

```
In [1]: # Import major Libraries

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

```
In [2]: # Import dataset from local system

df = pd.read_csv('house pricing delhi- MagicBricks.csv')
```

```
In [3]: df.head()
```

```
Out[3]:
```

	Area	BHK	Bathroom	Furnishing	Location	District	Locality	Parking	Status	Transa
0	950.0	2	2.0	Furnished	Karol Bagh	Central Delhi	DDA MIG Flats Prasad Nagar Phase 2, Prasad Nag...	1.0	Ready_to_move	F
1	535.0	2	1.0	Furnished	Karol Bagh	Central Delhi	Dev Nagar, Karol Bagh	1.0	Ready_to_move	New_Prc
2	1280.0	3	3.0	Furnished	Karol Bagh	Central Delhi	Karol Bagh	2.0	Ready_to_move	F
3	1135.0	3	3.0	Furnished	Karol Bagh	Central Delhi	The Amaryllis, Karol Bagh	2.0	Almost_ready	F
4	1135.0	3	3.0	Furnished	Karol Bagh	Central Delhi	The Amaryllis, Karol Bagh	2.0	Almost_ready	F

```
In [4]: df.tail()
```

Out[4]:

	Area	BHK	Bathroom	Furnishing	Location	District	Locality	Parking	Status	Tra
1209	4688.0	5	4.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	New_
1210	3901.0	4	4.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	New_
1211	2306.0	3	3.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	New_
1212	675.0	2	1.0	Unfurnished	Uttam Nagar	West Delhi	Uttam Nagar	1.0	Ready_to_move	
1213	585.0	2	2.0	Unfurnished	Uttam Nagar	West Delhi	Uttam Nagar	1.0	Ready_to_move	

In [5]: `df.shape`

Out[5]: (1214, 13)

In [6]: `# Data Discription and Data Information``df.describe() # It will work with only numeric values or Continious Values`

Out[6]:

	Area	BHK	Bathroom	Parking	Per_Sqft	Price
count	1214.000000	1214.000000	1212.000000	1182.000000	973.000000	1.214000e+03
mean	1451.850751	2.778418	2.523927	1.708122	15574.885920	2.079898e+07
std	1586.472855	0.946811	1.017723	5.717177	21574.389007	2.561308e+07
min	28.000000	1.000000	1.000000	1.000000	1259.000000	1.000000e+06
25%	800.000000	2.000000	2.000000	1.000000	6154.000000	5.600000e+06
50%	1150.000000	3.000000	2.000000	1.000000	10838.000000	1.400000e+07
75%	1620.000000	3.000000	3.000000	2.000000	17647.000000	2.500000e+07
max	24300.000000	10.000000	7.000000	114.000000	183333.000000	2.400000e+08

In [7]: `# Information regarding to the Dataset.``df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1214 entries, 0 to 1213
Data columns (total 13 columns):
#   Column          Non-Null Count  Dtype
---  ---
0   Area            1214 non-null  float64
1   BHK             1214 non-null  int64
2   Bathroom        1212 non-null  float64
3   Furnishing      1214 non-null  object
4   Location        1214 non-null  object
5   District        1214 non-null  object
6   Locality        1214 non-null  object
7   Parking         1182 non-null  float64
8   Status          1214 non-null  object
9   Transaction     1214 non-null  object
10  Type            1209 non-null  object
11  Per_Sqft        973 non-null   float64
12  Price           1214 non-null  int64
dtypes: float64(4), int64(2), object(7)
memory usage: 123.4+ KB
```

In [8]: *# Checking Missing Values in a Dataset.*

```
df.isnull().sum()
```

Out[8]:

Area	0
BHK	0
Bathroom	2
Furnishing	0
Location	0
District	0
Locality	0
Parking	32
Status	0
Transaction	0
Type	5
Per_Sqft	241
Price	0

dtype: int64

In [9]: *# Filling Null Values with appropriate data.*

```
df.Bathroom = df.Bathroom.fillna(df.Bathroom.median())
df.Parking = df.Parking.fillna(df.Parking.median())
df.Type = df.Type.fillna('Unknown')
df.Per_Sqft = df.Per_Sqft.fillna(df.Per_Sqft.median())
```

