

Unit III

RESEARCH SAMPLING

Compiled by
Dr. V.SETHURAMALINGAM
Professor and Head
Department of Social Work
Bharathidasan University
Tiruchirappalli-620 023

Note: This material is used for only study purpose and nor for any other purpose

SAMPLING

Need for sampling

Usually, researcher would collect information from every individual of the population that researcher decided to study

In order to **save time** and to **reduce cost of expenditure** towards the proposed research sampling is necessary.

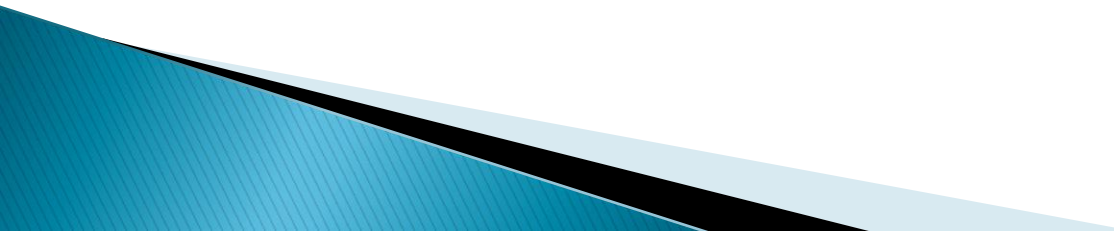
Hence, researchers have to work with a sample of subjects rather than the full population.

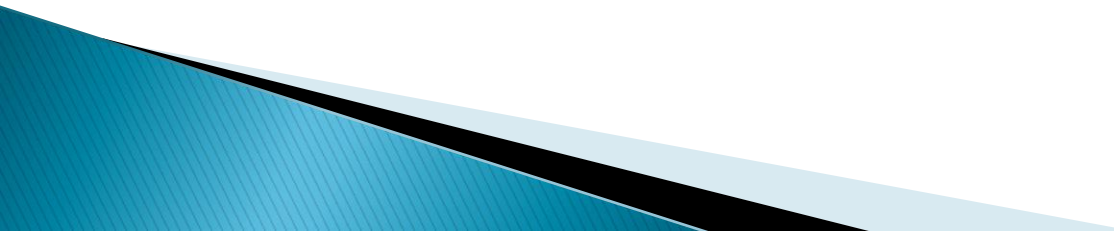
- ▶ To generalize from the sample to the population, the sample has to be representative of the population.
- ▶ The safest way to ensure the accuracy in collecting data from the sample representative is to use a random selection procedure

Sample

- ▶ Sample is a representative part of the population.
- ▶ Sampling is a smaller representation of a large whole (population) (Goode & Hatt).
- ▶ Any subset (sub group) of sampling units from a population (Nacmias & Nacmias)

Sampling

- ▶ Sampling is the process of selecting a group of people, events, behaviors, or other elements with which to conduct a study (umsl.edu)
 - ▶ Sample is a part of the research that indicates how cases are to be selected for research.
- 

- ▶ Sample is the selected elements (people or objects) chosen for participation in a study; people are referred to as subjects or participants (**umsl.edu**) or respondents
 - ▶ A sample is a miniature picture or cross section of the entire group or aggregate from which the sample is taken. (P.V. Young,)
- 

Population /universe

Before knowing the concept of sampling one must know the meaning of ‘**population**’

The entire group from which a sample is chosen is known as the 'population' or 'Universe' (P.V. Young,)

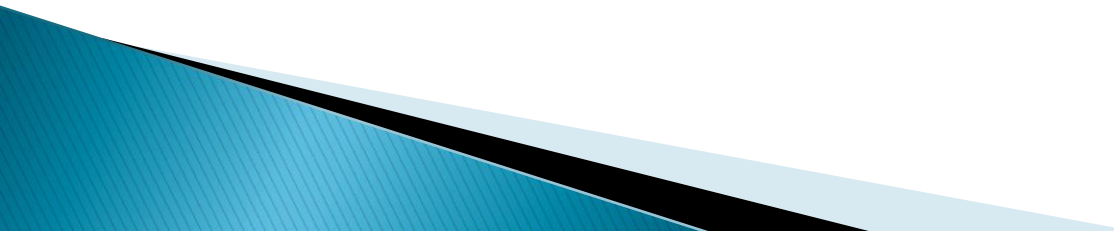
A population can be defined as ‘including all people or items with the characteristic one wish to understand.’

A research population is also known as a well-defined collection of individuals or objects known to have similar characteristics.

All individuals or objects within a certain population usually have a common, binding characteristic or trait.

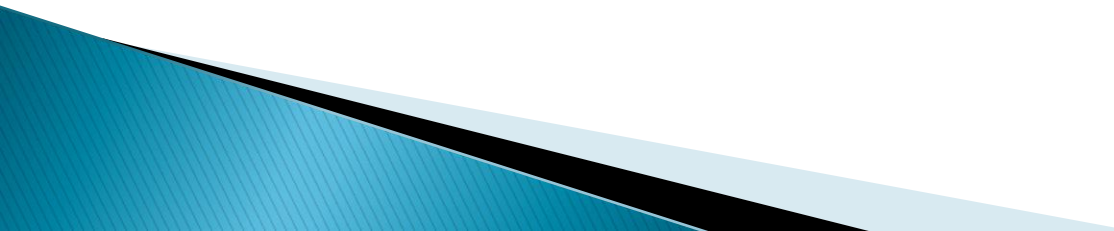
A large collection of individuals or objects that is the main focus of a scientific query

A complete set of elements (persons or objects) that possess some common characteristic defined by the sampling criteria established by the researcher
(Explore, Com)

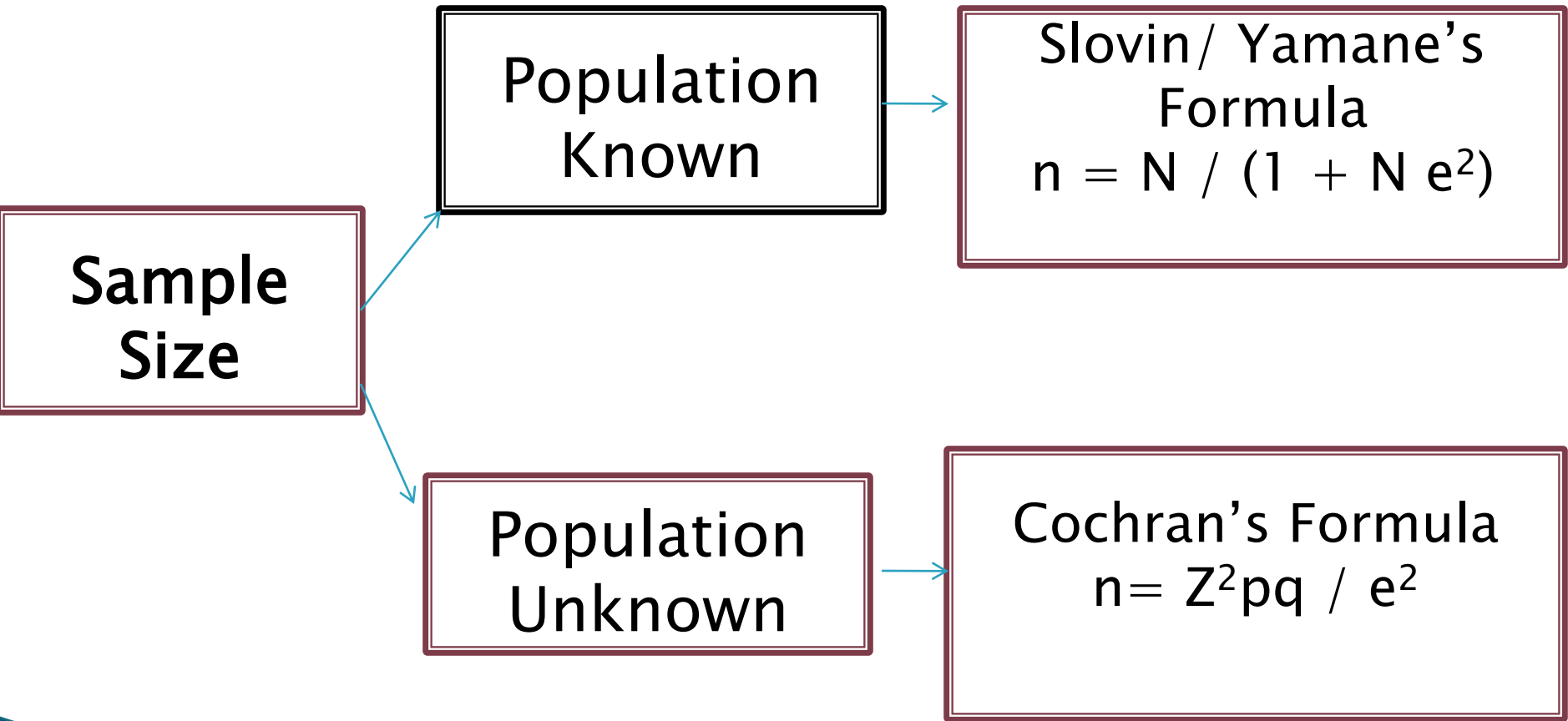


Principles behind choosing a sample

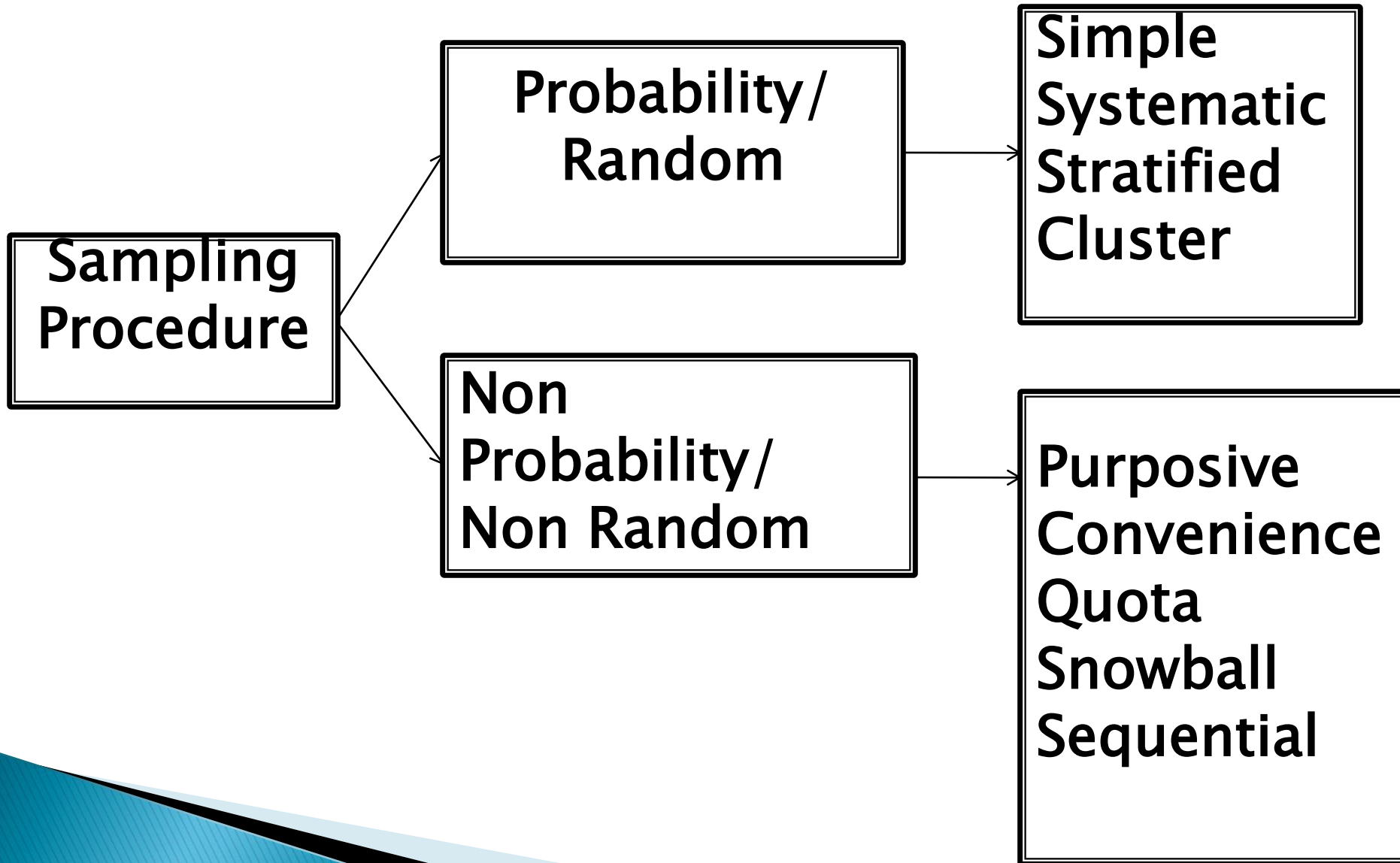
The idea behind selecting a sample is to be able to **generalise your findings** to the whole population, which means that your sample must be:

1. **Representative of the population.**
 2. **Large enough to give you adequate information to avoid errors.**
- 

How big should the sample be?



How is a sample taken correctly?



Types of sampling

There are two types of sampling procedure.

- 1. Probability / Random Sampling and**
 - 2. Non Probability / Non Random Sampling**
- 

SAMPLING

Random (Probability) sampling

Non Random (Non Probability) sampling

Simple random sampling

Systematic random sampling

Purposive/ judgement
Convenience/ accidental
Quota sampling
Snowball sampling
Sequential Sampling

Lottery method

Random number table

Linear system sampling

Circular system sampling

Cluster sampling

Stratified random sampling

Single stage cluster

Multi stage cluster

Proportionate

Disproportionate

Simple random (Lot/Number table) or Systematic random (Linear/Circular)

PROBABILITY SAMPLING / Random Sampling:

Probability or Random means the **chance of being included in the sample** is known as Probability Sampling.

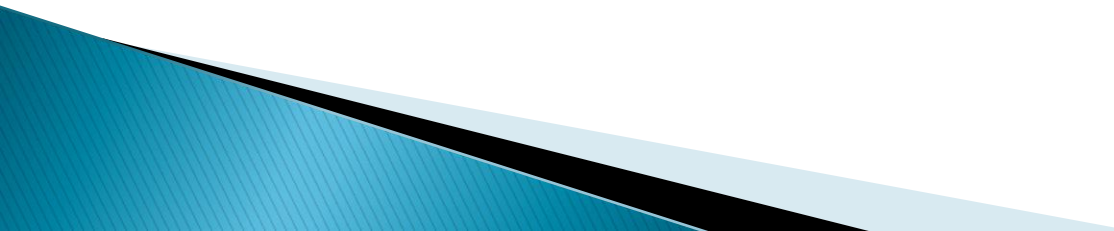
It means each individual in the population should be given equal chance of being included

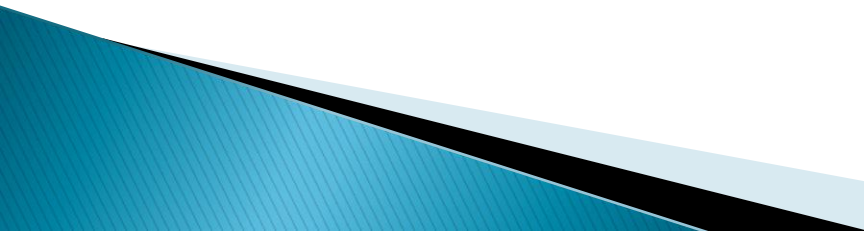
It means that every single individual in the sampling frame has a known and non-zero chance of being selected into the survey sample (The World Bank, 2018).

It means every individual from a population has an equal chance of being selected.



There are four types of Probability or Random

1. Simple Random Sampling
 2. Systematic Random Sampling
 3. Cluster Sampling
 4. Stratified Random Sampling
- 

- ▶ **1. Simple Random Sampling** (also known as un restricted random sampling)
 - ▶ It means Sample is taken from **a part** from **the total population** by giving EQUAL OPPORTUNITY for all the individuals
 - ▶ Simple random sampling can be carried out in two ways – the lottery method and using random number tables
 - ▶ **a. Lottery Method:**
 - ▶ All items of the universe are numbered or named on separate slips of paper of identical size and shape. These slips are folded and mixed up in a container and make a blind fold selection of sample
- 

b. Random Number Tables:

Random numbers are strings of digits that have been generated by the lottery method and can be found in books of statistical tables.

