Name____ Date __________

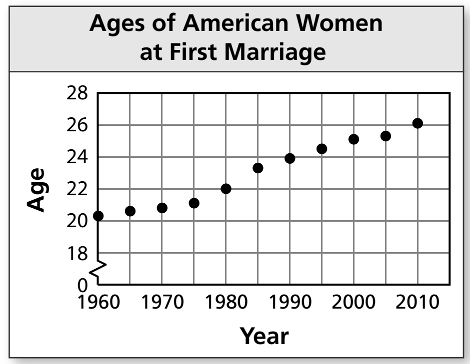
4.5 Analyzing Lines of Fit
For use with Exploration 4.5

Essential Question How can you *analytically* find a line of best fit for a scatter plot?

1 EXPLORATION: Finding a Line of Best Fit

Go to *BigIdeasMath.com* for an interactive tool to investigate this exploration.

Work with a partner.
The scatter plot shows the median ages of American women at their first marriage for selected years from 1960 through 2010. In Exploration 2 in Section 4.4, you approximated a line of fit graphically. To find the line of *best* fit, you can use a computer, spreadsheet, or graphing calculator that has a *linear regression* feature.



a. The data from the scatter plot is shown in the table. Note that 0, 5, 10, and so on represent the numbers of years since 1960. What does the ordered pair (25, 23.3) represent?

L1	L2	L3
0	20.3	
5	20.6	
10	20.8	
15	21.1	
20	22	
25	23.3	
30	23.9	
35	24.5	
40	25.1	
45	25.3	
50	26.1	

L1(55)=

b. Use the *linear regression* feature to find an equation of the line of best fit. You should obtain results such as those shown below.

```
LinReg
 y=ax+b
 a=.1261818182
 b=19.84545455
 r²=.9738676804
 r=.986847344
```

c. Write an equation of the line of best fit. Compare your result with the equation you obtained in Exploration 2 in Section 4.4.

Copyright © Big Ideas Learning, LLC
All rights reserved.

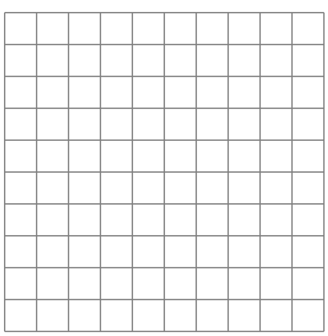
Communicate Your Answer

2. How can you *analytically* find a line of best fit for a scatter plot?

3. The data set relates the number of chirps per second for striped ground crickets and the outside temperature in degrees Fahrenheit. Make a scatter plot of the data. Then find an equation of the line of best fit. Use your result to estimate the outside temperature when there are 19 chirps per second.

Chirps per second	20.0	16.0	19.8	18.4	17.1
Temperature (°F)	88.6	71.6	93.3	84.3	80.6

Chirps per second	14.7	15.4	16.2	15.0	14.4
Temperature (°F)	69.7	69.4	83.3	79.6	76.3



Copyright © Big Ideas Learning, LLC
All rights reserved.

Name____ Date __________

4.5 Practice

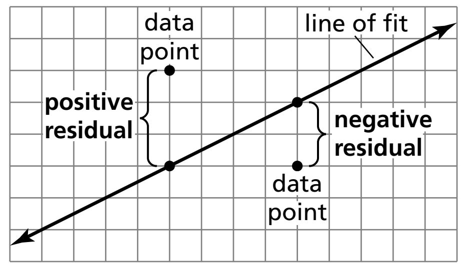
For use after Lesson 4.5

Notes:

Core Concepts

Residuals

A **residual** is the difference of the *y*-value of a data point and the corresponding *y*-value found using the line of fit. A residual can be positive, negative, or zero.



A scatter plot of the residuals shows how well a model fits a data set. If the model is a good fit, then the absolute values of the residuals are relatively small, and the residual points will be more or less evenly dispersed about the horizontal axis. If the model is not a good fit, then the residual points will form some type of pattern that suggests the data are not linear. Wildly scattered residual points suggest that the data might have no correlation.

Notes:

Worked-Out Examples

Example #1

Use residuals to determine whether the model is a good fit for the data in the table. Explain.

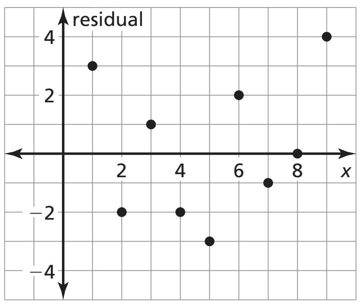
$y = 6x + 4$

x	1	2	3	4	5	6	7	8	9
y	13	14	23	26	31	42	45	52	62

x	y	y-Value from model	Residual
1	13	10	$13 - 10 = 3$
2	14	16	$14 - 16 = -2$
3	23	22	$23 - 22 = 1$
4	26	28	$26 - 28 = -2$
5	31	34	$31 - 34 = -3$
6	42	40	$42 - 40 = 2$
7	45	46	$45 - 46 = -1$
8	52	52	$52 - 52 = 0$
9	62	58	$62 - 58 = 4$

Copyright © Big Ideas Learning, LLC
All rights reserved.

4.5 Practice (continued)



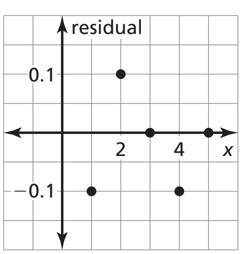
The points are evenly dispersed about the horizontal axis. So, the equation $y = 6x + 4$ is a good fit.

Example #2

ANALYZING RESIDUALS The table shows the growth *y* (in inches) of an elk's antlers during week *x*. The equation $y = -0.7x + 6.8$ models the data. Is the model a good fit? Explain.

Week, *x*	1	2	3	4	5
Growth, *y*	6.0	5.5	4.7	3.9	3.3

x	*y*	*y*-Value from model	Residual
1	6	6.1	$6 - 6.1 = -0.1$
2	5.5	5.4	$5.5 - 5.4 = 0.1$
3	4.7	4.7	$4.7 - 4.7 = 0$
4	3.9	4	$3.9 - 4 = -0.1$
5	3.3	3.3	$3.3 - 3.3 = 0$



The points are evenly dispersed about the horizontal axis. So, the equation $y = -0.7x + 6.8$ is a good fit.

Copyright © Big Ideas Learning, LLC
All rights reserved.

Practice A

In Exercises 1 and 2, use residuals to determine whether the model is a good fit for the data in the table. Explain.

1. $y = -3x + 2$

x	–4	–3	–2	–1	0	1	2	3	4
y	13	11	8	6	3	0	–4	–8	–10

2. $y = -0.5x + 1$

x	0	1	2	3	4	5	6	7	8
y	2	0	–3	–5	–7	–6	–4	–3	–1

3. The table shows the numbers *y* of visitors to a particular beach and the average daily temperatures *x*.

Average Daily Temperature (°F)	Number of Beach Visitors
80	100
82	150
83	145
85	190
86	215
88	263
89	300
90	350

a. Use a graphing calculator to find an equation of the line of best fit. Then plot the data and graph the equation in the same viewing window.

b. Identify and interpret the correlation coefficient.

c. Interpret the slope and *y*-intercept of the line of best fit.

Copyright © Big Ideas Learning, LLC
All rights reserved.

Name ____ Date __________

Practice B

In Exercises 1 and 2, use residuals to determine whether the model is a good fit for the data in the table. Explain.

1. $y = \frac{3}{2}x - 10$

x	2	4	6	8	10	12	14	16	18
y	−1	−1	1	2	5	6	8	10	14

2. $y = -2x + 56$

x	1	2	3	4	5	6	7	8	9
y	52	50	48	47	45	42	41	38	35

In Exercises 3 and 4, use a graphing calculator to find an equation of the line of best fit for the data. Identify and interpret the correlation coefficient.

3.

x	−12	−8	−4	0	4	8	12	16	20
y	48	42	37	31	29	24	19	14	7

4.

x	3	4	5	6	7	8	9	10	11
y	20	36	15	32	12	28	17	16	24

5. The table shows the average number of minutes y per kilometer for runners and the total distance of a running race, x (in kilometers).

x	3.1	6.2	9.3	12.4	15.5	18.6	21.7	24.8	27.9
y	5.4	5.6	5.7	5.9	6.0	6.1	6.3	6.5	6.9

a. Use a graphing calculator to find an equation of the line of best fit.

b. Identify and interpret the correlation coefficient.

c. Interpret the slope and y-intercept of the line of best fit.

d. Approximate the average number of minutes per kilometer when the distance of a race is 31 kilometers.

Copyright © Big Ideas Learning, LLC
All rights reserved.