

# Tanzanian Waterpoint Status Prediction

Enabling directed local action  
through machine learning

# Current Problem

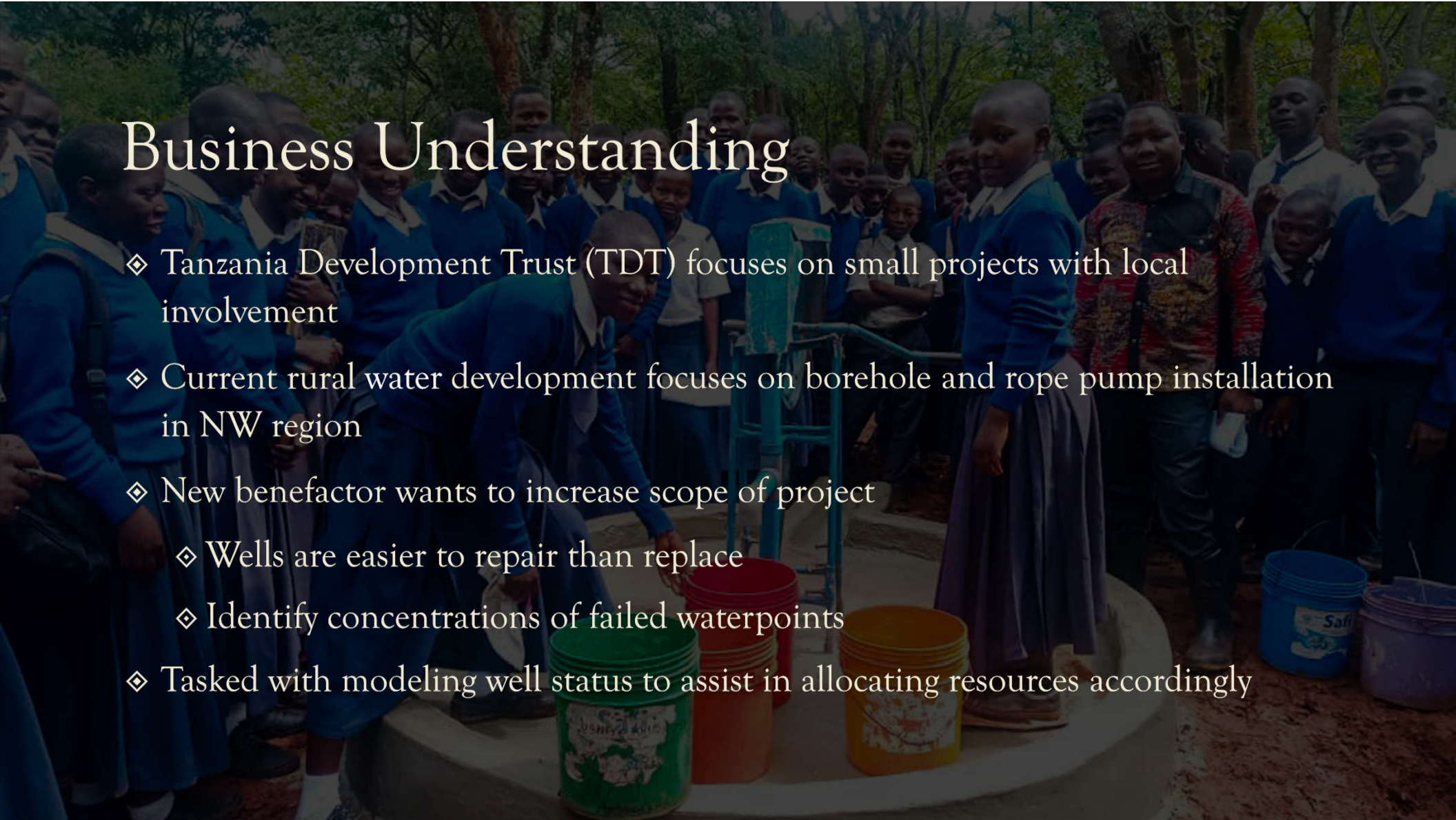
- 70% of the country's population lives in rural areas, where access to clean water covers only about 45% of the population
- International efforts are focused on improving this access through maintenance and new installations





# Business Understanding

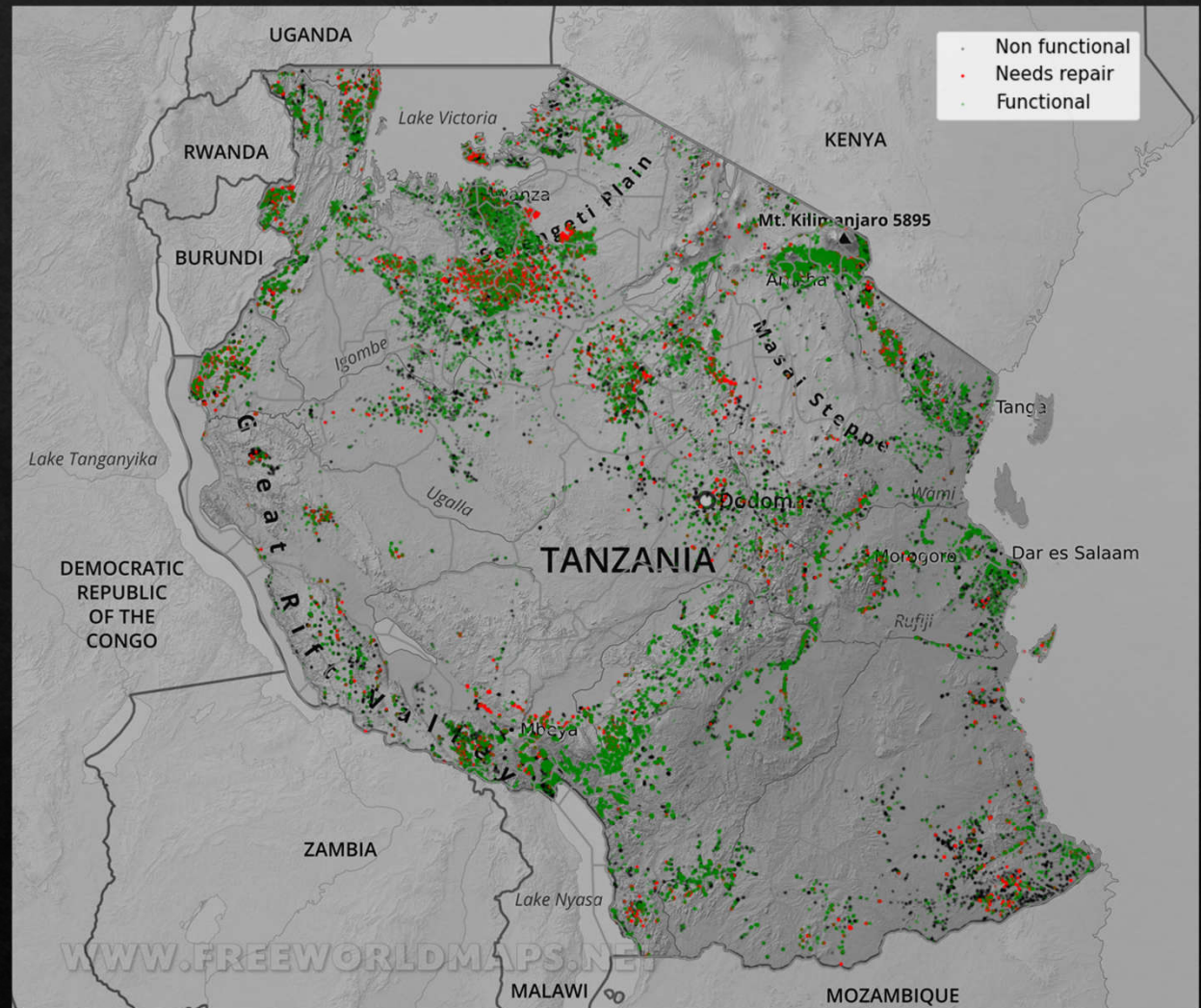
- ◆ Tanzania Development Trust (TDT) focuses on small projects with local involvement
- ◆ Current rural water development focuses on borehole and rope pump installation in NW region
- ◆ New benefactor wants to increase scope of project
  - ◆ Wells are easier to repair than replace
  - ◆ Identify concentrations of failed waterpoints
- ◆ Tasked with modeling well status to assist in allocating resources accordingly





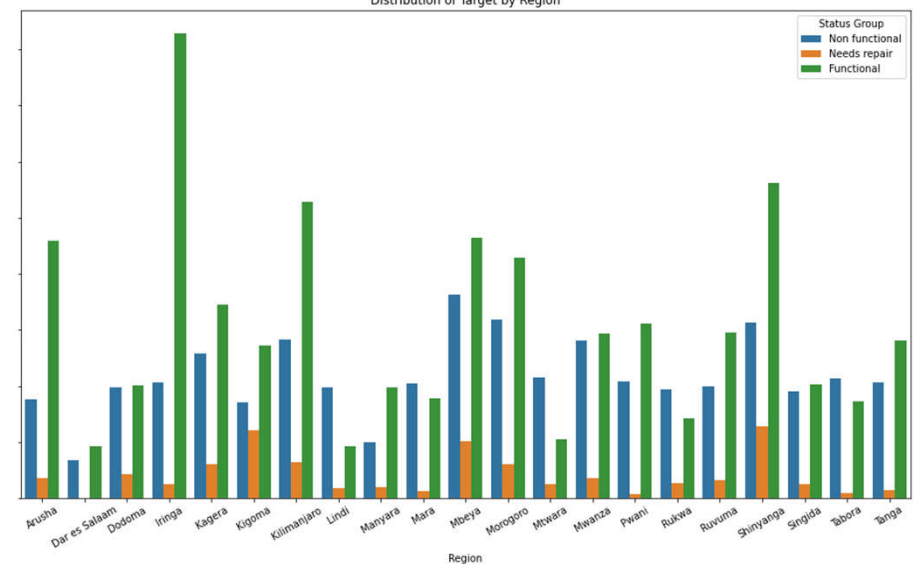
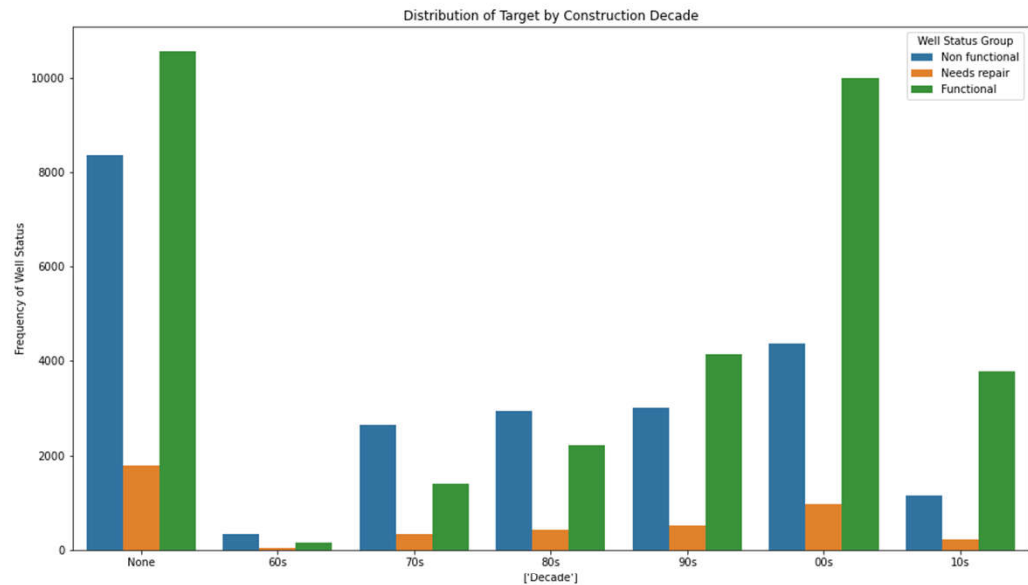
# Data Brief

