



Micro Economics Notes PDF

On

Collection, Organisation and Presentation of Data

(Class - 11)

DATA IS A COLLECTION OF FACTS AND MEASUREMENT

Data is a tool which helps in reaching a sound conclusion by providing information therefore. For statistical investigation, collection of data is the first and foremost.

SOURCES OF DATA

There are two sources of data:

- **Primary Source of Data:** It implies collection of data from its source of origin.
- **Secondary Source of Data:** It implies collection of data from some agency or institution which already happens to have collected the data through statistical survey.

Primary Data– Data originally collected in the process of investigation are known as primary data. This is the original form of data which is collected for the first time. It is collected directly from its source of origin.

Methods of collecting primary data:

There are three basic ways of collecting data :

- (i) Personal interview OR Direct Personal Investigation
- (ii) Mailing (questionnaire surveys)
- (iii) Telephone interviews
- (iv) Indirect verbal investigation
- (v) Information from local sources
- (vi) Enumerator method

Secondary data: It refers to collection of data by some agency, which already collected the data and processed it. The data thus collected is called secondary data.

Collection of Secondary Data:

There are two main sources of secondary data:

- Published sources
- Unpublished sources

1. Published Sources:

Some of the published source of secondary data is:

- Government publication
- Semi-government publication
- Reports of committees and commissions
- Publications of trade associations
- Publication of research institutions
- Journals and papers
- Publication of research scholars
- International publication

2. Unpublished Sources:

These data are collected by the government organisations and others, generally for their self use or office record:

- In order to assess the reliability, suitability and adequacy of the data, the following points must be kept in mind
- Ability of the collecting organisation
- Objective and scope

- Method of collection
- Time and condition of organisation
- Definition of the unit
- Accuracy

3. Other source: web-site.

Important points to be kept in mind while drafting the questionnaire:

- Introduction and purpose of investigation
- Reasonable number questions.
- Questions should be small & clear.
- Questions should be arranged logically.
- Instructions should be clear.
- Proper space for answer.
- Questions should be relevant to the investigation.
- Personal questions should be avoided.
- Avoid questions of calculations.
- Cross Verification.
- The question should not be ambiguous.
- The question should not use double negatives like “don’t you”.
- The question should not indicate alternative answers.

Methods of Sampling:

(i) Random Sampling Random sampling is that method of sampling in which each and every item of the universe has an equal chance of being selected in the sample. Random sampling may be done in any of the following ways:

- Lottery method
- Tables of random number

(ii) Purposive or Deliberate Sampling It is that method in which the investigator himself makes the choice of the samples items which in his opinion are the best representative of the universe.

(iii) Stratified or Mixed Sampling According to this method of sampling, the population is divided into different strata having different characteristics and some of the items are selected from each strata, so the entire population gets represented.

(iv) Systematic Sampling According to this methods, units of the population are numerically, geographically and alphabetically arranged. Every nth item of the numbered is selected as a sample item.

(v) Quota Sampling In this method, the population is divided into different groups or classes according to different characteristics of the population.

(vi) Convenience Sampling In this method, sampling is done by the investigator in such a manner that suits his convenience.

Reliability of Sampling Data:

It depends mainly on the following factors

- Size of the sample
- Method of sampling
- Bias of correspondents and enumerators
- Training of enumerators

Important agencies at the national level which collect, process and tabulate the statistical data. NSSO (National Sample Survey Organisation), RGI (Registrar General of India), DGCIS (Directorate General of Commercial Intelligence and Statistics) and Labour Bureaus.

Census Method: Census method is that method in which data are collected covering every item of the universe or population relating to the problem under investigation.

Sample Method: It is that method in which data is collected about the sample on a group of items taken from the populations for examination and conclusions are drawn on their basis.

Population or universe:

In Statistics, population or universe simply refers to an aggregate of items to be studied for an investigation.

Sample: A group of items taken from the population for investigation and representative of all the items.

Sampling Errors: Sampling error is the difference between the result of studying a sample and the result of the census of the whole population.

- Biased errors
- Unbiased errors

Non-Sampling Error: Can occur in any type of survey whether it be a census or sample survey.

- Error in data acquisition
- Non Response error
- Sampling Bias

ORGANIZATION OF DATA

It refers to the systematic arrangement of collected figures (raw data), so that the data becomes easy to understand and more convenient for further statistical treatment.

Classification: It is the process of arranging data into sequences and groups according to their common characteristics of separating them into different but related parts.

Objectives of Classification:

- Simplification and Briefness
- Utility
- Distinctiveness
- Comparability
- Scientific arrangement
- Attractive and effective

Characteristic of a Good Classification:

- Comprehensiveness
- Clarity
- Homogeneity
- Suitability
- Stability
- Elastic

BASIS OF CLASSIFICATION

Raw data can be classified as:

- 1. Chronological classification:** In such a classification data are classified either in ascending or in descending order with reference to time such as years, quarters, month's weeks etc.
- 2. Geographical/Spatial classification:** The data are classified with reference to geographical location/place such as countries, states, cities, districts, block etc.
- 3. Qualitative classification:** Data are classified with reference to descriptive characteristics like sex, caste, religion literacy etc.
- 4. Quantitative classification:** Data are classified on the basis of some measurable characteristics such as height, age, weight, income, marks of students.
- 5. conditional classification:** When data are classified with respect to condition, the type of classification is called conditional classification.

A mass of data in its original form is called **raw data**. It is an unorganized mass of various items.

CONCEPT OF VARIABLE

A characteristic or a phenomenon which is capable of being measured and changes its value overtime is called a variable.

