| TEAM ID | PNT2022TMID16214 |
|---|---|
| **PROJECT NAME** | Efficient Water Quality Analysis and Prediction using Machine Learning |

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

Out[4]:

| | STATION CODE | LOCATIONS | STATE | Temp | D.O. (mg/l) | PH | CONDUCTIVITY (µmhos/cm) | B.O.D. (mg/l) | NITRATENAN N+ NITRITENANN (mg/l) | FECAL COLIFORM (MPN/100ml) | TOTAL COLIFORM (MPN/100ml)Mean | year |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1393 | DAMANGANGA AT D/S OF MADHUBAN, DAMAN | DAMAN & DIU | 30.6 | 6.7 | 7.5 | 203 | NAN | 0.1 | 11 | 27 | 2014 |
| 1 | 1399 | ZUARI AT D/S OF PT. WHERE KUMBARJRIA CANAL JOI... | GOA | 29.8 | 5.7 | 7.2 | 189 | 2 | 0.2 | 4953 | 8391 | 2014 |
| 2 | 1475 | ZUARI AT PANCHAWADI | GOA | 29.5 | 6.3 | 6.9 | 179 | 1.7 | 0.1 | 3243 | 5330 | 2014 |
| 3 | 3181 | RIVER ZUARI AT BORIM BRIDGE | GOA | 29.7 | 5.8 | 6.9 | 64 | 3.8 | 0.5 | 5382 | 8443 | 2014 |
| 4 | 3182 | RIVER ZUARI AT MARCAIM JETTY | GOA | 29.5 | 5.8 | 7.3 | 83 | 1.9 | 0.4 | 3428 | 5500 | 2014 |

Activate Windows

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

Out[5]:

| | 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 [6]: 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 [7]: data.shape

Out[7]: (1991, 12)
```