Assignment 2: Physical Properties of Lakes

Tristen Townsend

OVERVIEW

This exercise accompanies the lessons in Hydrologic Data Analysis on the physical properties of lakes.

Directions

1. Change "Student Name" on line 3 (above) with your name.
2. Work through the steps, **creating code and output** that fulfill each instruction.
3. Be sure to **answer the questions** in this assignment document.
4. When you have completed the assignment, **Knit** the text and code into a single PDF file.
5. After Knitting, submit the completed exercise (PDF file) to the dropbox in Sakai. Add your last name into the file name (e.g., "Salk_A02_LakePhysical.Rmd") prior to submission.

The completed exercise is due on 11 September 2019 at 9:00 am.

Setup

1. Verify your working directory is set to the R project file,
2. Load the tidyverse, lubridate, and cowplot packages
3. Import the NTL-LTER physical lake dataset and set the date column to the date format
4. Set your ggplot theme (can be theme_classic or something else)

```
getwd()
```

```
## [1] "/Users/Tristen/OneDrive - Duke University/Fall 2019/Hydrologic Data Analysis/Hydrologic_Data_Ana
```

```
library(tidyverse)
```

```
## -- Attaching packages ---------------------------- tidyverse 1.2.1 --

## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   0.8.3     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0

## -- Conflicts ------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':
##
##     date
```

```
library(cowplot)
```

```
##
## ********************************************************
```

```
## Note: As of version 1.0.0, cowplot does not change the
##   default ggplot2 theme anymore. To recover the previous
##   behavior, execute:
##   theme_set(theme_cowplot())
## ********************************************************
##
## Attaching package: 'cowplot'
## The following object is masked from 'package:lubridate':
##
##     stamp
```

```
library(RColorBrewer)

NTLdata <- read.csv("./Data/Raw/NTL-LTER_Lake_ChemistryPhysics_Raw.csv")
NTLdata$sampledate <- as.Date(NTLdata$sampledate, "%m/%d/%y")

theme_set(theme_classic())
```

Creating and analyzing lake temperature profiles

Single lake, multiple dates

5. Choose either Peter or Tuesday Lake. Create a new data frame that wrangles the full data frame so that it only includes that lake during two different years (one year from the early part of the dataset and one year from the late part of the dataset).

```
#Choose Tuesday
Tuesdaydata <- filter(NTLdata, lakename == "Tuesday Lake")

#Filter for years 1985 and 2015
Tuesdaydata.1985.2015 <- filter(Tuesdaydata, year4 %in% c("1985", "2015"))
```

6. Create three graphs: (1) temperature profiles for the early year, (2) temperature profiles for the late year, and (3) a `plot_grid` of the two graphs together. Choose `geom_point` and color your points by date.

