

# **TANZANIA MINISTRY OF WATER PREDICTIVE** **MODELLING**

PREDICTING FUNCTIONALITY OF WATER PUMPS IN TANZANIA

By : Cynthia Adisa

## OBJECTIVES

- 1) TO PREDICT THE CONDITION OF A WATER PUMP BASED ON ITS AGE.
- 2) TO UNDERSTAND WHICH REGIONS ARE IN NEED OF WATER PUMPS.
- 3) UNDERSTANDING THE FACTORS THAT AFFECT THE LIFESPAN OF THE PUMPS.



- 1) BUSINESS UNDERSTANDING
- 2) DATA UNDERSTANDING
- 3) DATA PREPARATION
- 4) MODELLING
- 5) RESULTS & RECOMMENDATION
- 6) CONCLUSION

**CONTENT**

## PROJECT OVERVIEW

- TANZANIA, AS A DEVELOPING COUNTRY, STRUGGLES WITH PROVIDING CLEAN WATER TO ITS POPULATION OF OVER 57,000,000. THERE ARE MANY WATER POINTS ALREADY ESTABLISHED IN THE COUNTRY, BUT SOME ARE IN NEED OF REPAIR WHILE OTHERS HAVE FAILED ALTOGETHER.



## BUSINESS UNDERSTANDING

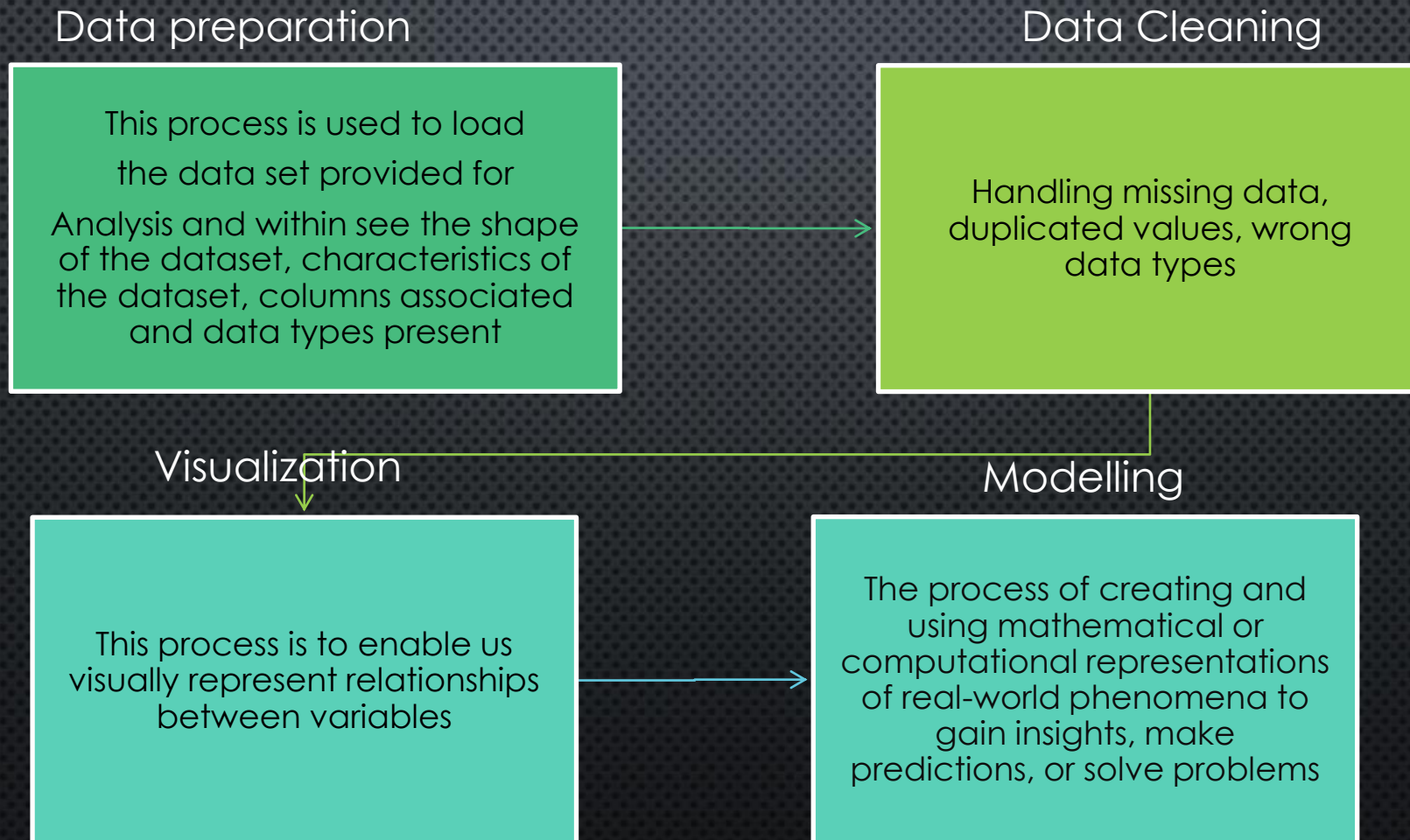
- TANZANIA MINISTRY OF WATER NEEDS TO IDENTIFY WHICH PUMPS ARE FUNCTIONAL, WHICH NEED SOME REPAIRS, AND WHICH DON'T WORK AT ALL THAT WILL INFLUENCE THEIR DECISION TO EITHER GET NEW PUMPS OR REPAIR THE ONES EXISTING.
- THEY ALSO NEED TO DEVELOP A MODEL TO IDENTIFY AND IMPROVE ON MAINTENANCE OPERATIONS OF THE PUMPS.

## DATA UNDERSTANDING

- THE DATA WAS OBTAINED FROM TAARIFA AND THE TANZANIAN MINISTRY OF WATER CONTAINING THE FOLLOWING :
  - a) TEST SET VALUES :THE INDEPENDENT VARIABLES THAT NEED PREDICTIONS
  - b) TRAINING SET LABELS: THE DEPENDENT VARIABLE (STATUS\_GROUP) FOR EACH OF THE ROWS IN TRAINING SET VALUES
  - c) TRAINING SET LABELS: THE INDEPENDENT VARIABLES FOR THE TRAINING SET



# DATA PREPARATION



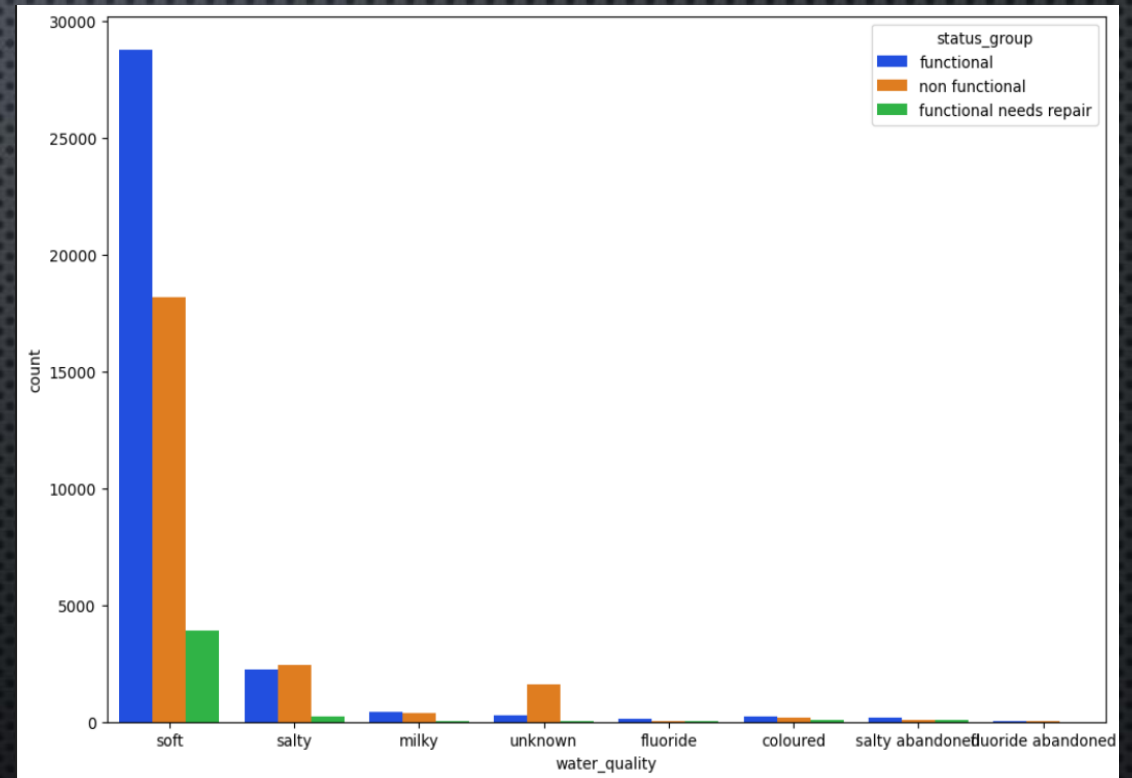
# MODELLING

- ABLE TO DO A TRAIN TEST SPLIT IN THE MODELS: RANDOM FOREST CLASSIFIER, LINEAR REGRESSION AND K-NEAREST NEIGHBORS MODELS OPTIMIZED BY ONE HOT ENCODING AND SCALING.
- FOR THE DATASET PROVIDED I WAS THE METHOD ADOPTED WAS THE **KNN** MODEL AS IT PRODUCED THE MOST FAVORABLE PREDICTIVE RESULTS OF THE THREE MODELS WITH AN ACCURACY OF 66%.



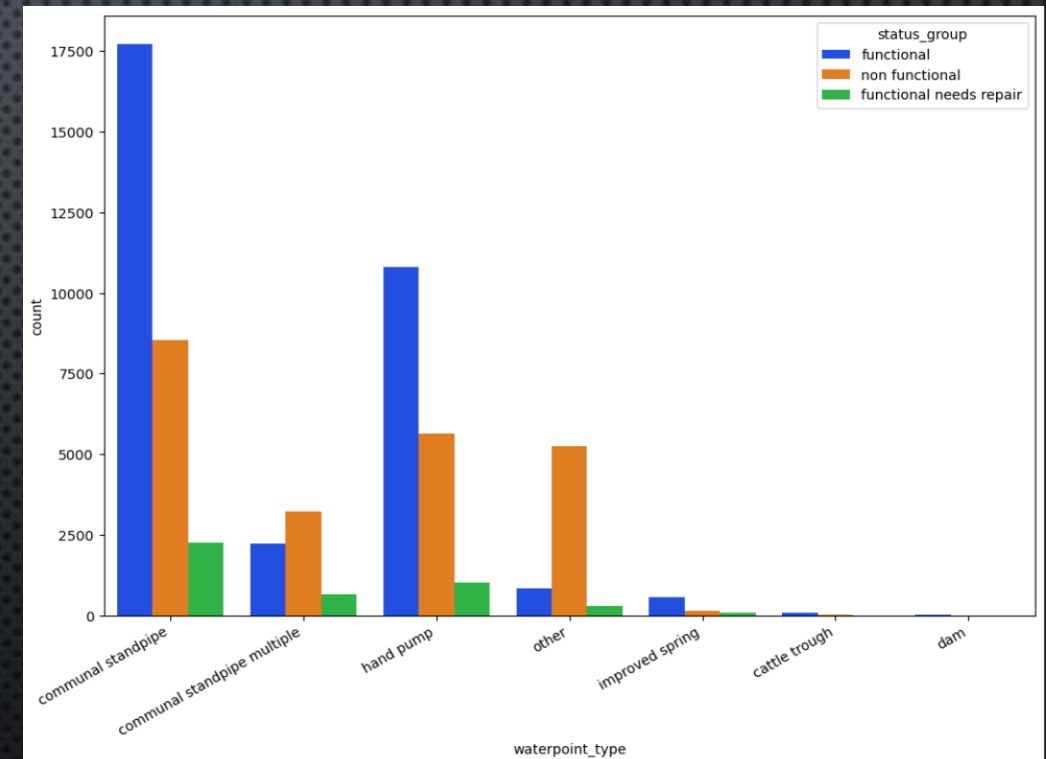
## RESULTS & RECOMMENDATIONS

- THE TANZANIA WATER MINISTRY SHOULD CONSIDER INVESTING MORE IN AREAS WITH SOFT WATER AS OPPOSED THE OTHER WATER TYPES I.E. SALTY, MILKY, FLUORIDE SEEM TO BE AFFECTING THE LIFESPAN OF THE PUMPS.
- THE MINISTSRY SHOULD ALSO FOCUS ON DRAWING WATER FROM SPRINGS AS WELL.



## RESULTS & RECOMMENDATIONS CONT...

- THEY CAN ALSO INVEST MORE ON COMMUNAL SANDPIPES AND HANPIPED PUMPS AS THEY ARE THE MOST FUNCTIONAL





## CONCLUSION

- THE MODEL SHOWS HOW VARIABLES SUCH AS WATER TYPE, LOCATION, PUMP TYPE QUITE SIGNIFICANTLY AFFECT THE PERFORMANCE OF A PUMP.

# THANK

# YOU

CYNTHIA.ADISA@STUDENT@MORINGASCHOOL.COM