Sampling Methods

Problem: You're the **quality manager** at a company that just received a **truckload of 10,000 mangoes** from farmers living in East, West, North, South of country.
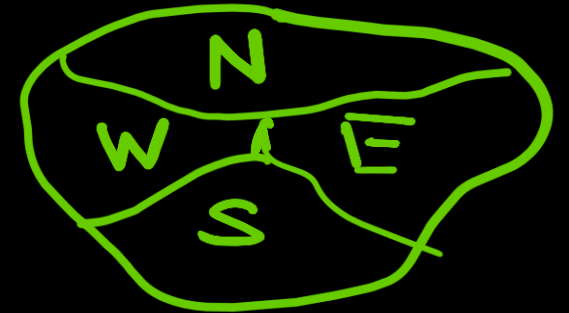Your boss says:
"Check if the shipment is good enough to send to the export market."
You can't cut open **all 10,000 mangoes** to test sweetness and ripeness.
You only have time to test **50 mangoes**.
How will you pick those 50 mangoes so that your decision represents the **whole truckload** ?
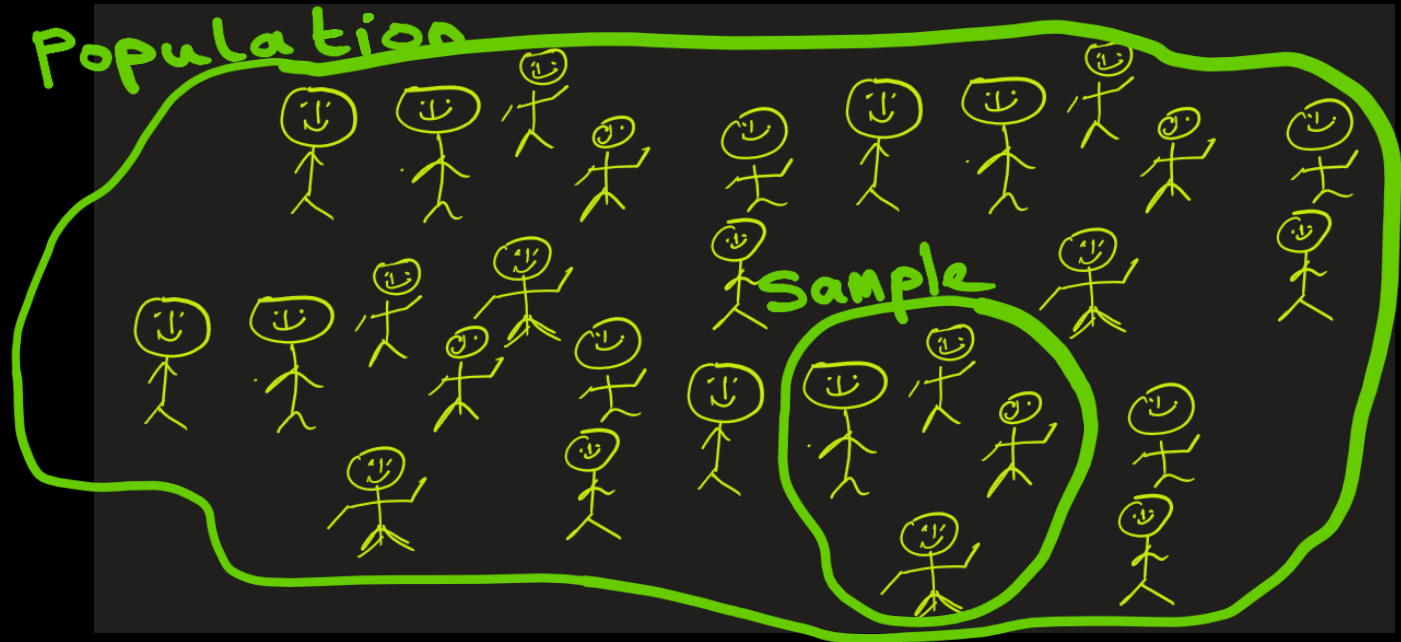
A) Simple Random Sampling
B) Systematic Sampling
C) Stratified Sampling
D) Cluster Sampling
E) Judgment Sampling
F) Convenience Sampling

What is Sampling and Why Do We Need It?

