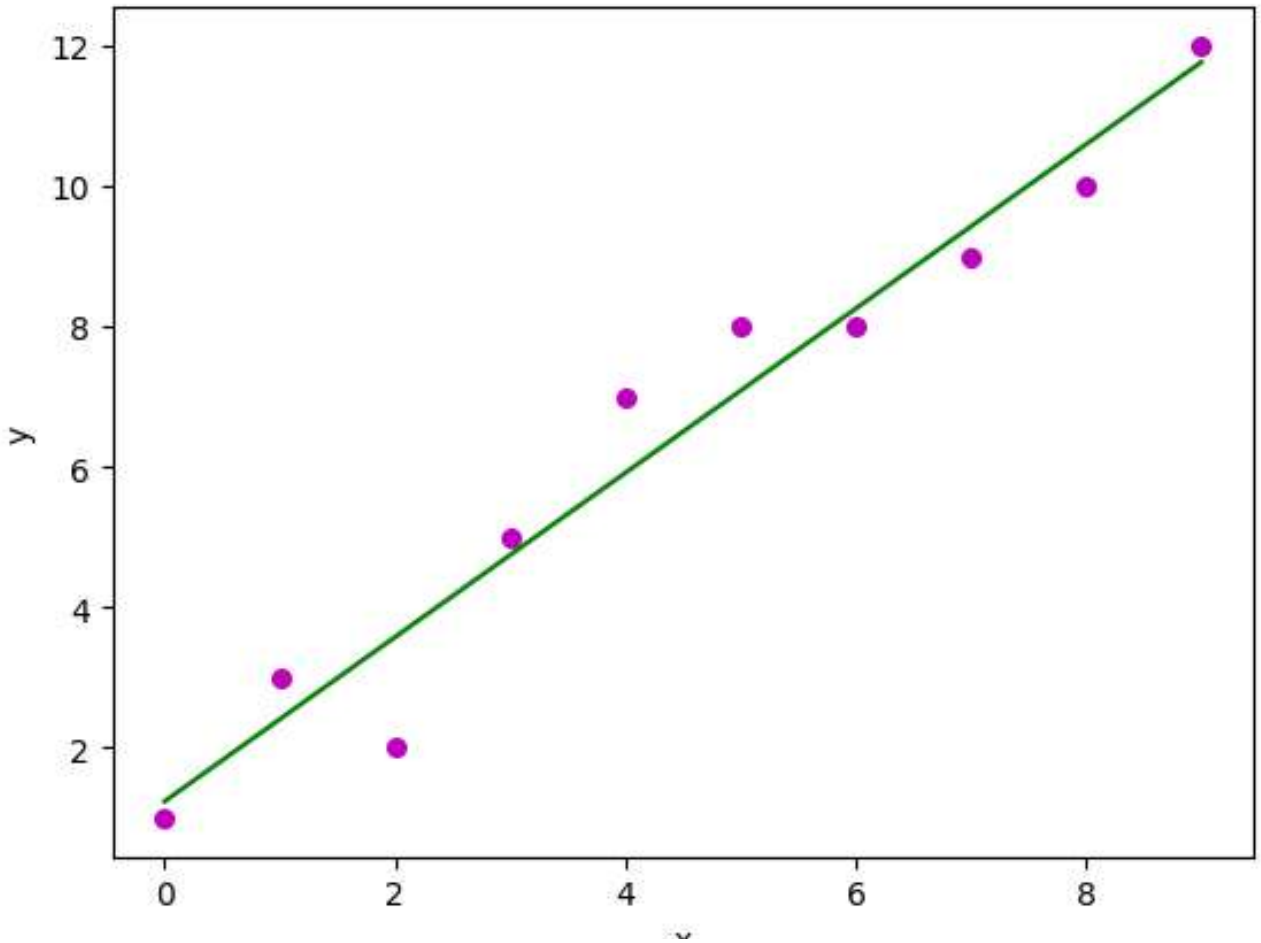
```
import numpy as np
import matplotlib.pyplot as plt

def estimate_coef(x,y):
  n = np.size(x)

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

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

  b_1 = SS_xy / SS_xx
  b_0 = m_y - b_1*m_x

  return(b_0,b_1)

def plot_regression_line(x,y,b):
    plt.scatter(x,y, color = "m", marker="o" , s=30)

    y_pred = b[0] + b[1]*x

    plt.plot(x,y_pred,color ="g")

    plt.xlabel('x')
    plt.ylabel('y')

    plt.show()

def main():
      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])

      b = estimate_coef(x,y)
      print("Estimated coefficient:\nb_0 = {} \nb_1 = {}".format(b[0],b[1]))

      plot_regression_line(x,y,b)

main()
```

```
Estimated coefficient:
b_0 = 1.2363636363636363
b_1 = 1.1696969696969697
```




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


data = pd.read_csv("/content/Salary_Data.csv")


x = data.iloc[:,:-1].values
y = data.iloc[:,1].values


print(data.iloc[:,:-1])
```

```
    YearsExperience
0               1.1
1               1.3
2               1.5
3               2.0
4               2.2
5               2.9
```

```
6               3.0
7               3.2
8               3.2
9               3.7
10              3.9
11              4.0
12              4.0
13              4.1
14              4.5
15              4.9
16              5.1
17              5.3
18              5.9
19              6.0
20              6.8
21              7.1
22              7.9
23              8.2
24              8.7
25              9.0
26              9.5
27              9.6
28             10.3
29             10.5
```

```
print(data.iloc[:,1].values)
```

```
[ 39343.  46205.  37731.  43525.  39891.  56642.  60150.  54445.  64445.
  57189.  63218.  55794.  56957.  57081.  61111.  67938.  66029.  83088.
  81363.  93940.  91738.  98273. 101302. 113812. 109431. 105582. 116969.
 112635. 122391. 121872.]
```

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(x, y,test_size=0.25,random_st
```

```
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
```

```
X_train = scaler.fit_transform(X_train)
X_test = scaler.fit_transform(X_test)
```

```
from sklearn.linear_model import LinearRegression
classifier = LinearRegression()
classifier.fit(X_train,y_train)
```

```
LinearRegression
LinearRegression()
```

```
y_pred = classifier.predict(X_train)
```

+ Code + Text

```
y_pred
```

```
array([103377.33953858,  89618.50845114, 100429.01859127, 108291.2077841 ,
        80773.5456092 ,  44410.92059238,  72911.35641638,  74876.90371458,
        62100.84627624,  67014.71452176,  54238.65708341, 116153.39697693,
       125981.13346796,  62100.84627624,  92566.82939845, 117136.17062603,
        42445.37329417,  63083.61992534,  52273.1097852 ,  37531.50504865,
        33600.41045224, 124015.58616976])
```

```
plt.scatter(X_train, y_train,color="green")
plt.plot(X_train,y_pred,color="red")
plt.title("Salary vs Experience (Training Dataset)")
plt.xlabel("Years of Experience")
plt.ylabel("Salary (In Rupees)")
```

```
Text(0, 0.5, 'Salary (In Rupees)')
```