ESS 330 Final Project

# 1. Title:

# The Effects of Sea Level Rise on Freshwater Quality Within the State of Florida from 2010 - 2024

# 2. Introduction:

# Sea Level rise has had significant impacts on water quality. As the sea level increases, saltwater intrusion can contaminate freshwater aquifers which reduce the availability of clean drinking water (Werner & Simmons, 2009). The intrusion of saltwater also degrades wetlands, which act as a natural water filter and are habitats for diverse aquatic species. The decrease in wetlands diminishes their ability to buffer against storm surges and filter pollutants, further compromising the water quality (Dessu et al., 2018). Also, rising sea levels lead to more frequent and severe coastal flooding which can result in an overflow of sewage systems and carry pollutants into waterways. This increase in contamination can endanger public health and harm aquatic ecosystems. Understanding and addressing the effects of sea level rise on water quality is significant for protecting freshwater resources, preserving ecosystems and ensuring the sustainability of coastal communities. Our hypothesis to address this research project is: If sea level rise has increased from 2010 to 2024 in the state of Florida, then the salinity levels in freshwater sources will also have increased, degrading overall freshwater quality. This will require two data sets: (1) sea level data from NOAA’s Tides and Current database and (2) freshwater quality (specifically, salinity) data from the USGS National Water Information System. From this data, we will construct a spearman correlation test to determine correlation, a linear regression model to determine the strength of the relationship, and a scatter plot to visualize the relationship.

# 3. Data:

# The data used in this project cover 2 environmental variables that were measured between 2010 and 2024 across Florida. The data we are using for this project come from 2 main sources: the salinity measurements from USGS and the sea level trends from NOAA. The first comes from water quality data available in the USGS website for salinity of water between 2010 and 2024 from three different gauges near Virginia Key, Pensacola, and Panama City Beach. The salinity data were downloaded from the USGS database. The sea level data comes from NOAA, specifically from the Virginia Key, Pensacola, and Panama City Beach stations. The first variable is sea level (in meters), which shows the trend in sea level rise throughout the years. The second is salinity (ppth), which shows fluctuations of salinity of the area in this timeframe. Salinity data are presented in the form of a time series that varies over time, while the data on sea level show an increasing trend. Sites and gauges were selected based on the nearest distance in km. Three gauges from the USGS water quality data set were selected to be paired with the Virginia Key, Pensacola, and Panama City Beach sites from the NOAA sea level rise data set. We will need to clean the sea level rise data by altering the format of the CSV (code given by Prof. Johnson) and merging all three sites’ datasets together. We will need to clean the USGS water quality data set by selecting for only the three specified locations and only the variables we need (such as salinity, date, latitude, longitude, and site ID).

# Sea level rise visual

# The sea level in Virginia Key shows an increasing trend from 2010 through 2024. This trend is highlighted in the following image, which shows a regression line and the monthly fluctuations on sea level average of Virginia Key, Florida.

library(tidyverse)

── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
✔ dplyr 1.1.4 ✔ readr 2.1.5  
✔ forcats 1.0.0 ✔ stringr 1.5.1  
✔ ggplot2 3.5.1 ✔ tibble 3.2.1  
✔ lubridate 1.9.4 ✔ tidyr 1.3.1  
✔ purrr 1.0.2   
── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
✖ dplyr::filter() masks stats::filter()  
✖ dplyr::lag() masks stats::lag()  
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(lubridate)  
  
slr\_data <- read\_csv("8723214\_SLR\_VirginiaKey.csv", skip = 4,   
 col\_names = c("Year", "Month", "Monthly\_MSL", "Linear\_Trend", "High\_Conf", "Low\_Conf"))

Warning: One or more parsing issues, call `problems()` on your data frame for details,  
e.g.:  
 dat <- vroom(...)  
 problems(dat)

Rows: 1064 Columns: 6  
── Column specification ────────────────────────────────────────────────────────  
Delimiter: ","  
chr (6): Year, Month, Monthly\_MSL, Linear\_Trend, High\_Conf, Low\_Conf  
  
ℹ Use `spec()` to retrieve the full column specification for this data.  
ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

slr\_data <- slr\_data %>%  
 filter(!is.na(Year)) %>%  
 mutate(  
 Monthly\_MSL = as.numeric(Monthly\_MSL), # convert to numeric!  
 date = make\_date(Year, Month, 1)  
 ) %>%  
 filter(date >= as.Date("2010-01-01") & date <= as.Date("2024-12-31"))

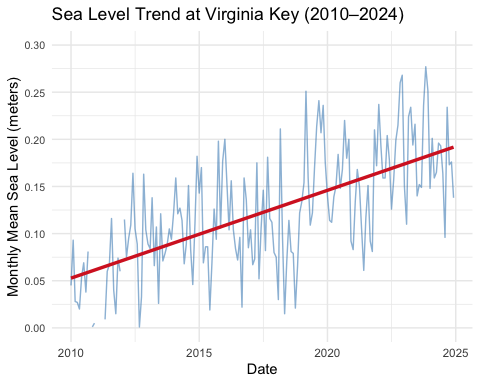
Warning: There were 3 warnings in `mutate()`.  
The first warning was:  
ℹ In argument: `Monthly\_MSL = as.numeric(Monthly\_MSL)`.  
Caused by warning:  
! NAs introduced by coercion  
ℹ Run `dplyr::last\_dplyr\_warnings()` to see the 2 remaining warnings.

ggplot(slr\_data, aes(x = date, y = Monthly\_MSL)) +  
 geom\_line(color = "#1F77B4", alpha = 0.5) +  
 geom\_smooth(method = "lm", color = "#D62728", se = FALSE, size = 1.2) +  
 scale\_y\_continuous(  
 limits = c(0, 0.30),  
 breaks = seq(0, 0.30, by = 0.05),  
 labels = scales::number\_format(accuracy = 0.01)  
 ) +  
 labs(  
 title = "Sea Level Trend at Virginia Key (2010–2024)",  
 x = "Date",  
 y = "Monthly Mean Sea Level (meters)"  
 ) +  
 theme\_minimal() +  
 theme(  
 axis.text.y = element\_text(size = 8)  
 )

Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.  
ℹ Please use `linewidth` instead.

`geom\_smooth()` using formula = 'y ~ x'

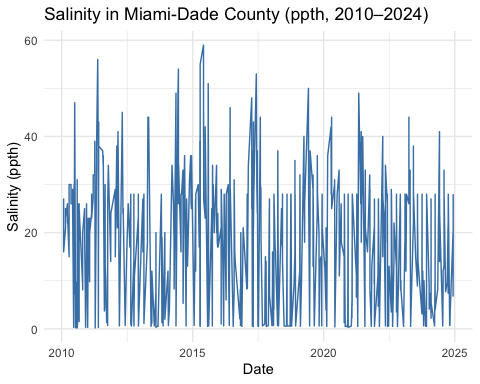
Warning: Removed 6 rows containing non-finite outside the scale range  
(`stat\_smooth()`).



# Salinity levels visual

## Salinity in Miami-Dade county, Florida shows fluctuations throughout the time frame used, with noticeable increases in certain periods of time. This trend can be seen in the following image, which shows how salinity levels have changed between 2010 and 2024.

salinity\_data <- read.csv("salinity copy.csv")  
  
salinity\_data <- salinity\_data %>%  
 mutate(Activity\_StartDate = as.Date(Activity\_StartDate)) %>%  
 filter(Activity\_StartDate >= as.Date("2010-01-01") & Activity\_StartDate <= as.Date("2024-12-31"),  
 Location\_CountyName == "Miami-Dade County")  
  
ggplot(salinity\_data, aes(x = Activity\_StartDate, y = Result\_Measure)) +  
 geom\_line(color = "steelblue", size = 0.5) +  
 labs(title = "Salinity in Miami-Dade County (ppth, 2010–2024)",  
 x = "Date",  
 y = "Salinity (ppth)") +  
 theme\_minimal()



# 4. Methods:

# We will be using a spearman correlation test, linear regression model, and ggplot to analyze our data to answer our research question. I think we do have everything we need for this: sea level rise data (2010-2024 for 18 sites in Florida), water quality (salinity) data (2010-2024 for all gauges in Florida). Latitude and longitudes were included for each point of the water quality data. We do need supplemental latitude and longitude data for the sea level rise points to match them with water quality data, but I was able to collect this from NOAA and record them in an excel sheet. I think the most challenging part of this process will be selecting which sites to use from the two data sets, as we will need to determine which three sites align most closely by location between the sets of data. The way their lat and longitudes were recorded are not identical, and the location data for sea level rise was recorded separately from the actual sea level data which makes this more difficult. I think some potential challenges in the analysis portion of the project that may arise are avoiding other confounding variables as hydrological conditions vary greatly across sites. Also, since this is an observational study, our analysis will only be able to determine correlation instead of causation, which limits the meaning of our results. Our selected tests, however, will be useful in determining our final results. The spearman’s correlation test will first determine if there is a correlation between the two variables (sea level rise and salinity). Once we’ve determined a correlation, the linear regression model will allow us to test the strength of this correlation to determine significance. Finally, employing a scatter plot will be useful for easier visualization of this relationship to confirm our results.

Dessu, S. B., Price, R. M., Troxler, T. G., & Kominoski, J. S. (2018). Effects of sea-level rise and freshwater management on long-term water levels and water quality in the florida coastal everglades. *Journal of Environmental Management*, *211*, 164–176.

Werner, A. D., & Simmons, C. T. (2009). Impact of sea-level rise on sea water intrusion in coastal aquifers. *Groundwater*, *47*(2), 197–204.