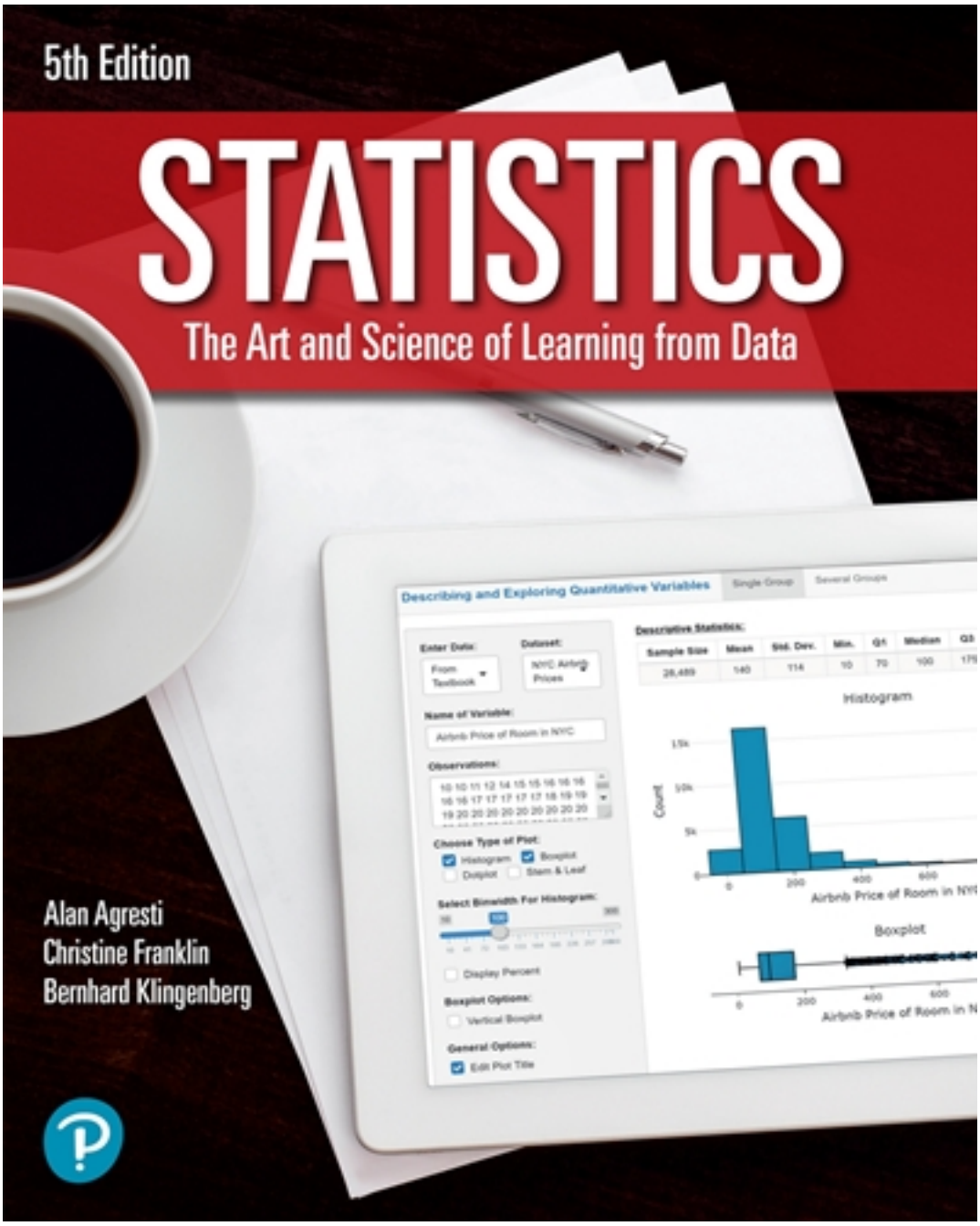
Test Bank for Statistics 5th Edition by Agresti

CLICK HERE TO ACCESS COMPLETE Test Bank



Test Bank

TEST BANK

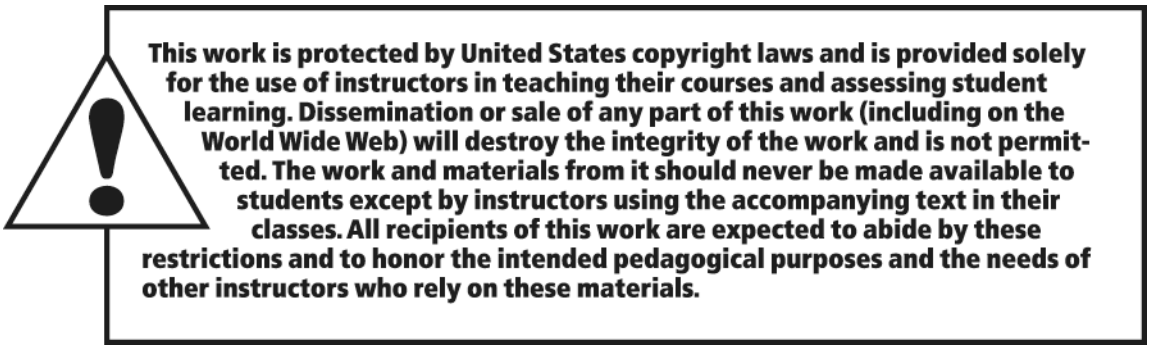
STATISTICS:
THE ART AND SCIENCE OF LEARNING FROM DATA
FIFTH EDITION

Alan Agresti
University of Florida

Christine Franklin
University of Georgia

Bernhard Klingenberg
Williams College and New College of Florida

Pearson



Reproduced by Pearson from electronic files supplied by the author.

ISBN-13: 978-0-13-660885-1
ISBN-10: 0-13-660885-X

***Statistics: The Art and Science of Learning from Data, 5e* (Agresti/Franklin/Klingenberg)**

Ch. 1 Statistics: The Art and Science of Learning From Data

1.1 Using Data to Answer Statistical Questions

1 Identify aspects of a study.

Answer true or false.

1) The information we gather with experiments and with surveys is collectively called data.
A) True B) False
Answer: A
Objective: (1) Identify aspects of a study.

Fill in the blank.

2) The information we gather with experiments and with surveys is collectively called _____.
Answer: data
Objective: (1) Identify aspects of a study.

3) _____ is the art and science of learning from data.
Answer: Statistics
Objective: (1) Identify aspects of a study.

4) The three main aspects of statistics are _____, _____, and _____.
Answer: design, description, inference
Objective: (1) Identify aspects of a study.

Provide an appropriate response.

5) What is statistics?
Answer: Statistics is the art and science of learning from data.
Objective: (1) Identify aspects of a study.

Select the most appropriate answer.

6) The following statement refers to which aspect of a statistical study: "A meteorologist constructs a graph showing the total precipitation in Phoenix, Arizona in each of the months of a given year"?
A) Description B) Design C) Inference
Answer: A
Objective: (1) Identify aspects of a study.

7) The following statement refers to which aspect of a statistical study: "The average age of the students in a statistics class is 25 years"?
A) Description B) Design C) Inference
Answer: A
Objective: (1) Identify aspects of a study.

8) The following statement refers to which aspect of a statistical study: "Based on a study of 25 hospitals nationwide, researchers have concluded that there is a relationship between smoking cigarettes and contracting emphysema"?
A) Inference B) Design C) Description
Answer: A
Objective: (1) Identify aspects of a study.

9) The following statement refers to which aspect of a statistical study: "Based on previous clients, a marriage counselor concludes that the majority of marriages that begin with cohabitation before marriage will result in divorce"?
A) Inference B) Design C) Description
Answer: A
Objective: (1) Identify aspects of a study.

10) Planning the methods for data collection to study the effects of Vitamin E on athletic strength would be classified as which aspect of statistics: design, inference or description?
A) Design B) Description C) Inference
Answer: A
Objective: (1) Identify aspects of a study.

11) The following statement refers to which aspect of a statistical study: "From past figures, it is predicted that 47% of the registered voters in Virginia will vote in the June primary"?
A) Inference B) Design C) Description
Answer: A
Objective: (1) Identify aspects of a study.

12) Planning how to obtain data to answer the questions of interest refers to which aspect of statistics?
A) Design
B) Description
C) Inference
D) Sampling
E) None of these
Answer: A
Objective: (1) Identify aspects of a study.

13) Summarizing the data that are obtained refers to which aspect of statistics?
A) Description
B) Design
C) Inference
D) Sampling
E) None of these
Answer: A
Objective: (1) Identify aspects of a study.

14) Making decisions and predictions based on the data refers to which aspect of statistics?
A) Inference
B) Description
C) Design
D) Sampling
E) None of these
Answer: A
Objective: (1) Identify aspects of a study.

The owners of a coffee shop conducted a taste test to determine whether its customers preferred a new coffee brand to the current one sold by the shop. Customers who were willing to participate were given small samples of each of the two brands in random order and were asked to select which one they preferred without knowing the brand. Of the 100 participating customers, 90% chose the new brand. Based on these results, the owners determined that a majority of their customers preferred the new brand and therefore switched their coffee supplier.
15) Predicting the preference of all of the coffee shop customers based on the taste test results refers to which aspect of statistics?
A) Inference
B) Description
C) Design
D) Investigation
E) None of these
Answer: A
Objective: (1) Identify aspects of a study.

16) Randomizing the order in which the samples of each brand were given to each customer refers to which aspect of statistics?
A) Design
B) Inference
C) Description
D) Investigation
E) None of these
Answer: A
Objective: (1) Identify aspects of a study.

17) Stating that 90% of the taste testers preferred the new brand is an example of which type of statistics?
A) Description
B) Design
C) Inference
D) Investigation
E) None of these
Answer: A
Objective: (1) Identify aspects of a study.

2 Identify and compare percentages.

Provide an appropriate response.

1) In a survey of students at a certain college, 677 students were asked whether they believed in life after death and whether they believed in reincarnation. The results are summarized in the table below.

Frequency Distribution

Response	Reincarnation (Column Percent, Number of cases)	Life After Death (Column Percent, Number of cases)
Yes, Definitely	9.5%, 64	37.7%, 255
Yes, Probably	19.9%, 135	31.0%, 210
No, Probably Not	43.0%, 291	19.2%, 130
No, Definitely Not	27.6%, 187	12.1%, 82
Column Total	100.0%, 677	100.0%, 677

When asked whether or not they believed in life after death, what percentage of those surveyed said yes, definitely; yes, probably; no, probably not; and no, definitely not?

A) 37.7%, 31.0%, 19.2%, 12.1% B) 9.5%, 19.9%, 43.0%, 27.6%

C) 255%, 210%, 130%, 82% D) 9.5%, 37.7%, 19.9%, 31.0%

Answer: A

Objective: (2) Identify and compare percentages.

2) In a survey of students at a certain college, 677 students were asked whether they believed in life after death and whether they believed in reincarnation. The results are summarized in the table below.

Frequency Distribution

Response	Reincarnation (Column Percent, Number of cases)	Life After Death (Column Percent, Number of cases)
Yes, Definitely	9.5%, 64	37.7%, 255
Yes, Probably	19.9%, 135	31.0%, 210
No, Probably Not	43.0%, 291	19.2%, 130
No, Definitely Not	27.6%, 187	12.1%, 82
Column Total	100.0%, 677	100.0%, 677

How does the percentage of "yes, definitely" responses for life after death compare with the percentage of "yes, definitely" responses for reincarnation?
A) The percentage who responded "yes, definitely" when asked whether they believed in life after death (37.7%) was roughly four times the percentage who responded "yes, definitely" when asked whether the believed in reincarnation (9.5%).
B) The percentage who responded "yes, definitely" when asked whether they believed in life after death (37.7%) was slightly higher than the percentage who responded "yes, definitely" when asked whether the believed in reincarnation (31.0%).
C) The percentage who responded "yes, definitely" when asked whether they believed in life after death (19.9%) was roughly twice the percentage who responded "yes, definitely" when asked whether the believed in reincarnation (9.5%).
D) The percentage who responded "yes, definitely" when asked whether they believed in life after death (255%) was roughly four times the percentage who responded "yes, definitely" when asked whether the believed in reincarnation (64%).
Answer: A
Objective: (2) Identify and compare percentages.

1.2 Sample Versus Population

1 Answer questions about the sample and population involved in a given study.

Answer true or false.
1) Based on 12,000 responses from 55,000 questionnaires sent to its alumni, a major university estimated that the annual salary of its alumni was $98,500 per year. This sample was collected using random sampling.
A) False B) True
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

2) A lobbyist for a major airspace firm assigns a number to each legislator and then uses a computer to randomly generate ten numbers. The lobbyist contacts the legislators corresponding to these numbers. This technique produces a random sample.
A) True B) False
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

3) A lobbyist for a major airspace firm wants to get the opinion of state legislators on an issue of importance to his industry. The lobbyist contacts 10 state legislators at a restaurant during a lunch break and each one is polled about the issue. This technique produces a random sample.
A) False B) True
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

4) A large company would like to know what percentage of it's employees would utilize an on-site daycare facility. Rather than asking it's 5000 employees, each employee is assigned a number from 1 to 5000 and then a computer program is used to randomly select 100 numbers between 1 and 5000. These employees are then contacted and asked whether they would use an on-site daycare facility. The method used to collect the sample of 100 responses produces a random sample.
A) True B) False
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

5) To estimate the percentage of defective items a machine is producing, a quality control analyst inspects the first 100 items produced in a day. This technique produces a random sample.
A) False B) True
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

6) In a recent television survey, participants were asked to answer "yes" or "no" to the question "Are you in favor of the death penalty?". Six thousand five hundred responded "yes" while 3700 responded "no". There was a fifty-cent charge for the call. The sampling technique used produces a random sample.
A) False B) True
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

Determine whether the summary measure is better described as a parameter or a statistic.
7) The average height of horse jockeys
A) Parameter B) Statistic
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

8) The average GPA of a university's graduating class
A) Parameter B) Statistic
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

9) The proportion of teenagers in a nationwide survey who stated that they consumed alcohol on a regular basis
A) Statistic B) Parameter
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

10) The average weekly amount spent on groceries by a family of four based on a random sample of 100 families
A) Statistic B) Parameter
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

11) The proportion of regional flights for that were considered on-time during the month of December
A) Parameter B) Statistic
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

12) The average age of 25 congressional members who were selected at random from all U.S. Congressmen
A) Statistic B) Parameter
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

Select the most appropriate answer.
13) A survey of American adults asks "would you like to see more or less government spending on natural disasters?" Of the 1496 respondents, 723 responded "more" or "much more". The population of interest consists of
A) all American adults
B) the 1496 respondents
C) the 723 respondents who responded "more" or "much more"
D) the proportion of American adults who would respond "more" or "much more"
E) the proportion of respondents who responded "more" or "much more"
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

14) A survey of American adults asks "would you like to see more or less government spending on natural disasters?" Of the 1496 respondents, 723 responded "more" or "much more". The sample consists of
A) the 1496 respondents
B) all American adults
C) the 723 respondents who responded "more" or "much more"
D) the proportion of American adults estimated to respond "more" or "much more"
E) the proportion of respondents who responded "more" or "much more"
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

15) A survey of American adults asks "would you like to see more or less government spending on natural disasters?" Of the 1496 respondents, 723 responded "more" or "much more". The parameter of interest consists of
A) all American adults
B) the 1496 respondents
C) the 723 respondents who responded "more" or "much more"
D) the proportion of American adults estimated to respond "more" or "much more"
E) the proportion of respondents who responded "more" or "much more"
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

16) A survey asks "would you like to see more or less government spending on natural disasters?" Of the 1496 respondents, 723 responded "more" or "much more". The proportion of respondents who responded "more" or "much more" is an example of
A) a statistic
B) the population
C) the sample
D) random sampling
E) a parameter
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

17) In a survey of American adults, 71% of 1052 adults polled answered "Yes" to the question "Do you believe the theory that increased carbon dioxide and other gases released into the atmosphere will, if unchecked, lead to global warming and an increase in average temperatures?" The sample consists of
A) all American adults
B) the 1052 polled adults
C) the proportion of adults polled who responded "Yes" to the question
D) the proportion of American adults who are predicted to respond "Yes" to the question
E) 71% of American adults
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

18) In a survey, 71% of 1052 adults polled answered "Yes" to the question "Do you believe the theory that increased carbon dioxide and other gases released into the atmosphere will, if unchecked, lead to global warming and an increase in average temperatures?" The 71% responding "Yes" to the question is an example of
A) a statistic
B) the population
C) the sample
D) random sampling
E) a parameter
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

19) In a survey, 71% of 1052 adults polled nationwide answered "Yes" to the question "Do you believe the theory that increased carbon dioxide and other gases released into the atmosphere will, if unchecked, lead to global warming and an increase in average temperatures?" The predicted proportion of all American adults who would respond "Yes" to the question is an example of
A) a parameter
B) the population
C) the sample
D) random sampling
E) a statistic
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

20) 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-four hundred Florida residents were surveyed. Which of the following is the population of interest in the study?
A) all Florida residents
B) the 2400 Florida residents surveyed
C) all Florida residents who were willing to spend more tax dollars on protecting the beaches from environmental disasters.
D) all Florida residents who lived along the beaches
E) all residents of the United States
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

21) A high school guidance counselor analyzed data from a sample of 600 community colleges taken throughout the United States. One of his goals was to estimate the annual tuition costs of community colleges in the United States. Identify the population of interest to the guidance counselor.
A) all community colleges in the United States
B) all high school guidance counselors
C) the sample of 600 community colleges
D) all students attending community colleges in the United States
E) the average annual tuition cost of community colleges in the United States
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

22) A manufacturer of cellular phones has decided that an assembly line is operating satisfactorily if less than 6% of the phones manufactured per day are defective. To check the quality of a day's production, the company decides to randomly sample 30 phones from a day's production and test for defects. Define the population of interest to the manufacturer.
A) all cellular phones manufactured during the day in question
B) the 30 cellular phones that were sampled and tested
C) the 30 responses: defective or not defective
D) the 6% of the cellular phones that were defective
E) all defective cellular phones manufactured by the company
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

23) A manufacturer of cellular phones has decided that an assembly line is operating satisfactorily if less than 6% of the phones manufactured per day are defective. To check the quality of a day's production, the company decides to randomly sample 30 phones from a day's production and test for defects. Define the sample of interest to the manufacturer.
A) the 30 cellular phones that were sampled and tested
B) all cellular phones manufactured during the day in question
C) the 30 responses: defective or not defective
D) the 6% of the cellular phones that were defective
E) all defective cellular phones manufactured by the company
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

Provide an appropriate response.
24) A recent poll asked 1,245 registered voters, "Which party do you think can do a better job of handling immigration issues?" 31% of the respondents answered "Republicans." With a margin of error of ±3%, it is estimated that between 28 and 34 percent of registered voters nationwide feel that the Republican party can do a better job of handling immigration issues. Does the proportion of the sample answering "Republican" refer to a statistic or a parameter?
A) statistic B) parameter
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

25) The average salary of all automotive workers is $42,000. Is this value a parameter or a statistic?
A) parameter B) statistic
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

26) A survey of 1103 students was taken from a college with 14,500 students. Is this value a parameter or a statistic?
A) statistic B) parameter
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

27) A survey of 1500 American households found that 33% of the respondents own a computer. Is this value a parameter or a statistic?
A) statistic B) parameter
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

28) Every year the U.S. Census Bureau releases the Current Population Report based on a survey of 50,000 households within the United States. Identify the population, the sample, and the subjects for this study.
Answer: The population is all U.S. households. The sample is the 50,000 households surveyed within the United States. The subjects are the U.S. households.
Objective: (1) Answer questions about the sample and population involved in a given study.

29) A survey of 1004 Americans, 18 years and older, were asked, "Do you have a great deal of concern regarding global warming (the greenhouse effect)?" Identify the population, the sample, and the subjects for this study.
Answer: The population is all Americans 18 years and older. The sample is the 1004 surveyed Americans 18 years and older. The subjects are Americans 18 years and older.
Objective: (1) Answer questions about the sample and population involved in a given study.

30) A survey of 1500 American households found that 33% of the households own a computer. Identify the population.
A) The collection of all American households
B) The 1500 American households surveyed
C) 33% of American households
D) All American households owning a computer
E) The 33% of the 1500 households sampled that own a computer
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

31) A survey of 1500 American households found that 33% of the households own a computer. Identify the sample.
A) The 1500 American households surveyed
B) The collection of all American households
C) 33% of American households
D) All American households owning a computer
E) The 33% of the 1500 households sampled that own a computer
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

32) A study is conducted on a group of 10,000 high school students nationwide. Each week for a period of 13 weeks, participants must take an unannounced urine analysis to screen for illegal drug use. Identify the subjects of the study.
A) high school students
B) 13 weeks
C) researchers
D) illegal drugs detected by urine analysis
E) the urine analysis test
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

33) The average salary of all automotive workers is $42,000. Identify a subject of the study.
A) An automotive worker
B) The average salary of an automotive worker
C) All automotive wage workers
D) All automotive salary workers
E) All U.S. managers
Answer: A
Objective: (1) Answer questions about the sample and population involved in a given study.

2 Interpret the results of a survey.

Answer true or false.

1) Descriptive statistics refers to methods of making decisions or predictions about a population, based on data obtained from a sample of that population; while inferential statistics refers to methods for summarizing the data.

A) False B) True

Answer: A

Objective: (2) Interpret the results of a survey.

Select the most appropriate answer.

2) The estimation of the population average age of registered voters in the state of Ohio based on the sample average age of 1,000 registered voters in that state and its corresponding margin of error is an example of

A) inferential statistics.

B) descriptive statistics.

C) a parameter.

D) a statistic.

E) deductive statistics.

Answer: A

Objective: (2) Interpret the results of a survey.

Provide an appropriate response.

3) A recent poll asked 1,245 registered voters nationwide, "Which party do you think can do a better job of handling immigration issues?" 31% of the respondents answered "Republicans." With a margin of error of ±3%, it is estimated that between 28 and 34 percent of registered voters nationwide feel that the Republican party can do a better job of handling immigration issues. Identify which part of this example is inferential.

Answer: It is estimated that between 28 and 34 percent of registered voters nationwide feel that the Republican party can do a better job of handling immigration issues.

Objective: (2) Interpret the results of a survey.

4) A recent poll asked 1,245 registered voters nationwide, "Which party do you think can do a better job of handling immigration issues?" 31% of the respondents answered "Republicans." With a margin of error of ±3%, it is estimated that between 28 and 34 percent of registered voters nationwide feel that the Republican party can do a better job of handling immigration issues. Identify which part of this example is descriptive.

Answer: 31% of the respondents answered "Republican"

Objective: (2) Interpret the results of a survey.

5) Drug X, a drug that claims to treat male pattern scalp hair loss, was administered for 12 months to over 1800 men aged 18 to 41 with mild to moderate amounts of ongoing hair loss. Whether they were receiving drug X or a placebo (a pill containing no medication), all men were given a medicated shampoo. In general, men who took drug X maintained or increased the number of visible scalp hairs; while scalp hair counts in men who took the placebo continued to decrease. This concluded that drug X is effective in maintaining or increasing the amount of scalp hair in men. Which statement of this example can be referred to as descriptive statistics?
Answer: Scalp hair counts in the men who took drug X maintained or increased; scalp hair counts in the men who took the placebo decreased
Objective: (2) Interpret the results of a survey.

6) Drug X, a drug that claims to treat male pattern scalp hair loss, was administered for 12 months to over 1800 men aged 18 to 41 with mild to moderate amounts of ongoing hair loss. Whether they were receiving drug X or a placebo (a pill containing no medication), all men were given a medicated shampoo. In general, men who took drug X maintained or increased the number of visible scalp hairs; while, scalp hair counts in men who took the placebo continued to decrease. This concluded that drug X is effective in maintaining or increasing the amount of scalp hair in men. For this example, which statement can be characterized as inferential statistics?
Answer: Drug X is effective in maintaining or increasing the amount of scalp hair in men.
Objective: (2) Interpret the results of a survey.

7) From past figures, it is predicted that 39% of the registered voters in California will vote in the June primary. Which type of statistics does this statement reflect: inferential statistics or descriptive statistics?
A) Inferential statistics B) Descriptive statistics
Answer: A
Objective: (2) Interpret the results of a survey.

3 Find and interpret the margin of error.

Provide an appropriate response.

1) A polling service usually selects samples of size 5000 when conducting its polls. Estimate the corresponding margin of error.
A) 0.014
B) 0.0002
C) 0.045
D) 0.002
E) cannot be determined from the information given
Answer: A
Objective: (3) Find and interpret the margin of error.

2) A survey of 1004 American adults 18 years and older were asked, "Do you have a great deal of concern regarding global warming (the greenhouse effect)?" Of the 1004 adults surveyed, 40% said they worried about global warming a great deal.

a. Is the study above an example of an observational study or an experimental study? Explain.
b. Calculate an approximate margin of error for the results of this study. How is it interpreted?
Answer: a. The study above is an example of an observational study, because the study merely observed the values on the response variable and the explanatory variables for the sampled subjects without doing anything to them; b. The margin of error was approximately 3%. This suggests that between 37% and 43% of all American adults 18 years and older worry about global warming a great deal.
Objective: (3) Find and interpret the margin of error.

4 Concepts

Fill in the blank.

