In [69]:

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn.linear_model import LinearRegression
```

In [70]:

```python
df=pd.read_csv('Bengaluru_house_data.csv')
df.head()
```

Out[70]:

	area_type	availability	location	size	society	total_sqft	bath	balcony	price
0	Super built-up Area	19-Dec	Electronic City Phase II	2 BHK	Coomee	1056	2.0	1.0	39.07
1	Plot Area	Ready To Move	Chikka Tirupathi	4 Bedroom	Theanmp	2600	5.0	3.0	120.00
2	Built-up Area	Ready To Move	Uttarahalli	3 BHK	NaN	1440	2.0	3.0	62.00
3	Super built-up Area	Ready To Move	Lingadheeranahalli	3 BHK	Soiewre	1521	3.0	1.0	95.00
4	Super built-up Area	Ready To Move	Kothanur	2 BHK	NaN	1200	2.0	1.0	51.00

In [71]:

```python
df.shape
```

Out[71]:

```
(13320, 9)
```

In [72]:

```python
df1=df.drop(['area_type','availability','society','balcony'], axis='columns')
df1.head()
```

Out[72]:

	location	size	total_sqft	bath	price
0	Electronic City Phase II	2 BHK	1056	2.0	39.07
1	Chikka Tirupathi	4 Bedroom	2600	5.0	120.00
2	Uttarahalli	3 BHK	1440	2.0	62.00
3	Lingadheeranahalli	3 BHK	1521	3.0	95.00
4	Kothanur	2 BHK	1200	2.0	51.00

In [73]:

```python
df1.isnull().sum()
```

Out[73]:

```
location       1
size          16
total_sqft     0
bath          73
price          0
dtype: int64
```

In [74]:

```python
df2=df1.dropna()
df2.isnull().sum()
```

Out[74]:

```
location      0
size          0
total_sqft    0
bath          0
price         0
dtype: int64
```

In [75]:

```python
df2['size'].unique()
```

Out[75]:

```
array(['2 BHK', '4 Bedroom', '3 BHK', '4 BHK', '6 Bedroom', '3 Bedroom',
       '1 BHK', '1 RK', '1 Bedroom', '8 Bedroom', '2 Bedroom',
       '7 Bedroom', '5 BHK', '7 BHK', '6 BHK', '5 Bedroom', '11 BHK',
       '9 BHK', '9 Bedroom', '27 BHK', '10 Bedroom', '11 Bedroom',
       '10 BHK', '19 BHK', '16 BHK', '43 Bedroom', '14 BHK', '8 BHK',
       '12 Bedroom', '13 BHK', '18 Bedroom'], dtype=object)
```

In [76]:

```python
df2['bhk']=df2['size'].apply(lambda x: int(x.split(' ')[0]))
df2.head()
```

```
c:\users\invest 360\appdata\local\programs\python\python39-32\lib\site-packages\pandas\core\frame.py:3607: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  self._set_item(key, value)
```

Out[76]:

	location	size	total_sqft	bath	price	bhk
0	Electronic City Phase II	2 BHK	1056	2.0	39.07	2
1	Chikka Tirupathi	4 Bedroom	2600	5.0	120.00	4
2	Uttarahalli	3 BHK	1440	2.0	62.00	3
3	Lingadheeranahalli	3 BHK	1521	3.0	95.00	3
4	Kothanur	2 BHK	1200	2.0	51.00	2

In [77]:

```python
df2.total_sqft.unique()
```

Out[77]:

```
array(['1056', '2600', '1440', ..., '1133 - 1384', '774', '4689'],
      dtype=object)
```

In [78]:

```python
df2.total_sqft
```

Out[78]:

```
0        1056
1        2600
2        1440
3        1521
4        1200
         ...
13315    3453
13316    3600
13317    1141
13318    4689
13319     550
Name: total_sqft, Length: 13246, dtype: object
```

In [79]:

```python
def is_float(x):
    try:
        float(x)
    except:
        return False
    return True
```

In [80]:

```python
df2[df2['total_sqft'].apply(is_float)].head(100)
```

Out[80]:

	location	size	total_sqft	bath	price	bhk
0	Electronic City Phase II	2 BHK	1056	2.0	39.07	2
1	Chikka Tirupathi	4 Bedroom	2600	5.0	120.00	4
2	Uttarahalli	3 BHK	1440	2.0	62.00	3
3	Lingadheeranahalli	3 BHK	1521	3.0	95.00	3
4	Kothanur	2 BHK	1200	2.0	51.00	2
...	...	...	...	...	...	...
98	Devasthanagalu	2 BHK	1200	2.0	65.00	2
99	T Dasarahalli	3 Bedroom	1200	3.0	90.00	3
100	Yeshwanthpur	3 BHK	2502	3.0	138.00	3
101	Chandapura	2 BHK	650	1.0	17.00	2
102	Kothanur	3 Bedroom	2400	2.0	150.00	3

100 rows × 6 columns

In [81]:

```python
df2.loc[30]
```

Out[81]:

```
location        Yelahanka
size                4 BHK
total_sqft    2100 - 2850
bath                  4.0
price               186.0
bhk                     4
Name: 30, dtype: object
```

In [82]:

```python
def sqfttonum(x):
    tokens=x.split('-')
    if len(tokens)==2:
        return (float(tokens[0])+float(tokens[1]))/2
    try:
        return float(x)
    except:
        return None
```

In [83]:

```python
sqfttonum('2100-2850')
```

Out[83]:

```
2475.0
```

In [84]:

```python
df3=df2.copy()
df3['total_sqft']=df3['total_sqft'].apply(sqfttonum)
df3.head()
```

Out[84]:

	location	size	total_sqft	bath	price	bhk
0	Electronic City Phase II	2 BHK	1056.0	2.0	39.07	2
1	Chikka Tirupathi	4 Bedroom	2600.0	5.0	120.00	4
2	Uttarahalli	3 BHK	1440.0	2.0	62.00	3
3	Lingadheeranahalli	3 BHK	1521.0	3.0	95.00	3
4	Kothanur	2 BHK	1200.0	2.0	51.00	2

In [85]:

```python
df4=df3.copy()
len(df4.location.unique())
```

Out[85]:

```
1304
```

In [86]:

```python
df4['price_per_sqft']=(df3['price']*100000)/df3.total_sqft
df4.head()
```

Out[86]:

	location	size	total_sqft	bath	price	bhk	price_per_sqft
0	Electronic City Phase II	2 BHK	1056.0	2.0	39.07	2	3699.810606
1	Chikka Tirupathi	4 Bedroom	2600.0	5.0	120.00	4	4615.384615
2	Uttarahalli	3 BHK	1440.0	2.0	62.00	3	4305.555556
3	Lingadheeranahalli	3 BHK	1521.0	3.0	95.00	3	6245.890861
4	Kothanur	2 BHK	1200.0	2.0	51.00	2	4250.000000

In [87]:

```python
df4.location=df4['location'].apply(lambda x: x.strip())
location_stats=df4.groupby('location')['location'].agg('count')
location_stats.head()
```

Out[87]:

```
location
1 Annasandrapalya                                  1
1 Giri Nagar                                       1
1 Immadihalli                                      1
1 Ramamurthy Nagar                                 1
12th cross srinivas nagar banshankari 3rd stage    1
Name: location, dtype: int64
```

In [88]:

```python
len(location_stats[location_stats<=10])
```

Out[88]:

```
1052
```

In [89]:

```python
location_lessthan_10=location_stats[location_stats<=10]
location_lessthan_10
```

Out[89]:

```
location
1 Annasandrapalya                                  1
1 Giri Nagar                                       1
1 Immadihalli                                      1
1 Ramamurthy Nagar                                 1
12th cross srinivas nagar banshankari 3rd stage    1
                                                  ..
t.c palya                                          1
tc.palya                                           4
vinayakanagar                                      1
white field,kadugodi                               1
whitefiled                                         1
Name: location, Length: 1052, dtype: int64
```

In [90]:

```python
df4.location=df4.location.apply(lambda x: 'others' if x in location_lessthan_10 else x)
len(df4.location.unique())
```

Out[90]:

```
242
```

In [91]:

```python
df4.head(10)
```

Out[91]:

	location	size	total_sqft	bath	price	bhk	price_per_sqft
0	Electronic City Phase II	2 BHK	1056.0	2.0	39.07	2	3699.810606
1	Chikka Tirupathi	4 Bedroom	2600.0	5.0	120.00	4	4615.384615
2	Uttarahalli	3 BHK	1440.0	2.0	62.00	3	4305.555556
3	Lingadheeranahalli	3 BHK	1521.0	3.0	95.00	3	6245.890861
4	Kothanur	2 BHK	1200.0	2.0	51.00	2	4250.000000
5	Whitefield	2 BHK	1170.0	2.0	38.00	2	3247.863248
6	Old Airport Road	4 BHK	2732.0	4.0	204.00	4	7467.057101
7	Rajaji Nagar	4 BHK	3300.0	4.0	600.00	4	18181.818182
8	Marathahalli	3 BHK	1310.0	3.0	63.25	3	4828.244275
9	others	6 Bedroom	1020.0	6.0	370.00	6	36274.509804

In [92]:

```python
df4[df4['total_sqft']/df4.bhk<300].head()
```

Out[92]:

	location	size	total_sqft	bath	price	bhk	price_per_sqft
9	others	6 Bedroom	1020.0	6.0	370.0	6	36274.509804
45	HSR Layout	8 Bedroom	600.0	9.0	200.0	8	33333.333333
58	Murugeshpalya	6 Bedroom	1407.0	4.0	150.0	6	10660.980810
68	Devarachikkanahalli	8 Bedroom	1350.0	7.0	85.0	8	6296.296296
70	others	3 Bedroom	500.0	3.0	100.0	3	20000.000000

In [93]:

```python
df5=df4[~(df4['total_sqft']/df4.bhk<300)]
df5.shape
```

Out[93]:

```
(12502, 7)
```

In [94]:

```python
df5.price_per_sqft.describe()
```

Out[94]:

```
count     12456.000000
mean       6308.502826
std        4168.127339
min         267.829813
25%        4210.526316
50%        5294.117647
75%        6916.666667
max      176470.588235
Name: price_per_sqft, dtype: float64
```

In [95]:

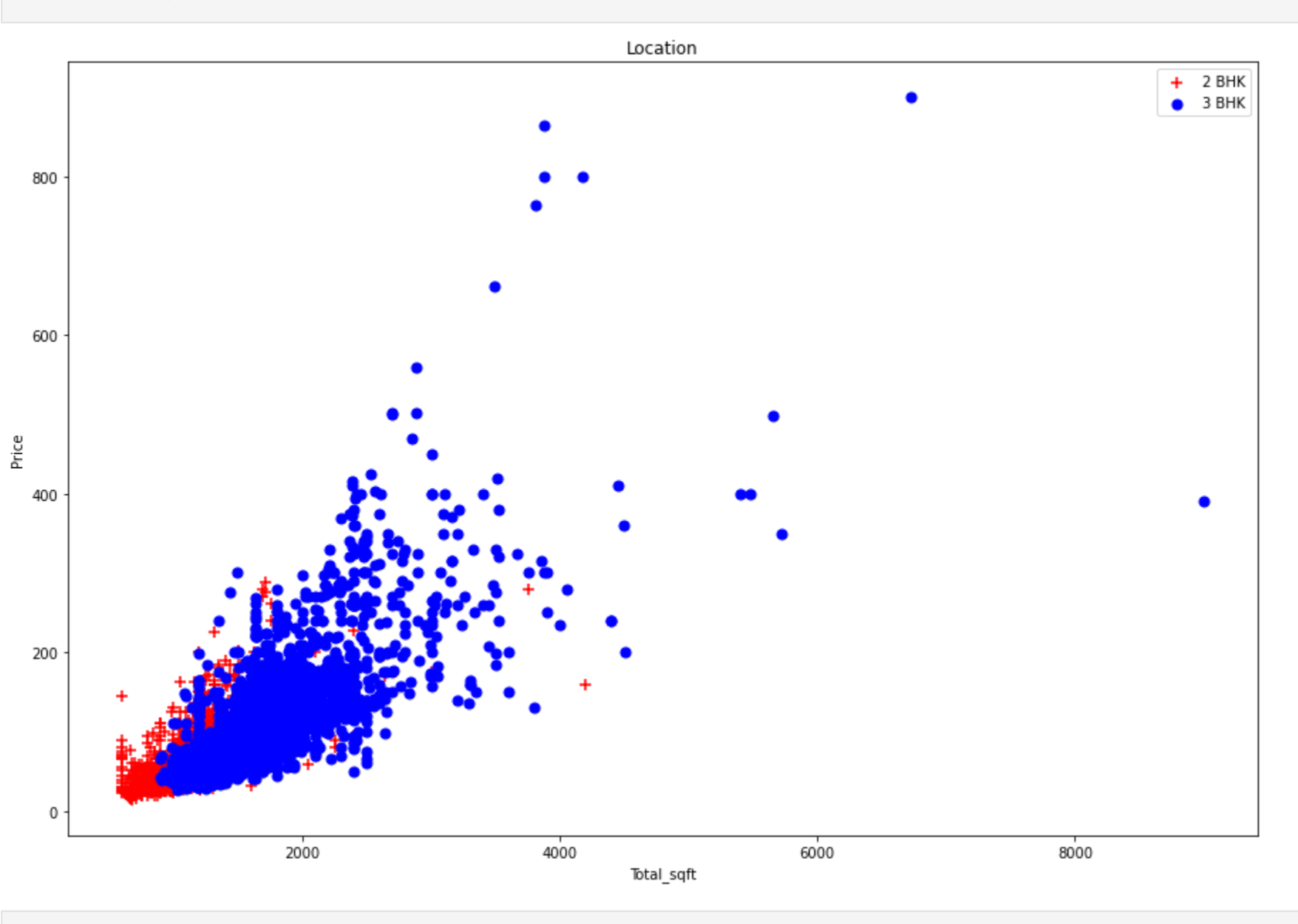
```python
def remove_outliers(df):
    df_out=pd.DataFrame()
    for key,subdf in df.groupby('location'):
        m=np.mean(subdf.price_per_sqft)
        st=np.std(subdf.price_per_sqft)
        reduced_df=subdf[(subdf.price_per_sqft>(m-st))&(subdf.price_per_sqft<=(m+st))]
        df_out=pd.concat([df_out, reduced_df], ignore_index=True)
    return df_out
df6=remove_outliers(df5)
df6.head()
```

Out[95]:

	location	size	total_sqft	bath	price	bhk	price_per_sqft
0	1st Block Jayanagar	4 BHK	2850.0	4.0	428.0	4	15017.543860
1	1st Block Jayanagar	3 BHK	1630.0	3.0	194.0	3	11901.840491
2	1st Block Jayanagar	3 BHK	1875.0	2.0	235.0	3	12533.333333
3	1st Block Jayanagar	3 BHK	1200.0	2.0	130.0	3	10833.333333
4	1st Block Jayanagar	2 BHK	1235.0	2.0	148.0	2	11983.805668

In [96]:

```python
def plot_chart(df,location):
    bhk2=df[(df.location==df.location)&(df.bhk==2)]
    bhk3=df[(df.location==df.location)&(df.bhk==3)]
    plt.rcParams['figure.figsize']=(15,10)
    plt.scatter(bhk2.total_sqft,bhk2.price,color='red',marker='+', s=50, label='2 BHK')
    plt.scatter(bhk3.total_sqft,bhk3.price,color='blue', s=50, label='3 BHK')
    plt.xlabel('Total_sqft')
    plt.ylabel('Price')
    plt.title('Location')
    plt.legend()
plot_chart(df6,'Hebbal')
```



In [97]:

```python
def remo_bhk_outliers(df):
    exclude_indices=np.array([])
    for location,location_df in df.groupby('location'):
        bhk_stats= {}
        for bhk,bhk_df in location_df.groupby('bhk'):
            bhk_stats[bhk]={
                'mean':np.mean(bhk_df.price_per_sqft),
                'std':np.std(bhk_df.price_per_sqft),
                'count':bhk_df.shape[0]
                }
        for bhk,bhk_df in location_df.groupby('bhk'):
                stats=bhk_stats.get(bhk-1)
                if stats and stats['count']>5:
                  exclude_indices=np.append(exclude_indices,bhk_df[bhk_df.price_per_sqft<(stats['mean'])].index.values)
    return df.drop(exclude_indices,axis='index')
```

In [98]:

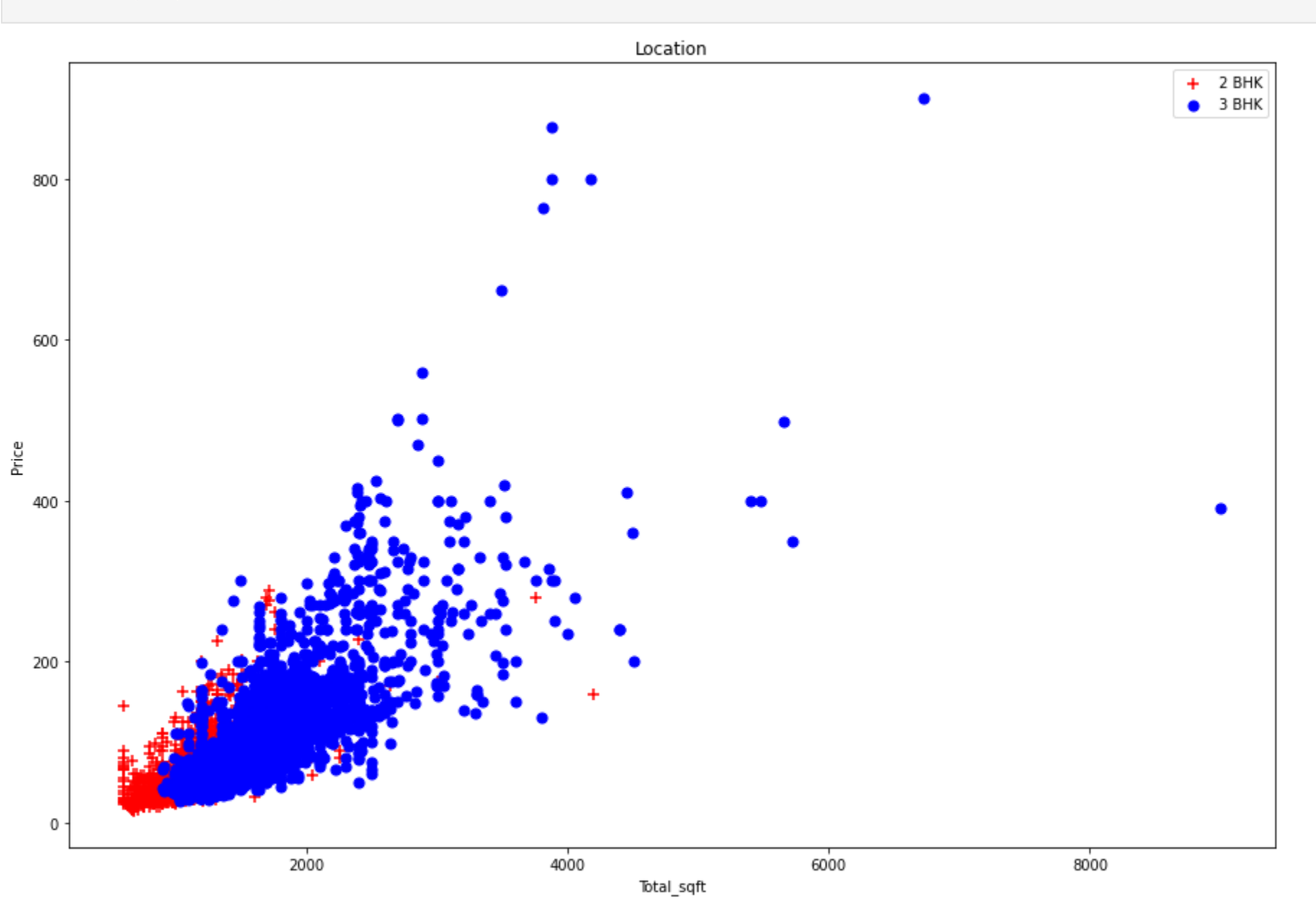
```python
df7=remo_bhk_outliers(df6)
df7.shape
```

Out[98]:

```
(7329, 7)
```

In [99]:

```python
plot_chart(df6,'Hebbal')
```

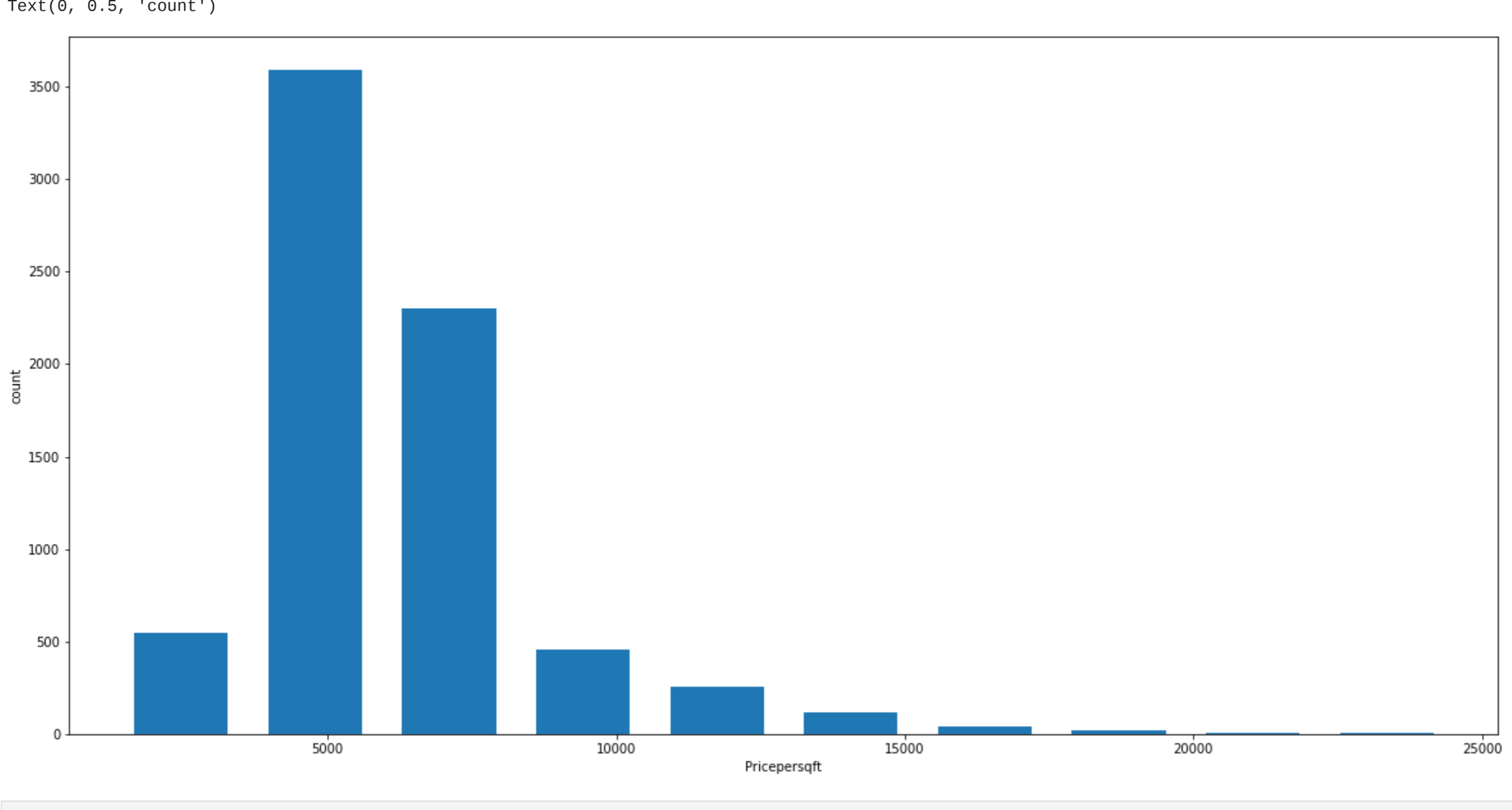


In [100…

```python
import matplotlib
matplotlib.rcParams['figure.figsize']=(20,10)
plt.hist(df7.price_per_sqft,rwidth=0.7)
plt.xlabel('Pricepersqft')
plt.ylabel('count')
```

Out[100…

```
Text(0, 0.5, 'count')
```



In [101…

```python
df7.bath.unique()
```

Out[101…

```
array([ 4.,  3.,  2.,  5.,  8.,  1.,  6.,  7.,  9., 12., 16., 13.])
```

In [102…

```python
df7[df7.bath>10]
```

Out[102…

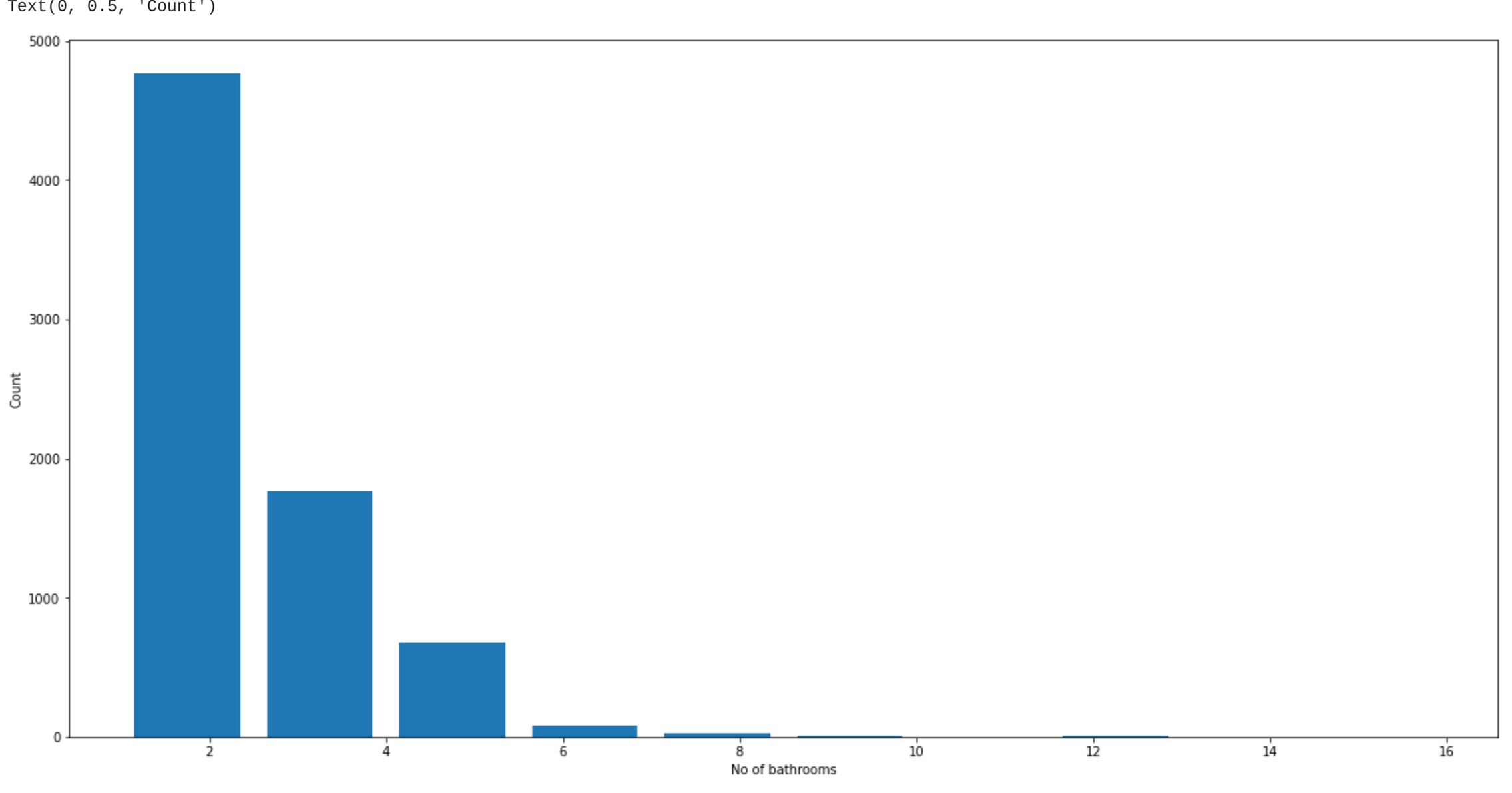
	location	size	total_sqft	bath	price	bhk	price_per_sqft
5277	Neeladri Nagar	10 BHK	4000.0	12.0	160.0	10	4000.000000
8486	others	10 BHK	12000.0	12.0	525.0	10	4375.000000
8575	others	16 BHK	10000.0	16.0	550.0	16	5500.000000
9308	others	11 BHK	6000.0	12.0	150.0	11	2500.000000
9639	others	13 BHK	5425.0	13.0	275.0	13	5069.124424

In [103…

```python
matplotlib.rcParams['figure.figsize']=(20,10)
plt.hist(df7.bath,rwidth=0.8)
plt.xlabel('No of bathrooms')
plt.ylabel('Count')
```

Out[103…

```
Text(0, 0.5, 'Count')
```



In [104…

```python
df7[df7.bath>(df7.bhk+2)]
```

Out[104…

	location	size	total_sqft	bath	price	bhk	price_per_sqft
1626	Chikkabanavar	4 Bedroom	2460.0	7.0	80.0	4	3252.032520
5238	Nagasandra	4 Bedroom	7000.0	8.0	450.0	4	6428.571429
6711	Thanisandra	3 BHK	1806.0	6.0	116.0	3	6423.034330
8411	others	6 BHK	11338.0	9.0	1000.0	6	8819.897689

In [105…

```python
df8=df7[df7.bath<(df7.bhk+2)]
df8.shape
```

Out[105…

```
(7251, 7)
```

In [106…

```python
df9=df8.drop(['size','price_per_sqft'],axis='columns')
df9.head()
```

Out[106…

	location	total_sqft	bath	price	bhk
0	1st Block Jayanagar	2850.0	4.0	428.0	4
1	1st Block Jayanagar	1630.0	3.0	194.0	3
2	1st Block Jayanagar	1875.0	2.0	235.0	3
3	1st Block Jayanagar	1200.0	2.0	130.0	3
4	1st Block Jayanagar	1235.0	2.0	148.0	2

In [107…

```python
dum=pd.get_dummies(df9.location)
dum
```

Out[107…

	1st Block Jayanagar	1st Phase JP Nagar	2nd Phase Judicial Layout	2nd Stage Nagarbhavi	5th Block Hbr Layout	5th Phase JP Nagar	6th Phase JP Nagar	7th Phase JP Nagar	8th Phase JP Nagar	9th Phase JP Nagar	...	Vishveshwarya Layout	Vishwapriya Layout	Vittasandra	Whitefield	Yelachenahalli	Yelahanka	Yelahanka New Town	Yelenahalli	Yeshwanthpur	others
0	1	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
4	1	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
10232	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1
10233	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1
10236	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1
10237	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1
10240	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1

7251 rows × 242 columns

In [108…

