Implementation of Univariate Linear Regression

Aim:

To implement univariate Linear Regression to fit a straight line using least squares.

Equipment's required:

1. Hardware – PCs
2. Anaconda – Python 3.7 Installation / Moodle-Code Runner

Algorithm:

1. Get the independent variable X and dependent variable Y.
2. Calculate the mean of the X -values and the mean of the Y -values.
3. Find the slope m of the line of best fit using the formula.

$$m = \frac{\sum_{i=1}^{n}(x_i-\overline{X})(y_i-\overline{Y})}{\sum_{i=1}^{n}(x_i-\overline{X})^2}$$

4. Compute the y -intercept of the line by using the formula:

$$b = \overline{Y} - m\overline{X}$$

5. Use the slope m and the y -intercept to form the equation of the line.
6. Obtain the straight line equation Y=mX+b and plot the scatterplot.

Program

```
import numpy as np
import matplotlib.pyplot as plt
#Preprocessing the input data
X=np.array(eval(input()))
Y=np.array(eval(input()))
#Building the model
Xmean=np.mean(X)
```

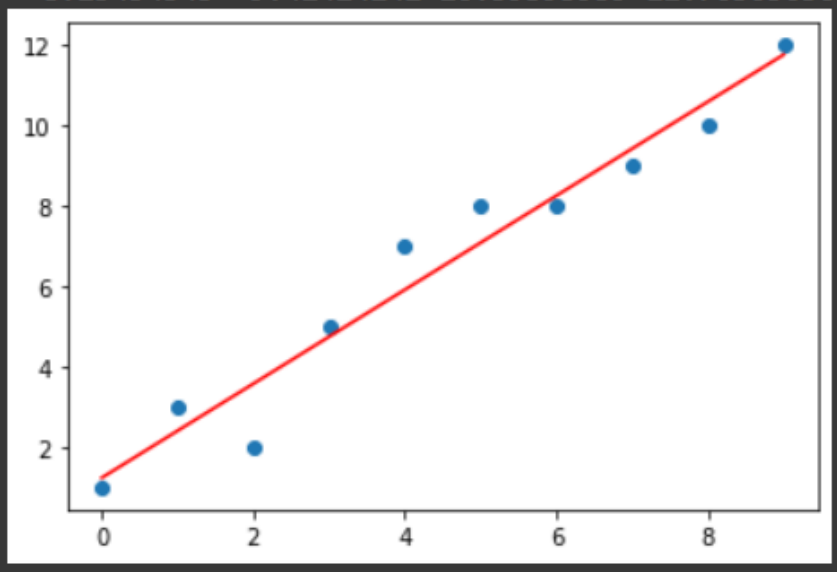
```
Ymean=np.mean(Y)
num,den=0,0
for i in range(len(X)):
  num+=(X[i]-Xmean)*(Y[i]-Ymean)
  den+=(X[i]-Xmean)**2
m=num/den
c=Ymean-m*Xmean
print (m,c)
#Predict the Output
Y_pred = m*X+c
print(Y_pred)
plt.scatter(X,Y)
plt.plot(X,Y_pred,color="red")
plt.show()
```

Sample Input and Output

Input	Result
[0, 1, 2, 3, 4, 5, 6, 7, 8, 9] [1, 3, 2, 5, 7, 8, 8, 9, 10, 12]	1.1696969696969697 1.2363636363636363 [1.23636364 2.40606061 3.57575758 4.74545455 5.91515152 7.08484848 8.25454545 9.42424242 10.59393939 11.76363636]

Output

```
[0, 1, 2, 3, 4, 5, 6, 7, 8, 9]
[1, 3, 2, 5, 7, 8, 8, 9, 10, 12]
1.1696969696969697 1.2363636363636363
[ 1.23636364  2.40606061  3.57575758  4.74545455  5.91515152  7.08484848
  8.25454545  9.42424242 10.59393939 11.76363636]
```



Result

Thus the univariate Linear Regression was implemented to fit a straight line using least squares.