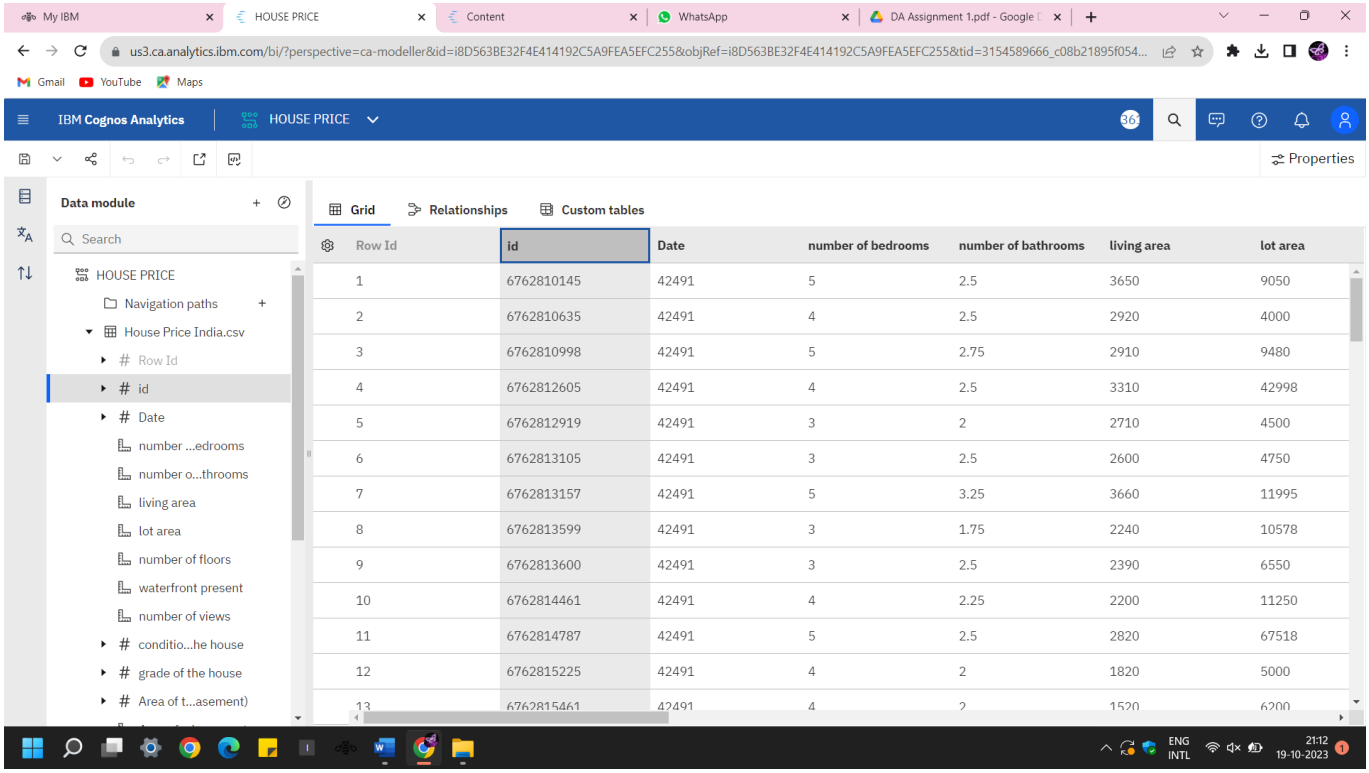
ASSIGNMENT 3



CODE:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
df = pd.read_csv('C:\\Users\\Downloads\\archive (4)\\House Price India.csv')
plt.figure(figsize=(8, 6))
sns.histplot(df['living area'], kde=True)
plt.title('Univariate Analysis ')
plt.show()
```

