In [1]:
```
#import data set
```

In [2]:
```
import pandas as pd
```

In [3]:
```
data=pd.read_csv('C:\\Users\\DELL\\Downloads\\Salary_Data.csv')
data
```

Out[3]:

	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

	YearsExperience	Salary
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

In [4]:
```
data.head()
```

Out[4]:

	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

In [5]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30 entries, 0 to 29
Data columns (total 2 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   YearsExperience  30 non-null     float64
```

```
 1   Salary           30 non-null     float64
dtypes: float64(2)
memory usage: 608.0 bytes
```

In [6]:
```
#correlation
data.corr()
```

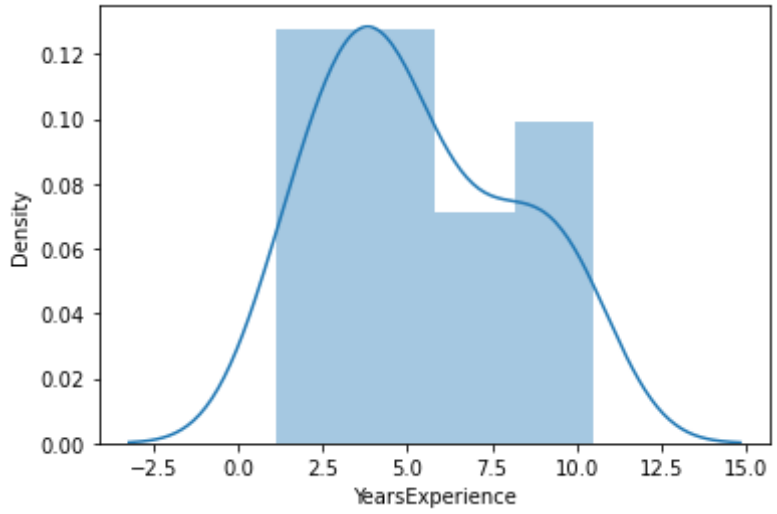
Out[6]:

	YearsExperience	Salary
YearsExperience	1.000000	0.978242
Salary	0.978242	1.000000

In [7]:
```
#histogram
import seaborn as sn
sn.distplot(data['YearsExperience'])
```

```
C:\Users\DELL\anaconda3\lib\site-packages\seaborn\distributions.py:2551: FutureWarning: `distplot` is a deprecated fu
nction and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level functi
on with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

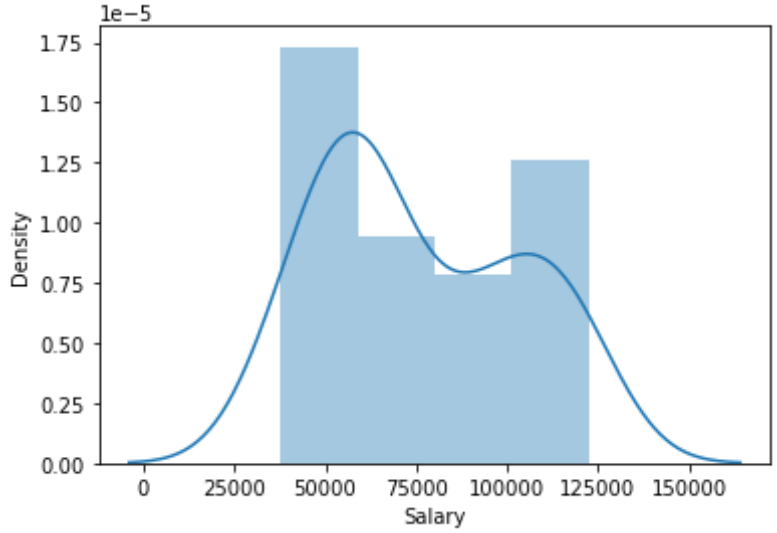
Out[7]:
```
<AxesSubplot:xlabel='YearsExperience', ylabel='Density'>
```



In [8]:
```
sn.distplot(data['Salary'])
```

```
C:\Users\DELL\anaconda3\lib\site-packages\seaborn\distributions.py:2551: FutureWarning: `distplot` is a deprecated fu
nction and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level functi
on with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Out[8]: <AxesSubplot:xlabel='Salary', ylabel='Density'>

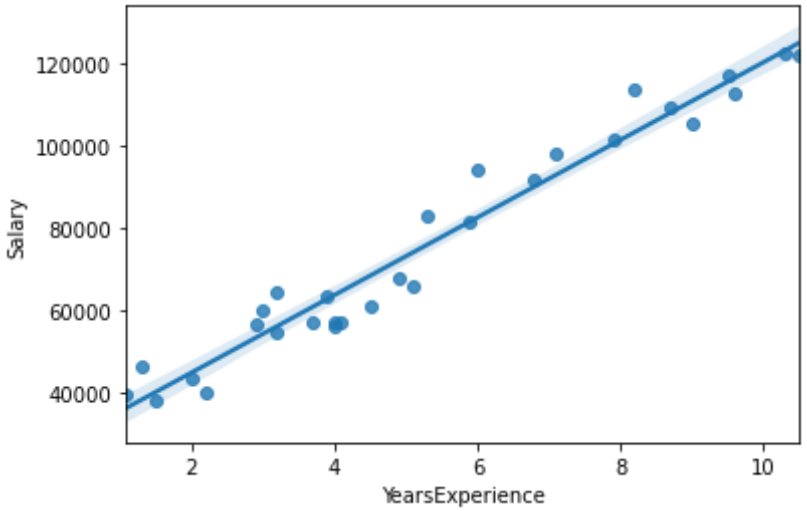


In [ ]:
```
#prediction for new model
```

In [9]:
```
#fitting yearsexperience and salary together
import statsmodels.formula.api as smf
model=smf.ols('YearsExperience~Salary',data=data).fit()
```

In [10]:
```
sn.regplot(x="YearsExperience",y="Salary",data=data)
```

Out[10]: <AxesSubplot:xlabel='YearsExperience', ylabel='Salary'>



In [12]:
```
#coefficients
model.params
```

Out[12]:
```
Intercept   -2.383161
Salary       0.000101
dtype: float64
```

In [13]:
```
#t and p values
print(model.tvalues,'\n',model.pvalues)
```

```
Intercept    -7.281283
Salary       24.950094
dtype: float64
 Intercept    6.300123e-08
Salary       1.143068e-20
dtype: float64
```

In [14]:
```
#r squared values
print(model.rsquared,model.rsquared_adj)
```

```
0.9569566641435086 0.9554194021486339
```

In [20]:
```
model.summary()
```

Out[20]:

OLS Regression Results

Dep. Variable:	YearsExperience	**R-squared:**	0.957
Model:	OLS	**Adj. R-squared:**	0.955
Method:	Least Squares	**F-statistic:**	622.5
Date:	Fri, 16 Apr 2021	**Prob (F-statistic):**	1.14e-20
Time:	01:04:34	**Log-Likelihood:**	-26.168
No. Observations:	30	**AIC:**	56.34
Df Residuals:	28	**BIC:**	59.14
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	-2.3832	0.327	-7.281	0.000	-3.054	-1.713
Salary	0.0001	4.06e-06	24.950	0.000	9.3e-05	0.000

Omnibus:	3.544	**Durbin-Watson:**	1.587
Prob(Omnibus):	0.170	**Jarque-Bera (JB):**	2.094
Skew:	-0.412	**Prob(JB):**	0.351
Kurtosis:	2.003	**Cond. No.**	2.41e+05

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 2.41e+05. This might indicate that there are
strong multicollinearity or other numerical problems.

In [ ]: