➢ DATA & INFORMATION

DATA

- Data are text and numerical values.
- Data doesn't rely on information.
- Data does not directly helps in decision making.
- Data does not have any specific purpose
- Data are the variables that help to develop ideas.
- Example: Student test score

INFORMATION

- Information is refined form of actual data.
- Information relies on Data.
- Information directly helps in decision making
- Information carries a meaning that has been assigned.
- Information is meaningful data.
- Example: Average score of class from a given data.

➢ HOW DATA USEFUL

- For informed decision making.
- For Problem Solving.
- For Greater Understanding.
- For Improving Process
- Reduce time

➢ BIG DATA

- Collection of data in huge volume.
- Data with huge size.
- No data management tool can store.
- Expands in every fraction of seconds.
- Examples: Social media, Online Platforms like Flipkart, Amazon.
- Types of Big Data :-

I. **STRUCTURED DATA**
II. **SEMI - STRUCTURED DATA**
III. **UN STRUCTURED DATA**

➢ STRUCTURED DATA

- Defined and Relational Data types.
- Uses Machine learning.
- Not flexible.
- Requires less storage.
- Example : SQL

➢ SEMI - STRUCTURED DATA

- Semi defined and Semi relational Data types.
- Natural language processing.
- Semi Flexible
- Requires Medium amount of storage.
- Example : Email

➢ UN STRUCTURED DATA

- Undefined and non relational Data types.
- Natural language processing.
- Flexible.
- Requires a lot of Storage.
- Example : JPEG, PDF

➢ QUANTITATIVE DATA & QUALITATIVE DATA

QUANTITATIVE DATA

- Quantitative = Quantity
- Measures of values or counts
- Data about numeric Variables
- Examples : Age, Height, Weight

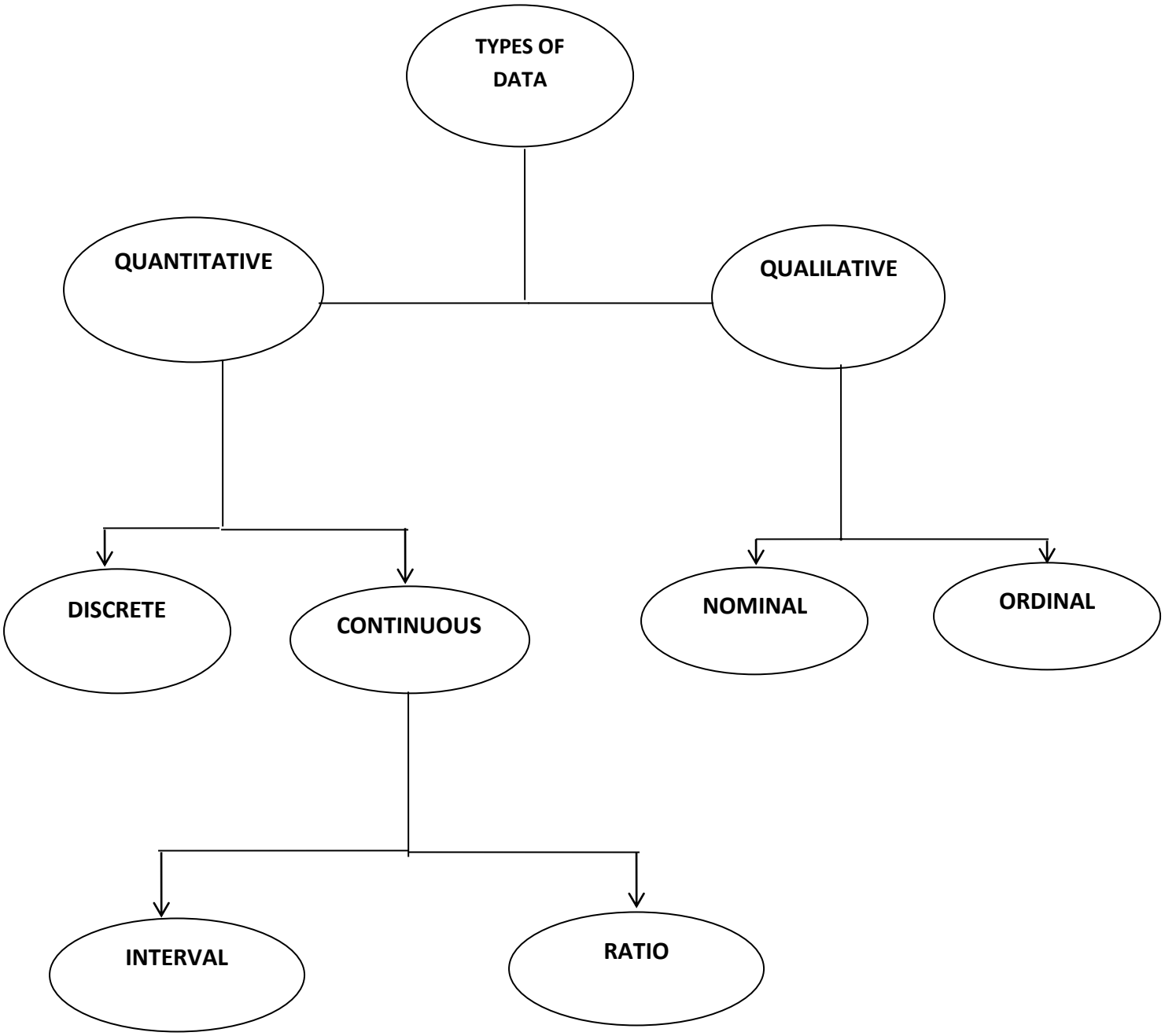
QUALITATIVE DATA

- Qualitative = Quality
- Measures of types
- Represent by name, symbol
- Examples : Gender, Religion, Qualifications

➢ DIFFERENT V's IN BIG DATA

- Value
- Volume
- Variety
- Velocity
- Veracity

➢ DIFFERENT TYPES OF DATA



➢ DISCRETE

- Means Discrete or separate.
- Data contain the values fall under whole numbers or integers.
- Can't be broken into Decimal or fraction values.
- Examples : No of students in a class, Cost of cell phone

➢ CONTINUOUS

- Data is in the form of fractional numbers.
- Data represent information that can be divided into smaller levels.
- Can take any value within a range.
- Examples: Temperature, Height, Weight

➢ NOMINAL

- Used to label variables without any order.
- Can't do any numerical task or can't give any order to sort the data.
- These data don't have any meaningful order.
- Examples: Nationality (Indian, American, Japan)

➢ ORDINAL

- Data have natural ordering where number is represented in some kind of order.
- Ordinal data only shows the sequences and can't use for statistical analysis.
- Ordinal data have some kind of order that is not present in nominal data.
- Examples: Grades in exam (A, B, C, D)

➢ INTERVAL

- Data appears in the form of number or numerical values.
- Measured along a scale in which each point is placed at equal distance from one another.
- Interval data also called as an integer
- Examples: Temperature measured in Fahrenheit and Celsius but not in Kelvin.

➢ RATIO

- Ratio data is a form of quantitative (numeric) data.
- Measures variables on a continuous scale.
- Ratio data can be used to calculate measures including frequency distribution.
- Examples: Height (5ft.8in, 5 ft.7in)

➢ TOOLS USED IN BIG DATA

- Apache Spark
- Apache Hadoop
- Apache Flink
- Google cloud Platform
- MongoDB