


```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt      #car price prediction
df=pd.read_csv('/content/train-data.csv')
df
```



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

6019 rows × 14 columns




Next steps:

Generate code with df

 View recommended plots


```
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

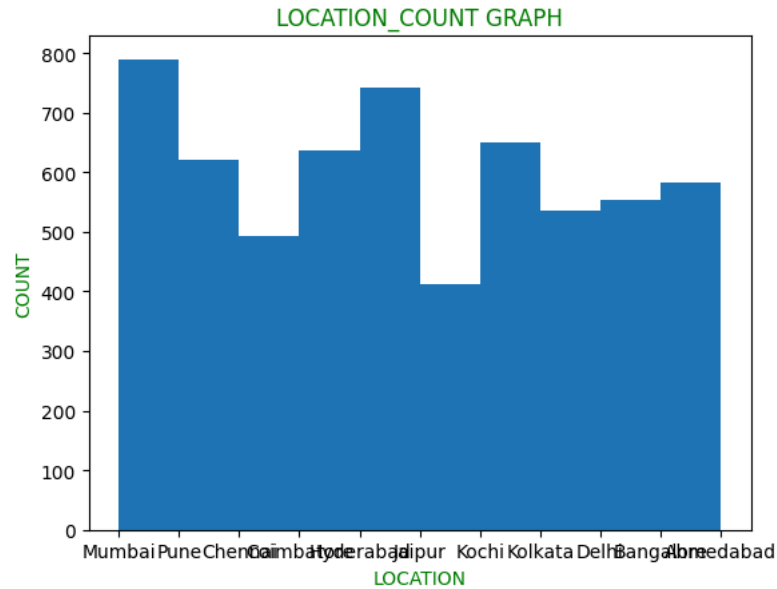
```
Seats          float64
New_Price      object
Price          float64
dtype: object
```

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

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

```
#graph
plt.hist(df['Location'])
plt.xlabel("LOCATION",color="g")
plt.ylabel("COUNT",color='g')
plt.title("LOCATION_COUNT GRAPH",color="g")
```

```
Text(0.5, 1.0, 'LOCATION_COUNT GRAPH')
```

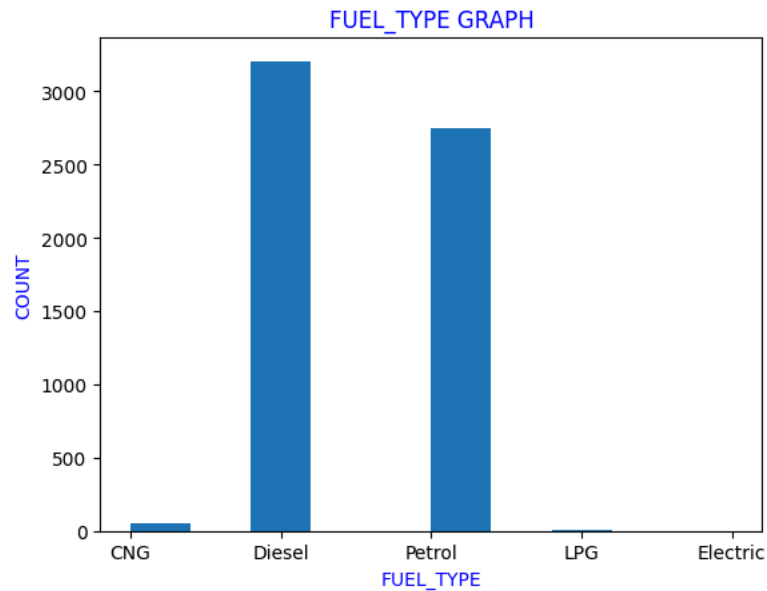


```
df['Fuel_Type'].value_counts()
```

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

```
plt.hist(df['Fuel_Type'])
plt.xlabel("FUEL_TYPE",color="b")
plt.ylabel("COUNT",color="b")
plt.title("FUEL_TYPE GRAPH",color="b")
```

```
Text(0.5, 1.0, 'FUEL_TYPE GRAPH')
```

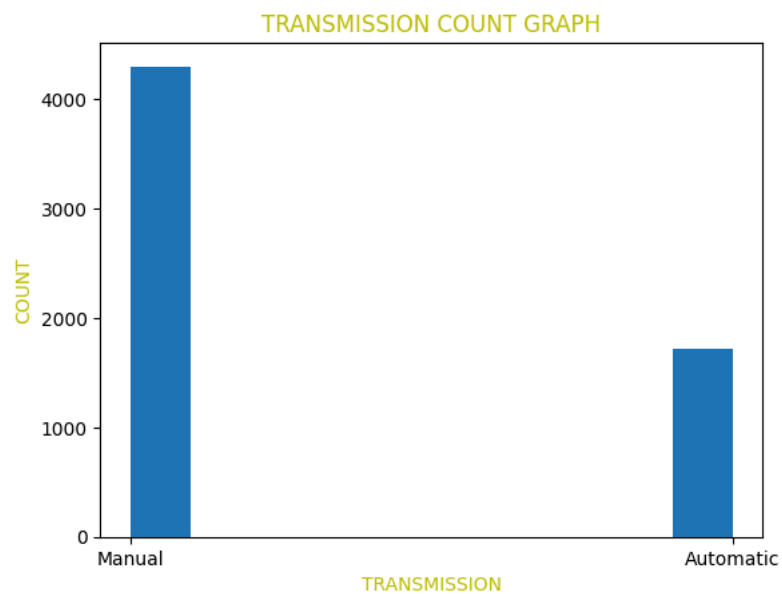


```
df['Transmission'].value_counts()
```

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

```
plt.hist(df['Transmission'])
plt.xlabel("TRANSMISSION",color="y")
plt.ylabel("COUNT",color="y")
plt.title("TRANSMISSION COUNT GRAPH",color="y")
```

```
Text(0.5, 1.0, 'TRANSMISSION COUNT GRAPH')
```



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

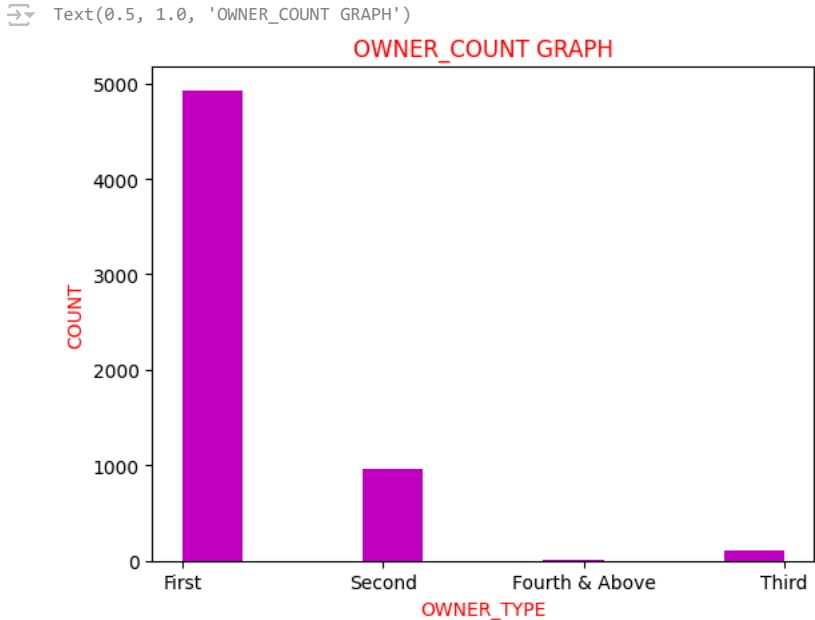
```
Owner_Type
First      4929
Second     968
Third      113
Fourth & Above 9
Name: count, dtype: int64
```

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

```
Name
Mahindra XUV500 W8 2WD      49
Maruti Swift VDI            45
Honda City 1.5 S MT         34
Maruti Swift Dzire VDI      34
Maruti Swift VDI BSIV       31
```

```
..
Ford Fiesta Titanium 1.5 TDCi      1
Mahindra Scorpio S10 AT 4WD       1
Hyundai i20 1.2 Era                1
Toyota Camry W4 (AT)              1
Mahindra Xylo D4 BSIV             1
Name: count, Length: 1878, dtype: int64
```

