Stastical Description of Data
Data

— 8 marks

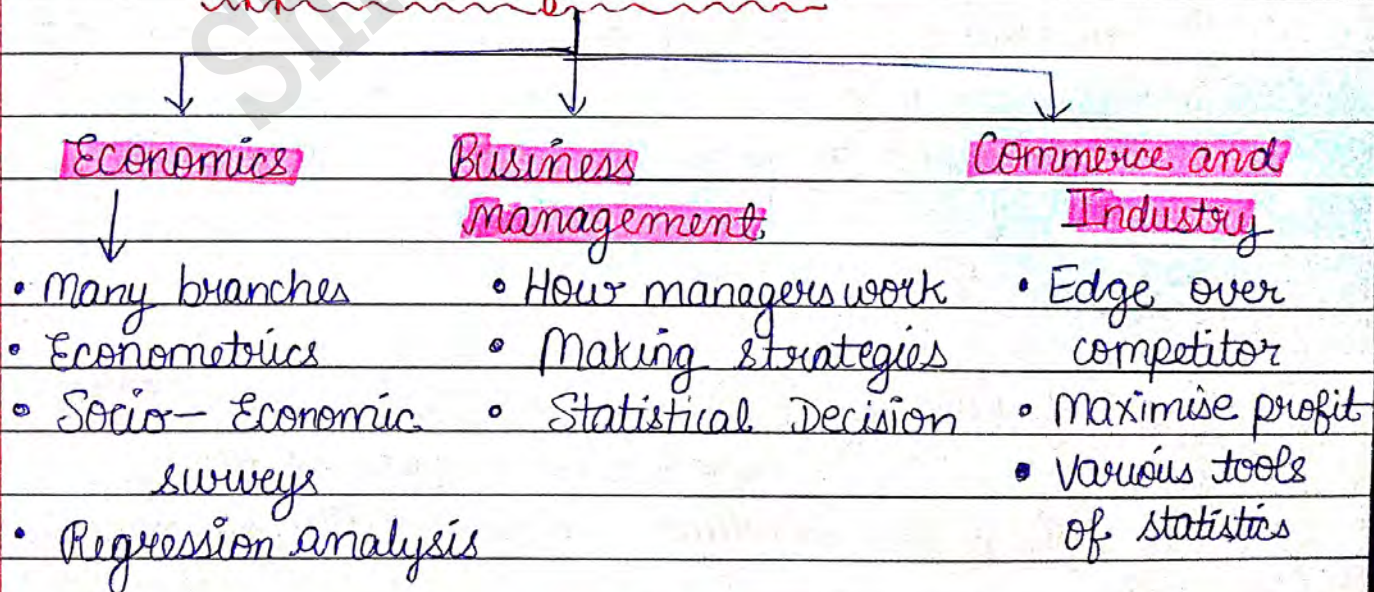### Statistics

— Latin word status
— Italian word statista
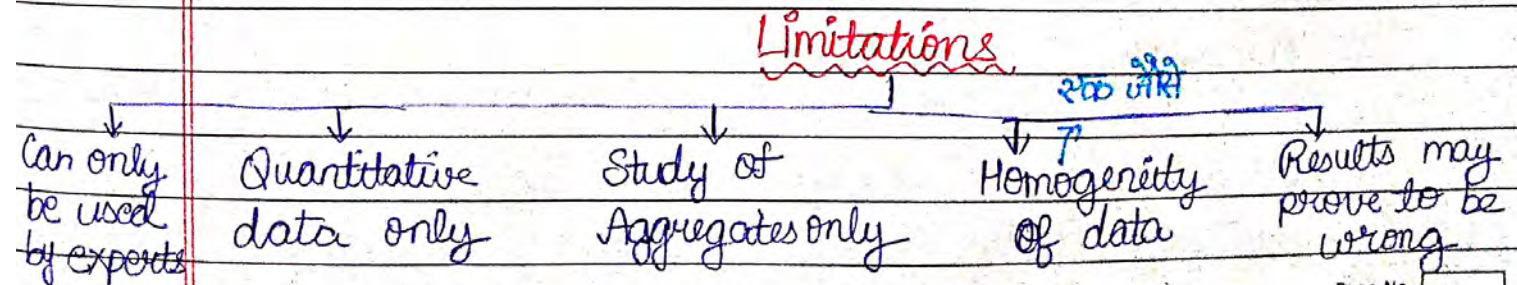— German word statistik
— French word statistique

✱ Plural sense : Collection of data for some specific purpose.

