Tanzanian Water-Well Predictive Modeling

Jessica Guzzo
Student
Email: jguzzo522@gmail.com

Tanzanian Water Crisis

- ❖ 1/3 population dry

- ❖ Water Borne Illnesses

- ❖ Gender Based Violence

Source: The Water Project, "Water in Crisis - Tanzania" (https://thewaterproject.org/water-crisis/water-in-crisis-tanzania)

Goal: Predict Non-Functional Water-Wells in Tanzania

Methodology

❖ Functionality of wells
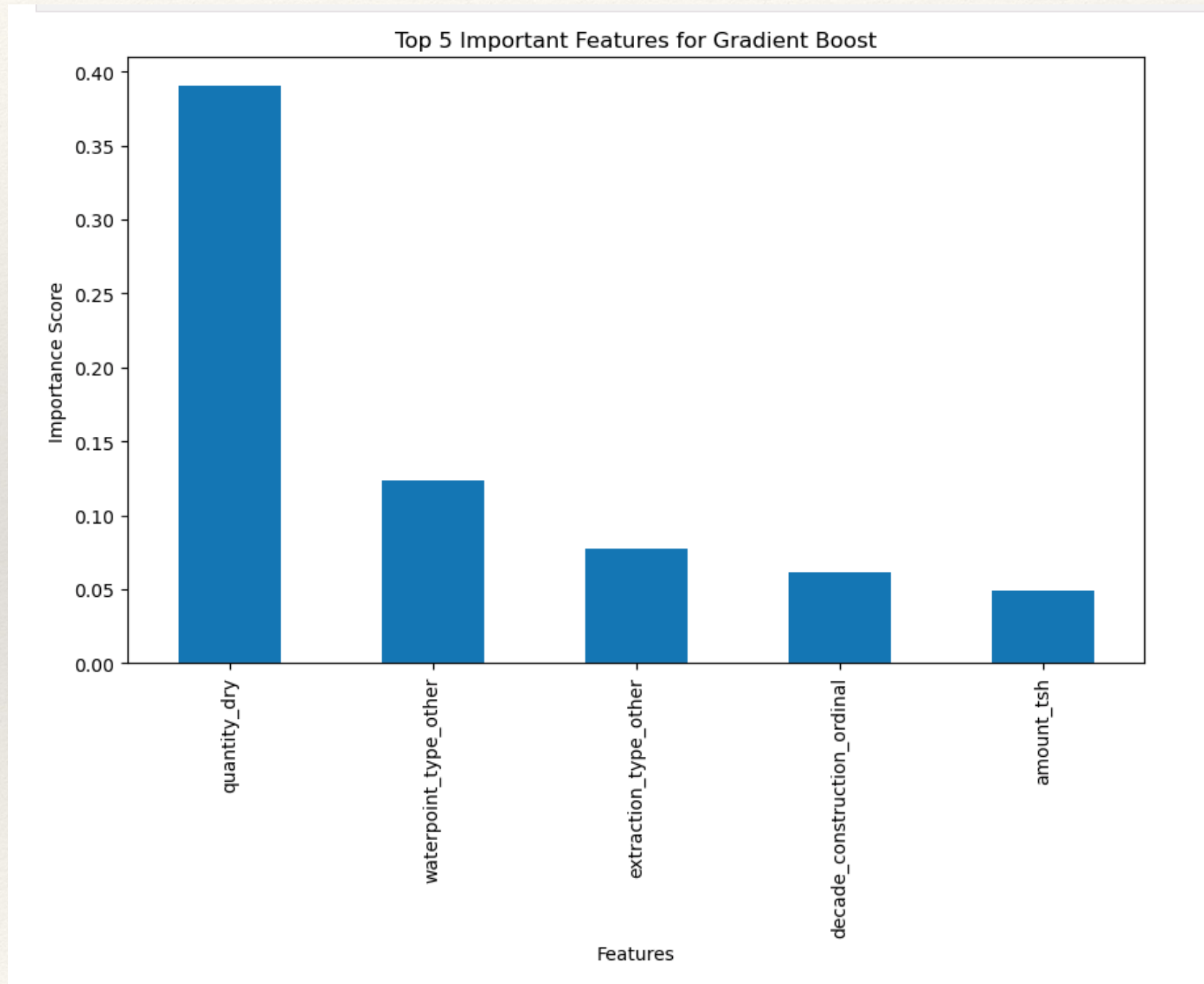
❖ Transformed data

❖ Gradient Boosting Classifier

Water Wells in Tanzania (Functional and Non-Functional)
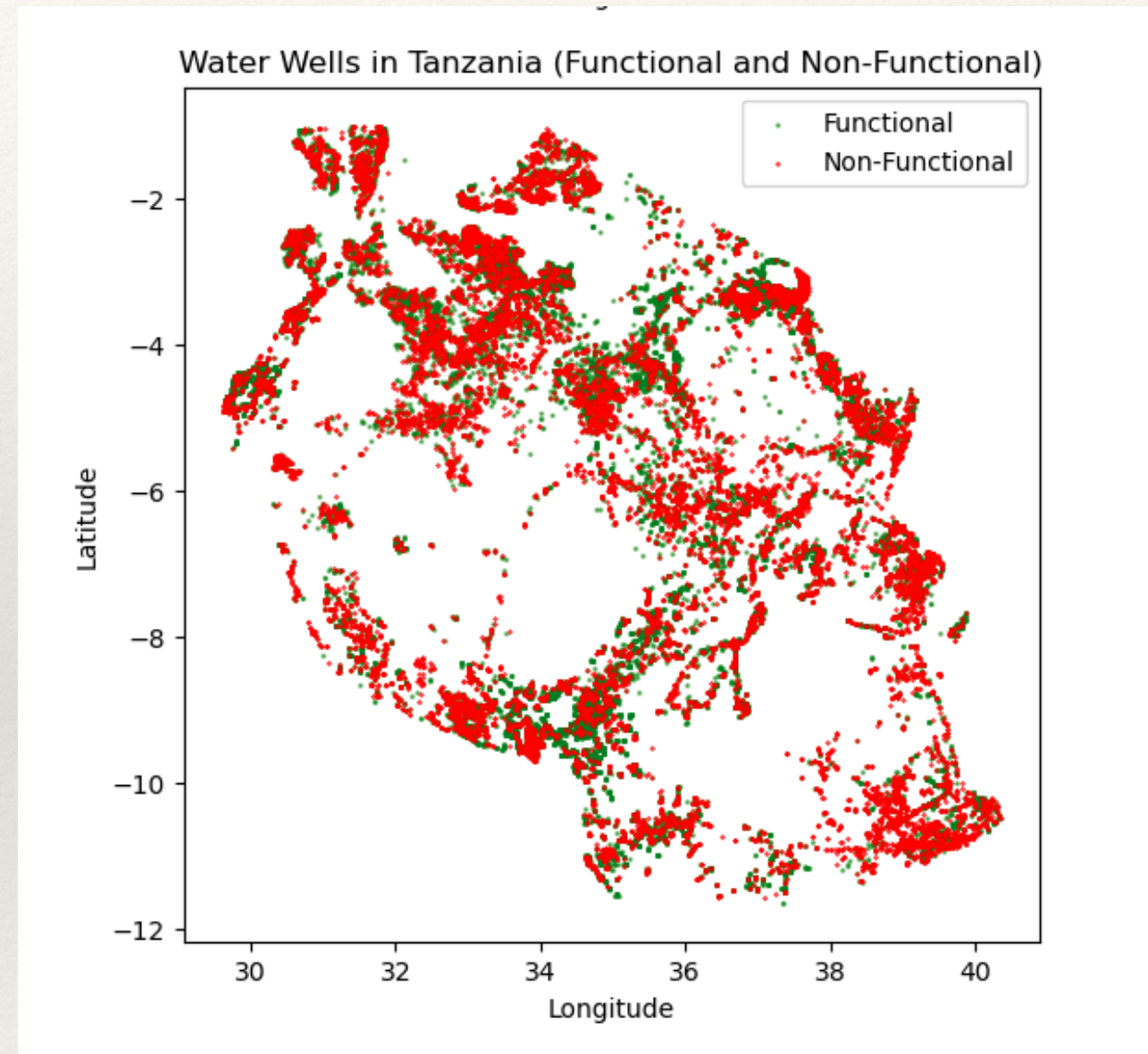
Model Results

- Accuracy 80%
- Low % of false negatives
- ROC AUC .756

Confusion Matrix for GBM Model
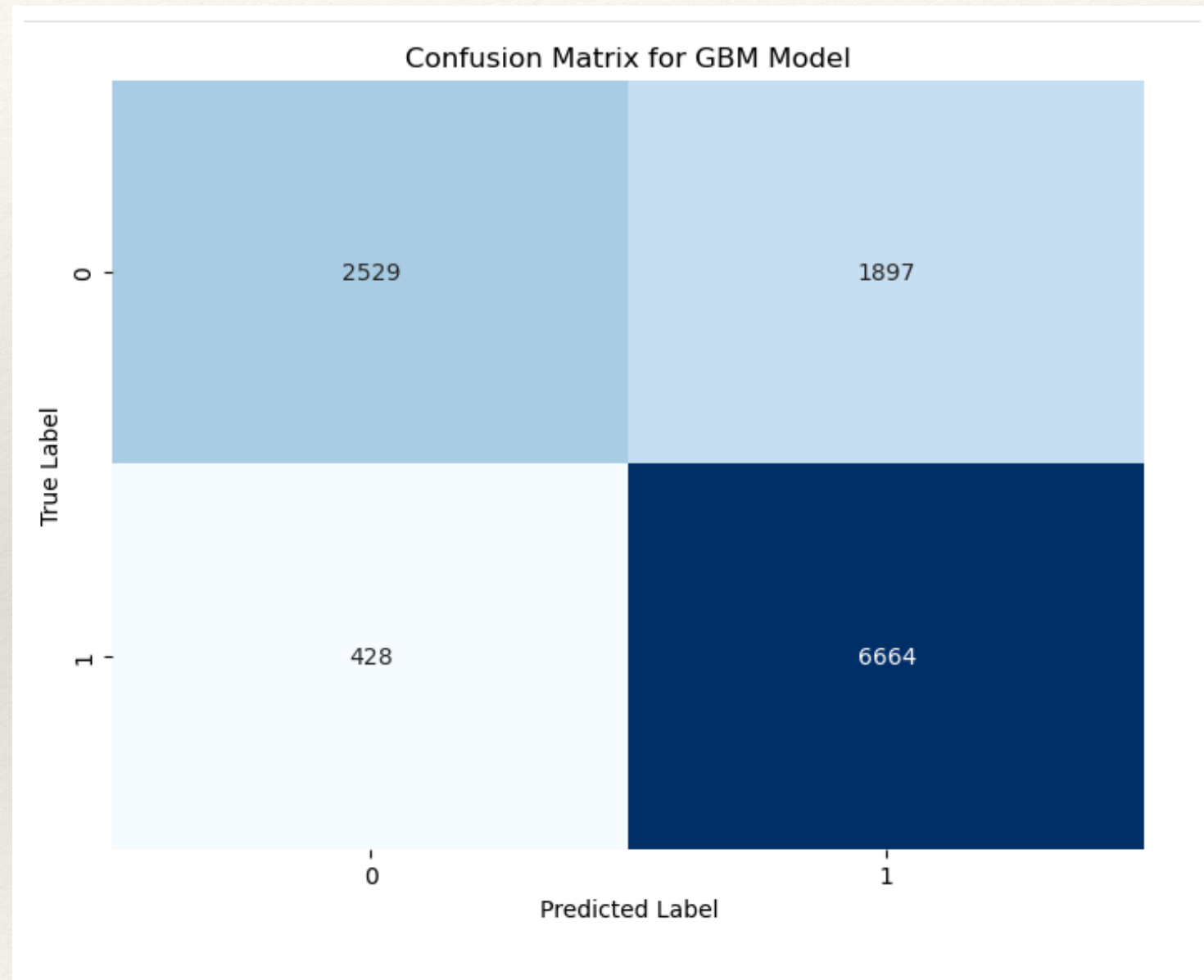
Key Features

- ❖ Water Quantity

- ❖ Water Point Type

- ❖ Decade of Construction

Top 5 Important Features for Gradient Boost

Water Quantity Dry

- Dry 5% Functionality

- Enough 75% functionality

- Unknown 20% Function

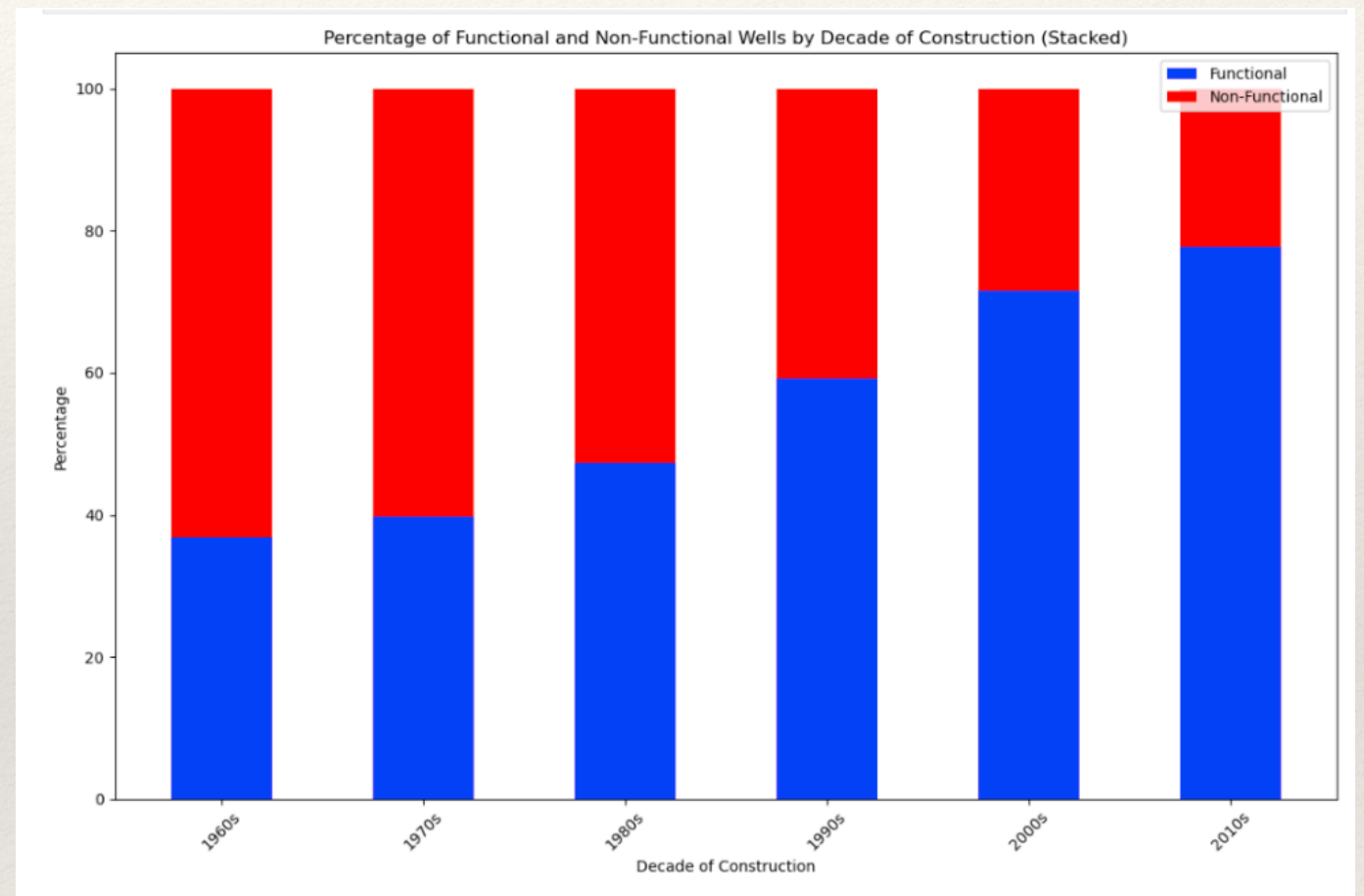Functional and Non-Functional Wells by Water Quantity Categories
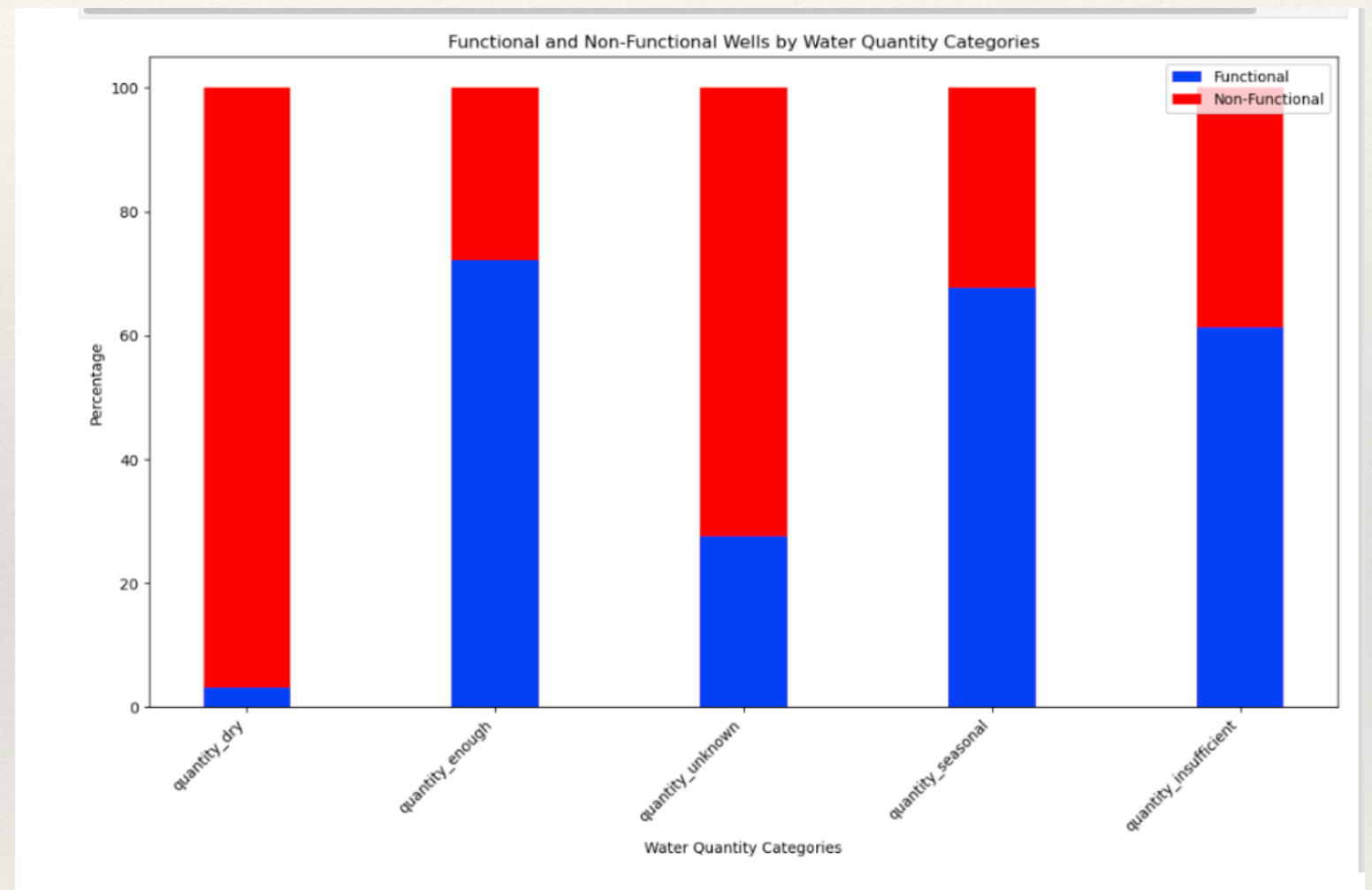
Water Point Type

- ❖ Other 20% function

- ❖ Dam most functional
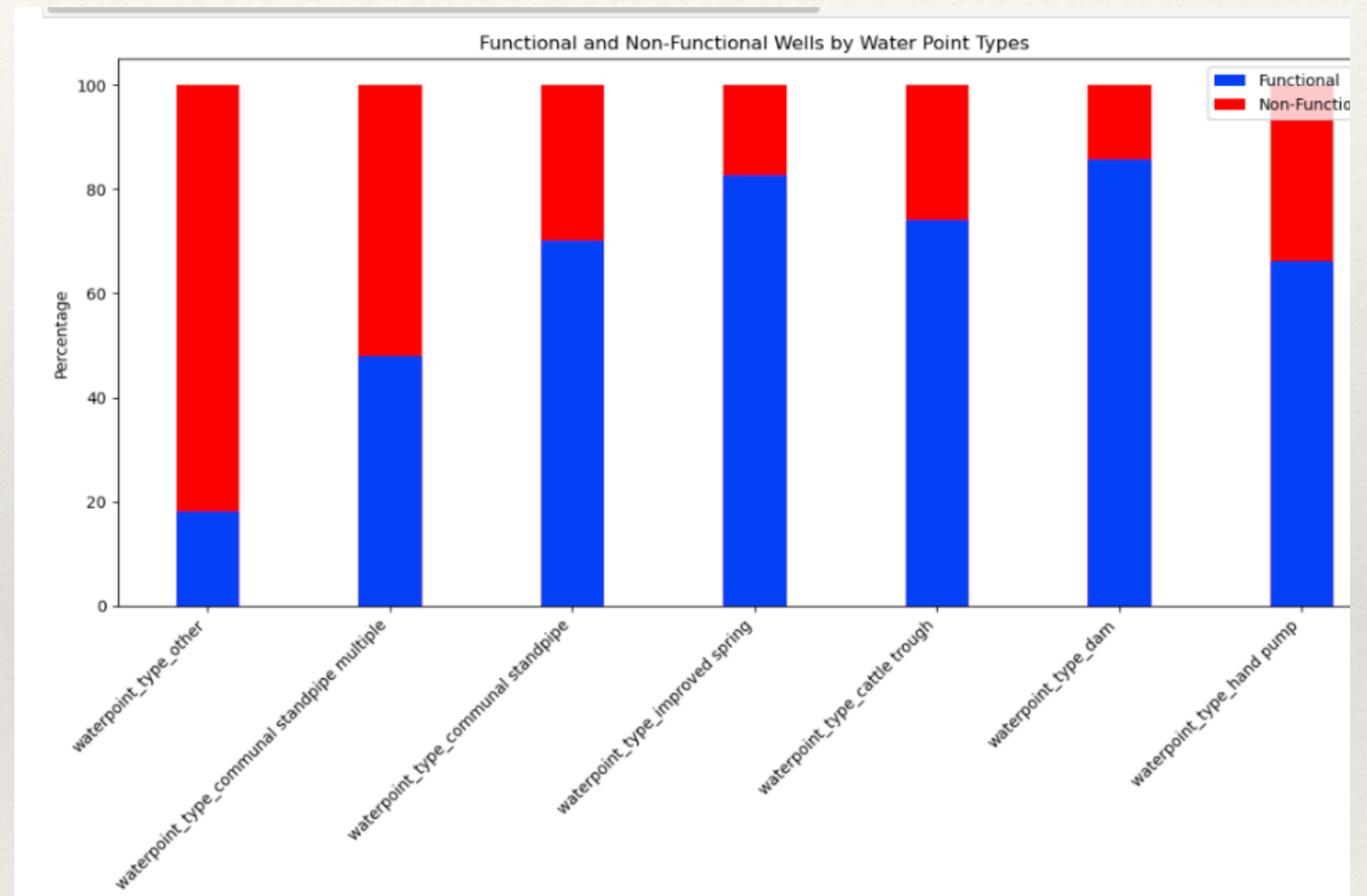
Decade of Construction

- 1960's least functional

- 2010's most functional

- Newer than 90's functional

Percentage of Functional and Non-Functional Wells by Decade of Construction (Stacked)

Recommendations

❖ Repair older wells

❖ Update water-point types

❖ Closely monitor dry wells

Limitations

- Examine funders, management companies
- Future projects geography
- Remote area alternatives
- City vs Remote Villages

Thank you

❖ Any questions?

Contact Info

Jessica Guzzo

Student

JGuzzo522@Gmail.com

Link to notebook [Tanwaterwell Notebook](https://github.com/jguzzo522/twaterwell/blob/main/tanwaterwell.ipynb).

Link to [README](https://github.com/jguzzo522/twaterwell/blob/main/README.md).