```
from google.colab import drive
drive.mount('/content/drive')
```

> Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.
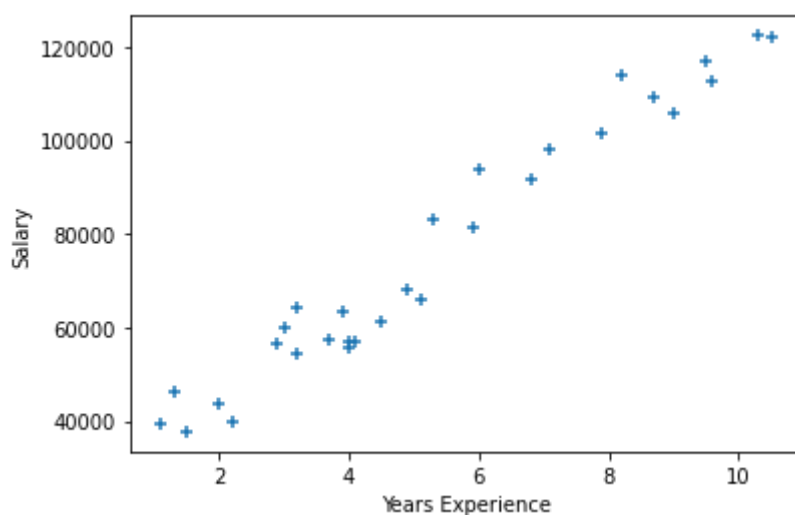
```
import pandas as pd
import numpy as np
```

```
df=pd.read_csv("/content/drive/MyDrive/Salary_Data.csv")
df.head()
```

	YearsExperience	Salary
0	1.1	39343.0
1	1.3	46205.0
2	1.5	37731.0
3	2.0	43525.0
4	2.2	39891.0

```
import matplotlib.pyplot as plt
plt.scatter(df.YearsExperience,df.Salary,marker='+')
plt.xlabel("Years Experience")
plt.ylabel("Salary")
plt.show()
```



```
X=df.iloc[:,:-1].values
Y=df.iloc[:,1:].values #we get an array
X
#X is independent var
#Y is depe var
```

> array([[1.1],

```
          [ 1.3],
          [ 1.5],
          [ 2. ],
          [ 2.2],
          [ 2.9],
          [ 3. ],
          [ 3.2],
          [ 3.2],
          [ 3.7],
          [ 3.9],
          [ 4. ],
          [ 4. ],
          [ 4.1],
          [ 4.5],
          [ 4.9],
          [ 5.1],
          [ 5.3],
          [ 5.9],
          [ 6. ],
          [ 6.8],
          [ 7.1],
          [ 7.9],
          [ 8.2],
          [ 8.7],
          [ 9. ],
          [ 9.5],
          [ 9.6],
          [10.3],
          [10.5]])
```

Y

```
array([[ 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.]])
```

```python
train_x=np.array(df[["YearsExperience"]])
train_y=np.array(df[["Salary"]])
print(train_x)
```

```
    [[ 1.1]
     [ 1.3]
     [ 1.5]
     [ 2. ]
     [ 2.2]
     [ 2.9]
     [ 3. ]
     [ 3.2]
     [ 3.2]
     [ 3.7]
     [ 3.9]
     [ 4. ]
     [ 4. ]
     [ 4.1]
     [ 4.5]
     [ 4.9]
     [ 5.1]
     [ 5.3]
     [ 5.9]
     [ 6. ]
     [ 6.8]
     [ 7.1]
     [ 7.9]
     [ 8.2]
     [ 8.7]
     [ 9. ]
     [ 9.5]
     [ 9.6]
     [10.3]
     [10.5]]
```

```python
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(X,Y,test_size=0.3,random_state=0)
```

```python
x_train
```

```
    array([[ 7.9],
           [ 2.9],
           [ 5.1],
           [ 3.2],
           [ 4.5],
           [ 8.2],
           [ 6.8],
           [ 1.3],
           [10.5],
           [ 3. ],
           [ 2.2],
           [ 5.9],
           [ 6. ],
```

```
       [ 3.7],
       [ 3.2],
       [ 9. ],
       [ 2. ],
       [ 1.1],
       [ 7.1],
       [ 4.9],
       [ 4. ]])
```

```python
#training the model
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(x_train,y_train)
```

```
    LinearRegression()
```

```python
#testing the model
y_pred = model.predict(x_test)
```

```python
plt.scatter(x_train,y_train,color='red')
plt.plot(x_train,model.predict(x_train),color='blue')
plt.title('Salary vs Experience (Training set)')
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.show()
```

✓ 0s   completed at 9:18 PM   ● ✕