Tanzania Water Wells

“Ajaye kisimani mbele hunywa maji maenge.”
“The first person coming at the well drinks clean water.”
~ Swahili Saying

Obatsa Davis

Project purpose

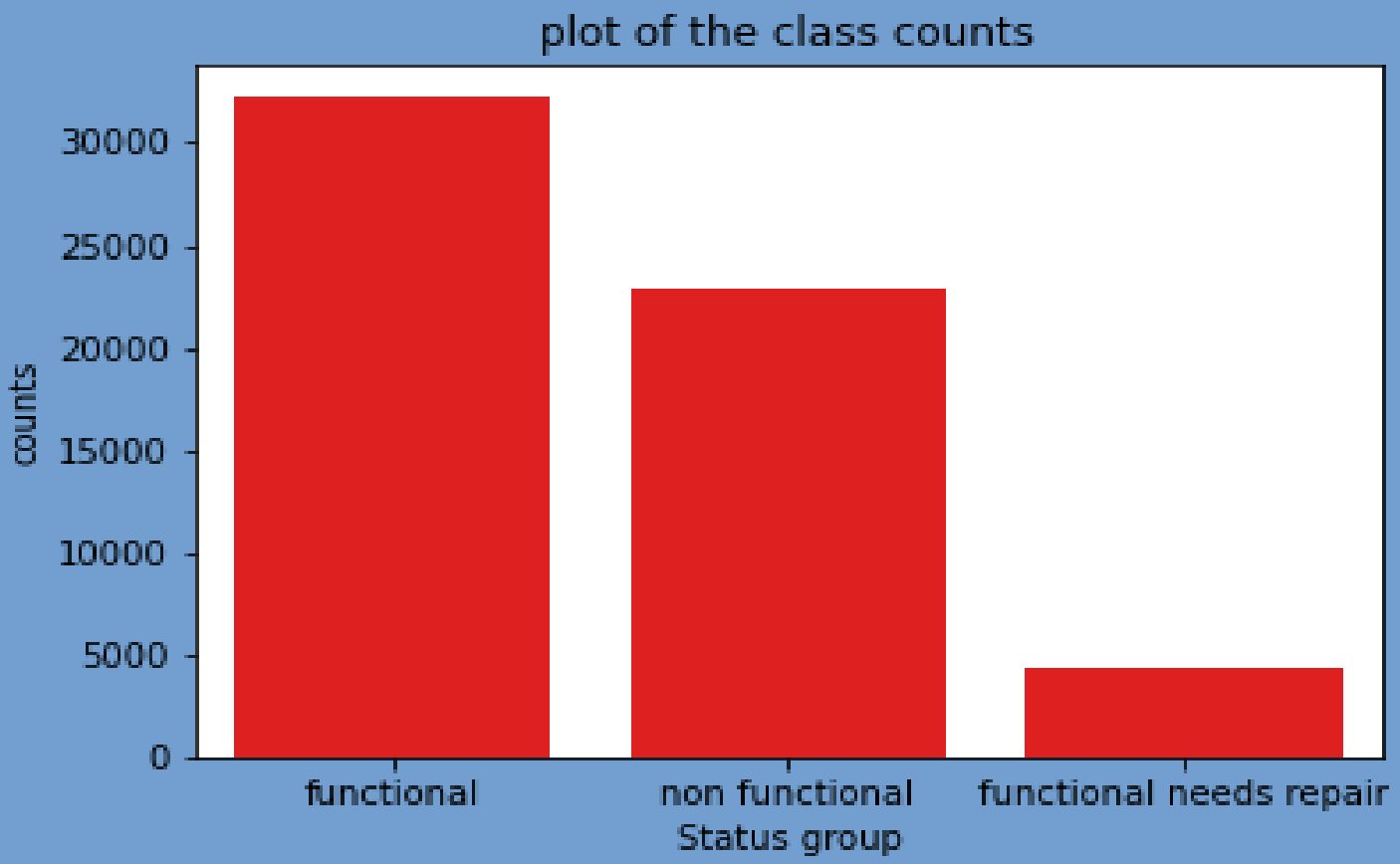
The purpose for this endeavor was to find come up with a model which could accurately predict whether a water point is functional or not.

Relationships

In the following slides, we will demonstrate some of the underlying relationships that were discovered during the data cleaning process.

Status of wells

During the data preparation phase, it was discovered that there were more functional wells compared to those that were either non functional but could be repaired and the completely non functional ones. With a small number of water points needing repair.

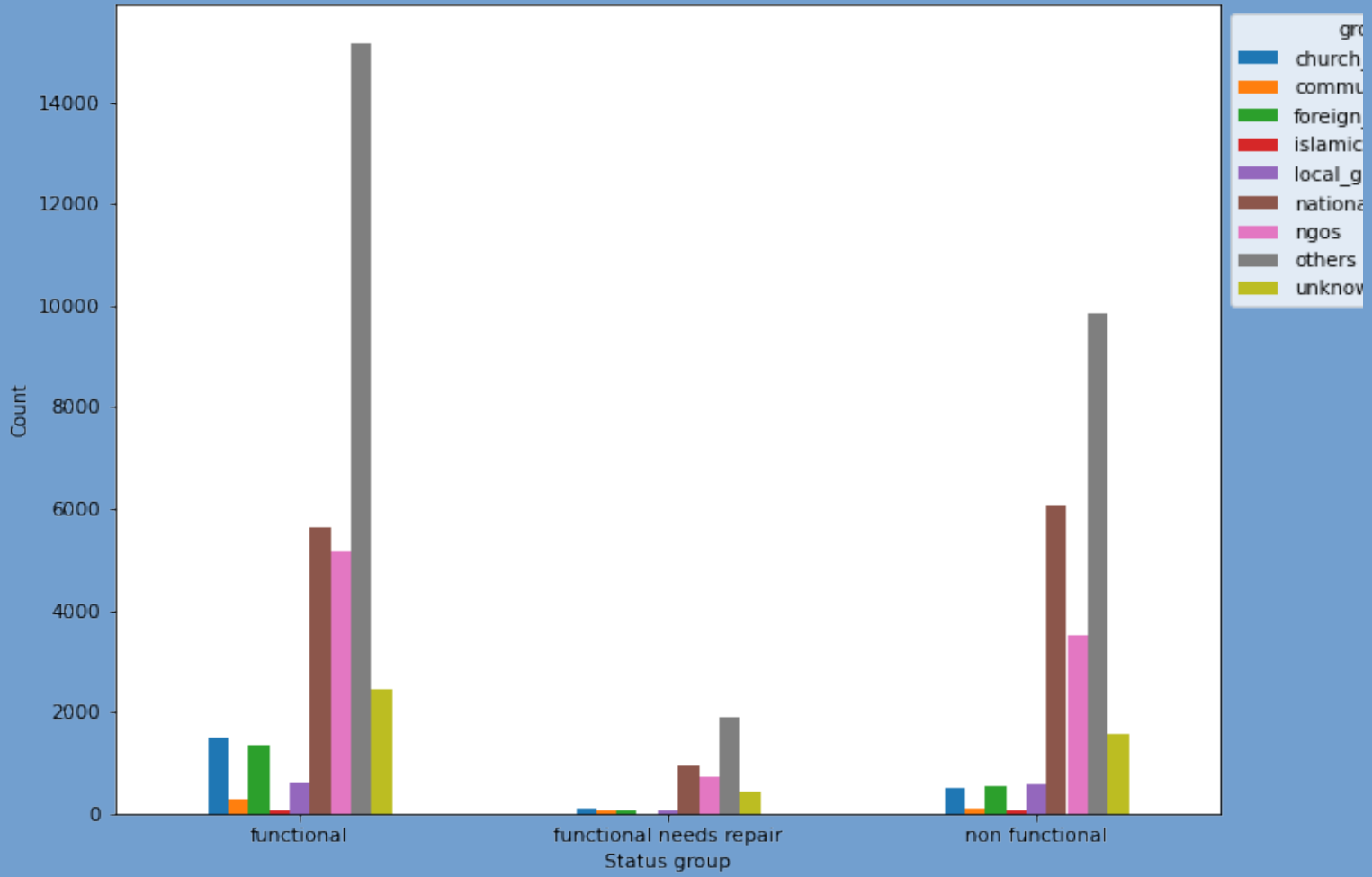

plot of the class counts
counts
30000
25000
20000
15000
10000
5000
0
functional
non functional
functional needs repair
Status group

Water point status by funding groups

Grouping water points according to the group that financed the implementation showed that:

- ***Most water points in Tanzania are funded by individuals.***
- ***Most functional water points in Tanzania are also funded by private individuals.***
- ***Most non functional water points are also funded by private individuals.***
- ***The national government funded water points are also the second most non functional water points.***

Distribution of status of water points wells according to financier groups

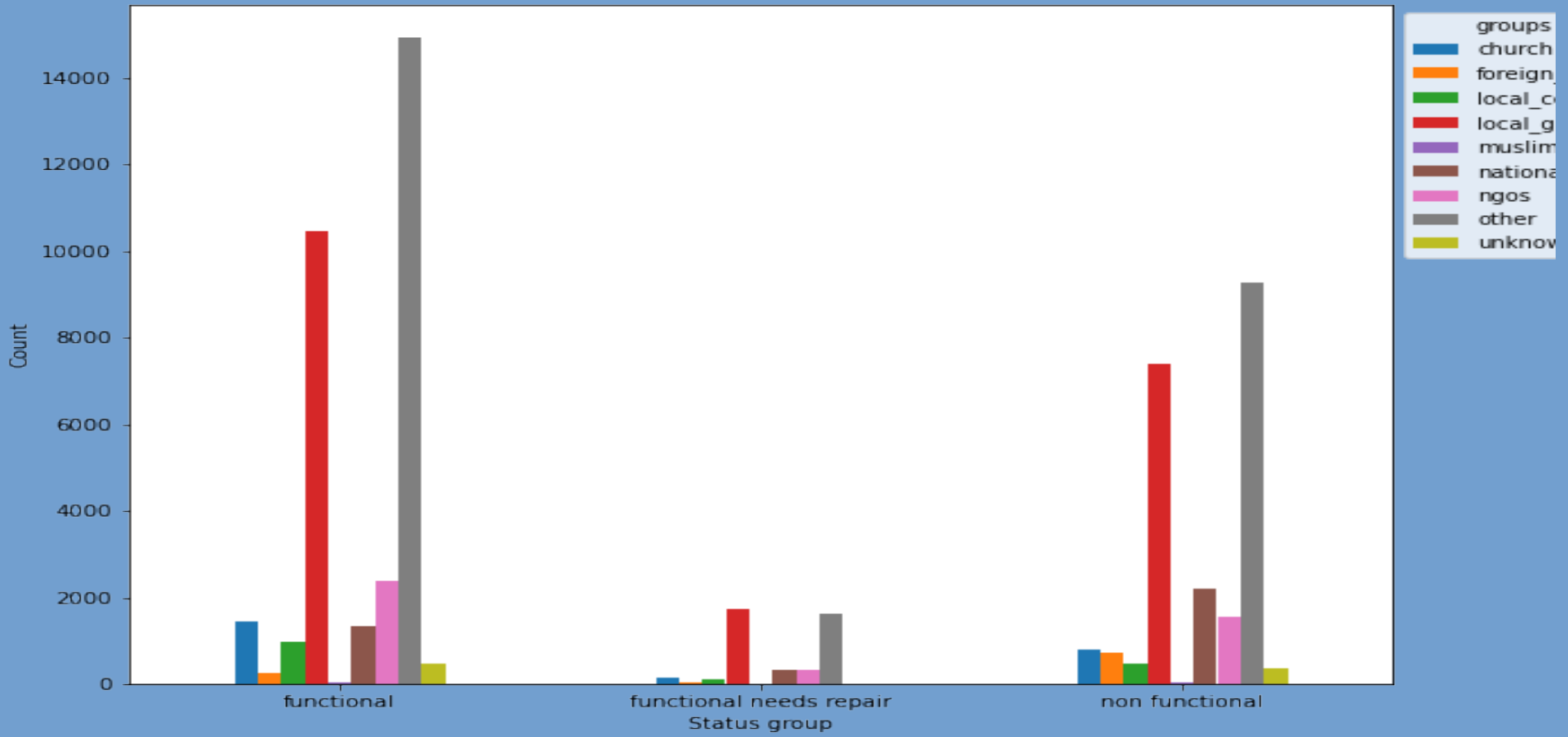


Water point status by installer groups

Grouping the water points according to the installer groups according to the current status showed that.