In [10]: *# Checking Null Values Again*

```
df.isnull().sum()
```

```
Out[10]: Area          0
        BHK           0
        Bathroom      0
        Furnishing    0
        Location      0
        District      0
        Locality      0
        Parking       0
        Status        0
        Transaction   0
        Type          0
        Per_Sqft      0
        Price         0
        dtype: int64
```

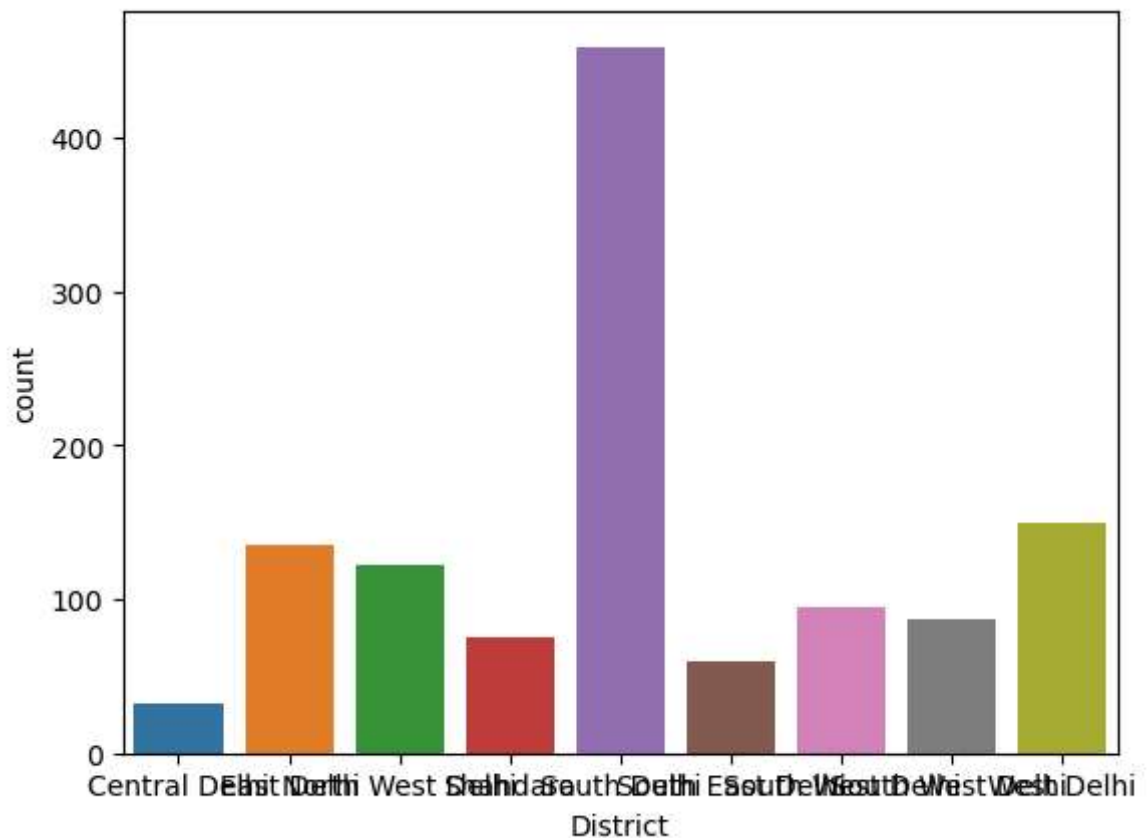
```
In [11]: # Saving a clean file in local system.
```

```
df.to_csv('magic_bricks_cleandata.csv')
```

```
In [12]: # Using Countplot with District Column
```

```
sns.countplot(df.District.sort_values())
```

