

# OUTLINE OF PRESENTATION

---

❖ SAMPLE

❖ SAMPLING

❖ SAMPLING METHOD

❖ TYPES OF SAMPLING METHOD

❖ SAMPLING ERROR

---

# SAMPLE

---

- It is a Unit that selected from population

- Representers of the population

- Purpose to draw the inference

---

# WHY SAMPLE ?

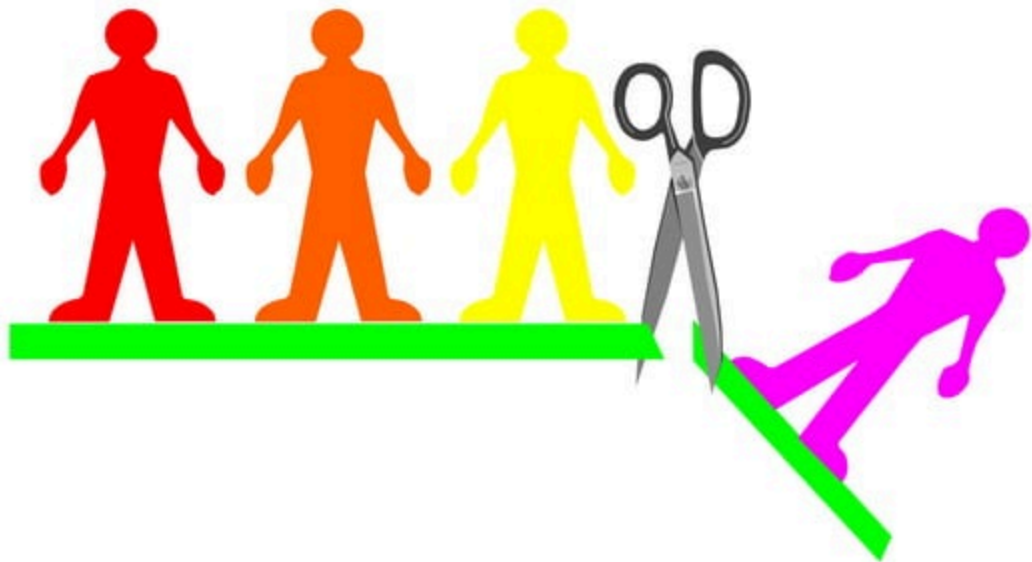
- ❖ Very difficult to study each and every unit of the population when population unit are heterogeneous

---

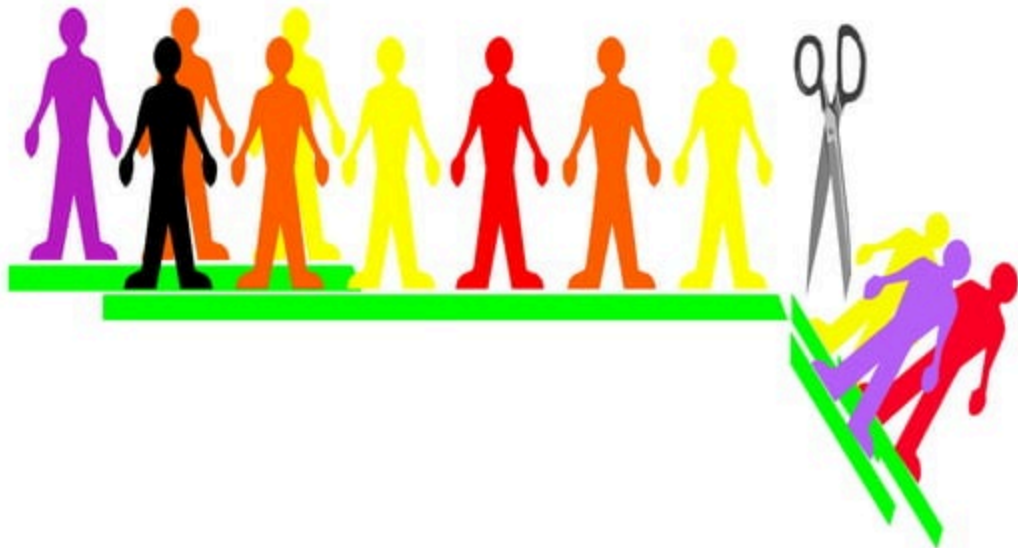
- ❖ Time Constraints

- ❖ Finance

It is very easy and convenient to draw the sample from homogenous population



The population having significant variations (Heterogeneous), observation of multiple individual needed to find all possible characteristics that may exist



## **Population**

The entire group of people of interest from whom the researcher needs to obtain information

---

## **Element (sampling unit)**

One unit from a population

## **Sampling**

The selection of a subset of the population through various sampling techniques

## **Sampling Frame**

Listing of population from which a sample is chosen. The sampling frame for any probability sample is a complete list of all the cases in the population from which your sample will be drawn

---

# Parameter

The variable of interest

---

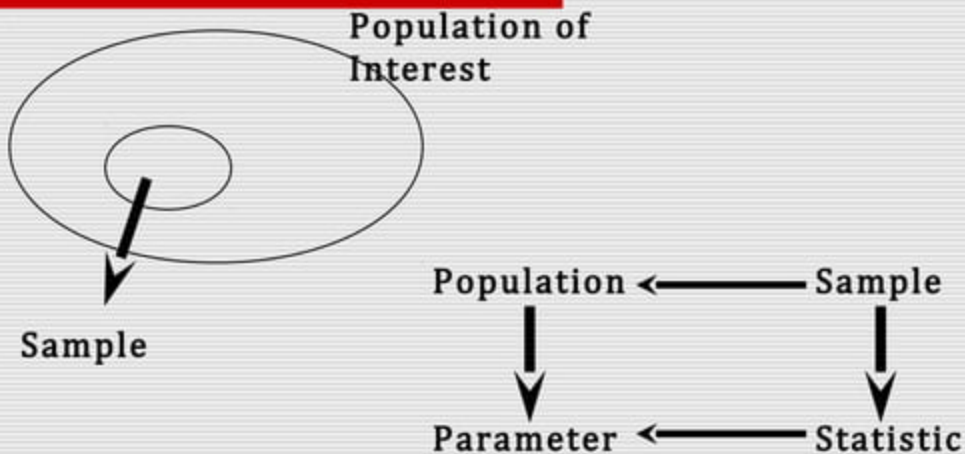
# Statistic

The information obtained from the sample about the parameter

---

# Population Vs. Sample

---



We measure the sample using statistics in order to draw inferences about the population and its parameters.

