Assignment 2: Physical Properties of Lakes

Keqi He

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/keqi/Desktop/Hydrologic_Data_Analysis"
```

```
# install.packages("tidyverse")
# install.packages("lubridate")
# install.packages("cowplot")

library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------------------

## 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 ---------------------------------------------------------------------------------- tid
## 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
```

```
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).

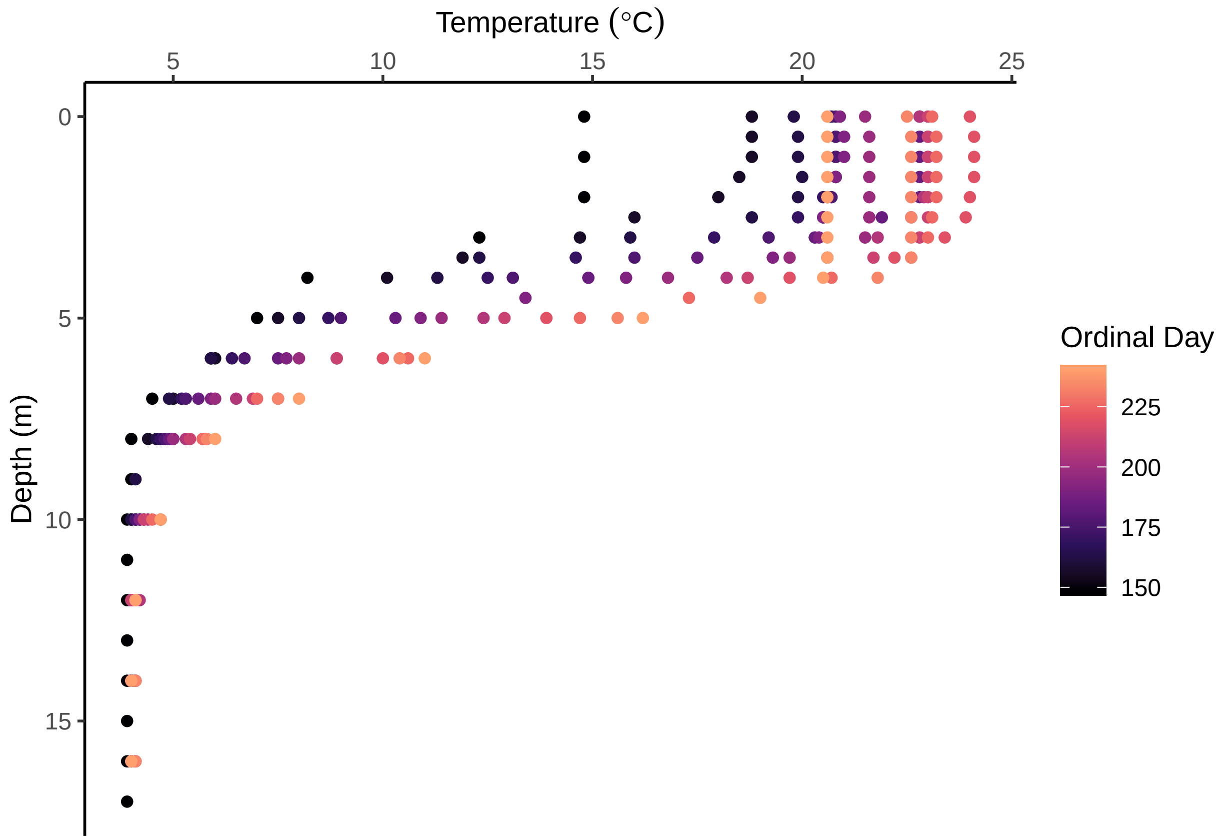
```
Peterdata <- filter(NTLdata, lakename == "Peter Lake" & (year4 == "1984" | year4 == "2016"))
```

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.

