DATA ANALYSIS

Japan’s Largest Online Used Car Marketplace

Contents

Table Of Figures 1
Introduction 2
Average Car price through the years/Correlation(Avg Price – Years) 2
Average Mileage of Cars through Years 3
Average Engine Capacity through the Years 3
Average Car Price by Brand 4
No.of Cars distribution according to the engine capacity 5
Average Price of Cars by Engine Capacity(Correlation Avg Price – Engine) 5
No. of Cars distribution by Mileage 6
Average Car Price by Mileage(Correlation Mileage – Avg Price) 7
Fuel-type of Cars 7
Transmission type of Cars 8
Average Car Price by Model(Correlation Average Price – Model) 8
Predictor Model For Average Car Price 9

Table Of Figures

Figure 1- Average Car price through the years 2
Figure 2- Average Mileage of Cars through Time 3
Figure 3- Average Engine Capacity through years 3
Figure 4- Average Car Price by Brand 4
Figure 5- No.of Cars by Engine Capacity 5
Figure 6- Average Car Price by Engine Capacity 5
Figure 7- Frequency of Cars by Mileage 6
Figure 8- Average Car Price by Mileage 7
Figure 9- Fuel-type distribution of Cars 7
Figure 10- Fuel-type through the years 8
Figure 11- Transmission distribution of Cars 8
Figure 12- Average Car price by Model 9
Figure 13- Predictor Model for Car Price 9

Introduction

This report provides a thorough analysis of the data acquired from the TC-V website which is the largest online marketplace for used cars in the Japanese market. This report not only summarizes the current findings i.e using historical data but also provides few insights related to the forecasting of car prices depending on the different variables or specifications that play an important role in determining the price of the car. The report is aided with aesthetic and comprehensive visualizations to help in the comprehension of the reader to gain logical and valuable insights for any possible business decision in the future.

Average Car price through the years/Correlation(Avg Price – Years)

The first and foremost graph that deserves a look is how the average price of the car has varied through time. This data comprises of car prices from about 1980 to about 2015. As it can be seen in the graph below, the general trend of the average price of the car has shown just a slight increase. This increase is not gradual but has emerged after a lot of highs and lows. The average car price saw a sharp decline in the time of 2008-2010(possibly due to the great recession). The average car price also saw a steep incline after the year 2011. The highest the average car price scored was about $1.1million in the year of 1990 and the lowest it scored was about $0.75million in the year of 2010. The graph is also aided by a subtle trend line which shows an increase but with a very gentle slope.

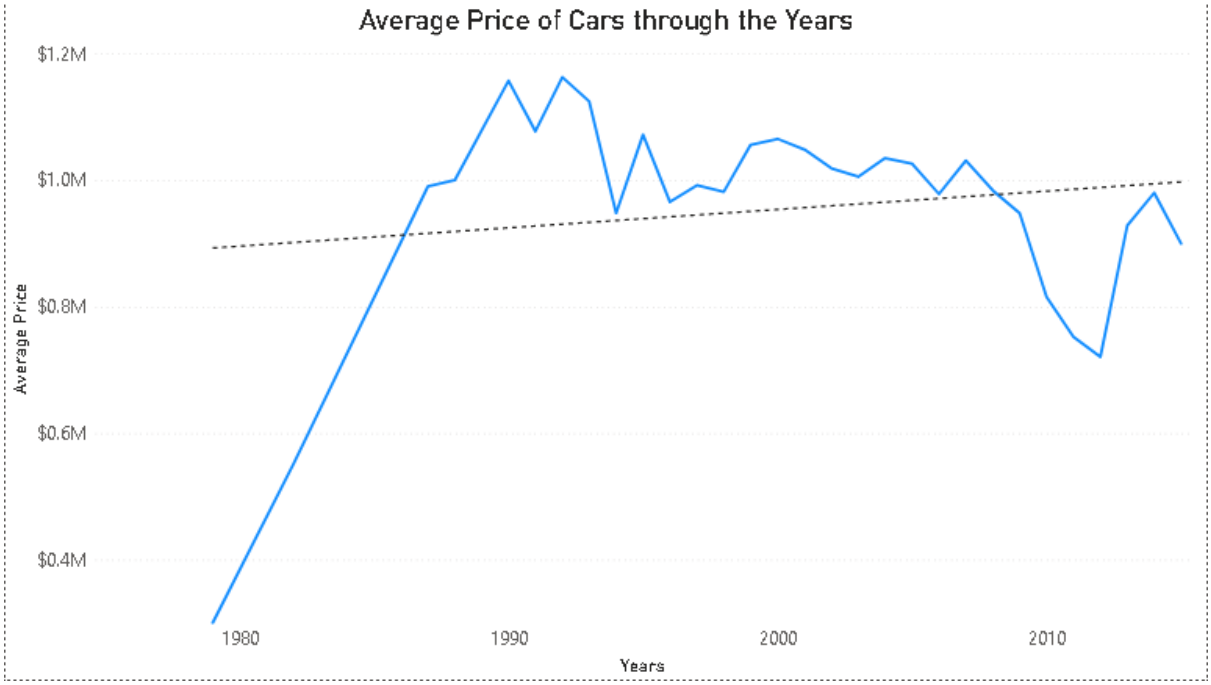

Figure 1- Average Car price through the years

The correlation co-efficient between the car price and years is -0.21109 which is skewed by the sharp declines of 2008-2009. This means that as years progress the average price of the car decreases. The negative sign indicates an opposite nature to variation and 0.21109 indicates a minimal magnitude that contributes to the change.

Average Mileage of Cars through Years

The visualization attached below is a representation of how the average mileage of cars has fared through the years. It is evident from the graph that the average mileage of the cars sold has followed a general increasing trend. This trend can also be visualized by the trend line fitted onto the graph. The year 1988 saw the highest car mileage of 137,000 Km whereas the year 1987 saw the least car mileage of 17,800Km.

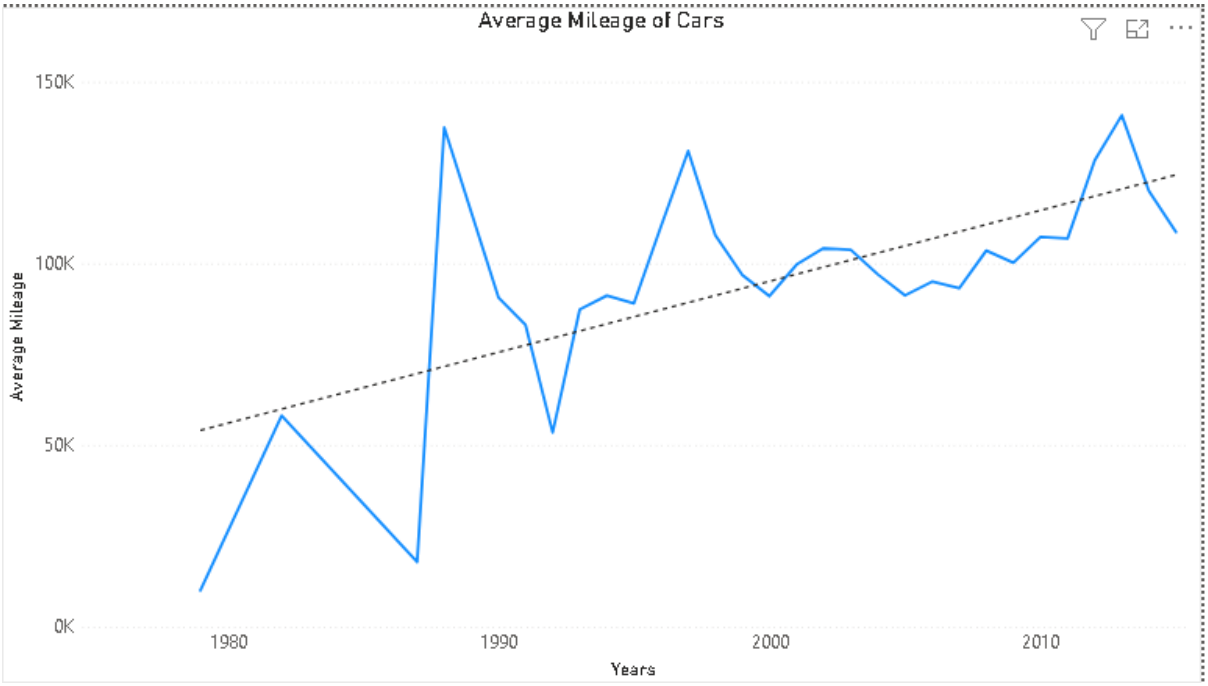

Figure 2- Average Mileage of Cars through Time

Average Engine Capacity through the Years

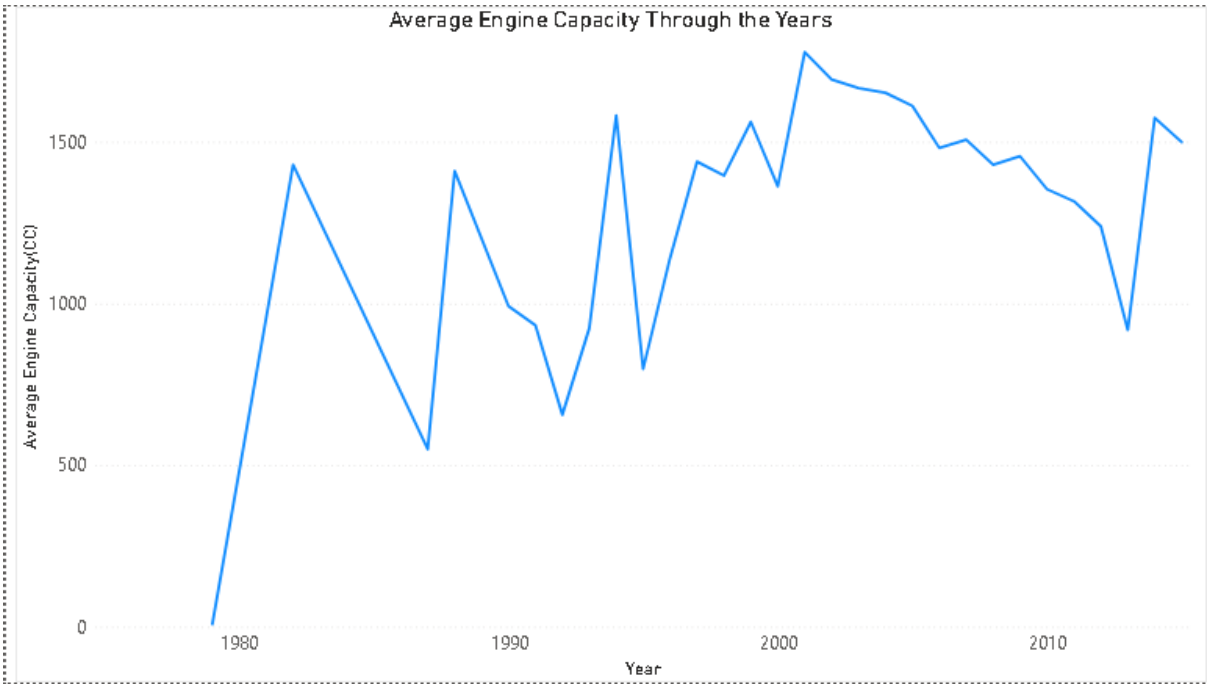

Figure 3- Average Engine Capacity through years

The visualization above shows the variation of engine capacity through the years. It can be clearly seen from the graph that the average engine capacity of the cars sold has increased through the years. It is currently at 1500cc. The highest engine capacity was recorded in the years of 1800cc in the year 2001 while the lowest engine capacity was recorded in the year 1979 of 9cc. The most recent engine capacity is rounded at 1500cc in the year of 2015.

Average Car Price by Brand

