In [6]:
```
import pandas as pd
import numpy as np
```

In [7]:
```
df=pd.read_csv("StudentsPerformance.csv")
```

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

Out[8]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.0	81.0	93.0	85.0	23	98.0	99	female
1	83.0	79.0	64.0	97.0	23	66.0	97	female
2	100.0	96.0	70.0	86.0	23	84.0	69	female
3	88.0	83.0	100.0	81.0	23	79.0	79	female
4	73.0	72.0	96.0	62.0	23	64.0	64	female
5	63.0	76.0	NaN	100.0	23	83.0	74	female
6	72.0	87.0	99.0	94.0	23	72.0	68	female
7	90.0	75.0	91.0	66.0	23	88.0	86	female
8	97.0	79.0	74.0	77.0	23	90.0	91	female
9	NaN	60.0	77.0	70.0	23	81.0	69	Male
10	77.0	65.0	61.0	90.0	23	60.0	67	Male
11	75.0	94.0	88.0	91.0	23	90.0	71	Male
12	85.0	79.0	63.0	93.0	23	99.0	80	Male
13	80.0	79.0	100.0	NaN	23	76.0	95	Male
14	81.0	81.0	88.0	61.0	23	91.0	84	Male
15	80.0	64.0	98.0	100.0	23	69.0	96	Male
16	66.0	85.0	94.0	84.0	23	60.0	62	Male
17	99.0	80.0	75.0	93.0	23	72.0	76	Male
18	85.0	NaN	81.0	99.0	23	63.0	73	Male
19	76.0	99.0	89.0	84.0	23	97.0	87	Male
20	34.0	23.0	45.0	21.0	23	45.0	65	female
21	43.0	23.0	34.0	56.0	23	56.0	56	female
22	45.0	65.0	67.0	43.0	23	45.0	76	female
23	12.0	24.0	NaN	10.0	23	56.0	87	female
24	23.0	21.0	43.0	54.0	23	54.0	23	female
25	11.0	12.0	13.0	11.0	23	NaN	16	Male
26	13.0	14.0	54.0	44.0	23	77.0	65	Male
27	43.0	54.0	65.0	76.0	23	43.0	43	Male
28	43.0	44.0	54.0	65.0	23	11.0	12	Male

In [5]:
```
df.isnull()
```

Out[5]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count
0	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False
5	False	False	True	False	False	False	False
6	False	False	False	False	False	False	False
7	False	False	False	False	False	False	False
8	False	False	False	False	False	False	False
9	True	False	False	False	False	False	False
10	False	False	False	False	False	False	False
11	False	False	False	False	False	False	False
12	False	False	False	False	False	False	False
13	False	False	False	True	False	False	False
14	False	False	False	False	False	False	False
15	False	False	False	False	False	False	False
16	False	False	False	False	False	False	False
17	False	False	False	False	False	False	False
18	False	True	False	False	False	False	False
19	False	False	False	False	False	False	False
20	False	False	False	False	False	False	False
21	False	False	False	False	False	False	False
22	False	False	False	False	False	False	False
23	False	False	True	False	False	False	False
24	False	False	False	False	False	False	False
25	False	False	False	False	False	True	False
26	False	False	False	False	False	False	False
27	False	False	False	False	False	False	False
28	False	False	False	False	False	False	False

In [16]:
```
series = pd.isnull(df["Math Score"])
df[series]
```

Out[16]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count
9	NaN	60.0	77.0	70.0	23	81.0	69

In [17]:
```
df.notnull()
```

Out[17]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count
0	True	True	True	True	True	True	True
1	True	True	True	True	True	True	True
2	True	True	True	True	True	True	True
3	True	True	True	True	True	True	True
4	True	True	True	True	True	True	True
5	True	True	False	True	True	True	True
6	True	True	True	True	True	True	True
7	True	True	True	True	True	True	True
8	True	True	True	True	True	True	True
9	False	True	True	True	True	True	True
10	True	True	True	True	True	True	True
11	True	True	True	True	True	True	True
12	True	True	True	True	True	True	True
13	True	True	True	False	True	True	True
14	True	True	True	True	True	True	True
15	True	True	True	True	True	True	True
16	True	True	True	True	True	True	True
17	True	True	True	True	True	True	True
18	True	False	True	True	True	True	True
19	True	True	True	True	True	True	True
20	True	True	True	True	True	True	True
21	True	True	True	True	True	True	True
22	True	True	True	True	True	True	True
23	True	True	False	True	True	True	True
24	True	True	True	True	True	True	True
25	True	True	True	True	True	False	True
26	True	True	True	True	True	True	True
27	True	True	True	True	True	True	True
28	True	True	True	True	True	True	True

