Tanzanian Waterpoint Status Prediction

Presenter: Melody Bass

Overview

- Tanzania is a developing country that struggles to provide access to safe drinking water for it’s 59 million residents
- Build a predictive model and provide insight on water pump failure
  - 60,000 waterpoints in Tanzania
  - Status of the waterpoints
  - 39 independent variables

Background

- According to WHO, 1 in 6 people in Tanzania lack access to safe drinking water
- 29 million don’t have access to improved sanitation
- Women walk 2 to 3 km per day carrying 20-25 liters on their head and sometimes wait hours at the water source

Business Problem

- The Tanzanian government has a severe water crisis on their hands
- They want to predict which pumps are functional, functional but need repairs, and non functional
- Taarifa and Tanzanian Ministry of Water have shared the dataset to aid understanding of pump failure
- I will build model to help the government improve maintenance operations and ensure clean drinking water is accessible to communities across Tanzania

Business Problem cont

- Precision is our main metric of model selection

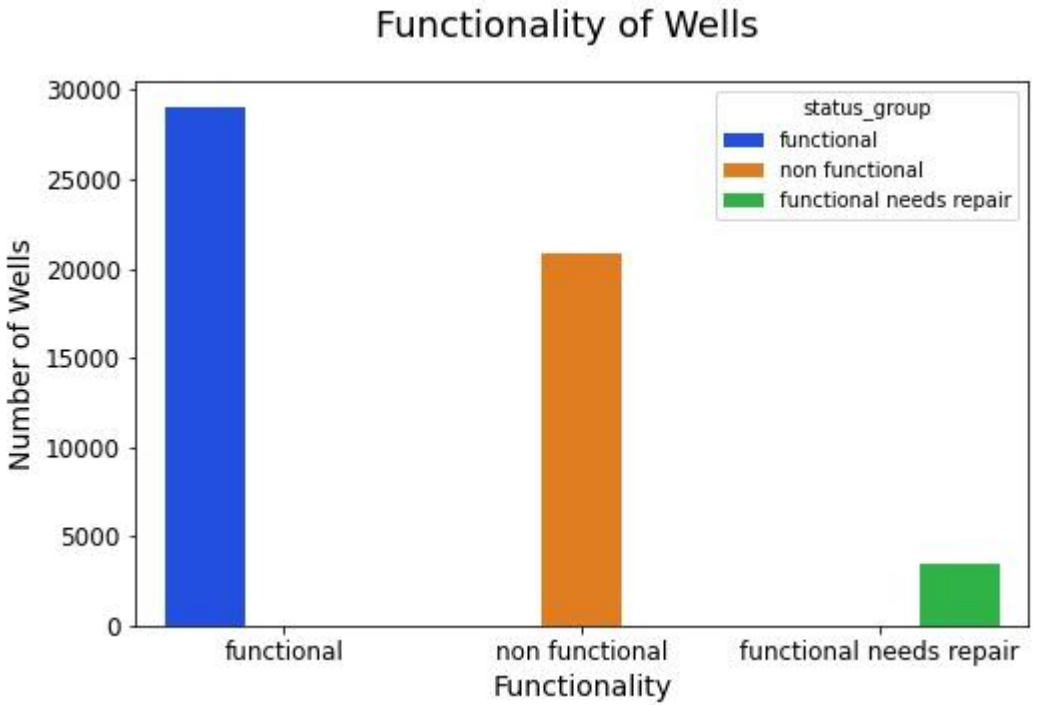
- A non functional well being predicted as a functional well is worse than the opposite case

Data

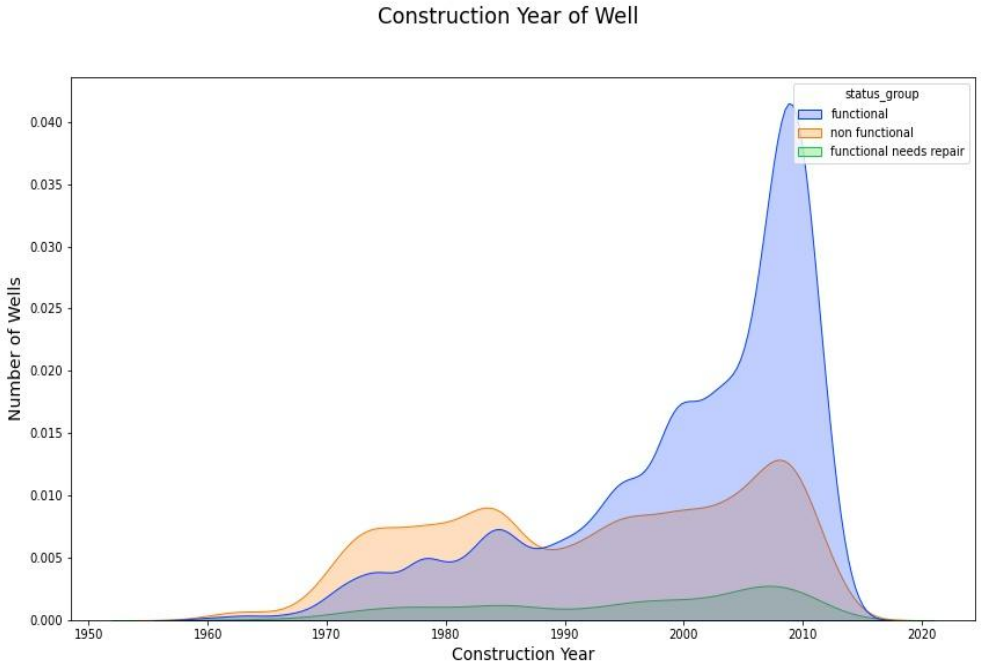
- Dataset contains information on 60,000 waterpoints in Tanzania
- 39 independent variables
- Pump Status
  - Functional
  - Functional needs repairs
  - Non functional
- Available for download on DrivenData
  - https://www.drivendata.org/competitions/7/pump-it-up-data-mining-the-water-table/page/23/

Distribution of Waterpoints

- 29,000 functional
- 21,000 non functional
- 3,500 functional needs repairs

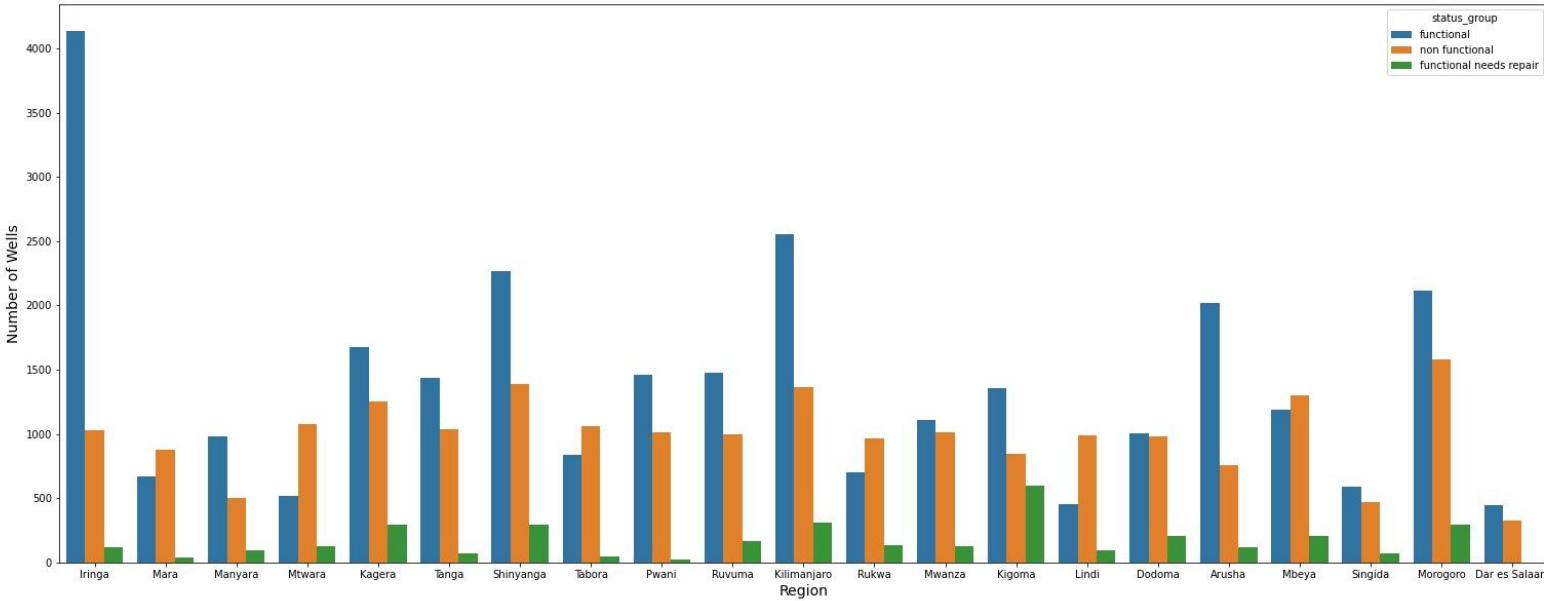


Construction Year of Waterpoint



- As expected, the older the pump installation year, the more non functional pumps there are.
- High rate of functioning pumps after 1988 peaking in 2000s

