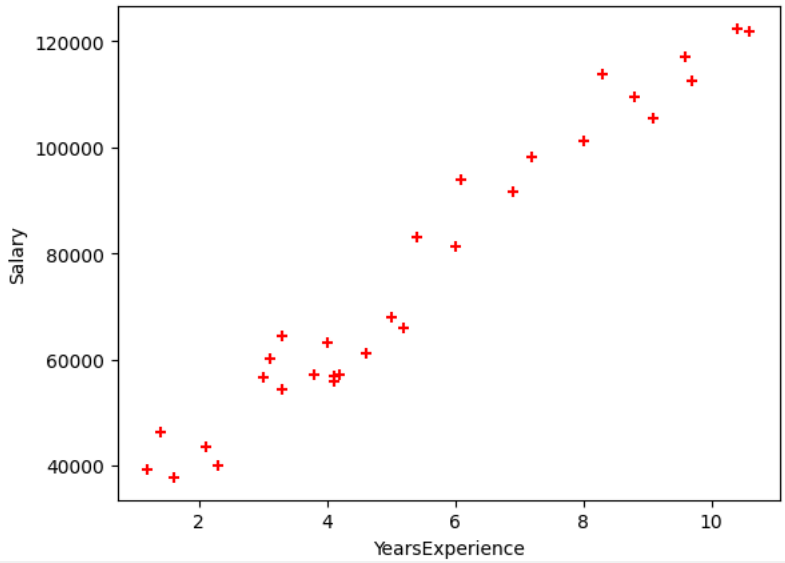
```
import pandas as pd
import numpy as np
from sklearn import linear_model
import matplotlib.pyplot as plt

df = pd.read_csv('/content/drive/MyDrive/Salary_dataset.csv')
df
```

| | Unnamed: 0 | YearsExperience | Salary |
|---|---|---|---|
| **0** | 0 | 1.2 | 39344 |
| **1** | 1 | 1.4 | 46206 |
| **2** | 2 | 1.6 | 37732 |
| **3** | 3 | 2.1 | 43526 |
| **4** | 4 | 2.3 | 39892 |
| **5** | 5 | 3.0 | 56643 |
| **6** | 6 | 3.1 | 60151 |
| **7** | 7 | 3.3 | 54446 |
| **8** | 8 | 3.3 | 64446 |
| **9** | 9 | 3.8 | 57190 |
| **10** | 10 | 4.0 | 63219 |
| **11** | 11 | 4.1 | 55795 |
| **12** | 12 | 4.1 | 56958 |
| **13** | 13 | 4.2 | 57082 |
| **14** | 14 | 4.6 | 61112 |
| **15** | 15 | 5.0 | 67939 |
| **16** | 16 | 5.2 | 66030 |
| **17** | 17 | 5.4 | 83089 |
| **18** | 18 | 6.0 | 81364 |
| **19** | 19 | 6.1 | 93941 |
| **20** | 20 | 6.9 | 91739 |
| **21** | 21 | 7.2 | 98274 |
| **22** | 22 | 8.0 | 101303 |
| **23** | 23 | 8.3 | 113813 |
| **24** | 24 | 8.8 | 109432 |
| **25** | 25 | 9.1 | 105583 |
| **26** | 26 | 9.6 | 116970 |
| **27** | 27 | 9.7 | 112636 |
| **28** | 28 | 10.4 | 122392 |
| **29** | 29 | 10.6 | 121873 |

```
plt.xlabel('YearsExperience')
plt.ylabel('Salary')
plt.scatter(df.YearsExperience,df.Salary,color='red',marker='+')
```

```
<matplotlib.collections.PathCollection at 0x7fc0aeb93bb0>
```




```
new_df = df.drop('Salary', axis='columns')
new_df

x= new_df.drop('Unnamed: 0',axis=1)
x

Salary = df.Salary
Salary
# # # Create linear regression object
reg = linear_model.LinearRegression()
reg.fit(x, Salary)
reg.predict([[10]])
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does not have valid feature names, but LinearRegression
  warnings.warn(
array([119347.82718107])
```

```
plt.xlabel('YearsExperience')
plt.ylabel('Salary')
plt.scatter(df.YearsExperience,df.Salary,color='red',marker='+')
plt.plot(df.YearsExperience,reg.predict(df[['YearsExperience']]),color='blue')
```

```
[<matplotlib.lines.Line2D at 0x7fc0ac77f580>]
```