```
In [ ]:  import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import seaborn as sns
         %matplotlib inline
```

```
In [ ]:  df = pd.read_csv("Salary_Data.csv")
         df.head()
```

Out[ ]:

	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 [ ]:  X = df.iloc[:, :-1].values
         y = df.iloc[:, -1].values
         print(X)
         print(y)
```

```
[[ 1.1]
 [ 1.3]
 [ 1.5]
 [ 2. ]
 [ 2.2]
 [ 2.9]
 [ 3. ]
 [ 3.2]
 [ 3.2]
 [ 3.7]
 [ 3.9]
 [ 4. ]
 [ 4. ]
 [ 4.1]
 [ 4.5]
 [ 4.9]
 [ 5.1]
 [ 5.3]
 [ 5.9]
 [ 6. ]
 [ 6.8]
 [ 7.1]
 [ 7.9]
 [ 8.2]
 [ 8.7]
 [ 9. ]
 [ 9.5]
 [ 9.6]
 [10.3]
 [10.5]]
[ 39343.  46205.  37731.  43525.  39891.  56642.  60150.  54445.  64445.
  57189.  63218.  55794.  56957.  57081.  61111.  67938.  66029.  83088.
  81363.  93940.  91738.  98273. 101302. 113812. 109431. 105582. 116969.
 112635. 122391. 121872.]
```

Splitting the dataset into the Training set and Test set

```
In [ ]:  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_state = 0)
```

```
In [ ]:  print(X_train)
```

```
[[ 9.6]
 [ 4. ]
 [ 5.3]
 [ 7.9]
 [ 2.9]
 [ 5.1]
 [ 3.2]
 [ 4.5]
 [ 8.2]
 [ 6.8]
 [ 1.3]
 [10.5]
 [ 3. ]
 [ 2.2]
 [ 5.9]
 [ 6. ]
 [ 3.7]
 [ 3.2]
 [ 9. ]
 [ 2. ]
 [ 1.1]
 [ 7.1]
 [ 4.9]
 [ 4. ]]
```

```
In [ ]:  print(X_test)
```

```
[[ 1.5]
 [10.3]
 [ 4.1]
 [ 3.9]
 [ 9.5]
 [ 8.7]]
```

Training Simple Linear Regression Model on the Training Set

```
In [ ]:  from sklearn.linear_model import LinearRegression
         regressor = LinearRegression()
         regressor.fit(X_train, y_train)
```

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

Predict the Test Set Result

```
In [ ]:  y_pred = regressor.predict(X_test).round(2)
```

Visualising the Training Set Result

```
In [ ]:  plt.scatter(X_train, y_train, color = "red")
         plt.plot(X_train, regressor.predict(X_train).round(2), color = "blue")
         plt.title("Salary vs Experience (Training Set)")
         plt.xlabel("Years of Experience")
         plt.ylabel("Salary")
         plt.show()
```


Visualising the Test Set Result

```
In [ ]:  plt.scatter(X_test, y_test, color = "red")
         plt.plot(X_test, y_pred, color = "blue")
         plt.title("Salary vs Experience (Test Set)")
         plt.xlabel("Years of Experience")
         plt.ylabel("Salary")
         plt.show()
```

Making a single prediction (for example the salary of an employee with 12 years of experience)

```python
predict = regressor.predict([[12]])
print(predict)
```

```
[138531.00067138]
```

Getting the final linear regression equation with the values of the coefficients

```python
# b0 (y-intersept (constant))
print(regressor.intercept_)
```

```
26780.099150628186
```

```python
# b1 (slope coefficient)
print(regressor.coef_)
```

```
[9312.57512673]
```

Salary = 26816.19 + 9345.94 × YearsExperience