**Practical 4**

**Aim:**

**Write a program to predict total payment for given number of claims on Swedish auto insurance dataset using linear regression.**

**Code:**

from statistics import mean

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

%matplotlib inline

def best\_fit\_slope(X,y):

slope\_m = ((mean(X)\*mean(y)) - mean(X\*y))/(mean(X)\*\*2 - mean(X\*\*2))

bias\_b = mean(y) - slope\_m\*mean(X)

return slope\_m, bias\_b

from google.colab import files

uploaded = files.upload()

df = pd.read\_excel('slr06.xls')

df.head()

df.describe()

X = np.array(df['X'], dtype=np.float64)

y = np.array(df['Y'], dtype=np.float64)

fig,ax = plt.subplots()

ax.scatter(X,y)

ax.set\_xlabel('X')

ax.set\_ylabel('y')

ax.set\_title('Input Data Scatter Plot')

m,b = best\_fit\_slope(X,y)

print('Slope: ',m)

print('Bias: ',b)

y\_hat = m\*X + b

print('y\_hat: ', y\_hat)

fig,ax = plt.subplots()

ax.scatter(X,y)

ax.set\_xlabel('X')

ax.set\_ylabel('y')

ax.plot(X,y\_hat)

ax.set\_title('Line fit to Input Data')

**Output:**





