

ARJUN COLLEGE OF TECHNOLOGY

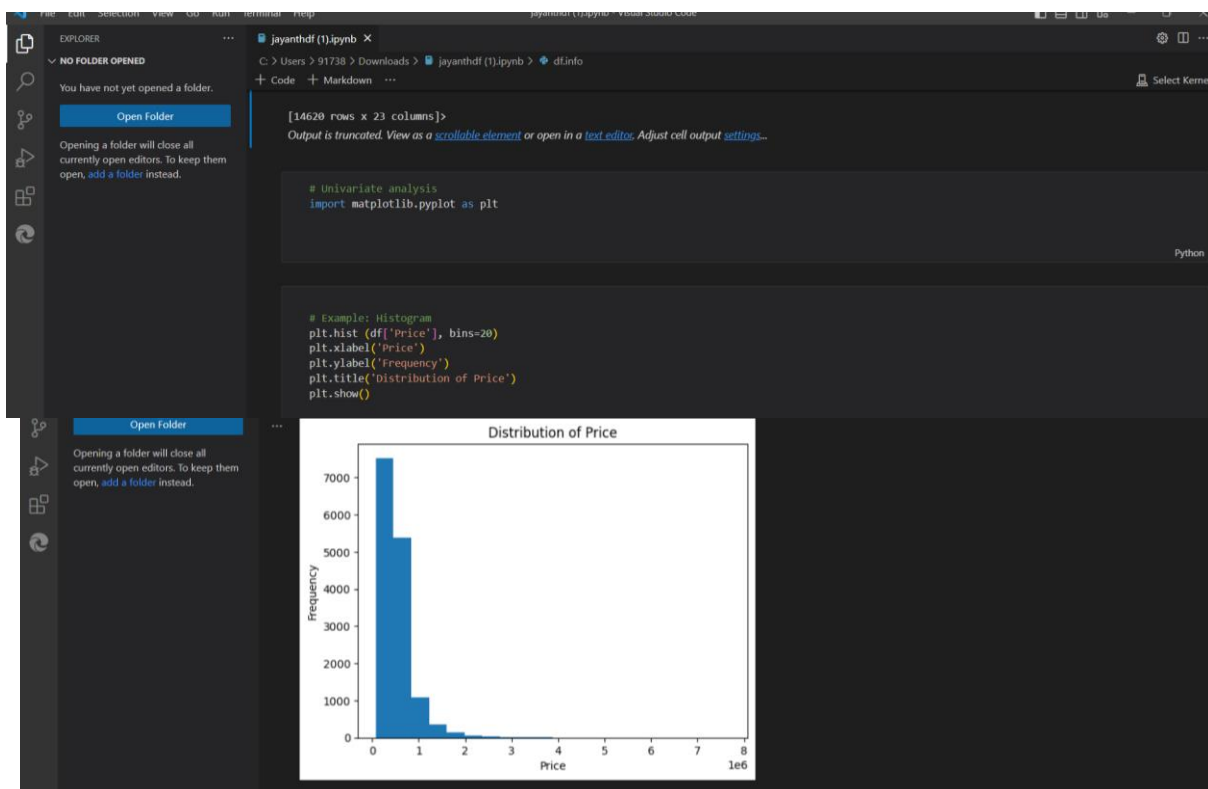
ASSIGNMENT – 3

NAAN MUDHALVAN

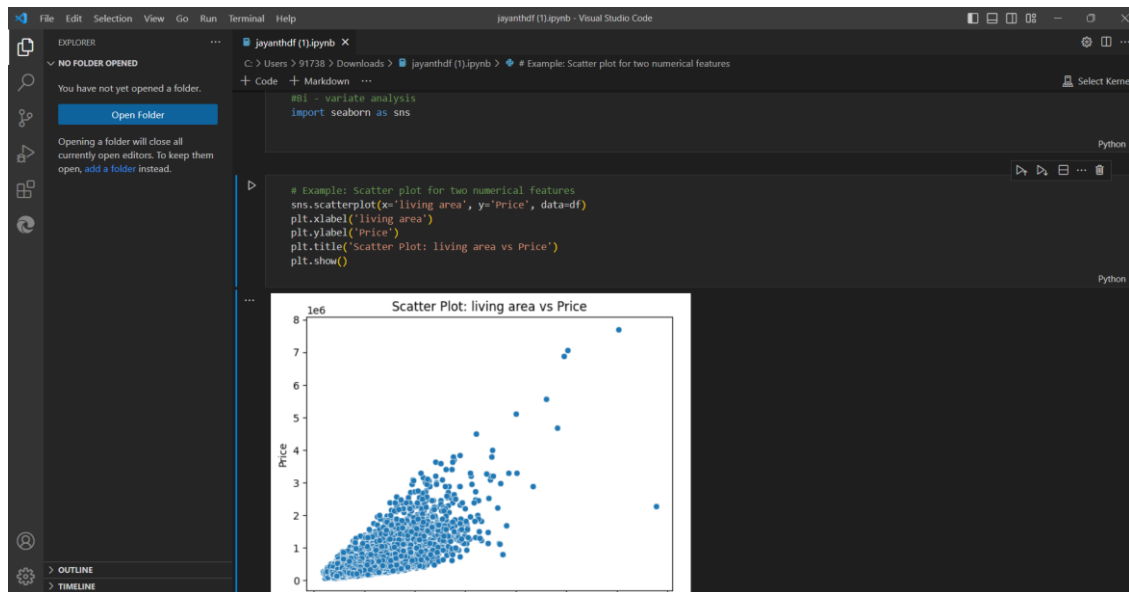
NAME : ROLLA JAYANTH REDDY

REG.NO : 723920104048

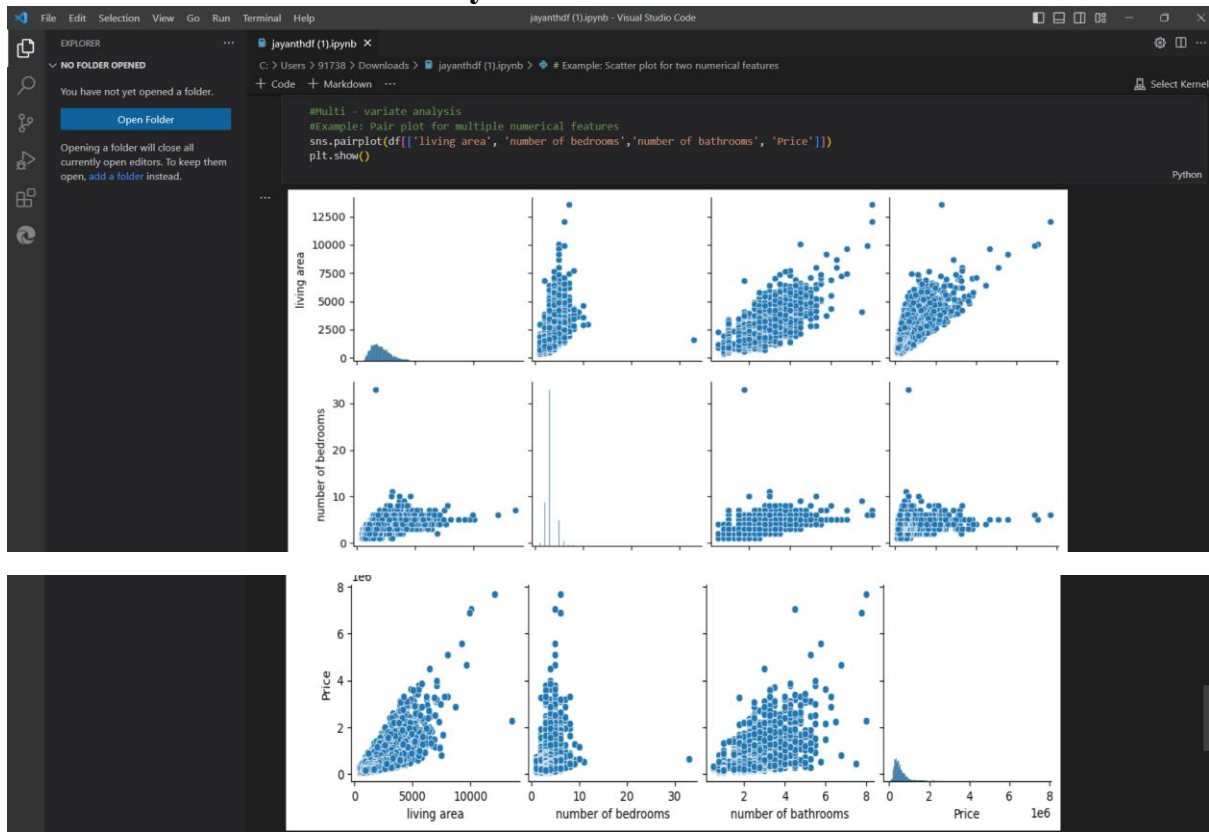
STEP-1: Univariate Analysis



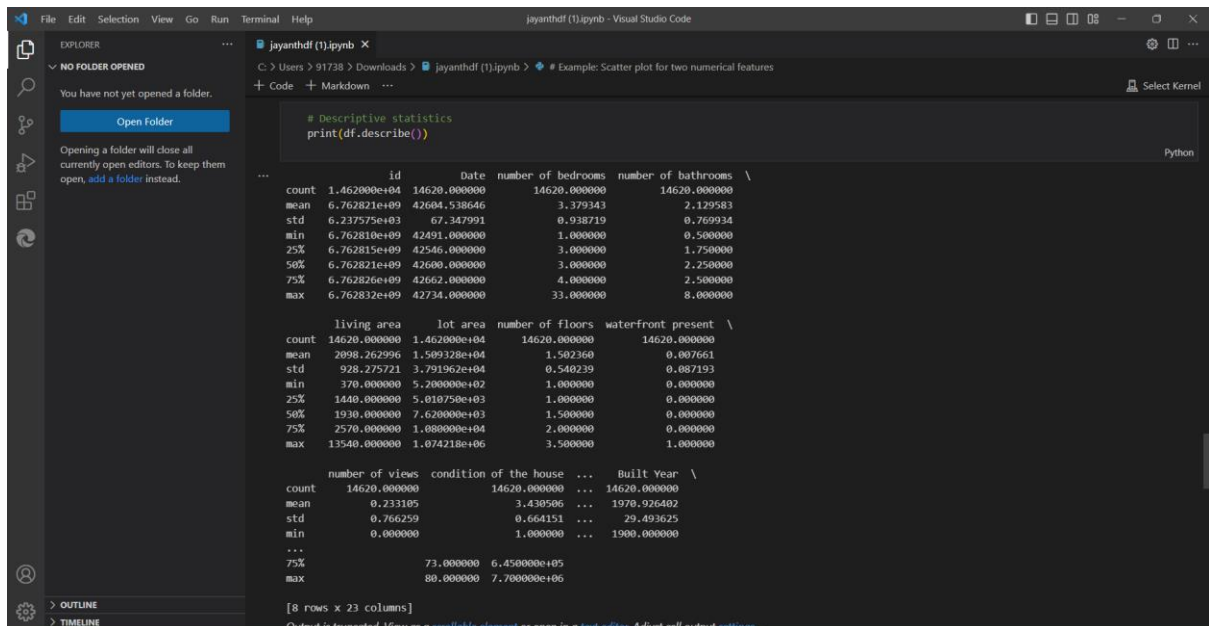
STEP-2: Bi - variate Analysis



STEP-3: Multi - variate Analysis



STEP-4: Descriptive statistics



The screenshot shows a Visual Studio Code window with a Jupyter Notebook titled 'jayanthdf (1).ipynb'. The notebook is open to a cell containing the following code:

```
# Descriptive statistics
print(df.describe())
```

The output of the code is displayed below the cell, showing the statistical summary for the dataset. The output is truncated, indicating that there are 8 rows x 23 columns in total. The output is as follows:

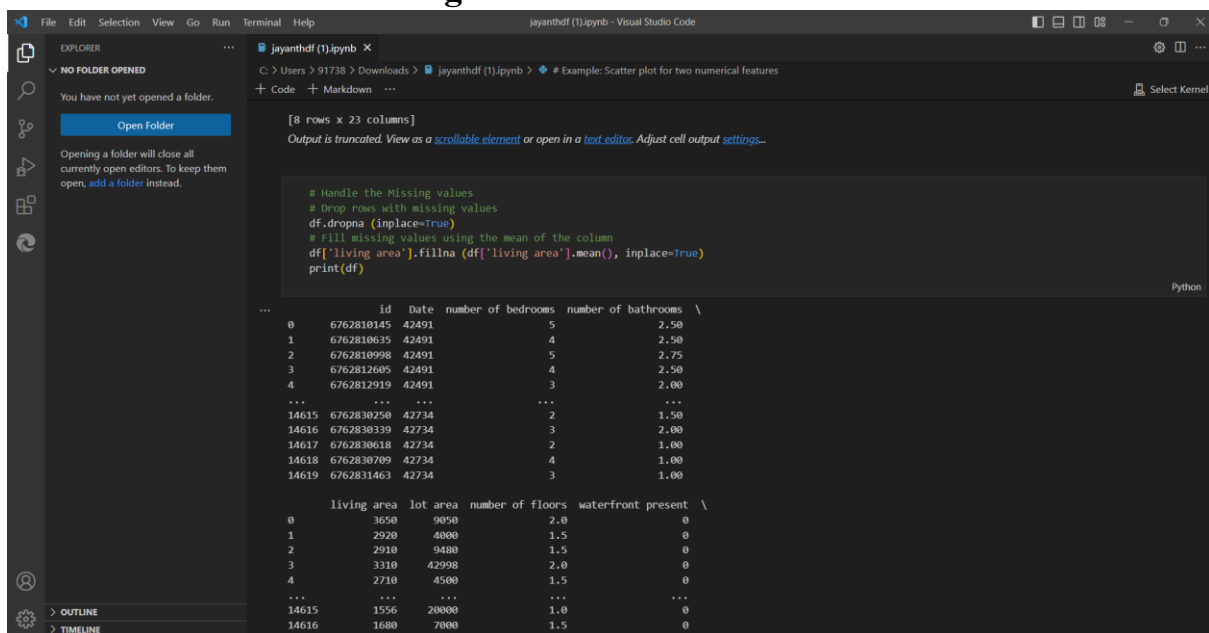
```
count    14620.000000    14620.000000    14620.000000    14620.000000    \
mean     6.762821e+09    42604.538646     3.379343     2.129583
std      6.237575e+03     67.347991     0.938719     0.769934
min      6.762810e+09    42491.000000     1.000000     0.500000
25%      6.762815e+09    42546.000000     3.000000     1.750000
50%      6.762815e+09    42600.000000     3.000000     2.250000
75%      6.762826e+09    42662.000000     4.000000     2.500000
max      6.762832e+09    42734.000000    33.000000     8.000000

count    14620.000000    14620.000000    14620.000000    14620.000000    \
mean     2898.262996    1.509328e+04     1.202360     0.007661
std      928.275721     3.791962e+04     0.548239     0.087193
min      370.000000     5.200000e+02     1.000000     0.000000
25%     1448.000000     5.010750e+03     1.000000     0.000000
50%     1930.000000     7.620000e+03     1.500000     0.000000
75%     2570.000000     1.080000e+04     2.000000     0.000000
max     13540.000000    1.074218e+06     3.500000     1.000000

count    14620.000000    14620.000000    14620.000000    14620.000000    \
mean      0.233105     3.430506     1970.926402
std      0.766259     0.664151     29.493625
min      0.000000     1.000000     1900.000000
...
75%      73.000000     6.450000e+05
max      80.000000     7.700000e+06

[8 rows x 23 columns]
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
```

STEP-4: Handle the Missing values



The screenshot shows a Visual Studio Code window with a Jupyter Notebook titled 'jayanthdf (1).ipynb'. The notebook is open to a cell containing the following code:

```
# Handle the Missing values
# Drop rows with missing values
df.dropna(inplace=True)
# Fill missing values using the mean of the column
df['living area'].fillna(df['living area'].mean(), inplace=True)
print(df)
```

The output of the code is displayed below the cell, showing the dataset after handling missing values. The output is truncated, indicating that there are 8 rows x 23 columns in total. The output is as follows:

```
count    14620.000000    14620.000000    14620.000000    14620.000000    \
mean     6.762821e+09    42604.538646     3.379343     2.129583
std      6.237575e+03     67.347991     0.938719     0.769934
min      6.762810e+09    42491.000000     1.000000     0.500000
25%      6.762815e+09    42546.000000     3.000000     1.750000
50%      6.762815e+09    42600.000000     3.000000     2.250000
75%      6.762826e+09    42662.000000     4.000000     2.500000
max      6.762832e+09    42734.000000    33.000000     8.000000

count    14620.000000    14620.000000    14620.000000    14620.000000    \
mean     2898.262996    1.509328e+04     1.202360     0.007661
std      928.275721     3.791962e+04     0.548239     0.087193
min      370.000000     5.200000e+02     1.000000     0.000000
25%     1448.000000     5.010750e+03     1.000000     0.000000
50%     1930.000000     7.620000e+03     1.500000     0.000000
75%     2570.000000     1.080000e+04     2.000000     0.000000
max     13540.000000    1.074218e+06     3.500000     1.000000

count    14620.000000    14620.000000    14620.000000    14620.000000    \
mean      0.233105     3.430506     1970.926402
std      0.766259     0.664151     29.493625
min      0.000000     1.000000     1900.000000
...
75%      73.000000     6.450000e+05
max      80.000000     7.700000e+06

[8 rows x 23 columns]
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
```