The variable may be either:

- **Discrete Variable:** These are those variables that increase in jumps or in complete numbers.
- **Continuous Variable:** Variable that assumes a range of values or increase not in jumps but continuously or in fractions are called continuous variables.

A frequency distribution is a comprehensive way to classify raw data of a quantitative variable. It shows how different values of a variable are distributed in different classes along with their corresponding class frequencies.

The class mid-point or class mark is the middle value of a class. It lies halfway between the lower class limit and the upper class limit of a class and can be ascertained in the following manner.

Class mid-point = upper class limit + lower class limit / 2.

Class frequency: It means the number of values in a particular class.

Class width:- It is the difference between the upper class limit and lower class limit

Class width = upper class Limit – Lower class Limit

Class Limits:- There are two ends of a class. The lowest value is called lower class limit and highest value is called upper class limit.

Types of series:

1. Individual series
2. Frequency series
 - Discrete series Or frequency array
 - Frequency distribution or continuous series

Individual series: These are those series in which the items are listed singly. For example:

Sr. No. of workers	Daily wages(in Rs.)
1	25
2	50
3	35
4	40
5	20
6	45

A discrete series or frequency array is that series in which data are prescribed in a way that exact measurements of items are clearly shown. The example in the following table illustrates a frequency array.

Frequency array of the size of household:

Size of the household	Number of household (Frequency)
1	5
2	15
3	25
4	35
5	10
6	5

A continuous series: It is that series in which items cannot be exactly measured. The items assume a range of values and are placed within the range of limits. In other words, data are classified into different classes with a range, the range is called class-intervals.

Frequency distribution or continuous series:

Marks	Frequency
10-20	4
20-30	5
30-40	8

40-50	5
50-60	4
60-70	3

THE PRESENTATION OF DATA

It means exhibition of data in such a clear and attractive manner that these can be easily understood and analysed.

Forms of Presentation of data:

- I. Textual/Descriptive Presentation
- II. Tabular Presentation
- III. Diagrammatic Presentation
- IV. Graphical Presentation

I. Textual/Descriptive Presentation of Data:

In this, data is presented in the form of text. This is suitable when quantity of data is not too large.

II. Tabulation:

It is the process of presenting data in the form of a table.

Parts or components of Table:

1. Table Number
2. Title
3. Caption Or Column Headings
4. Stubs Or Row Headings
5. Body of the Table
6. Unit of Measurement
7. Source
8. Head Note
9. Footnote

Features of a good table:

- (a) Compatible with the objective
- (b) Helpful in comparison
- (c) Ideal Size
- (d) Stubs
- (e) Headings
- (f) Percentage and ratio
- (g) Sources of Data
- (h) Simplicity

Kinds of Table:

1. According to Purpose
2. According to originality

3. According to construction

Classification of tabular presentation of data:

1. **Qualitative Classification:-** When classification is done according to attributes such as social status, nationality, etc. It is called qualitative classification.
2. **Quantitative Classification:-** In this, the data are classified on the basis of characteristics which are quantitative in nature. e.g., age, height, income, etc.
3. **Temporal classification:-** In this, time becomes the classifying variable and data are categorised according to time. Time may be in hours, weeks, years, etc.
4. **Spatial classification:-** When classification is done on the basis of place, it is called spatial classification. The place may be a village, town, state, country, etc.

III. Diagrammatic Presentation:

When data is presented in a simple and attractive manner in the form of diagrams is called diagrammatic presentation of data.

Types of Diagrammatic Presentation:

1. Geometric Form:

- a. Pie Diagram
- b. Bar Diagram
 - Simple
 - Multiple
 - Sub Divided
 - Percentage

2. Frequency Diagram:

- a. Histogram
- b. Frequency Polygon
- c. Frequency Curve
- d. Ogive curve

IV. Arithmetic Line Graph or Time series graph:

- **Bar diagram:-** Bar diagrams are those diagrams in which data are presented in the form of bars or rectangles.
- **Simple bar diagram:-** They are those diagrams which are based on a single set of numerical data. Different items are represented by different bars.
- **Multiple bar diagram:-** They are those diagrams which show two or more sets of data simultaneously. This type of diagram is, generally, used to make comparisons between two sets of series.
- **Sub divided bar diagram:-** These are those diagrams which present simultaneously, total values and parts there in a set of data.
- **Percentage bar diagram:-** They are those diagrams which show simultaneously different parts of the values of a set of data in terms of percentage.
- **Deviation bar diagram:-** These are used to compare the net deviation of related variables with respect to time and location. Bars which represent positive deviation and which represent negative deviation are drawn above and below the base line respectively.
- **Pie or circular diagram** is a circle divided into various segments showing the percent values of a series.
- **Histogram:** It is a graphical presentation of a frequency distribution of a continuous series. It can never be drawn for a discrete series.

- **Frequency polygon:** It is drawn by joining the mid points of the tops of rectangles in a histogram. It is constructed with the help of discrete as well as continuous series.
- **Frequency curve:** It is obtained by joining the points of a frequency polygon through a free hand smooth curve not by straight lines.
- **Cumulative frequency curve or Ogive curve:** Ogive curve is the curve which is constructed by plotting cumulative frequency data on graph paper in the form of a smooth curve.
- **Arithmetic line Graphs or Time Series Graphs:** In this graph, time(hour, day, date, week, month, and year) is plotted along X-axis and the corresponding value of variable along Y-axis.

[Follow on Facebook](#)[Follow on Instagram](#)[Join us on Telegram](#)A purple rectangular box with the word "Commerce" written in a white, cursive font. There are two small black dots on each of the four corners of the box.**CLASSES**