

Chapter 12

Chapter 12 Opener

Try It Yourself (p. 605)

1. *Answer should include, but is not limited to:* The student will conduct a survey to determine the after-school activities of students in the class. The results will be displayed in a circle graph.
2. *Answer should include, but is not limited to:* The student will conduct a survey to determine the number of pets owned by students in the class. The results will be displayed in a histogram.

Section 12.1

12.1 Activity (pp. 606–607)

1. a.

Stack	1	2	3	4	5	6	7	8	9
Coins	5	4	3	6	2	5	8	7	5

- b. Because there are 45 coins and 9 stacks, the mean is $45 \div 9 = 5$ coins.

To find the median and the mode, arrange the number of coins from least to greatest.

2, 3, 4, 5, 5, 5, 6, 7, 8

Because the median is the number in the middle, the median is 5. Because the mode is the number that occurs most often, the mode is 5.

- c. The mean will not change because the total number of coins and the number of stacks does not change.

Because the median and the mode are dependent on the number of coins in each stack, the median and the mode can be changed.

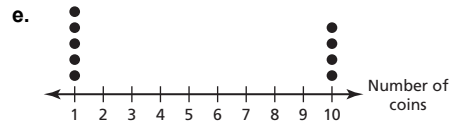
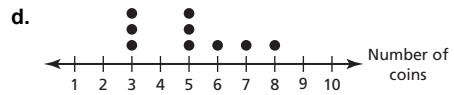
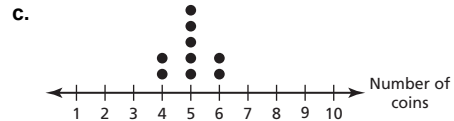
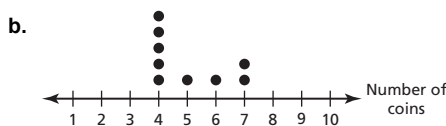
- d. Yes; yes; *Sample answer:* Let the number of coins in each stack be 1, 2, 3, 4, 6, 6, 7, 8, 8.

This is one way to arrange the coins so that the median is 6.

Let the number of coins in each stack be 1, 1, 1, 2, 8, 8, 8, 8, 8.

This is one way to arrange the coins so that the median is 8.

3. To distribute 45 coins into 9 stacks using a fair distribution, divide the number of coins by the number of stacks. So, each stack should have $45 \div 9 = 5$ coins. The number of coins in each stack is also equal to the mean, 5 coins.



The distributions in parts (b) and (c) seem the most fair because all of the values are near the mean of 5. The distribution in part (e) seems the least fair because a person has either 1 coin or 10 coins.

4. To distribute an amount evenly, each person should receive the same number of items, or the mean amount.
5. *Answer should include, but is not limited to:* Make sure the possible distributions produce the mean or median given in the article.

12.1 On Your Own (pp. 608–609)

1. **Mean:** The sum of the data is $58.8 + 6.99 = 65.79$ and the number of values is 9. So, the mean is $65.79 \div 9 = \$7.31$.

Median: 3.87, 6.99, 6.99, 7.25, **7.25**, 7.99, 8.25, 8.45, 8.75

The middle value is \$7.25, which is the median.

Mode: 3.87, **6.99**, **6.99**, **7.25**, **7.25**, 7.99, 8.25, 8.45, 8.75
Because 6.99 and 7.25 both occur the most often, the modes are \$6.99 and \$7.25.

By adding the value \$6.99, the mean decreases by \$0.04, the median decreases by \$0.37, and there are now 2 modes instead of 1.

Chapter 12

2. The value $4\frac{1}{5}$ miles is high compared to the other altitudes. It is the outlier. Find the mean, median, and mode with and without the outlier.

Mean with outlier: $\frac{11.1}{6} = 1.85$

Mean without outlier: $\frac{6.9}{5} = 1.38$

Median with outlier: $\frac{9}{10}, 1\frac{1}{5}, \underbrace{1\frac{2}{5}, 1\frac{1}{2}}_{\frac{2\frac{9}{10}}{2}}, 1\frac{9}{10}, 4\frac{1}{5}$
 $\frac{2\frac{9}{10}}{2} = \frac{9}{10} = 1.45$

Median without outlier: $\frac{9}{10}, 1\frac{1}{5}, 1\frac{2}{5}, 1\frac{1}{2}, 1\frac{9}{10}$

The median is $1\frac{2}{5}$.

Mode with outlier: $\frac{9}{10}, 1\frac{1}{5}, 1\frac{2}{5}, 1\frac{1}{2}, 1\frac{9}{10}, 4\frac{1}{5}$

There is no mode.

Mode without outlier: $\frac{9}{10}, 1\frac{1}{5}, 1\frac{2}{5}, 1\frac{1}{2}, 1\frac{9}{10}$

There is no mode.

By removing the outlier, the mean decreases $1.85 - 1.38 = 0.47$ mile, the median decreases $1.45 - 1.4 = 0.05$ mile, and the mode is unaffected because there is no mode.

3. Add $1\frac{1}{2}$ miles to each altitude. The new altitudes are $3\frac{2}{5}, 2\frac{9}{10}, 5\frac{7}{10}, 3, 2\frac{7}{10},$ and $2\frac{2}{5}$, respectively.

Find the mean, median, and mode with the increased altitudes.

Mean: $\frac{20.1}{6} = 3.35$

Median: $2\frac{2}{5}, 2\frac{7}{10}, \underbrace{2\frac{9}{10}, 3}_{\frac{5.9}{2}}, 3\frac{2}{5}, 5\frac{7}{10}$

The mean of the two middle values is

$$\frac{2\frac{9}{10} + 3}{2} = \frac{5.9}{2} = 2.95.$$

Mode: $2\frac{2}{5}, 2\frac{7}{10}, 2\frac{9}{10}, 3, 3\frac{2}{5}, 5\frac{7}{10}$

There is no mode.

By increasing each altitude by $1\frac{1}{2}$ miles, the mean and the median increase by $1\frac{1}{2}$ miles. There is still no mode.

12.1 Exercises (pp. 610–611)

Vocabulary and Concept Check

- A measure of central tendency is a measure that represents the center of a data set.
The mean of a data set is the sum of the data divided by the number of data values.
The median of a data set with an odd number of values is the middle value after the data has been ordered. The median of a data set with an even number of values is the mean of the two middle values after the data has been ordered.
The mode of a data set is the value or values that occur most often.
- Sample answer:* 1, 2, 2, 3, 6, 8, 8, 9, 12
The data set has 2 modes: 2 and 8.
- If the outlier is greater than the mean, the new mean will be lower when the outlier is removed. If the outlier is less than the mean, the new mean will be higher when the outlier is removed.

Practice and Problem Solving



Mean: $\frac{101}{9} = 11.\bar{2}$

Median: 7, 8, 10, 11, **11**, 12, 13, 14, 15
The middle value is 11.

Mode: 7, 8, 10, **11, 11**, 12, 13, 14, 15
The value 11 occurs the most often.

The mean is $11.\bar{2}$ strawberries, the median is 11 strawberries, and the mode is 11 strawberries.

5. **Mean:** $\frac{9}{9} = 1$

Median: -3, -2, -1, -1, **1**, 2, 3, 4, 6
The middle value is 1.

Mode: -3, -2, **-1, -1**, 1, 2, 3, 4, 6
The value -1 occurs most often.

The mean is 1, the median is 1, and the mode is -1.

Chapter 12

6. **Mean:** $\frac{1.8}{12} = 0.15$

Median: -3.01, -2.64, -2.41, -1.78, -0.38, **0.66**,
0.67, 1.05, 1.39, 2.03, 2.20, 4.02

The mean of the two middle values is

$$\frac{0.66 + 0.67}{2} = \frac{1.33}{2} = 0.665.$$

Mode: -3.01, -2.64, -2.41, -1.78, -0.38, 0.66,
0.67, 1.05, 1.39, 2.03, 2.20, 4.02

There is no value that occurs most often.

The mean is \$0.15, the median is \$0.665, and there is no mode.

7. **Mean:** $\frac{19\frac{2}{3}}{10} = 1\frac{29}{30}$

Median: $1\frac{1}{3}, 1\frac{2}{3}, 1\frac{2}{3}, 1\frac{2}{3}, 2, 2, 2, 2\frac{1}{3}, 2\frac{1}{3}, 2\frac{2}{3}$

The mean of the two middle values is

$$\frac{2 + 2}{2} = \frac{4}{2} = 2.$$

Mode: $1\frac{1}{3}, 1\frac{2}{3}, 1\frac{2}{3}, 1\frac{2}{3}, 2, 2, 2, 2\frac{1}{3}, 2\frac{1}{3}, 2\frac{2}{3}$

The values $1\frac{2}{3}$ and 2 occur most often.

The mean is $1\frac{29}{30}$ hours, the median is 2 hours, and the modes are $1\frac{2}{3}$ hours and 2 hours.

8. **Mean:** $\frac{628}{8} = 78.5$

Median: 65, 70, 75, 75, **80**, 84, 85, 94

The mean of the middle values is

$$\frac{75 + 80}{2} = \frac{155}{2} = 77.5.$$

Mode: 65, 70, **75, 75**, 80, 84, 85, 94

The value that occurs most often is 75.

The mean is 78.5 megabytes, the median is 77.5 megabytes, and the mode is 75 megabytes.

9. Because the mean is 6 and there are 6 data values, the sum of the values is $6 \cdot 6 = 36$. Find the value of x .

$$2 + 8 + 9 + 7 + 6 + x = 36$$

$$32 + x = 36$$

$$x = 4$$

So, the value of x is 4.

10. Because the mean is 0 and there are 5 data values, the sum of the values is $0 \cdot 5 = 0$. Find the value of x .

$$11.5 + 12.5 + (-10) + (-7.5) + x = 0$$

$$6.5 + x = 0$$

$$x = -6.5$$

So, the value of x is -6.5.

11. Because there is an even number of data values, the median is the mean of the two middle values, 12 and x . Find the value of x .

$$\frac{12 + x}{2} = 14$$

$$12 + x = 28$$

$$x = 16$$

So, the value of x is 16.

12. Because there is an even number of data values, the median is the mean of the two middle values, 45 and x . Find the value of x .

$$\frac{45 + x}{2} = 51$$

$$45 + x = 102$$

$$x = 57$$

So, the value of x is 57.

13. Because the mean is 410 and there are 8 data values, the sum of the values is $410 \cdot 8 = 3280$. Find the value of x .

$$455 + 364 + 262 + 553 + 471 + 352 + 358 + x = 3280$$

$$2815 + x = 3280$$

$$x = 465$$

So, the value of x is 465.

14. a. The value of 6 is high compared to the other values in the rookie season. 6 is the outlier in the rookie season. The value of 14 is high compared to the other values in this season. 14 is the outlier in this season.

b. Because both outliers are the highest values in each season, the mean would be most affected by removing the outlier.

Chapter 12

c. Rookie season:

Ordered data: 0, 0, 1, 2, 2, 6

$$\text{Mean: } \frac{0 + 0 + 1 + 2 + 2 + 6}{6} = \frac{11}{6} \approx 1.83$$

Median: The middle values are 1 and 2.

$$\frac{1 + 2}{2} = \frac{3}{2} = 1.5$$

Mode: The values 0 and 2 occur most often.

This season:

Ordered data: 4, 4, 6, 7, 8, 14

$$\text{Mean: } \frac{4 + 4 + 6 + 7 + 8 + 14}{6} = \frac{43}{6} \approx 7.17$$

Median: The middle values are 6 and 7.

$$\frac{6 + 7}{2} = \frac{13}{2} = 6.5$$

Mode: The value 4 occurs most often.

All the measures increased from the rookie season to this season. The mean increased by about 5.34, the median increased by 5, and the mode increased from 0 and 2 to 4.

15. a. Ordered data: 0.1, 1.8, 3.3, 5.3, 6.3, 8.8, 8.8, 14, 14.3, 16.7

$$\text{Mean: } \frac{79.4}{10} = 7.94$$

$$\text{Median: } \frac{6.3 + 8.8}{2} = \frac{15.1}{2} = 7.55$$

Mode: The value 8.8 occurs twice.

The mean is 7.94 miles, the median is 7.55 miles, and the mode is 8.8 miles.

- b. By skipping the first shelter, 0.1 mile is removed from the data values.

Ordered data: 1.8, 3.3, 5.3, 6.3, 8.8, 8.8, 14, 14.3, 16.7

$$\text{Mean: } \frac{79.3}{9} \approx 8.8$$

Median: The middle value is 8.8.

Mode: The value 8.8 occurs twice.

The mean is now about 8.8 miles, the median is now 8.8 miles, and the mode is now 8.8 miles.

So, the mean increases by about

$$8.8 - 7.94 = 0.86 \text{ mile, the median increases by}$$

$$8.8 - 7.55 = 1.25 \text{ miles, and the mode does not}$$

change.

16. Ordered data: 10, 11, 13, 14, 14, 15, 16, 17, 18

$$\text{Mean: } \frac{128}{9} = 14.\bar{2}$$

Median: The middle value is 14.

Mode: The value 14 occurs most often.

Because 3 is added to each value in the data set, all the measures increase by 3.

17. Adding k to each value in the data set will cause all the measures to increase by k .

18. Ordered data: $-3.06, -2.69, -2.46, -1.83, -0.43, 0.61, 0.62, 1.00, 1.34, 1.98, 2.15, 3.97$

$$\text{Mean: } \frac{1.2}{12} = 0.1$$

Median: The middle values are 0.61 and 0.62.

$$\frac{0.61 + 0.62}{2} = \frac{1.23}{2} = 0.615$$

Mode: There is no value that occurs most often.

Because \$0.05 is subtracted from each value in the data set, the mean and median decrease by 0.05 and there is still no mode.

19. Subtracting k from each value in a data set will cause all the measures to decrease by k .

20. a. Find the number of students for each age.

$$18 \text{ yr: } 0.35 \times 200 = 70 \text{ students}$$

$$19 \text{ yr: } 0.30 \times 200 = 60 \text{ students}$$

$$20 \text{ yr: } 0.14 \times 200 = 28 \text{ students}$$

$$21 \text{ yr: } 0.20 \times 200 = 40 \text{ students}$$

$$37 \text{ yr: } 0.01 \times 200 = 2 \text{ students}$$

$$\begin{aligned} \text{Mean: } & \frac{70 \cdot 18 + 60 \cdot 19 + 28 \cdot 20}{200} \\ & + \frac{40 \cdot 21 + 2 \cdot 37}{200} = \frac{3874}{200} \\ & = 19.37 \text{ years} \end{aligned}$$

Median: Because there are 200 data values, the median is the mean of the 100th and 101st data values.

$$\text{The median is } \frac{19 + 19}{2} = \frac{38}{2} = 19 \text{ years.}$$

Mode: Because the age with the most number of students is 18 years, the mode is 18 years.

The mean is 19.4 years, the median is 19 years, and the mode is 18 years.

- b. Because 37 years is much older than the other ages, the outliers are the 2 students who are 37 years old.

$$\begin{aligned} \text{Mean: } & \frac{70 \cdot 18 + 60 \cdot 19 + 28 \cdot 20 + 40 \cdot 21}{198} \\ & = \frac{3800}{198} \approx 19.19 \text{ years} \end{aligned}$$

Median: Because there are 198 data values, the median is the mean of the 98th and 99th data values.

$$\text{The median is } \frac{19 + 19}{2} = \frac{38}{2} = 19 \text{ years.}$$

Mode: Because the age with the most students is 18 years, the mode is 18 years.

By removing the outliers, the mean decreases about $19.37 - 19.19 = 0.18$ year. The median and the mode stay the same.

Chapter 12

21. $0.10(7.20) = 0.72$

$0.72 + 7.20 = 7.92$

With the information given, the increase in wages of 10% increases the mean and median hourly wage at the bagel shop to \$7.92, which is greater than these measures for the amusement park in Example 1.

So, the bagel shop is likely to pay a greater hourly wage.

