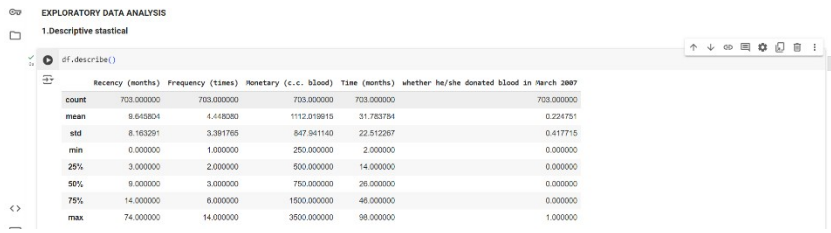
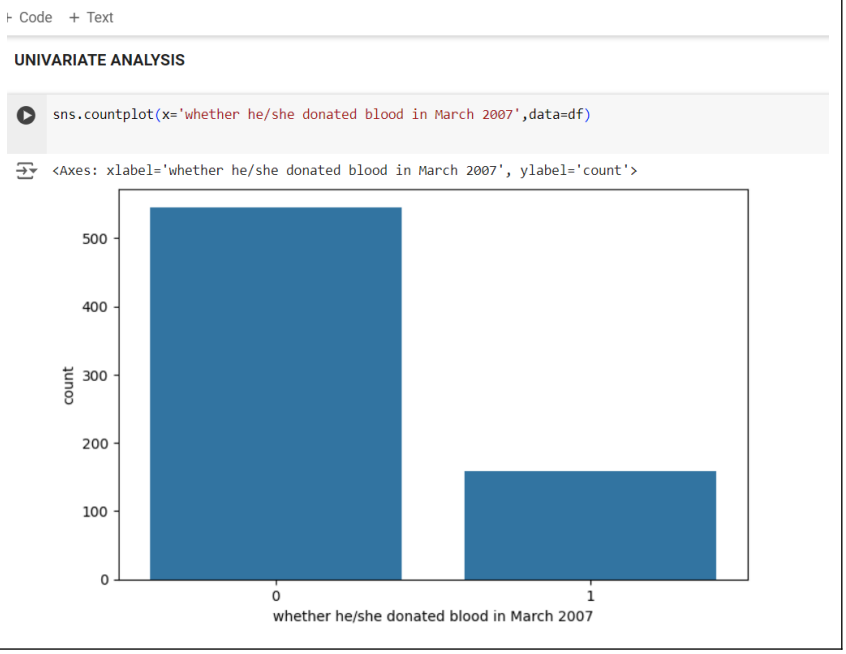
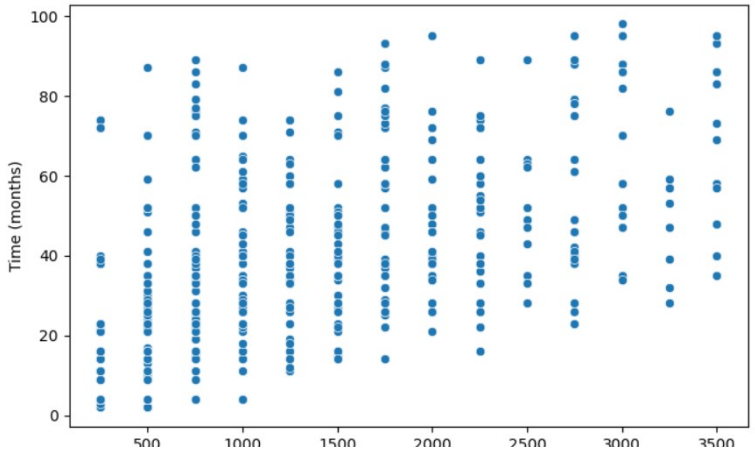
Data Collection and Preprocessing Phase

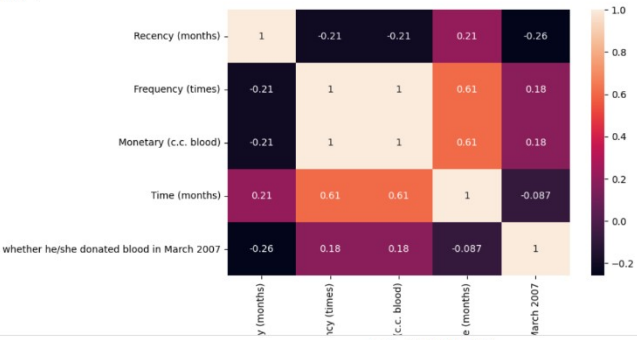
Date	14 July 2024
Team ID	739755
Project Title	Blood Donation Prediction
Maximum Marks	6 Marks

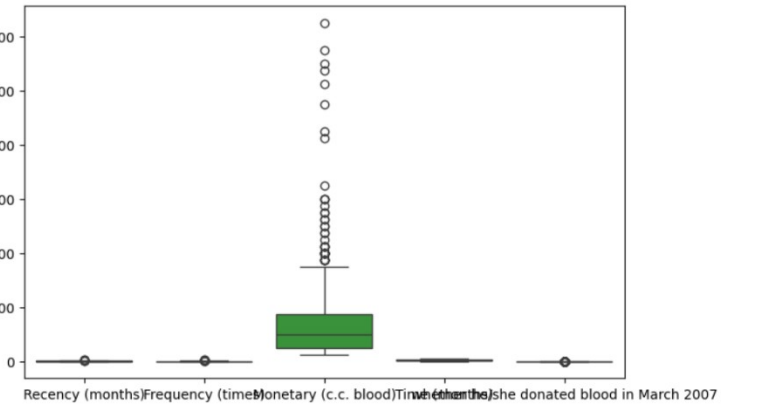
Data Exploration and Preprocessing Template

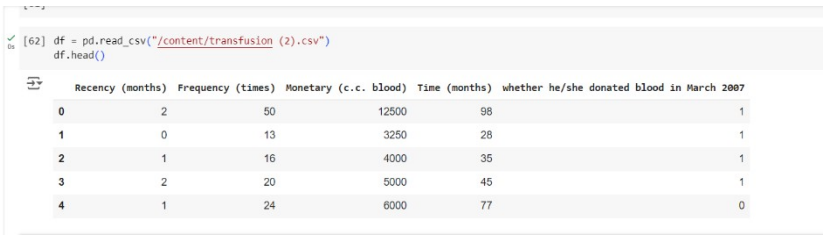
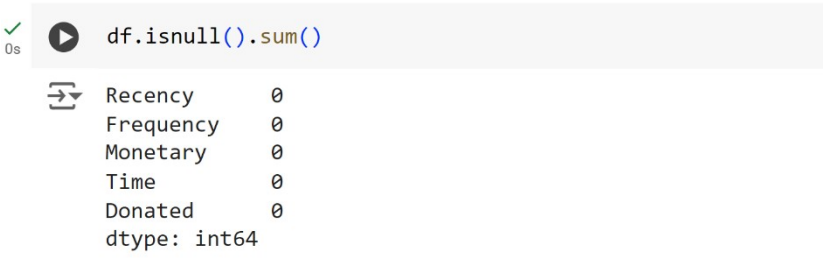
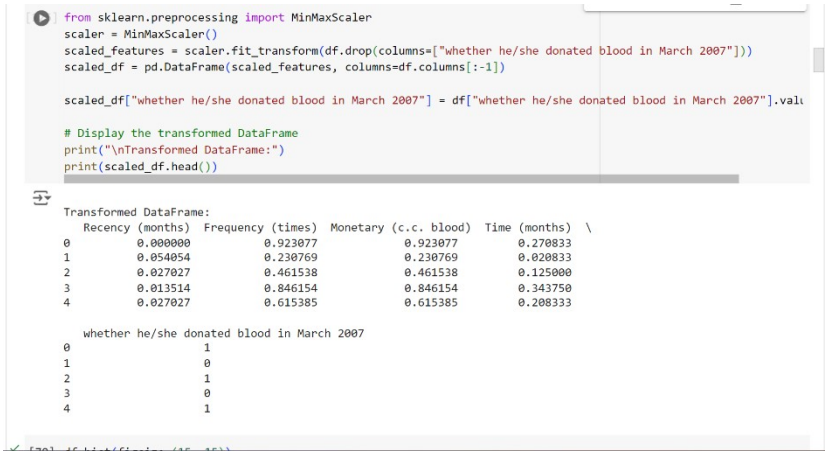
Identifies data sources, assesses quality issues like missing values and duplicates, and implements resolution plans to ensure accurate and reliable analysis.

Section	Description
Data Overview	
Univariate Analysis	

Bivariate Analysis	Code + Text
	BIVARIATE ANALYSIS
	`sns.scatterplot(x=df['Monetary (c.c. blood)'],y=df['Time (months)'])`
	`<Axes: xlabel='Monetary (c.c. blood)', ylabel='Time (months)'>`
	 

Multivariate Analysis	**MULTIVARIATE ANALYSIS**
	`[92] sns.heatmap(df.corr(),annot=True)`
	`<Axes: >`
	 
	✓ 0s completed at 19:08

Outliers and Anomalies	`sns.boxplot(df)`
	`<Axes: >`
	 

Data Preprocessing Code Screenshots	
Loading Data	
Handling Missing Data	
Data Transformation	
Feature Engineering	Attached the codes in final submission.
Save Processed Data	-