Assignment 2: Physical Properties of Lakes

Haoyu Zhang

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)

```
setwd("D:/William/Duke/Study/EOS 722/Hydrologic_Data_Analysis/Assignments/Assignment 2")
getwd()
```

```
## [1] "D:/William/Duke/Study/EOS 722/Hydrologic_Data_Analysis/Assignments/Assignment 2"
```

```
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
```

```r
rawdat <- read.csv("../../Data/Raw/NTL-LTER_Lake_ChemistryPhysics_Raw.csv",header = T)
head(rawdat)
```

```
##   lakeid  lakename year4 daynum sampledate depth temperature_C
## 1      L Paul Lake  1984    148    5/27/84  0.00          14.5
## 2      L Paul Lake  1984    148    5/27/84  0.25            NA
## 3      L Paul Lake  1984    148    5/27/84  0.50            NA
## 4      L Paul Lake  1984    148    5/27/84  0.75            NA
## 5      L Paul Lake  1984    148    5/27/84  1.00          14.5
## 6      L Paul Lake  1984    148    5/27/84  1.50            NA
##   dissolvedOxygen irradianceWater irradianceDeck comments
## 1             9.5            1750           1620     <NA>
## 2              NA            1550           1620     <NA>
## 3              NA            1150           1620     <NA>
## 4              NA             975           1620     <NA>
## 5             8.8             870           1620     <NA>
## 6              NA             610           1620     <NA>
```

```r
rawdat$date <- as.Date(rawdat$sampledate, "%m/%d/%y")
head(rawdat)
```

```
##   lakeid  lakename year4 daynum sampledate depth temperature_C
## 1      L Paul Lake  1984    148    5/27/84  0.00          14.5
## 2      L Paul Lake  1984    148    5/27/84  0.25            NA
## 3      L Paul Lake  1984    148    5/27/84  0.50            NA
## 4      L Paul Lake  1984    148    5/27/84  0.75            NA
## 5      L Paul Lake  1984    148    5/27/84  1.00          14.5
## 6      L Paul Lake  1984    148    5/27/84  1.50            NA
##   dissolvedOxygen irradianceWater irradianceDeck comments       date
## 1             9.5            1750           1620     <NA> 1984-05-27
## 2              NA            1550           1620     <NA> 1984-05-27
## 3              NA            1150           1620     <NA> 1984-05-27
## 4              NA             975           1620     <NA> 1984-05-27
## 5             8.8             870           1620     <NA> 1984-05-27
## 6              NA             610           1620     <NA> 1984-05-27
```

```r
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).

```r
Peterdata <- filter(rawdat, lakename == "Peter Lake")
head(Peterdata)
```

```
##   lakeid   lakename year4 daynum sampledate depth temperature_C
## 1      R Peter Lake  1984    149    5/28/84  0.00          14.8
## 2      R Peter Lake  1984    149    5/28/84  0.25            NA
## 3      R Peter Lake  1984    149    5/28/84  0.50            NA
## 4      R Peter Lake  1984    149    5/28/84  0.75            NA
## 5      R Peter Lake  1984    149    5/28/84  1.00          14.8
## 6      R Peter Lake  1984    149    5/28/84  1.50            NA
##   dissolvedOxygen irradianceWater irradianceDeck comments       date
## 1             9.2            1630           1540     <NA> 1984-05-28
## 2              NA            1530           1540     <NA> 1984-05-28
## 3              NA            1250           1540     <NA> 1984-05-28
## 4              NA            1050           1540     <NA> 1984-05-28
## 5             8.8             890           1540     <NA> 1984-05-28
## 6              NA             660           1540     <NA> 1984-05-28
```

```r
unique(Peterdata$year4)
```

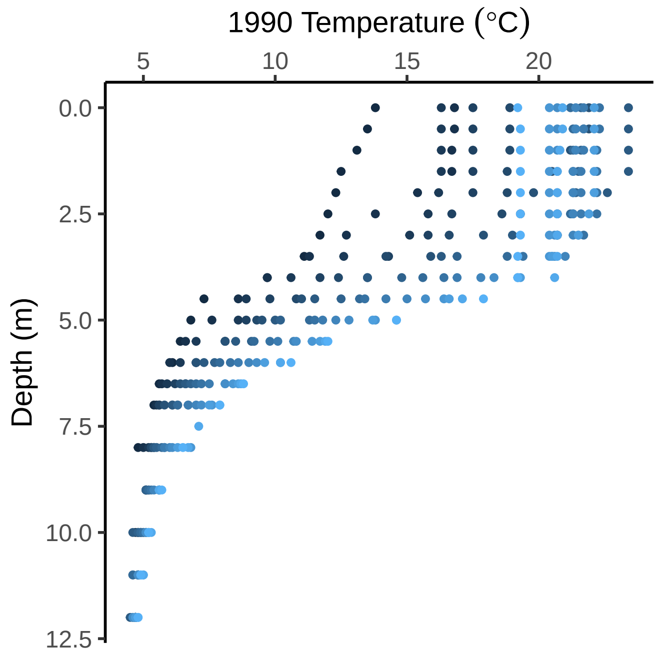
```
##  [1] 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997
## [15] 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## [29] 2012 2013 2014 2015 2016
```

```r
Peterdata2yrs <- filter(Peterdata, year4 %in% c(1990, 2010))
#Peterdata2yrs <- filter(Peterdata, year4 == 1990|year4 == 2010)
```

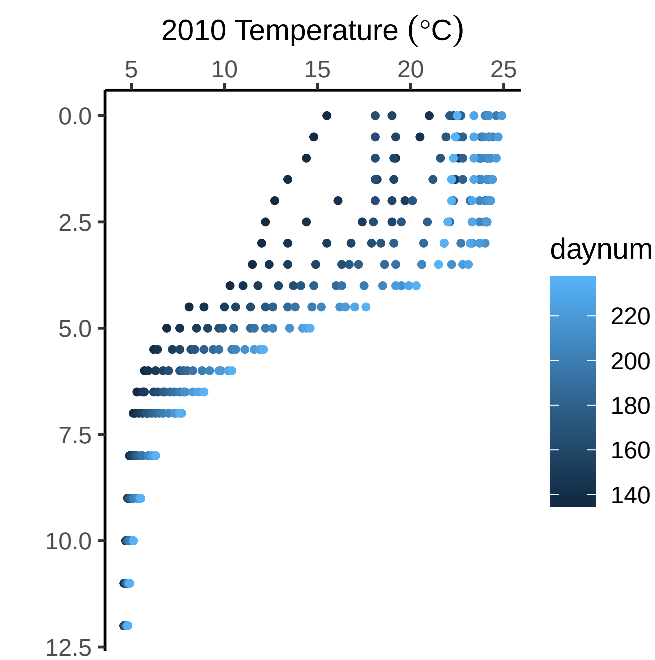
