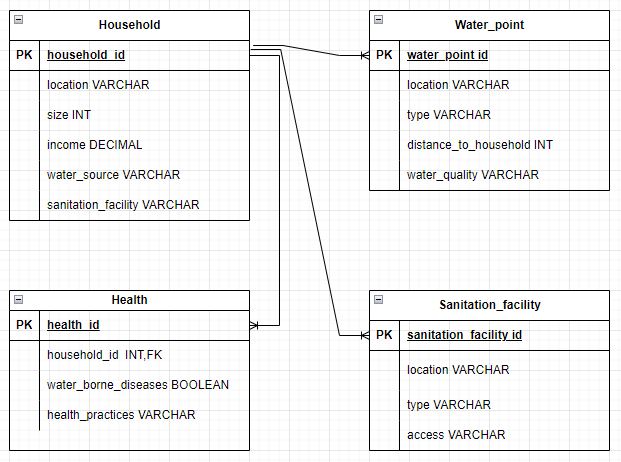
**Project Name:** Addressing Water and Sanitation Disparities in Kenya

**Part 1: SDG Selection and Problem Definition**

SDG 6: Ensure availability and sustainable management of water and sanitation for all.

**Problem:** Disparities in access to clean water and adequate sanitation in Kenya, leading to health issues and hindering socio-economic development.

**Part 2: Database Design (ERD Diagram)**



**SQL statements to create database schema**

CREATE TABLE Household (

household\_id INT PRIMARY KEY,

location VARCHAR(100),

size INT,

income DECIMAL(10,2),

water\_source VARCHAR(50),

sanitation\_facility VARCHAR(50)

);

CREATE TABLE Water\_point (

water\_point\_id INT PRIMARY KEY,

location VARCHAR(100),

type VARCHAR(50),

distance\_to\_household INT,

water\_quality VARCHAR(50)

);

CREATE TABLE Sanitation\_facility (

sanitation\_facility\_id INT PRIMARY KEY,

location VARCHAR(100),

type VARCHAR(50),

access VARCHAR(50)

);

CREATE TABLE Health (

health\_id INT PRIMARY KEY,

household\_id INT,

waterborne\_diseases BOOLEAN,

hygiene\_practices VARCHAR(255),

FOREIGN KEY (household\_id) REFERENCES Household(household\_id)

);

**Populating data to the database**

**Household table**

INSERT INTO household (household\_id, location, size, income, water\_source, sanitation\_facility)

VALUES

(1, 'Nairobi', 5, 30000, 'Tap Water', 'Flush Toilet'),

(2, 'Mombasa', 3, 15000, 'Borehole', 'Pit Latrine'),

(3, 'Kisumu', 8, 25000, 'Lake', 'None'),

(4, 'Nakuru', 6, 28000, 'Borehole', 'Pit Latrine'),

(5, 'Eldoret', 4, 35000, 'Tap Water', 'Flush Toilet'),

(6, 'Kisumu', 7, 22000, 'Lake', 'None'),

(7, 'Mombasa', 5, 18000, 'Borehole', 'Pit Latrine'),

(8, 'Nairobi', 3, 40000, 'Tap Water', 'Flush Toilet'),

(9, 'Malindi', 2, 16000, 'Rainwater', 'Pit Latrine'),

(10, 'Nyeri', 8, 29000, 'Spring', 'Flush Toilet');

**Water\_point table**

INSERT INTO water\_point (water\_point\_id, location, type, distance\_to\_household, water\_quality)

VALUES

(1, 'Nairobi', 'Public Tap', 500, 'Clean'),

(2, 'Mombasa', 'Borehole', 200, 'Slightly Salty'),

(3, 'Kisumu', 'Lake', 100, 'Contaminated'),

(4, 'Nakuru', 'Public Tap', 600, 'Clean'),

(5, 'Eldoret', 'Borehole', 300, 'Slightly Salty'),

(6, 'Kisumu', 'Lake', 150, 'Contaminated'),

(7, 'Mombasa', 'Borehole', 250, 'Slightly Salty'),

(8, 'Nairobi', 'Public Tap', 400, 'Clean'),

(9, 'Malindi', 'Rainwater', 800, 'Salty'),

(10, 'Nyeri', 'Spring', 700, 'Clean');

**Sanitation facility table**

INSERT INTO sanitation\_facility (sanitation\_facility\_id, location, type, access)

VALUES

(1, 'Nairobi', 'Flush Toilet', 'Public'),

(2, 'Mombasa', 'Pit Latrine', 'Private'),

(3, 'Kisumu', 'None', 'None'),

(4, 'Nakuru', 'Pit Latrine', 'Private'),

(5, 'Eldoret', 'Flush Toilet', 'Public'),

(6, 'Kisumu', 'None', 'None'),

(7, 'Mombasa', 'Pit Latrine', 'Private'),

(8, 'Nairobi', 'Flush Toilet', 'Public'),

(9, 'Malindi', 'Pit Latrine', 'Private'),

(10, 'Nyeri', 'Flush Toilet', 'Public');

**Health table**

INSERT INTO health (health\_id, household\_id, waterborne\_diseases, hygiene\_practices)

VALUES

(1, 1, FALSE, 'Good'),

(2, 2, TRUE, 'Fair'),

(3, 3, TRUE, 'Poor'),

(4, 4, TRUE, 'Fair'),

(5, 5, FALSE, 'Good'),

(6, 6, TRUE, 'Poor'),

(7, 7, FALSE, 'Fair'),

(8, 8, FALSE, 'Good'),

(9, 9, TRUE, 'Poor'),

(10, 10, FALSE, 'Good');

***Part 3: SQL Programming***

*Data Retrieval;*

**Count of households by water source**

SELECT water\_source, COUNT(\*) AS number\_of\_households

FROM Household

GROUP BY water\_source;

**Select all households with access to clean water.**

SELECT COUNT(\*) AS households\_with\_clean\_water

FROM Household h

INNER JOIN Water\_point w ON h.location = w.location

WHERE w.water\_quality = 'Clean';

**Number of households with waterborne diseases**

SELECT COUNT(\*) AS households\_with\_waterborne\_diseases

FROM Health

WHERE waterborne\_diseases = TRUE;

**Correlation between water source and waterborne diseases**

SELECT h.water\_source, AVG(s.waterborne\_diseases) AS average\_waterborne\_diseases

FROM Household h

INNER JOIN Health s ON h.household\_id = s.household\_id

GROUP BY h.water\_source;

Data Analysis Examples:

Calculate average distance to water points by region.

SELECT h.location, AVG(w.distance\_to\_household) AS average\_distance

FROM Household h

INNER JOIN Water\_point w ON h.location = w.location

GROUP BY h.location;

**Identify hotspots with high rates of waterborne diseases.**

SELECT h.location, AVG(s.waterborne\_diseases) AS waterborne\_disease\_rate

FROM Household h

INNER JOIN Health s ON h.household\_id = s.household\_id

GROUP BY h.location

ORDER BY waterborne\_disease\_rate DESC;