Import necessary library

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
import matplotlib.pyplot as plt
```

Data Exploration

```
data = pd.read_csv('Salary_data.csv')

data.head()
```

```
   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
```

Model Training

```
X = data.drop('Salary', axis=1)
y = data['Salary']

X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=4)

model = LinearRegression()
model.fit(X_train, y_train)

LinearRegression()
```

Model Evaluation

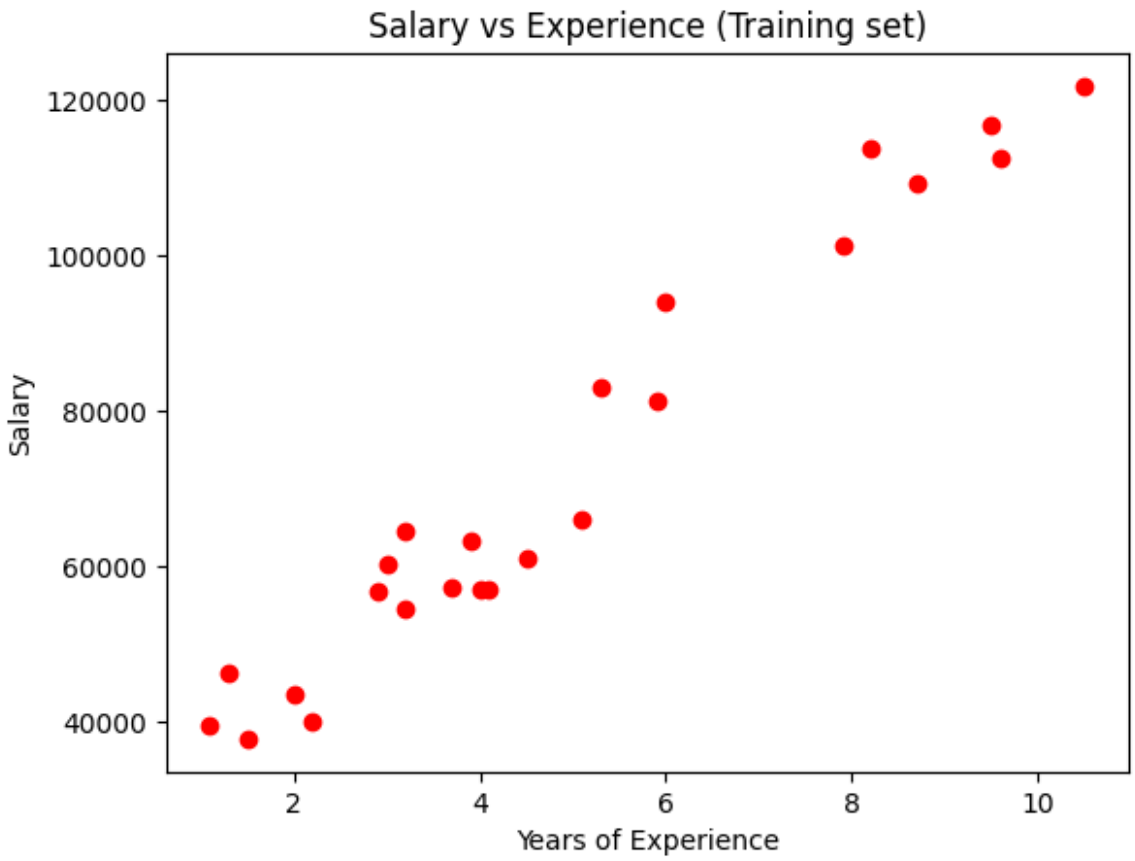
```
y_pred = model.predict(X_test)
result_dataframe = pd.DataFrame({'Actual': y_test, 'Predicted':
y_pred})

