Complete Solutions Manual
to Accompany

Introduction to Statistics and Data Analysis

6th Edition

Roxy Peck
California Polytechnic State University, San Luis Obispo

Tom Short
West Chester University of Pennsylvania

Chris Olsen
Grinnell College

Prepared by

Stephen Miller
Winchester Thurston School, Pittsburgh, PA

CENGAGE Learning®

Australia • Brazil • Mexico • Singapore • United Kingdom • United States

ISBN-13: 978-1-337-79417-6
ISBN-10: 1-337-79417-1

Cengage Learning
20 Channel Center Street
Boston, MA 02210
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil, and Japan. Locate your local office at: **www.cengage.com/global**.

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

To learn more about Cengage Learning Solutions or to purchase any of our products at our preferred online store, visit **www.cengage.com**.

Contents

Chapter 1
The Role of Statistics and the Data Analysis Process

1.1 *Descriptive statistics* is the branch of statistics that involves the organization and summary of the values in a data set. *Inferential statistics* is the branch of statistics concerned with reaching conclusions about a population based on the information provided by a sample.

1.2 The *population* is the entire collection of individuals or objects about which information is required. A *sample* is a subset of the population selected for study in some prescribed manner.

1.3 The proportions are stated as population values (although they were very likely calculated from sample results).

1.4 The sample is the set of 2121 children used in the study. The population is the set of all children between the ages of one and four.

1.5 **a** The population of interest is the set of all 15,000 students at the university.

b The sample is the 200 students who are interviewed.

1.6 The estimates given were computed using data from a sample.

1.7 The population is the set of all 7000 property owners. The sample is the 500 owners included in the survey.

1.8 The population is the set of all 2019 Toyota Camrys. The sample is the set of six cars that are tested.

1.9 The population is the set of 5000 used bricks. The sample is the set of 100 bricks she checks.

1.10 **a** The researchers wanted to know whether the new surgical approach would improve memory functioning in Alzheimer's patients. They hoped that the negative effects of the disease could be reduced by toxins being drained from the fluid filled space that cushions the brain.

b First, it is not stated that the patients were randomly assigned to the treatments (new approach and standard care); this would be necessary in a well designed study. Second, it would help if the experiment could have been designed so that the patients did not know whether they were receiving the new approach or the standard care; otherwise, it is possible that the patients' knowledge that they were receiving a new treatment might in itself have brought about an improvement in memory. Third, as stated in the investigators' conclusion, it would have been useful if the experiment had been conducted on a sufficient number of patients so that any difference observed between the two treatments could not have been attributed to chance.

1.11 **a** The researchers wanted to find out whether taking a garlic supplement reduces the likelihood that you will get a cold. They wanted to know whether a significantly lower proportion of people who took a garlic supplement would get a cold than those who did not take a garlic supplement.

b It is necessary that the participants were *randomly* assigned to the treatment groups. If this was the case, it seems that the study was conducted in a reasonable way.

1.12 **a** Numerical (discrete)

b Categorical

c Numerical (continuous)

d Numerical (continuous)

e Categorical

1.13 **a** Categorical

b Categorical

c Numerical (discrete)

d Numerical (continuous)

e Categorical

f Numerical (continuous)

1.14 **a** Discrete

b Continuous

c Discrete

d Discrete

1.15 **a** Continuous

b Continuous

c Continuous

d Discrete

1.16 For example:

a Ford, Toyota, Ford, General Motors, Chevrolet, Chevrolet, Honda, BMW, Subaru, Nissan.

b 3.23, 2.92, 4.0, 2.8, 2.1, 3.88, 3.33, 3.9, 2.3, 3.56, 3.32, 2.4, 2.8, 3.9, 3.12.

c 4, 2, 0, 6, 3, 3, 2, 4, 5, 0, 8, 2, 5, 3, 4, 7, 3, 2, 0, 1

d 50.27, 50.67, 48.98, 50.58, 50.95, 50.95, 50.21, 49.70, 50.33, 49.14, 50.83, 49.89

e In minutes: 10, 10, 18, 0, 17, 17, 0, 17, 12, 19, 12, 13, 15, 15, 15

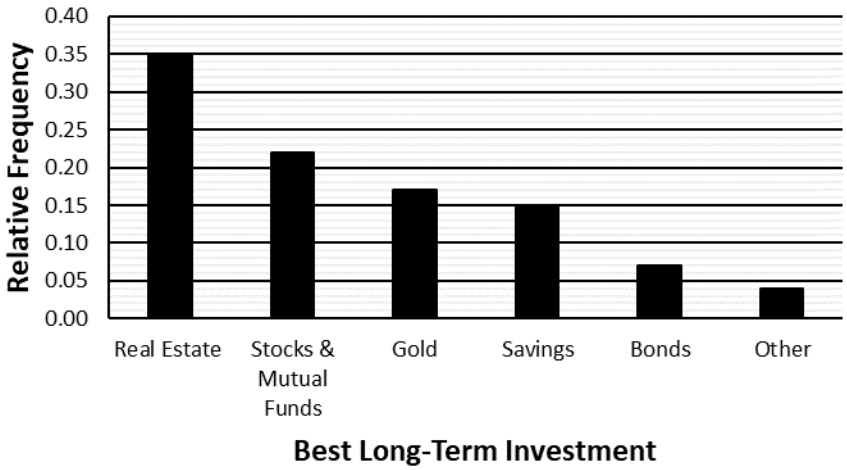
1.17 **a** Gender of purchaser, brand of motorcycle, telephone area code

b Number of previous motorcycles

c Bar chart

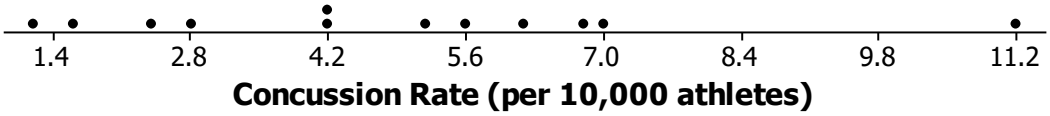
d Dotplot

1.18 **a**

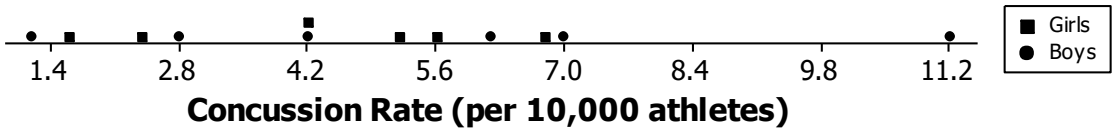


b Over half of the responses (57%) were from people who indicated that the best long-term investments were real estate (35%) and stocks & mutual funds (22%). The remaining 43% of respondents indicated that gold (17%), savings (15%), bonds (7%), and other (4%) were the best long-term investments.

1.19 **a** The dotplot below shows the concussion rate (concussions per 10,000 athletes).

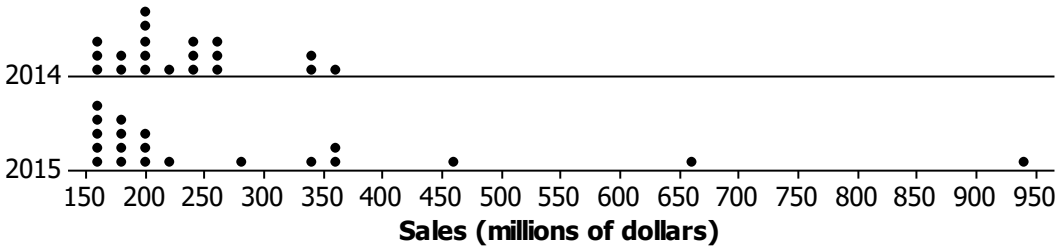


b The dotplot below shows the concussion rate (concussions per 10,000 athletes), with different symbols for boys and girls.



The sport with an unusually high (compared to all the other sports) concussion rate is football. Without considering football, the concussion rates for girls' sports is essentially the same as the concussion rate for boys' sports. However, if we consider football, the concussion rate for girls' sports tends to be lower than that for boys' sports.

