Data as in Statistics

Dr. Gayatri Nimbhore

Associate Professor

Dept of Organon of Medicine

Motiwala (National) Homoeopathic Medical College & Hospital ,Nashik

Data as in Statistics

What is data ???

Statistic or datum : measured or counted fact or piece of information

Statistics or data : plural of the same

Collective recording of the observations is called **data**

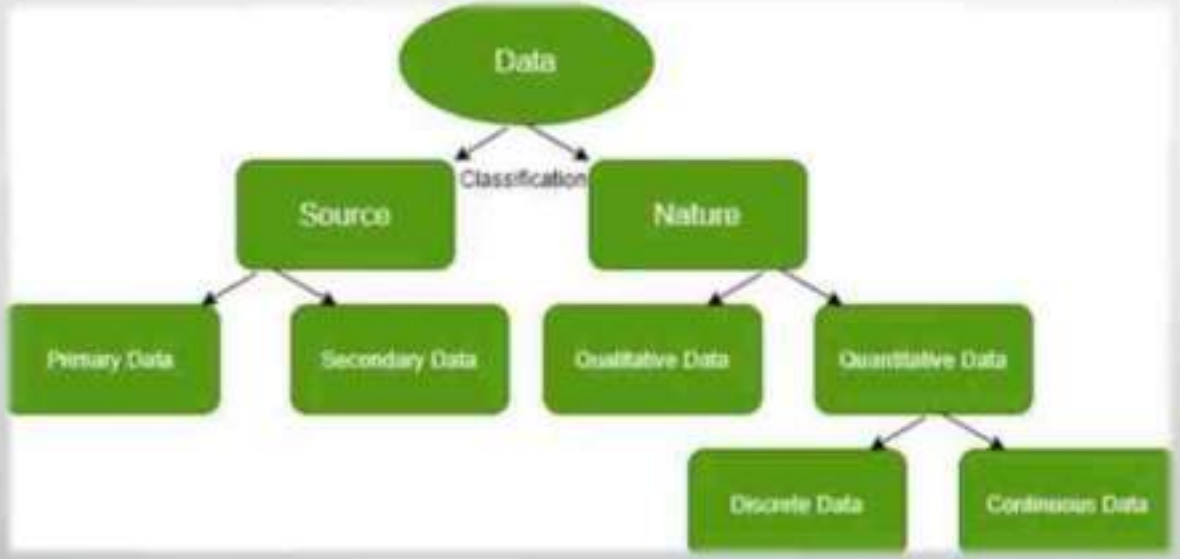
Data

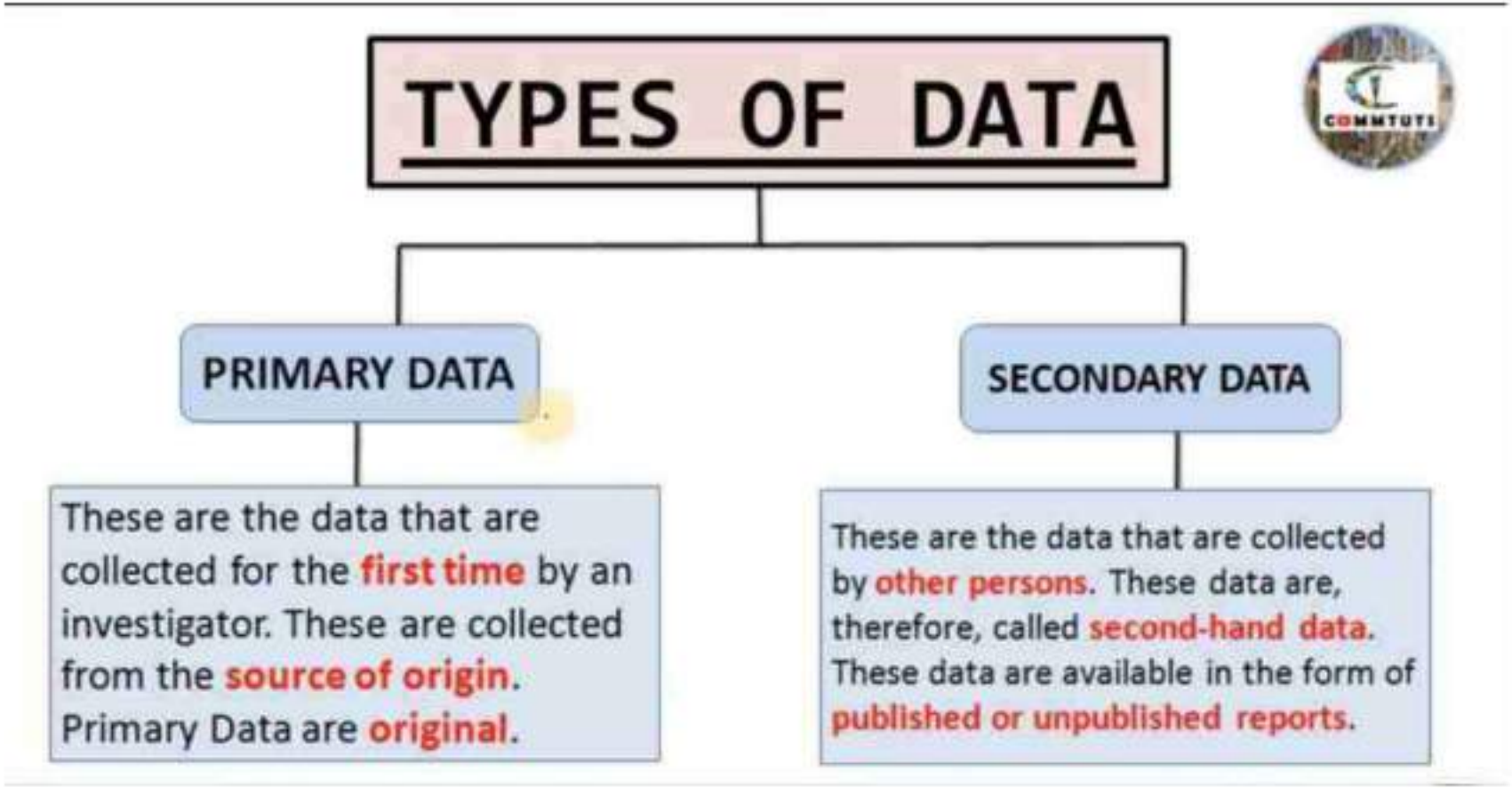
Data implies the **information in basic or unprocessed form**. It can be facts, details or statistics gathered by the researcher or any other person for the purpose of reference, reasoning and analysis.

Data

Data Collection is the process of acquiring information from different sources, about the topic under research. The function is performed by the researcher himself or his/her team.

Classification of data




TYPES OF DATA
PRIMARY DATA
These are the data that are collected for the first time by an investigator. These are collected from the source of origin. Primary Data are original.
SECONDARY DATA
These are the data that are collected by other persons. These data are, therefore, called second-hand data. These data are available in the form of published or unpublished reports.

TYPES OF DATA

DEPENDING UPON SOURCE OF DATA COLLECTION

- PRIMARY DATA
  - INTERVIEWS, EXAMINATION, QUESTIONNAIRES
- SECONDARY DATA
  - RECORDS, CENSUS DATA

	Primary	Secondary
Definitions	First-hand report of a study, experiment, procedure or event	Analyzes and interprets studies, events or procedures
Elements	Population, intervention, instruments, results, methods, implications and conclusion	Review, analyzes, parameters of included studies, table list of studies, procedures, interventions
Examples	Clinical Trials, Cohort, RCT, CCT Randomized Control	Systematic Review, Meta analysis, meta-synthesis, reviews, opinion

DIFFERENCE BETWEEN PRIMARY AND SECONDARY DATA

PRIMARY DATA	SECONDARY DATA
1) PRIMARY DATA are original because these are collected by the investigator from the source of their origin.	1) SECONDARY DATA are NOT original because these are already in existence.
2) PRIMARY DATA are costlier in terms of money, time and efforts.	2) SECONDARY DATA are less expensive.
3) PRIMARY DATA doesn't need adjustment for the concerned study.	3) SECONDARY DATA needs adjustment to suit the objective of study.
4) PRIMARY DATA are collected by investigator himself.	4) SECONDARY DATA are collected by other persons.

SOURCES OF DATA COLLECTION

PRIMARY SOURCE OF DATA

Primary Source of Data implies collection of data from its source of origin. It offers first-hand quantitative information.

SECONDARY SOURCE OF DATA

Secondary Source of Data implies collection of data from some agency or institution which already happens to have collected the data through statistical surveys.

