```
# Reading CSV file
import pandas as pd
data = pd.read_csv('salary_data.csv') # Giving the relative path of the file
# i.e. telling where my file is with respect to where i am


# Printing the contents of data
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

Next steps: Generate code with data | View recommended plots | New interactive sheet

```
# Printing information of data
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

```
# Getting X and y values
X = data.iloc[:, [0]].values # Getting the 0th position column i.e. YearsExperie
y = data.iloc[:, 1].values # Getting the 1th position column i.e. Salary


# Splitting the X and y to training set and testing set
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_


# Creating Model
```

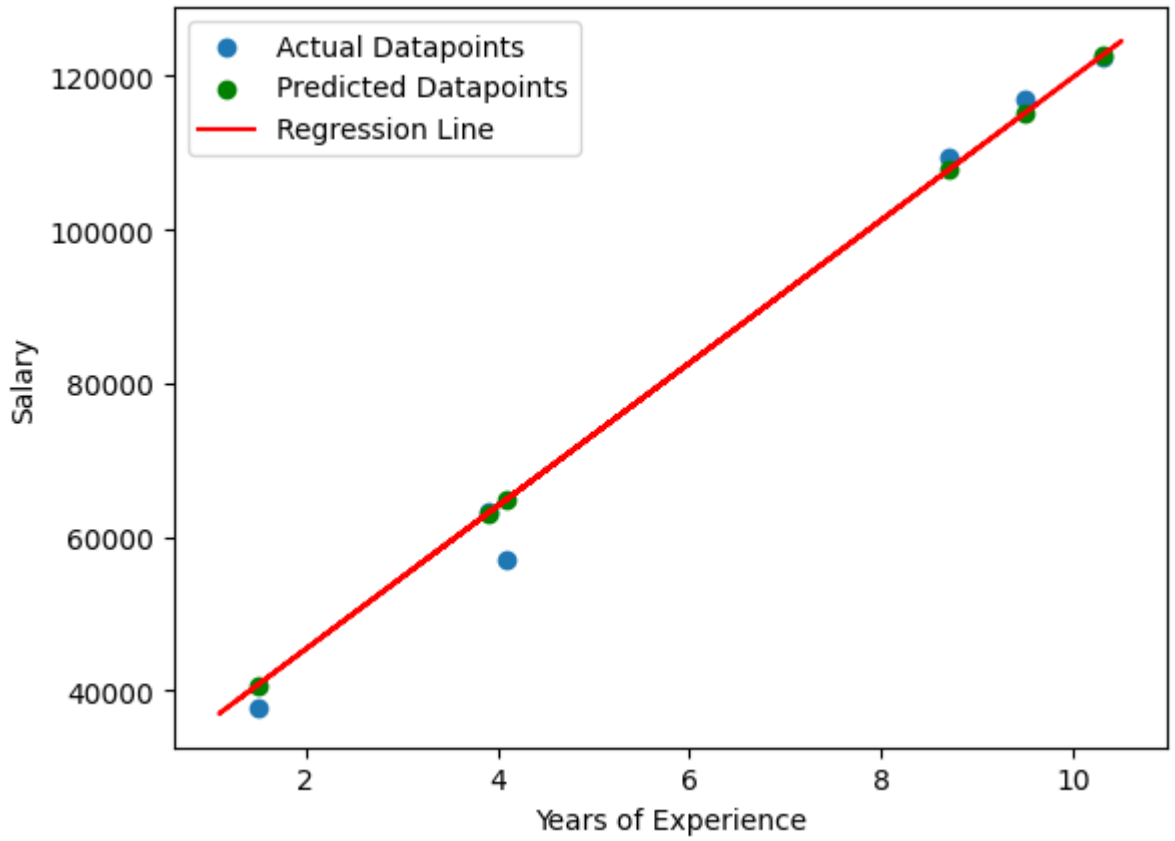
```
from sklearn.linear_model import LinearRegression
model = LinearRegression()

# Fitting/training the dataset
model.fit(X_train, y_train)
```

```
▾ LinearRegression
LinearRegression()
```

```
# Predicting the values using testing set
y_pred = model.predict(X_test)

# Plotting the Graph
import matplotlib.pyplot as plt
plt.scatter(X_test, y_test, label='Actual Datapoints')
plt.scatter(X_test, y_pred, label='Predicted Datapoints', c='green')
plt.plot(X_train, model.predict(X_train), label='Regression Line', color='red'
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.legend()
plt.show()
```




```
# Evaluating the Model
from sklearn.metrics import mean_squared_error, r2_score
mse = mean_squared_error(y_test, y_pred)
r2  = r2_score(y_test, y_pred)

print('Mean Squared Error is:', mse)
print('R^2 Score is:', r2)
```

```
print( R^2 Score is: , r2)
```

```
Mean Squared Error is: 12823412.298126549
R^2 Score is: 0.988169515729126
```