Vishwakarma Institute of Technology
Department of Artificial Intelligence and Data Science (AI&DS)

Name: **Mustansir Bohari**

Division: **AI-B: 2**

Roll Number: **36**

Subject: **AI**

Lab Assignment 2

```python
import seaborn as sns
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
import sklearn.metrics as sm
```
[1] ✓ 10.0s

```python
df=pd.read_csv("Salary_Data.csv")

X = df['Years']
y = df['Salary']
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=0)
```
[3] ✓ 0.0s

```python
meanX=X_train.mean()
meanX
meanY= y_train.mean()
data=X_train+y_train
# data=np.array(data)
data
```
[4] ✓ 0.0s

```
27    112644.6
11     55798.0
17     83093.3
22    101309.9
5      56644.9
16     66034.1
8      64448.2
14     61115.5
23    113820.2
20     91744.8
1      46206.3
29    121882.5
```

```python
sum=0
sum1=0
for i,j in zip(X_train,y_train):
    sum=sum+(i-meanX)*(j-meanY)
    sum1=sum1+(i-meanX)**2
```
[12] ✓ 0.0s

```python
n=data.shape[0]
cov=sum
cov
```
[13] ✓ 0.0s

```
1491114.0083333333
```
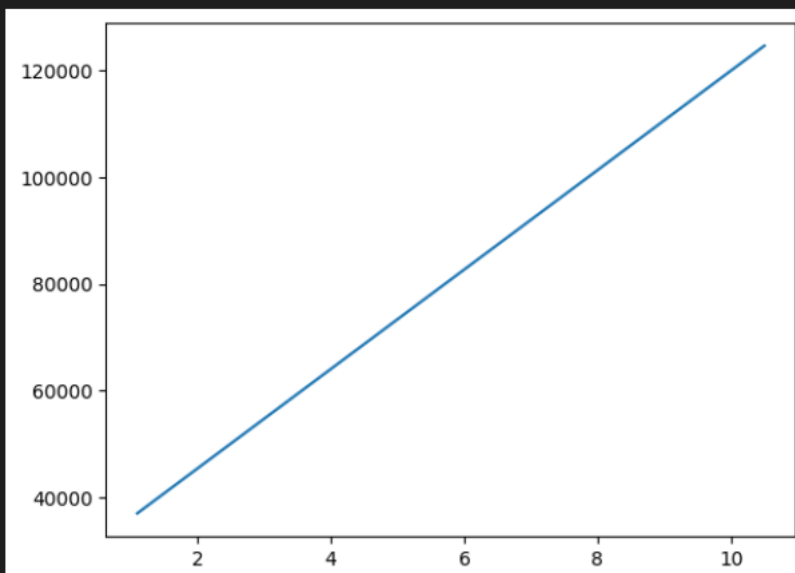
```python
var=(sum1)
var
```
[14] ✓ 0.0s

```
160.11833333333334
```

```
b=cov/var
b
```
[15] ✓ 0.0s

```
9312.575126729189
```

```
a=meanY-(b*meanX)
a
```
[16] ✓ 0.0s

```
26780.09915062818
```

```
y_pred=[]
for i in X:
    y_pred.append(b*i+a)
```
[17] ✓ 0.0s

```
import matplotlib.pyplot as plt
plt.plot(X,y_pred)
x_pred_mean=X.mean()
y_pred_mean=np.array(y_pred).mean()
# meanY**2-y_pred_mean**2
```
[18] ✓ 0.2s

```
print("R2 score =", round(sm.r2_score(y,y_pred), 2))
[19]   ✓ 0.0s

R2 score = 0.96
```