```
plt.figure(figsize=(8, 6))

sns.scatterplot(x='number of floors', y='condition of the house', data=df)

plt.title('Bi-Variate Analysis ')

plt.show()

sns.pairplot(df[['Area of the house(excluding basement)','grade of the house','Postal Code','living_area_renov']])

plt.title('Multi-Variate Analysis - Pair Plot')

plt.show()

descriptive_stats = df.describe()

print(descriptive_stats)

missing_values = df.isnull().sum()

print("Missing Values:")

print(missing_values)
```

OUTPUT:

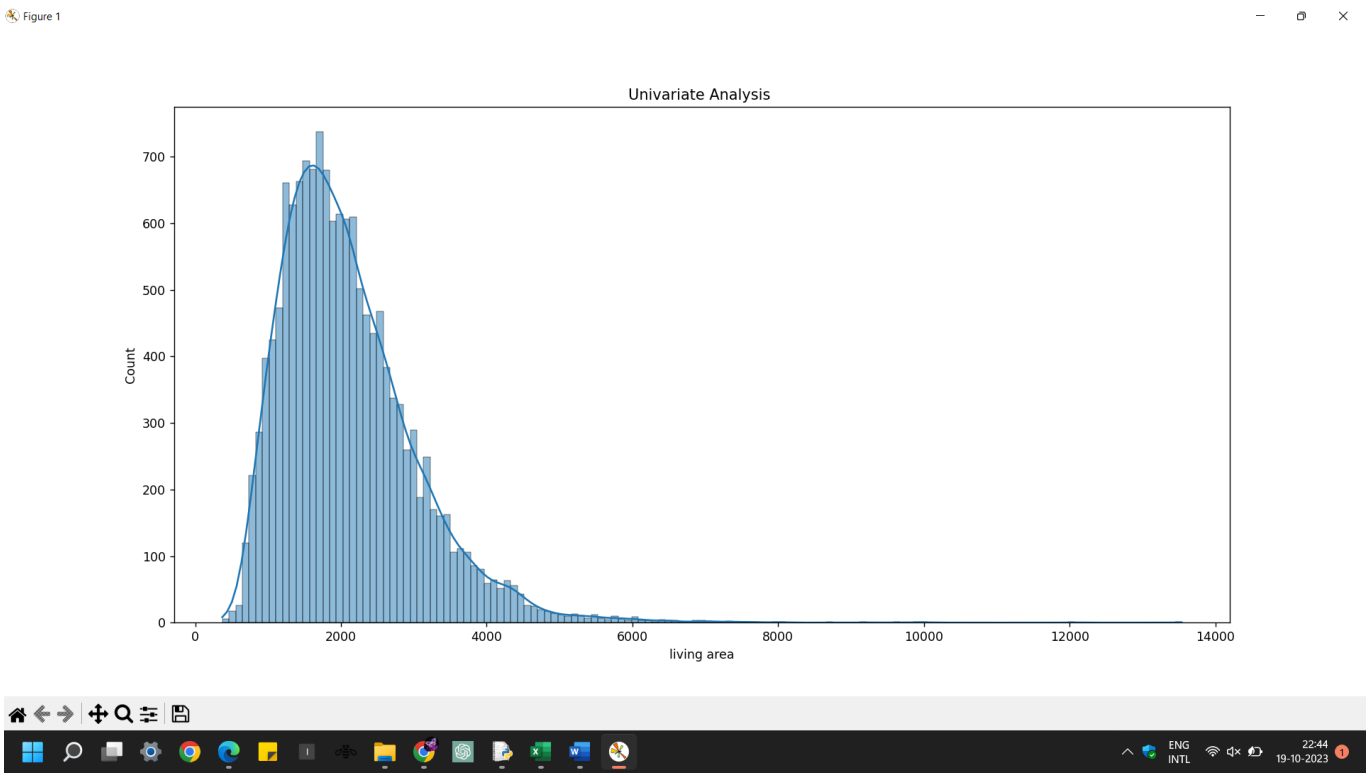


Figure 1
Univariate Analysis
Count
living area
0
100
200
300
400
500
600
700
0
2000
4000
6000
8000
10000
12000
14000

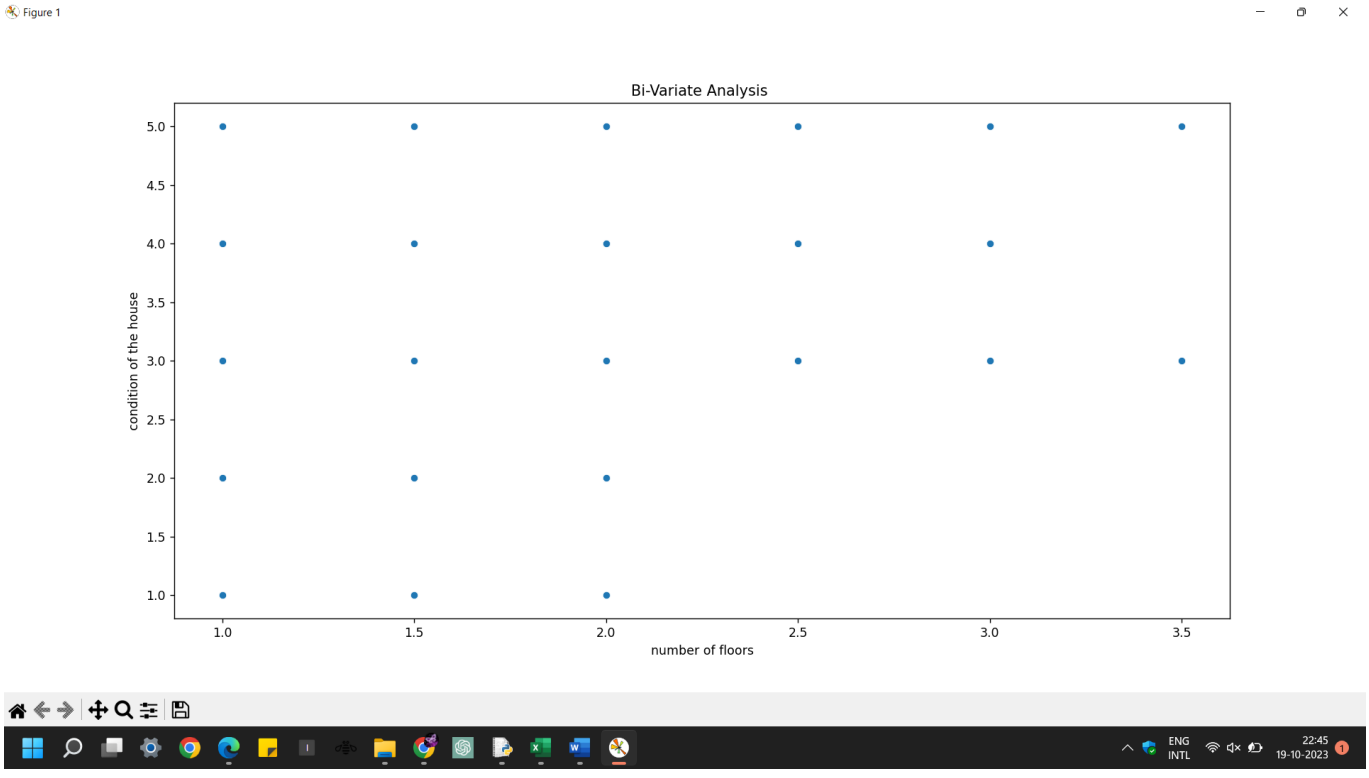


Figure 1
Bi-Variate Analysis
condition of the house
number of floors
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0

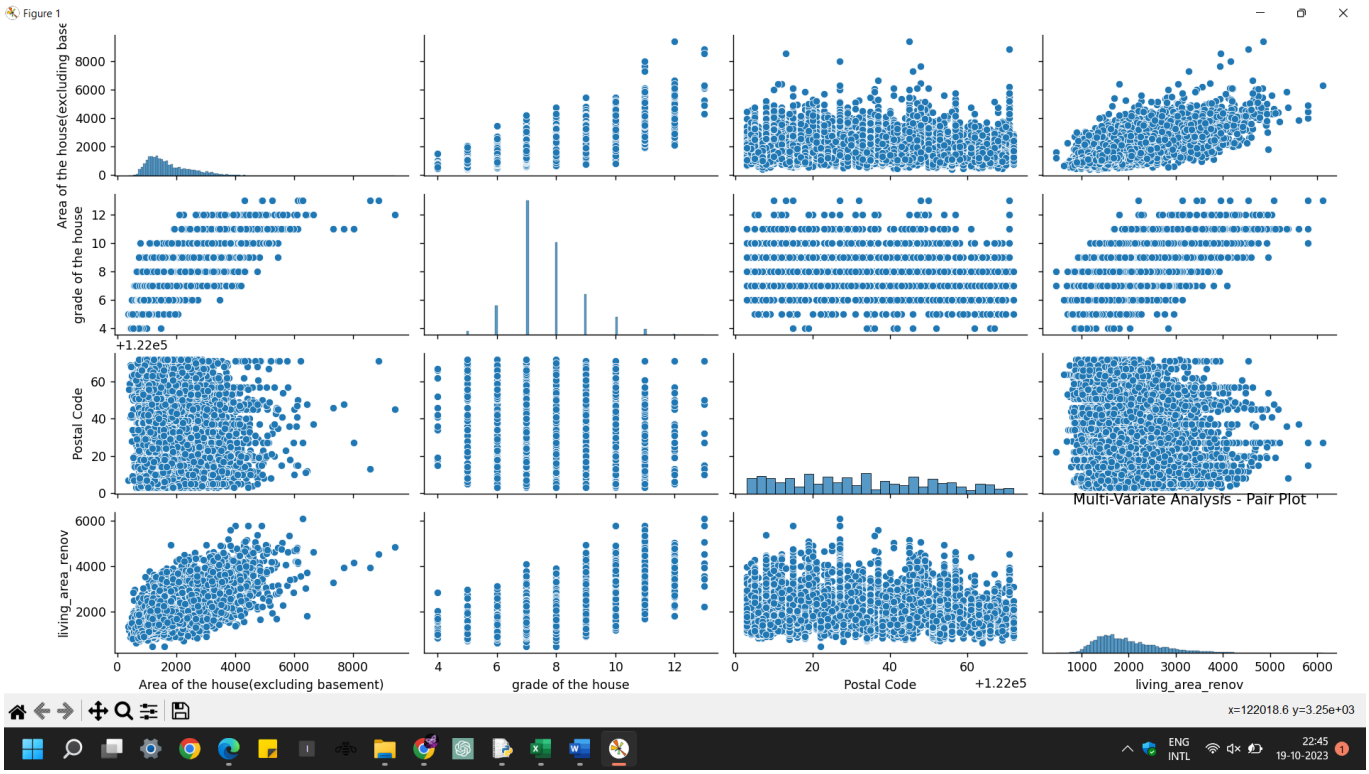
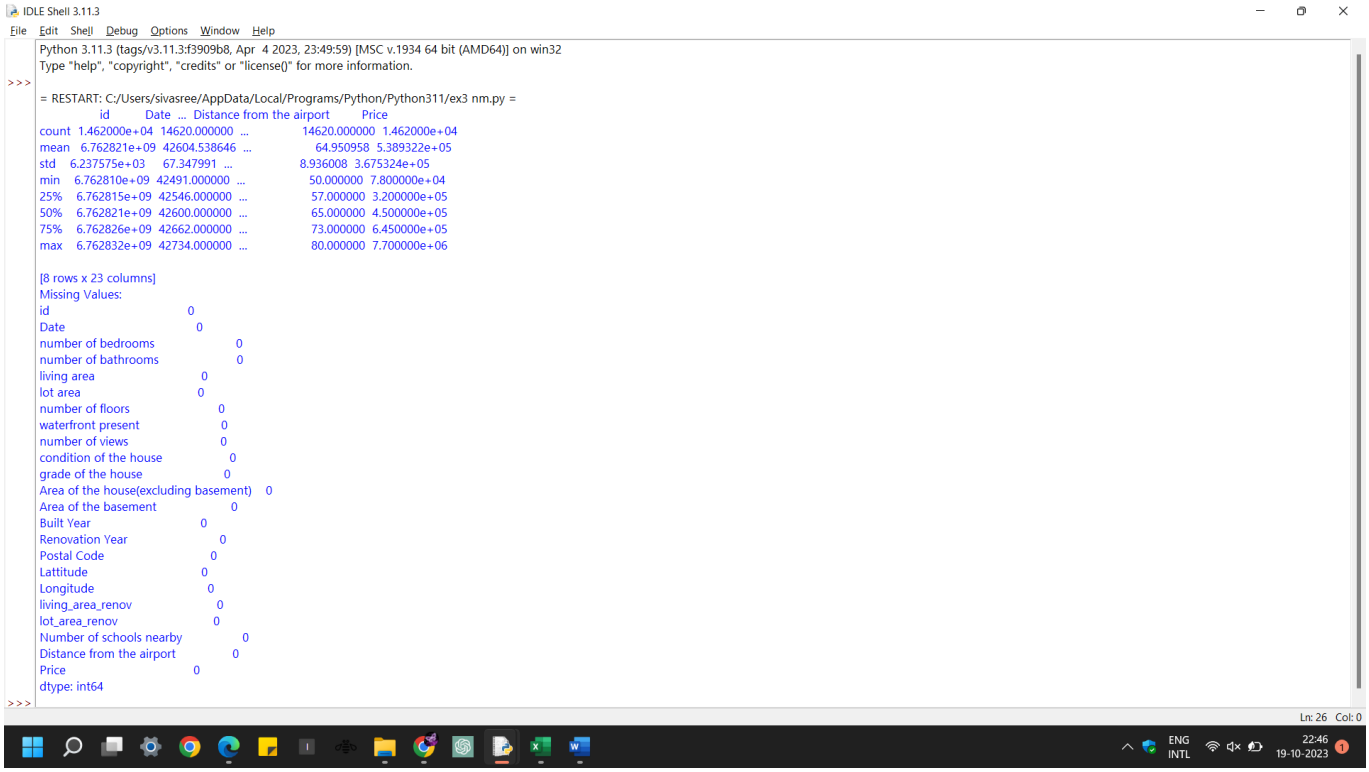


Figure 1
Multi-Variate Analysis - Pair Plot
Area of the house(excluding basement)
grade of the house
Postal Code
+1.22e5
living_area_renov
x=122018.6 y=3.25e+03


IDLE Shell 3.11.3
File Edit Shell Debug Options Window Help
