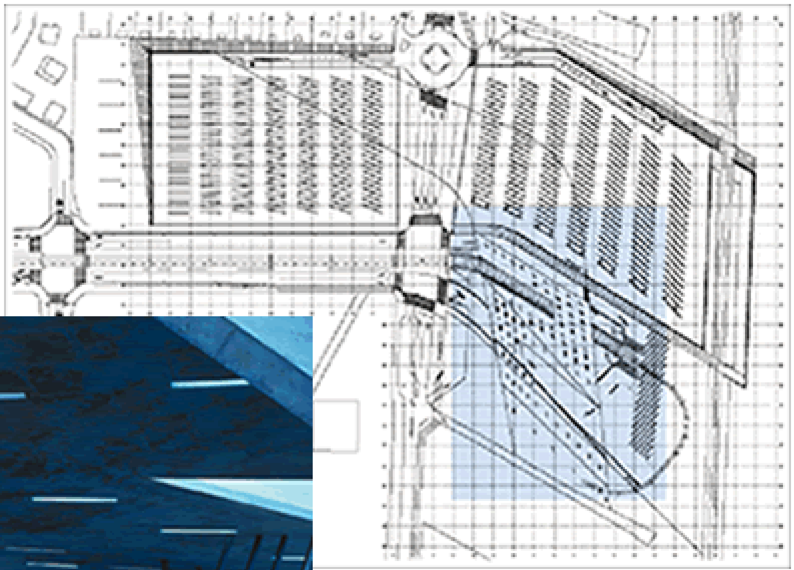
CHAPTER

3

Linear Models and Systems



Iraqi-British architect Zaha Hadid (b 1950) designed this commuter train station on the outskirts of Strasbourg, France. She worked with the concept of overlapping fields and lines to represent the interacting patterns of moving cars, trains, bicycles, and pedestrians. The roof of the waiting area in the station is punctured by angled lines that let sunlight shine on the floor, shifting throughout the day. You can see these lines in the photograph, which shows the region of the architectural plan highlighted in blue.

OBJECTIVES

In this chapter you will

- **review linear equations in intercept form and point-slope form**
- **explore connections between arithmetic sequences and linear equations**
- **find lines of fit for data sets that are approximately linear**
- **solve systems of linear equations**

LESSON 3.1

Linear Equations and Arithmetic Sequences

You can solve many rate problems by using recursion.

Matias wants to call his aunt in Chile on her birthday. He learned that placing the call costs \$2.27 and that each minute he talks costs \$1.37. How much would it cost to talk for 30 minutes?

Valparaiso, Chile

You can calculate the cost of Matias's phone call with the recursive formula

$u_0 = 2.27$

$u_n = u_{n-1} + 1.37 \quad \text{where } n \geq 1$

To find the cost of a 30-minute phone call, calculate the first 30 terms, as shown in the calculator screen.

As you learned in algebra, you or Matias can also find the cost of a 30-minute call by using the **linear equation**

$y = 2.27 + 1.37x$

where x is the length of the phone call in minutes and y is the cost in dollars. If the phone company always rounds up the length of the call to the nearest whole minute, then the costs become a sequence of discrete points, and you can write the relationship as an explicit formula,

$u_n = 2.27 + 1.37n$

where n is the length of the call in whole minutes and u_n is the cost in dollars. An **explicit formula** gives a direct relationship between two discrete quantities. How does the explicit formula differ from the recursive formula? How would you use each one for calculating the cost of a 15-minute call or an n-minute call?

In this lesson you will write and use explicit formulas for arithmetic sequences. You will also write linear equations for lines through the discrete points of arithmetic sequences.

EXAMPLE A

Consider the recursively defined arithmetic sequence

$u_0 = 2$

$u_n = u_{n-1} + 6 \quad \text{where } n \geq 1$

a. Find an explicit formula for the sequence.

b. Use the explicit formula to find u_{22}.

c. Find the value of n so that $u_n = 86$.

▶ Solution

a. Look for a pattern in the sequence.

u_0	2
u_1	$8 = 2 + 6 = 2 + 6 \cdot 1$
u_2	$14 = 2 + 6 + 6 = 2 + 6 \cdot 2$
u_3	$20 = 2 + 6 + 6 + 6 = 2 + 6 \cdot 3$

Notice that the common difference (or rate of change) between the terms is 6. You start with 2 and just keep adding 6. That means each term is equivalent to 2 plus 6 times the term number. In general, when you write the formula for a sequence, you use n to represent the number of the term and u_n to represent the term itself.

Term value = Initial value + Rate · Term number

$$u_n = 2 + 6 \cdot n$$

b. You can use the explicit formula to find u_{22} without calculating all of the previous terms. By substituting 22 for n, you get $u_{22} = 2 + 6 \cdot 22$. So, u_{22} equals 134.

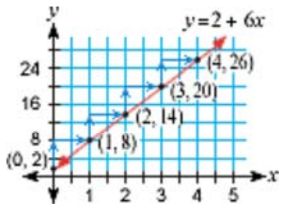
c. To find n so that $u_n = 86$, substitute 86 for u_n in the formula $u_n = 2 + 6n$.

$86 = 2 + 6n$	Substitute 86 for u_n.
$14 = n$	Solve for n.

So, 86 is the 14th term.

When you download photos from the Internet, sometimes the resolution improves as the percentage of download increases. The relationship between the percentage of download and the resolution could be modeled with a sequence, an explicit formula, or a linear equation.

You graphed sequences of points (n, u_n) in Chapter 1. The variable n stands for a term number, so it is a whole number: 0, 1, 2, 3, ... So, using different values for n will produce a set of discrete points. This graph shows the arithmetic sequence from Example A.



When n increases by 1, u_n increases by 6, the common difference. In terms of x and y, the number 6 is the change in the y-value, or function value, that corresponds to a unit change (a change of 1) in the x-value. So the points representing the sequence lie on a line with a slope of 6. In general, the common difference, or rate of change, between consecutive terms of an arithmetic sequence is the **slope** of the line through those points.

The pair (0, 2) names the starting value 2, which is the y-intercept. Using the intercept form of a linear equation, you can now write an equation of the line through the points of the sequence as $y = 2 + 6x$, or $y = 6x + 2$.

In this course you will use x and y to write linear equations. You will use n and u_n to write both recursive formulas and explicit formulas for sequences of discrete points.

In the investigation you will focus on this relationship between the formula for an arithmetic sequence and the equation of the line through the points representing the sequence.

Investigation
Match Point

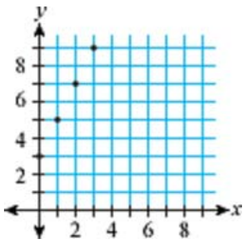
Below are three recursive formulas, three graphs, and three linear equations.

1. $u_0 = 4$
$u_n = u_{n-1} - 1$ where $n \geq 1$

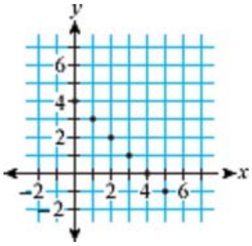
2. $u_0 = 2$
$u_n = u_{n-1} + 5$ where $n \geq 1$

3. $u_0 = -4$
$u_n = u_{n-1} + 3$ where $n \geq 1$

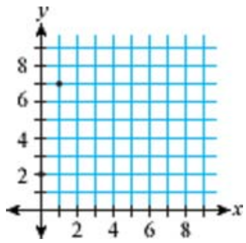
A.



B.



C.



i. $y = -4 + 3x$

ii. $y = 4 + x$

iii. $y = 2 + 5x$

Step 1 Match the recursive formulas, graphs, and linear equations that go together. (Not all of the appropriate matches are listed. If the recursive rule, graph, or equation is missing, you will need to create it.)

Step 2 Write a brief statement relating the starting value and common difference of an arithmetic sequence to the corresponding equation $y = a + bx$.

Step 3 Are points (n, u_n) of an arithmetic sequence always collinear? Write a brief statement supporting your answer.

EXAMPLE B

Retta typically spends \$2 a day on lunch. She notices that she has \$17 left after today's lunch. She thinks of this sequence to model her daily cash balance.

$u_1 = 17$

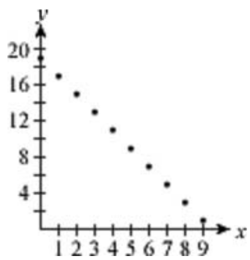
$u_n = u_{n-1} - 2$ where $n > 1$

a. Find the explicit formula that represents her daily cash balance and the equation of the line through the points of this sequence.

b. How useful is this formula for predicting how much money Retta will have each day?

▸ Solution

Use the common difference and starting term to write the explicit formula.



a. Each term is 2 less than the previous term, so the common difference of the arithmetic sequence and the slope of the line are both –2. The term u_1 is 17, so the previous term, u_0, or the y-intercept, is 19.

The explicit formula for the arithmetic sequence is $u_n = 19 - 2 \cdot n$, and the equation of the line containing these points is $y = 19 - 2x$.

b. We don't know whether Retta has any other expenses, when she receives her paycheck or allowance, or whether she buys lunch on the weekend. The formula could be valid for eight more days, until she has $1 left (on u_9), as long as she gets no more money and spends only $2 per day.

For both sequences and equations, it is important to consider the conditions for which the relationship is valid. For example, a phone company usually rounds up the length of a phone call to determine the charges, so the relationship between the length of the call and the cost of the call is valid only for a call length that is a positive integer. Also, the portion of the line left of the y-axis, where x is negative, is part of the mathematical model but has no relevance for the phone call scenario.

EXERCISES

You will need

Statistics software for Exercise **16**

Practice Your Skills

1. Consider the sequence

$u_0 = 18$

$u_n = u_{n-1} - 3 \quad$ where $n \geq 1$

a. Graph the sequence.

b. What is the slope of the line that contains the points? How is that related to the common difference of the sequence?

c. What is the y-intercept of the line that contains these points? How is it related to the sequence?

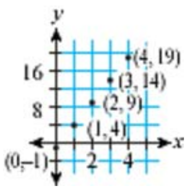
d. Write the equation of the line that contains these points.

2. Refer to the graph at right.

a. Write a recursive formula for the sequence. What is the common difference? What is the value of u_0?

b. What is the slope of the line through the points? What is the y-intercept?

c. Write the equation of the line that contains these points.



3. Write the equation of the line that passes through the points of an arithmetic sequence with $u_0 = 7$ and a common difference of 3.

4. Write a recursive formula for a sequence whose points lie on the line $y = 6 - 0.5x$.

Keymath.com More Practice Your Skills

5. Find the slope of each line.

a. $y = 2 + 1.7x$ **b.** $y = x + 5$ **c.** $y = 12 - 4.5x$ **d.** $y = 12$

Reason and Apply

6. An arithmetic sequence has a starting term, u_0, of 6.3 and a common difference of 2.5.

a. Write an explicit formula for the sequence.

b. Use the formula to figure out which term is 78.8.

7. Suppose you drive through Macon, Georgia, (which is 82 mi from Atlanta) on your way to Savannah, Georgia, at a steady 54 mi/h.

a. What is your distance from Atlanta two hours after you leave Macon?

b. Write an equation that represents your distance, y, from Atlanta x hours after leaving Macon.

c. Graph the equation.

d. Does this equation model an arithmetic sequence? Why or why not?

Savannah, Georgia

8. **APPLICATION** Melissa and Roy both sell cars at the same dealership and have to meet the same profit goal each week. Last week, Roy sold only three cars, and he was below his goal by \$2050. Melissa sold seven cars, and she beat her goal by \$1550. Assume that the profit is approximately the same for each car they sell.

a. Use a graph to find a few terms, u_n, of the sequence of sales amounts above or below the goal.

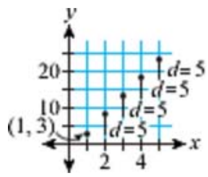
b. What is the real-world meaning of the common difference?

c. Write an explicit formula for this sequence of sales values in relation to the goal. Define variables and write a linear equation.

d. What is the real-world meaning of the horizontal and vertical intercepts?

e. What is the profit goal? How many more cars must Roy sell to be within \$500 of his goal?

9. The points on this graph represent the first five terms of an arithmetic sequence. The height of each point is its distance from the x-axis, or the value of the y-coordinate of the point.



a. Find u_0, the y-coordinate of the point preceding those given.

b. How many common differences (d's) do you need to get from the height of $(0, u_0)$ to the height of $(5, u_5)$?

c. How many d's do you need to get from the height of $(0, u_0)$ to the height of $(50, u_{50})$?

d. Explain why you can find the height from the x-axis to $(50, u_{50})$ using the equation $u_{50} = u_0 + 50d$.

e. In general, for an arithmetic sequence, the explicit formula is $u_n =$ _?_.

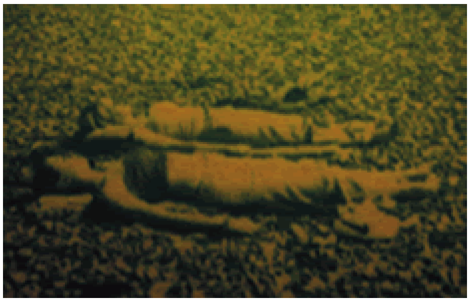
10. A gardener planted a new variety of ornamental grass and kept a record of its height over the first two weeks of growth.

Time (days)	0	3	7	10	14
Height (cm)	4.2	6.3	9.1	11.2	14

a. How much does the grass grow each day?

b. Write an explicit formula that gives the height of the grass after n days.

c. How long will it take for the grass to be 28 cm tall?



Heather Ackroyd and Dan Harvey print photographs on grass, such as this one, *Sunbathers 2000.* They position a photographic negative over growing grass, and the chlorophyll reacts to give a temporary image in shades of green and yellow. They use genetically modified grass to make their images last longer.

11. An arithmetic sequence of six numbers begins with 7 and ends with 27. Follow 11a-c to find the four missing terms.

a. Name two points on the graph of this sequence: (__?__, 7) and (__?__, 27).

b. Plot the two points you named in 11a and find the slope of the line connecting the points.

c. Use the slope to find the missing terms.

d. Plot all the points and write the equation of the line that contains them.

12. **APPLICATION** If an object is dropped, it will fall a distance of about 16 feet during the first second. In each second that follows, the object falls about 32 feet farther than in the previous second.

a. Write a recursive formula for the distance fallen each second under free fall.

b. Find an explicit formula for the distance fallen each second under free fall.

c. How far will the object fall during the 10th second?

d. During which second will the object fall 400 feet?

History CONNECTION

Leonardo da Vinci (1452-1519) was able to discover the formula for the velocity (directional speed) of a freely falling object by looking at a sequence. He let drops of water fall at equally spaced time intervals between two boards covered with blotting paper. When a spring mechanism was disengaged, the boards clapped together. By measuring the distances between successive blots and noting that these distances increased arithmetically, da Vinci discovered the formula $v = gt$, where v is the velocity of the object, t is the time since it was released, and g is a constant that represents any object's downward acceleration due to the force of gravity.

Review

13. **APPLICATION** Suppose a company offers a new employee a starting salary of \$18,150 with annual raises of \$1,000, or a starting salary of \$17,900 with a raise of \$500 every six months. At what point is one choice better than the other? Explain.

14. Suppose that you add 300 mL of water to an evaporating dish at the start of each day, and each day 40% of the water in the dish evaporates.

a. Write and solve an equation that computes the long-run water level in the dish.

b. Will a 1 L dish do the job? Explain why or why not.

15. **APPLICATION** Five stores in Tulsa, Oklahoma, sell the same model of a graphing calculator for \$89.95, \$93.49, \$109.39, \$93.49, and \$97.69.

a. What are the median price, the mean price, and the standard deviation?

b. If these stores are representative of all stores in the Tulsa area, of what importance is it to a consumer to know the median, mean, and standard deviation? Which is probably more helpful, the median or the mean?

