Statistics: Using Scatter Plots

GET READY for the Lesson

The table shows the number of Calories burned per hour by a 140-pound person running at various speeds. A linear function can be used to model these data.

Speed (mph)	Calories
5	508
6	636
7	731
8	858

Scatter Plots Data with two variables, such as speed and Calories, is called **bivariate data**. A set of bivariate data graphed as ordered pairs in a coordinate plane is called a **scatter plot**. A scatter plot can show whether there is a **positive**, **negative**, or **no correlation** between the data.

KEY CONCEPT   *Scatter Plots*

Positive Correlation   Negative Correlation   No Correlation
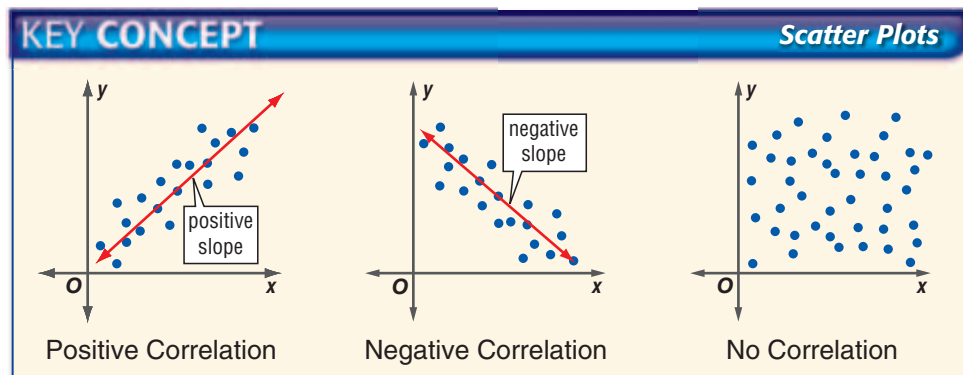
The more closely data can be approximated by a line, the stronger the correlation. Correlations are usually described as *strong* or *weak*.

Prediction Equations When you find a line that closely approximates a set of data, you are finding a **line of fit** for the data. An equation of such a line is often called a **prediction equation** because it can be used to predict one of the variables given the other variable.

To find a line of fit and a prediction equation for a set of data, select two points that appear to represent the data well. This is a matter of personal judgment, so your line and prediction equation may be different from someone else's.

**Real-World EXAMPLE**   **Find and Use a Prediction Equation**

HOUSING The table below shows the median selling price of new, privately-owned, one-family houses for some recent years.

Year	1994	1996	1998	2000	2002	2004
Price ($1000)	130.0	140.0	152.5	169.0	187.6	219.6

Source: U.S. Census Bureau and U.S. Department of Housing and Urban Development

a. Draw a scatter plot and a line of fit for the data. How well does the line fit the data?

Graph the data as ordered pairs, with the number of years since 1994 on the horizontal axis and the price on the vertical axis. The points (2, 140.0) and (8, 187.6) appear to represent the data well. Draw a line through these two points. Except for (10, 219.6), this line fits the data very well.
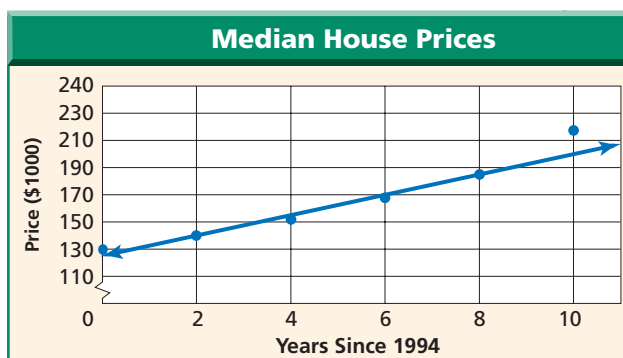
Median House Prices

Price ($1000) vs. Years Since 1994

b. Find a prediction equation. What do the slope and *y*-intercept indicate?

Find an equation of the line through (2, 140.0) and (8, 187.6).

$$m = \frac{y_2 - y_1}{x_2 - x_1} \quad \text{Slope formula}$$

$$= \frac{187.6 - 140.0}{8 - 2} \quad \text{Substitute.}$$

$$\approx 7.93 \quad \text{Simplify.}$$

$$y - y_1 = m(x - x_1) \quad \text{Point-slope form}$$

$$y - 140.0 = 7.93(x - 2) \quad \text{Substitute.}$$

$$y - 140.0 = 7.93x - 15.86 \quad \text{Distribute.}$$

$$y = 7.93x + 124.14 \quad \text{Simplify.}$$

One prediction equation is $y = 7.93x + 124.14$. The slope indicates that the median price is increasing at a rate of about $7930 per year. The *y*-intercept indicates that, according to the trend of the rest of the data, the median price in 1994 should have been about $124,140.

c. Predict the median price in 2014.

The year 2014 is 20 years after 1994, so use the prediction equation to find the value of *y* when *x* = 20.

$$y = 7.93x + 124.14 \quad \text{Prediction equation}$$

$$= 7.93(20) + 124.14 \quad x = 20$$

$$= 282.74 \quad \text{Simplify.}$$

The model predicts that the median price in 2014 will be about $282,740.

John Evans

d. How accurate does the prediction appear to be?

Except for the outlier, the line fits the data very well, so the predicted value should be fairly accurate.

1. The table shows the mean selling price of new, privately owned one-family homes for some recent years. Draw a scatter plot and a line of fit for the data. Then find a prediction equation and predict the mean price in 2014.

Year	1994	1996	1998	2000	2002	2004
Price ($1000)	154.5	166.4	181.9	207.0	228.7	273.5

Source: U.S. Census Bureau and U.S. Department of Housing and Urban Development

Online Personal Tutor at **algebra2.com**

ALGEBRA LAB

Head versus Height

COLLECT AND ORGANIZE THE DATA

Collect data from several of your classmates. Measure the circumference of each person's head and his or her height. Record the data as ordered pairs of the form (height, circumference).

ANALYZE THE DATA

1. Graph the data in a scatter plot and write a prediction equation.
2. Explain the meaning of the slope in the prediction equation.
3. Predict the head circumference of a person who is 66 inches tall.
4. Predict the height of an individual whose head circumference is 18 inches.

Study Tip

Outliers

If your scatter plot includes points that are far from the others on the graph, check your data before deciding that the point is an outlier. You may have made a graphing or recording mistake.

Example
(p. 87)

Complete parts a–c for each set of data in Exercises 1 and 2.
a. Draw a scatter plot and a line of fit, and describe the correlation.
b. Use two ordered pairs to write a prediction equation.
c. Use your prediction equation to predict the missing value.

1. **SCIENCE** The table shows the temperature in the atmosphere at various altitudes.

Altitude (ft)	0	1000	2000	3000	4000	5000
Temp (°C)	15.0	13.0	11.0	9.1	7.1	?

Source: NASA

2. **TELEVISION** The table shows the percentage of U.S. households with televisions that also had cable service in some recent years.

Year	1995	1997	1999	2001	2003	2015
Percent	65.7	67.3	68.0	69.2	68.0	?

Source: Nielsen Media Research

Complete parts a-c for each set of data in Exercises 3–6.

a. Draw a scatter plot and a line of fit, and describe the correlation.

b. Use two ordered pairs to write a prediction equation.

c. Use your prediction equation to predict the missing value.

3. SAFETY All states and the District of Columbia have enacted laws setting 21 as the minimum drinking age. The table shows the estimated cumulative number of lives these laws have saved by reducing traffic fatalities.

Year	1999	2000	2001	2002	2003	2015
Lives (1000s)	19.1	20.0	21.0	21.9	22.8	?

Source: National Highway Traffic Safety Administration

4. HOCKEY The table shows the number of goals and assists for some of the members of the Detroit Red Wings in a recent NHL season.

Goals	30	25	18	14	15	14	10	6	4	30	?
Assists	49	43	33	32	28	29	12	9	15	38	20

Source: www.detroitredwings.com

5. HEALTH The table shows the number of gallons of bottled water consumed per person in some recent years.

Year	1998	1999	2000	2001	2002	2003	2015
Gallons	15.0	16.4	17.4	18.8	20.7	22.0	?

**Source:** U.S. Department of Agriculture

6. THEATER The table shows the total revenue of all Broadway plays for recent seasons.

Season	1999–2000	2000–2001	2001–2002	2002–2003	2003–2004	2013–2014
Revenue ($ millions)	603	666	643	721	771	?

Source: The League of American Theatres and Producers, Inc.

MEDICINE For Exercises 7–9, use the graph that shows how much Americans spent on health care in some recent years and a prediction for how much they will spend in 2014.

7. Write a prediction equation from the data for 1999 to 2003.

8. Use your equation to predict the amount for 2014.

9. Compare your prediction to the one given in the graph.
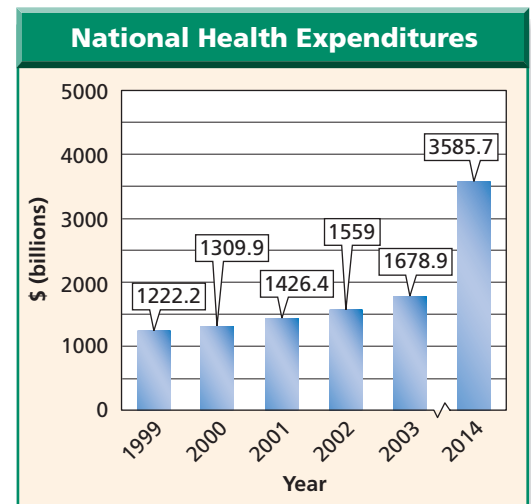
National Health Expenditures

Source: cms.hhs.gov

FINANCE For Exercises 10 and 11, use the following information.
Della has $1000 that she wants to invest in the stock market. She is considering buying stock in either Company 1 or Company 2. The values of the stocks at the end of each of the last 4 months are shown in the tables below.

