



# Predicting Functional Status Of Water Wells In Tanzania

Vivienne DiFrancesco



## The Problem

- ▶ People rely on the wells for clean water
- ▶ Need better understanding of when repairs are needed
- ▶ Classify wells as functional, non functional, or needs repair through data



## The Data

- ▶ From Taarifa and Tanzanian Ministry of Water
- ▶ Nearly 60,000 entries
- ▶ Contains information like:
  - ◀ Location
  - ◀ Type of water well
  - ◀ Water source
- ▶ Added information about closest cities and servicing water lab

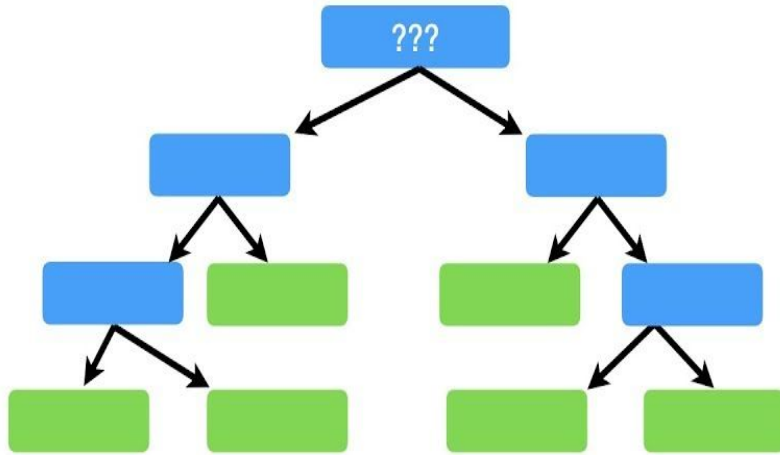


## Methods

- ▶ Obtain
- ▶ Scrubbing
  - ◁ Null values
- ▶ Exploring
  - ◁ Get to know the data
- ▶ Modeling
  - ◁ Different types of models
  - ◁ Used recall as the scoring metric
- ▶ Interpretation

