CAR PRICE PREDICTION (CAR DEKHO.COM)

BY: JASWINDER_SINGH

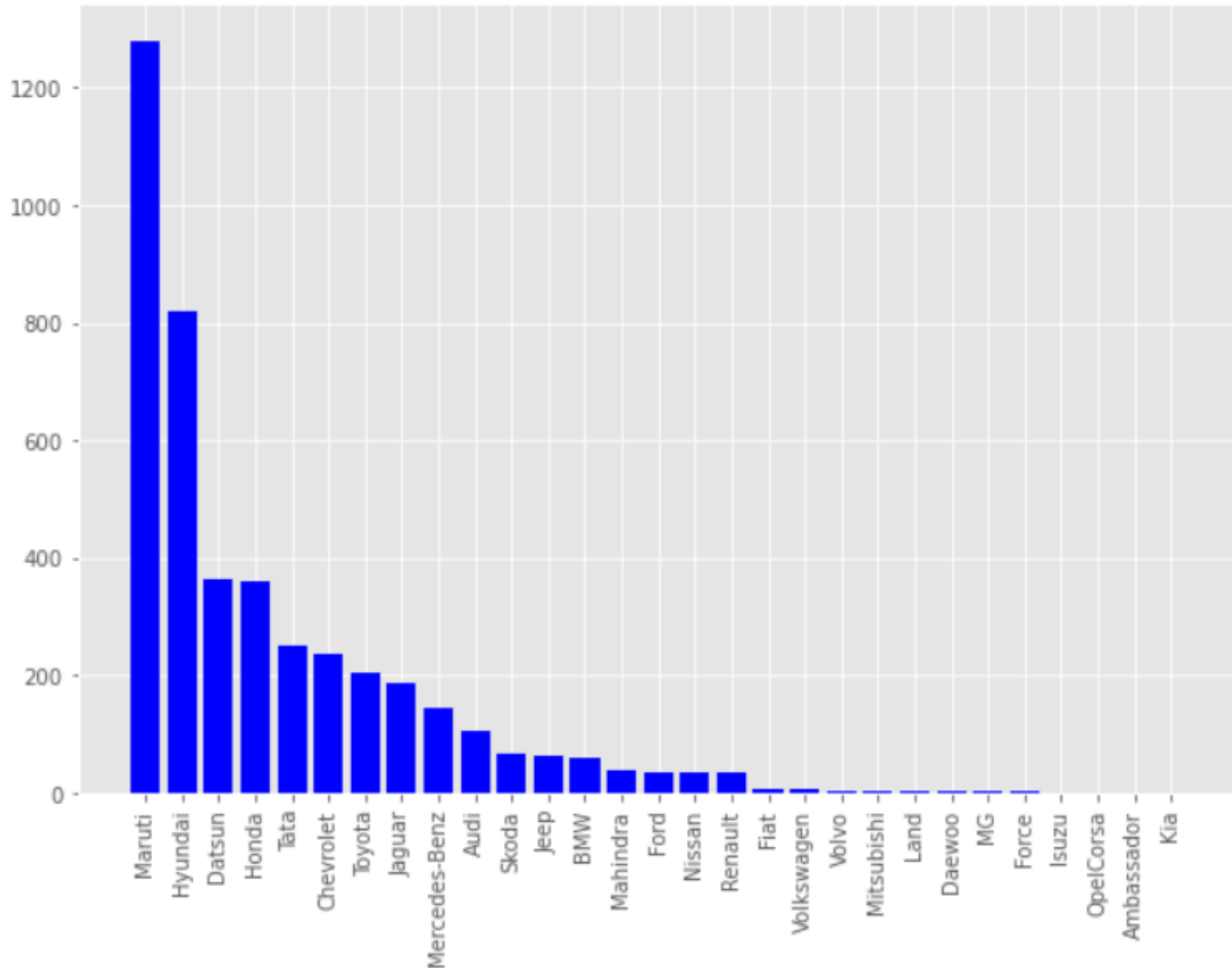# DATA

The data for this project belongs to KAGGLE.

BUSINESS PROBLEM: THE stake holder is new to the market, and want to understand the trend of price.

ALOS, seeking for some insights and recommendations.

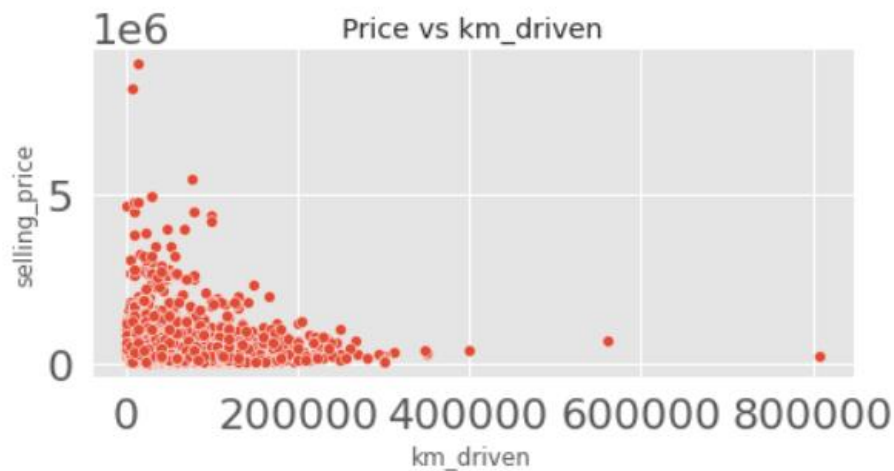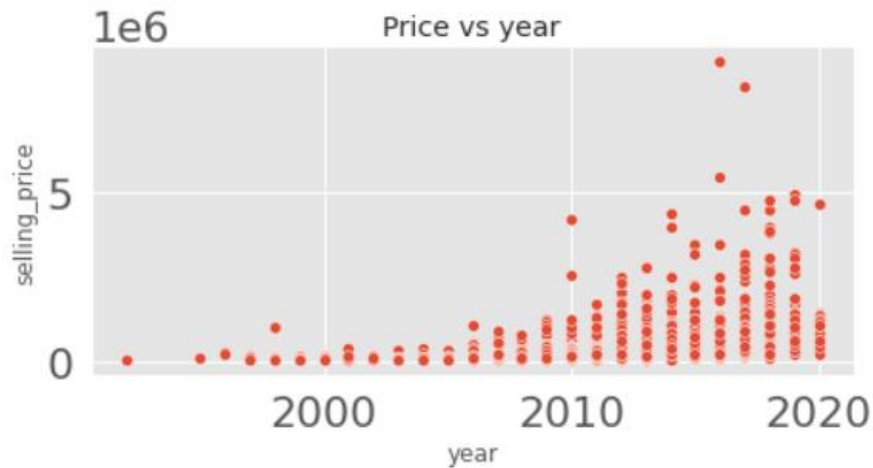|	name	year	selling_price	km_driven	fuel	seller_type	transmission	owner
0	Maruti 800 AC	2007	60000	70000	Petrol	Individual	Manual	First Owner
1	Maruti Wagon R LXI Minor	2007	135000	50000	Petrol	Individual	Manual	First Owner
2	Hyundai Verna 1.6 SX	2012	600000	100000	Diesel	Individual	Manual	First Owner
3	Datsun RediGO T Option	2017	250000	46000	Petrol	Individual	Manual	First Owner
4	Honda Amaze VX i-DTEC	2014	450000	141000	Diesel	Individual	Manual	Second Owner

•The data set is a combinations of all the attributes of cars
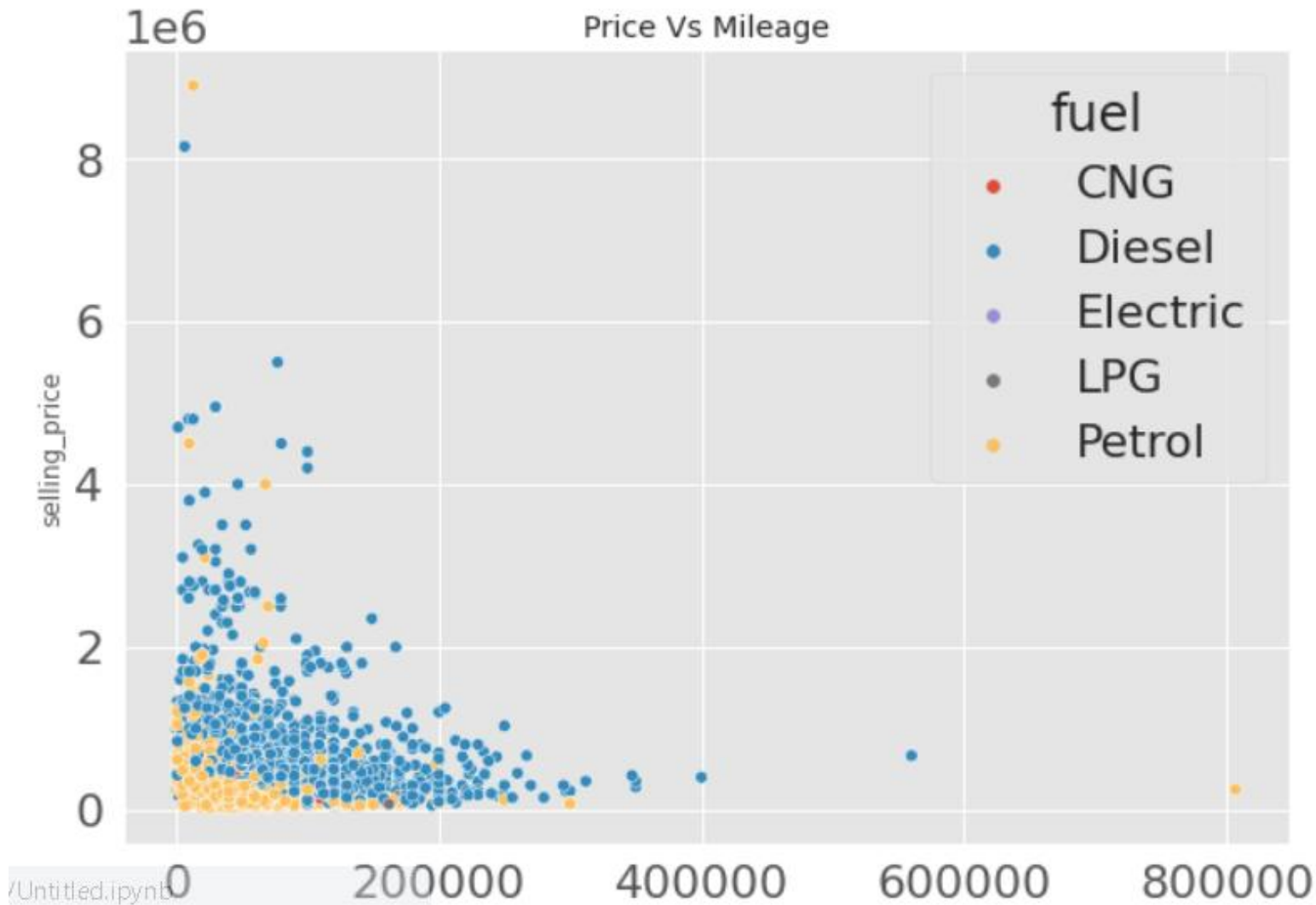•The data is generated by an INDIAN start-up company CAR DEKHO.COM

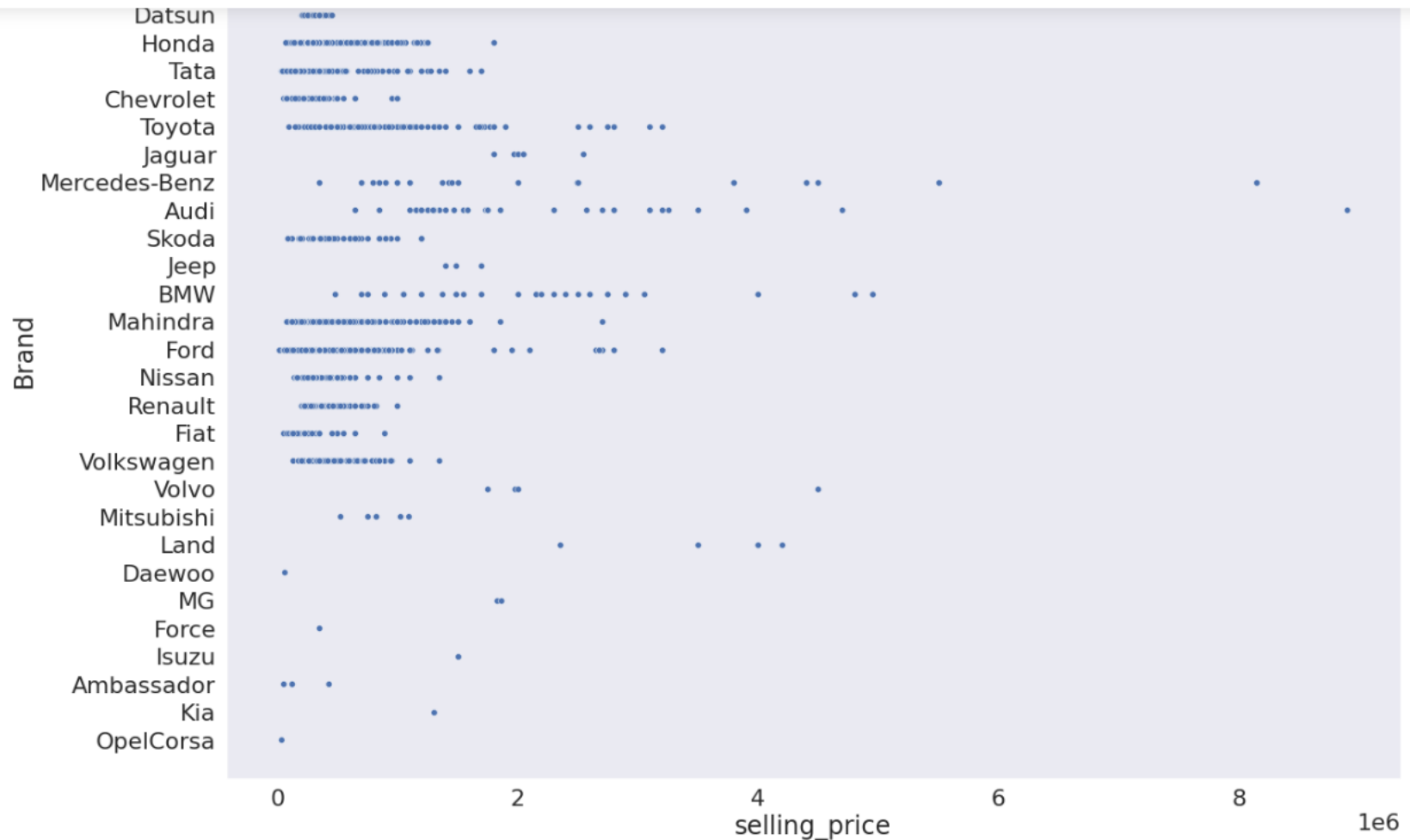# Brands to deal with

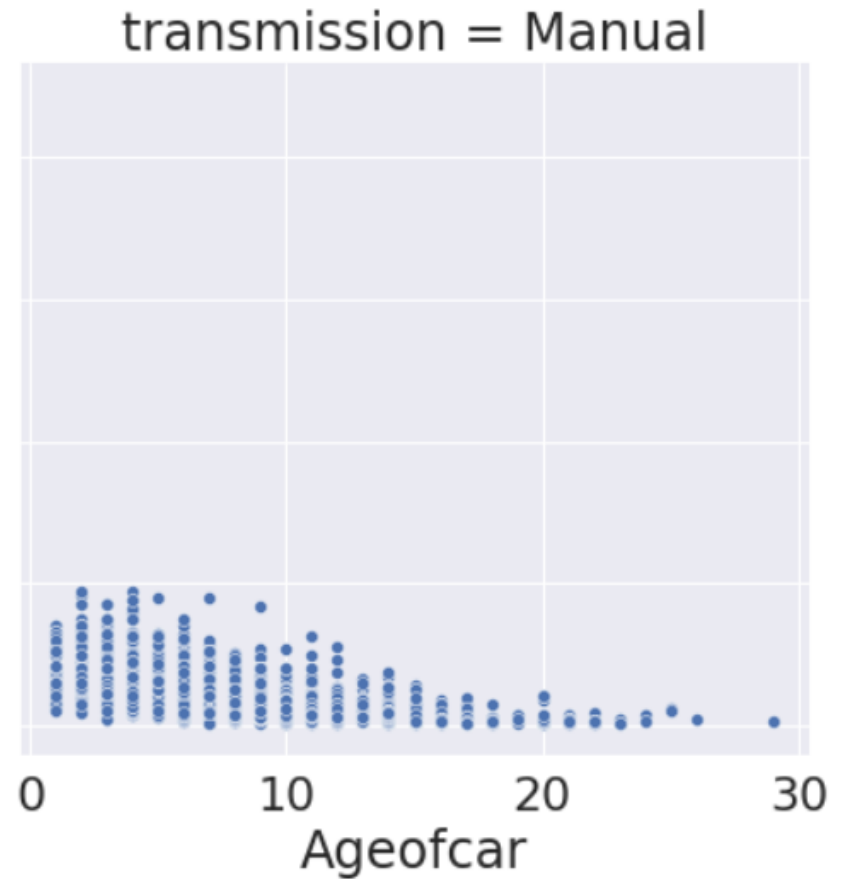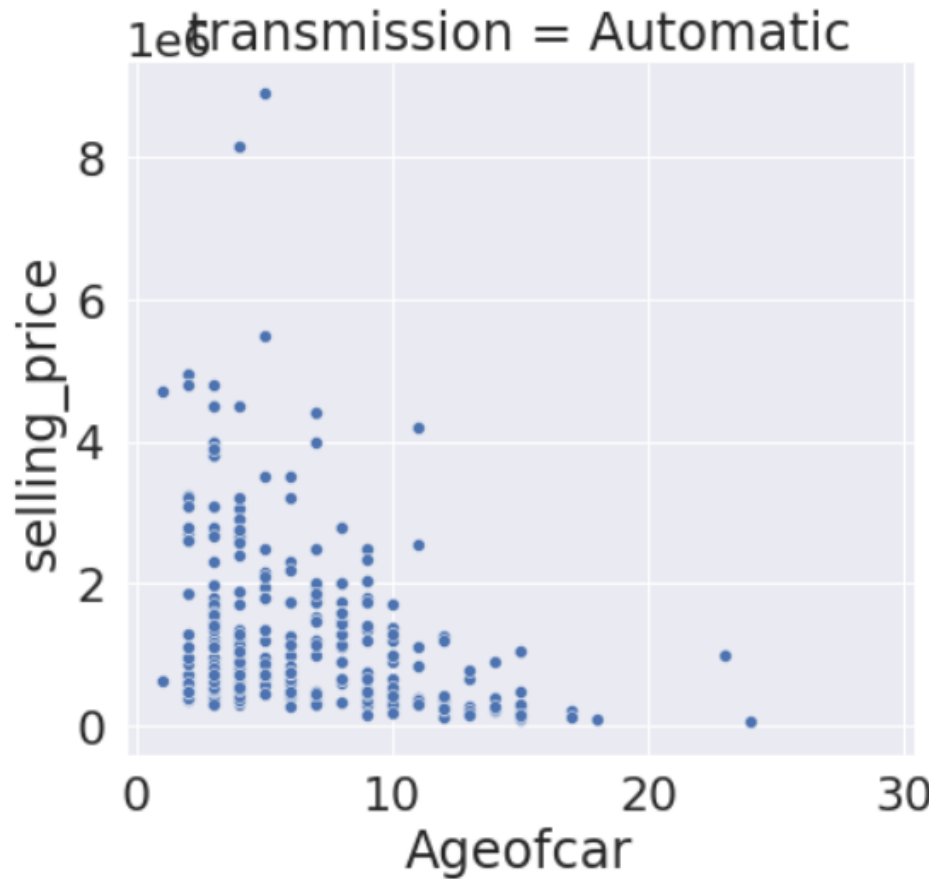# Relationship with price