Types of Random Number tables:

- i. Tippet Number: Consists of 41,600 grouped in to 10,400 sets of 4 digit random. Number
- ii. Fisher and Yales: consists of 15,000 grouped in to 1500 of 10 digit random number
- iii. Kendal & Babington Smith: Consists of 1,00,000 numbers grouped in to 5000 sets of 4 digits; and
- Iv. Rand Corporation : This table of random numbers consisting of 1,00,000 random digits grouped into 20,000 sets of five-digit random numbers;

2. Systematic Random Sampling (Quasi random sampling or interval random sampling) :

Systematic sampling is a type of probability sampling method in which sample members from a larger population are selected according to a random starting point and a fixed, periodic interval (investopedia.com , 2015)

Systematic sampling is a statistical method involving the selection of elements from an ordered sampling frame (SarkiAce,)

Systematic Sampling is the ordered sampling at fixed intervals from a list, starting from a randomly chosen point.



Types in Systematic Random Sampling

There are two types in this type of sampling

- a. Linear System Sampling
- b. Circular System Sampling



Linear System Sampling

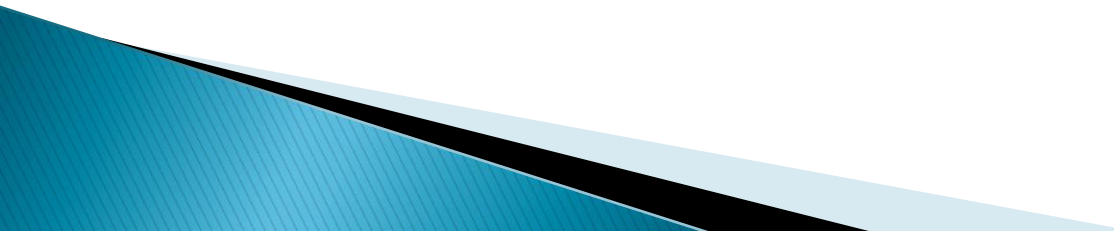
Linear system sampling is one that is not repeated after the end of the population. It goes linearly along the population and then stops (Mathcaptain.Com, 2015)

Example

If you want to collect data from 50 respondents among the total of 100 people, you can arrange the list of people and starts from 1, then 3,5, 7 and so on and end with 99

Or

You can start with 2,4,6,8,10 ,12 and so on and end with 100



Circular systematic sampling:

Circular system sampling starts from the first sample again after its end. Therefore, it continues in a circular shape (Mathcaptain.Com, 2015).

In Circular systematic sampling, the end of list is connected to the beginning of the list, making the list circular. This allows the random start from any number and end with the beginning number (Saifuddin, ND)


For example, There is list of 100 people . You have to select 50 samples from the list. To select the sample you can select the number 90 randomly there after you can select 92, 94, 96, 98, 100, 2, 4, 6, 8, 10 and so on and finally end with 88.

3.Cluster Sampling

A cluster is a naturally occurring unit or grouping within the population (O'Donnell, et al, 2010)

Cluster means, divide the population into separate groups.

Cluster sampling is useful because it avoids having to compile exhaustive lists of every single person in the population. Clusters should be as heterogeneous as possible within and as homogenous as possible between (note that this is the opposite criterion as that for strata). Clusters should be as small as possible



In cluster sampling, the population is broken into groups of cases, called Cluster sampling and the sample of clusters is selected at Random. Clusters, generally, consists of natural groupings, such as geographical , states, blocks, villages, etc. (Stattrek.Com, 2018)

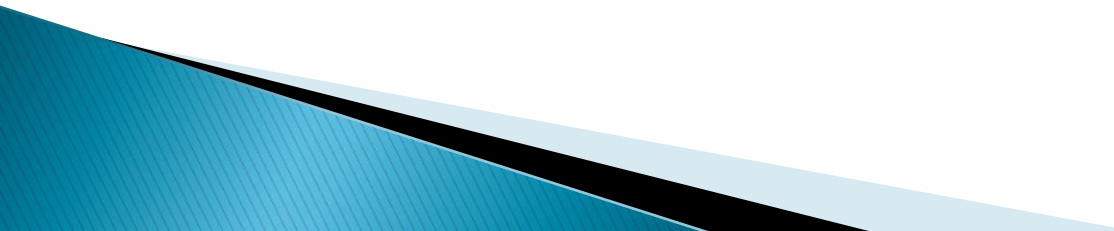
Note:

In stratified sampling, samples are drawn from each stratum.

In Cluster sampling, samples are drawn only from those clusters selected for the sample

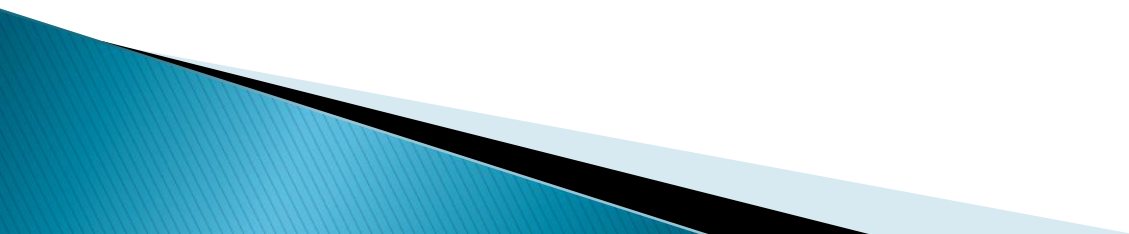
Types of Cluster sampling.

There are two types of Cluster Sampling

- a. Single –stage cluster sampling
 - b. Multi stage cluster sampling
- 

Single –stage cluster sampling:

If all the cases in each sampled cluster are included in the sample design - is called Single stage cluster sampling- in that case sampling occurs once - at the cluster level



Multi stage cluster sampling:

Cluster sampling involves sampling at two or more steps or stage.

The sampling units in the First Stage of multi stage sampling are called as Primary Sampling Units (PSU)

The sampling units in the Second Stage sample are termed as Secondary Sampling Units and so on.

Finally the researcher has to use simple or stratified random sampling to select the cases.

The principle reason for cluster sampling is to reduce the cost of data collection.

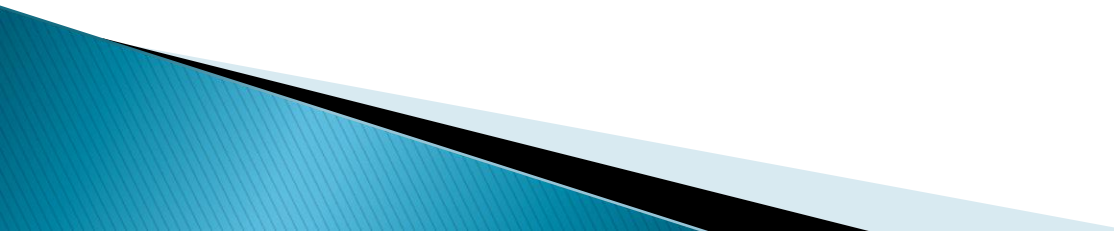
Used in larger size population group.



4. Stratified Random Sampling (also known as Restricted random sampling)

In this type of sampling, the sampling frame is divided into sub-sections comprising groups that are relatively homogeneous with respect to one or more characteristics and a random sample from each stratum is selected.

The Population is first sub divided in to *Stratum*. i.e. Divide the Heterogeneous population into Homogeneous unit called *stratum*.



Stratified random sampling divides the population into smaller groups, or strata, based on shared characteristics.

That is, the strata are formed based on members' shared attribute or characteristics.

Then apply the simple random techniques to obtain the representation representative sampling.

Types of Stratified Random Sampling

- a. Proportionate Stratified Random Sampling
 - b. Disproportionate Stratified Random Sampling
- 

a. Proportionate Stratified Random Sampling:

The Number of items drawn from each strata is proportional to the size of the strata i.e.

Sample size must be taken proportionately from each stratum.

For example there are 5 villages with 2000 households. The researcher has to select 5 per cent of sample from the total households.

Households	Sample size
800	40
600	30
400	20
200	10

2000

100

***b.* Disproportionate Stratified Random Sampling:**

An equal Number of sample / cases is taken from each stratum regardless of how the stratum is represented in the universe.

For example there are 5 villages with 1900 households. The researcher has to select 100 households (25 from each village) as sample from the five villages

Population	Sample size
500	25
1000	25
300	25
100	25

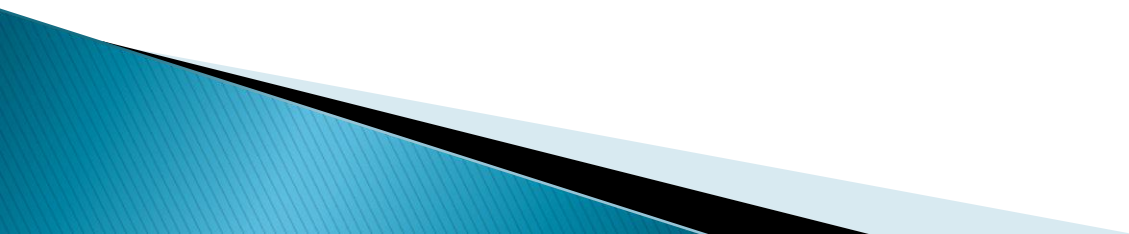
Total

1900

100

NON PROBABILITY /Non Random Sampling:

The chances of selecting any case are not known because cases are non-randomly selected.



Types of Non Probability Sampling

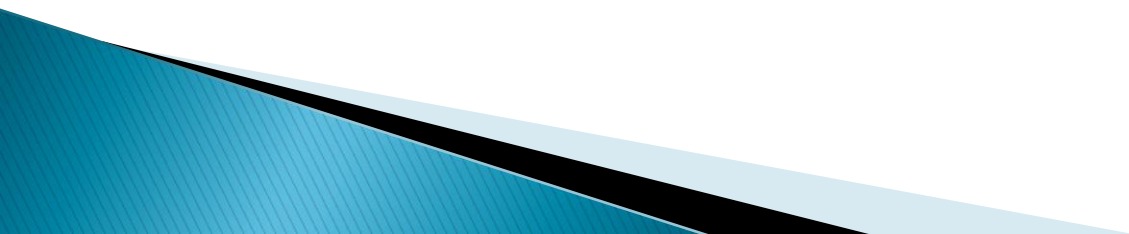
Purposive/ judgement

Convenience/ accidental

Quota sampling

Snowball sampling

Sequential Sampling



1. Purposive Sampling (Also known as *Judgment Sampling*)

Purposive sampling is where the researcher only approaches people who meet certain criteria, and then checks whether they meet other criteria.