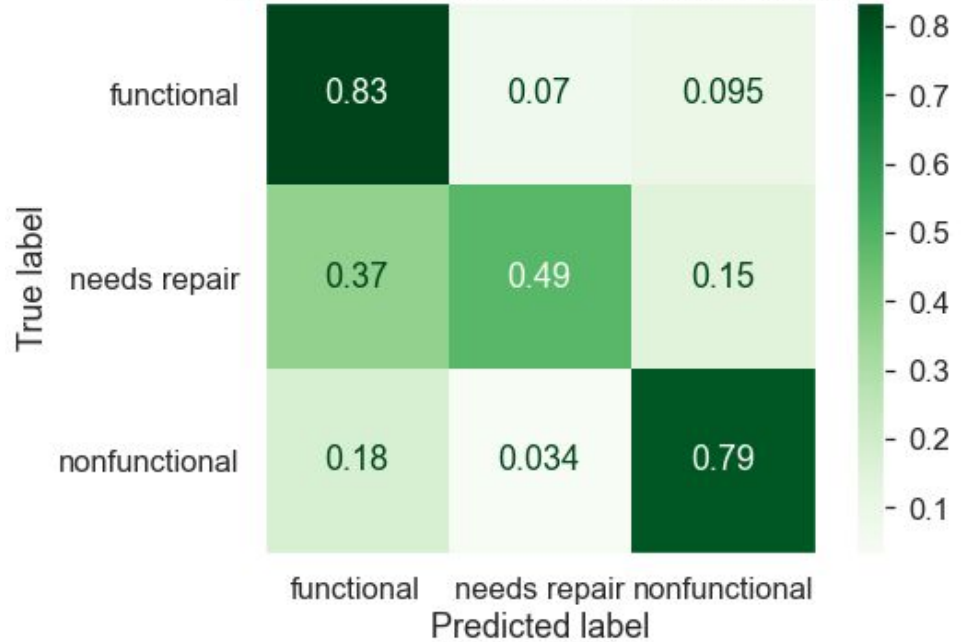
## Best Model

- ▶ XGBoost
- ▶ 79% accuracy



- ▶ Based on a decision tree
- ▶ Gradient descent is used to minimize loss

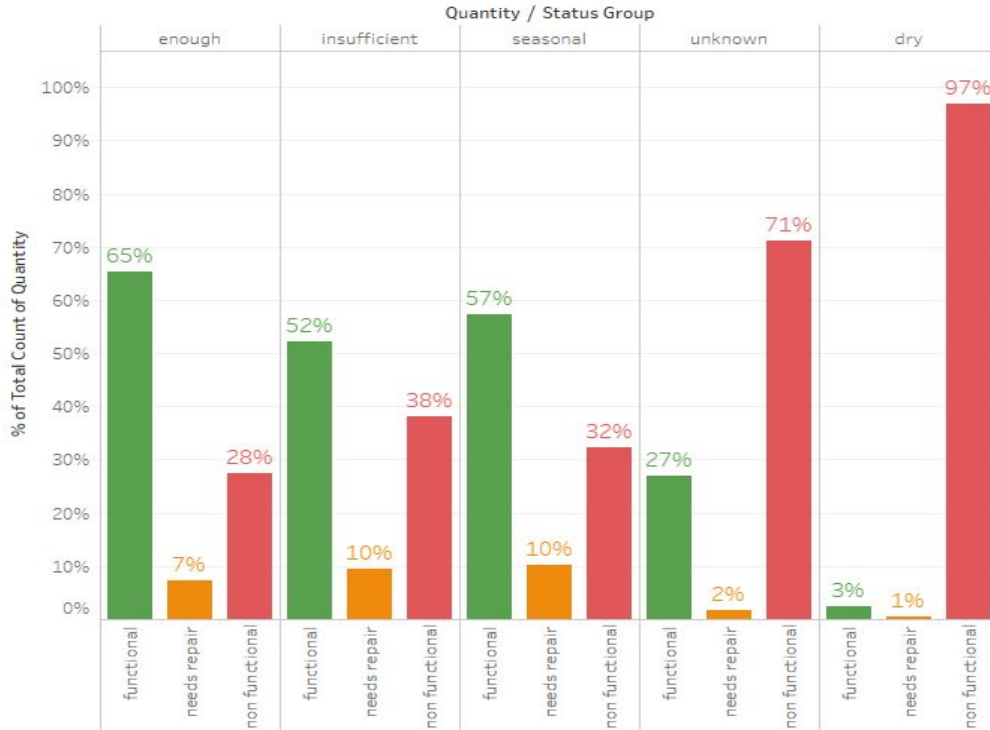
Confusion Matrix For XGBoost Tuned Model



# Recommendation 1:

- Service wells that are dry

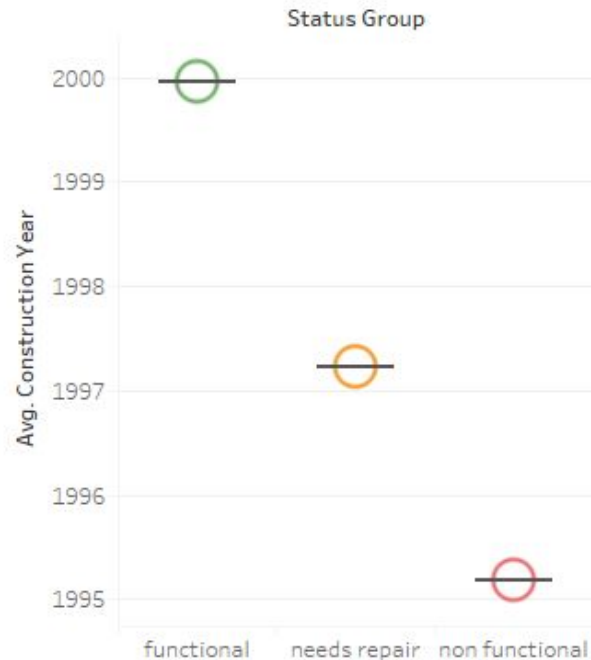
Quantity of Water in Wells



## Recommendation 2:

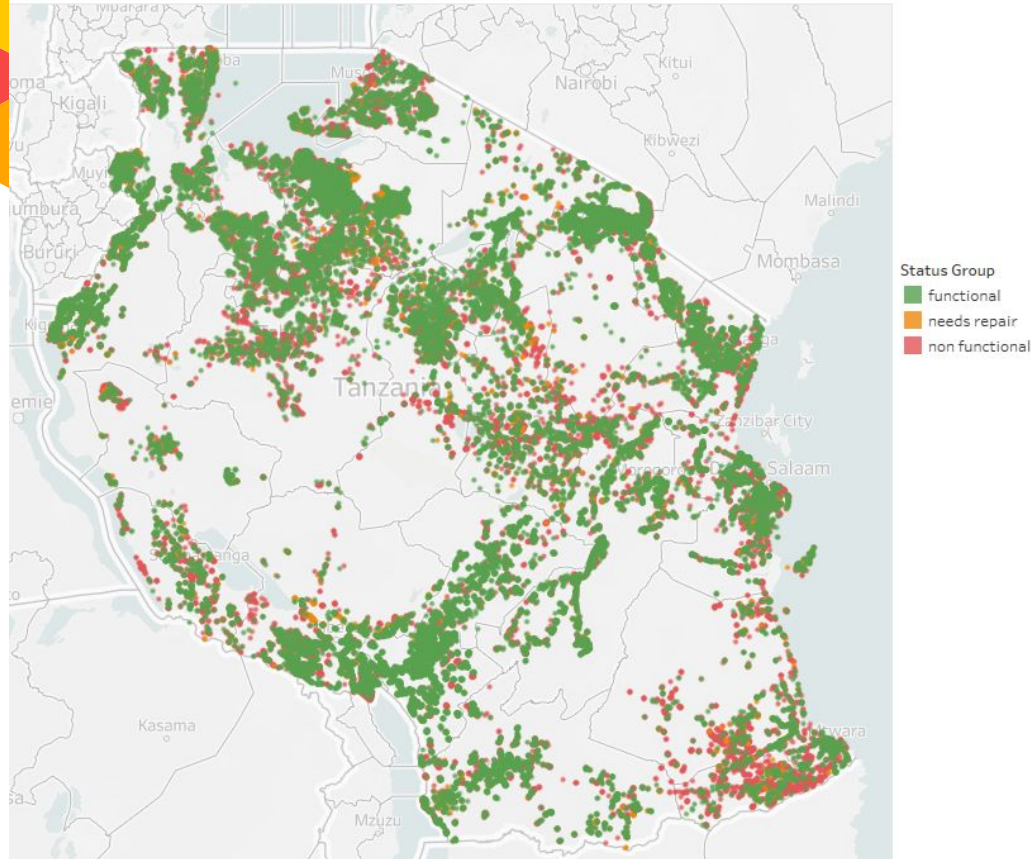
- ▶ Service wells that are older

Average Construction Year of Water Wells



# Recommendation 3:

Location of Water Wells By Functional Status



- Identify area pockets with more non functioning wells





## Recommendation Summary

- ▶ Service wells that are dry
- ▶ Service wells that are older
- ▶ Identify area pockets with more non functioning wells



## Future Work

- ▶ Add nearest city and servicing water lab to data collection
- ▶ Tracking new information to capture needs repair category
- ▶ More tuning or model types

Thank You For  
Your Time

**Any questions?**

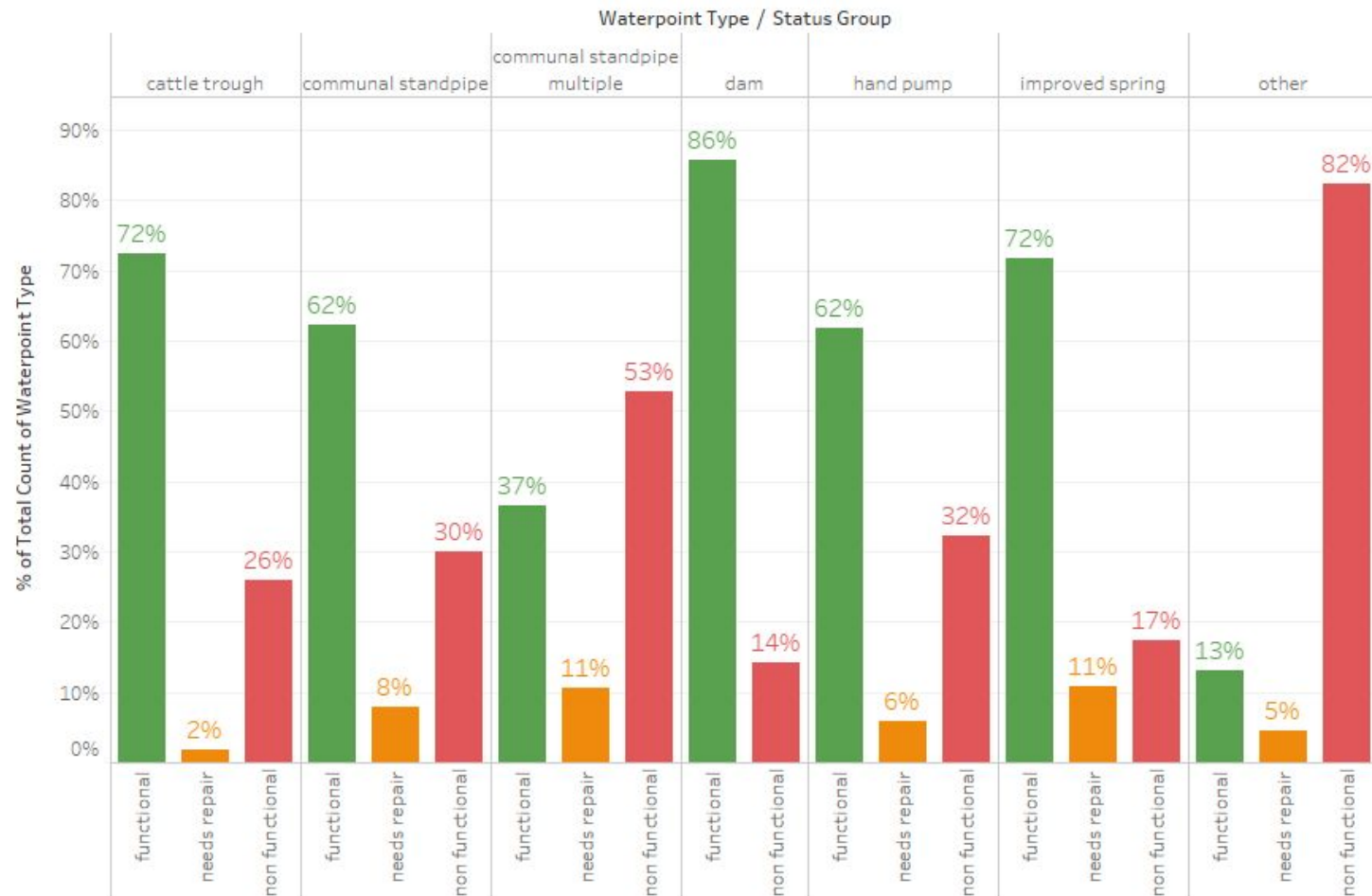




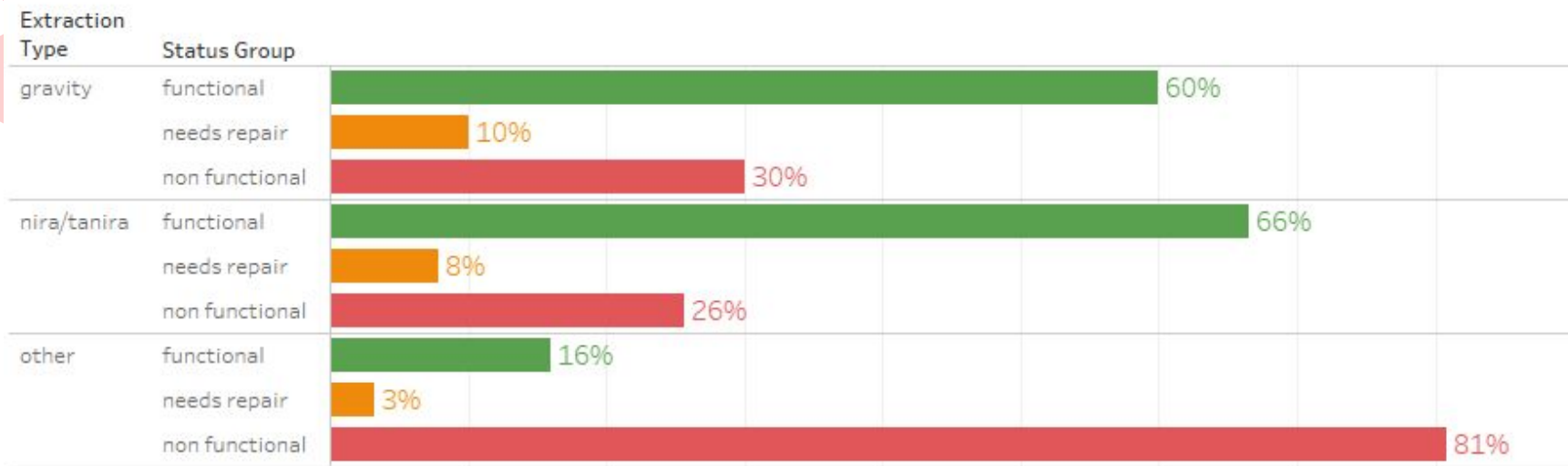
# Appendix

Extra visuals for those who want to know more

## Waterpoint Types of Water Wells

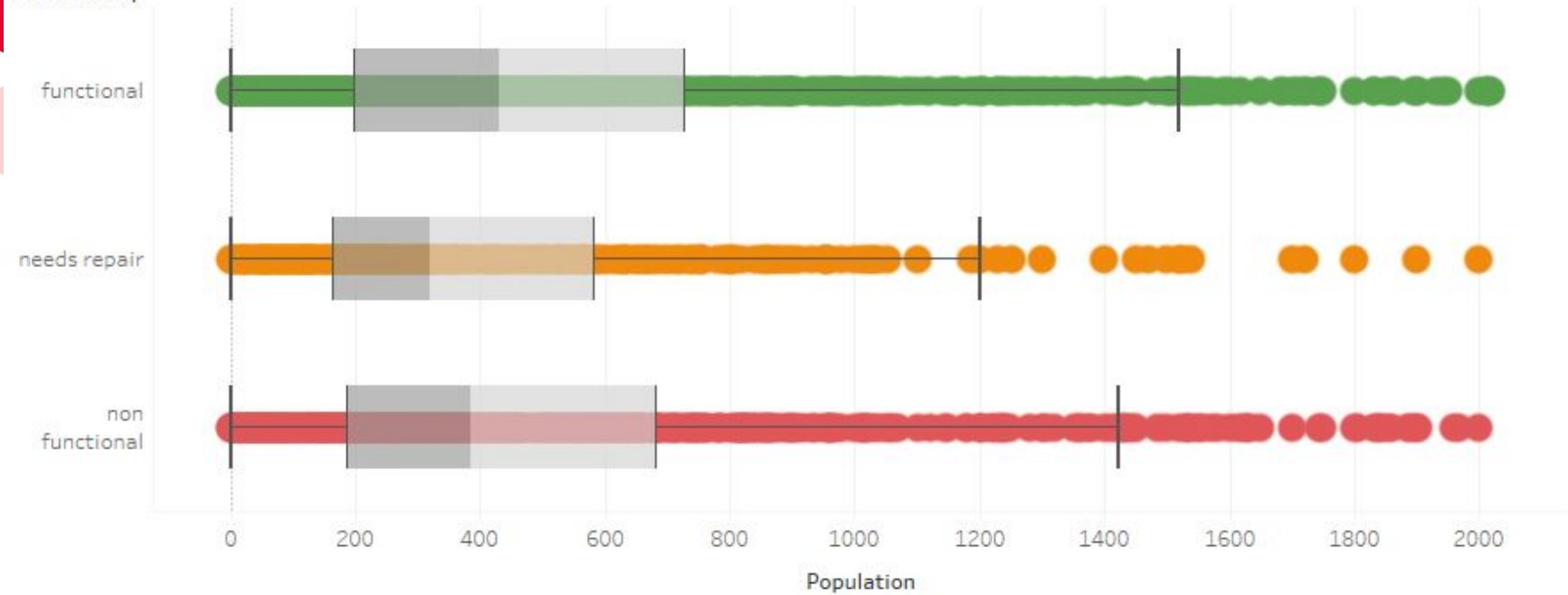


## Extraction Type of Water Wells



## Population Around Water Wells

Status Group



Location Density of Water Wells in Tanzania

