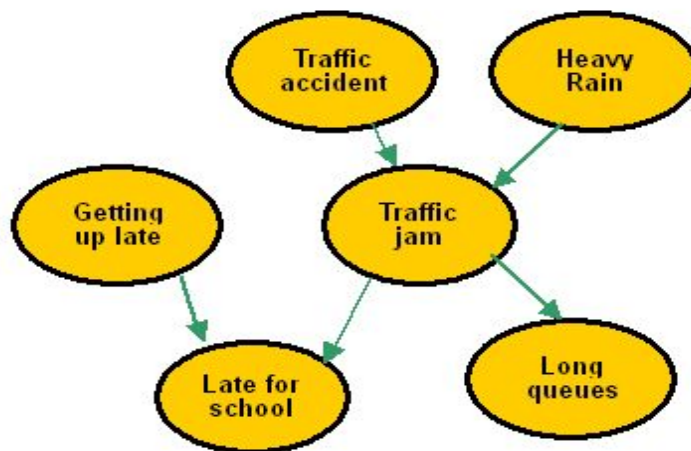


A **cause** is something that makes something else happen. Out of two events, it is the event that happens first. To determine the cause, ask the question "Why Did it Happen?"

An **effect** is what happens as a result of the cause. Of two related events, it's the one that happens second or last. To determine the effect, ask the question "What Happened?"

Causal Nets

Causal nets are used in representing causal connections and in calculating probabilistic relations between events. Here is a simple example of a causal net showing how some events are related to each other. The arrows indicate the direction of causation.



Correlation

Correlation means that two variables (sets of data) type of association with each other, such that as one variable increases, the other also increases (a positive correlation), decreases (a negative correlation).

A strong correlation does not prove that changes in one variable cause changes in the other.

Cause and Effect

It is tempting to assume that when two variables are positively correlated, then one causes the other (i.e., the variables have a "cause and effect" relationship) but this is not always the case.

Correlation and Causation

Data analysts often jump to unjustified conclusions by mistaking an observed correlation for a cause-effect relationship. A high sample correlation coefficient does not necessarily signify a causal relation between two variables. A classic example concerns an observed high positive correlation between the number of storks sighted and the number of births in a European city. Hopefully, no one would use this evidence to conclude that storks bring babies, or, worse yet, that killing storks would control population growth.

Another "classic" example: ice cream sales peak during the summer months; cases of heat stroke increase during the summer. Does this imply that eating ice cream causes heat stroke?

There are various types and degrees of casual relationships between variables.

1. **Cause-and-Effect Relationship:** A change in X produces a change in Y.
2. **Common-cause Factor:** A common-cause factor is one that influences two or more variables under study in the same way. When hidden, this can be considered to be a lurking variable. It is always a good idea to always ask "could there be a third factor that is actually causing both A and B?" Over the course of several weeks the needles from the pine trees along the Wombat River fell into the water. Shortly thereafter, many dead fish washed up on the river banks. When the EPA investigated, the owners of the Wombat River Chemical Company claimed that it was obvious that the pine needles had killed the fish. Many local environmentalists claimed that the chemical plant's toxic wastes caused both the trees and the fish to die and that the pine needles had no real effect on the fish.
3. **Reverse Cause and Effect Relationship:** This is a relationship in which the presumed dependent and independent variables are reversed in the process of establishing causality.
4. **Accidental Relationship**
This is a correlation between two variables that happens by random chance.
5. **Presumed Relationship:** This is a correlation that does not seem to be accidental even though no cause-and-effect relationship or common-cause factor is apparent.

Extraneous Variables

Extraneous Variables are **undesirable** variables that influence the relationship between the variables that an experimenter is examining. While these variables influence what is being studied, they are not the variables of interest. In order to reduce the effect of extraneous variables, results from an experimental group are compared to the results from a **control group**. These two groups should be as similar as possible, so that extraneous variables will have about the same effect on both groups, and any difference in the dependent variables for the two groups can then be attributed to the changes in the independent variable.

When designing a study or interpreting a correlation, background knowledge and insight are useful in recognizing the presence of a causal relationship.

Is the Correlation a result of a Cause-and-Effect Relationship?

Techniques that can help determine whether a correlation is the result of a cause-and-effect relationship:

- Use sampling methods that hold the extraneous variables constant
- Conduct similar investigations with different samples and check for consistency in the results
- Remove, or account for, possible common-cause factors