1.20 Dotplots (for parts **a** and **b**) drawn on the same scale for the 2014 and 2015 sales data are shown below.

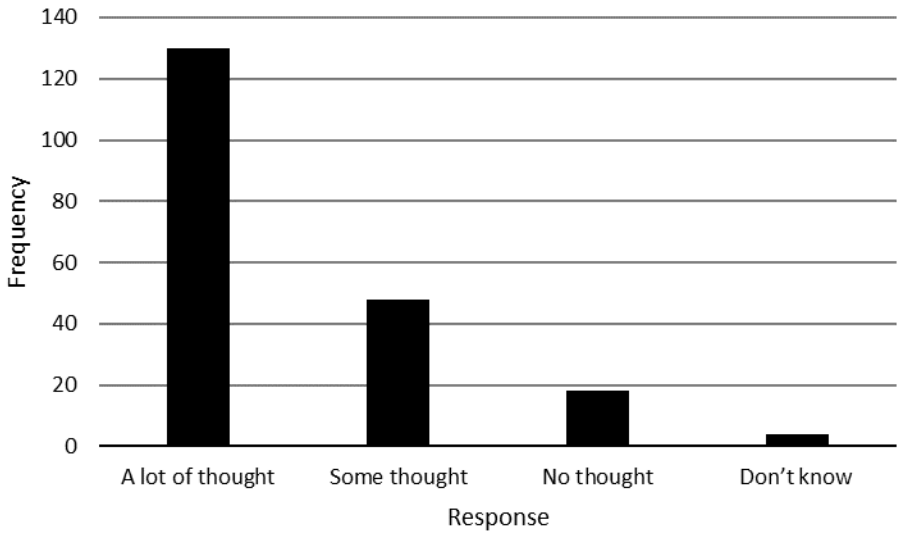


a A typical sales figure for 2014 was around 215 million dollars, with sales figures ranging from around 150 to around 350 million dollars. The greatest density of points was at the lower end of the distribution. There were no extreme observations in 2014.

b A typical sales figure for 2015 was around 200 million dollars, with sales figures ranging from around 155 to around 937 million dollars. The greatest density of points was at the lower end of the distribution. There are two extreme observations in 2015, namely, 936.7 million dollars and 652.3 million dollars.

c Sales figures were generally speaking higher in 2015 than in 2014. There were two extreme observations in 2015, and no extreme observations in 2014. If the extreme sales figures are taken into account, the variation in the sales figures (among the top 20 movies) was far greater in 2015 than in 2014. If the extreme sales figures are disregarded, the variation was still greater in 2015 than in 2014, but by not nearly as much. The distributions are similar in shape, with the greatest density of points being at the lower end of the distribution in both cases.

1.21 **a**



b The most common response was "A lot of thought", accounting for 130 (or 65%) of the students who started college but did not complete a degree. The next two most common responses were "Some thought" and "No thought", accounting for 48 (or 24%) and 18 (or 9%), respectively, of the students who started college but did not complete a degree. Finally,

4 of the 200 respondents (2%) indicated that don't know how much thought they have given to going back to school.

1.22 **a** Categorical

b Since the variable being graphed is categorical, a dotplot would not be suitable.

c If you add up the relative frequencies you get 107%. This total should be 100%, so a mistake has clearly been made.

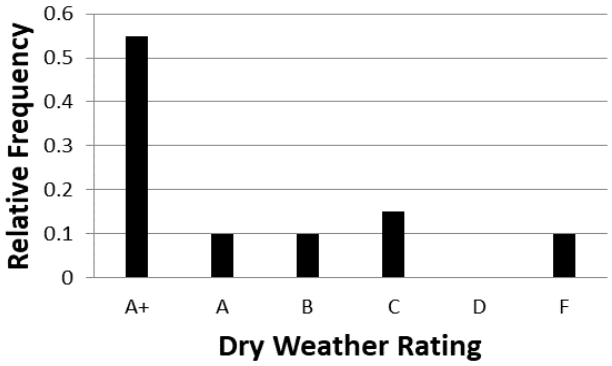
1.23 **a** The dotplot shows that there were three sites that received far greater numbers of visits than the remaining 6 sites. Also, it shows that the distribution of the number of visits has the greatest density of points for the smaller numbers of visits, with the density decreasing as the number of visits increases.

b It is clear from the dotplot that there were two sites that were used by far greater numbers of individuals (unique visitors) than the remaining 7 sites. However, these two sites are less far above the others in terms of the number of unique visitors than they are in terms of the total number of visits. As with the distribution of the total number of visits, the distribution of the number of unique visitors has the greatest density of points for the smaller numbers of visitors, with the density decreasing as the number of unique visitors increases. This is the case even when only the 7 less popular sites are considered.

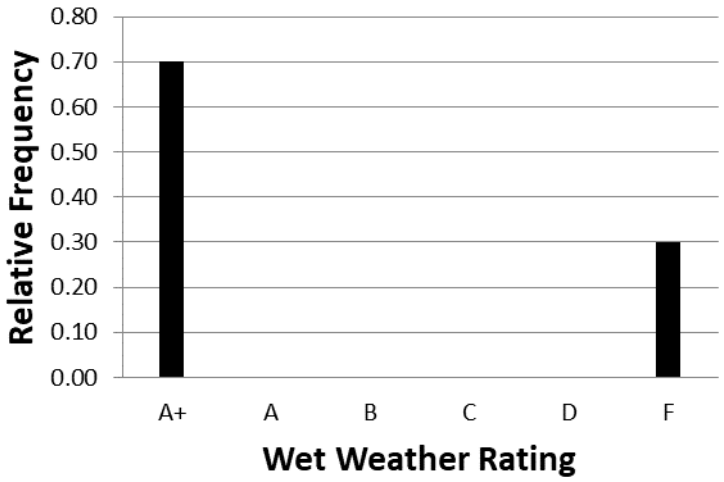
c The statistic "visits per unique visitor" tells us how heavily the individuals are using the sites. The table tells us that the most popular sites (Facebook and YouTube) in terms of total visits and unique visitors do not have the highest value of this statistic. The dotplot of visits per unique visitor shows that there are two individual sites are far ahead of the rest in this respect (Pinterest and Twitter).

1.24

Rating	Relative Frequency
A+	11/20 = 0.55
A	2/20 = 0.10
B	2/20 = 0.10
C	3/20 = 0.15
D	0/20 = 0.00
F	2/20 = 0.10

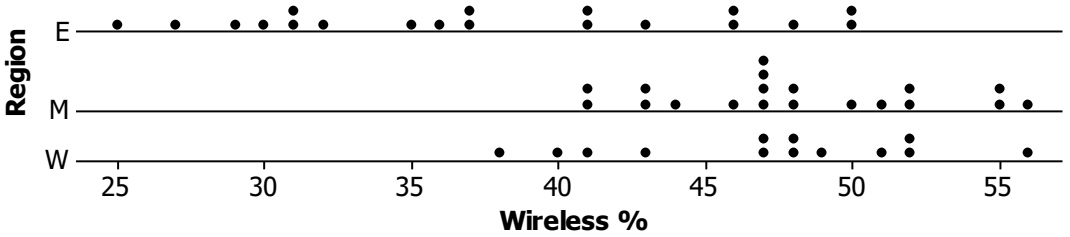


1.25 a



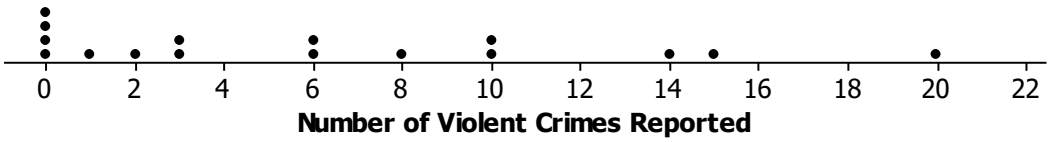
b Seventy-five percent (75%) of the dry weather ratings are B or higher, and 70% of wet weather ratings are B or higher, indicating that dry weather ratings are higher than wet weather ratings. Note that the wet weather ratings are only A+ or F, so the wet weather ratings are more extreme than dry weather ratings. If we only consider A+ ratings, then the wet weather ratings tend to be better than dry weather ratings because only 55% of dry weather ratings are A+, compared with 70% of wet weather ratings being A+.