Singular sense : Science of collecting, presenting, analysing.
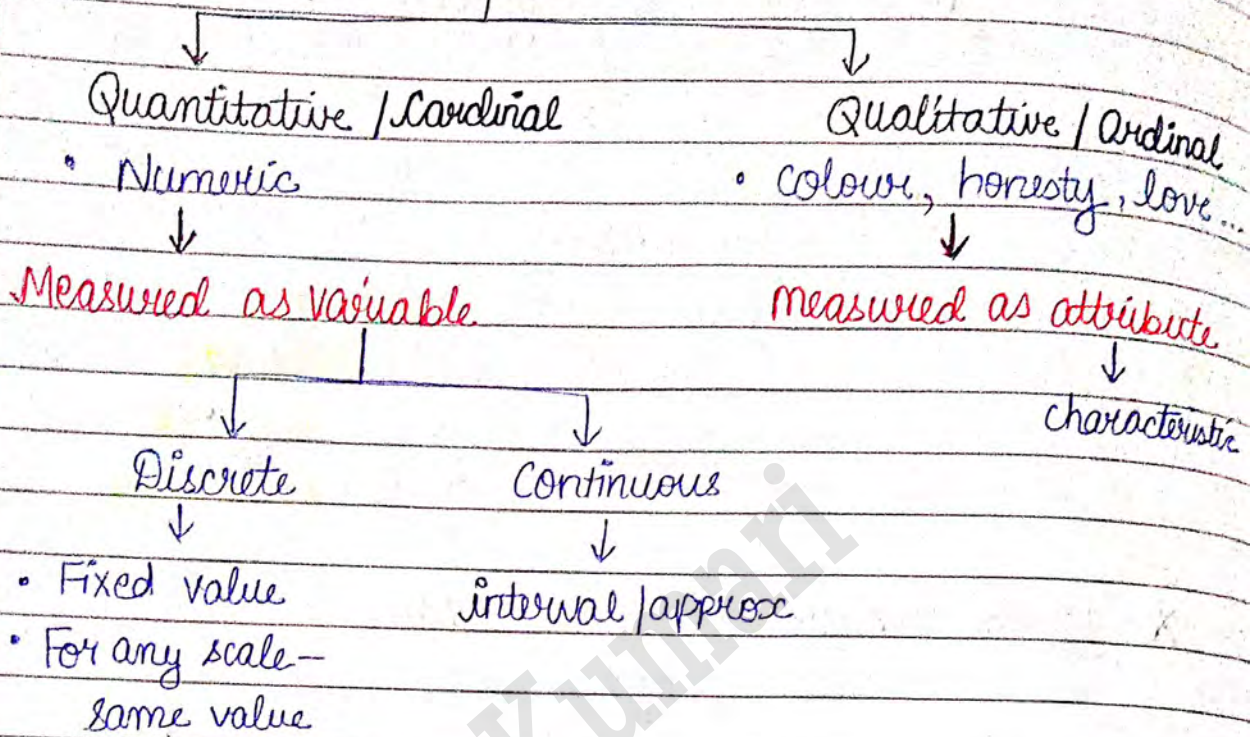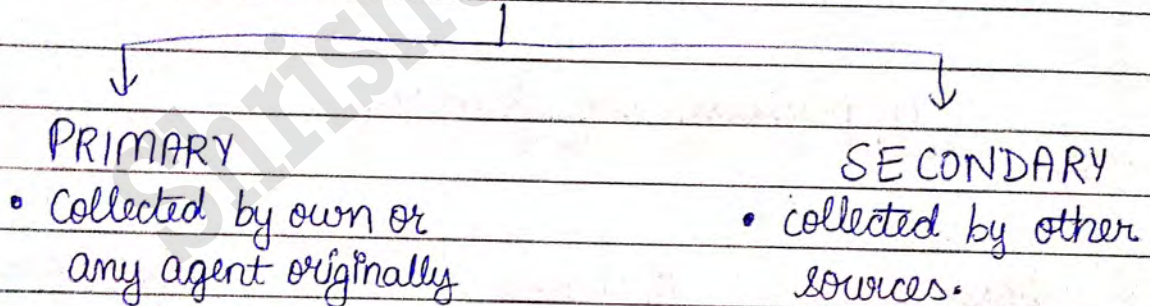
Application of Statistics

Economics	Business Management	Commerce and Industry
• Many branches	• How managers work	• Edge over competitor
• Econometrics	• Making strategies	• Maximise profit
• Socio- Economic surveys	• Statistical Decision	• Various tools of statistics
• Regression analysis		

Limitations

सब जोड़ी

Can only be used by experts	Quantitative data only	Study of Aggregates only	Homogenity of data	Results may prove to be wrong

Collection of Data

Quantitative / Cardinal
- Numeric

↓

Measured as variable

Discrete ↓
- Fixed value
- For any scale— same value

Continuous ↓
interval / approx

Qualitative / Ordinal
- colour, honesty, love...

↓

Measured as attribute

↓

characteristic

Source of data

PRIMARY
- Collected by own or any agent originally

↓

SECONDARY
- collected by other sources.

4 ways to collect data

1) Interview ⟨ Direct — Natural calamity
Indirect — Manmade accident (दुर्घटना) ⟩ Face to face
Telephone — Podcast (quickest method)

2) Mailed questionnaire — Google form (max. non responses)

3) Observation — PT teacher

4) Questionnaires filled and — Aadhar card
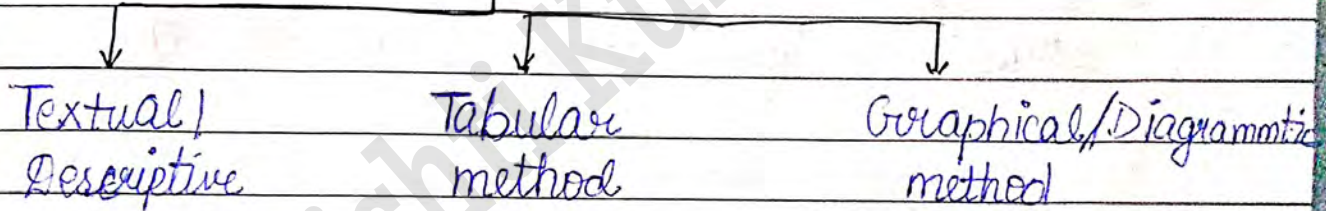sent by enumerators

Sources of Secondary data :-

- International sources like WHO, ILO, ...
- Govt. sources like ministry of Food, ...
- Private and quasi – govt. sources like NCERT ...
- Unpublished sources of various research institutes

* ## Scrutiny of data
  ↓
  For quality check

  Presentation of data
  ↓   ↓   ↓
  Textual/   Tabular   Graphical/Diagrammatic
  Descriptive   method   method

* ## Classification / Organising
- makes classes grouping similar things together

Objectives
— Simplification and Briefness
— Comparability
— Statistical Analysis
— makes data more understandable.

Types   → or Data Varying over the TIME.
— Data Chronological or Temporal or Time Series (Years)
— Geographical or Spatial Series Data (States)
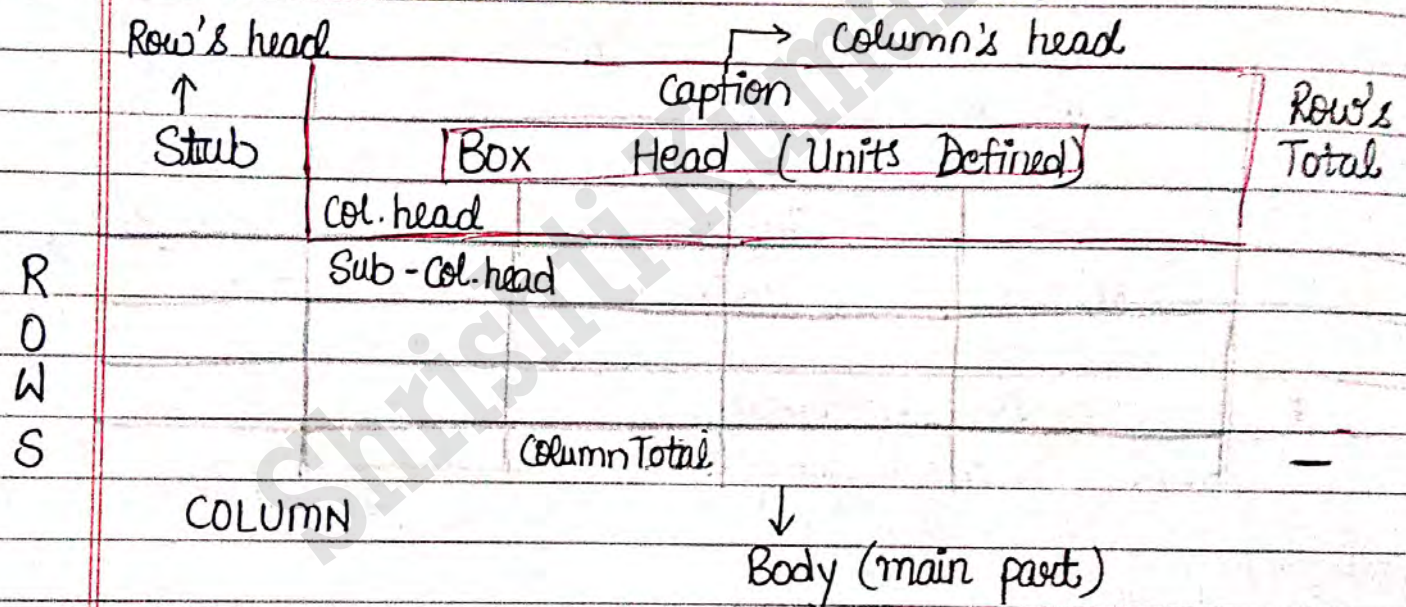   or Data Varying over the SPACE

✫ Frequency data → Time Series → Quantity / Qualitative
 • Number of times a data value occurs

✫ Non - Frequency data → Qualitative → Geographical / Time Series
 • Single data value

(TABULAR PRESENTATION)

Table no :- 01
Table name :- Jio Balance Sheet

Row's head → Column's head

Caption — Row's Total

Stub — Box Head (Units Defined)

Col. head
Sub - Col. head

R
O
W   Column Total
S

COLUMN   ↓
Body (main part)

Source :-
Footnote & Abbreviation

(DIAGRAMMATIC PRESENTATION)

i) Line diagram / Historiagram
Time — X axis
Fluctuations are wide. — use logarithm (Log Chart - Ratio Chart)

Potato
Tomato ✳ Multiple line chart (multiple data in one chart) (Same unit)
Potato
Milk   ✳ Multiple axis chart (multiple data in one chart with diff. unit)

ii) Bar diagram

➢ Vertical Bar diagram :- • Time series • Quantitative
➢ Horizontal Bar diagram :- • Qualitative
           • Data varying over space
➢ Multiple or Grouped Bar diagram :- 2011 (Jan, Feb, March)
➢ Component or Sub-divided Bar diagram :- Jan (Potato, Tomato, Onion Sale)

➢ Dividend Bar diagram/:- Compare diff. components
'/. Percentage Bar diagram Comparing diff. components with
              table value.

iii) Pie diagram <mark>(Comparing diff. comp. to total value.)</mark>

$$\text{Angle for a component} = \frac{\text{Value of Comp.}}{\text{Total value}} \times 360°$$

✱ Best Presentation — Tabular.
  Hidden trends — Diagrammatic.
  Most accurate — Tabular.

Frequency Distribution of data

Mutually exclusive : एक value एक class में आएगी,
Exhaustive : all data is put in one of the class

Two types

1) Ungrouped — Data is in discrete variable (no classes)

2) Grouped — making classes

* Range = Largest value − Smallest value

* No. of class = $\dfrac{\text{Range}}{\text{Class Length}}$

* No. of class × class length ≃ Range

* Class Interval = 45 − 49 → lower class limit (LCL), upper class limit (UCL)
  Class Length = 49 − 45 = 4

☆ For Continuous variable, <mark>class limit</mark> and <mark>class Boundaries</mark> are <mark>same</mark> and here we <mark>exclude</mark> the <mark>upper limit</mark> so it is mutually exclusive.

