NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

DATA VISUALIZATION USING PYTHON (CA-2 PROJECT)

DATASET- CAR DETAILS

PLOT USING SEABORN LIBRARY:

1.Box Plot

2.Line Plot

3.KDE Plot

4.Pie Plot

5.Strip Plot

6.Count Plot

7.Joint Plot

8.Bar Plot

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

```
import pandas as pd
import seaborn as sns
import io
df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
print(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1049 entries, 0 to 1048
Data columns (total 11 columns):
 #   Column         Non-Null Count  Dtype  
---  ------         --------------  -----  
 0   name           1049 non-null   object 
 1   year           1049 non-null   int64  
 2   selling_price  1049 non-null   int64  
 3   km_driven      1049 non-null   int64  
 4   fuel           1049 non-null   object 
 5   seller_type    1049 non-null   object 
 6   transmission   1049 non-null   object 
 7   owner          1049 non-null   object 
 8   mileage        1018 non-null   object 
 9   engine         1018 non-null   object 
 10  seats          1018 non-null   float64
dtypes: float64(1), int64(3), object(7)
memory usage: 90.3+ KB
None
```

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

```
[ ]  import pandas as pd
     import seaborn as sns
     import io
     df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
     df.isnull().sum()
```

```
name              0
year              0
selling_price     0
km_driven         0
fuel              0
seller_type       0
transmission      0
owner             0
mileage          31
engine           31
seats            31
dtype: int64
```

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

```
import pandas as pd
import seaborn as sns
import io
df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
print(df)
```

```
                                     name  year  selling_price  km_driven  \
0                   Maruti Swift Dzire VDI  2014         450000     145500
1               Skoda Rapid 1.5 TDI Ambition  2014         370000     120000
2                   Honda City 2017-2020 EXi  2006         158000     140000
3                  Hyundai i20 Sportz Diesel  2010         225000     127000
4                      Maruti Swift VXI BSIII  2007         130000     120000
...                                       ...   ...            ...        ...
1044  Toyota Innova 2.5 E Diesel MS 7-seater  2007         626000      80000
1045                  Toyota Qualis Fleet A3  2000         200000     100000
1046              Maruti Swift Dzire AMT VXI  2019         600000      10000
1047               Renault KWID RXT Optional  2019         290000      21000
1048           Maruti Alto K10 2010-2014 VXI  2013         215000      25000

         fuel seller_type transmission         owner     mileage   engine  \
0     Diesel  Individual       Manual   First Owner   23.4 kmpl  1248 CC
1     Diesel  Individual       Manual  Second Owner  21.14 kmpl  1498 CC
2     Petrol  Individual       Manual   Third Owner   17.7 kmpl  1497 CC
3     Diesel  Individual       Manual   First Owner   23.0 kmpl  1396 CC
4     Petrol  Individual       Manual   First Owner   16.1 kmpl  1298 CC
...      ...         ...          ...           ...          ...      ...
1044  Diesel  Individual       Manual  Second Owner   12.8 kmpl  2494 CC
1045  Diesel  Individual       Manual   First Owner         NaN      NaN
1046  Petrol  Individual    Automatic   First Owner  21.21 kmpl  1197 CC
1047  Petrol  Individual       Manual   First Owner  25.17 kmpl   799 CC
1048  Petrol  Individual       Manual   First Owner  20.92 kmpl   998 CC

      seats
0       5.0
1       5.0
2       5.0
3       5.0
4       5.0
...     ...
1044    7.0
1045    NaN
1046    5.0
1047    5.0
1048    5.0

[1049 rows x 11 columns]
```

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

```
[ ]  import pandas as pd
     import seaborn as sns
     import io
     df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
     df.describe()
```

	year	selling_price	km_driven	seats
count	1049.000000	1.049000e+03	1049.000000	1018.000000
mean	2013.697807	6.236611e+05	69625.546235	5.443026
std	3.979576	8.302837e+05	45707.191884	0.964052
min	1983.000000	4.000000e+04	1000.000000	4.000000
25%	2011.000000	2.600000e+05	35000.000000	5.000000
50%	2014.000000	4.150000e+05	60000.000000	5.000000
75%	2017.000000	6.500000e+05	100000.000000	5.000000
max	2020.000000	1.000000e+07	360003.000000	10.000000

