TANZANIA WATER WELL ANALYSIS

BY : ERIC W. NJERU
TM: SAMUEL KARU

Tanzania's Profile

Population 67,438,106(2023)

39% of Households lack safe water

Table of Contents

Project Overview

- The project is aimed to develop a predictive classifier for assessing the condition of water wells in Tanzania to address the issue of limited access to clean water. This project serves the dual purpose of assisting NGOs in locating wells requiring repair and providing insights to the Tanzanian government for more effective well construction.

Business Problem

- There is inadequate access to clean water in Tanzania due to the poor condition of water wells. There is need for a predictive solutions to identify wells in need of repair and influence future well construction.

Objectives

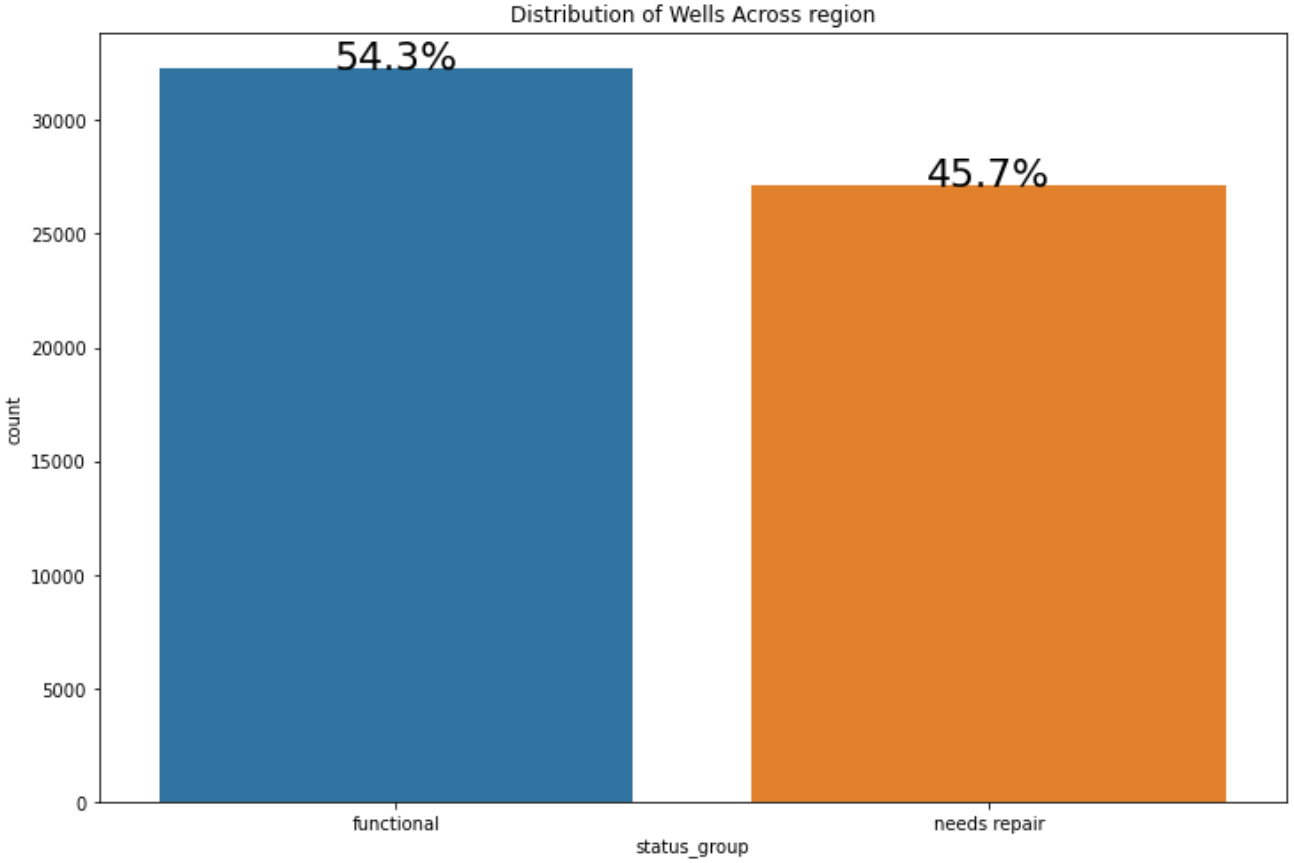
- Predict the state of Tanzanian water wells to help NGOs locate repair-worthy wells.

- Assist the Tanzanian government in improving future well construction.

Data Understanding..

- The data was collected from a sample of 59,400 wells distributed across Tanzania
- Features utilized included : construction_year, basin, region, extraction_type_class, management_group, payment, water_quality, quantity_group, source_type, source_class and waterpoint_type

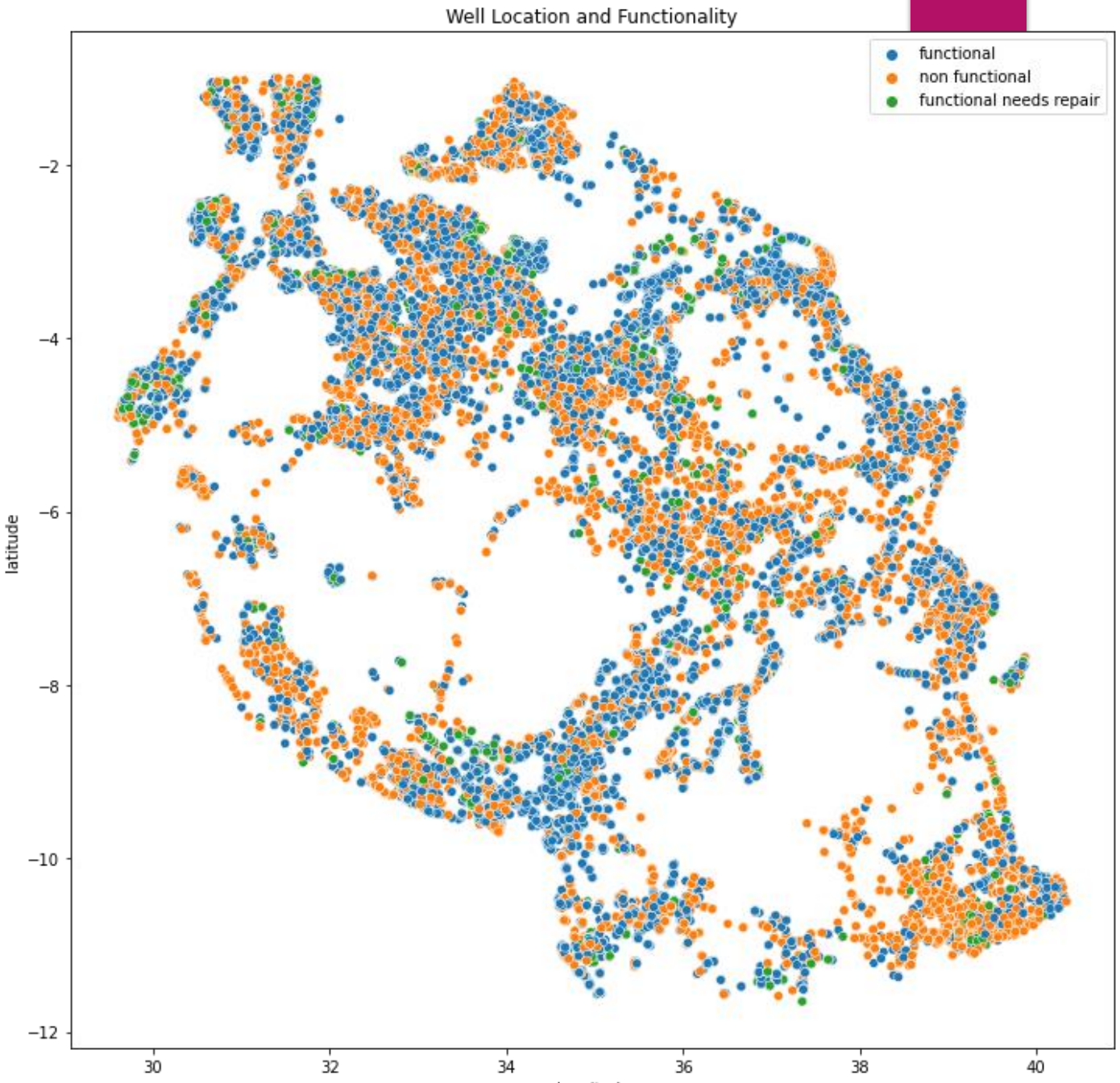
Dependent Variable



- The status_group was engineered to functional wells and those that needed repairs

