RESEARCH METHODOLOGY

Sampling Design

Raghu B. A.
Department of Computer Science and Engineering
raghubarao@pes.edu

RESEARCH METHODOLOGY

Ch. 4: Sampling Design

Raghu B. A.
Department of Computer Science and Engineering

CONTENTS

INTRODUCTION

- Items in field of enquiry – Universe / Population
- Enumeration of all items in population = Census Survey.
  - Ex: Avg monthly income of people of India
  - When all objects are covered -> highest accuracy
  - Possible? Cost, Bias?
- Census is impossible in situation where population is infinite.
- Select only few items = Sample
- Procedure for selection = Sampling Technique.

- A **sample** is “a smaller (but hopefully representative) collection of units from a population used to determine truths about that population” (Field, 2005)

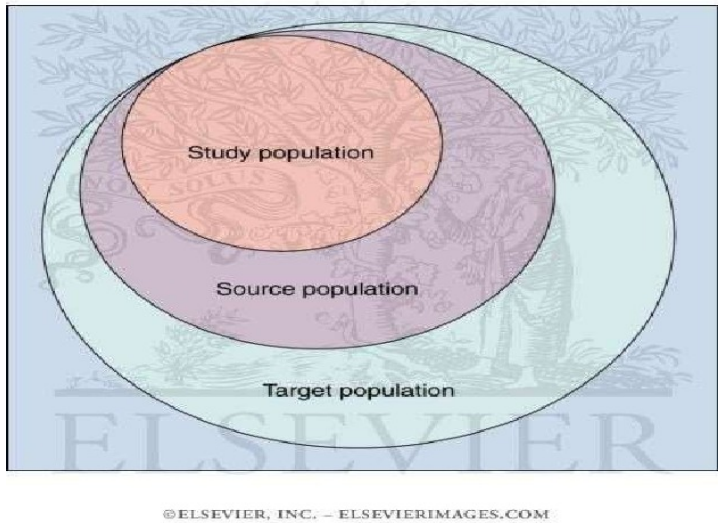
- The **sampling frame** A list of all elements or other sampling units containing the elements in a population.

POPULATION

The larger group from which individuals are selected to participate in a study.

TARGET POPULATION

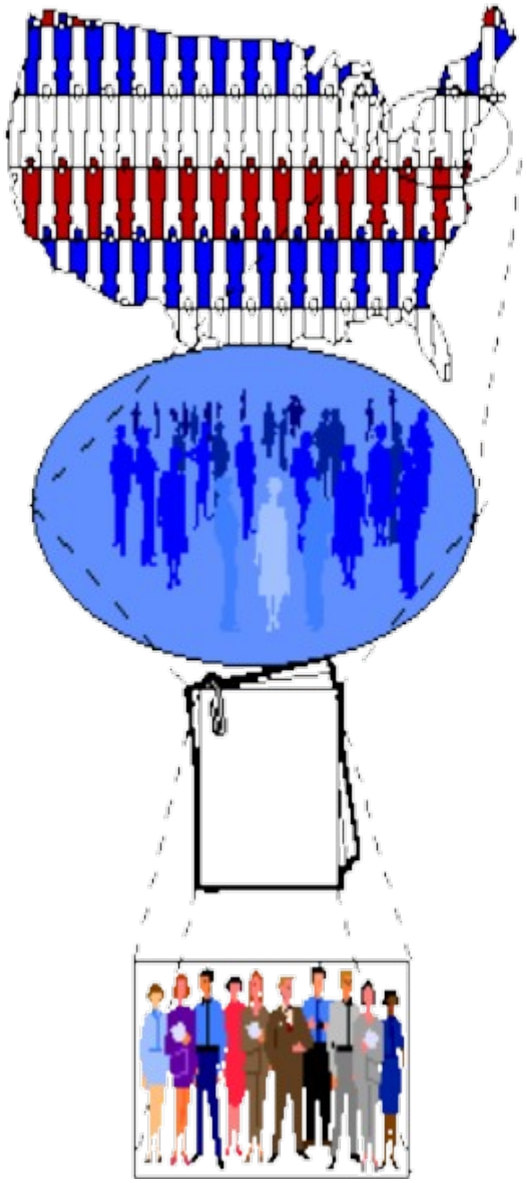
A set of elements larger than or different from the population sampled and to which the researcher would like to generalize study findings.



POPULATION TO SAMPLE

Who do you want to generalize to?

The Theoretical Population



What population can you get access to?

The Study Population

How can you get access to them?

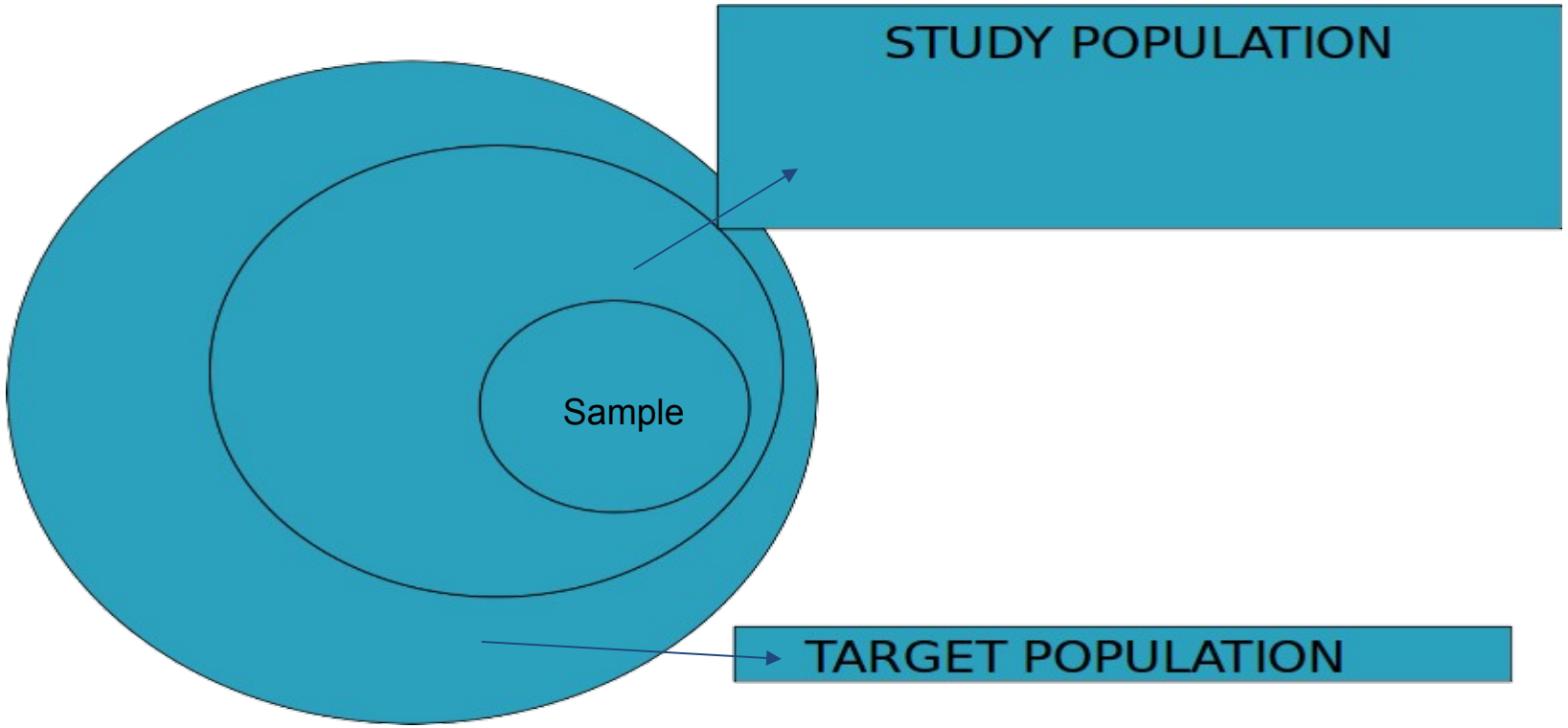
The Sampling Frame

Who is in your study?

