CLASS – XI	Subject-Economics	Chapter Name/Number- Collection of Data, chap-2

General Instructions:

1. These notes are from Chapter 2, Collection of Data

2. The notes need to be copied in the notebook

3. Reference book: Statistics for Economics by NM Shah and NCERT

LIMITATIONS OF SECONDARY DATA

- These may not have been collected by proper statistical techniques
- These may not be suitable for concerned enquiry
- These may not relate to the present times
- These may not be collected according to a reasonable standard of accuracy
- These may be biased and hence not reliable

PRECAUTIONS IN THE USE OF SECONDARY DATA

1. **Competency of the collecting agency:** The data should be used only if these are collected by a reliable, impartial and competent agency. So, the ability of the collecting organisation should be checked.

2. **Suitability of objective:** Data should be used only if the objective of the study as undertaken earlier matches with the objective and scope of current statistical enquiry.

3. **Method of collection:** The method of data collection used by the original investigator should match the nature of present statistical enquiry.

4. **Condition of collection:** The investigator should study conditions of investigation.

5. **Accuracy:** Accuracy of data should also be kept in mind. If the available data do not possess required level of accuracy, then such data should not be used.

CENSUS AND SAMPLE SURVEYS

Census or Complete Enumeration

A survey, which includes every element of the population, is known as Census or the Method of Complete Enumeration. If certain agencies are interested in studying the total population in India, they have to obtain information from all the households in rural and urban India.

Population and Sample Population or the Universe in statistics means totality of the items under study. Thus, the Population or the Universe is a group to which the results of the study are intended to apply.

A **population** is always all the individuals/items who possess certain characteristics (or a set of characteristics), according to the purpose of the survey.
The first task in selecting a sample is to identify the population.

Merits

- Accurate and reliable
- Extensive study of diverse items
- Study of different characteristics
- Indirect investigation

Demerits

- Expensive
- Large man power required
- Time consuming
- Not universally operative

Once the population is identified, the researcher selects a method of studying it. If the researcher finds that survey of the whole population is not possible, then he/ she may decide to select a Representative Sample.

SAMPLE METHOD

A sample refers to a group or section of the population from which information is to be obtained. A good sample (representative sample) is generally smaller than the population and is capable of providing reasonably accurate information about the population at a much lower cost and shorter time.

Merits –

- It is an economically viable method as it is less costly, saves time and requires less manpower to collect data.
- The result of the census method may be checked with the help of the sampling method.
- In cases where the population size is too large, the sampling method is easy and more practical.
- We can use it to make estimations about population characteristics without even surveying all units of the population.

Demerits –

- If the sampling is not properly conducted, it might lead to erroneous and unrepresentative results.
- Sampling normally generates an error due to leaving out of units from the population. If a crucial unit is left out of the sample, the resulting error will be large.
- If skilled personnel are not available to interpret the data, the results drawn will be unreliable.

Difference between Census method and sample method of data collection

Parameter	Census	Sample Survey
Definition	A statistical method that studies all the units or members of a population.	A statistical method that studies only a representative group of the population, and not all its members.
Calculation	Total/Complete	Partial
Time involved	It is a time-consuming process.	It is a quicker process.
Cost involved	It is a costly method.	It is a relatively inexpensive method.
Accuracy	The results obtained are accurate as each member is surveyed. So, there is a negligible error.	The results are relatively inaccurate due to leaving out of items from the sample. The resulting error is large.
Reliability	Highly reliable	Low reliability
Error	Not present	The smaller the sample size, the larger the error.
Relevance	This method is suited for	This method is suited for

	heterogeneous data.	homogeneous data.

Methods of Sampling

Random Sampling

Random sampling is a sampling method in which samples are gathered in a process that gives equal chance to all individuals in the population of being selected.

Methods of random sampling
Lottery Method: Under lottery method, chits or paper slips are made for each item in the population. Then the chits are shuffled in the container and without any biasness, some slips are drawn out. The chits then become the samples of the universe.

Tables of random number: Tippet has prepared a table consisting of 10400 numbers of four digits. In this method, all the items are arranged in an order and using Tippet's table, required numbers of items are selected.

