SDG6: Clean water and Sanitation.

Problem definition:

A key issue within the clean water and sanitation development goal is the monitoring of water quality in urban and rural areas. Many communities suffer from contaminated water sources, leading to health issues like waterborne diseases. The availability of real-time water quality data in different locations can help mitigate this problem by providing insights into which areas require intervention and allowing authorities to take preventive measures. Therefor saving countless lives in the process and improving standard of living within the regions, as better water, leads to better health and better ability to be economically viable.

Specific Problem:

Inadequate tracking of water contamination levels, such as pH, lead, arsenic, and bacterial contamination, in different regions. The goal is to create a system that collects, stores, and analyses water quality data from different water sources, helping decision-makers identify areas where water treatment is urgently needed.

SQL scripts to create tables:

Create 'Location' table

```
CREATE TABLE Location (
    Location_ID INT PRIMARY KEY AUTO_INCREMENT,
    Location_Name VARCHAR(100) NOT NULL,
   Region VARCHAR(100) NOT NULL,
     Country VARCHAR(100) NOT NULL );


-- Create 'Water_Sample' table
CREATE TABLE Water_Sample (
    Sample_ID INT PRIMARY KEY AUTO_INCREMENT,
    Location_ID INT NOT NULL,
    Sample_Date DATE NOT NULL,
    Sample_Time TIME NOT NULL,
    Collected_By VARCHAR(100),
    FOREIGN KEY (Location_ID) REFERENCES Location(Location_ID));
```

```sql
-- Create 'Water_Quality' table

CREATE TABLE Water_Quality (

    Quality_ID INT PRIMARY KEY AUTO_INCREMENT,

    Sample_ID INT NOT NULL,

    pH DECIMAL(3, 2) NOT NULL,

    Lead_Concentration DECIMAL(5, 3) NOT NULL,

    Arsenic_Concentration DECIMAL(5, 3) NOT NULL,

    Bacterial_Contamination BOOLEAN NOT NULL,

    Overall_Quality VARCHAR(50) NOT NULL,

    FOREIGN KEY (Sample_ID) REFERENCES Water_Sample(Sample_ID));


-- Create 'Users' table

CREATE TABLE Users (

    User_ID INT PRIMARY KEY AUTO_INCREMENT,

    Username VARCHAR(100) NOT NULL,

    Password VARCHAR(100) NOT NULL,

    Role ENUM('Admin', 'Field Agent', 'Public User') NOT NULL);
```

SQL scripts to input mock data:

```sql
-- Insert sample data into 'Location' table

INSERT INTO Location (Location_Name, Region, Country)

VALUES

('Tana River', 'Eastern Region', 'Kenya'),

('Vaal Dam', 'Western  Region', 'South Africa'),

('Lake Michigan', 'Western Region', 'United states of America'),

('Amazon River', 'North Region', 'Brazil'),

('River Thames', 'Southern Region', 'United Kingdom'),

('Lake Tanganyika', 'Eastern Region', 'Tanzania'),

('Nile River', 'North Region', 'Egypt'),

('Yangtze River', 'Central Region', 'China'),

('Mississippi River', 'Central Region', 'United States'),

('Ganges River', 'Northern Region', 'India'),
```

```
('Volga River', 'Western Region', 'Russia'),

('Danube River', 'Central Region', 'Germany'),

('Lake Victoria', 'Eastern Region', 'Kenya');


-- Insert sample data into 'Water_Sample' table

INSERT INTO Water_Sample (Location_ID, Sample_Date, Sample_Time, Collected_By)

VALUES

(1, '2023-05-20', '08:00:00', 'Tana and Athi Rivers development Authority'),

(2, '2023-06-21', '09:15:00', 'Dept of water and environmental affairs'),

(3, '2023-08-12', '07:45:00', 'Lake Michigan Lakewide Action and Management Plans'),

(4, '2023-07-15', '10:00:00', 'Ministry of Water Resources - Egypt'),

(5, '2023-09-12', '08:30:00', 'UK Environment Agency'),

(6, '2023-08-10', '09:45:00', 'Tanzania Water Authority'),

(7, '2023-07-22', '07:30:00', 'Nile Basin Initiative'),

(8, '2023-09-03', '06:50:00', 'Yangtze River Water Resource Commission'),

(9, '2023-06-15', '08:00:00', 'Mississippi Department of Environmental Quality'),

(10, '2023-07-18', '09:10:00', 'Central Pollution Control Board - India'),

(11, '2023-09-25', '10:25:00', 'Federal Service for Hydrometeorology and Environmental Monitoring - Russia'),

(12, '2023-08-30', '07:55:00', 'International Commission for the Protection of the Danube River'),

(13, '2023-06-19', '08:20:00', 'Kenya Water Resources Authority');
```

```sql
-- Insert sample data into 'Water_Quality' table

INSERT INTO Water_Quality (Sample_ID, pH, Lead_Concentration, Arsenic_Concentration,
Bacterial_Contamination, Overall_Quality)

VALUES

(1, 7.2, 0.001, 0.002, FALSE, 'Good'),

(2, 6.5, 0.015, 0.007, TRUE, 'Poor'),

(3, 7.0, 0.003, 0.001, FALSE, 'Moderate'),

(4, 7.5, 0.002, 0.004, FALSE, 'Good'),

(5, 6.9, 0.010, 0.003, TRUE, 'Moderate'),

(6, 7.1, 0.006, 0.002, FALSE, 'Good'),

(7, 7.3, 0.012, 0.008, TRUE, 'Poor'),

(8, 6.8, 0.007, 0.002, TRUE, 'Moderate'),

(9, 7.4, 0.003, 0.001, FALSE, 'Good'),

(10, 6.6, 0.015, 0.010, TRUE, 'Poor'),

(11, 7.0, 0.002, 0.005, FALSE, 'Good'),

(12, 6.9, 0.009, 0.006, TRUE, 'Moderate'),

(13, 7.2, 0.004, 0.003, FALSE, 'Good');
```

```sql
-- Insert sample data into 'Users' table
INSERT INTO Users (Username, Password, Role)
VALUES
('admin_user', 'admin123', 'Admin'),
('field_agent_1', 'Megan456', 'Field Agent'),
('public_user', 'pub789', 'Public User'),
('field_agent_2', 'JohnDoe123', 'Field Agent'),
('field_agent_3', 'JaneSmith456', 'Field Agent'),
('field_agent_4', 'MikeJohnson789', 'Field Agent'),
('lab_analyst_1', 'EmilyDavis123', 'Lab Analyst'),
('lab_analyst_2', 'ChrisBrown456', 'Lab Analyst'),
('lab_analyst_3', 'OliviaTaylor789', 'Lab Analyst'),
('data_scientist_1', 'NoahWilson123', 'Data Scientist'),
('data_scientist_2', 'AvaLee456', 'Data Scientist'),
('data_scientist_3', 'OliverCarter789', 'Data Scientist'),
('system_admin_2', 'SophiaKing123', 'System Admin');
```

