

## Cause and Effect

Usually, the main reason for a correlational study is to find evidence of a cause-and-effect relationship.

In a cause-and-effect relationship, a change in the independent variable (x) causes a change in the dependent variable (y). It is very difficult to prove causation. Correlation does not always imply a cause-and-effect relationship. Correlations can also result from:

**Common Cause Factors:** External variables that cause two variables to change in the same way.

For example, suppose that a town finds that its revenue from parking fees at the public beach each summer correlates with the local tomato harvest. It is very unlikely that the cars parked at the beach have any effect on the tomato crop. Instead, *good weather* is a common-cause factor that increases both the tomato crop and the number of people who park at the beach.

**Reverse Cause and Effect Relationships:**

The dependent and independent variables are reversed in the causal relationship.


For example, suppose that a researcher observes a positive linear correlation between the amount of coffee consumed by a group of 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 Relationships:** The correlation is by chance (no logical reason).

For example, the number of females enrolled in undergraduate engineering programs and the number of "reality" shows on TV have both increased in recent years. These two variables have a positive correlation but it is likely entirely coincidental.

**Presumed Relationships:** A correlation exists without any apparent cause-and-effect or common-cause factor.

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 physically fit person might like adventure movies but it would be difficult to find a common cause or prove causation in this relationship.



In addition to these factors, it is also important to be aware of any extraneous variables and outliers in a correlational study.



**Extraneous Variable:** An extraneous variable effects one or both of the variables in a correlational study.

For example, it could be shown that there is a positive linear correlation between a student's term mark and their final exam mark. An extraneous variable here would be the amount of time spent studying for the exam.

**Outlier:** A point that does not follow the pattern of the other points in a data set.

Outliers may be due to measurement error or due to too small a sample. If so , the point should be removed. If not, the study should comment on the significance of the outlier and how it influenced the line of best fit.

For example, in statistical studies involving Canadian provinces, the Territories are often excluded from regression analysis because their populations are too small in comparison to the provinces.



## Correlation is NOT Causation!

In your groups, read over each of these reported findings. As a group, first establish whether there is correlation between the two events and then decide whether the relationship is cause-effect, common-cause or accidental.

1. A higher number of ice cream sales corresponds to a higher number of shark attacks on swimmers.

Does this mean that increased ice cream sales CAUSES the number of shark attacks on swimmers to increase?

2. The number of cavities in elementary school children and vocabulary size have a strong positive correlation.

Does this mean that increasing the number of cavities in elementary school children CAUSES their vocabulary size to increase?

3. In a growing municipality, the traffic planner (who never completed Data Management class) observed that over a period of ten years the number of traffic accidents showed a high positive correlation with the number of traffic lights installed.

Was the planner correct in suggesting to the Mayor that they remove all the traffic lights to reduce the accident rate?

4. There is a strong, positive correlation between the number of fire engines responding to a fire and the damage caused by the fire. Does this suggest that reducing the number of responding fire engines will result in less fire damage? Why or why not?
5. Every time that I eat chocolate, I get acne. Does this mean that, for me, eating chocolate causes acne? Why or why not? Hint: Consider situations when I might want to eat chocolate and when I get acne (i.e. increased stress/anxiety).

6. It has been observed that the number of rings in a tree stump correspond roughly with the age of the tree.

7. Humans have 23 chromosome pairs. The earth's axis is tilted at approximately 23 degrees.

## Home work : Cause And Effect Relationships

### Communicate Your Understanding

1. Why does a strong linear correlation not imply cause and effect?
2. What is the key characteristic of a reverse cause-and-effect relationship?
3. Explain the difference between a common-cause factor and an extraneous variable.
4. Why are control groups used in statistical studies?

### Practise

1. Identify the most likely type of causal relationship between each of the following pairs of variables. Assume that a strong positive correlation has been observed with the first variable as the independent variable.

- a) alcohol consumption, incidence of automobile accidents
- b) score on physics examination, score on calculus examination
- c) increase in pay, job performance
- d) population of rabbits, consumer price index
- e) number of scholarships received, number of job offers upon graduation
- f) coffee consumption, insomnia
- e) funding for athletic programs, number of medals won at Olympic games

2. For each of the following common-cause relationships, identify the common-cause factor. Assume a positive correlation between each pair of variables.

- a) number of push-ups performed in one minute, number of sit-ups performed in one minute
- b) number of speeding tickets, number of accidents
- c) amount of money invested, amount of money spent

3. A civil engineer examining traffic flow problems in a large city observes that the number of traffic accidents is positively correlated with traffic density and concludes that traffic density is likely to be a major cause of accidents. What alternative conclusion should the engineer consider?

4. **Communication** An elementary school is testing a new method for teaching grammar. Two similar classes are taught the same material, one with the established method and the other with the new method. When both classes take the same test, the class taught with the established method has somewhat higher marks.

- a) What extraneous variables could influence the results of this study?
- b) Explain whether the study gives the school enough evidence to reject the new method.
- c) What further studies would you recommend for comparing the two teaching methods?

5. **Communication** An investor observes a positive correlation between the stock price of two competing computer companies. Explain what type of causal relationship is likely to account for this correlation.

6. **Application** A random survey of students at Statsville High School found that their interest in computer games is positively correlated with their marks in mathematics.

- a) How would you classify this causal relationship?
- b) Suppose that a follow-up study found that students who had increased the time they spent playing computer games tended to improve their mathematics marks. Assuming that this study held all extraneous variables constant, would you change your assessment of the nature of the causal relationship? Explain why or why not.

7. a) The net assets of Custom Industrial Renovations Inc., an industrial construction contractor, has a strong negative linear correlation with those of MuchMega-Fun, a toy distributor. How would you classify the causal relationship between these two variables?
- b) Suppose that the two companies are both subsidiaries of Diversified Holdings Ltd, which often shifts investment capital between them. Explain how this additional information could change your interpretation of the correlation in part a)

8. **Communication** Aunt Gisele simply cannot sleep unless she has her evening herbal tea. However, the package for the tea does not list any ingredients known to induce sleep. Outline how you would conduct a study to determine whether the tea really does help people sleep.

## ANSWERS :

### Practise

1. a) cause-and-effect relationship: alcohol consumption impairs driving ability
  - b) common-cause factor: achievement in physics and calculus requires similar skills
  - c) reverse cause-and-effect relationship: better job performance leads to increases in pay
  - d) accidental relationship: no causal relationship between the variables
  - e) presumed relationship: seems logical that a student who has obtained a number of scholarships would be attractive as an employee, but there are many other qualities that employers seek
  - f) cause-and-effect relationship: coffee consumption keeps people awake
  - g) reverse cause-and-effect relationship: higher number of medals won at Olympic games encourages investors to fund athletic programs
2. a) overall athletic ability
  - b) overall disregard of safe driving practices
  - c) a large income

### Apply, Solve, Communicate

3. Traffic accidents cause traffic congestion.
4. a) the teachers of the class
- b) no
- c) Conduct further trials giving teachers time to become comfortable with the new teaching method and examine the results of classes where the same teacher has used the different methods to minimize the effect of this possible extraneous variable.
5. The overall increase in computer use is likely to have caused parallel increases in the fortunes of the companies.
6. a) Accidental relationship is likely.
7. a) Accidental relationship is likely.
- b) cause-and-effect relationship
8. Answers may vary.