THE SPARKS FOUNDATION - GRIP November 2022

Data Science & Business Analytics

Name: Roshni Verma

Task 1: Prediction using Supervised ML

Dataset: https://bit.ly/w-data

Importing important libraries

In [3]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn import linear_model
```

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

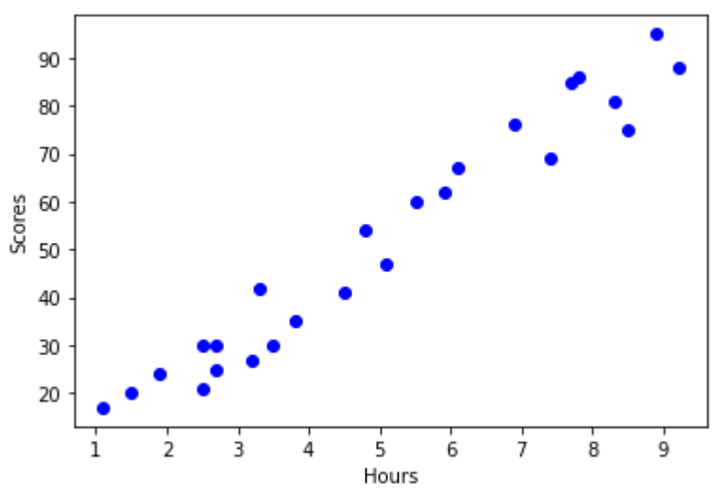
Out[2]:

	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

Visualization

In [6]:
```
%matplotlib inline
plt.xlabel('Hours')
plt.ylabel('Scores')
plt.scatter(df.Hours,df.Scores, color='blue')
```

Out[6]:
```
<matplotlib.collections.PathCollection at 0x12de9148>
```



Train the model

In [29]:
```
reg = linear_model.LinearRegression()
reg.fit(df[['Hours']],df.Scores)
```

Out[29]:
```
LinearRegression()
```

Prediction

In [31]:
```
reg.predict([[8]])
```

Out[31]:
```
array([80.69010053])
```