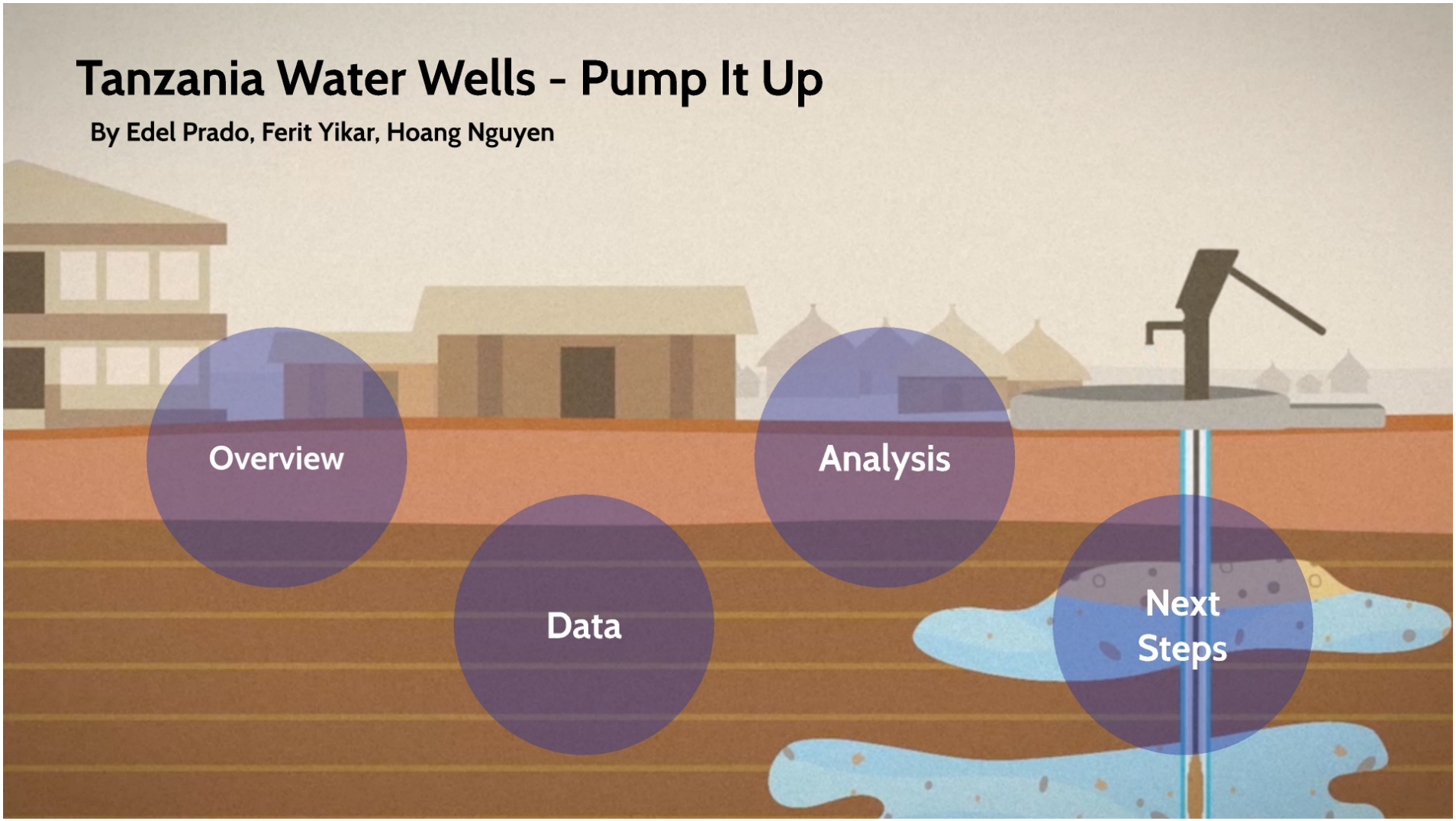

Tanzania Water Wells - Pump It Up
By Edel Prado, Ferit Yikar, Hoang Nguyen
Overview
Analysis
Data
Next
Steps

Overview

- Background
- Business Problem
- Data
- Modeling
- Further Analysis
- Recommendations

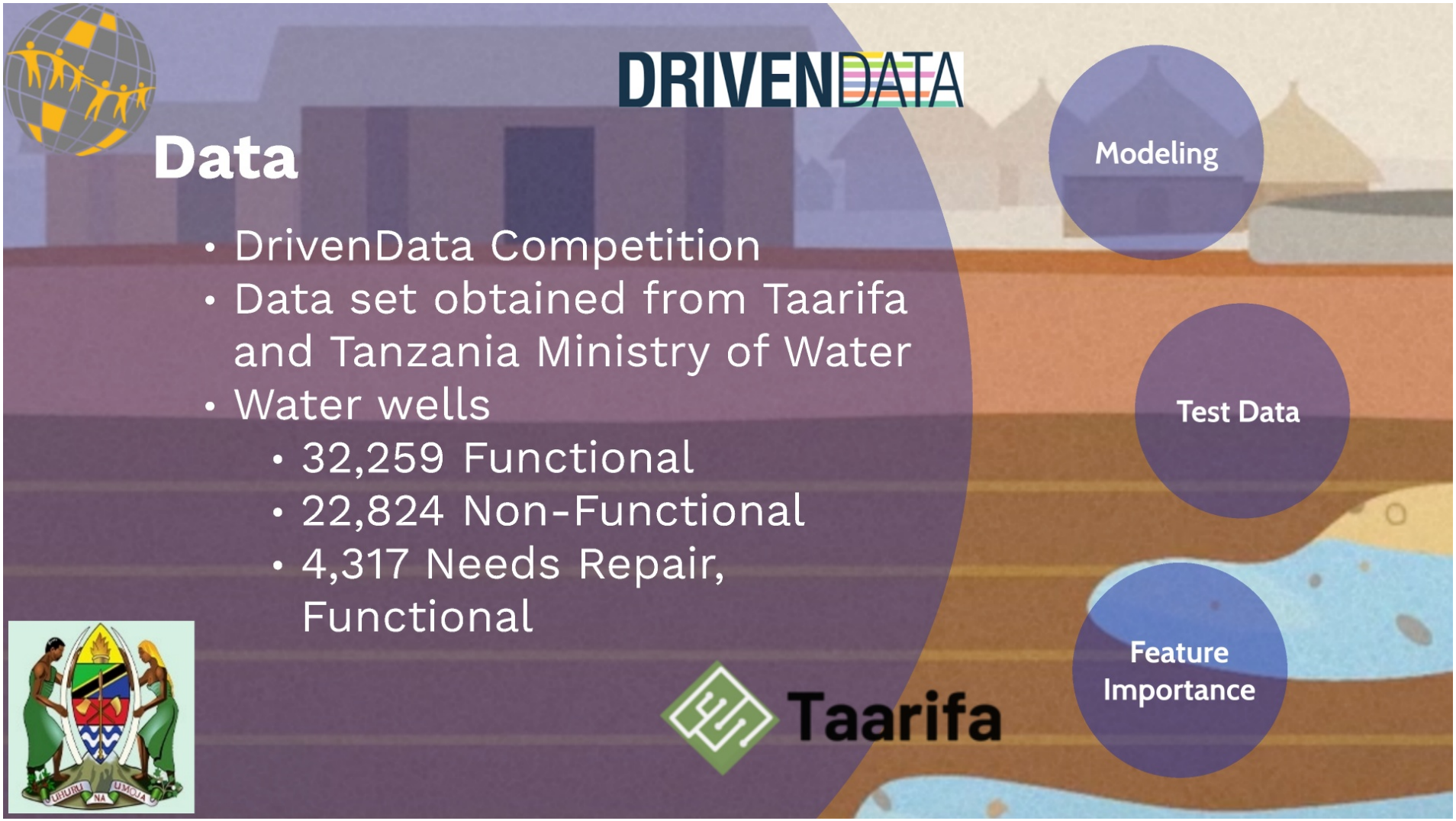
Background

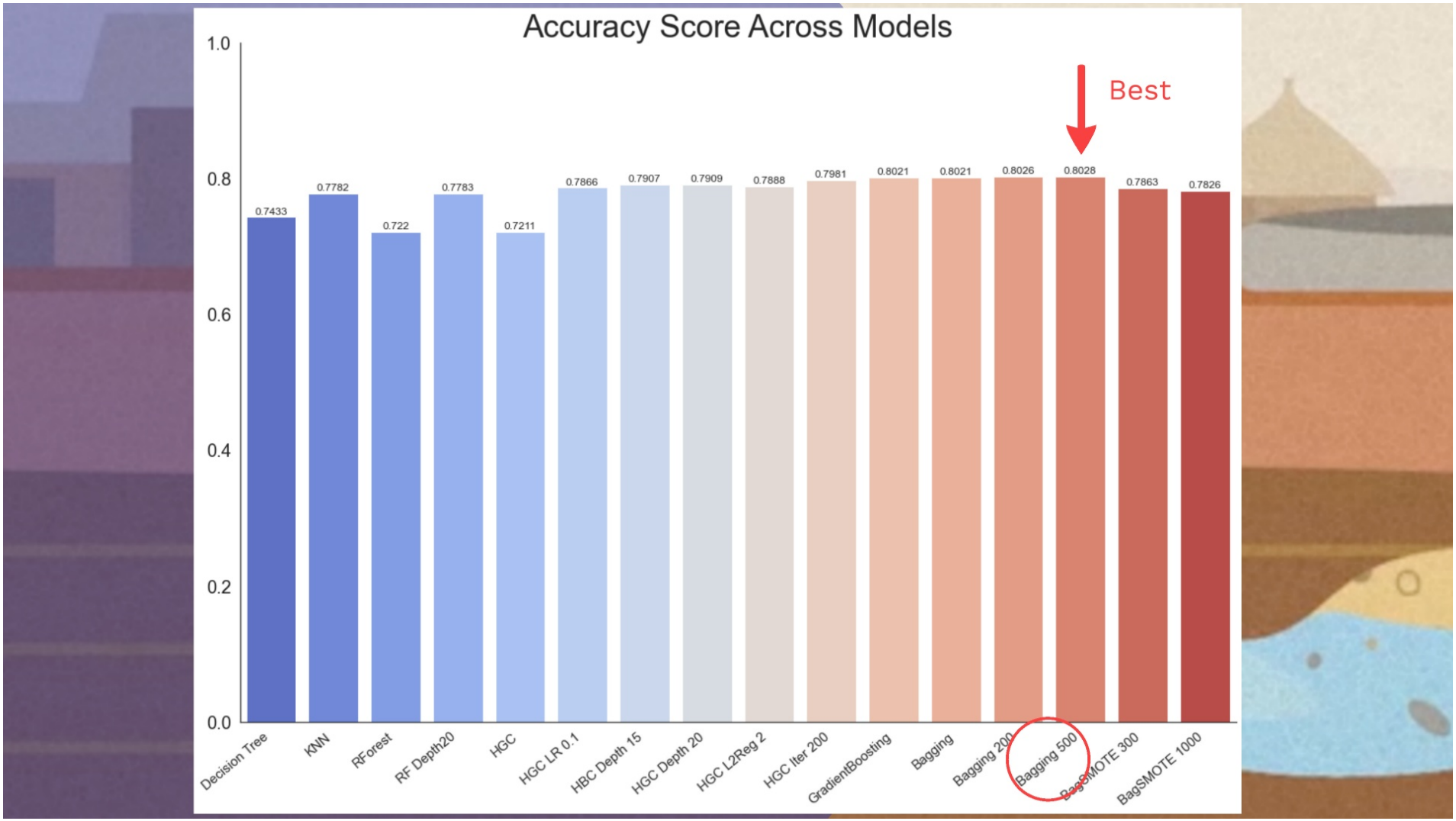
Business Understanding

Background

- Population of 59.73 million
- 36% of total population live on less than $3.20 a day
- 4 million lack access to an improved source of safe water
- 29 million do not have accessed to improved sanitation


unicef
Business Problem
UNICEF is planning on starting a water well campaign in Tanzania. They need to understand which wells need repairs and which need to be rebuilt. After that they will decide how many more wells they should build.


DRIVENDATA
Data
• DrivenData Competition
• Data set obtained from Taarifa and Tanzania Ministry of Water
• Water wells
• 32,259 Functional
• 22,824 Non-Functional
• 4,317 Needs Repair, Functional
Taarifa
Modeling
Test Data
Feature Importance

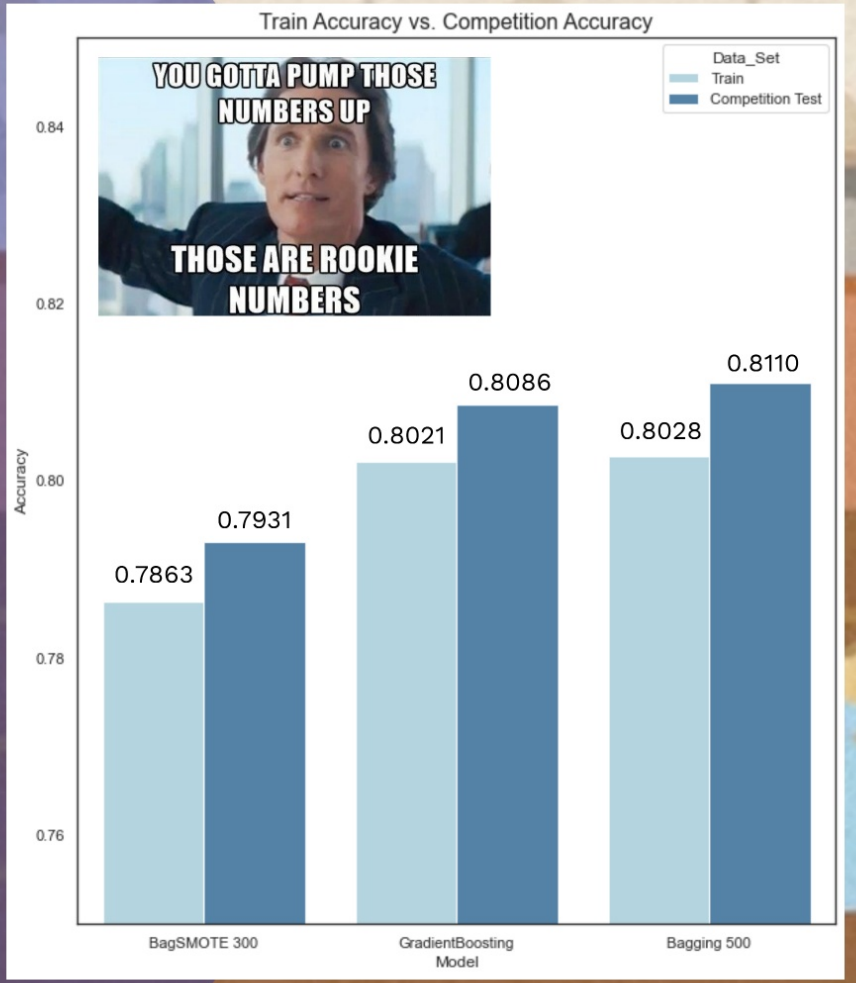

Accuracy Score Across Models
1.0
0.8
0.6
0.4
0.2
0.0
Best
0.7433
0.7782
0.722
0.7783
0.7211
0.7866
0.7907
0.7909
0.7888
0.7981
0.8021
0.8021
0.8026
0.8028
0.7863
0.7826
Decision Tree
KNN
RForest
RF Depth20
HGC
HGC LR 0.1
HBC Depth 15
HGC Depth 20
HGC L2Reg 2
HGC Iter 200
GradientBoosting
Bagging
Bagging 200
Bagging 500
BagSMOTE 300
BagSMOTE 1000

