

Objective:

The primary objective of this project is to analyze the relationships between various factors influencing real estate prices (e.g., property size, location, number of rooms, etc.) and identify trends and patterns that can help stakeholders make informed decisions.

Understanding Business Problems for Real Estate Data

The goal is to analyze real estate data to identify trends, property price influencers, and investment opportunities. Key questions include: What factors drive property prices? Which areas show high demand? Insights from the data will guide better decision-making for buyers, sellers, and investors.

Problem Statement:

Problem statement analyze real estate data to understand the key factors influencing property prices. By performing Exploratory Data Analysis (EDA) on the dataset, we aim to uncover trends, patterns, and relationships between the target variable, "price," and other features such as location. The insights derived will help in accurate price prediction and informed decision-making for stakeholders.

```
[6]: # lib for extraction ,manipulation,analysis
import numpy as np
import pandas as pd
# for visualization
import matplotlib.pyplot as plt
import seaborn as sns
# for stats
import scipy.stats
from scipy.stats import shapiro, chi2, normaltest, kstest, zscore
# train test split
from sklearn.model_selection import train_test_split
```

```
[7]: # importing dataset
data=pd.read_csv(r"C:\Users\deshm\OneDrive\Desktop\Python\Real_Estate_EDA\Real_Estate_properties.csv")
data
```

	Possession Status	Price	Covered Area	Carpet Area	City	Type of Property	bedroom	Bathroom	Parking	Location	Society	Ownership Type	isPrimeLocationProperty	Maintenance Chars
0	Under Construction	3150000.0	635.0	375.0	Thane	Apartment	1	2.0	1 Covered	Kalyan West, Thane	Y	Freehold		Y
1	Ready to Move	6300000.0	579.0	579.0	Thane	Apartment	2	2.0	1 Covered	Kalyan West, Thane	Y	Freehold		Y
2	Ready to Move	5400000.0	850.0	585.0	Thane	Apartment	2	2.0	1 Open	Kalyan West, Thane	Y	Freehold		Y
3	Under Construction	9000000.0	1050.0	815.0	Thane	Apartment	2	2.0	1 Covered, 1 Open	Kalyan West, Thane	Y	Freehold		320
4	Under Construction	4950000.0	561.0	419.0	Thane	Apartment	1	2.0	NaN	Kalyan West, Thane	Y	Freehold		N
...
12680	Dec '26	22200000.0	NaN	618.0	Mumbai	Apartment	2	2.0	1 Covered	Jogeshwari Vikhroli Link Road, Mumbai	N	NaN		Y
12681	Under Construction	31000000.0	1376.0	1076.0	Mumbai	Apartment	3	3.0	1 Covered	Jogeshwari Vikhroli Link Road, Mumbai	Y	Freehold		1
12682	Under Construction	18500000.0	970.0	618.0	Mumbai	Apartment	2	2.0	1 Covered	Jogeshwari Vikhroli Link Road, Mumbai	Y	Freehold		Y
12683	Ready to Move	22000000.0	NaN	821.0	Mumbai	Apartment	2	2.0	1 Open	Andheri East, Mumbai	Y	Freehold		400
12684	Under Construction	18500000.0	685.0	665.0	Mumbai	Apartment	2	2.0	1 Covered	Andheri East, Mumbai	Y	Freehold		1

12685 rows × 14 columns

```
[8]: #EDA for each data analysis
def eda(data):
    print("Shape:",data.shape)
    print(" "*50)
    print("Size:",data.size)
    print(" "*50)
    print("TNEO:",data.info)
```



```
Covered Area          1311
Carpet Area           2552
City                  0
Type of Property      0
bedroom               0
Bathroom              69
Parking               6159
Location              323
Society                0
Ownership Type        3260
isPrimeLocationProperty 0
Maintenance Charges   6402
dtype: int64
```

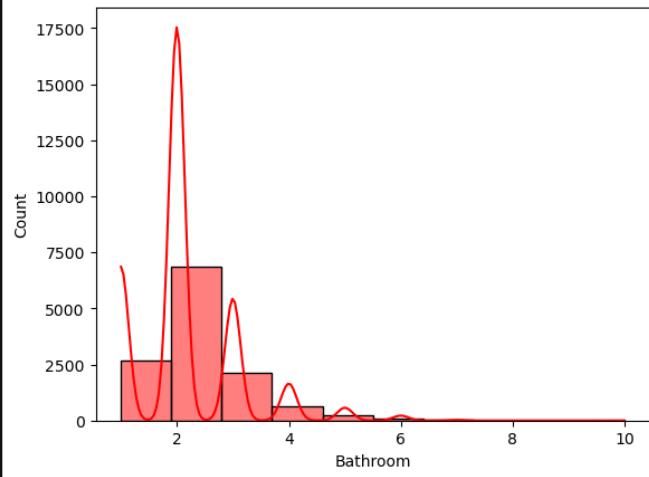
```
[9]: # Analysis for Numerical Columns
def Num_col(data, col):
    mean=data[col].mean()
    median=data[col].median()
    mode=data[col].mode()[0]
    var =data[col].var()
    std=data[col].std()
    skew=data[col].skew()
    Min=data[col].min()
    Max=data[col].max()
    Range=Max-Min
    print("Numerical Columns Analysis:")
    print(f"mean:{mean}\nmedian:{median}\nmode:{mode}\nvar:{var}\nstd:{std}\nskew:{skew}\nMIN:{Min}\nMAX:{Max}\nRange:{Range}")

[10]: Num_col(data,"Bathroom")
```

```
Numerical Columns Analysis:
mean:2.1433893468611287
median:2.0
mode:2.0
var:0.8592634698002165
std:0.9269646540188123
skew:1.3283395953866595
MIN:1.0
MAX:10.0
Range:9.0
```

```
[11]: sns.histplot(data=data,x="Bathroom",bins=10,kde=True,color="red")
```

```
[11]: <Axes: xlabel='Bathroom', ylabel='Count'>
```



```
[12]: Num_col(data,"Carpet Area")
Numerical Columns Analysis:
mean:819.2012237244646
median:671.0
mode:450.0
var:342269.2402985429
std:585.037810999053
skew:16.45746798406502
MIN:2.0
MAX:20870.0
Range:20868.0
```

```
[13]: sns.histplot(data=data,x="Carpet Area",bins=10,kde=True,color="Orange")
```

```
[13]: <Axes: xlabel='Carpet Area', ylabel='Count'>
```



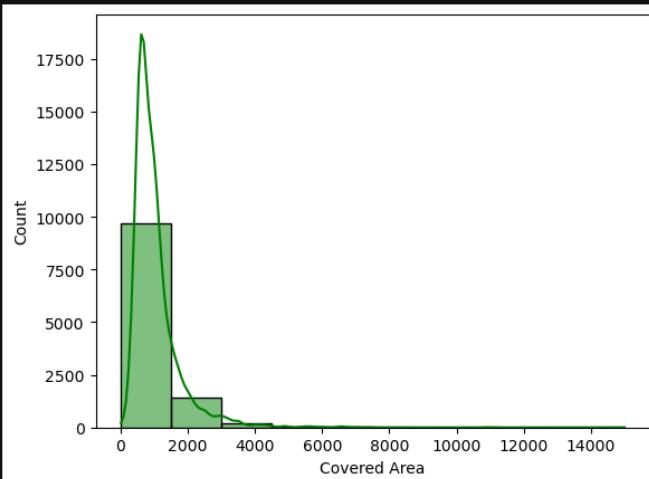


```
[14]: Num_col(data,"Covered Area")
```

```
Numerical Columns Analysis:  
mean:1037.7464392474064  
median:850.0  
mode:650.0  
var:529005.1313400863  
std:727.3273893784602  
skew:13.849281874575315  
MIN:10.0  
MAAX:15006.0  
Range:14996.0
```

```
[15]: sns.histplot(data=data,x="Covered Area",bins=10,kde=True,color="Green")
```

```
[15]: <Axes: xlabel='Covered Area', ylabel='Count'>
```

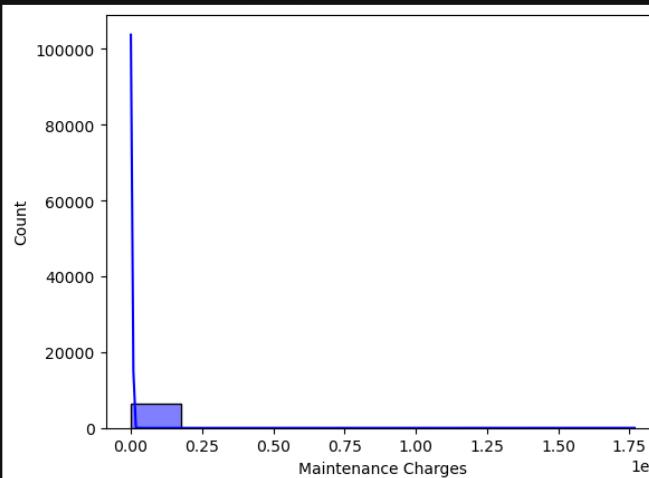


```
[16]: Num_col(data,"Maintenance Charges")
```

```
Numerical Columns Analysis:  
mean:9076.811873308929  
median:2500.0  
mode:0.0  
var:57599164677.88998  
std:239998.25973929474  
skew:167.4975874411053  
MIN:0.0  
MAAX:17680000.0  
Range:17680000.0
```

```
[17]: sns.histplot(data=data,x="Maintenance Charges",bins=10,kde=True,color="Blue")
```

```
[17]: <Axes: xlabel='Maintenance Charges', ylabel='Count'>
```



```
[18]: Num_col(data,"bedroom")
```

```
Numerical Columns Analysis:  
mean:1.9871501773748521  
median:2.0  
mode:2  
var:0.890011469482164  
std:0.943404191994653  
skew:0.900067500823549  
MIN:1  
MAAX:6  
Range:5
```

```
Rangetest
[19]: sns.histplot(data=data,x="bedroom",bins=10,kde=True,color="Purple")
[19]: <Axes: xlabel='bedroom', ylabel='Count'>

[20]: Num_col(data,"Price")
Numerical Columns Analysis:
mean:26402614.49019919
median:13500000.0
mode:9000000.0
var:1.3349145965572051e+17
std:365364830.8960791
skew:110.35727935416121
MIN:100000.0
MAX:408000000001.0
Range:407999000001.0
[21]: sns.histplot(data=data,x="Price",bins=10,kde=True,color="black")
[21]: <Axes: xlabel='Price', ylabel='Count'>

[22]: # checking and handling of outliers
def Checking_and_Handling_Of_Outliers(data, col):
    sns.boxplot(data[col], color = "Red")
    plt.title(f"Boxplot for {col}")
    plt.show()

    q1 = data[col].quantile(0.25)
    q3 = data[col].quantile(0.75)

    iqr = q3 - q1

    LowerTail = q1 - 1.5*iqr
    UpperTail = q3 + 1.5*iqr

    print(f"25% Quantile q1 = {q1}\n75% Quantile q3 = {q3}\nIQR = {iqr}\n")
    print("-"*80)
    print(f"Lower Tail = {LowerTail}\nUpper Tail = {UpperTail}")
    print("-"*80)

    # Checking for Outliers
    Outliers = data[(data[col] < LowerTail) | (data[col] > UpperTail)]
    print("\nOutliers :\n",Outliers)
    print("-"*80)

    #Handling of Outliers :
    data.loc[data[col] < LowerTail, col] = LowerTail # all outliers less than lowertail, assigned by lowertail value
    data.loc[data[col] > UpperTail, col] = UpperTail # all outliers greater than uppertail, assigned by uppertail value

    print("After handling of Outliers data:\n")
```

```
print(data.head())
```

```
[23]: Checking_and_Handling_Of_Outliers(data, "Bathroom")
```



```
25% Quantile q1 = 2.0  
75% Quantile q3 = 2.0  
IQR = 0.0
```

```
-----  
Lower Tail = 2.0  
Upper Tail = 2.0  
-----
```

Outliers :

```
Possession Status      Price   Covered Area   Carpet Area   City \
7     Ready to Move  4200000.0      650.0      585.0   Thane
8     Under Construction 3449000.0      592.0      370.0   Thane
9     Ready to Move  5000000.0      644.0      430.0   Thane
12    Ready to Move  4900000.0      595.0      NaN      Thane
13    Ready to Move  2187500.0      595.0      450.0   Thane
...     ...     ...
12671   Ready to Move  6250000.0      2210.0     1600.0  Mumbai
12672   Ready to Move  29040000.0     1600.0      967.0  Mumbai
12673     Dec '26  27500000.0      NaN      948.0  Mumbai
12674     Sep '26  10900000.0      NaN      426.0  Mumbai
12681   Under Construction 31000000.0     1376.0     1076.0  Mumbai
```

```
Type of Property  bedroom  Bathroom  Parking \
7       Apartment      1        1.0      NaN
8       Apartment      1        1.0      NaN
9       Apartment      1        1.0      NaN
12      Apartment      1        1.0      NaN
13      Apartment      1        1.0  1 Covered
...     ...
12671     Apartment      3        3.0  2 Covered
12672     Apartment      3        3.0  1 Covered
12673     Apartment      3        3.0  1 Covered
12674     Apartment      1        1.0      NaN
12681     Apartment      3        3.0  1 Covered
```

```
Location Society      Ownership Type \
7     Kalyan West, Thane      N     Freehold
8     Kalyan West, Thane      Y     Freehold
9     Kalyan West, Thane      N      NaN
12    Kalyan West, Thane      N  Co-operative Society
13    Kalyan West, Thane      Y     Freehold
...     ...
12671   Andheri East, Mumbai      Y     Freehold
12672   Andheri East, Mumbai      N     Freehold
12673  Jogeshwari Vikhroli Link Road, Mumbai      N      NaN
12674   Andheri East, Mumbai      N      NaN
12681  Jogeshwari Vikhroli Link Road, Mumbai      Y     Freehold
```

```
isPrimeLocationProperty  Maintenance Charges
7                      Y             1500.0
8                      Y              3.0
9                      Y              NaN
12                     Y              NaN
13                     Y            1000.0
...     ...
12671                    Y            5000.0
12672                    Y           12000.0
12673                    Y              NaN
12674                    Y              NaN
12681                    Y              12.0
```

```
[5765 rows x 14 columns]
```

```
-----  
After handling of Outliers data:
```

```
Possession Status      Price   Covered Area   Carpet Area   City \
0     Under Construction 3150000.0      635.0      375.0   Thane
1     Ready to Move  6300000.0      579.0      579.0   Thane
2     Ready to Move  5400000.0      850.0      585.0   Thane
3     Under Construction 9000000.0     1050.0     815.0   Thane
4     Under Construction 4950000.0      561.0      419.0   Thane
```

