SVKM'S NMIMS Deemed-to-be-University
Mukesh Patel School of Technology Management & Engineering
Department of Computer Engineering

Course Code	Computers	Program	B.Tech
Semester	IX	Year	V
Name of the Faculty	Artika Singh	Class	B and C
Course Title	Data Mining	Academic year	2022-23

PART B

Roll No: C035	Name: Krisha Goti
Class: B	Batch: B1
Date of Experiment:3/9/22	Date of Submission: 9/9/22
Grade	

B.1 Work done by student

CODE FOR SIMPLE LINEAR REGRESSION

```
#Krisha Goti C035

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.metrics import mean_squared_error, mean_absolute_error

from sklearn import preprocessing



#Krisha Goti C035

df = pd.read_csv('USA_cars_datasets.csv')

df.head()



#Krisha Goti C035

df_ = df.loc[:,['price','year']]
```

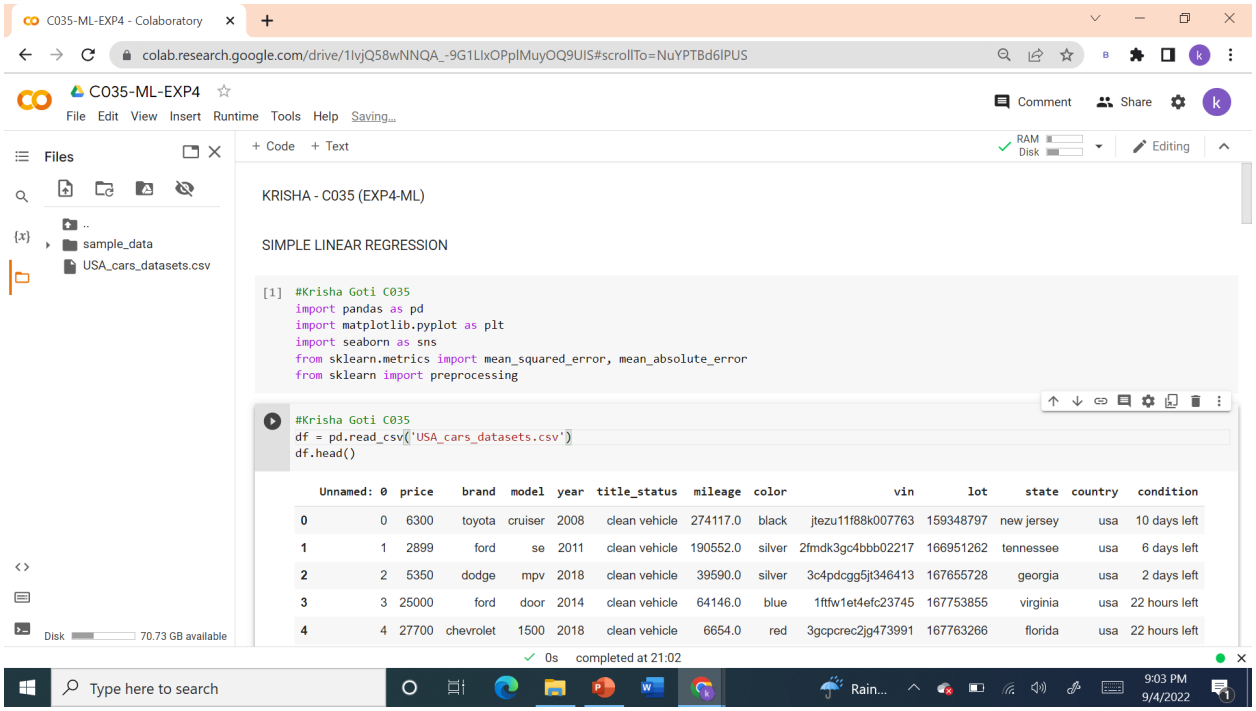
Course Code	Computers	Program	B.Tech
Semester	IX	Year	V
Name of the Faculty	Artika Singh	Class	B and C
Course Title	Data Mining	Academic year	2022-23

```
df.plot(x='year',y='price',style='o')

plt.xlabel('year')

plt.ylabel('price')

plt.show()

x= pd.DataFrame(df['year'])

y= pd.DataFrame(df['price'])

from sklearn.model_selection import train_test_split

x_train, x_test, y_train, y_test = train_test_split(x,y, train_size=0.7
,test_size=0.3, random_state=100)

x_train

y_train


#Krisha Goti C035

import statsmodels.api as sm

x_train_sm = sm.add_constant(x_train)

lr = sm.OLS(y_train, x_train_sm).fit()

lr.params

lr.summary()

plt.scatter(x_train, y_train)

plt.plot(x_train, -2.864e+06 + 1429.5888*x_train, 'r')

plt.xlabel('year')

plt.ylabel('price')
```

