4.7 Predict with Linear Models

Before You made scatter plots and wrote equations of lines of fit.

Now You will make predictions using best-fitting lines.

Why? So you can model trends, as in Ex. 21.

Key Vocabulary
- **best-fitting line**
- **linear regression**
- **interpolation**
- **extrapolation**
- **zero of a function**

The line that most closely follows a trend in data is called the **best-fitting line**. The process of finding the best-fitting line to model a set of data is called **linear regression**. You can perform linear regression using technology. Using a line or its equation to approximate a value between two known values is called **linear interpolation**.

EXAMPLE 1 Interpolate using an equation

COMMON CORE

CC.9-12.S.ID.6a Fit a function to the data; use functions fitted to data to solve problems in the context of the data.*

CD SINGLES The table shows the total number of CD singles shipped (in millions) by manufacturers for several years during the period 1993–1997.

Year	1993	1995	1996	1997
CD singles shipped (millions)	7.8	22	43	67

REVIEW REGRESSION
You may want to review performing a linear regression to find the best-fitting line.

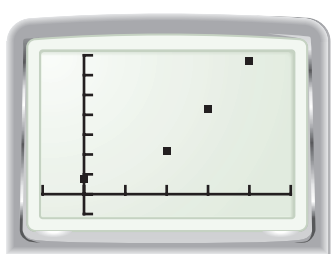
a. Make a scatter plot of the data.

b. Find an equation that models the number of CD singles shipped (in millions) as a function of the number of years since 1993.

c. Approximate the number of CD singles shipped in 1994.

Solution

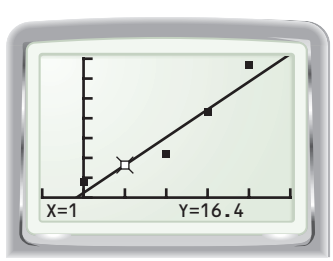
a. Enter the data into lists on a graphing calculator. Make a scatter plot, letting the number of years since 1993 be the x-values (0, 2, 3, 4) and the number of CD singles shipped be the y-values.



b. Perform linear regression using the paired data. The equation of the best-fitting line is approximately $y = 14x + 2.4$.

ANOTHER WAY
You can also estimate the number of CDs shipped in 1994 by evaluating $y = 14x + 2.4$ when $x = 1$.

c. Graph the best-fitting line. Use the *trace* feature and the arrow keys to find the value of the equation when $x = 1$.



▶ About 16 million CD singles were shipped in 1994.

Animated Algebra my.hrw.com

both, © Nick Ut/AP/Wide World Photos

EXTRAPOLATION Using a line or its equation to approximate a value outside the range of known values is called **linear extrapolation**.

EXAMPLE 2 Extrapolate using an equation

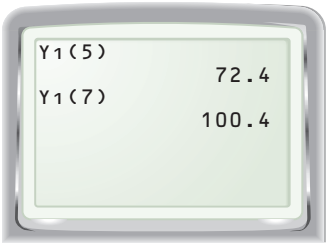
CD SINGLES Look back at Example 1.

a. Use the equation from Example 1 to approximate the number of CD singles shipped in 1998 and in 2000.

b. In 1998 there were actually 56 million CD singles shipped. In 2000 there were actually 34 million CD singles shipped. *Describe* the accuracy of the extrapolations made in part (a).

Solution

a. Evaluate the equation of the best-fitting line from Example 1 for $x = 5$ and $x = 7$.



The model predicts about 72 million CD singles shipped in 1998 and about 100 million CD singles shipped in 2000.

b. The differences between the predicted number of CD singles shipped and the actual number of CD singles shipped in 1998 and 2000 are 16 million CDs and 66 million CDs, respectively. The difference in the actual and predicted numbers increased from 1998 to 2000. So, the equation of the best-fitting line gives a less accurate prediction for the year that is farther from the given years.

ACCURACY As Example 2 illustrates, the farther removed an x-value is from the known x-values, the less confidence you can have in the accuracy of the predicted y-value. This is true in general but not in every case.

✓ GUIDED PRACTICE for Examples 1 and 2

1. **HOUSE SIZE** The table shows the median floor area of new single-family houses in the United States during the period 1995–1999.

Year	1995	1996	1997	1998	1999
Median floor area (square feet)	1920	1950	1975	2000	2028

a. Find an equation that models the floor area (in square feet) of a new single-family house as a function of the number of years since 1995.

b. Predict the median floor area of a new single-family house in 2000 and in 2001.

c. Which of the predictions from part (b) would you expect to be more accurate? *Explain* your reasoning.

EXAMPLE 3 Predict using an equation

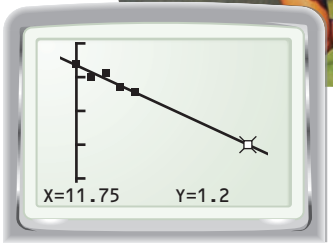
SOFTBALL The table shows the number of participants in U.S. youth softball during the period 1997–2001. Predict the year in which the number of youth softball participants reaches 1.2 million.

Year	1997	1998	1999	2000	2001
Participants (millions)	1.44	1.4	1.411	1.37	1.355

Solution

STEP 1 Perform linear regression. Let x represent the number of years since 1997, and let y represent the number of youth softball participants (in millions). The equation for the best-fitting line is approximately $y = -0.02x + 1.435$.

STEP 2 Graph the equation of the best-fitting line. Trace the line until the cursor reaches $y = 1.2$. The corresponding x-value is shown at the bottom of the calculator screen.



▸ There will be 1.2 million participants about 12 years after 1997, or in 2009.

ANOTHER WAY

You can also predict the year by substituting 1.2 for y in the equation and solving for x:

$y = -0.02x + 1.435$

$1.2 = -0.02x + 1.435$

$x = 11.75$

GUIDED PRACTICE for Example 3

2. **SOFTBALL** In Example 3, in what year will there be 1.25 million youth softball participants in the U.S?

ZERO OF A FUNCTION A **zero of a function** $y = f(x)$ is an x-value for which $f(x) = 0$ (or $y = 0$). Because $y = 0$ along the x-axis of the coordinate plane, a zero of a function is an x-intercept of the function's graph.

KEY CONCEPT *For Your Notebook*

Relating Solutions of Equations, x-Intercepts of Graphs, and Zeros of Functions

You have learned to solve an equation like $2x - 4 = 0$:

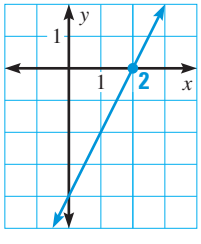
$2x - 4 = 0$

$2x = 4$

$x = 2$

The solution of $2x - 4 = 0$ is 2.

You have also found the x-intercept of the graph of a function like $y = 2x - 4$:



Now you are finding the zero of a function like $f(x) = 2x - 4$:

$f(x) = 0$

$2x - 4 = 0$

$x = 2$

The zero of $f(x) = 2x - 4$ is 2.

EXAMPLE 4 Find the zero of a function

SOFTBALL Look back at Example 3. Find the zero of the function. *Explain* what the zero means in this situation.

Solution

Substitute 0 for y in the equation of the best-fitting line and solve for x.

$y = -0.02x + 1.435$ **Write the equation.**

$0 = -0.02x + 1.435$ **Substitute 0 for *y*.**

$x \approx 72$ **Solve for *x*.**

▶ The zero of the function is about 72. The function has a negative slope, which means that the number of youth softball participants is decreasing. According to the model, there will be no youth softball participants 72 years after 1997, or in 2069.

✓ GUIDED PRACTICE for Example 4

3. **JET BOATS** The number y (in thousands) of jet boats purchased in the U.S. can be modeled by the function $y = -1.23x + 14$ where x is the number of years since 1995. Find the zero of the function. *Explain* what the zero means in this situation.

4.7 EXERCISES

HOMEWORK KEY

- ◯ = See **WORKED-OUT SOLUTIONS** Exs. 3 and 19
- ★ = **STANDARDIZED TEST PRACTICE** Exs. 2, 14, 16, and 21
- ◆ = **MULTIPLE REPRESENTATIONS** Exs. 22

SKILL PRACTICE

1. **VOCABULARY** Copy and complete: Using a linear function to approximate a value within a range of known data values is called __?__.

2. ★ **WRITING** *Explain* how extrapolation differs from interpolation.

EXAMPLE 1 for Exs. 3–4

LINEAR INTERPOLATION **Make a scatter plot of the data. Find the equation of the best-fitting line. Approximate the value of y for $x = 5$.**

3.

x	0	2	4	6	7
y	2	7	14	17	20

4.

x	2	4	6	8	10
y	6.2	22.5	40.2	55.4	72.1

EXAMPLE 2 for Exs. 5–6

LINEAR EXTRAPOLATION **Make a scatter plot of the data. Find the equation of the best-fitting line. Approximate the value of y for $x = 10$.**

5.

x	0	1	2	3	4
y	20	32	39	53	63

6.

x	1	3	5	7	9
y	0.4	1.4	1.9	2.3	3.2

EXAMPLE 4 for Exs. 7–13

ZERO OF A FUNCTION **Find the zero of the function.**

7. $f(x) = 7.5x - 20$

8. $f(x) = -x + 7$

9. $f(x) = \frac{1}{8}x + 2$

10. $f(x) = 17x - 68$

11. $f(x) = -0.5x + 0.75$

12. $f(x) = 5x - 7$

13. **ERROR ANALYSIS** *Describe* and correct the error made in finding the zero of the function $y = 2.3x - 2$.

