```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

```python
data=pd.read_csv('/content/student_scores.csv')
```

```python
data.head()
```

	Hours	Scores
0	2.5	21
1	5.1	47
2	3.2	27
3	8.5	75
4	3.5	30

```python
data.tail()
```

	Hours	Scores
20	2.7	30
21	4.8	54
22	3.8	35
23	6.9	76
24	7.8	86

```python
x=data.iloc[:,:-1].values
y=data.iloc[:,1].values
```

```python
print(x)
```

```
[[2.5]
 [5.1]
 [3.2]
 [8.5]
 [3.5]
 [1.5]
 [9.2]
 [5.5]
 [8.3]
 [2.7]
 [7.7]
 [5.9]
 [4.5]
 [3.3]
 [1.1]
 [8.9]
 [2.5]
 [1.9]
 [6.1]
 [7.4]
 [2.7]
 [4.8]
 [3.8]
 [6.9]
 [7.8]]
```

```python
print(y)
```

```
[21 47 27 75 30 20 88 60 81 25 85 62 41 42 17 95 30 24 67 69 30 54 35 76
 86]
```

```python
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
```

```python
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=1/3,random_state=0 )
```

```
regressor=LinearRegression()
regressor.fit(x_train,y_train)
```

    ▼ LinearRegression
    LinearRegression()

```
y_pred=regressor.predict(x_test)
print(y_pred)
```

    [17.04289179 33.51695377 74.21757747 26.73351648 59.68164043 39.33132858
     20.91914167 78.09382734 69.37226512]

```
print(y_test)
```

    [20 27 69 30 62 35 24 86 76]

```
#for train values
plt.scatter(x_train,y_train)
plt.plot(x_train,regressor.predict(x_train),color='black')
plt.title("Hours Vs Score(Training set)")
plt.xlabel("Hours")
plt.ylabel("Score")
plt.show()
```
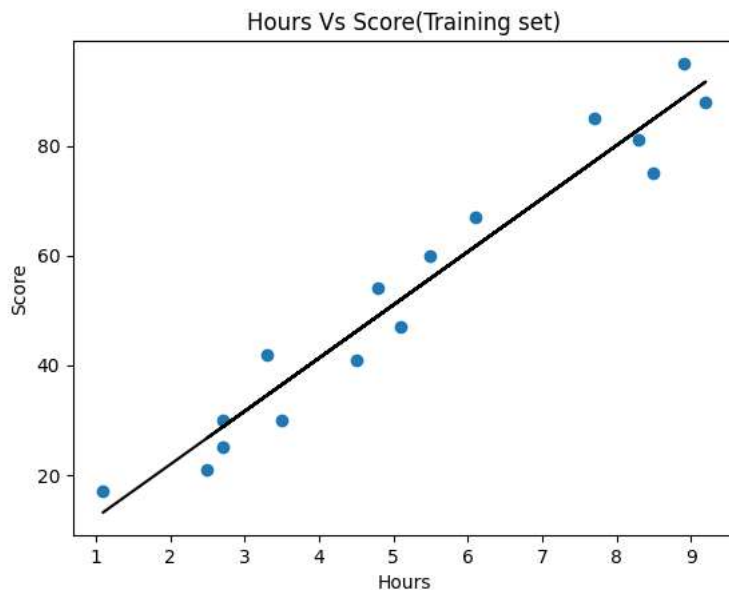


```
#for test values
y_pred=regressor.predict(x_test)
plt.scatter(x_test,y_test)
plt.plot(x_test,regressor.predict(x_test),color='black')
plt.title("Hours Vs Score(Test set)")
plt.xlabel("Hours")
plt.ylabel("Score")
plt.show()
```

Hours Vs Score(Test set)
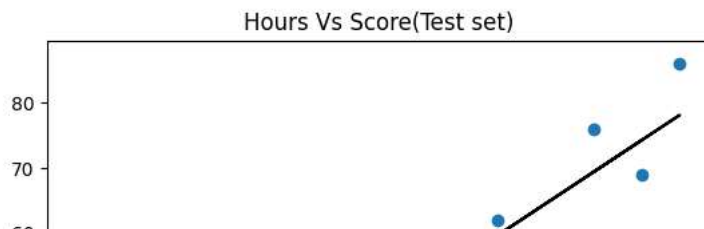

```
import sklearn.metrics as metrics

mae = metrics.mean_absolute_error(x, y)
mse = metrics.mean_squared_error(x, y)
rmse = np.sqrt(mse)

print("MAE:",mae)
print("MSE:", mse)
print("RMSE:", rmse)
```

```
    MAE: 46.468
    MSE: 2659.5692
    RMSE: 51.57101123693426
```

Hours