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

Dataset:

```python
data = {"YearsExperience":[1.1,1.3,1.5,2.0,2.2,2.9,3.0,3.2,3.2,3.7,3.9,4.0,4.5,4.9,5.1,5.3,5.9,6.0,6.8,7.1],"Salary
df = pd.DataFrame(data)
```

Prepare data

```python
X = df[['YearsExperience']]
y = df['Salary']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Train model

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

⇥ **Show hidden output**

Predict

```python
y_pred = model.predict(X_test)
```

Results:

```python
mse = mean_squared_error(y_test, y_pred)
print("Mean Squared Error:", mse)
```

⇥ Mean Squared Error: 34989569.24307344

```python
salary_5yrs = model.predict([[5]])
print("Predicted salary for 5 years experience:", salary_5yrs[0])
```

⇥ Predicted salary for 5 years experience: 68875.57933647823
   /usr/local/lib/python3.11/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feat
     warnings.warn(

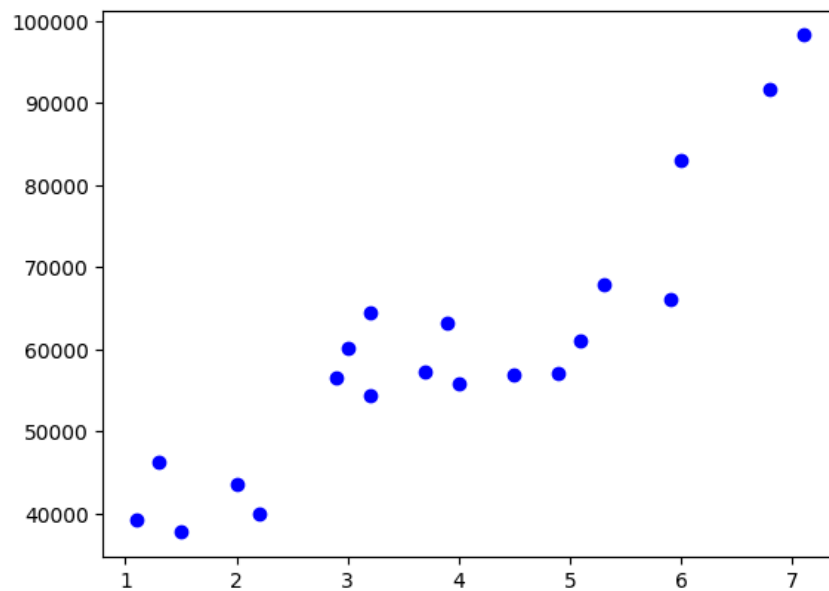Plot :

```python
plt.scatter(X, y, color='blue')
```
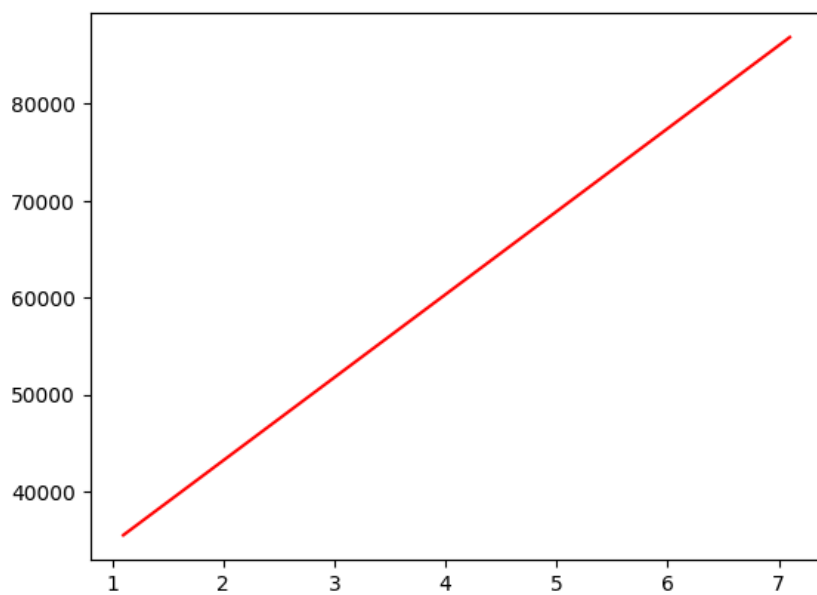
What can I help you build?

```
plt.plot(X, model.predict(X), color='red')
```

```
plt.title("Experience vs Salary")
```

Text(0.5, 1.0, 'Experience vs Salary')

Experience vs Salary


```
plt.xlabel("Years of Experience")
```

Text(0.5, 0, 'Years of Experience')

Years of Experience
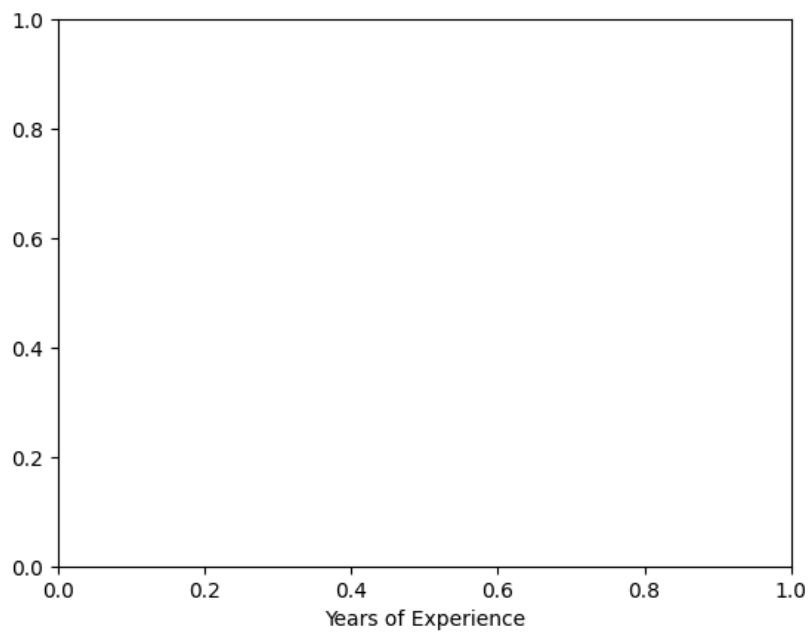
```
plt.scatter(X, y, color='blue')
plt.plot(X, model.predict(X), color='red')
plt.title("Experience vs Salary")
plt.xlabel("Years of Experience")
plt.ylabel("Salary")
plt.show()
```

Experience vs Salary