

Descriptive Statistics**2.1 FREQUENCY DISTRIBUTIONS AND THEIR GRAPHS****2.1 Try It Yourself Solutions**

1a. The number of classes is 7.

b. $\text{Min} = 26, \text{Max} = 86, \text{Class width} = \frac{\text{Range}}{\text{Number of classes}} = \frac{86 - 26}{7} = 8.57 \Rightarrow 9$

c.

Lower limit	Upper limit
26	34
35	43
44	52
53	61
62	70
71	79
80	88

de.

Class	Frequency, f
26-34	2
35-43	5
44-52	12
53-61	18
62-70	11
71-79	1
80-88	1

2a. See part (b).

b.

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
26-34	2	30	0.04	2
35-43	5	39	0.10	7
44-52	12	48	0.24	19
53-61	18	57	0.36	37
62-70	11	66	0.22	48
71-79	1	75	0.02	49
80-88	1	84	0.02	50
	$\sum f = 50$		$\sum \frac{f}{n} = 1$	

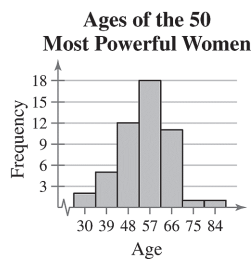
c. *Sample answer:* The most common age bracket for the 50 most powerful women is 53-61. Eighty-six percent of the 50 most powerful women are older than 43. Four percent of the 50 most powerful women are younger than 35.

3a.

Class Boundaries
25.5-34.5
34.5-43.5
43.5-52.5
52.5-61.5
61.5-70.5
70.5-79.5
79.5-88.5

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

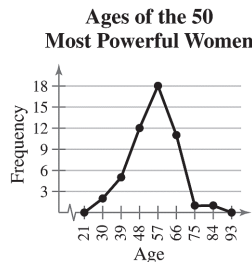
c.



d. Same as 2(c).

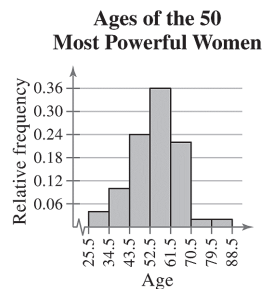
4a. Same as 3(b).

bc.



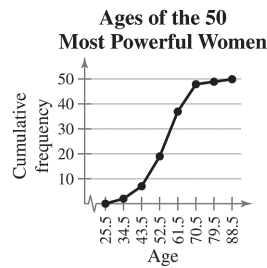
d. The frequency of ages increases up to 57 years old and then decreases.

5abc.



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

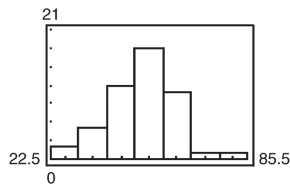
bc.



Sample answer: The greatest increase in cumulative frequency occurs between 52.5 and 61.5

- 7a. Enter data.

b.



2.1 EXERCISE SOLUTIONS

- 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.
- 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.
- Class limits determine which numbers can belong to that class.
Class boundaries are the numbers that separate classes without forming gaps between them.
- 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.
- The sum of the relative frequencies must be 1 or 100% because it is the sum of all portions or percentages of the data.
- A frequency polygon displays frequencies or relative frequencies whereas an ogive displays cumulative frequencies.
- False. Class width is the difference between the lower (or upper limits) of consecutive classes.
- True

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

10. True

$$11. \text{Class width} = \frac{\text{Range}}{\text{Number of classes}} = \frac{64 - 9}{7} \approx 7.9 \Rightarrow 8$$

Lower class limits: 9, 17, 25, 33, 41, 49, 57

Upper class limits: 16, 24, 32, 40, 48, 56, 64

$$12. \text{Class width} = \frac{\text{Range}}{\text{Number of classes}} = \frac{88 - 12}{6} \approx 12.7 \Rightarrow 13$$

Lower class limits: 12, 25, 38, 51, 64, 77

Upper class limits: 24, 37, 50, 63, 76, 89

$$13. \text{Class width} = \frac{\text{Range}}{\text{Number of classes}} = \frac{135 - 17}{8} = 14.75 \Rightarrow 15$$

Lower class limits: 17, 32, 47, 62, 77, 92, 107, 122

Upper class limits: 31, 46, 61, 76, 91, 106, 121, 136

$$14. \text{Class width} = \frac{\text{Range}}{\text{Number of classes}} = \frac{247 - 54}{10} = 19.3 \Rightarrow 20$$

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 = $31 - 20 = 11$

(b) and (c)

Class	Midpoint	Class boundaries
20-30	25	19.5-30.5
31-41	36	30.5-41.5
42-52	47	41.5-52.5
53-63	58	52.5-63.5
64-74	69	63.5-74.5
75-85	80	74.5-85.5
86-96	91	85.5-96.5

16. (a) Class width = $10 - 0 = 10$

(b) and (c)

Class	Midpoint	Class boundaries
0-9	4.5	-0.5-9.5
10-19	14.5	9.5-19.5
20-29	24.5	19.5-29.5
30-39	34.5	29.5-39.5
40-49	44.5	39.5-49.5
50-59	54.5	49.5-59.5
60-69	64.5	59.5-69.5

17.

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

18.

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

19. (a) Number of classes = 7 (b) Least frequency ≈ 10
 (c) Greatest frequency ≈ 300 (d) Class width = 10
20. (a) Number of classes = 7 (b) Least frequency = 1
 (c) Greatest frequency = 23 (d) Class width = 53
21. (a) 50 (b) 345.5-365.5 pounds
22. (a) 50 (b) 64-66 inches
23. (a) 15 (b) 385.5 pounds
 (c) $31 - 6 = 25$ (d) $50 - 42 = 8$
24. (a) 48 (b) 66 inches
 (c) $25 - 5 = 20$ (d) $50 - 44 = 6$
25. (a) Class with greatest relative frequency: 39-40 centimeters
 Class with least relative frequency: 34-35 centimeters
 (b) Greatest relative frequency ≈ 0.25
 Least relative frequency ≈ 0.02
 (c) Approximately 0.08

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

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

(c) Approximately 33%

27. Class with greatest frequency: 29.5-32.5
Classes with least frequency: 11.5-14.5 and 38.5-41.5

28. Class with greatest frequency: 7.75-8.25
Class with least frequency: 6.25-6.75

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

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

Classes with greatest frequency: 0-7, 8-15

Classes with least frequency: 16-23, 24-31, 32-39

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

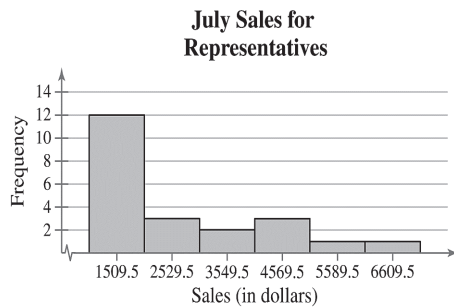
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
30-113	5	71.5	0.1724	5
114-197	7	155.5	0.2414	12
198-281	8	239.5	0.2759	20
282-365	2	323.5	0.0690	22
366-449	3	407.5	0.1034	25
450-533	4	491.5	0.1379	29
	$\sum f = 29$		$\sum \frac{f}{n} = 1$	

Class with greatest frequency: 198-281

Class with least frequency: 282-365

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

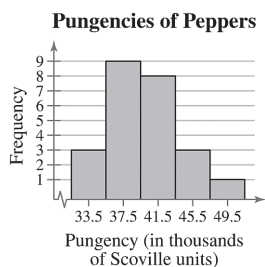
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
1000-2019	12	1509.5	0.5455	12
2020-3039	3	2529.5	0.1364	15
3040-4059	2	3549.5	0.0909	17
4060-5079	3	4569.5	0.1364	20
5080-6099	1	5589.5	0.0455	21
6100-7119	1	6609.5	0.0455	22
	$\sum f = 22$		$\sum \frac{f}{n} \approx 1$	



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

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

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
	$\sum f = 24$		$\sum \frac{f}{n} = 1$	

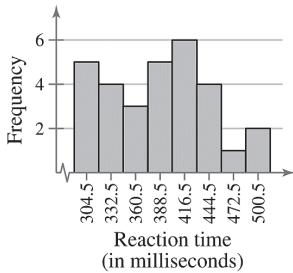


Sample answer: The graph shows that most of the pungencies of the peppers were between 36,000 and 43,000 Scoville units.

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

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

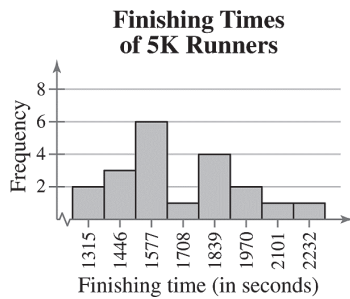
Reaction Times for Females



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

34. Class width = $\frac{\text{Range}}{\text{Number of classes}} = \frac{2296 - 1250}{8} = 130.75 \Rightarrow 131$

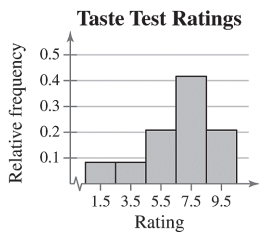
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
1250-1380	2	1315	0.10	2
1381-1511	3	1446	0.15	5
1512-1642	6	1577	0.30	11
1643-1773	1	1708	0.05	12
1774-1904	4	1839	0.20	16
1905-2035	2	1970	0.10	18
2036-2166	1	2101	0.05	19
2167-2297	1	2232	0.05	20
	$\sum f = 20$		$\sum \frac{f}{n} = 1$	



Sample answer: The graph shows that the most frequent finishing times were from 1381 to 1642 seconds and from 1774 to 2035 seconds.

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

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
1-2	2	1.5	0.0833	2
3-4	2	3.5	0.0833	4
5-6	5	5.5	0.2083	9
7-8	10	7.5	0.4167	19
9-10	5	9.5	0.2083	24
	$\sum f = 24$		$\sum \frac{f}{n} \approx 1$	

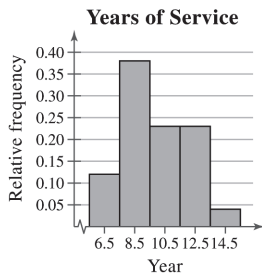


Class with greatest relative frequency: 7-8

Class with least relative frequency: 1-2 and 3-4

36. Class width = $\frac{\text{Range}}{\text{Number of classes}} = \frac{14 - 6}{5} = 1.6 \Rightarrow 2$

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

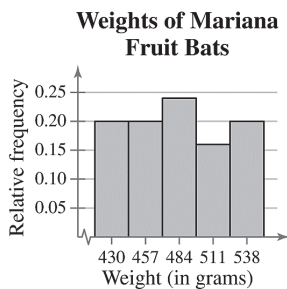


Class with greatest relative frequency: 8-9

Class with least relative frequency: 14-15

37. Class width = $\frac{\text{Range}}{\text{Number of classes}} = \frac{547 - 417}{5} = 26 \Rightarrow 27$

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
417-443	5	430	0.20	5
444-470	5	457	0.20	10
471-497	6	484	0.24	16
498-524	4	511	0.16	20
525-551	5	538	0.20	25
	$\sum f = 25$		$\sum \frac{f}{n} = 1$	