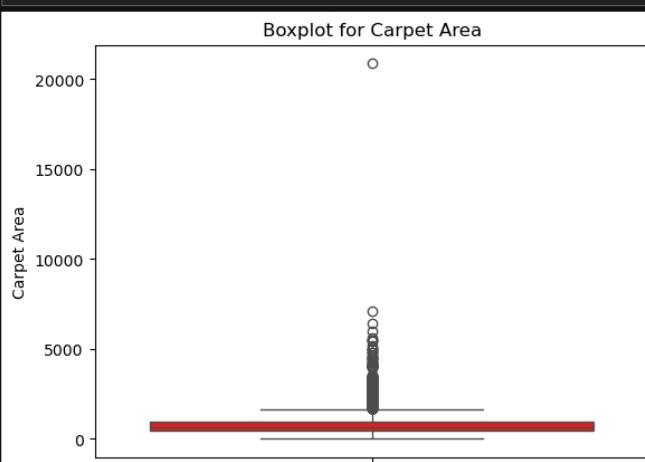
```
Type of Property  bedroom  Bathroom  Parking          Location \
0       Apartment      1        2.0  1 Covered  Kalyan West, Thane
1       Apartment      2        2.0  1 Covered  Kalyan West, Thane
2       Apartment      2        2.0  1 Open    Kalyan West, Thane
```

```
3      Apartment     2      2.0  1 Covered, 1 Open   Kalyan West, Thane
4      Apartment     1      2.0           NaN   Kalyan West, Thane
```

```
Society Ownership Type isPrimeLocationProperty Maintenance Charges
0       Y    Freehold          Y            3.0
1       Y    Freehold          Y            3.0
2       Y    Freehold          Y        1200.0
3       Y    Freehold          Y        3200.0
4       Y    Freehold          Y           NaN
```

```
[24]: data["Bathroom"] = data["Bathroom"].fillna(data["Bathroom"].median())
```

```
[25]: Checking_and_Handling_Of_Outliers(data, "Carpet Area")
```



```
-----  
Lower Tail = -255.0  
Upper Tail = 1673.0  
-----
```

Outliers :

```
Possession Status      Price  Covered Area  Carpet Area  City \
97      Ready to Move  25200000.0    2900.0    2115.0  Thane
380     Ready to Move  135000000.0    3300.0    2300.0  Mumbai
381     Ready to Move  120000000.0    4300.0    2800.0  Mumbai
397     Ready to Move  135000000.0    3300.0    2300.0  Mumbai
398     Ready to Move  120000000.0    4300.0    2800.0  Mumbai
...
12657    Ready to Move  53100000.0    3094.0    1864.0  Mumbai
12658    Ready to Move  100000000.0    3193.0    2065.0  Mumbai
12661    Under Construction  89900000.0    3500.0    2150.0  Mumbai
12665    Under Construction  73600000.0      NaN    2100.0  Mumbai
12669    Ready to Move  70000000.0    1714.0    1714.0  Mumbai
```

```
Type of Property  bedroom  Bathroom  Parking \
97      Apartment      5       2.0      NaN
380     Apartment      4       2.0      2 Open
381      Apartment      4       2.0  3 Covered
397      Apartment      4       2.0      2 Open
398      Apartment      4       2.0  3 Covered
...
12657    Apartment      4       2.0      1 Open
12658    Apartment      4       2.0  3 Covered
12661    Apartment      4       2.0  3 Covered
12665    Apartment      4       2.0  1 Covered
12669    Apartment      4       2.0  2 Covered
```

```
Location Society      Ownership Type \
97      Kalyan West, Thane      N  Co-operative Society
380      Oshiwara, Mumbai      Y    Freehold
381      Oshiwara, Mumbai      Y    Freehold
397      Oshiwara, Mumbai      Y    Freehold
398      Oshiwara, Mumbai      Y    Freehold
...
12657      Andheri East, Mumbai      Y    Freehold
12658  Jogeshwari Vikhroli Link Road, Mumbai      Y  Co-operative Society
12661      Andheri East, Mumbai      N    Freehold
12665      Andheri East, Mumbai      N      NaN
12669      Andheri East, Mumbai      Y    Freehold
```

```
isPrimeLocationProperty  Maintenance Charges
97                  N            NaN
380                 Y            14.0
381                 Y          30000.0
397                 Y            14.0
398                 Y          30000.0
...
12657                Y        15000.0
12658                Y        30000.0
12661                Y        20000.0
12665                Y            NaN
12669                Y            10.0
```

```
[727 rows x 14 columns]
```

```
-----  
After handling of Outliers data:
```

```
Possession Status      Price  Covered Area  Carpet Area  City \

```

```

0 Under Construction 3150000.0      635.0    375.0 Thane
1 Ready to Move 6300000.0       579.0    579.0 Thane
2 Ready to Move 5400000.0       850.0    585.0 Thane
3 Under Construction 9000000.0     1050.0   815.0 Thane
4 Under Construction 4950000.0      561.0    419.0 Thane

```

```

Type of Property bedroom Bathroom      Parking      Location \
0 Apartment         1     2.0      1 Covered Kalyan West, Thane
1 Apartment         2     2.0      1 Covered Kalyan West, Thane
2 Apartment         2     2.0      1 Open   Kalyan West, Thane
3 Apartment         2     2.0  1 Covered, 1 Open Kalyan West, Thane
4 Apartment         1     2.0           NaN Kalyan West, Thane

```

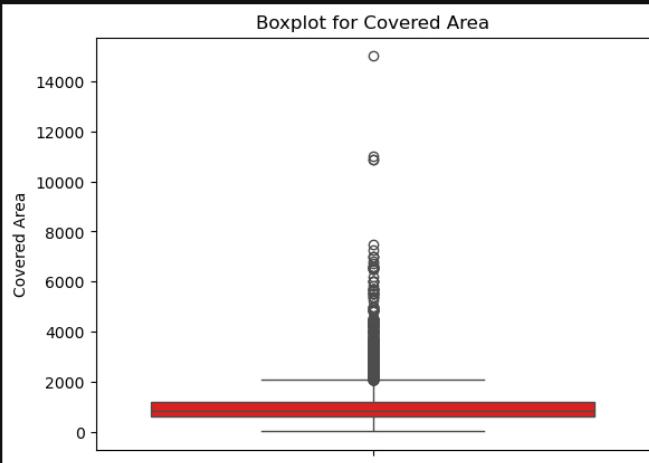
```

Society Ownership Type isPrimeLocationProperty Maintenance Charges
0 Y Freehold          Y             3.0
1 Y Freehold          Y             3.0
2 Y Freehold          Y            1200.0
3 Y Freehold          Y            3200.0
4 Y Freehold          Y             NaN

```

```
[26]: data["Carpet Area"] = data["Carpet Area"].fillna(data["Carpet Area"].median())
```

```
[27]: Checking_and_Handling_Of_Outliers(data, "Covered Area")
```



```

25% Quantile q1 = 605.5
75% Quantile q3 = 1200.0
IQR = 594.5

```

```

Lower Tail = -286.25
Upper Tail = 2091.75

```

```
Outliers :
```

```

Possession Status      Price  Covered Area  Carpet Area  City \
27 Under Construction 14796000.0    2160.0    1385.0 Thane
52 Under Construction 14796000.0    2160.0    1385.0 Thane
97 Ready to Move     25200000.0    2900.0    1673.0 Thane
367 Ready to Move     50000000.0    2500.0    1530.0 Mumbai
380 Ready to Move     135000000.0   3300.0    1673.0 Mumbai
...
...
12636 Ready to Move   51200000.0    2150.0    1673.0 Mumbai
12657 Ready to Move   53100000.0    3094.0    1673.0 Mumbai
12658 Ready to Move   100000000.0   3193.0    1673.0 Mumbai
12661 Under Construction 89900000.0   3500.0    1673.0 Mumbai
12671 Ready to Move   625000000.0   2210.0    1600.0 Mumbai

```

```

Type of Property bedroom Bathroom      Parking \
27 Apartment         3     2.0  2 Covered
52 Apartment         3     2.0  2 Covered
97 Apartment         5     2.0       NaN
367 Apartment         5     2.0       NaN
380 Apartment         4     2.0  2 Open
...
...
12636 Apartment       4     2.0  2 Covered
12657 Apartment       4     2.0  1 Open
12658 Apartment       4     2.0  3 Covered
12661 Apartment       4     2.0  3 Covered
12671 Apartment       3     2.0  2 Covered

```

```

Location Society      Ownership Type \
27 Kalyan West, Thane  Y             Freehold
52 Kalyan West, Thane  Y             Freehold
97 Kalyan West, Thane  N Co-operative Society
367 Jogeshwari West, Mumbai  Y             Freehold
380 Oshiwara, Mumbai   Y             Freehold
...
...
12636 Andheri East, Mumbai  Y             Freehold
12657 Andheri East, Mumbai  Y             Freehold
12658 Jogeshwari Vikhroli Link Road, Mumbai  Y Co-operative Society
12661 Andheri East, Mumbai   N             Freehold
12671 Andheri East, Mumbai   Y             Freehold

```

```

isPrimeLocationProperty Maintenance Charges
27 Y             7000.0
52 Y             7000.0
97 N             NaN
367 Y             NaN
380 Y             14.0
...
...
12636 Y             ...
12671 Y             10.0

```

```

12657          Y      15000.0
12658          Y      30000.0
12661          Y      20000.0
12671          Y      5000.0

```

[775 rows x 14 columns]

After handling of Outliers data:

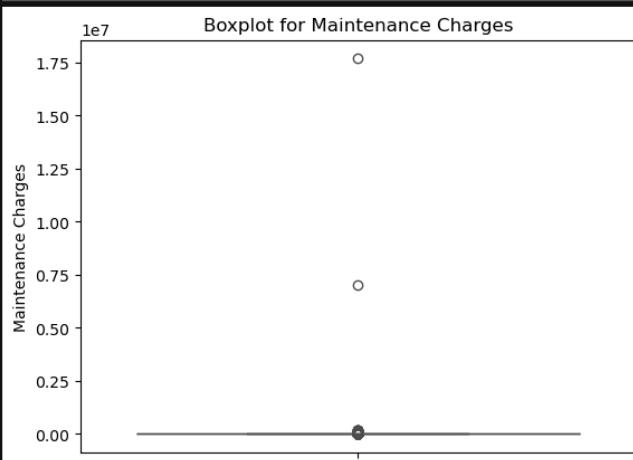
	Possession Status	Price	Covered Area	Carpet Area	City \
0	Under Construction	3150000.0	635.0	375.0	Thane
1	Ready to Move	6300000.0	579.0	579.0	Thane
2	Ready to Move	5400000.0	850.0	585.0	Thane
3	Under Construction	9000000.0	1050.0	815.0	Thane
4	Under Construction	4950000.0	561.0	419.0	Thane

	Type of Property	bedroom	Bathroom	Parking	Location \
0	Apartment	1	2.0	1 Covered	Kalyan West, Thane
1	Apartment	2	2.0	1 Covered	Kalyan West, Thane
2	Apartment	2	2.0	1 Open	Kalyan West, Thane
3	Apartment	2	2.0	1 Covered, 1 Open	Kalyan West, Thane
4	Apartment	1	2.0	NaN	Kalyan West, Thane

	Society	Ownership Type	isPrimeLocationProperty	Maintenance Charges
0	Y	Freehold	Y	3.0
1	Y	Freehold	Y	3.0
2	Y	Freehold	Y	1200.0
3	Y	Freehold	Y	3200.0
4	Y	Freehold	Y	NaN

[28]: `data["Covered Area"] = data["Covered Area"].fillna(data["Covered Area"].median())`

[29]: `Checking_and_Handling_Of_Outliers(data, "Maintenance Charges")`



25% Quantile q1 = 10.0
75% Quantile q3 = 6000.0
IQR = 5990.0

Lower Tail = -8975.0
Upper Tail = 14985.0

Outliers :

	Possession Status	Price	Covered Area	Carpet Area	City \
346	Under Construction	3200000.0	385.00	385.0	Thane
381	Ready to Move	12000000.0	2091.75	1673.0	Mumbai
398	Ready to Move	12000000.0	2091.75	1673.0	Mumbai
441	Under Construction	7500000.0	1980.00	1200.0	Mumbai
493	Ready to Move	3500000.0	1500.00	1150.0	Mumbai
...
12595	Under Construction	4090000.0	2041.00	1296.0	Mumbai
12606	Ready to Move	3750000.0	1377.00	890.0	Mumbai
12657	Ready to Move	5310000.0	2091.75	1673.0	Mumbai
12658	Ready to Move	10000000.0	2091.75	1673.0	Mumbai
12661	Under Construction	8990000.0	2091.75	1673.0	Mumbai

	Type of Property	bedroom	Bathroom	Parking \
346	Apartment	1	2.0	1 Covered
381	Apartment	4	2.0	3 Covered
398	Apartment	4	2.0	3 Covered
441	Apartment	3	2.0	2 Covered
493	Apartment	3	2.0	1 Covered, 1 Open
...
12595	Apartment	3	2.0	2 Covered
12606	Apartment	3	2.0	1 Covered
12657	Apartment	4	2.0	1 Open
12658	Apartment	4	2.0	3 Covered
12661	Apartment	4	2.0	3 Covered

