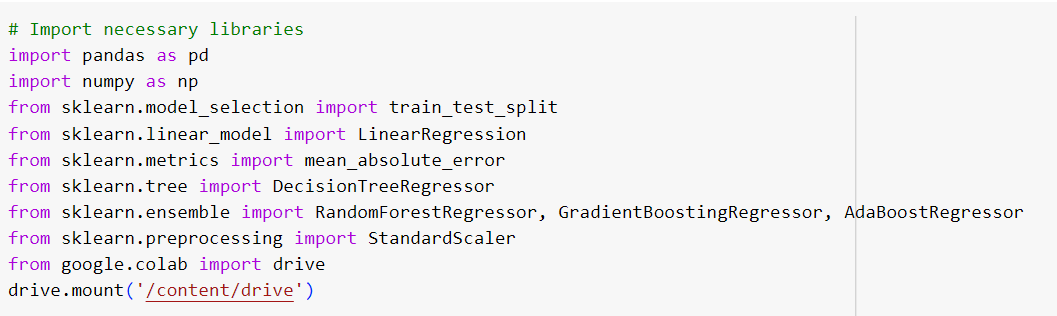
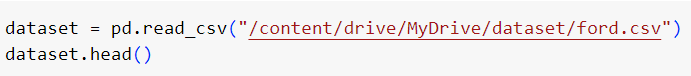
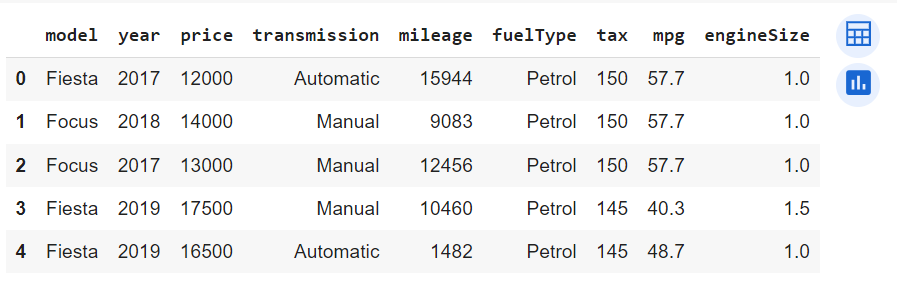
**Program:-**

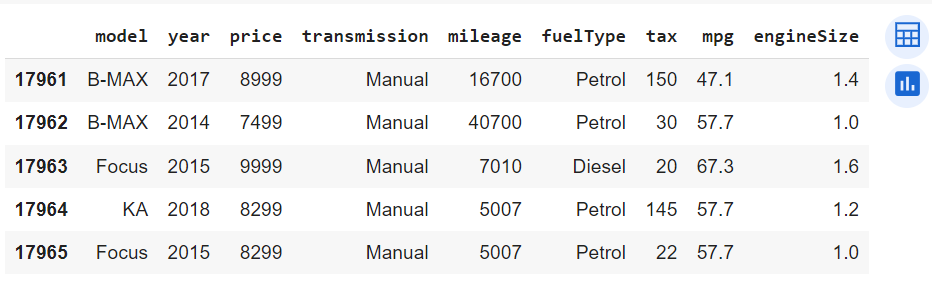
****

****

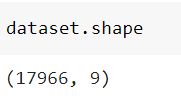
**#Dataset(head):-**

****

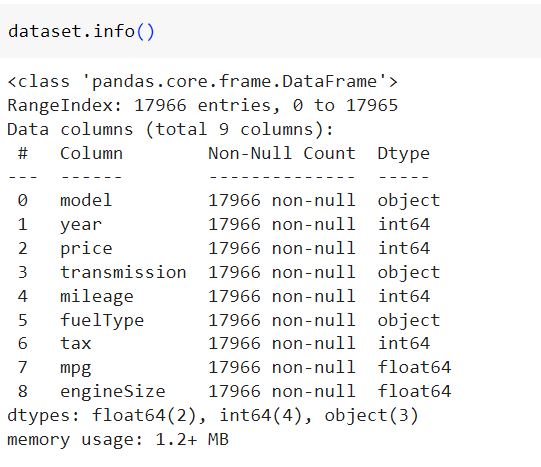
**#Dataset(tail):-**

****

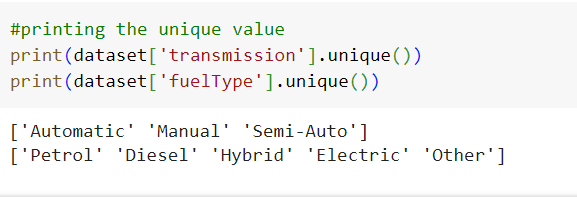
**#Rows & Columns of the Dataset:-**

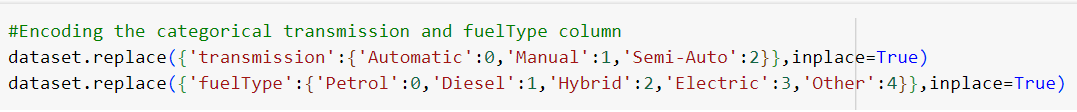
****

**#Info about the Dataset:-**

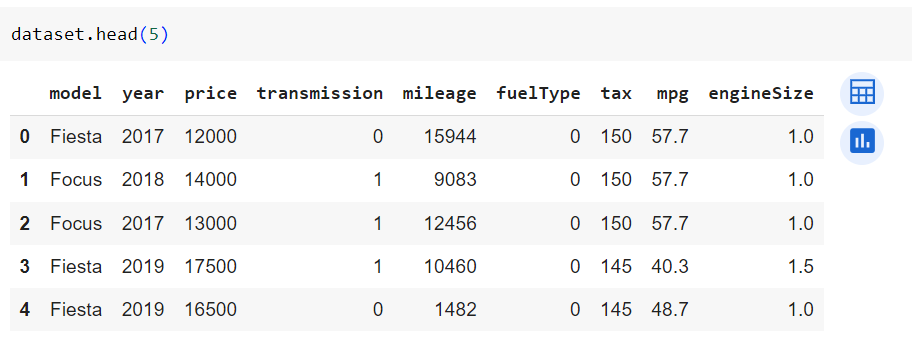
