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:
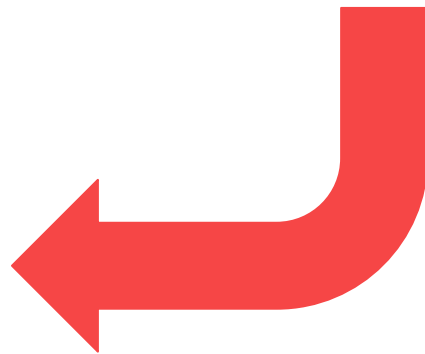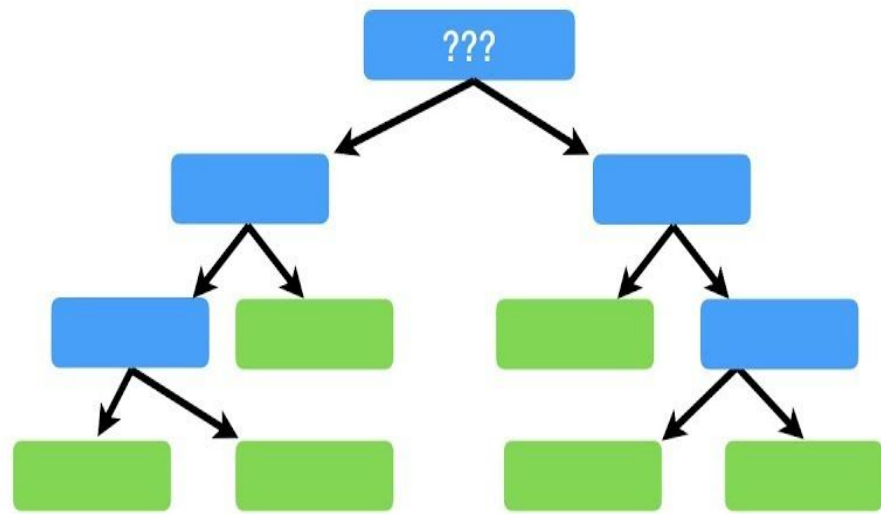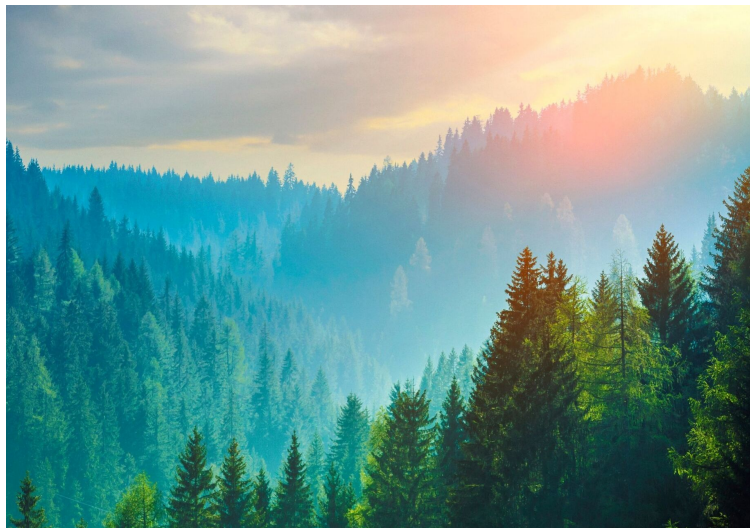 - Location
  - Type of water well
  - Water source

Methods

- Obtain
- Scrubbing
  - Null values
- Exploring
  - Get to know the data
- Modeling
  - Different types of models
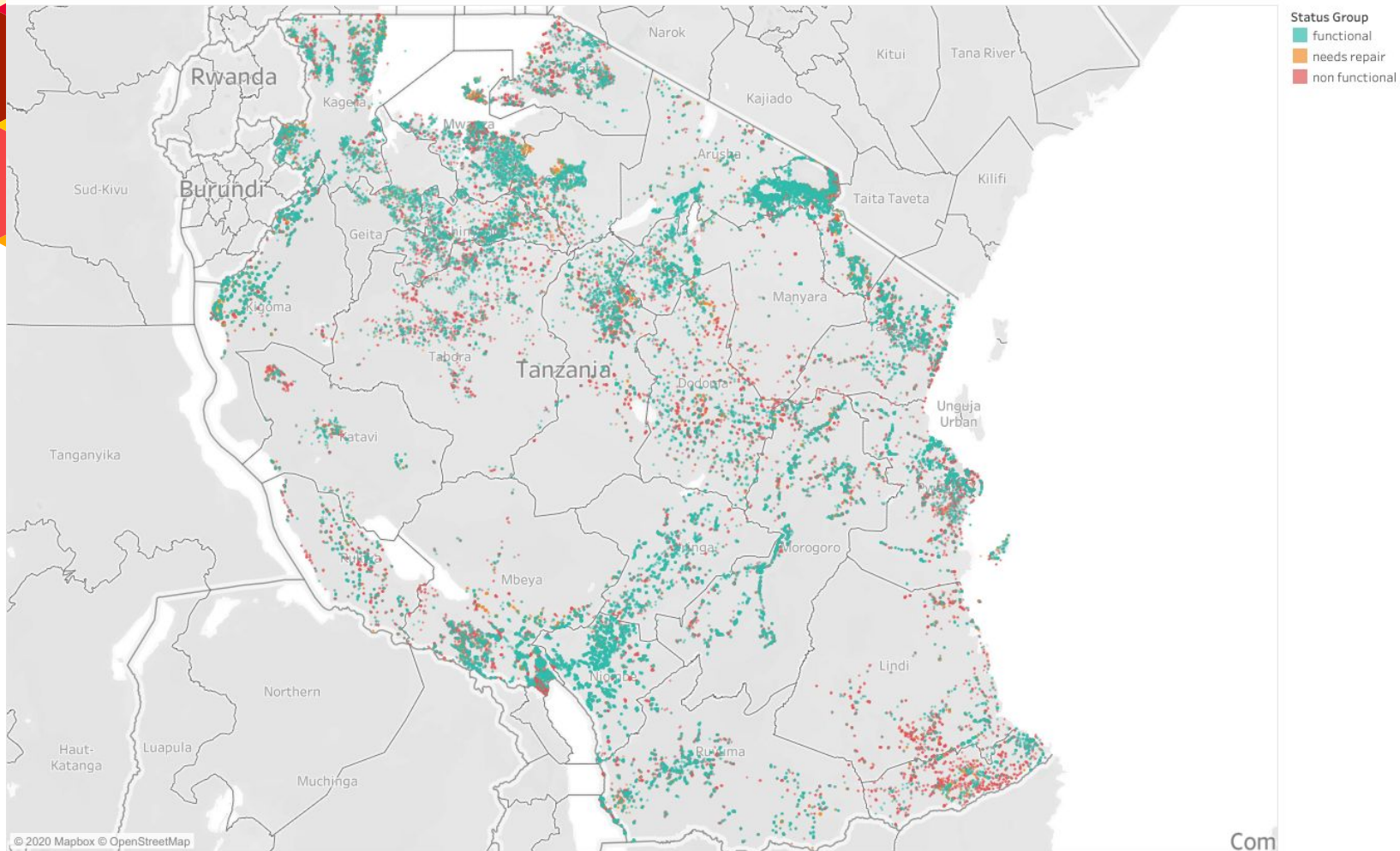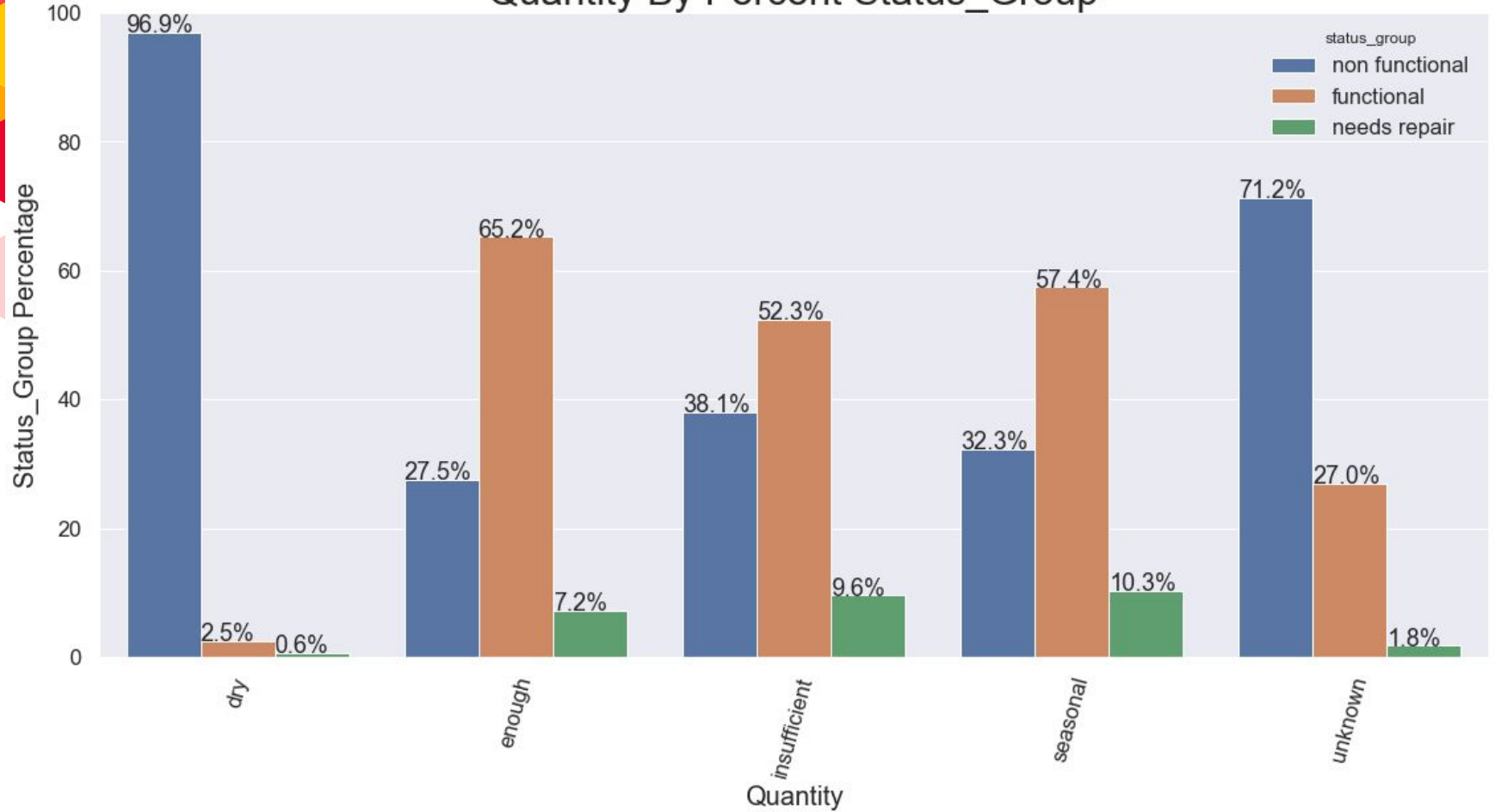- Interpretation

