What is the mean, median and mode in the dataset?

Mean:

- **Definition:** The average of all values in the dataset, calculated by summing them and dividing by the number of values.
- **Pros:** Sensitive to changes in all values, good for normally distributed data.
- **Cons:** Can be skewed by outliers, not robust to non-numerical data.

Median:

- **Definition:** The "middle" value when the data is arranged in ascending order (or the average of middle two values for even-sized datasets).
- **Pros:** Not affected by outliers, robust to non-numerical data.
- **Cons:** Doesn't give insight into spread or distribution, less informative than mean for normally distributed data.

Mode:

- **Definition:** The most frequently occurring value in the dataset.
- Pros: Easy to understand, useful for categorical data.
- **Cons:** Can have multiple modes, doesn't provide information about overall distribution.

Key differences:

- **Mean:** Reflects the "center of mass" of the data.
- **Median:** Represents the "typical" value, unaffected by outliers.
- **Mode:** Shows the most common value, good for understanding dominant categories.
-

Choosing the right measure:

The best measure depends on your data and research questions.

- **Normal distribution:** Mean is generally preferred.
- **Outliers present:** Median is more robust.
- **Categorical data:** Mode might be helpful.