In [1]:
```python
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

In [2]:
```python
s_data=pd.read_csv('studentscores.csv')
s_data
```

Out[2]:

	Hours	Scores
0	2.5	21
1	5.1	47
2	3.2	27
3	8.5	75
4	3.5	30
5	1.5	20
6	9.2	88
7	5.5	60
8	8.3	81
9	2.7	25
10	7.7	85
11	5.9	62
12	4.5	41
13	3.3	42
14	1.1	17
15	8.9	95
16	2.5	30
17	1.9	24
18	6.1	67
19	7.4	69
20	2.7	30
21	4.8	54
22	3.8	35
23	6.9	76
24	7.8	86

In [3]: `s_data.info`

Out[3]: ```
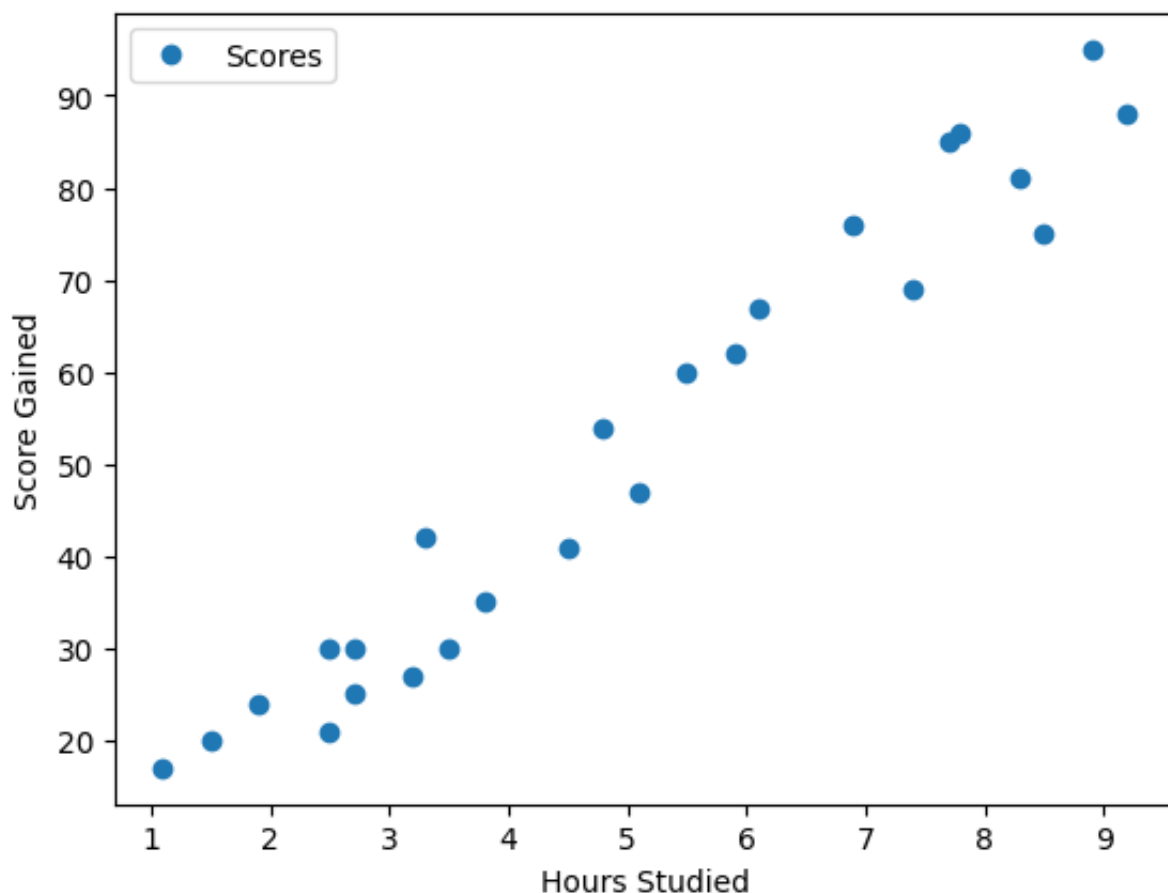<bound method DataFrame.info of   Hours Scores
0   2.5   21
1   5.1   47
2   3.2   27
3   8.5   75
4   3.5   30
5   1.5   20
6   9.2   88
7   5.5   60
8   8.3   81
9   2.7   25
10   7.7   85
11   5.9   62
12   4.5   41
13   3.3   42
14   1.1   17
15   8.9   95
16   2.5   30
17   1.9   24
18   6.1   67
19   7.4   69
20   2.7   30
21   4.8   54
22   3.8   35
23   6.9   76
24   7.8   86>
```

In [4]: `s_data.info`

```
Out[4]:  <bound method DataFrame.info of      Hours  Scores
         0     2.5     21
         1     5.1     47
         2     3.2     27
         3     8.5     75
         4     3.5     30
         5     1.5     20
         6     9.2     88
         7     5.5     60
         8     8.3     81
         9     2.7     25
         10    7.7     85
         11    5.9     62
         12    4.5     41
         13    3.3     42
         14    1.1     17
         15    8.9     95
         16    2.5     30
         17    1.9     24
         18    6.1     67
         19    7.4     69
         20    2.7     30
         21    4.8     54
         22    3.8     35
         23    6.9     76
         24    7.8     86>
```

```
In [5]:  s_data.plot(x='Hours',y='Scores',style='o')
         plt.xlabel('Hours Studied')
         plt.ylabel('Score Gained')
```

```
Out[5]:  Text(0, 0.5, 'Score Gained')
```

```
In [6]: x=s_data.iloc[:,:-1].values
        y=s_data.iloc[:,-1].values
```

```
In [7]: from sklearn.model_selection import train_test_split
```

```
In [8]: xtrain,xtest,ytrain,ytest=train_test_split(x,y,random_state=1,test_size=0
```

```
In [9]: from sklearn.linear_model import LinearRegression
```

```
In [10]: regressor=LinearRegression()
         regressor
```

Out[10]:
▼ **LinearRegression** ❶ ❷
LinearRegression()

```
In [11]: regressor.fit(xtest,ytest)
```

Out[11]:
▼ **LinearRegression** ❶ ❷
LinearRegression()

```
In [12]: print(regressor.coef_)
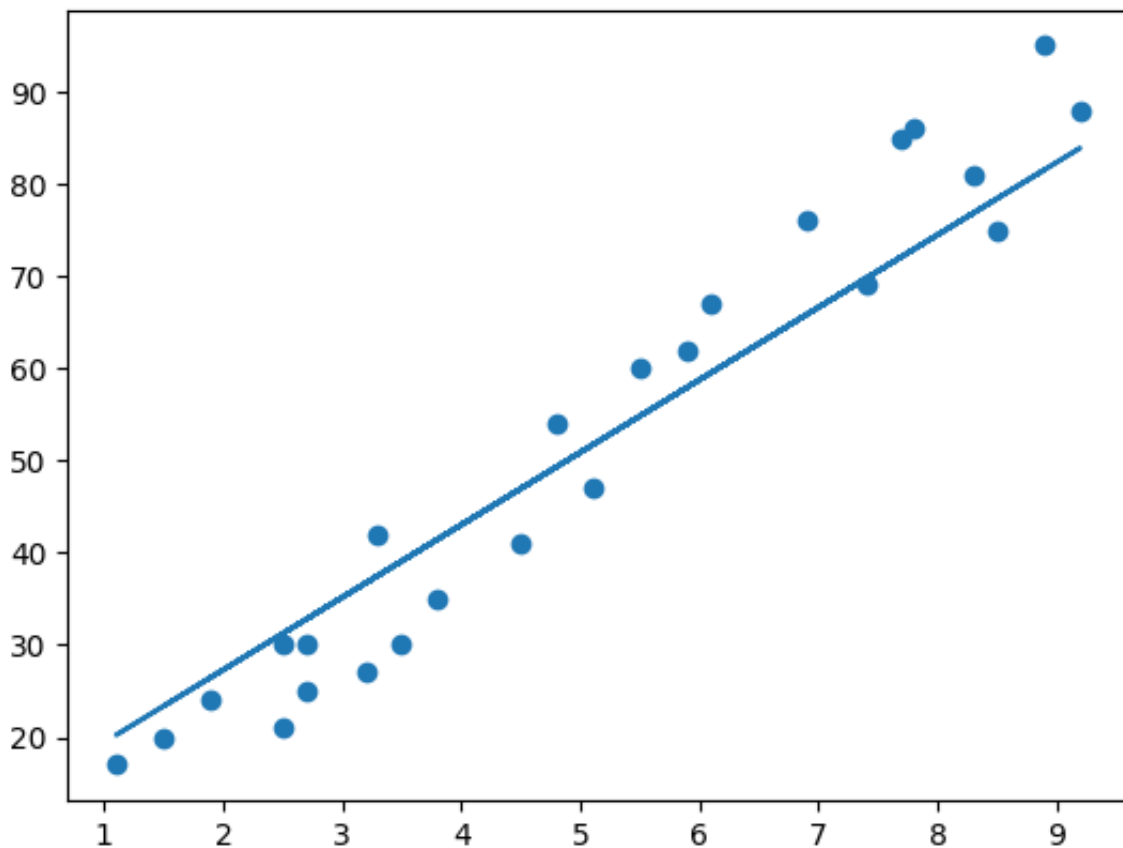```

```
                [7.85998595]
```

In [13]:
```
print(regressor.intercept_)
```

```
11.588855069070469
```

In [14]:
```
line=(regressor.coef_*x+regressor.intercept_)
plt.scatter(x,y)
plt.plot(x,line)
```

Out[14]:   [<matplotlib.lines.Line2D at 0x13aa73750>]

In [15]:
```
xtest
```

Out[15]:
```
array([[1.1],
       [3.3],
       [1.9],
       [8.5],
       [4.8]])
```

In [16]:
```
ypred=regressor.predict(xtest)
ypred
```

Out[16]:   `array([20.23483962, 37.52680871, 26.52282838, 78.39873566, 49.31678764])`

In [17]:
```
df=pd.DataFrame({'Actual' : ytest,'Predicted': ypred})
df
```

Out[17]:

	Actual	Predicted
0	17	20.234840
1	42	37.526809
2	24	26.522828
3	75	78.398736
4	54	49.316788

In [18]:
```python
from sklearn.metrics import mean_absolute_error
print(mean_absolute_error(ytest,ypred))
```

3.662561461016151

In [19]:
```python
from sklearn.metrics import mean_squared_error
print(mean_squared_error(ytest,ypred))
```

14.064434558651346

In [20]:
```python
from sklearn.metrics import r2_score
print(r2_score(ytest,ypred))
```

0.9677598694327633

In [ ]: