

Theorystat statistics statistical and sampling

* singular sense

• scientific method employed for collecting, analysing and presenting data to draw statistical inferences.

→ Plural sense

• data qualitative as well as quantitative that are collected, usually with view of having statistical data.

Latin → Status

Italian → Statista (Pasta)

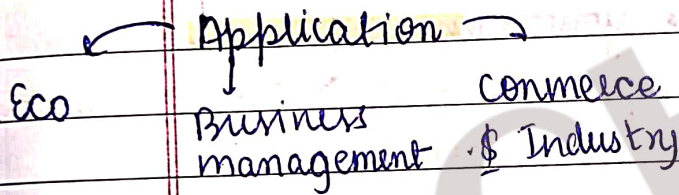
German → Statistik

French → statistique.

• Kautilya's Arthashastra
↳ birth and death.

• Ain-i-Akbari

↳ Statistical record on agriculture

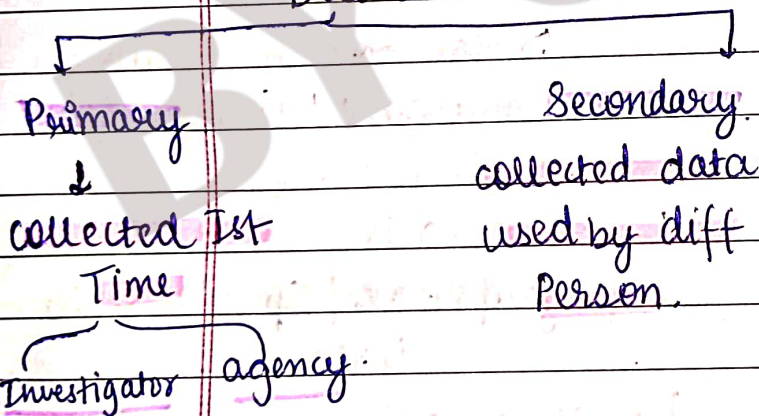
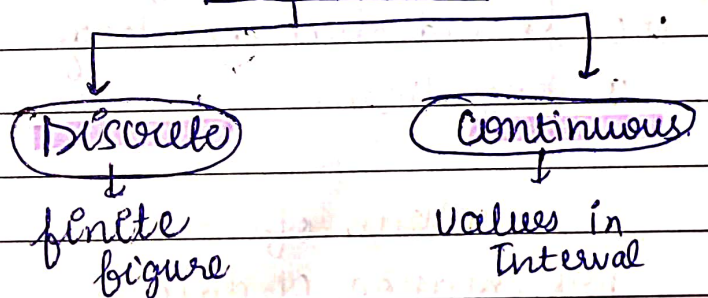
Limitation

• deal with aggregate not with individual.

• quantitative data only.

• can also with qualitative

• future project ✓ possible

DataVariablesAttribute

Qualitative characteristic is known attribute.

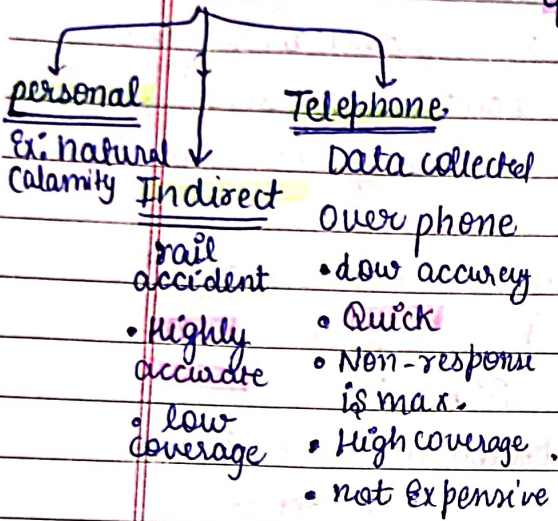
Ex:- Gender of baby

Ex: no. of petals in flowers.

Ex: height, weight

Collection of primary data

Interview Method



Mailed questionnaire

- coverage is wide.
- non response is max.