Fair Game Review

22. From least to greatest, the values are $-8, -5, -3, 1, 4,$ and $7.$

23. From least to greatest, the values are

$-4.7, -2.8, -\frac{2}{3}, 1.2, \frac{3}{2},$ and $5.4.$

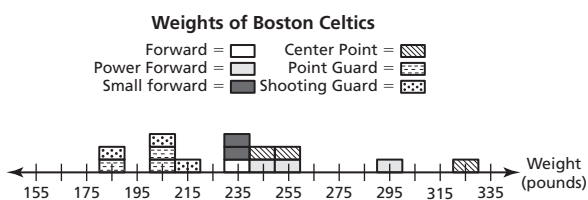
24. B; Because $y = 2x$ can be written in the form

$Ax + By = C,$ the equation is linear.

Section 12.2

12.2 Activity (pp. 612–613)

1. a. The data are very dispersed from the mean. The diagram shows the data spread out rather than all clustered near the mean.
- b. Yes, each position seems to be focused at a more specific weight range.
Yes, in most sports, there is some correlation between position and weight, but the correlation is most distinctive in football.
2. a. The weights seem to be clustered more toward the mean. The data are not very dispersed.
- b. The weights of the players for a National Football League team are more dispersed than the weights of the players for a Major League Baseball team.
- c. No, the various positions are spread throughout the data.
3. You can measure the dispersion of a data set based on how the data lie relative to the mean.



The weights appear to be rather dispersed.

Apart from a couple of outliers, the positions seem to be grouped around specific weights. So, the weight seems to be correlated to the position.

12.2 On Your Own (pp. 614–615)

1. Show A:

Mean: $\frac{275}{11} = 25$

Range: Ordered data: 19, 20, 21, 22, 22, 24, 29, 29, 30, 31
The least value is 19. The greatest value is 31.

So, the range is $31 - 19,$ or 12 years.

The mean and range remained the same.

Show B:

Mean: $\frac{264}{11} = 24$

Range: Ordered data: 19, 20, 21, 22, 24, 25, 25, 27, 27, 32

The least value is 19. The greatest value is 32.

So, the range is $32 - 19,$ or 13 years.

The mean and range both decreased.

The mean and range for Show A did not change because the data value which was removed, 25, is the mean of the data set. The mean and range for Show B decreased because the data value which was removed, 48, is an outlier that is higher than the other data values.

2.

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
25	26	-1	1
19	26	-7	49
20	26	-6	36
27	26	1	1
22	26	-4	16
25	26	-1	1
27	26	1	1
22	26	-4	16
48	26	22	484
21	26	-5	25
32	26	6	36
24	26	-2	4

$$\frac{1 + 49 + 36 + 1 + 16 + 1 + 1 + 16 + 484 + 25 + 36 + 4}{12} \approx 55.8$$

$$\sqrt{55.8} \approx 7.5$$

So, the standard deviation is about 7.5. This means that the typical age of a contestant on Show B differs from the mean by about 7.5 years.

Chapter 12

3. The standard deviation for Show B is greater than the standard deviation for Show A. This means that the ages of the contestants on Show B are more spread out than the ages of the contestants on Show A.

12.2 Exercises (pp. 616–617)

Vocabulary and Concept Check

1. A measure of central tendency represents a typical value in the data set.
A measure of dispersion represents how spread out the data set is.

2. An advantage to using the range is that it is easier to calculate.

The standard deviation is considered more reliable because it uses all of the values in the data set in its calculation. This means it is less affected by outliers than the range, which only uses two values from the data set in its calculation.

Practice and Problem Solving

3. The mean is 85. The data are relatively dispersed. The dot plot shows that the data are spread out, clustered away from the mean.

4. The mean is 85. The data are not very dispersed. The dot plot shows that the data are clustered around the mean.

5. Tigers:

$$\text{Mean: } \frac{552}{8} = 69$$

The mean is 69 inches.

Range: Ordered data: 65, 67, 67, 68, 69, 70, 72, 74

The least value is 65. The greatest value is 74.

The range is $74 - 65$, or 9 inches.

Centaurs:

$$\text{Mean: } \frac{552}{8} = 69$$

Range: Ordered data: 63, 63, 65, 68, 71, 73, 74, 75

The least value is 63. The greatest value is 75.

The range is $75 - 63$, or 12 inches.

The means of the heights of both teams are equal. The range of the heights of the Centaurs is more than the range of the heights of the Tigers. So, the heights of the Centaurs are more spread out than the heights of the Tigers.

6. Crew A:

$$\text{Mean: } \frac{728}{7} = 104$$

The mean is 104 fish.

Range: Ordered data: 75, 80, 100, 106, 112, 120, 135

The least value is 75. The greatest value is 135.

The range is $135 - 75$, or 60 fish.

Crew B:

$$\text{Mean: } \frac{1001}{7} = 143 \quad \text{The mean is 143 fish.}$$

Range: Ordered data: 104, 135, 140, 140, 158, 159, 165

The least value is 104. The greatest value is 165.

The range is $165 - 104$, or 61 fish.

The mean number of fish for Crew B is much more than the mean number of fish for Crew A. This means Crew B generally catches more fish. The ranges are about the same. So, the spreads of each data set are very similar.

7. **Mean:** $\frac{40}{8} = 5$ The mean is 5.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
4	5	-1	1
2	5	-3	9
7	5	2	4
3	5	-2	4
6	5	1	1
5	5	0	0
5	5	0	0
8	5	3	9

$$\frac{1 + 9 + 4 + 4 + 1 + 0 + 0 + 9}{8} = 3.5$$

$$\sqrt{3.5} \approx 1.9$$

The standard deviation is about 1.9.

8. **Mean:** $\frac{63}{7} = 9$ The mean is 9.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
12	9	3	9
4	9	-5	25
8	9	-1	1
7	9	-2	4
9	9	0	0
13	9	4	16
10	9	1	1

$$\frac{9 + 25 + 1 + 4 + 0 + 16 + 1}{7} = 8$$

$$\sqrt{8} \approx 2.8$$

The standard deviation is about 2.8.

Chapter 12

9. Mean: $\frac{210}{10} = 21$

The mean is 21.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
18	21	-3	9
20	21	-1	1
20	21	-1	1
20	21	-1	1
21	21	0	0
21	21	0	0
21	21	0	0
22	21	1	1
23	21	2	4
24	21	3	9

$$\frac{9 + 1 + 1 + 1 + 0 + 0 + 0 + 1 + 4 + 9}{10} = 2.6$$

$$\sqrt{2.6} \approx 1.6$$

The standard deviation is about 1.6.

10. Mean: $\frac{342}{9} = 38$

The mean is 38.

Standard deviation:

x	\bar{x}	$(x - \bar{x})$	$(x - \bar{x})^2$
36	38	-2	4
37	38	-1	1
37	38	-1	1
38	38	0	0
38	38	0	0
38	38	0	0
39	38	1	1
39	38	1	1
40	38	2	4

$$\frac{4 + 1 + 1 + 0 + 0 + 0 + 1 + 1 + 4}{9} \approx 1.33$$

$$\sqrt{1.33} \approx 1.2$$

The standard deviation is about 1.2.

11. Mean: $\frac{504}{8} = 63$

The mean is 63.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
40	63	-23	529
52	63	-11	121
61	63	-2	4
64	63	1	1
65	63	2	4
67	63	4	16
73	63	10	100
82	63	19	361

$$\frac{529 + 121 + 4 + 1 + 4 + 16 + 100 + 361}{8} = 142$$

$$\sqrt{142} \approx 11.9$$

The standard deviation is about 11.9.

12. Mean: $\frac{174}{12} = 14.5$

The mean is 14.5.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
13	14.5	-1.5	2.25
13	14.5	-1.5	2.25
13	14.5	-1.5	2.25
13	14.5	-1.5	2.25
14	14.5	-0.5	0.25
14	14.5	-0.5	0.25
14	14.5	-0.5	0.25
15	14.5	0.5	0.25
15	14.5	0.5	0.25
16	14.5	1.5	2.25
17	14.5	2.5	6.25
17	14.5	2.5	6.25

$$\frac{4(2.25) + 5(0.25) + 2.25 + 2(6.25)}{12} \approx 2.083$$

$$\sqrt{2.083} \approx 1.4$$

The standard deviation is about 1.4.

Chapter 12

13. a. Kirsten:

Mean: $\frac{900}{10} = 90$

Range: Ordered data: 83, 84, 87, 88, 89, 90, 91, 95, 95, 98

The least value is 83. The greatest value is 98.

The range is $98 - 83$, or 15.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
83	90	-7	49
88	90	-2	4
84	90	-6	36
95	90	5	25
91	90	1	1
89	90	-1	1
90	90	0	0
87	90	-3	9
98	90	8	64
95	90	5	25

$$\frac{49 + 4 + 36 + 25 + 1 + 1 + 0 + 9 + 64 + 25}{10} = 21.4$$

$$\sqrt{21.4} \approx 4.6$$

The standard deviation is about 4.6.

Leah:

Mean: $\frac{910}{10} = 91$

Range: Ordered data: 87, 88, 89, 89, 91, 92, 92, 93, 94, 95

The least value is 87. The greatest value is 95.

The range is $95 - 87$, or 8.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
89	91	-2	4
87	91	-4	16
93	91	2	4
95	91	4	16
92	91	1	1
94	91	3	9
88	91	-3	9
91	91	0	0
89	91	-2	4
92	91	1	1

$$\frac{4 + 16 + 4 + 16 + 1 + 9 + 9 + 0 + 4 + 1}{10} = 6.4$$

$$\sqrt{6.4} \approx 2.5$$

The standard deviation is about 2.5.

The mean score is about the same for both golfers. The range and standard deviation of Leah's scores are less than the range and standard deviation of Kirsten's scores.

- b.** Leah; Because Leah's range and standard deviation are lower, Kirsten's scores are more spread out. So, Leah is more consistent.

Chapter 12

14. Kirsten:

Mean: $\frac{990}{11} = 90$

Range: Ordered data: 83, 84, 87, 88, 89, 90, 90, 91, 95, 95, 98

The least value is 83. The greatest value is 98.

The range is $98 - 83$, or 15.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
83	90	-7	49
88	90	-2	4
84	90	-6	36
95	90	5	25
91	90	1	1
89	90	-1	1
90	90	0	0
87	90	-3	9
98	90	8	64
95	90	5	25
90	90	0	0

$$\frac{49 + 4 + 36 + 25 + 1 + 1 + 0 + 9 + 64 + 25 + 0}{11} \approx 19.45$$

$$\sqrt{19.45} \approx 4.4$$

The standard deviation is about 4.4.

Leah:

Mean: $\frac{990}{11} = 90$

Range: Ordered data: 80, 87, 88, 89, 89, 91, 92, 92, 93, 94, 95

The least value is 80. The greatest value is 95.

The range is $95 - 80$, or 15.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
89	90	-1	1
87	90	-3	9
93	90	3	9
95	90	5	25
92	90	2	4
94	90	4	16
88	90	-2	4
91	90	1	1
89	90	-1	1
92	90	2	4
80	90	-10	100

$$\frac{1 + 9 + 9 + 25 + 4 + 16 + 4 + 1 + 1 + 4 + 100}{11} \approx 15.82$$

$$\sqrt{15.82} \approx 4.0$$

The standard deviation is about 4.0.

Kirsten's mean and range remained the same and her standard deviation decreased slightly.

Leah's mean decreased and her range increased. Her mean and range are now the same as Kirsten's. Her standard deviation increased.

Leah's statistics changed more than Kirsten's because her new score was an outlier, much lower than her other scores, while Kirsten's new score was equal to her mean score.

Chapter 12

15. **Mean:** $\frac{32}{5} = 6.4$

The mean is 6.4.

Range: Ordered data: 2.3, 4.1, 6.4, 8.7, 10.5

The least value is 2.3. The greatest value is 10.5.

The range is $10.5 - 2.3 = 8.2$.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
4.1	6.4	-2.3	5.29
2.3	6.4	-4.1	16.81
8.7	6.4	2.3	5.29
10.5	6.4	4.1	16.81
6.4	6.4	0	0

$$\frac{5.29 + 16.81 + 5.29 + 16.81 + 0}{5} = 8.84$$

$$\sqrt{8.84} \approx 3.0$$

The standard deviation is about 3.0.

16. **Mean:** $\frac{-8}{8} = -1$

The mean is -1.

Range: Ordered data: -5, -4, -3, -2, 0, 1, 2, 3

The least value is -5. The greatest value is 3.

The range is $3 - (-5) = 8$.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
-2	-1	-1	1
0	-1	1	1
-5	-1	-4	16
1	-1	2	4
3	-1	4	16
-4	-1	-3	9
2	-1	3	9
-3	-1	-2	4

$$\frac{1 + 1 + 16 + 4 + 16 + 9 + 9 + 4}{8} = 7.5$$

$$\sqrt{7.5} \approx 2.7$$

The standard deviation is about 2.7.

17. No, the following sets are a counterexample.

Set A: 1, 4, 5, 5, 5, 6, 10

Set B: 1, 2, 2, 5, 9, 9, 10

Set A: **Mean:** $\frac{36}{7} \approx 5.1$ The mean is about 5.1.

Range: $10 - 1 = 9$ The range is 9.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
1	5.14	-4.14	17.14
4	5.14	-1.14	1.30
5	5.14	-0.14	0.02
5	5.14	-0.14	0.02
5	5.14	-0.14	0.02
6	5.14	0.86	0.74
10	5.14	4.86	23.62

$$\frac{42.86}{7} = 6.12$$

$$\sqrt{6.12} \approx 2.5$$

The standard deviation is about 2.5.

Set B: **Mean:** $\frac{38}{7} \approx 5.4$ The mean is about 5.4.

Range: $10 - 1 = 9$ The range is 9.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
1	5.43	-4.43	19.62
2	5.43	-3.43	11.76
2	5.43	-3.43	11.76
5	5.43	-0.43	0.18
9	5.43	3.57	12.74
9	5.43	3.57	12.74
10	5.43	4.57	20.88

$$\frac{89.68}{7} \approx 12.81$$

$$\sqrt{12.81} \approx 3.6$$

The standard deviation is about 3.6.

Both ranges are 9, but the standard deviation for Set B is higher than that of Set A.

Chapter 12

18. Data set (b) has the highest standard deviation, because the data is more spread out away from the mean. Data set (c) has the lowest standard deviation, because the data is all clustered by the mean.

19. a. *Answer should include, but is not limited to:* The student will measure the height of the students in the class. The data will be used to find the mean, range, and standard deviation using a calculator.

b. The height of the new student should be an outlier, much higher than the other heights. This will cause the mean, range, and standard deviation to increase.

The student will verify this by adding 7 feet to the data set of heights and recalculating the mean, range, and standard deviation.

20. Doctor's Office B; The standard deviation for Office B is more than the standard deviation for Office A, while the means are about the same. This means more wait times are further away from the mean at Doctor's Office B, which could cause waiting times over 20 minutes.

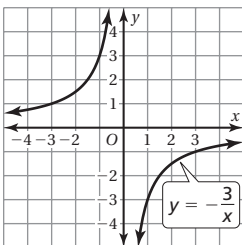
21. Yes; A data set in which all the values are the same has a standard deviation of 0.

No; The equation used to calculate the standard deviation involves a square root, which can never be negative. Also, the standard deviation is a measure of how the data differs from the mean, so it cannot be negative.

Fair Game Review

22. $y = -\frac{3}{x}$

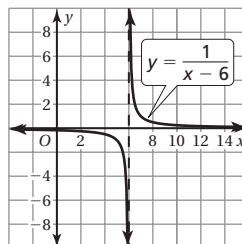
x	-3	-1	0	1	3
y	1	3	undef.	-3	-1



The domain and range are both all real numbers except 0.

23. $y = \frac{1}{x - 6}$

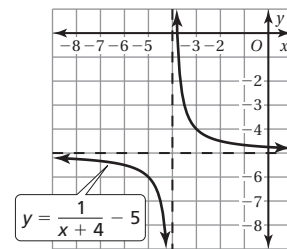
x	3	4	5	6	7	8	9
y	$-\frac{1}{3}$	$-\frac{1}{2}$	-1	undef.	1	$\frac{1}{2}$	$\frac{1}{3}$



The domain is all real numbers except 6. The range is all real numbers except 0.

24. $y = \frac{1}{x + 4} - 5$

x	-8	-6	-5	-4	-3	-2	-1
y	$-\frac{21}{4}$	$-\frac{11}{2}$	-6	undef.	-4	$-\frac{9}{2}$	$-\frac{14}{3}$



The domain is all real number except -4 . The range is all real numbers except -5 .

25. B;

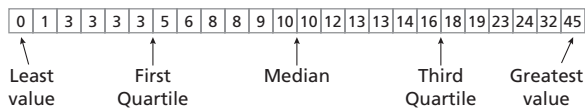
$$\begin{aligned} (x + 5) \div \frac{x^2 + 4x - 5}{x + 5} &= (x + 5) \cdot \frac{x + 5}{x^2 + 4x - 5} \\ &= \frac{x + 5}{1} \cdot \frac{x + 5}{(x + 5)(x - 1)} \\ &= \frac{x + 5}{1} \cdot \frac{1}{x - 1} \\ &= \frac{x + 5}{x - 1} \end{aligned}$$

Chapter 12

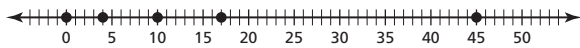
Section 12.3

12.3 Activity (pp. 618–619)

1. a–b.



c.



d. The data set is divided into four equal parts by the first quartile, median, and third quartile. The box-and-whisker plot shows the least value and the greatest value.

2. *Answer should include, but is not limited to:* Students should follow the steps in Activity 1 to draw their box-and-whisker plots.

3. a. *Sample answer:* The upper box-and-whisker plot shows higher scores. The scores for the lower box-and-whisker plot are spread out more than the upper box-and-whisker plot.

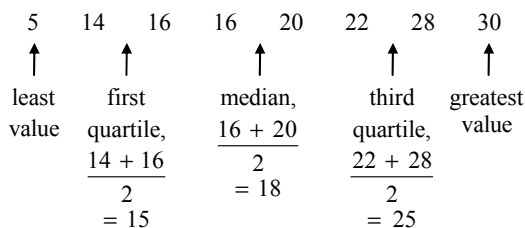
b. The upper box-and-whisker plot is the results of the spring test and the lower box-and-whisker plot is the results of the fall test. This is because the students will have learned more during the school year and will do better on the spring test.

4. You can describe the distribution of the data, the median, and the ranges where one-quarter or one-half of the data fall.

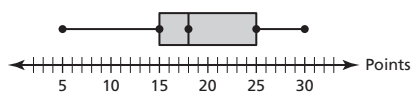
5. *Sample answer:* Teachers would be interested in the test score distributions because if they know why scores are lower, they can determine the reason and try to fix it.

12.3 On Your Own (pp. 620–622)

1. Order the data. Find the median and the quartiles.



Draw a box-and-whisker plot.



2. The least value is 18. The greatest value is 30.

$$\text{Range} = 30 - 18 = 12$$

So, the range is 12 years. This means that no ages vary by more than 12 years.

$$\begin{aligned} \text{Interquartile range} &= \text{third quartile} - \text{first quartile} \\ &= 29 - 23 \\ &= 6 \end{aligned}$$

So, the interquartile range is 6 years. This means that the middle half of the ages vary by no more than 6 years.

3. 25% of the ages are between the least value and the first quartile.

50% of the ages are between the first quartile and the third quartile.

25% of the ages are between the third quartile and the greatest value.

25% of the ages are between 18 and 23 years.

50% of the ages are between 23 and 29 years.

25% of the ages are between 29 and 30 years.

4. For both shops, the right whisker is longer than the left whisker and most of the data are on the left. So, the distribution for both shops is skewed right.

The range of the prices of Shop A is greater than the range of the prices of Shop B. Also, the box for Shop A is longer, meaning the interquartile range is greater. So, the prices of Shop A are more spread out than the prices of Shop B.

12.3 Exercises (pp. 623–625)

Vocabulary and Concept Check

1. Each whisker represents 25% of the data and the box represents 50% of the data.

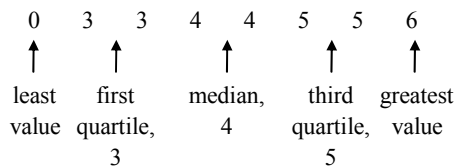
2. Divide the data in half by finding the median. The median of the first half is the first quartile.

3. The length gives the range of the data set. This tells how much the data vary.

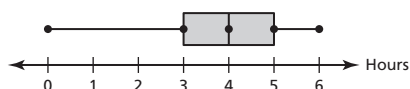
Practice and Problem Solving

4. *Sample answer:* Both sales reps have the same median number of cars sold, but Sales Rep B's numbers are more spread out than Sales Rep A's numbers.

5. Order the data. Find the median and the quartiles.

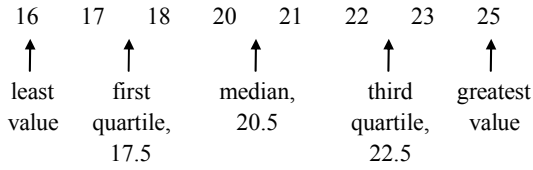


Draw a box-and-whisker plot.

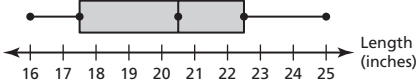


Chapter 12

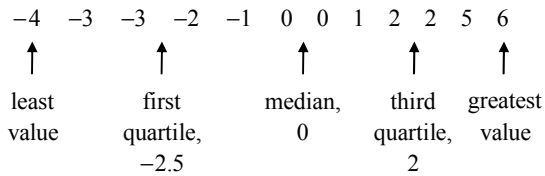
6. Order the data from least to greatest.



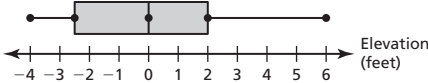
Draw a box-and-whisker plot.



7. Order the data from least to greatest.

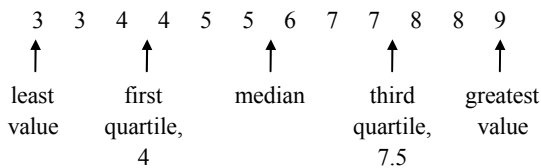


Draw a box-and-whisker plot.

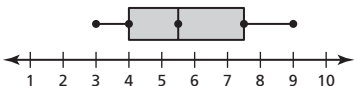


8. The median should be 5.5, not 6.

Order the data from least to greatest.



The median is $\frac{5 + 6}{2} = \frac{11}{2} = 5.5$.



9. a. Ordered data: 7, 8, 8, 9, 10, 10, 12, 12, 13, 14, 14, 14

The middle values are 10 and 12.

The median is $\frac{10 + 12}{2} = 11$ inches.

b. The first half of the data is 7, 8, 8, 9, 10, 10.

The middle values of the first half of the data are 8 and 9.

The median of the first half of the data is

$$\frac{8 + 9}{2} = 8.5.$$

So, the first quartile is 8.5 inches.

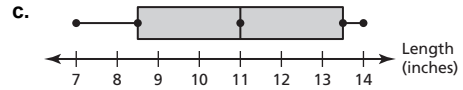
The second half of the data is 12, 12, 13, 14, 14, 14.

The middle values of the second half of the data are 13 and 14.

The median of the second half of the data is

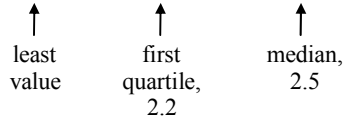
$$\frac{13 + 14}{2} = 13.5.$$

So, the third quartile is 13.5 inches.

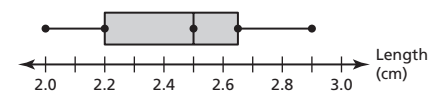
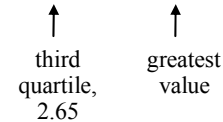


10. a. Ordered data:

2.0 2.1 2.1 2.3 2.4 2.5 2.5 2.6



2.6 2.7 2.8 2.9



b. The least value is 2.0. The greatest value is 2.9.

So, the range is $2.9 - 2.0 = 0.9$ centimeter. This means the lengths vary by no more than 0.9 centimeter.

c. 25% of the lengths are between 2.0 and 2.2 centimeters.

50% of the lengths are between 2.2 and 2.65 centimeters.

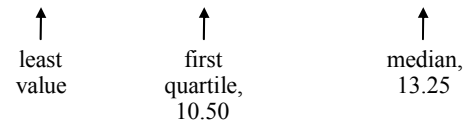
25% of the lengths are between 2.65 and 2.9 centimeters.

d. Interquartile range = third quartile – first quartile
 $= 2.65 - 2.2$
 $= 0.45$

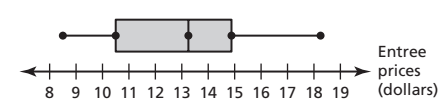
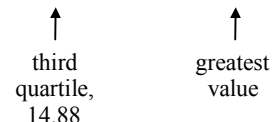
So, the interquartile range is 0.45 centimeter. This means the middle half of the lengths vary by no more than 0.45 centimeter.

11. a. Ordered data:

8.50 9.00 10.00 11.00 12.00 12.50 14.00



14.00 14.75 15.00 17.00 18.25



Chapter 12

b. Interquartile range = third quartile – first quartile
 $= 14.88 - 10.50$
 $= 4.38$

So, the inter quartile range is \$4.38. This means the middle half of the prices vary by no more than \$4.38.

- c. 25% of the prices are between \$8.50 and \$10.50.
 50% of the prices are between \$10.50 and \$14.88.
 25% of the prices are between \$14.88 and \$18.25.

d. Mean: $\frac{156}{12} = 13$

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
8.50	13.00	-4.50	20.25
9.00	13.00	-4.00	16.00
10.00	13.00	-3.00	9.00
11.00	13.00	-2.00	4.00
12.00	13.00	-1.00	1.00
12.50	13.00	-0.50	0.25
14.00	13.00	1.00	1.00
14.00	13.00	1.00	1.00
14.75	13.00	1.75	3.06
15.00	13.00	2.00	4.00
17.00	13.00	4.00	16.00
18.25	13.00	5.25	27.56

$$\frac{20.25 + 16 + 9 + 4 + 1 + 0.25 + 1 + 1 + 3.06 + 4 + 16 + 27.56}{12} \approx 8.59$$

$$\sqrt{8.59} \approx 2.93$$

So, the standard deviation is about 2.93. This means most of the prices differ from the mean by about \$2.93.

12. The range is 36. The interquartile range is 12. The range represents the spread of all of the data. The interquartile range represents the spread of the middle half of the data. So, the interquartile range will be less than the range.
13. The whiskers are about the same length and the median is in the middle of the data. So, the box-and-whisker plot is symmetric.
14. The right whisker is longer than the left whisker and most of the data are on the left. So, the box-and-whisker plot is skewed right.
15. The left whisker is longer than the right whisker and most of the data are on the right. So, the box-and-whisker plot is skewed left.

16. The whiskers are about the same length and the median is in the middle of the data. So, the box-and-whisker plot is symmetric.

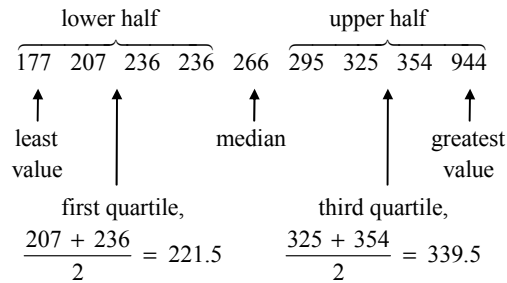
17. There are not more data values to the right of the median. The number of data values on each side of the median is the same.

18. *Answers will vary. Sample answer:*

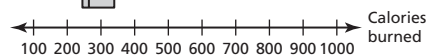
Data representing the heights of the students in the class would tend to be symmetric. Heights tend to distribute evenly around the average height.

Data representing test scores may not be symmetric if a student does very poorly on a test that most students do well on, or a student does well on a test that most students do poorly on.

19. a. Order the data from least to greatest.

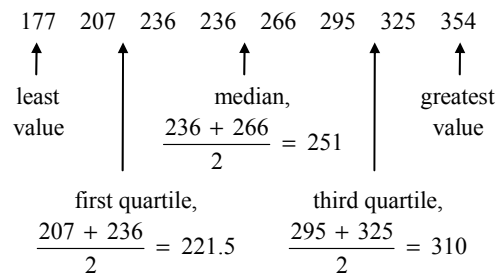


Draw a box-and-whisker plot.

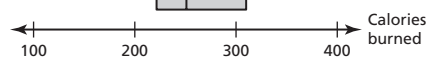


b. Because 944 is much greater than the other data values, it is the outlier.

c. Order the data from least to greatest without the outlier.



Draw a box-and-whisker plot.



d. The outlier makes the right whisker longer, increases the length of the box, increases the third quartile, and increases the median. In this case, the first quartile and left whisker were not affected.

Chapter 12

20. a. The whiskers of the box-and-whisker plot for Brand A are about the same length, and the median is in the middle of the distribution.

So, the distribution of Brand A is symmetric.

The right whisker of the box-and-whisker plot for Brand B is longer than the left whisker and most of the data is on the left.

So, the distribution of Brand B is skewed right.

- b. The upper 75% includes the data values between the first quartile and the greatest value. The range of Brand A is $7 - 3.5 = 3.5$ hours. The range of Brand B is about $6 - 3.3 = 2.7$ hours.

- c. Interquartile range = third quartile – first quartile

Brand A: $5 - 3.5 = 1.5$

The interquartile range of Brand A is 1.5. hours.

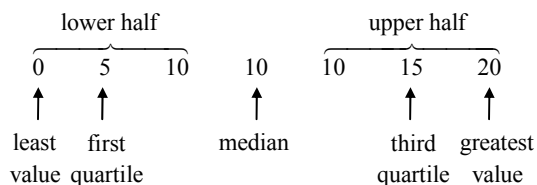
Brand B: $4.25 - 3.25 = 1$

The interquartile range of Brand B is 1 hour.

The interquartile range of Brand A is greater than the interquartile range of Brand B. This means the middle half of the battery lives of Brand A vary more than the middle half of the battery lives of Brand B.

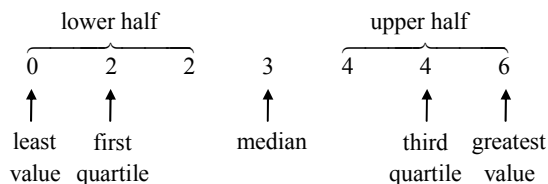
- d. Brand A; The battery lives for Brand A are much more spread out than the battery lives of Brand B. Also, the middle half of the battery lives for Brand A are more spread out than the middle half of the battery lives for Brand B.

21. *Sample answer:*



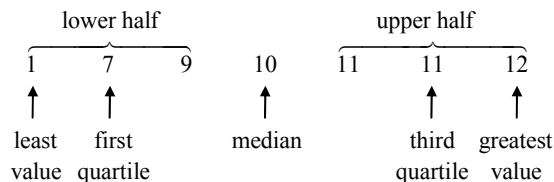
Because there is a difference of 5 between the least value, first quartile, median, third quartile, and greatest value, they are all equally spaced.

22. *Sample answer:*



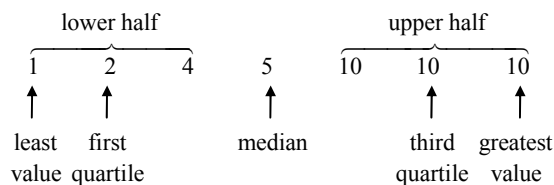
The length of the left whisker is $2 - 0 = 2$, the length of the box is $4 - 2 = 2$, and the length of the right whisker is $6 - 4 = 2$.

23. *Sample answer:*



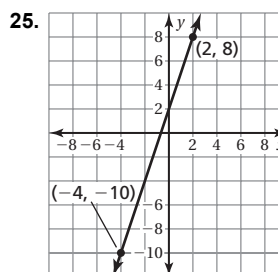
The length of the box between the first quartile and the median is $10 - 7 = 3$, and the length of the box between the median and the third quartile is $11 - 10 = 1$, and $3 = 3(1)$.

24. *Sample answer:*



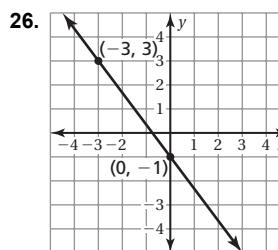
Because the distance between the third quartile and the greatest value is $10 - 10 = 0$, there is no right whisker.

Fair Game Review



$$\text{slope} = \frac{\text{rise}}{\text{run}} = \frac{18}{6} = 3$$

Because the line crosses the y -axis at $(0, 2)$, the y -intercept is 2. So, an equation of the line is $y = 3x + 2$.

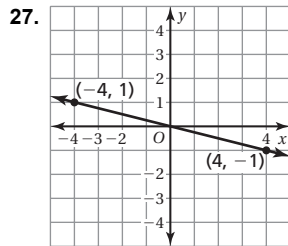


$$\text{slope} = \frac{\text{rise}}{\text{run}} = \frac{-4}{3} = -\frac{4}{3}$$

Because the line crosses the y -axis at $(0, -1)$, the y -intercept is -1 . So, an equation of the line is

$$y = -\frac{4}{3}x - 1.$$

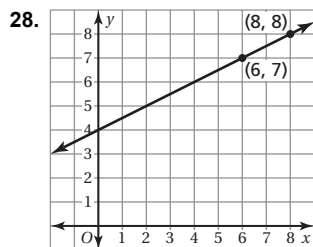
Chapter 12



$$\text{slope} = \frac{\text{rise}}{\text{run}} = \frac{-2}{8} = -\frac{1}{4}$$

Because the line crosses the y -axis at $(0, 0)$, the

y -intercept is 0. So, an equation of the line is $y = -\frac{1}{4}x$.



$$\text{slope} = \frac{\text{rise}}{\text{run}} = \frac{1}{2}$$

Because the line crosses the y -axis at $(0, 4)$, the y -intercept is 4. So, an equation of the line is

$$y = \frac{1}{2}x + 4.$$

29. B;

$$\begin{array}{r} 2z - 7 \\ z - 3 \overline{) 2z^2 - 13z + 21} \\ \underline{2z^2 - 6z} \\ -7z + 21 \\ \underline{-7z + 21} \\ 0 \end{array}$$

Section 12.4

12.4 Activity (pp. 626–627)

1. a. $40 - 2 = 38$

$$40 + 2 = 42$$

The chest sizes within 1 standard deviation of the mean are 38 to 42 inches.

$$749 + 1073 + 1079 + 934 + 658 = 4493$$

4493 chest sizes lie within 1 standard deviation of the mean.

$$\frac{4493}{5738}(100\%) \approx 78.3\%$$

So, about 78.3% of the chest sizes lie within 1 standard deviation of the mean.

b. $40 - 2(2) = 36$

$$40 + 2(2) = 44$$

The chest sizes within 2 standard deviations of the mean are 36 to 44 inches.

$$\begin{aligned} 185 + 420 + 749 + 1073 + 1079 + 934 + 658 \\ + 370 + 92 = 5560 \end{aligned}$$

5560 chest sizes lie within 2 standard deviations of the mean.

$$\frac{5560}{5738}(100\%) \approx 96.9\%$$

So, about 96.9% of the chest sizes lie within 2 standard deviations of the mean.

c. $40 - 3(2) = 34$

$$40 + 3(2) = 46$$

The chest sizes within 3 standard deviations of the mean are 34 to 46 inches.

$$\begin{aligned} 18 + 81 + 185 + 420 + 749 + 1073 + 1079 + 934 \\ + 658 + 370 + 92 + 50 + 21 = 5730 \end{aligned}$$

5730 chest sizes lie within 3 standard deviations of the mean.

$$\frac{5730}{5738}(100\%) = 99.9\%$$

So, about 99.9% of the chest sizes lie within 3 standard deviations of the mean.

2. a. Adult female heights have a smaller standard deviation.

This means that adult female heights differ from the average adult female height less than adult male heights differ from the average adult male height.

b. $19 + 32 + 25 + 32 + 30 + 26 + 23 = 187$

$$\frac{187}{250}(100\%) \approx 74.8\%$$

About 74.8% of male heights are between 67 and 73 inches.

3. *Answers will vary. Sample answer:*

A histogram shows how the data is distributed. It shows if most of the data are clustered in the middle or to the left or right. It also shows how the data are spread out and how the data differ from the mean.

4. a. In a symmetric distribution, the data are evenly distributed on each side of the highest bar in the histogram.

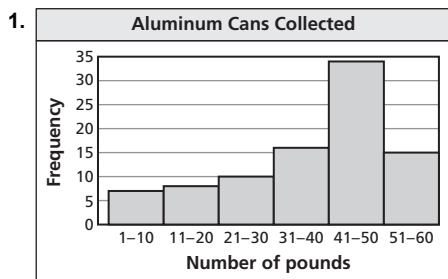
b. A curve drawn to represent the general shape of the histogram of a symmetric distribution looks like a bell.

c. *Answers will vary. Sample answer:*

IQ scores and body temperature both have symmetric distributions.

Chapter 12

12.4 On Your Own (pp. 628–630)



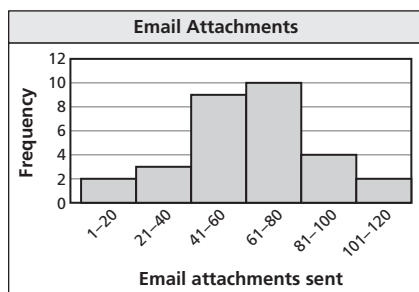
The “tail” of the graph extends to the left and most of the data are on the right.

So, the distribution is skewed left.

2. The median is the best measure of central tendency, because the distribution is skewed left.

3. a.

Number of Attachments	Frequency
1–20	2
21–40	3
41–60	9
61–80	10
81–100	4
101–120	2



- b. The mean and standard deviation best represent the data, because the distribution is symmetric.

12.4 Exercises (pp. 631–633)

Vocabulary and Concept Check

- A skewed distribution will have a tail on either the right or left. A symmetric distribution is even, or symmetrical, with respect to the mean.
- The center of a skewed distribution is best described by the median. The center of a symmetric distribution is best described by the mean.

Practice and Problem Solving

3. $4 - 2 = 2$

$4 + 2 = 6$

The numbers of glasses within 2 standard deviations of the mean are 2 to 6 glasses.

$3 + 23 + 28 + 25 + 5 = 84$

84 people are within 2 standard deviations of the mean water consumption.

$2 + 3 + 23 + 28 + 25 + 5 + 2 = 88$

There are 88 total data values.

$\frac{84}{88}(100\%) \approx 95\%$

About 95% of people are within 2 standard deviations of the mean water consumption.

4. $8 - 2 = 6$

$8 + 2 = 10$

The numbers of hours of sleep within 2 standard deviations of the mean are 6 to 10 hours.

$4 + 10 + 18 + 6 + 2 = 40$

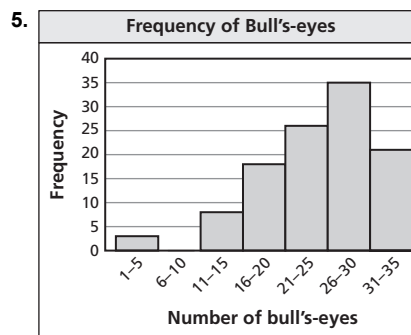
40 people are within 2 standard deviations of the mean sleep duration.

$1 + 4 + 10 + 18 + 6 + 2 + 1 = 42$

There are 42 total data values.

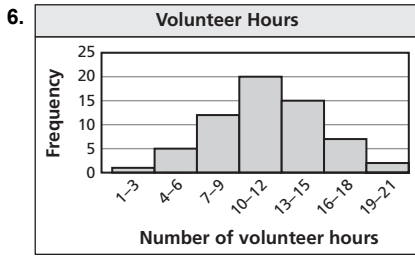
$\frac{40}{42}(100\%) \approx 95\%$

About 95% of people are within 2 standard deviations of the mean sleep duration.

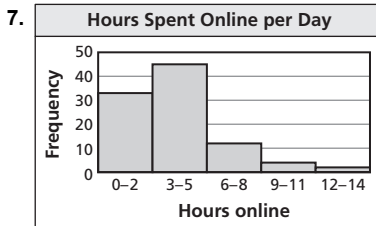


The “tail” of the graph extends to the left and most of the data are on the right. So, the distribution is skewed left.

Chapter 12



The data are evenly distributed on each side of the highest bar. So, the distribution is symmetric.



The “tail” of the graph extends to the right and most of the data are on the left. So, the distribution is skewed right.

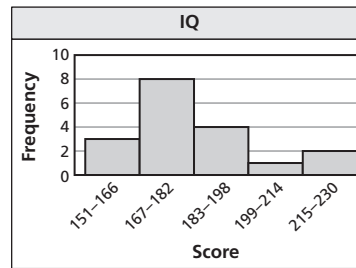
8. The distribution is skewed right. So, the center of the data is best represented by the median and the dispersion of the data is best represented by the five-number summary.
9. The distribution is skewed left. So, the center of the data is best represented by the median and the dispersion of the data is best represented by the five-number summary.
10. The distribution is symmetric. So, the center of the data is best represented by the mean and the dispersion of the data is best represented by the standard deviation.
11. The distribution is skewed left. So, the center of the data is best represented by the median and the dispersion of the data is best represented by the five-number summary.
12. C;
The distribution is symmetric.
13. A;
The distribution is skewed right.
14. B;
The distribution is skewed left.

15. *Answers will vary. Sample answer:*

Histogram; A histogram uses intervals, which makes it easy to show the distribution of a large data set. A stem-and-leaf plot uses every data value, which makes displaying large data sets very tedious.

16. a.

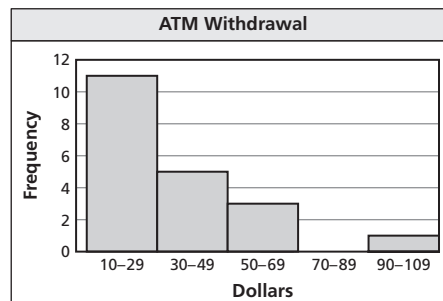
IQ Score	Frequency
151–166	2
167–182	8
183–198	4
199–214	1
215–230	2



- b. The distribution is skewed right. So, the center of the data is best represented by the median and the dispersion of the data is best represented by the five-number summary.
- c. As more IQ scores are included in the data set, there will be more scores below the mean. This will cause the distribution to become more symmetric.
17. *Answers will vary. Sample answer:*

The withdrawals range from \$10 to \$100. To keep a smaller number of intervals, use 5 intervals starting with 10–29.

Withdrawal Amount (dollars)	Frequency
10–29	11
30–49	5
50–69	3
70–89	0
90–109	1



Chapter 12

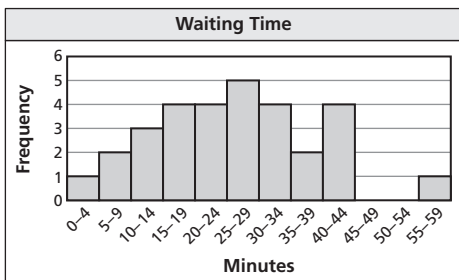
18. a.

Time (minutes)	Frequency
0–9	3
10–19	7
20–29	9
30–39	6
40–49	4
50–59	1



b.

Time (minutes)	Frequency
0–4	1
5–9	2
10–14	3
15–19	4
20–24	4
25–29	5
30–34	4
35–39	2
40–44	4
45–49	0
50–54	0
55–59	1



- c. The distribution flattens out as the number of intervals increases.
- d. The histogram with 6 intervals, from part (a), best represents the data. It shows the distribution better by showing that most of the data values tend to be in the middle of the distribution, with a slight skew to the right.

Fair Game Review

19. $(2, 2); y = x + 3$

Slope: 1

Slope of the perpendicular line: -1

$$y - y_1 = m(x - x_1)$$

$$y - 2 = -1(x - 2)$$

$$y - 2 = -x + 2$$

$$y = -x + 4$$

So, an equation of the perpendicular line is $y = -x + 4$.

20. $(-3, 4); y = 3x - 1$

Slope: 3

Slope of the perpendicular line: $-\frac{1}{3}$

$$y - y_1 = m(x - x_1)$$

$$y - 4 = -\frac{1}{3}(x - (-3))$$

$$y - 4 = -\frac{1}{3}x - 1$$

$$y = -\frac{1}{3}x + 3$$

So, an equation of the perpendicular line is

$$y = -\frac{1}{3}x + 3.$$

21. $(1, -5); y = -\frac{1}{2}x + 4$

Slope: $-\frac{1}{2}$

Slope of the perpendicular line: 2

$$y - y_1 = m(x - x_1)$$

$$y - (-5) = 2(x - 1)$$

$$y + 5 = 2x - 2$$

$$y = 2x - 7$$

So, an equation of the perpendicular line is $y = 2x - 7$.

Chapter 12

22. C;

$$(5, -2), (1, -3)$$

$$\begin{aligned} m &= \frac{y_2 - y_1}{x_2 - x_1} \\ &= \frac{-3 - (-2)}{1 - 5} \\ &= \frac{-1}{-4} \\ &= \frac{1}{4} \end{aligned}$$

$$\text{Slope: } \frac{1}{4}$$

$$y - y_1 = m(x - x_1)$$

$$y - 0 = \frac{1}{4}(x - 0)$$

$$y = \frac{1}{4}x$$

Study Help (p. 634)

Available at *BigIdeasMath.com*.

Quiz 12.1–12.4 (p. 635)

1. **Mean:** $\frac{180}{9} = 20$

Median: -20, -10, 0, 10, **30**, 40, 40, 40, 50

The value in the middle is 30.

Mode: -20, -10, 0, 10, 30, **40, 40, 40**, 50

The value that occurs most often is 40.

The mean is \$20, the median is \$30, and the mode is \$40.

2. **Mean:** $\frac{18}{6} = 3$

Median: $\frac{1}{2}, 2\frac{1}{2}, 3, 3\frac{1}{2}, 3\frac{1}{2}, 5$

The mean of the two middle values is

$$\frac{3 + 3\frac{1}{2}}{2} = \frac{6\frac{1}{2}}{2} = 3\frac{1}{4}$$

Mode: $\frac{1}{2}, 2\frac{1}{2}, 3, 3\frac{1}{2}, 3\frac{1}{2}, 5$

The value that occurs most often is $3\frac{1}{2}$.

The mean is 3 hours, the median is $3\frac{1}{4}$ hours, and the

mode is $3\frac{1}{2}$ hours.

3. Girls:

Mean: $\frac{19}{5} = 3.8$

The mean is 3.8 students.

Range: Ordered data: 2, 3, 4, 4, 6

The range is $6 - 2$, or 4 students.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
2	3.8	-1.8	3.24
3	3.8	-0.8	0.64
4	3.8	0.2	0.04
4	3.8	0.2	0.04
6	3.8	2.2	4.84

$$\frac{0.64 + 0.64 + 0.04 + 0.04 + 4.84}{5} = 1.76$$

$$\sqrt{1.76} \approx 1.3$$

The standard deviation is about 1.3 students.

Boys:

Mean: $\frac{29}{5} = 5.8$

The mean is 5.8 students.

Range: Ordered data: 3, 5, 5, 6, 9

The range is $9 - 3$, or 6 students.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
3	5.6	-2.6	6.76
5	5.6	-0.6	0.36
5	5.6	-0.6	0.36
6	5.6	0.4	0.16
9	5.6	3.4	11.56

$$\frac{6.76 + 0.36 + 0.36 + 0.16 + 11.56}{5} \approx 3.74$$

$$\sqrt{3.74} \approx 1.9$$

The standard deviation is about 1.9 students.

The mean, range, and standard deviation for the boys are greater than for the girls. This means the boys are more likely to be absent, and the number of boys absent each day varies more than the number of girls absent each day.

Chapter 12

4. Juniors: **Mean:** $\frac{132}{8} = 16.5$ The mean is 16.5 points.

Range: Ordered data: 10, 14, 15, 15, 18, 19, 20, 21

The range is $21 - 10$, or 11 points.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
19	16.5	2.5	6.25
15	16.5	-1.5	2.25
20	16.5	3.5	12.25
10	16.5	-6.5	42.25
14	16.5	-2.5	6.25
21	16.5	4.5	20.25
18	16.5	1.5	2.25
15	16.5	-1.5	2.25

$$\frac{6.25 + 2.25 + 12.25 + 42.25 + 6.25 + 20.25 + 2.25 + 2.25}{8} = 11.75$$

$$\sqrt{11.75} \approx 3.4$$

The standard deviation is about 3.4 points.

Seniors: **Mean:** $\frac{192}{8} = 24$ The mean is 24 points.