Primary Data

Survey

Questionnaire

Experiment

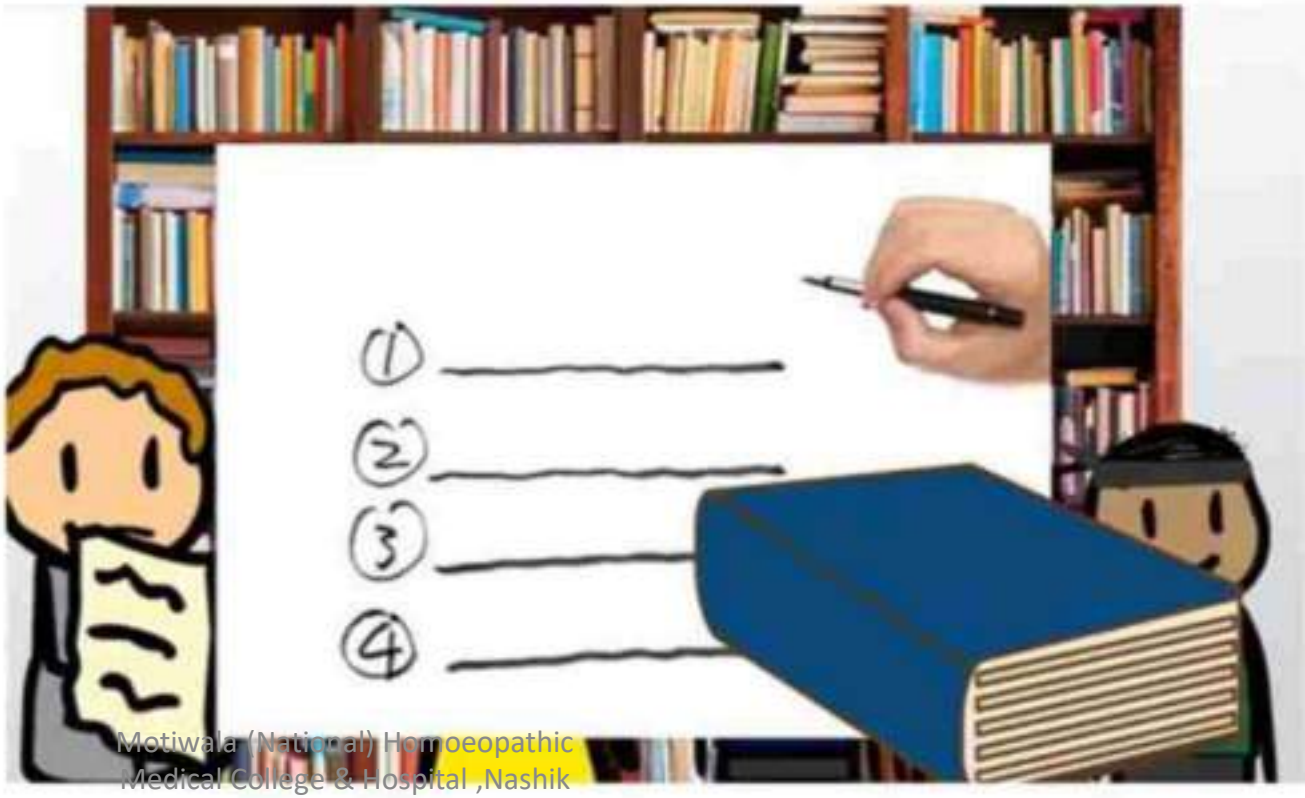
Direct Interview

Observation

Telephonic Interview

SOURCE OF SECONDARY DATA

B) Secondary data



Secondary Data

Newspaper

Websites

Government Publications

Records

Books

Data

Qualitative

"It was great fun"