Observation

- more accurate
- Time consuming
- laborious
- cover only small area

Questionnaire filled by Enumerator

- used in case of survey and census.

Security of data

- checking accuracy and consistency of data.

Source of Secondary data

- International sources
- Government sources
- private and Quasi-government sources.
- unpublished sources

Internal consistency

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

Mode of Presentation of data.

- Textual → Paragraph.
- Tabular → ~~graph~~ Table
- Diagrammatic → Graph.

Data classification

- chronological / Temporal / Time series
 - ↳ Time basis
- Geographical / Spatial series Data
 - ↳ Area basis
- Qualitative / Ordinal Data
 - ↳ Gender, smoking etc.
- Quantitative / cardinal Data
 - ↳ number.

Tabular Presentation

- **Box head** → entire upper part of table include column and sub-column.
↳ unit of measurement comes in Box head.
- **Caption** → The upper part of table describe
↳ column
↳ sub-column.
- **Stub** → left part of table
↳ provide description of rows.
- **Body** → main part of table that contain the numerical figures
- **Footnote** → Source of data at the bottom of table.

Best Method of presentation of data

↳ Textual.

- Any hidden trend present in the given data can be noticed only in this mode of representation
(Diagrammatic representation of data)

Types of diagram

- line diagram | Histogram → ^{use for} ↑ time series.
→ wide fluctuation → log Chart or ratio chart is used
→ two or more series of same unit - Multiple line chart is used
→ two or more series of diff. unit → Multiple axis Chart

Bar Diagram → Bars i.e. rectangles of equal width and usually of varying lengths drawn either horizontally or vertically.

Pie Chart : It is used for circular presentation of relative data.

$$\text{Segment angle} \rightarrow \frac{(\text{Segment value} \times 360^\circ)}{(\text{total value})}$$

Frequency and distribution :

→ number of times a particular observation is repeated.

→ **Frequency distribution Table**

• It is a table which contains observations or class intervals in one column and corresponding frequency in the other.

Types

• simple Frequency distribution

When there are limited no. of distinct observ. frequency can be assigned to each one of them.

• **Grouped Frequency distribution**

10 - 20 20 - 30

Class intervals

Class limit

minimum value

(lower class limit)

maximum value

(upper class limit)

Overlapping

0 - 10, 10 - 20

mutually Exclusive classification

Non-overlapping

0 - 9, 10 - 19

mutually Inclusive classification

class limit = class boundary

Cumulative Frequency \rightarrow Two Types

less than c.f

more than c.f

less than c.f + more than c.f = Total frequency

Frequency Density \Rightarrow $\frac{\text{class frequency}}{\text{class length of class (UCB - LCB)}}$ upper class boundary
lower class boundary.

used in construction of histogram.

Relative Frequency \Rightarrow $\frac{\text{class frequency}}{\text{Total frequency}}$

Relative Frequency add up to unity MCQ

• Percentage Frequency \Rightarrow $\frac{\text{class frequency}}{\text{Total frequency}} \times 100$

Histogram
(Area Diagram)

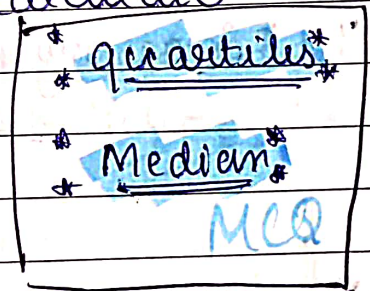
very convenient

way to represent a frequency distribution

Frequency polygon

c.f graph -
(ogive)

Calculate



v. important \rightarrow

Calculate Mode MCQ

\hookrightarrow Frequency curve: limiting form of a histogram or frequency polygon.

4 Types of curve

Bell shaped \rightarrow most commonly used distribution. MCQ
U shaped curve
J shaped curve
Mixed curve

Bar diagram \rightarrow 1 dimensional
Pie diagram \rightarrow 2 "
cube diagram \rightarrow 3 "

difference b/w

maximum value - minimum value = range

upper class - lower class \Rightarrow class interval.

histogram represent \rightarrow continuous series.

Ideographs also known as Pictographs