6. Create three graphs: (1) temperature profiles for the early year, (2) temperature profiles for the late year, and (3) a `plot_grid` of the two graphs together. Choose `geom_point` and color your points by date.

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

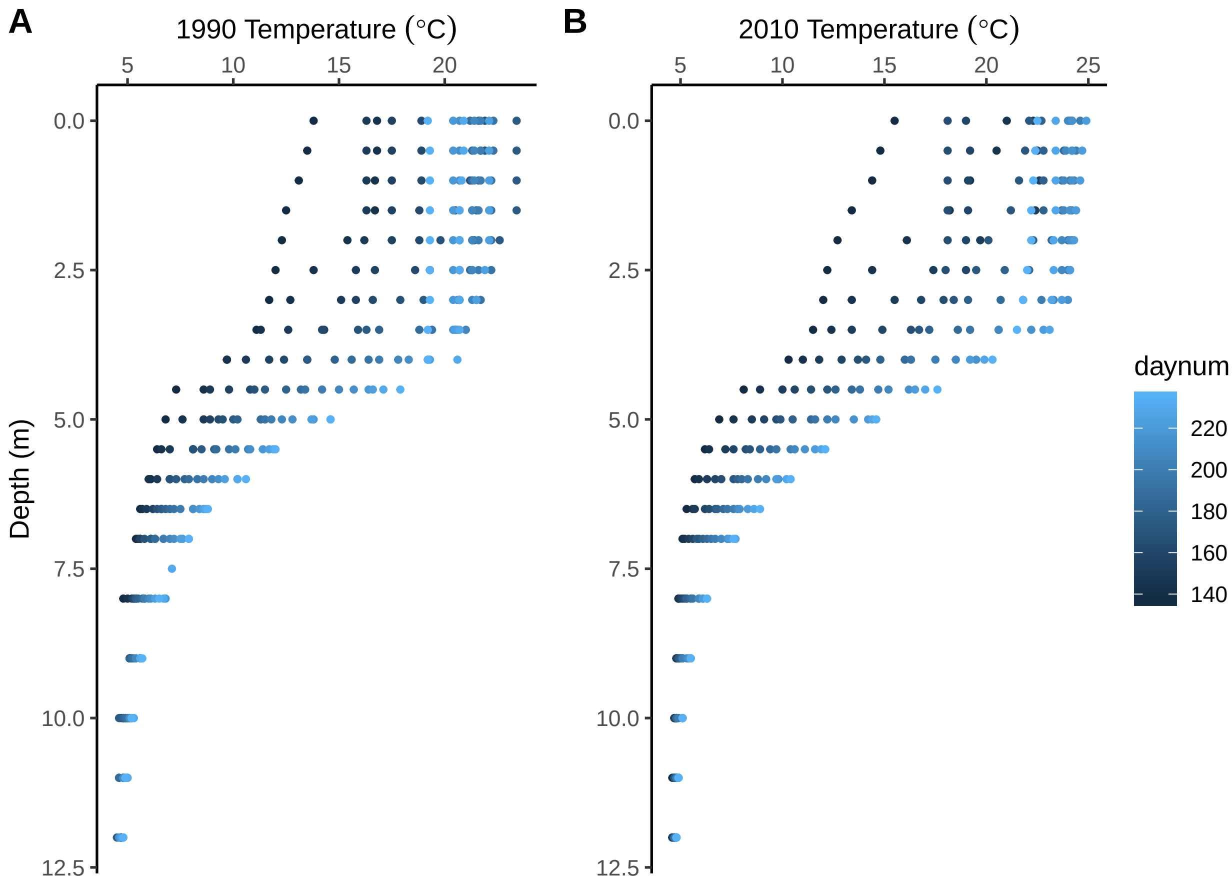
```r
Peterdata1990 <- filter(Peterdata2yrs, year4 == 1990)
Peterdata2010 <- filter(Peterdata2yrs, year4 == 2010)

p1 <- ggplot(Peterdata1990, aes(x = temperature_C, y = depth, color = daynum)) +
  geom_point(na.rm = T, size = 0.9) +
  scale_x_continuous(position = "top") +
  scale_y_reverse() +
  labs(x = expression("1990 Temperature "(degree*C)), y = "Depth (m)") +
  theme(legend.position = "none")
print(p1)
```



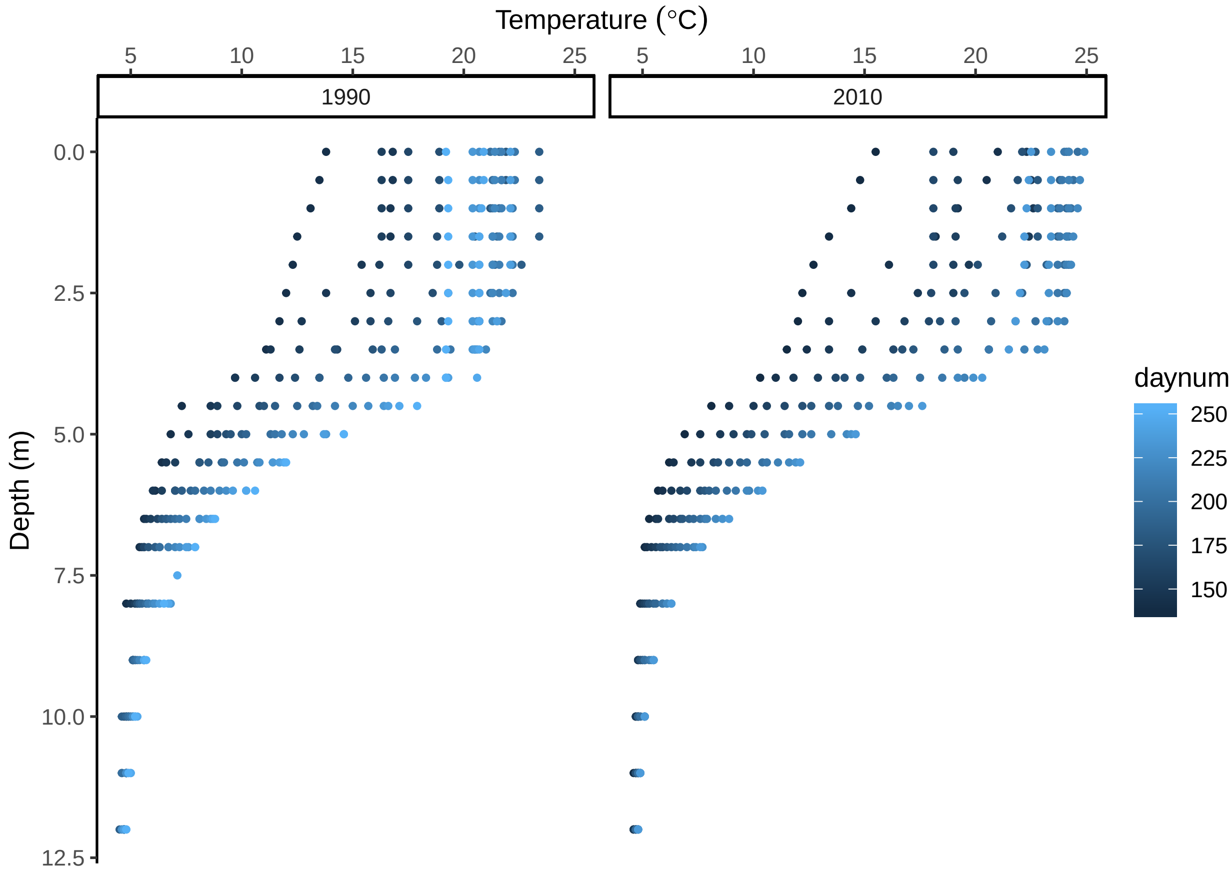

