In [26]: import numpy as np
 import pandas as pd
 import seaborn as sns
 df = pd.read_excel('C:/Users/SRIVENKATESH/Downloads/Salary_Data new.xlsx')

In [2]: df

Out[2]:

:		Experience	Salary
	0	10	39343
	1	20	46205
	2	30	37731
	3	40	38500
	4	50	39891
	5	60	56642
	6	70	60150
	7	80	62000
	8	90	64445
	9	100	57189
	10	90	63218
	11	80	60000
	12	70	56957
	13	60	57081
	14	50	61111
	15	40	65000
	16	30	66029
	17	20	83088
	18	10	81363
	19	20	93940
	20	30	91738
	21	40	98273
	22	50	101302
	23	60	113812
	24	70	109431
	25	80	105582
	26	90	116969
	27	100	112635
	28	90	122391
	29	80	121872

```
In [3]: df=pd.DataFrame(df,columns=['Experience','Salary'])
In [4]:
         df.plot.scatter(x='Experience',y='Salary')
         <AxesSubplot:xlabel='Experience', ylabel='Salary'>
Out[4]:
           120000
           100000
            80000
            60000
            40000
                        20
                                   40
                                             60
                                                       80
                                                                100
                                       Experience
In [5]:
         df.columns
         Index(['Experience', 'Salary'], dtype='object')
Out[5]:
         X=df[['Experience']]
In [6]:
In [7]:
```

Out[7]:		Experience
	0	10
	1	20
	2	30
	3	40
	4	50
	5	60
	6	70
	7	80
	8	90
	9	100
	10	90
	11	80
	12	70
	13	60
	14	50
	15	40
	16	30
	17	20
	18	10
	19	20
	20	30
	21	40
	22	50
	23	60
	24	70
	25	80
	26	90
	27	100
	28	90
	29	80

```
In [8]: Y=df[['Salary']]
In [9]: Y
```

 39343 46205 37731 38500 39891 556642 660150 62000 64445 57189 63218 60000 56957 57081 6111 65000 66029 83088 81363 93940 91738 98273 198273 1993940 91738 98273 101302 113812 109431 105582 116969 112635 12391 121872 	Out[9]:		Salary
2 37731 3 38500 4 39891 5 56642 6 60150 7 62000 8 64445 9 57189 10 63218 11 60000 12 56957 13 57081 14 61111 15 65000 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391		0	39343
3 38500 4 39891 5 56642 6 60150 7 62000 8 64445 9 57189 10 63218 11 60000 12 56957 13 57081 14 61111 15 65000 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391		1	46205
4 39891 5 56642 6 60150 7 62000 8 64445 9 57189 10 63218 11 60000 12 56957 13 57081 14 61111 15 65000 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391		2	37731
5 56642 6 60150 7 62000 8 64445 9 57189 10 63218 11 60000 12 56957 13 57081 14 61111 15 65000 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391		3	38500
6 60150 7 62000 8 64445 9 57189 10 63218 11 60000 12 56957 13 57081 14 61111 15 65000 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391		4	39891
7 62000 8 64445 9 57189 10 63218 11 60000 12 56957 13 57081 14 61111 15 65000 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391		5	56642
 8 64445 9 57189 10 63218 11 60000 12 56957 13 57081 14 61111 15 65000 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391 		6	60150
9 57189 10 63218 11 60000 12 56957 13 57081 14 61111 15 65000 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391		7	62000
10632181160000125695713570811461111156500016660291783088188136319939402091738219827322101302231138122410943125105582261169692711263528122391		8	64445
11 60000 12 56957 13 57081 14 61111 15 65000 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391		9	57189
12 56957 13 57081 14 61111 15 65000 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391		10	63218
13570811461111156500016660291783088188136319939402091738219827322101302231138122410943125105582261169692711263528122391		11	60000
1461111156500016660291783088188136319939402091738219827322101302231138122410943125105582261169692711263528122391		12	56957
15 65000 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391		13	57081
 16 66029 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391 		14	61111
 17 83088 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391 		15	65000
 18 81363 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391 		16	66029
 19 93940 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391 		17	83088
 20 91738 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391 		18	81363
 21 98273 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391 		19	93940
 22 101302 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391 		20	91738
 23 113812 24 109431 25 105582 26 116969 27 112635 28 122391 		21	98273
 24 109431 25 105582 26 116969 27 112635 28 122391 		22	101302
25 10558226 11696927 11263528 122391		23	113812
26 11696927 11263528 122391		24	109431
27 11263528 122391		25	105582
28 122391		26	116969
		27	112635
29 121872		28	122391
		29	121872

```
In [10]: from sklearn.model_selection import train_test_split
In [11]: xtrain,xtest,ytrain,ytest=train_test_split(X,Y,test_size=0.2,random_state=0)
In [12]: xtrain
```

Out[12]:		Experience
	27	100
	11	80
	17	20
	22	50
	5	60
	16	30
	8	90
	14	50
	23	60
	20	30
	1	20
	29	80
	6	70
	4	50
	18	10
	19	20
	9	100
	7	80
	25	80
	3	40
	0	10
	21	40
	15	40
	12	70

In [13]: xtest

Out[13]:

	Experience
2	30
28	90
13	60
10	90
26	90
24	70

In [14]: ytrain

Out[14]:		Salary
	27	112635
	11	60000
	17	83088
	22	101302
	5	56642
	16	66029
	8	64445
	14	61111
	23	113812
	20	91738
	1	46205
	29	121872
	6	60150
	4	39891
	18	81363
	19	93940
	9	57189
	7	62000
	25	105582
	3	38500
	0	39343
	21	98273
	15	65000
	12	56957

In [15]: ytest

```
Out[15]: Salary

2 37731

28 122391

13 57081

10 63218

26 116969
```

In [16]: from sklearn.linear_model import LinearRegression

```
In [17]: lm=LinearRegression()
```

24 109431

```
lm.fit(xtrain,ytrain)
In [18]:
          LinearRegression()
Out[18]:
          predictions=lm.predict(xtest)
In [19]:
          predictions
In [20]:
          array([[70170.46143123],
Out[20]:
                 [80132.16775093],
                 [75151.31459108],
                 [80132.16775093],
                 [80132.16775093],
                 [76811.5989777 ]])
          predictions=pd.DataFrame(predictions, columns=['predictions'])
In [21]:
          predictions
In [22]:
              predictions
Out[22]:
            70170.461431
          1 80132.167751
          2 75151.314591
          3 80132.167751
             80132.167751
            76811.598978
In [28]:
          sns.regplot(x='Experience',y='Salary',data=df,fit_reg=True)
          <AxesSubplot:xlabel='Experience', ylabel='Salary'>
Out[28]:
            120000
            100000
             80000
             60000
             40000
                          20
                                                        80
                                                                  100
                                              60
                                        Experience
```