1.26 a



b Looking at the dotplot we can see that Eastern states have, on average, lower wireless percents than states in the other two regions. The West and Middle states regions have, on average, roughly equal wireless percents.

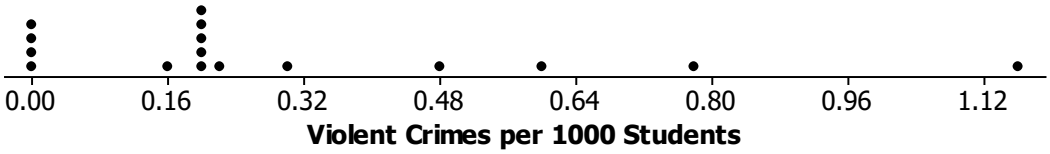
1.27 a



Three schools seem to stand out from the rest, these being, in increasing order of number of crimes, University of Central Florida (14 crimes reported), Florida International University (15 crimes reported), and Florida State University (20 crimes reported).

b

University/College	Violent Crime Rate Per 1000 Students
Florida A&M University	0.60435
Florida Atlantic University	0.19750
Florida Gulf Coast University	0.20225
Florida International University	0.30131
Florida South Western State College	0.00000
Florida State University, Tallahassee	0.48984
New College of Florida	1.16144
Pensacola State College	0.00000
Santa Fe College	0.00000
Tallahassee Community College	0.16071
University of Central Florida	0.22239
University of Florida	0.19745
University of North Florida	0.19139
University of South Florida, St. Petersburg	0.00000
University of South Florida, Tampa	0.19017
University of West Florida	0.78351



The colleges that stand out in violent crimes per 1000 students are, in increasing order of crime rate, Florida State University, Florida A&M University, University of West Florida, and New College of Florida. Only Florida State University stands out in both dotplots.

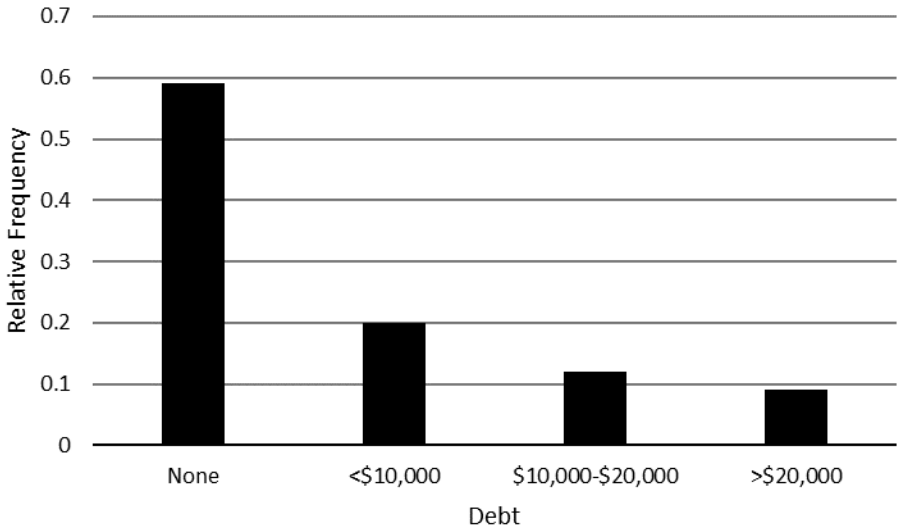
c For the number of violent crimes, there are three schools that stand out by having high numbers of crimes, with the majority of the schools having similar, and low (10 or fewer), numbers of crimes. There seems to be greater consistency for crime rate (per 1000 students) among the 16 schools than there is for number of crimes, with four schools standing out as having high crime rates, and four schools with crime rates that stand out as being noticeably low.

1.28 **a** When ranking the airlines according to delayed flights, one airline would be ranked above another if the probability of a randomly chosen flight being delayed is smaller for the first airline than it is for the second airline. These probabilities are estimated using the *rate per 10,000 flights* values, and so these are the data that should be used for this ranking. (Note that the *total number of flights* values are not suitable for this ranking. Suppose that one airline had a larger number of delayed flights than another airline. It is possible that this could be accounted for merely through the first airline having more flights than the second.)

b There are two airlines, ExpressJet and Continental, which, with 4.9 and 4.1 of every 10,000 flights delayed, stand out as the worst airlines in this regard. There are two further airlines that stand out above the rest: Delta and Comair, with rates of 2.8 and 2.7 delayed flights per

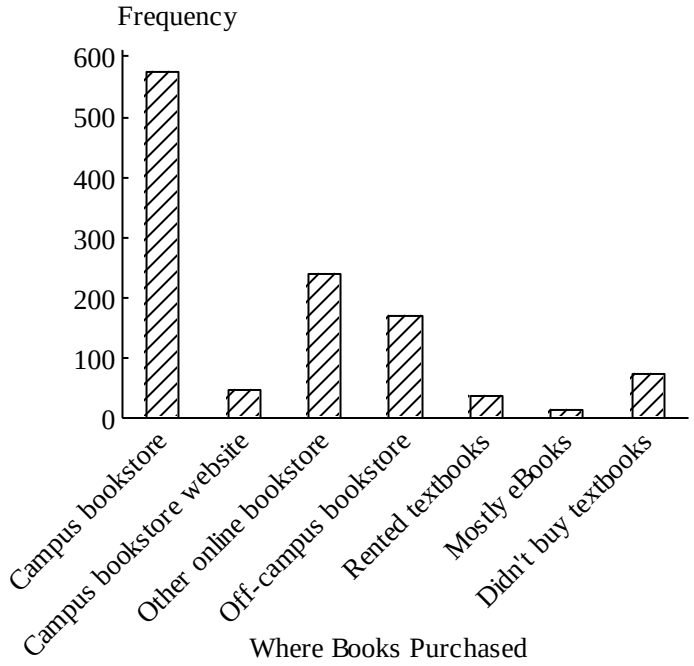
10,000 flights. All the other airlines have rates below 1.6, with the best rating being for Southwest, with a rate of only 0.1 delayed flights per 10,000.

1.29 a



b Most public community college graduates have no debt at all, and a debt of $10,000 or less accounts for 79% of the graduates. Among the 21% of the graduates who have a debt of more than $10,000, nearly 43% (9% of all graduates) have a debt of more than $20,000.

1.30 a



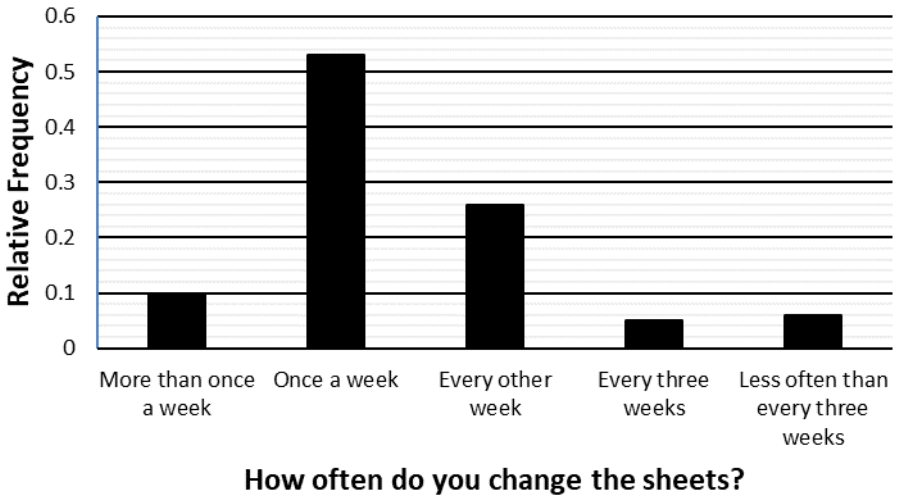
b By far the most popular place to buy books is the campus bookstore, with half of the students in the sample buying their books from that source. The next most popular sources are online bookstores other than the online version of the campus bookstore and off-campus bookstores,

with these two sources accounting for around 35% of students. Purchasing mostly eBooks was the least common response.

