

```
import numpy as np
```

```
import pandas as pd
```

```
df=pd.read_csv('https://raw.githubusercontent.com/YBI-Foundation/Dataset/main/Car%20Price.
```

```
df.head()
```

	Brand	Model	Year	Selling_Price	KM_Driven	Fuel	Seller_Type	Transmissi
0	Maruti	Maruti 800 AC	2007	60000	70000	Petrol	Individual	Mar
1	Maruti	Maruti Wagon R LXI Minor	2007	135000	50000	Petrol	Individual	Mar
		Hundai						

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4340 entries, 0 to 4339
Data columns (total 9 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Brand           4340 non-null  object
1   Model           4340 non-null  object
2   Year            4340 non-null  int64
3   Selling_Price   4340 non-null  int64
4   KM_Driven       4340 non-null  int64
5   Fuel            4340 non-null  object
6   Seller_Type     4340 non-null  object
7   Transmission    4340 non-null  object
8   Owner           4340 non-null  object
dtypes: int64(3), object(6)
memory usage: 305.3+ KB
```

```
df.describe()
```

	Year	Selling_Price	KM_Driven
count	4340.000000	4.340000e+03	4340.000000

```
df[['Brand']].value_counts()
```

```

Brand
Maruti      1280
Hyundai     821
Mahindra    365
Tata        361
Honda       252
Ford        238
Toyota      206
Chevrolet   188
Renault     146
Volkswagen  107
Skoda       68
Nissan       64
Audi        60
BMW         39
Fiat        37
Datsun      37
Mercedes-Benz 35
Mitsubishi  6
Jaguar      6
Land        5
Ambassador  4
Volvo       4
Jeep        3
OpelCorsa   2
MG          2
Isuzu       1
Force       1
Daewoo      1
Kia         1
dtype: int64

```

```
df[['Model']].value_counts()
```

```

Model
Maruti Swift Dzire VDI      69
Maruti Alto 800 LXI         59
Maruti Alto LXi             47
Hyundai EON Era Plus       35
Maruti Alto LX              35
..
Mahindra KUV 100 G80 K4 Plus 1
Mahindra KUV 100 mFALCON D75 K8 1
Mahindra KUV 100 mFALCON D75 K8 AW 1
Mahindra KUV 100 mFALCON G80 K2 Plus 1
Volvo XC60 D5 Inscription    1
Length: 1491, dtype: int64

```

```
df[['Fuel']].value_counts()
```

```
Fuel
Diesel      2153
Petrol      2123
CNG         40
LPG         23
Electric     1
dtype: int64
```

```
df[['Seller_Type']].value_counts()
```

```
Seller_Type
Individual    3244
Dealer        994
Trustmark Dealer  102
dtype: int64
```

```
df[['Transmission']].value_counts()
```

```
Transmission
Manual      3892
Automatic    448
dtype: int64
```

```
df[['Owner']].value_counts()
```

```
Owner
First Owner      2832
Second Owner     1106
Third Owner       304
Fourth & Above Owner  81
Test Drive Car    17
dtype: int64
```

```
df[['Fuel','Seller_Type','Transmission','Owner']].value_counts()
```

Fuel	Seller_Type	Transmission	Owner	
Petrol	Individual	Manual	First Owner	892
Diesel	Individual	Manual	First Owner	835
Petrol	Individual	Manual	Second Owner	454
Diesel	Individual	Manual	Second Owner	453
	Dealer	Manual	First Owner	336
Petrol	Dealer	Manual	First Owner	306
Diesel	Individual	Manual	Third Owner	142
Petrol	Individual	Manual	Third Owner	130
Diesel	Dealer	Automatic	First Owner	114
Petrol	Dealer	Automatic	First Owner	80
Diesel	Individual	Automatic	First Owner	73
Petrol	Individual	Automatic	First Owner	67
Diesel	Dealer	Manual	Second Owner	56
Petrol	Trustmark Dealer	Manual	First Owner	49
	Dealer	Manual	Second Owner	44
	Individual	Manual	Fourth & Above Owner	41
Diesel	Trustmark Dealer	Manual	First Owner	34
	Individual	Manual	Fourth & Above Owner	32
		Automatic	Second Owner	31
Petrol	Individual	Automatic	Second Owner	20

CNG	Individual	Manual	First Owner	14
			Second Owner	12
Diesel	Trustmark Dealer	Automatic	First Owner	12
Petrol	Individual	Automatic	Third Owner	10
Diesel	Dealer	Automatic	Second Owner	10
LPG	Individual	Manual	Second Owner	10
Diesel	Individual	Automatic	Third Owner	9
Petrol	Dealer	Manual	Test Drive Car	9
LPG	Individual	Manual	First Owner	9
Petrol	Dealer	Automatic	Second Owner	9
CNG	Dealer	Manual	First Owner	7
Diesel	Dealer	Manual	Test Drive Car	7
Petrol	Dealer	Manual	Third Owner	4
	Trustmark Dealer	Automatic	First Owner	3
CNG	Individual	Manual	Fourth & Above Owner	3
Diesel	Dealer	Manual	Third Owner	3
LPG	Individual	Manual	Third Owner	2
Petrol	Trustmark Dealer	Automatic	Second Owner	2
CNG	Dealer	Manual	Second Owner	2
Diesel	Trustmark Dealer	Automatic	Second Owner	2
CNG	Individual	Manual	Third Owner	2
Diesel	Dealer	Automatic	Third Owner	2
		Manual	Fourth & Above Owner	1
Petrol	Individual	Automatic	Fourth & Above Owner	1
LPG	Individual	Manual	Fourth & Above Owner	1
Diesel	Individual	Automatic	Fourth & Above Owner	1
Petrol	Dealer	Manual	Fourth & Above Owner	1
LPG	Dealer	Manual	First Owner	1
Petrol	Dealer	Automatic	Test Drive Car	1
Electric	Dealer	Automatic	Second Owner	1

dtype: int64

```
#df[['Fuel','Seller_Type','Transmission','Owner']].value_counts()
```

Fuel	Seller_Type	Transmission	Owner	
Petrol	Individual	Manual	First Owner	892
Diesel	Individual	Manual	First Owner	835
Petrol	Individual	Manual	Second Owner	454
Diesel	Individual	Manual	Second Owner	453
	Dealer	Manual	First Owner	336
Petrol	Dealer	Manual	First Owner	306
Diesel	Individual	Manual	Third Owner	142
Petrol	Individual	Manual	Third Owner	130
Diesel	Dealer	Automatic	First Owner	114
Petrol	Dealer	Automatic	First Owner	80
Diesel	Individual	Automatic	First Owner	73
Petrol	Individual	Automatic	First Owner	67
Diesel	Dealer	Manual	Second Owner	56
Petrol	Trustmark Dealer	Manual	First Owner	49
	Dealer	Manual	Second Owner	44
	Individual	Manual	Fourth & Above Owner	41
Diesel	Trustmark Dealer	Manual	First Owner	34
	Individual	Manual	Fourth & Above Owner	32
		Automatic	Second Owner	31
Petrol	Individual	Automatic	Second Owner	20
CNG	Individual	Manual	First Owner	14
			Second Owner	12
Diesel	Trustmark Dealer	Automatic	First Owner	12
Petrol	Individual	Automatic	Third Owner	10

Diesel	Dealer	Automatic	Second Owner	10
LPG	Individual	Manual	Second Owner	10
Diesel	Individual	Automatic	Third Owner	9
Petrol	Dealer	Manual	Test Drive Car	9
LPG	Individual	Manual	First Owner	9
Petrol	Dealer	Automatic	Second Owner	9
CNG	Dealer	Manual	First Owner	7
Diesel	Dealer	Manual	Test Drive Car	7
Petrol	Dealer	Manual	Third Owner	4
	Trustmark Dealer	Automatic	First Owner	3
CNG	Individual	Manual	Fourth & Above Owner	3
Diesel	Dealer	Manual	Third Owner	3
LPG	Individual	Manual	Third Owner	2
Petrol	Trustmark Dealer	Automatic	Second Owner	2
CNG	Dealer	Manual	Second Owner	2
Diesel	Trustmark Dealer	Automatic	Second Owner	2
CNG	Individual	Manual	Third Owner	2
Diesel	Dealer	Automatic	Third Owner	2
		Manual	Fourth & Above Owner	1
Petrol	Individual	Automatic	Fourth & Above Owner	1
LPG	Individual	Manual	Fourth & Above Owner	1
Diesel	Individual	Automatic	Fourth & Above Owner	1
Petrol	Dealer	Manual	Fourth & Above Owner	1
LPG	Dealer	Manual	First Owner	1
Petrol	Dealer	Automatic	Test Drive Car	1
Electric	Dealer	Automatic	Second Owner	1

dtype: int64