Sampling is the process of selecting a **small subset (sample)** from a **larger group (population)** to draw conclusions about the entire population. You could be drawing conclusion about mean, variance, etc.

- **Population** → The entire group you want to study e.g., all college students in Country

- **Sample** → A smaller group chosen from the population e.g., 500 students selected from different colleges

Why sampling ? Because collecting data from the entire population is usually **impractical or expensive**, we rely on **samples**. **For example,** suppose a **drug manufacturer** would like to research the adverse side effects of a drug on the country's population. In that case, it is almost impossible to conduct a research study that involves everyone. In this case, the manufacturer decides on a sample of people from each demographic and then researches side effects of drugs.

Fun Fact About Sampling Methods

In 1936, The Literary Digest magazine mailed out 10 million surveys to predict who would win the Presidential election — and they were wrong.

They predicted Alf Landon would win by a landslide.
But the actual winner was Franklin D. Roosevelt, who won by huge margin!

Why was the magazine wrong?
Because the magazine sent surveys only to people who had cars, telephones, or magazine subscriptions — mostly wealthier Citizens — and ignored poorer voters.

So even with millions of responses, the sample was biased.

What are Types **Sampling Methods**?

Sampling Methods

Probability Sampling:

Every individual in the population has a **known, non-zero chance** of being selected - It helps avoid bias and supports statistical inference.

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

Non-Probability Sampling:

Not everyone has a known or equal chance of being selected — can lead to bias but often used when probability sampling isn't possible.

- **Convenience Sampling**
- **Judgment Sampling**
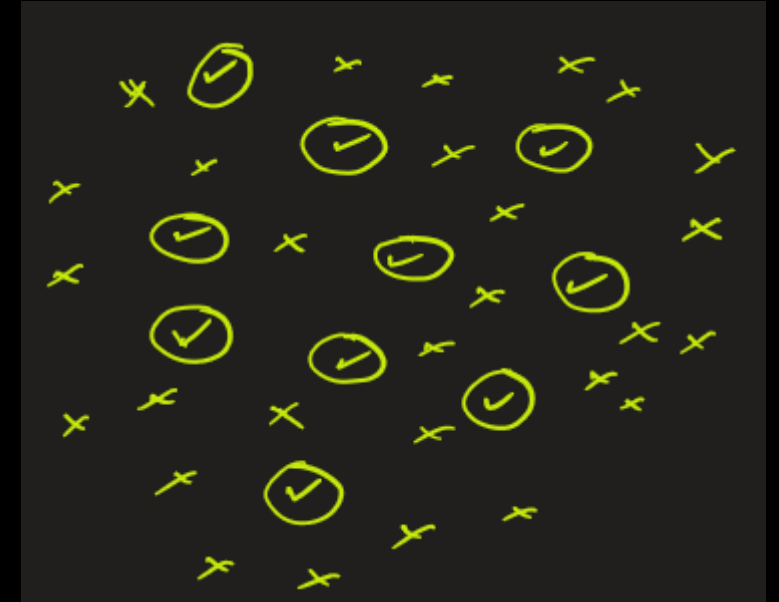- **Quota Sampling**
- **Snowball Sampling**

