

What is the difference between population and sample?

Authored by
stats writer

April 22, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the difference between population and sample?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138187>

The difference between population and sample is that a population refers to the entire group of individuals or objects that a researcher is interested in studying, while a sample is a smaller subset of the population that is selected for the purpose of conducting research. The population represents the total number of elements that possess certain characteristics, while the sample is a representative portion of the population that is used to make inferences about the entire population. A population is typically larger and more diverse, while a sample is smaller and more manageable for research purposes. Additionally, the characteristics of a population are usually more stable and consistent, while a sample may vary in its composition and characteristics. It is important for researchers to carefully select a sample that accurately represents the population in order to draw valid conclusions.

Population vs. Sample: What's the Difference?

Often in statistics we're interested in collecting data so that we can answer some research question.

For example, we might want to answer the following questions:

- 1. What is the median household income in Miami, Florida?**
- 2. What is the mean weight of a certain population of turtles?**
- 3. What percentage of residents in a certain county support a certain law?**

In each scenario, we are interested in answering some question about a population, which represents every

possible individual element that we're interested in measuring.

However, instead of collecting data on every individual in a population we instead collect data on a sample of the population, which represents a portion of the population.

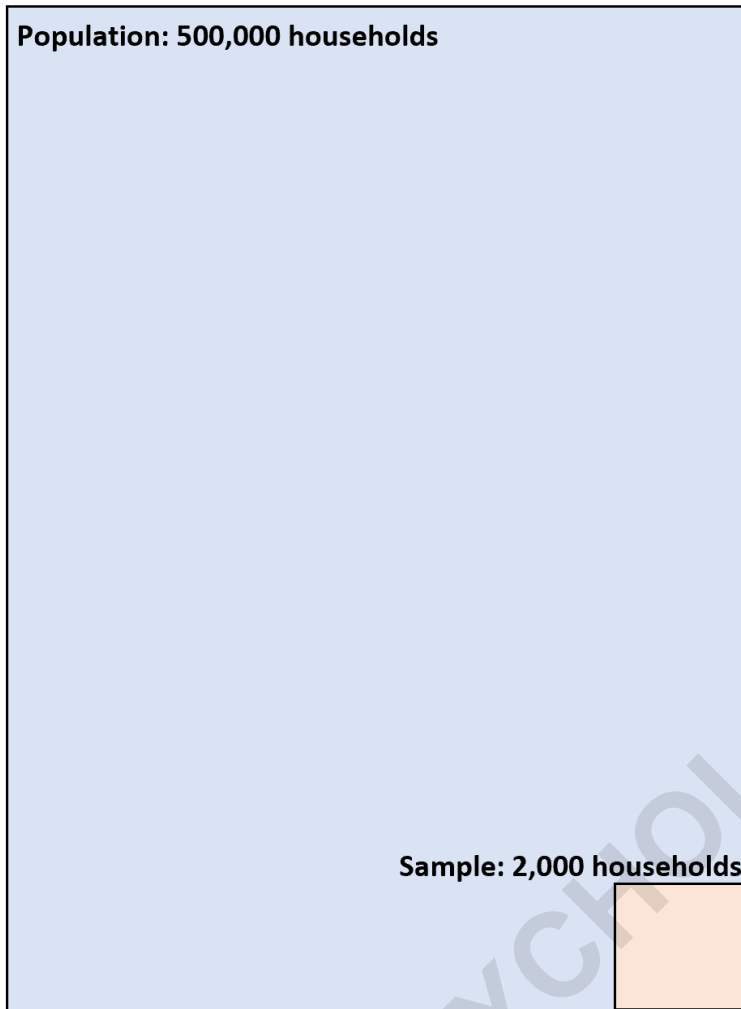
Population: Every possible individual element that we are interested in measuring.

Sample: A portion of the population.

Here is an example of a population vs. a sample in the three intro examples.