In [18]:
```
series1 = pd.notnull(df["Math Score"])
df[series1]
```

Out[18]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count
0	95.0	81.0	93.0	85.0	23	98.0	99
1	83.0	79.0	64.0	97.0	23	66.0	97
2	100.0	96.0	70.0	86.0	23	84.0	69
3	88.0	83.0	100.0	81.0	23	79.0	79
4	73.0	72.0	96.0	62.0	23	64.0	64
5	63.0	76.0	NaN	100.0	23	83.0	74
6	72.0	87.0	99.0	94.0	23	72.0	68
7	90.0	75.0	91.0	66.0	23	88.0	86
8	97.0	79.0	74.0	77.0	23	90.0	91
10	77.0	65.0	61.0	90.0	23	60.0	67
11	75.0	94.0	88.0	91.0	23	90.0	71
12	85.0	79.0	63.0	93.0	23	99.0	80
13	80.0	79.0	100.0	NaN	23	76.0	95
14	81.0	81.0	88.0	61.0	23	91.0	84
15	80.0	64.0	98.0	100.0	23	69.0	96
16	66.0	85.0	94.0	84.0	23	60.0	62
17	99.0	80.0	75.0	93.0	23	72.0	76
18	85.0	NaN	81.0	99.0	23	63.0	73
19	76.0	99.0	89.0	84.0	23	97.0	87
20	34.0	23.0	45.0	21.0	23	45.0	65
21	43.0	23.0	34.0	56.0	23	56.0	56
22	45.0	65.0	67.0	43.0	23	45.0	76
23	12.0	24.0	NaN	10.0	23	56.0	87
24	23.0	21.0	43.0	54.0	23	54.0	23
25	11.0	12.0	13.0	11.0	23	NaN	16
26	13.0	14.0	54.0	44.0	23	77.0	65
27	43.0	54.0	65.0	76.0	23	43.0	43
28	43.0	44.0	54.0	65.0	23	11.0	12

In [4]:
```
import pandas as pd
import numpy as np
```

In [5]:
```
df=pd.read_csv("StudentsPerformance.csv")
```

In [6]:
```
df
```

Out[6]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.0	81.0	93.0	85.0	23	98.0	99	female
1	83.0	79.0	64.0	97.0	23	66.0	97	female
2	100.0	96.0	70.0	86.0	23	84.0	69	female
3	88.0	83.0	100.0	81.0	23	79.0	79	female
4	73.0	72.0	96.0	62.0	23	64.0	64	female
5	63.0	76.0	NaN	100.0	23	83.0	74	female
6	72.0	87.0	99.0	94.0	23	72.0	68	female
7	90.0	75.0	91.0	66.0	23	88.0	86	female
8	97.0	79.0	74.0	77.0	23	90.0	91	female
9	NaN	60.0	77.0	70.0	23	81.0	69	Male
10	77.0	65.0	61.0	90.0	23	60.0	67	Male
11	75.0	94.0	88.0	91.0	23	90.0	71	Male
12	85.0	79.0	63.0	93.0	23	99.0	80	Male
13	80.0	79.0	100.0	NaN	23	76.0	95	Male
14	81.0	81.0	88.0	61.0	23	91.0	84	Male
15	80.0	64.0	98.0	100.0	23	69.0	96	Male
16	66.0	85.0	94.0	84.0	23	60.0	62	Male
17	99.0	80.0	75.0	93.0	23	72.0	76	Male
18	85.0	NaN	81.0	99.0	23	63.0	73	Male
19	76.0	99.0	89.0	84.0	23	97.0	87	Male
20	34.0	23.0	45.0	21.0	23	45.0	65	female
21	43.0	23.0	34.0	56.0	23	56.0	56	female
22	45.0	65.0	67.0	43.0	23	45.0	76	female
23	12.0	24.0	NaN	10.0	23	56.0	87	female
24	23.0	21.0	43.0	54.0	23	54.0	23	female
25	11.0	12.0	13.0	11.0	23	NaN	16	Male
26	13.0	14.0	54.0	44.0	23	77.0	65	Male
27	43.0	54.0	65.0	76.0	23	43.0	43	Male
28	43.0	44.0	54.0	65.0	23	11.0	12	Male

In [7]:
```
from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
df['gender'] = le.fit_transform(df['gender'])
newdf=df
df
```

