# DESCRIPTIVE VS. INFERENTIAL STATISTICS:

## DESCRIPTIVE STATISTICS

INVOLVE THE ORGANIZATION, SUMMARIZATION, AND PRESENTATION OF DATA TO DESCRIBE ITS ESSENTIAL FEATURES. THIS INCLUDES MEASURES OF CENTRAL TENDENCY (MEAN, MEDIAN, MODE) AND MEASURES OF DISPERSION (RANGE, VARIANCE, STANDARD DEVIATION). FOR INSTANCE, CONSIDER A DATASET OF EXAM SCORES WHERE WE CALCULATE THE MEAN SCORE TO REPRESENT THE AVERAGE PERFORMANCE OF STUDENTS.

## Inferential statistics

**extrapolate insights beyond the observed data to make inferences about populations based on sample data. It encompasses hypothesis testing, estimation, and prediction. For example, we might use inferential statistics to determine whether a new teaching method improves student performance by comparing the exam scores of students taught using the new method with those taught using the traditional method.**

# QUANTITATIVE VS. QUALITATIVE DATA:

## Quantitative data

are numerical and measurable, allowing for mathematical operations and statistical analysis. Discrete quantitative data are countable and often represent whole numbers, such as the number of customers visiting a store. Continuous quantitative data are measured on a continuous scale and can take any value within a range, such as temperature or weight.

## QUALITATIVE DATA

**are descriptive and categorical, providing insights into attributes, characteristics, or qualities. Examples include survey responses categorizing customer satisfaction levels (e.g., "very satisfied," "satisfied," "neutral," "unsatisfied").**

# SCALES OF MEASUREMENT:

**Scales of measurement categorize variables based on the level of measurement and the operations applicable to them.**

## 1-NOMINAL SCALE:

Nominal data are categorical and represent qualitative differences without any inherent order. Examples include gender (male, female) or marital status (married, single, divorced).

## 2-ORDINAL SCALE:

Ordinal data possess a meaningful order or ranking, but the intervals between categories may not be equal. For instance, Likert scales measuring agreement levels (e.g., strongly agree, agree, neutral, disagree, strongly disagree) represent ordinal data.

## 3-INTERVAL SCALE:

**Interval data have equal intervals between values, but zero does not indicate the absence of the attribute being measured. Examples include temperature measured in Celsius or Fahrenheit.**

## 4-RATIO SCALE:

**Ratio data have equal intervals between values, with a true zero point indicating the absence of the attribute being measured. Examples include height, weight, and time.**

# HYPOTHESIS TESTING AND NULL HYPOTHESIS:

Hypothesis testing involves making statistical inferences about populations based on sample data. It typically begins with the formulation of a null hypothesis (H0), which represents the status quo or the absence of an effect, difference, or relationship. For example, in a clinical trial evaluating the efficacy of a new drug, the null hypothesis might state that there is no difference in the recovery rates between patients receiving the new drug and those receiving a placebo.

# ALTERNATIVE HYPOTHESES:

**The alternative hypothesis (H1 or Ha) proposes a specific effect, difference, or relationship that the researcher seeks to detect. It encompasses various scenarios, including one-tailed (directional) and two-tailed (non-directional) hypotheses. For instance, a one-tailed alternative hypothesis might assert that the mean exam score of students taught using a new teaching method is higher than that of students taught using the traditional method.**

# P-VALUE:

**The p-value quantifies the strength of evidence against the null hypothesis. It represents the probability of observing the data or more extreme data under the assumption that the null hypothesis is true. A smaller p-value indicates stronger evidence against the null hypothesis, suggesting that the observed results are unlikely to have occurred by random chance alone.**

For example, suppose we conduct a hypothesis test to determine whether a new advertising campaign has a significant impact on sales. If the resulting p-value is less than a predetermined significance level (e.g., $\alpha = 0.05$), we reject the null hypothesis and conclude that the advertising campaign has a statistically significant effect on sales.

# CONFIDENCE INTERVALS:

**Confidence intervals provide a range of values within which the true population parameter is likely to fall with a specified level of confidence. They offer insights into the precision and variability of estimates derived from sample data. The width of the confidence interval reflects the uncertainty associated with the estimation process.**

For instance, consider estimating the average height of adult males in a population. By calculating a 95% confidence interval, we can infer that we are 95% confident that the true average height of adult males lies within the calculated interval.

# REGRESSION ANALYSIS:

**Regression analysis explores relationships between variables and is commonly used for prediction and modeling purposes. It involves fitting a regression model to the data and assessing the strength, direction, and significance of the relationships between the independent (predictor) and dependent (outcome) variables.**

**For example, in a study investigating the relationship between employee satisfaction (independent variable) and job performance (dependent variable), regression analysis can help determine the extent to which changes in employee satisfaction predict variations in job performance.**