

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_csv('/content/train-data.csv')
df
```

	Unnamed: 0	Name	Location	Year	Kilometers_Driven	Fuel_Type	T
0	0	Maruti Wagon R LXI CNG	Mumbai	2010	72000	CNG	
1	1	Hyundai Creta 1.6 CRDi SX Option	Pune	2015	41000	Diesel	
2	2	Honda Jazz V	Chennai	2011	46000	Petrol	
3	3	Maruti Ertiga VDI	Chennai	2012	87000	Diesel	
4	4	Audi A4 New 2.0 TDI Multitronic	Coimbatore	2013	40670	Diesel	
...

▼ TRAINING DATA

```
df.head()
```

	Unnamed: 0	Name	Location	Year	Kilometers_Driven	Fuel_Type	Trans
0	0	Maruti Wagon R LXI CNG	Mumbai	2010	72000	CNG	
1	1	Hyundai Creta 1.6 CRDi SX Option	Pune	2015	41000	Diesel	

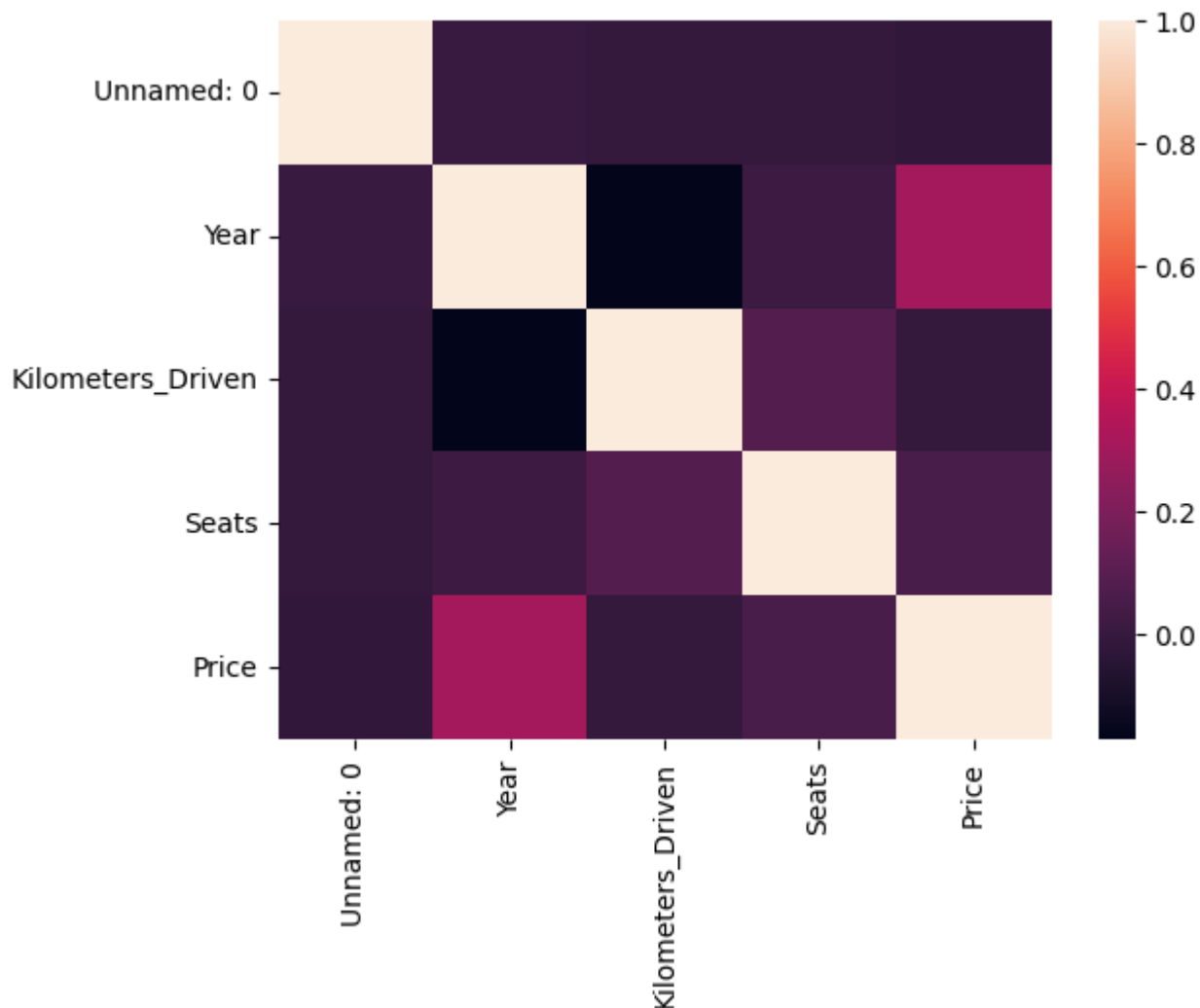
```
df.tail()
```

	Unnamed: 0	Name	Location	Year	Kilometers_Driven	Fuel_Type	Tr
6014	6014	Maruti Swift VDI	Delhi	2014	27365	Diesel	
6015	6015	Hyundai Xcent 1.1 CRDi S	Jaipur	2015	100000	Diesel	
6016	6016	Mahindra Xylo D4	Jaipur	2012	55000	Diesel	

DATA VISUALIZATION

```
sns.heatmap(df.corr())
```

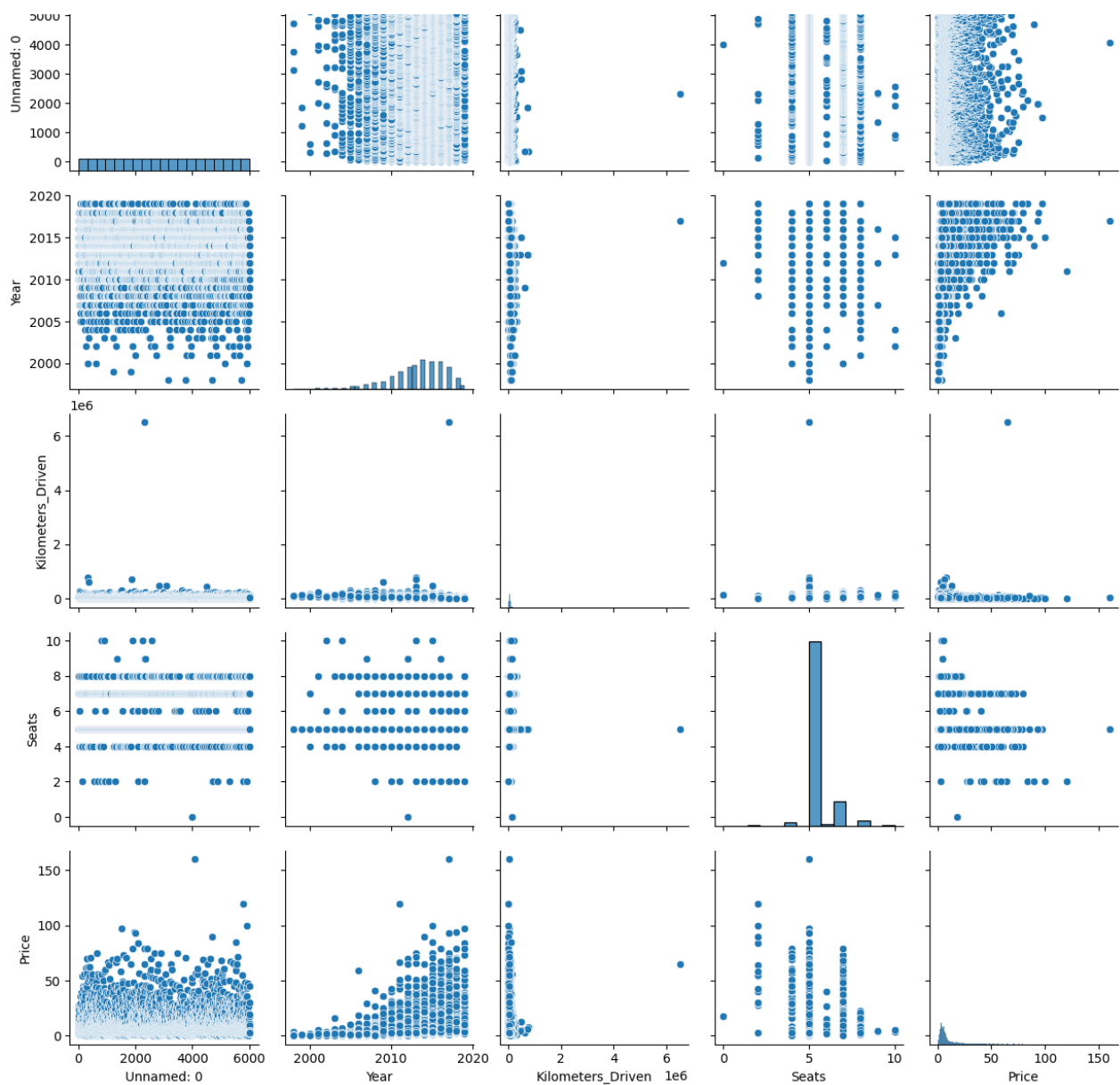
```
<ipython-input-689-74a73639b876>:2: FutureWarning: The default value of num
sns.heatmap(df.corr())
<Axes: >
```



```
sns.pairplot(df)
```

```
<seaborn.axisgrid.PairGrid at 0x7d5325912530>
```





```
df.isna().sum()
```

```
Unnamed: 0      0
Name            0
Location        0
Year            0
Kilometers_Driven  0
Fuel_Type       0
Transmission    0
Owner_Type      0
Mileage         2
Engine          36
Power           36
Seats           42
New_Price      5195
Price           0
dtype: int64
```

```
df.dtypes
```

```
Unnamed: 0      int64
Name          object
Location       object
Year          int64
Kilometers_Driven  int64
Fuel_Type      object
Transmission   object
Owner_Type     object
Mileage        object
Engine         object
Power          object
Seats         float64
New_Price     object
Price         float64
dtype: object
```

```
df.columns
```

```
Index(['Unnamed: 0', 'Name', 'Location', 'Year', 'Kilometers_Driven',
      'Fuel_Type', 'Transmission', 'Owner_Type', 'Mileage', 'Engine',
      'Power',
      'Seats', 'New_Price', 'Price'],
      dtype='object')
```

```
df.shape
```

```
(6019, 14)
```

```
df1=df['Location'].groupby(df['Location']).count()
```

```
df1
```

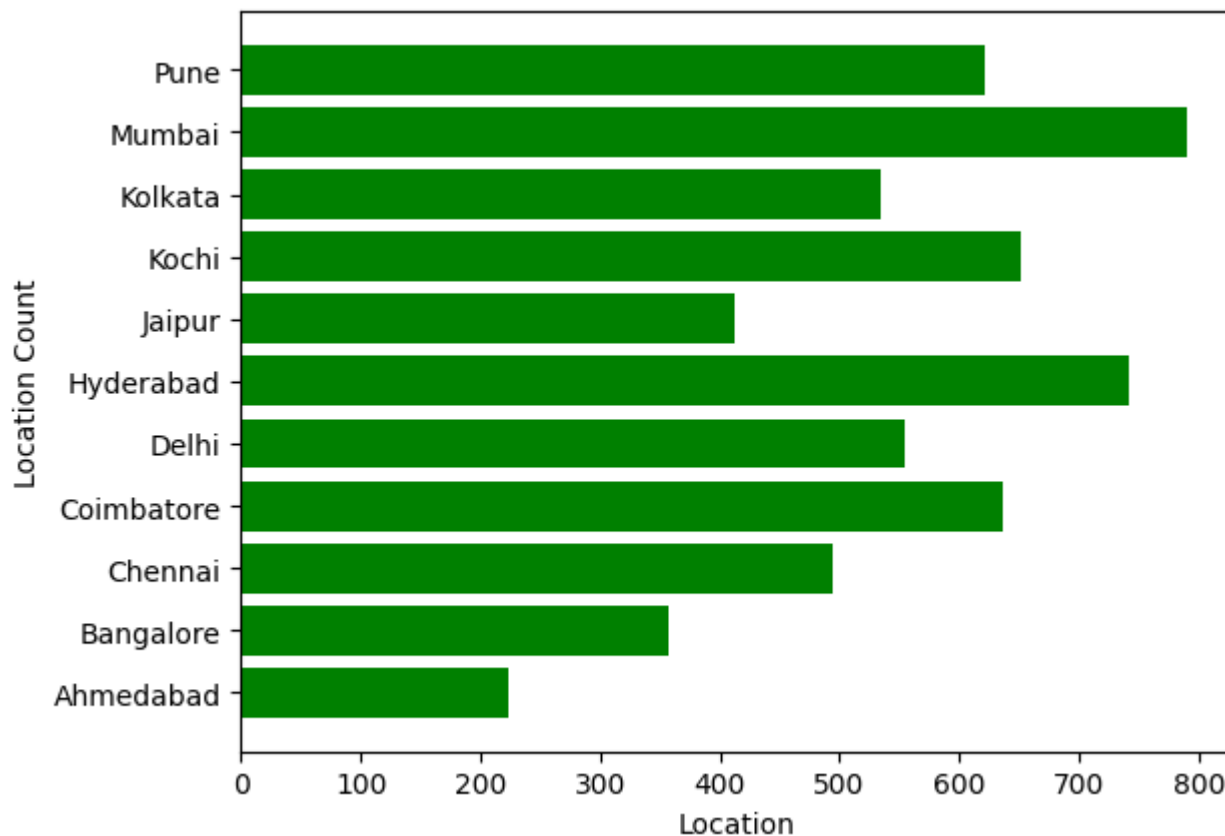
```
Location
Ahmedabad    224
Bangalore    358
Chennai      494
Coimbatore   636
Delhi        554
Hyderabad    742
Jaipur       413
Kochi        651
Kolkata      535
Mumbai       790
Pune         622
Name: Location, dtype: int64
```

```
plt.barh(df1.index,df1,color='green')
```

```
plt.xlabel('Location')
```

```
plt.ylabel('Location Count')
```

```
Text(0, 0.5, 'Location Count')
```

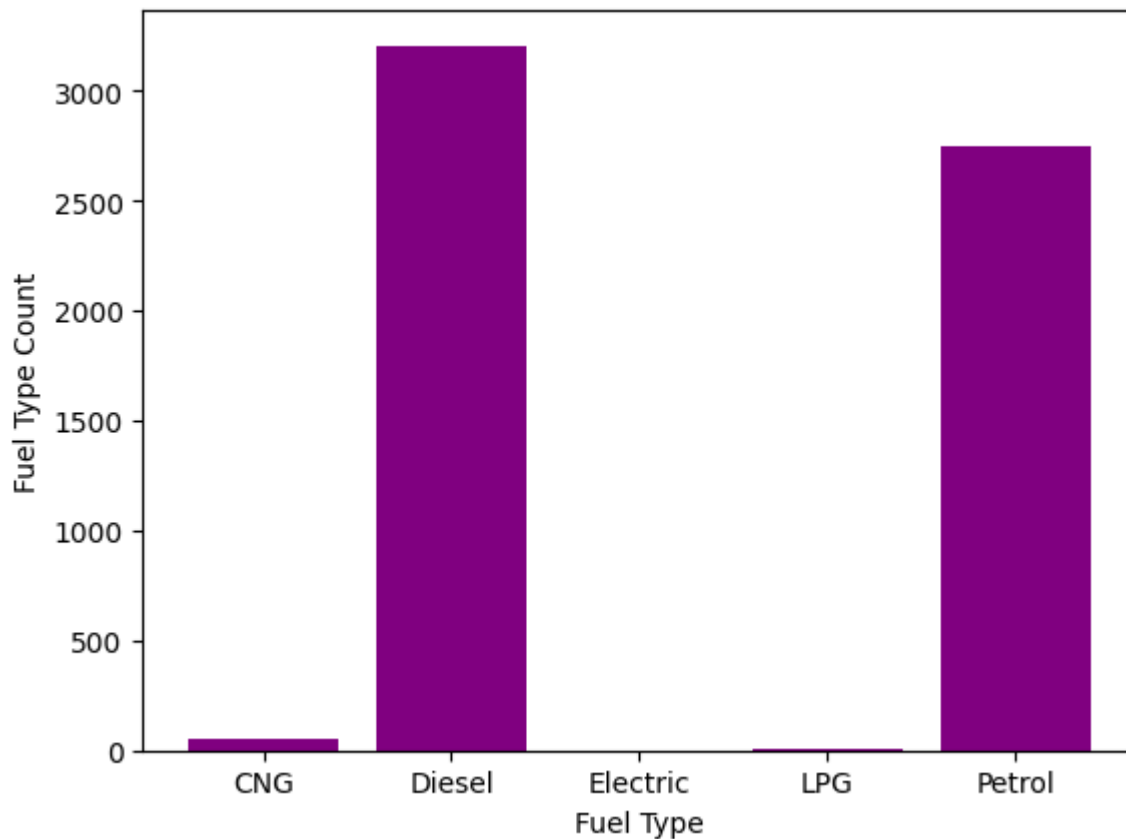


```
df2=df['Fuel_Type'].groupby(df['Fuel_Type']).count()
```

```
df2
```

```
Fuel_Type
CNG          56
Diesel      3205
Electric       2
LPG          10
Petrol      2746
Name: Fuel_Type, dtype: int64
```

```
plt.bar(df2.index,df2,color='purple')
plt.xlabel('Fuel Type')
plt.ylabel('Fuel Type Count')
Text(0, 0.5, 'Fuel Type Count')
```



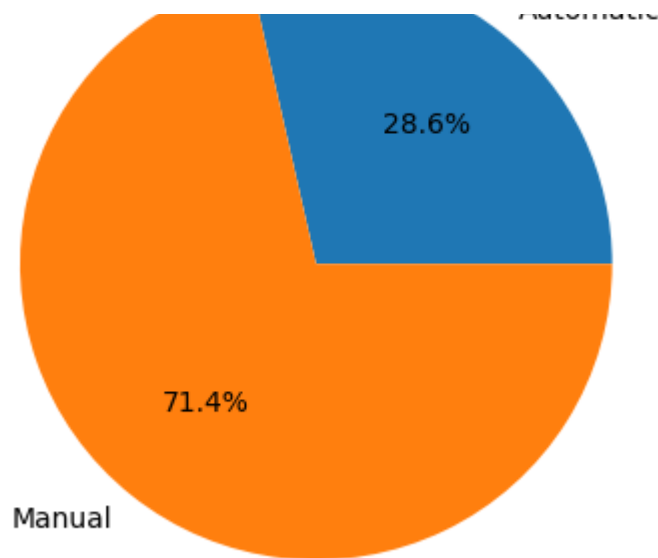
```
df3=df['Transmission'].groupby(df['Transmission']).count
df3
```

```
Transmission
Automatic    1720
Manual       4299
Name: Transmission, dtype: int64
```

```
plt.pie(df3,labels=['Automatic','Manual'],autopct='%1.1')
plt.legend(loc='upper left')
```

```
<matplotlib.legend.Legend at 0x7d5325a3bbe0>
```



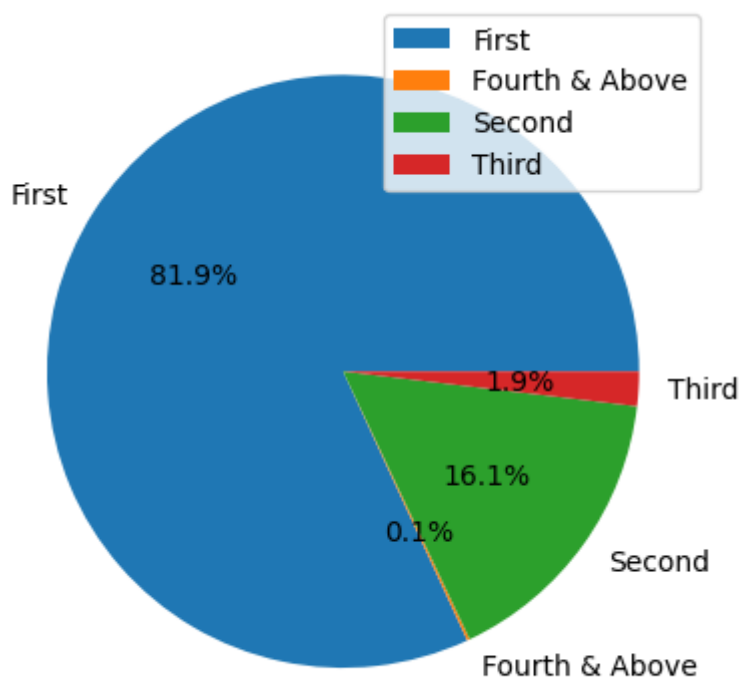


```
df4=df[ 'Owner_Type' ].groupby(df[ 'Owner_Type' ]).count()
df4
```

```
Owner_Type
First      4929
Fourth & Above    9
Second     968
Third      113
Name: Owner_Type, dtype: int64
```

```
plt.pie(df4,labels=['First','Fourth & Above','Second','Third'],
# plt.figure(figsize=(10,10))
plt.legend(loc='upper right')
```

<matplotlib.legend.Legend at 0x7d53258879a0>



```
df['Brand']=df['Name'].apply(lambda x:x.split()[0])
df
```

	Unnamed: 0	Name	Location	Year	Kilometers_Driven	Fuel_Type	T
0	0	Maruti Wagon R LXI CNG	Mumbai	2010	72000	CNG	
1	1	Hyundai Creta 1.6 CRDi SX Option	Pune	2015	41000	Diesel	
2	2	Honda Jazz V	Chennai	2011	46000	Petrol	
3	3	Maruti Ertiga VDI	Chennai	2012	87000	Diesel	
4	4	Audi A4 New 2.0 TDI Multitronic	Coimbatore	2013	40670	Diesel	
...

DATA PREPROCESSING

```
df5=pd.get_dummies(df[['Location','Fuel_Type','Transmis
df5
```

	Location_Bangalore	Location_Chennai	Location_Coimbatore	Location_D
0	0	0	0	
1	0	0	0	
2	0	1	0	
3	0	1	0	
4	0	0	1	
...
6014	0	0	0	
6015	0	0	0	
6016	0	0	0	
6017	0	0	0	

6018

0

0

0

6019 rows × 48 columns

```
df6=pd.concat([df,df5],axis=1)
df6
```

	Unnamed: 0	Name	Location	Year	Kilometers_Driven	Fuel_Type	T
0	0	Maruti Wagon R LXI CNG	Mumbai	2010	72000	CNG	
1	1	Hyundai Creta 1.6 CRDi SX Option	Pune	2015	41000	Diesel	
2	2	Honda Jazz V	Chennai	2011	46000	Petrol	
3	3	Maruti Ertiga VDI	Chennai	2012	87000	Diesel	
4	4	Audi A4 New 2.0 TDI Multitronic	Coimbatore	2013	40670	Diesel	
...
6014	6014	Maruti Swift VDI	Delhi	2014	27365	Diesel	
6015	6015	Hyundai Xcent 1.1 CRDi S	Jaipur	2015	100000	Diesel	
6016	6016	Mahindra Xylo D4 BSIV	Jaipur	2012	55000	Diesel	
6017	6017	Maruti Wagon R VXI	Kolkata	2013	46000	Petrol	
6018	6018	Chevrolet Beat Diesel	Hyderabad	2011	47000	Diesel	

6019 rows × 63 columns

df6.columns

Index(['Unnamed: 0', 'Name', 'Location', 'Year', 'Kilometers Driven',

```

    'Fuel_Type', 'Transmission', 'Owner_Type', 'Mileage', 'Engine',
    'Power',
    'Seats', 'New_Price', 'Price', 'Brand', 'Location_Bangalore',
    'Location_Chennai', 'Location_Coimbatore', 'Location_Delhi',
    'Location_Hyderabad', 'Location_Jaipur', 'Location_Kochi',
    'Location_Kolkata', 'Location_Mumbai', 'Location_Pune',
    'Fuel_Type_Diesel', 'Fuel_Type_Electric', 'Fuel_Type_LPG',
    'Fuel_Type_Petrol', 'Transmission_Manual', 'Owner_Type_Fourth &
Above',
    'Owner_Type_Second', 'Owner_Type_Third', 'Brand_Audi', 'Brand_BMW',
    'Brand_Bentley', 'Brand_Chevrolet', 'Brand_Datsun', 'Brand_Fiat',
    'Brand_Force', 'Brand_Ford', 'Brand_Honda', 'Brand_Hyundai',
    'Brand_ISUZU', 'Brand_Isuzu', 'Brand_Jaguar', 'Brand_Jeep',
    'Brand_Lamborghini', 'Brand_Land', 'Brand_Mahindra',
    'Brand_Maruti',
    'Brand_Mercedes-Benz', 'Brand_Mini', 'Brand_Mitsubishi',
    'Brand_Nissan',
    'Brand_Porsche', 'Brand_Renault', 'Brand_Skoda', 'Brand_Smart',
    'Brand_Tata', 'Brand_Toyota', 'Brand_Volkswagen', 'Brand_Volvo'],
dtype='object')

```

```

lst=['Unnamed: 0', 'Name', 'Location', 'Fuel_Type', 'Transm
dfe=df6.drop(lst,axis=1)
dfe

```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Price	Locatio
0	2010	72000	26.6 km/kg	998 CC	58.16 bhp	5.0	1.75	
1	2015	41000	19.67 kmpl	1582 CC	126.2 bhp	5.0	12.50	
2	2011	46000	18.2 kmpl	1199 CC	88.7 bhp	5.0	4.50	
3	2012	87000	20.77 kmpl	1248 CC	88.76 bhp	7.0	6.00	
4	2013	40670	15.2 kmpl	1968 CC	140.8 bhp	5.0	17.74	
...	
6014	2014	27365	28.4 kmpl	1248 CC	74 bhp	5.0	4.75	
6015	2015	100000	24.4 kmpl	1120 CC	71 bhp	5.0	4.00	
6016	2012	55000	14.0 kmpl	2498 CC	112 bhp	8.0	2.90	
6017	2013	46000	18.9 kmpl	998 CC	67.1 bhp	5.0	2.65	
6018	2011	47000	25.44 kmpl	936 CC	57.6 bhp	5.0	2.50	
----	--	.						

6019 rows × 50 columns

```

dfe['Mileage']=dfe['Mileage'].str.replace('km/kg','')
dfe['Mileage']=dfe['Mileage'].str.replace('kmpl','')
dfe['Engine']=dfe['Engine'].str.replace('CC','')
dfe['Power']=dfe['Power'].str.replace('bhp','')
dfe['Mileage']=dfe['Mileage'].str.replace('null','0')
dfe['Engine']=dfe['Engine'].str.replace('null','0')
dfe['Power']=dfe['Power'].str.replace('null','0')
dfe

```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Price	Locatio
0	2010	72000	26.6	998	58.16	5.0	1.75	
1	2015	41000	19.67	1582	126.2	5.0	12.50	
2	2011	46000	18.2	1199	88.7	5.0	4.50	
3	2012	87000	20.77	1248	88.76	7.0	6.00	
4	2013	40670	15.2	1968	140.8	5.0	17.74	
...	
6014	2014	27365	28.4	1248	74	5.0	4.75	
6015	2015	100000	24.4	1120	71	5.0	4.00	
6016	2012	55000	14.0	2498	112	8.0	2.90	
6017	2013	46000	18.9	998	67.1	5.0	2.65	
6018	2011	47000	25.44	936	57.6	5.0	2.50	

6019 rows × 50 columns

dfe.dtypes

```

Year                int64
Kilometers_Driven   int64
Mileage             object
Engine             object
Power              object
Seats              float64
Price              float64
Location_Bangalore   uint8
Location_Chennai     uint8
Location_Coimbatore  uint8
Location_Delhi       uint8
Location_Hyderabad   uint8
Location_Jaipur      uint8
Location_Kochi       uint8
Location_Kolkata     uint8

```

```

Location_Mumbai      uint8
Location_Pune        uint8
Fuel_Type_Diesel     uint8
Fuel_Type_LPG        uint8
Fuel_Type_Petrol     uint8
Transmission_Manual  uint8
Owner_Type_Fourth & Above  uint8
Owner_Type_Second    uint8
Owner_Type_Third     uint8
Brand_BMW            uint8
Brand_Bentley        uint8
Brand_Chevrolet      uint8
Brand_Datsun         uint8
Brand_Fiat           uint8
Brand_Ford           uint8
Brand_Honda          uint8
Brand_Hyundai        uint8
Brand_ISUZU          uint8
Brand_Isuzu          uint8
Brand_Jaguar         uint8
Brand_Jeep           uint8
Brand_Land           uint8
Brand_Mahindra       uint8
Brand_Maruti         uint8
Brand_Mercedes-Benz  uint8
Brand_Mini           uint8
Brand_Mitsubishi     uint8
Brand_Nissan         uint8
Brand_Porsche        uint8
Brand_Renault        uint8
Brand_Skoda          uint8
Brand_Tata           uint8
Brand_Toyota         uint8
Brand_Volkswagen     uint8
Brand_Volvo          uint8
dtype: object

```

```
#Data type conversion
```

```
#Mileage ==>float
```

```
#Engine ==>float
```

```
#Power ==>float
```

```
dfe['Mileage']=dfe['Mileage'].astype(float)
```

```
dfe['Engine']=dfe['Engine'].astype(float)
```

```
dfe['Power']=dfe['Power'].astype(float)
```

```
dfe
```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Price	Locati
0	2010	72000	26.60	998.0	58.16	5.0	1.75	
1	2015	41000	19.67	1582.0	126.20	5.0	12.50	
2	2011	46000	18.20	1199.0	88.70	5.0	4.50	
3	2012	87000	20.77	1248.0	88.76	7.0	6.00	

4	2013	40670	15.20	1968.0	140.80	5.0	17.74
...
6014	2014	27365	28.40	1248.0	74.00	5.0	4.75
6015	2015	100000	24.40	1120.0	71.00	5.0	4.00
6016	2012	55000	14.00	2498.0	112.00	8.0	2.90
6017	2013	46000	18.90	998.0	67.10	5.0	2.65
6018	2011	47000	25.44	936.0	57.60	5.0	2.50

6019 rows × 50 columns

dfe.dtypes

Year	int64
Kilometers_Driven	int64
Mileage	float64
Engine	float64
Power	float64
Seats	float64
Price	float64
Location_Bangalore	uint8
Location_Chennai	uint8
Location_Coimbatore	uint8
Location_Delhi	uint8
Location_Hyderabad	uint8
Location_Jaipur	uint8
Location_Kochi	uint8
Location_Kolkata	uint8
Location_Mumbai	uint8
Location_Pune	uint8
Fuel_Type_Diesel	uint8
Fuel_Type_LPG	uint8
Fuel_Type_Petrol	uint8
Transmission_Manual	uint8
Owner_Type_Fourth & Above	uint8
Owner_Type_Second	uint8
Owner_Type_Third	uint8
Brand_BMW	uint8
Brand_Bentley	uint8
Brand_Chevrolet	uint8
Brand_Datsun	uint8
Brand_Fiat	uint8
Brand_Ford	uint8
Brand_Honda	uint8
Brand_Hyundai	uint8
Brand_ISUZU	uint8
Brand_Isuzu	uint8
Brand_Jaguar	uint8
Brand_Jeep	uint8
Brand_Land	uint8
Brand_Mahindra	uint8
Brand_Maruti	uint8
Brand_Mercedes-Benz	uint8

```

-----
Brand_Mini                uint8
Brand_Mitsubishi          uint8
Brand_Nissan               uint8
Brand_Porsche             uint8
Brand_Renault             uint8
Brand_Skoda               uint8
Brand_Tata                uint8
Brand_Toyota              uint8
Brand_Volkswagen          uint8
Brand_Volvo               uint8
dtype: object

```

```
dfe.isna().sum()
```

```

Year                0
Kilometers_Driven  0
Mileage             2
Engine             36
Power              36
Seats             42
Price              0
Location_Bangalore  0
Location_Chennai    0
Location_Coimbatore  0
Location_Delhi       0
Location_Hyderabad   0
Location_Jaipur       0
Location_Kochi        0
Location_Kolkata      0
Location_Mumbai       0
Location_Pune         0
Fuel_Type_Diesel      0
Fuel_Type_LPG         0
Fuel_Type_Petrol      0
Transmission_Manual   0
Owner_Type_Fourth & Above  0
Owner_Type_Second     0
Owner_Type_Third      0
Brand_BMW            0
Brand_Bentley         0
Brand_Chevrolet       0
Brand_Datsun          0
Brand_Fiat            0
Brand_Ford            0
Brand_Honda           0
Brand_Hyundai         0
Brand_ISUZU           0
Brand_Isuzu           0
Brand_Jaguar          0
Brand_Jeep            0
Brand_Land            0
Brand_Mahindra        0
Brand_Maruti          0
Brand_Mercedes-Benz   0
Brand_Mini            0
Brand_Mitsubishi      0
Brand_Nissan          0
Brand_Porsche         0

```

```

Brand_Renault      0
Brand_Skoda         0
Brand_Tata          0
Brand_Toyota        0
Brand_Volkswagen    0
Brand_Volvo         0
dtype: int64

```

```

dfe.loc[dfe.Mileage==0, 'Mileage']=np.NaN
dfe.loc[dfe.Power==0, 'Power']=np.NaN
dfe.loc[dfe.Engine==0, 'Engine']=np.NaN
dfe

```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Price	Locati
0	2010	72000	26.60	998.0	58.16	5.0	1.75	
1	2015	41000	19.67	1582.0	126.20	5.0	12.50	
2	2011	46000	18.20	1199.0	88.70	5.0	4.50	
3	2012	87000	20.77	1248.0	88.76	7.0	6.00	
4	2013	40670	15.20	1968.0	140.80	5.0	17.74	
...	
6014	2014	27365	28.40	1248.0	74.00	5.0	4.75	
6015	2015	100000	24.40	1120.0	71.00	5.0	4.00	
6016	2012	55000	14.00	2498.0	112.00	8.0	2.90	
6017	2013	46000	18.90	998.0	67.10	5.0	2.65	
6018	2011	47000	25.44	936.0	57.60	5.0	2.50	

6019 rows × 50 columns

```

dfe.isna().sum()

```

```

Year      0
Kilometers_Driven  0
Mileage    70
Engine     36
Power     143
Seats      42
Price      0
Location_Bangalore  0
Location_Chennai    0
Location_Coimbatore  0
Location_Delhi       0
Location_Hyderabad   0
Location_Jaipur       0
Location_Kochi        0
Location_Kolkata      0
Location_Mumbai       0

```

```

Location_Pune      0
Fuel_Type_Diesel   0
Fuel_Type_LPG      0
Fuel_Type_Petrol   0
Transmission_Manual 0
Owner_Type_Fourth & Above 0
Owner_Type_Second  0
Owner_Type_Third   0
Brand_BMW          0
Brand_Bentley      0
Brand_Chevrolet    0
Brand_Datsun       0
Brand_Fiat         0
Brand_Ford         0
Brand_Honda        0
Brand_Hyundai      0
Brand_ISUZU        0
Brand_Isuzu        0
Brand_Jaguar       0
Brand_Jeep         0
Brand_Land         0
Brand_Mahindra     0
Brand_Maruti       0
Brand_Mercedes-Benz 0
Brand_Mini         0
Brand_Mitsubishi   0
Brand_Nissan       0
Brand_Porsche      0
Brand_Renault      0
Brand_Skoda        0
Brand_Tata         0
Brand_Toyota       0
Brand_Volkswagen   0
Brand_Volvo        0
dtype: int64

```

```

dfe['Mileage']=dfe['Mileage'].fillna(dfe['Mileage'].mean())
dfe['Engine']=dfe['Engine'].fillna(dfe['Engine'].mean())
dfe['Power']=dfe['Power'].fillna(dfe['Power'].mean())
dfe['Seats']=dfe['Seats'].fillna(dfe['Seats'].mode()[0])
dfe

```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Price	Locati
0	2010	72000	26.60	998.0	58.16	5.0	1.75	
1	2015	41000	19.67	1582.0	126.20	5.0	12.50	
2	2011	46000	18.20	1199.0	88.70	5.0	4.50	
3	2012	87000	20.77	1248.0	88.76	7.0	6.00	
4	2013	40670	15.20	1968.0	140.80	5.0	17.74	
...	
6014	2014	27365	28.40	1248.0	74.00	5.0	4.75	

6015	2015	100000	24.40	1120.0	71.00	5.0	4.00
6016	2012	55000	14.00	2498.0	112.00	8.0	2.90
6017	2013	46000	18.90	998.0	67.10	5.0	2.65
6018	2011	47000	25.44	936.0	57.60	5.0	2.50

6019 rows × 50 columns

```
dfe.isna().sum()
```

```

Year                                0
Kilometers_Driven                   0
Mileage                             0
Engine                             0
Power                               0
Seats                               0
Price                               0
Location_Bangalore                  0
Location_Chennai                    0
Location_Coimbatore                 0
Location_Delhi                      0
Location_Hyderabad                  0
Location_Jaipur                     0
Location_Kochi                      0
Location_Kolkata                    0
Location_Mumbai                     0
Location_Pune                       0
Fuel_Type_Diesel                    0
Fuel_Type_LPG                       0
Fuel_Type_Petrol                    0
Transmission_Manual                 0
Owner_Type_Fourth & Above            0
Owner_Type_Second                   0
Owner_Type_Third                    0
Brand_BMW                           0
Brand_Bentley                       0
Brand_Chevrolet                     0
Brand_Datsun                        0
Brand_Fiat                          0
Brand_Ford                          0
Brand_Honda                         0
Brand_Hyundai                       0
Brand_ISUZU                         0
Brand_Isuzu                         0
Brand_Jaguar                        0
Brand_Jeep                          0
Brand_Land                          0
Brand_Mahindra                      0
Brand_Maruti                        0
Brand_Mercedes-Benz                 0
Brand_Mini                          0
Brand_Mitsubishi                    0
Brand_Nissan                        0
Brand_Porsche                       0
Brand_Renault                       0
_._._                               -

```

```

Brand_Skoda      0
Brand_Tata       0
Brand_Toyota     0
Brand_Volkswagen 0
Brand_Volvo      0
dtype: int64

```

dfe.dtypes

```

Year              int64
Kilometers_Driven int64
Mileage           float64
Engine           float64
Power            float64
Seats            float64
Price            float64
Location_Bangalore      uint8
Location_Chennai        uint8
Location_Coimbatore      uint8
Location_Delhi           uint8
Location_Hyderabad       uint8
Location_Jaipur          uint8
Location_Kochi           uint8
Location_Kolkata         uint8
Location_Mumbai          uint8
Location_Pune            uint8
Fuel_Type_Diesel         uint8
Fuel_Type_LPG            uint8
Fuel_Type_Petrol         uint8
Transmission_Manual      uint8
Owner_Type_Fourth & Above uint8
Owner_Type_Second        uint8
Owner_Type_Third         uint8
Brand_BMW                uint8
Brand_Bentley            uint8
Brand_Chevrolet          uint8
Brand_Datsun             uint8
Brand_Fiat               uint8
Brand_Ford               uint8
Brand_Honda              uint8
Brand_Hyundai            uint8
Brand_ISUZU              uint8
Brand_Isuzu              uint8
Brand_Jaguar             uint8
Brand_Jeep               uint8
Brand_Land               uint8
Brand_Mahindra           uint8
Brand_Maruti             uint8
Brand_Mercedes-Benz      uint8
Brand_Mini               uint8
Brand_Mitsubishi         uint8
Brand_Nissan             uint8
Brand_Porsche            uint8
Brand_Renault            uint8
Brand_Skoda              uint8
Brand_Tata               uint8
Brand_Toyota             uint8
Brand_Volkswagen         uint8
Brand_Volvo              uint8

```

