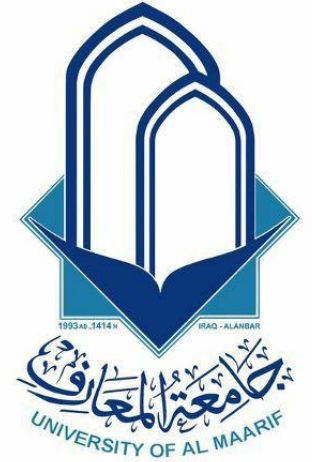




COMMUNITY (lecture 7)

Statistics

Dr. Mustafa Ibrahim
B.D.S M.Sc. Oral Medicine



Statistics: It is a method of collecting, classifying, summarizing and displaying a set of data, then analyzing and interpretation to find the results

Biostatistics:It is a branch of statistics responsible for the proper interpretation of scientific data generated in **clinical medicine, biology, public health and other health sciences.**

Epidemiology and biostatistics are sister sciences. Epidemiology collects facts relating to groups of population in places, times and situations, while biostatistics converts all facts into figures and at the end translates them into information.

Use of statistics in dentistry:

- To plan oral health measures.
- To indicate the basic factors and causation of oral diseases by diagnosing the community and solution of problems.
- For comparison and research.
- To determine success or failure (evaluate) of specific program of oral health care.

Data: Are any fact can be collected

Data are simply facts or figures — bits of information, but not information itself. **When data are processed, interpreted, organized, structured or presented so as to make them meaningful or useful, they are called information.** Information provides context for data

Variables: is a type of information/ characteristic can be measured and varies among the studied cases (study units).

- Types of data **depending on the nature of the variable** , data is classified into two categories:

1) **Quantitative data:or numerical data**, is information about quantities; that is information which can be measured and written down using **numbers**, like arch length, arch width, fluoride concentration in water supply, number of cigarette smoking per day,etc. it can be classified into two kinds :

- **Discrete:** when the variable under observation takes only **fixed values** like **whole numbers**, the data is discrete. e.g. DMF index.
- **Continuous:** if the variable can take **any value in a given range**, decimal or frictional, it is called as continuous data like arch length, height, weight etc.

2) Qualitative data

- when the data is collected on the basis of qualities (not in term of number), information that can't actually. Like gender, race , religion, color of eye, cavity types etc
- Qualitative data is a **categorical** measurement expressed not in terms of numbers, but rather by means of categories.

Qualitative data are classified as:

- 1-Nominal: if there is no natural order between the categories (e.g. eye color : blue, green, brown etc).
2. ordinal: if an ordering exists, Exam results: pass or fail, gingivitis: (mild, moderate, sever).

Methods of data collection:

The main sources of data are

1. Surveys
2. Experiments
3. past medical records

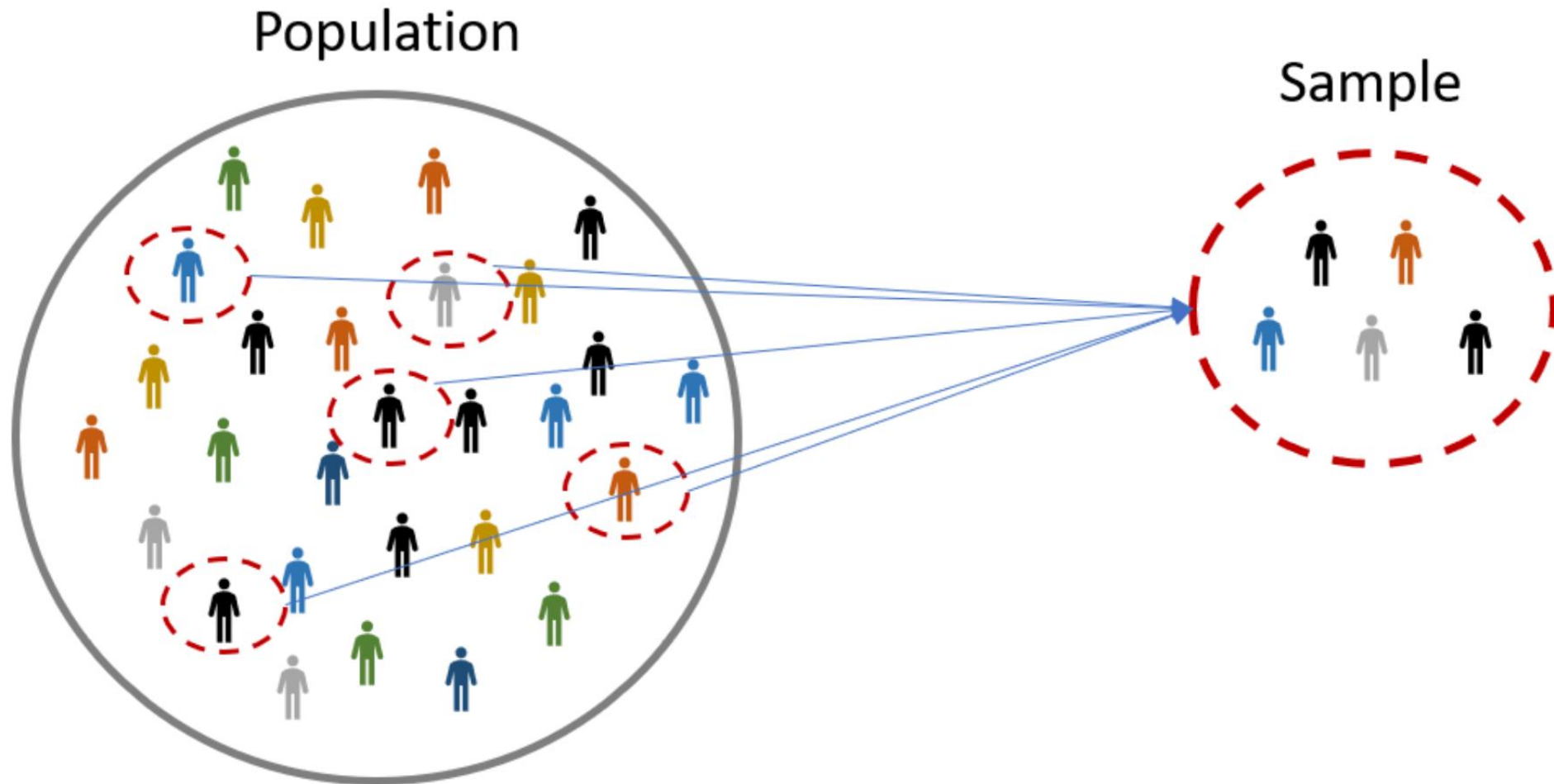
Data can be collected through:

a. **Primary source:** When the data is obtained by the investigator himself. It can be obtained using any one of the following methods:

1. Direct personal (face to face) interview
2. Oral health examination
3. Questionnaire method

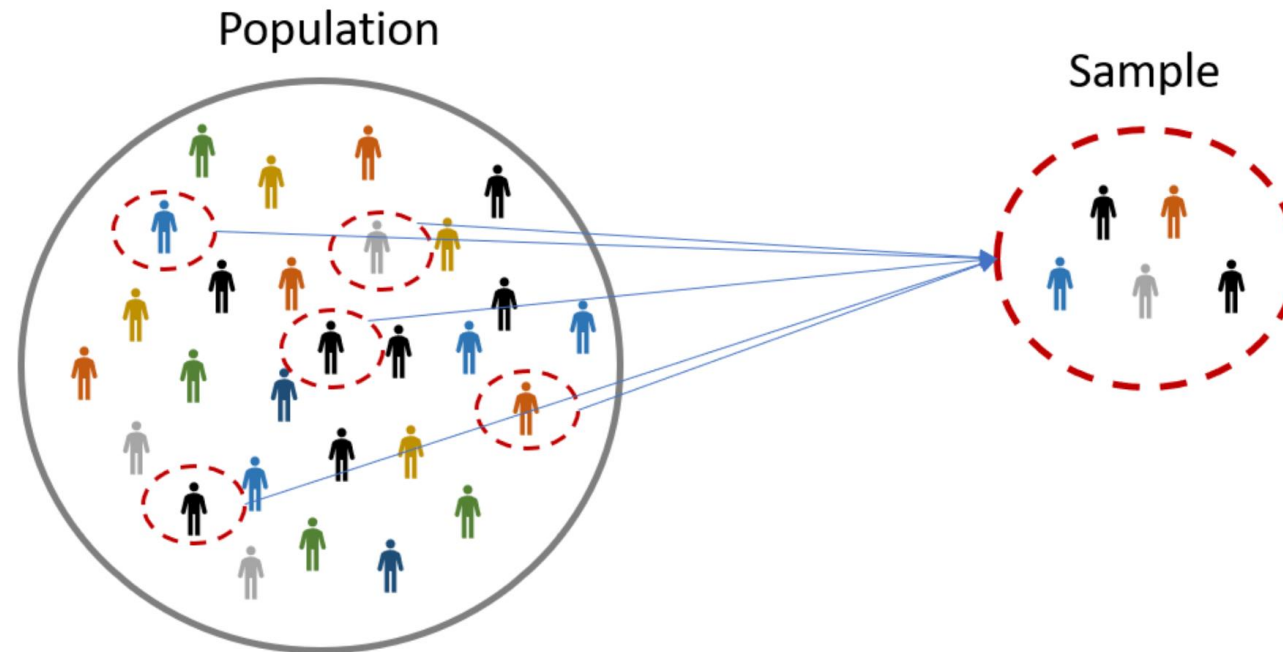
b. **Secondary source:** when the data is already recorded and it is used to serve the purpose of the objectives of the study

Sample: group of individuals who are actually available for the investigation.



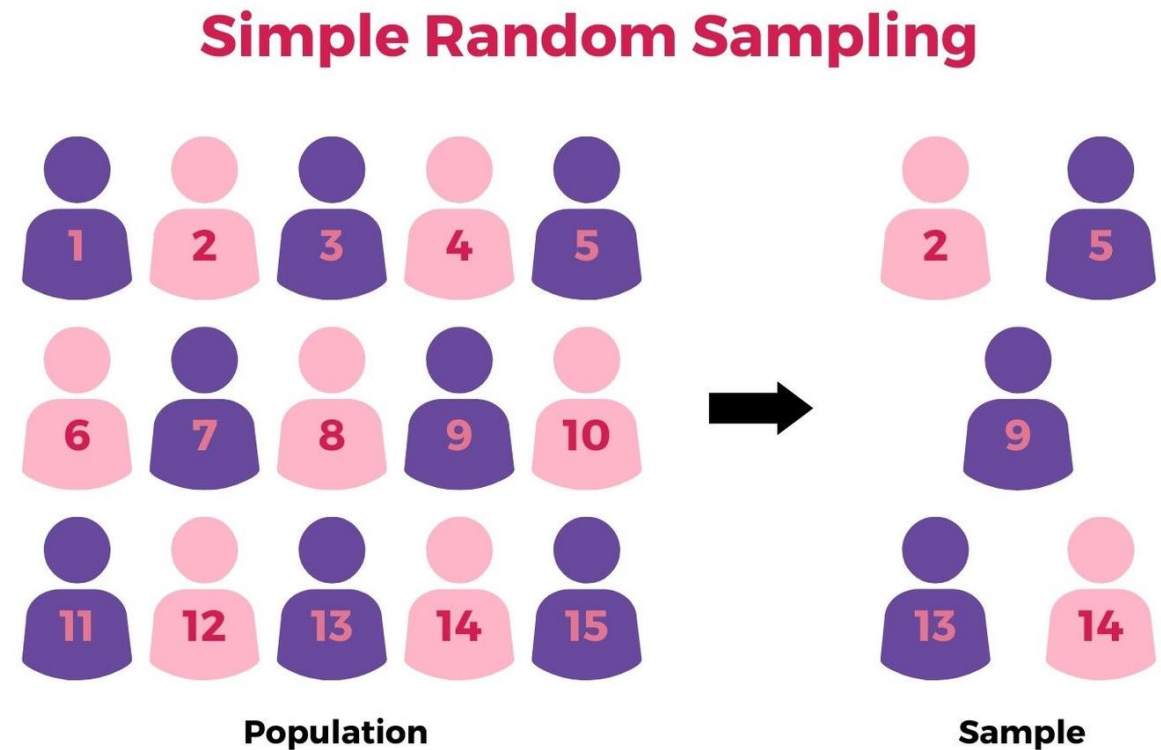
Sampling Technique:

- is the process of selecting a sample of appropriate characteristics and adequate size.
- It is difficult to study all members of population because it needs more time and resources. So, the sample that has been drawn needs to be **representative** of population from which it drawn.



The most sampling designs are:

- 1- **Simple Random Sampling:** In this method of sampling, each subject in the population has an **equal chance of being included in the sample**. In this method, selection of unit is determined by **chance only**.
- There are many ways to obtain a simple random sample. One way would be the **lottery method**. Each of N population members is assigned a **unique number (n)**. The numbers are placed in a bowl and thoroughly mixed. Then, a blind-folded researcher selects n numbers. Population members having the selected numbers are included in the sample.



2. Systematic Random Sampling

- Here, we create a **list of every member of the population**. From the list, we randomly select the first sample element randomly on the population list. Thereafter, we select the next element on the list to represent the interval so other elements will be selected at evenly spaced intervals till the sample size has been completed. (The interval determined by dividing the number of population (N) on the number of selected sample units (n)). (N/n)
- Ex: suppose there are 200 patients in the clinic and it is decided to select a sample of size 20 then $200/20$ gives quotient 10 so the interval will be 10, the first number should be selected randomly by simple random sampling, suppose the number will be 4 then the next one will be 14, 24, 34, 44 and so on till 20 numbers which is the size of the sample

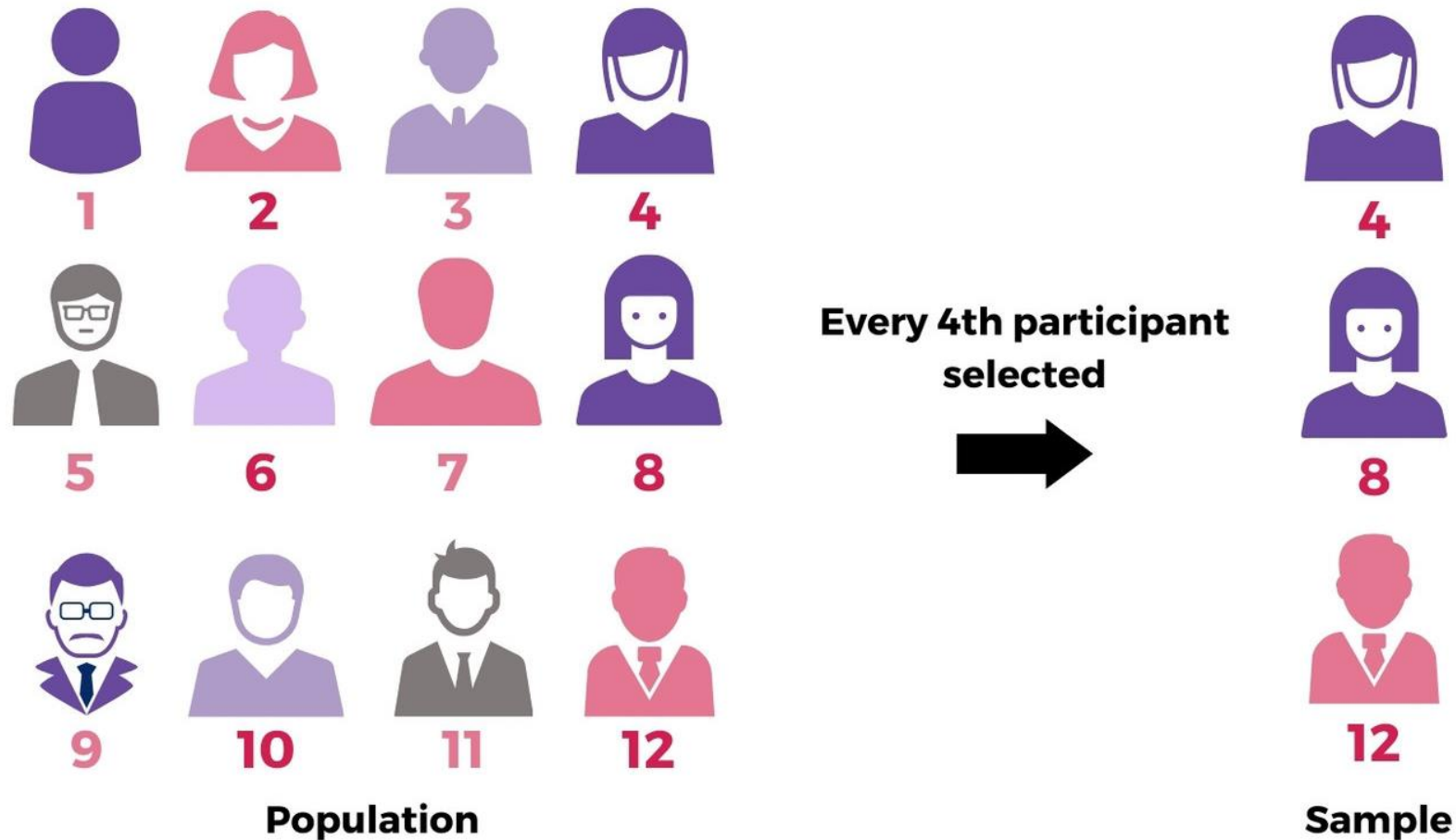
N=12

we want 3 sample out of 12

