William Norfolk MADA Course Project Water Quality

Content by William Norfolk and Template Structure by Andreas Handel

2019-09-08

# Summary/Abstract

*Write a summary of your project.*

# Introduction (required for part 1)

## General Background Information

*Provide enough background on your topic that others can understand the why and how of your analysis*

## Description of data and data source

These data are water quality measurements collected in Key Largo, Florida by the Marine Resources Development Foundation from 2010 to 2019. The Marine Resources Development Foundation is an environmental education non-profit that provides an immersive experience into the field of marine science for students ranging from fourth grade to undergraduates. Marinelab students take a variety of courses to educate them about the local ecosystems, and complement their laboratory and classroom time with daily field trips to the ecosystem of interest. Many courses within the Marinelab curriculum contain integrative data collection programs which task students with the collection of citizen science data on the health of local ecosystems. All data collected in the field on paper data sheets and is entered into a master raw database by a Marinelab staff member. Specific types of citizen science data collected through the programs are passed onto other agencies for further processing and based on individual need and interest.

These data are raw water quality data collected from various sampling sites frequented by Marinelab vessels. Water quality data is characterized by 9 distinct variables: location, instructor name, group name, pH, ammonia, dissolved oxygen, water temperature, salinity, and equipment. The Marine Resources Development Foundation has a desire to learn the large-scale patterns of the local water quality to better educate students enrolled in the program. Though this data has been collected for some time, no formal analysis of the data has ever been conducted on a large-scale with the master data.

## Questions/Hypotheses to be addressed

*State the research questions you plan to answer with this analysis*

# Methods and Results

*In most research papers, results and methods are separate. You can combine them here if you find it easier. You are also welcome to structure things such that those are separate sections.*

## Data aquisition

*As applicable, explain where and how you got the data. If you directly import the data from an online source, you can combine this section with the next.*

## Data import and cleaning

*Write code that reads in the file and cleans it so it’s ready for analysis. Since this will be fairly long code for most datasets, it might be a good idea to have it in one or several R scripts. If that is the case, explain here briefly what each file does. The files themselves should be commented well so everyone can follow along.*

## Univariate analysis

*Use a combination of text/tables/figures to explore and describe your data. You should produce plots or tables or other summary quantities for most of your variables. You definitely need to do it for the important variables, i.e. if you have main exposure or outcome variables, those need to be explored. Depending on the total number of variables in your dataset, explore all or some of the others.*

## Bivariate analysis

*Create plots or tables and compute simple statistics (e.g. t-tests, simple regression model with 1 predictor, etc.) to look for associations between your outcome(s) and each individual predictor variable*

## Full analysis

*Use one or several suitable statistical/machine learning methods to analyze your data and to produce meaningful figures, tables, etc. This might again be code that is best placed in one or several separate R scripts that need to be well documented. You can then load the results produced by this code*

# Discussion

## Summary and Interpretation

*Summarize what you did, what you found and what it means.*

## Strengths and Limitations

*Discuss what you perceive as strengths and limitations of your analysis.*

## Conclusions

*What are the main take-home messages?*

*Include citations in your Rmd file using bibtex, the list of references will automatically be placed at the end*

# References