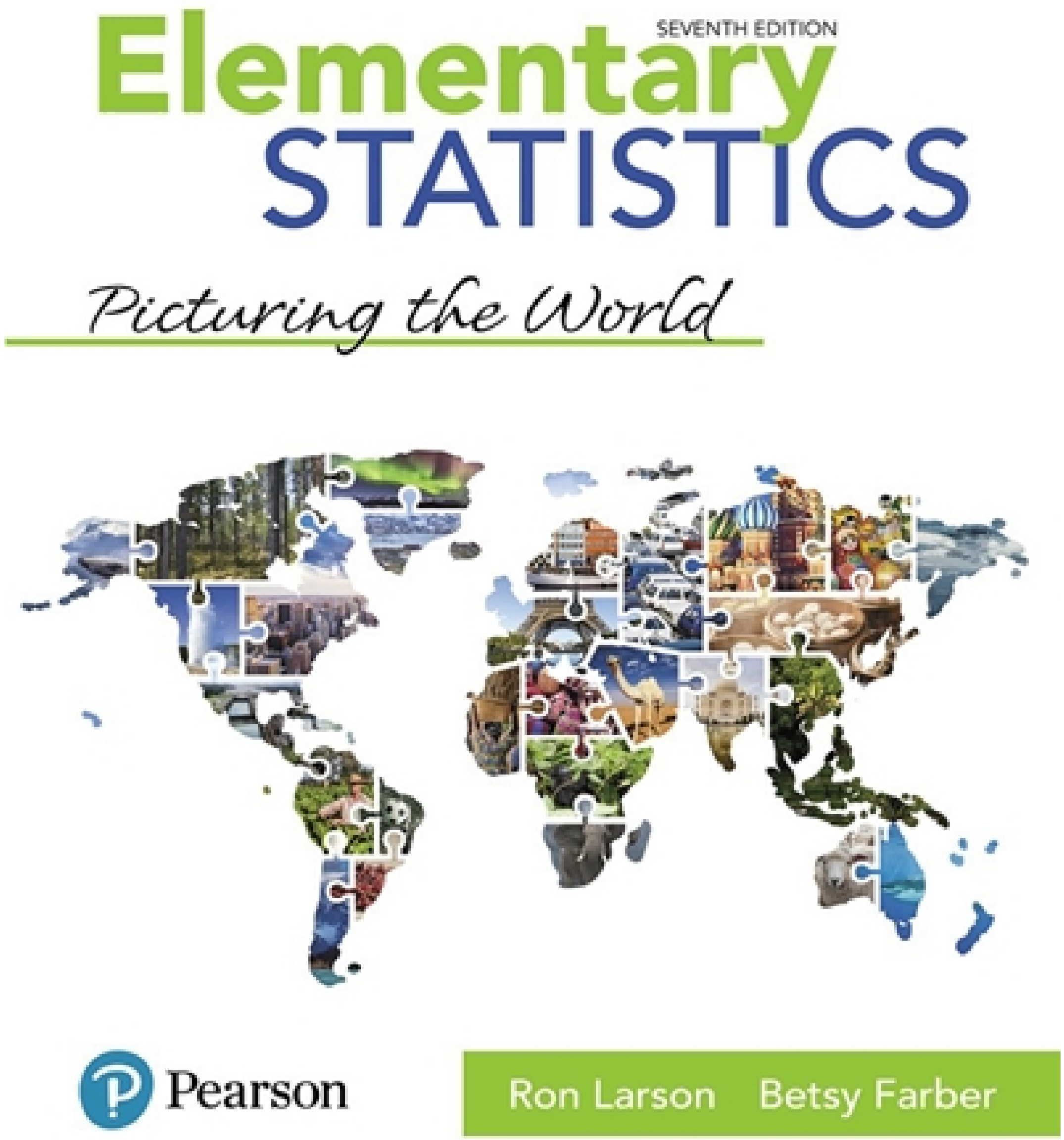
Solutions for Elementary Statistics Picturing the World 7th Edition by Larson

CLICK HERE TO ACCESS COMPLETE Solutions



Solutions

CHAPTER 2

Descriptive Statistics

2.1 FREQUENCY DISTRIBUTIONS AND THEIR GRAPHS

2.1 TRY IT YOURSELF SOLUTIONS

1. The number of classes is 6.

$$\text{Min} = 14,\ \text{Max} = 55,\ \text{Class width} = \frac{\text{Range}}{\text{Number of classes}} = \frac{55-14}{6} = 6.83 \Rightarrow 7$$

The minimum data entry is a convenient lower limit for the first class. Then add the class width to get the lower limits of the other classes. The upper limits are one less than the lower limit of the next class.

Lower limit	Upper limit
14	20
21	27
28	34
35	41
42	48
49	55

Make a tally mark for each entry in the appropriate class. The number of tally marks for a class is the frequency of that class.

Class	Frequency, f
14-20	8
21-27	15
28-34	14
35-41	7
42-48	4
49-55	3

2. Find each midpoint, relative frequency, and cumulative frequency.

$$\text{Midpoint} = \frac{(\text{Lower class limit}) + (\text{Upper class limit})}{2}$$

$$\text{Relative frequency} = \frac{\text{Class frequency}}{\text{Sample size}} = \frac{f}{n}$$

The cumulative frequency of a class is the sum of the frequencies of that class and all previous classes.

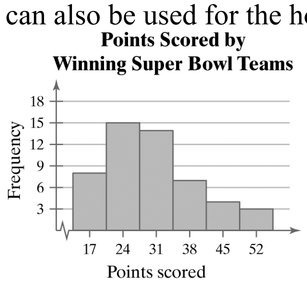
Class	f	Midpoint	Relative frequency	Cumulative frequency
14–20	8	$\frac{14+20}{2}=17$	$\frac{8}{51}\approx 0.1569$	8
21–27	15	$\frac{21+27}{2}=24$	$\frac{15}{51}\approx 0.2941$	$8+15=23$
28–34	14	$\frac{28+34}{2}=31$	$\frac{14}{51}\approx 0.2745$	$23+14=37$
35–41	7	$\frac{35+41}{2}=38$	$\frac{7}{51}\approx 0.1373$	$37+7=44$
42–48	4	$\frac{42+48}{2}=45$	$\frac{4}{51}\approx 0.0784$	$44+4=48$
49–55	3	$\frac{49+55}{2}=52$	$\frac{3}{51}\approx 0.0588$	$48+3=51$
	$\sum f=51$		$\sum \frac{f}{n}=1$	

Sample answer: The most common range of points scored by winning teams is 21 to 27. About 14% of the winning teams scored more than 41 points.

3. Find the class boundaries. Because the data entries are integers, subtract 0.5 from each lower limit to find the lower class boundaries and add 0.5 to each upper limit to find the upper class boundaries.

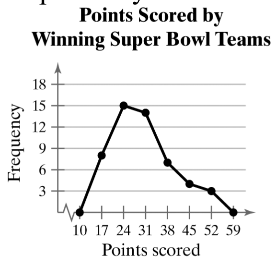
Class	Class Boundaries	Frequency, f
14–20	13.5–20.5	8
21–27	20.5–27.5	15
28–34	27.5–34.5	14
35–41	34.5–41.5	7
42–48	41.5–48.5	4
49–55	48.5–55.5	3

Use class midpoints for the horizontal scale and frequency for the vertical scale. (Class boundaries can also be used for the horizontal scale.)



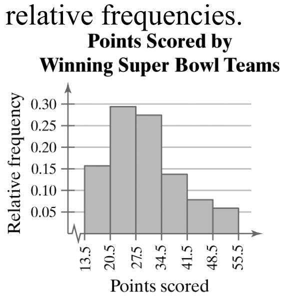
Sample answer: The most common range of points scored by winning teams is 21 to 27. About 14% of the winning teams scored more than 41 points.

4. To construct the frequency polygon, use the same horizontal and vertical scales that were used in the histogram labeled with the class midpoints in Try It Yourself 3. Then plot the points that represent the midpoint and frequency of each class and connect the points with line segments. Extend the left side and right side to one class width before the first class midpoint and after the last class midpoint, respectively.

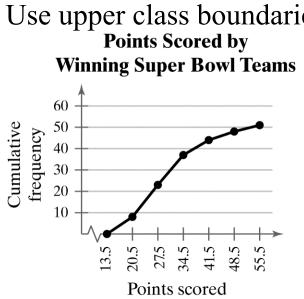


Sample answer: The frequency of points scored increases up to 24 points and then decreases.

5. Notice the shape of the relative frequency histogram is the same as the shape of the frequency histogram constructed in Try It Yourself 3. The only difference is that the vertical scale measures the relative frequencies.

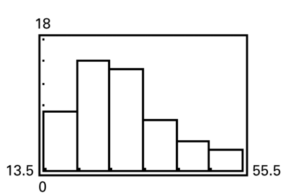


6. Use upper class boundaries for the horizontal scale and cumulative frequency for the vertical scale.



Sample answer: The greatest increase in cumulative frequency occurs between 20.5 and 27.5.

7.



2.1 EXERCISE SOLUTIONS

1. Organizing the data into a frequency distribution may make patterns within the data more evident. Sometimes it is easier to identify patterns of a data set by looking at a graph of the frequency distribution.

2. If there are too few or too many classes, it may be difficult to detect patterns because the data are too condensed or too spread out.

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. Relative frequency of a class is the portion, or percentage, of the data that falls in that class.
Cumulative frequency of a class is the sum of the frequencies of that class and all previous classes.

5. The sum of the relative frequencies must be 1 or 100% because it is the sum of all portions or percentages of the data.

6. A frequency polygon displays frequencies or relative frequencies whereas an ogive displays cumulative frequencies.

7. False. Class width is the difference between the lower (or upper limits) of consecutive classes.

8. True

9. False. An ogive is a graph that displays cumulative frequencies.

10. True

11. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{64-9}{7} \approx 7.9 \Rightarrow 8$
Lower class limits: 9, 17, 25, 33, 41, 49, 57
Upper class limits: 16, 24, 32, 40, 48, 56, 64

12. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{88-12}{6} \approx 12.7 \Rightarrow 13$
Lower class limits: 12, 25, 38, 51, 64, 77
Upper class limits: 24, 37, 50, 63, 76, 89

13. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{135-17}{8} = 14.75 \Rightarrow 15$
Lower class limits: 17, 32, 47, 62, 77, 92, 107, 122
Upper class limits: 31, 46, 61, 76, 91, 106, 121, 136

14. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{247-54}{10} = 19.3 \Rightarrow 20$
Lower class limits: 54, 74, 94, 114, 134, 154, 174, 194, 214, 234
Upper class limits: 73, 93, 113, 133, 153, 173, 193, 213, 233, 253

15. (a) Class width $= 11 - 0 = 11$

(b) and (c)

$$\text{Midpoint} = \frac{(\text{Lower class limit}) + (\text{Upper class limit})}{2}$$

Find the class boundaries. Because the data entries are integers, subtract 0.5 from each lower limit to find the lower class boundaries and add 0.5 to each upper limit to find the upper class boundaries.

Class	Midpoint	Class boundaries
0 – 10	5	–0.5 – 10.5
11 – 21	16	10.5 – 21.5
22 – 32	27	21.5 – 32.5
33 – 43	38	32.5 – 43.5
44 – 54	49	43.5 – 54.5
55 – 65	60	54.5 – 65.5
66 – 76	71	65.5 – 76.5

16. (a) Class width $= 33 - 25 = 8$

(b) and (c)

$$\text{Midpoint} = \frac{(\text{Lower class limit}) + (\text{Upper class limit})}{2}$$

Find the class boundaries. Because the data entries are integers, subtract 0.5 from each lower limit to find the lower class boundaries and add 0.5 to each upper limit to find the upper class boundaries.

Class	Midpoint	Class boundaries
25–32	28.5	24.5–32.5
33–40	36.5	32.5–40.5
41–48	44.5	40.5–48.5
49–56	52.5	48.5–56.5
57–64	60.5	56.5–64.5
65–72	68.5	64.5–72.5
73–80	76.5	72.5–80.5

17. Relative frequency $= \dfrac{\text{Class frequency}}{\text{Sample size}} = \dfrac{f}{n}$

The cumulative frequency of a class is the sum of the frequencies of that class and all previous classes.

Class	Frequency f	Midpoint	Relative frequency	Cumulative frequency
0 – 10	188	5	0.15	188
11 – 21	372	16	0.30	560
22 – 32	264	27	0.22	824
33 – 43	205	38	0.17	1029
44 – 54	83	49	0.07	1112
55 – 65	76	60	0.06	1188
66 – 76	32	71	0.03	1220
	$\sum f = 1220$		$\sum \frac{f}{n} = 1$	

18. Relative frequency $= \dfrac{\text{Class frequency}}{\text{Sample size}} = \dfrac{f}{n}$

The cumulative frequency of a class is the sum of the frequencies of that class and all previous classes.

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
25 – 32	86	28.5	0.24	86
33 – 40	39	36.5	0.11	125
41 – 48	41	44.5	0.11	166
49 – 56	48	52.5	0.13	214
57 – 64	43	60.5	0.12	257
65 – 72	68	68.5	0.19	325
73 – 80	40	76.5	0.11	365
	$\sum f = 365$		$\sum \frac{f}{n} \approx 1$	

19. (a) Number of classes: 7

(b) Greatest frequency: about 300
Least frequency: about 10

(c) Class width: 10

(d) *Sample answer:* About half of the employee salaries are between \$50,000 and \$69,000.

20. (a) Number of classes: 6

(b) Greatest frequency: 37
Least frequency: 1

(c) Class width: 53

(d) *Sample answer:* The heights of most roller coasters are less than 231 feet.

21. Identify the highest point and its respective class. Class with greatest frequency: 506 – 510
Identify the lowest point (not including the points on the horizontal axis) and its respective class.
Class with least frequency: 474 – 478

22. Identify the highest point and its respective class. Class with greatest frequency: 3.5 – 4.5 miles
Identify the lowest point (not including the points on the horizontal axis) and its respective class.
Class with least frequency: 0.5 – 1.5 miles

23. (a) Identify the tallest bar and its respective class. Class with greatest relative frequency: 35 – 36 centimeters
Identify the shortest bar and its respective class. Class with least relative frequency: 39 – 40 centimeters

(b) Greatest relative frequency ≈ 0.25
Least relative frequency ≈ 0.01

(c) *Sample answer*: From the graph, 0.25 or 25% of females have a fibula length between 35 and 36 centimeters.

24. (a) Identify the tallest bar and its respective class. Class with greatest relative frequency: 11 – 12 minutes
Identify the shortest bar and its respective class. Class with least relative frequency: 14 – 15 minutes

(b) Greatest relative frequency ≈ 38%
Least relative frequency ≈ 4%

(c) *Sample answer*: From the graph, about 0.75 or 75% of campus security response times are between 11 and 13 minutes.

25. (a) Locate the cumulative frequency of the highest (right-most) point. The number in the sample is 75.

(b) Locate the neighboring points where the pitch between them is the steepest. The greatest increase in frequency is from 158.5 – 201.5 pounds.

26. (a) Locate the cumulative frequency of the highest (right-most) point. The number in the sample is 77.

(b) Locate the neighboring points where the pitch between them is the steepest. The greatest increase in frequency is from 68 – 70 inches.

27. (a) Locate 201.5 on the horizontal axis and find the corresponding cumulative frequency at the point on the ogive: 47

(b) Locate 68 on the vertical axis and find the corresponding weight at the point on the ogive: 287.5 pounds

(c) Subtract the cumulative frequency for each weight: 62 – 22 = 40

(d) Subtract the cumulative frequency for bears weighing 330.5 pounds from the number in the sample: 75 – 69 = 6

28. (a) Locate 72 on the horizontal axis and find the corresponding cumulative frequency at the point on the ogive: 71

(b) Locate 15 on the vertical axis and find the corresponding height at the point on the ogive: 68 inches

(c) Subtract the cumulative frequency for each height: 71 – 15 = 56

(d) Subtract the cumulative frequency for adult males that are 70 inches tall from the number in the sample: 77 – 47 = 30

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

Class	**Frequency, f**	**Midpoint**	**Relative frequency**	**Cumulative frequency**
0 – 7	8	3.5	0.33	8
8–15	7	11.5	0.29	15
16–23	3	19.5	0.13	18
24–31	3	27.5	0.13	21
32–39	3	35.5	0.13	24
	$\sum f = 24$		$\sum \frac{f}{n} \approx 1$	

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

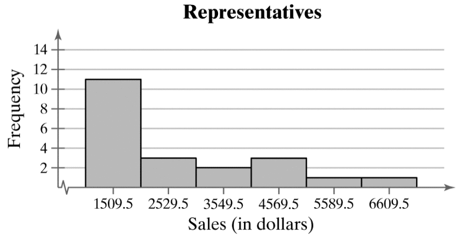
30. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{530-30}{6} \approx 83.3 \Rightarrow 84$

Class	**Frequency, f**	**Midpoint**	**Relative frequency**	**Cumulative frequency**
30–113	5	71.5	0.17	5
114–197	7	155.5	0.23	12
198–281	8	239.5	0.27	20
282–365	3	323.5	0.10	23
366–449	3	407.5	0.10	26
450–533	4	491.5	0.13	30
	$\sum f = 30$		$\sum \frac{f}{n} = 1$	

31. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{7119-1000}{6} \approx 1019.8 \Rightarrow 1020$

Class	Frequency, f	Mid-point	Relative frequency	Cumulative frequency
1000 –2019	11	1509.5	0.52	11
2020 –3039	3	2529.5	0.14	14
3040– 4059	2	3549.5	0.10	16
4060– 5079	3	4569.5	0.14	19
5080– 6099	1	5589.5	0.05	20
6100– 7119	1	6609.5	0.05	21
	$\sum f = 21$		$\sum \frac{f}{n} = 1$	

July Sales for Representatives

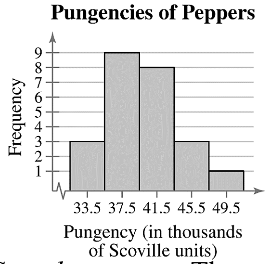


Sample answer: The graph shows that most of the sales representatives at the company sold from \$1000 to \$2019.

32. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{51-32}{5} = 3.8 \Rightarrow 4$

Class	Frequency, f	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
	$\sum f = 24$		$\sum \frac{f}{n} = 1$	

Pungencies of Peppers

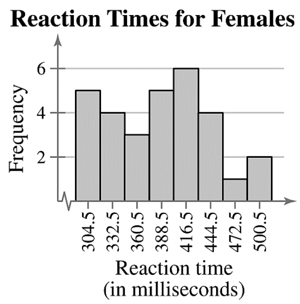


Sample answer: The graph shows that most of the pungencies of the peppers were between 36,000 and 43,000 Scoville units.

33. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{514-291}{8} = 27.875 \Rightarrow 28$

CLICK HERE TO ACCESS THE COMPLETE Solutions

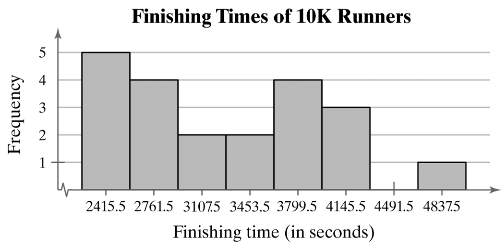
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
291-318	5	304.5	0.1667	5
319-346	4	332.5	0.1333	9
347-374	3	360.5	0.1000	12
375-402	5	388.5	0.1667	17
403-430	6	416.5	0.2000	23
431-458	4	444.5	0.1333	27
459-486	1	472.5	0.0333	28
487-514	2	500.5	0.0667	30
	$\sum f = 30$		$\sum \frac{f}{n}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



Sample answer: The graph shows that the most frequent reaction times were between 403 and 430 milliseconds.

34. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{5008 - 2243}{8} = 345.625 \Rightarrow 346$

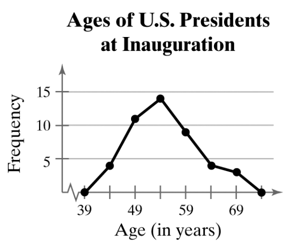
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
2243–2588	5	2415.5	0.2381	5
2589–2934	4	2761.5	0.1905	9
2935–3280	2	3107.5	0.0952	11
3281–3626	2	3453.5	0.0952	13
3627–3972	4	3799.5	0.1905	17
3973–4318	3	4145.5	0.1429	20
4319–4664	0	4491.5	0.0000	20
4665–5010	1	4837.5	0.0476	21
	$\sum f = 21$		$\sum \frac{f}{n} = 1$	



Sample answer: The graph shows that the most frequent finishing times were from 2243 to 2588 seconds.

35. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{70-42}{6} \approx 4.7 \Rightarrow 5$

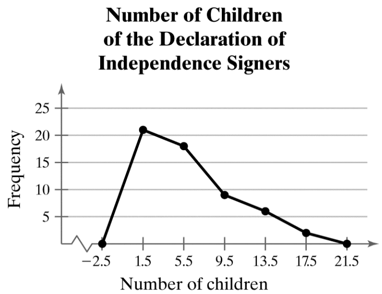
Class	**Frequency, *f***	**Midpoint**	**Relative frequency**	**Cumulative frequency**
42–46	4	44	0.0889	4
47–51	11	49	0.2444	15
52–56	14	54	0.3111	29
57–61	9	59	0.2000	38
62–66	4	64	0.0889	42
67–71	3	69	0.0667	45
	$\sum f = 45$		$\sum \frac{f}{n} = 1$	



Sample answer: The graph shows that the number of U.S. presidents who were 52 or older at inauguration was twice as many as those who were 51 and younger.

36. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{18-0}{5} = 3.6 \Rightarrow 4$

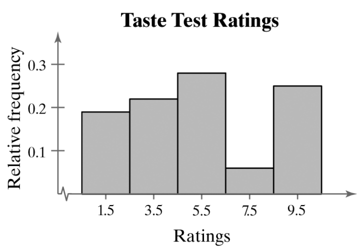
Class	**Frequency, *f***	**Midpoint**	**Relative frequency**	**Cumulative frequency**
0–3	21	1.5	0.3750	21
4–7	18	5.5	0.3214	39
8–11	9	9.5	0.1607	48
12–15	6	13.5	0.1071	54
16–19	2	17.5	0.0357	56
	$\sum f = 56$		$\sum \frac{f}{n} \approx 1$	



Sample answer: The graph shows that most of the signers of the Declaration of Independence had 7 or fewer children.

37. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{10-1}{5} = 1.8 \Rightarrow 2$

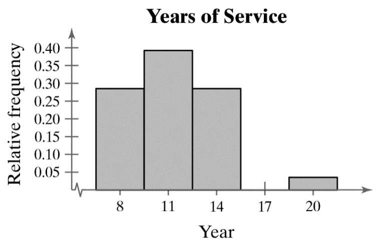
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
1–2	7	1.5	0.19	7
3–4	8	3.5	0.22	15
5–6	10	5.5	0.28	25
7–8	2	7.5	0.06	27
9–10	9	9.5	0.25	36
	$\sum f = 36$		$\sum \frac{f}{n} \approx 1$	



Class with greatest relative frequency: 5 – 6
Class with least relative frequency: 7 – 8

38. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{19-7}{5} = 2.4 \Rightarrow 3$

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
7–9	8	8	0.2857	8
10–12	11	11	0.3929	19
13–15	8	14	0.2857	27
16–18	0	17	0.0000	27
19–21	1	20	0.0357	28
	$\sum f = 28$		$\sum \frac{f}{n} \approx 1$	

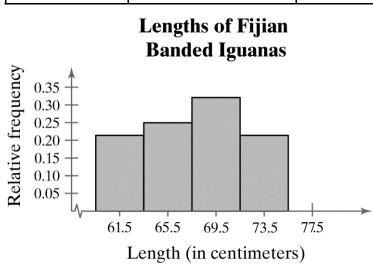


Class with greatest relative frequency: 10 – 12
Class with least relative frequency: 16 – 18

39. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{75-60}{5} = 3$

Notice that using a class width of 3 is not wide enough to include all the data with 5 classes. Therefore, use a class width of 4.

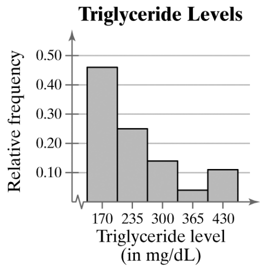
Class	Frequency ,*f*	Midpoint	Relative frequency	Cumulative frequency
60–63	6	61.5	0.2143	6
64–67	7	65.5	0.2500	13
68–71	9	69.5	0.3214	22
72–75	6	73.5	0.2143	28
76–79	0	77.5	0.0000	28
	$\sum f = 28$		$\sum \frac{f}{n} \approx 1$	



Class with greatest relative frequency: 68 – 71
Class with least relative frequency: 76 – 79

40. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{462-138}{5} = 64.8 \Rightarrow 65$

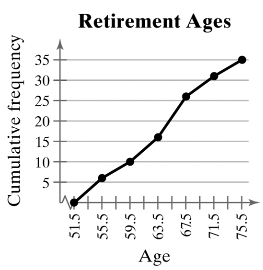
Class	Frequency, *f*	Midpoint	Relative frequency	Cumulative frequency
138–202	13	170	0.46	13
203–267	7	235	0.25	20
268–332	4	300	0.14	24
333–397	1	365	0.04	25
398–462	3	430	0.11	28
	$\sum f = 28$		$\sum \frac{f}{n} \approx 1$	



Class with greatest relative frequency: 138 – 202
Class with least relative frequency: 333 – 397

41. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{75-52}{6} \approx 3.8 \Rightarrow 4$

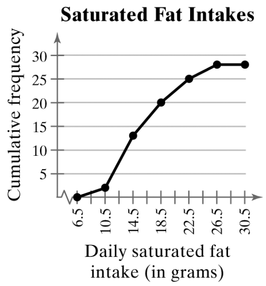
Class	Frequency, *f*	Relative frequency	Cumulative frequency
52–55	6	0.1714	6
56–59	4	0.1143	10
60–63	6	0.1714	16
64–67	10	0.2857	26
68–71	5	0.1429	31
72–75	4	0.1143	35
	$\sum f = 35$		$\sum \frac{f}{n} \approx 1$



Location of the greatest increase in frequency: 64 – 67

42. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{26-7}{6} \approx 3.2 \Rightarrow 4$

Class	Frequency, *f*	Relative frequency	Cumulative frequency
7–10	2	0.0714	2
11–14	11	0.3929	13
15–18	7	0.2500	20
19–22	5	0.1786	25
23–26	3	0.1071	28
27–30	0	0.0000	28
	$\sum f = 28$	$\sum \frac{f}{n} \approx 1$	

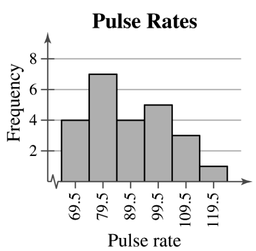


Location of the greatest increase in frequency: 11 – 14

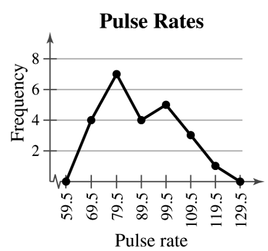
43. (a) Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{120-65}{6} \approx 9.2 \Rightarrow 10$

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
65-74	4	69.5	0.1667	4
75-84	7	79.5	0.2917	11
85-94	4	89.5	0.1667	15
95-104	5	99.5	0.2083	20
105-114	3	109.5	0.1250	23
115-124	1	119.5	0.0417	24
	$\sum f = 24$		$\sum \frac{f}{N} \approx 1$	

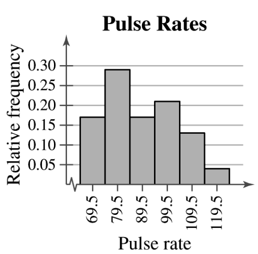
(b)



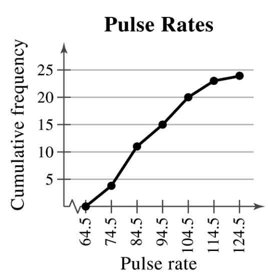
(c)



(d)



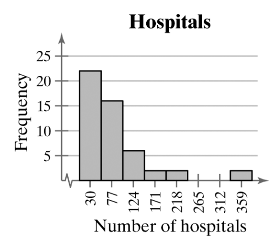
(e)



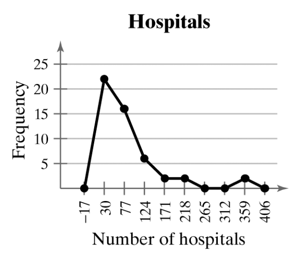
44. (a) Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{382-7}{8} = 46.875 \Rightarrow 47$

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
7–53	22	30	0.44	22
54–100	16	77	0.32	38
101–147	6	124	0.12	44
148–194	2	171	0.04	46
195–241	2	218	0.04	48
242–288	0	265	0.00	48
289–335	0	312	0.00	48
336–382	2	359	0.04	50
	$\sum f = 50$		$\sum \dfrac{f}{n} = 1$	

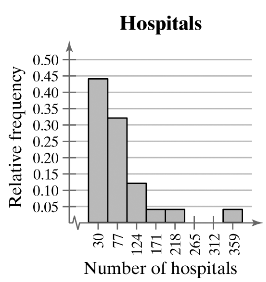
(b)



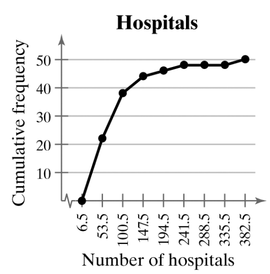
(c)



(d)



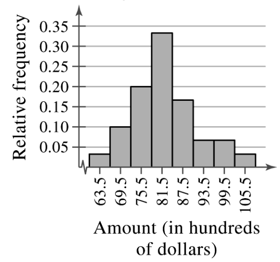
(e)



45. (a) Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{104-61}{8} = 5.375 \Rightarrow 6$

Class	Frequency, f	Midpoint	Relative frequency
61-66	1	63.5	0.033
67-72	3	69.5	0.100
73-78	6	75.5	0.200
79-84	10	81.5	0.333
85-90	5	87.5	0.167
91-96	2	93.5	0.067
97-102	2	99.5	0.067
103-108	1	105.5	0.033
	$\sum f = 30$		$\sum \dfrac{f}{n} = 1$

Daily Withdrawals



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

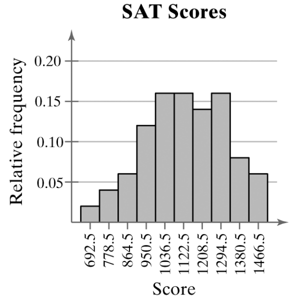
(c) \$9700, because the sum of the relative frequencies for the last two classes is 0.10.

46. (a) $\text{Class width} = \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{1500 - 650}{10} = 85$

Notice that using a class width of 85 is not wide enough to include all the data with 10 classes. Therefore, use a class width of 86.

Class	Frequency, f	Midpoint	Relative frequency
650-735	1	692.5	0.02
736-821	2	778.5	0.04
822-907	3	864.5	0.06
908-993	6	950.5	0.12
994-1079	8	1036.5	0.16
1080-1165	8	1122.5	0.16
1166-1251	7	1208.5	0.14
1252-1337	8	1294.5	0.16
1338-1423	4	1380.5	0.08
1424-1509	3	1466.5	0.06
	$\sum f = 50$		$\sum \frac{f}{n} = 1$

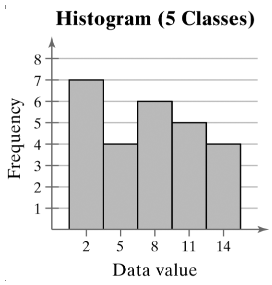
SAT Scores

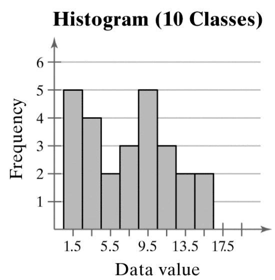


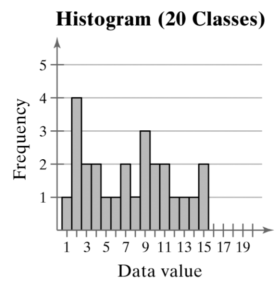
(b) 64%; The portion of the scores greater than or equal to 1070 is 0.64.

(c) A score of 908 or above, because the sum of the relative frequencies of the class starting with 908 and all classes with higher scores is 0.88.

47.







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. In 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

1. Because the data entries go from a low of 14 to a high of 55, use stem values from 1 to 5. List the stems to the left of a vertical line. For each data entry, list a leaf to the right of its stem.

1	4 6 6 6 7	Key: 1 | 4 = 14
2	0 0 0 1 1 1 3 3 4 4 4 4 6 7 7 7 7 7 8 9	
3	0 1 1 1 1 2 2 3 4 4 4 4 5 5 5 7 8 8 9	
4	2 3 6 8 9	
5	2 5	

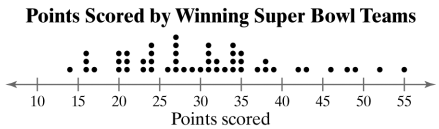
Sample answer: Most of the winning teams scored between 20 and 39 points.

2. Use the leaves 0, 1, 2, 3, and 4 in the first stem row and the leaves 5, 6, 7, 8, and 9 in the second stem row.

1	4	Key: 1 | 4 = 14
1	6 6 6 7	
2	0 0 0 1 1 1 3 3 4 4 4 4	
2	6 7 7 7 7 7 8 9	
3	0 1 1 1 1 2 2 3 4 4 4 4	
3	5 5 5 7 8 8 9	
4	2 3	
4	6 8 9	
5	2	
5	5	

Sample answer: Most of the winning teams scored from 20 to 35 points.

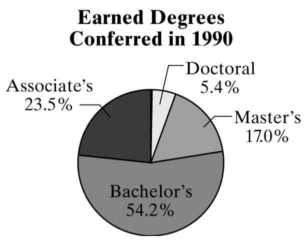
3. Choose the horizontal axis so that each data entry is included in the dot plot. For example, label the horizontal axis from 10 to 55.



Sample answer: Most of the points scored by winning teams cluster between 20 and 40.

4.

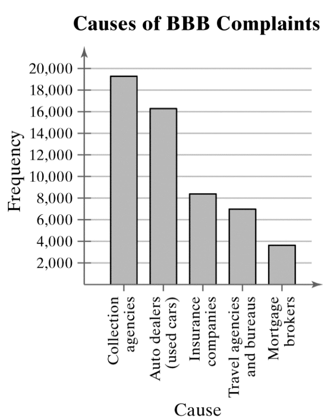
Type of Degree	f	Relative Frequency	Angle
Associate's	455	0.235	85°
Bachelor's	1051	0.542	195°
Master's	330	0.170	61°
Doctoral	104	0.054	19°
	$\sum f = 1940$	$\sum \frac{f}{n} \approx 1$	$\sum = 360°$



From 1990 to 2014, as percentages of the total degrees conferred, associate's degrees increased by 2.9%, bachelor's degrees decreased by 5.1%, master's degrees increased by 2.8%, and doctoral degrees decreased by 0.7%.

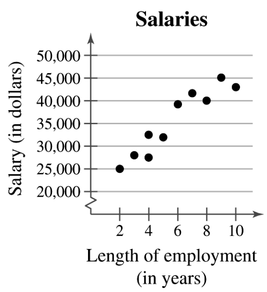
5.

Cause	Frequency, f
Auto dealers (used cars)	16,281
Insurance companies	8384
Mortgage brokers	3634
Collection agencies	19,277
Travel agencies and bureaus	6985



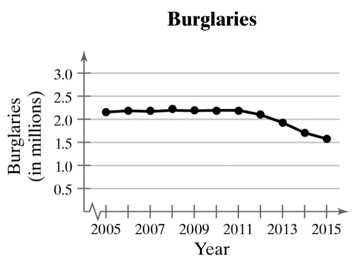
Collection agencies are the greatest cause of complaints.

6.



It appears that the longer an employee is with the company, the greater the employee's salary.

7. Let the horizontal axis represent the years and let the vertical axis represent the number of burglaries (in millions).



Sample answer: The number of burglaries remained about the same until 2012 and then decreased through 2015.

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 a 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 at 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. *Sample answer*: Facebook has the most users, and Pinterest has the least. Tumblr and Instagram have about the same number of users.

14. *Sample answer*: The year 2010 had the most motor vehicle thefts and 2013 had the least. Motor vehicle thefts decreased the most between 2011 and 2012.

15. *Sample answer*: The Texter is the least popular driver. The Left-Lane Hog is tolerated more than the Tailgater. The Speedster and the Drifter have the same popularity.

16. *Sample answer*: Food is the most costly aspect of pet care and live animal purchases is the least. The amounts spent on veterinarian care and supplies/OTC medicine are about the same.

17. Exam Scores 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

Sample answer: Most grades for the biology midterm were in the 80s and 90s.

18. Hours Worked by Nurses Key: $2|4 = 24$

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

Sample answer: Most nurses work between 30 and 40 hours per week.

19. Ice Thickness (in centimeters) 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

Sample answer: Most of the ice had a thickness of 5.8 centimeters to 7.2 centimeters.

20. Tomato Prices (in dollars per pound)

15	4 7	Key: 15 | 4 = 1.54
16	0 1 1 3 4 4 6 8	
17	1 4 7 8 8 9	
18	2 3 6 7 9	
19	1 3 7 8	
20	7 7 8	
21	1 3	

Sample answer: Most retail outlets charge $1.60 to $1.79 per pound of tomatoes.