Median: Ordered data: 15, 19, 19, 22, 26, 29, 30, 32

The range is $32 - 15$, or 17 points.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
22	24	-2	4
19	24	-5	25
29	24	5	25
32	24	8	64
15	24	-9	81
26	24	2	4
30	24	6	36
19	24	-5	25

$$\frac{4 + 25 + 25 + 64 + 81 + 4 + 36 + 25}{8} = 33$$

$$\sqrt{33} \approx 5.7$$

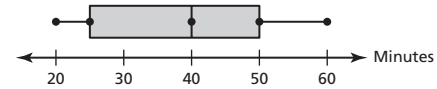
The standard deviation is about 5.7 points.

The mean, range, and standard deviation for the seniors are all greater than for the juniors. This means the seniors are more likely to score more points, and the number of points the seniors score vary more than the number of points the juniors score.

5. Order the data from least to greatest.

20 20 20 30 35 40 40 40 50 50 60 60
 ↑ ↑ ↑ ↑ ↑
 least first median, third greatest
 value quartile, quartile, value
 25 40 50

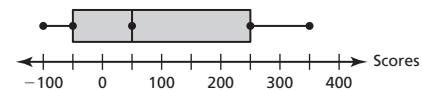
Draw a box-and-whisker plot.



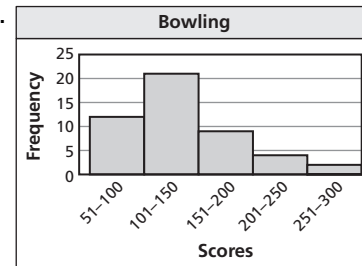
6. Order the data from least to greatest.

-100 -50 -50 0 100 200 300 350
 ↑ ↑ ↑ ↑ ↑
 least first median, third greatest
 value quartile, quartile, value
 -50 50 250

Draw a box-and-whisker plot.



7.



The "tail" of the graph extends to the right and most of the data are on the left. So, the distribution is skewed right.

8. a. **Mean:** $\frac{203.3}{12} \approx 16.9$

Median: Ordered data: 16.5, 16.5, 16.5, 16.7, 16.8, 17.0, 17.0, 17.0, 17.1, 17.3, 17.4, 17.5

The numbers in the middle are 17.0 and 17.0.

$$\frac{17.0 + 17.0}{2} = 17.0$$

Mode: The numbers that occur most often are 16.5 and 17.0.

So, the mean is approximately 16.9 centimeters, the median is 17.0 centimeters, and the modes are 16.5 and 17.0 centimeters.

Chapter 12

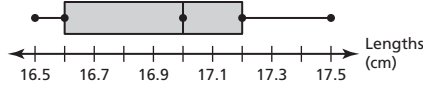
b. 16.5 16.5 16.5 16.7 16.8 17.0 17.0 17.0

↑ least value
↑ first quartile, 16.6
↑ median, 17.0

17.1 17.3 17.4 17.5

↑ third quartile, 17.2
↑ greatest value

Draw a box-and-whisker plot.



c. Interquartile range = third quartile – first quartile
= 17.2 – 16.6
= 0.6

The interquartile range is 0.6 centimeter. This means that the middle half of the lengths vary by no more than 0.6 centimeter.

d. 25% of the lengths are between 16.5 and 16.6 centimeters. 50% of the lengths are between 16.6 and 17.2 centimeters. 25% of the lengths are between 17.2 and 17.5 centimeters.

e. **Mean:** $\frac{211.3}{13} = 16.3$

Median: Ordered data: 8.0, 16.5, 16.5, 16.5, 16.7, 16.8, 17.0, 17.0, 17.0, 17.1, 17.3, 17.4, 17.5

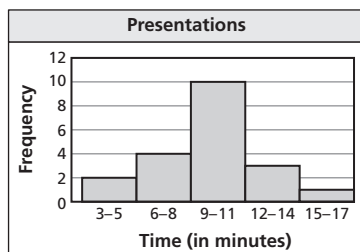
The number that is in the middle is 17.0.

Mode: The numbers that occur most often are 16.5 and 17.0.

The mean decreases to 16.3 centimeters. The median and mode remain the same.

9. a.

Time	Frequency
3–5	2
6–8	4
9–11	10
12–14	3
15–17	1



b. The distribution is symmetric. So, the center of the data is best represented by the mean and the dispersion of the data is best represented by the standard deviation.

Mean: $\frac{193}{20} = 9.65$

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
9	9.65	-0.65	0.4225
7	9.65	-2.65	7.0225
10	9.65	0.35	0.1225
12	9.65	2.35	5.5225
10	9.65	0.35	0.1225
11	9.65	1.35	1.8225
8	9.65	-1.65	2.7225
10	9.65	0.35	0.1225
10	9.65	0.35	0.1225
17	9.65	7.35	54.0225
11	9.65	1.35	1.8225
5	9.65	-4.65	21.6225
9	9.65	-0.65	0.4225
10	9.65	1.35	1.8225
4	9.65	-5.65	31.9225
12	9.65	2.35	5.5225
6	9.65	-3.65	13.3225
14	9.65	4.35	18.9225
8	9.65	-1.65	2.7255
10	9.65	0.35	0.1225

$\frac{167.55}{20} \approx 8.38$

$\sqrt{8.38} \approx 2.9$

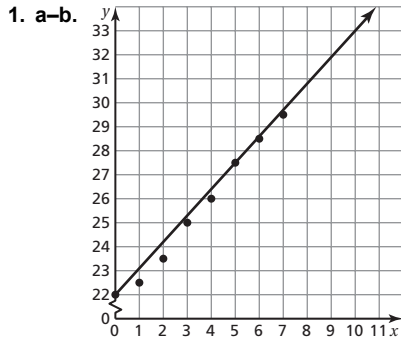
The mean is 9.65 minutes and the standard deviation is about 2.9 minutes.

c. Because the distribution is symmetric, most of the data falls between 9 and 11 minutes. So, most of the presentations are about as long a they are supposed to be.

Chapter 12

Section 12.5

12.5 Activity (pp. 636–637)



c. *Sample answer:* The line passes through (3, 25) and (6, 28.5).

$$\text{slope} = \frac{\text{rise}}{\text{run}} = \frac{3.5}{3} \approx 1.1$$

Because the line crosses the y -axis at about (0, 22), the y -intercept is 22.

So, the equation of a line of best fit is $y = 1.1x + 22$.

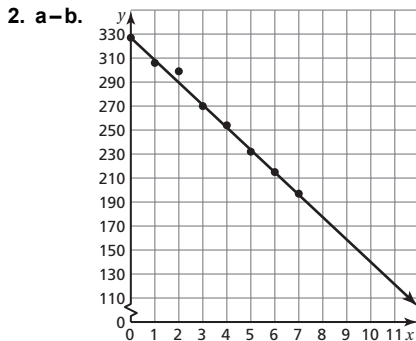
d. *Sample answer:* To predict the length next September, substitute 12 for x in the equation of the line of best fit.

$$y = 1.1x + 22$$

$$y = 1.1(12) + 22$$

$$y = 35.2$$

The length of the baby alligator next September should be about 35.2 inches.



c. *Sample answer:* The line passes through (4, 254) and (6, 215).

$$\text{slope} = \frac{\text{rise}}{\text{run}} = \frac{-39}{2} = -19.5$$

Because the line crosses the y -axis at (0, 327), the y -intercept is 327.

So, the equation of the line of best fit is $y = -19.5x + 327$.

d. *Sample answer:* In 3 years, it will be year 10.

$$\text{When } x = 10, y = -19.5(10) + 327 = 132.$$

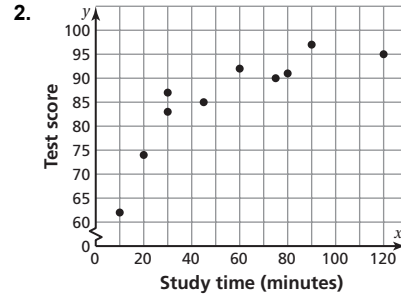
The number of bats in the mine in 3 years will be about 132,000.

3. You can plot data you know and then draw a line that best approximates the points. Then, you can write an equation and use it to predict an event.

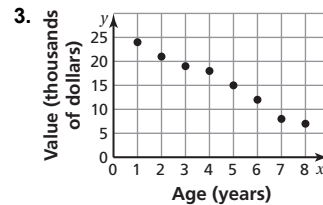
4. *Answer should include, but is not limited to:* Make sure that the data can be described by a line.

12.5 On Your Own (pp. 638–640)

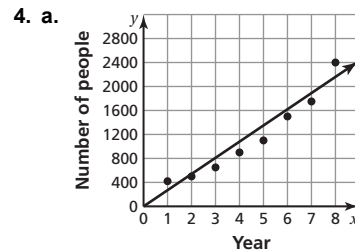
1. *Sample answer:* Draw a vertical line from the point that has a y -value just below 650. It crosses the x -axis at 33. So, a sandwich with 650 calories might have about 35 grams of fat.



As the study time increases, the test score increases. So, the scatter plot shows a positive relationship.



As the age of a car increases, the value decreases. So, the scatter plot shows a negative relationship.



b. *Sample answer:* The line passes through (0, 0) and (5, 1350).

$$\text{slope} = \frac{\text{rise}}{\text{run}} = \frac{1350}{5} = 270$$

Because the line crosses the y -axis at (0, 0), the y -intercept is 0.

So, the equation of a line of best fit is $y = 270x$.

c. The slope of the line of fit is 270. This means that the attendance is increasing by about 270 people each year.

Chapter 12

- d. *Sample answer:* To predict the number of people who will attend in year 10, substitute 10 for x in the equation of the line of best fit.

$$y = 270x$$

$$y = 270(10)$$

$$y = 2700$$

The number of people who will attend in year 10 should be about 2700.

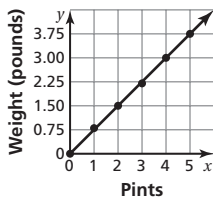
12.5 Exercises (pp. 641–643)

Vocabulary and Concept Check

- A set of coordinate points, or discrete data, is needed to create a scatter plot.
- Because a line of best fit is drawn close to most of the data points, the equation of the line can be used to estimate and predict values.

Practice and Problem Solving

3. a–b.



- c. *Sample answer:* The line passes through $(0, 0)$ and $(4, 3.0)$.

$$\text{slope} = \frac{\text{rise}}{\text{run}} = \frac{3.0}{4} = 0.75$$

Because the line crosses the y -axis at $(0, 0)$, the y -intercept is 0. So, the equation of a line of best fit is $y = 0.75x$.

- d. *Sample answer:* To predict the weight of 10 pints of blueberries, substitute 10 for x in the equation of the line of best fit.

$$y = 0.75x = 0.75(10) = 7.5$$

The weight of 10 pints of blueberries should be about 7.5 pounds.

- e. *Sample answer:* Because 10 pints of blueberries weigh 7.5 pounds, the cost of 10 pints of blueberries is $\$2.25 \cdot 7.5 \approx \16.88 .

- a. Draw a vertical line from the point that has a y -value of 1000. It crosses the x -axis at 2007. So, 1000 SUVs were sold in 2007.
- b. Draw a horizontal line from the point that has an x -value of 2009. It crosses the y -axis at about 875. So, in 2009, about 875 SUVs were sold.
- c. As the year increases, the number of SUVs sold decreases. So, the scatter plot shows a negative relationship.

- a. The point at $x = 3$ has a y -value below 70 and the point at $x = 4$ has a y -value above 70, so a value between 3 and 4 will yield a y -value of 70. You can estimate that the server should work about 3.5 hours to earn \$70.
- b. Draw a horizontal line from the point that has an x -value of 5. It crosses the y -axis at about 85. So, the server earned about \$85 for 5 hours of work.
- c. As the hours worked increases, the earnings increase. So, the scatter plot shows a positive relationship.
6. As x increases, y decreases. So, the scatter plot shows a negative relationship.
7. As x increases, y increases. So, the scatter plot shows a positive relationship.
8. The points show no pattern. So, the scatter plot shows no relationship.
9. As the year increases, the price per pound for honey increases. So, the data show a positive relationship.

10. *Answers will vary. Sample answer:*

As the number of weeks a movie has been in theaters increases, the weekly movie ticket revenue decreases.

11. *Answers will vary. Sample answer:*

This line is not a good representation of the data because too many points lie below the line.

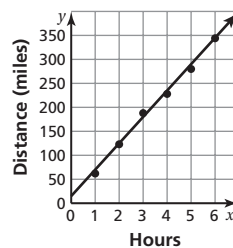
12. *Answers will vary. Sample answer:*

This line is not a good representation of the data because the line does not indicate the overall relationship of the data set.

13. *Answers will vary. Sample answer:*

This line is a good representation of the data because there are about the same number of points above the line as there are below the line. Also, the line indicates the overall positive relationship of the data set.

14. a–b.



The scatter plot shows a positive relationship.

- b. The line passes through $(2, 123)$ and $(6, 344)$.

$$\text{slope} = \frac{\text{rise}}{\text{run}} = \frac{221}{4} \approx 55$$

Because the line crosses the y -axis at about $(0, 15)$, the y -intercept is 15. So, the equation of a line of best fit is $y = 55x + 15$.

- c. The slope of the line of fit is 55. This means the distance you have traveled increases 55 miles each hour.

Chapter 12

- d. To predict the distance traveled in 7 hours, substitute 7 for x in the equation of the line of best fit.

$$y = 55x + 15 = 55(7) + 15 = 400$$

The distance traveled in 7 hours should be about 400 miles.

15. a. As the study time increases, the test scores increase. So, the scatter plot shows a positive relationship.
b. The more time you spend studying, the higher the test score you will receive.

16. No; Because there is no line that can be drawn close to most of the data points if the data set has no relationship, there is no line of best fit for the data.

17. *Answer should include, but is not limited to:* The student will find the height and arm span of three people, plot the data and draw the line of best fit, and predict his or her arm span using the line of best fit. The student will measure his or her arm span and compare the result with the prediction. The slope of the line of best fit should be approximately equal to 1.

18. In a scatter plot, an outlier is a data point that is far removed from the other data points in the data set.

19. *Answers will vary. Sample answer:*

The points follow a U-shaped pattern. There is a non-linear relationship between the variables.

20. *Answers will vary. Sample answer:*

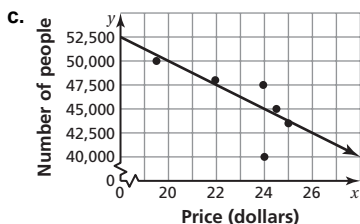
There are two clusters of data points separated by a gap. There seems to be a negative relationship between the variables.

21. *Answers will vary. Sample answer:*

There appear to be two outliers separated from the rest of the data. There seems to be a positive relationship between the variables.

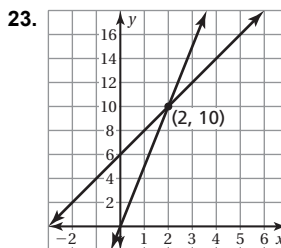
22. a. Because the yearly attendance when the price of admission was \$24.00 does not follow the same linear pattern as the other data, it is an outlier.

- b. Because the outlier is below the other values, it will increase the steepness of the line of best fit.

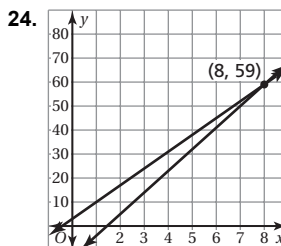


- d. Draw a horizontal line from the point that has an x -value of 27. It crosses the y -axis at about 41,000. So, the yearly attendance should be around 41,000 people.

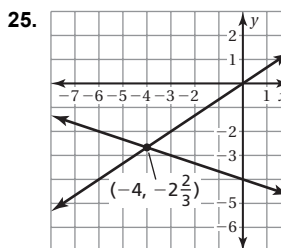
Fair Game Review



Because the lines cross at the point $(2, 10)$, the solution of the equation is $x = 2$.



Because the lines cross at the point $(8, 59)$, the solution of the equation is $x = 8$.



Because the lines cross at the point $(-4, -2\frac{2}{3})$, the solution of the equation is $x = -4$.

26. B; The percent of students that want to fly or want invisibility is $40\% + 22.5\% = 62.5\%$. Because $100\% - 62.5\% = 37.5\%$, the percent of students that want speed or strength is 37.5%.

$$2x + x = 37.5$$

$$3x = 37.5$$

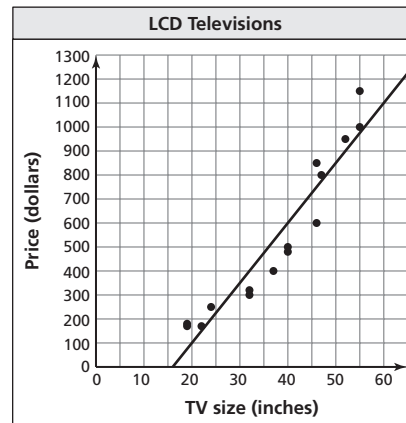
$$x = 12.5$$

So, 12.5% of students want strength as their super power.

Section 12.6

12.6 Activity (pp. 644–645)

1. a–b. *Answers will vary. Sample answer:*



As the TV size increases, the price increases. So, the scatter plot shows a positive relationship.

Chapter 12

c. *Answers will vary. Sample answer:*

An equation for the line of fit is $y = 25x - 400$.

d. *Answer should include, but is not limited to:* The students should compare their lines of fit. They should discuss the similarities and differences between the lines. They should determine which line they feel is a better fit and explain why. Reasons may include equal distribution of points above and below the line, the line of fit going through multiple points, or how close the line is to the points.

2. *Answer should include but is not limited to:* The student will compare his or her line of fit to those of the other students in the class. The student should determine any similarities or differences between the lines. Based on this, the student should choose a line of best fit and describe how he or she came to this conclusion. Criteria when choosing may include equal distribution of points below and above the line of fit, the line of fit passing through multiple points, or the line of fit being closer to the points.

3. c. *Answer should include, but is not limited to:* The student should compare his or her results from Activities 1 and 2 to the results from part (b). The lines of fit should be similar.

4. *Answers will vary. Sample answer:*

TV Size (in.), x	26	42	50	60
Price Using Line of Fit from Activity 2: $y = 25x - 400$	\$250.00	\$650.00	\$850.00	\$1100.00
Price Using Line of Best Fit: $y = 24.9x - 400$	\$247.40	\$645.80	\$845.00	\$1094.00
Difference Between the Prices	\$2.60	\$4.20	\$5.00	\$6.00

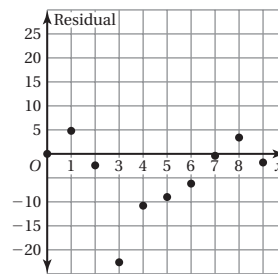
The predicted prices are only off by a few dollars, but the error in prediction increases as the TV size increases.

5. One method to find a line that best models a data set is to compare lines of fit to find the one that seems to fit best. Another method is to use a graphing calculator to find the line of best fit. The line of best fit is the most accurate line of fit.

12.6 On Your Own (pp. 647–648)

1.

x	y	y -value from Model	Residual
0	850	850	$850 - 850 = 0$
1	845	840.2	$845 - 840.2 = 4.8$
2	828	830.4	$828 - 830.4 = 2.4$
3	798	820.6	$798 - 820.6 = -22.6$
4	800	810.8	$800 - 810.8 = -10.8$
5	792	801	$792 - 801 = -9$
6	785	791.2	$785 - 791.2 = -6.2$
7	781	781.4	$781 - 781.4 = -0.4$
8	775	771.6	$775 - 771.6 = 3.4$
9	760	761.8	$760 - 761.8 = -1.8$



The points are randomly dispersed about the horizontal axis. So, the model $y = -9.8x + 850$ is a good fit.

2. Enter the data from the table into your calculator.

L1	L2	L3
0	850	
1	845	
2	828	
3	798	
4	800	
5	792	
6	785	
7	781	
8	775	
9	760	

Use the *linear regression* feature.

LinReg
$y = ax + b$
$a = -9.587878788$
$b = 844.5454545$
$r^2 = .9284574851$
$r = -.9635649875$

An equation of the line of best fit is $y = -9.6x + 845$.

The correlation coefficient is about -0.964 . This means that the relationship between the years and attendance is a strong negative relationship and the equation closely models the data.

3. There may be a negative correlation. A causal relationship is possible because the more time that is spent playing video games, the less time there is to study.

Chapter 12

12.6 Exercises (pp. 649–651)

Vocabulary and Concept Check

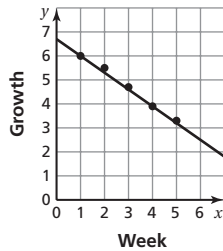
- The residual is positive when the actual y -value is greater than the value from the model.

The residual is negative when the actual y -value is less than the value from the model.

- When a model is a good fit, a scatter plot of the residual points will be randomly dispersed about the horizontal axis. If the model is not a good fit, a scatter plot of the residual points will form some type of pattern.
- A correlation coefficient of -0.98 indicates a stronger relationship because it is closer to -1 than 0.91 is to 1 .

Practice and Problem Solving

- Answers will vary. Sample answer:



An equation for the line of fit that passes through the points $(1, 6.0)$ and $(4, 3.9)$ is $y = -0.7x + 6.7$.

- Use the *linear regression* feature of a graphing calculator to find the line of best fit. The line of best fit is $y = -0.7x + 6.8$.

- Answers will vary. Sample answer:

The growth using the result in part (a):

$$\begin{aligned} y &= -0.7x + 6.7 \\ &= -0.7(6) + 6.7 \\ &= -4.2 + 6.7 \\ &= 2.5 \end{aligned}$$

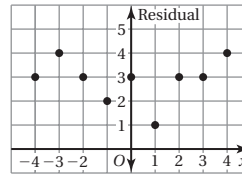
The growth is 2.5 inches.

The growth using the line of best fit:

$$\begin{aligned} y &= -0.7x + 6.8 \\ &= -0.7(6) + 6.8 \\ &= -4.2 + 6.8 \\ &= 2.6 \end{aligned}$$

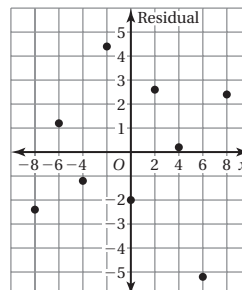
The growth is 2.6 inches.

x	y	y -value from Model	Residual
-4	-18	-21	$-18 - (-21) = 3$
-3	-13	-17	$-13 - (-17) = 4$
-2	-10	-13	$-10 - (-13) = 3$
-1	-7	-9	$-7 - (-9) = 2$
0	-2	-5	$-2 - (-5) = 3$
1	0	-1	$0 - (-1) = 1$
2	6	3	$6 - 3 = 3$
3	10	7	$10 - 7 = 3$
4	15	11	$15 - 11 = 4$



The residual points are all above the x -axis, so the equation does not model the data well.

x	y	y -value from Model	Residual
-8	9	11.4	$9 - 11.4 = -2.4$
-6	10	8.8	$10 - 8.8 = 1.2$
-4	5	6.2	$5 - 6.2 = -1.2$
-2	8	3.6	$8 - 3.6 = 4.4$
0	-1	1	$-1 - 1 = -2$
2	1	-1.6	$1 - (-1.6) = 2.6$
4	-4	-4.2	$-4 - (-4.2) = 0.2$
6	-12	-6.8	$-12 - (-6.8) = -5.2$
8	-7	-9.4	$-7 - (-9.4) = 2.4$



The residual points are randomly dispersed over the horizontal axis so the equation is a good fit.

Chapter 12

7. Use the *linear regression* feature of a graphing calculator.

```
LinReg
y=ax+b
a=0
b=3.5
r2=0
r=0
```

An equation of the line of best fit is $y = 3.5$. The correlation coefficient is 0. This means there is no correlation between x and y and the equation does not fit the data.

8. Use the *linear regression* feature of a graphing calculator.

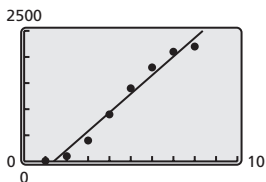
```
LinReg
y=ax+b
a=-1.2182
b=7.0182
r2=0.7797
r=-0.883
```

An equation of the line of best fit is $y = -1.2x + 7$. The correlation coefficient is about -0.883 . Because -0.883 is fairly close to -1 , there is a fairly strong negative correlation and the equation fits the data relatively well.

9. Use the *linear regression* feature of a graphing calculator.

```
LinReg
y=ax+b
a=357.5
b=-495
r2=.9748
r=.9873
```

An equation of the line of best fit is $y = 357.5x - 495$.



10. B; The variables have a strong positive relationship. So, r is close to 1.
11. A; The variables have a strong negative relationship. So, r is close to -1 .
12. C; The variables have little to no correlation. So, r is close to 0.
13. There is a negative correlation and causal relationship. The more time you spend talking on the phone, the more battery power you use, so there is less battery life left.
14. There is a positive correlation but not a causal relationship. The toddler's height will increase with age. The toddler's vocabulary will also increase with age. So, although both variables are increasing, the increase in one does not directly cause the increase in the other.

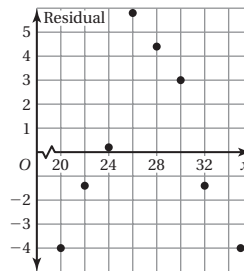
15. There is no likely correlation between the variables.
16. There is no likely correlation between the variables.
17. a. Use the *linear regression* feature of a graphing calculator.

```
LinReg
y=ax+b
a=0.2146
b=24.9282
r2=.0801
r=0.283
```

An equation for the line of best fit is $y = 0.2x + 25$. The correlation coefficient is about 0.283. Because 0.283 is relatively close to 0, there is a weak positive relationship between the price and the fuel economy of the automobiles, and the equation does not fit the data well.

b.

x	y	y -value from Model	Residual
24	30	29.8	$30 - 29.8 = 0.2$
32	30	31.4	$30 - 31.4 = -1.4$
30	34	31	$34 - 31 = 3$
28	35	30.6	$35 - 30.6 = 4.4$
35	28	32	$28 - 32 = -4$
20	25	29	$25 - 29 = -4$
22	28	29.4	$28 - 29.4 = -1.4$
26	36	30.2	$36 - 30.2 = 5.8$



The residuals form a \cup -shaped pattern. So, the equation does not model the data well.

18. a. Use the *linear regression* feature of a graphing calculator.

```
LinReg
y=ax+b
a=454.8286
b=-2811.5429
r2=0.9683
r=0.984
```

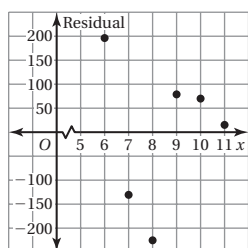
An equation for the line of best fit is $y = 454.8x - 2812$. The correlation coefficient is about 0.984. Because 0.984 is close to 1, there is a strong positive relationship between the year and the number of text messages sent, and the equation closely models the data.

Chapter 12

- b. The slope of the line of best fit is 454.8. This means that the number of text messages sent increases by 454.8 billion each year.

c.

x	y	y -value from Model	Residual
6	113	-83.2	$113 - (-83.2) = 196.2$
7	241	371.6	$241 - 371.6 = -130.6$
8	601	826.4	$601 - 826.4 = -225.1$
9	1360	1281.2	$1360 - 1281.2 = 78.8$
10	1806	1736	$1806 - 1736 = 70$
11	2206	2190.8	$2206 - 2190.8 = 15.2$



The residual points are randomly dispersed over the horizontal axis. So, the equation is a good fit for the data.

d. $y = 454.8x - 2812$
 $= 454.8(15) - 2812$
 $= 4010$

About 4010 billion, or 4 trillion, text messages will be sent in 2015.

19. a. Use the *linear regression* feature of a graphing calculator.

```
LinReg
y=ax+b
a=-0.0835
b=3.8290
r^2=-0.9312
r=-0.965
```

An equation for the line of best fit is $y = -0.08x + 3.8$. The correlation coefficient is about -0.965 . Because -0.965 is close to -1 , there is a strong negative relationship between the hours of TV watched and grade point average, and the equation closely models the data.

- b. The slope of the line of best fit is -0.08 . This means that the students' grade point averages decrease by 0.08 for each hour of television they watch each week. The y -intercept of the line of best fit is 3.8. This means that a student who does not watch any television has a grade point average of about 3.8.

c. $y = -0.08x + 3.8$
 $= -0.08(14) + 3.8$
 $= 2.68$

The student's grade point average is about 2.7.

- d. *Answers will vary. Sample answer:*

No; Watching television does not directly result in a lower grade point average. The cause may be studying less.

20. Use the *linear regression* feature on a graphing calculator. Include the point $(2, 2.4)$ with the data from Exercise 19.

```
LinReg
y=ax+b
a=-0.0557
b=3.4387
r^2=0.4449
r=-0.667
```

The correlation coefficient changed from about -0.965 to about -0.667 . This weakens the correlation. The change occurred because the point $(2, 2.4)$ is an outlier.

21. a.

x	1	2	3	4	5	6	7	8
y	10	90	300	500	500	400	300	100

- b. The y -values increase and then decrease as the x -values increase. The data does not show an approximately constant rate of change. So, a linear model would not fit the data well.

The way in which the y -values increase and then decrease means a quadratic model may fit the data well.

- c. Use the *quadratic regression* feature on a graphing calculator.

```
LinReg
y=a^2+bx+c
a=-35.3571
b=341.7857
c=-361.4286
R^2=.9216316858
```

The data can be modeled by

$$y = -35.36x^2 + 341.8x - 361.$$

Chapter 12

Fair Game Review

22. Mean: $\frac{469}{7} = 67$ The mean is 67.

Range: Ordered data: 58, 59, 62, 68, 70, 75, 77

The least number is 58. The greatest number is 77.

The range is $77 - 58 = 19$.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
59	67	-8	64
70	67	3	9
62	67	-5	25
68	67	1	1
75	67	8	64
77	67	10	100
58	67	9	81

$$\frac{64 + 9 + 25 + 1 + 64 + 100 + 81}{7} \approx 49.14$$

$$\sqrt{49.14} \approx 7.01$$

The standard deviation is about 7.0.

23. Mean: $\frac{135}{9} = 15$ The mean is 15.

Range: Ordered data: 11, 14, 14, 15, 15, 15, 16, 16, 19

The least value is 11. The greatest value is 19.

The range is $19 - 11 = 8$.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
15	15	0	0
14	15	-1	1
11	15	-4	16
15	15	0	0
16	15	1	1
19	15	4	16
14	15	-1	1
16	15	1	1
15	15	0	0

$$\frac{0 + 1 + 16 + 0 + 1 + 16 + 1 + 1}{9} = 4$$

$$\sqrt{4} = 2$$

The standard deviation is 2.

24. C;

interquartile range = third quartile – first quartile

$$= 50 - 35$$

$$= 15$$

The interquartile range is 15.

Section 12.7

12.7 Activity (pp. 652–653)

1. a.

		T-Shirt Size					Total
		S	M	L	XL	XXL	
Color	Blue/White	5	4	1	0	2	12
	Blue/Gold	3	6	5	2	0	16
	Red/White	4	2	4	1	3	14
	Black/White	3	4	1	2	1	11
	Black/Gold	5	2	3	0	2	12
Total		20	18	14	5	8	65

b. No; The entry in the “Black/Gold” row and “XL” column is 0. This means there are no black and gold XL T-shirts left in stock.

c.

		T-Shirt Size					Total
		S	M	L	XL	XXL	
Color	Blue/White	5	6	7	6	5	29
	Blue/Gold	5	6	7	6	5	29
	Red/White	5	6	7	6	5	29
	Black/White	5	6	7	6	5	29
	Black/Gold	5	6	7	6	5	29
Total		25	30	35	30	25	145

d. Answers will vary. Sample answer:

For every category of shirt, order the number of shirts sold last year minus the number remaining. If the result is a negative number, then do not order any shirts in that category. Also, order a couple extra of the shirts that were all sold.

2. a.

		Gender		Total
		Boys	Girls	
Hours	0 hours per week	160	160	320
	1–7 hours per week	130	140	270
	8+ hours per week	90	140	230
	Total	380	440	830

Chapter 12

b. *Answers will vary. Sample answer:*

More girls work part-time jobs than boys at Stevenson High School.

The students that do not work make up the largest portion of the student population.

c. No; The data does not support or oppose the claim because the data is not about students dropping out or working full-time jobs.

