

1.H: Sampling and Data (Homework)

1.2 Definitions of Statistics, Probability, and Key Terms

For each of the following eight exercises, identify: a. the population, b. the sample, c. the parameter, d. the statistic, e. the variable, and f. the data. Give examples where appropriate.

1.

A fitness center is interested in the mean amount of time a client exercises in the center each week.

2.

Ski resorts are interested in the mean age that children take their first ski and snowboard lessons. They need this information to plan their ski classes optimally.

3.

A cardiologist is interested in the mean recovery period of her patients who have had heart attacks.

4.

Insurance companies are interested in the mean health costs each year of their clients, so that they can determine the costs of health insurance.

5.

A politician is interested in the proportion of voters in his district who think he is doing a good job.

6.

A marriage counselor is interested in the proportion of clients she counsels who stay married.

7.

Political pollsters may be interested in the proportion of people who will vote for a particular cause.

8.

A marketing company is interested in the proportion of people who will buy a particular product.

Use the following information to answer the next three exercises: A Lake Tahoe Community College instructor is interested in the mean number of days Lake Tahoe Community College math students are absent from class during a quarter.

9.

What is the population she is interested in?

- a. all Lake Tahoe Community College students
- b. all Lake Tahoe Community College English students
- c. all Lake Tahoe Community College students in her classes
- d. all Lake Tahoe Community College math students

10.

Consider the following:

X = number of days a Lake Tahoe Community College math student is absent

In this case, X is an example of a:

- a. variable.
- b. population.
- c. statistic.
- d. data.

11.

The instructor's sample produces a mean number of days absent of 3.5 days. This value is an example of a:

- a. parameter.
- b. data.
- c. statistic.
- d. variable.

1.3 Data, Sampling, and Variation in Data and Sampling

For the following exercises, identify the type of data that would be used to describe a response (quantitative discrete, quantitative continuous, or qualitative), and give an example of the data.

12.

number of tickets sold to a concert

13.

percent of body fat

14.

favorite baseball team

15.

time in line to buy groceries

16.

number of students enrolled at Evergreen Valley College

17.

most-watched television show

18.

brand of toothpaste

19.

distance to the closest movie theatre

20.

age of executives in Fortune 500 companies

21.

number of competing computer spreadsheet software packages

Use the following information to answer the next two exercises: A study was done to determine the age, number of times per week, and the duration (amount of time) of resident use of a local park in San Jose. The first house in the neighborhood around the park was selected randomly and then every 8th house in the neighborhood around the park was interviewed.

22.

“Number of times per week” is what type of data?

- a. qualitative (categorical)
- b. quantitative discrete
- c. quantitative continuous

23.

“Duration (amount of time)” is what type of data?

- a. qualitative (categorical)
- b. quantitative discrete
- c. quantitative continuous

24.

Airline companies are interested in the consistency of the number of babies on each flight, so that they have adequate safety equipment. Suppose an airline conducts a survey. Over Thanksgiving weekend, it surveys six flights from Boston to Salt Lake City to determine the number of babies on the flights. It determines the amount of safety equipment needed by the result of that study.

- a. Using complete sentences, list three things wrong with the way the survey was conducted.
- b. Using complete sentences, list three ways that you would improve the survey if it were to be repeated.

25.

Suppose you want to determine the mean number of students per statistics class in your state. Describe a possible sampling method in three to five complete sentences. Make the description detailed.

26.

Suppose you want to determine the mean number of cans of soda drunk each month by students in their twenties at your school. Describe a possible sampling method in three to five complete sentences. Make the description detailed.

27.

List some practical difficulties involved in getting accurate results from a telephone survey.

28.

List some practical difficulties involved in getting accurate results from a mailed survey.

29.

With your classmates, brainstorm some ways you could overcome these problems if you needed to conduct a phone or mail survey.

30.

The instructor takes her sample by gathering data on five randomly selected students from each Lake Tahoe Community College math class. The type of sampling she used is

- a. cluster sampling
- b. stratified sampling
- c. simple random sampling
- d. convenience sampling

31.

A study was done to determine the age, number of times per week, and the duration (amount of time) of residents using a local park in San Jose. The first house in the neighborhood around the park was selected randomly and then every eighth house in the neighborhood around the park was interviewed. The sampling method was:

- a. simple random
- b. systematic
- c. stratified
- d. cluster

32.

Name the sampling method used in each of the following situations:

- a. A woman in the airport is handing out questionnaires to travelers asking them to evaluate the airport's service. She does not ask travelers who are hurrying through the airport with their hands full of luggage, but instead asks all travelers who are sitting near gates and not taking naps while they wait.
- b. A teacher wants to know if her students are doing homework, so she randomly selects rows two and five and then calls on all students in row two and all students in row five to present the solutions to homework problems to the class.
- c. The marketing manager for an electronics chain store wants information about the ages of its customers. Over the next two weeks, at each store location, 100 randomly selected customers are given questionnaires to fill out asking for information about age, as well as about other variables of interest.
- d. The librarian at a public library wants to determine what proportion of the library users are children. The librarian has a tally sheet on which she marks whether books are checked out by an adult or a child. She records this data for every fourth patron who checks out books.

