Sampling Process and Data Recording

Assoc. Prof. Thanitha Darbanandana
Fishery Dynamics Research Laboratory
Department of Fishery Biology
Faculty of Fisheries
Kasetsart Univerity, THAILAND
E-mail: thanitha.t@ku.th
Line i.d.: fishery_dynamics

CONTENT

KU

KASETSART
UNIVERSITY

- **Probability Sampling**
  - **Equal** Probability Sampling
  - **Unequal** Probability Sampling
- **Non-probability** Sampling

How to collect samples

KASETSART UNIVERSITY

Equal Probability Sampling

- **Lucky Draw**
  - Use in small population
  - With or Without replacement
- **Random number table**
  - Often used in large population and sampling frame is available
  - With or without replacement
- **Randomize by computer**
  - Uses a computer to generate a random number between the numbers 1 and N and
  - Use the command to select the desired number of numbers.

KASETSART
UNIVERSITY

Simple Random Sampling (SRS)

Advantage	Disadvantage
1. The most easiest sampling design	1. Only suitable for Homogeneous population
2. Easy estimation, methods can be adjusted.	2. Not suitable for an indefinite population because it is difficult to estimate the suitable sample size
	3. It is the most expensive plan as it may require a large sample size. To control the 8 discrepancies within the desired scope

KU

KASETSART
UNIVERSITY

Systematic Sampling

- **Taking the first sample unit randomly**

- **After that, every *k* next sample unit will be selected until the desired *n* unit of sample is obtained**

- **For convenience, *N = kn* in order to have *k* of all possible sample units**

- **There are 2 methods:**

  - **Linear systematic sampling**

  - **Circular systematic sampling**

Linear systematic sampling

Give rank to all units in the population from *1, 2, 3, …..N*

Let *n* be the specified sample size

Calculates the sampling interval using the symbol *I*, which $I = \frac{N}{n}$

Select the Random Start (*R*) where *R* is between *1* and *I*, where *R* may be random from lucky draw, random number table or computer
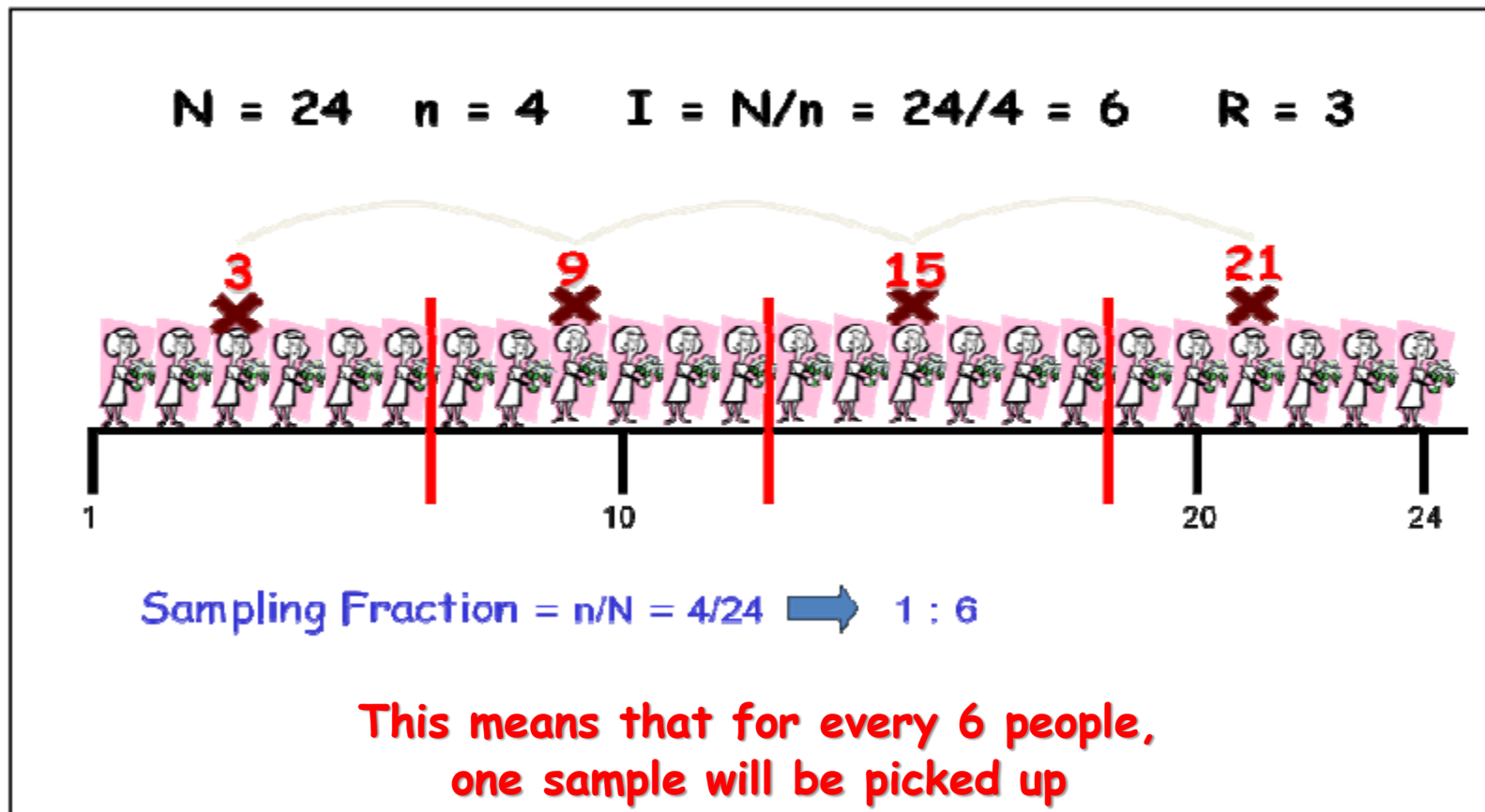
The unit selected as an example is a unit with ordinal numbers corresponding to the values *R, R + I, R + 2I, …, R + (n-1) I*

