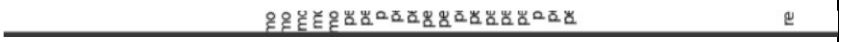
Section	Description
Data Overview	Summary of the dataset, including number of rows and columns, data types of each column, and brief descriptions of each column.
Univariate Analysis	Distribution analysis of individual variables using histograms, bar charts, and descriptive statistics (mean, median, mode, standard deviation).  **#Univariate Analysis**
Bivariate Analysis	Examination of relationships between pairs of variables using scatter plots, correlation matrices, and pairwise plots to identify patterns and trends.  **#Bivariate Analysis**
Multivariate Analysis	Investigation of interactions between multiple variables using heatmaps, PCA (Principal Component Analysis), and clustering to understand data structure. 
Outliers and Anomalies	Identification and description of outliers and anomalies, summarized in a table with details on detection method, number of outliers, description, and potential impact.
Data Preprocessing Code Screenshots	

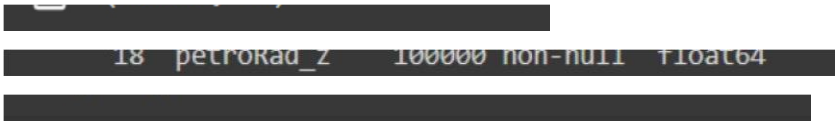
Loading Data	5 rows × 43 columns

Data Collection and Preprocessing Phase

Date	15 July 2024
Team ID	739858
Project Title	**SDSS galaxy classification using Machine Learning**
Maximum Marks	6 Marks

Data Exploration and Preprocessing Template

Exploration and Preprocessing Template for SDSS galaxy classification for Machine Learning: Load data, handle missing values, explore basic statistics, visualize distributions, encode categorical variables, normalize/scale features, identify outliers, and prepare for modeling

Handling Missing Data	 For checking the null values, . isnull() function is used. To sum those null values we use . sum() function. From the above image we found that there are no null values present in our dataset. So we can skip handling the missing values step.
Data Transformation	-
Feature Engineering	-
Save Processed Data	-