- ◆ Data source: Taarifa waterpoints dashboard
  - ◆ aggregate data from the Ministry of Water
- ◆ Features include geographic data, extraction information, operational and source data
- ◆ 59,400 entries recorded between 2011 and 2013

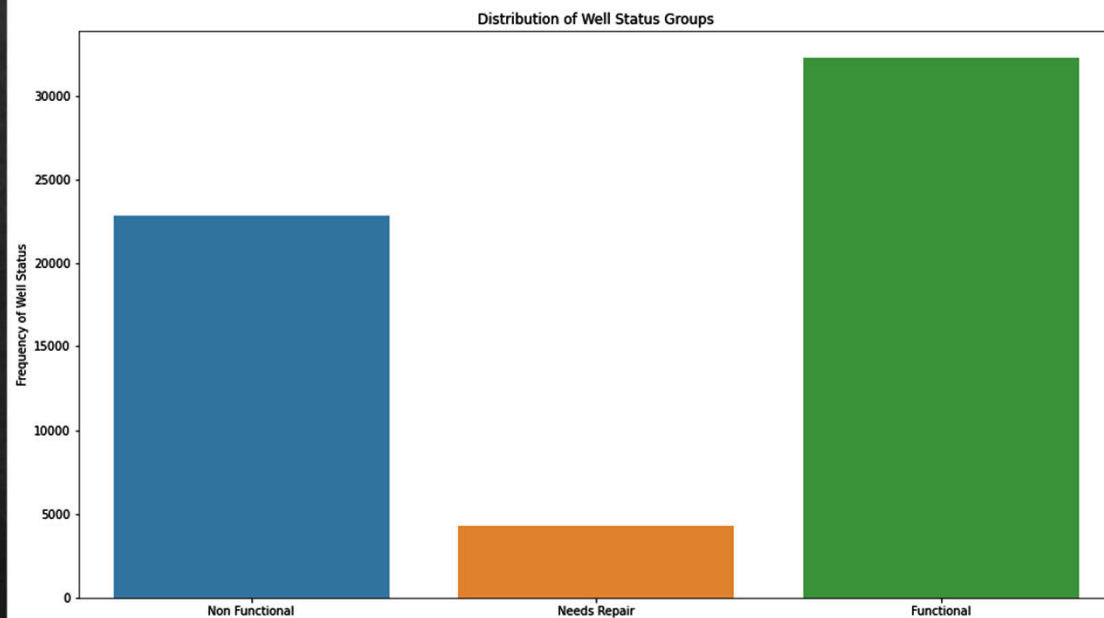


# Initial Insights

- Aging waterpoints have an increased share of non-functional status but a decreased needs repair status
- Regions in the northwest experience the highest share of waterpoints needing repair
- Southeast regions see the largest share of non-functional waterpoints



# Modeling Process



- Begin with a simple baseline model for comparison
- Iterative progression in successive models
  - GridSearch
  - Target weighting
- Classification reports and confusion matrices used for evaluating and comparing models



# Model Selection

Models struggled with identifying minority class

Functional wells misclassified as needing repair are still an opportunity for preventative maintenance



