Population or " universe "(unobserved) *: a set of data consist of all hypothetically possible observation of a given phenomenon to whom we limit any discussion or inference .*

Is a group of entities , having some quantifiable characteristic in common

The group (unit) may be people, machines, animal, bacteria ....Or it is not limited to population

They may be finite or infinite in No.

When might you sample the entire population?

1. population is very small

2. have extensive(large) resources

3. don't expect a very high response

4. Census

Sample (observed) *: a subset of population it should be representative to whole population.*

A sample is a smaller and used to determine truths about that population.

Its some finite No.

statistical inference: Making inference from a sample to a population

Why sample?

1. Resources (time, money) and workload

2. Gives results with known accuracy that can be calculated mathematically

The sampling frame: is the list from which the potential respondents are drawn.

which has the property that we can identify every single element and include any in our sample.

	Population	*Sample*
Size	Impossible & impractical	Possible & practical
Cost	High	Less
Observed	Not all be observed	All be observed
Staff	Large size	Smaller
Time	More	Less
Effort	More	Less
Accurate	Less	More
		• work lighter • uniform way • highly skilled • More precision (more trained)

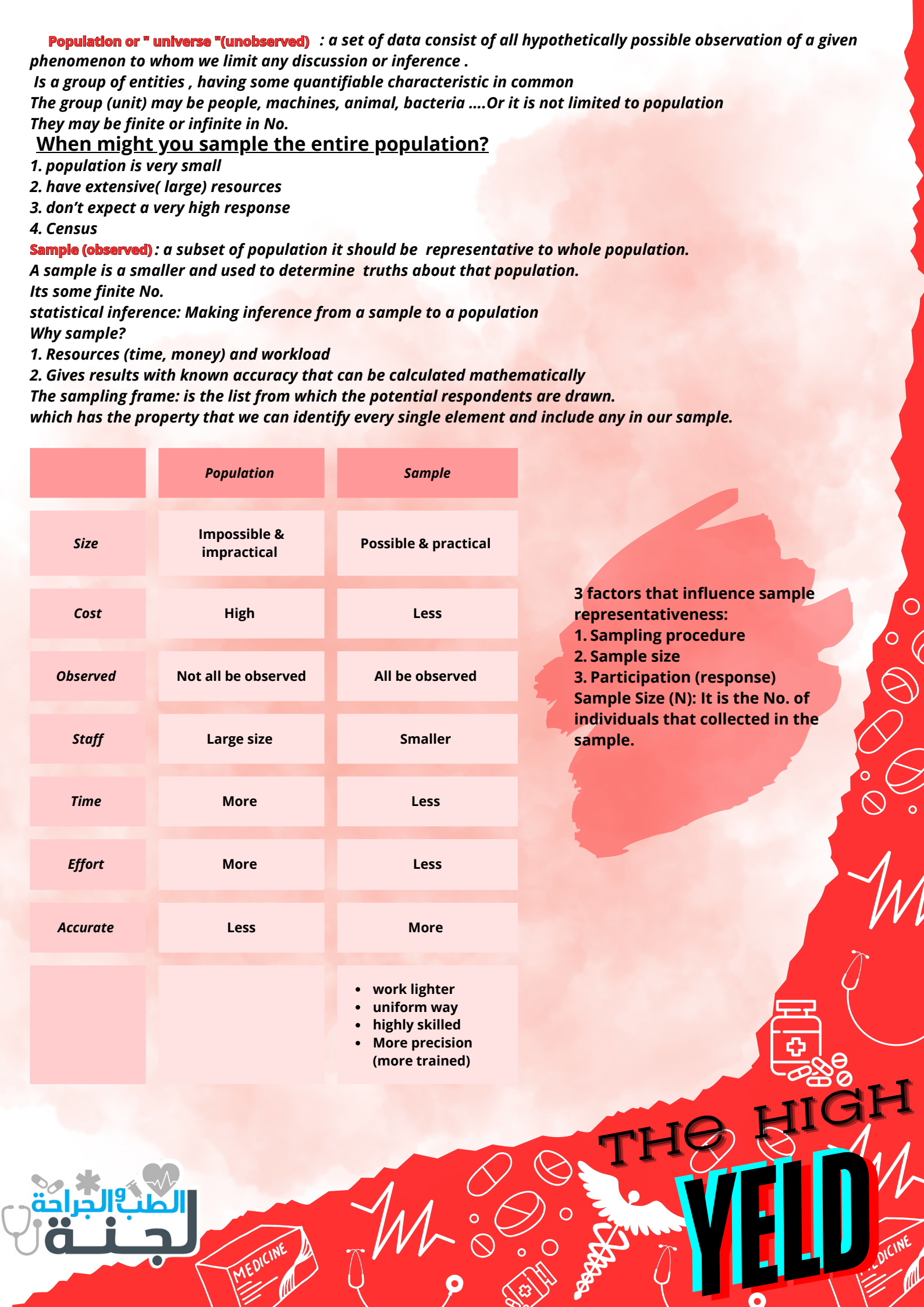
3 factors that influence sample representativeness:

1. Sampling procedure

2. Sample size

3. Participation (response)

Sample Size (N): It is the No. of individuals that collected in the sample.



Types of Sampling: 1. Non-probability 2. Probability .
Non-probability : one in which, the judgment of the experimenter, the methods in which the data are collected, or other factors could affect the results of the sample BIAS
Non-probability samples:
1) Convenience sample 2) Purposive sample 3) Quota
Probability: The chance of selection of each item of the population is known before the sample is picked NO BIAS

Probability (Random) Samples:

1. Simple random sample 2. Systematic random sample 3. Stratified random sample 4. Cluster sample 5. Multistage sample 6. Multiphase sample.

Random Sample: This is the most popular one, it is most commonly used in survey and research, it is the simplest and the best known way to avoid bias and to be representative to the population.
Every member of the population has the same probability of selection.
Every member in the population having independent and equal chance of appearing in the sample.
Applicable when population is small, homogeneous & readily available
Advantages:1.Estimates are easy to calculate. 2. Simple
Disadvantages:1. impracticable with sampling frame large.2. Need complete sampling frame. 3. Minority subgroups of interest in population may not be present in sample in sufficient numbers for study.

Systematic sampling : Sampling with system k=(population size/sample size).
All elements have the same probability of selection, It is not 'simple random sampling' because different subsets of the same size have different selection probabilitie
Advantages:1. Sample easy to select 2.Suitable sampling frame can be identified easily 3.Sample evenly spread over entire reference population
Disadvantages:1. Sample may be biased if hidden periodicity in population coincides with that of selection. 2. Difficult to assess precision of estimate from one survey.

Startified sampling : Where population enclose, a number of distinct categories, the frame can be organized into separate "strata"
Each stratum is sampled independently with random selection.
Same sampling fraction for all strata ensures proportionate representation.
Stratification allows representation of minority subgroups by varying sampling fractions.
Different sampling approaches can be used for each stratum.