3. To read a two-way table, note that each entry represents the number of data values that fit both the criteria for the row and column the entry is in. The entries in the Total row and column represent the number of data values that fit the one criterion in that column or row.

To make a two-way table for a data set, use one variable that describes the data for the rows for the table and another variable for the columns. Put the number of data points that fit both row and column criteria in each slot in the table. On the bottom and right edges of the table, sum up each column and row.

4. *Answer should include, but is not limited to:*

The student will collect data from a real-world example and use the data to make a two-way table.

12.7 On Your Own (pp. 654–655)

1. a. The number in the “Dance Attend” row and “Football Game Not Attend” column is 5.

So, 5 of the students in the survey are attending the dance but not the football game.

b.

		Football Game		
		Attend	Not Attend	Total
Dance	Attend	35	5	40
	Not Attend	16	20	36
Total		51	25	76

40 students are attending the dance.

36 students are not attending the dance.

51 students are attending the football game.

25 students are not attending the football game.

76 students were surveyed.

2. a.

		Grade			Total
		6	7	8	
Lunch	Pack	11	23	16	50
	Buy	9	27	14	50
Total		20	50	30	100

b.

		Grade		
		6	7	8
Lunch	Pack	55%	46%	53%
	Buy	45%	54%	47%

Sample answer:

$$\frac{9}{20} = 0.45$$

So, 45% of the 6th grade students in the survey buy a school lunch.

c. No; About half of the students in each grade buy a school lunch.

12.7 Exercises (pp. 656–657)

Vocabulary and Concept Check

- The joint frequencies are the entries in the two-way table. The marginal frequencies are the sums of the rows and columns of joint frequencies.
- A two-way table displays two categories of data collected from the same source.

Practice and Problem Solving

3.

		Fundraiser		Total
		No	Yes	
Gender	Female	22	51	73
	Male	30	29	59

Total number of females surveyed: 73

Total number of males surveyed: 59

4.

		Fundraiser	
		No	Yes
Gender	Female	22	51
	Male	30	29
Total		52	80

Total number of students not participating: 52

Total number of students participating: 80

- The entry in the “Female” row and “Yes” column is 51. So, 51 females will be participating.
- The joint frequency in the “Male” row and “No” column is 30. So, 30 males will not be participating.

Chapter 12

7.

		School Play		
		Attend	Not Attend	Total
Class	Junior	41	30	71
	Senior	52	23	75
Total		93	53	146

93 students surveyed are attending the school play.
 53 students surveyed are not attending the school play.
 71 students surveyed are juniors.
 75 students surveyed are seniors.
 146 students were surveyed.

8.

		Cell Phone Minutes		
		Limited	Unlimited	Total
Text Plan	Limited	78	0	78
	Unlimited	175	15	190
Total		253	15	268

253 people surveyed have limited cell phone minutes.
 15 people surveyed have unlimited cell phone minutes.
 78 people surveyed have limited text plans.
 190 people surveyed have unlimited text plans.
 268 people were surveyed.

9. a. The joint frequency in the “10th” row and “Sports” column is 19.

So, 19 students in the 10th grade chose sports as being most important to them.

b.

		Goal			Total
		Grades	Popularity	Sports	
Grade	9th	31	18	23	72
	10th	39	16	19	74
	11th	42	6	17	65
	Total	112	40	59	211

112 students surveyed chose grades as most important.
 40 students surveyed chose popularity as most important.
 59 students surveyed chose sports as most important.
 72 students in the 9th grade were surveyed.
 74 students in the 10th grade were surveyed.
 65 students in the 11th grade were surveyed.
 211 students were surveyed.

c. 9th graders who chose popularity: 18

Total students surveyed: 211

$$\frac{18}{211} \approx 0.085$$

About 8.5% of students in the survey are 9th graders who chose popularity.

10.

		Gender		
		Male	Female	Total
Pet	Own	18	6	24
	Do not own	25	14	39
Total		43	20	63

11. a.

		Eye Color			
		Green	Blue	Brown	Total
Gender	Male	5	16	27	48
	Female	3	19	18	40
Total		8	35	45	88

b. 8 students surveyed have green eyes.

35 students surveyed have blue eyes.

45 students surveyed have brown eyes.

48 males were surveyed.

40 females were surveyed.

88 students were surveyed.

c.

		Eye Color		
		Green	Blue	Brown
Gender	Male	63%	46%	60%
	Female	38%	54%	40%

Answers will vary. Sample answer:

3 of the 8 students with green eyes are female.

$$\frac{3}{8}(100\%) = 37.5\%$$

So, the joint frequency in the “Female” row and “Green” column tells us that 37.5% of the students with green eyes are female.

27 of the 45 students with brown eyes are male.

$$\frac{27}{45}(100\%) = 60\%$$

So, the joint frequency in the “Male” row and “Brown” column tells us that 60% of the students with brown eyes are male.

Chapter 12

12.

		Eye Color		
		Green	Blue	Brown
Gender	Male	10.4%	33.3%	56.3%
	Female	7.5%	47.5%	45%

Answers will vary. Sample answer:

5 of the 48 male students have green eyes.

$$\frac{5}{48}(100\%) \approx 10.4\%$$

So, the joint frequency in the “Green” row and “Male” column tells us that 10.4% of the male students have green eyes.

18 of the 40 female students have brown eyes.

$$\frac{18}{40}(100\%) = 45\%$$

So, the joint frequency in the “Brown” row and “Female” column tells us that 45% of the female students have brown eyes.

13. 40 out of 88 students are females.

8 out of 88 students have green eyes.

3 out of 88 students are females and have green eyes.

$$\frac{40}{88} + \frac{8}{88} - \frac{3}{88} = \frac{45}{88} \approx 0.511$$

So, about 51.1% of the students are females or have green eyes.

48 out of 88 students are males.

5 out of 88 students are males and have green eyes.

$$\frac{48}{88} - \frac{5}{88} = \frac{43}{88} \approx 0.489$$

So, about 48.9% of the students are males that do not have green eyes.

$$51.1\% + 48.9\% = 100\%$$

This means that the sum of the percentage of students who are female or have green eyes and the percentage of students who are male and do not have green eyes accounts for everyone who was surveyed.

Fair Game Review

14. $\frac{1-x}{x-1} = \frac{-(x-1)}{(x-1)} = -1$

The excluded value is $x = 1$.

15. $\frac{x^2 - x - 6}{x^2 + x - 12} = \frac{(x-3)(x+2)}{(x-3)(x+4)} = \frac{x+2}{x+4}$

The excluded values are $x = 3$ and $x = -4$.

16. $\frac{15x^3 - 6x^2}{21x^3 + 3x^2} = \frac{3x^2(5x - 2)}{3x^2(7x + 1)} = \frac{5x - 2}{7x + 1}$

The excluded values are $x = 0$ and $x = -\frac{1}{7}$.

17. B;

$$\frac{1}{x} = \frac{3}{x+4}$$

$$x + 4 = 3x$$

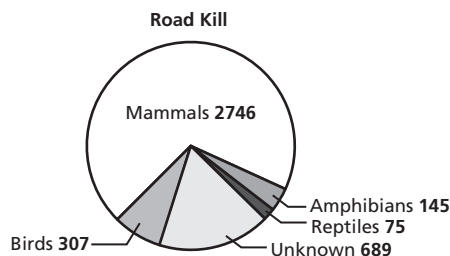
$$4 = 2x$$

$$2 = x$$

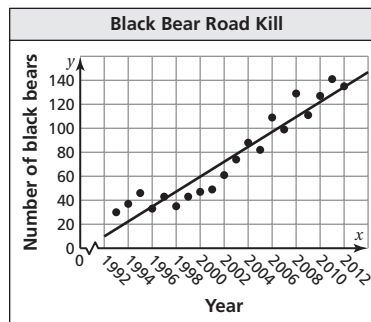
Section 12.8

12.8 Activity (pp. 658–659)

1. a. *Sample answer:* The data set represents parts of a whole, so a circle graph is an appropriate data display.



b. *Sample answer:* The data values change over time, so a line graph is an appropriate data display.



Chapter 12

c. *Sample answer:* A stem-and-leaf plot shows how the raccoon weights are distributed.

Raccoon Road Kill Weights

Stem	Leaf
9	4 5
10	
11	0
12	4 9
13	4 6 9
14	0 5 8 8
15	2 7
16	8
17	0 2 3 5
18	5 5 6 7
19	0 1 4
20	4
21	3 5 5 5
22	
23	
24	
25	4

Key: 9 | 4 = 9.4 pounds

d. *Sample answer:* Reduce the speed limit on some roads.

- Answer should include, but is not limited to:* The student will write a report about an animal species that is (or has been) endangered. See the example on p. 659. Make sure the students' graphical displays are appropriate.
- Sample answer:* A bar graph can be used to compare data in different categories, a circle graph can be used to view data that is part of a whole, a scatter plot shows the relationship between sets of data, a stem-and-leaf plot orders numerical data, and a box-and-whisker plot shows the variability of data.

Answer should include, but is not limited to: The student will find examples of each type of graph: bar graph, circle graph, scatter plot, stem-and-leaf plot, and box-and-whisker plot.

12.8 On Your Own (pp. 660–661)

- Sample answer:* A histogram shows frequencies of data values in intervals of the same size. So, a histogram is an appropriate data display.
- Sample answer:* You want to show percents of students in specific categories. So, a bar graph is an appropriate display.
- The tickets vary in width and the break in the vertical axis makes the difference in ticket prices appear to be greater.
- The bars become wider as the years progress, making the increase in profit appear larger.

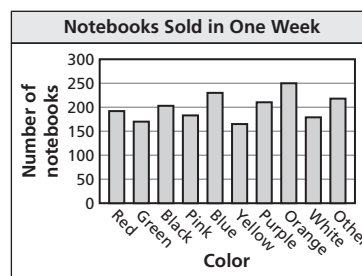
12.8 Exercises (pp. 662–663)

Vocabulary and Concept Check

- Yes; In many cases, more than one display could be appropriate, depending on what aspect of the data needs to be emphasized. Pictograms, bar graphs, line graphs, and circle graphs display and compare data values. Histograms, stem-and-leaf plots, box-and-whisker plots, and line plots show the distribution of data values.
- Sample answer:* The pictures of the objects or bars could be made larger so there appears to be more of them, or the vertical axis could be too small or too large.

Practice and Problem Solving

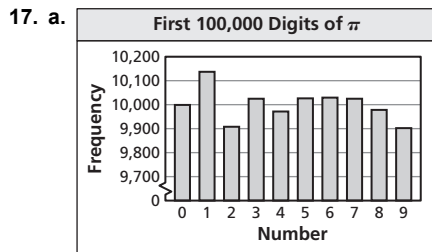
- Sample answer:* Because the data are grouped into specific categories, a bar graph is an appropriate display for the data.



- Sample answer:* A stem-and-leaf plot shows the distribution of data. So, a stem-and-leaf plot is an appropriate display.
- Sample answer:* A line graph shows changes over time. So, a line graph is an appropriate data display.
- Sample answer:* A line plot shows the number of times each outcome occurs. So, a line plot is an appropriate data display.
- Sample answer:* A line graph shows changes over time. So, a line graph is an appropriate data display.
- When the data are grouped into intervals of one category, as opposed to multiple categories, a histogram would be used instead of a bar graph.
- Each bicycle represents the same distance. Because the icons for Monday are the largest and Wednesday are the smallest, it looks like the distance is the same each day, but the distance on Wednesday is four times the distance traveled on Monday.
- The vertical axis on the left has a break and begins at 99.5. This makes it appear that the sales increased rapidly from April to July, but the difference is only $\$104 - \$100 = \$4$ million.
- The intervals of the histogram are not the same size, making it difficult to compare the data.
- The widths of the bars are unequal. This makes it appear that some months had more rainfall than they actually did.
- Sample answer:* Bar graph; Each bar can represent a different vegetable.

Chapter 12

14. Yes; The vertical axis of the line graph does not have a constant interval. The display looks like the data decreased constantly over 9 hours, but the data decreased more rapidly during the first hour.
15. *Sample answer:* Because a line plot shows the number of times each value occurs, the mode would be easily displayed. So, a line plot is an appropriate display.
16. a. Because the sum of the percentages is 110 and not 100, a circle graph is not appropriate because the data are not parts of a whole.
 b. *Sample answer:* A bar graph shows data in specific categories. So, a bar graph would be more appropriate.



b. Find the percent of each number.

$$0: \frac{9999}{100,000} \times 100 \approx 10.00\%$$

$$1: \frac{10,137}{100,000} \times 100 \approx 10.14\%$$

$$2: \frac{9908}{100,000} \times 100 \approx 9.91\%$$

$$3: \frac{10,025}{100,000} \times 100 \approx 10.03\%$$

$$4: \frac{9971}{100,000} \times 100 \approx 9.97\%$$

$$5: \frac{10,026}{100,000} \times 100 \approx 10.03\%$$

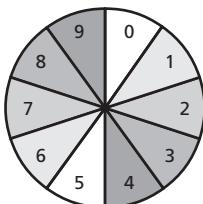
$$6: \frac{10,029}{100,000} \times 100 \approx 10.03\%$$

$$7: \frac{10,025}{100,000} \times 100 \approx 10.03\%$$

$$8: \frac{9978}{100,000} \times 100 \approx 9.98\%$$

$$9: \frac{9902}{100,000} \times 100 \approx 9.90\%$$

First 100,000 Digits of π



- c. Because the frequencies are all very close, the circle graph does not compare the data well. The bar graph is a more appropriate data display because it shows that some digits occur slightly more than others.
- d. Each digit occurs about 10% of the time.

Fair Game Review

18. $x + 3 = 5$

19. $8x = 24$

20. A; 20% of 25% of 400 = $0.20(0.25 \cdot 400)$
 $= 0.20(100)$
 $= 20$

So, 20 is 20% of 25% of 400.

Quiz 12.5–12.8 (p. 664)

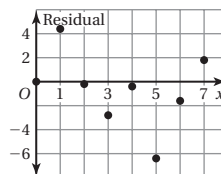
1. a. Draw a vertical line from the point that has a y -value of 150. It crosses the x -axis at 2005. So, the charity received \$150,000 in 2005.
 b. Draw a horizontal line from the point that has an x -value of 2008. It crosses the y -axis at 120. So, the charity received \$120,000 in 2008.
 c. As the year increases, the amount of money donated decreases. So, the scatter plot shows a negative relationship.
2. Use the linear regression feature of a calculator.

```
LinReg
y=ax+b
a=1.6190
b=12.0833
r2=0.8281
r=0.91
```

An equation for the line of best fit is $y = 1.6x + 12$.

The correlation coefficient is about 0.908. Because 0.908 is close to 1, there is a strong positive relationship between x and y , and the equation fits the data well.

x	y	y -value From Model	Residual
0	12	12	$12 - 12 = 0$
1	16	13.6	$16 - 13.6 = 4.4$
2	15	15.2	$15 - 15.2 = -0.2$
3	14	16.8	$14 - 16.8 = -2.8$
4	18	18.4	$18 - 18.4 = -0.4$
5	22	26.4	$20 - 26.4 = -6.4$
6	20	21.6	$20 - 21.6 = -1.6$
7	25	23.2	$25 - 23.2 = 1.8$



The residual points are randomly dispersed over the horizontal axis. So, the model is a good fit.

Chapter 12

3.

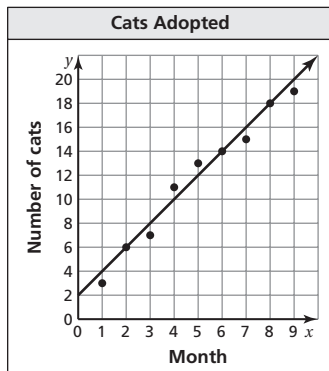
		Recycle		
		Yes	No	Total
Gender	Male	28	9	37
	Female	24	14	38
Total		52	23	75

52 of the people surveyed recycle.
 23 of the people surveyed do not recycle.
 37 of the people surveyed are male.
 38 of the people surveyed are female.
 75 people were surveyed.

4. *Sample answer:* A circle graph shows data as parts of a whole. Because percentages are parts of a whole, a circle graph is an appropriate data display.

5. *Sample answer:* A line graph compares data over time. So, a line graph is an appropriate data display.

6. a.



b. *Sample answer:* The line passes through (2, 6) and (8, 18).

$$\text{slope} = \frac{\text{rise}}{\text{run}} = \frac{12}{6} = 2$$

Because the line crosses at (0, 2), the y -intercept is 2.
 So, the equation for a line of best fit is $y = 2x + 2$.

c. The slope of the line of fit is 2. This means that the number of cats adopted increases by 2 each month.

d. *Sample answer:* To predict the number of cats adopted in month 10, substitute 10 for x in the equation of the line of best fit.

$$y = 2x + 2 = 2(10) + 2 = 22$$

So, about 22 cats will be adopted in month 10.

7. Yes; The vertical axis of the line graph has a break and begins at 500. This graph makes it appear that the amount of money raised increased more rapidly than it actually did.

Chapter 12 Review (pp. 665–669)

1. **Mean:** $\frac{44}{11} = 4$

Median: 2.0, 3.5, 3.9, 4.0, 4.0, **4.1**, 4.3, 4.3, 4.4, 4.5, 5.0
 The middle value is 4.1.

Mode: 2.0, 3.5, 3.9, **4.0, 4.0**, 4.1, **4.3, 4.3**, 4.4, 4.5, 5.0
 The values 4.0 and 4.3 occur most often.

By adding another value, the mean stays the same, the median decreases by $4.2 - 4.1 = 0.1$ kilometer, and the modes are 4.0 kilometers and 4.3 kilometers.

2. **Mean:** $\frac{17}{10} = 1.7$

Median: 0, 0, 1, 1, **1, 1**, 2, 2, 4, 5

The mean of the two middle values is $\frac{1+1}{2} = \frac{2}{2} = 1$.

Mode: The value 1 occurs most often.

The mean is 1.7 goals, the median is 1 goal, and the mode is 1 goal.

3. **Mean:** $\frac{36}{9} = 4$

Median: -9, -2, 0, 3, **3**, 10, 10, 10, 11

The middle value is 3.

Mode: -9, -2, 0, 3, 3, **10, 10, 10**, 11

The value 10 occurs most often.

The mean is 4°F, the median is 3°F, and the mode is 10°F.

4. Store A:

Mean: $\frac{1510}{8} = 188.75$ The mean is \$188.75.

Range: Ordered data: 130, 150, 160, 180, 190, 200, 250, 250.

The least value is 130. The greatest value is 250.

The range is $250 - 130 = \$120$.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
130	188.75	-58.75	3451.56
180	188.75	-8.75	76.56
200	188.75	11.25	126.56
250	188.75	61.25	3751.56
150	188.75	-38.75	1501.56
190	188.75	1.25	1.56
250	188.75	61.25	3751.56
160	188.75	71.25	826.56

$$\frac{13487.48}{8} = 1685.935$$

$$\sqrt{1685.935} \approx 41.1$$

The standard deviation is about 41.1.

Chapter 12

Store B:

Mean: $\frac{1910}{8} = 238.75$ The mean is \$238.75.

Range: Ordered data: 190, 200, 210, 225, 230, 260, 285, 310

The least value is 190. The greatest value is 310.

The range is $310 - 190 = \$120$.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
225	238.75	-13.75	189.06
310	238.75	71.25	5076.56
260	238.75	21.25	451.56
190	238.75	-48.75	2376.56
200	238.75	-38.75	1501.56
285	238.75	46.25	2139.06
210	238.75	-28.75	826.56
230	238.75	-8.75	76.56

$$\frac{12637.48}{8} = 1579.685$$

$$\sqrt{1579.685} \approx 39.7$$

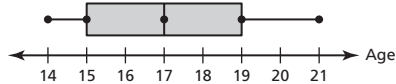
The standard deviation is about 39.7.

The mean at Store B is greater, the ranges are the same, and the standard deviations are about the same.

5. Order the data. Find the median and the quartiles.

14 14 16 17 17 18 20 21
 ↑ ↑ ↑ ↑ ↑
 least first median, third greatest
 value quartile, 17 quartile, value
 15 19

Draw a box-and-whisker plot.

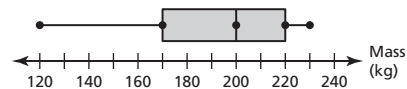


The right whisker is longer than the left whisker and most of the data are on the left. So, the box-and-whisker plot is skewed right.

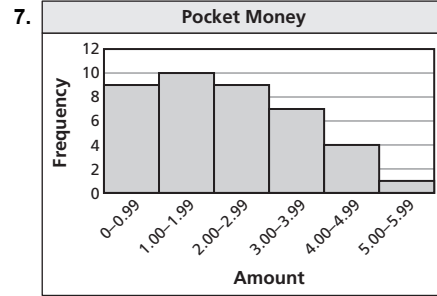
6. Order the data. Find the median and the quartiles.

120 160 180 200 200 210 230 230
 ↑ ↑ ↑ ↑ ↑
 least first median, third greatest
 value quartile, 200 quartile, value
 120 170 220 230

Draw a box-and-whisker plot.

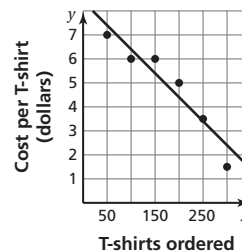


The left whisker is longer than the right whisker and most of the data is on the right. So, the distribution is skewed left.



The “tail” of the graph extends to the right and most of the data is on the left. So, the distribution is skewed right.

8. The distribution is skewed right. So, the center of the data is best represented by the median and the dispersion of the data is best represented by the five-number summary.
9. Answers will vary. Sample answer:



An equation for a line of fit is $y = -0.02x + 8.4$.

$$y = -0.02x + 8.4 = -0.02(275) + 8.4 = 2.9$$

The cost per T-shirt would be \$2.90.

10. Use the *linear regression* feature on a graphing calculator. Include the points (64, 7), (65, 9), and (71, 11) with the data from the example.

```
LinReg
y=ax+b
a=0.4881
b=-23.2071
r^2=0.9101
r=0.954
```

An equation for the line of best fit is $y = 0.49x - 23.2$.

The correlation coefficient decreased from about 0.971 to about 0.954. This is only a slight decrease and there is still a strong positive relationship between the variables.

- 11.

		Likes the New Food Court		Total
		Yes	No	
Age Group	Adult	21	79	100
	Teenager	96	4	100
Total		117	83	200

79 out of 100 adults surveyed disliked the new food court.

$$\frac{79}{100}(100\%) = 79\%$$

So, 79% of the adults surveyed disliked the new food court.

Chapter 12

12. *Sample answer:* A pictograph shows data using pictures. So, a pictograph is an appropriate display.

Chapter 12 Test (p. 670)

1. **Mean:** $\frac{9}{9} = 1$

Median: -6, -3, -3, -3, 0, 0, 3, 10, 11

The middle value is 0.

Mode: -6, -3, -3, -3, 0, 0, 3, 10, 11

The value -3 occurs most often.

The mean is 1 foot, the median is 0 feet, and the mode is -3 feet.

2. **Mean:** $\frac{588}{12} = 49$

Median: 35, 38, 40, 41, 48, 50, 54, 54, 54, 55, 59, 60

The mean of the two middle values is $\frac{50 + 54}{2} = 52$.

Mode: 35, 38, 40, 41, 48, 50, 54, 54, 54, 55, 59, 60

The value 54 occurs most often.

The mean is 49 minutes, the median is 52 minutes, and the mode is 54 minutes.

3. Pond A:

Mean: $\frac{80}{8} = 10$ The mean is 10 pounds.

Range: Ordered data: 6, 7, 7, 8, 12, 12, 13, 15

The least value is 6. The greatest value is 15.

The range is $15 - 6 = 9$ pounds.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
12	10	2	4
13	10	3	9
15	10	5	25
6	10	-4	16
7	10	-3	9
8	10	-2	4
12	10	2	4
7	10	-3	9

$\frac{80}{8} = 10$

$\sqrt{10} \approx 3.2$

The standard deviation is about 3.2.

Pond B:

Mean: $\frac{96}{8} = 12$ The mean is 12 pounds.

Range: Ordered data: 5, 8, 9, 12, 12, 15, 16, 19

The least value is 5. The greatest value is 19.

The range is $19 - 5 = 14$ pounds.

Standard deviation:

x	\bar{x}	$x - \bar{x}$	$(x - \bar{x})^2$
9	12	-3	9
12	12	0	0
5	12	-7	49
8	12	-4	16
12	12	0	0
15	12	3	9
16	12	4	16
19	12	7	49

$\frac{148}{8} = 18.5$

$\sqrt{18.5} \approx 4.3$

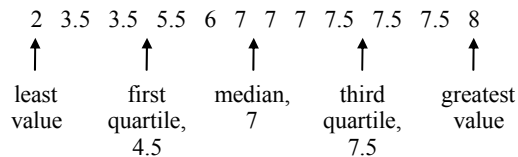
The standard deviation is about 4.3.

The mean weight of the turtles in Pond B is greater than the mean weight of the turtles in Pond A. The range and standard deviation of Pond B are also greater than Pond A. So, the turtles in Pond B tend to be heavier than the turtles in Pond A, and the weights of the turtles in Pond B are more spread out.

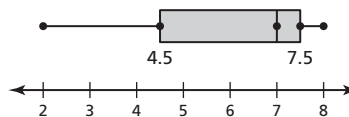
4. *Answers will vary. Sample answer:*

A box-and-whisker plot shows the spread of each data set. So, a box-and-whisker plot is an appropriate display.

5. a. Order the data. Find the median and quartiles.



Draw a box-and-whisker plot.



Chapter 12

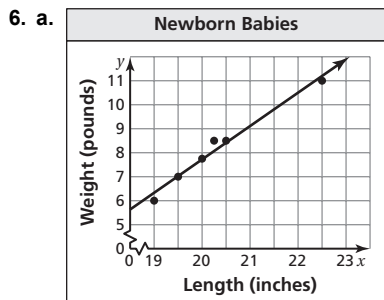
- b. The least value is 2. The greatest value is 8.

So, the range is $8 - 2$, or 6 hours.

$$\begin{aligned} \text{Interquartile range} &= \text{third quartile} - \text{first quartile} \\ &= 7.5 - 4.5 \\ &= 3 \end{aligned}$$

So, the interquartile range is 3.

- c. The distribution is skewed left. So, the center of the data is best represented by the median and the dispersion of the data is best represented by the five-number summary.



Sample answer: The line passes through (19.5, 7) and (20.5, 8.5).

$$\text{slope} = \frac{\text{rise}}{\text{run}} = \frac{1.5}{1} = 1.5$$

Using this slope, you can determine that the line crosses the y -axis at $(0, -22.25)$, and the y -intercept is -22.25 .

So, the equation of a line of best fit is

$$y = 1.5x - 22.25.$$

- b. Use the *linear regression* feature of a graphing calculator.

```
LinReg
y=ax+b
a=1.3909
b=-20.0981
r^2=0.9761
r=0.988
```

An equation for the line of best fit is

$y = 1.39x - 20.1$. The correlation coefficient is about 0.988. Because 0.988 is close to 1, there is a strong positive relationship between the length and weight of a newborn, and the equation closely models the data.

- c. $x = 21$ inches

$$y = 1.5x - 22.25 = 1.5(21) - 22.25 = 9.25$$

$$y = 1.39x - 20.1 = 1.39(21) - 20.1 = 9.09$$

The prediction using the equation from part (a) is 9.25 pounds. The prediction using the equation from part (b) is 9.09 pounds. The baby is slightly heavier according to the equation from part (a).

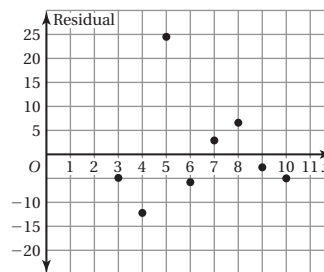
7. Use the *linear regression* feature of a graphing calculator.

```
LinReg
y=ax+b
a=20.3214
b=1350.2857
r^2=.951215
r=.975303
```

So, an equation for the line of best fit is

$$y = 20.3x + 1350.$$

x	y	y -value from Model	Residual
3	1406	1410.9	$1406 - 1410.9 = -4.9$
4	1419	1431.2	$1419 - 1431.2 = -12.2$
5	1476	1451.5	$1476 - 1451.5 = 24.5$
6	1466	1471.8	$1466 - 1471.8 = -5.8$
7	1495	1492.1	$1495 - 1492.1 = 2.9$
8	1519	1512.4	$1519 - 1512.4 = 6.6$
9	1530	1532.7	$1530 - 1532.7 = -2.7$
10	1548	1553	$1548 - 1553 = -5$



The residual points are randomly dispersed about the horizontal axis. So, the line of best fit models the data well.

- 8.

		Use reusable bags?		
		Yes	No	Total
Gender	Male	15	45	60
	Female	60	50	110
Total		75	95	170

Chapter 12

Chapter 12 Standardized Test Practice (pp. 671–673)

1. B;

$$\frac{36 + 28 + x + 48 + 42 + 57 + 63 + 52}{8} = 45.5$$

$$\frac{326 + x}{8} = 45.5$$

$$326 + x = 364$$

$$x = 38$$

2. G;

The graph of the function $y = 3x^2 - 1$ opens up and is narrower than the graph of $y = x^2$. The vertex of the graph of $y = 3x^2 - 1$ is a translation one unit down of the vertex of the graph of $y = x^2$.

3. 16;

$$-6 + \sqrt{x} = -2$$

$$\sqrt{x} = 4$$

$$(\sqrt{x})^2 = 4^2$$

$$x = 16$$

The solution is 16.

4. C;

$$\begin{aligned} (3z - 2)(2z + 4) &= 6z^2 + 12z - 4z - 8 \\ &= 6z^2 + 8z - 8 \end{aligned}$$

5. H;

$$\begin{aligned} \text{Interquartile range} &= \text{third quartile} - \text{first quartile} \\ &= 7 - 4 \\ &= 3 \end{aligned}$$

The interquartile range is 3. This means the middle half of the presentation lengths vary by no more than 3 minutes.

6. B;

$$\begin{aligned} \left(\frac{4}{5x}\right)^{-2} &= \left(\frac{5x}{4}\right)^2 \\ &= \frac{(5x)^2}{4^2} \\ &= \frac{25x^2}{16} \end{aligned}$$

7. I;

An equation shows inverse variation when it is of the form $y = \frac{k}{x}$. The equation $y = \frac{4}{x}$ is of the form

$y = \frac{k}{x}$. So, the equation $y = \frac{4}{x}$ shows inverse

variation.

8. 12;

The joint frequency in the "Male" row and "No" column is 12. So, 12 male students do not have a job.

9. D;

The first term is $-\frac{3}{4}$ and the common difference is $\frac{1}{2}$.

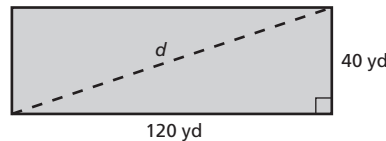
$$a_n = a_1 + (n - 1)d$$

$$a_n = -\frac{3}{4} + (n - 1)\frac{1}{2}$$

$$a_n = -\frac{3}{4} + \frac{1}{2}n - \frac{1}{2}$$

$$a_n = \frac{1}{2}n - \frac{5}{4}$$

10. About 126.5 yards;



A football field is in the shape of a rectangle. Drawing a line connecting opposite corners of the rectangle creates two right triangles with legs of 40 yards and 120 yards in length. The distance between opposite corners, d , is the length of the hypotenuse.

Use the Pythagorean Theorem.

$$a^2 + b^2 = d^2$$

$$40^2 + 120^2 = d^2$$

$$1600 + 14,400 = d^2$$

$$16,000 = d^2$$

$$\sqrt{16,000} = \sqrt{d^2}$$

$$126.5 \approx d$$

So, the distance between opposite corners of the football field is about 126.5 yards.

11. I;

When a scatter plot shows a negative relationship between x and y , x increases as y decreases. In the scatter plot shown in I, x increases as y decreases. So, the scatter plot shown in I shows a negative relationship between x and y .

12. C;

$$0.22(x + 6) = 0.2x + 1.8$$

$$0.22x + 1.32 = 0.2x + 1.8$$

$$0.02x + 1.32 = 1.8$$

$$0.02x = 0.48$$

$$x = 24$$