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

In [2]:
```
import numpy as np
```

In [7]:
```
df=pd.read_csv("/home/ubuntu/Downloads/StudentsPerformanceTest1.csv")
```

In [8]:
```
df
```

Out[8]:

| | gender | math score | reading score | writing score | Placement Score | placement offer count | Region |
|---|---|---|---|---|---|---|---|
| **0** | female | 72 | 72 | 74.0 | 78.0 | 1 | Pune |
| **1** | female | 69 | 90 | 88.0 | NaN | 2 | na |
| **2** | female | 90 | 95 | 93.0 | 74.0 | 2 | Nashik |
| **3** | male | 47 | 57 | NaN | 78.0 | 1 | Na |
| **4** | male | na | 78 | 75.0 | 81.0 | 3 | Pune |
| **5** | female | 71 | Na | 78.0 | 70.0 | 4 | na |
| **6** | male | 12 | 44 | 52.0 | 12.0 | 2 | Nashik |
| **7** | male | NaN | 65 | 67.0 | 49.0 | 1 | Pune |
| **8** | male | 5 | 77 | 89.0 | 55.0 | 0 | NaN |

In [9]:
```
df.isnull
```

Out[9]:
```
<bound method DataFrame.isnull of    gender math score reading scor
e  writing score  Placement Score  \
0  female         72             72           74.0              78.0
1  female         69             90           88.0               NaN
2  female         90             95           93.0              74.0
3    male         47             57            NaN              78.0
4    male         na             78           75.0              81.0
5  female         71             Na           78.0              70.0
6    male         12             44           52.0              12.0
7    male        NaN             65           67.0              49.0
8    male          5             77           89.0              55.0

   placement offer count  Region
0                      1    Pune
1                      2      na
2                      2  Nashik
3                      1      Na
4                      3    Pune
5                      4      na
6                      2  Nashik
7                      1    Pune
8                      0     NaN  >
```

In [10]:
```
series=pd.isnull(df["math score"])
```

In [11]:
```
df[series]
```

Out[11]:

| | gender | math score | reading score | writing score | Placement Score | placement offer count | Region |
|---|---|---|---|---|---|---|---|
| **7** | male | NaN | 65 | 67.0 | 49.0 | 1 | Pune |

In [12]:
```
df.notnull()
```

Out[12]:

| | gender | math score | reading score | writing score | Placement Score | placement offer count | Region |
|---|---|---|---|---|---|---|---|
| **0** | True | True | True | True | True | True | True |
| **1** | True | True | True | True | False | True | True |
| **2** | True | True | True | True | True | True | True |
| **3** | True | True | True | False | True | True | True |
| **4** | True | True | True | True | True | True | True |
| **5** | True | True | True | True | True | True | True |
| **6** | True | True | True | True | True | True | True |
| **7** | True | False | True | True | True | True | True |
| **8** | True | True | True | True | True | True | False |

In [17]:
```
from sklearn import preprocessing
le = preprocessing.LabelEncoder()
```

In [18]:
```
df['gender']=le.fit_transform(df['gender'])
```

In [19]:
```
newdf=df
```

In [20]:
```
df
```

Out[20]:

| | gender | math score | reading score | writing score | Placement Score | placement offer count | Region |
|---|---|---|---|---|---|---|---|
| **0** | 0 | 72 | 72 | 74.0 | 78.0 | 1 | Pune |
| **1** | 0 | 69 | 90 | 88.0 | NaN | 2 | na |
| **2** | 0 | 90 | 95 | 93.0 | 74.0 | 2 | Nashik |
| **3** | 1 | 47 | 57 | NaN | 78.0 | 1 | Na |
| **4** | 1 | na | 78 | 75.0 | 81.0 | 3 | Pune |
| **5** | 0 | 71 | Na | 78.0 | 70.0 | 4 | na |
| **6** | 1 | 12 | 44 | 52.0 | 12.0 | 2 | Nashik |
| **7** | 1 | NaN | 65 | 67.0 | 49.0 | 1 | Pune |
| **8** | 1 | 5 | 77 | 89.0 | 55.0 | 0 | NaN |

In [ ]:
```

```