The Sample

SAMPLING

Definition: "*The process of selecting a number of individuals for a study in such a way that the individuals represent the larger group from which they were selected.*"



SAMPLE DESIGN

- Plan for obtaining sample from a population
- Technique / procedure a researcher will adopt in selecting sample.
- Sample design is determined before data are collected.
- Many samplle designs are available:
  - Some more precise and easier to apply.
  - Selected sample design must be reliable and appropriate for the research study.

MAIN STEPS OF SAMPLING

1) Objective
2) Population
3) Sampling Unit and Frame
4) Size of Sample
5) Parameter of Interest
6) Data Collection
7) Non respondents
8) Selection of proper sampling design
9) Organizing field work
10) Pilot Survey
11) Budgetary Constraints

MAIN STEPS OF SAMPLING

Objective :

- Define the objective of study. It should be in proportion with money, manpower and time.

Population (or Universe):

- Should be clearly defined.

Sampling Units and Frames:

- Sampling unit – for study
  - geographic unit – state, district, state or
  - construction unit – flat, house, or
  - social unit – house, flat

Source List – same as **Sampling Frame**

MAIN STEPS OF SAMPLING

Size of Sample

- Number of units to be selected.
- Size should be neither too large nor too small
- Optimal size –> efficiency, flexibility, reliability.

Parameters of Interest

- Statistical constants of population – parameters.
- Ex - mean population, population proportion

MAIN STEPS OF SAMPLING

Data Collection:

- Only relevant information to be collected.
- Objective to be made clear

Non respondents

- Because of practical difficulties, data may not be collected for all sampled units.
- Non responses tend to change results

MAIN STEPS OF SAMPLING

Selection of Proper Sampling Design :

- Must decide technique in selecting the items for sample.
- Must yield less error

Organizing field work:

- Success depends on reliable field work.
- There should be efficient supervisory staff and trained personnel for field work.

MAIN STEPS OF SAMPLING

Pilot Survey

- Try research on small scale before going to field.

Budgetary Constraints

- Cost consideration, from practical viewpoint.
- Have major impact upon decision relating to not only size but also types of sample.

CHARACTERISTICS OF A GOOD SAMPLE DESIGN

(a) Sample design must result in a truly representative sample.
(b) Sample design must be such which results in a small sampling error.
(c) Sample design must be viable in the context of funds available for the research study.
(d) Sample design must be such that systematic bias can be controlled in a better way.
(e) Sample should be such that the results of the sample study can be applied, in general, for the universe with a reasonable level of confidence.

RESEARCH METHODOLOGY

Ch. 4: Sampling Design

Raghu B. A.
Department of Computer Science and Engineering

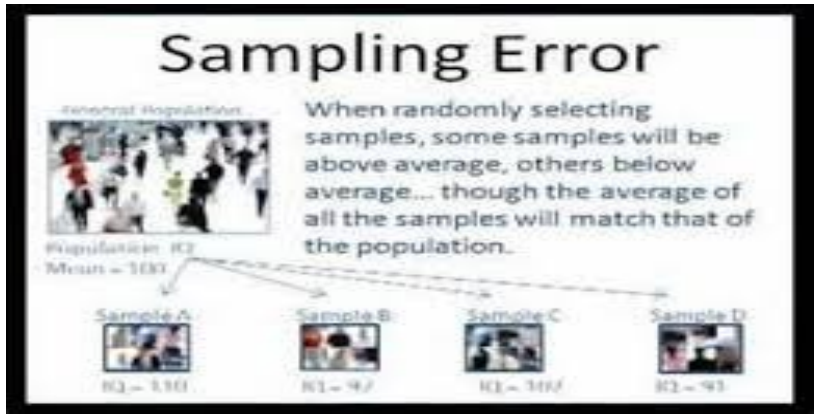
NEED FOR SAMPLING

Sampling --

1) Can save time & money – faster than census; less expensive, faster results.
2) Enables more accurate measurements – trained / experienced investigators.
3) Only way remaining when population is infinite.
4) Only choice left when test involves the destruction of the item (sample) under study.
5) Usually enables to estimate the sampling errors – helps to get info about some characteristic of the population.

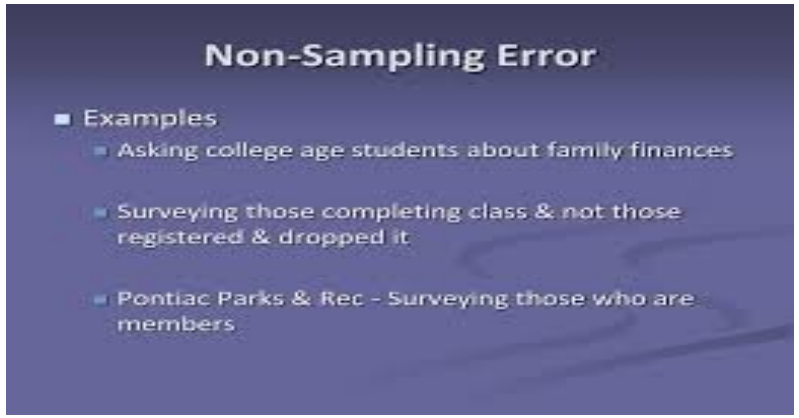
SAMPLING ERROR

- Only a part of population has been used for estimation and to draw inference.
- Sampling errors are absent in census survey
- Can be measured for a given size and sample design = ***precision*** of sampling plan.
- If we increase the sample size, precision can be improved.
- Also increases cost and systematic bias.
- Sample error = Frame error + Chance error + Response error