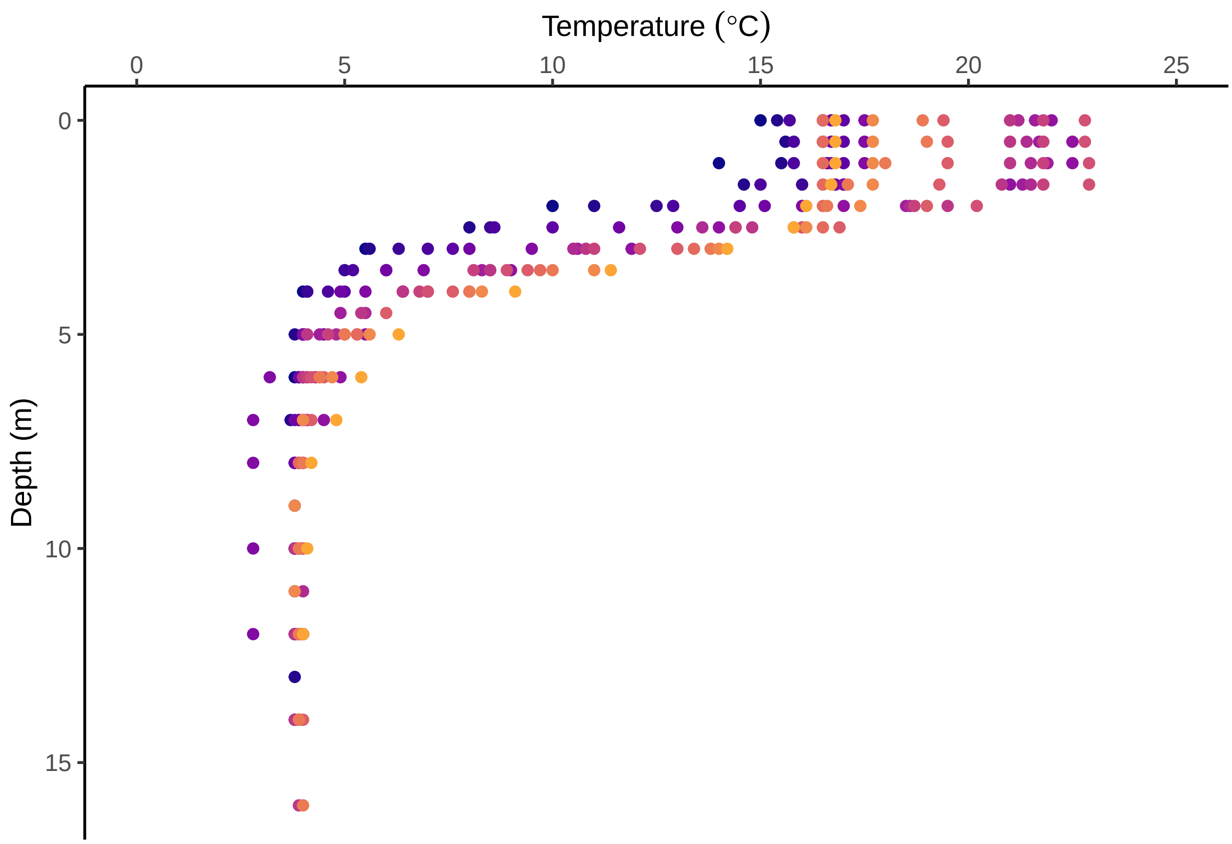
Remember to edit your graphs so they follow good data visualization practices.

```
#Temperature profile for 1985
TuesdayTemp1985 <-
  ggplot(filter(Tuesdaydata.1985.2015, year4 == 1985), aes(x = temperature_C, y = depth, color = daynum
  geom_point() +
  scale_y_continuous(breaks=seq(0, 17, 2.5)) +
  expand_limits(y=c(0,16)) +
  scale_y_reverse() +
  scale_x_continuous(position = "top") +
  expand_limits(x=c(0,25)) +
  scale_color_viridis_c(end = 0.8, option = "plasma") +  #end at 0.8 gets rid of yellow
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)") +
  theme(legend.position = "none")
```

```
## Scale for 'y' is already present. Adding another scale for 'y', which
## will replace the existing scale.
```

```
print(TuesdayTemp1985)
```

```
## Warning: Removed 36 rows containing missing values (geom_point).
```

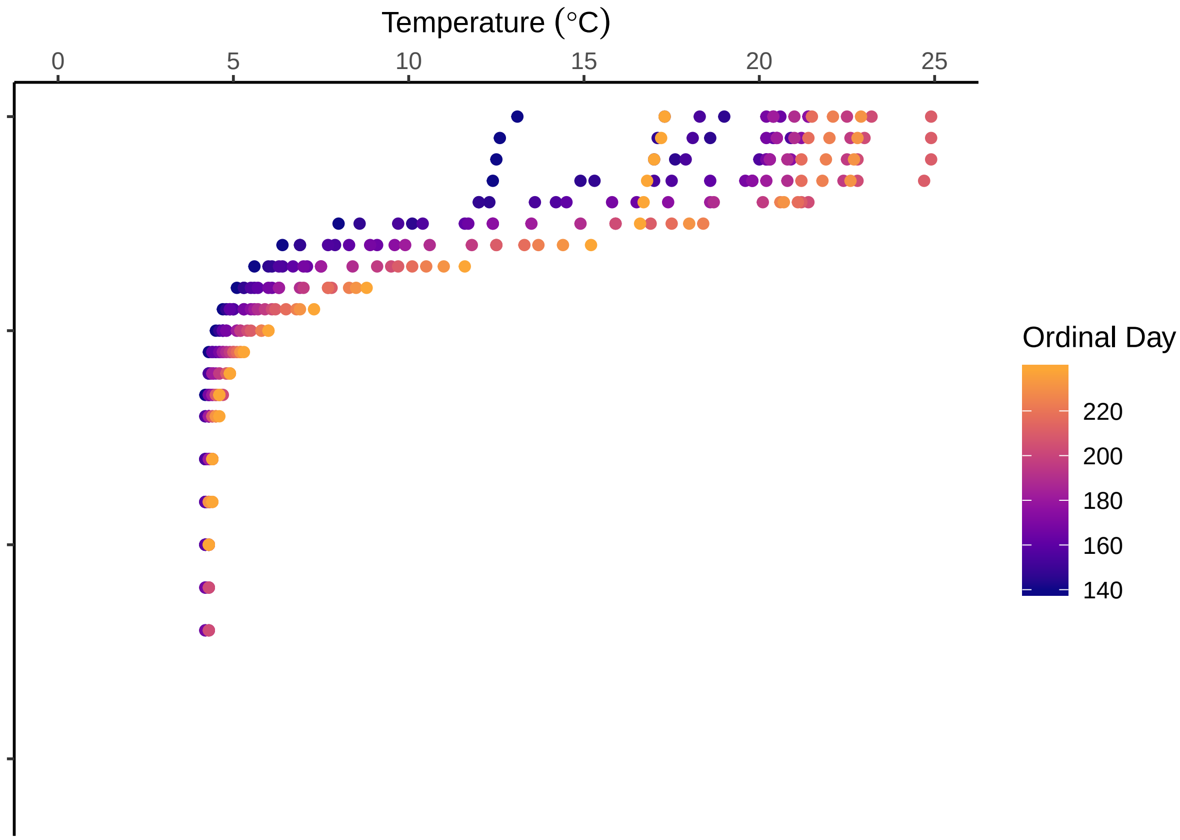



```
#Temperature profile for 2015
TuesdayTemp2015 <-
  ggplot(filter(Tuesdaydata.1985.2015, year4 == 2015), aes(x = temperature_C, y = depth, color = daynum
  geom_point() +
  scale_y_continuous(breaks=seq(0, 17, 2.5)) +
  scale_y_reverse() +
  expand_limits(y=c(0,16)) +
  scale_x_continuous(position = "top") +
  expand_limits(x=c(0,25)) +
  scale_color_viridis_c(end = 0.8, option = "plasma", name = "Ordinal Day") +  #end at 0.8 gets rid of
  labs(x = expression("Temperature "(degree*C)), y = "") +
  theme(legend.position = "right", axis.text.y=element_blank())
```

```
## Scale for 'y' is already present. Adding another scale for 'y', which
## will replace the existing scale.
```

```
print(TuesdayTemp2015)
```

```
## Warning: Removed 30 rows containing missing values (geom_point).
```