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

```
[ ]  import pandas as pd
     import seaborn as sns
     import io
     df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
     print(df.shape)
```

```
(1049, 11)
```

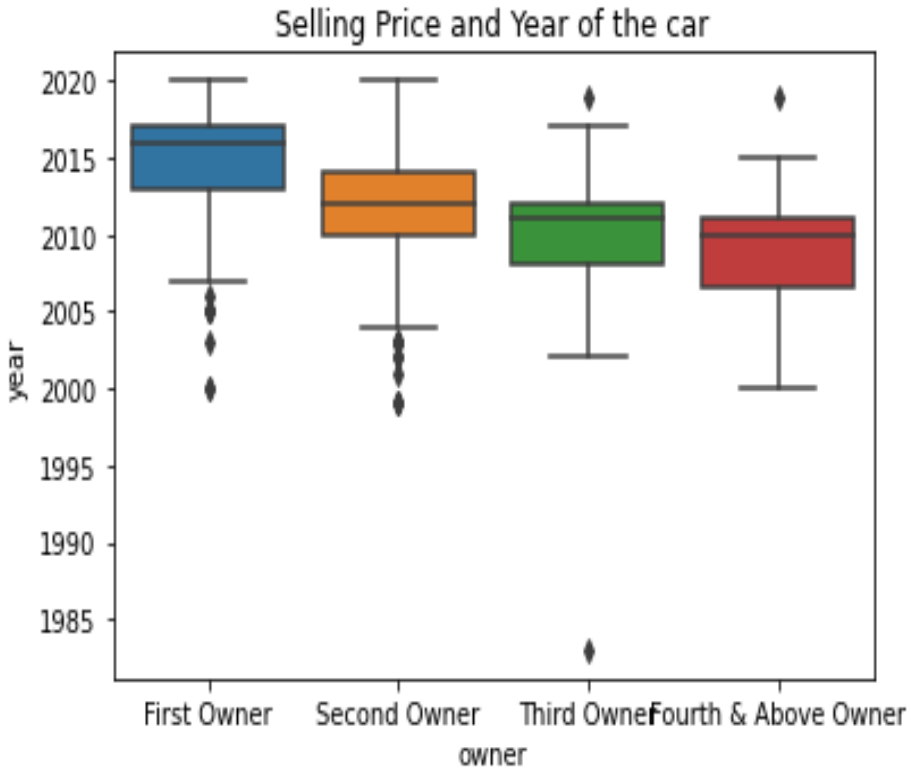
```
[ ]  import pandas as pd
     import seaborn as sns
     import io
     df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
     print(df.columns)
```

```
Index(['name', 'year', 'selling_price', 'km_driven', 'fuel', 'seller_type',
       'transmission', 'owner', 'mileage', 'engine', 'seats'],
      dtype='object')
```

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

BOX PLOT

```
import pandas as pd
import seaborn as sns
import io
import matplotlib.pyplot as plt
df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
plt.title('Selling Price and Year of the car')
sns.boxplot(data=df,y='year',x='owner')
```

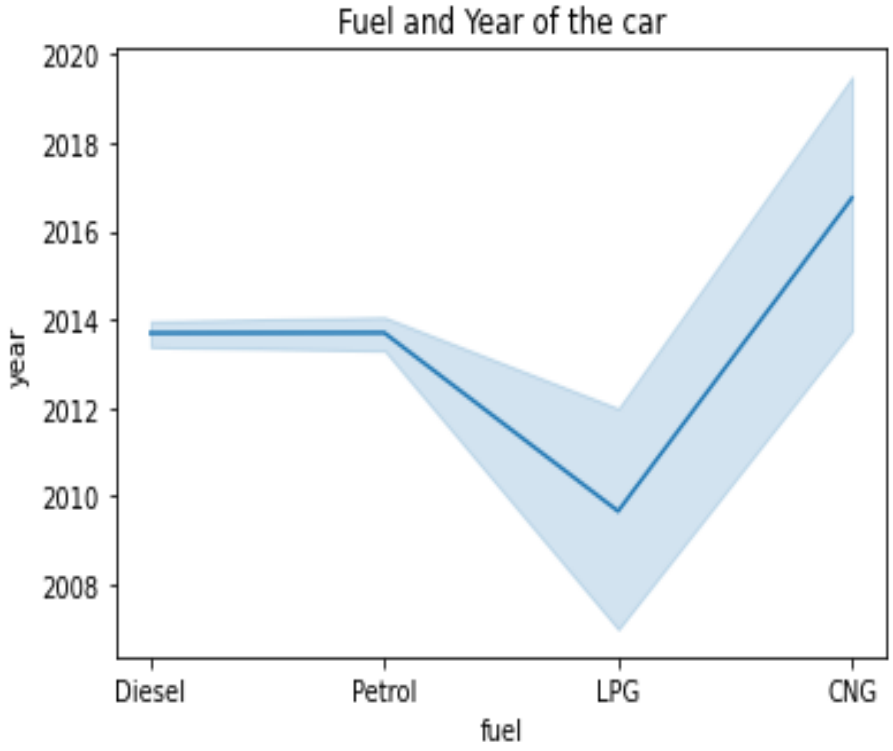


In the above box plot, the x-axis represents the no.of owner whereas y-axis represents the year in which car was bought.First owner took the car in year 2015 and fourth and above no.of owners took the car in 2010.

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

LINE PLOT

```
import pandas as pd
import seaborn as sns
import io
import matplotlib.pyplot as plt
df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
plt.title('Fuel and Year of the car')
sns.lineplot(data=df,x='fuel',y='year')
```

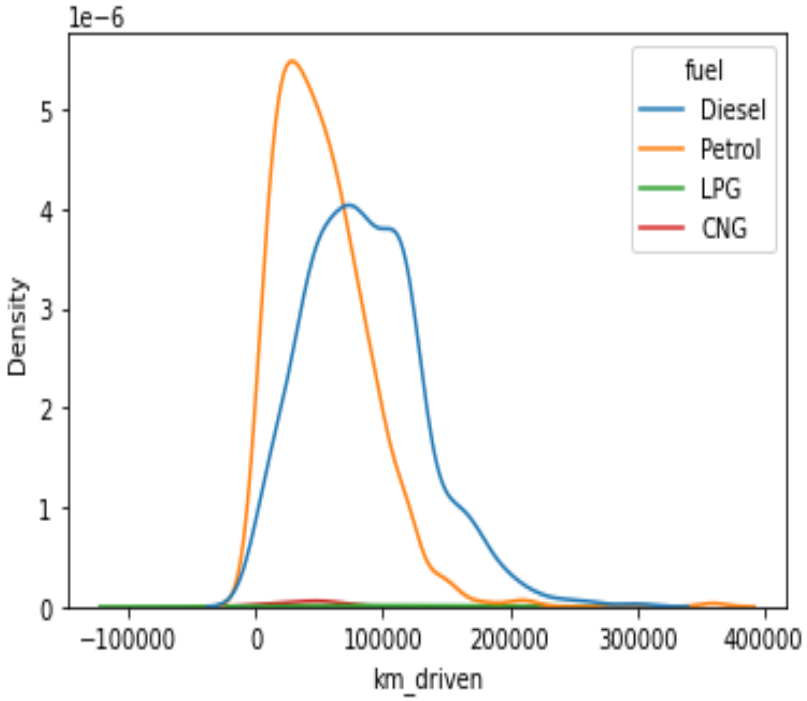


In this line plot , x-axis represents type of fuels whereas y-axis represents year of the car. As we can see in this line plot there is so much increase in the sales of cng type of fuel cars in the year 2015-2016.