- ***Most functional water points were installed by private individuals.***
- ***The second most functional water points were installed by local governments.***
- ***Also on the non functional water points, most of them were installed by private individuals followed by the local government.***

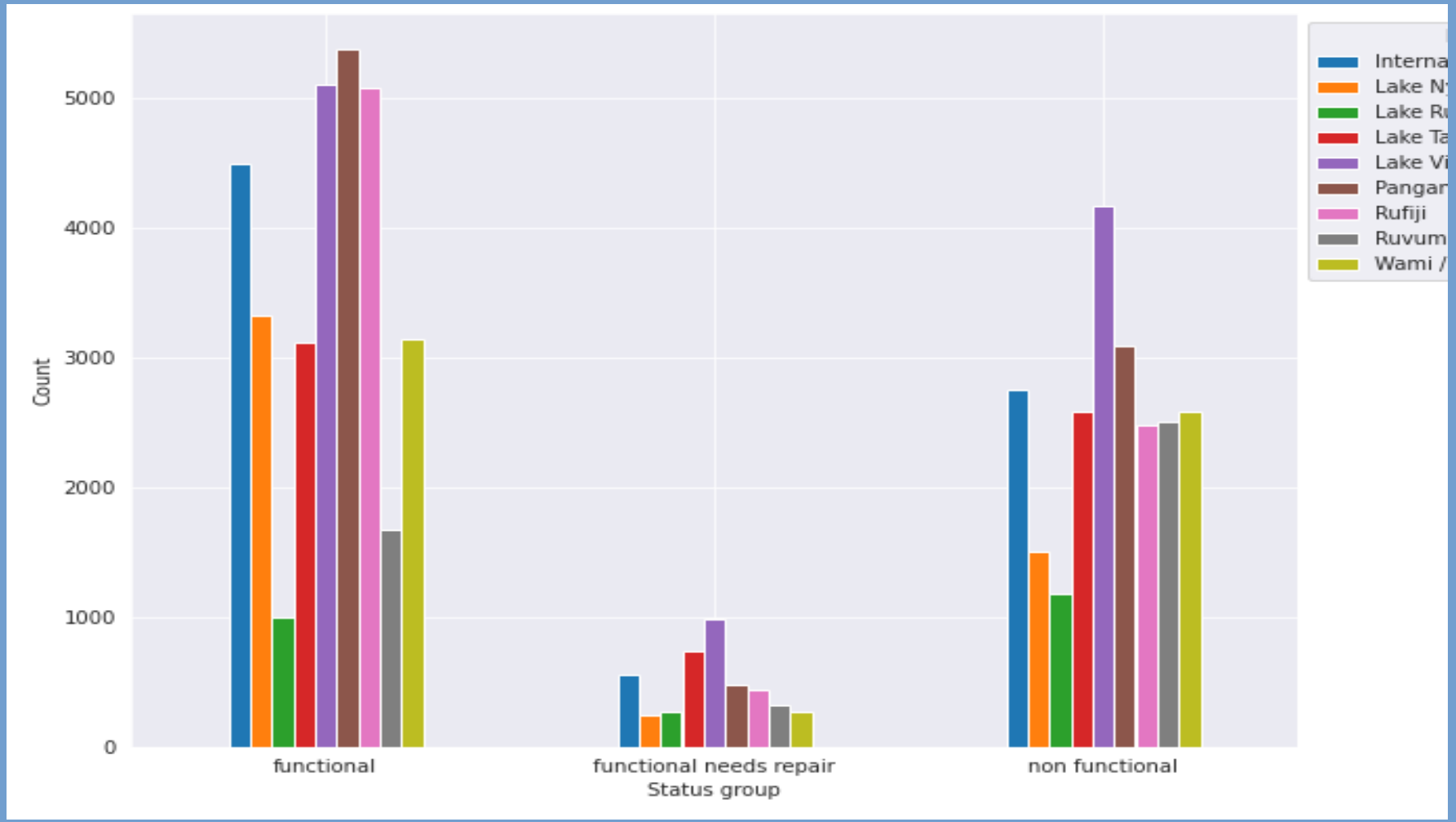
Graph of water points based on the installation groups



Water Basins

- ***It was observed that Pangani, Lake Victoria and Rufiji water basins had the highest functional water points.***
- ***Also internal water points had a high number of functional water points.***
- ***Lake Victoria has the highest number of non functional water points, followed by Pangani water basin.***

Distribution of water point status vs Water basin

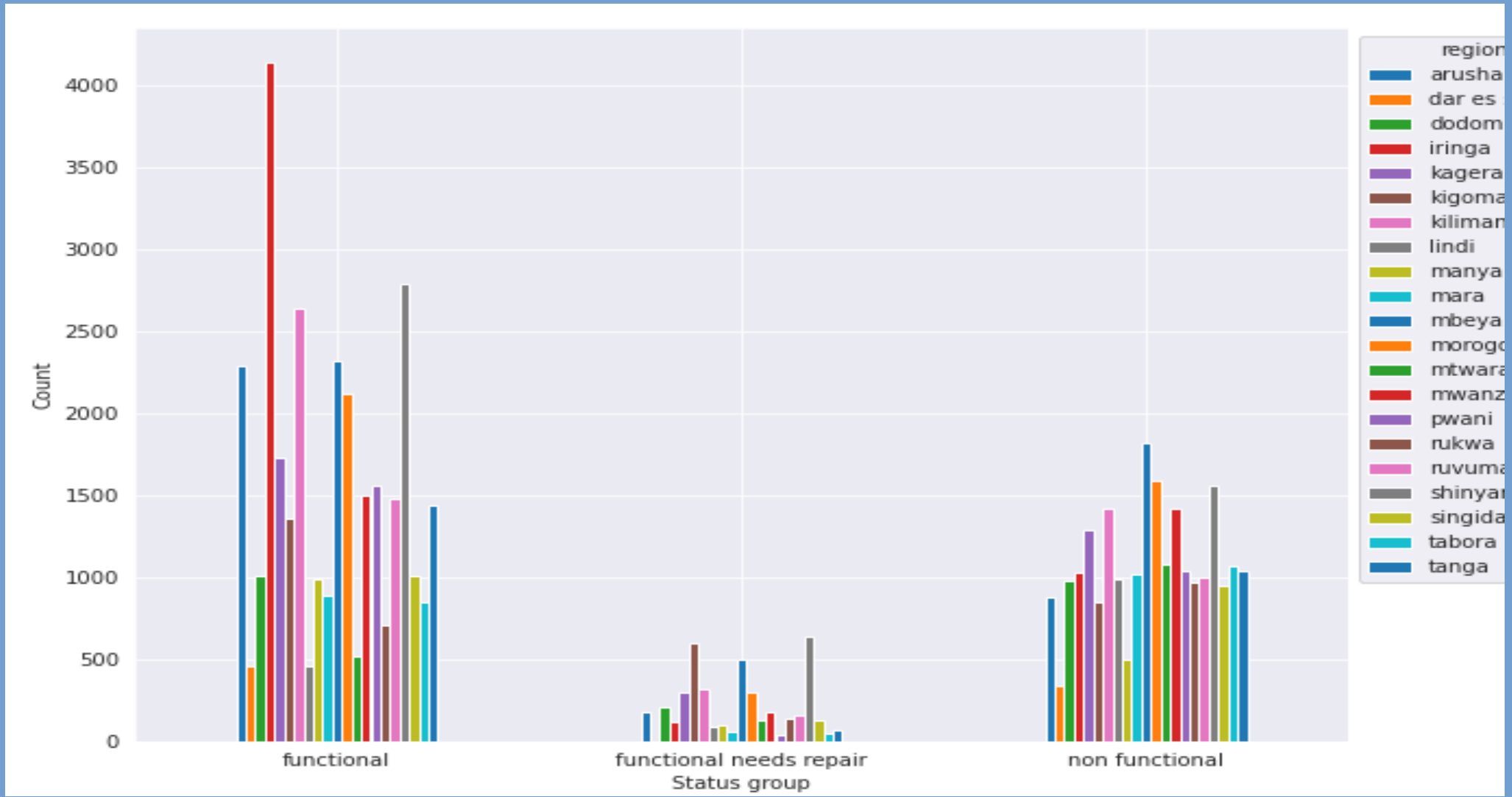


Region

The regions had varying number on the status of the water points.