Model

- Random forest
- 78% accuracy

Findings

- Most important features include many based on location
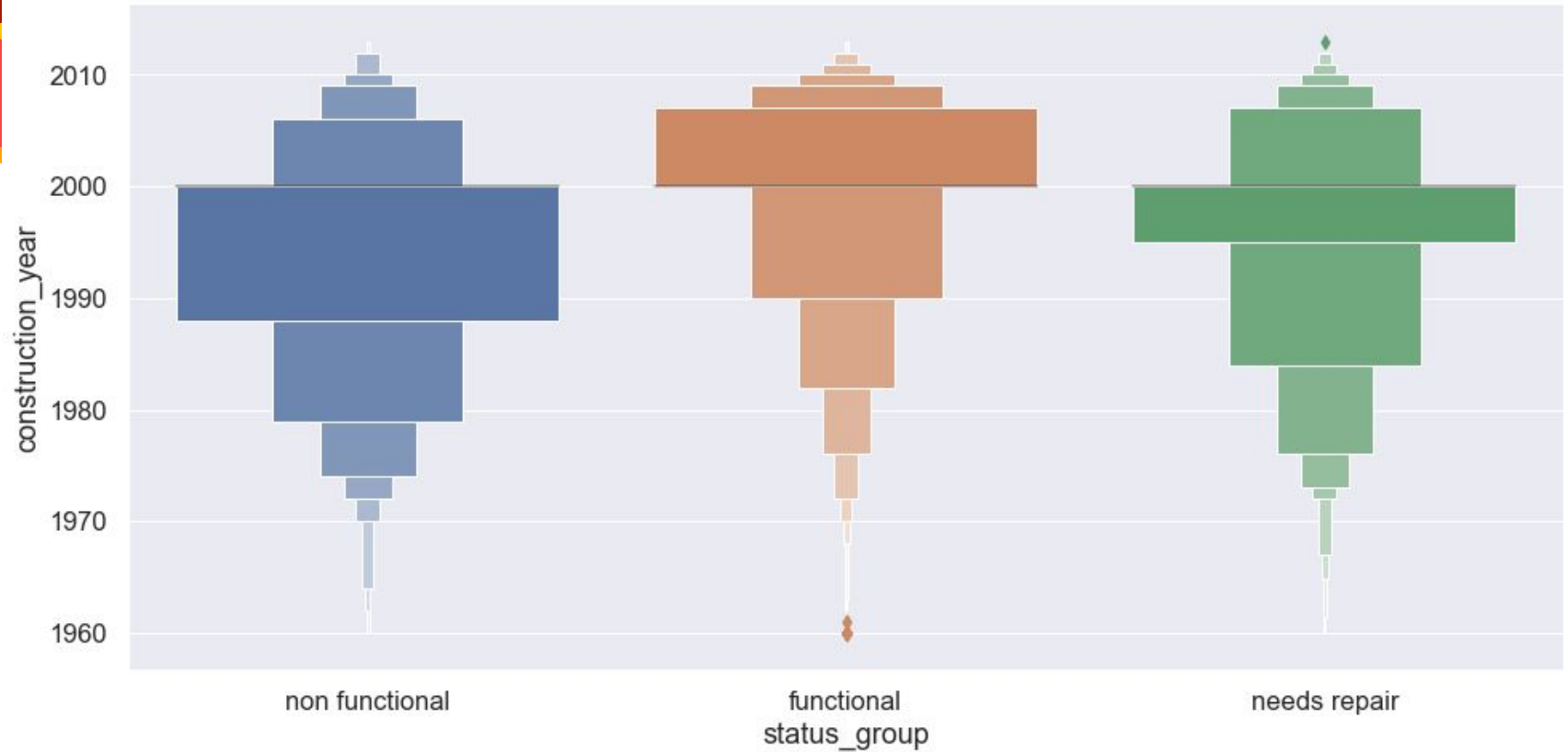- Construction year
- Water quantity

Water Well Status By Location

Status Group
- functional
- needs repair
- non functional

© 2020 Mapbox © OpenStreetMap

Quantity By Percent Status_Group

Construction_Year and Status_Group

Recommendations

- Nearest city and servicing water lab
- Disparity in location should be addressed
- Set up servicing schedule
- Adding features
    - Date well was last serviced
    - Repairs record
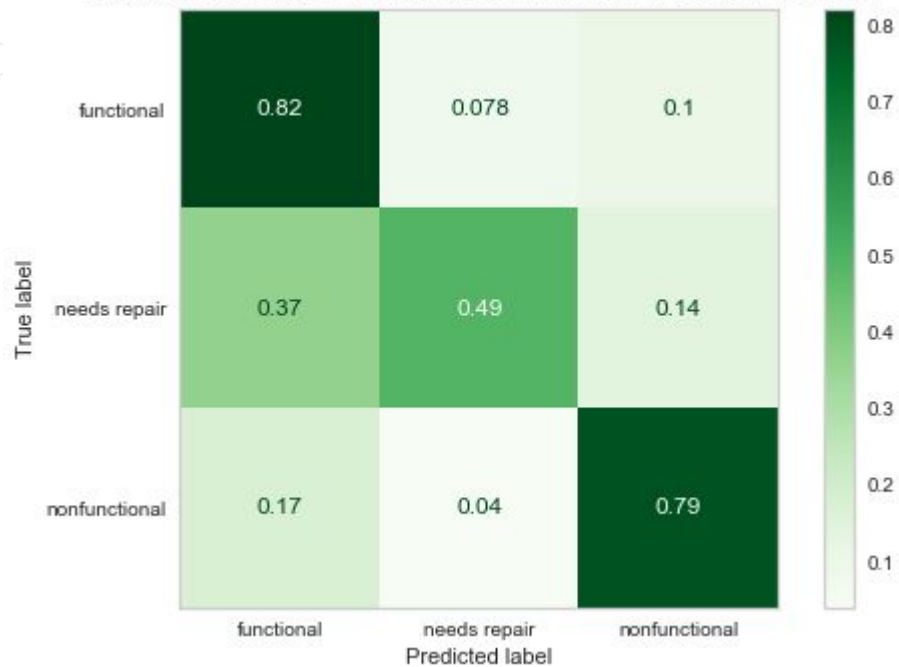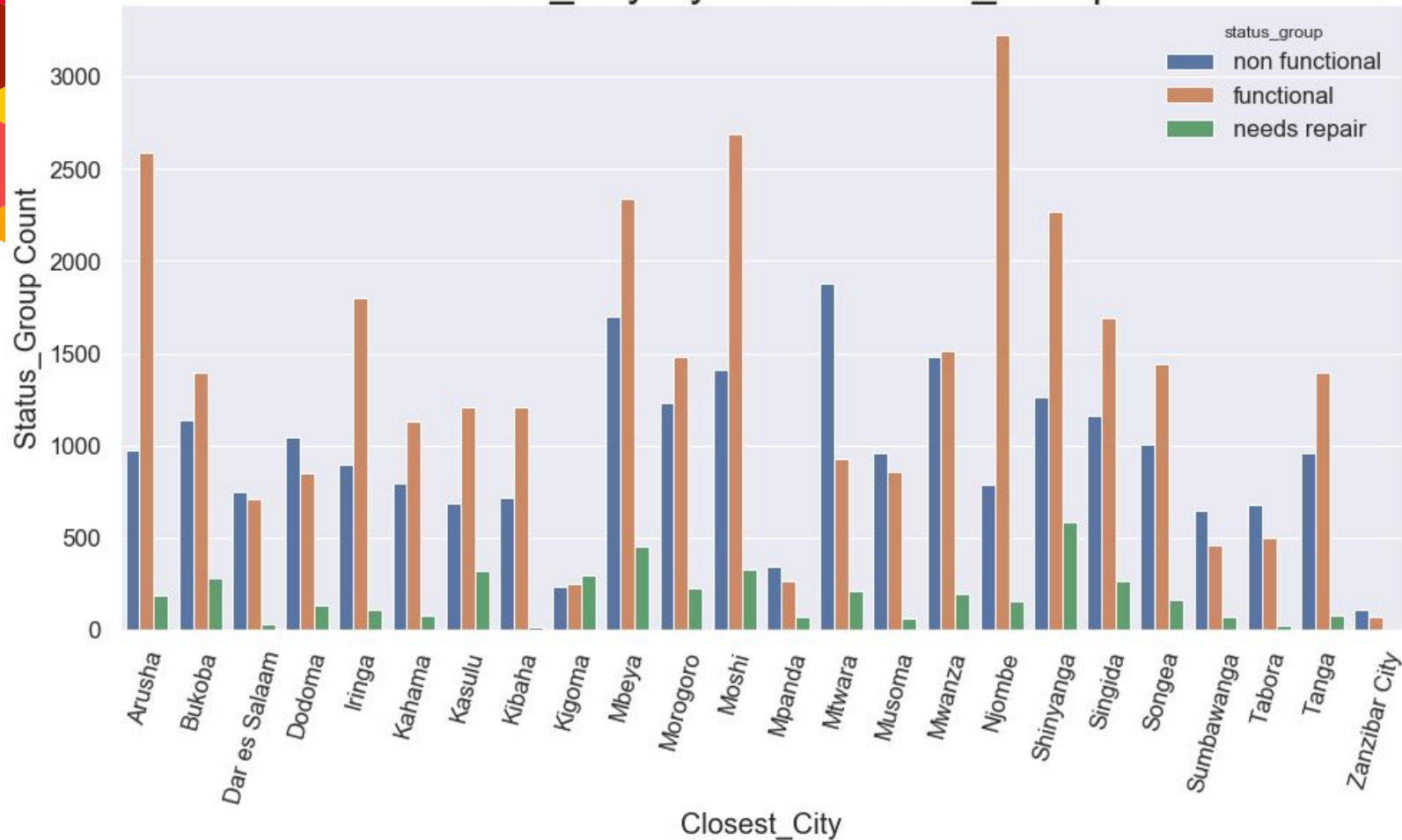
Thank You For Your Time

Any questions?