quantitative attributes

Price distribution with respect to fuel & mileage

Price VS brand

Comparing Automatic & Manual car's price

First Model Summary
(With only positively related variable

```
                          OLS Regression Results
=================================================================
Dep. Variable:      selling_price_log   R-squared:               0.513
Model:                            OLS   Adj. R-squared:          0.511
Method:                 Least Squares   F-statistic:             245.4
Date:                Sun, 12 Jun 2022   Prob (F-statistic):       0.00
Time:                        14:13:20   Log-Likelihood:        -2693.4
No. Observations:                3037   AIC:                     5415.
Df Residuals:                    3023   BIC:                     5499.
Df Model:                          13
Covariance Type:            nonrobust
=================================================================
                                  coef   std err       t     P>|t|    [0.025    0.975]
-----------------------------------------------------------------
const                          16.4916    0.180    91.767    0.000    16.139    16.844
km_driven_log                  -0.2442    0.013   -18.258    0.000    -0.270    -0.218
fuel_Diesel                     0.5360    0.106     5.077    0.000     0.329     0.743
fuel_Electric                  -0.4174    0.600    -0.696    0.486    -1.593     0.758
fuel_LPG                       -0.2872    0.185    -1.557    0.120    -0.649     0.075
fuel_Petrol                    -0.1569    0.105    -1.487    0.137    -0.364     0.050
seller_type_Individual         -0.0816    0.028    -2.959    0.003    -0.136    -0.028
seller_type_Trustmark Dealer    0.5276    0.074     7.140    0.000     0.383     0.672
transmission_Manual            -0.7914    0.039   -20.072    0.000    -0.869    -0.714
owner_Fourth & Above Owner     -0.7552    0.076    -9.884    0.000    -0.905    -0.605
owner_Second Owner             -0.3278    0.026   -12.530    0.000    -0.379    -0.276
owner_Test Drive Car            0.0824    0.170     0.485    0.628    -0.251     0.415
owner_Third Owner              -0.5366    0.044   -12.275    0.000    -0.622    -0.451
Brand_Class_Low                -0.4077    0.045    -8.968    0.000    -0.497    -0.319
=================================================================
Omnibus:                       52.817   Durbin-Watson:             1.969
Prob(Omnibus):                  0.000   Jarque-Bera (JB):         56.969
Skew:                          -0.297   Prob(JB):               4.26e-13
Kurtosis:                       3.311   Cond. No.                   620.
=================================================================
```
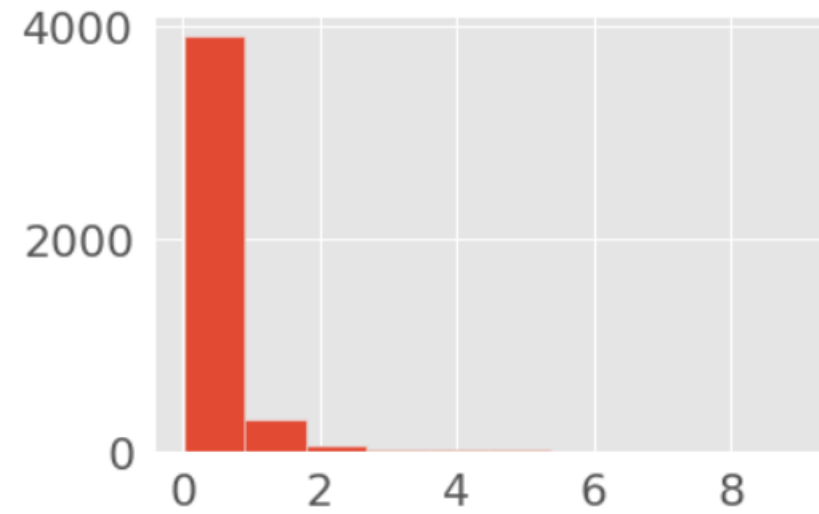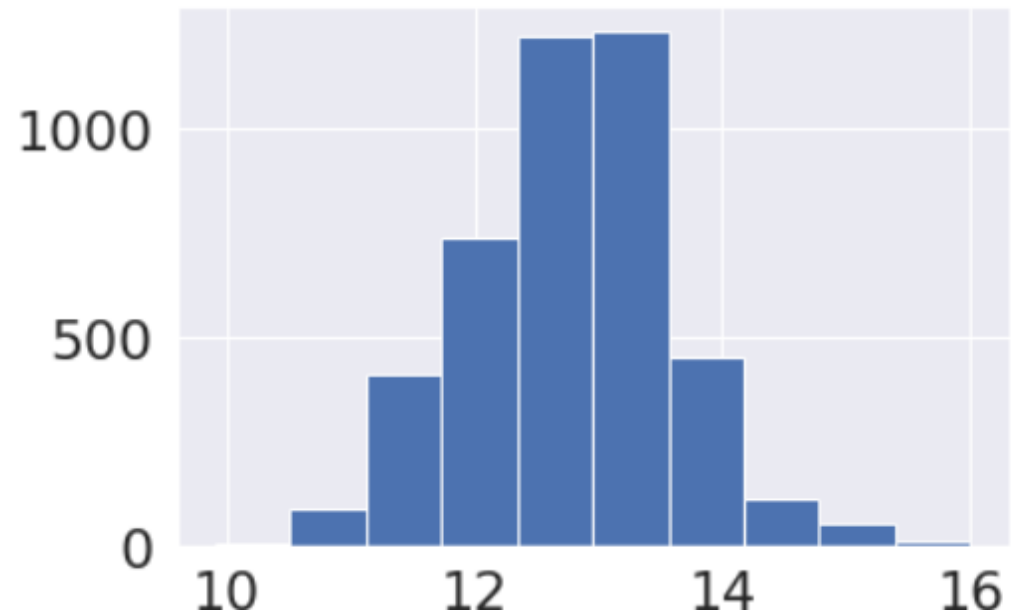
R2 value is less,
(can only predict 50% of the
true 342141values.
Most coef_ are in negative.