```
plt.hist(df['Owner_Type'],color='m')
plt.xlabel("OWNER_TYPE",color="red")
plt.ylabel("COUNT",color="red")
plt.title("OWNER_COUNT GRAPH",color="red")
```



```
df['Fuel_Type'].unique() #to find unique values

array(['CNG', 'Diesel', 'Petrol', 'LPG', 'Electric'], dtype=object)

#encoding

#new encoding technique get dummies

# get dummies is not a encoding in machine learning

# it is a pandas encoding technique

df1=pd.get_dummies(df[['Location','Transmission','Owner_Type']],drop_first=True)

df1
```

	Location_Bangalore	Location_Chennai	Location_Coimbatore	Location_Delhi	Loc.
0	False	False	False	False	
1	False	False	False	False	
2	False	True	False	False	
3	False	True	False	False	
4	False	False	True	False	
...	
6014	False	False	False	True	
6015	False	False	False	False	
6016	False	False	False	False	
6017	False	False	False	False	
6018	False	False	False	False	

6019 rows × 14 columns

Next steps:

[Generate code with df1](#)



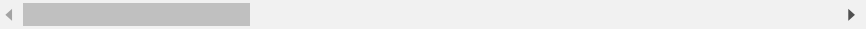
[View recommended plots](#)

```
dfe=pd.concat([df,df1],axis=1)
dfe
```



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

6019 rows × 28 columns



```
#tuesday'
#drop unwanted file.
dfe.drop(['Unnamed: 0', 'Name', 'Location', 'Transmission', 'Owner_Type', 'New_Price'],axis=1,inplace=True)
dfe
```



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

6019 rows × 22 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	Fuel_Type	Mileage	Engine	Power	Seats	Price	Locat
0	2010	72000	CNG	26.6	998	58.16	5.0	1.75	
1	2015	41000	Diesel	19.67	1582	126.2	5.0	12.50	
2	2011	46000	Petrol	18.2	1199	88.7	5.0	4.50	
3	2012	87000	Diesel	20.77	1248	88.76	7.0	6.00	
4	2013	40670	Diesel	15.2	1968	140.8	5.0	17.74	
...
6014	2014	27365	Diesel	28.4	1248	74	5.0	4.75	
6015	2015	100000	Diesel	24.4	1120	71	5.0	4.00	
6016	2012	55000	Diesel	14.0	2498	112	8.0	2.90	
6017	2013	46000	Petrol	18.9	998	67.1	5.0	2.65	
6018	2011	47000	Diesel	25.44	936	57.6	5.0	2.50	

6019 rows × 22 columns

dfe.dtypes



```
Year                int64
Kilometers_Driven   int64
Fuel_Type           object
Mileage             object
Engine             object
Power              object
Seats              float64
Price              float64
Location_Bangalore  bool
Location_Chennai    bool
Location_Coimbatore bool
Location_Delhi      bool
Location_Hyderabad  bool
```

```

Location_Jaipur          bool
Location_Kochi           bool
Location_Kolkata         bool
Location_Mumbai          bool
Location_Pune            bool
Transmission_Manual      bool
Owner_Type_Fourth & Above bool
Owner_Type_Second        bool
Owner_Type_Third         bool
dtype: object

```

```
#converting datatypes      #datatype conversion
```

```

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

```

```
dfe.dtypes
```

```

Year          int64
Kilometers_Driven int64
Fuel_Type      object
Mileage        float64
Engine         float64
Power          float64
Seats          float64
Price          float64
Location_Bangalore bool
Location_Chennai bool
Location_Coimbatore bool
Location_Delhi bool
Location_Hyderabad bool
Location_Jaipur bool
Location_Kochi bool
Location_Kolkata bool
Location_Mumbai bool
Location_Pune bool
Transmission_Manual bool
Owner_Type_Fourth & Above bool
Owner_Type_Second bool
Owner_Type_Third bool
dtype: object

```

```

#Replace a numerical value that is 0
# o t o NAN

```

```

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

```

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

```

Year          0
Kilometers_Driven 0
Fuel_Type      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
Transmission_Manual 0
Owner_Type_Fourth & Above 0
Owner_Type_Second 0
Owner_Type_Third 0
dtype: int64

```

```
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
```

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

```
Year      0
Kilometers_Driven  0
Fuel_Type  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
Transmission_Manual  0
Owner_Type_Fourth & Above  0
Owner_Type_Second  0
Owner_Type_Third  0
dtype: int64
```

```
from sklearn.preprocessing import LabelEncoder
end=LabelEncoder()
dfe['Fuel_Type'] =end.fit_transform(dfe['Fuel_Type'])
```

```
# seperate x and y
```

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

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

6019 rows × 21 columns

```
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
```

performing same operation in testing data

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



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


1234 rows × 13 columns

Next steps:

Generate code with data

View recommended plots

data.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

data.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

```

```
data['Location'].value_counts()
```

```

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

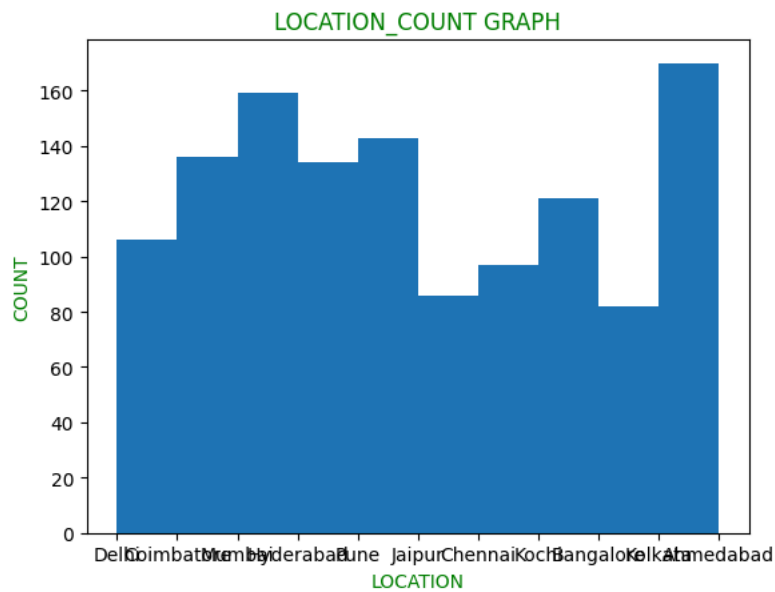
```

```

#graph
plt.hist(data['Location'])
plt.xlabel("LOCATION",color="g")
plt.ylabel("COUNT",color='g')
plt.title("LOCATION_COUNT GRAPH",color="g")

```

```
↳ Text(0.5, 1.0, 'LOCATION_COUNT GRAPH')
```



```
data['Fuel_Type'].value_counts()
```

```

↳ Fuel_Type
Diesel      647
Petrol      579
CNG         6
LPG         2
Name: count, dtype: int64

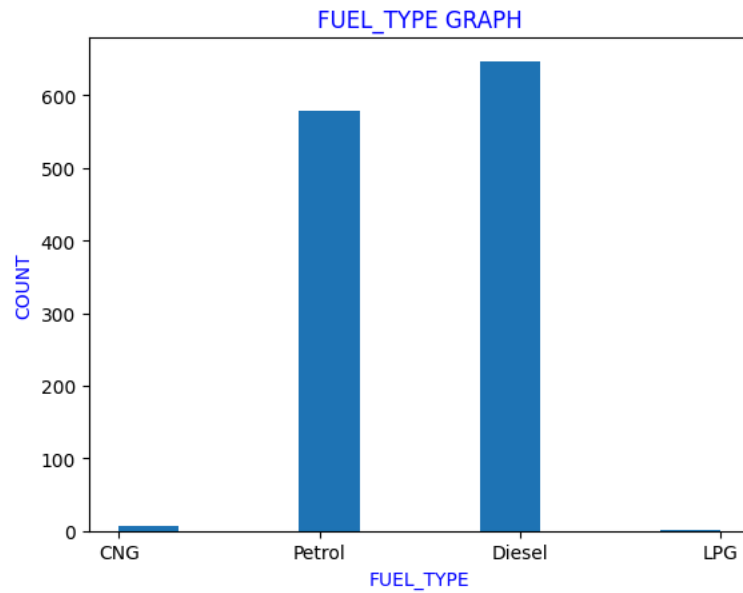
```

```

plt.hist(data['Fuel_Type'])
plt.xlabel("FUEL_TYPE",color="b")
plt.ylabel("COUNT",color="b")
plt.title("FUEL_TYPE GRAPH",color="b")

```

```
Text(0.5, 1.0, 'FUEL_TYPE GRAPH')
```

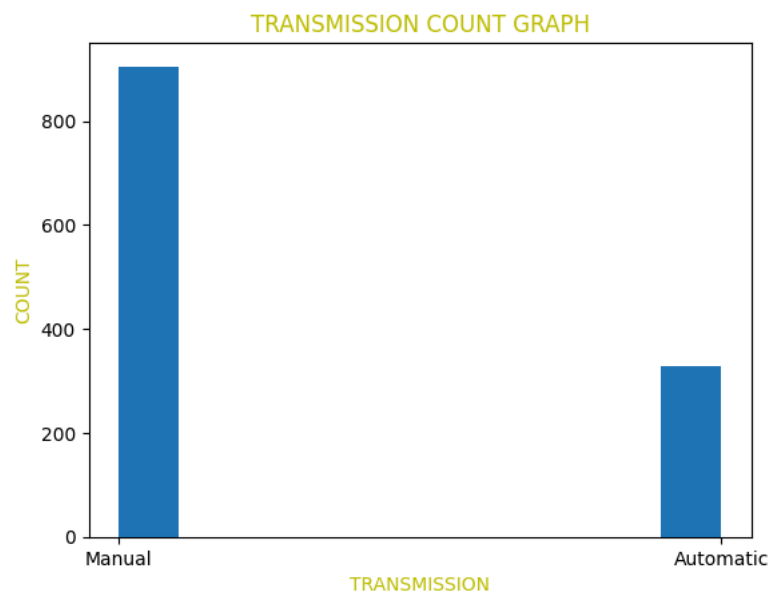


```
data['Transmission'].value_counts()
```

```
Transmission
Manual      905
Automatic   329
Name: count, dtype: int64
```

```
plt.hist(data['Transmission'])
plt.xlabel("TRANSMISSION",color="y")
plt.ylabel("COUNT",color="y")
plt.title("TRANSMISSION COUNT GRAPH",color="y")
```

```
Text(0.5, 1.0, 'TRANSMISSION COUNT GRAPH')
```

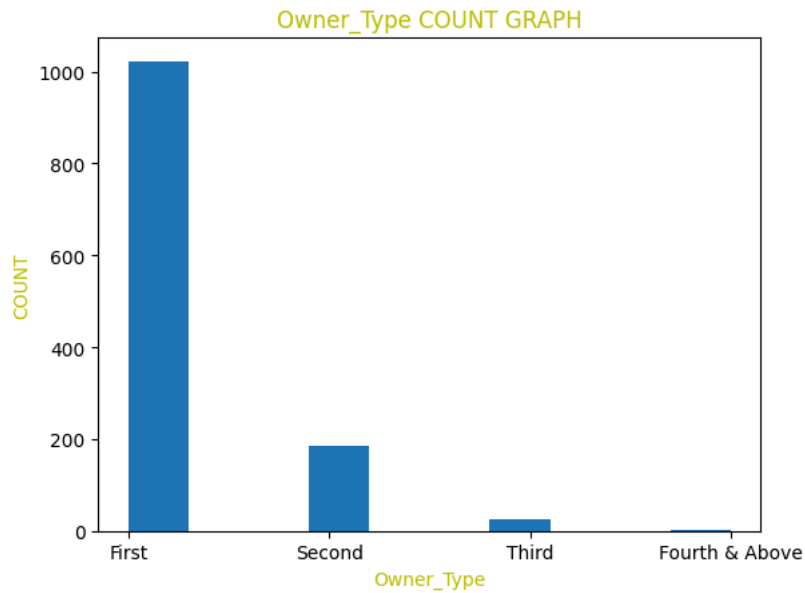


```
data['Owner_Type'].value_counts()
```