result_dataframe
```

```
      Actual      Predicted
11   55794.0   64027.973875
```

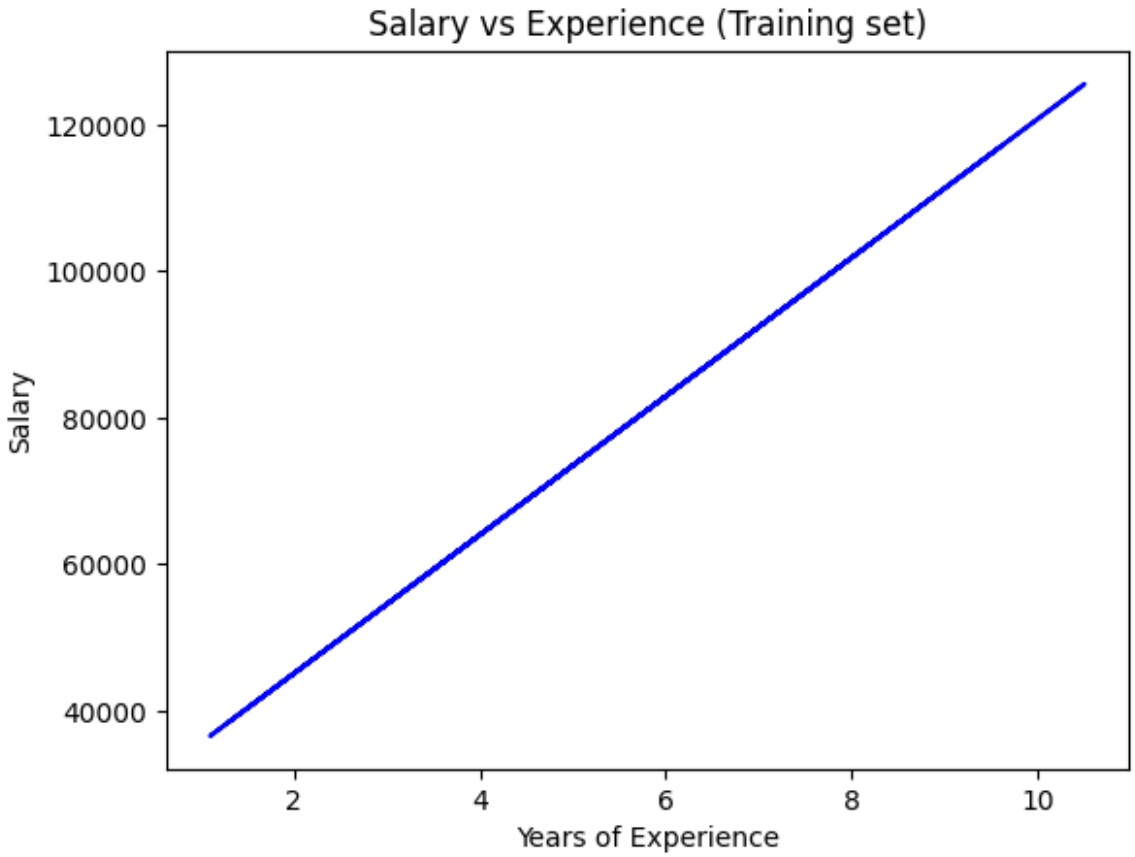
```
21   98273.0   93383.755269
28  122391.0  123686.497353
15   67938.0   72550.620086
20   91738.0   90542.873199
25  105582.0  111376.008381

accuracy = r2_score(y_test, y_pred)

