Assignment 2: Physical Properties of Lakes

Rachel Bash

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] "C:/Users/19524/Documents/DUKE/Hydrologic Data Analytics/Hydrologic_Data_Analysis/Assignments"
```

```
library(tidyverse)
library(lubridate)
library(cowplot)

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

theme_set(theme_classic())
```

Creating and analyzing lake temperature profiles

Single lake, multiple dates

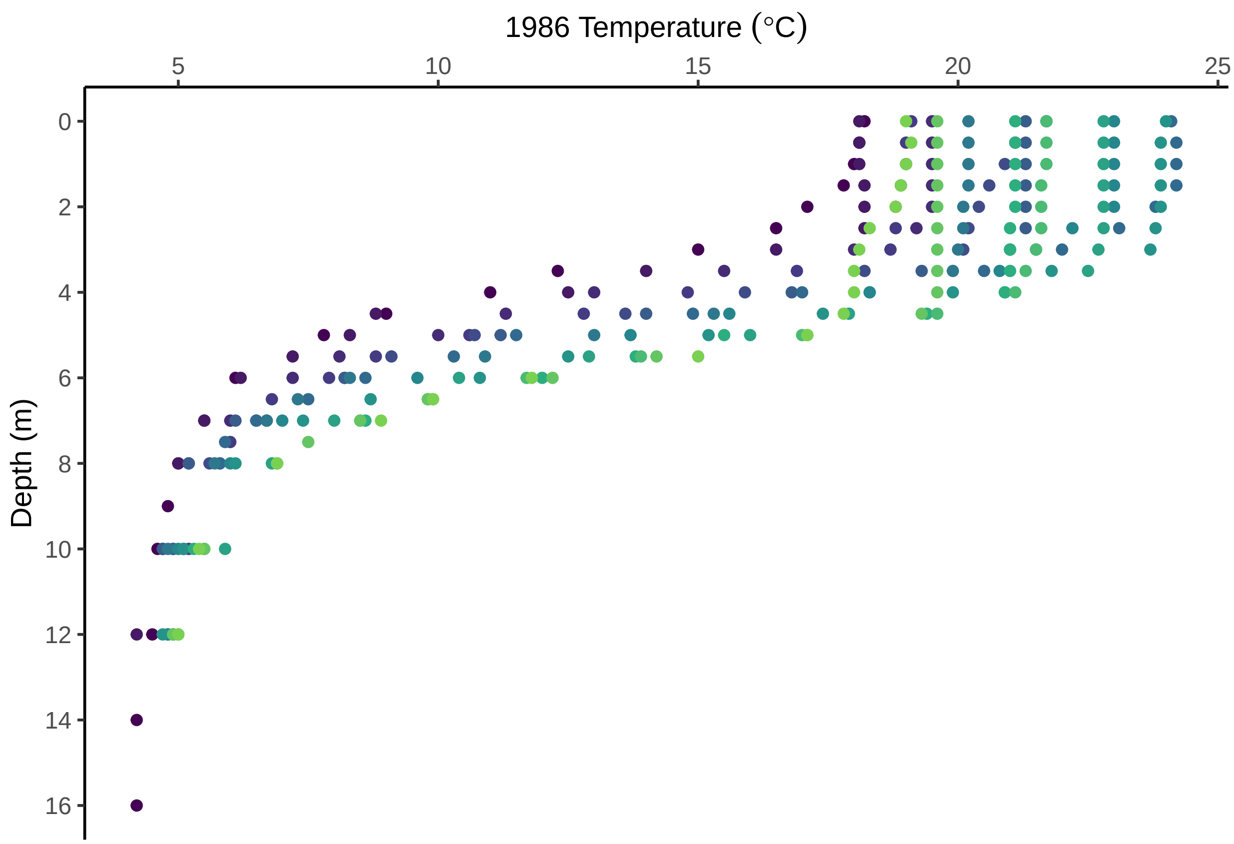
5. Choose either Peter or Tuesday Lake. Create a new data frame that wrangles the full data frame so that it only includes that lake during two different years (one year from the early part of the dataset and one year from the late part of the dataset).

```
Peterdata <- NTLdata %>% filter(lakename == "Peter Lake")
Peterdata <- Peterdata %>% filter(year4 == "1986" | year4 == "2016")
```

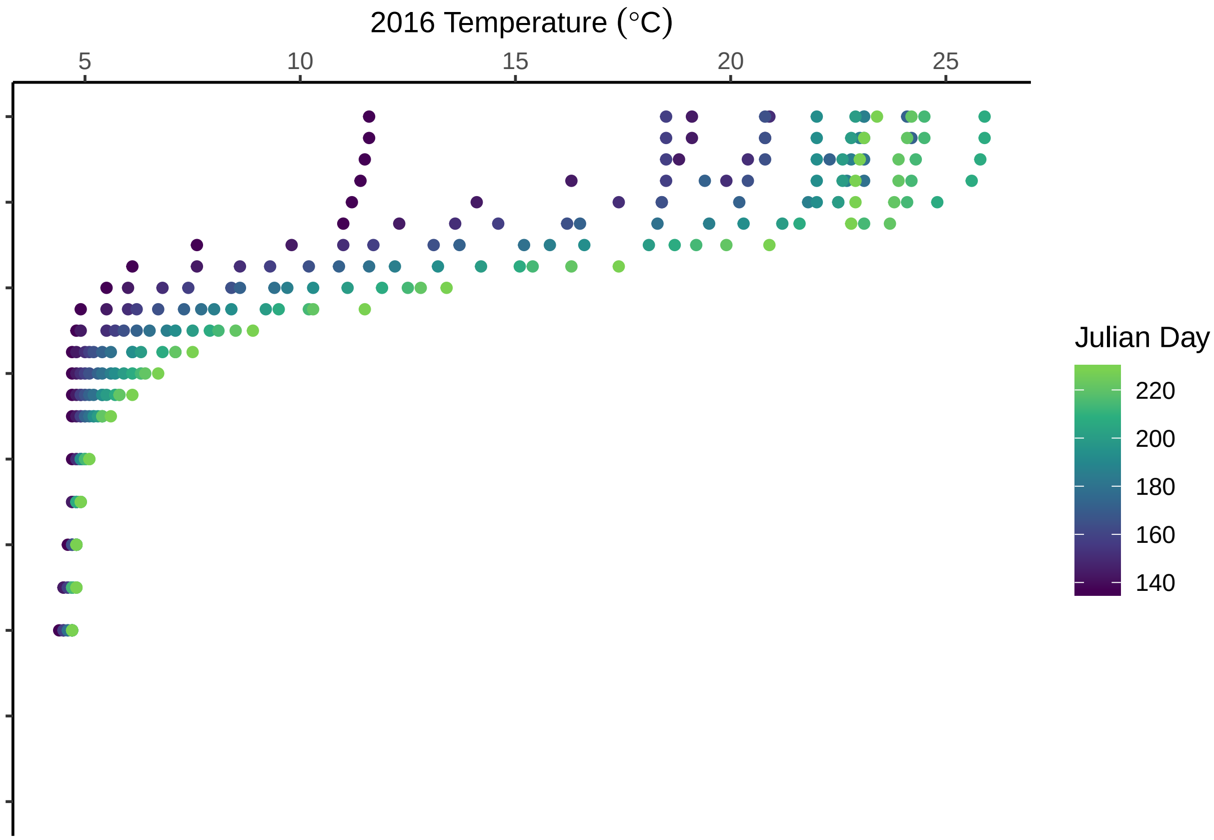
6. Create three graphs: (1) temperature profiles for the early year, (2) temperature profiles for the late year, and (3) a `plot_grid` of the two graphs together. Choose `geom_point` and color your points by date.

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

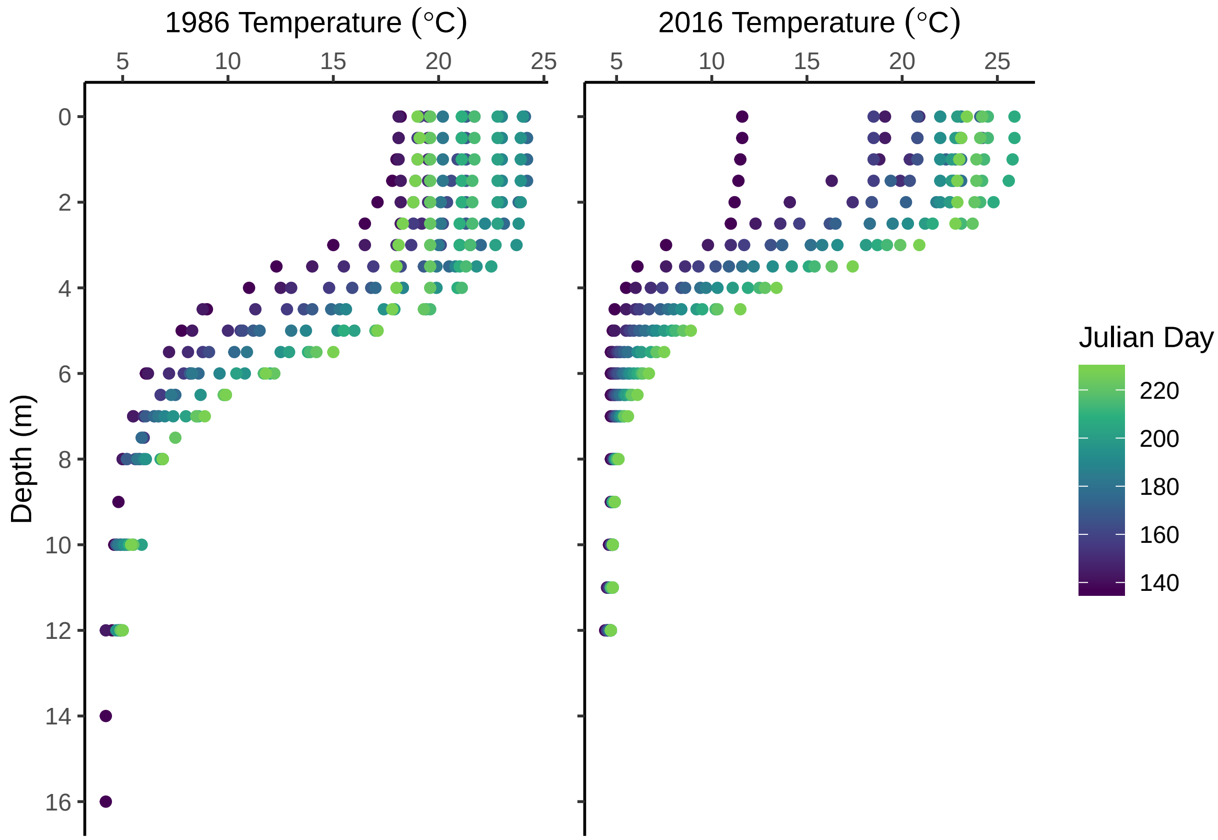
```r
Peterearly <- Peterdata %>% filter(year4=="1986")
Peterlate <- Peterdata %>% filter(year4=="2016")
Peterearly.plot <- ggplot(Peterearly, aes(x = temperature_C, y = depth, color = daynum)) +
  geom_point() +
  scale_y_reverse(breaks=c(0,2,4,6,8,10,12,14,16)) +
  scale_x_continuous(position = "top") +
  scale_color_viridis_c(end = "0.8", option = "viridis") +
  labs(x = expression("1986 Temperature "(degree*C)), y = "Depth (m)") +
  theme(legend.position = "none")

