

## 3.25: Causation and Lurking Variables (1 of 2)

### Learning Objectives

- Distinguish between association and causation. Identify lurking variables that may explain an observed relationship.

### Introduction

A common mistake people make when describing the relationship between two quantitative variables is that they confuse *association* and *causation*. This mistake is so common that we devote this entire section to clarifying the difference.

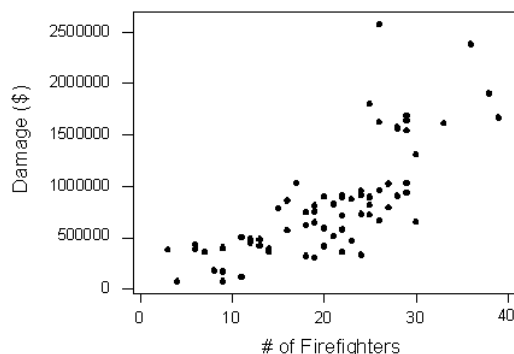
This confusion often occurs when there is a strong relationship between the two quantitative variables. In the case of a linear relationship, people mistakenly interpret an  $r$ -value that is close to 1 or -1 as evidence that the explanatory variable *causes* changes in the response variable. In this case, the *correct interpretation* is that there is a **statistical relationship** between the variables, not a causal link. In other words, the explanatory variable and the response variable vary together in a predictable way. There is an **association** between the variables. But this *should not* be interpreted as a cause-and-effect relationship.

Let's look at an example.

### Example

#### Fire Damage

The scatterplot below shows the relationship between the number of firefighters sent to fires ( $x$ ) and the amount of damage caused by fires ( $y$ ) in a certain city.

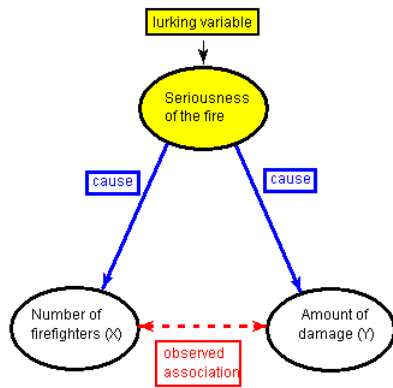


The scatterplot shows a positive association with a somewhat strong curvilinear form. An increase in the number of firefighters is associated with an increase in the damage done by the fire.

Can we conclude that the increase in firefighters causes the increase in damage? Of course not.

A third variable is at play in the background – the seriousness of the fire – and is responsible for the observed relationship. More serious fires require more firefighters and also result in more damage.

The following figure will help you visualize this situation:



The seriousness of the fire is a **lurking variable**. A lurking variable is a variable that is not measured in the study. It is a third variable that is neither the explanatory nor the response variable, but it affects your interpretation of the relationship between the explanatory and response variables.

In our example, the lurking variable has an effect on both the explanatory and the response variables. This common effect creates the observed association between the explanatory and response variables even though there is no cause-and-effect link between them.

Try It

<https://assessments.lumenlearning.co...sessments/3851>

Try It

<https://assessments.lumenlearning.co...sessments/3852>

### Contributors and Attributions

CC licensed content, Shared previously

- Concepts in Statistics. **Provided by:** Open Learning Initiative. **Located at:** <http://oli.cmu.edu>. **License:** *CC BY: Attribution*

This page titled [3.25: Causation and Lurking Variables \(1 of 2\)](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Lumen Learning](#).