CHAPTER 2

Descriptive Statistics

2.1 FREQUENCY DISTRIBUTIONS AND THEIR GRAPHS

2.1 Try It Yourself Solutions

1a. The number of classes is 8.

b. Min = 35, Max = 89, Class width = $\frac{\text{Range}}{\text{Number of classes}} = \frac{89-35}{8} = 6.75 \Rightarrow 7$

c.

Lower limit	Upper limit
35	41
42	48
49	55
56	62
63	69
70	76
77	83
84	90

d. See part (e).

e.

Class	Frequency, f
35-41	2
42-48	5
49-55	7
56-62	7
63-69	10
70-76	5
77-83	8
84-90	6

2a. See part (b).

b.

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
35-41	2	38	0.04	2
42-48	5	45	0.10	7
49-55	7	52	0.14	14
56-62	7	59	0.14	21
63-69	10	66	0.20	31
70-76	5	73	0.10	36
77-83	8	80	0.16	44
84-90	6	87	0.12	50
	$\sum f = 50$		$\sum \frac{f}{n} = 1$	

c. 72% of the 50 richest people are older than 55.
4% of the 50 richest people are younger than 42.
The most common age bracket for the 50 richest people is 63-69.

3a.

Class Boundaries
34.5-41.5
41.5-48.5
48.5-55.5
55.5-62.5
62.5-69.5
69.5-76.5
76.5-83.5
83.5-90.5

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. 72% of the 50 richest people are older than 55.
4% of the 50 richest people are younger than 42.
The most common age bracket for the 50 richest people is 63-69.

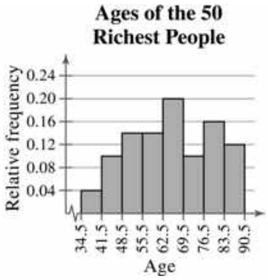
4a. Use class midpoints for the horizontal scale and frequency for the vertical scale. (Class boundaries can also be used for the horizontal scale.)

b. See part (c).

c.

d. The frequency of ages increases up to 66 and then decreases.

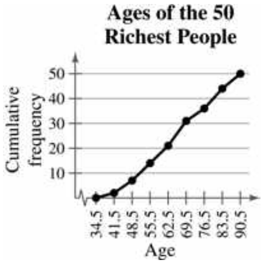
5abc.



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

b. See part (c).

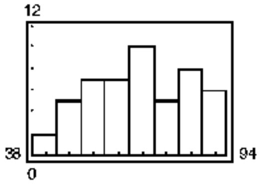
c.



d. Approximately 40 of the 50 richest people are 80 years or younger.

e. Answers will vary.

7ab.



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

12. 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. 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. 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

15a. Class width = 31 – 20 =11

b. and c.

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

16a. Class width = 10 – 0 = 10

b. and c.

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

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
	$\sum f = 1220$		$\sum \frac{f}{n} = 1$	

19a. Number of classes = 7
b. Least frequency ≈ 10
c. Greatest frequency ≈ 300
d. Class width = 10

20a. Number of classes = 7
b. Least frequency ≈ 100
c. Greatest frequency ≈ 900
d. Class width = 5

21a. 50
b. 22.5-23.5 pounds

22a. 50
b. 64-66 inches

23a. 42
b. 29.5 pounds
c. 35
d. 2
24a. 48
b. 66 inches
c. 20
d. 6

25a. Class with greatest relative frequency: 8-9 inches
Class with least relative frequency: 17-18 inches

b. Greatest relative frequency ≈ 0.195
Least relative frequency ≈ 0.005

c. Approximately 0.015

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

b. Greatest relative frequency ≈ 40%
Least relative frequency ≈ 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. $\text{Class width} = \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{39-0}{5} = 7.8 \Rightarrow 8$

Class	**Frequency, f**	**Midpoint**	**Relative frequency**	**Cumulative frequency**
0-7	8	3.5	0.32	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 $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{530-30}{6} \approx 83.3 \Rightarrow 84$

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
30-113	5	71.5	0.1724	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

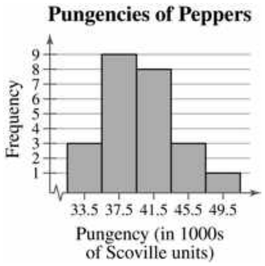
31. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{7119-1000}{6} \approx 1019.83 \Rightarrow 1020$

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
	$\sum f = 22$		$\sum \frac{f}{n} \approx 1$	

The graph shows that most of the sales representatives at the company sold between \$1000 and \$2019. (Answers will vary.)

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

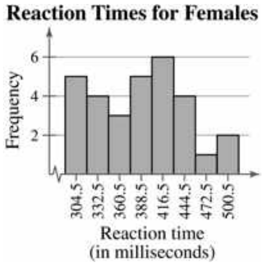
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
32-35	3	33.5	0.1250	3
36-39	9	37.5	0.3750	12
40-43	8	41.5	0.3333	20
44-47	3	45.5	0.1250	23
48-51	1	49.5	0.0417	24
	$\sum f = 24$		$\sum \frac{f}{n} = 1$	



The graph shows that most of the pungencies of the peppers were between 36 and 43 Scoville units. (Answers will vary.)

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

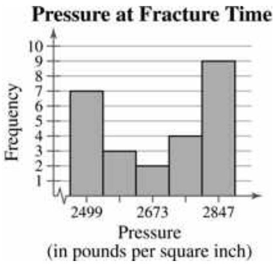
Class	Frequency, *f*	Midpoint	Relative frequency	Cumulative frequency
291-318	5	304.5	0.1667	5
319-346	4	332.5	0.1333	9
347-374	3	360.5	0.1000	12
375-402	5	388.5	0.1667	17
403-430	6	416.5	0.2000	23
431-458	4	444.5	0.1333	27
459-486	1	472.5	0.0333	28
487-514	2	500.5	0.0667	30
	$\sum f = 30$		$\sum \frac{f}{n}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



The graph shows that the most frequent reaction times were between 403 and 430 milliseconds. (Answers will vary.)

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

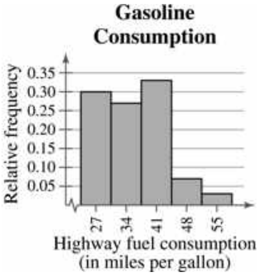
Class	Frequency, *f*	Midpoint	Relative frequency	Cumulative frequency
2456-2542	7	2499	0.28	7
2543-2629	3	2586	0.12	10
2630-2716	2	2673	0.08	12
2717-2803	4	2760	0.16	16
2804-2890	9	2847	0.36	25
	$\sum f = 25$		$\sum \frac{f}{n} = 1$	



The graph shows that the most common pressures at fracture time were between 2804 and 2890 pounds per square inch. (Answers will vary.)

35. Class width = $\dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{55-24}{5} = 6.2 \Rightarrow 7$

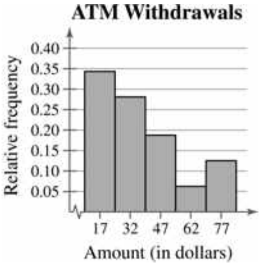
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
24-30	9	27	0.30	9
31-37	8	34	0.27	17
38-44	10	41	0.33	27
45-51	2	48	0.07	29
52-58	1	55	0.03	30
	$\sum f = 30$		$\sum \frac{f}{n} = 1$	



Class with greatest relative frequency: 38-44
Class with least relative frequency: 52-58

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

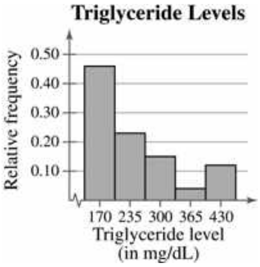
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
10-24	11	17	0.3438	11
25-39	9	32	0.2813	20
40-54	6	47	0.1875	26
55-69	2	62	0.0625	28
70-84	4	77	0.1250	32
	$\sum f = 32$		$\sum \frac{f}{n} \approx 1$	



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

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

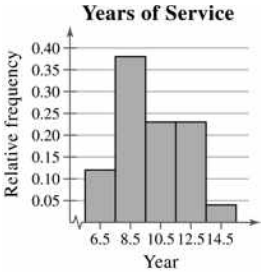
Class	**Frequency, *f***	**Midpoint**	**Relative frequency**	**Cumulative frequency**
138-202	12	170	0.46	12
203-267	6	235	0.23	18
268-332	4	300	0.15	22
333-397	1	365	0.04	23
398-462	3	430	0.12	26
	$\sum f = 26$		$\sum \frac{f}{n} = 1$	



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

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

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

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

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

Location of the greatest increase in frequency: 60-63

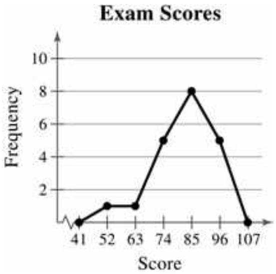
40. Class width = $\dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{57-16}{6} \approx 6.83 \Rightarrow 7$

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 = $\dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{98-47}{5} = 10.2 \Rightarrow 11$

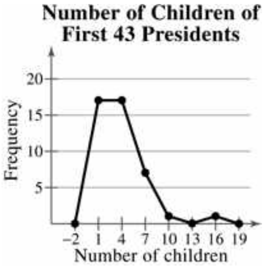
Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
47-57	1	52	0.05	1
58-68	1	63	0.05	2
69-79	5	74	0.25	7
80-90	8	85	0.40	15
91-101	5	96	0.25	20
	$\sum f = 20$		$\sum \frac{f}{N} = 1$	



The graph shows that the most frequent exam scores were between 80 and 90. (Answers will vary.)

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

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
	$\sum f = 43$		$\sum \frac{f}{N} = 1$	

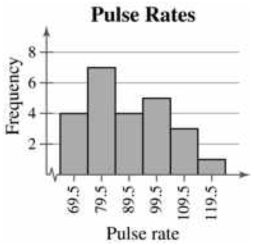


The graph shows that most of the first 43 presidents had fewer than 6 children. (Answers will vary.)

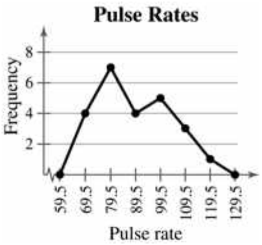
43a. Class width $= \dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{120-65}{6} \approx 9.2 \Rightarrow 10$

Class	Frequency, f	Midpoint	Relative frequency	Cumulative frequency
65-74	4	69.5	0.17	4
75-84	7	79.5	0.29	11
85-94	4	89.5	0.17	15
95-104	5	99.5	0.21	20
105-114	3	109.5	0.13	23
115-124	1	119.5	0.04	24
	$\sum f = 24$		$\sum \frac{f}{N} \approx 1$	

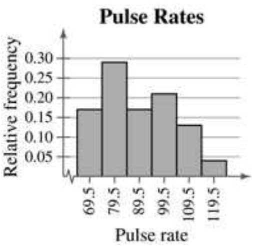
b.



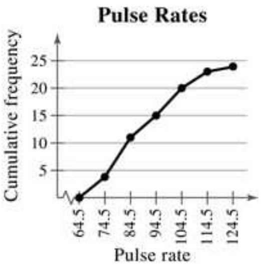
c.



d.



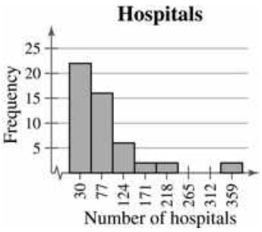
e.



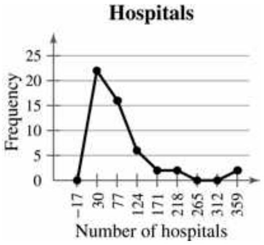
44a. Class width = $\dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{378-7}{8} = 46.375 \Rightarrow 47$

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	48
289-335	0	312	0	48
336-382	2	359	0.04	50
	$\sum f = 50$		$\sum \frac{f}{N} = 1$	

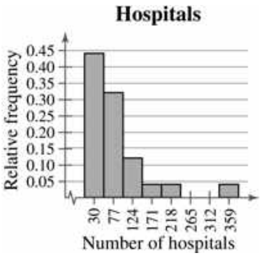
b.



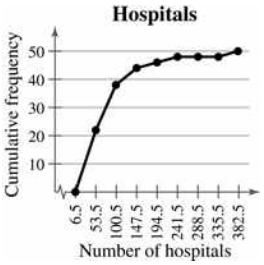
c.



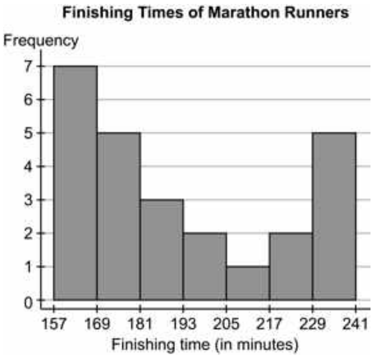
d.



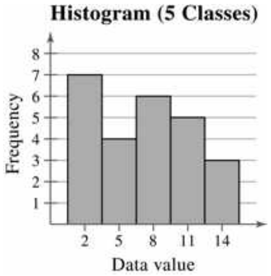
e.

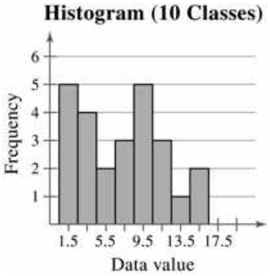


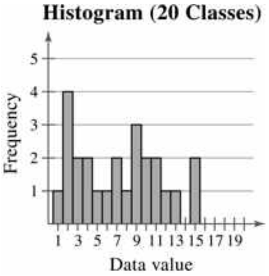
45.



46.



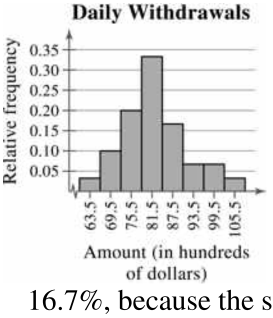




In general, a greater number of classes better preserves the actual values of the data set but is not as helpful for observing general trends and making conclusions. In choosing the number of classes, an important consideration is the size of the data set. For instance, you would not want to use 20 classes 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.

47a. 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	0.0333
67-72	3	69.5	0.1000
73-78	6	75.5	0.2000
79-84	10	81.5	0.3333
85-90	5	87.5	0.1667
91-96	2	93.5	0.0667
97-102	2	99.5	0.0667
103-108	1	105.5	0.0333
	$\sum f = 30$		$\sum \frac{f}{N} = 1$

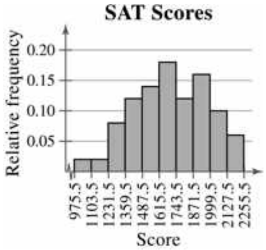


b. 16.7%, because the sum of the relative frequencies for the last three classes is 0.167.

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

48a. Class width = $\dfrac{\text{Range}}{\text{Number of classes}} = \dfrac{2250-976}{10} = 127.4 \Rightarrow 128$

Class	Frequency, f	Relative frequency
976-1103	1	0.02
1104-1231	1	0.02
1232-1359	4	0.08
1360-1487	6	0.12
1488-1615	7	0.14
1616-1743	9	0.18
1744-1871	6	0.12
1872-1999	8	0.16
2000-2127	5	0.10
2128-2255	3	0.06
	$\sum f = 50$	$\sum \frac{f}{N} = 1$



b. 62%, because the sum of the relative frequencies of the class starting with 1616 and all classes with higher scores is 0.62.

c. A score of 1360 or above, because the sum of the relative frequencies of the class starting with 1360 and all classes with higher scores is 0.88.

2.2 MORE GRAPHS AND DISPLAYS

2.2 Try It Yourself Solutions

1a.

3	
4	
5	
6	
7	
8	

b. Key $3|6 = 36$

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

c. Key $3|5 = 35$

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

d. More than 50% of the 50 richest people are older than 60. (Answers will vary.)

2a, b.

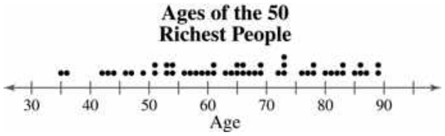
Key $3|5 = 35$

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

c. Most of the 50 richest people are older than 60. (Answers will vary.)

3a. Use the age for the horizontal axis.

b.

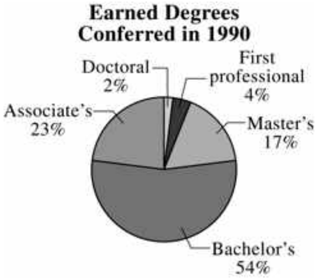


c. A large percentage of the ages are over 60. (Answers will vary.)

4a.

Type of Degree	f	Relative Frequency	Angle
Associate's	455	0.23	82.8°
Bachelor's	1052	0.54	194.4°
Master's	325	0.17	61.2°
First Professional	71	0.04	14.4°
Doctoral	38	0.02	7.2°
	$\sum f = 50$	$\sum \frac{f}{N} = 1$	$\sum \quad = 360°$

b.

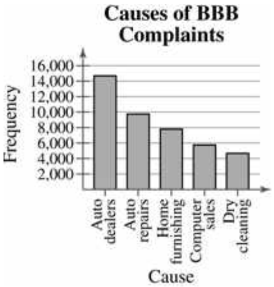


c. From 1990 to 2007, as percentages of total degrees conferred, associate's degrees increased by 1%, bachelor's degrees decreased by 3%, master's degrees increased by 3%, first professional degrees decreased by 1%, and doctoral degrees remained unchanged.

5a.

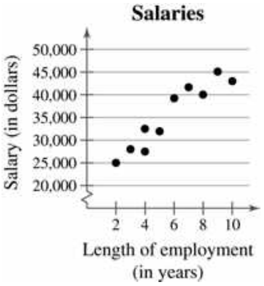
Cause	Frequency, *f*
Auto Dealers	14,668
Auto Repair	9,728
Home Furnishing	7,792
Computer Sales	5,733
Dry Cleaning	4,649

b.



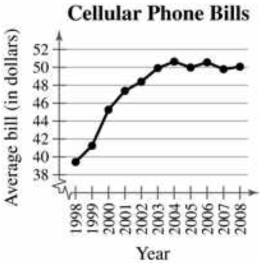
c. It appears that the auto industry (dealers and repair shops) account for the largest portion of complaints filed at the BBB. (Answers will very.)

6a, b.



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

7a, b.



c. The average bill increased from 1998 to 2004, then it hovered around $50.00 from 2004 to 2008.

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: Users spend the most amount of time on MySpace and the least amount of time on Twitter. Answers will vary.

14. Sample answer: Motor vehicle thefts decreased between 2003 and 2008. Answers will vary.

15. Answers will vary. Sample answer: Tailgaters irk drivers the most, while too cautious drivers irk drivers the least.

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

17. Key: $6|7 = 67$

$$\begin{array}{r|l}
6 & 7\ \ 8 \\
7 & 3\ \ 5\ \ 5\ \ 6\ \ 9 \\
8 & 0\ \ 0\ \ 2\ \ 3\ \ 5\ \ 5\ \ 7\ \ 7\ \ 8 \\
9 & 0\ \ 1\ \ 1\ \ 1\ \ 2\ \ 4\ \ 5\ \ 5
\end{array}$$

It appears that most grades for the biology midterm were in the 80s or 90s. (Answers will vary.)

18. Key: $4|9 = 49$

$$\begin{array}{r|l}
4 & 8\quad 9 \\
5 & 0\quad 0\quad 0\quad 1\quad 2 \\
5 & 5\quad 5\quad 5\quad 6\quad 7\quad 7\quad 8\quad 9\quad 9 \\
6 & 0\quad 0\quad 0\quad 1\quad 2\quad 2\quad 2\quad 3\quad 3\quad 4\quad 4\quad 4 \\
6 & 7 \\
7 & 4
\end{array}$$

It appears that most of the highest paid CEOs are between 55 and 65 years old. (Answers will vary.)

19. Key: $4|3 = 4.3$

$$\begin{array}{r|l}
4 & 3\quad 9 \\
5 & 1\quad 8\quad 8\quad 8\quad 9 \\
6 & 4\quad 8\quad 9\quad 9\quad 9 \\
7 & 0\quad 0\quad 2\quad 2\quad 2\quad 5 \\
8 & 0\quad 1
\end{array}$$

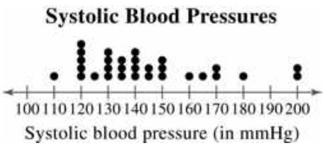
It appears that most ice had a thickness of 5.8 centimeters to 7.2 centimeters. (Answers will vary.)

20. Key: $17|5 = 17.5$

$$\begin{array}{r|l}
16 & 4\quad 8 \\
17 & 1\quad 1\quad 3\quad 4\quad 5\quad 5\quad 6\quad 7\quad 9 \\
18 & 1\quad 3\quad 4\quad 4\quad 6\quad 6\quad 6\quad 9 \\
19 & 0\quad 0\quad 2\quad 3\quad 3\quad 5\quad 6 \\
20 & 1\quad 8
\end{array}$$

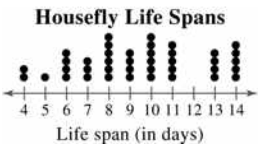
It appears that most farmers charge 17 to 19 cents per pound of apples. (Answers will vary.)

21.



It appears that systolic blood pressure tends to be between 120 and 150 millimeters of mercury. (Answers will vary.)

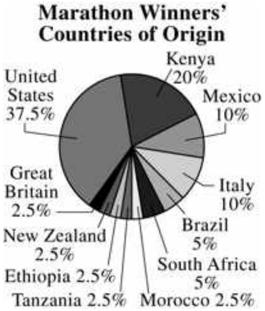
22.



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

23.

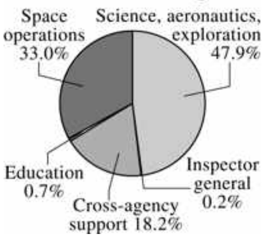
Category	Frequency, f	Relative Frequency	Angle
United States	15	0.375	135°
Italy	4	0.100	36°
Ethiopia	1	0.025	9°
South Africa	2	0.050	18°
Tanzania	1	0.025	9°
Kenya	8	0.200	72°
Mexico	4	0.100	36°
Morocco	1	0.025	9°
Great Britain	1	0.025	9°
Brazil	2	0.050	18°
New Zealand	1	0.025	9°
	$\sum f = 40$	$\sum \frac{f}{N} = 1$	$\sum \quad = 360°$



Most of the New York City Marathon winners are from the United States and Kenya. (Answers will vary.)

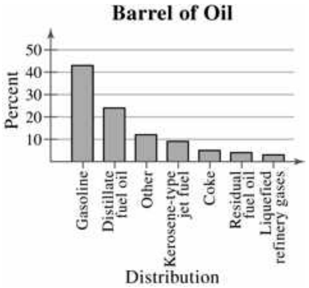
24.

Category	Frequency, f	Relative Frequency	Angle
Science, aeronautics, exploration	8947	0.479	172.4°
Space operations	6176	0.331	119.2°
Education	126	0.007	2.5°
Cross-agency support	3401	0.182	65.5°
Inspector general	36	0.002	0.7°
	$\sum f = 18{,}686$	$\sum \frac{f}{N} \approx 1$	$\sum \quad \approx 360°$



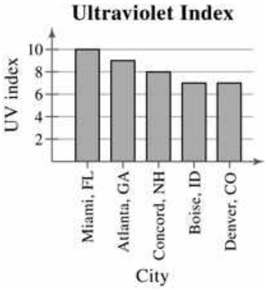
It appears that most of NASA's budget was spent on science, aeronautics, and exploration. (Answers will vary.)

25.



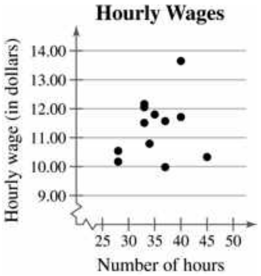
It appears that the largest portion of a 42-gallon barrel of crude oil is used for making gasoline. (Answers will vary.)

26.



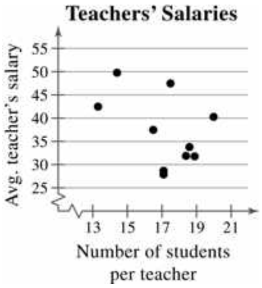
It appears that Boise, ID and Denver, CO have the same UV index. (Answers will vary.)

27.



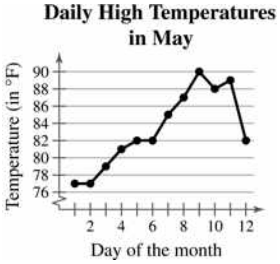
It appears that there is no relation between wages and hours worked. (Answers will vary.)

28.



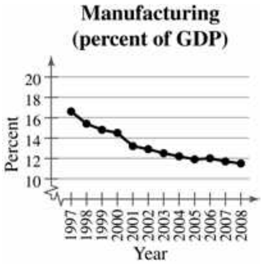
It appears that there is no relation between a teacher's average salary and the number of students per teacher. (Answers will vary.)

29.



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

30.



It appears that the largest decrease in manufacturing as a percent of GDP was from 2000 to 2001. (Answers will vary.)

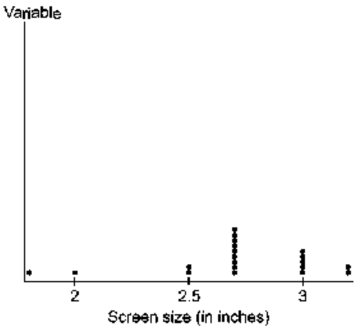
31. Variable: Scores

Key: $5|5 = 5.5$

5 | 5
6 | 2
6 | 8
7 | 0 1
7 | 5 6
8 | 0 2 3
8 | 5 6 7 8 8 9
9 | 0 3 3
9 | 5 5 8 9
10 | 0

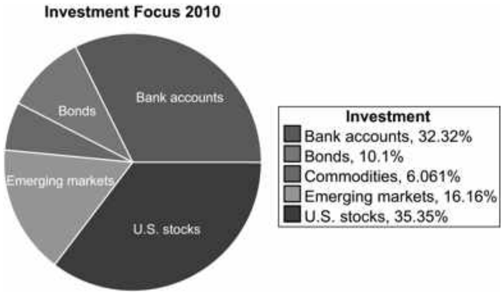
It appears that most scores on the final exam in economics were in the 80's and 90's. (Answers will vary.)

32.



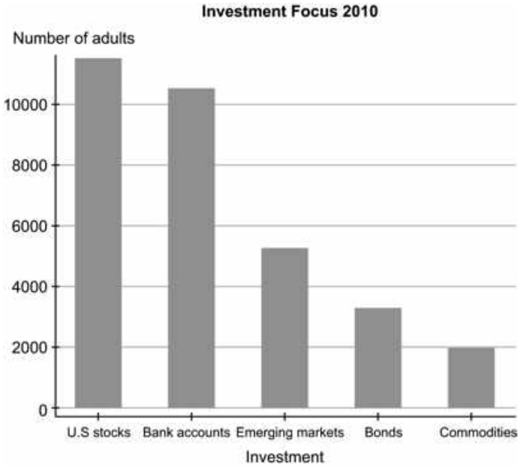
It appears that most screen sizes are between 2.5 and 3 inches. (Answers will vary.)

33a.



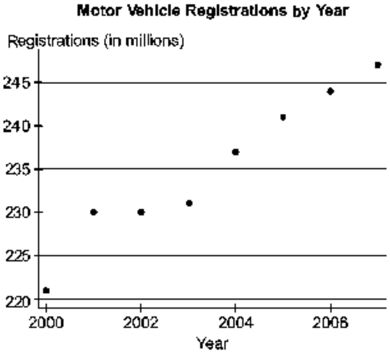
It appears that a large portion of adults said that the type of the investment that they would focus on in 2010 was U.S. stocks or bank accounts. (Answers will vary.)

b.