21. Incomes (in millions) of Highest Paid Athletes

3	3 4 4 4	Key: 3 | 3 = 33
3	5 6 7 7 8 8 8	
4	1 2 3 4 4	
4	5 5 5 6	
5	0 3 3 3	
5	6 6	
6		
6	8	
7		
7	7	
8	1	
8	8	

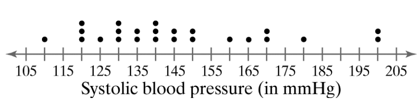
Sample answer: Most of the highest-paid athletes have an income of $33 million to $56 million.

22. Electoral Votes for the 50 States

0	3 3 3 3 3 3 3 4 4 4 4 4	Key: 0 | 3 = 3
0	5 5 5 6 6 6 6 6 6 7 7 7 8 8 9 9 9	
1	0 0 0 0 1 1 1 1 2 3 4	
1	5 6 6 8	
2	0 0	
2	9 9	
3		
3	8	
4		
4		
5		
5	5	

Sample answer: Over half the states have less than 10 electoral votes.

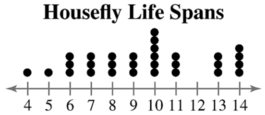
23.



Sample answer: Systolic blood pressure tends to be from 120 to 150 millimeters of mercury.

CLICK HERE TO ACCESS THE COMPLETE Solutions

24. **Housefly Life Spans**

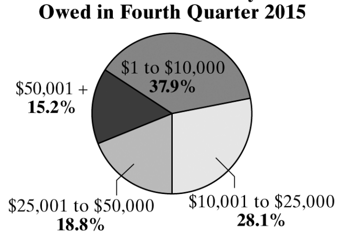


Sample answer: The lifespan of a housefly tends to be from 6 to 14 days.

25.

Balance Owed	f	Relative Frequency	Angle
\$1 to \$10,000	16.7	0.379	136°
\$10,001 to \$25,000	12.4	0.281	101°
\$25,001 to \$50,000	8.3	0.188	68°
\$50,001 +	6.7	0.152	55°
	$\sum f = 44.1$	$\sum \frac{f}{n} = 1$	$\sum \; = 360°$

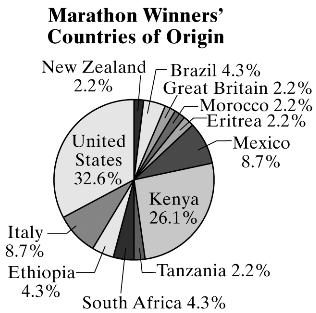
Student Loan Borrowers by Balance Owed in Fourth Quarter 2015



Sample answer: The majority of student loan borrowers owe \$25,000 or less.

26.

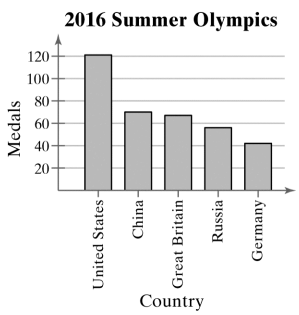
Category	Frequency, f	Relative Frequency	Angle
United States	15	0.326	117°
Italy	4	0.087	31°
Ethiopia	2	0.043	15°
South Africa	2	0.043	15°
Tanzania	1	0.022	8°
Kenya	12	0.261	94°
Mexico	4	0.087	31°
Morocco	1	0.022	8°
Great Britain	1	0.022	8°
Brazil	2	0.043	15°
New Zealand	1	0.022	8°
Eritrea	1	0.022	8°
	$\sum f = 46$	$\sum \frac{f}{n} \approx 1$	$\sum \; \approx 360°$



Sample answer: Most of the New York City Marathon winners are from the United States and Kenya.

27.

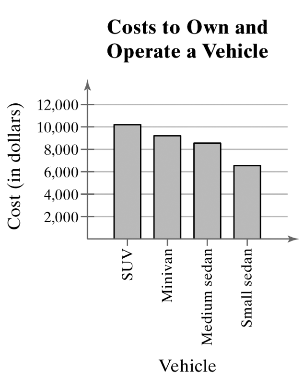
Country	**Medals**
Germany	42
Great Britain	67
United States	121
Russia	56
China	70



Sample answer: The United States won the most medals out of the five countries and Germany won the least.

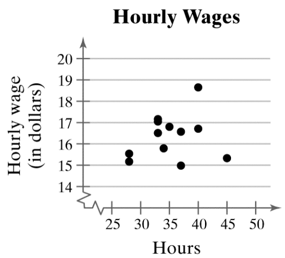
28.

Type of Vehicle	**Cost**
Small sedan	$6579
Medium sedan	$8604
SUV	$10,255
Minivan	$9262



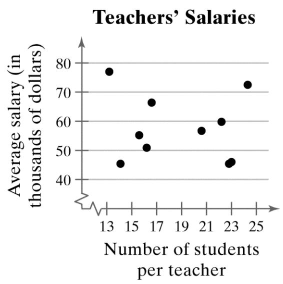
Sample answer: It costs the least to own and operate a small sedan.

29. Let the horizontal axis represent hours and let the vertical axis represent the hourly wage (in dollars).



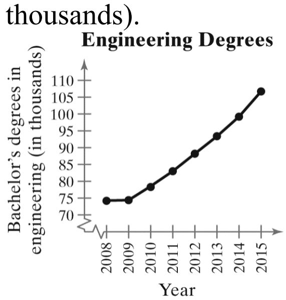
Sample answer: It appears that there is no relation between hourly wages and hours worked.

30. Let the horizontal axis represent the number of students per teacher and let the vertical axis represent the average salary (in thousands of dollars).



Sample answer: It appears that there is no relation between a teacher's average salary and the number of students per teacher.

31. Let the horizontal axis represent the years and let the vertical axis represent the number of degrees (in thousands).



Sample answer: The number of bachelor's degrees in engineering conferred in the U.S. has increased from 2008 to 2015.

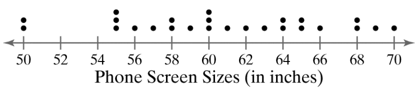
32. Let the horizontal axis represent the years and let the vertical axis represent the percent.

Sample answer: The percentage of the U.S. gross domestic product that comes from the construction sector decreased from 2007 to 2011 but then increased from 2012 to 2015.

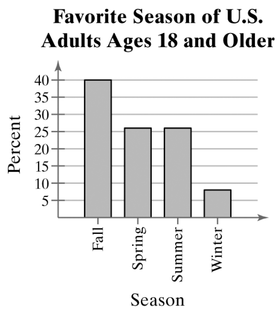
33. Heights (in inches)

7	2 2 4	Key: 7 | 2 = 72
7	5 5 5 5 6	
8	1 1 2 2 2 4	
8		

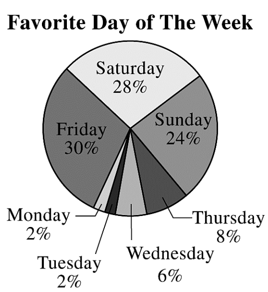
The dot plot helps you see that the data are clustered from 72 to 76 and 81 to 84, with 75 being the most frequent value. The stem-and-leaf plot helps you see that most values are 75 or greater.

34. 

The stem-and-leaf plot helps you see that most values are from 60 to 69. The dot plot helps you see that the values 55 and 60 occur most frequently.

35. 

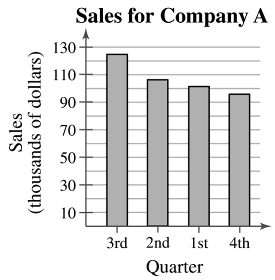
The pie chart helps you to see the percentages as parts of a whole, with fall being the largest. It also shows that while fall is the largest percentage, it makes up less than half of the pie chart. That means that a majority of U.S. adults ages 18 and older prefer a season other than fall. This means it would not be a fair statement to say that most U.S. adults ages 18 and older prefer fall. The Pareto chart helps you to see the rankings of the seasons.

36. 

The Pareto chart helps you see the order from the most favorite to least favorite day. The pie chart helps you visualize the data as parts of a whole and see that about 80% of people say their favorite day is Friday, Saturday, or Sunday.

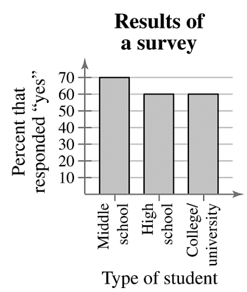
37. (a) The graph is misleading because the large gap from 0 to 90 makes it appear that the sales for the 3rd quarter are disproportionately larger than the other quarters.

(b)



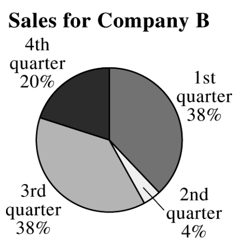
38. (a) The graph is misleading because the vertical axis has no break. The percent of middle schoolers that responded "yes" appears three times larger than either of the others when the difference is only 10%.

(b)



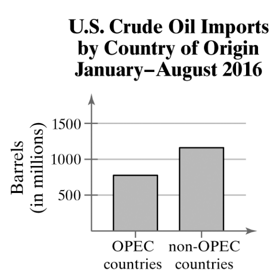
39. (a) The graph is misleading because the angle makes it appear as though the 3rd quarter had a larger percent of sales than the others, when the 1st and 3rd quarters have the same percent.

(b)



40. (a) The graph is misleading because the "non-OPEC countries" bar is wider than the "OPEC countries" bar.

(b)



41. (a) At Law Firm A, the lowest salary was $90,000 and the highest salary was $203,000. At Law Firm B, the lowest salary was $90,000 and the highest salary was $190,000. There are 30 lawyers at Law Firm A and 32 lawyers at Law Firm B.

(b) At Law Firm A, the salaries tend to be clustered at the far ends of the distribution range. At Law Firm B, the salaries are spread out.

42. (a)

3:00 P.M. Class		8:00 P.M. Class
	1	8 8 8 8 8 9 9 9 9 9
	2	0 0 0 2 3 4 4 5 5 8 9 9
8 5	3	1 1 9
0	4	3 4 4
9 7 5 3 1	5	6
9 8 8 8 8 4 2 0	6	
7 7 6 5 5 5 3 3	7	1
5 4	8	

Key: 5 | 3 | 1 = 35-year-old in 3:00 P.M. class and 31-year-old in 8:00 P.M. class

(b) In the 3:00 P.M. class, the lowest age is 35 years old and the highest age is 85 years old. In the 8:00 P.M. class, the lowest age is 18 years old and the highest age is 71 years old. There are 26 participants in the 3:00 P.M. class and there are 30 participants in the 8:00 P.M. class.

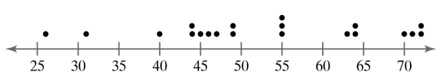
(c) *Sample answer:* The participants in each class are clustered at one of the ends of their distribution range. The 3:00 P.M. class mostly has participants over 50 years old and the 8:00 P.M. class mostly has participants under 50 years old.

43. (a)

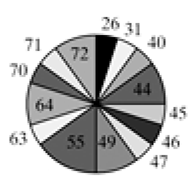
2	6
3	1
4	0 4 4 5 6 7 9 9
0	4 4 5 6 7 9 9
5	5 5 5
6	3 4 4
7	0 1 2 2

Key: 2 | 6 = 26

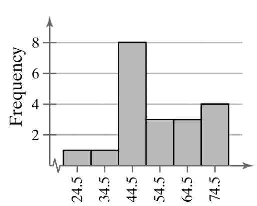
(b)

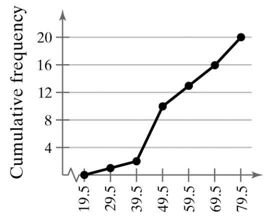


(c)



(d)



(e) 

Sample answer: The stem-and-leaf plot, dot plot, frequency histogram, and ogive display the data best because the data is quantitative.

2.3 MEASURES OF CENTRAL TENDENCY

2.3 TRY IT YOURSELF SOLUTIONS

1. $\sum x = 35 + 33 + 16 + 23 + 16 + ... + 28 + 24 + 34 = 1541$

$\bar{x} = \frac{\sum x}{n} = \frac{1541}{51} \approx 30.2$

The mean points scored by the 51 winning teams is about 30.2.

2. Order the data from smallest to largest.
Note: The stem-and-leaf plot from Try It Yourself 1 in Section 2.2 may be helpful to in ordering the data.
14, 16, 16, 16, 17, 20, 20, 20, 21, 21, 21, 23, 23, 24, 24, 24, 24, 26, 27, 27, 27, 27, 27, 28, 29, 30, 31, 31, 31, 31, 32, 32, 33, 34, 34, 34, 34, 35, 35, 35, 37, 38, 38, 39, 42, 43, 46, 48, 49, 52, 55
Because there are 51 entries (an odd number), the median is the middle, or 26th entry. So, the median is 30 points.

3. Order the data from smallest to largest.
17, 20, 21, 21, 24, 24, 27, 28, 29, 31, 31, 32, 34, 34, 43, 48
Because there are an even number of entries, the median is the mean of the two middle entries.

$\text{median} = \frac{28 + 29}{2} = \frac{57}{2} = 28.5$

The median points scored by the winning teams in the Super Bowls for the National Football League's 2001 through 2016 seasons is 28.5 points.

4. Look at the ordered data from Try It Yourself 1
14, 16, 16, 16, 17, 20, 20, 20, 21, 21, 21, 23, 23, 24, 24, 24, 24, 26, 27, 27, 27, 27, 27, 28, 29, 30, 31, 31, 31, 31, 32, 32, 33, 34, 34, 34, 34, 35, 35, 35, 37, 38, 38, 39, 42, 43, 46, 48, 49, 52, 55
The mode is the data entry that occurs with the greatest frequency.
The entry 27 occurs the most, so the mode is 27.

5. "some" occurs with the greatest frequency (578). The mode is "some".

6. $\bar{x} = \frac{\sum x}{n} = \frac{410}{19} \approx 21.6$

median = 21
mode = 20

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

7.

Final Grade	Points, x	Weight, w	$x \cdot w$
C	2	3	6
C	2	4	8
D	1	1	1
A	4	3	12
B	3	2	6
B	3	3	9
		$\sum w = 16$	$\sum(x \cdot w) = 42$

$$\bar{x} = \frac{\sum(x \cdot w)}{\sum w} = \frac{42}{16} \approx 2.6$$

The new weighted mean is about 2.6.

8.

Class	Midpoint, x	Frequency, f	$x \cdot f$
14-20	17	8	136
21-27	24	15	360
28-34	31	14	434
35-41	38	7	266
42-48	45	4	180
49-55	52	3	156
		$\sum f = 51 = n$	$\sum(x \cdot f) = 1532$

$$\bar{x} = \frac{\sum(x \cdot f)}{n} = \frac{1532}{51} \approx 30.0$$

This is very close to the mean found using the original data set.

2.3 EXERCISE SOLUTIONS

1. True

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

3. True

4. True

5. *Sample answer:* 1, 2, 2, 2, 3

6. *Sample answer:* 2, 4, 5, 5, 6, 8

7. *Sample answer:* 2, 5, 7, 9, 35

8. *Sample answer:* 1, 2, 3, 3, 3, 4, 5

9. The shape of the distribution is skewed right because the bars have a "tail" to the right.

10. The shape of the distribution is symmetric because a vertical line can be drawn down the middle, creating two halves that are approximately the same.

11. The shape of the distribution is uniform because the bars are approximately the same height.

12. The shape of the distribution is skewed left because the bars have a “tail” 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 and is skewed right due to the few vehicles that have much higher mileages than the majority of the vehicles.

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 approximately symmetric and the weights range from 80 to 160 pounds.

17. $\bar{x} = \frac{\sum x}{n} = \frac{209}{14} \approx 14.9$

$12\ 12\ 13\ 14\ 14\ 15\ \underbrace{15\ 15}\ 16\ 16\ 16\ 16\ 17\ 18$

$$\text{median} = \frac{15+15}{2} = 15$$

mode = 16 (occurs 4 times)

18. $\bar{x} = \frac{\sum x}{n} = \frac{1205}{7} \approx 172.1$

169 169 170 **172** 174 175 176

median = 172

mode = 169 (occurs 2 times)

The mode does not represent the center of the data because 169 is the smallest number in the data set.

19. $\bar{x} = \frac{\sum x}{n} = \frac{6316}{7} \approx 902.3$

650 662 709 **788** 803 1242 1462

median = 788

mode = none

The mode cannot be found because no data entry is repeated.

The mean does not represent the center of the data because it is influenced by the outliers of 1242 and 1462.

20. $\bar{x} = \frac{\sum x}{n} = \frac{523}{10} = 52.3$

$34\ 36\ 38\ 38\ \underbrace{59\ \ 60}\ 63\ 63\ 64\ 68$

$$\text{median} = \frac{59+60}{2} = 59.5$$

mode = 38, 63 (each occurs 2 times)

21. $\bar{x} = \frac{\sum x}{n} = \frac{697}{14} \approx 49.8$

$45\ 47\ 48\ 48\ 48\ 49\ \underbrace{50\ 51}\ 51\ 51\ 51\ 51\ 52\ 55$

$$\text{median} = \frac{50+51}{2} = 50.5$$

mode = 51 (occurs 5 times)

22. $\bar{x} = \frac{\sum x}{n} = \frac{2004}{10} = 200.4$

$154\ 171\ 173\ 181\ \underbrace{184\ 188}\ 203\ 235\ 240\ 275$

$$\text{median} = \frac{184+188}{2} = 186$$

mode = none; The mode cannot be found because no data entry is repeated.

23. $\bar{x} = \frac{\sum x}{n} = \frac{119}{16} \approx 7.4$

$1\ 2\ 2\ 3\ 3\ 5\ 6\ \underbrace{6\ 6}\ 8\ 10\ 10\ 10\ 11\ 17\ 19$

$$\text{median} = \frac{6+6}{2} = 6$$

mode = 6, 10 (both occur 3 times)

24. $\bar{x} = \frac{\sum x}{n} = \frac{1242}{21} \approx 59.1$

12, 18, 19, 26, 28, 31, 33, 40, 44, 45, <u>49</u>, 61, 63, 75, 80, 80, 89, 96, 103, 125, 125
The median is the middle value, 49.
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.

25. $\bar{x} = \frac{\sum x}{n} = \frac{100}{7} = 14.3$

6, 7, 8, <u>9</u>, 13, 15, 42
The median is the middle value, 9.
mode = none
The mode cannot be found because no data entry is repeated.
The mean does not represent the center of the data set because it is influenced by the outlier of 42.

26. $\bar{x} = \frac{\sum x}{n} = \frac{388}{33} \approx 11.8$

10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 11, 11, 11, 11, 12, 12, 12, 12, 12, 12, 12, 13, 13, 14, 14, 14, 14, 14, 14, 14, 16

The median is the middle value, 12.

mode = 10 (occurs 12 times)

27. $\bar{x}$ is not possible (nominal data)

median = not possible (nominal data)

mode = "Search and buy online"

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

28. $\bar{x}$ is not possible (nominal data)

median is not possible (nominal data)

mode = "Mental health", "Education"

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

29. $\bar{x}$ is not possible (nominal data)

median is not possible (nominal data)

mode = "Junior"

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

30. $\bar{x}$ is not possible (nominal data)

median is not possible (nominal data)

mode = "Yes, since 2014 or earlier"

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

31. $\bar{x} = \frac{\sum x}{n} = \frac{817}{28} \approx 29.2$

$5\ 8\ 10\ 11\ 13\ 16\ 21\ 23\ 23\ 23\ 26\ 27\ 27\ \underbrace{30\ 31}\ 32\ 34\ 34\ 34\ 35\ 37\ 38\ 43\ 44\ 45\ 46\ 49\ 52$

$$\text{median} = \frac{30+31}{2} = 30.5$$

mode = 23, 34 (both occur 3 times each)

32. $\bar{x} = \frac{\sum x}{n} = \frac{29.9}{12} \approx 2.49$

$0.8\ 1.5\ 1.6\ 1.8\ 2.1\ \underbrace{2.3\ 2.4}\ 2.5\ 3.0\ 3.9\ 4.0\ 4.0$

$$\text{median} = \frac{2.3+2.4}{2} = 2.35$$

mode = 4.0 (occurs 2 times)

The mode does not represent the center of the data set because it is the largest entry in the data set.

33. $\bar{x} = \frac{\sum x}{n} = \frac{292}{15} \approx 19.5$

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

median = 20

mode = 15 (occurs 3 times)

34. $\bar{x} = \dfrac{\sum x}{n} = \dfrac{3160}{16} \approx 197.5$

$$100\ 160\ 160\ 160\ 160\ 160\ 180\ \underbrace{200\ 200}\ 200\ 200\ 220\ 240\ 260\ 280\ 280$$

$$\text{median} = \frac{200+200}{2} = 200$$

mode = 160 (occurs 5 times)

35. Cluster around 275 – 425

36. Cluster around 450 – 1050, gap between 1950 and 2850, outlier at 3000

37. Mode, because the data are at the nominal level of measurement.

38. Mean, because the distribution is symmetric and there are no outliers.

39. Mean, because the distribution is symmetric and there are no outliers.

40. Median, because there is an outlier.

41.

Source	**Score, x**	**Weight, w**	$x \cdot w$
Homework	85	0.05	4.25
Quizzes	80	0.35	28
Project/Speech	100	0.35	35
Final exam	93	0.25	23.25
		$\sum w = 1$	$\sum (x \cdot w) = 90.5$

$\bar{x} = \dfrac{\sum (x \cdot w)}{\sum w} = \dfrac{90.5}{1} = 90.5$

42.

Source	**Score, x**	**Weight, w**	$x \cdot w$
Quizzes	100	20%	20
Midterm exam	89	30%	26.7
Student lecture	100	10%	10
Final exam	92	40%	36.8
		$\sum w = 100\%$	$\sum (x \cdot w) = 93.5$

$\bar{x} = \dfrac{\sum (x \cdot w)}{\sum w} = \dfrac{93.5}{1} = 93.5$

43.

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

$$\bar{x} = \frac{\sum(x \cdot w)}{\sum w} = \frac{18{,}382}{30} \approx \$612.73$$

44.

Balance, x	Days, w	$x \cdot w$
\$115.63	12	1387.56
\$637.19	6	3823.14
\$1225.06	7	8575.42
\$0	2	0
\$34.88	4	139.52
	$\sum w = 31$	$\sum(x \cdot w) = 13{,}925.64$

$$\bar{x} = \frac{\sum(x \cdot w)}{\sum w} = \frac{13{,}925.64}{31} \approx \$449.21$$

45.

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

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

46.

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

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

47.

Source	Score, x	Weight, w	$x \cdot w$
Homework	85	0.05	4.25
Quizzes	80	0.35	28
Project/Speech	100	0.35	35
Final exam	85	0.25	21.25
		$\sum w = 1$	$\sum(x \cdot w) = 88.5$

$$\bar{x} = \frac{\sum(x \cdot w)}{\sum w} = \frac{88.5}{1} = 88.5$$

48.

Grade	Points, x	Credits, w	$x \cdot w$
A	4	4	16
A	4	3	12
B	3	3	9
C	2	3	6
D	1	2	2
		$\sum w = 15$	$\sum (x \cdot w) = 45$

$$\bar{x} = \frac{\sum (x \cdot w)}{\sum w} = \frac{45}{15} = 3$$

49.

Class	Midpoint, x	Frequency, f	$x \cdot f$
29-33	31	11	341
34-38	36	12	432
39-43	41	2	82
44-48	46	5	230
		$n = 30$	$\sum (x \cdot f) = 1085$

$$\bar{x} = \frac{\sum (x \cdot f)}{n} = \frac{1085}{30} \approx 36.2 \text{ miles per gallon}$$

50.

Class	Midpoint, x	Frequency, f	$x \cdot f$
22-27	24.5	16	392
28-33	30.5	2	61
34-39	36.5	2	73
40-45	42.5	4	170
		$n = 24$	$\sum (x \cdot f) = 696$

$$\bar{x} = \frac{\sum (x \cdot f)}{n} = \frac{696}{24} = 29 \text{ miles per gallon}$$

51.

Class	Midpoint, x	Frequency, f	$x \cdot f$
0-9	4.5	78	351
10-19	14.5	97	1406.5
20-29	24.5	54	1323
30-39	34.5	63	2173.5
40-49	44.5	69	3070.5
50-59	54.5	86	4687
60-69	64.5	73	4708.5
70-79	74.5	53	3948.5
80-89	84.5	43	3633.5
90-99	94.5	15	1417.5
		$n = 631$	$\sum (x \cdot f) = 26{,}719.5$

$$\bar{x} = \frac{\sum (x \cdot f)}{n} = \frac{26{,}719.5}{631} \approx 42.3 \text{ years old}$$

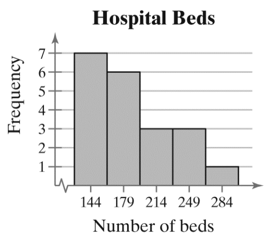
52.

Class	Midpoint, x	Frequency, f	$x \cdot f$
0-49	24.5	41	1004.5
50-99	74.5	9	670.5
100-149	124.5	6	747
150-199	174.5	2	349
200-249	224.5	1	224.5
250-299	274.5	2	549
300-349	324.5	0	0
350-399	374.5	1	374.5
400-449	424.5	2	849
		$n = 64$	$\sum(x \cdot f) = 4768$

$$\bar{x} = \frac{\sum(x \cdot f)}{n} = \frac{4768}{64} = 74.5 \text{ years old}$$

53. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{297 - 127}{5} = 34 \Rightarrow 35$

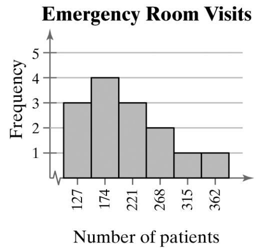
Class	Frequency, f	Midpoint
127–161	7	144
162–196	6	179
197–231	3	214
232–266	3	249
267–301	1	284



Positively skewed

54. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{382 - 104}{6} \approx 46.3 \Rightarrow 47$

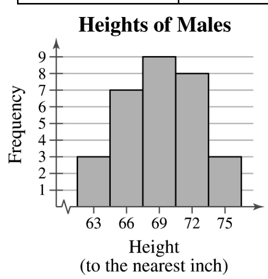
Class	Frequency, f	Midpoint
104–150	3	127
151–197	4	174
198–244	3	221
245–291	2	268
292–338	1	315
339–385	1	362



Positively skewed

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

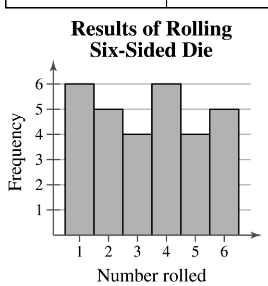
Class	Midpoint	Frequency, f
62-64	63	3
65-67	66	7
68-70	69	9
71-73	72	8
74-76	75	3
		$\sum f = 30$



Shape: Symmetric

56. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{6-1}{6} = 0.8333 \Rightarrow 1$

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



Shape: Uniform

57. (a) $\bar{x} = \frac{\sum x}{n} = \frac{9109}{6} \approx 1518.2$

$1502 \; 1511 \; \underbrace{1516 \; 1525} \; 1526 \; 1529$

$$\text{median} = \frac{1516 + 1525}{2} = 1520.5$$

(b) $\bar{x} = \frac{\sum x}{n} = \frac{9127}{6} = 1521.2$

$1511 \; 1516 \; \underbrace{1520 \; 1525} \; 1526 \; 1529$

$$\text{median} = \frac{1520 + 1525}{2} = 1522.5$$

(c) The mean was affected more.

58. (a) $\bar{x} = \frac{\sum x}{n} = \frac{841.7}{18} \approx 46.76$

$9.2 \; 9.3 \; 10.9 \; 12.5 \; 15 \; 15.5 \; 17.4 \; 17.7 \; \underbrace{21.7 \; 23.3} \; 28 \; 28.3 \; 30.4 \; 30.9 \; 60.7 \; 68.9 \; 74.8 \; 367.2$

$$\text{median} = \frac{21.7 + 23.3}{2} = 22.5$$

(b) $\bar{x} = \frac{\sum x}{n} = \frac{474.5}{17} \approx 27.91$

9.2, 9.3, 10.9, 12.5, 15, 15.5, 17.4, 17.7, <u>21.7</u>, 23.3, 28, 28.3, 30.4, 30.9, 60.7, 68.9, 74.8
The median is the middle value, 21.7.
The mean was affected more.

(c) $\bar{x} = \frac{\sum x}{n} = \frac{849}{19} = 44.68$

7.3, 9.2, 9.3, 10.9, 12.5, 15, 15.5, 17.4, 17.7, <u>21.7</u>, 23.3, 28, 28.3, 30.4, 30.9, 60.7, 68.9, 74.8, 367.2
The median is the middle value, 21.7.
The mean was affected more.

59. The data are skewed right.
A = mode, because it is the data entry that occurred most often.
B = median, because the median is to the left of the mean in a skewed right distribution.
C = mean, because the mean is to the right of the median in a skewed right distribution.

60. The data are skewed left.
A = mean, because the mean is to the left of the median in a skewed left distribution.
B = median, because the median is to the right of the mean in a skewed left distribution.
C = mode, because it is the data entry that occurred most often.

61. Increase one of the three-credit B classes to an A. The three-credit class is weighted more than the two-credit classes, so it will have a greater effect on the grade point average.

62. (a) $\bar{x} = \frac{\sum x}{n} = \frac{3222}{9} = 358$

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

median = 375

(b) $\bar{x} = \frac{\sum x}{n} = \frac{9666}{9} = 1074$

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

median = 1125

(c) The mean and median in part (b) are three times the mean and median in part (a).

(d) If you multiply the mean and median of the original data set by 36, you will get the mean and median of the data set in inches.

63. Car A

$\bar{x} = \frac{\sum x}{n} = \frac{152}{5} = 30.4$

28 28 **30** 32 34

median = 30

mode = 28 (occurs 2 times)

Car B

$\bar{x} = \frac{\sum x}{n} = \frac{151}{5} = 30.2$

29 29 **31** 31 31

median = 31

mode = 31 (occurs 3 times)

Car C

$\bar{x} = \frac{\sum x}{n} = \frac{151}{5} = 30.2$

28 29 **30** 32 32

median = 30

mode = 32 (occurs 2 times)

(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.

64. Car A: Midrange $= \frac{34+28}{2} = 31$

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

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

Car A because the midrange is the largest.

65. (a) $\bar{x} = \dfrac{\sum x}{n} = \dfrac{1477}{30} \approx 49.2$

11 13 22 28 36 36 36 37 37 37 38 41 43 44 $\underbrace{46\ 47}$ 51 51 51 53 61 62 63 64

72 72 74 76 85 90

$$\text{median} = \frac{46+47}{2} = 46.5$$

(b) Key: $3|6 = 36$

```
1|1  3
2|2  8                median
3|6  6  6  7  7  7  8
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
```

(c) The distribution is positively skewed.

66. (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.

$\bar{x} = \dfrac{\sum x}{n} = \dfrac{1180}{24} \approx 49.2$

10% trimmed mean ≈ 49.2

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

median = 46.5

mode = 36, 37, 51

$\text{midrange} = \dfrac{90+11}{2} = 50.5$

(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

1. Min = 23, or \$23,000 and Max = 58, or \$58,000,
Range = Max – Min = 58 – 23 = 35, or \$35,000
The range of the starting salaries for Corporation B is 35, or \$35,000. This is much larger than the range for Corporation A.

2. $\mu = 41.5$, or \$41,500

Salary, *x*	$x - \mu$	$(x - \mu)^2$
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
$\sum x = 415$	$\sum(x-\mu) = 0$	$\sum(x-\mu)^2 = 1102.5$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} = \frac{1102.5}{10} \approx 110.3$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1102.5}{10}} = 10.5, \text{ or } \$10,500$$

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

3 $\bar{x} = \frac{\sum x}{n} = \frac{316}{8} = 39.5$

Time, *x*	$x - \bar{x}$	$(x - \bar{x})^2$
43	3.5	12.25
57	17.5	306.25
18	-21.5	462.25
45	5.5	30.25
47	7.5	56.25
33	-6.5	42.25
49	9.5	90.25
24	-15.5	240.25

$\sum x = 316$	$\sum(x-\mu) = 0$	$\sum(x-\mu)^2 = 1240$

$$SS_x = \sum\left(x-\bar{x}\right)^2 = 1240$$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{1240}{7} \approx 177.1$$

$$s = \sqrt{s^2} = \sqrt{\frac{1240}{7}} \approx 13.3$$

4. Enter the data in a computer or a calculator.
$\bar{x} \approx 19.8, \quad s \approx 7.8$

5. *Sample answer:* 7, 7, 7, 7, 7, 13, 13, 13, 13, 13

Salary, x	$x-\mu$	$(x-\mu)^2$
7	–3	9
7	–3	9
7	–3	9
7	–3	9
7	–3	9
13	3	9
13	3	9
13	3	9
13	3	9
13	3	9
$\sum x = 100$	$\sum(x-\mu) = 0$	$\sum(x-\mu)^2 = 90$

$$\mu = \frac{\sum x}{N} = \frac{100}{10} = 10$$

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

6. 67.1 – 64.2 = 2.9 = 1 standard deviation
Because 67.1 is one standard deviation above the mean height, the percent of heights between 64.2 inches and 67.1 inches is 34.13%.
Approximately 34.13% of women ages 20-29 are between 64.2 and 67.1 inches tall.

7. 39.3 – 2(23.5) = –7.7
Because –7.7 does not make sense for an age, use 0.
39.3 + 2(23.5) = 86.3

$$1 - \frac{1}{k^2} = 1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = 0.75$$

At least 75% of the data lie within 2 standard deviations of the mean. At least 75% of the population of Iowa is between 0 and 86.3 years old. Because 80 < 86.3, and age of 80 lies within two standard deviations of the mean. So, the age is not unusual.

8.

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

$$\bar{x} = \frac{\sum xf}{n} = \frac{85}{50} = 1.7$$

$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
–1.7	2.89	28.90
–0.7	0.49	9.31
0.3	0.09	0.63
1.3	1.69	11.83
2.3	5.29	26.45
3.3	10.89	10.89
4.3	18.49	18.49
		$\sum(x - \bar{x})^2 f = 106.5$

$$s = \sqrt{\frac{\sum(x - \bar{x})^2 f}{n - 1}} = \sqrt{\frac{106.5}{49}} \approx 1.5$$

9.

Class	x	f	xf
1-99	49.5	380	18,810
100-199	149.5	230	34,385
200-299	249.5	210	52,395
300-399	349.5	50	17,475
400-499	449.5	60	26,970
500+	650	70	45,500
		$n =$ 1000	$\sum xf = 195{,}535$

$$\bar{x} = \frac{\sum xf}{n} = \frac{195{,}535}{1000} \approx 195.5$$

$x-\overline{x}$	$\left(x-\overline{x}\right)^2$	$\left(x-\overline{x}\right)^2 f$
–146.0	21,316	8,100,080
–46.0	2116	486,680
54.0	2916	612,360
154.0	23,716	1,185,800
254.0	64,516	3,870,960
454.5	206,570.25	14,459,917.5
		$\sum\left(x-\overline{x}\right)^2 f = 28,715,797.5$

$$s=\sqrt{\frac{\sum\left(x-\overline{x}\right)^2 f}{n-1}}=\sqrt{\frac{28,715,797.5}{999}}\approx 169.5$$

10. Los Angeles: $\overline{x}\approx 36.88$, $s\approx 17.39$

Dallas: $\overline{x}\approx 19.8$, $s\approx 7.8$

Los Angeles: $CV=\frac{s}{\overline{x}}\cdot 100\%=\frac{17.4}{36.9}\cdot 100\%\approx 47.2\%$

Dallas: $CV=\frac{s}{\overline{x}}\cdot 100\%=\frac{7.8}{19.8}\cdot 100\%\approx 39.4\%$

The office rental rates are more variable in Los Angeles than in Dallas.

2.4 EXERCISE SOLUTIONS

1. 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.

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

3. The units of variance are squared. Its units are meaningless (example: dollars^2). The units of standard deviation are the same as the data.

4. The standard deviation is the positive square root of the variance. The standard deviation and variance can never be negative because squared deviations can never be negative.

5. 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.

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

7. Similarity: Both estimate proportions of the data contained within k standard deviations of the mean. Difference: The Empirical Rule assumes the distribution is approximately symmetric and bell-shaped. Chebychev's Theorem makes no such assumption.

8. You must know that the distribution is approximately symmetric and bell-shaped.

9. Range = Max – Min = 75 – 40 = 35; Approximately 35, or $35,000

10. Range = Max – Min = 98 – 74 = 24

11. (a) Range = Max – Min = 38.5 – 20.7 = 17.8

(b) Range = Max – Min = 60.5 – 20.7 = 39.8

12. Changing the maximum value of the data set greatly affects the range.

13. Range = Max – Min = 14 – 10 = 4

$$\mu = \frac{\sum x}{N} = \frac{123}{11} \approx 11.2$$

x	$x-\mu$	$(x-\mu)^2$
14	2.8	7.84
13	1.8	3.24
13	1.8	3.24
12	0.8	0.64
11	–0.2	0.04
10	–1.2	1.44
10	–1.2	1.44
10	–1.2	1.44
10	–1.2	1.44
10	–1.2	1.44
10	–1.2	1.44
$\sum x = 123$	$\sum(x-\mu) \approx 0$	$\sum(x-\mu)^2 = 23.64$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} = \frac{23.64}{11} \approx 2.1$$

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}} = \sqrt{\frac{23.64}{11}} \approx 1.5$$

14. Range = Max – Min = 7870 – 0.09 = 7869.91

$$\mu = \frac{\sum x}{N} = \frac{22{,}511.5}{10} = 2251.15$$

x	$x-\mu$	$(x-\mu)^2$
1.4	−2249.75	5,061,375
2330	78.85	6217
2700	448.85	201,466
7870	5618.85	31,571,475
1500	−751.15	564,226
970	−1281.15	1,641,345
900	−1351.15	1,825,606
1740	−511.15	261,274
4500	2248.85	5,057,326
.09	−2251.06	5,067,271
$\sum x = 22{,}511.5$	$\sum(x-\mu) \approx 0$	$\sum(x-\mu)^2 = 51{,}257{,}584$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} = \frac{51{,}257{,}584}{10} \approx 5{,}125{,}758.4$$

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}} = \sqrt{\frac{5{,}125{,}758.4}{10}} \approx 2264.0$$

15. Range = Max – Min = 23 – 17 = 6

$$\bar{x} = \frac{\sum x}{n} = \frac{380}{20} = 19$$

x	$x-\bar{x}$	$(x-\bar{x})^2$
19	0	0
20	1	1
17	−2	4
19	0	0
17	−2	4
21	2	4
23	4	16
21	2	4
17	−2	4
17	−2	4
19	0	0
19	0	0
17	−2	4
20	1	1
23	4	16
18	−1	1
18	−1	1
18	−1	1
18	−1	1
19	0	0
$\sum x = 380$	$\sum(x-\bar{x}) = 0$	$\sum(x-\bar{x})^2 = 66$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{66}{20-1} \approx 3.5$$

$$s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}} = \sqrt{\frac{66}{19}} \approx 1.9$$

16. Range = Max – Min = 299 – 264 = 35

$$\bar{x} = \frac{\sum x}{n} = \frac{5902}{21} \approx 281.0$$

x	$x-\bar{x}$	$\left(x-\bar{x}\right)^2$
277	–4	16
277	–4	16
267	–14	196
291	10	100
282	1	1
281	0	0
295	14	196
279	–2	4
286	5	25
280	–1	1
296	15	225
269	–12	144
268	–13	169
285	4	16
264	–17	289
278	–3	9
269	–12	144
299	18	324
291	10	100
293	12	144
275	–6	36
$\sum x = 5902$	$\sum\left(x-\bar{x}\right) \approx 0$	$\sum\left(x-\bar{x}\right)^2 = 2155$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{2155}{21-1} \approx 107.7$$

$$s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}} = \sqrt{\frac{2155}{20}} \approx 10.4$$

17. The data set in (a) has a standard deviation of 2.4 and the data set in (b) has a standard deviation of 5 because the data in (b) have more variability.

18. The data set in (a) has a standard deviation of 24 and the data set in (b) has a standard deviation of 16 because the data in (a) have more variability.

19. Company B; An offer of \$43,000 is two standard deviations from the mean of Company A's starting salaries, which makes it unlikely. The same offer is within one standard deviation of the mean of Company B's starting salaries, which makes the offer likely.

20. Company C; An offer of \$62,000 is three standard deviations from the mean of Company D's starting salaries, which makes it unlikely. The same offer is within two standard deviations of the mean of Company C's starting salaries, which makes the offer somewhat likely.

21. (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.

(c) Estimates will vary; (i) $s \approx 1.1$; (ii) $s \approx 1.3$; (iii) $s \approx 0.8$

22. (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.

(c) Estimates will vary; (i) $s \approx 1.6$; (ii) $s \approx 2.9$; (iii) $s \approx 0.8$

23. (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.

(c) Estimates will vary; (i) $s \approx 9.6$; (ii) $s \approx 9.0$; (iii) $s \approx 5.1$

24. (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 modes and standard deviations.

(c) Estimates will vary; (i) $s \approx 1.5$; (ii) $s \approx 1.8$; (iii) $s \approx 2.5$

25. *Sample answer:* 3,3,3,7,7,7

26. *Sample answer:* 3,3,3,3,9,9,9,9

27. *Sample answer:* 9,9,9,9,9,9,9

28. *Sample answer:* 5,5,5,9,9,9

29. $(63,\ 71) \to (67-1(4),\ 67+1(4)) \to (\bar{x}-s,\ \bar{x}+s)$
68% of the vehicles have speeds between 63 and 71 mph.

30. 95% of the data falls between $\bar{x}-2s$ and $\bar{x}+2s$.
$\bar{x}-2s = 70-2(8) = 54$
$\bar{x}+2s = 70+2(8) = 86$
95% of the households have monthly utility bills between \$54 and \$86.

31. (a) $n = 75$; 68%(75) = (0.68)(75) ≈ 51 vehicles have speeds between 63 and 71 mph.

(b) $n = 25$; 68%(25) = (0.68)(25) ≈ 17 vehicles have speeds between 63 and 71 mph.

32. (a) $n = 40$; 95%(40) = (0.95)(40) ≈ 38 households have monthly utility bills between \$54 and \$86.

(b) $n = 20$; 95%(20) = (0.95)(20) ≈ 19 households have monthly utility bills between \$54 and \$86.

33. 78, 76, and 82 are unusual; 82 is very unusual because it is more than 3 standard deviations from the mean.

34. \$52 and \$98 are unusual; \$98 is very unusual because it is more than 3 standard deviations from the mean.

35. $(\bar{x}-2s,\ \bar{x}+2s) \to (0,\ 4)$ are 2 standard deviations from the mean.
$1-\frac{1}{k^2} = 1-\frac{1}{(2)^2} = 1-\frac{1}{4} = 0.75 \Rightarrow$ At least 75% of the eruption times lie between 0 and 4.
If n = 40, at least (0.75)(40) = 30 households have between 0 and 4 pets.

36. $(\bar{x}-2s,\ \bar{x}+2s) \to (16.18,\ 186.94)$ are 2 standard deviations from the mean.
$1-\frac{1}{k^2} = 1-\frac{1}{(2)^2} = 1-\frac{1}{4} = 0.75 \Rightarrow$ At least 75% of the eruption times lie between 16.18 and 186.94 minutes.
If n = 100, at least (0.75)(100) = 75 eruptions will lie between 16.18 and 186.94 minutes.

37. $(\bar{x}-4s,\ \bar{x}+4s) \rightarrow (70,\ 94)$ are 4 standard deviations from the mean.

$$1-\frac{1}{k^2}=1-\frac{1}{4^2}=1-\frac{1}{16}=0.9375$$

At least 93.75% of the exam scores are from 70 to 94.

38. $1-\frac{1}{k^2}=1-\frac{1}{(2)^2}=1-\frac{1}{4}=0.75$

At least 75% of the runs lie within 2 standard deviations of the mean.

$(\bar{x}-2s,\ \bar{x}+2s) \rightarrow (-2.86,\ 10.58) \rightarrow (0,\ 10)$

At least 75% of the runs per game scored by the Chicago Cubs during the 2016 World Series are from 0 to 10 (note that -2.86 and 10.58 do not make sense in the context of the data).

39.

x	f	xf
0	3	0
1	4	4
2	3	6
3	9	27
4	3	12
5	3	15
6	8	48
7	5	35
8	6	48
9	6	54
	$\Sigma=50$	$\Sigma=249$

$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
−5	25	75
−4	16	64
−3	9	27
−2	4	36
−1	1	3
0	0	0
1	1	8
2	4	20
3	9	54
4	16	96
		$\Sigma=383$

$$\bar{x}=\frac{\sum xf}{n}=\frac{249}{50}\approx 5.0$$

$$s=\sqrt{\frac{\sum(x-\bar{x})^2 f}{n-1}}=\sqrt{\frac{383}{49}}\approx 2.8$$

40.

x	f	xf
0	30	0
1	20	20
	$\Sigma=50$	$\Sigma=20$

$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
−0.4	0.16	4.8
0.6	0.36	7.2
		$\Sigma=12$

$$\bar{x}=\frac{\sum xf}{n}=\frac{20}{50}\approx 0.4$$

$$s=\sqrt{\frac{\sum(x-\bar{x})^2 f}{n-1}}=\sqrt{\frac{12}{49}}\approx 0.5$$

41.

Class	x	f	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
15,000–17,499	16,249.5	9	146,245.5	−4759.62	22,653,982.54	203,885,842.9
17,500–19,999	18,749.5	10	187,495	−2259.62	5,105,882.54	51,058,825.4
20,000–22,499	21,249.5	16	339,992	240.38	57,782.54	924,520.64
22,500–24,999	23,749.5	11	261,244.5	2740.38	7,509,682.54	82,606,507.94
25,000 or more	26,249.5	6	157,497	5240.38	27,461,582.54	164,769,495.20
		$n=52$	$\Sigma xf = 1,092,474$		$\Sigma(x-\bar{x})^2 f = 503,245,192.1$	

$$\bar{x} = \frac{\sum xf}{n} = \frac{1,092,474}{52} \approx \$21,009.12$$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2 f}{n-1}} = \sqrt{\frac{503,245,192.1}{51}} \approx \$3141.27$$

42.

Class	**Midpoint, x**	f	xf
0-4	2	5	10
5-9	7	12	84
10-14	12	24	288
15-19	17	17	289
20-24	22	16	352
25-29	27	11	297
30+	32	5	160
		$n=90$	$\sum xf = 1480$

$$\bar{x} = \frac{\sum xf}{n} = \frac{1480}{90} \approx 16.4$$

$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
−14.44	208.5136	1042.5680
−9.44	89.1136	1069.3632
−4.44	19.7136	473.1264
0.56	0.3136	5.3312
5.56	30.9136	494.6176
10.56	111.5136	1226.6496
15.56	242.1136	1210.5680
		$\sum(x-\bar{x})^2 f = 5522.2240$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2 f}{n-1}} = \sqrt{\frac{5522.2240}{89}} \approx 7.9$$

43.

x	f	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
1	2	2	−1.9	3.61	7.22
2	18	36	−0.9	0.81	14.58
3	24	72	0.1	0.01	0.24
4	16	64	1.1	1.21	19.36
	$n=60$	$\Sigma xf=174$		$\Sigma(x-\bar{x})^2 f=41.4$	

$$\bar{x}=\frac{\sum xf}{n}=\frac{174}{60}\approx 2.9$$

$$s=\sqrt{\frac{\sum\left(x-\bar{x}\right)^2 f}{n-1}}=\sqrt{\frac{41.4}{59}}\approx 0.8$$

44.

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
	$n=50$	$\sum xf=5725$

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

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

$$s=\sqrt{\frac{\sum\left(x-\bar{x}\right)^2 f}{n-1}}=\sqrt{\frac{16,456}{49}}\approx 18.33$$

45. Denver: $\bar{x}=\frac{\sum x}{n}=\frac{552.3}{12}\approx 46.0$

$$s^2=\frac{\sum\left(x-\bar{x}\right)^2}{n-1}=\frac{220.89}{11}\approx 20.08$$

$$s=\sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}}=\sqrt{\frac{220.89}{11}}\approx 4.48$$

$$CV=\frac{s}{\bar{x}}\cdot 100\%=\frac{4.48}{46.0}\cdot 100\%\approx 9.7\%$$

Los Angeles: $\bar{x} = \dfrac{\sum x}{n} = \dfrac{634.5}{12} \approx 52.9$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{239.97}{11} \approx 21.82$$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{239.97}{11}} \approx 4.67$$

$$CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{4.67}{52.9} \cdot 100\% \approx 8.8\%$$

Salaries for entry level architects are more variable in Denver than in Los Angeles.

46. Raleigh: $\bar{x} = \dfrac{\sum x}{n} = \dfrac{551.0}{9} \approx 61.22$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{459.02}{9-1} \approx 57.38$$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{459.02}{8}} \approx 7.57$$

$$CV = \frac{s}{\bar{x}} = \frac{7.57}{61.22} \cdot 100\% \approx 12.4\%$$

Wichita: $\bar{x} = \dfrac{\sum x}{n} = \dfrac{577.3}{9} \approx 64.14$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{824.98}{9-1} \approx 103.12$$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{824.98}{8}} \approx 10.15$$

$$CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{10.15}{64.14} \cdot 100\% \approx 15.8\%$$

Salaries for entry level software engineers are more variable in Wichita than in Raleigh.

47. Ages: $\mu = \dfrac{\sum x}{N} = \dfrac{491}{22} \approx 22.32$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} = \frac{194.77}{22} \approx 8.85$$

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{N}} = \sqrt{\frac{194.77}{22}} \approx 2.98$$

$$CV = \frac{\sigma}{\mu} \cdot 100\% = \frac{2.98}{22.32} \cdot 100\% \approx 13.3\%$$

Heights: $\mu = \dfrac{\sum x}{N} = \dfrac{1546}{22} \approx 70.27$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} = \frac{134.36}{22} \approx 6.11$$

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{N}} = \sqrt{\frac{134.36}{22}} \approx 2.47$$

$$CV = \frac{\sigma}{\mu} \cdot 100\% = \frac{2.47}{70.27} \cdot 100\% \approx 3.5\%$$

Ages are more variable than heights for all members of the 2016 Women's U.S. Olympic swimming team.

48. Ages: $\mu = \frac{\sum x}{N} = \frac{263}{10} = 26.3$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} = \frac{112.1}{10} \approx 11.21$$

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{N}} = \sqrt{\frac{112.1}{10}} \approx 3.35$$

$$CV = \frac{\sigma}{\mu} \cdot 100\% = \frac{3.35}{26.3} \cdot 100\% \approx 12.7\%$$

Weights: $\mu = \frac{\sum x}{N} = \frac{853}{10} = 85.3$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} = \frac{5850.1}{10} \approx 585.01$$

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{N}} = \sqrt{\frac{5850.1}{10}} \approx 24.19$$

$$CV = \frac{\sigma}{\mu} \cdot 100\% = \frac{24.19}{85.3} \cdot 100\% \approx 28.4\%$$

Weight classes are more variable than ages for all members of the 2016 Men's U.S. Olympic wrestling team.

49. Male: $\bar{x} = \frac{\sum x}{n} = \frac{8760}{8} = 1095$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{359{,}400}{8-1} \approx 51{,}342.86$$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{359{,}400}{7}} \approx 226.6$$

$$CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{226.6}{1095} \cdot 100\% \approx 20.7\%$$

Female: $\bar{x} = \frac{\sum x}{n} = \frac{9120}{8} = 1140$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{282{,}000}{8-1} \approx 40{,}285.71$$

$$s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}} = \sqrt{\frac{282{,}000}{7}} = 200.7$$

$$CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{200.7}{1140} \cdot 100\% \approx 17.6\%$$

SAT scores are more variable for males than for females.

50. Male: $\bar{x} = \frac{\sum x}{n} = \frac{32}{10} = 3.2$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{4.76}{10-1} \approx 0.529$$

$$s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}} = \sqrt{\frac{4.76}{9}} \approx 0.727$$

$$CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{0.727}{3.2} \cdot 100\% \approx 22.7\%$$

Female: $\bar{x} = \frac{\sum x}{n} = \frac{32}{10} = 3.2$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{5.9}{10-1} \approx 0.656$$

$$s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}} = \sqrt{\frac{5.9}{9}} = 0.810$$

$$CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{0.810}{3.2} \cdot 100\% \approx 25.3\%$$

Grade point averages are more variable for females than for males.

51. (a) Answers will vary.

(b) Ages of students

$\sum x = 380$; $\sum x^2 = 7286$

$$s = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}} = \sqrt{\frac{7286 - \frac{(380)^2}{20}}{20-1}} = \sqrt{\frac{66}{19}} \approx 1.9$$

(c) The answer is the same as from Exercise 15.

52. (a) Ages of students $\bar{x} = 19$

$\sum\left|x-\bar{x}\right| = 28$

$$MAD = \frac{\sum|x-\bar{x}|}{n} = \frac{28}{20} = 1.4$$

From Exercise 15, the sample standard deviation is 1.9.
The mean absolute deviation is less than the sample standard deviation.

(b) Days: $\bar{x} \approx 281.0$

$\sum|x-\bar{x}| = 177$

$$MAD = \frac{\sum|x-\bar{x}|}{n} = \frac{177}{21} \approx 8.4$$

From Exercise 16, the sample standard deviation is 10.4.
The mean absolute deviation is less than the sample standard deviation.

53. (a) $\bar{x} \approx 42.1$; $s \approx 5.6$

(b) $\bar{x} \approx 44.3$; $s \approx 5.9$

(c) 3.5, 3, 3, 4, 4, 2.75, 4.25, 3.25, 3.25, 3.5, 3.25, 3.75, 3.5, 4.17

$\bar{x} \approx 3.5$; $s \approx 0.47$

(d) When each entry is multiplied by a constant k, the new sample mean is $k \cdot x$, and the new sample standard deviation is $k \cdot s$.

54. (a) $\bar{x} \approx 41.2$, $s \approx 6.0$

(b) $\bar{x} \approx 42.2$, $s \approx 6.0$

(c) $\bar{x} \approx 39.2$, $s \approx 6.0$

(d) Adding a constant k to, or subtracting it from, each entry makes the new sample mean $\bar{x} + k$, or $\bar{x} - k$, with the sample standard deviation being unaffected.

55. (a) $P = \dfrac{3(\bar{x} - \text{median})}{s} = \dfrac{3(17-19)}{2.3} \approx -2.61;$ The data are skewed left.

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

(c) $P = \dfrac{3(\bar{x} - \text{median})}{s} = \dfrac{3(9.2-9.2)}{1.8} = 0;$ The data are symmetric.

(d) $P = \dfrac{3(\bar{x} - \text{median})}{s} = \dfrac{3(42-40)}{6.0} = 1;$ The data are skewed right.

(e) $P = \dfrac{3(\bar{x} - \text{median})}{s} = \dfrac{3(155 - 175)}{20.0} = -3$; The data are skewed left.

56. $1 - \dfrac{1}{k^2} = 0.99 \Rightarrow 1 - 0.99 = \dfrac{1}{k^2} \Rightarrow k^2 = \dfrac{1}{0.01} \Rightarrow k = \sqrt{\dfrac{1}{0.01}} = 10$

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

2.5 MEASURES OF POSITION

2.5 TRY IT YOURSELF SOLUTIONS

1. Order the data from least to greatest. The median (or Q_2) is 30. This was also found in Section 2.3, Try It Yourself 2.

The first quartile is the median of the data entries to the left of Q_2 and the third quartile is the median of the data entries to the right of Q_2.

$Q_1 = 23,\ Q_2 = 30,\ Q_3 = 35$

About one-quarter of the winning scores were 23 points or less, about one-half were 30 points or less, and about three-quarters were 35 points or less.

2. Enter data

$Q_1 = 23.5,\ Q_2 = 30,\ Q_3 = 41$

About one-quarter of these universities charge tuition of \$23,500 or less; about one-half charge \$30,000 or less; and about three-quarters charge \$41,000 or less.

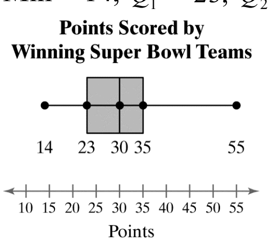
3. $Q_1 = 23,\ Q_3 = 35$

$\text{IQR} = Q_3 - Q_1 = 35 - 23 = 12$

$Q_1 - 1.5(\text{IQR}) = 23 - 1.5(12) = 5$; $Q_3 + 1.5(\text{IQR}) = 35 + 1.5(12) = 53$

The score 55 is greater than $Q_3 + 1.5(\text{IQR})$. So, 55 is an outlier.

4. Min = 14, $Q_1 = 23,\ Q_2 = 30,\ Q_3 = 35$, Max = 55



About 50% of the winning scores were between 23 and 35 points. About 25% of the winning scores were less than 23 points. About 25% of the winning scores were greater than 35 points.

5. The 10^{th} percentile is 19.5.

About 10% of the winning scores were 19.5 or less.

6 17,18,19,20,20,23,24,26,29,29,29,30,30,34,35,36,38,39,39,43,44,44,44,45,45

7 data entries are less than 26

$$\text{Percentile of } 26 = \frac{\text{number of data entries less than 26}}{\text{total number of data entries}} \cdot 100 = \frac{7}{25} \cdot 100 = 28^{\text{th}} \text{ percentile}$$

The tuition cost of $26,000 is greater than 28% of the other tuition costs.

7. $\mu = 70,\ \sigma = 8$

$$x = 60:\ z = \frac{x-\mu}{\sigma} = \frac{60-70}{8} = -1.25$$

$$x = 71:\ z = \frac{x-\mu}{\sigma} = \frac{71-70}{8} = 0.125$$

$$x = 92:\ z = \frac{x-\mu}{\sigma} = \frac{92-70}{8} = 2.75$$

From the *z*-scores, 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.

8 5 feet = 5(12) = 60 inches

Man: $z = \frac{x-\mu}{\sigma} = \frac{60-69.9}{3} = -3.3$; Woman: $z = \frac{x-\mu}{\sigma} = \frac{60-64.3}{2.6} \approx -1.7$

The *z*-score for the 5-foot-tall man is 3.3 standard deviations below the mean. This is an unusual height for a man. The *z*-score for the 5-foot-tall woman is 1.7 standard deviations below the mean. This is among the typical heights for a woman.

2.5 EXERCISE SOLUTIONS

1. The talk is longer in length than 75% of the lectures in the series.

2. The motorcycle's fuel efficiency is higher than 90% of the other vehicles in its class.

3. The student scored higher than 89% of the students who took the Fundamentals of Engineering exam.

4. The student has a higher IQ score than 91% of the students in the same age group.

5. The interquartile range of a data set can be used to identify outliers because data values that are greater than $Q_3 + 1.5(\text{IQR})$ or less than $Q_1 - 1.5(\text{IQR})$ are considered outliers.

6. Quartiles are special cases of percentiles. Q_1 is the 25th percentile, Q_2 is the 50th percentile, and Q_3 is the 75th percentile.

7. True

8. False. The second quartile is the median of an ordered data set.

9. False. An outlier is any number above $Q_3 + 1.5(\text{IQR})$ or below $Q_1 - 1.5(\text{IQR})$.

10. False. It is possible to have a *z*-score of zero when the *x*-value equals the mean.

11. (a) 51 54 56 **57** 59 60 60 **60** 60 62 63 **63** 63 65 80

Q_1 Q_2 Q_3

(b) $\text{IQR} = Q_3 - Q_1 = 63 - 57 = 6$

(c) $Q_1 - 1.5(IQR) = 57 - 1.5(6) = 48$; $Q_3 + 1.5(IQR) = 63 + 1.5(6) = 72$. The date entry 80 is an outlier.

12. (a) 19 20 20 21 **21** **21** 22 22 22 **22** **23** 23 23 23 **24** **24** 25 25 25 29

$Q_1 = 21$ $Q_2 = 22.5$ $Q_3 = 24$

(b) $\text{IQR} = Q_3 - Q_1 = 24 - 21 = 3$

(c) $Q_1 - 1.5(IQR) = 21 - 1.5(3) = 16.5$; $Q_3 + 1.5(IQR) = 24 + 1.5(3) = 28.5$. The data entry 29 is an outlier.

13. Min = 0, Q_1 = 2, Q_2 = 5, Q_3 = 8, Max = 10

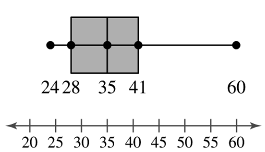
14. Min = 500, Q_1 = 580, Q_2 = 605, Q_3 = 630, Max = 720

15. (a) 24 26 27 **28** 30 32 35 **35** 36 39 39 **41** 50 51 60

Q_1 Q_2 Q_3

Min = 24, $Q_1 = 28$, $Q_2 = 35$, $Q_3 = 41$, Max = 60

(b)

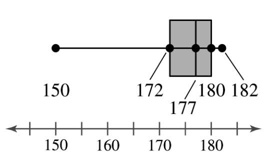


16. (a) $\overbrace{150\ \ 170\ \ 171\ \ 173\ \ 174\ \ 176}^{\text{lower half}}\ \overbrace{178\ \ 178\ \ 180\ \ 180\ \ 181\ \ 182}^{\text{upper half}}$

Q_1 Q_2 Q_3

Min = 150, $Q_1 = 172$, $Q_2 = 177$, $Q_3 = 180$, Max = 182

(b)

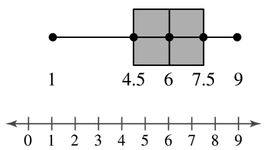


17. (a) $\overbrace{1\ 2\ 2\ 4\ 4\ 5\ 5\ 5\ 6\ 6}^{\text{lower half}}\ \overbrace{6\ 7\ 7\ 7\ 7\ 8\ 8\ 8\ 9\ 9}^{\text{upper half}}$

Q_1 Q_2 Q_3

Min = 1, $Q_1 = 4.5$, $Q_2 = 6$, $Q_3 = 7.5$, Max = 9

(b)

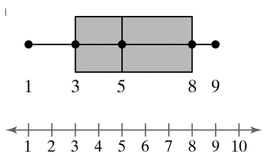


18. (a) 1 1 2 2 2 2 2 3 **3** 3 3 3 3 4 4 4 **5** **5** 5 5 5 6 7 7 7 **8** 8 8 9 9 9 9 9 9

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

Min = 1, $Q_1 = 3$, $Q_2 = 5$, $Q_3 = 8$, Max = 9

(b)



19. None. The Data are not skewed or symmetric.

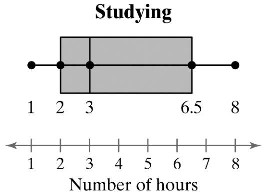
20. Skewed right. Most of the data lie to the left in the box-and-whisker plot.

21. Skewed left. Most of the data lie to the right in the box-and-whisker plot.

22. Symmetric. The data are evenly spaced to the left and to the right of the median.

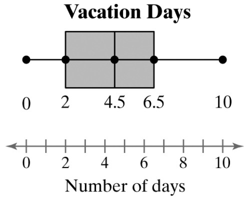
23. Min = 1, $Q_1 = 2$, $Q_2 = 3$, $Q_3 = 6.5$, Max = 8

Studying

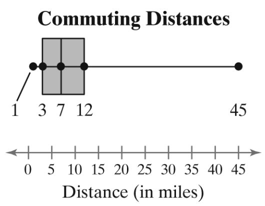


24. Min = 0, $Q_1 = 2$, $Q_2 = 4.5$, $Q_3 = 6.5$, Max = 10

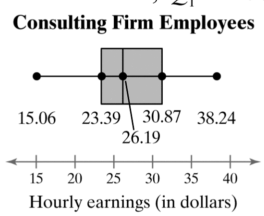
Vacation Days



25. Min = 1, $Q_1 = 3$, $Q_2 = 7$, $Q_3 = 12$, Max = 45



26. Min = 15.06, $Q_1 = 23.39$, $Q_2 = 26.19$, $Q_3 = 30.87$, Max = 38.24



27. (a) 6.5 hours (b) About 50% (c) About 25%

28. (a) $26.19 per hour (b) About 75% (c) About 75% (d) About 25%

29. About 158; About 70% of quantitative reasoning scores on the Graduate Record Examination are less than 158.

30. About 150; About 40% of quantitative reasoning scores on the Graduate Record Examination are less than 150.

31. About 8th percentile; About 8% of quantitative reasoning scores on the Graduate Record Examination are less than 140.

32. About 97th percentile; About 97% of quantitative reasoning scores on the Graduate Record Examination are less than 170.

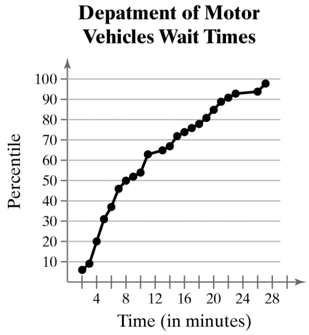
33. $\text{Percentile of } 40 = \dfrac{\text{number of data entries less than } 40}{\text{total number of data entries}} \cdot 100 = \dfrac{3}{30} \cdot 100 = 10^{\text{th}} \text{ percentile}$

34. $\text{Percentile of } 56 = \dfrac{\text{number of data entries less than } 56}{\text{total number of data entries}} \cdot 100 = \dfrac{21}{30} \cdot 100 = 70^{\text{th}} \text{ percentile}$

35. 75^{th} percentile $= Q_3 = 56$; Ages over 56 are 57,57,61,61,65,66

36. 25^{th} percentile $= Q_1 = 43$; Ages below 43 are 28,35,38,40,41,41,42

37.



38. The 50th percentile is about 8 minutes.
About 50% of wait times are less than 8 minutes.

39. A wait time of 20 minutes corresponds to about the 85th percentile.

40. The wait times between the 25th and 75th percentiles are about 4.5 minutes to 16 minutes.

41. $A \Rightarrow z = -1.43$
$B \Rightarrow z = 0$
$C \Rightarrow z = 2.14$
The z-score 2.14 is unusual because it is so large.

42. $A \Rightarrow z = -1.54$
$B \Rightarrow z = 0.77$
$C \Rightarrow z = 1.54$
None of the z-scores are unusual.

43. Christopher Froome: $x = 31 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{31-27.9}{3.3} \approx 0.94$

Not unusual; The z-score is 0.94, so the age of 31 is about 0.94 standard deviation above the mean.

44. Jan Ullrich: $x = 24 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{24-27.9}{3.3} \approx -1.18$

Not unusual; The z-score is -1.18, so the age of 24 is about 1.18 standard deviations below the mean.

45. Antonin Magne: $x = 27 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{27-27.9}{3.3} \approx -0.27$

Not unusual; The z-score is -0.27, so the age of 27 is about 0.27 standard deviation below the mean.

46. Firmin Lambot: $x = 36 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{36-27.9}{3.3} \approx 2.45$

Unusual; The z-score is 2.45, so the age of 36 is about 2.45 standard deviations above the mean.

47. Henri Cornet: $x = 20 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{20-27.9}{3.3} \approx -2.39$

Unusual; The z-score is -2.39, so the age of 20 is about 2.39 standard deviations below the mean.

48. Philippe Thys: $x = 28 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{28-27.9}{3.3} \approx 0.03$

Not unusual; The z-score is 0.03, so the age of 28 is about 0.03 standard deviation above the mean.

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

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

$x = 30{,}000 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{30{,}000-35{,}000}{2{,}250} \approx -2.22$

The tire with a life span of 30,000 miles has an unusually short life span.

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

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

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

50. (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 fruit fly with a life span of 42 days has an unusually long life span.

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

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

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

51. Robert Duvall: $x = 53 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{53-43.7}{8.7} \approx 1.07$

Jack Nicholson: $x = 46 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{46-50.4}{13.8} \approx -0.32$

The age of Robert Duvall was about 1 standard deviation above the mean age of Best Actor winners, and the age of Jack Nicholson was less than 1 standard deviation below the mean age of Best Supporting Actor winners. Neither actor's age is unusual.

52. Jamie Foxx: $x = 37 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{37-43.7}{8.7} \approx -0.77$

Morgan Freeman: $x = 67 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{67-50.4}{13.8} \approx 1.20$

The age of Jamie Foxx was less than 1standard deviation below the mean age of Best Actor winners, and the age of Morgan Freeman was between 1and 2 standard deviations above the mean age of Best Supporting Actor winners. Neither actor's age is unusual.

53. John Wayne: $x = 62 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{62-43.7}{8.7} \approx 2.10$

Gig Young: $x = 56 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{56-50.4}{13.8} \approx 0.41$

The age of John Wayne was more than 2 standard deviations above the mean age of Best Actor winners, which is unusual. The age of Gig Young was less than 1 standard deviation above the mean age of Best Supporting Actor winners, which is not unusual.

54. Henry Fonda: $x = 76 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{76-43.7}{8.7} \approx 3.71$

John Gielgud: $x = 77 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{77-50.4}{13.8} \approx 1.93$

The age of Henry Fonda was more than 3 standard deviations above the mean age of Best Actor winners, which is very unusual. The age of John Gielgud was less than 2 standard deviations above the mean age of Best Supporting Actor winners, which is not unusual.

55.

$$1\ 2\ \underset{\substack{\uparrow\\ Q_1}}{3}\ 3\ 5\ \underset{\substack{\uparrow\\ Q_2}}{5}\ 7\ \underset{\substack{\uparrow\\ Q_3}}{7}\ 8\ 10$$

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

56.

$$22\ 23\ 24\ \underset{\substack{\uparrow\\ Q_1}}{\ }\ 32\ 33\ 34\ \underset{\substack{\uparrow\\ Q_2}}{\ }\ 36\ 38\ 39\ \underset{\substack{\uparrow\\ Q_3}}{\ }\ 40\ 41\ 47$$

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

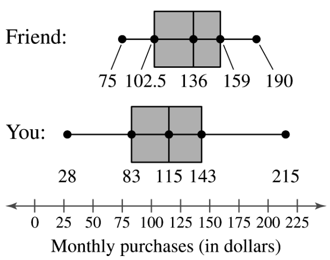
57. (a) The distribution of Concert 1 is symmetric. The distribution of Concert 2 is skewed right. Concert 1 has less variation.

(b) Concert 2 is more likely to have outliers because it has more variation.

(c) Concert 1, because 68% of the data should be between ± 16.3 of the mean.

(d) No, you do not know the number of songs played at either concert or the actual lengths of the songs.

58. **Credit Card Purchases**



Your distribution is symmetric and your friend's distribution slightly skewed to the right.

59. (a) $\overbrace{2\ \ 7\ \ 8\ \ 9\ \ 9\ \ 10\ \ 10}^{\text{lower half}}\ \ 11\ \ \overbrace{11\ \ 12\ \ 12\ \ 13\ \ 15\ \ 16\ \ 24}^{\text{upper half}}$

Q_1 Q_2 Q_3

$Q_1 = 9$, $Q_2 = 11$, $Q_3 = 13$

$\text{IQR} = Q_3 - Q_1 = 13 - 9 = 4$

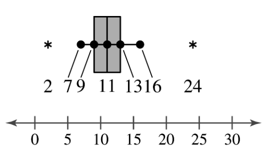
$1.5 \times \text{IQR} = 6$

$Q_1 - (1.5 \times \text{IQR}) = 9 - 6 = 3$

$Q_3 + (1.5 \times \text{IQR}) = 13 + 6 = 19$

Any values less than 6 or greater than 19 are outliers. So, 2 and 24 are outliers.

(b)



60. (a) $\overbrace{62\ \ 72\ \ 72\ \ 74\ \ 75\ \ 75}^{\text{lower half}}\ \ \overbrace{75\ \ 76\ \ 78\ \ 80\ \ 80\ \ 95}^{\text{upper half}}$

Q_1 Q_2 Q_3

$Q_1 = 73$, $Q_2 = 75$, $Q_3 = 79$

$\text{IQR} = Q_3 - Q_1 = 79 - 73 = 6$

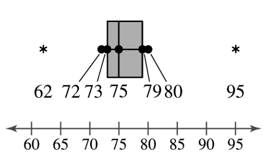
$1.5 \times \text{IQR} = 9$

$Q_1 - (1.5 \times \text{IQR}) = 73 - 9 = 64$

$Q_3 + (1.5 \times \text{IQR}) = 79 + 9 = 88$

Any values less than 64 or greater than 88 are outliers. So, 62 and 95 are outliers.

(b)



61. (a) 1, 23, 29, 35, 37, 46, **46**, 47, 49, 52, 53, 59, 83

$Q_1 = 32$, $Q_2 = 46$, $Q_3 = 52.5$

$\text{IQR} = Q_3 - Q_1 = 52.5 - 32 = 20.5$

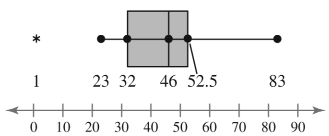
$1.5 \times \text{IQR} = 30.75$

$Q_1 - (1.5 \times \text{IQR}) = 32 - 30.75 = 1.25$

$Q_3 + (1.5 \times \text{IQR}) = 52.5 + 30.75 = 83.25$

Any values less than 1.25 or greater than 83.25 are outliers. So, 1 is an outlier.

(b)



62. (a) 19, 27, 30, 36, 38, 47, 47, 48, 50, 50, 53, 54, 56, 60, 62, 90

$Q_1 = 37$, $Q_2 = 49$, $Q_3 = 55$

$\text{IQR} = Q_3 - Q_1 = 55 - 37 = 18$

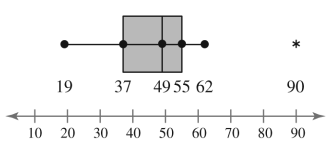
$1.5 \times \text{IQR} = 27$

$Q_1 - (1.5 \times \text{IQR}) = 37 - 27 = 10$

$Q_3 + (1.5 \times \text{IQR}) = 55 + 27 = 82$

Any values less than 10 or greater than 82 are outliers. So, 90 is an outlier.

(b)



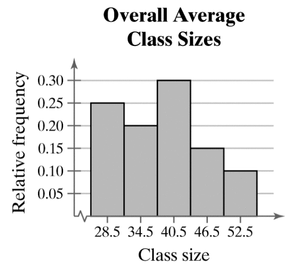
63. Answers will vary.

CHAPTER 2 REVIEW EXERCISE SOLUTIONS

1. $\text{Class width} = \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{55 - 26}{5} = 5.8 \Rightarrow 6$

Class	Midpoint	Class boundaries	Frequency, f	Relative frequency	Cumulative frequency
26–31	28.5	25.5–31.5	5	0.25	5
32–37	34.5	31.5–37.5	4	0.20	9
38–43	40.5	37.5–43.5	6	0.30	15
44–49	46.5	43.5–49.5	3	0.15	18
50–55	52.5	49.5–55.5	2	0.10	20
			$\sum f = 20$	$\sum \frac{f}{n} = 1$	

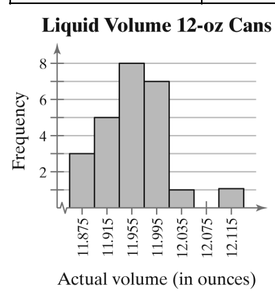
2.



Class with greatest relative frequency: 38 – 43
Class with least relative frequency: 50 – 55

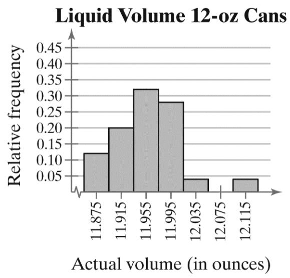
3. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{12.10 - 11.86}{7} \approx 0.03 \Rightarrow 0.04$

Class	Midpoint	Frequency, f	Relative frequency
11.86-11.89	11.875	3	0.12
11.90-11.93	11.915	5	0.20
11.94-11.97	11.955	8	0.32
11.98-12.01	11.995	7	0.28
12.02-12.05	12.035	1	0.04
12.06-12.09	12.075	0	0.00
12.10-12.13	12.115	1	0.04
		$\sum f = 25$	$\sum \frac{f}{n} = 1$



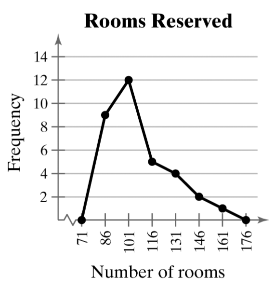
4. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{12.10 - 11.86}{7} \approx 0.03 \Rightarrow 0.04$

Class	Midpoint	Frequency, f	Relative frequency
11.86-11.89	11.875	3	0.12
11.90-11.93	11.915	5	0.20
11.94-11.97	11.955	8	0.32
11.98-12.01	11.995	7	0.28
12.02-12.05	12.035	1	0.04
12.06-12.09	12.075	0	0.00
12.10-12.13	12.115	1	0.04
		$\sum f = 25$	$\sum \frac{f}{n} = 1$

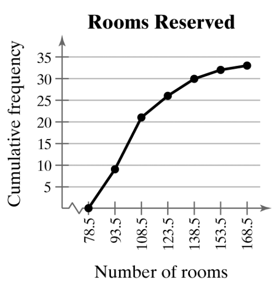


5. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{166-79}{6} = 14.5 \Rightarrow 15$

Class	Mid point	Frequency,
79–93	86	9
94–108	101	12
109–123	116	5
124–138	131	4
139–153	146	2
154–168	161	1
		$\sum f = 33$



6.



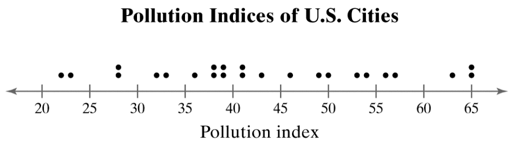
7. Because the data entries go from a low of 22 to a high of 65, use stem values from 2 to 6. List the stems to the left of a vertical line. For each data entry, list a leaf to the right of its stem.

Pollution Indices of U.S. Cities

2 | 2 3 8 8   Key: 2 | 2 = 22
3 | 2 3 6 8 8 9 9
4 | 1 1 3 6 9
5 | 0 3 4 6 7
6 | 3 5 5

Sample answer: Most U.S. cities have a pollution index from 32 to 57.

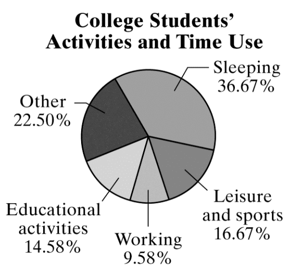
8.



Sample answer: Most U.S. cities have a pollution index from 32 to 57.

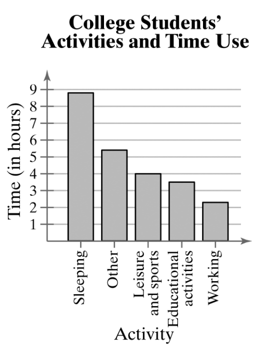
9.

Location	Frequency	Relative frequency	Degrees
Sleeping	8.8	0.3667	132°
Leisure and Sports	4.0	0.1667	60°
Working	2.3	0.0958	34°
Educational Activities	3.5	0.1458	52°
Other	5.4	0.2250	81°
	$\sum f = 24$	$\sum \frac{f}{n} = 1$	



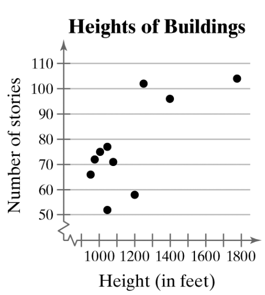
Sample answer: Full-time university and college students spend the least amount of time working.

10.



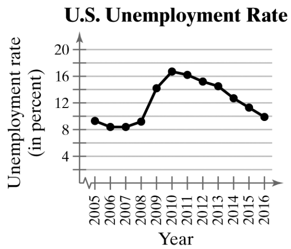
Sample answer: Full-time university and college students spend the most amount of time sleeping.

11.



Sample answer: The number of stories appears to increase with height.

12.



Sample answer: The real unemployment rate varied by a couple of percentage points from 2005 to 2008, then increased dramatically from 2008 to 2010, and then decreased from 2011 to 2016.

13. $\bar{x} = \frac{\sum x}{n} = \frac{295}{10} = 29.5$

21.0 24.0 26.0 28.0 **29.5** **29.5** 31.0 33.0 35.5 37.5

↑

median = 29.5

Mode = 29.5 (occurs 2 times)

14. $\bar{x}$ is not possible

median is not possible

mode = "$250-999"

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

15.

Source	**Score, *x***	**Weight, *w***	$\boldsymbol{x \cdot w}$
Test 1	78	0.15	11.7
Test 2	72	0.15	10.8
Test 3	86	0.15	12.9
Test 4	91	0.15	13.65
Test 5	87	0.15	13.05
Test 6	80	0.25	20
		$\sum w = 1$	$\sum (x \cdot w) = 82.1$

$$\bar{x} = \frac{\sum (x \cdot w)}{\sum w} = \frac{82.1}{1} = 82.1$$

16.

Source	**Score, x**	**Weight, w**	$x \cdot w$
Test 1	96	0.2	19.2
Test 2	85	0.2	17
Test 3	91	0.2	18.2
Test 4	86	0.4	34.4
		$\sum w = 1$	$\sum(x \cdot w) = 88.8$

$$\bar{x} = \frac{\sum(x \cdot w)}{\sum w} = \frac{88.8}{1} = 88.8$$

17.

Midpoint, x	**Frequency, f**	$x \cdot f$
28.5	5	142.5
34.5	4	138
40.5	6	243
46.5	3	139.5
52.5	2	105
	$n = 20$	$\sum(x \cdot f) = 768$

$$\bar{x} = \frac{\sum(x \cdot f)}{n} = \frac{768}{20} \approx 38.4$$

18.

x	f	$x \cdot f$
0	13	0
1	9	9
2	19	38
3	8	24
4	5	20
5	2	10
6	4	24
	$n = 60$	$\sum(x \cdot f) = 125$

$$\bar{x} = \frac{\sum(x \cdot f)}{n} = \frac{125}{60} \approx 2.1$$

19. Skewed right

20. Skewed right

21. Skewed right

22. Skewed left

23. Mean, because the mean is to the right of the median in a skewed right distribution.

24. Median, because the mean is to the left of the median in a skewed left distribution.

25. Range = Max – Min = 15 – 1 = 14

$$\mu = \frac{\sum x}{N} = \frac{96}{14} \approx 6.9$$

x	$x-\mu$	$(x-\mu)^2$
4	–2.9	8.41
2	–4.9	24.01
9	2.1	4.41
12	5.1	26.01
15	8.1	65.61
3	–3.9	15.21
6	–0.9	0.81
8	1.1	1.21
1	–5.9	34.81
4	–2.9	8.41
14	7.1	50.41
12	5.1	26.01
3	–3.9	15.21
3	–3.9	15.21
$\sum x = 96$	$\sum(x-\mu) \approx 0$	$\sum(x-\mu)^2 = 295.74$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} = \frac{295.74}{14} \approx 21.1$$

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}} = \sqrt{\frac{295.74}{14}} \approx 4.6$$

26. Range = Max – Min = 83 – 56 = 27

$$\mu = \frac{\sum x}{N} = \frac{554}{8} \approx 69.25$$

x	$x-\mu$	$(x-\mu)^2$
61	–8.25	68.063
80	10.75	115.563
68	–1.25	1.563
83	13.75	189.063
78	8.75	76.563
66	–3.25	10.563
62	–7.25	52.563
56	–13.25	175.563
$\sum x = 554$	$\sum(x-\mu) \approx 0$	$\sum(x-\mu)^2 = 689.5$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} = \frac{689.5}{8} \approx 86.19$$

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}} = \sqrt{\frac{689.5}{8}} \approx 9.28$$

27. Range = Max – Min = \$7439 – \$5395 = \$2044

$$\bar{x} = \frac{\sum x}{n} = \frac{100,269}{16} \approx \$6266.81$$

x	$x - \bar{x}$	$(x - \bar{x})^2$
5816	−450.81	203,230
7220	953.19	908,571
6045	−221.81	49,200
7439	1172.19	1,374,029
5612	−654.81	428,776
5395	−871.81	760,053
6341	74.19	5504
6908	641.19	411,125
6106	−160.81	25,860
5561	−705.81	498,168
7361	1094.19	1,197,252
5710	−556.81	310,037
6320	53.19	2829
5538	−728.81	531,164
6265	−1.81	3
6632	365.19	133,364
$\sum x = 100,269$	$\sum (x - \bar{x}) \approx 0$	$\sum (x - \bar{x})^2 = 6,839,165$

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1} = \frac{6,839,165}{15} \approx 455,944.3$$

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}} = \sqrt{\frac{6,839,165}{15}} \approx \$675.24$$

28. Range = Max – Min = 71,534 – 45,120 = \$26,414

$$\bar{x} = \frac{\sum x}{n} = \frac{433,020}{8} \approx \$54,127.50$$

x	$x - \bar{x}$	$(x - \bar{x})^2$
62,222	8094.5	65,520,930.25
56,719	2591.5	6,715,872.25
50,259	−3868.5	14,965,292.25
45,120	−9007.5	81,135,056.25
47,692	−6435.5	41,415,660.25
45,985	−8142.5	66,300,306.25
53,489	−638.5	407,682.25
71,534	17406.5	302,986,242.25
$\sum x = 433,020$	$\sum (x - \bar{x}) \approx 0$	$\sum (x - \bar{x})^2 = 579,447,042$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{579,447,042}{7} \approx 82,778,148.86$$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{579,447,042}{7}} \approx \$9098.25$$

29. 95% of the distribution lies within 2 standard deviations of the mean.

$\bar{x} - 2s = 110 - (2)(17.50) = 75$

$\bar{x} + 2s = 110 + (2)(17.50) = 145$

95% of the distribution lies between \$75 and \$145.

30. $(73.00,\ 102.00) \rightarrow (87.50 - 1(14.50),\ 87.50 + 1(14.50)) \rightarrow (\bar{x} - s, \bar{x} + s)$

68% of the satellite television charges lie between \$73.00 and \$102.00.

31. $(\bar{x} - 2s,\ \bar{x} + 2s) \rightarrow (24,\ 40)$ are 2 standard deviations from the mean.

$$1 - \frac{1}{k^2} = 1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = 0.75$$

At least (40)(0.75) = 30 customers have a mean sale between \$24 and \$40.

32. $(\bar{x} - 2s,\ \bar{x} + 2s) \rightarrow (2.3,\ 17.5)$ are 2 standard deviations from the mean.

$$1 - \frac{1}{k^2} = 1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = 0.75$$

At least $(135)(0.75) \approx 101$ shuttle flights lasted between 2.3 days and 17.5 days.

33. $\bar{x} = \frac{\sum xf}{n} = \frac{99}{40} \approx 2.5$

x	f	xf	$x - \bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
0	1	0	–2.5	6.25	6.25
1	8	8	–1.5	2.25	18.00
2	13	26	–0.5	0.25	3.25
3	10	30	0.5	0.25	2.50
4	5	20	1.5	2.25	11.25
5	3	15	2.5	6.25	18.75
	$n = 40$	$\sum xf = 99$			$\sum(x-\bar{x})^2 f = 60$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2 f}{n-1}} = \sqrt{\frac{60}{39}} \approx 1.2$$

34. $\bar{x} = \dfrac{\sum xf}{n} = \dfrac{61}{25} \approx 2.4$

x	f	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
0	4	0	–2.4	5.76	23.04
1	5	5	–1.4	1.96	9.80
2	2	4	–0.4	0.16	0.32
3	9	27	0.6	0.36	3.24
4	1	4	1.6	2.56	2.56
5	3	15	2.6	6.76	20.28
6	1	6	3.6	12.96	12.96
	$n = 25$	$\sum xf = 61$			$\sum (x-\bar{x})^2 f = 72.2$

$$s = \sqrt{\frac{\sum (x-\bar{x})^2 f}{n-1}} = \sqrt{\frac{72.2}{24}} \approx 1.7$$

35. Freshmen: $\bar{x} = \dfrac{\sum x}{n} = \dfrac{23.1}{9} \approx 2.567$

x	$x-\bar{x}$	$(x-\bar{x})^2$
2.8	0.233	0.0543
1.8	-0.767	0.5833
4.0	1.433	2.0535
3.8	1.233	1.5203
2.4	-0.167	0.0279
2.0	-0.567	0.3215
0.9	-1.667	2.7789
3.6	1.033	1.0671
1.8	-0.767	0.5883
		$\sum (x-\bar{x})^2 = 9.0000$

$$s = \sqrt{\frac{\sum (x-\bar{x})^2}{n-1}} = \sqrt{\frac{9.0000}{8}} \approx 1.061$$

$$CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{1.061}{2.567} \cdot 100\% \approx 41.3\%$$

Seniors: $\bar{x} = \dfrac{\sum x}{n} = \dfrac{26.6}{9} \approx 2.956$

x	$x-\bar{x}$	$(x-\bar{x})^2$
2.3	−0.656	0.4303
3.3	0.344	0.1183
1.8	−1.156	1.3363
4.0	1.044	1.0899
3.1	0.144	0.0207
2.7	−0.256	0.0655
3.9	0.944	0.8911
2.6	−0.356	0.1267
2.9	−0.056	0.0031
		$\sum(x-\bar{x})^2 = 4.0822$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{4.0822}{8}} \approx 0.714$$

$$CV = \frac{s}{\bar{x}} \cdot 100\% = \frac{0.714}{2.956} \cdot 100\% \approx 24.2\%$$

Grade point averages are more variable for freshmen than seniors.

36. Ages: $\mu = \frac{\sum x}{N} = \frac{406}{8} = 50.75$

x	$x-\mu$	$(x-\mu)^2$
66	15.25	232.5625
54	3.25	10.5625
47	−3.75	14.0625
61	10.25	105.0625
36	−14.75	217.5625
59	8.25	68.0625
50	−0.75	0.5625
33	−17.75	315.0625
$\sum x = 406$	$\sum(x-\mu) = 0$	$\sum(x-\mu)^2 = 963.5$

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{N}} = \sqrt{\frac{963.5}{8}} \approx 10.97$$

$$CV = \frac{\sigma}{\mu} \cdot 100\% = \frac{10.97}{50.75} \cdot 100\% \approx 21.6\%$$

Years of Experience: $\mu = \frac{\sum x}{N} = \frac{185}{8} \approx 23.1$

x	$x-\mu$	$(x-\mu)^2$
37	13.9	193.21
20	–3.1	9.61
23	–0.1	0.01
32	8.9	79.21
14	–9.1	82.81
29	5.9	34.81
22	–1.1	1.21
8	–15.1	228.01
$\sum x = 185$	$\sum(x-\mu)\approx 0$	$\sum(x-\mu)^2 = 628.88$

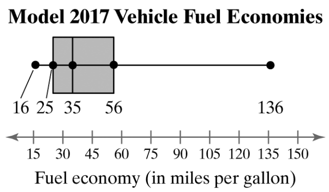
$$\sigma = \sqrt{\frac{\sum\left(x-\overline{x}\right)^2}{N}} = \sqrt{\frac{628.88}{8}} \approx 8.866$$

$$CV = \frac{\sigma}{\mu}\cdot 100\% = \frac{8.866}{23.13}\cdot 100\% \approx 38.3\%$$

Years of experience are more variable than ages for all lawyers at a firm.

37. 16, 16, 22, 22, 22, 22, **25**, 25, 30, 30, 34, 34, <u>35, 35</u>, 35, 41, 46, 50, 52, **56**, 58, 107, 112, 119, 124, 136
Min = 16, $Q_1 = 25$, $Q_2 = 35$, $Q_3 = 56$, Max = 136

38. $\text{IQR} = Q_3 - Q_1 = 56 - 25 = 31$ miles per gallon

39. 

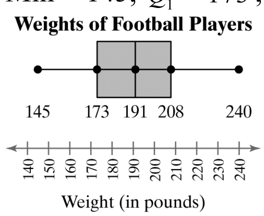
40. Count the number of entries that are 56 or below: 20 vehicles

41. 21.0 24.0 <u>26.0</u> 28.0 **29.5** **29.5** 31.0 <u>33.0</u> 35.5 37.5

median = 29.5

$\text{IQR} = Q_3 - Q_1 = 33 - 26 = 7$ inches

42. 145, 156, 167, 172, **173**, 184, 185, 190, <u>190, 192</u>, 195, 197, 205, **208**, 212, 227, 228, 240
Min = 145, $Q_1 = 173$, $Q_2 = 191$, $Q_3 = 208$, Max = 240



43. The 65th percentile means that 65% had a test grade of 75 or less. So, 35% scored higher than 75.

44. If there are 115 stations with a larger daily audience, then this station has the 721 – 115 = 606$^{\text{th}}$ largest audience. The percentile of $606 = \dfrac{606}{721} \cdot 100 = 84^{\text{th}}$ percentile

45. $z = \dfrac{16{,}500 - 11{,}830}{2370} \approx 1.97$

Not unusual; The z-score is 1.97, so a towing capacity of 16,500 pounds is about 1.97 standard deviations above the mean.

46. $z = \dfrac{5500 - 11{,}830}{2370} \approx -2.67$

Unusual; The z-score is –2.67, so a towing capacity of 5500 pounds is about 2.67 standard deviations below the mean.

47. $z = \dfrac{18{,}000 - 11{,}830}{2370} = 2.60$

Unusual; The z-score is 2.60, so a towing capacity of 18,000 pounds is about 2.60 standard deviations above the mean.

48. $z = \dfrac{11{,}300 - 11{,}830}{2370} = -0.22$

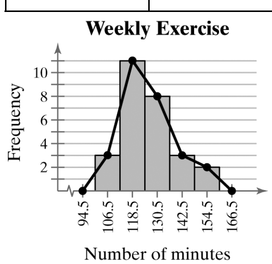
Not unusual; The z-score is –0.22, so a towing capacity of 11,300 pounds is about 0.22 standard deviation below the mean.

CHAPTER 2 QUIZ SOLUTIONS

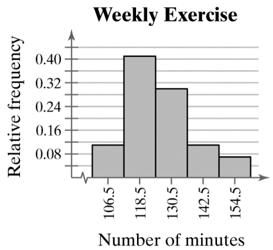
1. (a) Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{157 - 101}{5} = 11.2 \Rightarrow 12$

Class	Midpoint	Class boundaries	Frequency, f	Relative frequency	Cumulative frequency
101-112	106.5	100.5-112.5	3	0.11	3
113-124	118.5	112.5-124.5	11	0.41	14
125-136	130.5	124.5-136.5	8	0.30	22
137-148	142.5	136.5-148.5	3	0.11	25
149-160	154.5	148.5-160.5	2	0.07	27
			$\sum f = 27$	$\sum \dfrac{f}{n} = 1$	

(b)



(c)



(d) Skewed right

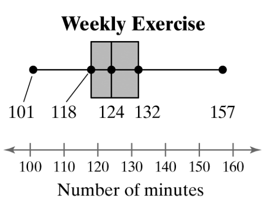
(e)

(f) **Weekly Exercise (in minutes)**

10	1 8	Key: 10 | 8 = 108
11	1 4 6 7 8 9 9	
12	0 0 3 3 4 7 7 8	
13	0 1 1 2 5 9 9	
14	2	
15	0 7	

(g) 101, 108, 111, 114, 116, 117, **118**, 119, 119, 120, 120, 123, 123, **124**, 127, 127, 128, 130, 131, 131, **132**, 135, 139, 139, 142, 150, 157

Min = 101, $Q_1 = 118$, $Q_2 = 124$, $Q_3 = 132$, Max = 157



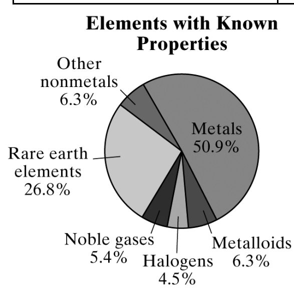
2. $\bar{x} = \dfrac{\sum xf}{n} = \dfrac{3403.5}{27} \approx 126.1$

Midpoint, x	**Frequency, f**	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
106.5	3	319.5	−19.6	384.16	1152.48
118.5	11	1303.5	−7.6	57.76	635.36
130.5	8	1044	4.4	19.36	154.88
142.5	3	427.5	16.4	268.96	806.88
154.5	2	309.0	28.4	806.56	1613.12
	$n=27$	$\sum xf = 3403.5$			$\sum(x-\bar{x})^2 f = 4362.72$

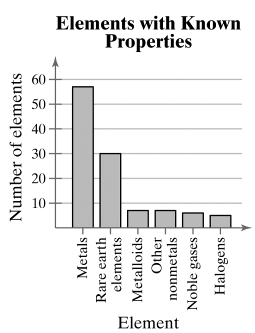
$$s=\sqrt{\frac{\sum(x-\bar{x})^2 f}{n-1}}=\sqrt{\frac{4362.72}{26}}\approx 13.0$$

3. (a)

Category	**Frequency**	**Relative frequency**	**Degrees**
Metals	57	0.5089	183°
Metalloids	7	0.0625	23°
Halogens	5	0.0446	16°
Noble gases	6	0.0536	19°
Rare earth elements	30	0.2679	96°
Other nonmetals	7	0.0625	23°
	$n=112$	$\sum \frac{f}{n}=1$	



(b)



4. (a) $\bar{x} = \dfrac{\sum x}{n} = \dfrac{16,262}{16} \approx 1016.4$

718, 720, 749, 790, 860, 891, 969, **976, 1062**, 1100, 1100, 1124, 1248, 1255, 1316, 1384

$$\text{median} = \frac{976 + 1062}{2} = 1019$$

mode = 1100 (occurs twice)

The mean or median best describes a typical salary because there are no outliers.

(b) Range = Max – Min = 1384 – 718 = 666

x	$x-\bar{x}$	$(x-\bar{x})^2$
1100	83.6	6989
749	–267.4	71,503
720	–296.4	87,853
1062	45.6	2079
1384	367.6	135,130
1248	231.6	53,639
1124	107.6	11,578
891	–125.4	15,725
1255	238.6	56,930
969	–47.4	2247
976	–40.4	1632
790	–226.4	51,257
718	–298.4	89,043
860	–156.4	24,461
1316	299.6	89,760
1100	83.6	6989
		$\sum(x-\bar{x})^2 = 706,815$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{706,815}{15} \approx 47,120.9$$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{706,815}{15}} \approx 217.1$$

(c) $CV = \dfrac{s}{\bar{x}} \cdot 100\% = \dfrac{217.1}{1016.4} \cdot 100\% \approx 21.4\%$

5. $\bar{x} - 2s = 180,000 - 2 \cdot 15,000 = \$150,000$

$\bar{x} + 2s = 180,000 + 2 \cdot 15,000 = \$210,000$

95% of the new home prices fall between \$150,000 and \$210,000.

6. (a) $x = 225,000: \; z = \dfrac{x-\bar{x}}{s} = \dfrac{225,000 - 180,000}{15,000} = 3.0$

Unusual; The z-score is 3, so a new home price of \$225,000 is about 3 standard deviations above the mean.

(b) $x = 80,000$: $z = \dfrac{x-\overline{x}}{s} = \dfrac{80,000-180,000}{15,000} \approx -6.67$

Unusual; The z-score is –6.67, so a new home price of \$80,000 is about 6.67 standard deviations below the mean.

(c) $x = 200,000$: $z = \dfrac{x-\overline{x}}{s} = \dfrac{200,000-180,000}{15,000} \approx 1.33$

Not unusual; The z-score is 1.33, so a new home price of \$200,000 is about 1.33 standard deviations above the mean.

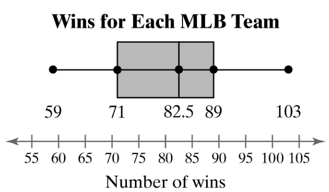
(d) $x = 147,000$: $z = \dfrac{x-\overline{x}}{s} = \dfrac{147,000-180,000}{15,000} = -2.2$

Unusual; The z-score is –2.2, so a new home price of \$147,000 is about 2.2 standard deviations below the mean.

7. 59 68 68 68 68 69 69 **71** 73 74 75 78 78 79 **81** **84** 84 86 86 86 87 87 **89** 89 91 93 94 95 95 103

$Q_1 = 71$ $Q_2 = 82.5$ $Q_3 = 89$

Min = 59, $Q_1 = 71$, $Q_2 = 82.5$, $Q_3 = 89$, Max = 103

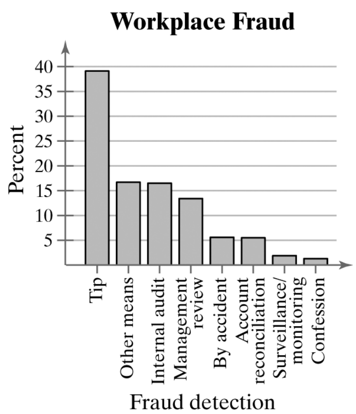


CUMULATIVE REVIEW FOR CHAPTERS 1 AND 2

1. Systematic sampling is used because every fortieth toothbrush from each assembly line is tested. It is possible for bias to enter into the sample if, for some reason, an assembly line makes a consistent error.

2. Simple random sampling is used because each telephone number has an equal chance of being dialed, and all samples of 1090 phone numbers have an equal chance of being selected. The sample may be biased because telephone sampling only samples those individuals who have telephones, who are available, and who are willing to respond.

3.



4. \$68,232 is a parameter. The median salary is based on all marketing account executives.

5. 88% is a statistic. The percent is based on a subset of the population.

6. (a) $\overline{x} = 86{,}500,\ s = 1500$

$(83{,}500,\ 89{,}500) = 86{,}500 \pm 2(1500) \Rightarrow 2$ standard deviations away from the mean.

Approximately 95% of the electrical engineers will have salaries between \$83,500 and \$89,500.

(b) $x = \$93{,}500:\ z = \dfrac{x - \overline{x}}{s} = \dfrac{93{,}500 - 86{,}500}{1500} \approx 4.67$

$x = \$85{,}600:\ z = \dfrac{x - \overline{x}}{s} = \dfrac{85{,}600 - 86{,}500}{1500} = -0.6$

$x = \$82{,}750:\ z = \dfrac{x - \overline{x}}{s} = \dfrac{82{,}750 - 86{,}500}{1500} = -2.5$

The salaries of \$93,500 and \$82,750 are unusual.

7. Population: Collection of opinions of all college and university admissions directors and enrollment officers
Sample: Collection of opinions of the 339 college and university admission directors and enrollment officers surveyed

8. Population: Reasons for pain reliever use of all Americans ages 12 or older
Sample: Reasons for pain reliever use of the 67,901 Americans ages 12 or older surveyed

9. Experiment. The study applies a treatment (digital device) to the subjects.

10. Observational study. The study does not attempt to influence the responses of the subjects.

11. Quantitative: The data are at the ratio level.

12. Qualitative: The data are at the nominal level.

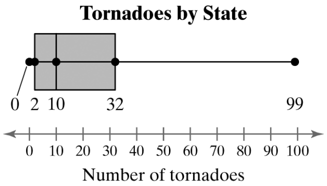
13. (a)

Q_1 Q_2

0 0 0 0 0 0 0 1 1 1 2 2 **2** 2 2 3 3 3 3 4 4 6 6 7 **9** **11** 11 12 15 16
16 23 23 27 31 31 32 **32** 40 44 45 46 47 48 50 55 67 87 90 99

Q_3

Min = 0, $Q_1 = 2$, $Q_2 = 10$, $Q_3 = 32$, Max = 99



(b) The distribution of the number of tornadoes is skewed right.

14.

Source	Score, x	Weight, w	$x \cdot w$
Test 1	85	0.15	12.75
Test 2	92	0.15	13.80
Test 3	84	0.15	12.60
Test 4	89	0.15	13.35
Test 5	91	0.40	36.40
		$\sum w = 1$	$\sum (x \cdot w) = 88.9$

$$\bar{x} = \frac{\sum (x \cdot w)}{\sum w} = \frac{88.9}{1} = 88.9$$

15. (a) $\bar{x} = \dfrac{49.4}{9} \approx 5.49$

3.4 3.9 4.2 4.6 5.4 6.5 6.8 7.1 7.5

median = 5.4

mode = none

Both the mean and median accurately describe a typical American alligator tail length. (Answers will vary.)

(b) Range – Max – Min – 7.5 – 3.4 = 4.1

x	$x - \bar{x}$	$(x - \bar{x})^2$
3.4	–2.09	4.3681
3.9	–1.59	2.5281
4.2	–1.29	1.6641
4.6	–0.89	0.7921
5.4	–0.09	0.0081
6.5	1.01	1.0201
6.8	1.31	1.7161
7.1	1.61	2.5921
7.5	2.01	4.0401
		$\sum (x - \bar{x})^2 = 18.7289$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{18.7289}{8} \approx 2.34$$

$$s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}} = \sqrt{\frac{18.7289}{8}} \approx 1.53$$

16. (a) An inference drawn from the study is that the life expectancies for Americans will continue to increase or remain stable.

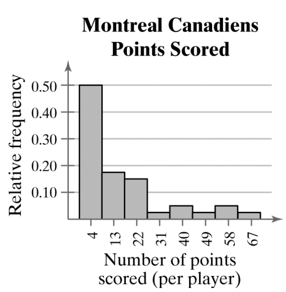
(b) This inference may incorrectly imply that women will have less of a chance of dying of heart disease in the future. The study was only conducted over the past 5 years and deaths may not decrease in the next year.

17. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{64-0}{8} = 8 \Rightarrow 9$

Class limits	Midpoint	Class boundaries	Frequency	Relative frequency	Cumulative frequency
0-8	4	–0.5-8.5	20	0.500	20
9-17	13	8.5-17.5	7	0.175	27
18-26	22	17.5-26.5	6	0.150	33
27-35	31	26.5-35.5	1	0.025	34
36-44	40	35.5-44.5	2	0.050	36
45-53	49	44.5-53.5	1	0.025	37
54-62	58	53.5-62.5	2	0.050	39
63-71	67	62.5-71.5	1	0.025	40
			$\sum f = 40$	$\sum \frac{f}{n} = 1$	

18. The distribution is skewed right.

19.



Class with greatest frequency: 0 – 8
Classes with least frequency: 27 – 35, 45 – 53, and 63 – 71

CHAPTER 2 TEST SOLUTIONS

1. (a) $\bar{x} = \dfrac{\sum x}{n} = \dfrac{964}{12} \approx 80.3$

$63\ 67\ 68\ 72\ 73\ \underbrace{81\ 85}\ 87\ 87\ 88\ 94\ 99$

median = 83

mode = 87 (occurs twice)

The median best represents the center of the data.

(b) Range = Max – Min = 99 – 63 = 36

x	$x - \bar{x}$	$(x - \bar{x})^2$
67	−13.3	176.89
72	−8.3	68.89
88	7.7	59.29
73	−7.3	53.29
99	18.7	349.69
85	4.7	22.09
81	0.7	0.49
87	6.7	44.89
63	−17.3	299.29
94	13.7	187.69
68	−12.3	151.29
87	6.7	44.89
	$\sum(x - \bar{x}) \approx 0$	$\sum(x - \bar{x})^2 = 1458.68$

$$s^2 = \frac{\sum(x - \bar{x})^2}{n-1} = \frac{1458.68}{11} \approx 132.6$$

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n-1}} = \sqrt{\frac{1458.68}{11}} \approx 11.5$$

(c) $CV = \dfrac{s}{\bar{x}} \cdot 100\% = \dfrac{11.5}{80.3} \cdot 100\% \approx 14.3\%$

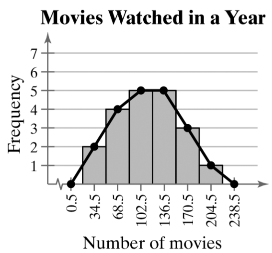
(d) **Points scored** Key: $6|3 = 63$

$$\begin{array}{r|l} 6 & 3\quad 7\quad 8 \\ 7 & 2\quad 3 \\ 8 & 1\quad 5\quad 7\quad 7\quad 8 \\ 9 & 4\quad 9 \end{array}$$

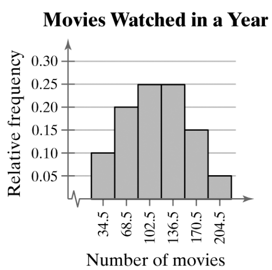
2. (a) Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{221-18}{6} \approx 33.8 \Rightarrow 34$

Class	Midpoint	Class boundaries	Frequency, f	Relative frequency	Cumulative frequency
18-51	34.5	17.5-51.5	2	0.10	2
52-85	68.5	51.5-85.5	4	0.20	6
86-119	102.5	85.5-119.5	5	0.25	11
120-153	136.5	119.5-153.5	5	0.25	16
154-187	170.5	153.5-187.5	3	0.15	19
188-221	204.5	187.5-221.5	1	0.05	20
			$\sum f = 20$	$\sum \frac{f}{n} = 1$	

(b)

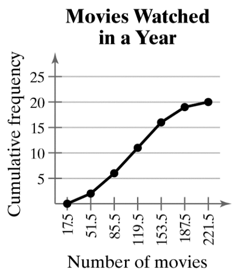


(c)



(d) The shape of the distribution is symmetric.

(e)



3. $\bar{x} = \dfrac{\sum xf}{n} = \dfrac{2254}{20} = 112.7$

Midpoint, x	Frequency, f	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
34.5	2	69.0	−78.2	6115.24	12,230.5
68.5	4	274.0	−44.2	1953.64	7814.6
102.5	5	512.5	−10.2	104.04	520.2
136.5	5	682.5	23.8	566.44	2832.2
170.5	3	511.5	57.8	3340.84	10,022.5
204.5	1	204.5	91.8	8427.24	8427.2
	$n = 20$	$\sum xf = 2254$			$\sum (x-\bar{x})^2 f \approx 41{,}847.2$

$$s = \sqrt{\frac{\sum (x-\bar{x})^2 f}{n-1}} = \sqrt{\frac{41{,}847.2}{19}} \approx 46.9$$

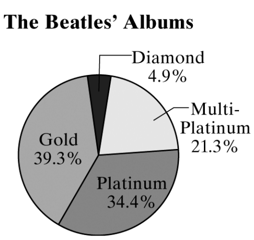
4. 149 is the 16th observation when the data are ordered. The percentile for 149

$$= \frac{\text{number of data entries less than 149}}{\text{total number of data entries}} \cdot 100 = \frac{15}{20} \cdot 100 = 75^{\text{th}} \text{ percentile}$$

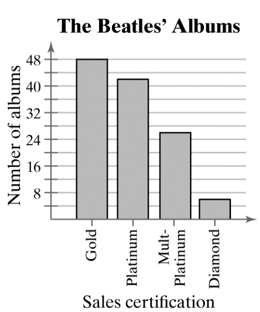
5.

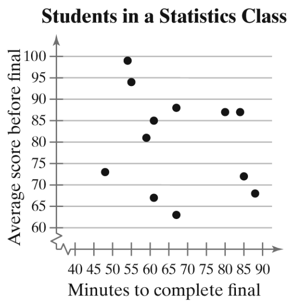
Certification	f	Relative Frequency	Angle
Diamond	6	0.049	18°
Multi-Platinum	26	0.213	77°
Platinum	42	0.344	124°
Gold	48	0.393	141°
	$\sum f = 122$	$\sum \frac{f}{n} \approx 1$	$\sum \quad = 360°$

(a)



(b)



6. 

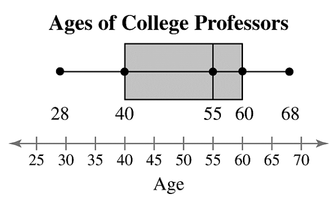
Sample answer: It appears that there is no relation between minutes to complete the final exam and average score before the final exam.

7. (a) Find the five-number summary.

28 30 37 $\underbrace{40}$ 42 46 51 $\underbrace{55}$ 56 58 59 $\underbrace{60}$ 62 65 68

$Q_1 = 40 \qquad Q_2 = 55 \qquad Q_3 = 60$

Min = 28, $Q_1 = 40$, $Q_2 = 55$, $Q_3 = 60$, Max = 68



(b) About 75% of the professors are over the age of 40.

8. (a) $(333.3,\ 354.1) \rightarrow (343.7 - 1(10.4),\ 343.7 + 1(10.4)) \rightarrow (\bar{x} - s,\ \bar{x} + s)$

About 68% of the gestational lengths are between 333.3 and 354.1 days. Thus, about $0.68(208) \approx 141$ gestational lengths are between 333.3 and 354.1 days.

(b) $z = \dfrac{x - \bar{x}}{s} = \dfrac{318.4 - 343.7}{10.4} \approx -2.43$; Since the z-score is about –2.43 (between 2 and 3 standard deviations below the mean), the gestational length is unusual.