	Location	Society	Ownership Type \
346	Chakki Naka, Thane	Y	Freehold
381	Oshiwara, Mumbai	Y	Freehold
398	Oshiwara, Mumbai	Y	Freehold
441	Oshiwara, Mumbai	Y	Freehold
493	Oshiwara, Mumbai	Y	Co-operative Society
...
12595	Andheri East, Mumbai	N	Freehold
12606	Jogeshwari Vikhroli Link Road, Mumbai	Y	Freehold
12657	Andheri East, Mumbai	Y	Freehold
12658	Jogeshwari Vikhroli Link Road, Mumbai	Y	Co-operative Society
12661	Andheri East, Mumbai	N	Freehold

```
12881        Andheri East, Mumbai      N      PREmboiu
```

```
    isPrimeLocationProperty  Maintenance Charges
346                  Y          72000.0
381                  Y          30000.0
398                  Y          30000.0
441                  Y          24000.0
493                  Y          15000.0
...
12595                 ...          ...
12606                 Y          15000.0
12657                 Y          20000.0
12658                 Y          15000.0
12661                 Y          30000.0
12661                 Y          20000.0
```

```
[524 rows x 14 columns]
```

```
-----  
After handling of Outliers data:
```

```
    Possession Status     Price   Covered Area   Carpet Area   City \
0  Under Construction  3150000.0       635.0       375.0  Thane
1  Ready to Move      6300000.0       579.0       579.0  Thane
2  Ready to Move      5400000.0       850.0       585.0  Thane
3  Under Construction  9000000.0      1050.0       815.0  Thane
4  Under Construction  4950000.0       561.0       419.0  Thane

    Type of Property  bedroom  Bathroom  Parking  Location \
0        Apartment      1.0      2.0      1 Covered  Kalyan West, Thane
1        Apartment      2.0      2.0      1 Covered  Kalyan West, Thane
2        Apartment      2.0      2.0      1 Open    Kalyan West, Thane
3        Apartment      2.0      2.0  1 Covered, 1 Open  Kalyan West, Thane
4        Apartment      1.0      2.0           NaN  Kalyan West, Thane

    Society Ownership Type isPrimeLocationProperty  Maintenance Charges
0            Y      Freehold             Y          3.0
1            Y      Freehold             Y          3.0
2            Y      Freehold             Y          1200.0
3            Y      Freehold             Y          3200.0
4            Y      Freehold             Y           NaN
```

```
[30]: data["Maintenance Charges"] = data["Maintenance Charges"].fillna(data["Maintenance Charges"].median())
```

```
[31]: Checking_and_Handling_Of_Outliers(data, "bedroom")
```



```
-----  
Lower Tail = -2.0  
Upper Tail = 6.0  
-----
```

```
Outliers :  
Empty DataFrame  
Columns: [Possession Status, Price, Covered Area, Carpet Area, City, Type of Property, bedroom, Bathroom, Parking, Location, Society, Ownership Type, isPrimeLocationProperty, Maintenance Charges]  
Index: []
```

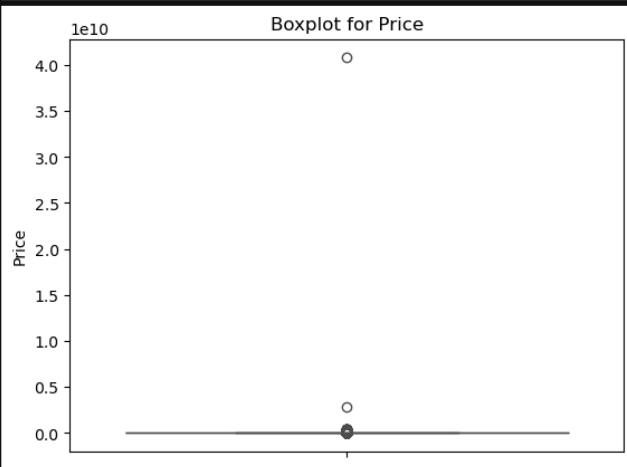
```
-----  
After handling of Outliers data:
```

```
    Possession Status     Price   Covered Area   Carpet Area   City \
0  Under Construction  3150000.0       635.0       375.0  Thane
1  Ready to Move      6300000.0       579.0       579.0  Thane
2  Ready to Move      5400000.0       850.0       585.0  Thane
3  Under Construction  9000000.0      1050.0       815.0  Thane
4  Under Construction  4950000.0       561.0       419.0  Thane

    Type of Property  bedroom  Bathroom  Parking  Location \
0        Apartment      1.0      2.0      1 Covered  Kalyan West, Thane
1        Apartment      2.0      2.0      1 Covered  Kalyan West, Thane
2        Apartment      2.0      2.0      1 Open    Kalyan West, Thane
3        Apartment      2.0      2.0  1 Covered, 1 Open  Kalyan West, Thane
4        Apartment      1.0      2.0           NaN  Kalyan West, Thane

    Society Ownership Type isPrimeLocationProperty  Maintenance Charges
0            Y      Freehold             Y          3.0
1            Y      Freehold             Y          3.0
2            Y      Freehold             Y          1200.0
3            Y      Freehold             Y          3200.0
4            Y      Freehold             Y          2500.0
```

```
[32]: Checking_and_Handling_Of_Outliers(data, "Price")
```



```
25% Quantile q1 = 7000000.0  
75% Quantile q3 = 25500000.0  
IQR = 18500000.0
```

```
-----  
Lower Tail = -20750000.0  
Upper Tail = 53250000.0  
-----
```

```
Outliers :
```

	Possession Status	Price	Covered Area	Carpet Area	City
380	Ready to Move	135000000.0	2091.75	1673.0	Mumbai
381	Ready to Move	120000000.0	2091.75	1673.0	Mumbai
397	Ready to Move	135000000.0	2091.75	1673.0	Mumbai
398	Ready to Move	120000000.0	2091.75	1673.0	Mumbai
438	Ready to Move	97500000.0	2091.75	1673.0	Mumbai
...
12661	Under Construction	89900000.0	2091.75	1673.0	Mumbai
12665	Under Construction	73600000.0	850.00	1673.0	Mumbai
12666	Under Construction	56100000.0	850.00	1650.0	Mumbai
12669	Ready to Move	70000000.0	1714.00	1673.0	Mumbai
12671	Ready to Move	62500000.0	2091.75	1600.0	Mumbai

	Type of Property	bedroom	Bathroom	Parking	Location
380	Apartment	4	2.0	2 Open	Oshiwara, Mumbai
381	Apartment	4	2.0	3 Covered	Oshiwara, Mumbai
397	Apartment	4	2.0	2 Open	Oshiwara, Mumbai
398	Apartment	4	2.0	3 Covered	Oshiwara, Mumbai
438	Apartment	6	2.0	4 Covered	Oshiwara, Mumbai
...
12661	Apartment	4	2.0	3 Covered	Andheri East, Mumbai
12665	Apartment	4	2.0	1 Covered	Andheri East, Mumbai
12666	Apartment	3	2.0	1 Covered	Andheri East, Mumbai
12669	Apartment	4	2.0	2 Covered	Andheri East, Mumbai
12671	Apartment	3	2.0	2 Covered	Andheri East, Mumbai

	Society	Ownership Type	isPrimeLocationProperty	Maintenance	Charges
380	Y	Freehold	Y		14.0
381	Y	Freehold	Y		14985.0
397	Y	Freehold	Y		14.0
398	Y	Freehold	Y		14985.0
438	Y	Freehold	Y		0.0
...
12661	N	Freehold	Y		14985.0
12665	N	NaN	Y		2500.0
12666	N	NaN	Y		2500.0
12669	Y	Freehold	Y		10.0
12671	Y	Freehold	Y		5000.0

```
[1277 rows x 14 columns]
```

```
-----  
After handling of Outliers data:
```

	Possession Status	Price	Covered Area	Carpet Area	City
0	Under Construction	3150000.0	635.0	375.0	Thane
1	Ready to Move	6300000.0	579.0	579.0	Thane
2	Ready to Move	5400000.0	850.0	585.0	Thane
3	Under Construction	9000000.0	1050.0	815.0	Thane
4	Under Construction	4950000.0	561.0	419.0	Thane

	Type of Property	bedroom	Bathroom	Parking	Location
0	Apartment	1	2.0	1 Covered	Kalyan West, Thane
1	Apartment	2	2.0	1 Covered	Kalyan West, Thane
2	Apartment	2	2.0	1 Open	Kalyan West, Thane
3	Apartment	2	2.0	1 Covered, 1 Open	Kalyan West, Thane
4	Apartment	1	2.0	NaN	Kalyan West, Thane

	Society	Ownership Type	isPrimeLocationProperty	Maintenance	Charges
0	Y	Freehold	Y		3.0
1	Y	Freehold	Y		3.0
2	Y	Freehold	Y		1200.0
3	Y	Freehold	Y		3200.0
4	Y	Freehold	Y		2500.0

```
[33]: data["Price"] = data["Price"].fillna(data["Price"].median())
```

```
[34]: data.isna().sum()
```

```
[34]: Possession Status
```

```
78
```

```

Price          0
Covered Area   0
Carpet Area    0
City           0
Type of Property 0
bedroom        0
Bathroom       0
Parking         6159
Location        323
Society          0
Ownership Type  3260
isPrimeLocationProperty 0
Maintenance Charges 0
dtype: int64

[35]: # Analysis of Categorical Columns(Variable)
def Cat_col(data, col):
    unique_values = data[col].unique() # Fixed typo: renamed to unique_values
    value_counts = data[col].value_counts()
    mode = data[col].mode()[0] # Fixed mode access by adding parentheses

    # Enhanced string formatting for clarity
    print(f"Unique Values in '{col}':\n{unique_values}\n")
    print(f"Value Counts in '{col}':\n{value_counts}\n")
    print(f"Mode of '{col}': {mode}\n")

    data[col].value_counts().plot.pie(autopct="%1.1f%%")
    plt.title(f"data-{col} (pie chart)")
    plt.show()

[36]: data.isna().sum()

[36]: Possession Status      78
      Price                0
      Covered Area         0
      Carpet Area          0
      City                 0
      Type of Property     0
      bedroom              0
      Bathroom             0
      Parking               6159
      Location              323
      Society               0
      Ownership Type        3260
      isPrimeLocationProperty 0
      Maintenance Charges  0
      dtype: int64

[37]: data.dtypes

[37]: Possession Status      object
      Price                float64
      Covered Area         float64
      Carpet Area          float64
      City                 object
      Type of Property     object
      bedroom              int64
      Bathroom             float64
      Parking               object
      Location              object
      Society               object
      Ownership Type        object
      isPrimeLocationProperty  object
      Maintenance Charges float64
      dtype: object

```

```

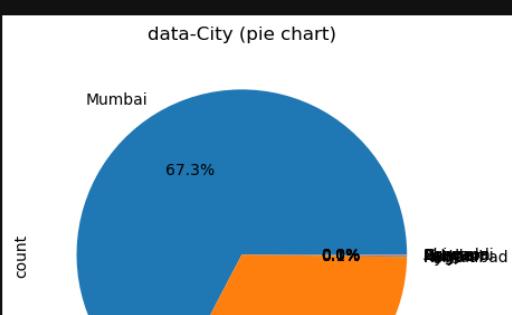
[38]: Cat_col(data, "City")

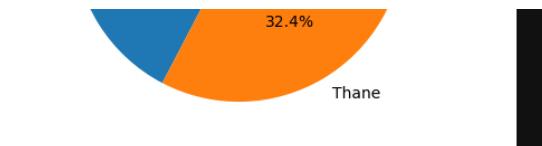
Unique Values in 'City':
['Thane' 'Kalyan' 'Mumbai' 'Agartala' 'Palghar' 'Bhiwandi' 'Gurgaon'
 'Hyderabad' 'Nagpur']

Value Counts in 'City':
City
Mumbai      8542
Thane       4113
Kalyan        9
Hyderabad     8
Nagpur        7
Palghar       2
Bhiwandi     2
Agartala      1
Gurgaon       1
Name: count, dtype: int64

Mode of 'City': Mumbai

```





```
[39]: Cat_col(data,"Location")
```

Unique Values in 'Location':

- ['Kalyan West, Thane' 'nan' 'Kalyan West, Kalyan' 'Bhoiwada, Thane'
- 'Rambaug Lane Number 4, Thane' 'Kalyan East, Thane' 'Chakki Naka, Thane'
- 'Nandivali Gaon, Thane' 'Kalyan East, Kalyan' 'Kalyan East, Mumbai'
- 'Tisgaon Naka, Kalyan' 'Tisgaon Naka, Thane' 'Oshiwara, Mumbai'
- 'Bandivali, Mumbai' 'Jogeshwari West, Mumbai' 'Malcom Baug, Mumbai'
- 'Amrut Nagar, Mumbai' 'Behram Baug, Mumbai' 'Indira Nagar, Mumbai'
- 'Jogeshwari Vikhroli Link Road, Mumbai' 'Jogeshwari East, Mumbai'
- 'Majes Wadi, Mumbai' 'Natwan Nagar, Mumbai' 'Banamalipur, Argatala'
- 'Hind Nagar, Mumbai' 'Shan Nagar, Mumbai' 'Dombivli East, Thane'
- 'Sonar Pada, Thane' 'Nandivali, Thane' 'Desai Village, Thane'
- 'Tilak Nagar, Thane' 'Tilak Nagar, Mumbai' 'Gandhi Nagar Dombivli, Thane'
- 'Usarghar Gaon, Thane' 'Dombivli East, Kalyan' 'Dahisar West, Mumbai'
- 'Kandarpada, Mumbai' 'Nhatre Wadi, Mumbai' 'Gomant Nagar, Mumbai'
- 'Leo Peter Wadi, Mumbai' 'Shanti Ashram, Mumbai' 'Navagaon, Mumbai'
- 'Ashok Van, Mumbai' 'Dahisar East, Mumbai' 'Anand Nagar, Mumbai'
- 'Vaishali Nagar Dahisar, Mumbai' 'Janta Nagar, Mumbai'
- 'Vaishali Nagar, Mumbai' 'Rawal Pada, Mumbai'
- 'Maratha Colony Dahisar, Mumbai' 'Shakti Nagar, Mumbai'
- 'Ketki Pada, Mumbai' 'NL Complex, Mumbai' 'Ghartan Pada, Mumbai'
- 'Balaji Nagar, Mumbai' 'Waghdevi Nagar, Mumbai' 'Virar West, Mumbai'
- 'Bolinj, Mumbai' 'Chikhaldongari, Mumbai' 'Agashi, Mumbai'
- 'Tirupati Nagar, Mumbai' 'Y K Nagar, Mumbai' 'Virar West, Thane'
- 'Gokul Township, Mumbai' 'Virat Nagar, Mumbai' 'Yashavant Nagar, Mumbai'
- 'Virar West, Palghar' 'Grant Road, Mumbai' 'Virar East, Mumbai'
- 'Virar East, Palghar' 'Bhalchandra Nagar, Mumbai' 'Dadar West, Mumbai'
- 'Shivaji Park, Mumbai' 'Kaka Sahab Dadarkar Chowk, Mumbai'
- 'Chandrakant Dhuru Wadi, Mumbai' 'Kabutar Khana Purandare Chowk, Mumbai'
- 'Prabodhankar Thakre Chowk, Mumbai' 'Bhawan Shankar, Mumbai'
- 'Jakhadevi Chowk, Mumbai'
- 'Suryavanshi Kshatriya Sabhagruha Chowk, Mumbai'
- 'Vile Parle West, Mumbai' 'Gulmohar Road, Mumbai' 'Suresh Colony, Mumbai'
- 'Irla, Mumbai' 'Sanyas Ashram Chowk, Mumbai' 'Kamla Nagar, Mumbai'
- 'Vile Parle East, Mumbai' 'Air India Colony, Mumbai'
- 'Shahaji Raje Marg, Mumbai' 'Parle Colony, Mumbai' 'Vasai West, Mumbai'
- 'Suyog Nagar, Mumbai' 'Chulne, Mumbai' 'Dindayal Nagar Vasai, Mumbai'
- 'Vasai East, Mumbai' 'Golani Naka, Mumbai' 'Dhumal Nagar, Mumbai'
- 'Pokharan Road Number 2, Thane' 'Naupada, Thane' 'Kolshet Road, Thane'
- 'Thane West, Thane' 'Owale, Thane' 'Kolshet, Thane' 'Dhokali, Thane'
- 'Balkum village, Thane' 'Majiwada, Thane' 'Manpada, Thane'
- 'Hiranandani Estate, Thane' 'Bhayandarpada, Thane'
- 'Wagle Industrial Estate, Thane' 'Vasant Vihar, Thane'
- 'Panch Pakhdhi, Thane' 'Teen Hath Naka, Thane' 'Patlipada, Thane'
- 'Brahmand, Thane' 'Waghbil, Thane' 'Thane West, Bhilwandi'
- 'Balkum Naka, Thane' 'Koliwada, Thane' 'Kolbad, Thane'
- 'Vartak Nagar, Thane' 'Naupada, Mumbai' 'Sawarkar Nagar, Thane'
- 'Vrindavan Society, Thane' 'Anand Nagar Thane West, Thane'
- 'Khopat, Thane' 'Pokharan Road Number 1, Thane' 'Kapurbaudi, Thane'
- 'Shivai Nagar, Thane' 'Charai, Thane' 'Ram Maruti Road Naupada, Thane'
- 'Rabodi, Thane' 'Lok Puram, Thane' 'Dhobi Ali, Thane'
- 'Mahagiri Koliwada, Thane' 'Lokmanya Nagar, Thane' 'Pawar Nagar, Thane'
- 'Uthalsar, Thane' 'Jambli Naka, Thane' 'Shivaji Nagar, Thane'
- 'Bhaskar Colony, Thane' 'Wishnu Nagar, Thane' 'Louis Wadi, Thane'
- 'Bawand Baug, Thane' 'Meenatai Thakrey Chowk, Thane' 'Raunak Park, Thane'
- 'Runwal Nagar, Thane' 'Ghantali, Thane' 'Tala Pali, Thane'
- 'Vijay Nagar, Thane' 'Harinivas Circle, Thane' 'Kopri, Thane'
- 'Thane East, Thane' 'Sion West, Mumbai' 'Chembur East, Mumbai'
- 'Charai, Mumbai' 'Sion East, Mumbai' 'Sion Koliwada, Mumbai'
- 'Sion East, Gunaon' 'Santa Cruz West, Mumbai' 'Hasmukh Nagar, Mumbai'
- 'Willingdon Colony, Mumbai' 'North Ave, Mumbai' 'Saraswat Colony, Mumbai'
- 'NSES Colony, Mumbai' 'Vithaldas Nagar, Mumbai' 'Dadar East, Mumbai'
- 'Hindu Colony, Mumbai' 'Dadar TT Circle, Mumbai'
- 'Mancharji Edhalji Joshi Chowk, Mumbai' 'L Tilak Colony, Mumbai'
- 'Santa Cruz East, Mumbai' 'Kalina, Mumbai' 'Vidya Nagari, Mumbai'
- 'Sen Nagar, Mumbai' 'Kranti Nagar, Mumbai' 'Sundar Nagar, Kalina, Mumbai'
- 'Prabhate Colony, Mumbai' 'Kolivary, Mumbai'
- 'Anand Nagar Santa Cruz East, Mumbai' 'Prabhadevi, Mumbai'
- 'Adv RD Koite Chowk, Mumbai' 'Century Bazaar, Mumbai'
- 'Shree Siddhivinayak Mandir Chowk, Mumbai' 'Babar Shaikh Chowk, Mumbai'
- 'Mira Road East, Mumbai' 'Mahajan Wadi, Mumbai'
- 'Geeta Nagar Mira Road, Mumbai' 'Penkarpada, Mumbai' 'Miragaon, Mumbai'
- 'Nalasopara West, Mumbai' 'Laxmiben Chheda Nagar, Mumbai'
- 'Samel Pada, Mumbai' 'Nalasopara West, Thane' 'Nalasopara East, Mumbai'
- 'Oswal Nagari, Mumbai' 'Naigaon West, Mumbai' 'Naigaon East, Mumbai'
- 'Goregaon East, Mumbai' 'Gokuldhham, Mumbai' 'Film City Road, Mumbai'
- 'Jayaprakash Nagar, Mumbai' 'Bimbisar Nagar, Mumbai'
- 'Ganesh Nagar, Mumbai' 'Kanya Pada, Mumbai'
- 'Nagar Niwara Parishad, Mumbai' 'Sai Baba Complex, Mumbai'
- 'Arey Milk Colony, Mumbai' 'NNP Colony, Mumbai' 'Shivaji Nagar, Mumbai'
- 'Royal Palms Estate, Mumbai' 'Laxmi Nagar, Mumbai'
- 'Pandurang Wadi, Mumbai' 'Subhash Nagar, Mumbai' 'Churi Wadi, Mumbai'
- 'Azad Nagar, Mumbai' 'Shreyas Colony, Mumbai'
- 'Durgeshwari Welfare Society, Mumbai' 'Film City, Mumbai'
- 'Yeshodham, Mumbai' 'Indian Oil Nagar- Shivaji Nagar, Mumbai'
- 'Ram Nagar, Mumbai' 'CANA Industrial Estate, Mumbai' 'Yashodham, Mumbai'
- 'ITT Bhatti, Mumbai' 'Vishveswar Nagar, Mumbai' 'Agricpada, Mumbai'
- 'Mumbai Central, Mumbai' 'RTO Colony, Mumbai' 'Churchgate, Mumbai'
- 'Natunga West, Mumbai' 'Joshi Wadi, Mumbai'
- 'Mahalaxmi Sindhi Colony, Mumbai' 'S Haralayya Chowk, Mumbai'
- 'Matunga East, Mumbai' 'Brhmanwada, Mumbai' 'Five Gardens, Mumbai'
- 'Kings Circle, Mumbai' 'Marine Lines East, Mumbai' 'Marine Lines, Mumbai'
- 'Churchgate station, Mumbai' 'Marine Drive, Mumbai' 'Malad East, Mumbai'
- 'Pathanwadi, Mumbai' 'Kurkar Village, Mumbai' 'Triveni Nagar, Mumbai'
- 'Pushpa Park, Mumbai' 'Shivaji Nagar Malad East, Mumbai'
- 'Primal Nagar, Mumbai' 'Dindoshi, Mumbai' 'Kokani Pada, Mumbai'

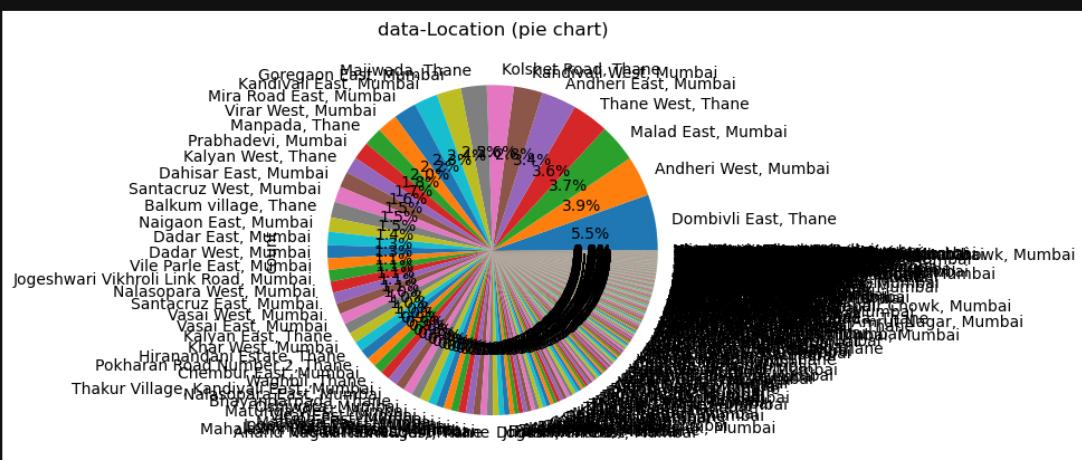
```
'Upper Govind Nagar, Mumbai' 'Dhanji Wadi, Mumbai'
'Pathan Wadi Chowk, Mumbai' 'Dombivli West, Thane' 'Kopargaon, Thane'
'Nahim West, Mumbai' 'Asavari, Mumbai' 'Nahim East, Mumbai'
'Nahalaxmi Race Course, Mumbai' 'Lower Parel West, Mumbai'
'Chembur West, Mumbai' 'Rahul Nagar Chembur, Mumbai'
'Lower Parel East, Mumbai' 'Khar West, Mumbai' 'Juhi Koliwada, Mumbai'
'Old Khar, Mumbai' 'Khar Danda, Mumbai' 'Danda, Mumbai' 'Golibar, Mumbai'
'RPF Colony, Mumbai' 'Khar East, Mumbai' 'Kandivali West, Mumbai'
'Mahavir Nagar, Mumbai' 'Mahatama Gandhi Road Amrut Nagar, Mumbai'
'Dhanukar Wadi, Mumbai' 'Charkop Sector 2, Mumbai' 'Irani Wada, Mumbai'
'Charkop, Mumbai' 'Bander Pakhdai, Mumbai' 'Kandivali West, Nagpur'
'Charkop Sector 9, Mumbai' 'Charkop Sector 3, Mumbai'
'Charkop Village, Mumbai' 'Charkop Sector 8, Mumbai'
'Charkop Sector 6, Mumbai' 'Charkop Sector 1, Mumbai'
'Charkop Sector 7, Mumbai' 'Jethava Nagar, Mumbai'
'Ekta Nagar Charkop, Mumbai' 'Shankar Pada, Mumbai' 'Goraswadi, Mumbai'
'Patel Nagar, Mumbai' 'Ganesh Nagar JJC Area, Mumbai'
'Fateh Baug, Mumbai' 'Kandivali East, Mumbai' 'Lokhandwala Twp, Mumbai'
'Alika Nagar, Mumbai' 'Thakur Village, Kandivali East, Mumbai'
'Akurli Nagar, Mumbai' 'Samata Nagar, Mumbai'
'Ashok Nagar Western Mumbai, Mumbai' 'Thakur Complex, Mumbai'
'Asha Nagar, Mumbai' 'Appa Pada, Mumbai' 'Damodarwadi, Mumbai'
'Dattani Park, Mumbai' 'Singh Agri Estate, Mumbai' 'Amboli, Mumbai'
'Laxmi Industrial Estate, Mumbai' 'Andheri West, Mumbai'
'Lokhandwala Complex, Mumbai' 'Four Bungalows, Mumbai'
'Seven Bungalows, Mumbai' 'Veera Desai Road, Mumbai' '4 Bunglows, Mumbai'
'Lailu Bhai Park, Mumbai' 'Azad Nagar 2, Mumbai' 'D.N. Nagar, Mumbai'
'Andheri West, Nagpur' 'Versova Harg, Mumbai' 'DN Nagar, Mumbai'
'Willat Nagar, Mumbai' 'Gilbert Hill JJC, Mumbai'
'SV Patel Nagar, Mumbai' 'Shastri Nagar D Phase, Mumbai'
'Shri Swami Samarth Nagar, Mumbai' 'Navneeth Colony, Mumbai'
'Sundervan Complex, Mumbai' 'Dattaguru nagar, Mumbai'
'Andheri East, Mumbai' 'Marol, Mumbai' 'Chakala, Mumbai'
'Marol Naka, Mumbai' 'Saki Vihar Road, Mumbai' 'JB Nagar, Mumbai'
'Poonam Nagar, Mumbai' 'Marol Maroshi Road, Mumbai'
'Sher E Punjab Society, Mumbai' 'Sakinaka Junction, Mumbai'
'Kadam Wadi Marol, Mumbai' 'Seepz Area, Mumbai' 'Chakala MIDC, Mumbai'
'D Charatsingh Colony, Mumbai' 'Andheri East, Nagpur'
'Malpa Dongri, Mumbai' 'Gondavalni Gaothan, Mumbai' 'Kajuwadi, Mumbai'
'Marol Maroshi, Mumbai' 'PNPG Colony, Mumbai'
'Mogra Village Road, Mumbai' 'Kanti Nagar, Mumbai'
'Bamandaya Pada, Mumbai' 'Koladongri, Mumbai' 'Mogra Village, Mumbai'
'Jijamata Road, Mumbai' 'Hanuman Nagar, Mumbai' 'Azad Nagar JJC, Mumbai'
'Parsi Wada, Mumbai' 'Marol Naka Junction, Mumbai'
'Bhavani Nagar, Mumbai' 'Maheshwari Nagar, Mumbai'
'Kondivita Village, Mumbai' 'Vijay Nagar, Hyderabad']
```

Value Counts in 'Location':

Location	Count
Dombivli East, Thane	674
Andheri West, Mumbai	488
Malad East, Mumbai	460
Thane West, Thane	443
Andheri East, Mumbai	421
...	...
Balaji Nagar, Mumbai	1
Waghdevi Nagar, Mumbai	1
Kopargaon, Thane	1
Virar West, Thane	1
Vijay Nagar, Hyderabad	1

Name: count, Length: 365, dtype: int64

Mode of 'Location': Dombivli East, Thane



[40]: `Cat_col(data,"Ownership Type")`

Unique Values in 'Ownership Type':

```
['Freehold' 'Co-operative Society' nan 'Power Of Attorney' 'Leasehold']
```

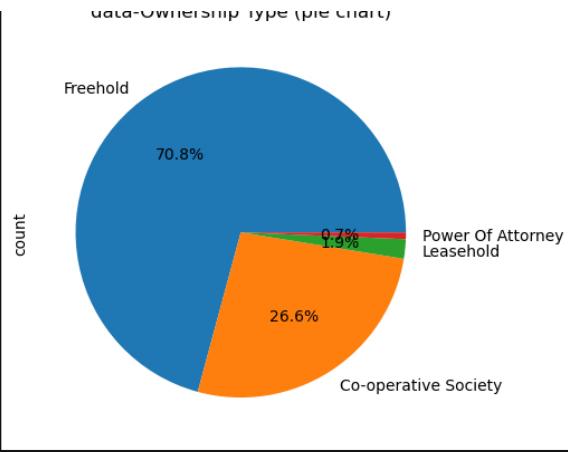
Value Counts in 'Ownership Type':

