Lesson 2.2.3: Modeling Relationships with Lines

Targets:

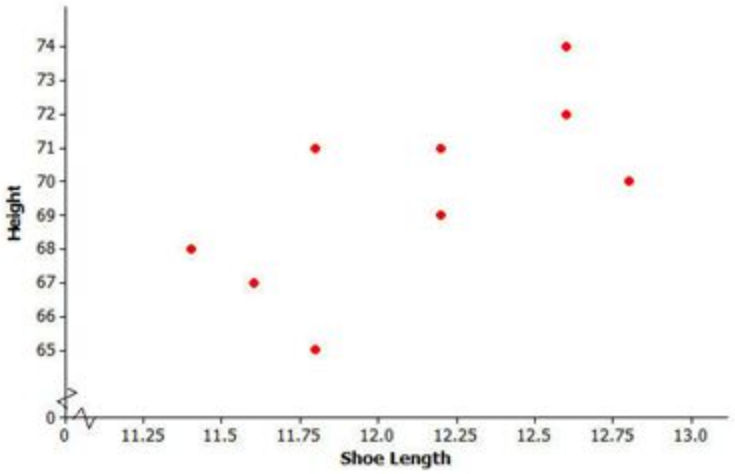
1. I can use technology to determine the equation for the line of best fit for a scatter plot.
2. I can use the line of best fit to make predictions.

Warm Up

Kendra likes to watch crime scene investigation shows on television. She watched a show where investigators used a shoe print to help identify a suspect in a case. She questioned how possible it is to predict someone's height is from his shoe print.

To investigate, she collected data on shoe length (in inches) and height (in inches) from 10 adult men. Her data appear in the table and scatter plot below.

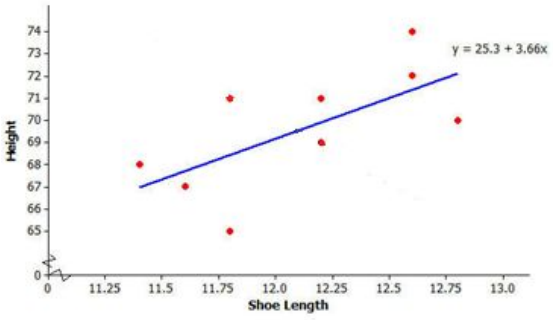
x = Shoe Length	y = Height
12.6	74
11.8	65
12.2	71
11.6	67
12.2	69
11.4	68
12.8	70
12.2	69
12.6	72
11.8	71



1. Is there a relationship between shoe length and height?

2. How would you describe the relationship? Do the men with longer shoe lengths tend be taller?

Practice 1



When two variables x and y are linearly related, you can use a line to describe their relationship. You can also use the equation of the line to predict the value of the y-variable based on the value of the x-variable. For example, the line $y = 25.3 + 3.66x$ might be used to describe the relationship between shoe length and height, where x represents shoe length and y represents height. To predict the height of a man with a shoe length of 12, you would substitute 12 in for x in the equation of the line and then calculate the value of y:

$y = 25.3 + 3.66x = 25.3 + 3.66(12) = 69.22$

You would predict a height of 69.22 inches for a man with a shoe length of 12 inches.

1. What do the red points on the graph represent? What does the blue line represent? Why aren't all the red points on the blue line?

2. One of the men in the sample has a shoe length of 11.8 in. and a height of 71 in. Circle this point the graph.

3. Suppose that you do not know this man's height, but do know that his shoe length is 11.8 inches. If you use the model $y = 25.3 + 3.66x$, what would you predict his height to be?

4. Why is there a difference between the actual height of the man and the predicted height of the man?

Practice 2

One way to think about how useful a line is for describing a relationship between two variables is to use the line to predict the values for the points in the scatter plot. These predicted values could then be compared to the actual y values.

