

Team ID	PNT2022TMID21972
Project Name	Efficient Water Quality Analysis and Prediction using Machine Learning

## Handling missing values 2

In [11]: data.dtypes

```
Out[11]: STATION CODE      object
LOCATIONS      object
STATE          object
Temp           object
D.O. (mg/l)    object
PH             object
CONDUCTIVITY (µmhos/cm) object
B.O.D. (mg/l)  object
NITRATENAN N+ NITRITENANN (mg/l) object
FECAL COLIFORM (MPN/100ml) object
TOTAL COLIFORM (MPN/100ml)Mean object
year           int64
dtype: object
```

```
In [12]: data['Temp']=pd.to_numeric(data['Temp'],errors='coerce')
data['D.O. (mg/l)']=pd.to_numeric(data['D.O. (mg/l)'],errors='coerce')
data['PH']=pd.to_numeric(data['PH'],errors='coerce')
data['B.O.D. (mg/l)']=pd.to_numeric(data['B.O.D. (mg/l)'],errors='coerce')
data['CONDUCTIVITY (µmhos/cm)']=pd.to_numeric(data['CONDUCTIVITY (µmhos/cm)'],errors='coerce')
data['NITRATENAN N+ NITRITENANN (mg/l)']=pd.to_numeric(data['NITRATENAN N+ NITRITENANN (mg/l)'],errors='coerce')
data['TOTAL COLIFORM (MPN/100ml)Mean']=pd.to_numeric(data['TOTAL COLIFORM (MPN/100ml)Mean'],errors='coerce')
data.dtypes
```

```
Out[12]: STATION CODE      object
LOCATIONS      object
STATE          object
Temp           float64
D.O. (mg/l)    float64
PH             float64
CONDUCTIVITY (µmhos/cm) float64
B.O.D. (mg/l)  float64
NITRATENAN N+ NITRITENANN (mg/l) float64
FECAL COLIFORM (MPN/100ml) object
TOTAL COLIFORM (MPN/100ml)Mean float64
year           int64
dtype: object
```

In [13]: data.isnull().sum()

```
Out[13]: STATION CODE      0
LOCATIONS      0
STATE          0
Temp           92
D.O. (mg/l)    31
PH             8
CONDUCTIVITY (µmhos/cm) 25
B.O.D. (mg/l)  43
NITRATENAN N+ NITRITENANN (mg/l) 225
FECAL COLIFORM (MPN/100ml) 0
TOTAL COLIFORM (MPN/100ml)Mean 132
year           0
```