CONTENTS

Introduction to Statistics

CHAPTER 1

1.1 AN OVERVIEW OF STATISTICS

1.1 Try It Yourself Solutions

1a. The population consists of the prices per gallon of regular gasoline at all gasoline stations in the United States.

b. The sample consists of the prices per gallon of regular gasoline at the 800 surveyed stations.

c. The data set consists of the 800 prices.

2a. Because the numerical measure of $2,326,706,685 is based on the entire collection of player's salaries, it is from a population.

b. Because the numerical measure is a characteristic of a population, it is a parameter.

3a. Descriptive statistics involve the statement "76% of women and 60% of men had a physical examination within the previous year."

b. An inference drawn from the study is that a higher percentage of women had a physical examination within the previous year.

1.1 EXERCISE SOLUTIONS

1. A sample is a subset of a population.

2. It is usually impractical (too expensive and time consuming) to obtain all the population data.

3. A parameter is a numerical description of a population characteristic. A statistic is a numerical description of a sample characteristic.

4. Descriptive statistics and inferential statistics.

5. False. A statistic is a numerical measure that describes a sample characteristic.

6. True

7. True

8. False. Inferential statistics involves using a sample to draw conclusions about a population.

9. False. A population is the collection of *all* outcomes, responses, measurements, or counts that are of interest.

10. True

11. The data set is a population because it is a collection of the ages of all the members of the House of Representatives.

12. The data set is a sample because only every fourth person is measured.

13. The data set is a sample because the collection of the 500 spectators is a subset within the population of the stadium's 42,000 spectators.

14. The data set is a population because it is a collection of the annual salaries of all lawyers at a firm.

15. Sample, because the collection of the 20 patients is a subset within the population.

16. The data set is a population since it is a collection of the number of televisions in all U.S. households.

17. Population: Party of registered voters in Warren County.

Sample: Party of Warren County voters responding to phone survey.

18. Population: Major of college students at Caldwell College.

Sample: Major of college students at Caldwell College who take statistics.

19. Population: Ages of adults in the United States who own computers.

Sample: Ages of adults in the United States who own Dell computers.

20. Population: Income of all homeowners in Texas.

Sample: Income of homeowners in Texas with mortgages.

21. Population: All adults in the United States that take vacations.

Sample: Collection of 1000 adults surveyed that take vacations.

22. Population: Collection of all infants in Italy.

Sample: Collection of the 33,043 infants in the study.

23. Population: Collection of all households in the U.S.

Sample: Collection of 1906 households surveyed.

24. Population: Collection of all computer users.

Sample: Collection of 1000 computer users surveyed.

25. Population: Collection of all registered voters.

Sample: Collection of 1045 registered voters surveyed.

26. Population: Collection of all students at a college.

Sample: Collection of 496 college students surveyed.

27. Population: Collection of all women in the U.S.

Sample: Collection of the 546 U.S. women surveyed.

28. Population: Collection of all U.S. vacationers.

Sample: Collection of the 791 U.S. vacationers surveyed.

29. Statistic. The value $68,000 is a numerical description of a sample of annual salaries.

30. Statistic. 43% is a numerical description of a sample of high school students.

31. Parameter. The 62 surviving passengers out of 97 total passengers is a numerical description of all of the passengers of the Hindenburg that survived.

32. Parameter. 44% is a numerical description of the total number of governors.

33. Statistic. 8% is a numerical description of a sample of computer users.

34. Parameter. 12% is a numerical description of all new magazines.

35. Statistic. 53% is a numerical description of a sample of people in the United States.

36. Parameter. 21.1 is a numerical description of ACT scores for all graduates.

37. The statement "56% are the primary investor in their household" is an application of descriptive statistics.

An inference drawn from the sample is that an association exists between U.S. women and being the primary investor in their household.

38. The statement "spending at least $2000 for their next vacation" is an application of descriptive statistics.

An inference drawn from the sample is that U.S. vacationers are associated with spending more than $2000 for their next vacation.

39. Answers will vary.

40. (a) The volunteers in the study represent the sample.

(b) The population is the collection of all individuals who completed the math test.

(c) The statement "three times more likely to answer correctly" is an application of descriptive statistics.

(d) An inference drawn from the sample is that individuals who are not sleep deprived will be three times more likely to answer math questions correctly than individuals who are sleep deprived.

41. (a) An inference drawn from the sample is that senior citizens who live in Florida have better memory than senior citizens who do not live in Florida.

(b) It implies that if you live in Florida, you will have better memory.

42. (a) An inference drawn from the sample is that the obesity rate among boys ages 2 to 19 is increasing.

(b) It implies the same trend will continue in future years.

43. Answers will vary.

1.2 DATA CLASSIFICATION

1.2 Try It Yourself Solutions

1a. One data set contains names of cities and the other contains city populations.

b. City: Nonnumerical
Population: Numerical

c. City: Qualitative
Population: Quantitative

2a. (1) The final standings represent a ranking of basketball teams.

(2) The collection of phone numbers represents labels. No mathematical computations can be made.

b. (1) Ordinal, because the data can be put in order.

(2) Nominal, because you cannot make calculations on the data.

3a. (1) The data set is the collection of body temperatures.

(2) The data set is the collection of heart rates.

b. (1) Interval, because the data can be ordered and meaningful differences can be calculated, but it does not make sense writing a ratio using the temperatures.

(2) Ratio, because the data can be ordered, can be written as a ratio, you can calculate meaningful differences, and the data set contains an inherent zero.

1.2 EXERCISE SOLUTIONS

1. Nominal and ordinal

2. Ordinal, Interval, and Ratio

3. False. Data at the ordinal level can be qualitative or quantitative.

4. False. For data at the interval level, you can calculate meaningful differences between data entries. You cannot calculate meaningful differences at the nominal or ordinal level.

5. False. More types of calculations can be performed with data at the interval level than with data at the nominal level.

6. False. Data at the ratio level can be placed in a meaningful order.

7. Qualitative, because telephone numbers are merely labels.

8. Quantitative, because the daily high temperature is a numerical measure.

9. Quantitative, because the lengths of songs on an MP3 player are numerical measures.

10. Qualitative, because the player numbers are merely labels.

11. Qualitative, because the poll results are merely responses.

12. Quantitative, because the diastolic blood pressure is a numerical measure.

13. Qualitative. Ordinal. Data can be arranged in order, but differences between data entries make no sense.

14. Qualitative. Nominal. No mathematical computations can be made and data are categorized using names.

15. Qualitative. Nominal. No mathematical computations can be made and data are categorized using names.

16. Quantitative. Ratio. A ratio of two data values can be formed so one data value can be expressed as a multiple of another.

17. Qualitative. Ordinal. The data can be arranged in order, but differences between data entries are not meaningful.

18. Quantitative. Ratio. The ratio of two data values can be formed so one data value can be expressed as a multiple of another.

19. Ordinal	**20.** Ratio	**21.** Nominal	**22.** Ratio
23. (a) Interval	(b) Nominal	(c) Ratio	(d) Ordinal
24. (a) Interval	(b) Nominal	(c) Interval	(d) Ratio

25. An inherent zero is a zero that implies "none." Answers will vary.

26. Answers will vary.

1.3 EXPERIMENTAL DESIGN

1.3 Try It Yourself Solutions

1a. (1) Focus: Effect of exercise on relieving depression.

(2) Focus: Success of graduates.

b. (1) Population: Collection of all people with depression.

(2) Population: Collection of all university graduates.

c. (1) Experiment

(2) Survey

2a. There is no way to tell why people quit smoking. They could have quit smoking either from the gum or from watching the DVD.

b. Two experiments could be done; one using the gum and the other using the DVD.

3a. Example: start with the first digits 92630782 . . .

b. 92|63|07|82|40|19|26

c. 63, 7, 40, 19, 26

4a. (1) The sample was selected by only using available students.

(2) The sample was selected by numbering each student in the school, randomly choosing a starting number, and selecting students at regular intervals from the starting number.

b. (1) Because the students were readily available in your class, this is convenience sampling.

(2) Because the students were ordered in a manner such that every 25th student is selected, this is systematic sampling.

1.3 EXERCISE SOLUTIONS

1. In an experiment, a treatment is applied to part of a population and responses are observed. In an observational study, a researcher measures characteristics of interest of part of a population but does not change existing conditions.

2. A census includes the entire population; a sample includes only a portion of the population.

3. Assign numbers to each member of the population and use a random number table or use a random number generator.

4. Replication is the repetition of an experiment using a large group of subjects. It is important because it gives validity to the results.

5. True

6. False. A double-blind experiment is used to decrease the placebo effect.

7. False. Using stratified sampling guarantees that members of each group within a population will be sampled.

8. False. A census is a count of an entire population.

9. False. To select a systematic sample, a population is ordered in some way and then members of the population are selected at regular intervals.

10. True

11. In this study, you want to measure the effect of a treatment (using a fat substitute) on the human digestive system. So, you would want to perform an experiment.

12. It would be nearly impossible to ask every consumer whether he or she would still buy a product with a warning label. So, you should use a survey to collect these data.

13. Because it is impractical to create this situation, you would want to use a simulation.

14. Because the U.S. Congress keeps accurate financial records of all members, you could take a census.

15. (a) The experimental units are the 30–35 year old females being given the treatment.

    (b) One treatment is used.

    (c) A problem with the design is that there may be some bias on the part of the researchers if he or she knows which patients were given the real drug. A way to eliminate this problem would be to make the study into a double-blind experiment.

    (d) The study would be a double-blind study if the researcher did not know which patients received the real drug or the placebo.

16. (a) The experimental units are the people with early signs of arthritis.

    (b) One treatment is used.

    (c) A problem with the design is that the sample size is small. The experiment could be replicated to increase validity.

    (d) In a placebo-controlled double-blind experiment, neither the subject nor the experimenter knows whether the subject is receiving a treatment or a placebo. The experimenter is informed after all the data have been collected.

    (e) The group could be randomly split into 20 males or 20 females in each treatment group.

17. Each U.S. telephone number has an equal chance of being dialed and all samples of 1599 phone numbers have an equal chance of being selected, so this is a simple random sample. Telephone sampling only samples those individuals who have telephones, are available, and are willing to respond, so this is a possible source of bias.

18. Because the persons are divided into strata (rural and urban), and a sample is selected from each stratum, this is a stratified sample.

19. Because the students were chosen due to their convenience of location (leaving the library), this is a convenience sample. Bias may enter into the sample because the students sampled may not be representative of the population of students. For example, there may be an association between time spent at the library and drinking habits.

20. Because the disaster area was divided into grids and thirty grids were then entirely selected, this is a cluster sample. Certain grids may have been much more severely damaged than others, so this is a possible source of bias.

21. Because a random sample of out-patients were selected and all samples of 1210 patients had an equal chance of being selected, this is a simple random sample.

22. Because every twentieth engine part is sampled from an assembly line, this is a systematic sample. It is possible for bias to enter into the sample if, for some reason, the assembly line performs differently on a consistent basis.

23. Because a sample is taken from each one-acre subplot (stratum), this is a stratified sample.

24. Because a sample is taken from members of a population that are readily available, this is a convenience sample. The sample may be biased if the teachers sampled are not representative of the population of teachers. For example, some teachers may frequent the lounge more often than others.

25. Because every ninth name on a list is being selected, this is a systematic sample.

26. Each telephone has an equal chance of being dialed and all samples of 1012 phone numbers have an equal chance of being selected, so this is a simple random sample. Telephone sampling only samples those individuals who have telephones, are available, and are willing to respond, so this is a possible source of bias.

27. Answers will vary.

28. Answers will vary.

29. Census, because it is relatively easy to obtain the salaries of the 50 employees.

30. Sampling, because the population of students is too large to easily record their color. Random sampling would be advised since it would be easy to randomly select students then record their favorite car color.

31. Question is biased because it already suggests that drinking fruit juice is good for you. The question might be rewritten as “How does drinking fruit juice affect your health?”

32. Question is biased because it already suggests that drivers who change lanes several times are dangerous. The question might be rewritten as “Are drivers who change lanes several times dangerous?”

33. Question is unbiased because it does not imply how many hours of sleep are good or bad.

34. Question is biased because it already suggests that the media has a negative effect on teen girls’ dieting habits. The question might be rewritten as “Do you think the media has an effect on teen girls’ dieting habits?”

35. The households sampled represent various locations, ethnic groups, and income brackets. Each of these variables is considered a stratum.

36. Stratified sampling ensures that each segment of the population is represented.

37. Open Question

Advantage: Allows respondent to express some depth and shades of meaning in the answer.

Disadvantage: Not easily quantified and difficult to compare surveys.

Closed Question

Advantage: Easy to analyze results.

Disadvantage: May not provide appropriate alternatives and may influence the opinion of the respondent.

38. (a) Advantage: Usually results in a savings in the survey cost.

(b) Disadvantage: There tends to be a lower response rate and this can introduce a bias into the sample.

Sampling Technique: Convenience sampling

39. Answers will vary.

40. If blinding is not used, then the placebo effect is more likely to occur.

41. The Hawthorne effect occurs when a subject changes behavior because he or she is in an experiment. However, the placebo effect occurs when a subject reacts favorably to a placebo he or she has been given.

42. Both a randomized block design and a stratified sample split their members into groups based on similar characteristics.

43. Answers will vary.

CHAPTER 1 REVIEW EXERCISE SOLUTIONS

1. Population: Collection of all U.S. adults.

Sample: Collection of the 1000 U.S. adults that were sampled.

2. Population: Collection of all nurses in San Francisco area.

Sample: Collection of 38 nurses in San Francisco area that were sampled.

3. Population: Collection of all credit cards.

Sample: Collection of 146 credit cards that were sampled.

4. Population: Collection of all physicians in the U.S.

Sample: Collection of 1205 physicians that were sampled.

5. The team payroll is a parameter since it is a numerical description of a population (entire baseball team) characteristic.

6. Since 42% is describing a characteristic of the sample, this is a statistic.

7. Since "10 students" is describing a characteristic of a population of math majors, it is a parameter.

8. Since 19% is describing a characteristic of a sample of Indiana ninth graders, this is a statistic.

9. The average late fee of $27.46 charged by credit cards is representative of the descriptive branch of statistics. An inference drawn from the sample is that all credit cards charge a late fee of $27.46.

10. 60% of all physicians surveyed consider leaving the practice of medicine because they are discouraged over the state of U.S. healthcare is representative of the descriptive branch of statistics. An inference drawn from the sample is that 60% of all physicians surveyed consider leaving the practice of medicine because they are discouraged over the state of U.S. healthcare.

11. Quantitative because monthly salaries are numerical measurements.

12. Qualitative because Social Security numbers are merely labels for employees.

13. Quantitative because ages are numerical measurements.

14. Qualitative because zip codes are merely labels for the customers.

15. Interval. It makes no sense saying that 100 degrees is twice as hot as 50 degrees.

16. Ordinal. The data are qualitative but could be arranged in order of car size.

17. Nominal. The data are qualitative and cannot be arranged in a meaningful order.

18. Ratio. The data are numerical, and it makes sense saying that one player is twice as tall as another player.

19. Because CEOs keep accurate records of charitable donations, you could take a census.

20. Because it is impractical to create this situation, you would want to perform a simulation.

21. In this study, you want to measure the effect of a treatment (fertilizer) on a soybean crop. You would want to perform an experiment.

22. Because it would be nearly impossible to ask every college student about his/her opinion on environmental pollution, you should take a survey to collect the data.

23. The subjects could be split into male and female and then be randomly assigned to each of the five treatment groups.

24. Number the volunteers and then use a random number generator to randomly assign subjects to one of the treatment groups or the control group.

25. Because random telephone numbers were generated and called, this is a simple random sample.

26. Because the student sampled a convenient group of friends, this is a convenience sample.

27. Because each community is considered a cluster and every pregnant woman in a selected community is surveyed, this is a cluster sample.

28. Because every third car is stopped, this is a systematic sample.

29. Because grade levels are considered strata and 25 students are sampled from each stratum, this is a stratified sample.

30. Because of the convenience of surveying people waiting for their baggage, this is a convenience sample.

31. Telephone sampling only samples individuals who have telephones, are available, and are willing to respond.

32. Due to the convenience sample taken, the study may be biased toward the opinions of the student's friends.

33. The selected communities may not be representative of the entire area.

32. It may be difficult for the law enforcement official to stop every third car.

CHAPTER 1 QUIZ SOLUTIONS

1. Population: Collection of all individuals with anxiety disorders.

   Sample: Collection of 372 patients in study.

2. (a) Statistic. 19% is a characteristic of a sample of Internet users.

   (b) Parameter. 84% is a characteristic of the entire company (population).

   (c) Statistic. 40% is a characteristic of a sample of Americans.

3. (a) Qualitative, since post office box numbers are merely labels.

   (b) Quantitative, since a final exam is a numerical measure.

4. (a) Nominal. Badge numbers may be ordered numerically, but there is no meaning in this order and no mathematical computations can be made.

   (b) Ratio. It makes sense to say that the number of candles sold during the 1st quarter was twice as many as sold in the 2nd quarter.

   (c) Interval because meaningful differences between entries can be calculated, but a zero entry is not an inherent zero.

5. (a) In this study, you want to measure the effect of a treatment (low dietary intake of vitamin C and iron) on lead levels in adults. You want to perform an experiment.

   (b) Because it would be difficult to survey every individual within 500 miles of your home, sampling should be used.

6. Randomized Block Design

7. (a) Because people were chosen due to their convenience of location (on the campground), this is a convenience sample.

   (b) Because every tenth part is selected from an assembly line, this is a systematic sample.

   (c) Stratified sample because the population is first stratified and then a sample is collected from each stratum.

8. Convenience

Descriptive Statistics

2.1 FREQUENCY DISTRIBUTIONS AND THEIR GRAPHS

2.1 Try It Yourself Solutions

1a. The number of classes (8) is stated in the problem.

b. Min $= 15$ Max $= 89$ Class width $= \dfrac{(89 - 15)}{8} = 9.25 \Rightarrow 10$

c.

Lower limit	Upper limit
15	24
25	34
35	44
45	54
55	64
65	74
75	84
85	94

d. See part (e).

e.

