```
# LINER REGRSSION
# Dataset
# to create dataframe
import pandas as pd

df = pd.read_csv('https://raw.githubusercontent.com/ameenmanna8824/DATASETS/main/Salary_Da
print(df)

# data visualisation

import matplotlib.pyplot as plt
#import numpy as np
plt.plot(df['YearsExperience'],df['Salary'])
plt.title('Salary vs YearsExperience')
#plt.xlable('YearsExperience')
#plt.ylable('Salary')
#divide into input and output(x -i/p, y-i\p)
x = df.iloc[:,0:101].values
y = df.iloc[:,1].values
print(x)
print(y)
```

```
    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
5               2.9   56642.0
6               3.0   60150.0
7               3.2   54445.0
8               3.2   64445.0
9               3.7   57189.0
10              3.9   63218.0
11              4.0   55794.0
12              4.0   56957.0
13              4.1   57081.0
14              4.5   61111.0
15              4.9   67938.0
16              5.1   66029.0
17              5.3   83088.0
18              5.9   81363.0
19              6.0   93940.0
20              6.8   91738.0
21              7.1   98273.0
22              7.9  101302.0
23              8.2  113812.0
24              8.7  109431.0
25              9.0  105582.0
26              9.5  116969.0
27              9.6  112635.0
28             10.3  122391.0
29             10.5  121872.0
[[1.10000e+00 3.93430e+04]
 [1.30000e+00 4.62050e+04]
 [1.50000e+00 3.77310e+04]
 [2.00000e+00 4.35250e+04]
 [2.20000e+00 3.98910e+04]
 [2.90000e+00 5.66420e+04]
 [3.00000e+00 6.01500e+04]
 [3.20000e+00 5.44450e+04]
 [3.20000e+00 6.44450e+04]
 [3.70000e+00 5.71890e+04]
 [3.90000e+00 6.32180e+04]
 [4.00000e+00 5.57940e+04]
 [4.00000e+00 5.69570e+04]
 [4.10000e+00 5.70810e+04]
 [4.50000e+00 6.11110e+04]
 [4.90000e+00 6.79380e+04]
 [5.10000e+00 6.60290e+04]
 [5.30000e+00 8.30880e+04]
 [5.90000e+00 8.13630e+04]
 [6.00000e+00 9.39400e+04]
 [6.80000e+00 9.17380e+04]
 [7.10000e+00 9.82730e+04]
```

```
#Run a classifer
from sklearn.linear_model import LinearRegression
model = LinearRegression()
#fit the model
model.fit(x,y)
```

```
#predict the output
y_pred=model.predict(x)
y_pred
```

```
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.])
```

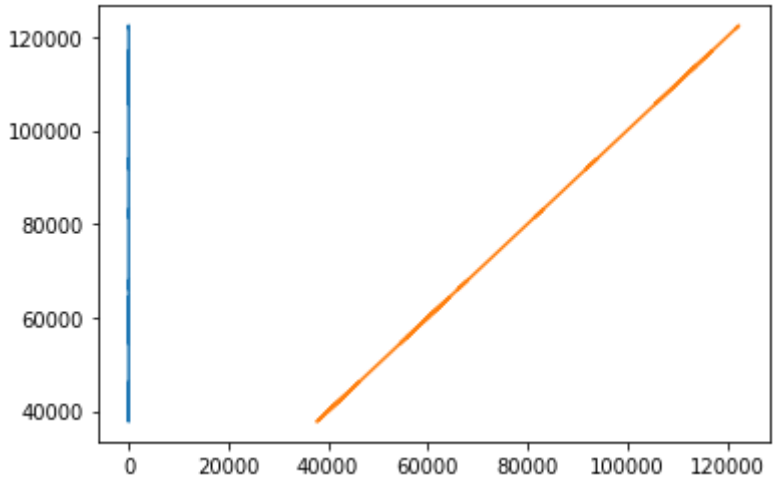
Double-click (or enter) to edit

```
model.coef_
model.intercept_
model.predict([[15000.67,10]])
```

```
array([10.00000002])
```

```
#post Model visualization
#plt.scatter(x,y)
plt.plot(x,y_pred)
```

```
[<matplotlib.lines.Line2D at 0x7f2e6f202b50>,
 <matplotlib.lines.Line2D at 0x7f2e6f202450>]
```



✓ 0s completed at 11:49 PM