Quantitative

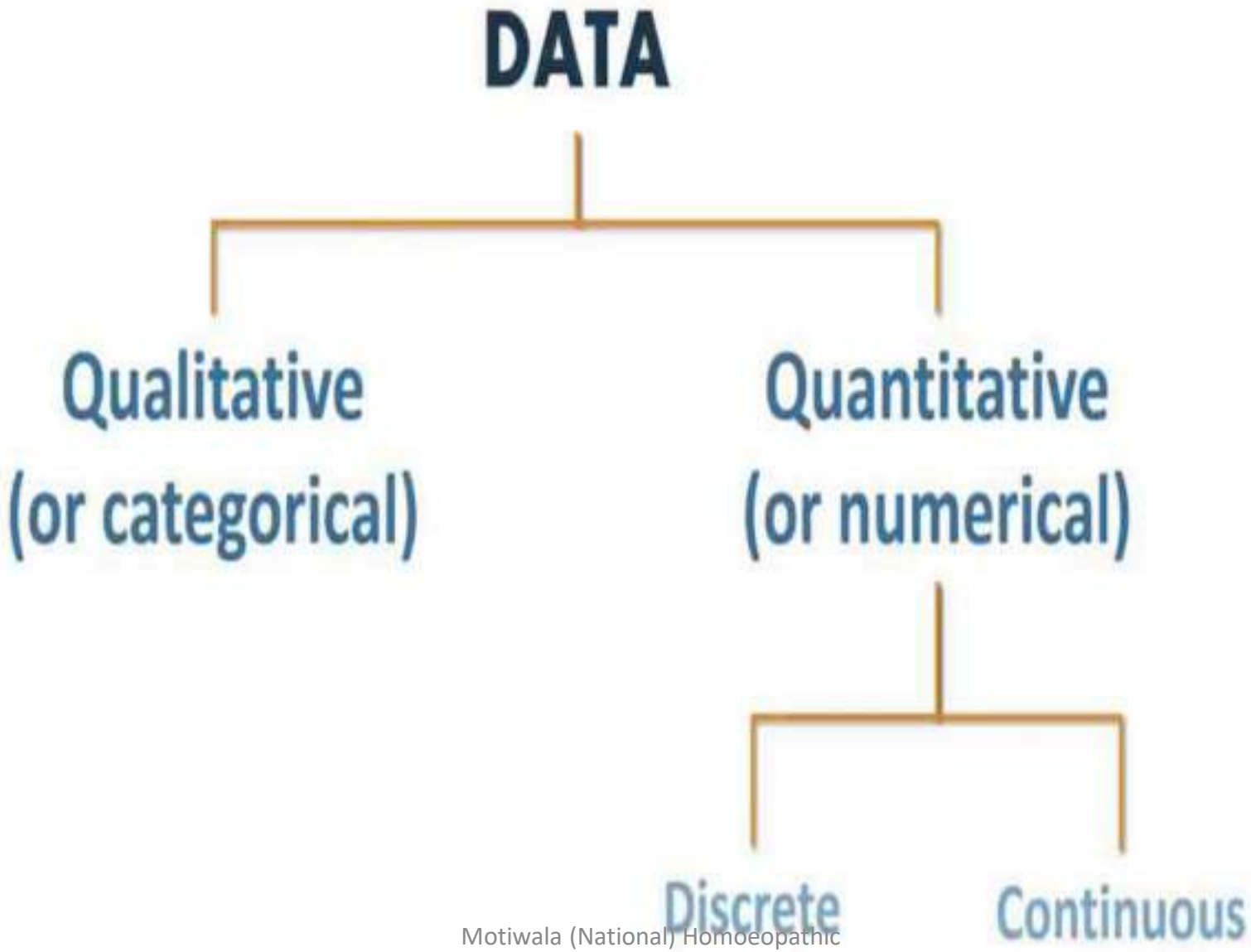
Discrete

5

Continuous

3.265...

Qualitative Data	Quantitative Data
Overview: • Deals with descriptions. • Data can be observed but not measured. • Colors, textures, smells, tastes, appearance, beauty, etc. • Qualitative → Quality	**Overview:** • Deals with numbers. • Data which can be measured. • Length, height, area, volume, weight, speed, time, temperature, humidity, sound levels, cost, members, ages, etc. • Quantitative → Quantity


DATA
Qualitative
(or categorical)
Quantitative
(or numerical)
Discrete
Continuous

Qualitative (Categorical) Data vs. Quantitative (Numerical) Data

Consists of *names* or *labels*.

Consists of *measurable quantities*.

Examples of Quantitative (Numerical) Data

Measurable quantities

The height (in inches) of a person

The weight (in pounds) of a cat

The number of siblings a person has

Examples of Qualitative (Categorical) Data

The major of a student

The color of a car

The area code of a phone number

CAUTION Categorical data are sometimes coded with numbers that are replacing names.