---



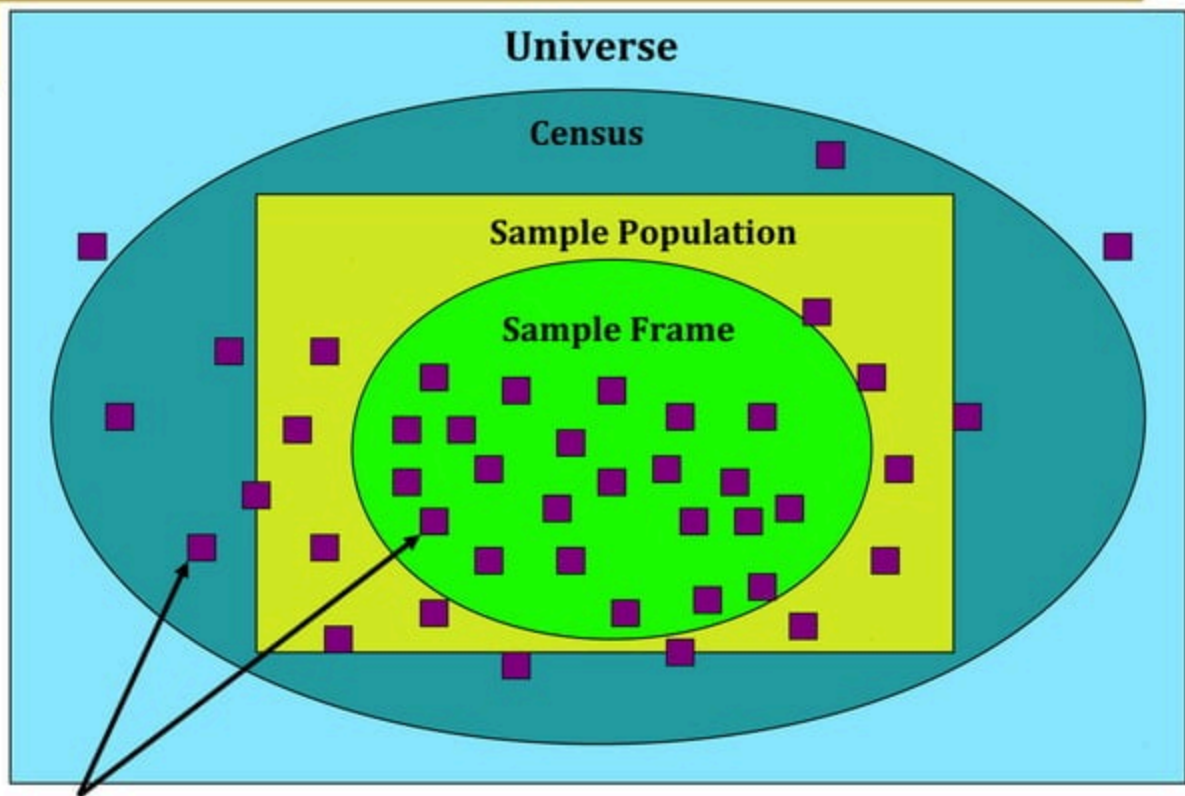
**Universe**

**Census**

**Sample Population**

**Sample Frame**

**Elements**



# Characteristics of Good Samples

---

❖ Representative

❖ Accessible

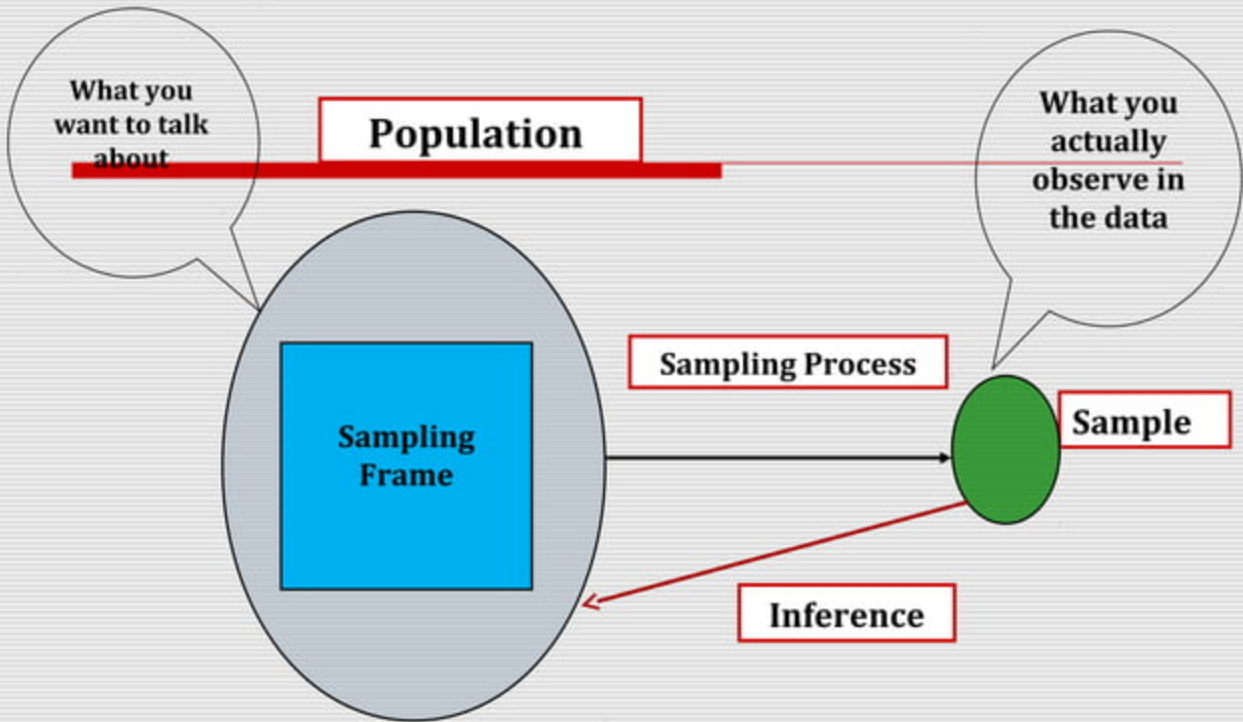
❖ Low cost

# SAMPLING

---

Process by which the sample are taken from population to obtain the information

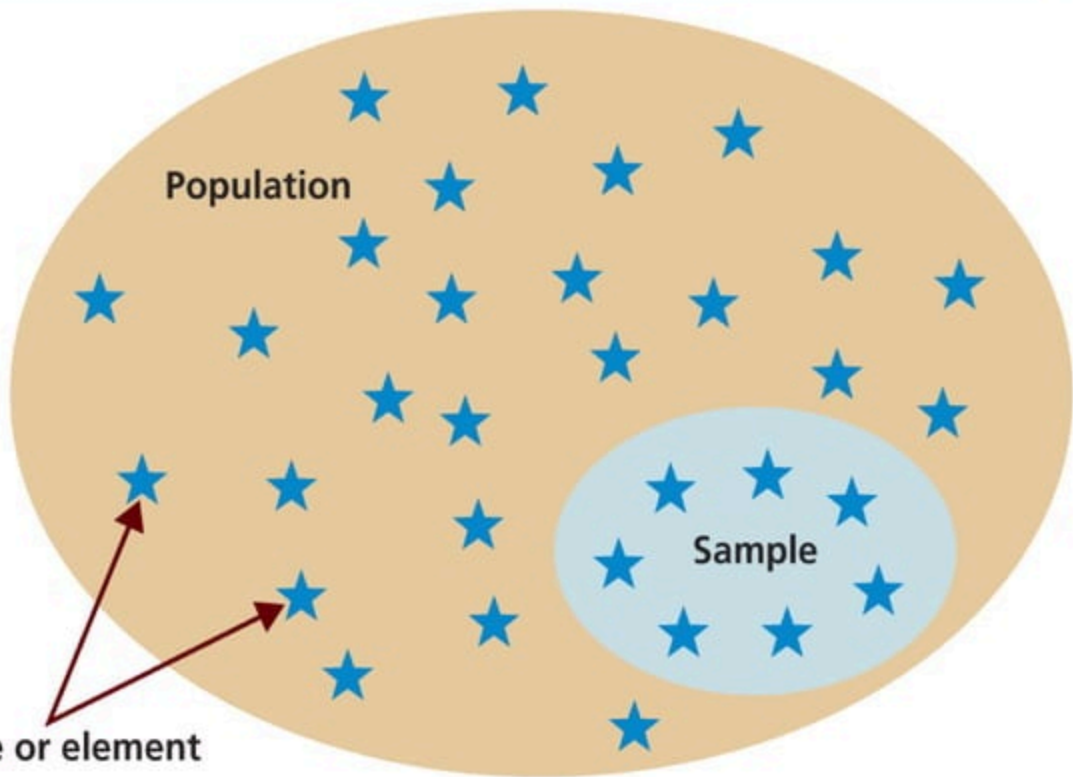
Sampling is the process of selecting observations (a sample) to provide an adequate description and inferences of the population



Population

Sample

Case or element



# Steps in Sampling Process

---

- ❖ Define the population
- ❖ Identify the sampling frame
- ❖ Select a sampling design or procedure
- ❖ Determine the sample size
- ❖ Draw the sample

# Sampling Design Process

Define Population

---

Determine Sampling Frame

Determine Sampling Procedure

**Probability Sampling**  
~~Simple Random Sampling~~  
Stratified Sampling  
Cluster Sampling  
Systematic Sampling  
Multistage Sampling

