

1.1: What is Statistics?

Learning Objectives

- Identify situations in which statistics can be misleading
- Define statistics

Introduction to Statistics

Statistics includes numerical facts and figures. For instance:

- The [largest earthquake](#) measured 9.5 on the Richter scale.
- In 2023, 83% of [all adult homicide victims](#) were male.
- in 2023, about a [quarter of women of reproductive age in South Africa were HIV positive](#).
- By the year 2050, there will be [82 million people aged 65 and over in the United States](#), a 47% increase since 2022.

The study of statistics involves math and various calculations, but as a body of knowledge, statistics is built upon much more. Statistics includes theoretical frameworks that guide the formulation of questions and the data collection, analysis, and interpretation needed to answer those questions. Consider the following three scenarios where interpretations are given based on presented statistical measures. You will find that the numbers may be correct, but the interpretation may be wrong. Try to identify a major flaw with each interpretation on your own, and then check your response.

Text Exercise 1.1.1

A new advertisement for Ben and Jerry's ice cream was introduced in late May of last year. The following three months saw a 30% increase in ice cream sales; thus, the advertisement was effective.

Answer

A major flaw is that ice cream sales generally increase in the months of June, July, and August, regardless of advertisements. This effect is called a history effect and leads people to interpret outcomes as the result of one variable when another variable (in this case, the time of the year) is actually responsible.

Text Exercise 1.1.2

The more churches there are in a city, the more crime there is. Thus, churches lead to crime.

Answer

A major flaw is that both increased churches and increased crime rates can be explained by larger populations. In bigger cities, there are both more churches and more crime. This problem refers to the third-variable problem; namely, people [erroneously](#) believe that there is a causal relationship between the two primary variables rather than recognize that a third variable can cause both.

Text Exercise 1.1.3

At Harvard, the percentage of seniors that graduated with a GPA of 4.0 increased nearly 78% from 2020 to 2023. Thus, grade inflation is a real epidemic.

Answer

A major flaw is that we don't have the information that we need. What are the actual rates of occurrence? Suppose only 1% of students earned a 4.0 GPA in 2020 and 1.78% of graduating seniors earned such a GPA in 2023. 1.78 is 78% higher than 1. But this latter number is hardly evidence suggesting an epidemic in grade inflation. In addition, the statistic provided does not rule out the possibility that the number of 4.0 GPAs had seen dramatic fluctuations in those years due to a variety of different causes. Again, there is simply not enough information to fully understand the impact of the statistic. If you're interested, read more about the statistics of graduating seniors [here](#).

As a whole, these examples show that "statistics" are not only facts and figures. In the broadest sense, **statistics** refers to a large range of techniques and procedures for analyzing, interpreting, displaying, and making decisions based on data.

1.1: What is Statistics? is shared under a [Public Domain](#) license and was authored, remixed, and/or curated by The Math Department at Fort Hays State University.

- 1.1: What are Statistics? by [David Lane](#) is licensed [Public Domain](#). Original source: <https://onlinestatbook.com>.