1) The entities that we measure in a study are called the _____.
A) subjects B) statistics C) parameters D) data
Answer: A
Objective: (4) Concepts

2) The basis of _____ is that each subject in the population has the same chance of being in the sample.
Answer: random sampling
Objective: (4) Concepts

3) The _____ is the set of all subjects of interest; while the _____ is a subset of the entire set of interest.
Answer: population; sample
Objective: (4) Concepts

4) A _____ is a numerical summary of the population; while a _____ is a numerical summary of the sample.
Answer: parameter; statistic
Objective: (4) Concepts

Answer true or false.

5) Parameter values are usually known.
A) False B) True
Answer: A
Objective: (4) Concepts

6) Inferential statistics are used when data are available only for a sample; however, descriptive statistics are used when data are available for either a sample or a population.
A) True B) False
Answer: A
Objective: (4) Concepts

7) Random sampling enables the sample to be a good reflection of the population.
A) True B) False
Answer: A
Objective: (4) Concepts

Select the most appropriate answer.

8) A summary measure that is computed from a sample to describe a characteristic of interest is called
A) a statistic
B) a data file
C) a poll
D) a parameter
E) a variable
Answer: A
Objective: (4) Concepts

9) A numerical summary of the population is called
A) a parameter
B) a census
C) a data file
D) a statistic
E) a variable
Answer: A
Objective: (4) Concepts

10) Using statistics to draw conclusions about parameters is called
A) inferential statistics.
B) sampling.
C) designing a study.
D) descriptive statistics.
E) deductive statistics.
Answer: A
Objective: (4) Concepts

Provide an appropriate response.

11) Define the terms population and sample.
Answer: The population is the total set of subjects in which we are interested. A sample is the subset of the population for whom we have data.
Objective: (4) Concepts

12) Briefly explain why survey data typically consists of data for a sample and not for an entire population.
Answer: Collecting data for an entire population, a census, is too expensive and time consuming.
Objective: (4) Concepts

1.3 Organizing Data, Statistical Software, and the New Field of Data Science

1 Construct a data file.

1) Which of these could be a data file for two characteristics for a sample of four people, where one characteristic takes numerical values and the other takes values that are categories?

A)

Class	Daily Screen Time (hours)
Fresh	5.5
Soph	4.75
Fresh	9.25
Jr	7.0

B)

Class	Grade
Fresh	A-
Fresh	B
Soph	C+
Jr	B+

C)

Age	Daily Screen Time (hours)
18	7.25
19	5.25
18	3.75
18	9.5

D)

Movie Genre	Number that Enjoy
Comedy	4
Drama	3
Horror	2
Sci-Fi	2

Answer: A
Objective: (1) Construct a data file.

2 Clean up messy data.

1) A postal inspector has the following data set for a sample of three packages.

Package	Length	Width	Weight	Domestic or International
1	10 in.	20 cm	2.7 lb	Domestic
2	27 cm	22 cm	1.3 lb	International
3	18 cm	14 cm	0.5 lb	Domestic

What must be done to clean up these data?
A) Convert the lengths to all be in inches or all in centimeters
B) Delete the "Package" column
C) Convert the lengths, widths, and weights to all use standard units or all use metric units
D) Change to use 0 or 1 to represent "Domestic" and the other to represent "International"
Answer: A
Objective: (2) Clean up messy data.

3 Identify training data and test data.

1) A researcher wants to create a device to detect the presence of explosive materials in closed luggage as they pass by on a conveyor belt. Which of these data sets would be useful as training data?
A) A data set that contains whether the device detected explosive materials in each piece of luggage and whether the luggage contained explosive materials
B) A data set that contains whether the device detected explosive materials in each piece of luggage
C) A data set that contains whether each piece of luggage contained explosive materials
D) A data set that contains whether the device detected explosive materials in randomly selected pieces of luggage that may or may not contain explosive materials
Answer: A
Objective: (3) Identify training data and test data.

4 Answer questions on ethics and bias in data analysis.

1) A center for the performing arts wants to identify which types of performances would maximize revenue. They use a data set that includes, for each of 100 randomly selected sustaining members for each of last 137 performances, whether the member attended, the individual ticket price, the date, the genre, and the local weather conditions. For which of these reasons could the data lead to algorithmic bias?
A) The sample includes only sustaining members.
B) The data include only the local weather conditions.
C) The data do not include the number of times each was performed.
D) The performances are not randomly selected.
Answer: A
Objective: (4) Answer questions on ethics and bias in data analysis.