**Non-Probability Sampling**  
~~Convenient~~  
Judgmental  
Quota  
Snow ball Sampling

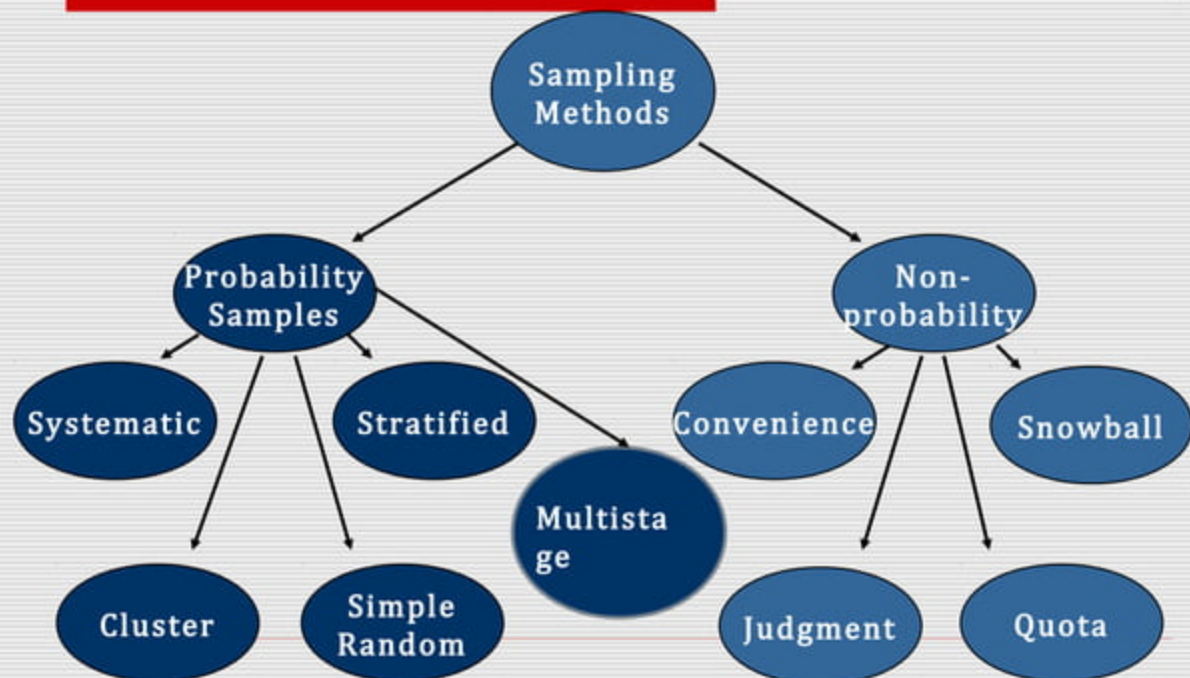
Determine Appropriate  
Sample Size

Execute Sampling  
Design

---

# Classification of Sampling Methods

---





# Probability Sampling

---

- ❖ Each and every unit of the population has the equal chance for selection as a sampling unit
- ❖ Also called formal sampling or random sampling
- ❖ Probability samples are more accurate
- ❖ Probability samples allow us to estimate the accuracy of the sample

## **Types of Probability Sampling**

---

- ❖ **Simple Random Sampling**
- ❖ **Stratified Sampling**
- ❖ **Cluster Sampling**
- ❖ **Systematic Sampling**
- ❖ **Multistage Sampling**

# Simple Random Sampling

---

- ❖ The purest form of probability sampling
- ❖ Assures each element in the population has an equal chance of being included in the sample
- ❖ Random number generators

# Simple random sampling

1	Albert D.
2	Richard D.
3	Belle H.
4	Raymond L.
5	Stéphane B.
6	Albert T.
7	Jean William V.
8	André D.
9	Jeremy W.
10	Anthony Q.
11	James B.
12	Denis G.
13	Amanda L.
14	Jennifer L.
15	Philippe K.
16	Eve F.
17	Priscilla O.
18	Robert D.
19	Brian F.
20	Hélène H.
21	Isabelle R.
22	Jean T.
23	Samanta D.
24	Berthe L.



25	Monique Q.
26	Régine D.
27	Lucille L.
28	Jérémy W.
29	Gilles D.
30	Renaud S.
31	Pierre K.
32	Etienne M.
33	Marie M.
34	Gaétan Z.
35	Fidèle D.
36	Maria P.
37	Anne Marie G.
38	Michel K.
39	Gaston C.
40	Alain M.
41	Olivier P.
42	Geneviève M.
43	Berthe D.
44	Jean Pierre P.
45	Jacques B.
46	François P.
47	Dominique M.
48	Antoine C.

# Types of Simple Random Sample

---

❖ With replacement

❖ Without replacement

## ❖ With replacement

---

The unit once selected has the chance for again selection

## ❖ Without replacement

The unit once selected can not be selected again

---

# Methods of SRS

---

- ❖ Tippet method
  - ❖ Lottery Method
  - ❖ Random Table
-

## Random numbers of table

6 8 4 2 5 7 9 5 4 1 2 5 6 3 2 1

4 0

5 8 2 0 3 2 1 5 4 7 8 5 9 6 2 0

2 4

3 6 2 3 3 3 2 5 4 7 8 9 1 2 0 3

2 5

9 8 5 2 6 3 0 1 7 4 2 4 5 0 3 6



# Advantages of SRS

---

- ❖ Minimal knowledge of population needed
- ❖ External validity high; internal validity high; statistical estimation of error
- ❖ Easy to analyze data

# Disadvantage

---

- High cost; low frequency of use
  - Requires sampling frame
  - Does not use researchers' expertise
  - Larger risk of random error than stratified
-

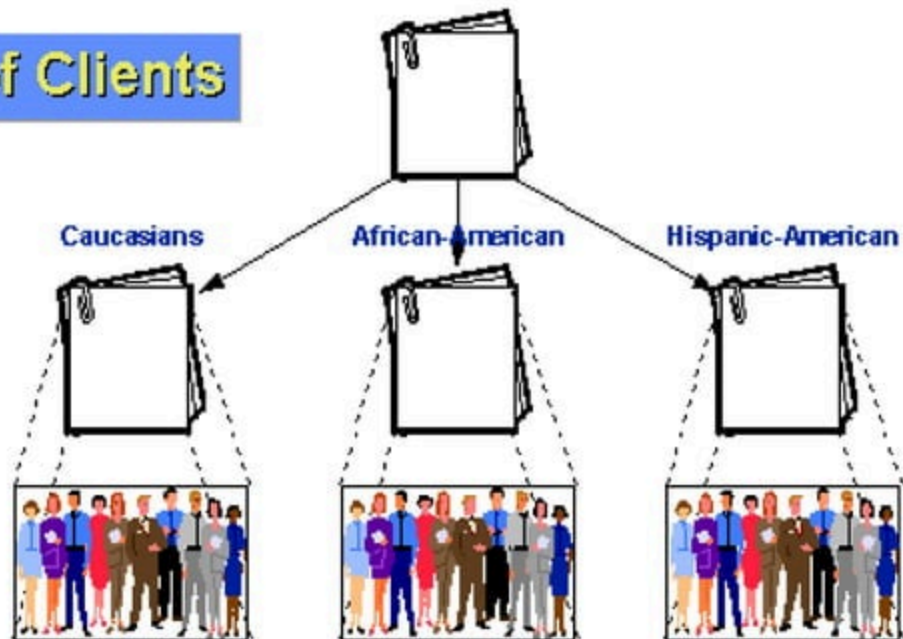
# Stratified Random Sampling

Population is divided into two or more groups called strata, according to some criterion, such as geographic location, grade level, age, or income, and subsamples are randomly selected from each strata.

