[1]:
```
import pandas as pd
```

[2]:
```
sal=pd.read_csv(r"C:\Users\kamaleshv\OneDrive\Desktop\Salary_Data.csv")
```

[3]:
```
sal
```

[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

[4]:
```
sal.shape
```

[4]:
```
(30, 2)
```

[5]:
```
sal.dtypes
```

[5]:
```
YearsExperience    float64
Salary             float64
dtype: object
```

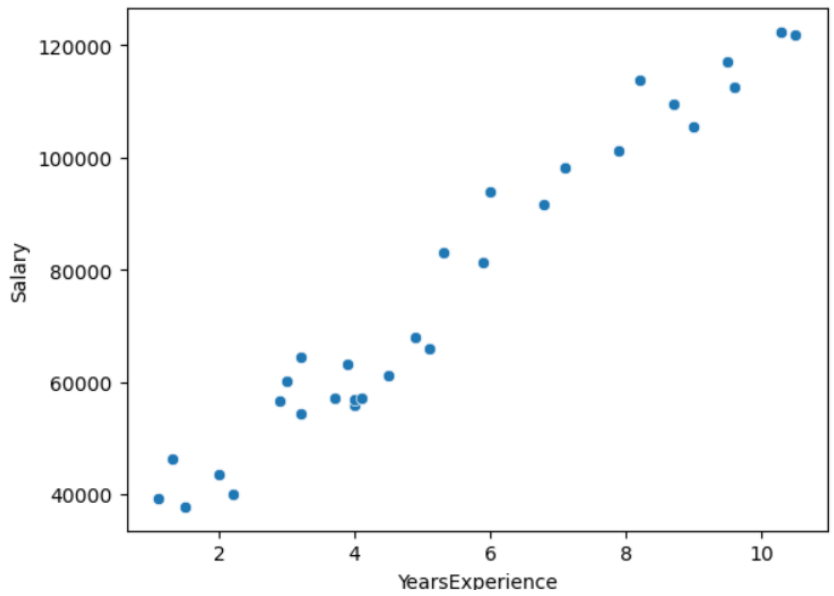
[7]:
```
sal.isna().sum()
```

[7]:
```
YearsExperience    0
Salary             0
dtype: int64
```

```
import seaborn as sns
from matplotlib import pyplot as plt
```

```
sns.scatterplot(x=sal['YearsExperience'],y=sal['Salary'])
```

```
<Axes: xlabel='YearsExperience', ylabel='Salary'>
```




```
x=sal[['YearsExperience']]#for feature we use data model so we used [[]] for data
x
```

	YearsExperience
0	1.1
1	1.3
2	1.5
3	2.0
4	2.2
5	2.9
6	3.0
7	3.2
8	3.2

```
y=sal['Salary']
y
```

```
0      39343.0
1      46205.0
2      37731.0
3      43525.0
4      39891.0
5      56642.0
6      60150.0
7      54445.0
8      64445.0
9      57189.0
10     63218.0
11     55794.0
12     56957.0
```

```
from sklearn.linear_model import LinearRegression
```

```
slr = LinearRegression()#model building
```

```
slr.fit(x,y)#model training
```

```
LinearRegression
LinearRegression()
```

```
y_pred = slr.predict(x) #model testing
```

```
y_pred.round(2)
```

```
array([ 36187.16,  38077.15,  39967.14,  44692.12,  46582.12,  53197.09,
        54142.09,  56032.08,  56032.08,  60757.06,  62647.05,  63592.05,
        63592.05,  64537.05,  68317.03,  72097.02,  73987.01,  75877.  ,
        81546.98,  82491.97,  90051.94,  92886.93, 100446.9 , 103281.89,
       108006.87, 110841.86, 115566.84, 116511.84, 123126.81, 125016.8 ])
```

```
from sklearn.metrics import r2_score,mean_squared_error,mean_absolute_error,mean_absolute_percentage_error,root_mean_squared_error
```

```
r2_score(y,y_pred)
```

```
0.9569566641435086
```

```
mean_squared_error(y,y_pred)
```

```
31270951.722280968
```

```
mean_absolute_error(y,y_pred)
```

```
4644.2012894435375
```

```
mean_absolute_percentage_error(y,y_pred)
```

```
0.07048034398306607
```

```
root_mean_squared_error(y,y_pred)
```

```
5592.043608760662
```