In [28]:
```
#importing libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
```

In [29]:
```
#importing Dataset
df = pd.read_csv("Salary_Data.csv")
```

In [30]:
```
#View The Data
df.head()
```

Out[30]:

	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 [31]:
```
#View The Data Info
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30 entries, 0 to 29
Data columns (total 2 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   YearsExperience  30 non-null     float64
 1   Salary           30 non-null     float64
dtypes: float64(2)
memory usage: 608.0 bytes
```

In [32]:
```
#View The Shape of Data
df.shape
```

Out[32]:
```
(30, 2)
```

In [33]:
```
#Check if There is Any NULl Values in Data
df.isnull().sum()
```

Out[33]:
```
YearsExperience    0
Salary             0
dtype: int64
```

In [34]:
```
#Defining Features & Label of Data
X = df.iloc[:, :-1].values
y = df.iloc[:, 1].values
```

In [35]:
```
X
```

Out[35]:
```
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 [36]:
```
y
```

Out[36]:
```
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 [37]:
```
#Spliting Data into Train Test

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=0)
```

In [38]:
```
#Import Linear Regression
from sklearn.linear_model import LinearRegression

regressor = LinearRegression()
```

In [39]:
```
#Fit Data into Linear Regression
regressor.fit(X_train, y_train)
```

Out[39]:
```
LinearRegression

LinearRegression()
```

In [40]:
```
#Predicting The Test Set Results
y_pred = regressor.predict(X_test)
```

In [41]:
```
y_pred
```

Out[41]:
```
array([ 40817.78327049, 123188.08258899,  65154.46261459,  63282.41035735,
       115699.87356004, 108211.66453108, 116635.89968866,  64218.43648597,
        76386.77615802])
```

In [42]:
```
y_test
```

Out[42]:
```
array([ 37731., 122391.,  57081.,  63218., 116969., 109431., 112635.,
        55794.,  83088.])
```

In [43]:
```
#Visualization of 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 [44]:
```
#Visualization of Training 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()
```

In [ ]: