

|                     |  |
|---------------------|--|
| <b>TEAM ID</b>      | PNT2022TMID16214   |
| <b>PROJECT NAME</b> | Efficient Water Quality Analysis and Prediction using Machine Learning |

## Handling Missing Value 3

```
In [19]: data['Temp'].fillna(data['Temp'].mean(),inplace=True)
data['D.O. (mg/l)'].fillna(data['D.O. (mg/l)'].mean(),inplace=True)
data['PH'].fillna(data['PH'].mean(),inplace=True)
data['CONDUCTIVITY (umhos/cm)'].fillna(data['CONDUCTIVITY (umhos/cm)'].mean(),inplace=True)
data['B.O.D. (mg/l)'].fillna(data['B.O.D. (mg/l)'].mean(),inplace=True)
data['NITRATENAN N+ NITRITENANN (mg/l)'].fillna(data['NITRATENAN N+ NITRITENANN (mg/l)'].mean(),inplace=True)
data['TOTAL COLIFORM (MPN/100ml)Mean'].fillna(data['TOTAL COLIFORM (MPN/100ml)Mean'].mean(),inplace=True)
```

```
In [21]: data.drop(['FECAL COLIFORM (MPN/100ml)'],axis=1,inplace=True)
```

```
In [22]: data=data.rename(columns={'PH':'ph'})
data=data.rename(columns={'D.O. (mg/l)':'do'})
data=data.rename(columns={'CONDUCTIVITY (umhos/cm)':'co'})
data=data.rename(columns={'B.O.D. (mg/l)':'bod'})
data=data.rename(columns={'NITRATENAN N+ NITRITENANN (mg/l)':'na'})
data=data.rename(columns={'TOTAL COLIFORM (MPN/100ml)Mean':'tc'})
data=data.rename(columns={'STATION CODE':'station'})
data=data.rename(columns={'LOCATIONS':'location'})
data=data.rename(columns={'STATE':'state'})
```