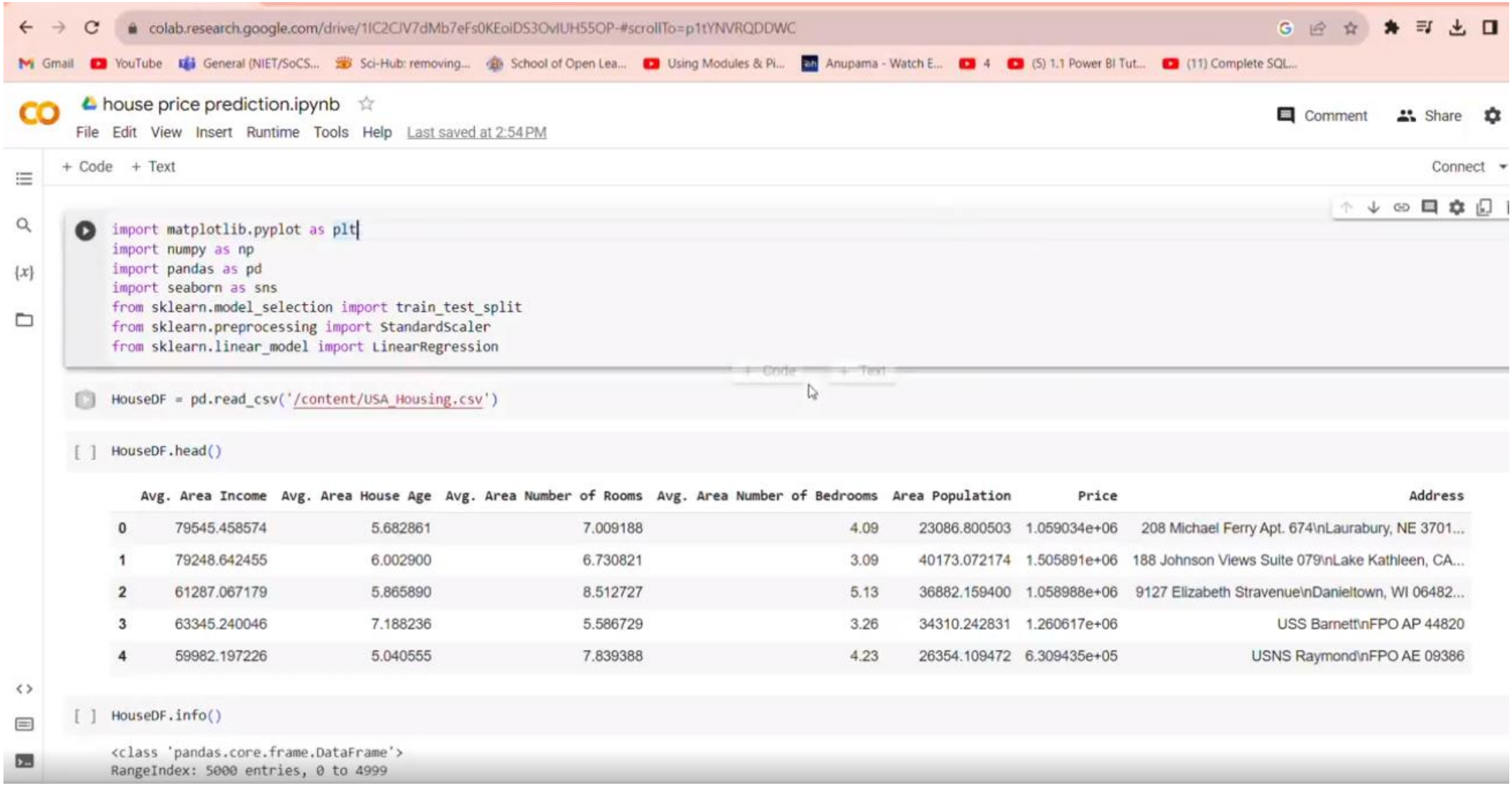
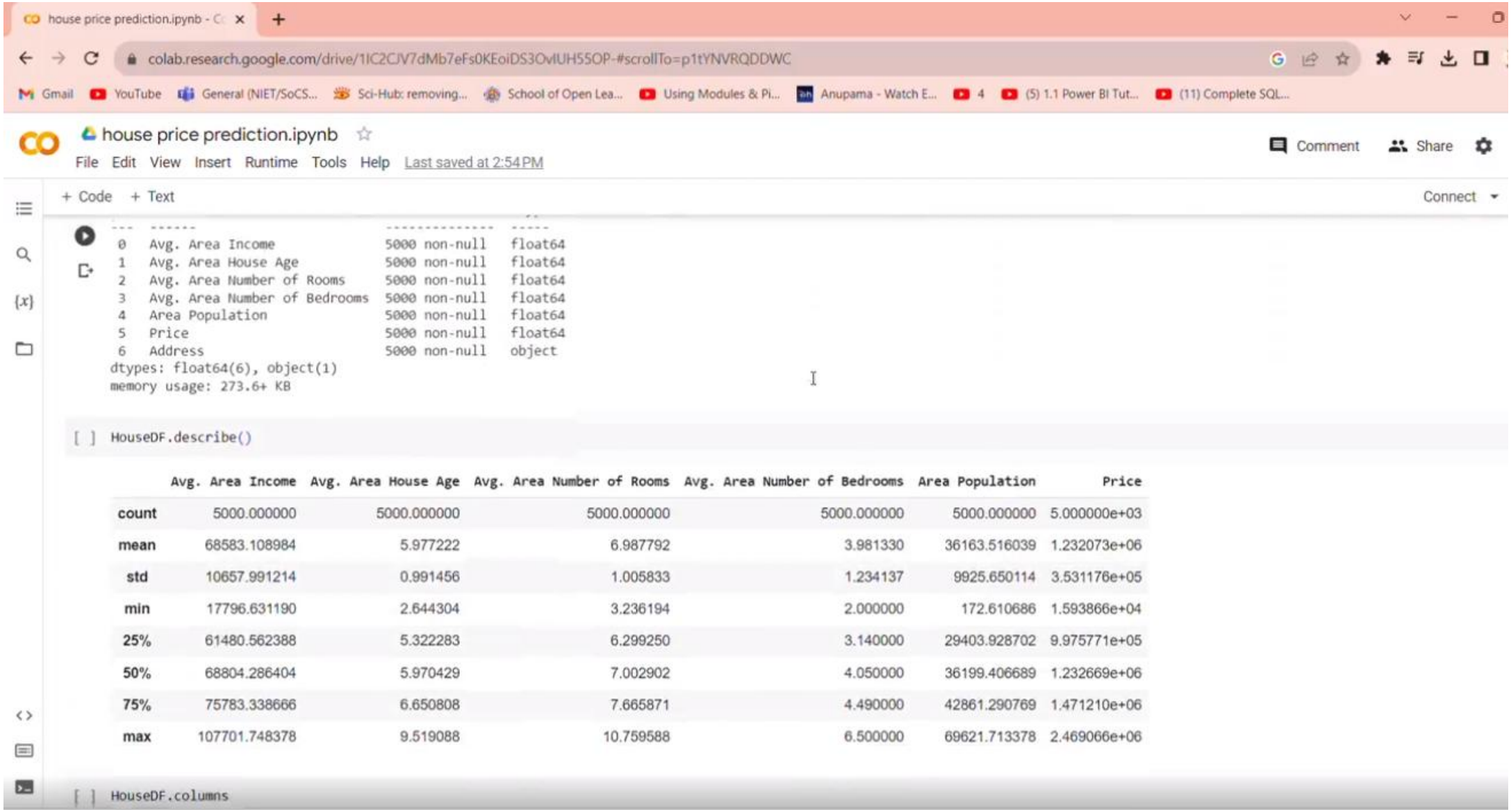

