**Helen Toledo**

**BIO670**

**R Package Choice and Project Update**

**What is due** (10 points each):

**The Data:** If you are in the process of collecting data, provide a brief 1-2 sentences explaining the progress in collecting the full dataset.

I will use the Lake Erie Fish Community Data from 2013-2019.

Datasets Used (please see attached CSV files):

The WB\_Catch data table contains data on fish captured via bottom trawl (i.e. species, life stage, and catch) and effort (i.e. date, time, location, fishing depth, and trawl metrics)

The WB\_WaterQuality dataset contains depth specific water quality data (i.e. temperature, conductivity, pH, turbidity, chlorophyll, and dissolved oxygen) from trawl locations.

The data used for this project will be from the USGS Science-Base Catalog:

<https://www.sciencebase.gov/catalog/item/5dee70cee4b02caea0f2fd1d>

Attribute Identifiers (metadata link):

<https://www.sciencebase.gov/catalog/file/get/5dee70cee4b02caea0f2fd1d?f=__disk__47%2Ff3%2F1f%2F47f31f31a763e61f81e793a5404133ae17528f88&transform=1&allowOpen=true#Entity%20and%20Attribute%20Information>

**The Proposed Analyses:** Based on the lectures and practical exercises in the course so far, what analyses and plotting techniques do you anticipate using?

I will create a linear regression model for the analysis to predict the fish population from the average daily turbidity in the lake.

I will plot the data by creating a scatterplot with a linear regression trend line.

R Package Description – 10 Points

As part of the work in this class we will be utilizing R packages that allow increased functionality when analyzing data. While we will download multiple packages for analysis in labs, you are tasked with finding, learning about, and utilizing an R package for a presentation.    
In this presentation you will describe what the package is, and a few of its main features and functions. In doing so you should make it clear to the class what the package might be used for.

**What is due (10 points)**: Your choice of the R package you have chosen to explain

I will explain the XTS package because it will allow me to easily manipulate and subset data from specific dates and times. The datasets I have chosen have values for each day for multiple years. The XTS package is specifically for time series data.