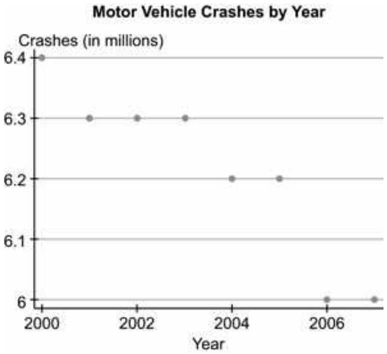
It appears that most adults said that the type of investment that they would focus on in 2010 was U.S. stocks or bank accounts. (Answers will vary.)

34a.



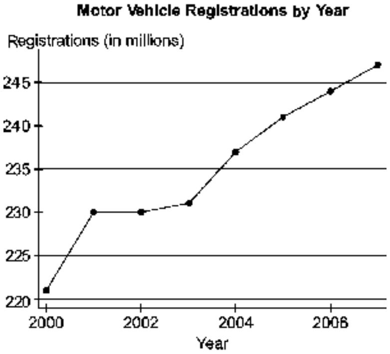
It appears that the number of registrations is increasing over time. (Answers will vary.)

b.



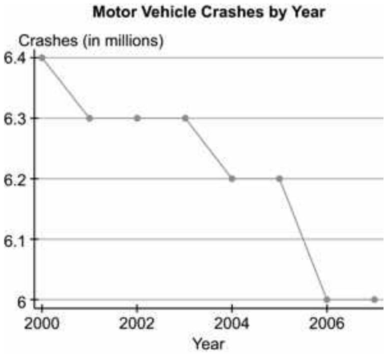
It appears that the number of crashes is decreasing over time. (Answers will vary.)

c.



It appears that the number of registrations is increasing over time. (Answers will vary.)

d.



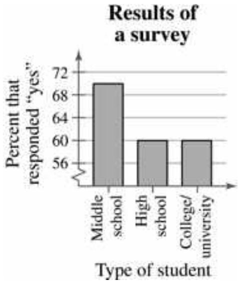
It appears that the number of crashes is decreasing over time. (Answers will vary.)

35a. 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. (Answers will vary.)

b.

36a. 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%. (Answers will vary.)

b.

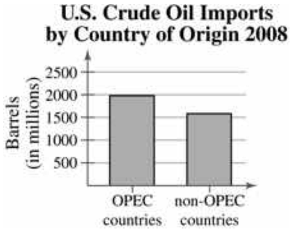


37a. The graph is misleading because the angle makes it appear as though the 3rd quarter had a larger percent of sales than the others, when the 1st and 3rd quarters have the same percent.

b.

38a. The graph is misleading because the "OPEC countries" bar is wider than the "non-OPEC countries" bar.

b.



39a. 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.

b. There are 30 lawyers at Law Firm A and 32 lawyers at Law Firm B.

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

40a. Key: $5|3|1 = 35$-year-old in 3:00 P.M. class and 31-year old in 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	

b. In the 3:00 P.M. class, the youngest participant is 35 years old and the oldest participant is 85 years old. In the 8:00 P.M. class, the youngest participant is 18 years old and the oldest participant is 71 years old.

c. In the 3:00 P.M. class, there are 26 participants and in the 8:00 P.M. class there are 30 participants.

d. 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. (Answers will vary.)

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, 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. 25, 60, 80, 97, 100, 130, 140, 200, 220, 250

b. median $= \frac{100 + 130}{2} = 115$

c. The median price of the sample of digital photo frames is $115.

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. “Yes” occurs with the greatest frequency (510).

b. The mode of the responses to the survey is "Yes." In this sample, there were more people who thought public cell phone conversations were rude than people who did not or had no opinion.

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
		$\sum w = 1$	$\sum(x \cdot w) = 91.8$

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***	**$x \cdot f$**
35-41	38	2	76
42-48	45	5	225
49-55	52	7	364
56-62	59	7	413
63-69	66	10	660
70-76	73	5	365
77-83	80	8	640
84-90	87	6	552
		$N = 50$	$\sum(x \cdot f) = 3265$

d. $\mu = \frac{\sum(x \cdot f)}{N} = \frac{3265}{50} = 65.3$

The mean age of the 50 richest people is 65.3

2.3 EXERCISE SOLUTIONS

1. True

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

3. True

4. True

5. 1, 2, 2, 2, 3 (Answers will vary.)

6. 2, 4, 5, 5, 6, 8 (Answers will vary.)

7. 2, 5, 7, 9, 35 (Answers will vary.)

8. 1, 2, 3, 3, 3, 4, 5 (Answers will vary.)

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

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

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

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

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

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

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

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

17. $\bar{x} = \frac{\sum x}{n} = \frac{64}{13} \approx 4.9$

$2\ 2\ 3\ 4\ 4\ 4\ \underset{\text{median} = 5}{\textcircled{5}}\ 5\ 6\ 6\ 7\ 8\ 8$

mode = 4 (occurs 3 times)

18. $\bar{x} = \frac{\sum x}{n} = \frac{396}{10} = 39.6$

$35\ 37\ 38\ 38\ \underbrace{39\ 39}\ 39\ 40\ 40\ 51$

$\text{median} - \frac{39+39}{2} - 39$

mode = 39 (occurs 3 times)

19. $\bar{x} = \frac{\sum x}{n} = \frac{76.8}{7} = 11.0$

$9.7\ \ 10.3\ \ 10.7\ \ \underset{\text{median} = 11.0}{\textcircled{11.0}}\ \ 11.7\ \ 11.7\ \ 11.7$

mode = 11.7 (occurs 3 times)

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

$$154\ \ 171\ \ 173\ \ 181\ \ \underbrace{184\ \ 188}_{\text{median} = \frac{184+188}{2} = 186}\ \ 203\ \ 235\ \ 240\ \ 275$$

mode = none

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

21. $\bar{x} = \frac{\sum x}{n} = \frac{686.8}{32} = 21.46$

$$10.9\ \ 12.3\ \ 15.2\ \ 15.3\ \ 16.1\ \ 16.4\ \ 16.6\ \ 17.5\ \ 18.1\ \ 18.4\ \ 19.1\ \ 19.7\ \ 20.4\ \ 20.4\ \ 20.6$$

$$\underbrace{21.8\ \ 22.1}_{\text{median} = \frac{21.8+22.1}{2} = 21.95}\ \ 22.5\ \ 22.6\ \ 22.7\ \ 23.0\ \ 23.4\ \ 24.2\ \ 24.4\ \ 25.1\ \ 26.0\ \ 26.7\ \ 26.8\ \ 28.4\ \ 28.8\ \ 29.4\ \ 31.9$$

mode = 20.4 (occurs 2 times)

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

$$12\ \ 18\ \ 26\ \ 28\ \ 31\ \ 33\ \ 40\ \ 44\ \ 45\ \ \underbrace{49\ \ 61}_{\text{median} - \frac{49+61}{2} = 55}\ \ 63\ \ 75\ \ 80\ \ 80\ \ 89\ \ 96\ \ 103\ \ 125\ \ 125$$

mode = 80, 125

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

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

median = not possible (nominal data)

mode = "Eyeglasses"

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

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

median = not possible (nominal data)

mode = "Money needed"

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

25. $\bar{x} = \frac{\sum x}{n} = \frac{1194.4}{7} \approx 170.63$

$$155.7\ \ 158.1\ \ 162.2\ \ \underset{\uparrow\ \text{median} = 169.3}{\boxed{169.3}}\ \ 180\ \ 181.8\ \ 187.3$$

mode = none

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

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

median = not possible (nominal data)

mode = "Mashed"

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

27. $\bar{x} = \dfrac{\sum x}{n} = \dfrac{1687}{10} = 168.7$

$$125\quad 125\quad 132\quad 140\quad \underbrace{155\quad 170}_{\text{median} = \frac{155+170}{2} - 162.5}\quad 175\quad 210\quad 225\quad 230$$

mode = 125 (occurs 2 times)

The mode does not represent the center of the data because 125 is the smallest number in the data set.

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

$$1\quad 10\quad \underset{\uparrow\ \text{median} = 15}{\textcircled{15}}\quad 25.5\quad 31.5$$

mode = none

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

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

$$1.5\quad 2.5\quad 2.5\quad 5\quad 10.5\quad 11\quad \underbrace{13\quad 15.5}_{\text{median} - \frac{13+15.5}{2} = 14.25}\quad 16.5\quad 17.5\quad 20\quad 26.5\quad 27\quad 28.5$$

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.

30. $\bar{x} = \dfrac{\sum x}{n} = \dfrac{3605}{15} \approx 240.3$

$$9\quad 28\quad 28\quad 32\quad 38\quad 52\quad 110\quad \underset{\uparrow\ \text{median} = 136}{\textcircled{136}}\quad 142\quad 350\quad 354\quad 409\quad 537\quad 625\quad 755$$

mode = 28 (occurs 2 times)

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

31. $\bar{x} = \dfrac{\sum x}{n} = \dfrac{835}{28} = 29.82$

$$6\ 7\ 12\ 15\ 18\ 19\ 20\ 24\ 24\ 24\ 25\ 28\ 29\ \underbrace{32\ 32}_{\text{median} = \frac{32+32}{2} = 32}\ 33\ 35\ 35\ 35\ 36\ 38\ 39\ 40\ 41\ 42\ 47\ 48\ 51$$

mode = 24, 35 (occurs 3 times each)

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

$$0.8\ 1.5\ 1.6\ 1.8\ 2.1\ \underbrace{2.3\ 2.4}_{\text{median} = \frac{2.3+2.4}{2} - 2.35}\ 2.5\ 3.0\ 3.9\ 4.0\ 4.0$$

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. $\bar{x}=\frac{\sum x}{n}=\frac{292}{15}\approx 19.47$

$$5\ 8\ 10\ 15\ 15\ 15\ 17\ \underset{\uparrow\ \text{median} = 20}{\boxed{20}}\ 21\ 22\ 22\ 25\ 28\ 32\ 37$$

mode = 15 (occurs 3 times)

34. $\bar{x}=\frac{\sum x}{n}=\frac{3110}{15}\approx 207.3$

$$170\ 180\ 190\ 200\ 200\ 210\ 210\ \underset{\uparrow\ \text{median} = 210}{\boxed{210}}\ 210\ 210\ 220\ 220\ 220\ 220\ 240$$

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 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***	$x \cdot w$
MBAs	92,500	8	740,000
BAs	68,000	17	1,156,000
		$\sum w = 25$	$\sum(x \cdot w) = 1{,}896{,}000$

$$\bar{x} = \frac{\sum(x \cdot w)}{\sum w} = \frac{1{,}896{,}000}{25} = 75{,}840$$

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
	$\sum w = 31$	$\sum(x \cdot w) = 30{,}448$

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

45.

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

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

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

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

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	55	247.5
10-19	14.5	70	1015
20-29	24.5	35	857.5
30-39	34.5	56	1932
40-49	44.5	74	3293
50-59	54.5	42	2289
60-69	64.5	38	2451
70-79	74.5	17	1266.5
80-89	84.5	10	845
		$n=397$	$\sum(x \cdot f)=14{,}196.5$

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

52.

Class	Midpoint, x	Frequency, f	$x \cdot f$
1-5	3	12	36
6-10	8	26	208
11-15	13	20	260
16-20	18	7	126
21-25	23	11	253
26-30	28	7	196
31-35	33	4	132
36-40	38	4	152
41-45	43	1	43
		$n = 92$	$\sum(x \cdot f) = 1406$

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

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

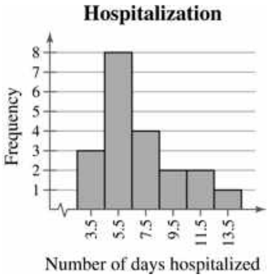
Class	Midpoint	Frequency, f
127-161	144	9
162-196	179	8
197-231	214	3
232-266	249	3
267-301	284	1
		$\sum f = 24$

Shape: Positively skewed

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

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
		$\sum f = 20$

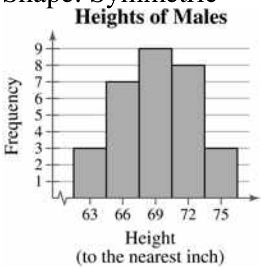
Shape: Positively skewed



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
		$\sum f = 30$

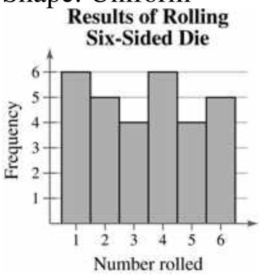
Shape: Symmetric



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



57a. $\bar{x} = \frac{\sum x}{n} = \frac{36.03}{6} = 6.005$

$5.59 \quad 5.99 \quad \underbrace{6 \quad 6.02} \quad 6.03 \quad 6.40$

median $= \frac{6 + 6.02}{2} - 6.01$

b. $\bar{x} = \frac{\sum x}{n} = \frac{35.67}{6} = 5.945$

$5.59 \quad 5.99 \quad \underbrace{6 \quad 6.02} \quad 6.03 \quad 6.04$

median $= \frac{6 + 6.02}{2} - 6.01$

c. The mean was affected more.

58a. $\bar{x} = \frac{\sum x}{n} = \frac{966.6}{19} \approx 50.87$

$9.1 \quad 12.5 \quad 12.9 \quad 15.5 \quad 22.0 \quad 22.2 \quad 24.9 \quad 27.9 \quad 28.8 \quad \enclose{circle}{28.9} \quad 32.3 \quad 34.7$

$39.7 \quad 53.6 \quad 54.5 \quad 65.1 \quad 69.7 \quad 151.2 \quad 261.1$

median = 28.9

b. $\bar{x} = \frac{\sum x}{n} = \frac{705.5}{18} \approx 39.19$

$9.1 \quad 12.5 \quad 12.9 \quad 15.5 \quad 22.0 \quad 22.2 \quad 24.9 \quad 27.9 \quad \underbrace{28.8 \quad 28.9} \quad 32.3 \quad 34.7$

$39.7 \quad 53.6 \quad 54.5 \quad 65.1 \quad 69.7 \quad 151.2$

median $= \frac{28.8 + 28.9}{2} = 28.85$

The mean was affected more.

c. $\bar{x} = \frac{\sum x}{n} = \frac{984.3}{20} \approx 49.22$

$9.1 \quad 12.5 \quad 12.9 \quad 15.5 \quad 17.7 \quad 22.0 \quad 22.2 \quad 24.9 \quad 27.9 \quad \underbrace{28.8 \quad 28.9} \quad 32.3 \quad 34.7$

$39.7 \quad 53.6 \quad 54.5 \quad 65.1 \quad 69.7 \quad 151.2 \quad 261.1$

median $= \frac{28.8 + 28.9}{2} = 28.85$

The mean was affected more.

59. Summary Statistics:

Column	n	Mean	Median	Min	Max
Amount (in dollars)	11	112.11364	105.25	79	151.5

60. Summary Statistics:

Column	n	Mean	Median	Min	Max
Price (in dollars)	30	47.26433	43.585	15.9	132.39

61a. $\bar{x} = \frac{\sum x}{n} = \frac{3222}{9} = 358$

$147 \quad 177 \quad 336 \quad 360 \quad \enclose{circle}{375} \quad 393 \quad 408 \quad 504 \quad 522$

median = 375

b. $\bar{x} = \dfrac{\sum x}{n} = \dfrac{9666}{9} = 1074$

$441\ \ 531\ \ 1008\ \ 1080\ \ \boxed{1125}\ \ 1179\ \ 1224\ \ 1512\ \ 1566$

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 from part (b) by 3, you will get the mean and median of the data set in inches.

62. Car A

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

$28\ \ 28\ \ \boxed{30}\ \ 32\ \ 34$

median = 30

mode = 28 (occurs 2 times)

Car B

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

$29\ \ 29\ \ \boxed{31}\ \ 31\ \ 31$

median = 31

mode = 31 (occurs 3 times)

Car C

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

$28\ \ 29\ \ \boxed{30}\ \ 32\ \ 32$

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.

63. Car A: Midrange = $\dfrac{34+28}{2} = 31$

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

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

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

64a. $\bar{x} = \dfrac{\sum x}{n} = \dfrac{1477}{30} \approx 49.2$

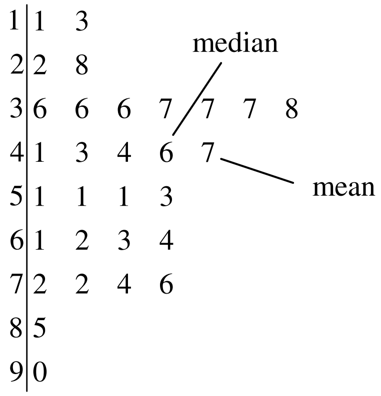
$11\ \ 13\ \ 22\ \ 28\ \ 36\ \ 36\ \ 36\ \ 37\ \ 37\ \ 37\ \ 38\ \ 41\ \ 43\ \ \underbrace{44\ \ 46}\ \ 47\ \ 51\ \ 51\ \ 51$

$53\ \ 61\ \ 62\ \ 63\ \ 64\ \ 72\ \ 72\ \ 74\ \ 76\ \ 85\ \ 90$

median = $\dfrac{46+47}{2} = 46.5$

b. Key: $3|6 = 36$

c. Positively skewed



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

c. The distribution is approximately symmetric.

65a. Order the data values.
11 13 22 28 36 36 36 37 37 37 38 41 43 44 46
47 51 51 51 53 61 62 63 64 72 72 74 76 85 90
Delete the lowest 10%, smallest 3 observations (11, 13, 22).
Delete the highest 10%, largest 3 observations (76, 85, 90).
Find the 10% trimmed mean using the remaining 24 observations.
10% trimmed mean ≈ 49.2

b. $\bar{x} \approx 49.2$
median = 46.5
mode = 36, 37, 51
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.

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

b.

Salary, x (1000s of dollars)	**Deviation, $x - \mu$ (100s of dollars)**
23	–18.5
29	–12.5
32	–9.5
40	–1.5
41	–0.5
41	–0.5
49	7.5
50	8.5
52	10.5
58	16.5
$\sum x = 415$	$\sum(x-\mu)=0$

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

Salary, x	**$x - \mu$**	**$(x - \mu)^2$**
23	–18.5	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
$\sum x = 415$	$\sum(x-\mu)=0$	$\sum(x-\mu)^2 = 1102.5$

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

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

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

4a. From 3ab, $SS_x = \Sigma\left(x - \bar{x}\right)^2 = 1102.5.$

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

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

d. The sample variance is 122.5 and the sample standard deviation is 11.1, or $11,100.

5a. Enter the data in a computer or a calculator.

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

6a. 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
13	3	9
13	3	9
13	3	9
13	3	9
13	3	9
$\sum x = 100$	$\sum(x-\mu)=0$	$\sum(x-\mu)^2 = 90$

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

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

7a. 66.92 – 64.3 = 2.62 = 1 standard deviation

b. 34%

c. Approximately 34% of women ages 20-29 are between 64.3 and 66.92 inches tall.

8a. 31.6 – 2(19.5) = –7.4

Because –7.4 does not make sense for an age, use 0.

b. 31.6 + 2(19.5) = 70.6

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 70.6 years old.

9a.

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

10a.

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

c.

$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
–146	21,316	8,100,080
–46	2116	486,680
54	2916	612,360
154	23,716	1,185,800
254	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. $s = \sqrt{\frac{\sum(x-\bar{x})^2 f}{n-1}} = \sqrt{\frac{28{,}715{,}797.5}{999}} \approx 169.5$

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)

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

5. {9, 9, 9, 9, 9, 9, 9}
$n = 7$
$\bar{x}=\frac{\sum \bar{x}}{n}=\frac{63}{7}=9$

x	$x-\bar{x}$	$(x-\bar{x})^2$
9	0	0
9	0	0
9	0	0
9	0	0
9	0	0
9	0	0
9	0	0
$\sum x=63$	$\sum(x-\bar{x})=0$	$\sum(x-\bar{x})^2=0$

$$s=\sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}=\sqrt{\frac{0}{6}}=0$$

6. {3, 3, 3, 7, 7, 7}
$n = 6$
$\mu=\frac{\sum x}{n}=\frac{30}{6}=5$

x	$x-\mu$	$(x-\mu)^2$
3	–2	4
3	–2	4
3	–2	4
7	2	4
7	2	4
7	2	4
$\sum x=30$	$\sum(x-\mu)=0$	$\sum(x-\mu)^2=24$

$$\sigma=\sqrt{\frac{\Sigma(x-\mu)^2}{N}}=\sqrt{\frac{24}{6}}=\sqrt{4}=2$$

7. 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.

8. When given a data set one 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.

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

10. You must know that the distribution is bell-shaped.

11. Range = Max – Min = 12 – 5 = 7

$$\mu = \frac{\sum x}{N} = \frac{90}{10} = 9$$

x	$x-\mu$	$(x-\mu)^2$
9	0	0
5	–4	16
9	0	0
10	1	1
11	2	4
12	3	9
7	–2	4
7	–2	4
8	–1	1
12	3	9
$\sum x = 90$	$\sum(x-\mu) = 0$	$\sum(x-\mu)^2 = 48$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} = \frac{48}{10} = 4.8$$

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}} = \sqrt{4.8} \approx 2.2$$

12. Range = Max – Min = 25 – 15 = 10

$$\mu = \frac{\sum x}{N} = \frac{266}{14} = 19$$

x	$x-\mu$	$(x-\mu)^2$
18	–1	1
20	1	1
19	0	0
21	2	4
19	0	0
17	–2	4
15	–4	16
17	–2	4
25	6	36
22	3	9
19	0	0
20	1	1
16	–3	9
18	–1	1
$\sum x = 90$	$\sum(x-\mu) = 0$	$\sum(x-\mu)^2 = 86$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} = \frac{86}{14} \approx 6.1$$

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}} = \sqrt{\frac{86}{14}} \approx 2.5$$

13. Range = Max – Min = 19 – 4 = 15

$$\bar{x} = \frac{\sum x}{n} = \frac{108}{9} = 12$$

x	$x-\bar{x}$	$(x-\bar{x})^2$
4	–8	64
15	3	9
9	–3	9
12	0	0
16	4	16
8	–4	16
11	–1	1
19	7	49
14	2	4
$\sum x = 108$	$\sum(x-\bar{x}) = 0$	$\sum(x-\bar{x})^2 = 168$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{168}{9-1} = 21$$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{21} \approx 4.6$$

14. Range = Max – Min = 28 – 7 = 21

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

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

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

$$s = \sqrt{\frac{\Sigma\left(x - \bar{x}\right)^2}{n-1}} = \sqrt{\frac{530.77}{12}} \approx 6.7$$

15. Range = Max – Min = 96 – 23 = 73

16. Range = Max – Min = 34 – 24 = 10

17. Range = Max – Min = 98 – 74 = 24

18. Range = Max – Min = 6.7 – 0.5 = 6.2

19a. Range = Max – Min = 38.5 – 20.7 = 17.8
b. Range = Max – Min = 60.5 – 20.7 = 39.8

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

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

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

23. Company B. An offer of $33,000 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.

24. Player B. A smaller standard deviation means that Player B’s scores tend to fall within a smaller interval of values than Player A’s scores.

25a. Dallas:

$$\bar{x}=\frac{\sum x}{n}=\frac{398.5}{9}\approx 44.28$$

38.7 39.9 40.5 41.6 44.7 45.8 47.8 49.5 50.0

median = 44.7

Range = Max – Min = 50.0 – 38.7 = 11.3

x	$x-\bar{x}$	$\left(x-\bar{x}\right)^2$
38.7	–5.58	31.1364
39.9	–4.38	19.1844
40.5	–3.78	14.2884
41.6	–2.68	7.1824
44.7	0.42	0.1764
45.8	1.52	2.3104
47.8	3.52	12.3904
49.5	5.22	27.2484
50.0	5.72	32.7184
		$\sum\left(x-\bar{x}\right)^2=146.6356$

$$s^2=\frac{\sum\left(x-\bar{x}\right)^2}{n-1}=\frac{146.6356}{8}\approx 18.33$$

$$s=\sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}}=\sqrt{\frac{146.6356}{8}}\approx 4.28$$

New York City:

$$\bar{x}=\frac{\sum x}{n}=\frac{458.2}{9}\approx 50.91$$

41.5 42.3 45.6 47.2 50.6 55.1 57.6 59.0 59.3

median = 50.6

Range = Max – Min = 59.3 – 41.5 = 17.8

x	$x-\bar{x}$	$\left(x-\bar{x}\right)^2$
41.5	–9.41	88.5481
42.3	–8.61	74.1321
45.6	–5.31	28.1961
47.2	–3.71	13.7641
50.6	–0.31	0.0961
55.1	4.19	17.5561
57.6	6.69	44.7561
59.0	8.09	65.4481
59.3	8.39	70.3921
		$\sum\left(x-\bar{x}\right)^2=402.8889$

$$s^2 = \frac{\Sigma\left(x-\overline{x}\right)^2}{n-1} = \frac{402.8889}{9-1} \approx 50.36$$

$$s = \sqrt{\frac{\Sigma\left(x-\overline{x}\right)^2}{n-1}} = \sqrt{\frac{402.8889}{8}} \approx 7.10$$

b. It appears from the data that the annual salaries in New York City are more variable than the annual salaries in Dallas. The annual salaries in Dallas have a lower mean and a lower median than the annual salaries in New York City.

26a. Boston:

$$\overline{x} = \frac{\Sigma x}{n} = \frac{667.4}{9} \approx 74.16$$

58.5 64.5 69.9 70.4 $\boxed{71.6}$ 79.9 80.1 84.2 88.3

median = 71.6

Range = Max – Min = 88.3 – 58.5 = 29.8

x	$x-\overline{x}$	$\left(x-\overline{x}\right)^2$
58.5	–15.66	245.2356
64.5	–9.66	93.3156
69.9	–4.26	18.1476
70.4	–3.76	14.1376
71.6	–2.56	6.5536
79.9	5.74	32.9476
80.1	5.94	35.2836
84.2	10.04	100.8016
88.3	14.14	199.9396
		$\sum\left(x-\overline{x}\right)^2 = 746.3624$

$$s^2 = \frac{\Sigma\left(x-\overline{x}\right)^2}{n-1} = \frac{746.3624}{9-1} \approx 93.30$$

$$s = \sqrt{\frac{\Sigma\left(x-\overline{x}\right)^2}{n-1}} = \sqrt{\frac{746.3624}{8}} \approx 9.66$$

Chicago:

$$\bar{x} = \frac{\sum x}{n} = \frac{599.5}{9} \approx 66.61$$

$$59.9\quad 60.9\quad 62.9\quad 65.4\quad \underset{\uparrow\ \text{median} = 68.5}{\boxed{68.5}}\quad 69.4\quad 70.1\quad 70.9\quad 71.5$$

Range = Max – Min = 71.5 – 59.9 = 11.6

x	$x-\bar{x}$	$\left(x-\bar{x}\right)^2$
59.9	–6.71	45.0241
60.9	–5.71	32.6041
62.9	–3.71	13.7641
65.4	–1.21	1.4641
68.5	1.89	3.5721
69.4	2.79	7.7841
70.1	3.49	12.1801
70.9	4.29	18.4041
71.5	4.89	23.9121
		$\sum\left(x-\bar{x}\right)^2 = 158.7089$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{158.7089}{9-1} \approx 19.84$$

$$s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}} = \sqrt{\frac{158.7089}{8}} \approx 4.45$$

b. It appears from the data that the annual salaries in Boston are more variable than the annual salaries in Chicago. The annual salaries in Boston have higher mean and median than the annual salaries in Chicago.

27a. Male:

$$\bar{x} = \frac{\sum x}{n} = \frac{13{,}144}{8} = 1643$$

$$1033\quad 1380\quad 1520\quad \underbrace{1645\quad 1714}_{\text{median} = \frac{1645 + 1714}{2} = 1679.5}\quad 1750\quad 1982\quad 2120$$

Range = Max – Min = 2120 – 1033 = 1087

x	$x-\bar{x}$	$\left(x-\bar{x}\right)^2$
1033	–610	372,100
1380	–263	69,169
1520	–123	15,129
1645	2	4
1714	71	5041
1750	107	11,449
1982	339	114,921
2120	477	227,529
		$\sum\left(x-\bar{x}\right)^2 = 815{,}342$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{815{,}342}{8-1} \approx 116{,}477.4$$

$$s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}} = \sqrt{\frac{815{,}342}{7}} \approx 341.3$$

Female:

$$\bar{x} = \frac{\sum x}{n} = \frac{13{,}673}{8} \approx 1709.1$$

$$1263 \quad 1497 \quad 1507 \quad \underbrace{1588 \quad 1785}_{\text{median} = \frac{1588+1785}{2} = 1686.5} \quad 1871 \quad 1952 \quad 2210$$

Range = Max – Min = 2210 – 1263 = 947

x	$x-\bar{x}$	$\left(x-\bar{x}\right)^2$
1263	–446.1	199,005.21
1497	–212.1	44,986.41
1507	–202.1	40,844.41
1588	–121.1	14,665.21
1785	75.9	5760.81
1871	161.9	26,211.61
1952	242.9	59,000.41
2210	500.9	250,900.81
		$\sum\left(x-\bar{x}\right)^2 = 641{,}374.88$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{641{,}374.88}{8-1} \approx 91{,}625.0$$

$$s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}} = \sqrt{\frac{641{,}374.88}{7}} = 302.7$$

b. It appears from the data that the SAT scores for males are more variable than the SAT scores for females. The SAT scores for males have a lower mean and median than the SAT scores for females.

28a. Team A:

$$\bar{x} = \frac{\sum x}{n} = \frac{2.694}{9} \approx 0.2993$$

0.235 0.256 0.272 0.295 $\boxed{0.297}$ 0.310 0.320 0.325 0.384

median = 0.297

Range = Max – Min = 0.384 – 0.235 = 0.149

x	$x-\bar{x}$	$\left(x-\bar{x}\right)^2$
0.235	–0.0643	0.00413449
0.256	–0.0433	0.00187489
0.272	–0.0273	0.00074529
0.295	–0.0043	0.00001849
0.297	–0.0023	0.00000529
0.310	0.0107	0.00011449
0.320	0.0207	0.00042849
0.325	0.0257	0.00066049
0.384	0.0847	0.00717409
		$\sum\left(x-\bar{x}\right)^2 = 0.01515601$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{0.01515601}{9-1} \approx 0.0019$$

$$s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}} = \sqrt{\frac{0.01515601}{8}} = 0.0435$$

Team B:

$$\bar{x} = \frac{\sum x}{n} = \frac{2.69}{9} \approx 0.2989$$

0.268 0.270 0.285 0.290 $\boxed{0.292}$ 0.305 0.315 0.330 0.335

median = 0.292

Range = Max – Min = 0.335 – 0.268 = 0.067

x	$x-\bar{x}$	$\left(x-\bar{x}\right)^2$
0.268	–0.0309	0.00095481
0.270	–0.0289	0.00083521
0.285	–0.0139	0.00019321
0.290	–0.0089	0.00007921
0.292	–0.0069	0.00004761
0.305	0.0061	0.00003721
0.315	0.0161	0.00025921
0.330	0.0311	0.00096721
0.335	0.0361	0.00130321
		$\sum\left(x-\bar{x}\right)^2 = 0.00467689$

$$s^2 = \frac{\sum\left(x-\bar{x}\right)^2}{n-1} = \frac{0.00467689}{9-1} \approx 0.0006$$

$$s = \sqrt{\frac{\Sigma\left(x-\bar{x}\right)^2}{n-1}} = \sqrt{\frac{0.00467689}{8}} \approx 0.02418$$

b. It appears form the data that the batting averages for Team A are more variable than the batting averages for Team B. The batting averages for Team A have a higher mean and a higher median than those for Team B.

29a. 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.

30a. 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.

31a. Greatest sample standard deviation: (ii)
Data set (ii) has more entries that are farther away from the mean.
Least sample standard deviation: (iii)
Data set (iii) has more entries that are close to the mean.

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

32a. 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.

33. $(1300,\ 1700) \rightarrow (1500-1(200),\ 1500+1(200)) \rightarrow \left(\bar{x}-s,\ \bar{x}+s\right)$

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

34. 95% of the data falls between $\bar{x}-2s$ and $\bar{x}+2s$.

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

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

95% of the farms have values between \$1500 and \$3300 per acre.

35a. $n = 75$
68%(75) = (0.68)(75) ≈ 51 farms have values between \$1300 and \$1700 per acre.

b. $n = 25$
68%(25) = (0.68)(25) ≈ 17 farms have values between \$1300 and \$1700 per acre.

36a. $n = 40$
95%(40) = (0.95)(40) ≈ 38 farms have values between \$1500 and \$3300 per acre.
b. $n = 20$
95%(20) = (0.95)(20) ≈ 19 farms have values between \$1500 and \$3300 per acre.

37. $\bar{x} = 1500 \qquad s = 200$
{\$950, \$1000, \$2000, \$2180} are outliers. They are more than 2 standard deviations from the mean (1100, 1900). \$2180 is very unusual because it is more than 3 standard deviations from the mean.

38. $\bar{x} = 2400 \qquad s = 450$
{\$1045, \$1490, \$3325, \$3800} are outliers. They are more than 2 standard deviations from the mean (1500, 3300). \$1045 and \$3800 are very unusual because they are more than 3 standard deviations from the mean.

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

$1 - \frac{1}{k^2} = 1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = 0.75 \Rightarrow$ At least 75% of the eruption times lie between 1.14 and 5.5 minutes.
If n = 32, at least (0.75)(32) = 24 eruptions will lie between 1.14 and 5.5 minutes.

40. $1 - \frac{1}{k^2} = 1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = 0.75$

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

41.

x	f	xf	$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f$
0	5	0	–2.1	4.41	22.05
1	11	11	–1.1	1.21	13.31
2	7	14	–0.1	0.01	0.07
3	10	30	0.9	0.81	8.10
4	7	28	1.9	3.61	25.27
	$n = 40$	$\sum xf = 83$			$\sum (x - \bar{x})^2 f = 68.8$

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

$$s = \sqrt{\frac{\sum (x - \bar{x})^2 f}{n - 1}} = \sqrt{\frac{68.8}{39}} \approx 1.3$$

42.

x	f	xf	$x-\bar{x}$	$\left(x-\bar{x}\right)^2$	$\left(x-\bar{x}\right)^2 f$
0	3	0	–1.7	2.89	8.67
1	15	15	–0.7	0.49	7.35
2	24	48	0.3	0.09	2.16
3	8	24	1.3	1.69	13.52
	$n = 50$	$\sum xf = 87$			$\sum\left(x-\bar{x}\right)^2 f = 31.7$

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

$$s = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2 f}{n-1}} = \sqrt{\frac{31.7}{49}} \approx 0.8$$

43. Class width = $\dfrac{\text{Max} - \text{Min}}{5} = \dfrac{14-1}{5} = \dfrac{13}{5} = 2.6 \Rightarrow 3$

Class	Midpoint, x	f	xf
1-3	2	3	6
4-6	5	6	30
7-9	8	13	104
10-12	11	7	77
13-15	14	3	42
		$N = 32$	$\sum xf = 259$

$$\mu = \frac{\sum xf}{N} = \frac{259}{32} \approx 8.1$$

$x-\mu$	$\left(x-\mu\right)^2$	$\left(x-\mu\right)^2 f$
–6.1	37.21	111.63
–3.1	9.61	57.66
–0.1	0.01	0.13
2.9	8.41	58.87
5.9	34.81	104.43
		$\sum\left(x-\mu\right)^2 f = 332.72$

$$\sigma = \sqrt{\frac{\sum\left(x-\mu\right)^2 f}{N}} = \sqrt{\frac{332.72}{32}} \approx 3.2$$

44. Class width = $\dfrac{\text{Max} - \text{Min}}{5} = \dfrac{244-145}{5} = \dfrac{99}{5} = 19.8 \Rightarrow 20$

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

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

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

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

45.

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

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

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

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

46.

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

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

$x-\bar{x}$	$\left(x-\bar{x}\right)^2$	$\left(x-\bar{x}\right)^2 f$
–1.9	3.61	3.61
–0.9	0.81	7.29
0.1	0.01	0.13
1.1	1.21	6.05
2.1	4.41	8.82
		$\sum\left(x-\bar{x}\right)^2 f = 25.9$

$$s=\sqrt{\frac{\Sigma\left(x-\bar{x}\right)^2 f}{n-1}}=\sqrt{\frac{25.9}{29}}\approx 0.9$$

47.

Class	Midpoint, x	f	xf
0-4	2.0	22.1	44.20
5-14	9.5	43.4	412.30
15-19	17.0	21.2	360.40
20-24	22.0	22.3	490.60
25-34	29.5	44.5	1312.75
35-44	39.5	41.3	1631.35
45-64	54.5	83.9	4572.55
65+	70.0	46.8	3276.00
		$n = 325.5$	$\sum xf = 12{,}100.15$

$$\bar{x}=\frac{\sum xf}{n}=\frac{12{,}100.15}{325.5}\approx 37.17$$

$x-\bar{x}$	$\left(x-\bar{x}\right)^2$	$\left(x-\bar{x}\right)^2 f$
–35.17	1236.9289	27,336.12869
–27.67	765.6289	33,228.29426
–20.17	406.8289	8624.77268
–15.17	230.1289	5131.87447
–7.67	58.8289	2617.88605
2.33	5.4289	224.21357
17.33	300.3289	25,197.59471
32.83	1077.8089	50,441.45642
		$\sum\left(x-\bar{x}\right)^2 f = 152{,}802.22085$

$$s=\sqrt{\frac{\Sigma\left(x-\bar{x}\right)^2 f}{n-1}}=\sqrt{\frac{152{,}802.22085}{324.5}}\approx 21.70$$

48.

Midpoint, x	f	xf
4.5	35.4	159.30
14.5	35.3	511.85
24.5	33.5	820.75
34.5	33.6	1159.20
44.5	28.8	1281.60
54.5	21.9	1193.55
64.5	13.9	896.55
74.5	7.2	536.40
84.5	2.6	219.70
94.5	0.3	28.35
	$n = 212.5$	$\sum xf = 6807.25$

$$\bar{x} = \frac{\sum xf}{n} = \frac{6807.25}{212.5} \approx 32.03$$

$x - \bar{x}$	$\left(x - \bar{x}\right)^2$	$\left(x - \bar{x}\right)^2 f$
–27.53	757.9009	26,829.69186
–17.53	307.3009	10,847.72117
–7.53	56.7009	1899.48015
2.47	6.1009	204.99024
12.47	155.5009	4478.42592
22.47	504.9009	11,057.32971
32.47	1054.3009	14,654.78251
42.47	1803.7009	12,986.64648
52.47	2753.1009	7158.06234
62.47	3902.5009	1170.75027
		$\sum\left(x - \bar{x}\right)^2 f = 91{,}287.88065$

$$s = \sqrt{\frac{\sum\left(x - \bar{x}\right)^2 f}{n - 1}} = \sqrt{\frac{91{,}287.88065}{211.5}} \approx 20.78$$

49. Summary Statistics:

Column	n	**Mean**	**Variance**
Amount (in dollars)	15	58.8	239.74286

Std. Dev.	**Median**	**Range**	**Min**	**Max**
15.483632	60	59	30	89

50. Summary Statistics:

Column	n	**Mean**	**Variance**
Price (in dollars)	12	216.65666	14,442.424

Std. Dev.	**Median**	**Range**	**Min**	**Max**
120.176636	189.99	410	89.99	499.99

51. Heights:

$$\mu = \frac{\sum x}{N} = \frac{873}{12} = 72.75$$

x	$x-\mu$	$(x-\mu)^2$
68	–4.75	22.5625
69	–3.75	14.0625
69	–3.75	14.0625
70	–2.75	7.5625
72	–0.75	0.5625
72	–0.75	0.5625
73	0.25	0.0625
74	1.25	1.5625
74	1.25	1.5625
76	3.25	10.5625
77	4.25	18.0625
79	6.25	39.0625
		$\sum(x-\mu)^2 = 130.25$

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}} = \sqrt{\frac{130.25}{12}} \approx 3.29$$

$$\text{CV}_{\text{heights}} = \frac{\sigma}{\mu}\cdot 100\% = \frac{3.29}{72.75}\cdot 100 \approx 4.5\%$$

Weights:

$$\mu = \frac{\sum x}{N} = \frac{2254}{12} = 187.83$$

x	$x-\mu$	$(x-\mu)^2$
162	–25.83	667.1889
168	–19.83	393.2289
171	–16.83	283.2489
174	–13.83	191.2689
180	–7.83	61.3089
185	–2.83	8.0089
189	1.17	1.3689
192	4.17	17.3889
197	9.17	84.0889
201	13.17	173.4489
210	22.17	491.5089
225	37.17	1381.6089
		$\sum(x-\mu)^2 = 3753.6668$

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}} = \sqrt{\frac{3753.6668}{12}} \approx 17.69$$

$$\text{CV}_{\text{weights}} = \frac{\sigma}{\mu}\cdot 100\% = \frac{17.69}{187.83}\cdot 100 \approx 9.4\%$$

It appears that weight is more variable than height.