Course Code	Computers	Program	B.Tech
Semester	IX	Year	V
Name of the Faculty	Artika Singh	Class	B and C
Course Title	Data Mining	Academic year	2022-23

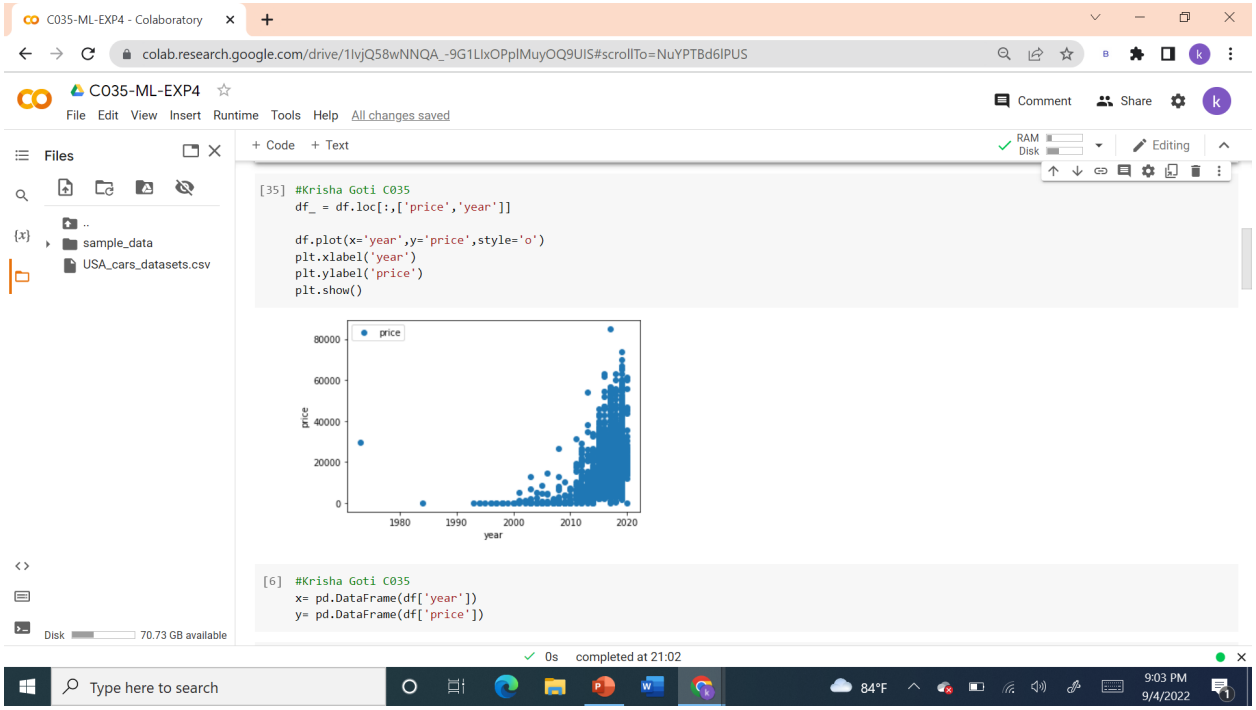
```
