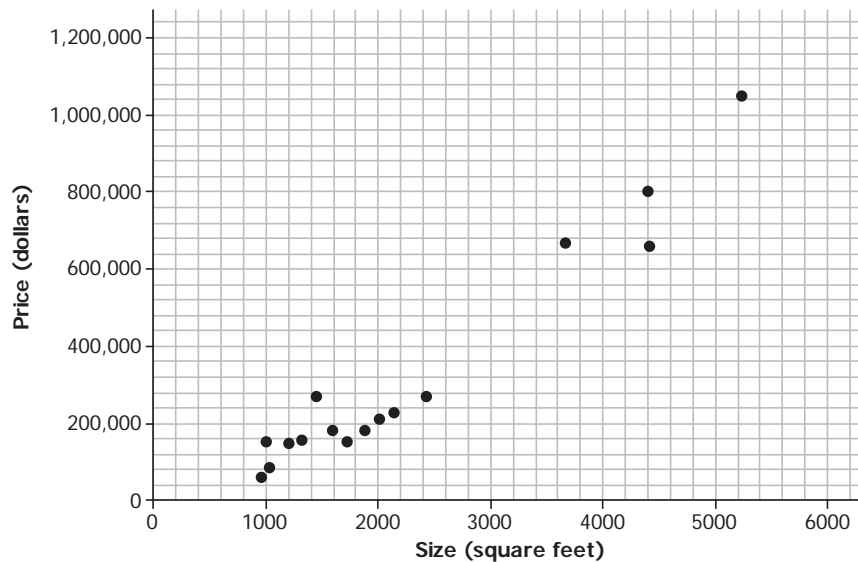# Lesson 8: Informally Fitting a Line

Classwork

Example 1: Housing Costs

Let's look at some data from one midwestern city that indicate the sizes and sale prices of various houses sold in this city.

Size (square feet)	Price (dollars)	Size (square feet)	Price (dollars)
5,232	1,050,000	1,196	144,900
1,875	179,900	1,719	149,900
1,031	84,900	956	59,900
1,437	269,900	991	149,900
4,400	799,900	1,312	154,900
2,000	209,900	4,417	659,999
2,132	224,900	3,664	669,000
1,591	179,900	2,421	269,900

Data Source: http://www.trulia.com/for_sale/Milwaukee,WI/5_p, accessed in 2013

A scatter plot of the data is given below.

Exercises 1–6

1. What can you tell about the price of large homes compared to the price of small homes from the table?

2. Use the scatter plot to answer the following questions.

   a. Does the scatter plot seem to support the statement that larger houses tend to cost more? Explain your thinking.

   b. What is the cost of the most expensive house, and where is that point on the scatter plot?

   c. Some people might consider a given amount of money and then predict what size house they could buy. Others might consider what size house they want and then predict how much it would cost. How would you use the scatter plot in Example 1?

   d. Estimate the cost of a 3,000-square-foot house.

   e. Do you think a line would provide a reasonable way to describe how price and size are related? How could you use a line to predict the price of a house if you are given its size?

3. Draw a line in the plot that you think would fit the trend in the data.

4. Use your line to answer the following questions:
   a. What is your prediction of the price of a 3,000-square-foot house?

   b. What is the prediction of the price of a 1,500-square-foot house?

5. Consider the following general strategies students use for drawing a line. Do you think they represent a good strategy for drawing a line that fits the data? Explain why or why not, or draw a line for the scatter plot using the strategy that would indicate why it is or why it is not a good strategy.
   a. Laure thought she might draw her line using the very first point (farthest to the left) and the very last point (farthest to the right) in the scatter plot.

   b. Phil wants to be sure that he has the same number of points above and below the line.

   c. Sandie thought she might try to get a line that had the most points right on it.

   d. Maree decided to get her line as close to as many of the points as possible.

6. Based on the strategies discussed in Exercise 5, would you change how you draw a line through the points? Explain your answer.

Example 2: Deep Water

Does the current in the water go faster or slower when the water is shallow? The data on the depth and velocity of the Columbia River at various locations in Washington State listed below can help you think about the answer.

**Depth and Velocity in the Columbia River, Washington State**

Depth (feet)	Velocity (feet/second)
0.7	1.55
2.0	1.11
2.6	1.42
3.3	1.39
4.6	1.39
5.9	1.14
7.3	0.91
8.6	0.59
9.9	0.59
10.6	0.41
11.2	0.22

Data Source: www.seattlecentral.edu/qelp/sets/011/011.html

a. What can you tell about the relationship between the depth and velocity by looking at the numbers in the table?

b. If you were to make a scatter plot of the data, which variable would you put on the horizontal axis, and why?

Exercises 7–9

7. A scatter plot of the Columbia River data is shown below.
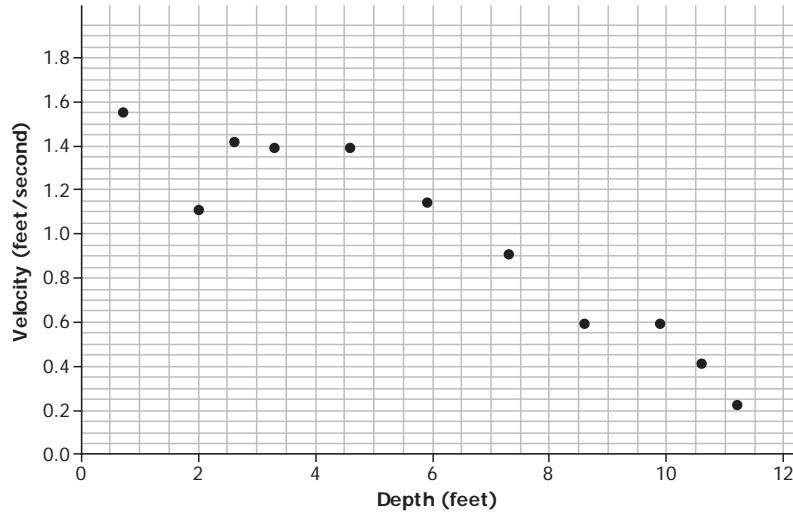
a. Choose a data point in the scatter plot, and describe what it means in terms of the context.

b. Based on the scatter plot, describe the relationship between velocity and depth.

c. How would you explain the relationship between the velocity and depth of the water?

d. If the river is two feet deep at a certain spot, how fast do you think the current would be? Explain your reasoning.

8. Consider the following questions:

   a.   If you draw a line to represent the trend in the plot, will it make it easier to predict the velocity of the water if you know the depth? Why or why not?

   b.   Draw a line that you think does a reasonable job of modeling the trend on the scatter plot in Exercise 7. Use the line to predict the velocity when the water is 8 feet deep.

9. Use the line to predict the velocity for a depth of 8.6 feet. How far off was your prediction from the actual observed velocity for the location that had a depth of 8.6 feet?

Lesson Summary

- When constructing a scatter plot, the variable that you want to predict (i.e., the dependent or response variable) goes on the vertical axis. The independent variable (i.e., the variable not related to other variables) goes on the horizontal axis.

- When the pattern in a scatter plot is approximately linear, a line can be used to describe the linear relationship.

- A line that describes the relationship between a dependent variable and an independent variable can be used to make predictions of the value of the dependent variable given a value of the independent variable.

- When informally fitting a line, you want to find a line for which the points in the scatter plot tend to be closest.

Problem Set

1. The table below shows the mean temperature in July and the mean amount of rainfall per year for 14 cities in the Midwest.

City	Mean Temperature in July (degrees Fahrenheit)	Mean Rainfall per Year (inches)
Chicago, IL	73.3	36.27
Cleveland, OH	71.9	38.71
Columbus, OH	75.1	38.52
Des Moines, IA	76.1	34.72
Detroit, MI	73.5	32.89
Duluth, MN	65.5	31.00
Grand Rapids, MI	71.4	37.13
Indianapolis, IN	75.4	40.95
Marquette, MI	71.6	32.95
Milwaukee, WI	72.0	34.81
Minneapolis–St. Paul, MN	73.2	29.41
Springfield, MO	76.3	35.56
St. Louis, MO	80.2	38.75
Rapid City, SD	73.0	33.21

Data Source: http://countrystudies.us/united-states/weather/

a. What do you observe from looking at the data in the table?

b. Look at the scatter plot below. A line is drawn to fit the data. The plot in the Exit Ticket had the mean July temperatures for the cities on the horizontal axis. How is this plot different, and what does it mean for the way you think about the relationship between the two variables—temperature and rain?

**July Rainfall and Temperatures in Selected Midwestern Cities**
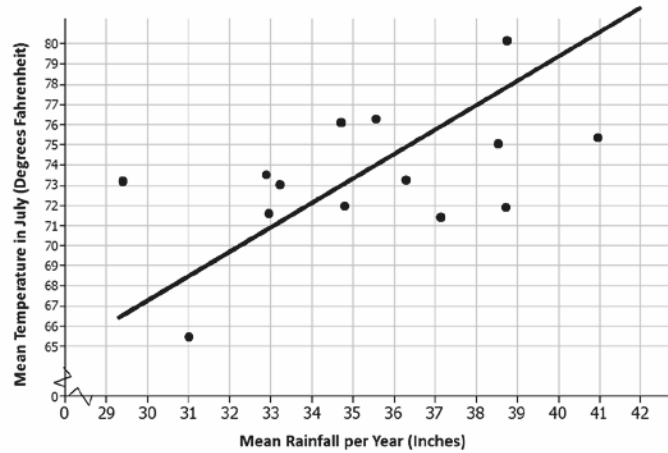
c. The line has been drawn to model the relationship between the amount of rain and the temperature in those midwestern cities. Use the line to predict the mean July temperature for a midwestern city that has a mean of 32 inches of rain per year.

d. For which of the cities in the sample does the line do the worst job of predicting the mean temperature? The best? Explain your reasoning with as much detail as possible.

