import pandas as pd

import numpy as np

data = pd.read\_csv('assignment.csv')

print(data.describe())

X = data['year experience']

Y = data['salary']

mean\_x = np.mean(X)

mean\_y = np.mean(Y)

size = len(X)

numerator = 0

denominator = 0

for i in range(size):

numerator += (X[i] - mean\_x)\*(Y[i] - mean\_y)

denominator += (X[i]-mean\_x) \*\* 2

m = numerator/denominator

c = mean\_y - (m \* mean\_x)

c

max\_x = np.max(X) + 100

min\_x = np.min(X) - 100

x = np.linspace(min\_x, max\_x, 1000)

y = m\*x + c

import matplotlib.pyplot as plt

plt.plot(x, y, color='Blue', label='Regression Line')

plt.scatter(X, Y, c='#ef5423', label='Scatter Plot')

plt.xlabel('Experience')

plt.ylabel('Salary')

plt.legend()

plt.show

n = 0

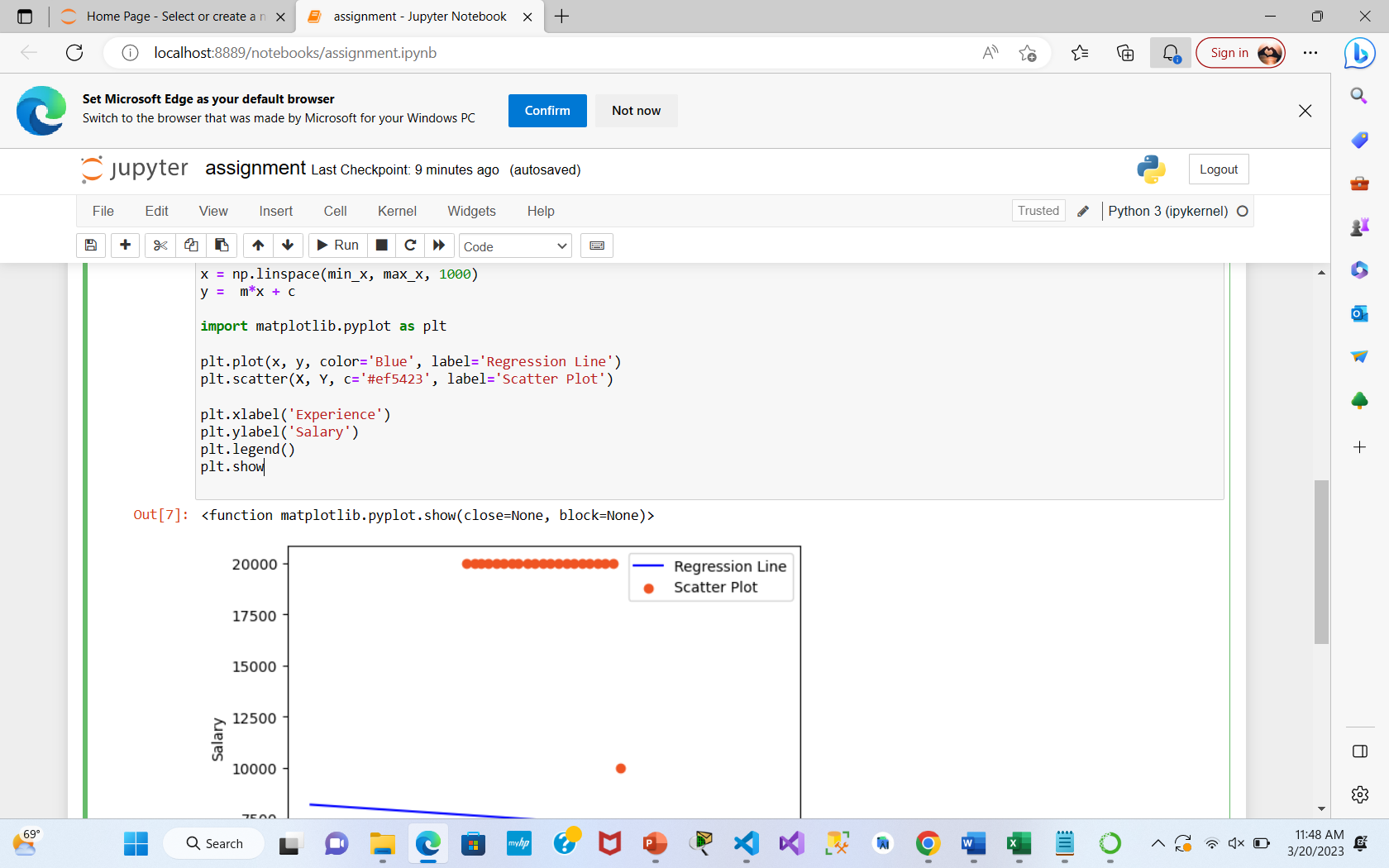
d = 0

for i in range(size):

