Assignment 2: Physical Properties of Lakes

Yutao Gong

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/gongy/Documents/Hydrologic_Data_Analysis/Assignments"
```

```
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(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```
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())
```

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

```
library(lubridate)
```

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

```
## The following object is masked from 'package:cowplot':
##
##     stamp
```

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

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

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).

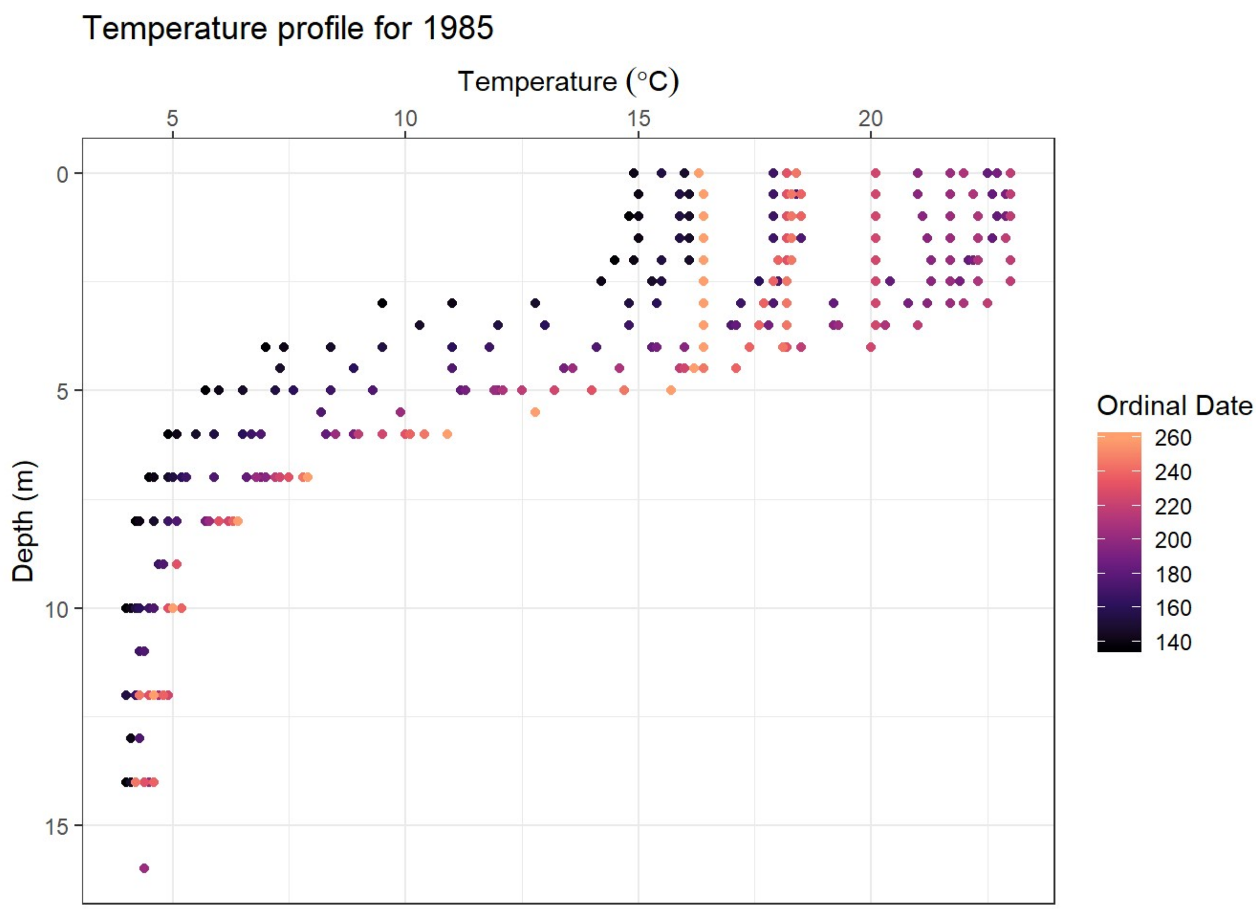
```
Peterdata.og <- filter(NTLdata, lakename == "Peter Lake")
Peterdata <- filter(Peterdata.og, year4 == "1985" | year4 == "2015")
```

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.

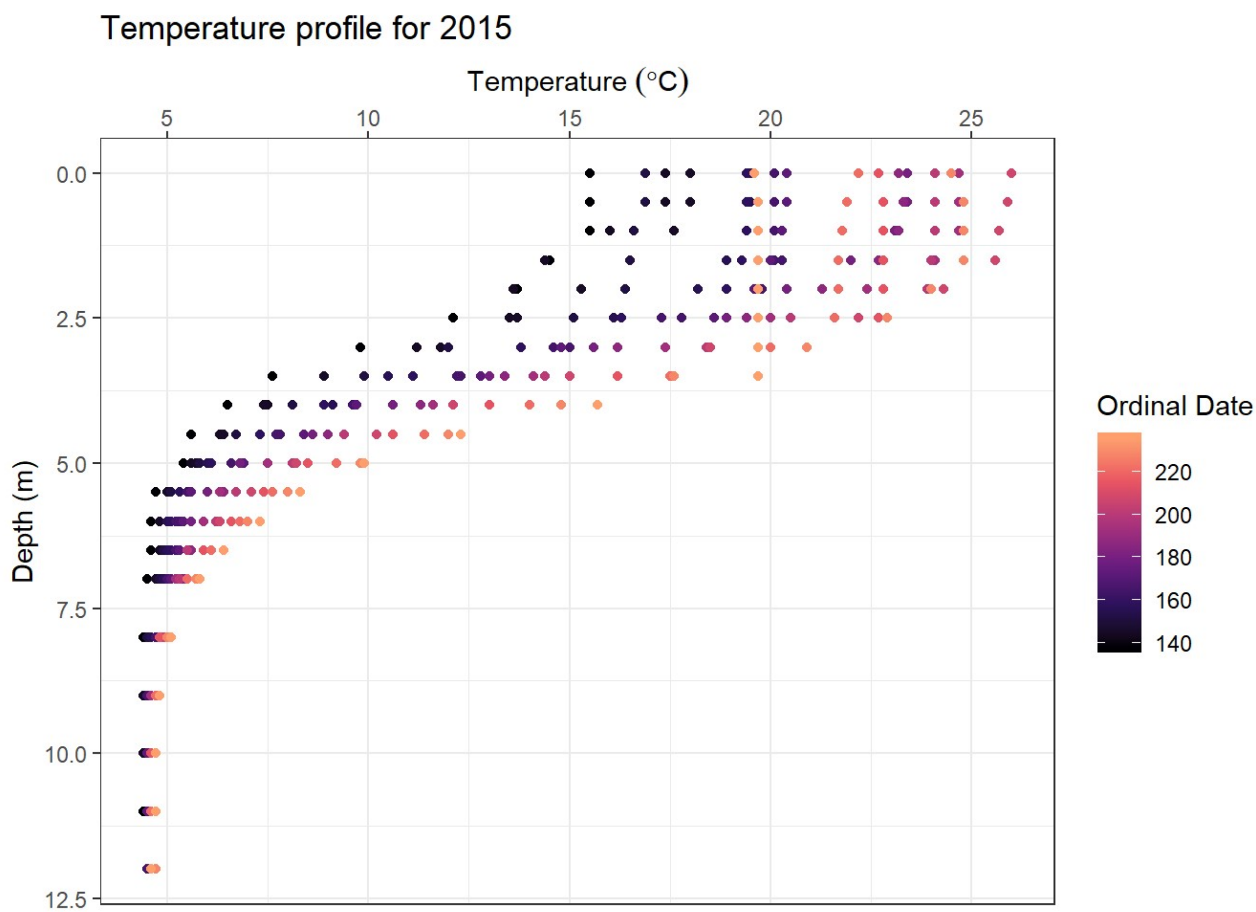
```
Tempprofiles_Peter1985 <-
  ggplot(subset(Peterdata, year4 == "1985"), aes(x = temperature_C, y = depth, color =
daynum)) +
  geom_point() +
  scale_y_reverse() +
  scale_x_continuous(position = "top") +
  scale_color_viridis_c(end = 0.8, option = "magma") +
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)", color = "Ordinal Dat
e") +
  ggtitle("Temperature profile for 1985")
print(Tempprofiles_Peter1985)
```

```
## Warning: Removed 57 rows containing missing values (geom_point).
```




```
Tempprofiles_Peter2015 <-
  ggplot(subset(Peterdata, year4 == "2015"), aes(x = temperature_C, y = depth, color =
daynum)) +
  geom_point() +
  scale_y_reverse() +
  scale_x_continuous(position = "top") +
  scale_color_viridis_c(end = 0.8, option = "magma") +
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)", color = "Ordinal Dat
e") +
  ggtitle("Temperature profile for 2015")
print(Tempprofiles_Peter2015)
```

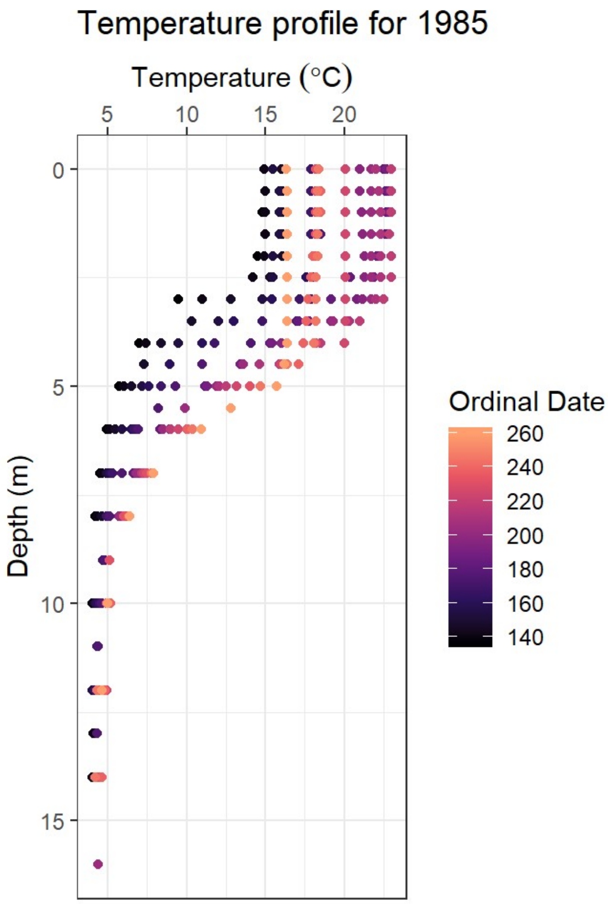
```
## Warning: Removed 30 rows containing missing values (geom_point).
```

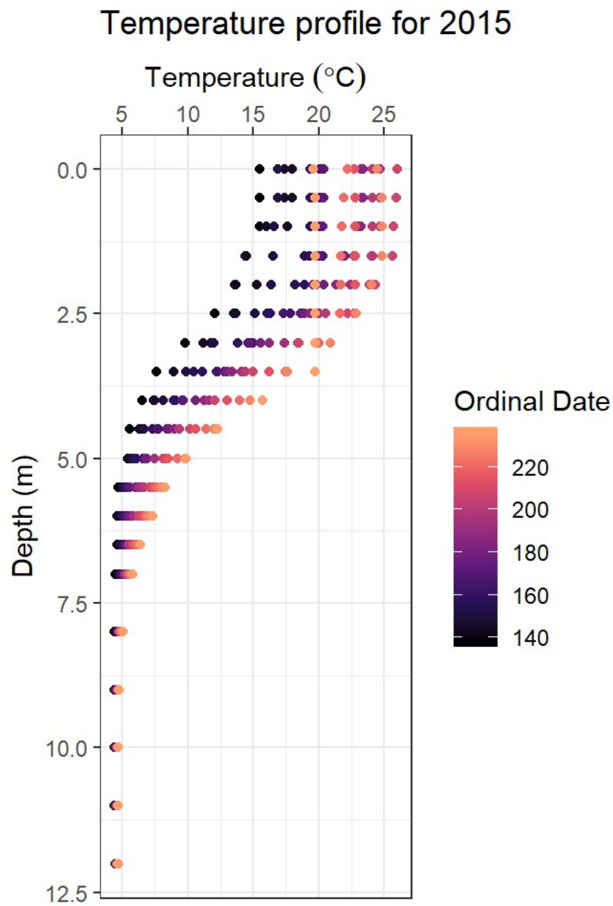



```
plot_grid(Tempprofiles_Peter1985, Tempprofiles_Peter2015)
```

```
## Warning: Removed 57 rows containing missing values (geom_point).
```

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





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

> From the graphs we can see that from May to August the temperature goes up and then down, because during this time period, the season goes from spring to summer and to early fall, and atmoshpheric temperature increases from spring to summer and then decreases from summer to fall. As of stratification patterns, surface/epilimnion temperature seems to change the most along with seasonal change, while hypolimnion temperature remains relatively stable. Combining these two, we can notice that temperature changes more rapidly within thermocline during warmer time. Comparing the two years, we can see that the average temperature is higher in 2015, especially in the warmest months, which is probably due to global warming.

Multiple lakes, single date

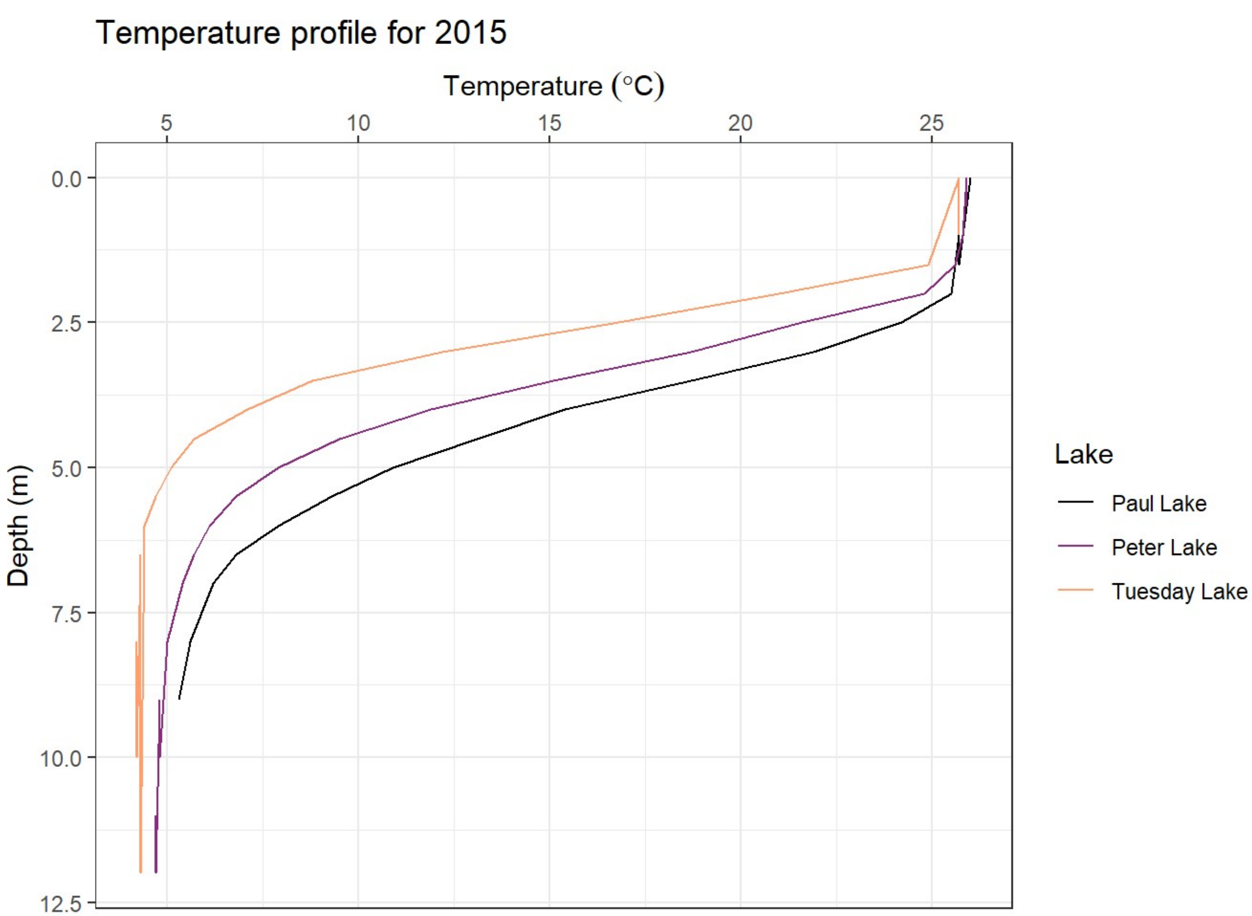
8. On July 25, 26, and 27 in 2016, all three lakes (Peter, Paul, and Tuesday) were sampled. Wrangle your data frame to include just these three dates.

```
samedate_data <- filter(NTLdata, year4 == "2016" & daynum  %in% c("207","208","209"))
```

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

```
Tempprofiles_samedate <-
  ggplot(samedate_data, aes(x = temperature_C, y = depth, color = lakename)) +
  geom_line() +
  scale_y_reverse() +
  scale_x_continuous(position = "top") +
  scale_color_viridis_d(end = 0.8, option = "magma") +
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)", color = "Lake") +
  ggtitle("Temperature profile for 2015")
print(Tempprofiles_samedate)
```

```
## 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?

> For epilimnion, the approximate depth range of Tuesday Lake is 0.0~1.3m, of Peter Lake is 0.0~1.5m, of Paul Lake is 0.0~2m. The approximate depth range of the thermocline for Tuesday Lake is 1.3m~6m, for Peter Lake is 1.5m~8m, for Paul Lake is 2m~8.3m. The approximate depth range of the hypolimnion for Tuesday Lake is below 6m (data ends at around 12m), for Peter Lake is below 8m (data ends at around 12m), while for Paul Lake data ends at around 8.3m so it's hard to estimate the depth range of hypolimnion.

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.

```
Petersurface <- Peterdata.og %>%
  mutate(Month = month(sampledate)) %>%
  filter(depth == 0 & Month < 9 & Month > 4) # Noticed that there is one row of data f
rom February of 2009 but does not have valid temperature data, so still only including
from May to August

Petersurface.May <- filter(Petersurface, Month == 5)
Petersurface.June <- filter(Petersurface, Month == 6)
Petersurface.July <- filter(Petersurface, Month == 7)
Petersurface.August <- filter(Petersurface, Month == 8)

Temptrend.May <- lm(data = Petersurface.May, temperature_C ~ year4)
summary(Temptrend.May)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Petersurface.May)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.2960 -1.4047 -0.0873  1.5605  6.5344
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -35.69596   70.65403  -0.505    0.615
## year4         0.02609    0.03531   0.739    0.463
##
## Residual standard error: 2.895 on 61 degrees of freedom
## Multiple R-squared:  0.008867,   Adjusted R-squared:  -0.007381
## F-statistic: 0.5457 on 1 and 61 DF,  p-value: 0.4629
```

```
Temptrend.June <- lm(data = Petersurface.June, temperature_C ~ year4)
summary(Temptrend.June)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Petersurface.June)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.1170 -1.5844 -0.0472  1.9932  6.9684
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.696621  43.622161   0.131    0.896
## year4        0.007286   0.021813   0.334    0.739
##
## Residual standard error: 2.572 on 142 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.000785,   Adjusted R-squared:  -0.006252
## F-statistic: 0.1116 on 1 and 142 DF,  p-value: 0.7389
```

```
Temptrend.July <- lm(data = Petersurface.July, temperature_C ~ year4)
summary(Temptrend.July)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Petersurface.July)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.4170 -1.0943  0.0191  1.2393  3.7861
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -104.75590   30.74789  -3.407 0.000846 ***
## year4          0.06392    0.01538   4.157 5.44e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.817 on 148 degrees of freedom
## Multiple R-squared:  0.1045, Adjusted R-squared:  0.09849
## F-statistic: 17.28 on 1 and 148 DF,  p-value: 5.444e-05
```