Appendix

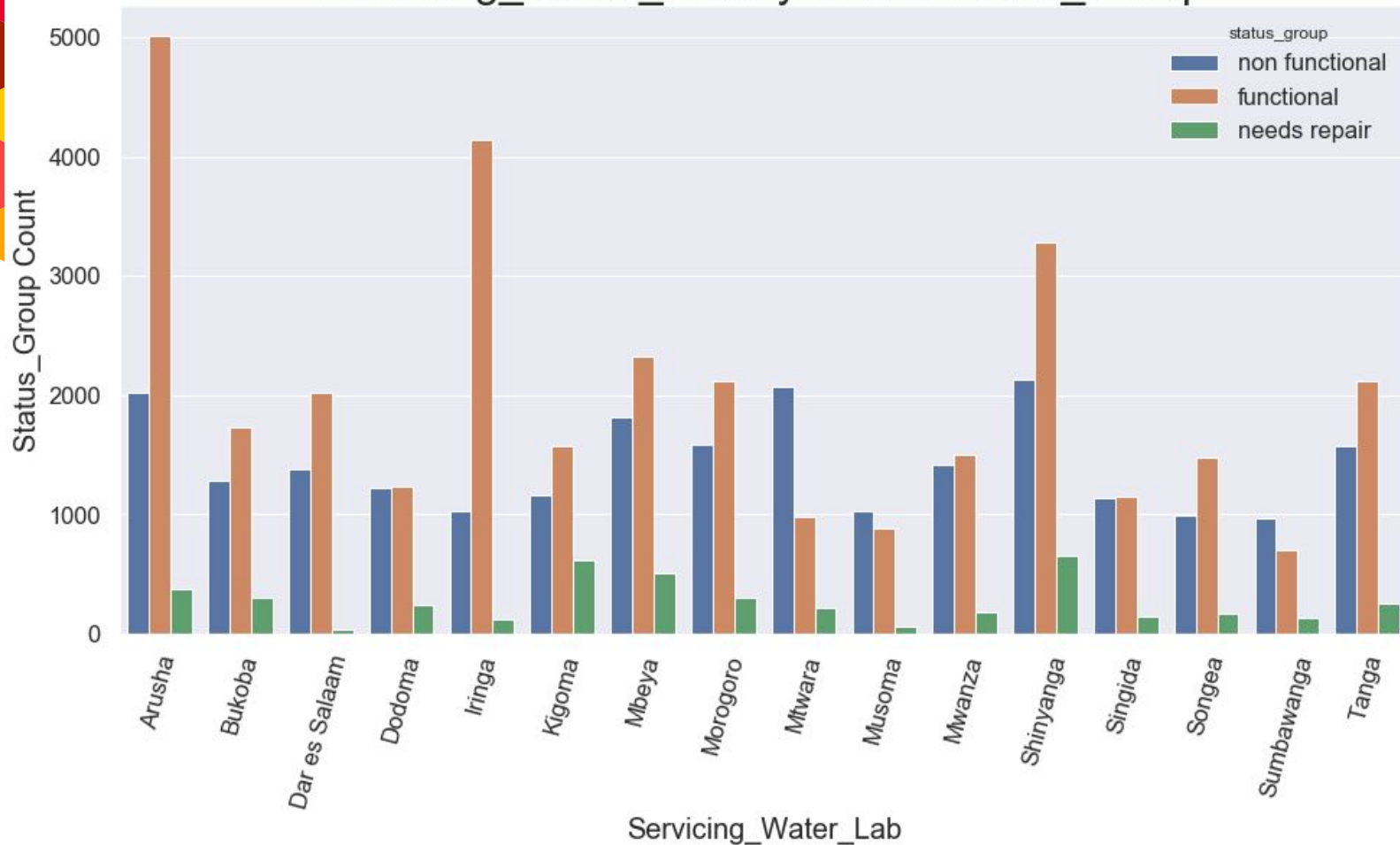Extra visuals for those who want to know more

Confusion Matrix For SMOTE Random Forest Model

Closest_City By Count Status_Group

Servicing_Water_Lab By Count Status_Group

Waterpoint_Type By Percent Status_Group