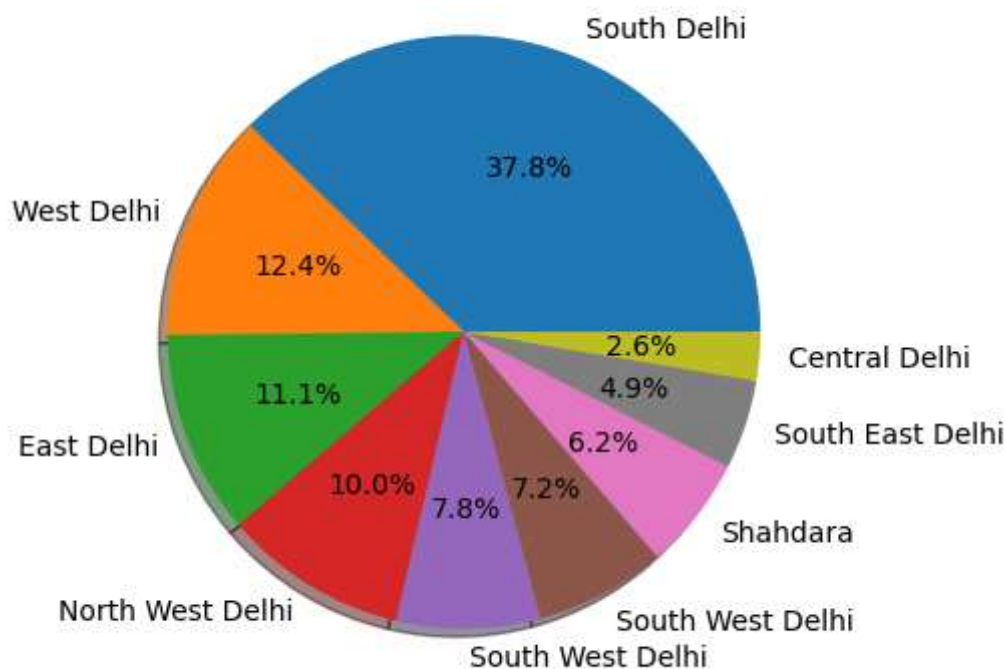
```
Out[12]: <AxesSubplot:xlabel='District', ylabel='count'>
```



```
In [13]: # Pie Chart with Distric to see the % of Data.
```

```
plt.pie(df.District.value_counts(), labels=df.District.value_counts().index, autopct='%1
```

```
Out[13]: ([<matplotlib.patches.Wedge at 0x2433f114430>,
<matplotlib.patches.Wedge at 0x2433f114e20>,
<matplotlib.patches.Wedge at 0x2433f1247f0>,
<matplotlib.patches.Wedge at 0x2433f1331c0>,
<matplotlib.patches.Wedge at 0x2433f133b50>,
<matplotlib.patches.Wedge at 0x2433f140520>,
<matplotlib.patches.Wedge at 0x2433f140eb0>,
<matplotlib.patches.Wedge at 0x2433f14e880>,
<matplotlib.patches.Wedge at 0x2433f15b250>],
[Text(0.411070020131639, 1.0203045812643272, 'South Delhi'),
Text(-1.0224184282933728, 0.40578388026893014, 'West Delhi'),
Text(-1.0296009597182176, -0.3871974480124132, 'East Delhi'),
Text(-0.5712251398707956, -0.9400541684283892, 'North West Delhi'),
Text(0.017078746572777605, -1.0998674085613698, 'South West Delhi'),
Text(0.5142950307378011, -0.9723685625103294, 'South West Delhi '),
Text(0.8655640918347755, -0.6788216282104164, 'Shahdara'),
Text(1.0447457644052096, -0.3442474222959039, 'South East Delhi'),
Text(1.0962305570630642, -0.09098662407851114, 'Central Delhi')],
[Text(0.22422001098089397, 0.5565297715987239, '37.8%'),
Text(-0.5576827790691125, 0.22133666196487095, '12.4%'),
Text(-0.5616005234826641, -0.2111986080067708, '11.1%'),
Text(-0.31157734902043394, -0.5127568191427576, '10.0%'),
Text(0.009315679948787785, -0.5999276773971107, '7.8%'),
Text(0.2805245622206188, -0.5303828522783615, '7.2%'),
Text(0.47212586827351377, -0.37026634266022707, '6.2%'),
Text(0.5698613260392053, -0.1877713212523112, '4.9%'),
Text(0.5979439402162167, -0.04962906767918789, '2.6%')])
```



```
In [14]: df.District.value_counts()
```

```
Out[14]: South Delhi      459
          West Delhi      150
          East Delhi       135
          North West Delhi 122
          South West Delhi  95
          South West Delhi  87
          Shahdara          75
          South East Delhi  59
          Central Delhi     32
          Name: District, dtype: int64
```

```
In [15]: df
```

