EXPERIMENT 4(A)

IMPLEMENTATION OF LINEAR REGRESSION

NAME : SREENIDHI GANACHARI

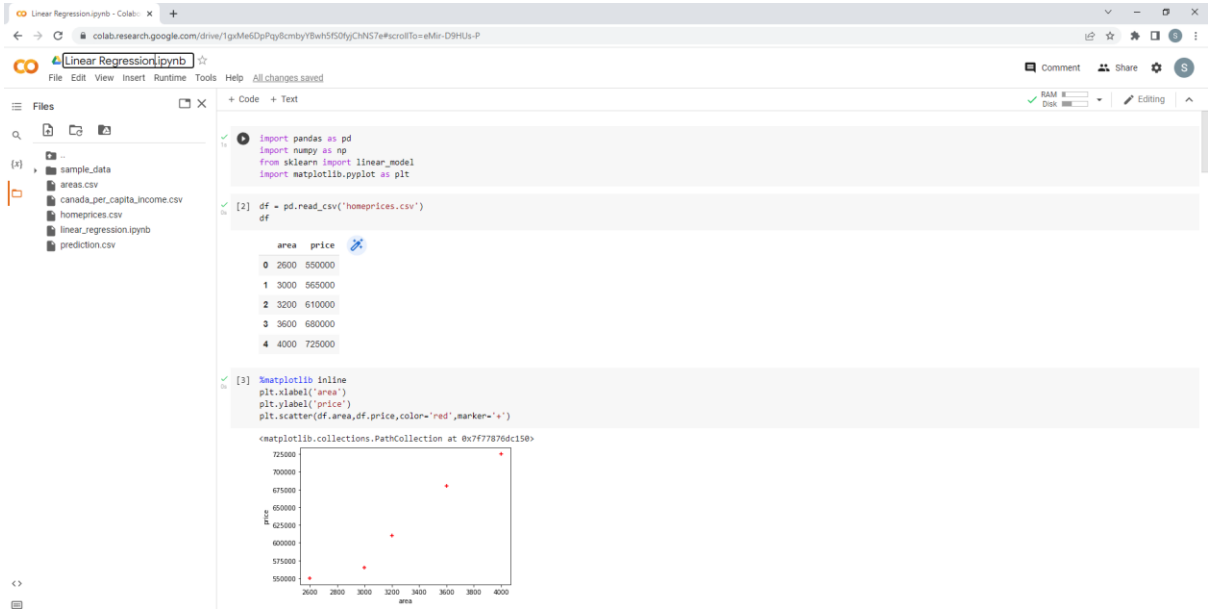
REGISTRATION NUMBER : 19BCE7230

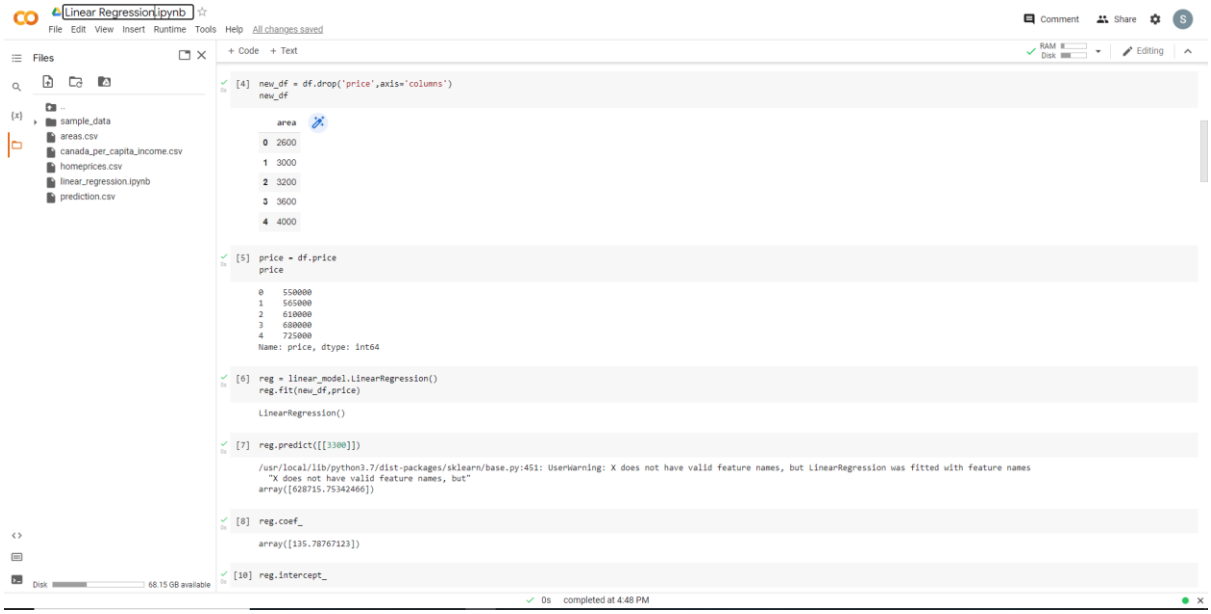
SLOT NO : L-23+24

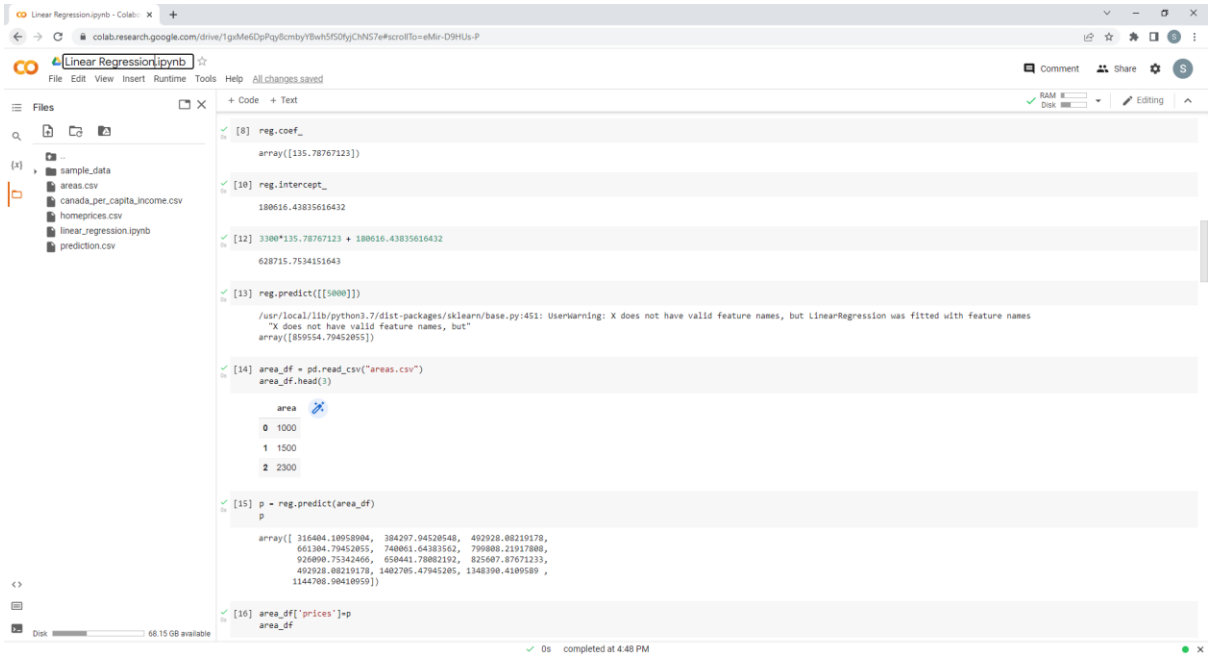
CODE –

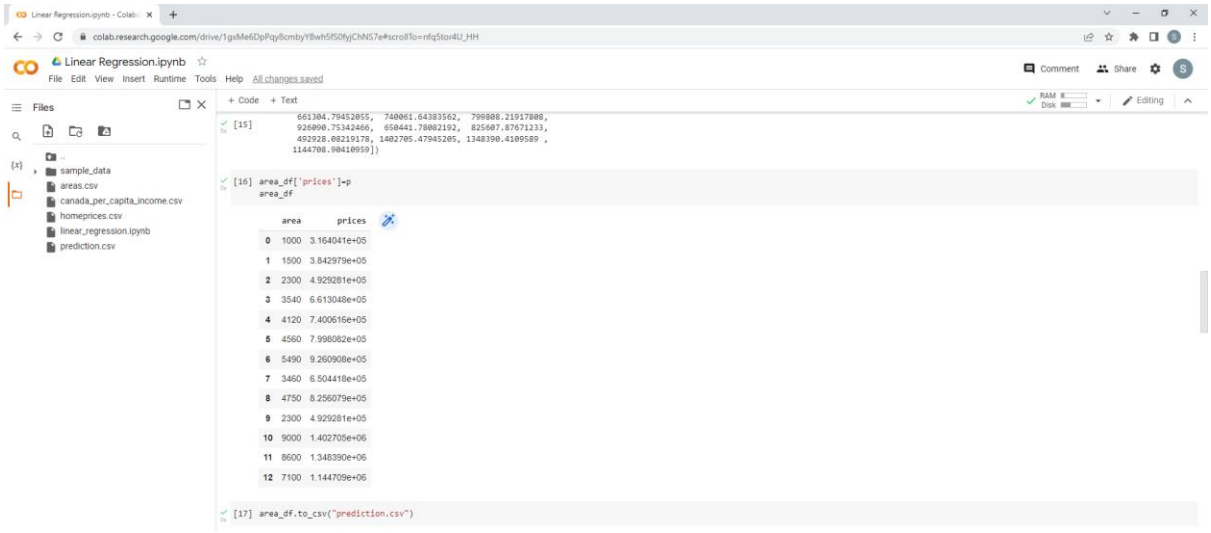
```
import pandas as pd
import numpy as np
from sklearn import linear_model
import matplotlib.pyplot as plt
df = pd.read_csv('homeprices.csv')
df
%matplotlib inline
plt.xlabel('area')
plt.ylabel('price')
plt.scatter(df.area,df.price,color='red',marker='+')
new_df = df.drop('price',axis='columns')
new_df
price = df.price
price
reg = linear_model.LinearRegression()
reg.fit(new_df,price)
reg.predict([[3300]])
reg.coef_
reg.intercept_
reg.predict([[5000]])
area_df = pd.read_csv("areas.csv")
area_df.head(3)
p = reg.predict(area_df)
p
area_df['prices']=p
area_df
area_df.to_csv("prediction.csv")
```

OUTPUT -


Linear Regression.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
sample_data
areas.csv
canada_per_capita_income.csv
homeprices.csv
linear_regression.ipynb
prediction.csv
+ Code + Text
Editing
import pandas as pd
import numpy as np
from sklearn import linear_model
import matplotlib.pyplot as plt
[2] df = pd.read_csv('homeprices.csv')
df
area price
0 2600 550000
1 3000 565000
2 3200 610000
3 3600 680000
4 4000 725000
[3] %matplotlib inline
plt.xlabel('area')
plt.ylabel('price')
plt.scatter(df.area,df.price,color='red',marker='+')
<matplotlib.collections.PathCollection at 0x7f77876dc150>
price
area


Linear Regression.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
sample_data
areas.csv
canada_per_capita_income.csv
homeprices.csv
linear_regression.ipynb
prediction.csv
+ Code + Text
Editing
[4] new_df = df.drop('price',axis='columns')
new_df
area
0 2600
1 3000
2 3200
3 3600
4 4000
[5] price = df.price
price
0 550000
1 565000
2 610000
3 680000
4 725000
Name: price, dtype: int64
[6] reg = linear_model.LinearRegression()
reg.fit(new_df,price)
LinearRegression()
[7] reg.predict([[3300]])
/usr/local/lib/python3.7/dist-packages/sklearn/base.py:451: UserWarning: X does not have valid feature names, but LinearRegression was fitted with feature names
"X does not have valid feature names, but"
array([628715.75342466])
[8] reg.coef_
array([135.78767123])
[10] reg.intercept_
Disk 68.15 GB available
0s completed at 4:48 PM





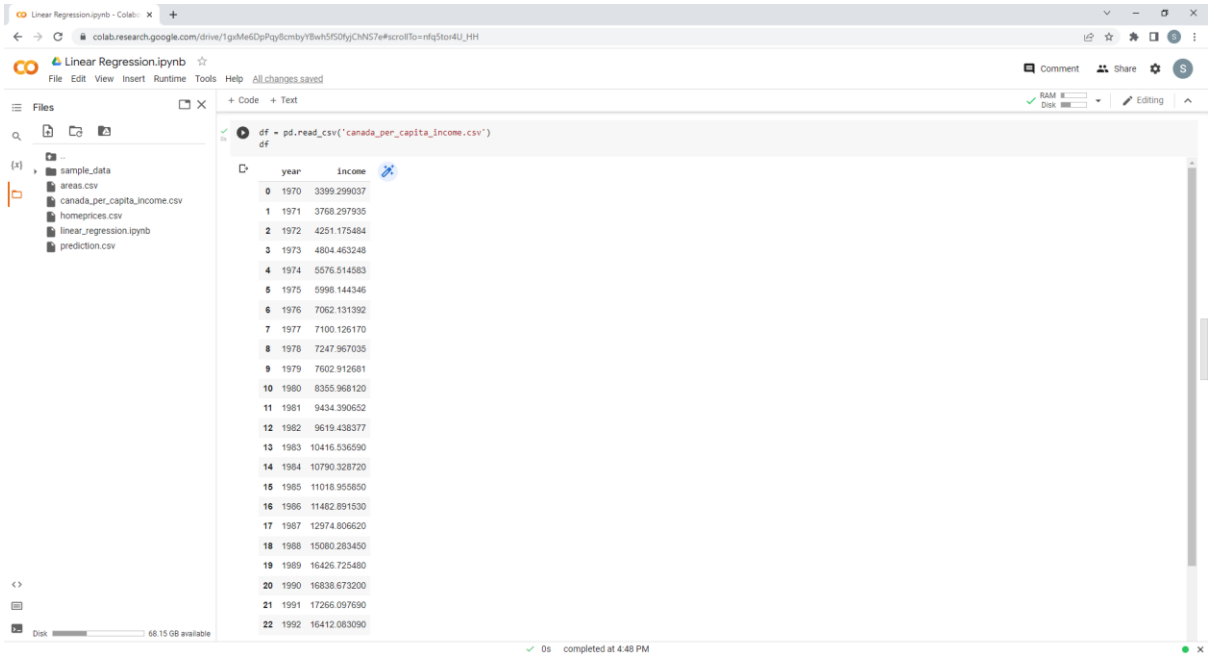
EXERCISE - CANADA INCOME

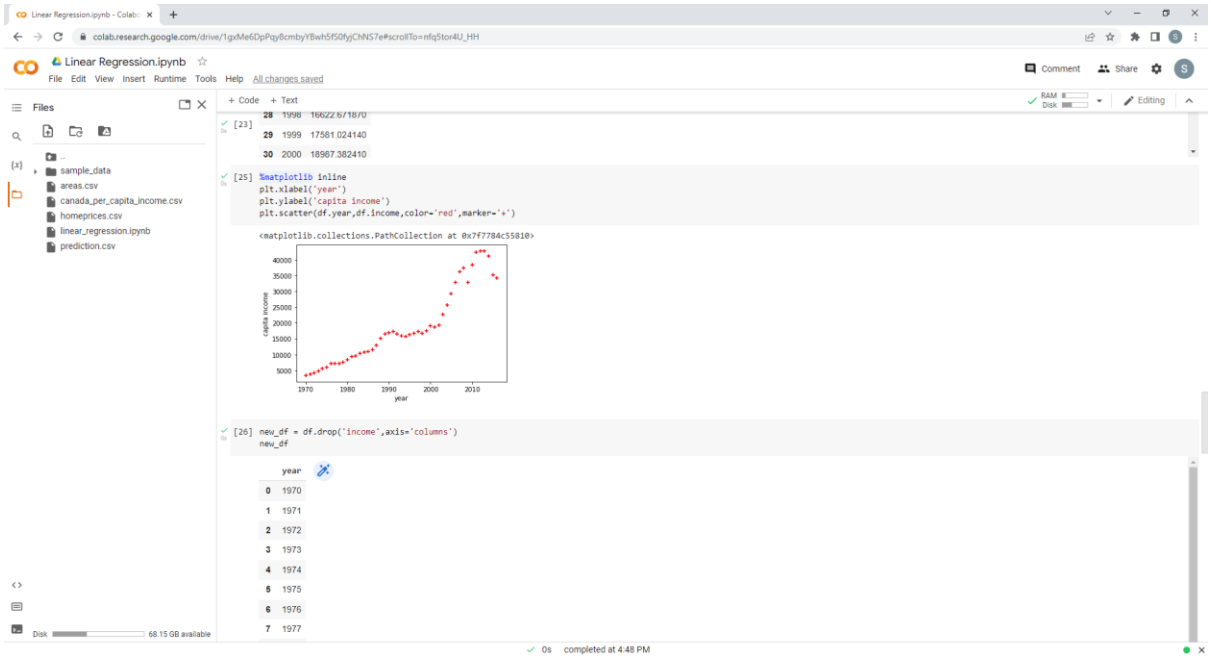
CODE -

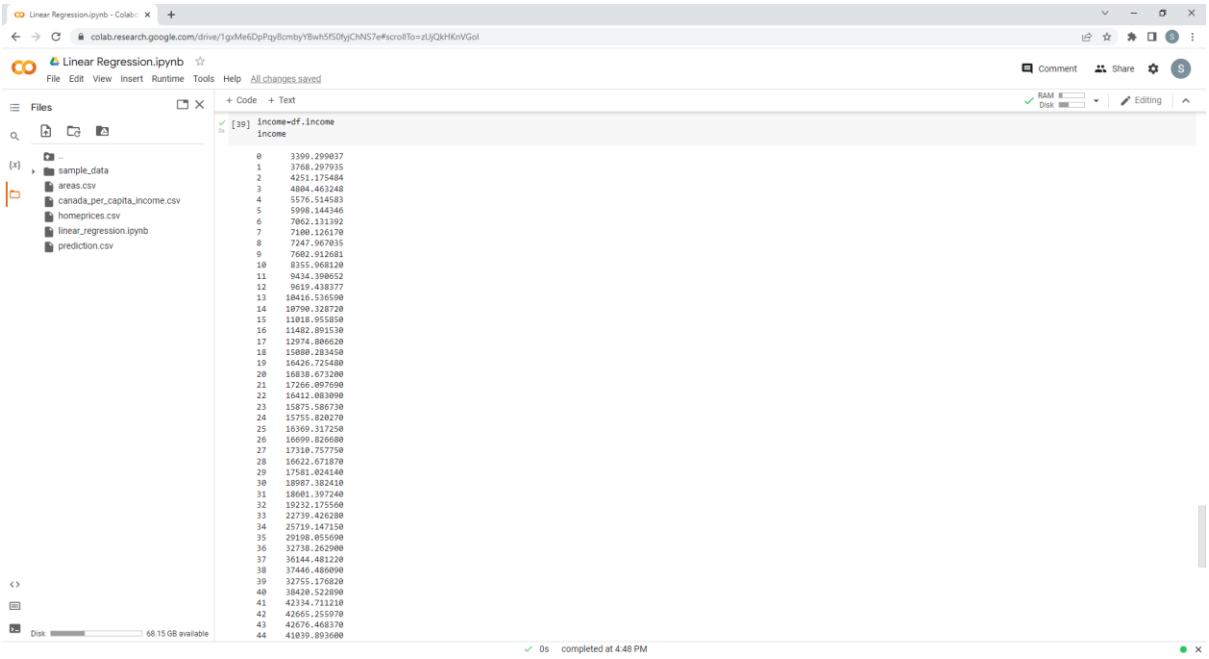
```
df = pd.read_csv('canada_per_capita_income.csv')
df
%matplotlib inline
plt.xlabel('year')
plt.ylabel('capita income')
plt.scatter(df.year,df.income,color='red',marker='+')
new_df = df.drop('income',axis='columns')
new_df
income=df.income
income
reg = linear_model.LinearRegression()
```

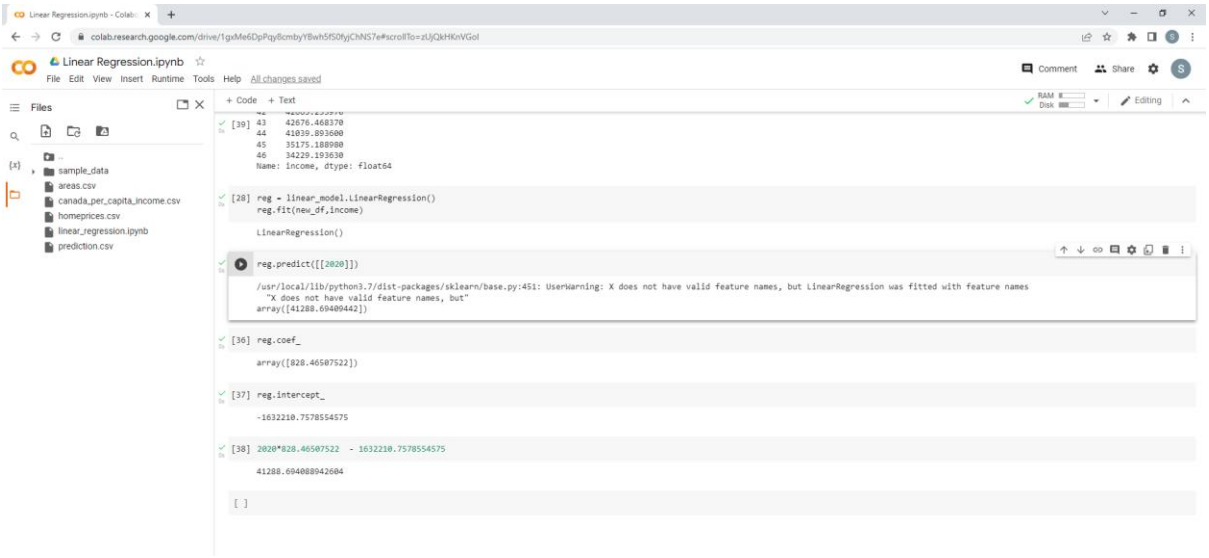
```
reg.fit(new_df,income)
reg.predict([[2020]])
reg.coef_
reg.intercept_
2020*828.46507522  - 1632210.7578554575
```

OUTPUT -






Linear Regression.ipynb - Colabo
colab.research.google.com/drive/1gxMe6DpPqy8cmbyYBwh5f50fyjChNS7e#scrollTo=zUjQkHKnVGol
Linear Regression.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
Editing
sample_data
areas.csv
canada_per_capita_income.csv
homeprices.csv
linear_regression.ipynb
prediction.csv
Disk 68.15 GB available
[39] income=df.income
income
0 3399.299037
1 3768.297935
2 4251.175484
3 4804.463248
4 5576.514583
5 5998.144346
6 7062.131392
7 7100.126170
8 7247.967035
9 7602.912681
10 8355.968120
11 9434.390652
12 9619.438377
13 10416.536590
14 10790.328720
15 11018.955850
16 11482.891530
17 12974.806620
18 15080.283450
19 16426.725480
20 16838.673200
21 17266.097690
22 16412.083090
23 15875.586730
24 15755.820270
25 16369.317250
26 16699.826680
27 17310.757750
28 16622.671870
29 17581.024140
30 18987.382410
31 18601.397240
32 19232.175560
33 22739.426280
34 25719.147150
35 29198.055690
36 32738.262900
37 36144.481220
38 37446.486090
39 32755.176820
40 38420.522890
41 42334.711210
42 42665.255970
43 42676.468370
44 41039.893600
0s completed at 4:48 PM


Linear Regression.ipynb - Colabo
colab.research.google.com/drive/1gxMe6DpPqy8cmbyYBwh5f50fyjChNS7e#scrollTo=zUjQkHKnVGol
Linear Regression.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
Editing
sample_data
areas.csv
canada_per_capita_income.csv
homeprices.csv
linear_regression.ipynb
prediction.csv
[39] 43 42676.468370
44 41039.893600
45 35175.188980
46 34229.193630
Name: income, dtype: float64
[28] reg = linear_model.LinearRegression()
reg.fit(new_df,income)
LinearRegression()
reg.predict([[2020]])
/usr/local/lib/python3.7/dist-packages/sklearn/base.py:451: UserWarning: X does not have valid feature names, but LinearRegression was fitted with feature names
"X does not have valid feature names, but"
array([41288.69409442])
[36] reg.coef_
array([828.46507522])
[37] reg.intercept_
-1632210.7578554575
[38] 2020*828.46507522 - 1632210.7578554575
41288.694088942604
[ ]