Selected XGBoost for classification modeling

Speed of processing makes it ideal for eventual production on a dashboard



Best metrics for finding repair sites

Other models offer higher total accuracy at the cost of misidentifying repair

# Feature Understanding



THE QUANTITY OF WATER  
AVAILABLE TO A WELL IS THE  
MOST IMPACTFUL FEATURE



PHYSICAL LOCATION ALSO HAS  
A LARGE IMPACT ON THE  
STATUS OF A WATERPOINT

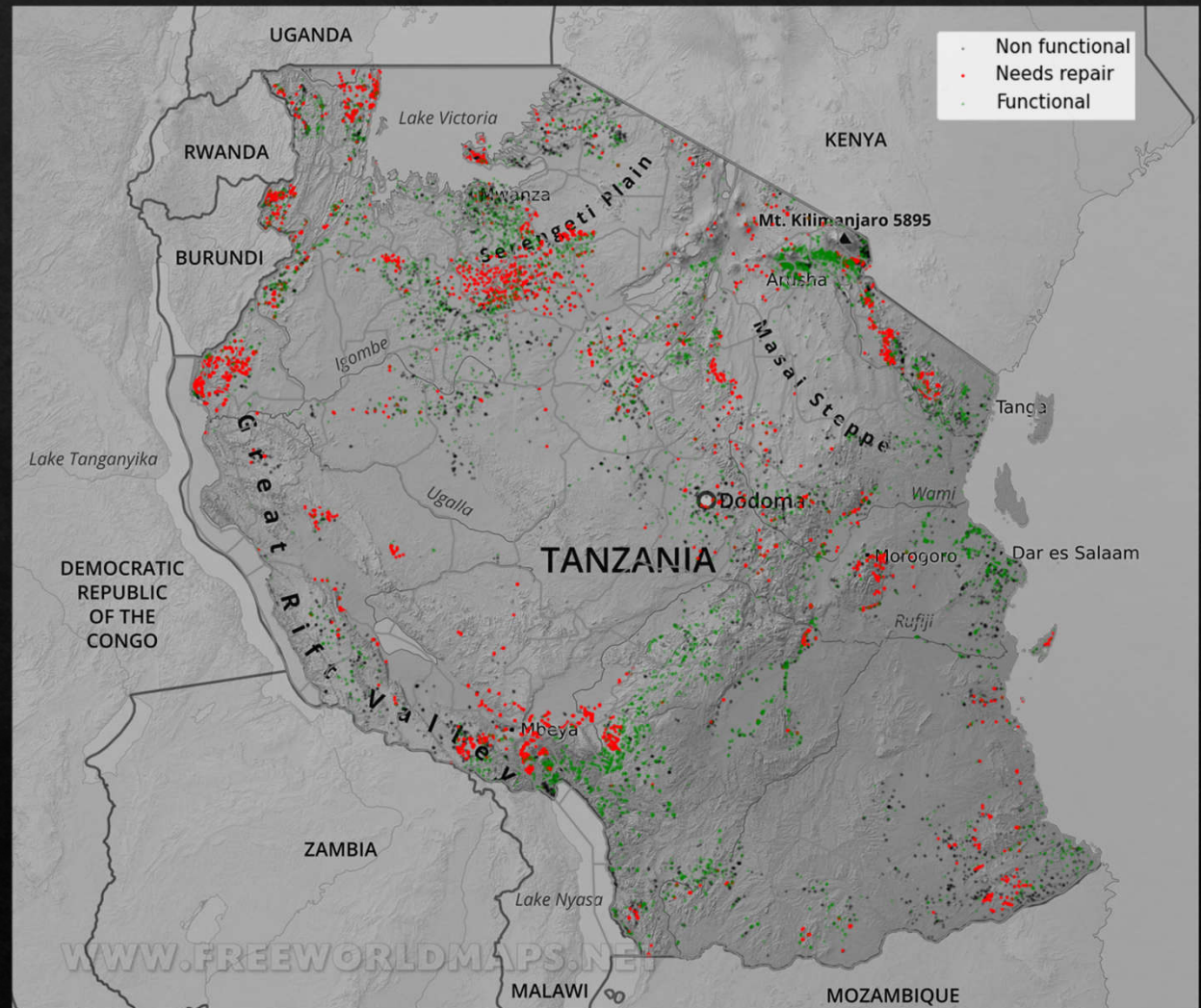


OTHER IMPACTFUL FEATURES  
INCLUDE EXTRACTION TYPE  
AND THE WELL HEIGHT



# Recommendations

- ◆ Expand well repair operations to Shinyanga
  - ◆ Future project expansion to Mbeya
- ◆ New well construction in Mara and Mwanza
  - ◆ Install handpump wells
  - ◆ Communal standpipes with multiple waterpoints



# Questions?



[HTTPS://GITHUB.COM/BRTRACY/PHASE\\_3\\_PROJECT](https://github.com/brtracy/phase_3_project)



[BRTRACY1984@GMAIL.COM](mailto:BRTRACY1984@GMAIL.COM)