Linear systematic sampling

Example Want to select a sample of 4 people out of a total of 24 people, it can proceed as follows:

$N = 24$ $n = 4$ $I = N/n = 24/4 = 6$ $R = 3$

Sampling Fraction $= n/N = 4/24$ ⟹ $1 : 6$

This means that for every 6 people, one sample will be picked up

Circular systematic sampling

Let the numbers in order with every unit in the population are 1,2,3, ... N

Let n be the specified sample size

Calculates the sampling interval using the symbol I, which $I = \dfrac{N}{n}$

Select the Random Start (R) where R is between 1 and I, where R may be random from lucky draw, random number table or computer
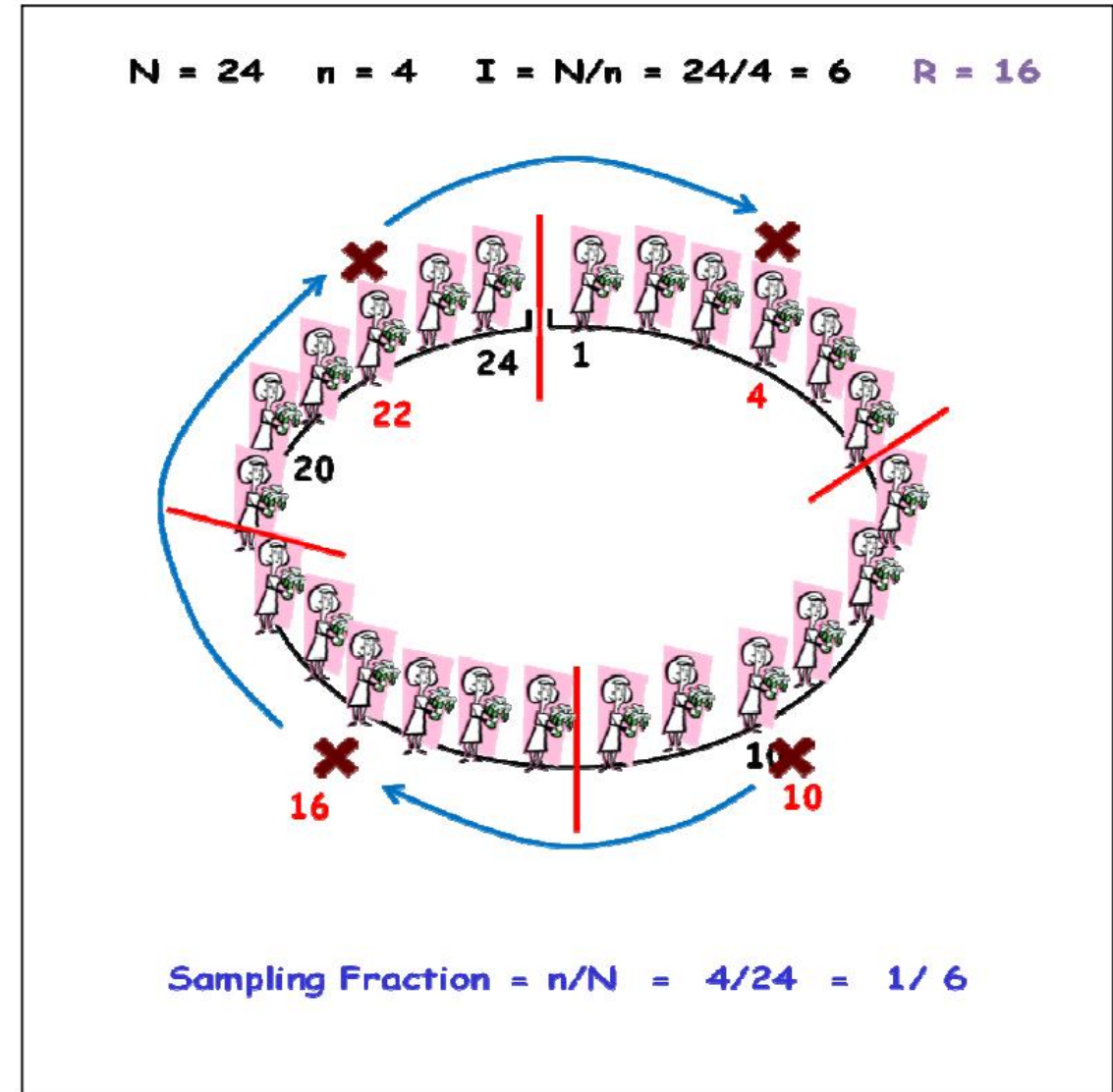
The unit selected as an example is a unit with ordinal numbers corresponding to the values $R, R + I, R + 2I, ..., R + (n-1) I$

Circular systematic sampling

In case that the value of **R + I** or **R + 2I** or ...., is greater than **N**, then **N** is subtracted. The result that matches any number of units is an example


$N = 24$ $n = 4$ $I = N/n = 24/4 = 6$ $R = 16$

Sampling Fraction $= n/N = 4/24 = 1/6$

Systematic Sampling

Advantage	Disadvantage
1. Easy, time-consuming and low-cost method	1. Get a bias estimator in the case that N is not equal to kn
2. Will be very effective when a population has a good order of sample units. Sampling can be selected more thoroughly than SRS	2. Cannot calculate unbiased estimators of the variance of the approximator from a single random sample (after every k units have been selected) then it is necessary to calculate the estimates from SRS
	3. If the data is not sorted well, the estimates are less effective than SRS

Stratified Sampling

population

strata

sampling

Total samples

Cluster Sampling

population
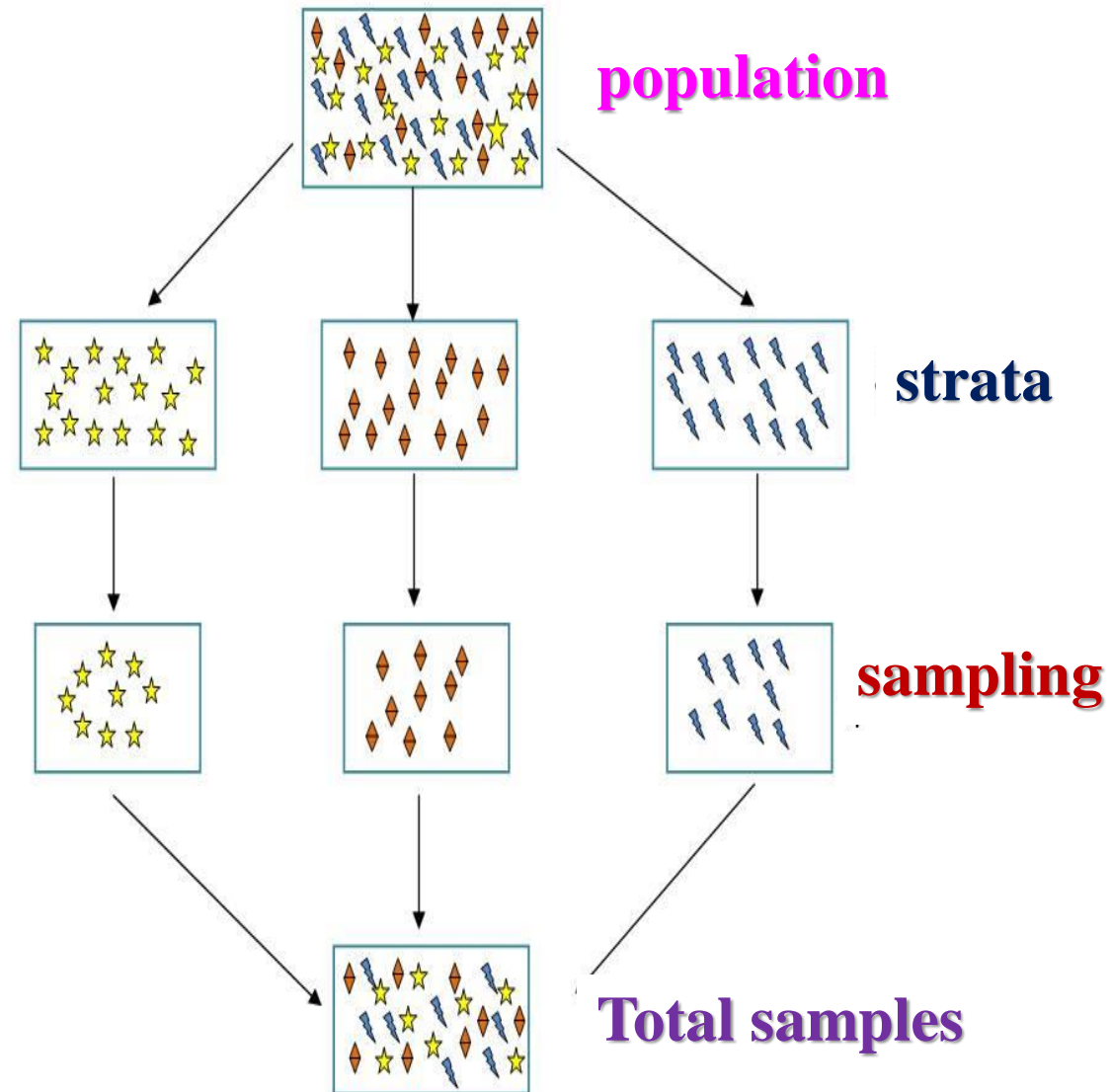
Clustered

Sampling
3 clusters
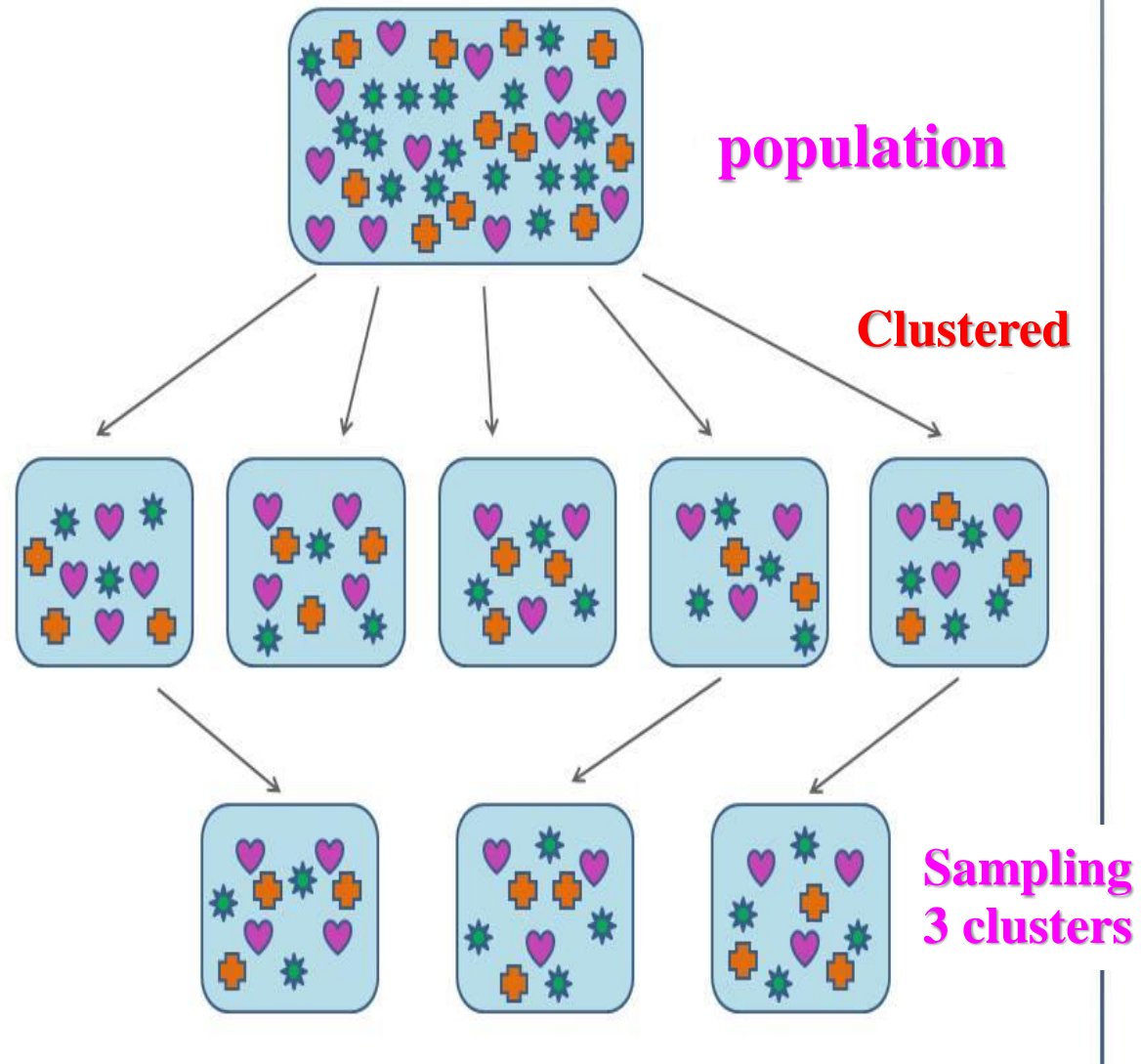
Unequal Probability Sampling

Example

Village A
N = 100 houses
10% sampling
N= 10 houses

Equal Probability Sampling

Village B
N = 1,000 houses
10% sampling
N= 100 houses

Un-equal Probability Sampling

- Select n of the total samples with unequal probability

- Number of samples depend on population size

- Probability proportional to size (PPS)

- Many methods are available

- However, PPS is suitable for replacement sampling

- While fisheries data almost use without replacement sampling

KASETSART
UNIVERSITY

Non-probability sampling

Non-parametric Data Analyses

Accessible sampling

Accidental sampling

Purposive sampling

Volunteer sampling

Quota sampling

Snowball sampling

Accessible sampling

- Convenient sample unit selection

- Sampling from the fish tank

- There is no opportunity to choose the ones below

- Or have already selected the size

Accidental sampling

- Accidental selection of sample units
- Don't fore plan which sample unit to choose
- Unexpected sample selection or easy to operate
- Select fish samples by swing
- Select passers-by for interviews or data collection

KU

KASETSART
UNIVERSITY

Purposive Sampling

- most common in research in fisheries biology / management
- Use one's judgment and experience how to choose which sample units to represent the population

- Often a specific sample unit that attractive to the researcher, or
- The sample unit has a good familiarity and happy in providing information
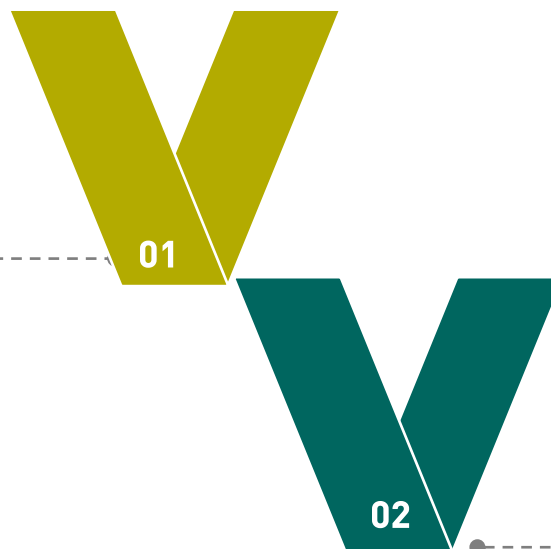
- convenient by studying a subset of the population
- Suitable for small sampling and
- Has limited operating time or
- Prelim. Study
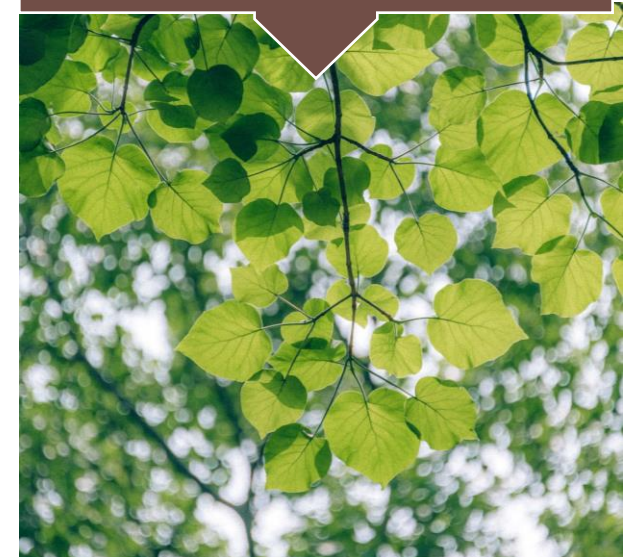
Volunteer Sampling