Elements *within* each strata are homogeneous, but are heterogeneous *across* strata

# Stratified Random Sampling

List of Clients



Strata

Random Subsamples of  $n/N$

# Types of Stratified Random Sampling

---

## ❖ Proportionate Stratified Random Sampling

Equal proportion of sample unit are selected from each strata

## ❖ Disproportionate Stratified Random Sampling

Also called as equal allocation technique and sample unit decided according to analytical consideration

---

# Advantage

---

- ❖ Assures representation of all groups in sample population needed
- ❖ Characteristics of each stratum can be estimated and comparisons made
- ❖ Reduces variability from systematic

# Disadvantage

---

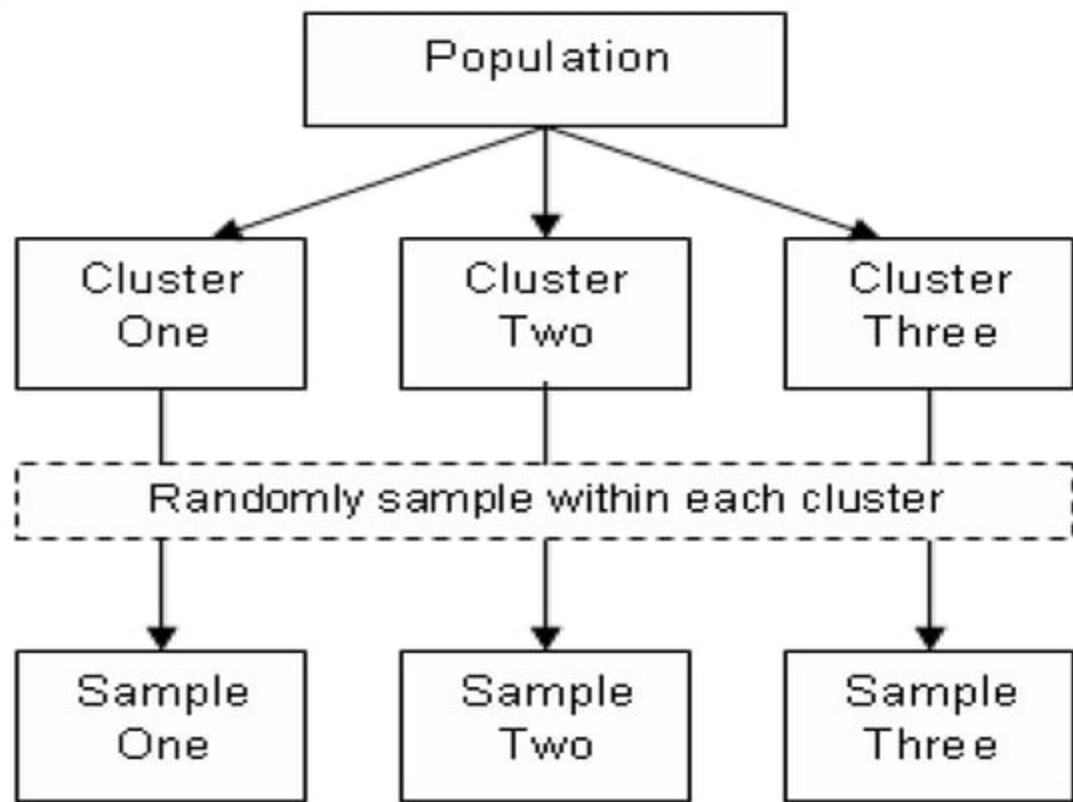
- ❖ Requires accurate information on proportions of each stratum
  
- ❖ Stratified lists costly to prepare

# Cluster Sampling

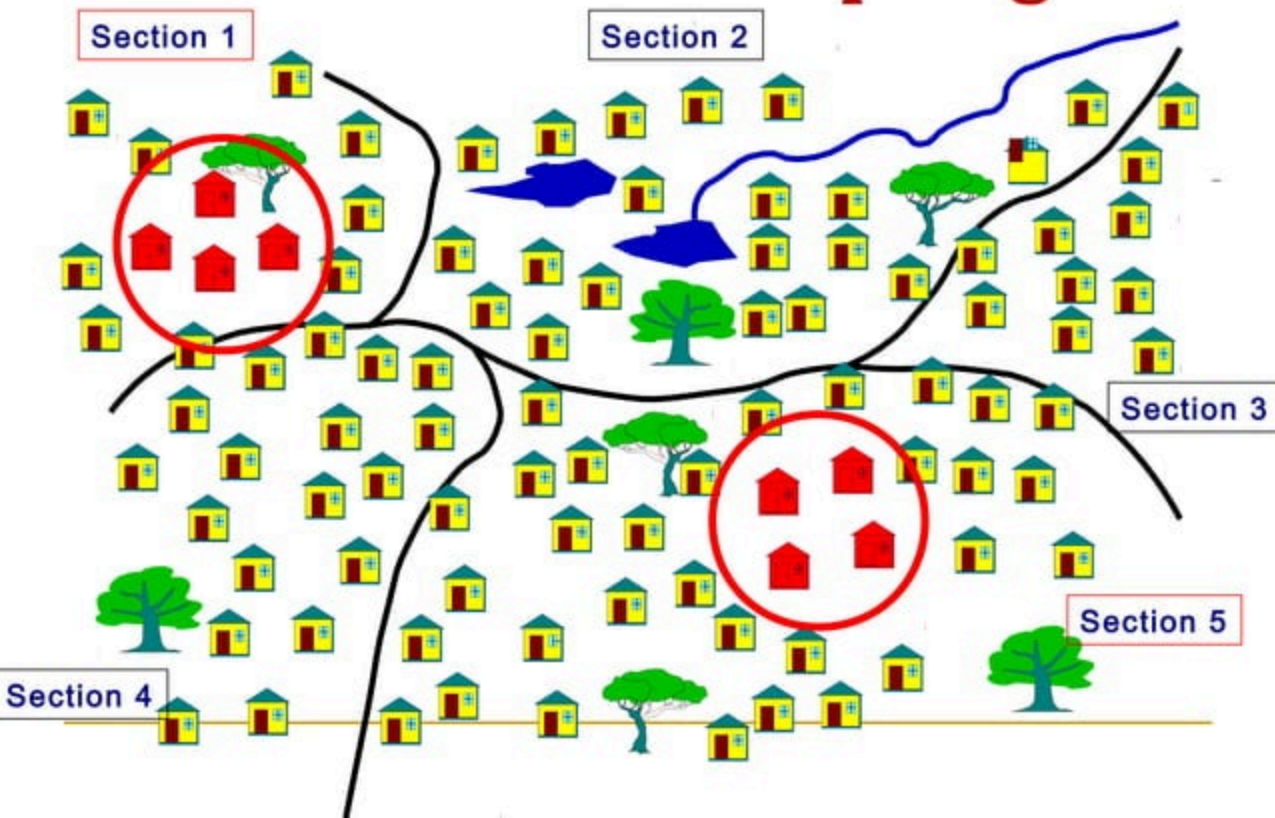
---

The population is divided into subgroups (clusters) like families. A simple random sample is taken of the subgroups and then all members of the cluster selected are surveyed.