NORMALIZATION

BEFORE NORMALIZATION

AFTER NORMALIZATION

SECOND(FINAL MODEL)

Dep. Variable:	selling_price_log	R-squared (uncentered):	0.996
Model:	OLS	Adj. R-squared (uncentered):	0.996
Method:	Least Squares	F-statistic:	7.326e+04
Date:	Sun, 12 Jun 2022	Prob (F-statistic):	0.00
Time:	14:13:23	Log-Likelihood:	-5193.2
No. Observations:	4339	AIC:	1.042e+04
Df Residuals:	4324	BIC:	1.051e+04
Df Model:	15		
Covariance Type:	nonrobust		

R2 value is appox. 1, mean it can predict almost all the values right Coef_ are in positive.

	coef	std err	t	P>|t|	[0.025	0.975]
km_driven	-1.46e-05	3.6e-07	-40.492	0.000	-1.53e-05	-1.39e-05
Ageofcar	-0.1467	0.004	-39.889	0.000	-0.154	-0.139
km_driven_log	1.1018	0.012	89.104	0.000	1.078	1.126
fuel_Diesel	3.7638	0.114	33.092	0.000	3.541	3.987
fuel_Electric	3.6455	0.812	4.489	0.000	2.053	5.238
fuel_LPG	3.0706	0.202	15.189	0.000	2.674	3.467
fuel_Petrol	3.4075	0.113	30.274	0.000	3.187	3.628
seller_type_Individual	-0.1007	0.031	-3.234	0.001	-0.162	-0.040
seller_type_Trustmark Dealer	0.3937	0.084	4.687	0.000	0.229	0.558
transmission_Manual	-0.5973	0.044	-13.431	0.000	-0.685	-0.510

selling_price_log

IMPORTANT POINTS & INSIGHTS

- Based on Analysis, we can to divide our cars into 3 segment Low, Medium and High budget.

- Brands like Maruti, Hyundai ,Honda are low budget and very popular brands in used car market.

- Brands like BMW, Bentley, Jaguar, Land Rover, Mercedes Benz,Porche,Mini Cooper are high budget cars and are mostly bought by car enthusiast who are ready to buy a two user owned car at higher price as well.

- Brands like Toyota,Volvo can be Medium budget cars.

- Need to acquire more Automatic cars to earn more profits, as this car sell at higher prices.

- With Increasing petrol rates diesel car are in more demand in recent years, acquiring and selling them can high profits

- Along with this we can include scheme like take a test drive for half day to pursue customer to buy.

- We can provide Car maintenance packages where customers pays a small upfront fees and can bring the car for servicing anytime in a year to attract more customers.

THANK YOU