plt.show()
```

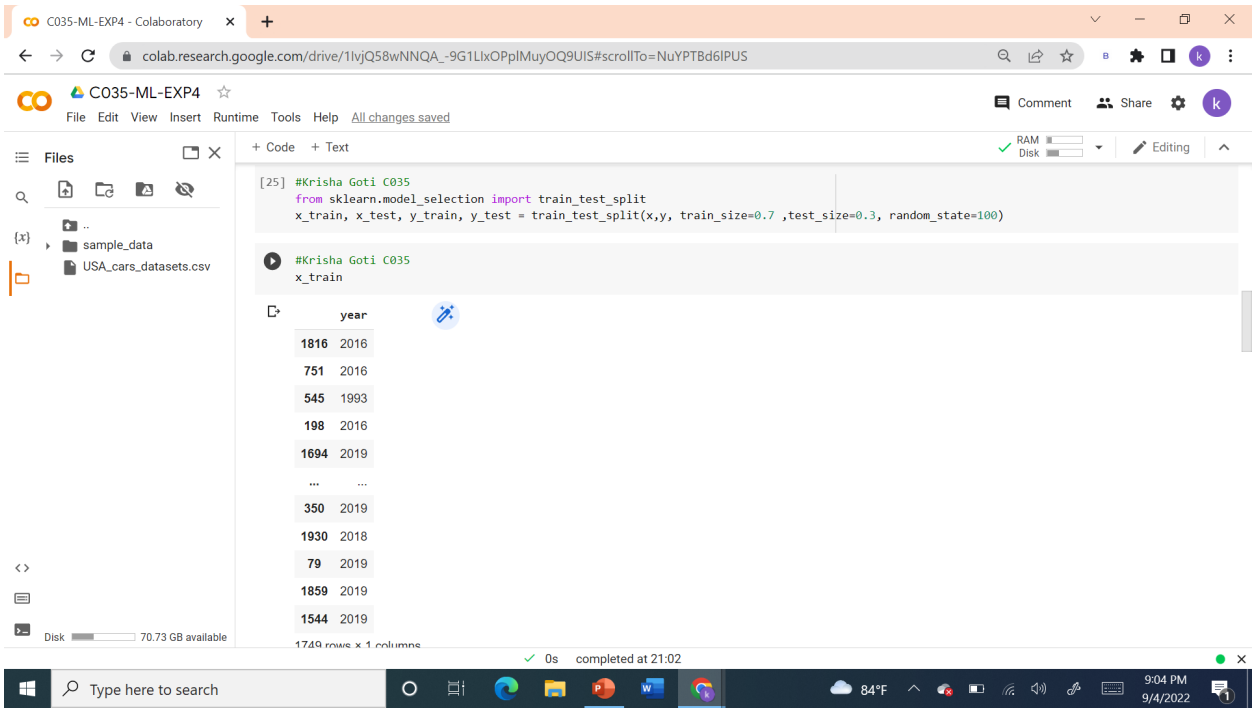
OUTPUT



SVKM'S NMIMS Deemed-to-be-University
Mukesh Patel School of Technology Management & Engineering
Department of Computer Engineering