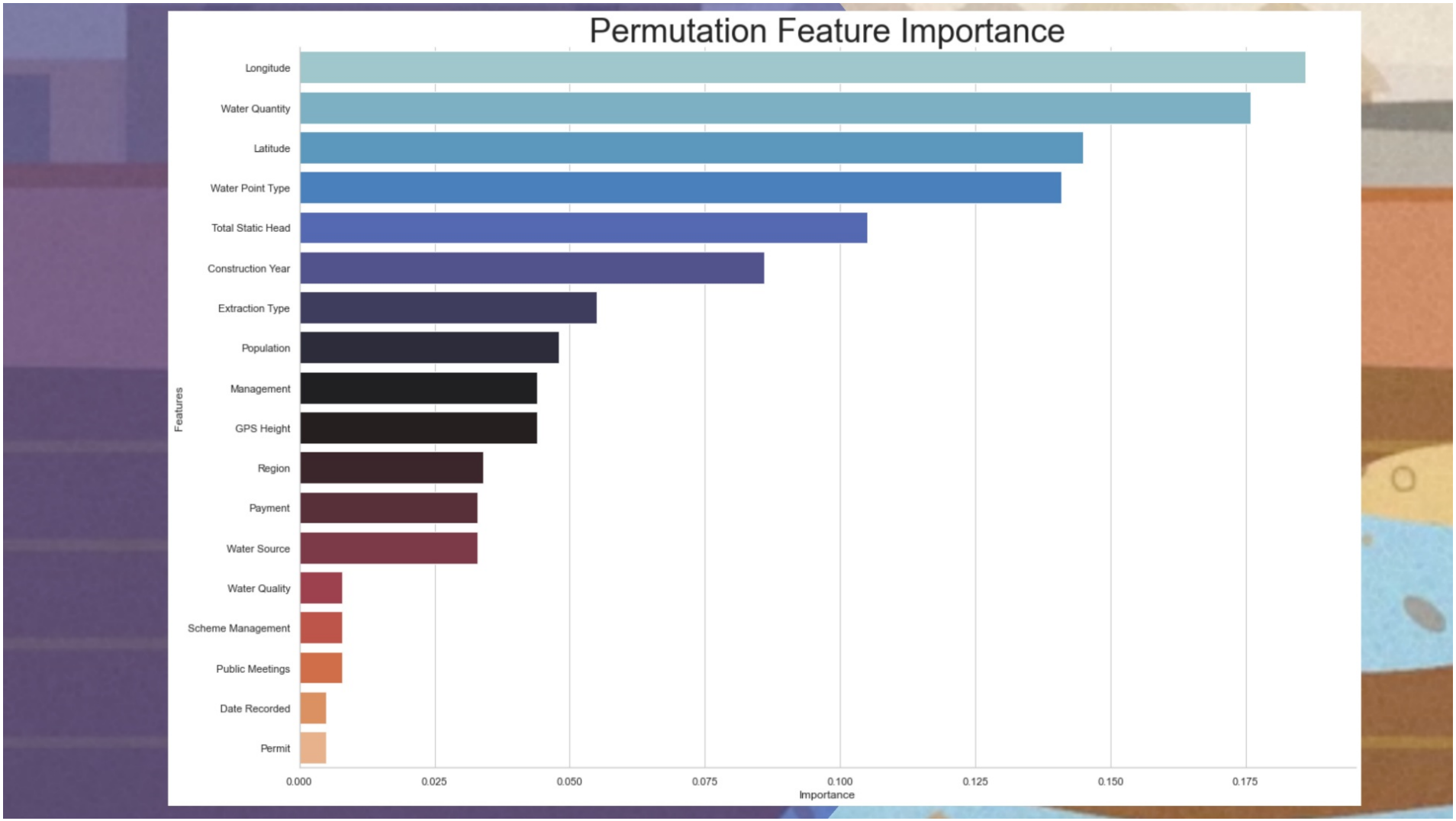
Scoring

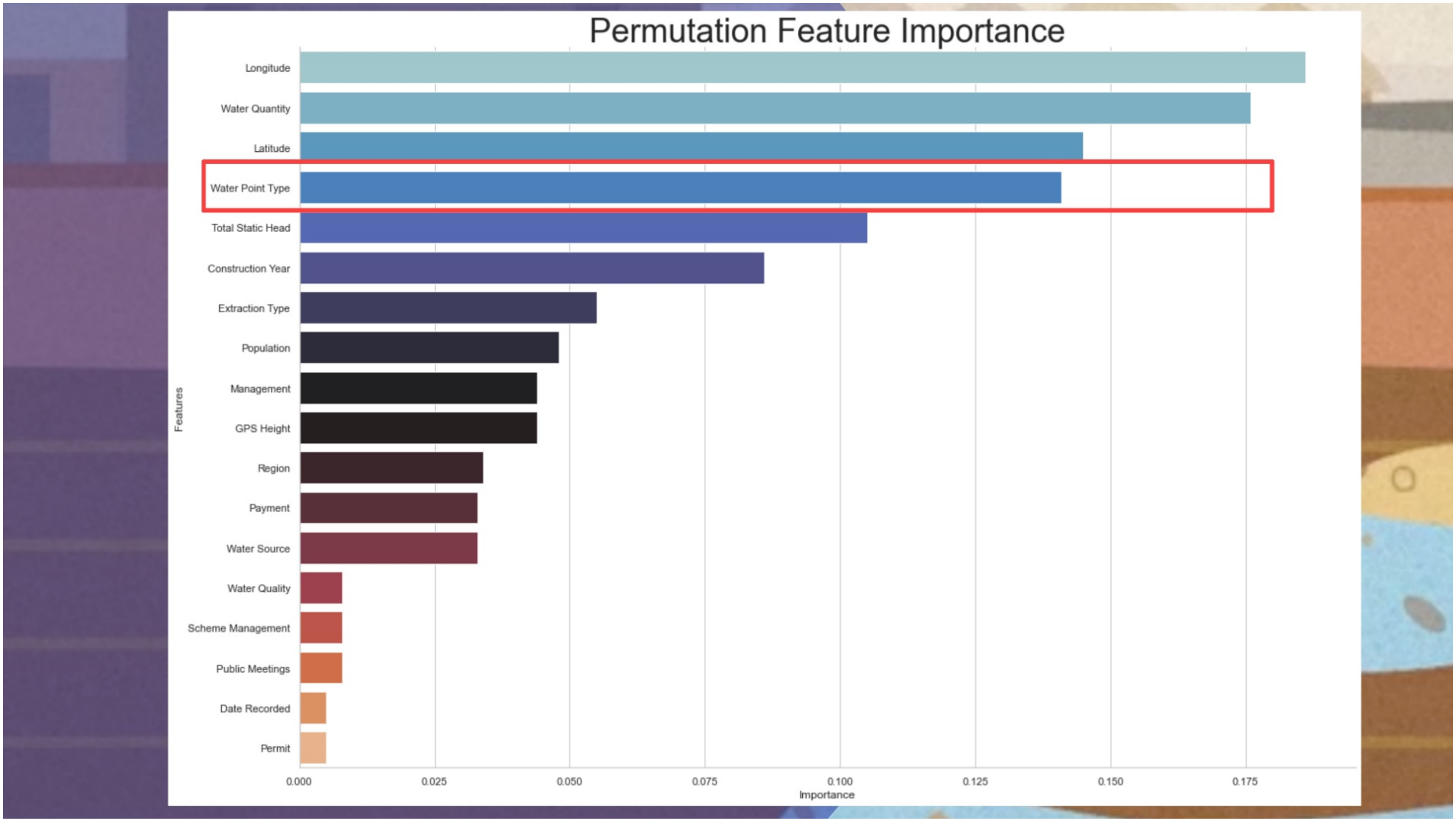
Bootstrap Aggregating

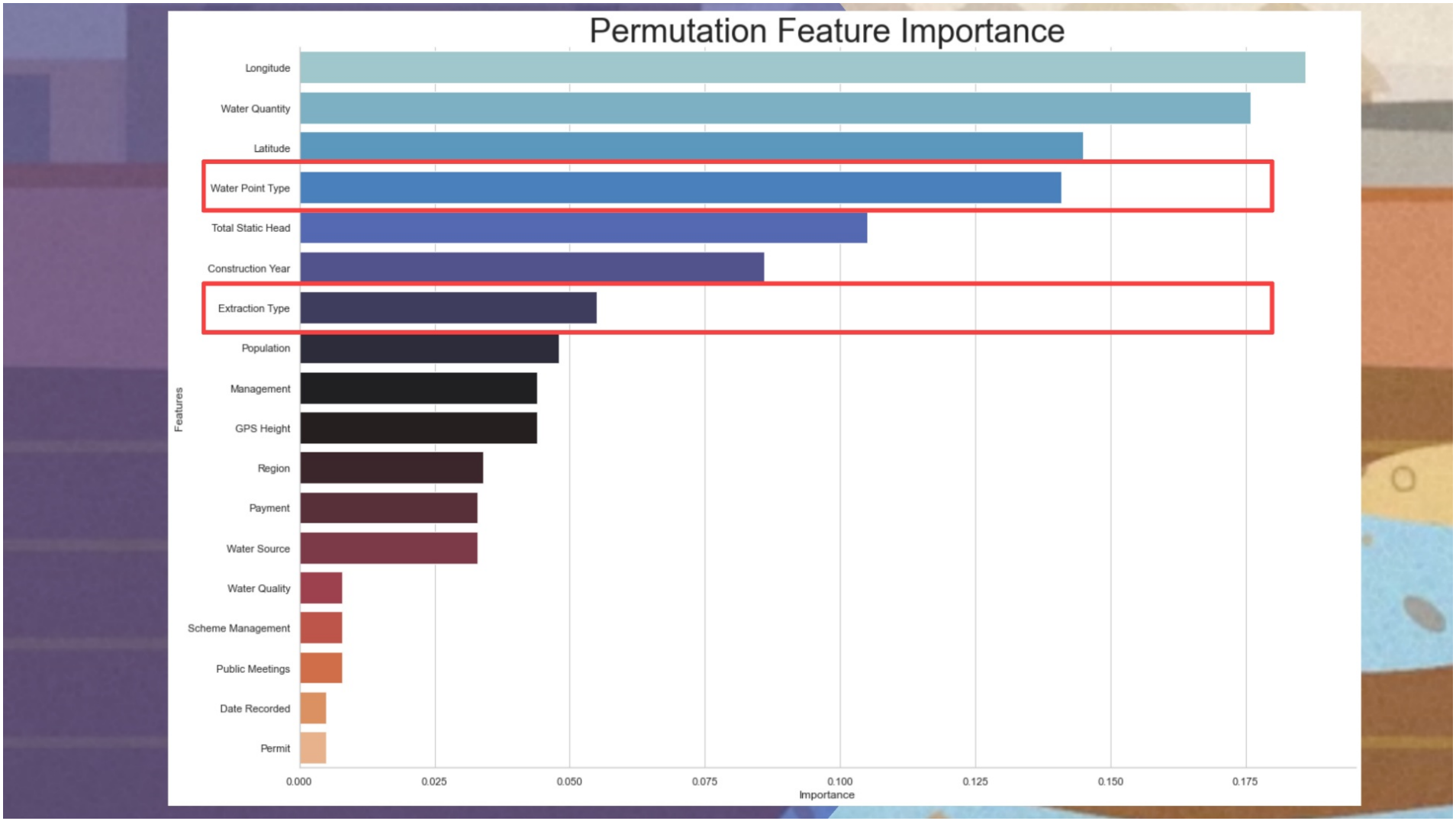
n_estimators = 500

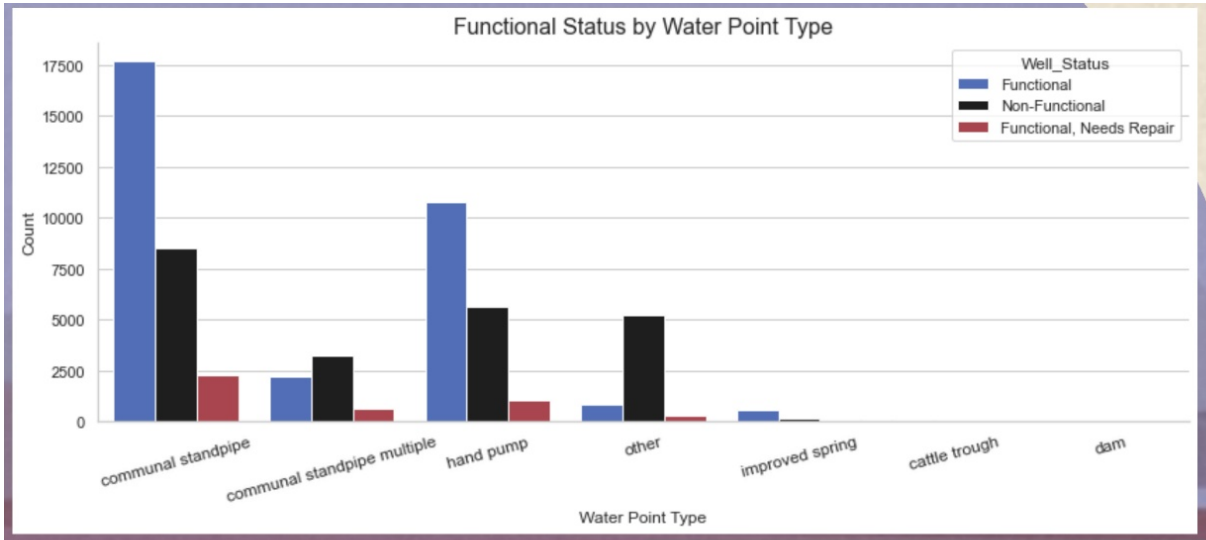
BEST	CURRENT RANK	# COMPETITORS
0.8110	2336	12877

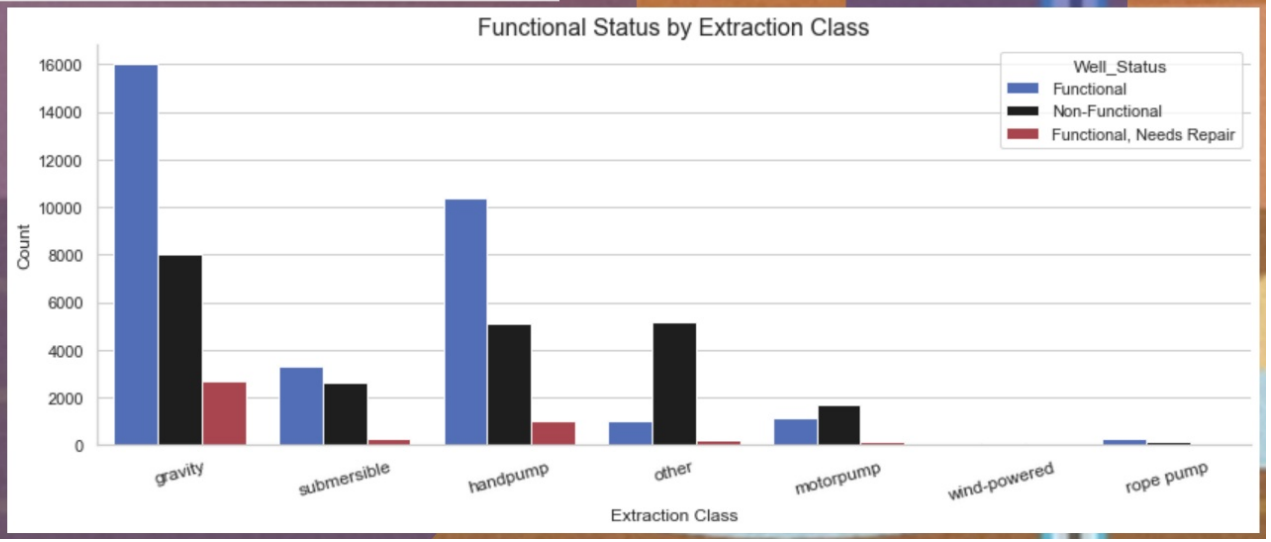


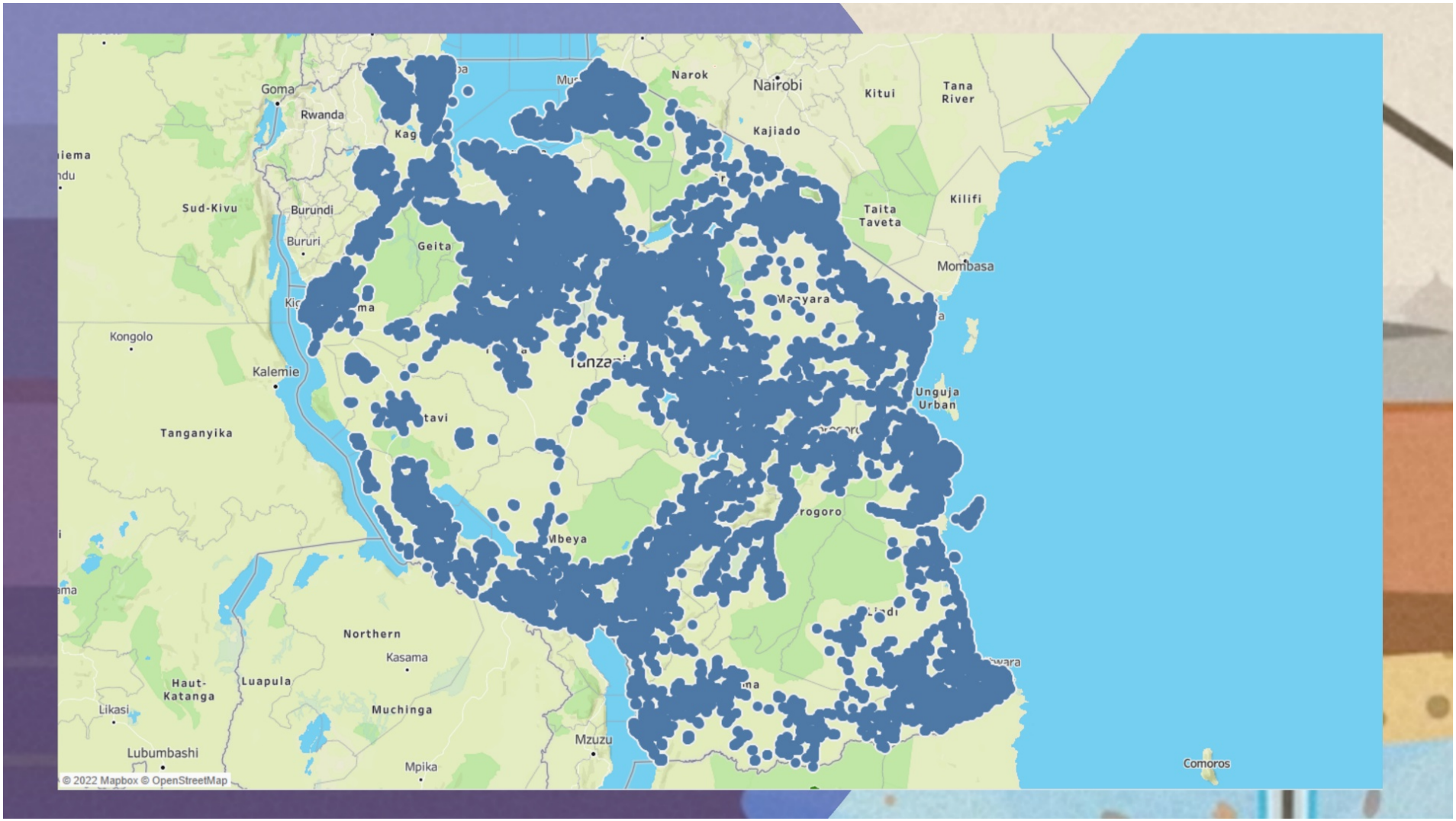

Permutation Feature Importance
Longitude
Water Quantity
Latitude
Water Point Type
Total Static Head
Construction Year
Extraction Type
Population
Management
GPS Height
Region
Payment
Water Source
Water Quality
Scheme Management
Public Meetings
Date Recorded
Permit
Features
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
Importance

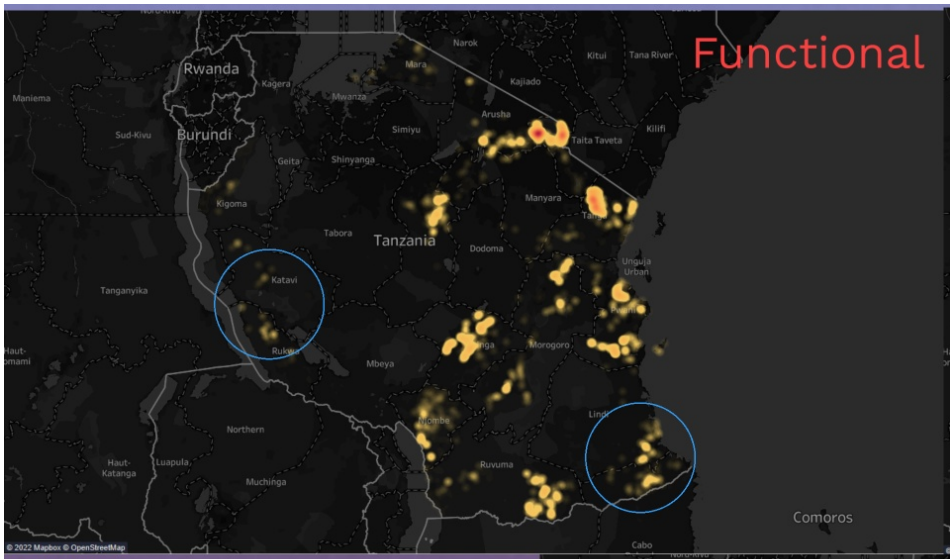

Permutation Feature Importance
Longitude
Water Quantity
Latitude
Water Point Type
Total Static Head
Construction Year
Extraction Type
Population
Management
GPS Height
Region
Payment
Water Source
Water Quality
Scheme Management
Public Meetings
Date Recorded
Permit
Features
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
Importance

