DESCRIPTIVE STATISTICS

INTRODUCTION

prepared by:

Gyro A. Madrona
Electronics Engineer

TOPIC OUTLINE

Descriptive Statistics

Types of Data

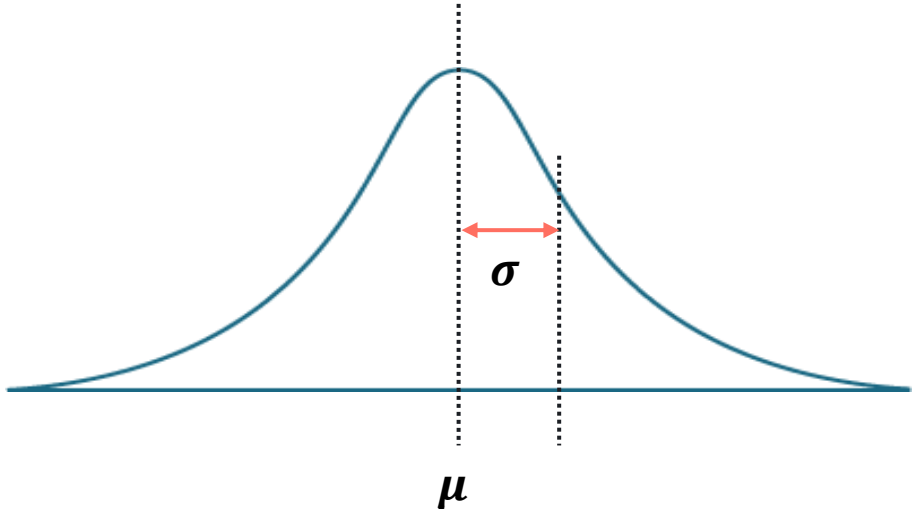
Levels of Measurement

Software Tools

DESCRIPTIVE STATISTICS

DESCRIPTIVE STATISTICS

Normal Distribution



Descriptive statistics is a branch of statistics that summarizes and organizes data to provide meaningful insights. It focuses on presenting raw data in a more understandable form using measures of **central tendency**, measures of **variability**, and **graphical** representations.

POPULATION VS SAMPLE

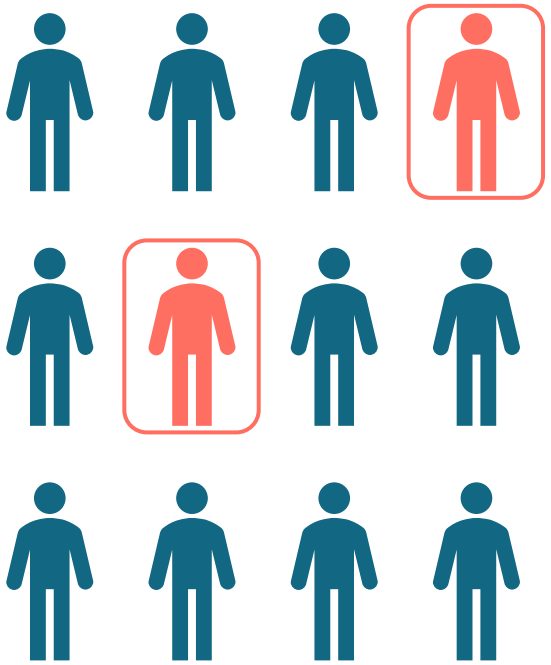
Population (N) refers to the **entire** group of individuals or observations being studied.

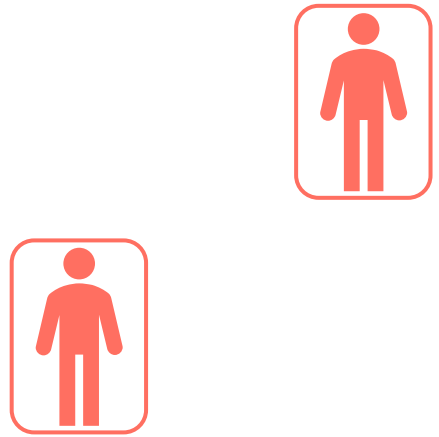
Sample (n) refers to **subset** of the population selected for analysis.

POPULATION VS SAMPLE

Population (N) refers to the **entire** group of individuals or observations being studied.

Sample (n) refers to **subset** of the population selected for analysis.



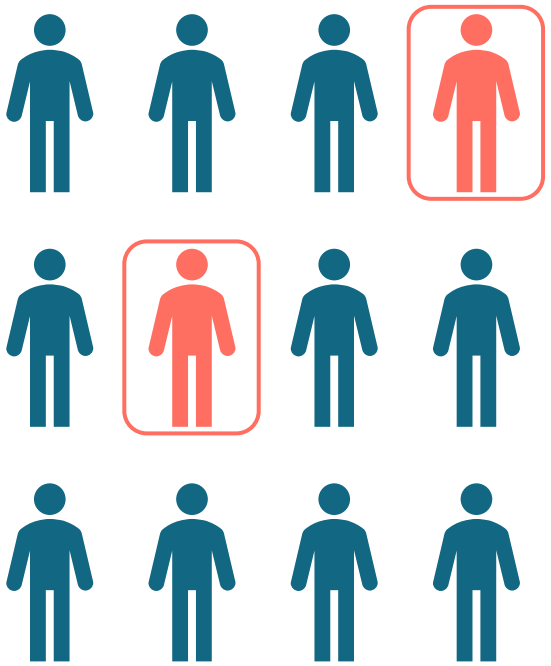


PARAMETERS VS STATISTICS

Characteristics of a **population** are called **parameters**.

μ – population mean

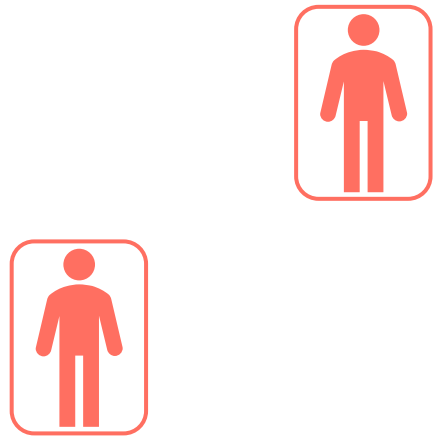
σ – population standard deviation



Characteristics of a **sample** are called **statistics**.

