

Data Analytics Project

12. Types of Statistics



Statistics, Populations & Samples

Statistics: It is the science of learning from data, and of measuring, controlling, and communicating uncertainty. It involves studies of Populations & Samples.

Populations: It is an entire group under study, often large and challenging to examine entirely.

Samples: It is the subset of the population, representing the whole for analysis.



TWO Types of Statistics

A. Descriptive Statistics:

1. Summarize and describe data.
2. Simplifies datasets.
3. Provides insights into trends.

Examples: Total data points, value ranges, frequency.

B. Inferential Statistics:

1. Draw conclusions from observed data.
2. Predicts based on samples.
3. Uses hypothesis testing, regression, and probability.



DESCRIPTIVE Statistics (1/2)

Descriptive statistics **summarize** and describe the **values and observations** of a dataset.

USAGE:

1. To **simplify** and **describe** data sets.
2. Providing insights into **trends** and **occurrences**.

EXAMPLES:

1. **Total number** of data points.
2. **Range** of numeric values.
3. **Frequency** of specific values.



DESCRIPTIVE Statistics (2/2)

REPRESENTATION:

Descriptive statistics are often presented in graphical formats like pie charts, bar charts, or histograms.

LIMITATIONS:

This cannot be used for group comparisons, conclusions, or predictions beyond the observed population.



INFERENTIAL Statistics (1/2)

Inferential statistics involve collecting, analyzing, and interpreting data from a sample to make generalizations or predictions about a population.

USAGE:

1. Generalizing findings from a sample to a larger population.
2. Testing hypotheses and drawing conclusions.

EXAMPLES:

1. Predicting election outcomes from a sample.
2. Testing a new drug's effectiveness on patients.



INFERENTIAL Statistics (2/2)

METHODOLOGICAL CONCERNS:

1. Ensuring the chosen sample accurately represents the population.
2. Addressing potential biases in sampling techniques.

SAMPLING TECHNIQUES:

1. Random sampling.
2. Stratified sampling.
3. Cluster sampling.
4. Convenience sampling.



Statistics & Big Data

Descriptive Statistics in Big Data:

1. Descriptive statistics describe a sample and help determine data quality.
2. Graphical representations aid in assessing sample quality.

EXAMPLES:

1. Tweets for analysis may contain text or both text and images.
2. Analysis goals dictate whether to include tweets with images.



Inferential Analysis in Big Data

1. **CLUSTER ANALYSIS:** Identifies groups of similar observations.
2. **ASSOCIATION ANALYSIS:** Reveals co-occurrences of variable values.
3. **REGRESSION ANALYSIS:** Quantifies relationships between variables' variations.



THANK YOU!!! FOR YOUR SUPPORT! For Now...

Keep Learning, Keep Sharing & Keep Following
Aagam Deolasi.



Aagam Deolasi
@Aagam Deolasi

