

Car Price Prediction

Submitted by:

Akshat Gaur

**ACKNOWLEDGMENT**

I want to thank kaggle, cars24 and olx for providing the data set and my mentor sajid chaudhary for guiding me

**INTRODUCTION**

* Business Problem Framing

The problem is we have the price of used cars and this will be very useful for the used cars industry.

* Review of Literature

The research is brief done with the scrapped data of used cars.

We predicted the growth and price prediction of the industry with the help of regression techniques.

* Motivation for the Problem Undertaken

The objective was to learn and explore the techniques to build a model that helps the real world to solve the problem of car industries.

**Analytical Problem Framing**

* Mathematical/ Analytical Modeling of the Problem

We were given a project for predicting used car prices, it was a regression model so i used regression techniques to create and build my model.

* Data Sources and their formats

I collected data from kaggle, olx and cars 24 through webscrapping and it was in CSV format.

* Data Preprocessing Done

First I collected the data from the above website, then i did some preprocessing by dropping null values, unnecessary columns that were not important for the model.

Then i skewed the data and performed the techniques.

* Data Inputs- Logic- Output Relationships

We were clear about the problem statement and the data had to be scrapped of the internet so i tried every possible way to scrape data which was relevant to the problem statement. The data was in CSV format.

* State the set of assumptions (if any) related to the problem under consideration

Used cars have a huge market and this business is blooming and with the help of data science and analytics we can predict and help it grow even more.

* Hardware and Software Requirements and Tools Used

This project was made on MacBook Air and the software used were Safari, Anaconda and Jupyter notebook.

**Model/s Development and Evaluation**

* Identification of possible problem-solving approaches (methods)

First i cleaned the data, then i have plotted some graphs to show the relation and correlation of the data and to show it easily i used some visualizing techniques.

After EDA, i have split the data and used it for testing and training and used regression techniques to train the model.

* Testing of Identified Approaches (Algorithms)

Listing down all the algorithms used for the training and testing.

