DATA REPORT
Relational Database (SQL) Data Analytics Project

OKURAME PATRICIA EFEOGHENE

AXIA AFRICA- 13TH MARCH, 2025.

OUTLINE

1. INTRODUCTION
2. INFORMATION OBTAINABLE FROM THE DATASET
3. DATABASE CREATION AND DATASET IMPORT
4. ANSWERS TO KEY QUESTIONS
5. OBSERVATION AND RECOMMENDATION
6. CONCLUSION

1. INTRODUCTION

The primary goal of this case study is to analyze data, identify patterns, and propose informed, data-driven recommendations that governments and stakeholders can implement to effectively improve water access and sanitation in African communities. This dataset consists of 2000 records. This report presents insights from the dataset on water supply, sanitation, and population data of various African communities. The analysis was conducted using MySQL, answering eight key questions and providing recommendations based on the findings.

2. INFORMATION OBTAINABLE FROM THE DATASET

From the dataset, we can extract the following key insights:

1. Water Availability and Population Distribution: This dataset provides insights into the availability of water resources in relation to community populations.

2. Water Source Type and Functionality: This dataset explores the distribution of different water sources (borehole, well, river) and their operational status (functional vs. non-functional water points).

3. Sanitation Infrastructure and Maintenance: This dataset examines the type of sanitation facilities available (toilets vs. latrines) and their associated annual maintenance costs.

4. Support from Governments and NGOs: This dataset assesses the level of governmental and NGO intervention in water and sanitation projects across different communities.

5. Health and Community Well-being: This dataset analyzes the correlation between water access, sanitation conditions, and waterborne disease incidence rates.

6. Community Satisfaction and Accessibility: This dataset provides insights into the satisfaction levels of communities based on water access, sanitation quality, and distance to the nearest water source.

3. DATABASE CREATION AND DATASET IMPORT

- **Creating the database (schema).**
- After opening MySQL Workbench and making sure it was connected to MySQL server, in the navigator panel on the left, I located the "Schemas" section.
- I clicked on the "Create a new schema" icon (a database symbol with a plus sign).
- I entered the schema name; Water Supply Sanitation.
- I double clicked to make it the default schema.
- **Importing the dataset (CSV file) in MySQL.**
- I opened MySQL workbench.
- I selected the database (Water Supply Sanitation).
- I clicked on Table Data Import Wizard.
- I selected the CSV file.
- I clicked next until the import was completed.

4. ANSWERS TO KEY QUESTIONS

```
 1  1. Average water availability (liters per capita per day) for each country.

 2

 3  SELECT Country, AVG(`Water Availability (liters per capita per day)`) AS Average_Water_Availability
 4  FROM water_supply_sanitation_africa
 5  GROUP BY Country;

 6

 7  2. Details of communities where at least one water point is non-functional.

 8

 9  SELECT *
10  FROM water_supply_sanitation_africa
11  WHERE `Number of Non-Functional Water Points` > 0;

12
```

```
13  3. Information for the top five communities with the highest annual sanitation maintenance costs.

14

15  SELECT `Community Name`, Country, `Annual Maintenance Cost (USD)`
16  FROM water_supply_sanitation_africa
17  ORDER BY `Annual Maintenance Cost (USD)` DESC
18  LIMIT 5;

19

20  4. Total number of functional and non-functional water points per country.

21

22  SELECT Country,
23  SUM(`Number of Functional Water Points`) AS Total_Functional_Water_Points,
24  SUM(`Number of Non-Functional Water Points`) AS Total_Non_Functional_Water_Points
25  FROM water_supply_sanitation_africa
26  GROUP BY Country;

27
```

```
28  5. Communities with a high incidence of waterborne diseases (>20%)

29

30  SELECT `Community Name`, Country, `Waterborne Diseases Incidence Rate (%)`
31  FROM water_supply_sanitation_africa
32  WHERE `Waterborne Diseases Incidence Rate (%)` > 20;

33

34  6. Average distance to the water source per region

35

36  SELECT Region, AVG(`Average Distance to Water Source (km)`) AS Average_Distance_km
37  FROM water_supply_sanitation_africa
38  GROUP BY Region;

39
```

```
40 ⊗   7. Communities that receive both government and NGO support
41
42      SELECT `Community Name`, Country
43      FROM water_supply_sanitation_africa
44      WHERE `Government Support` = 'Yes' AND `NGO Support` = 'Yes';
45
46 ⊗   8. Community with the highest population per country
47
48      SELECT Country, `Community Name`, Population
49   ⊖  FROM (
50      SELECT Country, `Community Name`, Population,
51      RANK() OVER (PARTITION BY Country ORDER BY Population DESC) AS rnk
52      FROM water_supply_sanitation_africa
53      ) ranked
54      WHERE rnk = 1;
```

OVERVIEW OF SQL QUERIES USED TO ANSWER KEY BUSINESS QUESTIONS

The queries answer the following questions;

1. Calculate the average water availability (liters per capita per day) for each country.

.

Country	Average_Water_Availability
Zambia	53.61970443349754
Malawi	54.418269230769205
Tanzania	55.205357142857096
Nigeria	58.60197044334978
Rwanda	56.66857142857141
Ethiopia	54.987500000000004
Kenya	54.746086956521715
Uganda	56.138942307692304
Senegal	57.89
Ghana	55.180232558139515

Result 1 ×

Output

Action Output ▼

#	Time	Action	Message
✓	1 08:48:12	SELECT Country, AVG('Water Availability (liters per capita per day)') AS Average_Water_Availa...	10 row(s) returned

AVERAGE WATER AVAILABITY

2. Retrieve details of communities where at least one water point is non-functional.

Community Name	Population	Water Source Type	Water Availability (liters per capita per day)	Number of Functional Water Points	
Northern Community 2	14820	Borehole	37.8	33	
Dar Community 3	14128	Well	82.9	28	
Eastern Community 5	14147	Borehole	35.7	12	
Northern Community 6	14944	Borehole	91.3	26	
Eastern Community 7	14456	Well	38.1	18	
Northern Community 8	7467	Well	51.2	10	
Northern Community 9	4851	Borehole	49.7	19	
Eastern Community 10	12970	Borehole	74.1	24	

water_supply_sanitation_africa 2 ×

Output

Action Output

#	Time	Action	Message
✓ 1	08:48:12	SELECT Country, AVG('Water Availability (liters per capita per day)') AS Average_Water_Availa...	10 row(s) returned
✓ 2	08:54:12	SELECT * FROM water_supply_sanitation_africa WHERE `Number of Non-Functional Water P...	1000 row(s) returned

OVERVIEW OF COMMUNITIES WITH AT LEAST ONE NON-FUNCTIONAL WATER POINT

3. Retrieve the information for the top five communities with the highest annual sanitation maintenance costs.

Community Name	Country	Annual Maintenance Cost (USD)
Dar Community 329	Tanzania	49959
Eastern Community 1658	Nigeria	49957
Eastern Community 1021	Nigeria	49949
Western Community 1259	Uganda	49904
Northern Community 638	Ghana	49894

COMMUNITIES WITH HIGHEST ANNUAL SANITATION MAINTENANCE COST (USD)

4. Calculate the total number of functional and non-functional water points per country.

Country	Total_Functional_Water_Points	Total_Non_Functional_Water_Points
Zambia	5253	1020
Malawi	5172	1029
Tanzania	4354	879
Nigeria	5271	985
Rwanda	5491	966
Ethiopia	5296	1088
Kenya	6049	1116
Uganda	5287	959
Senegal	5039	946
Ghana	4371	886