10. Based only on these data, which stock should Della buy? Explain.

11. Do you think investment decisions should be based on this type of reasoning? If not, what other factors should be considered?

Company 1	
Month	Share Price ($)
Aug.	25.13
Sept.	22.94
Oct.	24.19
Nov.	22.56

Company 2	
Month	Share Price ($)
Aug.	31.25
Sept.	32.38
Oct.	32.06
Nov.	32.44

PLANETS For Exercises 12–15, use the table below that shows the average distance from the Sun and average temperature for eight of the planets.
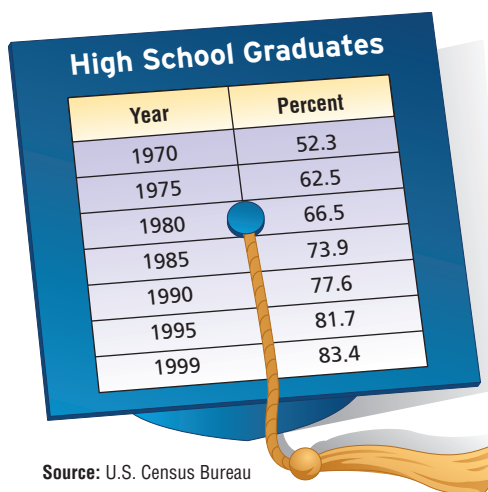
Planet	Average Distance from the Sun (million miles)	Average Temperature (°F)
Mercury	36	333
Venus	67.2	867
Earth	93	59
Mars	141.6	−85
Jupiter	483.8	−166
Saturn	890.8	−200
Uranus	1784.8	−320
Pluto	3647.2	−375

Source: World Meteorological Association

12. Draw a scatter plot with average distance as the independent variable.

13. Write a prediction equation.

14. Predict the average temperature for Neptune, which has an average distance from the Sun of 2793.1 million miles.

15. Compare your prediction to the actual value of −330°F.

16. **RESEARCH** Use the Internet or other resource to look up the population of your community in several past years. Organize the data as ordered pairs. Then use an equation to predict the population in some future year.

H.O.T. Problems

CHALLENGE For Exercises 17 and 18, use the table that shows the percent of people ages 25 and over with a high school diploma over the last few decades.

17. Use a prediction equation to predict the percent in 2015.

18. Do you think your prediction is accurate? Explain.

High School Graduates

Year	Percent
1970	52.3
1975	62.5
1980	66.5
1985	73.9
1990	77.6
1995	81.7
1999	83.4

Source: U.S. Census Bureau

19. OPEN ENDED Write a different prediction equation for the data in the example on page 87.

20. *Writing in Math* Use the information on page 86 to explain how a linear equation can model the number of Calories you burn while exercising. Include a scatter plot, a description of the correlation, and a prediction equation for the data. Then predict the number of Calories burned in an hour by a 140-pound person running at 9 miles per hour and compare your predicted value with the actual value of 953.

STANDARDIZED TEST PRACTICE

21. ACT/SAT Which line best fits the data in the graph?

A $y = x$
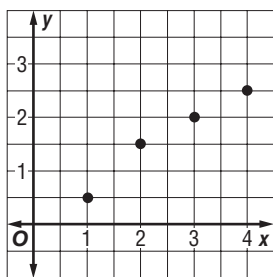
B $y = -0.5x + 4$

C $y = -0.5x - 4$

D $y = 0.5 + 0.5x$

22. REVIEW Anna took brownies to a club meeting. She gave half of her brownies to Sarah. Sarah gave a third of her brownies to Rob. Rob gave a fourth of his brownies to Trina. If Trina has 3 brownies, how many brownies did Anna have in the beginning?

F 12

G 36

H 72

J 144

Spiral Review

Write an equation in slope-intercept form that satisfies each set of conditions. (Lesson 2-4)

23. slope 4, passes through $(0, 6)$ **24.** passes through $(5, -3)$ and $(-2, 0)$

TELEPHONES For Exercises 25 and 26, use the following information. (Lesson 2-4)
Namid is examining the calling card portion of his phone bill. A 4-minute call at the night rate cost $2.65. A 10-minute call at the night rate cost $4.75.

25. Write a linear equation to model this situation.

26. How much would it cost to talk for half an hour at the night rate?

Find the slope of the line that passes through each pair of points. (Lesson 2-3)

27. $(5, 4), (-3, 8)$ **28.** $(-1, -2), (4, -2)$ **29.** $(3, -4), (3, 16)$

30. PROFIT Kara is planning to set up a booth at a local festival to sell her paintings. She determines that the amount of profit she will make is determined by the function $P(x) = 11x - 100$, where x is the number of paintings she sells. How much profit will Kara make if she sells 35 of her paintings? (Lesson 2-1)

GET READY for the Next Lesson

PREREQUISITE SKILL Find each absolute value. (Lesson 1-4)

31. $|-3|$ **32.** $|11|$ **33.** $|0|$ **34.** $\left|-\frac{2}{3}\right|$ **35.** $|-1.5|$