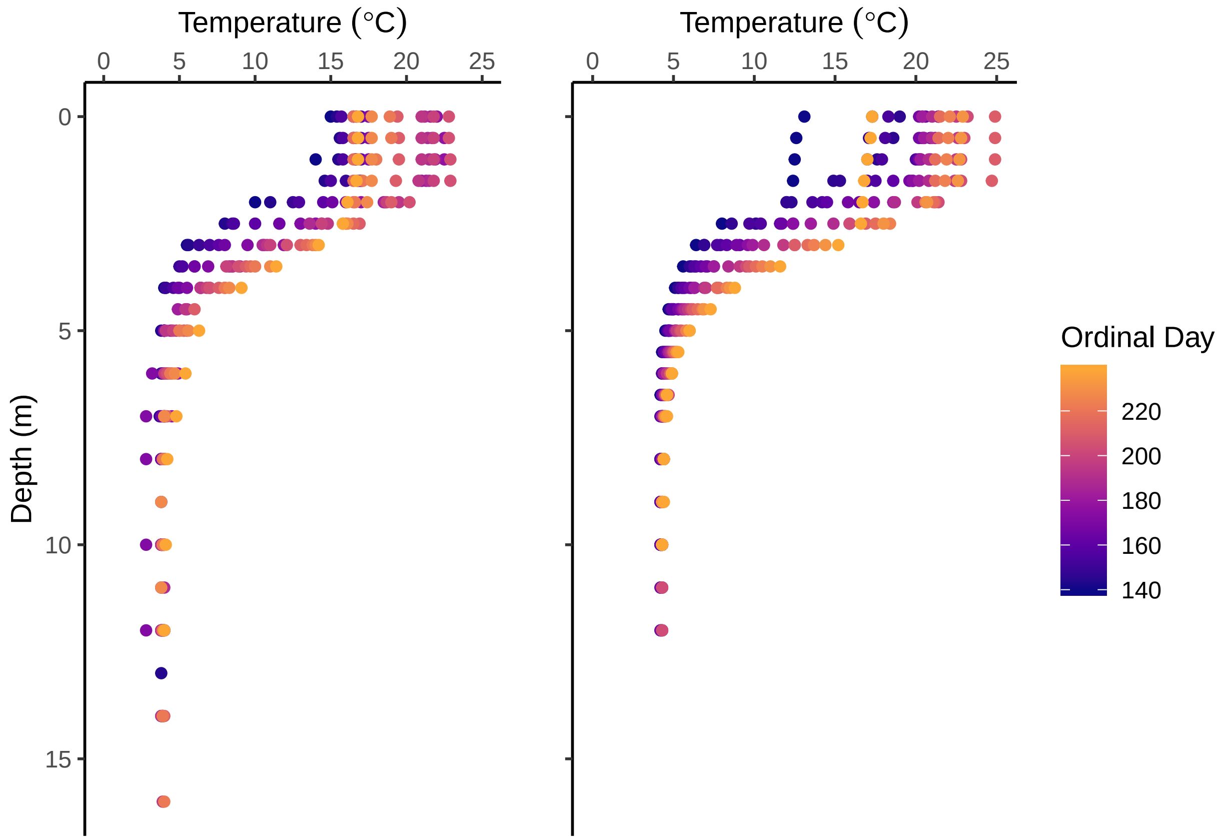



```r
Tuesday_cow <-
  plot_grid(TuesdayTemp1985, TuesdayTemp2015,
            ncol = 2, rel_widths = c(1, 1.4))
```

```
## Warning: Removed 36 rows containing missing values (geom_point).
```

```
## Warning: Removed 30 rows containing missing values (geom_point).
```

```r
print(Tuesday_cow)
```



7. Interpret the stratification patterns in your graphs in light of seasonal trends. In addition, do you see differences between the two years?

   It appears that surface temperatures have increased.

