



SIMATS ENGINEERING

Saveetha Institute of Medical and Technical Sciences
Chennai- 602105



Student Name: Cheredla Jasvina

Reg. No.: 192324026

Course Code: DSA0613

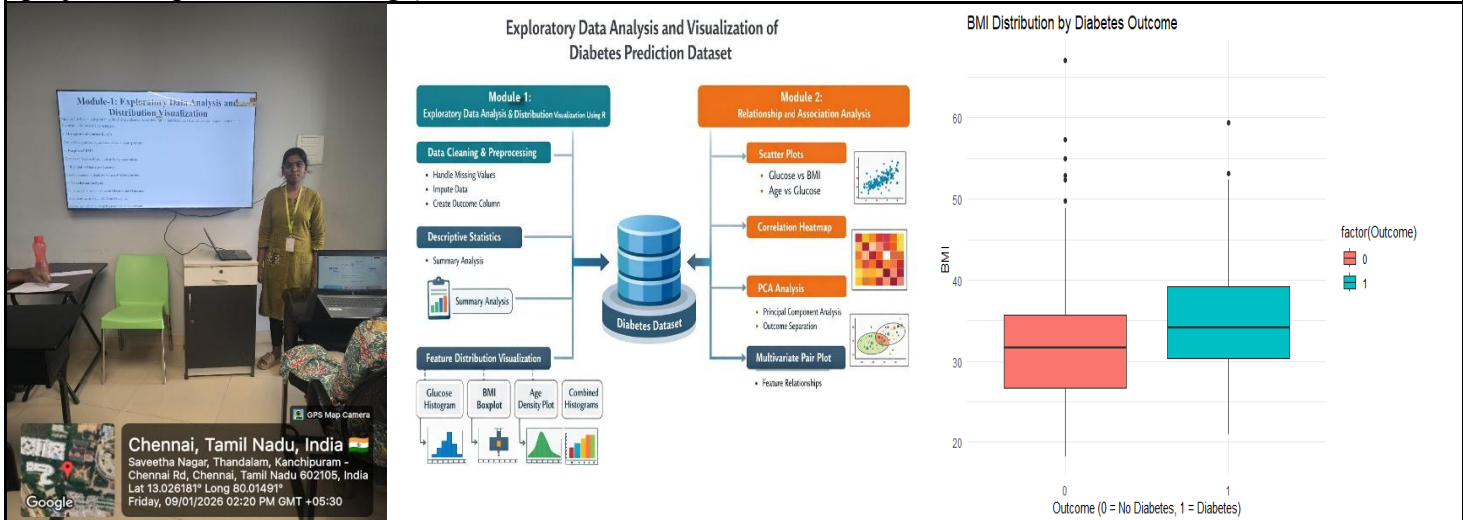
Slot: A

Course Name: Data Handling and Visualization for Data Analytics

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

Project Title: Exploratory Data Analysis and Visualization of Diabetes Prediction Dataset

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



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

Exploratory data analysis and visualization of a diabetes prediction dataset focuses on understanding medical data through statistical and visual techniques to uncover meaningful patterns related to diabetes. The title emphasizes the importance of analyzing healthcare data before applying predictive or advanced analytical models. By using visualization-driven analysis, the project aims to improve interpretability, identify trends, and support data-driven understanding of diabetes risk factors. This approach highlights the role of EDA as a foundational step in healthcare analytics.

Module 1 concentrates on the exploratory data analysis and distribution visualization of the diabetes dataset to ensure data quality and reliability. This module involves cleaning and preprocessing the dataset by handling missing and invalid values commonly found in medical records. Descriptive statistics and visualization techniques such as histograms, boxplots, and density plots are applied to analyze the distributions of key variables including Glucose, BMI, Age, Insulin, and Blood Pressure. Through this process, the module helps identify outliers, understand variability, and reveal underlying patterns that form a strong foundation for further analysis.

Student Signature

Guide Signature