|  |
| --- |
|  |

Lesson 6: Scatter Plots

Student Outcomes

* Students construct scatter plots.
* Students use scatter plots to investigate relationships.
* Students understand that a trend in a scatterplot does not establish cause-and-effect.

Lesson Notes

This lesson is the first in a set of lessons dealing with relationships between numerical variables. In this lesson, students learn how to construct a scatter plot and look for patterns which suggest that there is a statistical relationship between two numerical variables.

Classwork

*Scaffolding:*

* Point out to students that the word *bivariate* is composed of the prefix *bi-* and the stem *variate*.
* *Bi-* means *two*.
* *Variate* indicates a variable.
* The focus in this lesson is on two numerical variables.

**Example 1 (5 minutes)**

Spend a few minutes introducing the context of this example. Make sure that students understand that in this context, an *observation* can be thought of as an ordered pair consisting of the value for each of two variables.

Example 1

A bivariate data set consists of observations on two variables. For example, you might collect data on different car models. Each observation in the data set would consist of an pair.

weight (in pounds, rounded to the nearest pounds)

and

fuel efficiency (in miles per gallon, mpg)

The table below shows the weight and fuel efficiency for car models with automatic transmissions manufactured in 2009 by Chevrolet.

*Scaffolding:*

* English language learners new to the curriculum may be familiar with the metric system (kilometers, kilograms, and liters) but unfamiliar with the English system (miles, pounds, and gallons).
* It may be helpful to provide conversions:

kg lb.

lb. kg

km mi.

mi. km

|  |  |  |
| --- | --- | --- |
| Model | Weight (pounds) | Fuel Efficiency (mpg) |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Exercises 1–3 (10–12 minutes)

After students have had a chance to think about Exercise 1, make sure that everyone understands what an observation (an ordered pair) represents in the context of this example. Relate plotting the point that corresponds to the first observation to students’ previous work with plotting points in a rectangular coordinate system. As a way of encouraging the need to look at a graph of the data, consider asking students to try to determine if there is a relationship between weight and fuel efficiency by just looking at the table. Allow students time to complete the scatter plot and complete Exercise 3. Have students share their answers to Exercise 3.

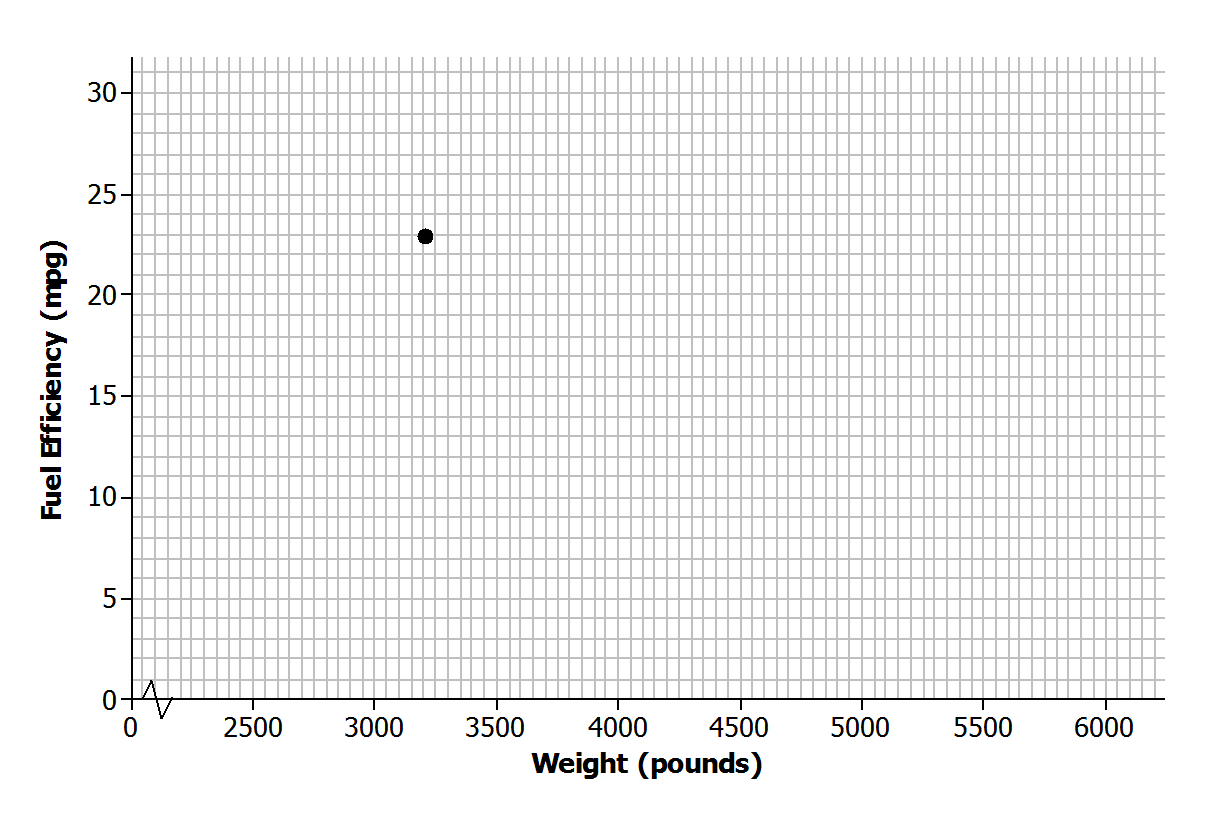
Exercises 1–8

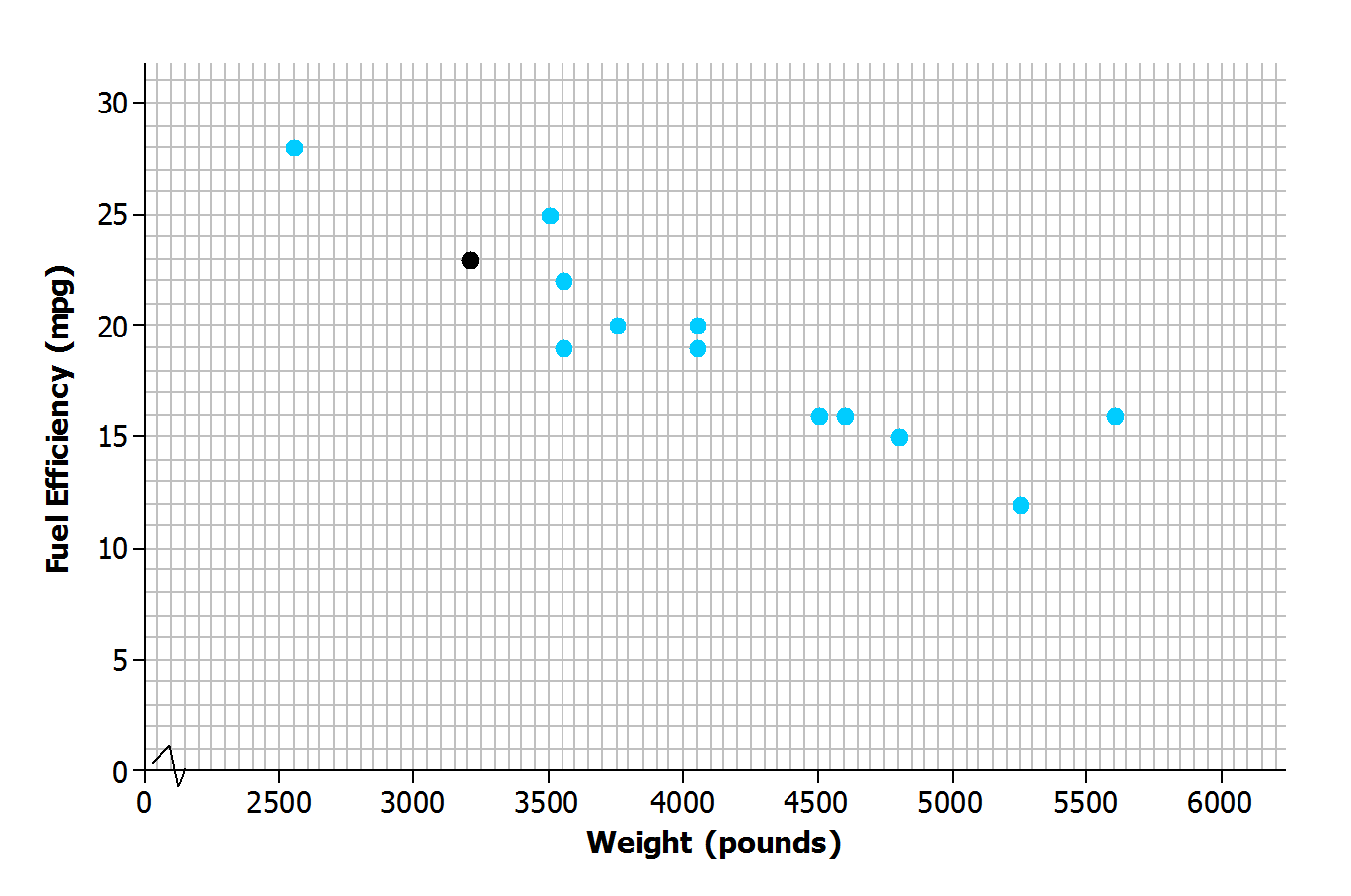
1. In the table above, the observation corresponding to Model 1 is (). What is the fuel efficiency of this car? What is the weight of this car?

The fuel efficiency is miles per gallon, and the weight is pounds.

One question of interest is whether there is a relationship between the car weight and fuel efficiency. The best way to begin to investigate is to construct a graph of the data. A *scatter plot* is a graph of the pairs in the data set. Each pair is plotted as a point in a rectangular coordinate system.

For example, the observationwould be plotted as a point located above on the -axis and across from on the -axis, as shown below.



1. Add the points corresponding to the other observations to the scatter plot.

1. Do you notice a pattern in the scatter plot? What does this imply about the relationship between weight () and fuel efficiency ()?

MP.7

There does seem to be a pattern in the plot. Higher weights tend to be paired with lesser fuel efficiencies, so it looks like heavier cars generally have lower fuel efficiency.

Exercises 4–8 (6–8 minutes)

These exercises give students additional practice creating a scatter plot and identifying a pattern in the plot. Students should work individually on these exercises and then discuss their answers to Exercises 7 and 8 with a partner. However, some English language learners may benefit from paired or small group work, particularly if their English literacy is not strong.

Is there a relationship between price and the quality of athletic shoes? The data in the table below are from the Consumer Reports website.

price (in dollars)

and

Consumer Reports quality rating

The quality rating is on a scale of to , with being the highest quality.

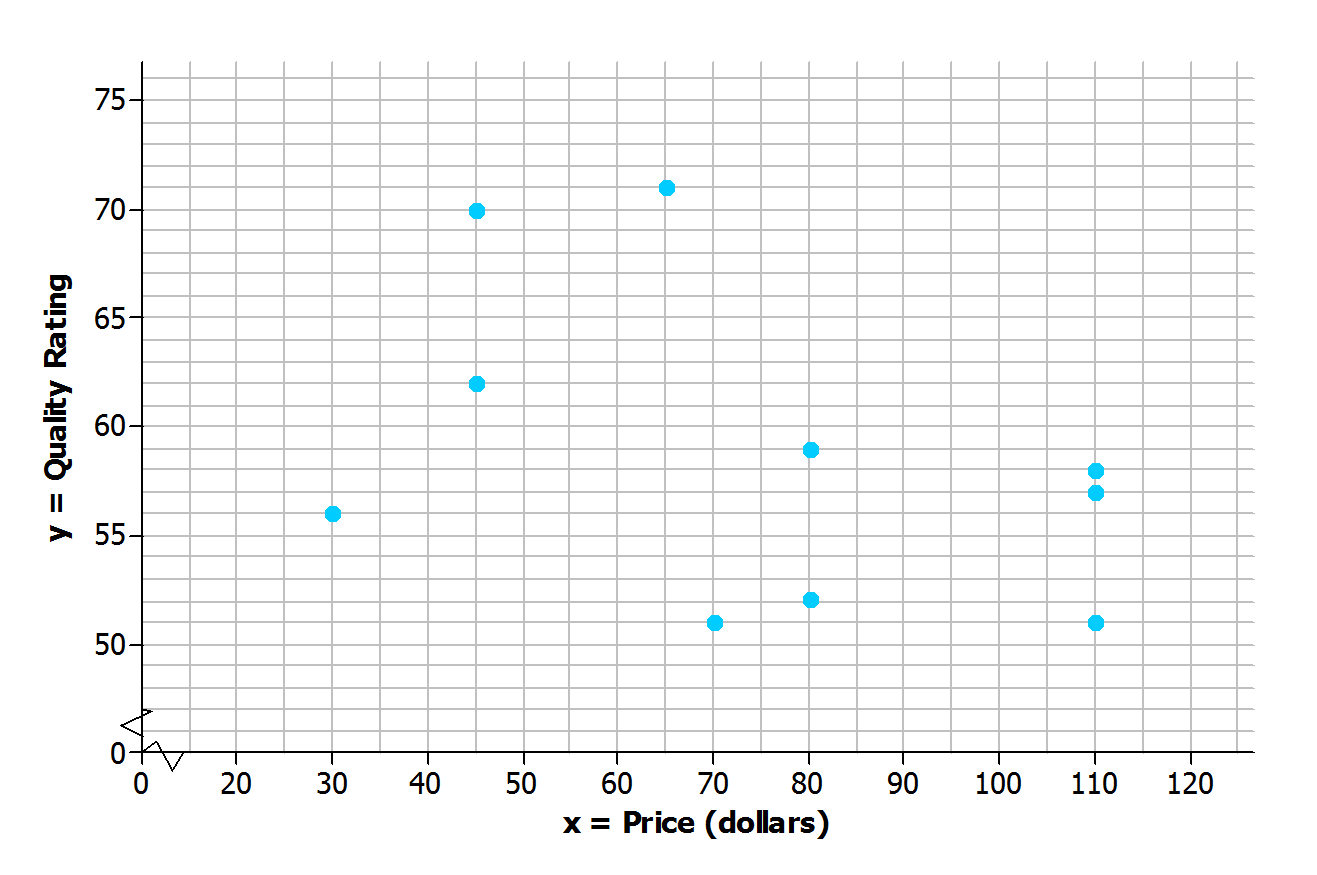
|  |  |  |
| --- | --- | --- |
| Shoe | Price (dollars) | Quality Rating |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

1. One observation in the data set is. What does this ordered pair represent in terms of cost and quality?

The pair represents a shoe that costs with a quality rating of .

1. To construct a scatter plot of these data, you need to start by thinking about appropriate scales for the axes of the scatter plot. The prices in the data set range from to , so one reasonable choice for the scale of the -axis would range fromto , as shown below. What would be a reasonable choice for a scale for the -axis?

Sample response: The smallest -value is , and the -largest value is . So, the -axis could be scaled from to .

1. Add a scale to the -axis. Then, use these axes to construct a scatter plot of the data.
2. Do you see any pattern in the scatter plot indicating that there is a relationship between price and quality rating for athletic shoes?

*Scaffolding:*

For more complicated and reflective answers, consider allowing English language learners to use one or more of the following options: collaborate with a same-language peer, illustrate their response, or provide a first-language narration or response.

Answers will vary. Students may say that they do not see a pattern, or they may say that they see a slight downward trend.

1. Some people think that if shoes have a high price, they must be of high quality. How would you respond?

Answers will vary. The data do not support this. Students will either respond that there does not appear to be a relationship between price and quality, or if they saw a downward trend in the scatter plot, they might even indicate that the higher-priced shoes tend to have lower quality. Look for consistency between the answer to this question and how students answered the previous question.

**Example 2 (5–10 minutes): Statistical Relationships**

This example makes a very important point. As you discuss this example with the class, make sure students understand the distinction between a statistical relationship and a cause-and-effect relationship. After discussing the example, ask students if they can think of other examples of numerical variables that might have a statistical relationship but which probably do not have a cause-and-effect relationship.

Example 2: Statistical Relationship

A pattern in a scatter plot indicates that the values of one variable tend to vary in a predictable way as the values of the other variable change. This is called a *statistical relationship*. In the fuel efficiency and car weight example, fuel efficiency tended to decrease as car weight increased.

This is useful information, but be careful not to jump to the conclusion that increasing the weight of a car *causes* the fuel efficiency to go down. There may be some other explanation for this. For example, heavier cars may also have bigger engines, and bigger engines may be less efficient. You cannot conclude that changes to one variable *cause* changes in the other variable just because there is a statistical relationship in a scatter plot.

Exercises 9–10 (5 minutes)

Students can work individually or with a partner on these exercises. Then, confirm answers as a class.

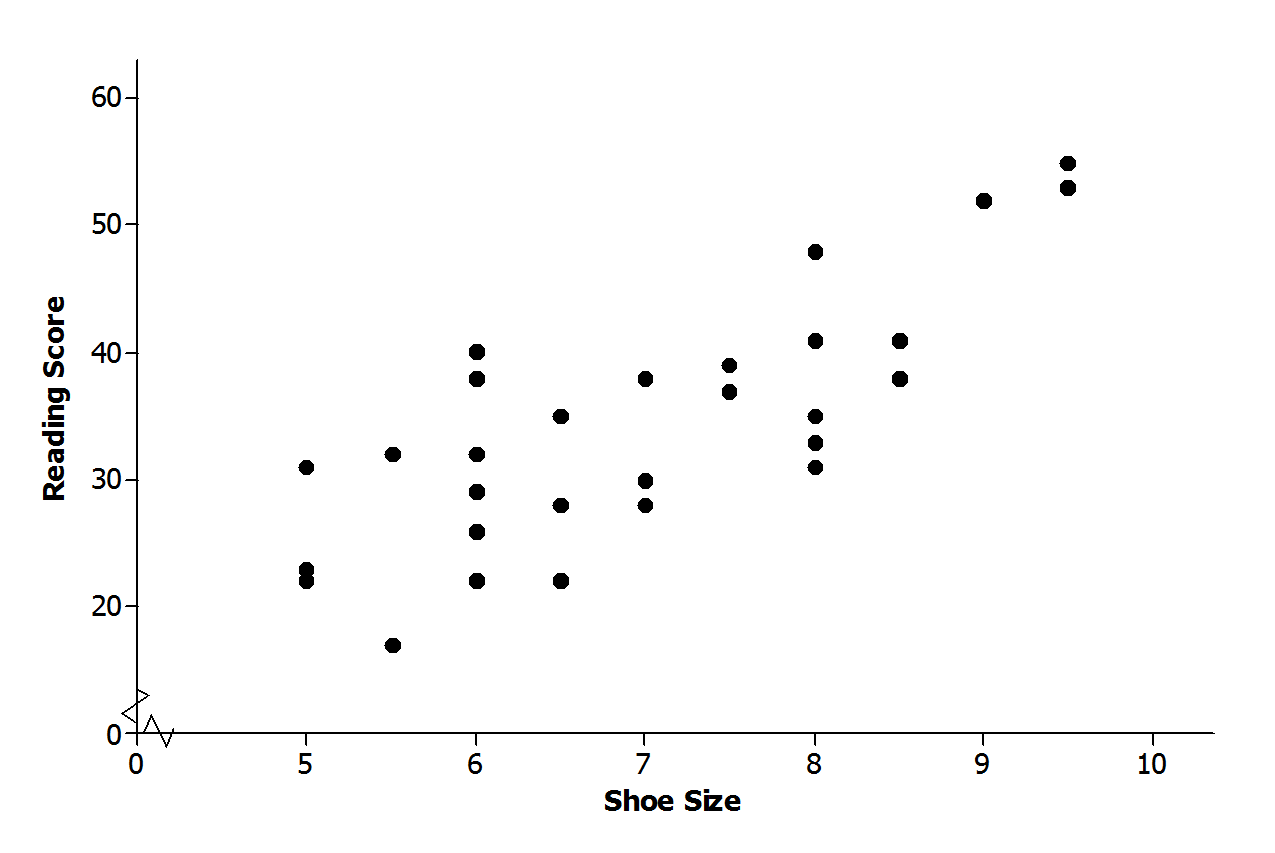
Exercises 9–10

1. Data were collected on

shoe size

and

score on a reading-ability test

for elementary school students. The scatter plot of these data is shown below. Does there appear to be a statistical relationship between shoe size and score on the reading test?

Possible response: The pattern in the scatter plot appears to follow a line. As shoe sizes increase, the reading scores also seem to increase. There does appear to be a statistical relationship because there is a pattern in the scatter plot.

1. Explain why it is not reasonable to conclude that having big feet causes a high reading score. Can you think of a different explanation for why you might see a pattern like this?

Possible response: You cannot conclude that just because there is a statistical relationship between shoe size and reading score that one causes the other. These data were for students completing a reading test for younger elementary school children. Older children, who would have bigger feet than younger children, would probably tend to score higher on a reading test for younger students.

Closing (3 minutes)

Consider posingthe following questions; allow a few student responses for each.

* Why is it helpful to make a scatter plot when you have data on two numerical variables?
  + *A scatter plot makes it easier to see patterns in the data and to see if there is a statistical relationship between the two variables.*
* Can you think of an example of two variables that would have a statistical relationship but not a cause-and-effect relationship?
  + *One famous example is the number of people who must be rescued by lifeguards at the beach and the number of ice cream sales. Both of these variables have higher values when the temperature is high and lower values when the temperature is low. So, there is a statistical relationship between them—they tend to vary in a predictable way. However, it would be silly to say that an increase in ice cream sales causes more beach rescues.*

Lesson Summary

* **A scatter plot is a graph of numerical data on two variables.**
* **A pattern in a scatter plot suggests that there may be a relationship between the two variables used to construct the scatter plot.**
* **If two variables tend to vary together in a predictable way, we can say that there is a statistical relationship between the two variables.**
* **A statistical relationship between two variables does not imply that a change in one variable causes a change in the other variable (a cause-and-effect relationship).**

Exit Ticket (5 minutes)

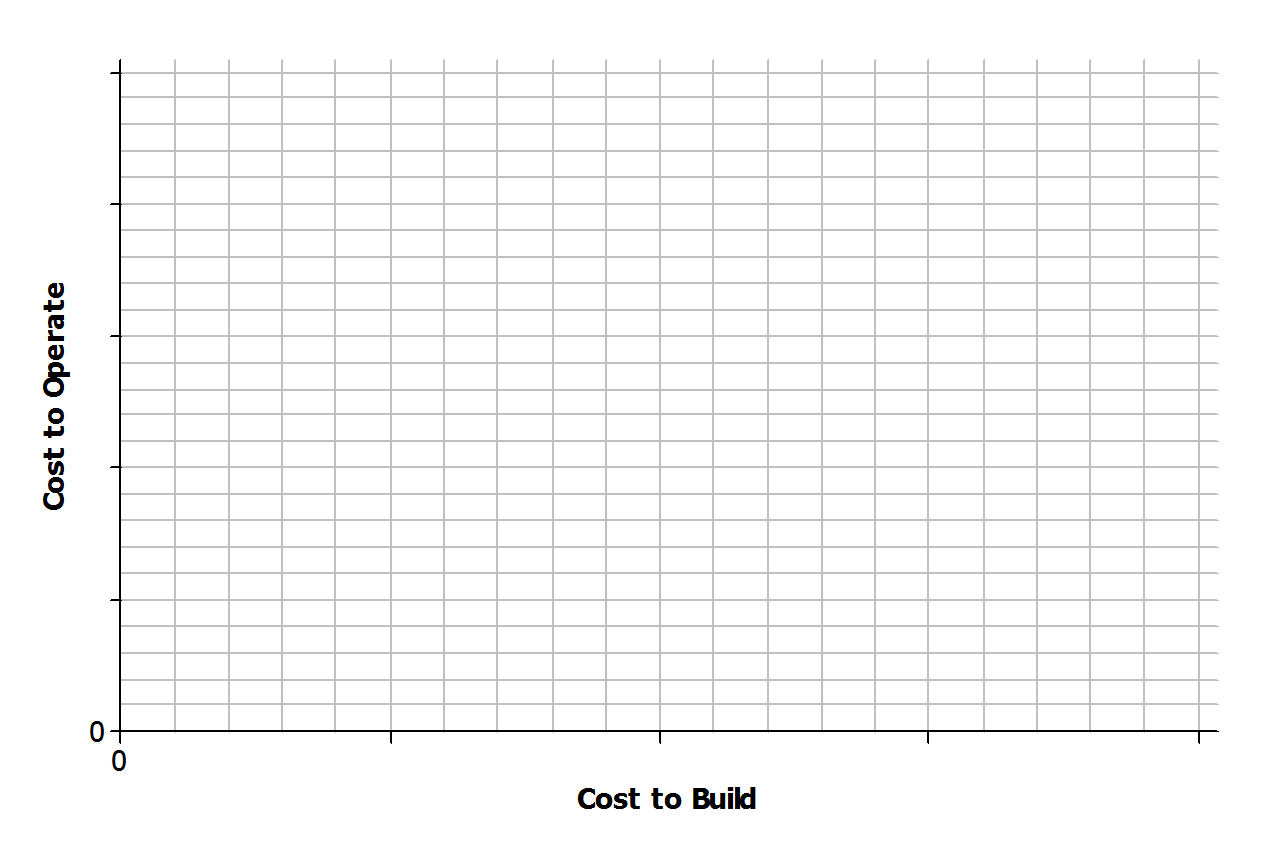
Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Lesson 6: Scatter Plots

