Tanzanian Water Wells Functionality Prediction.

—

Milena Afeworki
July 8, 2021

Outlines

- Overview
- Objectives
- Data
- Methods
- Results
- Recommendation
- Further study

Overview

Tanzania, as a developing country, struggles with providing clean water to its population of over 57,000,000. There are many water points already established in the country, but some are in need of repair while others have failed altogether. Using data from Taarifa and the Tanzanian Ministry of Water, we need to predict which pumps are functional, which need some repairs, and which don't work at all. A smart understanding of which water points will fail can improve maintenance operations and ensure that clean, potable water is available to communities across Tanzania.

Project objective:

Building an optimum model for predicting functionality of water wells.

Data

- Locational data
- Static Head
- Extraction type
- Water Quality
- Water source
- Administrative organizations
- Funders
- Installation Company

Methods

1. ## Obtain
   - Data from Taarifa and Tanzanian Ministry of water.

2. ## Scrub
   - Remove unnecessary features and replace missing values.

3. ## Explore
   - Look for trends and understand data.

4. ## Model
   - Create, optimize, and test models.

5. ## Interpret
   - Classification
   - Accuracy
   - Performance

Results

-Data points = 59,400 wells.

-Majority of wells are fed from spring water, rivers and shallow wells.

- Class imbalance in the representation of functionality of the data.

- Functional (54.5%)

- Non-functional (38.7%)

- Functional needs repair (6.8%)

Functionality Status of Wells

functional
non functional
functional needs repair
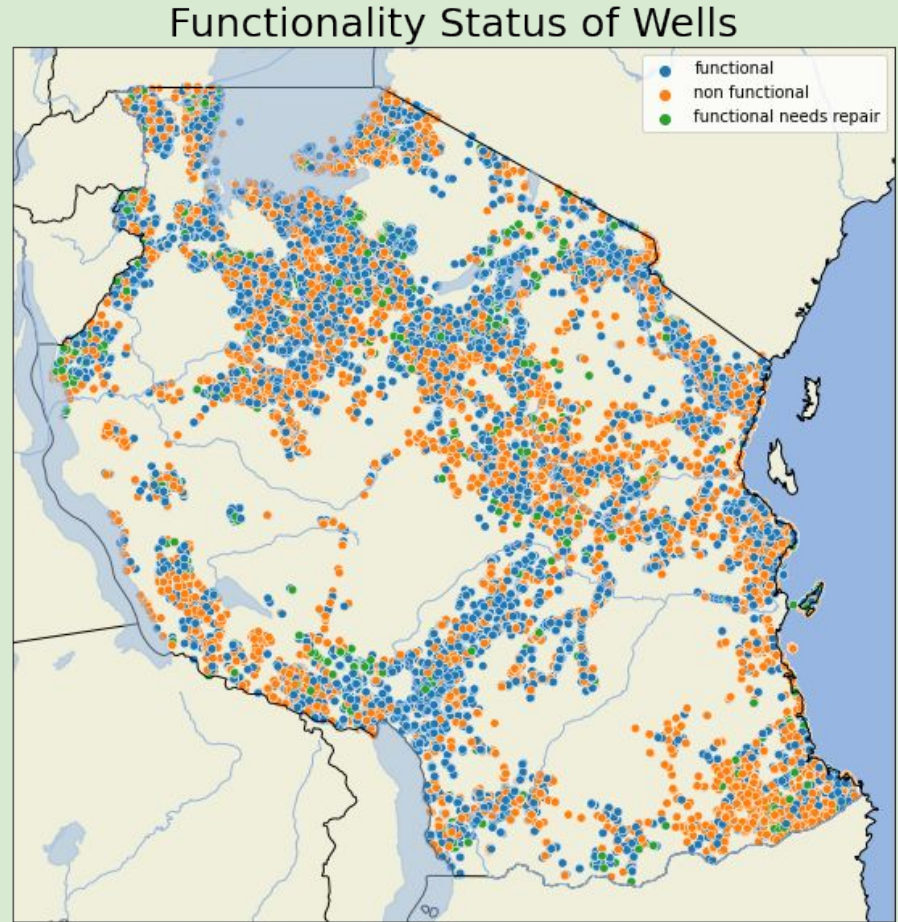
Results

Most Important features.

- Construction year
- Quantity of water
- Extraction type
- Location of Well
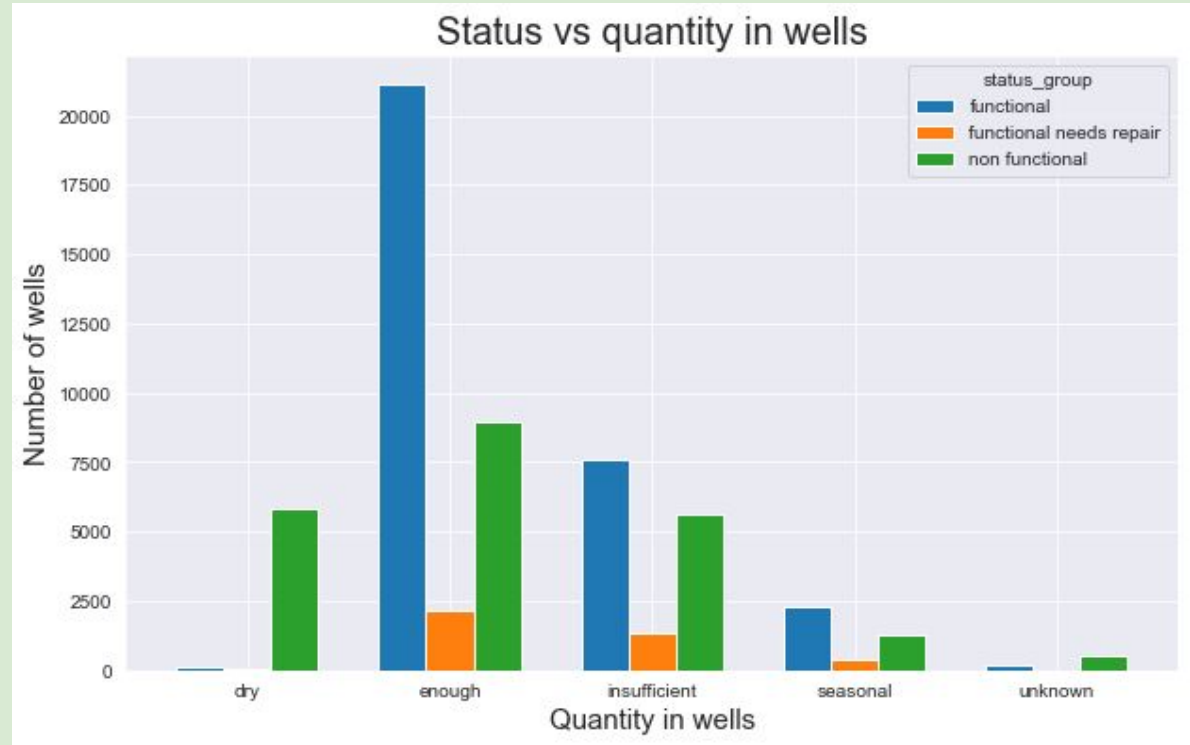- Altitude (GPS height)

Pump by Construction period
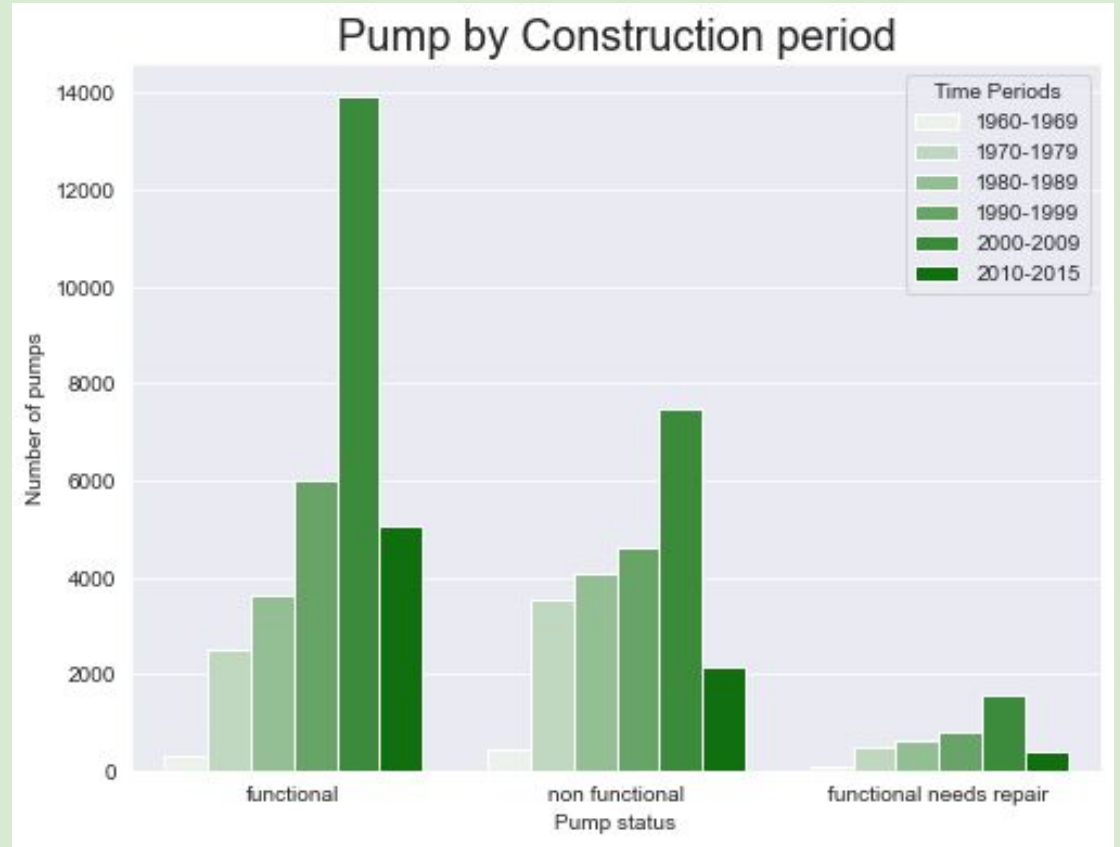
Results

GPS heights & Extraction type.

Location of Wells and their GPS height

Extraction type vs functionality of wells
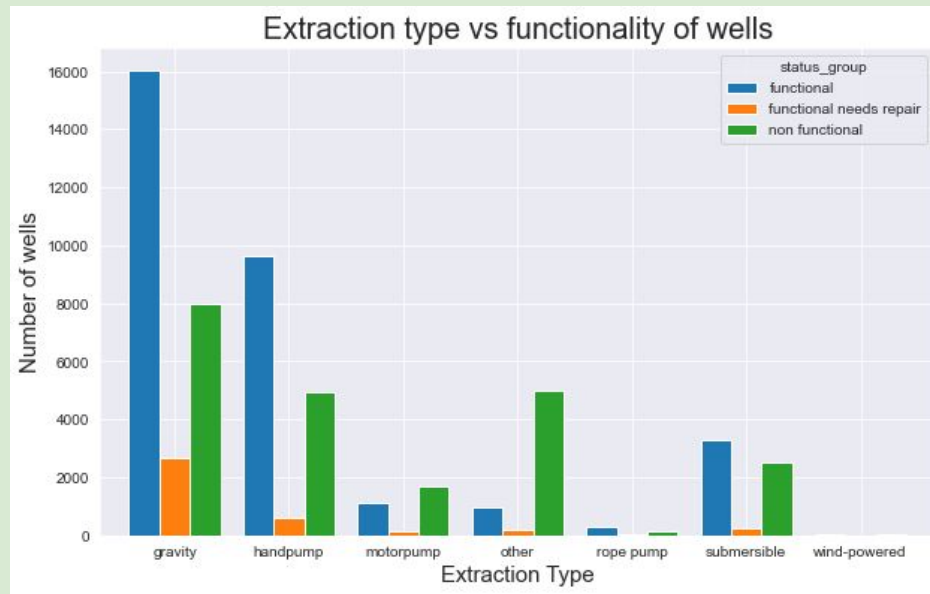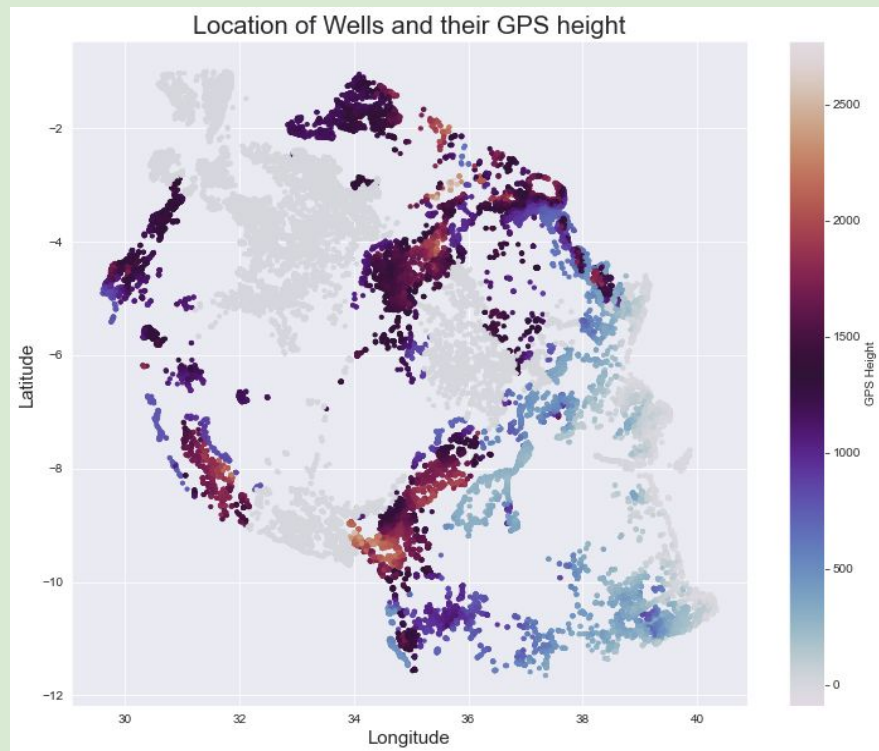
Results

Quantity of water in wells.

Enough   = 56%

Insufficient = 26%

Dry   = 10%

Seasonal   = 7%

Status vs quantity in wells

Target audience

The collaborators:
- Tanzanian Government
- Pump installation Companies
- Project funders
- Scheme managers

Recommendation

Model Performance

The final optimized (RandomForest) model was able to capture and label:

- 69% of the Functional.

- 79% of the Non-functional.

- 55% of the Needs repair.

Further Study

1. Identify and indicate wells that are no longer functional due to being past their life span.
2. Review which extraction types last longer in supply and quality of water.
3. Identify companies with a record of poor installation, and lack of maintenance.
4. Include date records of maintenance, accurate population size and Construction year.

Thank you!

Email : milena22peter@gmail.com
GitHub: @Milenaafeworki