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

KDE PLOT

```
import pandas as pd
import seaborn as sns
import io
import matplotlib.pyplot as plt
df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
sns.kdeplot(data=df,hue='fuel',x='km_driven')
```

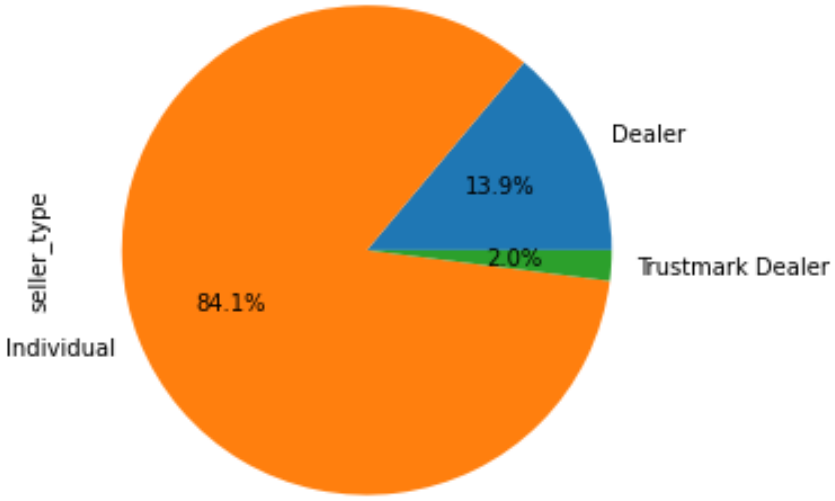


In the above kde plot , x-axis represents kilometers travelled by the car whereas y-axis represents fuel type of the car. As we can see there are more petrol cars as compared to any other fuel type of cars.

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

PIE PLOT

```
[ ] import pandas as pd
    import seaborn as sns
    import io
    import matplotlib.pyplot as plt
    df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
    print(df['seller_type'].groupby(df['seller_type']).count().plot.pie(figsize=(5,5),autopct='%1.1f%%',startangle=0,explode=[0,0,0]))
```

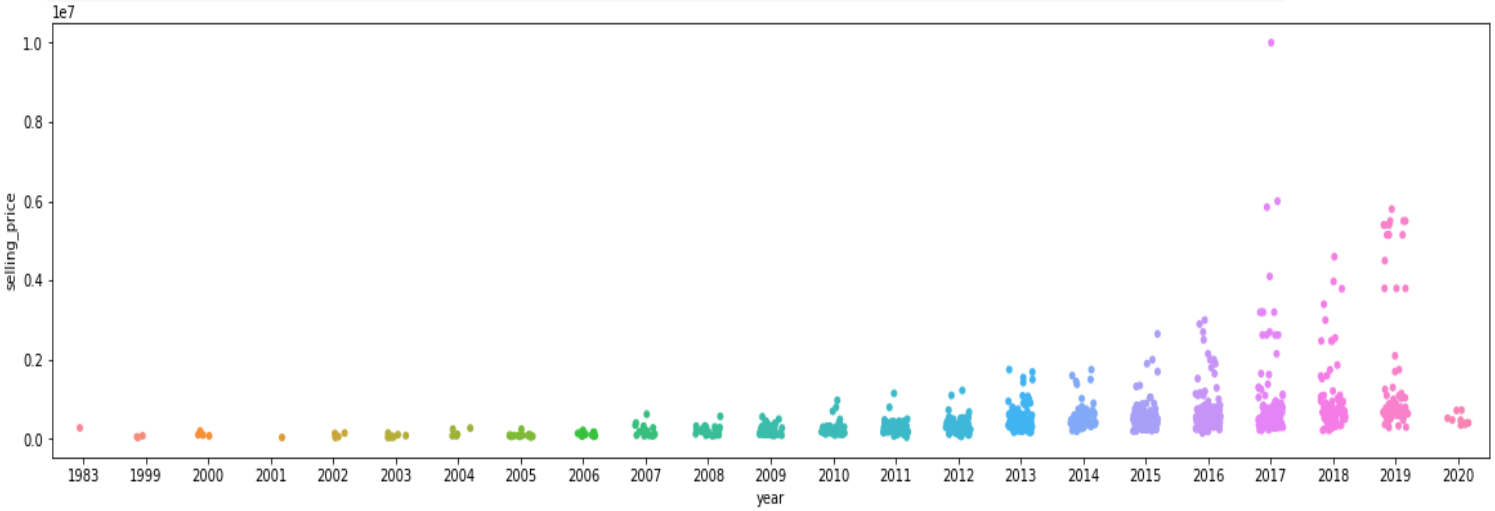


In this pie plot, it represents the type of sellers for example individual sellers, dealers, trustmark dealers. As we can see there are more individual types of sellers than any other type of sellers.

