

## 2-5 Scatter Plots and Lines of Regression

1. **OCEANS** The table shows the temperature in the ocean at various depths.

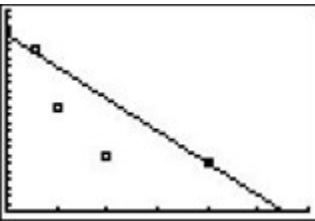
Depth (in meters)	0	300	500	1000	2000	2500
Temp (°C)	22	20	13	7	6	?

Source: NOAA

- Make a scatter plot and a line of fit, and describe the correlation.
- Use two ordered pairs to write a prediction equation.
- Use your prediction equation to predict the missing value.

**ANSWER:**

**a.**



[0, 3000] scl 500 by [0, 25] scl 1

**b.** Sample answer using (0, 22) and (2000, 6):  $y = -0.008x + 22$

**c.** 2°C

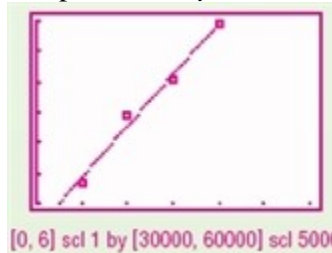
2. **CCSS TOOLS** The table shows the median income of families in North Carolina by family size in a recent year. Use a graphing calculator to make a scatter plot of the data. Find an equation for and graph a line of regression. Then use the equation to predict the median income of a North Carolina family of 9.

Family Size	Income (\$)
1	33,265
2	44,625
3	50,528
4	59,481

Source: U.S. Department of Justice

**ANSWER:**

Sample answer:  $y = 8455.1x + 25,837$ ; \$101,932;



## 2-5 Scatter Plots and Lines of Regression

For Exercises 3–6, complete parts a–c.

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

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

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

3. **COMPACT DISC SALES** The table shows the number of CDs sold in recent years at Jerome's House of Music. Let  $x$  be the number of years since 2000.

Year	2004	2005	2006	2007	2008	2017
Number of CDs sold	49,300	47,280	43,450	40,125	35,792	?

ANSWER:

a.



strong negative correlation

b. Sample answer, using (4, 49,300) and (8, 35,792):

$$y = -3377x + 62808$$

c. Sample answer: 5399 CDs

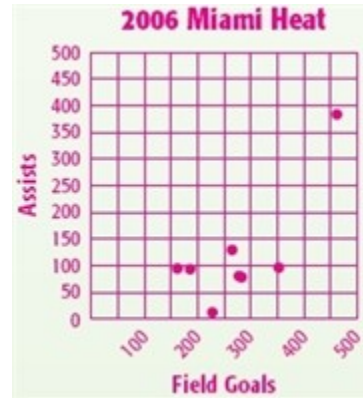
4. **BASKETBALL** The table shows the number of field goals and assists for some of the members of the Miami Heat in a recent NBA season.

Field Goals	472	353	278	283	238	265	186	162	144
Assists	394	97	81	79	18	130	94	95	?

Source: NBA

ANSWER:

a.



no correlation

b. Sample answer: No equation can be written because there is no correlation

c. unpredictable

## 2-5 Scatter Plots and Lines of Regression

5. **ICE CREAM** The table shows the amount of ice cream Sunee's Homemade Ice Creams sold for eight months. Let  $x = 1$  for January.

Month	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept
Gallons sold	37	44	72	80	105	110	119	131	?

ANSWER:

a.



strong positive correlation

b. Sample answer using (1,37) and (8, 131):

$$y = \frac{94}{7}x + \frac{165}{7}$$

c. about 144 gallons

6. **DRAMA CLUB** The table shows the total revenue of all of Central High School's plays in recent school years. Let  $x$  be the number of years since 2003.

School Year	2002	2003	2004	2005	2006	2007
Revenue (\$)	603	666	643	721	771	?

ANSWER:

a.



weak positive correlation

b. Sample answer using (2, 603) and (6, 771):  $y = 42x + 519$

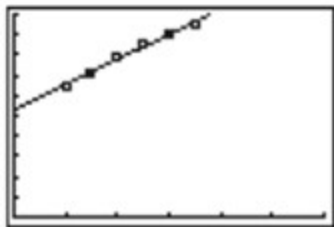
c. Sample answer: \$1065

## 2-5 Scatter Plots and Lines of Regression

7. **SALES** The table shows the sales of Chayton's Computers. Use a graphing calculator to make a scatter plot of the data. Find an equation for and graph a line of regression. Then use the function to predict the sales in 2018.

Year	Sales (\$ thousands)
2004	640
2005	715
2006	791
2007	852
2008	910
2009	944

ANSWER:



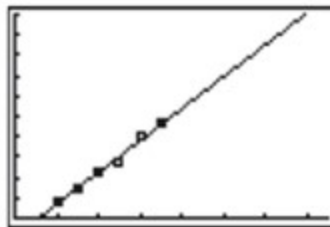
[0,12] scl: 1 by [0, 1000] scl: 100

$61.9x + 530.2$  ( $x$  is the number of years since 2002);  
\$1.52 million in sales.

8. **CCSS TOOLS** The table shows the number of employees of a small company. Use a graphing calculator to make a scatter plot of the data. Find an equation for and graph a line of regression. Then use the function to predict the number of employees in 2025.

Year	Number of Employees
2002	4
2003	7
2004	11
2005	14
2006	20
2007	23

ANSWER:



[0,15] scl: 2 by [0, 50] scl: 5

$y = 3.91x - 4.45$ ; about 93 employees

## 2-5 Scatter Plots and Lines of Regression

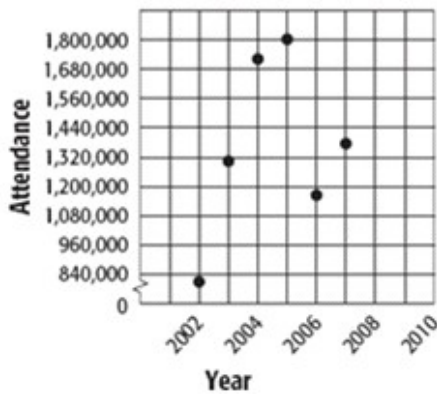
9. **BASEBALL** The table at the right shows the total attendance for the Florida Marlins in some recent years.

- Make a scatter plot of the data.
- Find a regression equation for the data.
- Predict the attendance in 2020.
- How reasonable is your prediction? Explain.

Year	Attendance
2007	1,370,511
2006	1,164,134
2005	1,852,608
2004	1,723,105
2003	1,303,215
2002	813,118

**ANSWER:**

a.



- $y = \$71,406.4x - 141,763,070.9$
- 2,477,857
- Sample answer: Unreasonable; the attendance will not increase without bound because attendance is largely dependent on the team's winning status.

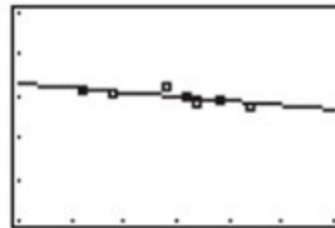
10. **CLASS SIZE** The table at the right shows the relationship between the number of students in a mathematics class and the average grade for each class.

- Make a scatter plot of the data, and find a regression equation for the data. Then sketch a graph of the regression line.
- What is the correlation coefficient of the data?
- Describe the correlation. How accurate is the regression equation?

Class Size	Class Average
16	81.2
19	80.6
24	82.5
26	79.9
27	78.6
29	79.3
32	77.7

**ANSWER:**

a.  $y = -0.21x + 85.1$



[10, 40] scl: 5 by [50, 100] scl: 10

- $r = -0.712$
- Sample answer: relatively accurate with a negative correlation

## 2-5 Scatter Plots and Lines of Regression

11. **CCSS TOOLS** Jocelyn is analyzing the sales of her company. The table at the right shows the total sales for each of six years.

- Find a regression equation and correlation coefficient for the data. Let  $x$  be the years.
- Use the regression equation to predict the sales in 2015.
- Remove the outlier from the data set and find a new regression equation and correlation coefficient.
- Use the new regression equation to predict the sales in 2015.
- Compare the correlation coefficients for the two regression equations. Which function fits the data better? Which prediction should Jocelyn expect to be more accurate?

Year	Sales (\$ millions)
2003	31.2
2004	34.6
2005	18.9
2006	37.7
2007	41.3
2008	45.1

**ANSWER:**

- $y = 3.1x - 6177$ ;  $r = 0.63$
- about \$64.2 million
- $y = 2.6x - 5170$ ;  $r = 0.986$
- about \$62.4 million
- Sample answer: The new equation has a correlation coefficient, 0.986, that is extremely close to 1, so this equation should accurately represent the data.

12. **REASONING** What is the relevance of the correlation coefficient of a linear regression line? Explain your reasoning.

**ANSWER:**

Sample answer: The correlation coefficient is very valuable to a linear regression line because it determines how close the actual data points are to the regression line. The closer the points are to the line, or the closer the correlation coefficient is to 1 or  $-1$ , the more accurate the regression line is.

13. **CHALLENGE** If statements  $a$  and  $b$  have a positive correlation,  $b$  and  $c$  have a negative correlation, and  $c$  and  $d$  have a positive correlation, what can you determine about the correlation between statements  $a$  and  $d$ ? Explain your reasoning.

**ANSWER:**

Sample answer: If  $a$  and  $b$  have a positive correlation, then they are both increasing. If  $b$  and  $c$  have a negative correlation and  $b$  is increasing, then  $c$  must be decreasing. If  $c$  and  $d$  have a positive correlation and  $c$  is decreasing, then  $d$  must be decreasing. If  $a$  is increasing and  $d$  is decreasing, then they must have a negative correlation.