Class	Frequency, f
15–24	16
25–34	34
35–44	30
45–54	23
55–64	13
65–74	2
75–84	0
85–94	1

2a. See part (b).

b.

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
15–24	16	19.5	0.13	16
25–34	34	29.5	0.29	50
35–44	30	39.5	0.25	80
45–54	23	49.5	0.19	103
55–64	13	59.5	0.11	116
65–74	2	69.5	0.02	118
75–84	0	79.5	0.00	118
85–94	1	89.5	0.01	119
	$\sum f = 119$		$\sum \frac{f}{n} = 1$	

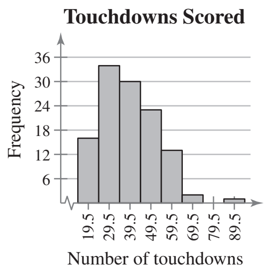
c. 86% of the teams scored fewer than 55 touchdowns. 3% of the teams scored more than 65 touchdowns.

3a.

Class Boundaries
14.5–24.5
24.5–34.5
34.5–44.5
44.5–54.5
54.5–64.5
64.5–74.5
74.5–84.5
84.5–94.5

b. Use class midpoints for the horizontal scale and frequency for the vertical scale.

c.

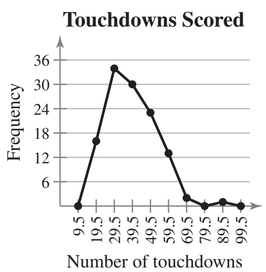


d. 86% of the teams scored fewer than 55 touchdowns. 3% of the teams scored more than 65 touchdowns.

4a. Use class midpoints for the horizontal scale and frequency for the vertical scale.

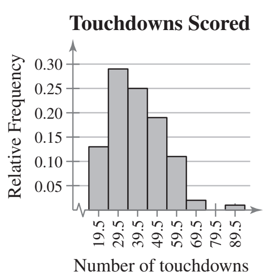
b. See part (c).

c.



d. The number of touchdowns increases until 34.5 touchdowns, then decreases afterward.

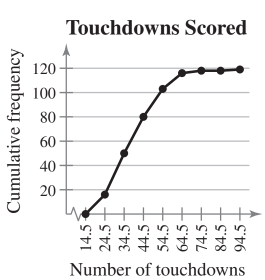
5abc.



6a. Use upper class boundaries for the horizontal scale and cumulative frequency for the vertical scale.

b. See part (c).

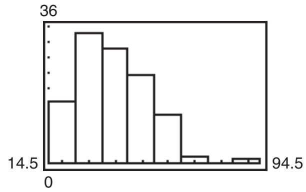
c.



d. Approximately 80 teams scored 44 or fewer touchdowns.

e. Answers will vary.

7ab.



2.1 EXERCISE SOLUTIONS

1. By organizing the data into a frequency distribution, patterns within the data may become more evident.

2. Sometimes it is easier to identify patterns of a data set by looking at a graph of the frequency distribution.

3. Class limits determine which numbers can belong to that class.

   Class boundaries are the numbers that separate classes without forming gaps between them.

4. Cumulative frequency is the sum of the frequency for that class and all previous classes. Relative frequency is the proportion of entries in each class.

5. False. Class width is the difference between the lower and upper limits of consecutive classes.

6. True

7. False. An ogive is a graph that displays cumulative frequency.

8. True

9. $\text{Width} = \dfrac{\text{Max} - \text{Min}}{\text{Classes}} = \dfrac{58 - 7}{6} = 8.5 \Rightarrow 9$

   Lower class limits: 7, 16, 25, 34, 43, 52

   Upper class limits: 15, 24, 33, 42, 51, 60

10. $\text{Width} = \dfrac{\text{Max} - \text{Min}}{\text{Classes}} = \dfrac{94 - 11}{8} = 10.375 \Rightarrow 11$

    Lower class limits: 11, 22, 33, 44, 55, 66, 77, 88

    Upper class limits: 21, 32, 43, 54, 65, 76, 87, 98

11. $\text{Width} = \dfrac{\text{Max} - \text{Min}}{\text{Classes}} = \dfrac{123 - 15}{6} = 18 \Rightarrow 19$

    Lower class limits: 15, 34, 53, 72, 91, 110

    Upper class limits: 33, 52, 71, 90, 109, 128

12. $\text{Width} = \dfrac{\text{Max} - \text{Min}}{\text{Classes}} = \dfrac{171 - 24}{10} = 14.7 \Rightarrow 15$

    Lower class limits: 24, 39, 54, 69, 84, 99, 114, 129, 144, 159

    Upper class limits: 38, 53, 68, 83, 98, 113, 128, 143, 158, 173

13. (a) Class width $= 31 - 20 = 11$

(b) and (c)

Class	Frequency, f	Midpoint	Class boundaries
20–30	19	25	19.5–30.5
31–41	43	36	30.5–41.5
42–52	68	47	41.5–52.5
53–63	69	58	52.5–63.5
64–74	74	69	63.5–74.5
75–85	68	80	74.5–85.5
86–96	24	91	85.5–96.5
	$\sum f = 365$		

14a. Class width $= 10 - 0 = 10$

bc.

Class	Frequency	Midpoint	Class boundaries
0–9	188	4.5	−0.5–9.5
10–19	372	14.5	9.5–19.5
20–29	264	24.5	19.5–29.5
30–39	205	34.5	29.5–39.5
40–49	83	44.5	39.5–49.5
50–59	76	54.5	49.5–59.5
60–69	32	64.5	59.5–69.5
	$\Sigma f = 1220$		

15.

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
20–30	19	25	0.05	19
31–41	43	36	0.12	62
42–52	68	47	0.19	130
53–63	69	58	0.19	199
64–74	74	69	0.20	273
75–85	68	80	0.19	341
86–96	24	91	0.07	365
	$\sum f = 365$		$\sum \frac{f}{n} = 1$	

16.

Class	Frequency	Midpoint	Relative frequency	Cumulative frequency
0–9	188	4.5	0.15	188
10–19	372	14.5	0.30	560
20–29	264	24.5	0.22	824
30–39	205	34.5	0.17	1029
40–49	83	44.5	0.07	1112
50–59	76	54.5	0.06	1188
60–69	32	64.5	0.03	1220
	$\Sigma f = 1220$		$\sum \frac{f}{n} = 1$	

17. (a) Number of classes $= 7$

(b) Least frequency ≈ 10

(c) Greatest frequency ≈ 300

(d) Class width $= 10$

18. (a) Number of classes $= 7$

(b) Least frequency ≈ 100

(c) Greatest frequency ≈ 900

(d) Class width $= 5$

19. (a) 50 (b) 22.5–24.5 lbs

20. (a) 50 (b) 64–66 inches

21. (a) 24 (b) 29.5 lbs

22. (a) 44 (b) 66 inches

23. (a) Class with greatest relative frequency: 8–9 inches.
Class with least relative frequency: 17–18 inches.

(b) Greatest relative frequency ≈ 0.195
Least relative frequency ≈ 0.005

(c) Approximately 0.015

24. (a) Class with greatest relative frequency: 19–20 minutes.
Class with least relative frequency: 21–22 minutes.

(b) Greatest relative frequency $\approx 40\%$
Least relative frequency $\approx 2\%$

(c) Approximately 33%

25. Class with greatest frequency: 500–550
Class with least frequency: 250–300 and 700–750

26. Class with greatest frequency: 7.75–8.25
Class with least frequency: 6.25–6.75

27. $\text{Class width} = \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{39 - 0}{5} = 7.8 \Rightarrow 8$

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
0–7	8	3.5	0.32	8
8–15	8	11.5	0.32	16
16–23	3	19.5	0.12	19
24–31	3	27.5	0.12	22
32–39	3	35.5	0.12	25
	$\sum f = 25$		$\sum \frac{f}{n} = 1$	

Class with greatest frequency: 0–7, 8–15
Class with least frequency: 16–23, 24–31, 32–39

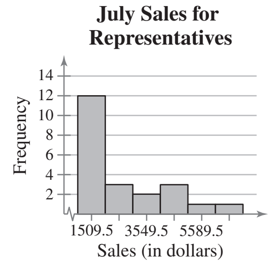
28. $\text{Class width} = \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{530 - 30}{6} = 83.3 \Rightarrow 84$

Class	Frequency	Midpoint	Relative frequency	Cumulative frequency
30–113	5	71.5	0.1724	5
114–197	7	155.5	0.2414	12
198–281	8	239.5	0.2759	20
282–365	2	323.5	0.0690	22
366–449	3	407.5	0.1034	25
450–533	4	491.5	0.1379	29
	$\Sigma f = 29$		$\Sigma \frac{f}{n} = 1$	

Class with greatest frequency: 198–281
Class with least frequency: 282–365

29. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{7119 - 1000}{6} = 1019.83 \Longrightarrow 1020$

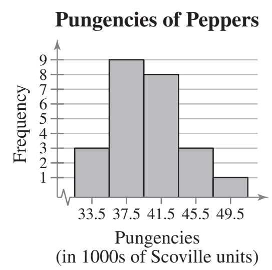
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
1000–2019	12	1509.5	0.5455	12
2020–3039	3	2529.5	0.1364	15
3040–4059	2	3549.5	0.0909	17
4060–5079	3	4569.5	0.1364	20
5080–6099	1	5589.5	0.0455	21
6100–7119	1	6609.5	0.0455	22
	$\sum f = 22$		$\sum \frac{f}{n} = 1$	



Class with greatest frequency: 1000–2019
Class with least frequency: 5080–6099; 6100–7119

30. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{51 - 32}{5} = 3.8 \Longrightarrow 4$

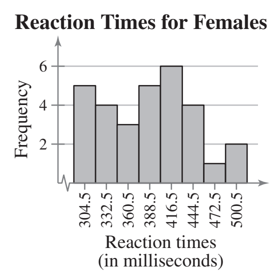
Class	Frequency	Midpoint	Relative frequency	Cumulative frequency
32–35	3	33.5	0.1250	3
36–39	9	37.5	0.3750	12
40–43	8	41.5	0.3333	20
44–47	3	45.5	0.1250	23
48–51	1	49.5	0.0417	24
	$\sum f = 24$		$\sum \frac{f}{n} = 1$	



Class with greatest frequency: 36–39

31. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{514 - 291}{8} = 27.875 \Longrightarrow 28$

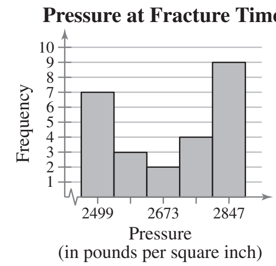
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
291–318	5	304.5	0.1667	5
319–346	4	332.5	0.1333	9
347–374	3	360.5	0.1000	12
375–402	5	388.5	0.1667	17
403–430	6	416.5	0.2000	23
431–458	4	444.5	0.1333	27
459–486	1	472.5	0.0333	28
487–514	2	500.5	0.0667	30
	$\sum f = 30$		$\sum \frac{f}{n} = 1$	



Class with greatest frequency: 403–430

32. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{2888 - 2456}{5} = 86.4 \Rightarrow 87$

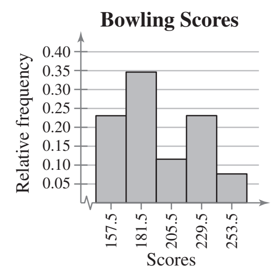
Class	Frequency	Midpoint	Relative frequency	Cumulative frequency
2456–2542	7	2499	0.28	7
2543–2629	3	2586	0.12	10
2630–2716	2	2673	0.08	12
2717–2803	4	2760	0.16	16
2804–2890	9	2847	0.36	25
	$\sum f = 25$		$\sum \frac{f}{n} = 1$	



Class with greatest frequency: 2804–2890
Class with least frequency: 2630–2716

33. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{264 - 146}{5} = 23.6 \Rightarrow 24$

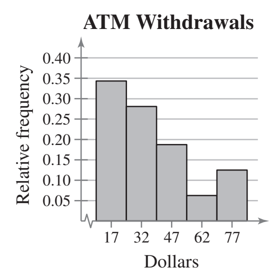
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
146–169	6	157.5	0.2308	6
170–193	9	181.5	0.3462	15
194–217	3	205.5	0.1154	18
218–241	6	229.5	0.2308	24
242–265	2	253.5	0.0769	26
	$\sum f = 26$		$\sum \frac{f}{n} = 1$	



Class with greatest relative frequency: 170–193
Class with least relative frequency: 242–265

34. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{80 - 10}{5} = 14 \Rightarrow 15$

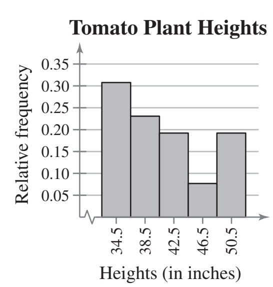
Class	Frequency	Midpoint	Relative frequency	Cumulative frequency
10–24	11	17	0.3438	11
25–39	9	32	0.2813	20
40–54	6	47	0.1875	26
55–69	2	62	0.0625	28
70–84	4	77	0.1250	32
	$\sum f = 32$		$\sum \frac{f}{n} = 1$	



Class with greatest relative frequency: 10–24
Class with least relative frequency: 55–69

35. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{52 - 33}{5} = 3.8 \Rightarrow 4$

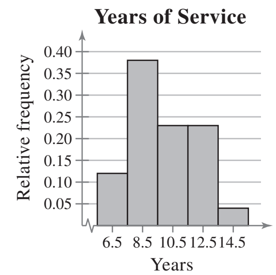
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
33–36	8	34.5	0.3077	8
37–40	6	38.5	0.2308	14
41–44	5	42.5	0.1923	19
45–48	2	46.5	0.0769	21
49–52	5	50.5	0.1923	26
	$\sum f = 26$		$\sum \frac{f}{n} = 1$	



Class with greatest relative frequency: 33–36
Class with least relative frequency: 45–48

36. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{16 - 7}{5} = 1.8 \Rightarrow 2$

Class	Frequency	Midpoint	Relative frequency	Cumulative frequency
6–7	3	6.5	0.12	3
8–9	10	8.5	0.38	13
10–11	6	10.5	0.23	19
12–13	6	12.5	0.23	25
14–15	1	14.5	0.04	26
	$\sum f = 26$		$\sum \frac{f}{n} = 1$	



Class with greatest relative frequency: 8–9
Class with least relative frequency: 14–15

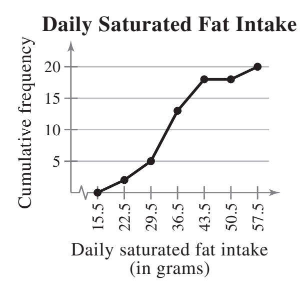
37. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{73 - 52}{6} = 3.5 \Rightarrow 4$

Class	Frequency, f	Relative frequency	Cumulative frequency
52–55	3	0.125	3
56–59	3	0.125	6
60–63	9	0.375	15
64–67	4	0.167	19
68–71	4	0.167	23
72–75	1	0.042	24
	$\sum f = 24$	$\sum \frac{f}{n} \approx 1$	

Location of the greatest increase in frequency: 60–63

38. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{57 - 16}{6} = 6.83 \Rightarrow 7$

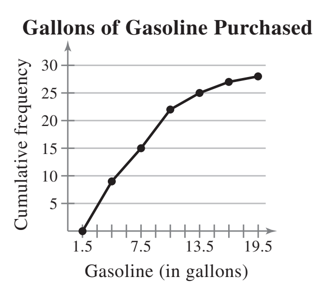
Class	Frequency, f	Relative frequency	Cumulative frequency
16–22	2	0.10	2
23–29	3	0.15	5
30–36	8	0.40	13
37–43	5	0.25	18
44–50	0	0.00	18
51–57	2	0.10	20
	$\sum f = 20$	$\sum \frac{f}{n} = 1$	



Location of the greatest increase in frequency: 30–36

39. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{18 - 2}{6} = 2.67 \Rightarrow 3$

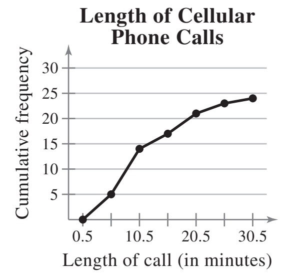
Class	Frequency, f	Relative frequency	Cumulative frequency
2–4	9	0.3214	9
5–7	6	0.2143	15
8–10	7	0.2500	22
11–13	3	0.1071	25
14–16	2	0.0714	27
17–19	1	0.0357	28
	$\sum f = 28$	$\sum \frac{f}{n} \approx 1$	



Location of the greatest increase in frequency: 2–4

40. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{29 - 1}{6} = 4.67 \Longrightarrow 5$

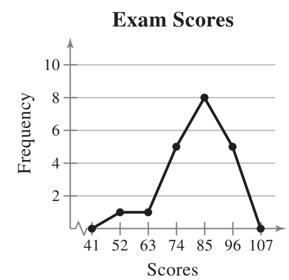
Class	Frequency, f	Relative frequency	Cumulative frequency
1–5	5	0.2083	5
6–10	9	0.3750	14
11–15	3	0.1250	17
16–20	4	0.1667	21
21–25	2	0.0833	23
26–30	1	0.0417	24
	$\sum f = 24$	$\sum \frac{f}{n} = 1$	



Location of the greatest increase in frequency: 6–10

41. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{98 - 47}{5} = 10.2 \Longrightarrow 11$

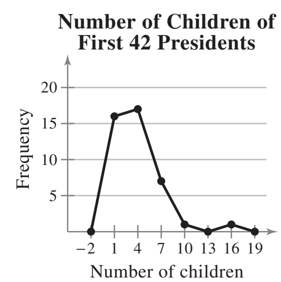
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
47–57	1	52	0.05	1
58–68	1	63	0.05	2
69–79	5	74	0.25	7
80–90	8	85	0.40	15
91–101	5	96	0.25	20
	$\sum f = 20$		$\sum \frac{f}{N} = 1$	



Class with greatest frequency: 80–90
Classes with least frequency: 47–57 and 58–68

42.

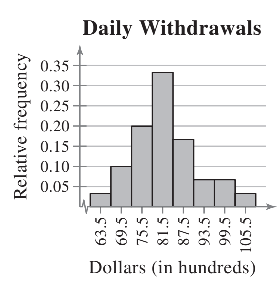
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
0–2	16	1	0.3810	16
3–5	17	4	0.4048	33
6–8	7	7	0.1667	40
9–11	1	10	0.0238	41
12–14	0	13	0.0000	41
15–17	1	16	0.0238	42
	$\sum f = 42$		$\sum \frac{f}{N} = 1$	



Classes with greatest frequency: 0–2
Classes with least frequency: 15–17

43. (a) Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{104 - 61}{8} = 5.375 \Longrightarrow 6$

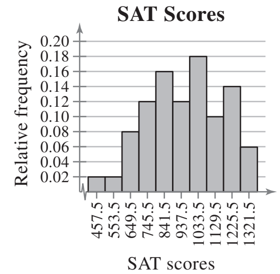
Class	Frequency, f	Midpoint	Relative frequency
61–66	1	63.5	0.0333
67–72	3	69.5	0.1000
73–78	6	75.5	0.2000
79–84	10	81.5	0.3333
85–90	5	87.5	0.1667
91–96	2	93.5	0.0667
97–102	2	99.5	0.0667
103–108	1	105.5	0.0333
	$\sum f = 30$		$\sum \frac{f}{N} = 1$



(b) 16.7%, because the sum of the relative frequencies for the last three classes is 0.167.

(c) \$9600, because the sum of the relative frequencies for the last two classes is 0.10.

44. (a) $\text{Class width} = \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{1359 - 410}{10} = 94.9 \Rightarrow 95$

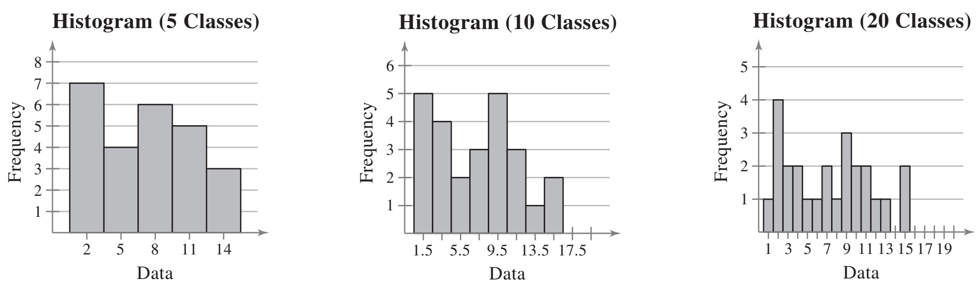


Class	Frequency	Midpoint	Relative frequency
410–505	1	457.5	0.02
506–601	1	553.5	0.02
602–697	4	649.5	0.08
698–793	6	745.5	0.12
794–889	8	841.5	0.16
890–985	6	937.5	0.12
986–1081	9	1033.5	0.18
1082–1177	5	1129.5	0.10
1178–1273	7	1225.5	0.14
1274–1369	3	1321.5	0.06
	$\sum f = 50$		$\sum \frac{f}{N} = 1$

(b) 48%, because the sum of the relative frequencies for the last four classes is 0.48.

(c) 698, because the sum of the relative frequencies for the last seven classes is 0.88.

45.



In general, a greater number of classes better preserves the actual values of the data set, but is not as helpful for observing general trends and making conclusions. When choosing the number of classes, an important consideration is the size of the data set. For instance, you would not want to use 20 classes if your data set contained 20 entries. In this particular example, as the number of classes increases, the histogram shows more fluctuation. The histograms with 10 and 20 classes have classes with zero frequencies. Not much is gained by using more than five classes. Therefore, it appears that five classes would be best.

2.2 MORE GRAPHS AND DISPLAYS

2.2 Try It Yourself Solutions

1a.

```
1 |
2 |
3 |
4 |
5 |
6 |
7 |
8 |
```

b. Key: 1 | 7 = 17

```
1 | 7 5 8 8 5 5
2 | 7 6 8 9 8 9 7 9 8 7 5 3 4 6 2 5 0 1 1 2 1 4 1
3 | 9 9 7 7 8 8 6 4 7 6 5 5 5 1 4 5 9 8 2 5 2 2 2 3 3 3 2 4 1 1 0 4 2 1 2 1 0
4 | 9 8 6 8 7 8 6 4 8 5 4 6 6 7 1 1 5 4 5 3 2 2 8 3 0 4 0 5 3 0
5 | 9 4 5 4 3 5 5 9 0 2 3 5 7 0 5
6 | 8 5 1 3 3 1 1 0
7 |
8 | 9
```

c. Key: 1 | 7 = 17

1	5 5 5 7 8 8
2	0 1 1 1 1 2 2 3 4 4 5 5 6 6 7 7 7 8 8 8 9 9 9
3	0 0 1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 4 4 4 4 5 5 5 5 5 6 6 7 7 7 8 8 8 9 9 9
4	0 0 0 1 1 2 2 3 3 3 4 4 4 4 5 5 5 5 6 6 6 6 7 7 8 8 8 8 8 9
5	0 0 2 3 3 4 4 5 5 5 5 7 9 9
6	0 1 1 1 3 3 5 8
7	
8	9

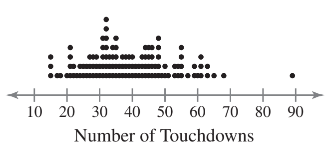
d. It seems that most teams scored under 54 touchdowns.

2ab. Key: 1 | 7 = 17

1	
1	5 5 5 7 8 8
2	0 1 1 1 1 2 2 3 4 4
2	5 5 6 6 7 7 7 8 8 8 9 9 9
3	0 0 1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 4 4 4 4
3	5 5 5 5 5 6 6 7 7 7 8 8 8 9 9 9
4	0 0 0 1 1 2 2 3 3 3 4 4 4 4
4	5 5 5 5 6 6 6 6 7 7 8 8 8 8 8 9
5	0 0 2 3 3 4 4
5	5 5 5 5 7 9 9
6	0 1 1 1 3 3
6	5 8
7	
7	
8	
8	9

3a. Use number of touchdowns for the horizontal axis.

b. **Touchdowns Scored**

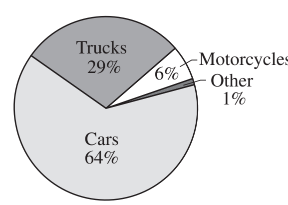


c. It appears that a large percentage of teams scored under 50 touchdowns.

4a.

Vehicle type	Killed (frequency)	Relative frequency	Central angle
Cars	22,423	0.64	$(0.64)(360°) \approx 230°$
Trucks	10,216	0.29	$(0.29)(360°) \approx 104°$
Motorcycles	2,227	0.06	$(0.06)(360°) \approx 22°$
Other	425	0.01	$(0.01)(360°) \approx 4°$
	$\sum f = 35{,}291$	$\sum \frac{f}{n} \approx 1$	$\sum = 360°$

b. **Motor Vehicle Occupants Killed in 1995**

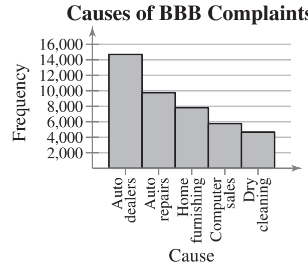


c. As a percentage of total vehicle deaths, car deaths decreased by 15%, truck deaths increased by 8%, and motorcycle deaths increased by 6%.

5a.

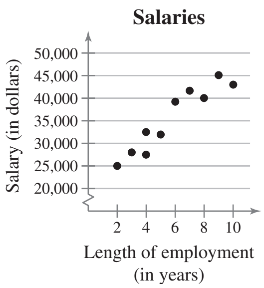
Cause	Frequency, *f*
Auto Dealers	14,668
Auto Repair	9,728
Home Furnishing	7,792
Computer Sales	5,733
Dry Cleaning	4,649

b.



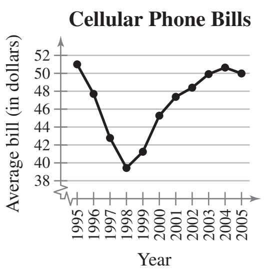
c. It appears that the auto industry (dealers and repair shops) account for the largest portion of complaints filed at the BBB.

6ab.



c. It appears that the longer an employee is with the company, the larger his/her salary will be.

7ab.



c. It appears that the average monthly bill for cellular telephone subscribers decreased significantly from 1995 to 1998, then increased from 1998 to 2004.

2.2 EXERCISE SOLUTIONS

1. Quantitative: Stem-and-Leaf Plot, Dot Plot, Histogram, Time Series Chart, Scatter Plot

Qualitative: Pie Chart, Pareto Chart

2. Unlike the histogram, the stem-and-leaf plot still contains the original data values. However, some data are difficult to organize in a stem-and-leaf plot.

3. Both the stem-and-leaf plot and the dot plot allow you to see how data are distributed, determine specific data entries, and identify unusual data values.

4. In the pareto chart, the height of each bar represents frequency or relative frequency and the bars are positioned in order of decreasing height with the tallest bar positioned to the left.

5. b **6.** d **7.** a **8.** c

9. 27, 32, 41, 43, 43, 44, 47, 47, 48, 50, 51, 51, 52, 53, 53, 53, 54, 54, 54, 54, 55, 56, 56, 58, 59, 68, 68, 68, 73, 78, 78, 85

Max: 85 Min: 27

10. 12.9, 13.3, 13.6, 13.7, 13.7, 14.1, 14.1, 14.1, 14.1, 14.3, 14.4, 14.4, 14.6, 14.9, 14.9, 15.0, 15.0, 15.0, 15.1, 15.2, 15.4, 15.6, 15.7, 15.8, 15.8, 15.8, 15.9, 16.1, 16.6, 16.7

Max: 16.7 Min: 12.9

11. 13, 13, 14, 14, 14, 15, 15, 15, 15, 15, 16, 17, 17, 18, 19

Max: 19 Min: 13

12. 214, 214, 214, 216, 216, 217, 218, 218, 220, 221, 223, 224, 225, 225, 227, 228, 228, 228, 228, 230, 230, 231, 235, 237, 239

Max: 239 Min: 214

13. Anheuser-Busch is the top sports advertiser spending approximately $190 million. Honda spends the least. (Answers will vary.)

14. The value of the stock portfolio has increased fairly steadily over the past five years with the greatest increase happening between 2003 and 2006. (Answers will vary.)

15. Tailgaters irk drivers the most, while too cautious drivers irk drivers the least. (Answers will vary.)

16. The most frequent incident occurring while driving and using a cell phone is swerving. Twice as many people "sped up" than "cut off a car." (Answers will vary.)

17. Key: $6|7 = 67$

6	7 8
7	3 5 5 6 9
8	0 0 2 3 5 5 7 7 8
9	0 1 1 1 2 4 5 5

Most grades of the biology midterm were in the 80s or 90s.

18. Key: $4|0 = 40$

4	0 7 9 9
5	0 1 2 4 6 8 9 9
6	1 2 3 7
7	1 3 6 8 9
8	0 4 4 7

It appears that most of the world's richest people are over 49 years old. (Answers will vary.)

19. Key: $4|3 = 4.3$

4	3 9
5	1 8 8 8 9
6	4 8 9 9 9
7	0 0 2 2 2 5
8	0 1

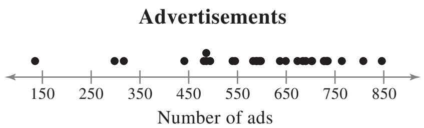
It appears that most ice had a thickness of 5.8 centimeters to 7.2 centimeters. (Answers will vary.)

20. Key: $17|5 = 17.5$

16	4 8
17	1 1 3 4 5 5 6 7 9
18	1 3 4 4 6 6 6 9
19	0 0 2 3 3 5 6
20	1 8

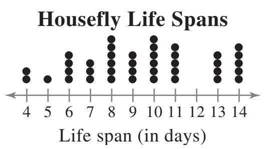
It appears that most farmers charge 17 to 19 cents per pound of apples. (Answers will vary.)

21.



It appears that most of the 30 people from the U.S. see or hear between 450 and 750 advertisements per week. (Answers will vary.)

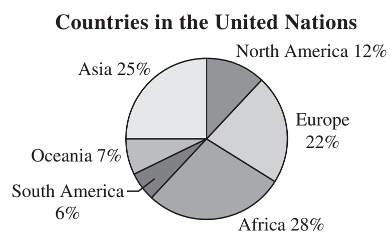
22.



It appears that the lifespan of a fly tends to be between 8 and 11 days. (Answers will vary.)

23.

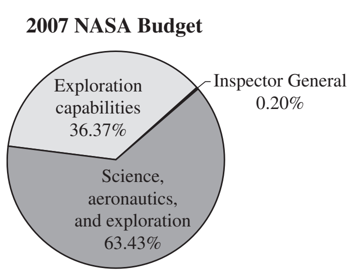
Category	Frequency	Relative frequency	Angle
North America	23	0.12	$(0.12)(360°) \approx 43°$
South America	12	0.06	$(0.06)(360°) \approx 23°$
Europe	43	0.22	$(0.22)(360°) \approx 81°$
Oceania	14	0.07	$(0.07)(360°) \approx 26°$
Africa	53	0.28	$(0.28)(360°) \approx 99°$
Asia	47	0.25	$(0.25)(360°) \approx 88°$
	$\sum f = 192$	$\sum \frac{f}{n} = 1$	



Most countries in the United Nations come from Africa and the least amount come from South America. (Answers will vary.)

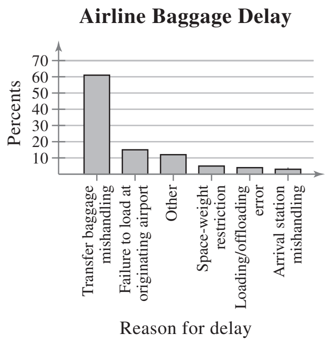
24.

Category	Budget frequency	Relative frequency	Angle
Science, aeronautics, and exploration	10,651	0.6343	$(0.6343)(360°) \approx 228°$
Exploration capabilities	6,108	0.3637	$(0.3637)(360°) \approx 131°$
Inspector General	34	0.0020	$(0.0020)(360°) \approx 0.7°$
	$\sum f = 16{,}793$	$\sum \frac{f}{N} = 1$	



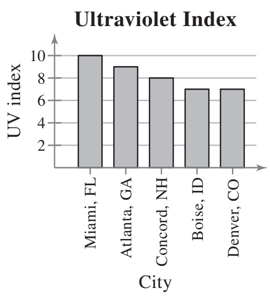
It appears that 63.4% of NASA's budget went to science, aeronautics, and exploration. (Answers will vary.)

25.



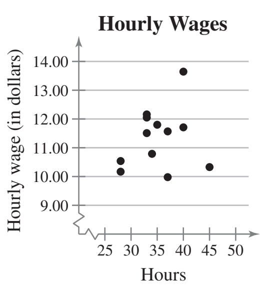
It appears that the biggest reason for baggage delay comes from transfer baggage mishandling. (Answers will vary.)

26.



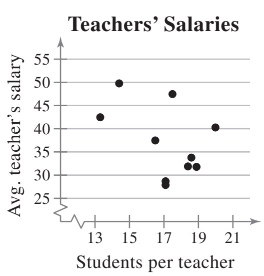
It appears that Boise, ID and Denver, CO have the same UV index. (Answers will vary.)

27.



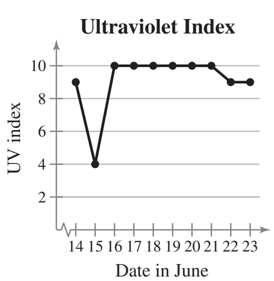
It appears that hourly wage increases as the number of hours worked increases. (Answers will vary.)

28.



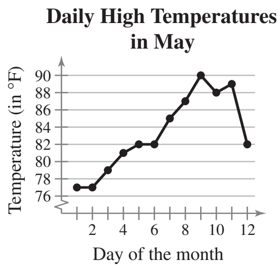
It appears that a teacher's average salary decreases as the number of students per teacher increases. (Answers will vary.)

29.



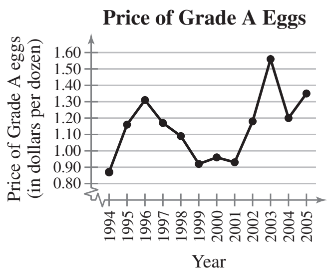
Of the period from June 14–23, 2001, in Memphis, TN, the ultraviolet index was highest from June 16–21. (Answers will vary.)

30.



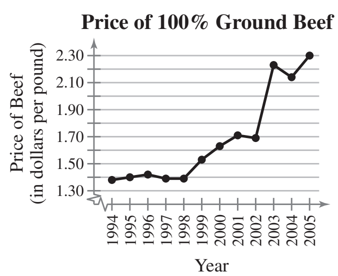
It appears that it was hottest from May 7 to May 11. (Answers will vary.)

31.



It appears the price of eggs peaked in 2003. (Answers will vary.)

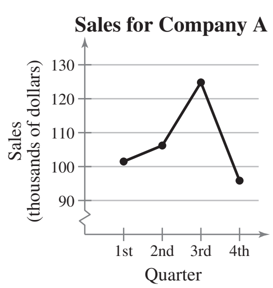
32.



It appears that the greatest increases in dollars per pound was 2002 to 2003. (Answers will vary.)

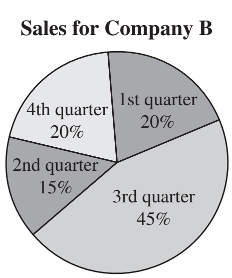
33. (a) When data are taken at regular intervals over a period of time, a time series chart should be used. (Answers will vary.)

(b)



34. (a) The pie chart should be displaying all four quarters, not just the first three.

(b)



35. (a) At law firm A, the lowest salary was \$90,000 and the highest was \$203,000; at law firm B, the lowest salary was \$90,000 and the highest salary was \$190,000.

(b) There are 30 lawyers at law firm A and 32 lawyers at law firm B.

(c) At Law Firm A, the salaries tend to be clustered at the far ends of the distribution range and at Law Firm B, the salaries tend to fall in the middle of the distribution range.

36. (a) At the 3:00 P.M. class, the youngest participant is 35 years old and the oldest participant is 85 years old. At the 8:00 P.M. class, the youngest participant is 18 years old and the oldest participant is 71 years old.

(b) In the 3:00 P.M. class, there are 26 particpants and in the 8:00 P.M. class, there are 30 particpants.

(c) The participants in each class are clustered at one of the ends of their distribution range. The 3:00 P.M. class mostly has particpants over 50 and the 8:00 P.M. class mostly has participants under 50. (Answers will vary.)

2.3 MEASURES OF CENTRAL TENDENCY

2.3 Try It Yourself Solutions

1a. $\Sigma x = 578$

b. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{578}{14} = 41.3$

c. The mean age of an employee in a department is 41.3 years.

2a. 18 18, 19, 19, 19, 20, 21, 21, 21, 21, 23, 24, 24, 26, 27, 27, 29, 30, 30, 30, 33, 33, 34, 35, 38

b. median = middle entry = 24

c. The median age for the sample of fans at the concert is 24.

3a. 70, 80, 100, 130, 140, 150, 160, 200, 250, 270

b. median = mean of two middle entries $\{140, 150\} = 145$

c. The median price of the sample of MP3 players is \$145.

4a. 0, 0, 1, 1, 1, 2, 3, 3, 3, 4, 5, 5, 5, 7, 9, 10, 12, 12, 13, 13, 13, 13, 13, 15, 16, 16, 17, 17, 18, 18, 18, 19, 19, 19, 20, 20, 21, 22, 23, 23, 24, 24, 25, 25, 26, 26, 26, 29, 33, 36, 37, 39, 39, 39, 39, 40, 40, 41, 41, 41, 42, 44, 44, 45, 47, 48, 49, 49, 49, 51, 53, 56, 58, 58, 59, 60, 67, 68, 68, 72

b. The age that occurs with the greatest frequency is 13 years old.

c. The mode of the ages is 13 years old.

5a. "Yes" occurs with the greatest frequency (171).

b. The mode of the responses to the survey is "Yes".

6a. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{410}{19} \approx 21.6$

median = 21

mode = 20

b. The mean in Example 6 ($\bar{x} \approx 23.8$) was heavily influenced by the age 65. Neither the median nor the mode was affected as much by the age 65.

7ab.

Source	Score, x	Weight, w	$x \cdot w$
Test Mean	86	0.50	(83)(0.50) = 43.0
Midterm	96	0.15	(96)(0.15) = 14.4
Final	98	0.20	(98)(0.20) = 19.6
Computer Lab	98	0.10	(98)(0.10) = 9.8
Homework	100	0.05	(100)(0.05) = 5.0
		$\sum w = 1.00$	$\sum(x \cdot w) = 91.8$

c. $\bar{x} = \dfrac{\Sigma(x \cdot w)}{\Sigma w} = \dfrac{91.8}{1.00} = 91.8$

d. The weighted mean for the course is 91.8. So, you did get an A.

8abc.

Class	Midpoint, x	Frequency, f	$x \cdot f$
15–24	19.5	16	312
25–34	29.5	34	1003
35–44	39.5	30	1185
45–54	49.5	23	1138.5
55–64	59.5	13	773.5
65–74	69.5	2	139
75–84	79.5	0	0
85–94	89.5	1	89.5
		$N = 119$	$\sum(x \cdot f) = 4640.5$

d. $\mu = \dfrac{\Sigma(x \cdot f)}{N} = \dfrac{4640.5}{119} \approx 39.0$

The average number of touchdowns is approximately 39.0.

2.3 EXERCISE SOLUTIONS

1. True.

2. False. Not all data sets must have a mode.

3. False. All quantitative data sets have a median.

4. False. The mode is the only measure of central tendency that can be used for data at the nominal level of measurement.

5. False. When each data class has the same frequency, the distribution is uniform.

6. False. When the mean is greater than the median, the distrubution is skewed right.

7. Answers will vary. A data set with an outlier within it would be an example. For instance, the mean of the prices of existing home sales tends to be "inflated" due to the presence of a few very expensive homes.

8. Any data set that is symmetric has the same median and mode.

9. Skewed right because the "tail" of the distribution extends to the right.

10. Symmetric because the left and right halves of the distribution are approximately mirror images.

11. Uniform because the bars are approximately the same height.

12. Skewed left because the tail of the distribution extends to the left.

13. (11), because the distribution values range from 1 to 12 and has (approximately) equal frequencies.

14. (9), because the distribution has values in the thousands of dollars and is skewed right due to the few executives that make a much higher salary than the majority of the employees.

15. (12), because the distribution has a maximum value of 90 and is skewed left due to a few students scoring much lower than the majority of the students.

16. (10), because the distribution is rather symmetric due to the nature of the weights of seventh grade boys.

17. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{81}{13} \approx 6.2$

5 5 5 5 5 5 6 6 7 8 9 9

middle value $\Rightarrow$ median $= 6$

mode $= 5$ (occurs 6 times)

18. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{252}{10} = 25.2$

19 20 21 22 22 23 25 30 35 35

two middle values $\Rightarrow$ median $= \dfrac{22 + 23}{2} = 22.5$

mode $= 22, 35$ (occurs 2 times each)

19. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{32}{7} \approx 4.57$

3.7 4.0 4.8 4.8 4.8 4.8 5.1

middle value $\Rightarrow$ median $= 4.8$

mode $= 4.8$ (occurs 4 times)

20. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{2004}{10} = 200.4$

154 171 173 181 184 188 203 235 240 275

two middle values $\Rightarrow$ median $= \dfrac{184 + 188}{2} = 186$

mode $=$ none

The mode cannot be found because no data points are repeated.

21. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{661.2}{32} \approx 20.66$

10.5, 13.2, 14.9, 16.2, 16.7, 16.9, 17.6, 18.2, 18.6, 18.8, 18.8, 19.1, 19.2, 19.6, 19.8,
$\underbrace{19.9, 20.2}$, 20.7, 20.9, 22.1, 22.1, 22.2, 22.9, 23.2, 23.3, 24.1, 24.9, 25.8, 26.6, 26.7, 26.7, 30.8

two middle values $\Rightarrow$ median $= \dfrac{19.9 + 20.2}{2} = 20.05$

mode $=$ 18.8, 22.1, 26.7 (occurs 2 times each)

22. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{1223}{20} = 61.2$

12 18 26 28 31 33 40 44 45 $\underbrace{49 \quad 61}$ 63 75 80 80 89 96 103 125 125

two middle values $\Rightarrow$ median $= \dfrac{49 + 61}{2} = 55$

mode $=$ 80, 125

The modes do not represent the center of the data set because they are large values compared to the rest of the data.

23. $\bar{x} =$ not possible (nominal data)

median $=$ not possible (nominal data)

mode $=$ "Worse"

The mean and median cannot be found because the data are at the nominal level of measurement.

24. $\bar{x} =$ not possible (nominal data)

median $=$ not possible (nominal data)

mode $=$ "Watchful"

The mean and median cannot be found because the data are at the nominal level of measurement.

25. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{1194.4}{7} \approx 170.63$

155.7, 158.1, 162.2, 169.3, 180, 181.8, 187.3

middle value $\Rightarrow$ median $=$ 169.3

mode $=$ none

The mode cannot be found because no data points are repeated.

26. $\bar{x} =$ not possible (nominal data)

median $=$ not possible (nominal data)

mode $=$ "Domestic"

The mean and median cannot be found because the data are at the nominal level of measurement.

27. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{226}{10} = 22.6$

14, 14, 15, 177, 18, 20, 22, 25, 40, 41

$\underbrace{18,\ 20}$ two middle values $\Rightarrow$ median $= \dfrac{18 + 20}{2} = 19$

mode = 14 (occurs 2 times)

28. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{83}{5} = 16.6$

1, 10, 15, 25.5, 31.5

middle value $\Rightarrow$ median = 15

mode = none

The mode cannot be found because no data points are repeated.

29. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{197.5}{14} \approx 14.11$

1.5, 2.5, 2.5, 5, 10.5, 11, 13, 15.5, 16.5, 17.5, 20, 26.5, 27, 28.5

$\underbrace{13,\ 15.5}$ two middle values $\Rightarrow$ median $= \dfrac{13 + 15.5}{2} = 14.25$

mode = 2.5 (occurs 2 times)

30. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{3455}{11} \approx 314.1$

25, 35, 93, 110, 356, 374, 380, 445, 458, 480, 699

middle value $\Rightarrow$ median = 374

mode = none

The mode cannot be found because no data points are repeated.

31. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{578}{14} = 41.3$

10, 12, 21, 24, 27, 37, 38, 41, 45, 45, 50, 57, 65, 106

$\underbrace{38,\ 41}$ two middle values $\Rightarrow$ median $= \dfrac{38 + 41}{2} = 39.5$

mode = 4.5 (occurs 2 times)

32. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{29.9}{12} \approx 2.49$

0.8, 1.5, 1.6, 1.8, 2.1, 2.3, 2.4, 2.5, 3.0, 3.9, 4.0, 4.0

$\underbrace{2.3,\ 2.4}$ two middle values $\Rightarrow$ median $= \dfrac{2.3 + 2.4}{2} = 2.35$

mode = 4.0 (occurs 2 times)

33. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{292}{15} \approx 19.5$

5, 8, 10, 15, 15, 15, 17, 20, 21, 22, 22, 25, 28, 32, 37

middle value $\Longrightarrow$ median = 20

mode = 15 (occurs 3 times)

34. $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{2987}{14} \approx 213.4$

205, 208, 210, 212, 212, 214, $\underbrace{214, 214,}$ 215, 215, 217, 217, 217, 217

two middle values $\Longrightarrow$ median $= \dfrac{214 + 214}{2} = 214$

mode = 217 (occurs 4 times)

35. A = mode (data entry that occurred most often)

B = median (left of mean in skewed-right dist.)

C = mean (right of median in skewed-right dist.)

36. A = mean (left of median in skewed-left dist.)

B = median (right of mean in skewed-left dist.)

C = mode (data entry that occurred most often)

37. Mode because the data is nominal.

38. Mean because the data are symmetric.

39. Mean because the data does not contain outliers.

40. Median because the data are skewed.

41.

Source	Score, x	Weight, w	$x \cdot w$
Homework	85	0.05	(85)(0.05) = 4.25
Quiz	80	0.35	(80)(0.35) = 28
Project	100	0.20	(100)(0.20) = 20
Speech	90	0.15	(90)(0.15) = 13.5
Final Exam	93	0.25	(93)(0.25) = 23.25
		$\Sigma w = 1$	$\Sigma(x \cdot w) = 89$

$\bar{x} = \dfrac{\Sigma(x \cdot w)}{\Sigma w} = \dfrac{89}{1} = 89$

42.

Source	Score, x	Weight, w	$x \cdot w$
MBAs	\$45,500	8	(45,500)(8) = 364,000
Bas	\$32,000	17	(32,000)(17) = 544,000
		$\Sigma w = 25$	$\Sigma(x \cdot w) = 908{,}000$

$\bar{x} = \dfrac{\Sigma(x \cdot w)}{\Sigma w} = \dfrac{908{,}000}{25} = \$36{,}320$

43.

Balance, x	Days, w	$x \cdot w$
\$523	24	(523)(24) = 12,552
\$2415	2	(2415)(2) = 4830
\$250	4	(250)(4) = 1000
	$\Sigma w = 30$	$\Sigma(x \cdot w) = 18{,}382$

$\bar{x} = \dfrac{\Sigma(x \cdot w)}{\Sigma w} = \dfrac{18{,}382}{30} = \$612.73$

44.

Balance, x	Days, w	$x \cdot w$
\$759	15	$(759)(15) = 11{,}385$
\$1985	5	$(1985)(5) = 9925$
\$1410	5	$(1410)(5) = 7050$
\$348	6	$(348)(6) = 2080$
	$\Sigma w = 31$	$\Sigma(x \cdot w) = 30{,}448$

$$\bar{x} = \frac{\Sigma(x \cdot w)}{\Sigma w} = \frac{30{,}448}{31} = \$982.19$$

45.

Grade	Points, x	Credits, w	$x \cdot w$
B	3	3	$(3)(3) = 9$
B	3	3	$(3)(3) = 9$
A	4	4	$(4)(4) = 16$
D	1	2	$(1)(2) = 2$
C	2	3	$(2)(3) = 6$
		$\Sigma w = 15$	$\Sigma(x \cdot w) = 42$

$$\bar{x} = \frac{\Sigma(x \cdot w)}{\Sigma w} = \frac{42}{15} = 2.8$$

46.

Source	Score, x	Weight, w	$x \cdot w$
Engineering	85	9	$(85)(9) = 765$
Business	81	13	$(81)(13) = 1053$
Math	90	5	$(90)(5) = 450$
		$\Sigma(x \cdot w) = 27$	$\Sigma w = 2268$

$$\bar{x} = \frac{\Sigma(x \cdot w)}{\Sigma w} = \frac{2268}{27} = 84$$

47.

Midpoint, x	Frequency, f	$x \cdot f$
61	4	$(61)(4) = 244$
64	5	$(64)(5) = 320$
67	8	$(67)(8) = 536$
70	1	$(70)(1) = 70$
	$n = 18$	$\Sigma(x \cdot f) = 1170$

$$\bar{x} = \frac{\Sigma(x \cdot f)}{n} = \frac{1170}{18} \approx 65 \text{ inches}$$

48.

Midpoint, x	Frequency, f	$x \cdot f$
64	3	$(64)(3) = 192$
67	6	$(67)(6) = 402$
70	7	$(70)(7) = 490$
73	4	$(73)(4) = 292$
76	3	$(76)(3) = 228$
	$\Sigma(x \cdot f) = 23$	$\Sigma(x \cdot f) = 1604$

$$\bar{x} = \frac{\Sigma(x \cdot f)}{n} = \frac{1604}{23} \approx 69.7 \text{ inches}$$

49.

Midpoint, x	Frequency, f	$x \cdot f$
4.5	55	$(4.5)(55) = 247.5$
14.5	70	$(14.5)(70) = 1015$
24.5	35	$(24.5)(35) = 857.5$
34.5	56	$(34.5)(56) = 1932$
44.5	74	$(44.5)(74) = 3293$
54.5	42	$(54.5)(42) = 2289$
64.5	38	$(64.5)(38) = 2451$
74.5	17	$(74.5)(17) = 1266.5$
84.5	10	$(84.5)(10) = 845$
	$n = 397$	$\Sigma(x \cdot f) = 14{,}196.5$

$$\bar{x} = \frac{\Sigma(x \cdot f)}{n} = \frac{14{,}196.5}{397} \approx 35.8 \text{ years old}$$

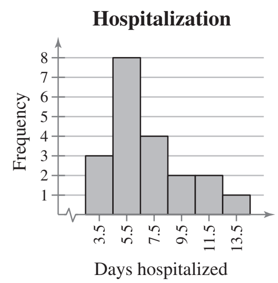
50.

Midpoint, x	Frequency, f	$x \cdot f$
3	12	$(3)(12) = 36$
8	26	$(8)(26) = 208$
13	20	$(13)(20) = 260$
18	7	$(18)(7) = 126$
23	11	$(23)(11) = 253$
28	7	$(28)(7) = 196$
33	4	$(33)(4) = 132$
38	4	$(38)(4) = 152$
43	1	$(43)(1) = 43$
	$n = 92$	$\Sigma(x \cdot f) = 1406$

$$\bar{x} = \frac{\Sigma(x \cdot f)}{n} = \frac{1406}{92} \approx 15.3 \text{ minutes}$$

51. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{14 - 3}{6} = 1.83 \Rightarrow 2$

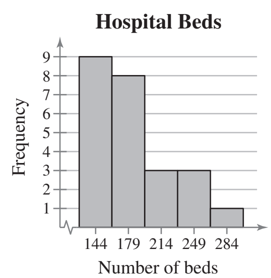
Class	Midpoint, x	Frequency, f
3–4	3.5	3
5–6	5.5	8
7–8	7.5	4
9–10	9.5	2
11–12	11.5	2
13–14	13.5	1
		$\Sigma f = 20$



Shape: Positively skewed

52. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{297 - 127}{5} = 34$

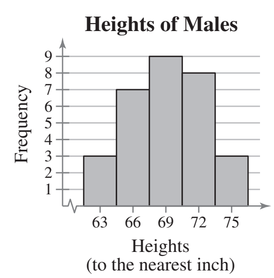
Class	Midpoint, x	Frequency, f
127–161	144	9
162–196	179	8
197–231	214	3
232–266	249	3
267–301	284	1
		$\Sigma f = 24$



Shape: Positively skewed

53. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{76 - 62}{5} = 2.8 \Rightarrow 3$

Class	Midpoint, x	Frequency, f
62–64	63	3
65–67	66	7
68–70	69	9
71–73	72	8
74–76	75	3
		$\Sigma f = 30$

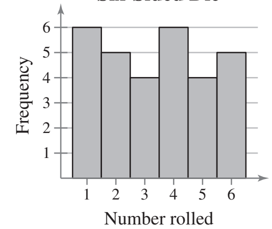


Shape: Symmetric

54. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{6 - 1}{6} = 0.8333 \Rightarrow 1$

Class	Frequency, f
1	6
2	5
3	4
4	6
5	4
6	5
	$\Sigma f = 30$

Results of Rolling Six-Sided Die



Shape: Uniform

55. (a) $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{36.03}{6} = 6.005$

$5.59,\ 5.99,\ \underbrace{6,\ 6.02,}\ 6.03,\ 6.4$

two middle values $\Rightarrow$ median $= \dfrac{6 + 6.02}{2} = 6.01$

(b) $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{35.67}{6} = 5.945$

$5.59,\ 5.99,\ \underbrace{6,\ 6.02,}\ 6.03,\ 6.4$

two middle values $\Rightarrow$ median $= \dfrac{6 + 6.02}{2} = 6.01$

(c) mean

56. (a) $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{815.4}{19} = 42.92$

7.8, 8.2, 12.6, 12.6, 14.4, 17.8, 19.2, 21.3, 23, 24.2, 24.7, 31.1, 32.5, 41.3, 45.4, 55.2,
59.6, 134.2, 230.3

middle value $\Rightarrow$ median $= 24.2$

(b) $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{585.1}{18} \approx 32.51$

$7.8,\ 8.2,\ 12.6,\ 12.6,\ 14.4,\ 17.8,\ 19.2,\ 21.3,\ \underbrace{23,\ 24.2,}\ 24.7,\ 31.1,\ 32.5,\ 41.3,\ 45.4,$
$55.5,\ 59.6,\ 134.2$

two middle values $\Rightarrow$

median $= \dfrac{23 + 24.2}{2} = 23.6$

(c) mean

57. (a) $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{3222}{9} = 358$

147, 177, 336, 360, 375, 393, 408, 504, 522

middle value $\Rightarrow$ median $= 375$

(b) $\bar{x} = \dfrac{\Sigma x}{n} = \dfrac{9666}{9} = 1074$

441, 531, 1008, 1080, 1125, 1179, 1224, 1512, 1566

middle value $\Rightarrow$ median = 1125

(c) $\bar{x}$ from part (b) is 3 times $\bar{x}$ from part (a). Median from part (b) is 3 times median from part (a).

(d) Multiply part (b) answers by 12.

58. (a) Mean should be used because Car A has the highest mean of the three.

(b) Median should be used because Car B has the highest median of the three.

(c) Mode should be used because Car C has the highest mode of the three.

59. Car A because it has the highest midrange of the three.

Car A: Midrange $= \dfrac{34 + 28}{2} = 32$

Car B: Midrange $= \dfrac{31 + 29}{2} = 30$

Car C: Midrange $= \dfrac{32 + 28}{2} = 30$

60. (a) $\bar{x} = 49.2$, median = 46.5

(b) Key: $3|6 = 36$ (c) Positively skewed

1	1 3	
2	2 8	
3	6 6 6 7 7 7 8	median
4	1 3 4 6 7	
5	1 1 1 3	mean
6	1 2 3 4	
7	2 2 4 6	
8	5	
9	0	

61. (a) Order the data values.

11	13	22	28	36	36	36	37	37	37	38	41	43	44	46
47	51	51	51	53	61	62	63	64	72	72	74	76	85	90

Delete the lowest 10%, smallest 3 observations (11, 13, 22).

Delete the highest 10%, largest 3 observations (76, 85, 90).

Find the 10% trimmed mean using the remaining 24 observations.

10% trimmed mean = 49.2

(b) $\bar{x} = 49.2$

median $= 46.5$

mode $= 36,\ 37,\ 51$

midrange $= 50.5$

(c) Using a trimmed mean eliminates potential outliers that may affect the mean of all the observations.

2.4 MEASURES OF VARIATION

2.4 Try It Yourself Solutions

1a. Min $= 23$ or \$23,000 and Max $= 58$ or \$58,000

b. Range $=$ max $-$ min $= 58 - 23 = 35$ or \$35,000

c. The range of the starting salaries for Corporation B is 35 or \$35,000 (much larger than range of Corporation A).

2a. $\mu = \dfrac{\Sigma x}{N} = \dfrac{415}{10} = 41.5$ or \$41,500

b.

Salary, x (1000s of dollars)	Deviation, $x - \mu$ (1000s of dollars)
23	$23 - 41.5 = -18.5$
29	$29 - 41.5 = -12.5$
32	$32 - 41.5 = -9.5$
40	$40 - 41.5 = -1.5$
41	$41 - 41.5 = -0.5$
41	$41 - 41.5 = -0.5$
49	$49 - 41.5 = 7.5$
50	$50 - 41.5 = 8.5$
52	$52 - 41.5 = 10.5$
58	$58 - 41.5 = 16.5$
$\sum x = 415$	$\sum(x - \mu) = 0$

3ab. $\mu = 41.5$ or \$41,500

Salary, x	$x - \mu$	$(x - \mu)^2$
23	-18.5	$(-18.5)^2 = 342.25$
29	-12.5	$(-12.5)^2 = 156.25$
32	-9.5	$(-9.5)^2 = 90.25$
40	-1.5	$(-1.5)^2 = 2.25$
41	-0.5	$(-0.5)^2 = 0.25$
41	-0.5	$(-0.5)^2 = 0.25$
49	7.5	$(7.5)^2 = 56.25$
50	8.5	$(8.5)^2 = 72.25$
52	10.5	$(10.5)^2 = 110.25$
58	16.5	$(16.5)^2 = 272.25$
$\sum x = 415$	$\sum(x - \mu) = 0$	$\sum(x - \mu)^2 = 1102.5$

c. $\sigma^2 = \dfrac{\Sigma(x-\mu)^2}{N} = \dfrac{1102.5}{10} \approx 110.25$

d. $\sigma = \sqrt{\sigma^2} = \sqrt{\dfrac{1102.5}{10}} = 10.5$ or \$10,500

e. The population variance is 110.3 and the population standard deviation is 10.5 or \$10,500.

4a.

Salary, x	$x - \bar{x}$	$(x - \bar{x})^2$
23	−18.5	342.25
29	−12.5	156.25
32	−9.5	90.25
40	−1.5	2.25
41	−0.5	0.25
41	−0.5	0.25
49	7.5	56.25
50	8.5	72.25
52	10.5	110.25
58	16.5	272.25
$\sum x = 415$	$\sum(x - \bar{x}) = 0$	$\sum(x - \bar{x})^2 = 1102.5$

$SS_x = \Sigma(x - \bar{x})^2 = 1102.5$

b. $s^2 = \dfrac{\Sigma(x-\bar{x})^2}{(n-1)} = \dfrac{1102.5}{9} = 122.5$

c. $s = \sqrt{s^2} = \sqrt{122.5} \approx 11.1$ or \$11,100

5a. (Enter data in computer or calculator)

b. $\bar{x} = 37.89, \qquad s = 3.98$

6a. 7, 7, 7, 7, 7, 13, 13, 13, 13, 13

b.

x	$x - \mu$	$(x - \mu)^2$
7	$7 - 10 = -3$	$(-3)^2 = 9$
7	$7 - 10 = -3$	$(-3)^2 = 9$
7	$7 - 10 = -3$	$(-3)^2 = 9$
7	$7 - 10 = -3$	$(-3)^2 = 9$
7	$7 - 10 = -3$	$(-3)^2 = 9$
13	$13 - 10 = 3$	$(3)^2 = 9$
13	$13 - 10 = 3$	$(3)^2 = 9$
13	$13 - 10 = 3$	$(3)^2 = 9$
13	$13 - 10 = 3$	$(3)^2 = 9$
13	$13 - 10 = 3$	$(3)^2 = 9$
$\Sigma x = 100$	$\sum(x - \mu) = 0$	$\sum(x - \mu)^2 = 90$

$\mu = \dfrac{\Sigma x}{N} = \dfrac{100}{10} = 10$

$\sigma = \sqrt{\dfrac{\Sigma(x-\mu)^2}{N}} = \sqrt{\dfrac{90}{10}} = \sqrt{9} = 3$

7a. $64 - 61.25 = 2.75 = 1$ standard deviation

b. 34%

c. The estimated percent of the heights that are between 61.25 and 64 inches is 34%.

8a. $31.6 - 2(19.5) = -7.4 = \ > 0$

b. $31.6 + 2(19.5) = 70.6$

c. $1 - \dfrac{1}{k^2} = 1 - \dfrac{1}{(2)^2} = 1 - \dfrac{1}{4} = 0.75$

At least 75% of the data lie within 2 standard deviations of the mean. At least 75% of the population of Alaska is between 0 and 70.6 years old.

9a.

x	f	xf
0	10	$(0)(10) = 0$
1	19	$(1)(19) = 19$
2	7	$(2)(7) = 14$
3	7	$(3)(7) = 21$
4	5	$(4)(5) = 20$
5	1	$(5)(1) = 5$
6	1	$(6)(1) = 6$
	$n = 50$	$\sum xf = 85$

b. $\bar{x} = \dfrac{\Sigma xf}{n} = \dfrac{85}{50} = 1.7$

c.

$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 \cdot f$
$0 - 1.7 = -1.70$	$(-1.70)^2 = 2.8900$	$(2.8900)(10) = 28.90$
$1 - 1.7 = -0.70$	$(-0.70)^2 = 0.4900$	$(0.4900)(19) = 9.31$
$2 - 1.7 = 0.30$	$(0.30)^2 = 0.0900$	$(0.0900)(7) = 0.63$
$3 - 1.7 = 1.30$	$(1.30)^2 = 1.6900$	$(1.6900)(7) = 11.83$
$4 - 1.7 = 2.30$	$(2.30)^2 = 5.2900$	$(5.2900)(5) = 26.45$
$5 - 1.7 = 3.30$	$(3.30)^2 = 10.9800$	$(10.9800)(1) = 10.89$
$6 - 1.7 = 4.30$	$(4.30)^2 = 18.4900$	$(18.4900)(1) = 18.49$
		$\sum (x - \bar{x})^2 f = 106.5$

d. $s = \sqrt{\dfrac{\Sigma(x - \bar{x})^2 f}{(n - 1)}} = \sqrt{\dfrac{106.5}{49}} = \sqrt{2.17} \approx 1.5$

10a.

Class	x	f	xf
0–99	49.5	380	$(49.5)(380) = 18{,}810$
100–199	149.5	230	$(149.5)(230) = 34{,}385$
200–299	249.5	210	$(249.5)(210) = 52{,}395$
300–399	349.5	50	$(349.5)(50) = 17{,}475$
400–499	449.5	60	$(449.5)(60) = 26{,}970$
500+	650.0	70	$(650.0)(70) = 45{,}500$
		$n = 1000$	$\sum xf = 195{,}535$

b. $\bar{x} = \dfrac{\Sigma xf}{n} = \dfrac{195{,}535}{1000} \approx 195.5$

c.

$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 \cdot f$
$49.5 - 195.5 = -146$	$(-146)^2 = 21{,}316$	$(21{,}316)(380) = 8{,}100{,}080$
$149.5 - 195.5 = -46$	$(-46)^2 = 2116$	$(2116)(230) = 486{,}680$
$249.5 - 195.5 = 54$	$(54)^2 = 2916$	$(2916)(210) = 612{,}360$
$349.5 - 195.5 = 154$	$(154)^2 = 23{,}716$	$(23{,}716)(50) = 1{,}185{,}800$
$449.5 - 195.5 = 254$	$(254)^2 = 64{,}516$	$(64{,}516)(60) = 3{,}870{,}960$
$650 - 195.5 = 454.5$	$(454.5)^2 = 206{,}570.25$	$(206{,}570.25)(70) = 14{,}459{,}917.5$
		$\sum(x - \bar{x})^2 f = 28{,}715{,}797.5$

d. $s = \sqrt{\dfrac{\Sigma(x - \bar{x})^2 f}{n - 1}} = \sqrt{\dfrac{28{,}715{,}797.5}{999}} = \sqrt{28{,}744.542} \approx 169.5$

2.4 EXERCISE SOLUTIONS

1. Range = Max − Min = 12 − 4 = 8

$$\mu = \frac{\Sigma x}{N} = \frac{79}{10} = 7.9$$

x	$x - \mu$	$(x - \mu)^2$
12	$12 - 7.9 = 4.1$	$(4.1)^2 = 16.81$
9	$9 - 7.9 = 1.1$	$(1.1)^2 = 1.21$
7	$7 - 7.9 = -0.9$	$(-0.9)^2 = 0.81$
5	$5 - 7.9 = -2.9$	$(-2.9)^2 = 8.41$
7	$7 - 7.9 = -0.9$	$(-0.9)^2 = 0.81$
8	$8 - 7.9 = 0.1$	$(0.1)^2 = 0.01$
10	$10 - 7.9 = 2.1$	$(2.1)^2 = 4.41$
4	$4 - 7.9 = -3.9$	$(-3.9)^2 = 15.21$
11	$11 - 7.9 = 3.1$	$(3.1)^2 = 9.61$
6	$6 - 7.9 = -1.9$	$(-1.9)^2 = 3.61$
$\Sigma x = 79$	$\Sigma(x - \mu) = 0$	$\Sigma(x - \mu)^2 = 60.9$

$$\sigma^2 = \frac{\Sigma(x - \mu)^2}{N} = \frac{60.9}{10} = 6.09 \approx 6.1$$

$$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}} = \sqrt{6.09} \approx 2.5$$

2. Range = Max − Min = 24 − 14 = 10

$$\mu = \frac{\Sigma x}{N} = \frac{264}{14} = 18.9$$

x	$x - \mu$	$(x - \mu)^2$
15	$15 - 18.9 = -3.9$	$(-3.9)^2 = 15.21$
24	$24 - 18.9 = 5.1$	$(5.1)^2 = 26.01$
17	$17 - 18.9 = -1.9$	$(-1.9)^2 = 3.61$
19	$19 - 18.9 = 0.1$	$(0.1)^2 = 0.01$
20	$20 - 18.9 = 1.1$	$(1.1)^2 = 1.21$
18	$18 - 18.9 = -0.9$	$(-0.9)^2 = 0.81$
20	$20 - 18.9 = 1.1$	$(1.1)^2 = 1.21$
16	$16 - 18.9 = -2.9$	$(-2.9)^2 = 8.41$
21	$21 - 18.9 = 2.1$	$(2.1)^2 = 4.41$
23	$23 - 18.9 = 4.1$	$(4.1)^2 = 16.81$
17	$17 - 18.9 = -1.9$	$(-1.9)^2 = 3.61$
18	$18 - 18.9 = -0.9$	$(-0.9)^2 = 0.81$
22	$22 - 18.9 = 3.1$	$(3.1)2 = 9.61$
14	$14 - 18.9 = -4.9$	$(-4.9)^2 = 24.01$
$\Sigma x = 264$	$\Sigma(x - \mu) = 0$	$\Sigma(x - \mu)^2 = 115.74$

$$\sigma^2 = \frac{\Sigma(x - \mu)^2}{N} = \frac{115.74}{14} = 8.27 \approx 8.3$$

$$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}} = \sqrt{8.27} \approx 2.9$$

3. Range = Max − Min = 18 − 6 = 12

$$\bar{x} = \frac{\Sigma x}{n} = \frac{107}{9} = 11.9$$

x	$(x - \bar{x})$	$(x - \bar{x})^2$
17	$17 - 11.9 = 5.1$	$(5.1)^2 = 26.01$
8	$8 - 11.9 = -3.9$	$(-3.9)^2 = 15.21$
13	$13 - 11.9 = 1.1$	$(1.1)^2 = 1.21$
18	$18 - 11.9 = 6.1$	$(6.1)^2 = 37.21$
15	$15 - 11.9 = 3.1$	$(3.1)^2 = 9.61$
9	$9 - 11.9 = -2.9$	$(-2.9)^2 = 8.41$
10	$10 - 11.9 = -1.9$	$(-1.9)^2 = 3.61$
11	$11 - 11.9 = -0.9$	$(-0.9)^2 = 0.81$
6	$6 - 11.9 = -5.9$	$(-5.9)^2 = 34.81$
$\Sigma x = 107$	$\Sigma(x - \bar{x}) = 0$	$\Sigma(x - \bar{x})^2 = 136.89$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} = \frac{136.89}{9 - 1} = 17.1$$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}} = \sqrt{17.1} \approx 4.1$$

4. Range = Max − Min = 28 − 7 = 21

$$\bar{x} = \frac{\Sigma x}{n} = \frac{238}{13} = 18.3$$

x	$(x - \bar{x})$	$(x - \bar{x})^2$
28	$28 - 18.3 = 9.7$	$(9.7)^2 = 94.09$
25	$25 - 18.3 = 6.7$	$(6.7)^2 = 44.89$
21	$21 - 18.3 = 2.7$	$(2.7)^2 = 7.29$
15	$15 - 18.3 = -3.3$	$(-3.3)^2 = 10.89$
7	$7 - 18.3 = -11.3$	$(-11.3)^2 = 127.69$
14	$14 - 18.3 = -4.3$	$(-4.3)^2 = 18.49$
9	$9 - 18.3 = -9.3$	$(-9.3)^2 = 86.49$
27	$27 - 18.3 = 8.7$	$(8.7)^2 = 75.69$
21	$21 - 18.3 = 2.7$	$(2.7)^2 = 7.29$
24	$24 - 18.3 = 5.7$	$(5.7)^2 = 32.49$
14	$14 - 18.3 = -4.3$	$(-4.3)^2 = 18.49$
17	$17 - 18.3 = -1.3$	$(-1.3)^2 = 1.69$
16	$16 - 18.3 = -2.3$	$(-2.3)^2 = 5.29$
$\Sigma x = 238$	$\Sigma(x - \bar{x}) = 0$	$\Sigma(x - \bar{x})^2 = 530.77$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} = \frac{530.77}{13 - 1} = 44.23 \approx 44.2$$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}} = \sqrt{44.23} \approx 6.7$$

5. Range = Max − Min = 96 − 23 = 73

6. Range = Max − Min = 34 − 24 = 10

7. The range is the difference between the maximum and minimum values of a data set. The advantage of the range is that it is easy to calculate. The disadvantage is that it uses only two entries from the data set.

8. The deviation, $(x - \mu)$, is the difference between an observation, x, and the mean of the data, μ. The sum of the deviations is always zero.

9. The units of variance are squared. Its units are meaningless. (Ex: dollars^2)

10. The standard deviation is the positive square root of the variance.

Because squared deviations can never be negative, the standard deviation and variance can never be negative.

$\{7, 7, 7, 7, 7\} \rightarrow n = 5$

$\bar{x} = 7$

$s = 0$

11. (a) Range $=$ Max $-$ Min $= 45.6 - 21.3 = 24.3$

(b) Range $=$ Max $-$ Min $= 65.6 - 21.3 = 44.3$

(c) The range has increased substantially.

12. $\{3, 3, 3, 7, 7, 7\} \rightarrow n = 6$

$\mu = 5$

$s = 2$

13. Graph (a) has a standard deviation of 24 and graph (b) has a standard deviation of 16 because graph (a) has more variability.

14. Graph (a) has a standard deviation of 2.4 and graph (b) has a standard deviation of 5. Graph (b) has more variability.

15. When calculating the population standard deviation, you divide the sum of the squared deviations by N, then take the square root of that value. When calculating the sample standard deviation, you divide the sum of the squared deviations by $n - 1$, then take the square root of that value.

16. When given a data set, one would have to determine if it represented the population or was a sample taken from a population. If the data are a population, then σ is calculated. If the data are a sample, then s is calculated.

17. Company B. Due to the larger standard deviation in salaries for company B, it would be more likely to be offered a salary of $33,000.

18. Player B. Due to the smaller standard deviation in number of strokes, player B would be the more consistent player.

19. (a) Los Angeles: range = Max − Min = 35.9 − 18.3 = 17.6

x	$(x - \bar{x})$	$(x - \bar{x})^2$
20.2	−6.06	36.67
26.1	−0.16	0.02
20.9	−5.36	28.68
32.1	5.84	34.16
35.9	9.64	93.02
23.0	−3.64	10.60
28.2	1.94	3.78
31.6	5.34	28.56
18.3	−7.96	63.29
$\Sigma x = 236.3$		$\Sigma(x - \bar{x})^2 = 298.78$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{236.3}{9} = 26.26$$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{(n - 1)} = \frac{298.78}{8} \approx 37.35$$

$$s = \sqrt{s^2} \approx 6.11$$

Long Beach: range = Max − Min = 26.9 − 18.2 = 8.7

x	$(x - \bar{x})$	$(x - \bar{x})^2$
20.9	−1.98	3.91
18.2	−4.68	21.88
20.8	−2.08	4.32
21.1	−1.78	3.16
26.5	3.62	13.12
26.9	4.02	16.18
24.2	1.32	1.75
25.1	2.22	4.94
22.2	−0.68	0.46
$\Sigma x = 205.9$		$\Sigma(x - \bar{x})^2 = 69.72$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{205.9}{9} = 22.88$$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{(n - 1)} = \frac{69.72}{8} \approx 8.71$$

$$s = \sqrt{s^2} \approx 2.95$$

(b) It appears from the data that the annual salaries in Los Angeles are more variable than the salaries in Long Beach.

20. (a) Dallas: range = Max − Min = 34.9 − 16.8 = 18.1

x	$x - \bar{x}$	$(x - \bar{x})^2$
34.9	8.92	79.61
25.7	−0.28	0.08
17.3	−8.68	75.30
16.8	−9.18	84.23
26.8	0.82	0.68
24.7	−1.28	1.63
29.4	3.42	11.71
32.7	6.72	45.19
25.5	−0.48	0.23
$\Sigma x = 233.8$		$\Sigma(x - \bar{x})^2 = 298.66$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{233.8}{9} = 25.98$$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{(n - 1)} = \frac{298.66}{8} \approx 37.33$$

$$s = \sqrt{s^2} \approx 6.11$$

Houston: range = Max − Min = 31.3 − 18.3 = 13

x	$x - \bar{x}$	$(x - \bar{x})^2$
25.6	−0.03	0.00
23.2	−2.43	5.92
26.7	1.07	1.14
27.7	2.07	4.27
25.4	−0.23	0.05
26.4	0.77	0.59
18.3	−7.33	53.78
26.1	0.47	0.22
31.3	5.67	32.11
$\Sigma x = 230.7$		$\Sigma(x - \bar{x})^2 = 98.08$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{230.7}{9} = 25.63$$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{(n - 1)} = \frac{98.08}{8} = 12.26$$

$$s = \sqrt{s^2} \approx 3.50$$

(b) It appears from the data that the annual salaries in Dallas are more variable than the salaries in Houston.

21. (a) Male: range = Max − Min = 1328 − 923 = 405

x	$(x - \bar{x})$	$(x - \bar{x})^2$
1059	−51.13	2,613.77
1328	217.8	47,469.52
1175	64.88	4,208.77
1123	12.88	165.77
923	−187.13	35,015.77
1017	−93.13	8,672.77
1214	103.88	10,790.02
1042	−68.13	4,641.02
$\Sigma x = 8881$		$\Sigma(x - \bar{x})^2 = 113{,}576.88$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{8881}{8} = 1110.13$$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{(n - 1)} = \frac{113{,}576.9}{7} \approx 16{,}225.3$$

$$s = \sqrt{s^2} \approx 127.4$$

Female: range = Max − Min = 1393 − 841 = 552

x	$(x - \bar{x})$	$(x - \bar{x})^2$
1226	92.50	8,556.25
965	−168.50	28,392.25
841	−292.50	85,556.25
1053	−80.50	6,480.25
1056	−77.50	6,006.25
1393	259.50	67,340.25
1312	178.50	31,862.25
1222	88.50	7,832.25
$\Sigma x = 9068$		$\Sigma(x - \bar{x})^2 = 242{,}026.00$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{9068}{8} = 1133.50$$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{(n - 1)} = \frac{242{,}026}{7} \approx 34{,}575.1$$

$$s = \sqrt{s^2} \approx 185.9$$

(b) It appears from the data, the SAT scores for females are more variable than the SAT scores for males.

22. (a) Public: range = Max − Min = 39.9 − 34.8 = 5.1

x	$(x - \bar{x})$	$(x - \bar{x})^2$
38.6	1.23	1.50
38.1	0.73	0.53
38.7	1.33	1.76
36.8	−0.58	0.33
34.8	−2.58	6.63
35.9	−1.48	2.18
39.9	2.53	6.38
36.2	−1.18	1.38
$\Sigma x = 299$		$\Sigma(x - \bar{x})^2 = 20.68$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{299}{8} = 37.38$$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{(n-1)} = \frac{20.68}{7} \approx 2.95$$

$$s = \sqrt{s^2} \approx 1.72$$

Private: range = Max − Min = 21.8 − 17.6 = 4.2

x	$(x - \bar{x})$	$(x - \bar{x})^2$
21.8	2.26	5.12
18.4	−1.14	1.29
20.3	0.76	0.58
17.6	−1.94	3.75
19.7	0.16	0.03
18.3	−1.24	1.53
19.4	−0.14	0.02
20.8	1.26	1.59
$\Sigma x = 156.3$		$\Sigma(x - \bar{x})^2 = 13.92$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{156.3}{8} = 19.54$$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{(n-1)} = \frac{13.92}{7} \approx 1.99$$

$$s = \sqrt{s^2} \approx 1.41$$

(b) It appears from the data that the annual salaries for public teachers are more variable than the salaries for private teachers.

23. (a) Greatest sample standard deviation: (ii)

Data set (ii) has more entries that are farther away from the mean.

Least sample standard deviation: (iii)

Data set (iii) has more entries that are close to the mean.

(b) The three data sets have the same mean but have different standard deviations.

24. (a) Greatest sample standard deviation: (i)

Data set (i) has more entries that are farther away from the mean.

Least sample standard deviation: (iii)

Data set (iii) has more entries that are close to the mean.

(b) The three data sets have the same mean, median, and mode, but have different standard deviations.

25. (a) Greatest sample standard deviation: (ii)

Data set (ii) has more entries that are farther away from the mean.

Least sample standard deviation: (iii)

Data set (iii) has more entries that are close to the mean.

(b) The three data sets have the same mean, median, and mode, but have different standard deviations.

26. (a) Greatest sample standard deviation: (iii)

Data set (iii) has more entries that are farther away from the mean.

Least sample standard deviation: (i)

Data set (i) has more entries that are close to the mean.

(b) The three data sets have the same mean and median but have different standard deviations.

27. Similarity: Both estimate proportions of the data contained within k standard deviations of the mean.

Difference: The Empirical Rule assumes the distribution is bell-shaped, Chebychev's Theorem makes no such assumption.

28. You must know the distribution is bell-shaped.

29. $(1300, 1700) \rightarrow (1500 - 1(200), 1500 + 1(200)) \rightarrow (\bar{x} - s,\ \bar{x} + s)$

68% of the farms value between \$1300 and \$1700 per acre.

30. 95% of the data falls between $\bar{x} - 2s$ and $\bar{x} + 2s$.

$\bar{x} - 2s = 2400 - 2(450) = 1500$

$\bar{x} + 2s = 2400 + 2(450) = 3300$

95% of the farm values lie between \$1500 and \$3300 per acre.

31. (a) $n = 75$

$68\%(75) = (0.68)(75) = 51$ farm values will be between \$1300 and \$1700 per acre.

(b) $n = 25$

$68\%(25) = (0.68)(25) = 17$ of these farm values will be between \$1300 and \$1700 per acre.

32. (a) $n = 40$

95% of the data lie within 2 standard deviations of the mean.

$(95\%)(40) = (0.95)(40) = 38$ farm values lie between \$1500 and \$3300 per acre.

(b) $n = 60$

$(95\%)(20) = (0.95)(20) = 19$ of these farm values lie between \$1500 and \$3300 per acre.

33. $\bar{x} = 1500$ {1000, 2000} are outliers. They are more than 2 standard deviations from the mean (1100, 1900).

$s = 200$

34. $\bar{x} = 2400$ {3325, 1490} are outliers. They are more than 2 standard deviations from the mean (1500, 3300).

$s = 450$

35. $(\bar{x} - 2s, \bar{x} + 2s) \rightarrow (1.14, 5.5)$ are 2 standard deviations from the mean.

$1 - \frac{1}{k^2} = 1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = 0.75 \Longrightarrow$ At least 75% of the eruption times lie between 1.14 and 5.5 minutes.

If $n = 32$, at least $(0.75)(32) = 24$ eruptions will lie between 1.14 and 5.5 minutes.

36. $1 - \frac{1}{k^2} = 1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = .75 \rightarrow$ At least 75% of the 400-meter dash times lie within 2 standard deviations of mean.

$(\bar{x} - 2s, \bar{x} + 2s) \rightarrow (54.97, 59.17) \rightarrow$ At least 75% of the 400-meter dash times lie between 54.97 and 59.17 seconds.

37.

x	f	xf	$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
0	5	0	−2.08	4.31	21.53
1	11	11	−1.08	1.16	12.71
2	7	14	−0.08	0.01	0.04
3	10	30	0.93	0.86	8.56
4	7	28	1.93	3.71	25.94
	$n = 40$	$\sum xf = 83$			$\sum(x - \bar{x})^2 f = 68.78$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{83}{40} \approx 2.1$$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2 f}{n - 1}} = \sqrt{\frac{68.78}{39}} = \sqrt{1.76} \approx 1.3$$

38.

x	f	xf	$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
0	3	0	−1.74	3.03	9.08
1	15	15	−0.74	0.55	8.21
2	24	48	0.26	0.07	1.62
3	8	24	1.26	1.59	12.70
	$n = 50$	$\sum xf = 87$			$\sum(x - \bar{x})^2 f = 31.62$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{87}{50} \approx 1.7$$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2 f}{n - 1}} \approx \sqrt{\frac{31.62}{49}} \approx \sqrt{0.645} \approx 0.8$$

39. Class width $= \dfrac{\text{Max} - \text{Min}}{5} = \dfrac{14 - 2}{5} = \dfrac{12}{5} = 2.4 \Rightarrow 3$

Class	Midpoint, x	f	xf
2–4	3	4	12
5–7	6	8	48
8–10	9	15	135
11–13	12	4	48
14–16	15	1	15
		$N = 32$	$\sum xf = 258$

$$\mu = \frac{\Sigma xf}{N} = \frac{258}{32} \approx 8.1$$

$x - \mu$	$(x - \mu)^2$	$(x - \mu)^2 f$
-5.1	26.01	104.04
-2.1	4.41	35.28
0.9	0.81	12.15
3.9	15.21	50.84
6.9	47.61	47.61
		$\sum(x - \mu)^2 f = 249.92$

$$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}} = \sqrt{\frac{249.92}{32}} \approx 2.8$$

40. Class width $= \dfrac{\text{Max} - \text{Min}}{5} = \dfrac{244 - 145}{5} = 19.8 \Rightarrow 20$

Class	Midpoint, x	f	xf
145–164	154.5	8	1236.0
165–184	174.5	7	1221.5
185–204	194.5	3	583.5
205–224	214.5	1	214.5
225–244	234.5	1	234.5
		$N = 20$	$\Sigma xf = 3490.0$

$$\mu = \frac{\Sigma xf}{N} = \frac{3490}{20} = 174.5$$

$x - \mu$	$(x - \mu)^2$	$(x - \mu)^2 f$
-20	400	3200
0	0	0
20	400	1200
40	1600	1600
60	3600	3600
		$\Sigma(x - \mu)^2 f = 9600$

$$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2 f}{N}} = \sqrt{\frac{9600}{20}} = \sqrt{480} \approx 21.9$$

41.

Midpoint, x	f	xf
70.5	1	70.5
92.5	12	1110.0
114.5	25	2862.5
136.5	10	1365.0
158.5	2	317.0
	$n = 50$	$\Sigma xf = 5725$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{5725}{50} = 114.5$$

$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
−44	1936	1936
−22	484	5808
0	0	0
22	484	4840
44	1936	3872
		$\Sigma(x - \bar{x})^2 f = 16{,}456$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}} = \sqrt{\frac{16{,}456}{49}} = \sqrt{335.83} \approx 18.33$$

42.

Class	f	xf
0	1	0
1	9	9
2	13	26
3	5	15
4	2	8
	$n = 30$	$\Sigma xf = 58$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{58}{30} \approx 1.9$$

$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
−1.93	3.72	3.72
−0.93	0.86	7.74
0.07	0.00	0.00
1.07	1.14	5.70
2.07	4.28	8.56
		$\Sigma(x - \bar{x})^2 f = 25.72$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2 f}{n - 1}} = \sqrt{\frac{25.72}{29}} \approx \sqrt{0.89} \approx 0.9$$

43.

Class	Midpoint, x	f	xf
0–4	2.0	20.3	40.60
5–13	9.0	35.5	319.50
14–17	15.5	16.5	255.75
18–24	21.0	30.4	638.40
25–34	29.5	39.4	1162.30
35–44	39.5	39.0	1540.50
45–64	54.5	80.8	4403.60
65+	70.0	40.4	2828.00
		$n = 302.3$	$\Sigma xf = 11{,}188.65$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{11{,}188.65}{302.3} \approx 37.01$$

$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
−35.01	1225.70	24,881.77
−28.01	784.56	27,851.88
−21.51	462.68	7634.22
−16.01	256.32	7792.13
−7.51	56.40	2222.16
2.49	6.20	241.80
17.49	305.70	24,716.72
32.99	1088.34	43,968.94
		$\sum(x - \bar{x})^2 f = 139{,}309.56$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}} = \sqrt{\frac{139{,}309.56}{301.3}} = \sqrt{462.36} \approx 21.50$$

44.

Midpoint, x	f	xf
5	11.3	56.5
15	12.1	181.5
25	12.8	320.0
35	16.5	577.5
45	18.3	823.5
55	15.2	836.0
65	17.8	1157.0
75	13.4	1005.0
85	7.3	620.5
95	1.5	142.5
	$n = 126.2$	$\Sigma xf = 5720$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{5720}{126.2} \approx 45.32$$

$(x - \bar{x})$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
−40.32	1625.70	18,370.41
−30.32	919.30	11,123.53
−20.32	412.92	5285.12
−10.32	106.50	1757.25
−0.32	0.10	1.83
9.68	93.70	1424.34
19.68	387.30	6893.94
29.68	880.90	11,804.06
39.68	1574.50	11,493.85
49.68	2468.10	3702.15
		$\sum(x - \bar{x})^2 f = 71{,}856.38$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2 f}{n - 1}} = \sqrt{\frac{71{,}856.38}{125.2}} \approx \sqrt{573.93} \approx 23.96$$

45. $CV_{\text{heights}} = \dfrac{\sigma}{\mu} \cdot 100\% = \dfrac{3.44}{72.75} \cdot 100 \approx 4.7$

$CV_{\text{weights}} = \dfrac{\sigma}{\mu} \cdot 100\% = \dfrac{18.47}{187.83} \cdot 100 \approx 9.8$

It appears that weight is more variable than height.

46. (a)

x	x^2
1059	1,121,481
1328	1,763,584
1175	1,380,625
1123	1,261,129
923	851,929
1017	1,034,289
1214	1,473,796
1042	1,085,764
$\sum x = 8881$	$\sum x^2 = 9{,}972{,}597$

Male: $s = \sqrt{\dfrac{\Sigma x^2 - [(\Sigma x)^2/n]}{n-1}} = \sqrt{\dfrac{9{,}972{,}597 - [(8881)^2/8]}{7}}$

$= \sqrt{\dfrac{113{,}576.875}{7}} = \sqrt{16{,}225.268} \approx 127.4$

x	x^2
1226	1,503,076
965	931,225
841	707,281
1053	1,108,809
1056	1,115,136
1393	1,940,449
1312	1,721,344
1222	1,493,284
$\sum x = 9068$	$\sum x^2 = 10{,}520{,}604$

Female: $s = \sqrt{\dfrac{\Sigma x^2 - [(\Sigma x)^2/n]}{n-1}} = \sqrt{\dfrac{10{,}520{,}604 - [(9068)^2/8]}{7}}$

$= \sqrt{\dfrac{242{,}026}{7}} = \sqrt{34{,}575.143} \approx 185.9$

(b) The answers are the same as from Exercise 21.

47. (a) $\bar{x} \approx 41.5 \qquad s \approx 5.3$

(b) $\bar{x} \approx 43.6 \qquad s \approx 5.6$

(c) $\bar{x} \approx 3.5 \qquad s \approx 0.4$

(d) By multiplying each entry by a constant k, the new sample mean is $k \cdot \bar{x}$ and the new sample standard deviation is $k \cdot s$.

48. (a) $\bar{x} \approx 41.7,\ s \approx 6.0$

(b) $\bar{x} \approx 42.7,\ s \approx 6.0$

(c) $\bar{x} \approx 39.7,\ s \approx 6.0$

(d) By adding or subtracting a constant k to each entry, the new sample mean will be $\bar{x} + k$ with the sample standard deviation being unaffected.

