

# Statistical Description of Data

Statistics is a method by which we collect data, and then analyse 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, medieval period, as well as modern period."

STATISTICS is an universal  
Application.

## Definition :-

In singular sense - STATISTICS is a scientific method  
by which we collect, analyse,  
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.

## Application of Statistics :-

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

Statistics is concerned with both Qualitative & Quantitative information. But statistics

Deals with Quantitative information.

"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 deal 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 information." about some Particular Characteristics under consideration.

## Division of Data :-

### Internally Division

- Qualitative Data
- Quantitative Data

### Externally Division

- Primary Data
- Secondary Data

## 1. Collection of Primary Data

- (i) Interview Method
- (ii) Indirect Interview Method
- (iii) Telephonic Interview Method.

(i) Interview Method :- In Natural Clarity we collect the necessary Data much more quickly and accurately by this method.

(ii) Indirect Interview :- In case of Practical Problem in reaching respondents directly, like in the case of rail accident, we collect data from other persons not directly with injured.

(iii) Telephonic Interview :- It is a quick and cheap way of collecting the Primary Data. It covers large area. "Wide coverage."

2. Observation method :- using instrument like :-  
Height, BP, sugar.

3. 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."

4. Questionnaire filled and sends by Enumerator  
(or)

"Survey method"

b) sources of secondary Data :-

(i) International sources like :- Reports of WHO, world Bank, IMF etc.

(ii) Govt. sources like

(iii) Private & Quasi govt. sources like :- NCERT

(iv) unpublished sources like :- research institute.

SCRUPITTY OF DATA

"it means verification of accuracy and consistency of Data."

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

$$\therefore \text{Density} = \frac{\text{Population}}{\text{Area}} \text{ holds.}$$

## Presentation of Data

Classification of Data :-

1. Time series Data or Chronological Data.
2. Geographical or Spatial Data.
3. Quantitative Data. (OR) Cardinal Data.
4. Qualitative Data. (OR) ordinal Data.

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 or.

2.) When Data arranged "Region wise" then it is known as "geographical Data."

3.) When Data arranged in respect of an attribute then it referred to as "Qualitative Data". ex:- Nationality of person, gender, smoking habit of group of individuals, sweetms. etc.

4.) When Data are classified in respect of a variable, say height, weight, Profits, 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) Diagrammatic .

→ (a) Textual Presentation :- when Data presents in Paragraph (or) no. of Paragraphs.  
it is first step to present any Data.

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

→ Head Note  
Growth of Population in India

→ Box Head (or) caption

Location	worker	Non-worker	Total
Rural Areas			
Urban Areas			
Total			

Stub ←

BODY

\* Source :-

\* Foot note :-

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)

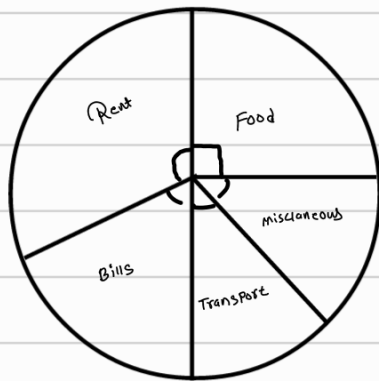
- when two or more series are given then we use "multiple line-diagram". (which have some unit)
- when two or more series are given (which have different units) then we use "multiple axis line Diagram."
- If there is wide Range of fluctuation in the Data then we use  $\log(YT)$  instead of  $YT$ .  
and the Diagram is known as  
"Ratio-chart of Logarithm - Chart."

## 2.) Bar Diagram :-

Vertical 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.

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

$$\text{Rent} = 3000$$

$$\text{Food} = 2500$$

$$\text{Transport} = 1000$$

$$\text{Bills} = 1500$$

$$\text{Miscellaneous} = \frac{2000}{10,000}$$

Used for individual Data Series.

$$= \frac{2500}{10,000} \times 360^\circ = 90^\circ \underline{\underline{90}}$$

\* Ogive :- Ogive 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, Quartiles, Deciles, Percentiles.

(\*) Both ogive (less than or more than) are intersect each other at  $(\text{Median } N/2)$ .



Snow Depth (x)	Frequency (f)	End point	Cumulative frequency
300 - 310	1	310	1
311 - 320	2	320	3
321 - 330	5	330	8
331 - 340	3	340	11
341 - 350	7	350	18
351 - 360	2	360	20

### ⊗ 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.

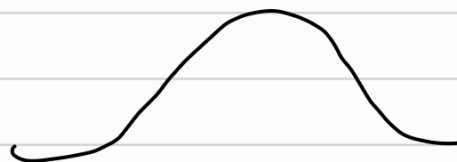
⊗ To Draw Histogram class interval must be "Uniform."

"When class interval is not uniform then to Draw histogram we use frequency Density instead of frequency."

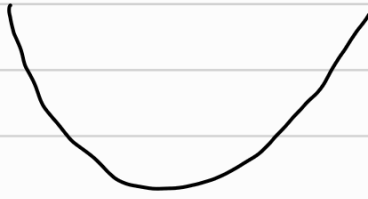
Histogram is helpful in finding "MODE"

\* Frequency curve :-

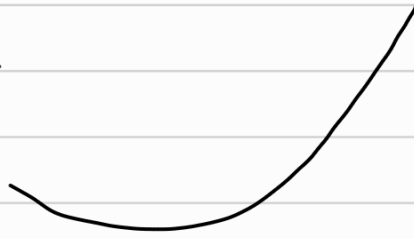
(a) Bell Shaped



(b) U-shaped



(c) J-shaped



(d) Mixed shape curve



$$\text{Relative frequency} = \frac{\text{frequency of class}}{\text{Total frequency}}$$

\* Frequency Distribution :- frequency distribution 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 Boundry

Lower class Boundry (LCB)  
or

Upper class Boundry (UCB)

$$LCB = LCL - \frac{D}{2}$$

$$UCB = UCL + \frac{D}{2} \quad (\because D = \text{Difference})$$

ex:-

$$\begin{aligned} LCB &= 44 - \frac{(49 - 48)}{2} \text{ kgs.} \\ &= 43.50 \text{ kg.} \end{aligned}$$

$$\begin{aligned} UCB &= 48 + \frac{(49 - 48)}{2} \\ &= 48.50 \text{ kgs.} \end{aligned}$$

## 3. mid-point / mid-value / class mark

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

(OR)

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

