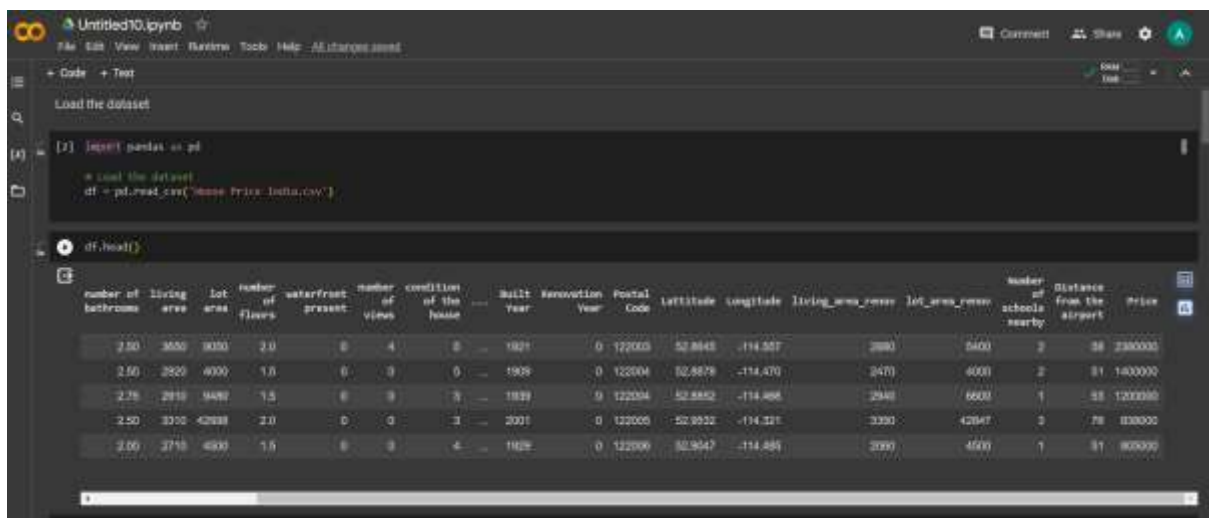


## DA ASSIGNMENT – 3

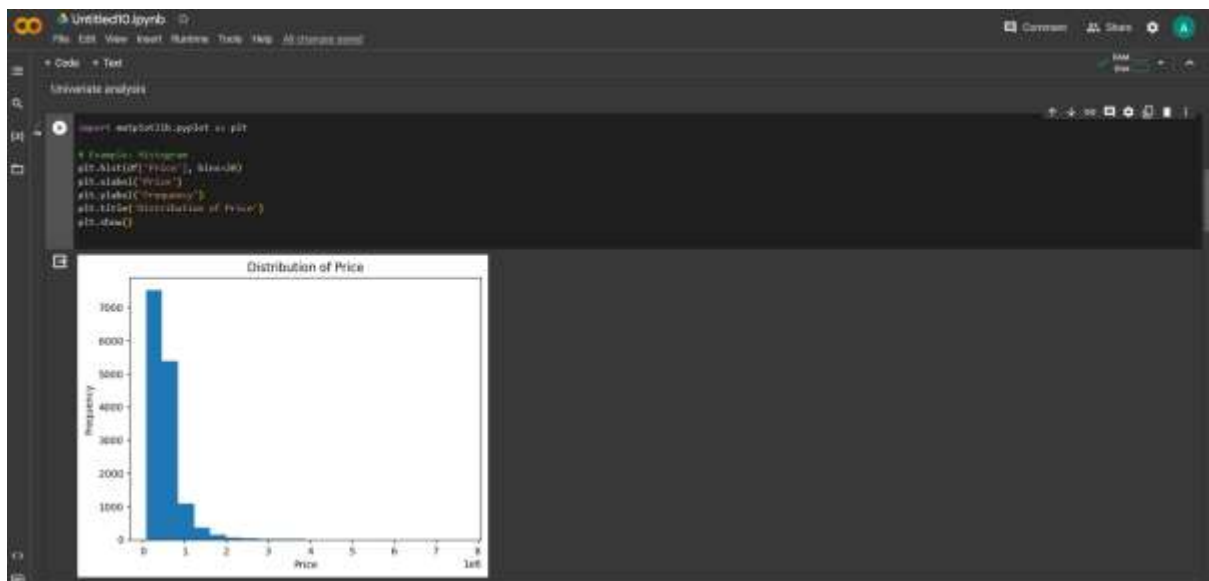
NAME : PREM KUMAR R

REGISTER NO : 723920243040