For example, the first data point in the table represents a man with a shoe length of 12.6 inches and height of 74 inches. If you use the line $y = 25.3 + 3.66x$ to predict this man's height, you would get:

$$y = 25.3 + 3.66x = 25.3 + 3.66(12.6) = 71.42 \text{ inches}$$

Because his actual height was 74 inches, you can calculate the prediction error by subtracting the predicted value from the actual value. This prediction error is called a *residual*. For the first data point, the residual is calculated as follows:

$$\text{Residual} = \text{actual } y \text{ value} - \text{predicted } y \text{ value}$$
$$2.58 \text{ inches} = 74 - 41.42$$

1. For the line $y = 25.3 + 3.66x$, calculate the missing values and add them to complete the table.

x = Shoe Length	y = Height	Predicted y-value	Residual
12.6	74	71.42	2.58
11.8	65		-3.49
12.2	71		
11.6	67	67.76	-0.76
12.2	69	69.95	-0.95
11.4	68	67.02	
12.8	70	72.15	-2.15
12.2	69		-0.95
12.6	72	71.42	0.58
11.8	71	68.49	2.51

2. Why is the residual in the table's first row positive and the residual in the second row negative?

3. What is the sum of the residuals? Why did you get a number close to zero for this sum? Does this mean that all of the residuals were close to 0?

Practice 3

- The line that best fits the data of a scatter plot has a number of names. The most common name is the "Line of Best Fit." This line can also be called the *best-fit line* or regression line.
- For the shoe-length and height data for the sample of 10 men, the line $y = 25.3 + 3.66x$ is the Line of Best Fit. No other line would have a smaller sum of residuals.
- There are equations that can be used to calculate the value for the slope and the intercept of the line of best fit, but these formulas require a lot of tedious calculations. Fortunately, a graphing calculator can be used to find the equation of the line of best fit.
- The next video will show you how to enter data and obtain the equation of the line of best fit using your Smarter-Balanced Calculator.

Practice 3 (Continued)

1. Enter the shoe-length and height data and then use your calculator to find the equation of the line of best fit. Did you get $y = 25.3 + 3.66x$? (The slope and -intercept here have been rounded to the nearest hundredth.)

2. Assuming that the 10 men in the sample are representative of adult men in general, what height would you predict for a man whose shoe length is 12.5 inches? What height would you predict for a man whose shoe length is 11.9 inches?

Once you have found the equation of the line of best fit, the values of the slope and -intercept of the line often reveals something interesting about the relationship you are modeling. The slope of the line of best fit is the change in the predicted value of the y variable associated with an increase of one in the value of the x-variable.

3. Give an interpretation of the slope of the line of best fit $y = 25.3 + 3.66x$ for predicting height from shoe size for adult men.

The y-intercept of a line is the predicted value of y when x equals zero. When using a line as a model for the relationship between two numerical variables, it often does not make sense to interpret the y-intercept because an x-value of zero may not make any sense.

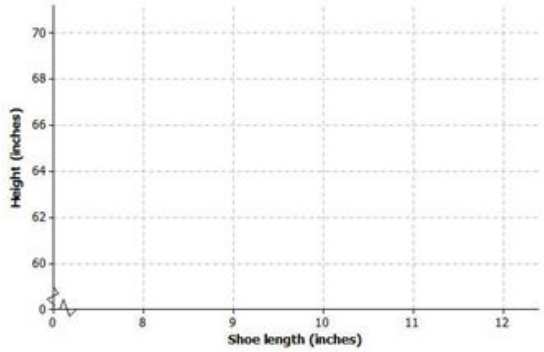
4. Explain why it does not make sense to interpret the *y*-intercept of 25.3 as the predicted height for an adult male whose shoe length is zero.

Exit Ticket

Kendra wondered if the relationship between shoe length and height might be different for men and women. To investigate, she also collected data on shoe length (in inches) and height (in inches) for 12 women.

x = Shoe Length (Women)	y = Height (Women)
8.9	61
9.6	61
9.8	66
10.0	64
10.2	64
10.4	65
10.6	65
10.6	67
10.5	66
10.8	67
11.0	67
11.8	70

1. Construct a scatter plot of these data.



2. Is there a relationship between shoe length and height for these 12 women?

3. Find the equation of the line of best fit. (Round values to the nearest hundredth.)

4. Suppose that these 12 women are representative of adult women in general. Based on the line of best fit, what would you predict for the height of a woman whose shoe length is 10.5 inches? What would you predict for the height of a woman whose shoe length is 11.5 inches?