

Applied Statistics in Business and Economics, 6e (Doane)

Chapter 2 Data Collection

- 1) Categorical data have values that are described by words rather than numbers.
- 2) Numerical data can be either discrete or continuous.
- 3) Categorical data are also referred to as nominal or qualitative data.
- 4) The number of checks processed at a bank in a day is an example of categorical data.
- 5) The number of planes per day that land at an airport is an example of discrete data.
- 6) The weight of a bag of dog food is an example of discrete data.
- 7) In last year's annual report, Thompson Distributors indicated that it had 12 regional warehouses. This is an example of ordinal level data.
- 8) Nominal data refer to data that can be ordered in a natural way.
- 9) This year, Oxnard University produced two football All-Americans. This is an example of continuous data.
- 10) The type of statistical test that we can perform is independent of the level of measurement of the variable of interest.
- 11) Your weight recorded at your annual physical would *not* be ratio data, because you cannot have zero weight.
- 12) The level of measurement for categorical data is nominal.
- 13) Temperature measured in degrees Fahrenheit is an example of interval data.
- 14) The closing price of a stock is an example of ratio data.
- 15) The *Statistical Abstract of the United States* is a huge annual compendium of data for the United States, and it is available online free of charge.
- 16) Ordinal data can be treated as if it were nominal data but not vice versa.
- 17) Responses on a seven-point Likert scale are usually treated as ratio data.
- 18) Likert scales are especially important in opinion polls and marketing surveys.
- 19) Ordinal data are data that can be ranked based on some natural characteristic of the items.
- 20) Ratio data are distinguished from interval data by the presence of a zero reference point.

- 21) It is better to attempt a census of a large population instead of relying on a sample.
- 22) Judgment sampling and convenience sampling are nonrandom sampling techniques.
- 23) A problem with judgment sampling is that the sample may not reflect the population.
- 24) When the population is large, a sample estimate is usually preferable to a census.
- 25) Sampling error is avoidable by choosing the sample scientifically.
- 26) A sampling frame is used to identify the target population in a statistical study.
- 27) By taking a systematic sample, in which we select every 50th shopper arriving at a specific store, we are approximating a random sample of shoppers.
- 28) A worker collecting data from every other shopper who leaves a store is taking a simple random sample of customer opinion.
- 29) Creating a list of people by taking the third name listed on every 10th page of the phone book is an example of convenience sampling.
- 30) Internet surveys posted on popular websites have no bias since anyone can reply.
- 31) Analysis of month-by-month changes in stock market prices during the most recent recession would require the use of time series data.
- 32) A cluster sample is a type of stratified sample that is based on geographical location.
- 33) An advantage of a systematic sample is that no list of enumerated data items is required.
- 34) Telephone surveys often have a low response rate and fail to reach the desired population.
- 35) Mail surveys are attractive because of their high response rates.
- 36) A problem with convenience sampling is that the target population is not well-defined.
- 37) If you randomly sample 50 students about their favorite places to eat, the data collected would be referred to as cross-sectional data.
- 38) The number of FedEx shipping centers in each of 50 cities would be ordinal level data.
- 39) Internet surveys posted on popular websites suffer from nonresponse bias.
- 40) Different variables are usually shown as *columns* of a multivariate data set.
- 41) Each *row* in a multivariate data matrix is an observation (e.g., an individual response).

- 42) A bivariate data set has only two observations on a variable.
- 43) Running times for 3,000 runners in a 5k race would be a multivariate data set.
- 44) Running times for 500 runners in a 5k race would be a univariate data set.
- 45) A list of the salaries, ages, and years of experience for 50 CEOs is a multivariate data set.
- 46) The daily closing price of Apple stock over the past month would be a time series.
- 47) The number of words on 50 randomly chosen textbook pages would be cross-sectional data.
- 48) A Likert scale with an even number of scale points between "Strongly Agree" and "Strongly Disagree" is intended to prevent "neutral" choices.
- 49) Private statistical databases (e.g., CRSP) are usually free.
- 50) An investment firm rates bonds for Aard Co Inc. as "B+," while bonds of Deva Corp. are rated "AA." Which level of measurement would be appropriate for such data?
- A) Nominal
 - B) Ordinal
 - C) Interval
 - D) Ratio
- 51) Which variable is *least* likely to be regarded as ratio data?
- A) Length of time required for a randomly chosen vehicle to cross a toll bridge (minutes)
 - B) Weight of a randomly chosen student (pounds)
 - C) Number of fatalities in a randomly chosen traffic disaster (persons)
 - D) Student's evaluation of a professor's teaching (Likert scale)
- 52) Which of the following is numerical data?
- A) Your gender
 - B) The brand of cell phone you own
 - C) Whether you have an American Express card
 - D) The fuel economy (MPG) of your car
- 53) Measurements from a sample are called
- A) statistics.
 - B) inferences.
 - C) parameters.
 - D) variables.

- 54) Quantitative variables use which two levels of measurement?
- A) Ordinal and ratio
 - B) Interval and ordinal
 - C) Nominal and ordinal
 - D) Interval and ratio
- 55) Temperature in degrees Fahrenheit is an example of a(n) _____ variable.
- A) nominal
 - B) ordinal
 - C) interval
 - D) ratio
- 56) Using a sample to make generalizations about an aspect of a population is called
- A) data mining.
 - B) descriptive statistics.
 - C) random sampling.
 - D) statistical inference.
- 57) Your telephone area code is an example of a(n) _____ variable.
- A) nominal
 - B) ordinal
 - C) interval
 - D) ratio
- 58) Which is *least* likely to be regarded as a ratio variable?
- A) A critic's rating of a restaurant on a 1 to 4 scale
 - B) Automobile exhaust emission of nitrogen dioxide (milligrams per mile)
 - C) Number of customer complaints per day at a cable TV company office
 - D) Cost of an eBay purchase
- 59) Automobile exhaust emission of CO₂ (milligrams per mile) is _____ data.
- A) nominal
 - B) ordinal
 - C) interval
 - D) ratio
- 60) Your rating of the food served at a local restaurant using a three-point scale of 0 = gross, 1 = decent, 2 = yummy is _____ data.
- A) nominal
 - B) ordinal
 - C) interval
 - D) ratio

- 61) The number of passengers "bumped" on a particular airline flight is _____ data.
- A) nominal
 - B) ordinal
 - C) interval
 - D) ratio
- 62) Which should *not* be regarded as a continuous random variable?
- A) Tonnage carried by a randomly chosen oil tanker at sea
 - B) Wind velocity at 7 o'clock this morning
 - C) Number of personal fouls by the Miami Heat in a game
 - D) Length of time to play a Wimbledon tennis match
- 63) Which of the following is *not* true?
- A) Categorical data have values that are described by words rather than numbers.
 - B) Categorical data are also referred to as nominal or qualitative data.
 - C) The number of checks processed at a bank in a day is categorical data.
 - D) Numerical data can be either discrete or continuous.
- 64) Which of the following is true?
- A) The type of charge card used by a customer (Visa, MasterCard, AmEx) is ordinal data.
 - B) The duration (minutes) of a flight from Boston to Minneapolis is ratio data.
 - C) The number of Nobel Prize-winning faculty at Oxnard University is continuous data.
 - D) The number of regional warehouses owned by Jankord Industries is ordinal data.
- 65) Which statement is *correct*?
- A) Judgment sampling is preferred to systematic sampling.
 - B) Sampling without replacement introduces bias in our estimates of parameters.
 - C) Cluster sampling is useful when strata characteristics are unknown.
 - D) Focus groups usually work best without a moderator.
- 66) A Likert scale
- A) yields interval data if scale distances are equal.
 - B) must have an odd number of scale points.
 - C) must have a verbal label on each scale point.
 - D) is rarely used in marketing surveys.
- 67) Which is most nearly correct regarding sampling error?
- A) It can be eliminated by increasing the sample size.
 - B) It cannot be eliminated by any statistical sampling method.
 - C) It can be eliminated by using Excel's =RANDBETWEEN() function.
 - D) It can be eliminated by utilizing systematic random sampling.
- 68) Which statement is *false*?
- A) Random dialing phone surveys have low response and are poorly targeted.
 - B) Selection bias means that many respondents dislike the interviewer.
 - C) Simple random sampling requires a list of the population.
 - D) Web surveys are economical but suffer from nonresponse bias.

- 69) Judgment sampling is sometimes preferred over random sampling, for example, when
- A) the desired sample size is much larger than the population.
 - B) the sampling budget is large and the population is conveniently located.
 - C) time is short and the sampling budget is limited.
 - D) the population is readily accessible and sampling is nondestructive.
- 70) An advantage of convenience samples is that
- A) the required sample size is easier to calculate.
 - B) sampling error can be reduced.
 - C) computation of statistics is easier.
 - D) they are often quicker and cheaper.
- 71) Before deciding whether to assess heavy fines against noisy airlines, which sampling method would the Federal Aviation Administration *probably* use to measure the peak noise from departing jets as measured by a ground-level observer at a point one mile from the end of the departure runway?
- A) Radio survey of pilots.
 - B) Simple random sample.
 - C) Judgment sample.
 - D) Stratified sample.
- 72) Professor Hardtack chose a sample of 7 students from his statistics class of 35 students by picking every student who was wearing red that day. Which kind of sample is this?
- A) Simple random sample
 - B) Judgment sample
 - C) Systematic sample
 - D) Convenience sample
- 73) Thirty work orders are selected from a filing cabinet containing 500 work order folders by choosing every 15th folder. Which sampling method is this?
- A) Simple random sample
 - B) Systematic sample
 - C) Stratified sample
 - D) Cluster sample
- 74) Which of the following is *not* a likely reason for sampling?
- A) The destructive nature of certain tests
 - B) The physical impossibility of checking all the items in the population
 - C) Prohibitive cost of studying the entire population
 - D) The expense of obtaining random numbers
- 75) Comparing a census of a large population to a sample drawn from it, we expect that the
- A) sample is usually a more practical method of obtaining the desired information.
 - B) accuracy of the observations in the census is surely higher than in the sample.
 - C) sample must be a large fraction of the population to be accurate.

- 76) A stratified sample is sometimes recommended when
- A) the sample size is very large.
 - B) the population is small compared to the sample.
 - C) distinguishable strata can be identified in the populations.
 - D) the population is spread out geographically.
- 77) A *random sample* is one in which the
- A) probability that an item is selected for the sample is the same for all population items.
 - B) population items are selected haphazardly by experienced workers.
 - C) items to be selected from the population are specified based on expert judgment.
 - D) probability of selecting a population item depends on the item's data value.
- 78) An advantage of convenience samples over random samples is that
- A) they are easy to analyze.
 - B) it is easier to determine the sample size needed.
 - C) it is easier to calculate the sampling errors involved.
 - D) data collection cost is reduced.
- 79) To measure satisfaction with its cell phone service, AT&T takes a stratified sample of its customers by age, gender, and location. Which is an advantage of this type of sampling, as opposed to other sampling methods?
- A) It is less intrusive on customers' privacy.
 - B) It does not require random numbers.
 - C) It gives faster results.
 - D) It can give more accurate results.
- 80) An accounting professor wishing to know how many MBA students would take a summer elective in international accounting did a survey of the class she was teaching. Which kind of sample is this?
- A) Simple random sample
 - B) Cluster sample
 - C) Systematic sample
 - D) Convenience sample
- 81) A binary variable (also called a dichotomous variable or dummy variable) has
- A) only two possible values.
 - B) continuous scale values.
 - C) rounded data values.
 - D) ordinal or interval values.
- 82) A population has groups that have a small amount of variation within them, but large variation among or between the groups themselves. The proper sampling technique is
- A) simple random.
 - B) stratified.
 - C) cluster.
 - D) judgment.

83) A manager chose two people from her team of eight to give an oral presentation because she felt they were representative of the whole team's views. What sampling technique did she use in choosing these two people?

- A) Convenience
- B) Simple random
- C) Judgment
- D) Cluster

84) Sampling bias can best be reduced by

- A) using appropriate data coding.
- B) having a computer tabulate the results.
- C) utilizing random sampling.
- D) taking a judgment sample.

85) A sampling technique used when groups are defined by their geographical location is

- A) cluster sampling.
- B) convenience sampling.
- C) judgment sampling.
- D) random sampling.

86) If we choose 500 random numbers using Excel's function =RANDBETWEEN(1,99), we would *most likely* find that

- A) numbers near the mean (50) would tend to occur more frequently.
- B) numbers near 1 and 99 would tend to occur less frequently.
- C) some numbers would occur more than once.
- D) the numbers would have a clear pattern.

87) A problem with nonrandom sampling is that

- A) larger samples need to be taken to reduce the sampling error inherent in this approach.
- B) not every item in the population has the same chance of being selected, as it should.
- C) it is usually more expensive than random sampling.
- D) it generally provides lower response rates than random sampling.

88) From its 32 regions, the FAA selects 6 regions, and then randomly audits 25 departing commercial flights in each region for compliance with legal fuel and weight requirements. This is an example of

- A) simple random sampling.
- B) stratified random sampling.
- C) cluster sampling.
- D) judgment sampling.

89) Which of the following is a *correct* statement?

- A) Choosing the third person listed on every fifth page of the phone book is stratified sampling.
- B) An advantage of a systematic sample is that no list of enumerated data items is required.
- C) Convenience sampling is used to study shoppers in convenience stores.
- D) Judgment sampling is an example of true random sampling.

