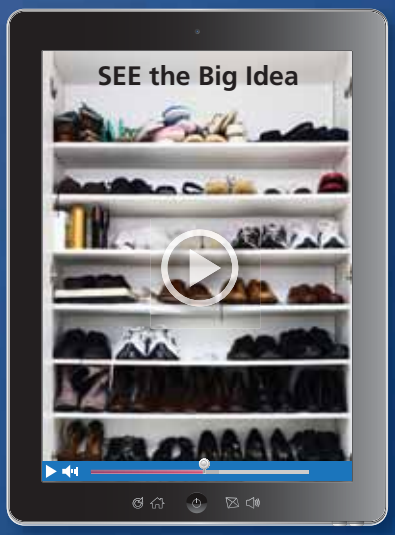
11 Data Analysis and Displays

11.1 Measures of Center and Variation
11.2 Box-and-Whisker Plots
11.3 Shapes of Distributions
11.4 Two-Way Tables
11.5 Choosing a Data Display

Watching Sports on TV *(p. 616)*



Shoes *(p. 603)*

Backpacking *(p. 594)*

Bowling Scores *(p. 591)*

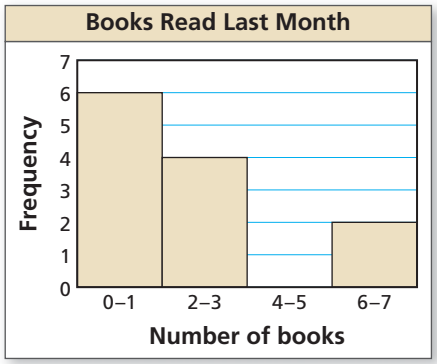
Altitudes of Airplanes *(p. 589)*

Maintaining Mathematical Proficiency

Displaying Data

Example 1 The frequency table shows the numbers of books that 12 people read last month. Display the data in a histogram.

Number of books	Frequency
0–1	6
2–3	4
4–5	0
6–7	2



Example 2 The table shows the results of a survey. Display the data in a circle graph.

Class trip location	Water park	Museum	Zoo	Other
Students	25	11	5	4

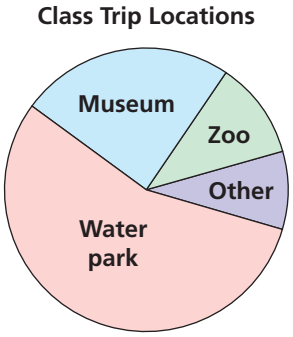
A total of 45 students took the survey.

Water park: $\frac{25}{45} \cdot 360° = 200°$

Museum: $\frac{11}{45} \cdot 360° = 88°$

Zoo: $\frac{5}{45} \cdot 360° = 40°$

Other: $\frac{4}{45} \cdot 360° = 32°$



The table shows the results of a survey. Display the data in a histogram.

1.

After-school activities	Frequency
0–1	11
2–3	8
4–5	6
6–7	1

2.

Pets	Frequency
0–1	10
2–3	18
4–5	2

The table shows the results of a survey. Display the data in a circle graph.

3.

Favorite subject	Math	Science	English	History
Students	8	5	7	4

4. ABSTRACT REASONING Twenty people respond "yes" or "no" to a survey question. Let a and b represent the frequencies of the responses. What must be true about the sum of a and b? What must be true about the sum when "maybe" is an option for the response?

Mathematical Practices

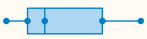
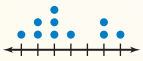
Mathematically proficient students use diagrams and graphs to show relationships between data. They also analyze data to draw conclusions.

Using Data Displays

Core Concept

Displaying Data Graphically

When solving a problem involving data, it is helpful to display the data graphically. This can be done in a variety of ways.

Data Display	What does it do?	
Pictograph	shows data using pictures	
Bar Graph	shows data in specific categories	
Circle Graph	shows data as parts of a whole	
Line Graph	shows how data change over time	
Histogram	shows frequencies of data values in intervals of the same size	
Stem-and-Leaf Plot	orders numerical data and shows how they are distributed	
Box-and-Whisker Plot	shows the variability of a data set using quartiles	
Dot Plot	shows the number of times each value occurs in a data set	
Scatter Plot	shows the relationship between two data sets using ordered pairs in a coordinate plane	

Monitoring Progress

1. The table shows the estimated populations of males and females by age in the United States in 2012. Use a spreadsheet, graphing calculator, or some other form of technology to make two different displays for the data.
2. Explain why you chose each type of data display in Monitoring Progress Question 1. What conclusions can you draw from your data displays?

U.S. Population by Age and Gender		
Ages (years)	**Males**	**Females**
0–14	31,242,542	29,901,556
15–29	33,357,203	31,985,028
30–44	30,667,513	30,759,902
45–59	31,875,279	33,165,976
60–74	19,737,347	22,061,730
75–89	6,999,292	10,028,195

Mathematical Practices

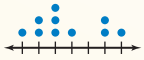
Mathematically proficient students identify important quantities in a practical situation and ***map their relationships using tools such as diagrams and graphs. They can analyze those relationships mathematically to draw conclusions.*** **(MP4)**

Using Data Displays

Core Concept

Displaying Data Graphically

When solving a problem involving data, it is helpful to display the data graphically. This can be done in a variety of ways.

Data Display	What does it do?	
Pictograph	shows data using pictures	
Bar Graph	shows data in specific categories	
Circle Graph	shows data as parts of a whole	
Line Graph	shows how data change over time	
Histogram	shows frequencies of data values in intervals of the same size	
Stem-and-Leaf Plot	orders numerical data and shows how they are distributed	
Box-and-Whisker Plot	shows the variability of a data set using quartiles	
Dot Plot	shows the number of times each value occurs in a data set	
Scatter Plot	shows the relationship between two data sets using ordered pairs in a coordinate plane	

Monitoring Progress

1. The table shows the estimated populations of males and females by age in the United States in 2012. Use a spreadsheet, graphing calculator, or some other form of technology to make two different displays for the data.
2. Explain why you chose each type of data display in Monitoring Progress Question 1. What conclusions can you draw from your data displays?

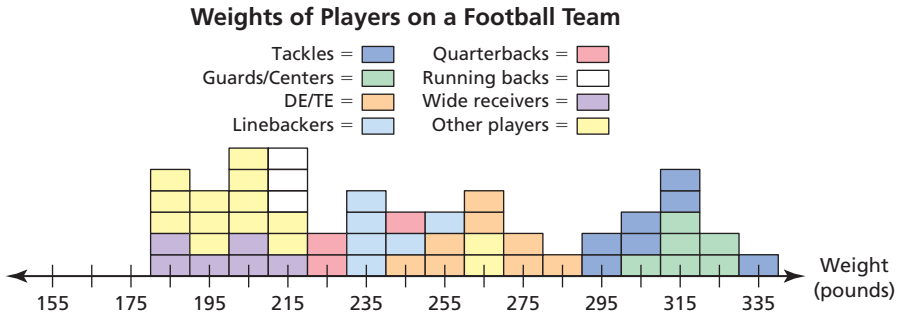
U.S. Population by Age and Gender		
Ages (years)	**Males**	**Females**
0–14	31,242,542	29,901,556
15–29	33,357,203	31,985,028
30–44	30,667,513	30,759,902
45–59	31,875,279	33,165,976
60–74	19,737,347	22,061,730
75–89	6,999,292	10,028,195

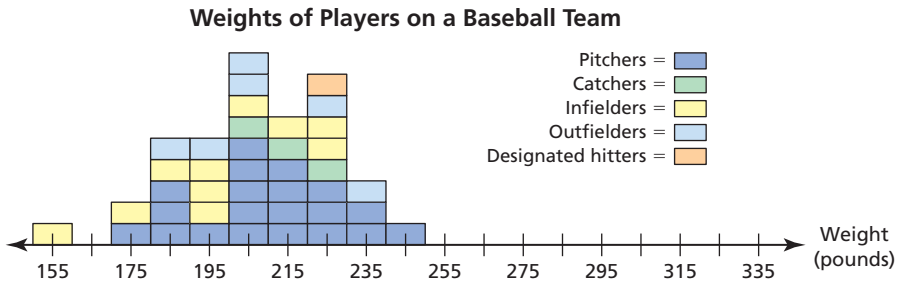
11.1 Measures of Center and Variation

Essential Question How can you describe the variation of a data set?

EXPLORATION 1 Describing the Variation of Data

Work with a partner. The graphs show the weights of the players on a professional football team and a professional baseball team.





CONSTRUCTING VIABLE ARGUMENTS

To be proficient in math, you need to reason inductively about data, making plausible arguments that take into account the context from which the data arose.

a. Describe the data in each graph in terms of how much the weights vary from the mean. Explain your reasoning.

b. Compare how much the weights of the players on the football team vary from the mean to how much the weights of the players on the baseball team vary from the mean.

c. Does there appear to be a correlation between the body weights and the positions of players in professional football? in professional baseball? Explain.

EXPLORATION 2 Describing the Variation of Data

Work with a partner. The weights (in pounds) of the players on a professional basketball team by position are as follows.

Power forwards: 235, 255, 295, 245; small forwards: 235, 235;
centers: 255, 245, 325; point guards: 205, 185, 205; shooting guards: 205, 215, 185

Make a graph that represents the weights and positions of the players. Does there appear to be a correlation between the body weights and the positions of players in professional basketball? Explain your reasoning.

Communicate Your Answer

3. How can you describe the variation of a data set?

11.1 Lesson

Core Vocabulary

measure of center, *p. 586*
mean, *p. 586*
median, *p. 586*
mode, *p. 586*
outlier, *p. 587*
measure of variation, *p. 587*
range, *p. 587*
standard deviation, *p. 588*
data transformation, *p. 589*

What You Will Learn

- Compare the mean, median, and mode of a data set.
- Find the range and standard deviation of a data set.
- Identify the effects of transformations on data.

Comparing the Mean, Median, and Mode

A **measure of center** is a measure that represents the center, or typical value, of a data set. The *mean*, *median*, and *mode* are measures of center.

Core Concept

Mean

The **mean** of a numerical data set is the sum of the data divided by the number of data values. The symbol $\bar{x}$ represents the mean. It is read as "x-bar."

Median

The **median** of a numerical data set is the middle number when the values are written in numerical order. When a data set has an even number of values, the median is the mean of the two middle values.

Mode

The **mode** of a data set is the value or values that occur most often. There may be one mode, no mode, or more than one mode.

EXAMPLE 1 Comparing Measures of Center

Students' Hourly Wages	
$16.50	$8.25
$8.75	$8.45
$8.65	$8.25
$9.10	$9.25

An amusement park hires students for the summer. The students' hourly wages are shown in the table.

a. Find the mean, median, and mode of the hourly wages.

b. Which measure of center best represents the data? Explain.

SOLUTION

a. Mean $\bar{x} = \dfrac{16.5 + 8.75 + 8.65 + 9.1 + 8.25 + 8.45 + 8.25 + 9.25}{8} = 9.65$

Median 8.25, 8.25, 8.45, 8.65, 8.75, 9.10, 9.25, 16.50 Order the data.

$\dfrac{17.4}{2} = 8.7$ Mean of two middle values

Mode 8.25, 8.25, 8.45, 8.65, 8.75, 9.10, 9.25, 16.50 8.25 occurs most often.

▶ The mean is $9.65, the median is $8.70, and the mode is $8.25.

b. The median best represents the data. The mode is less than most of the data, and the mean is greater than most of the data.

STUDY TIP

Mode is the only measure of center that can represent a nonnumerical data set.

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

1. **WHAT IF?** The park hires another student at an hourly wage of $8.45.
   (a) How does this additional value affect the mean, median, and mode? Explain.
   (b) Which measure of center best represents the data? Explain.

An **outlier** is a data value that is much greater than or much less than the other values in a data set.

EXAMPLE 2 Removing an Outlier

Consider the data in Example 1. (a) Identify the outlier. How does the outlier affect the mean, median, and mode? (b) Describe one possible explanation for the outlier.

SOLUTION

a. The value \$16.50 is much greater than the other wages. It is the outlier. Find the mean, median, and mode without the outlier.

Mean $\bar{x} = \frac{8.75 + 8.65 + 9.1 + 8.25 + 8.45 + 8.25 + 9.25}{7} \approx 8.67$

Median 8.25, 8.25, 8.45, 8.65, 8.75, 9.10, 9.25 — The middle value is 8.65.

Mode 8.25, 8.25, 8.45, 8.65, 8.75, 9.10, 9.25 — The mode is 8.25.

▶ When you remove the outlier, the mean decreases \$9.65 − \$8.67 = \$0.98, the median decreases \$8.70 − \$8.65 = \$0.05, and the mode is the same.

b. The outlier could be a student who is hired to maintain the park's website, while the other students could be game attendants.

STUDY TIP

Outliers usually have the greatest effect on the mean.

Annual Salaries	
\$32,000	\$42,000
\$41,000	\$38,000
\$38,000	\$45,000
\$72,000	\$35,000

Monitoring Progress Help in English and Spanish at *BigIdeasMath.com*

2. The table shows the annual salaries of the employees of an auto repair service. (a) Identify the outlier. How does the outlier affect the mean, median, and mode? (b) Describe one possible explanation for the outlier.

Finding the Range and Standard Deviation

A **measure of variation** is a measure that describes the spread, or distribution, of a data set. One measure of variation is the *range*. The **range** of a data set is the difference of the greatest value and the least value.

EXAMPLE 3 Finding a Range

Two reality cooking shows select 12 contestants each. The ages of the contestants are shown in the tables. Find the range of the ages for each show. Compare your results.

Show A	
Ages	
20	29
19	22
25	27
27	29
30	20
21	31

Show B	
Ages	
25	19
20	27
22	25
27	22
48	21
32	24

SOLUTION

Show A 19, 20, 20, 21, 22, 25, 27, 27, 29, 29, 30, 31 — Order the data.

So, the range is 31 − 19, or 12 years.

Show B 19, 20, 21, 22, 22, 24, 25, 25, 27, 27, 32, 48 — Order the data.

So, the range is 48 − 19, or 29 years.

▶ The range of the ages for Show A is 12 years, and the range of the ages for Show B is 29 years. So, the ages for Show B are more spread out.

Monitoring Progress Help in English and Spanish at *BigIdeasMath.com*

3. After the first week, the 25-year-old is voted off Show A and the 48-year-old is voted off Show B. How does this affect the range of the ages of the remaining contestants on each show in Example 3? Explain.

A disadvantage of using the range to describe the spread of a data set is that it uses only two data values. A measure of variation that uses all the values of a data set is the *standard deviation*.

Core Concept

Standard Deviation

The **standard deviation** of a numerical data set is a measure of how much a typical value in the data set differs from the mean. The symbol σ represents the standard deviation. It is read as "sigma." It is given by

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}}$$

where n is the number of values in the data set. The deviation of a data value x is the difference of the data value and the mean of the data set, $x - \bar{x}$.

Step 1 Find the mean, $\bar{x}$.

Step 2 Find the deviation of each data value, $x - \bar{x}$.

Step 3 Square each deviation, $(x - \bar{x})^2$.

Step 4 Find the mean of the squared deviations. This is called the *variance*.

Step 5 Take the square root of the variance.

REMEMBER

An ellipsis " · · · " indicates that a pattern continues.

A small standard deviation means that the data are clustered around the mean. A large standard deviation means that the data are more spread out.

EXAMPLE 4 Finding a Standard Deviation

Find the standard deviation of the ages for Show A in Example 3. Use a table to organize your work. Interpret your result.

SOLUTION

Step 1 Find the mean, $\bar{x}$.

$$\bar{x} = \frac{300}{12} = 25$$

Step 2 Find the deviation of each data value, $x - \bar{x}$, as shown.