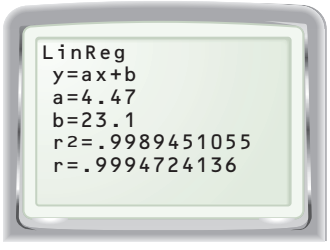
$y = 2.3(0) - 2$
$y = -2$

14. ★ **MULTIPLE CHOICE** Given the function $y = 12.6x + 3$, for what x-value does $y = 66$?

Ⓐ 0.2 Ⓑ 5 Ⓒ 5.5 Ⓓ 78.6

15. **ERROR ANALYSIS** *Describe* and correct the error in finding an equation of the best-fitting line using a graphing calculator.

Equation of the best-fitting line is $y = 23.1x + 4.47$.



16. ★ **OPEN-ENDED** Give an example of a real-life situation in which you can use linear interpolation to find the zero of a function. *Explain* what the zero means in this situation.

17. **CHALLENGE** A quantity increases rapidly for 10 years. During the next 10 years, the quantity decreases rapidly.

a. Can you fit a line to the data? *Explain*.

b. How could you model the data using more than one line? *Explain* the steps you could take.

PROBLEM SOLVING

EXAMPLE 1 for Ex. 18

18. **SAILBOATS** Your school's sailing club wants to buy a sailboat. The table shows the lengths and costs of sailboats.

Length (feet)	11	12	14	14	16	22	23
Cost (dollars)	600	500	1900	1700	3500	6500	6000

a. Make a scatter plot of the data. Let x represent the length of the sailboat. Let y represent the cost of the sailboat.

b. Find an equation that models the cost (in dollars) of a sailboat as a function of its length (in feet).

c. Approximate the cost of a sailboat that is 20 feet long.

EXAMPLE 2 for Ex. 19

19. FARMING The table shows the living space recommended for pigs of certain weights.

Weight (pounds)	40	60	80	100	120	150	230
Area (square feet)	2.5	3	3.5	4	5	6	8

a. Make a scatter plot of the data.

b. Write an equation that models the recommended living space (in square feet) as a function of a pig's weight (in pounds).

c. About how much living space is recommended for a pig weighing 250 pounds?

EXAMPLE 3 for Ex. 20

20. TELEVISION STATIONS The table shows the number of UHF and VHF broadcast television stations each year from 1996 to 2002.

Year	1996	1997	1998	1999	2000	2001	2002
Television stations	1551	1563	1583	1616	1730	1686	1714

a. Find an equation that models the number of broadcast television stations as a function of the number of years since 1996.

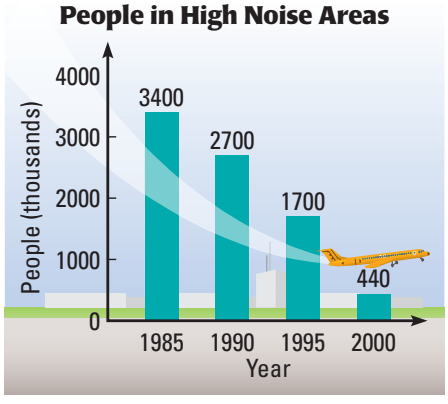
b. Approximate the year in which there were 1790 television stations.

EXAMPLE 4 for Exs. 21–22

21. ★ SHORT RESPONSE The table shows the number of people who lived in high noise areas near U.S. airports for several years during the period 1985–2000.

a. Find an equation that models the number of people (in thousands) living in high noise areas as a function of the number of years since 1985.

b. Find the zero of the function from part (a). *Explain* what the zero means in this situation. Is this reasonable?



22. MULTIPLE REPRESENTATIONS An Internet search for used cars of a given make, model, and year in your local area found cars with different mileages and different selling prices, as shown.

Mileage (thousands of miles)	22	14	18	30	8	24
Price (thousands of dollars)	16	17	17	14	18	15

a. Making a Graph Draw two scatter plots of the data, one by hand and one using a graphing calculator.

b. Writing an Equation Draw a line of fit on your hand-drawn scatter plot. Use the line to write an equation that models the selling price as a function of the mileage. Then use a graphing calculator to find the best-fitting line. *Compare* your models.

c. Describing in Words Identify the slope and y-intercept of the best-fitting line. *Explain* their meanings in the context of the situation.

○ = See WORKED-OUT SOLUTIONS in Student Resources ★ = STANDARDIZED TEST PRACTICE = MULTIPLE REPRESENTATIONS

23. CHALLENGE The table shows the estimated populations of mallard ducks and all ducks in North America for several years during the period 1975–2000.

Year	1975	1980	1985	1990	1995	2000
Mallards (thousands)	7727	7707	4961	5452	8269	9470
All ducks (thousands)	37,790	36,220	25,640	25,080	35,870	41,840

a. Make two scatter plots where x is the number of years since 1975 and y is the number of mallards (in thousands) for one scatter plot, while y is the number of ducks (in thousands) for the other scatter plot. *Describe* the correlation of the data in each scatter plot.

b. Can you use the mallard duck population to predict the total duck population? *Explain*.

QUIZ

1. PARALLEL LINES Write an equation of the line that passes through $(-6, 8)$ and is parallel to the line $y = 3x - 15$.

PERPENDICULAR LINES Write an equation of the line that passes through the given point and is perpendicular to the given line.

2. $(5, 5)$, $y = -x + 2$ **3.** $(10, -3)$, $y = 2x + 24$ **4.** $(2, 3)$, $x + 2y = -7$

5. CASSETTE TAPES The table shows the number of audio cassette tapes shipped for several years during the period 1994–2002.

Year	1994	1996	1998	2000	2002
Tapes shipped (millions)	345	225	159	76	31

a. Write an equation that models the number of tapes shipped (in millions) as a function of the number of years since 1994.

b. At about what rate did the number of tapes shipped change over time?

c. Approximate the year in which 125 million tapes were shipped.

d. Find the zero of the function from part (a). *Explain* what the zero means in this situation.

my.hrw.com
Keystrokes

Model Data from the Internet

MATHEMATICAL PRACTICES
Model with mathematics.

QUESTION How can you find reliable data on the Internet and use it to predict the total U.S voting-age population in 2010?

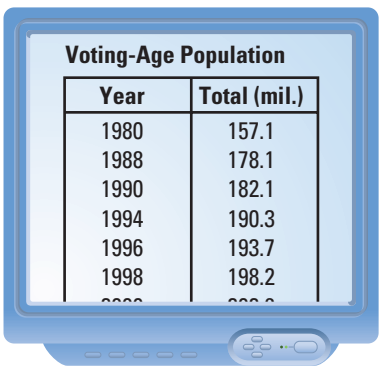
EXAMPLE 1 Collect and analyze data

Find data for the total U.S. voting-age population over several years. Use an equation that models the data to predict the total U.S. voting-age population in 2010.

STEP 1 Find a data source

Reliable data about the U.S. population can be found in the online *Statistical Abstract*. Go to the address shown below. Click on a link to the most recent version of the *Statistical Abstract*.

Address http://www.census.gov

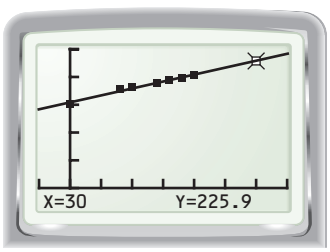


STEP 2 Find an appropriate data set

Choose the most recent "Elections" document. In this document, find the table of data entitled "Voting-Age Population."

STEP 3 Find a model

Use a graphing calculator to make a scatter plot. Let x represent the number of years since 1980. Let y represent the total U.S. voting-age population (in millions). Find an equation that models the total U.S. voting-age population (in millions) as a function of the number of years since 1980.



▸ $y = 2.23x + 159$

STEP 4 Predict

Use the model to predict the total voting-age population in 2010. You can either evaluate the equation for $x = 30$ or trace the graph of the equation, as shown.

▸ The total U.S. voting-age population will be about 225.9 million in 2010.

DRAW CONCLUSIONS

1. In the online *Statistical Abstract*, find data for the total value of agricultural imports over several years beginning with 1990.
2. Make a scatter plot of the data you found in Exercise 1. Find an equation that models the total value of agricultural imports (in millions of dollars) as a function of the number of years since 1990.
3. Predict the year in which the total value of agricultural imports will be $45,000 million. *Describe* the method you used.