



# Predicting Functionality of Tanzanian Water Wells

Evan Holder

Competition: [datadriven.org](https://datadriven.org)

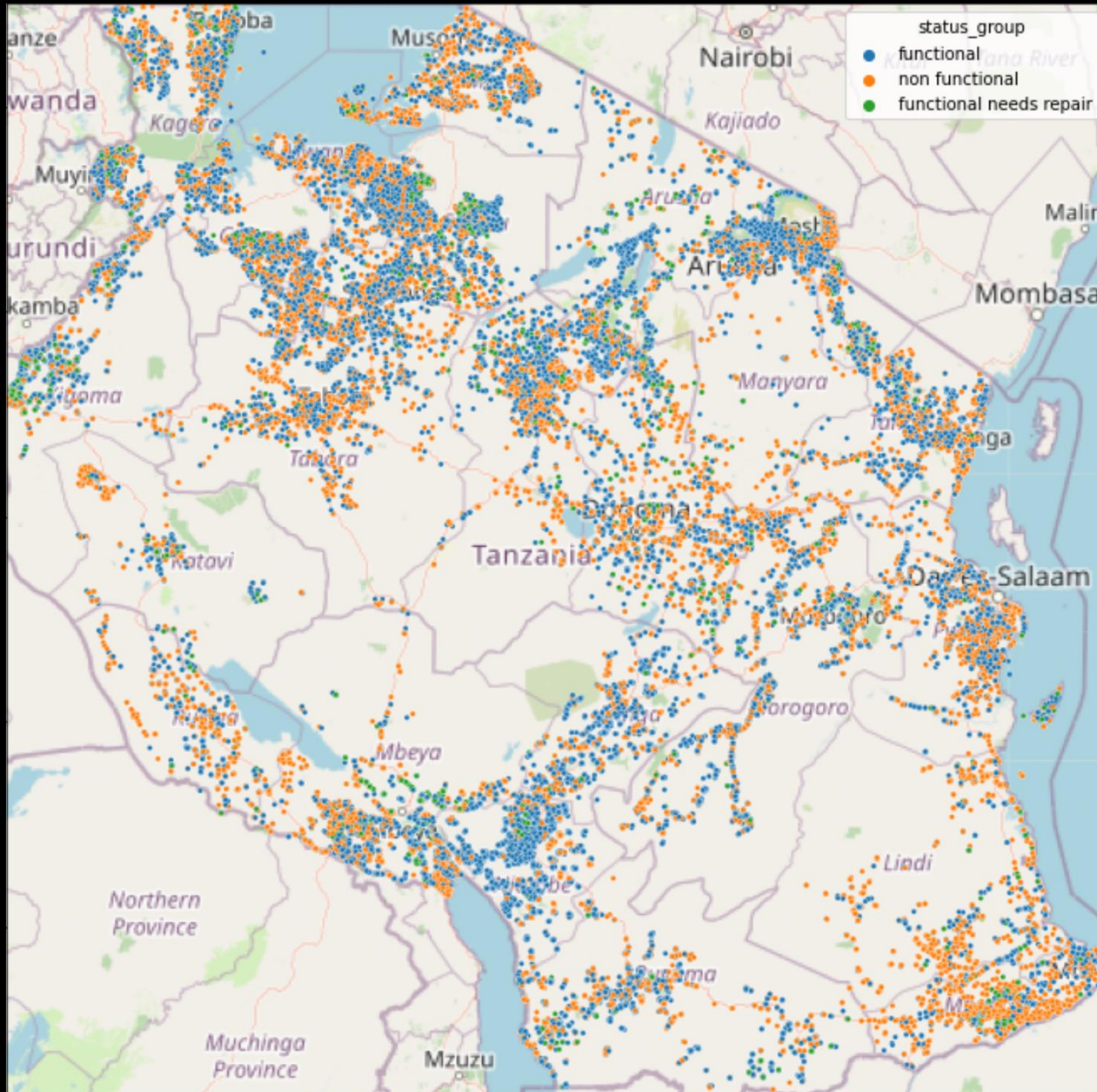




# Pump It Up: Data Mining the Water Table

- Goal





Functional = 54.3%

Non-functional = 38.4%

Needs Repair = 7.3%



# The Features

## Pump Design / Details

- Static Head (pump pressure)
- Extraction Type
- Waterpoint Type
- Scheme Name

## Management/Ops

- Management Group
- Local Government Authority

## Water Source

- Water Quality
- Water Quantity
- Source

## Location

- GPS Height
- Longitude
- Latitude
- Ward

## Pump Quality

- Installer
- Age
- Permit
- Payment Type

## Other

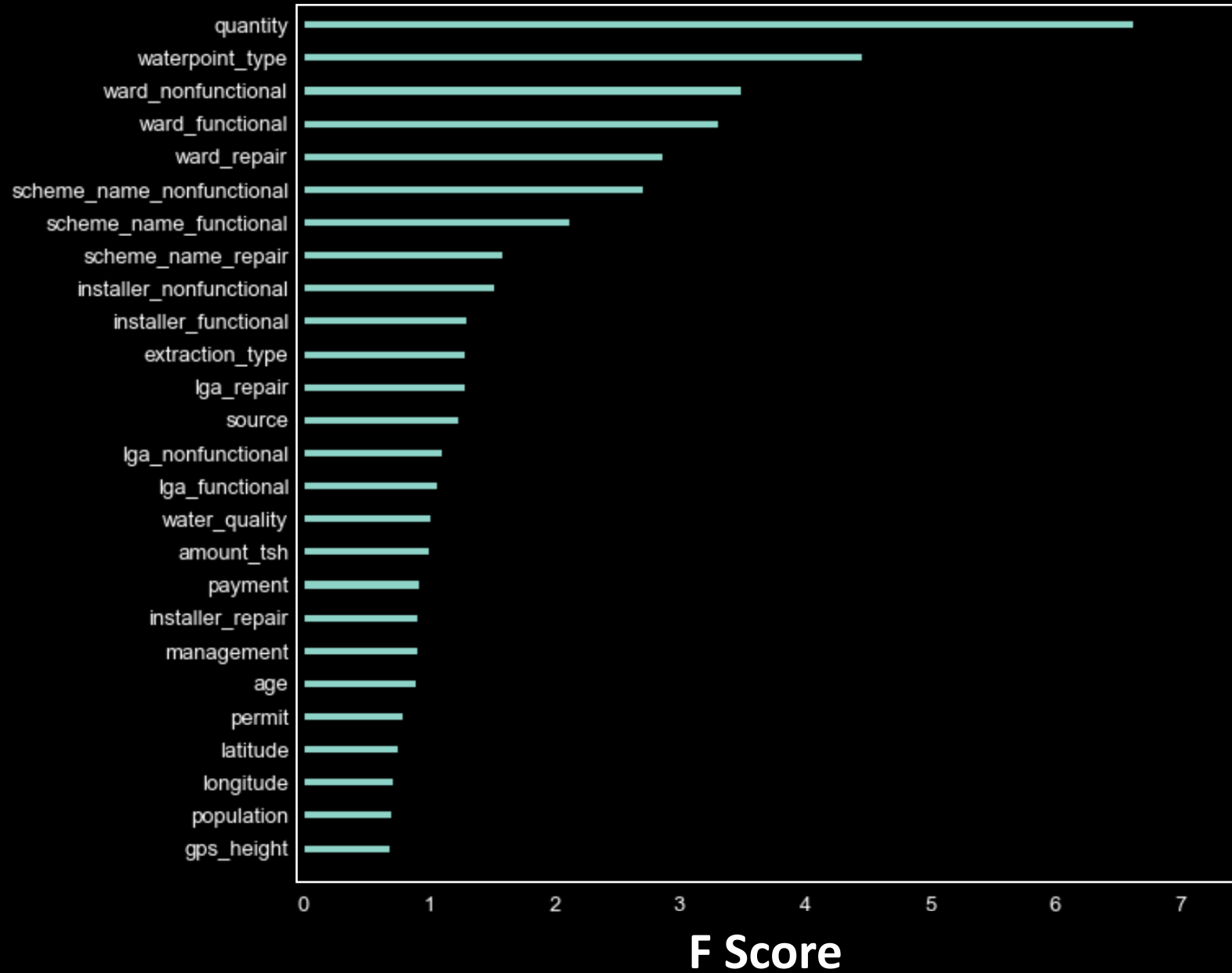
- Surrounding Population





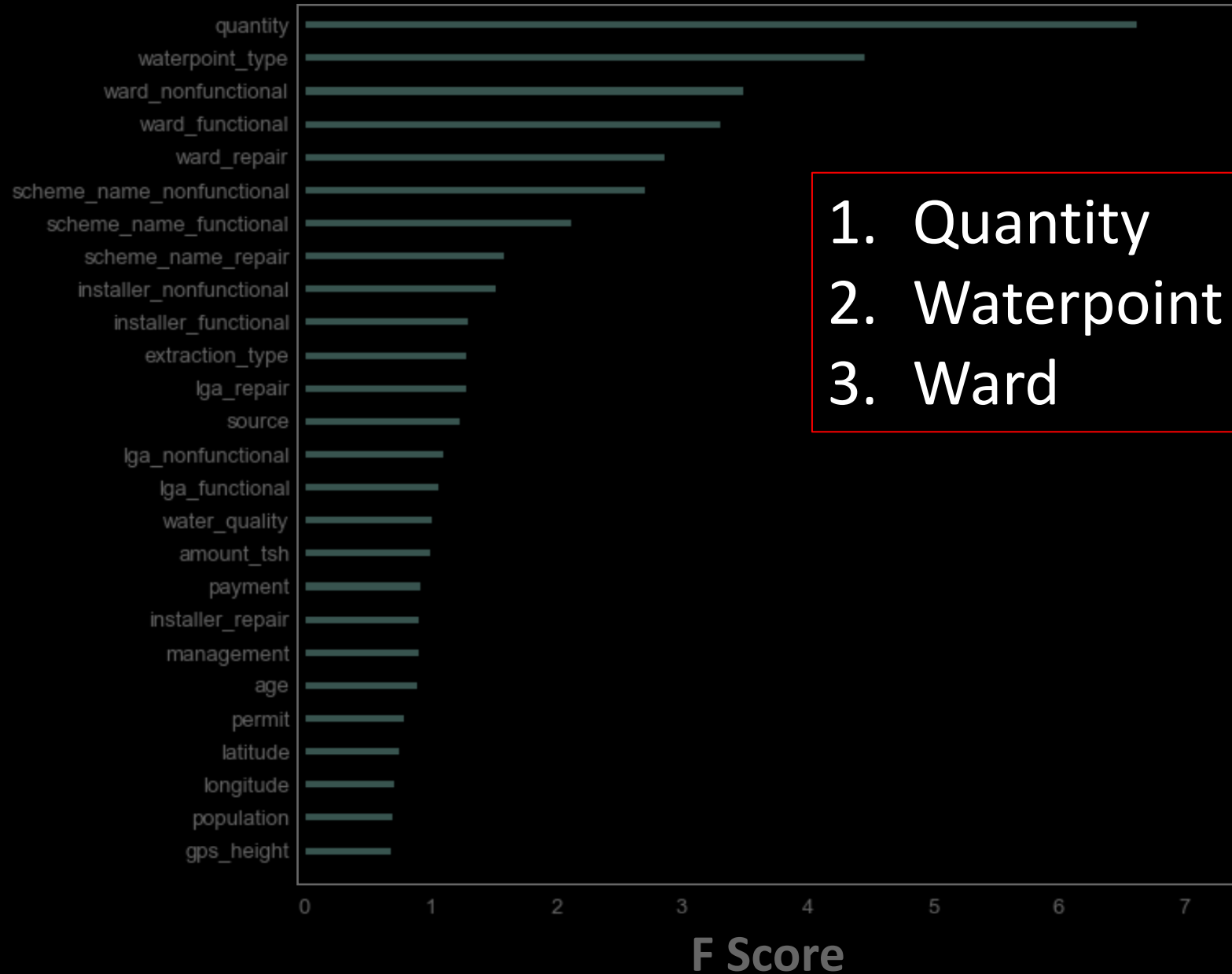
# Model Feature Importance

Features



# Model Feature Importance

Features

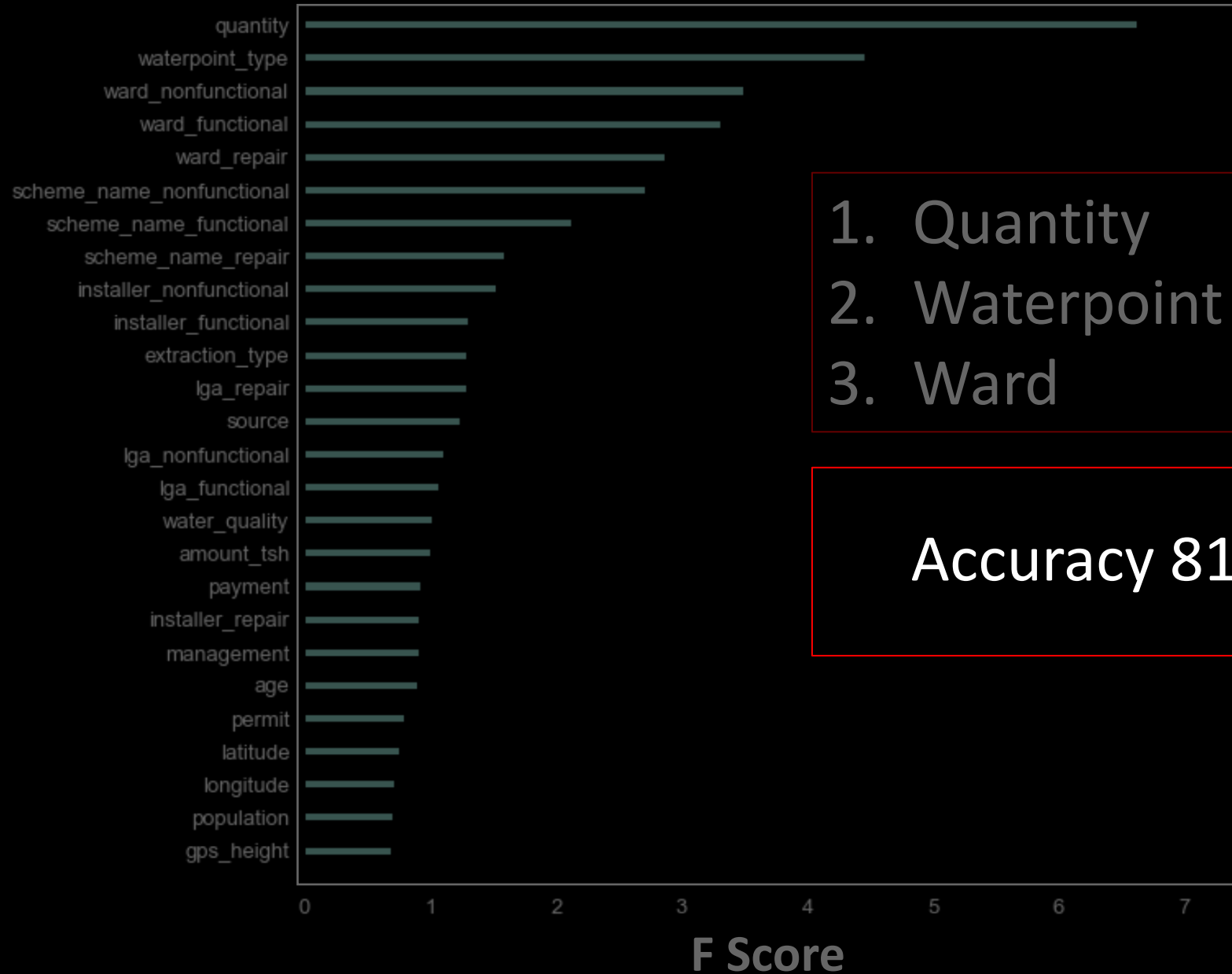


1. Quantity
2. Waterpoint Type
3. Ward



# Model Feature Importance

Features



1. Quantity
2. Waterpoint Type
3. Ward

Accuracy 81.6%



# Model Predictions Vs. True Labels

True: Functional	7000	244	752
True: Needs Repair	533	396	168
True: Non-Functional	1153	118	4486
	Predicted: Functional	Predicted: Needs Repair	Predicted:Non-Functional





# Model Predictions Vs. True Labels

True: Functional	7000	244	752
True: Needs Repair	533	396	168
True: Non-Functional	1153	118	4486
	Predicted: Functional	Predicted: Needs Repair	Predicted:Non-Functional

Existing Functional Wells: 7992 ~ 53.6%



# Model Predictions Vs. True Labels

True: Functional	7000	244	752
True: Needs Repair	533	396	168
True: Non-Functional	1153	118	4486
	Predicted: Functional	Predicted: Needs Repair	Predicted: Non-Functional

Existing Functional Wells: 7992 ~ 53.6%  
Wells to be fixed: 396



# Model Predictions Vs. True Labels

True: Functional	7000	244	752
True: Needs Repair	533	396	168
True: Non-Functional	1153	118	4486
	Predicted: Functional	Predicted: Needs Repair	Predicted:Non-Functional

Existing Functional Wells: 7992 ~ 53.6%  
Wells to be fixed: 396  
Wells to replace: 4486



# Model Predictions Vs. True Labels

True: Functional	7000	244	752
True: Needs Repair	533	396	168
True: Non-Functional	1153	118	4486
	Predicted: Functional	Predicted: Needs Repair	Predicted:Non-Functional

Existing Functional Wells: 7992 ~ 53.6%  
Wells to be fixed: 396  
Wells to replace: 4486





# Model Predictions Vs. True Labels

True: Functional	7000	244	752
True: Needs Repair	533	396	168
True: Non-Functional	1153	118	4486
	Predicted: Functional	Predicted: Needs Repair	Predicted:Non-Functional

Existing Functional Wells: 7992 ~ 53.6%  
Wells that can be fixed: 396  
Wells to replace: 4486  
168



# Model Predictions Vs. True Labels

True: Functional	7000	244	752
True: Needs Repair	533	396	168
True: Non-Functional	1153	118	4486
	Predicted: Functional	Predicted: Needs Repair	Predicted:Non-Functional

Existing Functional Wells: 7992 ~ 53.6%  
Wells that can be fixed: 396  
Wells to replace: 4486  
168

New Functional Wells: 12878 ~ 86.7%



# Summary

- Model Accuracy 81.6%
- Increase total well functionality to 86% (net 33% increase)

## Future Work

- Minimize false prediction of functional wells.



Any Questions?

