DATE :

Theory

~~stat~~ ~~stasticks~~ statistical and sampling

* **singular sense**
  - scientific method employed for collecting, analysing and presenting data to draw stastical inferences.

→ **Plural sense**
  - data qualitative as well as quantitative that are collected, usually with view of having stastical data.

- Latin → Status
- Italian → Statista (Pasta)
- German → Statistik
- French → statistique.

- Kautilya's Arthashastra
  - ↳ birth and death.
- Ain-i-Akbari
  - ↳ Stastical record on agriculture

Application

- Eco
- Business management
- Commerce & Industry

Limitation

- deal with aggregate not with Individual.
- quantative data ~~only~~.
- can also with qualitative
- future project ✓ possible

Data

- Primary
  - ↓ collected Ist Time
  - Investigator / agency.
- Secondary
  - collected data used by diff Person.

Variables

- Discrete
  - ↓ finite figure
  - Ex: no. of petal in flowers.
- Continuous
  - ↓ Values in Interval
  - Ex: height, weight

Attribute

Qualitative characteristic is know Attribute.

Ex:- Gender of baby

SPB

DATE :

Collection of primary data

- Interview Method
  - personal
    - Ex. natural Calamity
  - Indirect
    - rail accident
    - Highly accurate
    - low coverage
  - Telephone
    - Data collected over phone
    - low accuracy
    - Quick
    - Non-response is max.
    - High coverage
    - not expensive
- Mailed questionnaire
  - coverage is wide.
  - non response is max.
- observation
  - more accurate
  - Time consuming
  - laborious
  - cover only small area
- questionaire filled by Enumerator
  - used in case of survey and census.

Securuting of data

- checking accuracy and consistency of data.

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.

Source of Secondary data

- International sources
- Goverment sources.
- private and quasi-goverment sources.
- unpublished sources.

Data classification

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

DATE :

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 raws.
- **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.

MCQ

- 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

SPB

DATE :

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.

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

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

Overlapping: 0 - 10, 10 - 20 → Mutually Exclusive classification

Non-overlapping: 0 - 9, 10 - 19 → mutually Inclusive classification

Class limit

minimum value (lower class limit)

maximum value (upper class limit)

Class limit = class boundary

DATE :

Cumulative Frequency → Two Types
less than c.f | more than c.f

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

Frequency Density ⇒ $\frac{\text{class Frequency}}{\text{class length of class (UCB - LCB)}}$ upper class boundry, lower class boundry.
used in construction of histogram.

Relative Frequency ⇒ $\frac{\text{class Frequency}}{\text{Total Frequency}}$

Relative Frequency add up to unity. MCQ.

% Frequency ⇒ $\frac{\text{class Frequency}}{\text{Total Frequency}} \times 100$.

Histogram (Area Diagram) very convenient way to represent a frequency distribution

Frequency Polygon

c.f graph (Ogive) Calculate
* quartiles *
* Median *
MCQ

Calculate * Mode * MCQ. ← v. Important →

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

4 Types of curve:
- Bell shaped → most commonly used distribution. MCQ
- U shaped curve
- J shaped curve
- Mixed curve

Bar diagram → 1 dimensional
Pie diagram → 2 "
cube diagram → 3 "

difference b|w

maximum value − minimum value = range

upper class − lower class ⇒ class interval.

histogram represent → continous series.

Ideographs also known as Pictographs