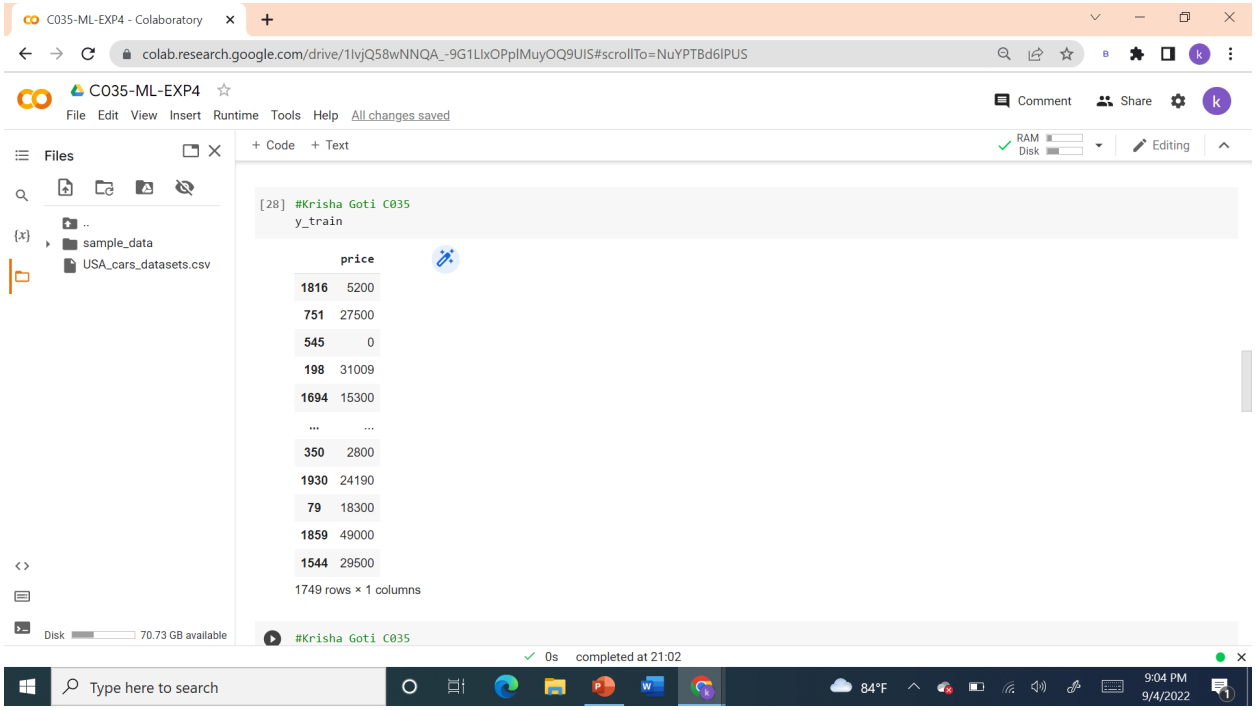
Course Code	Computers	Program	B.Tech
Semester	IX	Year	V
Name of the Faculty	Artika Singh	Class	B and C
Course Title	Data Mining	Academic year	2022-23

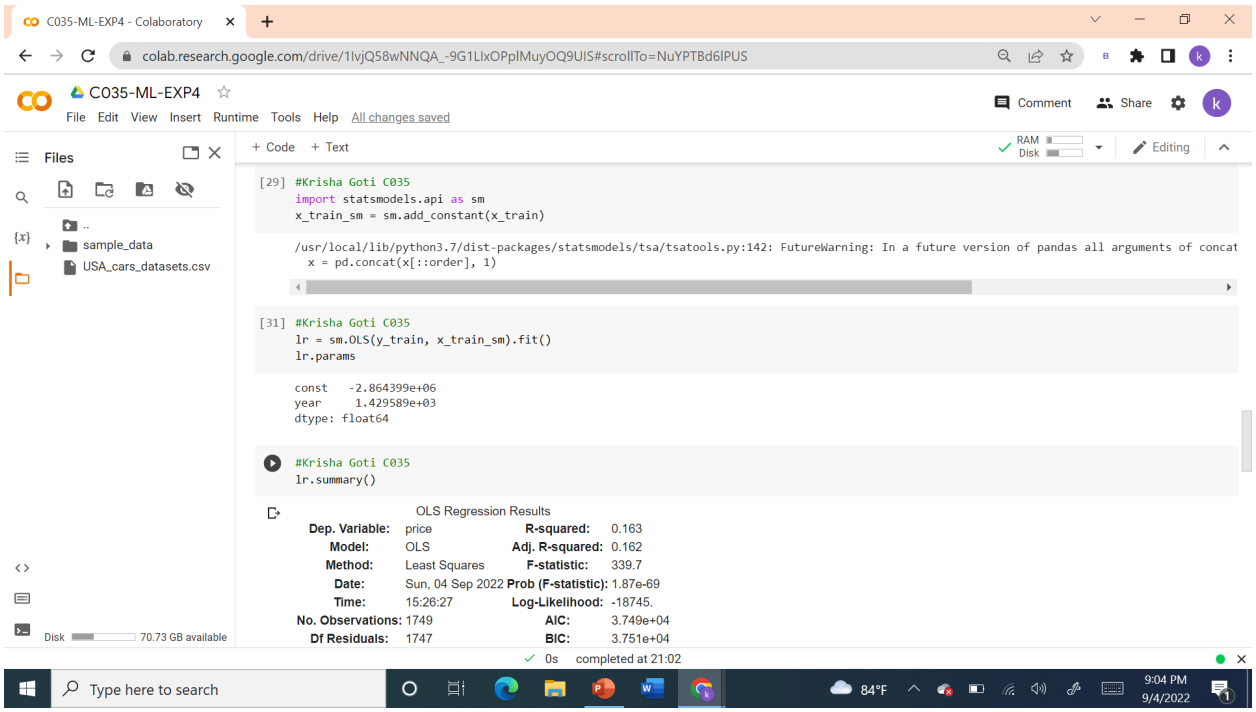




SVKM'S NMIMS Deemed-to-be-University
Mukesh Patel School of Technology Management & Engineering
Department of Computer Engineering