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

STRIP PLOT

```
[ ]  import pandas as pd
     import seaborn as sns
     import io
     import matplotlib.pyplot as plt
     fig=plt.figure(figsize=(20,5))
     sns.stripplot(data=df,y='selling_price',x='year',jitter=0.2)
```

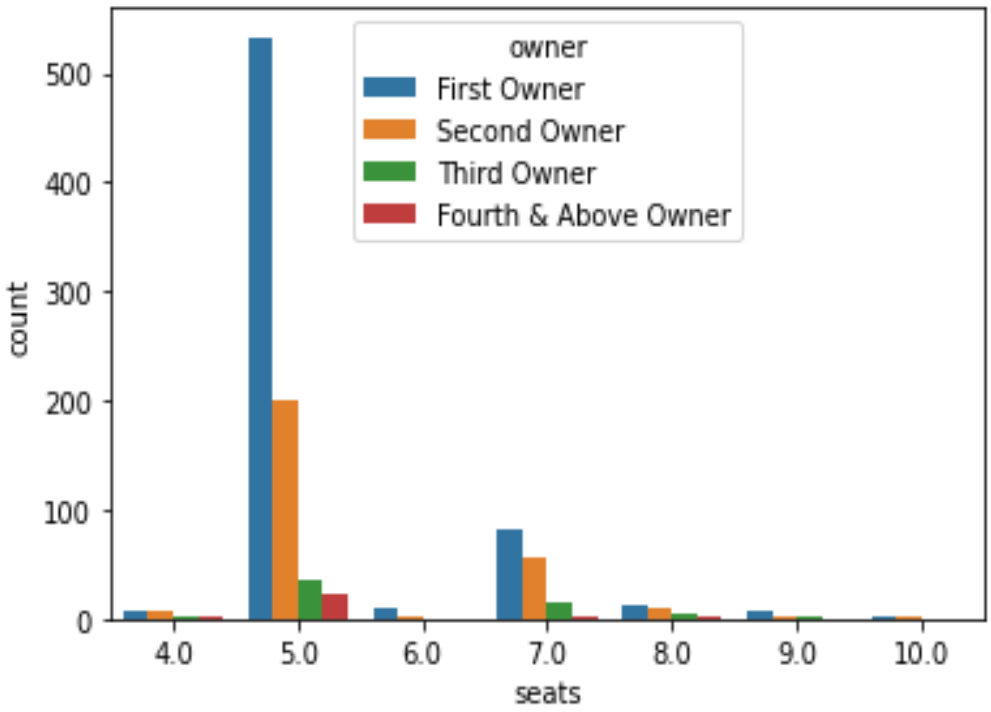


In the above strip plot , x-axis represents year of the car whereas y-axis represents the selling price of the cars.as we can there is so much rise in the selling price of the cars in the year 2019.

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

COUNT PLOT

```
[ ]  import pandas as pd
     import seaborn as sns
     df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
     sns.countplot(data=df,hue='owner',x='seats')
```