print(Peterearly.plot)
```




```r
Peterlate.plot <- ggplot(Peterlate,aes(x = temperature_C, y = depth, color = daynum)) +
  geom_point() +
  scale_y_reverse(breaks = c(0,2,4,6,8,10,12,14,16)) +
  coord_cartesian(ylim = c(0, 16)) +
  scale_x_continuous(position = "top") +
  scale_color_viridis_c(end = "0.8", option = "viridis") +
  labs(x = expression("2016 Temperature "(degree*C)), y = "Depth (m)",  color = "Julian Day") +
  theme(axis.text.y = element_blank(), axis.title.y = element_blank())
print(Peterlate.plot)
```




```
Peterearlylate.grid <- plot_grid(Peterearly.plot, Peterlate.plot, ncol=2,
                                 rel_widths = c(1,1.2))
print(Peterearlylate.grid)
```



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

   The relative temperature of the Peter Lake, specifically in the epilimnion and the metalimnion, increase as the season progresses, for both years. The lake depth in 1986 looks to be deeper than in 2016, as they took samples as deep as 16 feet, whereas in 2016 the deepest lake sample recorded was 12 feet. In 2016, it appears that the beginning of the year was particularly cold in the epilimnion, hovering around 12 degrees Celcius, until it jumps pretty substantially not too long later.In 1986, it looks like perhaps it was a colder fall, as the temperature in the epilimnion decreases later in the season. It is interesting to note that even in the hypolimnion in 2016, the temperature did increase a little, which perhaps means that the shallower lake received a little bit of light or some mixing behaviors near the end of the summer.

Multiple lakes, single date

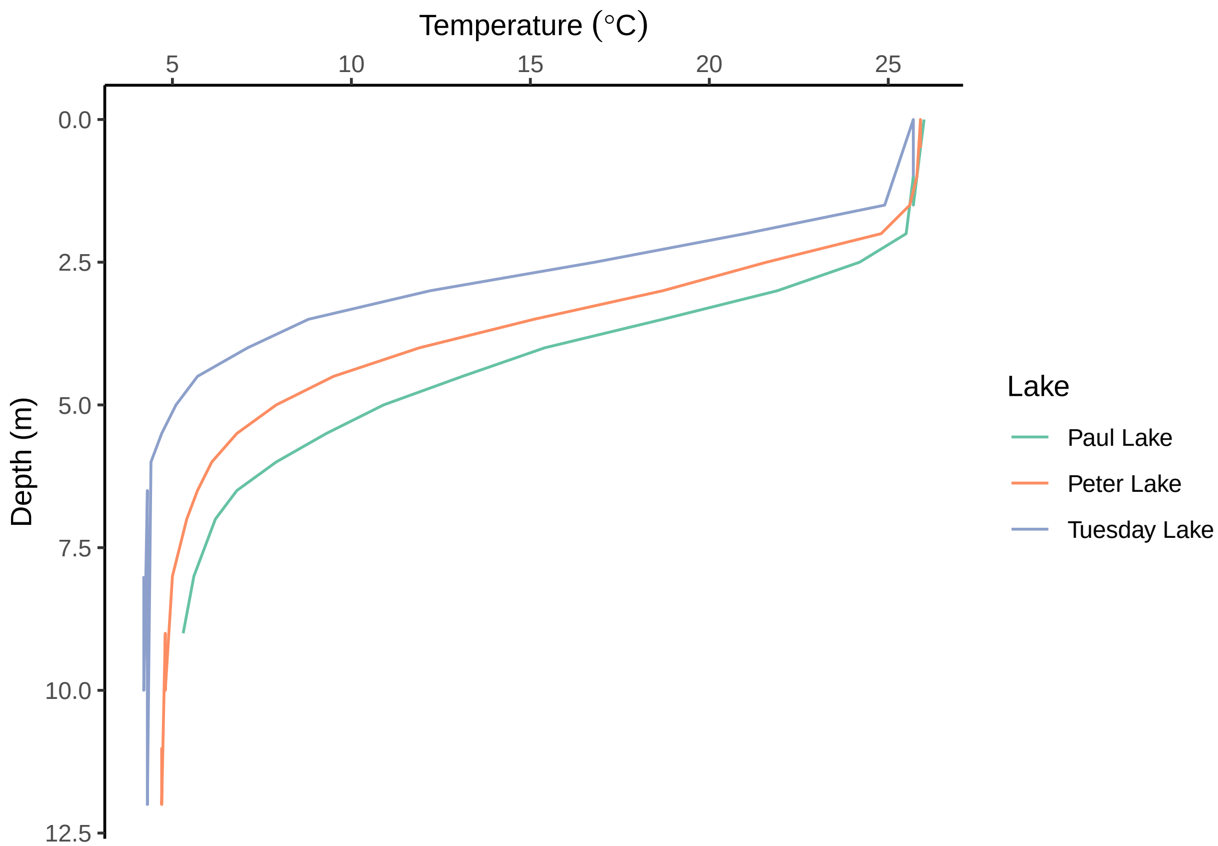
8. On July 25, 26, and 27 in 2016, all three lakes (Peter, Paul, and Tuesday) were sampled. Wrangle your data frame to include just these three dates.