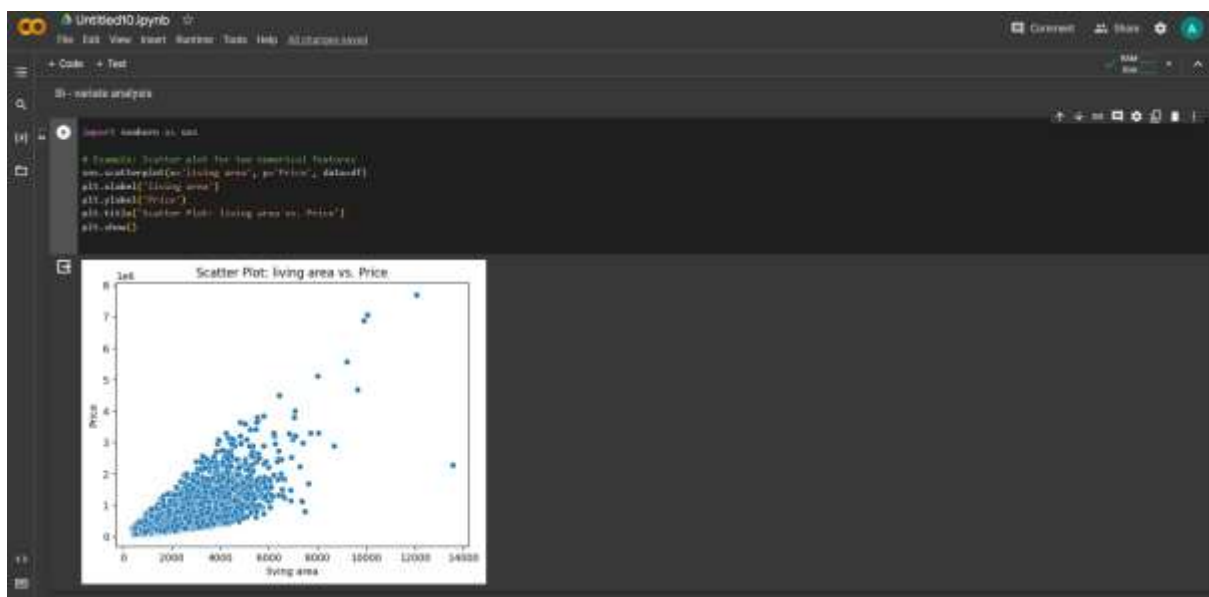
Load the dataset :



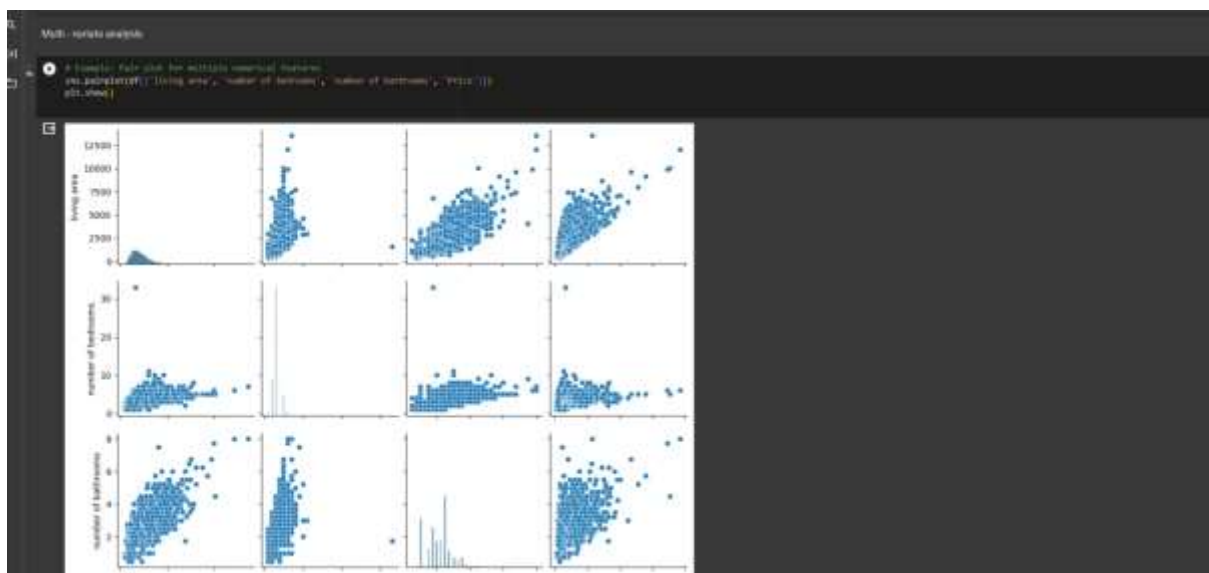
Univariate Analysis :



## Bi - Variate Analysis :



## Multi-Variate Analysis :



Descriptive statistics on the dataset :

```
descriptive statistics on the dataset

[0] # Descriptive statistics
print(df.describe())
```

	id	Date	number of bedrooms	number of bathrooms
count	1.462000e+04	14620.000000	14620.000000	14620.000000
mean	6.761821e+09	42684.538646	3.379343	2.129583
std	6.237575e+03	67.347991	0.958719	0.769934
min	6.762810e+09	42401.000000	1.000000	0.500000
25%	6.762815e+09	42566.000000	3.000000	1.750000
50%	6.762821e+09	42600.000000	3.000000	2.250000
75%	6.762825e+09	42662.000000	4.000000	2.500000
max	6.762832e+09	42734.000000	39.000000	8.000000

	living area	lot area	number of floors	waterfront present
count	14620.000000	1.462000e+04	14620.000000	14620.000000
mean	2096.262996	1.589320e+04	1.502360	0.007661
std	828.375721	3.791462e+04	0.548239	0.087193
min	370.000000	5.200000e+02	1.000000	0.000000
25%	1440.000000	5.010750e+03	1.000000	0.000000
50%	1930.000000	7.620000e+03	1.500000	0.000000
75%	2570.000000	1.000000e+04	2.000000	0.000000
max	13540.000000	1.074210e+06	3.500000	1.000000

	number of views	condition of the house	...	Built Year
count	14620.000000	14620.000000	...	14620.000000
mean	0.233105	5.430506	...	1970.926402
std	0.766259	0.664191	...	29.493625
min	0.000000	1.000000	...	1900.000000
25%	0.000000	3.000000	...	1951.000000
50%	0.000000	3.000000	...	1971.000000
75%	0.000000	4.000000	...	1987.000000

Handle the Missing values :

```
Untitled10.py
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Test

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)

[10] print(df)
```

	id	Date	number of bedrooms	number of bathrooms
0	6762810000	42401	3	2.75
1	6762812000	42401	4	2.50
4	6762812000	42401	3	2.00
...	...	...	...	...
14615	6762830250	42734	2	1.50
14616	6762830320	42734	3	2.00
14617	6762830610	42734	2	1.00
14618	6762830700	42734	4	1.00
14619	6762831403	42734	3	1.00

	living area	lot area	number of floors	waterfront present
0	3050	9070	2.0	0
1	2920	4000	1.5	0
2	2910	9000	1.5	0
3	3310	42900	2.0	0
4	2230	7000	1.5	0
...	...	...	...	...
14615	2550	20000	1.0	0
14616	1600	7000	1.5	0
14617	1070	6120	1.0	0
14618	1030	6021	1.0	0
14619	980	4770	1.0	0

	number of views	condition of the house	...	Built Year
0	0.000000	5.000000	...	1900.000000
1	0.000000	5.000000	...	1900.000000
2	0.000000	5.000000	...	1900.000000
3	0.000000	5.000000	...	1900.000000
4	0.000000	5.000000	...	1900.000000
...	...	...	...	...
14615	0.000000	5.000000	...	1900.000000
14616	0.000000	5.000000	...	1900.000000
14617	0.000000	5.000000	...	1900.000000
14618	0.000000	5.000000	...	1900.000000
14619	0.000000	5.000000	...	1900.000000