In [1]:
```
import pandas as pd
import numpy as np
```

In [2]:
```
dataset = pd.read_csv("student_scores.csv")
```

In [3]:
```
dataset.describe()
```

Out[3]:

	Hours	Scores
count	25.000000	25.000000
mean	5.012000	51.480000
std	2.525094	25.286887
min	1.100000	17.000000
25%	2.700000	30.000000
50%	4.800000	47.000000
75%	7.400000	75.000000
max	9.200000	95.000000

In [5]:
```
x = dataset.iloc[:, :-1].values
y = dataset.iloc[:, 1].values
```

In [6]:
```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(x, y, test_size = 0.2, random_state = 0)
```

In [7]:
```
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train, y_train)
```

Out[7]:
```
LinearRegression()
```

In [8]:
```
X_train.shape
```

Out[8]:
```
(20, 1)
```

In [9]:
```
X_test
```

Out[9]:
```
array([[1.5],
       [3.2],
       [7.4],
       [2.5],
       [5.9]])
```

In [11]:
```
y_pred = regressor.predict(X_test)
```

In [12]:
```
y_pred
```

Out[12]:
```
array([16.88414476, 33.73226078, 75.357018  , 26.79480124, 60.49103328])
```

In [13]:
```
from sklearn import metrics
print('Mean Absolute Error:', metrics.mean_absolute_error(y_test, y_pred))
print('Mean Squared Error:', metrics.mean_squared_error(y_test, y_pred))
print('Root Mean Squared Error:', np.sqrt(metrics.mean_absolute_error(y_test, y_pred)))
```

```
Mean Absolute Error: 4.183859899002975
Mean Squared Error: 21.5987693072174
Root Mean Squared Error: 2.0454485813637495
```

In [14]:
```
regressor.score(X_test,y_test)
```

Out[14]:
```
0.9454906892105356
```