- e. A political party wants to know the reaction of voters to a debate between the candidates. The day after the debate, the party's polling staff calls 1,200 randomly selected phone numbers. If a registered voter answers the phone or is available to come to the phone, that registered voter is asked whom he or she intends to vote for and whether the debate changed his or her opinion of the candidates.

33.

A "random survey" was conducted of 3,274 people of the "microprocessor generation" (people born since 1971, the year the microprocessor was invented). It was reported that 48% of those individuals surveyed stated that if they had \$2,000 to spend, they would use it for computer equipment. Also, 66% of those surveyed considered themselves relatively savvy computer users.

- Do you consider the sample size large enough for a study of this type? Why or why not?
- Based on your "gut feeling," do you believe the percents accurately reflect the U.S. population for those individuals born since 1971? If not, do you think the percents of the population are actually higher or lower than the sample statistics? Why?
Additional information: The survey, reported by Intel Corporation, was filled out by individuals who visited the Los Angeles Convention Center to see the Smithsonian Institute's road show called "America's Smithsonian."
- With this additional information, do you feel that all demographic and ethnic groups were equally represented at the event? Why or why not?
- With the additional information, comment on how accurately you think the sample statistics reflect the population parameters.

34.

The Well-Being Index is a survey that follows trends of U.S. residents on a regular basis. There are six areas of health and wellness covered in the survey: Life Evaluation, Emotional Health, Physical Health, Healthy Behavior, Work Environment, and Basic Access. Some of the questions used to measure the Index are listed below.

Identify the type of data obtained from each question used in this survey: qualitative(categorical), quantitative discrete, or quantitative continuous.

- Do you have any health problems that prevent you from doing any of the things people your age can normally do?
- During the past 30 days, for about how many days did poor health keep you from doing your usual activities?
- In the last seven days, on how many days did you exercise for 30 minutes or more?
- Do you have health insurance coverage?

35.

In advance of the 1936 Presidential Election, a magazine titled Literary Digest released the results of an opinion poll predicting that the republican candidate Alf Landon would win by a large margin. The magazine sent post cards to approximately 10,000,000 prospective voters. These prospective voters were selected from the subscription list of the magazine, from automobile registration lists, from phone lists, and from club membership lists. Approximately 2,300,000 people returned the postcards.

- Think about the state of the United States in 1936. Explain why a sample chosen from magazine subscription lists, automobile registration lists, phone books, and club membership lists was not representative of the population of the United States at that time.
- What effect does the low response rate have on the reliability of the sample?
- Are these problems examples of sampling error or nonsampling error?
- During the same year, George Gallup conducted his own poll of 30,000 prospective voters. These researchers used a method they called "quota sampling" to obtain survey answers from specific subsets of the population. Quota sampling is an example of which sampling method described in this module?

36.

Crime-related and demographic statistics for 47 US states in 1960 were collected from government agencies, including the FBI's *Uniform Crime Report*. One analysis of this data found a strong connection between education and crime indicating that higher levels of education in a community correspond to higher crime rates.

Which of the potential problems with samples discussed in [Example 1.H.4](#) could explain this connection?

37.

YouPolls is a website that allows anyone to create and respond to polls. One question posted April 15 asks:

“Do you feel happy paying your taxes when members of the Obama administration are allowed to ignore their tax liabilities?” (lastbaldeagle. 2013. On Tax Day, House to Call for Firing Federal Workers Who Owe Back Taxes. Opinion poll posted online at: <http://www.youpolls.com/details.aspx?id=12328> (accessed May 1, 2013).)

As of April 25, 11 people responded to this question. Each participant answered “NO!”

Which of the potential problems with samples discussed in this module could explain this connection?

38.

A scholarly article about response rates begins with the following quote:

“Declining contact and cooperation rates in random digit dial (RDD) national telephone surveys raise serious concerns about the validity of estimates drawn from such research.” (Scott Keeter et al., “Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey,” Public Opinion Quarterly 70 no. 5 (2006), <http://poq.oxfordjournals.org/content/70/5/759.full> (accessed May 1, 2013).)

The Pew Research Center for People and the Press admits:

“The percentage of people we interview – out of all we try to interview – has been declining over the past decade or more.” (Frequently Asked Questions, Pew Research Center for the People & the Press, <http://www.people-press.org/methodol...wer-your-polls> (accessed May 1, 2013).)

- a. What are some reasons for the decline in response rate over the past decade?
- b. Explain why researchers are concerned with the impact of the declining response rate on public opinion polls.

1.H: Sampling and Data (Homework) is shared under a [not declared](#) license and was authored, remixed, and/or curated by LibreTexts.

- **1.8: Homework** by OpenStax is licensed CC BY 4.0. Original source: <https://openstax.org/details/books/introductory-business-statistics>.