

# Group 8 Final Presentation

Russell and Frances

September 2022

## Group 8

### **Frances Smith**

email: frances.j.smith.nz@gmail.com

ORCHID ID: 0000-0002-5168-3134

### **Russell Syder**

email: russellsyder@gmail.com

ORCHID ID: 0000-0002-4582-5909



## Textual Description of the Dataset

Our dataset contains information from 3801 lakes in New Zealand. This dataset was extracted from Stats NZ.

<https://www.stats.govt.nz/indicators/modelled-lake-water-quality/>

### Variables

For analysis we split the dataset into two main categories; the lake health variables and the lake dimension variable. The lake health variables measure as a whole give an indication of the “health” of an individual lake. The five lake health variables are Clarity, Ammoniacal Nitrogen, Total Nitrogen, Total phosphorus, and Chlorophyll-A. The lake dimension variables measure the dimensions of the lake. The three lake dimension variables are depth, area, perimeter. Additional variables that we examined were:

- Region; which New Zealand region the lake was located in, and
- Dominant landcover; split into four types; Exotic Forest, Native, Pastoral and Urban area. The category Urban area is applied if urban cover exceeds 15 percent of catchment area. Pastoral is applied if pastoral exceeds 25 percent of catchment area and not already assigned urban. The other three categories; Exotic forest, Native, or Other were assigned according to the largest land cover type by area, if not already assigned urban or pastoral.

## Lake Health Variables

*Ammoniacal Nitrogen* is a form of nitrogen that supports algae and plant growth, but in large concentrations can be toxic to aquatic life. This is measured in milligrams per litre. The national bottom line for this measure is 1.3mg/L, which none of the observations exceed.

*Chlorophyll-a* is an organic molecule found in plant cells that allows plants to photosynthesize. The variable Chlorophyll-a is a measure of the concentration of phytoplankton biomass in milligrams per cubic metre. High concentrations of chlorophyll is a symptom of degraded water quality. The national bottom line for this measure is 12.

*Total Phosphorus* is the sum of all phosphorus forms in the water, including phosphorus bound to sediment. Large amounts of phosphorus in lakes can reduce dissolved oxygen in the water. This can cause low oxygen areas in the lake, where some aquatic life cannot survive. Total Phosphorus is measured in milligrams per cubic metre and has a national bottom line of 50mg/m<sup>3</sup>.

*Clarity* is measured in Secchi depth. This is the maximum depth (in metres) a black and white Secchi disk is visible from the surface of the lake.

# Exploratory Data Analysis

Perform an EDA of the variables that we grouped together. At this point in time, we did not have a specific leading question that we were interested in. As such, we thought it would be prudent to make the goal of our EDA to extract as much interesting information from our data as we could; then use the results to come up with leading questions.

RussBuss - Today at 1:00 PM Once the EDA was performed we came up with some interesting questions that we could investigate further for the Preliminary Report. WE then conducted more analyses. The results of this were then refined. The results of these were then refined into our Final Report.

There are 12 lakes with no entry for dominant land cover, however in the description of the dataset by Stats NZ, it states all lakes have been categorised, and indicated these empty entries should be another category called 'other' that includes 'Gorse and/or Broom', 'Surface mines and dumps', 'Mixed exotic shrubland', and 'Transport infrastructure' so we have assigned these to the 'other' category.