Multiple lakes, single date

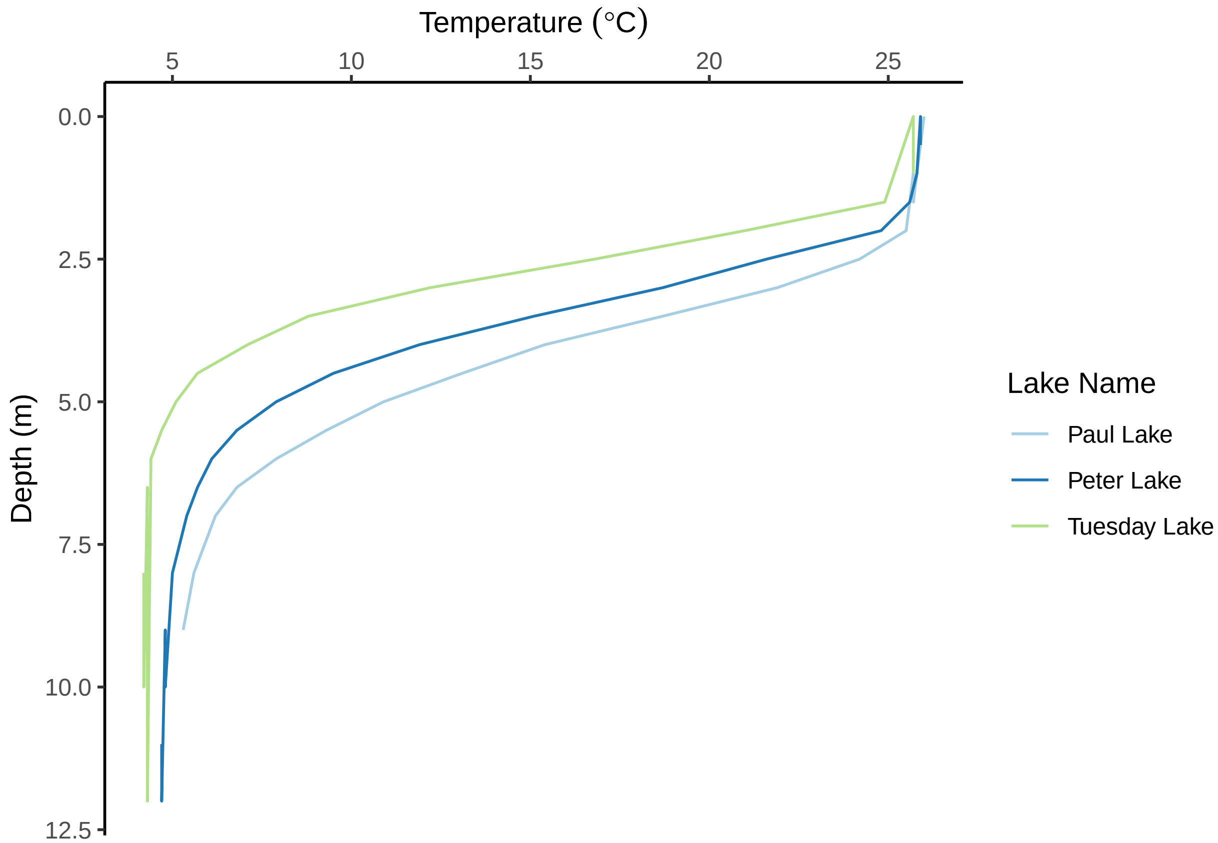
8. On July 25, 26, and 27 in 2016, all three lakes (Peter, Paul, and Tuesday) were sampled. Wrangle your data frame to include just these three dates.

```
#Wrangling for just July 25, 26, 27
Julydata <- NTLdata %>% filter(sampledate == "2016-07-25" | sampledate == "2016-07-26" | sampledate ==
```

9. Plot a profile line graph of temperature by depth, one line per lake. Each lake can be designated by a separate color.

```
Julyplot <- ggplot(Julydata, aes(x=temperature_C, y=depth, color=lakename)) +
  geom_line() +
  scale_y_reverse() +
  scale_x_continuous(position = "top") +
  scale_color_brewer(palette="Paired", name="Lake Name") +
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)")
print(Julyplot)
```

```
## Warning: Removed 6 rows containing missing values (geom_path).
```



10. What is the depth range of the epilimnion in each lake? The thermocline? The hypolimnion?

    The depth ranges of the epilimnion, thermocline, and hypolimnion are the surface to 1.3-2.5m, 2.5-6.25m, and 6.25 to the bottom, respectively.

Trends in surface temperatures over time.

11. Run the same analyses we ran in class to determine if surface lake temperatures for a given month have increased over time ("Long-term change in temperature" section of day 4 lesson in its entirety), this time for either Peter or Tuesday Lake.

```
#Add a month column to Tuesday dataframe
Tuesdaydata$Month <- month(Tuesdaydata$sampledate)

#Filter for surface depth
Tuesdaydata.surf <- Tuesdaydata %>%
  filter(depth == 0.00)

#Create separate data frames for each month
Tuesday.may <- Tuesdaydata.surf %>%
  filter(Month == 5)

Tuesday.june <- Tuesdaydata.surf %>%
  filter(Month == 6)

Tuesday.july <- Tuesdaydata.surf %>%
```

```r
  filter(Month == 7)

Tuesday.aug <- Tuesdaydata.surf %>%
  filter(Month == 8)

#Run linear regression
Tuesday.may.lm <- lm(temperature_C ~ year4, Tuesday.may)
Tuesday.june.lm <- lm(temperature_C ~ year4, Tuesday.june)
Tuesday.july.lm <- lm(temperature_C ~ year4, Tuesday.july)
Tuesday.aug.lm <- lm(temperature_C ~ year4, Tuesday.aug)

summary(Tuesday.may.lm) # p = 0.5558 DO NOT REJECT NULL
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesday.may)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6223 -1.4411  0.0314  1.5604  5.2216
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -27.15303   73.73032  -0.368    0.715
## year4         0.02196    0.03689   0.595    0.556
##
## Residual standard error: 2.522 on 32 degrees of freedom
## Multiple R-squared:  0.01095,    Adjusted R-squared:  -0.01995
## F-statistic: 0.3544 on 1 and 32 DF,  p-value: 0.5558
```

```r
summary(Tuesday.june.lm) # p = 0.9921 DO NOT REJECT NULL
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesday.june)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.0339 -1.5343 -0.0279  1.9180  6.7676
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 21.1373460 50.7897026   0.416    0.678
## year4       -0.0002531  0.0254253  -0.010    0.992
##
## Residual standard error: 2.621 on 80 degrees of freedom
## Multiple R-squared:  1.239e-06,  Adjusted R-squared:  -0.0125
## F-statistic: 9.912e-05 on 1 and 80 DF,  p-value: 0.9921
```

```r
summary(Tuesday.july.lm) # p = 0.05691 DO NOT REJECT NULL
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesday.july)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0561 -1.3275 -0.2047  1.4031  4.2161
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -49.18776   37.36614  -1.316   0.1916
## year4         0.03612    0.01871   1.931   0.0569 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.953 on 84 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.04248,    Adjusted R-squared:  0.03109
## F-statistic: 3.727 on 1 and 84 DF,  p-value: 0.05691
```

```r
summary(Tuesday.aug.lm) # p = 0.1547 DO NOT REJECT NULL
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesday.aug)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.9656 -1.1055 -0.0787  1.2820  3.8677
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -37.70343   41.36954  -0.911    0.365
## year4         0.02976    0.02072   1.436    0.155
##
## Residual standard error: 2.025 on 81 degrees of freedom
## Multiple R-squared:  0.02484,    Adjusted R-squared:  0.0128
## F-statistic: 2.063 on 1 and 81 DF,  p-value: 0.1547
```

12. How do your results compare to those we found in class for Paul Lake? Do similar trends exist for both lakes?

    For Paul Lake, we found that there were significant change in surface temperatures for July and August only for the time span reflected in the data. However, Tuesday Lake does not have any significant changes in surface temperatures.