**House price prediction using machine learning**

**Program:**

In [1]:

Model\_rf=RandomForestRegression(n\_estimators=50)

In[2]:

Model\_rf.fit(X\_train\_scal,Y\_train)

In[3]:

Print(r2\_score(Y\_test,Prediction2))

Print(mean\_absolute\_error(Y\_test, Prediction2))

Print (mean\_squared\_error(Y\_test, Prediction2))

Output:

Out[3]:

-0.0006222175925689744

286137.81086908665

128209033251.4034

**Program:**

Data.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...

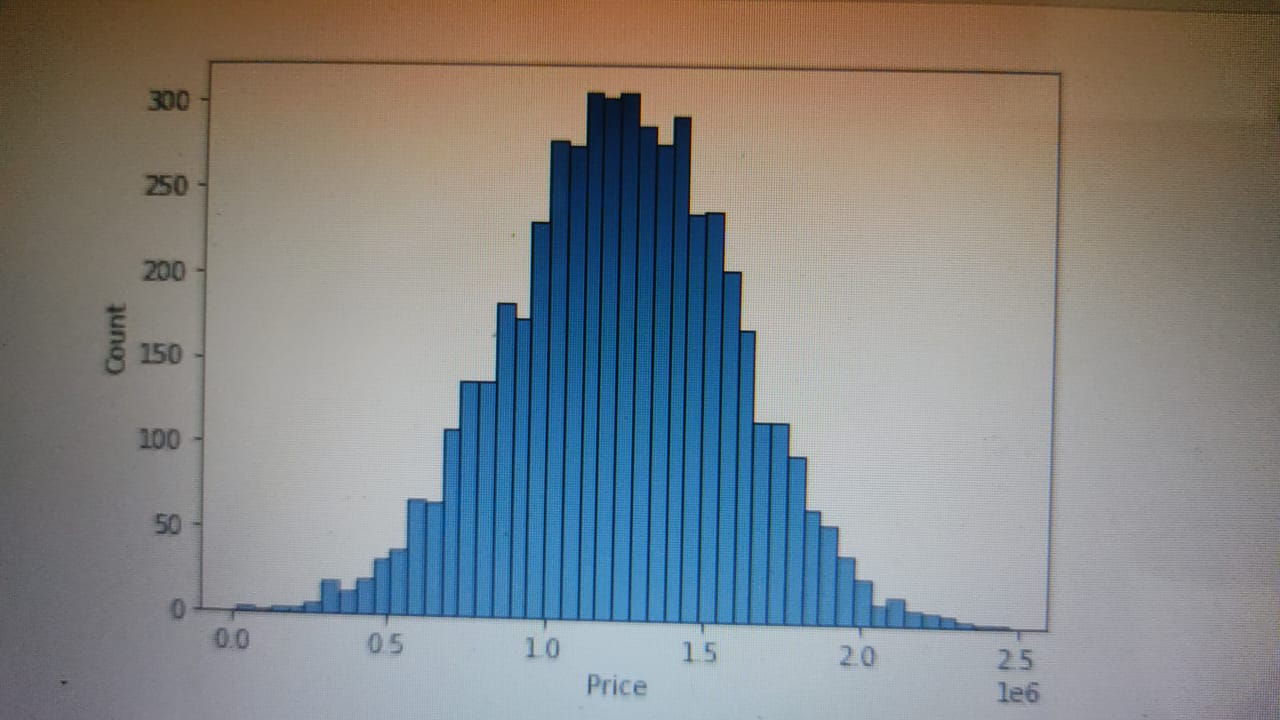
2 61287.067179 5.865890 8.512727 5.13 36882.159400 1.058988e+06 9127 Elizabeth Stravenue\nDanieltown, WI 06482...

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

Step:1

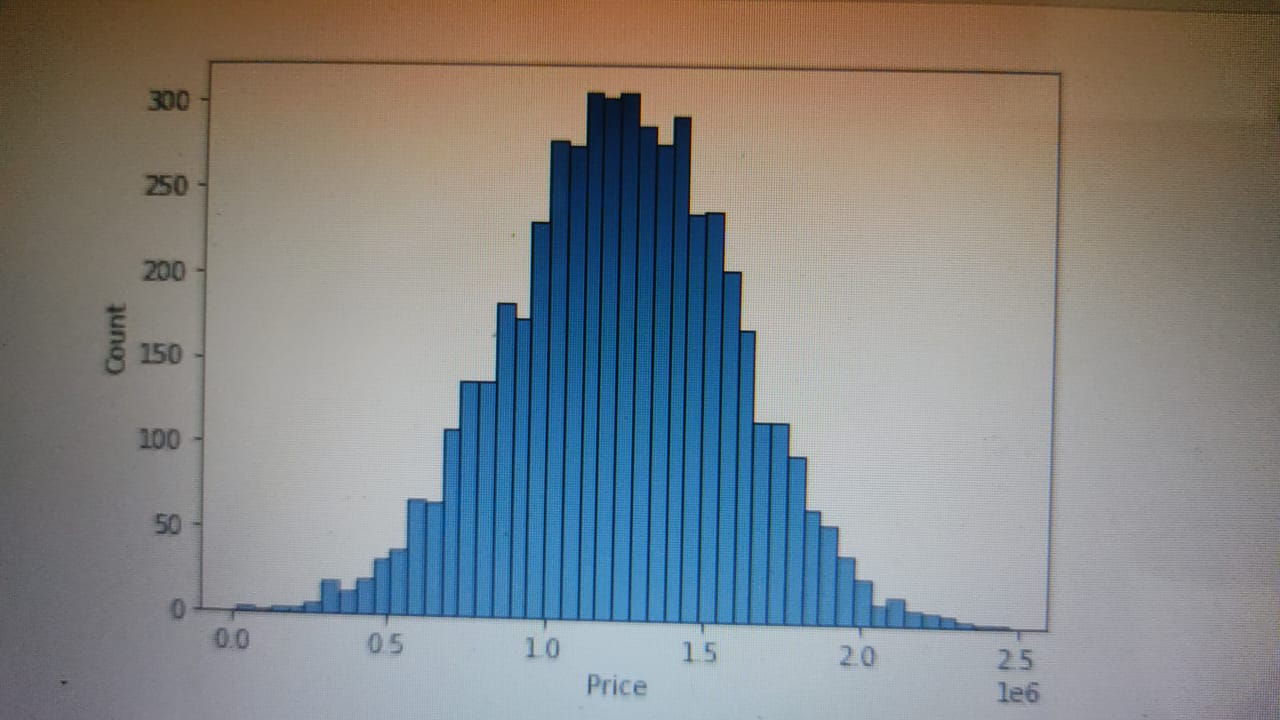
b.histplot(x = data[‘Price’]);

Output 1:

Step:2

Sb.histplot(x = data[‘Avg. Area Income’]);

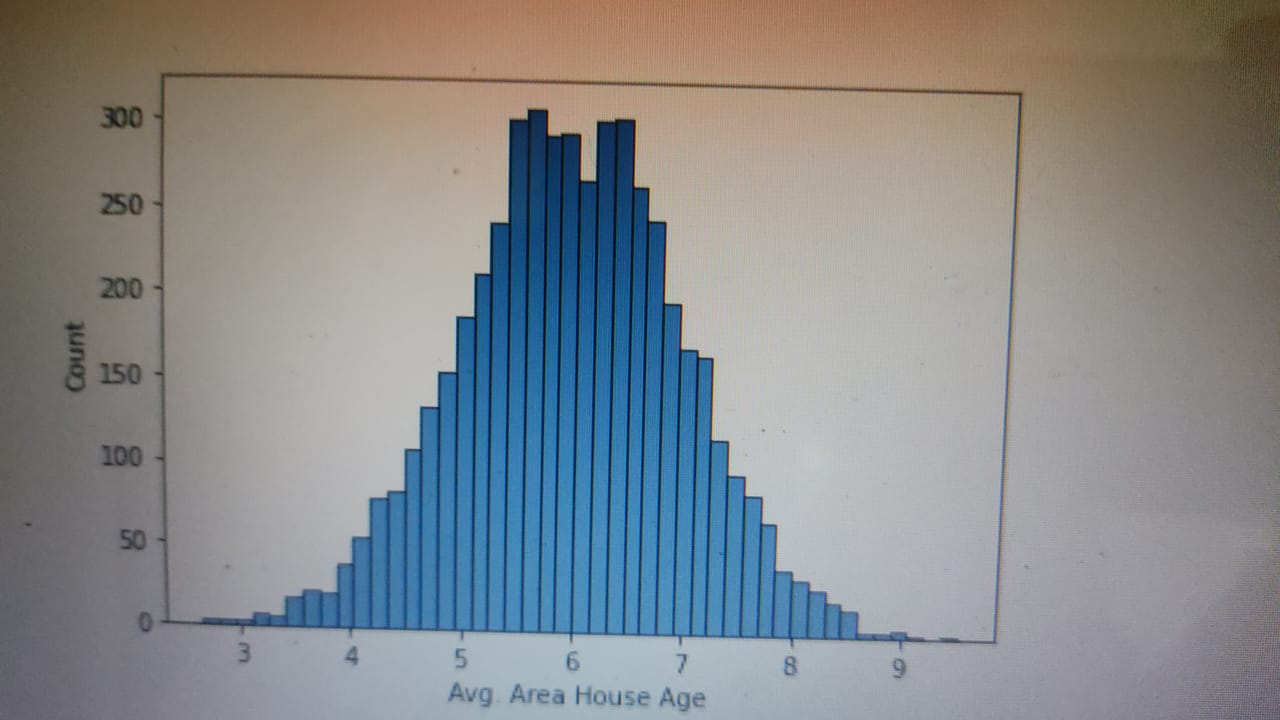
Output 2:



Step 3:

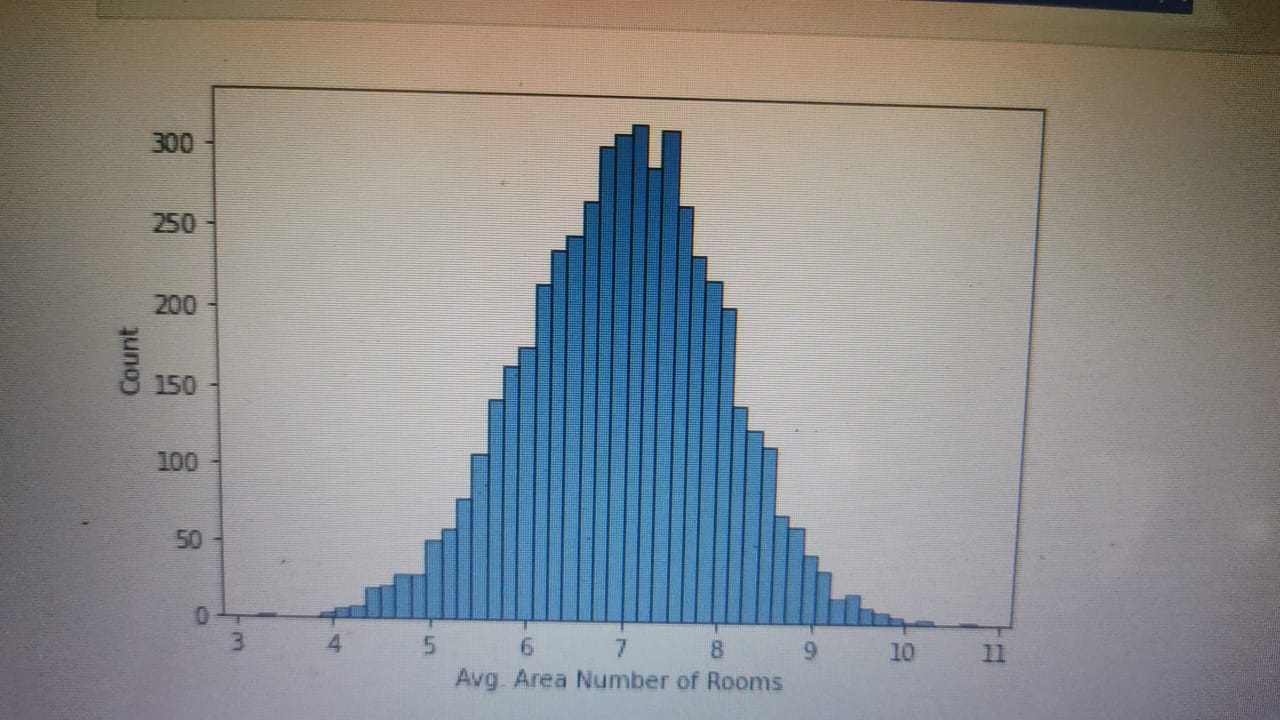
Sb.histplot(x = data[‘Avg. Area House Age’])

Output 3:



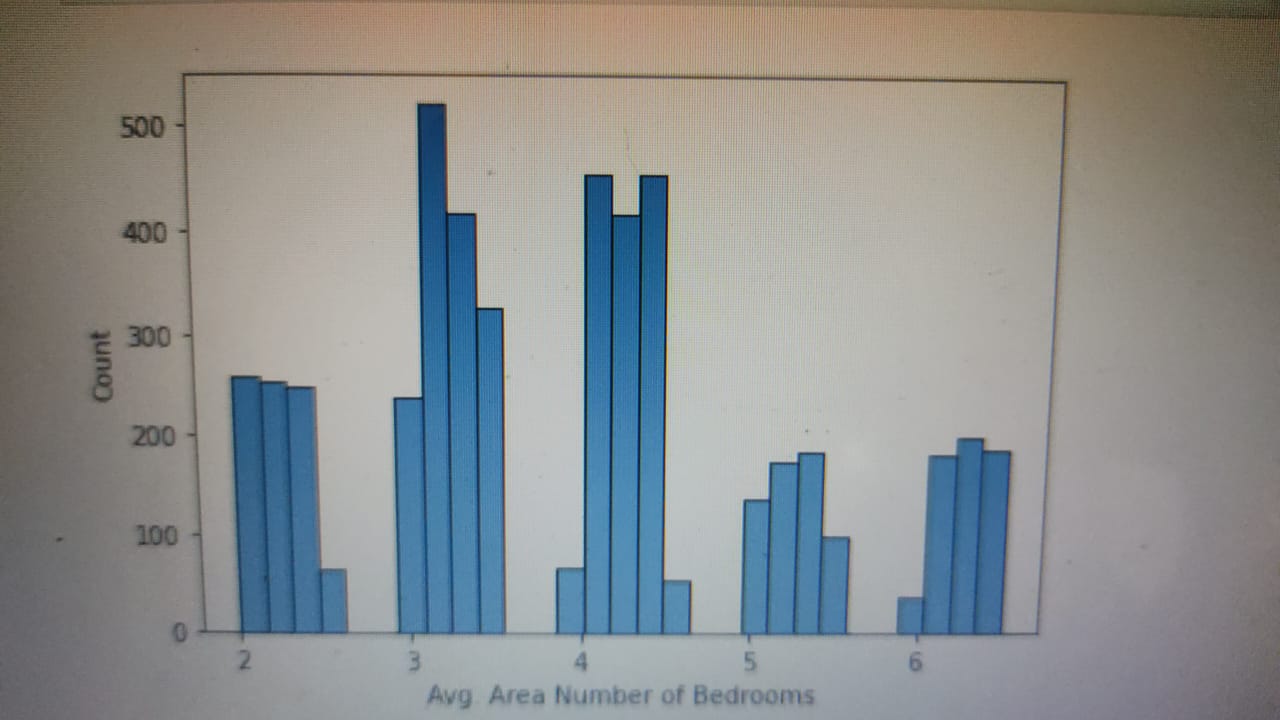
Step4:

Sb.histplot(x = data[‘Avg. Area Number of Rooms’]);

 Output 4:

Step:5

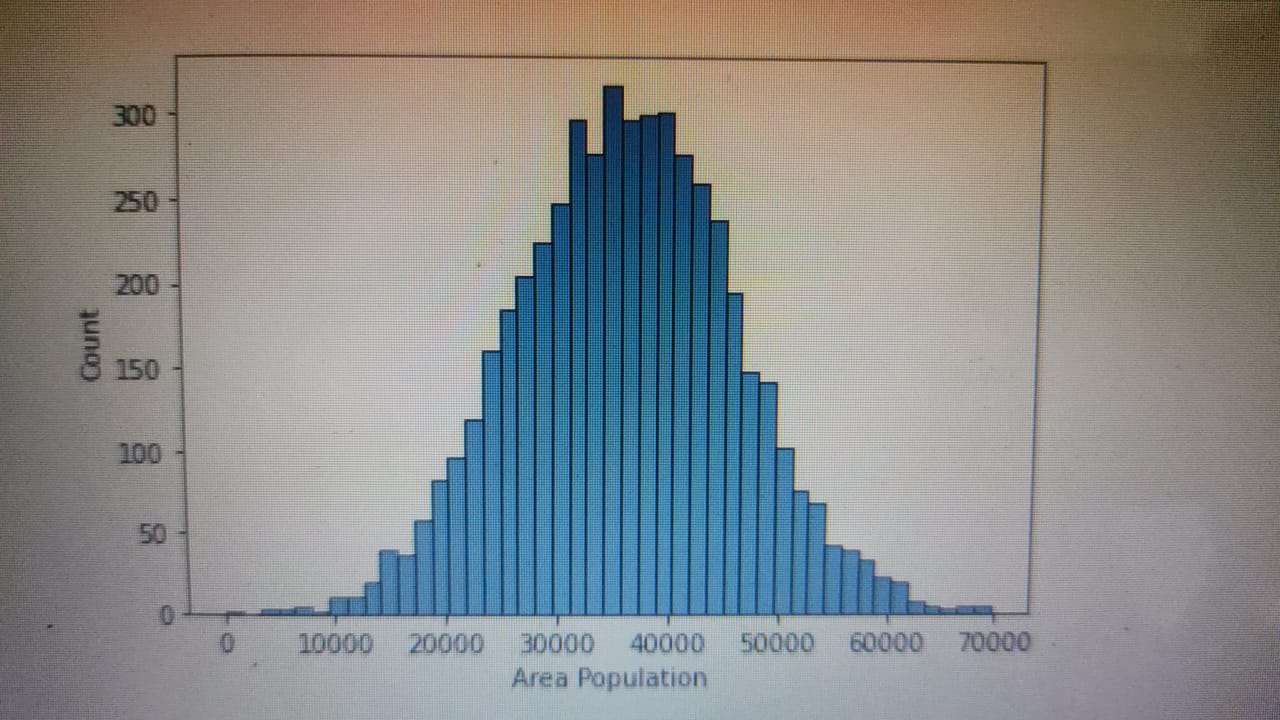
Sb.histplot(x = data[‘Avg. Area Number of Bedrooms’]);

 Output 5:

Step6:

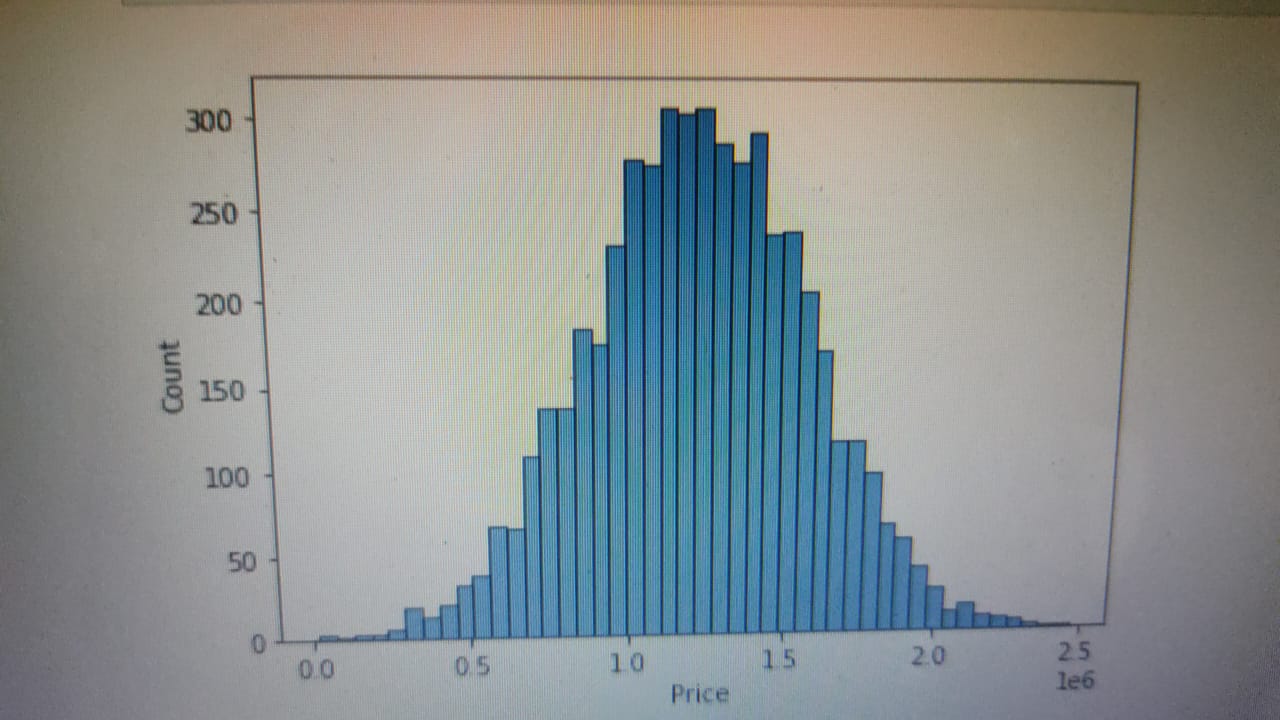
Sb.histplot(x = data[‘Area Population’]);

Output 6:



Step7:

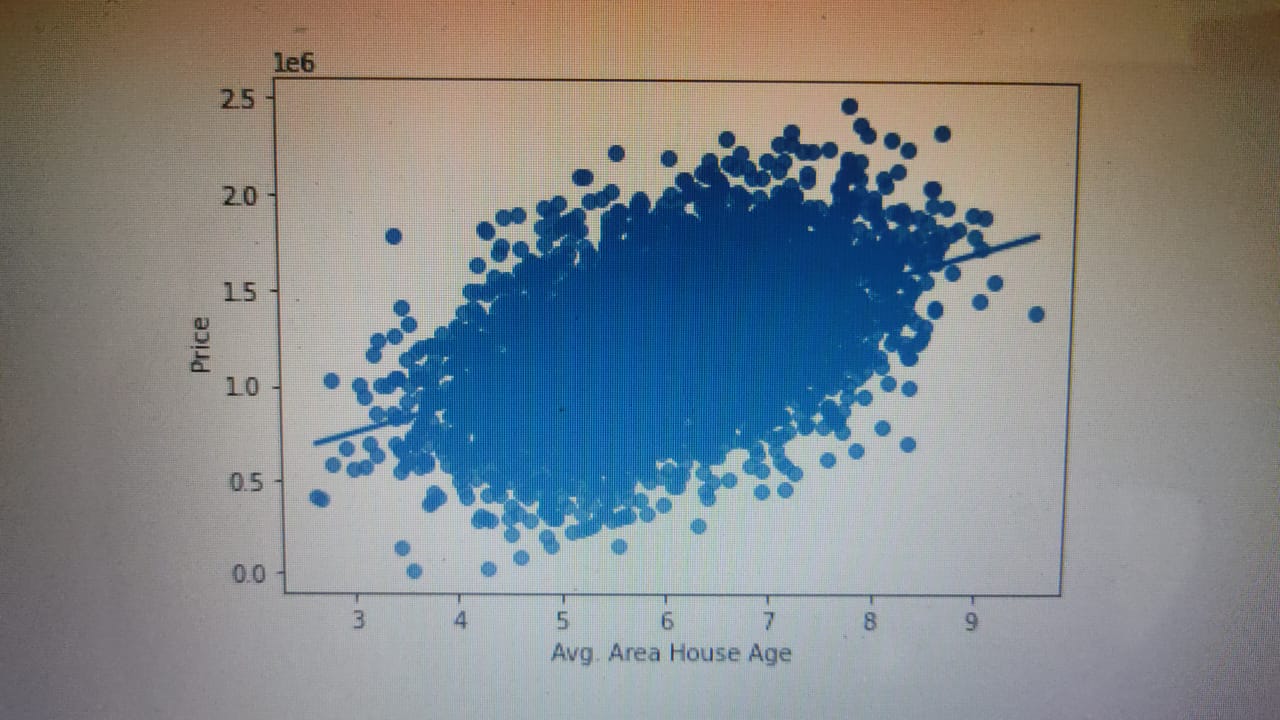
Sb.histplot(x = data[‘Price’]);

 Output 7:

Step8:

Sb.regplot(x = data[‘Avg. Area House Age’], y = data[‘Price’])

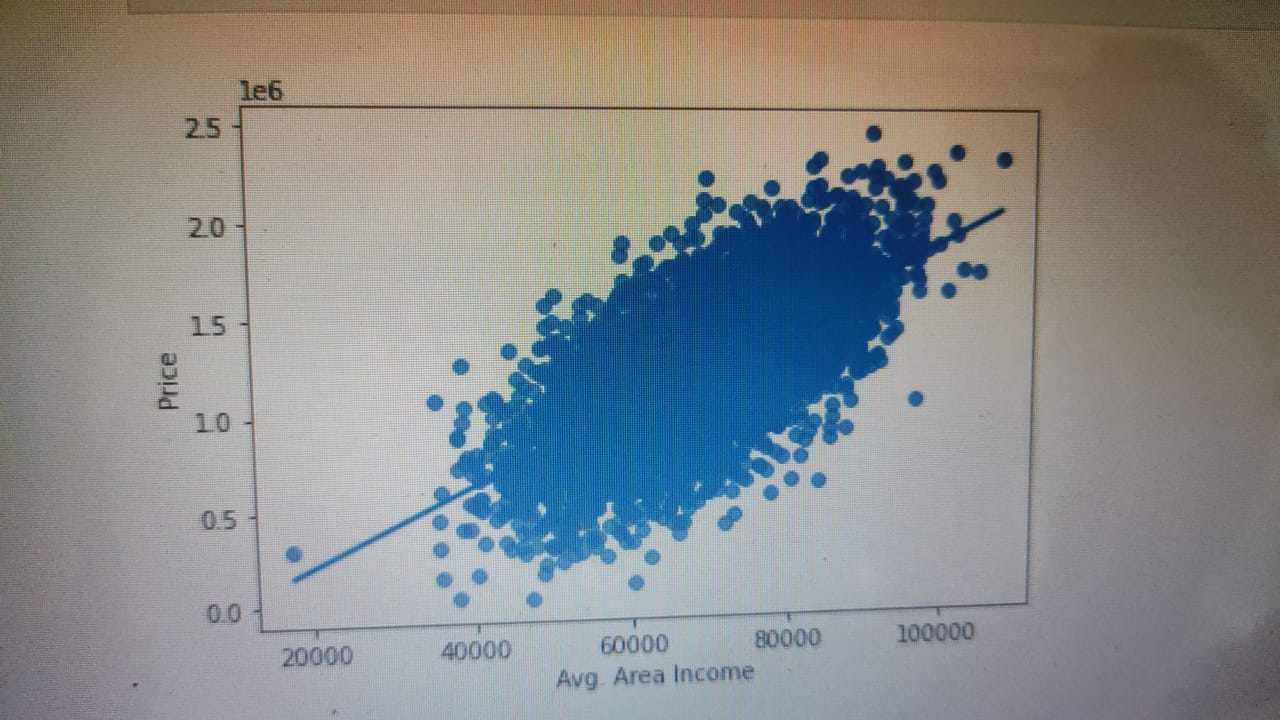
Output 8:



Step9:

Sb.regplot(x = data[‘Avg. Area Income’], y = data[‘Price’]);

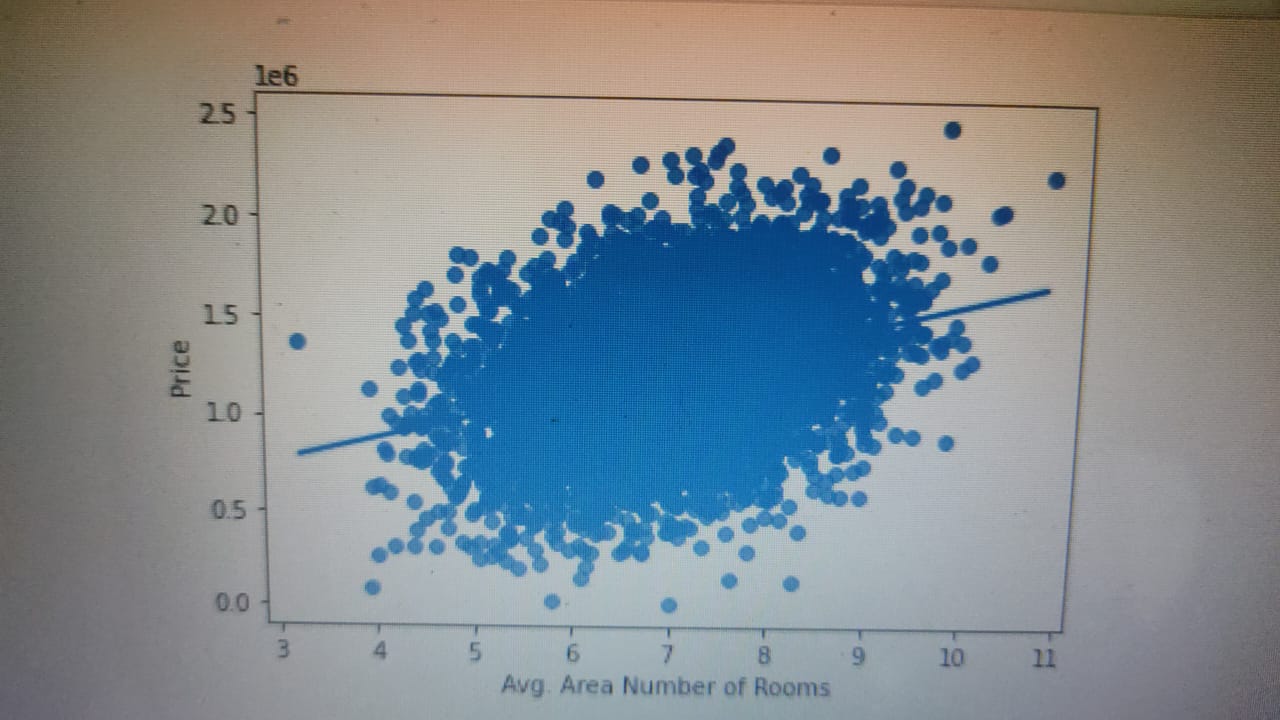
Output 9:



Step 10:

Sb.regplot(x = data[‘Avg. Area Number of Rooms’], y = data[‘Price’]);

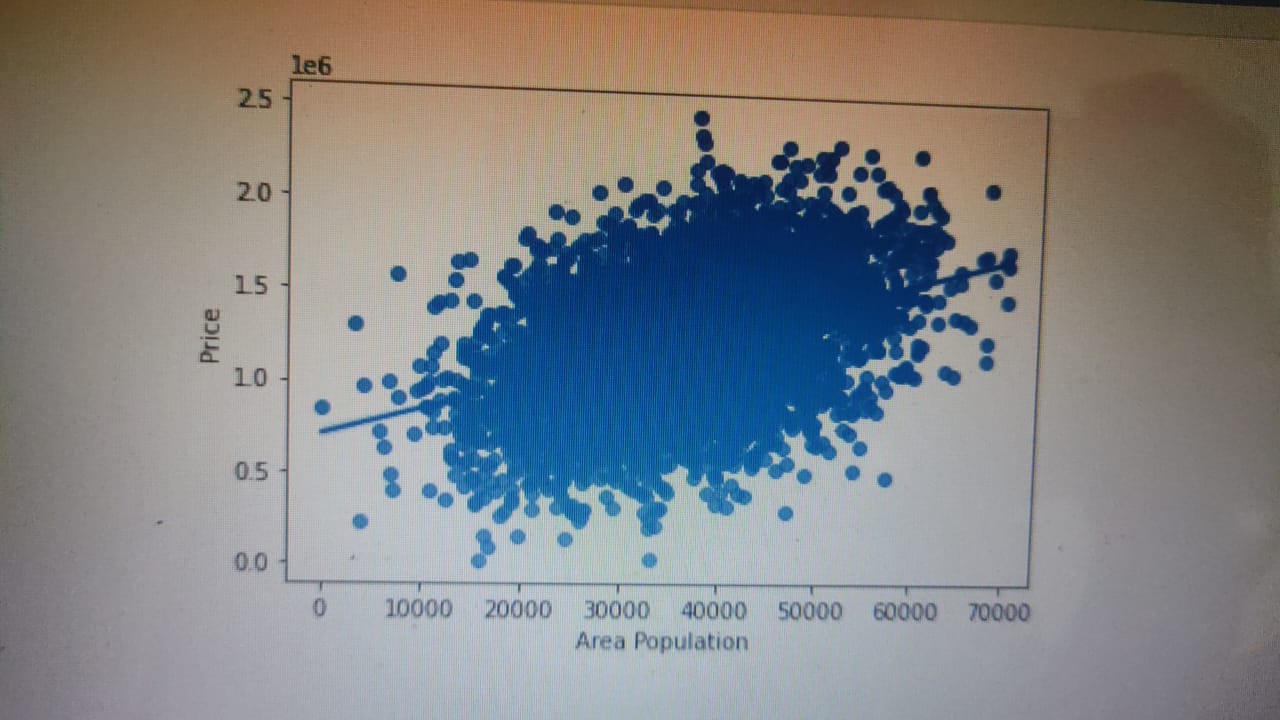
Output 10:



Step11:

Sb.regplot(x = data[‘Area Population’], y = data[‘Price’]);

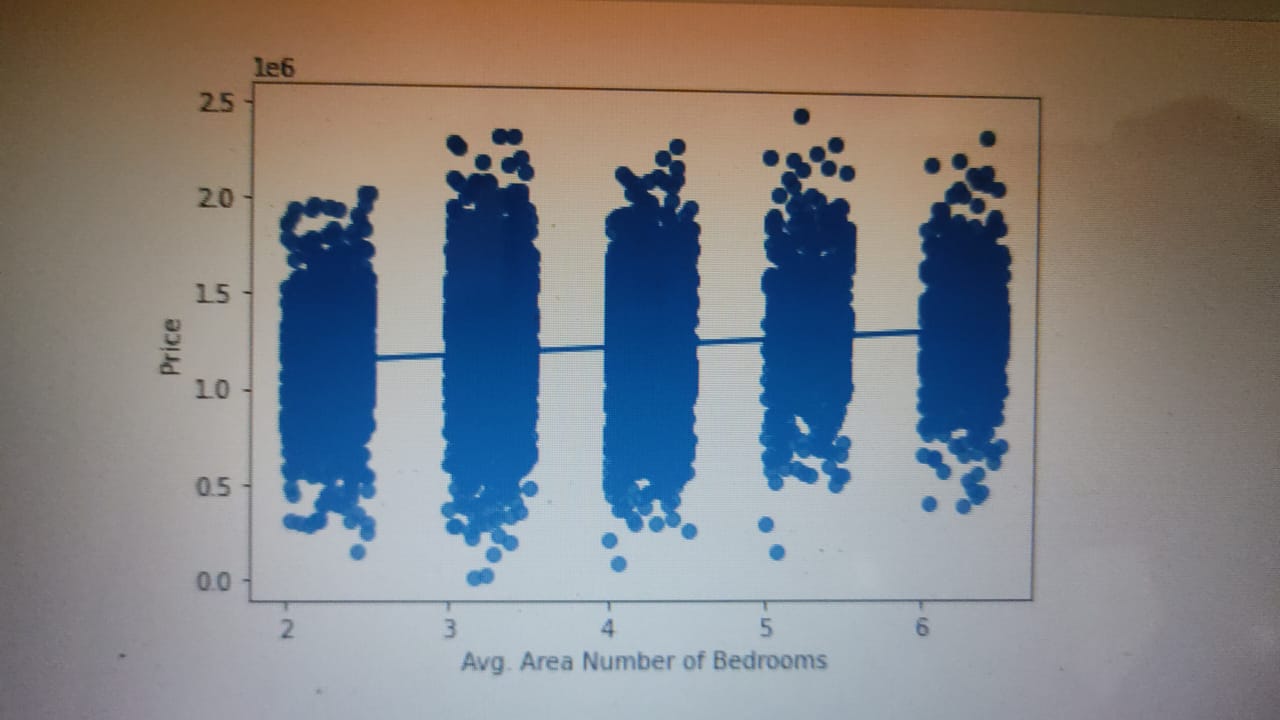
Output 11:



Step 12:

Sb.regplot(x = data[‘Avg. Area Number of Bedrooms’], y = data[‘Price’]);

Output 12:



Analysis

Price increases with all the variables

Price increases sharply with increase in Average Area Income

Plt.figure(figsize = (15, 10))

Sb.heatmap(data.corr(), annot = True, cmap = ‘mako’)

Y = data[‘Price’]

X = data.drop([‘Price’, ‘Address’], axis = 1)

From sklearn.model\_selection import train\_test\_split

From sklearn.metrics import mean\_absolute\_error

From sklearn.metrics import mean\_squared\_error

From sklearn.ensemble import RandomForestRegressor

Train\_X, val\_X, train\_y, val\_y = train\_test\_split(X, y, random\_state = 42)

Model = RandomForestRegressor(random\_state = 1)

Model.fit(train\_X, train\_y)

Preds = model.predict(val\_X)

Print(“MAE: “, mean\_absolute\_error(preds, val\_y))

Print(“RMSE: “, np.sqrt(mean\_squared\_error(preds, val\_y)))

MAE: 93812.37073246129

RMSE: 118380.48325186648

Output:

