In [1]:

```
import pandas as pd
import numpy as np
```

In [2]:

```
df = pd.read_csv("C:/Users/ameya/OneDrive/Desktop/DSBDAL/UpdatedStudentsPerformance.csv"
```

In [3]:

```
df.head()
```

Out[3]:

	gender	race/ethnicity	parental level of education	lunch	test preparation course	math score	reading score	writing score
0	female	group B	bachelor's degree	standard	none	72.0	72.0	74.0
1	female	group C	some college	standard	completed	69.0	90.0	88.0
2	female	group B	master's degree	standard	none	90.0	95.0	93.0
3	male	group A	associate's degree	free/reduced	none	47.0	57.0	44.0
4	male	group C	some college	standard	none	76.0	78.0	75.0

In [4]:

```
df.isnull().sum()
```

Out[4]:

```
gender                         0
race/ethnicity                 0
parental level of education    0
lunch                          0
test preparation course        0
math score                    10
reading score                 15
writing score                 11
dtype: int64
```

In [5]:

```
math_mean = df['math score'].mean()
```

In [6]:

```
math_mean
```

Out[6]:

```
66.20808080808081
```

In [7]:

```
reading_median = df['reading score'].median()
```

In [8]:

```
reading_median
```

Out[8]:

```
70.0
```

In [9]:

```
df['math score'].fillna(math_mean,inplace=True)
```

In [10]:

```
df.isnull().sum()
```

Out[10]:

```
gender                         0
race/ethnicity                 0
parental level of education    0
lunch                          0
test preparation course        0
math score                     0
reading score                 15
writing score                 11
dtype: int64
```

In [11]:

```
df['reading score'].fillna(reading_median,inplace=True)
```

In [12]:

```
df.isnull().sum()
```

Out[12]:

```
gender                         0
race/ethnicity                 0
parental level of education    0
lunch                          0
test preparation course        0
math score                     0
reading score                  0
writing score                 11
dtype: int64
```

In [26]:

```
df.dropna(inplace = True)
```

In [28]:

```
df.isnull().sum()
```

Out[28]:

```
gender                        0
race/ethnicity                0
parental level of education   0
lunch                         0
test preparation course       0
math score                    0
reading score                 0
writing score                 0
dtype: int64
```

In [30]:

```
df.shape
```

Out[30]:

```
(989, 8)
```

In [32]:

```
col = ["math score","reading score","writing score"]
```

In [33]:

```
from sklearn.preprocessing import MinMaxScaler
scalar = MinMaxScaler()
df_scaled = scalar.fit_transform(df[col].to_numpy())
df_scaled = pd.DataFrame(df_scaled,columns=col)
print("Scaled Dataset using MinMaxScaler")
df_scaled.head()
```

```
Scaled Dataset using MinMaxScaler
```

Out[33]:

	math score	reading score	writing score
0	0.72	0.662651	0.711111
1	0.69	0.879518	0.866667
2	0.90	0.939759	0.922222
3	0.47	0.481928	0.377778
4	0.76	0.734940	0.722222

In [ ]:

```

```