$\bar{x}$ – sample mean

s – sample standard deviation



TYPES OF DATA

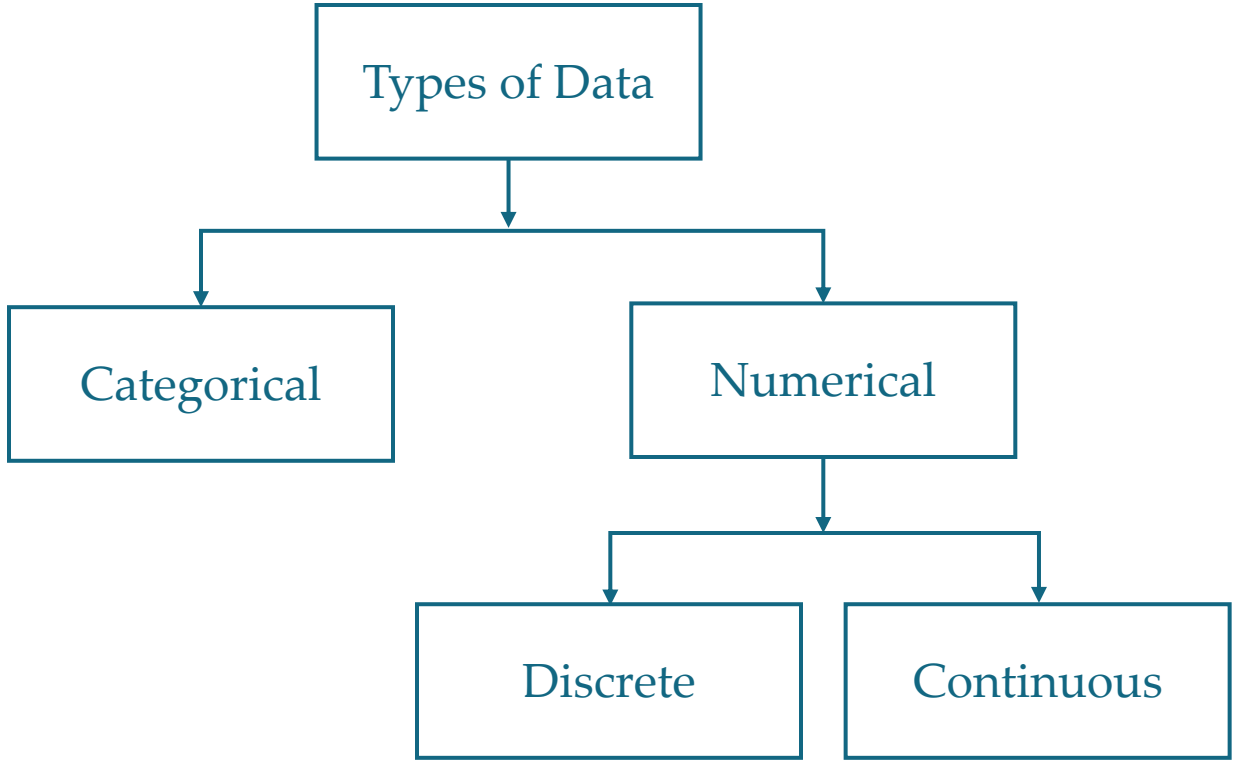
DATA

Data refers to the raw facts, figures, and **information** collected from various sources that can be analyzed to extract insights, identify patterns, and support decision-making.

Voltage Response

Measurement No.	Instrument A	Instrument B
1	12	2.8
2	5	4.5
3	9.1	6
4	3.3	9
5	24	11.7
6	18.5	14.8
7	15.2	17.3
8		20

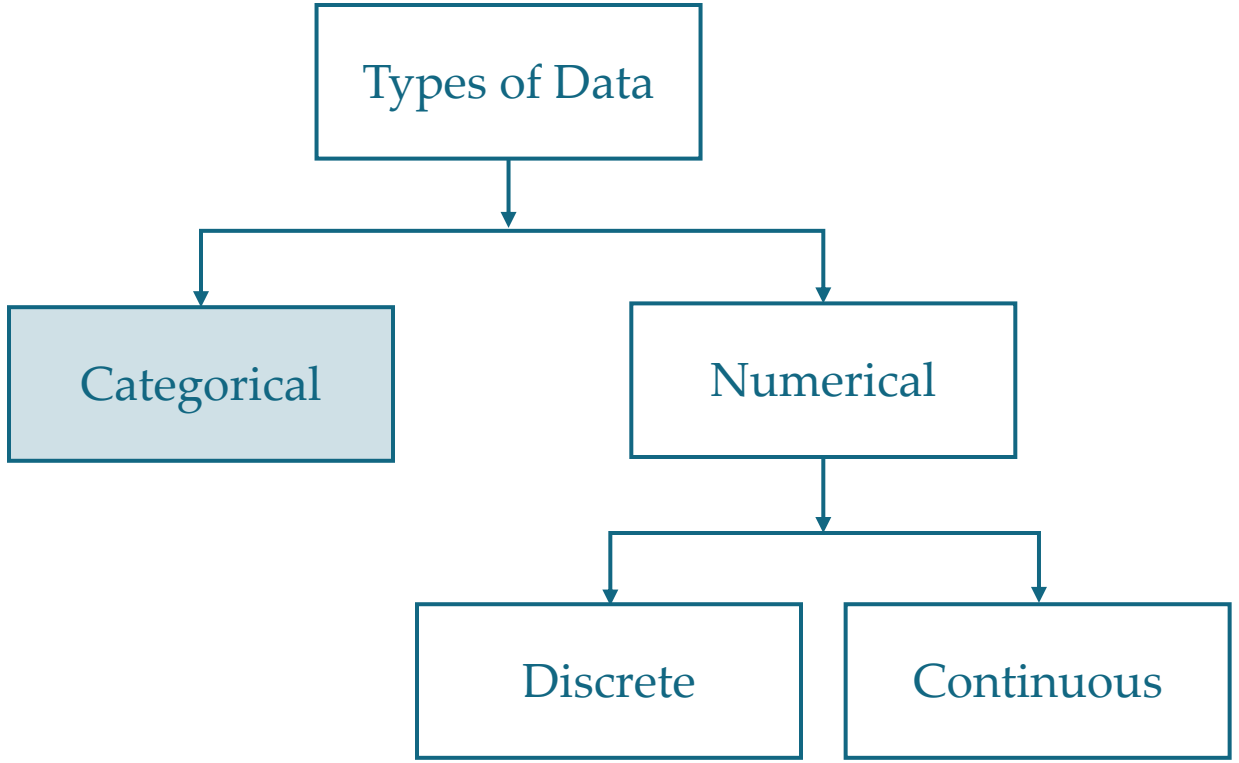
TYPES OF DATA



This classification focuses on the nature of the data and how it can be **categorized** or **measured**.

CATEGORICAL

Categorical data represents groups or categories.



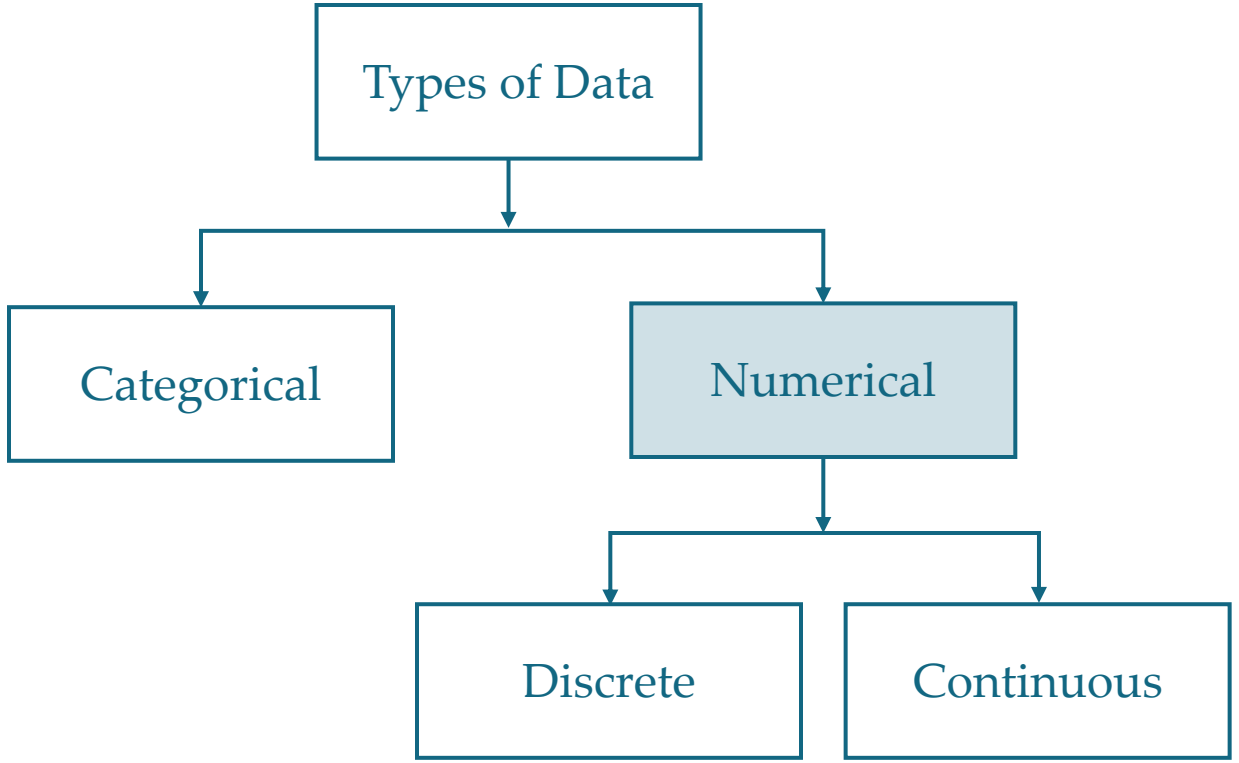
example

Gender (Male, Female)

Colors (Red, Blue, Green)

Types of Cars (Sedan, SUV, Truck)

NUMERICAL

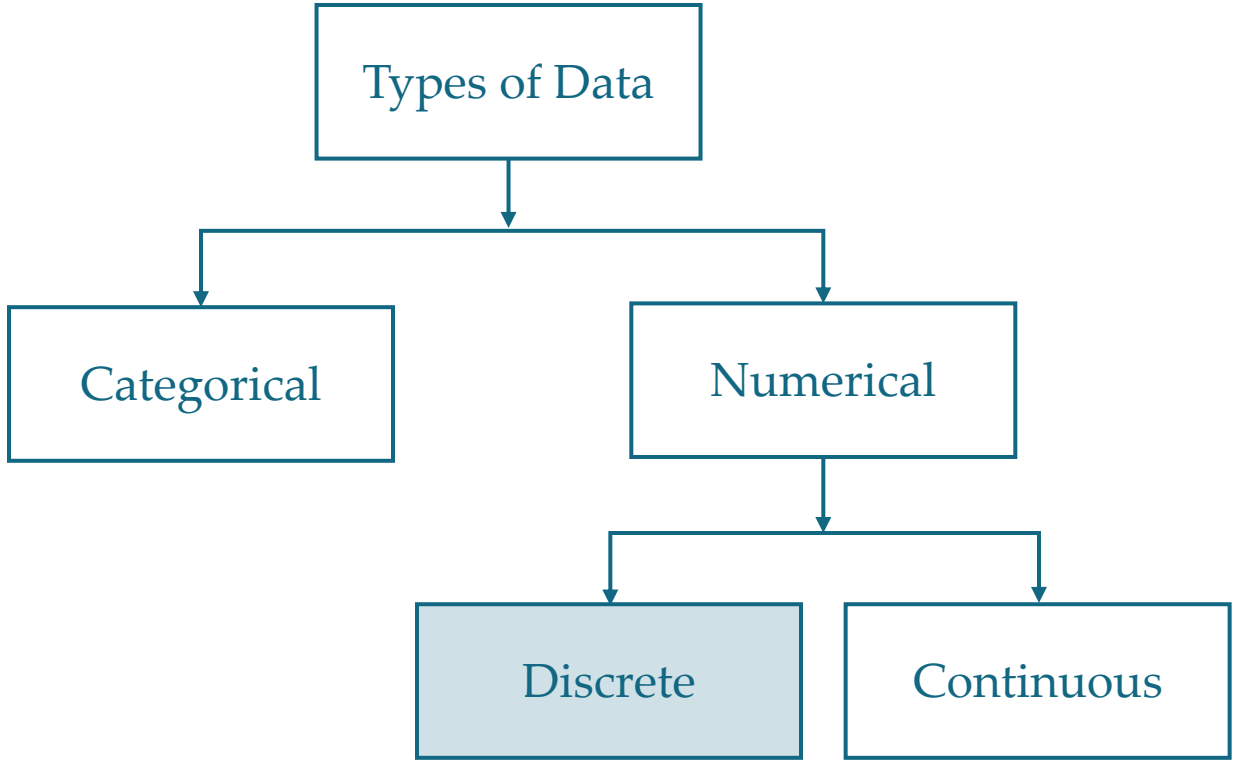


Numerical data represents quantities or numbers.

DISCRETE

Discrete data are data that can be counted in **finite** manner.

"You can imagine each member of the dataset."



example

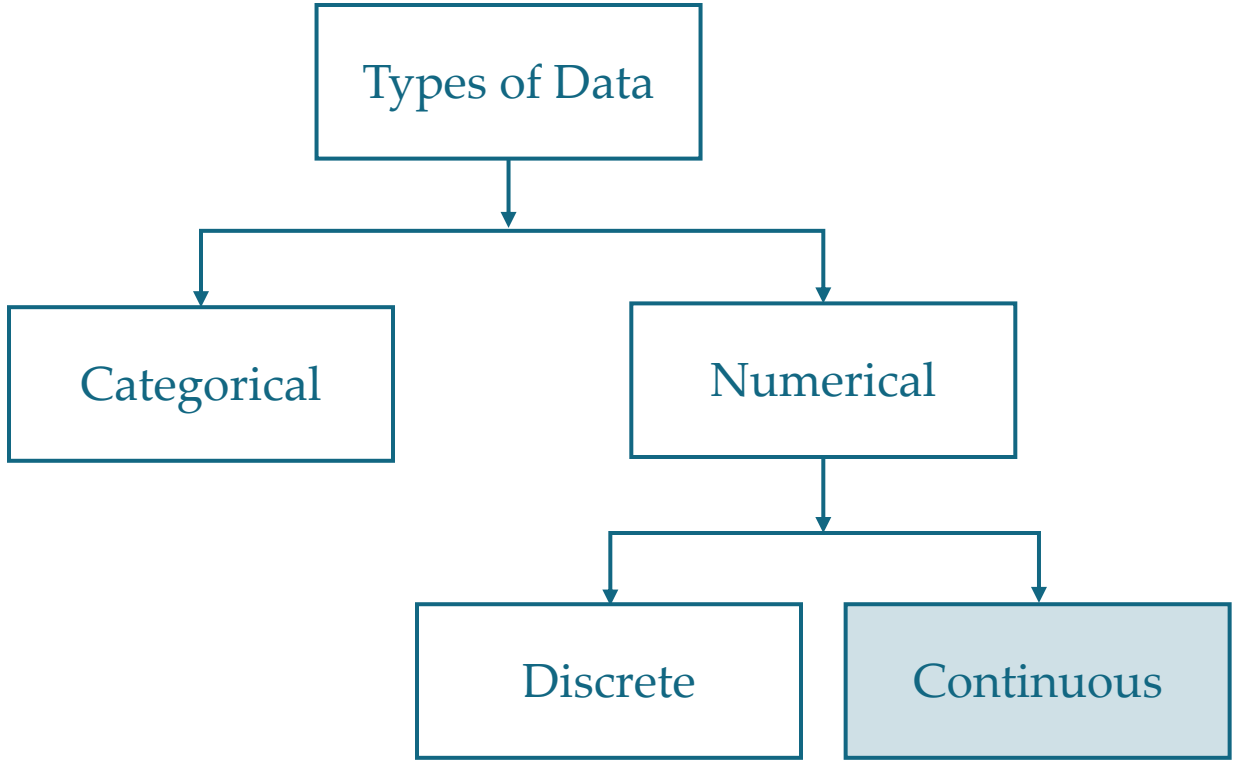
Number of students in a class

Number of correct answers

Count of defective products

CONTINUOUS

Continuous data can vary by infinitely small amounts, making them impossible to count directly. Measuring continuous data requires an **instrument** (e.g., ruler, thermometer) to obtain precise values.



example

Weight

Temperature

Voltage

EXERCISE

Determine if the given example is **discrete** or **continuous** data.

