dataset-eda-feature-engineering

August 10, 2024

1 Data Analysis on Laptop Dataset

[195]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')


from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split

from sklearn.linear_model import LinearRegression
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import RandomForestRegressor
from sklearn.ensemble import AdaBoostRegressor
from sklearn.ensemble import GradientBoostingRegressor
```

[196]:
```
df=pd.read_csv("laptop_data.csv")
```

[197]:
```
df
```

[197]:
```
      Unnamed: 0 Company            TypeName  Inches  \
0              0   Apple           Ultrabook    13.3
1              1   Apple           Ultrabook    13.3
2              2      HP            Notebook    15.6
3              3   Apple           Ultrabook    15.4
4              4   Apple           Ultrabook    13.3
…            …     …                       …     …
1298        1298  Lenovo  2 in 1 Convertible    14.0
1299        1299  Lenovo  2 in 1 Convertible    13.3
1300        1300  Lenovo            Notebook    14.0
1301        1301      HP            Notebook    15.6
1302        1302    Asus            Notebook    15.6

                              ScreenResolution  \
```

```
0             IPS Panel Retina Display 2560x1600
1                                       1440x900
2                             Full HD 1920x1080
3             IPS Panel Retina Display 2880x1800
4             IPS Panel Retina Display 2560x1600
…                                           …
1298   IPS Panel Full HD / Touchscreen 1920x1080
1299  IPS Panel Quad HD+ / Touchscreen 3200x1800
1300                                    1366x768
1301                                    1366x768
1302                                    1366x768

                                     Cpu   Ram               Memory  \
0                     Intel Core i5 2.3GHz   8GB            128GB SSD
1                     Intel Core i5 1.8GHz   8GB  128GB Flash Storage
2                Intel Core i5 7200U 2.5GHz   8GB            256GB SSD
3                     Intel Core i7 2.7GHz  16GB            512GB SSD
4                     Intel Core i5 3.1GHz   8GB            256GB SSD
…                                       …   …                …
1298             Intel Core i7 6500U 2.5GHz   4GB            128GB SSD
1299             Intel Core i7 6500U 2.5GHz  16GB            512GB SSD
1300  Intel Celeron Dual Core N3050 1.6GHz   2GB   64GB Flash Storage
1301             Intel Core i7 6500U 2.5GHz   6GB              1TB HDD
1302  Intel Celeron Dual Core N3050 1.6GHz   4GB            500GB HDD

                               Gpu       OpSys  Weight        Price
0     Intel Iris Plus Graphics 640       macOS  1.37kg   71378.6832
1           Intel HD Graphics 6000       macOS  1.34kg   47895.5232
2            Intel HD Graphics 620       No OS  1.86kg   30636.0000
3               AMD Radeon Pro 455       macOS  1.83kg  135195.3360
4     Intel Iris Plus Graphics 650       macOS  1.37kg   96095.8080
…                              …         …     …           …
1298         Intel HD Graphics 520  Windows 10   1.8kg   33992.6400
1299         Intel HD Graphics 520  Windows 10   1.3kg   79866.7200
1300             Intel HD Graphics  Windows 10   1.5kg   12201.1200
1301            AMD Radeon R5 M330  Windows 10  2.19kg   40705.9200
1302             Intel HD Graphics  Windows 10   2.2kg   19660.3200

[1303 rows x 12 columns]
```

[198]:
```
df.shape
```

[198]:
```
(1303, 12)
```

[199]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 1303 entries, 0 to 1302
Data columns (total 12 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Unnamed: 0        1303 non-null   int64
 1   Company           1303 non-null   object
 2   TypeName          1303 non-null   object
 3   Inches            1303 non-null   float64
 4   ScreenResolution  1303 non-null   object
 5   Cpu               1303 non-null   object
 6   Ram               1303 non-null   object
 7   Memory            1303 non-null   object
 8   Gpu               1303 non-null   object
 9   OpSys             1303 non-null   object
 10  Weight            1303 non-null   object
 11  Price             1303 non-null   float64
dtypes: float64(2), int64(1), object(9)
memory usage: 122.3+ KB
```

[200]:
```
df.describe()
```

[200]:

	Unnamed: 0	Inches	Price
count	1303.00000	1303.000000	1303.000000
mean	651.00000	15.017191	59870.042910
std	376.28801	1.426304	37243.201786
min	0.00000	10.100000	9270.720000
25%	325.50000	14.000000	31914.720000
50%	651.00000	15.600000	52054.560000
75%	976.50000	15.600000	79274.246400
max	1302.00000	18.400000	324954.720000

[201]:
```
df.describe(include=object)
```

[201]:

	Company	TypeName	ScreenResolution	Cpu	Ram	Memory	Gpu	OpSys	Weight
count	1303	1303	1303	1303	1303	1303	1303	1303	1303
unique	19	6	40	118	9	39	110	9	179
top	Dell	Notebook	Full HD 1920x1080	Intel Core i5 7200U 2.5GHz	8GB	256GB SSD	Intel HD Graphics 620	Windows 10	2.2kg
freq	297	727	507	190	619	412	281	1072	121

[202]:
```
df.isnull().sum()/len(df)*100
```

[202]:
```
Unnamed: 0          0.0
Company             0.0
TypeName            0.0
Inches              0.0
ScreenResolution    0.0
Cpu                 0.0
Ram                 0.0
Memory              0.0
Gpu                 0.0
OpSys               0.0
Weight              0.0
Price               0.0
dtype: float64
```

[203]:
```
df.duplicated().sum()/len(df)*100
```

[203]:
```
0.0
```

[204]:
```
df.head()
```

[204]:

	Unnamed: 0	Company	TypeName	Inches	ScreenResolution	Cpu	Ram	Memory	Gpu	OpSys	Weight	Price
0	0	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 2.3GHz	8GB	128GB SSD	Intel Iris Plus Graphics 640	macOS	1.37kg	71378.6832
1	1	Apple	Ultrabook	13.3	1440x900	Intel Core i5 1.8GHz	8GB	128GB Flash Storage	Intel HD Graphics 6000	macOS	1.34kg	47895.5232
2	2	HP	Notebook	15.6	Full HD 1920x1080	Intel Core i5 7200U 2.5GHz	8GB	256GB SSD	Intel HD Graphics 620	No OS	1.86kg	30636.0000
3	3	Apple	Ultrabook	15.4	IPS Panel Retina Display 2880x1800	Intel Core i7 2.7GHz	16GB	512GB SSD	AMD Radeon Pro 455	macOS	1.83kg	135195.3360
4	4	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 3.1GHz	8GB	256GB SSD	Intel Iris Plus Graphics 650	macOS	1.37kg	96095.8080

2 Data Cleaning

[205]:
```
df.drop(columns=['Unnamed: 0'],inplace=True)
```

[206]:
```
df
```

[206]:

	Company	TypeName	Inches	ScreenResolution	Cpu	Ram	Memory	Gpu	OpSys	Weight	Price
0	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 2.3GHz	8GB	128GB SSD	Intel Iris Plus Graphics 640	macOS	1.37kg	71378.6832
1	Apple	Ultrabook	13.3	1440x900	Intel Core i5 1.8GHz	8GB	128GB Flash Storage	Intel HD Graphics 6000	macOS	1.34kg	47895.5232
2	HP	Notebook	15.6	Full HD 1920x1080	Intel Core i5 7200U 2.5GHz	8GB	256GB SSD	Intel HD Graphics 620	No OS	1.86kg	30636.0000
3	Apple	Ultrabook	15.4	IPS Panel Retina Display 2880x1800	Intel Core i7 2.7GHz	16GB	512GB SSD	AMD Radeon Pro 455	macOS	1.83kg	135195.3360
4	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 3.1GHz	8GB	256GB SSD	Intel Iris Plus Graphics 650	macOS	1.37kg	96095.8080
…	…	…	…	…	…	…	…	…	…	…	…
1298	Lenovo	2 in 1 Convertible	14.0	IPS Panel Full HD / Touchscreen 1920x1080	Intel Core i7 6500U 2.5GHz	4GB	128GB SSD	Intel HD Graphics 520	Windows 10	1.8kg	33992.6400
1299	Lenovo	2 in 1 Convertible	13.3	IPS Panel Quad HD+ / Touchscreen 3200x1800	Intel Core i7 6500U 2.5GHz	16GB	512GB SSD				
1300	Lenovo	Notebook	14.0	1366x768	Intel Celeron Dual Core N3050 1.6GHz	2GB	64GB Flash Storage				
1301	HP	Notebook	15.6	1366x768	Intel Core i7 6500U 2.5GHz	6GB	1TB HDD				
1302	Asus	Notebook	15.6	1366x768	Intel Celeron Dual Core N3050 1.6GHz	4GB	500GB HDD				

```
1299         Intel HD Graphics 520  Windows 10   1.3kg   79866.7200
1300             Intel HD Graphics  Windows 10   1.5kg   12201.1200
1301           AMD Radeon R5 M330  Windows 10  2.19kg   40705.9200
1302             Intel HD Graphics  Windows 10   2.2kg   19660.3200

[1303 rows x 11 columns]
```

[207]:
```
df['Ram']=df['Ram'].str.replace("GB","")
```

[208]:
```
df['Ram']=df['Ram'].astype(int)
```

[209]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1303 entries, 0 to 1302
Data columns (total 11 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Company           1303 non-null   object
 1   TypeName          1303 non-null   object
 2   Inches            1303 non-null   float64
 3   ScreenResolution  1303 non-null   object
 4   Cpu               1303 non-null   object
 5   Ram               1303 non-null   int32
 6   Memory            1303 non-null   object
 7   Gpu               1303 non-null   object
 8   OpSys             1303 non-null   object
 9   Weight            1303 non-null   object
 10  Price             1303 non-null   float64
dtypes: float64(2), int32(1), object(8)
memory usage: 107.0+ KB
```

[210]:
```
df['Weight']=df['Weight'].str.replace("kg","")
```

[211]:
```
df['Weight']=df['Weight'].astype(float)
```

[212]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1303 entries, 0 to 1302
Data columns (total 11 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Company           1303 non-null   object
 1   TypeName          1303 non-null   object
 2   Inches            1303 non-null   float64
 3   ScreenResolution  1303 non-null   object
```

```
 4   Cpu               1303 non-null   object
 5   Ram               1303 non-null   int32
 6   Memory            1303 non-null   object
 7   Gpu               1303 non-null   object
 8   OpSys             1303 non-null   object
 9   Weight            1303 non-null   float64
 10  Price             1303 non-null   float64
dtypes: float64(3), int32(1), object(7)
memory usage: 107.0+ KB
```

[213]:
```
df
```

[213]:
```
      Company            TypeName  Inches  \
0       Apple           Ultrabook    13.3
1       Apple           Ultrabook    13.3
2          HP            Notebook    15.6
3       Apple           Ultrabook    15.4
4       Apple           Ultrabook    13.3
…       …                     …     …
1298   Lenovo  2 in 1 Convertible    14.0
1299   Lenovo  2 in 1 Convertible    13.3
1300   Lenovo            Notebook    14.0
1301       HP            Notebook    15.6
1302     Asus            Notebook    15.6

                               ScreenResolution  \
0             IPS Panel Retina Display 2560x1600
1                                      1440x900
2                               Full HD 1920x1080
3             IPS Panel Retina Display 2880x1800
4             IPS Panel Retina Display 2560x1600
…                                           …
1298   IPS Panel Full HD / Touchscreen 1920x1080
1299  IPS Panel Quad HD+ / Touchscreen 3200x1800
1300                                   1366x768
1301                                   1366x768
1302                                   1366x768

                                    Cpu  Ram               Memory  \
0                 Intel Core i5 2.3GHz    8            128GB SSD
1                 Intel Core i5 1.8GHz    8  128GB Flash Storage
2               Intel Core i5 7200U 2.5GHz    8            256GB SSD
3                 Intel Core i7 2.7GHz   16            512GB SSD
4                 Intel Core i5 3.1GHz    8            256GB SSD
…                                    …  …                    …
1298            Intel Core i7 6500U 2.5GHz    4            128GB SSD
1299            Intel Core i7 6500U 2.5GHz   16            512GB SSD
```

```
1300  Intel Celeron Dual Core N3050 1.6GHz    2   64GB Flash Storage
1301            Intel Core i7 6500U 2.5GHz    6              1TB HDD
1302  Intel Celeron Dual Core N3050 1.6GHz    4            500GB HDD

                               Gpu       OpSys  Weight        Price
0     Intel Iris Plus Graphics 640       macOS    1.37   71378.6832
1           Intel HD Graphics 6000       macOS    1.34   47895.5232
2            Intel HD Graphics 620       No OS    1.86   30636.0000
3               AMD Radeon Pro 455       macOS    1.83  135195.3360
4     Intel Iris Plus Graphics 650       macOS    1.37   96095.8080
…                              …         …      …          …
1298         Intel HD Graphics 520  Windows 10    1.80   33992.6400
1299         Intel HD Graphics 520  Windows 10    1.30   79866.7200
1300             Intel HD Graphics  Windows 10    1.50   12201.1200
1301            AMD Radeon R5 M330  Windows 10    2.19   40705.9200
1302             Intel HD Graphics  Windows 10    2.20   19660.3200

[1303 rows x 11 columns]
```

[214]:
```
df['OpSys'].value_counts()
```

[214]:
```
Windows 10      1072
No OS             66
Linux             62
Windows 7         45
Chrome OS         27
macOS             13
Windows 10 S       8
Mac OS X           8
Android            2
Name: OpSys, dtype: int64
```

3 Exploratory data analysis

[215]:
```
df
```

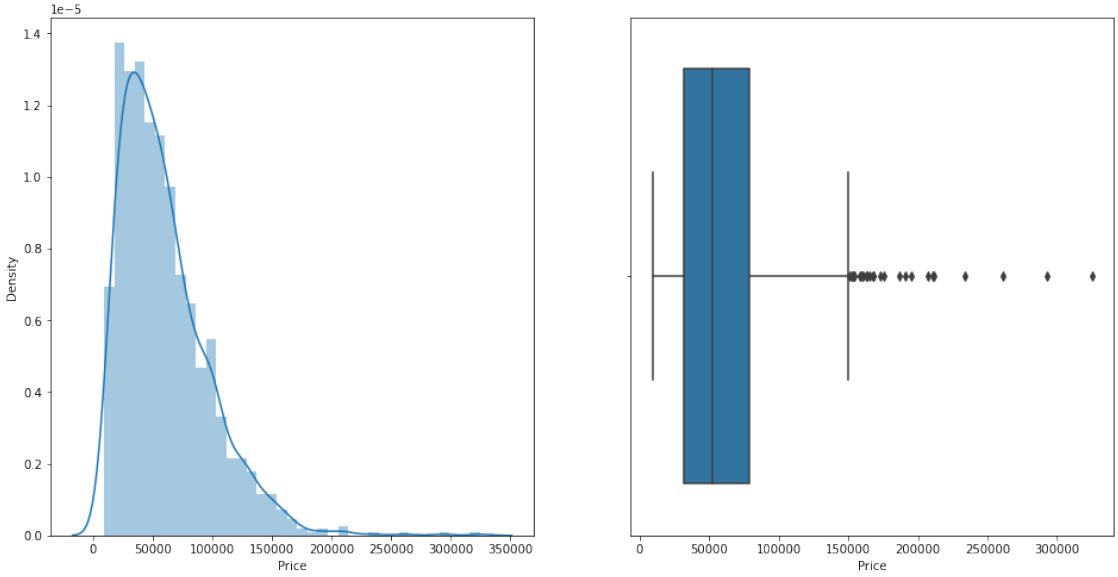
[215]:
```
      Company            TypeName  Inches  \
0       Apple           Ultrabook    13.3
1       Apple           Ultrabook    13.3
2          HP            Notebook    15.6
3       Apple           Ultrabook    15.4
4       Apple           Ultrabook    13.3
…       …                     …     …
1298  Lenovo  2 in 1 Convertible    14.0
1299  Lenovo  2 in 1 Convertible    13.3
1300  Lenovo            Notebook    14.0
```

```
1301      HP            Notebook    15.6
1302    Asus            Notebook    15.6

                               ScreenResolution  \
0             IPS Panel Retina Display 2560x1600
1                                     1440x900
2                            Full HD 1920x1080
3             IPS Panel Retina Display 2880x1800
4             IPS Panel Retina Display 2560x1600
…                                           …
1298   IPS Panel Full HD / Touchscreen 1920x1080
1299  IPS Panel Quad HD+ / Touchscreen 3200x1800
1300                                  1366x768
1301                                  1366x768
1302                                  1366x768

                                    Cpu  Ram               Memory  \
0                     Intel Core i5 2.3GHz    8            128GB SSD
1                     Intel Core i5 1.8GHz    8  128GB Flash Storage
2                Intel Core i5 7200U 2.5GHz    8            256GB SSD
3                     Intel Core i7 2.7GHz   16            512GB SSD
4                     Intel Core i5 3.1GHz    8            256GB SSD
…                                       …  …                    …
1298             Intel Core i7 6500U 2.5GHz    4            128GB SSD
1299             Intel Core i7 6500U 2.5GHz   16            512GB SSD
1300  Intel Celeron Dual Core N3050 1.6GHz    2   64GB Flash Storage
1301             Intel Core i7 6500U 2.5GHz    6              1TB HDD
1302  Intel Celeron Dual Core N3050 1.6GHz    4            500GB HDD

                               Gpu       OpSys  Weight        Price
0     Intel Iris Plus Graphics 640       macOS    1.37   71378.6832
1           Intel HD Graphics 6000       macOS    1.34   47895.5232
2            Intel HD Graphics 620       No OS    1.86   30636.0000
3               AMD Radeon Pro 455       macOS    1.83  135195.3360
4     Intel Iris Plus Graphics 650       macOS    1.37   96095.8080
…                              …         …     …           …
1298         Intel HD Graphics 520  Windows 10    1.80   33992.6400
1299         Intel HD Graphics 520  Windows 10    1.30   79866.7200
1300             Intel HD Graphics  Windows 10    1.50   12201.1200
1301            AMD Radeon R5 M330  Windows 10    2.19   40705.9200
1302             Intel HD Graphics  Windows 10    2.20   19660.3200

[1303 rows x 11 columns]
```

4 1. Target Feature

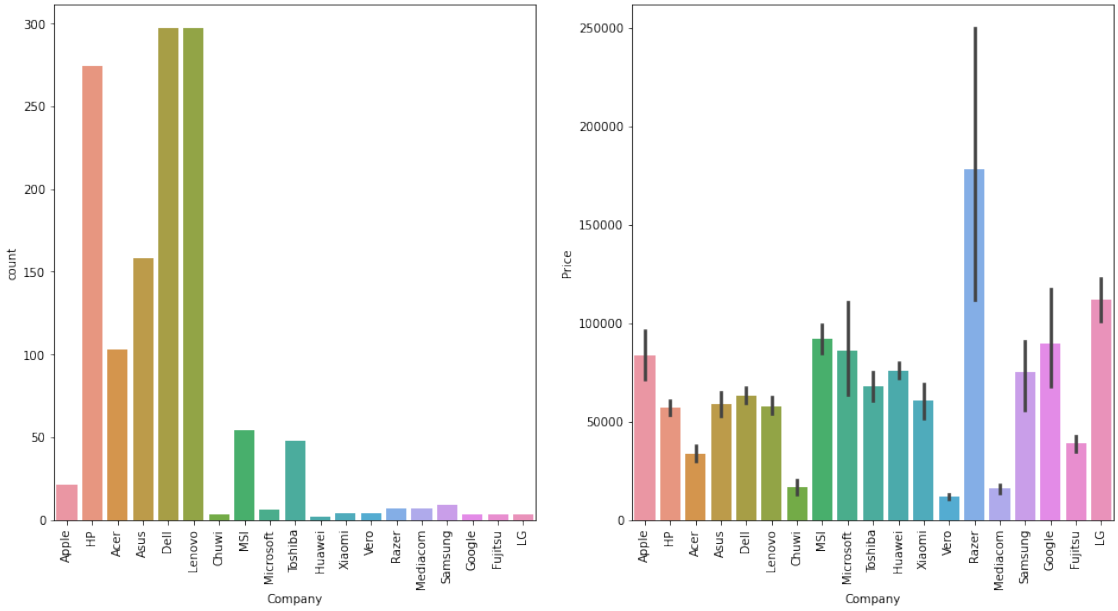
[216]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.distplot(df['Price'])