1.31 **a**

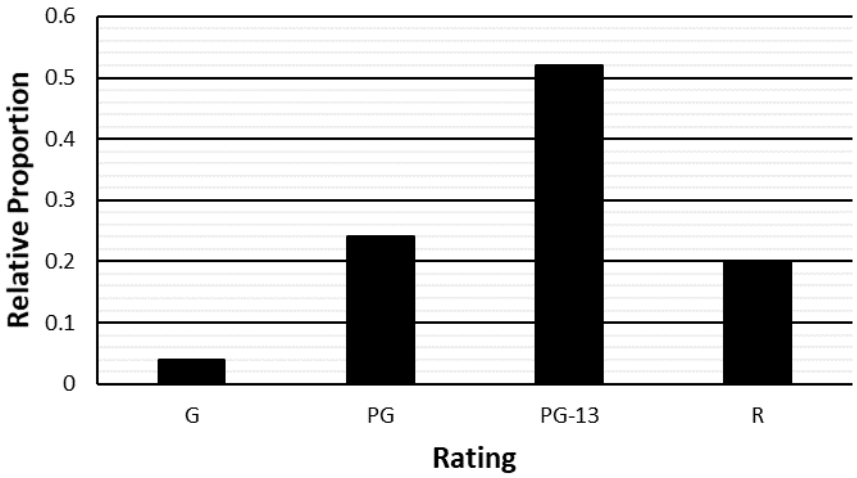
How Often	Relative Frequency
More than once a week	0.10
Once a week	0.53
Every other week	0.26
Every three weeks	0.05
Less often than every three weeks	0.06

b



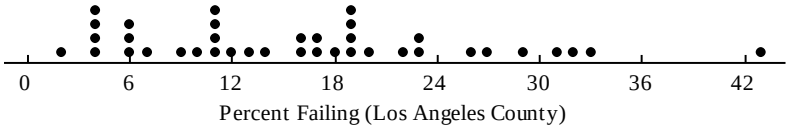
1.32 The relative frequency distribution is:

Rating	Relative Frequency
G	1/25 = 0.04
PG	6/25 = 0.24
PG-13	13/25 = 0.52
R	5/25 = 0.20



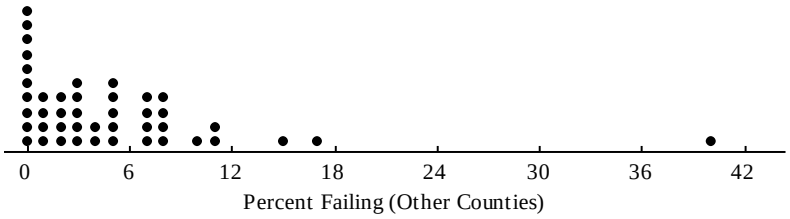
PG-13 is the rating with the highest relative proportion (0.52), followed by PG (0.24), R (0.20), and G (0.04). Seventy-two percent (72%) of the top 25 movies of 2015 are PG-13 or R, and the remaining 28% are rated G or PG.

1.33 **a** The dotplot for Los Angeles County is shown below.



A typical percent of tests failing for Los Angeles County is around 16. There is one value that is unusually high (43), with the other values ranging from 2 to 33. There is a greater density of points toward the lower end of the distribution than toward the upper end.

b The dotplot for the other counties is shown below.

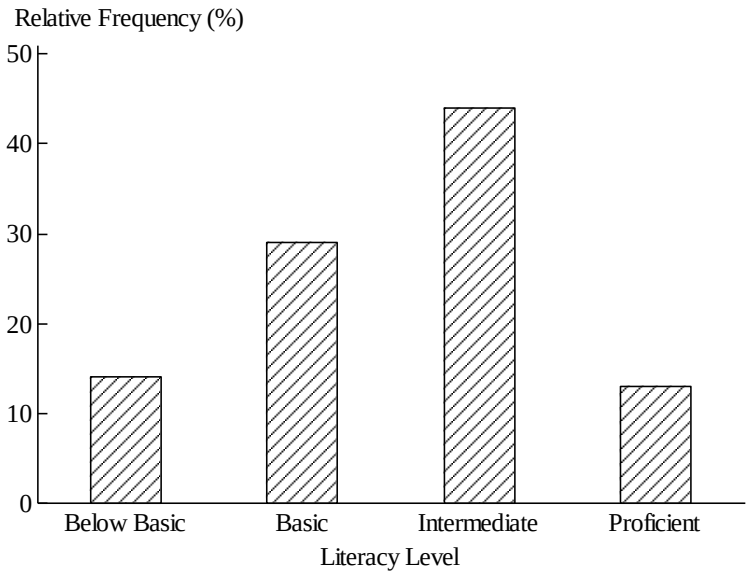


A typical percent of tests failing for the other counties is around 3. There is one extreme result at the upper end of the distribution (40); the other values range from 0 to 17. The density of points is highest at the left hand end of the distribution and decreases as the percent failing values increase.

c The typical value for Los Angeles County (around 16) is greater than for the other counties (around 3) and, disregarding the one extreme value in each case, there is a greater variability in the values for Los Angeles County than for the other counties. In the distribution for Los Angeles County the points are closer to being uniformly distributed than in the distribution for the other counties, where there is a clear tail-off of density of points as you move to the right of the distribution.

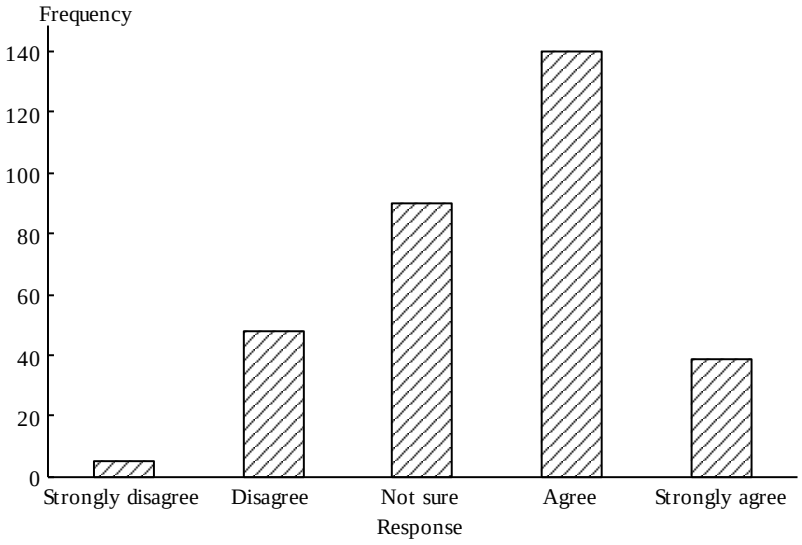
1.34 **a** Categorical

b

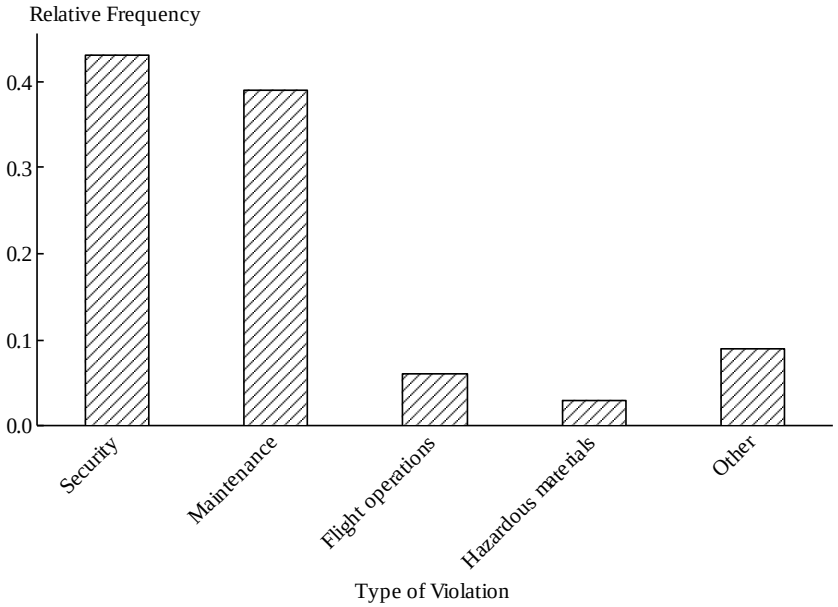


c No, since dotplots are used for numerical data.