14. **OPEN ENDED** Provide real-world quantities that represent each of the following.

- positive correlation
- negative correlation
- no correlation

**ANSWER:**

- Sample answer: years and height of a teenager
- Sample answer: time and capacity of a standard battery
- Sample answer: a person's weight and his or her income

## 2-5 Scatter Plots and Lines of Regression

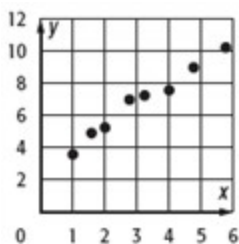
15. **CHALLENGE** Draw a scatter plot for the following data set.

<b>x</b>	1.0	1.5	2.0	2.8	3.2	4.0	4.8	5.8
<b>y</b>	3.5	4.7	5.1	6.8	7.1	7.5	8.8	10.3

Which of the following best represents the correlation coefficient  $r$  for the data? Justify your answer.

- a. 0.99
- b.  $-0.98$
- c. 0.62
- d. 0.08

**ANSWER:**



**a;** Sample answer: The data show a strong positive correlation which means that the correlation coefficient  $r$  should be close to 1.

16. **WRITING IN MATH** What are the strengths and weaknesses of using a regression equation to approximate data?

**ANSWER:**

A strength of using regression equations is that the regression equation can be used to make predictions when the values fall close to the domain of the original data set. A weakness of using regression equations is that they assume that a trend in the original data set will continue, and they are very sensitive to outliers. Both weaknesses can make predictions inaccurate.

17. **SHORT RESPONSE** What is the value of the expression below?

$$17 - 3[-1 + 2(7 - 4)]$$

**ANSWER:**

2

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

**A** 12

**B** 36

**C** 72

**D** 144

**ANSWER:**

C

19. **GEOMETRY** Which is always true?

**F** A parallelogram is a square.

**G** A parallelogram is a rectangle.

**H** A quadrilateral is a trapezoid.

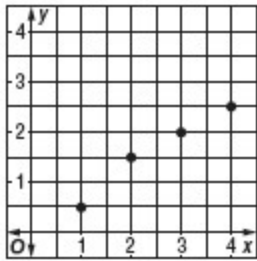
**J** A square is a rectangle.

**ANSWER:**

J

## 2-5 Scatter Plots and Lines of Regression

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

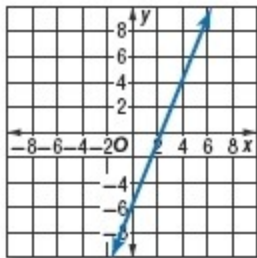


- A  $y = x$
- B  $y = -0.5x + 4$
- C  $y = -0.5x - 4$
- D  $y = 0.5x + 0.5$
- E  $y = 1.5x - 1.5$

ANSWER:

D

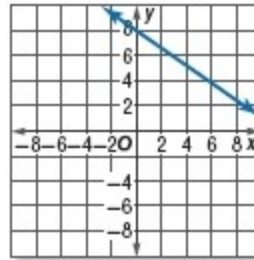
Write an equation in slope-intercept form for each graph.



21.

ANSWER:

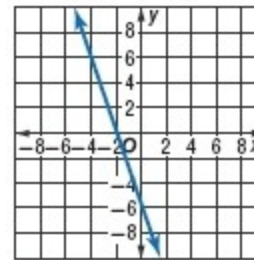
$$y = 2.5x - 6$$



22.

ANSWER:

$$y = -\frac{2}{3}x + 8$$



23.

ANSWER:

$$y = -3x - 6$$

Find the rate of change for each set of data.

24.

Time (day)	3	6	9	12	15
Height (mm)	12	24	36	48	60

ANSWER:

4 mm/day

25.

Time (h)	2	4	6	8
Distance (mi)	35	70	105	140

ANSWER:

17.5 mi/hr



## 2-5 Scatter Plots and Lines of Regression

26. 

Time (sec)	12	16	20	24	28
Volume (cm <sup>3</sup> )	45	60	75	90	105

ANSWER:

$$3.75 \text{ cm}^3/\text{sec}$$

27. 

Force (N)	32	40	48	56	64
Work (J)	48	60	72	84	96

ANSWER:

$$1.5 \text{ J/N}$$

28. **RECREATION** Ramona estimates that she will need 50 tennis balls for every player that signs up for the tennis club and at least 150 more just in case. Write an inequality to express the situation.

ANSWER:

$$t \geq 50p + 150$$

29. **DODGEBALL** Six teams played in a dodge ball tournament. In how many ways can the top three teams finish?

ANSWER:

$$120$$

**Solve each equation.**

30.  $-4|x - 2| = -12$

ANSWER:

$$5, -1$$

31.  $|3x + 4| = 21$

ANSWER:

$$\frac{17}{3}, -\frac{25}{3}$$

32.  $2|4x - 1| + 3 = 9$

ANSWER:

$$1, -\frac{1}{2}$$