```
p2 <- ggplot(Peterdata2010, aes(x = temperature_C, y = depth, color = daynum)) +
  geom_point(na.rm = T, size = 0.9) +
  scale_x_continuous(position = "top") +
  scale_y_reverse() +
  labs(x = expression("2010 Temperature "(degree*C)), y = "") +
  theme(legend.position = "right", legend.margin = margin(0,0,0,0,"cm"))
print(p2)
```




```
plot_grid(p1, p2, ncol = 2, rel_widths = c(1, 1.25), labels = "AUTO")
```




```r
# Use ggplot facet
ggplot(Peterdata2yrs, aes(x = temperature_C, y = depth, color = daynum)) +
  geom_point(na.rm = T, size = 0.9) +
  scale_x_continuous(position = "top") +
  scale_y_reverse() +
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)") +
  facet_grid(cols = vars(year4)) +
  theme(legend.position = "right", legend.margin = margin(0,0,0,0,"cm"))
```



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

   In both years, the temperature between 0 to about 3 m does not change very much with depth, suggesting the hypolimnion. By contrast, temperature changes considerably with depth from about 3 m until about 8 m, suggesting the metalimnion within the range and the hypolimnion below. Also, temperature increases with ordinal date during the time period of samples. Despite the similarity in trend between the two years, the overall water temperature was higher in 2010 than 1990, especially near the surface.

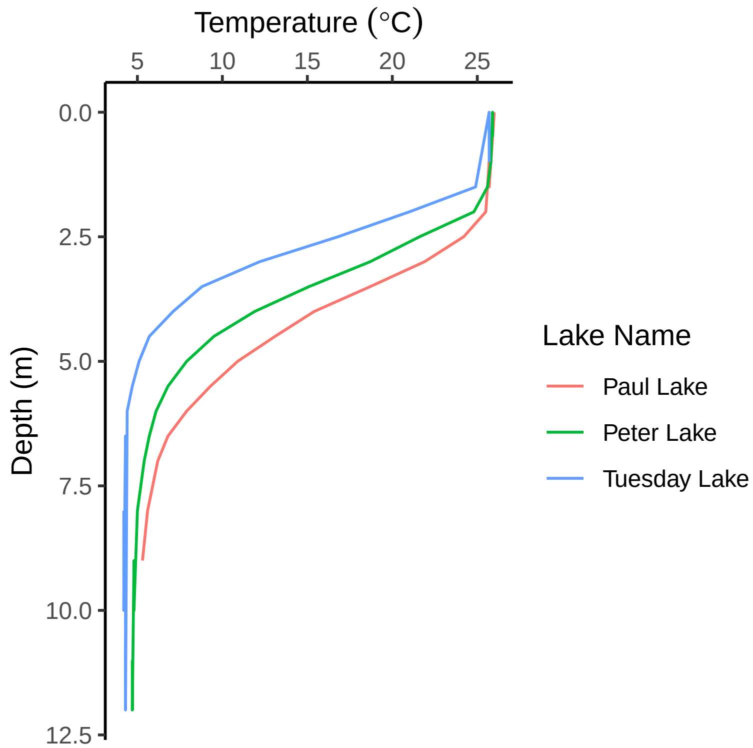
Multiple lakes, single date

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.

```
July2016data <-
  filter(rawdat, between(date, as.Date("2016-7-25"), as.Date("2016-7-27")))
```

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

```
ggplot(July2016data, aes(x = temperature_C, y = depth, color = lakename)) +
  geom_line(na.rm = T) +
  scale_x_continuous(position = "top") +
  scale_y_reverse() +
  labs(x = expression("Temperature "(degree*C)),y = "Depth (m)", color = "Lake Name")+
  theme(legend.position = "right", legend.margin = margin(0,0,0,0,"cm"))
```



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

- In Paul Lake, the epilimnion ranges from 0 to about 2 m, the thermocline from 2 to about 7 m, and the hypolimnion below 7 m.
- In Peter Lake, the epilimnion ranges from 0 to about 1.5 m, the thermocline from 1.5 to about 5 m, and the hypolimnion below 5 m.
- In Tuesday Lake, the epilimnion ranges from 0 to about 1.5 m, the thermocline from 1.5 to about 3.5 m, and the hypolimnion below 3.5 m.

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 <- filter(rawdat, lakename == "Peter Lake")
Petersurface <- Peterdata %>%
  mutate(month = month(date)) %>%
  filter(depth == 0 & between(month, 5, 8))

Petersurface5 <- filter(Petersurface, month == 5)
Petersurface6 <- filter(Petersurface, month == 6)
Petersurface7 <- filter(Petersurface, month == 7)
Petersurface8 <- filter(Petersurface, month == 8)

Peter5.lm <- lm(temperature_C ~ year4, data = Petersurface5)
summary(Peter5.lm)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Petersurface5)
##
## 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
```

```
Peter6.lm <- lm(temperature_C ~ year4, data = Petersurface6)
summary(Peter6.lm)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Petersurface6)
##
## 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
```

```
Peter7.lm <- lm(temperature_C ~ year4, data = Petersurface7)
summary(Peter7.lm)
```

