Task 4 Summary

Descriptive Statistics

In contrast with Inferential Statistics, Descriptive statistics use gathered real life data, as opposed to using sample data to draw conclusions and make inferences

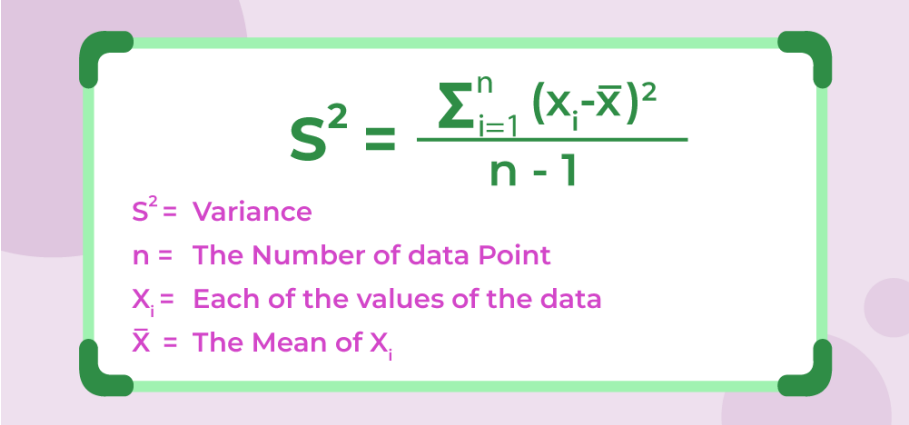
1. Organizing and summarizing data
2. Data Summary / Shape of Graph and Skewness
3. Measures of Central Tendency
4. Measures of Variability

Measures of Central Tendency

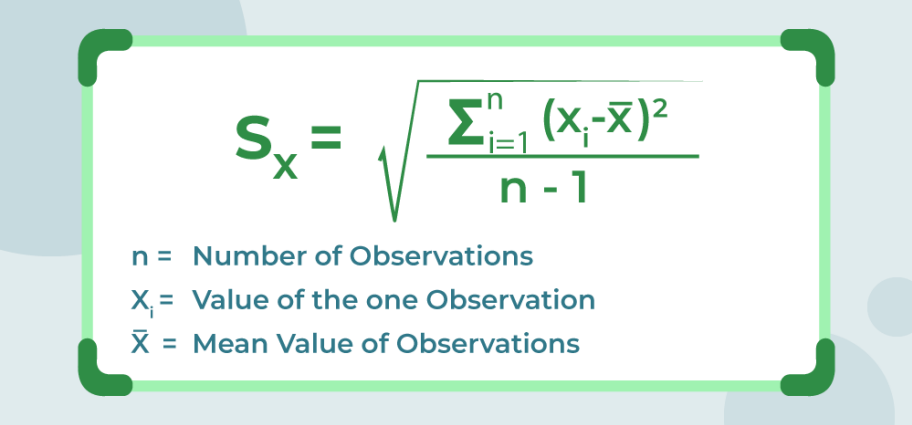
1. Mean:
   - Average of values
   - Sensitive to outliers
   - Outliers pull mean to a direction, which causes skewness
   - Only good for symmetrical distribution
2. Median:
   - Middle value of sorted data
   - Non-sensitive to outliers
   - Good for non-symmetrical data
3. Mode:
   - Most frequent value in a dataset
   - points out the peak of a distribution

Measures of Dispersion or Variability

1. **Range:** **Difference between the highest and lowest value in a distribution**
2. **Variance (s^2):** **Sum of square differences between each data point and the sample mean divided by n minus 1**



3. **Standard Deviation:** **It is the square root of the variance**



Quartiles

The median of the data is the second quartile (50%)

The median of the lower half data is the first quartile (25%)

The median of the upper half data is the third quartile (75%)

Interquartile Range (IQR): the middle 50% of the data Q3-Q1