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
```

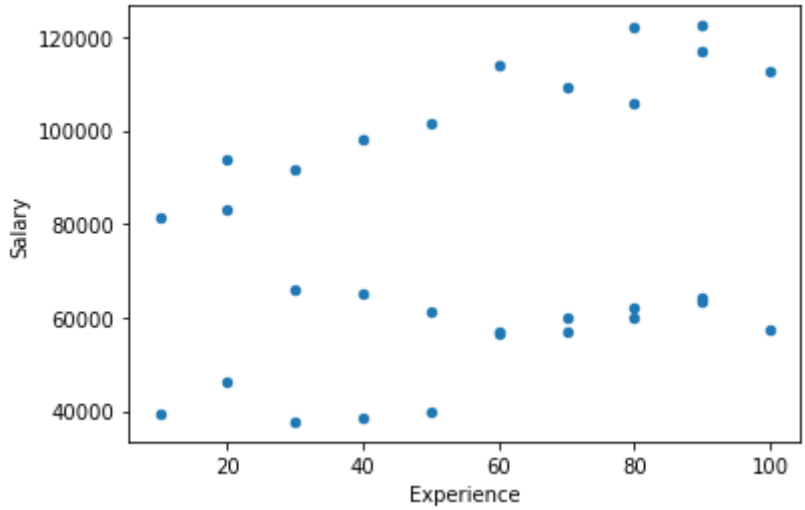
Out[2]:

	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')
```

Out[4]:
```
<AxesSubplot:xlabel='Experience', ylabel='Salary'>
```



In [5]:
```
df.columns
```

Out[5]:
```
Index(['Experience', 'Salary'], dtype='object')
```

In [6]:
```
X=df[['Experience']]
```

In [7]:
```
X
```

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
```

Out[9]:

	Salary
0	39343
1	46205
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
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
24	109431

In [16]:
```
from sklearn.linear_model import LinearRegression
```

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

In [18]:
```
lm.fit(xtrain,ytrain)
```

Out[18]:
```
LinearRegression()
```

In [19]:
```
predictions=lm.predict(xtest)
```

In [20]:
```
predictions
```

Out[20]:
```
array([[70170.46143123],
       [80132.16775093],
       [75151.31459108],
       [80132.16775093],
       [80132.16775093],
       [76811.5989777 ]])
```

In [21]:
```
predictions=pd.DataFrame(predictions,columns=['predictions'])
```

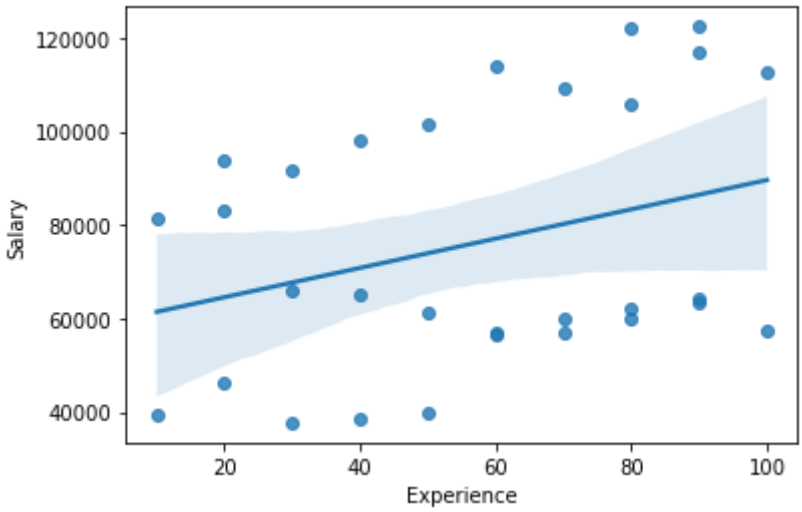
In [22]:
```
predictions
```

Out[22]:

	predictions
0	70170.461431
1	80132.167751
2	75151.314591
3	80132.167751
4	80132.167751
5	76811.598978

In [28]:
```
sns.regplot(x='Experience',y='Salary',data=df,fit_reg=True)
```

Out[28]:
```
<AxesSubplot:xlabel='Experience', ylabel='Salary'>
```



In [ ]:
```
```