```
brand_vocals
dtype: object
```

```
dfc
```

```
dfc
```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Price	Locati
0	2010	72000	26.60	998.0	58.16	5.0	1.75	
1	2015	41000	19.67	1582.0	126.20	5.0	12.50	
2	2011	46000	18.20	1199.0	88.70	5.0	4.50	
3	2012	87000	20.77	1248.0	88.76	7.0	6.00	
4	2013	40670	15.20	1968.0	140.80	5.0	17.74	
...	
6014	2014	27365	28.40	1248.0	74.00	5.0	4.75	
6015	2015	100000	24.40	1120.0	71.00	5.0	4.00	
6016	2012	55000	14.00	2498.0	112.00	8.0	2.90	
6017	2013	46000	18.90	998.0	67.10	5.0	2.65	
6018	2011	47000	25.44	936.0	57.60	5.0	2.50	

```
6019 rows × 50 columns
```

```
x=dfc.drop(['Price'],axis=1)
```

```
x
```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Location_Bang
0	2010	72000	26.60	998.0	58.16	5.0	
1	2015	41000	19.67	1582.0	126.20	5.0	
2	2011	46000	18.20	1199.0	88.70	5.0	
3	2012	87000	20.77	1248.0	88.76	7.0	
4	2013	40670	15.20	1968.0	140.80	5.0	
...	
6014	2014	27365	28.40	1248.0	74.00	5.0	
6015	2015	100000	24.40	1120.0	71.00	5.0	
6016	2012	55000	14.00	2498.0	112.00	8.0	
6017	2013	46000	18.90	998.0	67.10	5.0	
6018	2011	47000	25.44	936.0	57.60	5.0	

```
6019 rows × 49 columns
```

```
x.shape
```

```
(6019, 49)
```

```
y=dfe['Price']
```

```
y
```

```
0      1.75
1     12.50
2      4.50
3      6.00
4     17.74
```

```
...
6014    4.75
6015    4.00
6016    2.90
6017    2.65
6018    2.50
```

```
Name: Price, Length: 6019, dtype: float64
```

TESTING DATA

```
df_test=pd.read_csv('/content/test-data.csv')
```

```
df_test
```

	Unnamed: 0	Name	Location	Year	Kilometers_Driven	Fuel_Type
0	0	Maruti Alto K10 LXI CNG	Delhi	2014	40929	CNG
1	1	Maruti Alto 800 2016-2019 LXI	Coimbatore	2013	54493	Petrol
2	2	Toyota Innova Crysta Touring Sport 2.4 MT	Mumbai	2017	34000	Diesel
3	3	Toyota Etios Liva GD	Hyderabad	2012	139000	Diesel
4	4	Hyundai i20 Magna	Mumbai	2014	29000	Petrol

```
df_test.head()
```

	Unnamed: 0	Name	Location	Year	Kilometers_Driven	Fuel_Type	Trans
0	0	Maruti Alto K10 LXI CNG	Delhi	2014	40929	CNG	
1	1	Maruti Alto 800 2016-2019 LXI	Coimbatore	2013	54493	Petrol	

```
df_test.tail()
```

	Unnamed: 0	Name	Location	Year	Kilometers_Driven	Fuel_Type	Trans
1229	1229	Volkswagen Vento Diesel Trendline	Hyderabad	2011	89411	Diesel	
1230	1230	Volkswagen Polo GT TSI	Mumbai	2015	59000	Petrol	
		Nissan					

```
df_test.dtypes
```

```

Unnamed: 0      int64
Name           object
Location        object
Year           int64
Kilometers_Driven  int64
Fuel_Type       object
Transmission     object
Owner_Type      object
Mileage         object
Engine          object
Power           object
Seats          float64
New_Price       object
dtype: object

```

```
df_test.isna().sum()
```

```

Unnamed: 0      0
Name           0
Location        0
Year           0
Kilometers_Driven  0
Fuel_Type       0
Transmission     0
Owner_Type      0
Mileage         0
Engine          10
Power           10
Seats          11

```

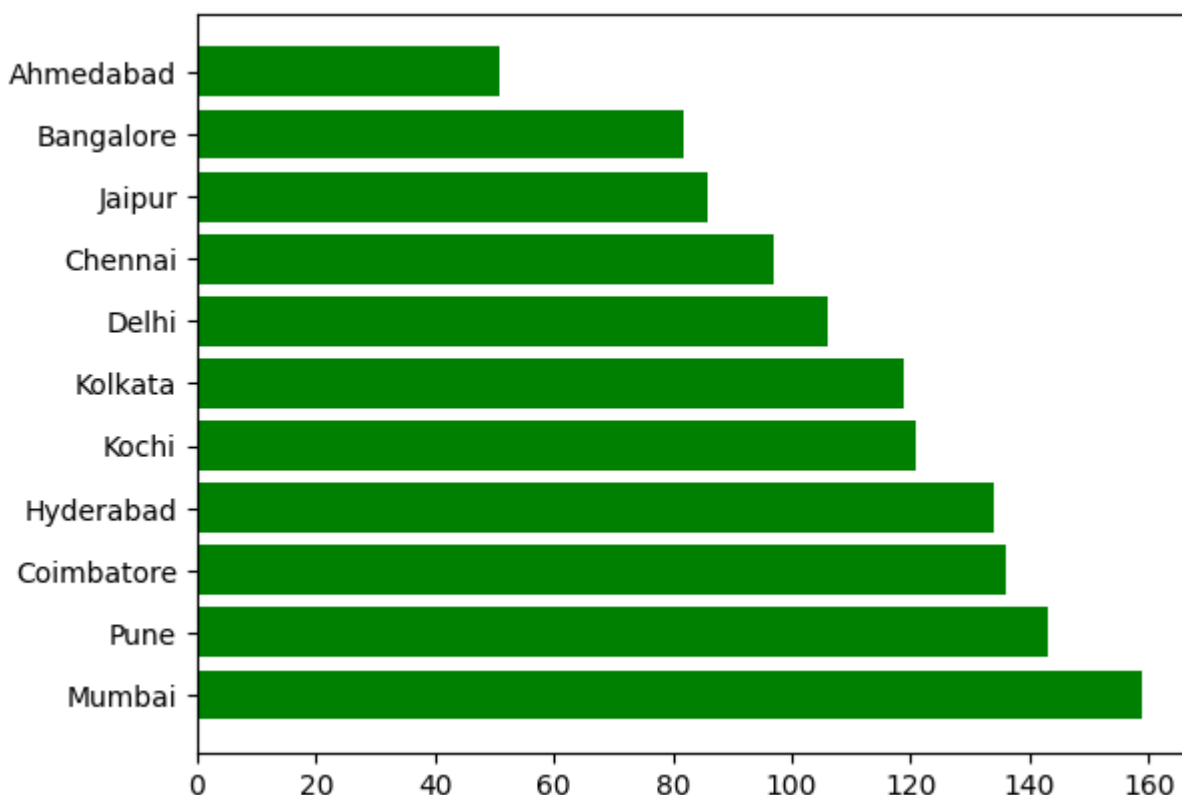
```
New_Price      1052  
dtype: int64
```

```
loc=df_test['Location'].value_counts()  
loc
```

```
Mumbai      159  
Pune        143  
Coimbatore   136  
Hyderabad    134  
Kochi        121  
Kolkata      119  
Delhi        106  
Chennai      97  
Jaipur       86  
Bangalore    82  
Ahmedabad    51  
Name: Location, dtype: int64
```

DATA VISUALIZATION

```
plt.barh(loc.index,loc,color='green')  
plt.show()
```



```
fuel=df_test['Fuel_Type'].value_counts()  
fuel
```

```
Diesel      647
```

```

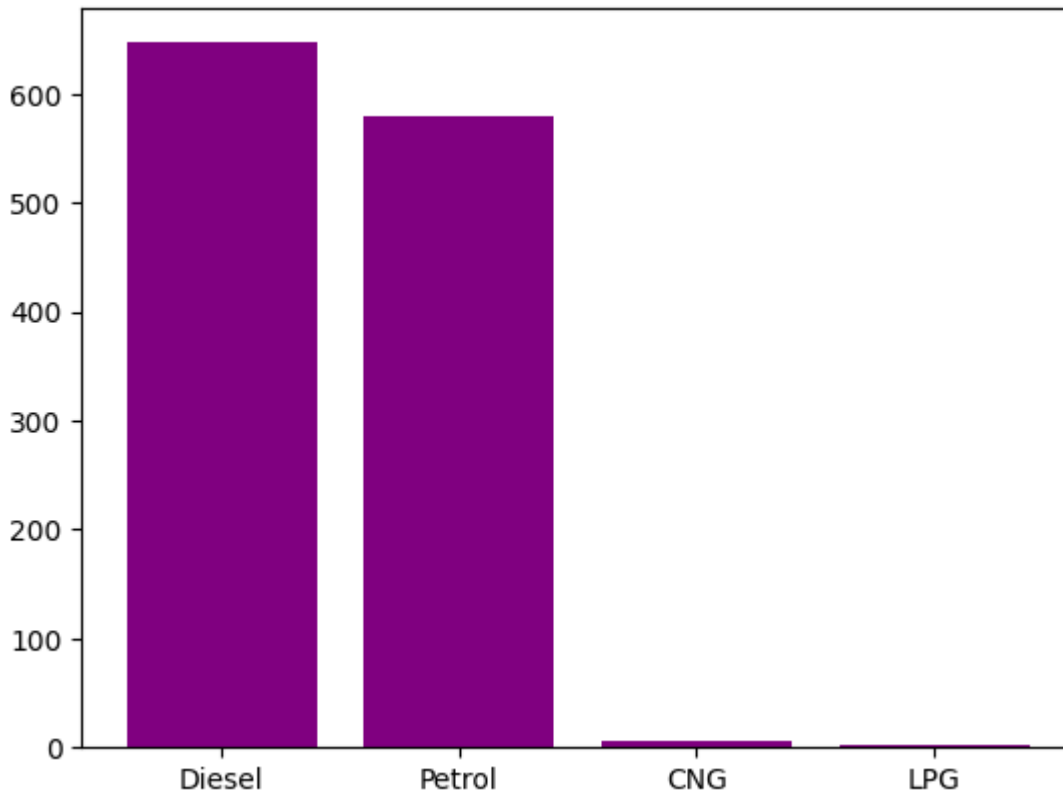
-----
Petrol      579
CNG          6
LPG          2
Name: Fuel_Type, dtype: int64

```

```

plt.bar(fuel.index, fuel, color='purple')
plt.show()

```



```
df_test.columns
```

```

Index(['Unnamed: 0', 'Name', 'Location', 'Year', 'Kilometers_Driven',
      'Fuel_Type', 'Transmission', 'Owner_Type', 'Mileage', 'Engine',
      'Power',
      'Seats', 'New_Price'],
      dtype='object')

```

```

trans=df_test['Transmission'].value_counts()
trans

```

```

Manual      905
Automatic   329
Name: Transmission, dtype: int64

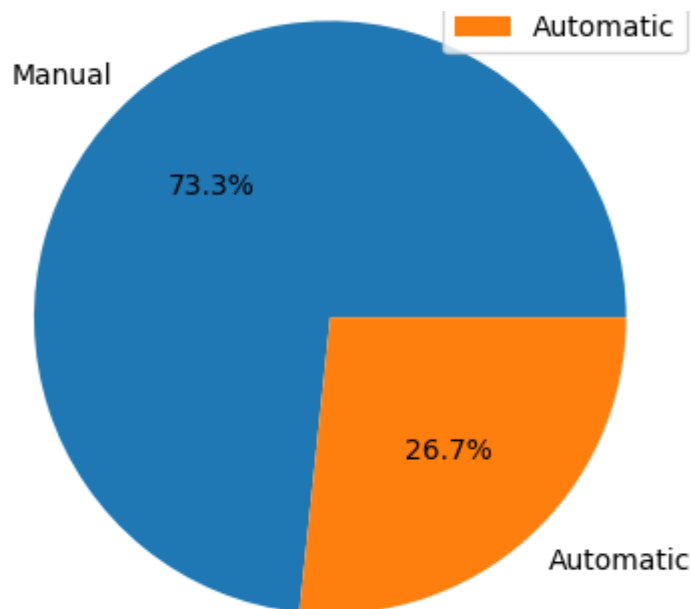
```

```

plt.pie(trans, labels=['Manual', 'Automatic'], autopct='%1
plt.legend(loc='upper right')
plt.show()

```

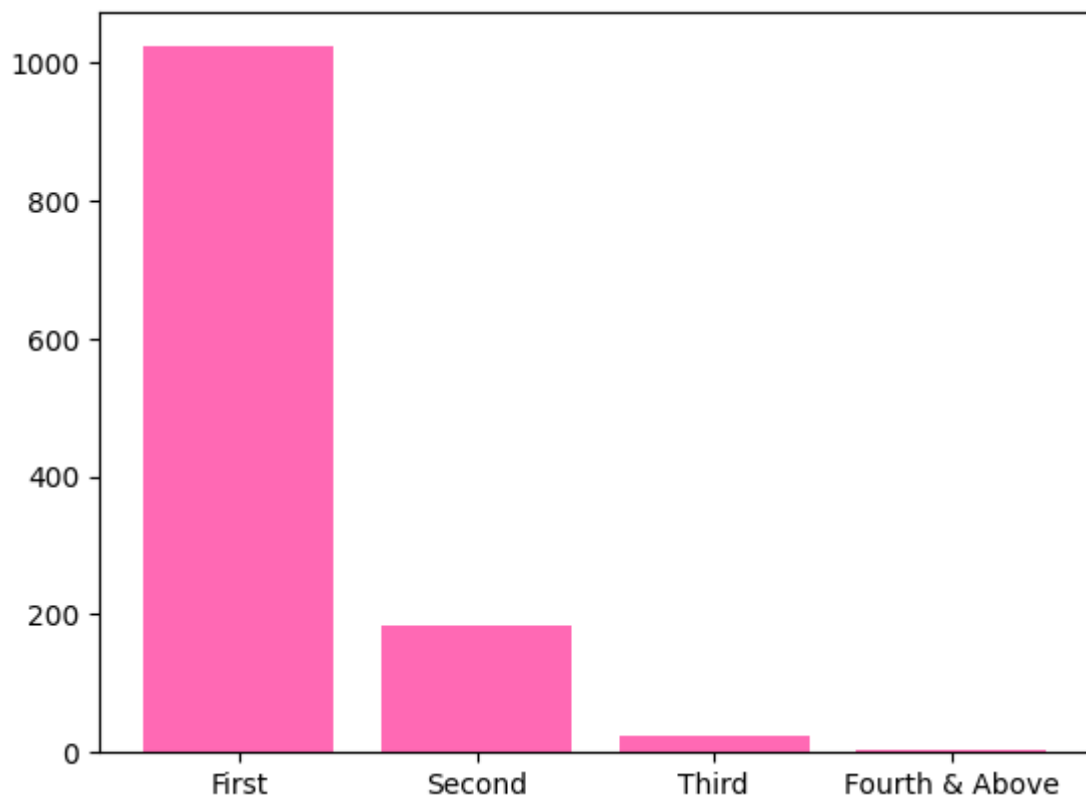




```
owner=df_test['Owner_Type'].value_counts()  
owner
```

```
First          1023  
Second         184  
Third           24  
Fourth & Above    3  
Name: Owner_Type, dtype: int64
```

```
plt.bar(owner.index,owner,color='hotpink')  
plt.show()
```




```
df_test['Brand']=df_test['Name'].apply(lambda x:x.split
df_test
```

	Unnamed: 0	Name	Location	Year	Kilometers_Driven	Fuel_Type
0	0	Maruti Alto K10 LXI CNG	Delhi	2014	40929	CNG
1	1	Maruti Alto 800 2016-2019 LXI	Coimbatore	2013	54493	Petrol
2	2	Toyota Innova Crysta Touring Sport 2.4 MT	Mumbai	2017	34000	Diesel
3	3	Toyota Etios Liva GD	Hyderabad	2012	139000	Diesel
4	4	Hyundai i20 Magna	Mumbai	2014	29000	Petrol

DATA PREPROCESSING

```
df_get=pd.get_dummies(df_test[['Location','Fuel_Type'],
df_get
```

	Location_Bangalore	Location_Chennai	Location_Coimbatore	Location_Delhi
0	0	0	0	0
1	0	0	1	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
...
1229	0	0	0	0
1230	0	0	0	0
1231	0	0	0	0
1232	0	0	0	0

1233

0

0

0

1234 rows × 45 columns

```
df_con=pd.concat([df_test,df_get],axis=1)
df_con
```

	Unnamed: 0	Name	Location	Year	Kilometers_Driven	Fuel_Type
0	0	Maruti Alto K10 LXI CNG	Delhi	2014	40929	CNG
1	1	Maruti Alto 800 2016-2019 LXI	Coimbatore	2013	54493	Petrol
2	2	Toyota Innova Crysta Touring Sport 2.4 MT	Mumbai	2017	34000	Diesel
3	3	Toyota Etios Liva GD	Hyderabad	2012	139000	Diesel
4	4	Hyundai i20 Magna	Mumbai	2014	29000	Petrol
...
1229	1229	Volkswagen Vento Diesel Trendline	Hyderabad	2011	89411	Diesel
1230	1230	Volkswagen Polo GT TSI	Mumbai	2015	59000	Petrol
1231	1231	Nissan Micra Diesel XV	Kolkata	2012	28000	Diesel
1232	1232	Volkswagen Polo GT TSI	Pune	2013	52262	Petrol
1233	1233	Mercedes- Benz E-Class 2009-2013 E 220 CDI Avan...	Kochi	2014	72443	Diesel

1234 rows × 59 columns

df_con.columns

```
Index(['Unnamed: 0', 'Name', 'Location', 'Year', 'Kilometers_Driven',
      'Fuel_Type', 'Transmission', 'Owner_Type', 'Mileage', 'Engine',
      'Power',
      'Seats', 'New_Price', 'Brand', 'Location_Bangalore',
      'Location_Chennai',
      'Location_Coimbatore', 'Location_Delhi', 'Location_Hyderabad',
      'Location_Jaipur', 'Location_Kochi', 'Location_Kolkata',
      'Location_Mumbai', 'Location_Pune', 'Fuel_Type_Diesel',
      'Fuel_Type_LPG',
      'Fuel_Type_Petrol', 'Transmission_Manual', 'Owner_Type_Fourth &
Above',
      'Owner_Type_Second', 'Owner_Type_Third', 'Brand_BMW',
      'Brand_Bentley',
      'Brand_Chevrolet', 'Brand_Datsun', 'Brand_Fiat', 'Brand_Ford',
      'Brand_Hindustan', 'Brand_Honda', 'Brand_Hyundai', 'Brand_ISUZU',
      'Brand_Isuzu', 'Brand_Jaguar', 'Brand_Jeep', 'Brand_Land',
      'Brand_Mahindra', 'Brand_Maruti', 'Brand_Mercedes-Benz',
      'Brand_Mini',
      'Brand_Mitsubishi', 'Brand_Nissan', 'Brand_OpelCorsa',
      'Brand_Porsche',
      'Brand_Renault', 'Brand_Skoda', 'Brand_Tata', 'Brand_Toyota',
      'Brand_Volkswagen', 'Brand_Volvo'],
      dtype='object')
```

```
lst=['Unnamed: 0', 'Name', 'Location','Fuel_Type', 'Tra
'Brand_OpelCorsa']
dfn=df_con.drop(lst,axis=1)
dfn
```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Location_Bangalore
0	2014	40929	32.26 km/kg	998 CC	58.2 bhp	4.0	
1	2013	54493	24.7 kmpl	796 CC	47.3 bhp	5.0	
2	2017	34000	13.68 kmpl	2393 CC	147.8 bhp	7.0	
3	2012	139000	23.59 kmpl	1364 CC	null bhp	5.0	
4	2014	29000	18.5 kmpl	1197 CC	82.85 bhp	5.0	
...	
1229	2011	89411	20.54 kmpl	1598 CC	103.6 bhp	5.0	
1230	2015	59000	17.21 kmpl	1197 CC	103.6 bhp	5.0	

1231	2012	28000	23.08 kmpl	1461 CC	63.1 bhp	5.0
1232	2013	52262	17.2 kmpl	1197 CC	103.6 bhp	5.0
1233	2014	72443	10.0 kmpl	2148 CC	170 bhp	5.0

1234 rows × 49 columns

dfn.columns