Ownership Type	Count
Freehold	6673
Co-operative Society	2511
Leasehold	179
Power Of Attorney	62

Name: count, dtype: int64

Mode of 'Ownership Type': Freehold



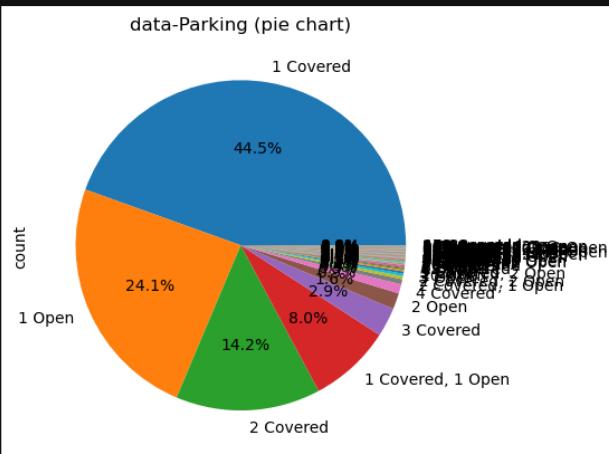


```
[41]: Cat_col(data,"Parking")
```

```
Unique Values in 'Parking':  
['1 Covered' '1 Open' '1 Covered', 1 Open' nan '2 Covered' '2 Open'  
'5 Open' '3 Covered' '4 Covered' '10 Covered' '2 Covered, 2 Open'  
'10 Open' '6 Open' '202 Covered' '306 Open' '8 Covered' '20 Open'  
'32 Covered' '180 Covered' '9 Open' '15 Open' '5 Covered' '100 Open'  
'50 Open' '10 Covered, 20 Open' '12 Covered' '7 Open' '40 Open'  
'503 Open' '1 Covered, 2 Open' '2 Covered, 1 Open' '9 Covered' '3 Open'  
'505 Covered' '47 Covered' '7 Covered, 1 Open' '4 Open' '14 Open'  
'20 Covered' '8 Open' '606 Covered, 606 Open' '10 Covered, 10 Open'  
'111 Open' '138 Covered' '392 Covered' '7 Covered' '18 Covered'  
'27 Covered' '181 Covered' '6 Covered, 6 Open' '801 Covered'  
'2 Covered, 4 Open' '16 Covered' '1 Covered, 3 Open' '12 Open'  
'50 Covered' '3 Covered, 1 Open' '25 Open' '408 Covered'  
'1 Covered, 10 Open' '45 Covered' '3 Covered, 2 Open' '107 Covered'  
'150 Covered' '200 Covered' '22 Covered' '16 Open'  
'250 Covered, 200 Open' '24 Covered, 21 Open' '76 Open' '904 Covered'  
'303 Open' '23 Open' '161 Open' '100 Covered' '43 Open' '34 Covered'  
'1 Covered, 504 Open' '5 Covered, 2 Open' '120 Covered' '35 Open'  
'30 Open' '502 Covered' '103 Covered']
```

```
Value Counts in 'Parking':  
Parking  
1 Covered           2905  
1 Open              1575  
2 Covered           925  
1 Covered, 1 Open   525  
3 Covered           186  
...  
16 Covered          1  
2 Covered, 4 Open    1  
505 Covered         1  
47 Covered          1  
103 Covered         1  
Name: count, Length: 83, dtype: int64
```

Mode of 'Parking': 1 Covered



```
[42]: Cat col(data,"Possession Status")
```

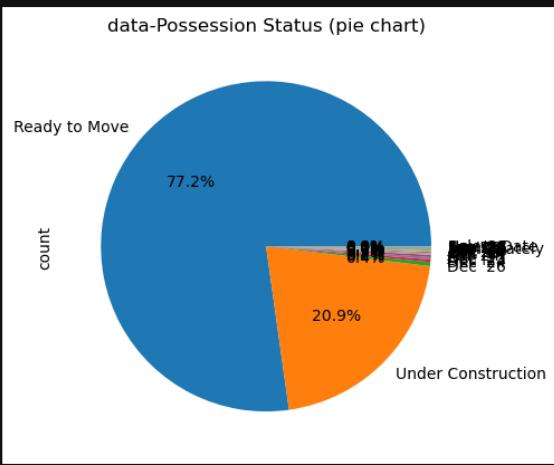
```
Unique Values in 'Possession Status':  
['Under Construction' 'Ready to Move' nan "Jun '26" "Dec '28" "Nov '25"  
"Dec '26" "Mar '25" "Dec '23" "Jun '24" "Dec '24" "Sep '23" "Mar '24"  
"Immediately" "Dec '27" "Dec '25" "Apr '28" "Nov '23" "Apr '24" "Jun '25"  
"Select Date" "Sep '26"]
```

Value Counts in 'Possession Status':

Possession Status	
Ready to Move	9735
Under Construction	2630
Dec '26	50
Dec '24	31
Dec '23	29
Dec '25	28
Jun '24	21
Dec '22	16

```
Jun '26      16
Dec '28     12
Mar '25      6
Mar '24      5
Apr '28      4
Sep '26      4
Immediately    4
Nov '25      3
Apr '24      3
Jun '25      3
Sep '23      3
Nov '23      2
Select Date   2
Name: count, dtype: int64
```

Mode of 'Possession Status': Ready to Move

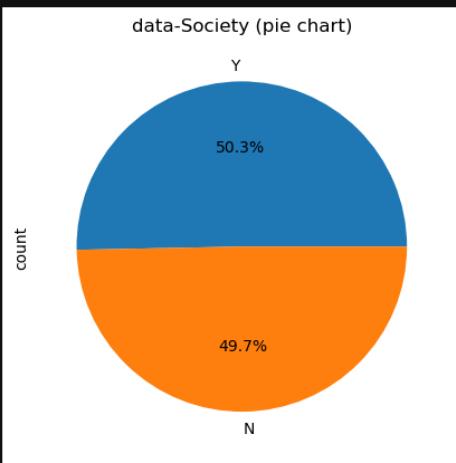


```
[43]: Cat_col(data,"Society")
```

Unique Values in 'Society':
['Y', 'N']

Value Counts in 'Society':
Society
Y 6383
N 6302
Name: count, dtype: int64

Mode of 'Society': Y

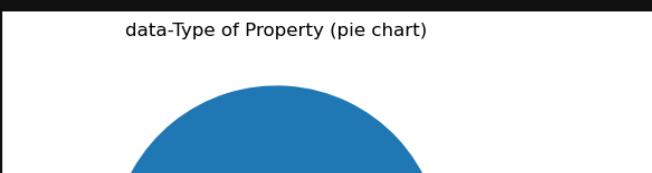


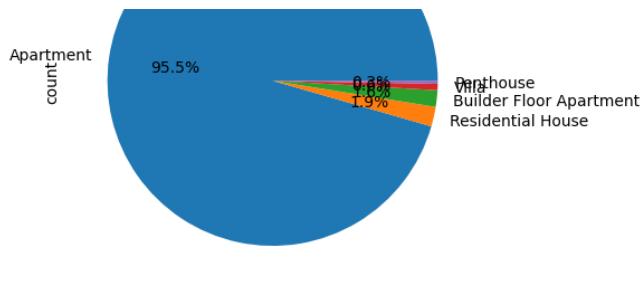
```
[44]: Cat_col(data,"Type of Property")
```

Unique Values in 'Type of Property':
['Apartment', 'Residential House', 'Builder Floor Apartment', 'Penthouse', 'Villa']

Value Counts in 'Type of Property':
Type of Property
Apartment 12119
Residential House 244
Builder Floor Apartment 205
Villa 79
Penthouse 38
Name: count, dtype: int64

Mode of 'Type of Property': Apartment

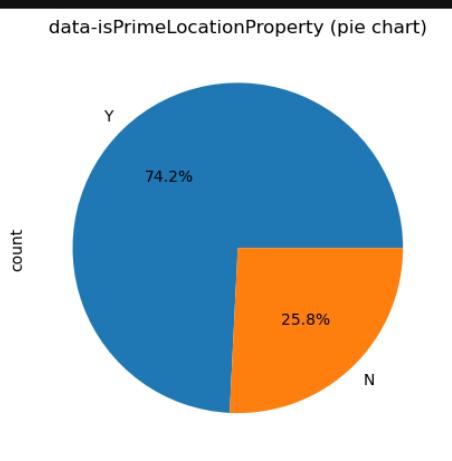




```
[45]: Cat_col(data,"isPrimeLocationProperty")
Unique Values in 'isPrimeLocationProperty':
['Y' 'N']

Value Counts in 'isPrimeLocationProperty':
isPrimeLocationProperty
Y    9414
N    3271
Name: count, dtype: int64

Mode of 'isPrimeLocationProperty': Y
```



```
[46]: data.isna().sum()
```

```
[46]: Possession Status      78
Price             0
Covered Area     0
Carpet Area       0
City              0
Type of Property  0
bedroom          0
Bathroom          0
Parking           6159
Location          323
Society            0
Ownership Type    3260
isprimeLocationProperty  0
Maintenance Charges  0
dtype: int64
```

```
[47]: # replace missing value of Categorical Column
data["Location"] = data["Location"].fillna(data["Location"].mode()[0])
data["Ownership Type"] = data["Ownership Type"].fillna(data["Ownership Type"].mode()[0])
data["Parking"] = data["Parking"].fillna(data["Parking"].mode()[0])
data["Possession Status"] = data["Possession Status"].fillna(data["Possession Status"].mode()[0])
```

```
[48]: data.isna().sum()
```

```
[48]: Possession Status      0
Price             0
Covered Area     0
Carpet Area       0
City              0
Type of Property  0
bedroom          0
Bathroom          0
Parking           0
Location          0
Society            0
Ownership Type    0
isprimeLocationProperty  0
Maintenance Charges  0
dtype: int64
```

```
[49]: data.dtypes
```

```
[49]: Possession Status      object
Price             float64
Covered Area     float64
Carpet Area       float64
City              object
Type of Property  object
bedroom          int64
```

```
Bathroom          float64
Parking           object
Location          object
Society           object
Ownership Type   object
isPrimeLocationProperty    object
Maintenance Charges float64
dtype: object
```

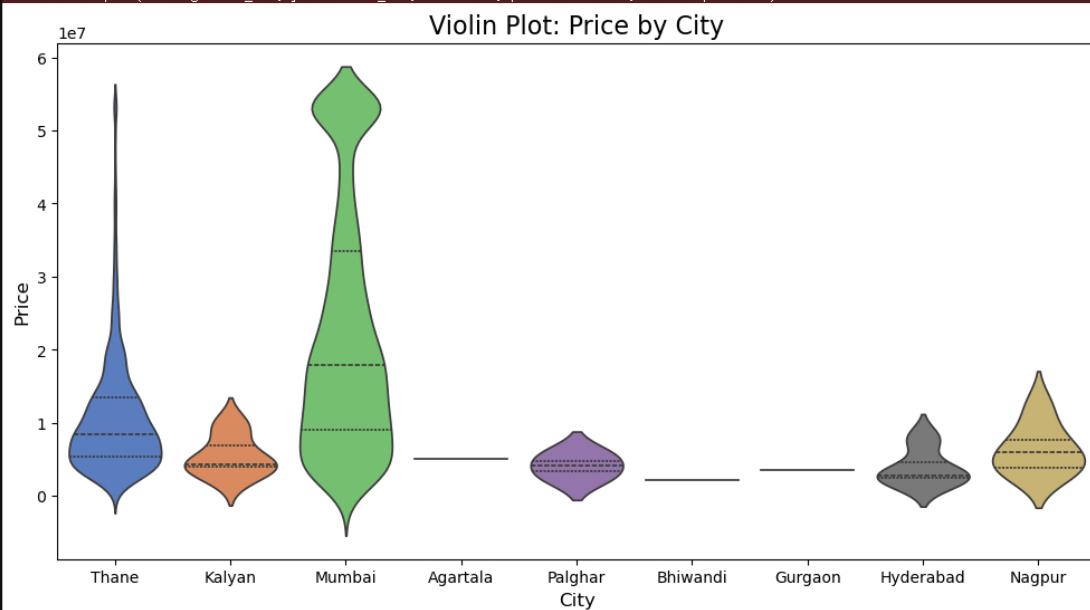
```
[50]: # Bivariate analysis of columns
def bivariate_violin_plot(data, categorical_col, numerical_col):
    plt.figure(figsize=(12, 6))
    sns.violinplot(x=categorical_col, y=numerical_col, data=data, palette="muted", inner="quartile")
    plt.title(f'Violin Plot: {numerical_col} by {categorical_col}', fontsize=16)
    plt.xlabel(categorical_col, fontsize=12)
    plt.ylabel(numerical_col, fontsize=12)
    plt.show()
```

```
[51]: bivariate_violin_plot(data,"City" , "Price")
```

```
C:\Users\deshm\AppData\Local\Temp\ipykernel_38268\3790340880.py:4: FutureWarning:
```

```
Passing 'palette' without assigning 'hue' is deprecated and will be removed in v0.14.0. Assign the 'x' variable to 'hue' and set 'legend=False' for the same effect.
```

```
sns.violinplot(x=categorical_col, y=numerical_col, data=data, palette="muted", inner="quartile")
```



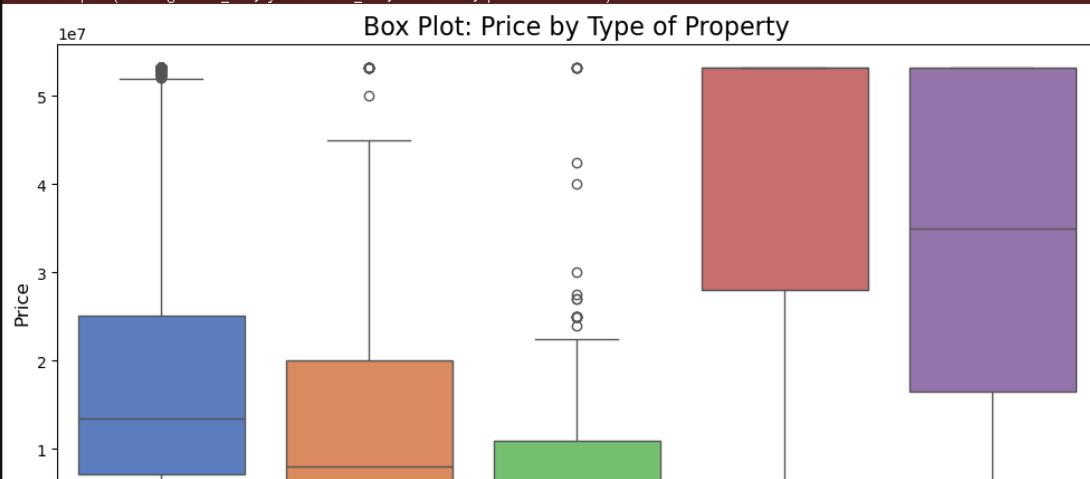
```
[52]: def bivariate_box_plot(data, categorical_col, numerical_col):
    plt.figure(figsize=(12, 6))
    sns.boxplot(x=categorical_col, y=numerical_col, data=data, palette="muted")
    plt.title(f'Box Plot: {numerical_col} by {categorical_col}', fontsize=16)
    plt.xlabel(categorical_col, fontsize=12)
    plt.ylabel(numerical_col, fontsize=12)
    plt.show()
```

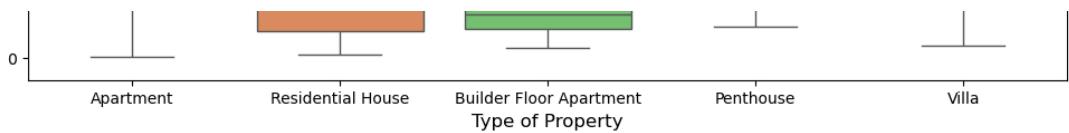
```
[53]: bivariate_box_plot(data,"Type of Property","Price")
```

```
C:\Users\deshm\AppData\Local\Temp\ipykernel_38268\3297854816.py:3: FutureWarning:
```

```
Passing 'palette' without assigning 'hue' is deprecated and will be removed in v0.14.0. Assign the 'x' variable to 'hue' and set 'legend=False' for the same effect.
```

```
sns.boxplot(x=categorical_col, y=numerical_col, data=data, palette="muted")
```





```
[54]: data.dtypes
```

```
[54]: Possession Status      object
Price             float64
Covered Area     float64
Carpet Area      float64
City              object
Type of Property  object
bedroom          int64
Bathroom         float64
Parking           object
Location          object
Society            object
Ownership Type    object
isPrimeLocationProperty  object
Maintenance Charges float64
dtype: object
```

```
[55]: # Encoding
```

```
#Bathroom
data["City"].unique()
```

```
[55]: array(['Thane', 'Kalyan', 'Mumbai', 'Agartala', 'Palghar', 'Bhiwandi',
       'Gurgaon', 'Hyderabad', 'Nagpur'], dtype=object)
```

```
[56]: data["City_encoded"] = data["City"].replace({'Thane':0, 'Kalyan':1, 'Mumbai':3, 'Agartala':4, 'Palghar':5, 'Bhiwandi':6,
       'Gurgaon':7, 'Hyderabad':8, 'Nagpur':9})
```

```
C:\Users\deshm\AppData\Local\Temp\ipykernel_38268\706628898.py:1: FutureWarning: Downcasting behavior in `replace` is deprecated and will be removed in a future version. To retain the old behavior, explicitly call `result.infer_objects(copy=False)`. To opt-in to the future behavior, set `pd.set_option('future.no_silent_downcasting', True)`
    data["City_encoded"] = data["City"].replace({'Thane':0, 'Kalyan':1, 'Mumbai':3, 'Agartala':4, 'Palghar':5, 'Bhiwandi':6,
```

```
[57]: #Location Column
```

```
data["Location"].unique()
```

```
[57]: array(['Kalyan West, Thane', 'Dombivli East, Thane',
       'Kalyan West, Kalyan', 'Bhoiwada, Thane',
       'Rambaug Lane Number 4, Thane', 'Kalyan East, Thane',
       'Chakki Naka, Thane', 'Nandivali Gaon, Thane',
       'Kalyan East, Kalyan', 'Kalyan East, Mumbai',
       'Tisgaon Naka, Kalyan', 'Tisgaon Naka, Thane', 'Oshiwara, Mumbai',
       'Bandivali, Mumbai', 'Jogeshwari West, Mumbai',
       'Malcom Baug, Mumbai', 'Amrut Nagar, Mumbai',
       'Behram Baug, Mumbai', 'Indira Nagar, Mumbai',
       'Jogeshwari Vikhroli Link Road, Mumbai', 'Jogeshwari East, Mumbai',
       'Najas Wadi, Mumbai', 'Natwan Nagar, Mumbai',
       'Banamalipur, Agartala', 'Hind Nagar, Mumbai',
       'Sham Nagar, Mumbai', 'Sonar Pada, Thane', 'Nandivali, Thane',
       'Desai Village, Thane', 'Tilak Nagar, Thane',
       'Tilak Nagar, Mumbai', 'Gandhi Nagar Dombivli, Thane',
       'Usanghar Gaon, Thane', 'Dombivli East, Kalyan',
       'Dahisar West, Mumbai', 'Kandarpada, Mumbai',
       'Nhatre Wadi, Mumbai', 'Gomant Nagar, Mumbai',
       'Leo Peter Wadi, Mumbai', 'Shanti Ashram, Mumbai',
       'Navagaon, Mumbai', 'Ashok Van, Mumbai', 'Dahisar East, Mumbai',
       'Anand Nagar, Mumbai', 'Vaishali Nagar Dahisar, Mumbai',
       'Janta Nagar, Mumbai', 'Vaishali Nagar, Mumbai',
       'Rawal Pada, Mumbai', 'Maratha Colony Dahisar, Mumbai',
       'Shakti Nagar, Mumbai', 'Ketki Pada, Mumbai', 'NL Complex, Mumbai',
       'Ghartan Pada, Mumbai', 'Balaji Nagar, Mumbai',
       'Waghdevi Nagar, Mumbai', 'Virar West, Mumbai', 'Bolinj, Mumbai',
       'Chikhaldongari, Mumbai', 'Agashi, Mumbai',
       'Tirupati Nagar, Mumbai', 'Y K Nagar, Mumbai', 'Virar West, Thane',
       'Gokul Township, Mumbai', 'Virat Nagar, Mumbai',
       'Yashwant Nagar, Mumbai', 'Virar West, Palghar',
       'Grant Road, Mumbai', 'Virar East, Mumbai', 'Virar East, Palghar',
       'Bhalchandra Nagar, Mumbai', 'Dadar West, Mumbai',
       'Shivaji Park, Mumbai', 'Kaka Sahab Dadarkar Chowk, Mumbai',
       'Chandrakant Dhuru Wadi, Mumbai',
       'Kabutarkhana Purandare Chowk, Mumbai',
       'Prabodhankar Thakre Chowk, Mumbai', 'Bhawani Shankar, Mumbai',
       'Jakhadevi Chowk, Mumbai',
       'Suryavanshi Kshatriya Sabhgruha Chowk, Mumbai',
       'Vile Parle West, Mumbai', 'Gulmohar Road, Mumbai',
       'Suresh Colony, Mumbai', 'Irla, Mumbai',
       'Sanyas Ashram Chowk, Mumbai', 'Kamla Nagar, Mumbai',
       'Vile Parle East, Mumbai', 'Air India Colony, Mumbai',
       'Shahaji Roje Marg, Mumbai', 'Parle Colony, Mumbai',
       'Vasai West, Mumbai', 'Suyog Nagar, Mumbai', 'Chulne, Mumbai',
       'Dindayal Nagar Vasai, Mumbai', 'Vasai East, Mumbai',
       'Golani Naka, Mumbai', 'Dhumal Nagar, Mumbai',
       'Pokharan Road Number 2, Thane', 'Naupada, Thane',
       'Kolshet, Thane', 'Dhokali, Thane', 'Balkum village, Thane',
       'Najiwada, Thane', 'Manpada, Thane', 'Hiranandani Estate, Thane',
       'Bhayandarpada, Thane', 'Wagle Industrial Estate, Thane',
       'Vasant Vihar, Thane', 'Panch Pakhdhi, Thane',
       'Tean Nath Naka, Thane', 'Patlipada, Thane', 'Brahmand, Thane',
       'Waghbil, Thane', 'Thane West, Bhiwandi', 'Balkum Naka, Thane',
       'Koliwada, Thane', 'Kolbad, Thane', 'Vartak Nagar, Thane',
       'Naupada, Mumbai', 'Savarkar Nagar, Thane',
       'Vrindavan Society, Thane', 'Anand Nagar Thane West, Thane',
       'Khopat, Thane', 'Pokharan Road Number 1, Thane',
       'Kapurbawdi, Thane', 'Shivai Nagar, Thane', 'Charai, Thane',
       'Ram Maruti Road Naupada, Thane', 'Rabodi, Thane',
       'Lok Puram, Thane', 'Dhobi Ali, Thane', 'Mahagiri Koliwada, Thane',
       'Lokmanya Nagar, Thane', 'Pawar Nagar, Thane', 'Ulhasagar, Thane'
```