Data Retrieval:

Retrieving all water quality data for a specified location;

```
SELECT

    L.Location_Name,

    WQ.pH,

    WQ.Lead_Concentration,

    WQ.Arsenic_Concentration,

    WQ.Bacterial_Contamination,

    WQ.Overall_Quality,

    WS.Sample_Date,

    WS.Sample_Time

FROM

    Water_Quality WQ

INNER JOIN

    Water_Sample WS ON WQ.Sample_ID = WS.Sample_ID

INNER JOIN

    Location L ON WS.Location_ID = L.Location_ID

WHERE

    L.Location_Name = 'Vaal Dam';
```

SQL command to list all locations with Moderate water quality:

```
SELECT

    L.Location_Name,

    L.Region,

    WQ.Overall_Quality

FROM

    Water_Quality WQ

INNER JOIN

    Water_Sample WS ON WQ.Sample_ID = WS.Sample_ID

INNER JOIN

    Location L ON WS.Location_ID = L.Location_ID

WHERE

    WQ.Overall_Quality = 'Moderate';
```

SQL Command to retrieve water quality data collected by a specific government organization:

```
SELECT

    L.Location_Name,

    WS.Sample_Date,

    WQ.pH,

    WQ.Lead_Concentration,

    WQ.Arsenic_Concentration,

    WQ.Bacterial_Contamination,

    WQ.Overall_Quality

FROM

    Water_Quality WQ

INNER JOIN

    Water_Sample WS ON WQ.Sample_ID = WS.Sample_ID

INNER JOIN

    Location L ON WS.Location_ID = L.Location_ID

WHERE

    WS.Collected_By = 'Tana and Athi Rivers development authority' ;
```

SQL command to retrieve average pH and contaminant for reach Region:

```
 SELECT

    L.Region,

    AVG(WQ.pH) AS Avg_pH,

    AVG(WQ.Lead_Concentration) AS Avg_Lead,

    AVG(WQ.Arsenic_Concentration) AS Avg_Arsenic

FROM

    Water_Quality WQ

INNER JOIN

    Water_Sample WS ON WQ.Sample_ID = WS.Sample_ID

INNER JOIN

    Location L ON WS.Location_ID = L.Location_ID

GROUP BY

    L.Region;
```

SQL command to identify and retrieve locations with high lead concentration:

```
SELECT

    L.Location_Name,

    L.Region,

    WQ.Lead_Concentration

FROM

    Water_Quality WQ

INNER JOIN

    Water_Sample WS ON WQ.Sample_ID = WS.Sample_ID

INNER JOIN

    Location L ON WS.Location_ID = L.Location_ID

WHERE

    WQ.Lead_Concentration > 0.01;
```

Count command in SQL:

This command is to count the total amount of samples that tested positive for bacterial contamination.

SELECT

    COUNT(*) AS Contaminated_Samples

FROM

    Water_Quality

WHERE

    Bacterial_Contamination = TRUE;