Merits
- Free from bias
- Equality
- Representation of population
- Simple

Demerits
- Ignores important items
- No proportionate representation

Stratified Random sampling

Stratified random sampling is a method of sampling that involves the division of a population into smaller sub-groups known as strata. In stratified random sampling or stratification, the strata are formed based on members' shared attributes or characteristics such as income or educational attainment.
Stratified random sampling is also called proportional random sampling or quota random sampling.

A stratified random sampling involves dividing the entire population into homogeneous groups called strata (plural for *stratum).* Random samples are then selected from each stratum.

Example: If there are 100 students in class X, out of which 70 have opted for Hindi and 30 for Sanskrit, then 100 students are divided into two strata. If 10 students have

to be selected, the 7 will be selected from 70 hindi students and remaining 3 from Sanskrit.

Merits

- It captures key population characteristics in the sample.
- This method of sampling produces characteristics in the sample that are proportional to the overall population.
- Stratified random sampling works well for populations with a variety of attributes
- Accuracy is achieved

Demerits

- This method of research cannot be used in every study.
- It has limited scope.
- It is an expensive method

Cluster Sampling

With cluster sampling, the researcher divides the population into separate groups, called clusters. Then, a simple random sample of clusters is selected from the population. The researcher conducts his analysis on data from the sampled clusters.

The most common cluster used in research is a geographical cluster. For example, a researcher wants to survey academic performance of high school students in India.

1. He can divide the entire population (population of India) into different clusters (states).
2. Then the researcher selects a number of clusters depending on his research through simple
3. Then, from the selected clusters (randomly selected states) the researcher can either include all the high school students as subjects or he can select a number of subjects from each cluster through simple random sampling

Merits

- It is the most time-efficient and cost-efficient probability design for large geographical areas
- This method is easy to be used from practicality viewpoint
- Larger sample size can be used due to increased level of accessibility of perspective sample group members

Demerits

- Requires group-level information to be known

- Commonly has higher sampling error than other sampling techniques
- Cluster sampling may fail to reflect the diversity in the sampling

Systematic sampling

Method in which items of the universe are arranged numerically, alphabetically and geographically and every n^{th} item of the numbered items is selected as a sample item is called systematic sampling. This method is considered to be a shortcut method of random sampling.

For example, if 10 items are to be selected from 100 items, then every 10th item from the selected universe would be selected randomly.

Merits

- It is a simple method
- This method is not biased
- It is less time consuming
- Satisfactory conclusions can be derived

Demerits

- This method is sometimes unfair
- Unsuitable if the items in the universe are homogeneous
- Unsuitable if size of universe is large

Non random sampling

Deliberate sampling

Method in which the investigator himself choses the sample from the universe, which in his opinion are the best representative of the population is deliberate sampling. This method is used when some items are of more importance than the others and need to be selected as a sample item.

For example if an investigation is to find out profit earned by mobile companies in india, then the inclusion Samsung and apple becomes essential.

Merits

- Flexible
- Simple
- Facilitates purpose of study

Demerits

- Biased
- Not reliable
- Inaccurate

Quota sampling

Method in which population is divided into different groups or different classes according to different characteristics of the population is called quota sampling. The quota of the units to be observed by the investigator is fixed in advance, according to specified characteristics like gender, age, annual income etc. It is a mixture of stratified and deliberate sampling.
For example, in a survey of newspaper readers, the interviewers may be told to interview 100 people living in a particular city out of which 25 percent people interviewed should be housewife, 45 percent to be employee and 30 percent should be children. Within the quotas, the interviewer has full freedom to select the people to the interviewed.

Merits

- Reliable
- Economical

Demerits

- Personal bias
- Impossible to detect sample error

Convenience sampling

Method in which sampling is done by the investigator in a manner that is according to his convenience is called convenience sampling. It is the simplest method as it serves the purpose of convenience of the investigator.

Merits

- Simple
- Inexpensive

Demerits

- Unreliable
- Unscientific
- Personal biased

ESSENTIAL OF A GOOD SAMPLE

- **Representative:** a good sample is one which represents the characteristics of all the items in the universe. This is possible only when each and every item of the universe has an equal chance of being selected.

- **Homogeneity:** the items that are selected as samples should be homogeneous in nature however these should not be contradictory to each other.

- **Independent:** selection of one item of the universe in the sample should not be depending on selection of the some other item in the sample. All items in the population should be independent of each other.

- **Sufficiency:** the number of items in the sample should be adequate enough to cover all characteristics of the universe.