*Gradient Boosting Model*

# 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

# Gradient Boosting Modeling Predictions

❖ Functionality of wells

❖ Transformed data

❖ KNN imputation
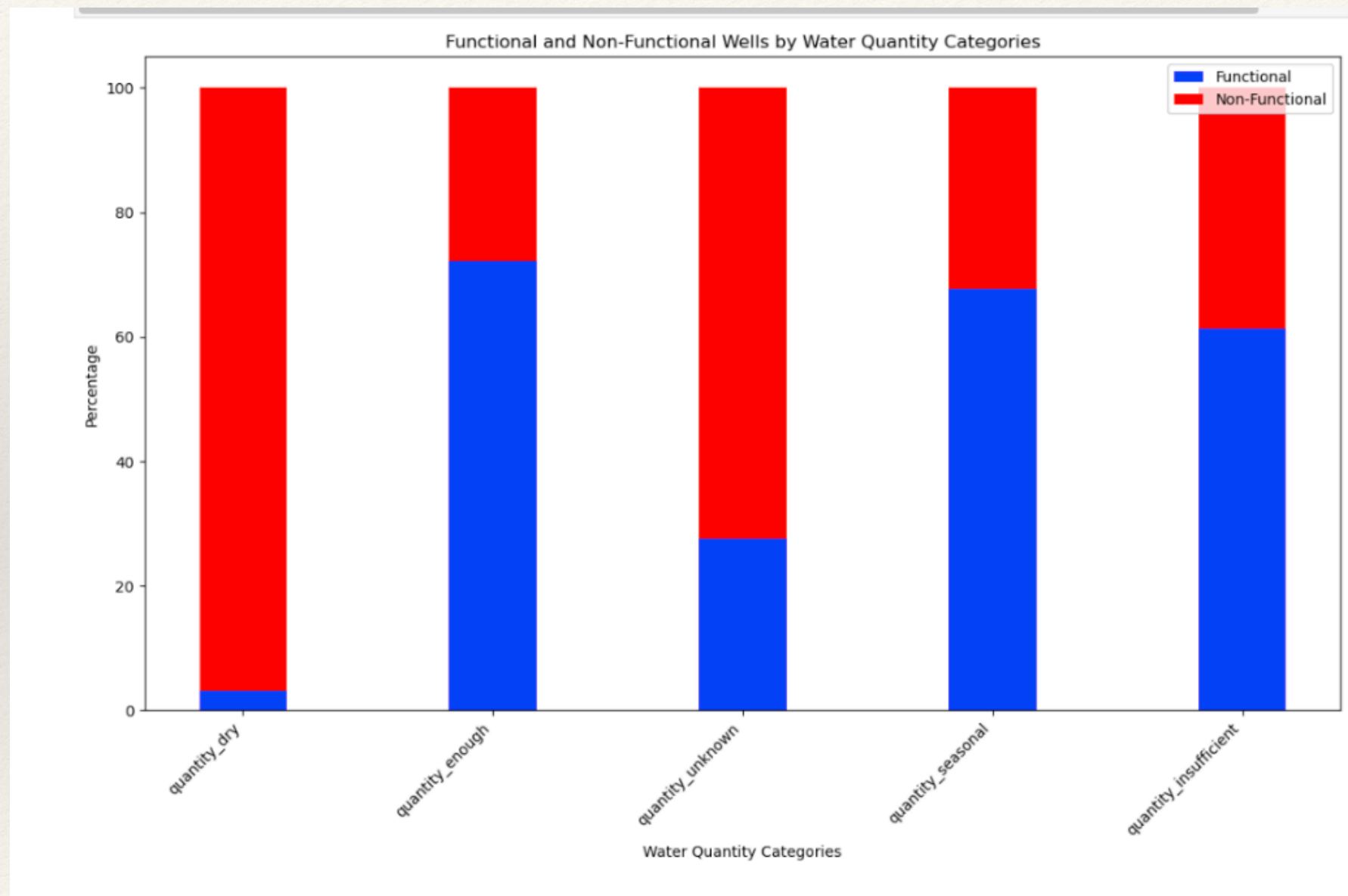
❖ Gradient Boosting Classifier

# Model Results

- ❖ Accuracy 80%

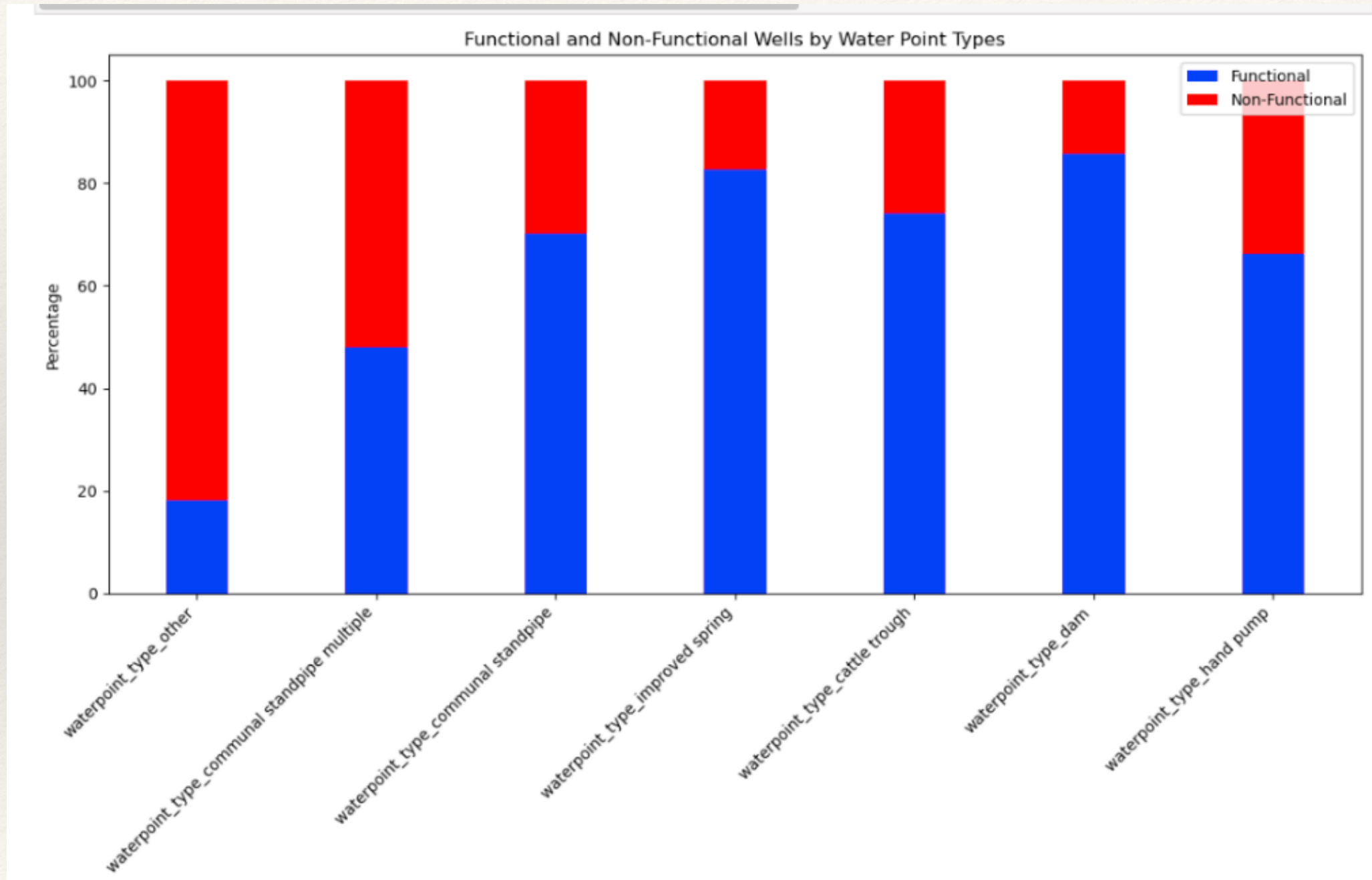- ❖ Low % of false negatives

- ❖ ROC AUC .756

# Key Features



❖ Water Quantity

❖ Water Point Type

❖ Decade of Construction
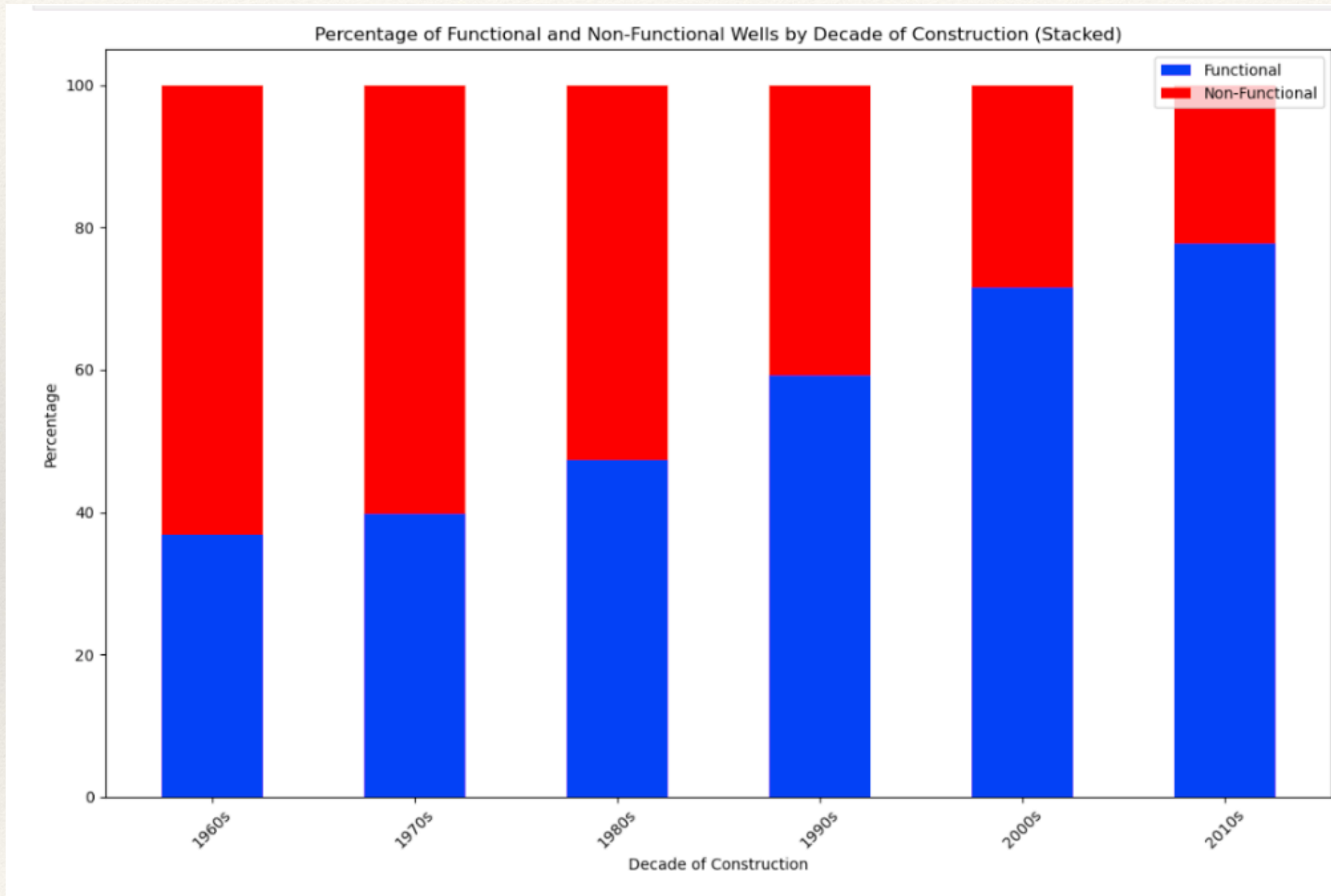
# Water Quantity Dry



Functional and Non-Functional Wells by Water Quantity Categories

# Water Point Type



Functional and Non-Functional Wells by Water Point Types

# Decade of Construction



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).