INTRODUCTION

Nowadays owning a four-wheeler is the number one priority for the majority of people. But most of them can’t afford it at the initial stage so, they go for a used vehicle.

A used car dealership specializes in selling cars from various brands. The mileage of these cars is a good predictor of their sale prices. What other factors might play a role in deciding the price a customer might be willing to pay?

For a better understanding of that it is necessary to do EDA and build a model to predict the multiple aspects that are most influencing the price of a vehicle.

DATA

The data consisted of around 7253 rows and 14 columns.

Each column contains different aspects of a car's data

Each row contains a different model of different brands sold car details

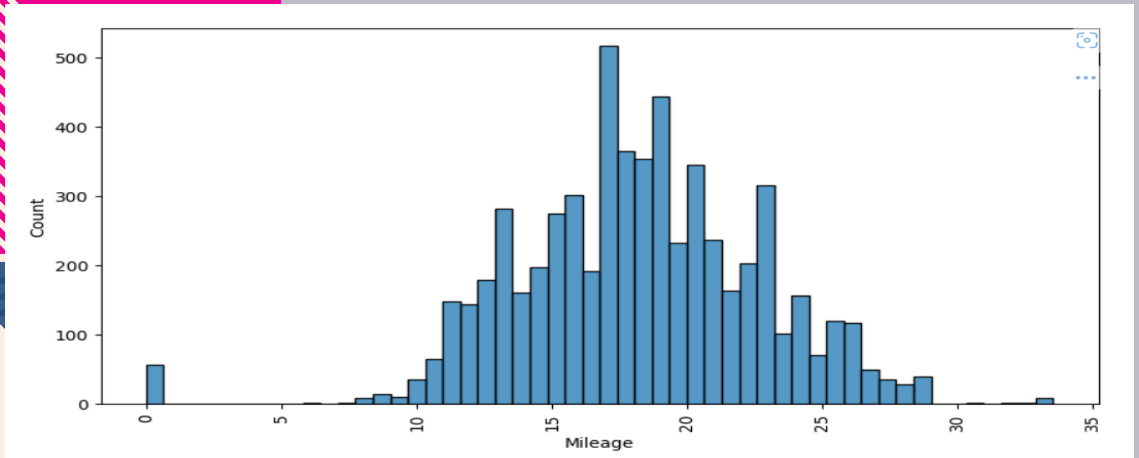
This data frame includes some outliers, null values, and some empty cells also.

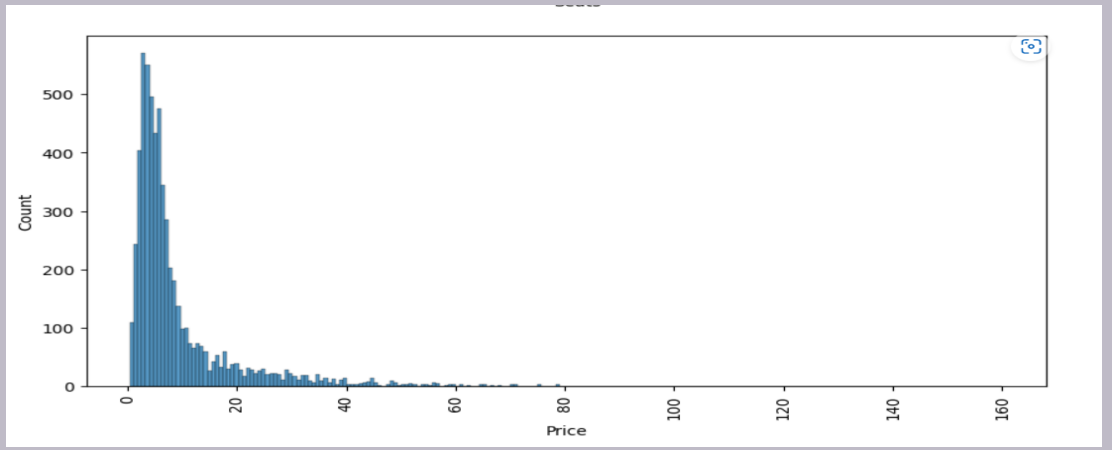
This has both Numerical and categorical values.

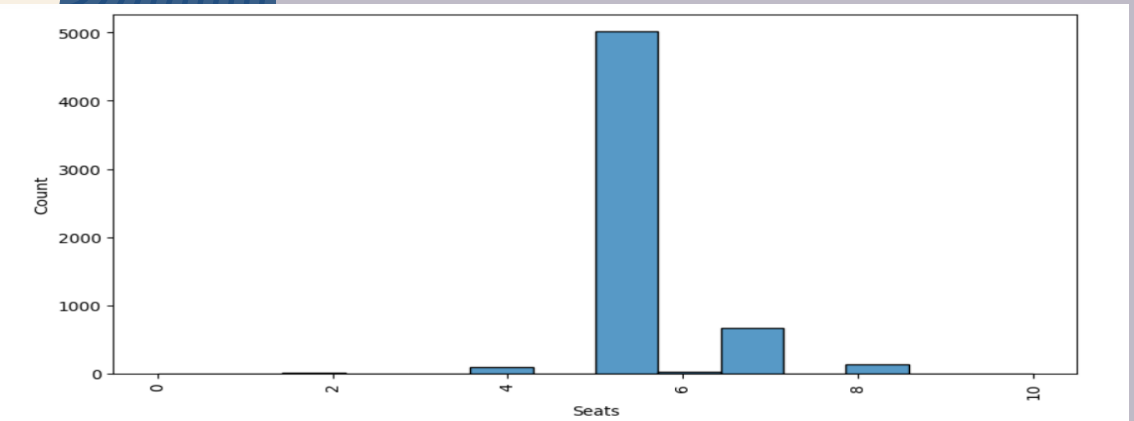
Name	Locatio	Year	Kilomet	Fuel_Ty	Transm	Owner_	Mileage	Engine	Power	Seats	New_Pr	Price
Maruti Wagon R LXI CNG	Mumbai	2010	72000	CNG	Manual	First	26.6 km/kg	998 CC	58.16 bhp	5		1.75
Hyundai Creta 1.6 CRDi SX Option	Pune	2015	41000	Diesel	Manual	First	19.67 kmpl	1582 CC	126.2 bhp	5		12.5
Honda Jazz V	Chennai	2011	46000	Petrol	Manual	First	18.2 kmpl	1199 CC	88.7 bhp	5	8.61 Lakh	4.5
Maruti Ertiga VDI	Chennai	2012	87000	Diesel	Manual	First	20.77 kmpl	1248 CC	88.76 bhp	7		6
Audi A4 New 2.0 TDI Multitronic	Coimbator	2013	40670	Diesel	Automatic	Second	15.2 kmpl	1968 CC	140.8 bhp	5		17.74
Hyundai EON LPG Era Plus Option	Hyderabac	2012	75000	LPG	Manual	First	21.1 km/kg	814 CC	55.2 bhp	5		2.35
Nissan Micra Diesel XV	Jaipur	2013	86999	Diesel	Manual	First	23.08 kmpl	1461 CC	63.1 bhp	5		3.5
Toyota Innova Crysta 2.8 GX AT 8S	Mumbai	2016	36000	Diesel	Automatic	First	11.36 kmpl	2755 CC	171.5 bhp	8	21 Lakh	17.5
Volkswagen Vento Diesel Comfortline	Pune	2013	64430	Diesel	Manual	First	20.54 kmpl	1598 CC	103.6 bhp	5		5.2
Tata Indica Vista Quadrajet LS	Chennai	2012	65932	Diesel	Manual	Second	22.3 kmpl	1248 CC	74 bhp	5		1.95
Maruti Ciaz Zeta	Kochi	2018	25692	Petrol	Manual	First	21.56 kmpl	1462 CC	103.25 bhp	5	10.65 Lakh	9.95

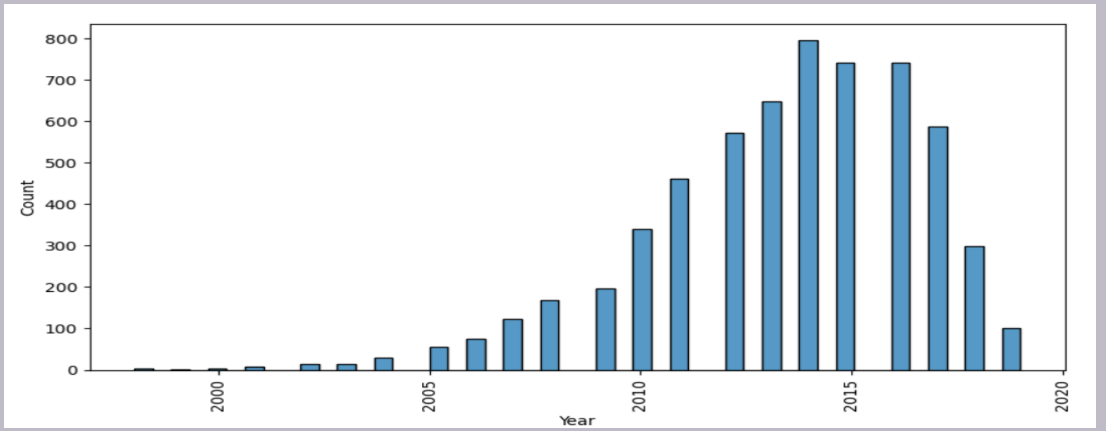
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5981 entries, 0 to 5980
Data columns (total 14 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   index              5981 non-null   int64
 1   Name               5981 non-null   object
 2   Location           5981 non-null   object
 3   Year               5981 non-null   int64
 4   Kilometers_Driven  5981 non-null   int64
 5   Fuel_Type          5981 non-null   object
 6   Transmission       5981 non-null   object
 7   Owner_Type         5981 non-null   object
 8   Mileage            5981 non-null   float64
 9   Engine             5981 non-null   float64
 10  Power              5981 non-null   float64
 11  Seats              5981 non-null   float64
 12  Price              5981 non-null   float64
 13  Brand              5981 non-null   object
dtypes: float64(5), int64(3), object(6)
memory usage: 654.3+ KB
```

EDA - HIGHLIGHTS









From these graphs, we came to know people mostly prefer used cars in the following categories

- 3L – 8L price range
- 13km/lit – 23km/lit mileage range
- 5 seater
- 2010 – 2014 model

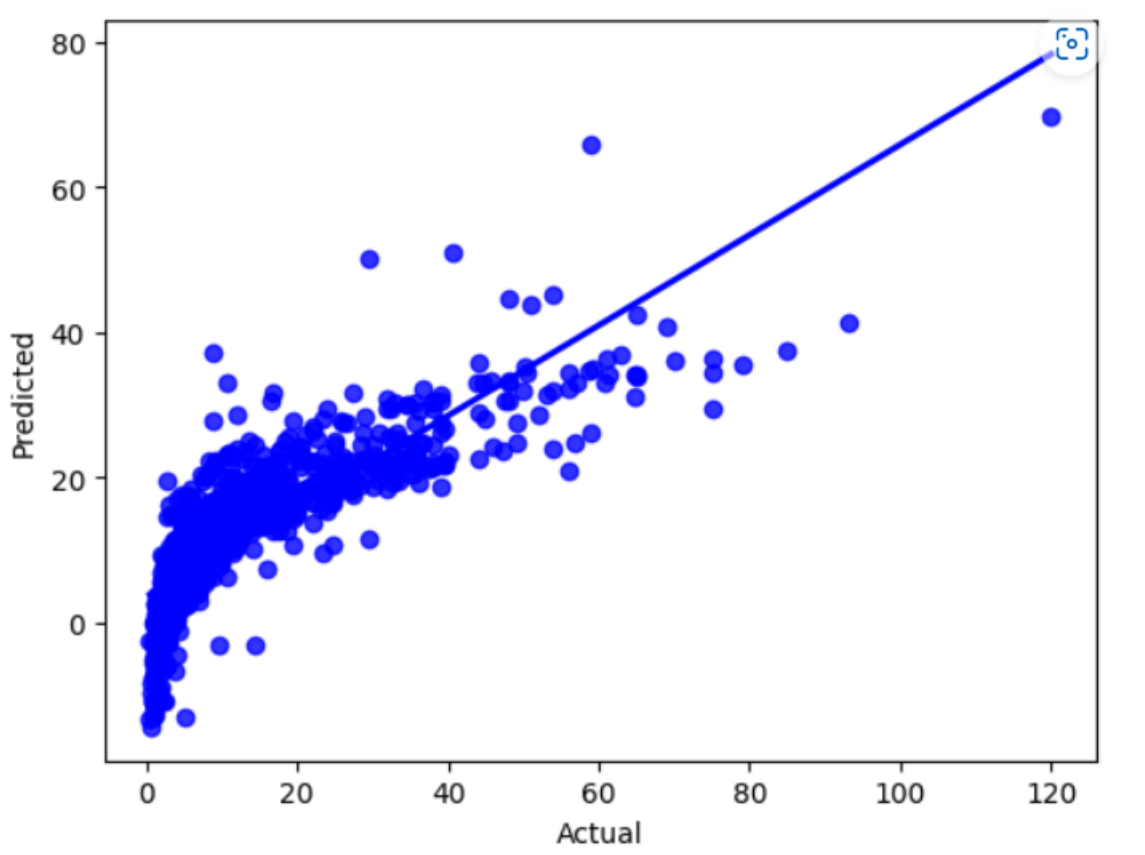
MODEL – MULTI-LINEAR REGRESSION

- Based on the findings from EDA, with respect to the selling price four aspects were considered for the model building
- Engine, Mileage, Power, Year
- Mean Squared Error is: 43.65
- Mean Absolute Error is: 4.03
- Root Mean Squared Error is: 6.60
- R2 Score is: 0.67

FINDINGS

Due to the price range of some high-end cars like Lamborghini, and Porsche the price range is getting affected by the mileage.
This kind of car has very less mileage and fewer seatings but, has a very high price range and good engine power.

CONCLUSION



This regression plot helps to understand the accuracy between the actual value and the value predicted by the model.
This model

This model has 67.8% accuracy.