TANZANIA WATER PUMP PROJECT

BY:BRIAN KIGEN

BUSINESS PROBLEM:

As an employee of the NGO i have been task to build machine learning models thats predict the functionality of all ground water pumps found throughout the country of Tanzania.If models are accurate, this could help save the Tanzanian government a lot of time and money.Accurate models can help to cut the cost on needing workers drive out to every water pump to inspect them.The government can use this study to know exactly which pumps are working, need repair and which ones aren’t working at all.

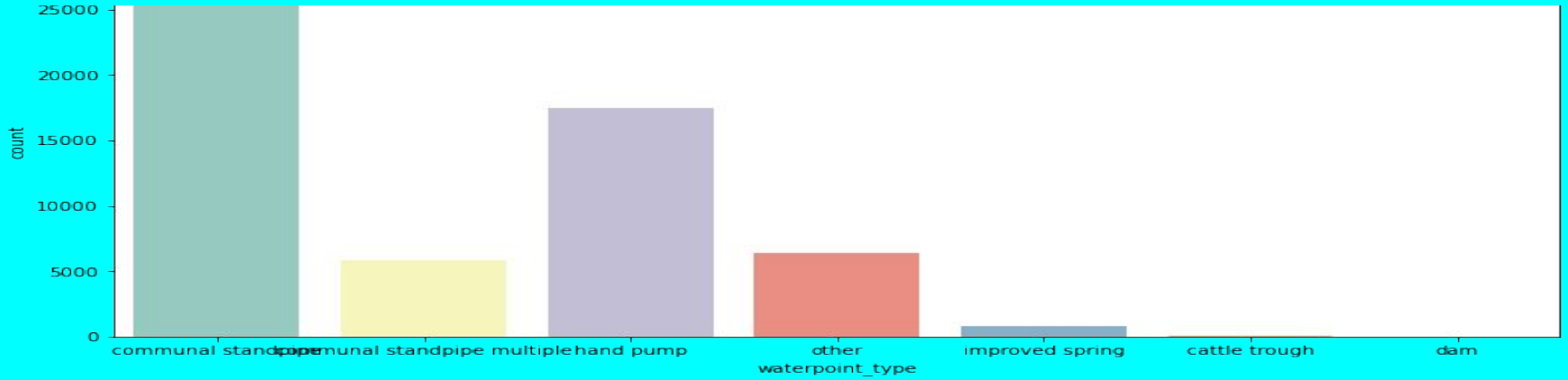
METHODOLOGY

- Data Understanding.
- Data preparation.
- build different predictive classification models for the functionality of the wells.
- Compare results from each Model.

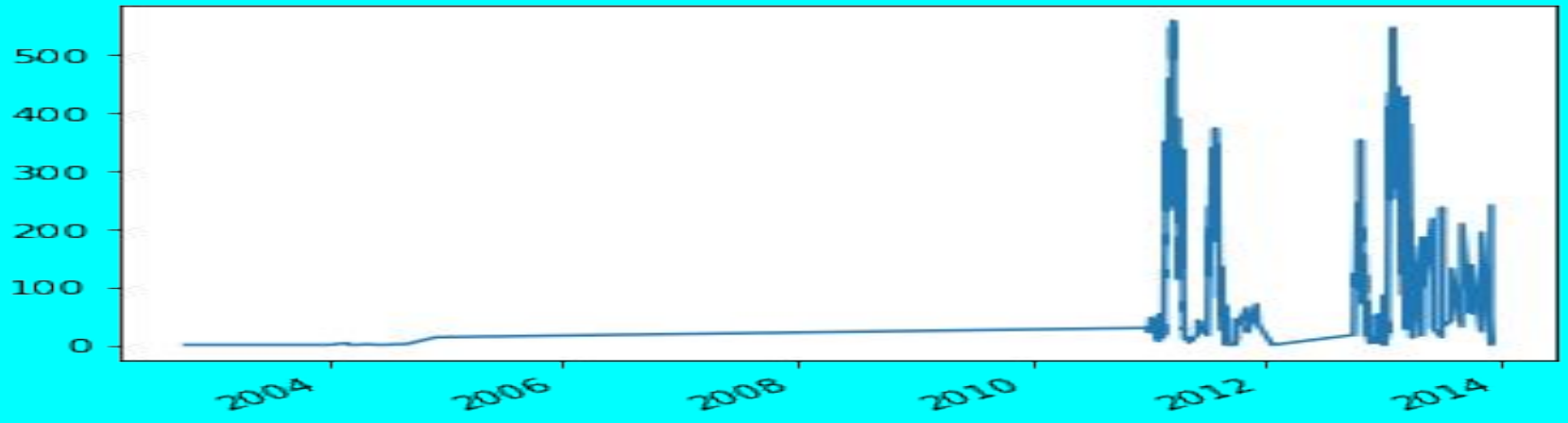
RESULTS:

From the plot below above 30,000 water points are still functional, about 20,00 are non-functional and about 5000 are functional but needs repair.

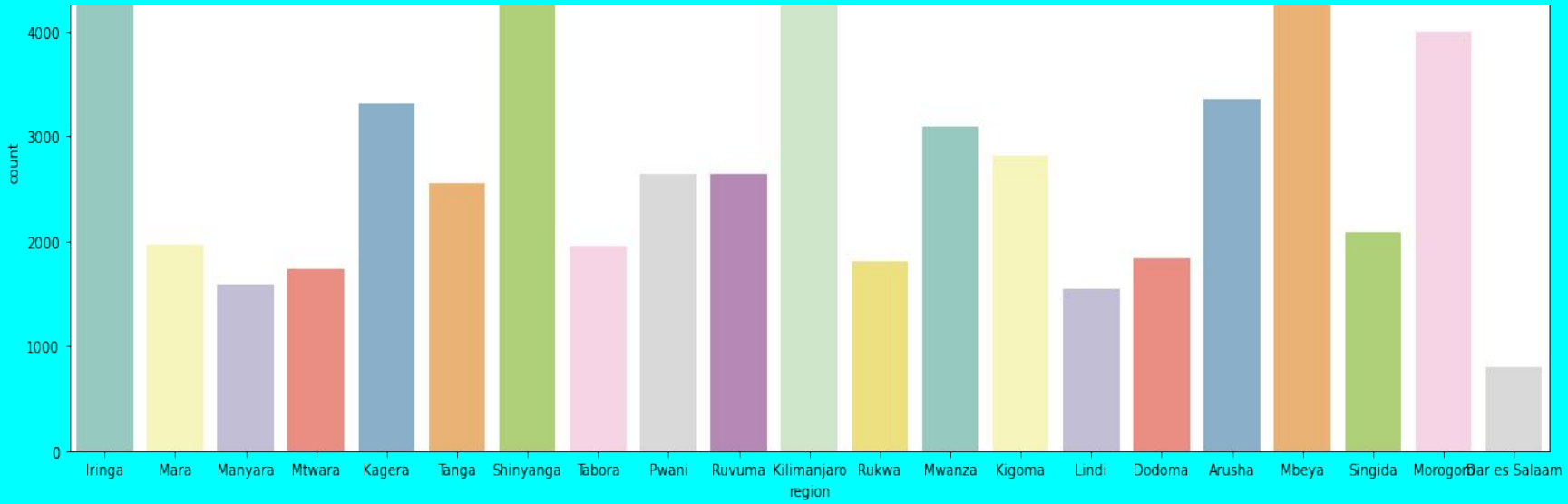
Most of the water pump used in tanzania are mostly community manual stand pump.Followed by the multiple pump and the least is dam



Most of the pumps were build from 2012 to 2014 .Those build from 2004 to 2010 are old most likely needs to be repaired.The recent water pump are most likely to be functional.



Iringa have the have the highest water pump in the country.The capital city of tanzania have the least water pump may because the have access to piped water.



Findings

Random forest was the best model for the analysis with an accuracy of 0.80 and f1 scores of 0.90, 0.28, and 0.77.

	precision	recall	f1-score	support
0	0.79	0.90	0.84	9523
1	0.61	0.28	0.39	1290
2	0.84	0.77	0.80	6896
accuracy			0.80	17709

RECOMMENDATION

- Population: Population was one of the most important features in the random forest classifier. Analysis showed that pumps in lower population areas may be more likely to be non functional or needing repairs. The government should focus resources on low population areas, as they may not be receiving enough.
- Location:Based on the analysis, pumps in lower altitude areas may be more likely to need repair or be non functional. The government should focus resources on lower altitude pumps.

QUESTION!!

THANK YOU