Population Vs Sample

There are two types of data sets available: Population and Sample. Whenever we do any mathematical calculation like mean, mode, or variance, we should know the dataset we are performing this operation, like the population type or sample type.

1. Population :

Population refers to the group of people in real life. But in statistics, it can be referred to as a group of individuals, objects, events, organisations, etc. So population is the entire set of items from which you draw data for a statistical study.

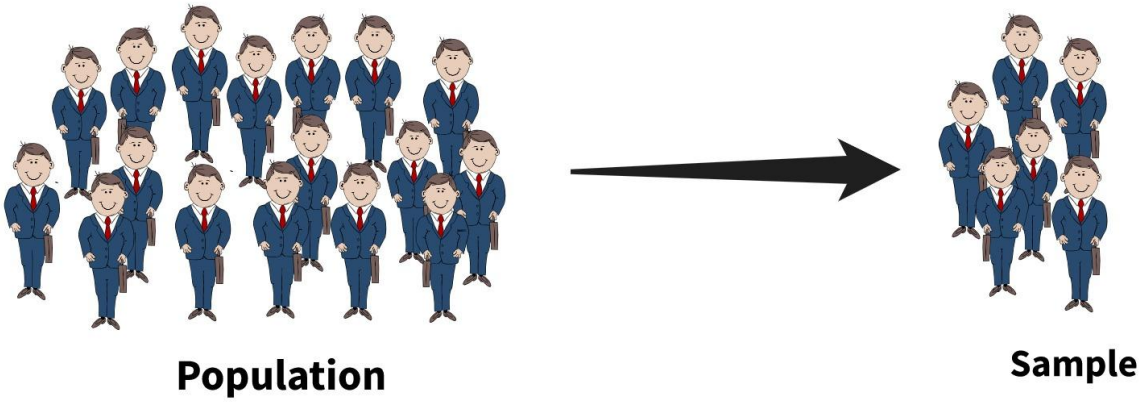
Let's suppose you want to collect the data from a company. The data consist of the employee details. This data will be easy to collect, and you will be able to give a correct outcome at the end. But what if you want to solve some problem and require the whole state population data. Which will be large in size.

So operation on the dataset is useful when we have medium-sized data available. You can perform the statistics to get the result, for example, School's Studnet Data, Employee detail, Single supermarket sales etc.

2. Sample :

It represents the set of people from the population. In other terms, you can say that they are the subset of the population that represents the whole dataset. You can perform the statistic operation on the sample dataset, and then you can generalise the same for the population.

Like, in the above example of a state dataset, you can take a small portion of the people from different age groups and then you could have generalised the result to the whole state population.



Sampling is the process of collecting data from a small subsection of the population and then using it to generalise over the entire set**.**