```python
df10=pd.concat([df9,dum.drop(['others'],axis='columns')], axis='columns')
df10.head()
```

Out[108…

	location	total_sqft	bath	price	bhk	1st Block Jayanagar	1st Phase JP Nagar	2nd Phase Judicial Layout	2nd Stage Nagarbhavi	5th Block Hbr Layout	...	Vijayanagar	Vishveshwarya Layout	Vishwapriya Layout	Vittasandra	Whitefield	Yelachenahalli	Yelahanka	Yelahanka New Town	Yelenahalli	Yeshwanthpur
0	1st Block Jayanagar	2850.0	4.0	428.0	4	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
1	1st Block Jayanagar	1630.0	3.0	194.0	3	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
2	1st Block Jayanagar	1875.0	2.0	235.0	3	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
3	1st Block Jayanagar	1200.0	2.0	130.0	3	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
4	1st Block Jayanagar	1235.0	2.0	148.0	2	1	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0

5 rows × 246 columns

In [109…

```python
df11=df10.drop(['location'],axis='columns')
df11.head(3)
```

Out[109…

	total_sqft	bath	price	bhk	1st Block Jayanagar	1st Phase JP Nagar	2nd Phase Judicial Layout	2nd Stage Nagarbhavi	5th Block Hbr Layout	5th Phase JP Nagar	...	Vijayanagar	Vishveshwarya Layout	Vishwapriya Layout	Vittasandra	Whitefield	Yelachenahalli	Yelahanka	Yelahanka New Town	Yelenahalli	Yeshwanthpur
0	2850.0	4.0	428.0	4	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
1	1630.0	3.0	194.0	3	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
2	1875.0	2.0	235.0	3	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0

3 rows × 245 columns

In [110…

```python
x=df11.drop(['price'],axis='columns')
x.head(4)
```

Out[110…

	total_sqft	bath	bhk	1st Block Jayanagar	1st Phase JP Nagar	2nd Phase Judicial Layout	2nd Stage Nagarbhavi	5th Block Hbr Layout	5th Phase JP Nagar	6th Phase JP Nagar	...	Vijayanagar	Vishveshwarya Layout	Vishwapriya Layout	Vittasandra	Whitefield	Yelachenahalli	Yelahanka	Yelahanka New Town	Yelenahalli	Yeshwanthpur
0	2850.0	4.0	4	1	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
1	1630.0	3.0	3	1	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
2	1875.0	2.0	3	1	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
3	1200.0	2.0	3	1	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0

4 rows × 244 columns

In [111…

```python
y=df11.price
y.shape[0]
```

Out[111…

```
7251
```

In [112…

```python
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,random_state=10,test_size=0.2)
```

In [113…

```python
df11.loc[6646]
```

Out[113…

```
total_sqft             1564.0
bath                      3.0
price                   105.0
bhk                       3.0
1st Block Jayanagar       0.0
                        ...
Yelachenahalli            0.0
Yelahanka                 0.0
Yelahanka New Town        0.0
Yelenahalli               0.0
Yeshwanthpur              0.0
Name: 6646, Length: 245, dtype: float64
```

In [114…

```python
lr=LinearRegression()
lr.fit(x_train,y_train)
lr.score(x_test,y_test)
```

Out[114…

```
0.8452277697874424
```

In [115…

```python
lr.coef_
```

Out[115…

```
array([ 7.94153739e-02,  5.07902494e+00, -1.77287776e+00,  1.20102684e+02,
        1.60975131e+00, -5.31632191e+01,  1.00744709e+02, -7.09814933e+01,
       -3.92159542e+01, -1.90172851e+01, -1.86570903e+01, -4.78597086e+01,
       -4.58072684e+01, -3.63103183e+01, -5.37188439e+01, -4.32014933e+01,
       -2.83334456e+01, -3.09802727e+01, -3.41349983e+01, -4.35542164e+01,
       -4.68557495e+01, -3.55443843e+01, -5.13413322e+01, -4.41161322e+01,
       -3.39107391e+01, -3.50994582e+01, -1.93206330e+01,  4.33396158e+00,
       -4.16896109e+01, -5.24666080e+01, -2.97715337e+01, -1.63570371e+01,
       -3.29128205e+01,  8.46006640e+01, -3.45699543e+01, -6.21719706e+01,
       -6.17763832e+01, -3.13725549e+01, -3.50075406e+01, -1.45173949e+01,
       -3.25890006e+01,  2.90158013e+01, -1.04900683e+00, -4.98796074e+01,
       -4.60568702e+01, -5.61764461e+01, -3.36588705e+01,  1.18623825e+02,
       -4.57177739e+01, -3.17196120e+01, -2.38864848e+01, -4.14547240e-01,
       -3.88800233e+01, -4.63934835e+01, -4.75920574e+01, -5.95786261e+01,
        3.25208099e+00, -2.09267632e+01, -4.01275019e+01, -3.20205448e+01,
        1.91028758e+01, -4.46079091e+01, -5.14124362e+01, -1.02311555e+02,
       -8.31265732e+01, -3.72810045e+01, -3.89472481e+01,  7.27249068e+01,
        9.94458762e-01,  4.47624334e+02, -2.66816146e+01, -4.04052596e+01,
       -4.02948650e+01, -4.43038796e+01, -4.14854563e+01, -2.08078765e+01,
       -4.30047003e+01, -4.55833710e+01,  1.07168386e+01, -4.96054853e+01,
       -2.58226568e+01, -3.33743079e+01, -5.11395493e+01, -3.51084916e+01,
        4.60973881e+01, -5.32602154e+01, -4.16434984e+01,  1.64923960e+02,
       -4.65426268e+01, -5.08237434e+01, -2.74152403e+01, -4.34883081e+01,
       -4.97030210e+01,  2.33425183e+02, -2.23252408e+01, -4.89670341e+00,
       -4.38638357e+01, -4.45479912e+01, -2.16456746e+01, -1.06640989e+01,
        1.90590524e+01, -2.87500766e+01, -4.70901373e+01, -3.05005432e+01,
       -3.46269695e+01, -4.70901538e+01, -4.94573250e+01, -4.01883681e+01,
       -3.30635317e+01,  2.81539302e+01, -5.13013920e+01, -3.04290466e+01,
       -2.99979381e+01, -4.86615552e+01, -4.64946705e+01, -2.30162538e+01,
        9.93888594e+01, -2.77165238e+01, -2.74209124e+01, -1.74992027e+01,
       -3.05562772e+01, -3.20674756e+01,  2.21809119e+01, -4.95667108e+01,
       -2.76507239e+01, -3.89497647e+01, -4.72755863e+01, -2.54297483e+01,
       -3.41210440e+01, -4.20426080e+01, -3.34676478e+01, -3.14378921e+01,
       -3.08037688e+01, -4.65986241e+01, -4.02464734e+01, -3.30268075e+01,
       -3.77307473e+01, -1.28561120e+01, -2.90027638e+01, -2.95971983e+01,
       -3.11762751e+01, -3.69746853e+01, -3.00638984e+01, -3.47666720e+01,
       -3.83382282e+01, -4.86746508e+01, -4.42698281e+01, -3.54421831e+01,
       -3.65089304e+00,  5.09688572e+01, -4.57987422e+01, -3.65202825e+00,
        5.36671613e+01, -5.04514008e+01, -4.46767965e+01, -4.04873619e+01,
       -3.77704397e+01, -6.10845688e+01, -1.23130697e+01, -3.48517220e+01,
       -1.39461204e+01, -1.93280870e+01, -3.01645870e+01, -4.88781723e+01,
       -4.34286244e+01,  2.21717757e+01, -4.17025712e+01, -4.31762828e+01,
        1.17607806e+02, -3.38825676e+01, -2.80298600e+01,  2.69692063e-01,
       -7.06227931e+01, -4.47353997e+01, -4.55407831e+01, -3.34670732e+01,
       -3.81431105e+01, -6.32217690e+01, -9.74296227e+00, -3.98491299e+01,
       -4.18421946e+01, -7.21167089e+00, -3.56274019e+01, -4.89912631e+01,
       -4.80480379e+01, -1.92567106e+01, -2.19494290e+01, -3.14754214e+01,
       -1.65345670e+01, -4.22128014e+01, -2.24931083e+01, -5.03012608e+01,
       -3.88289885e+01, -4.17080637e+01, -1.79203004e+01,  1.40276963e+00,
       -3.37171040e+01, -5.11325848e+01,  1.37251498e+02, -4.19181368e+01,
       -3.48556238e+01, -3.72129311e+01, -5.46353060e+01, -2.68088696e+01,
       -7.33738686e+00,  7.31156943e+01, -5.13374639e+01, -2.58895483e+01,
       -6.16937371e+01, -2.37749420e+01,  7.91553465e+00, -4.21070187e+01,
       -4.85997870e+01, -2.59869607e+00, -4.44684657e+01, -3.33066588e+01,
       -5.49871400e+01, -3.97346121e+01, -2.96330271e+01, -3.62434907e+01,
       -3.32482632e+01, -3.41924485e+01, -2.77396726e+01, -1.56509864e+01,
       -3.42013036e+01,  4.64181581e+00, -7.50654591e+01, -2.30437401e+01,
        8.88603688e+00, -4.78483731e+01, -4.20519818e+01, -4.57585945e+01,
       -4.77010848e+01, -3.72469830e+01, -1.95922017e+01, -8.16543671e+01,
       -3.65284509e+01, -3.69918205e+01, -2.85309074e+01, -3.07817516e+01,
       -3.52881228e+01, -2.48775394e+01, -5.36226957e+01, -1.25980855e+01])
```

In [116…

```python
lr.intercept_
```

Out[116…

```
-4.138372824959788
```

In [117…

```python
df11.head(1)
```

Out[117…

	total_sqft	bath	price	bhk	1st Block Jayanagar	1st Phase JP Nagar	2nd Phase Judicial Layout	2nd Stage Nagarbhavi	5th Block Hbr Layout	5th Phase JP Nagar	...	Vijayanagar	Vishveshwarya Layout	Vishwapriya Layout	Vittasandra	Whitefield	Yelachenahalli	Yelahanka	Yelahanka New Town	Yelenahalli	Yeshwanthpur
0	2850.0	4.0	428.0	4	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0

1 rows × 245 columns

In [118…

```python
lr.predict(x_train[0:1])
```

Out[118…

```
array([102.24604092])
```

In [119…

```python
df11.loc[6646]
```

Out[119…

```
total_sqft             1564.0
bath                      3.0
price                   105.0
bhk                       3.0
1st Block Jayanagar       0.0
                        ...
Yelachenahalli            0.0
Yelahanka                 0.0
Yelahanka New Town        0.0
Yelenahalli               0.0
Yeshwanthpur              0.0
Name: 6646, Length: 245, dtype: float64
```