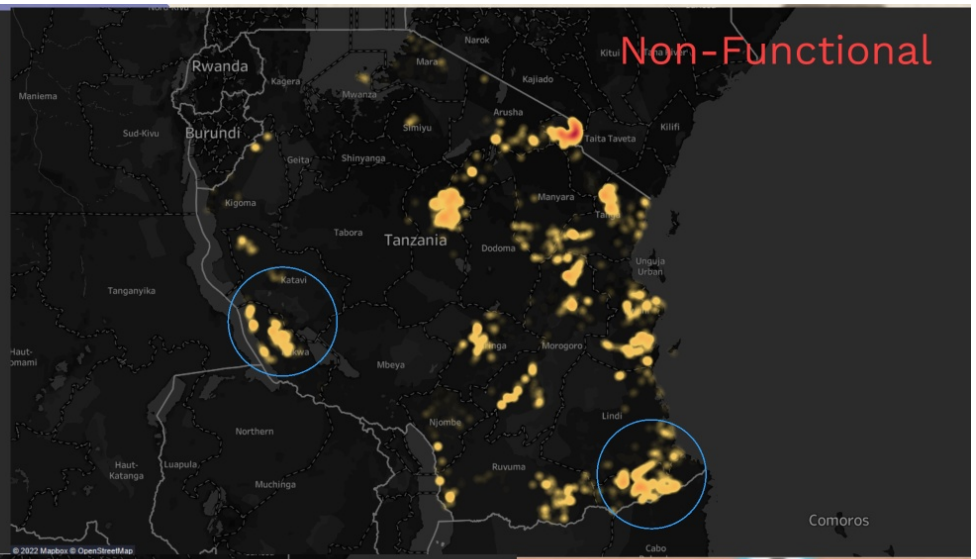

Permutation Feature Importance
Longitude
Water Quantity
Latitude
Water Point Type
Total Static Head
Construction Year
Extraction Type
Population
Management
GPS Height
Region
Payment
Water Source
Water Quality
Scheme Management
Public Meetings
Date Recorded
Permit
Features
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
Importance

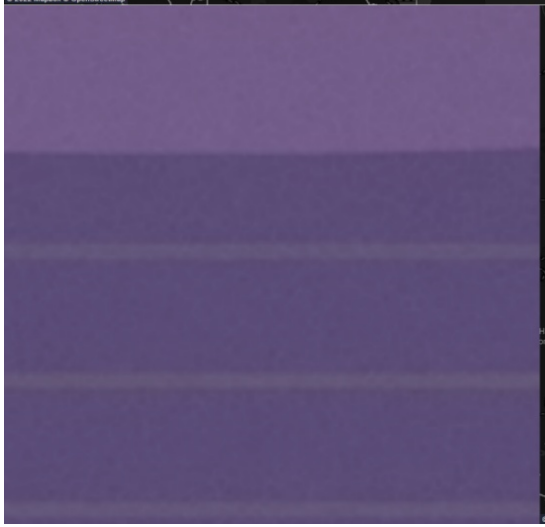

Functional Status by Water Point Type
Well_Status
Functional
Non-Functional
Functional, Needs Repair
Count
17500
15000
12500
10000
7500
5000
2500
0
communal standpipe
communal standpipe multiple
hand pump
other
improved spring
cattle trough
dam
Water Point Type

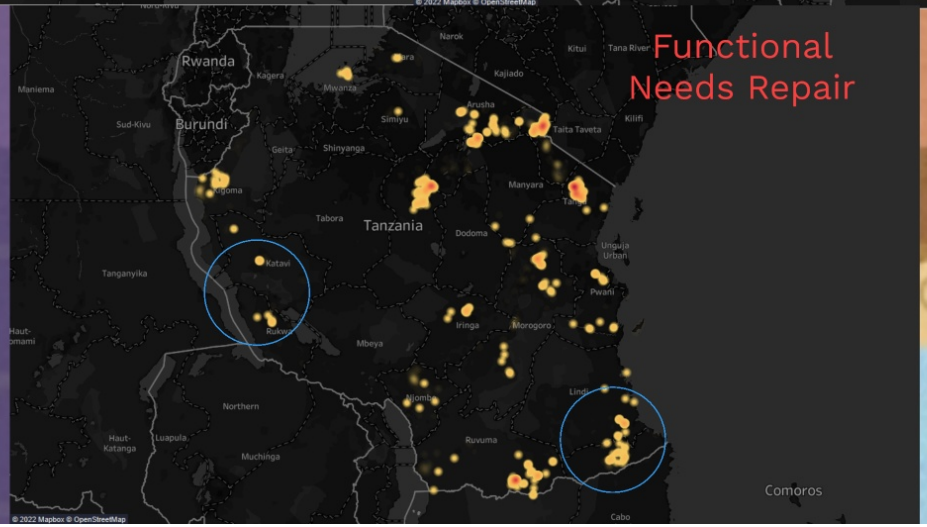

Functional Status by Extraction Class
Well_Status
Functional
Non-Functional
Functional, Needs Repair
Count
16000
14000
12000
10000
8000
6000
4000
2000
0
gravity
submersible
handpump
other
motorpump
wind-powered
rope pump
Extraction Class

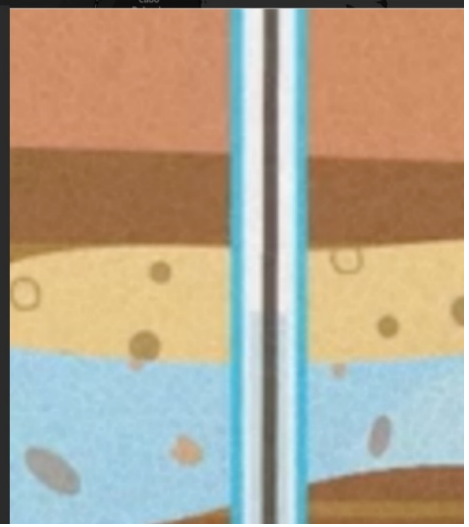

Goma
Rwanda
Narok
Nairobi
Kitui
Tana River
Kajiado
Kilifi
Taita Taveta
Mombasa
Sud-Kivu
Burundi
Bururi
Geita
Kongolo
Kalemie
Tanganyika
Unguja Urban
Northern
Kasama
Haut-Katanga
Luapula
Likasi
Muchinga
Lubumbashi
Mpika
Mzuzu
Comoros
© 2022 Mapbox © OpenStreetMap

