In [1]:
```
#importing the libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

In [8]:
```
#importing the dataset
dataset  = pd.read_csv("C://Users//pritespa//Machine Learning A-Z Template Fol
der//Part 2 - Regression//Section 4 - Simple Linear Regression//Salary_Data.cs
v")
print(dataset)
X = dataset.iloc[:,:-1].values
print(X)
y = dataset.iloc[:,1].values
print(y)
```

```
    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
5               2.9   56642.0
6               3.0   60150.0
7               3.2   54445.0
8               3.2   64445.0
9               3.7   57189.0
10              3.9   63218.0
11              4.0   55794.0
12              4.0   56957.0
13              4.1   57081.0
14              4.5   61111.0
15              4.9   67938.0
16              5.1   66029.0
17              5.3   83088.0
18              5.9   81363.0
19              6.0   93940.0
20              6.8   91738.0
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
[[  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.]
```

In [10]:
```
#splitting the dataset into Training and Test set
from sklearn.cross_validation import train_test_split
X_train, X_test, Y_train, Y_test = train_test_split(X,y,test_size=1/3,random_s
tate=0)
X_train, X_test, Y_train, Y_test
```

Out[10]:
```
(array([[  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. ]]), array([[  1.5],
        [ 10.3],
        [  4.1],
        [  3.9],
        [  9.5],
        [  8.7],
        [  9.6],
        [  4. ],
        [  5.3],
        [  7.9]]), array([  56642.,   66029.,   64445.,   61111.,  113812.,
91738.,
         46205.,  121872.,   60150.,   39891.,   81363.,   93940.,
         57189.,   54445.,  105582.,   43525.,   39343.,   98273.,
         67938.,   56957.]), array([  37731.,  122391.,   57081.,   63218.,
116969.,  109431.,
        112635.,   55794.,   83088.,  101302.]))
```

In [12]:
```
#fitting Simple Linear Regression to the Training Set
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train, Y_train)
```

Out[12]:
```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=1, normalize=False)
```

In [15]:
```
#predicting the Test Set results
y_pred = regressor.predict(X_test)
print(y_pred)
```

```
[  40835.10590871  123079.39940819   65134.55626083   63265.36777221
  115602.64545369  108125.8914992   116537.23969801   64199.96201652
   76349.68719258  100649.1375447 ]
```

In [17]:
```
#Visualising the Training Set results
plt.scatter(X_train, Y_train, color = 'red')
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()
```

In [18]:
```
#Visualising the Test Set results
plt.scatter(X_test, Y_test, color = 'red')
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()
```