CH-13

Statistical Discription of Data

Statistics is a method by which we collect data, and then analysis the data by using different Statistical Techniques

Conclusion :- Statistics = Collection of DATA
+
Analysis of DATA

Origin of Word "STATISTICS"

Latin Word → STATUS
Italian Word → STATISTA
German Word → STATISTIK
French Word → STATISTIQUE

"Statistics was used in ancient period, Medival Period, as well as Modern Period."

Statistics is an Universal Application

Definations :-

In Singular Sense → Stastics is a scientific Method by which we collect, analysis, and present data, it is also known as "scientific Counting" (or) "Science of Averages."

In Plural Sense → Data which is Qualitative as well as Quantitative which is collected with a view of having statistical analysis.

Statistics is concerned with both Qualitative & Quantitative Information. But statistics deals with Quantitative Info.

APPLICATION OF STATISTICS :-

1. Economics
2. Business Management
3. Commerce & Industry

"ECONOMETRICS is a Branch of Economics in which Statistics plays a very important Role."

LIMITATIONS OF STATISTICS :-

1. Statistics deals with aggregate. It does not deal with individual value except the fact that is part of an AGGREGATE.

2. Statistics concern with Quantitative DATA, it does not deals with Qualitative Data.
   If you want to deal with Qualitative Data then first you have to convert Qualitative data into "Quantitative Data" by giving it Numerical Value.

3. Statistical result is built upon "Random Sampling". So the rule of random sampling must be followed otherwise the result will be ERRONEOUS.

COLLECTION OF DATA :-

We may define Data as a "Quantitative Info." about some particular characteristics under consideration.

DIVISION OF DATA :-

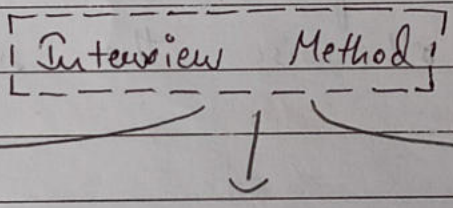
Internally Division	Externally Division
○ Qualitative Data	○ Primary Data
○ Quantitative Data	○ Secondary Data

1. Collection of Primary Data

(i) Interview Method
(ii) Indirect ?

Collection of Primary Data

(i) Interview Method – Personal, Indirect, Telephone
(ii) Mailed questionnaire Method
(iii) Observation Method
(iv) Questionnaires filled and sent by enumerators (OR)
SURVEY METHODS



Personal

In Natural Calamity we collect the necessary data much more quickly & accurately by this method.

Indirect

In case of Practicle problem in reaching respondents directly like in the case of rode accident, we collect data from other persons not directly with injured

Telephone

It is a quick and cheap way of collecting the Primary Data. It covers large area. "WIDE Coverage"

Observation Method → Using instrument like Height, BP, Sugar

Mailed Questionnaire Method → In this method group of Questions be asked and answered by respondent & Forward it.

"Widest Coverage Method"

"Amount of Non-Response is Maximum in this method"

"Least Accuracy"

Sources of Secondary Data

(i) International Sources like :- Reports of WHO, World Bank, IMF

(ii) Government Sources like :- CSO

(iii) Private & Quasi Government Sources like :- ICAR, ISI, NCERT

(iv) Unpublished ~~Reports~~ Sources like :- Reasearchers, Research Institute

Scrunity of Data

It means varification of accuracy and consistency of Data

When two or more related series are given then to check the data we use Internal consistency.

∴ Density = Population / Area holds.

Presentation of Data

- Classification of Data :-
  - (i) Chronological or Temporal or Time Series Data
  - (ii) Geographical or Spatial Series Data
  - (iii) Qualitative or Ordinal Data (attribute)
  - (iv) Quantitative or Cardinal Data (quantity)

① When data is classified in respect of successive "time" points or intervals then they known as time Series Data.
Ex: No. of student appear for CA final in last 12 years, Production of factory per month from 2000 to 2015

(2) When Data arranged "Region Wise" then it is known as "geographical data"

(3) When data arranged in respect of an attribute then it reffered to as "Quantitative Data" ex :- Nationality of Person, gender, smoking habit of group of individuals, swe

(4) When data are classified in respect of a variable, say height, weight, profit, salaries etc. they are known as Quantitative Data.

(5) Data may be classified as "frequency data" and "Non Frequency Data".

→ "Qualitative" as well as "Quantitative" data belongs to "Frequency group", where;

→ "Time Series Data" and "Geographical Data" belongs to "non - frequency group."

Mode of Presentation of Data :-

(a) Textual
(b) Tabular or Tabulation
(c) Diagramatic

→ (a) Textual Presentation → When data presents in pragraph or no. of paragraphs, it is first step to present any data. It is used to present in descriptive form.

(b) Tabular Presentation → This method is used to compare data internally and externally it is second step to present data.