plt.subplot(1,2,2)
sns.boxplot(df['Price'])
```

[216]: <AxesSubplot:xlabel='Price'>



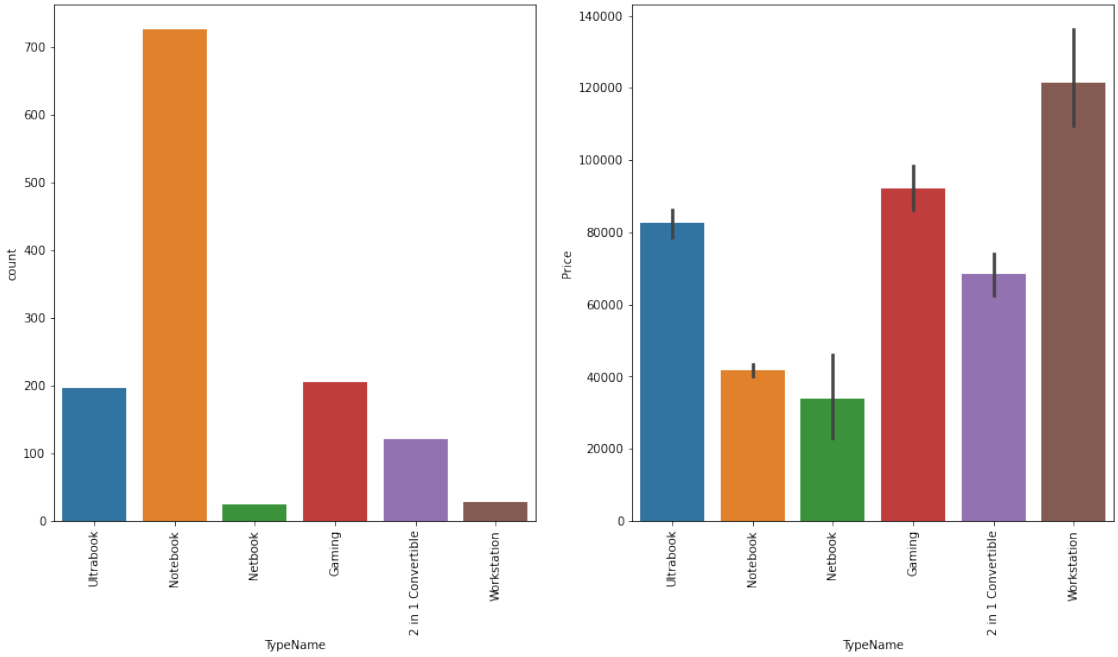
5 2 Company Feature

[217]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['Company'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['Company'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```



6 3 Type name feature

[218]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['TypeName'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['TypeName'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```

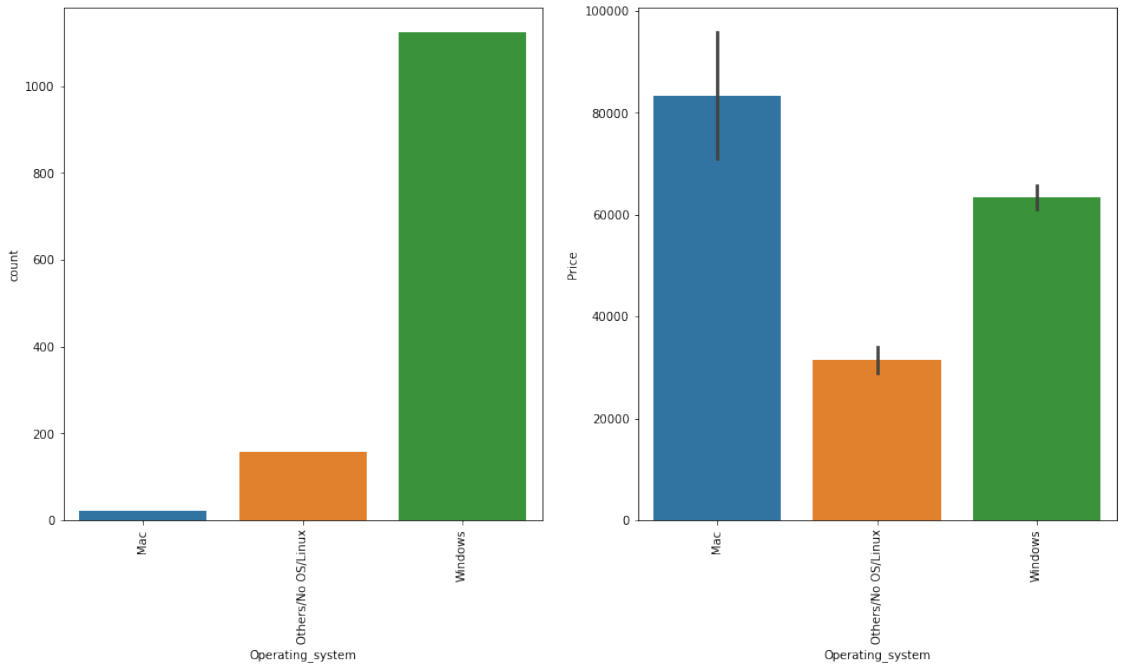


[219]:
```
def cat_os(inp):
    if inp == 'Windows 10' or inp == 'Windows 7' or inp == 'Windows 10 S':
        return 'Windows'
    elif inp == 'macOS' or inp == 'Mac OS X':
        return 'Mac'
    else:
        return 'Others/No OS/Linux'
```

[220]:
```
df['Operating_system']=df['OpSys'].apply(cat_os)
```

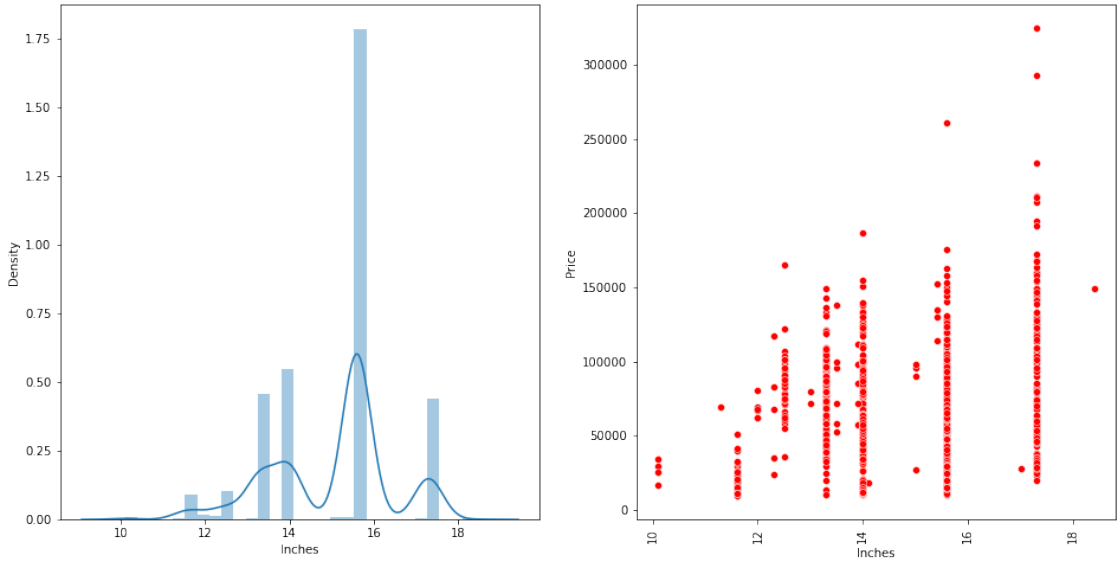
7 4 Operating_System Feature

[221]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['Operating_system'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['Operating_system'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```



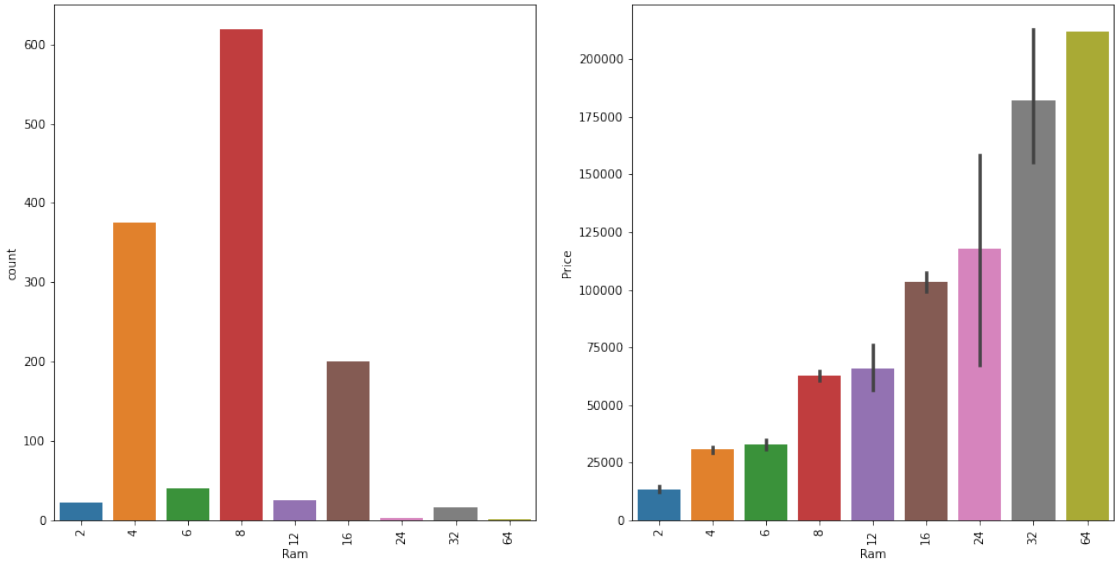
8 5 Inches Feature

[222]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.distplot(df['Inches'])

plt.subplot(1,2,2)
sns.scatterplot(df['Inches'],df['Price'],color='red',data=df)
plt.xticks(rotation=90)
plt.show()
```



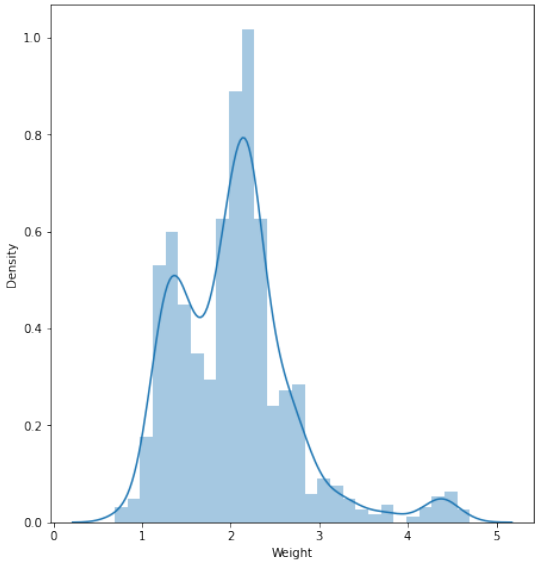
9 6 Ram Feature

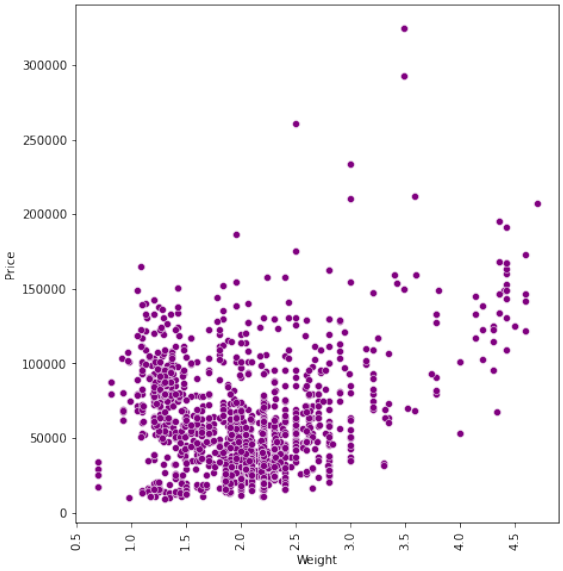
[223]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['Ram'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['Ram'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```



10 7 Weight Features

[224]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.distplot(df['Weight'])


plt.subplot(1,2,2)
sns.scatterplot(df['Weight'],df['Price'],color='purple',data=df)
plt.xticks(rotation=90)
plt.show()
```





11 Feature Engineering

[225]:
```
df['ScreenResolution'].value_counts()
```

[225]:
```
Full HD 1920x1080                                507
1366x768                                         281
IPS Panel Full HD 1920x1080                      230
IPS Panel Full HD / Touchscreen 1920x1080         53
Full HD / Touchscreen 1920x1080                   47
1600x900                                          23
Touchscreen 1366x768                              16
Quad HD+ / Touchscreen 3200x1800                  15
IPS Panel 4K Ultra HD 3840x2160                   12
IPS Panel 4K Ultra HD / Touchscreen 3840x2160     11
4K Ultra HD / Touchscreen 3840x2160               10
4K Ultra HD 3840x2160                              7
IPS Panel 1366x768                                 7
Touchscreen 2560x1440                              7
IPS Panel Quad HD+ / Touchscreen 3200x1800         6
IPS Panel Retina Display 2560x1600                 6
Touchscreen 2256x1504                              6
IPS Panel Retina Display 2304x1440                 6
IPS Panel Touchscreen 2560x1440                    5
1440x900                                           4
IPS Panel Retina Display 2880x1800                 4
IPS Panel Touchscreen 1920x1200                    4
IPS Panel 2560x1440                                4
```

```
IPS Panel Quad HD+ 2560x1440                        3
Touchscreen 2400x1600                               3
Quad HD+ 3200x1800                                  3
2560x1440                                           3
IPS Panel Touchscreen 1366x768                      3
1920x1080                                           3
IPS Panel Quad HD+ 3200x1800                        2
IPS Panel Full HD 2160x1440                         2
IPS Panel Touchscreen / 4K Ultra HD 3840x2160       2
IPS Panel Full HD 1920x1200                         1
IPS Panel Retina Display 2736x1824                  1
IPS Panel Full HD 2560x1440                         1
Touchscreen / Quad HD+ 3200x1800                    1
Touchscreen / Full HD 1920x1080                     1
Touchscreen / 4K Ultra HD 3840x2160                 1
IPS Panel Full HD 1366x768                          1
IPS Panel Touchscreen 2400x1600                     1
Name: ScreenResolution, dtype: int64
```

[226]:
```
df['TouchScreen']=df['ScreenResolution'].apply(lambda  x:1 if 'Touchscreen' in␣
↪x else 0)
```

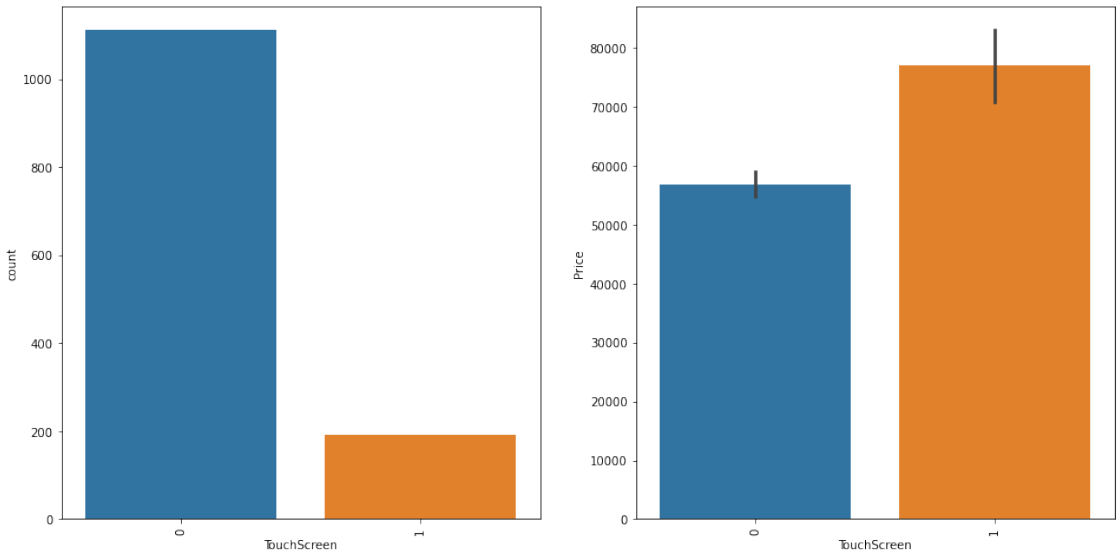
[227]:
```
df
```

[227]:
```
     Company            TypeName  Inches  \
0      Apple           Ultrabook    13.3
1      Apple           Ultrabook    13.3
2         HP            Notebook    15.6
3      Apple           Ultrabook    15.4
4      Apple           Ultrabook    13.3
…      …                     …     …
1298  Lenovo  2 in 1 Convertible    14.0
1299  Lenovo  2 in 1 Convertible    13.3
1300  Lenovo            Notebook    14.0
1301      HP            Notebook    15.6
1302    Asus            Notebook    15.6

                               ScreenResolution  \
0             IPS Panel Retina Display 2560x1600
1                                       1440x900
2                              Full HD 1920x1080
3             IPS Panel Retina Display 2880x1800
4             IPS Panel Retina Display 2560x1600
…                                             …
1298   IPS Panel Full HD / Touchscreen 1920x1080
1299  IPS Panel Quad HD+ / Touchscreen 3200x1800
1300                                    1366x768
```

```
1301                                  1366x768
1302                                  1366x768

                                        Cpu  Ram               Memory  \
0                       Intel Core i5 2.3GHz    8            128GB SSD
1                       Intel Core i5 1.8GHz    8  128GB Flash Storage
2                  Intel Core i5 7200U 2.5GHz    8            256GB SSD
3                       Intel Core i7 2.7GHz   16            512GB SSD
4                       Intel Core i5 3.1GHz    8            256GB SSD
…                                        …  …                    …
1298               Intel Core i7 6500U 2.5GHz    4            128GB SSD
1299               Intel Core i7 6500U 2.5GHz   16            512GB SSD
1300  Intel Celeron Dual Core N3050 1.6GHz    2   64GB Flash Storage
1301               Intel Core i7 6500U 2.5GHz    6              1TB HDD
1302  Intel Celeron Dual Core N3050 1.6GHz    4            500GB HDD

                               Gpu       OpSys  Weight        Price  \
0     Intel Iris Plus Graphics 640       macOS    1.37   71378.6832
1           Intel HD Graphics 6000       macOS    1.34   47895.5232
2            Intel HD Graphics 620       No OS    1.86   30636.0000
3               AMD Radeon Pro 455       macOS    1.83  135195.3360
4     Intel Iris Plus Graphics 650       macOS    1.37   96095.8080
…                              …         …     …           …
1298         Intel HD Graphics 520  Windows 10    1.80   33992.6400
1299         Intel HD Graphics 520  Windows 10    1.30   79866.7200
1300             Intel HD Graphics  Windows 10    1.50   12201.1200
1301            AMD Radeon R5 M330  Windows 10    2.19   40705.9200
1302             Intel HD Graphics  Windows 10    2.20   19660.3200

        Operating_system  TouchScreen
0                    Mac            0
1                    Mac            0
2     Others/No OS/Linux            0
3                    Mac            0
4                    Mac            0
…                      …          …
1298             Windows            1
1299             Windows            1
1300             Windows            0
1301             Windows            0
1302             Windows            0

[1303 rows x 13 columns]
```

12 8 TouchScreen Feature

[228]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['TouchScreen'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['TouchScreen'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```



[229]:
```
df['IPS']=df['ScreenResolution'].apply(lambda  x:1 if 'IPS' in x else 0)
```

[230]:
```
df
```

[230]:

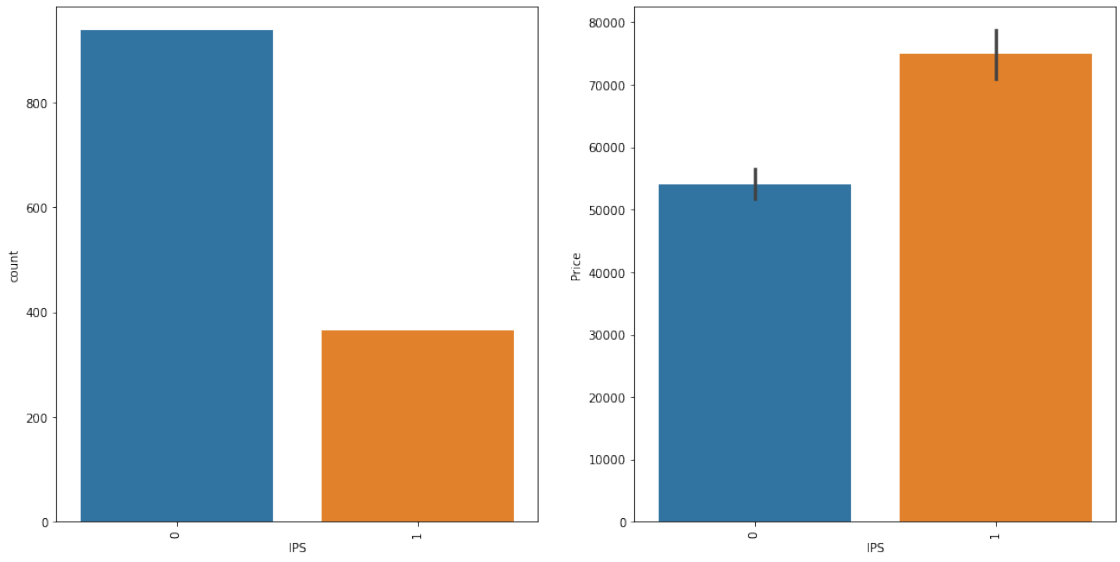
	Company	TypeName	Inches	\
0	Apple	Ultrabook	13.3	
1	Apple	Ultrabook	13.3	
2	HP	Notebook	15.6	
3	Apple	Ultrabook	15.4	
4	Apple	Ultrabook	13.3	
…	…	…	…	
1298	Lenovo	2 in 1 Convertible	14.0	
1299	Lenovo	2 in 1 Convertible	13.3	
1300	Lenovo	Notebook	14.0	
1301	HP	Notebook	15.6	
1302	Asus	Notebook	15.6	

```
                                ScreenResolution  \
0              IPS Panel Retina Display 2560x1600
1                                        1440x900
2                              Full HD 1920x1080
3              IPS Panel Retina Display 2880x1800
4              IPS Panel Retina Display 2560x1600
…                                              …
1298   IPS Panel Full HD / Touchscreen 1920x1080
1299  IPS Panel Quad HD+ / Touchscreen 3200x1800
1300                                     1366x768
1301                                     1366x768
1302                                     1366x768

                                       Cpu  Ram               Memory  \
0                       Intel Core i5 2.3GHz    8            128GB SSD
1                       Intel Core i5 1.8GHz    8  128GB Flash Storage
2                  Intel Core i5 7200U 2.5GHz    8            256GB SSD
3                       Intel Core i7 2.7GHz   16            512GB SSD
4                       Intel Core i5 3.1GHz    8            256GB SSD
…                                        …  …                    …
1298               Intel Core i7 6500U 2.5GHz    4            128GB SSD
1299               Intel Core i7 6500U 2.5GHz   16            512GB SSD
1300  Intel Celeron Dual Core N3050 1.6GHz    2   64GB Flash Storage
1301               Intel Core i7 6500U 2.5GHz    6              1TB HDD
1302  Intel Celeron Dual Core N3050 1.6GHz    4            500GB HDD

                                Gpu       OpSys  Weight        Price  \
0     Intel Iris Plus Graphics 640       macOS    1.37   71378.6832
1           Intel HD Graphics 6000       macOS    1.34   47895.5232
2            Intel HD Graphics 620       No OS    1.86   30636.0000
3               AMD Radeon Pro 455       macOS    1.83  135195.3360
4     Intel Iris Plus Graphics 650       macOS    1.37   96095.8080
…                              …         …     …          …
1298         Intel HD Graphics 520  Windows 10    1.80   33992.6400
1299         Intel HD Graphics 520  Windows 10    1.30   79866.7200
1300             Intel HD Graphics  Windows 10    1.50   12201.1200
1301            AMD Radeon R5 M330  Windows 10    2.19   40705.9200
1302             Intel HD Graphics  Windows 10    2.20   19660.3200

        Operating_system  TouchScreen  IPS
0                    Mac            0    1
1                    Mac            0    0
2     Others/No OS/Linux            0    0
3                    Mac            0    1
4                    Mac            0    1
…                    …          …  …
```

```
1298            Windows            1    1
1299            Windows            1    1
1300            Windows            0    0
1301            Windows            0    0
1302            Windows            0    0

[1303 rows x 14 columns]
```

13 9 IPS Feature

[231]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['IPS'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['IPS'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```



[232]:
```
df['4K_Ultra_HD']=df['ScreenResolution'].apply(lambda  x:1 if '4K Ultra HD' in␣
  ↪x else 0)
```

[233]:
```
df
```

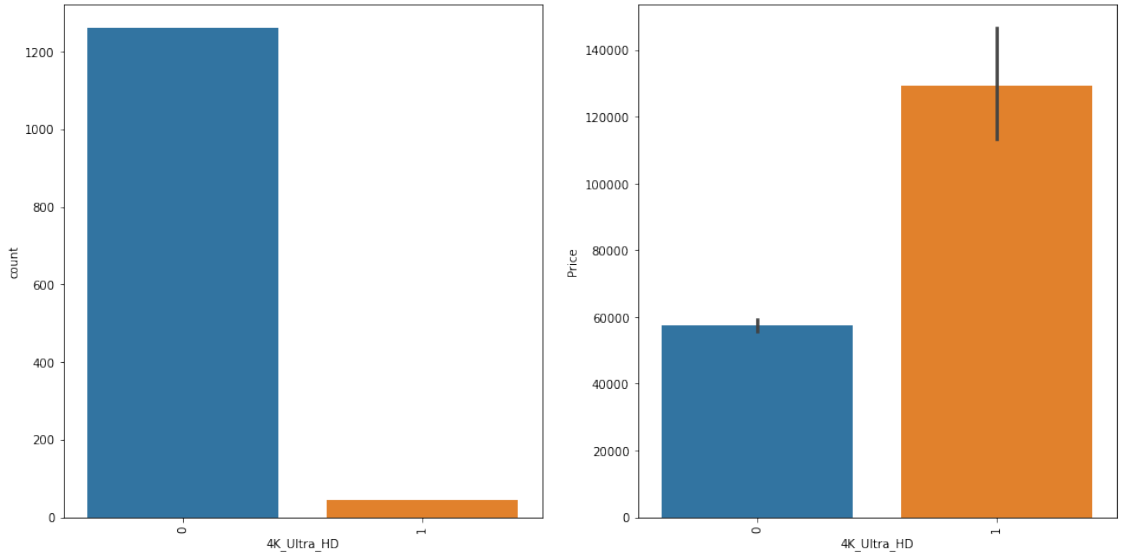
[233]:

	Company	TypeName	Inches	ScreenResolution	Cpu	Ram	Memory	Gpu	OpSys	Weight	Price
0	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 2.3GHz	8	128GB SSD	Intel Iris Plus Graphics 640	macOS	1.37	71378.6832
1	Apple	Ultrabook	13.3	1440x900	Intel Core i5 1.8GHz	8	128GB Flash Storage	Intel HD Graphics 6000	macOS	1.34	47895.5232
2	HP	Notebook	15.6	Full HD 1920x1080	Intel Core i5 7200U 2.5GHz	8	256GB SSD	Intel HD Graphics 620	No OS	1.86	30636.0000
3	Apple	Ultrabook	15.4	IPS Panel Retina Display 2880x1800	Intel Core i7 2.7GHz	16	512GB SSD	AMD Radeon Pro 455	macOS	1.83	135195.3360
4	Apple	Ultrabook	13.3	IPS Panel Retina Display 2560x1600	Intel Core i5 3.1GHz	8	256GB SSD	Intel Iris Plus Graphics 650	macOS	1.37	96095.8080
…	…	…	…	…	…	…	…	…	…	…	…
1298	Lenovo	2 in 1 Convertible	14.0	IPS Panel Full HD / Touchscreen 1920x1080	Intel Core i7 6500U 2.5GHz	4	128GB SSD	Intel HD Graphics 520	Windows 10	1.80	33992.6400
1299	Lenovo	2 in 1 Convertible	13.3	IPS Panel Quad HD+ / Touchscreen 3200x1800	Intel Core i7 6500U 2.5GHz	16	512GB SSD				
1300	Lenovo	Notebook	14.0	1366x768	Intel Celeron Dual Core N3050 1.6GHz	2	64GB Flash Storage				
1301	HP	Notebook	15.6	1366x768	Intel Core i7 6500U 2.5GHz	6	1TB HDD				
1302	Asus	Notebook	15.6	1366x768	Intel Celeron Dual Core N3050 1.6GHz	4	500GB HDD				

```
1299         Intel HD Graphics 520  Windows 10    1.30   79866.7200
1300             Intel HD Graphics  Windows 10    1.50   12201.1200
1301            AMD Radeon R5 M330  Windows 10    2.19   40705.9200
1302             Intel HD Graphics  Windows 10    2.20   19660.3200

        Operating_system  TouchScreen  IPS  4K_Ultra_HD
0                    Mac            0    1            0
1                    Mac            0    0            0
2     Others/No OS/Linux            0    0            0
3                    Mac            0    1            0
4                    Mac            0    1            0
…                      …          …  …          …
1298             Windows            1    1            0
1299             Windows            1    1            0
1300             Windows            0    0            0
1301             Windows            0    0            0
1302             Windows            0    0            0

[1303 rows x 15 columns]
```

14 10 4K_Ultra_Hd Features

[234]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['4K_Ultra_HD'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['4K_Ultra_HD'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```



[235]:
```
df["Cpu"].value_counts()
```

[235]:
```
Intel Core i5 7200U 2.5GHz              190
Intel Core i7 7700HQ 2.8GHz             146
Intel Core i7 7500U 2.7GHz              134
Intel Core i7 8550U 1.8GHz               73
Intel Core i5 8250U 1.6GHz               72
                                        ...
Intel Core i7 2.9GHz                      1
Intel Core i3 6100U 2.1GHz                1
Intel Pentium Dual Core 4405U 2.1GHz      1
Intel Atom x5-Z8300 1.44GHz               1
Intel Celeron Quad Core N3710 1.6GHz      1
Name: Cpu, Length: 118, dtype: int64
```

[236]:
```
df['CPU_Name']=df['Cpu'].apply(lambda x:" ".join(x.split()[0:3]))
```

[237]:
```
df
```

[237]:
```
      Company            TypeName  Inches  \
0       Apple           Ultrabook    13.3
1       Apple           Ultrabook    13.3
2          HP            Notebook    15.6
3       Apple           Ultrabook    15.4
4       Apple           Ultrabook    13.3
...       ...                 ...     ...
1298   Lenovo  2 in 1 Convertible    14.0
1299   Lenovo  2 in 1 Convertible    13.3
```

```
1300  Lenovo            Notebook    14.0
1301      HP            Notebook    15.6
1302    Asus            Notebook    15.6

                                ScreenResolution  \
0              IPS Panel Retina Display 2560x1600
1                                      1440x900
2                             Full HD 1920x1080
3              IPS Panel Retina Display 2880x1800
4              IPS Panel Retina Display 2560x1600
…                                            …
1298   IPS Panel Full HD / Touchscreen 1920x1080
1299  IPS Panel Quad HD+ / Touchscreen 3200x1800
1300                                   1366x768
1301                                   1366x768
1302                                   1366x768

                                      Cpu  Ram               Memory  \
0                     Intel Core i5 2.3GHz    8            128GB SSD
1                     Intel Core i5 1.8GHz    8  128GB Flash Storage
2                Intel Core i5 7200U 2.5GHz    8            256GB SSD
3                     Intel Core i7 2.7GHz   16            512GB SSD
4                     Intel Core i5 3.1GHz    8            256GB SSD
…                                       …  …                    …
1298             Intel Core i7 6500U 2.5GHz    4            128GB SSD
1299             Intel Core i7 6500U 2.5GHz   16            512GB SSD
1300  Intel Celeron Dual Core N3050 1.6GHz    2   64GB Flash Storage
1301             Intel Core i7 6500U 2.5GHz    6              1TB HDD
1302  Intel Celeron Dual Core N3050 1.6GHz    4            500GB HDD

                               Gpu       OpSys  Weight        Price  \
0     Intel Iris Plus Graphics 640       macOS    1.37   71378.6832
1           Intel HD Graphics 6000       macOS    1.34   47895.5232
2            Intel HD Graphics 620       No OS    1.86   30636.0000
3               AMD Radeon Pro 455       macOS    1.83  135195.3360
4     Intel Iris Plus Graphics 650       macOS    1.37   96095.8080
…                              …         …     …          …
1298         Intel HD Graphics 520  Windows 10    1.80   33992.6400
1299         Intel HD Graphics 520  Windows 10    1.30   79866.7200
1300             Intel HD Graphics  Windows 10    1.50   12201.1200
1301            AMD Radeon R5 M330  Windows 10    2.19   40705.9200
1302             Intel HD Graphics  Windows 10    2.20   19660.3200

        Operating_system  TouchScreen  IPS  4K_Ultra_HD            CPU_Name
0                    Mac            0    1            0       Intel Core i5
1                    Mac            0    0            0       Intel Core i5
2     Others/No OS/Linux            0    0            0       Intel Core i5
```

```
3                    Mac            0    1            0       Intel Core i7
4                    Mac            0    1            0       Intel Core i5
…                    …          …   …          …               …
1298             Windows            1    1            0       Intel Core i7
1299             Windows            1    1            0       Intel Core i7
1300             Windows            0    0            0  Intel Celeron Dual
1301             Windows            0    0            0       Intel Core i7
1302             Windows            0    0            0  Intel Celeron Dual

[1303 rows x 16 columns]
```

[238]:
```
def fetch_processor(text):
    if text== "Intel Core i7" or text=='Intel Core i5' or text=="Intel Core i3":
        return text
    else:
        if text.split()[0]=='Intel':
            return 'Intel Other Processor'
        else:
            return 'AMD Processor'
```

[239]:
```
df['CPU_BRAND']=df['CPU_Name'].apply(fetch_processor)
```

[240]:
```
df
```

[240]:
```
     Company            TypeName  Inches  \
0      Apple           Ultrabook    13.3
1      Apple           Ultrabook    13.3
2         HP            Notebook    15.6
3      Apple           Ultrabook    15.4
4      Apple           Ultrabook    13.3
…      …                     …     …
1298  Lenovo  2 in 1 Convertible    14.0
1299  Lenovo  2 in 1 Convertible    13.3
1300  Lenovo            Notebook    14.0
1301      HP            Notebook    15.6
1302    Asus            Notebook    15.6

                               ScreenResolution  \
0             IPS Panel Retina Display 2560x1600
1                                       1440x900
2                              Full HD 1920x1080
3             IPS Panel Retina Display 2880x1800
4             IPS Panel Retina Display 2560x1600
…                                             …
1298   IPS Panel Full HD / Touchscreen 1920x1080
1299  IPS Panel Quad HD+ / Touchscreen 3200x1800
1300                                    1366x768
```

```
1301                                  1366x768
1302                                  1366x768

                                        Cpu  Ram               Memory  \
0                       Intel Core i5 2.3GHz    8            128GB SSD
1                       Intel Core i5 1.8GHz    8  128GB Flash Storage
2                  Intel Core i5 7200U 2.5GHz    8            256GB SSD
3                       Intel Core i7 2.7GHz   16            512GB SSD
4                       Intel Core i5 3.1GHz    8            256GB SSD
…                                       …  …                    …
1298               Intel Core i7 6500U 2.5GHz    4            128GB SSD
1299               Intel Core i7 6500U 2.5GHz   16            512GB SSD
1300  Intel Celeron Dual Core N3050 1.6GHz    2   64GB Flash Storage
1301               Intel Core i7 6500U 2.5GHz    6              1TB HDD
1302  Intel Celeron Dual Core N3050 1.6GHz    4            500GB HDD

                               Gpu       OpSys  Weight        Price  \
0     Intel Iris Plus Graphics 640       macOS    1.37   71378.6832
1           Intel HD Graphics 6000       macOS    1.34   47895.5232
2            Intel HD Graphics 620       No OS    1.86   30636.0000
3               AMD Radeon Pro 455       macOS    1.83  135195.3360
4     Intel Iris Plus Graphics 650       macOS    1.37   96095.8080
…                              …         …     …           …
1298         Intel HD Graphics 520  Windows 10    1.80   33992.6400
1299         Intel HD Graphics 520  Windows 10    1.30   79866.7200
1300             Intel HD Graphics  Windows 10    1.50   12201.1200
1301            AMD Radeon R5 M330  Windows 10    2.19   40705.9200
1302             Intel HD Graphics  Windows 10    2.20   19660.3200

        Operating_system  TouchScreen  IPS  4K_Ultra_HD            CPU_Name  \
0                    Mac            0    1            0       Intel Core i5
1                    Mac            0    0            0       Intel Core i5
2     Others/No OS/Linux            0    0            0       Intel Core i5
3                    Mac            0    1            0       Intel Core i7
4                    Mac            0    1            0       Intel Core i5
…                      …          …  …          …                       …
1298             Windows            1    1            0       Intel Core i7
1299             Windows            1    1            0       Intel Core i7
1300             Windows            0    0            0  Intel Celeron Dual
1301             Windows            0    0            0       Intel Core i7
1302             Windows            0    0            0  Intel Celeron Dual

                CPU_BRAND
0             Intel Core i5
1             Intel Core i5
2             Intel Core i5
3             Intel Core i7
```

```
4             Intel Core i5
…                       …
1298          Intel Core i7
1299          Intel Core i7
1300  Intel Other Processor
1301          Intel Core i7
1302  Intel Other Processor

[1303 rows x 17 columns]
```

[241]:
```
df.drop(columns=['Cpu','ScreenResolution','CPU_Name'],inplace=True)
```

[242]:
```
df
```

[242]:
```
      Company            TypeName  Inches  Ram               Memory  \
0       Apple           Ultrabook    13.3    8            128GB SSD
1       Apple           Ultrabook    13.3    8  128GB Flash Storage
2          HP            Notebook    15.6    8            256GB SSD
3       Apple           Ultrabook    15.4   16            512GB SSD
4       Apple           Ultrabook    13.3    8            256GB SSD
…       …                     …      …  …                    …
1298  Lenovo  2 in 1 Convertible    14.0    4            128GB SSD
1299  Lenovo  2 in 1 Convertible    13.3   16            512GB SSD
1300  Lenovo            Notebook    14.0    2   64GB Flash Storage
1301      HP            Notebook    15.6    6              1TB HDD
1302    Asus            Notebook    15.6    4            500GB HDD

                              Gpu       OpSys  Weight        Price  \
0     Intel Iris Plus Graphics 640       macOS    1.37   71378.6832
1           Intel HD Graphics 6000       macOS    1.34   47895.5232
2            Intel HD Graphics 620       No OS    1.86   30636.0000
3               AMD Radeon Pro 455       macOS    1.83  135195.3360
4     Intel Iris Plus Graphics 650       macOS    1.37   96095.8080
…                              …         …      …           …
1298         Intel HD Graphics 520  Windows 10    1.80   33992.6400
1299         Intel HD Graphics 520  Windows 10    1.30   79866.7200
1300             Intel HD Graphics  Windows 10    1.50   12201.1200
1301            AMD Radeon R5 M330  Windows 10    2.19   40705.9200
1302             Intel HD Graphics  Windows 10    2.20   19660.3200

        Operating_system  TouchScreen  IPS  4K_Ultra_HD             CPU_BRAND
0                    Mac            0    1            0          Intel Core i5
1                    Mac            0    0            0          Intel Core i5
2     Others/No OS/Linux            0    0            0          Intel Core i5
3                    Mac            0    1            0          Intel Core i7
4                    Mac            0    1            0          Intel Core i5
…                    …          …  …          …                     …
```

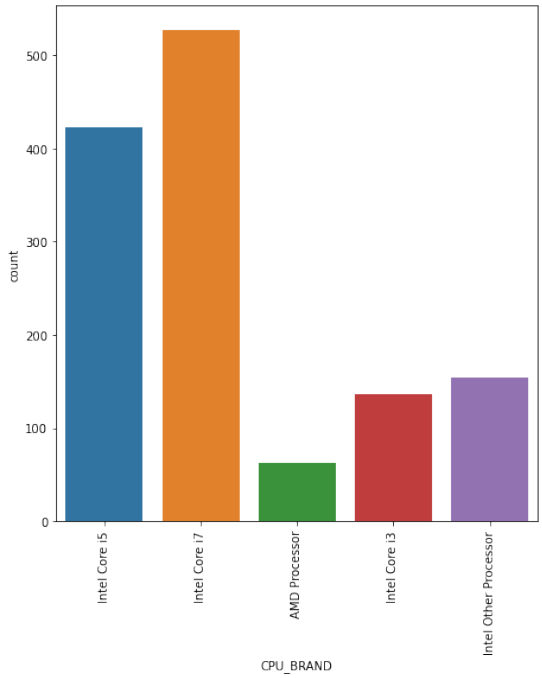
```
1298            Windows            1    1           0          Intel Core i7
1299            Windows            1    1           0          Intel Core i7
1300            Windows            0    0           0  Intel Other Processor
1301            Windows            0    0           0          Intel Core i7
1302            Windows            0    0           0  Intel Other Processor

[1303 rows x 14 columns]
```

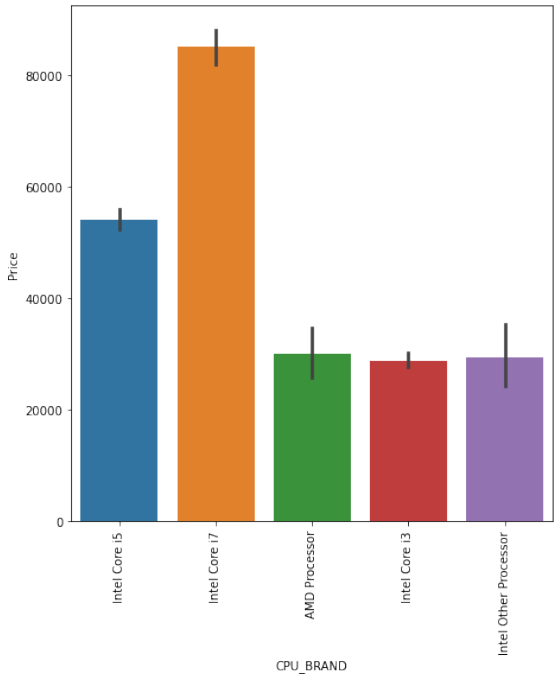
[243]:
```
df['CPU_BRAND'].value_counts()
```

[243]:
```
Intel Core i7            527
Intel Core i5            423
Intel Other Processor    154
Intel Core i3            136
AMD Processor             63
Name: CPU_BRAND, dtype: int64
```

15 11 CPU_BRAND feature

[244]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['CPU_BRAND'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['CPU_BRAND'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```





[245]:
```
df['Memory'].value_counts()
```

[245]:
```
256GB SSD                412
1TB HDD                  223
500GB HDD                132
512GB SSD                118
128GB SSD +  1TB HDD      94
128GB SSD                 76
256GB SSD +  1TB HDD      73
32GB Flash Storage        38
2TB HDD                   16
64GB Flash Storage        15
512GB SSD +  1TB HDD      14
1TB SSD                   14
256GB SSD +  2TB HDD      10
1.0TB Hybrid               9
256GB Flash Storage        8
16GB Flash Storage         7
32GB SSD                   6
180GB SSD                  5
128GB Flash Storage        4
512GB SSD +  2TB HDD       3
16GB SSD                   3
128GB SSD +  2TB HDD       2
1TB SSD +  1TB HDD         2
```

```
512GB Flash Storage                 2
256GB SSD +  500GB HDD              2
256GB SSD +  256GB SSD              2
64GB Flash Storage +  1TB HDD       1
8GB SSD                             1
512GB SSD +  256GB SSD              1
64GB SSD                            1
256GB SSD +  1.0TB Hybrid           1
128GB HDD                           1
1TB HDD +  1TB HDD                  1
240GB SSD                           1
508GB Hybrid                        1
1.0TB HDD                           1
32GB HDD                            1
512GB SSD +  1.0TB Hybrid           1
512GB SSD +  512GB SSD              1
Name: Memory, dtype: int64
```

[246]:
```
df['Memory'] = df['Memory'].astype(str).replace('\.0', '', regex=True)
df["Memory"] = df["Memory"].str.replace('GB', '')
df["Memory"] = df["Memory"].str.replace('TB', '000')
new = df["Memory"].str.split("+", n = 1, expand = True)

df["first"]= new[0]
df["first"]=df["first"].str.strip()

df["second"]= new[1]

df["Layer1HDD"] = df["first"].apply(lambda x: 1 if "HDD" in x else 0)
df["Layer1SSD"] = df["first"].apply(lambda x: 1 if "SSD" in x else 0)
df["Layer1Hybrid"] = df["first"].apply(lambda x: 1 if "Hybrid" in x else 0)
df["Layer1Flash_Storage"] = df["first"].apply(lambda x: 1 if "Flash Storage" in␣
 ↪x else 0)

df['first'] = df['first'].str.replace(r'\D', '')

df["second"].fillna("0", inplace = True)

df["Layer2HDD"] = df["second"].apply(lambda x: 1 if "HDD" in x else 0)
df["Layer2SSD"] = df["second"].apply(lambda x: 1 if "SSD" in x else 0)
df["Layer2Hybrid"] = df["second"].apply(lambda x: 1 if "Hybrid" in x else 0)
df["Layer2Flash_Storage"] = df["second"].apply(lambda x: 1 if "Flash Storage"␣
 ↪in x else 0)

df['second'] = df['second'].str.replace(r'\D', '')

df["first"] = df["first"].astype(int)
```

```
df["second"] = df["second"].astype(int)

df["HDD"]=(df["first"]*df["Layer1HDD"]+df["second"]*df["Layer2HDD"])
df["SSD"]=(df["first"]*df["Layer1SSD"]+df["second"]*df["Layer2SSD"])
df["Hybrid"]=(df["first"]*df["Layer1Hybrid"]+df["second"]*df["Layer2Hybrid"])
df["Flash_Storage"]=(df["first"]*df["Layer1Flash_Storage"]+df["second"]*df["Layer2Flash_Storag

df.drop(columns=['first', 'second', 'Layer1HDD', 'Layer1SSD', 'Layer1Hybrid',
       'Layer1Flash_Storage', 'Layer2HDD', 'Layer2SSD', 'Layer2Hybrid',
       'Layer2Flash_Storage'],inplace=True)
```

[247]:
```
df
```

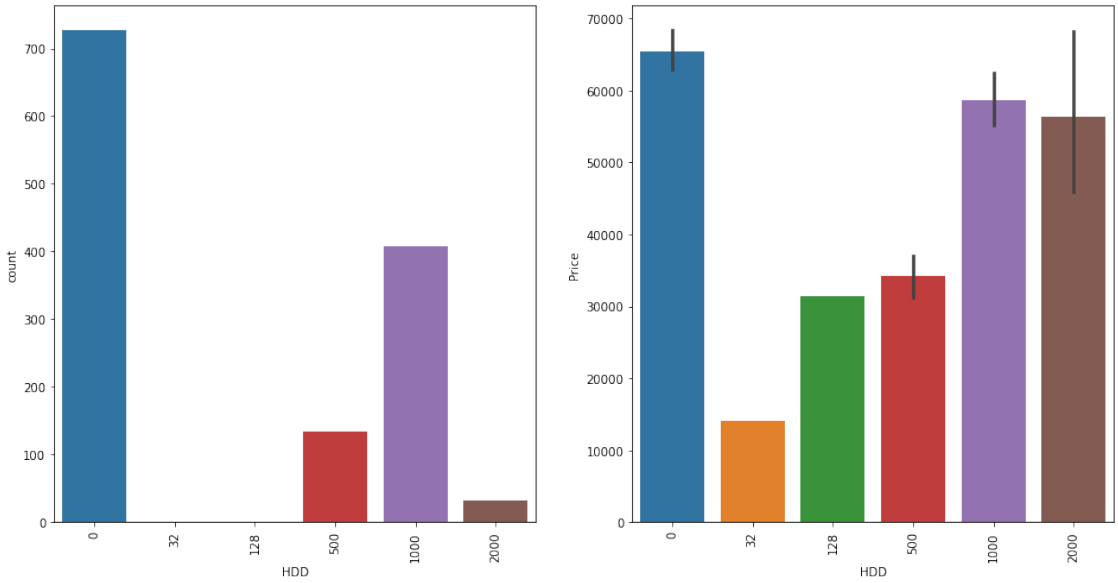
[247]:

	Company	TypeName	Inches	Ram	Memory	Gpu	OpSys	Weight	Price	Operating_system	TouchScreen	IPS	4K_Ultra_HD
0	Apple	Ultrabook	13.3	8	128 SSD	Intel Iris Plus Graphics 640	macOS	1.37	71378.6832	Mac	0	1	0
1	Apple	Ultrabook	13.3	8	128 Flash Storage	Intel HD Graphics 6000	macOS	1.34	47895.5232	Mac	0	0	0
2	HP	Notebook	15.6	8	256 SSD	Intel HD Graphics 620	No OS	1.86	30636.0000	Others/No OS/Linux	0	0	0
3	Apple	Ultrabook	15.4	16	512 SSD	AMD Radeon Pro 455	macOS	1.83	135195.3360	Mac	0	1	0
4	Apple	Ultrabook	13.3	8	256 SSD	Intel Iris Plus Graphics 650	macOS	1.37	96095.8080	Mac	0	1	0
…	…	…	…	…	…	…	…	…	…	…	…	…	…
1298	Lenovo	2 in 1 Convertible	14.0	4	128 SSD	Intel HD Graphics 520	Windows 10	1.80	33992.6400	Windows	1	1	0
1299	Lenovo	2 in 1 Convertible	13.3	16	512 SSD	Intel HD Graphics 520	Windows 10	1.30	79866.7200				
1300	Lenovo	Notebook	14.0	2	64 Flash Storage	Intel HD Graphics	Windows 10	1.50	12201.1200				
1301	HP	Notebook	15.6	6	1000 HDD	AMD Radeon R5 M330	Windows 10	2.19	40705.9200				
1302	Asus	Notebook	15.6	4	500 HDD	Intel HD Graphics	Windows 10	2.20	19660.3200				