'Somnath Nagar', 'Somnath Nagar', 'Thane',
'Jambli Naka, Thane', 'Shivaji Nagar, Thane',
'Bhaskar Colony, Thane', 'Vishnu Nagar, Thane',
'Louis Wadi, Thane', 'Gawand Baug, Thane',
'Meenatal Thakrey Chowk, Thane', 'Raunak Park, Thane',
'Runwal Nagar, Thane', 'Ghantali, Thane', 'Talav Pali, Thane',
'Vijay Nagar, Thane', 'Harinivas Circle, Thane', 'Kopri, Thane',
'Thane East, Thane', 'Sion West, Mumbai', 'Chembur East, Mumbai',
'Charai, Mumbai', 'Sion East, Mumbai', 'Sion Kolwada, Mumbai',
'Sion East, Gurgaon', 'Santacruz West, Mumbai',
'Hasmukh Nagar, Mumbai', 'Willingdon Colony, Mumbai',
'North Ave, Mumbai', 'Saraswat Colony, Mumbai',
'NSES Colony, Mumbai', 'Vithaldas Nagar, Mumbai',
'Dadar East, Mumbai', 'Hindu Colony, Mumbai',
'Dadar TT Circle, Mumbai', 'Nanchanjali Edhalji Joshi Chowk, Mumbai',
'L Tilak Colony, Mumbai', 'Santacruz East, Mumbai',
'Kalina, Mumbai', 'Vidya Nagar, Mumbai', 'Sen Nagar, Mumbai',
'Krantii Nagar, Mumbai', 'Sundai Nagar, Kalina, Mumbai',
'Prabhakar Colony, Mumbai', 'Kolivary, Mumbai',
'Anand Nagar Santacruz East, Mumbai', 'Prabhadevi, Mumbai',
'Adv RD Kowte Chowk, Mumbai', 'Century Bazaar, Mumbai',
'Shree Siddhivinayak Mandir Chowk, Mumbai',
'Babar Shaikh Chowk, Mumbai', 'Mira Road East, Mumbai',
'Nahajan Wadi, Mumbai', 'Geeta Nagar Mira Road, Mumbai',
'Penkarpada, Mumbai', 'Miragaon, Mumbai',
'Nalasopara West, Mumbai', 'Laxmibai Chheda Nagar, Mumbai',
'Samel Pada, Mumbai', 'Nalasopara West, Thane',
'Nalasopara East, Mumbai', 'Osval Nagar, Mumbai',
'Naigaon West, Mumbai', 'Naigaon East, Mumbai',
'Goregaon East, Mumbai', 'Gokuldharm, Mumbai',
'Film City Road, Mumbai', 'Jayprakash Nagar, Mumbai',
'Bimbisar Nagar, Mumbai', 'Ganesh Nagar, Mumbai',
'Kanya Pada, Mumbai', 'Nagari Niwaro Parishad, Mumbai',
'Sai Baba Complex, Mumbai', 'Aarey Milk Colony, Mumbai',
'NNP Colony, Mumbai', 'Shivaji Nagar, Mumbai',
'Royal Palms Estate, Mumbai', 'Laxmi Nagar, Mumbai',
'Pandurang Wadi, Mumbai', 'Subhash Nagar, Mumbai',
'Churi Wadi, Mumbai', 'Azad Nagar, Mumbai',
'Shreyas Colony, Mumbai', 'Durgeshwari Welfare Society, Mumbai',
'Film City, Mumbai', 'Yeshodham, Mumbai',
'Indian Oil Nagar- Shivaji Nagar, Mumbai', 'Ram Nagar, Mumbai',
'CANA Industrial Estate, Mumbai', 'Yashodham, Mumbai',
'IIT Bhatti, Mumbai', 'Vishveshwar Nagar, Mumbai',
'Agricpada, Mumbai', 'Mumbai Central, Mumbai', 'RTO Colony, Mumbai',
'Churchgate, Mumbai', 'Matunga West, Mumbai', 'Joshi Wadi, Mumbai',
'Mahalaxmi Sindhi Colony, Mumbai', 'S Haralaya Chowk, Mumbai',
'Natunga East, Mumbai', 'Brhmanwada, Mumbai',
'Five Gardens, Mumbai', 'Kings Circle, Mumbai',
'Marine Lines East, Mumbai', 'Marine Lines, Mumbai',
'Churchgate station, Mumbai', 'Marine Drive, Mumbai',
'Malad East, Mumbai', 'Pathanwadi, Mumbai',
'Kurkar Village, Mumbai', 'Triveni Nagar, Mumbai',
'Pushpa Park, Mumbai', 'Shivaji Nagar Malad East, Mumbai',
'Primal Nagar, Mumbai', 'Dindoshi, Mumbai', 'Kokani Pada, Mumbai',
'Upper Govind Nagar, Mumbai', 'Dhanji Wadi, Mumbai',
'Pathan Wadi Chowk, Mumbai', 'Dombivli West, Thane',
'Kopargaon, Thane', 'Mahim West, Mumbai', 'Asavari, Mumbai',
'Mahim East, Mumbai', 'Mahalaxmi Race Course, Mumbai',
'Lower Parel West, Mumbai', 'Chembur West, Mumbai',
'Rahul Nagar Chembur, Mumbai', 'Lower Parel East, Mumbai',
'Khar West, Mumbai', 'Juhu Kolwada, Mumbai', 'Old Khar, Mumbai',
'Khar Danda, Mumbai', 'Danda, Mumbai', 'Golibar, Mumbai',
'RPF Colony, Mumbai', 'Khar East, Mumbai',
'Kandivali West, Mumbai', 'Mahavir Nagar, Mumbai',
'Mahatama Gandhi Road Amrut Nagar, Mumbai',
'Dhanukar Wadi, Mumbai', 'Charkop Sector 2, Mumbai',
'Irani Wada, Mumbai', 'Charkop Sector 5, Mumbai',
'Charkop Industrial Estate, Mumbai', 'Charkop, Mumbai',
'Bander Pakhad, Mumbai', 'Kandivali West, Nagpur',
'Charkop Sector 9, Mumbai', 'Charkop Sector 3, Mumbai',
'Charkop Village, Mumbai', 'Charkop Sector 8, Mumbai',
'Charkop Sector 6, Mumbai', 'Charkop Sector 1, Mumbai',
'Charkop Sector 7, Mumbai', 'Jethava Nagar, Mumbai',
'Ekta Nagar Charkop, Mumbai', 'Shankar Pada, Mumbai',
'Goraswadi, Mumbai', 'Patel Nagar, Mumbai',
'Ganesh Nagar JJC Area, Mumbai', 'Fateh Baug, Mumbai',
'Kandivali East, Mumbai', 'Lokhandwala Twp, Mumbai',
'Alka Nagar, Mumbai', 'Thakur Village, Kandivali East, Mumbai',
'Akuli Nagar, Mumbai', 'Samata Nagar, Mumbai',
'Ashok Nagar Western Mumbai, Mumbai', 'Thakur Complex, Mumbai',
'Asha Nagar, Mumbai', 'Appa Pada, Mumbai', 'Damodarwadi, Mumbai',
'Dattani Park, Mumbai', 'Singh Aghri Estate, Mumbai',
'Amboli, Mumbai', 'Laxmi Industrial Estate, Mumbai',
'Andheri West, Mumbai', 'Lokhandwala Complex, Mumbai',
'Four Bungalows, Mumbai', 'Seven Bungalows, Mumbai',
'Veera Desai Road, Mumbai', '4 Bunglows, Mumbai',
'Lallu Bhai Park, Mumbai', 'Azad Nagar 2, Mumbai',
'D.N. Nagar, Mumbai', 'Andheri West, Nagpur',
'Versova Marg, Mumbai', 'DN Nagar, Mumbai', 'Millat Nagar, Mumbai',
'Gilbert Hill JJC, Mumbai', 'SV Patel Nagar, Mumbai',
'Shastry Nagar D Phase, Mumbai',
'Shri Swami Samarth Nagar, Mumbai', 'Navneeth Colony, Mumbai',
'Sundervan Complex, Mumbai', 'Dattaguru nagar, Mumbai',
'Andheri East, Mumbai', 'Marol, Mumbai', 'Chakala, Mumbai',
'Marol Naka, Mumbai', 'Saki Vihar Road, Mumbai',
'JB Nagar, Mumbai', 'Poornam Nagar, Mumbai',
'Marol Maroshi Road, Mumbai', 'Sher E Punjab Society, Mumbai',
'Sakinaka Junction, Mumbai', 'Kadam Wadi Marol, Mumbai',
'Seepz Area, Mumbai', 'Chakala MIDC, Mumbai',
'Dr Charatsingh Colony, Mumbai', 'Andheri East, Nagpur',
'Malpa Dongri, Mumbai', 'Gondvali Gaonthan, Mumbai',
'Kajiwadi, Mumbai', 'Marol Maroshi, Mumbai', 'PNGP Colony, Mumbai',
'Nogra Village Road, Mumbai', 'Kanti Nagar, Mumbai',
'Bamandaya Pada, Mumbai', 'Koladongri, Mumbai',
'Nogra Village, Mumbai', 'Ujjamatma Road, Mumbai',
'Hanuman Nagar, Mumbai', 'Azad Nagar JJC, Mumbai',

```

'Parsi Wada, Mumbai', 'Marol Naka Junction, Mumbai',
'Bhavani Nagar, Mumbai', 'Nareshwari Nagar, Mumbai',
'Kondivita Village, Mumbai', 'Vijay Nagar, Hyderabad'],
dtype=object)

[58]: # Define the mapping dictionary
Location_count=data[["Location"]].value_counts()
data[["Location_encoded"]]=data[["Location"]].map(Location_count)
print(data[["Location","Location_encoded"]].head())

   Location  Location_encoded
0  Kalyan West, Thane           199
1  Kalyan West, Thane           199
2  Kalyan West, Thane           199
3  Kalyan West, Thane           199
4  Kalyan West, Thane           199

[59]: data.drop(columns=["Location"],inplace=True)

[60]: #Ownership Type Column
data[["Ownership Type"]].unique()

[60]: array(['Freehold', 'Co-operative Society', 'Power Of Attorney',
       'Leasehold'], dtype=object)

[61]: data[["Ownership Type_encoded"]]=data[["Ownership Type"]].replace({'Freehold':0, 'Co-operative Society':1, 'Power Of Attorney':2,
       'Leasehold':3})

C:\Users\deshm\AppData\Local\Temp\ipykernel_38268\1978593805.py:1: FutureWarning: Downcasting behavior in 'replace' is deprecated and will be removed in a future version. To retain the old behavior, explicitly call `result.infer_objects(copy=False)`. To opt-in to the future behavior, set `pd.set_option('future.no_silent_downcasting', True)`
    data[["Ownership Type_encoded"]]=data[["Ownership Type"]].replace({'Freehold':0, 'Co-operative Society':1, 'Power Of Attorney':2},

[62]: data.drop(columns=["Ownership Type"],inplace=True)

[63]: #Ownership Type Column
data[["Parking"]].unique()

[63]: array(['1 Covered', '1 Open', '1 Covered, 1 Open', '2 Covered', '2 Open',
       '5 Open', '3 Covered', '4 Covered', '10 Covered', '102 Covered',
       '2 Covered, 2 Open', '10 Open', '6 Open', '32 Covered', '180 Covered',
       '306 Open', '8 Covered', '20 Open', '18 Open', '50 Open',
       '9 Open', '15 Open', '5 Covered', '100 Open', '50 Open',
       '10 Covered, 20 Open', '12 Covered', '7 Open', '40 Open',
       '503 Open', '1 Covered, 2 Open', '2 Covered, 1 Open', '9 Covered',
       '3 Open', '505 Covered', '47 Covered', '7 Covered, 1 Open',
       '4 Open', '14 Open', '20 Covered', '8 Open',
       '606 Covered, 606 Open', '10 Covered, 10 Open', '111 Open',
       '138 Covered', '392 Covered', '7 Covered', '18 Covered',
       '27 Covered', '181 Covered', '6 Covered, 6 Open', '801 Covered',
       '2 Covered, 4 Open', '16 Covered', '1 Covered, 3 Open', '12 Open',
       '50 Covered', '3 Covered, 1 Open', '25 Open', '408 Covered',
       '1 Covered, 10 Open', '45 Covered', '3 Covered, 2 Open',
       '107 Covered', '150 Covered', '200 Covered', '22 Covered',
       '16 Open', '250 Covered, 200 Open', '24 Covered, 21 Open',
       '76 Open', '904 Covered', '303 Open', '23 Open', '161 Open',
       '180 Covered', '43 Open', '34 Covered', '1 Covered, 504 Open',
       '5 Covered, 2 Open', '120 Covered', '35 Open', '30 Open',
       '502 Covered', '103 Covered'], dtype=object)

[64]: # Define the mapping dictionary
Parking_count=data[["Parking"]].value_counts()
data[["Parking_encoded"]]=data[["Parking"]].map(Parking_count)
print(data[["Parking","Parking_encoded"]].head())

   Parking  Parking_encoded
0      1 Covered          9064
1      1 Covered          9064
2        1 Open            1575
3  1 Covered, 1 Open          525
4      1 Covered          9064

[65]: data.drop(columns=["Parking"],inplace=True)

[66]: #Possession Status Column
data[["Possession Status"]].unique()

[66]: array(['Under Construction', 'Ready to Move', "Jun '26", "Dec '28",
       "Nov '25", "Dec '26", "Mar '25", "Dec '23", "Jun '24", "Dec '24",
       "Sep '23", "Mar '24", 'Immediately', "Dec '27", "Dec '25",
       "Apr '28", "Nov '23", "Apr '24", "Jun '25", 'Select Date',
       "Sep '26"], dtype=object)

[67]: data[["Possession Status_encoded"]]=data[["Possession Status"]].replace({'Under Construction':0, 'Ready to Move':1, "Jun '26":2, "Dec '28":3,
       "Nov '25":4, "Dec '26":5, "Mar '25":6, "Dec '23":7, "Jun '24":8, "Dec '24":9,
       "Sep '23":10, "Mar '24":11, 'Immediately':12, "Dec '27":13, "Dec '25":14,
       "Apr '28":15, "Nov '23":16, "Apr '24":17, "Jun '25":18, 'Select Date':19,
       "Sep '26":20})

C:\Users\deshm\AppData\Local\Temp\ipykernel_38268\3017873097.py:1: FutureWarning: Downcasting behavior in 'replace' is deprecated and will be removed in a future version. To retain the old behavior, explicitly call `result.infer_objects(copy=False)`. To opt-in to the future behavior, set `pd.set_option('future.no_silent_downcasting', True)`
    data[["Possession Status_encoded"]]=data[["Possession Status"]].replace({'Under Construction':0, 'Ready to Move':1, "Jun '26":2, "Dec '28":3},

[68]: data.drop(columns=["Possession Status"],inplace=True)

[69]: data.dtypes

[69]: Price          float64
Covered Area      float64
Carpet Area        float64
City              object
Type of Property   object
bedroom           int64
Bathroom          float64
balcony           float64

```

```
Society          object
isPrimeLocationProperty    object
Maintenance Charges      float64
City_encoded            int64
Location_encoded        int64
Ownership Type_encoded   int64
Parking_encoded          int64
Possession Status_encoded int64
dtype: object

[70]: #Society Column
data["Society"].unique()

[70]: array(['Y', 'N'], dtype=object)

[71]: data["Society_encoded"] = data["Society"].replace({'Y':0, 'N':1})

C:\Users\deshm\AppData\Local\Temp\ipykernel_38268\4023502697.py:1: FutureWarning: Downcasting behavior in `replace` is deprecated and will be removed in a future version. To retain the old behavior, explicitly call `result.infer_objects(copy=False)`. To opt-in to the future behavior, set `pd.set_option('future.no_silent_downcasting', True)`
    data["Society_encoded"] = data["Society"].replace({'Y':0, 'N':1})

[72]: data.drop(columns=["Society"], inplace=True)

[73]: # Type of Property Column
data["Type of Property"].unique()

[73]: array(['Apartment', 'Residential House', 'Builder Floor Apartment',
       'Penthouse', 'Villa'], dtype=object)

[74]: data["Type of Property_encoded"] = data["Type of Property"].replace({'Apartment':0, 'Residential House':1, 'Builder Floor Apartment':2,
       'Penthouse':3, 'Villa':4})

C:\Users\deshm\AppData\Local\Temp\ipykernel_38268\342787082.py:1: FutureWarning: Downcasting behavior in `replace` is deprecated and will be removed in a future version. To retain the old behavior, explicitly call `result.infer_objects(copy=False)`. To opt-in to the future behavior, set `pd.set_option('future.no_silent_downcasting', True)`
    data["Type of Property_encoded"] = data["Type of Property"].replace({'Apartment':0, 'Residential House':1, 'Builder Floor Apartment':2,
       'Penthouse':3, 'Villa':4})

[75]: data.drop(columns=["Type of Property"], inplace=True)
```

localhost:8888/notebooks/EDA_RealEstate.ipynb

jupyter EDA_RealEstate Last Checkpoint: 20 hours ago

File Edit View Run Kernel Settings Help Trusted JupyterLab Python 3 (pykernel)

```
[73]: data.drop(columns=["Type of Property"], inplace=True)

[74]: # Type of Property Column
data["isPrimeLocationProperty"].unique()

[75]: array(['Y', 'N'], dtype=object)

[76]: data["isPrimeLocationProperty_encoded"] = data["isPrimeLocationProperty"].replace({'Y':0, 'N':1})

C:\Users\deshm\AppData\Local\Temp\ipykernel_1896\1666895389.py:1: FutureWarning: Downcasting behavior in `replace` is deprecated and will be removed in a future version. To retain the old behavior, explicitly call `result.infer_objects(copy=False)`. To opt-in to the future behavior, set `pd.set_option('future.na_silent_downcasting', True)`
    data["isPrimeLocationProperty_encoded"] = data["isPrimeLocationProperty"].replace({'Y':0, 'N':1})

[77]: data.drop(columns=["isPrimeLocationProperty"], inplace=True)

[77]: data
```

	Price	Covered Area	Carpet Area	City	bedroom	Bathroom	Maintenance Charges	City_encoded	Location_encoded	Ownership Type_encoded	Parking_encoded	Possession Status_encoded	So
0	3150000.0	635.0	375.0	Thane	1	2.0	3.0	0	199	0	9064	0	
1	6300000.0	579.0	579.0	Thane	2	2.0	3.0	0	199	0	9064	1	
2	5400000.0	850.0	555.0	Thane	2	2.0	1200.0	0	199	0	1575	1	
3	9000000.0	1050.0	815.0	Thane	2	2.0	3200.0	0	199	0	525	0	

SH14 / Universit... Closed road

Search ENG IN 16:14 15-01-2023

Home EDA_RealEstate EDA_RealEstate.ipynb

jupyter EDA_RealEstate Last Checkpoint: 20 hours ago Trusted

File Edit View Run Kernel Settings Help JupyterLab Python 3 (pykernel)

	Price	Covered Area	Carpet Area	bedroom	Bathroom	Maintenance Charges	City_encoded	Location_encoded	Ownership Type_encoded	Parking_encoded	Possession Status_encoded	Society_encc
0	8150000.0	635.0	875.0	1	2.0	3.0	0	199	0	9064	0	
1	6300000.0	579.0	579.0	2	2.0	3.0	0	199	0	9064	1	
2	5400000.0	850.0	585.0	2	2.0	1200.0	0	199	0	1575	1	

12685 rows × 15 columns

[78]: data.drop(columns=["City"], inplace=True)

[79]: data

SH114 / Universit... Closed road ENG IN 16:14 15-01-2025

localhost:8888/notebooks/EDA_RealEstate.ipynb

jupyter EDA_RealEstate Last Checkpoint: 20 hours ago

File Edit View Run Kernel Settings Help Trusted

JupyterLab Python 3 (pykernel)

	3	9000000.0	1050.0	815.0	2	2.0	3200.0	0	199	0	325	0
4	4950000.0	561.0	419.0	1	2	2.0	2500.0	0	199	0	9064	0
...
12680	22000000.0	850.0	618.0	2	2	2.0	2500.0	3	140	0	9064	5
12681	31000000.0	1376.0	1076.0	3	2	2.0	12.0	3	140	0	9064	0
12682	18500000.0	970.0	616.0	2	2	2.0	9.0	3	140	0	9064	0
12683	22000000.0	850.0	821.0	2	2	2.0	4000.0	3	421	0	1575	1
12684	18500000.0	685.0	665.0	2	2	2.0	10.0	3	421	0	9064	0

12685 rows × 14 columns

```
[80]: # Dependent Column (target)
data = data.loc[:, data.columns.difference(["Price"]).tolist() + ["Price"]]

[81]: corr = data.iloc[:, 5:14].corr()
sns.pairplot(corr)

[81]: <seaborn.axisgrid.PairGrid at 0x237d8cd6720>
```

NFO - LTC Video highlight

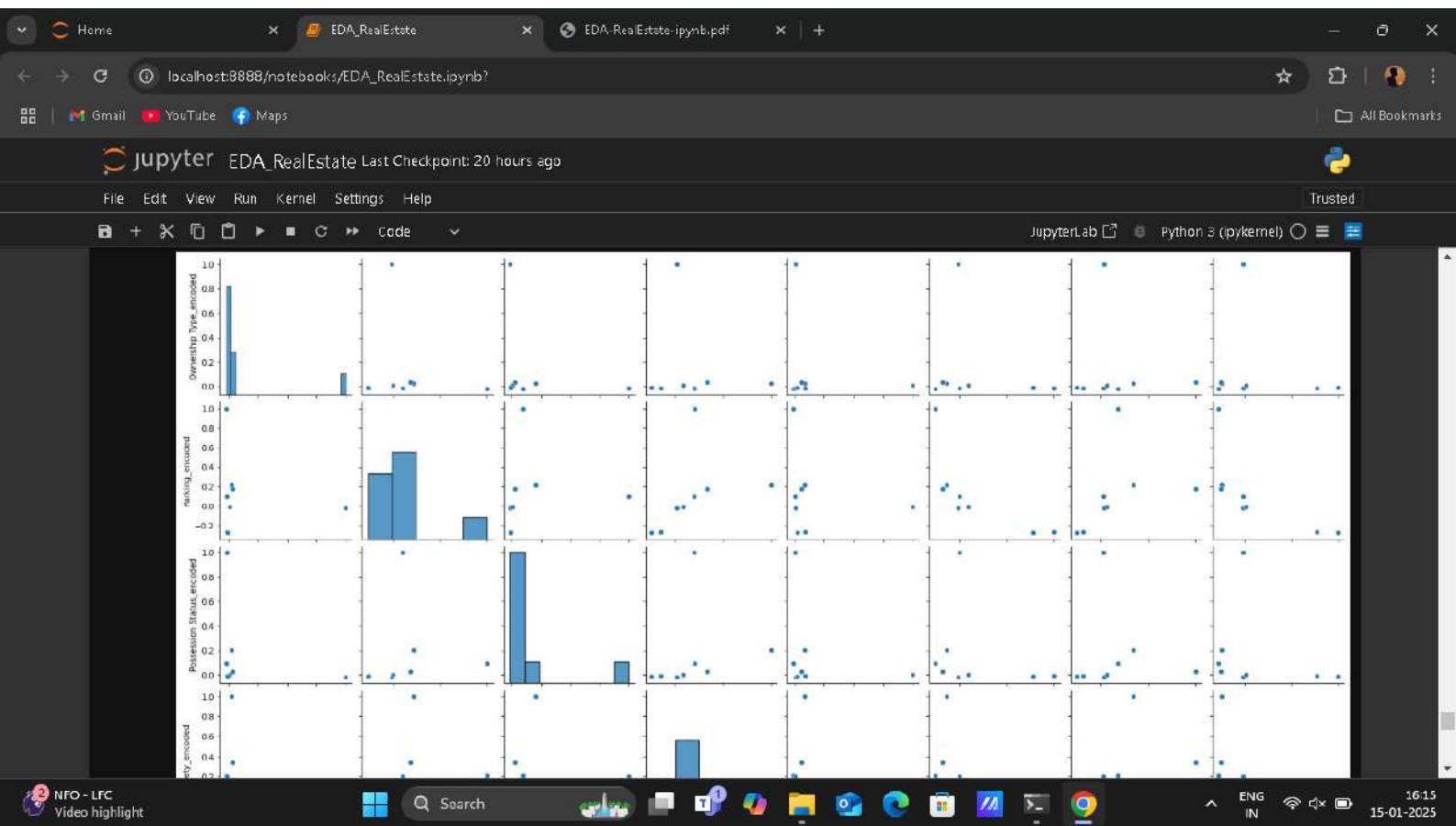
Search

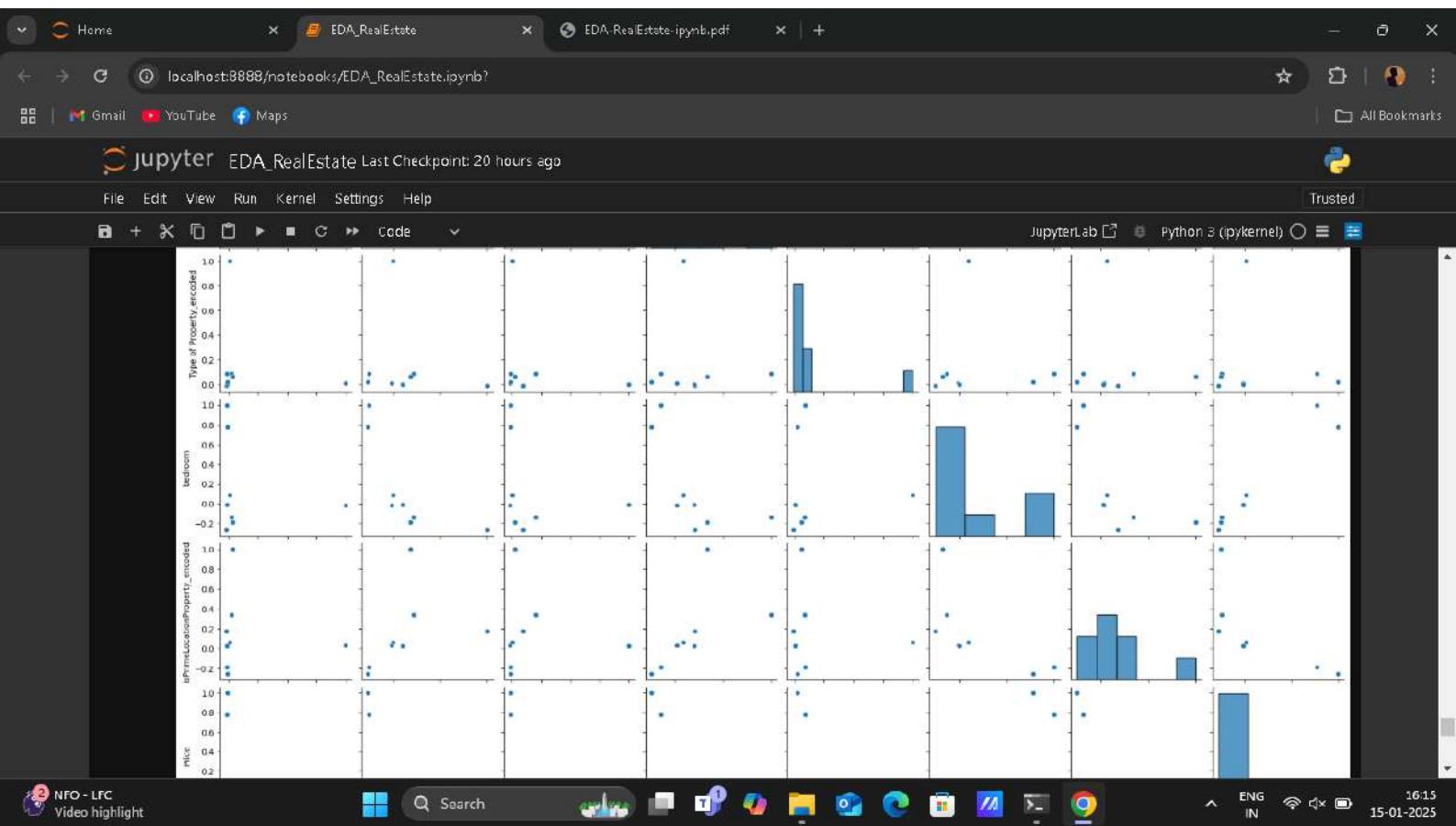
10

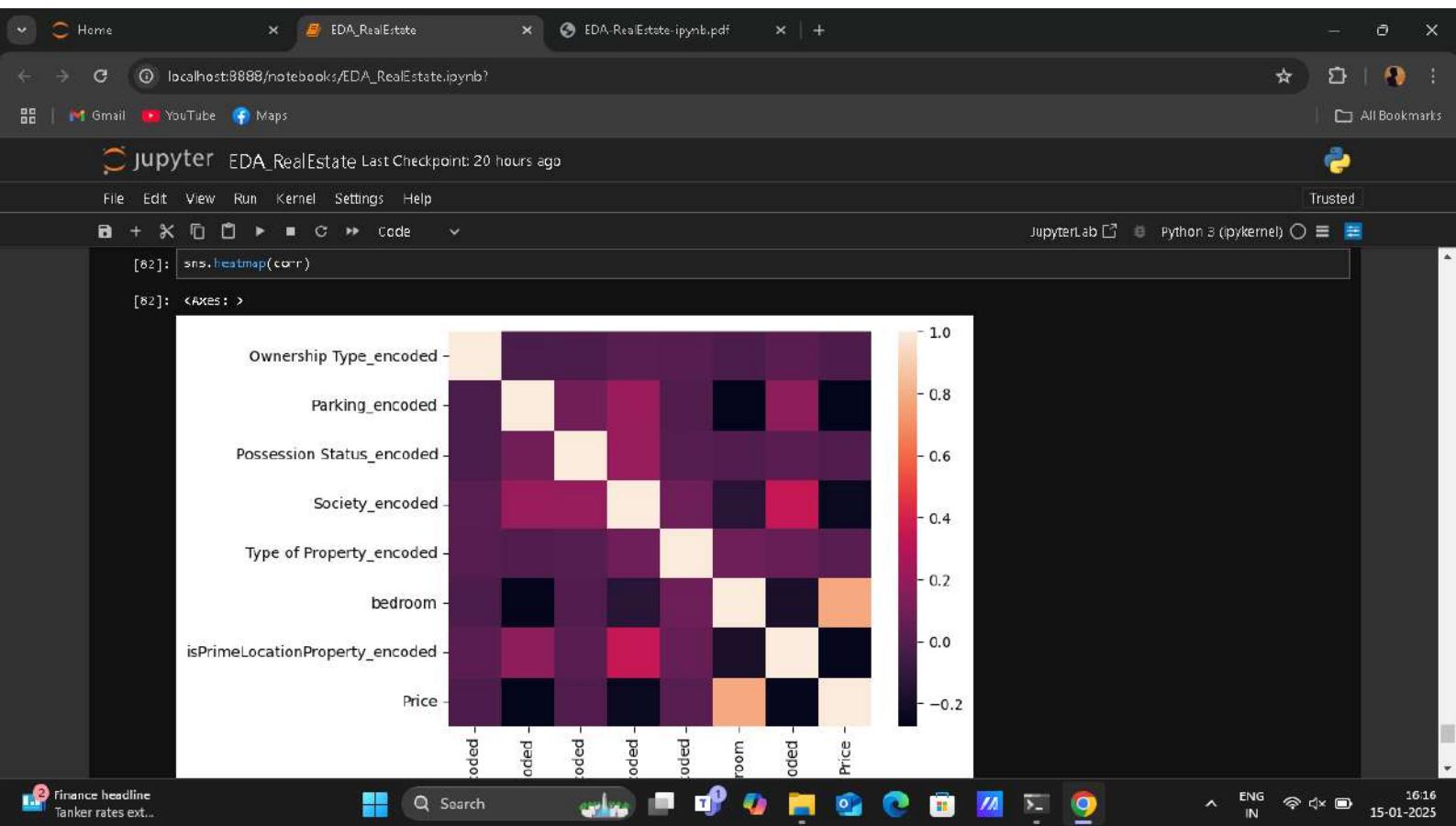
ENG IN

16:15

15-01-2025







localhost:8888/notebooks/EDA_RealEstate.ipynb

jupyter EDA_RealEstate Last Checkpoint: 20 hours ago

File Edit View Run Kernel Settings Help Trusted JupyterLab Python 3 (pykernel)

```
[83]: # Train-test split
X = data.drop(columns=['Price']) # Replace 'target_column' with your target variable
Y = data['Price']

xtrain, xtest, ytrain, ytest = train_test_split(X, Y, test_size=0.2, random_state=42)

print(f'Shape of X_train: {xtrain.shape}')
print(f'Shape of X_test: {xtest.shape}')
print(f'Shape of y_train: {ytrain.shape}')
print(f'Shape of y_test: {ytest.shape}')

Shape of X_train: (10148, 13)
Shape of X_test: (2537, 13)
Shape of y_train: (10148,)
Shape of y_test: (2537,)
```

```
[84]: data
```

	Bathroom	Carpet Area	City_encoded	Covered Area	Location_encoded	Maintenance Charges	Ownership Type_encoded	Parking_encoded	Possession Status_encoded	Society_encoded	Type of Property_encoded
0	2.0	375.0	0	635.0	199	3.0	0	9064	0	0	C
1	2.0	579.0	0	579.0	199	3.0	0	9064	1	0	C
2	2.0	585.0	0	850.0	199	1200.0	0	1575	1	0	C
3	2.0	815.0	0	1050.0	199	3200.0	0	525	0	0	C

Finance headline
Tanker rates ext...
Search ENG IN 16:16 15-01-2025

Home EDA_RealEstate EDA-RealEstate.ipynb.pdf

jupyter EDA_RealEstate Last Checkpoint: 20 hours ago Python 3 (pykernel)

File Edit View Run Kernel Settings Help Trusted

12681 2.0 1076.0 3 1376.0 140 12.0 0 9064 0 0 C
12682 2.0 618.0 3 970.0 140 9.0 0 9064 0 0 C
12683 2.0 821.0 3 850.0 421 4000.0 0 1575 1 0 C
12684 2.0 665.0 3 685.0 421 10.0 0 9064 0 0 C

12685 rows × 14 columns

[167]: data.dtypes

```
Bathroom          float64
Carpet Area       float64
City_encoded      int64
Covered Area     float64
Location_encoded int64
Maintenance Charges float64
Ownership Type_encoded int64
Parking_encoded   int64
Possession Status_encoded int64
Society_encoded   int64
Type of Property_encoded int64
bedroom           int64
isPrimeLocationProperty_encoded int64
Price             float64
dtype: object
```

Finance headline
Tanker rates ext...
Search ENG IN 16:16 15-01-2025