```
Index(['Year', 'Kilometers_Driven', 'Mileage', 'Engine', 'Power', 'Seats',
      'Location_Bangalore', 'Location_Chennai', 'Location_Coimbatore',
      'Location_Delhi', 'Location_Hyderabad', 'Location_Jaipur',
      'Location_Kochi', 'Location_Kolkata', 'Location_Mumbai',
      'Location_Pune', 'Fuel_Type_Diesel', 'Fuel_Type_LPG',
      'Fuel_Type_Petrol', 'Transmission_Manual', 'Owner_Type_Fourth &
      Above',
      'Owner_Type_Second', 'Owner_Type_Third', 'Brand_BMW',
      'Brand_Bentley',
      'Brand_Chevrolet', 'Brand_Datsun', 'Brand_Fiat', 'Brand_Ford',
      'Brand_Honda', 'Brand_Hyundai', 'Brand_ISUZU', 'Brand_Isuzu',
      'Brand_Jaguar', 'Brand_Jeep', 'Brand_Land', 'Brand_Mahindra',
      'Brand_Maruti', 'Brand_Mercedes-Benz', 'Brand_Mini',
      'Brand_Mitsubishi',
      'Brand_Nissan', 'Brand_Porsche', 'Brand_Renault', 'Brand_Skoda',
      'Brand_Tata', 'Brand_Toyota', 'Brand_Volkswagen', 'Brand_Volvo'],
      dtype='object')
```

```
dfn['Mileage']=dfn['Mileage'].str.replace('km/kg','')
dfn['Mileage']=dfn['Mileage'].str.replace('kmpl','')
dfn['Engine']=dfn['Engine'].str.replace('CC','')
dfn['Power']=dfn['Power'].str.replace('bhp','')
dfn
```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Location_Bangalore
0	2014	40929	32.26	998	58.2	4.0	
1	2013	54493	24.7	796	47.3	5.0	
2	2017	34000	13.68	2393	147.8	7.0	
3	2012	139000	23.59	1364	null	5.0	
4	2014	29000	18.5	1197	82.85	5.0	
...	
1229	2011	89411	20.54	1598	103.6	5.0	
1230	2015	59000	17.21	1197	103.6	5.0	
1231	2012	28000	23.08	1461	63.1	5.0	

```

1232 2013          52262      17.2    1197  103.6    5.0
1233 2014          72443      10.0    2148   170    5.0

```

1234 rows × 49 columns

```

lst1=['Mileage','Engine','Power']
for i in lst1:
    nulls=dfn[i].str.contains('null')
    print(nulls.value_counts())

```

```

False    1234
Name: Mileage, dtype: int64
False    1224
Name: Engine, dtype: int64
False    1202
True       22
Name: Power, dtype: int64

```

```

dfn['Mileage']=dfn['Mileage'].str.replace('null','0')
dfn['Engine']=dfn['Engine'].str.replace('null','0')
dfn['Power']=dfn['Power'].str.replace('null','0')
dfn

```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Location_Bang
0	2014	40929	32.26	998	58.2	4.0	
1	2013	54493	24.7	796	47.3	5.0	
2	2017	34000	13.68	2393	147.8	7.0	
3	2012	139000	23.59	1364	0	5.0	
4	2014	29000	18.5	1197	82.85	5.0	
...	
1229	2011	89411	20.54	1598	103.6	5.0	
1230	2015	59000	17.21	1197	103.6	5.0	
1231	2012	28000	23.08	1461	63.1	5.0	
1232	2013	52262	17.2	1197	103.6	5.0	
1233	2014	72443	10.0	2148	170	5.0	

1234 rows × 49 columns

```

lst1=['Mileage','Engine','Power']

```

```
for i in lst1:
    nulls=dfn[i].str.contains('null')
    print(nulls.value_counts())
```

```
False      1234
Name: Mileage, dtype: int64
False      1224
Name: Engine, dtype: int64
False      1224
Name: Power, dtype: int64
```

dfn.dtypes

```
Year                int64
Kilometers_Driven   int64
Mileage             object
Engine             object
Power              object
Seats             float64
Location_Bangalore   uint8
Location_Chennai     uint8
Location_Coimbatore  uint8
Location_Delhi       uint8
Location_Hyderabad   uint8
Location_Jaipur      uint8
Location_Kochi       uint8
Location_Kolkata     uint8
Location_Mumbai      uint8
Location_Pune        uint8
Fuel_Type_Diesel     uint8
Fuel_Type_LPG        uint8
Fuel_Type_Petrol     uint8
Transmission_Manual  uint8
Owner_Type_Fourth & Above uint8
Owner_Type_Second    uint8
Owner_Type_Third     uint8
Brand_BMW            uint8
Brand_Bentley        uint8
Brand_Chevrolet      uint8
Brand_Datsun         uint8
Brand_Fiat           uint8
Brand_Ford           uint8
Brand_Honda          uint8
Brand_Hyundai        uint8
Brand_ISUZU          uint8
Brand_Isuzu          uint8
Brand_Jaguar         uint8
Brand_Jeep           uint8
Brand_Land           uint8
Brand_Mahindra       uint8
Brand_Maruti         uint8
Brand_Mercedes-Benz  uint8
Brand_Mini           uint8
Brand_Mitsubishi     uint8
Brand_Nissan         uint8
Brand_Porsche        uint8
Brand_Renault        uint8
Brand_Skoda          uint8
```

```

Brand_Skoda          uint8
Brand_Tata            uint8
Brand_Toyota          uint8
Brand_Volkswagen      uint8
Brand_Volvo           uint8
dtype: object

```

```

dfn['Mileage']=dfn['Mileage'].astype(float)
dfn['Engine']=dfn['Engine'].astype(float)
dfn['Power']=dfn['Power'].astype(float)
dfn

```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Location_Bang
0	2014	40929	32.26	998.0	58.20	4.0	
1	2013	54493	24.70	796.0	47.30	5.0	
2	2017	34000	13.68	2393.0	147.80	7.0	
3	2012	139000	23.59	1364.0	0.00	5.0	
4	2014	29000	18.50	1197.0	82.85	5.0	
...	
1229	2011	89411	20.54	1598.0	103.60	5.0	
1230	2015	59000	17.21	1197.0	103.60	5.0	
1231	2012	28000	23.08	1461.0	63.10	5.0	
1232	2013	52262	17.20	1197.0	103.60	5.0	
1233	2014	72443	10.00	2148.0	170.00	5.0	

1234 rows x 49 columns

dfn.dtypes

```

Year                int64
Kilometers_Driven   int64
Mileage             float64
Engine              float64
Power               float64
Seats               float64
Location_Bangalore  uint8
Location_Chennai    uint8
Location_Coimbatore uint8
Location_Delhi       uint8
Location_Hyderabad   uint8
Location_Jaipur       uint8
Location_Kochi        uint8
Location_Kolkata      uint8
Location_Mumbai       uint8
Location_Pune         uint8
Fuel_Type_Diesel     uint8

```

```

Fuel_Type_LPG                uint8
Fuel_Type_Petrol              uint8
Transmission_Manual           uint8
Owner_Type_Fourth & Above     uint8
Owner_Type_Second             uint8
Owner_Type_Third              uint8
Brand_BMW                     uint8
Brand_Bentley                 uint8
Brand_Chevrolet               uint8
Brand_Datsun                  uint8
Brand_Fiat                    uint8
Brand_Ford                    uint8
Brand_Honda                   uint8
Brand_Hyundai                 uint8
Brand_ISUZU                   uint8
Brand_Isuzu                   uint8
Brand_Jaguar                  uint8
Brand_Jeep                    uint8
Brand_Land                    uint8
Brand_Mahindra                uint8
Brand_Maruti                  uint8
Brand_Mercedes-Benz           uint8
Brand_Mini                    uint8
Brand_Mitsubishi              uint8
Brand_Nissan                  uint8
Brand_Porsche                 uint8
Brand_Renault                 uint8
Brand_Skoda                   uint8
Brand_Tata                    uint8
Brand_Toyota                  uint8
Brand_Volkswagen              uint8
Brand_Volvo                   uint8
dtype: object

```

```

dfn.loc[dfn.Mileage==0, 'Mileage']=np.NaN
dfn.loc[dfn.Power==0, 'Power']=np.NaN
dfn.loc[dfn.Engine==0, 'Engine']=np.NaN
dfn

```

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Location_Bang
0	2014	40929	32.26	998.0	58.20	4.0	
1	2013	54493	24.70	796.0	47.30	5.0	
2	2017	34000	13.68	2393.0	147.80	7.0	
3	2012	139000	23.59	1364.0	NaN	5.0	
4	2014	29000	18.50	1197.0	82.85	5.0	
...	
1229	2011	89411	20.54	1598.0	103.60	5.0	
1230	2015	59000	17.21	1197.0	103.60	5.0	
1231	2012	28000	23.08	1461.0	63.10	5.0	

1232	2013	52262	17.20	1197.0	103.60	5.0
1233	2014	72443	10.00	2148.0	170.00	5.0

1234 rows × 49 columns

dfn.isna().sum()

Year	0
Kilometers_Driven	0
Mileage	13
Engine	10
Power	32
Seats	11
Location_Bangalore	0
Location_Chennai	0
Location_Coimbatore	0
Location_Delhi	0
Location_Hyderabad	0
Location_Jaipur	0
Location_Kochi	0
Location_Kolkata	0
Location_Mumbai	0
Location_Pune	0
Fuel_Type_Diesel	0
Fuel_Type_LPG	0
Fuel_Type_Petrol	0
Transmission_Manual	0
Owner_Type_Fourth & Above	0
Owner_Type_Second	0
Owner_Type_Third	0
Brand_BMW	0
Brand_Bentley	0
Brand_Chevrolet	0
Brand_Datsun	0
Brand_Fiat	0
Brand_Ford	0
Brand_Honda	0
Brand_Hyundai	0
Brand_ISUZU	0
Brand_Isuzu	0
Brand_Jaguar	0
Brand_Jeep	0
Brand_Land	0
Brand_Mahindra	0
Brand_Maruti	0
Brand_Mercedes-Benz	0
Brand_Mini	0
Brand_Mitsubishi	0
Brand_Nissan	0
Brand_Porsche	0
Brand_Renault	0
Brand_Skoda	0
Brand_Tata	0
Brand_Toyota	0
Brand_Volkswagen	0
Brand_Volvo	0