16. ***Technology*** Use statistics software to make a histogram of this data set: 44, 45, 51, 40, 28, 46, 34, 19.

a. Based on the histogram, predict what a box plot of the data will look like.

b. Use the software to create a box plot of the data. Were your predictions accurate?

IMPROVING YOUR REASONING SKILLS

Sequential Slopes

Here's a sequence that generates coordinate points. What is the slope between any two points of this sequence?

$(x_0, y_0) = (0, 0)$

$(x_n, y_n) = (x_{n-1} + 2,\ y_{n-1} + 3)$ where $n \geq 1$

Now match each of these recursive rules to the slope between points.

a. $(x_n, y_n) = (x_{n-1} + 2,\ y_{n-1} + 2)$

b. $(x_n, y_n) = (x_{n-1} + 1,\ y_{n-1} + 3)$

c. $(x_n, y_n) = (x_{n-1} + 3,\ y_{n-1} - 4)$

d. $(x_n, y_n) = (x_{n-1} - 2,\ y_{n-1} + 10)$

e. $(x_n, y_n) = (x_{n-1} + 1,\ y_{n-1})$

f. $(x_n, y_n) = (x_{n-1} + 9,\ y_{n-1} + 3)$

A. 0 **B.** 1 **C.** 3 **D.** -5 **E.** $\frac{1}{3}$ **F.** $\frac{-4}{3}$

In general, how do these recursive rules determine the slope between points?

LESSON 3.2

Keymath.com Links to Resources

Revisiting Slope

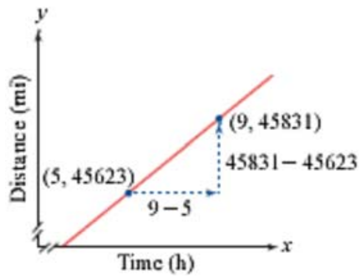
There is more to life than increasing its speed.

MOHANDAS GANDHI

Suppose you are taking a long trip in your car. At 5 P.M., you notice that the odometer reads 45,623 miles. At 9 P.M., you notice that it reads 45,831. You find your average speed during that time period by dividing the difference in distance by the difference in time.

$$\text{Average speed} = \frac{45831 \text{ miles} - 45623 \text{ miles}}{9 \text{ hours} - 5 \text{ hours}} = \frac{208 \text{ miles}}{4 \text{ hours}} = 52 \text{ miles per hour}$$

You can also write the rate 52 miles per hour as the ratio $\frac{52 \text{ miles}}{1 \text{ hour}}$. If you graph the information as points of the form (*time*, *distance*), the slope of the line connecting the two points is $\frac{52}{1}$, which also tells you the average speed.



Slope

The formula for the slope between two points, (x_1, y_1) and (x_2, y_2), is

$$slope = \frac{y_2 - y_1}{x_2 - x_1}$$

where $x_2 \neq x_1$.

The slope will be the same for any two points selected on the line. In other words, a line has only one slope. Two points on a line can have the same y-value; in that case, the slope of the line is 0. If they had the same x-value, the denominator would be 0 and the slope would be undefined. So the definition of slope specifies that the points cannot have the same x-value. What kinds of lines have a slope of 0? What kinds of lines have undefined slope?

Slope is another word for the steepness or rate of change of a line. If a linear equation is in **intercept form**, then the slope of the line is the coefficient of x.

Intercept Form of the Equation of a Line

You can write the equation of a line as

$$y = a + bx$$

where a is the y-intercept and b is the slope of the line.

Slope is often represented by the letter m. However, we will use the letter b in linear equations, as in the intercept form $y = a + bx$.

When you are using real-world data, choosing different pairs of points results in choosing lines with slightly different slopes. However, if the data are nearly linear, these slopes should not differ greatly.

Investigation

Balloon Blastoff

You will need

- paper
- tape
- a balloon
- a straw
- string
- a motion sensor

In this investigation you will launch a rocket and use your motion sensor's data to calculate the rocket's speed. Then you will write an equation for the rocket's distance as a function of time. Choose one person to be the monitor and one person to be the launch controller.

Procedure Note

1. Make a rocket of paper and tape. Design your rocket so that it can hold an inflated balloon and be taped to a drinking straw threaded on a string. Color or decorate your rocket if you like.
2. Tape your rocket to the straw on the string.
3. Inflate a balloon but do not tie off the end. The launch controller should insert it into your rocket and hold it closed.

Step 1 The monitor holds the sensor about 2.5 meters in front of the rocket and counts down to blastoff. When the monitor presses the trigger on the sensor and says, “blastoff,” the launch controller releases the balloon. [See Calculator Note 3C.]

Step 2 Retrieve the data from the sensor to each calculator in the group.

Step 3 Graph the data with time as the independent variable, x. What are the domain and range of your data? Explain.

Step 4 Select four points that you think will give the most accurate slope for the data. Indicate the points you selected on a graph of these data, and explain why you chose them. Use these points in pairs to calculate slopes. This should give six values for the slope.

Step 5 Are all six slope values that you calculated in Step 4 the same? Why or why not? Find the mean, median, and mode of your slopes. With your group, decide what value best represents the slope of your data. Explain why you chose this value.

Step 6 What is the real-world meaning of the slope, and how is this related to the speed of your rocket?

Step 7 Write an equation for the rocket's distance x seconds after it is released, and explain what each part of the equation means.

Step 8 Graph your line and label it “Our rocket.” Imagine each of the following scenarios. On the same graph, sketch lines to represent each of the rockets.

a. A rocket that was released at the same time as yours, but traveled at 75% the speed of your rocket. Label this line “Slower rocket.”

b. A rocket that was released two seconds before yours, and traveled at the same speed as yours. Label this line "Same speed."

c. A rocket that was released two seconds after yours, but got caught on the string and did not go anywhere. Label this line "No movement."

Step 9 Write the equation for each of the imaginary rockets in Step 8. Label each of the lines on your graph.

In many cases, when you try to model the steepness and trend of the points, you may have some difficulty deciding which points to use. In general, select two points that are far apart to minimize the error. They do not need to be data points. Disregard data points that you think might represent measurement errors.

When you analyze a relationship between two variables, you must decide which variable you will express in terms of the other. When one variable depends on the other variable, it is called the **dependent variable.** The other variable is called the **independent variable.** Time is usually considered an independent variable.

Next, you need to think about the domain and range. The set of possible x-values is called the **domain** and the set of y-values is called the **range.**

EXAMPLE

Daron's car gets 20 miles per gallon of gasoline. He starts out with a full tank, 16.4 gallons. As Daron drives, he watches the gas gauge to see how much gas he has left.

a. Identify the independent and dependent variables.

b. Write a linear equation in intercept form to model this situation.

c. How much gas will be left in Daron's tank after he drives 175 miles?

d. How far can he travel before he has less than 2 gallons remaining?

Solution

The two variables are the distance Daron has driven and the amount of gasoline remaining in his car's tank.

a. The amount of gasoline remaining in the car's tank depends on the number of miles Daron has driven. This means the amount of gasoline is the dependent variable and distance is the independent variable. So use x for the distance (in miles) and y for the amount of gasoline (in gallons).

b. Daron starts out with 16.4 gallons. He drives 20 miles per gallon, which means the amount of gasoline decreases $\frac{1}{20}$, or 0.05, gallon per mile. The equation is $y = 16.4 - 0.05x$.

c. You know the x-value is 175 miles. You can substitute 175 for x and solve for y.

$$y = 16.4 - 0.05 \cdot 175$$
$$= 7.65$$

He will have 7.65 gallons remaining.

d. You know the y-value is 2.0 gallons. You can substitute 2.0 for y and solve for x.

$$2.0 = 16.4 - 0.05x$$
$$-14.4 = -0.05x$$
$$288 = x$$

When he has traveled more than 288 miles, he will have less than 2 gallons in his tank.

EXERCISES

Practice Your Skills

1. Find the slope of the line containing each pair of points.

a. (3, -4) and (7, 2) **b.** (5, 3) and (2, 5) **c.** (-0.02, 3.2) and (0.08, -2.3)

2. Find the slope of each line.

a. $y = 3x - 2$ **b.** $y = 4.2 - 2.8x$ **c.** $y = 5(3x - 3) + 2$

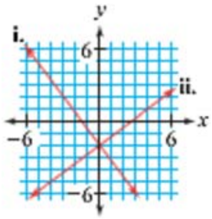
d. $y - 2.4x = 5$ **e.** $4.7x + 3.2y = 12.9$ **f.** $\frac{2}{3}y = \frac{2}{3}x + \frac{1}{2}$

3. Solve each equation.

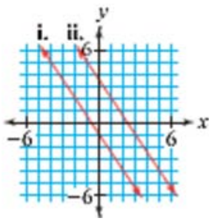
a. Solve $y = 4.7 + 3.2x$ for y if $x = 3$.
b. Solve $y = -2.5 + 1.6x$ for x if $y = 8$.
c. Solve $y = a - 0.2x$ for a if $x = 1000$ and $y = -224$.
d. Solve $y = 250 + bx$ for b if $x = 960$ and $y = 10$.

4. Find the equations of both lines in each graph.

a.



b.



5. Consider the equations and graphs of Exercise 4.

a. What do the equations in 4a have in common? What do you notice about their graphs?
b. What do the equations in 4b have in common? What do you notice about their graphs?

Reason and Apply

6. Use this graph to determine the speed of a balloon rocket. The independent variable is time in seconds, and the dependent variable is distance from the sensor in meters.

[0, 10, 1, 0, 8, 1]

7. **APPLICATION** Layton measures the voltage across different numbers of batteries placed end to end. He records his data in a table.

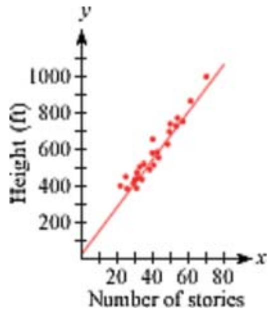
Number of batteries	1	2	3	4	5	6	7	8
Voltage (volts)	1.43	2.94	4.32	5.88	7.39	8.82	10.27	11.70

a. Let x represent the number of batteries, and let y represent the voltage. Find the slope of a line approximating these data. Be sure to include units with your answer.

b. Which points did you use and why? What is the real-world meaning of this slope?

c. Does it make sense that the y-intercept of the line is 0? Explain why or why not.

8. This graph shows the relationship between the height of some high-rise buildings and the number of stories in those buildings. A line is drawn to fit the data.



a. Estimate the slope. What is the meaning of the slope?

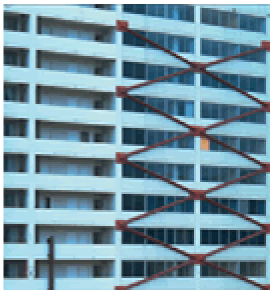
b. Estimate the y-intercept. What is the meaning of the y-intercept?

c. Explain why some of the points lie above the line and some lie below.

d. According to the graph, what are the domain and range of this relationship?

Engineering CONNECTION

Many earthquake-prone areas in the United States have strict building codes to ensure that high-rise buildings can withstand earthquakes. Some taller buildings have cross-bracing to make them more rigid and lessen the amount of shaking. The extent of damage that an earthquake can cause depends on several factors: strength of the earthquake, type of underlying soil, and building construction. The shaking increases with height, and the period of shaking (in seconds) is approximately equal to 0.1 times the number of stories in the high-rise. For more information on engineering and earthquakes, see the web links at www.keymath.com/DAA.



Steel braces will reduce earthquake damage to this concrete building.

9. This formula models Anita's salary for the last seven years: $u_n = 847n + 17109$. The variable n represents the number of years of experience she has, and u_n represents her salary in dollars.

a. What did she earn in the fifth year? What did she earn in her first year? (Think carefully about what n represents.)

b. What is the rate of change of her salary?

c. What is the first year Anita's salary will be more than $30,000?

10. **APPLICATION** This table shows how long it took to lay tile for hallways of different lengths.

Length of hallway (ft)	3.5	9.5	17.5	4.0	12.0	8.0
Time (min)	85	175	295	92	212	153

a. What is the independent variable? Why? Make a graph of the data.
b. Find the slope of the line through these data. What is the real-world meaning of this slope?
c. Which points did you use and why?
d. Find the y-intercept of the line. What is the real-world meaning of this value?

11. **APPLICATION** The manager of a concert hall keeps data on the total number of tickets sold and total sales income, or revenue, for each event. Two different ticket prices are offered.

Total tickets	448	601	297	533	523	493	320
Total revenue ($)	3357.00	4495.50	2011.50	3784.50	3334.50	3604.50	2353.50

a. Find the slope of a line approximating these data. What is the real-world meaning of this slope?
b. Which points did you use and why?

12. **APPLICATION** How much does air weigh? The following table gives the weight of a cubic foot of dry air at the same pressure at various temperatures in degrees Fahrenheit.

Temp. (°F)	0	12	32	52	82	112	152	192	212
Weight (lb)	0.0864	0.0842	0.0807	0.0776	0.0733	0.0694	0.0646	0.0609	0.0591

a. Make a scatter plot of the data.
b. What is the approximate slope of the line through the data?
c. Describe the real-world meaning of the slope.

Recreation CONNECTION

Hot air rises, cool air sinks. As air is heated up, it becomes less dense and lighter than the cooler air surrounding it. This simple law of nature is the principle behind hot-air ballooning. By heating the air in the balloon envelope, maintaining its temperature, or letting it cool, the balloon's pilot is able to climb higher, fly level, or descend. Joseph and Jacques Montgolfier developed the first hot-air balloon in 1783, inspired by the rising of a shirt that was drying above a fire.

Review

13. Rewrite each expression by eliminating parentheses and then combining like terms.

a. $2 + 3(x - 4)$ **b.** $(11 - 3x) - 2(4x + 5)$ **c.** $5.1 - 2.7(1 - (2x + 9.7))$

14. Solve each equation.

a. $12 = 6 + 2(x - 1)$ **b.** $27 = 12 - 2(x + 2)$

15. Charlotte and Emily measured the pulse rates of everyone in their class in beats per minute and collected this set of data.

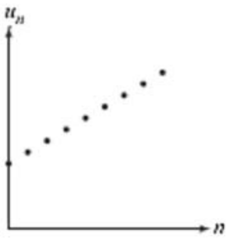
{62, 68, 68, 70, 74, 66, 82, 74, 76, 72, 70, 68, 80,
60, 84, 72, 66, 78, 70, 68, 66, 82, 76, 66, 66, 80}

a. What is the mean pulse rate for the class?

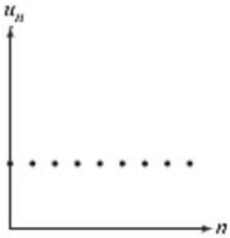
b. What is the standard deviation? What does this tell you?

16. Each of these two graphs was generated by a recursive formula in the form $u_0 = a$ and $u_n = (1 + r)u_{n-1} + p$ where $n \geq 1$. Describe the parameters a, r, and p that produce each graph. (There are two answers to 16b.)

a.

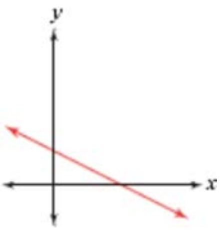


b.

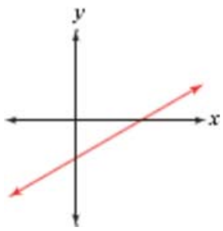


17. Each of these graphs was produced by a linear equation in the form $y = a + bx$. For each case, tell if a and b are greater than zero, equal to zero, or less than zero.

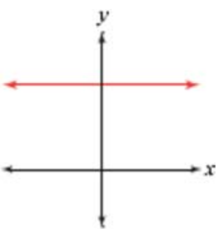
a.