1.35

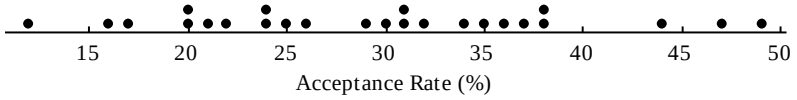


1.36 a



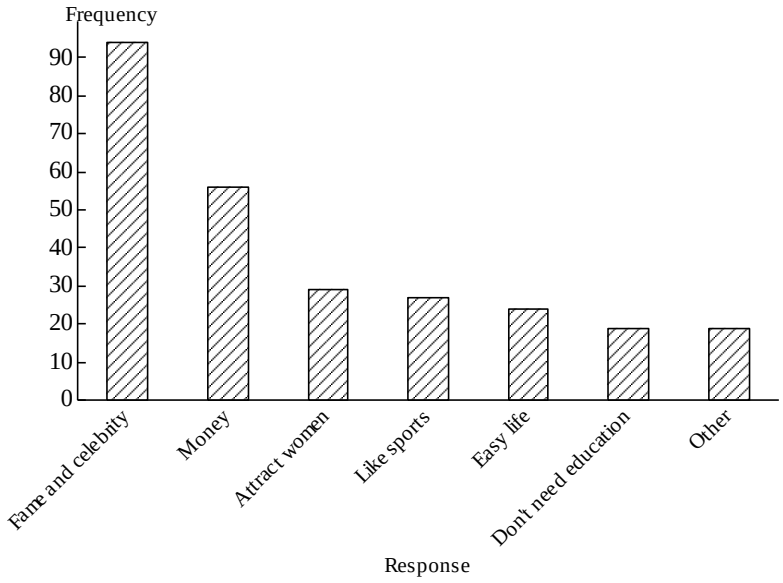
b By far the most frequently occurring violation categories were security (43%) and maintenance (39%). The least frequently occurring violation categories were flight operations (6%) and hazardous materials (3%).

1.37 a



b A typical acceptance rate for these top 25 schools is around 30, with the great majority of acceptance rates being between 19 and 39. There are no particularly extreme values. The pattern of the points is roughly symmetrical.

1.38


Frequency
90
80
70
60
50
40
30
20
10
0
Fame and celebrity
Money
Attract women
Like sports
Easy life
Don't need education
Other
Response

Chapter 2
Collecting Data Sensibly

2.1 This is an observational study. The treatments (length of stay) were determined by the condition of the patients. (In an experiment the patients would be assigned to the various lengths of stay by the investigators, usually using random assignment.)

2.2 This was an experiment, since the investigators (not the students) determined which discussion sections received the chocolate and which did not.

2.3 This was an experiment, since the professor (not the students) determined who was supplying a buying price and who was supplying a selling price.

2.4 **a** This is an observational study.

b No, it is not reasonable to conclude that getting less than 8 hours of sleep on school nights causes teenagers to fall asleep during school and to consume more caffeine. This is an observations study, so cause-and-effect conclusions cannot be drawn.

2.5 **a** This is an experiment since it was decided by the researchers (in this case by random assignment) which participants would receive which treatments.

b Yes. Since the participants were randomly assigned to the treatments it is reasonable to conclude that receiving either real or fake acupuncture was the cause of the observed reductions in pain.

2.6 **a** This is an observational study.

b Yes. Since the researchers looked at a random sample of publically accessible MySpace web profiles posted by 18-year-olds, it is reasonable to generalize the stated conclusion to all 18-year-olds with publically accessible MySpace profiles.

c No, it is not reasonable to generalize the stated conclusion to all 18-year-old MySpace users since no users without publically accessible profiles were included in the study.

d No, it is not reasonable to generalize the stated conclusion to all MySpace users with publically accessible profiles since only 18-year-olds were included in the study.

2.7 **a** This is an experiment.

b Yes. Since the participants were randomly assigned to the treatments the researcher can reasonably claim that hearing the cork pop was the cause of the higher rating.

2.8 It is quite possible, for example, that 3- and 4-year olds who drink something sweet once or twice a day generally speaking consume larger amounts of fat than those who do not, and that it is the fat that is causing the weight problems a year later, rather than the consumption of the sweet drinks.

2.9 We are told that moderate drinkers, as a group, tended to be better educated, wealthier, and more active than nondrinkers. It is therefore quite possible that the observed reduction in the risk of heart disease amongst moderate drinkers is caused by one of these attributes and not by the moderate drinking.

2.10 **a** No. It is quite possible, for example, that women who choose to watch *Oprah* generally speaking have a more health oriented outlook than those who watch other daytime talk shows, and it is this health oriented outlook that causes the decrease in craving for fattening foods, not the watching of *Oprah*.

b Neither generalization would be reasonable since the survey was conducted on the DietSmart.com website. It is unlikely that users of this website would be representative of the population of women in the United States or of the population of women who watch daytime talk shows.

2.11 **a** The data would need to be collected from a simple random sample of all adult American Internet users.

b No. Since the survey included only adult Americans who are internet users, the result cannot be generalized to all adult Americans.

2.12 I do agree with the statement. Observational studies cannot be used to draw cause-and-effect conclusions.

2.13 The following is one possible method. Use a computer list of all the students at the college. Assign to each student a 10-digit decimal number. Sort the students according to the numbers assigned, smallest to largest. The first 100 students on the sorted list will form the sample.

2.14 Method 1: Using a computer list of the graduates, number the graduates 1-140. Use a random number generator on a calculator or computer to randomly select a whole number between 1 and 140. The number selected represents the first graduate to be included in the sample. Repeat the number selection, ignoring repeated numbers, until 20 graduates have been selected.
Method 2: Using a computer list of the graduates, number the graduates 001-140. Take the first three digits from the left hand end of a row from a table of random digits. If the three-digit number formed is between 001 and 140 inclusive, the graduate with that number should be the first graduate in the sample. If the number formed is not between 001 and 140 inclusive, the number should be ignored. Repeat the process described for the next three digits in the random number table, and continue in the same way until 20 graduates have been selected. (Three-digit numbers that are repeats of numbers previously selected should be ignored.)

2.15 The following is one possible method. Number the signatures 1-500. Use a random number generator on a calculator or computer to randomly select a whole number between 1 and 500. The number selected represents the first signature to be included in the sample. Repeat the number selection, ignoring repeated numbers, until 30 signatures have been selected.

2.16 **a** The 716 bicycle fatalities constitute *all* bicycle fatalities in 2008, and so the group represents a census.

b The average age of 41 years is a population characteristic, since it is the average for *all* bicycle fatalities in 2008.

2.17 **a** The population is all American women.

b No. The sample included women only from Maryland, Minnesota, Oregon, and Pennsylvania. It is quite possible that the relationship between exercise and cognitive impairment is different for women in other states.

c No. As mentioned in Part (b), it is quite possible that the relationship between exercise and cognitive impairment is different for women in other states. For example, other states might contain a different racial mix from the four included in the study, and it is possible that women of different racial origins respond differently to exercise in terms of cognitive impairment.

d The inadequate sampling method used in this study is an example of selection bias. Women in the other 46 states were excluded from the study.

2.18 **a** Cluster sampling

b Stratified random sampling

c Convenience sampling

d Simple random sampling

e Systematic sampling

2.19 **a** Using the list, first number the part time students 1-3000. Use a random number generator on a calculator or computer to randomly select a whole number between 1 and 3000. The number selected represents the first part time student to be included in the sample. Repeat the number selection, ignoring repeated numbers, until 10 part time students have been selected. Then number the full time students 1-3500 and select 10 full time students using the same procedure.

b No. With 10 part time students being selected out of a total of 3000 part time students, the probability of any particular part time student being selected is $10/3000 = 1/300$. Applying a similar argument to the full time students, the probability of any particular full time student being selected is $10/3500 = 1/350$. Since these probabilities are different, it is not the case that every student has the same chance of being included in the sample.

2.20 Convenience samples are, by nature, very unlikely to be representative of the population.

2.21 **a** The pages of the book have already been numbered between 1 and the highest page number in the book. Use a random number generator on a calculator or computer to randomly select a whole number between 1 and the highest page number in the book. The number selected will be the first page to be included in the sample. Repeat the number selection, ignoring repeated numbers, until the required number of pages has been selected.

b Pages that include exercises tend to contain more words than pages that do not include exercises. Therefore, it would be sensible to stratify according to this criterion. Assuming that 20 non-exercise pages and 20 exercise pages will be included in the sample, the sample should be selected as follows. Use a random number generator to randomly select a whole number between 1 and the highest page number in the book. The number selected will be the

first page to be included in the sample. Repeat the number selection, ignoring repeated numbers and keeping track of the number of pages of each type selected, until 20 pages of one type have been selected. Then continue in the same way, but ignore numbers corresponding to pages of that type. When 20 pages of the other type have been selected, stop the process.

c Randomly select one page from the first 20 pages in the book. Include in your sample that page and every 20th page from that page onwards.

d Roughly speaking, in terms of the numbers of words per page, each chapter is representative of the book as a whole. It is therefore sensible for the chapters to be used as clusters. Using a random number generator randomly choose three chapters. Then count the number of words on each page in those three chapters.

2.22 Answers will vary.

2.23 Answers will vary.

2.24 It is not the *proportion* of voters that is important, but the *number* of voters in the sample – and 1000 is an adequate number.

2.25 The researchers should be concerned about nonresponse bias. Only a small proportion (20.7%) of the selected households completed the interview, and it is quite possible that those households who did complete the interview are different in some relevant way concerning Internet use from those who did not.

2.26 **a** This was a cluster sample, with each team being a cluster. Note that all student-athletes on a particular team are surveyed.

b The student-athletes on the randomly selected teams are unlikely to be representative of the set of all college students in the U.S. Also, since the survey was about illegal drug use (a sensitive subject), we cannot be sure that the students were giving truthful answers to the questions.

2.27 First, the participants in the study were all students in an upper-division communications course at one particular university. It is not reasonable to consider these students to be representative of all students with regard to their truthfulness in the various forms of communication. Second, the students knew during the week's activity that they were surveying themselves as to the truthfulness of their interactions. This could easily have changed their behavior in particular social contexts and therefore could have distorted the results of the study.

2.28 **a** No. The parents in the study were not randomly selected and therefore are not likely to be representative of the population of all parents.

b No. The sample is not likely to have been representative of the population of all parents and therefore it would not be reasonable to generalize the researchers' conclusion to all parents.

2.29 First, the people who responded to the print and online advertisements might be different in some way relevant to the study from the population of people who have online dating profiles. Second, only the *Village Voice* and Craigslist New York City were used for the recruitment. It is quite possible that people who read that newspaper or access those websites differ from the population

in some relevant way, particularly considering that they are both New York City based publications.

2.30 It is important to know if the sample of 800 college students was randomly selected, and from what population of college students was the sample selected. For example, were the students randomly selected from colleges in one particular state or region, or were all colleges considered in the sampling scheme.

2.31 **a** Yes. It is possible that students of a given class standing tend to be similar in the amounts of money they spend on textbooks.

b Yes. It is possible that students who pursue a certain field of study tend to be similar in the amounts of money they spend on textbooks.

c No. It is unlikely that stratifying in this way will produce groups that are homogeneous in terms of the students' spending on textbooks.

2.32 The individuals within each stratum should on the whole be similar in terms of the topic of the study. This is true of the proposed strata in Scheme 2, since it is likely that college students will on the whole be similar in their opinions regarding the possible tax increase; likewise nonstudents who work full time will on the whole be similar in their opinions regarding the possible tax increase, and nonstudents who do not work full time will on the whole be similar in their opinions regarding the possible tax increase. Scheme 1, however, is not suitable since we have no reason to believe that people within the proposed first-letter-of-last-name strata will be similar in terms of their attitudes to the possible tax increase. Similarly the suggested stratification in Scheme 3 is very unlikely to produce homogeneous groups.

2.33 It is not reasonable to generalize these results to the population of U.S. adults since the people who sent their hair for testing did so voluntarily. It is quite possible that people who would choose to participate in a study of this sort differ in their mercury levels from the population as a whole.

2.34 Different subsets of the population might have responded by different methods. For example, it is quite possible that younger people (who might generally be in favor of continuing the parade) chose to respond via the Internet while older people (who might on the whole be against the parade) chose to use the telephone to make their responses.

2.35 **a** Binding strength

b Type of glue

c The extraneous variables mentioned are the number of pages in the book and whether the book is bound as a hardback or a paperback. Further extraneous variables that might be considered include the weight of the material used for the cover and the type of paper used.

2.36 **a** Use two IQ tests (Test 1 and Test 2) of equal levels of difficulty. Randomly select 50 students from the school. Randomly assign the 50 students to two groups, Group A and Group B. The experiment will be conducted over two days. On the first day, students in Group A will do an IQ test without listening to any music and students in Group B will do an IQ test after listening to a Mozart piano sonata. On the second day the activities of the two groups will be switched. For each student decide randomly whether he/she will take Test 1 on the first day

and Test 2 on the second day, or vice versa. All conditions (temperature, time of day, amount of light, etc. – everything apart from the presence of the music or not) should be kept equal between the two days and between the two groups. In particular, the students taking the test without listening to the music should nonetheless sit quietly in a room for the length of time of the piano sonata. At the end of the experiment the after-music IQ scores should be compared to the no-music IQ scores.

b Yes, the fact that all conditions are kept the same is direct control.

c By having each student take IQ tests under *both* experimental conditions we are using a matched pairs design, and matched pairs is a form of blocking.

d The students were randomly assigned to the two orders of the treatments (no music then music, or music then no music). Also it was decided randomly for each student whether Test 1 or Test 2 would be taken first.

2.37 The following is one possible method. Write the names of the subjects on identical slips of paper, and place the slips in a hat. Mix the slips, and randomly select ten slips from the hat. The names on those ten slips are the people who will dry their hands by the first method. Randomly select a further ten slips from the hat. The names on those ten slips are the people who will dry their hands by the second method. The remaining ten people will dry their hands by the third method.

2.38 Random assignment should have been used to determine, for each cyclist, which drink would be consumed during which break.

2.39 **a** Blocking

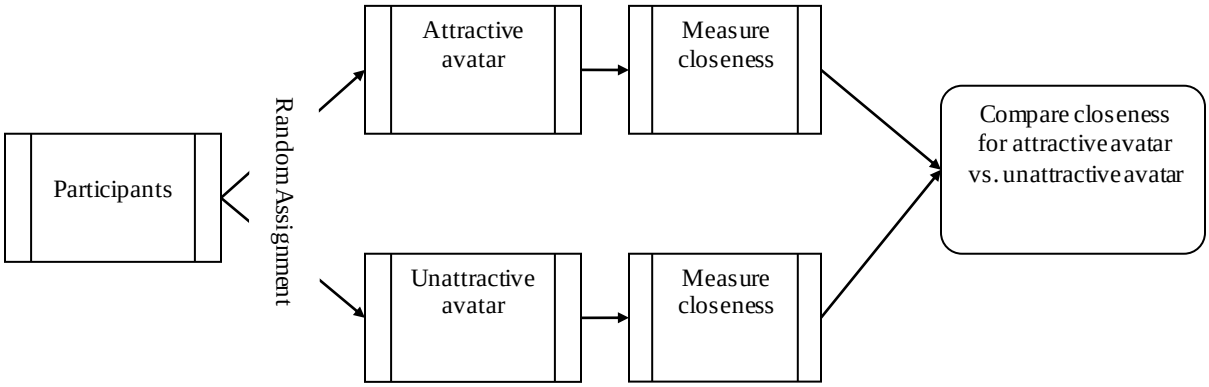
b Direct control

2.40 We rely on random assignment to produce comparable experimental groups. If the researchers had hand-picked the treatment groups, they might unconsciously have favored one group over the other in terms of some variable that affects the subjects' ability to deal with multiple inputs.

2.41 The figure shows that comparable groups in terms of age have been formed. (Differences between the age distributions in the groups can be seen: there are one or two children in the LR group who are younger than all those in the OR group, and also there is a greater number of children over 14 years old in the LR group. It is inevitable that there will be *some* differences between the groups under random assignment.)

