MORINGA SCHOOL

Tanzania Water Point Classification System

tom.mwambire@student.moringaschool.com

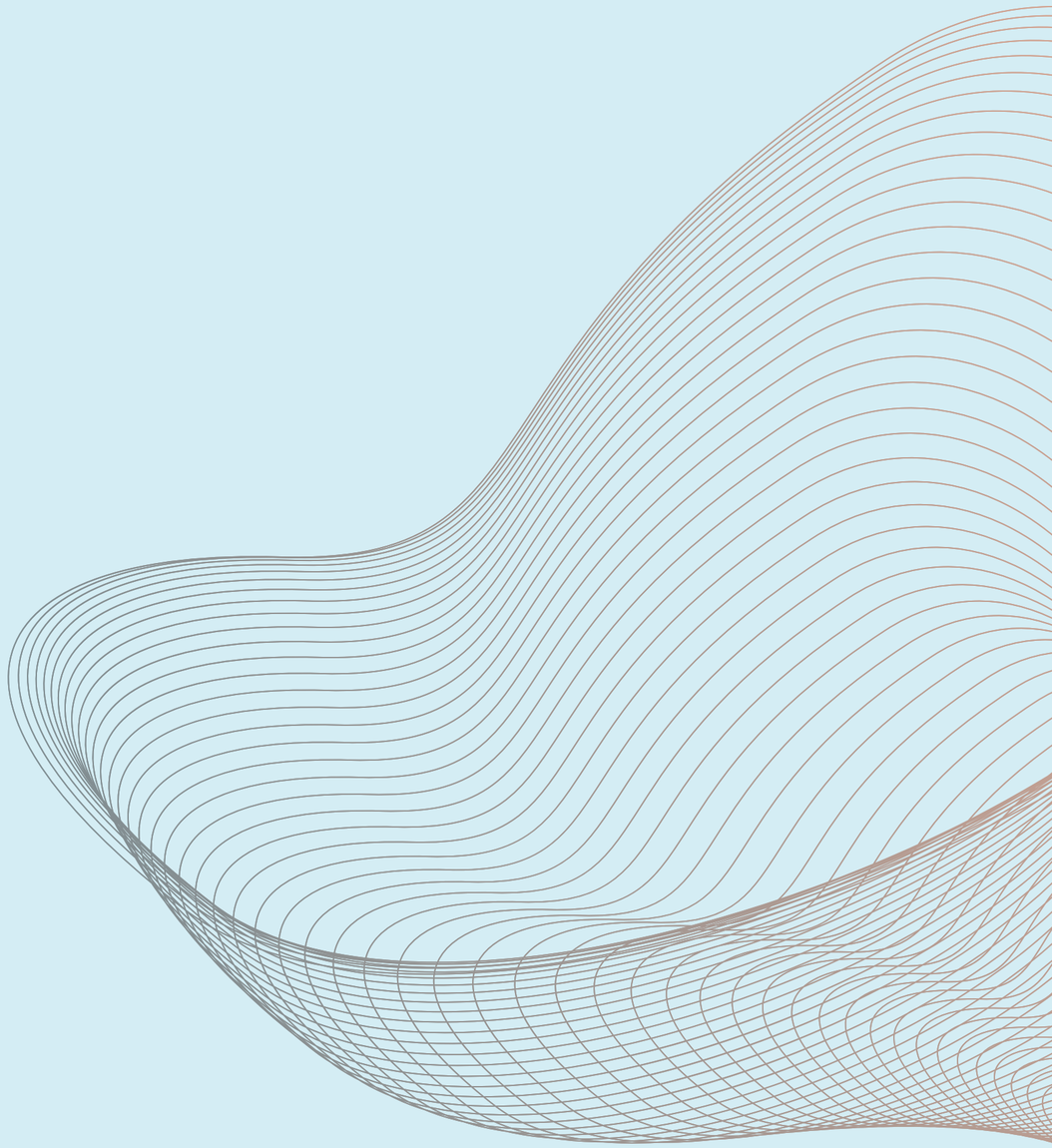


TABLE OF CONTENT

The following are the key topics in the presentation.

INTRODUCTION

Access to water and sanitation are recognized as human rights by the United Nations. The right to water entitles everyone to access sufficient, safe, acceptable, physically accessible, and affordable water for personal and domestic use.

In Tanzania especially, 88% percent of the population lacks access to safe water and this greatly affects the quality of life of the people. The problem is made considerably worse by the lack by the number of nonfunctioning water points. Due to the large number of water points in the country, it is a challenge to figure out which ones are not functioning but it doesn’t have to be.

Using data from the Tanzania Ministry of Water, we can predict the functioning status of water points efficiently and accurately

PROBLEM STATEMENT

Making accurate and efficient predictions about the functioning status of water points throughout the country is a huge challenge that Dutch Water Limited face. This can be mitigated by creating a system that analyses the data of the various details about the water points it has and learns from it and uses the information it has learned to predict the functioning status of other water points

TARGET AUDIENCE

The target is Dutch Water Limited (DWL), a world-leading organization that produces and sells affordable, healthy, and purified water to all people. They have looked at current United Nations studies and have visited the country (Tanzania) and they feel like they can make a considerable effort in alleviating some of the water accessibility crisis felt in Tanzania.

They have decided to help the government and its citizens by repairing the waterpoints throughout the country and this system is crucial to them because it helps them find out the statuses of water points in the country

OBJECTIVES

Main Objective

To analyze the various data that have been made available to them by the Ministry of Water in Tanzania and come up with a machine learning model that can predict the functioning status of a waterpoint given certain factors.

Specific Objective

1

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer vulputate vel ipsum ac fringilla. Nunc cursus, arcu nec pretium aliquet.

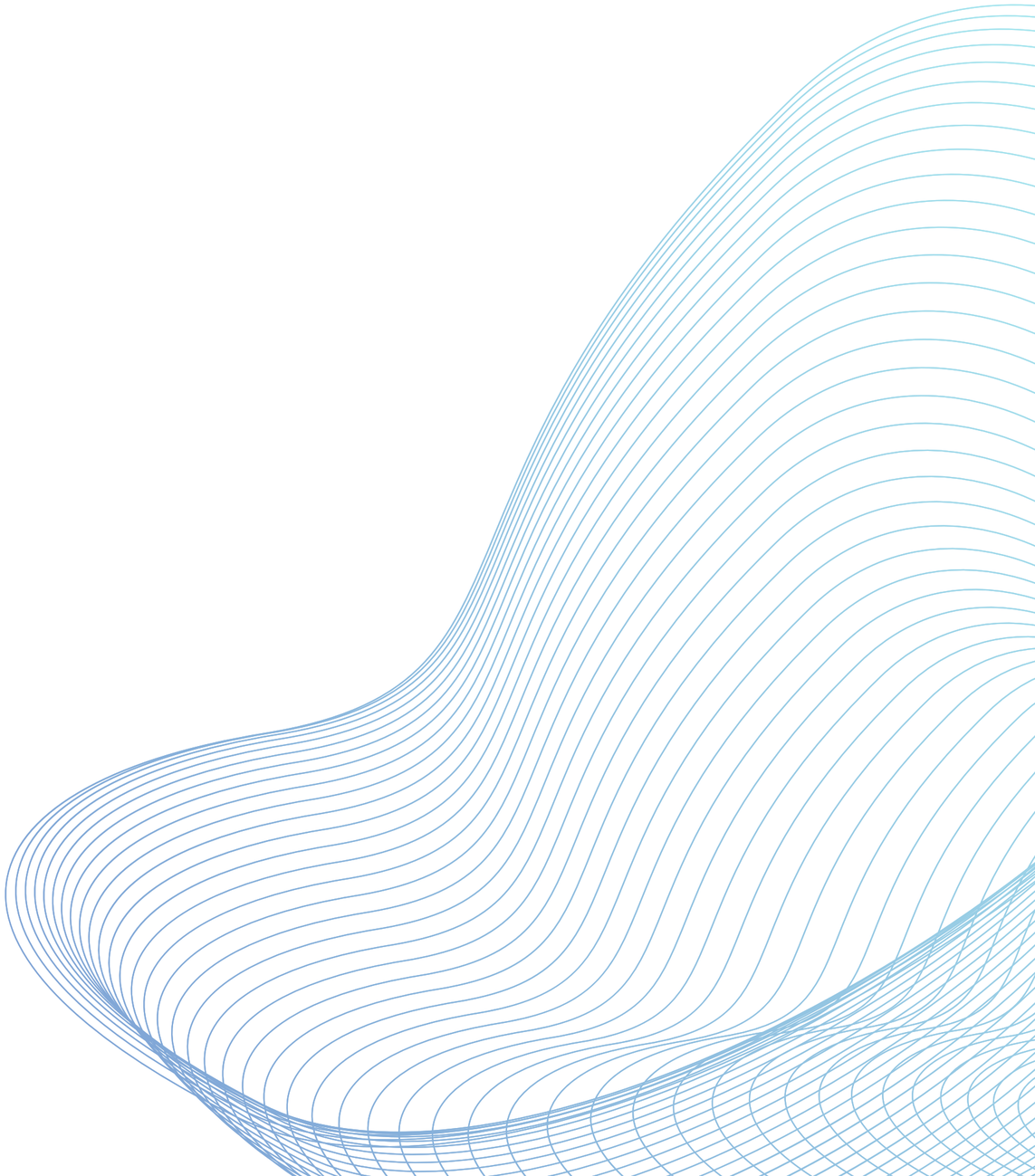
2

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer vulputate vel ipsum ac fringilla. Nunc cursus, arcu nec pretium aliquet.

3

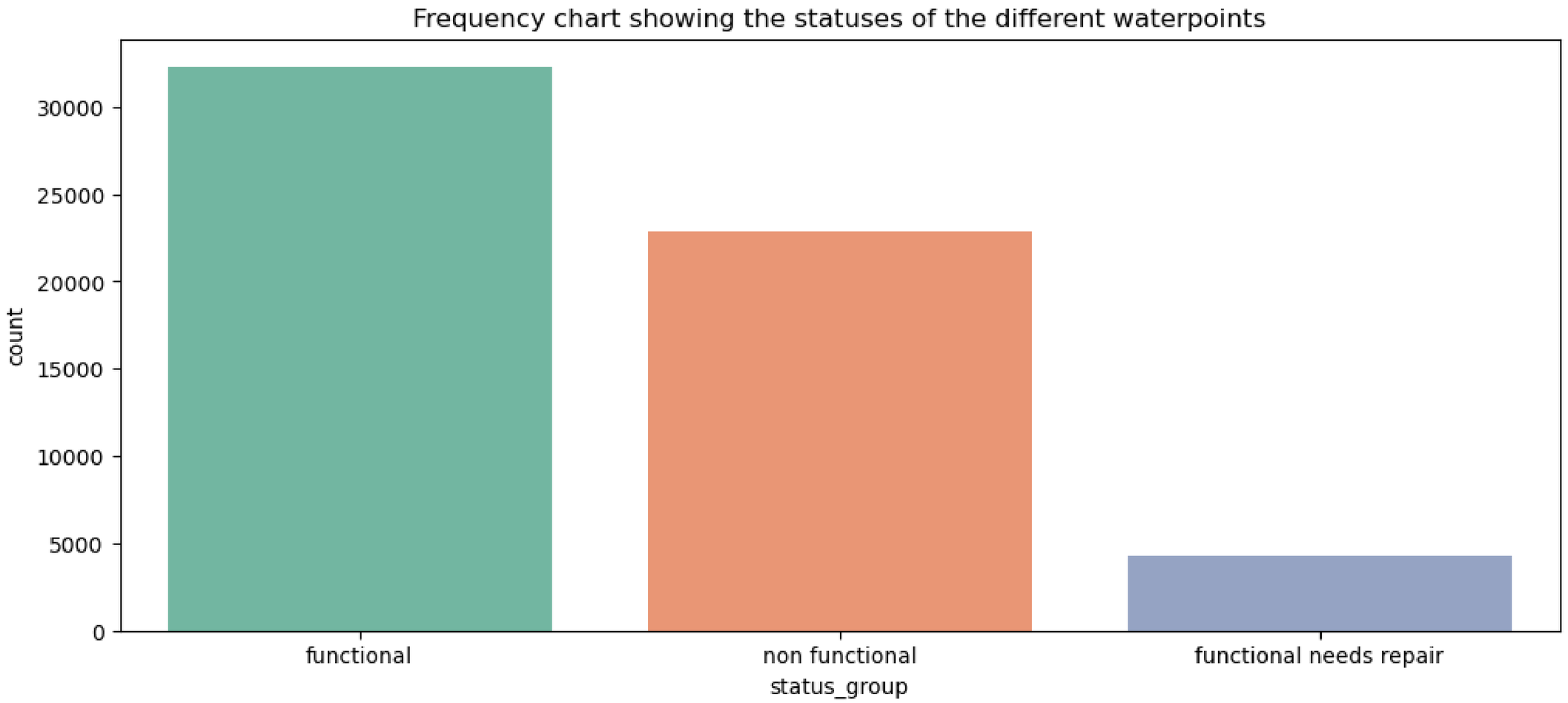
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer vulputate vel ipsum ac fringilla. Nunc cursus, arcu nec pretium aliquet.

PATTERNS AND FINDINGS



OPERATIONAL STATUSES OF WATER POINTS

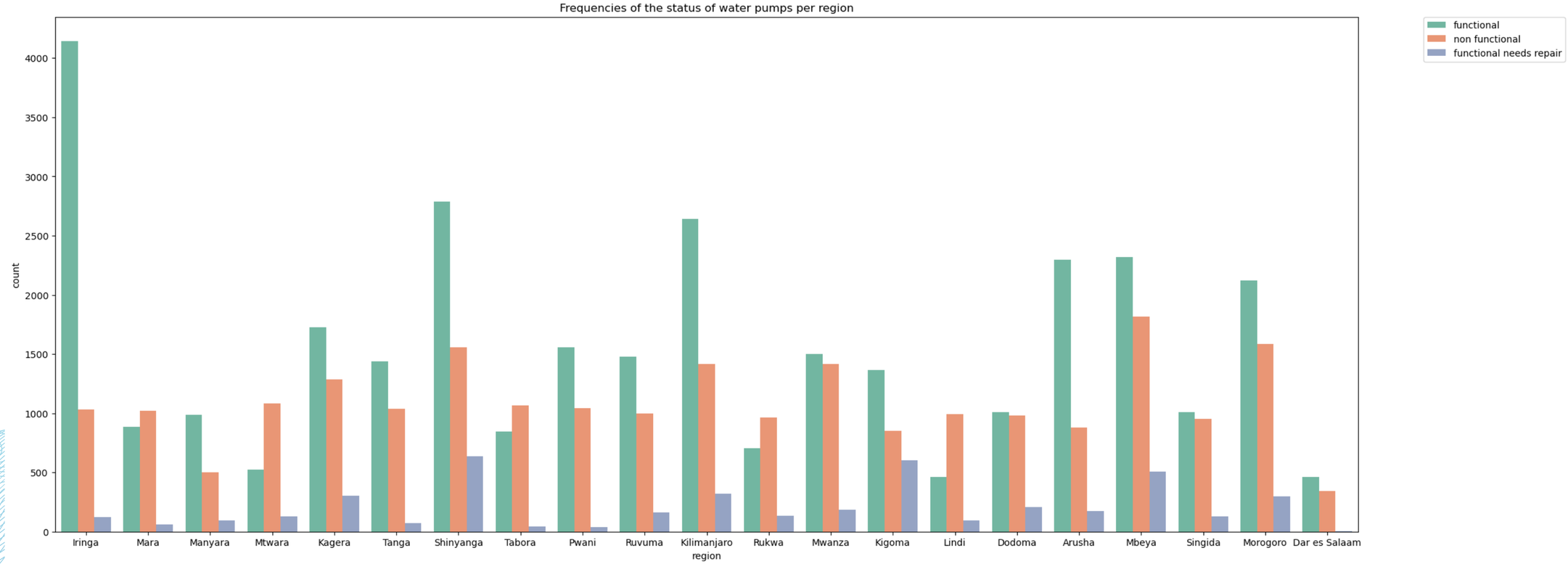
Majority of the water points are functional closely followed by non functional and then lastly functional needs repair



REGIONAL DISTRIBUTION OF WATER POINTS PER STATUS

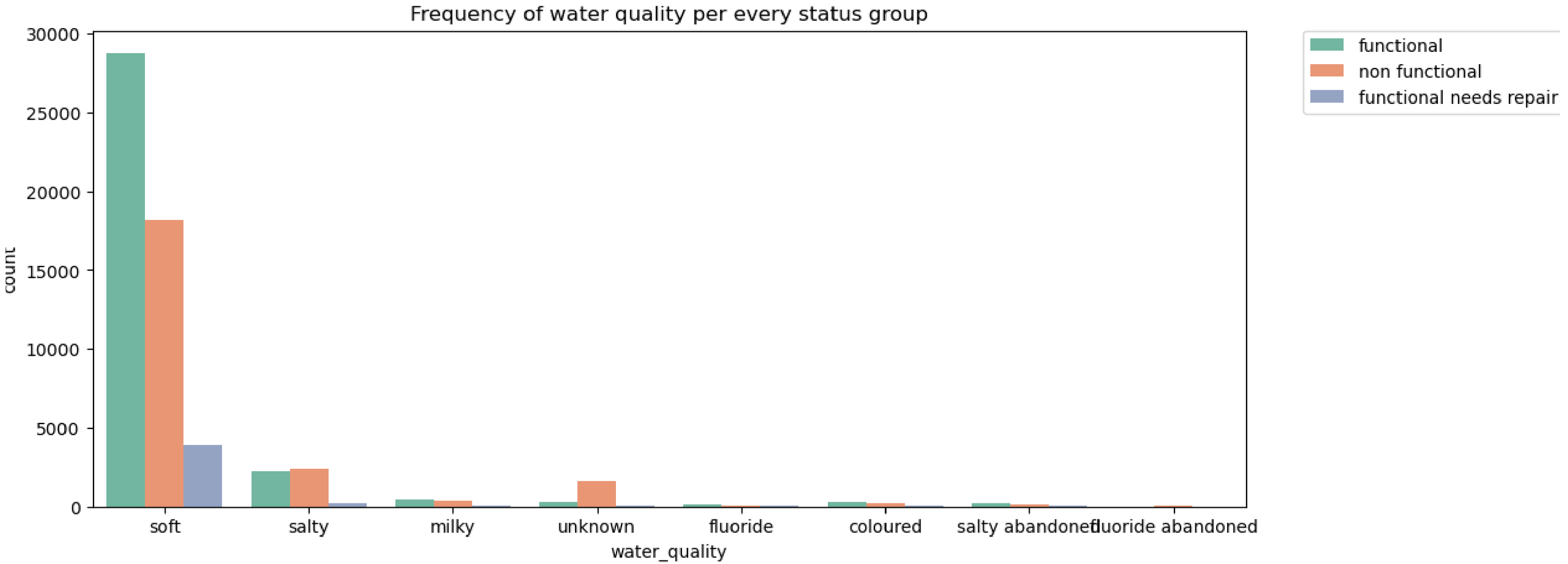
the regions of Mara, Mtwara, Tabora, Rukwa and Lindi, there may be a water accessibility problem.

This is because the non functional water points are more than the functional ones



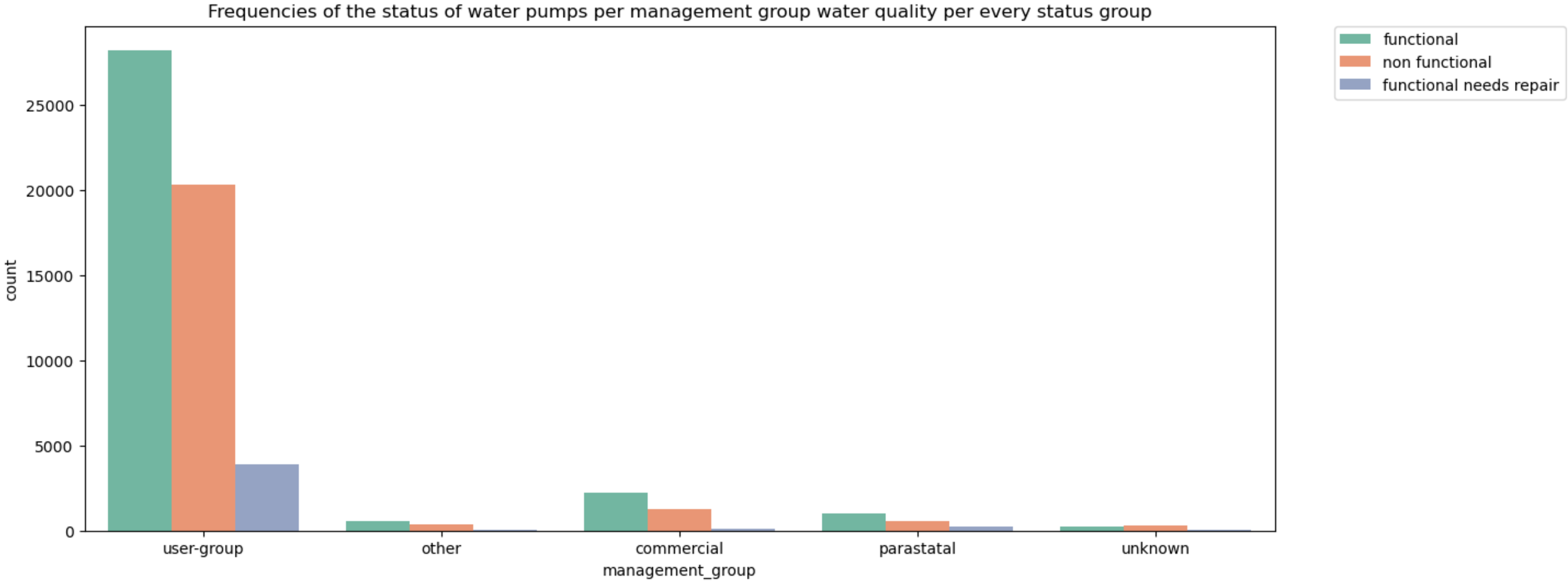
FREQUENCY OF WATER QUALITY AND OPERATIONAL STATUS

The majority of the water points produce soft water, and the second most common water produced is salty