- 90) Which of the following is *false*?
- A) Sampling error is the difference between the true parameter and the sample estimate.
 - B) Sampling error is a result of unavoidable random variation in a sample.
 - C) A sampling frame is chosen from the target population in a statistical study.
 - D) The target population must first be defined by a full list or data file of all individuals.
- 91) When we are choosing a random sample and we do not place chosen units back into the population, we are
- A) sampling with replacement.
 - B) sampling without replacement.
 - C) using a systematic sample.
 - D) using a voluntary sample.
- 92) Which method is likely to be used by a journalism student who is casually surveying opinions of students about the university's cafeteria food for an article that she is writing?
- A) Simple random sample
 - B) Systematic random sample
 - C) Cluster sample
 - D) Convenience sample
- 93) Which of the following is *false*?
- A) Mail surveys are cheap but have low response rates.
 - B) Coverage error is when respondents give untruthful answers.
 - C) Focus groups are nonrandom but can probe issues more deeply.
 - D) Surveys posted on popular websites suffer from selection bias.
- 94) Which is a time series variable?
- A) VISA balances of 30 students on December 31 of this year
 - B) Net earnings reported by Xena Corp. for the last 10 quarters
 - C) Dollar exchange rates yesterday against 10 other world currencies
 - D) Titles of the top 10 movies in total revenue last week
- 95) An *observation* in a data set would refer to
- A) only a variable whose value is recorded by visual inspection.
 - B) a data item whose value is numerical (as opposed to categorical).
 - C) a single row that contains one or more observed variables.
 - D) the values of all the variables in the entire data set.
- 96) A *multivariate* data set contains
- A) more than two observations.
 - B) more than two categorical variables.
 - C) more than two variables.
 - D) more than two levels of measurement.

97) The Centers for Disease Control and Prevention (CDC) wants to estimate the average extra hospital stay that occurs when heart surgery patients experience postoperative atrial fibrillation. They divide the United States into nine regions. In each region, hospitals are selected at random within each hospital size group (small, medium, large). In each hospital, heart surgery patients are sampled according to known percentages by age group (under 50, 50 to 64, 65 and over) and gender (male, female). This procedure combines which sampling methods?

- A) Systematic, simple random, and convenience
- B) Convenience, systematic, and judgment
- C) Cluster, stratified, and simple random
- D) Judgment, systematic, and simple random

98) Which statement is correct?

- A) Selecting every fifth shopper arriving at a store will approximate a random sample of shoppers.
- B) Selecting only shoppers who drive SUVs is a stratified sampling method.
- C) A census is preferable to a sample for most business problems.
- D) Stratified samples are usually cheaper than other methods.

99) Which is a categorical variable?

- A) The brand of jeans you usually wear
- B) The price you paid for your last pair of jeans
- C) The distance to the store where you purchased your last pair of jeans
- D) The number of pairs of jeans that you own

100) Which is a discrete variable?

- A) The time it takes to put on a pair of jeans
- B) The price you paid for your last pair of jeans
- C) The distance to the store where you purchased your last pair of jeans
- D) The number of pairs of jeans that you own

101) A section of the population we have targeted for analysis is

- A) a statistic.
- B) a frame.
- C) a sample.
- D) a coven.

102) Which is *not* a time series variable?

- A) Closing checkbook balances of 30 students on December 31 of this year
- B) Net earnings reported by Xena Corp. for the last 10 quarters
- C) Dollar/euro exchange rates at 12 noon GMT for the last 30 days
- D) Movie attendance at a certain theater for each Saturday last year

103) A good Likert scale may *not* have

- A) unequal distances between scale points.
- B) an odd number of scale points.
- C) a verbal label on each scale point.
- D) verbal anchors at its end points.

104) A Likert scale with an odd number of scale points between "Strongly Agree" and "Strongly Disagree"

- A) cannot have equal scale distances.
- B) cannot have a neutral middle point.
- C) must have a verbal label on each scale point.
- D) is often used in marketing surveys.

105) A Likert scale with an even number of scale points between "Strongly Agree" and "Strongly Disagree"

- A) cannot have equal scale distances.
- B) is intended to prevent "neutral" choices.
- C) must have a verbal label on each scale point.
- D) is rarely used in surveys.

106) Which statement is correct?

- A) Analysts rarely consult business periodicals (e.g., *Bloomberg Businessweek*).
- B) Web searches (e.g., Google) often yield unverifiable data.
- C) Government data sources (e.g., www.bls.gov) are often costly.
- D) Private statistical databases (e.g., CRSP) are usually free.

107) Which statement is correct?

- A) Analysts avoid business periodicals (e.g., *Bloomberg Businessweek*).
- B) Web searches (e.g., Google) yield reliable and easily verified data.
- C) Government data sources (e.g., www.bls.gov) usually are free.
- D) Private statistical databases (e.g., CRSP) usually are free.

108) A *valid* survey is one that

- A) measures what the researcher wants to measure.
- B) has been approved by top management.
- C) is administered by a professional statistician.
- D) has a large number of questions.

109) A *reliable* survey is one that

- A) is administered by trusted employees.
- B) has been approved by quality engineers.
- C) gives consistent measurements.
- D) has many easy questions.