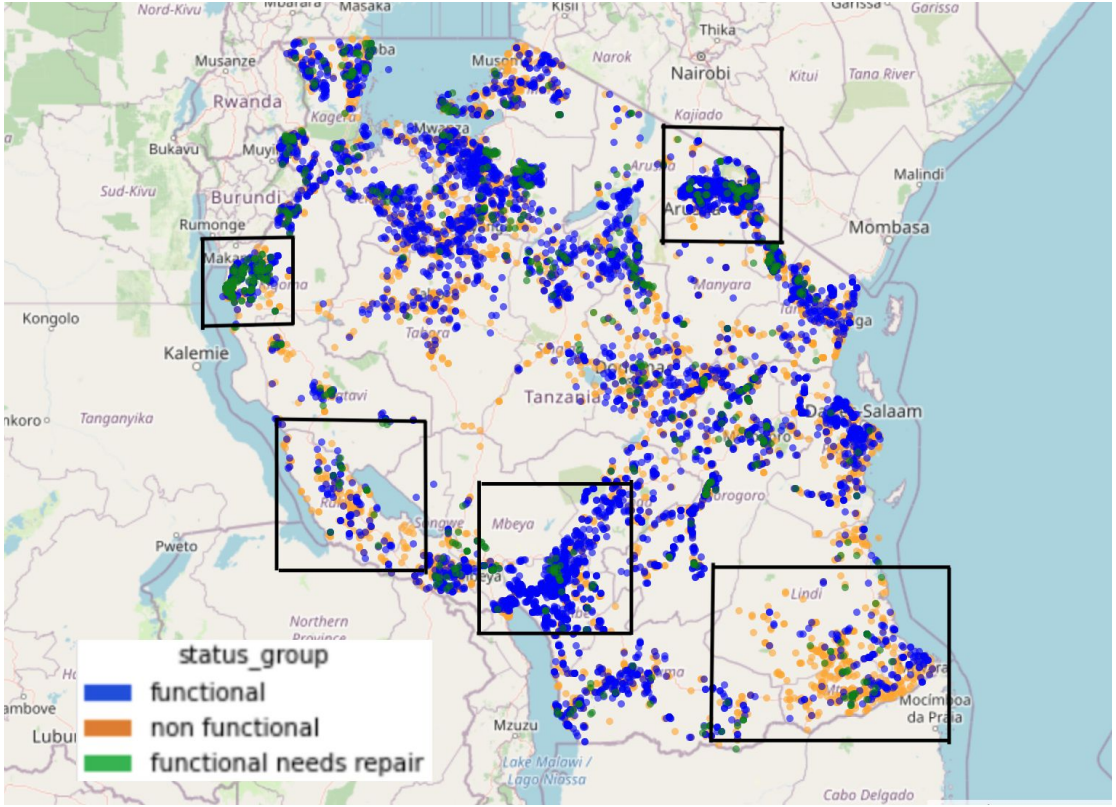
Well function by Region



- There are a high number of functional wells in Iringa, Shinyanga, Kilimanjaro, and Arusha.
- More non functional wells than functional in Mara, Mtwara, Lindi, and Rukwa