The visualization below shows the average price of cars by the brand that they belong to. It shows the distribution of all the brands where the highest average price of the car belongs to the brand Jaguar. This is about $300-$400k higher than the average price of all brands which is $1.1million. The brands that have average car price higher than the average are Jaguar,Hino,Ford,Hyundai,Volvo,Kia,BMW and Volkswagem. The brand Kubota has the lowest average car price of about $0.3million. The visualization is also aided by an average line denoting the average price of the total cars in the dataset.

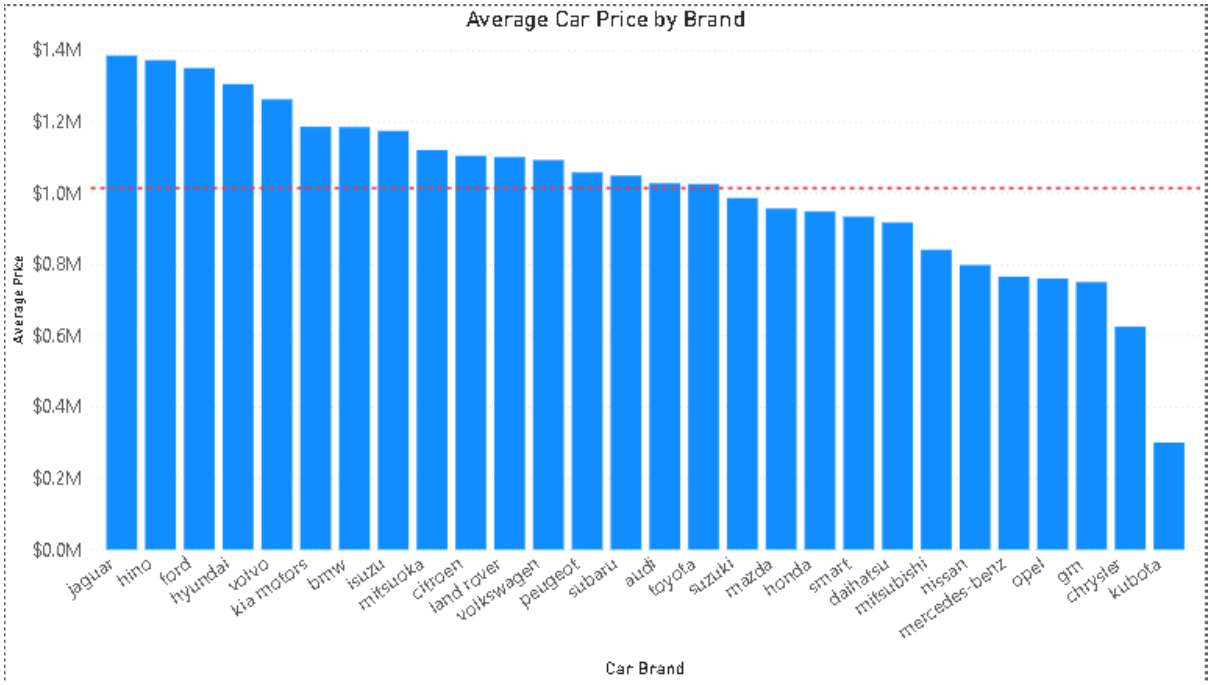

Figure 4- Average Car Price by Brand

The correlation between the average price and brand can be understood by looking at the graph above. It shows that if a car picked at random belongs to a brand such as Jaguar, Hino , Ford or Hyundai , its price would be at a much higher value then the average price of the car. Brands such as Peugeot, Subaru , Audi , Toyota and Suzuki would have car price at round about average car price whereas brands such as Kubota, Chrysler , GM would have car prices at lower than the average value.

No.of Cars distribution according to the engine capacity

The visualization below shows the number of cars according to its engine capacity in the dataset. The engine capacity between the range 1000cc to 1500cc has the highest number of cars, close to 1500. The engine capacity between 3500cc-5000cc is the lowest at about 10-15. The engine capacity with 1800cc to 2500cc has the frequency of 600 cars. The engine capacity from 0-900cc is about 200 cars.

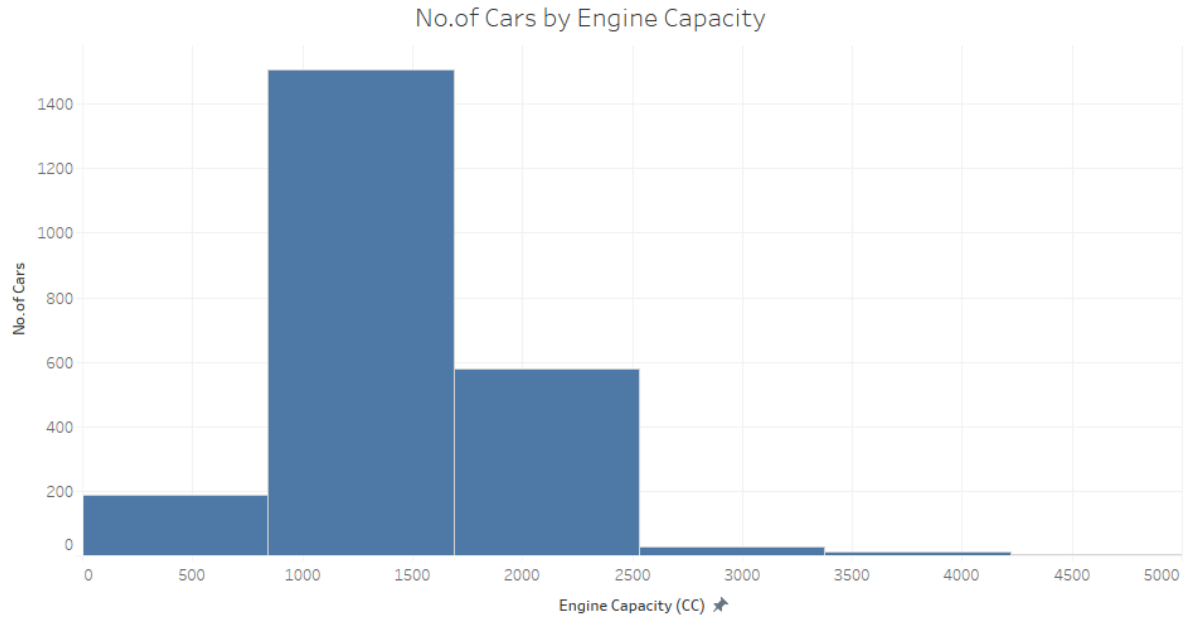

Figure 5- No.of Cars by Engine Capacity

Average Price of Cars by Engine Capacity(Correlation Avg Price – Engine)