Step 3 Square each deviation, $(x - \bar{x})^2$, as shown.

x	$\bar{x}$	$x - \bar{x}$	$(x - \bar{x})^2$
20	25	−5	25
29	25	4	16
19	25	−6	36
22	25	−3	9
25	25	0	0
27	25	2	4
27	25	2	4
29	25	4	16
30	25	5	25
20	25	−5	25
21	25	−4	16
31	25	6	36

Step 4 Find the mean of the squared deviations, or variance.

$$\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n} = \frac{25 + 16 + \cdots + 36}{12} = \frac{212}{12} \approx 17.7$$

Step 5 Use a calculator to take the square root of the variance.

$$\sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}} = \sqrt{\frac{212}{12}} \approx 4.2$$

▶ The standard deviation is about 4.2. This means that the typical age of a contestant on Show A differs from the mean by about 4.2 years.

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

4. Find the standard deviation of the ages for Show B in Example 3. Interpret your result.
5. Compare the standard deviations for Show A and Show B. What can you conclude?

Effects of Data Transformations

A **data transformation** is a procedure that uses a mathematical operation to change a data set into a different data set.

STUDY TIP

The standard deviation stays the same because the amount by which each data value deviates from the mean stays the same.

Core Concept

Data Transformations Using Addition

When a real number k is added to each value in a numerical data set

- the measures of center of the new data set can be found by adding k to the original measures of center.
- the measures of variation of the new data set are the *same* as the original measures of variation.

Data Transformations Using Multiplication

When each value in a numerical data set is multiplied by a real number k, where $k > 0$, the measures of center and variation can be found by multiplying the original measures by k.

EXAMPLE 5 Real-Life Application

Consider the data in Example 1. (a) Find the mean, median, mode, range, and standard deviation when each hourly wage increases by \$0.50. (b) Find the mean, median, mode, range, and standard deviation when each hourly wage increases by 10%.

SOLUTION

Students' Hourly Wages	
\$17.00	\$8.75
\$9.25	\$8.95
\$9.15	\$8.75
\$9.60	\$9.75

a. Method 1 Make a new table by adding \$0.50 to each hourly wage. Find the mean, median, mode, range, and standard deviation of the new data set.

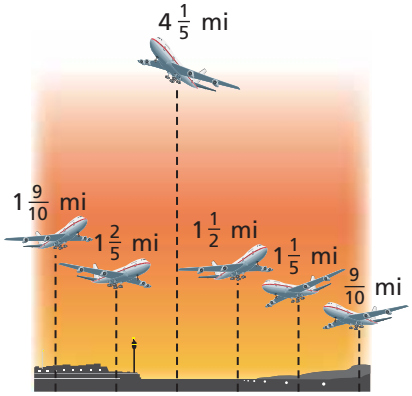
▶ Mean: \$10.15 Median: \$9.20 Mode: \$8.75
Range: \$8.25 Standard deviation: \$2.61

Method 2 Find the mean, median, mode, range, and standard deviation of the original data set.

Mean: \$9.65 Median: \$8.70 Mode: \$8.25 From Example 1
Range: \$8.25 Standard deviation: \$2.61

Add \$0.50 to the mean, median, and mode. The range and standard deviation are the same as the original range and standard deviation.

▶ Mean: \$10.15 Median: \$9.20 Mode: \$8.75
Range: \$8.25 Standard deviation: \$2.61



b. Increasing by 10% means to multiply by 1.1. So, multiply the original mean, median, mode, range, and standard deviation from Method 2 of part (a) by 1.1.

▶ Mean: \$10.62 Median: \$9.57 Mode: \$9.08
Range: \$9.08 Standard deviation: \$2.87

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

6. Find the mean, median, mode, range, and standard deviation of the altitudes of the airplanes when each altitude increases by $1\frac{1}{2}$ miles.

11.1 Exercises

Dynamic Solutions available at *BigIdeasMath.com*

Vocabulary and Core Concept Check

1. **VOCABULARY** In a data set, what does a measure of center represent? What does a measure of variation describe?
2. **WRITING** Describe how removing an outlier from a data set affects the mean of the data set.
3. **OPEN-ENDED** Create a data set that has more than one mode.
4. **REASONING** What is an advantage of using the range to describe a data set? Why do you think the standard deviation is considered a more reliable measure of variation than the range?

Monitoring Progress and Modeling with Mathematics

In Exercises 5–8, (a) find the mean, median, and mode of the data set and (b) determine which measure of center best represents the data. Explain. *(See Example 1.)*

5. 3, 5, 1, 5, 1, 1, 2, 3, 15
6. 12, 9, 17, 15, 10
7. 13, 30, 16, 19, 20, 22, 25, 31
8. 14, 15, 3, 15, 14, 14, 18, 15, 8, 16

9. **ANALYZING DATA** The table shows the lengths of nine movies.

Movie Lengths (hours)		
$1\frac{1}{3}$	$1\frac{2}{3}$	2
3	$2\frac{1}{3}$	$1\frac{2}{3}$
2	2	$1\frac{2}{3}$

   a. Find the mean, median, and mode of the lengths.

   b. Which measure of center best represents the data? Explain.

10. **ANALYZING DATA** The table shows the daily changes in the value of a stock over 12 days.

Changes in Stock Value (dollars)			
1.05	2.03	−13.78	−2.41
2.64	0.67	4.02	1.39
0.66	−0.28	−3.01	2.20

   a. Find the mean, median, and mode of the changes in stock value.

   b. Which measure of center best represents the data? Explain.

   c. On the 13th day, the value of the stock increases by $4.28. How does this additional value affect the mean, median, and mode? Explain.

In Exercises 11–14, find the value of x.

11. 2, 8, 9, 7, 6, x; The mean is 6.
12. 12.5, −10, −7.5, x; The mean is 11.5.
13. 9, 10, 12, x, 20, 25; The median is 14.
14. 30, 45, x, 100; The median is 51.

15. **ANALYZING DATA** The table shows the masses of eight polar bears. *(See Example 2.)*

Masses (kilograms)			
455	262	471	358
364	553	62	351

   a. Identify the outlier. How does the outlier affect the mean, median, and mode?

   b. Describe one possible explanation for the outlier.

16. **ANALYZING DATA** The sizes of emails (in kilobytes) in your inbox are 2, 3, 5, 2, 1, 46, 3, 7, 2, and 1.

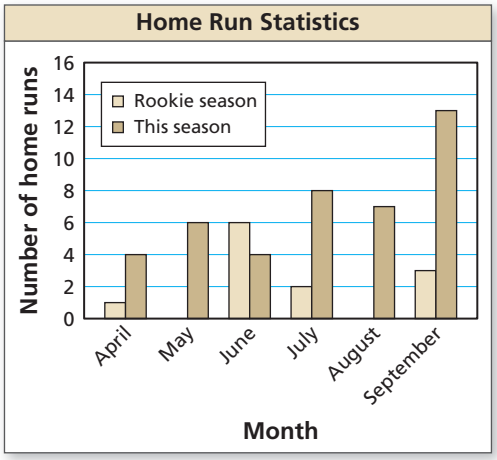
   a. Identify the outlier. How does the outlier affect the mean, median, and mode?

   b. Describe one possible explanation for the outlier.

17. **ANALYZING DATA** The scores of two golfers are shown. Find the range of the scores for each golfer. Compare your results. *(See Example 3.)*

Golfer A		Golfer B	
83	88	89	87
84	95	93	95
91	89	92	94
90	87	88	91
98	95	89	92

18. ANALYZING DATA The graph shows a player's monthly home run totals in two seasons. Find the range of the number of home runs for each season. Compare your results.



In Exercises 19–22, find (a) the range and (b) the standard deviation of the data set.

19. 40, 35, 45, 55, 60

20. 141, 116, 117, 135, 126, 121

21. 0.5, 2.0, 2.5, 1.5, 1.0, 1.5

22. 8.2, 10.1, 2.6, 4.8, 2.4, 5.6, 7.0, 3.3

23. ANALYZING DATA Consider the data in Exercise 17. *(See Example 4.)*

a. Find the standard deviation of the scores of Golfer A. Interpret your result.

b. Find the standard deviation of the scores of Golfer B. Interpret your result.

c. Compare the standard deviations for Golfer A and Golfer B. What can you conclude?

24. ANALYZING DATA Consider the data in Exercise 18.

a. Find the standard deviation of the monthly home run totals in the player's rookie season. Interpret your result.

b. Find the standard deviation of the monthly home run totals in this season. Interpret your result.

c. Compare the standard deviations for the rookie season and this season. What can you conclude?

In Exercises 25 and 26, find the mean, median, and mode of the data set after the given transformation.

25. In Exercise 5, each data value increases by 4.

26. In Exercise 6, each data value increases by 20%.

27. TRANSFORMING DATA Find the values of the measures shown when each value in the data set increases by 14. *(See Example 5.)*

Mean: 62 Median: 55 Mode: 49
Range: 46 Standard deviation: 15.5

28. TRANSFORMING DATA Find the values of the measures shown when each value in the data set is multiplied by 0.5.

Mean: 320 Median: 300 Mode: none
Range: 210 Standard deviation: 70.6

29. ERROR ANALYSIS Describe and correct the error in finding the median of the data set.

✗ 7, 4, 6, 2, 4, 6, 8, 8, 3
The median is 4.

30. ERROR ANALYSIS Describe and correct the error in finding the range of the data set after the given transformation.

✗ −13, −12, −7, 2, 10, 13
Add 10 to each value.
The range is 26 + 10 = 36.

31. PROBLEM SOLVING In a bowling match, the team with the greater mean score wins. The scores of the members of two bowling teams are shown.

Team A: 172, 130, 173, 212
Team B: 136, 184, 168, 192

a. Which team wins the match? If the team with the greater median score wins, is the result the same? Explain.

b. Which team is more consistent? Explain.

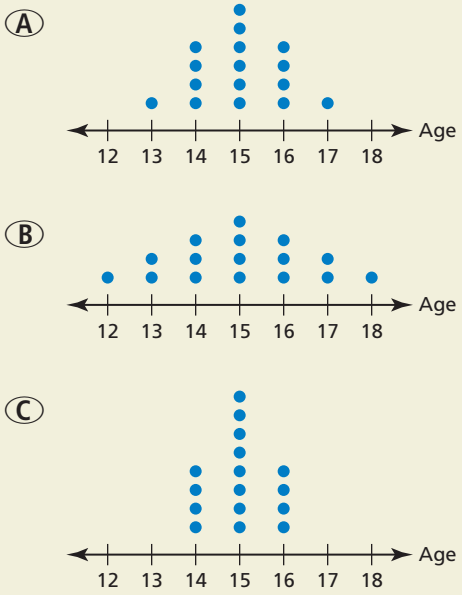
c. In another match between the two teams, all the members of Team A increase their scores by 15 and all the members of Team B increase their scores by 12.5%. Which team wins this match? Explain.

32. MAKING AN ARGUMENT Your friend says that when two data sets have the same range, you can assume the data sets have the same standard deviation, because both range and standard deviation are measures of variation. Is your friend correct? Explain.

33. **ANALYZING DATA** The table shows the results of a survey that asked 12 students about their favorite meal. Which measure of center (mean, median, or mode) can be used to describe the data? Explain.

Favorite Meal			
spaghetti	pizza	steak	hamburger
steak	taco	pizza	chili
pizza	chicken	fish	spaghetti

34. **HOW DO YOU SEE IT?** The dot plots show the ages of the members of three different adventure clubs. Without performing calculations, which data set has the greatest standard deviation? Which has the least standard deviation? Explain your reasoning.



35. **REASONING** A data set is described by the measures shown.

Mean: 27 Median: 32 Mode: 18
Range: 41 Standard deviation: 9

Find the mean, median, mode, range, and standard deviation of the data set when each data value is multiplied by 3 and then increased by 8.

36. **CRITICAL THINKING** Can the standard deviation of a data set be 0? Can it be negative? Explain.

37. **USING TOOLS** Measure the heights (in inches) of the students in your class.

a. Find the mean, median, mode, range, and standard deviation of the heights.

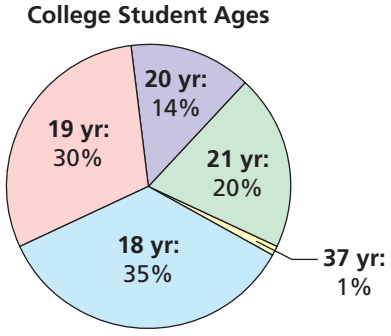
b. A new student who is 7 feet tall joins your class. How would you expect this student's height to affect the measures in part (a)? Verify your answer.

38. **THOUGHT PROVOKING** To find the arithmetic mean of n numbers, divide the sum of the numbers by n. To find the geometric mean of n numbers $a_1, a_2, a_3, \ldots, a_n$, take the nth root of the product of the numbers.

$$\text{geometric mean} = \sqrt[n]{a_1 \cdot a_2 \cdot a_3 \cdot \ldots \cdot a_n}$$

Compare the arithmetic mean to the geometric mean of n numbers.

39. **PROBLEM SOLVING** The circle graph shows the distribution of the ages of 200 students in a college Psychology I class.



a. Find the mean, median, and mode of the students' ages.

b. Identify the outliers. How do the outliers affect the mean, median, and mode?

c. Suppose all 200 students take the same Psychology II class exactly 1 year later. Draw a new circle graph that shows the distribution of the ages of this class and find the mean, median, and mode of the students' ages.

Maintaining Mathematical Proficiency

Reviewing what you learned in previous grades and lessons

Solve the inequality. *(Section 2.4)*

40. $6x + 1 \leq 4x - 9$ **41.** $-3(3y - 2) < 1 - 9y$ **42.** $2(5c - 4) \geq 5(2c + 8)$ **43.** $4(3 - w) > 3(4w - 4)$

Evaluate the function for the given value of x. *(Section 6.3)*

44. $f(x) = 4^x; x = 3$ **45.** $f(x) = 7^x; x = -2$ **46.** $f(x) = 5(2)^x; x = 6$ **47.** $f(x) = -2(3)^x; x = 4$

11.2 Box-and-Whisker Plots

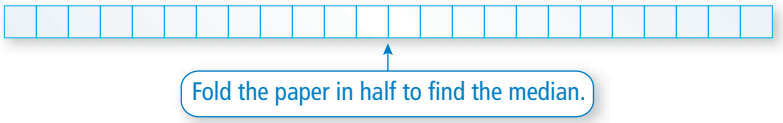
Essential Question How can you use a box-and-whisker plot to describe a data set?

EXPLORATION 1 Drawing a Box-and-Whisker Plot

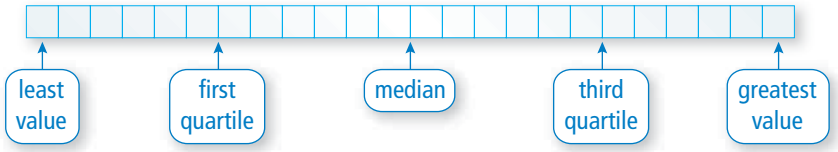
Work with a partner. The numbers of first cousins of the students in a ninth-grade class are shown. A *box-and-whisker plot* is one way to represent the data visually.

Numbers of First Cousins			
3	10	18	8
9	3	0	32
23	19	13	8
6	3	3	10
12	45	1	5
13	24	16	14

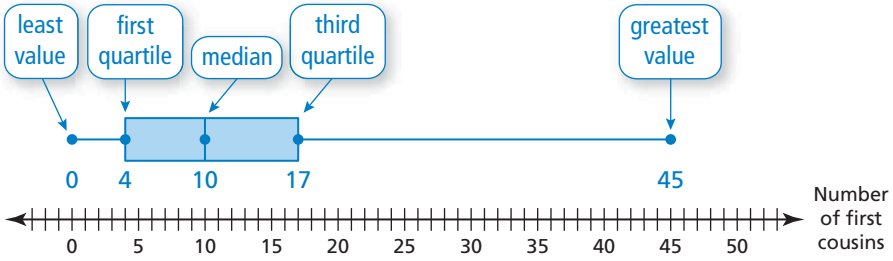
a. Order the data on a strip of grid paper with 24 equally spaced boxes.



b. Fold the paper in half again to divide the data into four groups. Because there are 24 numbers in the data set, each group should have 6 numbers. Find the least value, the greatest value, the first quartile, and the third quartile.



c. Explain how the box-and-whisker plot shown represents the data set.



MODELING WITH MATHEMATICS

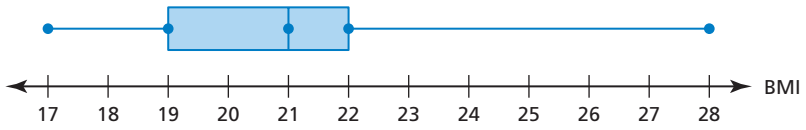
To be proficient in math, you need to identify important quantities in a practical situation.

Communicate Your Answer

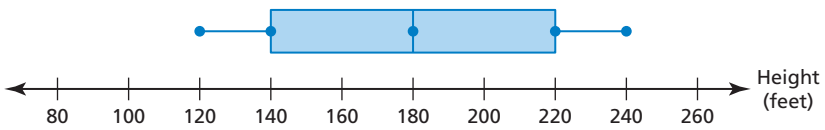
2. How can you use a box-and-whisker plot to describe a data set?

3. Interpret each box-and-whisker plot.

a. body mass indices (BMI) of students in a ninth-grade class



b. heights of roller coasters at an amusement park



11.2 Lesson

Core Vocabulary

box-and-whisker plot, *p. 594*
quartile, *p. 594*
five-number summary, *p. 594*
interquartile range, *p. 595*

What You Will Learn

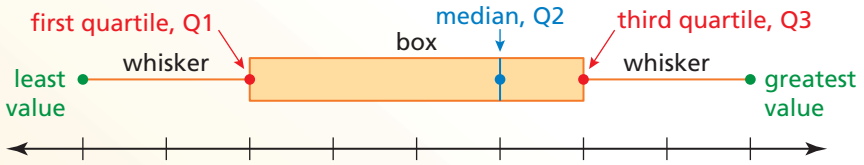
- Use box-and-whisker plots to represent data sets.
- Interpret box-and-whisker plots.
- Use box-and-whisker plots to compare data sets.

Using Box-and-Whisker Plots to Represent Data Sets

Core Concept

Box-and-Whisker Plot

A **box-and-whisker plot** shows the variability of a data set along a number line using the least value, the greatest value, and the *quartiles* of the data. **Quartiles** divide the data set into four equal parts. The median (second quartile, Q2) divides the data set into two halves. The median of the lower half is the first quartile, Q1. The median of the upper half is the third quartile, Q3.



The five numbers that make up a box-and-whisker plot are called the **five-number summary** of the data set.

STUDY TIP

Sometimes, the first quartile is called the *lower quartile* and the third quartile is called the *upper quartile*.

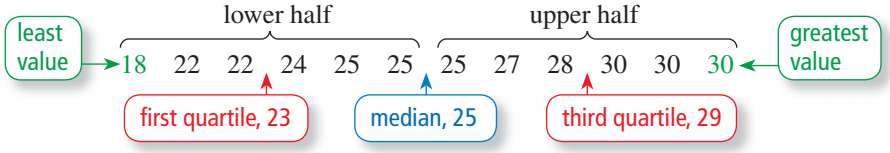
EXAMPLE 1 Making a Box-and-Whisker Plot

Make a box-and-whisker plot that represents the ages of the members of a backpacking expedition in the mountains.

24, 30, 30, 22, 25, 22, 18, 25, 28, 30, 25, 27

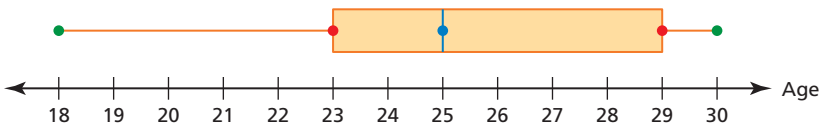
SOLUTION

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



Step 2 Draw a number line that includes the least and greatest values. Graph points above the number line for the five-number summary.

Step 3 Draw a box using Q1 and Q3. Draw a line through the median. Draw whiskers from the box to the least and greatest values.



Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

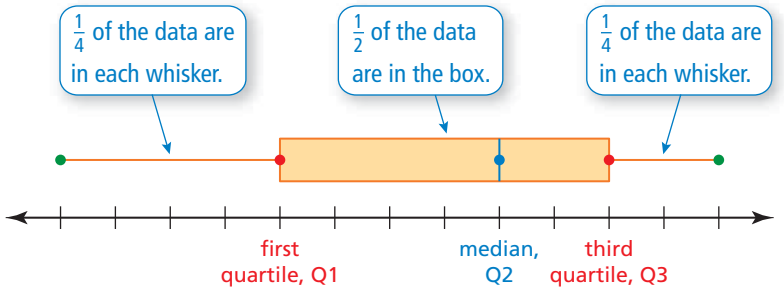
1. A basketball player scores 14, 16, 20, 5, 22, 30, 16, and 28 points during a tournament. Make a box-and-whisker plot that represents the data.

Interpreting Box-and-Whisker Plots

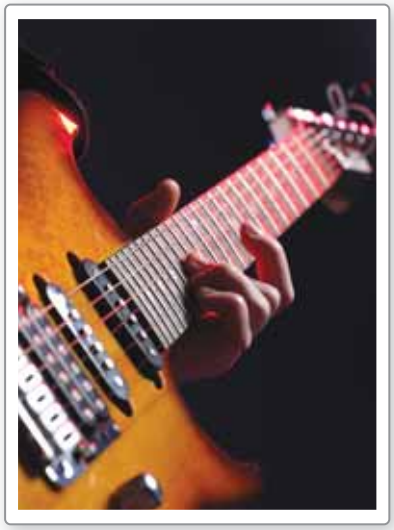
The figure shows how data are distributed in a box-and-whisker plot.

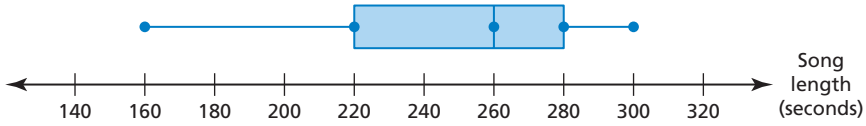
STUDY TIP

A long whisker or box indicates that the data are more spread out.



Another measure of variation for a data set is the **interquartile range** (IQR), which is the difference of the third quartile, Q3, and the first quartile, Q1. It represents the range of the middle half of the data.

EXAMPLE 2 Interpreting a Box-and-Whisker Plot

The box-and-whisker plot represents the lengths (in seconds) of the songs played by a rock band at a concert.



a. Find and interpret the range of the data.

b. Describe the distribution of the data.

c. Find and interpret the interquartile range of the data.

d. Are the data more spread out below Q1 or above Q3? Explain.

SOLUTION

a. The least value is 160. The greatest value is 300.

So, the range is $300 - 160 = 140$ seconds. This means that the song lengths vary by no more than 140 seconds.

b. Each whisker represents 25% of the data. The box represents 50% of the data. So,

- 25% of the song lengths are between 160 and 220 seconds.
- 50% of the song lengths are between 220 and 280 seconds.
- 25% of the song lengths are between 280 and 300 seconds.

c. $\text{IQR} = \text{Q3} - \text{Q1} = 280 - 220 = 60$

So, the interquartile range is 60 seconds. This means that the middle half of the song lengths vary by no more than 60 seconds.

d. The left whisker is longer than the right whisker.

So, the data below Q1 are more spread out than data above Q3.

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

Use the box-and-whisker plot in Example 1.

2. Find and interpret the range and interquartile range of the data.

3. Describe the distribution of the data.

Using Box-and-Whisker Plots to Compare Data Sets

A box-and-whisker plot shows the shape of a distribution.

STUDY TIP

If you can draw a line through the median of a box-and-whisker plot, and each side is approximately a mirror image of the other, then the distribution is symmetric.

Core Concept

Shapes of Box-and-Whisker Plots

Skewed left

- The left whisker is longer than the right whisker.
- Most of the data are on the right side of the plot.

Symmetric

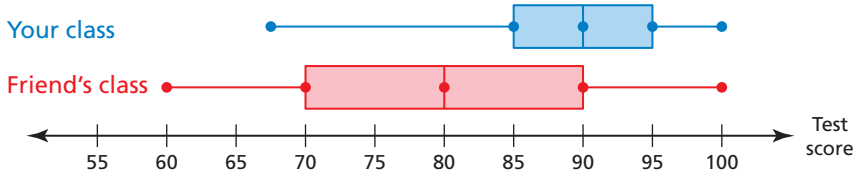
- The whiskers are about the same length.
- The median is in the middle of the plot.

Skewed right

- The right whisker is longer than the left whisker.
- Most of the data are on the left side of the plot.

EXAMPLE 3 Comparing Box-and-Whisker Plots

The double box-and-whisker plot represents the test scores for your class and your friend's class.



a. Identify the shape of each distribution.

b. Which test scores are more spread out? Explain.

SOLUTION

a. For your class, the left whisker is longer than the right whisker, and most of the data are on the right side of the plot. For your friend's class, the whisker lengths are equal, and the median is in the middle of the plot.

▶ So, the distribution for your class is skewed left, and the distribution for your friend's class is symmetric.

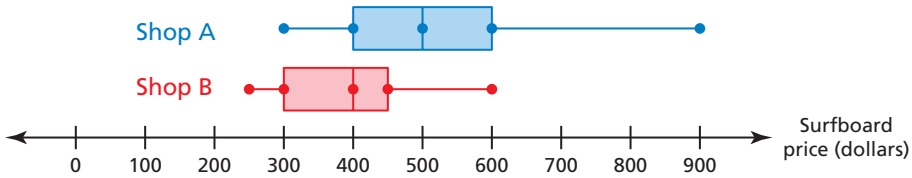
b. The range and interquartile range of the test scores in your friend's class are greater than the range and interquartile range in your class.

▶ So, the test scores in your friend's class are more spread out.

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

4. The double box-and-whisker plot represents the surfboard prices at Shop A and Shop B. Identify the shape of each distribution. Which shop's prices are more spread out? Explain.



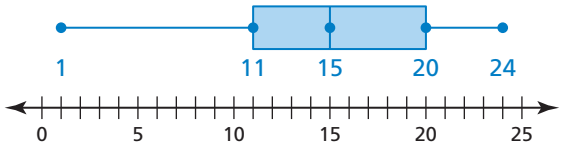
11.2 Exercises

Dynamic Solutions available at *BigIdeasMath.com*

Vocabulary and Core Concept Check

1. **WRITING** Describe how to find the first quartile of a data set.

2. **DIFFERENT WORDS, SAME QUESTION** Consider the box-and-whisker plot shown. Which is different? Find "both" answers.



Find the interquartile range of the data.

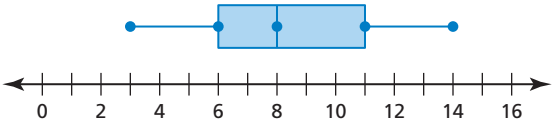
Find the range of the middle half of the data.

Find the difference of the greatest value and the least value of the data set.

Find the difference of the third quartile and the first quartile.

Monitoring Progress and Modeling with Mathematics

In Exercises 3–8, use the box-and-whisker plot to find the given measure.

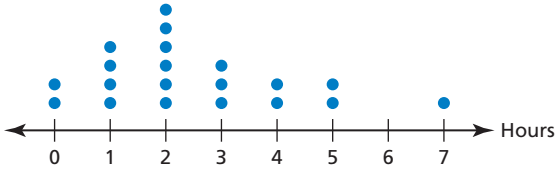


3. least value
4. greatest value
5. third quartile
6. first quartile
7. median
8. range

In Exercises 9–12, make a box-and-whisker plot that represents the data. *(See Example 1.)*

9. Hours of television watched: 0, 3, 4, 5, 2, 4, 6, 5
10. Cat lengths (in inches): 16, 18, 20, 25, 17, 22, 23, 21
11. Elevations (in feet): −2, 0, 5, −4, 1, −3, 2, 0, 2, −3, 6
12. MP3 player prices (in dollars): 124, 95, 105, 110, 95, 124, 300, 190, 114

13. **ANALYZING DATA** The dot plot represents the numbers of hours students spent studying for an exam. Make a box-and-whisker plot that represents the data.

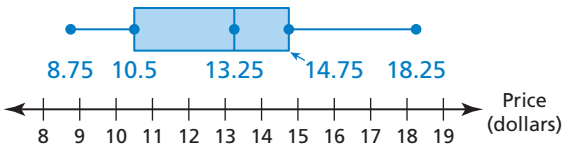


14. **ANALYZING DATA** The stem-and-leaf plot represents the lengths (in inches) of the fish caught on a fishing trip. Make a box-and-whisker plot that represents the data.

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

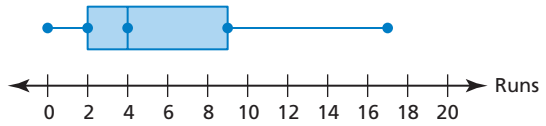
Key: 1|0 = 10 inches

15. **ANALYZING DATA** The box-and-whisker plot represents the prices (in dollars) of the entrées at a restaurant. *(See Example 2.)*



   a. Find and interpret the range of the data.
   b. Describe the distribution of the data.
   c. Find and interpret the interquartile range of the data.
   d. Are the data more spread out below Q1 or above Q3? Explain.

16. ANALYZING DATA A baseball player scores 101 runs in a season. The box-and-whisker plot represents the numbers of runs the player scores against different opposing teams.

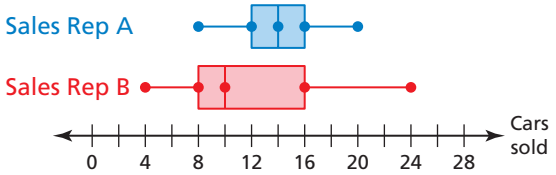


a. Find and interpret the range and interquartile range of the data.

b. Describe the distribution of the data.

c. Are the data more spread out between Q1 and Q2 or between Q2 and Q3? Explain.

17. ANALYZING DATA The double box-and-whisker plot represents the monthly car sales for a year for two sales representatives. *(See Example 3.)*

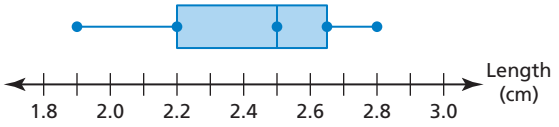


a. Identify the shape of each distribution.

b. Which representative's sales are more spread out? Explain.

c. Which representative had the single worst sales month during the year? Explain.

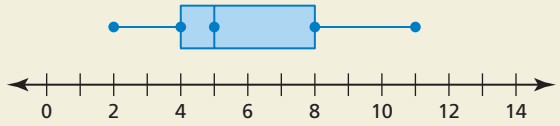
18. ERROR ANALYSIS Describe and correct the error in describing the box-and-whisker plot.



✗ The distribution is skewed left. So, most of the data are on the left side of the plot.

19. WRITING Given the numbers 36 and 12, identify which number is the range and which number is the interquartile range of a data set. Explain.

20. HOW DO YOU SEE IT? The box-and-whisker plot represents a data set. Determine whether each statement is always true. Explain your reasoning.



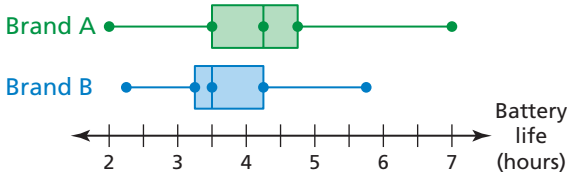
a. The data set contains the value 11.

b. The data set contains the value 6.

c. The distribution is skewed right.

d. The mean of the data is 5.

21. ANALYZING DATA The double box-and-whisker plot represents the battery lives (in hours) of two brands of cell phones.



a. Identify the shape of each distribution.

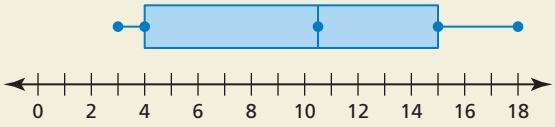
b. What is the range of the upper 75% of each brand?

c. Compare the interquartile ranges of the two data sets.

d. Which brand do you think has a greater standard deviation? Explain.

e. You need a cell phone that has a battery life of more than 3.5 hours most of the time. Which brand should you buy? Explain.

22. THOUGHT PROVOKING Create a data set that can be represented by the box-and-whisker plot shown. Justify your answer.



23. CRITICAL THINKING Two data sets have the same median, the same interquartile range, and the same range. Is it possible for the box-and-whisker plots of the data sets to be different? Justify your answer.

Maintaining Mathematical Proficiency

Reviewing what you learned in previous grades and lessons

Use zeros to graph the function. *(Section 8.5)*

24. $f(x) = -2(x + 9)(x - 3)$

25. $y = 3(x - 5)(x + 5)$

26. $y = 4x^2 - 16x - 48$

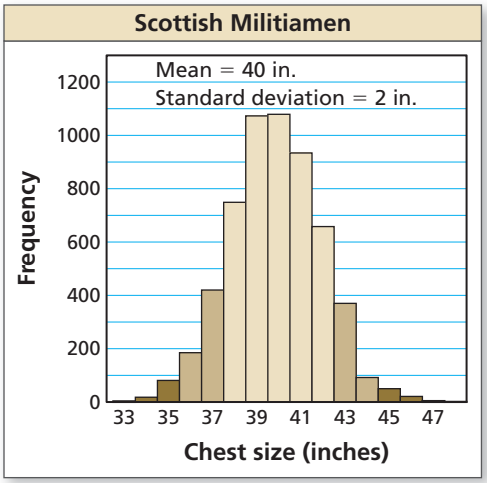
27. $h(x) = -x^2 + 5x + 14$

11.3 Shapes of Distributions

Essential Question How can you use a histogram to characterize the basic shape of a distribution?

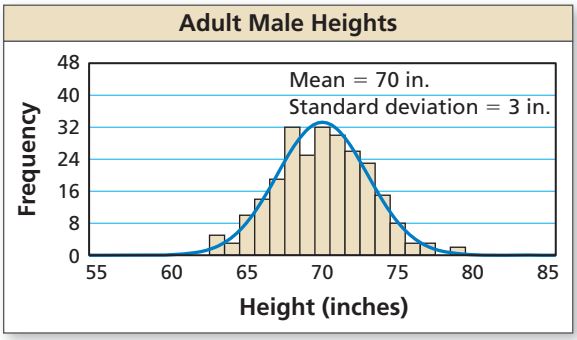
EXPLORATION 1 Analyzing a Famous Symmetric Distribution

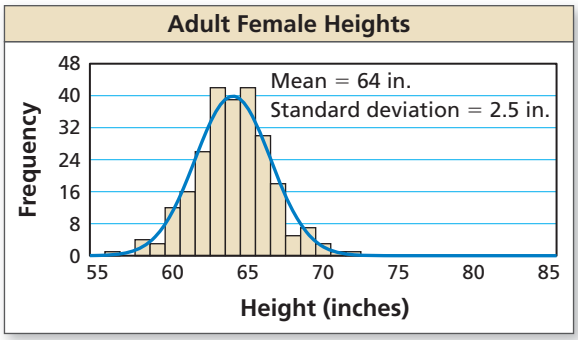
Work with a partner. A famous data set was collected in Scotland in the mid-1800s. It contains the chest sizes, measured in inches, of 5738 men in the Scottish Militia. Estimate the percent of the chest sizes that lie within (a) 1 standard deviation of the mean, (b) 2 standard deviations of the mean, and (c) 3 standard deviations of the mean. Explain your reasoning.



EXPLORATION 2 Comparing Two Symmetric Distributions

Work with a partner. The graphs show the distributions of the heights of 250 adult American males and 250 adult American females.





a. Which data set has a smaller standard deviation? Explain what this means in the context of the problem.

b. Estimate the percent of male heights between 67 inches and 73 inches.

ATTENDING TO PRECISION
To be proficient in math, you need to express numerical answers with a level of precision appropriate for the problem's context.

Communicate Your Answer

3. How can you use a histogram to characterize the basic shape of a distribution?

4. All three distributions in Explorations 1 and 2 are roughly symmetric. The histograms are called "bell-shaped."

a. What are the characteristics of a symmetric distribution?

b. Why is a symmetric distribution called "bell-shaped?"

c. Give two other real-life examples of symmetric distributions.

11.3 Lesson

What You Will Learn

- Describe the shapes of data distributions.
- Use the shapes of data distributions to choose appropriate measures.
- Compare data distributions.

Core Vocabulary

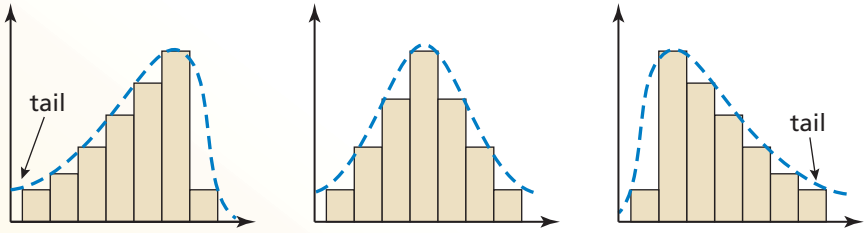
Previous
histogram
frequency table

Describing the Shapes of Data Distributions

Recall that a histogram is a bar graph that shows the frequency of data values in intervals of the same size. A histogram is another useful data display that shows the shape of a distribution.

Core Concept

Symmetric and Skewed Distributions



Skewed left

- The "tail" of the graph extends to the left.
- Most of the data are on the right.

Symmetric

- The data on the right of the distribution are approximately a mirror image of the data on the left of the distribution.

Skewed right

- The "tail" of the graph extends to the right.
- Most of the data are on the left.

STUDY TIP

If all the bars of a histogram are about the same height, then the distribution is a *flat*, or *uniform*, distribution. A uniform distribution is also symmetric.

Number of tickets sold	Frequency
1–8	5
9–16	9
17–24	16
25–32	25
33–40	20
41–48	8
49–56	7

Number of pounds	Frequency
1–10	7
11–20	8
21–30	10
31–40	16
41–50	34
51–60	15

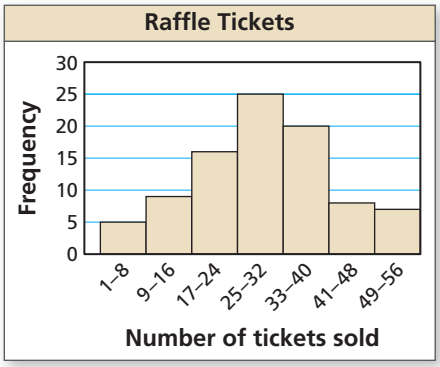
EXAMPLE 1 Describing the Shape of a Distribution

The frequency table shows the numbers of raffle tickets sold by students in your grade. Display the data in a histogram. Describe the shape of the distribution.

SOLUTION

Step 1 Draw and label the axes.

Step 2 Draw a bar to represent the frequency of each interval.



The data on the right of the distribution are approximately a mirror image of the data on the left of the distribution.

So, the distribution is symmetric.

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

1. The frequency table shows the numbers of pounds of aluminum cans collected by classes for a fundraiser. Display the data in a histogram. Describe the shape of the distribution.

Choosing Appropriate Measures

Use the shape of a distribution to choose the most appropriate measure of center and measure of variation to describe the data set.

STUDY TIP

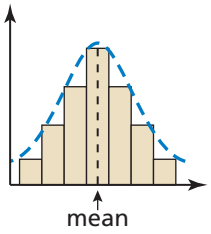
When a distribution is symmetric, the mean and median are about the same. When a distribution is skewed, the mean will be in the direction in which the distribution is skewed while the median will be less affected.

Core Concept

Choosing Appropriate Measures

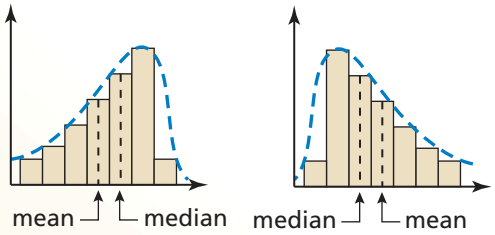
When a data distribution is symmetric,

- use the mean to describe the center and
- use the standard deviation to describe the variation.



When a data distribution is skewed,

- use the median to describe the center and
- use the five-number summary to describe the variation.



EXAMPLE 2 Choosing Appropriate Measures

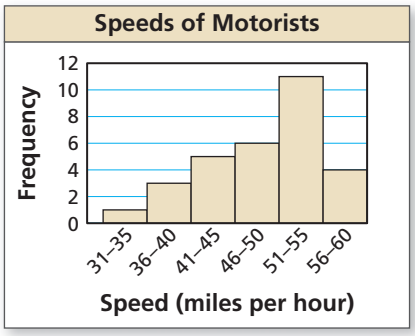
Speeds (mi/h)		
32	44	39
53	38	48
56	41	42
50	50	55
55	45	49
51	53	52
54	60	55
52	50	52
55	40	60
45	58	47

A police officer measures the speeds (in miles per hour) of 30 motorists. The results are shown in the table at the left. (a) Display the data in a histogram using six intervals beginning with 31–35. (b) Which measures of center and variation best represent the data? (c) The speed limit is 45 miles per hour. How would you interpret these results?

SOLUTION

a. Make a frequency table using the described intervals. Then use the frequency table to make a histogram.

Speed (mi/h)	Frequency
31–35	1
36–40	3
41–45	5
46–50	6
51–55	11
56–60	4



b. Because most of the data are on the right and the tail of the graph extends to the left, the distribution is skewed left. So, use the median to describe the center and the five-number summary to describe the variation.

c. Using the frequency table and the histogram, you can see that most of the speeds are more than 45 miles per hour. So, most of the motorists were speeding.

Email Attachments Sent				
74	105	98	68	64
85	75	60	48	51
65	55	58	45	38
64	52	65	30	70
72	5	45	77	83
42	25	95	16	120

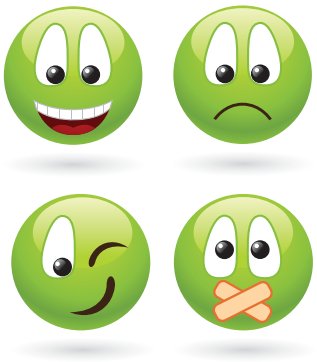
Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

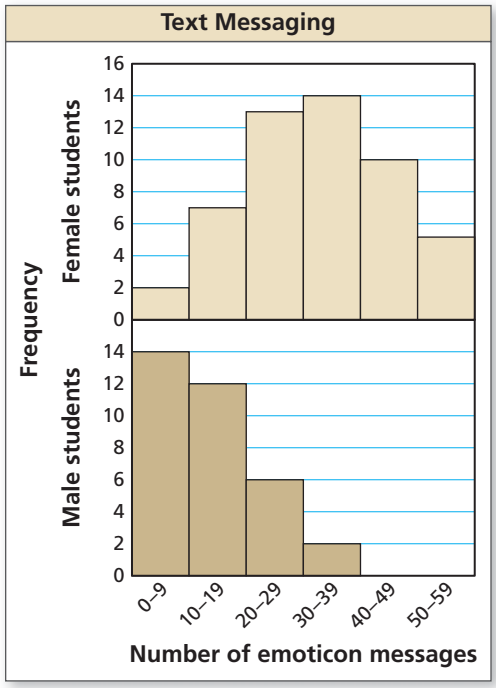
2. You record the numbers of email attachments sent by 30 employees of a company in 1 week. Your results are shown in the table. (a) Display the data in a histogram using six intervals beginning with 1–20. (b) Which measures of center and variation best represent the data? Explain.

Comparing Data Distributions

EXAMPLE 3 Comparing Data Distributions



Emoticons are graphic symbols that represent facial expressions. They are used to convey a person's mood in a text message. The double histogram shows the distributions of emoticon messages sent by a group of female students and a group of male students during 1 week. Compare the distributions using their shapes and appropriate measures of center and variation.



SOLUTION

Because the data on the right of the distribution for the female students are approximately a mirror image of the data on the left of the distribution, the distribution is symmetric. So, the mean and standard deviation best represent the distribution for female students.

Because most of the data are on the left of the distribution for the male students and the tail of the graph extends to the right, the distribution is skewed right. So, the median and five-number summary best represent the distribution for male students.

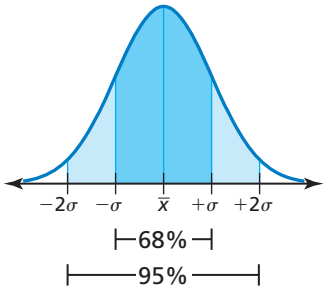
The mean of the female data set is probably in the 30–39 interval, while the median of the male data set is in the 10–19 interval. So, a typical female student is much more likely to use emoticons than a typical male student.

The data for the female students is more variable than the data for the male students. This means that the use of emoticons tends to differ more from one female student to the next.

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

3. Compare the distributions using their shapes and appropriate measures of center and variation.



Many real-life data sets have distributions that are bell-shaped and approximately symmetric about the mean. In a future course, you will study this type of distribution in detail. For now, the following rules can help you see how valuable the standard deviation can be as a measure of variation.

- About 68% of the data lie within 1 standard deviation of the mean.
- About 95% of the data lie within 2 standard deviations of the mean.
- Data values that are more than 2 standard deviations from the mean are considered unusual.

Because the data are symmetric, you can deduce that 34% of the data lie within 1 standard deviation to the left of the mean, and 34% of the data lie within 1 standard deviation to the right of the mean.

EXAMPLE 4 Comparing Data Distributions

The table shows the results of a survey that asked men and women how many pairs of shoes they own.

	Men	Women
Survey size	35	40
Minimum	2	5
Maximum	17	24
1st Quartile	5	12
Median	7	14
3rd Quartile	10	17
Mean	8	14
Standard deviation	3	4

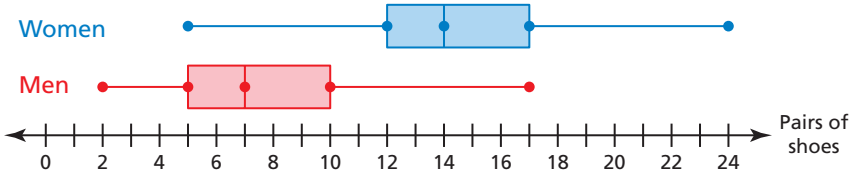
a. Make a double box-and-whisker plot that represents the data. Describe the shape of each distribution.

b. Compare the number of pairs of shoes owned by men to the number of pairs of shoes owned by women.

c. About how many of the women surveyed would you expect to own between 10 and 18 pairs of shoes?

SOLUTION

a.



▶ The distribution for men is skewed right, and the distribution for women is symmetric.

b. The centers and spreads of the two data sets are quite different from each other. The mean for women is twice the median for men, and there is more variability in the number of pairs of shoes owned by women.

c. Assuming the symmetric distribution is bell-shaped, you know about 68% of the data lie within 1 standard deviation of the mean. Because the mean is 14 and the standard deviation is 4, the interval from 10 to 18 represents about 68% of the data. So, you would expect about $0.68 \cdot 40 \approx 27$ of the women surveyed to own between 10 and 18 pairs of shoes.

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

4. Why is the mean greater than the median for the men?

5. If 50 more women are surveyed, about how many more would you expect to own between 10 and 18 pairs of shoes?

11.3 Exercises

Dynamic Solutions available at *BigIdeasMath.com*

Vocabulary and Core Concept Check

1. **VOCABULARY** Describe how data are distributed in a symmetric distribution, a distribution that is skewed left, and a distribution that is skewed right.

2. **WRITING** How does the shape of a distribution help you decide which measures of center and variation best describe the data?

Monitoring Progress and Modeling with Mathematics

3. **DESCRIBING DISTRIBUTIONS** The frequency table shows the numbers of hours that students volunteer per month. Display the data in a histogram. Describe the shape of the distribution. *(See Example 1.)*

Number of volunteer hours	1–2	3–4	5–6	7–8	9–10	11–12	13–14
Frequency	1	5	12	20	15	7	2

4. **DESCRIBING DISTRIBUTIONS** The frequency table shows the results of a survey that asked people how many hours they spend online per week. Display the data in a histogram. Describe the shape of the distribution.

Hours online	Frequency
0–3	5
4–7	7
8–11	12
12–15	14
16–19	26
20–23	45
24–27	33

In Exercises 5 and 6, describe the shape of the distribution of the data. Explain your reasoning.

5.

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

Key: 3|1 = 31

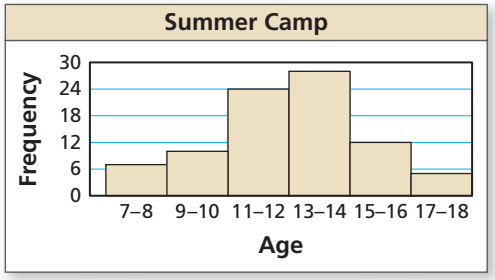
6.

Stem	Leaf
5	0 0 1
6	3 6 7 9
7	1 4 5 8 9
8	2 4 5 5 7
9	4 6 8 9
10	1 3 4

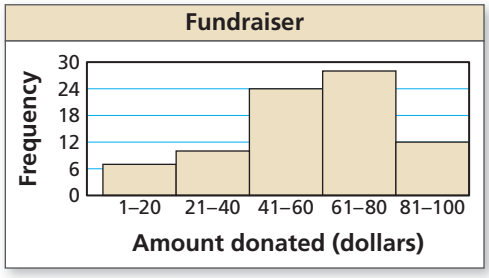
Key: 6|3 = 63

In Exercises 7 and 8, determine which measures of center and variation best represent the data. Explain your reasoning.

7.



8.



9. **ANALYZING DATA** The table shows the last 24 ATM withdrawals at a bank. *(See Example 2.)*

ATM Withdrawals (dollars)		
120	100	70
60	40	80
150	80	50
120	60	175
30	50	50
60	200	30
100	150	110
70	40	100

a. Display the data in a histogram using seven intervals beginning with 26–50.

b. Which measures of center and variation best represent the data? Explain.

c. The bank charges a fee for any ATM withdrawal less than $150. How would you interpret the data?

10. ANALYZING DATA Measuring an IQ is an inexact science. However, IQ scores have been around for years in an attempt to measure human intelligence. The table shows some of the greatest known IQ scores.

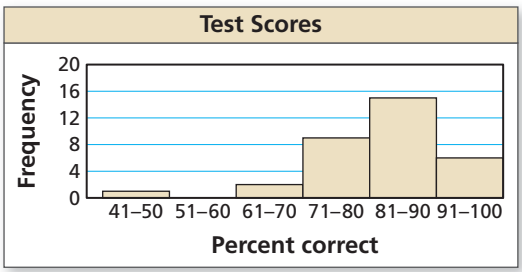
IQ Scores		
170	190	180
160	180	210
154	170	180
195	230	160
170	186	180
225	190	170

a. Display the data in a histogram using five intervals beginning with 151–166.

b. Which measures of center and variation best represent the data? Explain.

c. The distribution of IQ scores for the human population is symmetric. What happens to the shape of the distribution in part (a) as you include more and more IQ scores from the human population in the data set?

ERROR ANALYSIS In Exercises 11 and 12, describe and correct the error in the statements about the data displayed in the histogram.



11.

✗ Most of the data are on the right. So, the distribution is skewed right.

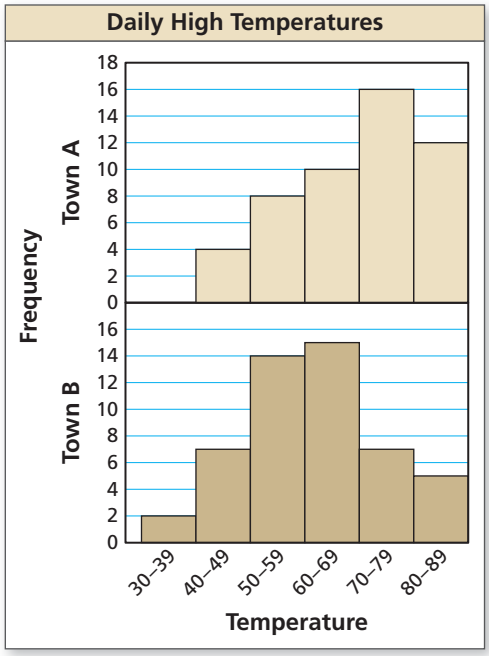
12.

✗ Because the distribution is skewed, use the standard deviation to describe the variation of the data.

13. USING TOOLS For a large data set, would you use a stem-and-leaf plot or a histogram to show the distribution of the data? Explain.

14. REASONING For a symmetric distribution, why is the mean used to describe the center and the standard deviation used to describe the variation? For a skewed distribution, why is the median used to describe the center and the five-number summary used to describe the variation?

15. COMPARING DATA SETS The double histogram shows the distributions of daily high temperatures for two towns over a 50-day period. Compare the distributions using their shapes and appropriate measures of center and variation. *(See Example 3.)*



16. COMPARING DATA SETS The frequency tables show the numbers of entrées in certain price ranges (in dollars) at two different restaurants. Display the data in a double histogram. Compare the distributions using their shapes and appropriate measures of center and variation.

Restaurant A	
Price range	Frequency
8–10	5
11–13	9
14–16	12
17–19	4
20–22	3
23–25	0

Restaurant B	
Price range	Frequency
8–10	0
11–13	2
14–16	5
17–19	7
20–22	8
23–25	6

17. OPEN-ENDED Describe a real-life data set that has a distribution that is skewed right.

18. OPEN-ENDED Describe a real-life data set that has a distribution that is skewed left.

19. COMPARING DATA SETS The table shows the results of a survey that asked freshmen and sophomores how many songs they have downloaded on their MP3 players. *(See Example 4.)*

	Freshmen	Sophomores
Survey size	45	54
Minimum	250	360
Maximum	2150	2400
1st Quartile	800	780
Median	1200	2000
3rd Quartile	1600	2200
Mean	1150	1650
Standard deviation	420	480

a. Make a double box-and-whisker plot that represents the data. Describe the shape of each distribution.

b. Compare the number of songs downloaded by freshmen to the number of songs downloaded by sophomores.

c. About how many of the freshmen surveyed would you expect to have between 730 and 1570 songs downloaded on their MP3 players?

d. If you survey 100 more freshmen, about how many would you expect to have downloaded between 310 and 1990 songs on their MP3 players?

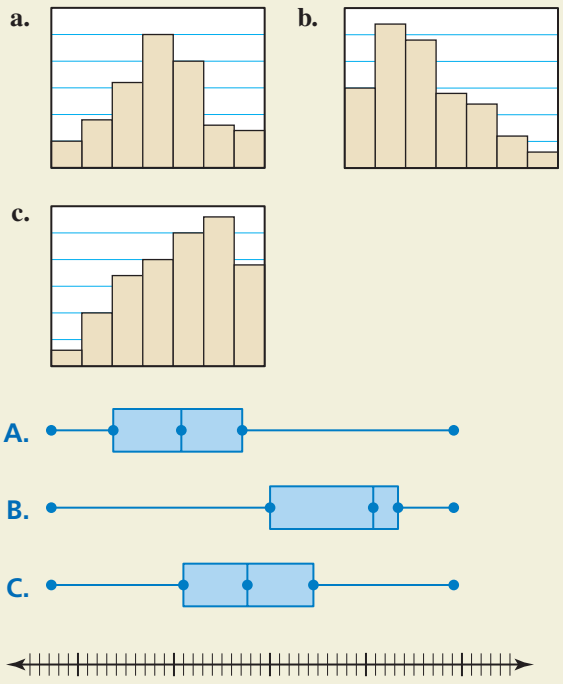
20. COMPARING DATA SETS You conduct the same survey as in Exercise 19 but use a different group of freshmen. The results are as follows.
Survey size: 60; minimum: 200; maximum: 2400; 1st quartile: 640; median: 1670; 3rd quartile: 2150; mean: 1480; standard deviation: 500

a. Compare the number of songs downloaded by this group of freshmen to the number of songs downloaded by sophomores.

b. Why is the median greater than the mean for this group of freshmen?

21. REASONING A data set has a symmetric distribution. Every value in the data set is doubled. Describe the shape of the new distribution. Are the measures of center and variation affected? Explain.

22. HOW DO YOU SEE IT? Match the distribution with the corresponding box-and-whisker plot.



23. REASONING You record the following waiting times at a restaurant.

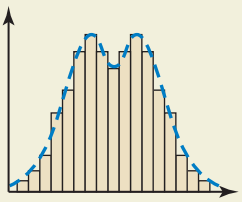
Waiting Times (minutes)									
26	38	15	8	22	42	25	20	17	18
40	35	24	31	42	29	25	0	30	13

a. Display the data in a histogram using five intervals beginning with 0–9. Describe the shape of the distribution.

b. Display the data in a histogram using 10 intervals beginning with 0–4. What happens when the number of intervals is increased?

c. Which histogram best represents the data? Explain your reasoning.

24. THOUGHT PROVOKING
The shape of a *bimodal* distribution is shown. Describe a real-life example of a bimodal distribution.



Maintaining Mathematical Proficiency

Reviewing what you learned in previous grades and lessons

Find the domain of the function. *(Section 10.1)*

25. $f(x) = \sqrt{x + 6}$

26. $f(x) = \sqrt{2x}$

27. $f(x) = \frac{1}{4}\sqrt{x - 7}$

11.1–11.3 What Did You Learn?

Core Vocabulary

measure of center, *p. 586*
mean, *p. 586*
median, *p. 586*
mode, *p. 586*
outlier, *p. 587*
measure of variation, *p. 587*
range, *p. 587*
standard deviation, *p. 588*
data transformation, *p. 589*
box-and-whisker plot, *p. 594*
quartile, *p. 594*
five-number summary, *p. 594*
interquartile range, *p. 595*

Core Concepts

Section 11.1

Measures of Center, *p. 586*
Measures of Variation, *p. 587*
Data Transformations Using Addition, *p. 589*
Data Transformations Using Multiplication, *p. 589*

Section 11.2

Box-and-Whisker Plot, *p. 594*
Shapes of Box-and-Whisker Plots, *p. 596*

Section 11.3

Symmetric and Skewed Distributions, *p. 600*
Choosing Appropriate Measures, *p. 601*

Mathematical Practices

1. Exercises 15 and 16 on page 590 are similar. For each data set, is the outlier much greater than or much less than the rest of the data values? Compare how the outliers affect the means. Explain why this makes sense.
2. In Exercise 18 on page 605, provide a possible reason for why the distribution is skewed left.

Study Skills

Studying for Finals

- Form a study group of three or four students several weeks before the final exam.
- Find out what material you must know for the final exam, even if your teacher has not yet covered it.
- Ask for a practice final exam or create one yourself and have your teacher look at it.
- Have each group member take the practice final exam.
- Decide when the group is going to meet and what you will cover during each session.
- During the sessions, make sure you stay on track.

11.1–11.3 Quiz

Find the mean, median, and mode of the data set. Which measure of center best represents the data? Explain. *(Section 11.1)*

1.

Hours Spent on Project		
$3\frac{1}{2}$	5	$2\frac{1}{2}$
3	$3\frac{1}{2}$	$\frac{1}{2}$

2.

Waterfall Height (feet)		
1000	1267	1328
1200	1180	1000
2568	1191	1100

Find the range and standard deviation of each data set. Then compare your results. *(Section 11.1)*

3. Absent students during a week of school
Female: 6, 2, 4, 3, 4
Male: 5, 3, 6, 6, 9

4. Numbers of points scored
Juniors: 19, 15, 20, 10, 14, 21, 18, 15
Seniors: 22, 19, 29, 32, 15, 26, 30, 19

Make a box-and-whisker plot that represents the data. *(Section 11.2)*

5. Ages of family members:
60, 15, 25, 20, 55, 70, 40, 30

6. Minutes of violin practice:
20, 50, 60, 40, 40, 30, 60, 40, 50, 20, 20, 35

7. Display the data in a histogram. Describe the shape of the distribution. *(Section 11.3)*

Quiz score	0–2	3–5	6–8	9–11	12–14
Frequency	1	3	6	16	4

8. The table shows the prices of eight mountain bikes in a sporting goods store. *(Section 11.1 and Section 11.2)*

Price (dollars)	98	119	95	211	130	98	100	125

a. Find the mean, median, mode, range, and standard deviation of the prices.

b. Identify the outlier. How does the outlier affect the mean, median, and mode?

c. Make a box-and-whisker plot that represents the data. Find and interpret the interquartile range of the data. Identify the shape of the distribution.

d. Find the mean, median, mode, range, and standard deviation of the prices when the store offers a 5% discount on all mountain bikes.

9. The table shows the times of 20 presentations. *(Section 11.3)*

Time (minutes)			
9	7	10	12
10	11	8	10
10	17	11	5
9	10	4	12
6	14	8	10

a. Display the data in a histogram using five intervals beginning with 3–5.

b. Which measures of center and variation best represent the data? Explain.

c. The presentations are supposed to be 10 minutes long. How would you interpret these results?

11.4 Two-Way Tables

Essential Question

How can you read and make a two-way table?

EXPLORATION 1 Reading a Two-Way Table

Work with a partner. You are the manager of a sports shop. The two-way tables show the numbers of soccer T-shirts in stock at your shop at the beginning and end of the selling season. (a) Complete the totals for the rows and columns in each table. (b) How would you alter the number of T-shirts you order for next season? Explain your reasoning.

Beginning of season

Color	T-Shirt Size: S	M	L	XL	XXL	Total
blue/white	5	6	7	6	5	
blue/gold	5	6	7	6	5	
red/white	5	6	7	6	5	
black/white	5	6	7	6	5	
black/gold	5	6	7	6	5	
Total						145

End of season

Color	T-Shirt Size: S	M	L	XL	XXL	Total
blue/white	5	4	1	0	2	
blue/gold	3	6	5	2	0	
red/white	4	2	4	1	3	
black/white	3	4	1	2	1	
black/gold	5	2	3	0	2	
Total						

MODELING WITH MATHEMATICS

To be proficient in math, you need to identify important quantities and map their relationships using tools such as graphs and two-way tables.

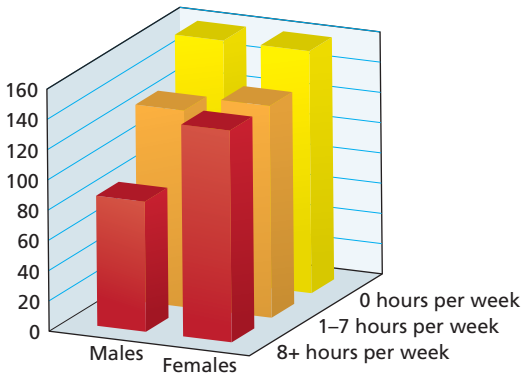
EXPLORATION 2 Making a Two-Way Table

Work with a partner. The three-dimensional bar graph shows the numbers of hours students work at part-time jobs.

a. Make a two-way table showing the data. Use estimation to find the entries in your table.

b. Write two observations that summarize the data in your table.

Part-Time Jobs of Students at a High School



Communicate Your Answer

3. How can you read and make a two-way table?

11.4 Lesson

Core Vocabulary

two-way table, *p. 610*
joint frequency, *p. 610*
marginal frequency, *p. 610*
joint relative frequency, *p. 611*
marginal relative frequency, *p. 611*
conditional relative frequency, *p. 612*

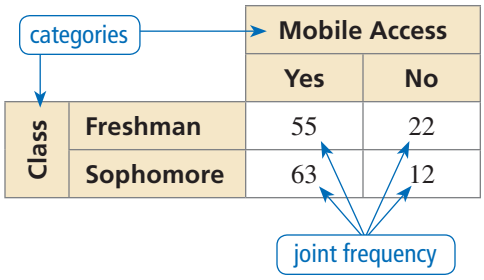
What You Will Learn

- Find and interpret marginal frequencies.
- Make two-way tables.
- Find relative and conditional relative frequencies.
- Use two-way tables to recognize associations in data.

Finding and Interpreting Marginal Frequencies

A **two-way table** is a frequency table that displays data collected from one source that belong to two different categories. One category of data is represented by rows, and the other is represented by columns. For instance, the two-way table below shows the results of a survey that asked freshmen and sophomores whether they access the Internet using a mobile device, such as a smartphone.

The two categories of data are *class* and *mobile access*. Class is further divided into *freshman* and *sophomore*, and mobile access is further divided into *yes* and *no*.



Each entry in the table is called a **joint frequency**. The sums of the rows and columns in a two-way table are called **marginal frequencies**.

REMEMBER
The frequency of an event is the number of times the event occurs.

EXAMPLE 1 Finding and Interpreting Marginal Frequencies

Find and interpret the marginal frequencies for the two-way table above.

SOLUTION

Create a new column and a new row for the marginal frequencies. Then add the entries in each row and column.

STUDY TIP
The sum of the "total" row should be equal to the sum of the "total" column. Place this sum of the marginal frequencies at the bottom right of your two-way table.

		Mobile Access		
		Yes	No	Total
Class	Freshman	55	22	77
	Sophomore	63	12	75
	Total	118	34	152

77 freshmen responded.
75 sophomores responded.
152 students were surveyed.
118 students access the Internet using a mobile device.
34 students do not access the Internet using a mobile device.

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

1. You conduct a technology survey to publish on your school's website. You survey students in the school cafeteria about the technological devices they own. The results are shown in the two-way table. Find and interpret the marginal frequencies.

		Tablet Computer	
		Yes	No
Cell Phone	Yes	34	124
	No	18	67

Making Two-Way Tables

EXAMPLE 2 Making a Two-Way Table

You conduct a survey that asks 286 students in your freshman class whether they play a sport or a musical instrument. One hundred eighteen of the students play a sport, and 64 of those students play an instrument. Ninety-three of the students do not play a sport or an instrument. Organize the results in a two-way table. Include the marginal frequencies.

SOLUTION

Step 1 Determine the two categories for the table: *sport* and *instrument*.

Step 2 Use the given joint and marginal frequencies to fill in parts of the table.

		Instrument		
		Yes	No	Total
Sport	Yes	64		118
	No		93	
	Total			286

Step 3 Use reasoning to find the missing joint and marginal frequencies. For instance, you can conclude that there are $286 - 118 = 168$ students who do not play a sport, and $118 - 64 = 54$ students who play a sport but do not play an instrument.

		Instrument		
		Yes	No	Total
Sport	Yes	64	54	118
	No	75	93	168
	Total	139	147	286

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

2. You survey students about whether they are getting a summer job. Seventy-five males respond, with 18 of them responding "no." Fifty-seven females respond, with 45 of them responding "yes." Organize the results in a two-way table. Include the marginal frequencies.

Finding Relative and Conditional Relative Frequencies

You can display entries of a two-way table as frequency counts (as in Examples 1 and 2) or as *relative frequencies*.

Core Concept

Relative Frequencies

A **joint relative frequency** is the ratio of a frequency that is not in the "total" row or the "total" column to the total number of values or observations.

A **marginal relative frequency** is the sum of the joint relative frequencies in a row or a column.

When finding relative frequencies in a two-way table, you can use the corresponding decimals or percents.

EXAMPLE 3 Finding Relative Frequencies

		Major in Medical Field	
		Yes	No
Class	Junior	124	219
	Senior	101	236

The two-way table shows the results of a survey that asked college-bound high school students whether they plan to major in a medical field. Make a two-way table that shows the joint and marginal relative frequencies.

SOLUTION

There are $124 + 219 + 101 + 236 = 680$ students in the survey. To find the joint relative frequencies, divide each frequency by 680. Then find the sum of each row and each column to find the marginal relative frequencies.

STUDY TIP

The sum of the marginal relative frequencies in the "total" row and the "total" column should each be equal to 1.

		Major in Medical Field		
		Yes	No	Total
Class	Junior	$\frac{124}{680} \approx 0.18$	$\frac{219}{680} \approx 0.32$	0.50
	Senior	$\frac{101}{680} \approx 0.15$	$\frac{236}{680} \approx 0.35$	0.50
	Total	0.33	0.67	1

About 50% of the students are juniors.

About 35% of the students are seniors and are not planning to major in a medical field.

Core Concept

Conditional Relative Frequencies

A **conditional relative frequency** is the ratio of a joint relative frequency to the marginal relative frequency. You can find a conditional relative frequency using a row total or a column total of a two-way table.

EXAMPLE 4 Finding Conditional Relative Frequencies

Use the survey results in Example 3 to make a two-way table that shows the conditional relative frequencies based on the column totals.

SOLUTION

Use the marginal relative frequency of each *column* to calculate the conditional relative frequencies.

STUDY TIP

When you use column totals, the sum of the conditional relative frequencies for each column should be equal to 1.

		Major in Medical Field	
		Yes	No
Class	Junior	$\frac{0.18}{0.33} \approx 0.55$	$\frac{0.32}{0.67} \approx 0.48$
	Senior	$\frac{0.15}{0.33} \approx 0.45$	$\frac{0.35}{0.67} \approx 0.52$

Given that a student is not planning to major in a medical field, the conditional relative frequency that he or she is a junior is about 48%.

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

3. Use the survey results in Monitoring Progress Question 2 to make a two-way table that shows the joint and marginal relative frequencies. What percent of students are not getting a summer job?

4. Use the survey results in Example 3 to make a two-way table that shows the conditional relative frequencies based on the row totals. Given that a student is a senior, what is the conditional relative frequency that he or she is planning to major in a medical field?

Recognizing Associations in Data

EXAMPLE 5 Recognizing Associations in Data

You survey students and find that 40% exercise regularly, 35% eat fruits and vegetables each day, and 52% do not exercise and do not eat fruits and vegetables each day. Is there an association between exercising regularly and eating fruits and vegetables each day?

SOLUTION

Use the given information to make a two-way table. Use reasoning to find the missing joint and marginal relative frequencies.

		Exercise Regularly		
		Yes	No	Total
Eats Fruit/ Vegetables	Yes	27%	8%	35%
	No	13%	52%	65%
	Total	40%	60%	100%

Use conditional relative frequencies based on the column totals to determine whether there is an association. Of the students who exercise regularly, 67.5% eat fruits and vegetables each day. Of the students who do not exercise regularly, only about 13% eat fruits and vegetables each day. It appears that students who exercise regularly are more likely to eat more fruits and vegetables than students who do not exercise regularly.

		Exercise Regularly	
		Yes	No
Eats Fruit/ Vegetables	Yes	$\frac{0.27}{0.4} = 0.675$	$\frac{0.08}{0.6} \approx 0.133$
	No	$\frac{0.13}{0.4} = 0.325$	$\frac{0.52}{0.6} \approx 0.867$

▶ So, there is an association between exercising regularly and eating fruits and vegetables each day.

You can also find the conditional relative frequencies by dividing each joint frequency by its corresponding column total or row total.

EXAMPLE 6 Recognizing Associations in Data

		Age			
		12–13	14–15	16–17	18–19
Share a Computer	Yes	40	47	42	22
	No	10	25	36	34

The two-way table shows the results of a survey that asked students whether they share a computer at home with other family members. Is there an association between age and sharing a computer?

SOLUTION

Use conditional relative frequencies based on column totals to determine whether there is an association. Based on this sample, 80% of students ages 12–13 share a computer and only about 39% of students ages 18–19 share a computer.

▶ The table shows that as age increases, students are less likely to share a computer with other family members. So, there is an association.

		Age			
		12–13	14–15	16–17	18–19
Share a Computer	Yes	$\frac{40}{50} = 0.8$	$\frac{47}{72} \approx 0.65$	$\frac{42}{78} \approx 0.54$	$\frac{22}{56} \approx 0.39$
	No	$\frac{10}{50} \approx 0.2$	$\frac{25}{72} \approx 0.35$	$\frac{36}{78} \approx 0.46$	$\frac{34}{56} \approx 0.61$

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

5. Using the results of the survey in Monitoring Progress Question 1, is there an association between owning a tablet computer and owning a cell phone? Explain your reasoning.

11.4 Exercises

Dynamic Solutions available at *BigIdeasMath.com*

Vocabulary and Core Concept Check

1. **COMPLETE THE SENTENCE** Each entry in a two-way table is called a(n) _____.
2. **WRITING** When is it appropriate to use a two-way table to organize data?
3. **VOCABULARY** Explain the relationship between joint relative frequencies, marginal relative frequencies, and conditional relative frequencies.
4. **WRITING** Describe two ways you can find conditional relative frequencies.

Monitoring Progress and Modeling with Mathematics

You conduct a survey that asks 346 students whether they buy lunch at school. In Exercises 5–8, use the results of the survey shown in the two-way table.

		Buy Lunch at School	
		Yes	No
Class	Freshman	92	86
	Sophomore	116	52

5. How many freshmen were surveyed?
6. How many sophomores were surveyed?
7. How many students buy lunch at school?
8. How many students do not buy lunch at school?

In Exercises 9 and 10, find and interpret the marginal frequencies. *(See Example 1.)*

9.

		Set Academic Goals	
		Yes	No
Gender	Male	64	168
	Female	54	142

10.

		Cat	
		Yes	No
Dog	Yes	104	208
	No	186	98

11. **USING TWO-WAY TABLES** You conduct a survey that asks students whether they plan to participate in school spirit week. The results are shown in the two-way table. Find and interpret the marginal frequencies.

		Participate in Spirit Week		
		Yes	No	Undecided
Class	Freshman	112	56	54
	Sophomore	92	68	32

12. **USING TWO-WAY TABLES** You conduct a survey that asks college-bound high school seniors about the type of degree they plan to receive. The results are shown in the two-way table. Find and interpret the marginal frequencies.

		Type of Degree		
		Associate's	Bachelor's	Master's
Gender	Male	58	126	42
	Female	62	118	48

USING STRUCTURE In Exercises 13 and 14, complete the two-way table.

13.

		Traveled on an Airplane		
		Yes	No	Total
Class	Freshman		62	
	Sophomore	184		
	Total	274		352

14.

		Plan to Attend School Dance		
		Yes	No	Total
Gender	Male	38		
	Female		24	112
	Total			196

15. MAKING TWO-WAY TABLES You conduct a survey that asks 245 students in your school whether they have taken a Spanish or a French class. One hundred nine of the students have taken a Spanish class, and 45 of those students have taken a French class. Eighty-two of the students have not taken a Spanish or a French class. Organize the results in a two-way table. Include the marginal frequencies.
(See Example 2.)

16. MAKING TWO-WAY TABLES A car dealership has 98 cars on its lot. Fifty-five of the cars are new. Of the new cars, 36 are domestic cars. There are 15 used foreign cars on the lot. Organize this information in a two-way table. Include the marginal frequencies.

In Exercises 17 and 18, make a two-way table that shows the joint and marginal relative frequencies.
(See Example 3.)

17.

		Exercise Preference	
		Aerobic	Anaerobic
Gender	Male	88	104
	Female	96	62

18.

		Meat	
		Turkey	Ham
Bread	White	452	146
	Wheat	328	422

19. USING TWO-WAY TABLES Refer to Exercise 17. What percent of students prefer aerobic exercise? What percent of students are males who prefer anaerobic exercise?

20. USING TWO-WAY TABLES Refer to Exercise 18. What percent of the sandwiches are on wheat bread? What percent of the sandwiches are turkey on white bread?

ERROR ANALYSIS In Exercises 21 and 22, describe and correct the error in using the two-way table.

		Participate in Fundraiser	
		Yes	No
Class	Freshman	187	85
	Sophomore	123	93

21. ✗ *One hundred eighty-seven freshmen responded to the survey.*

22. ✗ *The two-way table shows the joint relative frequencies.*

		Participate in Fundraiser	
		Yes	No
Class	Freshman	$\frac{187}{272} \approx 0.69$	$\frac{85}{272} \approx 0.31$
	Sophomore	$\frac{123}{216} \approx 0.57$	$\frac{93}{216} \approx 0.43$

23. USING TWO-WAY TABLES A company is hosting an event for its employees to celebrate the end of the year. It asks the employees whether they prefer a lunch event or a dinner event. It also asks whether they prefer a catered event or a potluck. The results are shown in the two-way table. Make a two-way table that shows the conditional relative frequencies based on the row totals. Given that an employee prefers a lunch event, what is the conditional relative frequency that he or she prefers a catered event?
(See Example 4.)

		Menu	
		Potluck	Catered
Meal	Lunch	36	48
	Dinner	44	72

24. USING TWO-WAY TABLES The two-way table shows the results of a survey that asked students about their preference for a new school mascot. Make a two-way table that shows the conditional relative frequencies based on the column totals. Given that a student prefers a hawk as a mascot, what is the conditional relative frequency that he or she prefers a cartoon mascot?

		Type		
		Tiger	Hawk	Dragon
Style	Realistic	67	74	51
	Cartoon	58	18	24

25. ANALYZING TWO-WAY TABLES You survey college-bound seniors and find that 85% plan to live on campus, 35% plan to have a car while at college, and 5% plan to live off campus and not have a car. Is there an association between living on campus and having a car at college? Explain. *(See Example 5.)*

26. ANALYZING TWO-WAY TABLES You survey students and find that 70% watch sports on TV, 48% participate in a sport, and 16% do not watch sports on TV or participate in a sport. Is there an association between participating in a sport and watching sports on TV? Explain.

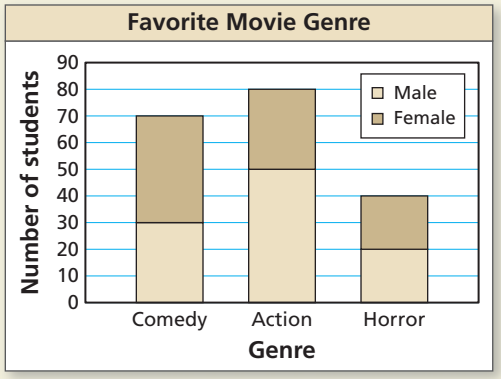
27. ANALYZING TWO-WAY TABLES The two-way table shows the results of a survey that asked adults whether they participate in recreational skiing. Is there an association between age and recreational skiing? *(See Example 6.)*

		Age				
		21–30	31–40	41–50	51–60	61–70
Ski	Yes	87	93	68	37	20
	No	165	195	148	117	125

28. ANALYZING TWO-WAY TABLES Refer to Exercise 12. Is there an association between gender and type of degree? Explain.

29. WRITING Compare Venn diagrams and two-way tables.

30. HOW DO YOU SEE IT? The graph shows the results of a survey that asked students about their favorite movie genre.



a. Display the given information in a two-way table.

b. Which of the data displays do you prefer? Explain.