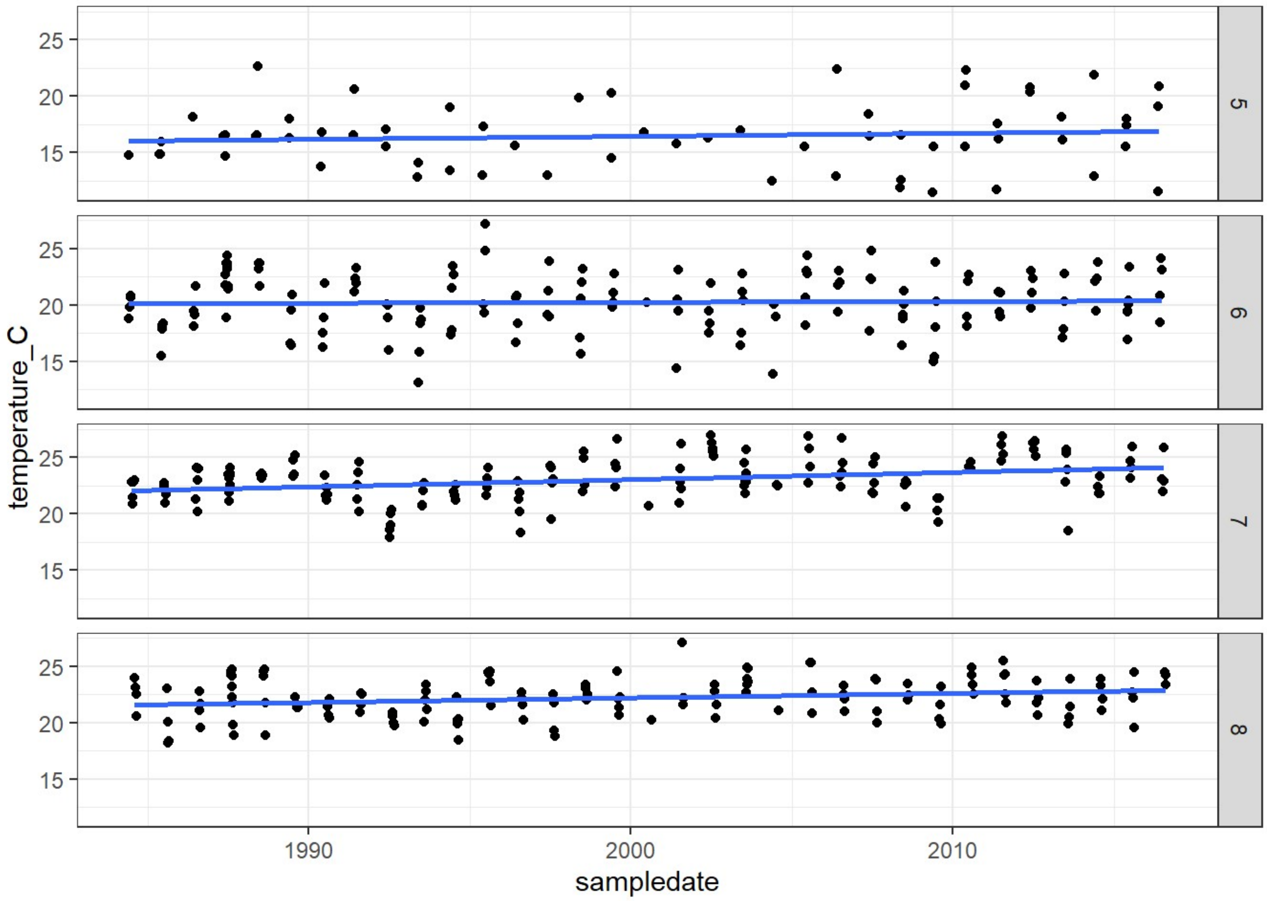
```
Temptrend.August <- lm(data = Petersurface.August, temperature_C ~ year4)
summary(Temptrend.August)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Petersurface.August)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -3.497 -1.218  0.016  1.185  4.819
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -59.00883   29.69264  -1.987  0.04888 *
## year4         0.04062    0.01485   2.735  0.00706 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.683 on 137 degrees of freedom
## Multiple R-squared:  0.05178,    Adjusted R-squared:  0.04486
## F-statistic: 7.481 on 1 and 137 DF,  p-value: 0.007061
```

```
ggplot(Petersurface, aes(x = sampledate, y = temperature_C)) +
  geom_point() +
  facet_grid(rows = vars(Month)) +
  geom_smooth(se = FALSE, method = lm)
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```



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

> The results are pretty similar to those of Paul Lake. The trend is that for both lakes the surface temperature has a slight increasing trend across the years. Among the months, July has the most significant increasing trend.