Python 3.11.3 (tags/v3.11.3:f3909b8, Apr 4 2023, 23:49:59) [MSC v.1934 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: C:/Users/sivasree/AppData/Local/Programs/Python/Python311/ex3 nm.py =
id Date ... Distance from the airport Price
count 1.462000e+04 14620.000000 ... 14620.000000 1.462000e+04
mean 6.762821e+09 42604.538646 ... 64.950958 5.389322e+05
std 6.237575e+03 67.347991 ... 8.936008 3.675324e+05
min 6.762810e+09 42491.000000 ... 50.000000 7.800000e+04
25% 6.762815e+09 42546.000000 ... 57.000000 3.200000e+05
50% 6.762821e+09 42600.000000 ... 65.000000 4.500000e+05
75% 6.762826e+09 42662.000000 ... 73.000000 6.450000e+05
max 6.762832e+09 42734.000000 ... 80.000000 7.700000e+06
[8 rows x 23 columns]
Missing Values:
id 0
Date 0
number of bedrooms 0
number of bathrooms 0
living area 0
lot area 0
number of floors 0
waterfront present 0
number of views 0
condition of the house 0
grade of the house 0
Area of the house(excluding basement) 0
Area of the basement 0
Built Year 0
Renovation Year 0
Postal Code 0
Lattitude 0
Longitude 0
living_area_renov 0
lot_area_renov 0
Number of schools nearby 0
Distance from the airport 0
Price 0
dtype: int64
>>>
Ln: 26 Col: 0