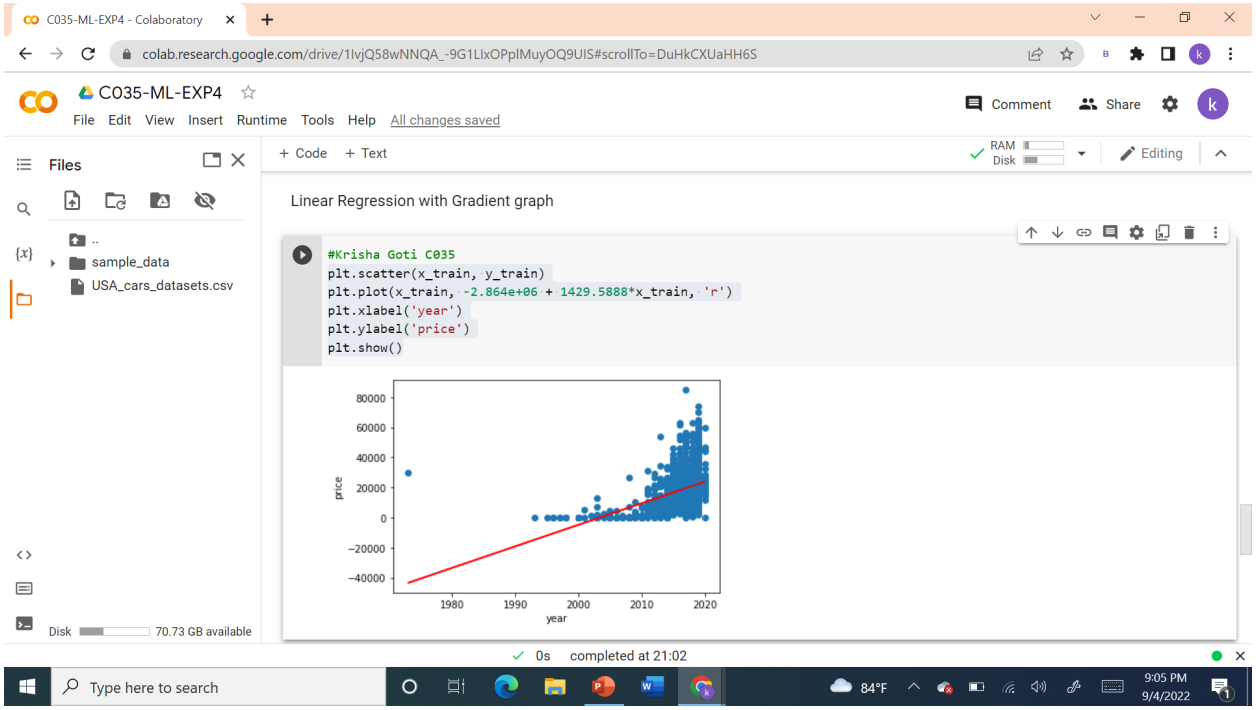
Course Code	Computers	Program	B.Tech
Semester	IX	Year	V
Name of the Faculty	Artika Singh	Class	B and C
Course Title	Data Mining	Academic year	2022-23





SVKM'S NMIMS Deemed-to-be-University
Mukesh Patel School of Technology Management & Engineering
Department of Computer Engineering

Course Code	Computers	Program	B.Tech
Semester	IX	Year	V
Name of the Faculty	Artika Singh	Class	B and C
Course Title	Data Mining	Academic year	2022-23



B.2 Conclusion

After successfully completing this experiment, I am able to:

Apply simple regression on real world dataset and estimate its parameters.Also learnt some new python libraries for plotting a scatter plot.Learnt how to predict dependent variable for dataset.