Linear Relationships: Introduction

Learning Objective: Interpret the value of the correlation coefficient, and be aware of its limitations as a numerical measure of the association between two quantitative variables.

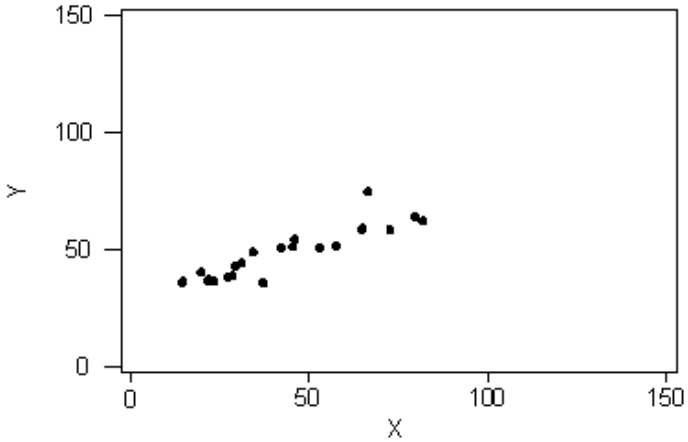
Introduction

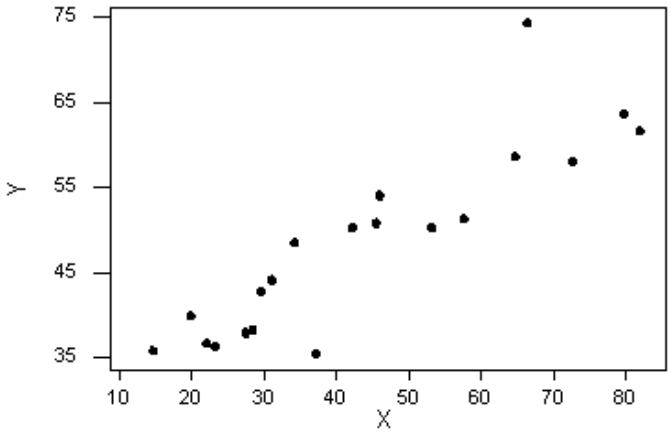
So far we have visualized relationships between two quantitative variables using scatterplots, and described the overall pattern of a relationship by considering its direction, form, and strength. We noted that assessing the strength of a relationship just by looking at the scatterplot is quite difficult, and therefore we need to supplement the scatterplot with some kind of numerical measure that will help us assess the strength.

In this part, we will restrict our attention to the **special case of relationships that have a linear form**, since they are quite common and relatively simple to detect. More importantly, there exists a numerical measure that assesses the strength of the linear relationship between two quantitative variables with which we can supplement the scatterplot. We will introduce this numerical measure here and discuss it in detail.

Even though from this point on we are going to focus only on linear relationships, it is important to remember that **not every relationship between two quantitative variables has a linear form.** We have actually seen several examples of relationships that are not linear. The statistical tools that will be introduced here are **appropriate only for examining linear relationships,** and as we will see, when they are used in nonlinear situations, these tools can lead to errors in reasoning.

Let's start with a motivating example. Consider the following two scatterplots.





We can see that in both cases, the direction of the relationship is **positive** and the form of the relationship is **linear**. What about the strength? Recall that the strength of a relationship is the extent to which the data follow its form.

Learn By Doing

1/1 point (graded)

Which of the two scatterplots do you think displays a stronger linear relationship?

- ○ The top scatterplot
- ○ The bottom scatterplot
- ◉ Both relationships are equally strong ✔

Answer

Correct:

Both scatterplots display the same data, but they look different, because different scales were used on the axes. Both scatterplots display a positive linear relationship.

Submit

The purpose of this example was to illustrate how assessing the strength of the linear relationship from a scatterplot alone is problematic, since our judgment might be affected by the scale on which the values are plotted. This example, therefore, provides a motivation for the **need** to supplement the scatterplot with a **numerical measure** that will **measure the strength** of the linear relationship between two quantitative variables.