NON- SAMPLING ERROR

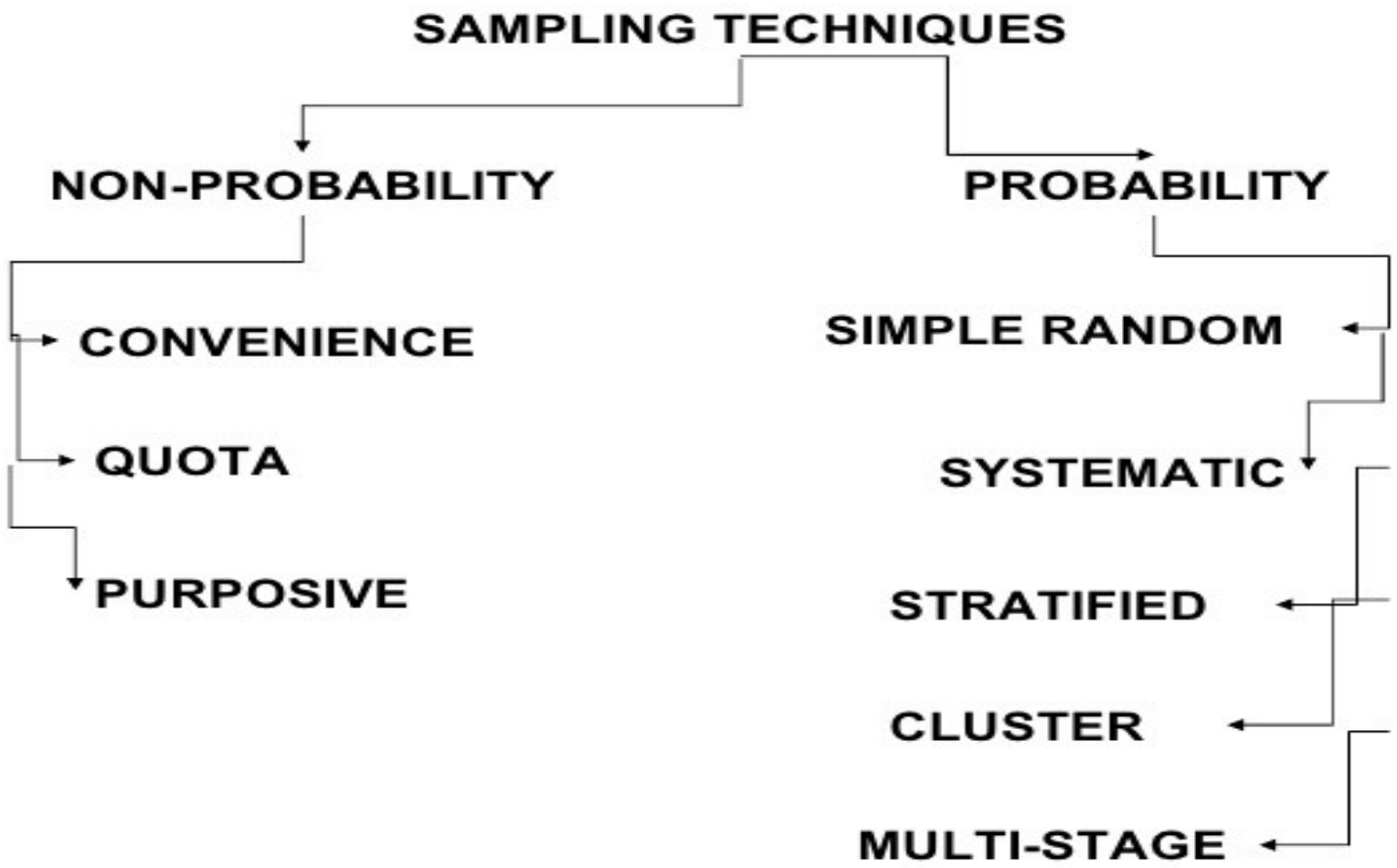
- Arises at the stage of collection and preparation of data.
- Present in both sample & census survey
- Can be reduced by defining the sampling unit, frame and population correctly.
- Total error = Sampling error + non sampling error



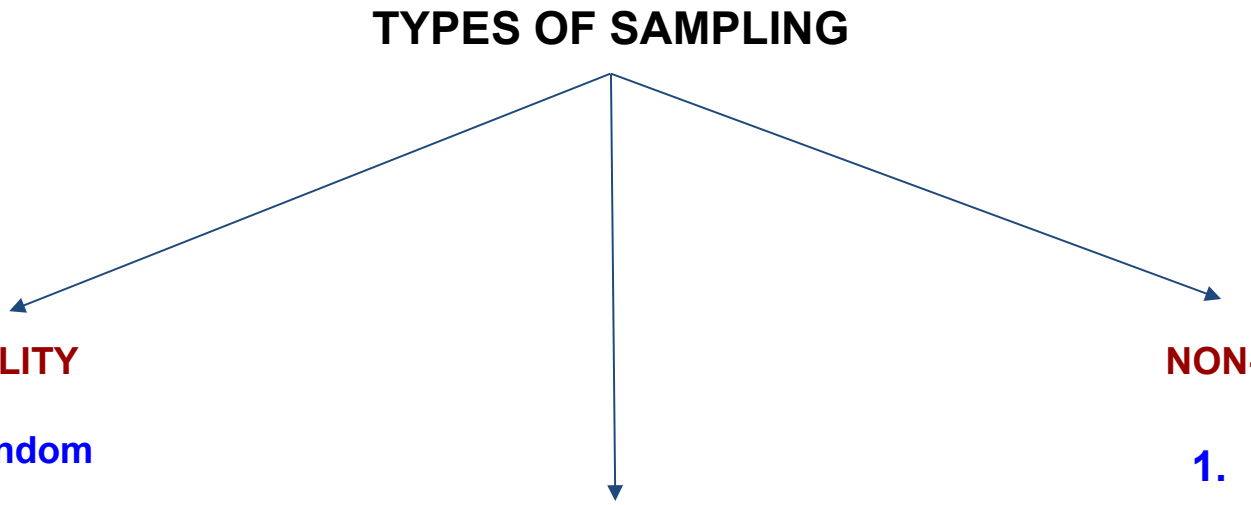
SAMPLING TECHNIQUES

PES UNIVERSITY ONLINE

TYPES OF SAMPLE DESIGNS



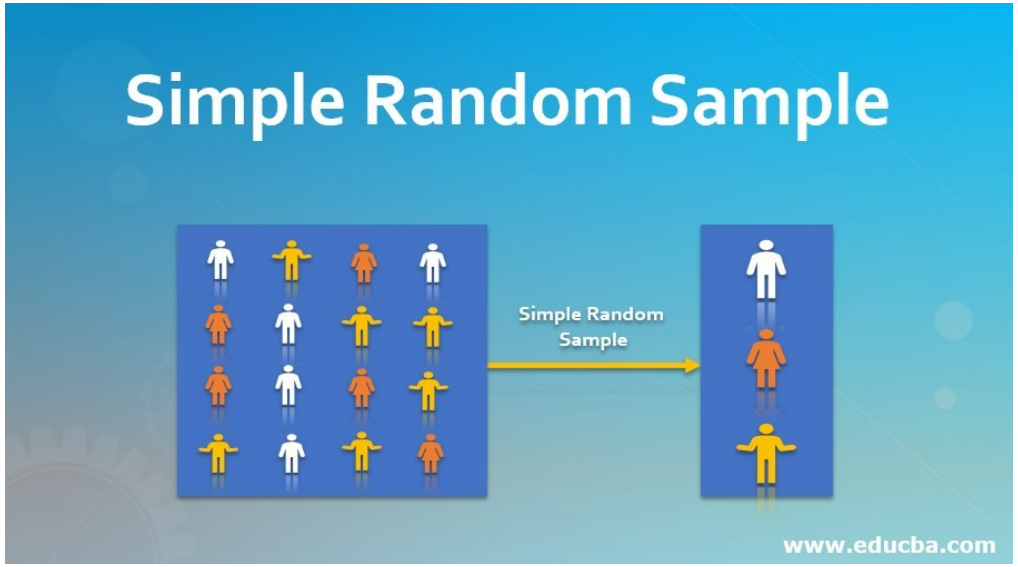
TYPES OF SAMPLE DESIGNS

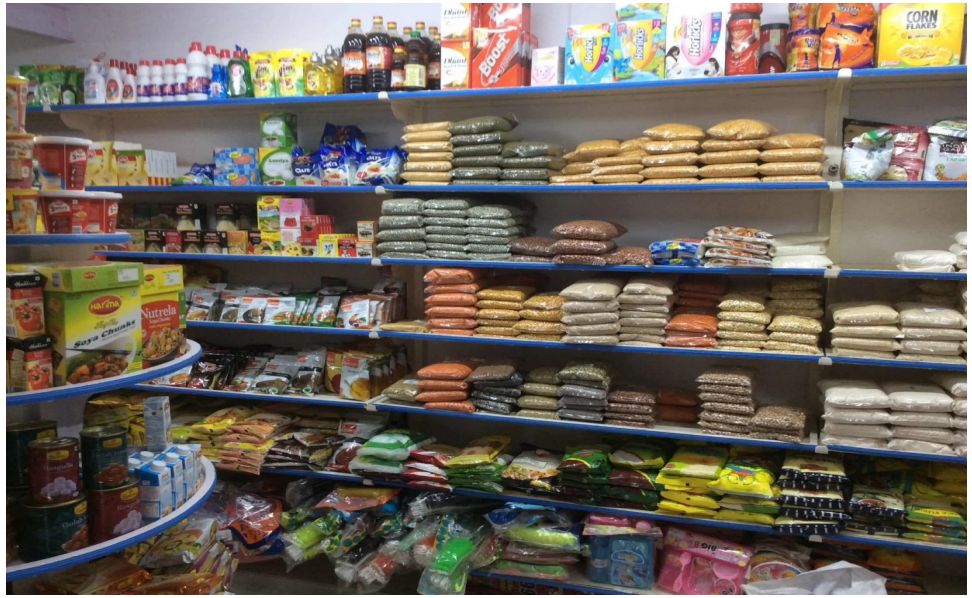


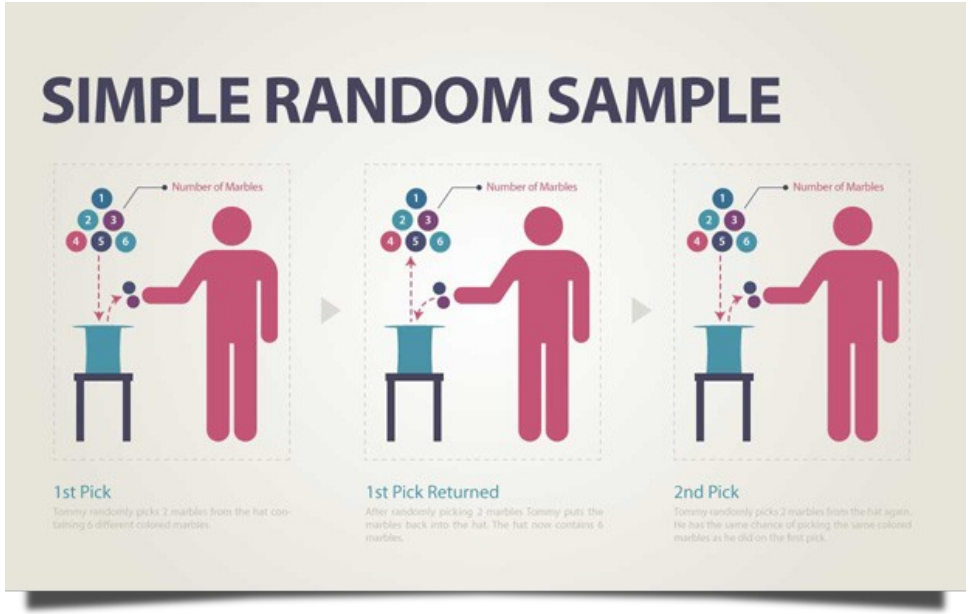
SIMPLE RANDOM SAMPLING

- It is the purest form of probability sampling.
- Each member of the finite population has an equal probability and known chance of being selected.
- Sampling without replacement
- When there are very large populations?
    - it is often difficult
    - impossible to identify every member of the population
    - The pool of available subjects becomes biased.
- Examples :
    - Lottery or Fishbowl
    - Survey of insect population living in woodland
    - Trees no. 1 to 1200. 10 trees are chosen at random
- Random number tables available [by Tippett, Yates, Fisher]

