 17 years of age, i.e., between 15 and 19.
 - The median age of those who attended the movie was 17 years old as well.
5. The headlines proclaimed: “Education Impacts Work-Life Earnings Five Times More Than Other Demographic Factors, Census Bureau Reports.” According to a U.S. Census Bureau study, education levels had more effect on earnings over a 40-year span in the workforce than any other demographic factor. www.census.gov/newsroom/releases/archives/education/cb11-153.html
- The article stated that the estimated impact on annual earnings between a professional degree and an eighth-grade education was roughly five times the impact of gender, which was \$13,000. What would the difference in annual earnings be with a professional degree and with an eighth-grade education?
 - Explain whether you think the data are from a population or a sample, and identify either the population characteristic or the sample statistic.

The History of the United States Census

The word *census* is Latin in origin and means to *tax*. The first U.S. census took place over 200 years ago, but the United States is certainly not the first country to implement a census. Based on archaeological records, it appears that the ancient Egyptians conducted a census as early as 3000 B.C.

The U.S. census is mandated by the U.S. Constitution in Article I, Section 2, which states, in part, “Representatives and direct Taxes shall be apportioned among the several States . . . according to their respective Numbers The Number of Representatives shall not exceed one for every thirty thousand, but each State shall have at Least one Representative” The Constitution then specifies how to calculate the number of people in each state and how often the census should take place.

The U.S. census has been conducted every ten years since 1790, but as time has passed, our census has evolved. Not only have the types of questions changed but also the manner in which the data are collected and tabulated. Originally, the census had only a few questions, the purpose of which was to discern the number of people in each household and their ages. Presumably, this data was used to determine the number of men in each state who were available to go to war. Federal marshals were charged with the task of conducting this first census. After collecting data from their respective jurisdictions, the marshals sent the data to President Washington.

As time has passed, more questions have been added to the U.S. census. Today, the census includes questions designed to collect data in various fields such as manufacturing, commerce, and transportation to name a few. Data that were once manually tabulated are now processed by computers. Home visits by census officials were once the norm, but now the census is conducted primarily through the U.S. Postal Service. Each household in the U.S. receives in the mail a copy of the census questionnaire to be completed by its head of household who then mails it back to the Census Bureau. Home visits are paid only to those individuals who do not return the questionnaire by the specified deadline.

The census is an important part of our Constitution. Today, the census not only tells us the population of each state, thereby determining the number of representatives that each state will have in the House of Representatives, but it also provides the U.S. government with very useful data that paint a picture of the current state of our population and how it has changed over the decades.

“U.S. Census History,” *essortment*, accessed November 4, 2014, <http://www.essortment.com/census-history-20901.html>.