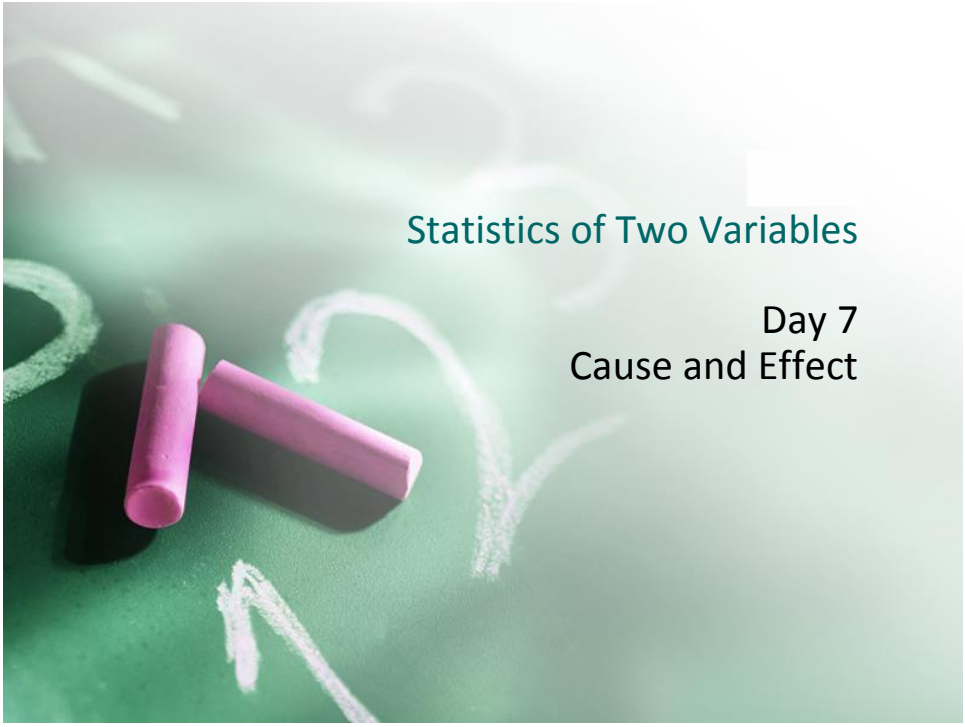


Statistics of Two Variables

Day 7
Cause and Effect

Cause and Effect

Usually the main reason for a correlational study is to prove that a change in X produces a change in Y.

- For example a school board may want to know whether calculators help students learn mathematics.

However a strong correlation does not prove that one variable causes the other to change.

- For example several years following World War II, you will find a strong positive correlation between the number of storks, and the annual number of human births in the city of Copenhagen.

Types of Causal Relationships

Cause and Effect Relationship

A change in X produces a change in Y.

- For example increasing the height from which you drop an object increases its impact velocity.
- OR increasing the speed of a production line increases the number of items produced (and, perhaps, the rate of defects).

Common-Cause Factor

An external variables causes both variables to change in the same way.

- For example a town finds that revenue from parking fees at the public beach each summer correlates with the sales at local restaurants.

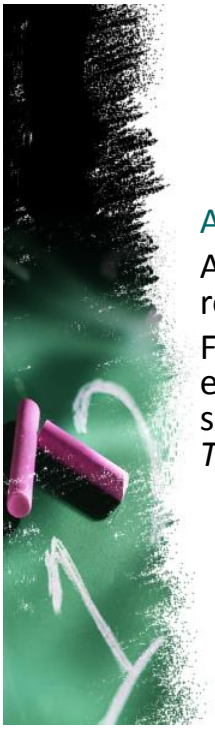
  Instead good weather is a common-cause factor that increases both variables.

Reverse Cause and Effect Relationship

The dependent and independent variables are reversed in the process of establishing causality.

- For example a researcher finds a positive linear correlation between the amount of coffee consumed by a group medical students and their levels of anxiety.

  The researcher theorizes that drinking coffee causes nervousness, but instead finds that nervous people are more likely to drink coffee.



Accidental Relationship

A correlation exists without any causal relationship.

For example, the number of females enrolled in engineering programs and the number of reality shows on TV both increased for several years. *The correlation is likely coincidental.*

Presumed Relationship

A correlation does not seem to be accidental even though no cause-and-effect relationship or no common-cause factor is apparent.

For example, suppose you found a correlation between people's level of fitness and the number of adventure movies they watched.

- It seems logical that a fit person might prefer action movies but it would be difficult to find a common cause or prove that one variable affects the other.

Example 1

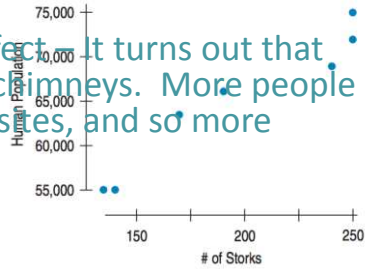
Classify the causal relationships in the following situations.

a) The rate of a chemical reaction increases with temperature.

Cause and Effect – higher temperatures cause faster reaction rates

b) A scatterplot of the number of storks in Oldenburg, Germany, plotted against the population of the town shows a strong positive correlation.

Reverse Cause and Effect – It turns out that storks nest on house chimneys. More people means, more nesting sites, and so more storks.



c) An increase in number of students enrolled in Data Management and increase in the number of BMW's on the road.

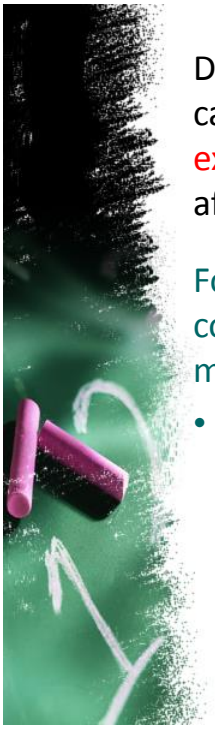
Accidental Relationship – The correlation between the number of students enrolled in data management and the number of BMW is coincidental.

d) The prices of butter and motorcycles have a strong positive correlation over many years.

Common-Cause Factor – Inflation has caused parallel increases in the prices of both items over the years.

e) Heart attack rates drop as fitness clubs bring in more revenue.

Presumed Relationship – A negative correlation seems logical, however there is no apparent common-cause factor or cause-and-effect relationship.



Determining the nature of causal relationships can be complicated by the presence of extraneous variables (or external variables) that affect one of the variables

For example you might see a strong positive correlation between term marks and final exam marks for students.

- Extraneous factors could affect exam results such as how much studying a student did, exam schedules, ability to work under pressure

To reduce the effect of extraneous variables researchers often compare and experimental group to a control group

- Groups should be as similar as possible
- Researchers vary the independent variable for the experimental group but not the control group

For example a researcher wants to test a new drug believed to help smokers overcome the addictive effects of nicotine. One group is given nicotine patches, while the second is given an ordinary patch (placebo).