49. (a) Male SAT Scores: $\bar{x} = 1110.125$

x	$\lvert x - \bar{x}\rvert$
1059	51.125
1328	217.88
1175	64.875
1123	12.875
923	187.13
1017	93.125
1214	103.88
1042	68.125
	$\Sigma\lvert x - \bar{x}\rvert = 799$

$\Sigma|x - \bar{x}| = 799 \Rightarrow \frac{\Sigma|x - \bar{x}|}{n} = \frac{799}{8} = 99.9$

$s = 127.4$

Female SAT Score: $\bar{x} = 1133.5$

x	$\lvert x - \bar{x}\rvert$
1226	92.5
965	168.5
841	292.5
1053	80.5
1056	77.5
1393	259.5
1312	178.5
1222	88.5
	$\Sigma\lvert x - \bar{x}\rvert = 1238$

$\Sigma|x - \bar{x}| = 799 \Rightarrow \frac{\Sigma|x - \bar{x}|}{n} = \frac{1238}{8} = 154.8$

$s = 185.9$

(b) Public Teachers: $\bar{x} = 37.375$

x	$\lvert x - \bar{x}\rvert$
38.6	1.225
38.1	0.725
38.7	1.325
36.8	0.575
34.8	2.575
35.9	1.475
39.9	2.525
36.2	1.175
	$\Sigma\lvert x - \bar{x}\rvert = 11.6$

$\Sigma|x - \bar{x}| = 11.6 \Rightarrow \frac{\Sigma|x - \bar{x}|}{n} = \frac{11.6}{8} = 1.45$

$s = 1.72$

Private Teachers: $\bar{x} = 19.538$

x	$\lvert x - \bar{x}\rvert$
21.8	2.262
18.4	1.138
20.3	0.762
17.6	1.938
19.7	0.162
18.3	1.238
19.4	0.138
20.8	1.262
	$\Sigma\lvert x - \bar{x}\rvert = 8.9$

$\Sigma|x - \bar{x}| = 8.9 \Rightarrow \frac{\Sigma|x - \bar{x}|}{n} = \frac{8.9}{8} = 1.11$

$s = 1.41$

50. $1 - \frac{1}{k^2} = 0.99 \Rightarrow 1 - 0.99 = \frac{1}{k^2} \Rightarrow k^2 = \frac{1}{0.01} \Rightarrow k = \sqrt{\frac{1}{0.01}} = 10$

At least 99% of the data in any data set lie within 10 standard deviations of the mean.

51. (a) $P = \frac{3(\bar{x} - \text{median})}{s} = \frac{3(17 - 19)}{2.3} \approx -2.61$; skewed left

(b) $P = \frac{3(\bar{x} - \text{median})}{s} = \frac{3(32 - 25)}{5.1} \approx 4.12$; skewed right

2.5 MEASURES OF POSITION

2.5 Try It Yourself Solutions

1a. 15, 15, 15, 17, 18, 18, 20, 21, 21, 21, 21, 22, 22, 23, 24, 24, 25, 25, 26, 26, 27, 27, 27, 28, 28, 28, 29, 29, 29, 30, 30, 31, 31, 31, 31, 31, 32, 32, 32, 32, 32, 32, 32, 33, 33, 33, 34, 34, 34, 34, 35, 35, 35, 35, 35, 36, 36, 37, 37, 37, 38, 38, 38, 39, 39, 39, 40, 40, 40, 41, 41, 42, 42, 43, 43, 43, 44, 44, 44, 44, 45, 45, 45, 45, 46, 46, 46, 46, 47, 47, 48, 48, 48, 48, 49, 50, 50, 52, 53, 53, 54, 54, 55, 55, 55, 55, 57, 59, 59, 60, 61, 61, 61, 63, 63, 65, 68, 89

b. $Q_2 = 37$

c. $Q_1 = 30 \qquad Q_3 = 47$

2a. (Enter the data)

b. $Q_1 = 17 \qquad Q_2 = 23 \qquad Q_3 = 28.5$

c. One quarter of the tuition costs is \$17,000 or less, one half is \$23,000 or less, and three quarters is \$28,500 or less.

3a. $Q_1 = 30 \qquad Q_3 = 47$

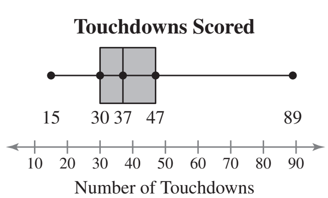
b. $\text{IQR} = Q_3 - Q_1 = 47 - 30 = 17$

c. The touchdowns in the middle half of the data set vary by 17 years.

4a. $\text{Min} = 15 \qquad Q_1 = 30 \qquad Q_2 = 37$

$Q_3 = 47 \qquad \text{Max} = 89$

bc.



d. It appears that half of the teams scored between 30 and 47 touchdowns.

5a. 50th percentile

b. 50% of the teams scored 40 or fewer touchdowns.

6a. $\mu = 70, \sigma = 8$

b. $x = 60\colon\ z = \dfrac{x - \mu}{\sigma} = \dfrac{60 - 70}{8} = -1.25$

$x = 71\colon\ z = \dfrac{x - \mu}{\sigma} = \dfrac{71 - 70}{8} = 0.125$

$x = 92\colon\ z = \dfrac{x - \mu}{\sigma} = \dfrac{92 - 70}{8} = 2.75$

c. From the z-score, the utility bill of \$60 is 1.25 standard deviations below the mean, the bill of \$71 is 0.125 standard deviation above the mean, and the bill of \$92 is 2.75 standard deviations above the mean.

7a. Best supporting actor: $\mu = 50.1, \sigma = 13.9$

Best supporting actress: $\mu = 39.7, \sigma = 1.4$

b. Alan Arkin: $x = 72 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{72 - 50.1}{13.9} = 1.58$

Jennifer Hudson: $x = 25 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{25 - 39.7}{14} = -1.05$

c. Alan Arkin's age is 1.58 standard deviations above the mean of the best support actors. Jennifer Hudson's age is 1.05 standard deviations below the mean of the best supporting actresses. Neither actor's age is unusual.

2.5 EXERCISE SOLUTIONS

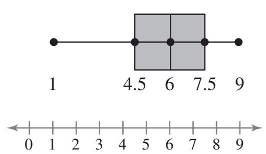
1. (a)

$$\overbrace{1\ 2\ 2\ 4\ 4\ 5\ 5\ 5\ 6\ 6}^{\text{lower half}}\ \overbrace{6\ 7\ 7\ 7\ 7\ 8\ 8\ 8\ 9\ 9}^{\text{upper half}}$$

Q_1 Q_2 Q_3

$Q_1 = 4.5 \qquad Q_2 = 6 \qquad Q_3 = 7.5$

(b)



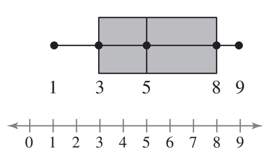
2. (a)

$$\overbrace{1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 4, 4, 4, 5,}^{\text{lower half}}\ 5,\ \overbrace{5, 5, 5, 6, 7, 7, 7, 7, 8, 8, 8, 9, 9, 9, 9, 9, 9}^{\text{upper half}}$$

Q_1 Q_2 Q_3

$Q_1 = 3 \qquad Q_2 = 5 \qquad Q_3 = 8$

(b)



3. The soccer team scored fewer points per game than 75% of the teams in the league.

4. The salesperson sold more hardware equipment than 80% of the other sales people.

5. The student scored higher than 78% of the students who took the actuarial exam.

6. The child's IQ is higher than 93% of the other children in the same age group.

7. True

8. False. The five numbers you need to graph a box-and-whisker plot are the minimum, the maximum, Q_1, Q_3, and the median (Q_2).

9. False. The 50th percentile is equivalent to Q_2.

10. False. Any score equal to the mean will have a corresponding z-score of zero.

11. (a) Min = 10 (b) Max = 20

(c) $Q_1 = 13$ (d) $Q_2 = 15$

(e) $Q_3 = 17$ (f) $\text{IQR} = Q_3 - Q_1 = 17 - 13 = 4$

12. (a) Min = 100 (b) Max = 320

(c) $Q_1 = 130$ (d) $Q_2 = 205$

(e) $Q_3 = 270$ (f) $\text{IQR} = Q_3 - Q_1 = 270 - 130 = 140$

13. (a) Min = 900 (b) Max = 2100

(c) $Q_1 = 1250$ (d) $Q_2 = 1500$

(e) $Q_3 = 1950$ (f) $\text{IQR} = Q_3 - Q_1 = 1950 - 1250 = 700$

14. (a) Min = 25 (b) Max = 85

(c) $Q_1 = 50$ (d) $Q_2 = 65$

(e) $Q_3 = 70$ (f) $\text{IQR} = Q_3 - Q_1 = 70 - 50 = 20$

15. (a) Min = -1.9 (b) Max = 2.1

(c) $Q_1 = -0.5$ (d) $Q_2 = 0.1$

(e) $Q_3 = 0.7$ (f) $\text{IQR} = Q_3 - Q_1 = 0.7 - (-0.5) = 1.2$

16. (a) Min = -1.3 (b) Max = 2.1

(c) $Q_1 = -0.3$ (d) $Q_2 = 0.2$

(e) $Q_3 = 0.4$ (f) $\text{IQR} = Q_3 - Q_1 = 0.4 - (-0.3) = 0.7$

17. None. The data are not skewed or symmetric.

18. Skewed right. Most of the data lie to the right.

19. Skewed left. Most of the data lie to the left.

20. Symmetric

21. $Q_1 = \text{B}$, $Q_2 = \text{A}$, $Q_3 = \text{C}$

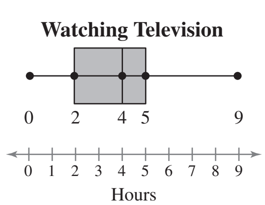
25% of the entries are below B, 50% are below A, and 75% are below C.

22. $P_{10} = \text{T}$, $P_{50} = \text{R}$, $P_{80} = \text{S}$

10% of the entries are below T, 50% are below R, and 80% are below S.

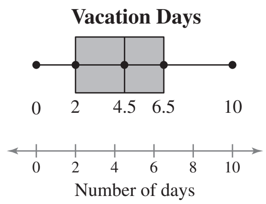
23. (a) $Q_1 = 2$, $Q_2 = 4$, $Q_3 = 5$

(b)



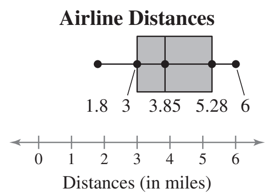
24. (a) $Q_1 = 2, \quad Q_2 = 4.5, \quad Q_3 = 6.5$

(b)



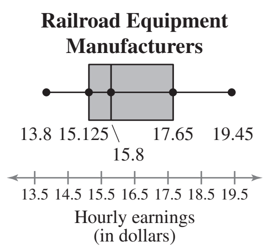
25. (a) $Q_1 = 3, \quad Q_2 = 3.85, \quad Q_3 = 5.28$

(b)



26. (a) $Q_1 = 15.125, \; Q_2 = 15.8, \; Q_3 = 17.65$

(b)



27. (a) 5 (b) 50% (c) 25%

28. (a) \$17.65 (b) 50% (c) 50%

29. $A \Rightarrow z = -1.43$

$B \Rightarrow z = 0$

$C \Rightarrow z = 2.14$

The z-score 2.14 is unusual because it is so large.

30. $A \rightarrow z = -1.54$

$B \rightarrow z = 0.77$

$C \rightarrow z = 1.54$

None of the z-scores are unusual.

31. (a) Statistics:

$$x = 73 \Rightarrow z = \frac{x - \mu}{\sigma} = \frac{73 - 63}{7} \approx 1.43$$

Biology:

$$x = 26 \Rightarrow z = \frac{x - \mu}{\sigma} = \frac{26 - 23}{3.9} \approx 0.77$$

(b) The student had a better score on the statistics test.

32. (a) Statistics: $x = 60 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{60 - 63}{7} \approx -0.43$

Biology: $x = 20 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{20 - 23}{3.9} \approx -0.77$

(b) The student had a better score on the statistics test.

33. (a) Statistics:

$x = 78 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{78 - 63}{7} \approx 2.14$

Biology:

$x = 29 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{29 - 23}{3.9} \approx 1.54$

(b) The student had a better score on the statistics test.

34. (a) Statistics: $x = 63 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{63 - 63}{7} = 0$

Biology: $x = 23 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{23 - 23}{3.9} = 0$

(b) The student performed equally well on the two tests.

35. (a) $x = 34{,}000 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{34{,}000 - 35{,}000}{2{,}250} \approx -0.44$

$x = 37{,}000 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{37{,}000 - 35{,}000}{2{,}250} \approx 0.89$

$x = 31{,}000 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{31{,}000 - 35{,}000}{2{,}250} \approx -1.78$

None of the selected tires have unusual life spans.

(b) $x = 30{,}500 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{30{,}500 - 35{,}000}{2{,}250} = -2 \Rightarrow$ 2.5th percentile

$x = 37{,}250 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{37{,}250 - 35{,}000}{2{,}250} = 1 \Rightarrow$ 84th percentile

$x = 35{,}000 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{35{,}000 - 35{,}000}{2{,}250} = 0 \Rightarrow$ 50th percentile

36. (a) $x = 34 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{34 - 33}{4} = 0.25$

$x = 30 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{30 - 33}{4} = -0.75$

$x = 42 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{42 - 33}{4} = 2.25$

The life span of 42 days is unusual due to a rather large z-score.

(b) $x = 29 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{29 - 33}{4} = -1 \Rightarrow$ 16th percentile

$x = 41 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{41 - 33}{4} = 2 \Rightarrow$ 97.5th percentile

$x = 25 \Rightarrow z = \dfrac{x - \mu}{\sigma} = \dfrac{25 - 33}{4} = -2 \Rightarrow$ 2.5th percentile

37. 68.5 inches

40% of the heights are below 68.5 inches.

38. 99th percentile

99% of the heights are below 76 inches.

39. $x = 74$: $z = \dfrac{x - \mu}{\sigma} = \dfrac{74 - 69.6}{3.0} \approx 1.47$

$x = 62$: $z = \dfrac{x - \mu}{\sigma} = \dfrac{62 - 69.6}{3.0} \approx -2.53$

$x = 80$: $z = \dfrac{x - \mu}{\sigma} = \dfrac{80 - 69.6}{3.0} \approx 3.47$

The height of 62 inches is unusual due to a rather small z-score. The height of 80 inches is very unusual due to a rather large z-score.

40. $x = 70$: $z = \dfrac{x - \mu}{\sigma} = \dfrac{70 - 69.6}{3.0} \approx 0.13$

$x = 66$: $z = \dfrac{x - \mu}{\sigma} = \dfrac{66 - 69.6}{3.0} \approx -1.20$

$x = 68$: $z = \dfrac{x - \mu}{\sigma} = \dfrac{68 - 69.6}{3.0} \approx -0.53$

None of the heights are unusual.

41. $x = 71.1$: $z = \dfrac{x - \mu}{\sigma} = \dfrac{71.1 - 69.6}{3.0} \approx 0.5$

Approximately the 70th percentile.

42. $x = 66.3$: $z = \dfrac{x - \mu}{\sigma} = \dfrac{66.3 - 69.6}{3.0} = -1.1$

Approximately the 12th percentile.

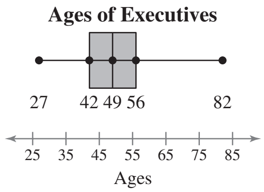
43. (a)

27	28	31	32	32	33	35	36	36	36	36	37	38	39	39	40	40	40	41	41
41	42	42	42	42	42	42	43	43	43	44	44	45	45	46	47	47	47	47	47
48	48	48	48	48	49	49	49	49	49	49	50	50	51	51	51	51	51	51	52
52	52	53	53	54	54	54	54	54	54	54	54	55	56	56	56	57	57	57	59
59	59	60	60	60	61	61	61	62	62	63	63	63	63	64	65	67	68	74	82

Q_1 $\quad$ Q_2 $\quad$ Q_3

$Q_1 = 42, \qquad Q_2 = 49, \qquad Q_3 = 56$

(b)



(c) Half of the ages are between 42 and 56 years.

(d) About 49 years old ($x = 49.62$ and $Q_2 = 49.00$), because half of the executives are older and half are younger.

(e) The age groups 20–29, 70–79, and 80–89 would all be considered unusual because they lie more than two standard deviations from the mean.

44. 1, 2, 3, 3, 5, 5, 7, 7, 8, 10

$\uparrow$ $\quad$ $\uparrow$ $\quad$ $\uparrow$

Q_1 $\quad$ Q_2 $\quad$ Q_3

$$\text{Midquartile} = \frac{Q_1 + Q_3}{2} = \frac{3 + 7}{2} = 5$$

45. 22 23 24 32 33 34 36 38 39 40 41 47

$\uparrow$ $\quad$ $\uparrow$ $\quad$ $\uparrow$

$Q_1 = 28$ $\quad$ Q_2 $\quad$ $Q_3 = 39.5$

$$\text{Midquartile} = \frac{Q_1 + Q_3}{2} = \frac{28 + 39.5}{2} = 33.75$$

46. 7.9, 8, 8.1, 9.7, 10.3, 11.2, 11.8, 12.2, 12.3, 12.7, 13.4, 15.4, 16.1

$\uparrow$ $\quad$ $\uparrow$ $\quad$ $\uparrow$

$Q_1 = 8.9$ $\quad$ Q_2 $\quad$ $Q_3 = 13.05$

$$\text{Midquartile} = \frac{Q_1 + Q_3}{2} = \frac{8.9 + 13.05}{2} = 10.975$$

47. 13.4 15.2 15.6 16.7 17.2 18.7 19.7 19.8 19.8 20.8 21.4 22.9 28.7 30.1 31.9

$\uparrow$ $\quad$ $\uparrow$ $\quad$ $\uparrow$

Q_1 $\quad$ Q_2 $\quad$ Q_3

$$\text{Midquartile} = \frac{Q_1 + Q_3}{2} = \frac{16.7 + 22.9}{2} = 19.8$$

48. (a) Disc 1: Symmetric

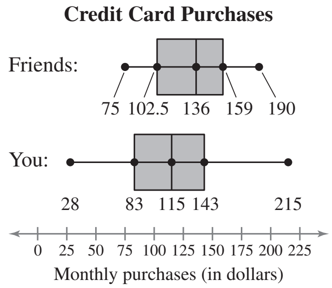
Disc 2: Skewed left

Disc 1 has less variation.

(b) Disc 2 is more likely to have outliers because its distribution is wider.

(c) Disc 1, because the distribution's typical distance from the mean is roughly 16.3.

49.