Out[15]:

	Area	BHK	Bathroom	Furnishing	Location	District	Locality	Parking	Status	Tr
0	950.0	2	2.0	Furnished	Karol Bagh	Central Delhi	DDA MIG Flats Prasad Nagar Phase 2, Prasad Nag...	1.0	Ready_to_move	
1	535.0	2	1.0	Furnished	Karol Bagh	Central Delhi	Dev Nagar, Karol Bagh	1.0	Ready_to_move	Nev
2	1280.0	3	3.0	Furnished	Karol Bagh	Central Delhi	Karol Bagh	2.0	Ready_to_move	
3	1135.0	3	3.0	Furnished	Karol Bagh	Central Delhi	The Amaryllis, Karol Bagh	2.0	Almost_ready	
4	1135.0	3	3.0	Furnished	Karol Bagh	Central Delhi	The Amaryllis, Karol Bagh	2.0	Almost_ready	
...	
1209	4688.0	5	4.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1210	3901.0	4	4.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1211	2306.0	3	3.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1212	675.0	2	1.0	Unfurnished	Uttam Nagar	West Delhi	Uttam Nagar	1.0	Ready_to_move	
1213	585.0	2	2.0	Unfurnished	Uttam Nagar	West Delhi	Uttam Nagar	1.0	Ready_to_move	

1214 rows × 13 columns

```
In [16]: # Compare Prices of Karol Bagh and Patel Nagar homes.  
df[df.Location=='Karol Bagh'].Price
```

```
Out[16]: 0      12500000
          1      3900000
          2     15000000
          3     25000000
          4     25000000
          5     25000000
          6     24500000
         176     6540000
         177     6800000
         178     7000000
         179    10000000
         180    20000000
         181     5500000
         182    14500000
         183    24000000
         184    25500000
         185    18500000
         186    39700000
         187    22000000
         188    25500000
         189    17000000
         190    27500000
         191    25500000
         192    20000000
         193    20900000
         194    43100000
        853    16000000
        854    16000000
        855     4100000
        856     5000000
        857    73700000
        858    52300000
Name: Price, dtype: int64
```

```
In [17]: df[df.Location=='Patel Nagar'].Price
```