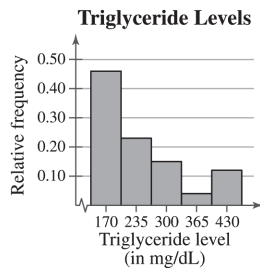


Class with greatest relative frequency: 471-497

Class with least relative frequency: 498-524

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

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
138-202	12	170	0.46	12
203-267	6	235	0.23	18
268-332	4	300	0.15	22
333-397	1	365	0.04	23
398-462	3	430	0.12	26
	$\sum f = 26$		$\sum \frac{f}{n} = 1$	



Class with greatest relative frequency: 138-202

Class with least relative frequency: 333-397

39. Class width = $\frac{\text{Range}}{\text{Number of classes}} = \frac{73 - 52}{6} = 3.5 \Rightarrow 4$

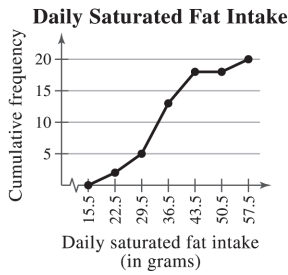
Class	Frequency, f	Relative frequency	Cumulative frequency
52-55	3	0.125	3
56-59	3	0.125	6
60-63	9	0.375	15
64-67	4	0.167	19
68-71	4	0.167	23
72-75	1	0.042	24
	$\sum f = 24$	$\sum \frac{f}{n} \approx 1$	



Location of the greatest increase in frequency: 60-63

40. Class width = $\frac{\text{Range}}{\text{Number of classes}} = \frac{57 - 16}{6} \approx 6.83 \Rightarrow 7$

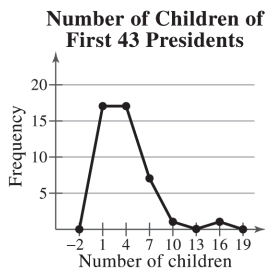
Class	Frequency, f	Relative frequency	Cumulative frequency
16-22	2	0.10	2
23-29	3	0.15	5
30-36	8	0.40	13
37-43	5	0.25	18
44-50	0	0.00	18
51-57	2	0.10	20
	$\sum f = 20$	$\sum \frac{f}{n} = 1$	



Location of the greatest increase in frequency: 30-36

41. Class width = $\frac{\text{Range}}{\text{Number of classes}} = \frac{15 - 0}{6} = 2.5 \Rightarrow 3$

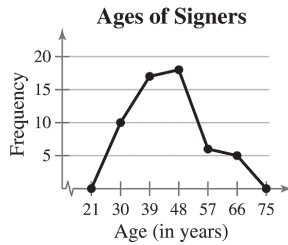
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
0-2	17	1	0.3953	17
3-5	17	4	0.3953	34
6-8	7	7	0.1628	41
9-11	1	10	0.0233	42
12-14	0	13	0.0000	42
15-17	1	16	0.0233	43
	$\sum f = 43$		$\sum \frac{f}{N} = 1$	



Sample answer: The graph shows that most of the first 43 presidents had fewer than 6 children.

42. Class width = $\frac{\text{Range}}{\text{Number of classes}} = \frac{70 - 26}{5} = 8.8 \Rightarrow 9$

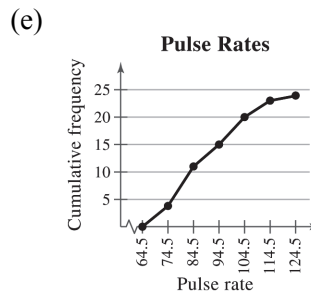
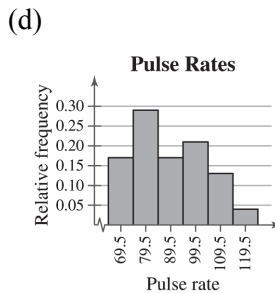
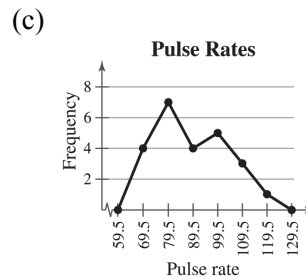
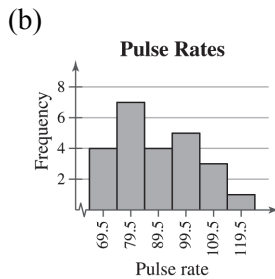
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
26-34	10	30	0.1786	11
35-43	17	39	0.3036	27
44-52	18	48	0.3214	45
53-61	6	57	0.1071	51
62-70	5	66	0.0893	56
	$\sum f = 56$		$\sum \frac{f}{N} \approx 1$	



Sample answer: The graph shows that majority of signers were from 35 to 52 years old.

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

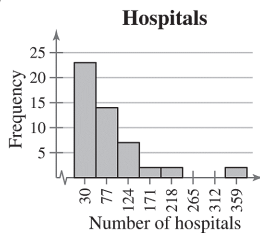
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
65-74	4	69.5	0.1667	4
75-84	7	79.5	0.2917	11
85-94	4	89.5	0.1667	15
95-104	5	99.5	0.2083	20
105-114	3	109.5	0.1250	23
115-124	1	119.5	0.0417	24
	$\sum f = 24$		$\sum \frac{f}{N} \approx 1$	



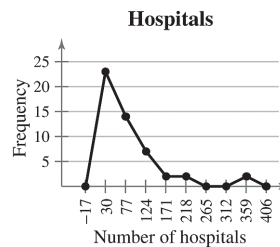
44. (a) Class width = $\frac{\text{Range}}{\text{Number of classes}} = \frac{377 - 7}{8} = 46.25 \Rightarrow 47$

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
7-53	23	30	0.46	23
54-100	14	77	0.28	37
101-147	7	124	0.14	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
	$\sum f = 50$		$\sum \frac{f}{N} = 1$	

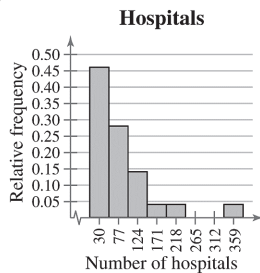
(b)



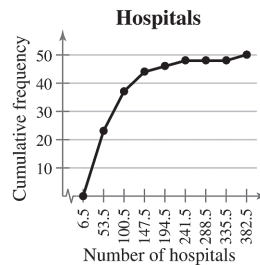
(c)



(d)

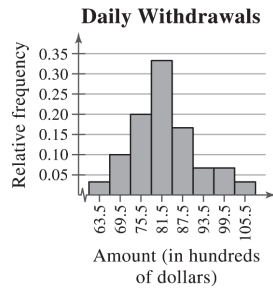


(e)



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

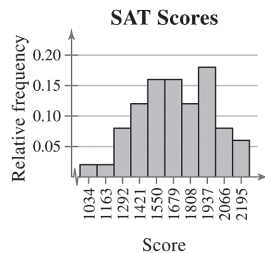
Class	Frequency, f	Midpoint	Relative frequency
61-66	1	63.5	0.033
67-72	3	69.5	0.100
73-78	6	75.5	0.200
79-84	10	81.5	0.333
85-90	5	87.5	0.167
91-96	2	93.5	0.067
97-102	2	99.5	0.067
103-108	1	105.5	0.033
	$\sum f = 30$		$\sum \frac{f}{N} = 1$



- (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.

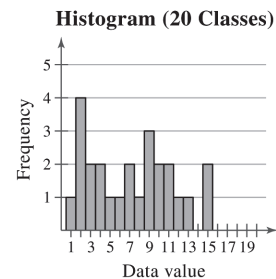
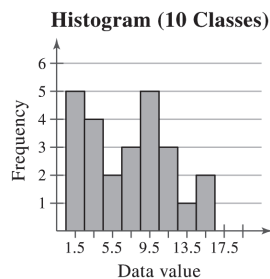
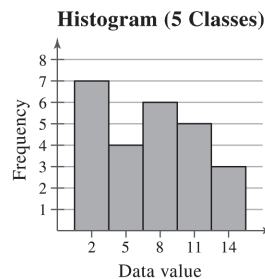
46. (a) Class width = $\frac{\text{Range}}{\text{Number of classes}} = \frac{2250 - 970}{10} = 128 \Rightarrow 129$

Class	Frequency, f	Midpoint	Relative frequency
970-1098	1	1034	0.02
1099-1227	1	1163	0.02
1228-1356	4	1292	0.08
1357-1485	6	1421	0.12
1486-1614	8	1550	0.16
1615-1743	8	1679	0.16
1744-1872	6	1808	0.12
1873-2001	9	1937	0.18
2002-2130	4	2066	0.08
2131-2259	3	2195	0.06
	$\sum f = 50$		$\sum \frac{f}{N} = 1$



- (b) 62%; The proportion of scores greater than or equal to 1610 is 0.62.
 (c) A score of 1357 or above, because the sum of the relative frequencies of the class starting with 1357 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

1a.	2	b.	2	6
3		3	1	5 7
4		4	3	3 3 4 5 7 8 8 9
5		5	0	1 1 1 1 2 4 4 4 4 5 5 5 6 7 7 7 8 8 8 8 9 9 9
6		6	2	2 3 4 5 5 5 6 6 7 7
7		7	2	
8		8	6	

Key $3|6 = 36$

c. *Sample answer:* Most of the most powerful women are between 40 and 70 years old.

2a, b.

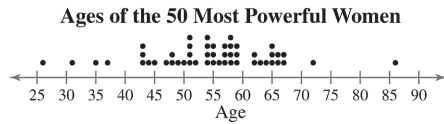
2	
2	6
3	1
3	5 7
4	3 3 3 4
4	5 7 8 8 9
5	0 1 1 1 1 2 4 4 4 4
5	5 5 5 6 7 7 7 8 8 8 8 9 9 9
6	2 2 3 4
6	5 5 5 6 6 7 7
7	2
7	
8	
8	6

Key $3|5 = 35$

c. *Sample answer:* Most of the 50 most powerful women are older than 50.

3a. Use the age for the horizontal axis.

b.

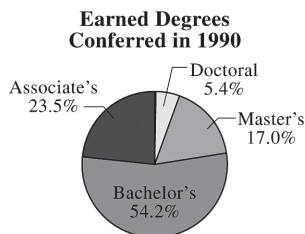


c. *Sample answer:* Most of the ages cluster between 43 and 67 years old. The age of 86 years old is an unusual data entry.

4a.

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°
	$\sum f = 1940$	$\sum \frac{f}{N} \approx 1$	$\sum = 360^\circ$

b.

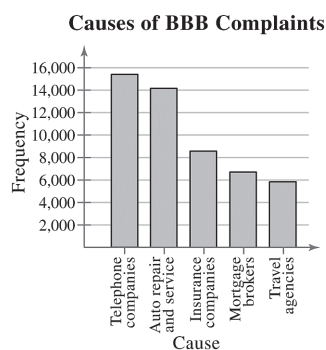


c. From 1990 to 2011, as percentages of total degrees conferred, associate's degrees increased by 3%, bachelor's degrees decreased by 5.9%, master's degrees increased by 3.6%, and doctoral degrees decreased by 0.8%.

5a.

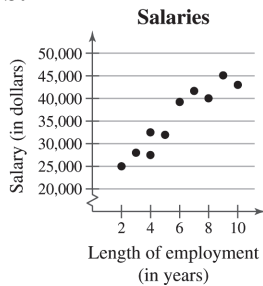
Cause	Frequency, f
Auto repair and service	14,156
Insurance companies	8,568
Mortgage brokers	6,712
Telephone companies	15,394
Travel agencies	5,841

b.



