

```
In [1]: import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

In [3]: car=pd.read_csv("C:/Users/77son/Downloads/car.csv")

In [4]: car.head(10)

Out[4]:
   Unnamed: 0  Name      Location  Year  Kilometers_Driven  Fuel_Type  Transmission  Owner_Type  Mileage  Engine  Power  Seats  New_Price  Price
0            0      Maruti Wagon R LXI CNG      Mumbai  2010.0           72000.0      CNG      Manual      First   26.6 km/kg   998 CC   58.16 bhp   5.0      NaN   1.75
1            1  Hyundai Creta 1.6 CRDI SX Option      Pune  2015.0           41000.0     Diesel      Manual      First  19.67 kmpl  1582 CC  126.2 bhp   5.0      NaN  12.50
2            2      Honda Jazz V      Chennai  2011.0           46000.0     Petrol      Manual      First  18.2 kmpl  1199 CC   88.7 bhp   5.0    8.61 Lakh   4.50
3            3      Maruti Ertiga VDI      Chennai  2012.0           87000.0     Diesel      Manual      First  20.77 kmpl  1248 CC   88.76 bhp   7.0      NaN   6.00
4            4  Audi A4 New 2.0 TDI Multitronic  Coimbatore  2013.0           40670.0     Diesel      Automatic  Second  15.2 kmpl  1968 CC  140.8 bhp   5.0      NaN  17.74
5            5  Hyundai EON LPG Era Plus Option  Hyderabad  2012.0           75000.0      LPG      Manual      First  21.1 km/kg   814 CC   55.2 bhp   5.0      NaN   2.35
6            6      Nissan Micra Diesel XV      Jaipur  2013.0           86999.0     Diesel      Manual      First  23.08 kmpl  1461 CC   63.1 bhp   5.0      NaN   3.50
7            7  Toyota Innova Crysta 2.8 GX AT 8S      Mumbai  2016.0           36000.0     Diesel      Automatic  First  11.36 kmpl  2755 CC  171.5 bhp   8.0    21 Lakh  17.50
8            8  Volkswagen Vento Diesel Comfortline  Pune  2013.0           64430.0     Diesel      Manual      First  20.54 kmpl  1598 CC  103.6 bhp   5.0      NaN   5.20
9            9      Tata Indica Vista Quadrajet LS      Chennai  2012.0           65932.0     Diesel      Manual      Second  22.3 kmpl  1248 CC   74 bhp   5.0      NaN   1.95

In [5]: car.columns

Out[5]:
Index(['Unnamed: 0', 'Name', 'Location', 'Year', 'Kilometers_Driven',
      'Fuel_Type', 'Transmission', 'Owner_Type', 'Mileage', 'Engine', 'Power',
      'Seats', 'New_Price', 'Price'],
      dtype='object')

In [6]: car.isnull().sum()

Out[6]:
Unnamed: 0      0
Name            6
Location        7
Year            7
Kilometers_Driven 7
Fuel_Type       7
Transmission    7
Owner_Type      7
Mileage         9
Engine         43
Power          43
Seats          49
New_Price      5202
Price          7
dtype: int64

In [7]: car.shape

Out[7]:
(6026, 14)

In [10]: car.head()

Out[10]:
   Unnamed: 0  Name      Location  Year  Kilometers_Driven  Fuel_Type  Transmission  Owner_Type  Mileage  Engine  Power  Seats  New_Price  Price
0            0      Maruti Wagon R LXI CNG      Mumbai  2010.0           72000.0      CNG      Manual      First  26.6 km/kg   998 CC   58.16 bhp   5.0      NaN   1.75
1            1  Hyundai Creta 1.6 CRDI SX Option      Pune  2015.0           41000.0     Diesel      Manual      First  19.67 kmpl  1582 CC  126.2 bhp   5.0      NaN  12.50
2            2      Honda Jazz V      Chennai  2011.0           46000.0     Petrol      Manual      First  18.2 kmpl  1199 CC   88.7 bhp   5.0    8.61 Lakh   4.50
3            3      Maruti Ertiga VDI      Chennai  2012.0           87000.0     Diesel      Manual      First  20.77 kmpl  1248 CC   88.76 bhp   7.0      NaN   6.00
4            4  Audi A4 New 2.0 TDI Multitronic  Coimbatore  2013.0           40670.0     Diesel      Automatic  Second  15.2 kmpl  1968 CC  140.8 bhp   5.0      NaN  17.74

In [11]: car.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6026 entries, 0 to 6025
Data columns (total 14 columns):
#   Column                Non-Null Count  Dtype
---  ---
0   Unnamed: 0            6026 non-null  object
1   Name                  6020 non-null  object
2   Location              6019 non-null  object
3   Year                  6019 non-null  float64
4   Kilometers_Driven     6019 non-null  float64
5   Fuel_Type             6019 non-null  object
6   Transmission          6019 non-null  object
7   Owner_Type           6019 non-null  object
8   Mileage               6017 non-null  object
9   Engine               5983 non-null  object
10  Power                5983 non-null  object
11  Seats                5977 non-null  float64
12  New_Price            824 non-null   object
13  Price                6019 non-null  float64
dtypes: float64(4), object(10)
memory usage: 659.2+ KB

In [12]: car.describe

Out[12]:
<bound method NDFrame.describe of      Unnamed: 0      Name      Location      Year \
0            0      Maruti Wagon R LXI CNG      Mumbai  2010.0
1            1  Hyundai Creta 1.6 CRDI SX Option      Pune  2015.0
2            2      Honda Jazz V      Chennai  2011.0
3            3      Maruti Ertiga VDI      Chennai  2012.0
4            4  Audi A4 New 2.0 TDI Multitronic  Coimbatore  2013.0
...      ...
6021  Footer navigation      NaN      NaN      NaN
6022  Terms              NaN      NaN      NaN
6023  Privacy            NaN      NaN      NaN
6024  Security           NaN      NaN      NaN
6025  Status             NaN      NaN      NaN

      Kilometers_Driven  Fuel_Type  Transmission  Owner_Type      Mileage \
0           72000.0      CNG      Manual      First   26.6 km/kg
1           41000.0     Diesel      Manual      First  19.67 kmpl
2           46000.0     Petrol      Manual      First  18.2 kmpl
3           87000.0     Diesel      Manual      First  20.77 kmpl
4           40670.0     Diesel      Automatic  Second  15.2 kmpl
...      ...
6021           NaN           NaN           NaN           NaN           NaN
6022           NaN           NaN           NaN           NaN           NaN
6023           NaN           NaN           NaN           NaN           NaN
6024           NaN           NaN           NaN           NaN           NaN
6025           NaN           NaN           NaN           NaN           NaN

      Engine      Power  Seats  New_Price  Price
0      998 CC   58.16 bhp   5.0      NaN   1.75
1     1582 CC  126.2 bhp   5.0      NaN  12.50
2     1199 CC   88.7 bhp   5.0    8.61 Lakh   4.50
3     1248 CC   88.76 bhp   7.0      NaN   6.00
4     1968 CC  140.8 bhp   5.0      NaN  17.74
...      ...
6021           NaN           NaN           NaN           NaN           NaN
6022           NaN           NaN           NaN           NaN           NaN
6023           NaN           NaN           NaN           NaN           NaN
6024           NaN           NaN           NaN           NaN           NaN
6025           NaN           NaN           NaN           NaN           NaN

[6026 rows x 14 columns]>

In [14]: car.Kilometers_Driven.unique()

Out[14]:
array([72000., 41000., 46000., ..., 70602., 27365., nan])

In [15]: car.Fuel_Type.value_counts()

Out[15]:
Diesel    3205
Petrol    2746
CNG        56
LPG        10
Electric    2
Name: Fuel_Type, dtype: int64

In [16]: car.Transmission.value_counts()

Out[16]:
Manual    4299
Automatic 1720
Name: Transmission, dtype: int64

In [17]: car.Mileage.value_counts()

Out[17]:
18.9 kmpl    172
17.0 kmpl    172
18.6 kmpl    119
20.36 kmpl    88
21.1 kmpl     86
...
