Lesson 7: Patterns in Scatter Plots

Classwork

Example 1

In the previous lesson, you learned that when data is collected on two numerical variables, a good place to start is to look at a scatter plot of the data.

When you look at a scatter plot, you should ask yourself the following questions:
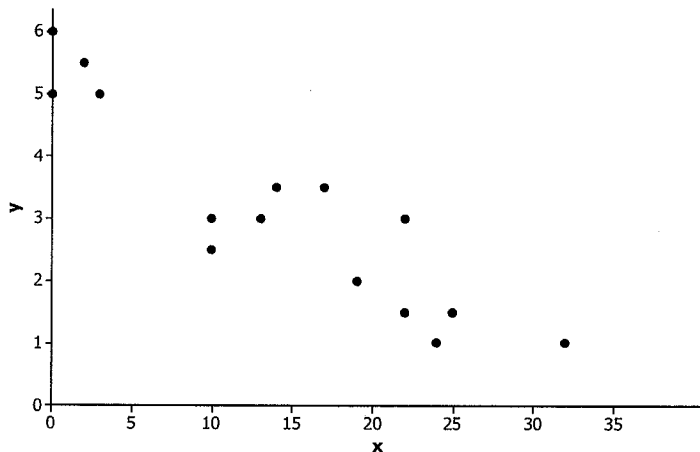
1. Does it look like there is a relationship between the two variables used to make the scatter plot?
2. If there is a relationship, does it appear to be linear?
3. If the relationship appears to be linear, is the relationship a positive linear relationship or a negative linear relationship?

To answer the first question, look for patterns in the scatter plot. Does there appear to be a general pattern to the points in the scatter plot, or do the points look as if they are scattered at random? If you see a pattern, you can answer the second question by thinking about whether the pattern would be well-described by a line. Answering the third question requires you to distinguish between a positive linear relationship and a negative linear relationship. A positive linear relationship is one that is described by a line with a positive slope. A negative linear relationship is one that is described by a line with a negative slope.

Exercises 1–5

Take a look at the following five scatter plots. Answer the three questions above for each scatter plot.
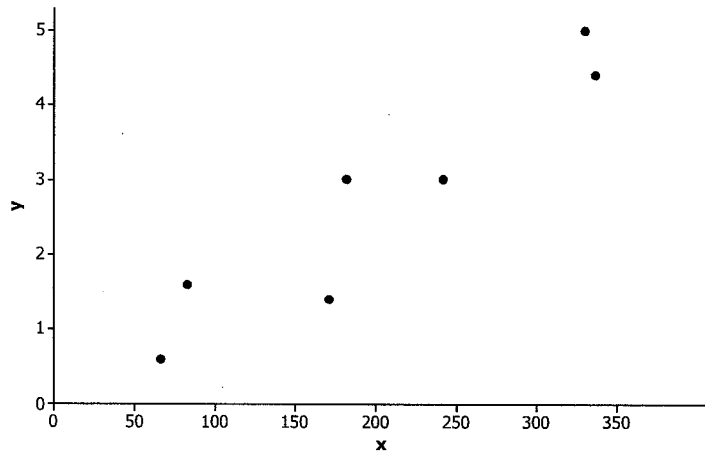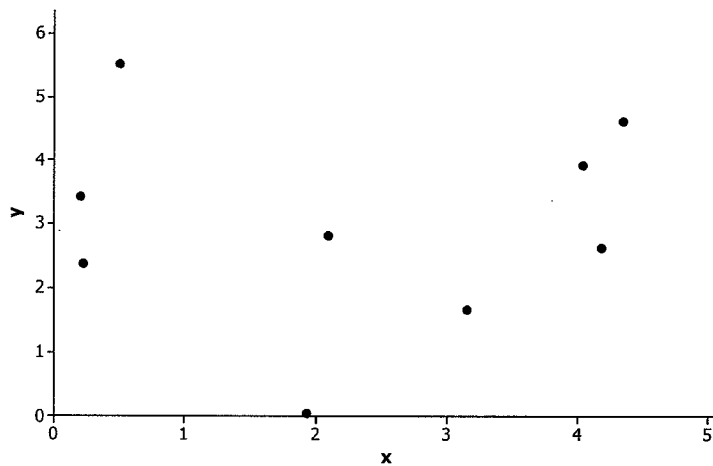
1. Scatter plot 1

Is there a relationship?

If there is a relationship, does it appear to be linear?

If the relationship appears to be linear, is it a positive or negative linear relationship?

2. Scatter plot 2

Is there a relationship?

If there is a relationship, does it appear to be linear?

If the relationship appears to be linear, is it a positive or negative linear relationship?

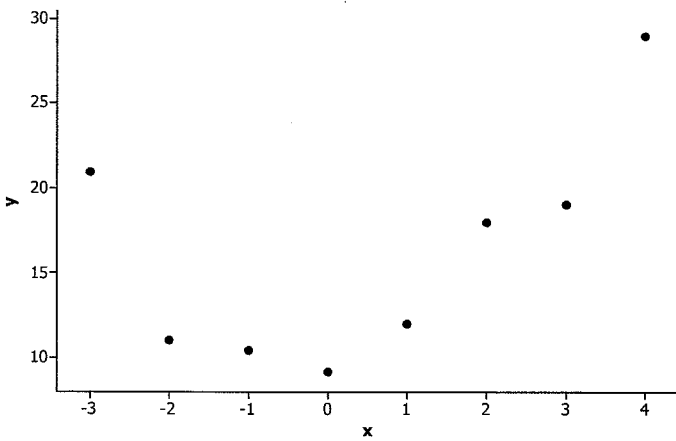3. Scatter plot 3

Is there a relationship?

If there is a relationship, does it appear to be linear?

If the relationship appears to be linear, is it a positive or negative linear relationship?

4.   Scatter plot 4

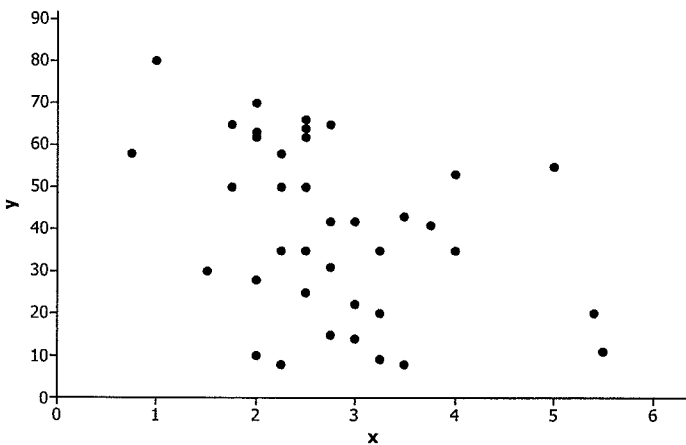

Is there a relationship?

If there is a relationship, does it appear to be linear?

If the relationship appears to be linear, is it a positive or negative linear relationship?

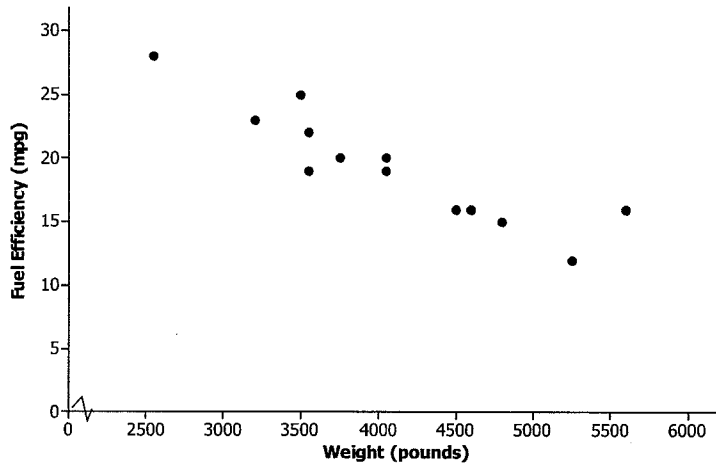5.   Scatter plot 5

Is there a relationship?

If there is a relationship, does it appear to be linear?

If the relationship appears to be linear, is it a positive or negative linear relationship?
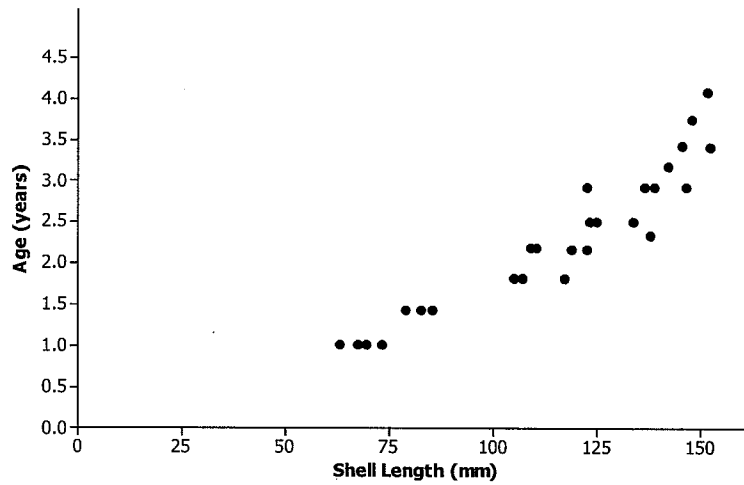
Exercises 6–9

6. Below is a scatter plot of data on weight (x) and fuel efficiency (y) for 13 cars. Using the questions at the beginning of this lesson as a guide, write a few sentences describing any possible relationship between x and y.


