Ch. 1   Statistics, Data, and Statistical Thinking

1.1   The Science of Statistics

1   Define Statistics

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Solve the problem.
   1) Which of the following is not the job of a statistician?
      A) implementing new procedures based on the results of a study
      B) determining what information is relevant in a given problem
      C) collecting numerical information in the form of data
      D) determining whether the conclusions drawn from a study are to be trusted

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

   2) What is statistics?

1.2   Types of Statistical Applications

1   Define Descriptive and Inferential Statistics

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Solve the problem.
   1) A recent report stated "Based on a sample of 170 truck drivers, there is evidence to indicate that, on average, independent truck drivers earn more than company–hired truck drivers." Does this statement describe descriptive or inferential statistics?
      A) Inferential statistics   B) Descriptive statistics

   2) A survey of high school teenagers reported that 96% of those sampled are interested in pursuing a college education. Does this statement describe descriptive or inferential statistics?
      A) Descriptive statistics   B) Inferential statistics

   3) The average age of the students in a statistics class is 22 years. Does this statement describe descriptive or inferential statistics?
      A) Descriptive statistics   B) Inferential statistics

   4) From past figures, it is predicted that 14% of the registered voters will vote in the March primary. Does this statement describe descriptive or inferential statistics?
      A) Inferential statistics   B) Descriptive statistics

   5) Parking at a large university has become a very big problem. University administrators are interested in determining the average parking time (e.g. the time it takes a student to find a parking spot) of its students. An administrator inconspicuously followed 250 students and carefully recorded their parking times. The university is interested in using the information from the sample of 250 students collected to learn information about the entire student parking population. Would this be an application of descriptive or inferential statistics?
      A) Inferential statistics   B) Descriptive statistics

6) As part of an economics class project, students were asked to randomly select 500 New York Stock Exchange (NYSE) stocks from the Wall Street Journal. As part of the project, students were asked to summarize the current prices (also referred to as the closing price of the stock for a particular trading date) of the collected stocks using graphical and numerical techniques. Would this be an application of descriptive or inferential statistics?

    A) Descriptive statistics   B) Inferential statistics

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

7) In a survey of 3000 high school students, 15% of those surveyed read at least one best–seller each month. Give an example of a descriptive statement and an inferential statement that could be made based on this information.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

8) Which of the following is not an element of descriptive statistical problems?
    A) predictions are made about a larger set of data
    B) information revealed in a data set is summarized
    C) data are displayed visually in graphs
    D) patterns in a data set are identified

Answer the question True or False.

9) When we take data obtained from a sample and make generalizations or predictions about the entire population, we are utilizing inferential statistics.
    A) True   B) False

10) Statistics involves two different processes, describing sets of data and drawing conclusions about the sets of data on the basis of sampling.
    A) True   B) False

1.3 Fundamental Elements of Statistics

1 Identify Elements of Statistics

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Solve the problem.

1) Parking at a university has become a problem. University administrators are interested in determining the average time it takes a student to find a parking spot. An administrator inconspicuously followed 120 students and recorded how long it took each of them to find a parking spot. Identify the population of interest to the university administration.
    A) the entire set of students who park at the university
    B) the 120 students about whom the data were collected
    C) the entire set of faculty, staff, and students who park at the university
    D) the students who park at the university between 9 and 10 AM on Wednesdays

2) Parking at a university has become a problem. University administrators are interested in determining the average time it takes a student to find a parking spot. An administrator inconspicuously followed 200 students and recorded how long it took each of them to find a parking spot. Identify the variable of interest to the university administration.
    A) time to find a parking spot   B) students who drive cars on campus
    C) number of empty parking spots   D) number of students who cannot find a spot

3) An assembly line is operating satisfactorily if fewer than 4% of the phones produced per day are defective. To check the quality of a day's production, the company randomly samples 50 phones from a day's production to test for defects. Define the population of interest to the manufacturer.
   A) all the phones produced during the day in question
   B) the 50 phones sampled and tested
   C) the 50 responses: defective or not defective
   D) the 4% of the phones that are defective

4) An insurance company conducted a study to determine the percentage of cardiologists who had been sued for malpractice in the previous ten years. The sample was randomly chosen from a national directory of doctors. What is the variable of interest in this study?
   A) the responses: have been sued/have not been sued for malpractice in the last ten years
   B) the doctor's area of expertise (i.e., cardiology, pediatrics, etc.)
   C) the number of doctors who are cardiologists
   D) all cardiologists in the directory

5) A study attempted to estimate the proportion of Florida residents who were willing to spend more tax dollars on protecting the Florida coastline from environmental disasters. Thirty–two hundred Florida residents were surveyed.Which of the following is the population used in the study?
   A) all Florida residents
   B) the 3200 Florida residents who were surveyed
   C) Florida residents willing to spend more tax dollars protecting the coastline from environmental disasters
   D) all Florida residents who lived along the coastline

6) A study attempted to estimate the proportion of Florida residents who were willing to spend more tax dollars on protecting the Florida beaches from environmental disasters. Twenty–five hundred Florida residents were surveyed.Which of the following describes the variable of interest in the study?
   A) the response to the question, "Are you willing to spend more tax dollars on protecting the Florida beaches from environmental disasters?"
   B) the response to the question "Do you live along the beach?"
   C) the response to the question "Do you use the beach?"
   D) the 2500 Florida residents surveyed

7) Parking at a large university has become a very big problem. University administrators are interested in determining the average parking time (e.g. the time it takes a student to find a parking spot) of its students. An administrator inconspicuously followed 250 students and carefully recorded their parking times. Identify the population of interest to the university administration.
   A) the 250 students that data was collected from
   B) the entire set of students that park at the university
   C) a single student that parks at the university
   D) the parking time, defined to be the amount of time the student spent finding a parking spot

8) Parking at a large university has become a very big problem. University administrators are interested in determining the average parking time (e.g. the time it takes a student to find a parking spot) of its students. An administrator inconspicuously followed 250 students and carefully recorded their parking times. Identify the sample of interest to the university administration.
   A) the 250 students that data was collected from
   B) the entire set of students that park at the university
   C) a single student that parks at the university
   D) the parking time, defined to be the amount of time the student spent finding a parking spot

9) Parking at a large university has become a very big problem. University administrators are interested in determining the average parking time (e.g. the time it takes a student to find a parking spot) of its students. An administrator inconspicuously followed 250 students and carefully recorded their parking times. Identify the experimental unit of interest to the university administration.
   A) the 250 students that data was collected from
   B) the entire set of students that park at the university
   C) a single student that parks at the university
   D) the parking time, defined to be the amount of time the student spent finding a parking spot

10) Parking at a large university has become a very big problem. University administrators are interested in determining the average parking time (e.g. the time it takes a student to find a parking spot) of its students. An administrator inconspicuously followed 250 students and carefully recorded their parking times. Identify the variable of interest to the university administration.
   A) the 250 students that data was collected from
   B) the entire set of students that park at the university
   C) a single student that parks at the university
   D) the parking time, defined to be the amount of time the student spent finding a parking spot

11) As part of an economics class project, students were asked to randomly select 500 New York Stock Exchange (NYSE) stocks from the Wall Street Journal. As part of the project, students were asked to summarize the current prices (also referred to as the closing price of the stock for a particular trading date) of the collected stocks using graphical and numerical techniques. Identify the population of interest for this study.
   A) the entire set of stocks that are traded on the NYSE
   B) the 500 NYSE stocks that current prices were collected from
   C) the current price (or closing price) of a NYSE stock
   D) a single stock traded on the NYSE

12) As part of an economics class project, students were asked to randomly select 500 New York Stock Exchange (NYSE) stocks from the Wall Street Journal. As part of the project, students were asked to summarize the current prices (also referred to as the closing price of the stock for a particular trading date) of the collected stocks using graphical and numerical techniques. Identify the sample of interest for this study.
   A) the entire set of stocks that are traded on the NYSE
   B) the 500 NYSE stocks that current prices were collected from
   C) the current price (or closing price) of a NYSE stock
   D) a single stock traded on the NYSE

13) As part of an economics class project, students were asked to randomly select 500 New York Stock Exchange (NYSE) stocks from the Wall Street Journal. As part of the project, students were asked to summarize the current prices (also referred to as the closing price of the stock for a particular trading date) of the collected stocks using graphical and numerical techniques. Identify the experimental unit of interest for this study.
   A) the entire set of stocks that are traded on the NYSE
   B) the 500 NYSE stocks that current prices were collected from
   C) the current price (or closing price) of a NYSE stock
   D) a single stock traded on the NYSE

14) As part of an economics class project, students were asked to randomly select 500 New Your Stock Exchange (NYSE) stocks from the Wall Street Journal. As part of the project, students were asked to summarize the current prices (also referred to as the closing price of the stock for a particular trading date) of the collected stocks using graphical and numerical techniques. Identify the variable of interest for this study.
   A) the entire set of stocks that are traded on the NYSE
   B) the 500 NYSE stocks that current prices were collected from
   C) the current price (or closing price) of a NYSE stock
   D) a single stock traded on the NYSE