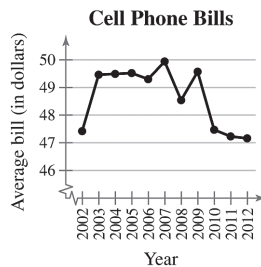
- c. *Sample answer:* Telephone companies and auto repair and service account for over half of all complaints received by the BBB.

6a, b.



- c. It appears that the longer an employee is with the company, the larger the employee's salary will be.

7ab.



- c. The average bill increased from 2002 to 2003, then it hovered from 2003 to 2009, and decreased from 2009 to 2012.

2.2 EXERCISE SOLUTIONS

- Quantitative: stem-and-leaf plot, dot plot, histogram, time series chart, scatter plot.
Qualitative: pie chart, Pareto chart
- 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.
- 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.
- 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.
- b
- d
- a
- c
- 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:* Users spend the most amount of time on Facebook and the least amount of time on LinkedIn.

14. *Sample answer:* Motor vehicle thefts decreased from 2006 and 2011.

15. *Sample answer:* Tailgaters irk drivers the most, while too-cautious drivers irk drivers the least.

16. *Sample answer:* Food is the most costly aspect of pet care. The actual price of the pet is the least costly aspect of pet care.

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. Apple Prices (in cents per pound) Key: 25|4 = 25.4

```

25|4 8
26|1 1 3 4 5 5 6 7 9
27|1 3 4 4 6 6 6 9
28|0 0 2 3 3 5 6
29|1 8
    
```

Sample answer: Most farmers charge 26 to 28 cents per pound of apples.

21. Ages of Highest-Paid CEOs Key: 5|0 = 50

```

5|0 2 3
55|5 5 6 6 7 7 8 8 8 8 9 9 9 9 9
60|1 1 1 3 4
65|5 5 6 6 7
72|
    
```

Sample answer: Most of the highest-paid CEOs have ages that range from 55 and 64 years old.

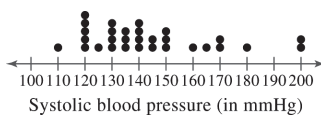
22. Super Bowl Winning Scores Key: 1|4 = 14

```

1|4
16|6 6 6 7
20|0 0 0 1 1 1 3 3 4 4 4
26|7 7 7 7 7 9
30|1 1 1 1 2 2 3 4 4 4
35|5 5 5 7 8 8 9
42|
46|8 9
52|
55|
    
```

23.

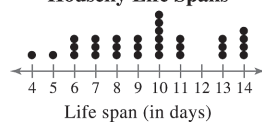
Systolic Blood Pressures



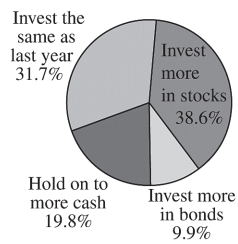
Sample answer: Systolic blood pressure tends to be between 120 and 150 millimeters of mercury.

24.

Housefly Life Spans



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

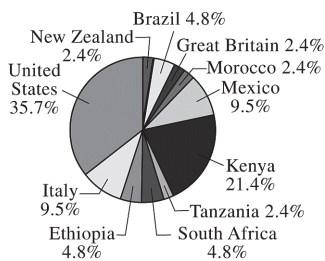
25. How Will You Invest in 2013?

Sample answer: The majority of people will either invest more in stocks or invest the same as last year.

26.

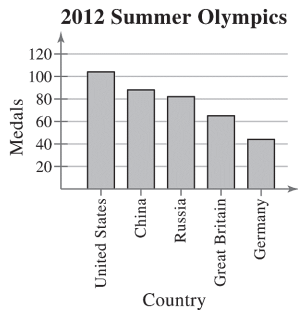
Category	Frequency, f	Relative Frequency	Angle
United States	15	0.357	129°
Italy	4	0.095	34°
Ethiopia	2	0.048	17°
South Africa	2	0.048	17°
Tanzania	1	0.024	9°
Kenya	9	0.214	77°
Mexico	4	0.095	34°
Morocco	1	0.024	9°
Great Britain	1	0.024	9°
Brazil	2	0.048	17°
New Zealand	1	0.024	9°
	$\sum f = 42$	$\sum \frac{f}{N} \approx 1$	$\sum \approx 360^\circ$

**Marathon Winners'
Countries of Origin**



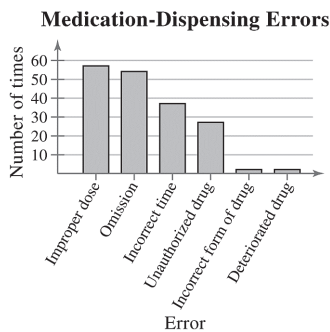
Sample answer: Most of the New York City Marathon winners are from the United States and Kenya.

27.



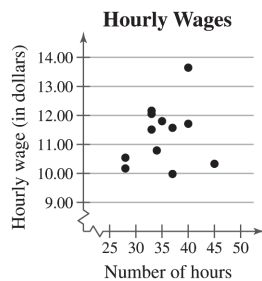
Sample answer: The United States won the most medals out of the five countries and Germany won the least.

28.



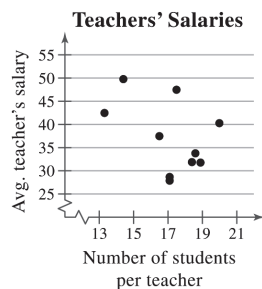
Sample answer: The greatest types of medication-dispensing errors are improper doses and omissions.

29.



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

30.



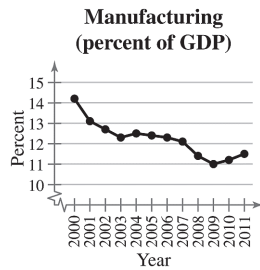
Sample answer: It appears that there is no relation between a teacher's average salary and the number of students per teacher.

31.



