Central Tendency

Mean (Average)

The mean, often called the average, is a measure of central tendency that represents the typical value in a set of numbers. It is calculated by adding all the numbers in a data set and then dividing by the total number of values.

Formula: $$\text{Mean} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

- $x_1 + x_2 + \ldots + x_n$: represents each individual value in the data set.
- n : the total number of values.

Example:

Suppose we have the numbers: 4, 7, 10, 13, 16

1. Add them up: 4 + 7 + 10 + 13 + 16 = 50
2. Divide by the total count (5 numbers): 50 ÷ 5 = 10

So, the mean is 10.

$$10 = \frac{4 + 7 + 10 + 13 + 16}{5}$$

Why Use the Mean?

The mean is useful for understanding the overall trend of a data set, but it can be influenced by extreme values (outliers). In such cases, other measures like the median might be more appropriate.

Median (Middle Value)

The median is the middle value of a data set when the numbers are arranged in ascending order. It is another measure of central tendency, often used when a data set has extreme values (outliers) that might distort the mean.

How to Find the Median:

1. Arrange the numbers in order (smallest to largest).
2. If there is an odd number of values, the median is the middle number.
3. If there is an even number of values, the median is the average of the two middle numbers.

Example 1 (Odd number of values):

Data set: 3, 7, 9, 12, 15

- The numbers are already in order.
- The middle number is 9, so the median is 9.

Example 2 (Even number of values):

Data set: 2, 5, 8, 11, 14, 18

- The middle numbers are 8 and 11.
- The median is the average: (8 + 11) ÷ 2 = 9.5

Why Use the Median?

- The median is not affected by extreme values (unlike the mean).
- It gives a better measure of central tendency when a data set has outliers or is skewed.

Central Tendency

Central tendency refers to statistical measures that identify the center or typical value of a data set. It helps summarize a large amount of data by providing a single value that represents the "average" or most typical data point.

Mode (Most Frequent Value)

The mode is the value that appears most frequently in a data set. Unlike the mean and median, which describe central tendency in terms of averages or middle values, the mode simply identifies the most common number.

How to Find the Mode:

1. List all the numbers in the data set.
2. Identify which number(s) appear most frequently.
3. If one number appears the most, it is the mode.
4. If multiple numbers appear with the same highest frequency, the data set is bimodal (two modes) or multimodal (more than two modes).
5. If no number repeats, there is no mode.

Example 1 (Single Mode - Unimodal):

Data set: 2, 3, 4, 4, 5, 6, 7

- The number 4 appears twice, more than any other number.
- Mode = 4

Example 2 (Two Modes - Bimodal):

Data set: 1, 2, 3, 3, 4, 5, 5, 6

- The numbers 3 and 5 both appear twice (most frequently).
- Modes = 3 and 5 (Bimodal)

Example 3 (No Mode):

Data set: 1, 2, 3, 4, 5, 6, 7

- No number repeats.
- No mode.

Why Use the Median?

- The mode is useful for categorical data (e.g., most popular car color, favorite sport, etc.).
- It helps understand frequency distributions in real-world scenarios (e.g., most common shoe size sold).