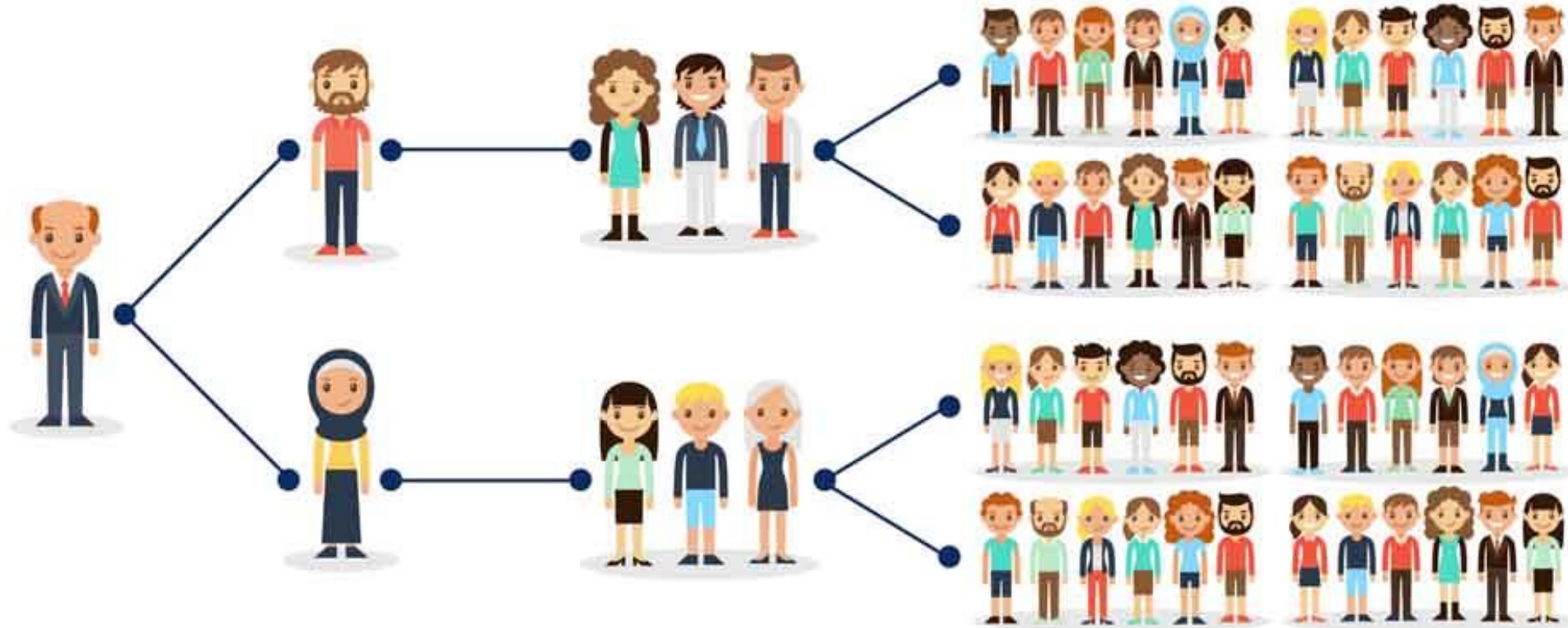- **Selection of representative sample units**

01

02

- **Impossible to select the originally sample unit due to the difficulty in selecting the sample unit and / or**
- **The occurrence of unforeseen circumstances**
- **Need to find an agent to compensate for the original samples**

KU

Quota Sampling

- Widely used in polling, market research.

- Pre-setting quotas for all strata.

- The sample will be collected gradually until meet the complete no. of samples

- The strata usually be geographic area, gender, age, income, etc.

- The most popular non-probability sampling because it seems to have a good principle.

SNOWBALL SAMPLING

10 MINUTES BREAK