

# PREDICTING THE CONDITION OF TANZANIA WATER WELLS

BY WILFRED KIVINDA

OCT 2023

# CONTENT

- 01** BUSINESS UNDERSTANDING
- 02** DATA UNDERSTANDING
- 03** DATA PREPARATION
- 04** MODELLING & EVALUATION
- 05** EXTERNAL VALIDATION
- 06** CONCLUSION
- 07** RECOMMENDATIONS

# **1. BUSINESS UNDERSTANDING**

# PROBLEM STATEMENT

LACK OF CLEAN AND POTABLE WATER IS A MAJOR ISSUE IN COMMUNITIES ACROSS TANZANIA. THE TANZANIA MINISTRY OF WATER HAS INSTALLED SEVERAL WATER WELLS.

THE AIM IS TO IMPROVE MAINTENANCE OPERATIONS AND ENSURE THAT CLEAN AND POTABLE WATER IS AVAILABLE TO COMMUNITIES ACROSS TANZANIA.

# PROJECT GOAL

*THE GOAL OF THIS PROJECT IS TO BUILD A PREDICTIVE MODEL THAT CAN ACCURATELY PREDICT THE CONDITION OF WATER WELLS IN TANZANIA BASED ON THE VARIABLES PROVIDED IN THE DATA.*

# OBJECTIVES

## MAIN OBJECTIVE:

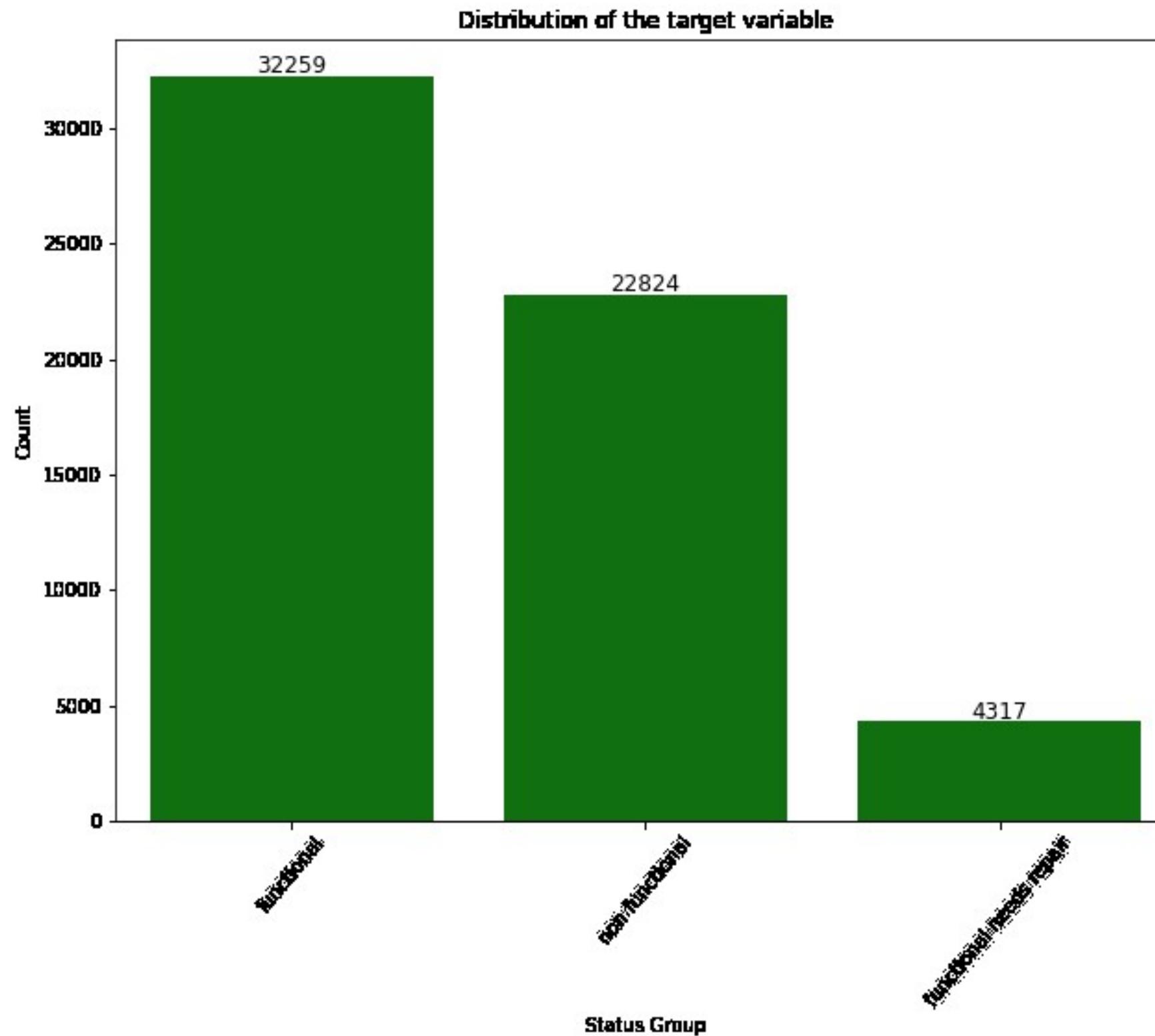
*TO PREDICT THE CONDITION OF WATER WELLS IN TANZANIA TO ENSURE THAT CLEAN AND POTABLE WATER IS AVAILABLE TO COMMUNITIES ACROSS TANZANIA.*

## SPECIFIC OBJECTIVES:

- 1. TO UNDERSTAND THE PROBLEM STATEMENT AND THE GOAL OF THE PROJECT.*
- 2. TO IDENTIFY THE VARIABLES THAT CAN IMPACT THE FUNCTIONALITY OF WATER WELLS.*
- 3. TO DETERMINE THE TARGET VARIABLE (FUNCTIONAL, NEED REPAIRS, OR NON-FUNCTIONAL).*

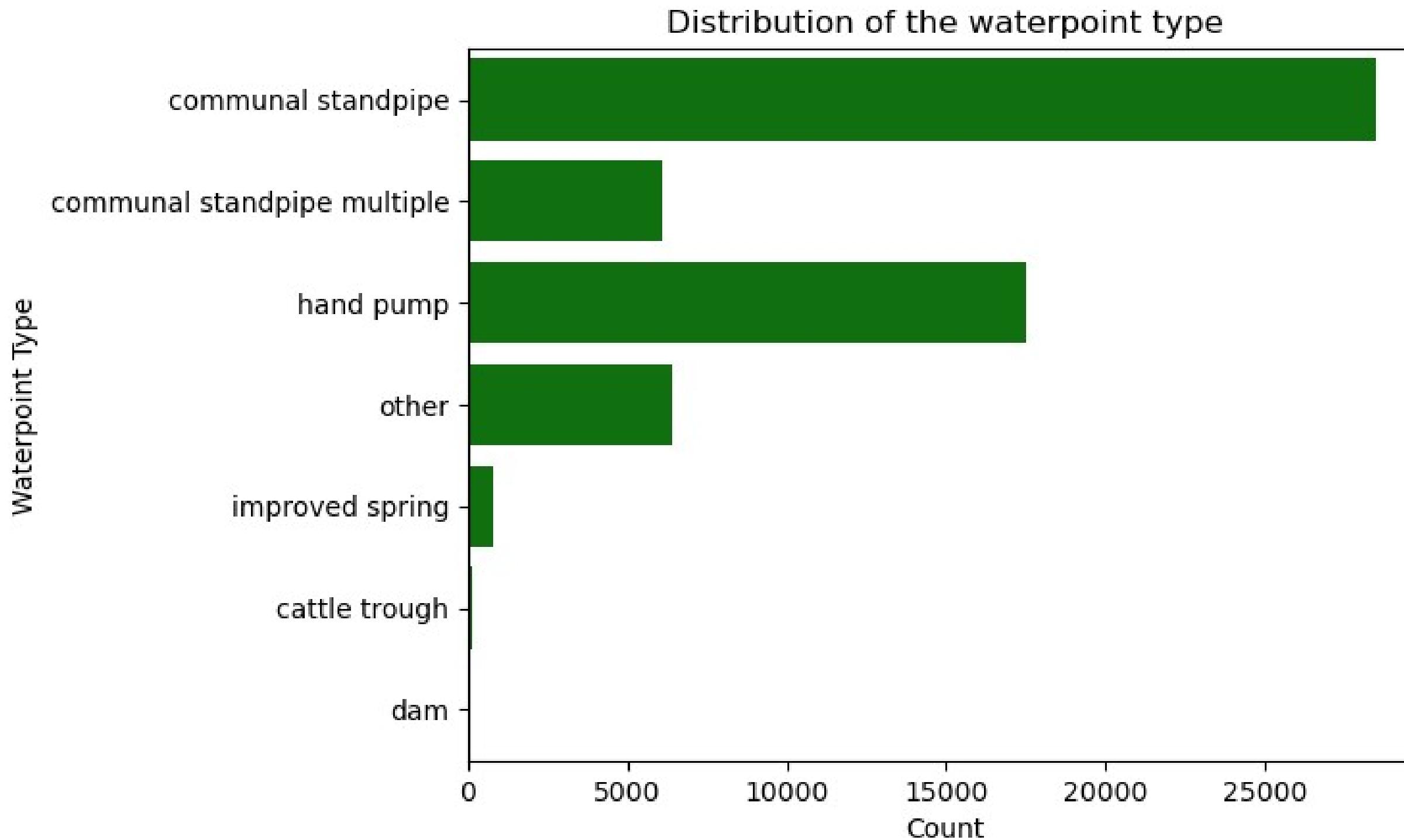
# 2. ANALYSIS

# STATUS GROUP



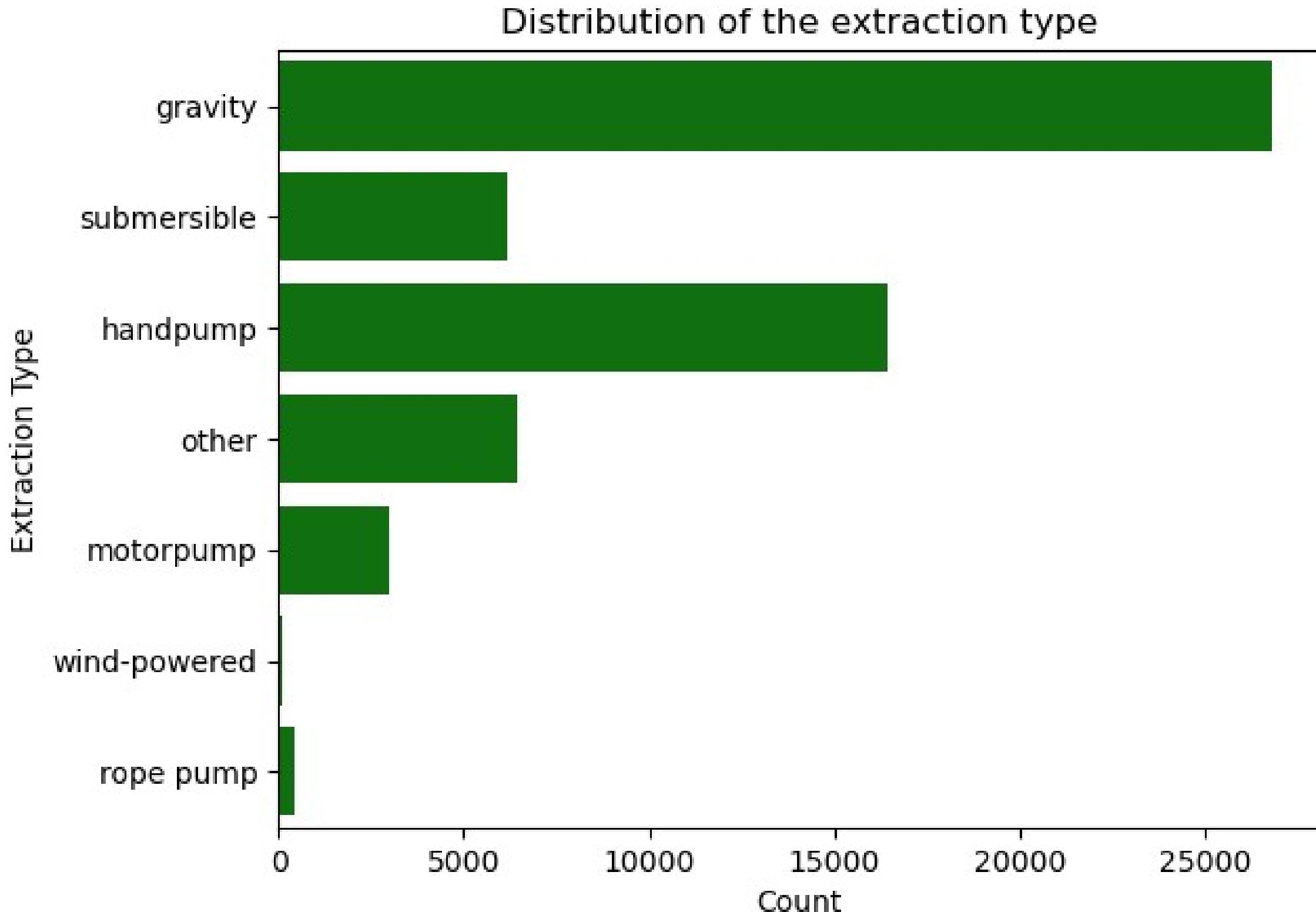
***THE MAJORITY CLASS IS  
THE FUNCTIONAL CLASS  
WHILE THE MINORITY IS  
THE FUNCTIONAL NEEDS  
REPAIR CLASS.***

