Statistics
Interpreting Categorical and Quantitative Data
Correlation vs. Causation
High School

Rationale

- Many students do not have a strong concept of correlation and may never have taken time to consider when circumstances are truly causal or merely correlated. The significance of the differences between correlation and causation has implications not only in the fields of mathematics and science but well beyond the classroom walls. This lesson is designed to help students question the existence of correlation and the possibility of correlation without causation. This will eventually lead to a deeper appreciation for the study of statistics.

Goal

- To create a concept of correlation and draw a distinction between causation, so that circumstance can be considered and applied in particular situations

Standards

- S-ID. 9. Distinguish between correlation and causation.

Objectives

- Students will understand what correlation means.
- Students will understand what causation means.
- Students will understand the difference between correlation and causation.
- Students will consider some situations and try to determine if there is merely a correlation or if there is truly causation.

Materials

- A class list
- Scrap paper

Procedures

- Initiate a discussion (see **Teacher Tips**) regarding the terms *correlation* and *causation.* Ask students to define the terms in their own words. Encourage students to consider the roots of each word, *correlate* and *cause*, to help direct them to an appropriate definition. Given the importance of the terms for this lesson, it may help to write the students' definitions on the board. Do not correct mistakes at this point; there will be an opportunity for revision.
- Share this situation with your students: "A survey is taken at an elementary school comparing shoe size and reading ability." Then ask, "Will there be a correlation?" For students who volunteer their answers, ask them to explain why the data is

correlated, or why it is not. Students should generally agree that students with larger shoes *will* have a higher reading ability, meaning that the events *are* correlated.

- After that discussion has subsided, ask, "Does a larger shoe size *cause* a better reading ability?" Students should insist that it does not! Then ask, "Does learning to read *cause* your feet to grow?" Again, students will likely insist that there is not a cause and effect relationship.
- Now ask your students if they would like to modify their definitions of correlation and causation. By referencing the example of the elementary school survey they should be able to paint a pretty clear picture of correlation and causation.
- Ask your students, "If you counted the number of ant traps in houses and compared it to the number of ants in the same houses, what relationship would you expect to see?" Students will likely respond that there would be a lot more ant traps in houses that have ant problems, but if there are no ants in the house, then you probably would not find any ant traps. Now ask, "Do ant traps cause ants?" and "Do ants cause ant traps?" Students will say, no! Finally, ask, "What do we need to be careful about when we find a correlation?" Students need to realize that a cause and effect will certainly be correlated, but a correlation may not be a result of a cause and effect. They need to make sure that the causation points in the right direction!
- Now mention the importance (and difficulty) in determining causation in real-world scenarios. Ask them whether smoking causes lung cancer. Students' answers will vary. Most will insist yes, but some may question it. Now say, "Scientists believed for a long time that smoking *did* cause lung cancer, but warnings said things like: 'Smoking may cause lung cancer.' or 'Cigarette Smoking May be Hazardous to Your Health.' Why did these warnings use the word *may*?" Students will discuss that there was a correlation, but they were unable to show causation. Then tell students, "In 1985 this message first appeared on boxes of cigarettes: 'SURGEON GENERAL'S WARNING: Smoking Causes Lung Cancer, Heart Disease, Emphysema, And May Complicate Pregnancy.' Why did the wording change in the warning?" Students will hopefully realize that researchers were able to definitively show that smoking greatly increased the risk of getting lung cancer.
- Tell your students, "Researchers found links between coffee and heart disease. Is this relationship a correlation or is it causation?" Once students have responded, ask them, "If this is merely a correlation, what may be leading to this correlation?" Once students have identified some behaviors, like links between coffee and other addictive behaviors, then ask them what the implications of a causation relationship between coffee and heart disease are. Students should point out that it would mean fewer coffee customers for health-conscious reasons. Now ask, "What could be done to show that it is merely a correlation, or to show definitively that coffee causes heart disease?" Allow students to discuss ideas like the scientific method, having control groups or looking at the medical records of coffee drinkers and non-coffee drinkers. Eventually you can point out that statistics would need to be studied and quantitative comparisons would need to be made. A discussion about climate change can highlight the difficulties that

scientists are having about the direct effect of humankind's impact on the temperature of the planet.

- Next you can ask students to think about and discuss some other situations where it is unclear whether there is a correlation or a causation relationship. Write down their ideas as topics for a possible extension activity.
- Finally, tell students, "It has been shown that areas with unhealthy foods are correlated with overweight residents. Does having this unhealthy food available cause obesity?" Encourage thoughtful, considerate conversations between students.
- Closing: Remind students again, "You cannot assume there is causation (even when it seems to make sense). Careful research and studies are required to *prove* causation!"

Teacher Tips

- Before the lesson, be sure to lay the ground rules for a discussion. Encourage students to develop their own set of rules for a collaborative, safe discussion, or you can provide your own guidelines. Here are some suggestions:
  - o Kindness and courtesy are first in every discussion.
  - o Wait patiently for someone to finish speaking before you speak.
  - o Allow people the opportunity to respond to a comment before you begin a new conversation.
  - o Write down anything that seems interesting to you.
  - o If you have a thought, then write it down and wait for an appropriate time to add it to the conversation.
  - o *Every* student should speak at least twice during the discussion.
- You can mention previous mistakes made considering causality. Spicy foods were often correlated with a higher ulcer rate. However, explain to students that even seemingly obvious causes may be a mere correlation. In 1982, doctors were able to show that *H. pylori*, a type of bacteria found in many peoples' stomachs, was the cause of most ulcers, rather than spicy foods. Spicy foods often caused the symptoms to be more painful and evident, though. This led people to falsely believe in a causal relationship for a long time before these bacteria were discovered.
- Make sure to get your quiet students involved in the conversation early. Address them personally on one of your questions in order to draw them in.
- You may want to require students to offer at least two comments or questions to the conversation and even attach a grade to their participation. You will need to keep careful track on a class list if this is the case.
- You want to encourage students to write their ideas or questions throughout the discussion so that they may be shared at an appropriate time. Tell your students that they must avoid interrupting their classmates, but they may write their comments and speak when their peer is finished speaking.

Extension Activities

- Ask students to describe events in their own lives where causation is occurs (such as studying and test scores) and where perhaps only correlation occurs (such as grade level and ownership of a driver's license).
- Have students research some of the topics discussed previously. They can prepare a brief report on what researchers were doing to show causation in particular situations.
- Have students research some creative correlation vs. causation examples they can find online. Many websites offer some circumstances that are correlated, but definitely not linked causally. Students can also come up with examples of causal relationships. It is often funny to take an obviously causal relationship, point out the correlation, and then ask if the causation makes sense in the opposite direction!