TOTAL NUMBER OF FUNCTIONAL AND NON-FUNCTIONAL POINTS PER COUNTRY

5. Identify communities with a high incidence of waterborne diseases (>20%)

Community Name	Country	Waterborne Diseases Incidence Rate (%)
Southern Community 1	Zambia	21.8
Eastern Community 4	Nigeria	28.1
Northern Community 6	Malawi	25
Eastern Community 11	Ethiopia	22.4
Eastern Community 12	Nigeria	28.7
Southern Community 13	Zambia	23.2
Western Community 18	Uganda	27.2
Western Community 21	Kenya	24.3
Eastern Community 23	Nigeria	23.2

OVERVIEW OF COMMUNITIES WITH A HIGH INCIDENCE OF WATERBORNE DISEASES (>20%)

6. Find the average distance to the water source per region.

Region	Average_Distance_km
Southern Zambia	5.221674876847291
Northern Malawi	5.414423076923076
Dar es Salaam	5.15059523809524
Eastern Nigeria	5.477832512315274
Eastern Rwanda	5.208571428571427
Eastern Ethiopia	5.00480769230769
Western Kenya	5.539565217391304
Western Uganda	5.0375
Central Senegal	5.255789473684208
Northern Ghana	5.187209302325583

AVERAGE DISTANCE TO THE WATER SOURCE PER REGION

7. List the communities that receive both government and NGO support.

Community Name	Country
Northern Community 2	Malawi
Northern Community 9	Malawi
Eastern Community 10	Rwanda
Western Community 19	Uganda
Eastern Community 23	Nigeria
Eastern Community 24	Ethiopia
Dar Community 29	Tanzania
Eastern Community 30	Nigeria
Eastern Community 31	Nigeria
Southern Community 33	Zambia

COMMUNITIES THAT RECEIVE BOTH GOVERNMENT AND NGO SUPPORT

8. Identify the community with the highest population per country.

Country	Community Name	Population
Ethiopia	Eastern Community 1877	14993
Ghana	Northern Community 1393	14992
Kenya	Western Community 1036	14976
Malawi	Northern Community 6	14944
Nigeria	Eastern Community 932	14997
Rwanda	Eastern Community 1440	14890
Senegal	Central Community 1395	14886
Tanzania	Dar Community 1199	14888
Uganda	Western Community 121	14948
Zambia	Southern Community 748	14952

COMMUNITY WITH THE HIGHEST POPULATION PER COUNTRY

5. OBSERVATIONS AND RECOMMENDATIONS

- **Observation:**

Some countries have significantly lower average water availability per capita, indicating potential water scarcity.

- **Recommendation:**

Improve water distribution systems – ensure water reaches all communities, especially rural areas. Implement water rationing & efficiency programs – reduce waste and promote conservation.

- **Observation:**

Of the five communities that have the highest annual sanitation maintenance costs, we see that two communities; Northen Community 638- Ghana and Eastern Community 1021- Nigeria have a high incidence of waterborne diseases (>20%).

- **Recommendation:**

Those allocating the funds i.e. government or NGOs should ensure proper monitoring of the funds and that they are properly channeled to the right resources that they are meant for. This would also prevent the embezzlement of funds.

- **Observation:**

Some communities receive both government and NGO support, while others rely on just one or none at all. The distribution of support is not uniform across all regions.

- **Recommendation:**

Expand Government & NGO Collaboration – Develop joint programs to ensure wider coverage.
Prioritize Unserved Communities – Direct resources to areas lacking both support types.
Monitor and Evaluate Impact – Regularly assess the effectiveness of aid and adjust funding.

- **Observation:**

Some communities lack both government and NGO support, putting them in the greatest danger. Many of these areas coincide with low water availability, worsening the situation.

- **Recommendation:**

Launch Special Aid Programs for High-Risk Areas – Governments and NGOs must target these communities first.
Encourage Private Sector Involvement – Companies and philanthropists can contribute to water access projects.
Set Up Monitoring Systems – Ensure no community is left out of future initiatives.

6. CONCLUSION:

Continuous monitoring and data-driven policies will help improve long-term water sustainability. By implementing these recommendations, millions of people in Africa can gain better access to clean water and sanitation, improving overall public health and quality of life.

LINK TO MySQL SCRIPT:
MySQL queries