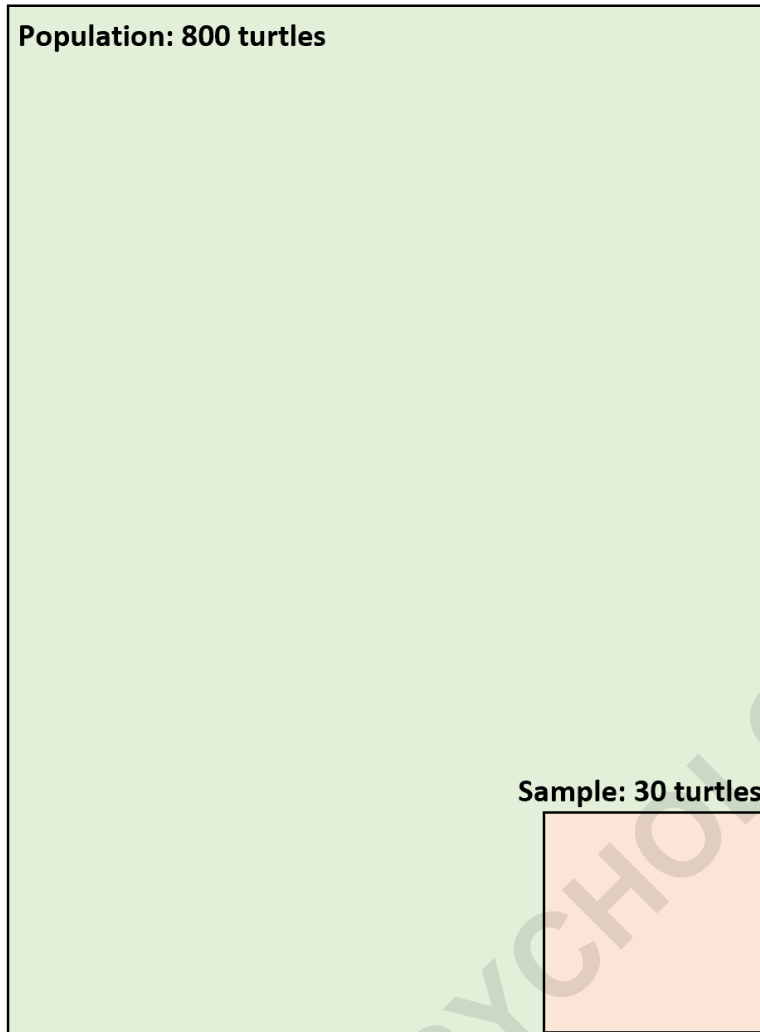
Example 1: What is the median household income in Miami, Florida?

The entire population might include 500,000 households, but we might only collect data on a sample of 2,000 total households.



2. What is the mean weight of a certain population of turtles?

The entire population might include 800 turtles, but we might only collect data on a sample of 30 turtles.



3. What percentage of residents in a certain county support a certain law?

The entire population might include 50,000 residents, but we might only collect data on a sample of 1,000 residents.

Why Use Samples?

There are several reasons that we typically collect data

on samples instead of entire populations, including:

1. It is too time-consuming to collect data on an entire population. For example, if we want to know the median household income in Miami, Florida, it might take months or even years to go around and gather income for each household. By the time we collect all of this data, the population may have changed or the research question of interest might no longer be of interest.

2. It is too costly to collect data on an entire population. It is often too expensive to go around and collect data for every individual in a population, which is why we instead choose to collect data on a sample instead.

3. It is unfeasible to collect data on an entire population. In many cases it's simply not possible to collect data for every individual in a population. For example, it may be extraordinarily difficult to track down and weigh every turtle in a certain population that we're interested in.

By collecting data on samples, we're able to gather information about a given population much faster and cheaper.

And if our sample is representative of the population, then we can generalize the findings from a sample to the larger population with a high level of confidence.

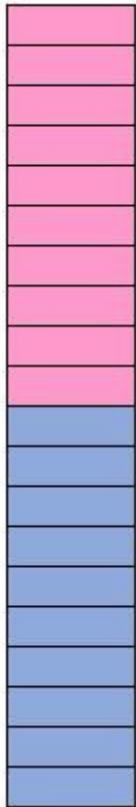
The Importance of Representative Samples

When we collect a sample from a population, we ideally want the sample to be like a "mini version" of our population.

For example, suppose we want to understand the movie preferences of students in a certain school district that has a population of 5,000 total students. Since it would take too long to survey every individual student, we might instead take a sample of 100 students and ask them about their preferences.

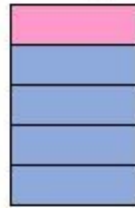
If the overall student population is composed of 50% girls and 50% boys, our sample would not be representative if it included 90% boys and only 10% girls.

Population



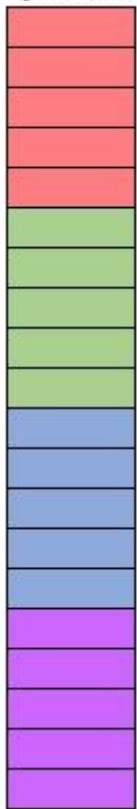
Sample is not representative of population

Sample



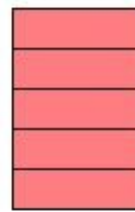
Or if the overall population is composed of equal parts freshman, sophomores, juniors, and seniors, then our sample would not be representative if it only included freshman.

Population



Sample is not representative
of population

Sample



A sample is representative of a population if the characteristics of the individuals in the sample closely matches the characteristics of the individuals in the overall population.

When this occurs, we can generalize the findings from the sample to the overall population with confidence.

How to Obtain Samples

There are many different methods we can use to obtain samples from populations.

To maximize the chances that we obtain a representative sample, we can use one of the three following methods:

Simple random sampling: Randomly select individuals through the use of a random number generator or some means of random selection.

Systematic random sampling: Put every member of a population into some order. Choose a random starting point and select every n th member to be in the sample.

Stratified random sampling: Split a population into groups. Randomly select some members from each group to be in the sample.

In each of these methods, every individual in the population has an equal probability of being included in the sample. This maximizes the chances that we obtain a sample that is a "mini version" of the population.