

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
```

```
In [2]: df = pd.read_csv('student_data.csv')
```

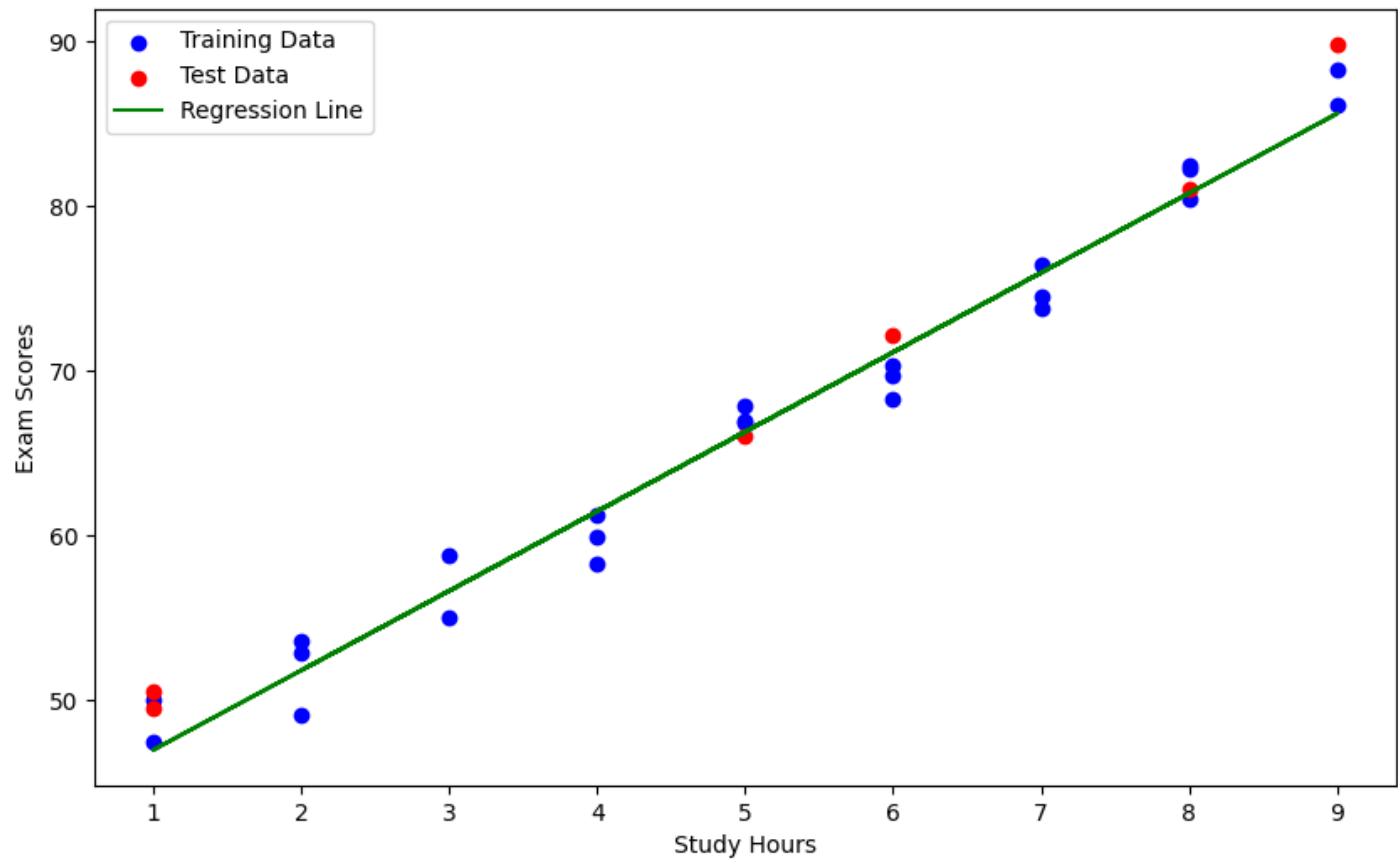
```
In [3]: X = df[['StudyHours']]
y = df['ExamScore']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
In [4]: model = LinearRegression()
model.fit(X_train, y_train)
```

```
Out[4]: ▾ LinearRegression ⓘ ?
LinearRegression()
```

```
In [5]: y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
```

```
In [6]: plt.figure(figsize=(10, 6))
plt.scatter(X_train['StudyHours'], y_train, label='Training Data', color='blue')
plt.scatter(X_test['StudyHours'], y_test, label='Test Data', color='red')
plt.plot(X_train, model.predict(X_train), color='green', label='Regression Line')
plt.xlabel('Study Hours')
plt.ylabel('Exam Scores')
plt.legend()
plt.show()
```



```
In [7]: print("Mean Squared Error:", mse)
        print("R-squared:", r2)
```

Mean Squared Error: 6.353087473532352
R-squared: 0.9710368865999279