

DATA SAMPLING

Population

Population is the collection of the elements which has some or the other characteristic in common. Number of elements in the population is the size of the population.

Sample

Sample is the subset of the population. The process of selecting a sample is known as sampling. Number of elements in the sample is the sample size.

Sampling

Sampling is the selection of a subset (a **statistical sample**) of individuals from within a **statistical population** to estimate characteristics of the whole population.



TYPES OF SAMPLING

There are lot of sampling techniques which are grouped into two categories as

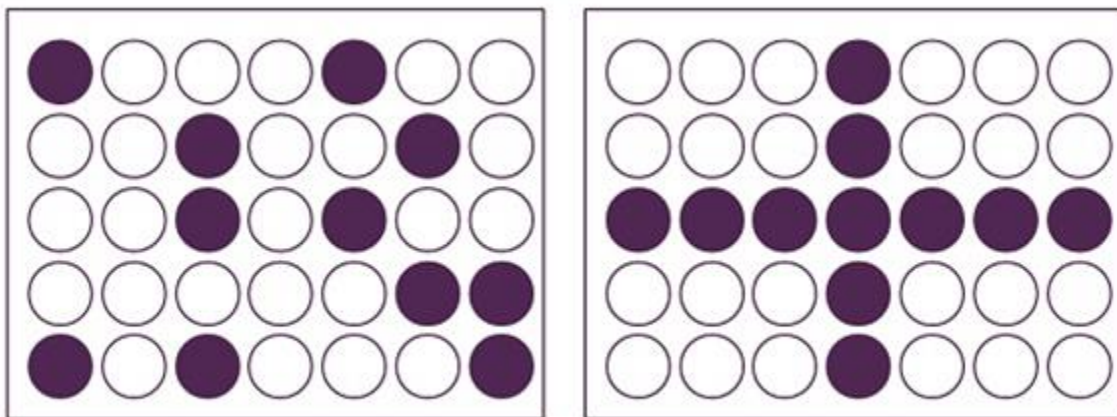
1-Probability Sampling

This Sampling technique uses randomization to make sure that every element of the population gets an equal chance to be part of the selected sample. It's alternatively known as random sampling.

2-Non-Probability Sampling

This Sampling is not random, it means that every elements are not equally-likely to be selected in sample.

The difference lies between the above two is whether the sample selection is based on randomization or not. With randomization, every element gets equal chance to be picked up and to be part of sample for study.



Probability Sampling Vs Non-Probability Sampling

TYPES OF PROBABILITY SAMPLING

1-Simple Random sampling without replacement

Let suppose population size is N and sample size is S where $S \leq N$. This sampling a subset of size S is selected randomly from the population. Each element have equal chance of getting selected. In other way one by one element can be selected but the element which is selected is removed from population.

2-Simple Random sampling with replacement

In this sampling a subset of size S is selected randomly one by one from the population. The element which is selected is remains in the population. It has equal chance of getting selected next time.

Application of Data Sampling

1-Data Reduction

Disadvantages of Data Sampling

1-Information Loss.

2-Class Imbalance problem