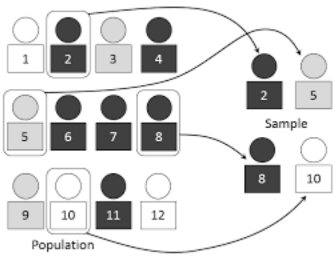
Definition 1: A ***population*** consists of the totality of the observations with which we are concerned. A ***sample*** is a subset of a population which is collected by some process for analysis.

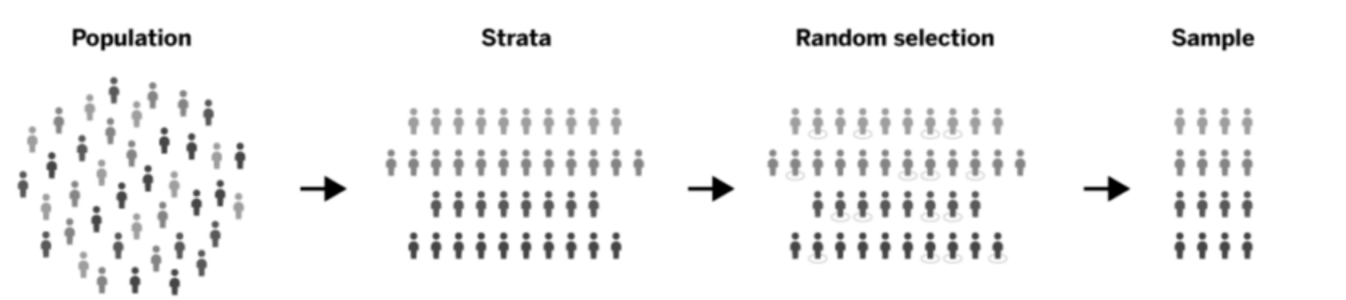
Definition 2: The ***population probability distribution*** is the probability distribution of the population data. A ***sampling distribution*** is a probability distribution of a statistic that is obtained through repeated sampling of a specific population. It describes a range of possible outcomes for a statistic, such as the mean or mode of some variable, of a population.

1. ***Simple Random Sample*** Let a finite population contains N units (called sampling units) all of which are distinguished from one another. The number of distinct samples of size n that can be drawn from the N units is given by ${}^N\mathrm{C}_n = \frac{N!}{n!(N-n)!}$. Then the Simple Random Sampling is a method of selecting n units out of N units such that every one of the ${}^N C_n$ samples has an equal chance of being chosen.

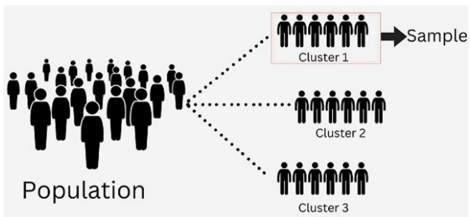
   In practice a Simple Random Sample is drawn units by unit. At any stage in the draw, the process gives an equal chance of selection to all units not previously drawn. The unit drawn from the population is not replaced since this might allow the same units to enter the sample more than once. This is described as the sampling ***without replacement***. Sampling is said to be ***with replacement*** if the sampling unit drawn from the population is returned to the population before the next unit is drawn.

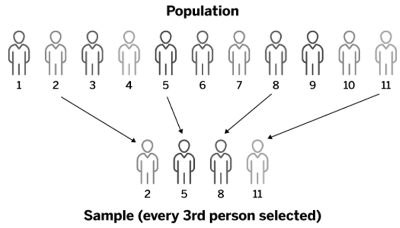
2. ***Stratified Random Sample*** In stratified sampling, the population is divided into relatively homogenous subpopulations. These subpopulations are non-overlapping and together they comprise the whole of the population. These subpopulations are called **strata** and the process is called **stratification**. Simple random samples are then drawn independently from each stratum and combined into a single sample. The whole procedure is described as ***Stratified Random Sampling***.

3. ***Cluster Sampling*** In cluster sampling the population is divided into many subgroups, each with only few elements, having heterogeneity within subgroups and homogeneity between subgroups. These subgroups are called **clusters** and the process is called **clustering**. Each cluster is taken as a sampling unit of the population. A random sample of these clusters is then drawn. This procedure is described as **cluster sampling**.

4. ***Systematic Sampling*** In systematic sampling all the N units of the population are serially numbered from 1 to N and every k^{th} unit is selected to get a sample. The first unit of the sample, say ith unit, is selected at random between the numbers 1 and k and every k^{th} unit thereafter. Such a process is called ***systematic sampling*** with random start and k is called the sampling interval.



Sampling Distribution: The value of the sample mean for any one sample will depend on the elements included in that sample. Consequently, the sample mean, $\mu_{\overline{x}}$, is a random variable.
The probability distribution of $\overline{x}$ is called its sampling distribution. It lists the various values that $\overline{x}$ can assume and the probability of each value of $\overline{x}$.

Mean and satandard error: The mean of the sampling distribution of a statistic is called mean of that statistic. The standard deviation of of the sampling distribution is of a statistic is called standard error of that statistic. The mean and standard error of any statistic are of great importance in the study of inferential statistics.

Sampling distribution of sample means: The sampling distribution of mean has the following properties:

1. $\mu_{\overline{x}} = \mu$
2. If the sampling is done with replacement from a polulation the the standard error is $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$
3. If the sampling is done without replacement from a finite polulation the the standard error is $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$. The factor $\sqrt{\frac{N-n}{N-1}}$ is called Finite Population Correction Factor.

Central Limit Theorem : *For a large sample size, the sampling distribution of $\overline{x}$ approach a normal distribution, irrespective of the shape of the population distribution. The mean and standard deviation of the sampling distribution of $\overline{x}$ are, respectively,*

$$\mu_{\overline{x}} = \mu, \sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

Remark: If the sample size is sufficiently large, the central limit theorem can be used to answer questions about sample means in the same manner that a normal distribution can be used to answer questions about individual values. The only difference is that a new formula must be used for the z values. It is

$$z = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}}$$

Example: *Hours That Children Watch Television:* A. C. Neilsen reported that children between the ages of 2 and 5 watch an average of 25 hours of television per week. Assume the variable is normally distributed and the standard deviation is 3 hours. If 20 children between the ages of 2 and 5 are randomly selected, find the probability that the mean of the number of hours they watch television will be greater than 26.3 hours.