

2.2: Who are your participants? Who is your population?

Research in psychology typically begins with a general question about a specific group (or groups) of individuals or animals. For example, a researcher might want to know how many homeless people live on the streets of Phoenix. Or a researcher might want to know how often married people have sex, as reported by partners separately. In the first example, the researcher is interested in the group of homeless people. In the second example, the researcher may study heterosexual couples and compare the group of men with the group of women. In statistics, we call the entire group that a researcher wishes to study a **population**. As you can well imagine, a population can be quite large; for example, any student enrolled in college. A researcher might be more specific, limiting the population for a study to college students who have successfully completed a statistics course and who live in the United States.

Populations can obviously vary in size from extremely large to very small, depending on how the researcher defines the population. The population being studied should always be identified by the researcher. In addition, the population can include more than people and animals. A population could be corporations, parts produced in a factory, or anything else a researcher wants to study. Because populations tend to be very large it usually is impossible for a researcher to examine every individual in the population of interest. It is typically not feasible to collect data from an entire population. Therefore, researchers typically select a smaller, more manageable group from the population and limit their studies to the individuals in the selected group. A smaller more manageable group, known as a **sample**, is used to measure populations.

The participants in the research are the **sample**, and the larger group the sample represents is the **population**. In statistical terms, a set of individuals selected from a population is called a **sample**. A sample is intended to be representative of its population, and a sample should always be identified in terms of the population from which it was selected. As with populations, samples can vary in size. For example, one study might examine a sample of only 10 autistic children, and another study might use a sample of more than 10,000 people who take specific cholesterol medication. The sample is intended to represent the population in a research study.

When describing data it is necessary to distinguish whether the data come from a population or a sample.

- If data describe a **sample** it is called a **statistic**.
- If data describe a **population** it is called a **parameter**.

If I had given a statistical attitudes survey to the class, the class would be my sample. I might be interested in all students taking a statistics class for the first time, generalizing my findings to all statistics students would be applying information from my sample to a population. While it might be convenient for me to ask my class, does my class best represent all students taking statistics? I would need to carefully consider selecting the best sample for a population or critically think about the limits for generalizing my findings to a population. While our results would be most accurate if we could study the entire population rather than a sample from it, in most research situations this is not practical. Moreover, research usually is to be able to make generalizations or predictions about events beyond your reach. Additionally, sampling is an important concept to consider with the big picture of understanding statistics.

Imagine that we wanted to see if statistics anxiety was related to procrastination. We could measure everyone's levels of statistics anxiety and procrastination and observe how strongly they were related to each other. This would, however, be prohibitively expensive. A more convenient way is to select a number of individuals randomly from the population and find the relationship between their statistics anxiety and procrastination levels. We could then generalize the findings from this sample to the population. We use statistics, more specifically inferential statistics, to help us generalize from a particular sample to the whole population. Understanding the relationship between populations and their samples is the first vital concept to grasp in this course. Remember that the research started with a general question about the population but to answer the question, a researcher studies a sample and then generalizes the results from the sample to the population.

As we move into further concepts in statistics, we will see that how you get your participants (sampling) and sample size are important. The general rule is to get a large enough sample size and have the sample be a good representation of your population.

2.2: Who are your participants? Who is your population? is shared under a [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/) license and was authored, remixed, and/or curated by LibreTexts.