Predict Salary using simple Linear Regression

```
import pandas as pd
import numpy as np
```

Step 1: Load the Dataset

```
dataset = pd.read_csv('Salary_Data.csv')
X = dataset.iloc[:, :-1].values
y = dataset.iloc[:,1].values
```

Training and Testing set

Step 2: Split data into training and testing

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=1/3, random_state=0)
```

Fit Simple Linear Regression model to training set

Step 3: Fit Simple Linear Regression to Training Data

```
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train, y_train)
```

```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

Predict the test Set

Step 3: Make prediction

```
y_pred = regressor.predict(X_test)
```

Visualizing the training set

Step 5 - Visualize training set results

```
import matplotlib.pyplot as plt
# plot the actual data points of training set
plt.scatter(X_train, y_train, color = 'red')
# plot the regression line
plt.plot(X_train, regressor.predict(X_train), color='blue')
plt.title('Salary vs Experience (Training set)')
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.show()
```

Visualizing the test set

Step 6 - Visualize test set results

```
import matplotlib.pyplot as plt
# plot the actual data points of test set
plt.scatter(X_test, y_test, color = 'red')
# plot the regression line (same as above)
plt.plot(X_train, regressor.predict(X_train), color='blue')
plt.title('Salary vs Experience (Test set)')
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.show()
```

Make new predictions

Step 7 - Make new prediction

```
new_salary_pred = regressor.predict([[15]])
print('The predicted salary of a person with 15 years experience is ',new_salary_pred)
```

```
The predicted salary of a person with 15 years experience is  [106904.54215178]
```