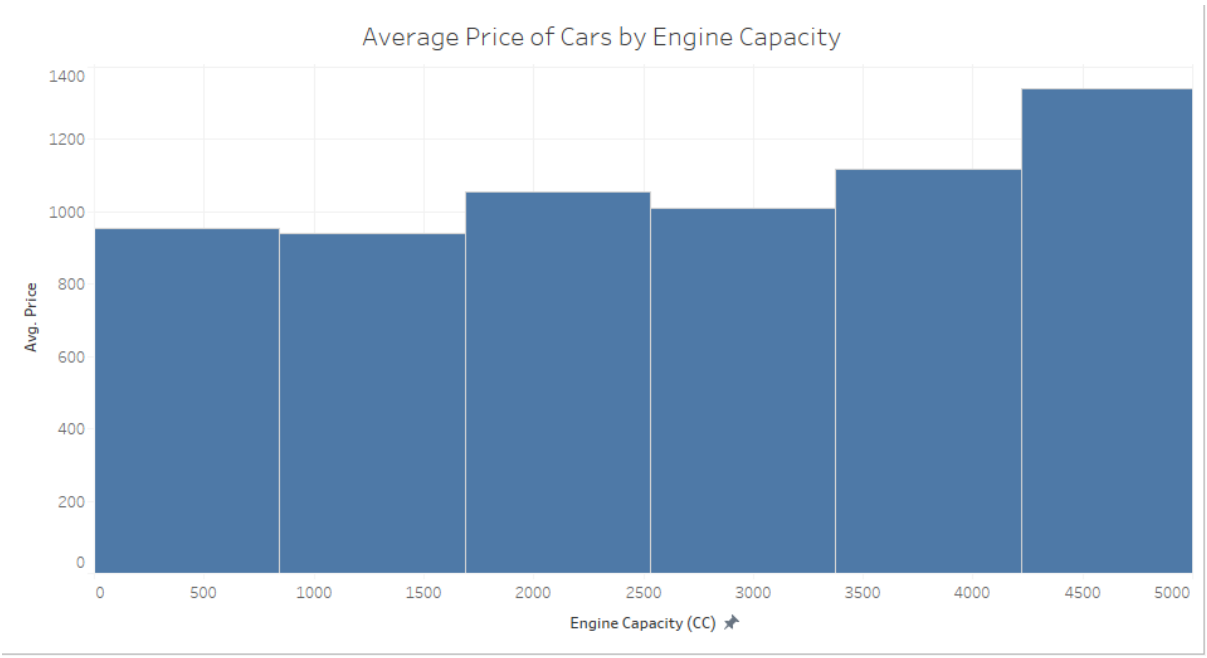

Figure 6- Average Car Price by Engine Capacity

The engine capacity with the highest average price is the 4500cc-5000cc at about $1.3million. This is not an outlier as it liest pretty close to the average range of $1.1million. The lowest average price is of the category of 1000-1500cc of $0.9million. The average price of engine capacity by 3500cc-4000cc is $1.1million. There is not much of difference between all these different engine capacities hence there is no outlier.

There is a very slight positive correlation between the engine capacity and price of the car. The correlation coefficient is at 0.174. This means that for each increase of 1 unit in engine capacity the price of the car is increased by 20%.

No. of Cars distribution by Mileage

The visualization below shows the frequency of cars in the dataset according to the mileage they have travelled. It is visible from the graph below that the highest frequency of cars in the dataset lie within the mileage range of 75,000 to 100,000km range. Least number of cars have mileage greater than 200,000km.

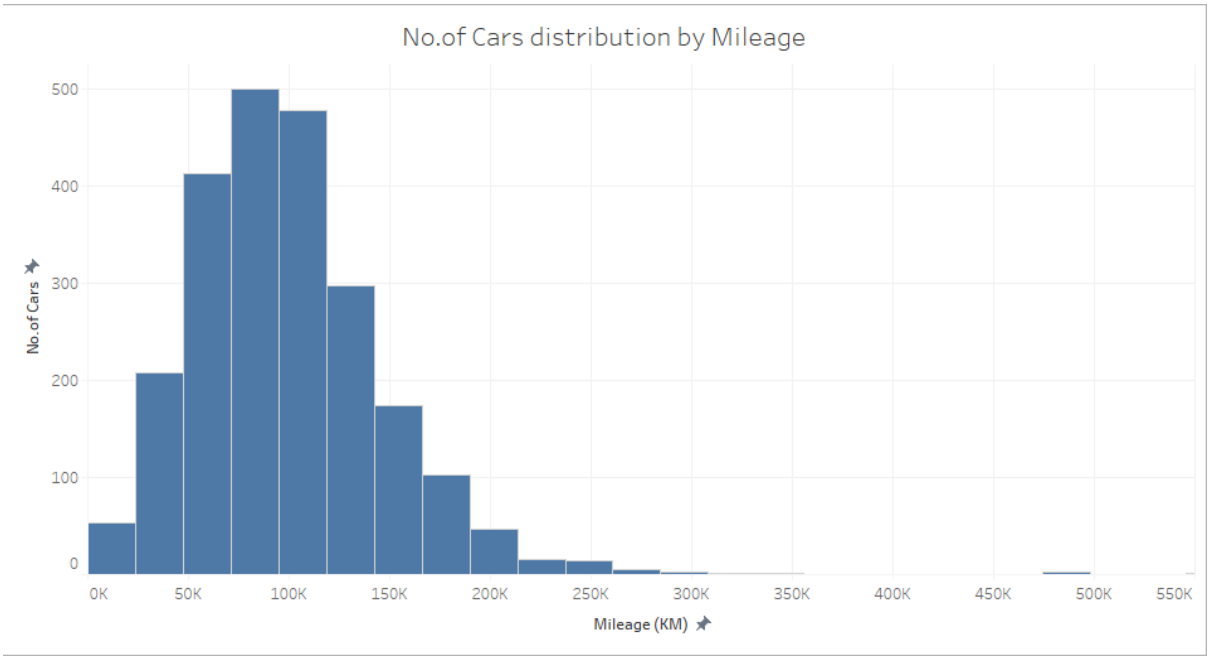

Figure 7- Frequency of Cars by Mileage

Average Car Price by Mileage(Correlation Mileage – Avg Price)

