

In [2]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from pandas.plotting import scatter_matrix
import seaborn as sns
```

In [3]:

```
url="https://raw.githubusercontent.com/AdiPersonalWorks/Random/master/student_scores%20-%20%20"
df1=pd.read_csv(url)
print("Data imported successfully")
```

Data imported successfully

In [4]:

```
df1
```

Out[4]:

| | 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]:

```
print("shape:(rows,cols)")  
df1.shape
```

```
shape:(rows,cols)
```

Out[5]:

```
(25, 2)
```

In [6]:

```
df1.head(15)
```

Out[6]:

| | 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 |

In [7]:

```
df1.describe()
```

Out[7]:

| | 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 [8]:

```
df1.info()
```

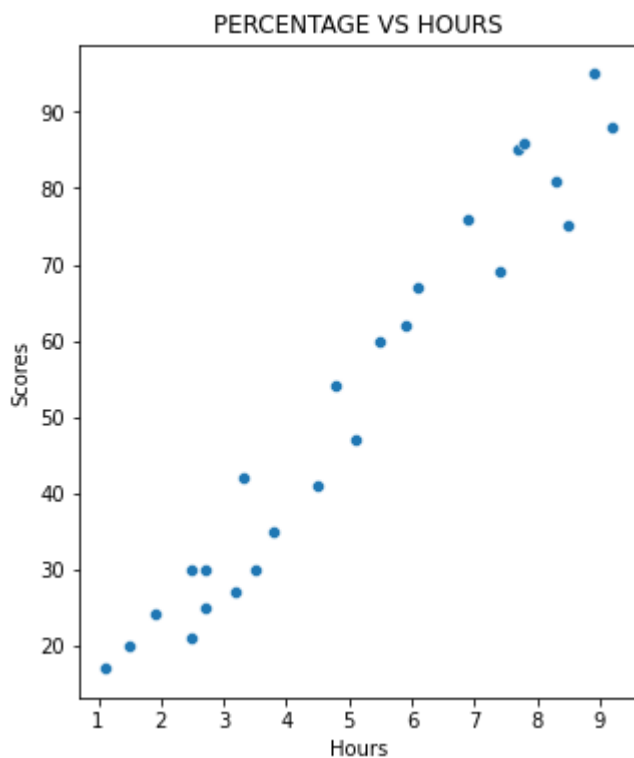
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 25 entries, 0 to 24
Data columns (total 2 columns):
#   Column  Non-Null Count  Dtype  
---  -
0   Hours    25 non-null      float64
1   Scores   25 non-null      int64   
dtypes: float64(1), int64(1)
memory usage: 528.0 bytes
```

In [9]:

```
plt.figure(figsize=(5,6))
sns.scatterplot(x="Hours",y="Scores",data=df1)
plt.title("PERCENTAGE VS HOURS")
```

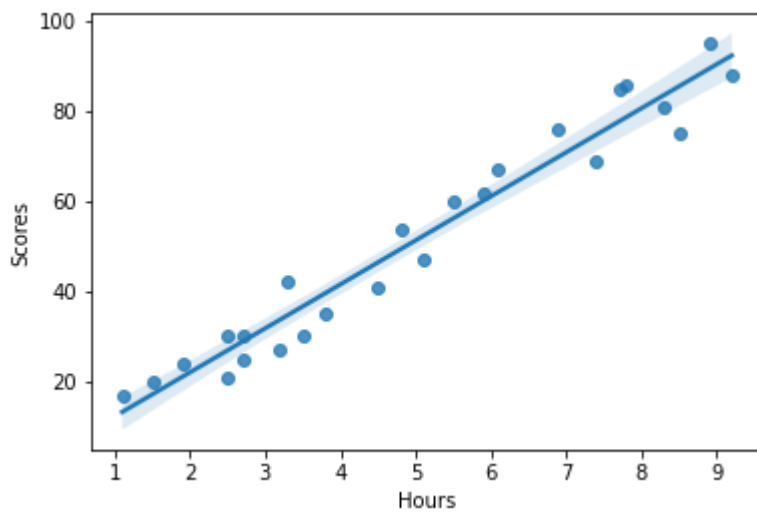
Out[9]:

```
Text(0.5, 1.0, 'PERCENTAGE VS HOURS')
```



