import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

dataset = pd.read\_csv(r'C:\Users\HP\Documents\lab\_\data.csv')

dataset

x = dataset.iloc[:,:-1]

y = dataset.iloc[:,1]

x

y

plt.xlabel('No of hours')

plt.ylabel('Risk Score')

plt.scatter(x,y,color='red',marker='\*')

mean\_x = np.mean(x)

mean\_y = np.mean(y)

print(mean\_x, mean\_y)

#initializing our inputs and outputs

X = dataset['No of hours'].values

Y = dataset['risk score'].values

#mean of our inputs and outputs

x\_mean = np.mean(X)

y\_mean = np.mean(Y)

#total number of values

n = len(X)

#using the formula to calculate the b1 and b0

numerator = 0

denominator = 0

for i in range(n):

numerator += (X[i] - x\_mean) \* (Y[i] - y\_mean)

denominator += (X[i] - x\_mean) \*\* 2

b1 = numerator / denominator

b0 = y\_mean - (b1 \* x\_mean)

#printing the coefficient

print(b1, b0)

plt.xlabel('No of hours')

plt.ylabel('Risk Score')

plt.scatter(x,y,color='red',marker='+')

y\_pred = b0 + b1\*x

plt.plot(x, y\_pred, color = "b")

ss\_t=0

ss\_r=0

#mean\_y = np.mean(y)

for i in range(n):

y\_pred = b0 + b1\*X[i]

ss\_t += (y[i] - y\_mean)\*\*2

ss\_t += (y[i] - y\_pred)\*\*2

r2=1-(ss\_r/ss\_t)

print(r2)