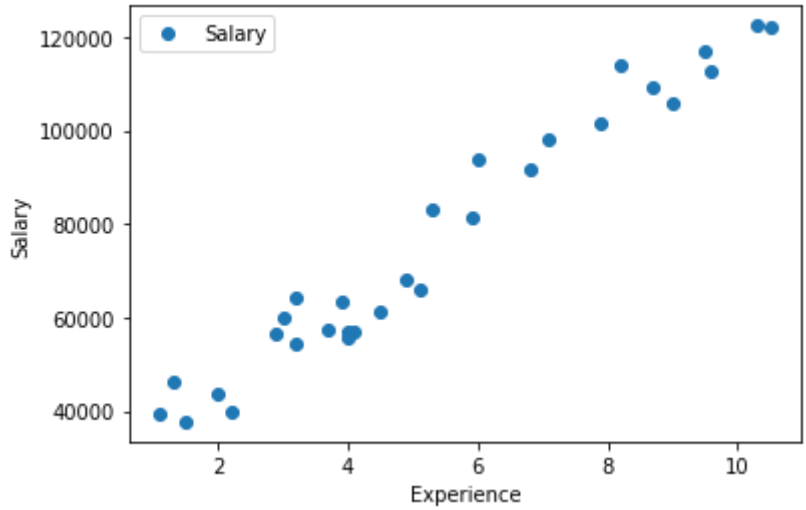
```
import pandas as pd
data = pd.read_csv("/content/Salary_Data.csv")
data.head()
```

	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

```
import matplotlib.pyplot as plt

data.plot(x='YearsExperience',y='Salary',style='o')
plt.xlabel('Experience')
plt.ylabel('Salary')
plt.show()
```




```
data.shape
```

```
(30, 2)
```

```
data.isnull().sum()
```

```
YearsExperience    0
Salary             0
dtype: int64
```

```
data.describe()
```

	YearsExperience	Salary
count	30.000000	30.000000
mean	5.313333	76003.000000
std	2.837888	27414.429785
min	1.100000	37731.000000
25%	3.200000	56720.750000
50%	4.700000	65237.000000
75%	7.700000	100544.750000
max	10.500000	122391.000000

```
data.columns
```

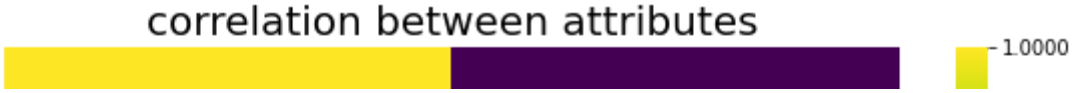
```
Index(['YearsExperience', 'Salary'], dtype='object')
```

```
corr=data.corr()
corr
```

	YearsExperience	Salary
YearsExperience	1.000000	0.978242
Salary	0.978242	1.000000

```
import seaborn as sns
plt.figure(figsize=(10,6))
plt.title("Correlation between Attributes",size=20)
sns.heatmap(corr,annot=True,cmap='viridis')
```

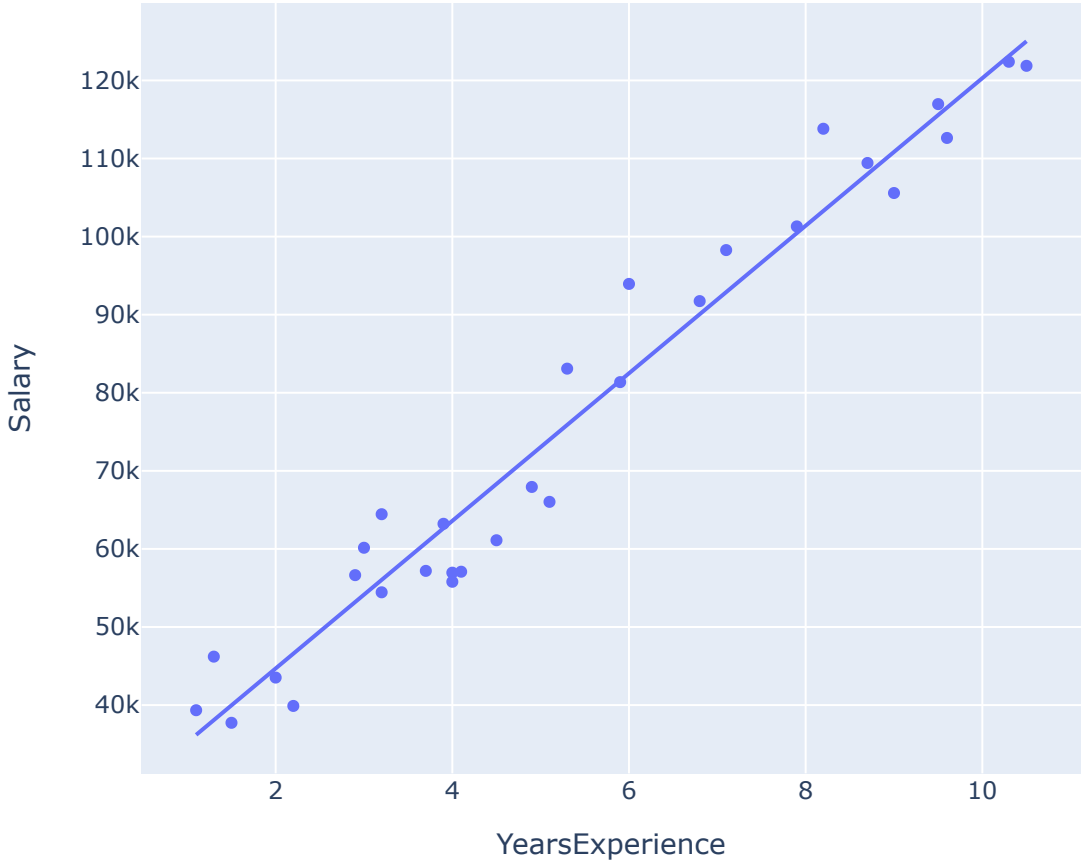
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f401d096550>
```




```
data
```

	YearsExperience	Salary
0	1.1	39343.0

```
import plotly.express as px
fig=px.scatter(data,x=data.YearsExperience,y=data.Salary,trendline='ols')
fig.show()
```



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

✓ 2s completed at 10:52 AM