Exit Ticket

Energy is measured in kilowatt hours. The table below shows the cost of building a facility to produce energy and the ongoing cost of operating the facility for five different types of energy.

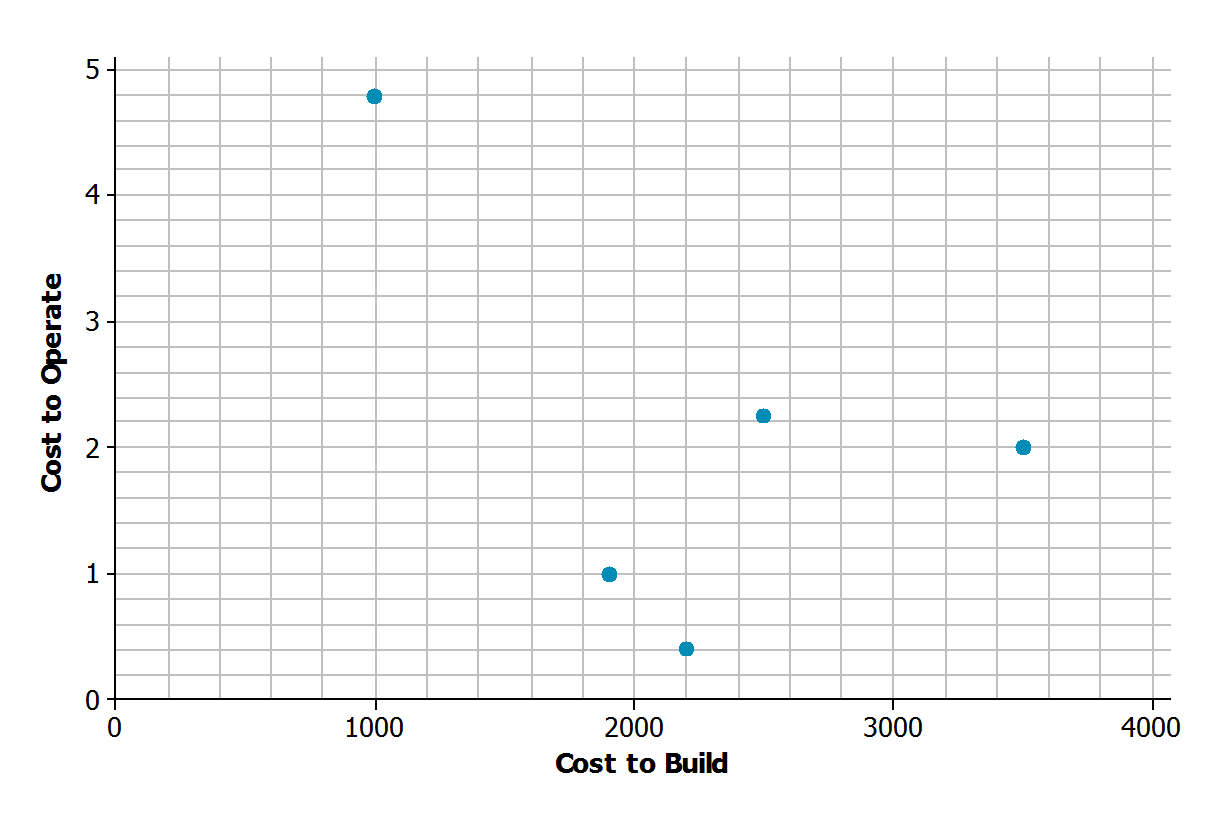
|  |  |  |
| --- | --- | --- |
| **Type of Energy** | **Cost to Operate**  **(cents per kilowatt hour)** | **Cost to Build**  **(dollars per kilowatt hour)** |
| Hydroelectric |  |  |
| Wind |  |  |
| Nuclear |  |  |
| Coal |  |  |
| Natural Gas |  |  |

1. Construct a scatter plot of the cost to build the facility () and the cost to operate the facility (). Use the grid below, and be sure to add an appropriate scale to the axes.
2. Do you think that there is a statistical relationship between building cost and operating cost? If so, describe the nature of the relationship.
3. Based on the scatter plot, can you conclude that decreased building cost is the cause of increased operating cost? Explain.

Exit Ticket Sample Solutions

Energy is measured in kilowatt hours. The table below shows the cost of building a facility to produce energy and the ongoing cost of operating the facility for five different types of energy.

|  |  |  |
| --- | --- | --- |
| Type of Energy | Cost to Operate  (cents per kilowatt hour) | Cost to Build  (dollars per kilowatt hour) |
| Hydroelectric |  |  |
| Wind |  |  |
| Nuclear |  |  |
| Coal |  |  |
| Natural Gas |  |  |

1. Construct a scatter plot of the cost to build the facility () and the cost to operate the facility (). Use the grid below, and be sure to add an appropriate scale to the axes.
2. Do you think that there is a statistical relationship between building cost and operating cost? If so, describe the nature of the relationship.

Answers may vary. Sample response: Yes, because it looks like there is a downward pattern in the scatter plot. It appears that the types of energy that have facilities that are more expensive to build are less expensive to operate.

1. Based on the scatter plot, can you conclude that decreased building cost is the cause of increased operating cost? Explain.

Sample response: No. Just because there may be a statistical relationship between cost to build and cost to operate does not mean that there is a cause-and-effect relationship.

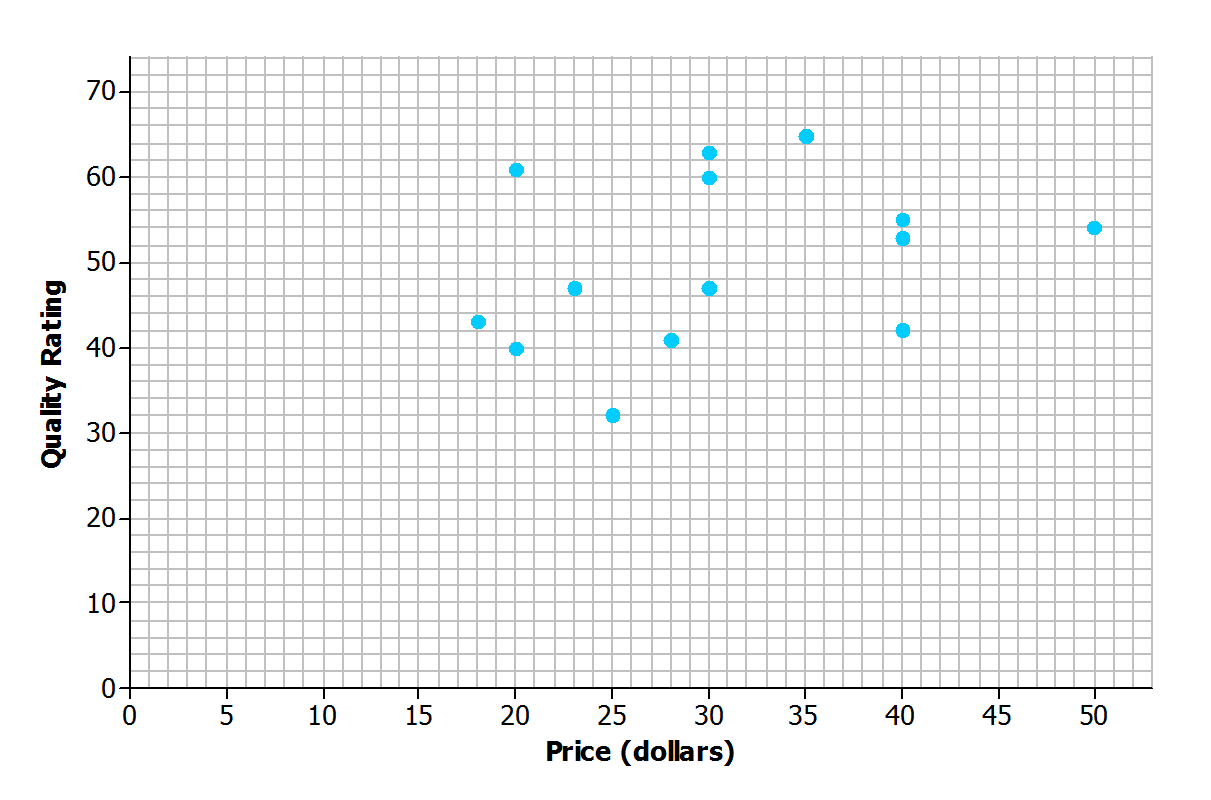
Problem Set Sample Solutions

The Problem Set is intended to reinforce material from the lesson and have students think about the meaning of points in a scatter plot, clusters, positive and negative linear trends, and trends that are not linear.

