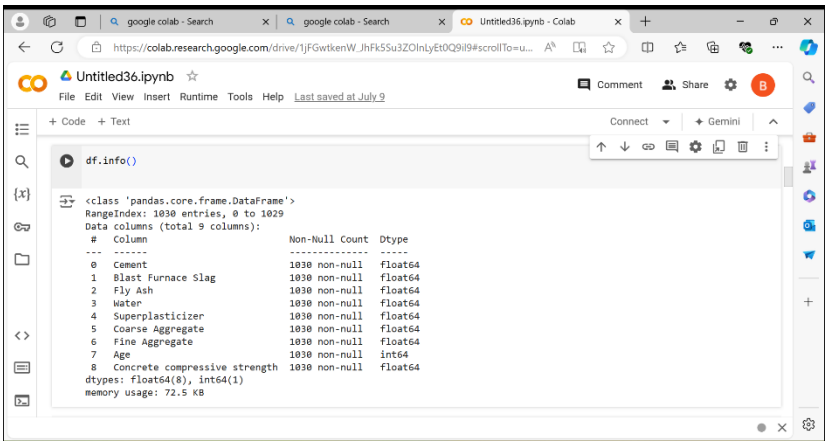
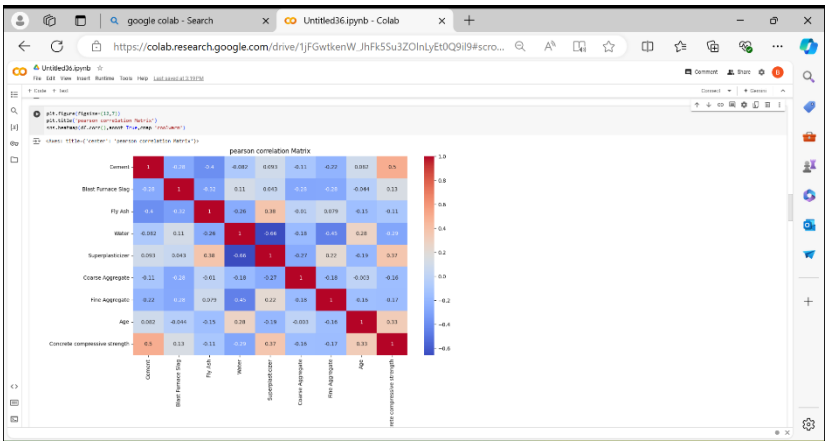
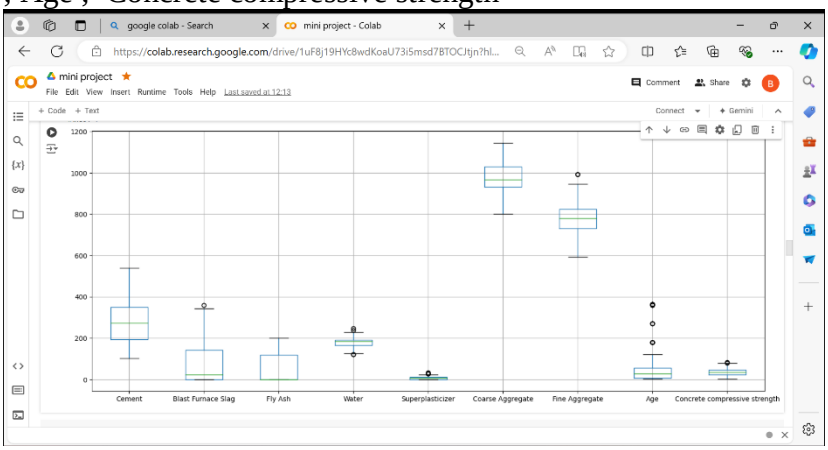
Data Collection and Preprocessing Phase

