

# Capstone 2 - Project Proposal

## Background

The UN publishes and reviews a list of *least developed countries* (LDC) every 3 years. An LDCs are described as *"low-income countries confronting severe structural impediments to sustainable development. They are highly vulnerable to economic and environmental shocks and have low levels of human assets."* Tanzania has been classified as an LDC since the UN published the first list of LDCs in 1971<sup>2</sup>. A common challenge of LDCs is a lack of infrastructure to support the development of the nation, including access to education and healthcare, waste management, and access to potable water.

According to UNICEF, as of 2017, more than 24 million Tanzanians lacked access to basic drinking water<sup>3</sup>. This corresponds to only 56.7% of the country's population having access to basic drinking water. Outside of developed urban areas, much of the potable water is accessed via water pumps.

[Taarifa](#) is an open-source platform for crowd-sourced reporting and triaging of infrastructure related issues. Together with the Tanzanian Ministry of Water, data has been collected for thousands of water pumps throughout Tanzania. The goal of this project is to be able to predict the condition of these water pumps to improve maintenance, reduce pump downtime, and ensure basic water access for millions of Tanzanians.

## Problem Statement

Predict the operating condition of waterpoints in Tanzania given various metadata on each water point.

## Data

The dataset is provided by Taarifa, together with the Tanzanian Ministry of Water and is hosted by DrivenData.org:

<https://www.drivendata.org/competitions/7/pump-it-up-data-mining-the-water-table/page/23/>

## Strategy

The strategy will be to implement a Random Forest model as well as a neural network model for predictions.

## Deliverables

1. Source code for project including data cleaning, model, and reporting.

2. Report summarizing findings.

## **References**

1. <https://www.un.org/development/desa/dpad/least-developed-country-category.html>
2. [https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/ldc\\_list.pdf](https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/ldc_list.pdf)
3. <https://washwatch.org/en/countries/tanzania/summary/statistics/>