In [10]:

```
ax = sns.regplot(x="Hours", y="Scores", data=df1)
```

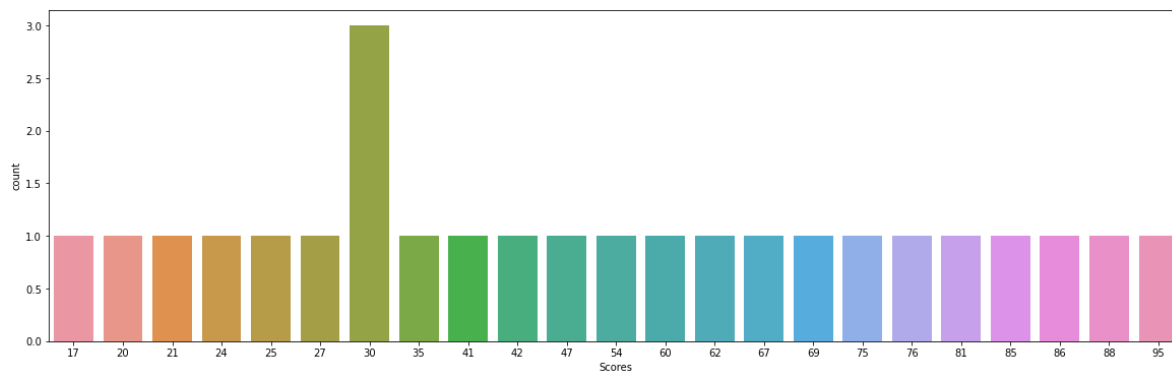


In [11]:

```
plt.figure(figsize=(20,6))  
sns.countplot(x=df1["Scores"])
```

Out[11]:

<AxesSubplot:xlabel='Scores', ylabel='count'>

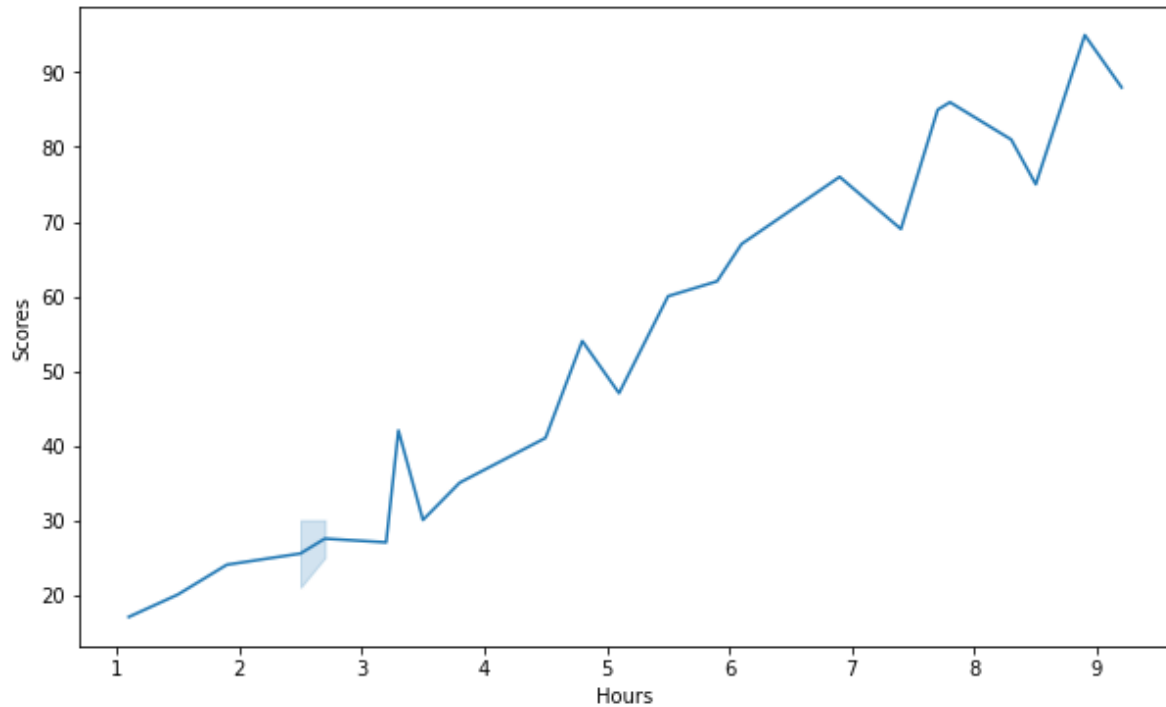


In [12]:

```
plt.figure(figsize=(10,6))  
sns.lineplot(x=df1["Hours"],y=df1["Scores"])
```

Out[12]:

<AxesSubplot:xlabel='Hours', ylabel='Scores'>



In [13]:

```
plt.figure(figsize=(15,6))
plt.subplot(1,2,1)
sns.boxplot(y=df1["Hours"], color="red")
plt.subplot(1,2,2)
sns.boxplot(y=df1["Scores"])
plt.show()
```