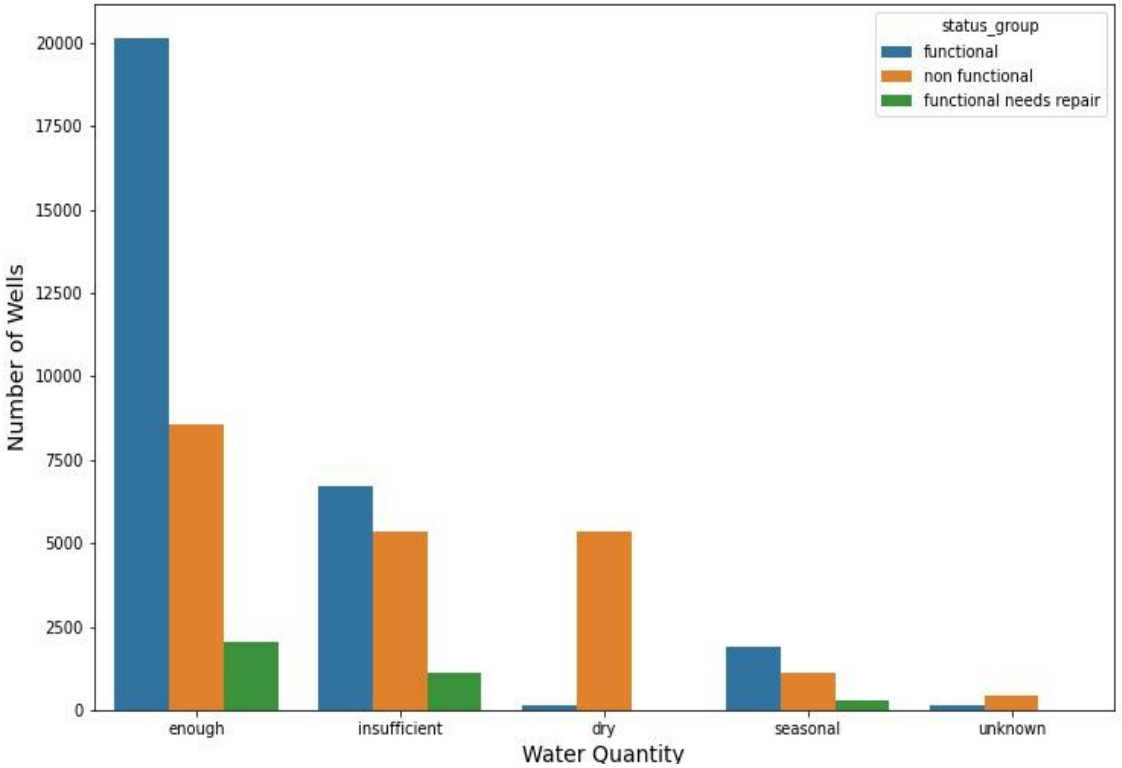
Insights

- High rate of non functional wells in southeast corner of Tanzania in Mtwara and Lindi, as well as Mara and Rukwa
- There is cluster of functional but needs repair wells in Kigoma



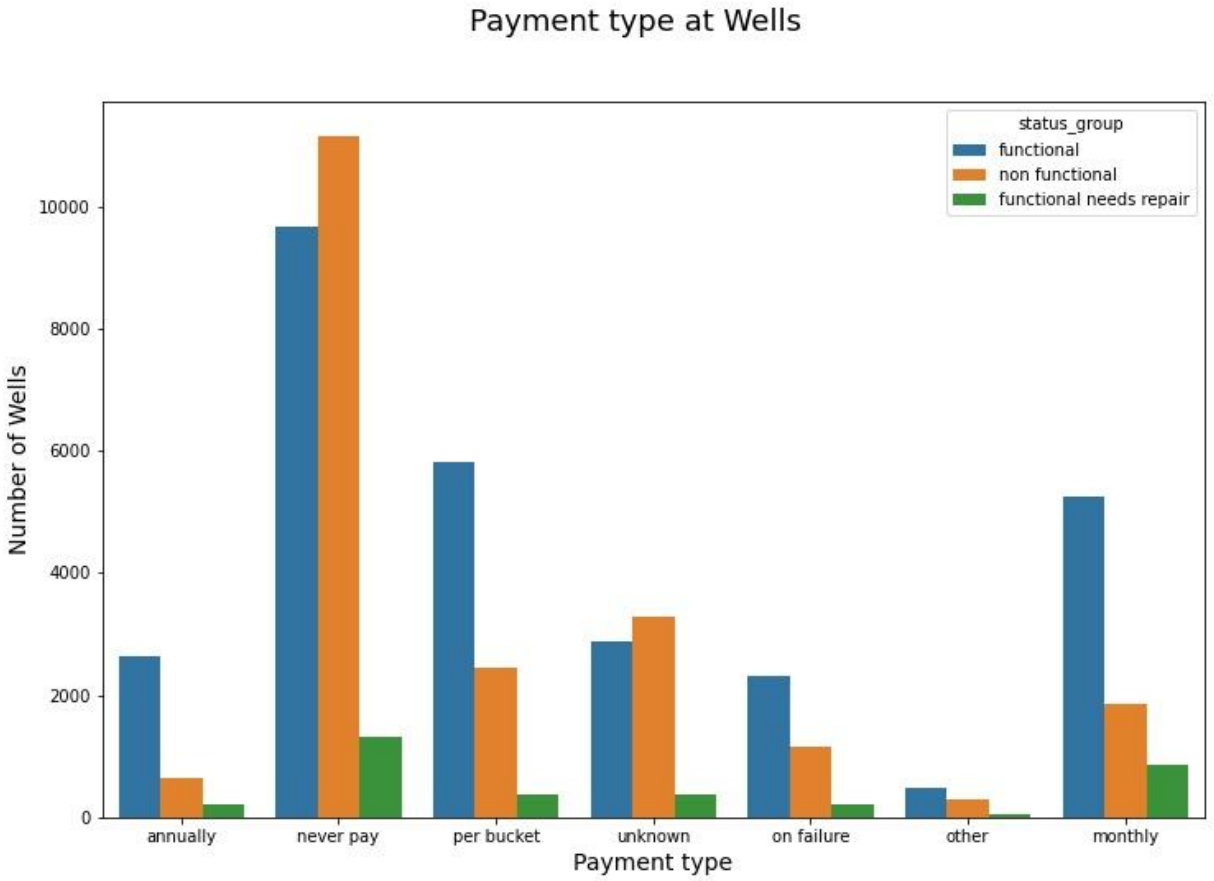
Water Quantity

- As expected, high number of non functional wells that are dry
- Over 8,000 waterpoints have enough water, but are non functional.
- 2,500 are functional but need repairs

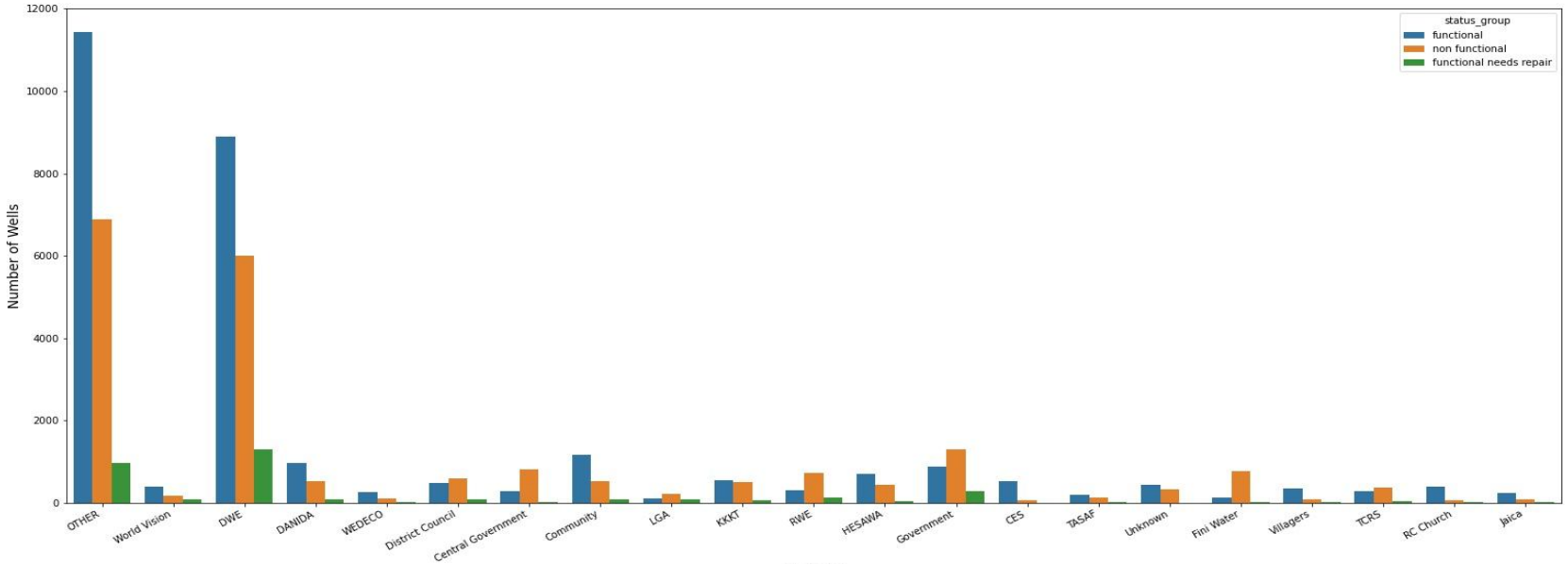


Payment

- Wells with no fee are more likely to be non functional
- Some form of payment increases functionality



Installer



The Government, District Council, and Fini Water all have a high rate of pump failure

Recommendations

- Location
  - Target repairs in areas like Lindi and Mtwara that have a high rate of non functional wells
  - Make repairs to functional wells in Kigoma to maximize cost effectiveness
- Repairs
  - Prioritize non functional and functional wells which need repair and have enough water
- Payment
  - Payment provides incentive and means to keep wells functional
- Installers
  - Avoid using installers with a high rate of pump failure

My model has 81% precision and 82% accuracy

Future Improvements

- Improve Data
  - Better data will build a better model
- Monitor Wells
  - Update regularly with new data to continuously improve strategy

Thank You!

Email: meljoy1099@email.com
GitHub: https://github.com/meljoy1099/waterwell-status-prediction