# Cluster sampling



# Advantage

---

- Low cost/high frequency of use
  - Requires list of all clusters, but only of individuals within chosen clusters
  - Can estimate characteristics of both cluster and population
  - For multistage, has strengths of used methods
  - Researchers lack a good sampling frame for a dispersed population
-

# Disadvantage

---

The cost to reach an element to sample is very high

Usually less expensive than SRS but not as accurate

Each stage in cluster sampling introduces sampling error—the more stages there are, the more error there tends to be

## Systematic Random Sampling

---

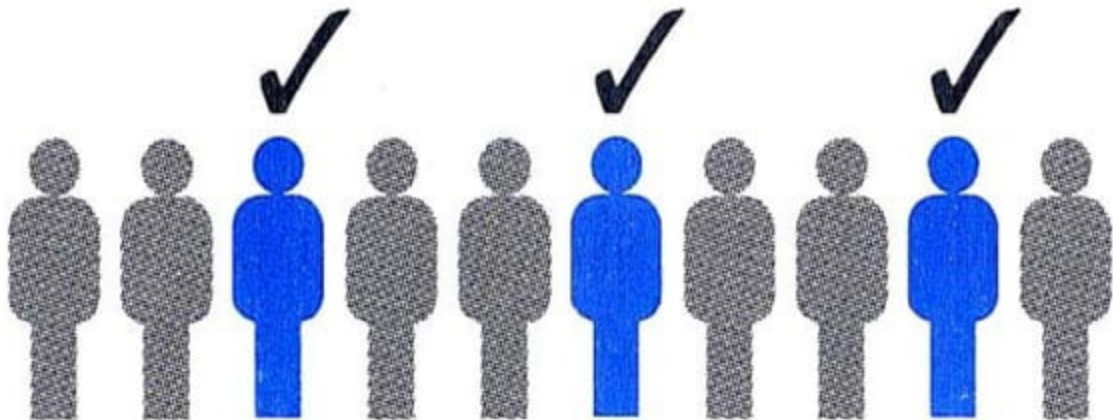
Order all units in the sampling frame based on some variable and then every  $n$ th number on the list is selected

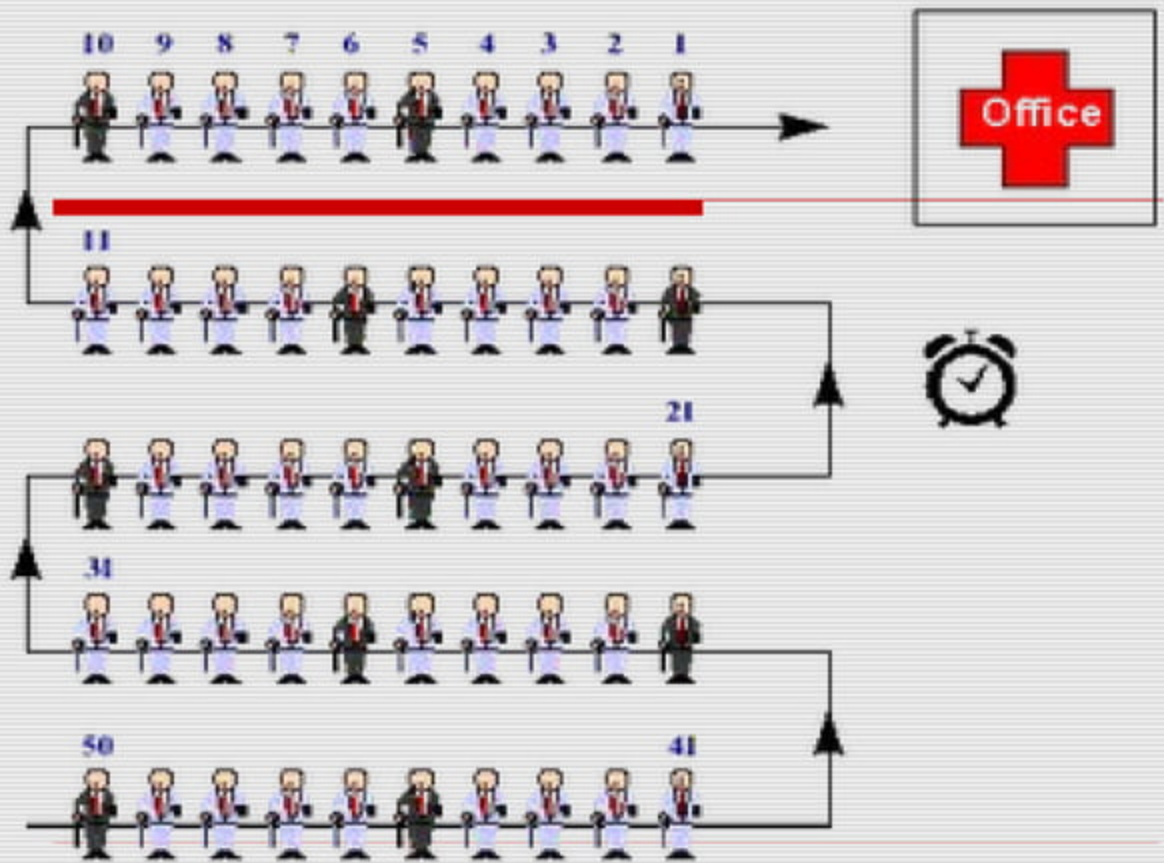
Gaps between elements are equal and Constant  $\rightarrow$  There is periodicity.