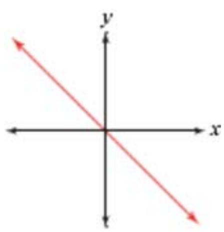
b.



c.



d.



LESSON 3.3

Fitting a Line to Data

Odd how the creative power at once brings the whole universe to order.

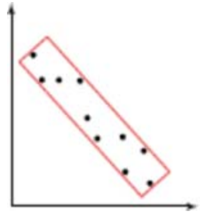
VIRGINIA WOOLF

All the points of an arithmetic sequence lie on a line. When you collect data and make a graph, sometimes the data will appear to have a linear relationship. However, the points will rarely lie on a single line. They will usually be scattered, and it is up to you to determine a reasonable location for the line that summarizes or gives the trend of the data set. A line that fits the data reasonably well is called a **line of fit**.

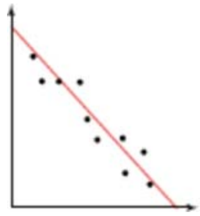
There is no single list of rules that will give the best line of fit in every instance, but you can use these guidelines to obtain a reasonably good fit.

Finding a Line of Fit

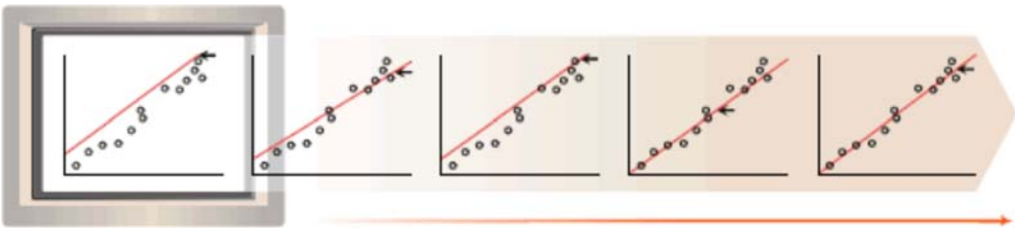
1. Determine the direction of the points. The longer side of the smallest rectangle that contains most of the points shows the general direction of the line.



2. The line should divide the points equally. Draw the line so that there are about as many points above the line as below the line. The points above the line should not be concentrated at one end, and neither should the points below the line. The line has nearly the same slope as the longer sides of the rectangle.



Some computer applications allow you to manually fit a line to data points. You can also graph the data by hand and draw the line of fit.



Once you have drawn a line of fit for your data, you can write an equation that expresses the relationship. To indicate that the line is a prediction line, the variable $\hat{y}$ (“*y* hat”) is used in place of *y*.

As you learned in algebra, there are several ways to write the equation of a line. You can find the slope and the y-intercept and write the intercept form of the equation. Often, it is easier to choose any two points on the line, use them to calculate the slope, and use the slope and either of the points to write the **point-slope form** of the equation. Either method should give almost exactly the same results. In general, using two points that are farther apart results in a more accurate calculation of the slope.

Point-Slope Form

The formula for the slope is $b = \frac{y_2 - y_1}{x_2 - x_1}$, so you can write the equation of a line with slope b and containing point (x_1, y_1) for any general point (x, y) as $b = \frac{y - y_1}{x - x_1}$ or, equivalently, as

$$y = y_1 + b(x - x_1)$$

This is called the point-slope form for a linear equation.

You can then use this equation to predict points for which data are not available.

EXAMPLE

On a barren lava field on top of the Mauna Loa volcano in Hawaii, scientists have been monitoring the concentration of CO_2 (carbon dioxide) in the atmosphere since 1959. This site is favorable because it is relatively isolated from vegetation and human activities that produce CO_2. The average concentrations for 17 different years, measured in parts per million (ppm), are shown here.

Year	CO_2 (ppm)	Year	CO_2 (ppm)	Year	CO_2 (ppm)
1980	337.84	1987	347.84	1995	359.98
1981	339.06	1988	350.25	1996	362.09
1982	340.57	1989	352.60	1997	363.23
1983	341.20	1990	353.50	1998	365.38
1984	343.52	1993	356.63	1999	368.24
1986	346.11	1994	358.96		

(Carbon Dioxide Information Analysis Center)

a. Find a line of fit to summarize the data.

b. Predict the concentration of CO_2 in the atmosphere in the year 2050.

Science CONNECTION

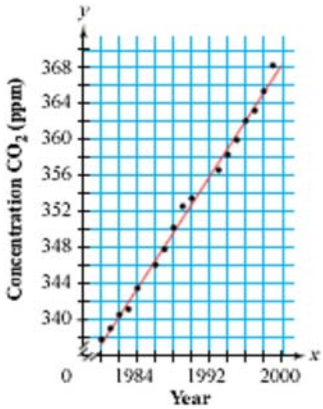
With a name meaning "long mountain," Mauna Loa has an area of 2035 mi^2, covers half of the island of Hawaii, and is Earth's largest active volcano. This active volcano's last eruption was in 1984. Volcanologists routinely monitor Mauna Loa for signs of eruption and specify hazardous areas of the mountain.


Caldera on Mauna Loa

Solution

Let x represent the year, and let y represent the concentration of CO_2 in parts per million. A graph of the data shows a linear pattern.

a. You can draw a line that seems to fit the trend of the data.



The line shown on the graph is a good fit because many of the points lie on or near the line, the data points above and below the line are roughly equal in number, and they are evenly distributed on both sides along the line. You do not know what the y-intercept should be because you do not know the concentration in the year 0; you have no reason to believe that the line passes through (0, 0).

Next, you need to find the equation for this line of fit. The y-intercept is not easily available, but you can choose two points and use the point-slope form. Note that the points do not need to be data points. You might choose the points (1984, 343.5) and (1994, 359.0). They are far enough apart, and their coordinates are easy to find on the graph.

You can use these points to find the slope.

$$\text{slope} = \frac{359.0 - 343.5}{1994 - 1984} = \frac{15.5}{10} = 1.55$$

This means the concentration of CO_2 in the atmosphere has been increasing at a rate of about 1.55 ppm each year.

Use either of the two points you chose for the slope calculation, and substitute its coordinates for x_1 and y_1 in the point-slope form of a linear equation.

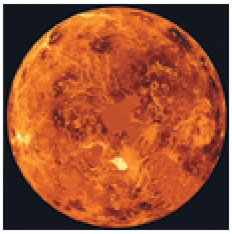
$$\hat{y} = 343.5 + 1.55(x - 1984)$$

Use point (1984, 343.5), substituting 1984 for x_1 and 343.5 for y_1.

b. You can substitute 2050 for x and solve for $\hat{y}$ to predict the CO_2 concentration in the year 2050. The prediction is 445.8 ppm.

$$\hat{y} = 343.5 + 1.55(2050 - 1984)$$
$$= 445.8$$

Environment CONNECTION



Cars and trucks emit carbon dioxide, methane, and nitrous oxide by burning fossil fuels. These gases trap heat from the sun, producing a "greenhouse effect" and causing global warming. Some scientists warn that if accelerated warming continues, higher sea levels will result. To learn how scientists are studying the greenhouse effect, see the web links at www.keymath.com/DAA.

The planet Venus suffers from an extreme greenhouse effect due to constant volcanic activity, which has created a dense atmosphere that is 97% carbon dioxide. As a result, surface temperatures reach 462°C.

Investigation

The Wave

You will need

- a stop watch or watch with second hand

Sometimes at sporting events, people in the audience stand up quickly in succession with their arms upraised and then sit down again. The continuous rolling motion that this creates through the crowd is called "the wave." You and your class will investigate how long it takes different-size groups to do the wave.

Step 1 Using different-size groups, determine the time for each group to complete the wave. Collect at least nine pieces of data of the form (*number of people, time*), and record them in a table.

Step 2 Plot the points, and find the equation of a reasonable line of fit. Write a paragraph about your results. Be sure to answer these questions:

- What is the slope of your line, and what is its real-world meaning?
- What are the x- and y-intercepts of your line, and what are their real-world meanings?
- What is a reasonable domain for this equation? Why?

Step 3 Can you use your line of fit to predict how long it would take to complete the wave if everyone at your school participated? Everyone in a large stadium? Explain why or why not.

Keep these data and the equation. You will be using them in a later lesson.

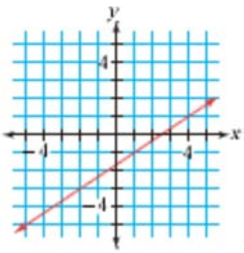
Finding a value between other values given in a data set is called **interpolation**. Using a model to extend beyond the first or last data points is called **extrapolation**. How would you use the Mauna Loa data in the example to estimate the CO_2 levels in 1960? In 1991?

EXERCISES

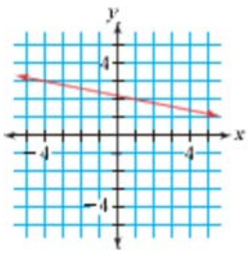
Practice Your Skills

1. Write the equation in point-slope form of each line shown.

a.



b.



2. Write the equation in point-slope form of each line described.

a. Slope $\frac{2}{3}$ passing through (5, –7)

b. Slope -4 passing through (1, 6)

c. Parallel to $y = -2 + 3x$ passing through (-2, 8)

d. Parallel to $y = -4 - \frac{3}{5}(x + 1)$ passing through (-4, 11)

3. Solve each equation.

a. Solve $u_n = 23 + 2(n - 7)$ for u_n if $n = 11$.

b. Solve $d = -47 - 4(t + 6)$ for t if $d = 95$.

c. Solve $y = 56 - 6(x - 10)$ for x if $y = 107$.

4. Consider the line $y = 5$.

a. Graph this line and identify two points on it.

b. What is the slope of this line?

c. Write the equation of the line that contains the points (3, -4) and (-2, -4).

d. Write three statements about horizontal lines and their equations.

5. Consider the line $x = -3$.

a. Graph it and identify two points on it.

b. What is the slope of this line?

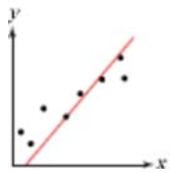
c. Write the equation of the line that contains the points (3, 5) and (3, 1).

d. Write three statements about vertical lines and their equations.

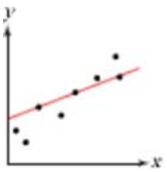
Reason and Apply

6. Of the graphs below, choose the *one* with the line that best satisfies the guidelines on page 128. For each of the other graphs, explain which guidelines the line violates.

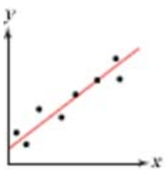
a.



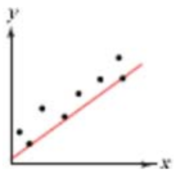
b.



c.

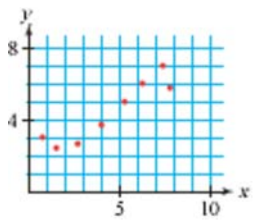


d.

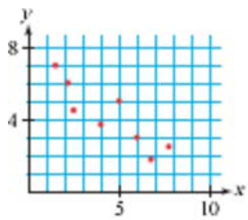


7. For each graph below, lay your ruler along your best estimate of the line of fit. Estimate the y-intercept and the coordinates of one other point on the line. Write an equation in intercept form for the line of fit.

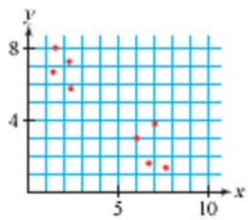
a.



b.



c.



8. **APPLICATION** A photography studio offers several packages to students posing for yearbook photos. Let x represent the number of pictures, and let y represent the price in dollars.

Number of pictures	44	31	24	15
Price ($)	19.00	16.00	13.00	10.00

a. Plot the data, and find an equation of a line of fit. Explain the real-world meaning of the slope of this line.
b. Find the y-intercept of your line of fit. Explain the real-world meaning of the y-intercept.
c. If the studio offers a 75-print package, what do you think it should charge?
d. How many prints do you think the studio should include in the package for a $7.99 special?

9. Use height as the independent variable and length of forearm as the dependent variable for the data collected from nine students.

a. Name a good graphing window for your scatter plot.
b. Write a linear equation that models the data.
c. Write a sentence describing the real-world meaning of the slope of your line.
d. Write a sentence describing the real-world meaning of the y-intercept. Explain why this doesn't make sense and how you might correct it.
e. Use your equation to estimate the height of a student with a 50 cm forearm and to estimate the length of a forearm of a student 158 cm tall.

Height (cm)	Forearm (cm)
185.9	48.5
172.0	44.5
155.0	41.0
191.5	50.5
162.0	43.0
164.3	42.5
177.5	47.0
180.0	48.0
179.5	47.5

10. **APPLICATION** This data set was collected by a college psychology class to determine the effects of sleep deprivation on students' ability to solve problems. Ten participants went 8, 12, 16, 20, or 24 hours without sleep and then completed a set of simple addition problems. The number of addition errors was recorded.

Hours without sleep	8	8	12	12	16	16	20	20	24	24
Number of errors	8	6	6	10	8	14	14	12	16	12

a. Define your variables and create a scatter plot of the data.
b. Write an equation of a line that approximates the data and sketch it on your graph.
c. Based on your model, how many errors would you predict a person to make if she or he hadn't slept in 22 hours?
d. In 10c, did you use interpolation or extrapolation? Explain.

Review

11. The 3rd term of an arithmetic sequence is 54. The 21st term is 81. Find the 35th term.

12. Write the first four terms of this sequence and describe its long-run behavior.

$u_1 = 56$

$u_n = \frac{u_{n-1}}{2} + 4$ where $n \geq 2$

13. Given the data set {20, 12, 15, 17, 21, 15, 30, 16, 14}:

a. Find the median.

b. Add as few elements as possible to the set in order to make 19.5 the median.

14. You start 8 meters from a marker and walk toward it at the rate of 0.5 m/s.

a. Write a recursive rule that gives your distance from the marker after each second.

b. Write an explicit formula that allows you to find your distance from the marker at any time.

c. Interpret the real-world meaning of a negative value for u_n in 14a or 14b.

Project

TALKIN' TRASH

Garbage piles up in New York City during a 1982 sanitation workers strike.

How much trash do you and your family generate each day? How much trash does that amount to for the entire U.S. population daily? Annually? How much land is needed to dump garbage? Research some data on U.S. population and waste production. Find linear models for your data. Use your equations to predict the population and the amount of waste that you might expect in the years 2010 and 2020. Use your graphs and equations to decide if the amount of waste is increasing because of the increase in population or because of other factors.

Your project should include

- Data and sources.
- Linear models for predicting population and waste, an explanation of how you found them, and real-world meanings for each part of your model.
- Domain and range for each model.
- Population and waste predictions for 2010 and 2020, including amount of waste per person per day.
- A complete analysis of your findings, in paragraph form.

LESSON 3.4

Keymath.com Links to Resources

The Median-Median Line

The growth of understanding follows an ascending spiral rather than a straight line.

JOANNA FIELD

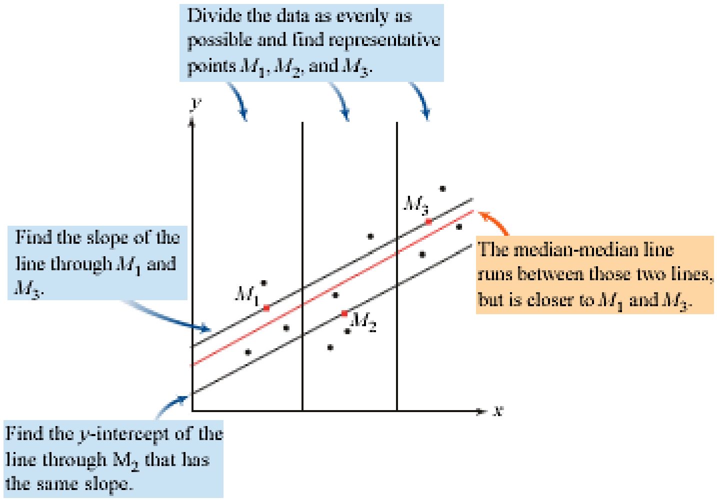
Have you noticed that you and your classmates frequently find different equations to model the same data? Some lines fit data better than others and can be used to make more accurate predictions. In this lesson you will learn a standard method for finding a line of fit that will enable each member of the class to get the same equation for the same set of data.

You can use many different methods to find a statistical line of fit. The **median-median line** is one of the simplest methods. The procedure for finding the median-median line uses three points (M_1, M_2, and M_3) to represent the entire data set, and the equation that best fits these three points is taken as the line of fit for the entire set of data.

To find the three points that will represent the entire data set, you first order all the data points by their domain value (the *x*-value) and then divide the data into three equal groups. If the number of points is not divisible by 3, then you split them so that the first and last groups are the same size. For example:

18 data points: split into groups of 6-6-6
19 data points: split into groups of 6-7-6
20 data points: split into groups of 7-6-7

You then order the *y*-values within each of the groups. The representative point for each group has the coordinates of the median *x*-value of that group and the median *y*-value of that group. Because a good line of fit divides the data evenly, the median-median line should pass between M_1, M_2, and M_3, but be closer to M_1 and M_3 because they represent two-thirds of the data. To accomplish this, you can find the *y*-intercept of the line through M_1 and M_3, and the *y*-intercept of the line through M_2 that has the same slope. The mean of the three *y*-intercepts of the lines through M_1, M_2, and M_3 gives you the *y*-intercept of a line that satisfies these requirements.



Don't forget to order the y-values in each group when finding the median y-value. Carefully study the following example to see how this is done.

EXAMPLE

Find the median-median line for these data.

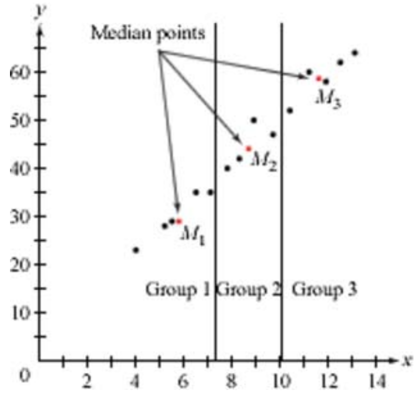
x	y
4.0	23
5.2	28
5.8	29
6.5	35
7.1	35
7.8	40
8.3	42

x	y
8.9	50
9.7	47
10.4	52
11.2	60
11.9	58
12.5	62
13.1	64

▸Solution

First, group the data and find M_1, M_2, and M_3.

x	y	(median x, median y)
4.0	23	
5.2	28	
5.8	29	(5.8, 29)
6.5	35	
7.1	35	
7.8	40	
8.3	42	(8.6, 44.5)
8.9	50	
9.7	47	
10.4	52	
11.2	60	
11.9	58	(11.9, 60)
12.5	62	
13.1	64	



The slope of the median-median line is determined by the slope of the line through M_1 and M_3.

$$\text{slope} = \frac{60 - 29}{11.9 - 5.8} \approx 5.082$$

Next, find the equation of the line containing points $M_1(5.8, 29)$ and $M_3(11.9, 60)$. You have already found that the slope is 5.082. Using M_1, you can write the point-slope form of the equation. Rewrite the equation in intercept form to find the y-intercept.

$y = 29 + 5.082(x - 5.8)$ Write the equation in point-slope form.

$y = 29 + 5.082x - 29.4756$ Distribute the 5.082.

$y = -0.476 + 5.082x$ Add like terms. This is the intercept form. The y-intercept is -0.476.

If you use M_3 instead of M_1, you should get the same equation, because your equation is the line through both M_1 and M_3.

Next, find the equation of the line that is parallel to the line through M_1 and M_3 and passes through the middle representative point, $M_2(8.6, 44.5)$.

$y = 44.5 + 5.082(x - 8.6)$	Write the equation in point-slope form.
$y = 0.795 + 5.082x$	Distribute and add like terms to find the intercept form of the equation.

The median-median line is parallel to both of these lines, so it will also have a slope of 5.082. To find the y-intercept of the median-median line, you find the mean of the y-intercepts of the lines through M_1, M_2, and M_3. Note that the y-intercept of the line through M_1 is the same as the y-intercept of the line through M_3.

$$\frac{-0.476 + (-0.476) + 0.795}{3} \approx -0.052$$ Find the mean of the y-intercepts.

So, finally, the equation of the median-median line is $\hat{y} = -0.052 + 5.082x$. Note that this line is one-third of the way from the first line to the second line.

Finding a Median-Median Line

1. Order your data by domain value first. Then, divide the data into three sets equal in size. If the number of points does not divide evenly into three groups, make sure that the first and last groups are the same size. Find the median x-value and the median y-value in each group. Call these points M_1, M_2, and M_3.
2. Find the slope of the line through M_1 and M_3. This is the slope of the median-median line.
3. Find the equation of the line through M_1 with the slope you found in Step 2. The equation of the line through M_3 will be the same.
4. Find the equation of the line through M_2 with the slope you found in Step 2.
5. Find the y-intercept of the median-median line by taking the mean of the y-intercepts of the lines through M_1, M_2, and M_3. The y-intercepts of the lines through M_1 and M_3 are the same. Finally, write the equation of the median-median line using this mean y-intercept and the slope from Step 2.

Japanese artist Yoshio Itagaki (b 1967) created *Tourist on the Moon #2* as a commentary on tourists' desire to document their visits to spectacular scenes. He is both amused by and critical of the human appetite for sensation and novelty. This work is a triptych-a piece in three panels.

Investigation

Spring Experiment

You will need

- a spring
- a mass holder
- small unit masses
- a ruler

In this investigation you will collect data on how a spring responds to various weights. You will find a median-median line to model this relationship.

Procedure Note

1. Attach the mass holder to the spring.
2. Hang the spring from a support, and the mass holder from the spring.
3. Measure the length of the spring (in centimeters) from the first coil to the last coil.

Step 1 Place different amounts of mass on the mass holder, recording the corresponding length of the spring each time. Collect about 10 data points of the form (*mass, spring length*).

Step 2 Graph the data on graph paper.

Step 3 Show the steps to calculate the median-median line through the data. Write the equation of this line. Use your calculator to check your work. [See Calculator Note 3D to learn how to find the median-median line with your calculator.]

Step 4 On your graph, mark the three representative points used in the median-median process. Add the line to this graph.

Step 5 Answer these questions about your data and model.

a. Use your median-median line to interpolate two points for which you did not collect data. What is the real-world meaning of each of these points?

b. Which two points differ the most from the value predicted by your equation? Explain why.

c. What is the real-world meaning of your slope?

d. Find the y-intercept of your median-median line. What is its real-world meaning?

e. What are the domain and range for your data? Why?

f. Compare the median-median line method to the method you used in Lesson 3.3 to find the line of fit. What are the advantages and disadvantages of each? In your opinion, which method produces a better line of fit? Why?

Step 6 Summarize what you learned in this investigation and describe any difficulties you had.

There are many ways to find a line of fit for linear data. Estimating the line of fit is adequate in many cases, but different people will estimate different lines of fit and may use their own bias to draw the line higher or lower. A median-median line is a systematic method, accepted by statisticians, that summarizes the overall trend in linear data.

EXERCISES

Practice Your Skills

1. How should you divide the following sets into three groups for the median-median line method?

a. Set of 51 elements **b.** Set of 50 elements **c.** Set of 47 elements **d.** Set of 38 elements

2. Find an equation in point-slope form of the line passing through

a. (8.1, 15.7) and (17.3, 9.5) **b.** (3, 47) and (18, 84)

3. Find an equation in point-slope form of the line parallel to $y = -12.2 + 0.75x$ that passes through the point (14.4, 0.9).

4. Find an equation of the line one-third of the way from $y = -1.8x + 74.1$ to $y = -1.8x + 70.5$. (Hint: The first line came from two points in a median-median procedure, and the other line came from the third point.)

Keymath.com More Practice Your Skills

5. Find the equation of the line one-third of the way from the line $y = 2.8 + 4.7x$ to the point (12.8, 64).

Reason and Apply

6. **APPLICATION** Follow these steps to find the equation for the median-median line for the data on life expectancy at birth for males in the United States for different years in the 20th century. Let x represent the year, and let y represent the male life expectancy in years.

a. How many points are there in each of the three groups?
b. What are the three representative points for these data? Graph these points.
c. Draw the line through the first and third points. What is the slope of this line? What is the real-world meaning of the slope?
d. Write the equation of the line through the first and third points. Rewrite this equation in intercept form.
e. Draw the line parallel to the line in 6d passing through the second point. What is the equation of this line? Rewrite this equation in intercept form.
f. Find the mean of the y-intercepts and write the equation of the median-median line. Graph this line. Remember to use the intercepts of all three lines.
g. The year 1978 is missing from the table. Using your model, what would you predict the life expectancy at birth to be for males born in 1978?
h. Use your model to predict the life expectancy at birth for males born in 1991.
i. Using this model, when would you predict the life expectancy at birth for males in the United States to exceed 80 years?

Year of birth	Male life expectancy (years)
1920	53.6
1925	56.3
1930	58.1
1935	59.4
1940	60.8
1945	62.8
1950	65.6
1955	66.2
1960	66.6
1965	66.8
1970	67.1
1975	68.8
1980	70.0
1985	71.2
1990	71.8
1995	72.5
1998	73.9

(*The World Almanac and Book of Facts 2001*)

7. Choose an investigation in this chapter. Find the **residual** for each data point (difference between the actual y-value and the model-predicted y-value). Display these residuals with a histogram or box plot. Using the information in your histogram or box plot, describe how good you think your model is. Justify your conclusion.

8. Refer to your data from the Investigation The Wave in Lesson 3.3 and find the equation for the median-median line. Compare this equation with the one you found previously. Which equation do you feel is a better model for the data? Why?

9. **APPLICATION** Use these data on world records for the 1-mile run to answer the questions below. Times are in minutes and seconds.

World Records for 1-Mile Run

Year	Runner	Time
1915	Norman Taber, U.S.	4:12.6
1923	Paavo Nurmi, Finland	4:10.4
1937	Sydney Wooderson, U.K.	4:06.4
1942	Gunder Haegg, Sweden	4:06.2
1945	Gunder Haegg, Sweden	4:01.4
1958	Herb Elliott, Australia	3:54.5
1967	Jim Ryun, U.S.	3:51.1
1979	Sebastian Coe, U.K.	3:49.0
1985	Steve Cram, U.K.	3:46.31
1993	Noureddine Morceli, Algeria	3:44.39

(*Time Almanac 1999*)

   a. Let x represent the year, and let y represent the time in seconds. What is the equation of the median-median line?
   b. What is the real-world meaning of the slope?
   c. Use the equation to interpolate and predict what new record might have been set in 1954. How does this compare with Roger Bannister's actual 1954 record of 3:59.4?
   d. Use the equation to extrapolate and predict what new record might have been set in 1875. How does this compare with Walter Slade's 1875 world record of 4:24.5?
   e. Describe some problems you might have with the meaning of 997.12 as a y-intercept.
   f. Has a new world record been set since 1993? Find more recent information on this subject and compare it with the predictions of your model.

10. Devise a mean-mean line procedure and use it on the data in Exercise 9. What problems might arise when using this method? Compare the advantages and disadvantages of this method and the median-median line method.

11. The number of deaths caused by automobile accidents, D, per hundred thousand population in the United States is given for various years, t.

t	1924	1925	1926	1927	1928	1929	1930	1931	1932	1933
D	15.5	17.1	18.0	19.6	20.8	23.3	24.5	25.2	21.9	23.3

(W. A. Wilson and J. I. Tracey, *Analytic Geometry,* 3rd ed., Boston: D. C. Heath, 1949, p. 246.)

   a. Make a scatter plot of the data.
   b. Find the three summary points.
   c. Write the equation of the median-median line.
   d. What does the slope mean?
   e. Give a possible explanation for the drop in the number of auto deaths in 1932.
   f. Would you use this model to predict the number of deaths from auto accidents in 2000? Explain why or why not.

Review

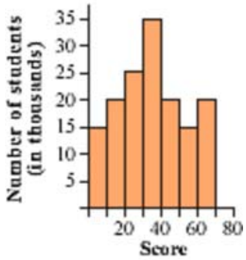
12. Create a data set of 9 values such that the median is 28, the minimum is 11, and there is no upper whisker on a box plot of the data.

13. What is the equation of the line that passes through a graph of the points of the sequence defined by

$u_1 = 4$

$u_n = u_{n-1} - 3$ where $n \geq 2$

14. The histogram at right shows the results of a statewide math test given to eleventh graders. If Ramon scored 35, what is the range of his percentile ranking?



15. Earl's science lab group made six measurements of mass and then summarized the results. Someone threw away the measurements. Help the group reconstruct the measurements from these statistics.

- The median and mean are both 3.2 g.
- The mode is 3.0 g.
- The *IQR* is 0.6 g.
- The largest deviation is –0.9 g.

16. Travis is riding with his parents on Interstate 15 across Utah. He records the digital speedometer reading in mi/h at 4:00 P.M. and every five minutes for the next hour. His record is

{61.3, 48.7, 62.4, 50.1, 60.3, 64.8, 67.1, 54.0, 60.2, 45.3, 52.3, 67.6, 63.9}

a. What is their mean speed?

b. What is the standard deviation?

c. Interpret the meaning of the standard deviation. Speculate about the traffic conditions.

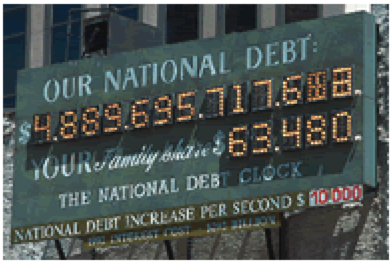
Project

COUNTING FOREVER

How long would it take you to count to 1 million? Collect data by recording the time it takes you to count to different numbers. Predict the time it would take to count to 1 million, 1 billion, 1 trillion, and to the amount of the federal debt.

Your project should include

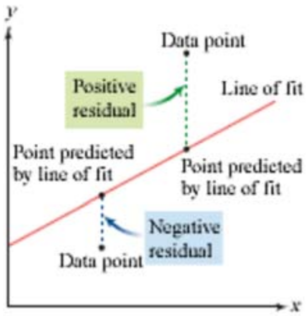
- An explanation of the procedure you used to collect these data and the approach you used to solve the problem.
- Your data, graphs, and equations.
- A summary of your predictions.



In 1992, this sign in New York City showed the increasing federal debt.

LESSON 3.5

Residuals



The median-median line method gives you a process by which anyone will get the same line of fit for a set of data. However, unless your data are perfectly linear, even the median-median line will not be perfect. In some cases you can find a more satisfactory model if you draw a line by hand. Having a line that "looks better" is not a convincing argument that it really is better.

One excellent method to evaluate your line's fit is to look at the **residuals,** or the vertical differences between the points in your data set and the points generated by your line of fit.

residual = y-value of data point – y-value of point on line

Similar to the deviation from the mean that you learned about in Chapter 2, a residual is a signed distance. Here, a positive residual indicates that the point is above the line, and a negative residual indicates that the point is below the line.

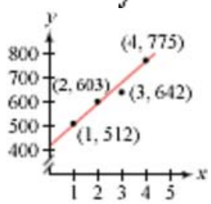
A good-fitting line should have about as many points above it as below. This means that the sum of the residuals should be near zero. In other words, if you connect each point to the line with a vertical segment, the sum of the lengths of the segments above the line should be about equal to the sum of those below the line.

