



# Understanding Data & Central Tendency

A beginner's guide to the building blocks of statistics. Let's explore data types and how to find the "center" of a dataset.

## Types of Data

In statistics, we primarily work with two types of data: Qualitative and Quantitative. Understanding the difference is the first step in any analysis.

### Qualitative Data (গুণবাচক ডেটা)

*Qualitative data describes qualities or characteristics. It's non-numerical and is often represented by names or labels. Think of it as **descriptive information**.*

**Key Idea:** You cannot perform mathematical operations like addition or subtraction on qualitative data.

#### Examples:

- Eye Color (Blue, Green, Brown)
- Favorite Food (Pizza, Burger, Pasta)
- Yes/No answers in a survey
- Car Brands (Toyota, Honda, Ford)

**Uses:** Great for understanding motivations, opinions, and experiences. Often used in social sciences and market research.

### Quantitative Data (সংখ্যাবাচক ডেটা)

*Quantitative data deals with numbers and things you can measure objectively. Think of it as **numerical information**.*

**Key Idea:** You can perform mathematical operations on this data. It has two sub-types:

1. **Discrete:** The data can only take certain values (like whole numbers). You can count it.
2. **Continuous:** The data can take any value within a range. You measure it.

**Examples:**

- **Discrete:** Number of students in a class (e.g., 30), number of cars in a parking lot (e.g., 50).
- **Continuous:** A person's height (e.g., 5.75 feet), the temperature of a room (e.g., 25.5°C).

**Uses:** Used in scientific experiments, financial reports, and any field where precise measurement is needed.

## Measures of Central Tendency

Central tendency is a single value that attempts to describe the center of a set of data. The three main measures are Mean, Median, and Mode.

### 1. Mean (গড়)

*The **Mean** is the most common measure of central tendency. It's simply the **average** of all the numbers in a dataset.*

**How to find it:** Add up all the numbers and then divide by how many numbers there are.

Mean =

**Example:** Find the mean of test scores {80, 85, 95, 70, 100}

**Sum:**  $80 + 85 + 95 + 70 + 100 = 430$

**Count:** There are 5 numbers.

**Mean:**  $430 / 5 = 86$

**Uses:** Very common and easy to calculate. However, it is sensitive to **outliers** (extremely high or low values).

## 2. Median (মধ্যক)

*The **Median** is the **middle value** in a dataset that has been arranged in order from smallest to largest.*

**How to find it:**

1. Arrange the data in ascending order.
2. If there's an **odd** number of values, the median is the middle one.
3. If there's an **even** number of values, it's the average of the two middle values.

**Example 1 (Odd count):** {7, 3, 9, 2, 5}

**Sorted:** {2, 3, 7, 9, 5} → Median is 7.

**Example 2 (Even count):** {10, 4, 8, 2}

**Sorted:** {2, 4, 8, 10} → Average:  $(4 + 8) / 2 = 6$ .

**Uses:** Excellent when your data has outliers, because it isn't affected by extreme values.

## 3. Mode (প্রচুরক)

The **Mode** is the value that appears **most frequently** in a dataset.

**How to find it:** A dataset can have one mode (unimodal), more than one mode (multimodal), or no mode at all.

Example 1: {🍏, 🍌, 🍏, 🍊, 🍇, 🍏}

The mode is 🍏 (appears 3 times).

Example 2: {5, 2, 8, 5, 9, 10, 8}

The modes are **5 and 8** (bimodal).

**Uses:** It's the only measure that can be used for **qualitative data**. Useful for finding the most popular item.

## Quick Summary

MEASURE	WHAT IT IS	WHEN TO USE
Mean	The average value	For numerical data with no major outliers.
Median	The middle value	For numerical data, especially when there are outliers.
Mode	The most frequent value	For categorical (qualitative) data or to find the most popular item.