1. The table below shows the price and overall quality rating for different brands of bike helmets.

Data Source: [www.consumerreports.org](http://www.consumerreports.org)

|  |  |  |
| --- | --- | --- |
| Helmet | Price (dollars) | Quality Rating |
| A |  |  |
| B |  |  |
| C |  |  |
| D |  |  |
| E |  |  |
| F |  |  |
| G |  |  |
| H |  |  |
| I |  |  |
| J |  |  |
| K |  |  |
| L |  |  |
| M |  |  |
| N |  |  |
| O |  |  |

Construct a scatter plot of price () and quality rating (). Use the grid below.

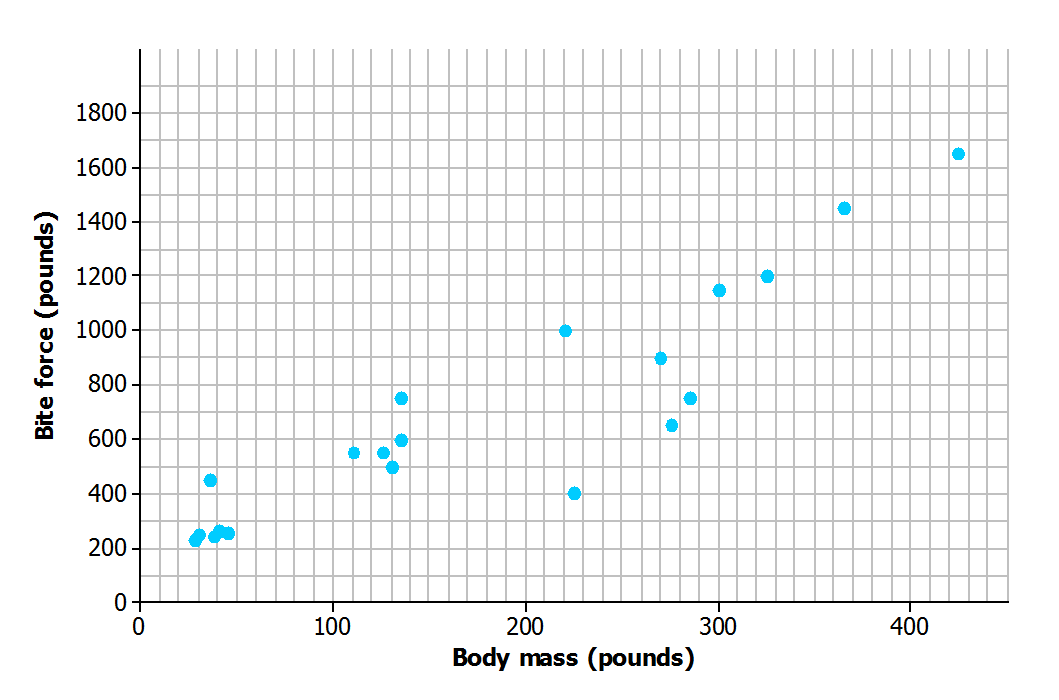
1. Do you think that there is a statistical relationship between price and quality rating? If so, describe the nature of the relationship.

Sample response: No. There is no pattern visible in the scatter plot. There does not appear to be a relationship between price and the quality rating for bike helmets.

1. Scientists are interested in finding out how different species adapt to finding food sources. One group studied crocodilian species to find out how their bite force was related to body mass and diet. The table below displays the information they collected on body mass (in pounds) and bite force (in pounds).

|  |  |  |
| --- | --- | --- |
| **Species** | **Body Mass (pounds)** | **Bite Force (pounds)** |
| **Dwarf crocodile** |  |  |
| **Crocodile F** |  |  |
| **Alligator A** |  |  |
| **Caiman A** |  |  |
| **Caiman B** |  |  |
| **Caiman C** |  |  |
| **Croc A** |  |  |
| **Nile crocodile** |  |  |
| **Croc B** |  |  |
| **Croc C** |  |  |
| **Croc D** |  |  |
| **Caiman D** |  |  |
| **Indian Gharial croc** |  |  |
| **Crocodile G** |  |  |
| **American croc** |  |  |
| **Croc D** |  |  |
| **Croc E** |  |  |
| **American Alligator** |  |  |
| **Alligator B** |  |  |
| **Alligator C** |  |  |

Data Source: PLoS One Greg Erickson biomechanics, Florida State University

Construct a scatter plot of body mass () and bite force (). Use the grid below, and be sure to add an appropriate scale to the axes.

1. Do you think that there is a statistical relationship between body mass and bite force? If so, describe the nature of the relationship.

Sample response: Yes, because it looks like there is an upward pattern in the scatter plot. It appears that alligators with larger body mass also tend to have greater bite force.

1. Based on the scatter plot, can you conclude that increased body mass causes increased bite force? Explain.

Sample response: No. Just because there is a statistical relationship between body mass and bite force does not mean that there is a cause-and-effect relationship.