2. The scatter plot below shows the results of a survey of eighth-grade students who were asked to report the number of hours per week they spend playing video games and the typical number of hours they sleep each night.

Mean Hours Sleep per Night Versus Mean Hours Playing Video Games per Week
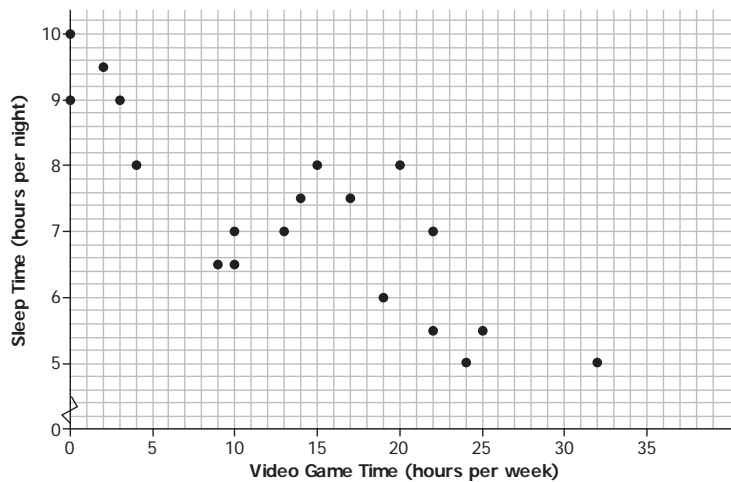
    a.   What trend do you observe in the data?

    b.   What was the fewest number of hours per week that students who were surveyed spent playing video games? The most?

    c.   What was the fewest number of hours per night that students who were surveyed typically slept? The most?

    d.   Draw a line that seems to fit the trend in the data, and find its equation. Use the line to predict the number of hours of sleep for a student who spends about 15 hours per week playing video games.

3.   Scientists can take very good pictures of alligators from airplanes or helicopters. Scientists in Florida are interested in studying the relationship between the length and the weight of alligators in the waters around Florida.

    a.   Would it be easier to collect data on length or weight? Explain your thinking.

    b.   Use your answer to decide which variable you would want to put on the horizontal axis and which variable you might want to predict.

4.   Scientists captured a small sample of alligators and measured both their length (in inches) and weight (in pounds). Torre used their data to create the following scatter plot and drew a line to capture the trend in the data. She and Steve then had a discussion about the way the line fit the data. What do you think they were discussing, and why?

Alligator Length (inches) and Weight (pounds)
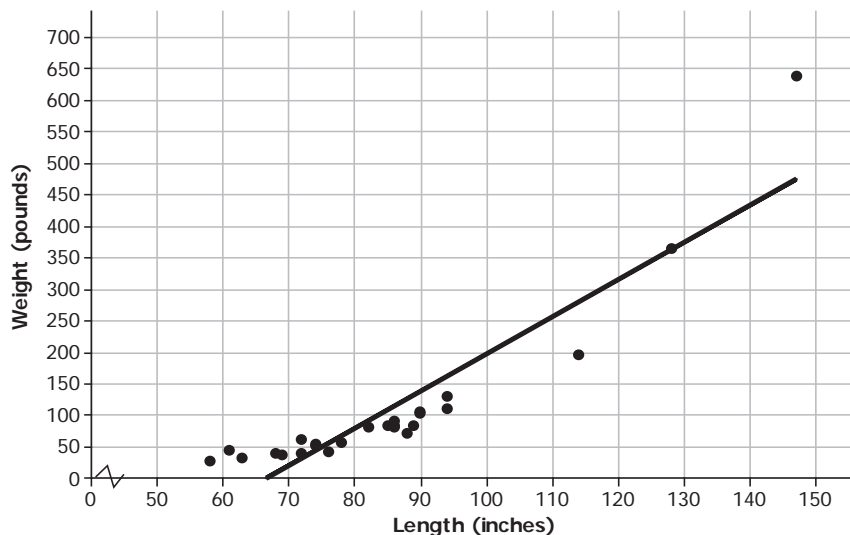
Data Source: James Landwehr and Ann Watkins, *Exploring Data*, Quantitative Literacy Series (Dale Seymour, 1987).