

Exercise: Markdown

Environmental Data Analytics | John Fay and Luana Lima

Set up the coding environment

```
#Load packages
library(tidyverse);library(here);library(lubridate)

## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.3      v readr      2.1.4
## v forcats    1.0.0      v stringr   1.5.0
## v ggplot2    3.4.3      v tibble    3.2.1
## v lubridate  1.9.2      v tidyr     1.3.0
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
## here() starts at /home/guest/R/EDA_Spring2024

#Load data
nutrient_data_raw <- read.csv(
  here('Data/Processed_KEY', 'NTL-LTER_Lake_Nutrients_PeterPaul_Processed.csv'),
  stringsAsFactors = TRUE)

#Fix dates
nutrient_data_raw$sampldate = ymd(nutrient_data_raw$sampldate)
```

Wrangle the data

```
#Subset columns and rows
nutrient_data <- nutrient_data_raw %>%
  select(-c(lakeid,depth_id,comments)) %>%
  filter(depth == 0) %>%
  drop_na()

#Compute summary stats for total nitrogen
nutrient_data_tn <- nutrient_data %>%
  group_by(lakename) %>%
  summarize(
    mean_tn_ug = mean(tn_ug),
    min_tn_ug = min(tn_ug),
    max_tn_ug = max(tn_ug),
```

```
sd_tn_ug = sd(tn_ug)  
)
```

Report the summary