

Tanzania Waterpoints

Data Analysis and Model

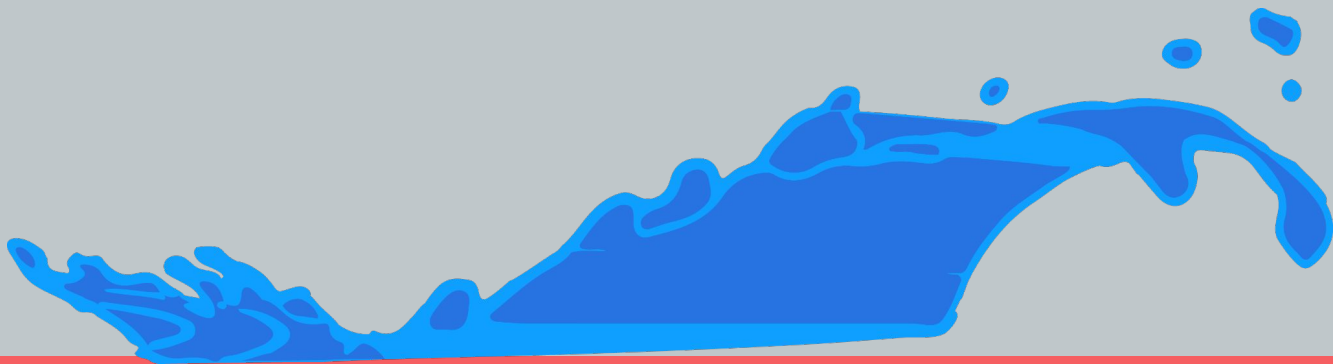
By: Mohammed Siddiqui
April 23, 2021



SUMMARY

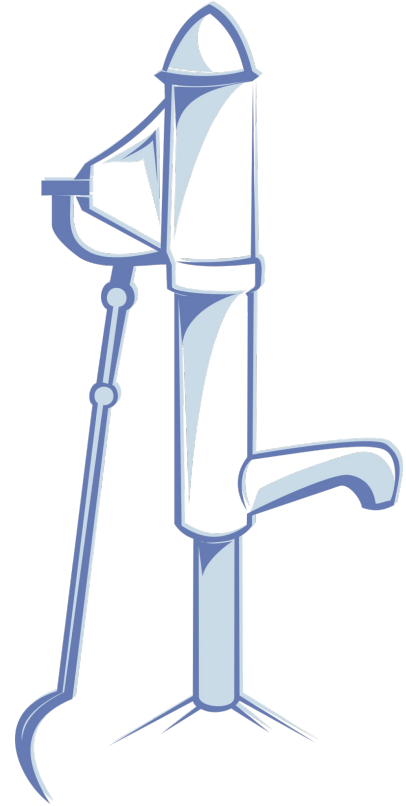
Descriptive analysis and modeling reveal which factors contribute most to repair needs:

- Water Quantity
- Payment
- Region and Local Government Authority



OUTLINE

- Business Problem
- Data and Methods
- Results
- Conclusions



BUSINESS PROBLEM

- WaterAid would like us to find areas with the highest repair needs
- Identify factors which contribute to non-functionality

DATA & METHODS

Data

- Tanzania Waterpoint Mapping data from 2012. datadriven.org
- 59.400 individual waterpoints
- Data son location, water quality, water quantity, payments and waterpoint types.

Methods

- Machine Learning
- Classification

RESULTS

Prediction Recall

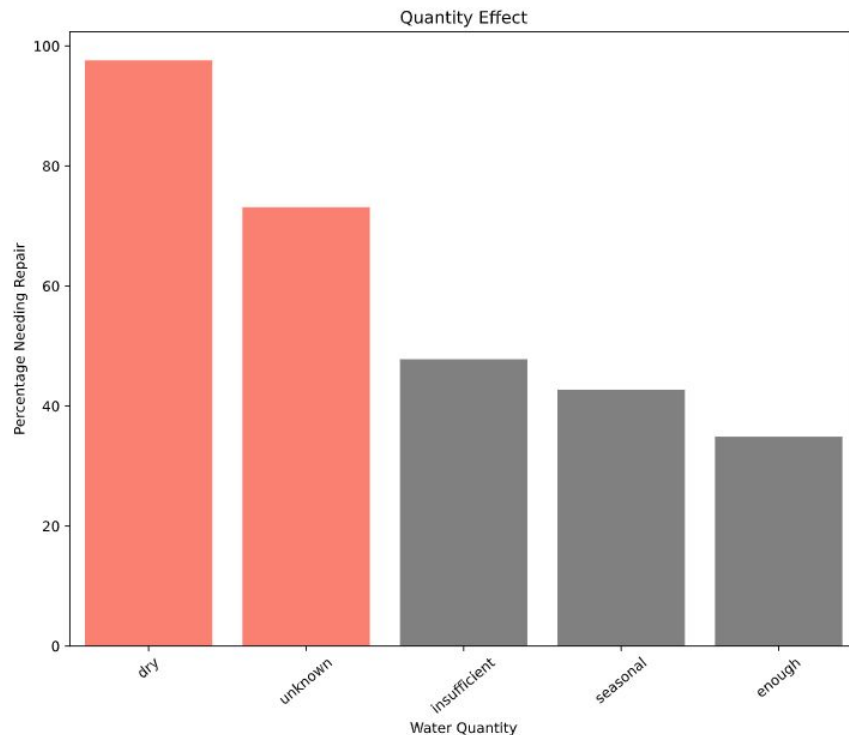
- Our model was able to correctly identify 77% of waterpoints in need of repair.

RECALL

77%

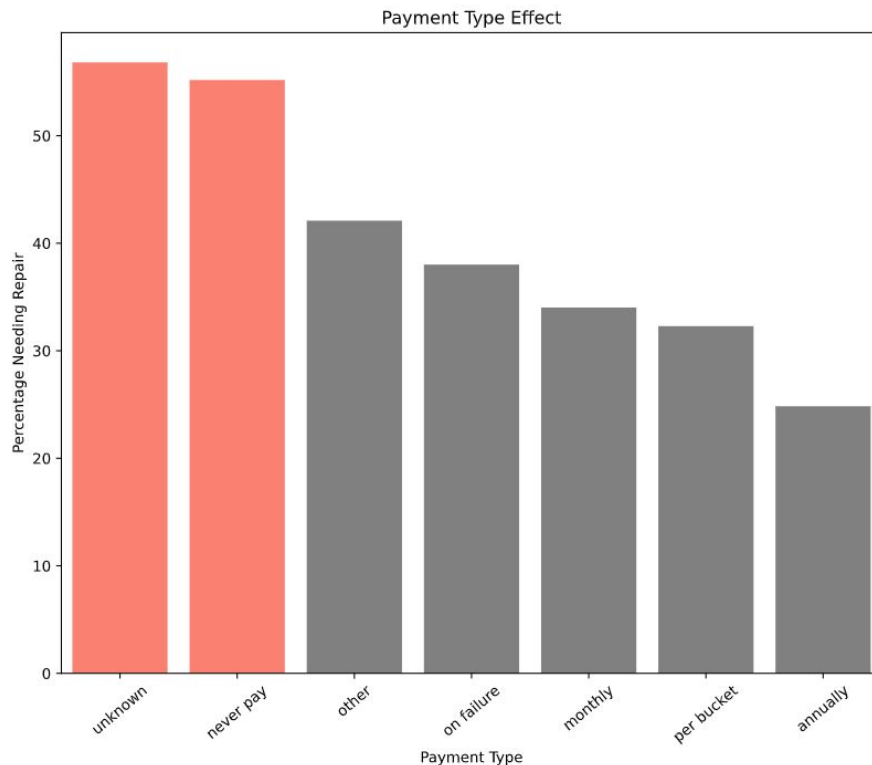
RESULTS

- Quantity of water plays a significant role.
- Waterpoints in 'dry' and 'unknown' areas need the most repair.
- Those areas should get priority



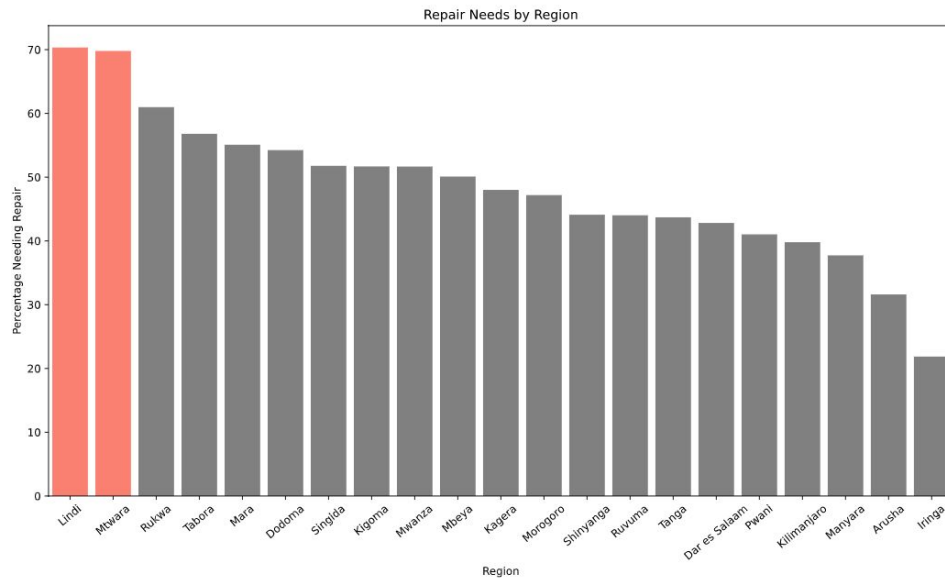
RESULTS

- Waterpoints that receive regular payments tend to stay more functional.
- Higher repair needs for 'never pay' and 'unknown'



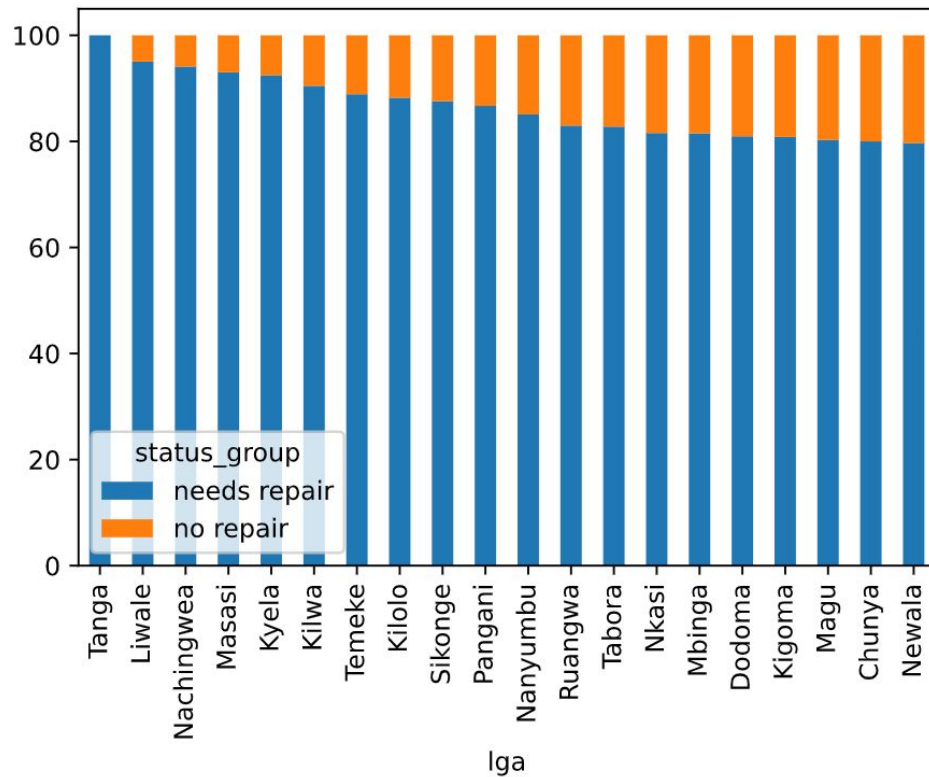
RESULTS

- Lindi and Mtwara have the highest needs.



RESULTS

- There are 126 Local Governing Authorities who are in charge of the waterpoints.
- Within areas of not enough water, these LGAs are most likely to need repair.
- These 20 have over 80% of their waterpoints in need of repair.



CONCLUSION

Recommendations

- Prioritize areas where water quantity is Dry
- Assist in setting up payment systems to fund repair and maintainance
- Regions of Lindi and Mtwara as well as the 20 LGAs shown require special attention.

NEXT STEPS

- Deal with missing data on latitude, longitude and elevation
- Differentiate between rural and urban waterpoints
- Obtain data on maintenance

Thank You

Questions?

Email: siddiqui@mohammed@gmail.com

GitHub: [@saifword](#)