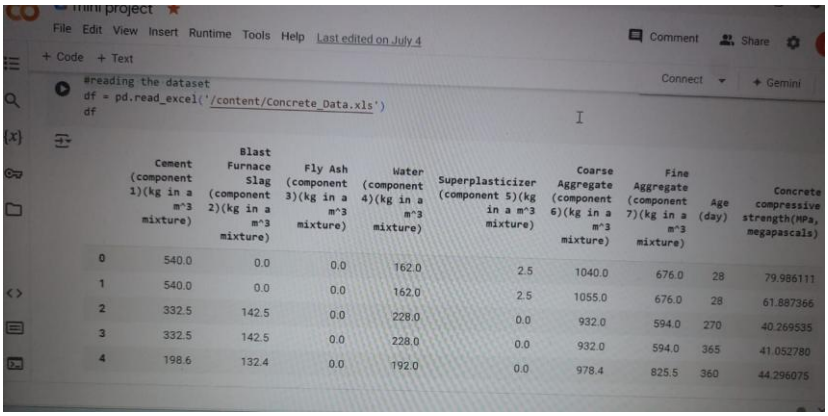
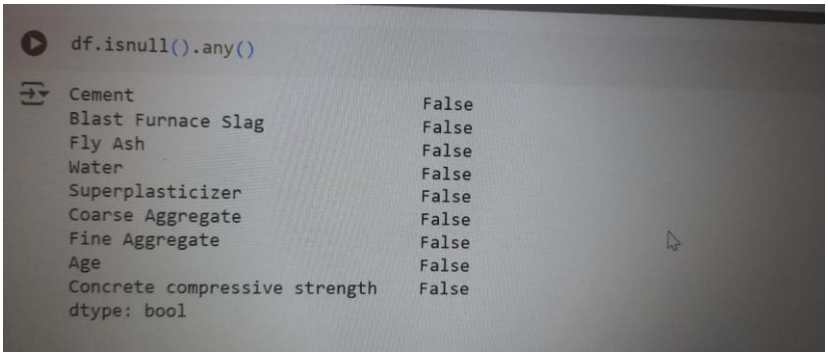
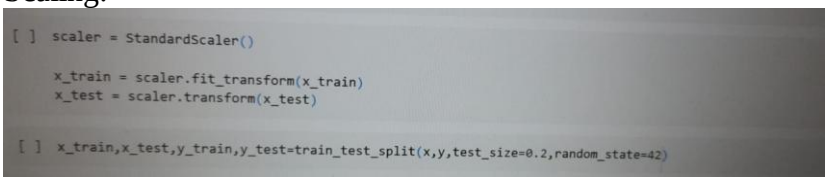
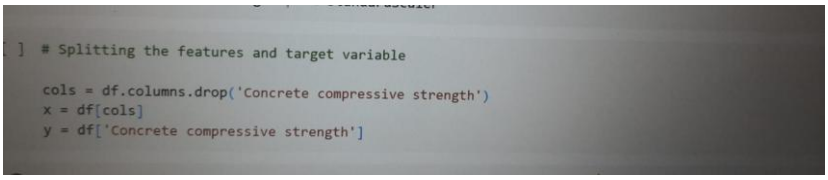
Date	6 July 2024
Team ID	739916
Project Title	Predicting the Compressive Strength of Concrete
Maximum Marks	6 Marks

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	1030 rows x 9 columns, dtypes: float64 
Univariate Analysis	Exploration of individual of accuracy_score,mean_squared_error,r2_score,mean_absolute_error
Bivariate Analysis	Relationships between two variables (correlation, scatter plots)

	 .
Multivariate Analysis	
Outliers and Anomalies	Identification of outliers in the dataset. Their is outliers in Blast Furnace slag, water, superplasticizer, Fine Aggregate , Age , Concrete compressive strength 
Data Preprocessing Code Screenshots	

Loading Data	
Handling Missing Data	
Data Transformation	Scaling: 
Feature Engineering	modifying existing ones 
Save Processed Data	Code to save the cleaned and processed data for future use.