52a. Male:

x	x^2
1520	2,310,400
1750	3,062,500
2120	4,494,400
1380	1,904,400
1982	3,928,324
1645	2,706,025
1033	1,067,089
1714	2,937,796
$\sum x = 13{,}144$	$\sum x^2 = 22{,}410{,}934$

$$s = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}} = \sqrt{\frac{22{,}410{,}934 - \frac{(13{,}144)^2}{8}}{8-1}} = \sqrt{\frac{815{,}342}{7}} \approx 341.3$$

Female:

x	x^2
1785	3,186,225
1507	2,271,049
1497	2,241,009
1952	3,810,304
2210	4,884,100
1871	3,500,641
1263	1,595,169
1588	2,521,744
$\sum x = 13{,}673$	$\sum x^2 = 24{,}010{,}241$

$$s = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}} = \sqrt{\frac{24{,}010{,}241 - \frac{(13{,}673)^2}{8}}{8-1}} = \sqrt{\frac{641{,}374.875}{7}} \approx 302.7$$

b. The answers are the same as from Exercise 27.

53a. $\bar{x} \approx 41.5$ $\quad s \approx 5.3$

b. $\bar{x} \approx 43.6$ $\quad s \approx 5.6$

c. $\bar{x} \approx 3.5$ $\quad s \approx 0.4$

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

54a. $\bar{x} \approx 41.7$ $\quad s \approx 6.0$

b. $\bar{x} \approx 42.7$ $\quad s \approx 6.0$

c. $\bar{x} \approx 39.7$ $\quad s \approx 6.0$

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

55a. Male: $\bar{x} = 1643$

x	$\lvert x - \bar{x} \rvert$
1520	123
1750	107
2120	477
1380	263
1982	339
1645	2
1033	610
1714	71
	$\sum \lvert x - \bar{x} \rvert = 1992$

$$\frac{\sum \lvert x - \bar{x} \rvert}{n} = \frac{1992}{8} = 249$$

$s = 341.3$

Female: $\bar{x} \approx 1709.1$

x	$\lvert x - \bar{x} \rvert$
1785	75.9
1507	202.1
1497	212.1
1952	242.9
2210	500.9
1871	161.9
1263	446.1
1588	121.1
	$\sum \lvert x - \bar{x} \rvert = 1963$

$$\frac{\sum \lvert x - \bar{x} \rvert}{n} = \frac{1963}{8} \approx 245.4$$

$s = 302.7$

The mean absolute deviation is less than the sample standard deviation.

b. Team A: $\bar{x} \approx 0.2993$

x	$\lvert x-\bar{x} \rvert$
0.295	0.0043
0.310	0.0107
0.325	0.0257
0.272	0.0273
0.256	0.0433
0.297	0.0023
0.320	0.0207
0.384	0.0847
0.235	0.0643
	$\sum \lvert x-\bar{x} \rvert = 0.2833$

$$\frac{\Sigma \lvert x-\bar{x} \rvert}{n} = \frac{0.2833}{9} \approx 0.0315$$

$s = 0.0435$

Team B: $\bar{x} \approx 0.2989$

x	$\lvert x-\bar{x} \rvert$
0.285	0.0139
0.305	0.0061
0.315	0.0161
0.270	0.0289
0.292	0.0069
0.330	0.0311
0.335	0.0361
0.268	0.0309
0.290	0.0089
	$\sum \lvert x-\bar{x} \rvert = 0.1789$

$$\frac{\Sigma \lvert x-\bar{x} \rvert}{n} = \frac{0.1789}{9} \approx 0.0199$$

$s = 0.0242$

The mean absolute deviation is less than the sample standard deviation.

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

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

57a. $P = \frac{3(\bar{x}-\text{median})}{s} = \frac{3(17-19)}{2.3} \approx -2.61;$ skewed left

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

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

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

2.5 MEASURES OF POSITION

2.5 Try It Yourself Solutions

1a. 35, 36, 42, 43, 44, 46, 47, 49, 51, 51, 53, 53, 54, 54, 56, 57, 58, 59, 60, 61, 61, 63, 64, 65, 65, 66, 66, 67, 68, 69, 69, 72, 73, 73, 73, 76, 77, 78, 78, 80, 81, 82, 83, 83, 85, 86, 86, 87, 89, 89

b. $Q_2=65.5$

c. $Q_1=54,\ Q_3=78$

d. About one fourth of the 50 richest people are 54 years old or younger; one half are 65.5 years old or younger; and about three fourths of the 50 richest people are 78 years old or younger.

2a. (Enter the data)

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

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

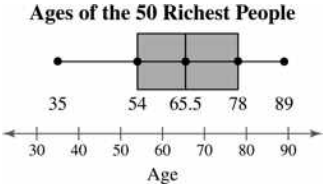
3a. $Q_1=54,\ Q_3=78$

b. IQR $=Q_3-Q_1=78-54=24$

c. The ages of the 50 richest people in the middle portion of the data set vary by at most 24 years.

4a. Min = 35, $Q_1=54,\ Q_2=65.5,\ Q_3=78,$ Max = 89

b,c.



d. It appears that half of the ages are between 54 and 78.

5a. 50th percentile

b. 50% of the 50 richest people are younger than 66.

6a. $x=60:\ z=\dfrac{x-\mu}{\sigma}=\dfrac{60-70}{8}=-1.25$

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

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

b. 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.

7a. Best actor: $\mu = 43.7,\ \sigma = 8.7$

Best actress: $\mu = 35.9,\ \sigma = 11.4$

b. Sean Penn: $x = 48:\ z = \dfrac{x-\mu}{\sigma} = \dfrac{48-43.7}{8.7} \approx 0.49$

Kate Winslet: $x = 33:\ z = \dfrac{x-\mu}{\sigma} = \dfrac{33-35.9}{11.4} \approx -0.25$

c. Sean Penn's age is 0.49 standard deviation above the mean of the best actors. Kate Winslet's age is -0.25 standard deviation below the mean of the best actresses. Neither actor's age is unusual.

2.5 EXERCISE SOLUTIONS

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

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

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

4. The child's IQ is higher than 93% of the other children 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. False. The median of a data set is a fractile, but the mean may or may not be fractile depending on the distribution of the data.

8. True

9. True

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

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

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

13. False. A z-score of -2.5 is considered unusual.

14. True

15a. Min = 10, $Q_1 = 13$, $Q_2 = 15$, $Q_3 = 17$, Max = 20

b. IQR = $Q_3 - Q_1 =$ 17 – 13 = 4

16a. Min = 100, $Q_1 = 130$, $Q_2 = 205$, $Q_3 = 270$, Max = 320

b. IQR = $Q_3 - Q_1 =$ 270 – 130 = 140

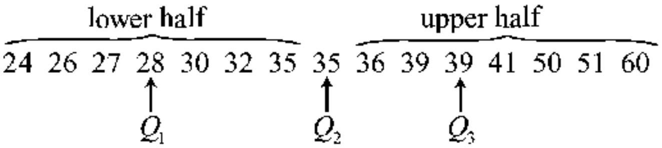
17a. Min = 900, $Q_1 = 1250$, $Q_2 = 1500$, $Q_3 = 1950$, Max = 2100
b. IQR = $Q_3 - Q_1 =$ 1950 – 1250 = 700

18a. Min = 25, $Q_1 = 50$, $Q_2 = 65$, $Q_3 = 70$, Max = 85
b. IQR = $Q_3 - Q_1 =$ 70 – 50 = 20

19a. Min = -1.9, $Q_1 = -0.5$, $Q_2 = 0.1$, $Q_3 = 0.7$, Max = 2.1
b. IQR = $Q_3 - Q_1 =$ $0.7-(-0.5)$ = 1.2

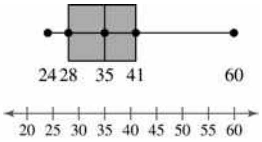
20a. Min = -1.3, $Q_1 = -0.3$, $Q_2 = 0.2$, $Q_3 = 0.4$, Max = 2.1
b. IQR = $Q_3 - Q_1 =$ $0.4-(-0.3)$ = 0.7

21a.

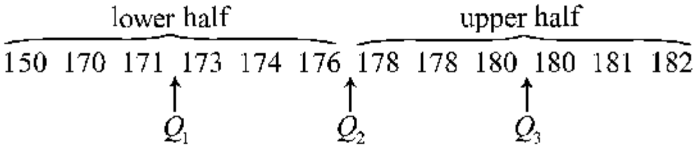


Min = 24, $Q_1 = 28$, $Q_2 = 35$, $Q_3 = 41$, Max = 60

b.

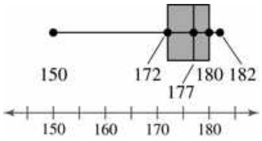


22a.

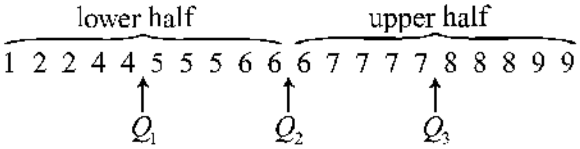


Min = 150, $Q_1 = 172$, $Q_2 = 177$, $Q_3 = 180$, Max = 182

b.

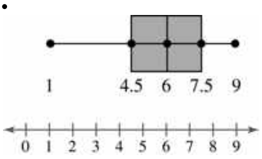


23a.



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

b.

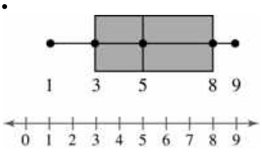


24a.

$$\overbrace{1\ 1\ 2\ 2\ 2\ 2\ 2\ 3\ \underset{\underset{Q_1}{\uparrow}}{3}\ 3\ 3\ 3\ 3\ 4\ 4\ 4\ 5}^{\text{lower half}}\ \underset{\underset{Q_2}{\uparrow}}{5}\ \overbrace{5\ 5\ 5\ 6\ 7\ 7\ 7\ 7\ \underset{\underset{Q_3}{\uparrow}}{8}\ 8\ 8\ 9\ 9\ 9\ 9\ 9\ 9}^{\text{upper half}}$$

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

b.



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

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

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

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

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

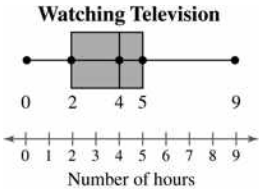
25% of the values are below B, 50% of the values are below A, and 75% of the values are below C.

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

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

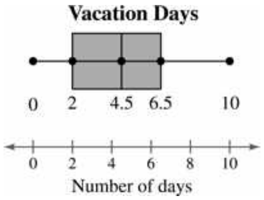
31a. $Q_1 = 2$, $Q_2 = 4$, $Q_3 = 5$

b.



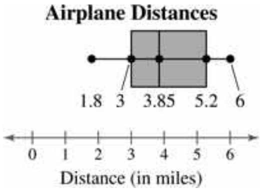
32a. $Q_1 = 2$, $Q_2 = 4.5$, $Q_3 = 6.5$

b.



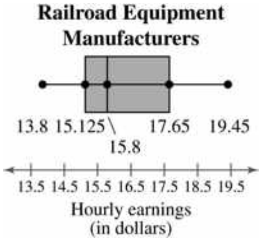
33a. $Q_1 = 3,\ Q_2 = 3.85,\ Q_3 = 5.2$

b.



34a. $Q_1 = 15.125,\ Q_2 = 15.8,\ Q_3 = 17.65$

b.



35a. 5 **b.** 50% **c.** 25%

36a. \$17.65 **b.** 50% **c.** 50%

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

38. $\text{A} \Rightarrow z = -1.54$
$\text{B} \Rightarrow z = 0.77$
$\text{C} \Rightarrow z = 1.54$
None of the z-scores are unusual.

39a. Statistics: $x = 75 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{75-63}{7} \approx 1.71$

Biology: $x = 25 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{25-23}{3.9} \approx 0.51$

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

40a. Statistics: $x = 60 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{60-63}{7} \approx -0.43$

Biology: $x = 22 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{22-23}{3.9} \approx -0.26$

b. The student had a better score on the biology test.

41a. Statistics: $x = 78 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{78-63}{7} \approx 2.14$

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

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

42a. Statistics: $x = 63 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{63-63}{7} = 0$

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

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

43a. $x = 34{,}000 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{34{,}000-35{,}000}{2{,}250} \approx -0.44$

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

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

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

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

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

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

44a. $x = 34 \Rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{34-33}{4} = 0.25$

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

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

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

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

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

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

45. 72 inches
60% of the heights are below 72 inches.

46. 98th percentile
98% of the heights are below 77 inches

47. $x=74 \Rightarrow z=\dfrac{x-\mu}{\sigma}=\dfrac{74-69.9}{3.0}\approx 1.37$

$x=62 \Rightarrow z=\dfrac{x-\mu}{\sigma}=\dfrac{62-69.9}{3.0}\approx -2.63$

$x=80 \Rightarrow z=\dfrac{x-\mu}{\sigma}=\dfrac{80-69.9}{3.0}\approx 3.37$

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

48. $x=70 \Rightarrow z=\dfrac{x-\mu}{\sigma}=\dfrac{70-69.9}{3.0}\approx 0.03$

$x=66 \Rightarrow z=\dfrac{x-\mu}{\sigma}=\dfrac{66-69.9}{3.0}=-1.3$

$x=68 \Rightarrow z=\dfrac{x-\mu}{\sigma}=\dfrac{68-69.9}{3.0}\approx -0.63$

49. $x=71.1 \Rightarrow z=\dfrac{x-\mu}{\sigma}=\dfrac{71.1-71.1}{3.0}=0.0$

Approximately the 50th percentile.

50. $x=66.3 \Rightarrow z=\dfrac{x-\mu}{\sigma}=\dfrac{66.3-69.9}{3.0}=-1.2$

Approximately the 10th percentile.

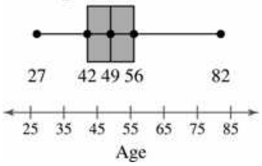
51a.

