PNT2022TMID21280_Efficient Water Quality Analysis & Prediction using Machine Learning.

Analyse the data:

In [4]:
```
data.describe()
```

Out[4]:

	year
count	1991.000000
mean	2010.038172
std	3.057333
min	2003.000000
25%	2008.000000
50%	2011.000000
75%	2013.000000
max	2014.000000

In [5]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1991 entries, 0 to 1990
Data columns (total 12 columns):
 #   Column                              Non-Null Count  Dtype
---  ------                              --------------  -----
 0   STATION CODE                        1991 non-null   object
 1   LOCATIONS                           1991 non-null   object
 2   STATE                               1991 non-null   object
 3   Temp                                1991 non-null   object
 4   D.O. (mg/l)                         1991 non-null   object
 5   PH                                  1991 non-null   object
 6   CONDUCTIVITY (µmhos/cm)             1991 non-null   object
 7   B.O.D. (mg/l)                       1991 non-null   object
 8   NITRATENAN N+ NITRITENANN (mg/l)    1991 non-null   object
 9   FECAL COLIFORM (MPN/100ml)          1991 non-null   object
 10  TOTAL COLIFORM (MPN/100ml)Mean      1991 non-null   object
 11  year                                1991 non-null   int64
dtypes: int64(1), object(11)
memory usage: 186.8+ KB
```

In [6]:
```
data.shape
```

Out[6]: (1991, 12)