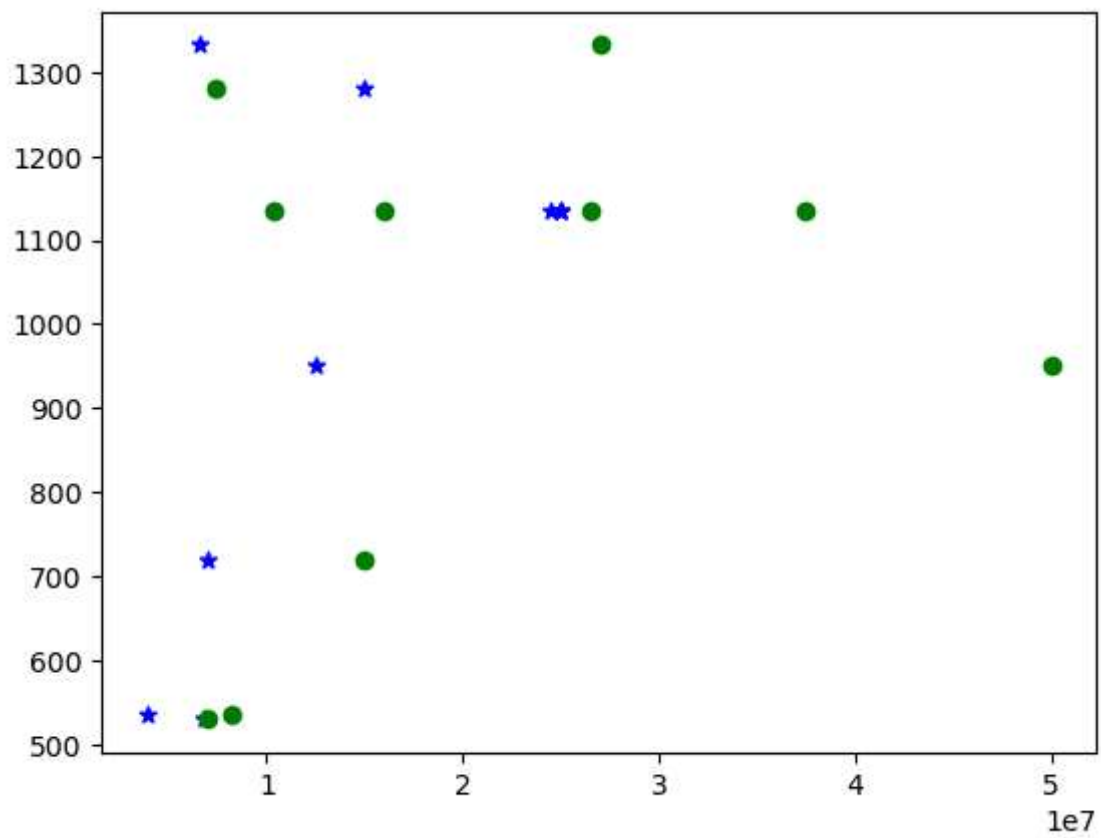
```
Out[17]: 165      50000000
166      82000000
167      74000000
168     104000000
808     375000000
809     265000000
810     160000000
811     270000000
812      70000000
813     150000000
814      95000000
815     200000000
816     150000000
817      80000000
818     195000000
819      59000000
820     395000000
839     553000000
1187      80000000
1188     170000000
1189     211000000
1190     120000000
1191     240000000
1192     125000000
1201     120000000
1203    1350000000
1204     930000000
1205     570000000
1206    1350000000
1207     930000000
1208     570000000
1209    1350000000
1210     930000000
1211     570000000
```

Name: Price, dtype: int64

```
In [18]: plt.scatter(x = df[df.Location=='Karol Bagh'].Price.head(10),y = df[df.Location=='Karol Bagh'].Price.head(10),
                    color='b',marker='*')

plt.scatter(x = df[df.Location=='Patel Nagar'].Price.head(10),y = df[df.Location=='Patel Nagar'].Price.head(10),
            color='g',marker='o')
```

```
Out[18]: <matplotlib.collections.PathCollection at 0x2433f1919a0>
```



In [19]: df

Out[19]:

	Area	BHK	Bathroom	Furnishing	Location	District	Locality	Parking	Status	Tr
--	------	-----	----------	------------	----------	----------	----------	---------	--------	----

0	950.0	2	2.0	Furnished	Karol Bagh	Central Delhi	DDA MIG Flats Prasad Nagar Phase 2, Prasad Nag...	1.0	Ready_to_move	
1	535.0	2	1.0	Furnished	Karol Bagh	Central Delhi	Dev Nagar, Karol Bagh	1.0	Ready_to_move	Nev
2	1280.0	3	3.0	Furnished	Karol Bagh	Central Delhi	Karol Bagh	2.0	Ready_to_move	
3	1135.0	3	3.0	Furnished	Karol Bagh	Central Delhi	The Amaryllis, Karol Bagh	2.0	Almost_ready	
4	1135.0	3	3.0	Furnished	Karol Bagh	Central Delhi	The Amaryllis, Karol Bagh	2.0	Almost_ready	
...	
1209	4688.0	5	4.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1210	3901.0	4	4.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1211	2306.0	3	3.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1212	675.0	2	1.0	Unfurnished	Uttam Nagar	West Delhi	Uttam Nagar	1.0	Ready_to_move	
1213	585.0	2	2.0	Unfurnished	Uttam Nagar	West Delhi	Uttam Nagar	1.0	Ready_to_move	

1214 rows × 13 columns



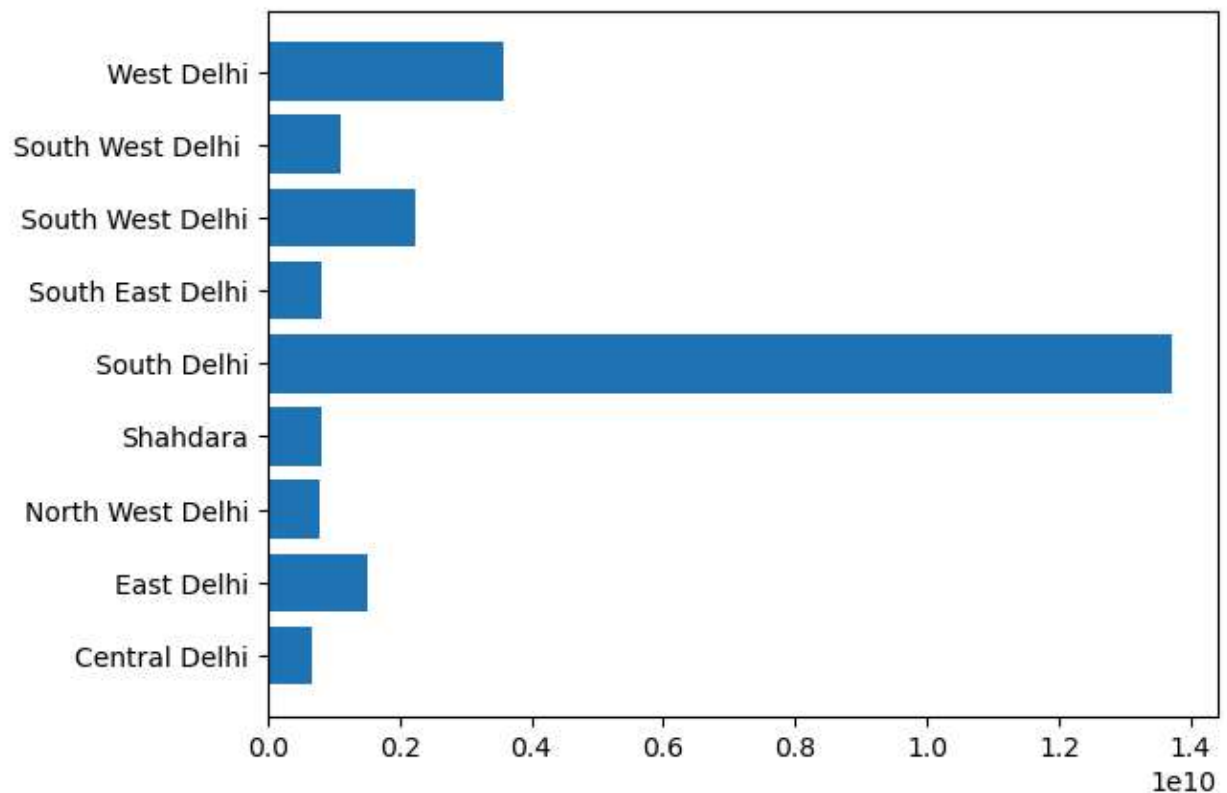
In [20]:

```
df.groupby('District')['Price'].sum()
```

```
Out[20]: District
Central Delhi      677540000
East Delhi         1511130000
North West Delhi   795550000
Shahdara           827720000
South Delhi        13717840000
South East Delhi   813500000
South West Delhi   2226100000
South West Delhi   1100370000
West Delhi         3580210000
Name: Price, dtype: int64
```

```
In [21]: plt.barh(df.groupby('District')['Price'].sum().index, df.groupby('District')['Price'].s
```

```
Out[21]: <BarContainer object of 9 artists>
```



```
In [22]: df[df.Location=='Karol Bagh'].Price.mean()
```

```
Out[22]: 21173125.0
```

```
In [23]: df
```

Out[23]:

	Area	BHK	Bathroom	Furnishing	Location	District	Locality	Parking	Status	Tr
0	950.0	2	2.0	Furnished	Karol Bagh	Central Delhi	DDA MIG Flats Prasad Nagar Phase 2, Prasad Nag...	1.0	Ready_to_move	
1	535.0	2	1.0	Furnished	Karol Bagh	Central Delhi	Dev Nagar, Karol Bagh	1.0	Ready_to_move	Nev
2	1280.0	3	3.0	Furnished	Karol Bagh	Central Delhi	Karol Bagh	2.0	Ready_to_move	
3	1135.0	3	3.0	Furnished	Karol Bagh	Central Delhi	The Amaryllis, Karol Bagh	2.0	Almost_ready	
4	1135.0	3	3.0	Furnished	Karol Bagh	Central Delhi	The Amaryllis, Karol Bagh	2.0	Almost_ready	
...	
1209	4688.0	5	4.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1210	3901.0	4	4.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1211	2306.0	3	3.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1212	675.0	2	1.0	Unfurnished	Uttam Nagar	West Delhi	Uttam Nagar	1.0	Ready_to_move	
1213	585.0	2	2.0	Unfurnished	Uttam Nagar	West Delhi	Uttam Nagar	1.0	Ready_to_move	

1214 rows × 13 columns



In [24]: df[(df.Location=='Karo1 Bagh') & (df.Parking==3)].Price.mean()

Out[24]: 73700000.0

```
In [25]: plt.title('Heatmap')
sns.heatmap(df[['Area', 'BHK', 'Bathroom', 'Parking', 'Per_Sqft', 'Price']].corr(),annot=Tr
```

```
Out[25]: <AxesSubplot:title={'center':'Heatmap'}>
```



```
In [26]: df[(df.District=='West Delhi') & (df.Bathroom==5)].Price.mean()
```

```
Out[26]: 50050000.0
```

```
In [27]: df
```

Out[27]:

	Area	BHK	Bathroom	Furnishing	Location	District	Locality	Parking	Status	Tr
0	950.0	2	2.0	Furnished	Karol Bagh	Central Delhi	DDA MIG Flats Prasad Nagar Phase 2, Prasad Nag...	1.0	Ready_to_move	
1	535.0	2	1.0	Furnished	Karol Bagh	Central Delhi	Dev Nagar, Karol Bagh	1.0	Ready_to_move	Nev
2	1280.0	3	3.0	Furnished	Karol Bagh	Central Delhi	Karol Bagh	2.0	Ready_to_move	
3	1135.0	3	3.0	Furnished	Karol Bagh	Central Delhi	The Amaryllis, Karol Bagh	2.0	Almost_ready	
4	1135.0	3	3.0	Furnished	Karol Bagh	Central Delhi	The Amaryllis, Karol Bagh	2.0	Almost_ready	
...	
1209	4688.0	5	4.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1210	3901.0	4	4.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1211	2306.0	3	3.0	Unfurnished	Patel Nagar	West Delhi	The Leela Sky Villas, Patel Nagar	1.0	Almost_ready	Nev
1212	675.0	2	1.0	Unfurnished	Uttam Nagar	West Delhi	Uttam Nagar	1.0	Ready_to_move	
1213	585.0	2	2.0	Unfurnished	Uttam Nagar	West Delhi	Uttam Nagar	1.0	Ready_to_move	

1214 rows × 13 columns

```
In [28]: print(df[df.Furnishing=='Unfurnished'].Price.mean())
print(df[df.Furnishing=='Furnished'].Price.mean())
print(df[df.Furnishing=='Semi-Furnished'].Price.mean())
```

20309556.7867036
17171363.636363637
22003028.064992614

```
In [29]: df.Furnishing.value_counts()
```

```
Out[29]: Semi-Furnished    677  
Unfurnished      361  
Furnished        176  
Name: Furnishing, dtype: int64
```

```
In [ ]:
```