Sampling Method1 : Probability Sampling Types

Simple Random Sampling: Every member has an equal chance of being chosen
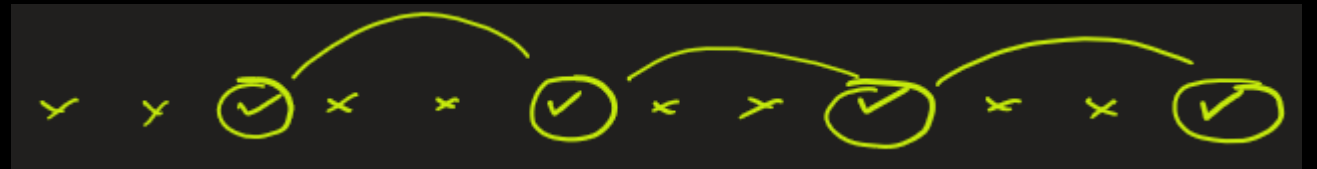
Example: Randomly pick 10 names from a list

Systematic Sampling: Select every *kth* item

Example: Pick every 3rd student from a list

Sampling Method1 : Probability Sampling Types

Stratified Sampling: Divide population into subgroups (strata) and take random samples from each

Example: You want a **sample of 40 students** from a population of 400 that represents all years **proportionally**.
Here the strategy is to
1) divide students by year (1st, 2nd, 3rd, 4th) and
2) then sample proportionally.

STEP1:
Divide by Year

Year	Number of Students
1st Year	120
2nd Year	100
3rd Year	80
4th Year	100
Total	**400**

STEP2:
Here 40 is 10% of total, so we take 10% of students from each year.

Year	Total Students	Sample Size
1st Year	120	$0.10 \times 120 =$ **12**
2nd Year	100	$0.10 \times 100 =$ **10**
3rd Year	80	$0.10 \times 80 =$ **8**
4th Year	100	$0.10 \times 100 =$ **10**
Total	400	**40**

Cluster Sampling: Divide population into clusters, then randomly select whole clusters. Each clusters should have similar characteristics within themselves.

Example: Suppose we want to conduct survey from students. Here divide student's population into **20 clusters based on colleges**. Randomly select 5 colleges (say C, F, K, N, S) and survey **all students** in those colleges.

College	Students
College A	450
College B	320
College C	**610**
College D	275
College E	390
College F	**520**
College G	205
College H	410
College I	360
College J	290
College K	**480**
College L	335
College M	255
College N	305
College O	225
College P	410
College Q	380
College R	300
College S	**420**
College T	455

Problem: You're the **quality manager** at a company that just received a **truckload of 10,000 mangoes** from farmers living in East, West, North, South of country.

Your boss says: "Check if the shipment is good enough to send to the export market."

You can't cut open **all 10,000 mangoes** to test sweetness and ripeness. You only have time to test **50 mangoes**.

How will you pick those 50 mangoes so that your decision represents the **whole truckload** ?

Ans: Since the mangoes came from **different regions** (E, W, N, S), their quality may differ by source.

So, you should d**ivide** the 10,000 mangoes into 4 strata (E, W, N, S).

Example:

    East: 4,000 mangoes
    West: 3,000 mangoes
    North: 3,000 mangoes
    South: 0 mangoes

Here 50 is 50/10,000 x 100 = 0.5%. So, **Sample proportionally** from each stratum, that is, 0.5% from each strata:

    East   → 4,000 × 0.5 = **20 mangoes**
    West → 3,000 × 0.5 = **15 mangoes**
    North → 3,000 × 0.5 = **15 mangoes**
    South → 0 × 0.5 = **0**   **mangoes**

Then **randomly select** those mangoes within each group. So, we use **Stratified Random Sampling**.

Sampling Method 2: Non-Probability Sampling Types

Method	Description	Example
Convenience Sampling	Choose who is easiest to access	Survey people in **your School** because they are easiest to access
Judgment Sampling	Researcher selects based on expertise	A doctor chooses patients she thinks represent her clinic. (Maybe patients are > 40 years of age. Choose samples of patients whose age > 40.)
Quota Sampling	Ensure specific groups are represented	Interview 50 males and 50 females, regardless of randomization
Snowball Sampling	Participants recruit others	Used for hard-to-reach groups like rare disease patients

STOP