7. Below is a scatter plot of data on price (x) and quality rating (y) for 14 bike helmets. Using the questions at the beginning of this lesson as a guide, write a few sentences describing any possible relationship between x and y.
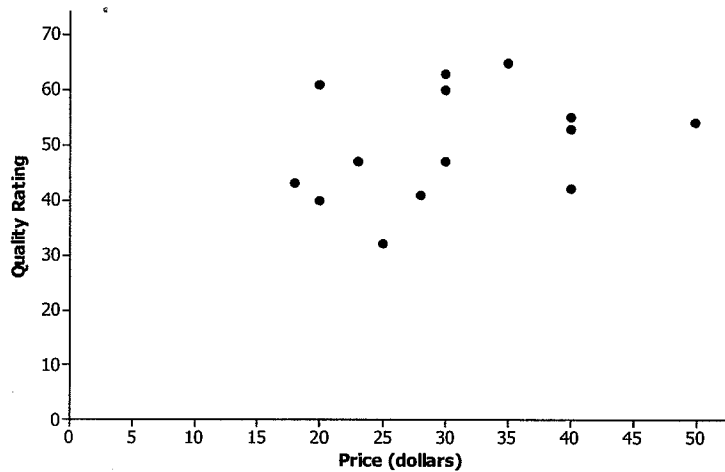
8. Below is a scatter plot of data on shell length (x) and age (y) for 27 lobsters of known age. Using the questions at the beginning of this lesson as a guide, write a few sentences describing any possible relationship between x and y.

9. Below is a scatter plot of data from crocodiles on body mass (x) and bite force (y). Using the questions at the beginning of this lesson as a guide, write a few sentences describing any possible relationship between x and y.
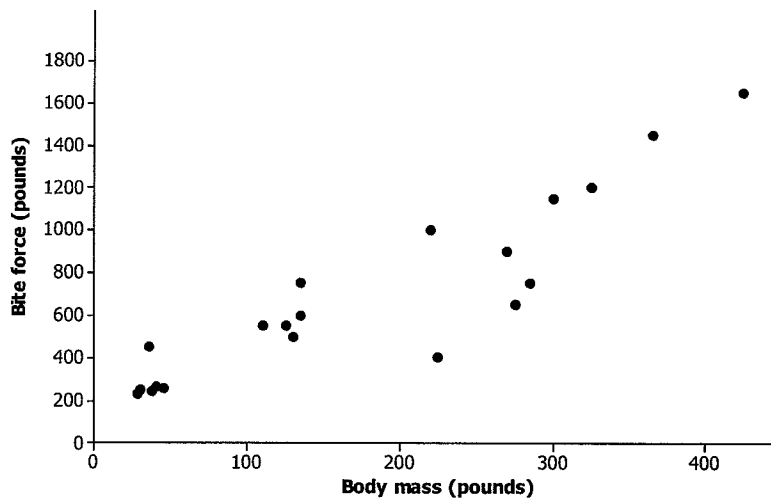
Example 2

In addition to looking for a general pattern in a scatter plot, you should also look for other interesting features that might help you understand the relationship between two variables. Two things to watch for are as follows:

- **Clusters:** Usually the points in a scatter plot form a single cloud of points, but sometimes the points may form two or more distinct clouds of points. These clouds are called *clusters*. Investigating these clusters may tell you something useful about the data.

- **Outliers:** An *outlier* is an unusual point in a scatter plot that does not seem to fit the general pattern or that is far away from the other points in the scatter plot.

The scatter plot below was constructed using data from a study of Rocky Mountain elk ("Estimating Elk Weight from Chest Girth," *Wildlife Society Bulletin*, 1996). The variables studied were chest girth in cm (x) and weight in kg (y).
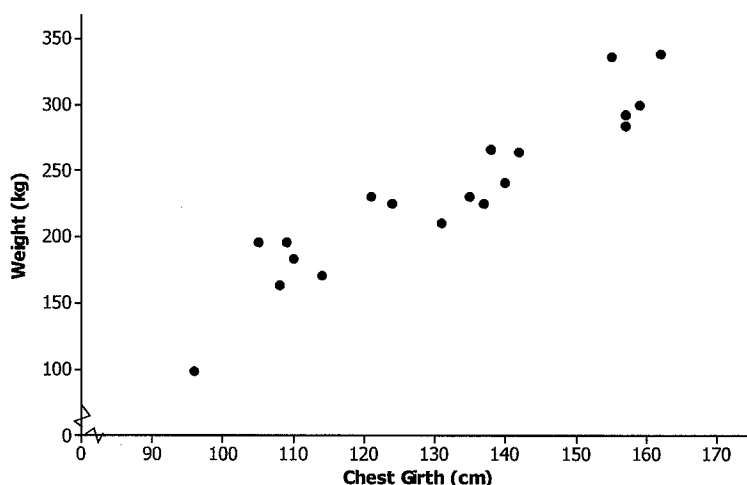
Exercises 10–12

10. Do you notice any point in the scatter plot of elk weight versus chest girth that might be described as an outlier? If so, which one?

11. If you identified an outlier in Exercise 10, write a sentence describing how this data observation differs from the others in the data set.

12. Do you notice any clusters in the scatter plot? If so, how would you distinguish between the clusters in terms of chest girth? Can you think of a reason these clusters might have occurred?