colab.research.google.com/drive/1IC2CJV7dMb7eFs0KEoiDS3OvIUH55OP-#scrollTo=p1tYNVRQDDWC
Gmail
YouTube
General (NIET/SoCS...
Sci-Hub: removing...
School of Open Lea...
Using Modules & Pi...
Anupama - Watch E...
4
(5) 1.1 Power BI Tut...
(11) Complete SQL...
house price prediction.ipynb
File Edit View Insert Runtime Tools Help Last saved at 2:54 PM
Comment
Share
+ Code + Text
Connect
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LinearRegression
HouseDF = pd.read_csv('/content/USA_Housing.csv')
[ ] HouseDF.head()
Avg. Area Income Avg. Area House Age Avg. Area Number of Rooms Avg. Area Number of Bedrooms Area Population Price Address
0 79545.458574 5.682861 7.009188 4.09 23086.800503 1.059034e+06 208 Michael Ferry Apt. 674\nLaurabury, NE 3701...
1 79248.642455 6.002900 6.730821 3.09 40173.072174 1.505891e+06 188 Johnson Views Suite 079\nLake Kathleen, CA...
2 61287.067179 5.865890 8.512727 5.13 36882.159400 1.058988e+06 9127 Elizabeth Stravenue\nDanieltown, WI 06482...
3 63345.240046 7.188236 5.586729 3.26 34310.242831 1.260617e+06 USS Barnett\nFPO AP 44820
4 59982.197226 5.040555 7.839388 4.23 26354.109472 6.309435e+05 USNS Raymond\nFPO AE 09386
[ ] HouseDF.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5000 entries, 0 to 4999


house price prediction.ipynb - C
colab.research.google.com/drive/1IC2CJV7dMb7eFs0KEoiDS3OvlUH55OP-#scrollTo=p1tYNVRQDDWC
Gmail
YouTube
General (NIET/SoCS...
Sci-Hub: removing...
School of Open Lea...
Using Modules & Pi...
Anupama - Watch E...
4
(5) 1.1 Power BI Tut...
(11) Complete SQL...
house price prediction.ipynb
File Edit View Insert Runtime Tools Help Last saved at 2:54 PM
Comment
Share
+ Code + Text
Connect
0 Avg. Area Income 5000 non-null float64
1 Avg. Area House Age 5000 non-null float64
2 Avg. Area Number of Rooms 5000 non-null float64
3 Avg. Area Number of Bedrooms 5000 non-null float64
4 Area Population 5000 non-null float64
5 Price 5000 non-null float64
6 Address 5000 non-null object
dtypes: float64(6), object(1)
memory usage: 273.6+ KB
[ ] HouseDF.describe()
Avg. Area Income Avg. Area House Age Avg. Area Number of Rooms Avg. Area Number of Bedrooms Area Population Price
count 5000.000000 5000.000000 5000.000000 5000.000000 5000.000000 5.000000e+03
mean 68583.108984 5.977222 6.987792 3.981330 36163.516039 1.232073e+06
std 10657.991214 0.991456 1.005833 1.234137 9925.650114 3.531176e+05
min 17796.631190 2.644304 3.236194 2.000000 172.610686 1.593866e+04
25% 61480.562388 5.322283 6.299250 3.140000 29403.928702 9.975771e+05
50% 68804.286404 5.970429 7.002902 4.050000 36199.406689 1.232669e+06
75% 75783.338666 6.650808 7.665871 4.490000 42861.290769 1.471210e+06
max 107701.748378 9.519088 10.759588 6.500000 69621.713378 2.469066e+06
[ ] HouseDF.columns

house price prediction.ipynb - C

colab.research.google.com/drive/1IC2CJV7dMb7eFs0KEoiDS3OvIUH55OP-#scrollTo=p1tYNVRQDDWC

Gmail YouTube General (NIET/SoCS... Sci-Hub: removing... School of Open Lea... Using Modules & Pi... Anupama - Watch E... 4 (5) 1.1 Power BI Tut... (11) Complete SQL...

house price prediction.ipynb

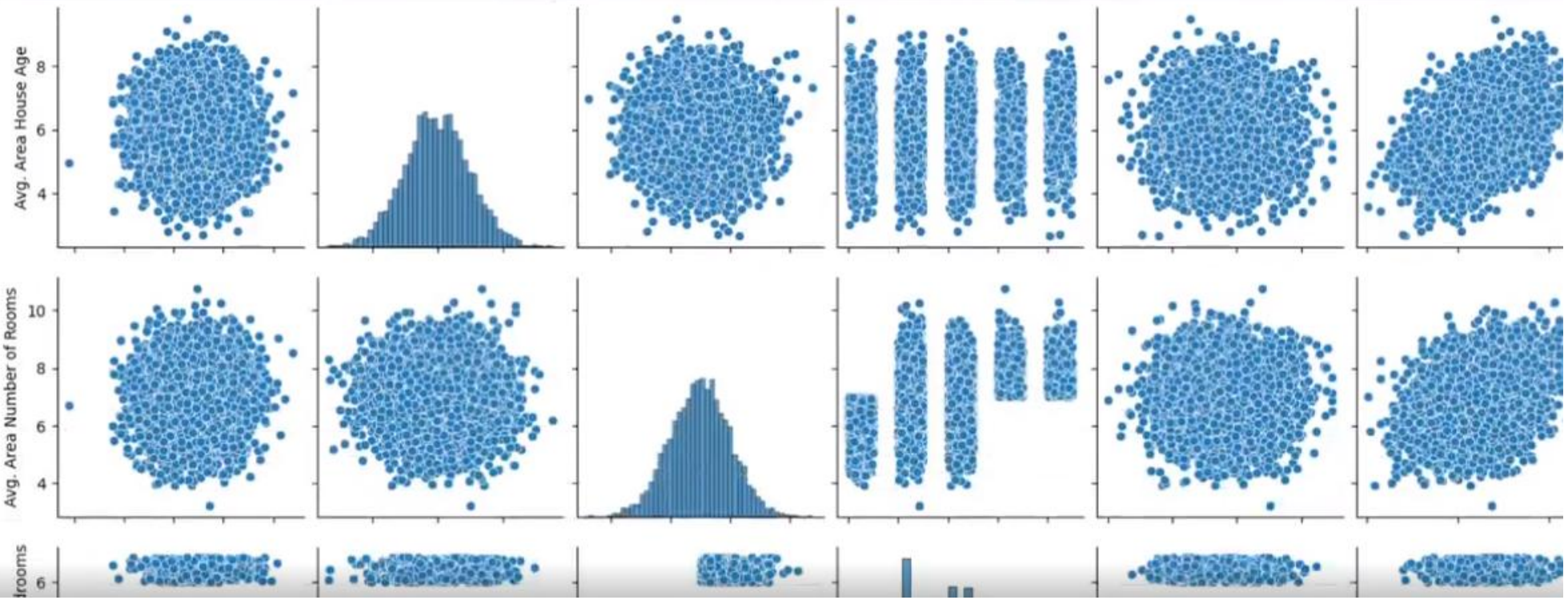
File Edit View Insert Runtime Tools Help Last saved at 2:54 PM

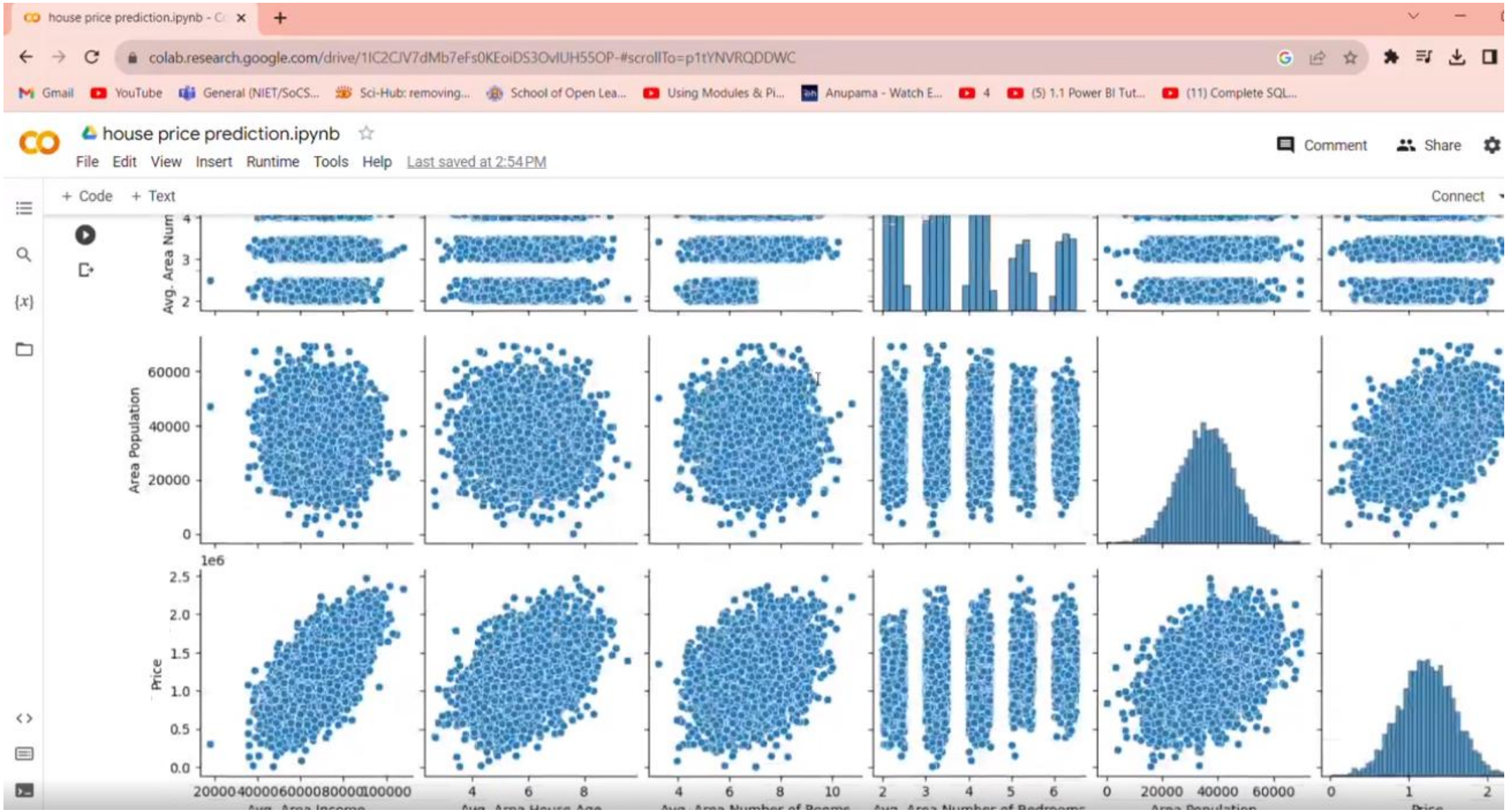
Comment Share

+ Code + Text Connect

```
dtype='object')
```

```
sns.pairplot(HouseDF)
```




house price prediction.ipynb - C
colab.research.google.com/drive/1IC2CJV7dMb7eFs0KEoiDS3OvIUH55OP-#scrollTo=p1tYNVRQDDWC
Gmail
YouTube
General (NIET/SoCS...
Sci-Hub: removing...
School of Open Lea...
Using Modules & Pi...
Anupama - Watch E...
4
(5) 1.1 Power BI Tut...
(11) Complete SQL...
house price prediction.ipynb
File Edit View Insert Runtime Tools Help Last saved at 2:54 PM
Comment
Share
+ Code
+ Text
Connect
Avg. Area Num
Area Population
Price
1e6
Avg. Area Income
Avg. Area House Age
Avg. Area Number of Rooms
Avg. Area Number of Bedrooms
Area Population
Price

house price prediction.ipynb - C

colab.research.google.com/drive/1IC2CJV7dMb7eFs0KEoiDS3OvIUH55OP-#scrollTo=p1tYNVRQDDWC

Gmail YouTube General (NIET/SoCS... Sci-Hub: removing... School of Open Lea... Using Modules & Pi... Anupama - Watch E... 4 (5) 1.1 Power BI Tut... (11) Complete SQL...

house price prediction.ipynb

File Edit View Insert Runtime Tools Help Last saved at 2:54 PM

Comment Share

+ Code + Text Connect

```
sns.heatmap(HouseDF.corr(), annot=True)
```

```
<ipython-input-22-72089517bde6>:1: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future version, it will default to False. Select only v
  sns.heatmap(HouseDF.corr(), annot=True)
<Axes: >
```

house price prediction.ipynb - C

colab.research.google.com/drive/1IC2CJV7dMb7eFs0KEoiDS3OvIUH55OP-#scrollTo=p1tYNVRQDDWC

Gmail YouTube General (NIET/SoCS... Sci-Hub: removing... School of Open Lea... Using Modules & Pi... Anupama - Watch E... 4 (5) 1.1 Power BI Tut... (11) Complete SQL...

house price prediction.ipynb

File Edit View Insert Runtime Tools Help Last saved at 2:54 PM

Comment Share

+ Code + Text Connect

```
[ ] X = HouseDF[['Avg. Area Income', 'Avg. Area House Age', 'Avg. Area Number of Rooms',
        'Avg. Area Number of Bedrooms', 'Area Population']]

    y = HouseDF['Price']
```

```
[ ] X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.40, random_state=101)
```

```
[ ] X_train
```

	Avg. Area Income	Avg. Area House Age	Avg. Area Number of Rooms	Avg. Area Number of Bedrooms	Area Population
1303	68091.179676	5.364208	7.502956	3.10	44557.379656
1051	75729.765546	5.580599	7.642973	4.21	29996.018448
4904	70885.420819	6.358747	7.250241	5.42	38627.301473
931	73386.407340	4.966360	7.915453	4.30	38413.490484
4976	75046.313791	5.351169	7.797825	5.23	34107.888619
...	...	...	...	...	...
4171	56610.642563	4.846832	7.558137	3.29	25494.740298
599	70596.850945	6.548274	6.539986	3.10	51614.830136
1361	55621.809104	3.735942	6.868291	2.30	63184.613147

house price prediction.ipynb - C

colab.research.google.com/drive/1IC2CJV7dMb7eFs0KEoiDS3OvlUH55OP-#scrollTo=p1tYNVRQDDWC

Gmail YouTube General (NIET/SoCS... Sci-Hub: removing... School of Open Lea... Using Modules & Pi... Anupama - Watch E... 4 (5) 1.1 Power BI Tut... (11) Complete SQL...

house price prediction.ipynb

File Edit View Insert Runtime Tools Help Last saved at 2:54 PM

Comment Share

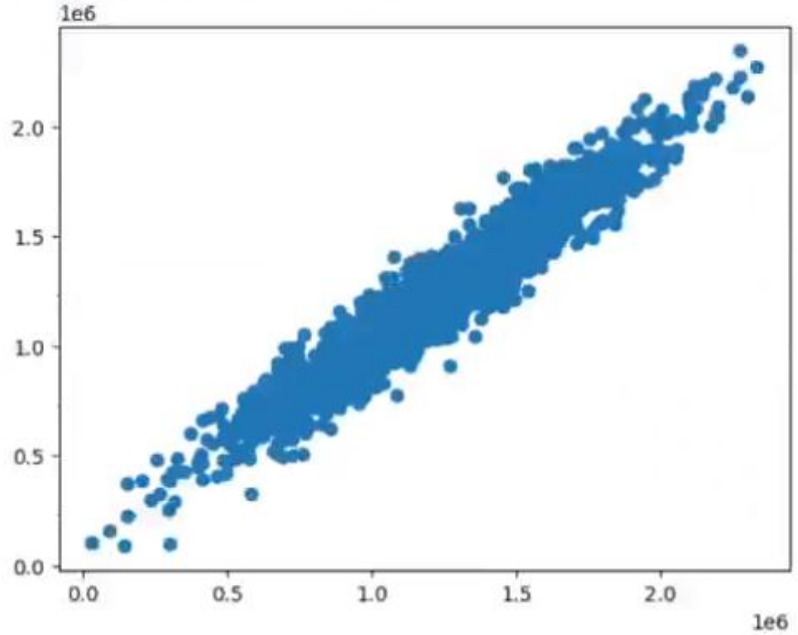
+ Code + Text Connect

predictions = lm.predict(X_test)

+ Code + Text

```
[ ] plt.scatter(y_test, predictions)
```

```
<matplotlib.collections.PathCollection at 0x7b5b9b70ae60>
```



house price prediction.ipynb - C

colab.research.google.com/drive/1IC2CJV7dMb7eFs0KEoiDS3OvlUH55OP-#scrollTo=p1tYNVRQDDWC

Gmail YouTube General (NIET/SoCS... Sci-Hub: removing... School of Open Lea... Using Modules & Pi... Anupama - Watch E... 4 (5) 1.1 Power BI Tut... (11) Complete SQL...

house price prediction.ipynb

File Edit View Insert Runtime Tools Help Last saved at 2:54 PM

Comment Share

+ Code + Text

Connect

```
`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot((y_test-predictions),bins=50);
```