Descriptive Statistics: Measures of Central Tendency

Adejumo Ridwan Suleiman

Descriptive Statistics and Measures of Central Tendency.

Descriptive statistics

- Branch of statistics that involves summarizing, organizing and presenting data in a meaningful way.
- Main objective is to provide a clear and concise summary of a data.
- Data can be described either numerically or graphically.

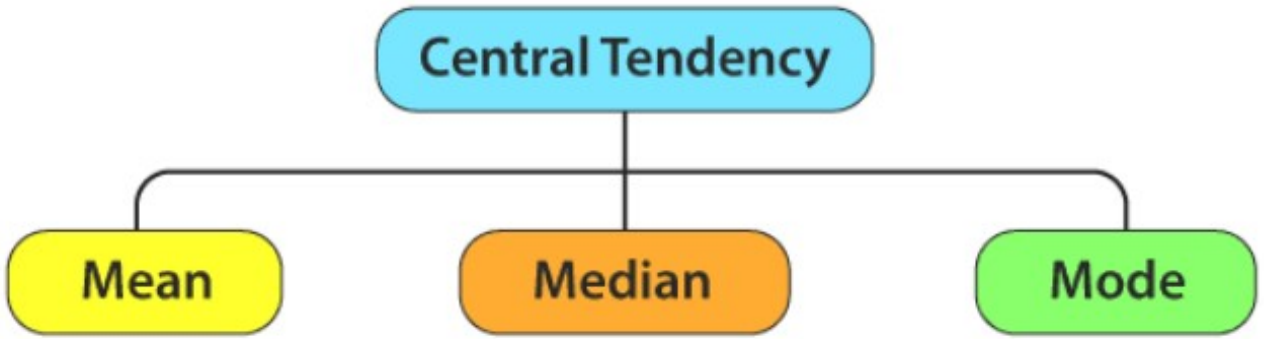
Type of Descriptive Statistics

- **Measures of Central Tendency:** Mean, Median, Mode
- **Measures of Variability:** Range, Variance, Standard Deviation, Inter-quartile range
- **Measures of Partition:** Quartiles, Percentiles, Deciles

Measures of Central Tendency

- Single value reflecting the center of a data distribution.
- Gives a summary of a variable.

Measures of Central Tendency



Source: BYJU

Mean

Definition

- Average of a given set of values
- Denoted mathematically as

Example

- The weight of 5 patients in a hospital is given as $56, 54, 79, 63, \text{and } 51$
- The average weight is the sum of all the weight divided by the number of patients in the hospital which is 60.6.

Merits of the Mean

- It is easy to calculate and understand.
- It is based upon all the observations of a variable.
- It is unique and can be used to compare different sets of data.

Demerits of the Mean

- Easily affected by extreme values
- Can’t be located graphically or by inspection
- Can only be used with numerical variables
- Can’t be used when there are presence of missing observations
- Easily affected by sample size, smaller sample size give a more accurate estimate of the true population mean.

Types of Mean

- Arithmetic Mean
- Geometric Mean
- Harmonic Mean

Arithmetic Mean

- Also known as average, is the ratio of the sum of all observations to the total number of observations.
- For example the average weight of patients in an hospital is 65kg

Geometric Mean

- n_{th} root of the product of all observations in a variable.
- Always less than the arithmetic mean

Harmonic mean

- Reciprocal of the arithmetic mean i.e reciprocal of the average of the reciprocals of the data values.
- Good at handling extreme values.

Example

The ages of children in a primary health care center is given as follows - $12, 10, 11, 8, 7$:

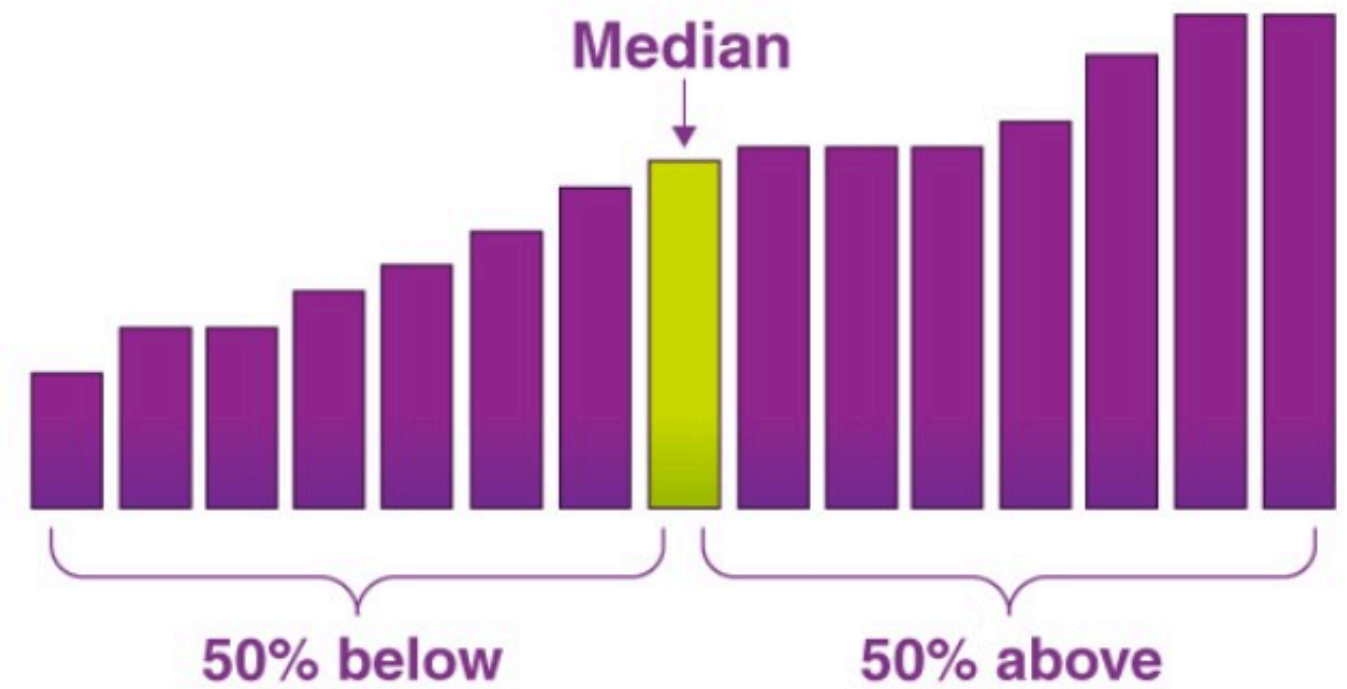
- The arithmetic mean is given as 9.6
- The geometric mean is given as 9.4
- The harmonic mean is given as 9.22

Median

Median

- The middle number when the data is sorted.
- Also used as an average in the presence of extreme values.
- A case when the number of observations is even, and the median values are two, take the average of the two numbers and divide by 2.

Median



Source: BYJU

Median - Example

- The age of women attending ANC were recorded as follows.
  $23, 19, 21, 20, 23, 21, 22, 24, 22, \text{ and } 22$
- **Step 1:** Arrange in ascending order.

  $19, 20, 21, 21, 22, 22, 22, 23, 23, 23, 24$
- **Step 2:** Pick the middle observation.

  The middle observation is 22 and 22, hence the median is
  $\frac{22+22}{2} = 22$

Mode

Mode

- The most occuring observation in a given set of data.
- Mode can either be **Bimodal**, **Trimodal**, or **MultiModal**.

Mode - Example

- If the temperature of 12 rabbits are $23, 19, 21, 20, 23, 21, 22, 24, 22,$ and 22.
- The mode is given as 22.

Skewness

Skewness - Definition

- Measure of the deviation of a given variable from a symmetric distribution.
- A data is symmetric if the Mean, Median and Mode are all equal.

Skewness - Types of Skewness

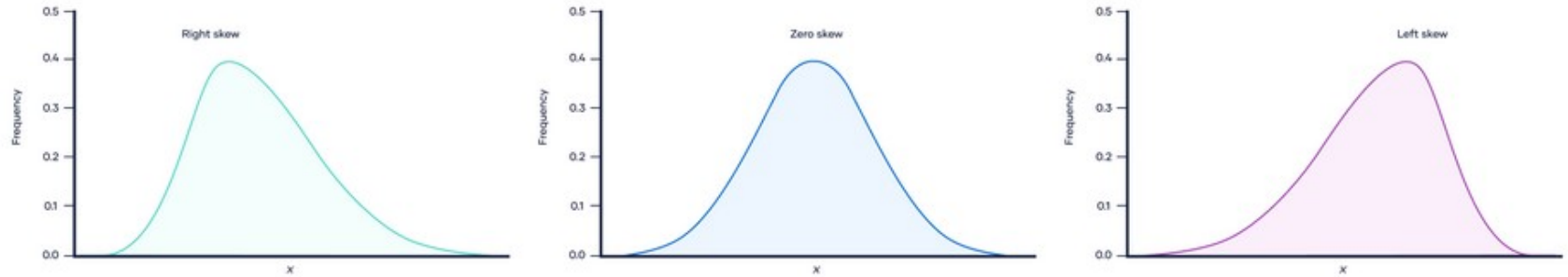
Positive Skewness

- Higher extreme values.
- Has a tail on the right
- Mean > Median > Mode

Negative Skewness

- Lower extreme values.
- Has a tail on the left.
- Mean < Median < Mode

Skewness



Source: Scribbr

Calculating and Interpreting Skewness

$$Skewness = \frac{\text{Mean} - \text{Median}}{StandardDeviation}$$

Interpretation

- Zero, if a perfect Symmetrical distribution.
- Negative, when negatively skewed or skewed to the left.
- Positive, when positively skewed or skewed to the right.

Comparison of Median and Mean

Median	Mean
Insensitive to extreme values	Sensitive to extreme values
Defines the central value of the data	Defines the midpoint of the data
Used with symmetric distributions	Used with skewed distributions
Affected by external factors	Robust and reliable