PROJECT DEVELOPMENT - DELIVERY OF SPRINT 2

Date:	14 November 2022
Team ID:	PNT2022TMID08065
Name:	Efficient Water Quality analysis & Prediction using Machine Learning.

PYTHON CODE

```python
import numpy as np

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

import warnings

from google.colab import drive

drive.mount('/content/drive')

data=pd.read_csv('/content/drive/My Drive/water_dataX.csv',encoding='ISO-8859-1',low_memory=False)

data.shape

data.head

data.describe

data.info()

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.dtypes

```
data['NITRATENAN N+
NITRITENANN
(mg/l)']=pd.to_numeric(d
ata['NITRATENAN N+
NITRITENANN
(mg/l)'],errors='coerce')
```

```
data['TOTAL
COLIFORM
(MPN/100ml)Mean']=pd.t
o_numeric(data['TOTAL
COLIFORM
(MPN/100ml)Mean'],erro
rs='coerce')
```

data.dtypes

```
# Data visualization
Sns.catplot(x='wbdo',
y='wec'
,kind='bar',data=data)
```

Output: