

1.5: Populations and Samples

Because social scientists want to make people's lives better, they see what works on a small group of people, and then apply it to everyone. The small group of people is the sample, and "everyone" is the population.

Definition: Sample

People who participate in a study; the smaller group that the data is gathered from.

A sample is the small group of people that scientists test stuff on. We want at least 30 people in each group, so a study that has two groups will need about 60 people in the sample.

Definition: Population

The biggest group that your sample can represent.

A population is the "everyone" that we want to apply the results to. Sometimes, "everyone" can be a pretty small group; if I measured the GPA of one of my Behavioral Statistics classes, then the sample would be the class and the population could be students in all Behavioral Statistics classes at the college. (GPA would be the DV.)

A sample is a concrete thing. You can open up a data file, and there's the data from your sample. A population, on the other hand, is a more abstract idea. It refers to the set of all possible people, or all possible observations, that you want to draw conclusions about, and is generally *much* bigger than the sample. In an ideal world, the researcher would begin the study with a clear idea of what the population of interest is, since the process of designing a study and testing hypotheses about the data that it produces does depend on the population about which you want to make statements. However, that doesn't always happen in practice: usually the researcher has a fairly vague idea of what the population is and designs the study as best he/she can on that basis.

Examples

In our Scientific Method example, the sample would be the class from which we got the data from, and the population would be the biggest group that they could represent. There's often more than one possible population, but I might say all college students could be a good population for this sample.

You might have heard the phrase "random sample." This means that everyone in the population has an equal chance of being chosen to be in the sample; this almost never happens.

? Exercise 1.5.1

Let's say I want to know if there's a relationship between intelligence and reading science fiction books. If I survey 100 of my Introduction to Psychology students on their intelligence and reading of science fiction:

1. Who is the Sample?
2. Who could be the population? In other words, what is the biggest group that this sample could represent?

Answer

Add texts here. Do not delete this text first.

1. Who is the Sample? 100 of my Introduction to Psychology students
2. Who could be the population? In other words, what is the biggest group that this sample could represent? There are many possible populations, but all Introduction to Psychology students could work, or all Introduction to Psychology students at my college might make sense, too.

Sometimes it's easy to state the population of interest. In most situations the situation is much less simple. In a typical a psychological experiment, determining the population of interest is a bit more complicated. Suppose Dr. Navarro ran an experiment using 100 undergraduate students as participants. Her goal, as a cognitive scientist, is to try to learn something about how the mind works. So, which of the following would count as "the population":

- All of the undergraduate psychology students her university in Australia?
- Undergraduate psychology students in general, anywhere in the world?
- Australians currently living?
- Australians of similar ages to my sample?
- Anyone currently alive?
- Any human being, past, present or future?
- Any biological organism with a sufficient degree of intelligence operating in a terrestrial environment?
- Any intelligent being?

Each of these defines a real group of mind-possessing entities, all of which might be of interest to me as a cognitive scientist, and it's not at all clear which one ought to be the true population of interest. Maybe surprisingly for you, there's no "right" answer! Although some the suggestions get a little vague, they all could potentially be a population that her sample represents. Irrespective of how the population is defined, the critical point is that the sample is a subset of the population. The goal of researchers is to use our knowledge of the sample to draw inferences about the properties of the population. More on that in later chapters!

? Exercise 1.5.1

Actual drug use is much higher than drug arrests suggest, so you might want to measure how many people use marijuana. If you send out a survey asking about their drug use to everyone with a driver's license in California, but only 30% fill it out:

1. Who is the Sample?
2. Who could be the population? In other words, what is the biggest group that this sample could represent?

Answer

Add texts here. Do not delete this text first.

1. Who is the Sample? 30% of Californians with driver's licenses
2. Who could be the population? In other words, what is the biggest group that this sample could represent? There are many possible populations, but Californians who have driver's licenses might make the most sense here.

This last example shows that sometimes our sample limits who we can generalize our results about, who could be our population.

In almost every situation of interest, what we have available to us as researchers is a sample of data. We might have run experiment with some number of participants; a polling company might have phoned some number of people to ask questions about voting intentions; etc. Regardless: the data set available to us is finite, and incomplete. We can't possibly get every person in the world to do our experiment; a polling company doesn't have the time or the money to ring up every voter in the country etc.

This page titled [1.5: Populations and Samples](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Michelle Oja](#).

- [10.1: Samples, Populations and Sampling](#) by [Danielle Navarro](#) is licensed [CC BY-SA 4.0](#). Original source: <https://bookdown.org/ekothe/navarro26/>.
- [Current page](#) by [Michelle Oja](#) is licensed [CC BY-SA 4.0](#).