# WATERPOINT TYPE



***THE TYPE OF WATERPOINT FOR  
MOST WELLS IS THE COMMUNAL  
STANDPIPE FOLLOWED BY HAND  
PUMP***

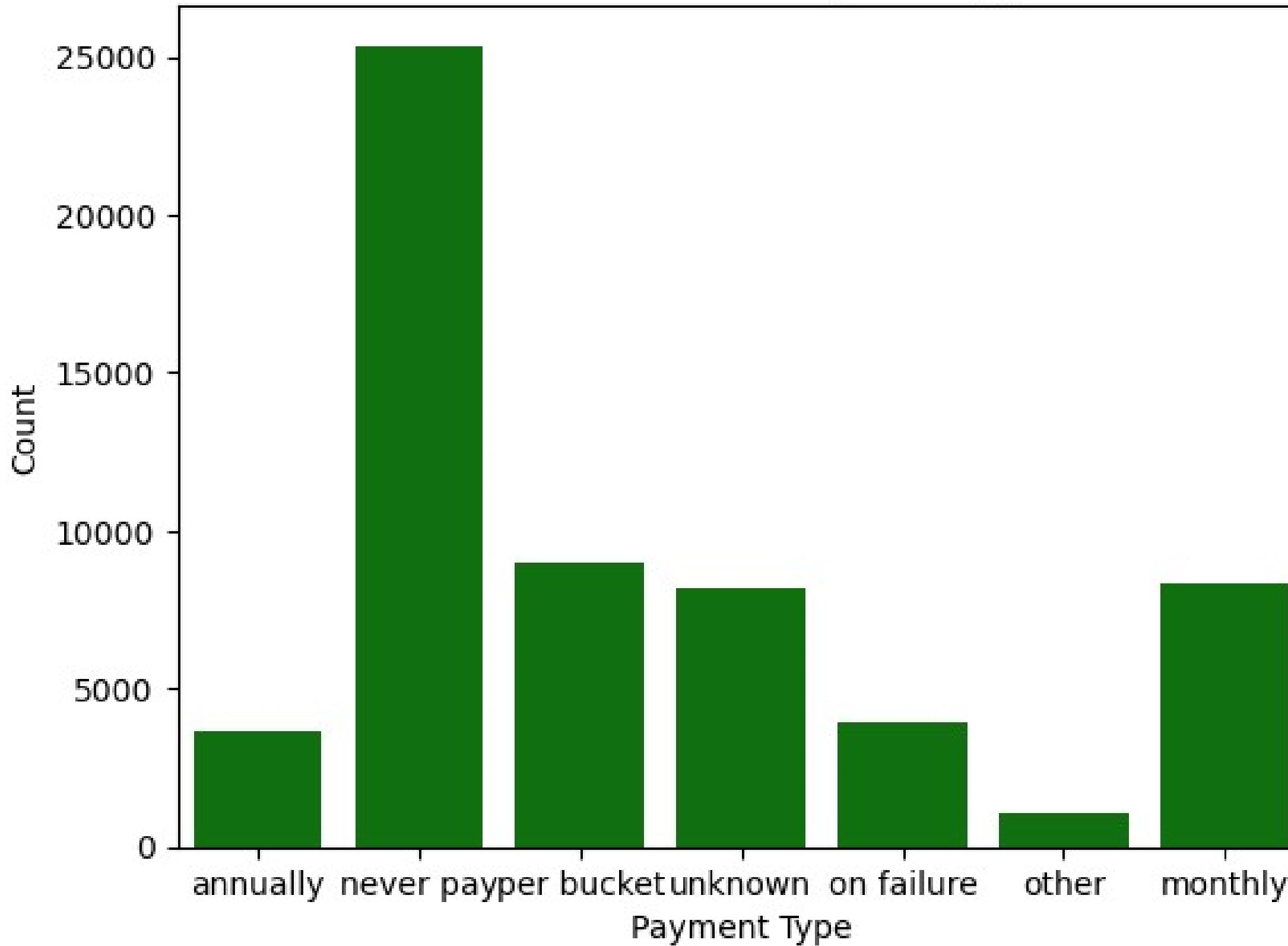
# EXTRACTION TYPE



***THE EXTRACTION TYPE FOR MOST  
WELLS IS THROUGH GRAVITY***

# PAYMENT TYPE

Distribution of the Payment type



***MOST WATER POINTS ARE NEVER PAID FOR.***

# **3. DATA PREPARATION**

**1. HANDLING MISSING VALUES**

**2. FEATURE SELECTION**

**3. ENCODING CATEGORICAL**

**4. VARIABLES**

**5. SCALING**

**6. HANDLING CLASS IMBALANCE**

# **4. MODELLING AND EVALUATION**

## MODELING AND EVALUATION

- *THE BEST MODEL (RANDOM FOREST) WAS SELECTED FROM FIVE CLASSIFIER MODELS.*
- *THIS MODEL WAS TUNED TO IMPROVE ITS PERFORMANCE.*
- *THE ACCURACY OF THE MODEL IS 0.8229, WHICH MEANS THAT THE MODEL CORRECTLY PREDICTS THE STATUS GROUP WITH AN ACCURACY OF 82.29%*

# **5. EXTERNAL VALIDATION**

# EXTERNAL VALIDATION

*THE CLASSIFICATION RATE FOR THE SUBMISSION IS 0.8148 WHICH MEANS IT WORKS WELL ON BOTH SEEN AND UNSEEN DATA TO PREDICT THE WATERWELLS CONDITION.*

# **6. CONCLUSION AND RECOMMENDATIONS**

# CONCLUSION

*THE MODEL COULD BE FURTHER IMPROVED BY INCORPORATING MORE DATA ESPECIALLY FOR THE FUNCTIONAL NEEDS REPAIR CLASS TO HANDLE IMBALANCE FOR THE CLASSES.*

# RECOMMENDATIONS

THE TANZANIA MINISTRY OF WATER SHOULD INVEST IN BETTER WATERPOINT TYPES SUCH AS COMMUNAL STANDPIPES AND HAND PUMPS.

THE TANZANIA MINISTRY OF WATER SHOULD ENSURE THAT THE EXTRACTION TYPE FOR THE WELLS IS MOSTLY THROUGH GRAVITY AND HAND PUMP.

THE TANZANIA MINISTRY OF WATER SHOULD ENSURE THAT THE GPS HEIGHT (ALTITUDE OF THE WELL) FOR MOST WATER POINTS IS HIGH ENOUGH.

THE TANZANIA MINISTRY OF WATER SHOULD ALSO ENSURE THAT THE PEOPLE USING THE WATERPOINTS PAY EITHER MONTHLY OR ANNUAL

**THANK YOU !!**