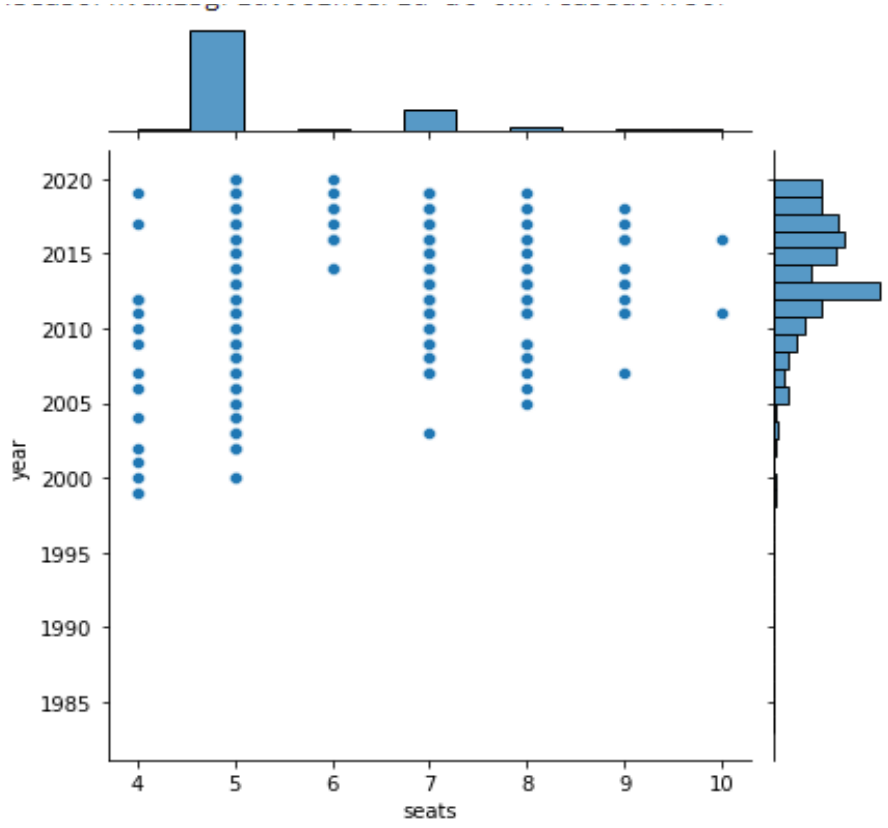


In the above count plot, x-axis represents seats in the car whereas y-axis represents the no.of car owners. As we can see there are more car of 5 seats which are bought by first owner.

JOINT PLOT

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

```
[ ]  import pandas as pd
     import seaborn as sns
     import io
     import matplotlib.pyplot as plt
     df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
     sns.jointplot(data=df,y='year',x='seats')
```



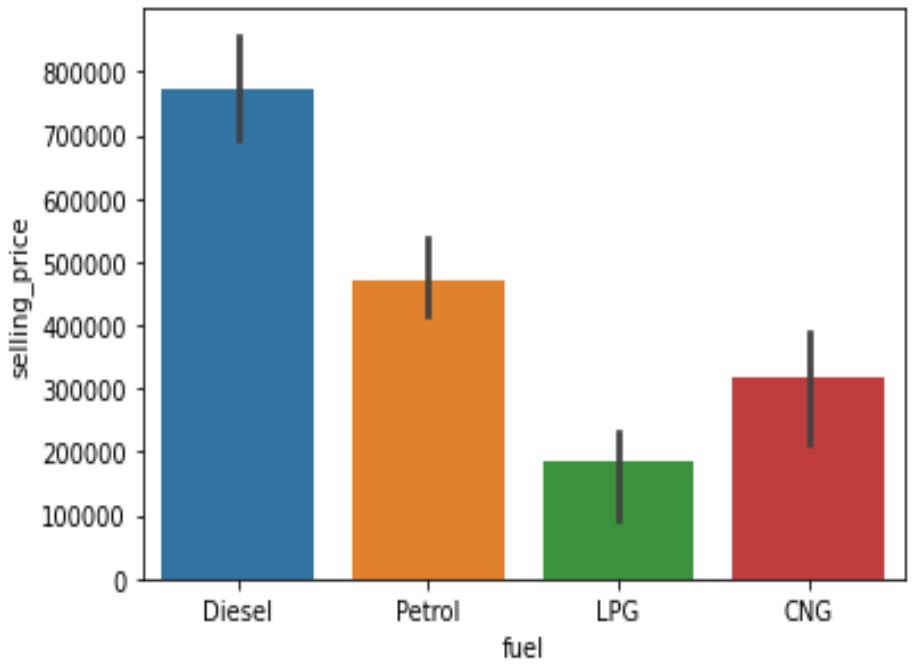
In the above joint plot, x-axis represents seats of the car whereas y-axis represents the year of the car. As we can see there is more sale of 5 seater car than any other car.

-

NAME-SARTHAK KADAM
CLASS- SYCS-A
ROLL NO.5863

BAR PLOT

```
import pandas as pd
import seaborn as sns
import io
import matplotlib.pyplot as plt
df=pd.read_csv(io.BytesIO(uploaded['cars details.csv']))
sns.barplot(data=df,x='fuel',y='selling_price')
```



In the above barplot, x-axis represents the fuel type of the car whereas y-axis represents selling price of the cars. As we can see the price of diesel based car is much more than any other fuel type of the car.