Grades

Area

Number of objects

Time

Time on the clock

Money

Physical Money

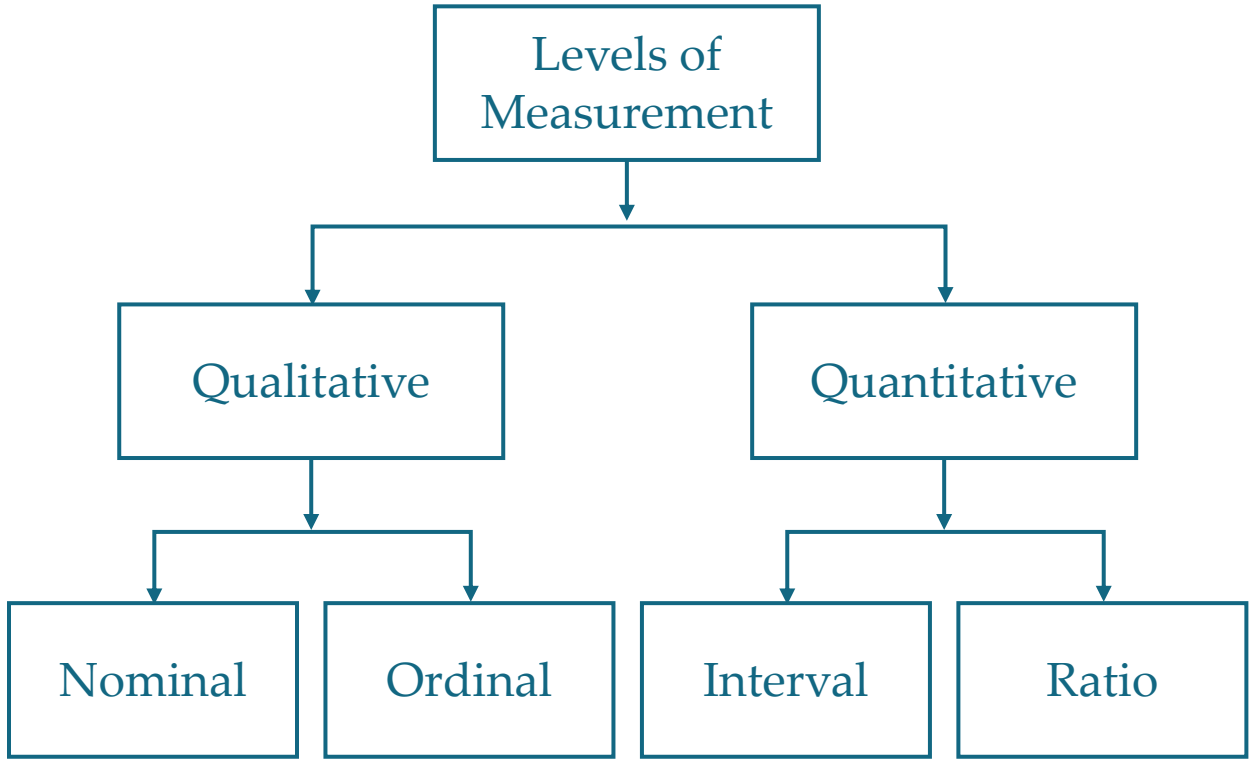
Height

Distance

LEVELS OF MEASUREMENT

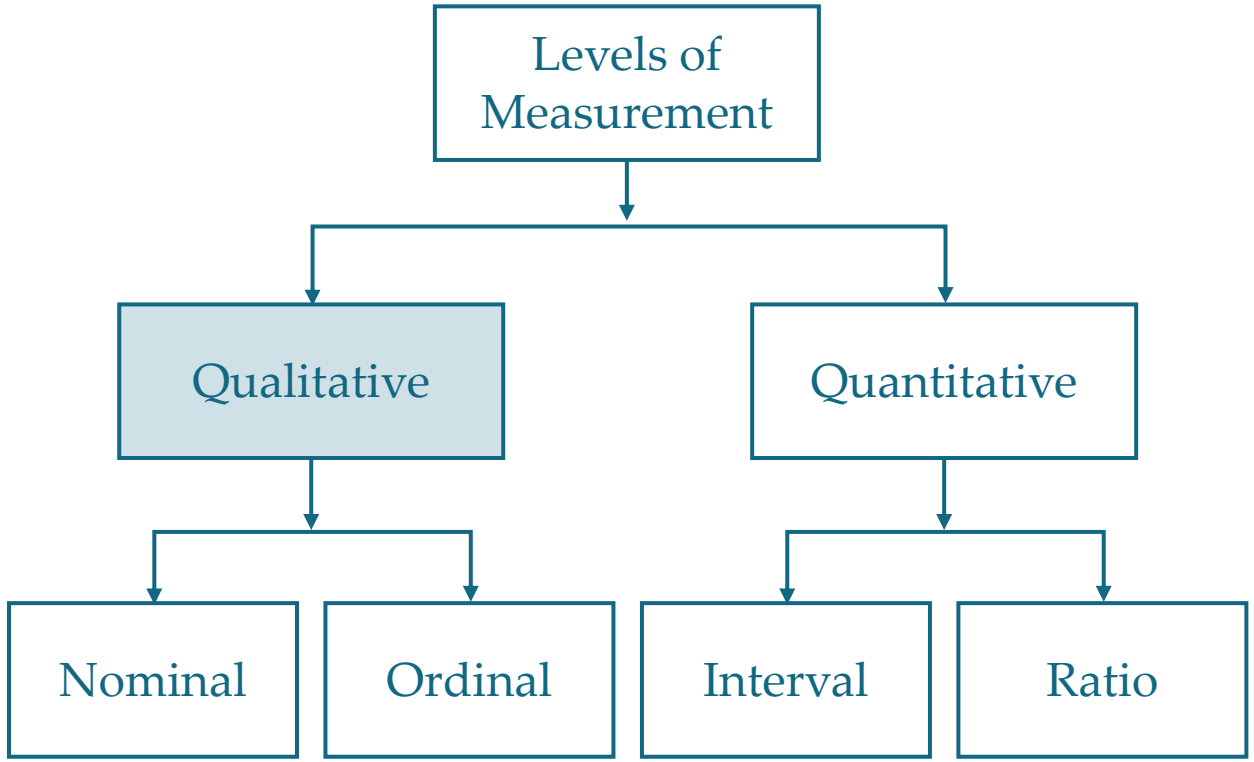
LEVELS OF MEASUREMENT

This classification focuses on the properties of the data and how it can be **analyzed mathematically**.



QUALITATIVE

Qualitative data refers to **non-numerical** information that describes characteristics or attributes (e.g., customer feedback). It is **subjective** in nature, as it is derived from personal perceptions, opinions, and interpretations.



NOMINAL

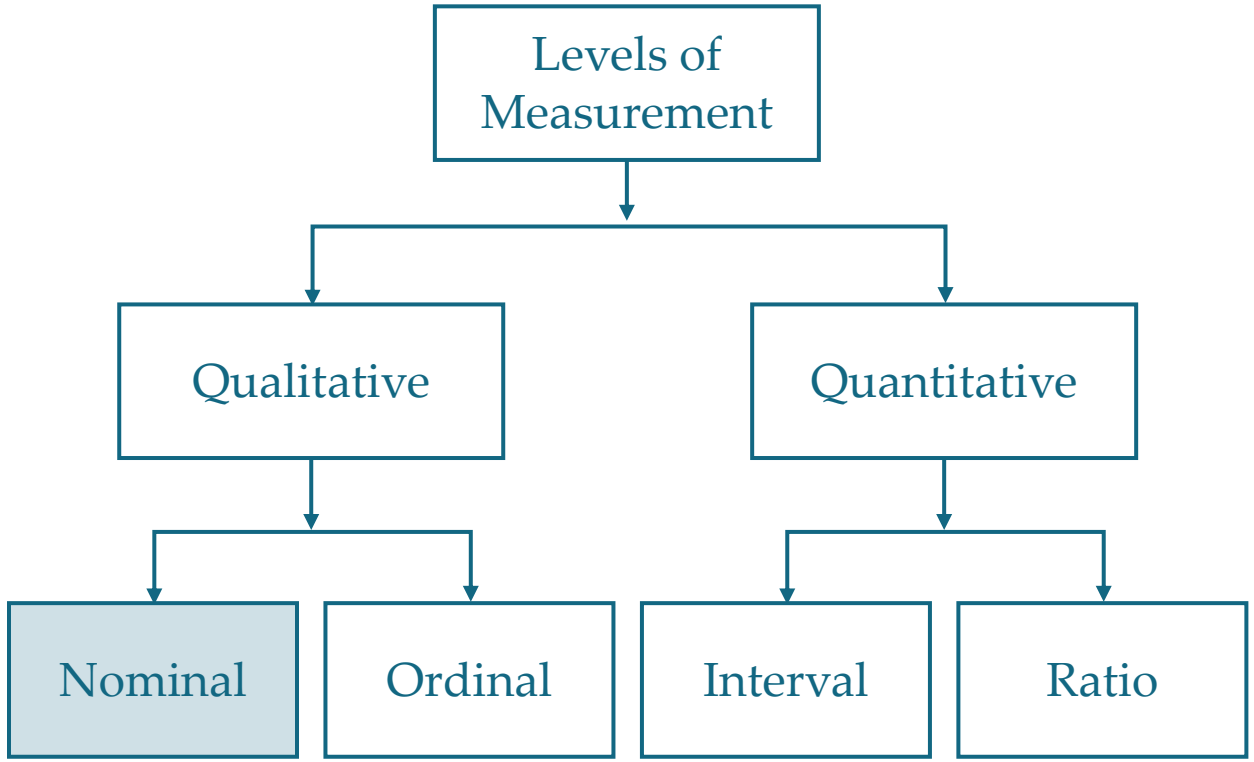
Nominal data can be categorized without any order or ranking.

example

Gender (Male, Female)

Color code (Red, Green, Blue)

Types of Cars (Sedan, SUV, Truck)



ORDINAL

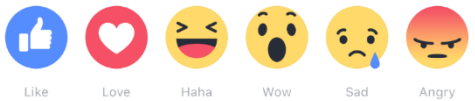
Ordinal data can be categorized with a **meaningful** order or ranking.

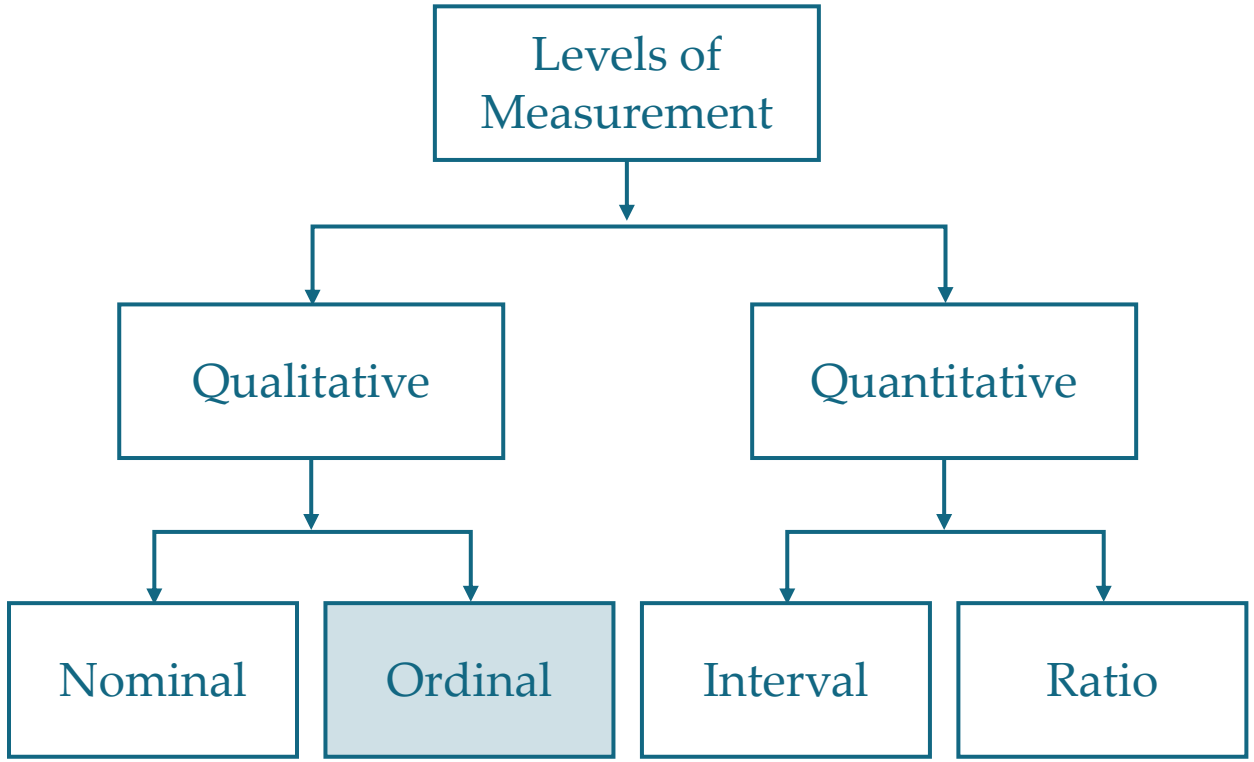
example

Customer feedback:

★★☆☆☆

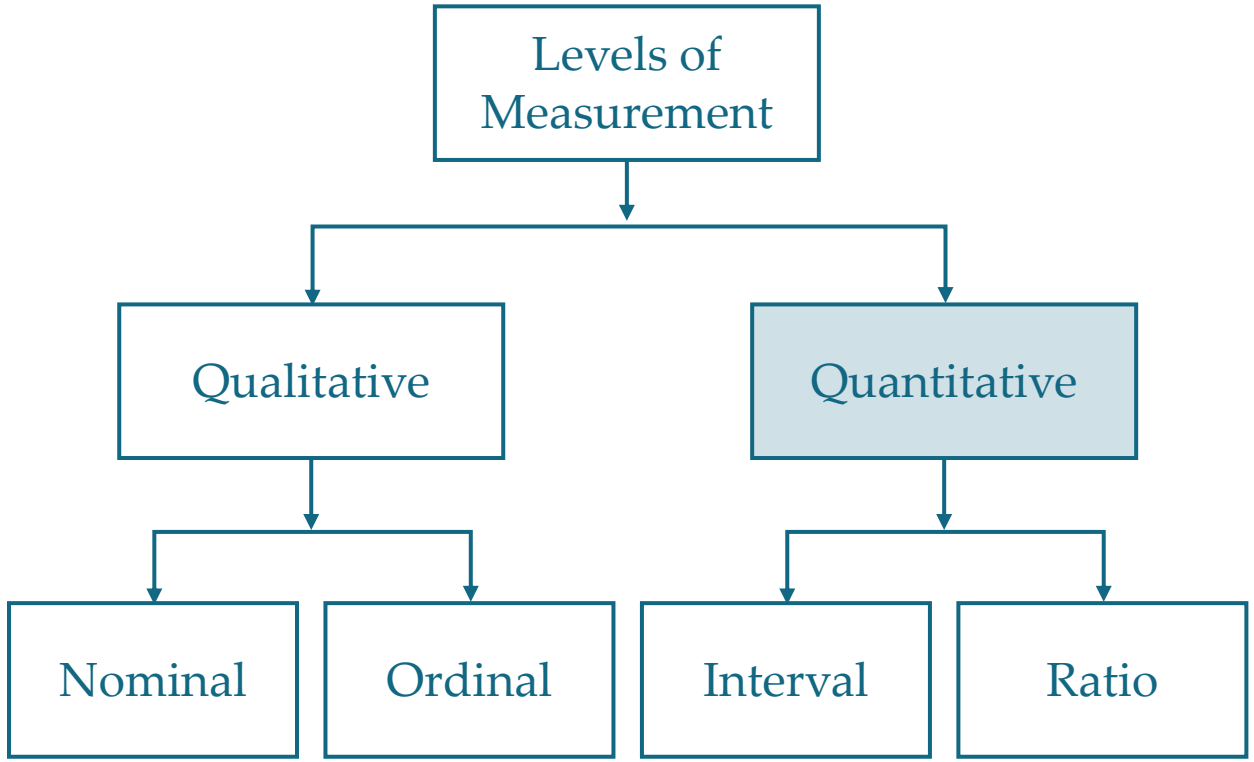
Facebook reactions:





QUANTITATIVE

Quantitative data refers to **numerical** data that can be measured and analyzed statistically (e.g., age, income, temperature).

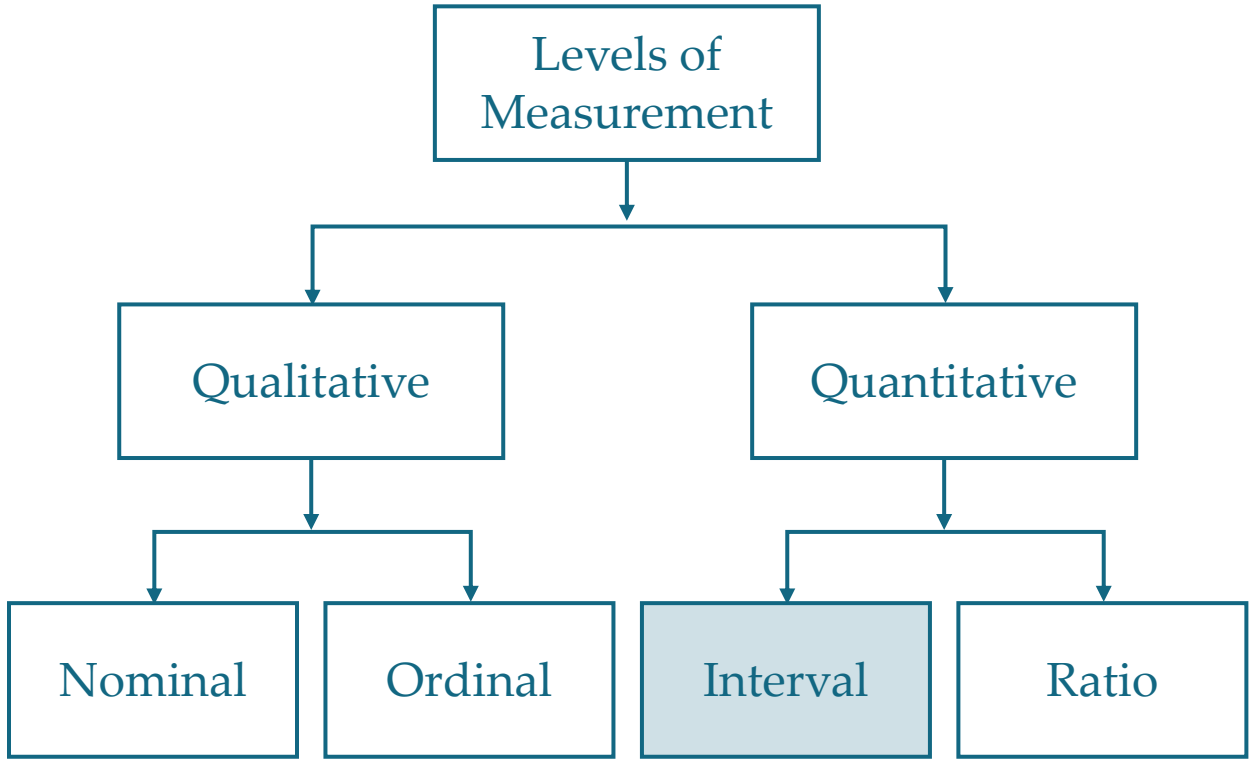


INTERVAL

Interval data has **no true zero**, meaning that the value of zero does not indicate the absence of the measured quantity.

example

Temperature in Celsius or Fahrenheit — 0°C or 0°F does not mean there is no temperature; it is just a reference point.