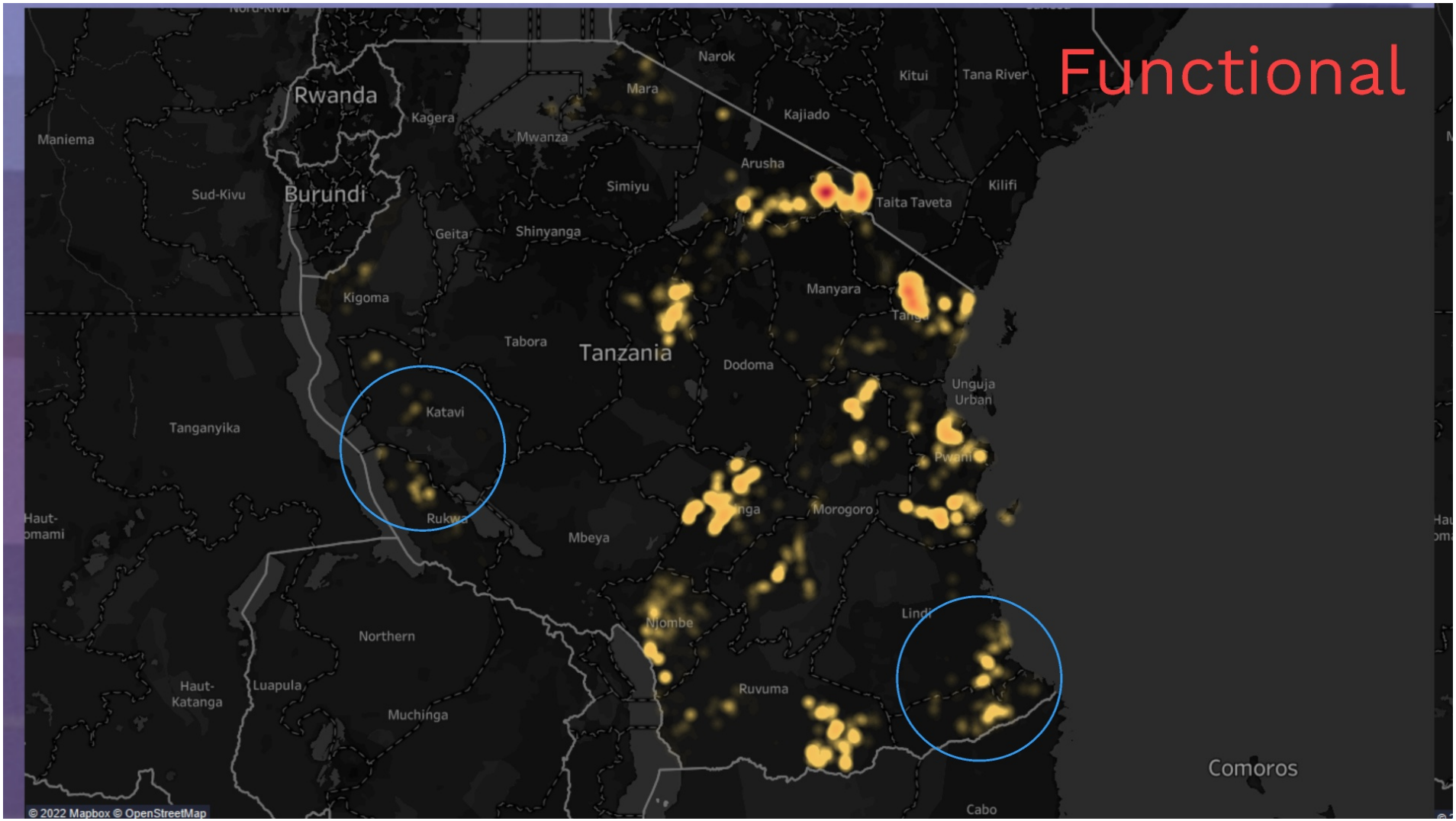

Functional
Rwanda
Burundi
Tanzania
Comoros
Kigoma
Katavi
Rukwa
Mbeya
Morogoro
Tabora
Dodoma
Manyara
Arusha
Taita Taveta
Kilifi
Kajiado
Tanganyika
Northern
Ruvuma
Muchinga
Luapula
Haut-Katanga
Sud-Kivu
Kagera
Mwanza
Simiyu
Shinyanga
Geita
Kitui
Tana River
Narok
Lindi
Cabo
© 2022 Mapbox © OpenStreetMap

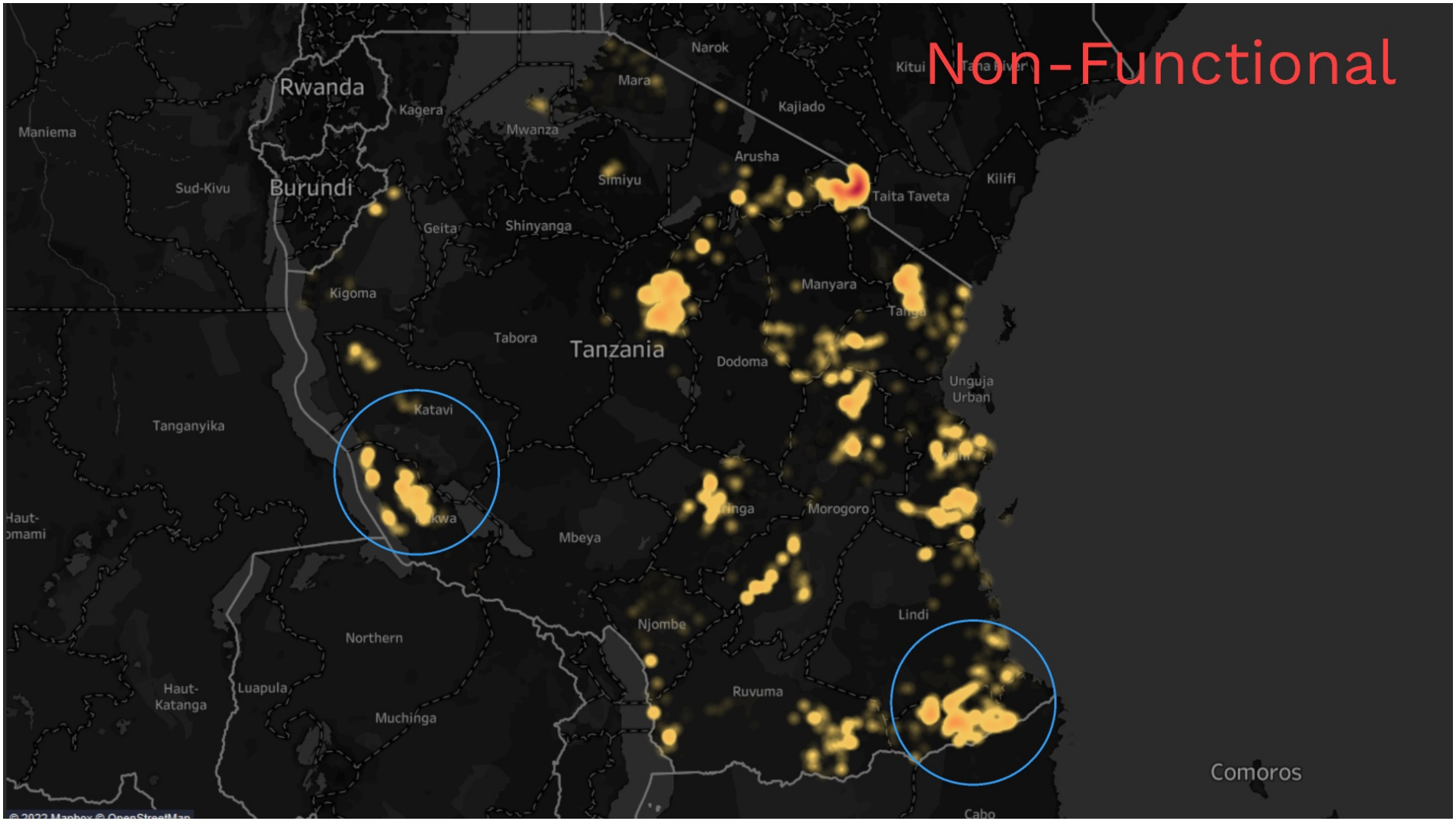

Non-Functional
Rwanda
Burundi
Tanzania
Comoros
Kigoma
Katavi
Mbeya
Morogoro
Tabora
Dodoma
Manyara
Arusha
Taita Taveta
Kilifi
Kajiado
Tanganyika
Northern
Njombe
Ruvuma
Muchinga
Luapula
Haut-Katanga
Sud-Kivu
Kagera
Mwanza
Simiyu
Shinyanga
Geita
Kitui
Narok
Lindi
Cabo
© 2022 Mapbox © OpenStreetMap