31. PROBLEM SOLVING A box office sells 1809 tickets to a play, 800 of which are for the main floor. The tickets consist of $2x + y$ adult tickets on the main floor, $x - 40$ child tickets on the main floor, $x + 2y$ adult tickets in the balcony, and $3x - y - 80$ child tickets in the balcony.

a. Organize this information in a two-way table.

b. Find the values of x and y.

c. What percent of tickets are adult tickets?

d. What percent of child tickets are balcony tickets?

32. THOUGHT PROVOKING Compare "one-way tables" and "two-way tables." Is it possible to have a "three-way table?" If so, give an example of a three-way table.

Maintaining Mathematical Proficiency

Reviewing what you learned in previous grades and lessons

Tell whether the table of values represents a *linear*, an *exponential*, or a *quadratic* function. *(Section 8.6)*

33.

x	0	1	2	3	4
y	144	24	4	$\frac{2}{3}$	$\frac{1}{9}$

34.

x	-1	0	1	2	3
y	3	0	-1	0	3

11.5 Choosing a Data Display

Essential Question How can you display data in a way that helps you make decisions?

EXPLORATION 1 Displaying Data

Work with a partner. Analyze the data and then create a display that best represents the data. Explain your choice of data display.

a. A group of schools in New England participated in a 2-month study and reported 3962 animals found dead along roads.

birds: 307	mammals: 2746	amphibians: 145
reptiles: 75	unknown: 689	

b. The data below show the numbers of black bears killed on a state's roads from 1993 to 2012.

1993: 30	2000: 47	2007: 99
1994: 37	2001: 49	2008: 129
1995: 46	2002: 61	2009: 111
1996: 33	2003: 74	2010: 127
1997: 43	2004: 88	2011: 141
1998: 35	2005: 82	2012: 135
1999: 43	2006: 109	

c. A 1-week study along a 4-mile section of road found the following weights (in pounds) of raccoons that had been killed by vehicles.

13.4	14.8	17.0	12.9	21.3	21.5	16.8	14.8
15.2	18.7	18.6	17.2	18.5	9.4	19.4	15.7
14.5	9.5	25.4	21.5	17.3	19.1	11.0	12.4
20.4	13.6	17.5	18.5	21.5	14.0	13.9	19.0

d. A yearlong study by volunteers in California reported the following numbers of animals killed by motor vehicles.

raccoons: 1693	gray squirrels: 715
skunks: 1372	cottontail rabbits: 629
ground squirrels: 845	barn owls: 486
opossum: 763	jackrabbits: 466
deer: 761	gopher snakes: 363

USING TOOLS STRATEGICALLY

To be proficient in math, you need to identify relevant external mathematical resources.

Communicate Your Answer

2. How can you display data in a way that helps you make decisions?

3. Use the Internet or some other reference to find examples of the following types of data displays.

bar graph	circle graph	scatter plot
stem-and-leaf plot	pictograph	line graph
box-and-whisker plot	histogram	dot plot

11.5 Lesson

Core Vocabulary

qualitative (categorical) data, *p. 618*
quantitative data, *p. 618*
misleading graph, *p. 620*

What You Will Learn

- Classify data as quantitative or qualitative.
- Choose and create appropriate data displays.
- Analyze misleading graphs.

Classifying Data

Data sets can consist of two types of data: *qualitative* or *quantitative*.

Core Concept

Types of Data

Qualitative data, or **categorical data**, consist of labels or nonnumerical entries that can be separated into different categories. When using qualitative data, operations such as adding or finding a mean do not make sense.

Quantitative data consist of numbers that represent counts or measurements.

STUDY TIP

Just because a frequency count can be shown for a data set does not make it quantitative. A frequency count can be shown for both qualitative and quantitative data.

EXAMPLE 1 Classifying Data

Tell whether the data are *qualitative* or *quantitative*.

a. prices of used cars at a dealership
b. jersey numbers on a basketball team
c. lengths of songs played at a concert
d. zodiac signs of students in your class

SOLUTION

a. Prices are numerical entries. So, the data are quantitative.

b. Jersey numbers are numerical, but they are labels. It does not make sense to compare them, and you cannot measure them. So, the data are qualitative.

c. Song lengths are numerical measurements. So, the data are quantitative.

d. Zodiac signs are nonnumerical entries that can be separated into different categories. So, the data are qualitative.

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

Tell whether the data are *qualitative* or *quantitative*. Explain your reasoning.

1. telephone numbers in a directory
2. ages of patients at a hospital
3. lengths of videos on a website
4. types of flowers at a florist

Qualitative and quantitative data can be collected from the same data source, as shown below. You can use these types of data together to obtain a more accurate description of a population.

Data Source	Quantitative Data	Qualitative Data
a student	How much do you earn per hour at your job? $10.50	What is your occupation? painter
a house	How many square feet of living space is in the house? 2500 ft^2	In what city is the house located? Chicago

Choosing and Creating Appropriate Data Displays

As shown on page 584, you have learned a variety of ways to display data sets graphically. Choosing an appropriate data display can depend on whether the data are qualitative or quantitative.

EXAMPLE 2 Choosing and Creating a Data Display

Analyze the data and then create a display that best represents the data. Explain your reasoning.

a.

Eye Color Survey	
Color	**Number of students**
brown	63
blue	37
hazel	25
green	10
gray	3
amber	2

b.

Speeds of Vehicles (mi/h)			
Interstate A		**Interstate B**	
65	67	67	72
68	71	70	78
72	70	65	71
68	65	71	80
65	68	84	81
75	82	77	79
68	59	69	70
62	68	66	69
75	80	73	75
77	75	84	79

SOLUTION

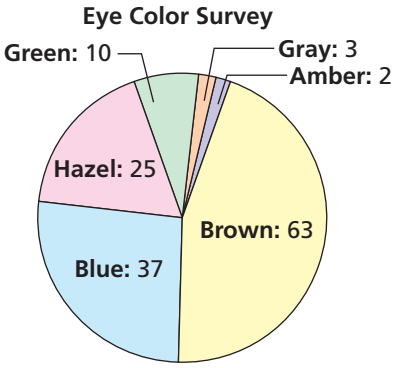
a. A circle graph is one appropriate way to display this qualitative data. It shows data as parts of a whole.

Step 1 Find the angle measure for each section of the circle graph by multiplying the fraction of students who have each eye color by 360°. Notice that there are $63 + 37 + 25 + 10 + 3 + 2 = 140$ students in the survey.

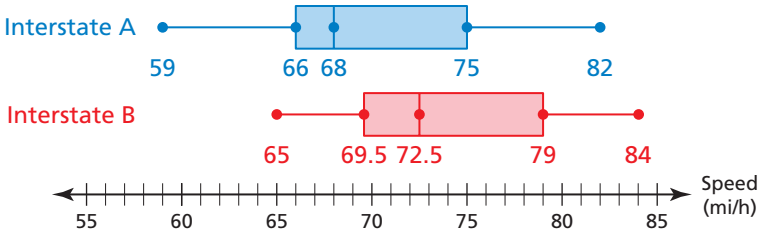
Brown: $\frac{63}{140} \cdot 360° \approx 162°$ **Blue:** $\frac{37}{140} \cdot 360° \approx 95°$ **Hazel:** $\frac{25}{140} \cdot 360° \approx 64°$

Green: $\frac{10}{140} \cdot 360° \approx 26°$ **Gray:** $\frac{3}{140} \cdot 360° \approx 8°$ **Amber:** $\frac{2}{140} \cdot 360° \approx 5°$

Step 2 Use a protractor to draw the angle measures found in Step 1 on a circle. Then label each section and title the circle graph, as shown.



b. A double box-and-whisker plot is one appropriate way to display this quantitative data. Use the five-number summary of each data set to create a double box-and-whisker plot.



Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

5. Display the data in Example 2(a) in another way.
6. Display the data in Example 2(b) in another way.

Analyzing Misleading Graphs

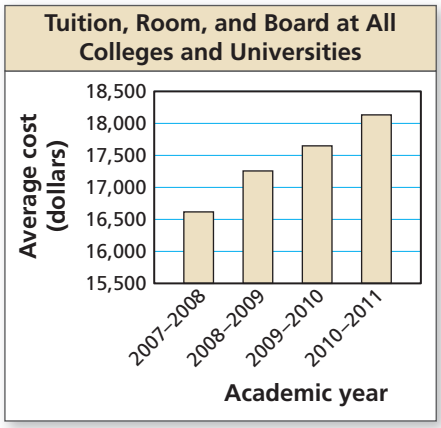
Just as there are several ways to display data accurately using graphs, there are several ways to display data that are misleading. A **misleading graph** is a statistical graph that is not drawn appropriately. This may occur when the creator of a graph wants to give viewers the impression that results are better than they actually are. Below are some questions you can ask yourself when analyzing a statistical graph that will help you recognize when a graph is trying to deceive or mislead.

- Does the graph have a title?
- Are the numbers of the scale evenly spaced?
- Does the scale begin at zero? If not, is there a break?
- Does the graph need a key?
- Are all the axes or sections of the graph labeled?
- Are all the components of the graph, such as the bars, the same size?

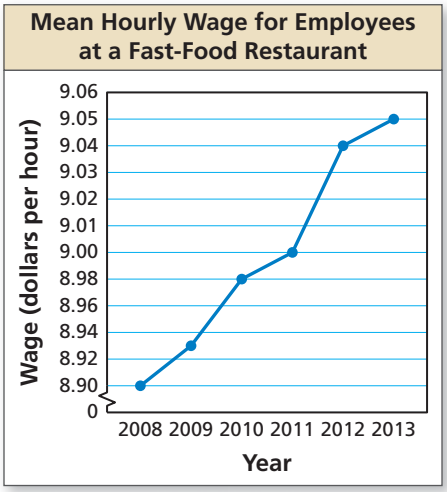
EXAMPLE 3 Analyzing Misleading Graphs

Describe how each graph is misleading. Then explain how someone might misinterpret the graph.

a.



b.



SOLUTION

a. The scale on the vertical axis of the graph starts at \$15,500 and does not have a break. This makes it appear that the average cost increased rapidly for the years given.

Someone might believe that the average cost more than doubled from 2007 to 2011, when actually, it increased by only about \$1500.

b. The scale on the vertical axis has very small increments that are not equal.

Someone might believe that the greatest increase in the mean hourly wage occurred from 2011 to 2012, when the greatest increase actually occurred from 2009 to 2010.

Monitoring Progress

Help in English and Spanish at *BigIdeasMath.com*

7. Redraw the graphs in Example 3 so they are not misleading.

11.5 Exercises

Dynamic Solutions available at *BigIdeasMath.com*

Vocabulary and Core Concept Check

1. **OPEN-ENDED** Describe two ways that a line graph can be misleading.

2. **WHICH ONE DOESN'T BELONG?** Which data set does *not* belong with the other three? Explain your reasoning.

ages of people attending a concert	heights of skyscrapers in a city
populations of counties in a state	breeds of dogs at a pet store

Monitoring Progress and Modeling with Mathematics

In Exercises 3–8, tell whether the data are *qualitative* or *quantitative*. Explain your reasoning. *(See Example 1.)*

3. brands of cars in a parking lot

4. weights of bears at a zoo

5. budgets of feature films

6. file formats of documents on a computer

7. shoe sizes of students in your class

8. street addresses in a phone book

In Exercises 9–12, choose an appropriate data display for the situation. Explain your reasoning.

9. the number of students in a marching band each year

10. a comparison of students' grades (out of 100) in two different classes

11. the favorite sports of students in your class

12. the distribution of teachers by age

In Exercises 13–16, analyze the data and then create a display that best represents the data. Explain your reasoning. *(See Example 2.)*

13.

Ages of World Cup Winners	
2010 Men's World Cup Winner (Spain)	**2011 Women's World Cup Winner (Japan)**
29 24 23 30 32 26 28 30 26 23 32 28 22 28 24 21 27 22 25 21 24 24 27	36 27 24 20 27 23 29 26 25 32 27 27 22 25 24 23 24 28 20 18 24

14.

Average Precipitation (inches)			
January	1.1	July	4.0
February	1.5	August	4.4
March	2.2	September	4.2
April	3.7	October	3.5
May	5.1	November	2.1
June	5.5	December	1.8

15.

Grades (out of 100) on a Test				
96	74	97	80	62
84	88	53	77	75
89	81	52	85	63
87	95	59	83	100

16.

Colors of Cars that Drive by Your House			
white	25	green	3
red	12	silver/gray	27
yellow	1	blue	6
black	21	brown/biege	5

17. **DISPLAYING DATA** Display the data in Exercise 13 in another way.

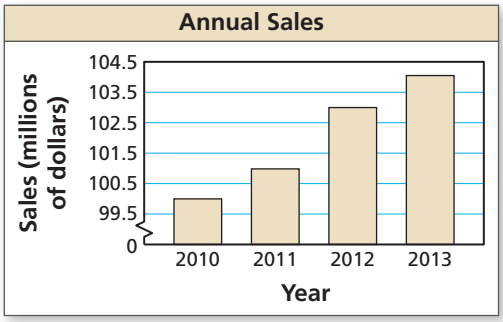
18. **DISPLAYING DATA** Display the data in Exercise 14 in another way.

19. **DISPLAYING DATA** Display the data in Exercise 15 in another way.

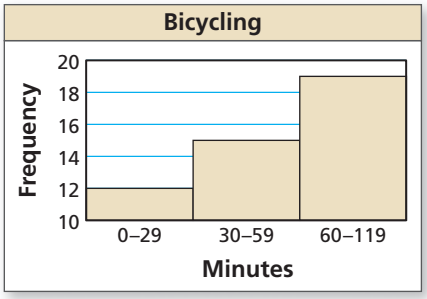
20. **DISPLAYING DATA** Display the data in Exercise 16 in another way.

In Exercises 21–24, describe how the graph is misleading. Then explain how someone might misinterpret the graph. *(See Example 3.)*

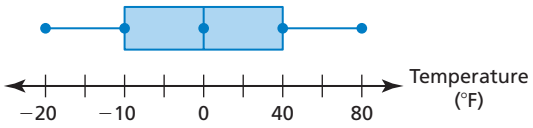
21.



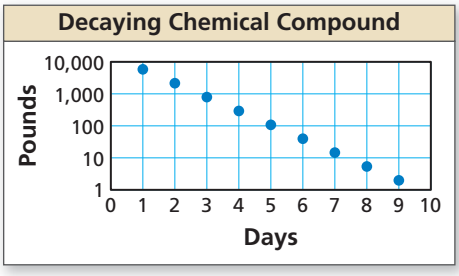
22.



23.



24.

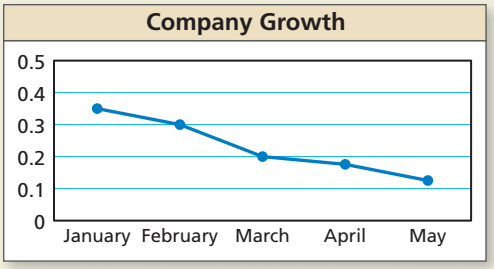


25. **DISPLAYING DATA** Redraw the graph in Exercise 21 so it is not misleading.

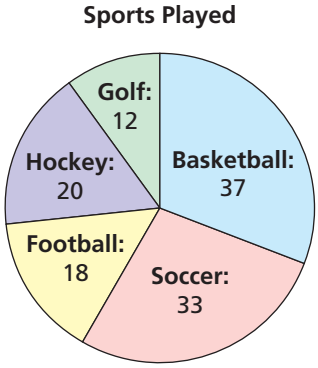
26. **DISPLAYING DATA** Redraw the graph in Exercise 22 so it is not misleading.

27. **MAKING AN ARGUMENT** A data set gives the ages of voters for a city election. Classmate A says the data should be displayed in a bar graph, while Classmate B says the data would be better displayed in a histogram. Who is correct? Explain.

28. **HOW DO YOU SEE IT?** The manager of a company sees the graph shown and concludes that the company is experiencing a decline. What is missing from the graph? Explain why the manager may be mistaken.



29. **REASONING** A survey asked 100 students about the sports they play. The results are shown in the circle graph.



a. Explain why the graph is misleading.

b. What type of data display would be more appropriate for the data? Explain.

30. **THOUGHT PROVOKING** Use a spreadsheet program to create a type of data display that is not used in this section.

31. **REASONING** What type of data display shows the mode of a data set?

Maintaining Mathematical Proficiency

Reviewing what you learned in previous grades and lessons

Determine whether the relation is a function. Explain. *(Section 3.1)*

32. $(-5, -1), (-6, 0), (-5, 1), (-2, 2), (3, 3)$

33. $(0, 1), (4, 0), (8, 1), (12, 2), (16, 3)$

11.4–11.5 What Did You Learn?

Core Vocabulary

two-way table, *p. 610*
joint frequency, *p. 610*
marginal frequency, *p. 610*
joint relative frequency, *p. 611*
marginal relative frequency, *p. 611*
conditional relative frequency, *p. 612*
qualitative (categorical) data, *p. 618*
quantitative data, *p. 618*
misleading graph, *p. 620*

Core Concepts

Section 11.4

Joint and Marginal Frequencies, *p. 610*
Making Two-Way Tables, *p. 611*
Relative Frequencies, *p. 611*
Conditional Relative Frequencies, *p. 612*
Recognizing Associations in Data, *p. 613*

Section 11.5

Types of Data, *p. 618*
Choosing and Creating Appropriate Data Displays, *p. 619*
Analyzing Misleading Graphs, *p. 620*

Mathematical Practices

1. Consider the data given in the two-way table for Exercises 5–8 on page 614. Your sophomore friend responded to the survey. Is your friend more likely to have responded "yes" or "no" to buying a lunch? Explain.
2. Use your answer to Exercise 28 on page 622 to explain why it is important for a company manager to see accurate graphs.

Performance Task

College Student Study Time

Data from a small survey at a state university could provide insight into the amount of study time necessary to be successful in college. Based on the information you find when you organize the data, what advice should you give your peers? How will you support your conclusions?

To explore the answers to these questions and more, go to ***BigIdeasMath.com***.

11 Chapter Review

Dynamic Solutions available at *BigIdeasMath.com*

11.1 Measures of Center and Variation *(pp. 585–592)*

The table shows the number of miles you ran each day for 10 days. Find the mean, median, and mode of the distances.

Miles Run	
3.5	4.1
4.0	4.3
4.4	4.5
3.9	2.0
4.3	5.0

Mean $\bar{x} = \dfrac{3.5 + 4.0 + 4.4 + 3.9 + 4.3 + 4.1 + 4.3 + 4.5 + 2.0 + 5.0}{10} = 4$

Median 2.0, 3.5, 3.9, 4.0, 4.1, 4.3, 4.3, 4.4, 4.5, 5.0 Order the data.

$\dfrac{8.4}{2} = 4.2$ Mean of two middle values

Mode 2.0, 3.5, 3.9, 4.0, 4.1, 4.3, 4.3, 4.4, 4.5, 5.0 4.3 occurs most often.

▶ The mean is 4 miles, the median is 4.2 miles, and the mode is 4.3 miles.

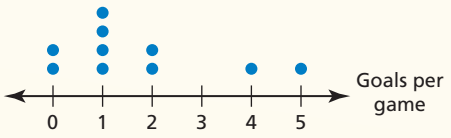
1. Use the data in the example above. You run 4.0 miles on Day 11. How does this additional value affect the mean, median, and mode? Explain.
2. Use the data in the example above. You run 10.0 miles on Day 11. How does this additional value affect the mean, median, and mode? Explain.

Find the mean, median, and mode of the data.

3.

Ski Resort Temperatures (°F)		
11	3	3
0	−9	−2
10	10	10

4.



Find the range and standard deviation of each data set. Then compare your results.

5.

Bowling Scores			
Player A		Player B	
205	190	228	205
185	200	172	181
210	219	154	240
174	203	235	235
194	230	168	192

6.

Tablet Prices			
Store A		Store B	
$140	$180	$225	$310
$200	$250	$260	$190
$150	$190	$190	$285
$250	$160	$160	$240

Find the values of the measures shown after the given transformation.

Mean: 109 Median: 104 Mode: 96 Range: 45 Standard deviation: 3.6

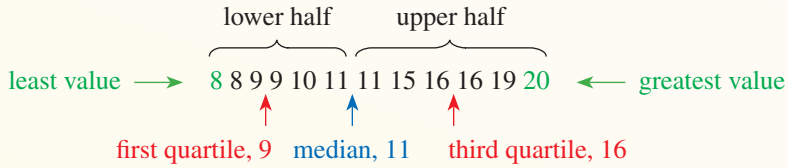
7. Each value in the data set increases by 25.
8. Each value in the data set is multiplied by 0.6.

11.2 Box-and-Whisker Plots *(pp. 593–598)*

Make a box-and-whisker plot that represents the weights (in pounds) of pumpkins sold at a market.

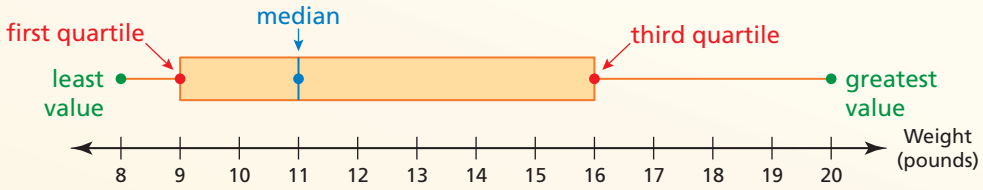
16, 20, 11, 15, 10, 8, 8, 19, 11, 9, 16, 9

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



Step 2 Draw a number line that includes the least and greatest values. Graph points above the number line for the five-number summary.

Step 3 Draw a box using Q_1 and Q_3. Draw a line through the median. Draw whiskers from the box to the least and greatest values.



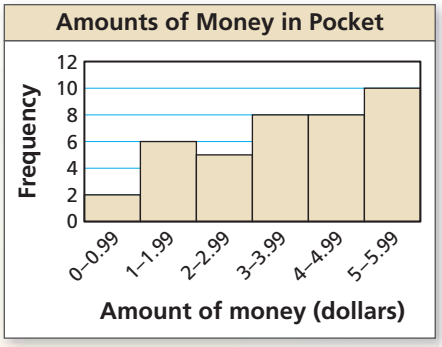
Make a box-and-whisker plot that represents the data. Identify the shape of the distribution.

9. Ages of volunteers at a hospital:
14, 17, 20, 16, 17, 14, 21, 18, 22

10. Masses (in kilograms) of lions:
120, 230, 180, 210, 200, 200, 230, 160

11.3 Shapes of Distributions *(pp. 599–606)*

The histogram shows the amounts of money a group of adults have in their pockets. Describe the shape of the distribution. Which measures of center and variation best represent the data?



▶ The distribution is skewed left. So, use the median to describe the center and the five-number summary to describe the variation.

Amount	Frequency
0–0.99	9
1–1.99	10
2–2.99	9
3–3.99	7
4–4.99	4
5–5.99	1

11. The frequency table shows the amounts (in dollars) of money the students in a class have in their pockets.

a. Display the data in a histogram. Describe the shape of the distribution.

b. Which measures of center and variation best represent the data?

c. Compare this distribution with the distribution shown above using their shapes and appropriate measures of center and variation.

11.4 Two-Way Tables *(pp. 609–616)*

You conduct a survey that asks 130 students about whether they have an after-school job. Sixty males respond, 38 of which have a job. Twenty-six females do not have a job. Organize the results in a two-way table. Find and interpret the marginal frequencies.

		After-School Job		
		Yes	No	Total
Gender	Males	38	22	60 ← 60 males responded.
	Females	44	26	70 ← 70 females responded.
	Total	82	48	130 ← 130 students were surveyed.

82 students have a job. 48 students do not have a job.

12. The two-way table shows the results of a survey that asked shoppers at a mall about whether they like the new food court.

a. Make a two-way table that shows the joint and marginal relative frequencies.

b. Make a two-way table that shows the conditional relative frequencies based on the column totals.

		Food Court	
		Like	Dislike
Shoppers	Adults	21	79
	Teenagers	96	4

11.5 Choosing a Data Display *(pp. 617–622)*

Analyze the data and then create a display that best represents the data.

Ages of U.S. Presidents at Inauguration										
57	61	57	57	58	57	61	54	68	51	49
64	50	48	65	52	56	46	54	49	51	47
55	55	54	42	51	56	55	51	54	51	60
62	43	55	56	61	52	69	64	46	54	47

A stem-and-leaf plot is one appropriate way to display this quantitative data. It orders numerical data and shows how they are distributed.

Ages of U.S. Presidents at Inauguration

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

Key: 5|0 = 50

13. Analyze the data in the table at the right and then create a display that best represents the data. Explain your reasoning.

Perfect Attendance	
Class	Number of students
freshman	84
sophomore	42
junior	67
senior	31

Tell whether the data are *qualitative* or *quantitative*. Explain.

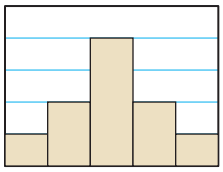
14. heights of the members of a basketball team

15. grade level of students in an elementary school

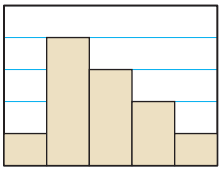
11 Chapter Test

Describe the shape of the data distribution. Then determine which measures of center and variation best represent the data.

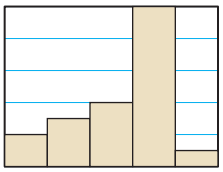
1.



2.



3.



4. Determine whether each statement is *always*, *sometimes*, or *never* true. Explain your reasoning.

a. The sum of the marginal relative frequencies in the "total" row and the "total" column of a two-way table should each be equal to 1.

b. In a box-and-whisker plot, the length of the box to the left of the median and the length of the box to the right of the median are equal.

c. Qualitative data are numerical.

5. Find the mean, median, mode, range, and standard deviation of the prices.

Prices of Shirts at a Clothing Store			
$15.50	$18.90	$10.60	$12.25
$7.80	$23.50	$9.75	$21.70

6. Repeat Exercise 5 when all the shirts in the clothing store are 20% off.

7. Which data display best represents the data, a histogram or a stem-and-leaf plot? Explain.

15, 21, 18, 10, 12, 11, 17, 18, 16, 12, 20, 12, 17, 16

8. The tables show the battery lives (in hours) of two brands of laptops.

a. Make a double box-and-whisker plot that represents the data.

b. Identify the shape of each distribution.

c. Which brand's battery lives are more spread out? Explain.

d. Compare the distributions using their shapes and appropriate measures of center and variation.

Brand A	
20.75	18.5
13.5	16.25
8.5	13.5
14.5	15.5
11.5	16.75

Brand B	
10.5	12.5
9.5	10.25
9.0	9.75
8.5	8.5
9.0	7.0

Preferred method of exercise	Number of students
walking	20
jogging	28
biking	17
swimming	11
lifting weights	10
dancing	14

9. The table shows the results of a survey that asked students their preferred method of exercise. Analyze the data and then create a display that best represents the data. Explain your reasoning.

10. You conduct a survey that asks 271 students in your class whether they are attending the class field trip. One hundred twenty-one males respond, 92 of which are attending the field trip. Thirty-one females are not attending the field trip.

a. Organize the results in a two-way table. Find and interpret the marginal frequencies.

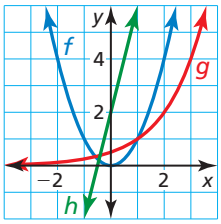
b. What percent of females are attending the class field trip?

11 Cumulative Assessment

1. You ask all the students in your grade whether they have a cell phone. The results are shown in the two-way table. Your friend claims that a greater percent of males in your grade have cell phones than females. Do you support your friend's claim? Justify your answer.

		Cell Phones	
		Yes	No
Gender	Male	27	12
	Female	31	17

2. Use the graphs of the functions to answer each question.



a. Are there any values of x greater than 0 where $f(x) > h(x)$? Explain.

b. Are there any values of x greater than 1 where $g(x) > f(x)$? Explain.

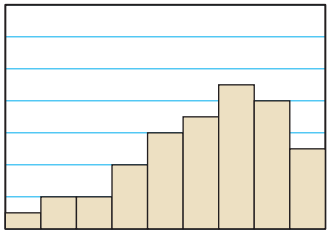
c. Are there any values of x greater than 0 where $g(x) > h(x)$? Explain.

3. Classify the shape of each distribution as *symmetric*, *skewed left*, or *skewed right*.

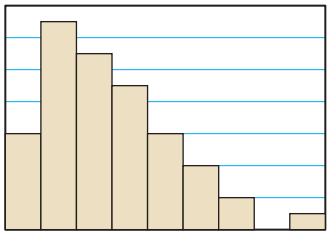
a.

b.

c.



d.



4. Complete the equation so that the solutions of the system of equations are $(-2, 4)$ and $(1, -5)$.

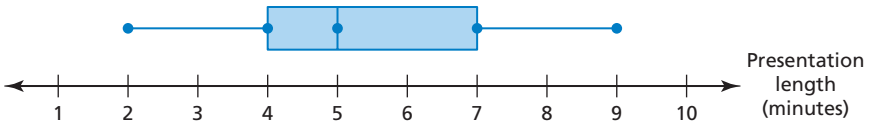
$$y = \square x + \square$$

$$y = 2x^2 - x - 6$$

5. Pair each function with its inverse.

$y = -3x^2, x \geq 0$	$y = -x + 7$	$y = 2x - 4$	$y = \sqrt{-\frac{1}{3}x}$
$y = \frac{1}{2}x + 2$	$y = x^2 - 5, x \geq 0$	$y = \sqrt{x + 5}$	$y = -x + 7$

6. The box-and-whisker plot represents the lengths (in minutes) of project presentations at a science fair. Find the interquartile range of the data. What does this represent in the context of the situation?



Ⓐ 7; The middle half of the presentation lengths vary by no more than 7 minutes.

Ⓑ 3; The presentation lengths vary by no more than 3 minutes.

Ⓒ 3; The middle half of the presentation lengths vary by no more than 3 minutes.

Ⓓ 7; The presentation lengths vary by no more than 7 minutes.

7. Scores in a video game can be between 0 and 100. Use the data set shown to fill in a value for x so that each statement is true.

a. When $x =$ ____, the mean of the scores is 45.5.

b. When $x =$ ____, the median of the scores is 47.

c. When $x =$ ____, the mode of the scores is 63.

d. When $x =$ ____, the range of the scores is 71.

Video Game Scores	
36	28
48	x
42	57
63	52

8. Select all the numbers that are in the range of the function shown.

$$y = \begin{cases} x^2 + 4x + 7, & \text{if } x \le -1 \\ \frac{1}{2}x + 2, & \text{if } x > -1 \end{cases}$$

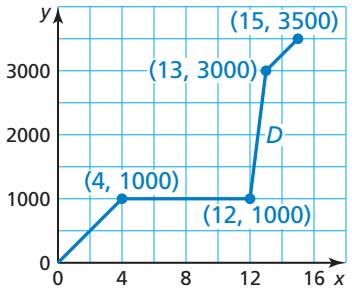
0	$\frac{1}{2}$	1	$1\frac{1}{2}$	2	$2\frac{1}{2}$	3	$3\frac{1}{2}$	4

9. A traveler walks and takes a shuttle bus to get to a terminal of an airport. The function $y = D(x)$ represents the traveler's distance (in feet) after x minutes.

a. Estimate and interpret $D(2)$.

b. Use the graph to find the solution of the equation $D(x) = 3500$. Explain the meaning of the solution.

c. How long does the traveler wait for the shuttle bus?

d. How far does the traveler ride on the shuttle bus?

e. What is the total distance that the traveler walks before and after riding the shuttle bus?