# Day_02_Data Representation

---

## 🌟 Story to Begin With

Imagine you are a **class monitor** 👩‍🏫 in a class of **50 students**.

- There are **20 girls** and **30 boys**.

- You want to show this information to your teacher in a **clear way**.

But how do you represent the data?

👉 That's where **Statistics** helps us!

---

## 🗂️ Data Types

1. **Categorical Data** 🎭

   - Data in **labels or categories**.

   - Example: Gender, Blood Group, Movie Genre.
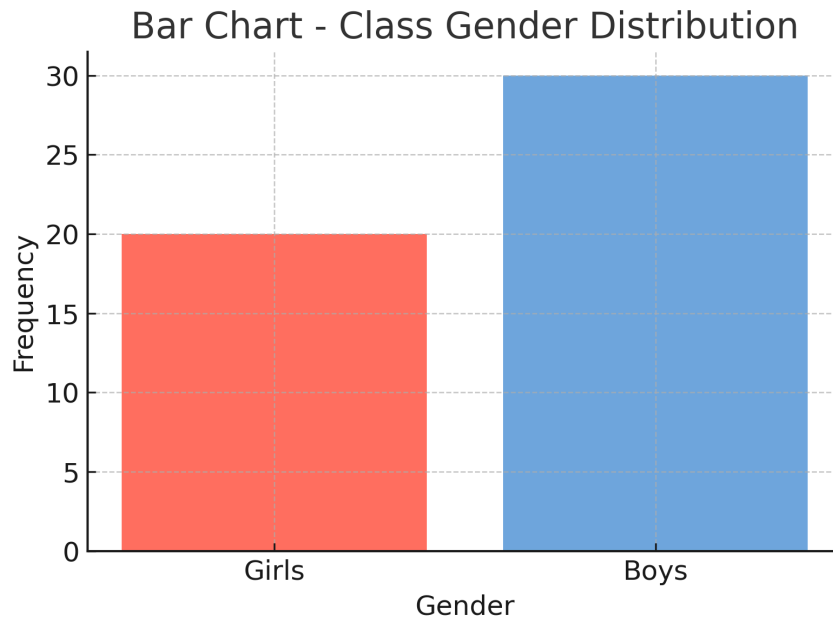
2. **Numerical Data** 🔢

   - Data in **numbers**.

   - Example: Marks, Age, Height.

---

## 📋 Representation of Categorical Data

### 1. Frequency Table

Let's organize the class data 👇

| Gender | Frequency |
|--------|-----------|
| Boys   | 30        |
| Girls  | 20        |

Bar Chart - Class Gender Distribution

## 2. Relative Frequency (Percentage Form)

We can also represent the data in **fractions or percentages**.

$$\text{Relative Frequency} = \frac{\text{Class Frequency}}{\text{Total Students}}$$
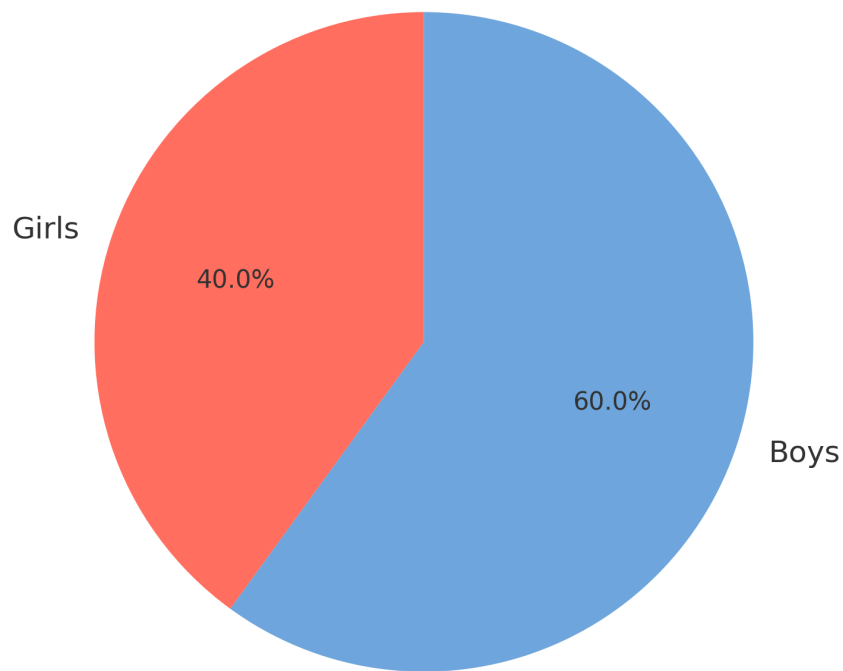
- Girls:

$$\frac{20}{50} = 0.4 = 40\%$$

- Boys:

$$\frac{30}{50} = 0.6 = 60\%$$

| Class Name | Class Frequency | Relative Frequency |
|------------|-----------------|--------------------|
| Girls      | 20              | 0.4 (40%)          |
| Boys       | 30              | 0.6 (60%)          |

Pie Chart - Class Gender Distribution

---

## 3. Graphical Representation

- **Bar Chart / Bar Graph** 📊
  - X-axis → Class names (Boys, Girls)
  - Y-axis → Frequency
- **Pie Chart** 🥧
  - Girls = 40% of the circle
  - Boys = 60% of the circle

👉 These visuals help us **understand data faster than raw numbers**.

---

# 📊 Representation of Numerical Data

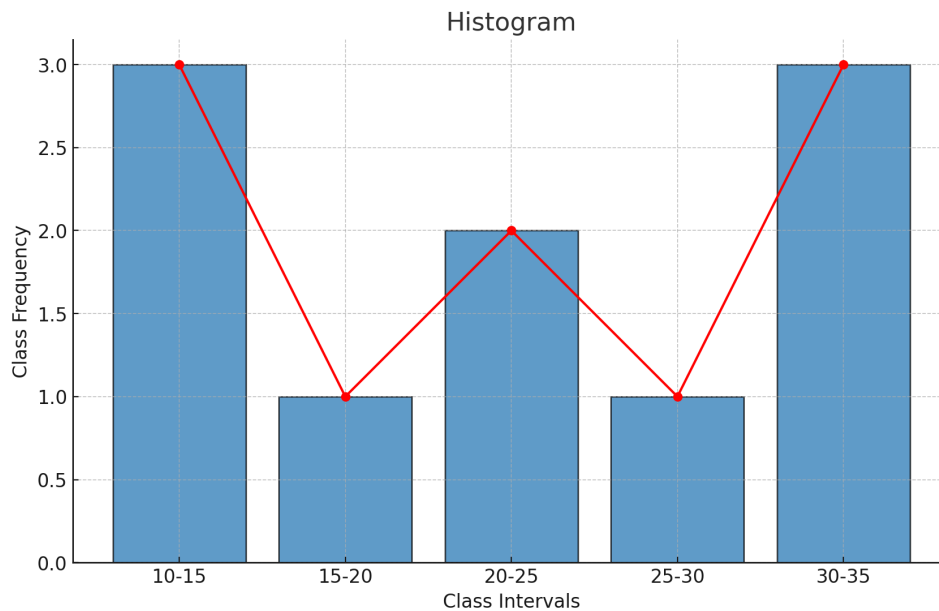Now suppose we collect the **marks of 10 students**:

**Marks →** 15, 10, 21, 35, 27, 12, 31, 24, 18, 33

Raw numbers are hard to read 😵, so we group them into **intervals (ranges)**.

## 1. Frequency Distribution Table

| Class Interval | Class Frequency |
| --- | --- |
| 10 − 15 | 3 |
| 15 − 20 | 1 |
| 20 − 25 | 2 |
| 25 − 30 | 1 |
| 30 − 35 | 3 |

👉 This table shows **how many students fall into each range of marks**.



## 2. Why Do We Use Intervals?

- Easier to **summarize large data**

- Quickly see **patterns & trends**

- Example: Most students here scored **between 30–35 marks**

## 3. Graphical Representation of Numerical Data

- **Histogram** 📈 (bars represent intervals)

- **Frequency Polygon** 📈 (line graph joining class frequencies)

---

## 🎯 Key Takeaways

- **Categorical Data** → Represent with **frequency tables, bar graphs, pie charts**.

- **Numerical Data** → Represent with **frequency distribution tables, histograms, polygons**.

- **Relative frequency** helps in comparing percentages.

- Grouping into **intervals** makes data more meaningful.