

## **SDG NUMBER 6: CLEAN WATER AND SANITATION**

**PART I: Specific Problem to Address:** Access to clean drinking water in rural areas.

### **Problem Definition:**

In many rural areas, access to clean and safe drinking water remains a significant challenge. Inequities in access can arise due to factors such as geographic location, infrastructure quality, and socioeconomic status. Addressing this issue requires data-driven solutions to identify areas with inadequate water access and to allocate resources effectively.

## **PART II: DATABASE DESIGN**

2.1) ERD at the bottom of the page

2.2) SQL statements to create database schema include the following:

### **Creating Country Table**

```
CREATE TABLE Country (  
    Country_ID INT PRIMARY KEY AUTO_INCREMENT,  
    Country_Name VARCHAR(100) NOT NULL,  
    ISO3_Code CHAR(3) NOT NULL UNIQUE  
);
```

### **Creating Year Table**

```
CREATE TABLE Year (  
    Year_ID INT PRIMARY KEY,  
    Year INT NOT NULL  
);
```

### **Creating Population Table**

```
CREATE TABLE Population (  
    Country_ID INT,
```

```

Population_ID INT PRIMARY KEY AUTO_INCREMENT,
Country_ID INT,
Year_ID INT,
School_Age_Population_Thousands INT,
Percent_Urban DECIMAL(5, 2),
Percent_Pre_Primary DECIMAL(5, 2),
Percent_Primary DECIMAL(5, 2),
Percent_Secondary DECIMAL(5, 2),
FOREIGN KEY (Country_ID) REFERENCES Country(Country_ID),
FOREIGN KEY (Year_ID) REFERENCES Year(Year_ID)
);

```

### 2.3 Sample Data Insertion

-- Insert into Country table

```
INSERT INTO Country (Country_Name, ISO3_Code) VALUES ('Kenya', 'KEN');
```

-- Insert into Year table

```
INSERT INTO Year (Year_ID, Year) VALUES (2000, 2000), (2001, 2001), (2002, 2002);
```

-- Insert into Population table

```
INSERT INTO Population (Country_ID, Year_ID, School_Age_Population_Thousands,
Percent_Urban, Percent_Pre_Primary, Percent_Primary, Percent_Secondary)
VALUES (1, 2000, 12985, 20, 22, 41, 36),
```

(1, 2001, 13254, 20, 23, 41, 37),

(1, 2002, 13545, 21, 23, 41, 37);

## **PART III: SQL SCRIPTS**

### **3.1) Retrieve Areas with the Lowest Access to Basic Water Services**

```
SELECT c.Country_Name, y.Year, p.Percent_Basic_Water_Services
FROM Population p
JOIN Country c ON p.Country_ID = c.Country_ID
JOIN Year y ON p.Year_ID = y.Year_ID
WHERE y.Year = 2023
ORDER BY p.Percent_Basic_Water_Services ASC
LIMIT 10; -- Adjust the LIMIT value to show more or fewer results
```

### **Compare Water Access Between Urban and Rural Areas**

```
SELECT c.Country_Name, y.Year,
       p.Percent_Basic_Water_Services AS Percent_Basic_Water_Urban,
       p.Percent_Limited_Water_Services AS Percent_Limited_Water_Urban,
       p.Percent_No_Water_Service AS Percent_No_Water_Service_Urban
FROM Population p
JOIN Country c ON p.Country_ID = c.Country_ID
JOIN Year y ON p.Year_ID = y.Year_ID
WHERE c.Country_Name = 'Kenya' AND y.Year = 2023;
```

### **Identify Rural Areas with No Water Services**

```
SELECT c.Country_Name, y.Year, p.Percent_No_Water_Service
FROM Population p
JOIN Country c ON p.Country_ID = c.Country_ID
JOIN Year y ON p.Year_ID = y.Year_ID
WHERE y.Year = 2023 AND p.Percent_No_Water_Service > 50
ORDER BY p.Percent_No_Water_Service DESC;
```

### **3.2) Distribution of Hygiene Services Across Regions**

```
SELECT c.SDG_Region, AVG(p.Percent_Basic_Hygiene_Services) AS
Avg_Percent_Basic_Hygiene_Services
FROM Population p
JOIN Country c ON p.Country_ID = c.Country_ID
JOIN Year y ON p.Year_ID = y.Year_ID
WHERE y.Year = 2023
GROUP BY c.SDG_Region
ORDER BY Avg_Percent_Basic_Hygiene_Services DESC;
```

### **Water Access in Countries with High Rural Population**

```
SELECT c.Country_Name,
       p.Percent_Basic_Water_Services_Rural AS Percent_Basic_Water_Rural,
       p.Percent_No_Water_Service_Rural AS Percent_No_Water_Service_Rural
```

```
FROM Population p
JOIN Country c ON p.Country_ID = c.Country_ID
JOIN Year y ON p.Year_ID = y.Year_ID
WHERE y.Year = 2023
ORDER BY p.Percent_Rural_Population DESC
LIMIT 10; -- Show top 10 countries with high rural populations
```

### **Historical Overview of Water Access for a Specific Region**

```
SELECT y.Year,
        AVG(p.Percent_Basic_Water_Services) AS Avg_Percent_Basic_Water_Services
FROM Population p
JOIN Country c ON p.Country_ID = c.Country_ID
JOIN Year y ON p.Year_ID = y.Year_ID
WHERE c.SDG_Region = 'Eastern and Southern Africa'
GROUP BY y.Year
ORDER BY y.Year;
```

### **PART V: Documentation process of importing data from database to Excel**

To import data related to SDG number 6 into an Excel spreadsheet while ensuring consistency, follow these steps:

#### **1. Data Collection:**

- Gather relevant data on SDG 6 (e.g., water and sanitation indicators) from reliable sources.

#### **2. Data Cleaning:**

- Review the data for errors, duplicates, and inconsistencies.
- Ensure all data entries are in the same format (e.g., consistent date formats, units of measurement).

### **3. Data Structuring:**

- Organize the data into columns, with each column representing a specific variable (e.g., country, year, water access percentage).
- Use a consistent header naming convention.

### **4. Import into Excel:**

- Open Excel and import the data using the "Import Data" feature or by copying and pasting from the source.
- If importing from a CSV file, ensure the delimiter settings match the data format.

### **5. Data Validation:**

- Use Excel's data validation tools to enforce consistency (e.g., restrict inputs to certain ranges or formats).
- Check for missing or incorrect data using Excel functions like ISBLANK or IFERROR.

### **6. Review and Save:**

- Once imported and validated, review the data for accuracy.
- Save the Excel file with a clear and descriptive name, indicating the content and version.

This process helps maintain data integrity, ensuring that the spreadsheet accurately reflects the collected SDG 6 data.

Country	
PK	<u>country_ID</u>
	country_name(var 50)
	Iso3_code

Year	
PK	<u>Year ID</u>
FK1	Year

Population	
PK	<u>population id</u>
FK1	country_id
FK1	year_id

