Car Price Prediction Using Machine Learning Techniques

Acknowledgment

In This Project, I Have Investigated All Factors That Affect The Cars Price And Buying Decision. I Have Analyzed All The Factors That Affect The Cars Price And There Buying Decision. For This Purpose, I Collect The 5000 Cars Data From Different Regions Of India. I Am Very Thankful To My Friends And Family Who Helped Me Through This Study.

Introduction

A car price prediction has been a high-interest research area, as it requires noticeable effort and knowledge of the field expert. Predicting the price for new vehicles is a more interesting and needed problem by many users and companies this increases the efficiency of prediction techniques. So the varying prediction algorithms from machine learning suit this topic more efficiently. While predicting the price of vehicles we need entire different features and factors. The most significant feature is the brand and model of the vehicle and also the mileage plays a major role in predicting the price of the vehicle. The most popular ingredient for vehicles is the type of fuel and the volume of fuel in which it consumes for each mile and steering type, number of sets in the car, fuel type and numbers of the owner, transmission type, gear type, drive type, and the most important factors car colors and many more. This particular data set might highly affect the price of a vehicle. And also we need to consider the price of fuel because it may changes frequently. Now we can consider four-wheelers also. Varying more features might require four-wheelers than two-wheelers. The variety of features such as exterior color, door number, type of transmission, dimensions, safety, air condition, interior, whether it has navigation or not will also influence the vehicle price. In this paper, we applied different methods and techniques in order to achieve higher precision of the used vehicle price prediction.

Problem Statement

The Automobiles Industries Emerged Rapidly From Past Decades And It Has Grown Faster During Past Years. The Competition Has Increased. In The Present Market Giants Like Tata Motors, Maruti Motors And Many More. Company Is Trying To Capturing Indian Market Rapidly.

- In Such Competitive Market It's Very Difficult To Increase The Sale And Retaining Customers And Winning Trust Is Very Big Challenge Because The Market Changed By Day To Day.
- Understanding Various Factors That Influence Price Of The Cars.
- Understanding Customer's Behaviors Before Buying A Car.

Dataset For This Project

For This Project, We Are Using The Dataset Is Collected By The Cardekho.Com Here We Collect The Around 5000 Cars Data Which Contains The 20 Features And We Collect The Data From Different States So The Model Shows The Right Prediction.

Conceptual Background Of The Domain Problem

Our Main Problem Statement Is To Investigate The Factors Affecting That Impact The Car Price And Buying Decision. Due To Competitive Market, It Becomes More Challenging For Company To Select The Fair Price. During Investigation We Will Study Factors Like Brand, Steering Type, Number Of Sets In The Car, Fuel Type And Numbers Of Owner, Transmission Type, Gear Type, Drive Type And The Most Important Factors Car Colors. Also, We Check The Most Important Features And Interest Area That Impact The Price And Buying Decision. Which Are More Important For Attracting Customers And Retaining.

Hardware And Software Requirements And Tools Used

Hardware Requirements: We Need A Laptop With Minimum 4 Gb Ram And 500 Gb Of Heard Dish.

Software Requirement: We Need Anaconda Installed In Our Hardware. It Comes With All The Packages Required For Data Analysis And Visualization. Anaconda Having Jupyter Notebook, The Best Environment For Data Analysis. Library Used: We Have Used Mainly Four Libraries For Data Analysis, Mathematical Calculation, And Visualization Of Data. Numpy Is Used For Numerical Calculation And Pandas Is Used For Making Data Frame And Pre-Processing Of The Data. For The Visualization Part, We Have Used Matplotlib And Seaborn Package. Both The Package Provide A Wide Variety Of The Graphs For Data Visualization And Analysis. Coding Language: Python

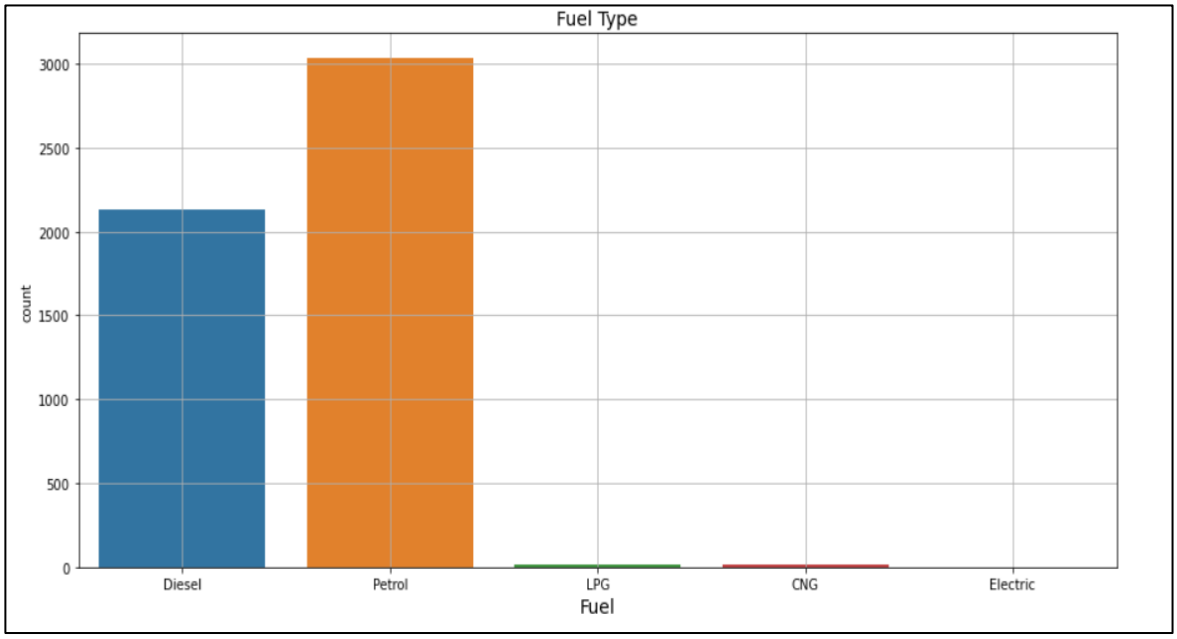
Data Analysis :

In This Section Of The Report, We Will Analyses The Data And Visualize It Using Tools Available In Python. In Jupyter Notebook We First Import All The Required Libraries. We Need Numpy For Numerical Calculation, Pandas Data Frame To Read And Process The Data Files. Our Dataset Has 5197 Rows And 20 Columns. Once We Check The Data Types Of The Columns, We Found 1 Column Has Float32, 3 Columns Are Float64, 5 Columns Are Int32. 1 Column Is Int64 And Rest 9 Columns Are Object Data Types. Data Having Some Null Values In Drive_Type, Steering_Type, Front_Break_Type, Rear_Break_Type And Acceleration Columns.

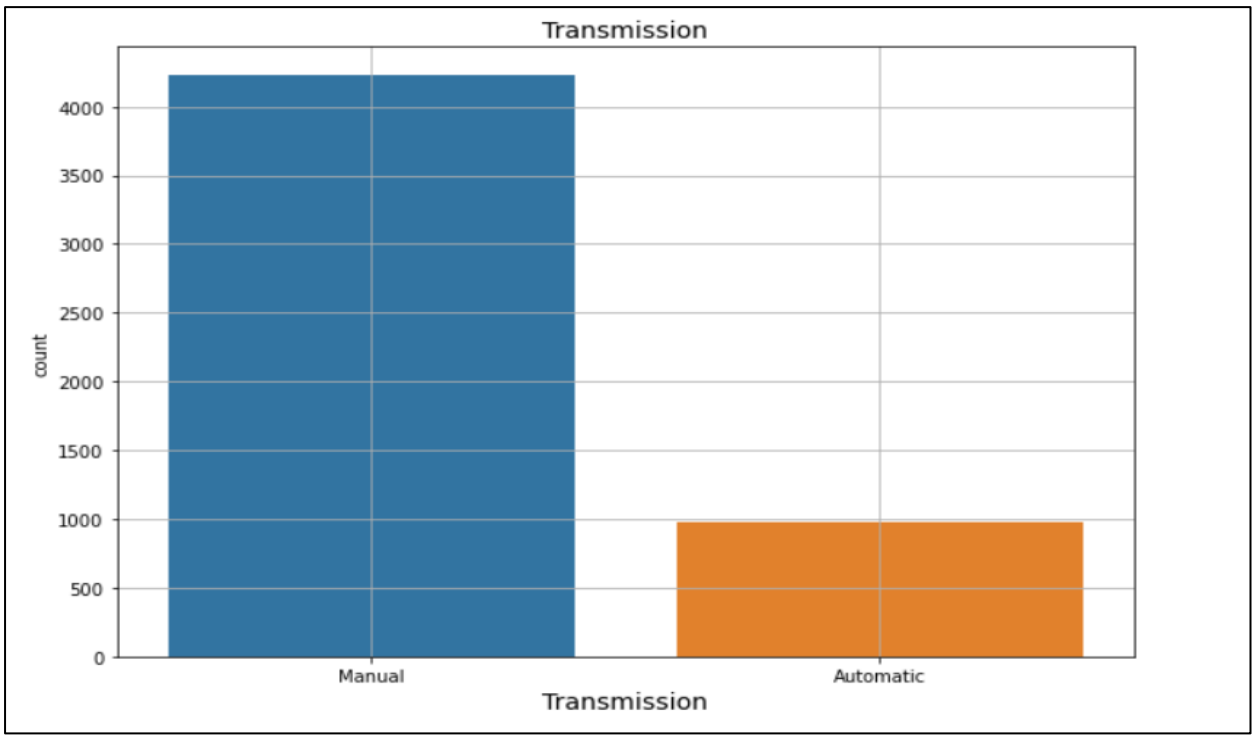
```
1 #Checking the informaion about the dataset
2 df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5197 entries, 0 to 5196
Data columns (total 19 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Brand                5197 non-null   object
 1   Make_Year            5197 non-null   int32
 2   Fuel                 5197 non-null   object
 3   KMs_Driven           5197 non-null   int32
 4   Engine_Displacement  5197 non-null   int32
 5   No_Of_Owner          5197 non-null   object
 6   Transmission         5197 non-null   object
 7   Mileage              5197 non-null   float32
 8   Max_Power            5197 non-null   float64
 9   Torque               5197 non-null   float64
 10  Seats                5197 non-null   int64
 11  Color                5197 non-null   object
 12  Gear_Box             5197 non-null   int32
 13  Drive_Type           5104 non-null   object
 14  Steering_Type        5172 non-null   object
 15  Front_Brake_Type     5173 non-null   object
 16  Rear_Brake_Type      5173 non-null   object
 17  Acceleration         4660 non-null   float64
 18  Price                5197 non-null   int32
dtypes: float32(1), float64(3), int32(5), int64(1), object(9)
memory usage: 649.8+ KB
```

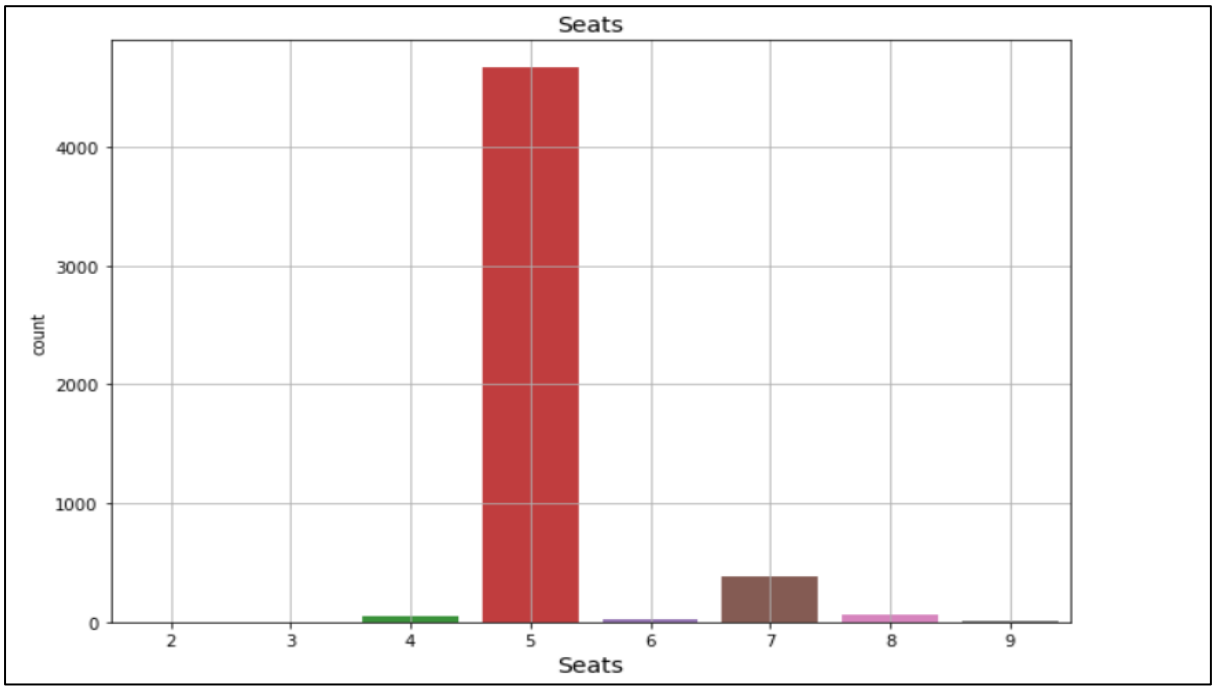
Here We See That The Petrol Fuel Cars Is Highly Demand In The Market.



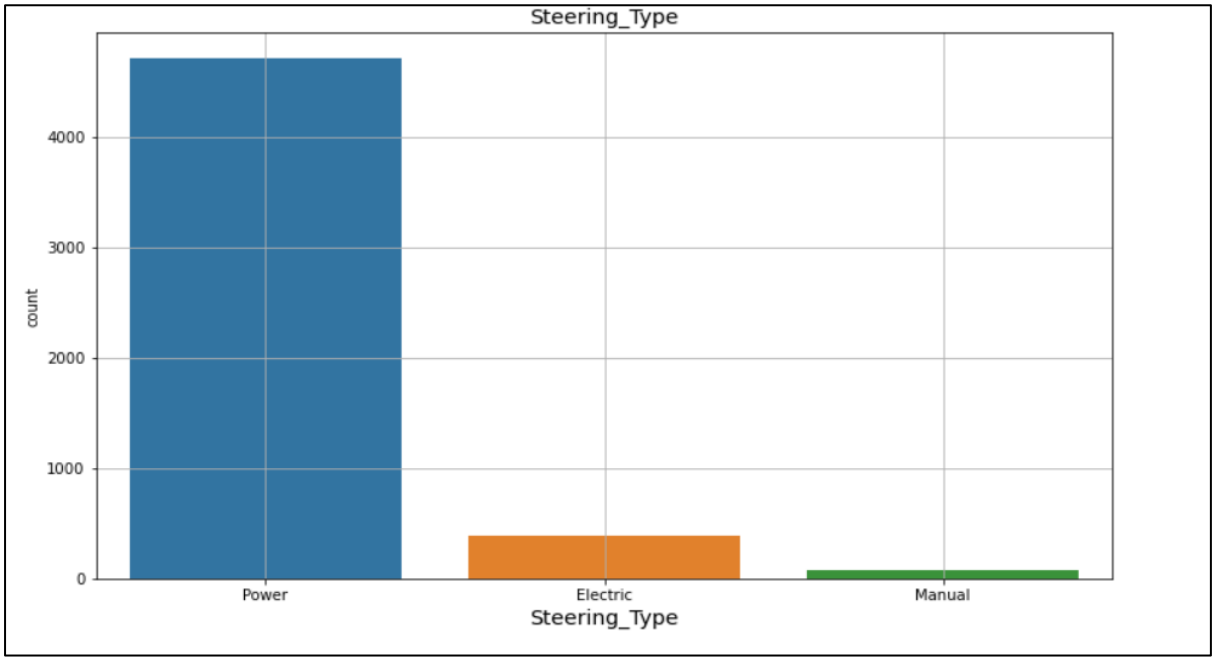
Here We See That The Transmission Of The Cars. Manual Transmission Cars I s Highly Demand In The Market



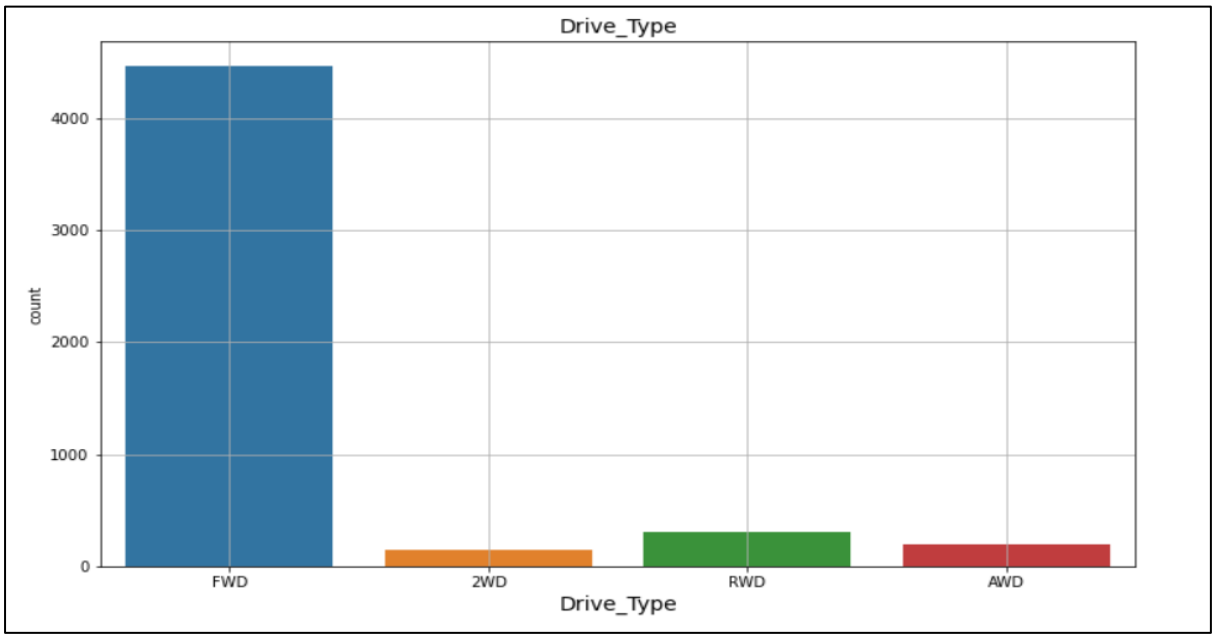
Number Of Sheets Is A Big Factor That Impact The Price Of The Car And Here We See That The 5-Seater Cars Is Highly Demanded In The Market.



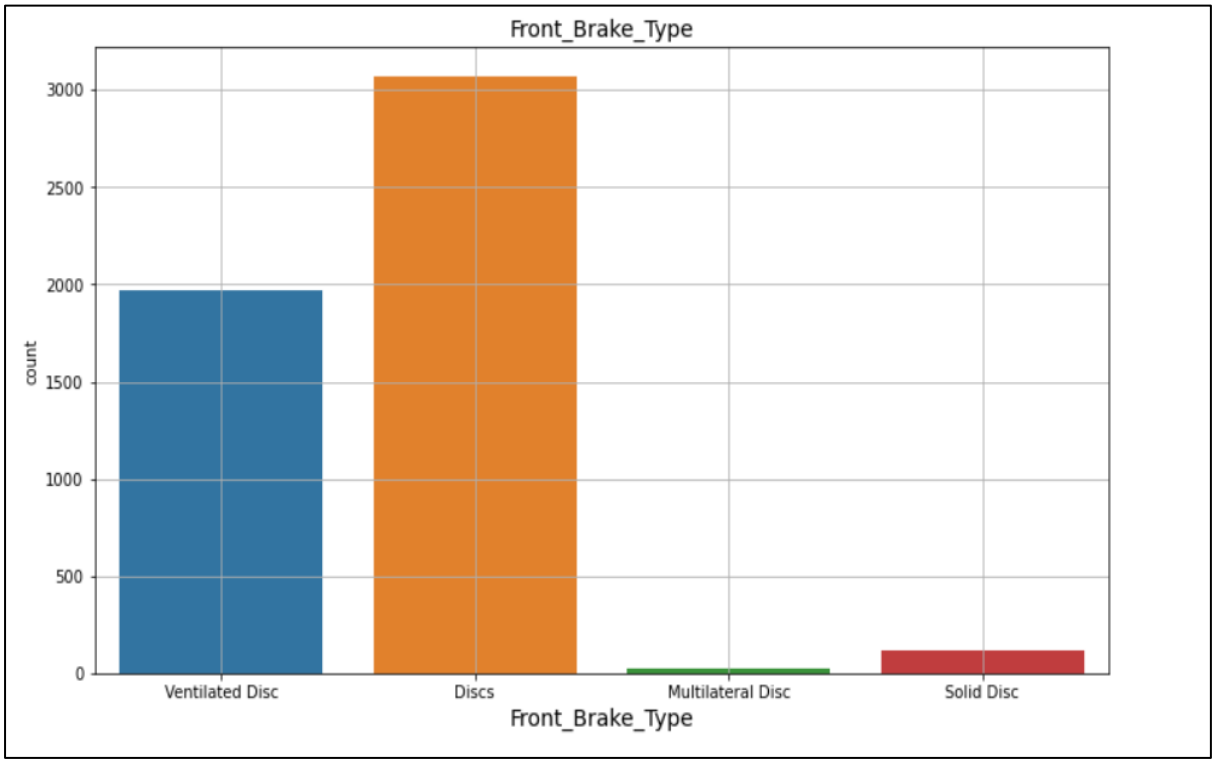
Here We See That The Power Steering Cars Is Highly Demand In The Market



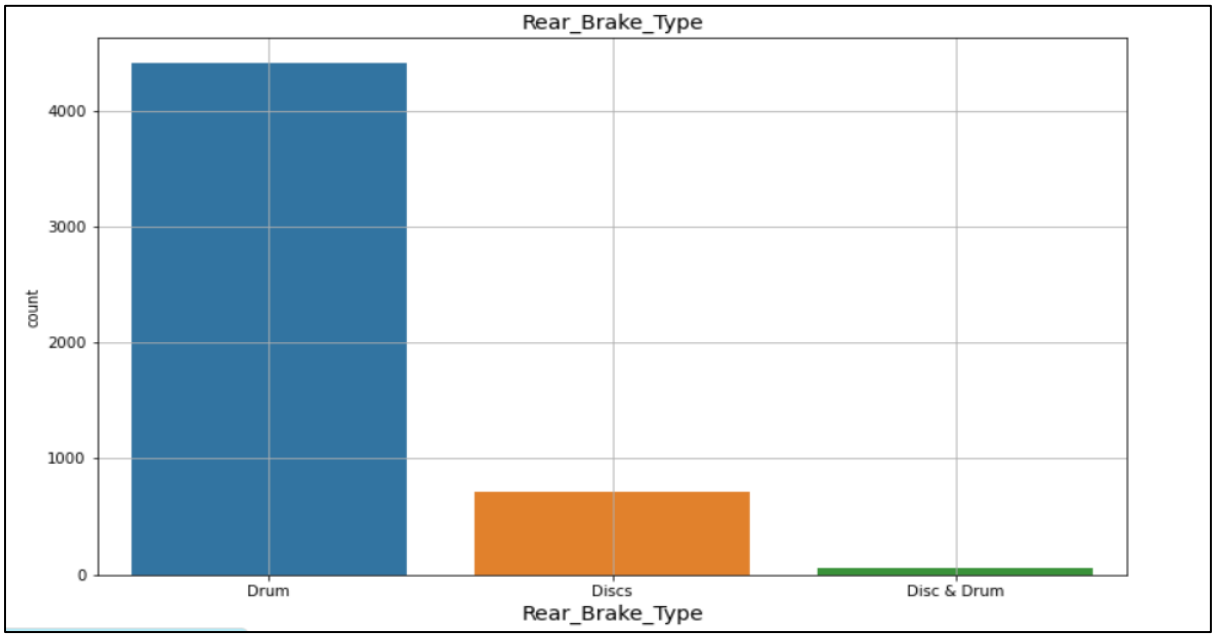
Here We See That The Front Wheel Drive Cars Is Highly Demand In The Market



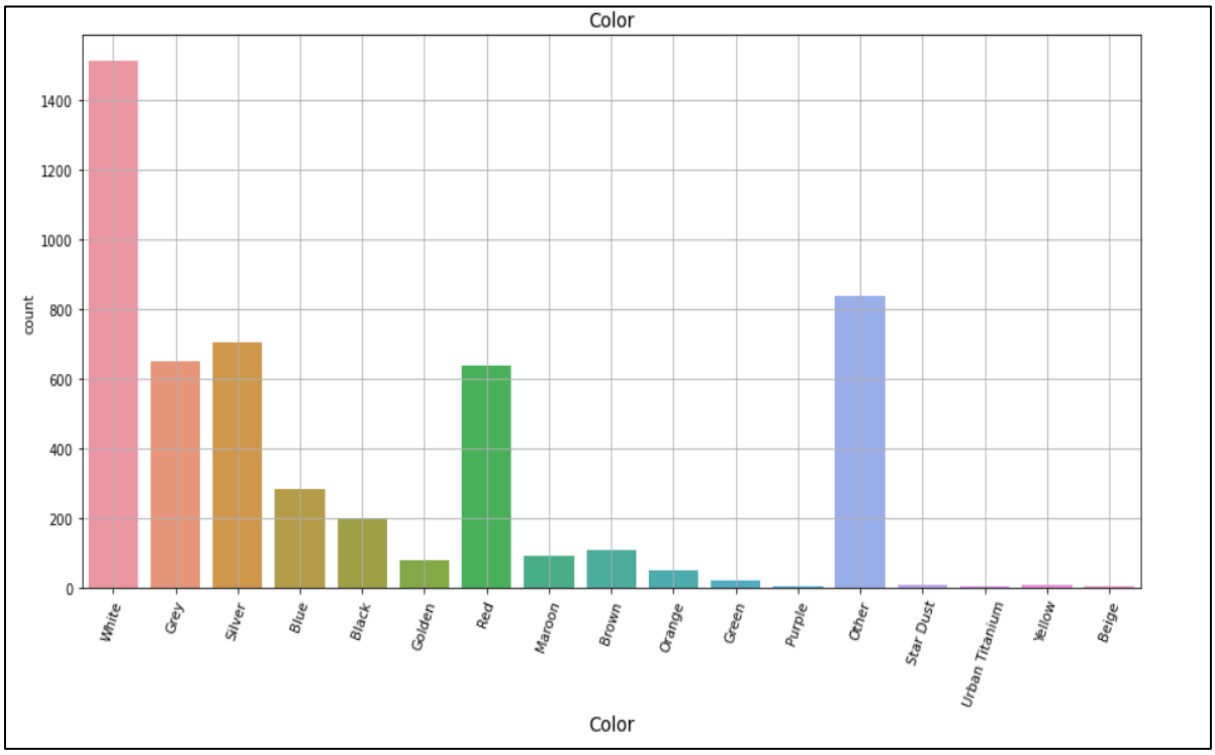
In This Chart, We See That The Front Brake Types Of Cars Is Highly Demand I n The Market And We See That The Discs Brake Types Is Highly Demanded.



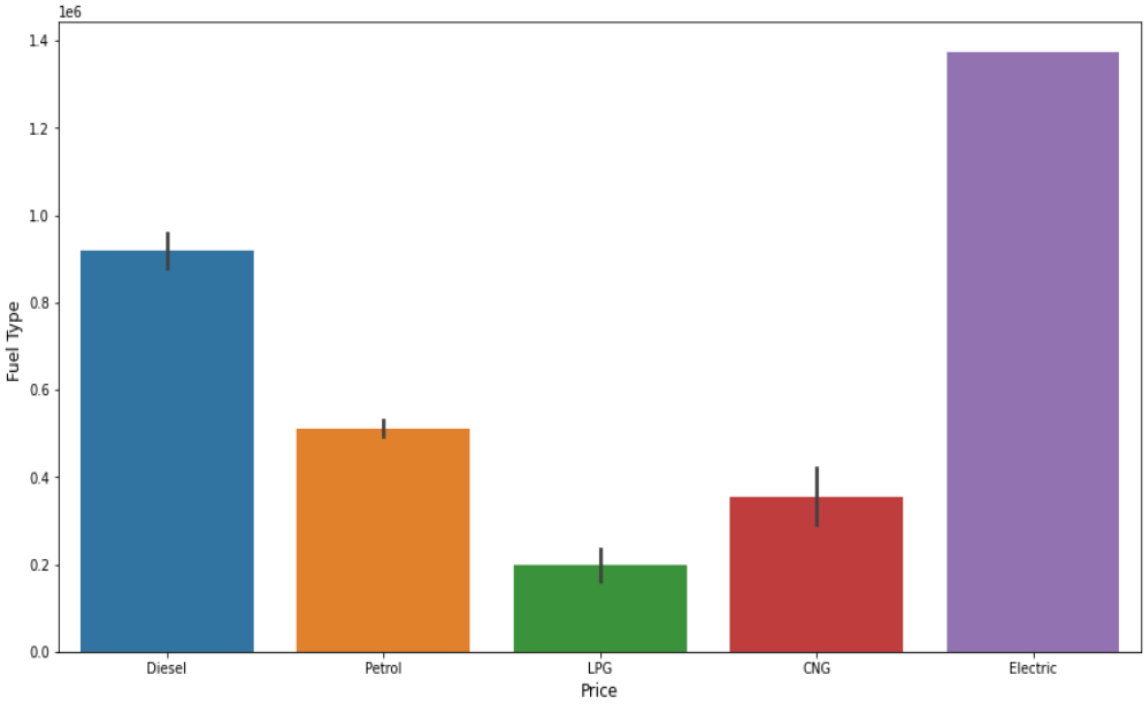
In This Chart, We See That The Rear Brake Types Of Cars Is Highly Demand I n The Market And We See That The Drum Brake Types Is Highly Demanded.