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

```
plt.hist(data['Owner_Type'])
plt.xlabel("Owner_Type",color="y")
plt.ylabel("COUNT",color="y")
plt.title("Owner_Type COUNT GRAPH",color="y")
```

```
Text(0.5, 1.0, 'Owner_Type COUNT GRAPH')
```



```
data['Name'].value_counts()
```

```
Name
Maruti Alto LXi                      9
Honda City 1.5 V MT                  8
Maruti Swift Dzire VDI               8
Volkswagen Polo 1.2 MPI Highline     8
Hyundai i10 Magna                    7
..
Hyundai Santro GLS I - Euro II       1
Honda City i DTec VX Option BL       1
Land Rover Discovery 4 SDV6 SE        1
Hyundai Verna CRDi 1.6 SX Option     1
Mercedes-Benz E-Class 2009-2013 E 220 CDI Avantgarde 1
Name: count, Length: 769, dtype: int64
```

```
#encoding
```

```
data1=pd.get_dummies(data[['Location','Transmission','Owner_Type']],drop_first=True)
```

```
data1
```

```
Location_Bangalore Location_Chennai Location_Coimbatore Location_Delhi Loc.
0 False False False True
1 False False True False
2 False False False False
3 False False False False
4 False False False False
...
1229 False False False False
1230 False False False False
1231 False False False False
1232 False False False False
1233 False False False False
1234 rows x 14 columns
```

Next steps:

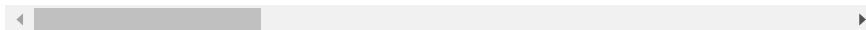
[Generate code with data1](#)[View recommended plots](#)

```
#concat
datacon=pd.concat([data,data1],axis=1)
datacon
```



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

1234 rows × 27 columns

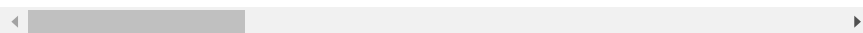


```
datacon.drop(['Unnamed: 0', 'Name', 'Location', 'Transmission', 'Owner_Type', 'New_Price'], axis=1, inplace=True)
datacon
```



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

1234 rows × 21 columns



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



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

1234 rows × 21 columns

datacon.dtypes



```
Year          int64
Kilometers_Driven  int64
Fuel_Type      object
Mileage        object
Engine         object
Power          object
Seats          float64
Location_Bangalore  bool
Location_Chennai   bool
Location_Coimbatore bool
Location_Delhi     bool
Location_Hyderabad bool
Location_Jaipur    bool
```

```

Location_Kochi          bool
Location_Kolkata        bool
Location_Mumbai         bool
Location_Pune           bool
Transmission_Manual     bool
Owner_Type_Fourth & Above bool
Owner_Type_Second       bool
Owner_Type_Third        bool
dtype: object

```

```
#converting datatypes      #datatype conversion
```

```

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

```

```
datacon.dtypes
```

```

Year          int64
Kilometers_Driven int64
Fuel_Type     object
Mileage       float64
Engine        float64
Power         float64
Seats         float64
Location_Bangalore bool
Location_Chennai bool
Location_Coimbatore bool
Location_Delhi bool
Location_Hyderabad bool
Location_Jaipur bool
Location_Kochi bool
Location_Kolkata bool
Location_Mumbai bool
Location_Pune bool
Transmission_Manual bool
Owner_Type_Fourth & Above bool
Owner_Type_Second bool
Owner_Type_Third bool
dtype: object

```

```

#Replace a numerical value that is 0
# o t o NAN

```

```

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

```

```
datacon.isna().sum()
```

```

Year          0
Kilometers_Driven 0
Fuel_Type     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
Transmission_Manual 0
Owner_Type_Fourth & Above 0
Owner_Type_Second 0
Owner_Type_Third 0
dtype: int64

```

```
data.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
```

#fillng missing values

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

datacon.isna().sum()

```
Year      0
Kilometers_Driven  0
Fuel_Type  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
Transmission_Manual  0
Owner_Type_Fourth & Above  0
Owner_Type_Second  0
Owner_Type_Third  0
dtype: int64
```

datacon

	Year	Kilometers_Driven	Fuel_Type	Mileage	Engine	Power	Seats	Location_Bangalore	Location_Chennai	Location_Coimbatore
0	2014	40929	CNG	32.26	998.0	58.20000	4.0	False	False	False