▷ To Convert Class Limit into Class Boundary :-

$m = \dfrac{\text{UCL of a class} - \text{LCL of a class}}{2}$

$UCB = UCL + m$
$LCB = LCL + m$

> Discrete variable → Grouped frequency distribution
  → Class Interval → Mutually Inclusive.

* $\dfrac{\text{Class Mid-point} /}{\text{Class mark}} = \dfrac{L.C.L + U.C.L}{2}$ or $\dfrac{L.C.B + U.C.B}{2}$

CUMULATIVE FREQUENCY

Less than   More than

Frequency Density = $\dfrac{\text{Frequency Of Class Interval}}{\text{Class Length}}$

Relative Density = $\dfrac{\text{Frequency}}{\text{Total Frequency}}$

Percentage Frequency = $\dfrac{\text{Frequency}}{\text{Total Frequency}} \times 100$

Sum of Relative Frequency = 1 [Lies between 0 - 1]

Sum of % Frequency = 100%

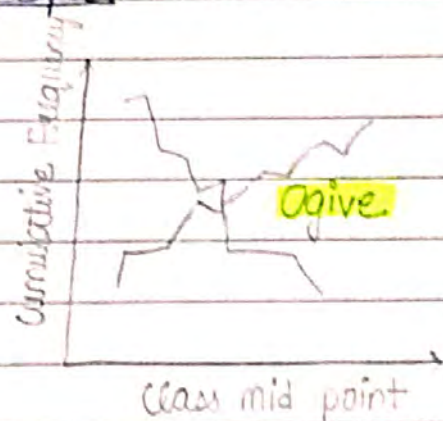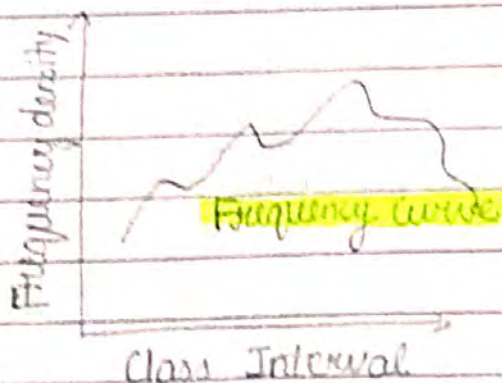## Graphical Representation of Frequency Distribution

Mode ↑

1) Histogram (Class boundaries)
2) Frequency Polygon (Joining the mid-points of Histogram and ends)
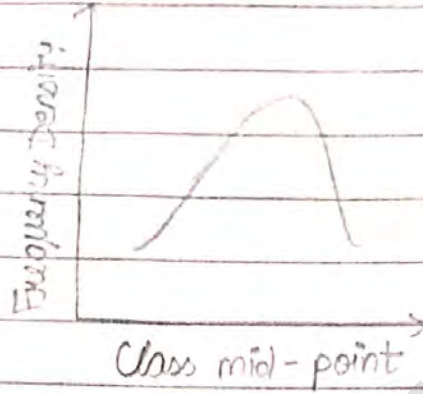3) Frequency Curve (limiting form of polygon or histogram)
4) Ogives / Cumulative Frequency graph. (Less than or More than)
   ↓
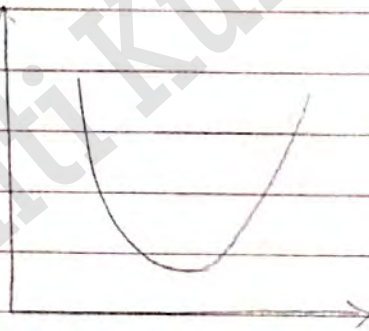   Median, Quartiles (Q₁,Q₂,Q₃)

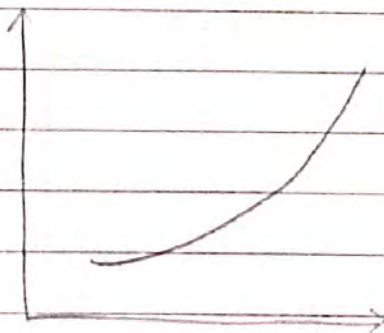## Four types of Frequency Curves :-

1) Bell - Shaped   (max. in centre)

2) U-shaped   (Traffic) (min in afternoon)

3) J- shaped   (Traffic)

4) Mixed Curve