27.28 kmpl     1
14.57 kmpl     1
22.8 km/kg     1
8.0 kmpl       1
17.24 kmpl     1
Name: Mileage, Length: 442, dtype: int64

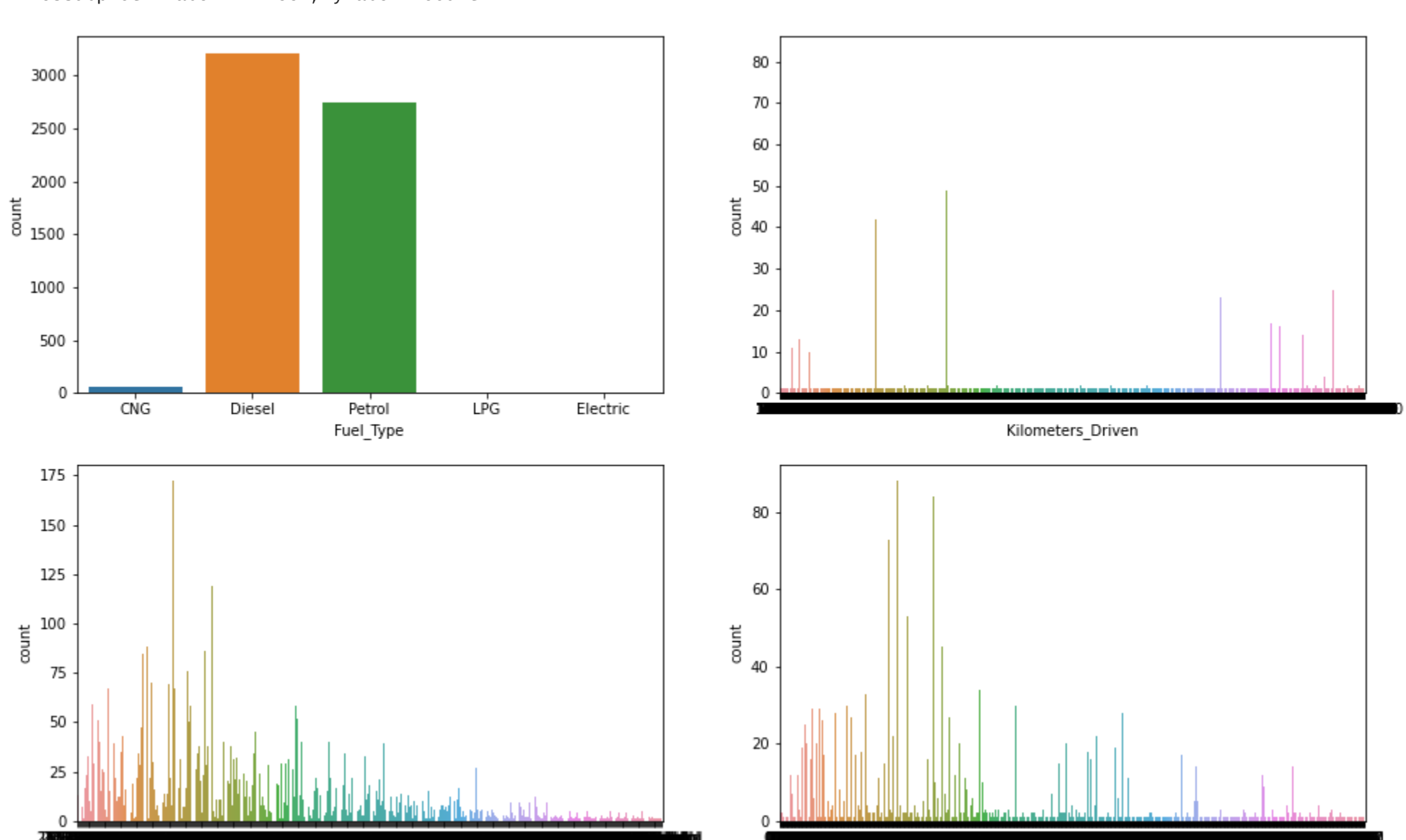
In [18]: car.Engine.value_counts()

Out[18]:
1197 CC     606
1248 CC     512
1498 CC     304
998 CC      259
2179 CC     240
...
2999 CC      1
2147 CC      1
2495 CC      1
3200 CC      1
1797 CC      1
Name: Engine, Length: 146, dtype: int64

In [22]: car.Fuel_Type.replace(regex=("gas""0","diesel""1"),inplace=True)
car.Engine.replace(regex=("gas""0","diesel""1"),inplace=True)

In [23]: #countplot
fig,ax = plt.subplots(2,2,figsize=(16,10))
sns.countplot(x='Fuel_Type',data=car, ax=ax[0][0])
sns.countplot(x='Kilometers_Driven',data=car, ax=ax[0][1])
sns.countplot(x='Mileage',data=car, ax=ax[1][0])
sns.countplot(x='Price',data=car, ax=ax[1][1])

Out[23]:
<AxesSubplot:xlabel='Price', ylabel='count'>



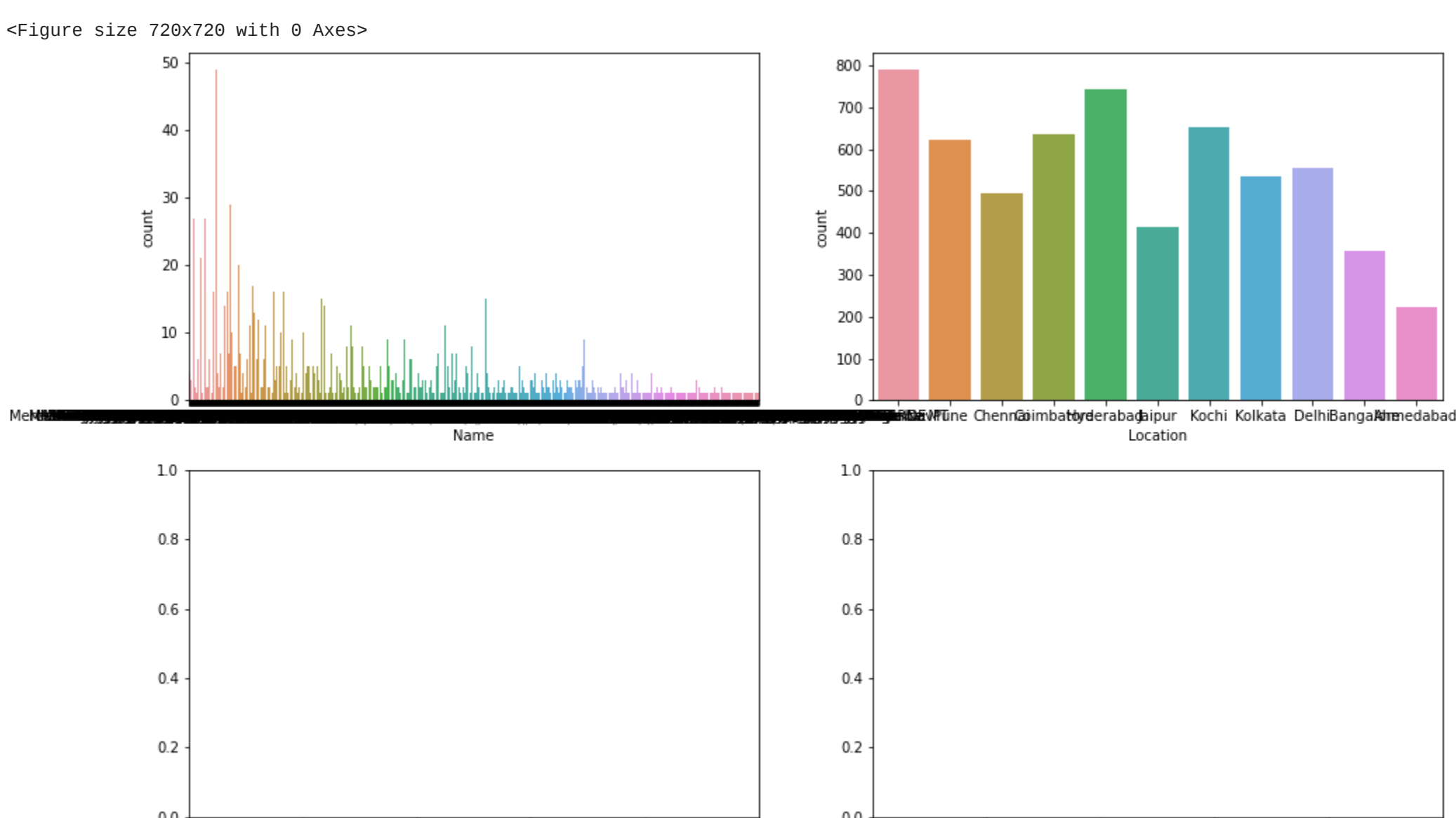
In [24]: car.corr()

Out[24]:
      Year  Kilometers_Driven  Seats  Price
Year  1.000000         -0.173048  0.012333  0.305327
Kilometers_Driven -0.173048         1.000000  0.083113 -0.011493
Seats    0.012333         0.083113  1.000000  0.052225
Price    0.305327        -0.011493  0.052225  1.000000

In [29]: #this is a Visual representation of our correlation
plt.figure(figsize = (10,10))
correlation_matrix = car.corr()
#countplot
fig,ax = plt.subplots(2,2,figsize=(16,10))
sns.countplot(x='Name',data=car, ax=ax[0][0])
sns.countplot(x='Location',data=car, ax=ax[0][1])
plt.show

Out[29]:
<function matplotlib.pyplot.show(close=None, block=None)>

<Figure size 720x720 with 0 Axes>


```