interval=N/n=12/3=4

if we choose the first sample 2 then the rest will be 6 and 10

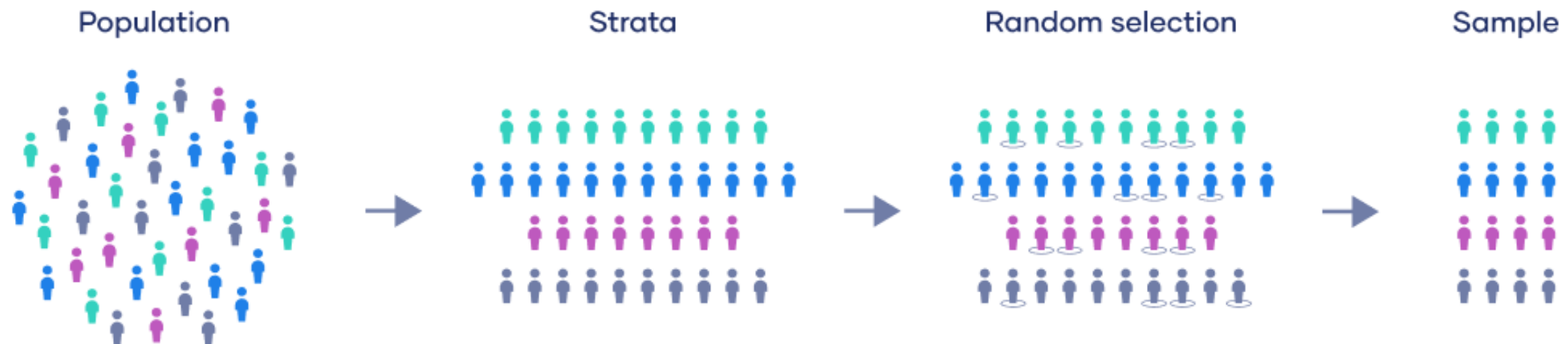
Systematic Random Sampling



Stratified Sampling

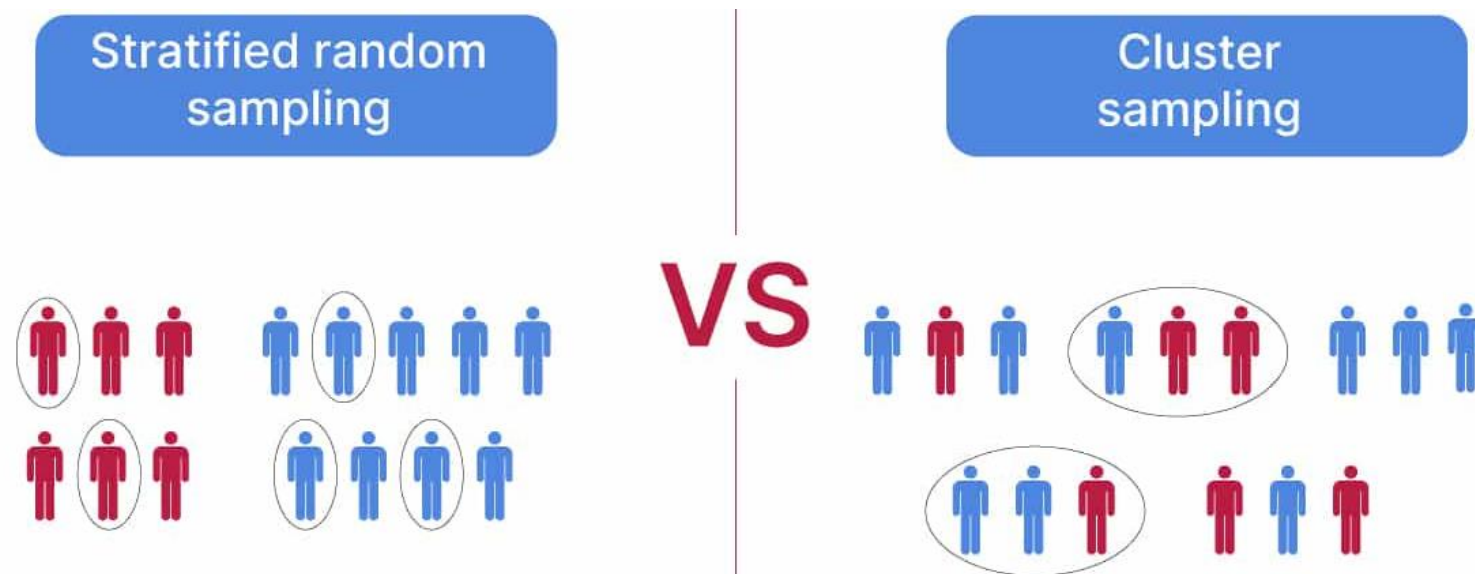
- the population is divided into groups, based on some characteristic (age, gender). Then, within each group, a simple random sample is selected. In stratified sampling, the groups are called **strata**.
- Ex: if we have national survey. We divide the population into groups or strata, based on geography to north, south, east and west. Then, within each stratum, we might randomly select the region to be surveyed

Stratified sampling



4 -Cluster sampling

- This method is used when the population forms **natural group or cluster, like villages, children of school, workers ect.**
- Here, first a sample of clusters is selected using simple random sampling and then **all the units** in each of the selected clusters are surveyed. This method is simpler and involves less time and cost.
- The difference between cluster sampling and stratified sampling, with stratified sampling, **the sample includes elements from each stratum. With cluster sampling, in contrast, the sample includes elements only from the sampled clusters.**



Multiphase sampling

- in this method, a part of the information is collected from the whole sample and a part from the sub- sample.
- For example, in a school health survey all the children in the school may be examined. From these, only the ones with oral health problems may be selected in the second phase.
- A section needing treatment may be selected in the third phase. Numbers in the subsamples in the third and fourth phase becomes smaller and smaller.

Sample size

- **It is the number of participants in a sample.**
- It is an important feature of any practical study in which the goal is to make inferences about a population from a sample. **A larger sample can yield more accurate results.**
- **The sample size should be defined before starting** a clinical study so as to **avoid bias** in interpreting results. If we have very few subjects in a study, the result cannot be generalized to the population as this sample will not represent the size of the target population. Further, the study then may not be able to detect the difference between test groups, making the study unethical.
- On the other hand, if the study include more subject than required, waste precious recourses including the researchers' time.

Errors Of Sampling

There are two types of errors that arise in sampling investigation :

1- *The sampling errors* which occur due to the sampling process, including:

- (i) Faulty sampling design.
- (ii) Small size of the sample.

2- *The non-sampling errors* arise due to:

a) Coverage error—due to

- (i) Non-response of the informant.
- (ii) Non-cooperation of the informant.

b) Observational error —due to

- 1-interviewers bias
- 2-imperfect experimental technique
- 3-Interaction of both. -

c) Processing error —due to errors in statistical analysis.

THANK YOU