Statistical Description of data

Singular Sense - Scientific method that is employed for collecting, analysing and presenting the data to draw statistical inferences.

Plural Sense - Data qualitative as well as quantitative.

language		Word
Latin	→	Status
Italian	→	Statista
German	→	Statistik
French	→	Statistique

> Statistics deals with the aggregate and not with individual data

> Statistics is concerned with quantitative data as well as qualitative data

Data

Primary
The data which are collected for the first time by investigator

Secondary
Collected data used by different person or agency

Variable

Discrete variable
- When a variable assumes a finite or contably infinited number of isolated values

Ex - marks of student, Salary of a person etc

Continuous variable
- When a variable assumes any value from a given interval

Ex - Height, profit scale, weight etc

Attribute - A qualitative characteristic is known as an attribute.
Ex - nationality of a person, colour

Collection of primary Data - Methods -

i) Interview method
- Telephone interview
- indirect interview
- Personal

ii) Mailed questionnaire

iii) Observation

iv) Questionnaire filled by enumerator

Scrunity → Checking accuracy & consistency of data.

Internal consistency -
When two or more series of related data are given, we should check consistency among them.

Popular Data Classification -

1) Chronological / Temporal / Time series → Time points
2) Geographical / spatial series Data. → Region wise
3) Qualitative / oridal Data. → Attribute
4) Quantitative / cardinal Data → Respect of a variable

Data

frequency (quantitave & quality) | Non-frequency (geographical & chronological)

Modes of Presentation of Data -

i) Textual Presentation → paragraphs. Ex - official Reports → manifold not recomended

ii) Tabular presentation → data with statistical table

iii) Diagrammatic → Attractive presentation

ii) Table → Stubs - left part of table (Rows)

→ Caption - upper part of table (columns)

→ Footnote - source of data

→ Body - contains the numerical figures

→ Box head → Columns or sub-columns & unit measurement

Q. Out of 1000 persons, 40% are female, others are male. In a marriage function 300 persons enjoyed the song 30% of ppl who had not enjoyed the song were female. What is no. of male, who did not enjoy the song in the function?

	M Enjoy	Not enjoy	
Female		210	400
Male	→	490	600
	300	700	1000

Ans → ~~700~~ 490

Types of Diagrams - Line

i) Line / histogram → time series

ii) Bar diagram

iii) Pie chart → Circular presentation of relative data

Q - Two sales - persons present their no. of sales per week for a month
↳ Adjcent bar chart

Pie chart → Segment = $\frac{(\text{segment value} \times 360)}{\text{total value}}$.

Frequency → No. of times a particular observation repeated.

frequency Distribution table → It contains intervals in one column & corresponding frequency in the other.

Types -

1) Ungrouped Frequency - When there are limited no. of distinct observations, frequency can be assigned to each one of them

2) Grouped frequency distribution -

When there are large no. of observation.

Non-overlaping / mutually inclusive classification

LCL ≠ LCB, so less 0.5

UCL ≠ UCB, so add 0.5

Overlaping / Mutually exclusive classification

LCB = LCL

* Class limit = class boundary.

* Class length = Upper class limit − Lower class limit

* Mid-pt = $\frac{LCL + UCL}{2}$

* No. of class interval = $\frac{\text{Range}}{\text{Class length}}$