

ACTIVITY 4: Correlation versus Causation

OVERVIEW: In this activity, students will be introduced to the concepts of correlation versus causation, lurking variables, confounding, and common response. They will then examine some unusual linear relationships on gapminder.org and use their knowledge and critical thinking skills to determine whether a common response or confounding is at play. Some examples include mobile cellular subscriptions (per 100 people) and babies per woman which shows a negative trend, and smoking (% of population) versus life expectancy which shows a moderate, positive trend. Finally, students will evaluate the relationships they explored in previous activities and decide whether there was a common response or confounding for their data. They will then add this information to a final slide in their PowerPoint presentation.

MATERIALS NEEDED TO CARRY OUT THE LESSON:

- In-Class Activity Worksheet
- Laptop with PowerPoint

CONCEPTS TO BE LEARNED/APPLIED:

- Students will understand the difficulties of establishing causation.
 - Students will understand that correlation does not imply causation.
 - Students will understand that a common response is when two variables appear to be related when in actuality there is an outside variable causing the correlation.
 - Students will understand that confounding is when multiple variables, including the independent variable studied, will have an impact on the dependent variable.

INSTRUCTIONAL PLAN:

Before handing out the activity worksheet, go over the following vocabulary.

Correlation is when there is a covariational relationship between two quantitative variables.

Causation is when one variable, the independent variable, causes a change in the other variable, the dependent variable.

A **lurking variable** is a variable that is not of interest in the study but can have an impact on the independent and/or dependent variable being studied.

A **common response** is when a lurking variable causes changes in both the independent variable and the dependent variable, creating a correlation (i.e. an association) between the two variables when there is no direct causal relationship.

Confounding is when multiple lurking variables, including the independent variable, have an impact on the dependent variable, thus weakening the relationship between the two variables of interest.

Once you have presented the vocabulary, the students can start working on Activity 4. After they have finished the activity have an all-class discussion.

After the discussion, the students can complete the final slide of their PowerPoint presentation, which will include a discussion of whether they had a common response or confounding and what some of the lurking variables might be.

The groups should be ready to present their data to the class. If you choose to have the groups present, a sample grading scheme for the presentation is provided in the CoRD.

MIP COMPONENTS OF INQUIRY

Active Learning

- In Question 1, students have to select the definitions of form, direction, and strength from their prior knowledge, evaluate whether the data meet their definition, and perform the aforementioned classifications of the scatterplot.
- In Question 4, students select the correct vocabulary from their prior knowledge of variables (e.g., lurking variables, confounding), evaluate whether the scenario meets their definition, and classify the possible “outside” variables.
- In Question 5, students interpret the information in a scatterplot by selecting the data they need, evaluating if a larger value in the independent variable appears to produce a larger value in the dependent variable, and then forming a conclusion.
- In Question 7, students select the correct vocabulary from their prior knowledge of variables (e.g., lurking variables, confounding), evaluate whether the scenario meets their definition, and classify possible “outside” variables.
- In Question 8, students have to select the definitions of form, direction, and strength from their prior knowledge, evaluate whether the data meet their definition, and perform the aforementioned classifications of the scatterplot.
- In Question 9, students interpret the information in a scatterplot by selecting the data they need, evaluating how the dependent variable changes as the independent variable increases, and then forming a conclusion.
- In Question 11, students select the correct vocabulary from their prior knowledge of variables (e.g., lurking variables, confounding), evaluate whether the scenario meets their definition, and classify possible “outside” variables.
- In Question 12, students have to select the definitions of form, direction, and strength from their prior knowledge, evaluate whether the data meet their definition, and perform the aforementioned classifications of the scatterplot.
- In Question 16, students have to select the definitions of form, direction, and strength from their prior knowledge, evaluate whether the data meet their definition, and perform the aforementioned classifications of the scatterplot.

Meaningful Applications

- In Question 2, students make and justify claims regarding the effect of an increase in the value of the independent variable on the dependent variable.
- In Question 3, students make and justify claims regarding the validity of the independent variable directly impacting the dependent variable.
- In Question 4, students use their critical thinking skills to make and justify claims regarding other potential variables that would impact the dependent variable.
- In Question 7, students make and justify claims regarding the reason no relationship was found between the two variables.
- In Question 11, students make and justify claims regarding whether a larger value in the independent variable causes a larger value in the dependent variable.
- In Question 13, students make and justify claims regarding whether an increase in the value of the independent variable causes a decrease in the value of the dependent variable.
- In Question 14, students use their critical thinking skills to make and justify claims regarding other potential variables that would impact the dependent variable.
- In Question 15, students make and justify claims to explain the moderate correlation between the two variables.
- In Question 17, students make and justify claims as to whether a larger value in the independent variable causes a larger value in the dependent variable.

Academic Success Skills

- In Question 6, students are asked to provide their thoughts on whether having more freedom in a country should impact the overall happiness score of the country. Allowing students to communicate their thoughts promotes mathematical creativity and promotes student confidence.
- In Question 14, students use their critical thinking skills to come up with variables that will likely impact the dependent variable. Students will be able to use personalized knowledge to answer this question and hence develop a sense of ownership and connection to the content, hence enhancing their identities as learners.
- Throughout this activity, students use their critical thinking skills to interpret the data of empirical examples and determine if other variables might impact the dependent variable. This allows students to use their creativity which may boost student confidence and hence create a positive change to their identity as mathematics learners.