- ***Iringa had the highest number of functional water points.***
- ***Morogoro had the lowest number of functional water points. It also had the lowest number of non functional water points.***
- ***The Tanga region had the highest number of non functional water points.***

Graph of water point status by region

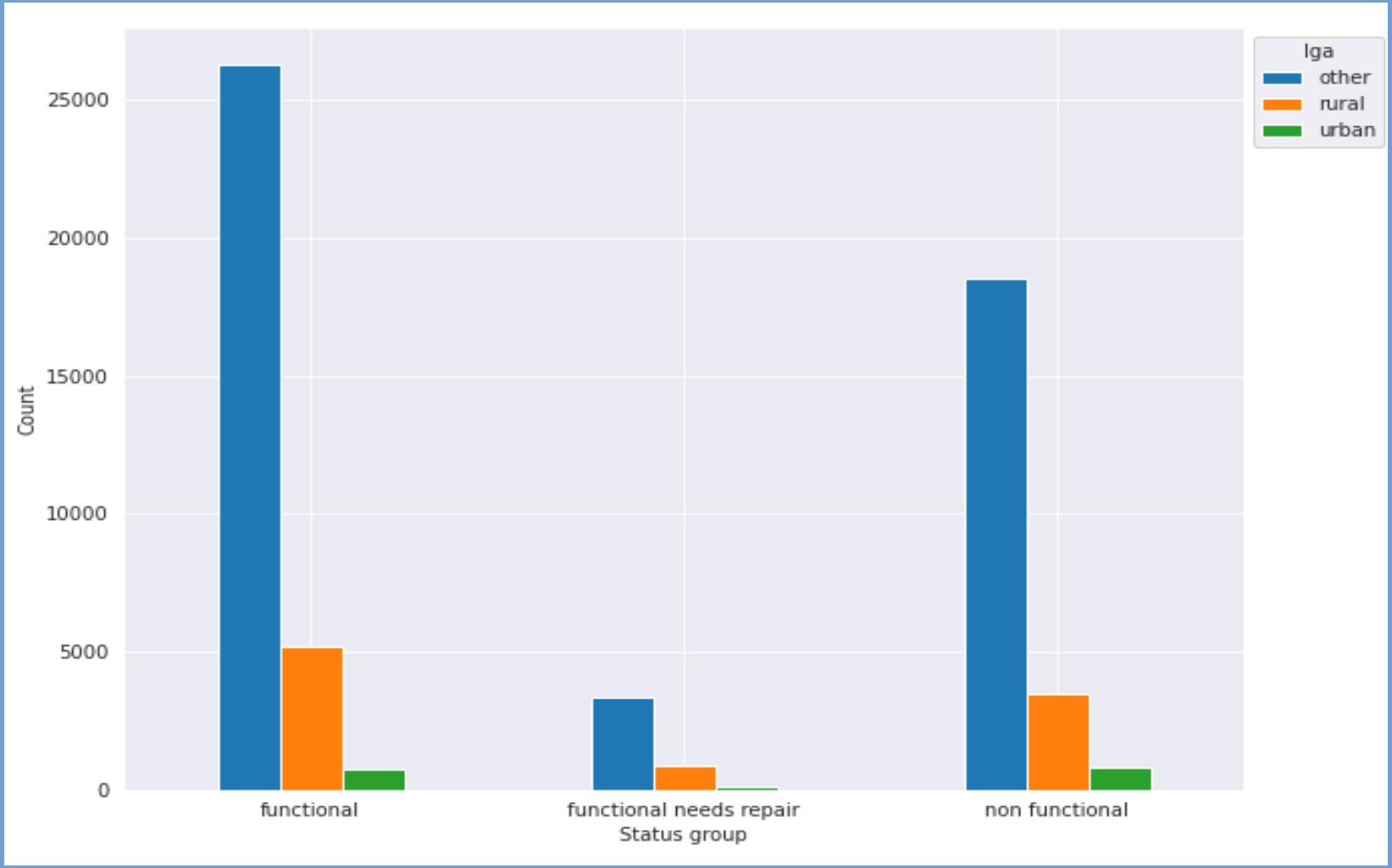


Urban, Rural and Semi Urban locations

It was observed that:

- ***Semi urban areas had the highest number of functional water points.***
- ***Urban areas had the least number of functional water points.***
- ***Rural areas still had a small number of functional water points compared to semi urban areas.***

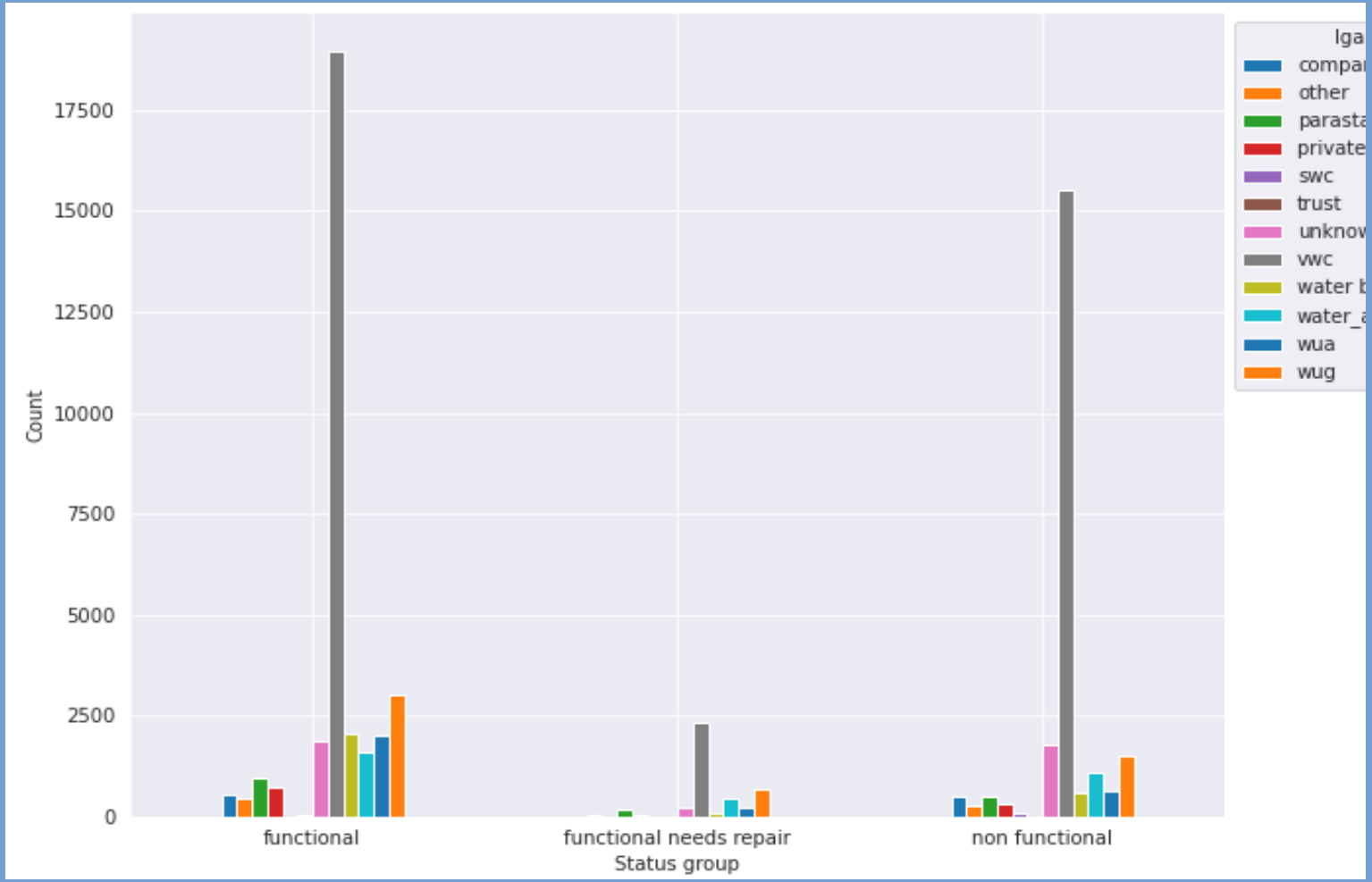
Graph of locations per water group



Scheme management

The VWC organization is managing the highest number of both functional and non functional water points.

Graph of town locations against water point status

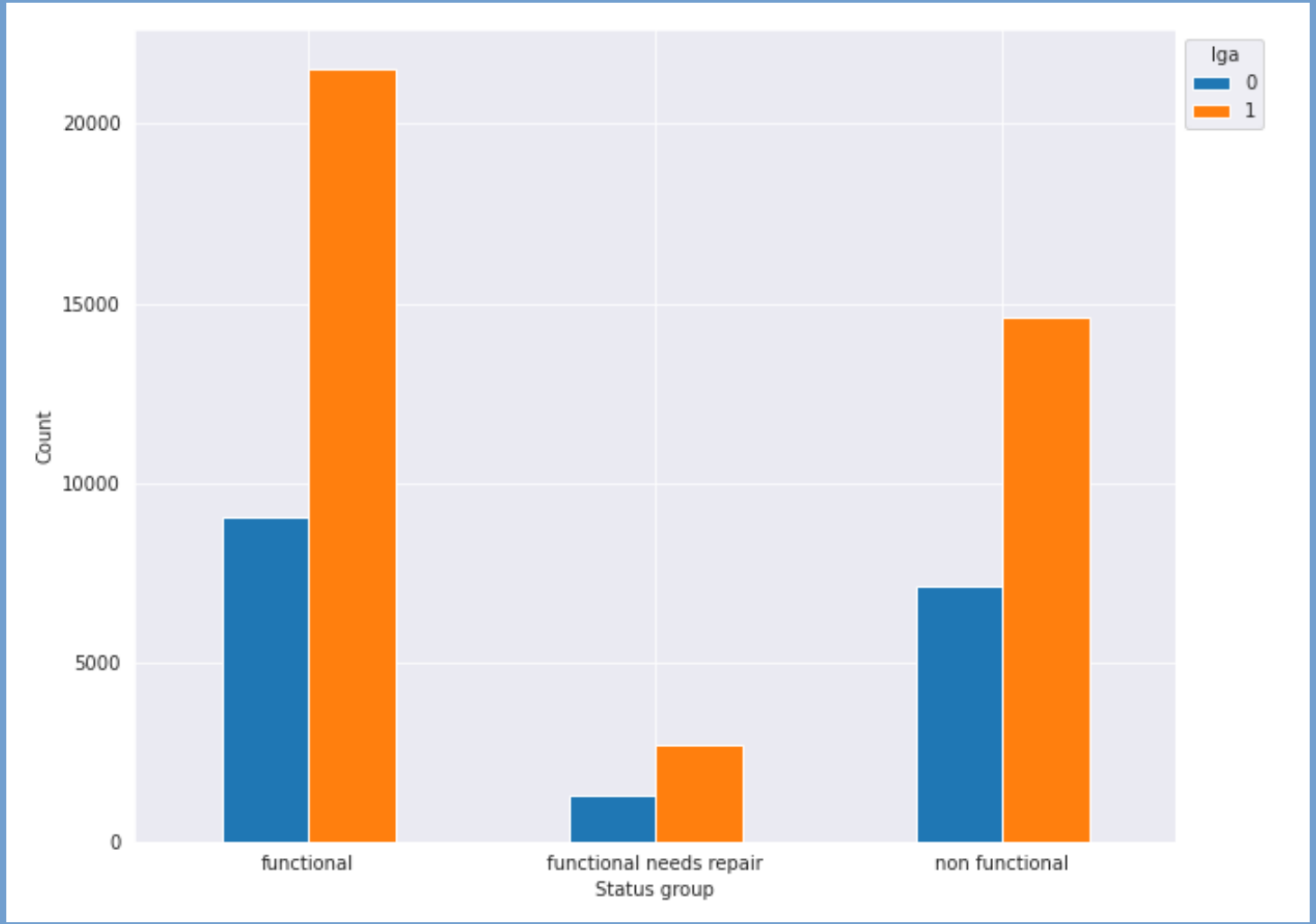


Permits

Observations

- ***There is a high number of functional water points which were formally approved.***
- ***Also there is a high number of non functional water points with formal approval.***

Graph of permits over the water point status

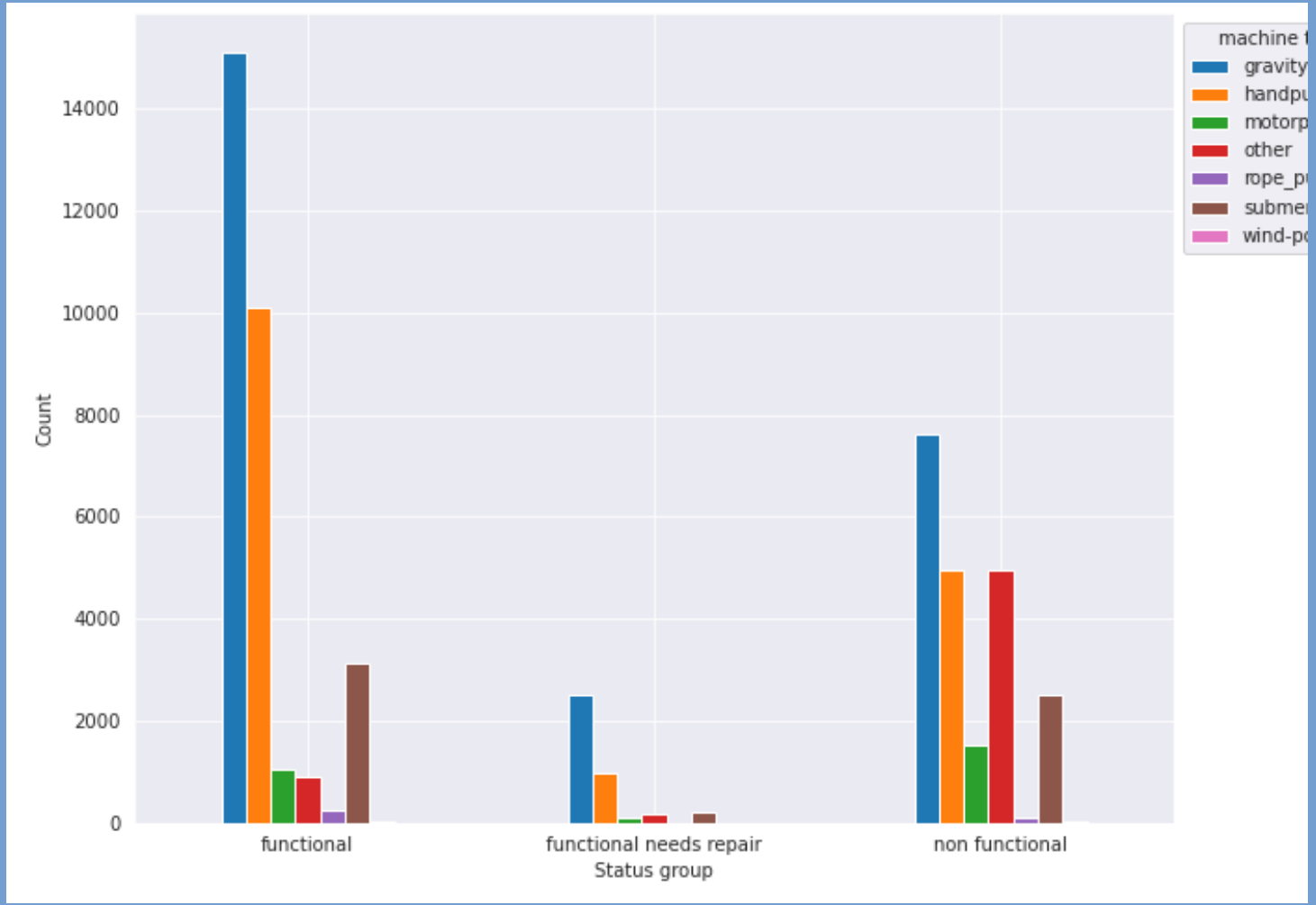


Water point extraction mechanism

Observations

- ***The gravity extraction method is the most popular extraction method.***
- ***Most functional and non functional water points are of gravity type extraction.***
- ***Hand pumps are the second most popular water extraction method***

Graph of extraction type per water point status

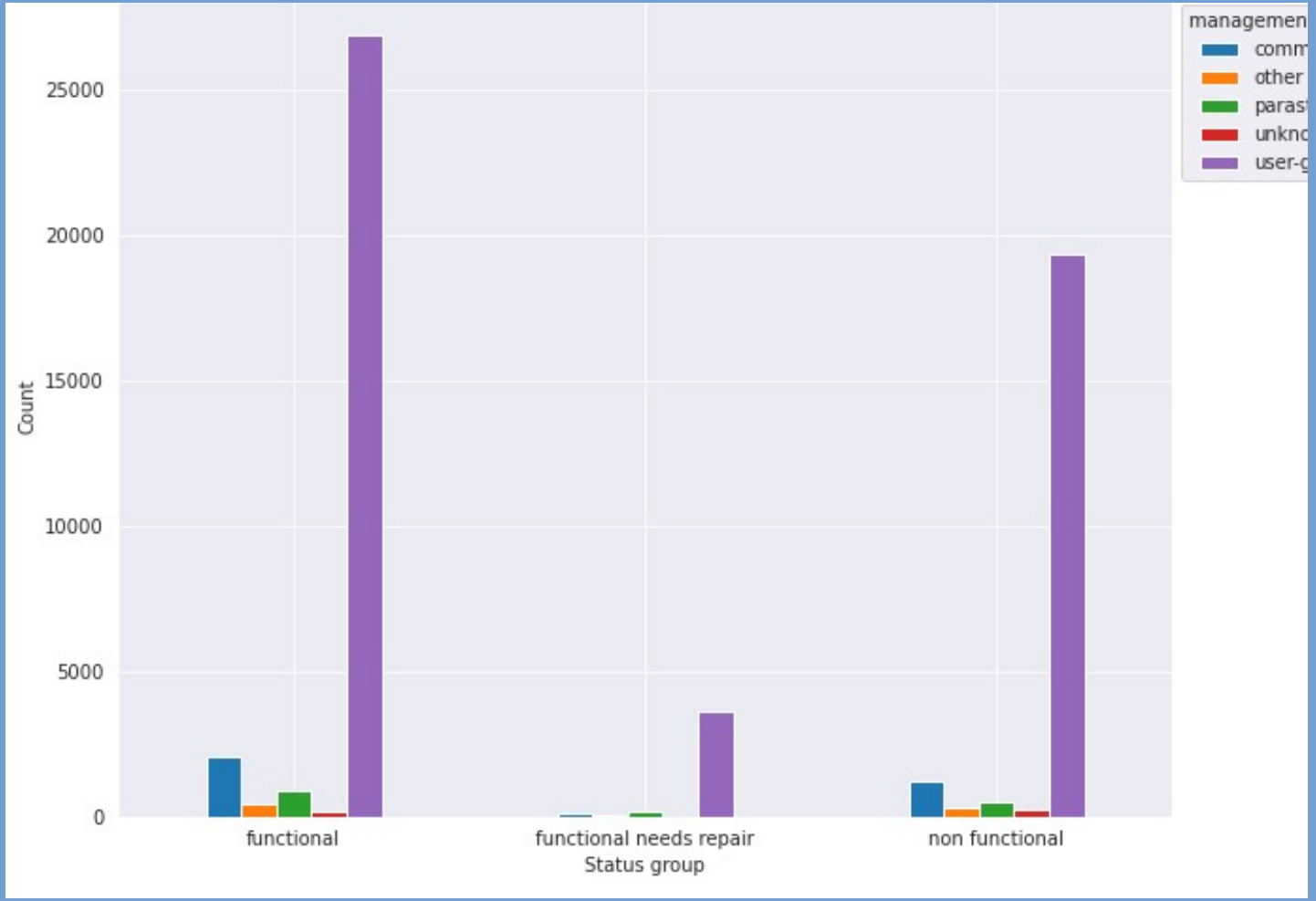


Management Group

Observations

- ***User group management groups is the most popular form of management of water points.***
- ***User group management have the most number of functional water points.***
- ***Also user group management groups have the most***

Graph of management based on water point status

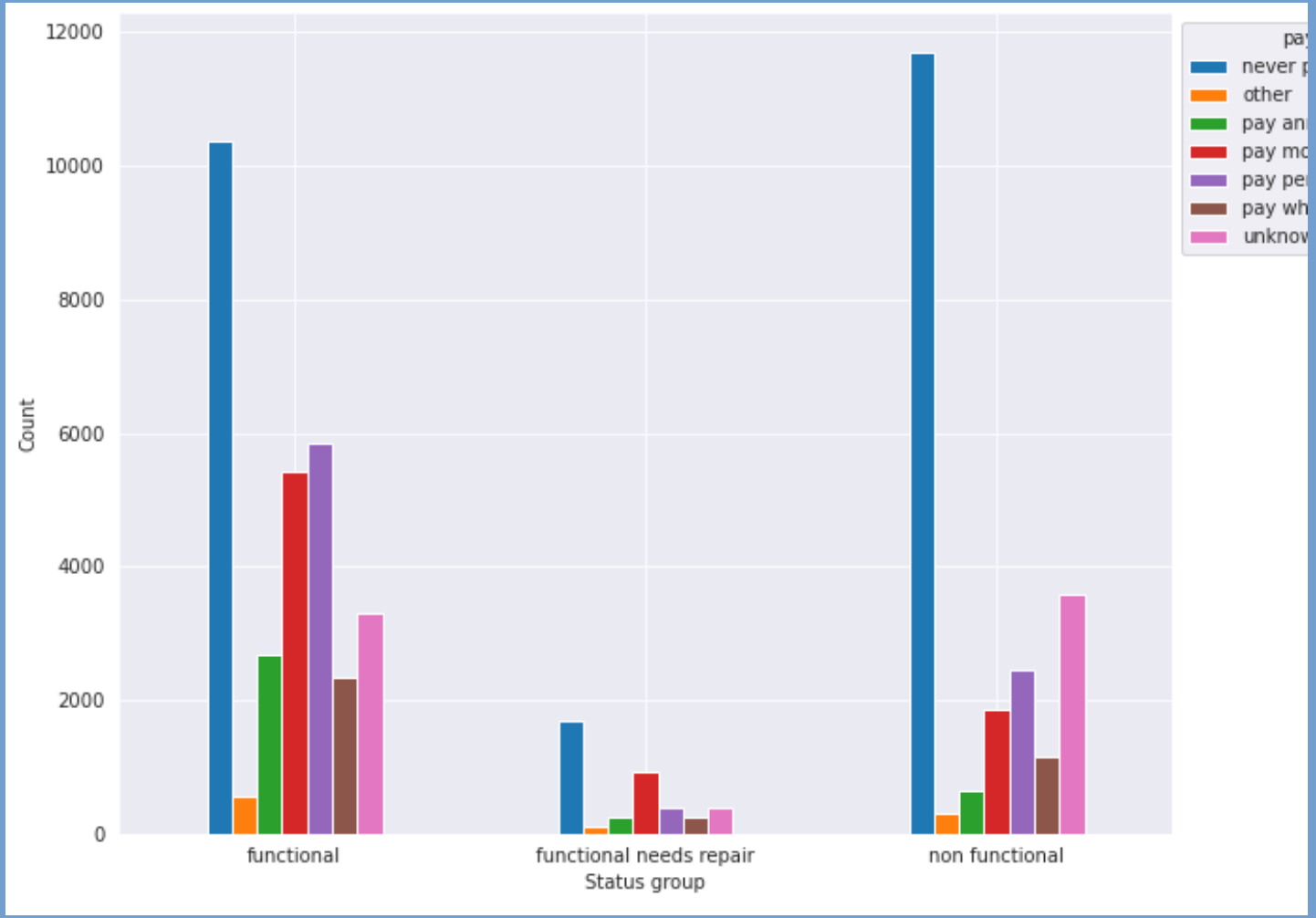


Payment

Observations

- ***Water points that are free are more popular on the whole.***
- ***Free water points still are the ones with the highest functional and non functional count.***
- ***We observed that with the water points that were not for free, there was a low count of non functionality in the water point.***

Graph of payment over water point status

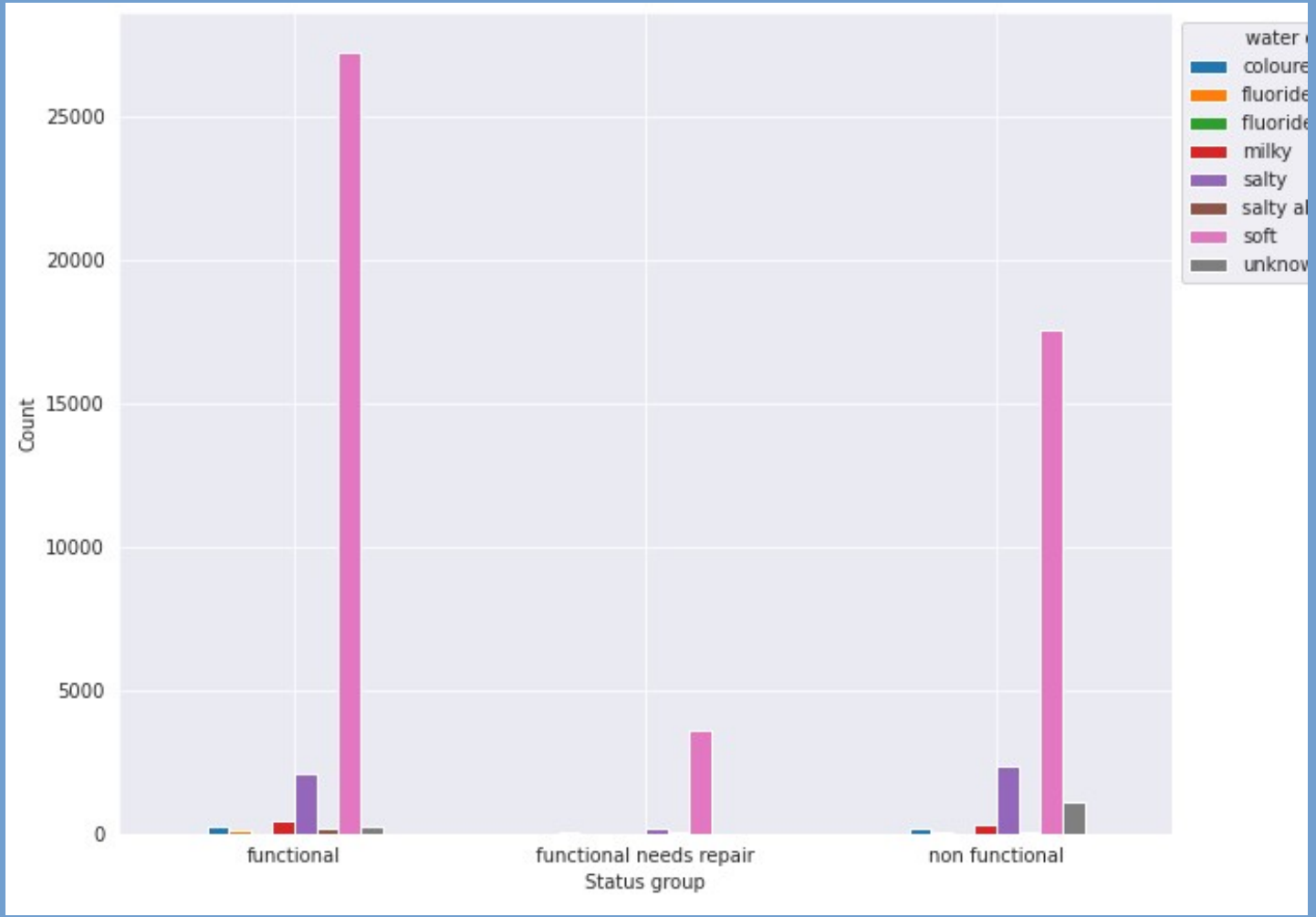


Water Quality

Observations

- ***Most functional water points have soft water.***
- ***Similarly non functional water points have soft water.***

Water Quality

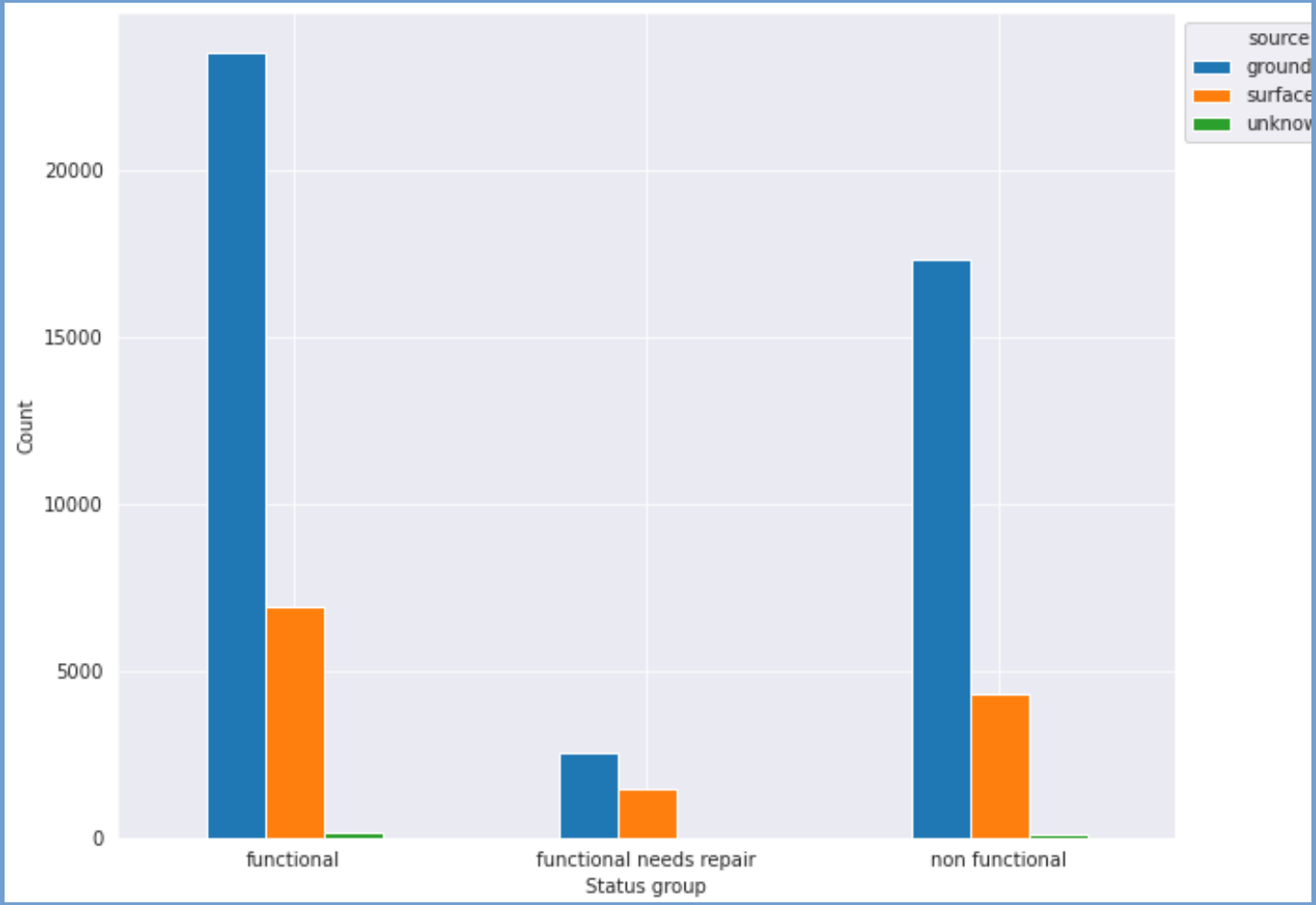


Water Source Class

Observations

- ***Ground water is the most common source of water.***
- ***Most functional and non functional water points extract water from the ground.***

Graph of water source over the water point status



Conclusions and Recommendations

I would recommend to the ministry to use the random forest model as it looks more promising compared to the other models. The model however struggles to predict the water points that are in need of repair, with about a less than half of a chance the model predicts if a water point needs repair correctly, which is more of guessing than anything else.

Model report

	Model precision
Functional	***81%***
Functional needs repair	***48%***
Non functional	***83%***

Accuracy ***80%***

Q & A