```
1299             Windows            1    1            0
1300             Windows            0    0            0
1301             Windows            0    0            0
1302             Windows            0    0            0

                  CPU_BRAND   HDD  SSD  Hybrid  Flash_Storage
0             Intel Core i5     0  128       0              0
1             Intel Core i5     0    0       0            128
2             Intel Core i5     0  256       0              0
3             Intel Core i7     0  512       0              0
4             Intel Core i5     0  256       0              0
…                       …   …  …     …            …
1298          Intel Core i7     0  128       0              0
1299          Intel Core i7     0  512       0              0
1300  Intel Other Processor     0    0       0             64
1301          Intel Core i7  1000    0       0              0
1302  Intel Other Processor   500    0       0              0

[1303 rows x 18 columns]
```

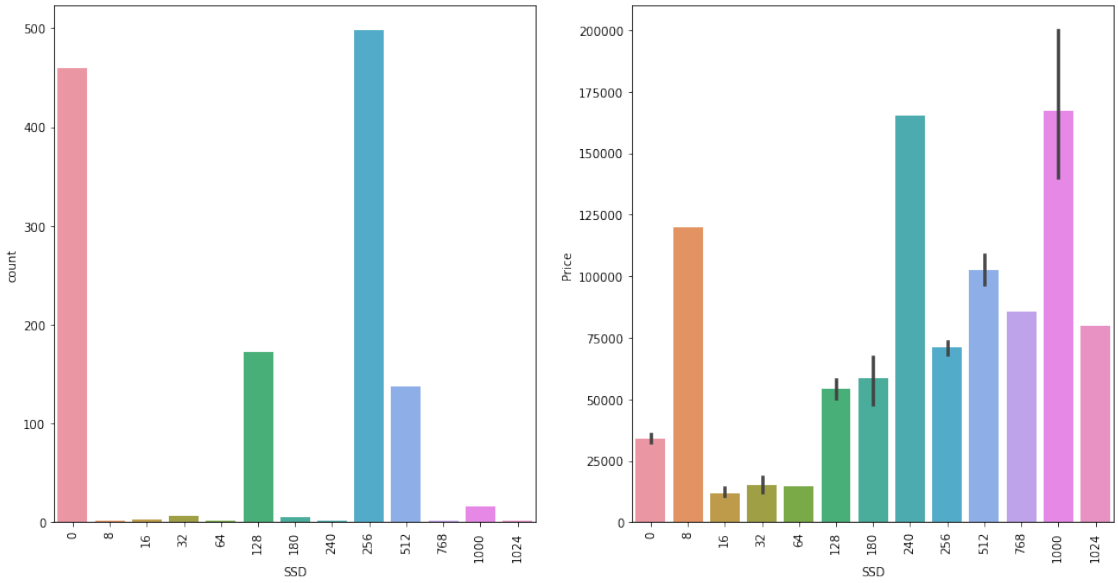
16 12 HDD Feature

[248]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['HDD'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['HDD'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```



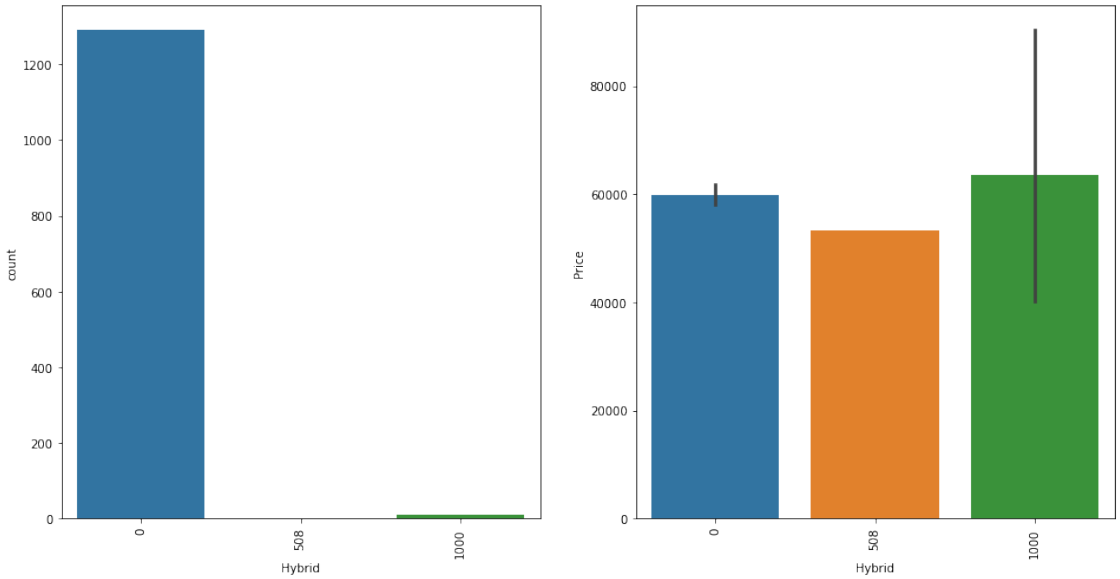
17 13 SSD Feature

[249]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['SSD'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['SSD'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```



18 14 Hybrid Feature

[250]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['Hybrid'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['Hybrid'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```



[251]:
```
df.drop(columns=['Memory'],inplace=True)
```

[252]:
```
df
```

[252]:

	Company	TypeName	Inches	Ram	Gpu	OpSys	Weight	Price	Operating_system	TouchScreen	IPS
0	Apple	Ultrabook	13.3	8	Intel Iris Plus Graphics 640	macOS	1.37	71378.6832	Mac	0	1
1	Apple	Ultrabook	13.3	8	Intel HD Graphics 6000	macOS	1.34	47895.5232	Mac	0	0
2	HP	Notebook	15.6	8	Intel HD Graphics 620	No OS	1.86	30636.0000	Others/No OS/Linux	0	0
3	Apple	Ultrabook	15.4	16	AMD Radeon Pro 455	macOS	1.83	135195.3360	Mac	0	1
4	Apple	Ultrabook	13.3	8	Intel Iris Plus Graphics 650	macOS	1.37	96095.8080	Mac	0	1
…	…	…	…	…	…	…	…	…	…	…	…
1298	Lenovo	2 in 1 Convertible	14.0	4	Intel HD Graphics 520	Windows 10	1.80	33992.6400	Windows	1	1
1299	Lenovo	2 in 1 Convertible	13.3	16	Intel HD Graphics 520	Windows 10	1.30	79866.7200	Windows	1	1
1300	Lenovo	Notebook	14.0	2	Intel HD Graphics	Windows 10	1.50	12201.1200	Windows	0	0
1301	HP	Notebook	15.6	6	AMD Radeon R5 M330	Windows 10	2.19	40705.9200	Windows	0	0
1302	Asus	Notebook	15.6	4	Intel HD Graphics						

```
1302  Windows 10    2.20   19660.3200             Windows            0    0

      4K_Ultra_HD              CPU_BRAND   HDD  SSD  Hybrid  Flash_Storage
0               0          Intel Core i5     0  128       0              0
1               0          Intel Core i5     0    0       0            128
2               0          Intel Core i5     0  256       0              0
3               0          Intel Core i7     0  512       0              0
4               0          Intel Core i5     0  256       0              0
…             …                      …   …  …      …             …
1298            0          Intel Core i7     0  128       0              0
1299            0          Intel Core i7     0  512       0              0
1300            0  Intel Other Processor     0    0       0             64
1301            0          Intel Core i7  1000    0       0              0
1302            0  Intel Other Processor   500    0       0              0

[1303 rows x 17 columns]
```

[253]:
```
df['Gpu'].value_counts()
```

[253]:
```
Intel HD Graphics 620      281
Intel HD Graphics 520      185
Intel UHD Graphics 620      68
Nvidia GeForce GTX 1050     66
Nvidia GeForce GTX 1060     48
                          …
Nvidia GeForce GTX 980       1
Nvidia GeForce GTX1080       1
Nvidia GeForce 960M          1
AMD FirePro W6150M           1
AMD Radeon R5 M315           1
Name: Gpu, Length: 110, dtype: int64
```

[254]:
```
df['GPU_BRAND']=df['Gpu'].apply(lambda x:x.split()[0])
```

[255]:
```
df
```

[255]:
```
     Company            TypeName  Inches  Ram                           Gpu  \
0      Apple           Ultrabook    13.3    8  Intel Iris Plus Graphics 640
1      Apple           Ultrabook    13.3    8        Intel HD Graphics 6000
2         HP            Notebook    15.6    8         Intel HD Graphics 620
3      Apple           Ultrabook    15.4   16            AMD Radeon Pro 455
4      Apple           Ultrabook    13.3    8  Intel Iris Plus Graphics 650
…      …                     …     …  …                             …
1298  Lenovo  2 in 1 Convertible    14.0    4         Intel HD Graphics 520
1299  Lenovo  2 in 1 Convertible    13.3   16         Intel HD Graphics 520
1300  Lenovo            Notebook    14.0    2             Intel HD Graphics
1301      HP            Notebook    15.6    6            AMD Radeon R5 M330
```

```
1302    Asus            Notebook    15.6    4             Intel HD Graphics

           OpSys  Weight        Price    Operating_system  TouchScreen  IPS  \
0          macOS    1.37   71378.6832                 Mac            0    1
1          macOS    1.34   47895.5232                 Mac            0    0
2          No OS    1.86   30636.0000  Others/No OS/Linux            0    0
3          macOS    1.83  135195.3360                 Mac            0    1
4          macOS    1.37   96095.8080                 Mac            0    1
…          …     …          …                  …          …  …
1298  Windows 10    1.80   33992.6400             Windows            1    1
1299  Windows 10    1.30   79866.7200             Windows            1    1
1300  Windows 10    1.50   12201.1200             Windows            0    0
1301  Windows 10    2.19   40705.9200             Windows            0    0
1302  Windows 10    2.20   19660.3200             Windows            0    0

      4K_Ultra_HD              CPU_BRAND   HDD  SSD  Hybrid  Flash_Storage  \
0               0          Intel Core i5     0  128       0              0
1               0          Intel Core i5     0    0       0            128
2               0          Intel Core i5     0  256       0              0
3               0          Intel Core i7     0  512       0              0
4               0          Intel Core i5     0  256       0              0
…           …                      …   …  …     …            …
1298            0          Intel Core i7     0  128       0              0
1299            0          Intel Core i7     0  512       0              0
1300            0  Intel Other Processor     0    0       0             64
1301            0          Intel Core i7  1000    0       0              0
1302            0  Intel Other Processor   500    0       0              0

     GPU_BRAND
0        Intel
1        Intel
2        Intel
3          AMD
4        Intel
…        …
1298     Intel
1299     Intel
1300     Intel
1301       AMD
1302     Intel

[1303 rows x 18 columns]
```

[256]:
```
df['GPU_BRAND'].value_counts()
```

[256]:
```
Intel     722
Nvidia    400
```

```
AMD       180
ARM         1
Name: GPU_BRAND, dtype: int64
```

[257]:
```
df = df[df['GPU_BRAND'] != 'ARM']
```

[258]:
```
df.head()
```

[258]:

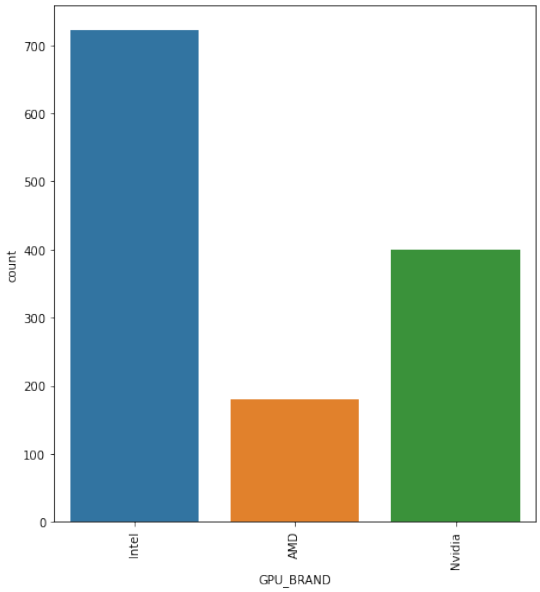
	Company	TypeName	Inches	Ram	Gpu	OpSys	\
0	Apple	Ultrabook	13.3	8	Intel Iris Plus Graphics 640	macOS	
1	Apple	Ultrabook	13.3	8	Intel HD Graphics 6000	macOS	
2	HP	Notebook	15.6	8	Intel HD Graphics 620	No OS	
3	Apple	Ultrabook	15.4	16	AMD Radeon Pro 455	macOS	
4	Apple	Ultrabook	13.3	8	Intel Iris Plus Graphics 650	macOS	

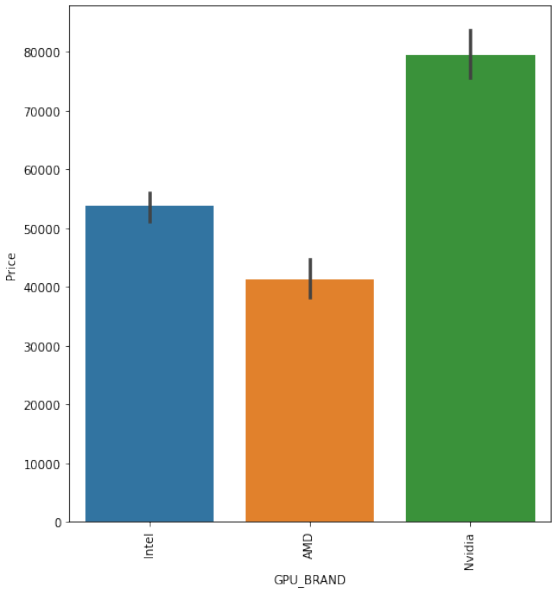
	Weight	Price	Operating_system	TouchScreen	IPS	4K_Ultra_HD	\
0	1.37	71378.6832	Mac	0	1	0	
1	1.34	47895.5232	Mac	0	0	0	
2	1.86	30636.0000	Others/No OS/Linux	0	0	0	
3	1.83	135195.3360	Mac	0	1	0	
4	1.37	96095.8080	Mac	0	1	0	

	CPU_BRAND	HDD	SSD	Hybrid	Flash_Storage	GPU_BRAND
0	Intel Core i5	0	128	0	0	Intel
1	Intel Core i5	0	0	0	128	Intel
2	Intel Core i5	0	256	0	0	Intel
3	Intel Core i7	0	512	0	0	AMD
4	Intel Core i5	0	256	0	0	Intel

19 15 GPU_BRAND Feature

[259]:
```
plt.figure(figsize=(16,8))
plt.subplot(1,2,1)
sns.countplot(df['GPU_BRAND'])
plt.xticks(rotation=90)


plt.subplot(1,2,2)
sns.barplot(df['GPU_BRAND'],df['Price'],data=df)
plt.xticks(rotation=90)
plt.show()
```





[260]:
```
df
```

[260]:

	Company	TypeName	Inches	Ram	Gpu	\
0	Apple	Ultrabook	13.3	8	Intel Iris Plus Graphics 640	
1	Apple	Ultrabook	13.3	8	Intel HD Graphics 6000	
2	HP	Notebook	15.6	8	Intel HD Graphics 620	
3	Apple	Ultrabook	15.4	16	AMD Radeon Pro 455	
4	Apple	Ultrabook	13.3	8	Intel Iris Plus Graphics 650	
…	…	…	…	…	…	
1298	Lenovo	2 in 1 Convertible	14.0	4	Intel HD Graphics 520	
1299	Lenovo	2 in 1 Convertible	13.3	16	Intel HD Graphics 520	
1300	Lenovo	Notebook	14.0	2	Intel HD Graphics	
1301	HP	Notebook	15.6	6	AMD Radeon R5 M330	
1302	Asus	Notebook	15.6	4	Intel HD Graphics	

	OpSys	Weight	Price	Operating_system	TouchScreen	IPS	\
0	macOS	1.37	71378.6832	Mac	0	1	
1	macOS	1.34	47895.5232	Mac	0	0	
2	No OS	1.86	30636.0000	Others/No OS/Linux	0	0	
3	macOS	1.83	135195.3360	Mac	0	1	
4	macOS	1.37	96095.8080	Mac	0	1	
…	…	…	…	…	…	…	
1298	Windows 10	1.80	33992.6400	Windows	1	1	
1299	Windows 10	1.30	79866.7200	Windows	1	1	
1300	Windows 10	1.50	12201.1200	Windows	0	0	
1301	Windows 10	2.19	40705.9200	Windows	0	0	
1302	Windows 10	2.20	19660.3200	Windows	0	0	

```
      4K_Ultra_HD              CPU_BRAND   HDD  SSD  Hybrid  Flash_Storage  \
0               0          Intel Core i5     0  128       0              0
1               0          Intel Core i5     0    0       0            128
2               0          Intel Core i5     0  256       0              0
3               0          Intel Core i7     0  512       0              0
4               0          Intel Core i5     0  256       0              0
…             …                      …   …  …     …             …
1298            0          Intel Core i7     0  128       0              0
1299            0          Intel Core i7     0  512       0              0
1300            0  Intel Other Processor     0    0       0             64
1301            0          Intel Core i7  1000    0       0              0
1302            0  Intel Other Processor   500    0       0              0

     GPU_BRAND
0        Intel
1        Intel
2        Intel
3          AMD
4        Intel
…        …
1298     Intel
1299     Intel
1300     Intel
1301       AMD
1302     Intel

[1302 rows x 18 columns]
```

[261]:
```
df.drop(columns=['Gpu'],inplace=True)
```

[262]:
```
df
```

[262]:
```
      Company            TypeName  Inches  Ram       OpSys  Weight  \
0       Apple           Ultrabook    13.3    8       macOS    1.37
1       Apple           Ultrabook    13.3    8       macOS    1.34
2          HP            Notebook    15.6    8       No OS    1.86
3       Apple           Ultrabook    15.4   16       macOS    1.83
4       Apple           Ultrabook    13.3    8       macOS    1.37
…       …                     …     …  …          …     …
1298  Lenovo  2 in 1 Convertible    14.0    4  Windows 10    1.80
1299  Lenovo  2 in 1 Convertible    13.3   16  Windows 10    1.30
1300  Lenovo            Notebook    14.0    2  Windows 10    1.50
1301      HP            Notebook    15.6    6  Windows 10    2.19
1302    Asus            Notebook    15.6    4  Windows 10    2.20

             Price    Operating_system  TouchScreen  IPS  4K_Ultra_HD  \
```

```
0      71378.6832                  Mac            0    1            0
1      47895.5232                  Mac            0    0            0
2      30636.0000  Others/No OS/Linux             0    0            0
3     135195.3360                  Mac            0    1            0
4      96095.8080                  Mac            0    1            0
…           …                     …          …   …          …
1298   33992.6400              Windows            1    1            0
1299   79866.7200              Windows            1    1            0
1300   12201.1200              Windows            0    0            0
1301   40705.9200              Windows            0    0            0
1302   19660.3200              Windows            0    0            0

                    CPU_BRAND   HDD  SSD  Hybrid  Flash_Storage GPU_BRAND
0             Intel Core i5     0  128       0              0     Intel
1             Intel Core i5     0    0       0            128     Intel
2             Intel Core i5     0  256       0              0     Intel
3             Intel Core i7     0  512       0              0       AMD
4             Intel Core i5     0  256       0              0     Intel
…                       …   …  …     …           …        …
1298          Intel Core i7     0  128       0              0     Intel
1299          Intel Core i7     0  512       0              0     Intel
1300  Intel Other Processor     0    0       0             64     Intel
1301          Intel Core i7  1000    0       0              0       AMD
1302  Intel Other Processor   500    0       0              0     Intel

[1302 rows x 17 columns]
```

[263]:
```
df['TypeName'].value_counts()
```

[263]:
```
Notebook              727
Gaming                205
Ultrabook             196
2 in 1 Convertible    120
Workstation            29
Netbook                25
Name: TypeName, dtype: int64
```

[264]:
```
df.drop(columns=['OpSys'],inplace=True)
```

20 Apply One hot Encoding on Categorical Columns For ML model

[267]:
```
df= pd.
↪get_dummies(columns=['Company','TypeName','Operating_system','CPU_BRAND','GPU_BRAND'],data=
```

[268]:
```
df
```

[268]:

	Inches	Ram	Weight	Price	TouchScreen	IPS	4K_Ultra_HD	HDD
0	13.3	8	1.37	71378.6832	0	1	0	0
1	13.3	8	1.34	47895.5232	0	0	0	0
2	15.6	8	1.86	30636.0000	0	0	0	0
3	15.4	16	1.83	135195.3360	0	1	0	0
4	13.3	8	1.37	96095.8080	0	1	0	0
…	…	…	…	…	…	…	…	…
1298	14.0	4	1.80	33992.6400	1	1	0	0
1299	13.3	16	1.30	79866.7200	1	1	0	0
1300	14.0	2	1.50	12201.1200	0	0	0	0
1301	15.6	6	2.19	40705.9200	0	0	0	1000
1302	15.6	4	2.20	19660.3200	0	0	0	500

	SSD	Hybrid	…	Operating_system_Others/No OS/Linux
0	128	0	…	0
1	0	0	…	0
2	256	0	…	1
3	512	0	…	0
4	256	0	…	0
…	…	…	…	…
1298	128	0	…	0
1299	512	0	…	0
1300	0	0	…	0
1301	0	0	…	0
1302	0	0	…	0

	Operating_system_Windows	CPU_BRAND_AMD Processor
0	0	0
1	0	0
2	0	0
3	0	0
4	0	0
…	…	…
1298	1	0
1299	1	0
1300	1	0
1301	1	0
1302	1	0

	CPU_BRAND_Intel Core i3	CPU_BRAND_Intel Core i5
0	0	1
1	0	1
2	0	1
3	0	0
4	0	1
…	…	…
1298	0	0

```
1299                         0                         0
1300                         0                         0
1301                         0                         0
1302                         0                         0

      CPU_BRAND_Intel Core i7  CPU_BRAND_Intel Other Processor  GPU_BRAND_AMD  \
0                            0                                 0              0
1                            0                                 0              0
2                            0                                 0              0
3                            1                                 0              1
4                            0                                 0              0
…                          …                                 …            …
1298                         1                                 0              0
1299                         1                                 0              0
1300                         0                                 1              0
1301                         1                                 0              1
1302                         0                                 1              0

      GPU_BRAND_Intel  GPU_BRAND_Nvidia
0                   1                 0
1                   1                 0
2                   1                 0
3                   0                 0
4                   1                 0
…                 …                 …
1298                1                 0
1299                1                 0
1300                1                 0
1301                0                 0
1302                1                 0

[1302 rows x 47 columns]
```