Analyzing the water pumps' functionality in wells of Tanzania

By: Esther Nyawera

INTRODUCTION:

Tanzania's has a population of 57+ million with majority facing water shortage. This project aims to **build a classifier model** leveraging pump types, water source, quality, and quantity data.

- Its aim is to aid **NGOs and the Tanzanian government** as the stakeholders in pinpointing well functionality status and those in need of repair and maintenance. Ultimately improving water accessibility, especially in remote or underprivileged regions that experience water scarcity.
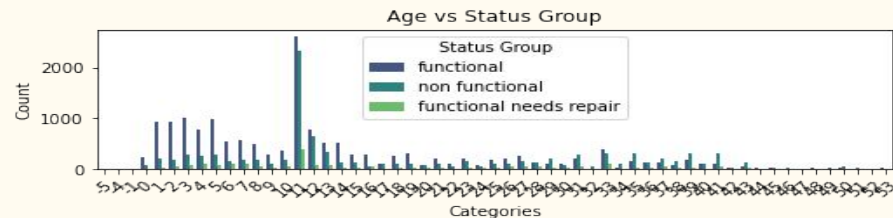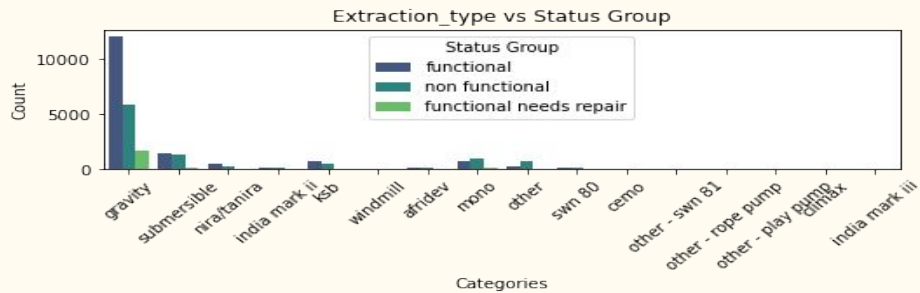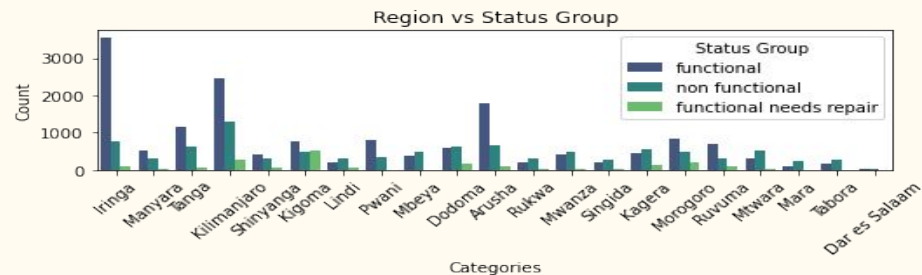
Pump functionality

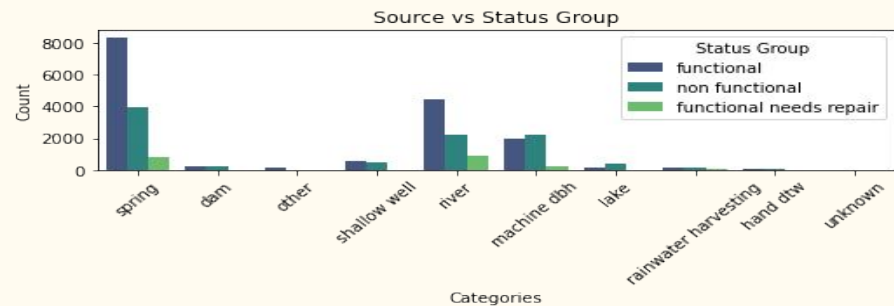Class Distribution of Tanzania Wells
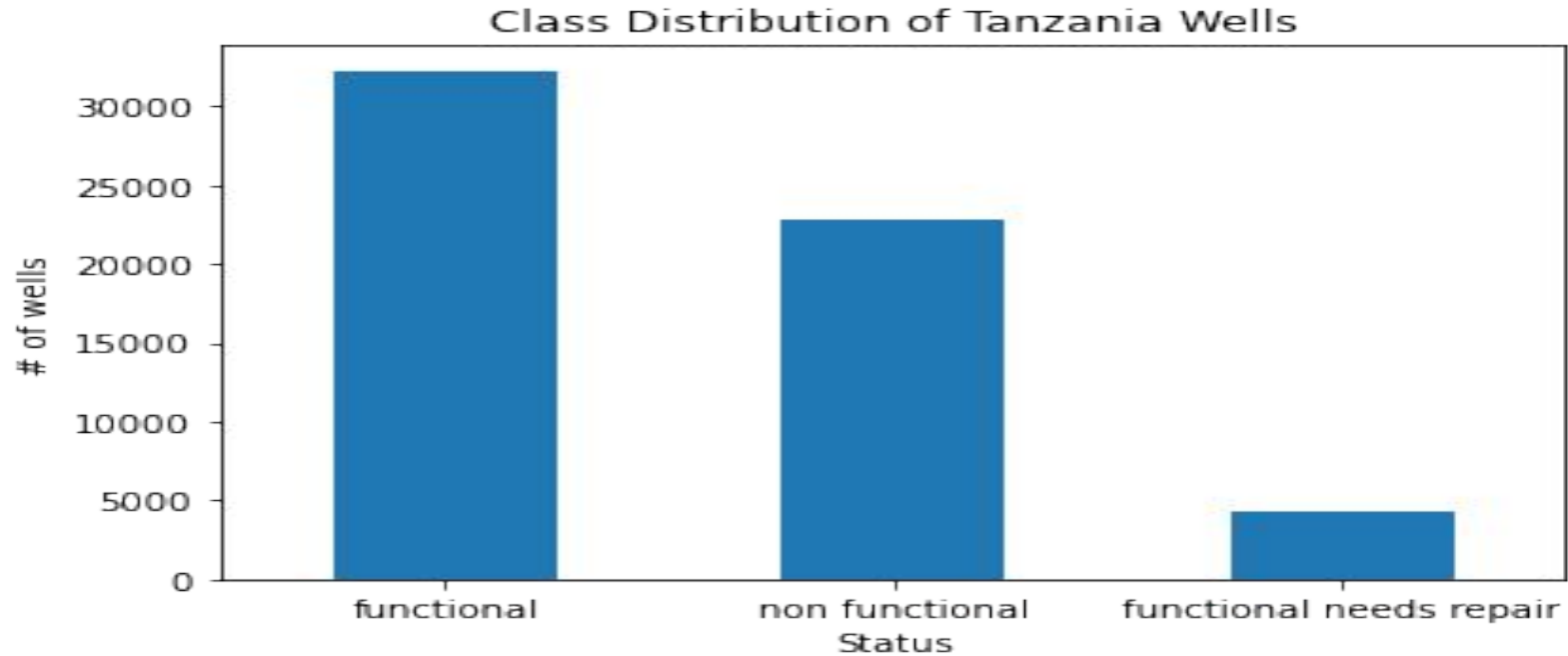
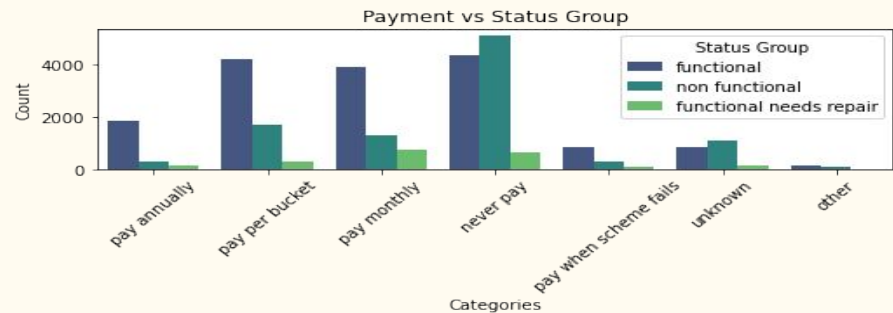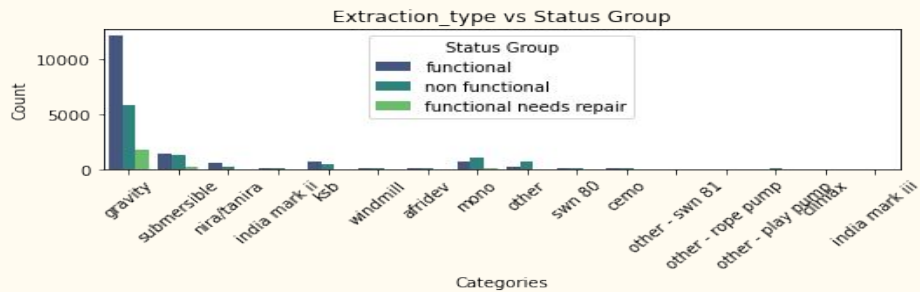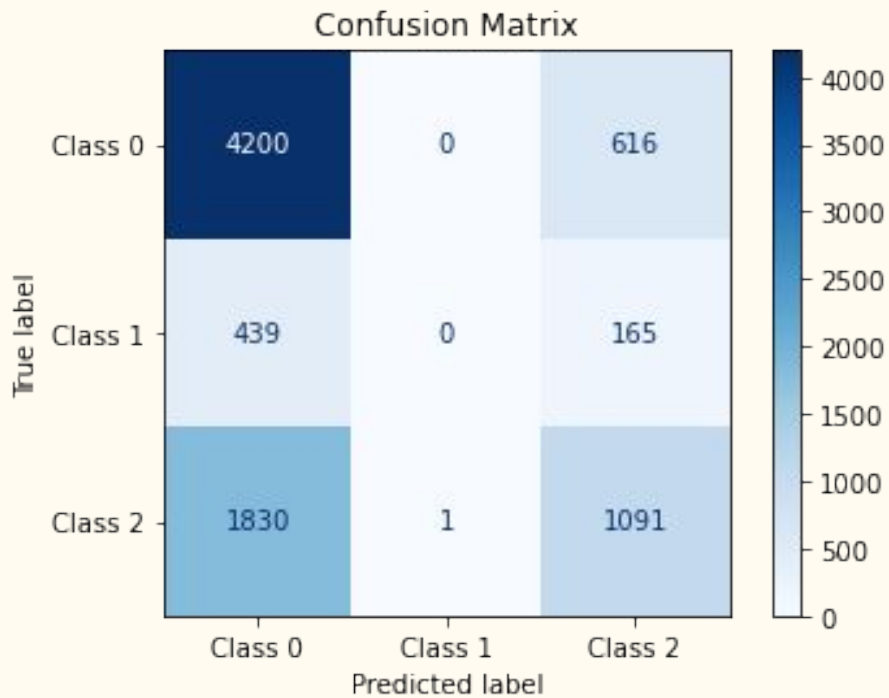# Features

Found to affect functionality

- Region
- Payments
- Source
- Water quality
- Etc.

Visualization

Base model

Confusion Matrix

Accuracy: 63%

Precision: 58%

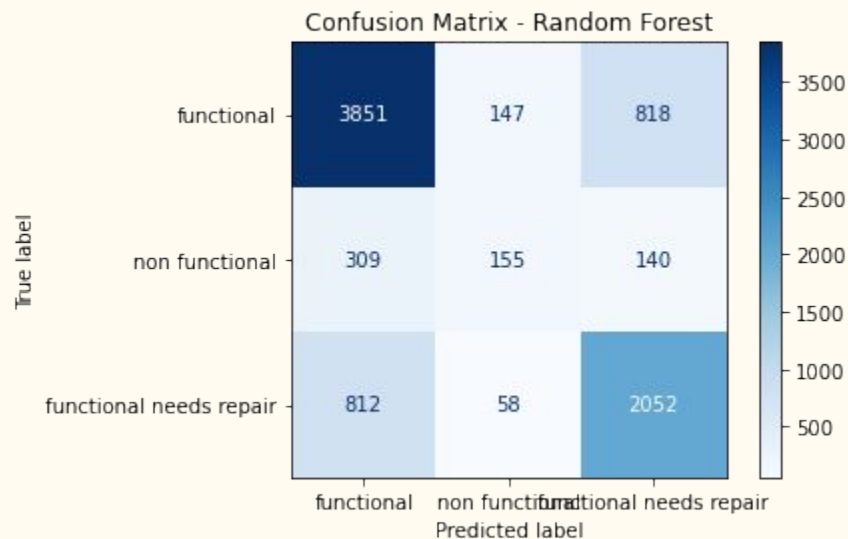Recall: 63%

Data Preprocessing:

Clean,organize and transform the data while still maintaining the important features so as to enhance the accuracy and performance of different machine learning models till we get the best results.
The best model can then be used in future to help predict functionality.

Best model

Confusion Matrix - Random Forest

Accuracy: 65%

Precision: 72%

Recall: 65%

Conclusion:

-Most of the functional wells are found in the city where people mostly pay to use them.

-Old wells are the most are non functional and in need of repairs this is to show they have been neglected.

-Wells that run on soft water sources the most functional.

RECOMMENDATION

- Build more wells resembling those in urban. This replication strategy might enhance the functionality of wells in other areas

- Increase attention and maintenance of older wells to prevent deterioration and improve their functionality.

R

Next steps:

1. Further exploration into wider scopes of data and analysis of different features on how they affect functionality
2. Develop a strong strategy on well maintenance and repairs
3. Implement recommendations

Thank you.

Any questions?