

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression

# Load and explore data
df_sal = pd.read_csv('Salary_Data.csv')
print(df_sal.head())
print(df_sal.describe())

# Salary distribution
plt.title('Salary Distribution Plot')
sns.histplot(df_sal['Salary'], kde=True)
plt.show()

# Scatter plot
plt.scatter(df_sal['YearsExperience'], df_sal['Salary'], color='lightcoral')
plt.title('YearsExperience vs Salary')
plt.xlabel('YearsExperience')
plt.ylabel('Salary')
plt.grid(True)
plt.show()

# Prepare data
x = df_sal.iloc[:, :1].values
y = df_sal.iloc[:, 1:].values

# Split data
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=0)

# Train model
regressor = LinearRegression()
regressor.fit(x_train, y_train)

# Predict
y_pred_test = regressor.predict(x_test)
y_pred_train = regressor.predict(x_train)

# Plot training results
plt.scatter(x_train, y_train, color='lightcoral')
plt.plot(x_train, y_pred_train, color='firebrick')
plt.title('YearsExperience vs Salary (Training Set)')
plt.xlabel('YearsExperience')
plt.ylabel('Salary')
plt.legend(['Regression Line', 'Training Data'], loc='best', facecolor='white')
plt.show()

# Output model coefficients
print(f'Coefficient: {regressor.coef_}')
print(f'Intercept: {regressor.intercept_}')

```