You	Friends
min = 28	min = 75
$Q_1 = 83$	$Q_1 = 102.5$
$Q_2 = 115$	$Q_2 = 136$
$Q_3 = 143$	$Q_3 = 159$
max = 215	max = 190

Your distribution is symmetric and your friend's distribution is uniform.

50. $$\text{Percentile} = \frac{\text{Number of data values less than } x}{\text{Total number of data values}} \cdot 100$$

$$= \frac{51}{80} \cdot 100 \approx 64\text{th percentile}$$

51. $$\text{Percentile} = \frac{\text{Number of data values less than } x}{\text{Total number of data values}} \cdot 100$$

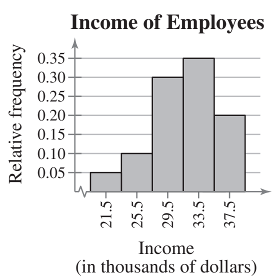
$$= \frac{75}{80} \cdot 100 \approx 94\text{th percentile}$$

CHAPTER 2 REVIEW EXERCISE SOLUTIONS

1.

Class	Midpoint	Boundaries	Frequency, f	Relative frequency	Cumulative frequency
20–23	21.5	19.5–23.5	1	0.05	1
24–27	25.5	23.5–27.5	2	0.10	3
28–31	29.5	27.5–31.5	6	0.30	9
32–35	33.5	31.5–35.5	7	0.35	16
36–39	37.5	35.5–39.5	4	0.20	20
			$\Sigma f = 20$	$\Sigma \frac{f}{n} = 1$	

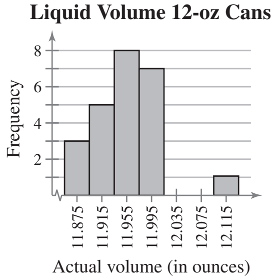
2.



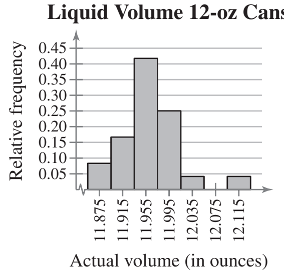
Greatest relative frequency: 32–35

Least relative frequency: 20–23

3.

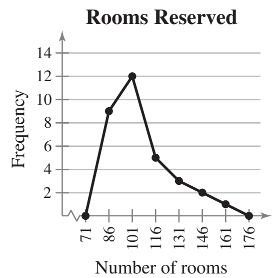


4.

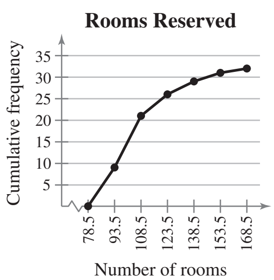


5.

Class	Midpoint, x	Frequency, f	Cumulative frequency
79–93	86	9	9
94–108	101	12	21
109–123	116	5	26
124–138	131	3	29
139–153	146	2	31
154–168	161	1	32
		$\Sigma f = 32$	



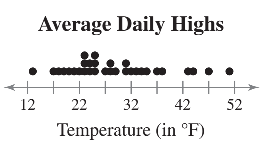
6.



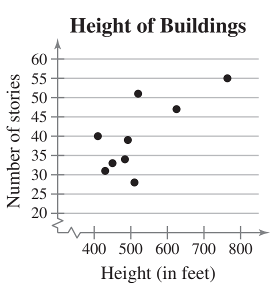
7.

1	3 7 8 9
2	0 1 2 3 3 3 4 4 5 5 5 7 8 8 9
3	1 1 2 3 4 5 7 8
4	3 4 7
5	1

8.

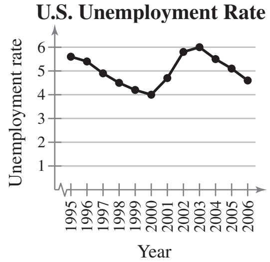


9.

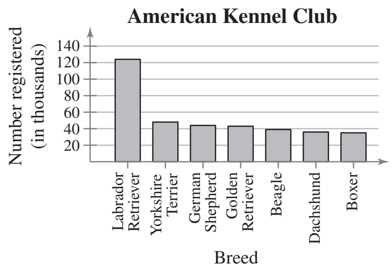


It appears as height increases, the number of stories increases.

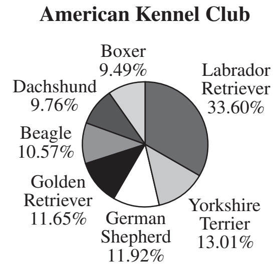
10.



11.



12.



13. $\bar{x} = 9.1$

median = 8.5

mode = 7

14. $\bar{x} = 40.6$

median = 42

mode = 42

15.

Midpoint, x	Frequency, f	xf
21.5	1	21.5
25.5	2	51.0
29.5	6	177.0
33.5	7	234.5
37.5	4	150.0
	$n = 20$	$\Sigma xf = 634$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{634}{20} = 31.7$$

16.

x	f	xf
0	13	0
1	9	9
2	19	38
3	8	24
4	5	20
5	2	10
6	4	24
	$n = 60$	$\Sigma xf = 125$

$$\bar{x} = \frac{\Sigma xf}{n} = \frac{125}{60} \approx 2.1$$

17. $$\bar{x} = \frac{\Sigma xw}{w} = \frac{(78)(0.15) + (72)(0.15) + (86)(0.15) + (91)(0.15) + (87)(0.15) + (80)(0.25)}{0.15 + 0.15 + 0.15 + 0.15 + 0.15 + 0.25}$$

$$= \frac{82.1}{1} = 82.1$$

18. $\bar{x} = \frac{\Sigma xw}{w} = \frac{(96)(0.20) + (85)(0.20) + (91)(0.20) + (86)(0.40)}{0.20 + 0.20 + 0.20 + 0.40}$

$= \frac{88.8}{1} = 88.8$

19. Skewed **20.** Skewed **21.** Skewed left

22. Skewed right **23.** Median **24.** Mean

25. Range = Max − Min = 8.26 − 5.46 = 2.8

26. Range = Max − Min = 19.73 − 15.89 = 3.84

27. $\mu = \frac{\Sigma x}{N} = \frac{96}{14} = 6.9$

$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}} = \sqrt{\frac{(4 - 6.9)^2 + (2 - 6.9)^2 + \cdots + (3 - 6.9)^2 + (3 - 6.9)^2}{12}}$

$= \sqrt{\frac{295.7}{14}} \approx \sqrt{21.12} \approx 4.6$

28. $\mu = \frac{\Sigma x}{N} = \frac{602}{9} \approx 66.9$

$\sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{N}}$

$= \sqrt{\frac{(52 - 66.9)^2 + (86 - 66.9)^2 + \cdots + (68 - 66.9)^2 + (56 - 66.9)^2}{9}}$

$\approx \sqrt{\frac{862.87}{9}} \approx \sqrt{95.87} \approx 9.8$

29. $\bar{x} = \frac{\Sigma x}{n} = \frac{36{,}801}{15} = 2453.4$

$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}} = \sqrt{\frac{(2445 - 2453.4)^2 + \cdots + (2.377 - 2453.4)^2}{14}}$

$= \sqrt{\frac{1{,}311{,}783.6}{14}} \approx \sqrt{93{,}698.8} \approx 306.1$

30. $\bar{x} = \frac{\Sigma x}{n} = \frac{416{,}659}{8} = 52{,}082.4$

$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}} = \sqrt{\frac{(49{,}632 - 52{,}082.3)^2 + \cdots + (49{,}924 - 52{,}082.3)^2}{7}}$

$= \sqrt{\frac{73{,}225{,}929.87}{7}} = \sqrt{10{,}460{,}847.12} \approx 3234.3$

31. 99.7% of the distribution lies within 3 standard deviations of the mean.

$\mu + 3\sigma = 49 + (3)(2.50) = 41.5$

$\mu - 3\sigma = 49 - (3)(2.50) = 56.5$

99.7% of the distribution lies between \$41.50 and \$56.50.

32. $(46.75, 52.25) \to (49.50 - (1)(2.75), 49.50 + (1)(2.75)) \to (\mu - \sigma, \mu + \sigma)$

68% of the cable rates lie between \$46.75 and \$52.25.

33. $n = 40 \qquad \mu = 36 \qquad \sigma = 8$

$(20, 52) \to (36 - 2(8), 36 + 2(8)) \Longrightarrow (\mu - 2\sigma, \mu + 2\sigma) \Longrightarrow k = 2$

$$1 = \frac{1}{k^2} = 1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = 0.75$$

At least $(40)(0.75) = 30$ customers have a mean sale between \$20 and \$52.

34. $n = 20 \qquad \mu = 7 \qquad \sigma = 2$

$(3, 11) \to (7 - 2(2), 7 + 2(2)) \to (\mu - 2\sigma, \mu + 2\sigma) \to k = 2$

$$1 - \frac{1}{k^2} = 1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = 0.75$$

At least $(20)(0.75) = 15$ shuttle flights lasted between 3 days and 11 days.

35. $\bar{x} = \dfrac{\Sigma xf}{n} = \dfrac{99}{40} \approx 2.5$

$$s = \sqrt{\frac{\Sigma\,(x - \bar{x})^2 f}{n - 1}} = \sqrt{\frac{(0 - 1.24)^2(1) + (1 - 1.24)^2(8) + \cdots + (5 - 1.24)^2(3)}{39}}$$

$$= \sqrt{\frac{59.975}{39}} \approx 1.2$$

36. $\bar{x} = \dfrac{\Sigma xf}{n} = \dfrac{61}{25} \approx 2.4$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2 f}{n - 1}}$$

$$= \sqrt{\frac{(0 - 2.44)^2(4) + (1 - 2.44)^2(5) + \cdots + (6 - 2.44)^2(1)}{24}}$$

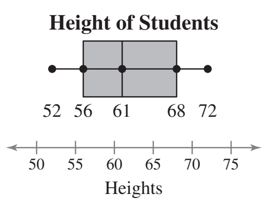
$$= \sqrt{\frac{72.16}{24}} \approx 1.7$$

37. $Q_1 = 56$ inches

38. $Q_3 = 68$ inches

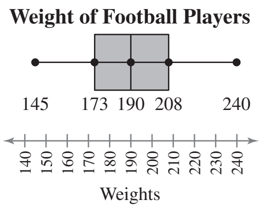
39. $\text{IQR} = Q_3 - Q_1 = 68 - 56 = 12$ inches

40.



41. $\text{IQR} = Q_3 - Q_1 = 33 - 29 = 4$

42.



43. 23% of the students scored higher than 68.

44. $\frac{84}{728} \approx 0.109 \rightarrow 11\%$ have larger audiences.

The station would represent the 89th percentile, P_{89}.

45. $x = 213 \Rightarrow z = \frac{x - \mu}{\sigma} = \frac{213 - 186}{18} \approx 1.5$

This player is not unusual.

46. $x = 141 \Rightarrow z = \frac{x - \mu}{\sigma} = \frac{141 - 186}{18} = -2.5$

This is an unusually light player.

47. $x = 178 \Rightarrow z = \frac{x - \mu}{\sigma} = \frac{178 - 186}{18} = -0.44$

This player is not unusual.

48. $x = 249 \Rightarrow z = \frac{x - \mu}{\sigma} = \frac{249 - 186}{18} \approx 3.5$

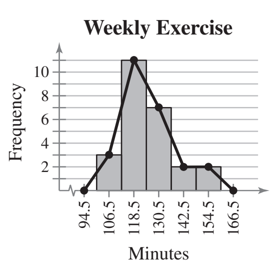
This is an unusually heavy player.

CHAPTER 2 QUIZ SOLUTIONS

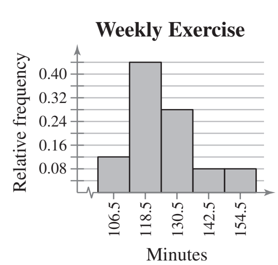
1. (a)

Class limits	Midpoint	Class boundaries	Frequency, f	Relative frequency	Cumulative frequency
101–112	106.5	100.5–112.5	3	0.12	3
113–124	118.5	112.5–124.5	11	0.44	14
125–136	130.5	124.5–136.5	7	0.28	21
137–148	142.5	136.5–148.5	2	0.08	23
149–160	154.5	148.5–160.5	2	0.08	25

(b) Frequency Histogram and Polygon

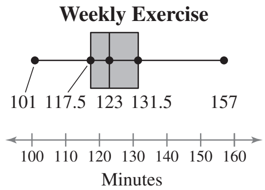


(c) Relative Frequency Histogram



(d) Skewed

(e)



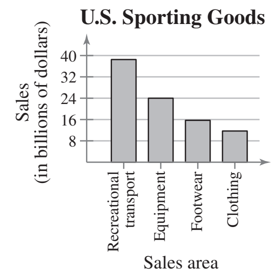
(f)

2. $\bar{x} = \dfrac{\Sigma xf}{n} = \dfrac{3130.5}{25} \approx 125.2$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2 f}{n - 1}} = \sqrt{\frac{4055.04}{24}} \approx 13.0$$

3. (a)

(b)



4. (a) $\bar{x} = \dfrac{\Sigma x}{n} = 751.6$

median $= 784.5$

mode $=$ (none)

The mean best describes a typical salary because there are no outliers.

(b) range $=$ Max $-$ Min $= 575$

$$s^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} = 48{,}135.1$$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}} = 219.4$$

5. $\bar{x} - 2s = 155{,}000 - 2 \cdot 15{,}000 = \$125{,}000$

$\bar{x} + 2s = 155{,}000 + 2 \cdot 15{,}000 = \$185{,}000$

95% of the new home prices fall between \$125,000 and \$185,000.

6. (a) $x = 200{,}000 \qquad z = \dfrac{x - \mu}{\sigma} = \dfrac{200{,}000 - 155{,}000}{15{,}000} = 3.0 \Rightarrow$ unusual price

(b) $x = 55{,}000 \qquad z = \dfrac{x - \mu}{\sigma} = \dfrac{55{,}000 - 155{,}000}{15{,}000} \approx -6.67 \Rightarrow$ very unusual price

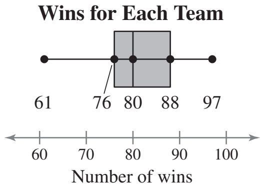
(c) $x = 175{,}000 \qquad z = \dfrac{x - \mu}{\sigma} = \dfrac{175{,}000 - 155{,}000}{15{,}000} \approx 1.33 \Rightarrow$ not unusual

(d) $x = 122{,}000 \qquad z = \dfrac{x - \mu}{\sigma} = \dfrac{122{,}000 - 155{,}000}{15{,}000} = -2.2 \Rightarrow$ unusual price

7. (a) $Q_1 = 76 \qquad Q_2 = 80 \qquad Q_3 = 88$

(b) $\text{IQR} = Q_3 - Q_1 = 88 - 76 = 12$

(c)



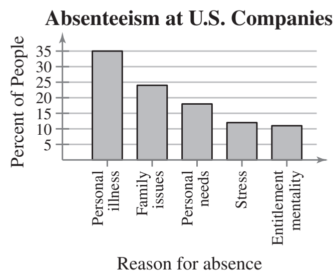
CUMULATIVE REVIEW FOR CHAPTERS 1 AND 2

1. Systematic sampling

2. Simple Random Sampling. However, all U.S. adults may not have a telephone.

3.

Personal Illness	35%
Family Issues	24%
Personal Needs	18%
Stress	12%
Entitlement Mentality	11%



4. \$42,500 is a parameter because it is describing the average salary of all 43 employees in a company.

5. 28% is a statistic because it is describing a proportion within a sample of 1000 adults.

6. (a) $\bar{x} = 83{,}500$, $s = \$1500$

$(80{,}500 \;\; 86{,}500) = 83{,}500 \pm 2(1500) \Rightarrow$ 2 standard deviations away from the mean.

Approximately 95% of the electrical engineers will have salaries between (80,500, 86,500).

(b) $40(.95) = 38$

7. Sample, because a survey of 1498 adults was taken.

8. Sample, because a study of 232,606 people was done.

9. Census, because all of the members of the Senate were included.

10. Experiment, because we want to study the effects of removing recess from schools.

11. Quantitative: Ratio

12. Qualitative: Nominal

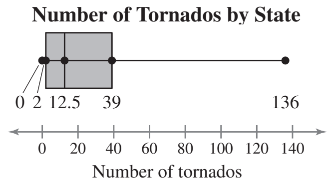
13. min $= 0$

$Q_1 = 2$

$Q_2 = 12.5$

$Q_3 = 39$

max $= 136$



14. $$\bar{x} = \frac{(0.15)(85) + (0.15)(92) + (0.15)(84) + (215)(89) + (0.40)(91)}{0.15 + 0.15 + 0.15 + 0.15 + 0.40} = 88.9$$

15. (a) $\bar{x} = 5.49$

median $= 5.4$

mode $=$ none

Median, because the distribution is not symmetric.

(b) Range $= 4.1$

$s^2 = 2.34$

$s = 1.53$

The standard deviation of tail lengths of alligators is 1.53 feet.

16. (a) The number of deaths due to heart disease for women is decreasing.

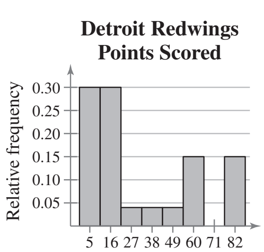
(b) The study was only conducted over the past 5 years and deaths may not decrease in the next year.

17. Class width $= \dfrac{87 - 0}{8} = 10.875 \Longrightarrow 11$

Class limits	Class boundaries	Class midpoint	Frequency, f	Relative frequency	Cumulative frequency
0–10	−0.5–10.5	5	8	0.296	8
11–21	10.5–21.5	16	8	0.296	16
22–32	21.5–32.5	27	1	0.037	17
33–43	32.5–43.5	38	1	0.037	18
44–54	43.5–54.5	49	1	0.037	19
55–65	54.5–65.5	60	4	0.148	23
66–76	65.5–76.5	71	0	0.000	23
77–87	76.5–87.5	82	4	0.148	27
			$n = 27$	$\sum \frac{f}{n} = 1$	

18. Skewed right

19.



Least relative frequency: 66–76

Greatest relative frequency: 0–10 and 11–21