Group 8 Presentation 1

Russell and Frances

September 10th 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. A high value in any of these variables is an indicator of possible poor lake health (with the exception of clarity where a lower value is an indicator of poor lake health). 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.
- Dominant landcover; split into Exotic Forest, Native, Pastoral, Urban Area and Other, Dominant landcover describes the landscape surrounding the lake.

Visualisation 1

Our first visualisation is an example of one of the histograms that we produced to visualise the distribution of the lake health variables (shown in figure 1).

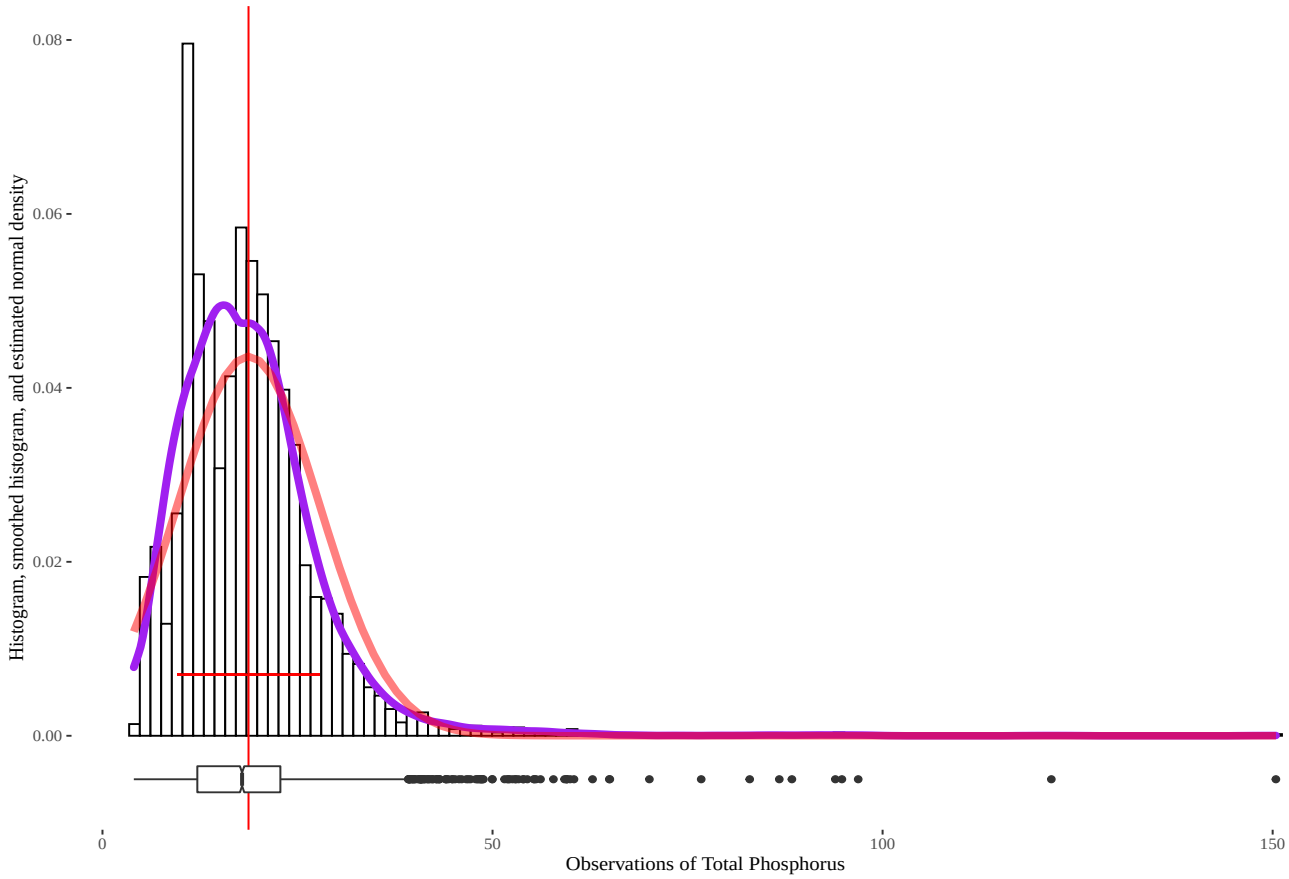

Figure 1: Histogram of Total Phosphorus

Visualisation 2

Our second visualisation is are side-by-side boxplot of the log-transformed lake health variables by dominant landcover in figure 2.

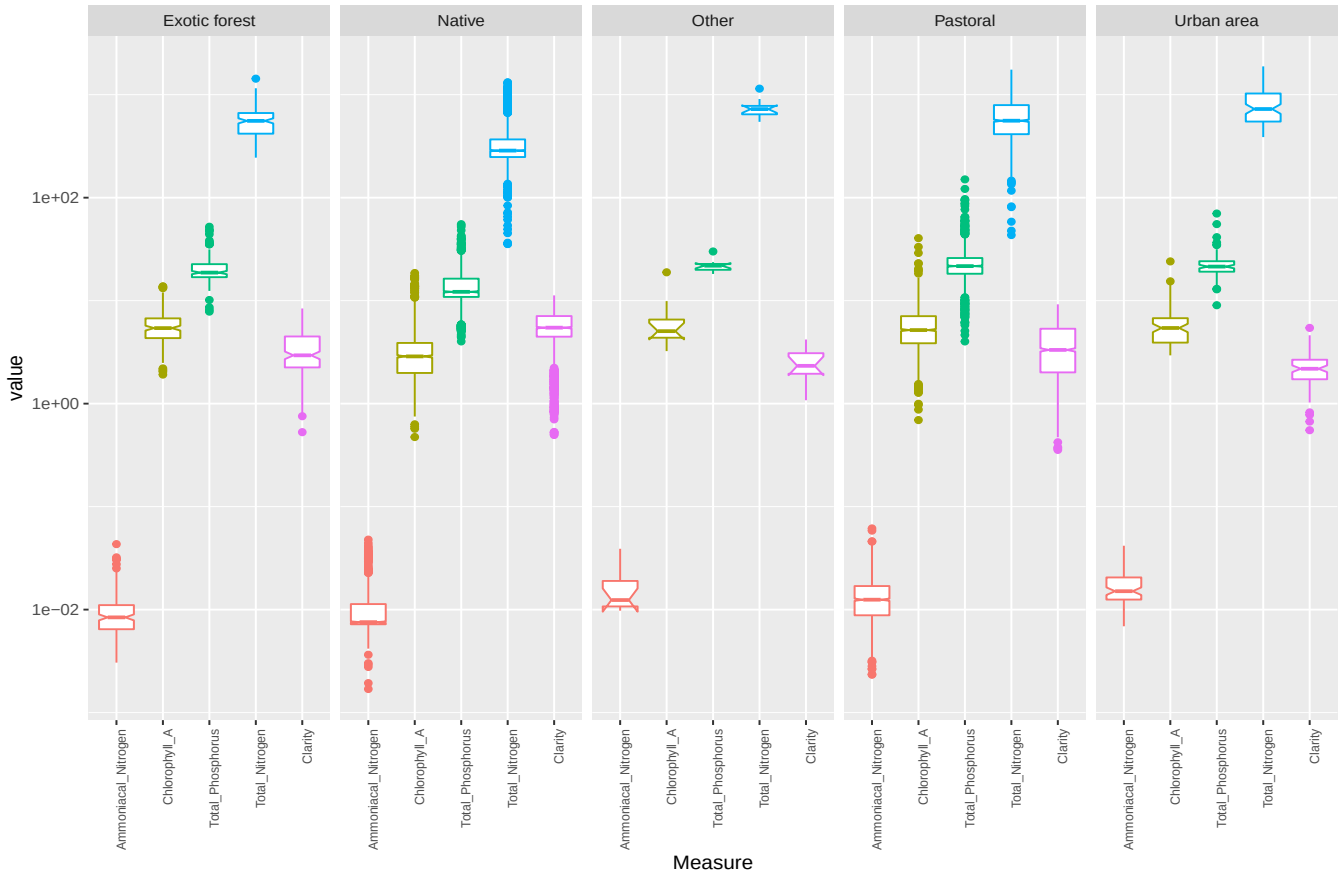

Figure 2: Box Plots of Ammoniacal Nitrogen, Chlorophyll-A, Total Phosphorus, Total Nitrogen and Clarity by Landcover

Interesting Observations

From our exploratory data analysis, we found each of the lake health variables differed between landcovers, especially between the Native and Urban landcovers. The data suggested lakes with predominantly Native landcover had lower levels of Ammoniacal Nitrogen, Chlorophyll-A, Total Phosphorus and Total Nitrogen, and were clearer than lakes in Urban areas. We also found that there was a relationship between the regions and the lake health variables. Specifically, lakes in the Waikato region tended to be of poorer health.

Future Steps

Our future steps include:
- Investigate the relationships between the Lake Health Variables and Lake Dimension variable.
- Multivariate testing for differences in lake health and dimension variables for each region and type of landcover. - Examining the relationship between depth and clarity relationship may impact some of our other findings. - We want to take our research outside of our data now. We have some indicators for example of Regions with poor lake health. We wopuld like to investigate what it is about individual regions that give them poorer lake health and what it is about healthy lakes that we can learn. This is so that we can begin to apply our analysis and make real world changes.