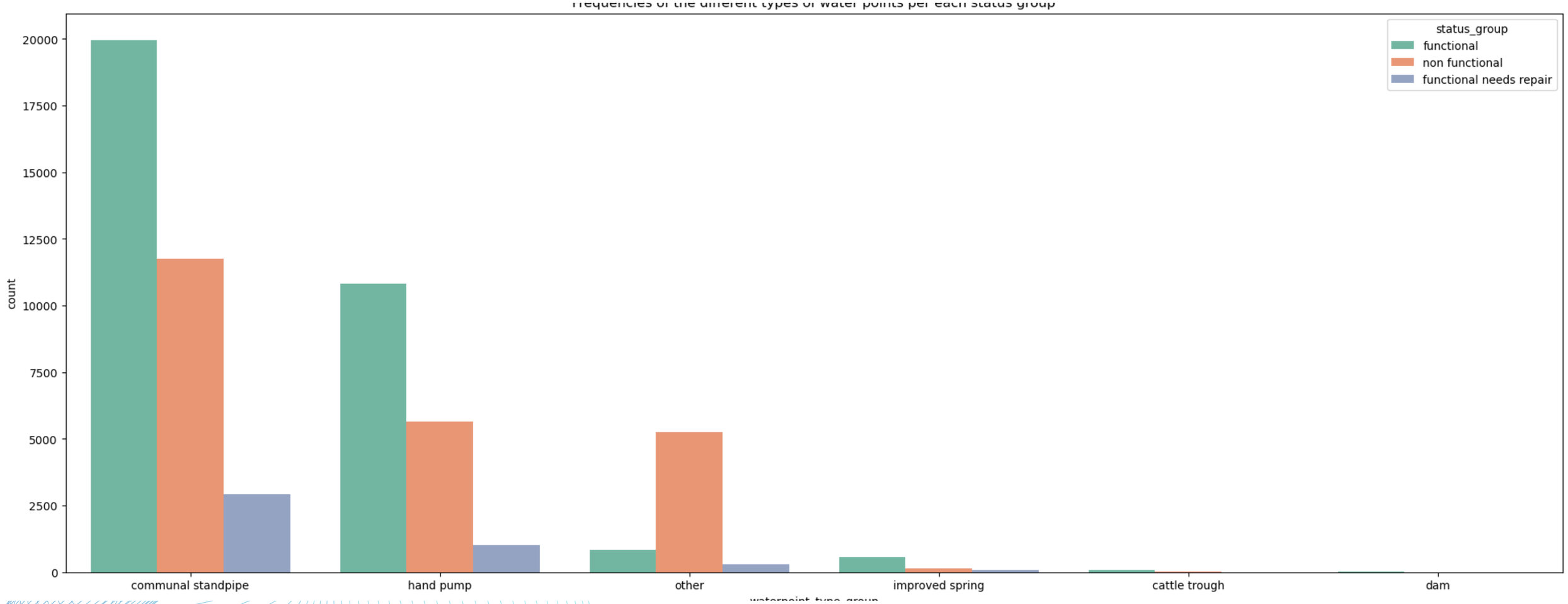
FREQUENCY OF MANAGEMENT AND OPERATIONAL STATUS

An overwhelming majority of the water points are managed by communal user groups.



FREQUENCY OF WATER POINT TYPE AND OPERATIONAL STATUS

From the image we can tell that most common type of water point is communal standpipe and the second is hand pump.



MODELLING

MODEL SELECTION

The following models were thoroughly tested to determine the best for our current scenario:

- Decision Tree Classifier
- K Nearest Neighbors
- Random Forest
- XGBoost Classifiers
- AdaBoost Classifier

We compared various metrics such as accuracy in training and testing, recall, precision and f1 score and XGBoost Classifier performed better than the rest in all metrics

After selecting XGBoost Classifier, we proceeded to fine tune the performance. This involves tuning its operational parameters.

We tried out different combination of parameters until found the perfect one. We used a method called GridSearch and this took a while to run but gave us the best operating parameters.

Once we identified the model and tuned it to its optimum set of parameters, we proceeded to test it and it produced an overall accuracy percentage of 85

CONCLUSION

Some of our takeaways from the project are:

- This project has met its main objective of creating a system that can predict the functioning status of a water point. This was achieved by using a tuned XGBoost Classifier model to ensure accurate and optimum results.
- The project has brought to light key factors such as the water quality, the management type, and the water point type that play a huge role in determining the functioning status of a water point.
- The project has brought to light regions such as Mara, Mtwara, Tabora, Rukwa, and Lindi which have the biggest water accessibility problems. This is due to the number of non-functioning water points being more than the functioning water points.
- For our prediction problem, the XGBoost Classifier algorithm proved to be the best since it had the highest accuracy and f1 score.
- The final model after hyperparameter tuning produced results with an accuracy level of 85%

RECOMMENDATIONS

- In order to improve the accuracy of the system we recommend acquiring more complete and accurate data. This will enable for better training of the model which would lead to higher accuracy level than the 85%
- DWL should place greater attention and emphasis in their efforts to repair water points in the regions of Mara, Mtwara, Tabora, Rukwa and Lindi. This is because they have more non functioning water points than functioning ones.
- For this massive exercise, DWL should consider using submersible ground water pumps since they are made with stainless steel and can handle a wide variety of conditions. They should get the best of these from either Honda or Grundfos
- DWL should invest in a community program to teach the community living near the water point on suitable ways they can maintain and better take care of the waterpoints.

THANK YOU