$N =$  Sampling Interval

# Systematic Random Sampling

---





# Advantage

---

- ❖ Moderate cost; moderate usage
- ❖ External validity high; internal validity high; statistical estimation of error
- ❖ Simple to draw sample; easy to verify



## Disadvantage

---

- **Periodic ordering**
- **Requires sampling frame**

## Multistage Random Sampling

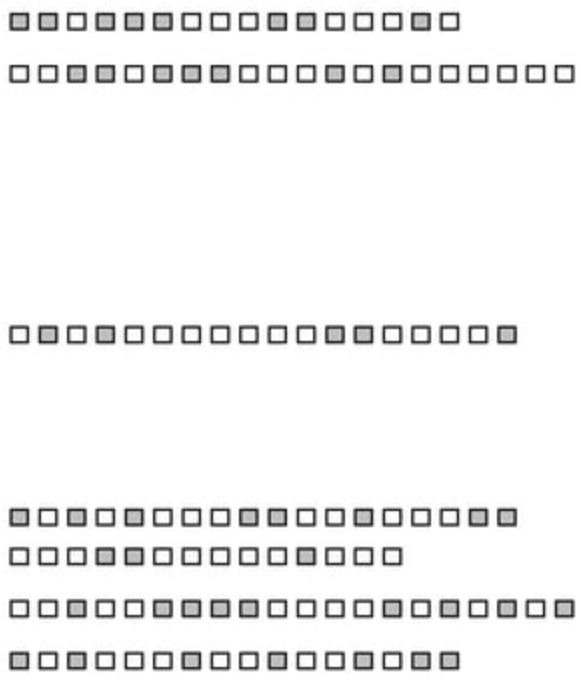
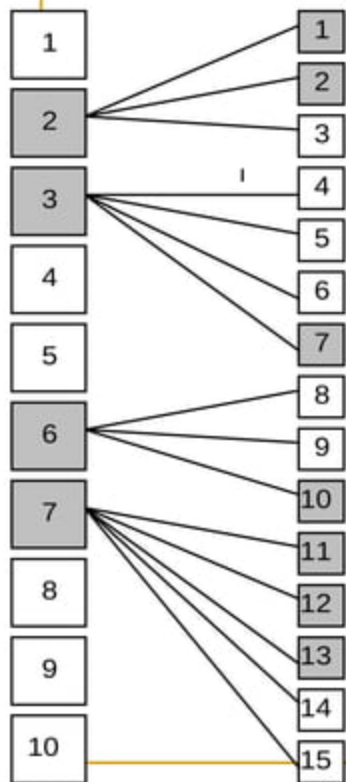
Multistage sampling refers to sampling plans where the sampling is carried out in stages using smaller and smaller sampling units at each stage.

Not all Secondary Units Sampled normally used to overcome problems associated with a geographically dispersed population

Primary  
Clusters

Secondary  
Clusters

Simple Random Sampling within Secondary Clusters



## Multistage Random Sampling

---

❖ Select all schools; then *sample* within schools

❖ Sample schools; then measure *all* students

❖ Sample schools; then *sample* students

# Non Probability Sampling

---

The probability of each case being selected from the total population is not known

Units of the sample are chosen on the basis of personal judgment or convenience

There are NO statistical techniques for measuring random sampling error in a non-probability sample. Therefore, generalizability is never statistically appropriate. 

---

# Non Probability Sampling

- ❖ Involves non random methods in selection of sample
- ❖ All have not equal chance of being selected
- ❖ Selection depend upon situation
- ❖ Considerably less expensive
- ❖ Convenient
- ❖ Sample chosen in many ways

## Types of Non probability Sampling

---

- ❖ Purposive Sampling
- ❖ Quota sampling (larger populations)
- ❖ Snowball sampling
- ❖ Self-selection sampling
- ❖ Convenience sampling

# Purposive Sampling

---

- ❖ Also called judgment Sampling
  - ❖ The sampling procedure in which an experienced research selects the sample based on some appropriate characteristic of sample members... to serve a purpose
  - ❖ When taking sample reject, people who do not fit for a particular profile
  - ❖ Start with a purpose in mind
-



# Advantage

---

Sample are chosen well based on the  
some criteria

There is a assurance of Quality  
response

Meet the specific objective

# Demerit

---

- ❖ Bias selection of sample may occur
- ❖ Time consuming process

# Quota Sampling

---

- ❖ The population is divided into cells on the basis of relevant control characteristics.
- ❖ A quota of sample units is established for each cell
- ❖ A convenience sample is drawn for each cell until the quota is met
- ❖ It is entirely non random and it is normally used for interview surveys

# Advantage

---

- Used when research budget limited
- Very extensively used/understood
- No need for list of population elements
- Introduces some elements of stratification

## Demerit

- Variability and bias cannot be measured or controlled
- Time Consuming
- Projecting data beyond sample not justified

# Snowball Sampling

---

- ❖ The research starts with a key person and introduce the next one to become a chain
- ❖ Make contact with one or two cases in the population
- ❖ Ask these cases to identify further cases.
- ❖ Stop when either no new cases are given or the sample is as large as manageable

# Advantage

---

- low cost
- Useful in specific circumstances
- Useful for locating rare populations

# Demerit

- Bias because sampling units not independent
- Projecting data beyond sample not justified

# Self selection Sampling

---

It occurs when you allow each case usually individuals, to identify their desire to take part in the research you therefore

Publicize your need for cases, either by advertising through appropriate media or by asking them to take part

Collect data from those who respond

# Advantage

---

- More accurate
- Useful in specific circumstances to serve the purpose

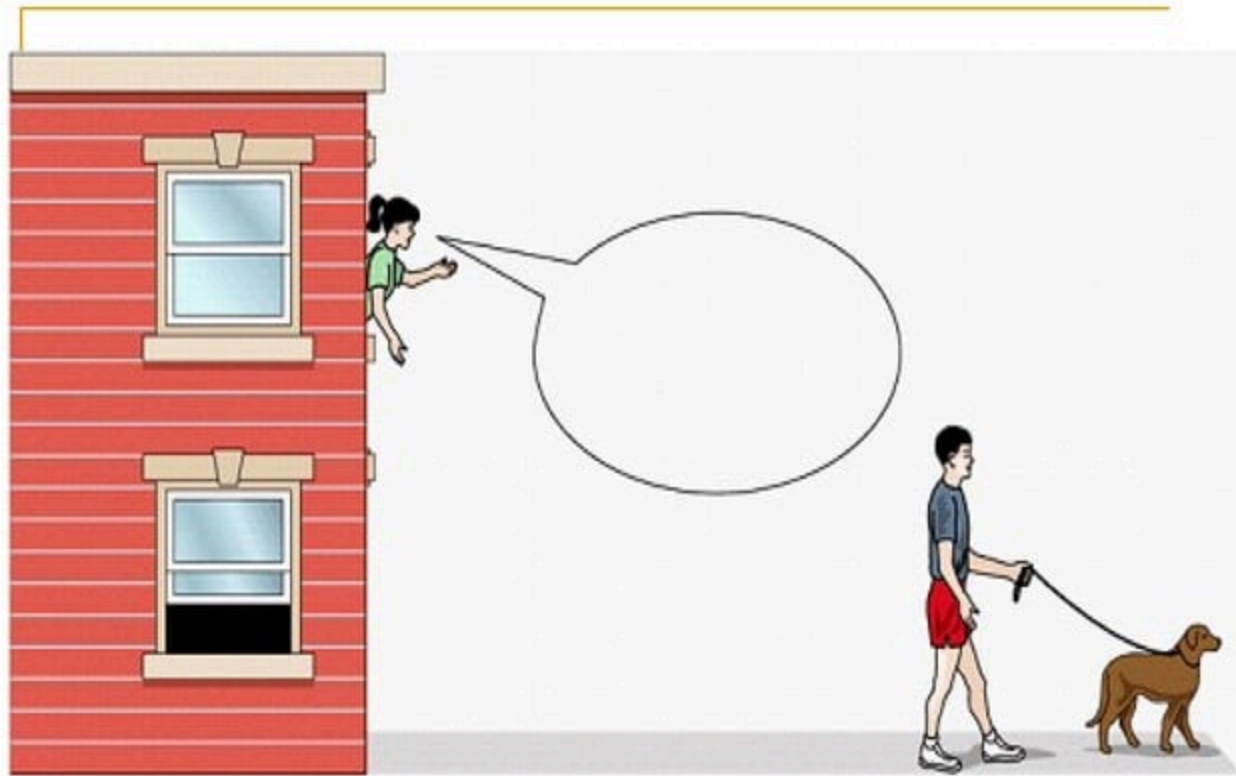
# Demerit

- More costly due to Advertizing
- Mass are left



# Convenience Sampling

- ❖ Called as Accidental / Incidental Sampling
- ❖ Selecting haphazardly those cases that are easiest to obtain
- ❖ Sample most available are chosen
- ❖ It is done at the “convenience” of the researcher



# Merit

---

- Very low cost
- Extensively used/understood
- No need for list of population elements

## Demerit

- Variability and bias cannot be measured or controlled
  - Projecting data beyond sample not justified
-

# Sampling Error

---

Sampling error refers to differences between the sample and the population that exist only because of the observations that happened to be selected for the sample

Increasing the sample size will reduce this type of error

**Error**



# Types of Sampling Error

---

❖ **Sample Errors**

❖ **Non Sample Errors**

# Sample Errors

---

- ❖ Error caused by the act of taking a sample
- ❖ They cause sample results to be different from the results of census
- ❖ Differences between the sample and the population that exist only because of the observations that happened to be selected for the sample
- ❖ Statistical Errors are sample error
- ❖ We have no control over

# Non Sample Errors

---

Not Control by Sample Size

❖ **Non Response Error**

❖ **Response Error**



# Non Response Error

---

A non-response error occurs when units selected as part of the sampling procedure do not respond in whole or in part

# Response Errors

A response or data error is any systematic bias that occurs during data collection, analysis or interpretation

- ❖ Respondent error (e.g., lying, forgetting, etc.)
- ❖ Interviewer bias
- ❖ Recording errors
- ❖ Poorly designed questionnaires
- ❖ Measurement error

# Respondent error

---

- ❖ respondent gives an incorrect answer, e.g. due to prestige or competence implications, or due to sensitivity or social undesirability of question
  - ❖ respondent misunderstands the requirements
  - ❖ lack of motivation to give an accurate answer
  - ❖ “lazy” respondent gives an “average” answer
  - ❖ question requires memory/recall
  - ❖ proxy respondents are used, i.e. taking answers from someone other than the respondent
-

# Interviewer bias

- ❖ Different interviewers administer a survey in different ways
- ❖ Differences occur in reactions of respondents to different interviewers, e.g. to interviewers of their own sex or own ethnic group
- ❖ Inadequate training of interviewers
- ❖ Inadequate attention to the selection of interviewers
- ❖ There is too high a workload for the interviewer

# Measurement Error

- ❖ The question is unclear, ambiguous or difficult to answer
- ❖ The list of possible answers suggested in the recording instrument is incomplete
- ❖ Requested information assumes a framework unfamiliar to the respondent
- ❖ The definitions used by the survey are different from those used by the respondent (e.g. how many part-time employees do you have? See next slide for an example)

# Key Points on Errors

---

Non-sampling errors are inevitable in production of national statistics. Important that:-

- ❑ At planning stage, all potential non-sampling errors are listed and steps taken to minimise them are considered.
  - ❑ If data are collected from other sources, question procedures adopted for data collection, and data verification at each step of the data chain.
  - ❑ Critically view the data collected and attempt to resolve queries immediately they arise.
  - ❑ Document sources of non-sampling errors so that results presented can be interpreted meaningfully.
-



**THANKS**