```
dtype: int64
```

```
for i in lst1:  
    dfn[i]=dfn[i].fillna(dfn[i].mean())
```

```
dfn['Seats']=dfn['Seats'].fillna(dfn['Seats'].mode()[0])
```

```
dfn.isna().sum()
```

Year	0
Kilometers_Driven	0
Mileage	0
Engine	0
Power	0
Seats	0
Location_Bangalore	0
Location_Chennai	0
Location_Coimbatore	0
Location_Delhi	0
Location_Hyderabad	0
Location_Jaipur	0
Location_Kochi	0
Location_Kolkata	0
Location_Mumbai	0
Location_Pune	0
Fuel_Type_Diesel	0
Fuel_Type_LPG	0
Fuel_Type_Petrol	0
Transmission_Manual	0
Owner_Type_Fourth & Above	0
Owner_Type_Second	0
Owner_Type_Third	0
Brand_BMW	0
Brand_Bentley	0
Brand_Chevrolet	0
Brand_Datsun	0
Brand_Fiat	0
Brand_Ford	0
Brand_Honda	0
Brand_Hyundai	0
Brand_ISUZU	0
Brand_Isuzu	0
Brand_Jaguar	0
Brand_Jeep	0
Brand_Land	0
Brand_Mahindra	0
Brand_Maruti	0
Brand_Mercedes-Benz	0
Brand_Mini	0
Brand_Mitsubishi	0
Brand_Nissan	0
Brand_Porsche	0
Brand_Renault	0
Brand_Skoda	0
Brand_Tata	0

```

Brand_Toyota      0
Brand_Volkswagen  0
Brand_Volvo       0
dtype: int64

```

`dfn.shape`

```
(1234, 49)
```

`dfe.shape`

```
(6019, 50)
```

`dfn`

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Location_F
0	2014	40929	32.26	998.0	58.20000	4.0	
1	2013	54493	24.70	796.0	47.30000	5.0	
2	2017	34000	13.68	2393.0	147.80000	7.0	
3	2012	139000	23.59	1364.0	110.38042	5.0	
4	2014	29000	18.50	1197.0	82.85000	5.0	
...	
1229	2011	89411	20.54	1598.0	103.60000	5.0	
1230	2015	59000	17.21	1197.0	103.60000	5.0	
1231	2012	28000	23.08	1461.0	63.10000	5.0	
1232	2013	52262	17.20	1197.0	103.60000	5.0	
1233	2014	72443	10.00	2148.0	170.00000	5.0	

1234 rows x 49 columns

`dfn.columns`

```

Index(['Year', 'Kilometers_Driven', 'Mileage', 'Engine', 'Power', 'Seats',
      'Location_Bangalore', 'Location_Chennai', 'Location_Coimbatore',
      'Location_Delhi', 'Location_Hyderabad', 'Location_Jaipur',
      'Location_Kochi', 'Location_Kolkata', 'Location_Mumbai',
      'Location_Pune', 'Fuel_Type_Diesel', 'Fuel_Type_LPG',
      'Fuel_Type_Petrol', 'Transmission_Manual', 'Owner_Type_Fourth &
      Above',
      'Owner_Type_Second', 'Owner_Type_Third', 'Brand_BMW',
      'Brand_Bentley',
      'Brand_Chevrolet', 'Brand_Datsun', 'Brand_Fiat', 'Brand_Ford',
      'Brand_Honda', 'Brand_Hyundai', 'Brand_ISUZU', 'Brand_Isuzu',
      'Brand_Jaguar', 'Brand_Jeep', 'Brand_Land', 'Brand_Mahindra',
      'Brand_Maruti', 'Brand_Mercedes-Benz', 'Brand_Mini',
      'Brand_Mitsubishi']

```

```
Brand_Mitsubishi',
      'Brand_Nissan', 'Brand_Porsche', 'Brand_Renault', 'Brand_Skoda',
      'Brand_Tata', 'Brand_Toyota', 'Brand_Volkswagen', 'Brand_Volvo'],
      dtype='object')
```

dfn.shape

```
(1234, 49)
```

dfe.columns

```
Index(['Year', 'Kilometers_Driven', 'Mileage', 'Engine', 'Power', 'Seats',
      'Price', 'Location_Bangalore', 'Location_Chennai',
      'Location_Coimbatore', 'Location_Delhi', 'Location_Hyderabad',
      'Location_Jaipur', 'Location_Kochi', 'Location_Kolkata',
      'Location_Mumbai', 'Location_Pune', 'Fuel_Type_Diesel',
      'Fuel_Type_LPG',
      'Fuel_Type_Petrol', 'Transmission_Manual', 'Owner_Type_Fourth &
      Above',
      'Owner_Type_Second', 'Owner_Type_Third', 'Brand_BMW',
      'Brand_Bentley',
      'Brand_Chevrolet', 'Brand_Datsun', 'Brand_Fiat', 'Brand_Ford',
      'Brand_Honda', 'Brand_Hyundai', 'Brand_ISUZU', 'Brand_Isuzu',
      'Brand_Jaguar', 'Brand_Jeep', 'Brand_Land', 'Brand_Mahindra',
      'Brand_Maruti', 'Brand_Mercedes-Benz', 'Brand_Mini',
      'Brand_Mitsubishi',
      'Brand_Nissan', 'Brand_Porsche', 'Brand_Renault', 'Brand_Skoda',
      'Brand_Tata', 'Brand_Toyota', 'Brand_Volkswagen', 'Brand_Volvo'],
      dtype='object')
```

dfe.shape

```
(6019, 50)
```

dfe

	Year	Kilometers_Driven	Mileage	Engine	Power	Seats	Price	Locati
0	2010	72000	26.60	998.0	58.16	5.0	1.75	
1	2015	41000	19.67	1582.0	126.20	5.0	12.50	
2	2011	46000	18.20	1199.0	88.70	5.0	4.50	
3	2012	87000	20.77	1248.0	88.76	7.0	6.00	
4	2013	40670	15.20	1968.0	140.80	5.0	17.74	
...	
6014	2014	27365	28.40	1248.0	74.00	5.0	4.75	
6015	2015	100000	24.40	1120.0	71.00	5.0	4.00	
6016	2012	55000	14.00	2498.0	112.00	8.0	2.90	
6017	2013	46000	18.90	998.0	67.10	5.0	2.65	
6018	2011	47000	25.44	998.0	57.60	5.0	2.50	

```
6018 2011 47000 25.44 936.0 57.60 5.0 2.50
```

```
6019 rows x 50 columns
```

dfe.columns

```
Index(['Year', 'Kilometers_Driven', 'Mileage', 'Engine', 'Power', 'Seats',
      'Price', 'Location_Bangalore', 'Location_Chennai',
      'Location_Coimbatore', 'Location_Delhi', 'Location_Hyderabad',
      'Location_Jaipur', 'Location_Kochi', 'Location_Kolkata',
      'Location_Mumbai', 'Location_Pune', 'Fuel_Type_Diesel',
      'Fuel_Type_LPG',
      'Fuel_Type_Petrol', 'Transmission_Manual', 'Owner_Type_Fourth &
      Above',
      'Owner_Type_Second', 'Owner_Type_Third', 'Brand_BMW',
      'Brand_Bentley',
      'Brand_Chevrolet', 'Brand_Datsun', 'Brand_Fiat', 'Brand_Ford',
      'Brand_Honda', 'Brand_Hyundai', 'Brand_ISUZU', 'Brand_Isuzu',
      'Brand_Jaguar', 'Brand_Jeep', 'Brand_Land', 'Brand_Mahindra',
      'Brand_Maruti', 'Brand_Mercedes-Benz', 'Brand_Mini',
      'Brand_Mitsubishi',
      'Brand_Nissan', 'Brand_Porsche', 'Brand_Renault', 'Brand_Skoda',
      'Brand_Tata', 'Brand_Toyota', 'Brand_Volkswagen', 'Brand_Volvo'],
      dtype='object')
```

dfn.columns

```
Index(['Year', 'Kilometers_Driven', 'Mileage', 'Engine', 'Power', 'Seats',
      'Location_Bangalore', 'Location_Chennai', 'Location_Coimbatore',
      'Location_Delhi', 'Location_Hyderabad', 'Location_Jaipur',
      'Location_Kochi', 'Location_Kolkata', 'Location_Mumbai',
      'Location_Pune', 'Fuel_Type_Diesel', 'Fuel_Type_LPG',
      'Fuel_Type_Petrol', 'Transmission_Manual', 'Owner_Type_Fourth &
      Above',
      'Owner_Type_Second', 'Owner_Type_Third', 'Brand_BMW',
      'Brand_Bentley',
      'Brand_Chevrolet', 'Brand_Datsun', 'Brand_Fiat', 'Brand_Ford',
      'Brand_Honda', 'Brand_Hyundai', 'Brand_ISUZU', 'Brand_Isuzu',
      'Brand_Jaguar', 'Brand_Jeep', 'Brand_Land', 'Brand_Mahindra',
      'Brand_Maruti', 'Brand_Mercedes-Benz', 'Brand_Mini',
      'Brand_Mitsubishi',
      'Brand_Nissan', 'Brand_Porsche', 'Brand_Renault', 'Brand_Skoda',
      'Brand_Tata', 'Brand_Toyota', 'Brand_Volkswagen', 'Brand_Volvo'],
      dtype='object')
```

MODEL CREATION

```
from sklearn.linear_model import LinearRegression
model=LinearRegression()
model.fit(x,y)
v pred=model.predict(dfn)
```

```

y_pred
array([ 2.4535847,  1.26972859, 16.72302279, ..., -0.03425225,
        5.77243299, 24.73047308])

print('slope is',list(zip(dfn,model.coef_)))
slope is [('Year', 1.0271549352849454), ('Kilometers_Driven', 1.55146875074

print('constant is',model.intercept_)
constant is -2061.9735471473764

```