Table has 4 main parts :-

1. Head Note
2. Box-Head
3. Caption & Stub
4. Body of Table
5. Source Note
6. Foot Note

(C) Diagramatic Presentation

"it is used to find out Hidden Trend Values"

This presentation is easily understood by a Lay-Man (or) Uneducated person.

★ Types of Diagramatic Presentation :-

1. **Line Diagram** → it is drawn for "time series data" (or) Non-Frequency data

- When two or more series are given then we use "multiple" line-diagram. (which have same unit)
- If When two or more series are given (which have different units) then we use "multiple axis line Diagram"

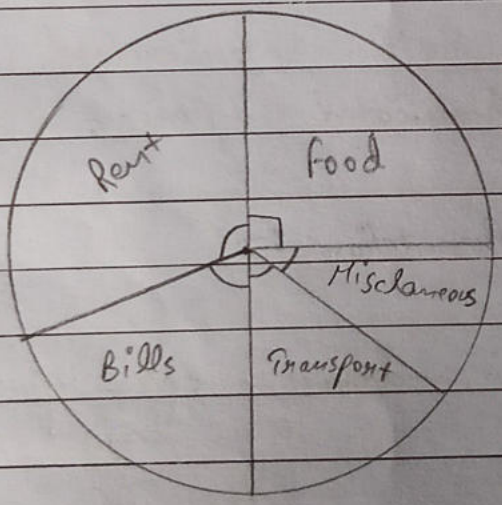
- If there is wide Range of fluctuation in the data then we use LOG (Yt) instead of Yt. and the Diagram is known as "Ratio - Chart of Logarithm - Chart"

2. Bar Diagram

Verticle Bar Diagram (Time Series Data & Quantitative Data)

Horizontal Bar Diagram (Geographical & Qualitative Data)

3. Pie Chart



it is used for individual data series which has only one head

$\therefore \frac{\text{Head Expenses}}{\text{Total Expenses}} \times 360^\circ$

Used for individual Data Series

Rent	=	3000
Food	=	2500
Transport	=	1000
Bills	=	1500
Misclaneous	=	2000
		10,000

$= \frac{2500}{10000} \times 360^\circ = 90^\circ$

★ OGIVE — It is a graph which is drawn on behalf of "Cumulative Frequency" (less than or more than)

- Ogive is always drawn for short-term Projection

- By using ogive we can obtain "Partition Value" – Median, Quantiles, Deciles, Percentiles

- Both ogive (less than or more than) are intersect each other at (Median $N/2$).

★ HISTOGRAM (or) Area Diagram
(Two Dimensional figure)

- Histogram is Drawn for Quantitative Data and Based on frequency of class interval and in histogram bars are of equal width but different in length and they are adjacent to each other.

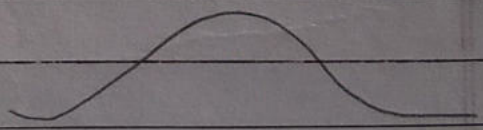
- To Draw Histogram class Interval must be "UNIFORM"

- When class Interval is not uniform then to Drew histogram we use frequency Density instead of Frequency.

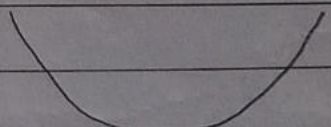
- Histogram is helpful in finding "MODE"

★ Frequency Curve

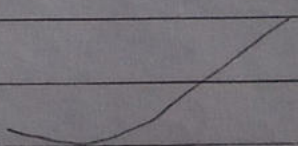
(a) Bell Shaped



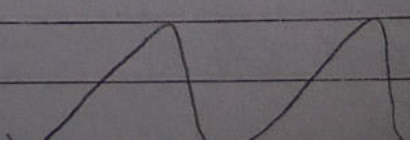
(b) U-Shaped



(c) J-Shaped



(d) Mixed Shape Curve



Relative Frequency = $\frac{\text{Frequency of Class}}{\text{Total Frequency}}$

★ Frequency Distribution → It may be defined as a tabular representation of Statistical Data usually in an ascending order.

Some important terms associated with a frequency distribution :-

1. Class Limit (CL)

Upper Class Limit (UCL)
Lower Class Limit (LCL)

2. Class Boundary

Lower Class Boundary (LCB)
or
Upper Class Boundary (UCB)

$$LCB = LCL - \frac{D}{2} \qquad (\because D = \text{Difference})$$

$$UCB = UCL + \frac{D}{2}$$

Example $LCB = 44 - \frac{(49-48)}{2}$ kgs
$= 43.50$ kgs

$UCB = 48 + \frac{(49-48)}{2}$ kgs
$= 48.50$ kgs

3. Mid - Point / Mid - Value / Class Mark

$$\frac{LCL + UCL}{2}$$

(or)

$$\frac{LCB + UCB}{2}$$