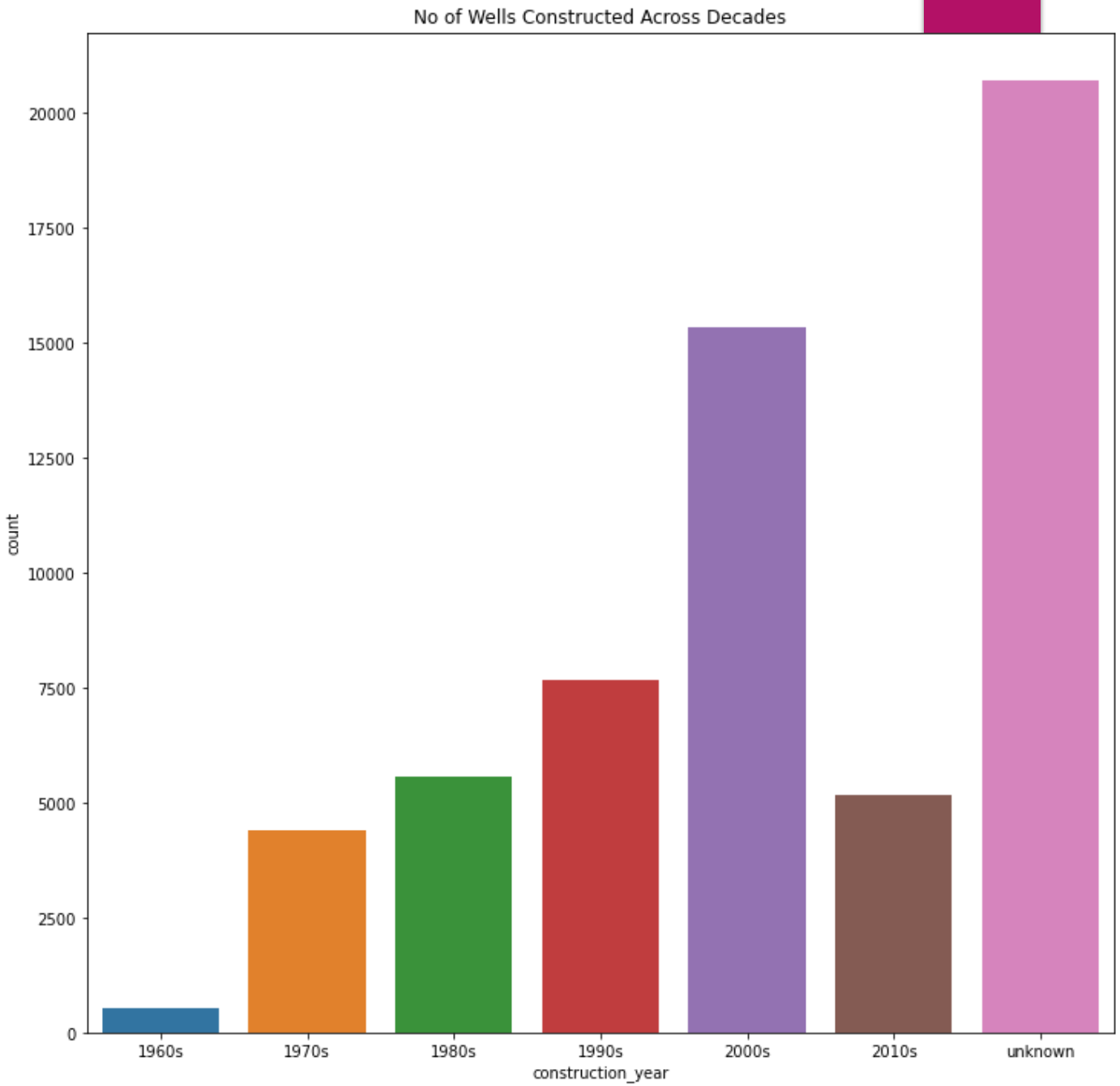
Well Location and Functionality

The wells locations are as per the map. Majority of the wells are located at the upper half of the map



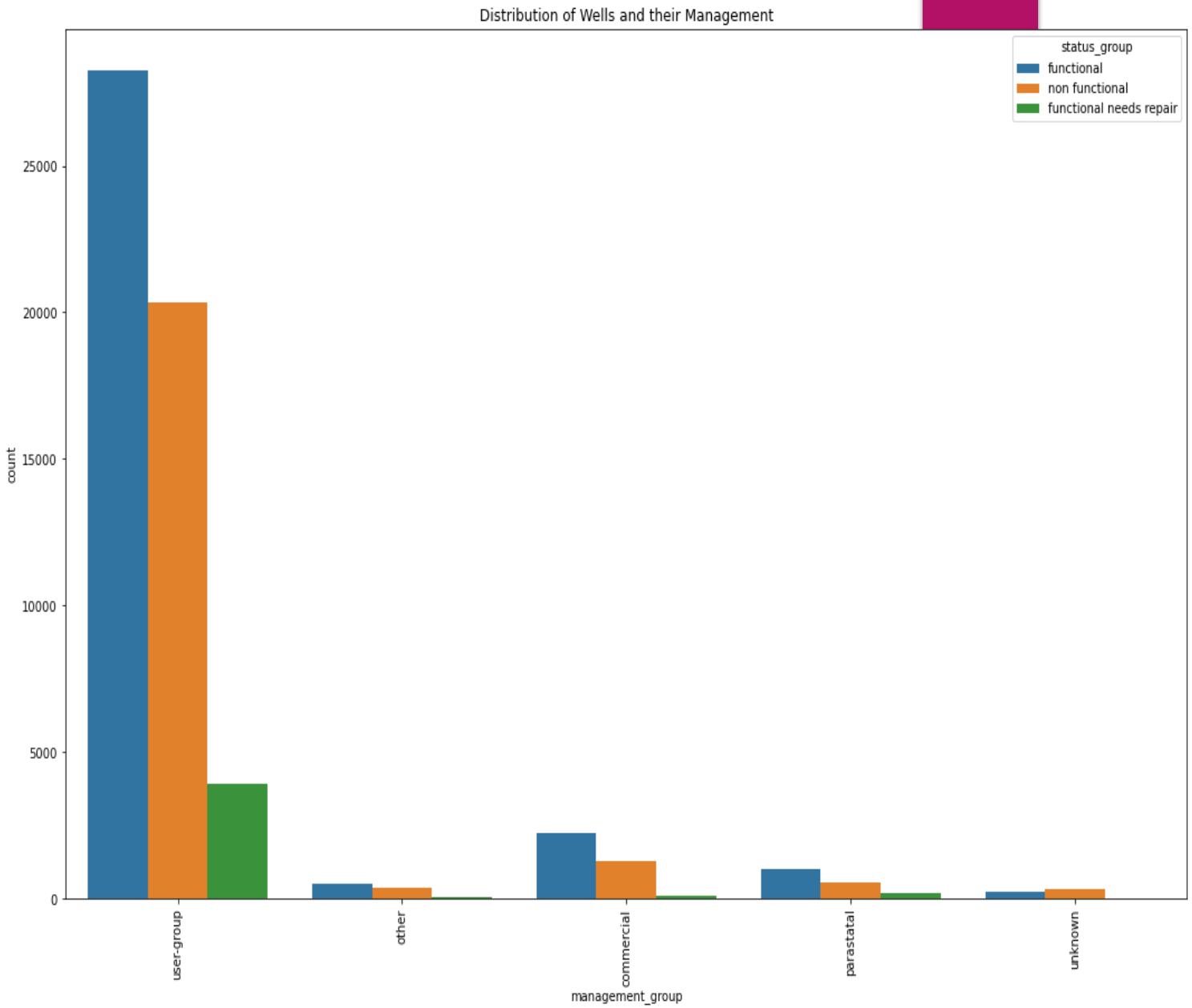
Year of Construction

Most of the wells had no known year of construction. Majority were constructed in the years between 2000-2009, with the least being constructed between 1960-1969.



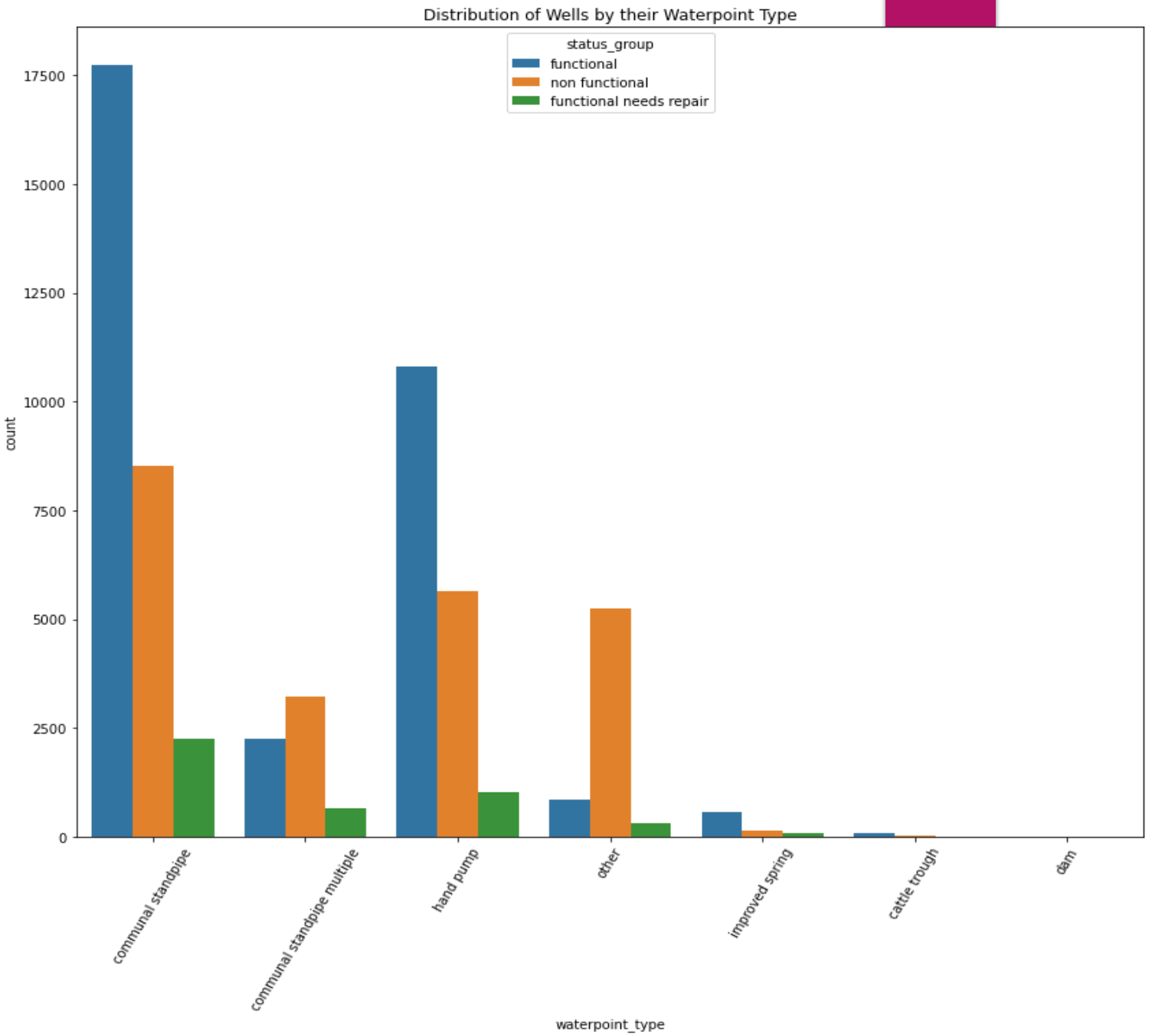
Management

Majority of the wells were user-group managed and they had the largest numbers of functional wells



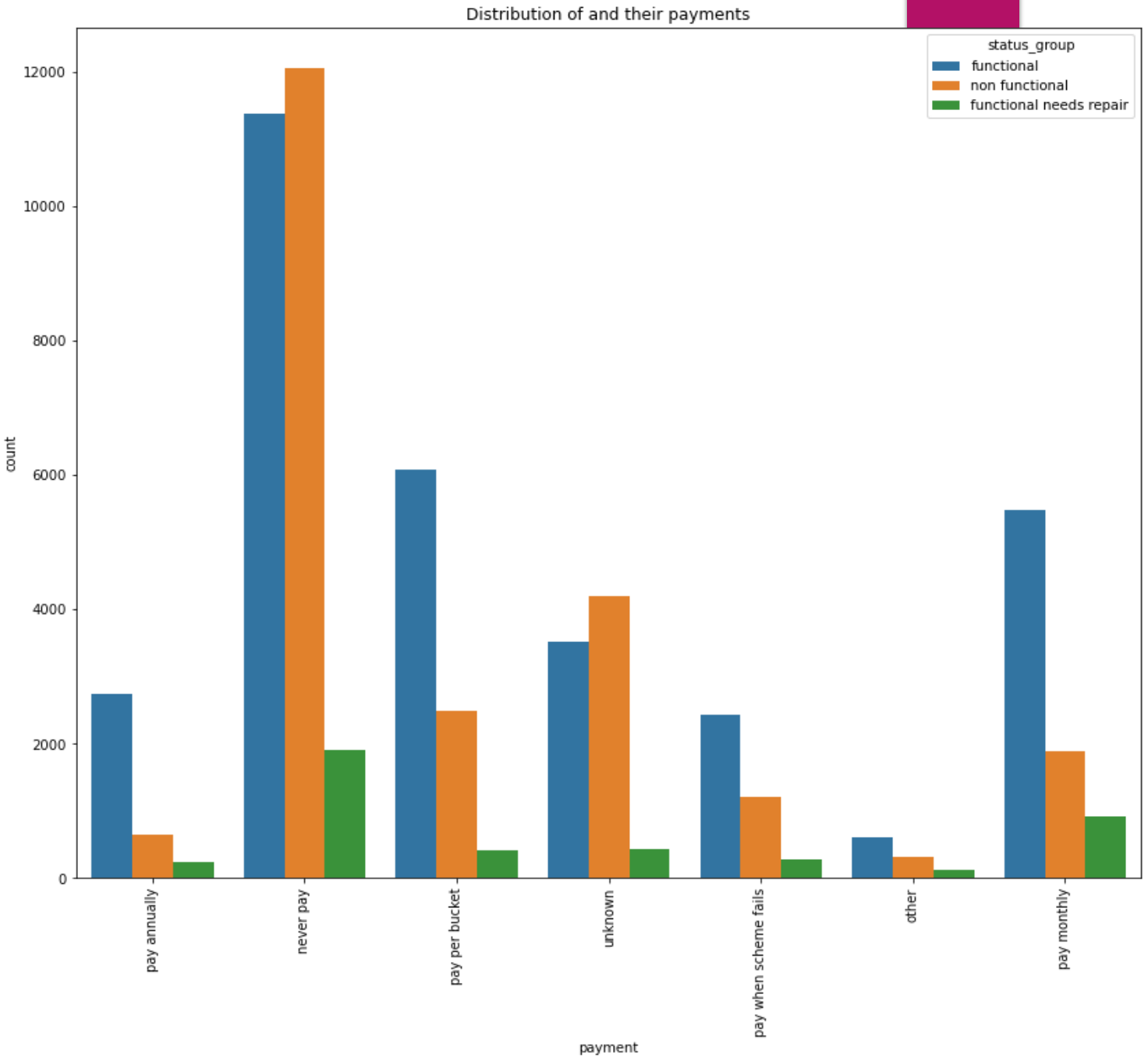
Water Point Type

Majority of the water points were communal standpipe followed by hand pumps with the least being dams.



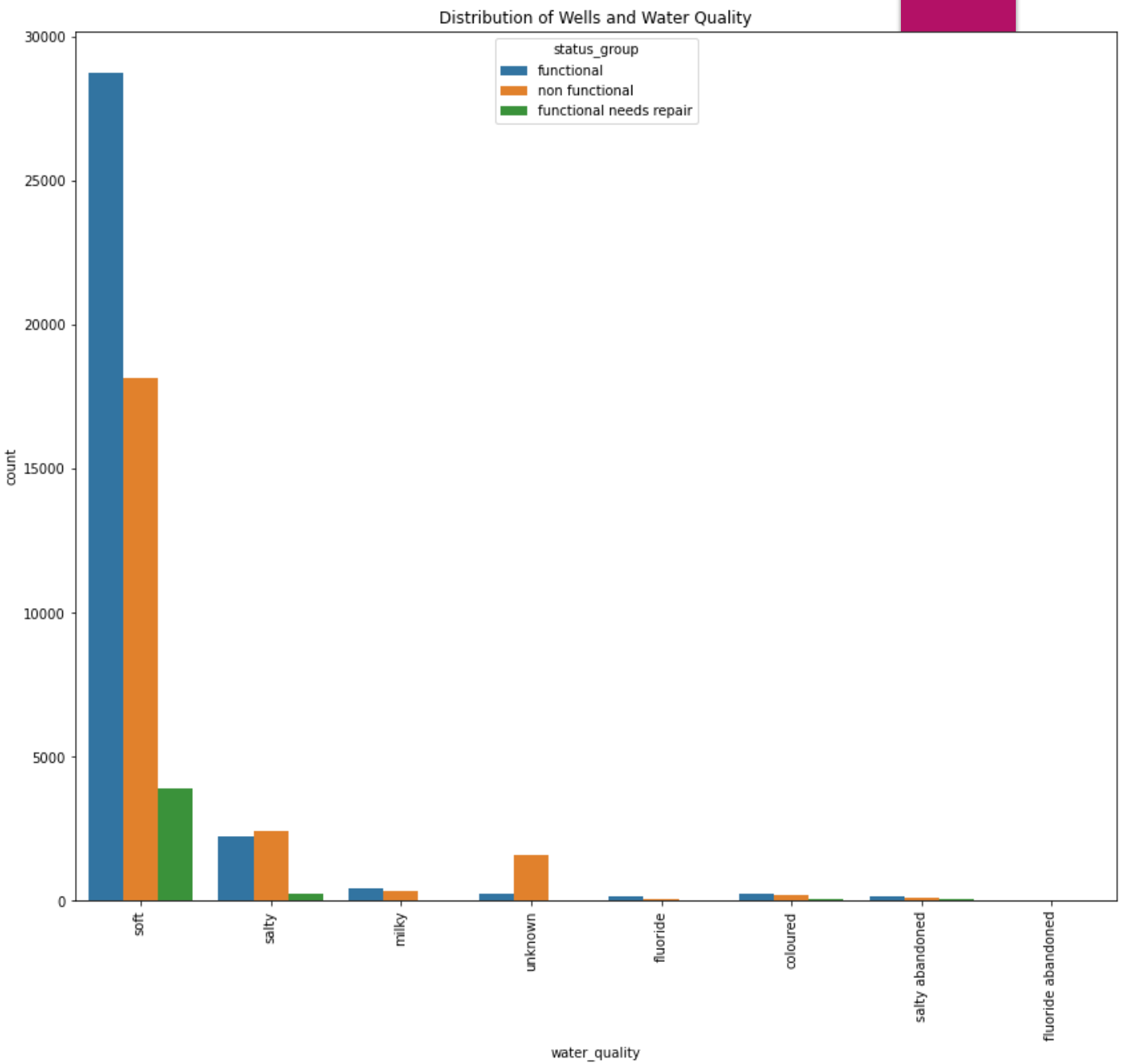
Water Payments

Most of the wells give water for free but also that category has the largest non-functional wells. Wells where one pays per bucket or monthly have a larger number of functional wells compared to non-functional ones



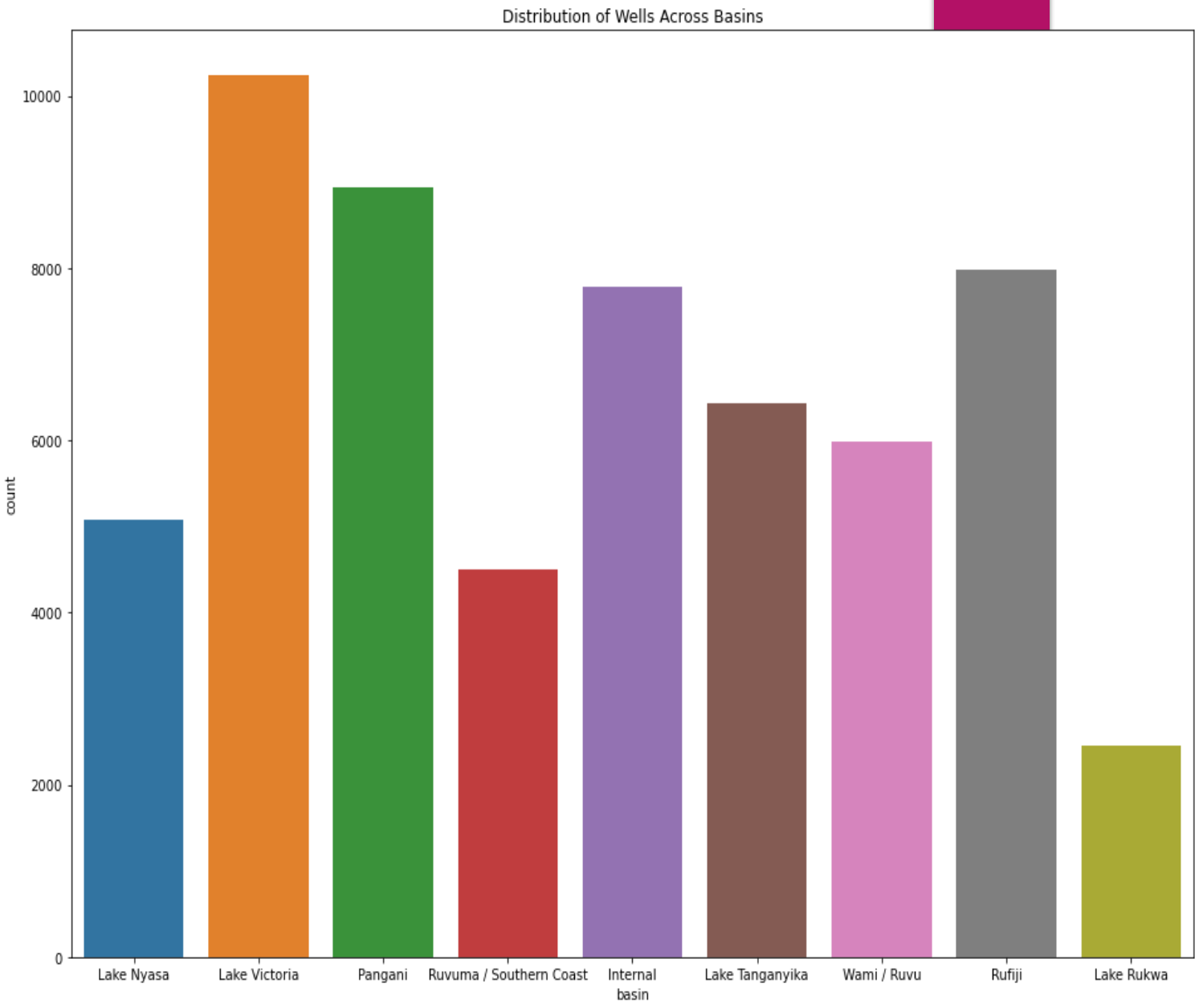
Water Quality

Most of the wells have soft water followed by the salty water while the well that had fluoride and were abandoned had the least numbers. The functional numbers were relatively high as compared to non-functional wells in the soft water wells



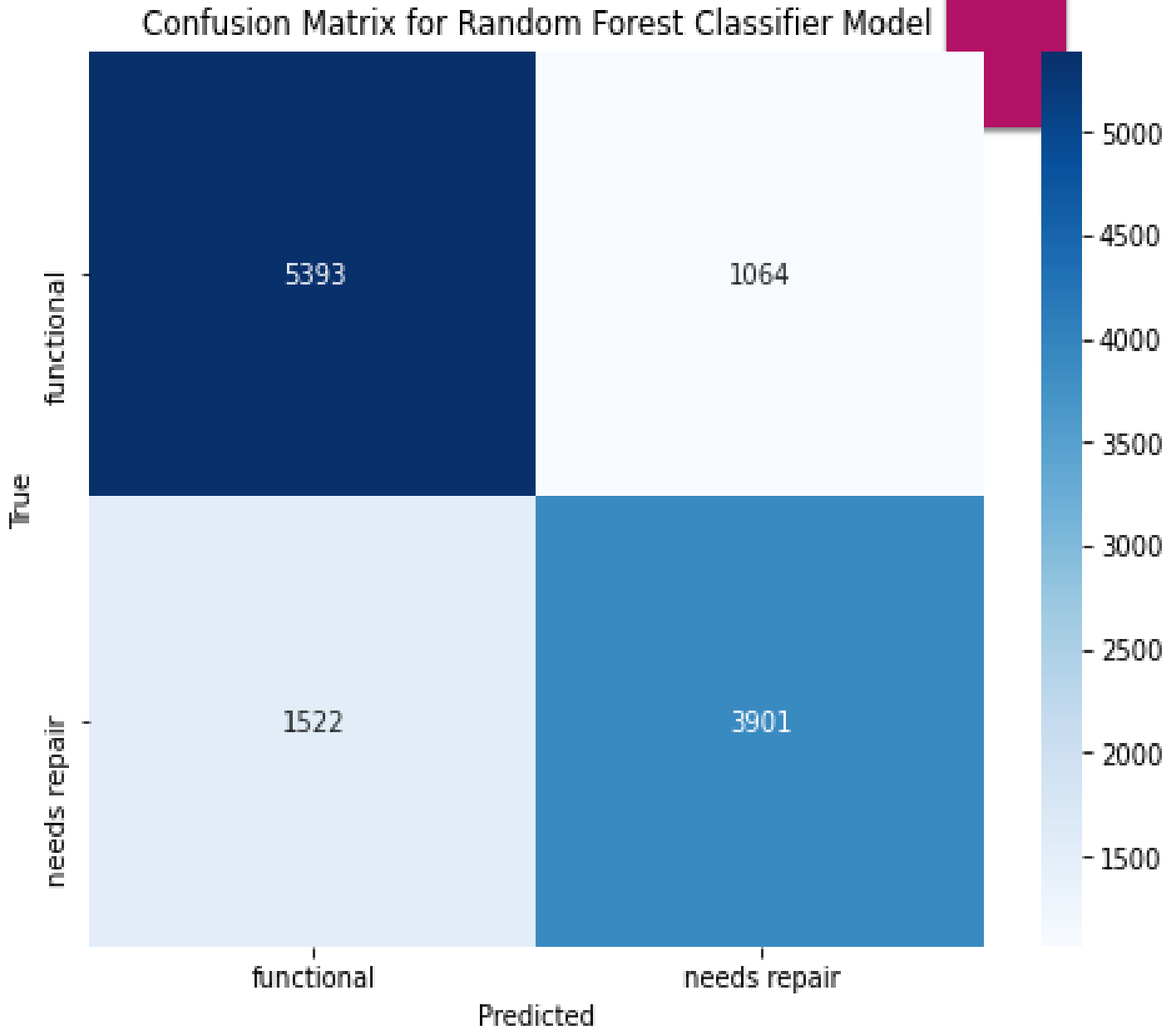
Water Basin

Most of the wells were found on the lake Victoria basin followed by Pangani. The least number of wells were found on the Lake Rukwa basin



Model

The RandomForestClassifier model gave an accuracy of 78% and a 80%accuracy after tuning. This would be an ideal model to adopt for the project. The confusion matrix has also improved and the model is able to predict both Positive and Negative classes



Conclusion

- Prioritizing the construction of wells with soft water
- Introducing payment plans, no matter how modest
- Entrusting well management to user groups has shown a consistent positive impact on functionality
- Constructing wells that guarantee a sufficient water supply is vital.
- Prioritizing the construction and maintenance of groundwater and Communal Stand Pipe wells is a strategic investment
- The RandomForestClassifier model, with an 80% accuracy rate after tuning, proves to be a valuable tool for assessing well conditions

Recommendations

- Prioritize Soft Water Wells
- Implement Payment Plans
- Utilize User Group Management
- Ensure Sufficient Water Quantity
- Invest in Ground Water and Communal Stand Pipes
- Utilize the RandomForestClassifier Model