



SIMATS ENGINEERING

Saveetha Institute of Medical and Technical Sciences
Chennai- 602105



Student Name: GOGIREDDY VENKATA GOUTHAM REDDY **Reg. No.:** 192324265

Course Code: DSA0613

Slot: A

Course Name: Data Handling and Visualization for Data Analytics

Course Faculty: Dr. Kumaragurubaran & T Dr. Senthilvadivu S

Project Title: Design & Implementation of an Automated Data Visualization System for Exploratory Data Analysis

Module 2: Automated Visualization Engine

Module Photographs: (3 photographs –Module Photo, Individual student contribution module work in the project and presentation image)



Abstract

Data visualization is essential for understanding large and complex datasets. Raw data often needs to be transformed into a format that is easier to interpret and analyze. This system aims to provide an automated solution for exploratory data analysis. It is a hybridized software system that automatically generates visualizations based on the characteristics of the dataset. The system uses various machine learning algorithms to identify patterns and relationships within the data. These insights help analysts make informed decisions quickly and efficiently.

Purpose:
To analyze the uploaded dataset and automatically select suitable visualizations for exploratory data analysis.

Key Implementation Details:

- Identify numeric, categorical, and date/time columns
- Apply rule-based logic to select appropriate plot types
- Generate default visualizations based on data characteristics

Techniques Used:

- Data type detection using Pandas
- Rule-based decision logic
- Matplotlib and Seaborn for plot generation

EDA Visualizer - Automated Data Visualization

Automated Data Visualization

Upload a CSV file to automatically generate visualizations

Upload Dataset

Supports CSV files only

Project Description: (here you write what you did in this project (contribution) including Model Description)

The design and implementation of an automated data visualization system play a crucial role in simplifying exploratory data analysis by reducing the need for manual intervention. Such a system is designed to accept raw datasets and automatically perform essential preprocessing tasks, including data cleaning, detection of missing values, and identification of variable types. This automation ensures consistency and saves time, allowing analysts to focus more on interpretation rather than preparation. Once preprocessing is complete, the system analyzes the dataset to compute basic statistical measures such as mean, median, variance, and correlations. Based on these insights, it intelligently selects suitable visualization techniques. Common visual outputs include histograms for distribution analysis, box plots for identifying outliers, scatter plots for relationship exploration, and heatmaps for correlation patterns. These visualizations help users quickly understand data trends and anomalies. The system is designed to be dynamic and adaptable, updating visualizations automatically when the underlying data changes. Its architecture emphasizes scalability and modularity, making it suitable for large and complex datasets.

Student Signature

Guide Signature