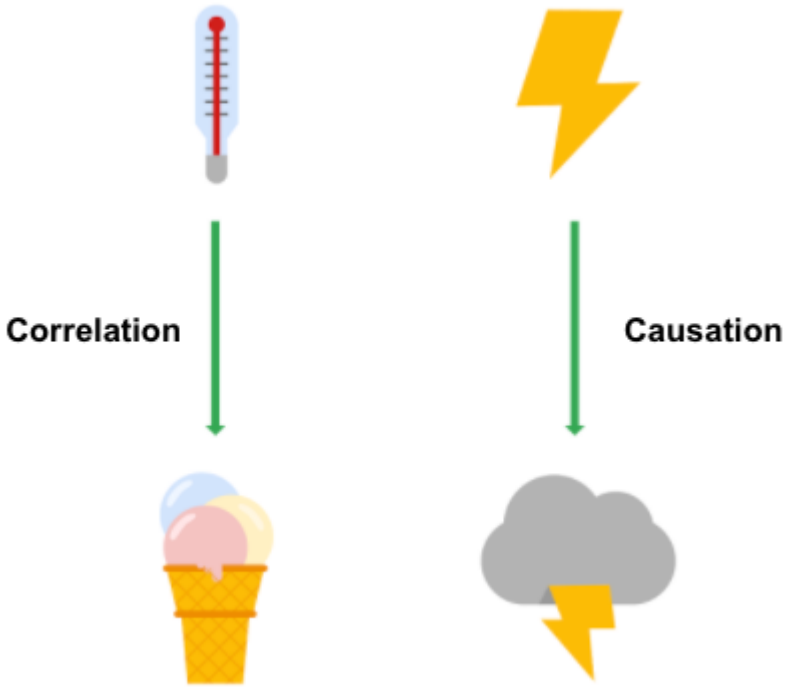
Correlation and causation

In this reading, you will examine correlation and causation in more detail. Let's review the definitions of these terms:

- **Correlation** in statistics is the measure of the degree to which two variables move in relationship to each other. An example of correlation is the idea that "As the temperature goes up, ice cream sales also go up." It is important to remember that correlation doesn't mean that one event causes another. But, it does indicate that they have a pattern with or a relationship to each other. If one variable goes up and the other variable also goes up, it is a positive correlation. If one variable goes up and the other variable goes down, it is a negative or inverse correlation. If one variable goes up and the other variable stays about the same, there is no correlation.
- **Causation** refers to the idea that an event leads to a specific outcome. For example, when lightning strikes, we hear the thunder (sound wave) caused by the air heating and cooling from the lightning strike. Lightning causes thunder.



Why is differentiating between correlation and causation important?

When you make conclusions from data analysis, you need to make sure that you don't assume a causal relationship between elements of your data when there is only a correlation. When your data