In [1]:
```
import numpy as nd
import pandas as pd
import matplotlib.pyplot as pt
```

In [2]:
```
dataset = pd.read_csv("datasets_9401_13260_Salary_Data.csv")
dataset
```

Out[2]:

	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

In [3]:
```
X = dataset.iloc[:,:-1].values
X
```

Out[3]:
```
array([[ 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]])
```

In [4]:
```
Y =dataset.iloc[:,1].values
Y
```

Out[4]:
```
array([ 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 [5]:
```
from sklearn.model_selection import train_test_split
```

In [6]:
```
X_train,X_test,Y_train,Y_test =  train_test_split(X,Y,test_size=1/3)
```

In [7]:
```
from sklearn.linear_model import LinearRegression
```

In [8]:
```
Regressor = LinearRegression()
```

In [9]:
```
Regressor.fit(X_train,Y_train)
```

Out[9]:
```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

In [10]:
```
y_pred = Regressor.predict(X_test)
```

In [11]:
```
y_pred
```

Out[11]:
```
array([111558.78949179, 117327.47689289,  93291.27938832,  35604.40537735,
        63486.39448265,  75985.21718503,  90406.93568777,  52910.46758064,
        63486.39448265,  62524.94658247])
```

In [12]:
```
Y_test
```

Out[12]:
```
array([105582., 112635.,  98273.,  39343.,  56957.,  83088.,  91738.,
        56642.,  55794.,  63218.])
```

In [13]:
```
pt.scatter(X_train,Y_train,color='red')
pt.plot(X_train,Regressor.predict(X_train),color='blue')
pt.title('Experience Vs Salary(Training Set)')
pt.xlabel('Year of Experience')
pt.ylabel('Salary')
pt.show()
```

In [14]:
```
pt.scatter(X_test,Y_test,color='red')
pt.plot(X_test,Regressor.predict(X_test),color='blue')
pt.title('Experience Vs Salary(Testing Set)')
pt.xlabel('Year of Experience')
pt.ylabel('Salary')
pt.show()
```