**WATER QUALITY ANALYSIS:**

**TEAM NUMBER 11**

*TEAM LEAD*

*1)SATHYA PRIYA-310821106079*

*https://github.com/sathpriya123/Water\_Analysis.git*

*TEAM MEMBERS*

*2)Reg no:310821106067*

*Name:Ramya R*

*3)Reg no:310821106082*

*Name: Sornasri M*

*4)Reg no:310821106084*

*Name: Sridevi K*

**CONCEPT**

Water quality analysis involves assessing the physical, chemical, and biological characteristics of water to determine its suitability for various purposes. The process typically includes.

**DATA SOURCE:**

<https://www.kaggle.com/datasets/adityakadiwal/water-potability>

1. **DATA COLLECTION:**

This step involves gathering water samples from different sources, such as rivers, lakes, or tap water. Proper sampling techniques are crucial to ensure the collected samples are representative of the water source.

Collect water samples from your source(s).Use appropriate instruments to measure pH, turbidity, salinity, and hardness. Ensure proper calibration of these instruments.Record the measurements along with relevant information such as date, time, and location.

**PHYSICAL & CHEMICAL ANALYSIS:**

This assesses physical properties like temperature, turbidity, colour, and odour. These parameters can provide insights into water's appearance and clarity.

This involves testing for chemical components, such as pH, dissolved oxygen, nutrients (nitrate, phosphate), heavy metals, and organic contaminants. These measurements help determine the presence of pollutants and the water's chemical balance.

**BIOLOGICAL ANALYSIS:**

It includes the examination of microorganisms and macroinvertebrates to assess water quality and the health of aquatic ecosystems. Bacteria, algae, and other organisms can indicate pollution or ecosystem disturbances.

**2)** **DATA PROCESSING:**

* Load the database into a data analysis library like Pandas in Python.
* Clean data using missing, duplicate and outlier values.
* Change the data type if needed (eg date to time object).
* Make sure the column names are consistent and in the correct format.

**3) DATA INTERPRETATION:**

After collecting data, it's crucial to interpret the results. Comparing the obtained values to water quality standards or guidelines helps determine if the water meets specific criteria for drinking, recreational, or ecological use.

**Step-1:**

% Load data from a CSV file

data = csvread('your\_data.csv');

**Step-2:**

% Visualise the data with a scatter plot

scatter(data(:,1), data(:,2));

xlabel('X-axis');

ylabel('Y-axis');

title('Scatter Plot of Data');

**Step-3:**

% Calculate mean and standard deviation

mean\_value = mean(data(:,1));

std\_deviation = std(data(:,2));

disp(['Mean: ', num2str(mean\_value)]);

disp(['Standard Deviation: ', num2str(std\_deviation)]);

**4) CALCULATE THE AVERAGE :**

- Calculate the level of SO2, NO2 and RSPM/PM10 for

each category (eg state or city/town/village/region). You

can use the "groupby" function in Pandas to group the data

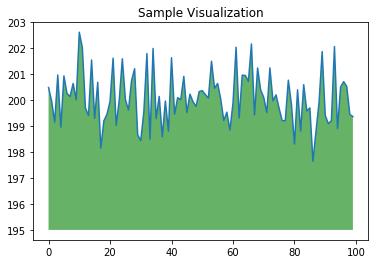
and the "mean()" function to calculate the mean.

**5) DATA VISUALISATION:**

Create data visualisations to present the results.

For this you can use a data visualisation library like Matplotlib or

Seaborn.



**CODING**

# Import necessary libraries

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# Load your dataset into a Pandas DataFrame

# Replace 'your\_data.csv' with the actual path to your

dataset

df = pd.read\_csv("/content/water\_Quality\_Dataset.csv")

# Calculate average PH, turbidity,hardness, dissolved O2, conductivity, temperature,10 levels across different areas

average\_data = df.groupby["Sample ID","pH","Temperature ( °C)","Turbidity (NTU)","Dissolved Oxygen,No2,(mg/L)","Conductivity ( µS/cm)"]'RSPM/PM10']].mean().reset\_index()

# Create visualisations

# Bar plot for average turbidity levels

plt.figure(figsize=(12, 6))

sns.barplot(x='turbidity', y='City/Town/Village/Area',

data=average\_data, hue='State')

plt.title('Average turbid Levels')

plt.xlabel('turbidity' (µg/m³)')

plt.ylabel('City/Town/Village/Area')

plt.legend(title='State')

plt.show()

# Bar plot for average 'dissolved O2' levels

plt.figure(figsize=(12, 6))

sns.barplot(x='dissolved O2',y='City/Town/Village/Area',

data=average\_data, hue='State')

plt.title('Average 'dissolved O2' Levels')

plt.xlabel('dissolved O2' (µg/m³)')

plt.ylabel('City/Town/Village/Area')

plt.legend(title='State')

plt.show()

# Bar plot for average NO2 levels

plt.figure(figsize=(12, 6))

sns.barplot(x='NO2', y='City/Town/Village/Area',

data=average\_data, hue='State')

plt.title('Average NO2 Levels')

plt.xlabel('NO2 (µg/m³)')

plt.ylabel('City/Town/Village/Area')

plt.legend(title='State')

plt.show()

# Bar plot for average 'hardness' levels

plt.figure(figsize=(12, 6))

sns.barplot(x='hardness',

y='City/Town/Village/Area', data=average\_data,

hue='State')

plt.title('Average 'hardness' Levels')

plt.xlabel('hardness' (µg/m³)')

plt.ylabel('City/Town/Village/Area')

plt.legend(title='State')

plt.show()

# Bar plot for average RSPM/PM10 levels

plt.figure(figsize=(12, 6))

sns.barplot(x='RSPM/PM10',

y='City/Town/Village/Area', data=average\_data,

hue='State')

plt.title('Average RSPM/PM10 Levels')

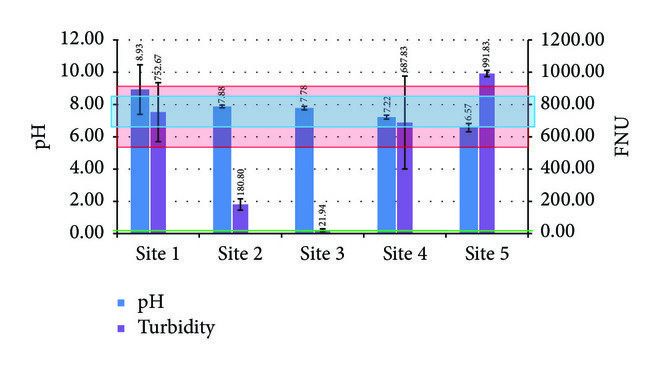
plt.xlabel('RSPM/PM10 (µg/m³)')

plt.ylabel('City/Town/Village/Area')

plt.legend(title='State')

plt.show()

**OUTPUT:**

****

**REPORT GENERATION:** The findings are typically summarised in a report, including data, analysis, and recommendations for any necessary actions to improve water quality or address issues.