Example **market research**- , if they are looking to **examine the shopping habits of men aged between 20 and 40**, they would only approach men, and then ask their age. (SkillsYouNeed.com (2018))

The researcher relies on his expert judgment to select the units that are “representatives” of the population

Sampling units are selected subjectively by the researcher (eg. Election forecast)

2. Convenience Sampling: Also called *Accidental Sampling/Consecutive sampling*

The Researcher simply selects a requisite number from cases that are **conveniently available**.

A convenience sample is chosen because of the group members' **accessibility and willingness** to participate in the group.

Convenience sampling selects a sample on the basis of **how easy it is to access**. Such samples are extremely easy to organise, but there is no way to guarantee whether they are representative.

Consecutive Sampling

Consecutive sampling is very **similar to convenience sampling** except that it seeks to include **ALL accessible subjects as part of the sample**. This non-probability sampling technique can be considered as the best of all non-probability samples because it includes all subjects that are available that makes the sample a better representation of the entire population. consecutive sampling, also known as total enumerative sampling (Sharma, 2014)

Note:

Convenience sampling is a quick, inexpensive method and is **generally used in pilot studies when researchers need to gather general data on trends or phenomena.**

However, when using statistics to analyze trends and cycles, Researchers **cannot rely on** convenience sampling **to make statistical inferences** about larger populations. Such a technique does not generate samples that adequately represent the population from which the samples are chosen, creating a high degree of sampling bias.

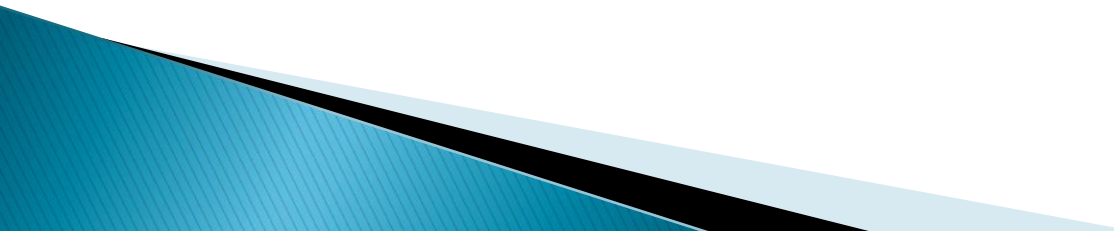
Eg. 1. Study of campus alcohol consumption

2. News papers and Televisions - Conducting Public Opinion on particular issues

3. Quota Sampling:

It is a form of purposive sampling like proportionate stratified random sampling. Quota sampling begins by dividing the population into relevant strata such as age, sex, race, geographical area, etc.

Quota sampling divides the **population into categories**, and then **selects from within categories** until a sample of the chosen size is obtained within that category (SkillsYouNeed.com , 2018)



The quota of cases with in a stratum may be filled in what ever way the researcher choose

(Difference between Stratified Random Sampling-
Requisite number of cases with in **each stratum must be drawn by simple Random sampling**)

Used on the basis of :

- Characteristics of Population
- Size of population
- The quota of the sample unit is divided and assigned for each group


4. **Snowball Sampling:**

Used when the universe is unknown.

Used in sensitive studies like - Prostitution, Drug abusers, etc.

Snowball sampling is where the researcher **starts with one person who meets their criteria, and then uses that person to identify others** (SkillsYouNeed.com, 2018).

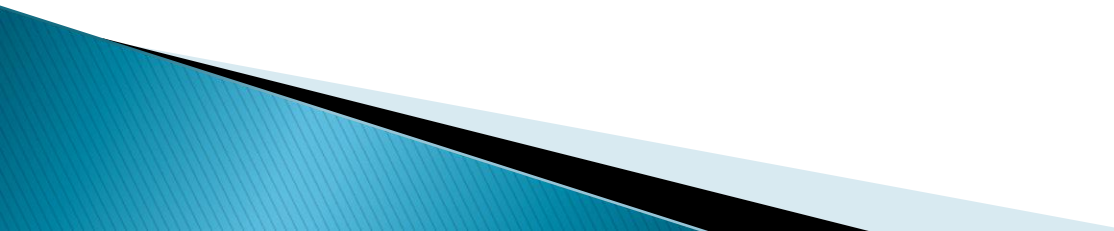
This works well when your sample has very specific criteria: for example, if you want to talk to workers with a particular set of responsibilities, you might approach one person with that set, and ask them to introduce you to others.



Warning: Non Probability Sampling

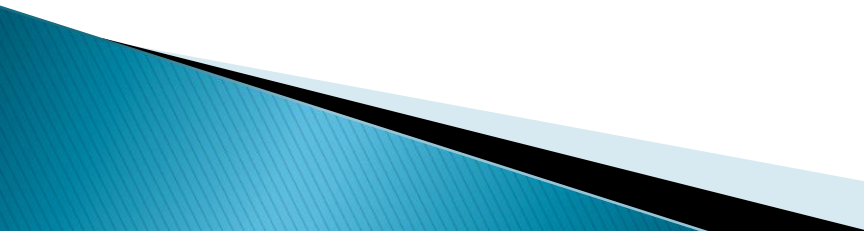
Non-probability sampling methods have generally been developed to address very specific problems. For example, snowball sampling deals with hard-to-find populations, and convenience sampling allows for speed and ease.

However, although some non-probability sampling methods, particularly quota and purposive sampling, ensure the sample draws from all categories in the population, samples taken using these methods may not be representative. (SkillsYouNeed.com., 2018)

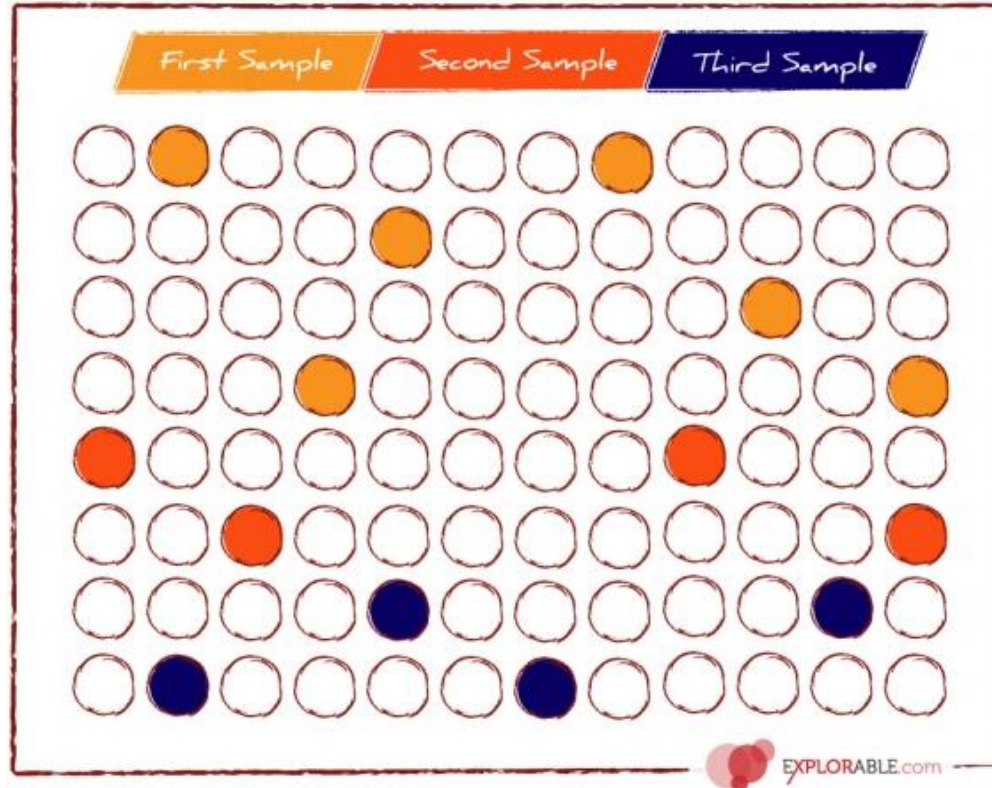


5. Sequential sampling

Sequential sampling is a non-probabilistic sampling technique, initially developed as a tool for **product quality control**. The sample size, n , is not fixed in advanced, nor is the timeframe of data collection. The process begins, first, with the sampling of a single observation or a group of observations. **These are then tested to see whether or not the null hypothesis can be rejected.** If the null is not rejected, then another observation or group of observations is sampled and the test is run again. In this way the test continues until the researcher is confident in his or her results (Bakeman, & Gottman, 1997) .



Sequential sampling is a non-probability sampling technique wherein the researcher picks a single or a group of subjects in a given time interval, conducts his study, analyzes the results then picks another group of subjects if needed and so on (Explorable.com, 2010)

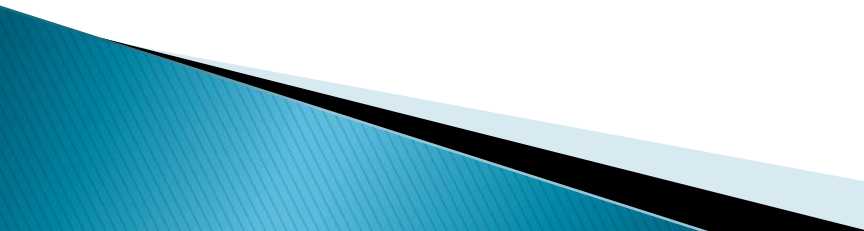


Source: Explorable.com, 2010

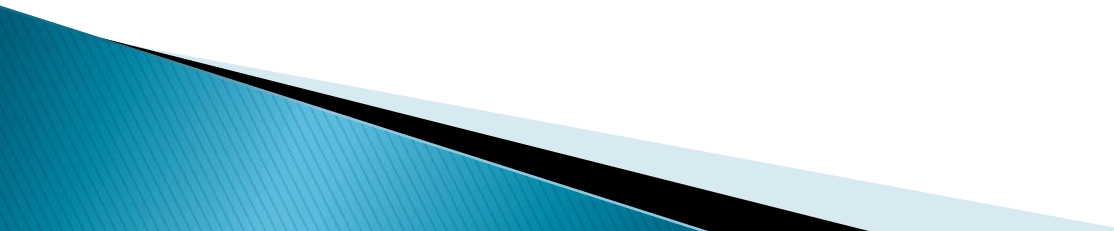
In sequential sampling, a sequence of one or more samples is taken from a group. Once the group has been sampled, a hypothesis test is performed to see if you can reach a conclusion.

A characteristic feature of sequential sampling is that the **sample size is not set in advance**, because you don't know at the outset how many times you'll be repeating the process (Statisticshowto.com , 2018)

The principle is “**can I make a decision if I stop right now?**”. Conventional sampling methods have a prefixed sample size. On the other hand, in sequential sampling, you as a researcher have to ask at the end of each measurement, the collected samples are good enough to make a decision without a sample size fixed in advance. (Thapliyal, 2018).



Sequential sampling is a sampling method in which samples are taken one at a time or in successive predetermined groups, until the cumulative result of their measurements (as assessed against predetermined limits) permits a decision to accept or reject the population or to continue sampling. The number of observations required is not determined in advance, but the decision to terminate the operation depends, at each stage, on the results of the previous observations. The plan may have a practical, automatic termination after a certain number of units have been examined (U.S. Census Bureau, 2013).



Outcomes in Sequential Sampling

With conventional sampling methods, the possible results of a hypothesis test has **either reject the null hypothesis, or do not.**

In sequential sampling technique, there exists another step, **a third option. :**

Reject the null hypothesis (end the experiment),

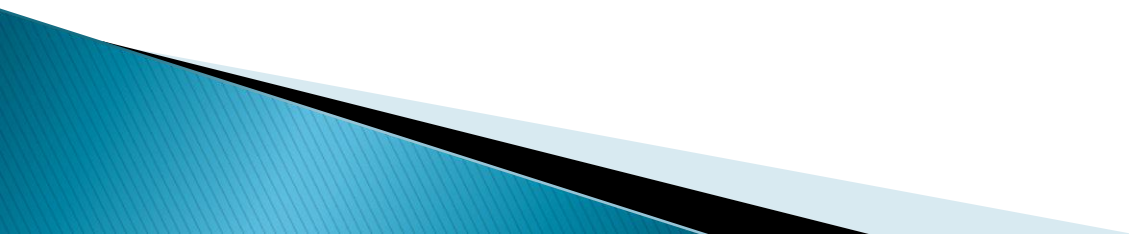
Do not reject the null hypothesis (end the experiment),

Fail to draw any conclusion (draw another sample and repeat the test)-
(statisticsshowto.com. 2018). This entails that the researcher can obtain limitless number of subjects before finally making a decision whether to accept his null or alternative hypothesis

Time-Sequential Sampling

In this variant, sometimes called time-sequential classification, you use time as your sampling frame instead of a physical population to sample from. For example, you might choose a sample member at 24-hour intervals (Statisticsshowto.com, 2018).

This sampling technique gives the researcher limitless chances of fine tuning his research methods and gaining a vital insight into the study that he is currently pursuing



References

Ahmed, S. (ND) Methods in Sample Surveys : Systematic Sampling , Johns Hopkins University Retrieved from <http://ocw.jhsph.edu/courses/StatMethodsForSampleSurveys/PDFs/Lecture2.pdf>

Bakeman, R & Gottman, J.M. (1997) Observing Interaction: An Introduction to Sequential Analysis. Cambridge: Cambridge University Press.

Bakeman, R & Gottman, J.M. (1997) Observing Interaction: An Introduction to Sequential Analysis Cambridge: Cambridge University Press.

Bakeman, R & Gottman, J.M. (1997). Observing Interaction: An Introduction to Sequential Analysis Cambridge: Cambridge University Press. Retrieved from <https://www.betterevaluation.org/en/evaluation-options/sequential>

BCPS.Org (2018) Develop a Research Proposal. Retrieved from https://www.bcps.org/offices/lis/researchcourse/develop_data_sampling_random.html

Ecourseonline (2012) Elementary Statistics. Retrieved from <http://ecourseonline.iasri.res.in/mod/page/view.php?id=34127>

Explorable.com (2010). Sequential Sampling Method. Retrieved from <https://explorable.com/sequential-sampling>

Explorable.com (2009). Quota Sampling. Retrieved from <https://explorable.com/quota-sampling>

Goode, W.J. and Hatt, P.J. (1952) Methods in Social Research. By William J. Goode and Paul K. Hatt. New York: McGraw-Hill Book Co.

Investopedia.Com (2015). Systematic Sampling <https://www.investopedia.com/terms/s/systematic-sampling.asp>

Josiah Kaplan (2014) Sequential Sampling. Oxford University, Oxford
Retrieved from <https://www.betterevaluation.org/en/evaluation-options/sequential>

Krebs, C. (2014). Ecological Methodology. 3rd. Ed. Pedigo, L. & Buntin, G. (1992). Handbook of Sampling Methods for Arthropods in Agriculture. CRC Press

Mathcaptain.Com (2015). Systematic Sampling. Retrieved from <http://www.Mathcaptain.com/statistics/systematic-sampling.html>

O'Donnell, O., Van Doorslaer, E., Wagstaff, A., & Lindelow, M. (2010). Analyzing health equity using household survey data: a guide to techniques and their implementation. 2008. Washington, DC: The World Bank Google Scholar.

Sharma, S. (2014). Nursing Research and Statistics. Elsevier Health Sciences. p. 224. I

SkillsYouNeed.com (2018). Sampling and Sample Design. Retrieved from <https://www.skillsyouneed.com/learn/sampling-sample-design.html>

Statisticshowto.com (2018). Sequential Sampling
<http://www.statisticshowto.com/sequential-sampling/>

Stattrek.Com (2018) Cluster Sampling. Retrieved from https://stattrek.com/statistics/dictionary.aspx?definition=cluster_sampling

Thapliyal, O. (2018). What is sequential sampling?. Retrieved from <https://www.quora.com/What-is-sequential-sampling>



The World Bank, (2018). Sampling Procedure, Retrieved from .
<http://microdata.worldbank.org/index.php/catalog/1739/sampling>

U.S. Census Bureau (2013). Sequential Sampling . Information Quality Glossary. Retrieved from https://definedterm.com/sequential_sampling

Young, P.V. (1949) Scientific Social Surveys and Research, New York.