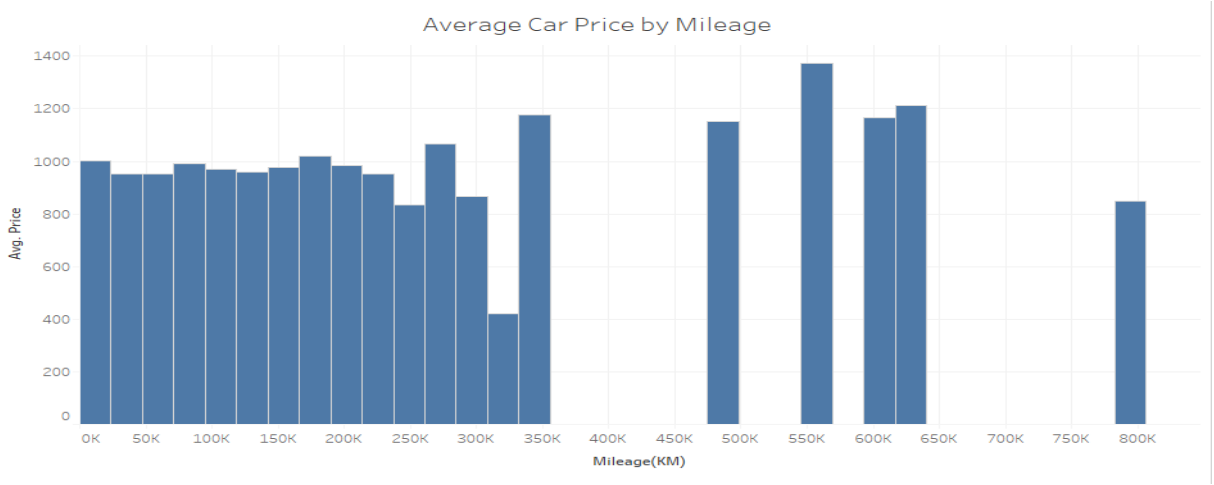

Figure 8- Average Car Price by Mileage

The graph above shows the average car price of each mileage category. The average price of mileage 550,000km-600,000km is the highest at about $1.4million.The lowest average price is of the mileage category 325,000km-350,000km at about $0.4million.

The visualization and correlation analysis shows that there is very little correlation between the two quantities. It measures in at about 0.022 which can be neglected as it is not that considerable in magnitude.

Fuel-type of Cars

The visualization below shows the distribution of fuels in the dataset. It is shown that about 99% of the cars are run gasoline and the rest of the fuel-types contribute only 1 % to the entire total. The major factor in that is diesel which stands at 0.78%.