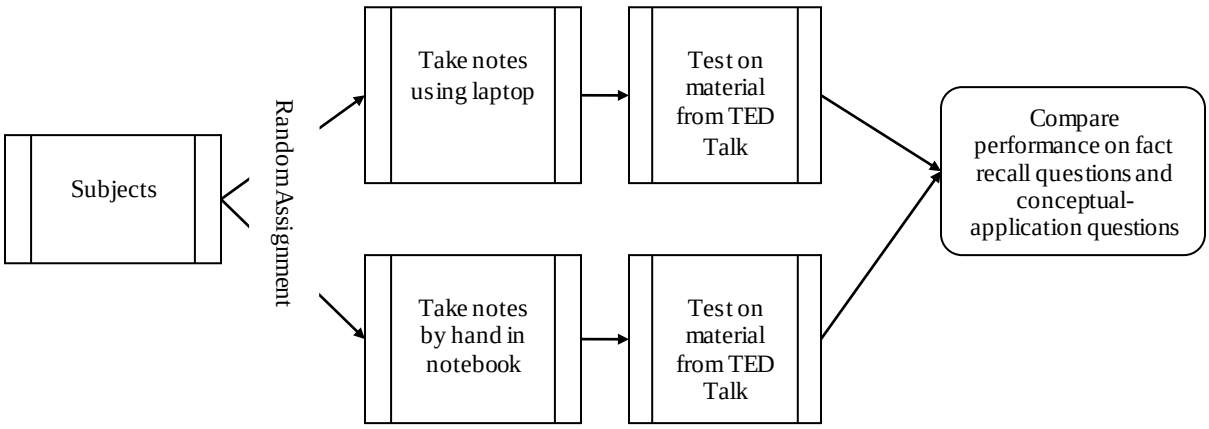
2.42 **a** If the participants had been able to choose their own avatars, then it is quite possible, for example, that people with a lot of self confidence would tend to choose the attractive avatar while those with less self confidence would tend to choose the unattractive avatar. Then, if the same result was obtained as the one described in the paper, it would be impossible to tell whether the greater closeness achieved by those with the attractive avatar came about as a result of the avatar or as a result of those people's greater self confidence.

b

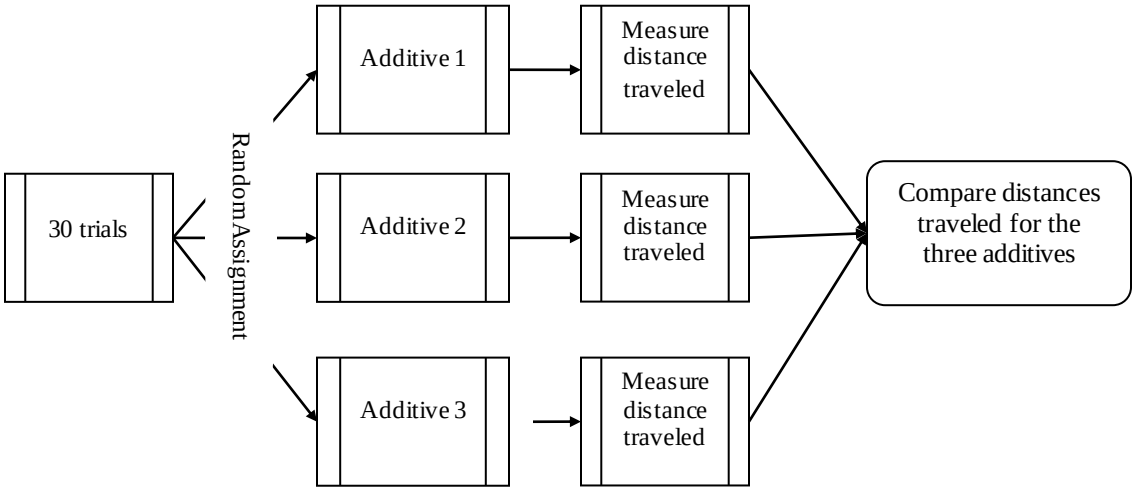


2.43 We rely on random assignment to produce comparable experimental groups. If the researchers had hand-picked the treatment groups, they might unconsciously have favored one group over the other in terms of some variable that affects the subjects' ability to learn through video gaming activity.

2.44



2.45



2.46 How many bottles of water were used in the experiment? Were the bottles identical? Was there a control group in which the bottles had identical labels to those used in the treatment group but without the positive words? Were the bottles randomly assigned to the groups? Was the state of the water measured both before and after the experiment? Did the people who measured the water's structure know which bottles were in which group?

2.47 **a** This is an experiment. The researchers assigned the students to the treatments.

b No. The subjects were randomly assigned to the treatments, but there was no random selection from a population of interest.

c Yes.

2.48 **a** If the performance was *significantly* better for the group that read the material in the landscape orientation then, yes, for the set of subjects used in the experiment, reasoning was improved by turning the screen to landscape orientation. The subjects were randomly assigned to the treatment groups, and a significant difference in the results of the two groups means that the observed difference is unlikely to have occurred by chance.

b No. The students used in the study were all undergraduates taking psychology at a large university, and therefore cannot be considered to be representative of any larger population.

2.49 **a** The treatments are the names – Ann Clark and Andrew Clark – given to the participants.

b The response variables are the participants' answers to the questions given.

2.50 Selecting a random sample of 1161 voters and giving them the female name, and then selecting a second random sample 1139 voters and giving them the male name, is exactly equivalent to selecting a random sample of 2300 voters, and then randomly assigning 1161 of them to the female name and the remainder to the male name. (It is assumed here that in the study given in the question it was ensured that there was no overlap between the two samples.)

2.51 **a** Red wine, yellow onions, black tea

b Absorption of flavonols

c Alcohol tolerance, amount of flavonols in the diet (other than from red wine, yellow onions, and black tea), gender

2.52 Suppose that an experiment is conducted in which people are given either a drug or a placebo, and that those who are given the drug do significantly better, on average, than those who are given the placebo. Since both groups experience the placebo effect (the psychological effect of taking a tablet) we are able to attribute the greater improvement of those who took the drug to the chemicals in the drug. However we are unable to tell just how much of a placebo effect is being experienced by all the subjects. By adding a control group (a group that is given nothing) and comparing the results for this group with the results for the placebo group we can measure the extent of the placebo effect.

2.53 "Blinding" is ensuring that the experimental subjects do not know which treatment they were given and/or ensuring that the people who measure the response variable do not know who was given which treatment. When this is possible to implement, it is useful that the subjects do not know which treatments they were given since, if a person knows what treatment he/she was given, this knowledge could influence the person's perception of the response variable, or even, through psychological processes, have a direct effect on the response variable. If the response variable is to be measured by a person other than the experimental subjects it is useful if this person doesn't know who received which treatment since, if this person *does* know who received which treatment, then this could influence the person's perception of the response variable.

2.54 Answers will vary.

2.55 **a** Allowing study participants to choose which group they want to be in could introduce systematic differences between the two experimental conditions (knee replacement surgery with exercise and exercise therapy alone), resulting in potential confounding. Those who chose knee replacement surgery plus exercise might, in some way, be different from those who chose exercise therapy alone. We would not know if differences in pain relief between the two groups were due to the knee replacement surgery with exercise, or due to some inherent differences in the subjects who chose their experimental groups.

b The researchers likely did not include a control group because the study participants needed some relief from their pain. Because the purpose of this experiment is to determine whether knee replacement surgery with exercise provided more pain relief than exercise therapy alone, a control group would not allow the study participants to have the opportunity to experience pain relief.

2.56 **a** It is very possible that the nurses might have preconceptions about the two forms of surgery in terms of the amount of pain and nausea caused. Therefore, if the nurses know which children have been given which form of surgery, this might affect the amounts of medication they give. By making sure that the nurses do not have this knowledge, this possible effect is avoided.

b Since the incisions made under the two procedures are different the patients and/or their parents would know which method had been used.

2.57 We will assume that only four colors will be compared, and that only headache sufferers will be included in the study.

Prepare a supply of "Regular Strength" Tylenol in four different colors: white (the current color of the medication, and therefore the "control"), red, green, and blue. Recruit 20 volunteers who suffer from headaches. Instruct each volunteer not to take any pain relief medication for a week. After that week is over, issue each volunteer a supply of all four colors. Give each volunteer an order in which to use the colors (this order would be determined randomly for each volunteer). Instruct the volunteers to use one fixed dose of the medication for each headache over a period of four weeks, and to note on a form the color used and the pain relief achieved (on a scale of 0-10, where 0 is no pain relief and 10 is complete pain relief). At the end of the four weeks gather the results and compare the pain relief achieved by the four colors.

2.58 **a** Randomly assigning 852 children to the book group and the rest to the control group consists of randomly selecting 852 to be in the book group and putting the remaining children in the control group.

b If no control group had been included in the study, then the only results available to the researchers would be the reading scores of the children who had been given the reading books. There would be no way of telling whether these scores were any better than the scores for children who were not given reading books.

2.59 Of the girls, randomly assign 350 to the book group and 350 to the no-book group. (You could do this by starting with a computer list of the names of the 700 girls. Assign to each name a random 10-digit number. Sort the names according to the numbers, from smallest to largest. The first 350 names on the sorted list are assigned to the book group, and the remainder to the no-book group.) Using a similar method, randomly assign 315 of the boys to the book group and 315 to the no-book group.

2.60 Suppose that the dog handlers and/or the experimental observers had known which patients did and did not have cancer. It would then be possible for some sort of (conscious or unconscious) communication to take place between these people and the dogs so that the dogs would pick up the conditions of the patients from these people rather than through their perception of the patients' breath. By making sure that the dog handlers and the experimental observers do not know who has the disease and who does not it is ensured that the dogs are getting the information from the patients.

2.61 **a** If the judges had known which chowder came from which restaurant then it is unlikely that Denny's chowder would have won the contest, since the judges would probably be conditioned by this knowledge to choose chowders from more expensive restaurants.

b In experiments, if the people measuring the response are not blinded they will often be conditioned to see different responses to some treatments over other treatments, in the same way as the judges would have been conditioned to favor the expensive restaurant chowders. It is therefore necessary that the people measuring the response should not know which subject received which treatment, so that the treatments can be compared on their own merits.

2.62 This describes the placebo effect. The implication is that the experiments have included patients who have been given a placebo in place of the antidepressants.

2.63 **a** A placebo group would be necessary if the mere thought of having amalgam fillings could produce kidney disorders. However, since the experimental subjects were sheep the researchers do not need to be concerned that this would happen.

b A resin filling treatment group would be necessary in order to provide evidence that it is the material in the amalgam fillings, rather than the process of filling the teeth, or just the presence of foreign bodies in the teeth, that is the cause of the kidney disorders. If the amalgam filling group developed the kidney disorders and the resin filling group did not, then this would provide evidence that it is some ingredient in the amalgam fillings that is causing the kidney problems.

c Since there is concern about the effect of amalgam fillings it would be considered unethical to use humans in the experiment.

2.64 Answers will vary.

2.65 **a** This is an observational study.

b In order to evaluate the study, we need to know whether the sample was a random sample.

2.66 **a** No. Since the sample used in the Healthy Steps study was known to be nationally representative, and since the paper states that, compared with the HS trial, parents in the study sample were disproportionately older, white, more educated, and married, it is clear that it is not reasonable to regard the sample as representative of parents of all children at age 5.5 years.

b The potential confounding variable mentioned is what the children watched.

c The quotation from Kamila Mistry makes a statement about cause and effect and therefore is inconsistent with the statement that the study can't show that TV was the cause of later problems.

2.67 Answers will vary.

2.68 Answers will vary.

2.69 **a** There are many possible designs. We will describe here a design that blocks for the day of the week and the section of the newspaper in which the advertisement appears. For the sake of argument we will assume that the mortgage lender is interested in advertising on only two days of the week (Monday and Tuesday) and that there are three sections in the newspaper (A, B, and C). We will refer to the three types of advertisement as Ad 1, Ad 1, and Ad 3.

The experimental units are 18 issues of the newspaper (that is, 18 dates) consisting of Mondays and Tuesdays over 9 weeks. Use a random process to decide which three Mondays will receive advertisements in Section A, which three Mondays will receive advertisements in Section B, and which three Mondays will receive advertisements in Section C. Do the same for the nine Tuesdays. We have now effectively split the 18 issues into the six blocks shown below. (There are 3 issues in each block.)

Mon, Sect A	Mon, Sect B	Mon, Sect C
Tue, Sect A	Tue, Sect B	Tue, Sect C

Now randomly assign the three issues in each block to the three advertisements. (Ad 1 is then appearing on three Mondays, once in each section, and on three Tuesdays, once in each section. The same applies to Ad 2 and Ad 3.) The response levels for the three advertisements can now be compared (as can the three different sections and the two different days).

b Within each block, the three issues of the newspaper were randomly assigned to the three advertisements.

2.70 Study 1

1. Observational study
2. No
3. No
4. No. The fact that calcium takers were more common among the heart attack patients implies mathematically that calcium takers were more likely to be heart attack patients than non-calcium takers. However, it is quite possible that people who take a calcium supplement very often also take another supplement, and it is this other supplement that is causing heart attacks, not the calcium.
5. No. The hospital at which this study was conducted cannot be considered to be representative of any larger population.

Study 2

1. Observational study
2. Yes
3. No
4. No. It is quite possible that people who take a calcium supplement very often also take another supplement, and it is this other supplement that is causing heart attacks, not the calcium.
5. The conclusions can only be generalized to the population of people living in Minneapolis who receive Social Security.

Study 3

1. Experiment
2. Yes
3. No
4. No
5. No reasonable conclusion can be drawn from the study.

Study 4

1. Experiment
2. No
3. Yes
4. Yes
5. No. The participants were volunteers.

2.71 By randomly selecting the phone numbers, calling back those where there are no answers, and asking for the adult in the household with the most recent birthday the researchers are avoiding selection bias. However, selection bias could come about as a result of the fact that not all

Californians have phones. Also, by selecting the adult with the most recent birthday certain birth dates are being favored, and it could be suggested that different attributes of people born at different times of the year could introduce further selection bias. Further to that, there is always the concern that people might not answer truthfully. This is response bias.

2.72 Answers will vary.

2.73 We rely on random assignment to produce comparable experimental groups. If the researchers had hand-picked the treatment groups, they might unconsciously have favored one group over the other in terms of some variable that affects the ability of the people at the centers to respond to the materials provided.

2.74 **a** Observational study

b There are many reasons the results of the study may not generalize to the population of all U.S. men. Here are two possible reasons. First, the study participants might not have been randomly selected. Second, if they were randomly chosen, we don't know from what population they were selected.

2.75 This is an observational study based on results of a survey (no consumers were assigned to different experimental conditions).

2.76 Nonresponse bias: it is possible that those who responded differed in some important way from those who did not. Also there is a possibility of response bias in that those who *did* respond might not have been answering truthfully.

2.77 **a** The design is good in that it includes several desired experimental design components. Specifically, the design includes a control group, random assignment of the subjects to the treatments, and includes blinding.

b Perhaps rather than taking photos of the top of the head of all the women, the expert who determined the change in hair density should have the opportunity to evaluate the women in person. Additionally, there should have been more than one expert doing the change in hair density evaluation.

2.78 **a** Answers will vary. The most basic experimental design would be as follows. First number the 100 locations in the kiln. Randomly assign 50 of the locations to receive the first type of clay. The remaining 50 locations will receive the second type of clay. Fire the tiles in the kiln. After firing is complete, compare the proportions of cracked tiles for the two types.

Alternatively, select 50 pairs of locations where the temperature is expected to be the same for the two locations in each pair. For each pair, randomly assign one location to the first type of clay and the other location to the other type of clay. Fire the tiles and compare the proportions of cracked tiles for the two types.

b In the first design, the extraneous variable temperature is dealt with by randomly assigning the locations to the clay types. In the second design it is dealt with by blocking by temperature.

Online Exercises

2.79 **a** "forests are being destroyed…80 acres per minute"

b **"**vanishing tropical forests"

c **"**man-made extinction"

d "destruction of tropical forests"

2.80 Answers will vary.

2.81 Answers will vary.

2.82 Answers will vary.

2.83 Answers will vary.

2.84 Answers will vary.