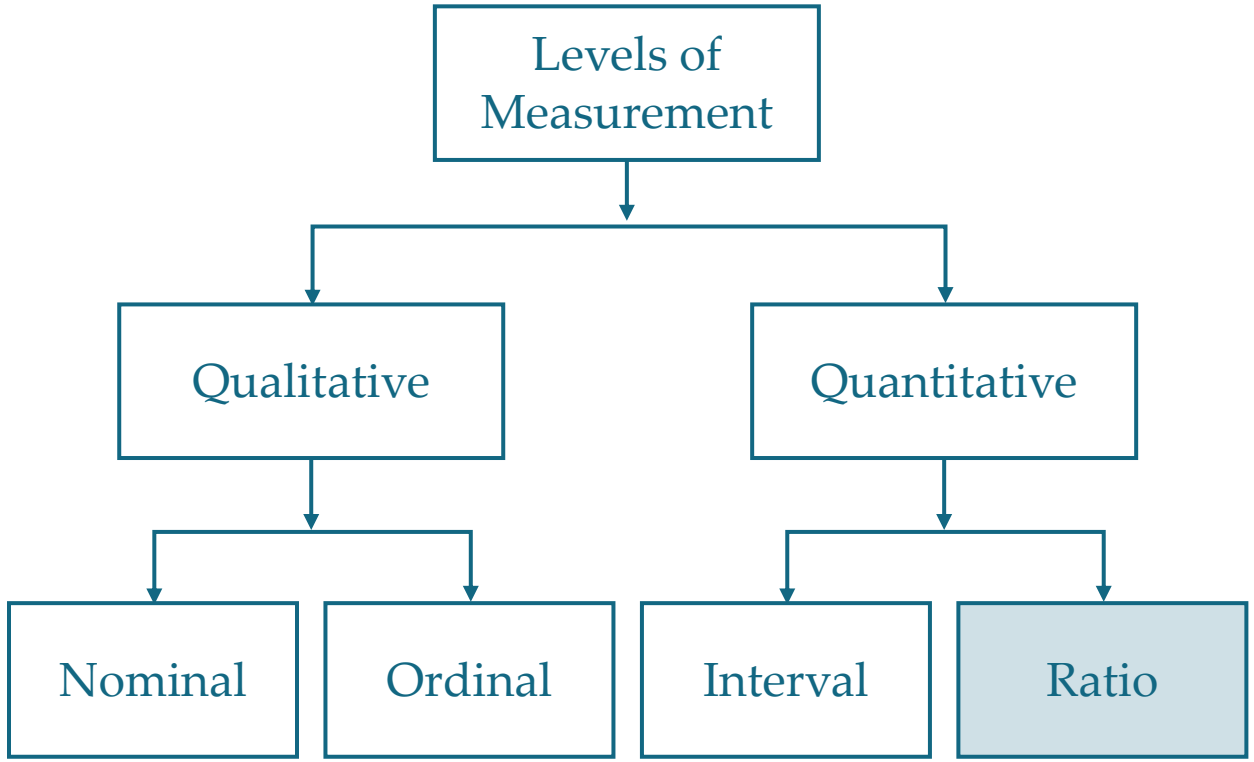
RATIO

Ratio data has a **true zero**, meaning that the value of zero represents the complete absence of the measured quantity.

example

Weight, if an object has a weight of 0 kg, it means it has no weight.

Temperature in Kelvin − 0°K represents absolute zero, where molecular motion stops.



EXERCISE

Determine if the given example is **interval** or **ratio** data.

Height of a building (m)

Time to complete a race (seconds)

Exam scores (out of 100)

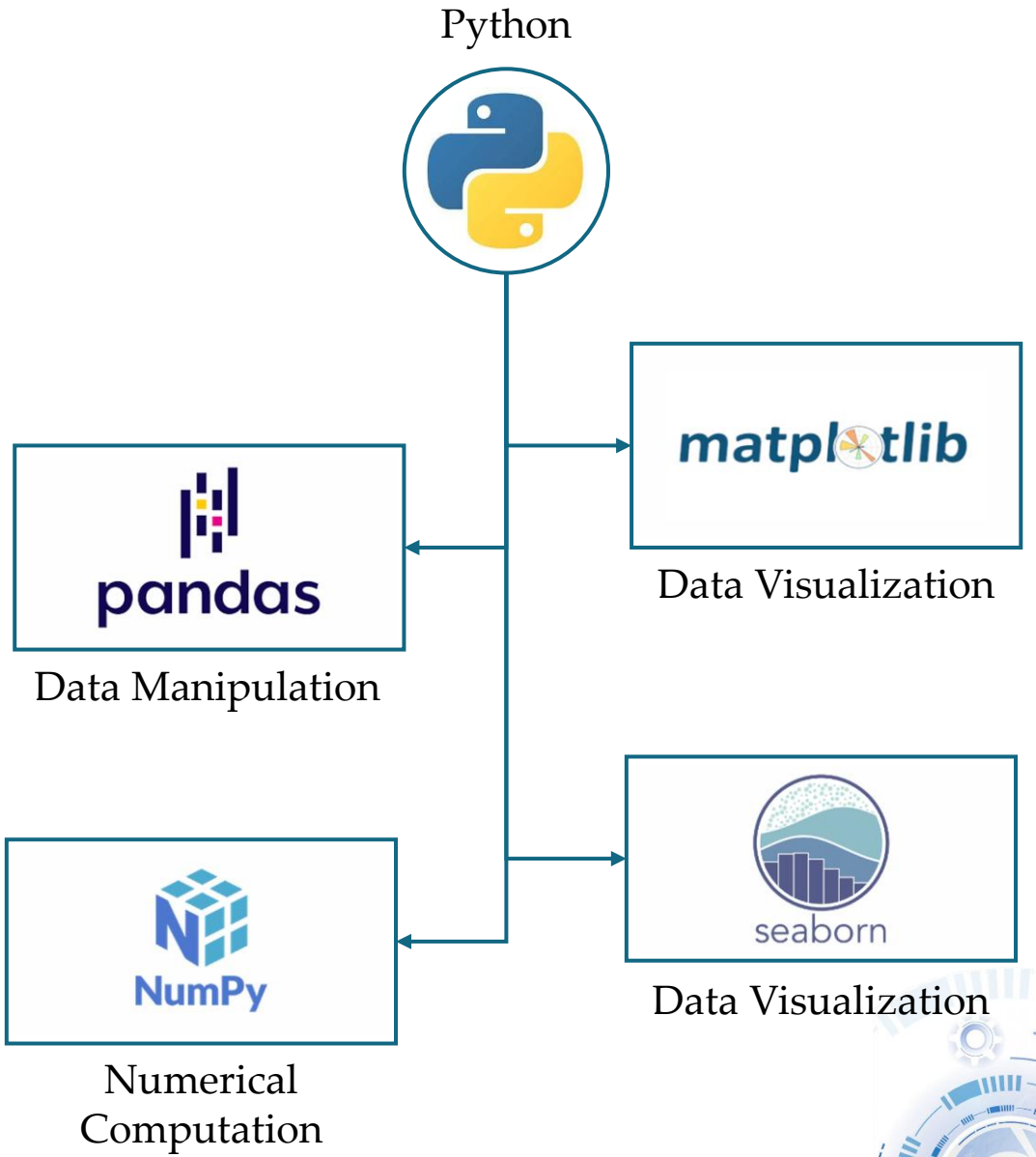
Year of historical event (1776, 1945, 2001)

Altitude above sea level (m)

SOFTWARE TOOLS

PYTHON

Python is a powerful and versatile programming language widely used in data analytics due to its simplicity, readability, and **extensive library support**. It enables data analysts to perform various tasks, from data cleaning and manipulation to statistical analysis and visualization.



https://www.python.org/

SOFTWARE TOOLS

1. **Visual Studio Code**

   https://code.visualstudio.com/Download

2. **Python**

   https://www.python.org/downloads/

3. Jupyter Notebook Extension on VS Code

4. numpy, pandas, scipy, matplotlib libraries

5. **Github Account**

   https://github.com/

6. **Git Bash**

   https://git-scm.com/downloads/win

LABORATORY