

Tanzania Water Points

By Chaquayla Halmon



Overview

- Business Problem
- Data
- Process
- Final Model
- Recommendations
- Future Research

Business Problem



Working with the
Tanzanian Ministry Of
Water to predict the
operating condition of a
waterfront.



Data Set

Over 59k data points

There are columns missing anywhere from 300 to 28k data points.

Location

A few columns dealing with regions, coordinates, and districts.

Descriptive

Columns that describe the type of water point or the amount of water.

Clerical

Columns that stated when a water point was constructed or who managed it.

Process

Explore

After acquiring data, take this chance to see what your data offers.

Clean

Get rid of missing and replace incorrect values. Try to maintain as much data as possible.

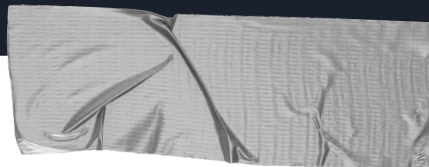
Model

Create many test models for your data. Like Random Forest or XGBoost.

Interpret

Interpret the models and pick the best one for your problem.





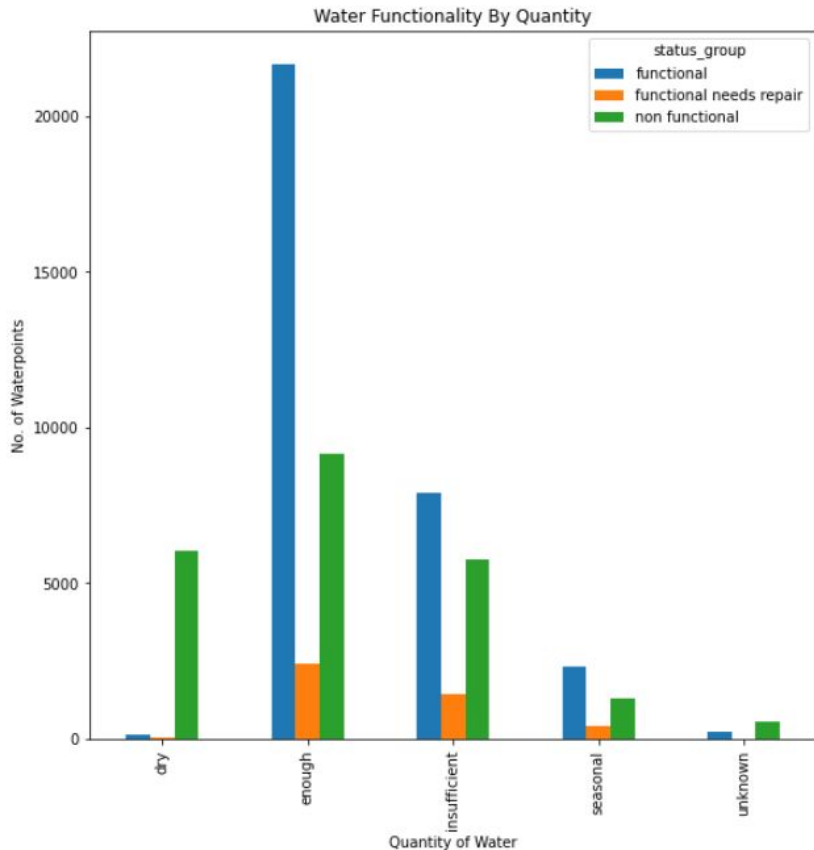
2. Best Model

XGBOOST

- The highest accuracy of 79% overall.
- Predicted 80% of wells that are non functional and functional, and only 56% of wells that need repairs.

Important Features

Weight	Feature
0.2173 ± 0.0022	quantity_enough
0.1120 ± 0.0009	quantity_insufficient
0.0809 ± 0.0031	longitude
0.0807 ± 0.0023	latitude
0.0653 ± 0.0016	construction_year
0.0437 ± 0.0017	gps_height
0.0435 ± 0.0014	payment
0.0378 ± 0.0005	quantity_seasonal
0.0324 ± 0.0010	population
0.0128 ± 0.0011	district_code
0.0098 ± 0.0008	wp_type_communal_standpipe
0.0097 ± 0.0009	et_submersible
0.0093 ± 0.0009	et_handpump
0.0087 ± 0.0015	source_river_lake
0.0081 ± 0.0012	amount_tsh
0.0063 ± 0.0007	source_shallow_well
0.0061 ± 0.0013	permit
0.0057 ± 0.0006	source_spring
0.0050 ± 0.0008	et_motorpump
0.0050 ± 0.0002	lga_Bariadi
... 159 more ...	

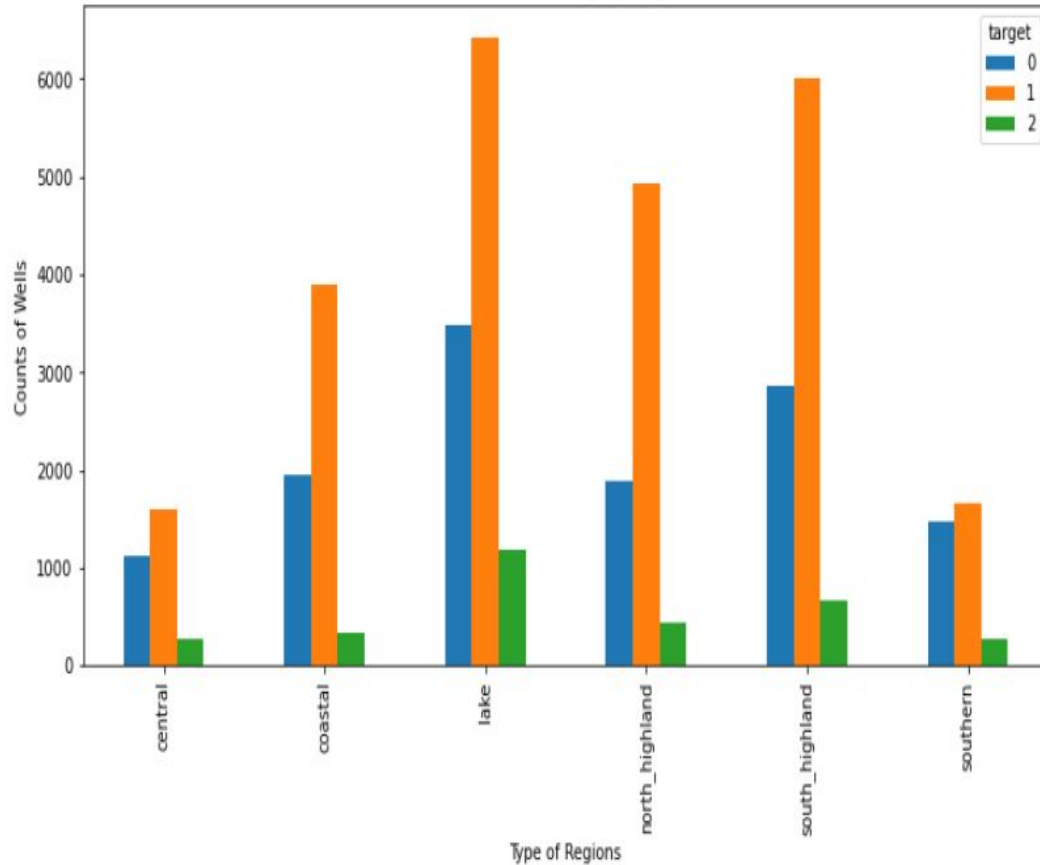


65% percent of 'enough' wells are functional noting this might be a great predictor for functional wells.

As for predicting non functioning wells, 'dry' wells make up 96%.

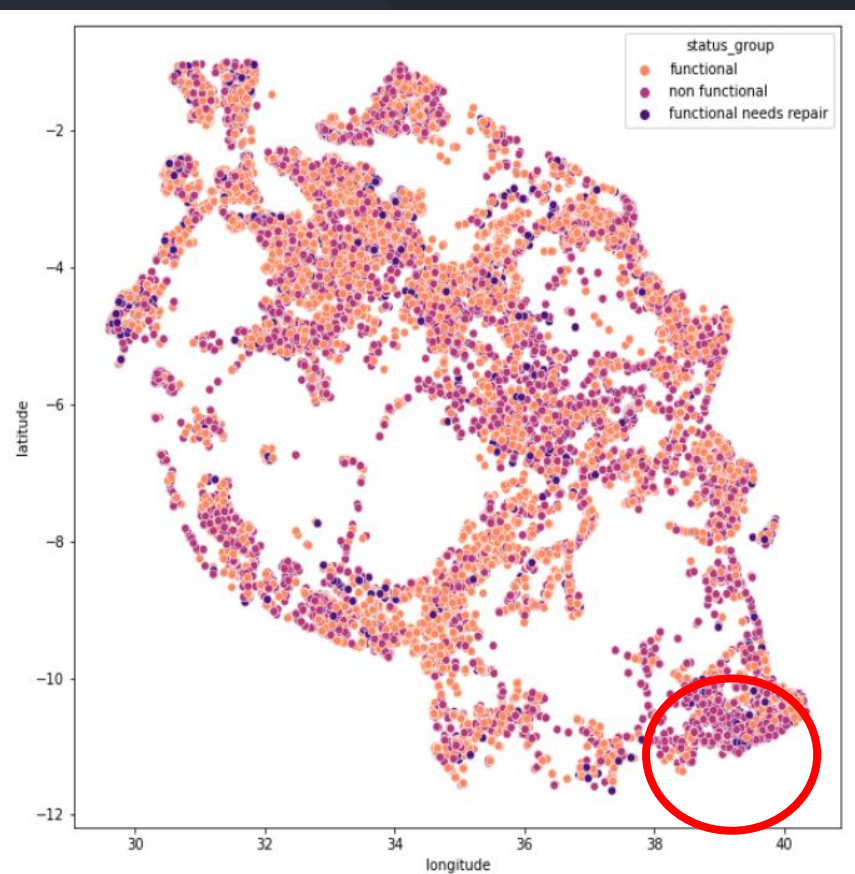
Dry wells makes up about 26% of all non functional wells.

Water Functionality By Regions



Locations

- Lake and Southern Highland zones have the most functioning wells.
- They are both located near bodies of water.
- Whereas, central is somewhat in the middle of Tanzania and doesn't have the same access.



Source: <https://www.mapsofworld.com/tanzania/>

Recommendations

1. Knowing the quantity of the wells can help identify the status of other wells
2. Focusing on areas that are far from bodies of water can help identify areas that has non functional wells due to the lack of access to water.

Future Research:

- Examine the longevity of different extraction types
- Review different management companies and styles for water points with less repairs and non functional water point
- Check to see if population plays a role in usage
- Review how different sources and quality of water can affect structures
- Narrow down the focus of the predictions





Until Next Time

. **Questions: Contact**

→ halmonchaquayla@gmail.com

→ Github: <https://github.com/halmonchaquayla>