Name _____ Date _____

**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.

**Ages of American Women at First Marriage**

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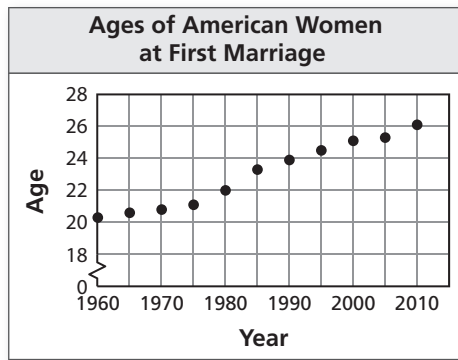
a. The data from the scatter plot is shown in the table. Note that 0, 5, 10, and so on represent the numbers of years since 1960. What does the ordered pair (25, 23.3) represent?

| 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.

---

**4.5** **Analyzing Lines of Fit** (continued)

## 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 |

Name _____ Date _____
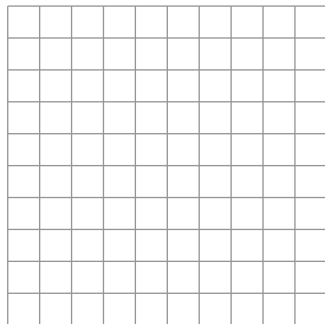
**In your own words, write the meaning of each vocabulary term.**

residual

linear regression

line of best fit

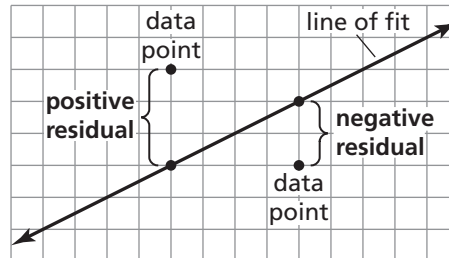correlation coefficient

interpolation

extrapolation

causation

**Notes:**

**4.5** **Notetaking with Vocabulary** (continued)

## 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:**

## Extra Practice

**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 |

**4.5** **Notetaking with Vocabulary** (continued)

**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 number of visitors $y$ to a particular beach for average daily temperatures $x$.

**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.

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

**b.** Identify and interpret the correlation coefficient.

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