Types of Statistical Data

Quantitative Data

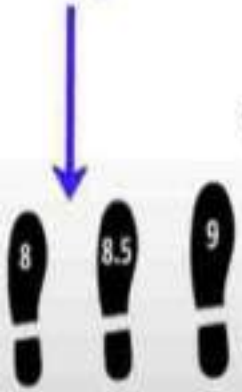
Discrete

Gaps between successive, possible values

Continuous

Can take on any value between two possible values

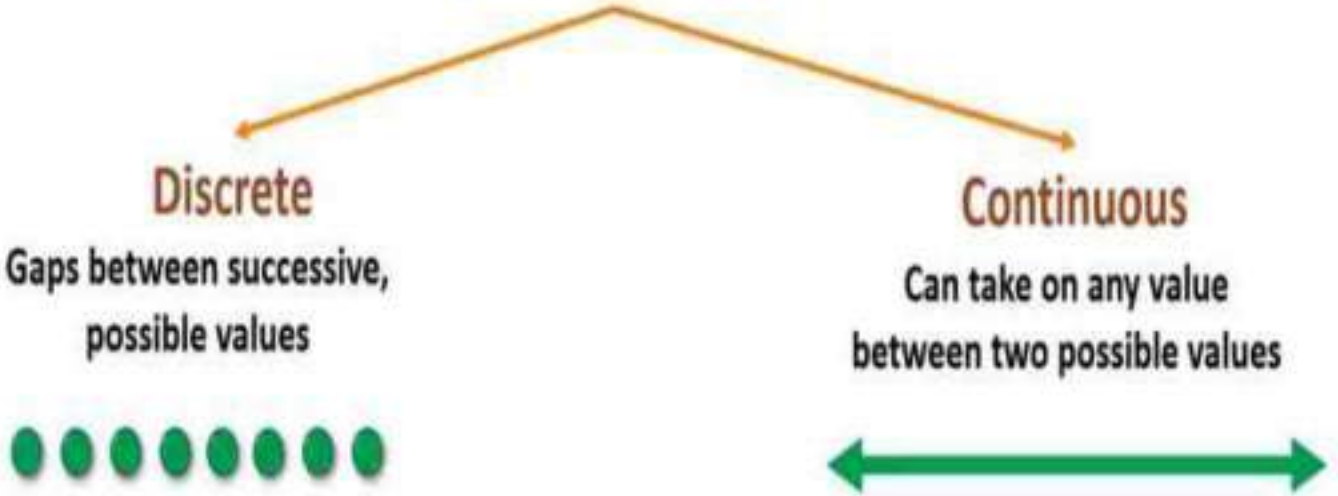
Gap



Examples:

- The number of pets a person has. (Ex: 0, 1, 2, 3,... and no values in between)
- The shoe size a person is wearing.

Quantitative Data



Discrete

Gaps between successive, possible values

Examples:

- The number of pets a person has. (Ex: 0, 1, 2, 3,... and no values in between)
- The shoe size a person is wearing.
- The cost of a gallon of milk.

Continuous

Can take on any value between two possible values

Examples:

- The speed (in miles per hour) of an airplane. (Ex: 400, 401, 402,... and every value in between. Ex: 400.25 mph.)
- The weight of a large pizza.
- The temperature (in degrees Fahrenheit) of a person.

Identifying Qualitative and Quantitative Data

Identify the type of data in each of the following:

1. Brand of toothpaste
2. Type of defect
3. Number of defects
4. Mass of a bearing
5. Color of hair
6. Weight of a person
7. Body Temperature

Identifying Qualitative and Quantitative Data

Identify the type of data in each of the following:

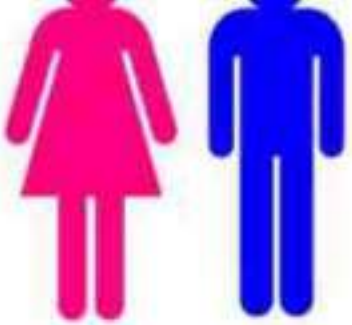
1. Brand of toothpaste — Qualitative
2. Type of defect — Qualitative
3. Number of defects — Quantitative
4. Mass of a bearing — Quantitative
5. Color of hair — Qualitative
6. Weight of a person — Quantitative
7. Body Temperature — Quantitative

Four Levels of Data

(in increasing order of information)

- Qualitative Levels
  1. Nominal
  2. Ordinal
- Quantitative Levels
  3. Interval
  4. Ratio

NOMINAL Data



- Use of Quality to depict the subjects
- Data on number of Males, females in an area
- Data on number of Rich, Poor and middle class subjects as per Socio economic status
- Unordered category

NOMINAL

****Nominal sounds like name****

Notes

- Lowest level of measurement
- Discrete categories
- No natural order
- Categorical or dichotomous
- May be referred to a qualitative or categorical

Examples

- Gender
  - 0 = Female
  - 1 = Male

  Dichotomous
- Group Membership
  - 1 = Experimental
  - 2 = Placebo
  - 3 = Routine

  Categorical
- Marital status, colour, religion, type of car ...