* Run and Evaluate selected models

All the algorithms used are

LinearRegression,

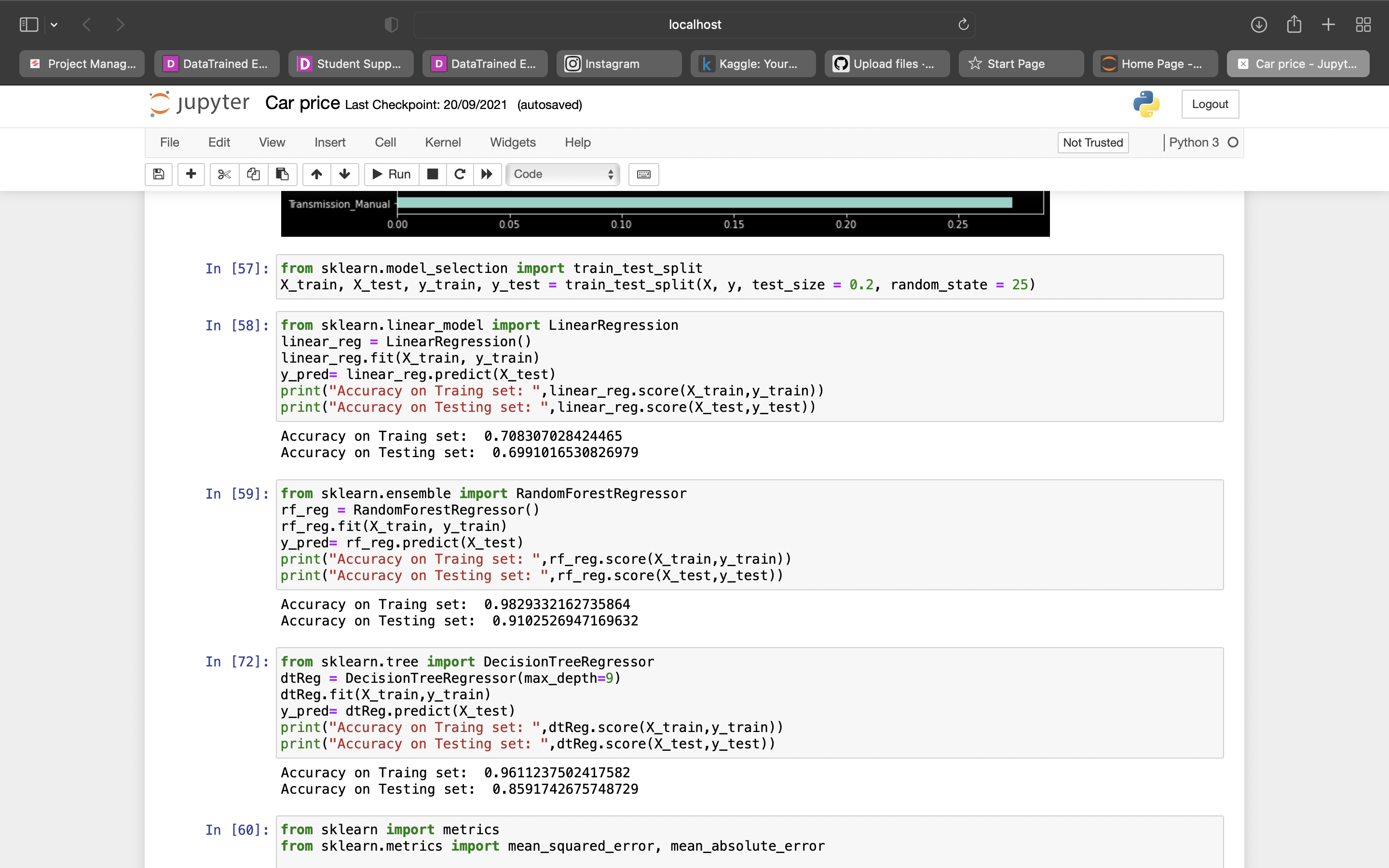
RandomForestRegressor,

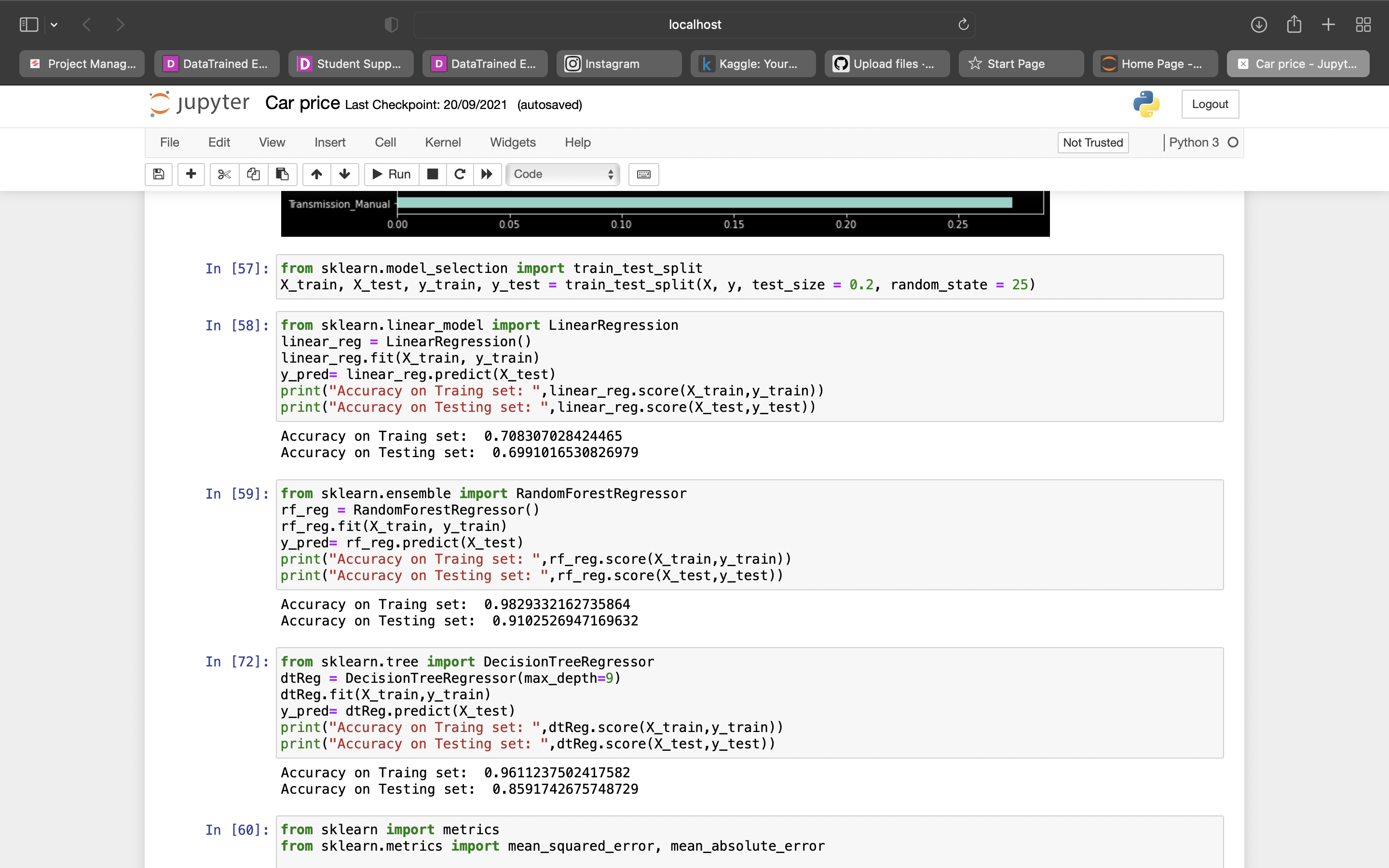
DecisionTreeRegressor,

mean\_squared\_error,

mean\_absolute\_error

Their codes are





.

* Visualizations

I have plotted various graphs for the visualization of the data to help us the data better and the interconnection of the features.