****

**#Datapreprocessing:-**

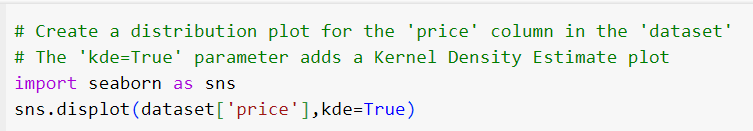
****

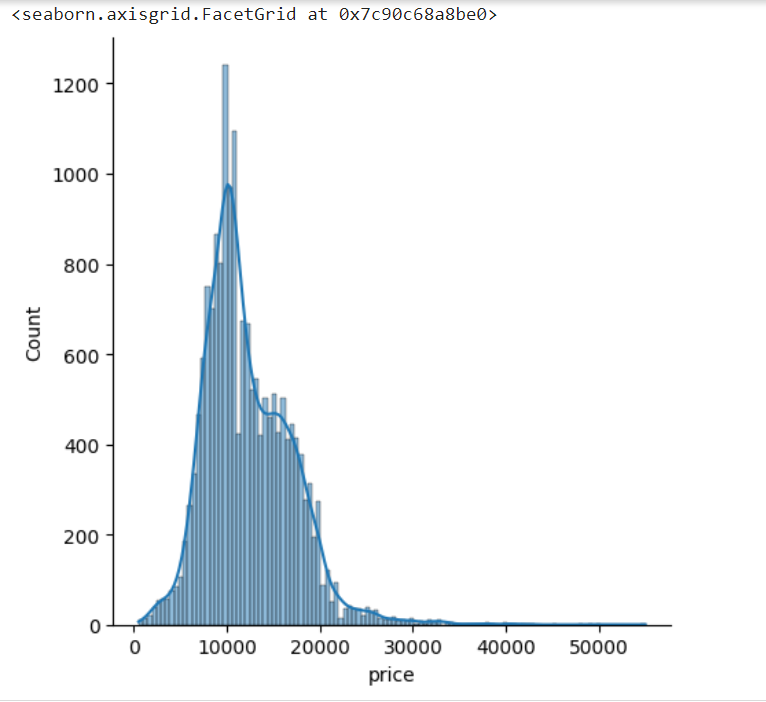
****

**#Checking for the conversion:-**

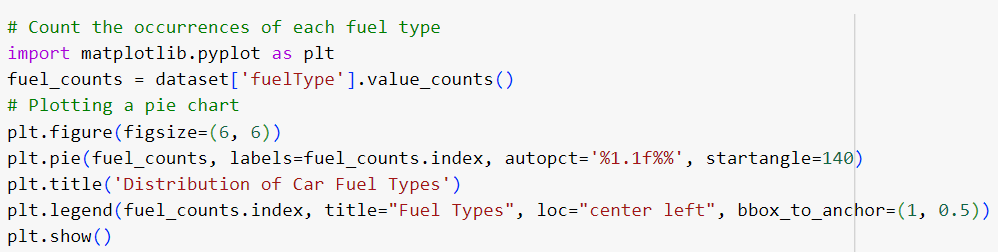
****

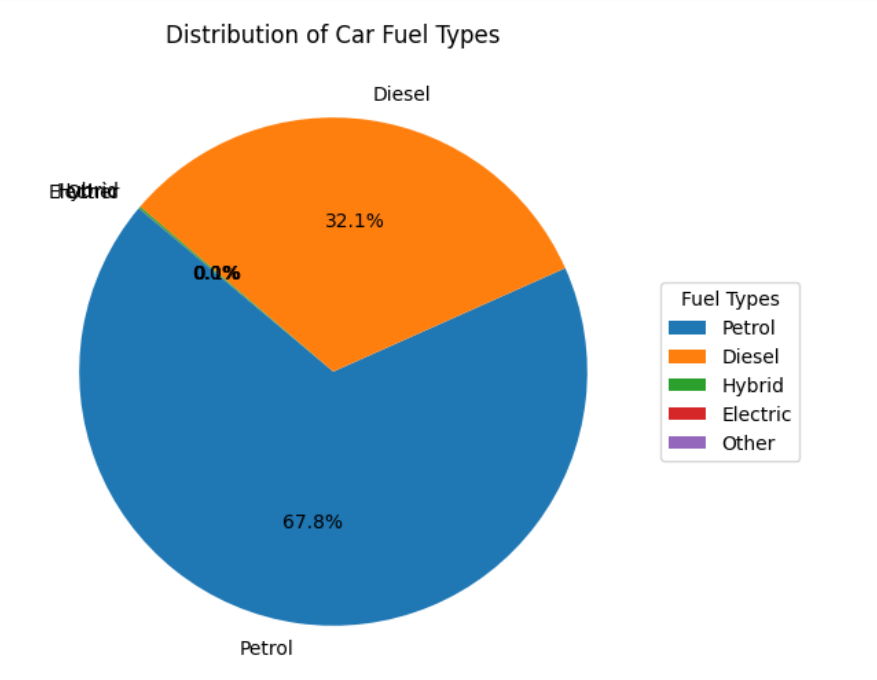
**#Creating a distribution plot (histogram) for the 'price' column in your dataset**

