Predicting the condition of water pumps in Tanzania

EM JAGER

Introduction

DATA SCIENCE

SUSTAINABLE DEVELOPMENT

Overview

- Reseach Context
- Data & Methods
- Recommendations
- Next Steps

Research Context

63.6 million Population of Tanzania (2021)

61% of households have access to a basic water supply

Water is a human right

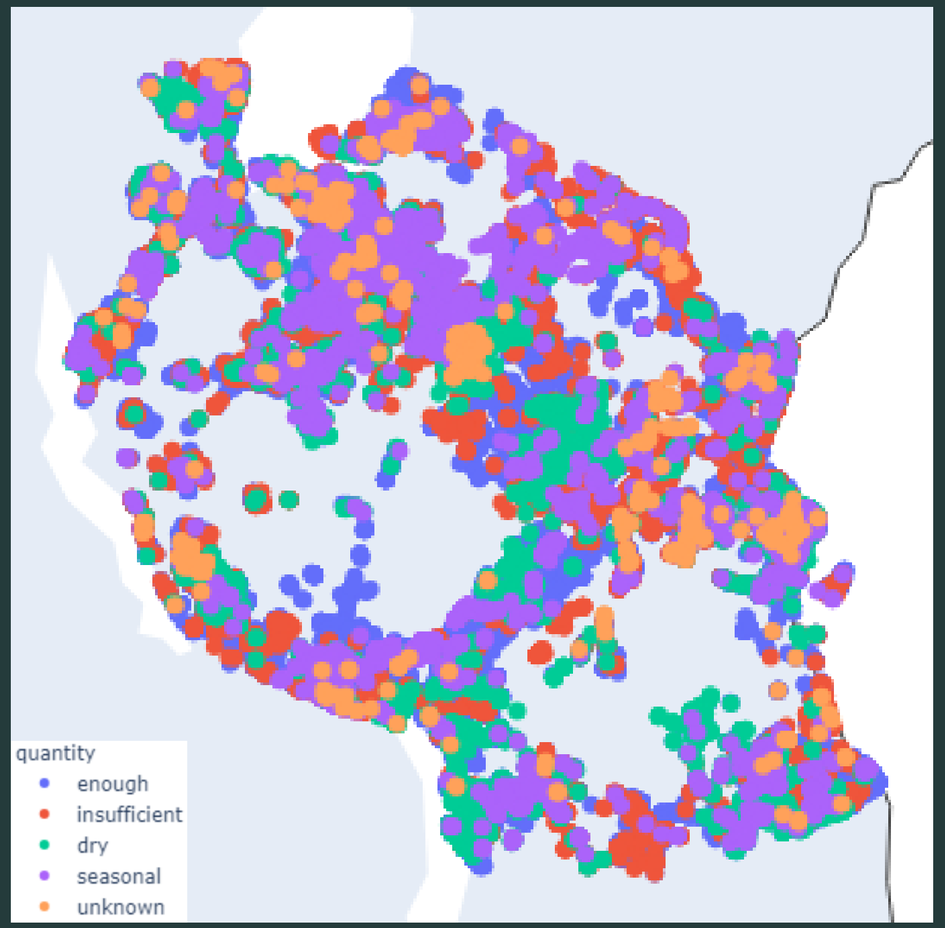
Project Goal:
To provide the Tanzanian government and local NGOs with a method to prioritize water pump sites that need repair

Lack of access to water impacts:

- Public health
- Gender equality
- Income inequality
- Economic growth and development
- Overall quality of life

The Data

- Source: TAARIFA
- 60,000 water wells with categorical data on their type, location, and functioning status
- From 1960-2013

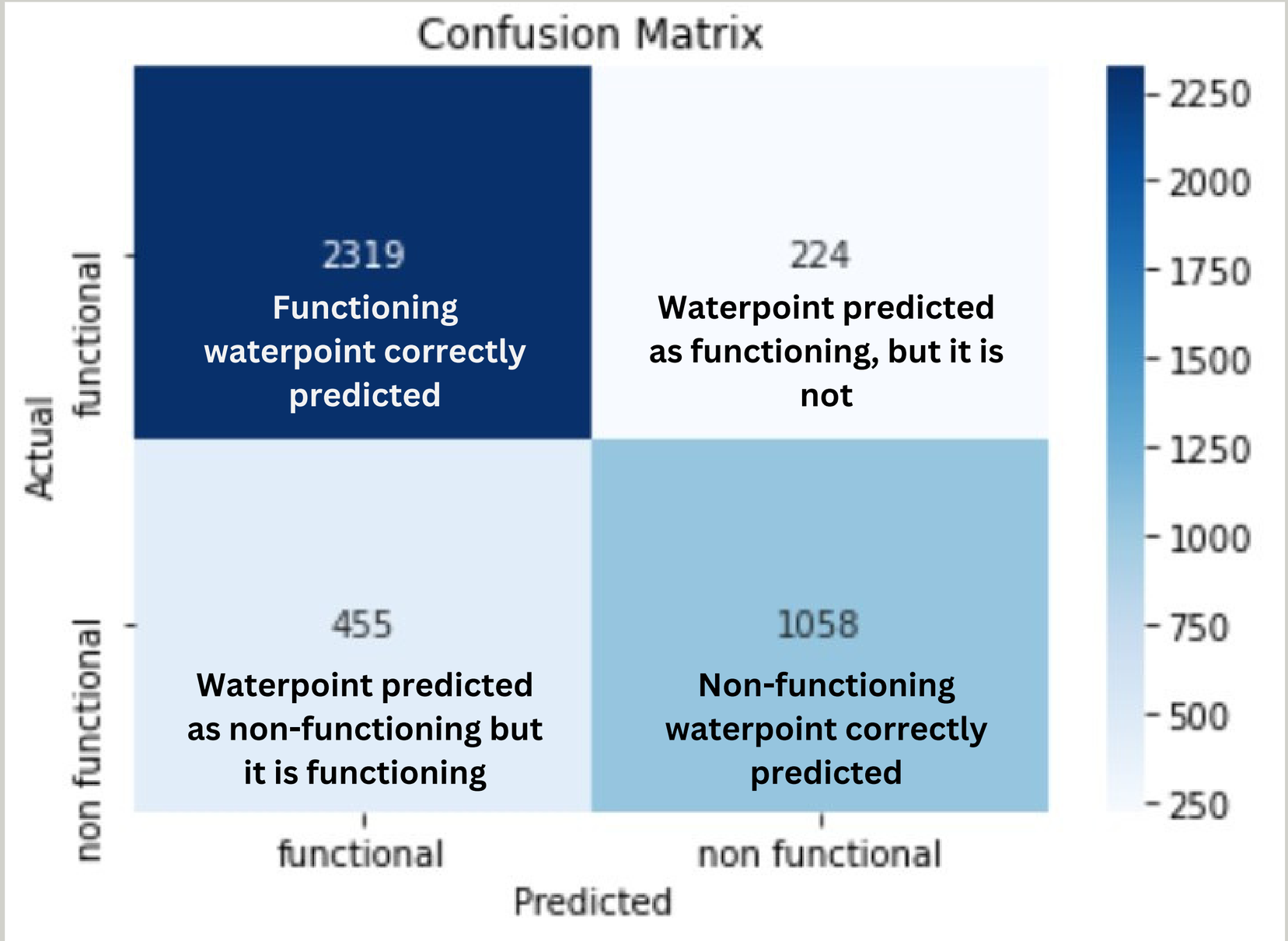


The Data

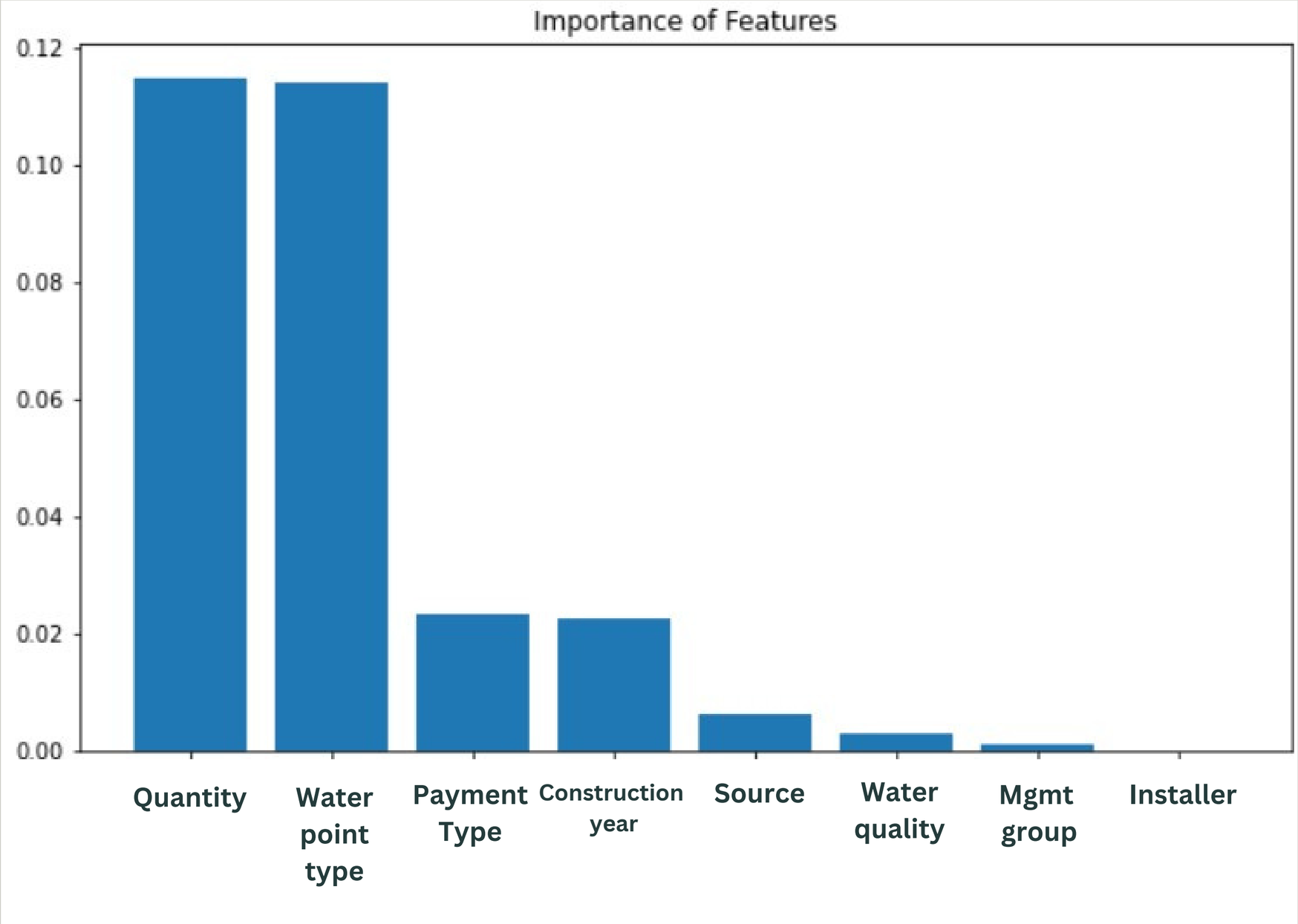
- Variables investigated:
  - Output quantity
  - Waterpoint type
  - Payment type
  - Construction year
  - Water source
  - Water quality
  - Waterpoint management group
  - Installer

- **Functioning or non-functioning?**

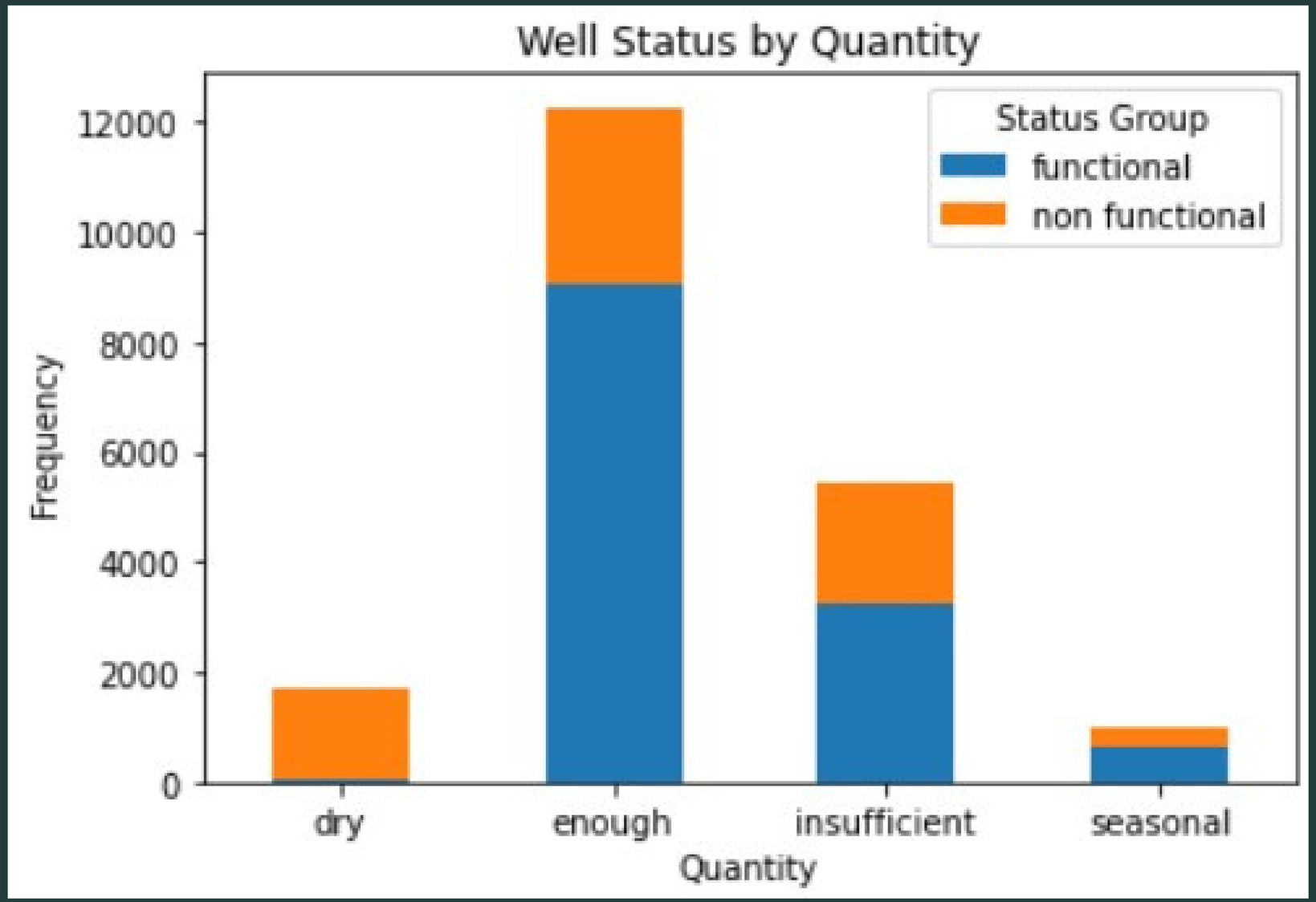
Decision Tree Model Predictions



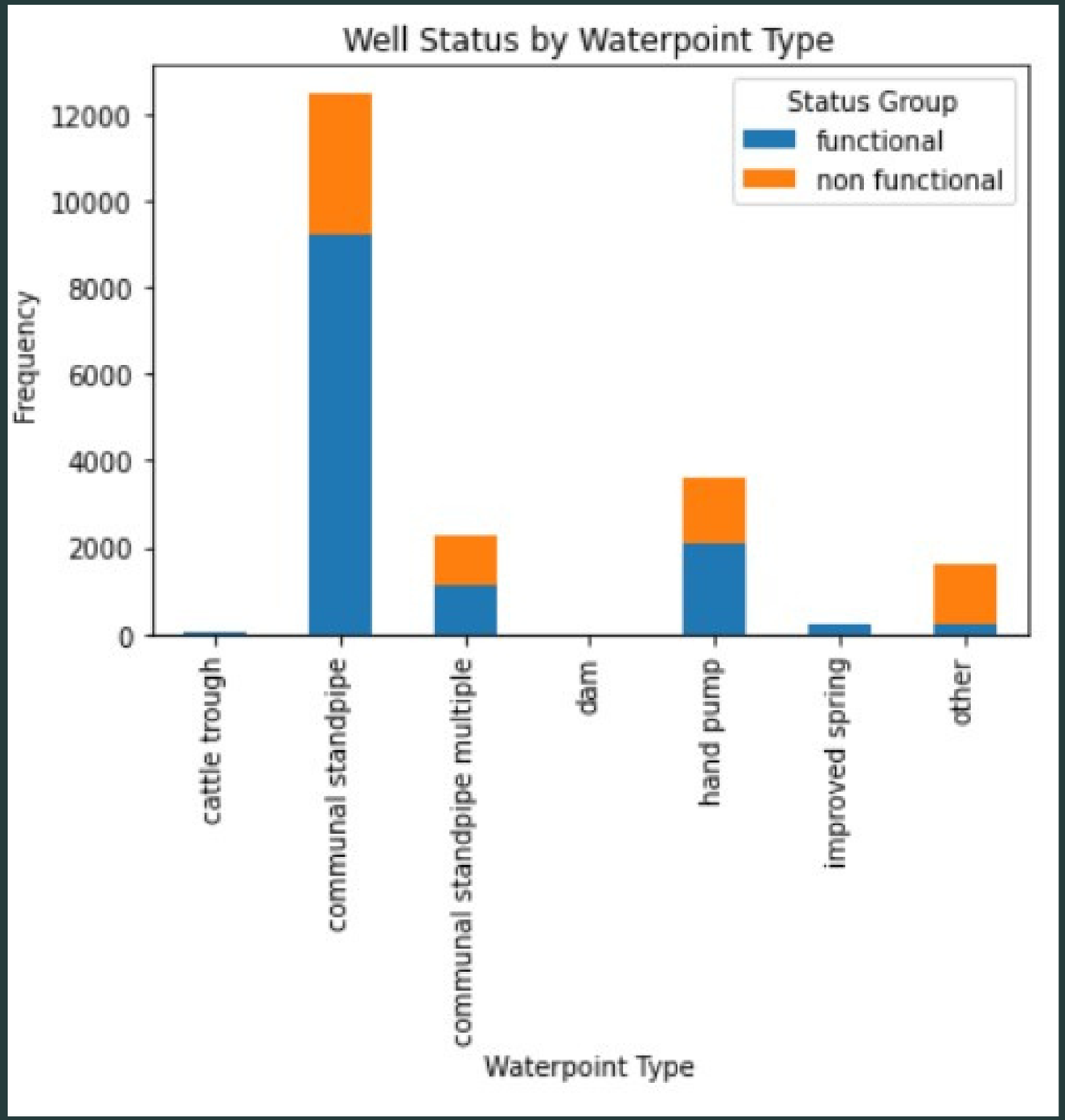
Top Predictors of Waterpoint Functioning Status



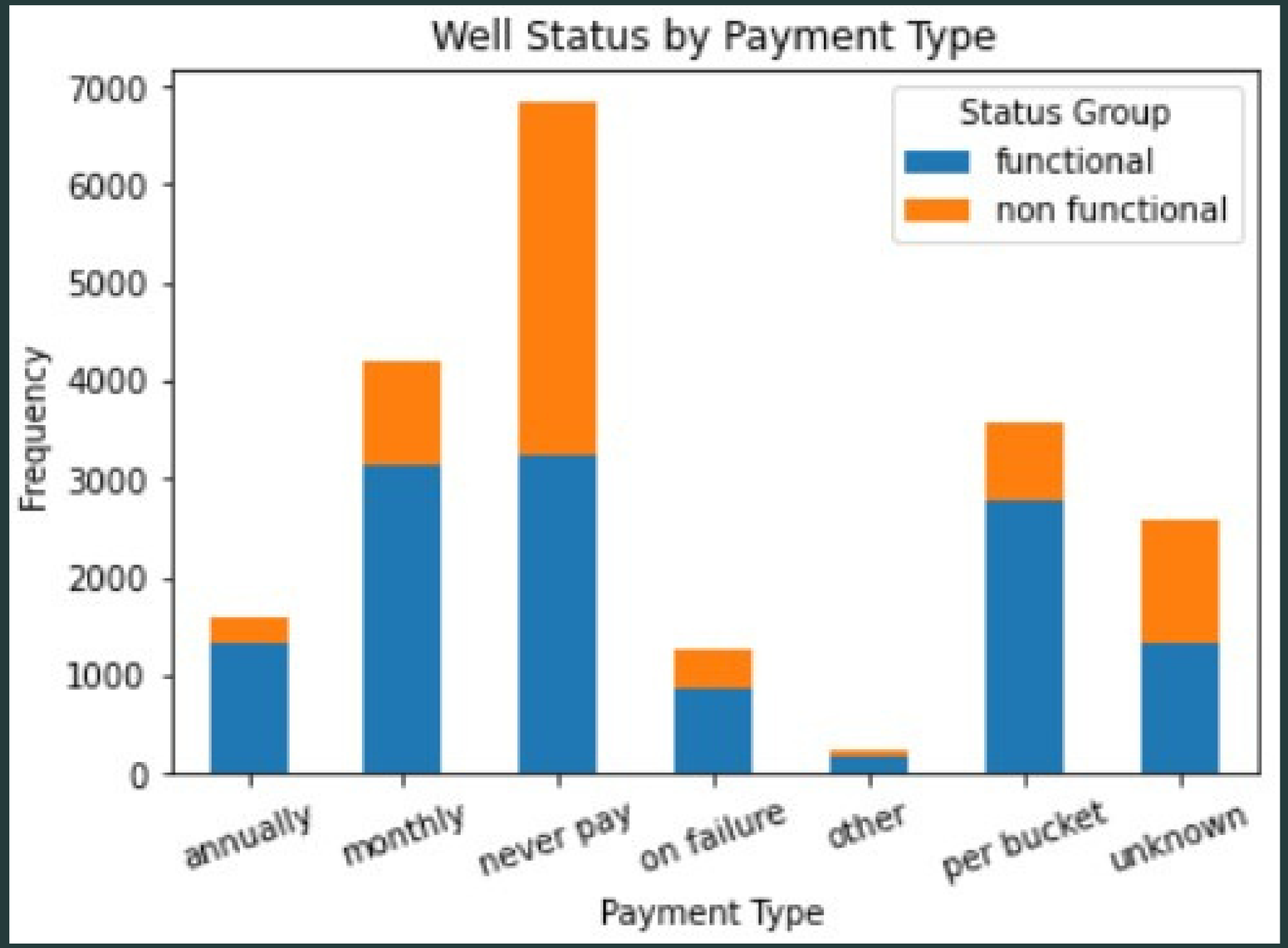
Quantity


Well Status by Quantity
Status Group
functional
non functional
Frequency
12000
10000
8000
6000
4000
2000
0
dry
enough
insufficient
seasonal
Quantity

Waterpoint type



Payment type



Next Steps

1. Collect more recent data
2. Use model to predict and prioritize waterpoint repairs
3. Use data to find what types of water pumps perform best and are functioning over time

Thank you!

Any questions?

Contact me:

For more information on the model and recommendations, check out my Github repository