I have plotted the graph for price growth of cars

I have plotted the graph showing the fuel type

I have plotted the graph showing year of the make

I have plotted the graph showing the type of ownership

I have plotted the graph showing the companies of the cars

I have plotted the price over different regions of the country

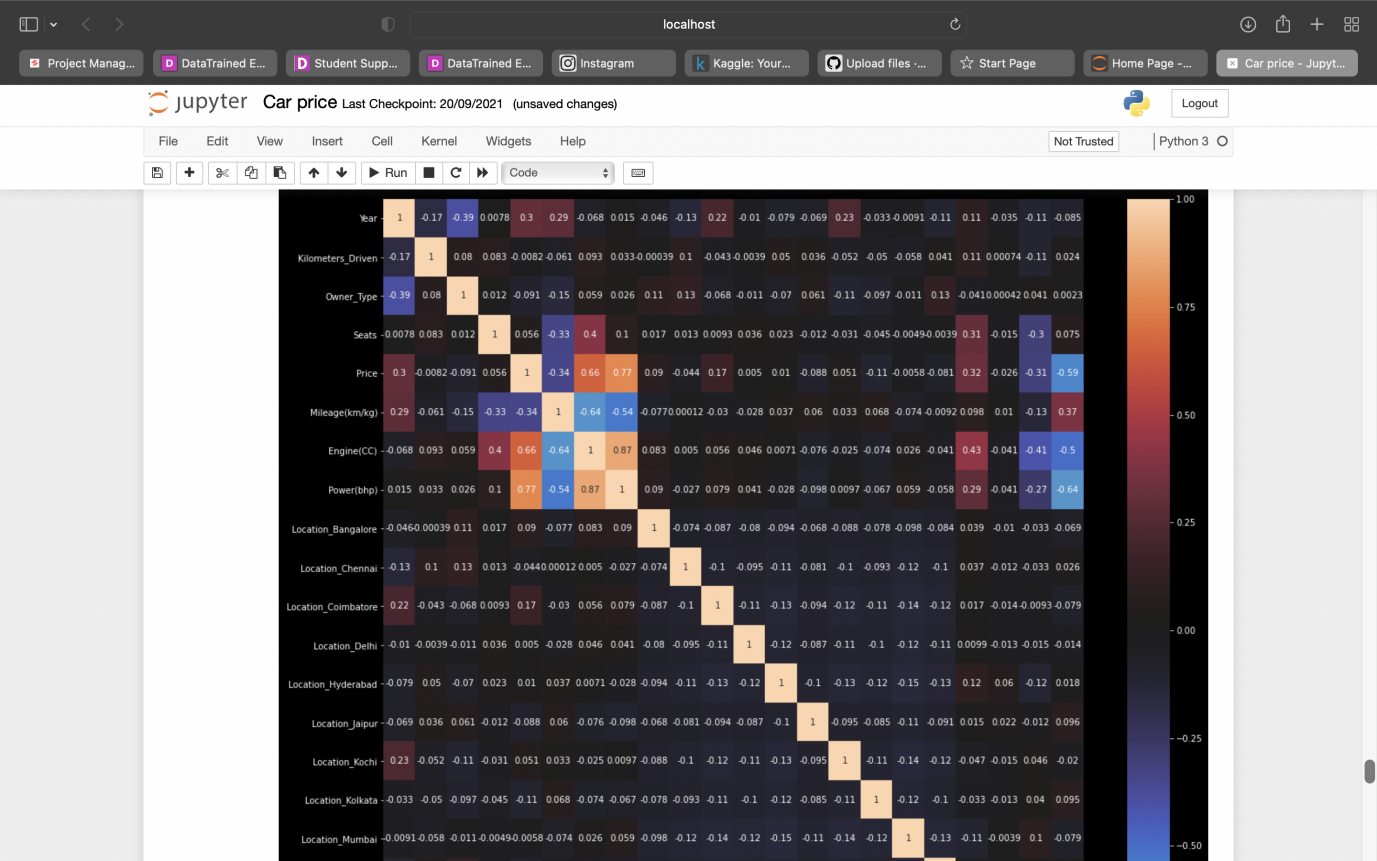
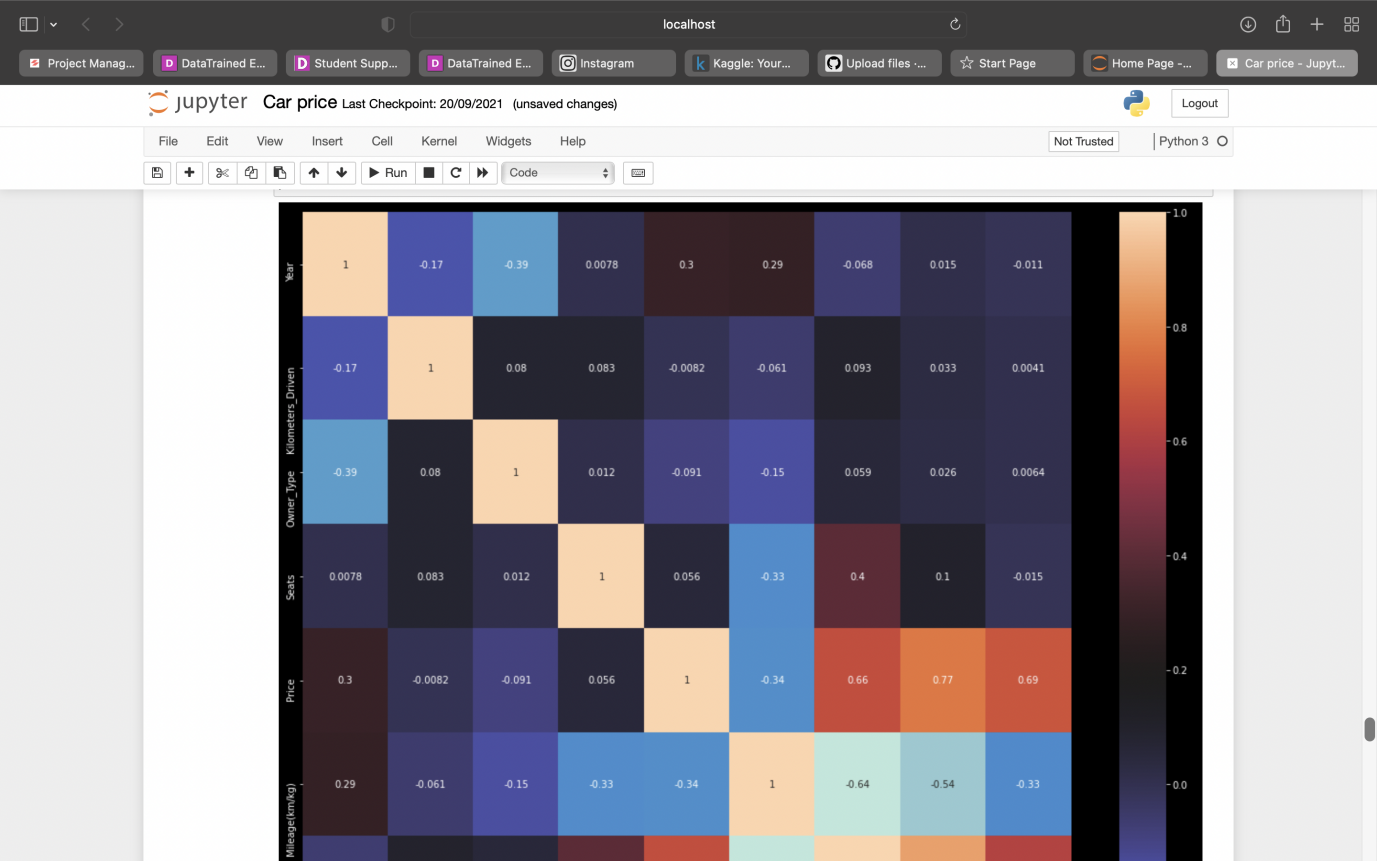
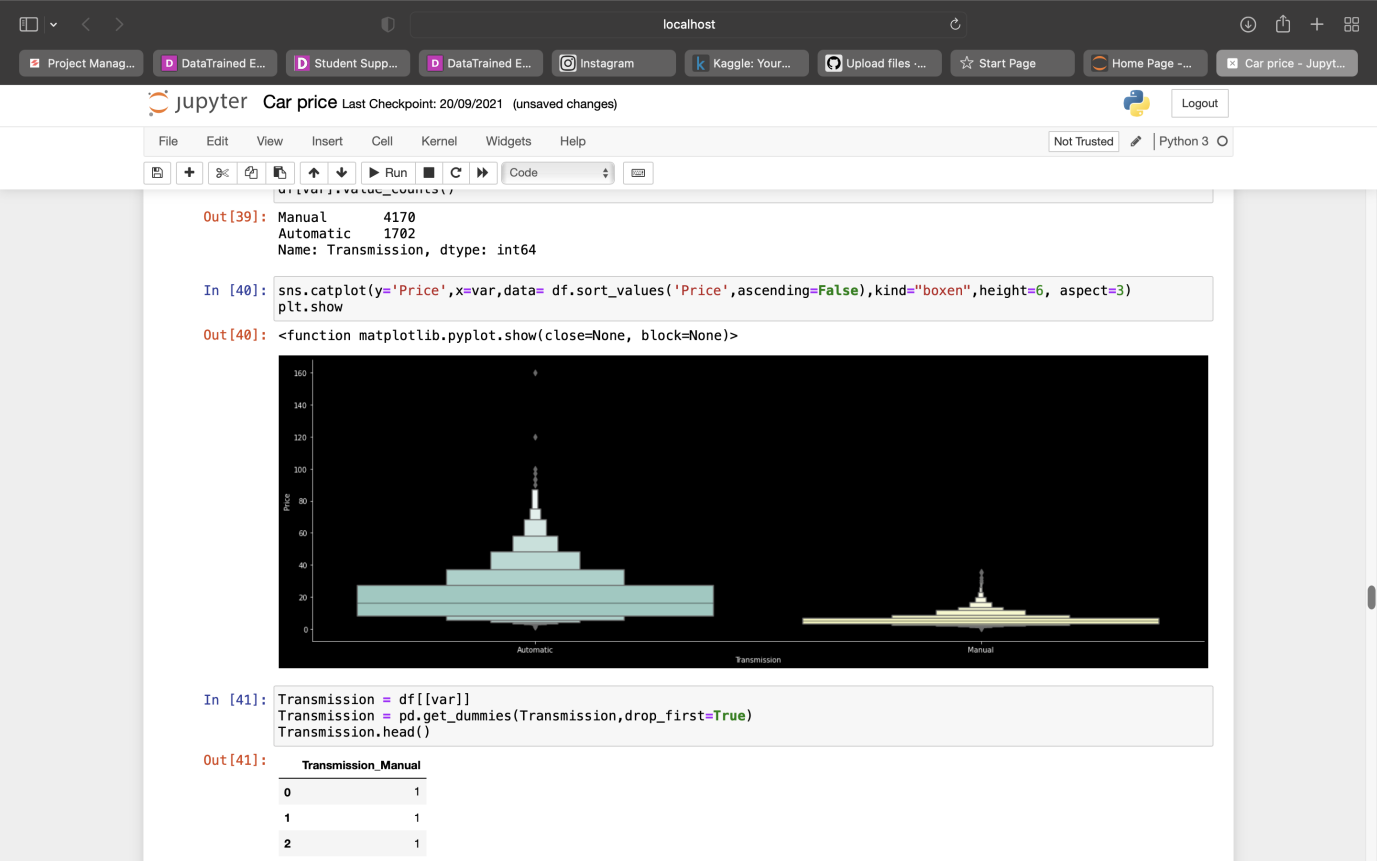
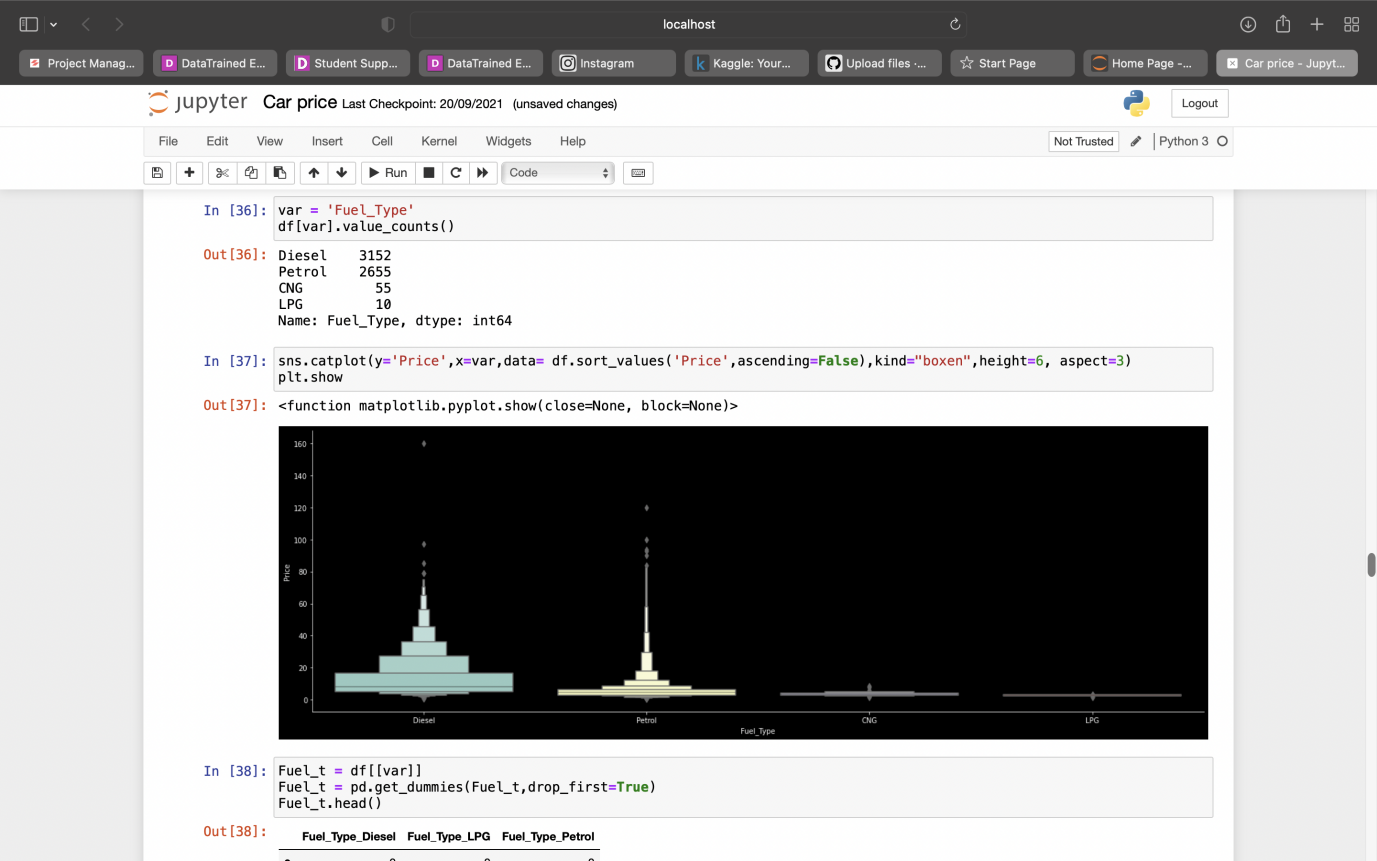
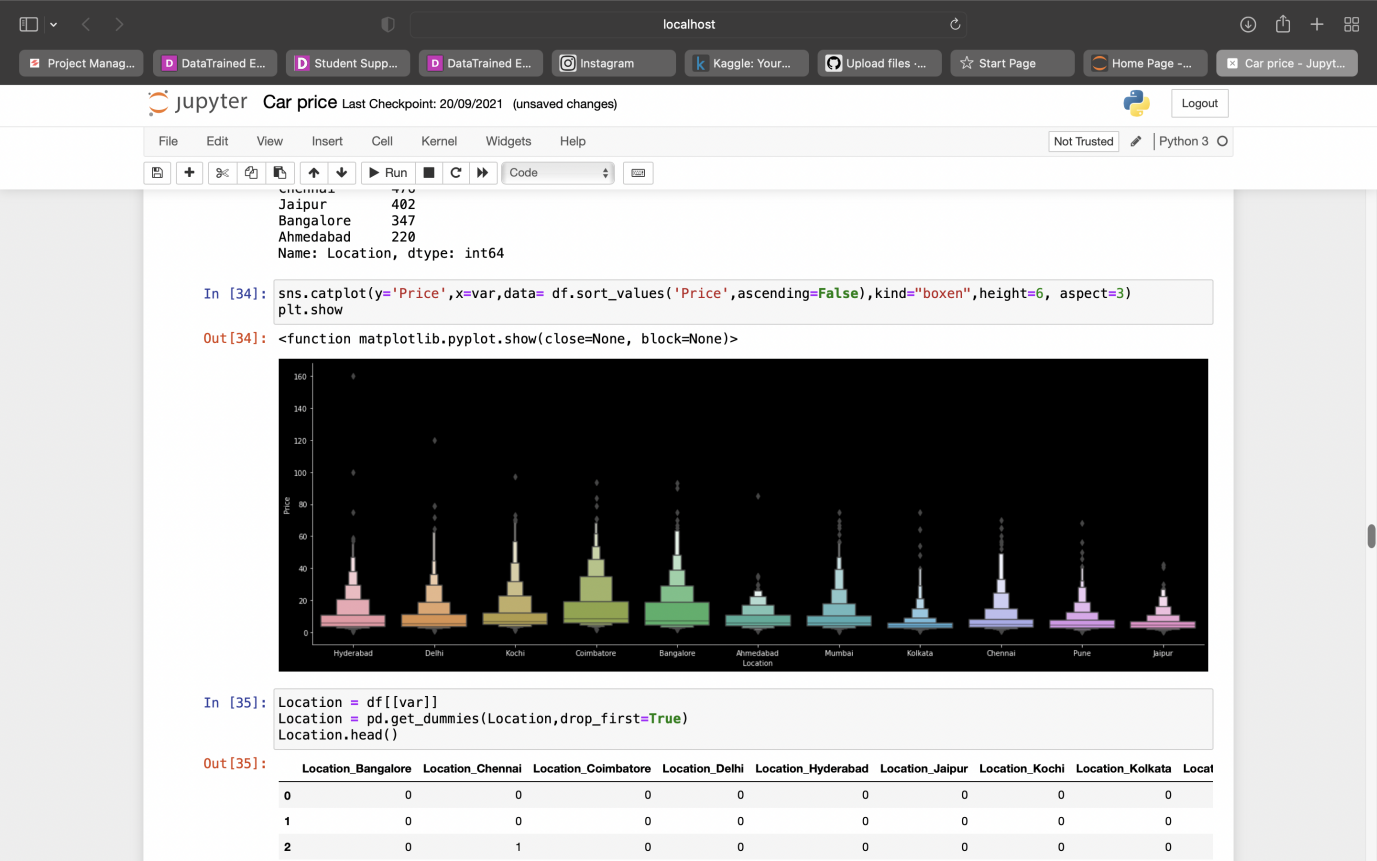
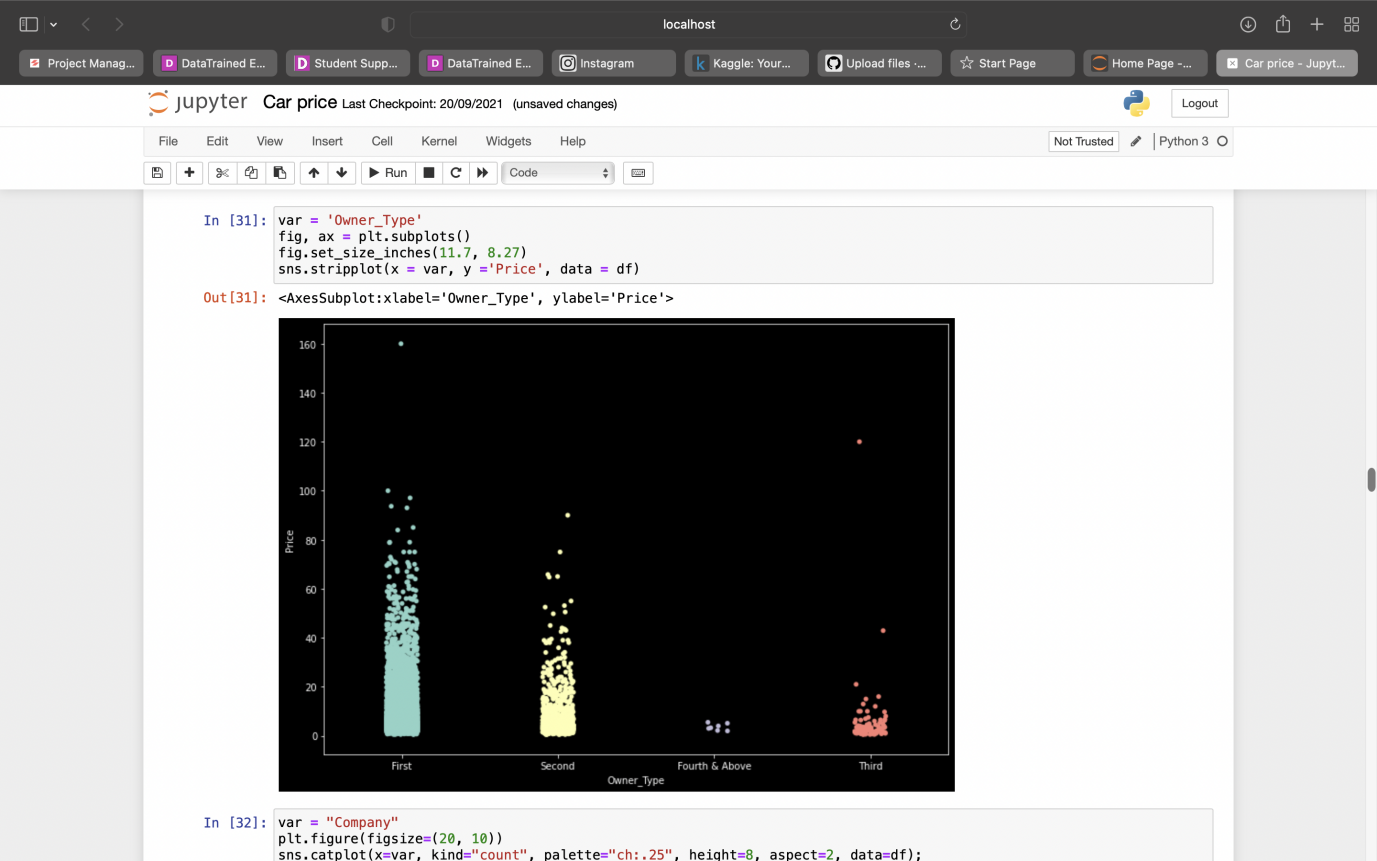
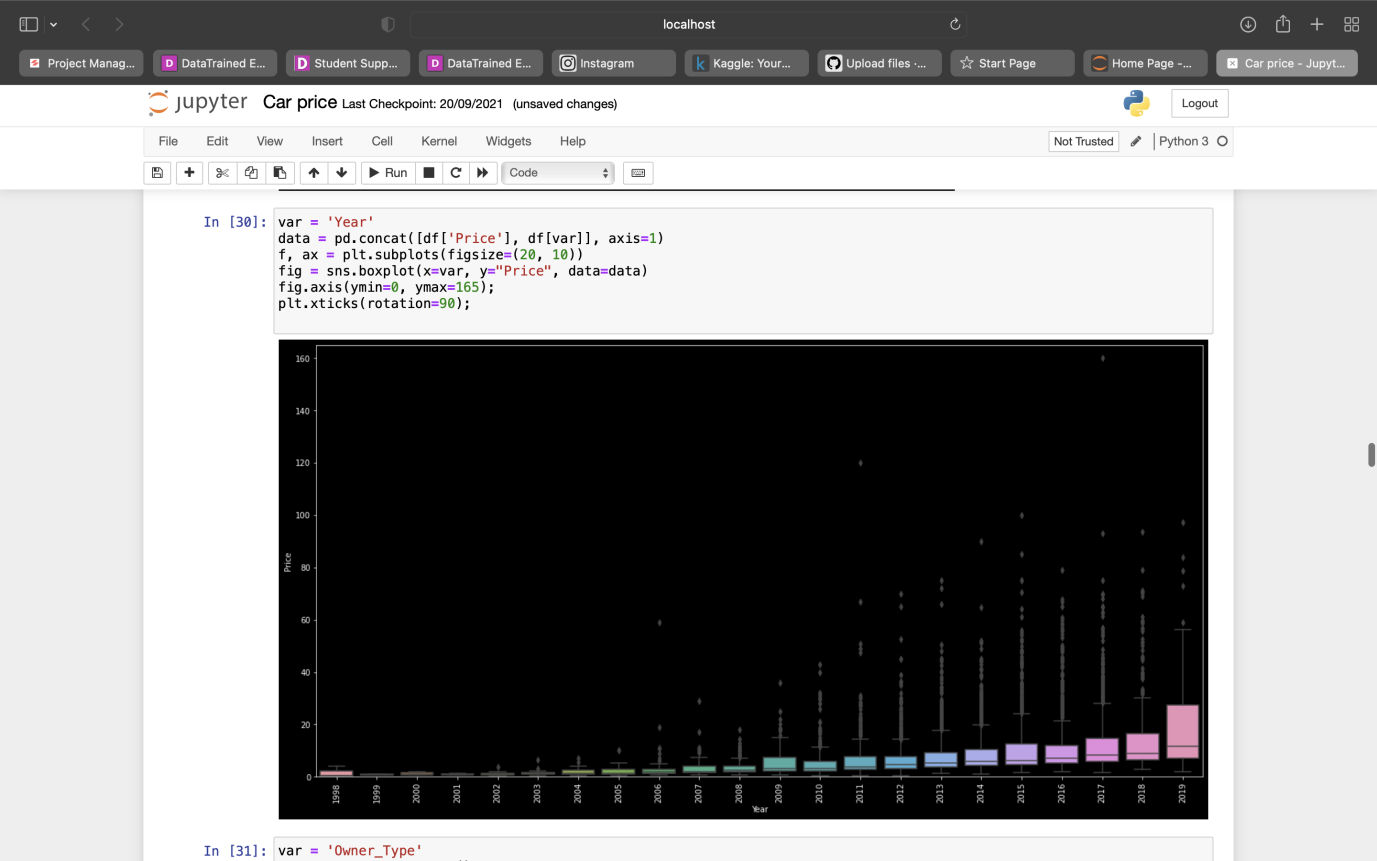
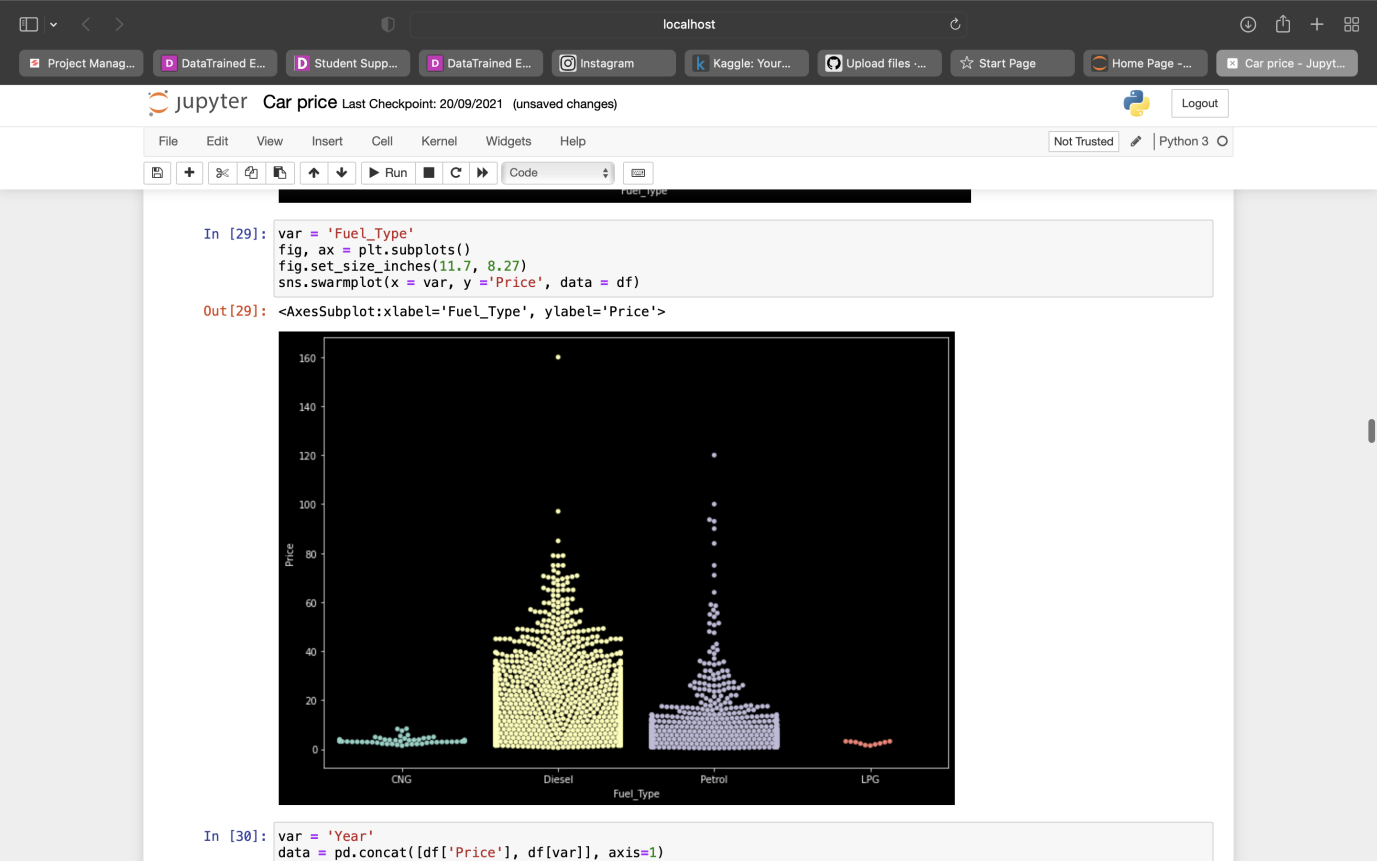
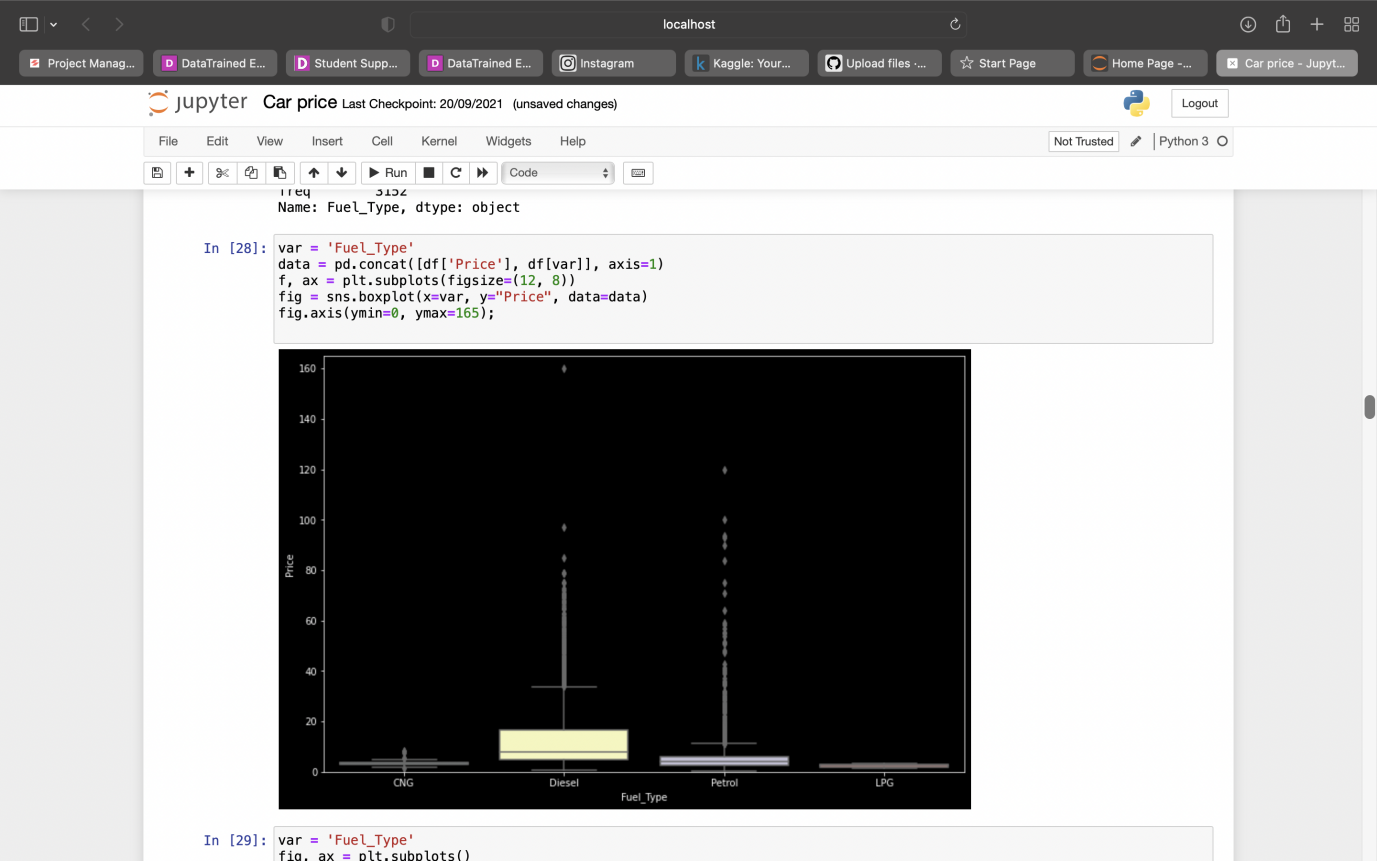
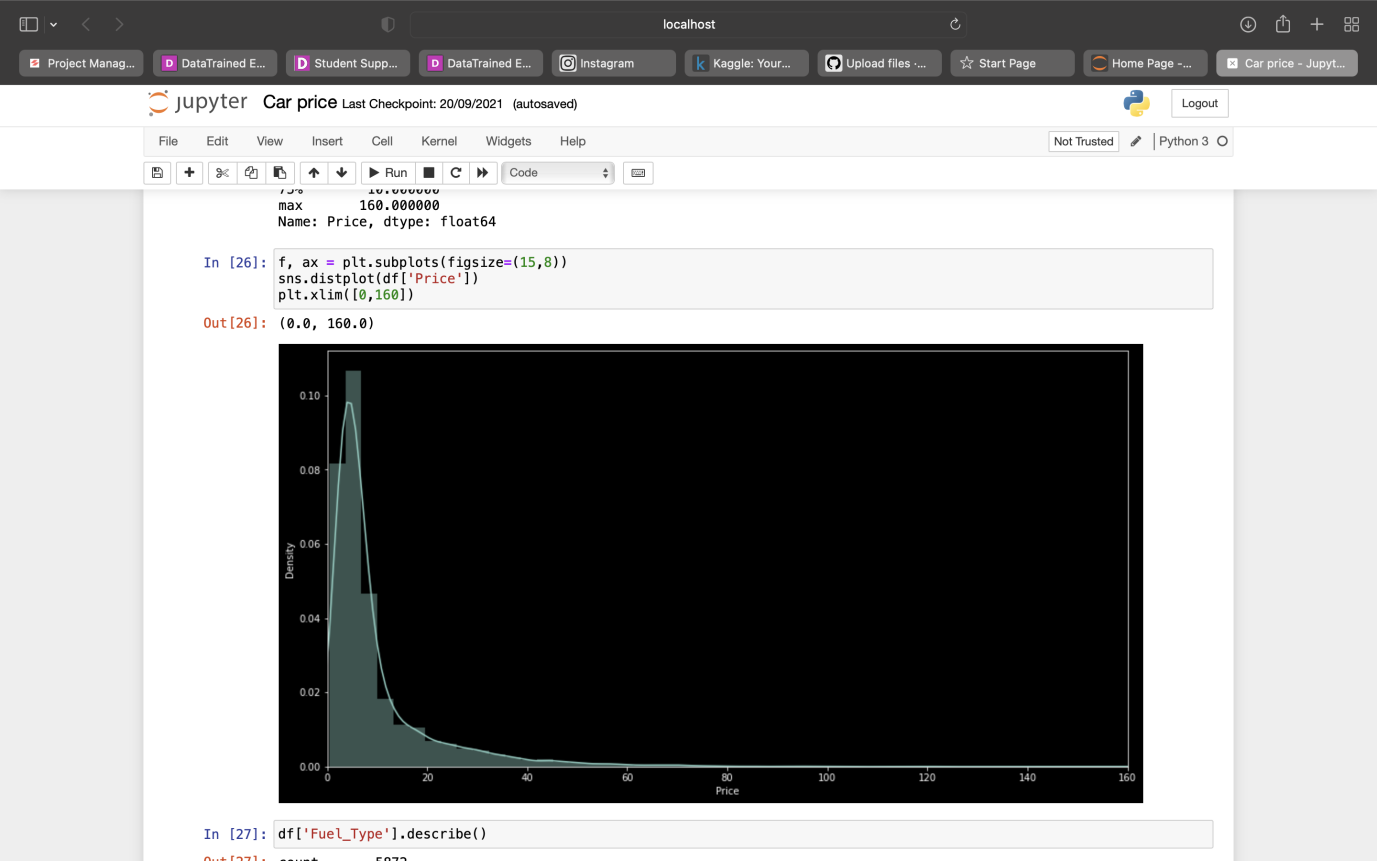
I have plotted the graph showing the price with respect to the fuel type

I have plotted the graph showing the price with respect to the transmission

I have plotted graph showing the correlation of the data

These graphs helped us getting to know the data well and also understand it on a very good and simple way.

Here are the pictures of the graphs



* Interpretation of the Results

Car prices depend upon various factors and can be assessed with the help of regression techniques as shown in the above slides.

It can be said that car industry can be better with the help of such techniques.

**CONCLUSION**

* Key Findings and Conclusions of the Study

There are many different many important factors such as region, transmission, fuel type, year of make, ownership

* Learning Outcomes of the Study in respect of Data Science

I have learned from this model that anything can be predicted with the help if data science techniques and working upon this made me believe in myself.

Visualizing the data helps in understanding the data better and also makes the model a clean and sophisticated look.

* Limitations of this work and Scope for Future Work

The limitations were the working environment, if we were to work together the result would have been better