In [ ]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
```

In [ ]:
```
df= pd.read_csv('ml_data_salary.csv')
```

In [ ]:
```
X = df[["YearsExperience"]]
y = df['Salary']
```

In [ ]:
```
X.head()
```

Out[ ]:

	YearsExperience
0	1.1
1	1.3
2	1.5
3	2.0
4	2.2

In [ ]:
```
y.head()
```

Out[ ]:
```
0    39343
1    46205
2    37731
3    43525
4    39891
Name: Salary, dtype: int64
```

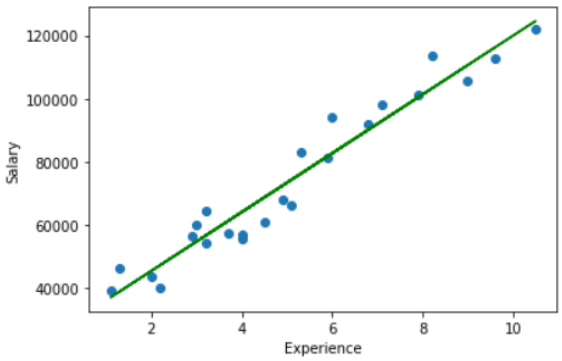
In [ ]:
```
x_train, x_test, y_train, y_test = train_test_split(X,y, test_size= 0.2, random_state=0)
```

In [ ]:
```
model = LinearRegression()
```

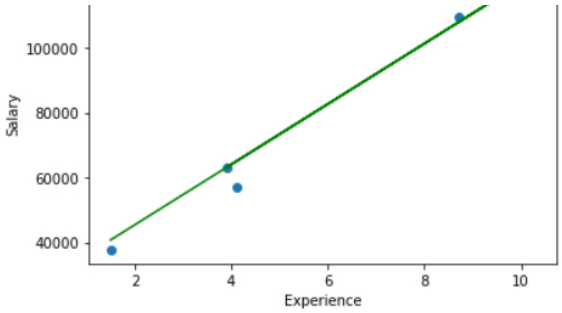
In [ ]:
```
model = model.fit(x_train, y_train)
model
```

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

In [ ]:
```
plt.scatter(x_train, y_train)
plt.plot(x_train, model.predict(x_train), color= 'Green')
plt.xlabel('Experience')
plt.ylabel("Salary")
plt.show()
```



In [ ]:
```
plt.scatter(x_test, y_test)
plt.plot(x_test, model.predict(x_test), color= 'Green')
plt.xlabel('Experience')
plt.ylabel("Salary")
plt.show()
```



In [ ]:
```
model.score(x_test, y_test)
```

Out[ ]:
```
0.988169515729126
```

In [ ]:
```
model.score(x_train, y_train)
```

Out[ ]:
```
0.9411949620562126
```

In [ ]:
```
model.predict([[5]])
```

Out[ ]:
```
array([73342.97478427])
```

In [ ]:
```
model.predict([[5],[6],[7]])
```

Out[ ]:
```
array([73342.97478427, 82655.549911  , 91968.12503773])
```

In [ ]:
```
x= ([10],[20],[30],[5])
model.predict(x)
```

Out[ ]:
```
array([119905.85041792, 213031.60168521, 306157.3529525 ,  73342.97478427])
```