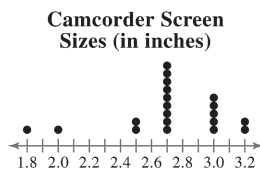
Sample answer: The number of motorcycle registrations has increased from 2000 to 2011.

32.



Sample answer: The percentage of the U.S. gross domestic product that comes from the manufacturing sector has decreased from 2000 to 2009.

33.



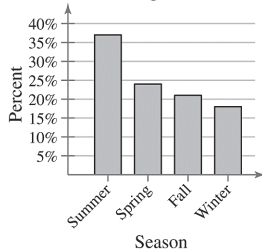
The stem-and-leaf plot helps you see that most values are from 2.5 to 3.2. The dot plot helps you see that the values 2.7 and 3.0 occur most frequently, with 2.7 occurring most frequently.

34. **Heights (in inches)** Key: $6|9 = 69$

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

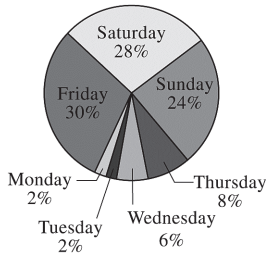
The dot plot helps you see that the data are clustered from 78 to 83 with 78 being the most frequent value. The stem-and-leaf plot helps you see that most values are in the 70s and 80s.

35. Favorite Season of U.S. Adults Ages 18 to 29



The pie chart helps you to see the percentages as parts of a whole, with summer being the largest. It also shows that while summer is the largest percentage, it only makes up about one-third of the pie chart. That means that about two-thirds of U.S. adults ages 18 to 29 prefer a season other than summer. This means it would not be a fair statement to say that most U.S. adults ages 18 to 29 prefer summer. The Pareto chart helps you to see the rankings of the seasons. It helps you to see that the favorite seasons in order from greatest to least percentage are summer, spring, fall, and winter.

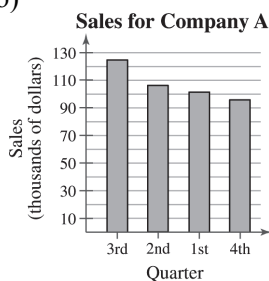
36. Favorite Day of The Week



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) 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.
- (c) There are 26 participants in the 3:00 P.M. class and there are 30 participants in the 8:00 P.M. class.
- (d) *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.

2.3 MEASURES OF CENTRAL TENDENCY

2.3 Try It Yourself Solutions

- 1a. $\sum x = 74 + 78 + 81 + 87 + 81 + 80 + 77 + 80 + 85 + 78 + 80 + 83 + 75 + 81 + 73 = 1193$
- b. $\bar{x} = \frac{\sum x}{n} = \frac{1193}{15} \approx 79.5$
- c. The mean height of the player is about 79.5 inches.

- 2a. 18, 18, 19, 19, 20, 21, 21, 21, 21, 23, 24, 24, 26, 27, 27, 29, 30, 30, 30, 33, 33, 34, 35, 38
- b. median = 24
- c. The median age of the sample of fans at the concert is 24.

- 3a. 10, 50, 50, 70, 70, 80, 100, 100, 120, 130
- b. median = $\frac{70 + 80}{2} = \frac{150}{2} = 75$
- c. The median price of the sample of digital photo frames is \$75.

- 4a. 324, 385, 450, 450, 462, 475, 540, 540, 564, 618, 624, 638, 670, 670, 670, 705, 720, 723, 750, 750, 825, 830, 912, 975, 980, 980, 1100, 1260, 1420, 1650
- b. The price that occurs with the greatest frequency is \$670 per square foot.
- c. The mode of the prices for the sample of South Beach, FL condominiums is \$670 per square foot.

- 5a. “Better prices” occurs with the greatest frequency (399).
- b. In this sample, there were more people who shop online for better prices than for any other reason.

- 6a. $\bar{x} = \frac{\sum x}{n} = \frac{410}{19} \approx 21.6$
 median = 21
 mode = 20
- b. The mean in Example 6 ($\bar{x} \approx 23.8$) was heavily influenced by the entry 65. Neither the median nor the mode was affected as much by the entry 65.

7a, b.

Source	Score, x	Weight, w	$x \cdot w$
Test mean	86	0.50	43.0
Midterm	96	0.15	14.4
Final exam	98	0.20	19.6
Computer lab	98	0.10	9.8
Homework	100	0.05	5.0
		$\sum w = 1$	$\sum (x \cdot w) = 91.8$

$$\text{c. } \bar{x} = \frac{\sum (x \cdot w)}{\sum w} = \frac{91.8}{1} = 91.8$$

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

8a, b, c.

Class	Midpoint, x	Frequency, f	$x \cdot f$
26-34	30	2	60
35-43	39	5	195
44-52	48	12	576
53-61	57	18	1026
62-70	66	11	726
71-79	75	1	75
80-88	84	1	84
		$N = 50$	$\sum (x \cdot f) = 2742$

$$\text{d. } \mu = \frac{\sum (x \cdot f)}{N} = \frac{2742}{50} \approx 54.8$$

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 of dollars and is skewed right due to the few executives that make a much higher salary than the majority of the employees.
15. (12), because the distribution has a maximum value of 90 and is skewed left due to a few students scoring much lower than the majority of the students.
16. (10), because the distribution is approximately symmetric and the weights range from 80 to 160 pounds.

$$17. \bar{x} = \frac{\sum x}{n} = \frac{192}{13} \approx 14.8$$

12 12 13 14 14 15 **15** 15 16 16 16 16 18

median = 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. $\bar{x} = \frac{\sum x}{n} = \frac{8249}{7} \approx 1178.4$

818 1125 1155 **1229** 1275 1277 1370

median = 1229

mode = none

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

20. $\bar{x} = \frac{\sum x}{n} = \frac{414}{9} = 46$

36 38 40 43 **43** 49 50 52 63

median = 43

mode = 43 (occurs 2 times)

$$21. \bar{x} = \frac{\sum x}{n} = \frac{603}{14} \approx 43.1$$

39 40 41 42 42 42 44 44 44 44 44 45 45 47

$$\text{median} = \frac{44 + 44}{2} = 44$$

mode = 44 (occurs 5 times)

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

154 171 173 181 184 188 203 235 240 275

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

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

$$23. \bar{x} = \frac{\sum x}{n} = \frac{481}{16} \approx 30.1$$

17 21 21 25 26 30 31 31 31 34 34 34 35 36 37 38

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

mode = 31 and 34 (both occur 3 times)

$$24. \bar{x} = \frac{\sum x}{n} = \frac{1223}{20} \approx 61.2$$

12 18 26 28 31 33 40 44 45 49 61 63 75 80 80 89 96 103 125 125

$$\text{median} = \frac{49 + 61}{2} = 55$$

mode = 80, 125

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

$$25. \bar{x} = \frac{\sum x}{n} = \frac{83}{5} = 16.6$$

1.0 10.0 **15.0** 25.5 31.5
 ▼ median = 43

mode = none

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

$$26. \bar{x} = \frac{\sum x}{n} = \frac{197.5}{14} \approx 14.11$$

1.5 2.5 2.5 5.0 10.5 11.0 13.0 15.5 16.5 17.5 20.0 26.5 27.0 28.5

$$\text{median} = \frac{13.0 + 15.5}{2} = 14.25$$

mode = 2.5 (occurs 2 times)

The mode does not represent the center of the data set because 2.5 is much smaller than most of the data in the set.

27. $\bar{x}$ is not possible (nominal data)
 median = not possible (nominal data)
 mode = “Eyeglasses”
 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 = “Money needed”
 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 = “on Facebook, find it valuable”
 The mean and median cannot be found because the data are at the nominal level of measurement.

31. $\bar{x} = \frac{\sum x}{n} = \frac{835}{28} \approx 29.8$
 6 7 12 15 18 19 20 24 24 24 25 28 29 32 32 33 35 35 36 38 39 40 41 42 47 48 51

$$\text{median} = \frac{32 + 32}{2} = 32$$

mode = 24, 35 (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 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 value 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. $\bar{x} = \frac{\sum x}{n} = \frac{3110}{15} \approx 207.3$

170 180 190 200 200 210 210 **210** 210 210 220 220 220 220 240

↖ median = 210

mode = 210 (occurs 5 times)

35. 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.

36. 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.

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

38. Mean, because the data are symmetric.

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
Quiz	80	0.35	28
Project	100	0.20	20
Speech	90	0.15	13.5
Final exam	93	0.25	23.25
		$\sum w = 1$	$\sum (x \cdot w) = 89$

$$\bar{x} = \frac{\sum (x \cdot w)}{\sum w} = \frac{89}{1} = 89$$

42.

Source	Score, x	Weight, w	$x \cdot w$
Article reviews	95	10%	9.5
Quizzes	100	10%	10
Midterm exam	89	30%	26.7
Student lecture	100	10%	10
Final exam	92	40%	36.8
		$\sum w = 100\%$	$\sum (x \cdot w) = 93$

$$\bar{x} = \frac{\sum (x \cdot w)}{\sum w} = \frac{93}{1} = 93$$

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	$x \cdot w$
\$759	15	11,385
\$1985	5	9925
\$1410	5	7050
\$348	6	2088
	$\sum w = 31$	$\sum (x \cdot w) = 30,448$

$$\bar{x} = \frac{\sum (x \cdot w)}{\sum w} = \frac{30,448}{31} \approx \$982.20$$

45.

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
		$\sum w = 15$	$\sum (x \cdot w) = 42$

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

46.

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

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

47.

Source	Score, x	Weight, w	$x \cdot w$
Homework	85	0.05	4.25
Quiz	80	0.35	28
Project	100	0.20	20
Speech	90	0.15	13.5
Final exam	85	0.25	21.25
		$\sum w = 1$	$\sum (x \cdot w) = 87$

$$\bar{x} = \frac{\sum (x \cdot w)}{\sum w} = \frac{87}{1} = 87$$

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	3	127.5
46-51	48.5	1	48.5
		$n = 24$	$\sum (x \cdot f) = 702$

$$\bar{x} = \frac{\sum (x \cdot f)}{n} = \frac{702}{24} \approx 29.3 \text{ miles per gallon}$$

51.

Class	Midpoint, x	Frequency, f	$x \cdot f$
0-9	4.5	44	198.0
10-19	14.5	66	957.0
20-29	24.5	32	784.0
30-39	34.5	53	1828.5
40-49	44.5	35	1557.5
50-59	54.5	31	1689.5
60-69	64.5	23	1483.5
70-79	74.5	13	968.5
80-89	84.5	2	169.0
		$n = 299$	$\sum(x \cdot f) = 9,635.5$

$$\bar{x} = \frac{\sum(x \cdot f)}{n} = \frac{9,635.5}{299} \approx 32.2 \text{ years old}$$

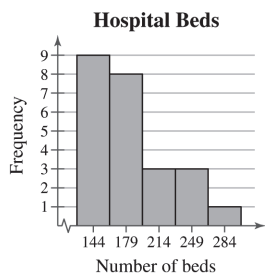
52.

Class	Midpoint, x	Frequency, f	$x \cdot f$
0-9	4.5	30	135.0
10-19	14.5	28	406.0
20-29	24.5	17	416.5
30-39	34.5	22	759.0
40-49	44.5	23	1023.5
50-59	54.5	46	2507.0
60-69	64.5	37	2386.5
70-79	74.5	18	1341.0
80-89	84.5	4	338.0
		$n = 225$	$\sum(x \cdot f) = 9312.5$

$$\bar{x} = \frac{\sum(x \cdot f)}{n} = \frac{9312.5}{225} \approx 41.4 \text{ years old}$$

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

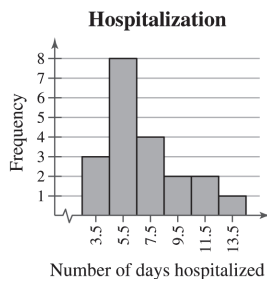
Class	Midpoint	Frequency, f
127-161	144	9
162-196	179	8
197-231	214	3
232-266	249	3
267-301	284	1
		$\sum f = 24$



Shape: Positively skewed

54. Class width = $\frac{\text{Range}}{\text{Number of classes}} = \frac{14 - 3}{6} \approx 1.83 \Rightarrow 2$

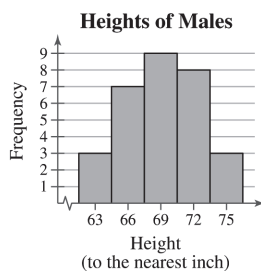
Class	Midpoint	Frequency, f
3-4	3.5	3
5-6	5.5	8
7-8	7.5	4
9-10	9.5	2
11-12	11.5	2
13-14	13.5	1
		$\Sigma f = 20$



Shape: Positively skewed

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

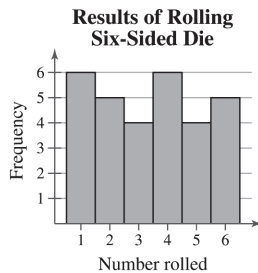
Class	Midpoint	Frequency, f
62-64	63	3
65-67	66	7
68-70	69	9
71-73	72	8
74-76	75	3
		$\Sigma f = 30$



Shape: Symmetric

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

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



Shape: Uniform

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

5.59 5.99 6.00 6.02 6.03 6.40

$$\text{median} = \frac{6.00 + 6.02}{2} = 6.01$$

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

5.59 5.99 6.00 6.02 6.03 6.04

$$\text{median} = \frac{6.00 + 6.02}{2} = 6.01$$

(c) The mean was affected more.

58. (a) $\bar{x} = \frac{\Sigma x}{n} = \frac{1104.3}{19} \approx 58.12$

10.9 13.8 14.2 16.0 24.4 25.9 27.5 27.8 29.9 **31.2** 42.4 42.9 43.4 49.2 55.9 65.7
103.9 198.4 280.9

↖
median = 31.2

(b) $\bar{x} = \frac{\Sigma x}{n} = \frac{823.4}{18} \approx 45.74$

10.9 13.8 14.2 16.0 24.4 25.9 27.5 27.8 29.9 31.2 42.4 42.9 43.4 49.2 55.9 65.7 103.9 198.4

$$\text{median} = \frac{29.9 + 31.2}{2} = 30.55$$

The mean was affected more.

$$(c) \bar{x} = \frac{\sum x}{n} = \frac{1125.8}{20} = 56.29$$

10.9 13.8 14.2 16.0 21.5 24.4 25.9 27.5 27.8 29.9 31.2 42.4 42.9 43.4 49.2 55.9 65.7 103.9 198.4 280.9

$$\text{median} = \frac{29.9 + 31.2}{2} = 30.55$$

The mean was affected more.

59. Clusters around 16-21 and around 36

60. Cluster around 18-27, gap between 27 and 72, outlier at 72

61. Sample answer: Option 2; The two clusters represent different types of vehicles which can be more meaningfully analyzed separately. For instance, suppose the mean gas mileage for cars is very far from the mean gas mileage for trucks, vans, and SUVs. Then, the mean gas mileage for all of the vehicles would be somewhere in the middle and would not accurately represent the gas mileages of either group of vehicles.

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

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)

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.

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

10% trimmed mean ≈ 49.2

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

median = 46.5

mode = 36, 37, 51

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

- (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

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

- b. Range = Max – Min = 58 – 23 = 35, or \$35,000

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

- 2ab.
- $\mu = 41.5$
- , 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
$\sum x = 415$	$\sum (x - \mu) = 0$	$\sum (x - \mu)^2 = 1102.5$

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

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

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

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

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

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

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

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

4a. Enter the data in a computer or a calculator.

b. $\bar{x} \approx 22.1$, $s \approx 5.3$

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

b.

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

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

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

6a. $67.1 - 64.2 = 2.9 = 1$ standard deviation

b. 34%

c. Approximately 34% of women ages 20-29 are between 64.2 and 67.1 inches tall.

7a. $35.3 - 2(21.1) = -6.9$

Because -6.9 does not make sense for an age, use 0.

b. $35.3 + 2(21.1) = 77.5$

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

At least 75% of the data lie within 2 standard deviations of the mean. At least 75% of the population of Alaska is between 0 and 77.5 years old.

8a.

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

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

c.

$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
-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
		$\sum (x - \bar{x})^2 f = 106.5$

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

9a.

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
		$n = 1000$	$\sum xf = 195,535$

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

56 CHAPTER 2 | DESCRIPTIVE STATISTICS

c.

$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
-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
		$\sum (x - \bar{x})^2 f = 28,715,797.5$

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

10a. Los Angeles: $\bar{x} \approx 31.0$, $s \approx 12.6$ Dallas/Fort Worth: $\bar{x} \approx 22.1$, $s \approx 5.3$ b. Los Angeles: $CV = \frac{s}{\bar{x}} = \frac{12.6}{31.0} \cdot 100\% \approx 40.6\%$ Dallas/Fort Worth: $CV = \frac{s}{\bar{x}} = \frac{5.3}{22.1} \cdot 100\% \approx 24.0\%$

c. The office rental rates are more variable in Los Angeles than in Dallas/Fort Worth.

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: dollars²). 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.