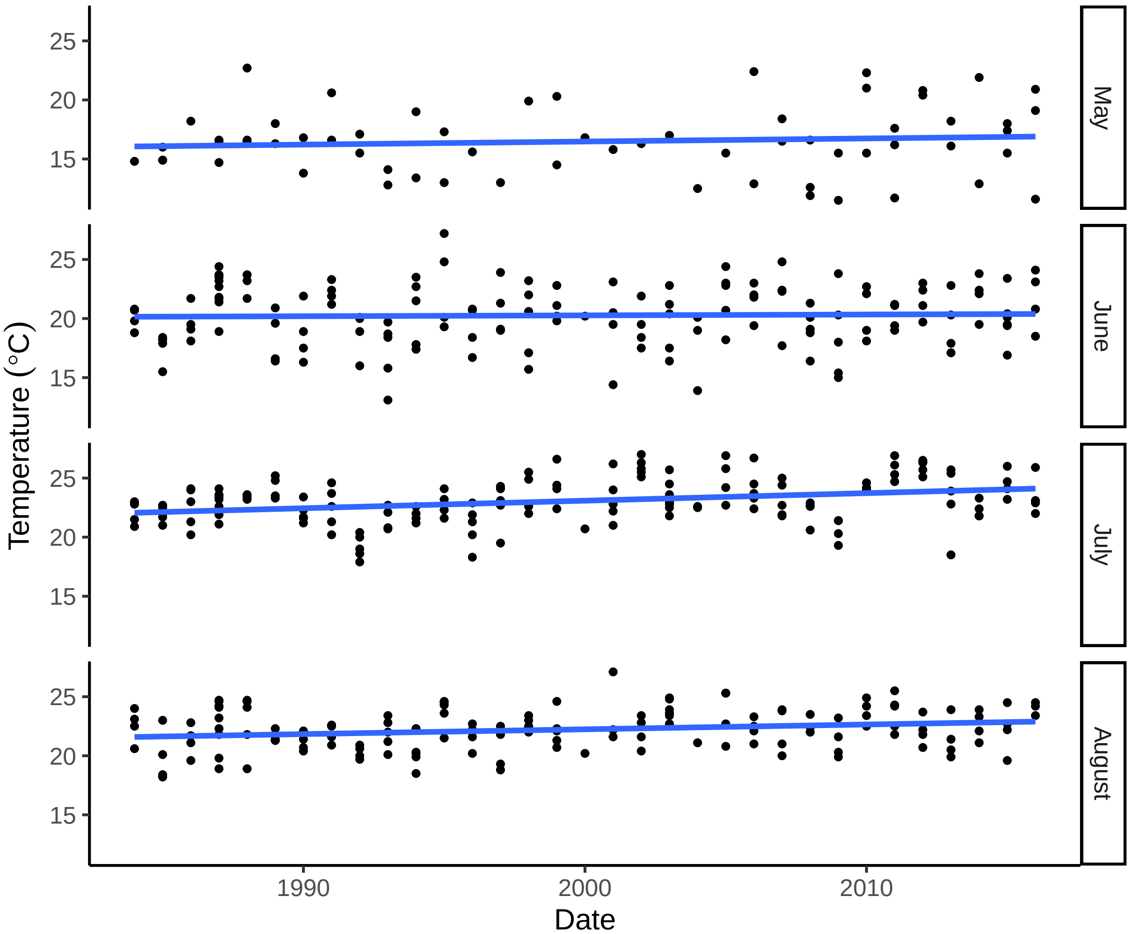
```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Petersurface7)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.4170 -1.0943  0.0191  1.2393  3.7861
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -104.75590   30.74789  -3.407 0.000846 ***
## 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
```

```
Peter8.lm <- lm(temperature_C ~ year4, data = Petersurface8)
summary(Peter8.lm)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Petersurface8)
##
## 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
```

```
p3 <- ggplot(data = Petersurface, aes(x = year4, y = temperature_C), na.rm = T) +
  geom_point(na.rm = T, size = 0.9) +
  geom_smooth(se = F, method = "lm") +
  labs(x = "Date", y = expression("Temperature "(degree*C)))

month.labs <- c("5" = "May", "6" = "June", "7" = "July", "8" = "August")
p3 + facet_grid(rows = vars(month) , labeller = labeller(month = month.labs))
```



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

    In Peter Lake, the increases in July and August over time are significant, while those in May and June are not. Similarly, the increases in Paul Lake in July and August are significant, while those in May and June not. Thus, the two lakes exhibit very similar trends in surface temperature over the last few decades.