

① People who took Vaccine	<u>Sampling frame</u> list numeric id's: Drivers licence. → telephone number
② Senior citizens in NJ	Drivers licence, telephone number SSN, resident sheet
③ Stat 400 Students.	UID, email.

Sampling

Sampling strategies critical to good statistics!
 {
 Key point

Randomness.

Want to minimize "bias" in the sample.

Sampling frame → a listing of individuals/objects in the population to be sampled.
 {
 might not always be available.

Simple Random Samples (SRS)

- (i) Every object in sampling frame has equal chance of being in a sample
- (ii) Every subset of size n has equal chance of being the chosen sample of size n .

Advantages

Stratified Sampling

- (i) Divide population into strata.
- (ii) Apply SRS to each strata.

The strata

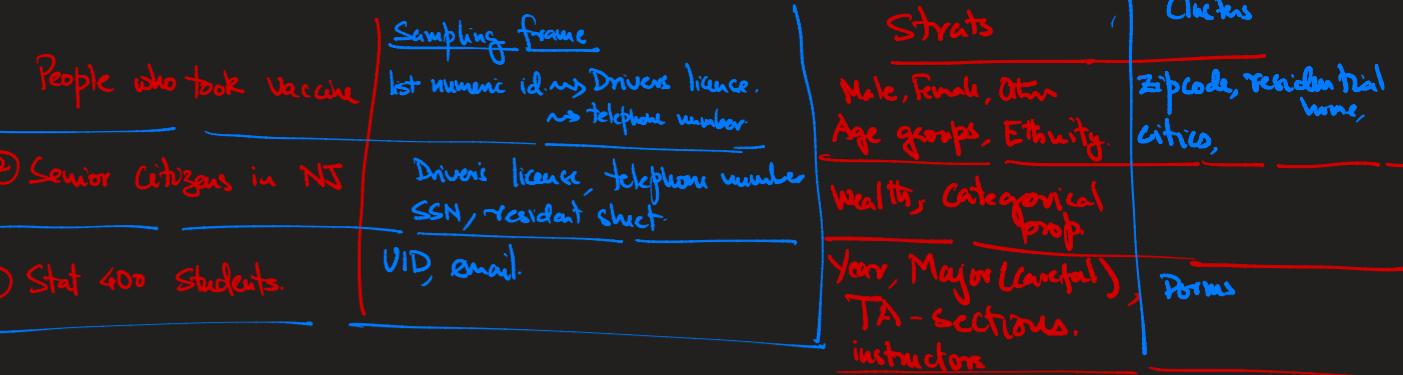
→ homogeneous w.r.t to a given attribute.

Sampling Techniques

Sampling frame → a listing of individuals/objects in the population to be sampled.

}

w might not always be available.

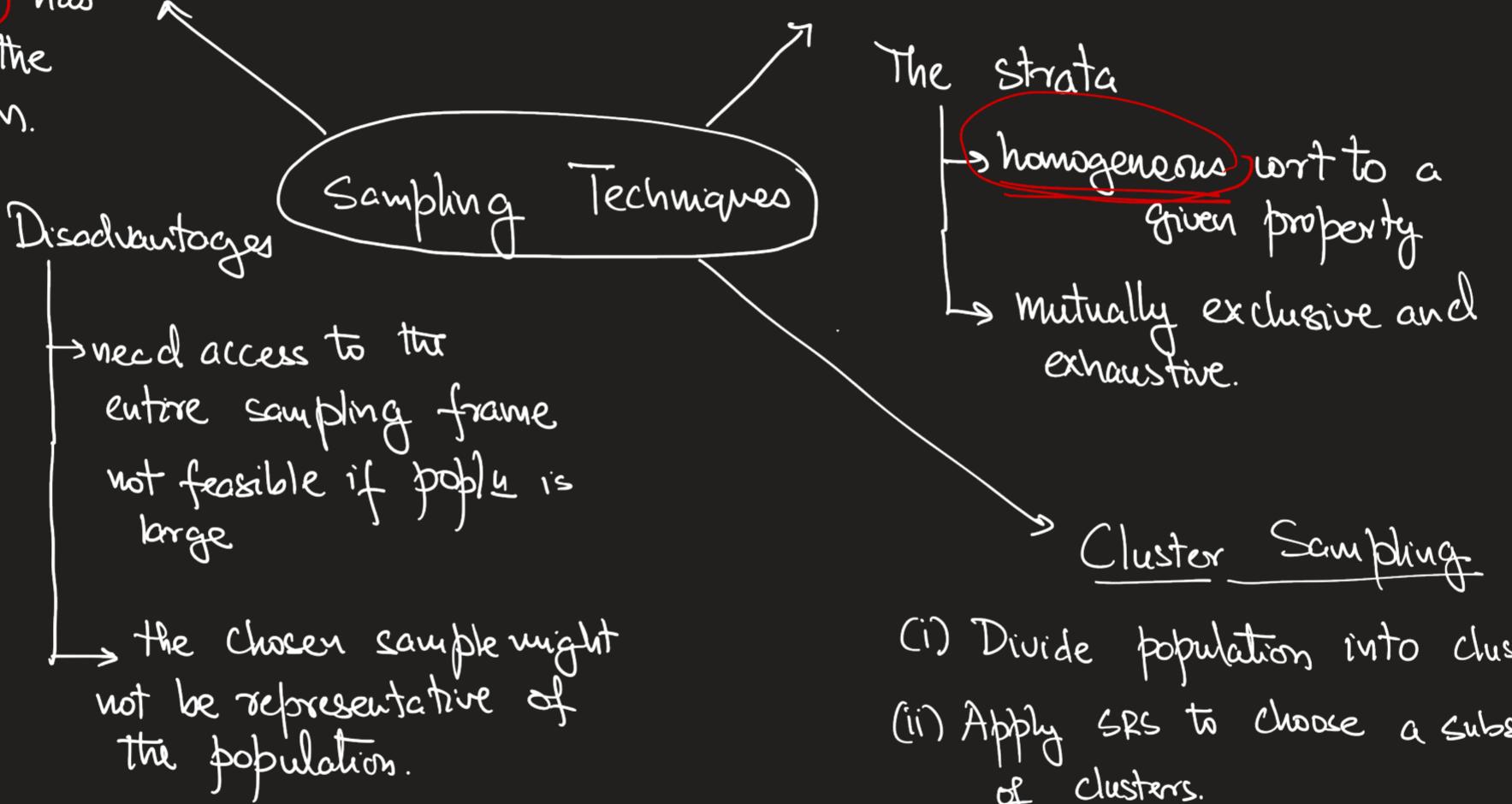


Simple Random Samples (SRS)

- (i) Every object in sampling frame has equal chance of being in a sample.
- (ii) Every subset of size n has equal chance of being the chosen sample of size n .

Advantages

- simple
- unbiased
- need minimum advance information about population
- works well when experiment involves repeated sampling and cost of sampling is low.



Stratified Sampling

- (i) Divide population into strata.
- (ii) Apply SRS to each strata.

Main Advantage

- Sample representative of population
- Improved precision.

Cluster Sampling

- (i) Divide population into clusters.
- (ii) Apply SRS to choose a subset of clusters.

Main Advantage:

- Cost effective → travelling costs
- research and listing costs.

Note: Each cluster must be representative of the population. Improved precision.

Stratified vs Cluster Sampling

stratified

cluster

Data:

5.13, 5.16, 5.20, 5.12, 5.10, 4.72, 5.26

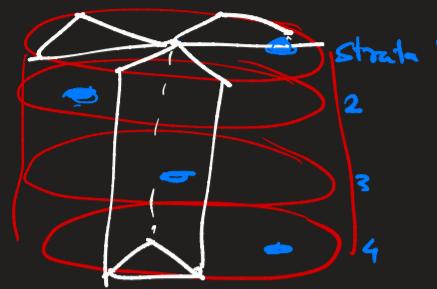
4.84, 5.02, 5.19, 5.04, 5.35, 5.21, 4.79,

5.44

Cost effective \rightarrow

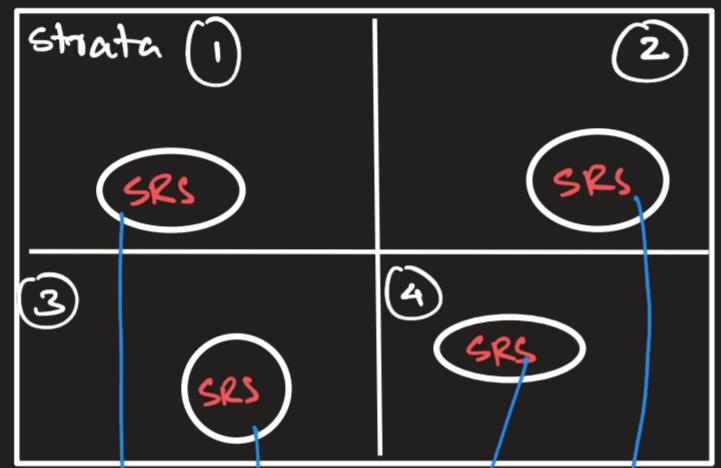
- travelling costs
- research and listing costs.

Note: Each cluster must be representative/improved precision of the population.

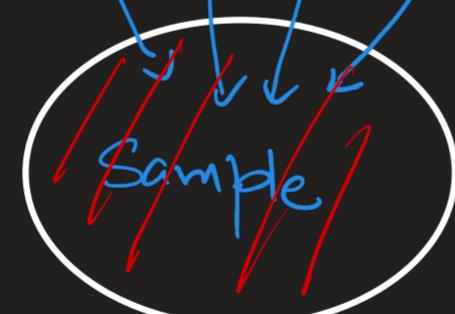


Stratified vs Cluster Sampling

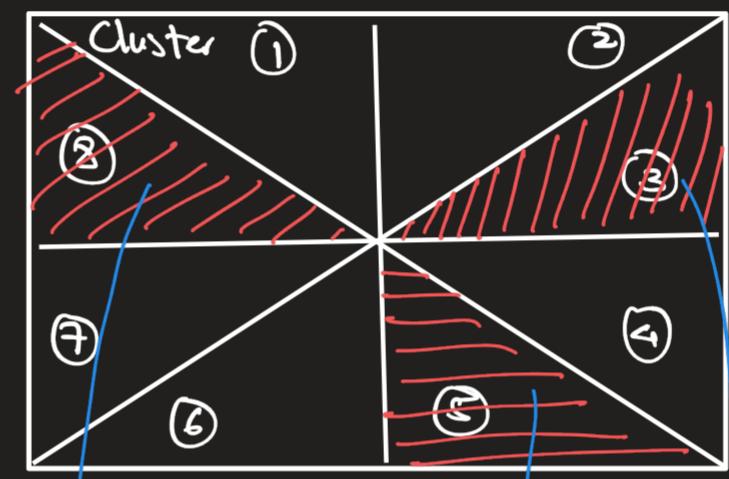
Stratified



We do SRS in each of the four strata.



Cluster.



We do SRS on the set
 $\{1, 2, 3, 4, 5, 6, 7, 8\}$

Suppose clusters 3, 5, 8 are selected.