```
df.columns
```

```
Index(['Brand', 'Model', 'Year', 'Selling_Price', 'KM_Driven', 'Fuel',
      'Seller_Type', 'Transmission', 'Owner'],
      dtype='object')
```

```
df.shape
```

```
(4340, 9)
```

```
#Get encoding of categorical Features
```

```
df.replace({'Fuel':{'Petrol':0,'Diesel':1,'CNG':2,'LPG':3,'Electric':4}},inplace=True)
```

```
df.replace({'Seller_Type':{'Individual':0,'Dealer':1,'Trustmark Dealer':2}},inplace=True)
```

```
df.replace({'Transmission':{'Manual':0,'Automatic':1}},inplace=True)
```

```
df.replace({'Owner':{'First Owner':0,'Second Owner':1,'Third Owner':2,'Fourth & Above Owne
```

```
#define y(dependent variable) and x(indepdent variable)
```

```
y=df['Selling_Price']
```

y.shape

(4340,)

y

```
0      60000
1     135000
2     600000
3     250000
4     450000
...
4335   409999
4336   409999
4337   110000
4338   865000
4339   225000
Name: Selling_Price, Length: 4340, dtype: int64
```


X=df[['Year','KM_Driven','Fuel','Seller_Type','Transmission','Owner']]

#X=df.drop(['Brand','Model','Selling_Price'],axis=1)

X.shape

(4340, 6)

X

	Year	KM_Driven	Fuel	Seller_Type	Transmission	Owner	
0	2007	70000	0	0	0	0	
1	2007	50000	0	0	0	0	
2	2012	100000	1	0	0	0	
3	2017	46000	0	0	0	0	
4	2014	141000	1	0	0	1	
...	
4335	2014	80000	1	0	0	1	
4336	2014	80000	1	0	0	1	
4337	2009	83000	0	0	0	1	
4338	2016	90000	1	0	0	0	
4339	2016	40000	0	0	0	0	

4340 rows × 6 columns

```
#Get Train test split
```

```
from sklearn.model_selection import train_test_split
```

```
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.3,random_state=2529)
```

```
X_train.shape,X_test.shape,y_train.shape,y_test.shape
```

```
((3038, 6), (1302, 6), (3038,), (1302,))
```

```
#Get model Train
```

```
from sklearn.linear_model import LinearRegression
```

```
lr=LinearRegression()
```

```
lr.fit(X_train,y_train)
```

```
LinearRegression()
```

```
#Get modelprediction
```

```
y_pred=lr.predict(X_test)
```

```
y_pred.shape #Get Midel Predict
```

```
(1302,)
```

```
y_pred
```

```
array([502458.82786413, 646333.17428704, 521962.74075836, ...,  
       620183.32683781, 315403.8278857 , 731862.54196037])
```

```
#Get Model evaluation
```

```
from sklearn.metrics import mean_squared_error,mean_absolute_error,r2_score
```

```
mean_squared_error(y_test,y_pred)
```

```
193242972302.19553
```

```
mean_absolute_error(y_test,y_pred)
```

```
228808.95522977872
```

```
r2_score(y_test,y_pred)
```

```
0.4075563394370839
```

```
#get model Evaluation
```

```
from sklearn.metrics import mean_squared_error,mean_absolute_error,r2_score
```

```
mean_squared_error(y_test,y_pred)
```

```
193242972302.19553
```

```
mean_absolute_error(y_test,y_pred)
```

```
228808.95522977872
```

```
r2_score(y_test,y_pred)
```

```
0.4075563394370839
```

```
#get visualization of actual predicted Result
```

```
import matplotlib.pyplot as plt
```

```
plt.scatter(y_test,y_pred)
plt.xlabel('Actual Prices')
plt.ylabel('Predicted Prices')
plt.title("Actual price VS Prediction")
plt.show()
```




```
#Get Fure Prediction
```

```
df_new=df.sample(1)
```

```
df_new
```

Brand	Model	Year	Selling_Price	KM_Driven	Fuel	Seller_Type	Transmission
Maruti							

```
df_new.shape
```

```
(1, 9)
```

```
X_new=df_new.drop({'Brand', 'Model', 'Selling_Price'},axis=1)
```

```
y_pred_new=lr.predict(X_new)
```

```
y_pred_new
```

```
array([633659.87693281])
```

✓ 0s completed at 11:46 PM

● ×