****

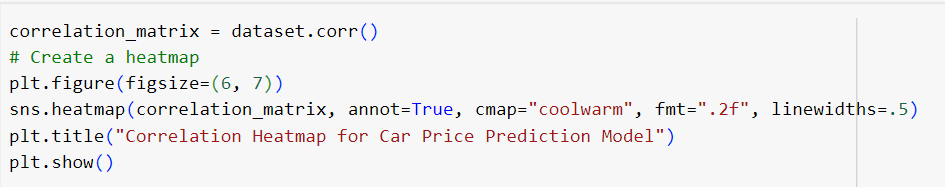
****

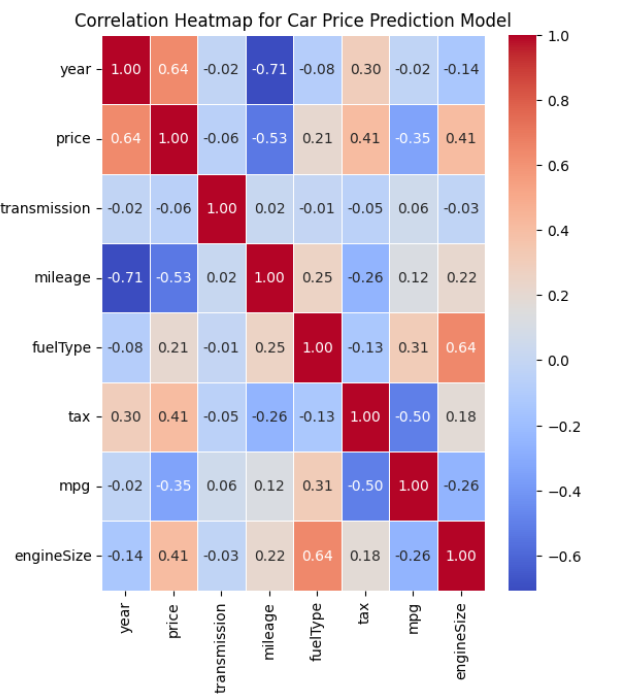
**#Distribution of Car’s Fuel Type**

****

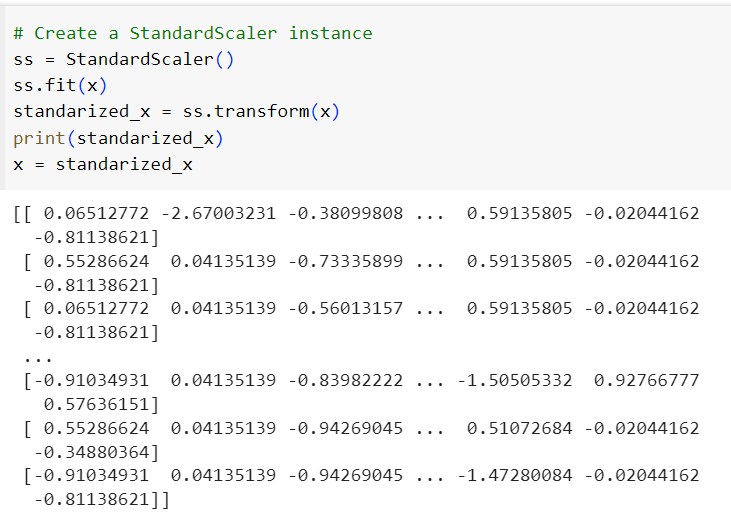
****

**# Calculate the correlation matrix for the dataset**

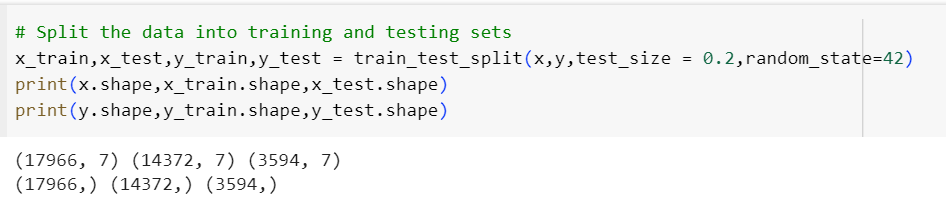
****

****

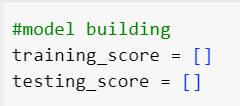
**# Transform (standardize) the data using the computed mean and standard deviation**

****

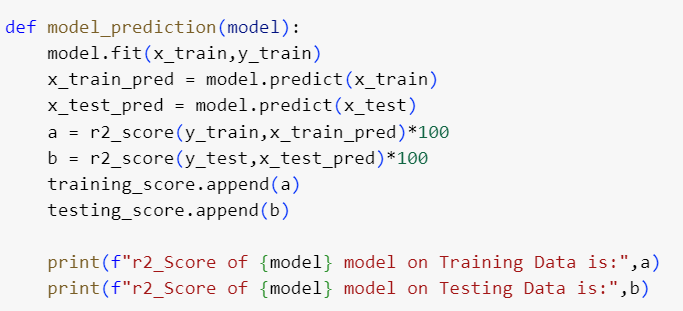
**#Splitting the dataset**

****

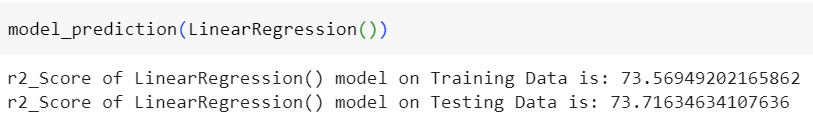
**#Model Building**

****

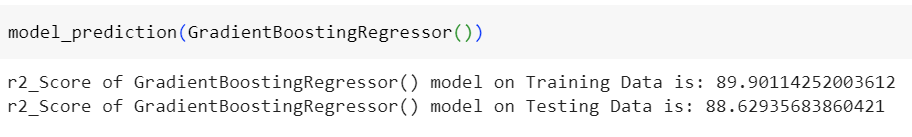
**#Function for Model\_Predictions**

****

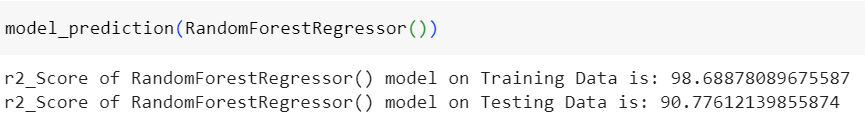
**#Model Prediction for Linear Regression**

****

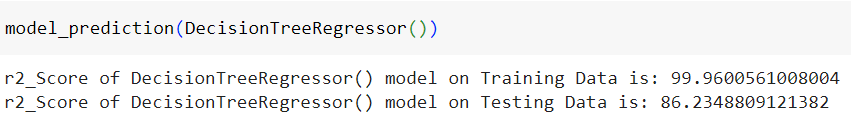
**#Model Prediction for Gradient Boosting Regressor**

****

**#Model Prediction for Random Forest Regressor**

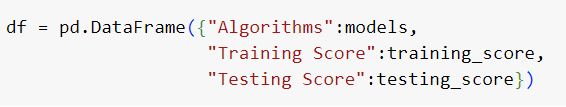
****

**#Model Prediction for Decision Tree Regressor**

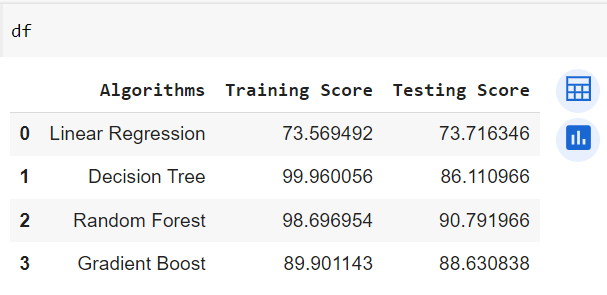
****

****

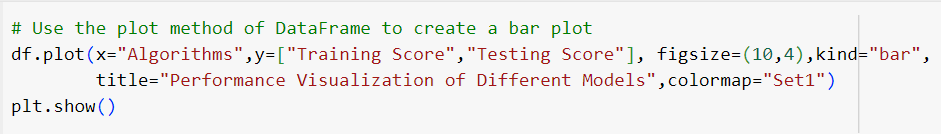
**#Making Dataframe for all Predicted models**

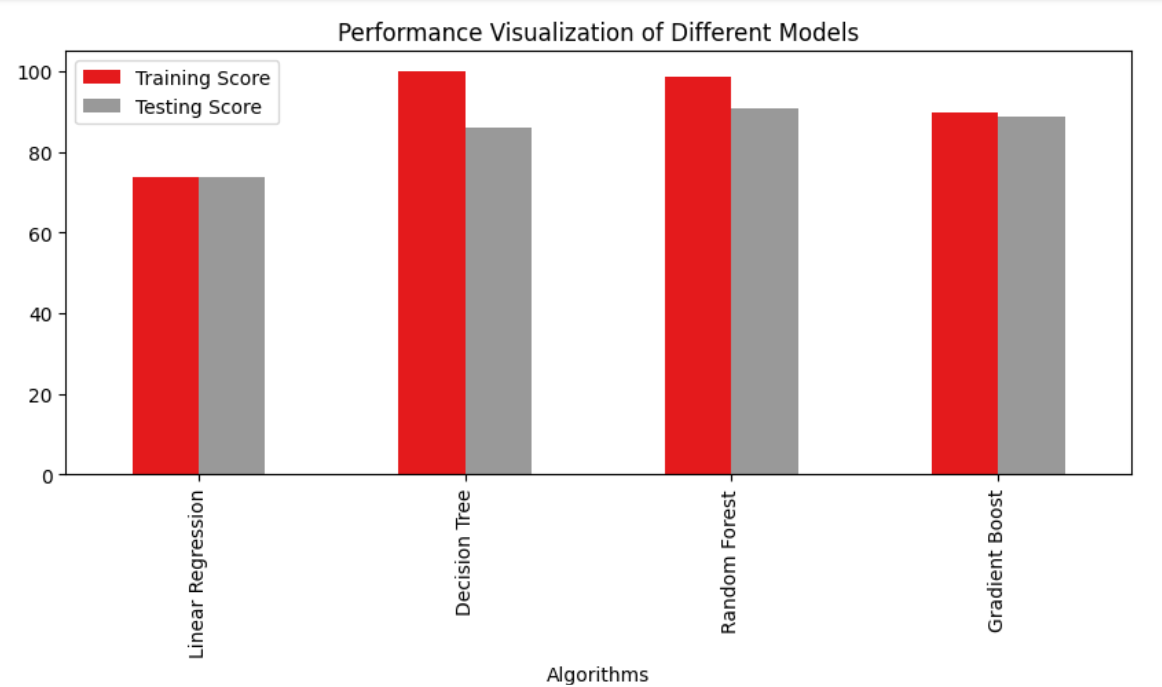
****

**#R2 Score of all Predicted Models**

****

**#Plotting Bar for Visualization**

****

****

**Conclusion :** Based on the various Model’s Prediction Decision Tree predicts better on this dataset

**CAR PRICE PREDICTION MODEL**

**Aim:** The aim of our project is to predict car prices based on various features using different regression techniques (Linear Regression, Random Forest Regressor, Gradient Boosting Regressor, Decision Tree Regressor).

**Steps Involved:**

1. Utilize the pandas read\_csv() function to import the car dataset.
2. Filter the data to include only relevant features and records, focusing on aspects that contribute significantly to car prices.
3. Check for any null values in the dataset and replace them with appropriate values, such as the mean or median for numerical features.
4. Create a heatmap to visualize the correlation between different features. Remove features with high correlation to avoid multicollinearity.
5. Split the dataset into training and testing sets to train and evaluate the regression models.
6. Regression Models: Apply various regression algorithms to predict car prices:
   * Linear Regression
   * Random Forest Regression
   * Gradient Boosting Regression
   * Decision Tree Regression
7. **Model Evaluation:** Assess the accuracy of each model using metrics R-squared metrics.

**Conclusion:** Summarize the findings and identify the most effective regression model for predicting car prices.

Dataset consist of 17966 rows \* 9 columns consists of the following :-

| **model** | **year** | **price** | **transmission** | **mileage** | **fuelType** | **tax** | **mpg** | **engineSize** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |

**MACHINE LEARNING**

**MINI PROJECT**



**Indira Gandhi Delhi Technical University for Women**

**SUBMITTED BY -**

Kusum Sharma

00602102023

1st Sem   M.Tech CSE AI

**SUBMITTED TO –**

**Prof. Seeja .K.R**