accuracy

0.9504404484884267
```

Visualization

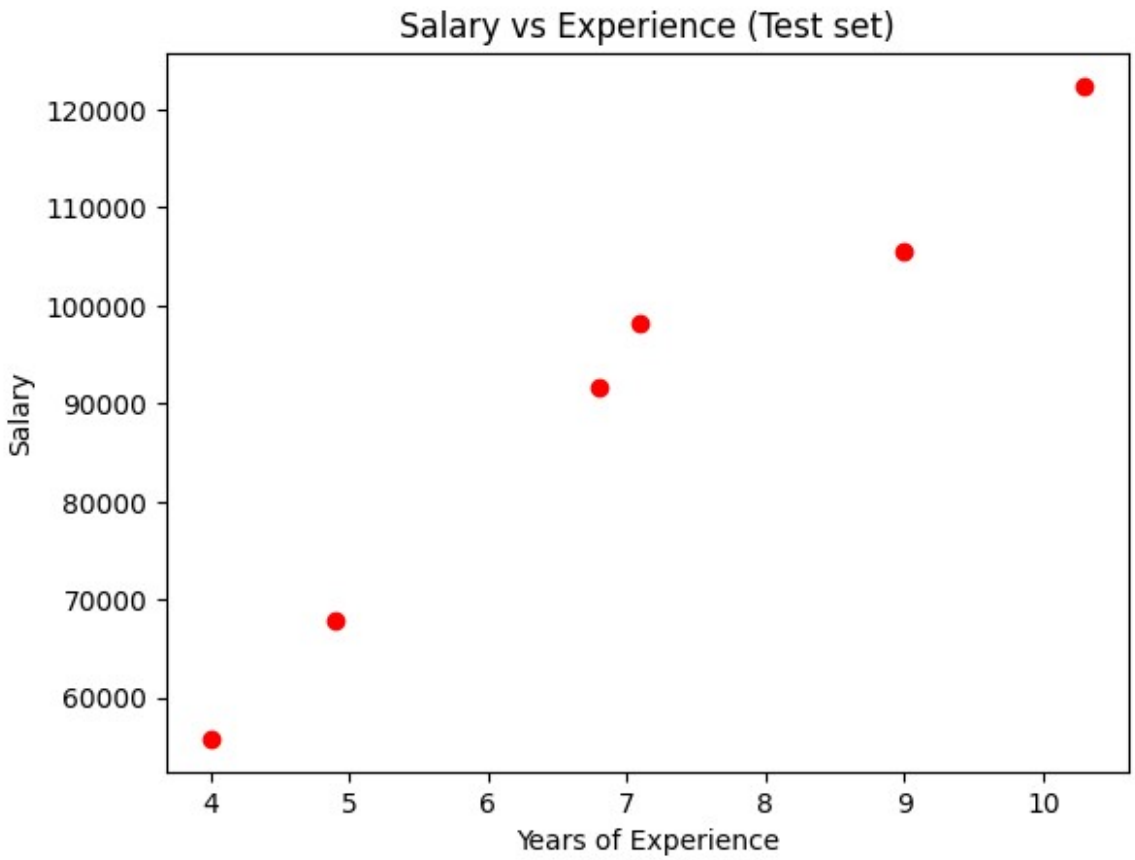
```
plt.scatter(X_train, y_train, color = 'red')
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.title('Salary vs Experience (Training set)')
plt.show()
```



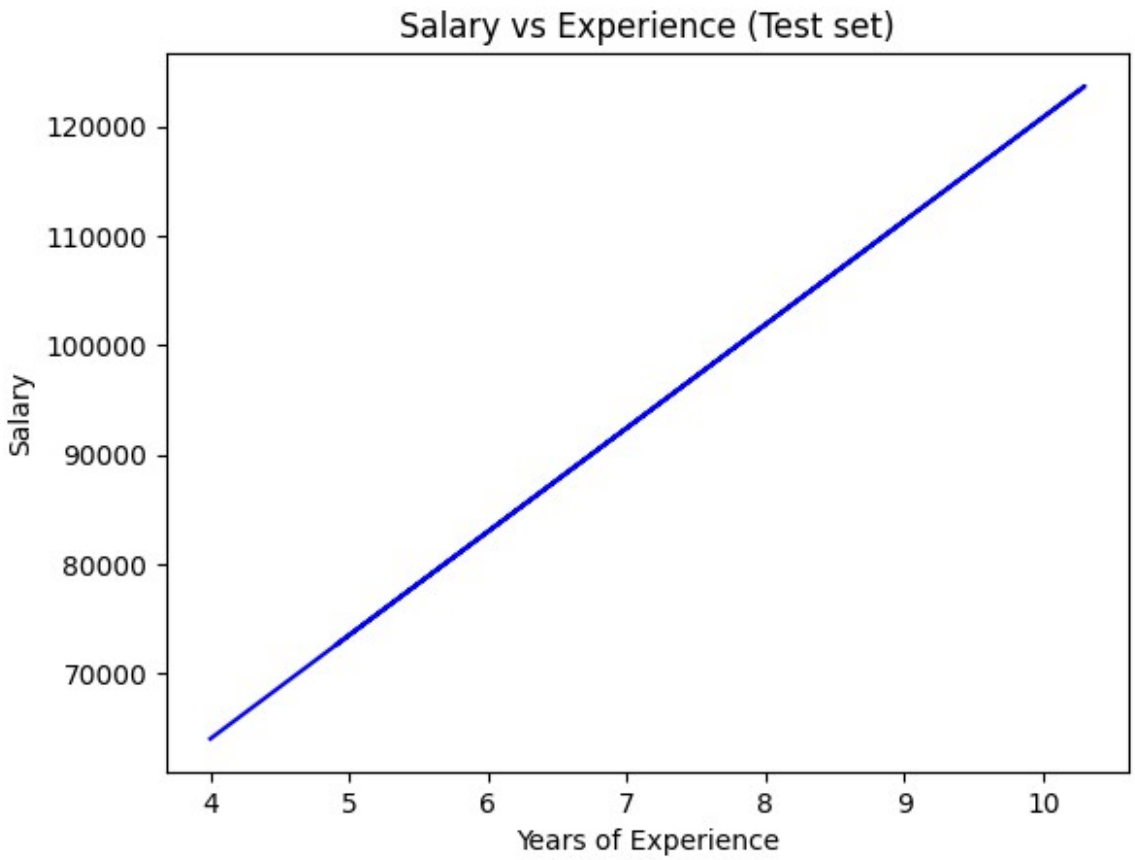
```
plt.plot(X_train, model.predict(X_train), color = 'blue')
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.title('Salary vs Experience (Training set)')
plt.show()
```




```
plt.scatter(X_test, y_test, color = 'red')
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.title('Salary vs Experience (Test set)')
plt.show()
```




```
plt.plot(X_test, model.predict(X_test), color = 'blue')
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.title('Salary vs Experience (Test set)')
plt.show()
```


Salary vs Experience (Test set)
120000
110000
100000
90000
80000
70000
Salary
4
5
6
7
8
9
10
Years of Experience