Out[7]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.0	81.0	93.0	85.0	23	98.0	99	1
1	83.0	79.0	64.0	97.0	23	66.0	97	1
2	100.0	96.0	70.0	86.0	23	84.0	69	1
3	88.0	83.0	100.0	81.0	23	79.0	79	1
4	73.0	72.0	96.0	62.0	23	64.0	64	1
5	63.0	76.0	NaN	100.0	23	83.0	74	1
6	72.0	87.0	99.0	94.0	23	72.0	68	1
7	90.0	75.0	91.0	66.0	23	88.0	86	1
8	97.0	79.0	74.0	77.0	23	90.0	91	1
9	NaN	60.0	77.0	70.0	23	81.0	69	0
10	77.0	65.0	61.0	90.0	23	60.0	67	0
11	75.0	94.0	88.0	91.0	23	90.0	71	0
12	85.0	79.0	63.0	93.0	23	99.0	80	0
13	80.0	79.0	100.0	NaN	23	76.0	95	0
14	81.0	81.0	88.0	61.0	23	91.0	84	0
15	80.0	64.0	98.0	100.0	23	69.0	96	0
16	66.0	85.0	94.0	84.0	23	60.0	62	0
17	99.0	80.0	75.0	93.0	23	72.0	76	0
18	85.0	NaN	81.0	99.0	23	63.0	73	0
19	76.0	99.0	89.0	84.0	23	97.0	87	0
20	34.0	23.0	45.0	21.0	23	45.0	65	1
21	43.0	23.0	34.0	56.0	23	56.0	56	1
22	45.0	65.0	67.0	43.0	23	45.0	76	1
23	12.0	24.0	NaN	10.0	23	56.0	87	1
24	23.0	21.0	43.0	54.0	23	54.0	23	1
25	11.0	12.0	13.0	11.0	23	NaN	16	0
26	13.0	14.0	54.0	44.0	23	77.0	65	0
27	43.0	54.0	65.0	76.0	23	43.0	43	0
28	43.0	44.0	54.0	65.0	23	11.0	12	0

In [9]:
```
missing_values = ["Na", "na"]
df = pd.read_csv("StudentsPerformance.csv", na_values =
missing_values)
df
```

Out[9]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.0	81.0	93.0	85.0	23	98.0	99	female
1	83.0	79.0	64.0	97.0	23	66.0	97	female
2	100.0	96.0	70.0	86.0	23	84.0	69	female
3	88.0	83.0	100.0	81.0	23	79.0	79	female
4	73.0	72.0	96.0	62.0	23	64.0	64	female
5	63.0	76.0	NaN	100.0	23	83.0	74	female
6	72.0	87.0	99.0	94.0	23	72.0	68	female
7	90.0	75.0	91.0	66.0	23	88.0	86	female
8	97.0	79.0	74.0	77.0	23	90.0	91	female
9	NaN	60.0	77.0	70.0	23	81.0	69	Male
10	77.0	65.0	61.0	90.0	23	60.0	67	Male
11	75.0	94.0	88.0	91.0	23	90.0	71	Male
12	85.0	79.0	63.0	93.0	23	99.0	80	Male
13	80.0	79.0	100.0	NaN	23	76.0	95	Male
14	81.0	81.0	88.0	61.0	23	91.0	84	Male
15	80.0	64.0	98.0	100.0	23	69.0	96	Male
16	66.0	85.0	94.0	84.0	23	60.0	62	Male
17	99.0	80.0	75.0	93.0	23	72.0	76	Male
18	85.0	NaN	81.0	99.0	23	63.0	73	Male
19	76.0	99.0	89.0	84.0	23	97.0	87	Male
20	34.0	23.0	45.0	21.0	23	45.0	65	female
21	43.0	23.0	34.0	56.0	23	56.0	56	female
22	45.0	65.0	67.0	43.0	23	45.0	76	female
23	12.0	24.0	NaN	10.0	23	56.0	87	female
24	23.0	21.0	43.0	54.0	23	54.0	23	female
25	11.0	12.0	13.0	11.0	23	NaN	16	Male
26	13.0	14.0	54.0	44.0	23	77.0	65	Male
27	43.0	54.0	65.0	76.0	23	43.0	43	Male
28	43.0	44.0	54.0	65.0	23	11.0	12	Male

In [10]:
```
ndf=df
ndf.fillna(0)
```

Out[10]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.0	81.0	93.0	85.0	23	98.0	99	female
1	83.0	79.0	64.0	97.0	23	66.0	97	female
2	100.0	96.0	70.0	86.0	23	84.0	69	female
3	88.0	83.0	100.0	81.0	23	79.0	79	female
4	73.0	72.0	96.0	62.0	23	64.0	64	female
5	63.0	76.0	0.0	100.0	23	83.0	74	female
6	72.0	87.0	99.0	94.0	23	72.0	68	female
7	90.0	75.0	91.0	66.0	23	88.0	86	female
8	97.0	79.0	74.0	77.0	23	90.0	91	female
9	0.0	60.0	77.0	70.0	23	81.0	69	Male
10	77.0	65.0	61.0	90.0	23	60.0	67	Male
11	75.0	94.0	88.0	91.0	23	90.0	71	Male
12	85.0	79.0	63.0	93.0	23	99.0	80	Male
13	80.0	79.0	100.0	0.0	23	76.0	95	Male
14	81.0	81.0	88.0	61.0	23	91.0	84	Male
15	80.0	64.0	98.0	100.0	23	69.0	96	Male
16	66.0	85.0	94.0	84.0	23	60.0	62	Male
17	99.0	80.0	75.0	93.0	23	72.0	76	Male
18	85.0	0.0	81.0	99.0	23	63.0	73	Male
19	76.0	99.0	89.0	84.0	23	97.0	87	Male
20	34.0	23.0	45.0	21.0	23	45.0	65	female
21	43.0	23.0	34.0	56.0	23	56.0	56	female
22	45.0	65.0	67.0	43.0	23	45.0	76	female
23	12.0	24.0	0.0	10.0	23	56.0	87	female
24	23.0	21.0	43.0	54.0	23	54.0	23	female
25	11.0	12.0	13.0	11.0	23	0.0	16	Male
26	13.0	14.0	54.0	44.0	23	77.0	65	Male
27	43.0	54.0	65.0	76.0	23	43.0	43	Male
28	43.0	44.0	54.0	65.0	23	11.0	12	Male

