

Tanzania Water Wells Status Prediction

15 November, 2023

MEET THE TEAM



Presentation Lead

Mytreyi Abburu

Email: abburumk@gmail.com

Github: myt-hue



Github Lead

Goknur Kaya

Email: goknurkaya@gmail.com

Github: goknurk



Technical Lead

Kari Primiano

Email: kkprim@gmail.com

Github: kkprim

BACKGROUND

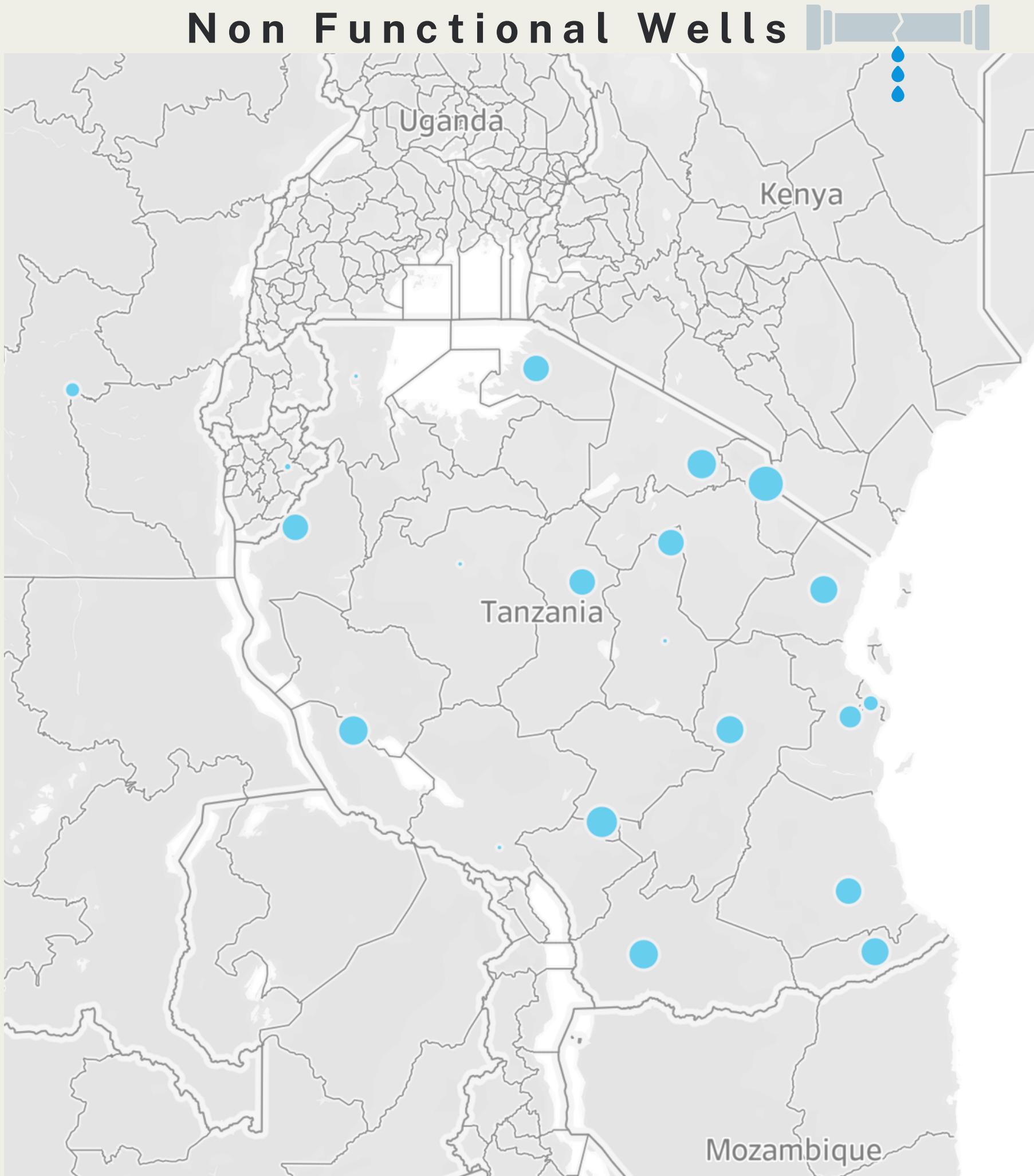
- Out of 59 million population, 55% of people do not have access to clean water. That is more than 30 million,
- Locals spend hours each day walking miles to collect buckets of water

A woman is shown from the side, carrying a large, round, light-colored water container balanced on her head. She is wearing a dark, patterned dress. The background is a soft-focus outdoor scene with trees and a building.

**TOGETHER, LET'S IMPROVE
TANZANIA WATER WELLS**

MAIN FINDING

Geographical indicators are the most important features in identifying non-functional wells at 21% MORE influential than other features



A G E N D A

- Business Problem
- Data Overview
- Analysis
- Modeling
- Recommendations
- Next Steps

BUSINESS PROBLEM



Tanzania faces challenges in ensuring access to clean water for its population due to non-functional water wells.

DATA OVERVIEW

DATA:

- Driven Data - Tanzanian Water Wells

FILTERS:

- Added Binary Target Column (1 = functional, 0 = non-functional)
- Removed Duplicates
- 58,706 datapoints
- 19 features

LIMITATIONS:

- We do not know what types of repairs are needed
- We were not able to calculate the well age, due to missing construction years
- Without full population information, we do not know the supply needs

ANALYSIS

TARGET:

- CLASS 0 = Non-Functional
- CLASS 1 = Functional & Needs Repair

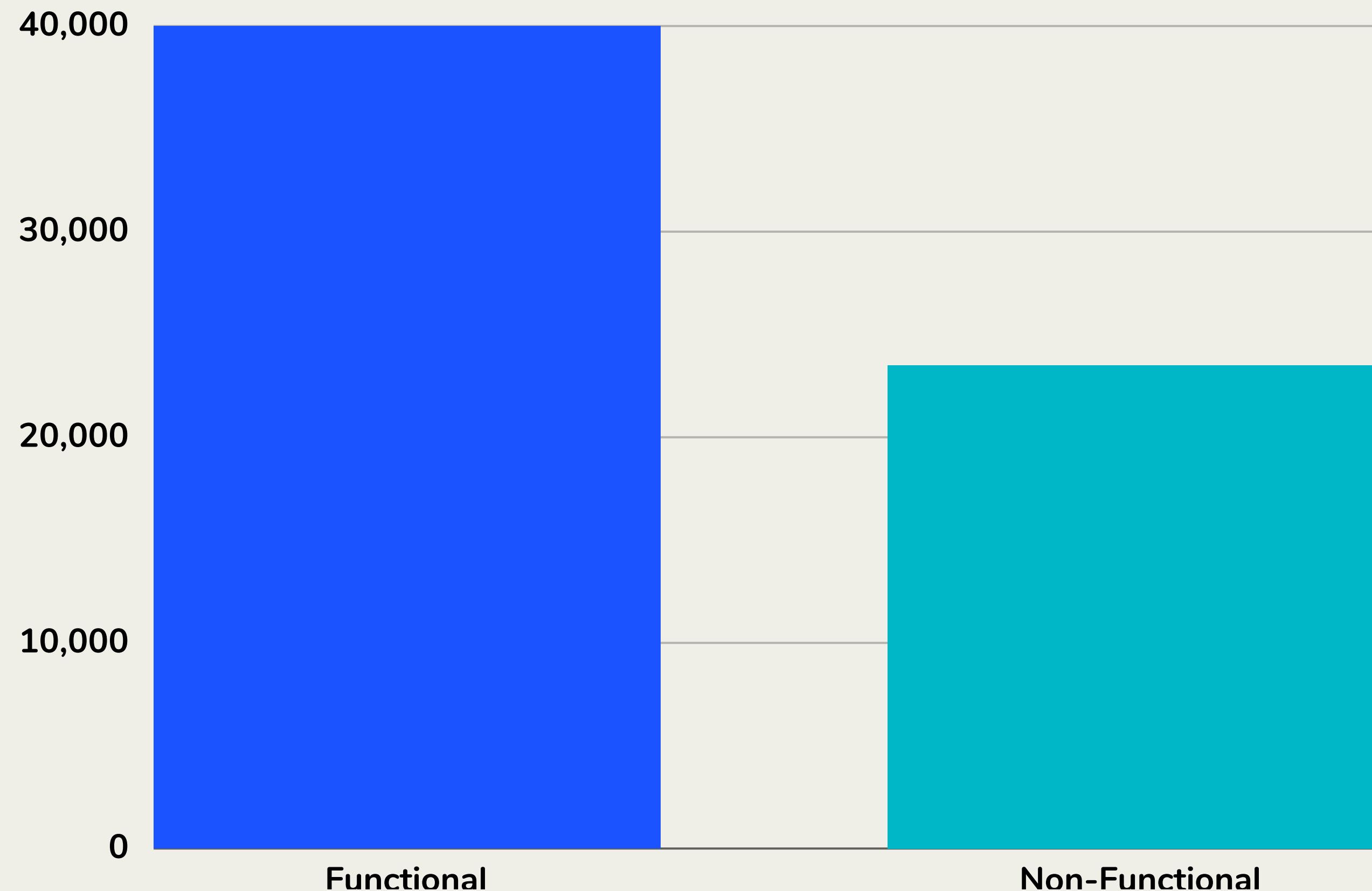
FEATURES:

- Assessed Well Type for Baseline Models

MODELS:

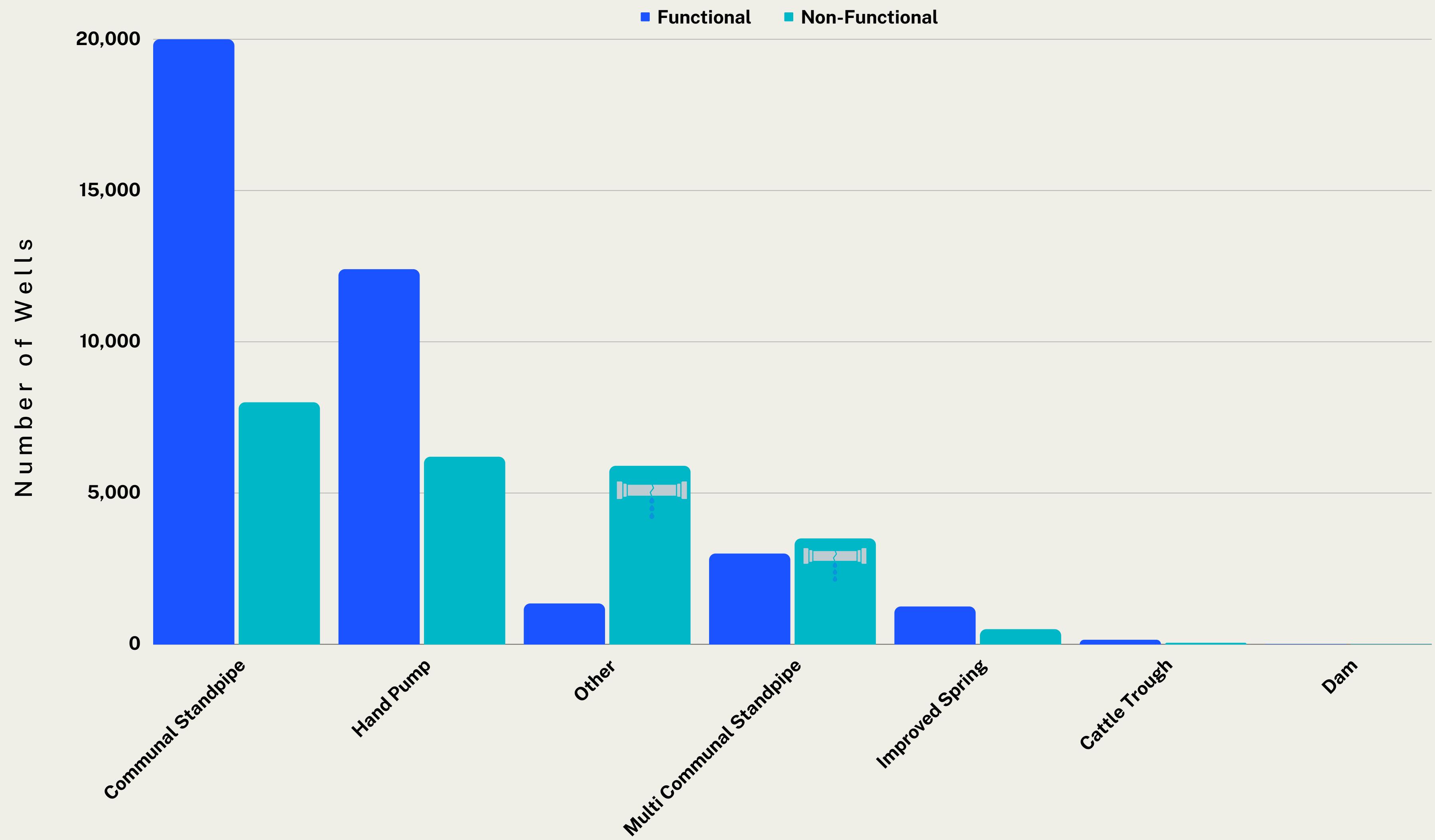
- Evaluated 4 Models to Predict Patterns in Non-Functional Wells

DISTRIBUTION OF WELL STATUS



TYPE OF WELLS

- “Other” has the highest percentage of non-functional wells at 81.38%
- Multi Communal Handpipes has the next highest percentage of non-functional wells at 53.85%
- Communal Standpipe is most likely to be a functional well



MODELING

- Recall is our main metric of model selection
- Our model predicts non-functionalities, while minimizing False negatives

BEST MODEL

- Random Forest Classifier

MODELS	PREDICTS ACCURATELY	PREDICTS NON-FUNCTIONAL	PREDICTS FUNCTIONAL
BASELINE 1: TARGET 1 = FUNCTIONAL 0 = NON/NEEDS REPAIR	65%	35%	90%
BASELINE 2: REVERSE TARGET 1 = FUNCTIONAL/NEEDS REPAIR 0 = NON-FUNCTIONAL	69%	37%	89%
RANDOM TREE	85%	77%	90%
LOGISTIC REGRESSION	78%	54%	93%

FEATURE IMPORTANCE

- Geographic indicators are the most important features in identifying non-functional wells
- 21% more influential than other features
- Potential causes include:
 - Climate
 - Altitude
 - Access to resources

= Highest number of non-functional wells

Region

Latitude

Longitude

Quantity

Dry

Enough

Insufficient

GPS Height

Altitude

Well Type

Other

Communal

Extraction Type

Other

Gravity

Payment Type

Not Paid

RECOMMENDATION

- Geographical indicators is the main predictor of non-functional wells
 - Mbeya
 - Morogoro
 - Kilimanjaro
- Communal Standpipe is most likely to be a functional well
- Invest in well infrastructure (payment type)

NEXT STEPS

- Investigate additional features that influence geography
 - Climate
 - Population
 - Amount of water available in the area (amount_tsh)
- Time-series analysis of well ages to predict lifetime of more robust well structures
- Local Government should look at what type of water wells needing repair

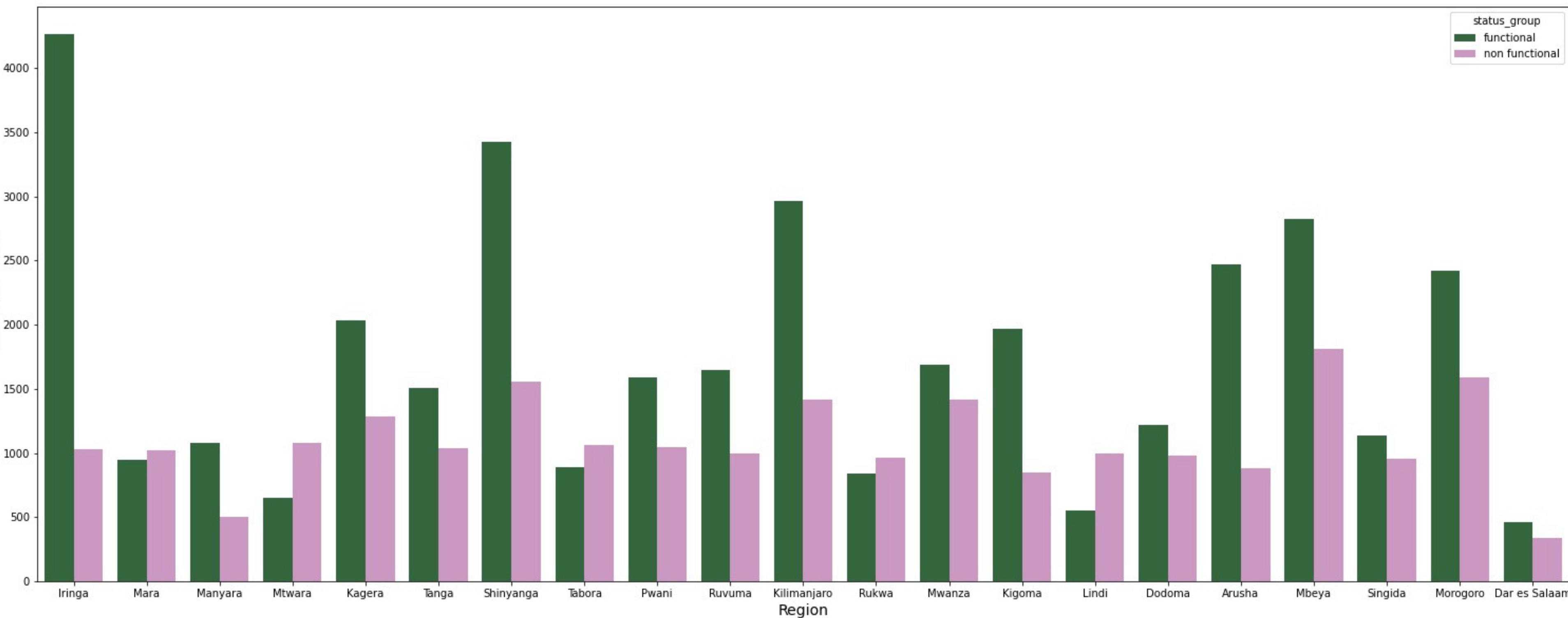
Thank you!

15 November, 2023

Appendix

15 November, 2023

Status of Wells by Region



EXTRACTION TYPE PER REGION