SIMPLE RANDOM SAMPLING





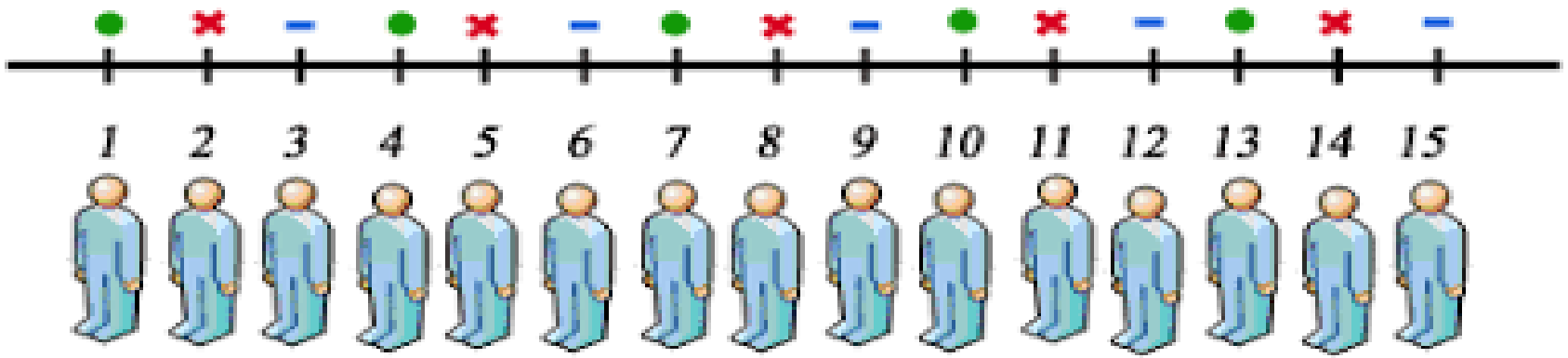


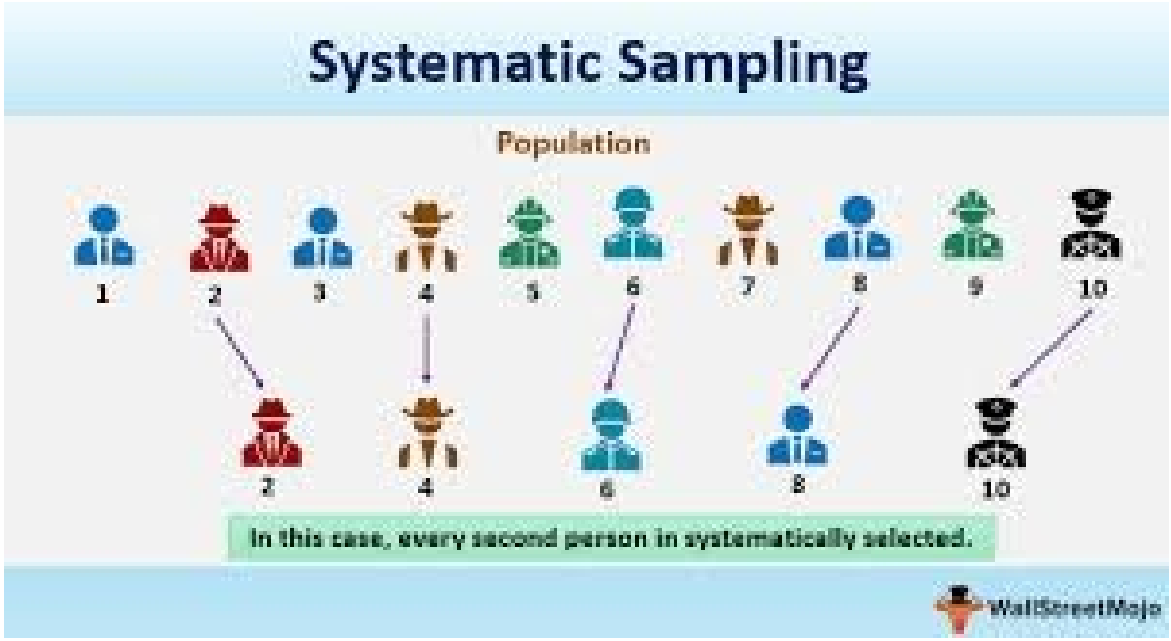
COMPLEX RANDOM SAMPLING DESIGNS

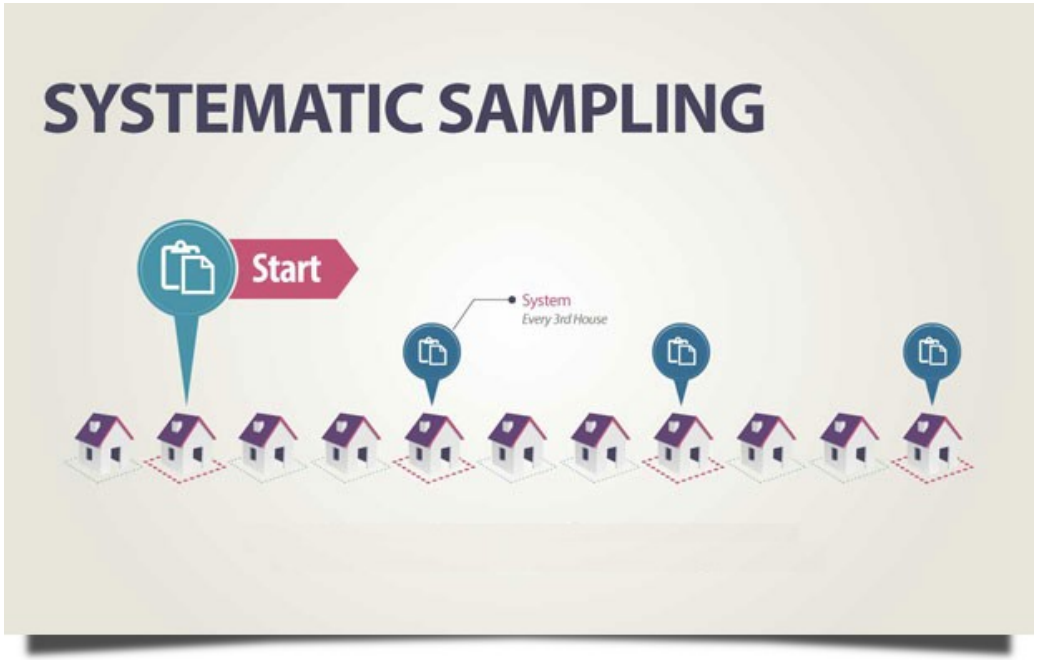
SYSTEMATIC SAMPLING

- It is often used instead of random sampling.
- It is also called an Nth name selection technique.
- After the required *sample size* has been calculated, every Nth record is selected from a list of population members.
- Randomness introduced in picking the start point.
- As long as the list *does not contain any hidden order*, this sampling method is as good as the random sampling method.
- It's only advantage over the random sampling technique is simplicity, easy, cheap. And applicable for large populations.
- Systematic sampling is frequently used to select a specified number of records from a computer file.

SYSTEMATIC SAMPLING







SYSTEMATIC SAMPLING

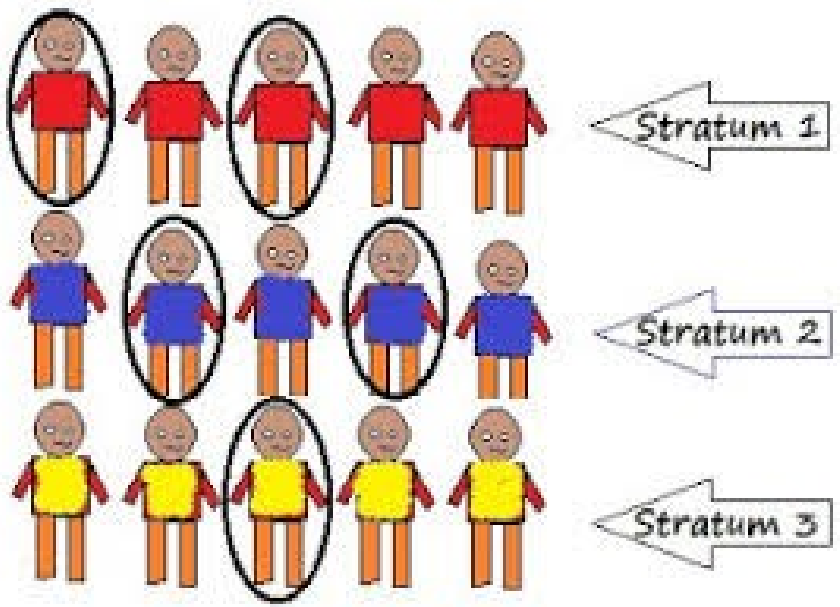
Example: to select a sample of 25 dorm rooms in your college hotel, make a list of all the room numbers in the dorm.

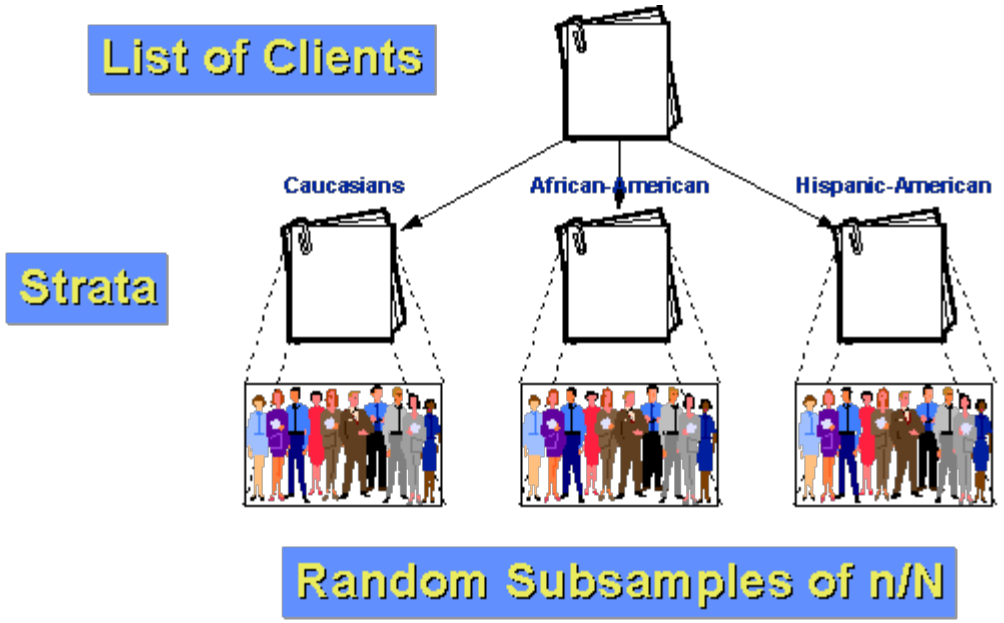
- For example if there are 100 rooms.
- divide the total number of rooms (100) by the number of rooms you want in the sample (25). The answer is 4.
- This means that you are going to *select every fourth* dorm room from the list.
- First of all, we have to determine the random starting point.
- This step can be done by picking any point on the table of room numbers, and read across or down until you come to a number between 1 and 4. This is your random starting point.
- For instance, your random starting point is "3". This means you select dorm room #3 as your first room, and then every fourth room down the list (3, 7, 11, 15, 19, etc.) until you have 25 rooms selected.

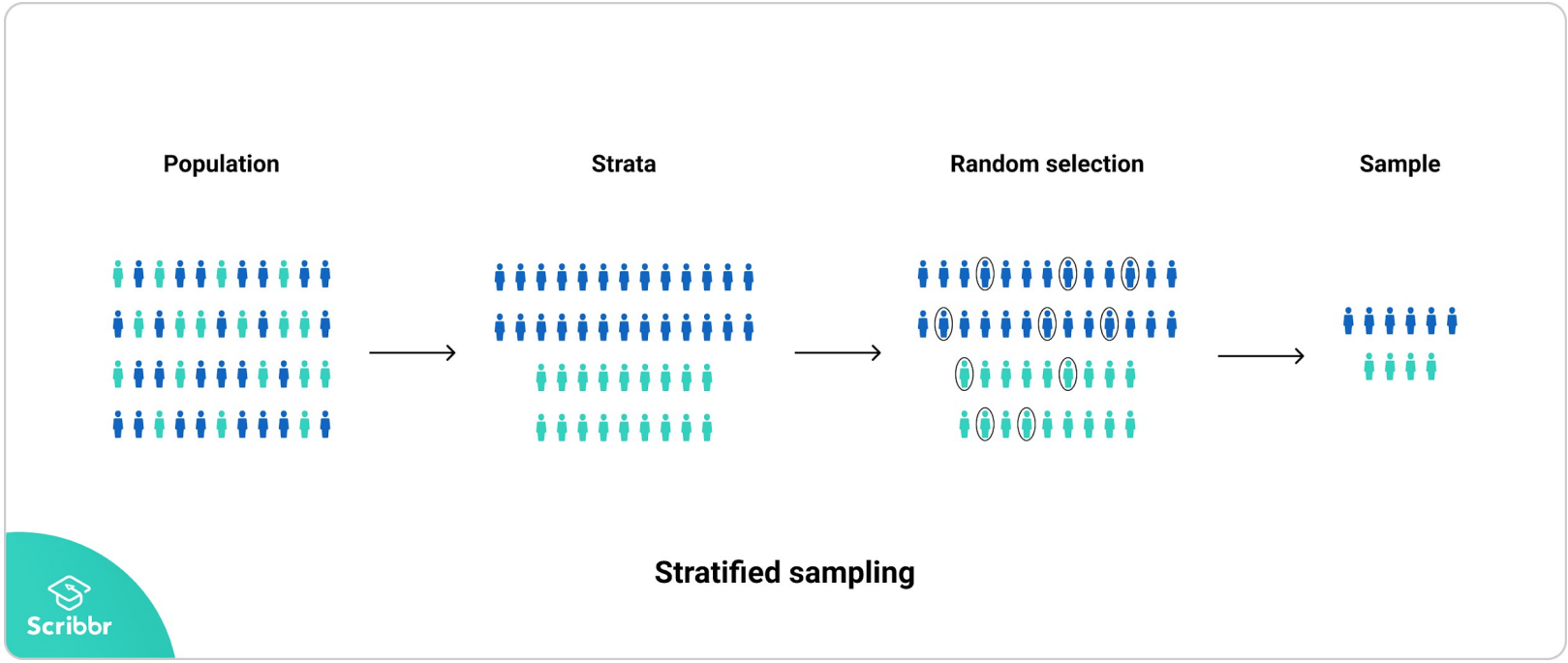
STRATIFIED SAMPLING

- Chosen when sample to be drawn doesn't constitute homogeneous group.
- Population is divided into sub-populations that are individually homogeneous - *strata*
- Stratified Sampling is possible when it makes sense to partition the population into groups based on a factor that may influence the variable that is being measured.
- These groups are then called *strata*. Based on one or more characteristic.
- An individual group is called a stratum. With stratified sampling one should:
    - partition the population into groups (strata)
    - obtain a simple random sample from each group (stratum)
    - collect data on each sampling unit that was randomly sampled from each group (stratum)
    - Called Stratified random sampling
- Stratified sampling works best when a heterogeneous population is split into fairly homogeneous groups.

STRATIFIED SAMPLING







STRATIFIED SAMPLING

	Example 1	Example 2	Example 3
Population	All people in US	All PSU intercollegiate athletes	All elementary students in the local school district
Groups (Strata)	4 Time Zones in the U.S. (Eastern,Central, Mountain,Pacific)	26 PSU intercollegiate teams	11 different elementary schools in the local school district
Obtain a Simple Random Sample	500 people from each of the 4 time zones	5 athletes from each of the 26 PSU teams	20 students from each of the 11 elementary schools
Sample	4 × 500 = 2000 selected people	26 × 5 = 130 selected athletes	11 × 20 = 220 selected students

CLUSTER SAMPLING

- Used when population is divided into groups or clusters
- Samples are selected from groups rather than individuals which is employed into large scale survey.

Advantages:

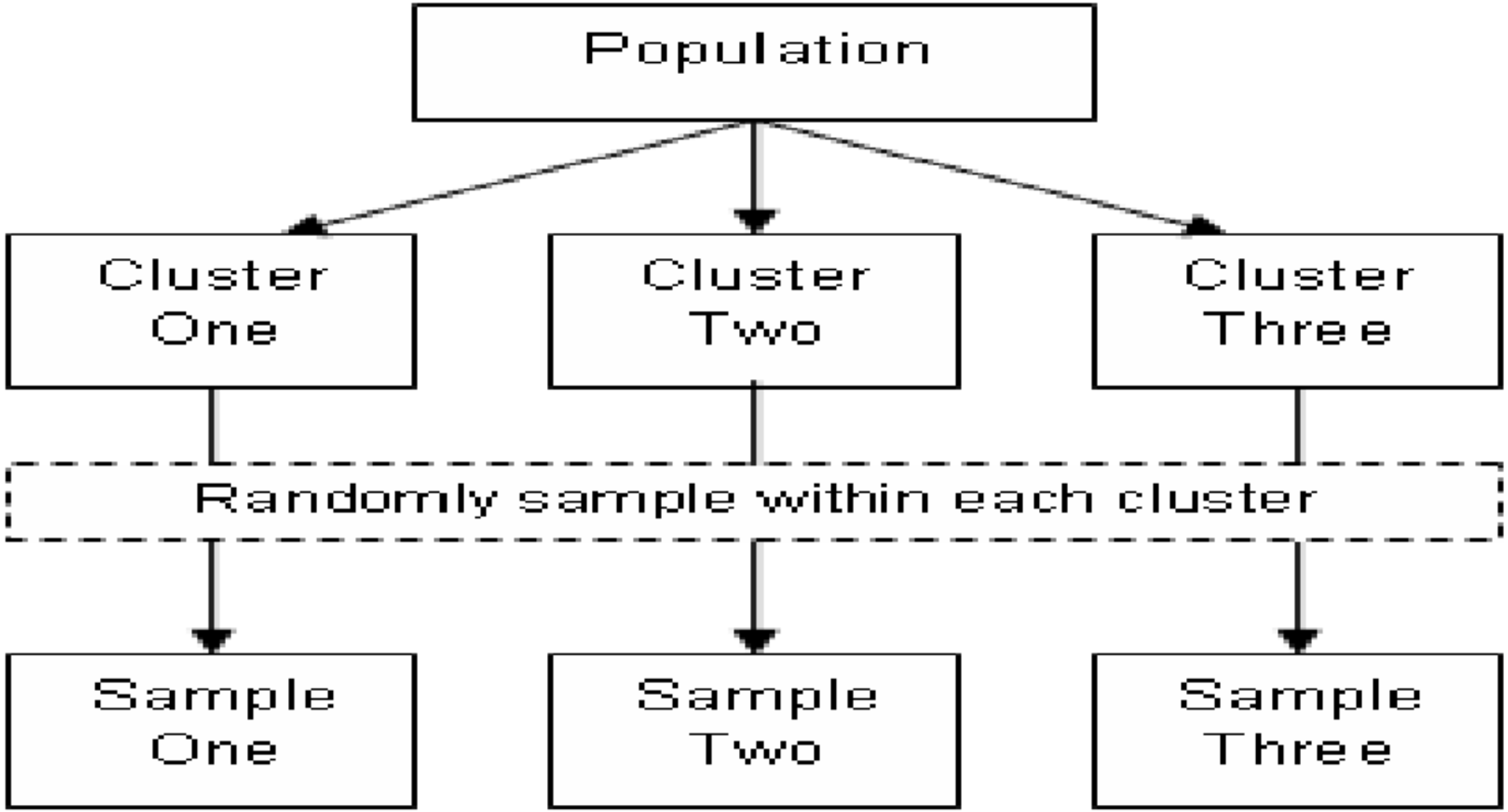
- Very useful when population is spread over large geographic area.
- Convenient and expedient
- Does not need names of everyone in population.
- Reduced cost

Disadvantages:

- Representation is likely to become an issue
- Less precise than random sampling.

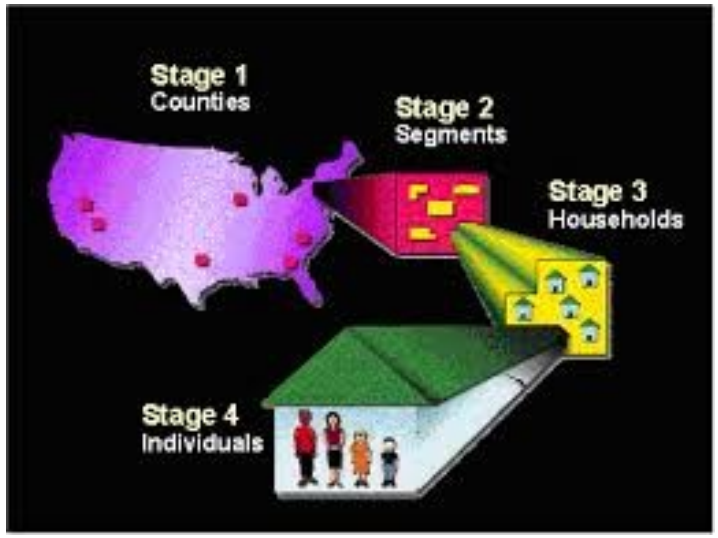
If clusters (or groups/subdivisions) are geographic areas, then this is called **Area Sampling**.

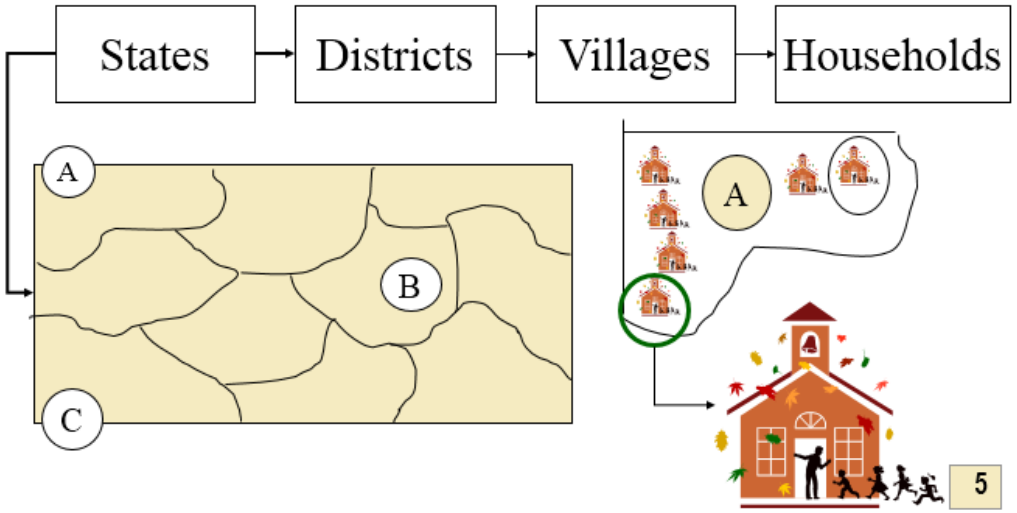
CLUSTER SAMPLING



MULTISTAGE SAMPLING

- Selects samples using more than 1 sampling technique.
- Complex - Hence rarely used.
- Requires lot of effort, time & cost.



