

SNS COLLEGE OF TECHNOLOGY

**An Autonomous Institution
Coimbatore-35**



DEPARTMENT OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE

23ADT202 – FUNDAMENTALS OF DATA SCIENCE AND ANALYTICS

II YEAR IV SEM

UNIT I – INTRODUCTION TO DATA SCIENCE

EXPLORATORY DATA ANALYSIS

EXPLORATORY DATA ANALYSIS

EMPATHY:

Exploratory Data Analysis (EDA) is an important step in data science and data analytics as it visualizes data to understand its main features, find patterns and discover how different parts of the data are connected.

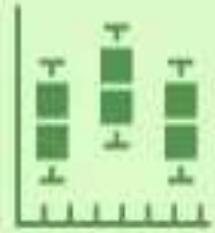
Exploratory Data Analysis (EDA)

DEFINE:



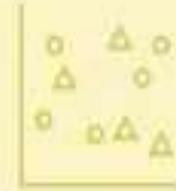
EDA

Explore and summarize data before modeling



Distribution

Use histograms and box plots



Relationships

Use scatter and pair plots



Spot Issues

Find outliers or missing data

Why is Exploratory Data Analysis Important

IDEATE:

- Helps to understand the dataset by showing how many features it has, what type of data each feature contains and how the data is distributed.
- Helps to identify hidden patterns and relationships between different data points which help us in and model building.
- Allows to identify errors or unusual data points (outliers) that could affect our results.
- The insights gained from EDA help us to identify most important features for building models and guide us on how to prepare them for better performance.
- By understanding the data it helps us in choosing best modeling techniques and adjusting them for better results.

Types of Exploratory Data Analysis

TESTING:

1. Univariate Analysis

Focuses on studying one variable to understand its characteristics. It helps to describe data and find patterns within a single feature.

2. Bivariate Analysis

Focuses on identifying relationship between two variables to find connections, correlations and dependencies..

3. Multivariate Analysis

Identify relationships between two or more variables in the dataset and aim to understand how variables interact with one another which is important for statistical modelling techniques.

Steps for Performing Exploratory Data Analysis

PROTOTYPING:

Step 1: Understanding the Problem and the Data

Step 2: Importing and Inspecting the Data

Step 3: Handling Missing Data

Step 4: Exploring Data Characteristics

Step 5: Performing Data Transformation

Step 6: Visualizing Relationship of Data

Step 7: Handling Outliers

Step 8: Communicate Findings and Insights