

Procedure:

- 1. Compute $\sum X$, $\sum Y$, $\sum X^2$, $\sum Y^2$ and $\sum XY$.
- 2. Calculate correlation coefficient by

$$\rho = \frac{N \sum XY - \sum X \sum Y}{\sqrt{N \sum X^2 - \left(\sum X\right)^2} \sqrt{N \sum Y^2 - \left(\sum Y\right)^2}}$$

- 3. Compute $\bar{X} = \frac{\sum X}{N}$ and $\bar{Y} = \frac{\sum Y}{N}$
- 4. Calculate regression coefficient by

$$b_{YX} = \frac{N\sum XY - \sum X\sum Y}{N\sum X^2 - \left(\sum X\right)^2}$$

The regression line Y on X is given by

$$Y = b_{YX}(X - \bar{X}) + \bar{Y}$$

6. Plot the given data and the Regression line in a graph.

Program:

```
þ
Developed By: GANESH PRABHU J
Reg No: 212223220023
Dept: IT
import numpy as np
import math
import matplotlib.pyplot as plt
x=[int(i) for i in input().split()]
y=[int(i) for i in input().split()]
N=len(x)
sx=sy=sxy=sx2=sy2=0
for i in range(0,N):
    sx=sx+x[i]
    sy=sy+y[i]
    sxy=sxy+x[i]*y[i]
    sx2=sx2+x[i]**2
    sy2=sy2+x[i]**2
r = (N*sxy-sx*sy)/(math.sqrt(N*sx2-sx*2)*math.sqrt(N*sy2-sy*2))
print("The coorelation coefficient is %0.3f"%r)
byx=(N*sxy-sx*sy)/(N*sx2-sx**2)
xmean=sx/N
ymean=sy/N
print("The regression line Y on X ::: y = \%0.3f \%0.3f (x-\%0.3f)"%(ymean,byx,xmean))
plt.scatter(x,y)
def Reg(x):
    return ymean+byx*(x-xmean)
x=np.linspace(0,80,51)
y1=Reg(x)
plt.plot(x,y1,'r')
plt.xlabel('x-data')
```

```
plt.ylabel('y-data')
plt.legend(['regression line','data points'])
```

```
import matplotlib.pyplot as plt
x=[int(i) for i in input().split()]
y=[int(i) for i in input().split()]
N=len(x)
Sx=0
Sy=0
Sxy=0
Sx2=0
Sy2=0
for i in range(0,N):
 Sx=Sx+x[i]
 Sy=Sy+y[i]
 Sxy=Sxy+x[i]*y[i]
 Sx2=Sx2+x[i]**2
 Sy2=Sy2+y[i]**2
r=(N*Sxy-Sx*Sy)/(math.sqrt(N*Sx2-Sx**2)*math.sqrt(N*Sy2-Sy**2))
print("The Correlation coefficient is %0.3f"%r)
byx=(N*Sxy-Sx*Sy)/(N*Sx2-Sx**2)
xmean=Sx/N
ymean=Sy/N
print("THe Regression line Y on X is ::: y = \%0.3f + \%0.3f (x-\%0.3f)"%(ymean,byx,xmean))
plt.scatter(x,y)
def Reg(x):
 return ymean + byx*(x-xmean)
x=np.linspace(0,80,51)
y1=Reg(x)
plt.plot(x,y1,'r')
plt.xlabel('x-data')
plt.ylabel('y-data')
plt.legend(['Regression Line','Data points'])
```

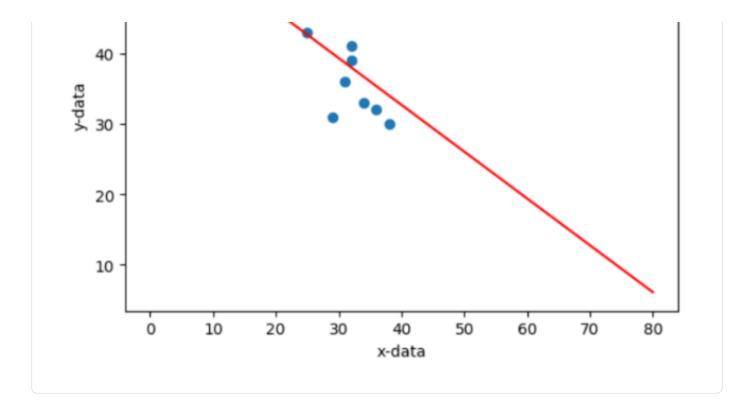
Result

Thus to analyse given data using coefficient of correlation and regression line is successfully completed

Output

```
25 28 35 32 31 36 29 38 34 32
43 46 49 41 36 32 31 30 33 39
The correlation cofficient of -0.394
The Regression line Y on X is ::: 38.000 + -0.664 (x-32.000)
<matplotlib.legend.Legend at 0x7fdc926d9f40>
```





Releases

No releases published

Packages

No packages published

5 of 5