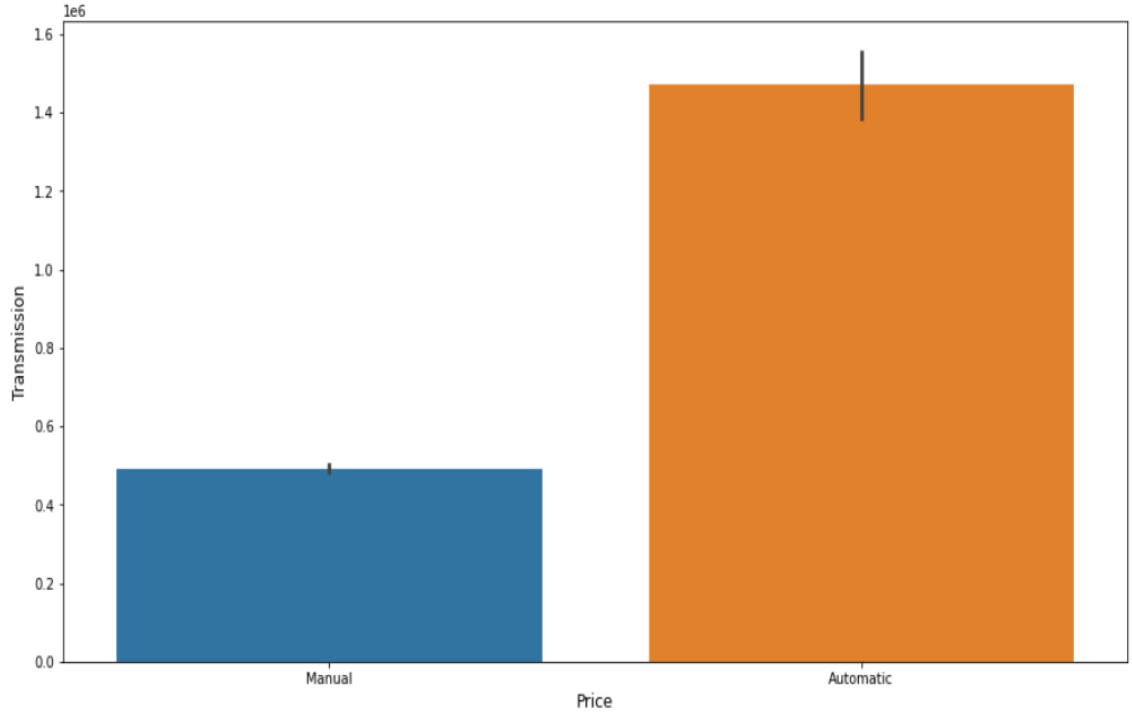
In This Chart, We See That The Color Types Of Cars Is Highly Demand In The Market And We See That The White Color Cars Is Highly Demanded.



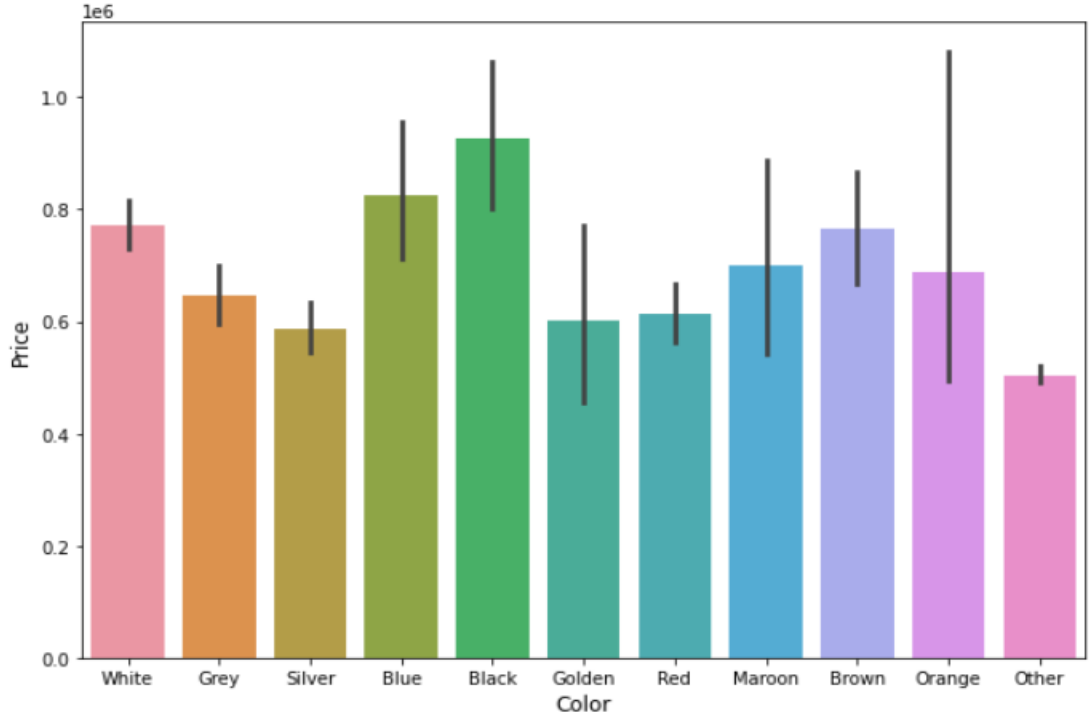
In This Chart, We See That The Electric Cars Is High Price In The Market.



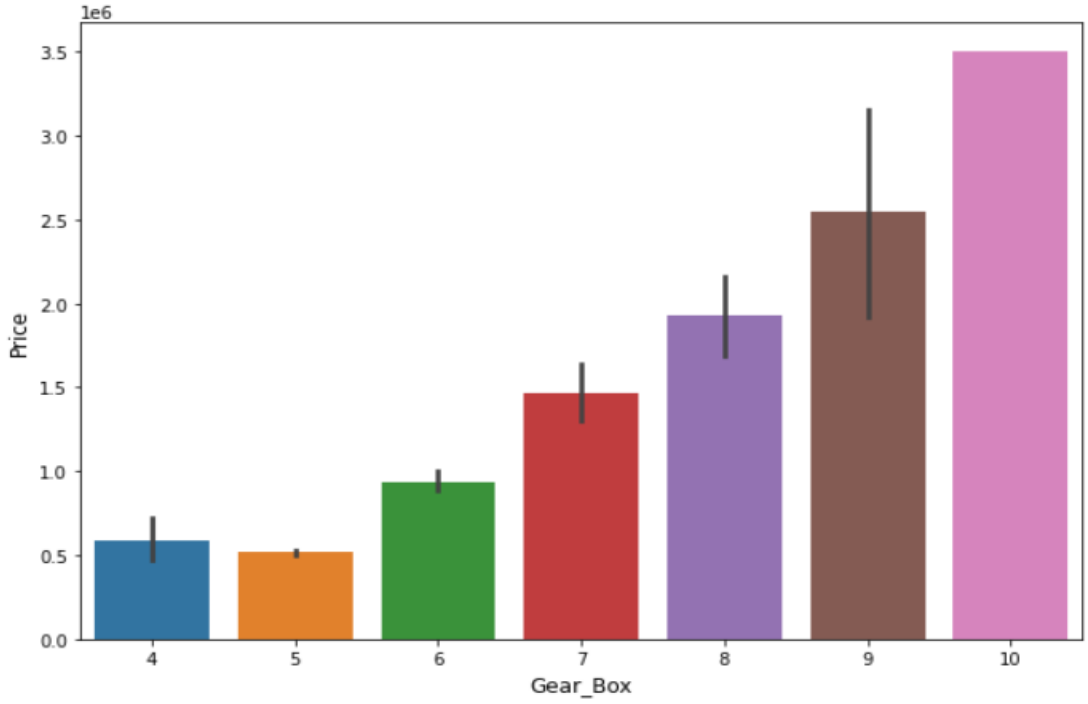
In This Chart, We See That The Automatic Cars Is High Price In The Market.



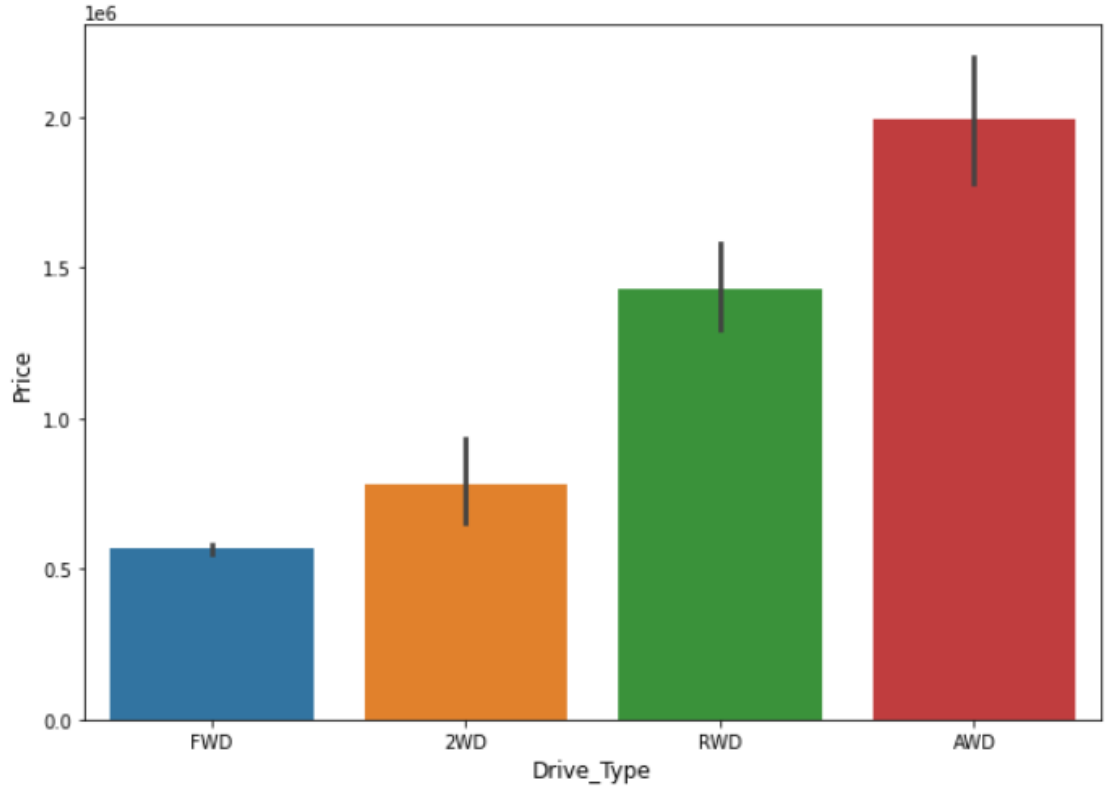
In This Chart, We See That The Black Color Cars Is High Price In The Market.



In This Chart, We See That The Which Cars Have Higher Gear Box That Cars Is High Price In The Market.

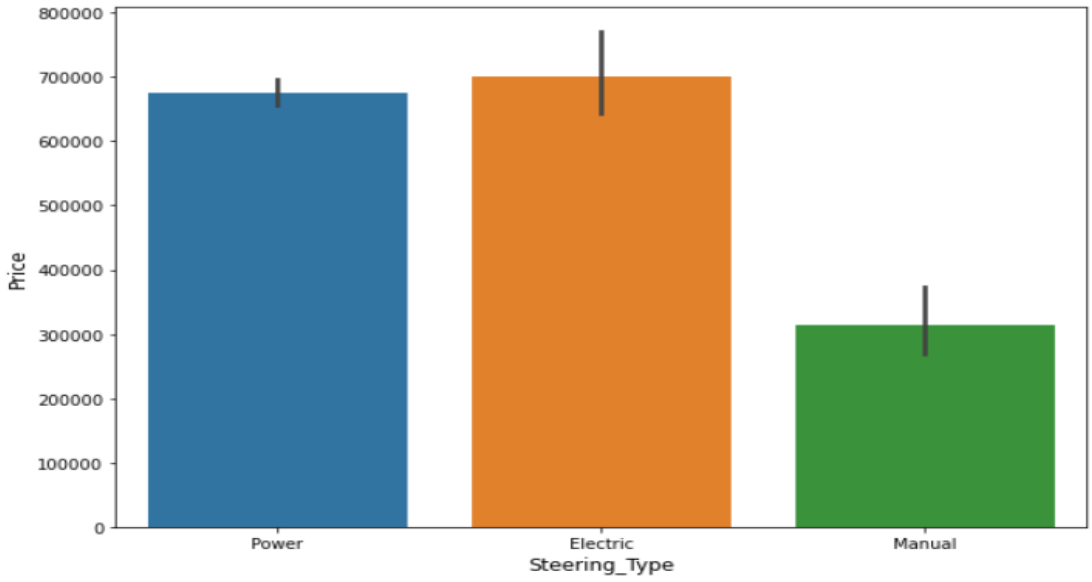


In This Chart, We See That The All-Wheel-Drive Cars Is High Price In The Ma rket.

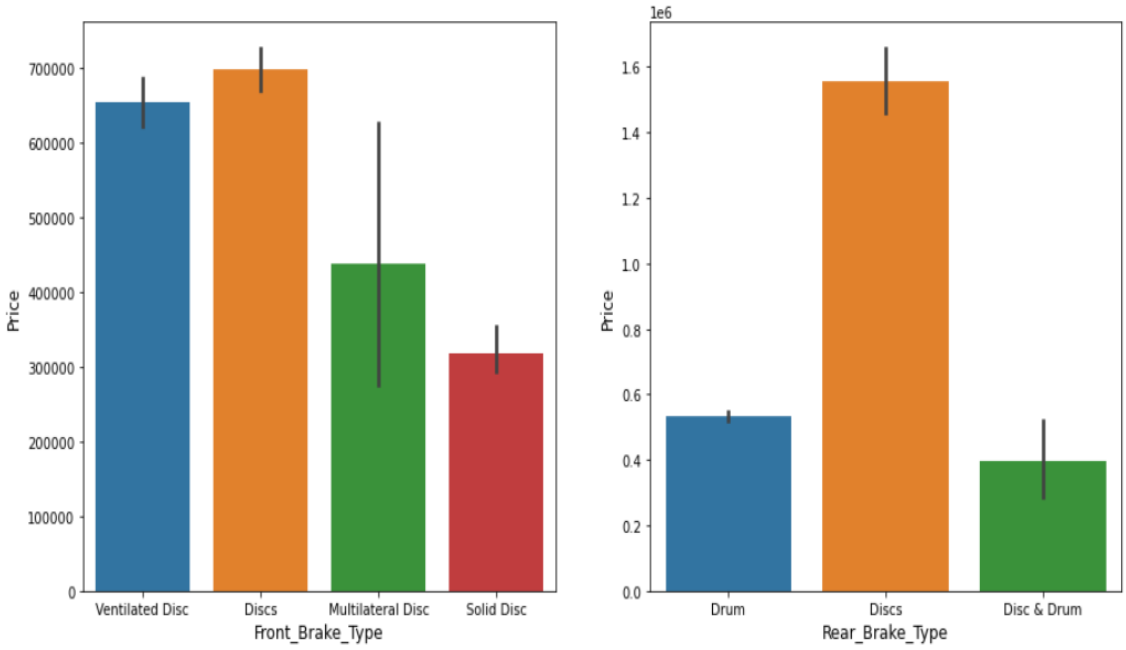


In This Chart, We See That The Electric Steering Type Cars Is High Price In The Market

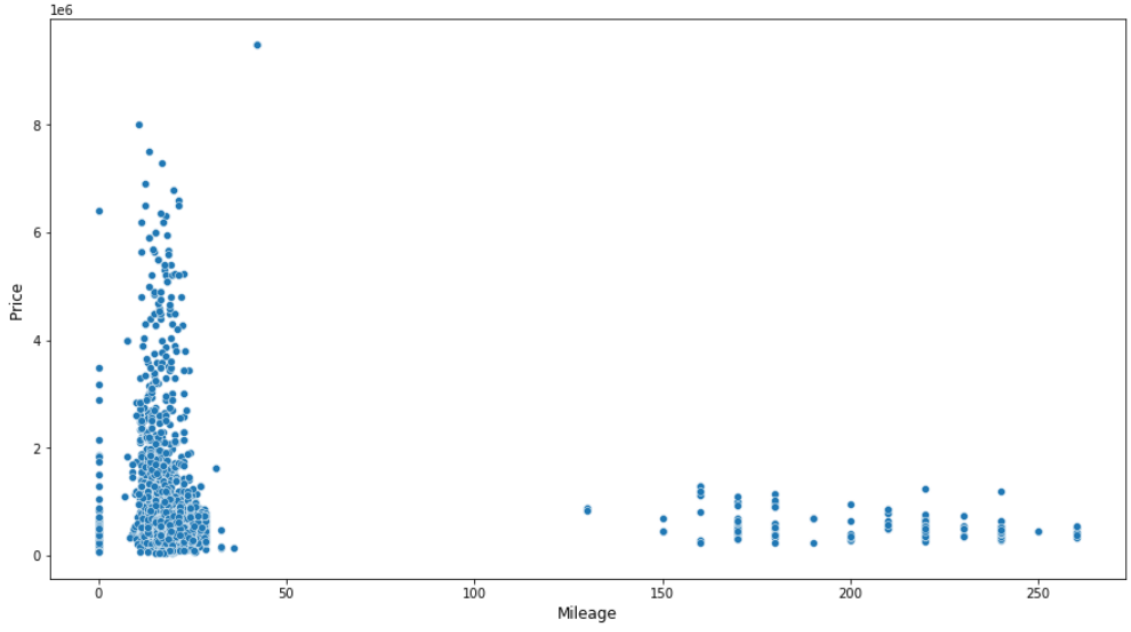
.



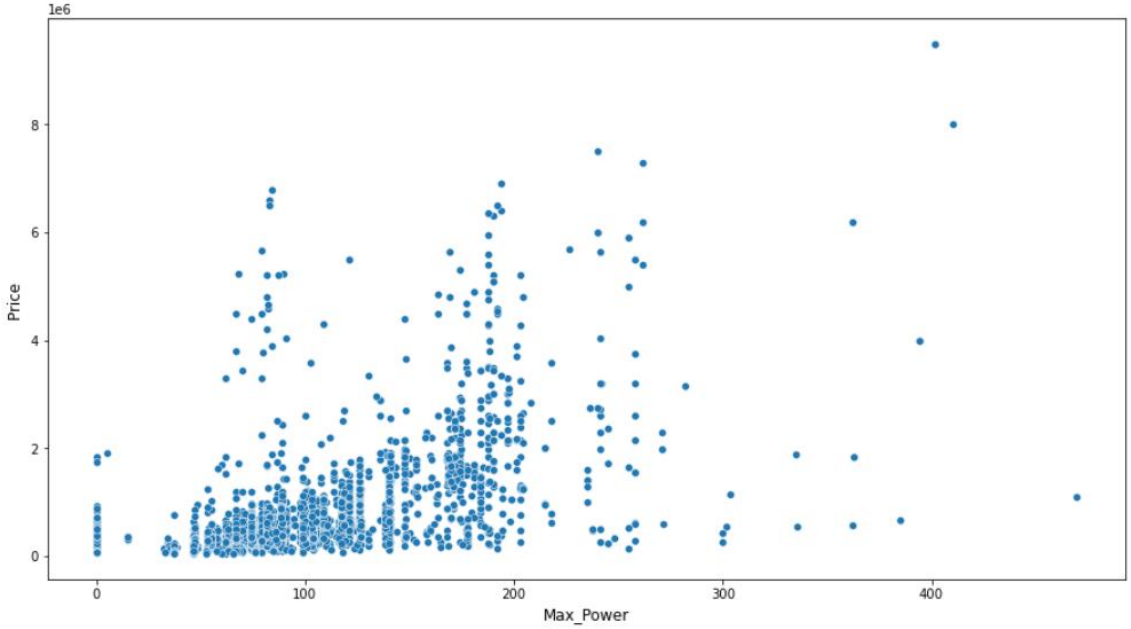
In This Chart, We See That Dics Brake Type Cars Is High Price In The Market



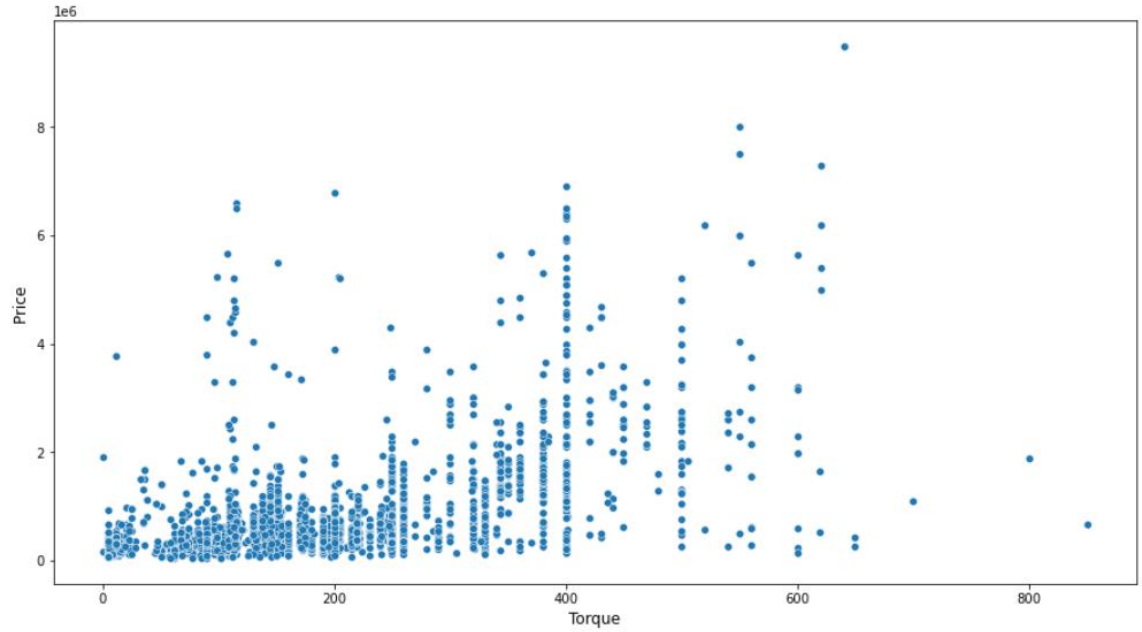
In This Chart, We See That Those Cars Have Low Mileage They Cars Is High Price In The Market.



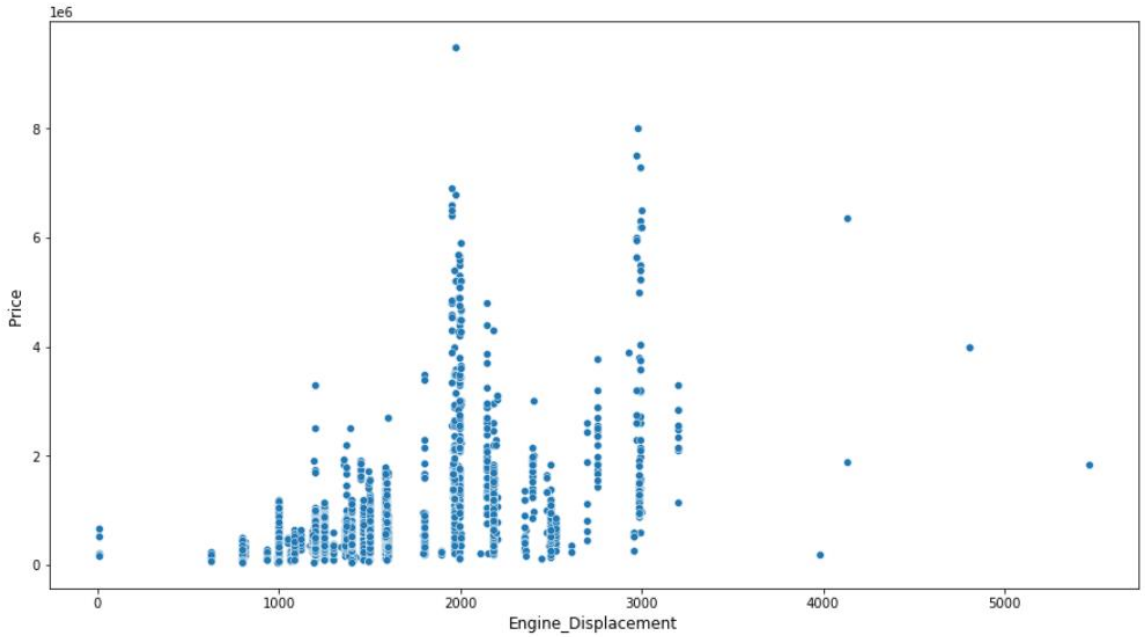
In This Chart, We See That Those Cars Have High Power They Cars Is High Price In The Market.



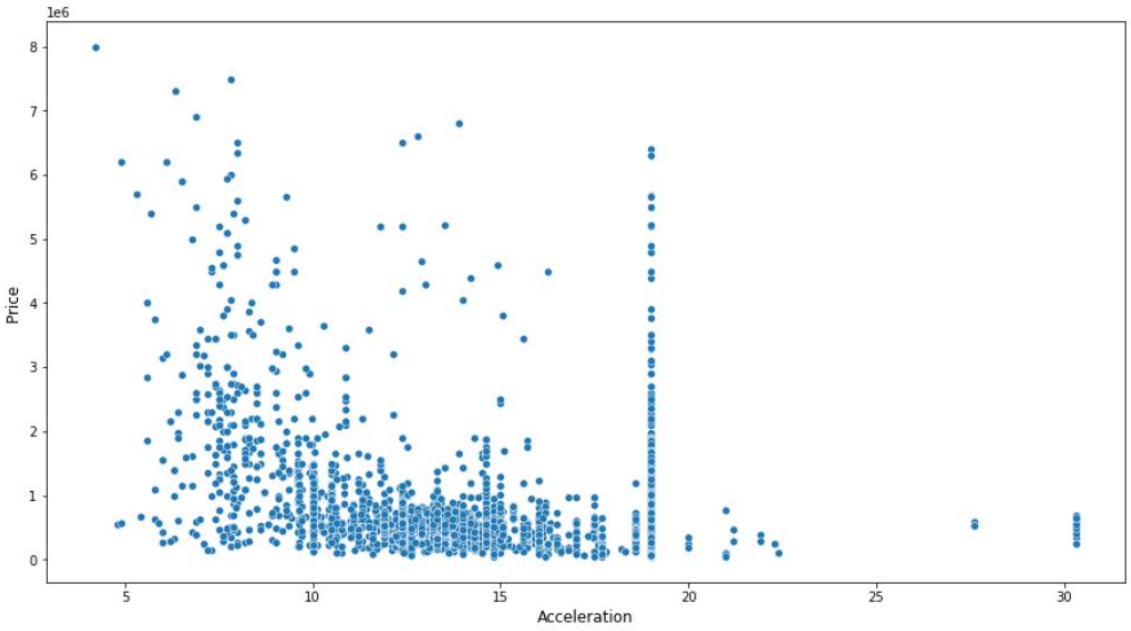
In This Chart, We See That Those Cars Have High Torque They Cars Is High Price In The Market.



In This Chart, We See That Those Cars Have Medium Engine Displacement They Cars Is High Price In The Market.



In This Chart, We See That Those Cars Have Low Acceleration They Cars Is High Price In The Market.



CONCLUSION

- Key Findings and Conclusions of the Study

  With the increase in automobiles, Every company trying to capture the marketplace, and automobiles company capture the market by launching new cars that is based on client feedback so they can increase the same and also retain the client and Give the best price. By performing different ML models, we aim to get a better result or less error with max accuracy

- Learning Outcomes of the Study in respect of Data Science

  It was very interesting to study and visualize the data using tools available in python. We have visualized the data very deeply and come to the conclusion of what customers need and what factors cause to select the car price and retention of the Client.

- Limitations of this work and Scope for Future Work

  The data set was very small and all the conclusions are based on this small dataset. For better and more clarity, we can perform the same steps on the big dataset to make some clear and more accurate decisions.