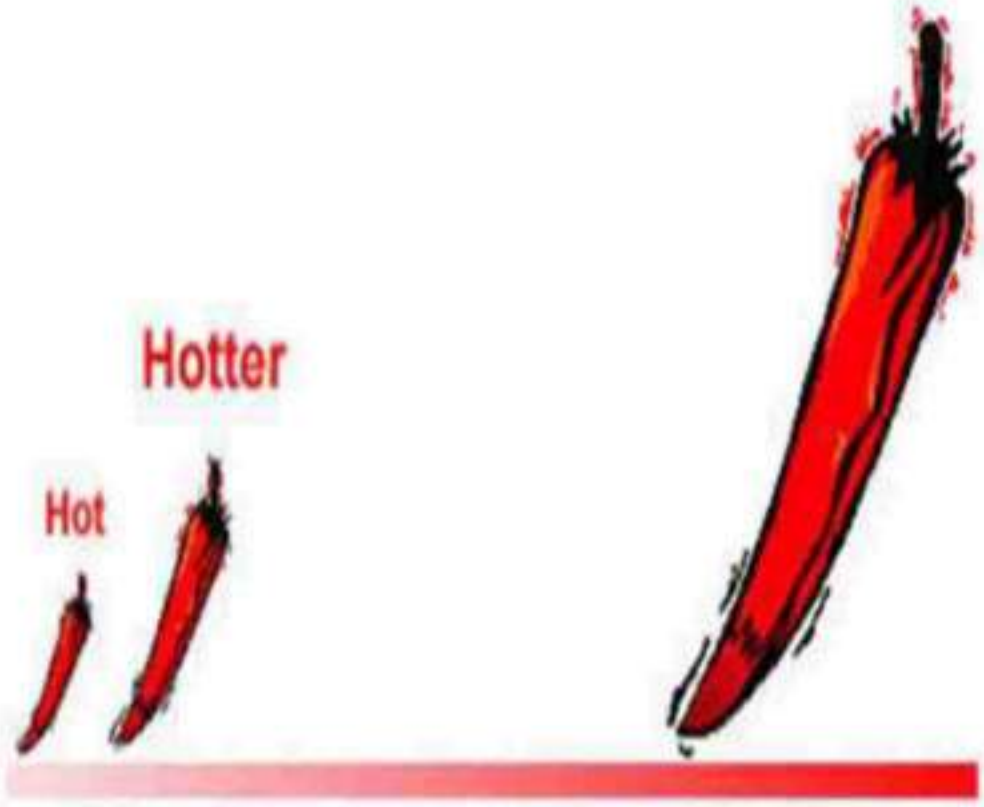
Nominal Data

- This type of data uses numbers as names
- Each number uniquely identifies a person or an object.
- Numbers on team jerseys are a good example.
- Arithmetic operations on nominal data don't make any sense.

Nominal Data

- This type of data is also called categorical data.
- If the data are in categories, it makes sense to talk about the frequency of each category.

Ordinal Data



- Numbers begin to take on meaning.
- A higher number means a greater amount of a property.
- The Mohs scale of hardness of minerals is an example.
- Differences between the numbers don't make any

Interval Data

- Differences between numbers begin to take on meaning.
- Celsius and Fahrenheit are good example.
- The difference between 20° and 30° is the same as the difference between 50° and 60°.
- But ratio statements don't make sense.

INTERVAL

Notes

- Ordered categories
- **Equal distance** between values
- An accepted unit of measurement
- **Zero is arbitrary**

