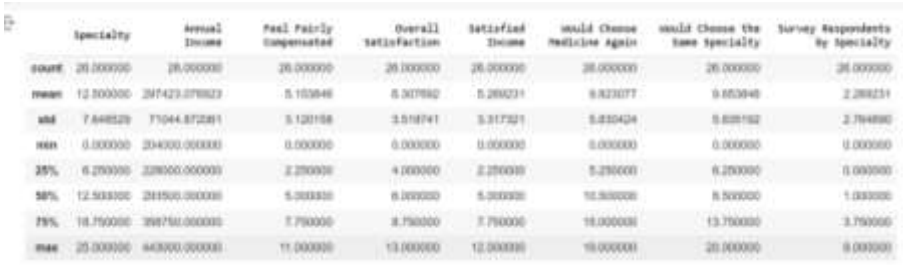
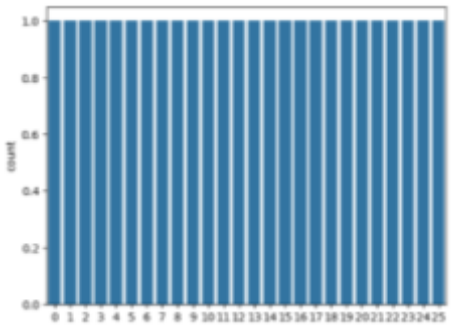
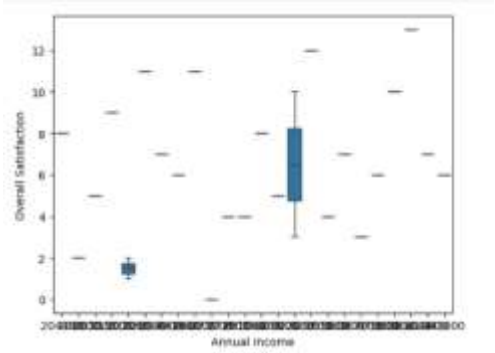
Data Collection and Preprocessing Phase

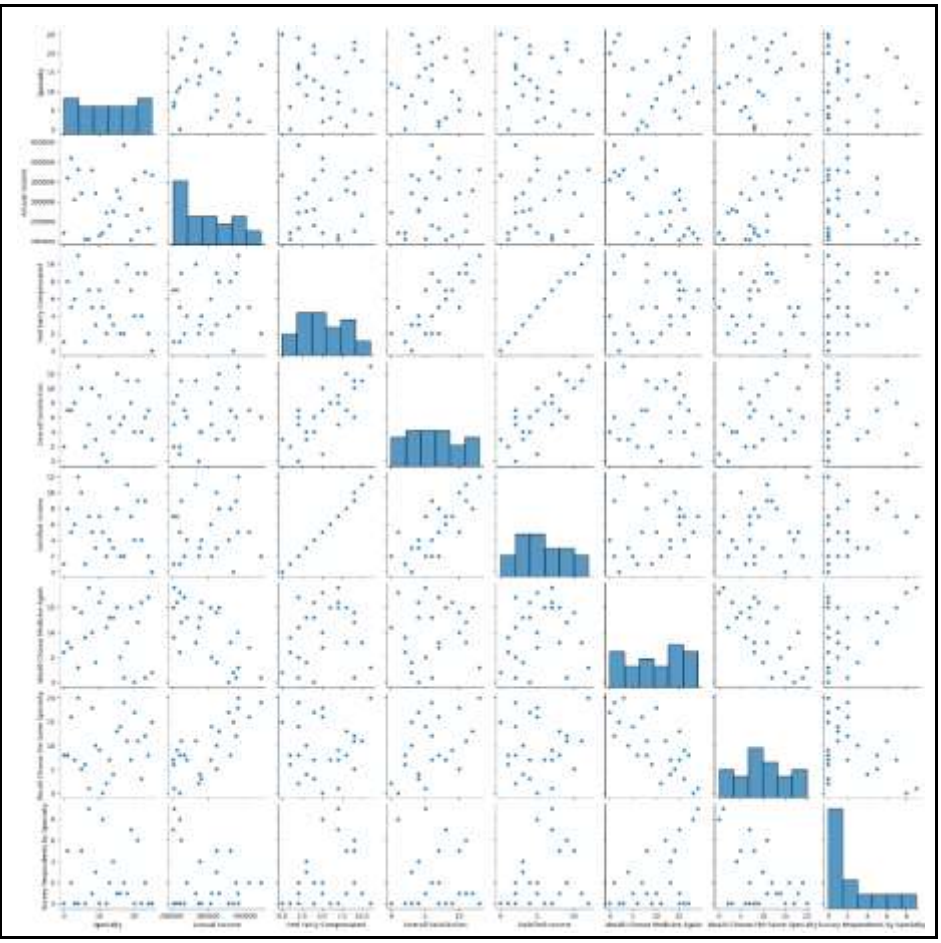
Date	15 July 2024
Team ID	740683
Project Title	Doctors Annual Salary prediction
Maximum Marks	6 Marks

Data Exploration and Preprocessing Report

Dataset variables will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling, and forming a strong foundation for insights and predictions.

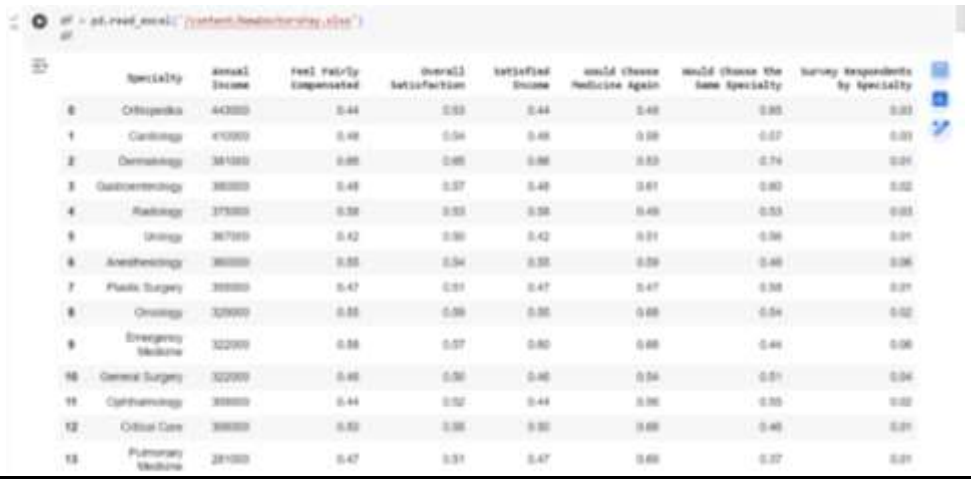
Section	Description
Data Overview	Dimension: 28 rows × 8 columns Descriptive statistics: 
Countplot Analysis	

	
Boxplot Analysis	

PairPlot Analysis	

Outliers and Anomalies	-

Data Preprocessing Code Screenshots

Loading Data	
Data Transformation	
Feature Engineering	Attached the codes in final submission.
Save Processed Data	-