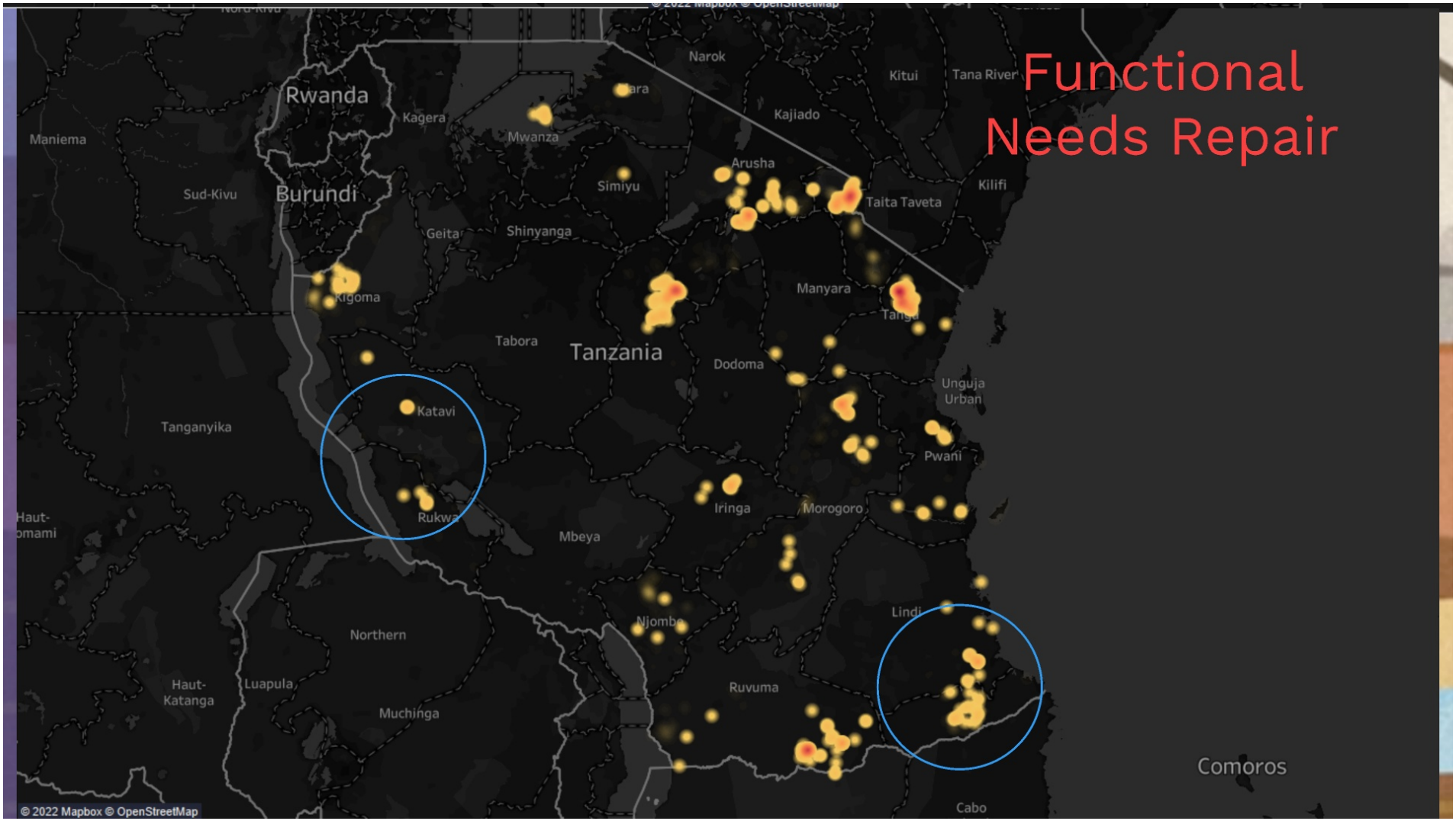


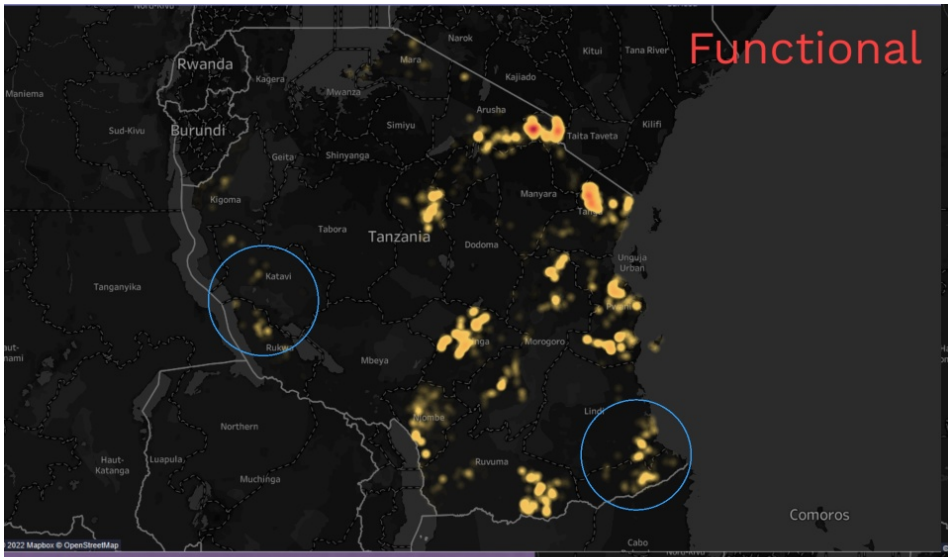

Functional
Needs Repair
Rwanda
Burundi
Tanzania
Comoros
Kigoma
Katavi
Rukwa
Mbeya
Iringa
Morogoro
Tabora
Dodoma
Manyara
Arusha
Taita Taveta
Kilifi
Kajiado
Pwani
Tanganyika
Northern
Njombe
Ruvuma
Muchinga
Luapula
Haut-Katanga
Sud-Kivu
Kagera
Mwanza
Simiyu
Shinyanga
Geita
Kitui
Tana River
Narok
Lindi
Cabo
© 2022 Mapbox © OpenStreetMap

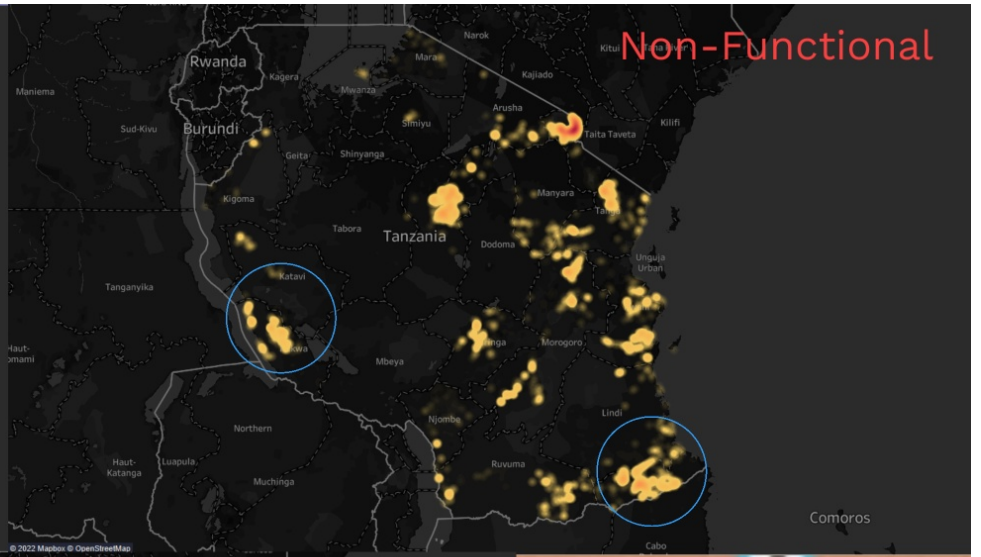


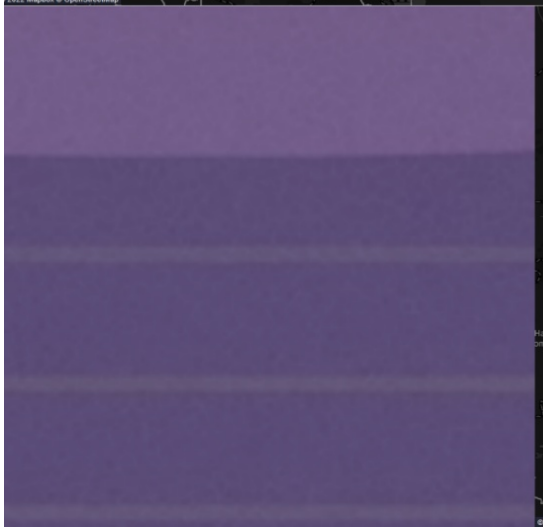

Functional
Rwanda
Burundi
Tanzania
Kagera
Mwanza
Mara
Narok
Kajiado
Kitui
Tana River
Arusha
Taita Taveta
Kilifi
Simiyu
Geita
Shinyanga
Kigoma
Tabora
Manyara
Dodoma
Unguja Urban
Pwani
Morogoro
Katavi
Rukwa
Mbeya
Njombe
Ruvuma
Lindi
Northern
Luapula
Muchinga
Haut-Katanga
Tanganyika
Sud-Kivu
Maniema
Comoros
Cabo
© 2022 Mapbox © OpenStreetMap

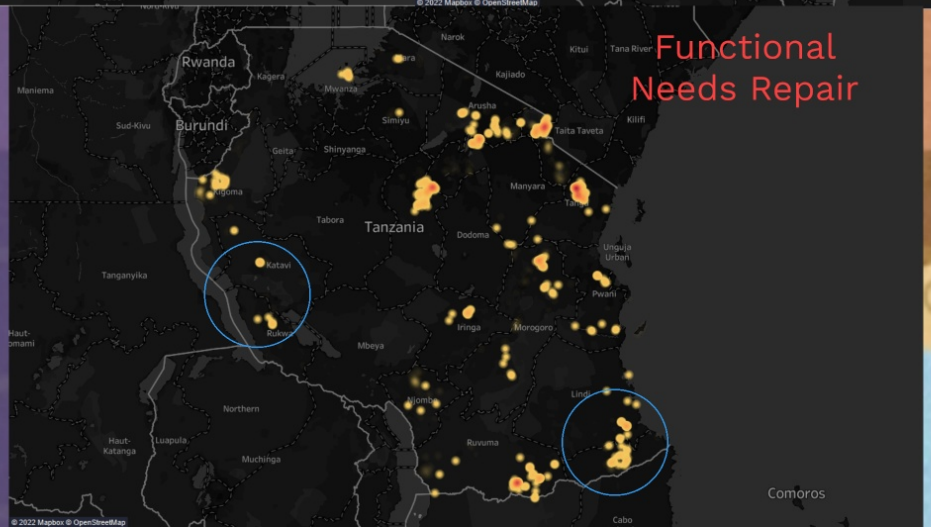

Non-Functional
Rwanda
Burundi
Tanzania
Kagera
Mwanza
Mara
Narok
Kajiado
Kitui
Arusha
Taita Taveta
Kilifi
Simiyu
Geita
Shinyanga
Kigoma
Tabora
Manyara
Dodoma
Unguja Urban
Katavi
Tanganyika
Morogoro
Mbeya
Njombe
Lindi
Ruvuma
Northern
Luapula
Muchinga
Haut-Katanga
Maniema
Sud-Kivu
Comoros