EXAMPLE A

The manager of Big K Pizza must order supplies for the month of November. The numbers of pizzas sold in November during the past four years were 512, 603, 642, and 775, respectively. How many pizzas should she plan for this November?

Solution

Let x represent the past four years, 1 through 4, and let y represent the number of pizzas. Graph the data. The median-median method gives the linear model $\hat{y} = 417.3 + 87.7x$.



Calculating the residuals is one way to evaluate this model before using it to make a prediction. Evaluate the linear equation when x is 1, 2, 3, and 4. When $x = 1$, for example, $417.3 + 87.7(1) = 505.0$. A table helps organize the information.

Year (x)	1	2	3	4
Number of pizzas (y)	512	603	642	775
y-value from line $\hat{y} = 417.3 + 87.7x$	505.0	592.7	680.4	768.1
Residual	7	10.3	-38.4	6.9

The sum of the residuals is 7 + 10.3 (-38.4) + 6.9, or -14.2 pizzas, which is fairly close to zero in relation to the large number of pizzas purchased. The linear model is therefore a pretty good fit. The sum is negative, which means that all together the points below the line are a little farther away than those above.

If the manager plans for 417.3 + 87.7(5), or approximately 856 pizzas, she will be very close to the linear pattern established over the past four years. Because the residuals range from -38.4 to 7, she may want to adjust her prediction higher or lower depending on factors such as whether or not the supplies are perishable or whether or not she can easily order more supplies.

The investigation gives you another opportunity to fit a line to data and to analyze the real-world meaning of the linear model. You'll also calculate residuals and explore their meaning.

Investigation

Airline Schedules

You will need

- an airline timetable for the continental United States that includes both flight times and distances
- a time zone map for the continental United States

In this investigation you will create a linear model that relates the distance and time of an airline flight. You'll use residuals to judge the fit of your line, and you'll consider factors that may make the fit less than perfect.

This painting by Julie Mehretu (b 1970), *Retopistics: A Renegade Excavation* (2001), is based on plans of international airports.

Step 1 Decide as a class which major city is going to be the starting point. Then work with your group and record in a table the flight times and distances to at least eight other cities from the chosen starting point. Choose only nonstop flights.

Step 2 Plot your data. Let x represent the flight time, and let y represent the flight distance.

Step 3 Find the median-median line for your data. When you have your linear model, answer these questions.

a. What is the real-world meaning of the slope?

b. What is the meaning of the y-intercept?

c. What is the value of the x-intercept? What is its real-world meaning?

Step 4 Use a table to organize your data and calculate the residuals.

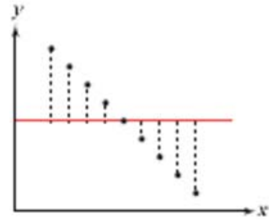
a. What is the sum of the residuals? Does it appear that your linear model is a good fit?

b. Find the greatest positive and negative residuals. What could the magnitude of these residuals indicate?

c. In general, why do you think two flights of the same distance might require different flight times? Are flight times from east to west the same as those from west to east? What other interesting observations can you share?

d. Use your model to predict the distance of a 147-minute flight. Based on the residuals, would you adjust the estimate higher or lower? Explain your reasoning.

The sum of the residuals for your line of fit should be close to 0. However, as shown at right, you could find a poorly fitting line and still get 0 for the sum of the residuals. Due to the residual sizes, you should realize that this model will not make very accurate predictions. A good-fitting line should also follow the direction of the points. This means that the individual residuals should be as close to 0 as possible.



Because the sum of the residuals does not give a complete picture, you need some way to judge how accurate predictions from your model will be. One useful measure of accuracy starts by squaring the residuals to make them all positive. For the pizza data and the linear model in Example A, this gives

Year (x)	1	2	3	4
Number of pizzas (y)	512	603	642	775
Residual	7.0	10.3	– 38.4	6.9
(Residual)^2	49	106.09	1474.56	47.61

The sum of the squares is quite large: 1677.26 pizzas^2. Perhaps some kind of average would give a better indication of the error in the predictions from this model. You could divide by 4 since there are four data points. However, this is not necessarily the best thing to do. Consider that it takes a minimum of two points to make any equation of a line. So two of the points can be thought of as defining the line. The other points determine the spread. So you should actually divide by 2 less than the number of data points.

$$\frac{1677.26}{4-2} = 838.63 \text{ pizzas}^2$$

The measure of the error should be measured in numbers of pizzas, rather than pizzas^2, so take the square root of this number.

$$\sqrt{838.63} \approx 29.0 \text{ pizzas}$$

This means that generally this line should predict values for the number of pizzas that are within 29.0 pizzas of the actual data. This value is called the **root mean square error.**

Root Mean Square Error

The root mean square error, s, is a measure of the spread of data points from a model.

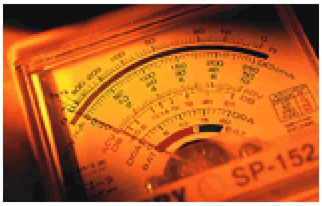
$$s=\sqrt{\frac{\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}{n-2}}$$

where y_i represents the y-values of the individual data pairs, $\hat{y}_i$ represents the respective y-values predicted from the model, and n is the number of data pairs.

You should notice that root mean square error is very similar to standard deviation, which you learned about in Chapter 2. Because of their similarities, both are represented by the variable s.

EXAMPLE B

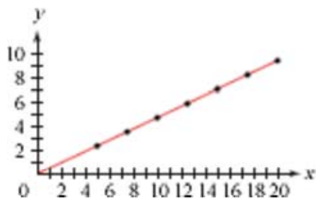
A scientist measures the current in milliamps through a circuit with constant resistance as the voltage in volts is varied. What is the root mean square error for the model $\hat{y} = 0.47x$? What is the real-world meaning of the root mean square error? Predict the current when the voltage is 25.000 volts.



Voltage (x)	5.000	7.500	10.000	12.500	15.000	17.500	20.000
Current (y)	2.354	3.527	4.698	5.871	7.053	8.225	9.403

▶ Solution

A graph of the data and the linear model shows a good fit. To calculate the root mean square error, first calculate the residual for each data point, take the sum of the squares of the residuals, divide by $n - 2$, and take a square root. [▸ See Calculator Note 3E for ways to calculate the root mean square error on your calculator.◂]



Voltage (x)	5.000	7.500	10.000	12.500	15.000	17.500	20.000
Current (y)	2.354	3.527	4.698	5.871	7.053	8.225	9.403
Prediction from line ($\hat{y}$)	2.350	3.525	4.700	5.875	7.050	8.225	9.400
Residual ($y - \hat{y}$)	0.004	0.002	–0.002	-0.004	0.003	0	0.003
(Residual)^2	1.6×10^{-5}	4.0×10^{-6}	4.0×10^{-6}	1.6×10^{-5}	9×10^{-6}	0	9.0×10^{-6}

$$s = \sqrt{\frac{5.8\times10^{-5}}{7-2}} = \sqrt{\frac{5.8\times10^{-5}}{5}} = \sqrt{1.16\times10^{-5}} \approx 0.0034$$

This means that values predicted by this equation will generally be within 0.0034 milliamp of the actual current.

For a current of 25.000 volts, the model gives 0.47(25.000), or 11.750 milliamps. Considering the root mean square error, the scientist could expect a reading between 11.7466 milliamps and 11.7534 milliamps.

You have seen many ways to find a line of fit for data. Calculating residuals and the root mean square error now gives you tools for evaluating how well your line fits the data.

EXERCISES

Practice Your Skills

1. The median-median line for a set of data is $\hat{y} = 2.4x + 3.6$. Find the residual for each of these data points.

a. (2, 8.2) **b.** (4, 12.8) **c.** (10, 28.2)

2. The median-median line for a set of data is $\hat{y} = -1.8x + 94$. This table gives the x-value and the residual for each data point. Determine the y-value for each data point.

x-value	5	8	12	20
Residual	-2.7	3.3	2.1	-1.1

3. Return to Exercise 6 in Lesson 3.4 about life expectancy for males. Use your median-median line equation to answer these questions.

a. Calculate the residuals.

b. Calculate the root mean square error for the median-median line.

c. What is the real-world meaning of the root mean square error?

4. Suppose the residuals for a data set are 0.4, -0.3, 0.2, 0.1, -0.2, -0.3, 0.2, 0.1. What is the root mean square error for this set of residuals?

Reason and Apply

5. **APPLICATION** This table gives the mean height in centimeters of boys ages 5 to 13 in the United States.

a. Define variables, plot the data, and find the median-median line.

b. Calculate the residuals.

c. What is the root mean square error for the median-median line?

d. What is the real-world meaning of the root mean square error?

e. If you use the median-median line to predict the mean height of boys age 15, what range of heights should be predicted?

Age	Height (cm)	Age	Height (cm)
5	109.2	10	138.8
6	115.7	11	143.7
7	122.0	12	149.3
8	128.1	13	156.4
9	133.7		

(National Center for Health Statistics)

6. Consider the residuals from Exercise 5b.

a. Make a box plot of these values.

b. Describe the information about the residuals that is shown in the box plot.

7. With a specific line of fit, the data point (6, 47) has a residual of 2.8. The slope of the line of fit is 2.4. What is the equation of the line of fit?

8. The following readings were taken from a display outside the First River Bank. The display alternated between °F and °C. However, there was an error within the system that calculated the temperatures.

°F	18	33	37	25	40	46	43	49	55	60	57
°C	–6	2	3	-3	5	8	7	10	12	15	13

a. Plot the data and the median-median line.

b. Calculate the residuals. You will notice that the residuals are generally negative for the lower temperatures and positive for the higher temperatures. How is this represented on the graph of the data and line?

c. Adjust your equation (by adjusting the slope or *y*-intercept) to improve this distribution of the residuals.

d. Calculate the root mean square error for the median-median line and for the equation you found in 8c. Compare the root mean square errors for the two lines and explain what this tells you.

e. Use your equation from 8c to predict what temperature will be paired with

i. 85°F **ii.** 0°C

9. Alex says, &8220;The formula for the root mean square error is long. Why do you have to square and then take the square root? Isn't that just doing a lot of work for nothing? Can't you just make them all positive, add them up, and divide?" Help Calista show Alex that his method does not give the same value as the root mean square error. Use the residual set -2, 1, -3, 4, -1 in your demonstration. Do you think Alex's method could be used as another measure of accuracy? Explain your reasoning.

10. Leajato experimented by turning the key of a wind-up car different numbers of times and recorded how far it traveled.

Turns	0.5	1	1.5	2	2.5	3	3.5	4
Distance (in.)	33	73	114	152	187	223	256	298

a. Graph the data and find the median-median line.

b. Calculate the root mean square error.

c. Predict how far the car will go if you turn the key five times. Use the root mean square error to describe the accuracy of your answer.

11. **APPLICATION** Since 1964, the total number of electors in the electoral college has been 538. In order to declare a winner in a presidential election, a majority, or 270 electoral votes, is needed. The table at right shows the number of electoral votes that the Democratic and Republican parties have received in the presidential elections since 1964.

a. Let x represent the electoral votes for the Democratic Party, and let y represent the votes for the Republican Party. Make a scatter plot of the data.

b. Why are the points nearly linear? What are some factors that make these data not perfectly linear?

c. Sketch the line $y = 270$ on the scatter plot. How are all the points above the line related?

d. Find the residuals for the line $y = 270$. What does a negative residual represent?

e. What does a residual value that is close to 0 represent?

Electoral Votes

Year	Democrats	Republicans
1964	486	52
1968	191	301
1972	17	520
1976	297	240
1980	49	489
1984	13	525
1988	111	426
1992	370	168
1996	379	159
2000	266	271

History CONNECTION

Article II, Section 1, of the U.S. Constitution instituted the electoral college as a means of electing the president. The number of electoral votes allotted to each state corresponds to the number of representatives that each state sends to Congress. The distribution of electoral votes among the states can change every 10 years depending on the results of the U.S. census. The actual process of selecting electors is left for each state to decide.

Election workers in Dade County, Florida, hand check ballots during the 2000 presidential election.

Review

12. Write an equation in point-slope form for each of these lines.

a. The slope is $\frac{3}{5}$ and the line passes through (2, 4.7).

b. The line has slope –7 and x-intercept 6.

c. The line passes through (3, 11) and (-6, -18).

13. Create a set of 7 values with median 47, minimum 28, and interquartile range 12.

14. Solve.

a. $3 + 5x = 17 - 2x$ **b.** $12 + 3(t - 5) = 6t + 1$

15. David deposits \$30 into his bank account at the end of each month. The bank pays 7% annual interest compounded monthly.

a. Write a recursive formula to show David's balance at the end of each month.

b. How much of the balance was deposited and how much interest is earned after

i. 1year **ii.** 10 years **iii.** 25 years **iv.** 50 years

c. What can you conclude about regular saving in a bank with compound interest?

EXPLORATION

Residual Plots and Least Squares

You have learned a couple of different ways to fit a line to data. You've also used residuals to judge how well a line fits and to give a range for predictions made with that line. In this activity, you will learn two graphical methods to judge how well your line fits the data. You will also use these methods to identify outliers.

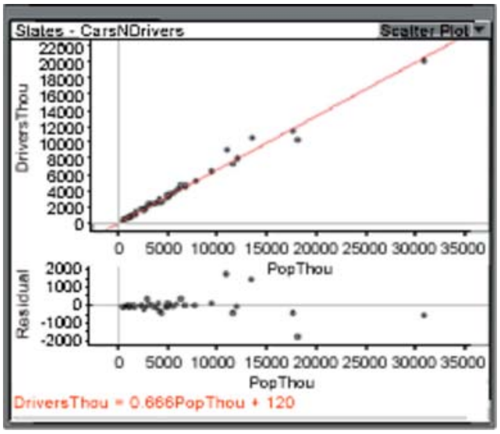
Activity

A Good Fit?

Step 1 Start Fathom and open the sample document titled **States-CarsNDrivers.ftm.** When the file opens you will see a collection and a case table. This collection gives you various data about the population, drivers, vehicles, and roadways in each U.S. state and the District of Columbia in the year 1992.

Step 2 Create a new graph. Drag the attribute PopThou (population in thousands) to the *x*-axis and drag the attribute DriversThou (licensed drivers in thousands) to the *y*-axis. Your graph will automatically become a scatter plot. Describe the trend in the data and give a possible explanation for the trend.

Step 3 With the graph window selected, choose **Movable Line** from the Graph menu. Drag the movable line until it fits the data well. What is the equation of your estimated line of fit? Based on your line, which states are outliers?



Step 4 Choose **Make Residual Plot** from the Graph menu. What does the residual plot show you? From the residual plot, which states are outliers? Are they the same states you selected in Step 3?

Step 5 Experiment by moving the movable line and observe how the residual plot changes.

Step 6 Return to your scatter plot and residual plot. With the graph window selected, choose **Show Squares** from the Graph menu. Move the movable line and observe how the squares change. Explain how each square is drawn.

Step 7 Move your line back to a position where it is a good fit. (You might want to turn off the squares while you do this.) Notice the size of the squares for the outliers that you identified in Step 4. How do they compare to the other squares?

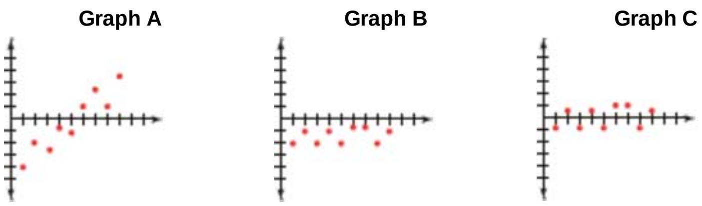
Step 8 In addition to the movable line, Fathom will graph a median-median line and a least-squares line. (You'll learn more about the least squares line in Chapter 13.) Try adding a median-median line and a least squares line to your graph, either alone or with your estimate of a line of fit. How do the equations compare? How do the residual plots compare? How do the squares compare?

Questions

1. How can you identify outliers from the squares? How are approaches interrelated? How did you identify outliers from looking at the line of fit? How did you identify outliers from the residual plot? Do you find one approach easier than the other?

2. As you change the slope of your line, what happens to the residual plot? As you change the *y*-intercept of your line, what happens to the residual plot? Explain how you can use the residual plot to adjust the fit of your line.

3. Graphs A, B, and C are residual plots for different lines of fit for the same data set. How would you adjust each line to be a better fit?

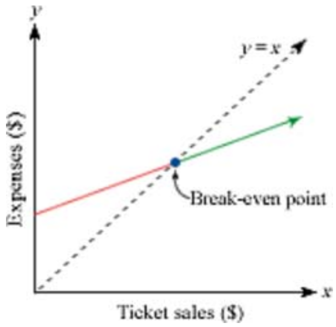


4. Explain how you can use the squares to adjust your moveable line to a better fit. Based on your explanation, how do you think the least-squares line got its name?

5. When you experimented with all three lines of fit in Step 8, did you get the same equation for all three? Give some reasons why this may or may not have happened.

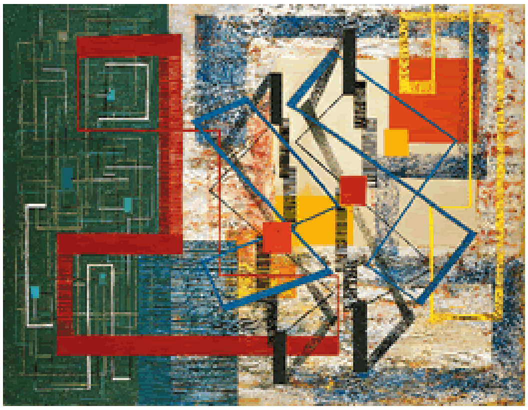
LESSON 3.6

Linear Systems



The number of tickets sold for a school activity, like a spaghetti dinner, helps determine the financial success of the event. Income from ticket sales can be less than, equal to, or greater than expenses. The break-even value is the intersection of the expense function and the line $y = x$. This equation models all the situations where the expenses, y, are equal to income, x. In this lesson, you will focus on mathematical situations involving two or more equations or conditions that must be satisfied at the same time. A set of two or more equations that share the same variables and are solved or studied simultaneously is called a **system of equations**.

American painter I. Rice Pereira (1907-1971) explored light and space in her works, characterized by intersecting lines in mazelike patterns.
This piece is titled *Green Mass* (1950).



EXAMPLE A

Minh and Daniel are starting a business together, and they need to decide between long-distance phone carriers. One company offers the Phrequent Phoner Plan, which costs 20¢ for the first minute of a phone call and 17¢ for each minute after that. A competing company offers the Small Business Plan, which costs 50¢ for the first minute and 11¢ for each additional minute. Which plan should Minh and Daniel choose?

Solution

The plan they choose depends on the nature of their business and the most likely length and frequency of their phone calls. Because the Phrequent Phoner Plan (PPP) costs less for the first minute, it is obviously better for very short calls. However, the Small Business Plan (SBP) probably will be cheaper for longer calls because the cost is less for additional minutes. There is a phone conversation length for which both plans cost the same. To find it, let x represent the call length in minutes, and let y represent the cost in cents.

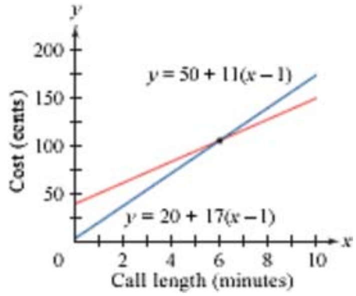
A cost equation that models the Phrequent Phoner Plan is

$$y = 20 + 17(x - 1)$$

A cost equation for the Small Business Plan is

$$y = 50 + 11(x - 1)$$

A graph of these equations shows the Phrequent Phoner Plan cost is below the Small Business Plan cost until the lines intersect. You may be able to estimate the coordinates of this point from the graph.



Or you can look in a table to find an answer. The point of intersection at (6, 105) tells you that a six-minute call will cost \$1.05 with either plan.

If Minh and Daniel believe their average business call will last less than six minutes, they should choose the Phrequent Phoner Plan. But if they think most of their calls will last more than six minutes, the Small Business Plan is the better option.

You can also use equations to solve for the point of intersection. What equations could you write for Example A?

Investigation

Population Trends

The table below gives the populations of San Jose, California, and Detroit, Michigan.

Populations

Year	1950	1960	1970	1980	1990
San Jose	95,280	204,196	459,913	629,400	782,224
Detroit	1,849,568	1,670,144	1,514,063	1,203,368	1,027,974

(*The World Almanac and Book of Facts 2001*)

Step 1 If the trends continue, when will San Jose be as large as Detroit? What will the two populations be at that time?

Step 2 Show the method you used to make this prediction. Choose a different method to check your answer. Discuss the pros and cons of each method.

In algebra, you studied different methods for finding the exact coordinates of an intersection point by solving systems of equations. One method is illustrated in the next example.

EXAMPLE B

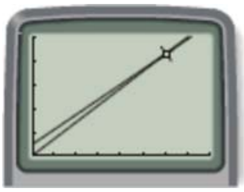
Justine and her little brother Evan are running a race. Because Evan is younger, Justine gives him a 50-foot head start. Evan runs at 12.5 feet per second and Justine runs at 14.3 feet per second. How far will they be from Justine's starting line before Justine passes Evan? What distance should Justine mark for a close race?

Solution

You can compare the time and distance that each person runs. Because distance, *d*, depends on time, *t*, you can write these equations:

$d = rt$	Distance equals the rate, or speed, times time.
$d = 14.3t$	Justine's distance equation.
$d = 50 + 12.5t$	Evan's distance equation.

Graphing these two equations shows that Justine eventually catches up to Evan and passes him, if the race is long enough. At that moment, they are at the same distance from the start, at the same time. You can estimate this point from the graph or scroll down until you find the answer in the table. You can also solve the system of equations.



[0, 40, 5, 0, 500, 100]

Because both equations represent distances and you want to know when those distances are equal, you can set the equations equal to each other and solve for the time, *t*, when the distances are equal.

$14.3t = 50 + 12.5t$	The right side of Justine's equation equals the right side of Evan's equation, because they are equal to the same distance, *d*.
$1.8t = 50$	Subtract 12.5*t* from both sides.
$t = \frac{50}{1.8} \approx 27.8$	Divide both sides by 1.8.

So, Justine passes Evan after 27.8 seconds. Now you can substitute this value back into either equation to find their distances from the starting line when Justine passes Evan.

$$d = 14.3t = 14.3 \cdot \frac{50}{1.8} \approx 397.2$$

If Justine marks a 400 ft distance, she will win, but it will be a close race.

The method of solving a system demonstrated in Example B uses one form of **substitution**. In this case you substituted one expression for distance in place of the distance, *d*, in the other equation, resulting in an equation with only one variable, *t*. When you have the two equations written in intercept form, substitution is a straightforward method for finding an exact solution. The solution to a system of equations with two variables is a pair of values that satisfies both equations. Sometimes a system will have many solutions or no solution.

EXERCISES

Practice Your Skills

1. Use a table to find the point of intersection for each pair of linear equations.

a. $\begin{cases} y = 3x - 17 \\ y = -2x - 8 \end{cases}$ **b.** $\begin{cases} y = 28 - 3(x - 5) \\ y = 6 + 7x \end{cases}$

2. Write a system of equations that has (2, 7.5) as its solution.

3. Write the equation of the line perpendicular to $y = 4 - 2.5x$ and passing through the point (1, 5).

4. Solve each equation.

a. $4 - 2.5(x - 6) = 3 + 7x$ **b.** $11.5 + 4.1t = 6 + 3.2(t - 4)$

5. Use substitution to find the point (x, y) where each pair of lines intersect. Use a graph or table to verify your answer.

a. $\begin{cases} y = -2 + 3(x - 7) \\ y = 10 - 5x \end{cases}$ **b.** $\begin{cases} y = 0.23x + 9 \\ y = 4 - 1.35x \end{cases}$ **c.** $\begin{cases} y = -1.5x + 7 \\ 2y = -3x + 14 \end{cases}$

Reason and Apply

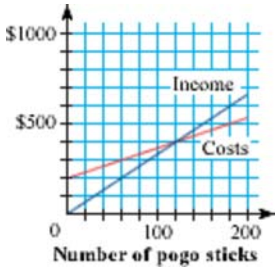
6. The equations $s_1 = 18 + 0.4m$ and $s_2 = 11.2 + 0.54m$ give the lengths of two different springs in centimeters, s_1 and s_2, as mass amounts in grams, m, are separately added to each.

a. When are the springs the same length?

b. When is one spring at least 10 cm longer than the other?

c. Write a statement comparing the two springs.

7. **APPLICATION** This graph shows the Kangaroo Company's production costs and revenue for its pogo sticks. Use the graph to estimate the answers to the questions below.



a. If 25 pogo sticks are sold, will the company earn a profit? Describe how you can use the graph to answer this question.

b. If the company sells 200 pogo sticks, will it earn a profit? If so, approximately how much?

c. How many pogo sticks must the company sell to break even? How do you know?

8. **APPLICATION** Winning times for men and women in the 1500 m Olympic speed skating event are given below, in minutes and seconds.

1500 m Olympic Speed Skating

Year	1964	1968	1972	1976	1980	1984	1988	1992	1994	1998
Men	2:10.3	2:03.4	2:02.96	1:59.38	1:55.44	1:58.36	1:52.06	1:54.81	1:51.29	1:47.87
Women	2:22.6	2:22.4	2:20.85	2:16.58	2:10.95	2:03.42	2:00.68	2:05.87	2:02.19	1:57.58

(*The World Almanac and Book of Facts 2001*)

a. Analyze the data and predict when the winning times for men and women will be the same if the current trends continue.

b. How reasonable do you think your prediction is? Explain your reasoning.

c. Predict the winning times for the 2002 Winter Olympics. How close are your predictions to the actual results?

d. Is it appropriate to use a linear model for these data? Why?

Derek Parra of the United States won the 2002 gold medal for the men's 1500 m speed skating event.

9. APPLICATION Suppose the long-distance phone companies in Example A calculate their charges so that a call of exactly 3 min will cost the same as a call of 3.25 min or 3.9 min, and there is no increase in cost until you have been connected for 4 min. Increases are calculated after each additional minute. A function that models this situation is the **greatest integer function**, $y = [x]$, which outputs the greatest integer less than or equal to x. [See Calculator Note 3F for instructions on using the greatest integer function on your calculator.]

a. Use the greatest integer function to write cost equations for the two companies in Example A.

b. Graph the two new equations representing the Phrequent Phoner Plan and the Small Business Plan.

c. Now determine when each plan is more desirable. Explain your reasoning.

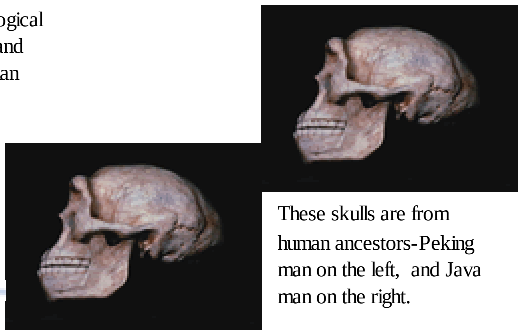
10. APPLICATION At an excavation site, anthropologists use various clues to draw conclusions about people and populations based on skeletal remains. For instance, when a partial skeleton is found, an anthropologist can use the lengths of certain bones to estimate the height of the living person. The humerus bone is the single large bone that extends from the elbow to the shoulder socket. The following formulas, attributed to the work of Mildred Trotter and G. C. Gleser, have been used to estimate a male's height, m, or a female's height, f, when the length, h, of the humerus bone is known: $m = 3.08h + 70.45$ and $f = 3.36h + 58.0$. All measurements are in centimeters.

a. Graph the two lines on the same set of axes.

b. If a humerus bone is found and it measures 42 cm, how tall would the person be according to the model if the bone was determined to come from a male? From a female?

Science CONNECTION

Physical anthropology is a science that deals with the biological evolution of human beings, the study of human ancestors and non-human primates, and the in-depth analysis of the human skeleton. By studying bones and bone fragments, physical anthropologists have developed methods that can provide a wealth of information on the age, sex, ancestry, height, and diet of a person who lived in ancient times, just by studying his or her skeleton.



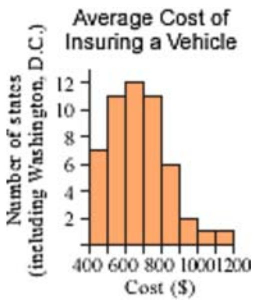
These skulls are from human ancestors-Peking man on the left, and Java man on the right.

11. Write a system of equations to model each situation, and solve for the values of the appropriate variables.
  a. The perimeter of a rectangle is 44 cm. Its length is 2 cm more than twice its width.
  b. The perimeter of an isosceles triangle is 40 cm. The base length is 2 cm less than the length of a leg of the triangle.
  c. The Fahrenheit reading on a dual thermometer is 0.4 degrees less than three times the Celsius reading. (Hint: Your second equation needs to be a conversion formula between degrees Fahrenheit and degrees Celsius.)

Review

12. Use the model $\hat{y} = 343.5 + 1.55(x - 1984)$ that was found in the example in Lesson 3.3. Recall that x represents the year, and y represents the concentration of CO_2 in parts per million (ppm) at Mauna Loa.
  a. Predict the concentration of CO_2 in the year 1984.
  b. Use the model to predict the concentration of CO_2 in the year 2010.
  c. According to this model, when will the level of CO_2 be double the preindustrial level of 280 ppm?

13. The histogram shows the average annual cost of insuring a motor vehicle in the United States.
  a. How many jurisdictions are included in the histogram?
  b. Mississippi is the median jurisdiction. Mississippi is in what bin?
  c. In what percentage of the jurisdictions is the average cost less than $600?

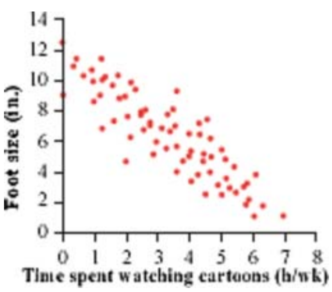


14. Solve these equations for y.
  a. $3x - 8y = 12$
  b. $5x + 2y = 12$
  c. $-3x + 4y = 5$

IMPROVING YOUR REASONING SKILLS

Cartoon Watching Causes Small Feet

Lisa did a study for her health class about the effects of cartoon watching on foot size. Based on a graph of her data, she finds that there was an inverse relationship between foot size and hours spent watching cartoons per week. She concludes that "cartoon watching causes small feet." Is this true? Explain any flaws in Lisa's reasoning.



LESSON 3.7

Substitution and Elimination

A place for everything and everything in its place.

ENGLISH PROVERB

A solution to a system of equations in two variables is a pair of values that satisfies both equations and represents the intersection of their graphs. In Lesson 3.6, you reviewed solving a system of equations using substitution, when both equations are in intercept form. Suppose you want to solve a system and one or both of the equations are not in intercept form. You can rearrange them into intercept form, or you can use a different method.

If one equation is in intercept form, you can still use substitution.

In *Modern Warrior Series, Shirt #1* (1998), northern Cheyenne artist Bently Spang (b 1960) reflects on identity. He presents opposing influences as an equivalence of forms -the modern and the traditional, the spiritual and the mundane, the Cheyenne and the non-Cheyenne -all intersecting in an arrangement of family photographs.

EXAMPLE A

Solve this system for x and y.

$$\begin{cases} y = 15 + 8x \\ -10x - 5y = -30 \end{cases}$$

▸Solution

You can solve the second equation for y so that both equations will be in intercept form and substitute the right side of one equation for y in the other equation. Or you can simply substitute the right side of the first equation for y in the second equation.

$-10x - 5y = -30$	Original form of the second equation.
$-10x - 5(15 + 8x) = -30$	Substitute the right side of the first equation for y.
$-10x - 75 - 40x = -30$	Distribute –5.
$-50x = 45$	Add 75 to both sides and combine like terms.
$x = -0.9$	Divide both sides by –50.

Now that you know the value of x, you can substitute it in either equation to find the value of y.

$y = 15 + 8(-0.9)$	Substitute –0.9 for x in the first equation.
$y = 7.8$	Multiply and combine like terms.

Write your solution as an ordered pair. The solution to this system is $(-0.9, 7.8)$.

The substitution method relies on the substitution property, which says that if $a = b$, then a may be replaced by b in an algebraic expression. Substitution is a powerful mathematical tool that allows you to rewrite expressions and equations in forms that are easier to use and solve. Notice that substituting an expression for y, as you did in Example A, eliminates y from the equation, allowing you to solve a single equation for a single variable, x.

A third method for solving a system of equations is the **elimination** method. The elimination method uses the addition property of equality, which says that if $a = b$, then $a + c = b + c$. In other words, if you add equal quantities to both sides of an equation, the equation is still true. If necessary, you can also use the multiplication property of equality, which says that if $a = b$, then $ac = bc$, or if you multiply both sides of an equation by equal quantities, then the equation is still true.

EXAMPLE B

Solve these systems for x and y.

a. $\begin{cases} 4x + 3y = 14 \\ 3x - 3y = 13 \end{cases}$ **b.** $\begin{cases} -3x + 5y = 6 \\ 2x + y = 6 \end{cases}$

Solution

Because neither of these equations is in intercept form, it is probably easier to solve the systems using the elimination method.

a. You can solve the system without changing either equation to intercept form by adding the two equations.

$4x + 3y = 14$	Original equations.
$\underline{3x - 3y = 13}$	Add equal quantities to both sides of the equation.
$7x = 27$	The variable y is eliminated.
$x = \frac{27}{7}$	Solve for x.
$4\left(\frac{27}{7}\right) + 3y = 14$	Substitute $\frac{27}{7}$ for x back into either equation.
$y = -\frac{10}{21}$	Solve for y.

The solution to this system is $\left(\frac{27}{7}, -\frac{10}{21}\right)$. You can substitute the coordinates back into both equations to check that the point is a solution for both.

$$4\left(\frac{27}{7}\right) + 3\left(-\frac{10}{21}\right) = \frac{108}{7} - \frac{30}{21} = \frac{294}{21} = 14$$

$$3\left(\frac{27}{7}\right) - 3\left(-\frac{10}{21}\right) = \frac{81}{7} + \frac{30}{21} = \frac{273}{21} = 13$$

b. Adding the equations as they are written will not eliminate either of the variables. You need to multiply one or both equations by some value so that if you add the equations together, one of the variables will be eliminated.

The easiest choice is to multiply the second equation by –5, and then add it to the first equation.

$-3x + 5y = 6$	$\rightarrow -3x + 5y = 6$	Original form of the first equation.
$-5(2x + y) = -5(6)$	$\rightarrow \underline{-10x - 5y = -30}$	Multiply both sides of the second equation by –5.
	$-13x = -24$	Add the equations.

This eliminates the y-variable and gives $x = \frac{24}{13}$. Substituting this x-value back into either of the original equations gives the y-value. Or you can use the same process to eliminate the x-variable.

$$\begin{array}{rcrl} -3x + 5y = 6 & \rightarrow & -6x + 10y = 12 & \text{Multiply both sides by 2.} \\ 2x + \ \ y = 6 & \rightarrow & 6x + \ \ 3y = 18 & \text{Multiply both sides by 3.} \\ \hline & & 13y = 30 & \\ & & y = \frac{30}{13} & \end{array}$$

The solution to this system is $\left(\frac{24}{13}, \frac{30}{13}\right)$. You can use your calculator to verify the solution.

It would take a lot of effort to solve this last system using a table on your calculator. If you had used the substitution method to solve the systems in Example B, you would have had to work with fractions to get an accurate answer. To solve these systems, the easiest method to use is the elimination method.

Investigation

It All Adds Up

In this investigation you may discover some interesting things that happen when you multiply both sides of an equation by the same number or add equations. Work with a partner and follow the steps below. [You can also use the program in Calculator Note 3G to do this investigation.]

Step 1 Graph each of these equations on the same coordinate axes. Where do the lines appear to intersect?

$$\begin{cases} 7x + 2y = -3 & \text{(Equation1)} \\ 3x + 4y = 5 & \text{(Equation2)} \end{cases}$$

Step 2 One partner needs to select any number, M, and the other partner a different number, N. The first partner multiplies Equation 1 by the number M he or she chose while the other partner multiplies Equation 2 by the number N he or she chose.

Step 3 Add the two new equations to form Equation 3. Graph Equation 3 on the same coordinate axis as the two original lines, and describe the location of this new line in relation to the original lines.

Step 4 Next, one partner multiplies the original Equation 1 by $M = -3$ while the other partner multiplies the original Equation 2 by $N = 7$.

Step 5 Add the two new equations from Step 4 to form Equation 4. Graph Equation 4 and describe the location of it in relation to the two original lines.

Step 6 How does Equation 4 differ from the other equations?

Step 7 Repeat this process with different systems of equations, and make a conjecture based on your observations. Summarize your findings.

The elimination method uses a combination technique to eliminate one of the variables in a system of equations. Solving for the variable that remains gives you the *x*- or *y*-coordinate of the point of intersection, if there is a point of intersection.

EXERCISES

Practice Your Skills

1. Solve each equation for the specified variable.

a. $w - r = 11$, for w

b. $2p + 3h = 18$, for h

c. $w - r = 11$, for r

d. $2p + 3h = 18$, for p

2. Multiply both sides of each equation by the given value. What is the relationship between the graphs of the new equation and the original equation?

a. $j + 5k = 8$, by -3

b. $2p + 3h = 18$, by 5

c. $6f - 4g = 22$, by 0.5

d. $\frac{5}{6}a + \frac{3}{4}b = \frac{7}{2}$, by 12

3. Add each pair of equations. What is the relationship between the graphs of the new equation and the original pair?

a. $\begin{cases} 3x - 4y = 7 \\ 2x + 2y = 5 \end{cases}$

b. $\begin{cases} 5x - 7y = 3 \\ -5x + 3y = 5 \end{cases}$

4. Graph each system and find an approximate solution. Then choose a method and find the exact solution. List each solution as an ordered pair.

a. $\begin{cases} y = 3.2x + 44.61 \\ y = -5.1x + 5.60 \end{cases}$

b. $\begin{cases} y = \frac{2}{3}x - 3 \\ y = -\frac{5}{6}x + 7 \end{cases}$

c. $\begin{cases} y = 4.7x + 25.1 \\ 3.1x + 2y = 8.2 \end{cases}$

d. $\begin{cases} -6x - 7y = 20 \\ -5x + 4y = -5 \end{cases}$

e. $\begin{cases} 2.1x + 3.6y = 7 \\ -6.3x + y = 8.2 \end{cases}$

5. Solve each system of equations.

a. $\begin{cases} 5.2x + 3.6y = 7 \\ -5.2x + 2y = 8.2 \end{cases}$

b. $\begin{cases} \frac{1}{4}x - \frac{2}{5}y = 3 \\ \frac{3}{8}x + \frac{2}{5}y = 2 \end{cases}$

c. $\begin{cases} 4x + 9y = 12 \\ 3x - 8y = 10 \end{cases}$

d. $\begin{cases} s = 7 - 3n \\ 7n + 2s = 40 \end{cases}$

e. $\begin{cases} f = 3d + 5 \\ 10d - 4f = 16 \end{cases}$

f. $\begin{cases} \frac{1}{4}x - \frac{4}{5}y = 7 \\ \frac{3}{4}x + \frac{2}{5}y = 2 \end{cases}$

Keymath.com More Practice Your Skills

g. $\begin{cases} 2x + 3y = 4 \\ 1.2x + 1.8y = 2.6 \end{cases}$

Reason and Apply

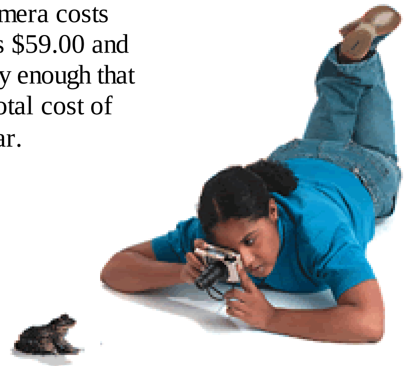
6. Solve each problem.

a. If $4x + y = 6$, then what is $(4x + y - 3)^2$?

b. If $4x + 3y = 14$ and $3x - 3y = 13$, what is $7x$?

7. The formula to convert between Fahrenheit and Celsius is $C = \frac{5}{9}(F - 32)$. What reading on the Fahrenheit scale is three times the equivalent temperature on the Celsius scale?

8. **APPLICATION** Ellen must decide between two cameras. The first camera costs $47.00 and uses two alkaline AA batteries. The second camera costs $59.00 and uses one $4.95 lithium battery. She plans to use the camera frequently enough that she probably would replace the AA batteries six times a year for a total cost of $11.50 per year. The lithium battery, however, will last an entire year.



a. Let x represent the number of years, and let y represent the cost in dollars. Write an equation to represent the overall expense for each camera.

b. Which camera is less expensive in the short term? In the long term? How long will it take until the overall cost of the less expensive camera is equal to the overall cost of the other camera?

c. Carefully describe three different ways to verify your solution.

9. Write a system of two equations that has a solution of $(-1.4, 3.6)$.

10. The two sequences below have one term that is the same. Determine which term this is and find its value.

$u_1 = 12$

$u_n = u_{n-1} + 0.3 \quad \text{where } n \geq 2$

$v_1 = 15$

$v_n = v_{n-1} + 0.2 \quad \text{where } n \geq 2$

11. Formulas play an important part in many fields of mathematics and science. You can create a new formula using substitution to combine formulas.

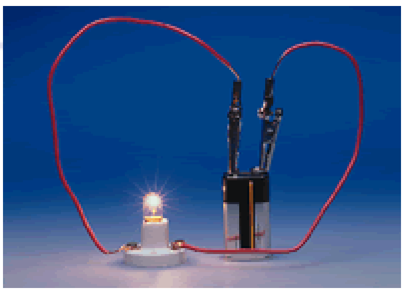
a. Using the formulas $A = s^2$ and $d = s\sqrt{2}$, write a formula for A in terms of d.

b. Using the formulas $P = IE$ and $E = IR$, write a formula for P in terms of I and R.

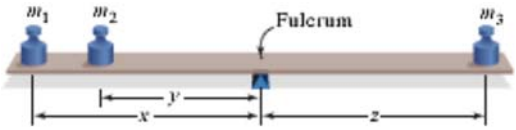
c. Using the formulas $A = \pi r^2$ and $C = 2\pi r$, write a formula for A in terms of C.

Science CONNECTION

In this simple circuit, a 9-volt battery lights a small light bulb. The power P supplied to the bulb in watts is a product of the voltage E of the battery times the electrical current I in the wire. The electrical current, in turn, can be calculated as the battery voltage E divided by the resistance R of the circuit, which depends on the length and gauge, or diameter, of the wire. It is important to know basic functions and relations when designing a circuit, so that there will be enough power supplied to the components, but not excessive power, which would damage the components.



12. **APPLICATION** A support bar will be in equilibrium (balanced) at the fulcrum, *O*, if $m_1x + m_2y = m_3z$, where m_1, m_2, and m_3 represent masses and x, y, and z represent the distance of the masses to the fulcrum. Draw a diagram for each question and calculate the answer.

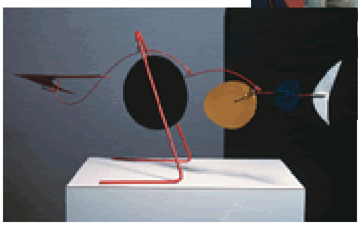


a. A 40 in. bar is in equilibrium when a weight of 6 lb is hung from one end and a weight of 9 lb is hung from the other end. Find the position of the fulcrum.

b. While in the park, Michael and his two sons, Justin and Alden, go on a 16 ft seesaw. Michael, who weighs 150 lb, sits at the edge of one end while Justin and Alden move to the other side and try to balance. The seesaw balances with Justin at the other edge and Alden 3 ft from him. After some additional experimentation the see saw balances once again with Alden at the edge and Justin 5.6 ft from the fulcrum. How much does each boy weigh?

Art CONNECTION

Alexander Calder (1898-1976) was an artistic pioneer who revolutionized the mobile as an art form. After getting a degree in mechanical engineering, he went to art school, supporting himself through school by working as an illustrator. Once out of school, he began creating small three-dimensional sculptures made from wire, wood, and cloth that balanced perfectly, whether or not they were symmetrical. Eventually, he designed sculptures with painted elements that moved mechanically, and then went on to produce pieces that moved with the air. These free-moving, hanging sculptures became known as "mobiles." He also designed "stabiles," essentially mobiles that balance on a fixed support.



Alexander Calder builds a mobile in this studio. At left is one of his stabiles, *Boomerang and Sickle Moon.*

Review

13. Classify each statement as true or false. If the statement is false, change the right side to make it true.

a. $x^2 + 8x + 15 = (x + 3)(x + 5)$

b. $x^2 - 16 = (x - 4)(x - 4)$

14. Consider the equation $3x + 2y - 7 = 0$

a. Solve the equation for y.

b. Graph this equation.

c. What is the slope?

d. What is the y-intercept?

e. Write an equation for a line perpendicular to this one and having the same y-intercept. Graph this equation.

15. **APPLICATION** This table gives the mean price for a gallon of gasoline in the United States from 1950 through 2000.

Year	1950	1960	1970	1980	1990	2000
Price ($)	0.27	0.31	0.36	1.19	1.15	1.56

(*The New York Times Almanac 2002*)

a. Make a scatter plot of the data. Let x represent the year, and y represent the price in dollars.

b. Find the median-median line of the data.

c. Assuming that the same trend continues, predict the price of gas in 2010. From what you know (or can find out) about current world fossil fuel supplies, is the pattern likely to continue? Explain.

16. **APPLICATION** This table shows the normal monthly precipitation in inches for Pittsburgh, Pennsylvania, and Portland, Oregon.

Month	J	F	M	A	M	J	J	A	S	O	N	D
Pittsburgh	2.5	2.4	3.4	3.1	3.6	3.7	3.8	3.2	3.0	2.4	2.9	2.9
Portland	5.4	3.9	3.6	2.4	2.1	1.5	0.7	1.1	1.8	2.7	5.3	6.1

(*The New York Times Almanac 2002*)

a. Display the data in two box plots on the same axis.

b. Give the five-number summary of each data set.

c. Describe the differences in living conditions with respect to precipitation.

d. Which city generally has more rain annually?

17. Consider these three sequences.

i. 243, –324, 432, –576, . . .

ii. 22, 26, 31, 37, 44, . . .

iii. 24, 25.75, 27.5, 29.25, 31, . . .

a. Find the next two terms in each sequence.

b. Identify each sequence as arithmetic, geometric, or other.

c. If a sequence is arithmetic or geometric, write a recursive routine to generate the sequence.

d. If a sequence is arithmetic, give an explicit formula that generates the sequence.