Examples

Interval Data - Properties

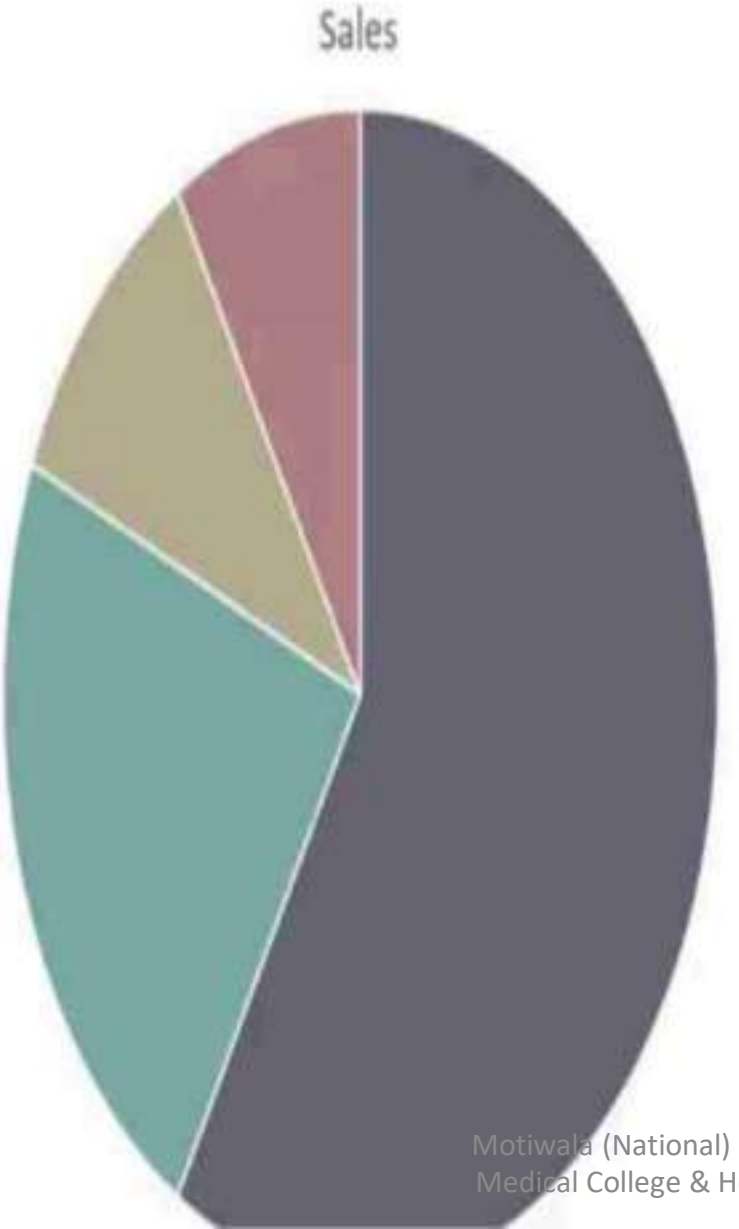
Measured

Ordered

Equidistant

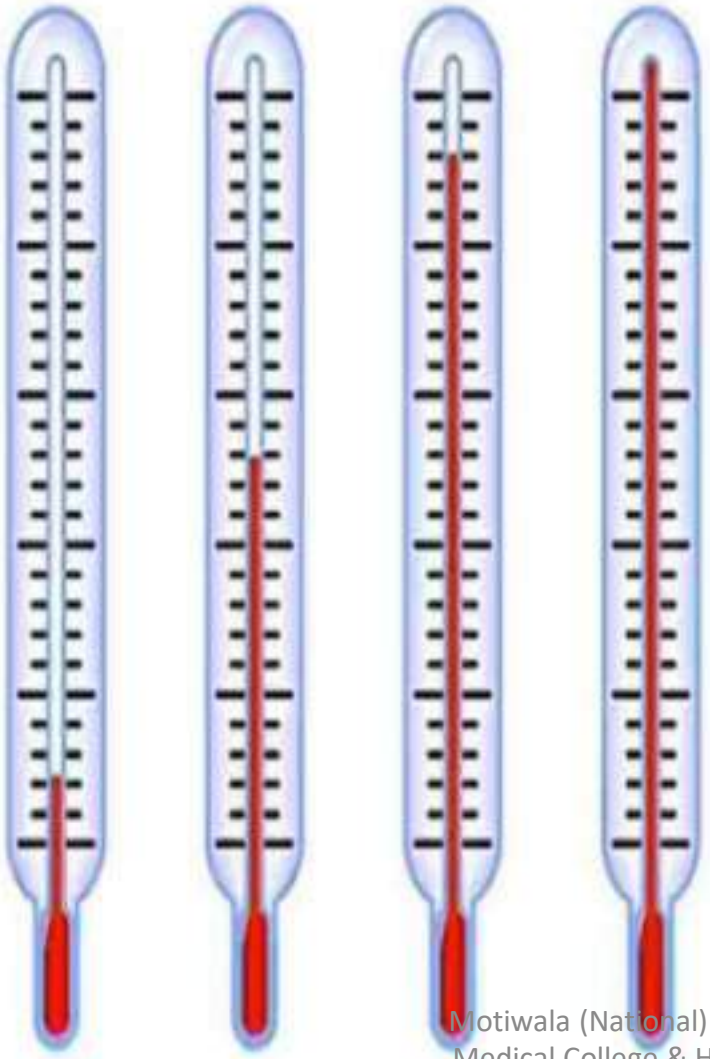
Meaningful 0

Negative



Ratio Data

- We can make statements like "twice as much as" or "half as much as."
- This type of data includes a zero point-the complete absence of the measured property.
- So it makes sense to say that 4 inches is twice as long as 2



Ratio Data

- Continuing with temperature.....
- The kelvin scale does have meaningful zero.
- This is the absence of heat no molecular motion
- So 200 k really is twice as 100 k.