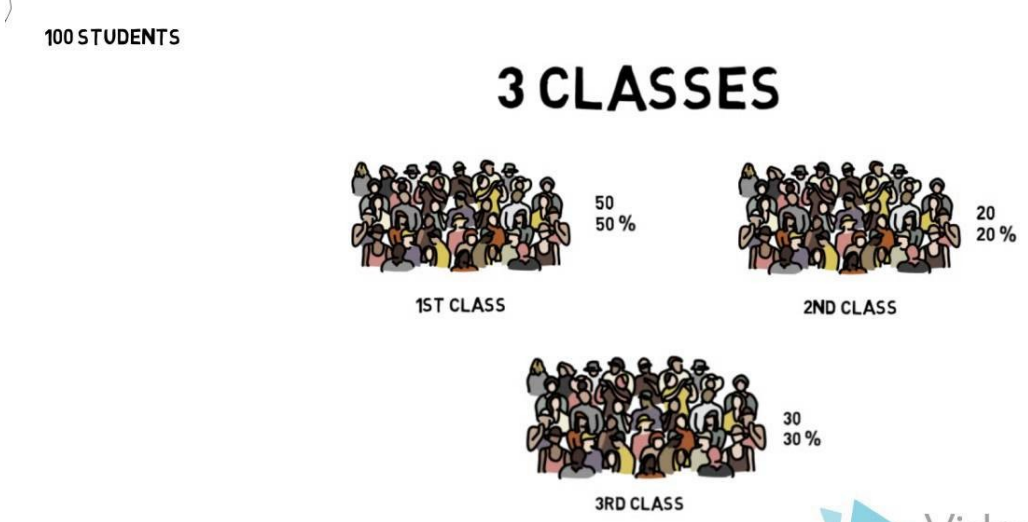


SEQUENTIAL SAMPLING

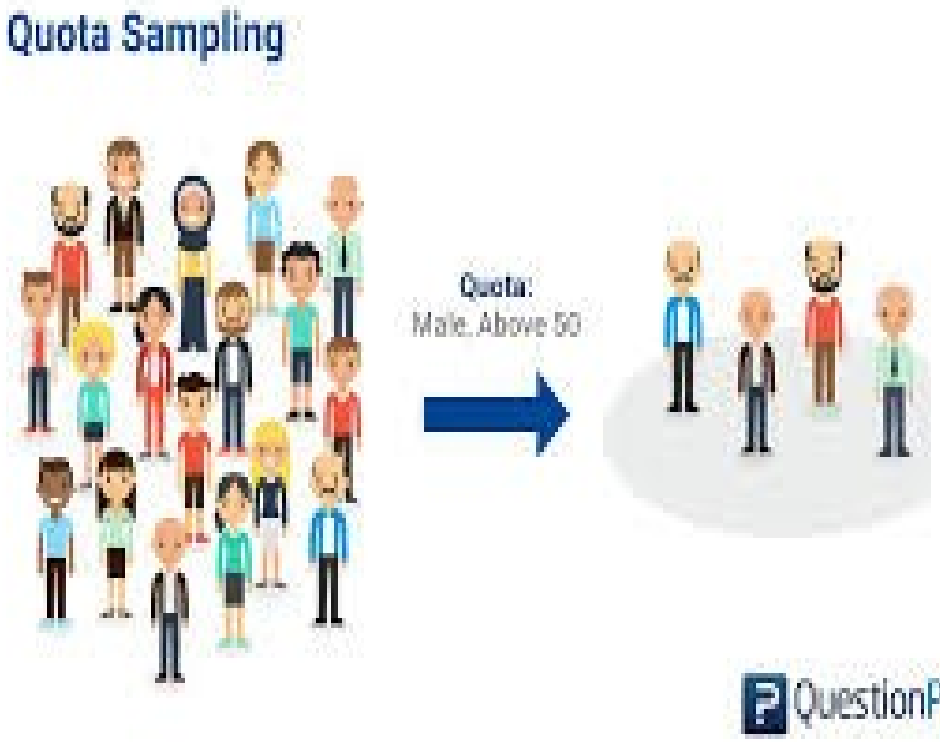
- Complex Sample design
- Size is not fixed in advance.
- Adopted in case of acceptance sampling.
    - When a particular lot to be accepted / rejected on basis of single sample - single sampling
    - When decision taken on basis of 2 samples - double sampling
- When number of samples are more than two, but neither certain nor decided in advance - Sequential Sampling

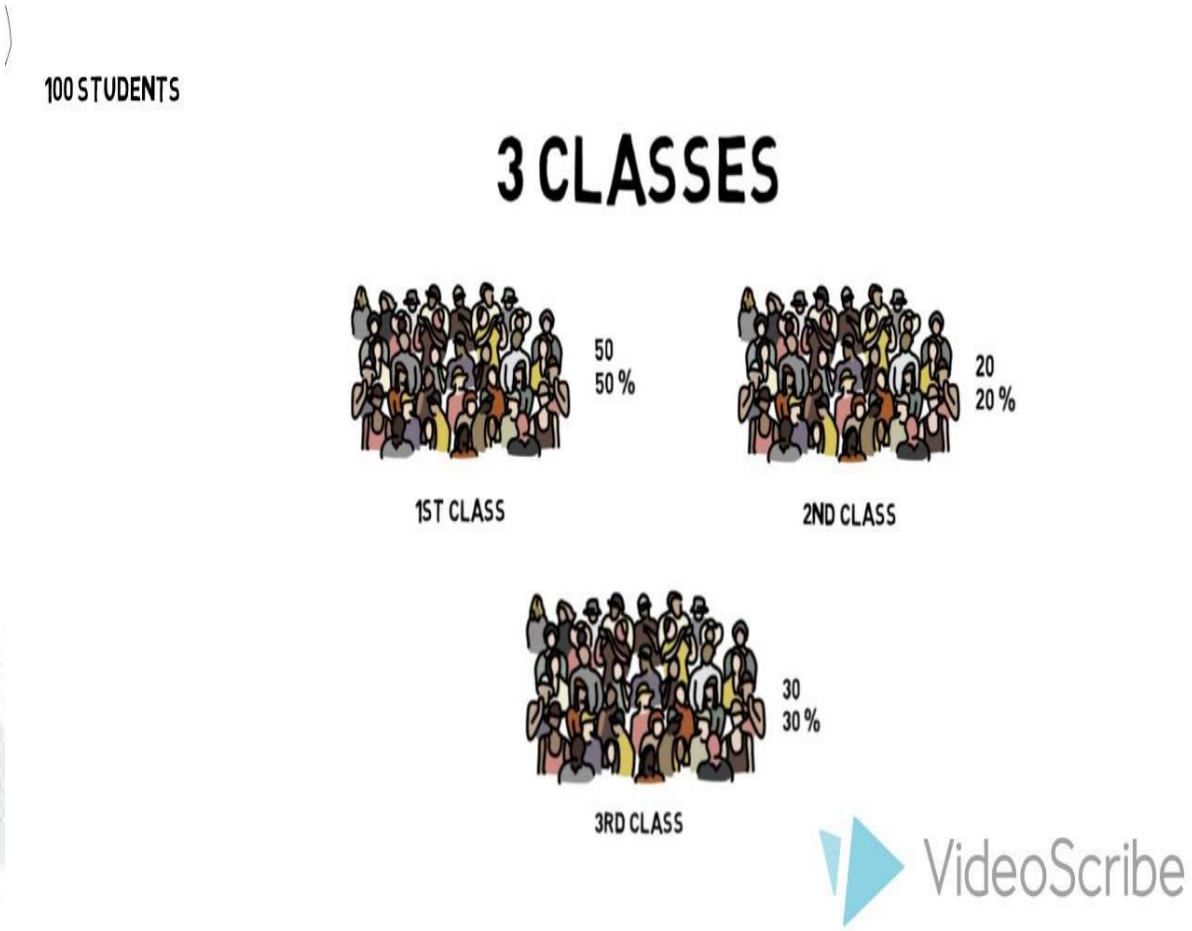
QUOTA SAMPLING

- This is a type of non probability sampling.
- Population is divided into mutually exclusive sub groups as in stratified sampling.
- Judgement is used to select the subject or unit from each segment based on specified portion.
- Quota sampling is judgement samples rather than random samples.



QUOTA SAMPLING





THANK YOU

Raghu B. A.
Department of Computer Science and Engineering

raghubarao@pes.edu