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

Q_1 Q_2 Q_3

Min = 27, $Q_1=42$, $Q_2=49$, $Q_3=56$, Max = 82

b.



c. Half of the executives are between 42 and 56 years old.
d. About 49 years old because half of the executives are older and half are younger.
e. The age groups 20-29, 70-79, and 80-89 would all be considered unusual because they are more than two standard deviations from the mean.

52.

1 2 3 3 5 5 7 7 8 10

↑ ↑ ↑

Q_1 Q_2 Q_3

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

53.

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

↑ ↑ ↑

Q_1 Q_2 Q_3

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

54.

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

↑ ↑ ↑

$Q_1 = 8.9$ Q_2 $Q_3 = 13.05$

$$\text{Midquartile} = \frac{Q_1 + Q_3}{2} = \frac{8.9+13.05}{2} = 10.975$$

55.

13.4 15.2 15.6 16.7 17.2 18.7 19.7 19.8 19.8 20.8 21.4 22.9 28.7 30.1 31.9

↑ ↑ ↑

Q_1 Q_2 Q_3

$$\text{Midquartile} = \frac{Q_1 + Q_3}{2} = \frac{16.7+22.9}{2} = 19.8$$

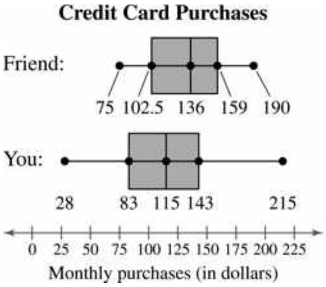
56a. Concert 1: Symmetric
Concert 2: Skewed right
Concert 1 has less variation.

b. Concert 2 is more likely to have outliers because its distribution is wider.

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.

57.



Your distribution is symmetric and your friend’s distribution is uniform.

58. $$\text{Percentile} = \frac{\text{Number of data values less than } x}{\text{Total number of data values}} \cdot 100$$
$$= \frac{3}{73} \cdot 100 \approx \text{4th percentile}$$

59. $$\text{Percentile} = \frac{\text{Number of data values less than } x}{\text{Total number of data values}} \cdot 100$$
$$= \frac{30}{73} \cdot 100 \approx \text{41st percentile}$$

60a.

$$\overbrace{2\ \ 7\ \ 8\ \underset{\underset{Q_1}{\uparrow}}{9}\ \ 9\ \ 10\ \ 10}^{\text{lower half}}\ \ \underset{\underset{Q_2}{\uparrow}}{11}\ \ \overbrace{11\ \ 12\ \ 12\ \underset{\underset{Q_3}{\uparrow}}{13}\ \ 15\ \ 16\ \ 24}^{\text{upper half}}$$

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

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

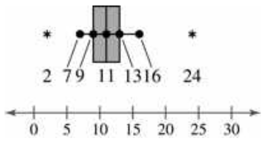
$1.5 \times \text{IQR} = 6$

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

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

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

b.



61a.

$$\overbrace{62\ \ 72\ \ 72\ \underset{\underset{Q_1}{\uparrow}}{\ }\ 74\ \ 75\ \ 75}^{\text{lower half}}\underset{\underset{Q_2}{\uparrow}}{\ }\overbrace{75\ \ 76\ \ 78\ \underset{\underset{Q_3}{\uparrow}}{\ }\ 80\ \ 80\ \ 95}^{\text{upper half}}$$

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

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

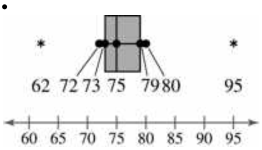
$1.5 \times \text{IQR} = 9$

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

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

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

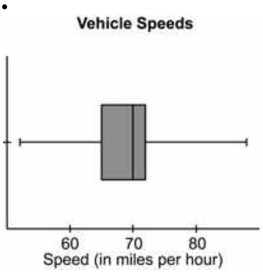
b.



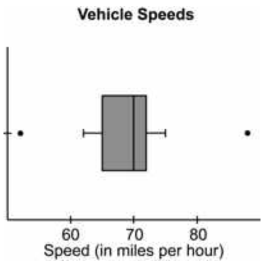
62a. Summary statistics:

Column	Min	Q_1	Median	Q_3	Max
Speed (in miles per hour)	52	65	70	72	88

b.



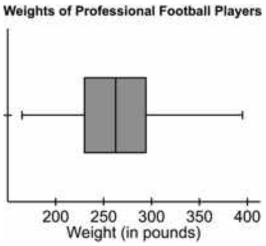
c.



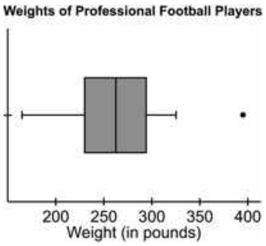
63a. Summary statistics:

Column	Min	Q_1	Median	Q_3	Max
Weight (in pounds)	165	230	262.5	294	395

b.



c.

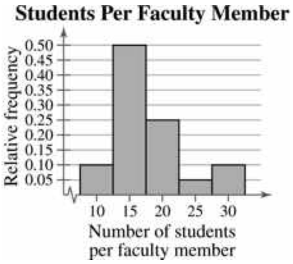


CHAPTER 2 REVIEW EXERCISE SOLUTIONS

1. Class width = $\dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{30-8}{5} = 4.4 \Rightarrow 5$

Class	Midpoint, x	Boundaries	Frequency, f	Relative frequency	Cumulative frequency
8-12	10	7.5-12.5	2	0.10	2
13-17	15	12.5-17.5	10	0.50	12
18-22	20	17.5-22.5	5	0.25	17
23-27	25	22.5-27.5	1	0.05	18
28-32	30	27.5-32.5	2	0.10	20
			$\sum f = 20$	$\sum \frac{f}{n} = 1$	

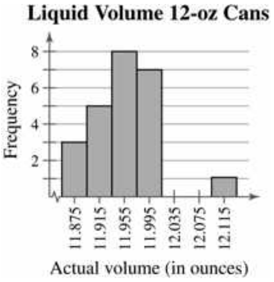
2.



Class with greatest relative frequency: 13-17
Class with least relative frequency: 23-27

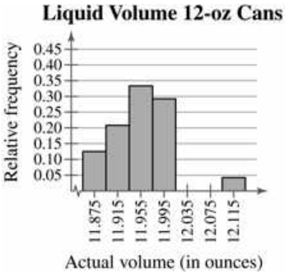
3. Class width = $\dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{12.10-11.86}{7} \approx 0.03 \Rightarrow 0.04$

Class	**Midpoint**	**Frequency, f**	**Relative frequency**
11.86-11.89	11.875	3	0.125
11.90-11.93	11.915	5	0.208
11.94-11.97	11.955	8	0.333
11.98-12.01	11.995	7	0.292
12.02-12.05	12.035	0	0
12.06-12.09	12.075	0	0
12.10-12.13	12.115	1	0.042
		$\sum f = 24$	$\sum \frac{f}{n} = 1$



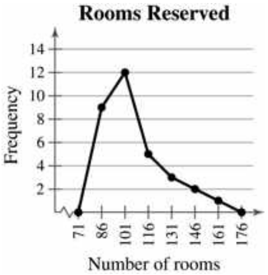
4. Class width = $\dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{12.10-11.86}{7} \approx 0.03 \Rightarrow 0.04$

Class	**Midpoint**	**Frequency, f**	**Relative frequency**
11.86-11.89	11.875	3	0.125
11.90-11.93	11.915	5	0.208
11.94-11.97	11.955	8	0.333
11.98-12.01	11.995	7	0.292
12.02-12.05	12.035	0	0
12.06-12.09	12.075	0	0
12.10-12.13	12.115	1	0.042
		$\sum f = 24$	$\sum \frac{f}{n} = 1$

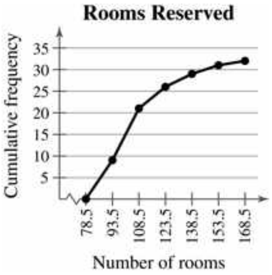


5.

Class	Midpoint	Frequency, f	Cumulative frequency
79-93	86	9	9
94-108	101	12	21
109-123	116	5	26
124-138	131	3	29
139-153	146	2	31
154-168	161	1	32
		$\sum f = 32$	



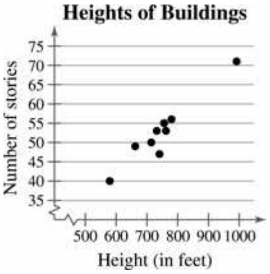
6.



7. 1 | 0 0 Key: 1 | 0 = 10
2 | 0 0 2 5 5
3 | 0 3 4 5 5 8
4 | 1 2 4 4 7 8
5 | 2 3 3 7 9
6 | 1 1 5
7 | 1 5
8 | 9

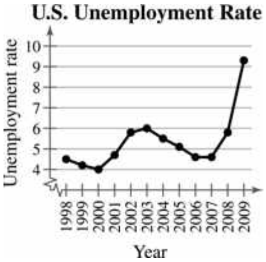
8.

9.

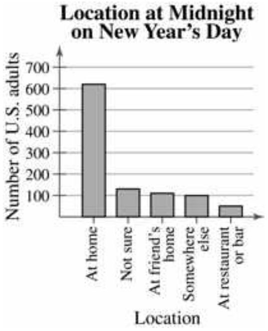


The number of stories appears to increase with height.

10.



11.



12.

Location	Frequency	Relative frequency	Degrees
At home	620	0.6139	221°
At friend's home	110	0.1089	39°
At restaurant or bar	50	0.0495	18°
Somewhere else	100	0.0990	36°
Not sure	130	0.1287	46°
		$\sum \frac{f}{n} = 1$	

13. $\bar{x} = \dfrac{\sum x}{n} = \dfrac{291.5}{10} = 29.15$

$$25.0\ \ 26.0\ \ 27.0\ \ 27.5\ \ \underbrace{29.5\ \ 29.5}_{\text{median} = 29.5}\ \ 30.5\ \ 31.5\ \ 32.0\ \ 33.0$$

Mode = 29.5 (occurs 2 times)

14. $\bar{x} =$ not possible
median = not possible
mode = "Approved"
The mean and median cannot be found because the data are at the nominal level of measurement.

15.

Midpoint, *x*	**Frequency, *f***	**$x \cdot f$**
10	2	20
15	10	15
20	5	100
25	1	25
30	2	60
	$n = 20$	$\sum(x \cdot f) = 355$

$\bar{x} = \dfrac{\sum(x \cdot f)}{n} = \dfrac{355}{20} \approx 17.8$

16.

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
	$n = 60$	$\sum(x \cdot f) = 125$

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

17.

Source	**Score, x**	**Weight, w**	$x \cdot w$
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
		$\sum w = 1$	$\sum (x \cdot w) = 82.1$

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

18.

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
		$\sum w = 1$	$\sum (x \cdot w) = 88.8$

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

19. Skewed

20. Skewed

21. Skewed left

22. Skewed right

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

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

25. Range = Max – Min = 8.26 – 5.46 = \$2.80

26. Range = Max – Min = 19.73 – 15.89 = \$3.84

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

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
$\sum x = 96$	$\sum(x-\mu) \approx 0$	$\sum(x-\mu)^2 = 295.74$

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

28. $\mu = \frac{\sum x}{N} = \frac{612}{9} = 68$

x	$x-\mu$	$(x-\mu)^2$
55	–13	169
89	21	441
73	5	25
73	5	25
61	–7	49
76	8	64
71	3	9
59	–9	81
55	–13	169
$\sum x = 612$	$\sum(x-\mu) = 0$	$\sum(x-\mu)^2 = 1032$

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}} = \sqrt{\frac{1032}{9}} \approx 10.7$$

29. $\bar{x}=\dfrac{\sum x}{n}=\dfrac{36{,}801}{15}=2453.4$

x	$x-\bar{x}$	$\left(x-\bar{x}\right)^2$
2445	–8.4	70.56
2940	486.6	236,779.56
2399	–54.4	2959.36
1960	–493.4	243,443.56
2421	–32.4	1049.76
2940	486.6	236,779.56
2657	203.6	41,452.96
2153	–300.4	90,240.16
2430	–23.4	547.56
2278	–175.4	30,765.16
1947	–506.4	256,440.96
2383	–70.4	4956.16
2710	256.6	65,843.56
2761	307.6	94,617.76
2377	–76.4	5836.96
$\sum x=36{,}801$	$\sum\left(x-\bar{x}\right)=0$	$\sum\left(x-\bar{x}\right)^2=1{,}311{,}783.6$

$$s=\sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}}=\sqrt{\frac{1{,}311{,}783.6}{14}}\approx 306.1$$

30. $\bar{x}=\dfrac{\sum x}{n}=\dfrac{416{,}659}{8}\approx 52{,}082.4$

x	$x-\bar{x}$	$\left(x-\bar{x}\right)^2$
49,632	–2450.4	6,004,460.16
54,619	2536.6	6,434,339.56
58,298	6215.6	38,633,683.36
48,250	–3832.4	14,687,289.76
51,842	–240.4	57,792.16
50,875	–1207.4	1,457,814.76
53,219	1136.6	1,291,859.56
49,924	–2158.4	4,658,690.56
$\sum x=416{,}659$	$\sum\left(x-\bar{x}\right)\approx 0$	$\sum\left(x-\bar{x}\right)^2=73{,}225{,}929.88$

$$s=\sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n-1}}=\sqrt{\frac{73{,}225{,}929.88}{7}}\approx 3234.3$$

31. 99.7% of the distribution lies within 3 standard deviations of the mean.

$\mu-3\sigma=49-(3)(2.50)=41.5$

$\mu+3\sigma=49+(3)(2.50)=56.5$

99.7% of the distribution lies between \$41.50 and \$56.50.

32. $(46.75,\ 52.25) \to \left(49.50 - 1(2.75),\ 49.50 + 1(2.75)\right) \to \left(\bar{x} - s,\ \bar{x} + s\right)$

68% of the cable rates lie between \$46.75 and \$52.25.

33. $\left(\bar{x} - 2s,\ \bar{x} + 2s\right) \to (20,\ 52)$ 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 \$20 and \$52.

