

DSF-PT07: Tanzania Wells

Phase 3 Project
Nduku Kiteng'e





OUTLINE

1. Problem Statement
2. Data Process
3. Model Review
4. Limitations
5. Recommendations



Problem Statement

The goal of this project is to :

1. Predict which pumps are functional, which need some repairs and which dont work at all.
2. Predict one of these three classes based on a number of variables about what kind of pump is operating, when it was installed, and how it is managed.
3. Provide a smart understanding of which waterpoints will fail can improve maintenance operations and ensure that clean, potable water is available to communities across Tanzania.

The challenge from DrivenData. (2015). Pump it Up: Data Mining the Water Table. Retrieved [Month Day Year] from

<https://www.drivendata.org/competitions/7/pump-it-up-data-mining-the-water-table> with data from Taarifa and the Tanzanian Ministry of Water.



The Data Process

Data Understanding

Session to load and review the data to understand the information it can provide.

Data Preparation

Includes data cleaning for missing values, data types and null values and feature engineering for potential analysis methods and variables.



Recommendations

Develop a list of recommendations for the clients especially potential buyers and home owners.

Modelling and interpretation

Using the cleaned and prepared data, we review the potential models based on the identified objectives.



Modelling

I was not able to get to this section, as my EDA took way longer than planned.

**Thank you.
Questions?**