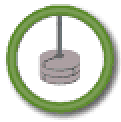
18. Melina and Angus are designing a fixed speed hot-rod car for a remote-control rally. They need to construct a car that travels at a constant 0.50 m/s. In order to qualify for the rally they must show time trial results with a root mean square error of less than 0.05 m. This table shows their time trial results. Will Melina and Angus qualify or do they need to go back to the drawing board?

Time (s)	1	2	3	4	5	6	7	8
Distance (m)	0.48	0.95	1.6	2.0	2.52	2.93	3.49	4.05

CHAPTER 3 REVIEW

Keymath.com Links to Resources

In this chapter you analyzed sets of two-variable data. A plot of two-variable data may be linear, or it may appear nearly linear over a short domain. If it is linear, you can find a **line of fit** to model the data. You can write a **linear equation** for this line and use it to **interpolate** or **extrapolate** points for which data are not available. You can also solve a **system of equations** to find the intersection of two lines, using the methods of **substitution**, **elimination**, or using a graph or table of values.



To write the equation of a line, you can use its slope and y-intercept to write the equation in **intercept form**, or you can use the coordinates of two points to write the equation in **point-slope form**.

You learned two methods for finding a line of fit for a set of data: estimating by looking at the trend of the data and fitting a line based on certain criteria, and using the more systematic **median-median** method. Regardless of the method you use, you can examine the **residuals** to determine whether your model is a good fit. With a good model, the residuals should be randomly positive and negative, the sum of the residuals should be zero, and each residual should be as small as possible. You also learned about **root mean square error**, a single measure of good fit.

EXERCISES

1. Find the slope of the line containing the points (16, 1300) and (–22, 3250).

2. Consider the line $y = -5.02 + 23.45x$.

a. What is the slope of this line?

b. Write an equation for a line that is parallel to this line.

c. Write an equation for a line that is perpendicular to this line.

3. Find the point on each line where y is equal to 740.0.

a. $y = 16.8x + 405$ **b.** $y = -7.4 + 4.3(x - 3.2)$

4. Consider the system of equations

$$\begin{cases} y = 3.2x - 4 & \text{(Equation 1)} \\ y = 3.1x - 3 & \text{(Equation 2)} \end{cases}$$

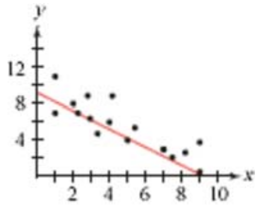
a. Substitute the y-value from Equation 1 into Equation 2 to obtain a new equation. Solve the new equation for x.

b. Subtract Equation 2 from Equation 1 and solve for the remaining variable.

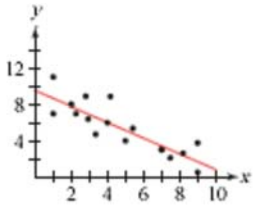
c. Explain why the solutions to 4a and 4b are the same.

5. The graphs below show three different lines of fit for the same set of data. For each graph, decide whether the line is a good line of fit or not, and explain why.

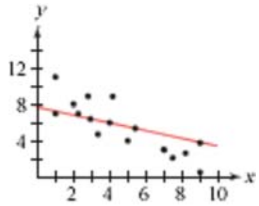
a.



b.



c.



6. Solve each system.

a. $\begin{cases} y = 6.2x + 18.4 \\ y = -2.1x + 7.40 \end{cases}$

b. $\begin{cases} y = \frac{3}{4}x - 1 \\ \frac{7}{10}x + \frac{2}{5}y = 8 \end{cases}$

c. $\begin{cases} 3x + 2y = 4 \\ -3x + 5y = 3 \end{cases}$

7. Find the point (or points) where each pair of lines intersect.

a. $\begin{cases} 5x - 4y = 5 \\ 2x + 10y = 2 \end{cases}$

b. $\begin{cases} y = \frac{1}{4}(x - 8) + 5 \\ y = 0.25x + 3 \end{cases}$

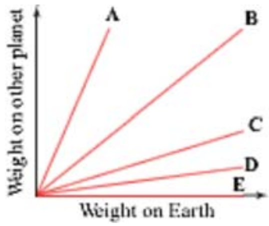
c. $\begin{cases} \frac{3}{5}x - \frac{2}{5}y = 3 \\ 0.6x - 0.4y = -3 \end{cases}$

8. The ratio of the weight of an object on Mercury to its weight on Earth is 0.38.

a. Explain why you can use the equation $m = 0.38e$ to model the weight of an object on Mercury.

b. How much would a 160-pound student weigh on Mercury?

c. The ratios for the Moon and Jupiter are 0.17 and 2.54 respectively. The equations $y_1 = 0.38x$, $y_2 = 1x$, $y_3 = 0.17x$, and $y_4 = 2.54x$ are graphed at right. Match each planet with its graph and equation.



Science CONNECTION

Planning a trip to outer space? This table gives the ratio of an object's weight on each of the planets and the Moon to its weight on Earth.

You can calculate your weight on each of these celestial bodies. Some of these calculations will make you seem like a light-weight, and others will make you seem heavy! But, of course, your body will still be the same. Your *mass* won't change. Your *weight* on other planets depends on your mass, *m*, the planet's mass, *M*, and the distance, *r*, from the center of the planet.

Mercury	0.38
Venus	0.91
Earth	1
Moon	0.17
Mars	0.38
Jupiter	2.54
Saturn	0.93
Uranus	0.80
Neptune	1.19
Pluto	0.06

Architecture CONNECTION

In 1173 C.E., the Tower of Pisa was built on soft ground and, ever since, has been leaning to one side as it sinks in the soil. This 8-story, 191-foot tower was built with only a 7-foot-deep foundation. The tower was completed in the mid-1300s even though it started to lean after the first 3 stories were completed. The tower's structure consists of a cylindrical body, arches, and columns. Rhombuses and rectangles decorate the surface.

The Tower of Pisa, in Pisa, Italy

9. Read the Architecture Connection. The table lists the amount of lean, measured in millimeters, for thirteen different years.

Tower of Pisa

Year	Lean	Year	Lean
1910	5336	1960	5414
1920	5352	1970	5428
1930	5363	1980	5454
1940	5391	1990	5467
1950	5403		

a. Make a scatter plot of the data. Let x represent the year, and let y represent the amount of lean in millimeters.

b. Find a median-median line for the data.

c. What is the slope of the median-median line? Interpret the slope in the context of the problem.

d. Find the amount of lean predicted by your equation for 1992 (the year work was started to secure the foundation).

e. Find the root mean square error of the median-median line. What does this error tell you about your answer to 8d?

f. What are the domain and range for your linear model? Give an explanation for the numbers you chose.

10. The 4th term of an arithmetic sequence is 64. The 54th term is – 61. Find the 23rd term.

MIXED REVIEW

11. State whether each recursive formula defines a sequence that is arithmetic, geometric, shifted geometric, or none of these. State whether a graph of the sequence would be linear or curved. Then list the first 5 terms of the sequence.

a. $u_1 = 4$ and $u_n = 3u_{n-1}$ where $n \geq 2$ **b.** $u_0 = 20$ and $u_n = 2u_{n-1} + 7$ where $n \geq 1$

12. **APPLICATION** You receive a $500 gift for high school graduation and deposit it into a savings account on June 15. The account has an annual interest rate of 5.9% compounded annually.

a. Write a recursive formula for this problem.

b. List the first 3 terms of the sequence.

c. What is the meaning of the value of u_3?

d. How much money will you have in your account when you retire, 35 years later?

e. If you deposit an additional $100 in your account each year on June 15, how much will you have in savings 35 years after graduation?

13. **APPLICATION** An Internet web site gives the current world population and projects the population for future years. Its projections for the number of people on Earth on January 1 in the years 2005 through 2009 are given in the table at right.

Population Projections

Year	World population
2005	6,486,915,022
2006	6,581,251,218
2007	6,676,959,301
2008	6,774,059,223
2009	6,872,843,077

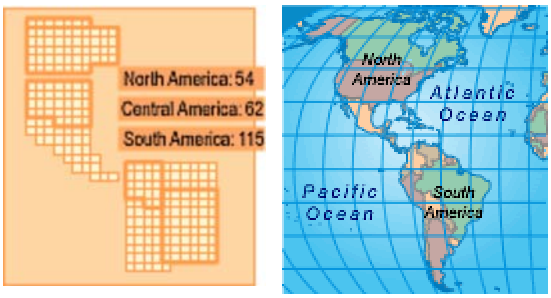
(*www.ibiblio.org*)

a. Find a recursive formula to model the population growth. What kind of sequence is this?

b. Predict the population on January 1, 2010.

c. In what year will the world's population exceed 10 billion?

d. Is this a realistic model that could predict world population in the next millennium? Explain.



The United Nations estimates that by 2025, the human population will increase by 31 percent. This diagram shows projected growth in the Western Hemisphere. Each square represents an increase of 1 million people. The outlined countries are Canada, the United States, Mexico, Colombia, and Brazil.

14. **APPLICATION** Jonah must take an antibiotic every 12 hours. Each pill is 25 milligrams, and after every 12 hours, 50% of the drug remains in his body. What is the amount of antibiotic in his body over the first 2 days? What amount will there be in his body in the long run?

15. Create a box-and-whisker plot that has this five-number summary: 5, 7, 12, 13, 17.

a. Are the data skewed left, skewed right, or symmetric?

b. What is the median of the data?

c. What is the *IQR*?

d. What percentage of data values are above 12? Above 13? Below 5?

16. The table shows high school dropout rates reported by states and the District of Columbia in 1998–1999. Data are unavailable for some states.

High School Dropout Rates, 1998–1999

State	Rate (%)
Alabama	4.4
Alaska	5.3
Arizona	8.4
Arkansas	6.0
Connecticut	3.3
Delaware	4.1
District of Columbia	8.2
Georgia	7.4
Idaho	6.9
Illinois	6.5
Iowa	2.5
Kentucky	4.9
Louisiana	10.0

State	Rate (%)
Maine	3.3
Maryland	4.4
Massachusetts	3.6
Minnesota	4.5
Mississippi	5.2
Missouri	4.8
Montana	4.5
Nebraska	4.2
Nevada	7.9
New Jersey	3.1
New Mexico	7.0
North Dakota	2.4
Ohio	3.9

State	Rate (%)
Oklahoma	5.2
Oregon	6.5
Pennsylvania	3.8
Rhode Island	4.5
South Dakota	4.5
Tennessee	4.6
Utah	4.7
Vermont	4.6
Virginia	4.5
West Virginia	4.9
Wisconsin	2.6
Wyoming	5.2

(National Center for Education Statistics)

a. What are the mean, median, mode, and standard deviation of the data?

b. Do any states lie more than 2 standard deviations above or below the mean?

c. Draw a histogram from the data, using an appropriate bin width.

17. Use an appropriate method to solve each system of equations.

a. $\begin{cases} 2.1x - 3y = 4 \\ 5x + 3y = 7 \end{cases}$

b. $\begin{cases} y = \frac{1}{3}x + 5 \\ y = -3x + \frac{1}{2} \end{cases}$

c. $\begin{cases} 3x + 4y = 12 \\ 2x - 6y = 5 \end{cases}$

18. Consider this data on the median age of U.S. women who married for the first time in these years between 1972 and 1990. Approximately 0.08% of Americans get married each year.

a. Create a scatter plot of this data. Do these data seem linear?

b. Find a median-median line for the data.

c. Use your median-median line to predict the median age of women at first marriage in 2000.

d. Use your median-median line to make a prediction for the year 2100. Does your prediction seem reasonable?

e. Calculate the residuals for each data point, and find the root mean square for your equation. What does this tell you about your model?

Women's Median Age at First Marriage

Year	Age	Year	Age
1972	20.5	1982	22.3
1974	20.6	1984	22.8
1976	21.0	1986	23.3
1978	21.4	1988	23.7
1980	21.8	1990	24.0

(Centers for Disease Control and Prevention, National Center for Health Statistics)

19. Consider the arithmetic sequence 6, 13, 20, 27, 34, ... Let u_1 represent the first term.

a. Write a recursive formula that describes this sequence.

b. Write an explicit formula for this sequence.

c. What is the slope of your equation in 19b? What relationship does this have to the arithmetic sequence?

d. Determine the value of the 32nd term. Is it easier to use your formula from 19a or 19b for this?

20. For an arithmetic sequence $u_1 = 12$, and $u_{10} = 52.5$.

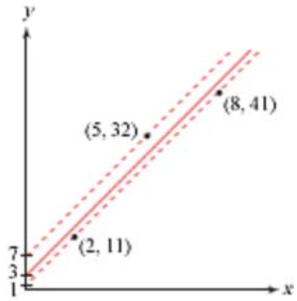
a. What is the common difference of the sequence?

b. Find the equation of the line through the points (1, 12) and (10, 52.5).

c. What is the relationship between 20a and 20b?

TAKE ANOTHER LOOK

1. The three representative points shown here are used to find the two parallel lines and, finally, the median-median line for data points that are not shown.
The median-median line is two-thirds of the vertical distance from the top line to the bottom line.
Find the centroid, or balance point, of the triangle formed by the three representative points. Remember, the centroid is the point $(\bar{x}, \bar{y})$. Will the three representative points always form a triangle?
Write an equation of a line that passes through the centroid and has the same slope as the line through (2, 11) and (8, 41).
Compare this equation with the median-median equation given by your calculator. Describe your results. Make some conjectures based on your observations.



2. The data at right show the average price of a movie ticket for selected years. Find a median-median line for the years 1935–2001. Does your line seem to fit the data well? Which years are not predicted well by your equation? Consider whether or not two or more line segments would fit the data better. Sketch several connected line segments that fit the data. Which model, the single median-median line or the connected segments, do you think is more accurate for predicting the price of a ticket in 2010? Is there another line or curve you might draw that you think would be better? Why do you think these data might not be modeled best by a single linear equation?

Year x	Average ticket price ($) y	Year x	Average ticket price ($) y
1935	0.25	1974	1.89
1940	0.28	1978	2.33
1948	0.38	1982	2.93
1954	0.50	1986	3.70
1958	0.69	1990	4.21
1963	0.87	1994	4.10
1967	1.43	1998	4.68
1970	1.56	2001	5.65

(Motion Picture Association of America)

3. In this chapter you learned three methods for solving a system of linear equations—graphing, substitution, and elimination. These methods also can be applied to systems of nonlinear equations. Use all three methods to solve this system:

$$\begin{cases} y = x^2 - 4 \\ y = -2x^2 + 2 \end{cases}$$

Did you find the same solution(s) with all three methods? Describe how the process of solving this system was different from solving a system of linear equations. If you were given another system similar to this one, which method of solution would you choose? What special things would you look out for?

Assessing What You've Learned

In Chapters 0, 1, and 2, you were introduced to different ways to assess what you learned. Maybe you have tried all six ways—writing in your journal, giving a presentation, organizing your notebook, doing a performance assessment, keeping a portfolio, or writing test items. By now you should realize that assessment is more than just taking tests and more than your teacher giving you a grade.

In the working world, performance in only a few occupations can be measured with tests. All employees, however, must communicate and demonstrate to their employers, coworkers, clients, patients, or customers that they are skilled in their field. Assessing your own understanding and demonstrating your ability to apply what you've learned gives you practice in this important life skill. It also helps you develop good study habits, and that, in turn, will help you advance in school and give you the best possible opportunities in your life.

WRITE IN YOUR JOURNAL Use one of these prompts to write a paragraph in your journal.

- Find an exercise from this chapter that you could not fully solve. Write out the problem and as much of the solution as possible. Then clearly explain what is keeping you from solving the problem. Be as specific as you can.
- Compare and contrast arithmetic sequences and linear equations. How do you decide which to use?

PERFORMANCE ASSESSMENT Show a classmate, family member, or teacher different ways to find a line of fit for a data set. You may want to go back and use one of the data sets presented in an example or exercise, or you may want to research your own data. Discuss how well the line fits and whether you think a linear model is a good choice for the data.