EXAMPLE 6    **Winning a Lottery Twice: Fraud?**

The likelihood of winning a major state lottery is very small. Yet some people have won state lotteries more than once. Does this indicate fraud? Explain your reasoning.

> ### Lottery Luck Strikes Twice in Three Months
>
> ST. LOUIS - You are as likely to win the lottery as you are to be struck by lightning. Ernest Pullen, a Missouri resident, has won the lottery twice in a matter of three months. In June, he won his first cash prize for $1 million. In September, he won his second cash prize for $2 million. One lottery official called it an amazing coincidence.

### SOLUTION

This problem is difficult because there is not enough information to analyze it completely. Even so, consider a simplified example.

- There is a drawing every day with one winner.

- The same million people play each day.

- The probability of winning with one ticket is one-millionth.

With these assumptions, the mathematical strategy for the solution is the same as that in the "Birthday Problem" in Example 3 on page 386.

<div align="center">365 factors</div>

Probability that all
365 tickets have     $= \left(\dfrac{1,000,000}{1,000,000}\right)\left(\dfrac{999,999}{1,000,000}\right)\left(\dfrac{999,998}{1,000,000}\right) \cdots \left(\dfrac{999,636}{1,000,000}\right)$
*different* owners

$\approx 0.936$

Probability that at
least 2 tickets have $= 1 - 0.936 = 0.064$
the *same* owner

So, there is about a 6.4% chance that during 1 year, someone will win the lottery twice. At the end of 2 years, this probability increases to 23.4%. By the end of 3 years, it increases to 45.1%.

Be sure you see that this is not the probability that *you* (or any other given person) will win the lottery twice. It is simply the probability that *someone* will win it twice. The point is this—having someone win twice is not an indication of fraud.

✓ **Checkpoint**                          Help at *Math*.and**Y**O**U**.com

DATA    You win a state lottery for $1 million and you have the choice of receiving $50,000 a year for 20 years or $500,000 in a lump-sum payment now.
(a) Suppose you live in California and take the lump-sum payment. How much state and federal tax do you have to pay on your winnings? (See Section 5.2.)
(b) Canadian citizens do not have to pay income tax on lottery winnings. Do you agree that Americans should have to pay state income tax on a lottery sponsored by the state? Explain.
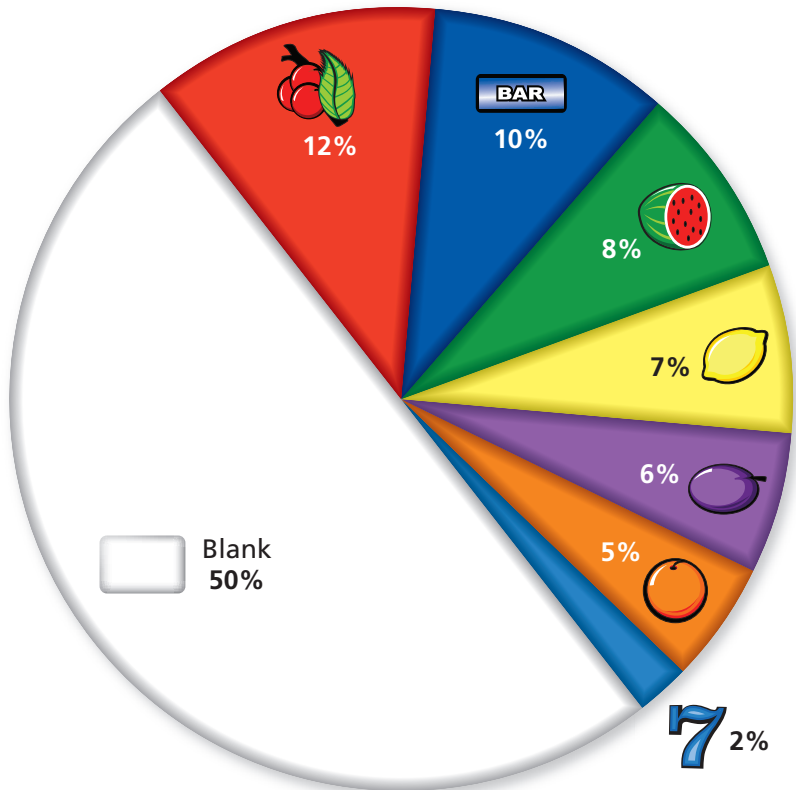
## 8.4 Exercises

**Computer-Generated Sequence** **A computer randomly generates a sequence of symbols.**
**The circle graph shows the probability that the computer generates each of the symbols.**
**In Exercises 1–7, find the probability that the sequence occurs.** *(See Examples 1 and 2.)*

1.

2. BAR BAR BAR

3.

4.

5.

6.

7.

**Probabilities of Symbols**

Around 1900, Charles Fey invented a slot machine called the Liberty Bell. It had three spinning reels with a total of five symbols: horseshoes, diamonds, spades, hearts, and a Liberty Bell. Three bells produced the biggest payoff, which was 10 nickels.

8. **Slot Machine** Complete the spreadsheet to find the expected value of each of the events in Exercises 1–7. Then find the probability that a given spin is a "no win." Suppose these symbols represent a slot machine game. What is the expected value?

| | A | B | C | D |
|---|---|---|---|---|
| 1 | **Event** | **Probability** | **Payoff** | **Expected Value** |
| 2 | 3 cherries | | 150 | |
| 3 | 3 bars | | 200 | |
| 4 | 3 watermelons | | 250 | |
| 5 | 3 lemons | | 400 | |
| 6 | 3 plums | | 1,000 | |
| 7 | 3 oranges | | 7,500 | |
| 8 | 3 sevens | | 12,000 | |
| 9 | No win | | -1 | |
| 10 | | | | |

**Dice** In Exercises 9–12, find the probability of the event. *(See Examples 3 and 4.)*

9. You roll three dice. What is the probability of rolling at least two of the same number?

10. You roll five dice. What is the probability of rolling at least two of the same number?

11. You roll six dice. What is the probability of rolling at least one odd number?

12. You roll seven dice. What is the probability of rolling at least two of the same number?

**Marbles** There are 26 marbles in a bag. The graph shows the color distribution of the marbles. In Exercises 13–16, use the graph. Assume that you randomly draw one marble from the bag and you put the marble back before drawing another. *(See Examples 3 and 4.)*



Distribution of Marbles
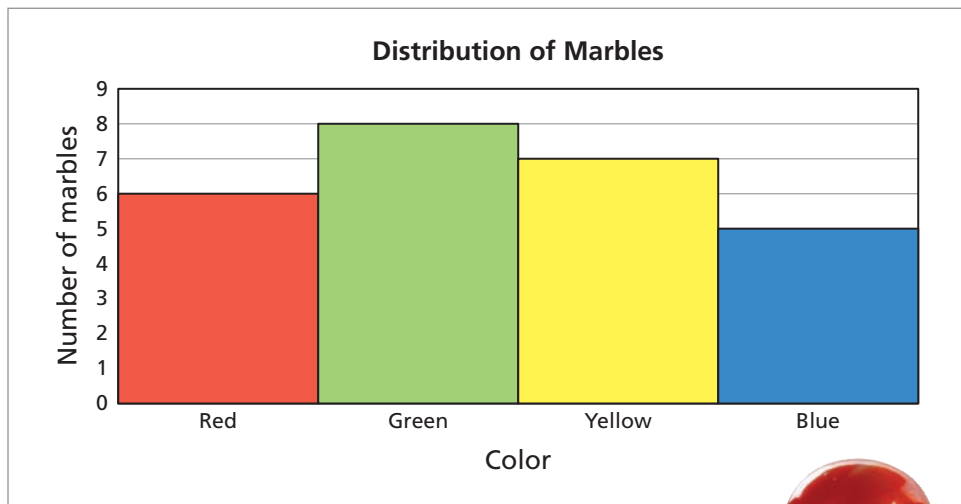
13. You randomly draw three marbles from the bag. What is the probability that none of the marbles are red?

14. You randomly draw five marbles from the bag. What is the probability that none of the marbles are yellow?

15. You randomly draw four marbles from the bag. What is the probability that at least one marble is green?

16. You randomly draw six marbles from the bag. What is the probability that at least one marble is blue?

**17. Coin Conundrum** Your friend flips two coins and tells you that at least one coin landed heads up. What is the probability that both coins landed heads up? *(See Example 5.)*

**18. Boy or Girl Paradox** Your friend has two children. At least one of the children is a boy. Considering the order that the children were born, what is the probability that the other child is a girl? *(See Example 5.)*

**19. Treasure Beyond Measure** You are a contestant on the fictional game show *Treasure Beyond Measure*. The host presents you with five unopened treasure chests. One chest contains $1,000,000. The other four contain nothing. You randomly choose two of the chests. The host, knowing which chest contains the money, opens two of the remaining chests and shows that they contain nothing. The host then asks you, "Do you want to switch your two chests for the one chest that I didn't open?" Based on probability, what should you do? *(See Example 5.)*

**20. Ultimate Monty Hall** You are a contestant on the fictional game show *Ultimate Monty Hall*. The host presents you with 101 unopened doors. Behind one door is a car. Behind the other 100 doors are goats. You randomly choose 50 of the doors. The host, knowing the location of the car, reveals the goats behind 50 of the remaining doors. The host then asks you, "Do you want to switch your 50 doors for the 1 door that I didn't open?" Based on probability, what should you do? *(See Example 5.)*

**Card Puzzle** **In Exercises 21 and 22, use the information below.** *(See Examples 5 and 6.)*

You deal two of the cards shown to your friend.

**21.** Your friend says, "I have at least one ace." What is the probability that your friend's other card is an ace?

**22.** Your friend says, "I have the ace of spades." What is the probability that your friend's other card is an ace?

# ▶ Extending Concepts

**Bayes' Theorem**  **In Exercises 23–26, use the information below.**

For any two events with probabilities greater than 0,

$$\text{Probability of event 1 given event 2} = \frac{\left(\begin{array}{c}\text{probability of event 2}\\\text{given event 1}\end{array}\right)\left(\begin{array}{c}\text{probability}\\\text{of event 1}\end{array}\right)}{\text{probability of event 2}}.$$

**23.** You have the following information about students at a college.

- 49% of the students are male.
- 11% of the students are nursing majors.
- 9% of the nursing majors are male.

What is the probability that a student is a nursing major given that the student is male?

**24.** You have the following information about students at a college.

- 51% of the students are female.
- 10% of the students are history majors.
- 60% of the history majors are female.

What is the probability that a student is a history major given that the student is female?

**25.** You have the following information about voters in a local mayoral election.

- 61% of voters were registered Republican.
- 53% of voters voted Republican.
- 86% of voters who voted Republican were registered Republican.

What is the probability that a voter voted Republican given that the voter was registered Republican?

**26.** You have the following information about voters in a local congressional election.

- 74% of voters were registered Democrat.
- 62% of voters voted Democrat.
- 79% of voters who voted Democrat were registered Democrat.

What is the probability that a voter voted Democrat given that the voter was registered Democrat?

## 8.3–8.4   Quiz

**Multiple-Choice Quiz**  **You take a 10-question multiple-choice quiz. Each question has four choices. In Exercises 1–7, suppose you guess when answering the questions.**

1. Suppose you gain 1 point when you answer a question correctly and 0 points when you answer a question incorrectly. What is the expected value of the quiz?

2. Suppose you gain 1 point when you answer a question correctly and lose ¼ point when you answer a question incorrectly. What is the expected value of the quiz?

3. What is the probability that you answer every question incorrectly?

4. What is the probability that you answer every question correctly?

5. What is the probability that you answer only the first question correctly?

6. What is the probability that you answer only the first five questions correctly?

7. What is the probability that you answer at least one question correctly?

Name _____     Date _____

1. Who was the first president of the United States?
   Ⓐ George Washington    Ⓑ Thomas Jefferson    Ⓒ John Adams    Ⓓ Gerald Ford

2. Who wrote the *Declaration of Independance*?
   Ⓐ Thomas Jefferson    Ⓑ Benjamin Franklin    Ⓒ Samuel Adams    Ⓓ John Hancock

3. Who wrote the *The Wealth of Nations*?
   Ⓐ John Maynard Keynes  Ⓑ Adam Smith    Ⓒ Thomas Paine    Ⓓ Thomas Johnson

4. Who was the president of the Confederate States?
   Ⓐ Robert E. Lee    Ⓑ Jefferson Davis    Ⓒ John C. Calhoun    Ⓓ Edwin Stanton

5. Who was Thomas Jefferson's first vice president?
   Ⓐ George Clinton    Ⓑ Aaron Burr    Ⓒ James Madison    Ⓓ Spiro Agnew

6. In what year did Alaska officially become a state?
   Ⓐ 1859    Ⓑ 1911    Ⓒ 1945    Ⓓ 1959

7. Which president was never married?
   Ⓐ Grover Cleveland    Ⓑ Andrew Jackson    Ⓒ James Buchanan    Ⓓ Ulysses S. Grant

8. Which president was called "the Great Communicator"?
   Ⓐ Franklin Roosevelt    Ⓑ John Kennedy    Ⓒ Jimmy Carter    Ⓓ Ronald Reagan

9. Which president's foreign policy was to "speak softly and carry a big stick"?
   Ⓐ William McKinley    Ⓑ Theodore Roosevelt  Ⓒ William Taft    Ⓓ James Polk

10. Which president's 1972 visit to China helped improve the relationship between the U.S. and China?
   Ⓐ Lyndon Johnson    Ⓑ Harry Truman    Ⓒ Richard Nixon    Ⓓ Dwight Eisenhower

# Chapter 8    Summary

| | *Section Objectives* | *How does it apply to you?* |
|---|---|---|
| **Section 1** | Use probability to describe the likelihood of an event. | Probabilities are numbers between 0 and 1, including 0 and 1. An event with a probability of 0 is impossible. An event with a probability of 1 is certain. *(See Examples 1 and 2.)* |
| | Analyze the likelihood of a risk. | You can assess the risk involved in a given situation. *(See Examples 3 and 4.)* |
| | Use likelihood to describe actuarial data. | Actuaries use probabilities to calculate the costs of risks. *(See Examples 5 and 6.)* |

| | | |
|---|---|---|
| **Section 2** | Find a theoretical probability. | You can find the probability of an event in cases where all possible outcomes of an event are known and can be counted. *(See Examples 1 and 2.)* |
| | Find an experimental probability. | You can find the probability of an event in cases where a representative sample can be taken and counted. *(See Examples 3 and 4.)* |
| | Estimate a probability using historical results. | You can find the probability of an event in cases where past occurrences are assumed to be representative of future occurrences. *(See Examples 5 and 6.)* |

| | | |
|---|---|---|
| **Section 3** | Find an expected value involving two events. | You can determine the long-run average of an experiment. *(See Examples 1 and 2.)* |
| | Find an expected value involving multiple events. | You can extend the concept of expected value to situations in which there are more than two events. *(See Examples 3 and 4.)* |
| | Use expected value to make investment decisions. | You can use expected value to compare different investments. *(See Examples 5 and 6.)* |

| | | |
|---|---|---|
| **Section 4** | Find the probability of independent events. | You can find the probability that two events occur when the occurrence of one does not affect the occurrence of the other. *(See Examples 1 and 2.)* |
| | Find the probability that an event does not occur. | If you know the probability that an event occurs, then you can calculate the probability that the event does not occur. *(See Examples 3 and 4.)* |
| | Find counterintuitive probabilities. | You can solve problems in which the probability of an event is counterintuitive. *(See Examples 5 and 6.)* |

## Chapter 8   Review Exercises

### Section 8.1

1. **Describing Likelihood**  The table shows the probabilities of several events.
   Complete the table. Then describe the likelihood of each event in words.

| Event | Fraction | Decimal | Percent |
|---|---|---|---|
| Being an organ donor | $\frac{4}{15}$ | | |
| Eats breakfast | | 0.61 | |
| Having a dream that comes true | | | 42.9% |
| Household with television | | | 98.2% |

**Snowfall in Washington**  In Exercises 2–5, describe the likelihood
that at least 1 inch of snow will accumulate in the city.

2. Olympia

3. Packwood

4. Richland

5. Seattle

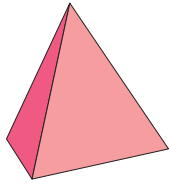**Probability of 1 Inch (or more) of Snow**
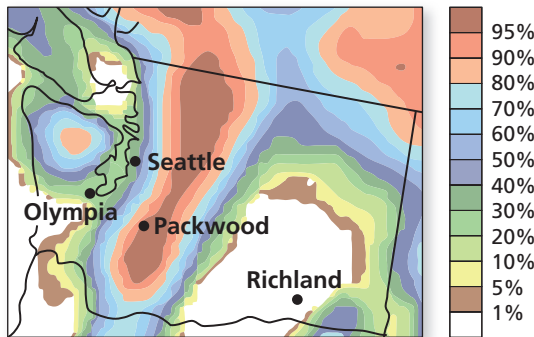


6. **Deaths**  The graph
   shows the death rate
   per 100,000 people
   in the United States
   for several causes.
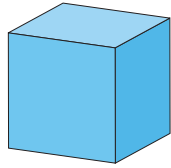   Compare the rates for
   the causes of death.
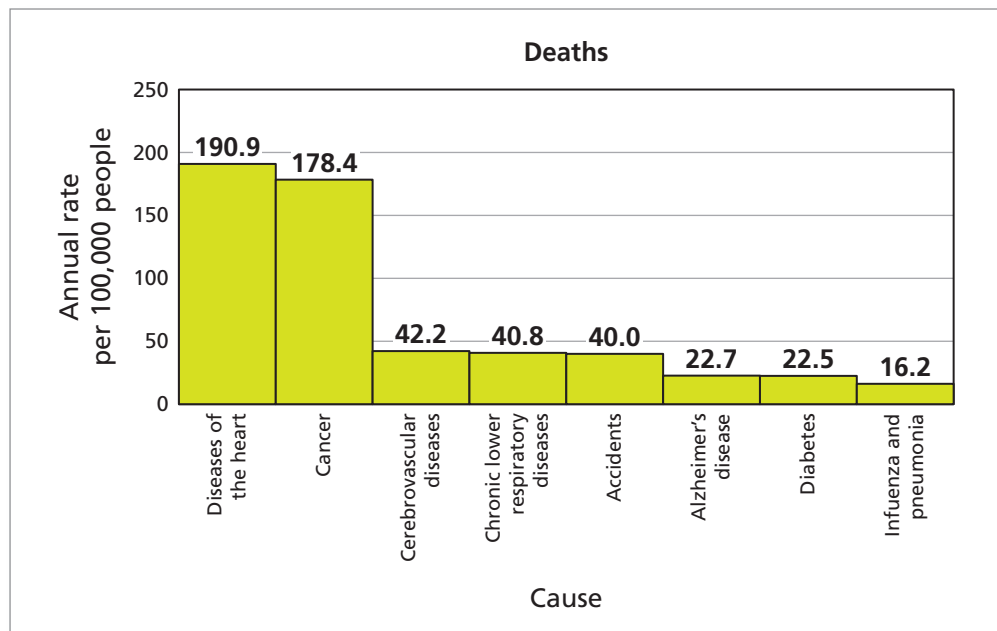
## Section 8.2

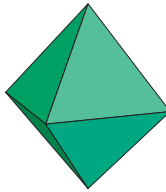**Platonic Solids  In Exercises 7–10, use the information below.**

There are five Platonic solids: tetrahedron, cube, octahedron, dodecahedron, and icosahedron.
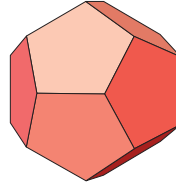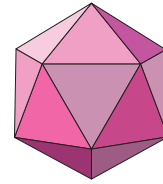
| Tetrahedron | Cube | Octahedron | Dodecahedron | Icosahedron |

**7.** Why do these solids make natural shapes for dice?

**8.** Which shape is the best for dice? Why? Which is the worst? Why?

**9.** A die has the shape of an icosahedron, with consecutively numbered sides starting at 1. What is the probability of rolling a number that is greater than 5?

**10.** There are only five possible Platonic solids. What is the definition of a Platonic solid?

**MP3 Player  The tables show the results of a survey that asked adults whether they own an MP3 player. In Exercises 11 and 12, use the tables.**

| Age group | Yes | No |
|-----------|-----|-----|
| 18–29 | 630 | 212 |
| 30–49 | 479 | 343 |
| 50–64 | 250 | 553 |
| 65 and over | 52 | 482 |

| Geographic location | Yes | No |
|---------------------|-----|-----|
| Rural | 120 | 270 |
| Suburban | 700 | 750 |
| Urban | 591 | 615 |

**11.** Find the probability that an adult in each age group owns an MP3 player. Then describe the likelihood.

**12.** Find the probability that an adult in each geographic location owns an MP3 player. Then describe the likelihood.

## Section 8.3

**Home Appliance Company** In Exercises 13–16, assume that you are the owner of a home appliance company.

13. You take out a $500,000 fire insurance policy on your factory. The annual premium is $2750. The probability of a fire is 0.005. What is the expected value?

14. You take out a $250,000 flood insurance policy on your factory. The annual premium is $3770. The probability of a flood is 0.014. What is the expected value?

15. Your company is considering developing one of two toaster models. Use a decision tree to decide which model your company should develop.

**Toaster A:** Cost of development: $500,000

| Projected Sales | |
| --- | --- |
| Probability | Net sales (in thousands) |
| 25% | $1000 |
| 65% | $600 |
| 10% | $250 |

**Toaster B:** Cost of development: $750,000

| Projected Sales | |
| --- | --- |
| Probability | Net sales (in thousands) |
| 30% | $1200 |
| 55% | $800 |
| 15% | $500 |

16. Your company is considering developing one of two microwave models. Use a decision tree to decide which model your company should develop.

**Microwave A:** Cost of development: $1,000,000

| Projected Sales | |
| --- | --- |
| Probability | Net sales (in thousands) |
| 20% | $1600 |
| 65% | $1400 |
| 15% | $800 |

**Microwave B:** Cost of development: $900,000

| Projected Sales | |
| --- | --- |
| Probability | Net sales (in thousands) |
| 35% | $1500 |
| 45% | $1300 |
| 20% | $750 |

17. **Option Comparison** Compare the two options.

| | Probability | Gain |
| --- | --- | --- |
| Option 1 | 100% | $1000 |
| | 0% | $0 |
| Option 2 | 50% | $500 |
| | 50% | $2000 |

18. **Investment Comparison** You want to invest $1000. Find the expected values for the two investments.

**Speculative investment**
- Complete loss: 30% chance
- No gain or loss: 25% chance
- 100% gain: 25% chance
- 200% gain: 15% chance
- 500% gain: 5% chance

**Conservative investment**
- Complete loss: 5% chance
- No gain or loss: 30% chance
- 25% gain: 60% chance
- 50% gain: 5% chance

## Section 8.4

**Summer Weather**  The table shows the probability of sunshine or rain on any given day in July for a town. In Exercises 19–22, use the table.

| Weather | Probability |
|---------|-------------|
| ☀️ | 75% |
| 🌧️ | 25% |

**19.** What is the probability of rain three days in a row?

**20.** What is the probability of sunshine five days in a row?

**21.** What is the probability of rain at least once during a week?

**22.** What is the probability of sunshine at least once during a week?

**Winter Weather**  The probability of snow on any given day in December for a town is 30%. In Exercises 23–26, use this information.

**23.** What is the probability of snow four days in a row?

**24.** What is the probability of no snow six days in a row?

**25.** What is the probability of snow at least once during a week?

**26.** What is the probability of no snow at least once during a week?

# 9 The Mathematics of Description

## 9.1 Information Design

▶ Use stacked area graphs to represent the changing parts of a whole.

▶ Use a radar graph and an area graph to represent data.

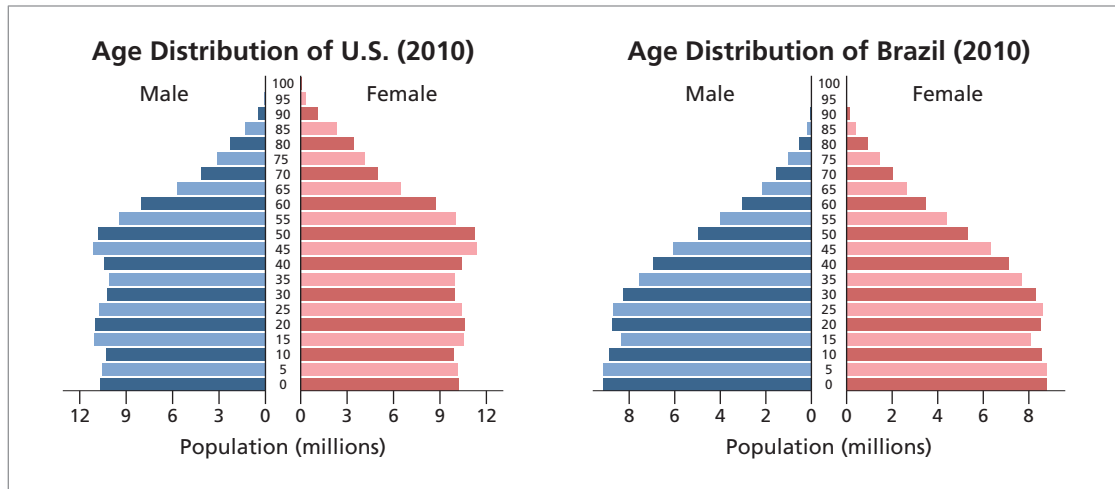▶ Graphically represent data sets that have several variables.

## 9.2 Describing "Average"

▶ Use mean, median, and mode to describe the average value of a data set.

▶ Read and understand box-and-whisker plots and histograms.

▶ Understand the effect of outliers on averages.

## 9.3 Describing Dispersion

▶ Use standard deviation to describe the dispersion of a data set.

▶ Use standard deviation to describe a data set that is normally distributed.

▶ Compare different types of distributions.

## 9.4 Describing by Sampling

▶ Use a randomly chosen sample to describe a population.

▶ Determine whether a sample is representative of a population.

▶ Determine a sample size to obtain valid inferences.

**Age Distribution of U.S. (2010)**

Male — Female

Population (millions)
12  9  6  3  0  0  3  6  9  12

**Age Distribution of Brazil (2010)**

Male — Female

Population (millions)
8  6  4  2  0  0  2  4  6  8

Example 1 on page 412 uses population pyramids to describe
the distribution of ages in the United States and in Brazil.
How do the two distributions differ?
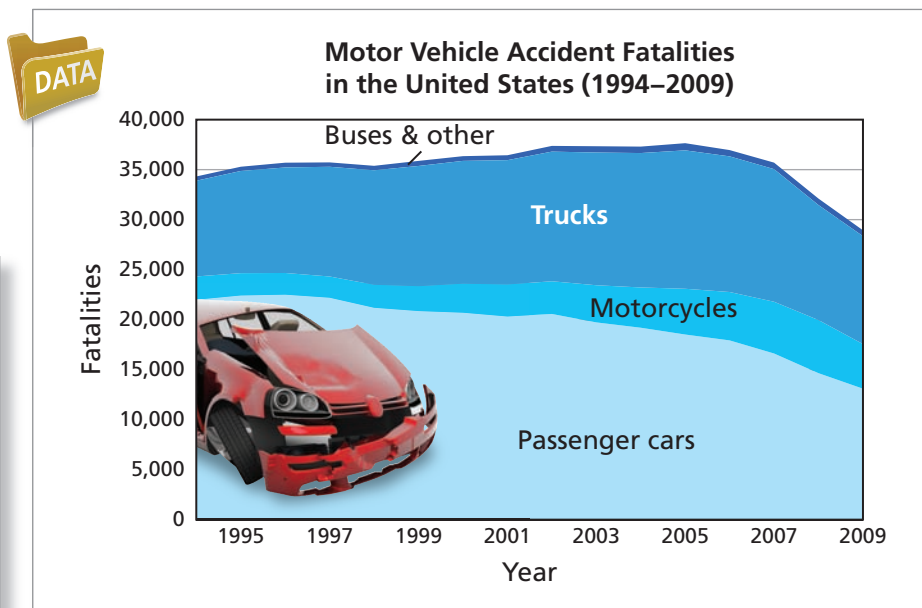
# 9.1 Information Design

▶ Use stacked area graphs to represent the changing parts of a whole.

▶ Use a radar graph and an area graph to represent data.

▶ Graphically represent data sets that have several variables.

## Stacked Area Graphs

**Information design** is the presentation of data and information so that people can understand and use it. Throughout this text, you have seen many types of information design: bar graphs, circle graphs, scatter plots, line graphs, and bubble graphs. In this section, you will look at several more graphical ways to organize and present data.

### EXAMPLE 1    Reading a Stacked Area Graph

Describe the information presented in the *stacked area graph*.

DATA

**Motor Vehicle Accident Fatalities in the United States (1994–2009)**

Buses & other

Trucks

Motorcycles

Passenger cars

Fatalities (y-axis): 0, 5,000, 10,000, 15,000, 20,000, 25,000, 30,000, 35,000, 40,000

Year (x-axis): 1995, 1997, 1999, 2001, 2003, 2005, 2007, 2009

The picture of data and information that you show people has everything to do with their understanding of the data and information.
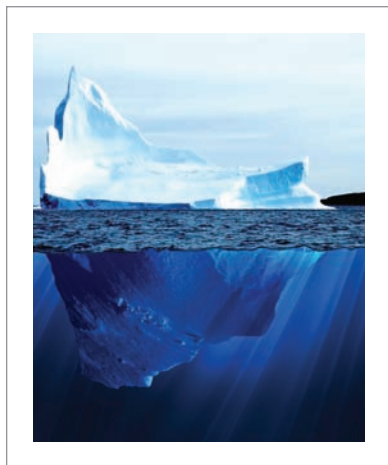
## SOLUTION

Here are some observations.

- The total number of fatalities was relatively constant from 1994 to 2007.
- The number of passenger car fatalities has decreased.
- In 1994, passenger cars accounted for about two-thirds of all motor vehicle fatalities. By 2009, they accounted for only about one-half of all motor vehicle fatalities.
- The number of motorcycle fatalities has increased.

✓ **Checkpoint**                    Help at *Math*.**and**Y☻U.com

The data for Example 1 is available at *Math.andYou.com*. Use the data to compare motorcycle fatalities in 1994 and 2009.
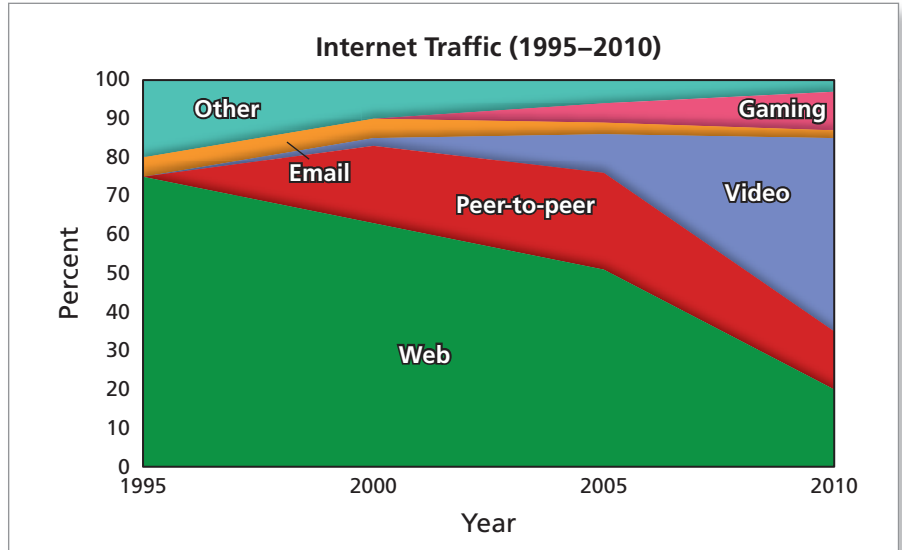
**EXAMPLE 2**    **Interpreting a Stacked Area Graph**

A household began using the Internet in 1995. The stacked area graph shows the composition of Internet traffic for the household from 1995 to 2010. Suppose you see this graph online. What comments can you make about the graph?

**Internet Traffic (1995–2010)**



### SOLUTION

Here are some comments.

COMMENT:

> What does the vertical axis represent? Time? Bits? Users?

This is a good question. The answer is "bits."

Here are some other comments.

COMMENT:

> In the graph, the term Web does not mean Internet. It means HTML data. Someone looking at the graph may be misled and assume that the term Web is referring to the Internet.

COMMENT:

> One minute of streaming video is about 2 MB.  This is about 100 times the size of a typical e-mail and about 10 times the size of a typical web page. So, you would have to send 100 e-mails or visit 10 web pages to use the same amount of bandwidth as a single minute of streaming video.

This graph illustrates that you must be careful with how you present data. Information design is a powerful tool. When it is used incorrectly, the results can be very misleading.

✓ **Checkpoint**                    Help at *Math*.and**Y🌐U**.com

From the graph, can you conclude that the household's Web usage, in total bits downloaded, declined from 1995 to 2010? Explain.

## Radar Graphs and Area Graphs

EXAMPLE 3    **Reading a Radar Graph**

Describe the information presented in the *radar graph*.



During the 1989 revolution against Communist dictator Nicolae Ceausescu, a fire destroyed about 500,000 books in Bucharest's Central University Library. After the fall of communism, Romania moved into a period of economic reform. The library was repaired and reopened in 2001. Romania joined the European Union in 2007.
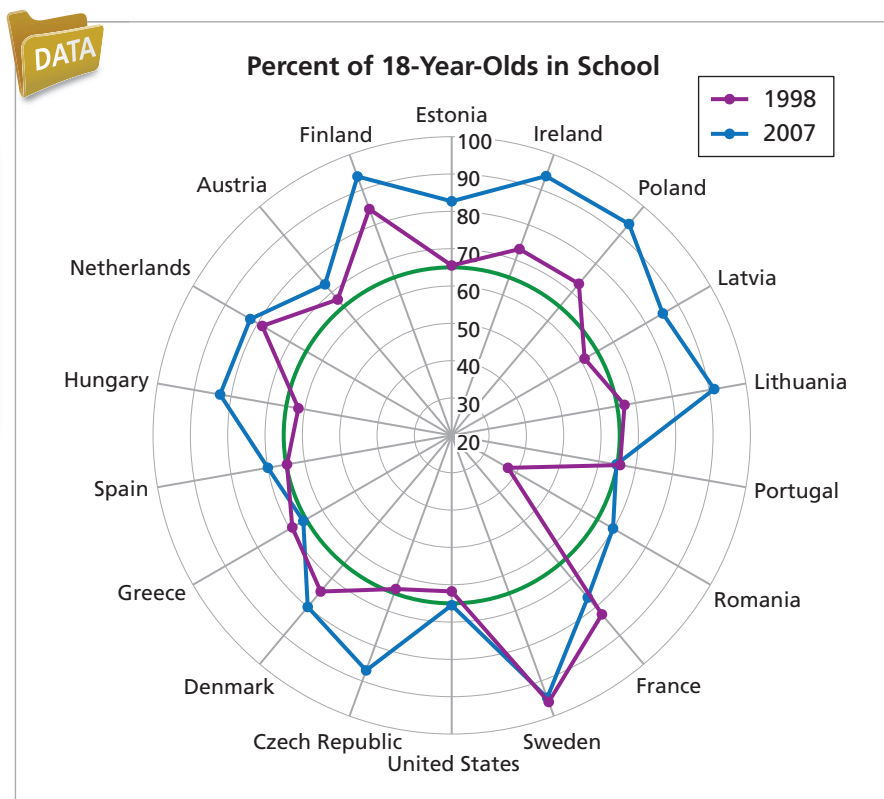
### SOLUTION

Here are some observations.

- The percent increased in all countries (the **blue graph** is farther away from the center than the **purple graph**) except for France, Greece, Portugal, and Sweden.

- In France and Greece, the percent decreased.

- In Portugal and Sweden, the percent stayed about the same.

- By drawing a **circle** at 65%, you can see that in 2007, the United States, Portugal, and Greece had the least percent of 18-year-olds in school.

- The country that made the greatest increase from 1998 to 2007 was Romania. You might assume that Romania had a great cultural change during these 10 years.
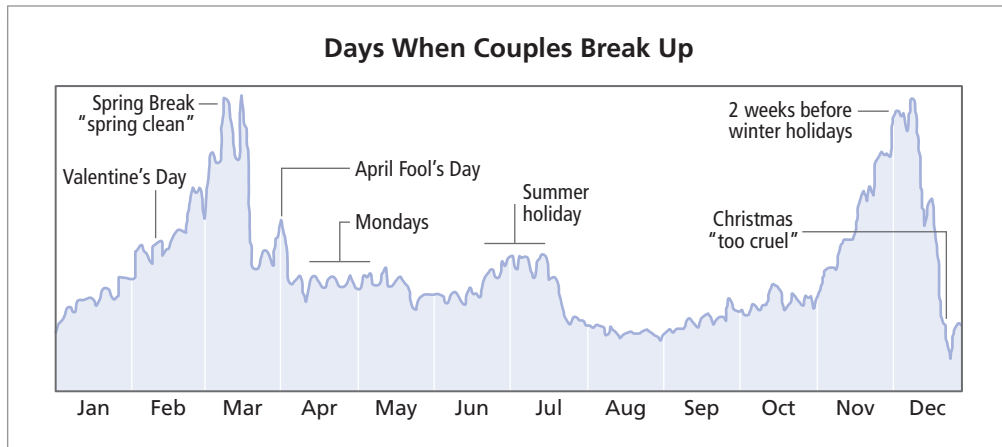
✓ **Checkpoint**                                    Help at *Math*.and**Y♥U**.com
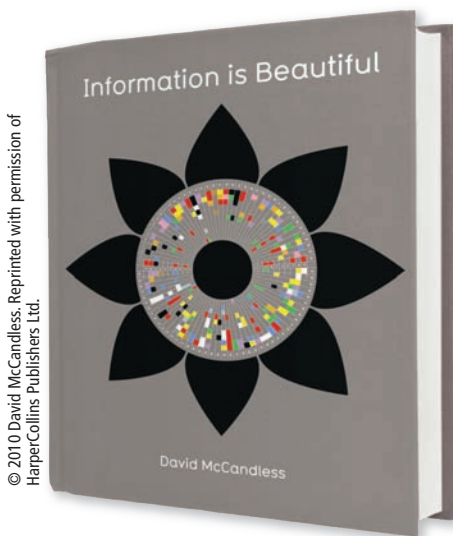
What other observations can you make about the radar graph?

EXAMPLE 4    **Drawing an Area Graph**

The following area graph is taken from *Information is Beautiful* by David McCandless. Describe the information and patterns presented in the graph.

**Days When Couples Break Up**

Spring Break —
"spring clean"

Valentine's Day

April Fool's Day

Mondays

Summer holiday

2 weeks before
winter holidays

Christmas
"too cruel"

Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct   Nov   Dec

COPYRIGHT © 2009 BY DAVID MCCANDLESS. Reprinted with permission of HarperCollins Publishers

Information is Beautiful

David McCandless

David McCandless's book on information design is "a visual guide to the way the world really works."

**SOLUTION**

The graph shows the results of tracking 10,000 Facebook status updates to determine when people in relationships broke up.
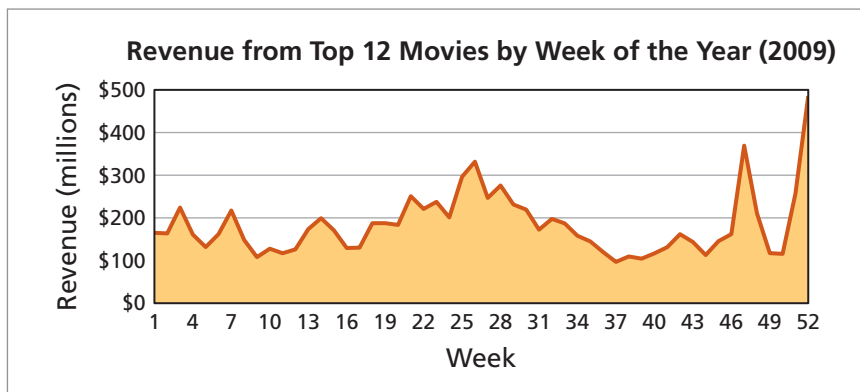
There are several patterns that you can observe in the graph.

- The two peak times during the year when people break up are in the month leading up to Spring Break and the month leading up to Christmas.

- The spike in the number of breakups on April Fool's Day may be a result of people changing their relationship statuses as a joke.

- The frequency of breakups is relatively low from late July through October.

- The day of the year when people are least likely to break up is Christmas.

✓ **Checkpoint**                    Help at *Math*.and**YOU**.com

What observations can you make about the area graph below?

**Revenue from Top 12 Movies by Week of the Year (2009)**

$500
$400
$300
$200
$100
$0

Revenue (millions)

1   4   7   10   13   16   19   22   25   28   31   34   37   40   43   46   49   52

Week

## Information Design with Several Variables

**EXAMPLE 5** **Animating Information**

At *Gapminder.org/world*, you can watch "Wealth & Health of Nations" to see how countries have changed from 1800 through the present. Describe some of the variables displayed in the graph.
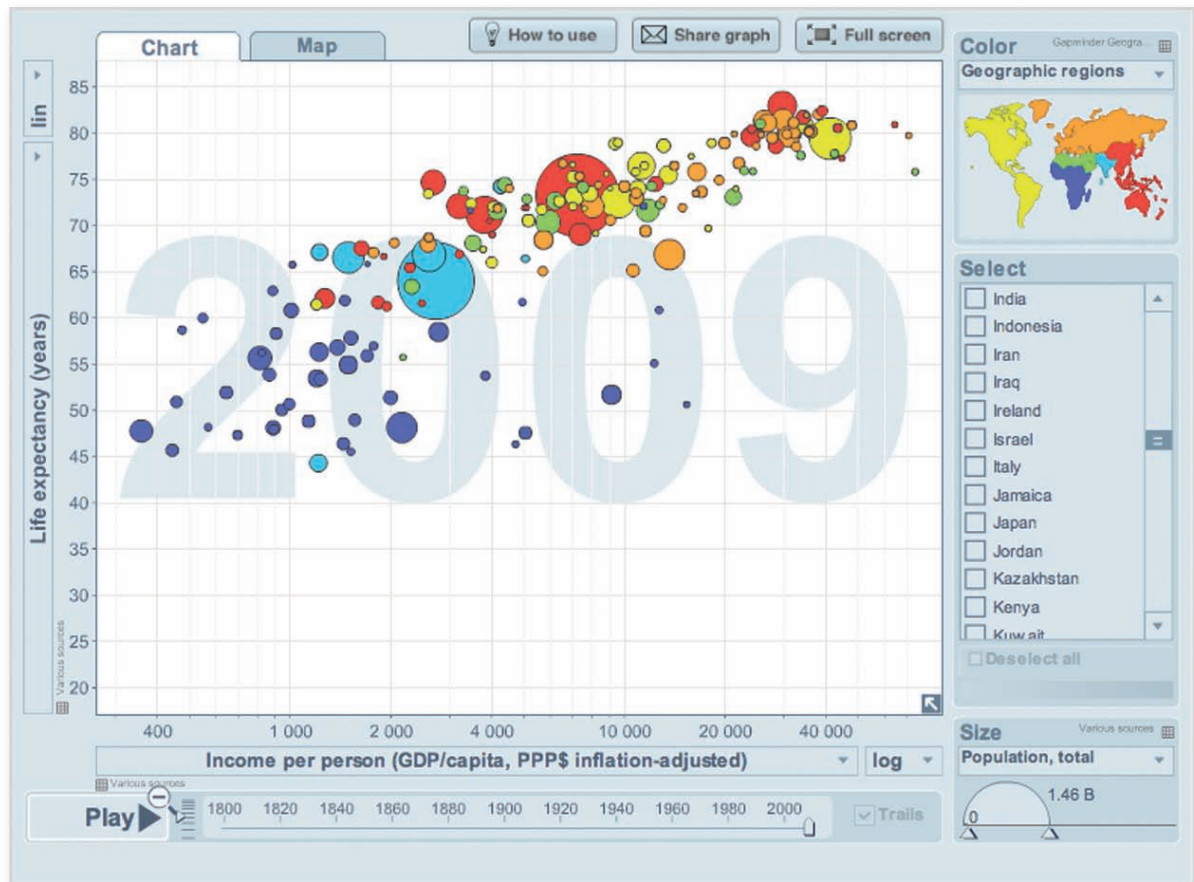


### SOLUTION

Here are some of the variables.

- The vertical axis measures *life expectancy*. The higher the bubble, the greater the life expectancy.
- The horizontal axis measures *income per person*. The farther to the right the bubble, the greater the income per person.
- The area of the bubble measures *population size*. The bigger the bubble, the larger the population.
- The color of the bubble indicates the *geographic region*.
- The date on the coordinate plane indicates the *year*.



You can watch a short lecture called "200 years that changed the world" at *Gapminder.org*.
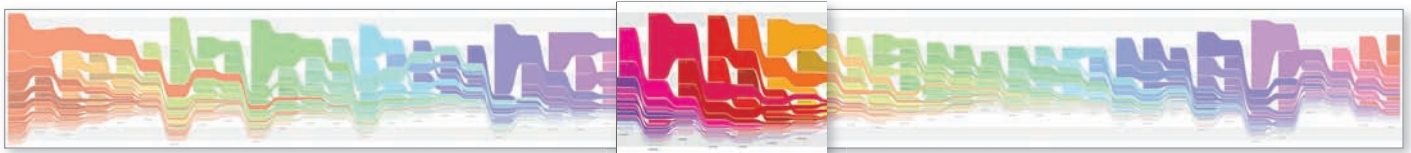
✓ **Checkpoint** Help at *Math*.and**Y**○**U**.com

● Watch the position of China in the animation from 1800 through the present. Describe the changes during the 200+ years.
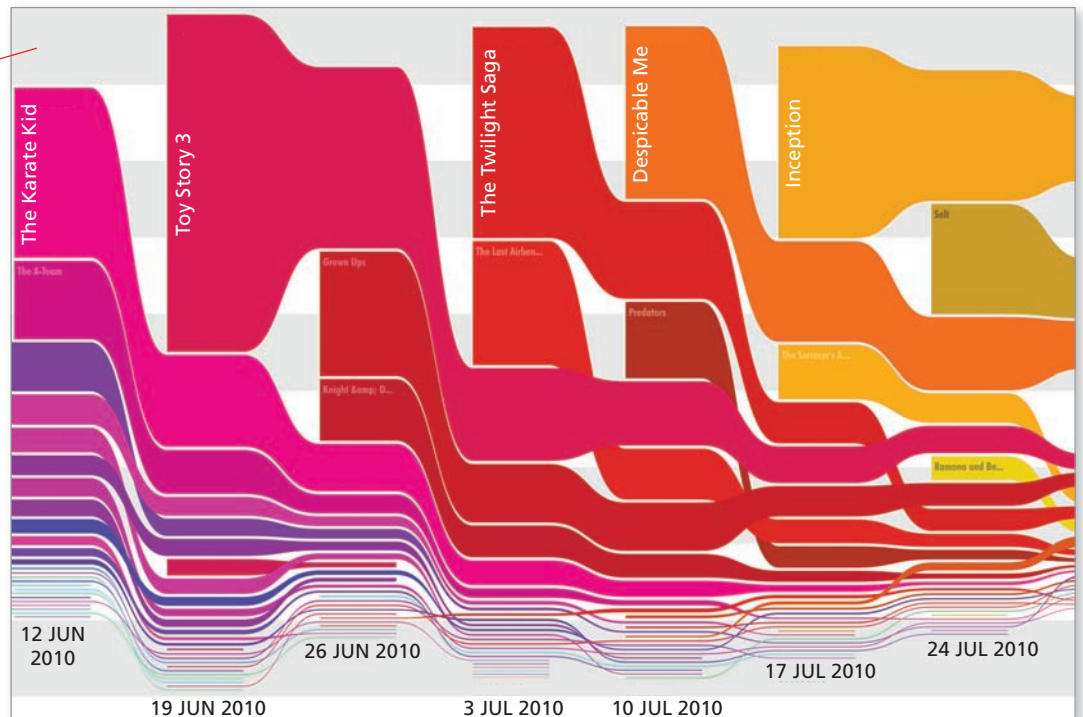
EXAMPLE 6     **Reading a Design with Several Variables**

The stream graph shows trends in the top 25 movies at the box office for each weekend in 2010. Describe the variables displayed in the graph.



**Revenue for Top 25 Movies in 2010**

Each horizontal section repesents $25 million in revenue.

The Karate Kid

Toy Story 3

The Twilight Saga

Despicable Me

Inception

The A-Team

Grown Ups

Knight &amp;amp; D...

The Last Airben...

Predators

Salt

The Sorcerer's A...

Ramona and Be...

12 JUN 2010     19 JUN 2010     26 JUN 2010     3 JUL 2010     10 JUL 2010     17 JUL 2010     24 JUL 2010

© 2010 Disney/Pixar

*Toy Story 3* is a 3D computer-animated film by Disney/Pixar. During 2010, it had the greatest box office receipts for the weekends of June 19 and June 26.

**SOLUTION**

Here are the variables.

- The color identifies the *title* of the movie.

- The horizontal axis shows the *weekend* during the year.

- The height of the color shows the *revenue* for the weekend.

- The position of the color for each weekend indicates the *order* from highest box office receipts to lowest box office receipts.

✓ **Checkpoint**                     Help at *Math*.and**YOU**.com

Describe how the colors in the graph can help you identify movies that have particularly long runs at the box office.

# 9.1 Exercises

**USAF Fighter Force** The stacked area graph shows the composition of the aircraft
in the United States Air Force (USAF) fighter force. In Exercises 1–6, use the graph.
*(See Examples 1 and 2.)*

**USAF Fighter Force (1950–2009)**
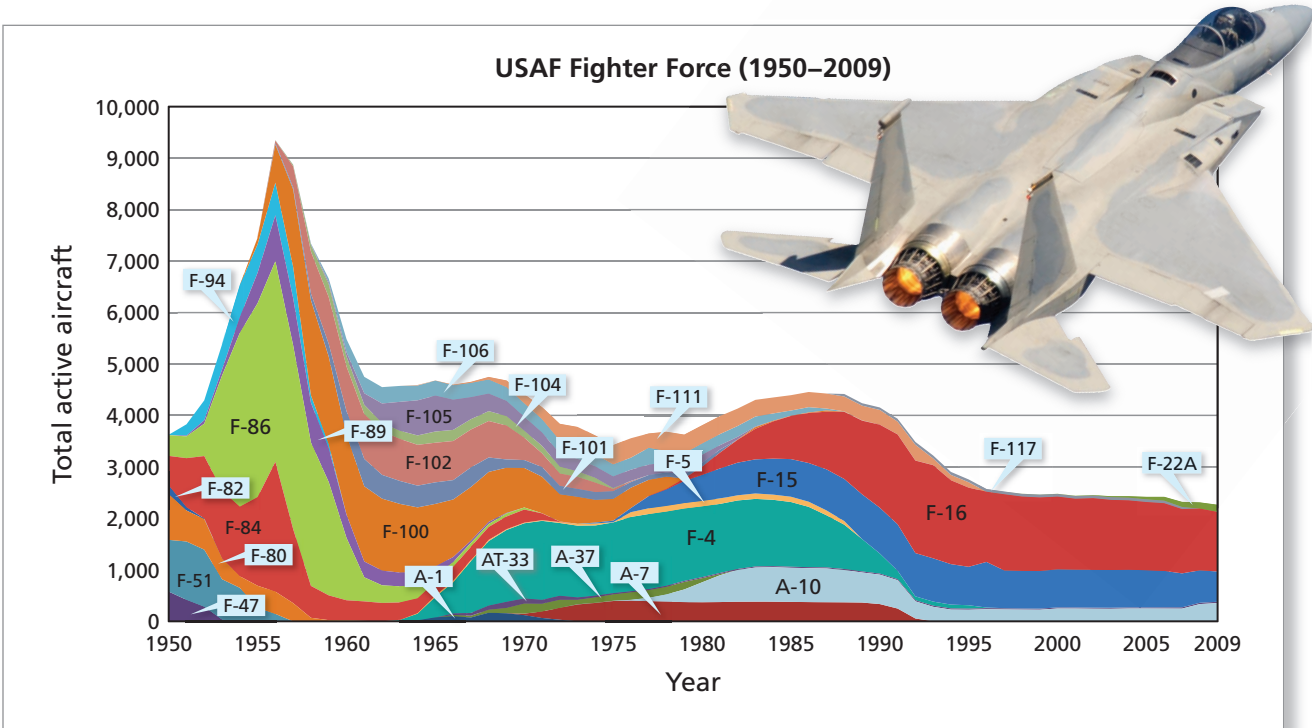
1. Describe the information presented in the stacked area graph.

2. Which aircraft was the largest component of the USAF fighter force
   in the 1970s?

3. Find the size of the USAF fighter force in 2009 as a percent of the
   peak size in the 1950s.

4. In which decade did the USAF fighter force have the least variety of aircraft?
   Explain your reasoning.

5. Do you agree with the comment shown? Explain your reasoning.

COMMENT:

The F-16 is losing favor as the most popular fighter jet in the USAF
fighter force.

6. What other comments can you make about the stacked
   area graph?

**Competitor Profiles**  In Exercises 7–10, use the radar graph. *(See Example 3.)*



**Competitor Profiles**

Revenue (millions)

Company A
Company B
Company C
Company D
Company E
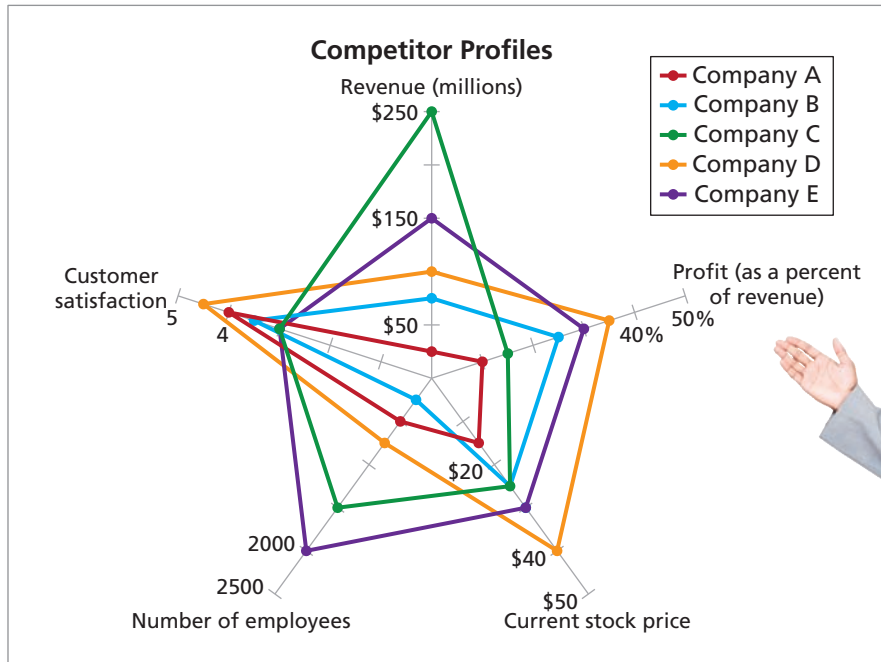
7. Which competitor has the greatest customer satisfaction?

8. Which competitor has the most employees?

9. Which competitor has the greatest profit percent? Which competitor makes the most profit? Explain your reasoning.

10. When do you think it would be beneficial to use a radar graph in information design?

**Stock Price**  The area graph shows the stock price of company C during 2010. In Exercises 11–14, use the area graph. *(See Example 4.)*

11. Describe the information and patterns presented in the graph.

12. What was the stock price in week 31?

13. Suppose you bought 100 shares of stock in week 15 and sold them in week 51. Did you lose money or earn a profit?

14. What was the highest percent return an investor could have earned during 2010? Explain your reasoning.



**2010 Stock Price for Company C**
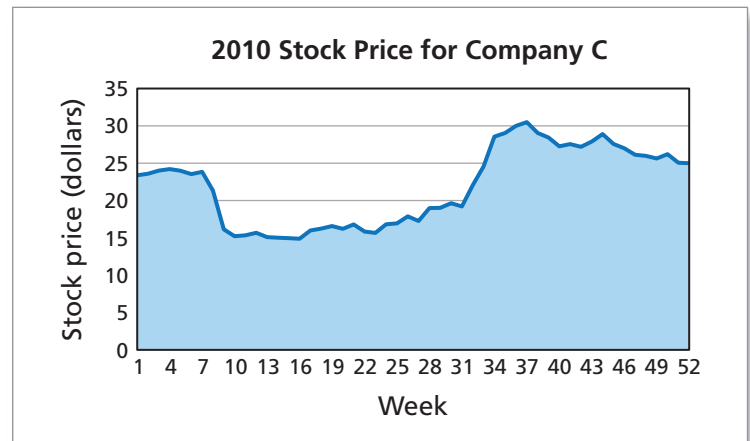
**Candlestick Chart** A candlestick chart can be used to monitor the movements of a stock price. In Exercises 15–22, use the chart. *(See Examples 5 and 6.)*
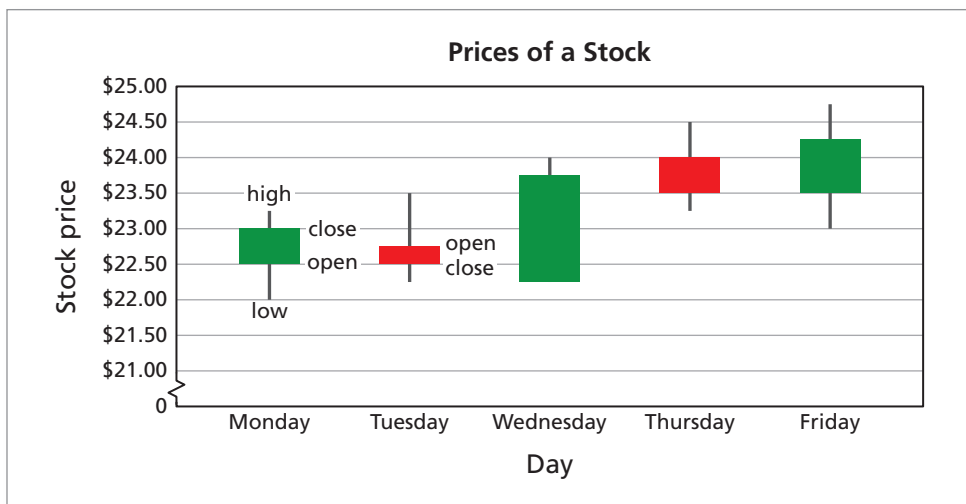


**15.** Describe the variables displayed in the graph.

**16.** What does it mean when the candlestick is green? red?

**17.** What is the opening price on Monday? Tuesday?

**18.** What is the closing price on Wednesday? Thursday?

**19.** On which day did the stock have the greatest gain? loss?

**20.** Explain why the candlestick for Wednesday does not have a lower shadow.

**21.** What is the highest percent return you could have earned during this week? Explain your reasoning.

**22.** Use the Internet to find other ways to display stock prices. Which type of design do you prefer?
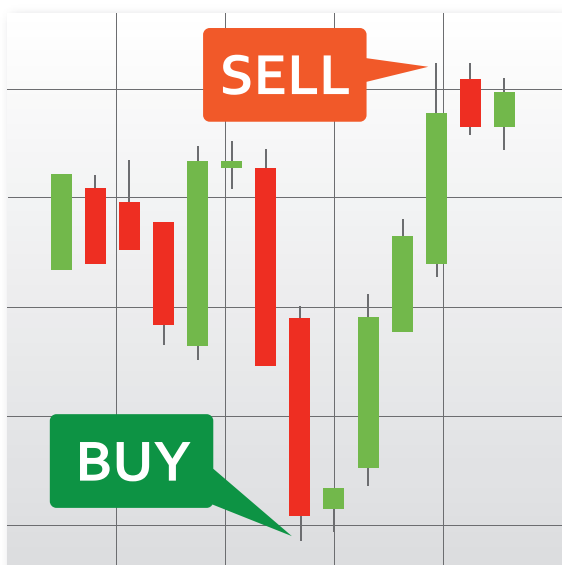


The techniques used in candlestick charting originated in 17th-century Japan. Homma, a rice trader from Japan, is credited with early versions of candlestick charting, which have evolved over many years.

# ▶ Extending Concepts

**Life Chart**  **In Exercises 23–28, use the following information.**

Your habits, behaviors, and activities indicate who you are as a person. A life chart gives a visual representation of your daily information over time. An example of a life chart is shown.
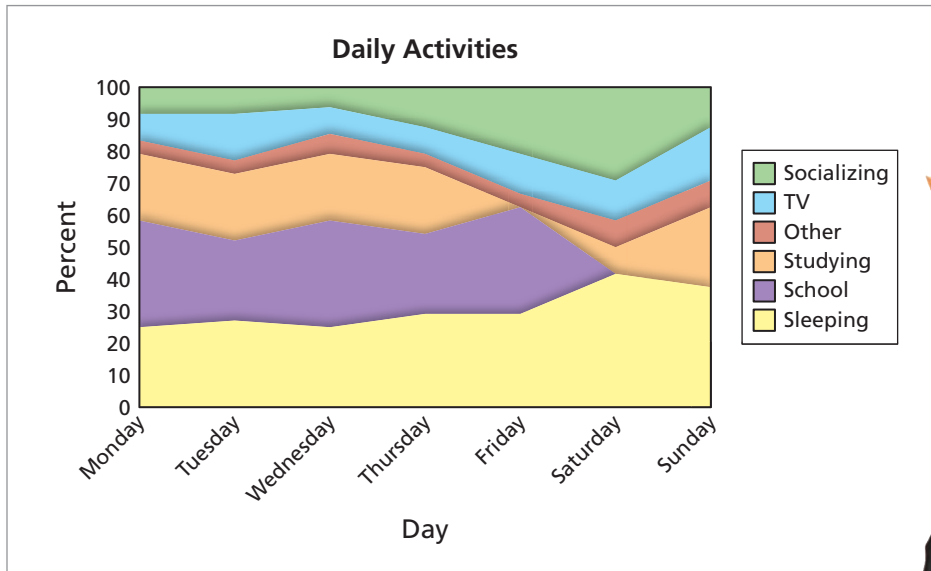


**DATA**  **23.** Estimate the amount of time you allocate to your daily activities for a typical week. Use the data to create a life chart like the one shown.

**24.** How might your life chart for a year differ from your life chart for a week?

**25.** Explain why a chart like the one in Example 2 is better suited for the information displayed in a life chart than a chart like the one in Example 1.

**26.** Many people try to make lifestyle changes such as getting more sleep, saving more money, and losing weight. Do you think a life chart can help you make lifestyle changes over time? Explain your reasoning.

**27.** Other than time allocation, what are some other personal data that you can track to help manage your life?

**28.** Organize some data from one of the topics you listed in Exercise 27. Create an information design to present the data in a way that you have not seen in this text.

# 9.2 Describing "Average"

▶ Use mean, median, and mode to describe the average value of a data set.

▶ Read and understand box-and-whisker plots and histograms.

▶ Understand the effect of outliers on averages.

## Mean, Median, and Mode

Some data sets have typical values that are representative of the entire set. For instance, a typical adult thumb is about 1 inch wide. (Historically, this is how inches were measured.) In such data sets, there are three basic ways to describe the "average" of the data set. These measures are called *measures of central tendency*. This is part of a field called *descriptive statistics*.

## Mean, Median, and Mode

To find the **mean,** add all the values in the data set and divide by the number of values in the set.

To find the **median,** arrange the values in order. The number in the middle or the mean of the two middle values is the median.

To find the **mode,** look for the value that occurs most often in the data set.

### EXAMPLE 1    Estimating the Mean, Median, and Mode

The population pyramid shows the age distributions of males and females in the United States. Estimate the mean, median, and mode for males and females.

**Age Distribution of U.S. (2010)**



#### SOLUTION

You can use a spreadsheet and the actual data to determine the mean, median, and mode.

**Mean:** The mean age for males is about 36, and the mean age for females is about 39.
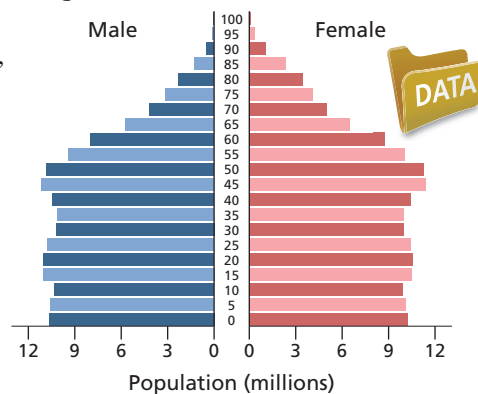
**Median:** The median age for both males and females is between 35 and 39.

**Mode:** The mode age for both males and females is between 45 and 49.
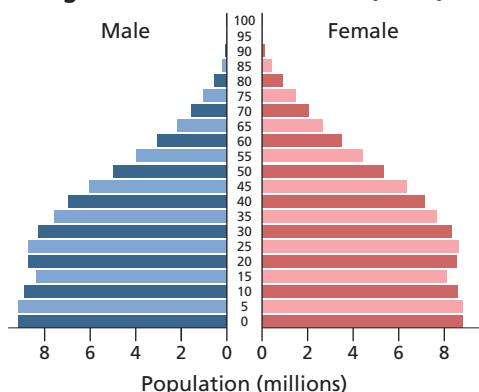
**Age Distribution of Brazil (2010)**



✓ **Checkpoint**    Help at *Math*.and**YOU**.com

How does the population pyramid of the United States differ from the population pyramid of Brazil? Explain your reasoning.
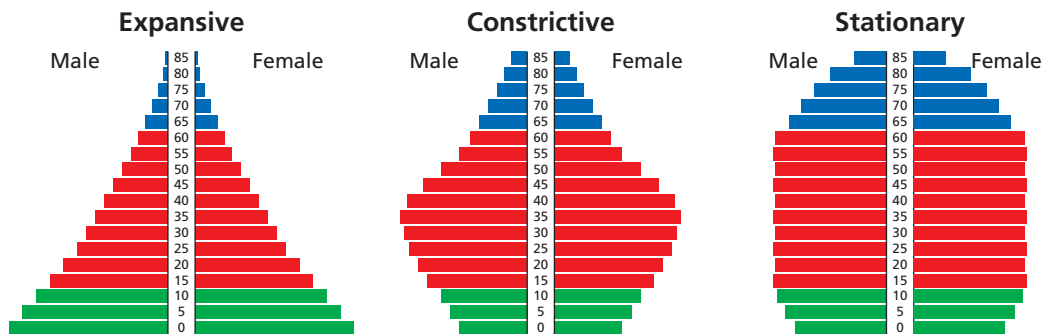
EXAMPLE 2    **Comparing Population Pyramids**

Compare the mean, median, and mode of the three types of population pyramids. Describe the economic repercussions of each type.

**Expansive**

Male    Female

**Constrictive**

Male    Female

**Stationary**

Male    Female

## SOLUTION

**Expansive:** A large percent of the population is in the **young age group** and **working age group**. The mean, median, and mode lie toward the beginning of the age spectrum. The working age group must support a relatively large population of youths.

**Constrictive:** A large percent of the population is in the **working age group**. The mean, median, and mode are similar to those of the United States in Example 1. This group must provide for the young and the old.

**Stationary:** There is a larger percent of the population in the **old age group** than in the other two population pyramids. The mean and the median lie toward the middle of the age spectrum. The working age group must support a relatively large population of old people.

✓ **Checkpoint**

Help at *Math*.andY♥U.com

Match the population pyramid with Afghanistan, Canada, and Mexico. Explain your reasoning.

**a.**

Male    Female

4.8  3.6  2.4  1.2  0    0  1.2  2.4  3.6  4.8
Population (millions)

**b.**

Male    Female

1.6  1.2  0.8  0.4  0    0  0.4  0.8  1.2  1.6
Population (millions)

**c.**

Male    Female

2.4  1.8  1.2  0.6  0    0  0.6  1.2  1.8  2.4
Population (millions)
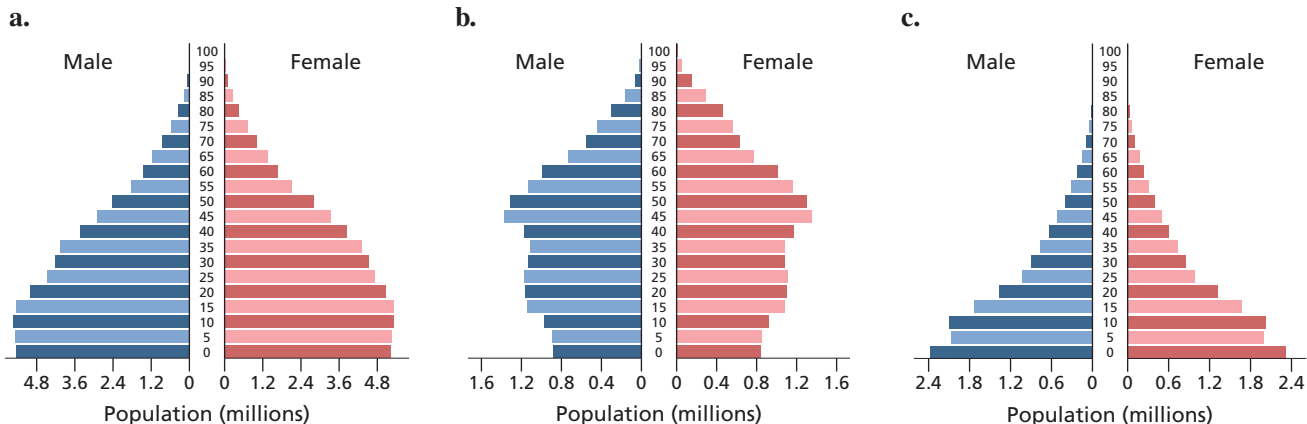
## Box-and-Whisker Plots and Histograms

**EXAMPLE 3**     **Reading a Box-and-Whisker Plot**

The box-and-whisker plot shows the weights of the 59 players on the 2010 Chicago Bears football team. Use the graph to analyze the players' weights.

**Weights of Chicago Bears**

Minimum: 183     Median: 233     Mean: 241.3     Maximum: 328

Lower quartile: 202.5     Upper quartile: 276.5

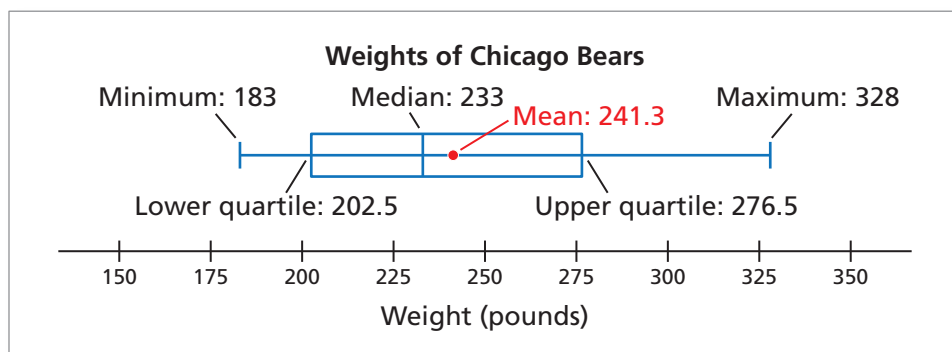| 150 | 175 | 200 | 225 | 250 | 275 | 300 | 325 | 350 |

Weight (pounds)

### SOLUTION

A box-and-whisker plot conveys a wealth of information. Here are some observations you can make about the weights of the players.

● About 15 players weighed between 183 and 202.5 pounds.

● About 15 players weighed between 202.5 and 233 pounds.

● About 15 players weighed between 233 and 276.5 pounds.

● About 15 players weighed between 276.5 and 328 pounds.

● There was a greater range of weights for the heaviest 25% of the players (about 50 pounds) than there was for the lightest 25% of the players (about 20 pounds).

You can make your own box-and-whisker plots at *Math.andYou.com*.

The mean weight of National Football League (NFL) quarterbacks is about 24 pounds less than the mean weight of the players in general.

✓ **Checkpoint**          Help at *Math*.and**YOU**.com

The box-and-whisker plot shows the weights of the 40-man roster of the 2010 Los Angeles Angels of Anaheim baseball team. Use the graph to analyze the players' weights.

**Weights of Los Angeles Angels of Anaheim**

Minimum: 170   Median: 200     Mean: 202.5          Maximum: 240

Lower quartile: 188.75     Upper quartile: 215

| 150 | 175 | 200 | 225 | 250 | 275 | 300 | 325 | 350 |

Weight (pounds)

EXAMPLE 4    **Reading Histograms**
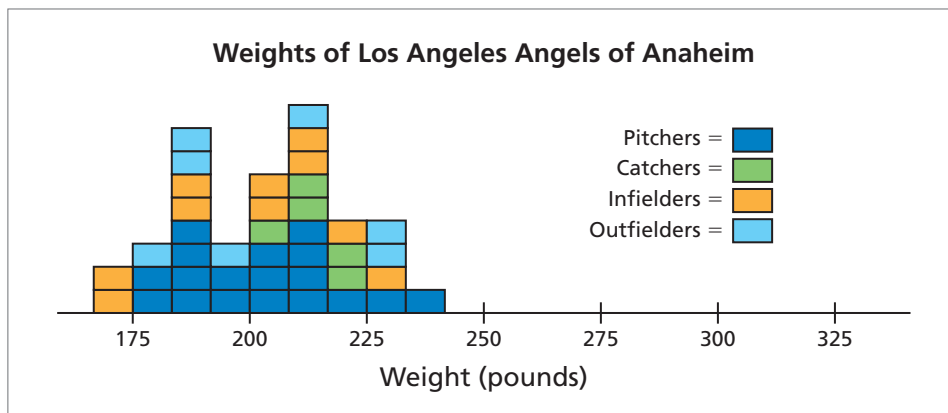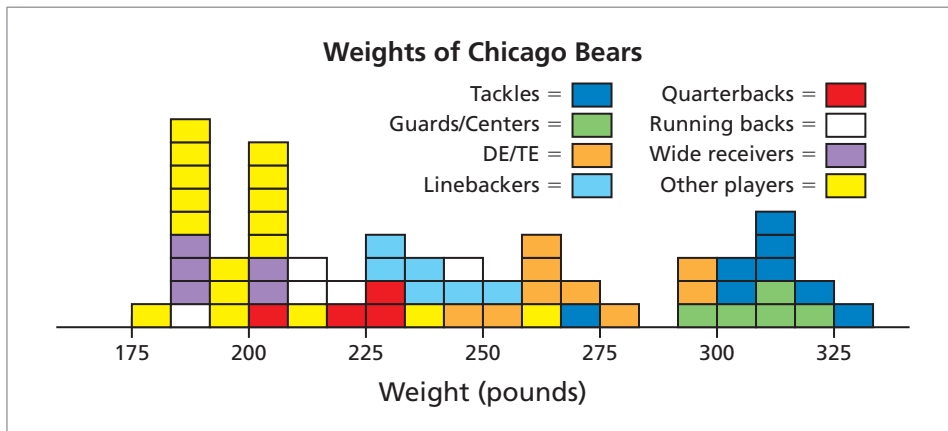
The color-coded histograms show the data from Example 3 in a different way. Compare the histograms to the box-and-whisker plots. What new information do the histograms provide?

**Weights of Chicago Bears**

Tackles = (blue)    Quarterbacks = (red)
Guards/Centers = (green)    Running backs = (white)
DE/TE = (orange)    Wide receivers = (purple)
Linebackers = (light blue)    Other players = (yellow)

Weight (pounds)

**Weights of Los Angeles Angels of Anaheim**

Pitchers = (blue)
Catchers = (green)
Infielders = (orange)
Outfielders = (light blue)

Weight (pounds)

The mean weight of Major League Baseball (MLB) pitchers is not very different from the mean weight of the players in general.

**SOLUTION**

The histograms do not show the medians or quartiles of the data sets.

On the other hand, the histograms show how the data are distributed. You can make the following observations that are not evident in the box-and-whisker plots.

● For *football players*, there is a correlation between weight and position. Tackles, guards, centers, tight ends, and defensive ends tend to weigh considerably more than the other players.

● For *baseball players*, there is not a strong correlation between weight and position.

✓ **Checkpoint**                    Help at *Math*.and**Y◉U**.com

Make histograms showing the weights of players on another NFL team and another MLB team. Are the histograms similar to the two above? Explain your reasoning.

## The Effect of Outliers on Averages

An **outlier** in a data set is a value that lies outside (is much smaller or larger than) most of the other values in the data set.

> **EXAMPLE 5** **Analyzing Outliers and Averages**

The 2009 U.S. Census report on income in the United States uses median income to describe the average income. Why does the report use median income instead of mean income to represent the average income?

**U.S. Male and Female Earnings**
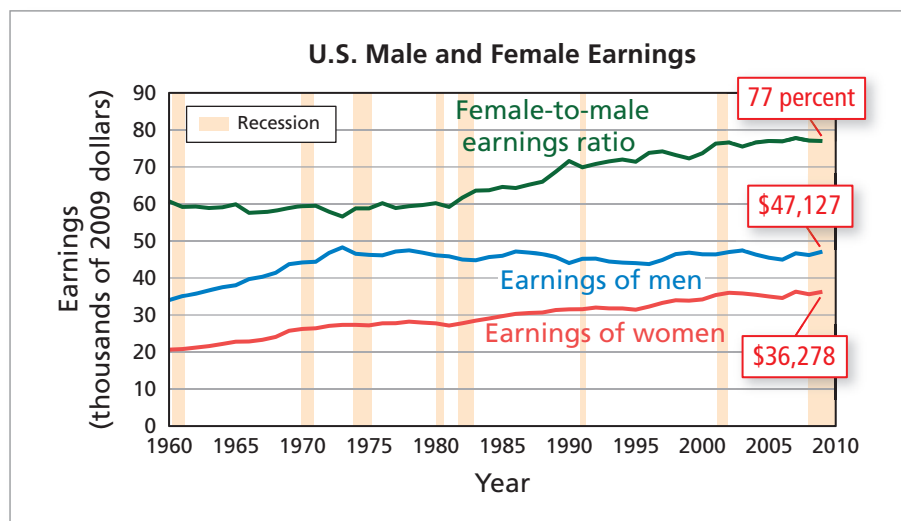


### SOLUTION

The mean annual household income in the United States was about $67,976 in 2009. The median household income was $49,777. The circle graph below helps explain the discrepancy between these two averages. The very high incomes in the top 20% caused the mean to overestimate the typical household income. Because the median is not as strongly affected by outliers as the mean, it is used to measure average income.

**Mean Household Income**



Each year, the U.S. Census Bureau issues a report called *Income, Poverty, and Health Insurance Coverage in the United States.* This one was issued in September 2010.
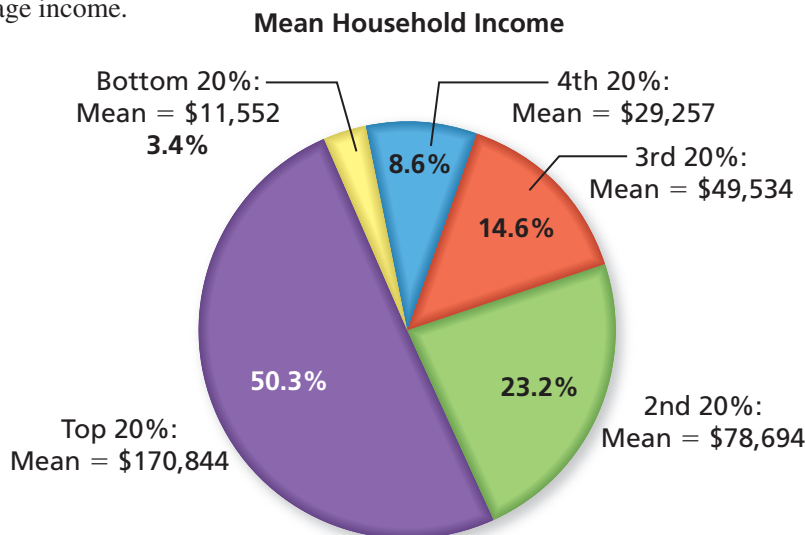
✓ **Checkpoint**

Help at *Math*.and**YOU**.com

Discuss other instances where the median is a better measure of the average value than the mean.

In Malcolm Gladwell's book *Outliers: The Story of Success*, the position of being an "outlier" is enviable. When using the mean as an average, however, outliers can cause undesired effects.
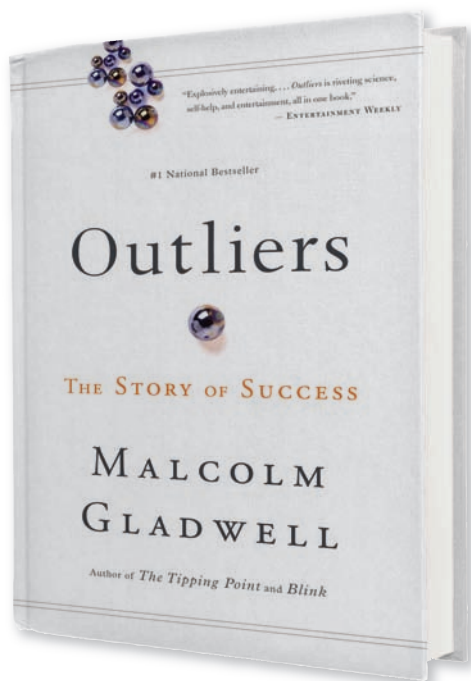
EXAMPLE 6    **Analyzing Data**

The following gives the estimated IQs of 12 recent presidents of the United States. It also gives their placements in the 2010 Siena College Research Institute Presidential Ranking Survey (with 43 being the lowest and 1 being the highest).

| President | IQ | Poll Rating |
|---|---|---|
| Franklin Roosevelt | 140 | 1 |
| Harry Truman | 128 | 9 |
| Dwight Eisenhower | 132 | 10 |
| John Kennedy | 151 | 11 |
| Lyndon Johnson | 128 | 16 |
| Richard Nixon | 131 | 30 |
| Gerald Ford | 127 | 28 |
| Jimmy Carter | 145 | 32 |
| Ronald Reagan | 130 | 18 |
| George H. Bush | 130 | 22 |
| William Clinton | 149 | 13 |
| George W. Bush | 125 | 39 |

Assuming these data are valid, which of the following statements are valid?

**a.** On average, U.S. presidents have above average IQs.

**b.** As president, the higher your IQ, the more popular you will be.

### SOLUTION

**a.** This statement is certainly valid. Standardized IQ scores follow a distribution with an average IQ of 100. The median of the above IQs is about 131. The mean is about 135.

**b.** The scatter plot compares estimated IQs with poll ratings. If the statement were true, the scatter plot would show a pattern that moved from the lower left to the upper right. So, according to the above list, this conclusion is not valid.



Poll Ratings and IQs of 12 Recent U.S. Presidents
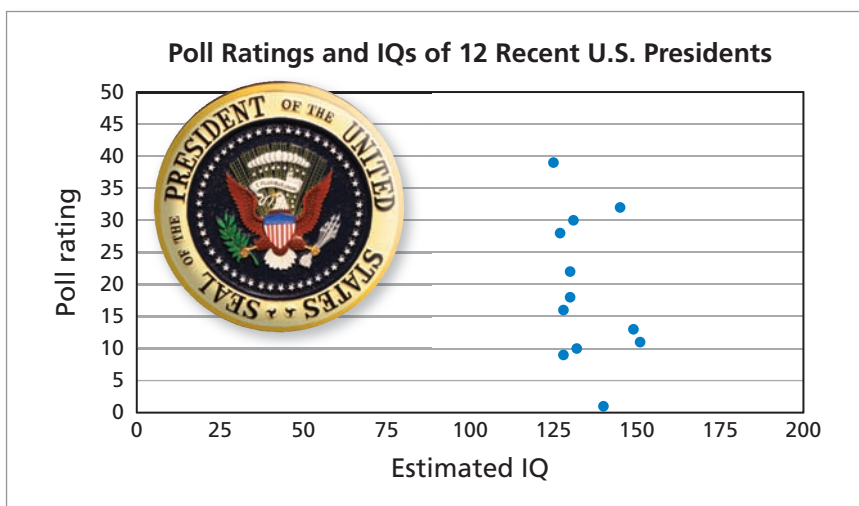
✓ **Checkpoint**                    Help at *Math*.and**YOU**.com

Suppose a president with an IQ of 200 is elected. How does this affect the mean and the median of the data?

# 9.2  Exercises

**Height**  The double stem-and-leaf plot shows the heights of players on a women's basketball team and players on a men's basketball team. In Exercises 1–8, use the double stem-and-leaf plot. *(See Examples 1 and 2.)*

**Player Height**

| Women's Team | Stem | Men's Team |
|---|---|---|
| 9  8  5 | 6 | |
| 7  6  5  5  4  4  3  2 | 7 | 2  4  4  6  7  7  8  9  9 |
| | 8 | 0  1  2  3  4  4  5 |

**Key:**

Women's:   5|6| = 65-inch-tall woman

Men's:   |7|2 = 72-inch-tall man

1. Compare the mean of each data set.

2. What is the mode of each data set?

3. Compare the median of each data set.

4. What percent of the players on the men's team are at least 7 feet tall?

5. Which measure of central tendency best describes the average height of the players on the women's basketball team?

6. Which measure of central tendency best describes the average height of the players on the men's basketball team?

7. Suppose each player's height on the women's team is 2 inches greater. How does this affect the mean and the median of the data?

8. The 5 starting players on the men's team have a mean height of 80.2 inches. Find the heights of 5 possible starters.

**Annual Cost of College**  The box-and-whisker plot shows the annual cost for undergraduate tuition, room, and board at 21 public and private colleges. In Exercises 9–12, use the box-and-whisker plot. *(See Example 3.)*

**Annual Cost of College**

Minimum: $4850   Median: $16,200                    Maximum: $41,350

—Mean: $18,713

Lower quartile: $12,500        Upper quartile: $24,500

$0        $10,000        $20,000        $30,000        $40,000        $50,000
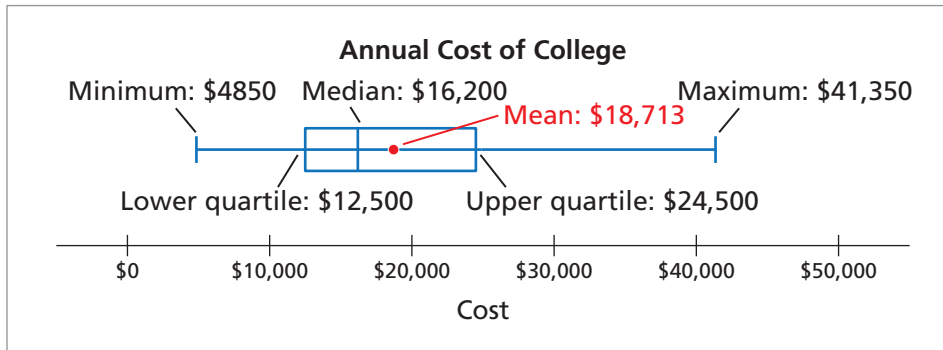
Cost

9.  How many colleges have an annual cost between $12,500 and $24,500?

10. Are there more colleges above or below the mean annual cost? Explain your reasoning.

11. The annual costs are divided into four quartiles, each of which contains about 25% of the data. In which quartile are the data most spread out? least spread out? Explain your reasoning.

12. The annual cost to attend college A is $10,780. In which quartile does this college belong?

**Annual Cost of College**  The color-coded histogram shows the data from the box-and-whisker plot above in a different way. In Exercises 13–16, use the histogram. *(See Example 4.)*

**Annual Cost of College**

Public 4 yr =
Public 2 yr =
Private 4 yr =
Private 2 yr =

$0        $10,000        $20,000        $30,000        $40,000        $50,000
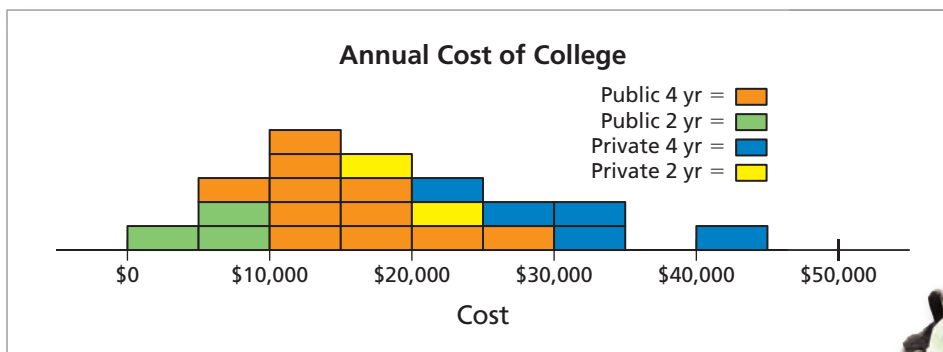
Cost

13. Can you use the histogram to find the mean, median, or mode of the annual costs? Explain your reasoning.

14. Which annual cost interval contains the most data values?

15. How many of the colleges have an annual cost less than $20,000?

16. Which type of college seems to be the most expensive? Explain your reasoning.

**Deadliest Earthquakes**  The double line graph shows the world's deadliest earthquake for each year from 1990 through 2010. In Exercises 17–22, use the double line graph.  *(See Examples 5 and 6.)*



**Deadliest Earthquakes (1990–2010)**
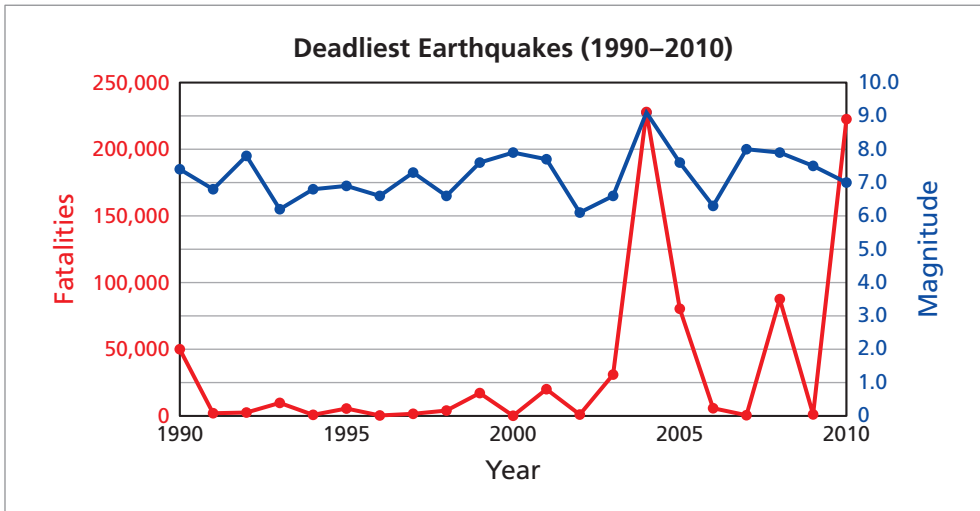
**17.** Which years appear to be outliers when considering the numbers of fatalities?

**18.** Which years appear to be outliers when considering the magnitudes of the earthquakes?

**DATA 19.** Find the effect of outliers on averages.

    **a.** Use a spreadsheet to find the mean and the median number of fatalities.

    **b.** Use a spreadsheet to find the mean and the median number of fatalities without the two most significant outliers.

    **c.** Do these outliers have a greater effect on the mean or the median? Explain your reasoning.

**20.** How do outliers affect the mode of a data set?

**21.** Does the mean or the median number of fatalities better describe the average number of fatalities each year? Explain your reasoning.

**22.** Determine whether each statement is valid.

    **a.** On average, the most fatal earthquakes have a magnitude around 6.0.

    **b.** The number of fatalities from an earthquake depends on the earthquake's magnitude.

# ▶ Extending Concepts

**Printer Speeds**  **A consumer testing service is determining the printing speed, in pages per minute (ppm), of three printers. The table shows the results of five print jobs for each printer. In Exercises 23 and 24, use the table.**

|  | Job 1 | Job 2 | Job 3 | Job 4 | Job 5 |
|---|---|---|---|---|---|
| **Printer A (ppm)** | 28 | 32 | 30 | 27 | 32 |
| **Printer B (ppm)** | 30 | 28 | 34 | 28 | 32 |
| **Printer C (ppm)** | 28 | 31 | 32 | 27 | 31 |

**23.**  Each printer has a different manufacturer, and each manufacturer advertises that its printer has a faster average printing speed than its competitors' printers. Which measure of central tendency (mean, median, or mode) would each of the following manufacturers use to support its claim? Explain your reasoning.

   **a.**  the manufacturer of printer A

   **b.**  the manufacturer of printer B

   **c.**  the manufacturer of printer C

**24.**  The midrange is $\dfrac{(\text{maximum data value}) + (\text{minimum data value})}{2}$ .

   **a.**  Which printer has the fastest average speed according to the midrange?

   **b.**  Why do you think the midrange is rarely used as a measure of central tendency?

**Printer Prices**  **The table below shows the prices of all the printers in stock at an electronics store. In Exercises 25–28, use the table.**

| Prices of Printers in Stock | | | | | |
|---|---|---|---|---|---|
| $45 | $89 | $72 | $118 | $68 | $89 |
| $105 | $76 | $95 | $182 | $120 | $48 |
| $68 | $89 | $128 | $52 | $108 | $99 |

**25.**  Make a box-and-whisker plot of the printer prices.

**26.**  Describe the shape of the distribution.

**27.**  Make a histogram of the printer prices.

**28.**  What new information does the histogram give you about the distribution?

## 9.1–9.2 Quiz

**World Education Rankings** In Exercises 1–8, use the chart and the following information.
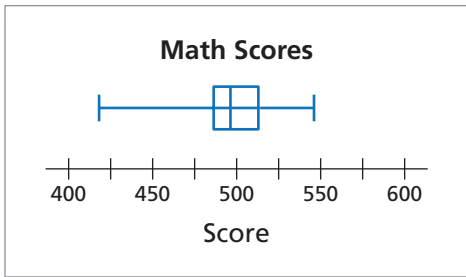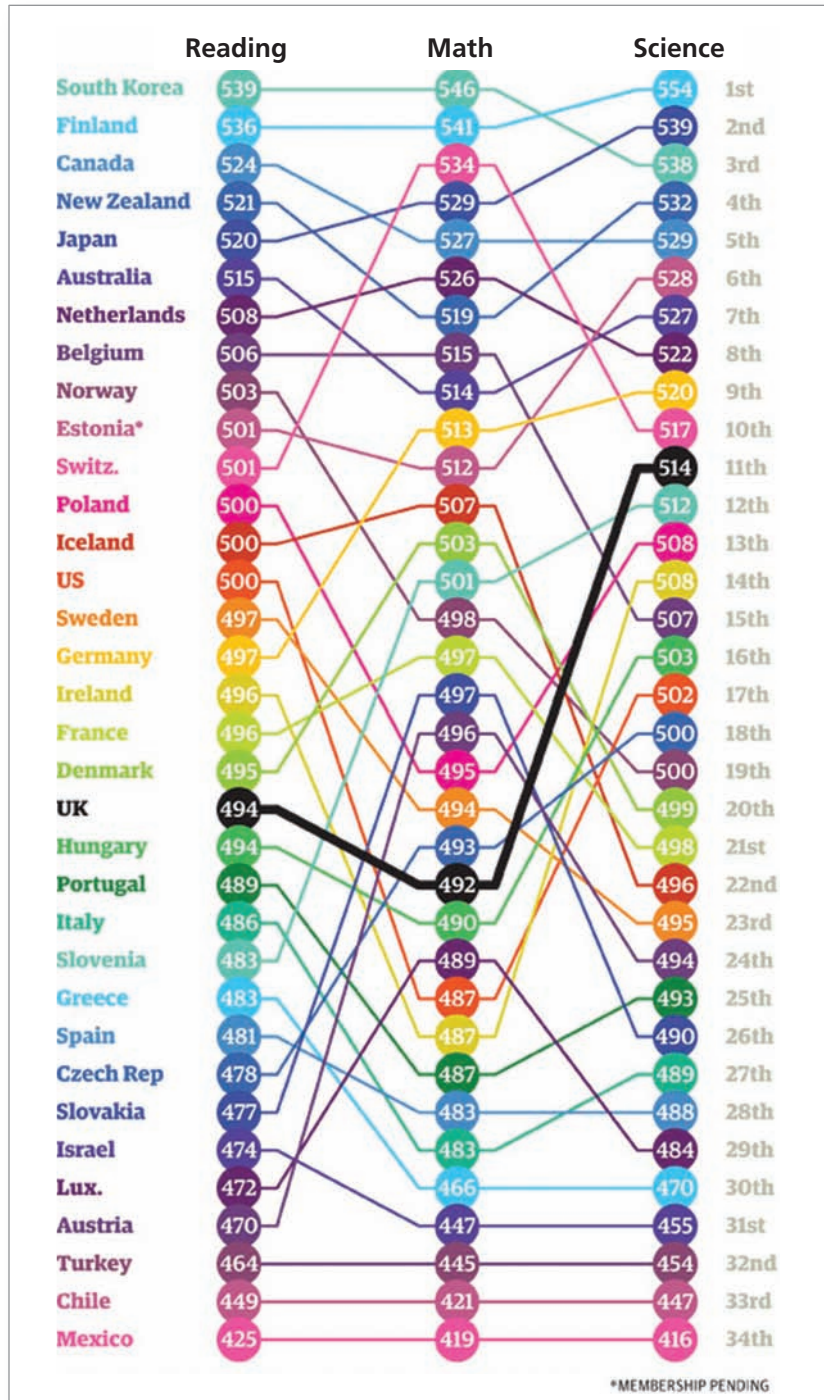
The chart shows the world education rankings from the Organisation for Economic Co-operation and Development (OECD).

1. Describe the information presented in the rankings chart.

2. Are the countries listed in order of greatest total score? Explain your reasoning.

3. Do you agree with the following comment? Explain your reasoning.

   COMMENT:

   Overall, the United States scored higher than Germany.

4. Did Austria or Poland score higher in math? Explain your reasoning.

5. Do the following for each column in the chart.
   a. Use the chart to find the median and mode.

   DATA   b. Use a spreadsheet to find the mean.

6. Which measure of central tendency best describes the average of the math scores? Explain your reasoning.

7. Describe the distribution of math scores.

**Math Scores**

(box plot, Score axis: 400, 450, 500, 550, 600)

| | Reading | Math | Science | |
|---|---|---|---|---|
| South Korea | 539 | 546 | 554 | 1st |
| Finland | 536 | 541 | 539 | 2nd |
| Canada | 524 | 534 | 538 | 3rd |
| New Zealand | 521 | 529 | 532 | 4th |
| Japan | 520 | 527 | 529 | 5th |
| Australia | 515 | 526 | 528 | 6th |
| Netherlands | 508 | 519 | 527 | 7th |
| Belgium | 506 | 515 | 522 | 8th |
| Norway | 503 | 514 | 520 | 9th |
| Estonia* | 501 | 513 | 517 | 10th |
| Switz. | 501 | 512 | 514 | 11th |
| Poland | 500 | 507 | 512 | 12th |
| Iceland | 500 | 503 | 508 | 13th |
| US | 500 | 501 | 508 | 14th |
| Sweden | 497 | 498 | 507 | 15th |
| Germany | 497 | 497 | 503 | 16th |
| Ireland | 496 | 497 | 502 | 17th |
| France | 496 | 496 | 500 | 18th |
| Denmark | 495 | 495 | 500 | 19th |
| UK | 494 | 494 | 499 | 20th |
| Hungary | 494 | 493 | 498 | 21st |
| Portugal | 489 | 492 | 496 | 22nd |
| Italy | 486 | 490 | 495 | 23rd |
| Slovenia | 483 | 489 | 494 | 24th |
| Greece | 483 | 487 | 493 | 25th |
| Spain | 481 | 487 | 490 | 26th |
| Czech Rep | 478 | 487 | 489 | 27th |
| Slovakia | 477 | 483 | 488 | 28th |
| Israel | 474 | 483 | 484 | 29th |
| Lux. | 472 | 466 | 470 | 30th |
| Austria | 470 | 447 | 455 | 31st |
| Turkey | 464 | 445 | 454 | 32nd |
| Chile | 449 | 421 | 447 | 33rd |
| Mexico | 425 | 419 | 416 | 34th |

*MEMBERSHIP PENDING

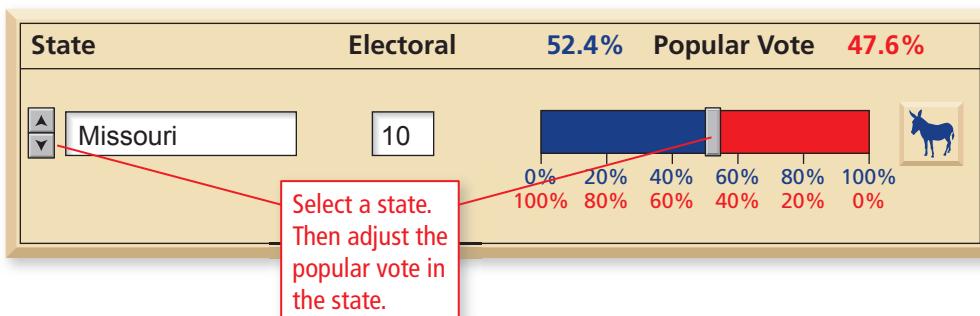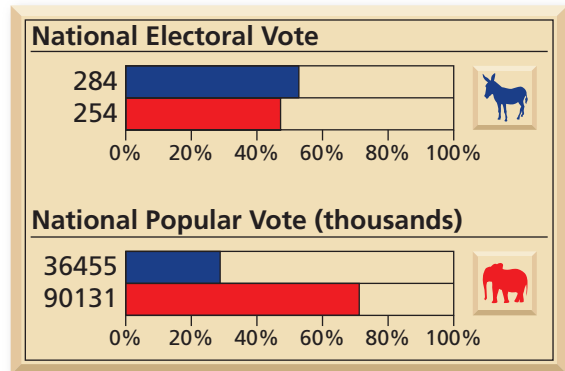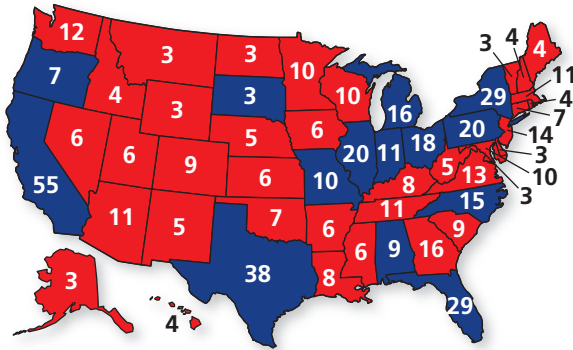8. How might a politician or policymaker use this chart?

# Math & the Electoral College
## PROJECT: Modeling a Presidential Election

The Electoral College consists of the electors who formally elect the president of the United States. Since the 1964 election, there have been 538 electors. The map below shows the number of electors for each state. These numbers were determined by the 2010 Census. If a candidate wins the popular vote in a state, he or she gets all the electoral votes for that state.

1. Use the *Electoral College Simulator* at *Math.andYou.com*. It is possible to win the presidential election and not win the popular vote. This has happened four times in the United States (see page 114). The simulator below shows an imaginary election in which the winner receives 52.8% of the electoral votes, but only 28.8% of the popular vote. Use the simulator to find the minimum popular vote that a candidate can receive and still win.

**National Electoral Vote**

284
254

0%   20%   40%   60%   80%   100%

**National Popular Vote (thousands)**

36455
90131

0%   20%   40%   60%   80%   100%

| State | Electoral | **52.4%** | Popular Vote | **47.6%** |
|---|---|---|---|---|

Missouri    10

0%    20%   40%   60%   80%   100%
100%   80%   60%   40%   20%   0%

Select a state. Then adjust the popular vote in the state.

2. Is it possible to become president and not have a single person west of the Mississippi River vote for you? Use the simulator to verify your answer.

3. Is it possible to win the election without carrying one of the "big" states (California, Texas, Florida, New York, Pennsylvania, Illinois, Ohio, Michigan, Georgia)? Use the simulator to verify your answer.

4. Why was the Electoral College created? Do you agree with this system? If not, describe a different system that you think would be better. For instance, do you think the system used in Canada is better? Explain your reasoning.

# 9.3 Describing Dispersion

▶ Use standard deviation to describe the dispersion of a data set.

▶ Use standard deviation to describe a data set that is normally distributed.

▶ Compare different types of distributions.

## Standard Deviation

**Standard deviation** is a measurement that shows how much variation or dispersion there is from the mean. A small standard deviation indicates that the data are clustered tightly around the mean. A large standard deviation indicates that the data are spread out over a large range of values.

### Standard Deviation and Dispersion

In any data set, at least 75% of the values lie within 2 standard deviations of the mean. At least 89% lie within 3 standard deviations, and at least 94% lie within 4 standard deviations.

**Study Tip**

In statistics, the rule at the right is called Chebyshev's inequality. Its estimates are conservative. On page 426, you will see a stronger result for data sets that follow a normal distribution.

**EXAMPLE 1** **Describing Dispersion**

The histogram shows the distribution of the lengths of 45 cuckoo eggs.

**a.** Describe the dispersion.

**b.** What percent of the lengths lie within 2 standard deviations of the mean?

DATA

**Cuckoo Eggs**

Mean = 22 mm
Standard deviation = 1 mm

Frequency (vertical axis): 0, 2, 4, 6, 8, 10, 12
Egg length (mm) (horizontal axis): 20, 21, 22, 23

Rather than building its own nest, the common cuckoo lays its eggs in the nests of other birds. When it hatches, the cuckoo chick eventually pushes the other eggs and chicks from the nest.
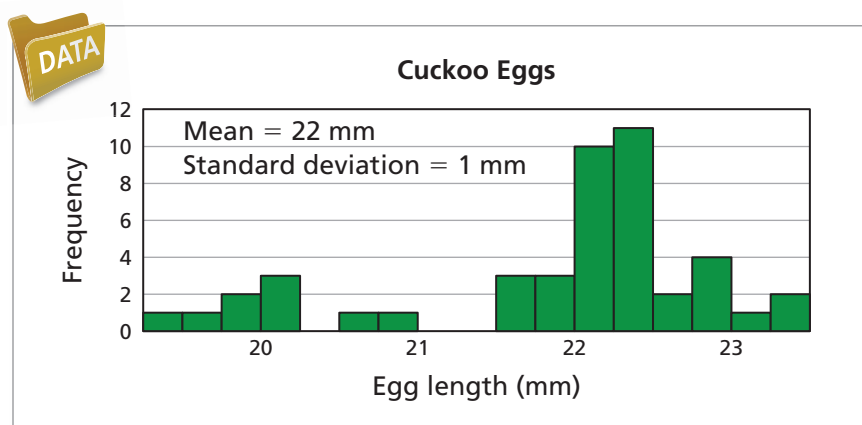
## SOLUTION

**a.** 27 of the lengths lie within $\frac{1}{2}$ of a standard deviation of the mean.

**b.** 41 of the lengths, or about 91%, lie within 2 standard deviations of the mean.

✓ **Checkpoint**

Help at *Math*.and**YOU**.com

The data set for Example 1 is available at *Math.andYou.com*. Use the *Histogram Generator* at *Math.andYou.com* to display the data. Experiment with different interval widths. The histogram above uses an interval width of 0.25. Try using an interval width of 0.5.

EXAMPLE 2    **Describing Dispersion**

The histogram shows the distribution of the mean daily temperatures at McGuire Air Force Base from 1955 through 2010. Describe the dispersion.

**Mean Daily Temperature (1955–2010)**

Mean = 53.8
Standard deviation = 17.2
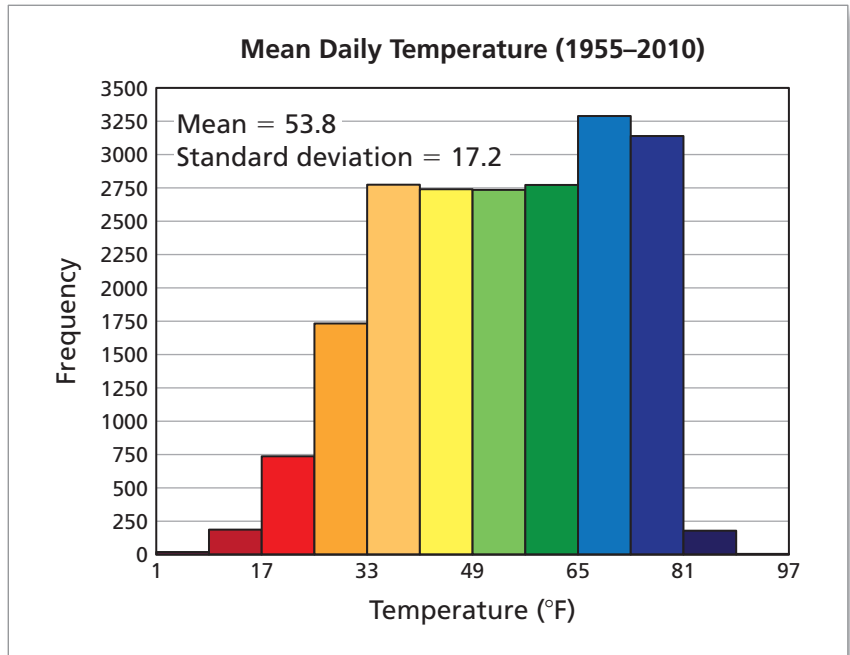


**SOLUTION**

The vast majority of the temperatures lie within 1 standard deviation of the mean. That is, they range from 36.6°F to 71°F.
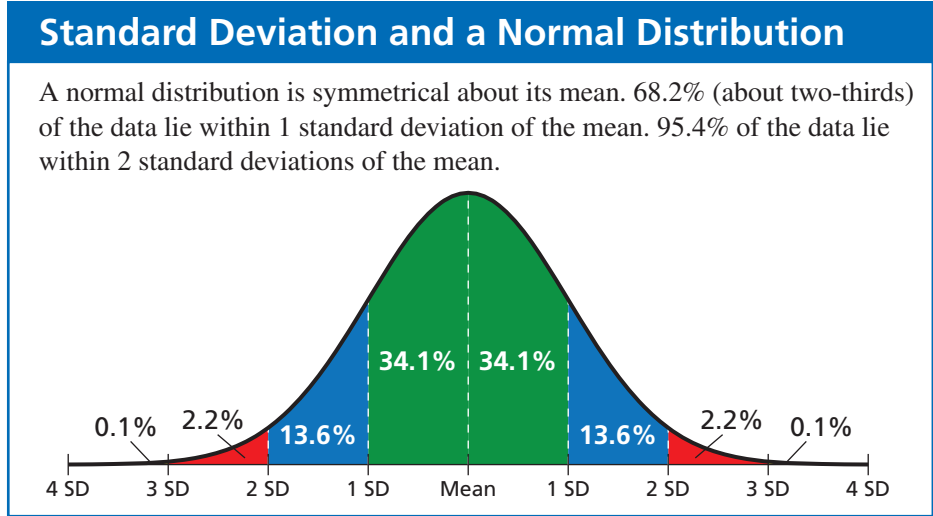
✓ **Checkpoint**                    Help at *Math*.and**Y⊕U**.com

The line graph shows 101-day moving averages of the data represented in the histogram. The temperatures are plotted from 1955 through 2010. Suppose you are doing a study on global warming. Which of the two graphic displays would be more helpful? Explain your reasoning.

**101-Day Moving Averages**
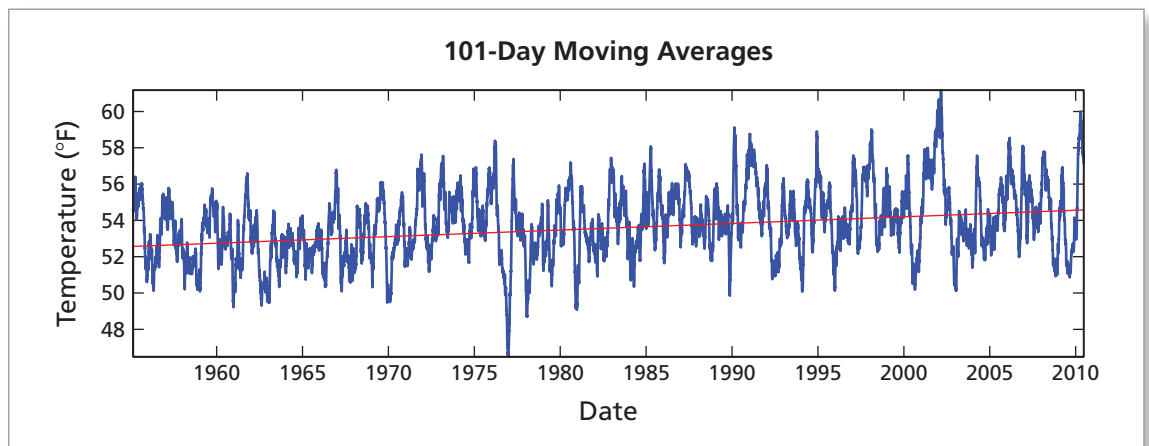
## Standard Deviation and Normal Distribution

In many naturally occurring data sets, a histogram of the data is often bell shaped. In statistics, such data sets are said to have a **normal distribution.**

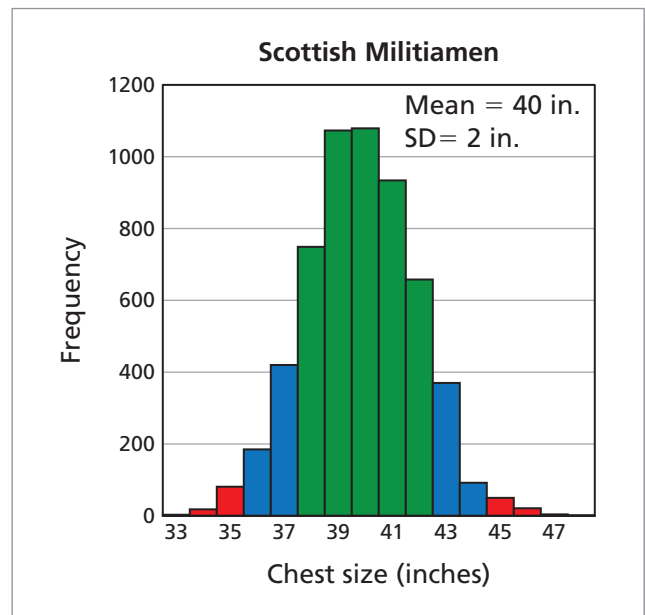### Standard Deviation and a Normal Distribution

A normal distribution is symmetrical about its mean. 68.2% (about two-thirds) of the data lie within 1 standard deviation of the mean. 95.4% of the data lie within 2 standard deviations of the mean.



| | | |
|---|---|---|
| 0.1% | 2.2% | 13.6% 34.1% 34.1% 13.6% 2.2% 0.1% |

4 SD   3 SD   2 SD   1 SD   Mean   1 SD   2 SD   3 SD   4 SD

---

**EXAMPLE 3**   **Analyzing a Famous Normal Distribution**

A famous data set was collected in Scotland in the mid-1800s. It contains the chest sizes (in inches) of 5738 men in the Scottish Militia. What percent of the chest sizes lie within 1 standard deviation of the mean?

| Chest Size | Number of Men |
|---|---|
| 33 | 3 |
| 34 | 18 |
| 35 | 81 |
| 36 | 185 |
| 37 | 420 |
| 38 | 749 |
| 39 | 1073 |
| 40 | 1079 |
| 41 | 934 |
| 42 | 658 |
| 43 | 370 |
| 44 | 92 |
| 45 | 50 |
| 46 | 21 |
| 47 | 4 |
| 48 | 1 |



Scottish Militiamen

Mean = 40 in.
SD = 2 in.

*The Thin Red Line* is a painting by Robert Gibb. It was painted in 1881. Only the left portion of the painting is shown above.

### SOLUTION

The number of chest sizes within 1 standard deviation of the mean is 749 + 1073 + 1079 + 934 + 658 = 4493. This is about 78.3% of the total, which is somewhat more than the percent predicted by a normal distribution.
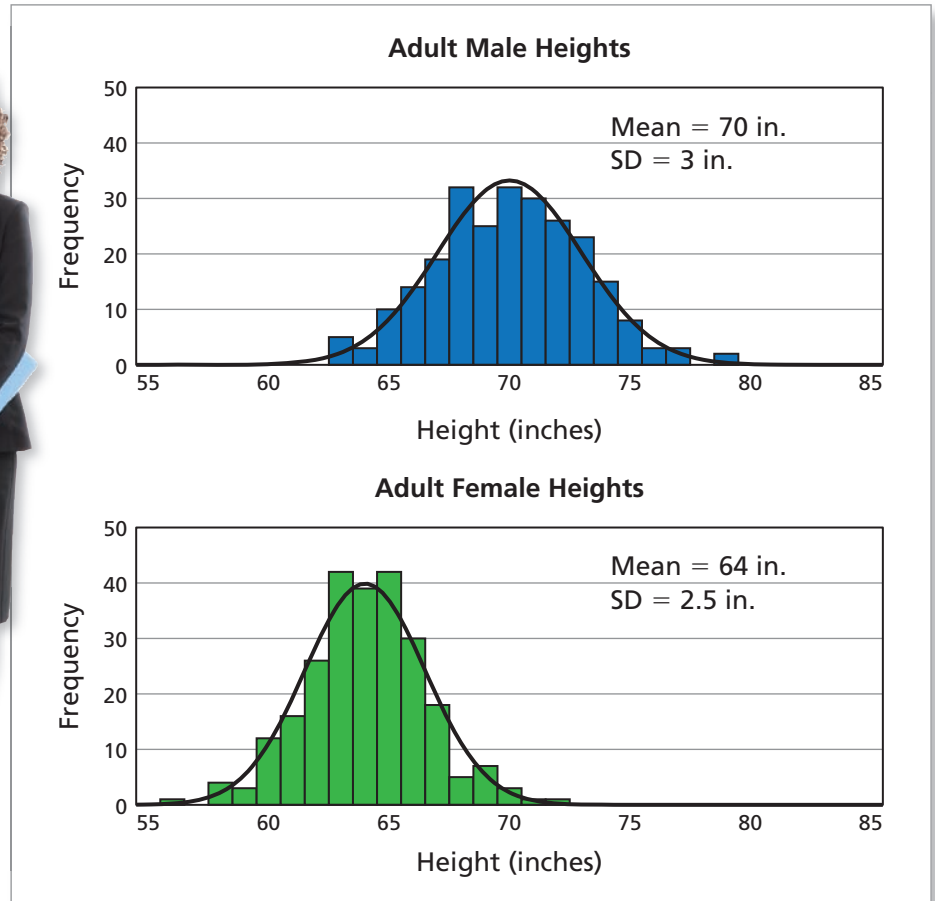
✓ **Checkpoint**

Help at *Math*.and**YOU**.com

What percent of the chest sizes lie within 2 standard deviations of the mean?

EXAMPLE 4    **Comparing Distributions**

The graphs show the distributions of samples of heights of adult American males and females. There are 250 people in each sample.

**a.** What is the significance of the smaller standard deviation for females?

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

**Adult Male Heights**

Mean = 70 in.
SD = 3 in.

Height (inches)

**Adult Female Heights**

Mean = 64 in.
SD = 2.5 in.

Height (inches)

**SOLUTION**

**a.** Standard deviation is a measure of dispersion. The larger the standard deviation, the more the data are spread out. So, if these two samples are representative of male and female heights in the United States, you can conclude that male heights have a greater variation than female heights.

**b.** The data for male heights appear to be normally distributed. If this is true, then you can conclude that about 68% of adult male heights are between 67 inches (5' 7") and 73 inches (6' 1").
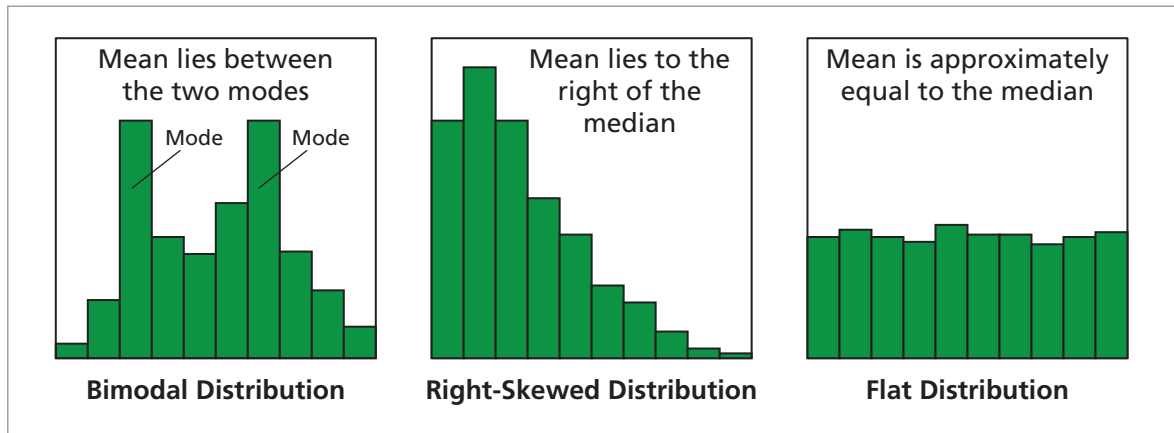
✓ **Checkpoint**

Help at *Math*.and**YOU**.com

In a study, Timothy Judge and Daniel Cable found that each additional inch of height is worth an extra $789 per year in salary. According to Judge, "Height matters for career success." Do you agree with Judge's claim? Explain.

## Comparing Different Types of Distributions

While normal distributions (bell-shaped) do occur commonly in real life, there are other types of distributions. Here are three of them.

| Mean lies between the two modes | Mean lies to the right of the median | Mean is approximately equal to the median |
|---|---|---|
| Mode   Mode | | |
| **Bimodal Distribution** | **Right-Skewed Distribution** | **Flat Distribution** |

EXAMPLE 5   **Comparing Distributions**

The histograms show the distributions of game lengths (in number of turns) when there are two players. Describe the differences in the distributions.

**Normalized Distribution of Game Lengths**

Cootie

Chutes and Ladders

Candyland

Number of turns
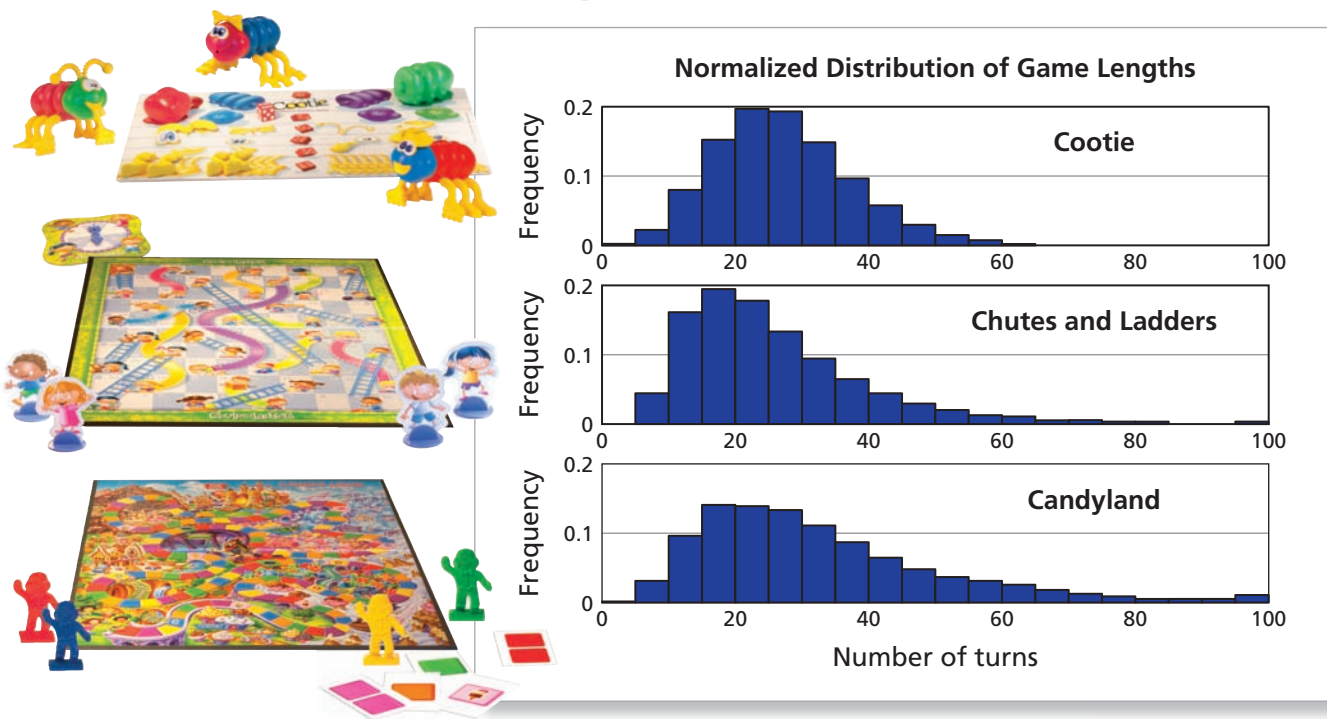
### SOLUTION

*Candyland* tends to take the most turns to finish. *Chutes and Ladders* tends to take the least.

✓ **Checkpoint**     Help at *Math*.and**Y⊙U**.com

Suppose you work in the marketing department for a game company. How would you use the information in Example 5 to target different age groups? Explain.

EXAMPLE 6    **Analyzing Bimodal Distribution**

The graph shows the distribution of the full-time salaries of 22,665 people who graduated from law school in 2006. How would you explain the bimodal distribution?

**Full-Time Salaries**

← Median Salary $62,000

(Graph: Percent vs Annual Salary, bimodal distribution with peaks near 45,000 and 135,000)
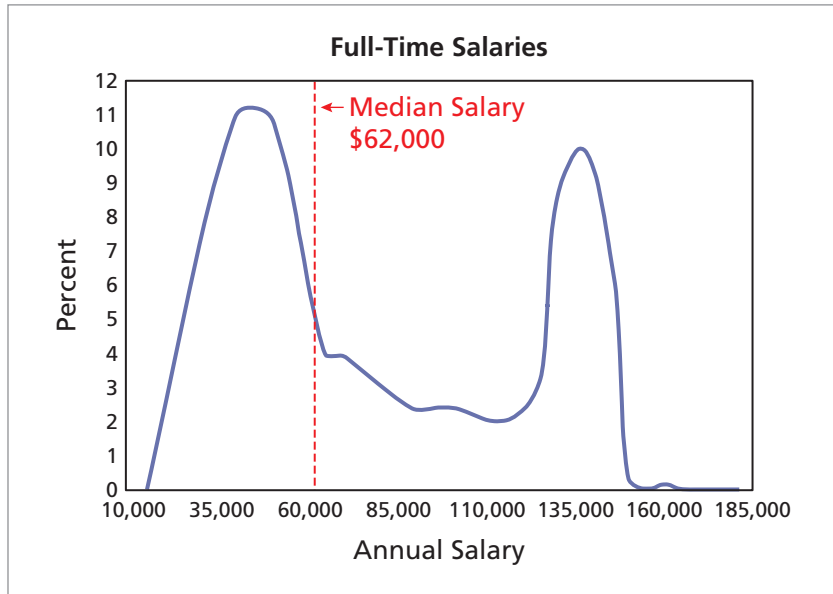
Employed graduates from the law school class of 2006 took the following types of jobs: private practice (55.8%), business (14.2%), other government (10.6%), judicial clerk (9.8%), public interest (5.4%), academic (1.7%), military (1.1%), and unknown (1.3%).

## SOLUTION

According to the National Association for Law Placement (NALP), 71% of employed graduates from the class of 2006 took jobs at small firms (50 or fewer lawyers) or in nonfirm settings. Only 20% took jobs at large firms (more than 100 lawyers).

The bimodal distribution represents a cluster of lawyers at small and midsize firms earning between $40,000 and $50,000 and a cluster of lawyers at large firms earning between $135,000 and $145,000.
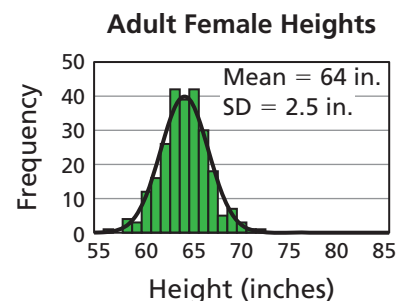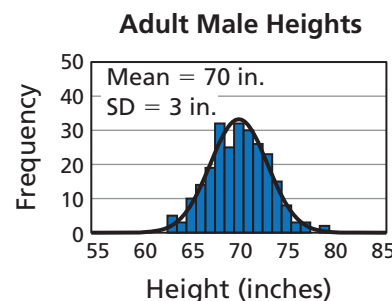
✓ **Checkpoint**                                      Help at *Math*.and**Y●U**.com

Suppose that you combine the data from the two data sets in Example 4 into one data set. What will the histogram look like? Explain how you can discover two (or more) subpopulations within a given larger population.

**Adult Male Heights**

Mean = 70 in.
SD = 3 in.

(Histogram: Frequency vs Height (inches), peak near 70 in.)

**Adult Female Heights**

Mean = 64 in.
SD = 2.5 in.

(Histogram: Frequency vs Height (inches), peak near 64 in.)

Law schools are ranked each year. One of the recent rankings by *U.S. News and World Report* listed the following as the top five law schools in the United States (with full-time tuition).

1. Yale University ($50,750)
2. Harvard University ($45,450)
3. Stanford University ($46,581)
4. Columbia University ($50,428)
5. University of Chicago ($45,405)

# 9.3 Exercises

**Rainbow Trout Eggs** The histogram shows the distribution of the diameters of 100 rainbow trout eggs. In Exercises 1–3, use the histogram. *(See Examples 1 and 2.)*

**Rainbow Trout**

Mean = 4.31 mm
SD = 0.25 mm
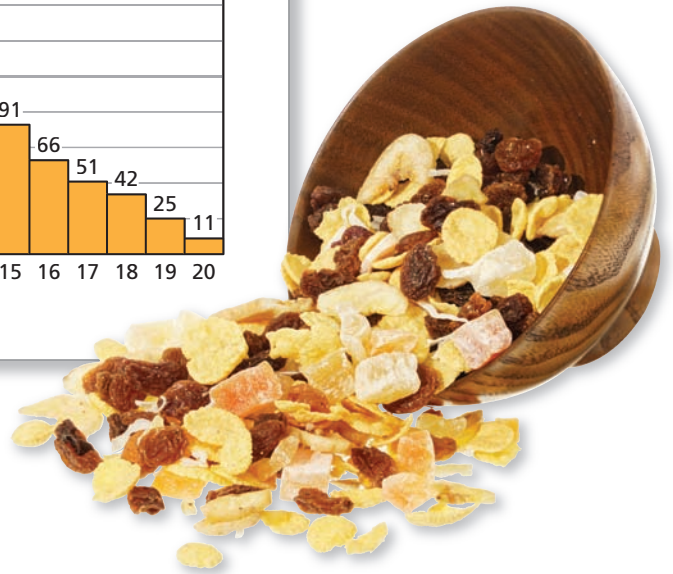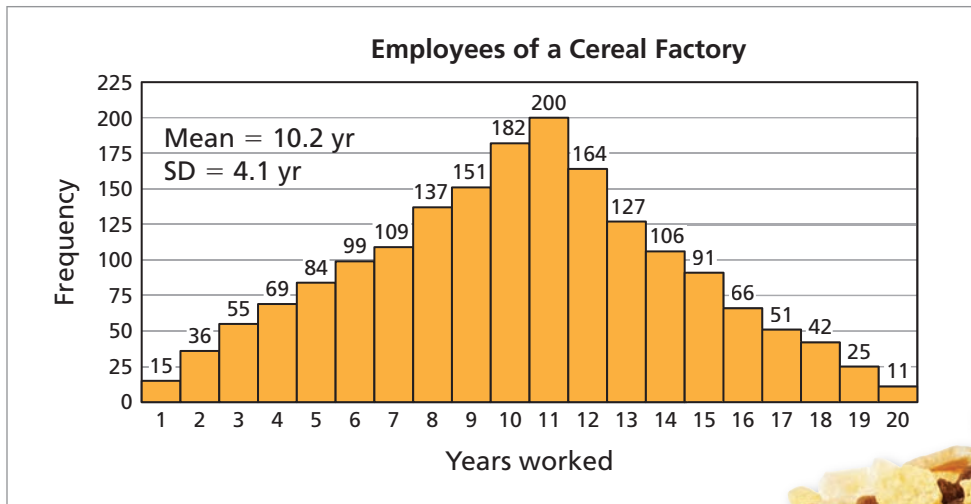
Frequency

Egg diameter (mm)
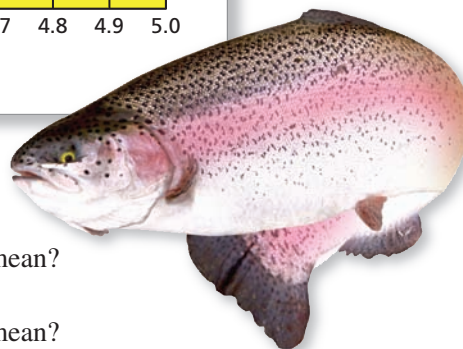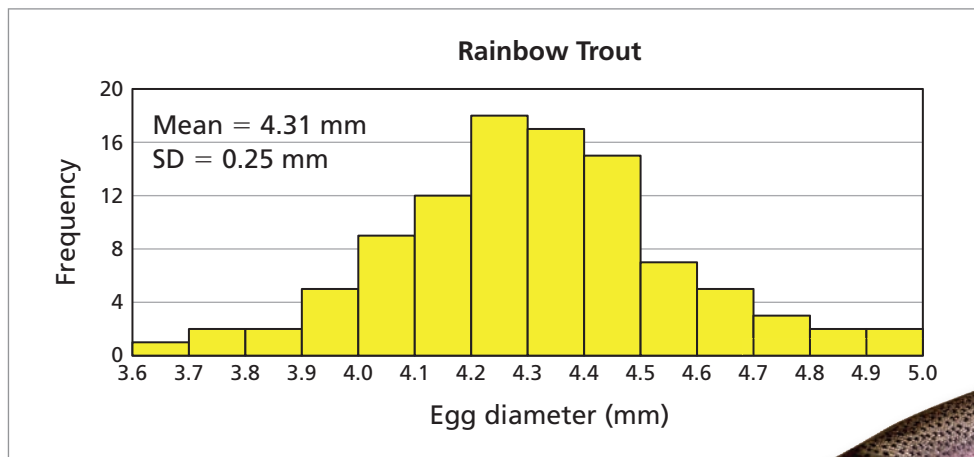
1. Describe the dispersion.

2. What percent of the diameters lie within 2 standard deviations of the mean?

3. What percent of the diameters lie within 3 standard deviations of the mean?

**Temperatures** The histogram shows the distribution of the daily high temperatures in Pittsburgh for 2010. In Exercises 4–6, use the histogram. *(See Examples 1 and 2.)*
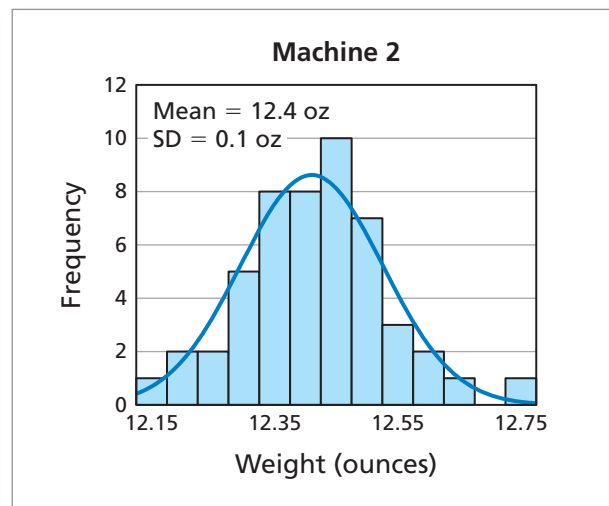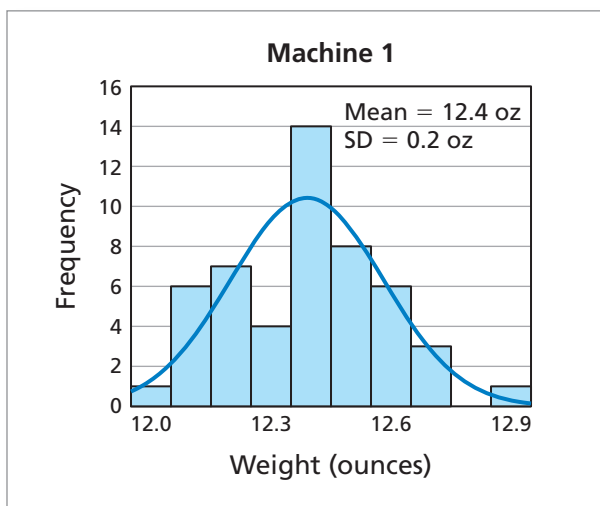
**Pittsburgh Daily High Temperatures**

Mean = 61.3°F
SD = 21.5°F

Frequency
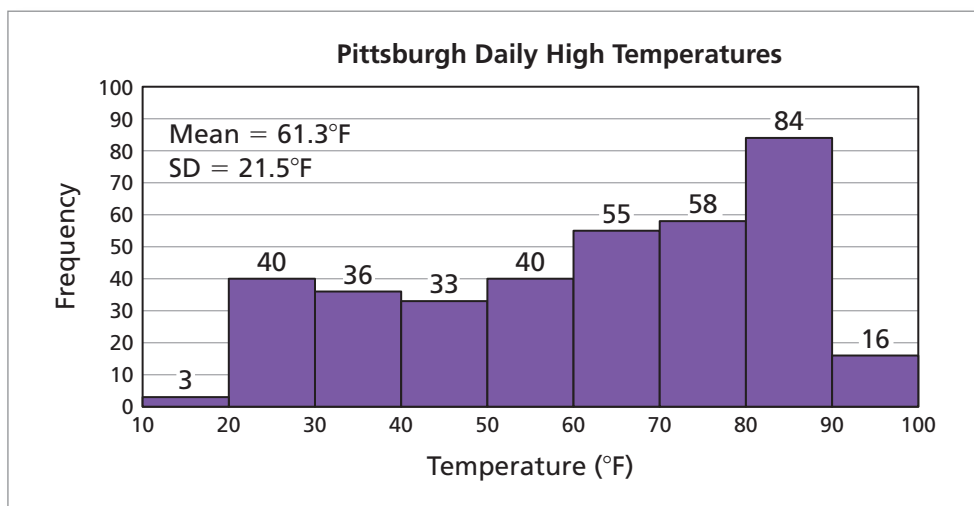
Temperature (°F)

4. Describe the dispersion.

5. What percent of the temperatures lie within 1 standard deviation of the mean?

6. What percent of the tematures lie within 2 standard deviations of the mean?

**Factory Employees** The graph shows the distribution of the years worked by the 1820 employees of a cereal factory. In Exercises 7–10, use the histogram. *(See Example 3.)*
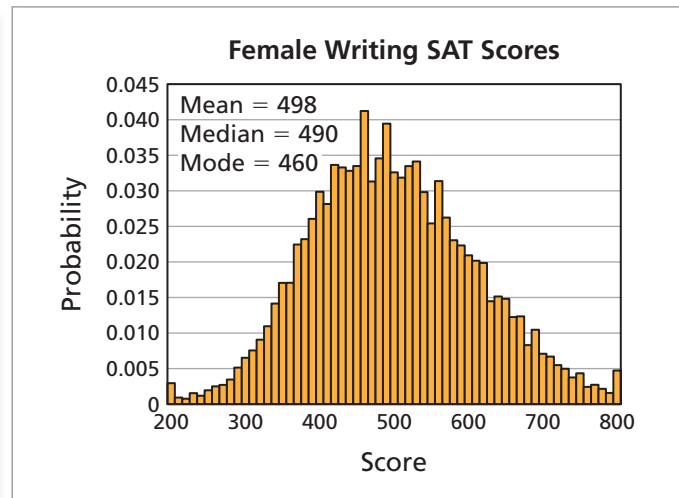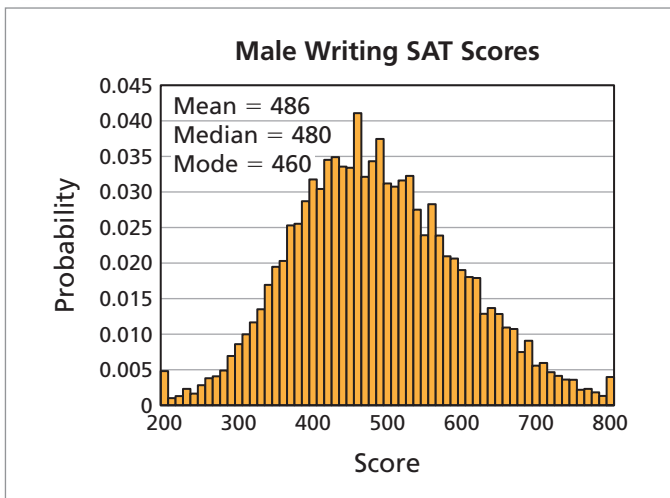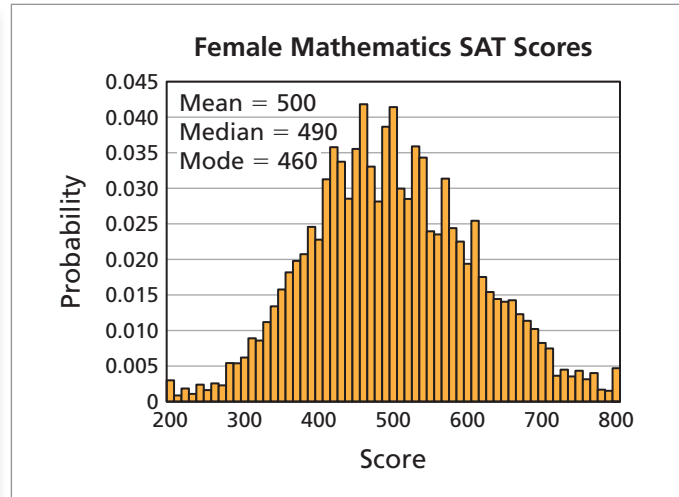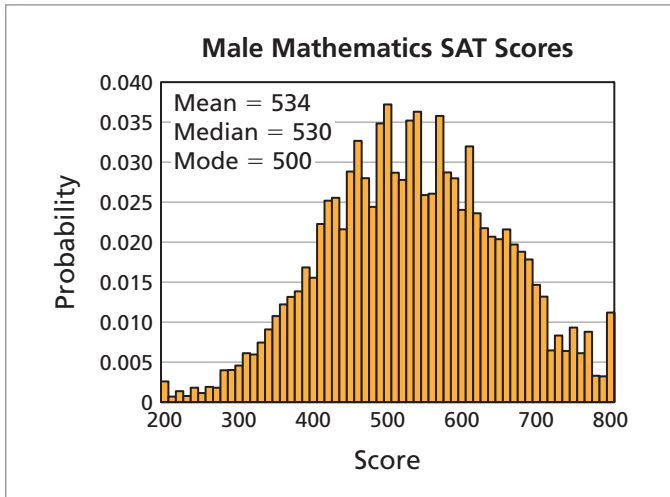
**Employees of a Cereal Factory**

Mean = 10.2 yr
SD = 4.1 yr

Frequency / Years worked

15, 36, 55, 69, 84, 99, 109, 137, 151, 182, 200, 164, 127, 106, 91, 66, 51, 42, 25, 11

7. Does this data set have a normal distribution? Explain your reasoning.

8. What percent of the employees lie within 1 standard deviation of the mean?

9. What percent of the employees lie within 2 standard deviations of the mean?

10. Compare the percents in Exercises 8 and 9 with the percents given by the normal distribution.

**Cereal** The graphs show the distributions of samples of weights of boxes of cereal filled by two machines. In Exercises 11 and 12, use the histograms. *(See Example 4.)*

**Machine 1**

Mean = 12.4 oz
SD = 0.2 oz

Frequency / Weight (ounces)

**Machine 2**

Mean = 12.4 oz
SD = 0.1 oz

Frequency / Weight (ounces)

11. What is the significance of the smaller standard deviation for machine 2?

12. The boxes must weigh at least 12.2 ounces. Estimate the percent of boxes from each machine that pass the weight requirement.

**SAT Scores** The histograms show the distributions of SAT scores for mathematics and writing for males and females in a recent year. In Exercises 13 and 14, use the histograms. *(See Example 5.)*



Male Mathematics SAT Scores

Mean = 534
Median = 530
Mode = 500



Female Mathematics SAT Scores

Mean = 500
Median = 490
Mode = 460



Male Writing SAT Scores

Mean = 486
Median = 480
Mode = 460



Female Writing SAT Scores

Mean = 498
Median = 490
Mode = 460

**13.** Describe the differences in the distributions of mathematics SAT scores for males and females. Describe the differences in the distributions of writing SAT scores for males and females.

**14.** Describe the differences in the distributions of mathematics SAT scores and writing SAT scores for males. Describe the differences in the distributions of mathematics SAT scores and writing SAT scores for females.

**Test Preparation** The histogram shows the distribution of the total numbers of hours students at a school study for the SAT over 12 months. In Exercises 15 and 16, use the histogram. *(See Example 6.)*



SAT Study Hours at a School
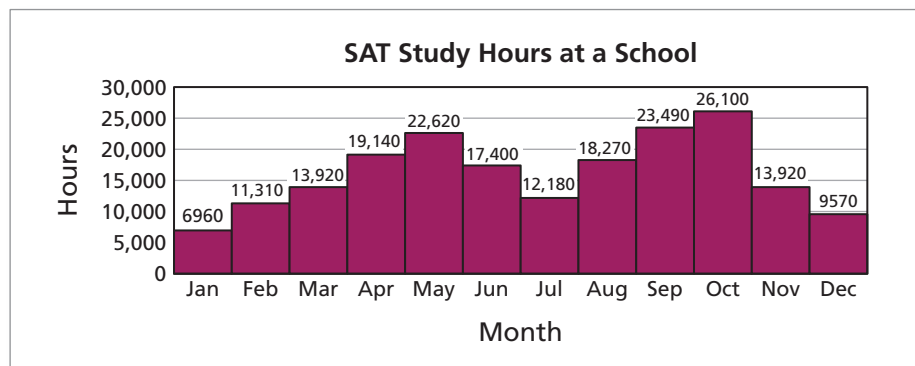
**15.** What are the two modes of this bimodal distribution? How would you explain the bimodal distribution?

**16.** What percent of the study hours are in October?

# ▶ Extending Concepts

**Hodgkin Lymphoma**  The histogram shows the distribution of the ages at which people were diagnosed with Hodgkin lymphoma in a recent year. The table shows the population of the United States for each age group in the same year. In Exercises 17 and 18, use the histogram and the table.
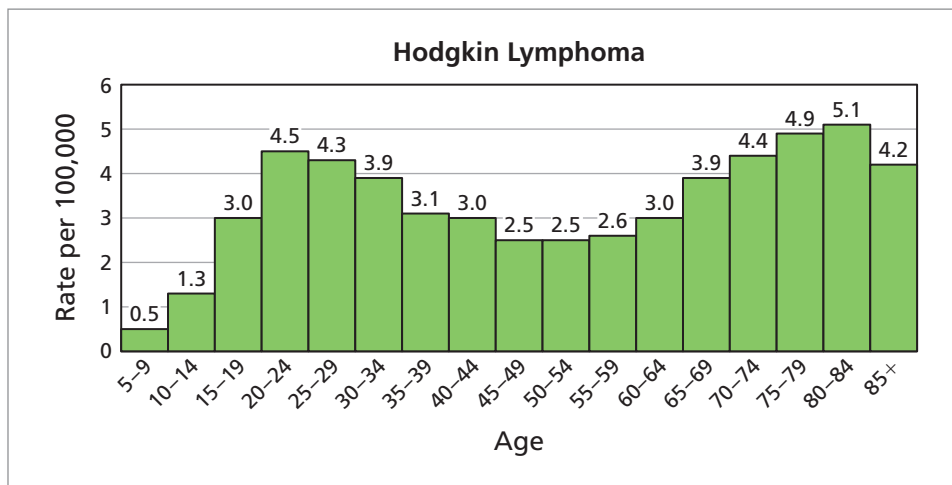
| Age | Population |
|-----|------------|
| 5–9 | 20,313,416 |
| 10–14 | 20,103,526 |
| 15–19 | 21,628,099 |
| 20–24 | 21,322,253 |
| 25–29 | 21,441,834 |
| 30–34 | 19,515,671 |
| 35–39 | 20,846,774 |
| 40–44 | 21,394,190 |
| 45–49 | 22,802,020 |
| 50–54 | 21,431,624 |
| 55–59 | 18,541,377 |
| 60–64 | 15,081,608 |
| 65–69 | 11,332,535 |
| 70–74 | 8,806,448 |
| 75–79 | 7,385,163 |
| 80–84 | 5,825,975 |
| 85+ | 5,449,770 |

**Hodgkin Lymphoma**

(Histogram — Rate per 100,000 vs Age; values: 5–9: 0.5, 10–14: 1.3, 15–19: 3.0, 20–24: 4.5, 25–29: 4.3, 30–34: 3.9, 35–39: 3.1, 40–44: 3.0, 45–49: 2.5, 50–54: 2.5, 55–59: 2.6, 60–64: 3.0, 65–69: 3.9, 70–74: 4.4, 75–79: 4.9, 80–84: 5.1, 85+: 4.2)

**17.** The distribution of the ages in the graph is bimodal. How would you explain this?

**18.** Use the rates and the populations to create a histogram that estimates the number of people diagnosed with Hodgkin lymphoma for each age group. Is the histogram bimodal? Explain.

**Standard Deviation**  The formula for the sample standard deviation of a data set is given by

$$s = \sqrt{\dfrac{\sum x^2 - \dfrac{(\sum x)^2}{n}}{n-1}}$$

where $x$ represents each value in the data set and $n$ is the number of values in the data set. The symbol $\Sigma$ indicates a sum of values. So, $\sum x$ is the sum of the data values, and $\sum x^2$ is the sum of the squared data values. In Exercises 19 and 20, complete the table and then use the formula to find the sample standard deviation of the data set.

**19.**

| $x$ | $x^2$ |
|-----|-------|
| 16 | |
| 12 | |
| 23 | |
| 20 | |
| 18 | |
| 15 | |
| 18 | |
| 19 | |
| $\sum x =$ | $\sum x^2 =$ |

**20.**

| $x$ | $x^2$ |
|-----|-------|
| 33 | |
| 24 | |
| 27 | |
| 26 | |
| 30 | |
| 29 | |
| 26 | |
| 31 | |
| 34 | |
| 20 | |
| $\sum x =$ | $\sum x^2 =$ |

# 9.4 Describing by Sampling

▶ Use a randomly chosen sample to describe a population.

▶ Determine whether a sample is representative of a population.

▶ Determine a sample size to obtain valid inferences.

## Inferring from a Sample

You can best describe any population when you have data for the entire population. Every 10 years, the U.S. Census Bureau attempts to do this. It is costly. The 2010 Census cost the United States $13.1 billion!

A complete census is often unpractical. So, governments, researchers, and businesses attempt to describe populations by taking a **representative sample.** You can be assured that a sample is representative if it is *randomly chosen* and *large enough*.

**Study Tip**

To obtain smaller margins of error, increase the size of the sample.

### Inferring from a Sample

If a sample is randomly taken from a population, then the sample mean can be used to estimate the population mean, given the following limitations.

Estimated mean of population

Margin of error

90%, 95%, or 99% confidence level
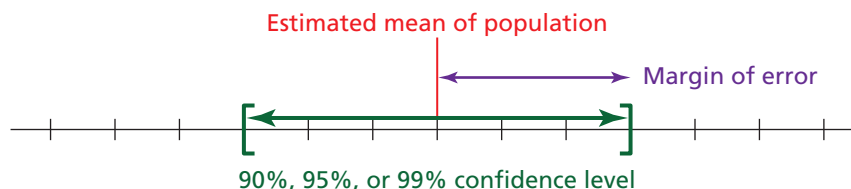
**EXAMPLE 1** **Estimating a Population Mean by Sampling**

The histogram shows the distribution of the shoulder heights of a sample of 40 male black bears in Great Smoky Mountains National Park. Use a 90% confidence level and a 95% confidence level to estimate the population mean shoulder height.

**Male Black Bears**

Mean = 80 cm
SD = 12 cm

Shoulder height (cm)

Shoulder height
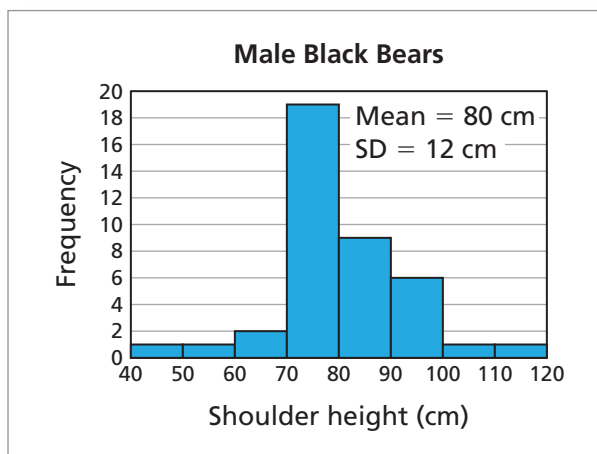
### SOLUTION

Using a confidence interval calculator, you can obtain the following.

● **90% Confidence level:** The population mean is 80 cm ± 3.1 cm.

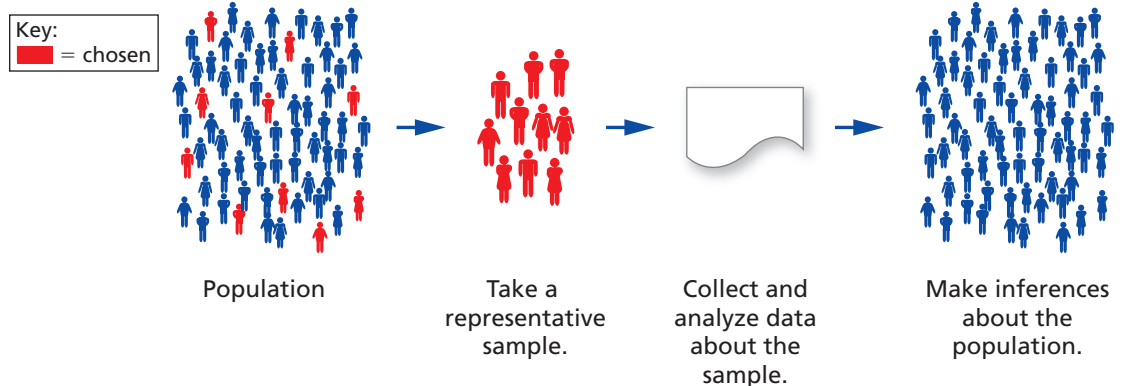● **95% Confidence level:** The population mean is 80 cm ± 3.7 cm.

✓ **Checkpoint**

Help at *Math*.and**YOU**.com

Use the *Confidence Interval Calculator* at *Math.andYou.com* and a 99% confidence level to estimate the population mean shoulder height.
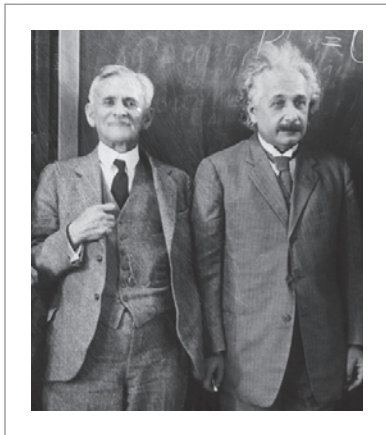
Here is the basic idea of statistical sampling.

1. Take a representative sample of the population.
2. Find the mean and the standard deviation of the sample.
3. Determine how confident you want to be of your *inference* (90%, 95%, 99%).
4. Use a confidence interval calculator to determine the margin of error.
5. *Infer* that the mean of the population is equal to the mean of the sample (± the margin of error).
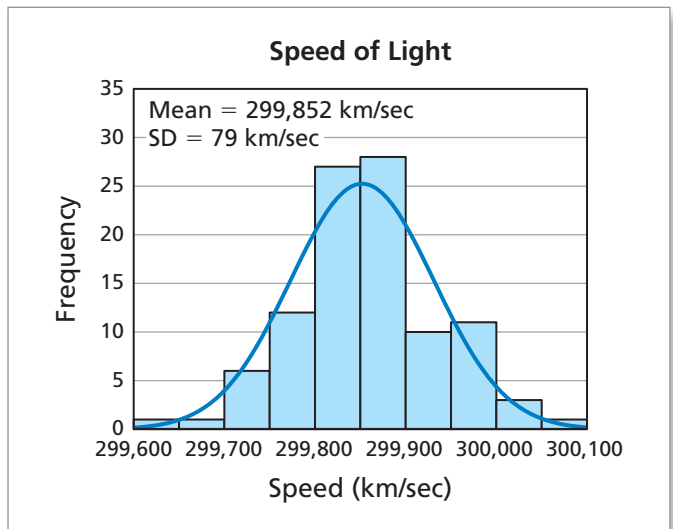
This process is called **statistical inference.**

Key:
■ = chosen

| Population | Take a representative sample. | Collect and analyze data about the sample. | Make inferences about the population. |

---

### EXAMPLE 2    **Estimating by Sampling**

In 1879, Albert Michelson conducted an experiment to measure the speed of light. He conducted the experiment 100 times and obtained the data represented in the histogram. Use a 95% confidence level to estimate the speed of light.

**Speed of Light**

Mean = 299,852 km/sec
SD = 79 km/sec

Frequency vs. Speed (km/sec)

299,600  299,700  299,800  299,900  300,000  300,100

#### SOLUTION

Using a confidence interval calculator and a 95% confidence level, you can infer that the speed of light is

299,852 kilometers per second ± 15 kilometers per second.

In 1926, Michelson improved his estimate to 299,796 ± 4 kilometers per second. The accepted value today is 299,792.458 kilometers per second.

Albert Michelson and Albert Einstein met at Mount Wilson Observatory in 1931, shortly before Michelson's death that same year.

✓ **Checkpoint**                    Help at *Math*.and**Y**⊖**U**.com

A researcher records the temperature of 130 people. The mean of the sample is 98.25°F and the standard deviation is 0.73°F. Does this study allow for the accepted human body temperature of 98.6°F? Explain your reasoning.

## Determining Whether a Sample is Representative

Statistical inference only applies when the sample is randomly chosen from the population. Even then, there may be other factors that make the sample unrepresentative. For a sample to be random, each member of the population must have an equal chance of being chosen.

Key:
■ = chosen

Random Sample          Nonrandom Sample
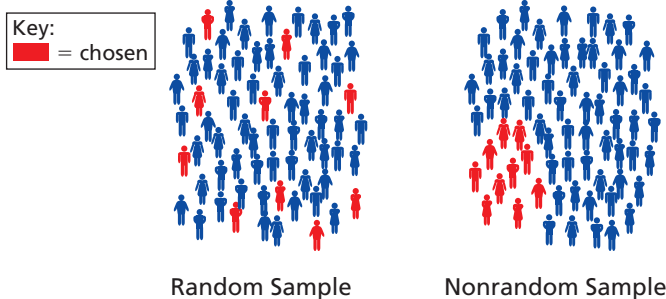
Obtaining a sample that is truly representative is a major problem faced by researchers, polling organizations, and market research departments.

**EXAMPLE 3** **Analyzing How a Poll Can Go Wrong**

Comment on the following article from NBC News.

> "When there's plentiful public polling and stable national trends to base their assumptions upon, their prognostications are usually more or less close to the mark. But politics is unpredictable, and pollsters use different methods to determine who's up and who's down. Every once in a while, they get it wrong.
>
> The famous erroneous 1948 'Dewey defeats Truman' banner headline in the first edition of the *Chicago Tribune* was the result of polls and conventional wisdom that turned out to be dramatically off base. After Los Angeles Mayor Tom Bradley's unanticipated loss in the 1982 California governors' race, many blamed racial bias undetected in public opinion polls. After Barack Obama's dramatic win in the Iowa caucuses in 2008, the media's overwhelming assumption that he would continue his march to victory days later in New Hampshire was due to polling. Those assumptions were shattered when Hillary Clinton wound up winning—prompting weeks of media navel-gazing and questions of 'how did we get it so wrong?'"

On election night in 1948, the press deadline for the *Chicago Tribune* required that the first post-election issue go to press before the states had reported many of their results. Conventional wisdom, supported by polls, predicted that Dewey would win. So, the *Tribune* went to press with the incorrect headline "DEWEY DEFEATS TRUMAN."

### SOLUTION

There are many reasons a poll can go wrong. Here are a few.

- The people being polled do not represent a random selection from the population.

- The questions in the poll can be leading or confusing, as in "Do you plan to vote for Proposition 4, which is against ownership of pit bulls?"

- The people being polled may not know how they will react to a product or how they plan to vote.

✓ **Checkpoint**                    Help at *Math*.and**YOU**.com

What other reasons might cause a poll to go wrong?

When the population consists of people, obtaining a representative sample can be difficult. For various reasons, the sample can be **biased,** which means that it is not representative of the entire population.

EXAMPLE 4    **Analyzing Biased Samples**
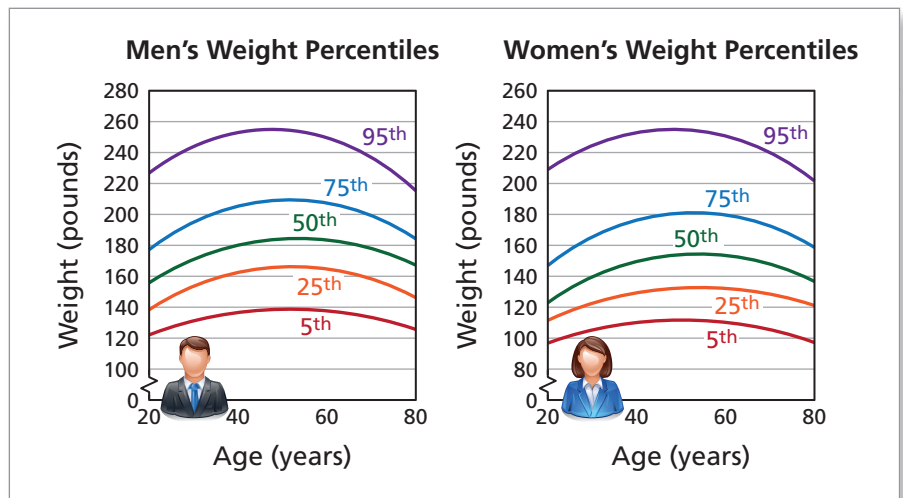
Why might the following samples be biased?

**a.** In 1936, *The Literary Digest* used the results of more than two million responses to a survey to predict that the Republican presidential candidate, Alfred Landon, would defeat the incumbent, Franklin Roosevelt, by a large margin.

**b.** In 2010, the Pew Research Center conducted a pre-election survey of likely voters and found Republican candidates leading by 12 points, 51% to 39%.

**c.** In the 1940s, the Metropolitan Life Insurance Company introduced its standard height-weight tables for men and women. To create the tables, the company used the weights of people who were insured with the company and lived the longest.

## SOLUTION

**a.** *The Literary Digest* primarily collected its sample from people who, during the depression, could afford cars, phones, and magazine subscriptions. This sample overrepresented higher income people, who were more likely to vote for the Republican candidate. Also, Landon supporters may have been more likely to return the survey than Roosevelt supporters, creating what is called a nonresponse bias. By contrast, a poll by George Gallup's organization successfully predicted the result because its sample was more representative of the voting population.

**b.** In a sample that included both landline *and* cell phone interviews, the Republican lead was only 6 points, 48% to 42%. Based on this and other surveys, the Pew Research Center concluded that landline-only surveys tend to be biased toward Republicans.

**c.** The sample may have been biased because it contained only the weights of people who were insured with the company.

Alfred Mossman "Alf" Landon (1887–1987) was the 26th governor of Kansas. He was the Republican nominee in the 1936 presidential election.



**Men's Weight Percentiles** / **Women's Weight Percentiles**

✔ **Checkpoint**    Help at *Math*.and**Y☉U**.com

Give another example of a biased sample.

## Sampling and Market Research

The following are five types of sampling used by market research departments.

1. **Surveys:** The wording of the questions is critical. Surveys can be in person, or via telephone, mail, or Internet.

2. **Focus groups:** Focus groups involve a moderator and questions that lead to a discussion among participants.

3. **Personal interviews:** Personal interviews are similar to focus groups and involve open-ended questions.

4. **Observations:** Observations include purchasing habits and product usage of consumers.

5. **Field trials:** Field trials involve selling a product at selected locations to get an idea of how it will perform in the market.

---

**EXAMPLE 5** **Determining a Sample Size**

You work for the market research department of a cosmetics company. You want to identify characteristics of men who are likely to buy men's personal care products. You send a free sample of two of the products to men and ask them to respond to an online survey. How many samples should you send to obtain reliable results?

### SOLUTION

Suppose you want a confidence level of 90% and a margin of error of 5%. Using the *Sample Size Calculator* at *Math.andYou.com*, you can determine that you need 271 completed surveys.

Suppose you think that 10% of the men who receive the sample will take the time to use it and respond to the online survey. This means that you should send the sample products to about 3000 men. (You still have to decide how to randomly choose 3000 men to obtain a representative sample.)

✓ **Checkpoint**

Help at *Math*.and**YOU**.com

How would you conduct market research on the men's personal care products?

EXAMPLE 6   **Analyzing a Historical Example of Sampling**
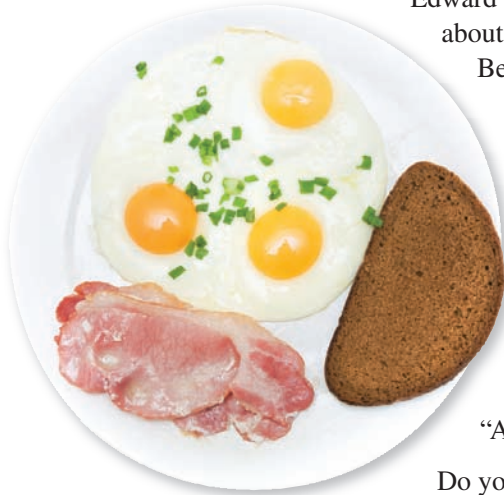
Edward Bernays, a nephew of Sigmund Freud, used some of his uncle's ideas about psychology to influence public opinion. In one well-known campaign, Bernays took an assignment from a large bacon producer, Beechnut Packing Company. Rather than try to take away business from Beechnut's competitors, Bernays decided to try to change America's attitude toward breakfast. At the time, Americans tended to eat small breakfasts, often consisting of juice, toast, and coffee.

In the 1920s, Bernays persuaded a well-known physician in New York to write to his colleagues and ask whether they recommended light breakfasts or hearty breakfasts. The result was "hearty." Newspapers spread the message. As a result of Bernays's campaign, the phrase "bacon and eggs" became synonymous with "American breakfast."

Do you think that his claim that doctors recommend a hearty breakfast is valid?

## SOLUTION

There are two ways to view this question.

- Is Bernays's implied claim that a big breakfast is healthy true?

- Is the claim statistically valid based on his survey of doctors?

Today, most people in the fitness and health fields recommend that people eat breakfast, but there is no consensus on what size it should be.

Is the claim statistically valid based on his survey? The answer is unknown because you do not know the sample size, the form of the questions, or the responses. However, based on other campaigns that Bernays ran, it is clear that he was not as concerned with the truth as he was with the results.

✓ **Checkpoint**                                    Help at *Math*.and**Y⊙U**.com

Which of the following questions might produce more people saying they believe Australians are more sports minded than Americans? Explain.

**a.** Do you agree that Australians are more sports minded than Americans?

　　　　○ Agree　　　　○ Disagree

**b.** Do you think that Australians are more sports minded than Americans?

　　　　○ Yes　　　　○ No

**c.** Do you believe that Australians are more sports minded than Americans, less sports minded, or about the same?

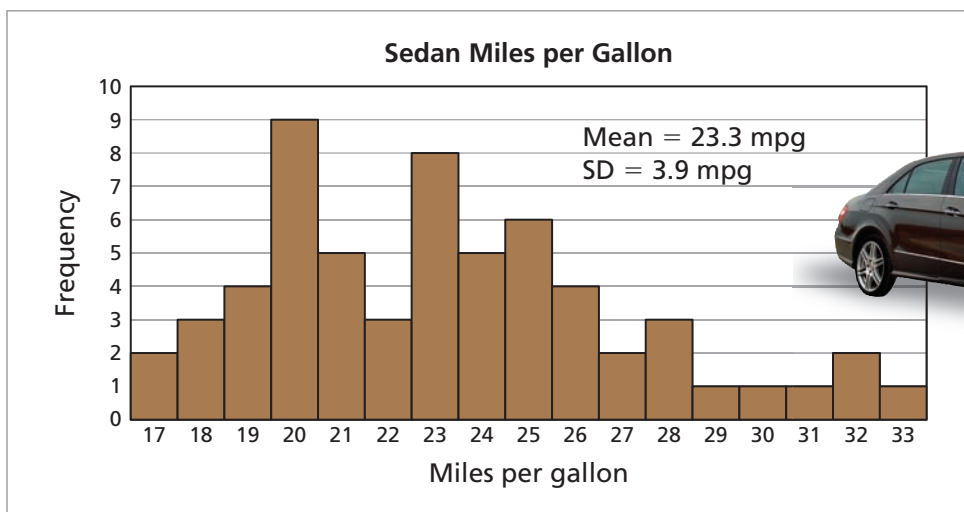　　　　○ More　　　　○ Less　　　　○ About the same

# 9.4 Exercises

**IQ Scores** The histogram shows the distribution of the IQ scores of 50 adults. In Exercises 1 and 2, use the histogram. *(See Examples 1 and 2.)*
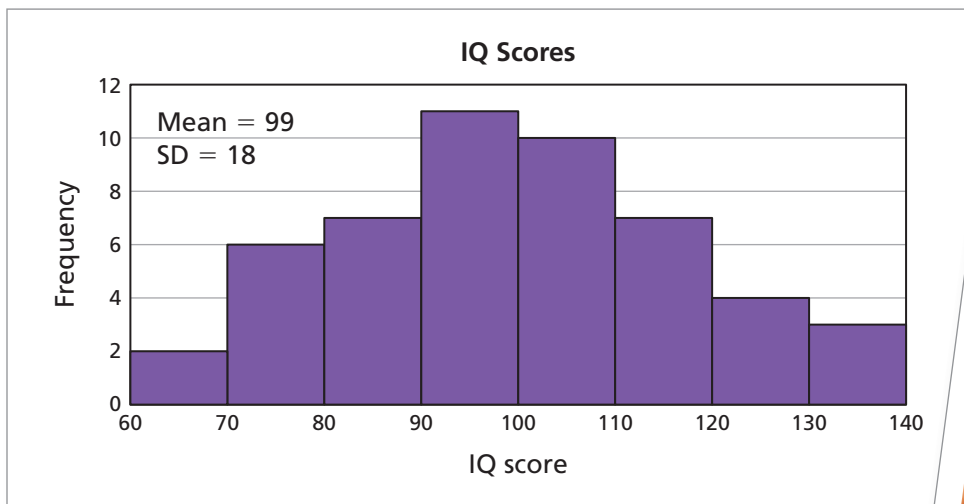
**IQ Scores**

Mean = 99
SD = 18

*(histogram: Frequency vs IQ score; x-axis 60 to 140, y-axis 0 to 12)*

1. Use the *Confidence Interval Calculator* at *Math.andYou.com* and a 95% confidence level to estimate the population mean IQ score.

2. Repeat Exercise 1 using a sample size of 100. Compare the confidence intervals. What happens when the sample size increases? Explain.

**Fuel Efficiency** The graph shows the distribution of the fuel efficiencies of 60 sedans. In Exercises 3 and 4, use the histogram. *(See Examples 1 and 2.)*

**Sedan Miles per Gallon**

Mean = 23.3 mpg
SD = 3.9 mpg

*(histogram: Frequency vs Miles per gallon; x-axis 17 to 33, y-axis 0 to 10)*

3. Use the *Confidence Interval Calculator* at *Math.andYou.com* and a 90% confidence level to estimate the population mean fuel efficiency.

4. Repeat Exercise 3 using a 99% confidence level. Compare the confidence intervals. What happens when the confidence level increases? Explain.

**Biased Samples** In Exercises 5–12, explain why the sample may be biased.
Then explain how to find an unbiased sample. *(See Examples 3 and 4.)*

5. A research company wants to determine how many people floss
   their teeth. The company asks a dentist to conduct a survey by
   randomly selecting patients.

6. You want to estimate the number of students in a high school
   who ride the school bus. You randomly survey 60 seniors.

7. You want to estimate the number of defective items produced
   by a factory. You randomly inspect 200 items from one of the
   machines in the factory.

8. A college wants to determine whether to renovate the gym or
   the science lab. The college asks you to conduct a survey. You
   randomly survey 30 students leaving a science club meeting.

9. A radio station wants to determine how many people in the
   listening area support gun control laws. The station asks
   listeners to call in and answer the survey.

10. A city wants to determine whether the residents of the city favor using tax dollars to
    build a new baseball stadium. The city asks you to conduct a survey. You randomly
    survey people entering a sporting goods store.

11. A mayor wants to determine whether the residents of a city support a bill providing
    insurance for nursing home care. The mayor asks you to conduct a survey. You randomly
    survey residents of five nursing homes in the city.

12. A research company wants to determine how many people in the United States spend at
    least 1 week at the beach each year. The company surveys residents of California, Florida,
    and North Carolina.

**Alternative Energy**  **You work for a research company. You want to determine whether U.S. adults support more funding for alternative energy. In Exercises 13–15, use the *Sample Size Calculator* at *Math.andYou.com*. (See Example 5.)**

13. Choose a 95% confidence level and a 3% margin of error. How many people should you survey to obtain reliable results?

14. Repeat Exercise 13 using a 99% confidence level. Compare the sample sizes. What happens when the confidence level increases? Explain.

15. Repeat Exercise 13 choosing a 5% margin of error. Compare the sample sizes. What happens when the margin of error increases? Explain.

**Biased Questions**  **In Exercises 16–18, determine which question might produce biased results. Explain.  (See Example 6.)**

16.

a. Do you think solar panels should be installed at city hall using taxpayer money?

  ○ Yes          ○ No

b. Do you agree with most of your neighbors that it is a waste of taxpayer money to install solar panels at city hall?

  ○ Agree        ○ Disagree

17.

a. Do you agree that the unfair policy of requiring students to do a time-consuming community service project should be changed?

  ○ Agree        ○ Disagree

b. Do you think the policy of requiring students to do a community service project should be kept?

  ○ Yes          ○ No

18.

a. Do you think people should recycle phone books?

  ○ Yes          ○ No

b. Do you think people should recycle old, out-of-date phone books?

  ○ Yes          ○ No

# ▶ Extending Concepts

**Confidence Intervals**  **In Exercises 19–22, use the confidence interval to find the sample mean and the margin of error.**

**19.** An electronics magazine reports that a 90% confidence interval for the mean price of GPS navigation systems is $178.75 to $211.87.

**20.** A state agency reports that a 95% confidence interval for the mean annual salaries of employees in Colorado is $45,832 to $47,890.

**21.** A hospital reports that a 99% confidence interval for the mean length of stay (in days) of patients is 5.1 to 5.9.

**22.** A company reports that a 95% confidence interval for the mean weight (in ounces) of filled paint cans is 159.97 to 160.03.

**Minimum Sample Size**  **For a 95% confidence level, the minimum sample size $n$ needed to estimate the population mean is**

$$n = \left(\frac{1.96s}{E}\right)^2$$

**where $E$ is the margin of error and $s$ is the population standard deviation. In Exercises 23–26, find the minimum sample size. If necessary, round your answer up to a whole number.**

**23.** You want to estimate the mean weight of newborns within 0.25 pound of the population mean. Assume the population standard deviation is 1.3 pounds.

**24.** You want to estimate the mean number of text messages sent per day by 18- to 24-year-olds within 5 messages of the population mean. Assume the population standard deviation is 30 messages.

**25.** You want to estimate the mean number of hours of television watched per person per day within 0.1 hour of the population mean. Assume the population standard deviation is 1.5 hours.

**26.** You want to estimate the mean number of minutes waiting at a department of motor vehicles office within 0.5 minute of the population mean. Assume the population standard deviation is 7 minutes.

# 9.3–9.4   Quiz

**Ambulance Response Times**  The graph shows the distribution of the ambulance response times for 250 emergency calls in a city. In Exercises 1–6, use the histogram.
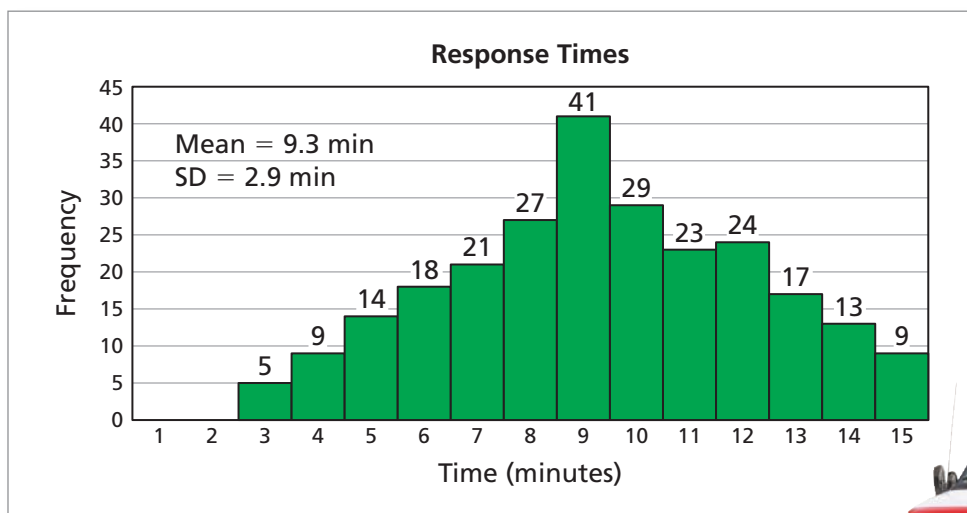
**Response Times**

Mean = 9.3 min
SD = 2.9 min



1.  Does this data set have a normal distribution? Explain your reasoning.

2.  What percent of the response times lie within 1 standard deviation of the mean?

3.  What percent of the response times lie within 2 standard deviations of the mean?

4.  Compare the percents in Exercises 2 and 3 with the percents given by the normal distribution.

5.  Use the *Confidence Interval Calculator* at *Math.andYou.com* and a 95% confidence level to estimate the population mean response time.

6.  Use the *Confidence Interval Calculator* at *Math.andYou.com* and a 99% confidence level to estimate the population mean response time.

7.  **Fire Station**  A city wants to know whether residents will favor a tax increase for the renovation of a fire station. You randomly survey 100 people in the neighborhood around the fire station. Explain why the sample may be biased. Then explain how to find an unbiased sample.

8.  **Hospital**  You work for a research company. You want to estimate the percent of U.S. adults who have contributed to a hospital fundraiser in the past 12 months. Use the *Sample Size Calculator* at *Math.andYou.com*. Choose a 99% confidence level and a 5% margin of error. How many people should you survey to obtain reliable results?

# Chapter 9    Summary

| **Section Objectives** | **How does it apply to you?** |
|---|---|

### Section 1

| Use stacked area graphs to represent the changing parts of a whole. | You can see how related variables graphically change over time. *(See Examples 1 and 2.)* |
| Use a radar graph and an area graph to represent data. | There are many creative and unique ways to help people understand and use information. *(See Examples 3 and 4.)* |
| Graphically represent data sets that have several variables. | Not all information is represented with simple types of information design, such as bar graphs, circle graphs, and scatter plots. *(See Examples 5 and 6.)* |

### Section 2

| Use mean, median, and mode to describe the average value of a data set. | Measures of central tendency are used to compare and communicate data in real life. *(See Examples 1 and 2.)* |
| Read and understand box-and-whisker plots and histograms. | You can use box-and-whisker plots and histograms to analyze the variability and distribution of data sets. *(See Examples 3 and 4.)* |
| Understand the effect of outliers on averages. | Outliers can significantly change measures of central tendency. This could lead to using bad information for decision making. *(See Examples 5 and 6.)* |

### Section 3

| Use standard deviation to describe the dispersion of a data set. | Standard deviation allows you to determine whether data values are clustered around the mean or spread out over a large range. *(See Examples 1 and 2.)* |
| Use standard deviation to describe a data set that is normally distributed. | Many naturally occurring data sets have a normal distribution. You can use standard deviation to determine whether a data set is normal. *(See Examples 3 and 4.)* |
| Compare different types of distributions. | Not all data distributions are normal. There are other types of distributions. *(See Examples 5 and 6.)* |

### Section 4

| Use a randomly chosen sample to describe a population. | If a sample is randomly taken from a population, then you can use the sample mean to estimate the population mean. *(See Examples 1 and 2.)* |
| Determine whether a sample is representative of a population. | If your sample is too small, not random, or biased, the statistics from the sample will not represent the population. *(See Examples 3 and 4.)* |
| Determine a sample size to obtain valid inferences. | The minimum sample size depends on the confidence level and the margin of error chosen. *(See Examples 5 and 6.)* |

# Chapter 9   Review Exercises

## Section 9.1

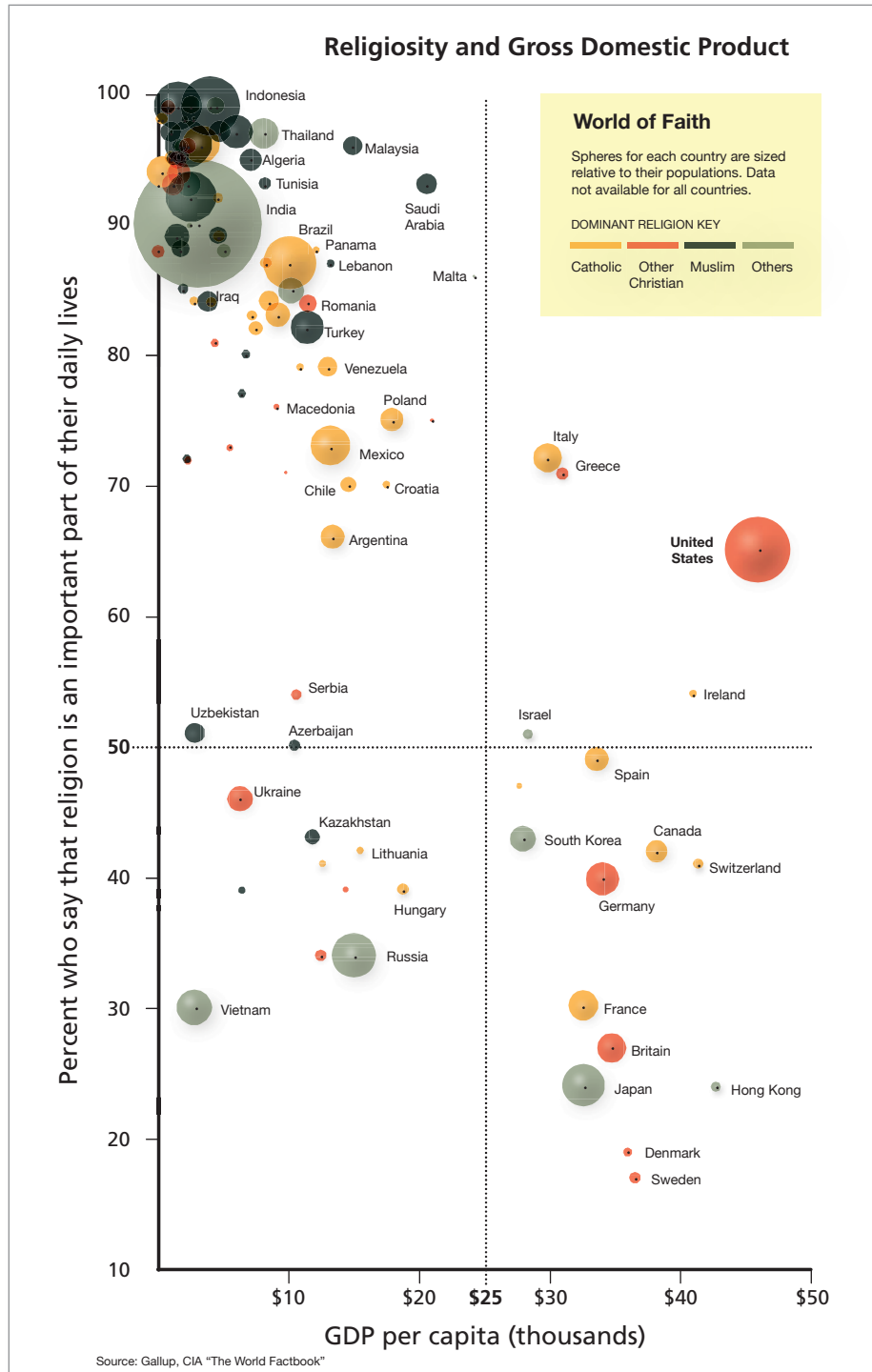**Religiosity**  In Exercises 1–8, use the information below.

The graph shows the data from a Gallup survey relating religiosity to gross domestic product (GDP) per capita. GDP is the market value of the goods and services produced by a country over a certain time period, usually by year. It is often used as a comparative statistic to gauge a country's standard of living.

**1.** Describe the variables displayed in the graph.

**2.** Describe the information presented in the graph.

**3.** Which country has the highest GDP per capita?

**4.** What is the dominant religion in France?

**5.** Does the graph support the comment below? Explain your reasoning.

COMMENT:

> Religiosity is highly correlated to poverty. Richer countries, in general, are less religious.

**6.** Are there more religious people in Argentina or Russia? Explain your reasoning.

**7.** Compare the religiosity of the United States with that of other countries with high GDPs per capita.

**8.** Could you use a stacked area graph to show the information presented in this graph? Explain your reasoning.

### Religiosity and Gross Domestic Product

**World of Faith**

Spheres for each country are sized relative to their populations. Data not available for all countries.

DOMINANT RELIGION KEY

Catholic     Other     Muslim     Others
               Christian

*y-axis:* Percent who say that religion is an important part of their daily lives

*x-axis:* GDP per capita (thousands)

Source: Gallup, CIA "The World Factbook"

## Section 9.2

DATA    **Unemployment Rate**  In Exercises 9–12, use a spreadsheet and the line graph shown.

**Historical Unemployment Rates (1991–2010)**
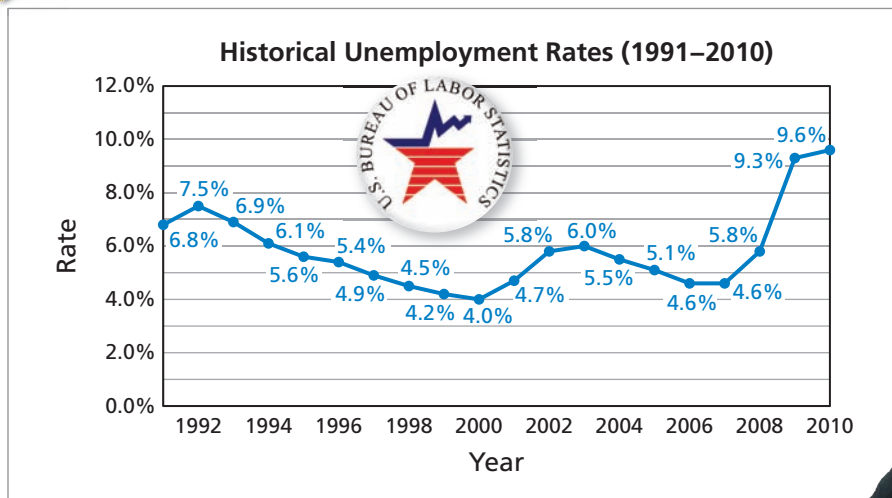


9. What are the mean, median, and mode of the data presented in the line graph?

10. Which measure of central tendency best describes the average of the data? Explain your reasoning.

11. Which two data values are outliers?

12. Do the outliers have a greater effect on the mean or the median? Explain your reasoning.

13. **Box-and-Whisker Plot**  Use the box-and-whisker plot to analyze the unemployment rates.

**Historical Unemployment Rates**
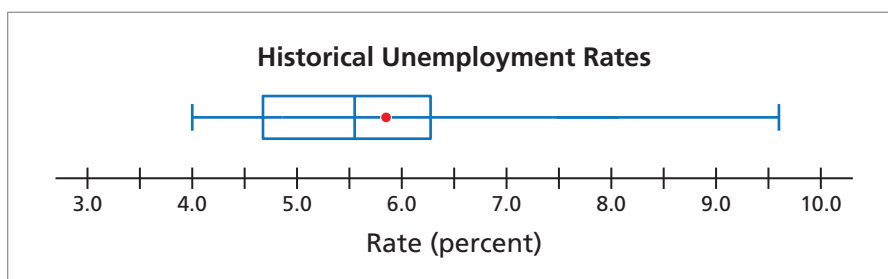


14. **Histogram**  Explain whether the following statement is supported by the histogram: "The unemployment rate during most years is between 4.0% and 7.0%."

**Historical Unemployment Rates**
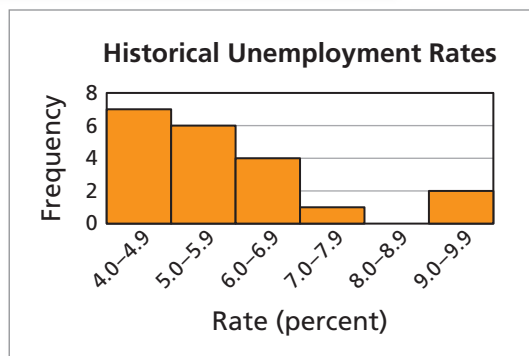
## Section 9.3

**Hotel Rooms** The histogram shows the distribution of the numbers of hotel rooms occupied for 150 days in a year. In Exercises 15–18, use the histogram.
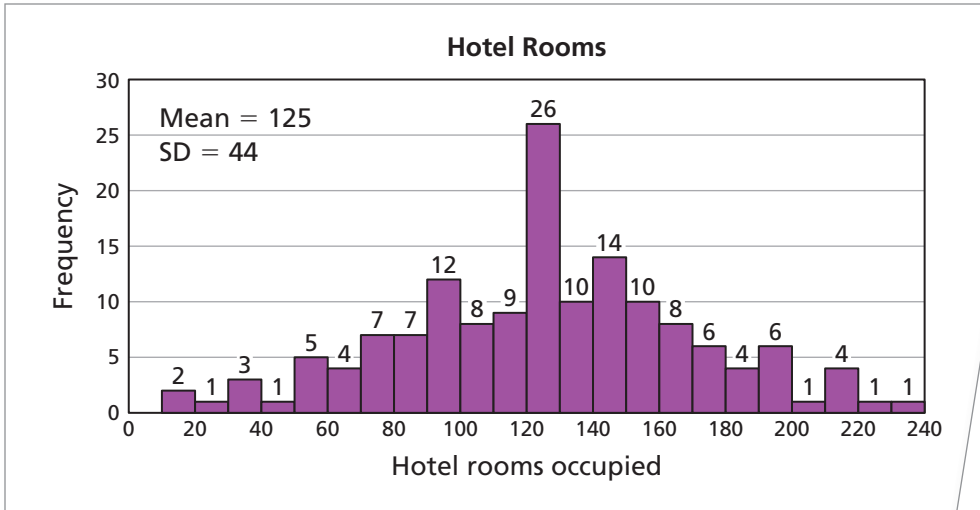
**Hotel Rooms**

Mean = 125
SD = 44



**15.** Does this data set have a normal distribution? Explain your reasoning.

**16.** What percent of the days lie within 1 standard deviation of the mean?

**17.** What percent of the days lie within 2 standard deviations of the mean?

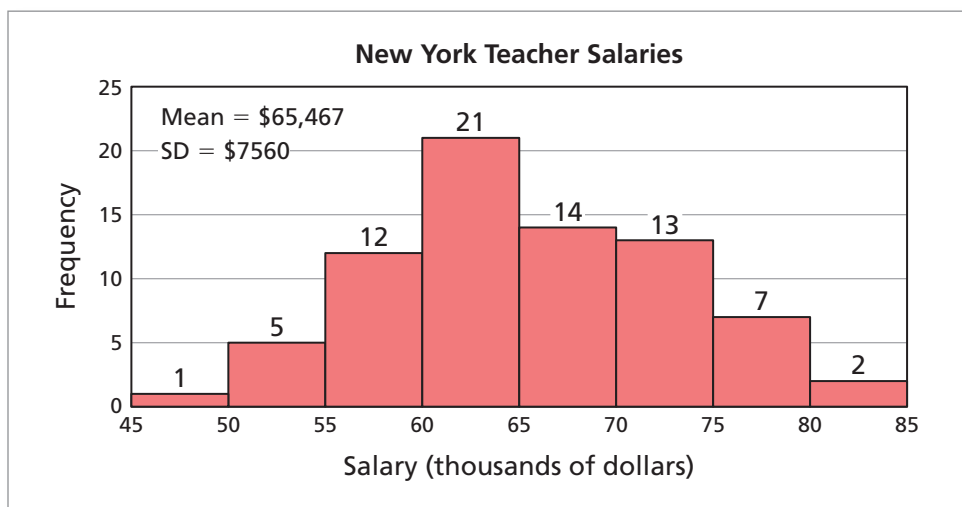**18.** Compare the percents in Exercises 16 and 17 with the percents given by the normal distribution.

**Television** The histograms show the distributions of samples of hours of television watched per day by men and women. In Exercises 19 and 20, use the histograms.

**Men**

Mean = 3.1 hr
SD = 0.8 hr

**Women**

Mean = 2.5 hr
SD = 0.6 hr
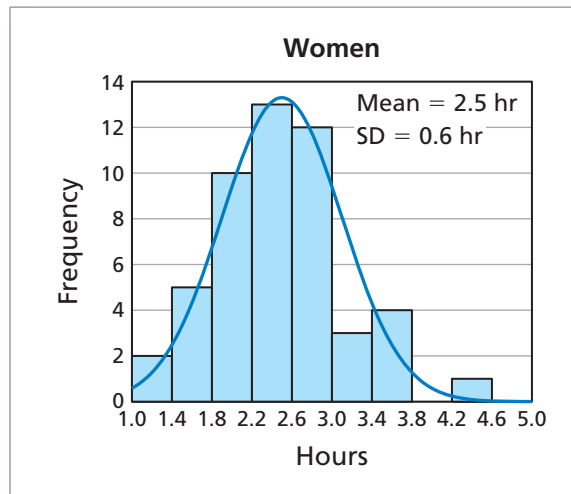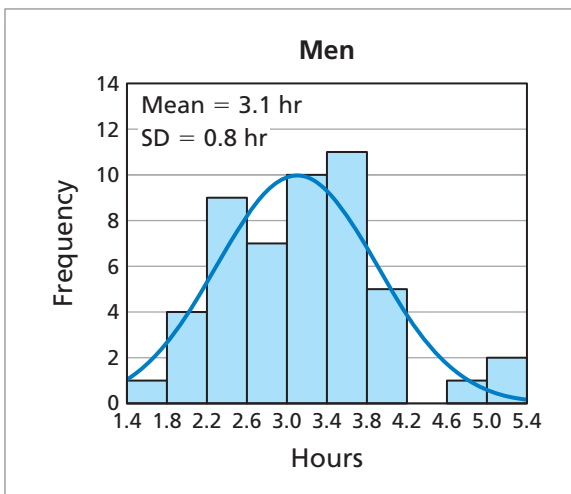
**19.** What is the significance of the smaller standard deviation for women?

**20.** Estimate the percents of men and women that watch at least 4 hours of television per day.

## Section 9.4

**Teacher Salaries**  The histogram shows the distribution of the salaries of
75 New York public school teachers. In Exercises 21 and 22, use the histogram.

**New York Teacher Salaries**

Mean = $65,467
SD = $7560

Frequency values by salary bin (thousands of dollars):
- 45–50: 1
- 50–55: 5
- 55–60: 12
- 60–65: 21
- 65–70: 14
- 70–75: 13
- 75–80: 7
- 80–85: 2

Salary (thousands of dollars)

**21.** Use the *Confidence Interval Calculator* at *Math.andYou.com* and a
95% confidence level to estimate the population mean salary.

**22.** Repeat Exercise 21 for a 99% confidence level.

**Biased Samples**  In Exercises 23 and 24, explain why the sample may
be biased. Then explain how to find an unbiased sample.

**23.** A politically conservative magazine wants to estimate the percent
of U.S. adults who approve of the job the president is doing. The
magazine sends the survey to readers on its mailing list.

**24.** A pollster wants to determine the percent of registered voters who
voted for the Republican candidate in an election. The pollster
conducts an exit poll in a primarily Republican neighborhood.

**25. Political Campaign**  You work for a political campaign. You want to
estimate the percent of registered voters who plan to vote for your
candidate. Use the *Sample Size Calculator* at *Math.andYou.com*.
Choose a 90% confidence level and a 2% margin of error. How many
people should you survey to obtain reliable results?

**26. Biased Questions**  Determine which question could produce biased results. Explain.

  **a.** Do you agree with most of the students that the senator should resign?

  ○ Agree      ○ Disagree

  **b.** Do you think the senator should resign?

  ○ Yes      ○ No

# 10 The Mathematics of Fitness & Sports

**Calories Burned During Physical Activities**



The graph in Example 6 on page 489 compares the number of calories burned per minute for eight activities. Why does cross-country skiing consume so much energy?

# 10.1 Health & Fitness

▶ Compare a person's weight, height, and body fat percentage.
▶ Interpret and use a person's heart rate and metabolism.
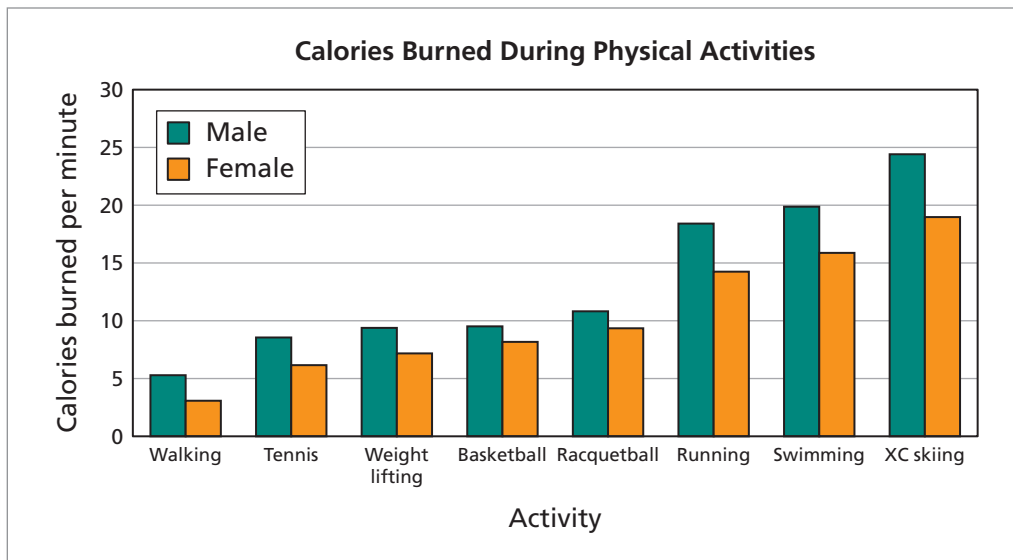▶ Determine factors for cardiovascular health.

## Comparing Weight, Height, and Fat Percentage

There are many different opinions about the "ideal" weight and body fat percentage for a person of a given height. Some of the opinions are expressed by tables, and some are expressed by formulas.

**Study Tip**

The Metropolitan Life Insurance Company *Height-Weight Tables* are available at *Math.andYou.com*. When using the tables, you are wearing shoes with 1-inch heels.

### EXAMPLE 1 Analyzing the MetLife Height-Weight Tables

In a criticism of the MetLife tables, Steven B. Halls, MD, said, "For very tall men and women, the MetLife tables suggest impossibly low weights." As part of the criticism, Halls presented the following graphs. What do Halls's graphs show?



**Weight for Height of Men**

**Weight for Height of Women**
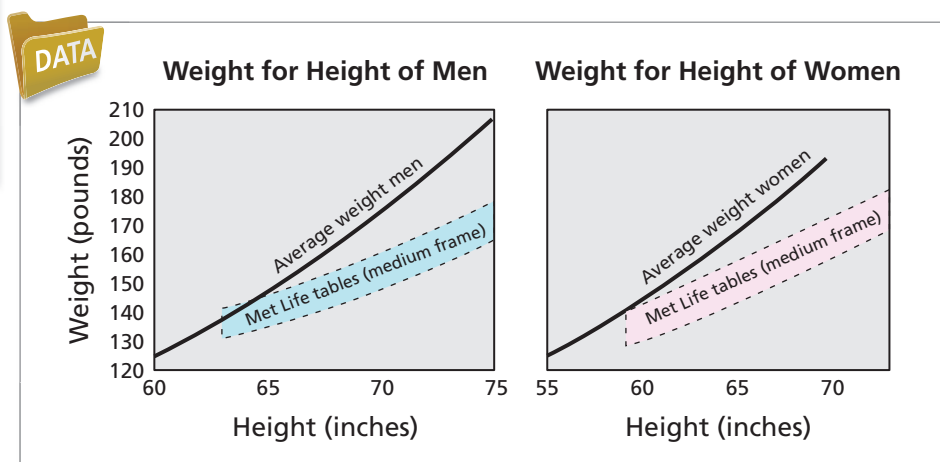
### SOLUTION

The graphs are roughly linear. The rate of increase of each graph is about 5 pounds per inch of height, as indicated by the following formulas.

Medium-framed men:     Weight = 125 pounds + 5 pounds per inch over 5 feet
Medium-framed women:  Weight = 145 pounds + 5 pounds per inch over 5 feet

Compared to average weights, the MetLife tables recommend unrealistically low weights, especially for tall people.

✓ **Checkpoint**                          Help at *Math.andYOU.com*

Here is another formula by Dr. G.J. Hamwi that first appeared in 1964. Graphically compare this formula with the average weights of Americans.

Medium-framed men:     Weight = 106 pounds + 6 pounds per inch over 5 feet
Medium-framed women:  Weight = 100 pounds + 5 pounds per inch over 5 feet

*Ideal Body Weight* calculators are available at *Math.andYou.com*.

Weight: 145 lb
Wrist: 6.5 in.
Waist: 28 in.
Hips: 36 in.
Forearm: 10 in.

The formulas on this page only give estimates for body fat percentage. One way to find an accurate measurement is to find your weight in and out of water.



**EXAMPLE 2**   **Finding a Body Fat Percentage**

The human body is composed of many things including muscles, bones, internal organs, and fat. You can find your body fat percentage by dividing your body fat weight by your total body weight. You can use the following formulas to approximate your body fat percentage. (Use pounds and inches.)

**Body Fat Formula for Women**

**Factor 1:**  (total body weight × 0.732) + 8.987
**Factor 2:**  wrist circumference (at fullest point)/3.140
**Factor 3:**  waist circumference (at naval) × 0.157
**Factor 4:**  hip circumference (at fullest point) × 0.249
**Factor 5:**  forearm circumference (at fullest point) × 0.434

Lean body mass = factor 1 + factor 2 − factor 3 − factor 4 + factor 5
Body fat weight = total body weight − lean body mass
Body fat percentage = (body fat weight × 100)/total body weight

**Body Fat Formula for Men**

**Factor 1:**  (total body weight × 1.082) + 94.42
**Factor 2:**  waist circumference × 4.15

Lean body mass = factor 1 − factor 2
Body fat weight = total body weight − lean body mass
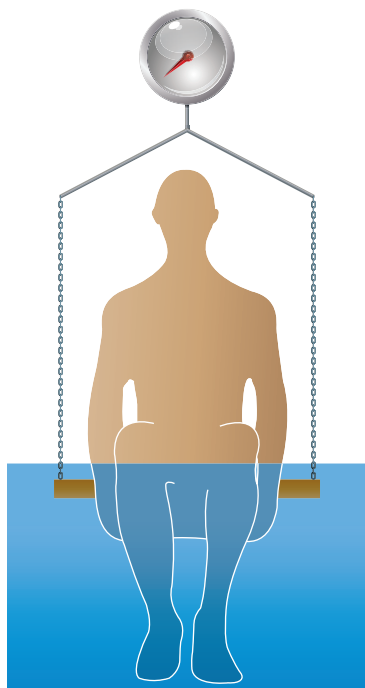Body fat percentage = (body fat weight × 100)/total body weight

Find the body fat percentage for the woman shown.

**SOLUTION**

Using a spreadsheet, you can see that the woman's body fat percentage is about 25.4%.

| | A | B | C |
|---|---|---|---|
| 1 | | | **Factor** |
| 2 | Weight | 145 | 115.13 |
| 3 | Wrist | 6.5 | 2.07 |
| 4 | Waist | 28 | 4.40 |
| 5 | Hips | 36 | 8.96 |
| 6 | Forearm | 10 | 4.34 |
| 7 | | | |
| 8 | Lean body mass | | 108.18 |
| 9 | Body fat weight | | 36.82 |
| 10 | Body fat percentage | | 25.4% |

✓ **Checkpoint**     Help at *Math*.andY☺U.com

A study was conducted to determine the relationship between body fat percentage (as a percent) and body mass index (BMI), taking age and gender (males = 1, females = 0) into account. The study consisted of 521 males and 708 females, with wide ranges in BMI and age. The researchers determined that the formula for adults is

BMI = (0.83 × body fat %) − (0.19 × age) + (9 × gender) + 4.5.

Find the BMI for the 24-year-old woman in Example 2 and for a 26-year-old man who weighs 210 pounds and has a 36-inch waist.

**Heart Rate Levels for
a 20-Year-Old**

| | |
|---|---|
| 200 | |
| 180 | MAXIMUM |
| 160 | |
| 140 | HIGH TARGET |
| | LOW TARGET |
| 120 | |
| | ABOVE NORMAL |
| 100 | |
| 80 | AT REST |
| 60 | |
| | BELOW NORMAL |
| 40 | |
| 20 | DANGEROUSLY LOW |
| 0 | |
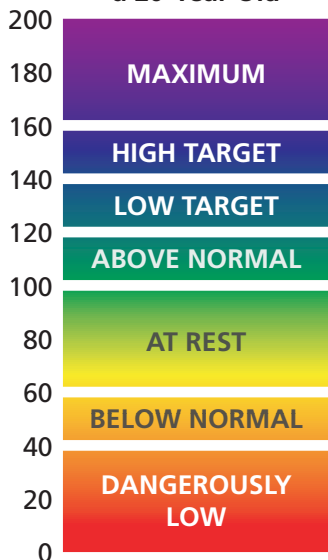
## Heart Rate and Metabolism

Your **heart rate** is the number of times your heart beats in 1 minute. Your heart rate is lower when you are at rest and increases when you exercise because your body needs more oxygen-rich blood when you exercise. Here is the normal heart rate for a person at rest.

Ages 1–10: 70–120 beats per minute

Ages 11+: 60–100 beats per minute

### EXAMPLE 3  Finding a Target Heart Rate

An estimate for your maximum heart rate is

Maximum heart rate (MHR) = 220 − (your age).

Many fitness specialists recommend staying within 60–80% of your maximum heart rate during exercise. This range is called your target heart rate zone. Create a table showing the target heart rate zones for different ages.

### SOLUTION

DATA

| | A Age | B Maximum Heart Rate | C Low Target Rate | D High Target Rate |
|---|---|---|---|---|
| 1 | Age | Maximum Heart Rate | Low Target Rate | High Target Rate |
| 2 | 20 | 200 | 120 | 160 |
| 3 | 30 | 190 | 114 | 152 |
| 4 | 40 | 180 | 108 | 144 |
| 5 | 50 | 170 | 102 | 136 |
| 6 | 60 | 160 | 96 | 128 |
| 7 | 70 | 150 | 90 | 120 |
| 8 | 80 | 140 | 84 | 112 |

For instance, when a 30-year-old exercises, his or her heart rate should be between 114 and 152 beats per minute.

✓ **Checkpoint**

Help at *Math*.and**YOU**.com

*The 24/5 Complete Personal Training Manual* suggests that there is a "fat burning zone." This zone is 60–65% of your maximum heart rate, as indicated in the following table.

Fat burning zone

| Estimates for a 130-pound woman during exercise | 60–65% MHR | 80–85% MHR |
|---|---|---|
| Total calories expended per minute | 4.86 | 6.86 |
| Fat calories expended per minute | 2.43 | 2.70 |

With more fat calories being burned at the higher rate, why is the 60–65% zone called the "fat burning zone"?

**Incoming    Outgoing**



| Activity |
| Intake | BMR |
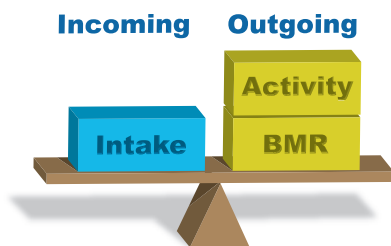
If your incoming calories are equal to your outgoing calories, you should maintain your weight. To lose weight, decrease your incoming calories and/or increase your outgoing calories. To gain weight, do the opposite.

Your *basal metabolic rate* (BMR) is the number of calories you expend per day while in a state of rest. BMR decreases with age and increases with the gain of lean body mass.

To determine your BMR accurately, you need a fairly sophisticated test. However, there are several formulas for approximating your BMR. Here is one that was developed by Mark Mifflin and Sachiko St. Jeor in 1990.

### Formula for Basal Metabolic Rate (BMR) in Calories Per Day

**Factor 1:**  $4.545 \times$ weight (lb)
**Factor 2:**  $15.875 \times$ height (in.)
**Factor 3:**  $5 \times$ age (yr)
**Factor 4:**  5 for males and $-161$ for females

$$\text{BMR} = \text{Factor 1} + \text{Factor 2} - \text{Factor 3} + \text{Factor 4}$$

**EXAMPLE 4**    **Using a Basal Metabolic Rate**

To determine your daily calorie needs, you must multiply your basal metabolic rate by a number determined by your activity level.

Weight: 198 lb
Height: 72 in.
Age: 32 yr



| Outgoing Calories | |
|---|---|
| **Activity level** | **BMR multiplier** |
| Sedentary | 1.200 |
| Lightly active | 1.375 |
| Moderately active | 1.550 |
| Very active | 1.725 |
| Extra active | 1.900 |

| Incoming Calories | |
|---|---|
| **Food** | **Calories per gram** |
| Carbohydrate | 4 |
| Protein | 4 |
| Fat | 9 |
| Alcohol | 7 |

Assume the man shown is very active and his daily calorie intake is about 3500 calories. Would you expect him to be losing weight or gaining weight? Explain your reasoning.

### SOLUTION

DATA

| | A | B | C |
|---|---|---|---|
| 1 | | | **Factor** |
| 2 | Weight | 198 | 900 |
| 3 | Height | 72 | 1143 |
| 4 | Age | 32 | 160 |
| 5 | Gender | M | 5 |
| 6 | | | |
| 7 | BMR | | 1888 |
| 8 | Activity level | VA | 1.725 |
| 9 | Outgoing calories | | 3257 |

| Incoming calories: | 3500 |
|---|---|
| Outgoing calories: | $-3257$ |
| **Balance:** | 243 |

The man should be gaining weight.

✓ **Checkpoint**

Find your daily calorie balance.

Help at *Math*.and**Y❂U**.com

## Cardiovascular Health

On page 454, you saw that the 60–65% target of your maximum heart rate is called the "fat burning zone." The 80–85% target is called the "cardio zone." Both of these target zones are classified as *aerobic exercise* because they involve increased oxygen to the lungs. *Anaerobic exercise* (without oxygen) includes activities like weight lifting, which are designed to increase muscle mass.

Cardio zone

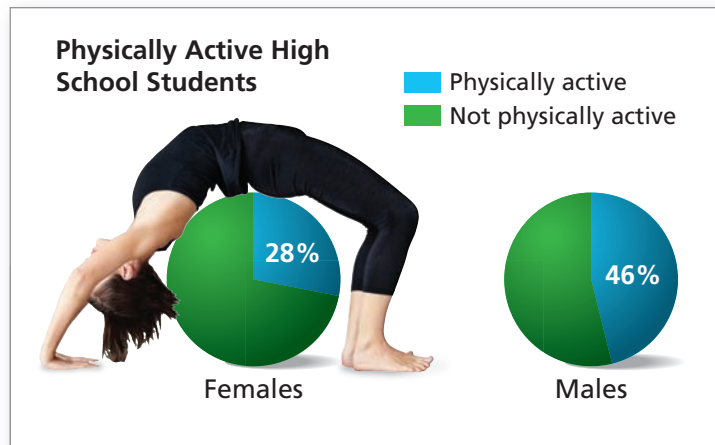|  | 60–65% MHR | 80–85% MHR |
|---|---|---|
| Total calories expended per minute | 4.86 | 6.86 |
| Fat calories expended per minute | 2.43 | 2.70 |

### EXAMPLE 5    Graphically Representing Health Claims

Graphically represent one of the statements.

- Cardiovascular disease accounts for about 34% of all deaths in America.
- Lack of physical activity is a risk factor for cardiovascular disease. About 49% of Americans 18 years old and older are not physically active.
- About 46% of male high school students and 28% of female high school students are physically active.
- Even low-to-moderate intensity activities, such as walking, when done for as little as 30 minutes a day, bring benefits.

### SOLUTION

Here is one possibility.



**Physically Active High School Students**

■ Physically active
■ Not physically active

28%  Females

46%  Males

Your cardiovascular system consists of your heart and your blood vessels. Cardiovascular fitness refers to the ability of your heart, lungs, and other organs to transport and use oxygen. The maximum volume of oxygen your body can transport and use is called your *aerobic capacity*. When you exercise regularly, you can increase your cardiovascular fitness as your heart becomes more efficient at pumping blood and oxygen to the body, and as your body becomes more efficient at using that oxygen.

✓ **Checkpoint**                    Help at *Math*.and**Y☺U**.com

Choose a different claim and represent it graphically.

Cholesterol is a fat-like substance that occurs naturally in all areas of the human body. Your body needs some cholesterol to help it work properly. About 75% of the cholesterol in your body is produced by your liver. The rest comes from foods like meats, eggs, and dairy products. The biggest influence on your blood cholesterol level is fats in your diet, not the amount of cholesterol in the food you eat.

| | HDL Cholesterol | | LDL Cholesterol | Total Cholesterol | Triglycerides | |
|---|---|---|---|---|---|---|
| Desirable ▶ | 60+ | | 160+ | 240+ | 200+ | ◀ High |
| Borderline low ▶ | 40–59 | | 130–159 | 200–239 | 150–199 | ◀ Borderline high |
| Low ▶ | 0–39 | | 0–129 | 0–199 | 0–149 | ◀ Desirable |

Milligrams per deciliter                    Milligrams per deciliter

**EXAMPLE 6**    **Describing Insulin and Glucose Interaction**

Use the graph to describe the interaction between glucose and insulin in a typical daily diet.
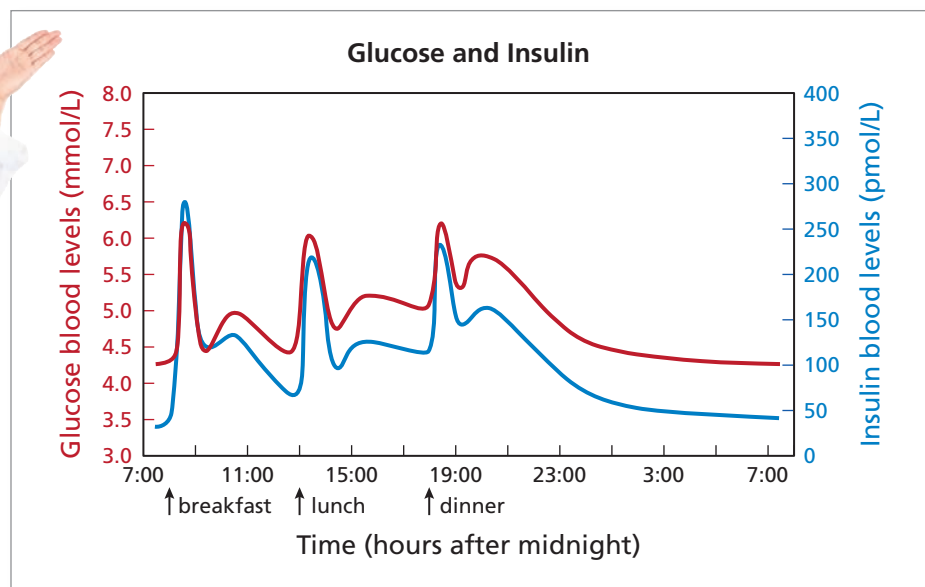


**SOLUTION**

Food that you eat is broken down into glucose, the simple sugar that is the main source of energy for your body's cells. But, your cells cannot use glucose without insulin, a hormone secreted by the pancreas. Insulin helps the cells take in glucose and convert it to energy.

Throughout the day, each time you eat, your pancreas receives a signal to produce insulin. The graph shows that the amount of insulin a healthy person produces is proportional to the amount of glucose in the blood.

✔ **Checkpoint**                    Help at *Math*.and**Y**🌐**U**.com

Use the Internet to describe the relationship between cholesterol, glucose, and cardiovascular health.

Testing your cholesterol and glucose levels is a standard part of a physical examination by your doctor. Both of these tests require you to fast for several hours before the test.

# 10.1 Exercises

**The Robinson Formula** In 1983, Dr. J.D. Robinson published a formula for ideal weight. In Exercises 1–6, use the formula. *(See Example 1.)*
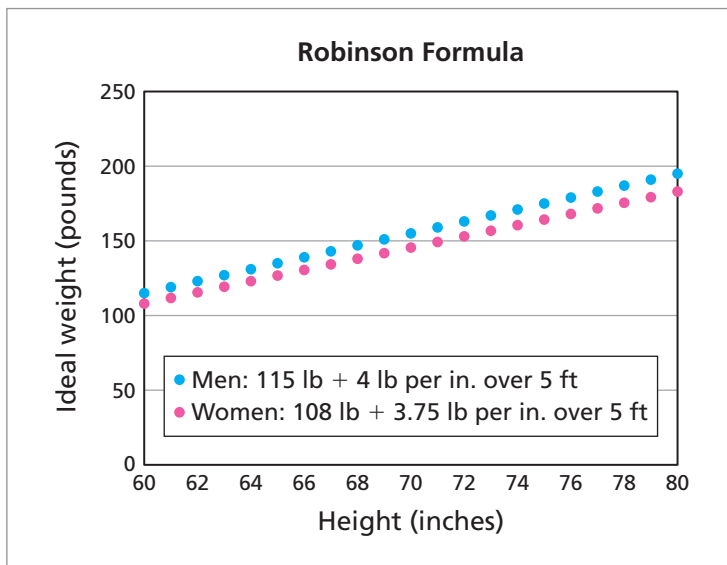
1. Find the ideal weight of a woman who is 5 feet 2 inches tall.

2. Find the ideal weight of a man who is 5 feet 3 inches tall.

3. Find the ideal weight of a man who is 6 feet tall.

4. Find the ideal weight of a woman who is 6 feet 4 inches tall.

**Robinson Formula**

Ideal weight (pounds) vs Height (inches)

● Men: 115 lb + 4 lb per in. over 5 ft
● Women: 108 lb + 3.75 lb per in. over 5 ft

5. How do you think you should calculate the ideal weight of a person who is shorter than 5 feet tall? Explain your reasoning. Then find the ideal weight of a woman who is 4 feet 8 inches tall.

6. Compare the ideal weights for men and women given by the Robinson formula.

DATA **Body Fat Percentage** In Exercises 7 and 8, find (a) the body fat percentage and (b) the body mass index for the person shown. *(See Example 2.)*
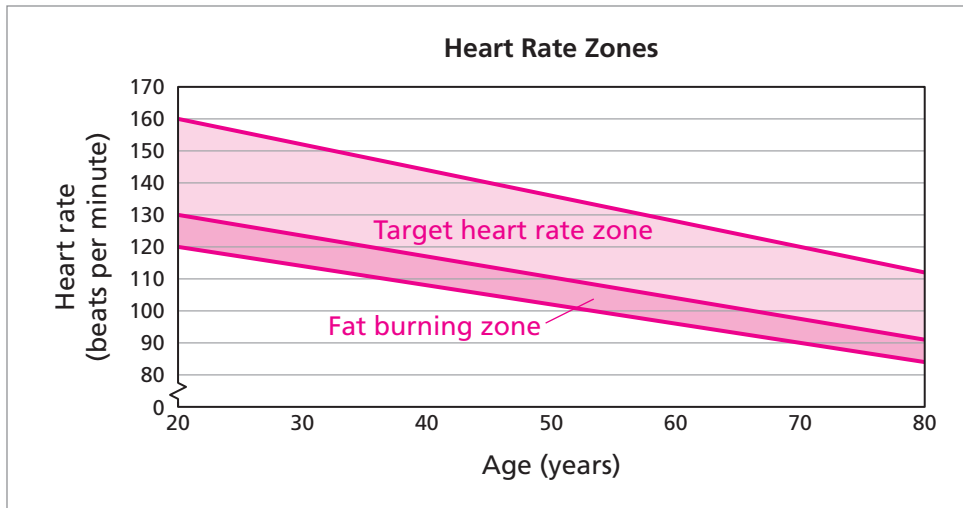
7.

8.

Weight: 160 lb
Waist: 34 in.
Age: 21 yr

Weight: 125 lb
Wrist: 4.5 in.
Waist: 24 in.
Hips: 30 in.
Forearm: 7 in.
Age: 18 yr

**Heart Rate Zones**  In Exercises 9–12, use the graph.  *(See Example 3.)*

### Heart Rate Zones



9. Describe any trends in the graph.

10. Is a 25-year-old with a heart rate of 150 beats per minute in the target heart rate zone?

11. Is a 42-year-old with a heart rate of 130 beats per minute in the fat burning zone?

12. Is a 40-year-old with a heart rate of 110 beats per minute in the fat burning zone?

**DATA** **Daily Calorie Balance**  In Exercises 13 and 14, determine whether you would expect the person to be losing weight or gaining weight. Explain your reasoning.  *(See Example 4.)*

13.
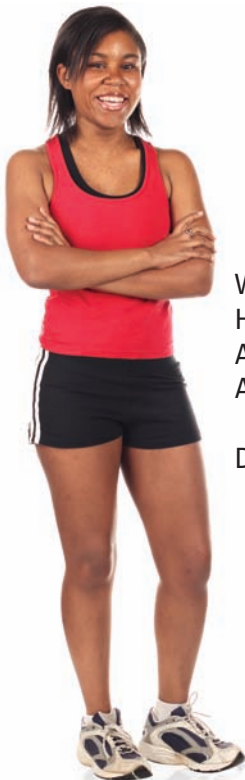


Weight: 121 lb
Height: 64 in.
Age: 24 yr
Activity level:
    moderately active
Daily calorie intake:
    1500 calories

14.



Weight: 154 lb
Height: 68 in.
Age: 26 yr
Activity level:
    lightly active
Daily calorie intake:
    3000 calories

**Graphical Representation**  **In Exercises 15–18, graphically represent the statement.**  *(See Example 5.)*

15. About 15% of adults 20 years old and older have total cholesterol greater than or equal to 240 milligrams per deciliter.

16. About 13% of males 20 years old and older and 16% of females 20 years old and older have total cholesterol greater than or equal to 240 milligrams per deciliter.

17. The average total cholesterol of adults 20 years old and older is 198 milligrams per deciliter.

18. The average total cholesterol of males 20 years old and older is 195 milligrams per deciliter, and the average total cholesterol of females 20 years old and older is 200 milligrams per deciliter.

**Cholesterol and Coronary Heart Disease**  **The graph shows the results from a study. In Exercises 19–22, use the graph.**  *(See Example 6.)*



**Coronary Heart Disease Mortality for 356,222 Men Aged 35 to 57 Years Over a Six-Year Period**
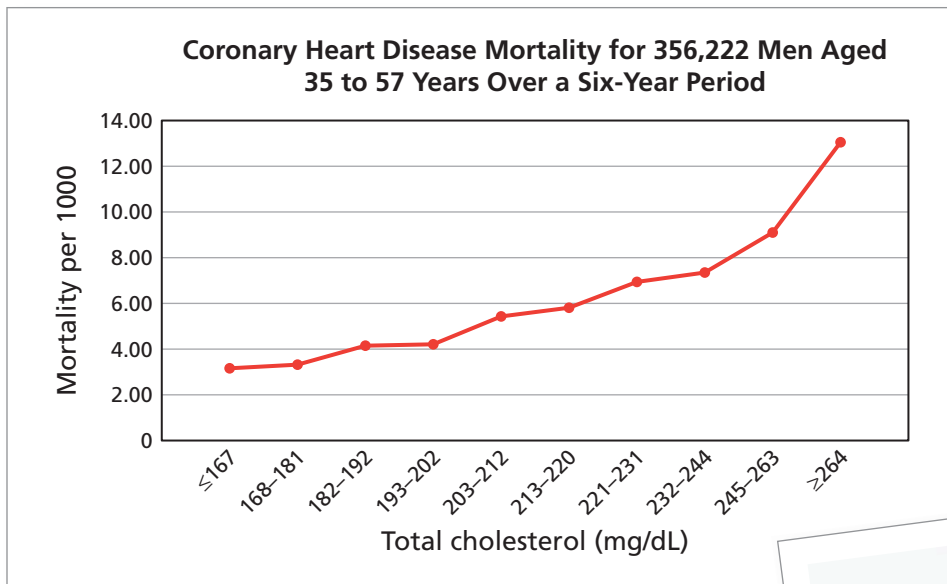
19. Describe any trends in the graph.

20. Compare the information in the graph to the information about total cholesterol in the chart on page 457.

21. What is the 6-year mortality rate for men with total cholesterols between 182 and 192 milligrams per deciliter?

22. How many times greater is the risk of dying from coronary heart disease for a man with total cholesterol greater than or equal to 264 milligrams per deciliter than for a man with total cholesterol less than or equal to 167 milligrams per deciliter?

# ▶ Extending Concepts

**DATA**  **The Katch-McArdle Formula**  The Katch-McArdle formula for basal metabolic rate applies to both men and women. In Exercises 23–30, use the formula.

$$\frac{\text{Basal metabolic rate}}{\text{(calories per day)}} = 370 + 9.8 \times \text{lean body mass (lb)}$$

**23.** Make a graph for the Katch-McArdle formula. Describe any trends in the graph.

**24.** How much does one additional pound of lean body mass increase a person's basal metabolic rate?

**25.** Find the basal metabolic rate for a person with 100 pounds of lean body mass.

**26.** Find the basal metabolic rate for a person with 130 pounds of lean body mass.

**27.** Find the basal metabolic rate for a person who weighs 160 pounds and has a body fat percentage of 25%.

**28.** Find the basal metabolic rate for a person who weighs 175 pounds and has a body fat percentage of 20%.

**29.** How many calories per day does this person need to take in to maintain his weight?

**30.** How many calories per day does this person need to take in to maintain her weight?

Weight: 180 lb
Waist: 36 in.
Activity level: sedentary

Weight: 130 lb
Wrist: 4.5 in.
Waist: 26 in.
Hips: 33 in.
Forearm: 8 in.
Activity level: extra active

# 10.2 The Olympics

▶ Analyze winning times and heights in the Summer Olympics.

▶ Analyze winning times in the Winter Olympics.

▶ Understand Olympic scoring.

## The Summer Olympics

The Summer Olympics are held every 4 years. The games feature thousands of athletes with over 30 sports and about 400 different events.

EXAMPLE 1 **Analyzing Winning Times**

The graph shows the winning times for the men's 200-meter freestyle swimming event from 1968 through 2008. Describe the pattern in the graph.

**Winning Times for Men's 200-Meter Freestyle (1968–2008)**



*3 wins by USA (1968–2008)*
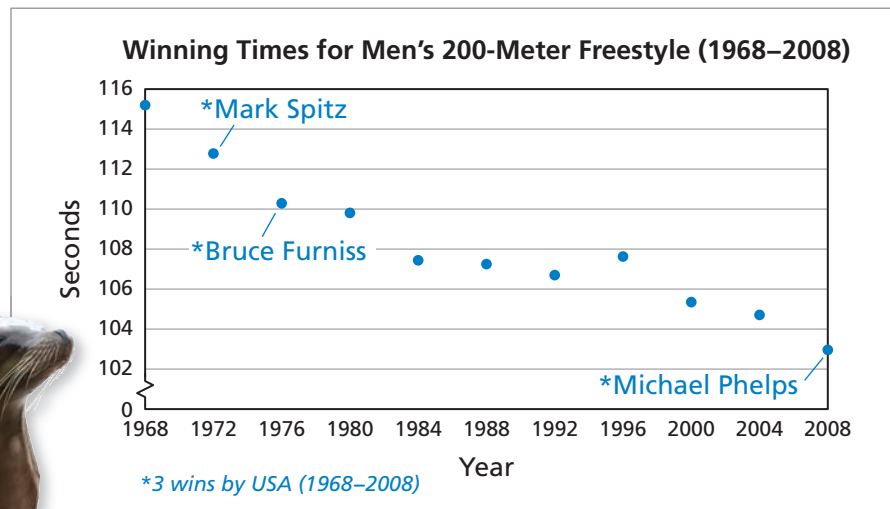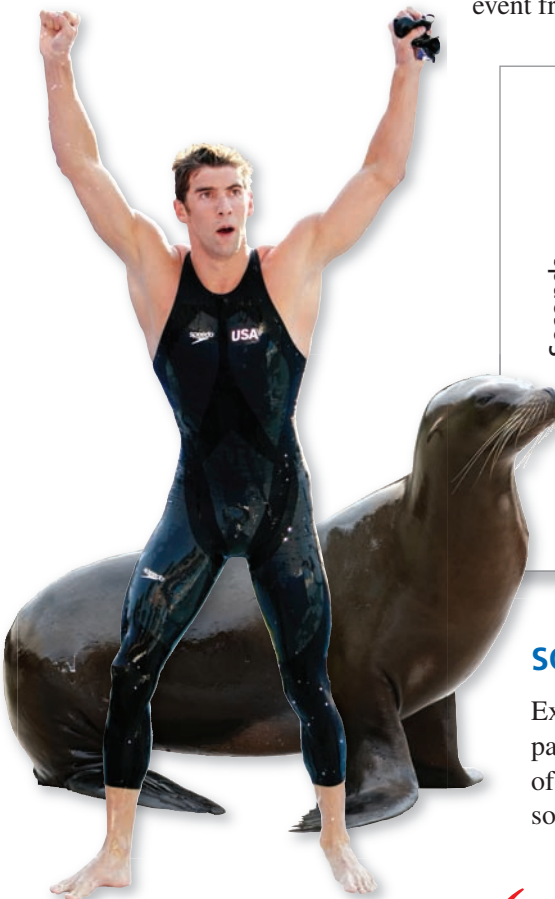
**SOLUTION**

Except for 1996, the winning time has decreased each year. The pattern for the times is roughly linear. At the 2008 Olympics, the use of full-body suits coincided with many new records, which caused some controversy.

At the 2008 Olympics, Michael Phelps's speed in the 200-meter freestyle swimming event was about 4.3 miles per hour. A California sea lion, which is larger than a human, can swim up to 25 miles per hour.

✓ **Checkpoint** Help at Math.andYOU.com

Sketch a graph of the winning times (in seconds) for the women's 200-meter freestyle swimming event from 1968 through 2008. Describe any patterns in the graph.
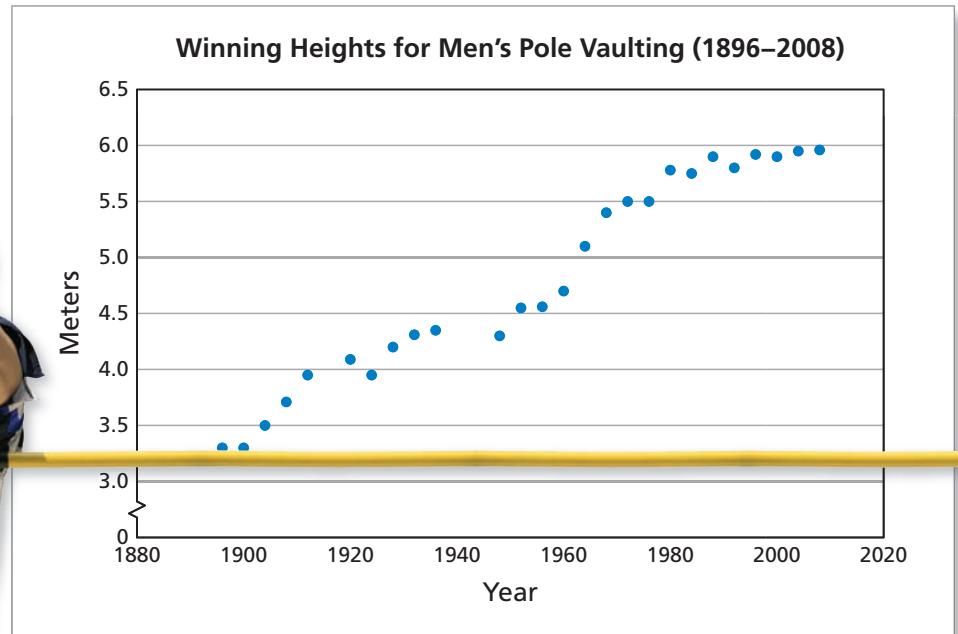
DATA (1968, 130.50), (1972, 123.56), (1976, 119.26), (1980, 118.33), (1984, 119.23), (1988, 117.65), (1992, 117.90), (1996, 118.16), (2000, 118.24), (2004, 118.03), (2008, 114.82)

EXAMPLE 2   **Analyzing Winning Heights**

Pole vaulting has been an event at the Olympic Games since 1896 for men and since 2000 for women. Early Olympians used solid wood or bamboo poles. Describe the pattern of the winning heights for men's pole vaulting from 1896 through 2008. When do you think fiberglass poles were first used?

**Winning Heights for Men's Pole Vaulting (1896–2008)**

Meters (vertical axis): 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5
Year (horizontal axis): 1880, 1900, 1920, 1940, 1960, 1980, 2000, 2020

**SOLUTION**

The winning heights increased until around 1950 when it appeared as though they would level off at around 4.25 meters. Then in the 1950s, pole-vault technology improved with fiberglass poles. With that, the heights started rising again. Now they seem to have leveled off at around 6 meters.

Many years ago, pole vaulters went over the bar with their feet pointing downward. Today, they do a complicated gymnastic maneuver, turning upside down as the jump takes place.

✓ **Checkpoint**

Help at *Math*.and*YOU*.com

Today, high school students can pole vault at heights that would have broken Olympic records in 1950. How is this possible?

## The Winter Olympics

Like the Summer Olympics, the Winter Olympics are also held every 4 years. The 2010 Winter Olympics were held in Vancouver, British Columbia.

**EXAMPLE 3**    **Analyzing Winning Times**

The graph shows the winning times for the men's 1500-meter speed skating event from 1924 through 2010. Describe the pattern in the graph.

**Winning Times for Men's 1500-Meter Speed Skating (1924–2010)**
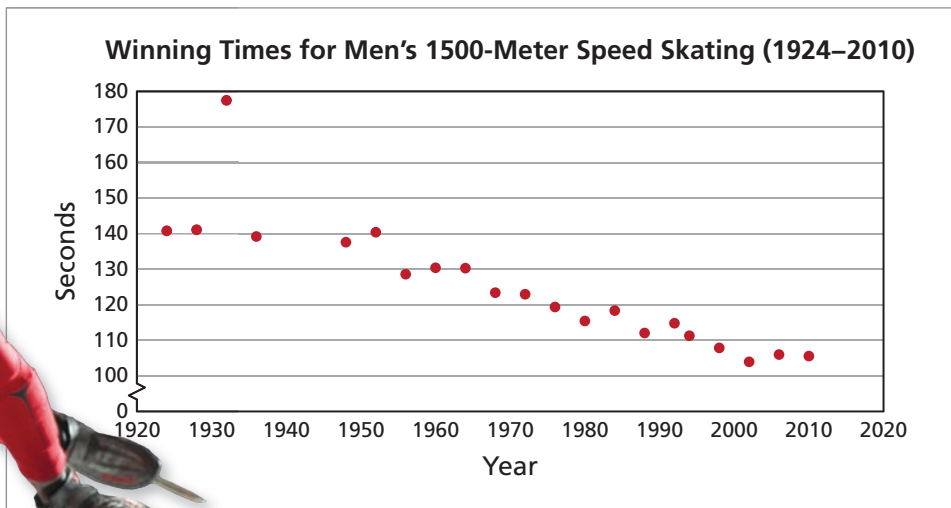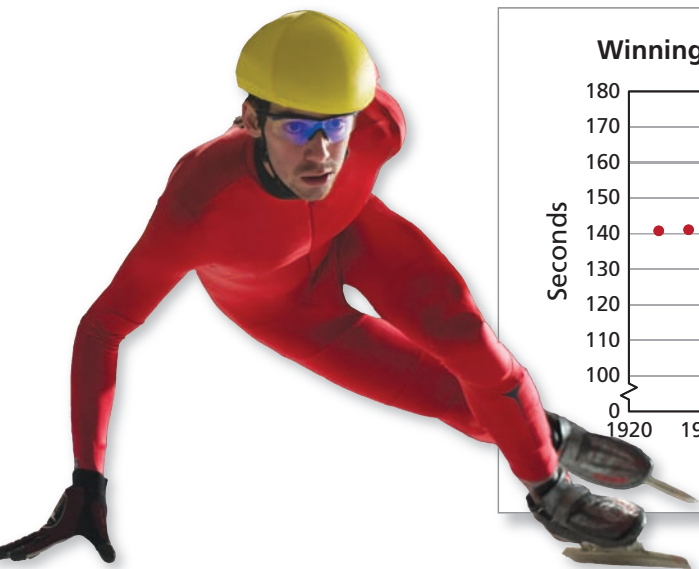


### SOLUTION

One thing that stands out in the graph is that the 1932 winning time is an outlier. The 1500-meter speed skating event at the 1932 Olympics (held at Lake Placid) was not typical. Instead of racing against the clock individually, skaters raced against each other in small packs. This accounts for the time that was longer than the other winning times.

From 1948 through 2010, the winning time generally decreased. The biggest change in technology occurred in the 1998 Olympics when the "clap skate" was first allowed. It has allowed speed skaters to reach speeds that skaters on regular skates could not reach.

Regular Skate          Clap Skate

✓ **Checkpoint**                    Help at *Math*.and**YOU**.com

Sketch a graph of the winning times (in seconds) for the women's 1500-meter speed skating event from 1960 through 2010. Describe any patterns in the graph.

(1960, 145.2), (1964, 142.6), (1968, 142.4), (1972, 140.85),
(1976, 136.58), (1980, 130.95), (1984, 123.42), (1988, 120.68),
(1992, 125.87), (1994, 122.19), (1998, 117.58), (2002, 114.02),
(2006, 115.27), (2010, 116.89)

EXAMPLE 4    **Analyzing Winning Times**

Downhill skiing has been an event at the Winter Olympics since 1948. Before 1964, the skiers' times were measured to the nearest tenth of a second. Since 1964, the times have been measured to the nearest one-hundredth of a second. Describe the pattern of the winning times.

**Winning Times for Men's Downhill Skiing (1948–2010)**



For downhill skiing, remember that the local skiing and weather conditions may affect skiers' times.
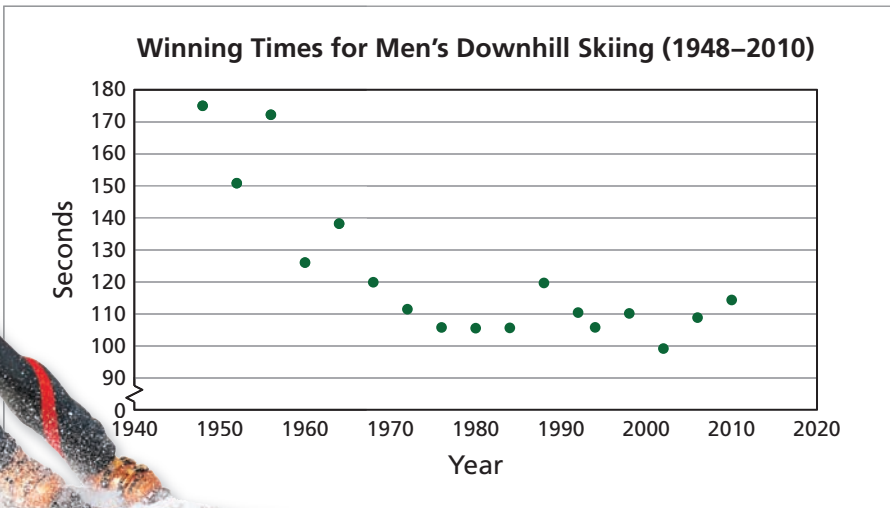
**SOLUTION**

From 1948 through 1976, the winning times decreased dramatically. Since 1968, the winning times seem to have clustered between 105 seconds and 120 seconds. The one time since 1968 that it was out of this range was 2002, when the winning time by Fritz Strobl (Austria) was only 99.13 seconds. To travel about 2 miles in 100 seconds, Strobl had an average speed of about

$$\frac{2 \text{ mi}}{100 \text{ sec}} = \frac{2 \text{ mi}}{100 \text{ sec}} \times \frac{60 \text{ sec}}{1 \text{ min}} \times \frac{60 \text{ min}}{1 \text{ hr}} = 72 \text{ mph}.$$

✓ **Checkpoint**                                    Help at *Math*.and**Y☺U**.com

In 2010, the men's downhill skiing course had a vertical drop of 2799 feet. The course length was 1.929 miles. The time differences between the medalists were the closest in the history of the event at the Olympics. Graphically represent the top 15 times (shown in seconds).


114.31,


114.38,


114.40,


114.52,


114.64,


114.67,


114.82,


114.84,


114.87,


114.88,


115.02,


115.19,


115.19,


115.26,


115.29

## Olympic Scoring

In both the Summer Olympics and the Winter Olympics, the scores for some events are based on subjective decisions by judges. Some examples are gymnastics, diving, figure skating, and ski jumping. Example 5 describes the judging system for diving.

**EXAMPLE 5**   **Analyzing Olympic Scoring**

For a women's diving event, a panel of seven judges evaluate a dive. Each judge awards a score between 0 and 10, as shown in the table. The degree of difficulty of the dive is 1.8. To determine the diver's final score, discard the two highest and two lowest scores awarded by the judges. Then add the remaining scores and multiply by the degree of difficulty. What is the diver's score? Is this a good score?
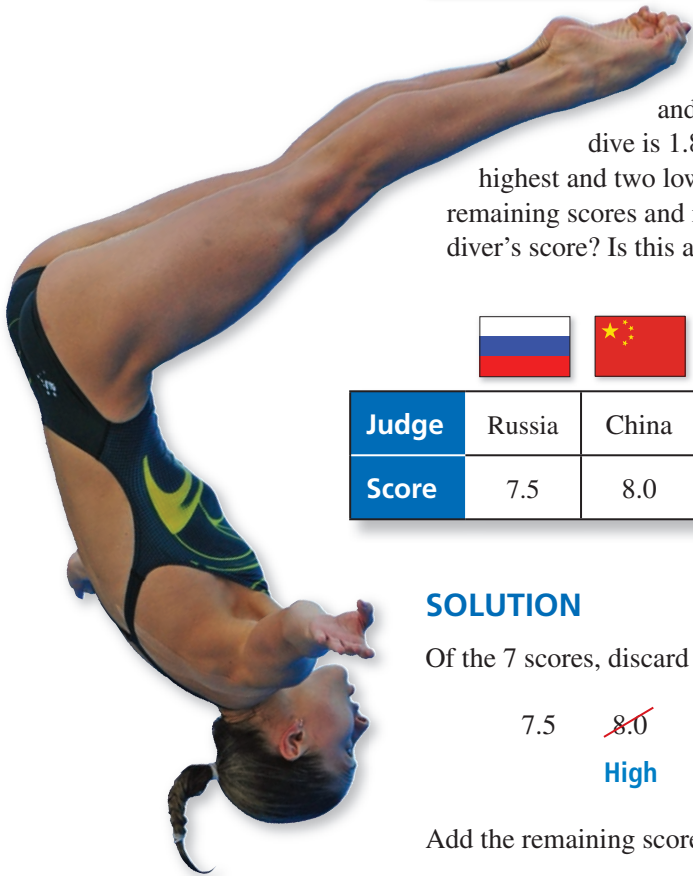
| Judge | Russia | China | Mexico | Germany | Italy | Japan | Brazil |
|-------|--------|-------|--------|---------|-------|-------|--------|
| Score | 7.5 | 8.0 | 6.5 | 8.5 | 7.0 | 7.5 | 7.0 |

### SOLUTION

Of the 7 scores, discard 8.5, 8.0, 7.0, and 6.5.

7.5    ~~8.0~~    ~~6.5~~    ~~8.5~~    ~~7.0~~    7.5    7.0
     **High**   **Low**   **High**   **Low**

Add the remaining scores and multiply by the degree of difficulty.

$$1.8(7.5 + 7.5 + 7.0) = 1.8(22) = 39.6$$

The diver's score is 39.6. This is not particularly good.

✓ **Checkpoint**          Help at *Math*.and**YOU**.com

**VIDEO**   Watch videos of five dives at *Math.andYou.com*. Award a score between 0 and 10 to each dive. (You must award a whole number or a half, as in 6.0 or 8.5. You cannot award a score like 8.3.) Use the following criteria when awarding a score.

- Rate the approach, the height above the diving board, and the acrobatics.
- Toes pointed = good       • Feet touching = good
- No splash = high points   • Ripples = lower points
- Diver straight up and down on entry into water = good

EXAMPLE 6    **Analyzing Olympic Scoring**

Olympic ski jumping scores are based on a point system that combines *style points* and *distance points*.

**Style points:** For style points, a jump is divided into three parts: flight, landing, and outrun. A panel of five judges evaluate the jump. Each judge awards a maximum of 20 points. The highest and lowest scores are discarded, and the remaining three scores are totaled.

**Distance points:** A "K-point" is marked in the landing area. When the K-point is at 120 meters, jumpers are awarded 60 points plus 1.8 points for each meter that they exceed the K-point, or 60 points minus 1.8 points for each meter that they are short of the K-point.

In the 2014 Winter Olympics, women will be allowed to compete in ski jumping for the first time.

Suppose the K-point is at 120 meters. Find the score for a ski jumper with the following style points and distance.

Style points: 17.0, 18.0, 18.5, 20.0, 19.0          Distance: 125 meters

**SOLUTION**

Style points: Discard the 20.0 and 17.0.        $18.0 + 18.5 + 19.0 = 55.5$ points
Distance points: $60 + 1.8(5) = 69$ points
Total points: $55.5 + 69 = 124.5$ points

✓ **Checkpoint**                    Help at *Math*.and**Y**❂**U**.com

Sketch a graph of the winning points for Olympic ski jumping from 1932 through 2010. Describe any patterns in the graph.

(1932, 228.1), (1936, 232.0), (1948, 228.1), (1952, 226.0), (1956, 227.0),
(1960, 227.2), (1964, 230.7), (1968, 231.3), (1972, 219.9), (1976, 234.8),
(1980, 271.0), (1984, 231.2), (1988, 224.0), (1992, 239.5), (1994, 274.5),
(1998, 272.3), (2002, 281.4), (2006, 276.9), (2010, 283.6)

## 10.2 Exercises

**Men's Discus Throw** In Exercises 1–4, use the graph. *(See Examples 1 and 2.)*

**Winning Distances for Men's Discus Throw (1896–2008)**
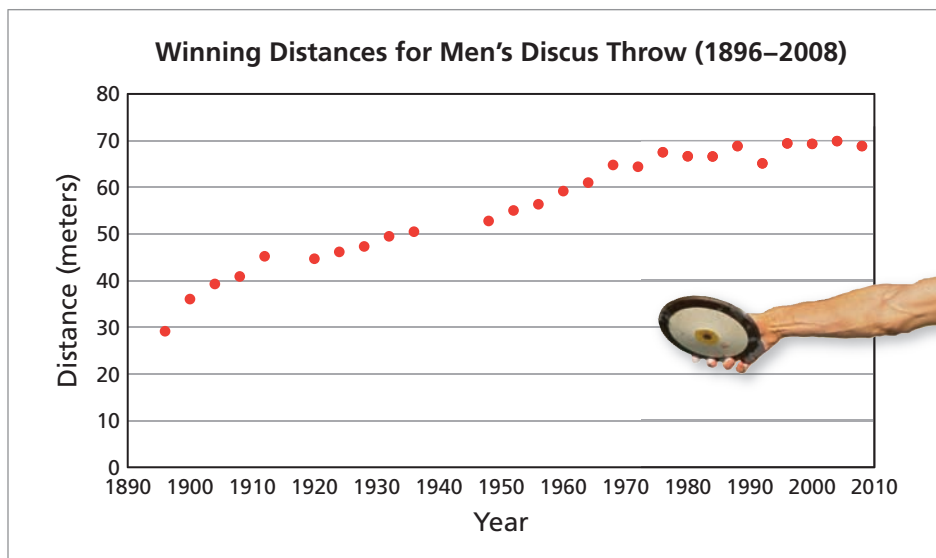


1. Describe any patterns in the graph.

2. The concrete throwing circle was introduced to the Olympic discus throw in 1954. Discuss the impact of the concrete throwing circle on the winning distances.

3. The Olympics were not held in 1940 or 1944 because of World War II. Based on the graph, what do you think the winning distance would have been in 1940 if the Olympics had been held? Explain your reasoning.

4. What is a reasonable expectation for the winning distance in the 2012 Olympics? Explain your reasoning.

**DATA** **Women's Discus Throw** The winning distances for the women's discus throw are shown. In Exercises 5 and 6, use the data. *(See Example 1.)*

(1928, 39.62), (1932, 40.58), (1936, 47.63), (1948, 41.92), (1952, 51.42), (1956, 53.69), (1960, 55.10), (1964, 57.27), (1968, 58.28), (1972, 66.62), (1976, 69.00), (1980, 69.96), (1984, 65.36), (1988, 72.30), (1992, 70.06), (1996, 69.66), (2000, 68.40), (2004, 67.02), (2008, 64.74)

5. Sketch a graph of the winning distances. Describe any patterns in the graph.

6. Sketch a graph that shows the Olympic record for each of the years in the data set.

**Women's 1000-Meter Speed Skating**  In Exercises 7–10, use the graph.  *(See Examples 3 and 4.)*



**Winning Times for Women's
1000-Meter Speed Skating (1960–2010)**
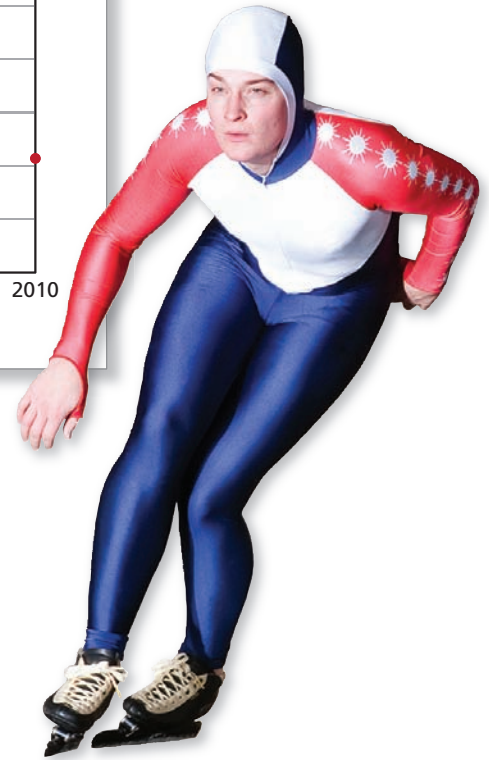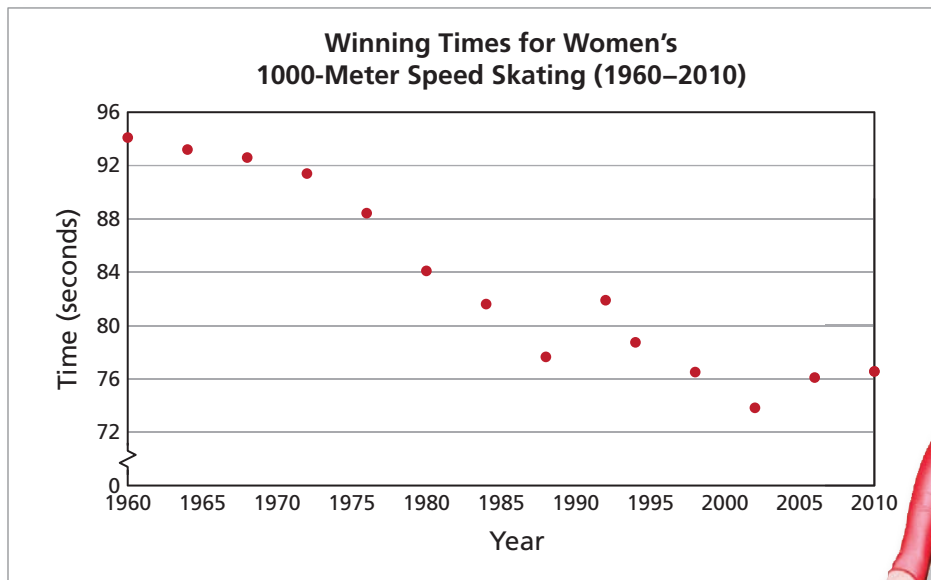
7. Describe any patterns in the graph.

8. What is the Olympic record for the women's 1000-meter speed skating event?

9. What percent of the winning times were Olympic records?

DATA  **10.** Sketch a graph that shows the Olympic record for each year in the data set.

DATA  **11.** **Finishing Times**  Graphically represent the top 12 times for the 2010 Olympic women's 1000-meter speed skating event (shown in seconds).  *(See Example 4.)*

76.56,    76.58,    76.72,    76.78,    76.80,    76.94,

77.08,    77.28,    77.37,    77.43,    77.46,    77.53
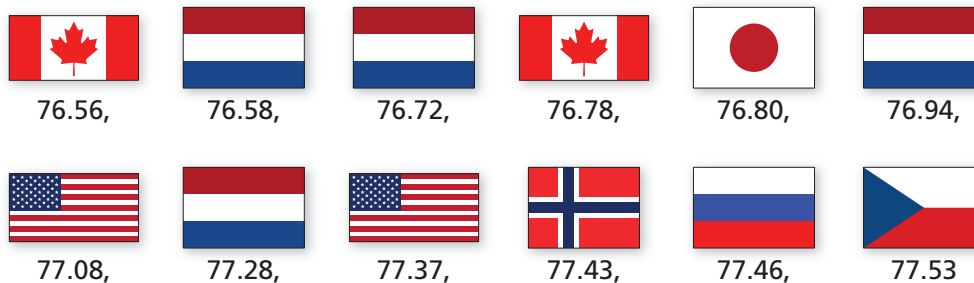
DATA  **12.** **Men's 1000-Meter Speed Skating**  The winning times (in seconds) for the men's 1000-meter speed skating event are shown. Sketch a graph and describe any patterns.  *(See Example 3.)*

(1976, 79.32),  (1980, 75.18),  (1984, 75.80),  (1988, 73.03),  (1992, 74.85),
(1994, 72.43),  (1998, 70.64),  (2002, 67.18),  (2006, 68.90),  (2010, 68.94)

**Diving Scores** In Exercises 13–16, find the diver's score. *(See Example 5.)*

**13.** Degree of difficulty: 1.7

Judges' scores:

| Judge | Russia | China | Mexico | Germany | Italy | Japan | Brazil |
|-------|--------|-------|--------|---------|-------|-------|--------|
| Score | 7.0 | 8.5 | 8.5 | 9.0 | 9.5 | 9.0 | 8.5 |

**14.** Degree of difficulty: 3.6

Judges' scores:

| Judge | Russia | China | Mexico | Germany | Italy | Japan | Brazil |
|-------|--------|-------|--------|---------|-------|-------|--------|
| Score | 7.0 | 7.5 | 7.0 | 8.5 | 7.5 | 8.0 | 6.5 |

**15.** Degree of difficulty: 2.4

Judges' scores:

| Judge | Russia | China | Mexico | Germany | Italy | Japan | Brazil |
|-------|--------|-------|--------|---------|-------|-------|--------|
| Score | 9.0 | 7.5 | 8.0 | 8.5 | 6.5 | 7.0 | 7.5 |

**16.** Degree of difficulty: 2.6

Judges' scores:

| Judge | Russia | China | Mexico | Germany | Italy | Japan | Brazil |
|-------|--------|-------|--------|---------|-------|-------|--------|
| Score | 8.0 | 9.5 | 9.5 | 9.0 | 8.5 | 9.0 | 10.0 |

**Ski Jumping Scores** In Exercises 17–20, find the score for a ski jumper with the given style points and distance. Assume the K-point is at 120 meters. *(See Example 6.)*

**17.** Style Points: 17.0, 18.5, 17.5, 17.5, 19.0     Distance: 124 meters

**18.** Style Points: 18.0, 19.0, 20.0, 19.5, 19.5     Distance: 130 meters

**19.** Style Points: 16.5, 17.0, 18.5, 17.0, 19.0     Distance: 110 meters

**20.** Style Points: 18.0, 19.5, 19.5, 20.0, 19.5     Distance: 115 meters

# ▶ Extending Concepts

DATA    **Men's Super Combined**  In the super combined, skiers compete in two events. The finishing times for the events are added together to determine the skier with the fastest combined time. The graphs show the finishing times for both events for the skiers with the top 6 combined times in 2010. In Exercises 21–23, use the graphs.
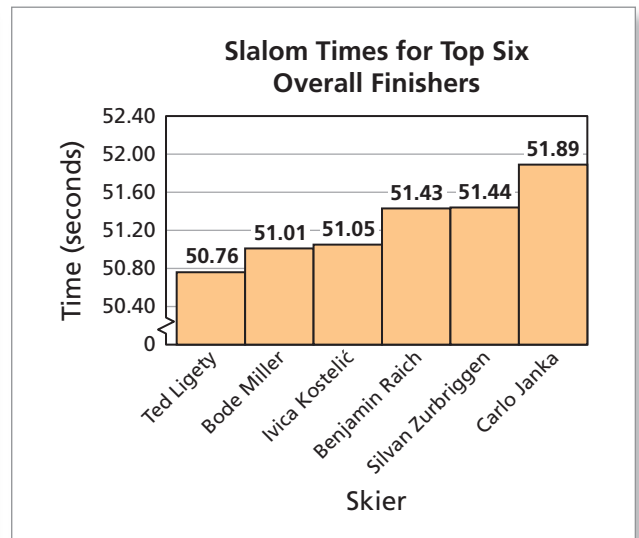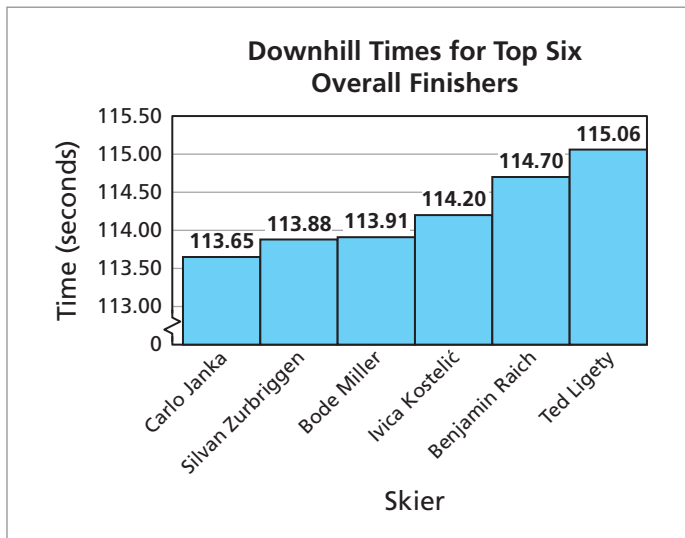


**Downhill Times for Top Six Overall Finishers**



**Slalom Times for Top Six Overall Finishers**

**21.** Find the combined time for each skier and make a bar graph showing the combined times from least to greatest. Who won the gold medal? the silver medal? the bronze medal?

**22.** Determine how many seconds the gold medalist finished ahead of each of the other skiers in the top six. Then graphically represent the information.

**23.** Aksel Lund Svindal had the fastest finishing time in the downhill portion at 113.15 seconds, but he skied off course in the slalom and did not finish that portion. What finishing time in the slalom would have won him the gold medal? Compare this time to the finishing times in the slalom graph.

DATA    **Women's Super Combined**  The table shows the finishing times for the 2 events in the women's super combined for the skiers with the top 6 combined times in 2010. In Exercises 24–26, use the table.

**24.** Find the combined time for each skier and make a bar graph showing the combined times from least to greatest. Who won the gold medal? the silver medal? the bronze medal?

**25.** Determine how many seconds the gold medalist finished ahead of each of the other skiers in the top six. Then graphically represent the information.

**26.** Lindsey Vonn had the fastest finishing time in the downhill portion at 84.16 seconds, but she missed a gate and fell in the slalom and did not finish that portion. What finishing time in the slalom would have won her the gold medal? Compare this time to the slalom finishing times for the top six overall finishers.

| Times for Top Six Overall Finishers in Women's Super Combined, 2010 | | |
|---|---|---|
| | Time (seconds) | |
| **Skier** | **Downhill** | **Slalom** |
| Julia Mancuso | 84.96 | 45.12 |
| Tina Maze | 85.97 | 44.56 |
| Anja Pärson | 85.57 | 44.62 |
| Maria Riesch | 84.49 | 44.65 |
| Fabienne Suter | 85.29 | 45.56 |
| Kathrin Zettel | 86.01 | 44.49 |

# 10.1–10.2   Quiz

**DATA**

**Health Calculations**  In Exercises 1–5, use the information about the javelin thrower.

1.  Find the woman's body fat percentage. ← Use the Body Fat Formula for women on page 453.

2.  Find the woman's body mass index. ← Use the Body Mass Index Formula on page 453.

3.  Find the woman's maximum heart rate. ← Use the Maximum Heart Rate Formula on page 454.

4.  Find the woman's basal metabolic rate. ← Use the Basal Metabolic Rate Formula on page 455.

5.  The woman is extra active and has a daily calorie intake of 2400 calories. Would you expect the woman to be losing weight or gaining weight? Explain your reasoning.

**Women's Javelin Throw**  In Exercises 6–8, use the graph.

Weight: 135 lb
Height: 64 in.
Wrist: 5.5 in.
Waist: 29 in.
Hips: 34 in.
Forearm: 7.5 in.
Age: 22 yr

**Winning Distances for Women's Javelin Throw (1932–2008)**
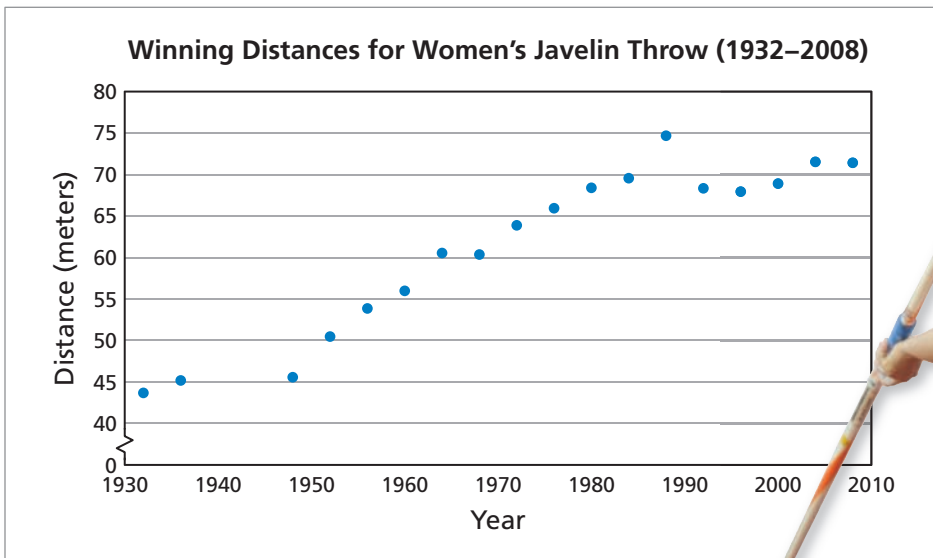


6.  Describe any patterns in the graph.

7.  What percent of the winning distances were Olympic records?

8.  What is the Olympic record for the women's javelin throw?
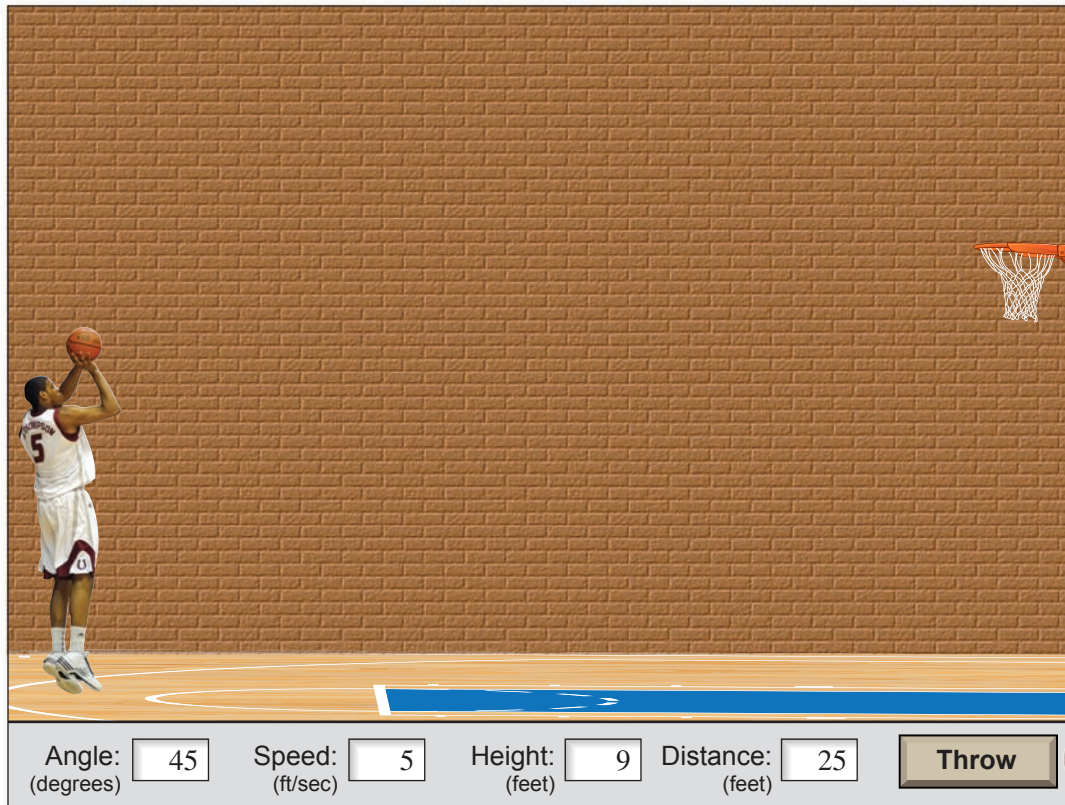
# *Math & Projectiles*

## PROJECT: Modeling the Path of a Basketball

Basketball was invented by James Naismith in 1891. At the time, he was an instructor at the YMCA in Springfield, Massachusetts. He called the game "Basket Ball." The first official game was played on December 21, 1891.

**1.** The path that a basketball takes depends on several things. Here are three of them.

- The height from which the basketball is tossed
- The angle at which the basketball is tossed (as measured from the horizontal)
- The initial speed of the basketball

Use the *Basketball Simulator* at *Math.andYou.com* to find a height, angle, and speed that successfully make a basket. Set the distance at 25 feet.

| Angle: (degrees) | 45 | Speed: (ft/sec) | 5 | Height: (feet) | 9 | Distance: (feet) | 25 | **Throw** |
|---|---|---|---|---|---|---|---|---|

**2.** You are 25 feet from the basket and release the ball from a height of 9 feet. Is it possible to make a basket by shooting with an initial speed of 5 feet per second? Use the simulator to verify your answer.

**3.** You are 25 feet from the basket and release the ball from a height of only 7 feet. Is it possible to make a basket by shooting with an initial speed of 5 feet per second? Use the simulator to verify your answer. Compare your answer with your answer in Exercise 2.

**4.** Is it true that "the greater the speed, the smaller the angle you should throw the ball"? Explain your answer.
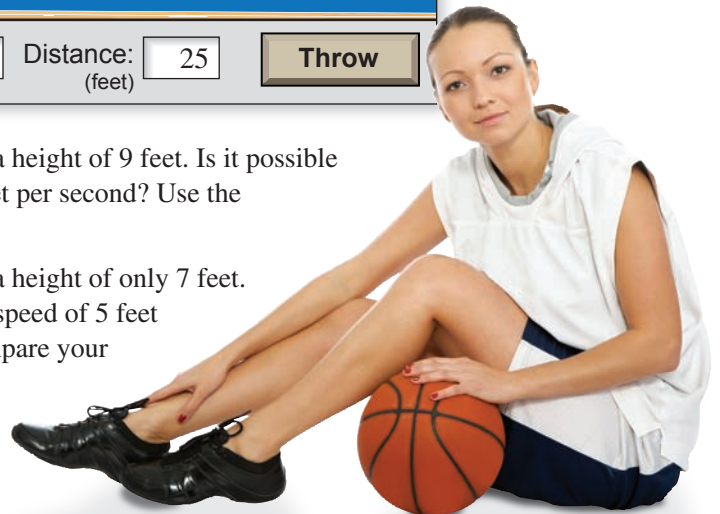
# 10.3 Professional Sports

▶ Use mathematics to analyze baseball statistics.

▶ Use mathematics to analyze football statistics.

▶ Use mathematics to analyze statistics in other professional sports.

## Baseball Statistics

The statistics of baseball seem to have attracted more interest than the statistics of any other professional sport. There are dozens of different types of statistics, including batting, pitching, fielding, and baserunning statistics.
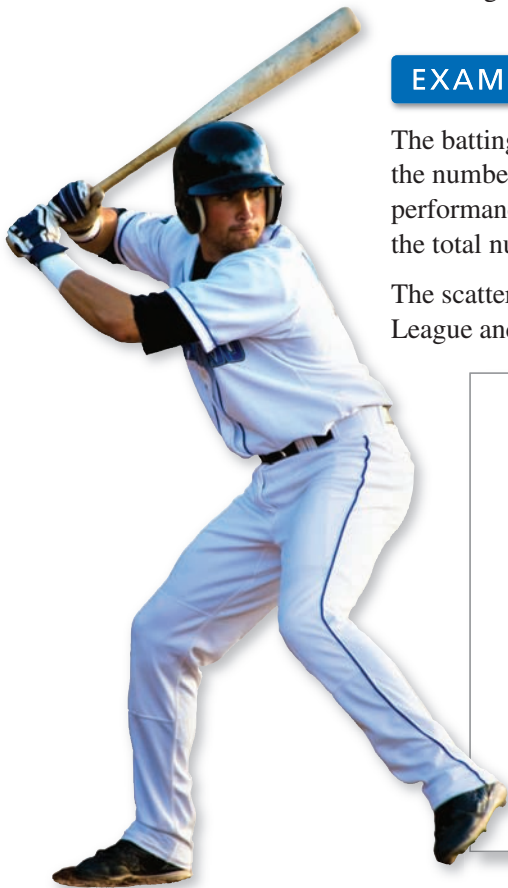
### EXAMPLE 1 Analyzing Batting Statistics

The batting average of a *player* or of a team is the ratio of the number of hits to the number of "at bats." The earned run average (ERA) is a measure of a *pitcher's* performance obtained by dividing the total number of earned runs allowed by the total number of innings pitched, and then multiplying by nine.

The scatter plot compares the batting averages and the ERAs of the American League and the National League from 2000 through 2010. What can you conclude?



**MLB Stats (2000–2010)**
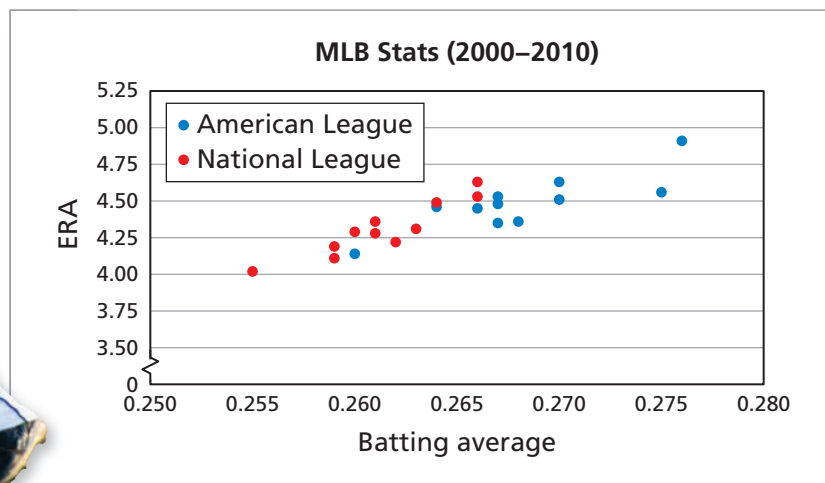
### SOLUTION

The American League had better (higher) batting averages. The National League had better (lower) earned run averages.

✔ **Checkpoint**      Help at *Math*.and**Y⊙U**.com

It is possible for 1 player to have a higher batting average than another player 2 years in a row, but to have a lower batting average when the 2 years are combined. Here is an example. How can you explain this?

|  | 1995 | | 1996 | | Combined | |
|---|---|---|---|---|---|---|
| **Derek Jeter** | 12/48 | .250 | 183/582 | .314 | 195/630 | .310 |
| **David Justice** | 104/411 | .253 | 45/140 | .321 | 149/551 | .270 |

## EXAMPLE 2   Analyzing the Strike Zone

The strike zone is an imaginary box above home plate. The dimensions of the strike zone vary according to the batter's height and stance. It is defined as "that area over home plate the upper limit of which is a horizontal line at the midpoint between the top of the shoulders and the top of the uniform pants, and the lower level is a line at the hollow beneath the kneecap."

In his book *The Science of Hitting*, Ted Williams figured that for a man of his height, the official strike zone would be 7 baseballs wide by 11 baseballs high, as shown. The numbers inside the circles are Williams's estimates of his batting average on balls pitched in that part of the strike zone.

Which part of the strike zone did Williams consider his weakest? Explain your reasoning.

Ted Williams was voted the American League Most Valuable Player twice. He led the league in batting 6 times and had a career batting average of .344.

### SOLUTION

Ted Williams was a left-handed batter. The lowest batting averages are in the lower right corner of the diagram. So, Williams believed that his weakest area was farthest from his body and down, as shown.

This might not be true for all batters, but Williams must have believed that when he had to extend his arms downward, he was not able to obtain the power he wanted.

Weakest area

## ✓ Checkpoint

Which part of the strike zone did Williams consider his strongest? Explain.

Help at Math.andYOU.com

## Football Statistics

### EXAMPLE 3    Calculating Quarterback Ratings

The NFL uses a complicated formula to rate a quarterback's passing efficiency. It involves the following four categories.

**a.** Completion percentage:  $a = 5\left(\dfrac{\text{completions}}{\text{attempts}} - 0.3\right)$

**b.** Average yards per attempt:  $b = 0.25\left(\dfrac{\text{yards}}{\text{attempts}} - 3\right)$

**c.** Touchdown pass percentage:  $c = 20\left(\dfrac{\text{touchdowns}}{\text{attempts}}\right)$

**d.** Interception percentage:  $d = 2.375 - 25\left(\dfrac{\text{interceptions}}{\text{attempts}}\right)$

If the result of any of these is greater than 2.375, the result is lowered to 2.375. If the result is negative, it is raised to 0. The quarterback's passer rating is

$$\text{Passer rating} = 100\left(\frac{a + b + c + d}{6}\right).$$

On November 25, 2010, Tom Brady completed a game with a perfect passer rating. He had 21 completions, 27 attempts, 341 yards, 4 touchdowns, and 0 interceptions. What was Brady's passer rating in that game?

Tom Brady is a quarterback for the New England Patriots. As of 2011, he played in four Super Bowls, winning three of them. In 2005, he was named *Sports Illustrated*'s Sportsman of the Year.

### SOLUTION

$a = 5\left(\dfrac{\text{completions}}{\text{attempts}} - 0.3\right) = 5\left(\dfrac{21}{27} - 0.3\right) \approx 2.39$    <span style="color:red">Lowered to 2.375.</span>

$b = 0.25\left(\dfrac{\text{yards}}{\text{attempts}} - 3\right) = 0.25\left(\dfrac{341}{27} - 3\right) \approx 2.41$    <span style="color:red">Lowered to 2.375.</span>

$c = 20\left(\dfrac{\text{touchdowns}}{\text{attempts}}\right) = 20\left(\dfrac{4}{27}\right) \approx 2.96$    <span style="color:red">Lowered to 2.375.</span>

$d = 2.375 - 25\left(\dfrac{\text{interceptions}}{\text{attempts}}\right) = 2.375 - 25(0) = 2.375$

$\text{Passer rating} = 100\left(\dfrac{2.375 + 2.375 + 2.375 + 2.375}{6}\right) \approx 158.3$

So, Brady's rating was 158.3. This is the highest possible rating in this rating system. That is why it is called a perfect passer rating.

### ✓ Checkpoint

Help at *Math*.and**Y⊙U**.com

Which of the following had perfect passer ratings? Explain.

|  |  | CMP | ATT | YDS | TD | INT |
|---|---|---|---|---|---|---|
| Ben Roethlisberger | December 20, 2007 | 16 | 20 | 261 | 3 | 0 |
| Kurt Warner | September 14, 2008 | 19 | 24 | 361 | 3 | 0 |
| Eli Manning | October 11, 2009 | 8 | 10 | 173 | 2 | 0 |
| Drew Brees | November 30, 2009 | 18 | 23 | 371 | 5 | 0 |

EXAMPLE 4    **Describing the Probability of Winning**

Statistics from NFL games have been analyzed to produce the following graph, which compares the probability of winning for the team with possession of the ball, the point difference, and the minutes remaining in the game. The graph is based on all of the regular season NFL games from 2000 through 2007.

**Win Probability by Lead**



At any given time in a football game, a team's probability of winning depends on its lead. When there is not much time left in the game, the team with the lead is likely to win even if it is only ahead by 7 points or less. Surprises do, however, occur. For instance, on November 11, 2010, the Atlanta Falcons scored a touchdown and beat the Baltimore Ravens with only 20 seconds left in the game.
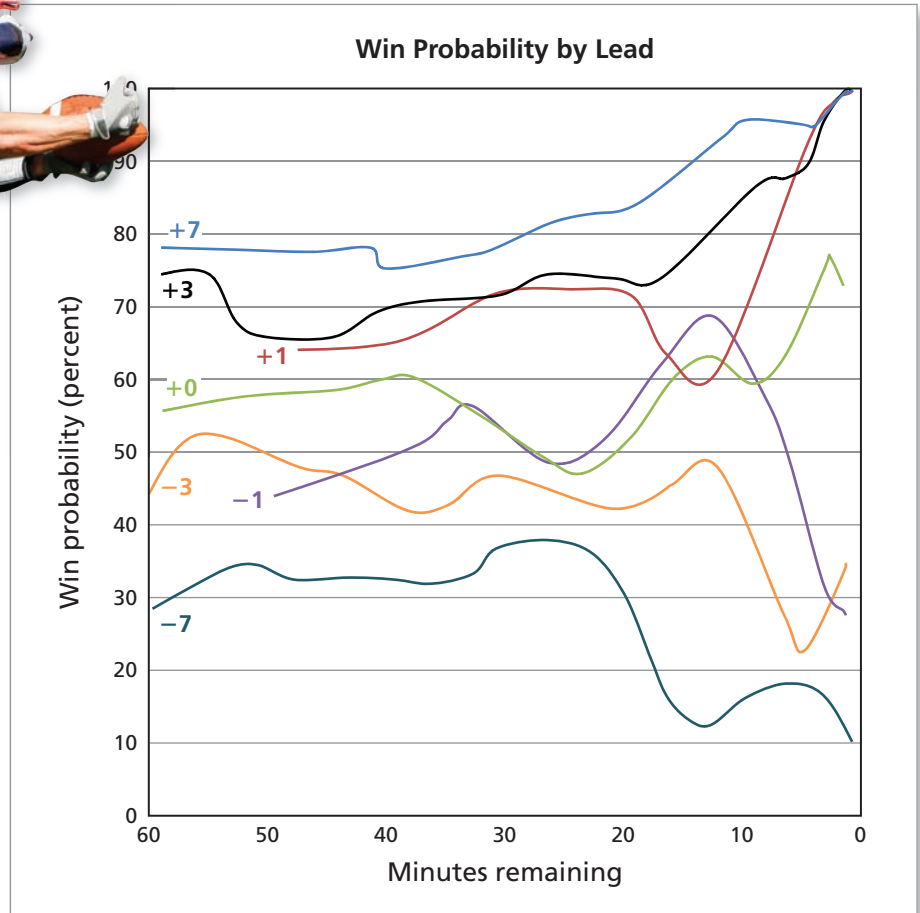
Describe the implications of the graph for a team that is 3 points ahead.

**SOLUTION**

A team that is 3 points ahead and has possession of the ball is likely to win, regardless of the number of minutes remaining.

• With 20 to 30 minutes remaining, the team's chance of winning is about 70%.

• With 10 minutes remaining, the team's chance of winning increases to about 85%.

• With 5 minutes remaining, the team's chance of winning is about 90%.

✓ **Checkpoint**                                    Help at *Math*.and*YOU*.com

Suppose you are the coach of a football team. Explain how you could use the graph above to make decisions about how your team plays. Describe the different strategies you would use depending on the point difference between your team and your opponent.
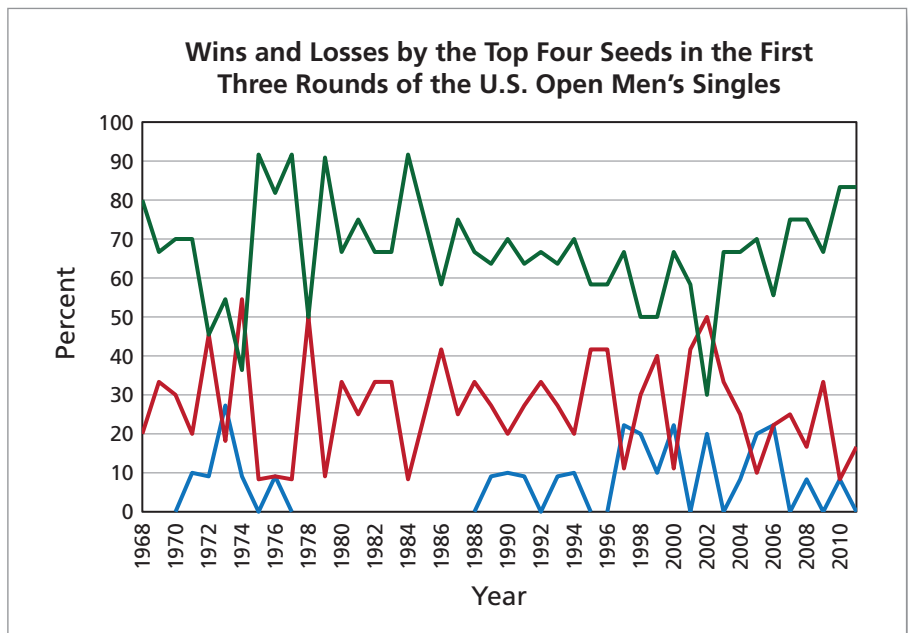
## Other Professional Sports

**EXAMPLE 5**    **Analyzing U.S. Open Wins and Losses**

The graph shows the wins and losses by the "top four seeds" (players with the best rankings prior to the tournament) in the first three rounds of the U.S. Open Tennis Men's Singles Tournament. The **green line** shows straight set wins, the **red line** shows wins when they dropped one or two sets, and the **blue line** shows losses.

> In men's tennis, a player must win 3 out of 5 sets. A "straight set win" means that the player won the first 3 sets. (In women's tennis, a player must win 2 out of 3 sets. A "straight set win" means that the player won the first 2 sets.)

> In men's tennis, a player can drop 1 or 2 sets and still win the match. This can happen in 9 different orders: L-W-W-W, W-L-W-W, W-W-L-W, L-L-W-W-W, L-W-L-W-W, L-W-W-L-W, W-L-L-W-W, W-L-W-L-W, and W-W-L-L-W. (In women's tennis, a player can drop 1 set and still win the match. This can happen in only 2 orders: L-W-W and W-L-W.)

**Wins and Losses by the Top Four Seeds in the First Three Rounds of the U.S. Open Men's Singles**



What observations can you make from this graph?

## SOLUTION

The most obvious observation is that the "seeding system" really works! The players who entered the U.S. Open as 1 of the top 4 ranked players were matched against players with lower rankings, and the top 4 seeds appeared to win easily. Another observation is that the green line always exceeded the red line, except in 1972, 1974, and 2002.

Seeding is a system of ranking players in a tournament draw to avoid the highest-ranked players playing against each other in the early stages of the event.

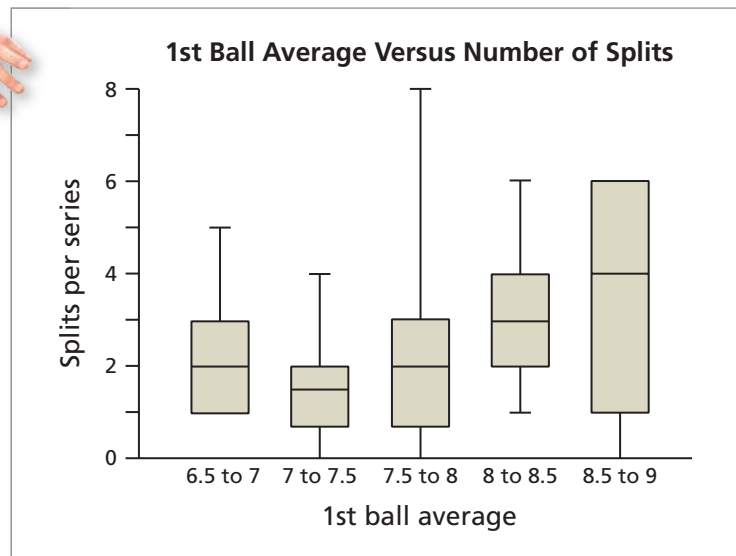✓ **Checkpoint**                           Help at *Math*.and**Y☺U**.com

**a.** Do the 3 lines in the graph always total 100%? Explain.

**b.** In the years in which there were no losses, are the green and red lines mirror images of each other? Explain.

EXAMPLE 6    **Analyzing Splits and Strikes**

The graphs show that the greater a bowler's skill, the tighter the range of locations where the bowling ball hits the pins.



Skill:    **Low**                **Medium**                **High**

Professional bowlers almost always get 8, 9, or 10 pins down on their first ball. The box-and-whisker plots compare the number of splits a bowler gets to his or her first-ball average. What does this graph show?



**1st Ball Average Versus Number of Splits**

**SOLUTION**

Overall, the box-and-whisker plots show that the median number of splits a bowler gets in a series (3 games) increases as his or her first-ball average increases. This makes sense because you cannot have a high pin average on the first ball unless you are hitting the head pin almost every time. The bad news is that when you hit the head pin, you have an increased likelihood of ending up with a split—a result in which the standing pins are "split" with a space between them.

✓ **Checkpoint**                    Help at *Math*.and**YOU**.com

Bowling pins are numbered from 1 through 10. The "7-10 split" is considered to be the worst. List three other possible splits. Include a diagram with your answer.

As of 2010, Pete Weber won 35 Professional Bowlers Association (PBA) Tour events, including 8 major titles. His 35th tour win on April 4, 2010, secured him the position of 3rd place on the all-time PBA Tour titles list.

# 10.3 Exercises

**DATA**

**Hitting Statistics** **The 2010 hitting statistics of 10 Major League Baseball players are shown. In Exercises 1–6, use a spreadsheet.** *(See Examples 1 and 2.)*

| PLAYER | TEAM | POS | G | AB | R | H | 2B | 3B | HR | RBI | TB | BB | SO | SB | CS | SF | HBP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| M Cabrera | DET | 1B | 150 | 548 | 111 | 180 | 45 | 1 | 38 | 126 | 341 | 89 | 95 | 3 | 3 | 8 | 3 |
| S Choo | CLE | OF | 144 | 550 | 81 | 165 | 31 | 2 | 22 | 90 | 266 | 83 | 118 | 22 | 7 | 2 | 11 |
| P Fielder | MIL | 1B | 161 | 578 | 94 | 151 | 25 | 0 | 32 | 83 | 272 | 114 | 138 | 1 | 0 | 1 | 21 |
| J Hamilton | TEX | OF | 133 | 518 | 95 | 186 | 40 | 3 | 32 | 100 | 328 | 43 | 95 | 8 | 1 | 4 | 5 |
| J Mauer | MIN | C | 137 | 510 | 88 | 167 | 43 | 1 | 9 | 75 | 239 | 65 | 53 | 1 | 4 | 6 | 3 |
| J Morneau | MIN | 1B | 81 | 296 | 53 | 102 | 25 | 1 | 18 | 56 | 183 | 50 | 62 | 0 | 0 | 2 | 0 |
| A Pujols | STL | 1B | 159 | 587 | 115 | 183 | 39 | 1 | 42 | 118 | 350 | 103 | 76 | 14 | 4 | 6 | 4 |
| J Thome | MIN | DH | 108 | 276 | 48 | 78 | 16 | 2 | 25 | 59 | 173 | 60 | 82 | 0 | 0 | 2 | 2 |
| J Votto | CIN | 1B | 150 | 547 | 106 | 177 | 36 | 2 | 37 | 113 | 328 | 91 | 125 | 16 | 5 | 3 | 7 |
| K Youkilis | BOS | 1B | 102 | 362 | 77 | 111 | 26 | 5 | 19 | 62 | 204 | 58 | 67 | 4 | 1 | 5 | 10 |

1. Which player had the greatest batting average (AVG) in 2010?

2. Slugging percentage (SLG) is a statistic used to measure the power of a hitter. The formula to calculate slugging percentage is shown. Find the slugging percentage for each player.

$$SLG = \frac{TB}{AB}$$

3. On-base percentage (OBP) is a statistic used to measure how often a player reaches base. The formula to calculate on-base percentage is shown. Find the on-base percentage for each player.

$$OBP = \frac{H + BB + HBP}{AB + BB + HBP + SF}$$

4. Considering AVG, SLG, and OBP, which player would you choose as the most valuable player (MVP)? Explain your reasoning.

5. Is it possible for a player to have an OBP that is less than his AVG? Explain your reasoning.

6. Albert Pujols had a stolen base percentage (SB%) of 77.8%. Write a formula using abbreviations to calculate SB%.
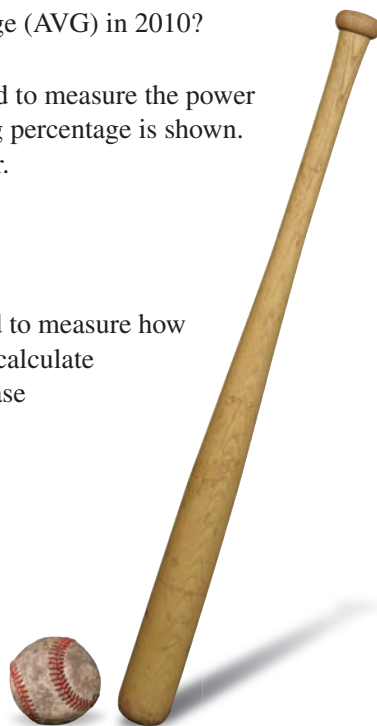
**Abbreviations**

| | |
|---|---|
| 2B | Doubles |
| 3B | Triples |
| AB | At Bats |
| AVG | Batting Average |
| BB | Bases on Balls (Walks) |
| CS | Caught Stealing |
| G | Games Played |
| H | Hits |
| HBP | Hit by Pitch |
| HR | Home Runs |
| OBP | On-base Percentage |
| R | Runs Scored |
| RBI | Runs Batted In |
| SB% | Stolen Base Percentage |
| SB | Stolen Bases |
| SF | Sacrifice Flies |
| SLG | Slugging Percentage |
| SO | Strikeouts |
| TB | Total Bases |
| TPA | Total Plate Appearances |

**3rd Down**  The graphs show 3rd down statistics for all of the regular season NFL games from 2000 through 2007. Plays within 2 minutes of the end of a half and all plays within field goal range are excluded. In Exercises 7–14, use the graphs.  *(See Examples 3 and 4.)*

**7.** Do NFL coaches call more pass plays or run plays on 3rd down? Explain your reasoning.

**8.** Describe the pattern in the "3rd Down Success" graph.

**9.** For what "to go" distances are the conversion rates for running and for passing about equal on 3rd down?

**10.** With 2 yards to go on 3rd down, should the defense expect a pass play or a run play? Explain your reasoning.

**11.** With 2 yards to go on 3rd down, should the coach call a pass play or a run play? Explain your reasoning.

**12.** How might a quarterback's passing efficiency influence a coach's play selection on 3rd down and 4 yards to go?

**13.** The optimum mix for running and passing occurs when the conversion rates are about equal. When this happens, the overall conversion rate will be greatest. Is the pass/run ratio at an optimum mix on 3rd down and 1 yard to go? Explain your reasoning.

**14.** Suppose you are the coach of an NFL team. Your team is facing a 3rd down with 5 yards to go. Would you call a pass play or a run play? Explain your reasoning.

**3rd Down Play Balance**

**3rd Down Success**
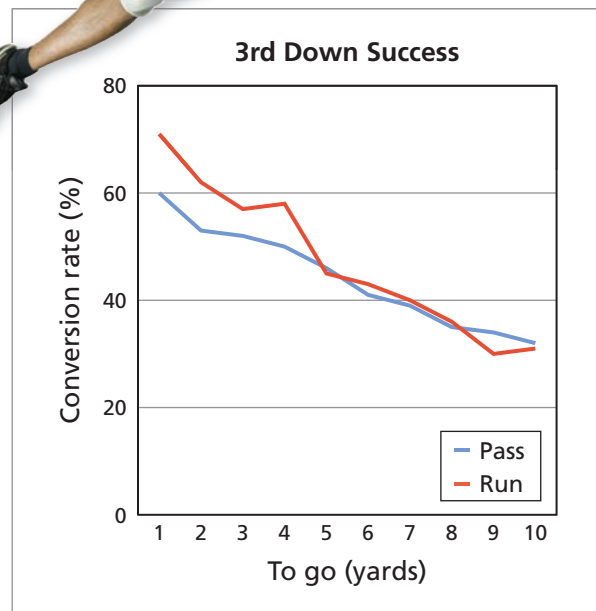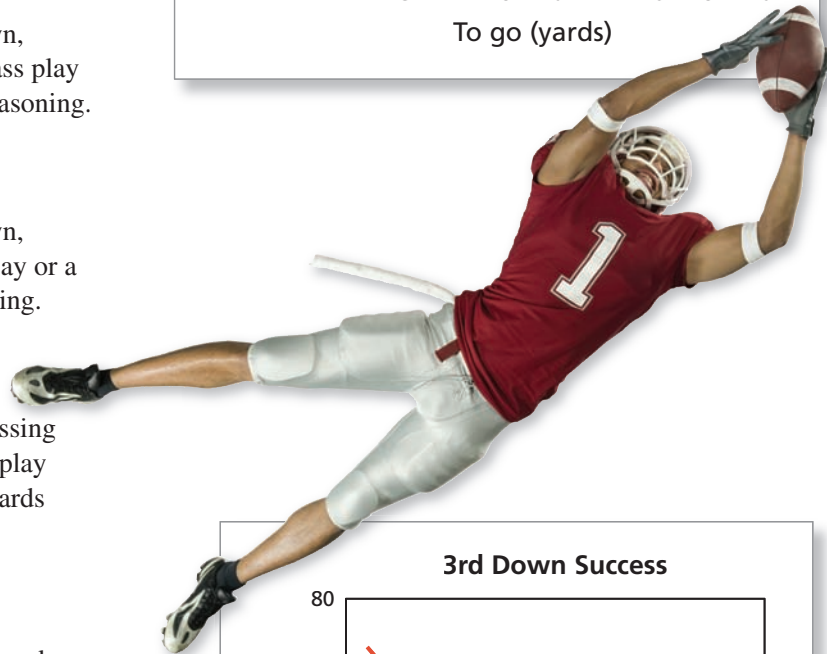
**NBA Shot Chart** The Dallas Mavericks hosted the Oklahoma City Thunder in game 1 of the Western Conference finals on May 17, 2011. The shot chart shows the final field goal data for Kevin Durant and Dirk Nowitzki, the top scorers on each team. In Exercises 15–21, use the shot chart. *(See Examples 5 and 6.)*



Durant (Oklahoma City)    Nowitzki (Dallas)
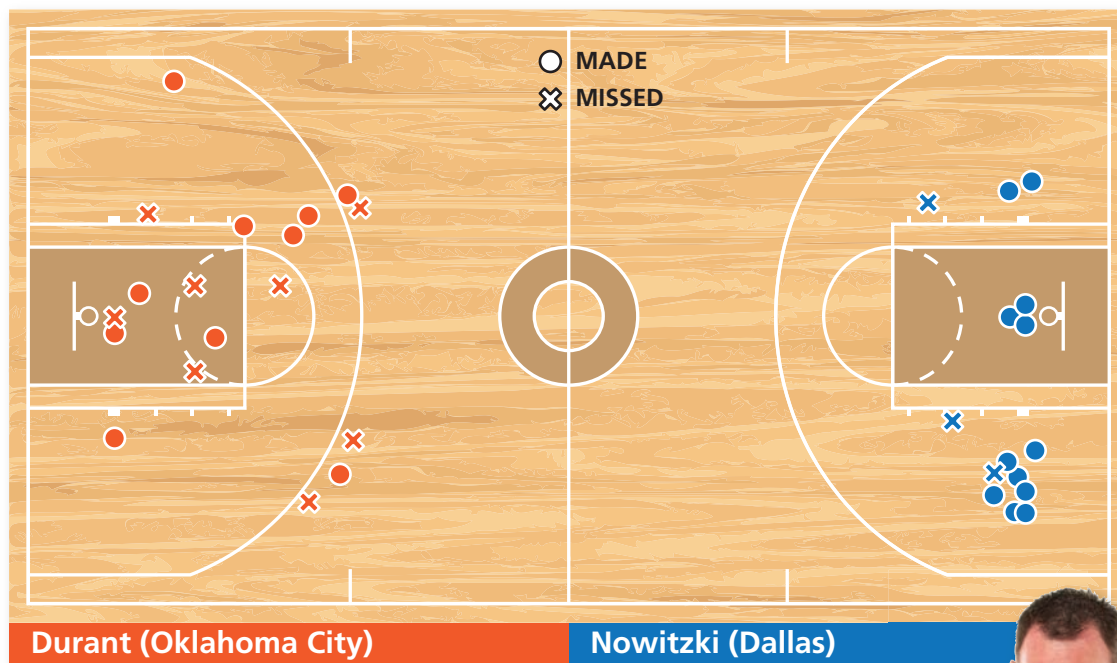
15. Which player attempted more field goals? Explain your reasoning.

16. Field goal percentage is the number of field goals made divided by the number of field goals attempted. Which player had a greater field goal percentage?

17. Nowitzki had one of the most efficient playoff games in NBA history. Not only did he shoot a high percentage from the floor, he also made 24 out of 24 free throws (an NBA playoff record). How many points did Nowitzki score?

18. The Thunder's final score was 112. Nowitzki accounted for 39.7% of the Mavericks' final score. Which team won the game? Explain your reasoning.

19. What was Durant's 3-point field goal percentage?

20. Durant accounted for 35.7% of the Thunder's 112 points. How many free throws did he make during the game?

21. What observations can you make about the variability of shot selection for each player?

22. **Long Range** Would you rather shoot 40% from 3-point range or 50% from inside the 3-point line? Explain your reasoning.

# ▶ Extending Concepts

**Voting Habits of Sports Fans**  The bubble graph shows the results of a study about the voting habits of sports fans in the United States. In Exercises 23–28, use the graph.

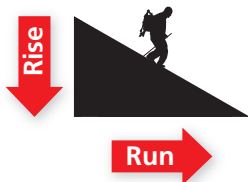

**Voting Habits of Sports Fans in the United States**
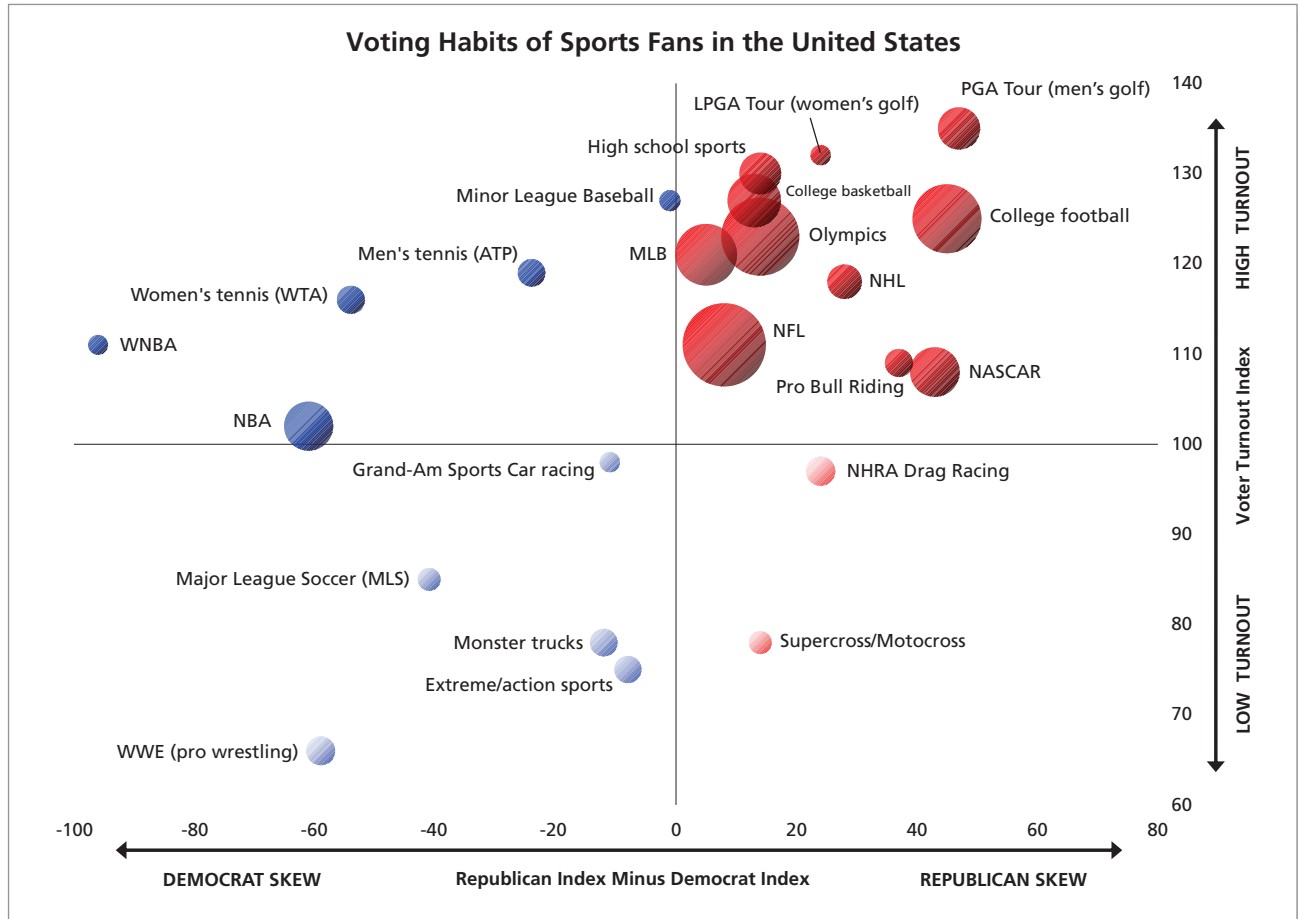
23. What do the size and color of a bubble represent in the graph?

24. Fans of which sport are most likely to vote? least likely to vote?

25. Are sports fans more likely to vote Republican or Democrat? Explain your reasoning.

26. How would you rate the voter turnout of sports fans?

27. Television ad buyers tend to focus on sporting events because sports fans usually watch sporting events live rather than on a DVR machine. This means viewers are unable to skip the ads. Which sporting events should ad buyers target to reach Republican voters? Democratic voters?

28. You are in charge of promoting a Democratic campaign. Would you rather advertise during an NBA game or a women's tennis (WTA) match? Explain your reasoning.

# 10.4 Outdoor Sports

▶ Use mathematics to analyze hiking and mountain climbing.

▶ Use mathematics to analyze kayaking and sailing.

▶ Use mathematics to analyze bicycling and cross-country skiing.

## Hiking and Mountain Climbing

Hiking is an outdoor sport, often done on trails created by the National Forest Service or by a state agency. There are thousands of trails in the United States, such as Glacier Gorge Trail in Estes Park, Colorado, and the Appalachian Trail in the Eastern United States.

Negative Grade

Positive Grade

$$\text{Grade} = \frac{\text{Rise}}{\text{Run}}$$

### EXAMPLE 1 Comparing Grade and Speed

The scatter plot shown was created by a hiker in California. The hiker used a GPS to compare the grade of the trail with his speed. Describe the scatter plot. What does the green line represent?

**Hiking Speed Versus Trail Grade**

- Will Rogers Trail Run
- Iron Mt Hike
- Isoenergy Line

Speed (miles per hour) vs Grade

*hikingscience.blogspot.com*

## SOLUTION

As a general observation, you can see that the hiker was traveling faster on level ground. As the grade changed to uphill or downhill, the hiker's speed decreased. In the Internet post about this experiment, the hiker called the green line his "isoenergy line" because he figured he was exerting the same amount of energy at every point on the line. For instance, walking at 2 miles per hour at a grade of 0.4 uses the same energy as running at 5 miles per hour at a grade of 0.1.

✓ **Checkpoint**                    Help at *Math*.and**YOU**.com

The suggestion for a casual hiking pace on a trail is 1 hour to walk 1.8 miles with a change in elevation of 0.2 mile. The suggestion for a fast hiking pace is 30 minutes. How do these speeds compare to the scatter plot shown in Example 1?
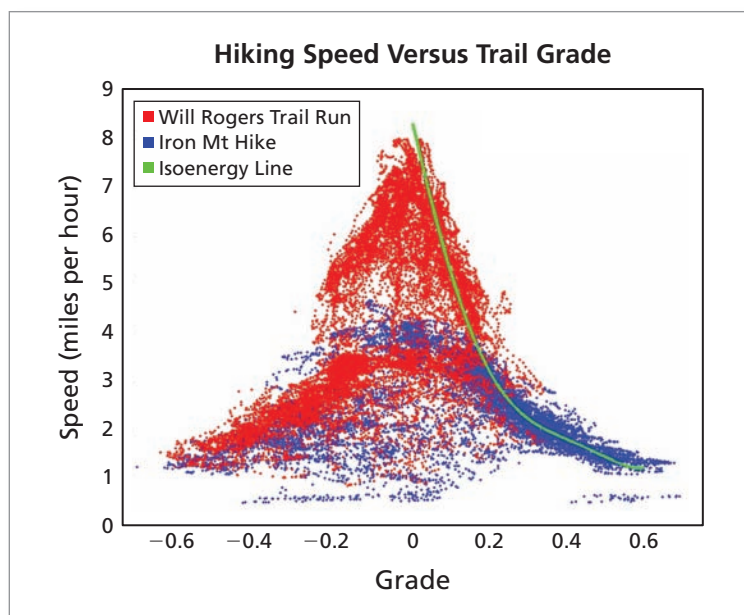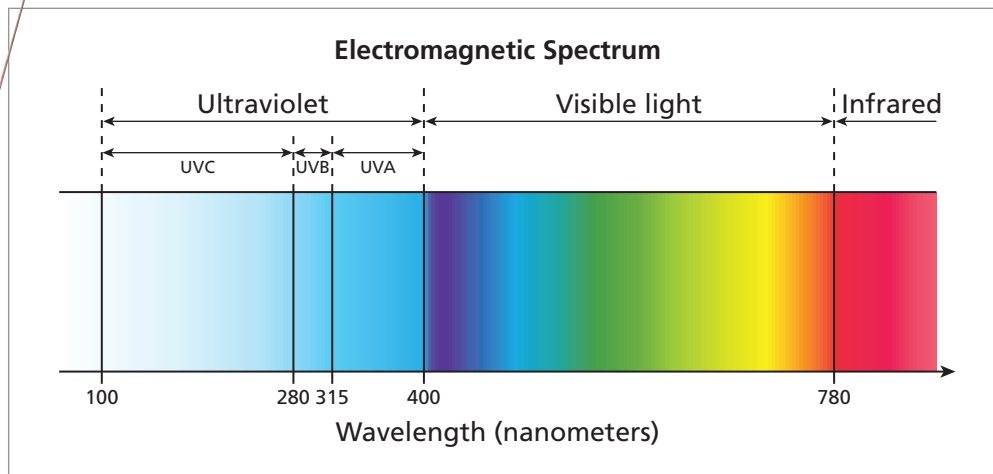
The American Hiking Society works "toward ensuring that hiking trails and natural places are cherished and preserved" for all generations.

| EXAMPLE 2 | **Analyzing UV Radiation** |

Sunlight consists of visible and invisible light. Ultraviolet (UV) light is invisible and is classified according to its wavelength, measured in nanometers (one-billionth of a meter). UV radiation is dangerous. It causes premature aging of the skin and can also cause various forms of skin cancer.

**Electromagnetic Spectrum**

| Ultraviolet | Visible light | Infrared |

UVC | UVB | UVA

100    280 315    400    780

**Wavelength (nanometers)**

UV radiation increases with elevation. It increases about 5% for every 1000 feet. Compare the UV radiation at the following elevations.

- Seattle, Washington (0 feet)
- Denver, Colorado (5000 feet)
- Mount McKinley, Alaska (20,000 feet)

## SOLUTION

- For the sake of comparison, assume that the amount of UV radiation in Seattle is 1.

- Because Denver has an elevation of 5000 feet, the amount of UV radiation in Denver is

$$(1.05)^5 \approx 1.276 \qquad \text{5000 feet elevation}$$

  or about 28% more than the UV radiation in Seattle.

- The peak of Mount McKinley has an elevation of 20,000 feet. The amount of UV radiation near the peak is

$$(1.05)^{20} \approx 2.653 \qquad \text{20,000 feet elevation}$$

  or about 165% more than the UV radiation in Seattle.

Mountain climbers need special protection for their skin and eyes. They should wear goggles and sunscreen that block both UVA and UVB rays.

✓ **Checkpoint**

Help at *Math*.and**Y☺U**.com

Compare the UV radiation at the following elevations.

- Reno, Nevada (4000 feet)
- Mount Whitney, California (14,000 feet)
- Mount Everest, Nepal (29,000 feet)

## Kayaking and Sailing

A *kayak* is a boat in which a paddler faces forward, legs in front, using a double-bladed paddle. A *canoe*, on the other hand, is a boat in which a paddler faces forward and sits or kneels in the boat, using a single-bladed paddle.

### EXAMPLE 3   Analyzing a Graph

The graph shows the discharge (amount of water) flowing through Rock Creek in Oregon. You want to kayak in Rock Creek only when the discharge is 50 cubic feet per second or greater. Estimate the percent of days from January 26 through May 26 that the creek meets your criteria.



**Rock Creek near Mill City, OR**

From a blog by a kayaker: "The East Fork of Rock Creek (aka *The Rock Creek* or simply 'The Rock') is a full-on, bare-knuckled brawler and one of my personal favorites. This creek has an outlandish gradient created by a non-stop series of ledges, falls, and boulder gardens that thunder down through Rock Creek Canyon."
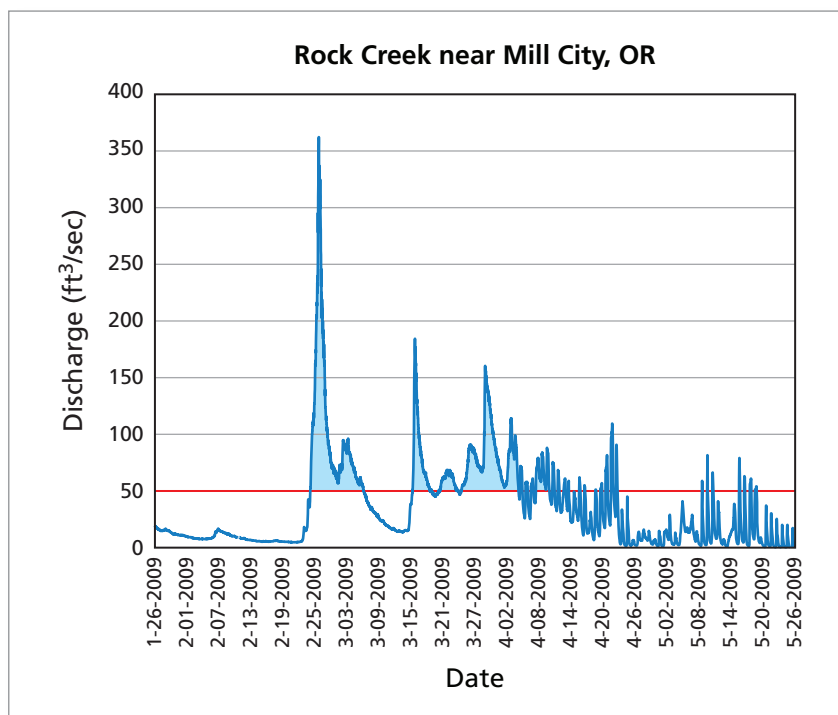
### SOLUTION

There are only 2 time periods when the creek has a discharge that is comfortably above 50 cubic feet per second. The first is about 11 days from February 23 through March 5. The second is about 23 days between March 14 and April 8. So, the percent of the days from January 26 through May 26 (121 days) that the creek water is high enough is

$$\frac{\text{days above 50 ft}^3/\text{sec}}{\text{total days}} = \frac{34}{121} \approx 0.281.$$

The creek meets your criteria about 28% of the time.

✓ **Checkpoint**                                    Help at *Math*.and**Y**☺**U**.com

You want to kayak in Rock Creek only when the discharge is 150 cubic feet per second or greater. Estimate the percent of days from January 26 through May 26 that the creek meets your criteria.

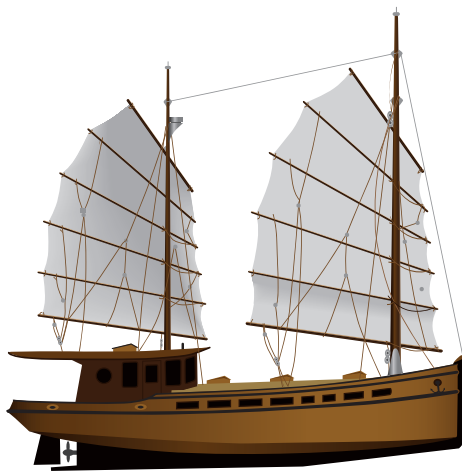EXAMPLE 4   **Analyzing Sail Designs**

The sail design of a modern yacht is quite different from the sail design used during most of the past 5000 years. Why is that?

Ancient Egyptian Design          Viking Design               Classic Chinese Design          Modern Yacht

### SOLUTION

The sail design of a modern yacht is different from the sail design used during most of the past 5000 years because of advancements in sail design. The following is a summary of some of the advancements.

> To sail against the wind, a boat can sail at an angle to the wind and zigzag. This is called "tacking." Early sailboats with square sails were not very effective when sailing against the wind. They were most effective when sailing with the wind. The development of triangular sails enabled boats to sail against the wind more effectively. This was because the boats were able to tack at a smaller angle with the wind than with square sails.
>
> When a triangular sail is positioned correctly, the air flow creates a pressure differential. This differential generates a force called the lift, which pulls a ship forward. A modern yacht like the one shown above has two sails—the mainsail and the jib. This design increases the pressure differential and generates more lift.

✓ **Checkpoint**                                    Help at *Math*.and**Y**☺**U**.com

Buoyancy

Center of gravity

Center of buoyancy

Weight
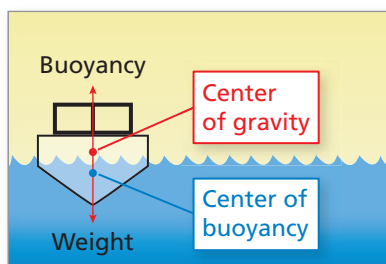
In an answer to the question, "Why doesn't a sailboat tip over?," the following explanation is given. "It's a balancing act between the boat's center of gravity and its center of buoyancy." Explain what this means.

## Bicycling and Cross-Country Skiing

Bicycling is one of the most popular outdoor sports in the United States. People not only ride bicycles for exercise and for pleasure, but many commute to work by bicycle.

---

**EXAMPLE 5** **Describing a Bicycle Trip**

The following graph was created by a cyclist using two devices: one that measured his heart rate and another that measured the elevation. Describe his cycling trip.



**Bicycle Trip**
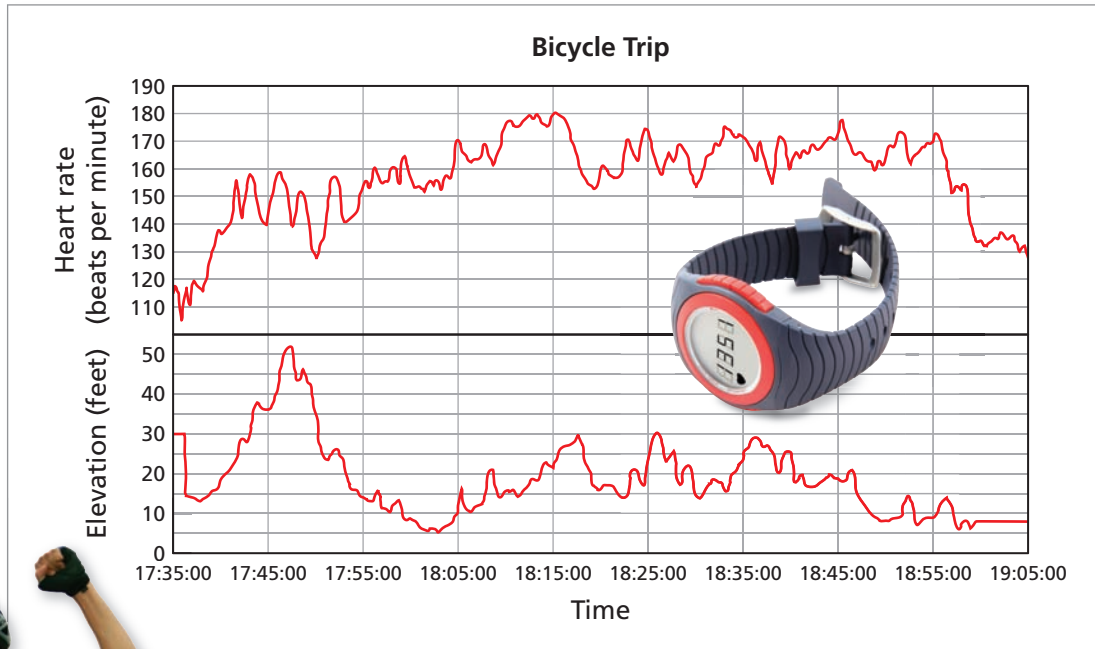
### SOLUTION

- The trip lasted from 17:35 (5:35 P.M.) to 19:05 (7:05 P.M.), which is 90 minutes.
- The trip did not start and stop at the same elevation, so the person did not make a round trip—beginning and ending at the same location.
- The elevation is low, at times only 5 feet above sea level. So, the person must have been riding near the ocean. On the other hand, there was 1 point on the trip where the elevation was 50 feet above sea level.
- The person's heart rate was elevated during most of the trip. In fact, during much of the trip, his heart rate was above 160 beats per minute, which is the maximum target heart rate for cardio exercise for a 20-year-old (see page 454).

✓ **Checkpoint** Help at *Math*.and**Y☺U**.com

In a survey, bicyclists in the United States were asked why they ride bicycles. Graphically represent the results.

| | | | |
|---|---|---|---|
| Recreation | 26.0% | Visit friend/relative | 10.1% |
| Exercise | 23.6% | Go on a bicycle ride | 2.3% |
| Commute to school/work | 19.2% | Other | 4.9% |
| Personal errand | 13.9% | | |