In [12]:
```
m_v=df['Math Score'].mean()
df['Math Score'].fillna(value=m_v, inplace=True)
df
```

Out[12]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.000000	81.0	93.0	85.0	23	98.0	99	female
1	83.000000	79.0	64.0	97.0	23	66.0	97	female
2	100.000000	96.0	70.0	86.0	23	84.0	69	female
3	88.000000	83.0	100.0	81.0	23	79.0	79	female
4	73.000000	72.0	96.0	62.0	23	64.0	64	female
5	63.000000	76.0	NaN	100.0	23	83.0	74	female
6	72.000000	87.0	99.0	94.0	23	72.0	68	female
7	90.000000	75.0	91.0	66.0	23	88.0	86	female
8	97.000000	79.0	74.0	77.0	23	90.0	91	female
9	65.428571	60.0	77.0	70.0	23	81.0	69	Male
10	77.000000	65.0	61.0	90.0	23	60.0	67	Male
11	75.000000	94.0	88.0	91.0	23	90.0	71	Male
12	85.000000	79.0	63.0	93.0	23	99.0	80	Male
13	80.000000	79.0	100.0	NaN	23	76.0	95	Male
14	81.000000	81.0	88.0	61.0	23	91.0	84	Male
15	80.000000	64.0	98.0	100.0	23	69.0	96	Male
16	66.000000	85.0	94.0	84.0	23	60.0	62	Male
17	99.000000	80.0	75.0	93.0	23	72.0	76	Male
18	85.000000	NaN	81.0	99.0	23	63.0	73	Male
19	76.000000	99.0	89.0	84.0	23	97.0	87	Male
20	34.000000	23.0	45.0	21.0	23	45.0	65	female
21	43.000000	23.0	34.0	56.0	23	56.0	56	female
22	45.000000	65.0	67.0	43.0	23	45.0	76	female
23	12.000000	24.0	NaN	10.0	23	56.0	87	female
24	23.000000	21.0	43.0	54.0	23	54.0	23	female
25	11.000000	12.0	13.0	11.0	23	NaN	16	Male
26	13.000000	14.0	54.0	44.0	23	77.0	65	Male
27	43.000000	54.0	65.0	76.0	23	43.0	43	Male
28	43.000000	44.0	54.0	65.0	23	11.0	12	Male

In [13]:
```
ndf.replace(to_replace = np.nan, value = -99)
```

Out[13]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.000000	81.0	93.0	85.0	23	98.0	99	female
1	83.000000	79.0	64.0	97.0	23	66.0	97	female
2	100.000000	96.0	70.0	86.0	23	84.0	69	female
3	88.000000	83.0	100.0	81.0	23	79.0	79	female
4	73.000000	72.0	96.0	62.0	23	64.0	64	female
5	63.000000	76.0	-99.0	100.0	23	83.0	74	female
6	72.000000	87.0	99.0	94.0	23	72.0	68	female
7	90.000000	75.0	91.0	66.0	23	88.0	86	female
8	97.000000	79.0	74.0	77.0	23	90.0	91	female
9	65.428571	60.0	77.0	70.0	23	81.0	69	Male
10	77.000000	65.0	61.0	90.0	23	60.0	67	Male
11	75.000000	94.0	88.0	91.0	23	90.0	71	Male
12	85.000000	79.0	63.0	93.0	23	99.0	80	Male
13	80.000000	79.0	100.0	-99.0	23	76.0	95	Male
14	81.000000	81.0	88.0	61.0	23	91.0	84	Male
15	80.000000	64.0	98.0	100.0	23	69.0	96	Male
16	66.000000	85.0	94.0	84.0	23	60.0	62	Male
17	99.000000	80.0	75.0	93.0	23	72.0	76	Male
18	85.000000	-99.0	81.0	99.0	23	63.0	73	Male
19	76.000000	99.0	89.0	84.0	23	97.0	87	Male
20	34.000000	23.0	45.0	21.0	23	45.0	65	female
21	43.000000	23.0	34.0	56.0	23	56.0	56	female
22	45.000000	65.0	67.0	43.0	23	45.0	76	female
23	12.000000	24.0	-99.0	10.0	23	56.0	87	female
24	23.000000	21.0	43.0	54.0	23	54.0	23	female
25	11.000000	12.0	13.0	11.0	23	-99.0	16	Male
26	13.000000	14.0	54.0	44.0	23	77.0	65	Male
27	43.000000	54.0	65.0	76.0	23	43.0	43	Male
28	43.000000	44.0	54.0	65.0	23	11.0	12	Male

In [14]:
```
ndf.dropna()
```

Out[14]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.000000	81.0	93.0	85.0	23	98.0	99	female
1	83.000000	79.0	64.0	97.0	23	66.0	97	female
2	100.000000	96.0	70.0	86.0	23	84.0	69	female
3	88.000000	83.0	100.0	81.0	23	79.0	79	female
4	73.000000	72.0	96.0	62.0	23	64.0	64	female
6	72.000000	87.0	99.0	94.0	23	72.0	68	female
7	90.000000	75.0	91.0	66.0	23	88.0	86	female
8	97.000000	79.0	74.0	77.0	23	90.0	91	female
9	65.428571	60.0	77.0	70.0	23	81.0	69	Male
10	77.000000	65.0	61.0	90.0	23	60.0	67	Male
11	75.000000	94.0	88.0	91.0	23	90.0	71	Male
12	85.000000	79.0	63.0	93.0	23	99.0	80	Male
14	81.000000	81.0	88.0	61.0	23	91.0	84	Male
15	80.000000	64.0	98.0	100.0	23	69.0	96	Male
16	66.000000	85.0	94.0	84.0	23	60.0	62	Male
17	99.000000	80.0	75.0	93.0	23	72.0	76	Male
19	76.000000	99.0	89.0	84.0	23	97.0	87	Male
20	34.000000	23.0	45.0	21.0	23	45.0	65	female
21	43.000000	23.0	34.0	56.0	23	56.0	56	female
22	45.000000	65.0	67.0	43.0	23	45.0	76	female
24	23.000000	21.0	43.0	54.0	23	54.0	23	female
26	13.000000	14.0	54.0	44.0	23	77.0	65	Male
27	43.000000	54.0	65.0	76.0	23	43.0	43	Male
28	43.000000	44.0	54.0	65.0	23	11.0	12	Male

In [15]:
```
ndf.dropna(how = 'all')
```

Out[15]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.000000	81.0	93.0	85.0	23	98.0	99	female
1	83.000000	79.0	64.0	97.0	23	66.0	97	female
2	100.000000	96.0	70.0	86.0	23	84.0	69	female
3	88.000000	83.0	100.0	81.0	23	79.0	79	female
4	73.000000	72.0	96.0	62.0	23	64.0	64	female
5	63.000000	76.0	NaN	100.0	23	83.0	74	female
6	72.000000	87.0	99.0	94.0	23	72.0	68	female
7	90.000000	75.0	91.0	66.0	23	88.0	86	female
8	97.000000	79.0	74.0	77.0	23	90.0	91	female
9	65.428571	60.0	77.0	70.0	23	81.0	69	Male
10	77.000000	65.0	61.0	90.0	23	60.0	67	Male
11	75.000000	94.0	88.0	91.0	23	90.0	71	Male
12	85.000000	79.0	63.0	93.0	23	99.0	80	Male
13	80.000000	79.0	100.0	NaN	23	76.0	95	Male
14	81.000000	81.0	88.0	61.0	23	91.0	84	Male
15	80.000000	64.0	98.0	100.0	23	69.0	96	Male
16	66.000000	85.0	94.0	84.0	23	60.0	62	Male
17	99.000000	80.0	75.0	93.0	23	72.0	76	Male
18	85.000000	NaN	81.0	99.0	23	63.0	73	Male
19	76.000000	99.0	89.0	84.0	23	97.0	87	Male
20	34.000000	23.0	45.0	21.0	23	45.0	65	female
21	43.000000	23.0	34.0	56.0	23	56.0	56	female
22	45.000000	65.0	67.0	43.0	23	45.0	76	female
23	12.000000	24.0	NaN	10.0	23	56.0	87	female
24	23.000000	21.0	43.0	54.0	23	54.0	23	female
25	11.000000	12.0	13.0	11.0	23	NaN	16	Male
26	13.000000	14.0	54.0	44.0	23	77.0	65	Male
27	43.000000	54.0	65.0	76.0	23	43.0	43	Male
28	43.000000	44.0	54.0	65.0	23	11.0	12	Male

In [16]:
```
ndf.dropna(axis = 1)
```

Out[16]:

	Math Score	Club Join Year	Offer Count	gender
0	95.000000	23	99	female
1	83.000000	23	97	female
2	100.000000	23	69	female
3	88.000000	23	79	female
4	73.000000	23	64	female
5	63.000000	23	74	female
6	72.000000	23	68	female
7	90.000000	23	86	female
8	97.000000	23	91	female
9	65.428571	23	69	Male
10	77.000000	23	67	Male
11	75.000000	23	71	Male
12	85.000000	23	80	Male
13	80.000000	23	95	Male
14	81.000000	23	84	Male
15	80.000000	23	96	Male
16	66.000000	23	62	Male
17	99.000000	23	76	Male
18	85.000000	23	73	Male
19	76.000000	23	87	Male
20	34.000000	23	65	female
21	43.000000	23	56	female
22	45.000000	23	76	female
23	12.000000	23	87	female
24	23.000000	23	23	female
25	11.000000	23	16	Male
26	13.000000	23	65	Male
27	43.000000	23	43	Male
28	43.000000	23	12	Male

In [17]:
```
new_data = ndf.dropna(axis = 0, how ='any')
new_data
```

Out[17]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.000000	81.0	93.0	85.0	23	98.0	99	female
1	83.000000	79.0	64.0	97.0	23	66.0	97	female
2	100.000000	96.0	70.0	86.0	23	84.0	69	female
3	88.000000	83.0	100.0	81.0	23	79.0	79	female
4	73.000000	72.0	96.0	62.0	23	64.0	64	female
6	72.000000	87.0	99.0	94.0	23	72.0	68	female
7	90.000000	75.0	91.0	66.0	23	88.0	86	female
8	97.000000	79.0	74.0	77.0	23	90.0	91	female
9	65.428571	60.0	77.0	70.0	23	81.0	69	Male
10	77.000000	65.0	61.0	90.0	23	60.0	67	Male
11	75.000000	94.0	88.0	91.0	23	90.0	71	Male
12	85.000000	79.0	63.0	93.0	23	99.0	80	Male
14	81.000000	81.0	88.0	61.0	23	91.0	84	Male
15	80.000000	64.0	98.0	100.0	23	69.0	96	Male
16	66.000000	85.0	94.0	84.0	23	60.0	62	Male
17	99.000000	80.0	75.0	93.0	23	72.0	76	Male
19	76.000000	99.0	89.0	84.0	23	97.0	87	Male
20	34.000000	23.0	45.0	21.0	23	45.0	65	female
21	43.000000	23.0	34.0	56.0	23	56.0	56	female
22	45.000000	65.0	67.0	43.0	23	45.0	76	female
24	23.000000	21.0	43.0	54.0	23	54.0	23	female
26	13.000000	14.0	54.0	44.0	23	77.0	65	Male
27	43.000000	54.0	65.0	76.0	23	43.0	43	Male
28	43.000000	44.0	54.0	65.0	23	11.0	12	Male

In [25]:
```
col = ['Math Score', 'Reading Score', 'Writing Score','Placement Score']
df.boxplot(col)
```

Out[25]: <Axes: >

In [22]:
```
print(np.where(df['Math Score']>90))
print(np.where(df['Reading Score']<25))
print(np.where(df['Writing Score']<30))
```

```
(array([ 0,  2,  8, 17], dtype=int64),)
(array([20, 21, 23, 24, 25, 26], dtype=int64),)
(array([25], dtype=int64),)
```

In [26]:
```
df=pd.read_csv("StudentsPerformance.csv")
```

In [27]:
```
df
```

Out[27]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.0	81.0	93.0	85.0	23	98.0	99	female
1	83.0	79.0	64.0	97.0	23	66.0	97	female
2	100.0	96.0	70.0	86.0	23	84.0	69	female
3	88.0	83.0	100.0	81.0	23	79.0	79	female
4	73.0	72.0	96.0	62.0	23	64.0	64	female
5	63.0	76.0	NaN	100.0	23	83.0	74	female
6	72.0	87.0	99.0	94.0	23	72.0	68	female
7	90.0	75.0	91.0	66.0	23	88.0	86	female
8	97.0	79.0	74.0	77.0	23	90.0	91	female
9	NaN	60.0	77.0	70.0	23	81.0	69	Male
10	77.0	65.0	61.0	90.0	23	60.0	67	Male
11	75.0	94.0	88.0	91.0	23	90.0	71	Male
12	85.0	79.0	63.0	93.0	23	99.0	80	Male
13	80.0	79.0	100.0	NaN	23	76.0	95	Male
14	81.0	81.0	88.0	61.0	23	91.0	84	Male
15	80.0	64.0	98.0	100.0	23	69.0	96	Male
16	66.0	85.0	94.0	84.0	23	60.0	62	Male
17	99.0	80.0	75.0	93.0	23	72.0	76	Male
18	85.0	NaN	81.0	99.0	23	63.0	73	Male
19	76.0	99.0	89.0	84.0	23	97.0	87	Male
20	34.0	23.0	45.0	21.0	23	45.0	65	female
21	43.0	23.0	34.0	56.0	23	56.0	56	female
22	45.0	65.0	67.0	43.0	23	45.0	76	female
23	12.0	24.0	NaN	10.0	23	56.0	87	female
24	23.0	21.0	43.0	54.0	23	54.0	23	female
25	11.0	12.0	13.0	11.0	23	NaN	16	Male
26	13.0	14.0	54.0	44.0	23	77.0	65	Male
27	43.0	54.0	65.0	76.0	23	43.0	43	Male
28	43.0	44.0	54.0	65.0	23	11.0	12	Male

In [21]:
```
import pandas as pd
import matplotlib.pyplot as plt
import sklearn as sk
```

In [22]:
```
import numpy as np
```

In [23]:
```
df=pd.read_csv("StudentsPerformance.csv")
```

In [24]:
```
df
```

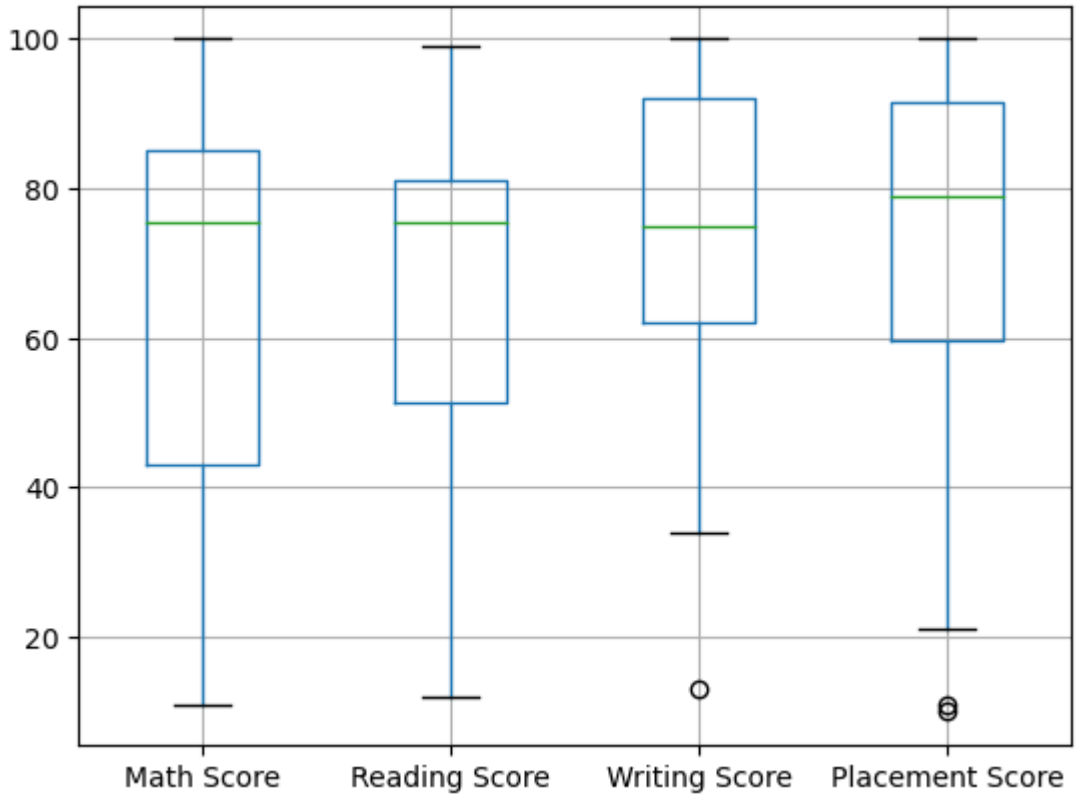
Out[24]:

	Math Score	Reading Score	Writing Score	Placement Score	Club Join Year	Placement	Offer Count	gender
0	95.0	81.0	93.0	85.0	23	98.0	99	female
1	83.0	79.0	64.0	97.0	23	66.0	97	female
2	100.0	96.0	70.0	86.0	23	84.0	69	female
3	88.0	83.0	100.0	81.0	23	79.0	79	female
4	73.0	72.0	96.0	62.0	23	64.0	64	female
5	63.0	76.0	NaN	100.0	23	83.0	74	female
6	72.0	87.0	99.0	94.0	23	72.0	68	female
7	90.0	75.0	91.0	66.0	23	88.0	86	female
8	97.0	79.0	74.0	77.0	23	90.0	91	female
9	NaN	60.0	77.0	70.0	23	81.0	69	Male
10	77.0	65.0	61.0	90.0	23	60.0	67	Male
11	75.0	94.0	88.0	91.0	23	90.0	71	Male
12	85.0	79.0	63.0	93.0	23	99.0	80	Male
13	80.0	79.0	100.0	NaN	23	76.0	95	Male
14	81.0	81.0	88.0	61.0	23	91.0	84	Male
15	80.0	64.0	98.0	100.0	23	69.0	96	Male
16	66.0	85.0	94.0	84.0	23	60.0	62	Male
17	99.0	80.0	75.0	93.0	23	72.0	76	Male
18	85.0	NaN	81.0	99.0	23	63.0	73	Male
19	76.0	99.0	89.0	84.0	23	97.0	87	Male
20	34.0	23.0	45.0	21.0	23	45.0	65	female
21	43.0	23.0	34.0	56.0	23	56.0	56	female
22	45.0	65.0	67.0	43.0	23	45.0	76	female
23	12.0	24.0	NaN	10.0	23	56.0	87	female
24	23.0	21.0	43.0	54.0	23	54.0	23	female
25	11.0	12.0	13.0	11.0	23	NaN	16	Male
26	13.0	14.0	54.0	44.0	23	77.0	65	Male
27	43.0	54.0	65.0	76.0	23	43.0	43	Male
28	43.0	44.0	54.0	65.0	23	11.0	12	Male

In [36]:
```
!pip install matplotlib
```

```
Requirement already satisfied: matplotlib in c:\users\system21\anaconda3\li
b\site-packages (3.7.2)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\system21\anacon
da3\lib\site-packages (from matplotlib) (1.0.5)
Requirement already satisfied: cycler>=0.10 in c:\users\system21\anaconda3
\lib\site-packages (from matplotlib) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\system21\anaco
nda3\lib\site-packages (from matplotlib) (4.25.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\system21\anaco
nda3\lib\site-packages (from matplotlib) (1.4.4)
Requirement already satisfied: numpy>=1.20 in c:\users\system21\anaconda3\l
ib\site-packages (from matplotlib) (1.24.3)
Requirement already satisfied: packaging>=20.0 in c:\users\system21\anacond
a3\lib\site-packages (from matplotlib) (23.1)
Requirement already satisfied: pillow>=6.2.0 in c:\users\system21\anaconda3
\lib\site-packages (from matplotlib) (9.4.0)
Requirement already satisfied: pyparsing<3.1,>=2.3.1 in c:\users\system21\a
naconda3\lib\site-packages (from matplotlib) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\system21\an
aconda3\lib\site-packages (from matplotlib) (2.8.2)
Requirement already satisfied: six>=1.5 in c:\users\system21\anaconda3\lib
\site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)
```

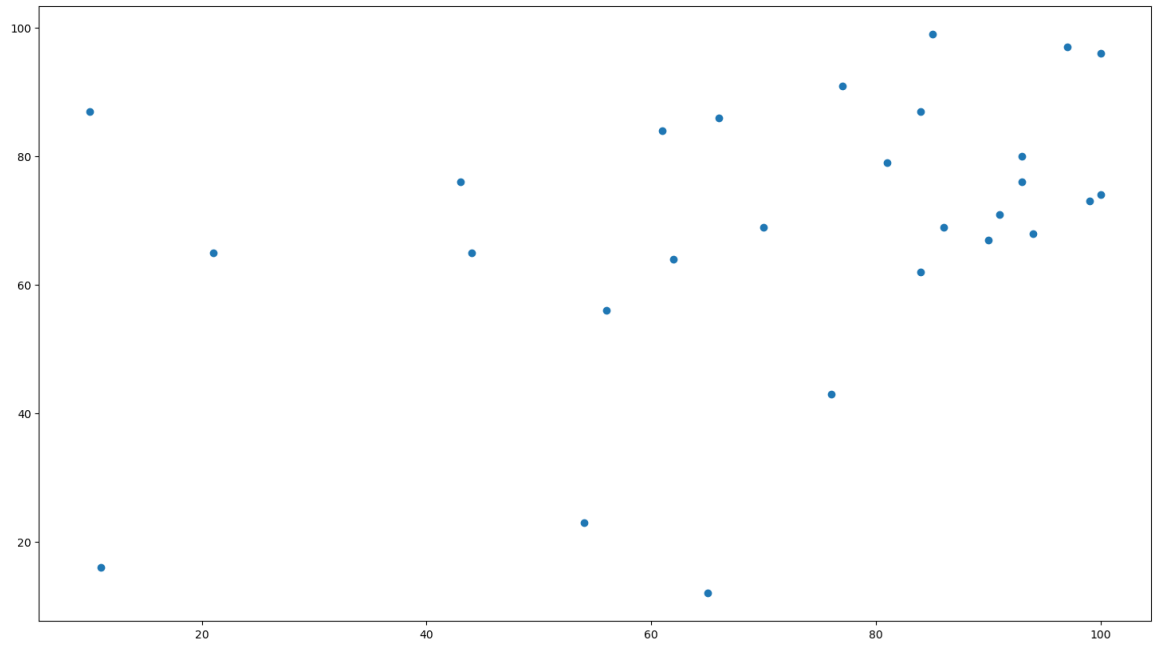
In [25]:
```
col = ['Math Score', 'Reading Score' , 'Writing Score','Placement Score']
df.boxplot(col)
plt.show()
```



In [26]:
```
print(np.where(df['Math Score']>90))
print(np.where(df['Reading Score']<25))
print(np.where(df['Writing Score']<30))
```

```
(array([ 0,  2,  8, 17], dtype=int64),)
(array([20, 21, 23, 24, 25, 26], dtype=int64),)
(array([25], dtype=int64),)
```

In [31]:
```
fig, ax = plt.subplots(figsize = (18,10))
ax.scatter(df['Placement Score'], df['Offer Count'])
plt.show()
```



In [ ]:
```
Name-Anurag Jadhav
Roll No-12371
Practical-2
Class-TE A
```