```
Julyzoom <- NTLdata %>% filter(lakename == "Peter Lake" | lakename == "Paul Lake" | lakename == "Tuesda
```

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

```
Julyzoom.plot <- ggplot(Julyzoom, aes(x=temperature_C, y=depth, color=lakename)) +
  geom_line() +
  scale_y_reverse() +
  scale_x_continuous(position = "top") +
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)",  color = "Lake" ) +
```

```
  scale_color_brewer(type = "div", palette = "Set2")
print(Julyzoom.plot)
```



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

    Epilimnion: for both Tuesday and Peter, the epilimnion is from about surface level (0) to about 1.5m, and for Paul, it is about 0 to 2.25m. Thermocline: Peter and Tuesday lakes have about the same thermocline as well, between 1.5m to about 6m. Paul's thermocline ends a bit deeper, going from 2.25 to about 7m. Hypolimnion: Tuesday and Peter - 6m and below to the bottom which reaches to about 12m. Paul Lake has a hypolimnion from 7 to about 9m, the bottom of that lake.

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.

```
Peterdata.all <- NTLdata %>% filter(lakename == "Peter Lake") %>%  mutate(Month = month(sampledate)) %>%

Peterdata.May <- Peterdata.all %>% filter(Month=="5")
Peterdata.Jun <- Peterdata.all %>% filter(Month=="6")
Peterdata.Jul <- Peterdata.all %>% filter(Month=="7")
Peterdata.Aug <- Peterdata.all %>% filter(Month=="8")

Peterdata.May.test <- lm(data=Peterdata.May, temperature_C ~ year4)
summary(Peterdata.May.test)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Peterdata.May)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.240 -2.064  0.998  2.154  3.773
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.028e+00  1.623e+02   0.049    0.961
## year4       6.303e-03  8.121e-02   0.078    0.939
##
## Residual standard error: 3.006 on 15 degrees of freedom
## Multiple R-squared:  0.0004014,  Adjusted R-squared:  -0.06624
## F-statistic: 0.006023 on 1 and 15 DF,  p-value: 0.9392
```

```r
Peterdata.Jun.test <- lm(data=Peterdata.Jun, temperature_C ~ year4)
summary(Peterdata.Jun.test)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Peterdata.Jun)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.2988 -2.1762  0.5796  2.1197  3.1462
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.985e+00  1.246e+02   0.024    0.981
## year4       9.006e-03  6.237e-02   0.144    0.888
##
## Residual standard error: 2.698 on 12 degrees of freedom
## Multiple R-squared:  0.001735,   Adjusted R-squared:  -0.08145
## F-statistic: 0.02085 on 1 and 12 DF,  p-value: 0.8876
```

```r
Peterdata.Jul.test <- lm(data=Peterdata.Jul, temperature_C ~ year4)
summary(Peterdata.Jul.test)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Peterdata.Jul)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0234 -1.4940  0.0323  1.2400  5.1224
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -66.60715  148.89186  -0.447    0.661
## year4         0.04427    0.07445   0.595    0.562
##
## Residual standard error: 2.378 on 14 degrees of freedom
## Multiple R-squared:  0.02463,    Adjusted R-squared:  -0.04504
```

```
## F-statistic: 0.3536 on 1 and 14 DF,  p-value: 0.5616
```

```r
Peterdata.Aug.test <- lm(data=Peterdata.Aug, temperature_C ~ year4)
summary(Peterdata.Aug.test)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Peterdata.Aug)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.0680 -2.2719 -0.4855  2.7149  4.9429
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -10.33735  185.37054  -0.056    0.957
## year4         0.01583    0.09273   0.171    0.868
##
## Residual standard error: 3.573 on 10 degrees of freedom
## Multiple R-squared:  0.002906,   Adjusted R-squared:  -0.0968
## F-statistic: 0.02914 on 1 and 10 DF,  p-value: 0.8678
```

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

    In the analyses above, we find that there is no significant temperature change across years for any of the summer months in Peter Lake. Every linear regression that I ran had p-values greater than 0.05, indicating no significant change over time. These results are different than the ones we found in Paul lake in class. It is possible that these lakes have differences that have caused there to be different results when I run the same test on each lake. This doesn't make either of my analyses wrong, it simply indicates that there may be more things to discover about each of these lakes.