34. $\left(\bar{x} - 2s,\ \bar{x} + 2s\right) \to (3,\ 11)$ 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 (20)(0.75) = 15 shuttle flights lasted between 3 days and 11 days.

35.

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
	$n = 40$	$\sum xf = 99$			$\sum (x - \bar{x})^2 f = 60$

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

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

36.

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
	$n = 25$	$\sum xf = 61$			$\sum (x - \bar{x})^2 f = 72.2$

$$\bar{x} = \frac{\sum xf}{n} = \frac{61}{25} \approx 2.4$$

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

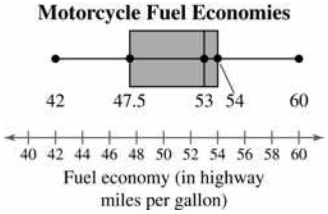
37.

$$\overbrace{42\ \ 42\ \ 42\ \ 47\ \ 47 \underset{\overset{\uparrow}{Q_1}}{\,} 48\ \ 48\ \ 50\ \ 53\ \ 53}^{\text{lower half}}\ \ \underset{\overset{\uparrow}{Q_2}}{53}\ \ \overbrace{53\ \ 54\ \ 54\ \ 54\ \ 54 \underset{\overset{\uparrow}{Q_3}}{\,} 54\ \ 57\ \ 57\ \ 60\ \ 60}^{\text{upper half}}$$

Min = 42, $Q_1 = 47.5$, $Q_2 = 53$, $Q_3 = 54$, Max = 60

38. IQR = $Q_3 - Q_1$ = 54 – 47.5 = 6.5

39.



40. 15

41.

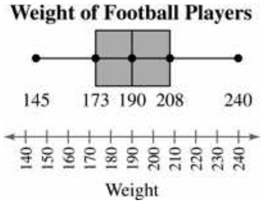
$$\overbrace{25.0\ \ 26.0\ \ \underset{\overset{\uparrow}{Q_1}}{27.0}\ \ 27.5\ \ 29.5}^{\text{lower half}} \underset{\overset{\uparrow}{Q_2}}{\,} \overbrace{29.5\ \ 30.5\ \ \underset{\overset{\uparrow}{Q_3}}{31.5}\ \ 32.0\ \ 33.0}^{\text{upper half}}$$

IQR = $Q_3 - Q_1$ = 31.5 – 27.0 = 4.5

42.

$$\overbrace{145\ \ 156\ \ 167\ \ 172\ \ \underset{\overset{\uparrow}{Q_1}}{173}\ \ 184\ \ 185\ \ 185\ \ 190}^{\text{lower half}}\ \ \underset{\overset{\uparrow}{Q_2}}{190}\ \ \overbrace{192\ \ 195\ \ 197\ \ 205\ \ \underset{\overset{\uparrow}{Q_3}}{208}\ \ 212\ \ 227\ \ 228\ \ 240}^{\text{upper half}}$$

Min = 145, $Q_1 = 173$, $Q_2 = 190$, $Q_3 = 208$, Max = 240



The distribution is symmetric.

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

44. 14th percentile

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

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

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

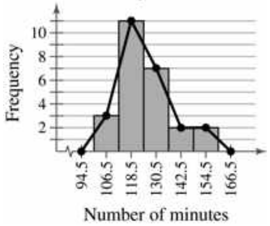
48. $z=\dfrac{11{,}300-11{,}830}{2370}=-0.22$

CHAPTER 2 QUIZ SOLUTIONS

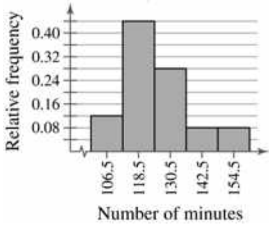
1a. Class width = $\dfrac{\text{Max} - \text{Min}}{\text{Number of classes}}=\dfrac{157-101}{5}=11.2\Rightarrow 12$

Class	Midpoint	Class boundaries	Frequency, f	Relative frequency	Cumulative frequency
101-112	106.5	100.5-112.5	3	0.12	3
113-124	118.5	112.5-124.5	11	0.44	14
125-136	130.5	124.5-136.5	7	0.28	21
137-148	142.5	136.5-148.5	2	0.08	23
149-160	154.5	148.5-160.5	2	0.08	25
			$\sum f=25$	$\sum\frac{f}{n}=1$	

b. Frequency histogram and polygon



c. Relative frequency histogram



d. Skewed

e. 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 | 1 1 2 5 9 9

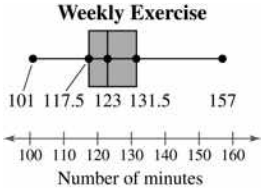
14 |

15 | 0 7

f.

$$\overbrace{101\ \ 108\ \ 111\ \ 114\ \ 116\ \ 117\ \underset{\underset{Q_1}{\uparrow}}{}\ 118\ \ 119\ \ 119\ \ 120\ \ 120\ \ 123}^{\text{lower half}}\ \ \underset{\underset{Q_2}{\uparrow}}{123}\ \ \overbrace{124\ \ 127\ \ 127\ \ 128\ \ 131\ \ 131\ \underset{\underset{Q_3}{\uparrow}}{}\ 132\ \ 135\ \ 139\ \ 139\ \ 150\ \ 157}^{\text{upper half}}$$

Min = 101, $Q_1=117.5$, $Q_2=123$, $Q_3=131.5$, Max = 157



g.

2.

Midpoint, x	Frequency, f	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
106.5	3	319.5	–18.7	349.69	1049.07
118.5	11	1303.5	–6.7	44.89	493.79
130.5	7	913.5	5.3	28.09	196.63
142.5	2	285.0	17.3	299.29	598.58
154.5	2	309.0	29.3	858.49	1716.98
	$n=25$	$\sum xf = 3130.5$			$\sum(x-\bar{x})^2 f = 4055.05$

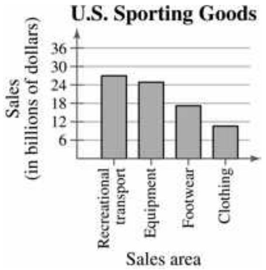
$$\bar{x} = \frac{\sum xf}{n} = \frac{3130.5}{25} \approx 125.2$$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2 f}{n-1}} = \sqrt{\frac{4055.05}{24}} \approx 13.0$$

3a.

Category	Frequency	Relative frequency	Degrees
Clothing	10.6	0.1330	48°
Footwear	17.2	0.2158	78°
Equipment	24.9	0.3124	112°
Rec. Transport	27.0	0.3388	122°
	$n=79.7$	$\sum \frac{f}{n} = 1$	

b.



4a. $\bar{x} = \frac{\sum x}{n} = \frac{6013}{8} \approx 751.6$

$$\underbrace{444\ \ 446\ \ 667\ \ 774\ \ 795}\ \ 908\ \ 960\ \ 1019$$

$$\text{median} = \frac{774 + 795}{2} - 784.5$$

mode = none

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

b. Range = Max – Min = 1019 – 444 = 575

x	$x - \bar{x}$	$(x - \bar{x})^2$
774	22.4	501.76
446	–305.6	93,391.36
1019	267.4	71,502.76
795	43.4	1883.56
908	156.4	24,460.96
667	–84.6	7157.16
444	–307.6	94,617.76
960	208.4	43,430.56
		$\sum(x - \bar{x})^2 = 336{,}945.88$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{336{,}945.88}{7} \approx 48{,}135.1$$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{336{,}945.88}{7}} \approx 219.4$$

5. $\bar{x} - 2s = 155{,}000 - 2 \cdot 15{,}000 = \$125{,}000$

$\bar{x} + 2s = 155{,}000 + 2 \cdot 15{,}000 = \$185{,}000$

95% of the new home prices fall between \$125,000 and \$185,000.

6a. $x = 200{,}000\text{:}\ z = \frac{x - \mu}{\sigma} = \frac{200{,}000 - 155{,}000}{15{,}000} = 3.0 \Rightarrow$ unusual price

b. $x = 55{,}000\text{:}\ z = \frac{x - \mu}{\sigma} = \frac{55{,}000 - 155{,}000}{15{,}000} = -6.67 \Rightarrow$ very unusual price

c. $x = 175{,}000\text{:}\ z = \frac{x - \mu}{\sigma} = \frac{175{,}000 - 155{,}000}{15{,}000} \approx 1.33 \Rightarrow$ not unusual

d. $x = 122{,}000: \ z = \frac{x-\mu}{\sigma} = \frac{122{,}000-155{,}000}{15{,}000} = -2.2 \Rightarrow$ unusual price

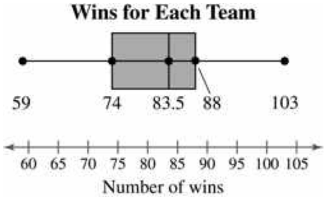
7a.

$$59\ 62\ 64\ 65\ 65\ 70\ 70\ \underset{\overset{\uparrow}{Q_1}}{74}\ 75\ 75\ 75\ 78\ 79\ 80\ 83\ \underset{Q_2}{\uparrow}\ 84\ 85\ 86\ 86\ 87\ 87\ 87\ \underset{\overset{\uparrow}{Q_3}}{88}\ 91\ 92\ 93\ 95\ 95\ 97\ 103$$

Min = 59, $Q_1 = 74$, $Q_2 = 83.5$, $Q_3 = 88$, Max = 103

b. IQR = $Q_3 - Q_1 = 88 - 74 = 14$

c.



CUMULATIVE REVIEW FOR CHAPTERS 1 AND 2

1. Systematic sampling. A bias may enter this study if the machine makes a consistent error.

2. Simple Random Sampling. A bias of this type of study is that the researchers did not include people without telephones.

3.

4. \$2,996,106 is a parameter because it is describing the average salary of all Major League Baseball players.

5. 19% is a statistic because it is describing a proportion within a sample of 100 adults.

6a. $\bar{x} = 83{,}500, \ s = \$1500$

$(80{,}500\ 86{,}500) = 83{,}500 \pm 2(1500) \Rightarrow$ 2 standard deviations away form the mean.

Approximately 95% of the electrical engineers will have salaries between (80,500 86,500).

b. 40(0.95) = 38

c. $x = \$90{,}500\text{: } z = \dfrac{x-\mu}{\sigma} = \dfrac{90{,}500 - 83{,}500}{1500} \approx 4.67$

$x = \$79{,}750\text{: } z = \dfrac{x-\mu}{\sigma} = \dfrac{79{,}750 - 83{,}500}{1500} = -2.5$

$x = \$82{,}600\text{: } z = \dfrac{x-\mu}{\sigma} = \dfrac{82{,}600 - 83{,}500}{1500} = -0.6$

The salaries of $90,500 and $79,750 are unusual.

7. Population: Collection of the career interests of all college and university students
Sample: Collection of the career interests of the 195 college and university students whose career counselors were surveyed

8. Population: Collection of the life spans of all people
Sample: Collection of the life spans of the 232,606 people in the study

9. Census, because there are only 100 members in the Senate.

10. Experiment, because we want to study the effects of removing recess from schools.

11. Quantitative: The data are at the ratio level.

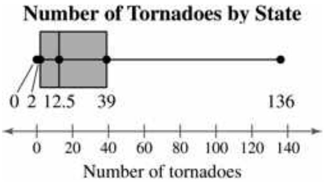
12. Qualitative: The data are at the nominal level.

13.

Q_1
↓
0 0 0 0 0 0 0 0 1 1 1 1 2 2 4 4 4 4 5 6 6 7 8 10
11 14 14 17 19 21 23 23 23 24 27 30 34 39 40 46 53 54 56 63 69 71 81 105 105 136
↑ ↑
Q_2 Q_3

Min = 0, $Q_1 = 2$, $Q_2 = 12.5$, $Q_3 = 39$, Max = 136

b.



c. 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
		$\sum w = 1$	$\sum(x \cdot w) = 88.9$

$$\bar{x} = \frac{\sum(x \cdot w)}{\sum w} = \frac{88.9}{1} = 88.9$$

15a. $\bar{x} = \dfrac{49.4}{9} \approx 5.49$

$3.4\ \ 3.9\ \ 4.2\ \ 4.6\ \textcircled{5.4}\ 6.5\ \ 6.8\ \ 7.1\ \ 7.5$

$\uparrow$ 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
		$\sum(x - \bar{x})^2 = 18.7289$

$$s^2 = \frac{\sum(x - \bar{x})^2}{n-1} = \frac{18.7289}{8} \approx 2.34$$

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n-1}} = \sqrt{\frac{18.7289}{8}} \approx 1.53$$

The maximum difference in alligator tail lengths is 4.1 feet and the standard deviation of tail lengths is about 1.53 feet.

16a. The number of deaths due to heart disease for women will continue to decrease.

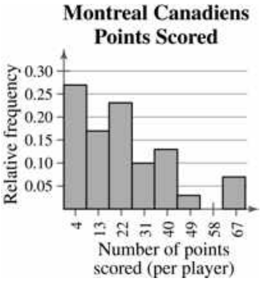
b. The study was only conducted over the past 5 years and deaths may not decrease in the next year.

17. Class width $= \dfrac{\text{Max} - \text{Min}}{\text{Number of classes}} = \dfrac{65-0}{8} = 8.125 \Rightarrow 9$

Class limits	**Midpoint**	**Class boundaries**	**Frequency**	**Relative frequency**	**Cumulative frequency**
0-8	4	–0.5-8.5	8	0.27	8
9-17	13	8.5-17.5	5	0.17	13
18-26	22	17.5-26.5	7	0.23	20
27-35	31	26.5-35.5	3	0.10	23
36-44	40	35.5-44.5	4	0.13	27
45-53	49	44.5-53.5	1	0.03	28
54-62	58	53.5-62.5	0	0.00	28
63-71	67	62.5-71.5	2	0.07	30
			$\sum f = 30$	$\sum \frac{f}{n} = 1$	

18. The distribution is skewed right.

19.



Class with greatest frequency: 0-8
Class with least frequency: 54-62