Ratio Data - Properties

Measured

Ordered

Equidistant

Meaningful 0

Negative

Eg : There is nothing like -5 km in distance.

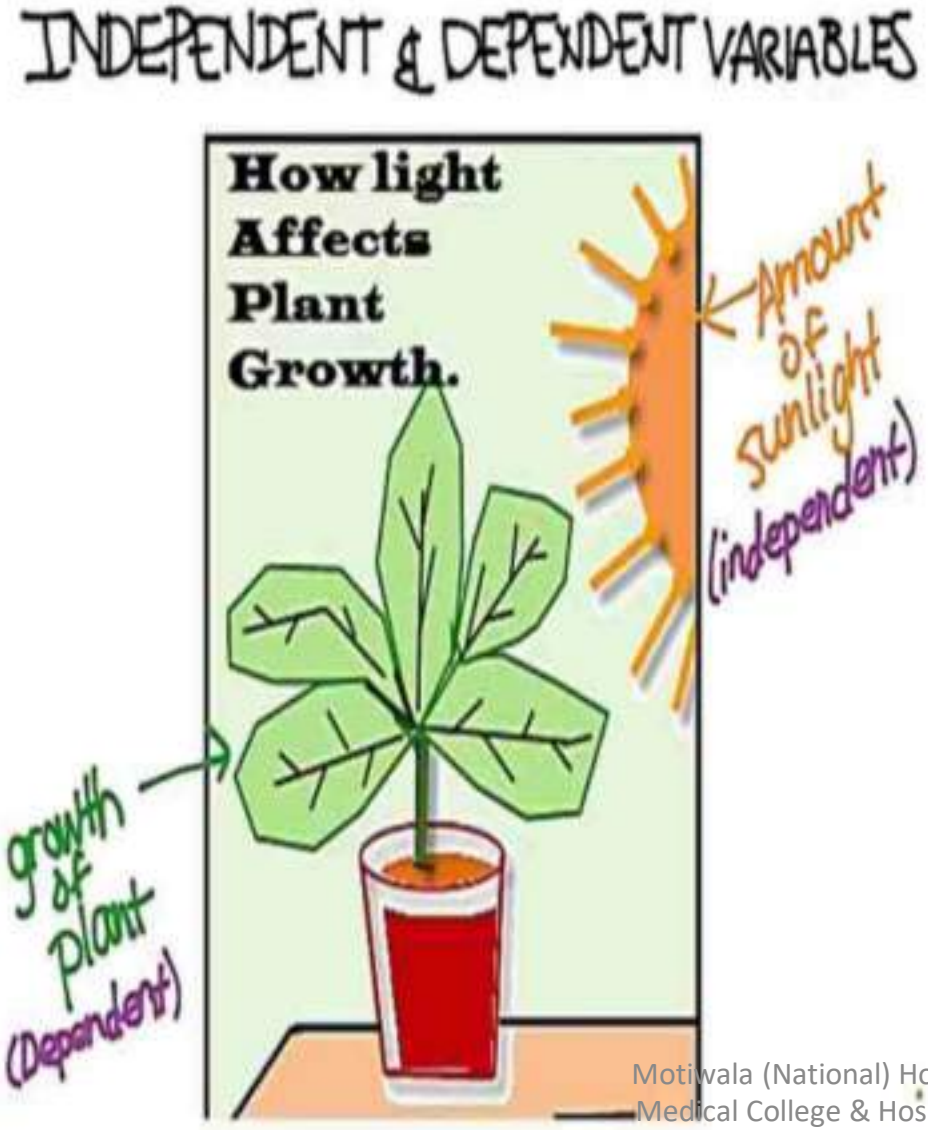
VARIABLE

What varies is called Variable

Ex. Age, serum calcium, intensity of pain, intra pulpal pressure, CAL, lumen of root canal, depth of palate, nucleus cytoplasmic ratio, daily OP, BMR.

What does not vary is called Attribute

Ex. Sex, religion, race.



DEPENDENT VARIABLES

Depends on the independent variables what a researcher measures

THANK YOU