15) A study in the state of Georgia was conducted to determine the percentage of all community college students who have taken at least one online class. 1500 community college students were contacted and asked if they had taken at least one online class during their time at their community college. These responses were then used to estimate the percentage of all community college students who have taken at least one online class. Identify the population of interest in this study.
   A) the response (Yes/No) to the question, "Have you taken at least one online class?"
   B) the 1500 community college students contacted
   C) all community college students in the state of Georgia
   D) the number of online classes a student has taken

16) A study in the state of Georgia was conducted to determine the percentage of all community college students who have taken at least one online class. 1500 community college students were contacted and asked if they had taken at least one online class during their time at their community college. These responses were then used to estimate the percentage of all community college students who have taken at least one online class. Identify the variable of interest in this study.
   A) the response (Yes/No) to the question, "Have you taken at least one online class?"
   B) the 1500 community college students contacted
   C) all community college students in the state of Georgia
   D) the number of online classes a student has taken

17) Which of the following is not typically an element of inferential statistical problems?
   A) census   B) sample
   C) variable of interest   D) measure of reliability

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

18) Parking at a university has become a problem. University administrators are interested in determining the average time it takes a student to find a parking spot. An administrator inconspicuously followed 140 students and recorded how long it took each of them to find a parking spot. Identify the population, sample, and variable of interest to the administrators.

19) A quality inspector tested 76 copiers in an attempt to estimate the average failure rate of the copier model. His study indicated that the number of failures decreased from two years ago, indicating an increase in the reliability of the copiers. Describe the variable of interest to the inspector.

20) A high school guidance counselor analyzed data from a sample of 300 community colleges throughout the United States. One of his goals was to estimate the annual tuition costs of community colleges in the United States. Describe the population and variable of interest to the guidance counselor.

21) Explain why it is not necessary to provide a measure of reliability when a census is used rather than a sample.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Answer the question True or False.

22) A variable is a characteristic or property of a population.
   A) True   B) False

23) Measurement is the process of assigning numbers to variables of individual population units.
   A) True   B) False

24) A census is feasible when the population of interest is small.
   A) True   B) False

25) The process of using information from a sample to make generalizations about the larger population is called statistical inference.

    A) True   B) False

26) A measure of reliability is an important element of a descriptive statistical problem.

    A) True   B) False

1.4 Types of Data

1 Classify Data as Quantitative or Qualitative

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Solve the problem.

1) The amount of television viewed by today's youth is of primary concern to Parents Against Watching Television (PAWT). 250 parents of elementary school–aged children were asked to estimate the number of hours per week that their child watches television. Identify the type of data collected by PAWT.

    A) quantitative   B) qualitative

2) The manager of a car dealership records the colors of automobiles on a used car lot. Identify the type of data collected.

    A) qualitative   B) quantitative

3) A postal worker counts the number of complaint letters received by the United States Postal Service in a given day. Identify the type of data collected.

    A) quantitative   B) qualitative

4) An usher records the number of unoccupied seats in a movie theater during each viewing of a film. Identify the type of data collected.

    A) quantitative   B) qualitative

5) A fan observes the numbers on the shirts of a girl's soccer team. Identify the type of data collected.

    A) qualitative   B) quantitative

6) Parking at a large university has become a very big problem. University administrators are interested in determining the average parking time (e.g. the time it takes a student to find a parking spot) of its students. An administrator inconspicuously followed 250 students and carefully recorded their parking times. What type of variable is the administration interested in collecting?

    A) quantitative data   B) qualitative data

7) As part of an economics class project, students were asked to randomly select 500 New York Stock Exchange (NYSE) stocks from the Wall Street Journal. As part of the project, students were asked to summarize the current prices (also referred to as the closing price of the stock for a particular trading date) of the collected stocks using graphical and numerical techniques. What type of variable is being collected?

    A) quantitative data   B) qualitative data

8) A study in the state of Georgia was conducted to determine the percentage of all community college students who have taken at least one online class. 1500 community college students were contacted and asked if they had taken at least one online class during their time at their community college. These responses were then used to estimate the percentage of all community college students who have taken at least one online class. What type of variable is being collected?

    A) quantitative data   B) qualitative data

9) Which data about paintings would *not* be qualitative?

    A) the value   B) the artist   C) the style   D) the theme

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

10) Gender is one variable of interest in a study of the effectiveness of a new medication. For data entry purposes, the researcher conducting the study assigns 1 for Male and 2 for Female. Is the gender data quantitative or qualitative?

1.5   Collecting Data

1   Identify Data Collection Method

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Solve the problem.

1) The amount of television viewed by today's youth is of primary concern to Parents Against Watching Television (PAWT). 330 parents of elementary school–aged children were asked to estimate the number of hours per week that their child watches television. Identify how the data were collected in this study.
   A) from a survey   B) from a published source
   C) from a designed experiment   D) observationally

2) A personnel director studied the eating habits of a company's employees. The director noted whether employees brought their own lunch to work, ate at the company cafeteria, or went out to eat lunch. This type of data collection would best be considered as a(n) _____.
   A) observational study   B) designed experiment

3) A student worked on her statistics project in the library and found a reference book that contained the median family incomes for all 50 states. On her project, she would report her data as being collected _____.
   A) from a published source   B) from a designed experiment
   C) observationally   D) from a survey

4) What method of data collection would you use to collect data for a study where a drug was given to 40 patients and a placebo to another group of 40 patients to determine if the drug has an effect on a patient's illness?
   A) designed experiment   B) published source
   C) observational study   D) survey

5) What method of data collection would you use to collect data for a study where a political pollster wishes to determine if his candidate is leading in the polls?
   A) survey   B) published source
   C) designed experiment   D) observational study

6) Parking at a large university has become a very big problem. University administrators are interested in determining the average parking time (e.g. the time it takes a student to find a parking spot) of its students. An administrator inconspicuously followed 250 students and carefully recorded their parking times. Identify the data collection method used by the administration in this study.
   A) data from a published source
   B) data from a designed experiment
   C) data collected observationally

7) As part of an economics class project, students were asked to randomly select 500 New York Stock Exchange (NYSE) stocks from the Wall Street Journal. As part of the project, students were asked to summarize the current prices (also referred to as the closing price of the stock for a particular trading date) of the collected stocks using graphical and numerical techniques. Identify the data collection method used in this study.
   A) data from a published source
   B) data from a designed experiment
   C) data collected observationally

8) Does online teaching help or hinder student learning? To help answer this question, a statistics teacher decided to teach his three sections of a particular class using three different teaching models – a traditional face–to–face section, a completely online section, and a hybrid or blended section that incorporated both a face–to–face and online component in the section. Students were randomly assigned to the different sections, taught identical information using the different teaching formats, and given identical examinations to measure student learning. The goal was to identify if the teaching method used affected student learning performance. Identify the data collection method used in this study.
A) data from a published source
B) data from a designed experiment
C) data collected observationally

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

9) What is meant by a representative sample?

10) What is the most common way to satisfy the representative sample requirement?

11) Three female students and two male students are to be chosen from a group of 30 female students and 20 male students. Does this sample of five students satisfy the conditions to be a random sample of the 50 students in the group? Explain.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Answer the question True or False.
12) In an observational study, the researcher exerts strict control over the units in the study.
A) True   B) False

13) When using data from a published source, it is not important to know how the data were collected and whether randomization was used.
A) True   B) False

1.6   The Role of Statistics in Critical Thinking and Ethics

1   Identify Bias

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Solve the problem.
1) Define statistical thinking.

2) What is meant by selection bias?

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

3) Which type of problem has occurred when inaccuracies exist in the values of the data recorded?
A) measurement error   B) nonresponse bias   C) selection bias

4) A watchdog group is investigating how people are treated during the foreclosure process. Surveys were mailed to a random sample of 300 people who had recently been threatened with foreclosure. 75 of the surveys were returned by the postal service because the intended recipients had moved and left no forwarding address. What type of problem has occurred?
A) nonresponse bias   B) selection bias   C) measurement error

5) A university was interested in student reaction to a proposal to spend more on athletic scholarships and less on academic scholarships. 35 student athletes were surveyed. What type of problem has occurred?
    A) selection bias   B) nonresponse bias   C) measurement error

6) The way in which an interviewer asks a question about political party affiliation causes respondents to answer that they have no affiliation when they actually do. What type of problem has occurred?
    A) measurement error   B) selection bias   C) nonresponse bias

7) A student completing a research project for a criminal justice class obtained a radar gun for determining automobile speeds and recorded the speeds of automobiles passing a fixed location over a period of several hours. The student was unaware that the device needed to be recharged after two hours of use and that the speeds recorded after two hours were not reliable. What type of problem has occurred?
    A) measurement error   B) selection bias   C) nonresponse bias

8) Because of the possible legal consequences, many people in a sample of the U.S. population chose not to participate in a survey regarding illegal drug use. What type of problem has occurred?
    A) nonresponse bias   B) selection bias   C) measurement error

9) A middle school was interested in surveying their students to find out opinions about the schools media center. To facilitate data collection, the homeroom period was extended 30 minutes to allow everyone in the school ample time to respond to a short questionnaire. Unfortunately, it was learned after the surveys had been completed that all honors students in the middle school were on an all–day field trip and away from school for the entire day. The exclusion of their input into the survey would be considered which type of sampling problem?
    A) selection bias   B) nonresponse bias   C) measurement error

10) A county planning commission is attempted to survey 1500 households from the counties 400,000 households. A random sample was selected and surveys were mailed to the randomly selected households, but only 1075 were returned. The inability to collect data from the 425 households that didn't return the survey would be considered which type of sampling problem?
    A) selection bias   B) nonresponse bias   C) measurement error

2   Unethical Statistics

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Solve the problem.

11) Give an example of unethical statistical practice.

12) A health food company has the following statement on their new product packaging: "Prevents all types of cancer!" (Fact: Past studies have shown that some ingredients in the new product have been know to possibly reduce the risk of many types of cancer). Discuss why it is unethical to make this statement.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

13) A researcher studying malnutrition among children in a developing country collected weights of a random sample of children using a scale that she had set to give weights 2.5 kilograms less than the actual weight. Which statement best describes this situation?
    A) Measurement error has occurred, and the researcher is guilty of unethical statistical practice.
    B) Measurement error has occurred, but the researcher is not guilty of unethical statistical practice.
    C) Measurement error has not occurred, but the researcher is guilty of unethical statistical practice.
    D) Measurement error has not occurred, and the researcher is not guilty of unethical statistical practice.

Ch. 1   Statistics, Data, and Statistical Thinking
Answer Key

1.1   The Science of Statistics
1   Define Statistics
   1) A
   2) Statistics is the science of data that involves collecting, classifying, summarizing, organizing, analyzing, and interpreting numerical information.

1.2   Types of Statistical Applications
1   Define Descriptive and Inferential Statistics
   1) A
   2) A
   3) A
   4) A
   5) A
   6) A
   7) Descriptive: 15% of the students sampled (or 450) read at least one best–seller each month.

      Inferential: Based on the survey, we estimate that about 15% of all high school students read at least one best–seller each month.
   8) A
   9) A
  10) A

1.3   Fundamental Elements of Statistics
1   Identify Elements of Statistics
   1) A
   2) A
   3) A
   4) A
   5) A
   6) A
   7) B
   8) A
   9) C
  10) D
  11) A
  12) B
  13) D
  14) C
  15) C
  16) A
  17) A
  18) The population of interest are all students at the university who park. The sample is the parking times of the 140 students that were collected by the university administrator. The variable of interest to the administrators is the parking time variable.
  19) The variable of interest to the researcher is the failure rate of the copiers.
  20) The population of interest to the guidance counselor is all community colleges in the United States. The variable of interest is the annual tuition cost of the community college.
  21) When a census is used, there should be no error.
  22) B
  23) A
  24) A
  25) A
  26) B

1.4   Types of Data
1   Classify Data as Quantitative or Qualitative
    1) A
    2) A
    3) A
    4) A
    5) A
    6) A
    7) A
    8) B
    9) A
    10) Qualitative; The numbers are arbitrarily selected numerical codes for the categories and have no utility beyond that.

1.5   Collecting Data
1   Identify Data Collection Method
    1) A
    2) A
    3) A
    4) A
    5) A
    6) C
    7) A
    8) B
    9) a sample that exhibits characteristics typical of those possessed by the population of interest
    10) selecting a random sample
    11) No; not every sample of 5 students from the group has an equal chance of selection; for example, a sample consisting of 5 males has no chance of being selected.
    12) B
    13) B

1.6   The Role of Statistics in Critical Thinking and Ethics
1   Identify Bias
    1) Statistical thinking involves applying rational thought and the science of statistics to critically assess data and make inferences.
    2) Selection bias is when a subset of the experimental units in the population is excluded so that these units have no possibility of being selected in the sample.
    3) A
    4) A
    5) A
    6) A
    7) A
    8) A
    9) A
    10) B

2   Unethical Statistics
    11) Researchers select a biased sample, with the intention of misleading the public.
    12) Answers may vary. One possible answer is that the past studies show that the ingredients only have possible cancer reducing effects on many, not all, types of cancer.
    13) A