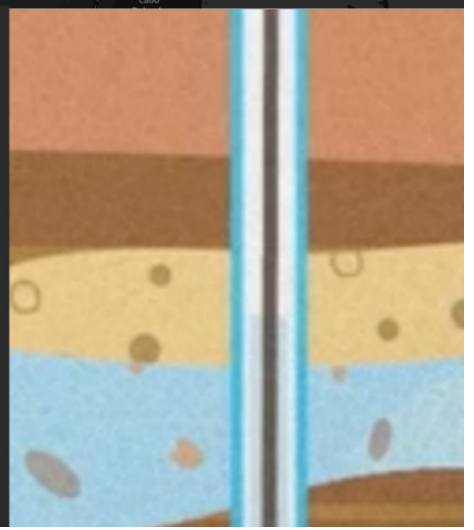

Functional
Needs Repair
Rwanda
Burundi
Tanzania
Kagera
Mwanza
Narok
Kajiado
Kitui
Tana River
Kilifi
Taita Taveta
Arusha
Manyara
Simiyu
Geita
Shinyanga
Kigoma
Tabora
Dodoma
Unguja
Urban
Pwani
Iringa
Morogoro
Mbeya
Njombe
Lindi
Ruvuma
Katavi
Rukwa
Tanganyika
Northern
Luapula
Haut-
Katanga
Muchinga
Maniema
Sud-Kivu
Comoros
Cabo
© 2022 Mapbox © OpenStreetMap


Functional
Rwanda
Burundi
Tanzania
Comoros


Non-Functional
Rwanda
Burundi
Tanzania
Comoros




Functional
Needs Repair
Rwanda
Burundi
Tanzania
Comoros

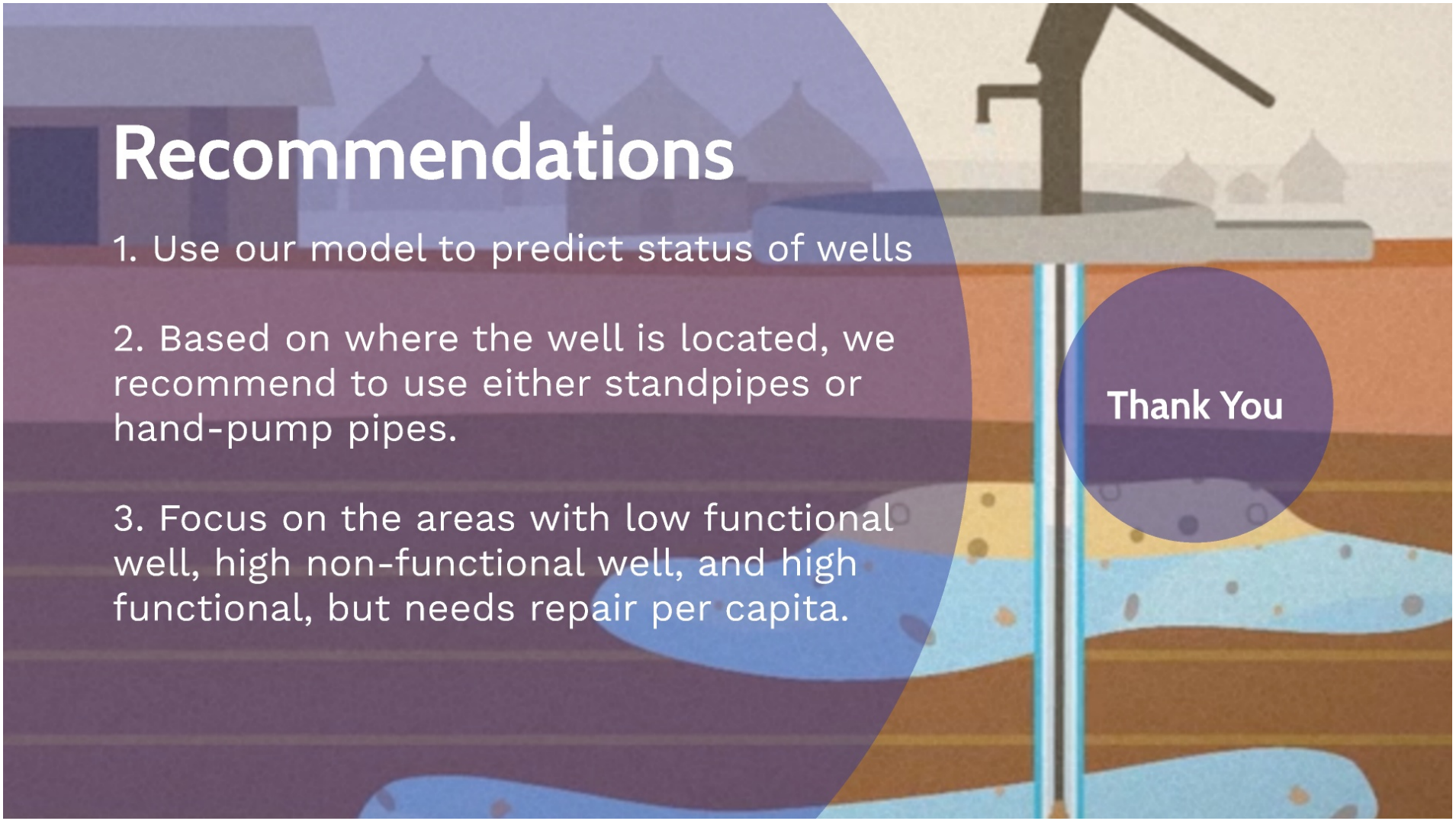


Recommendations

1. Use our model to predict status of wells

2. Based on where the well is located, we recommend to use either standpipes or hand-pump pipes.

3. Focus on the areas with low functional well, high non-functional well, and high functional, but needs repair per capita.

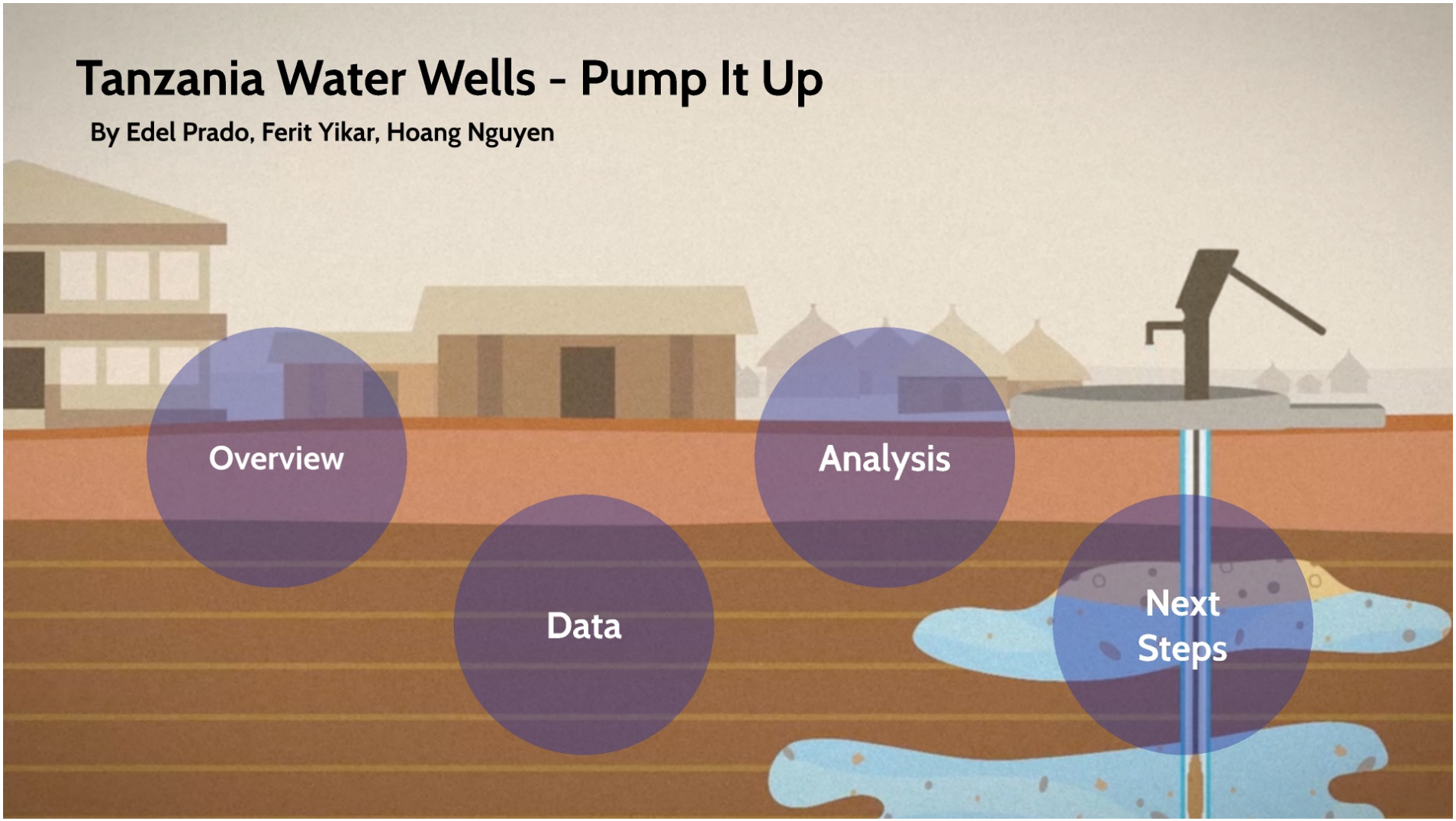


Thank You!

GitHub @ FeritYikar
yikarferit@gmail.com

GitHub @ epradojr
edel.prado.jr@gmail.com

GitHub @ DataDoggo215
hvnguyen90@gmail.com


Tanzania Water Wells - Pump It Up
By Edel Prado, Ferit Yikar, Hoang Nguyen
Overview
Analysis
Data
Next
Steps