## Program no:7

Aim: Program to implement linear regression techniques using any standard dataset available in the public domain and evaluate its performance(without using inbuilt function)

```
Program
```

```
import numpy as np import matplotlib.pyplot as plt
```

```
def estimate_coef(x, y):
    # number of observations/points
    n = np.size(x)
```

# mean of x and y vector

$$m_x = np.mean(x)$$
  
 $m_y = np.mean(y)$ 

# calculating cross-deviation and deviation about x

$$SS_xy = np.sum(y*x) - n*m_y*m_x$$
  
 $SS_xx = np.sum(x*x) - n*m_x*m_x$ 

# calculating regression coefficients

return (b\_0, b\_1)

```
def plot_regression_line(x, y, b):
    # plotting the actual points as scatter plot
    plt.scatter(x, y, color = "m",
        marker = "o", s = 30)
```

```
# predicted response vector
 y_pred = b[0] + b[1]*x
 # plotting the regression line
 plt.plot(x, y_pred, color = "g")
 # putting labels
 plt.xlabel('x')
 plt.ylabel('y')
 # function to show plot
 plt.show()
def main():
 # observations / data
 x = np.array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9])
 y = np.array([1, 3, 2, 5, 7, 8, 8, 9, 10, 12])
 # estimating coefficients
 b = estimate_coef(x, y)
 print("Estimated coefficients:\nb_0 = {} \
   nb_1 = {}".format(b[0], b[1])
 # plotting regression line
 plot_regression_line(x, y, b)
if __name__ == "__main__":
 main()
```

OUTPUT