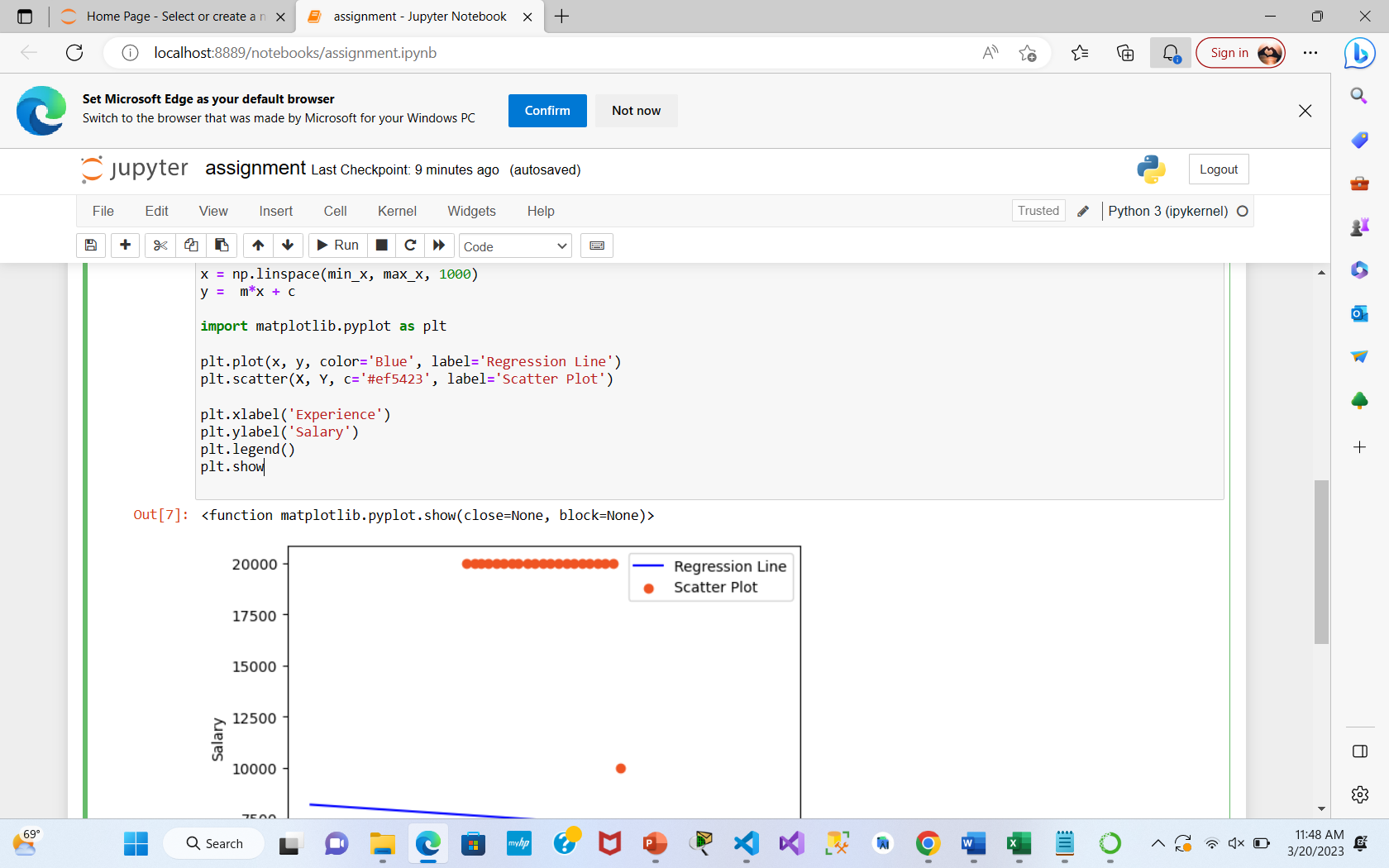
y\_pred = m \* X[i] + c

n += (y\_pred-mean\_y) \*\*2

d += (Y[i] - Meany) \*\*2

square = n/d ..square

.r\_square

.r\_square

