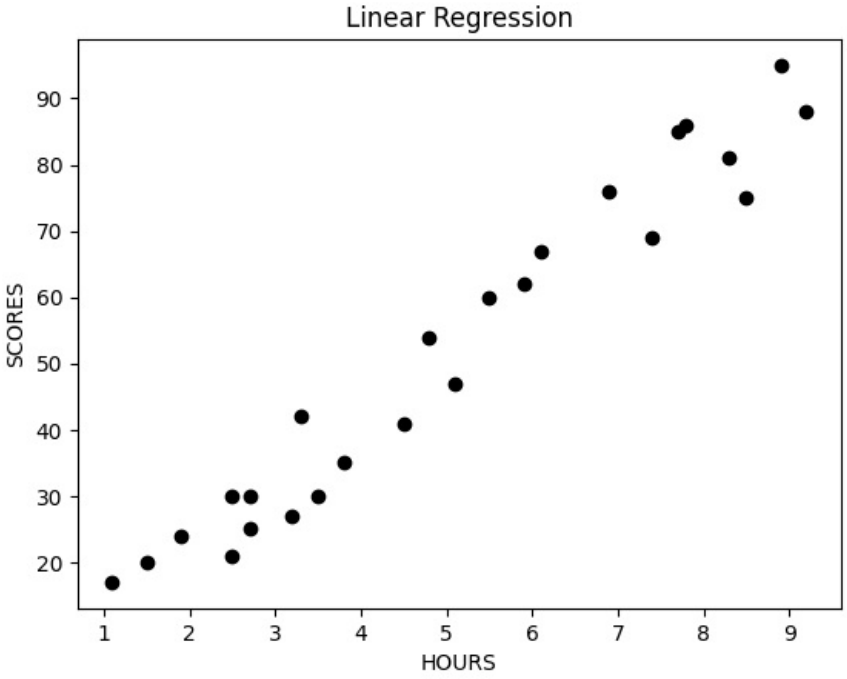
In [18]:
```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

df = pd.read_csv("STUDENTSCORE_DATASET.csv")

df.head()
x = df.iloc[:,:-1]
y = df.iloc[:,1]
plt.scatter(x,y, color='black')
plt.title('Linear Regression')
plt.xlabel('HOURS')
plt.ylabel('SCORES')
plt.show()

from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.33, random_state = 40)
model = LinearRegression()
model.fit(x_train, y_train)

y_pred = model.predict(x_test)
print(y_test, y_pred)

plt.plot(x_test, y_pred, color = 'blue', linewidth=2, label='Predicted Line')
plt.scatter(x,y, color='red')

from sklearn.metrics import mean_squared_error
mse = mean_squared_error(y_test, y_pred)
print(f"Mean Squared Error: {mse}")
```




```
15    95
24    86
14    17
19    69
13    42
3     75
21    54
2     27
11    62
Name: Scores, dtype: int64 [90.05323389 79.0394557  11.95553405 75.03444546 33.98309041 86.04822364
 49.00187884 32.98183785 60.01565703]
Mean Squared Error: 42.8684601892673
```