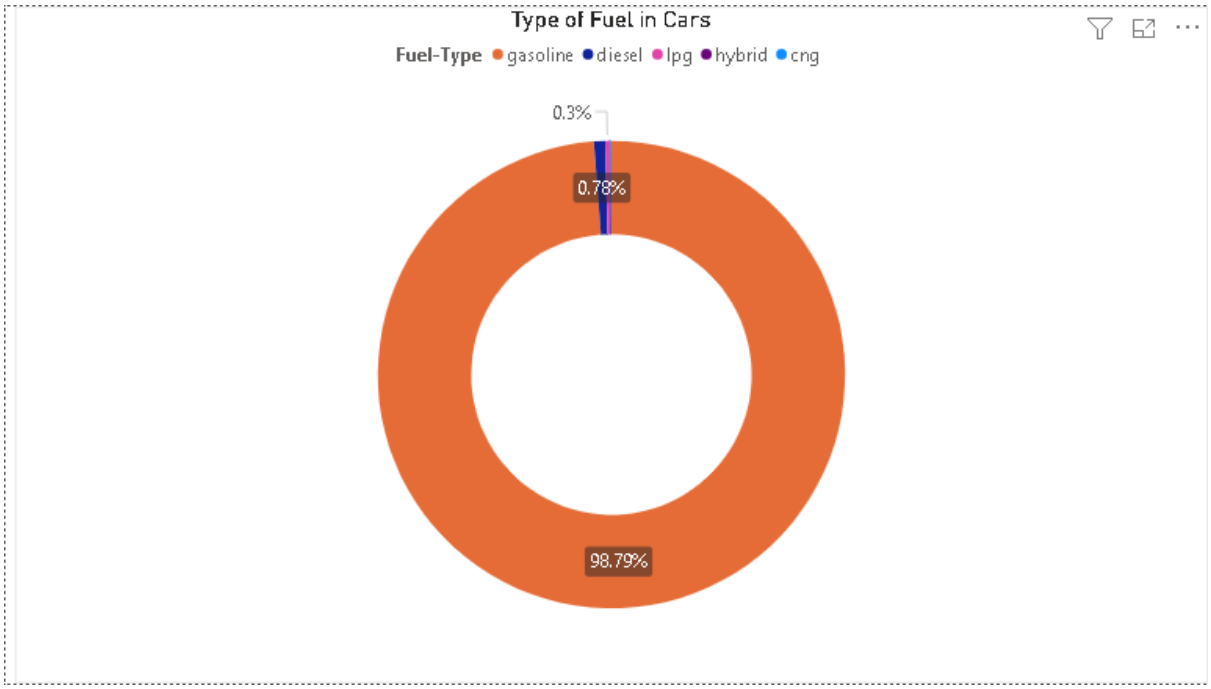

Figure 9- Fuel-type distribution of Cars

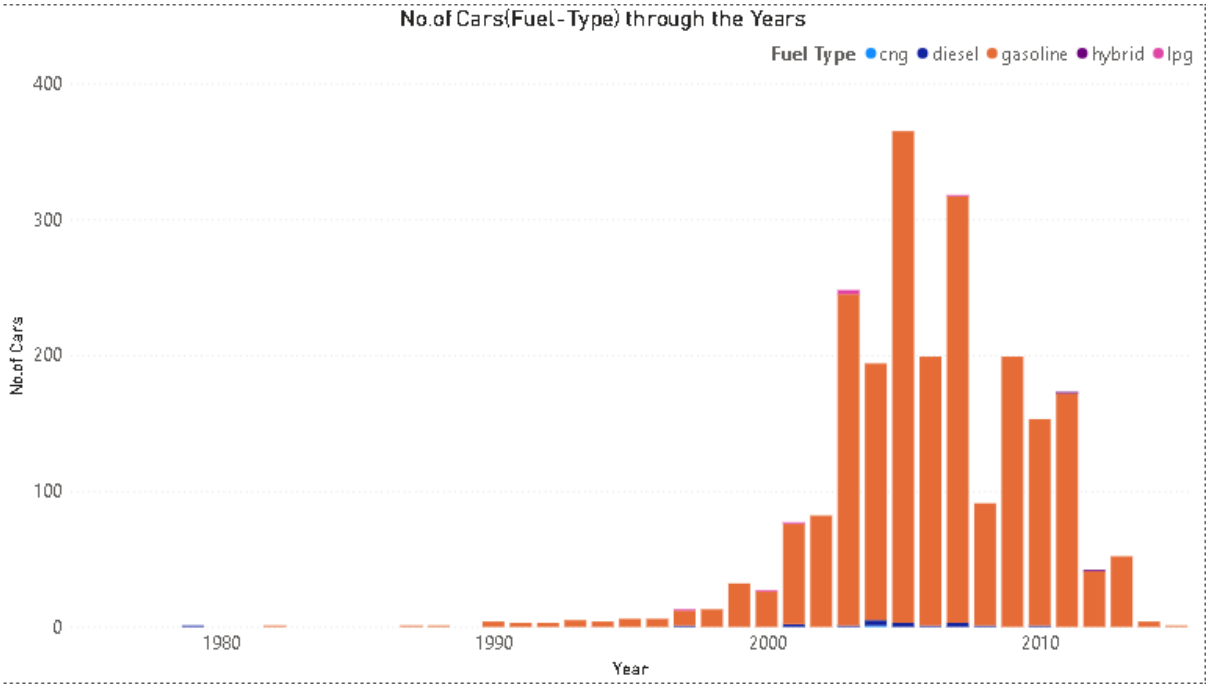

Figure 10- Fuel-type through the years

Transmission type of Cars

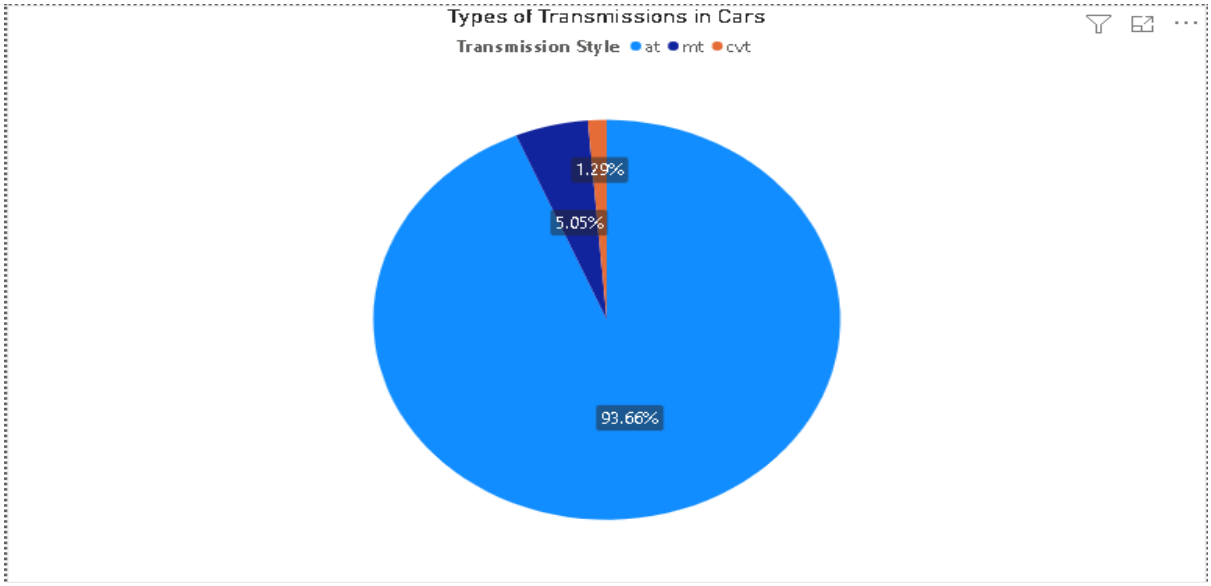

Figure 11- Transmission distribution of Cars

The visualization above shows that the about 94% of the transmission is automatic while only 5% of the transmission is 5%. The CVT transmission is only at 1.29%.

Average Car Price by Model(Correlation Average Price – Model)

The image below gives a visualization of the average car price by the car model. The highest average price was of the model estima t at about $1.4million. The lowest average price is at the $1.2million of the car zest.

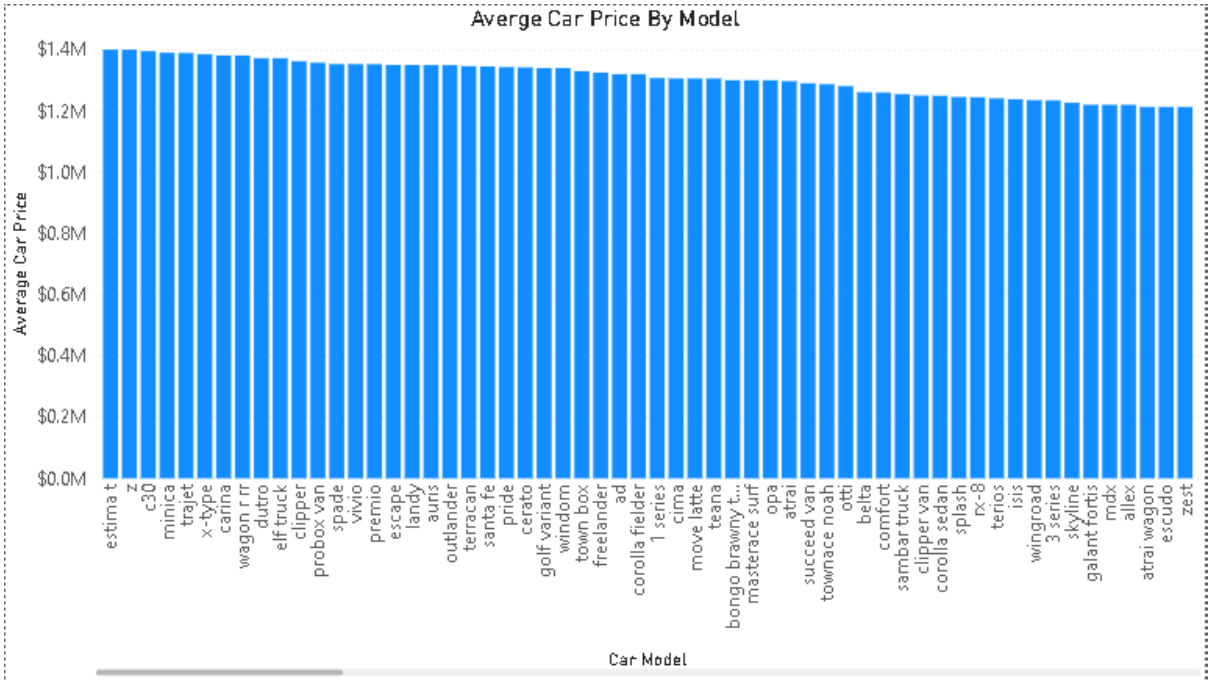

Figure 12- Average Car price by Model

The visualization above shows that the model does not play a vital role in determining the average price of the car. It has about 0 correlation with average price.

Predictor Model For Average Car Price

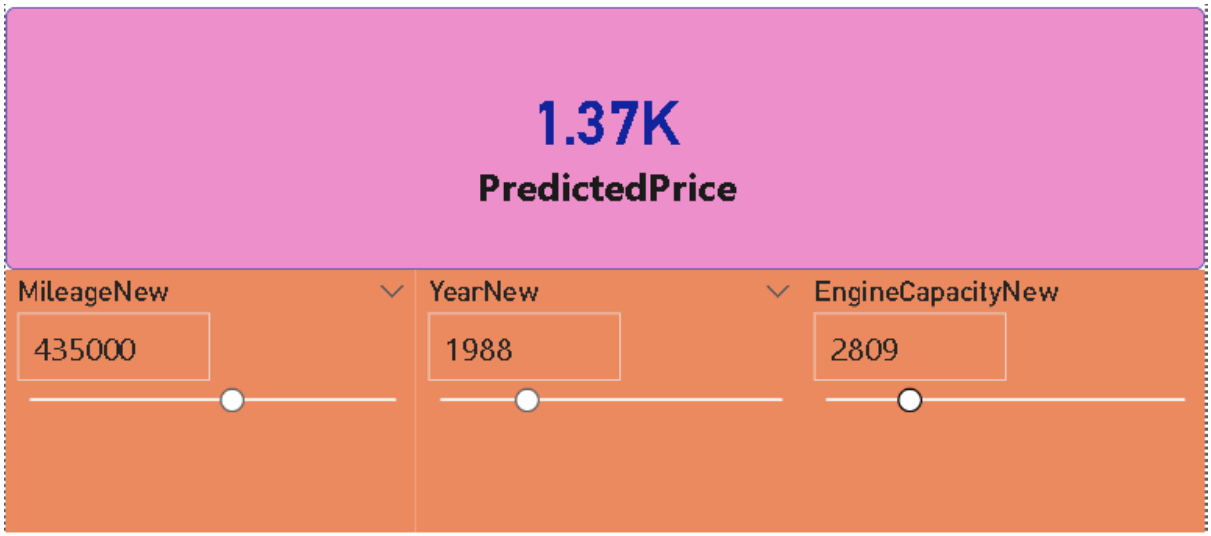

Figure 13- Predictor Model for Car Price

The image above shows the predictor model that is prepared for deducing the average car price. It is designed in an intuitive way in which only the year , engine capacity and mileage of the car is input and it yields an average car price.

It is built on the principle of multiple linear regression and is based off on this equation given below:

$$Predicted\ Price = 30947.59 - 15.01 * Year + 0.0001 * Mileage + 0.0755 * Engine$$