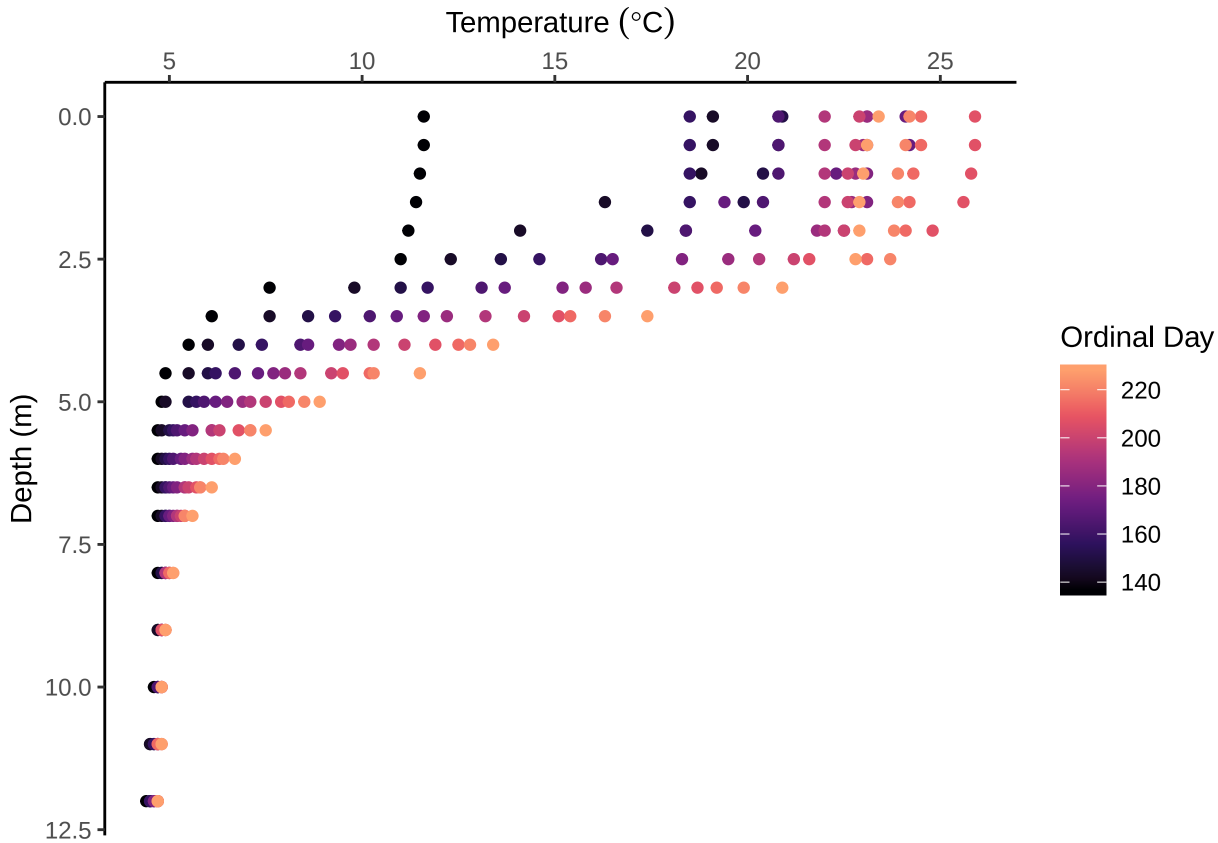
```
Peterdata1984 <- filter(Peterdata, year4 == "1984")
Peterdata2016 <- filter(Peterdata, year4 == "2016")
Tempprofiles1984 <-
  ggplot(Peterdata1984, aes(x = temperature_C, y = depth, color = daynum)) +
  geom_point() +
  scale_y_reverse() +
  scale_x_continuous(position = "top") +
  scale_color_viridis_c(end = 0.8, option = "magma") +
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)") +
  theme(legend.position = "right") +
  labs(color = "Ordinal Day")
print(Tempprofiles1984)
```

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




```
Tempprofiles2016 <-
  ggplot(Peterdata2016, 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)") +
  theme(legend.position = "right") +
  labs(color = "Ordinal Day")
print(Tempprofiles2016)
```

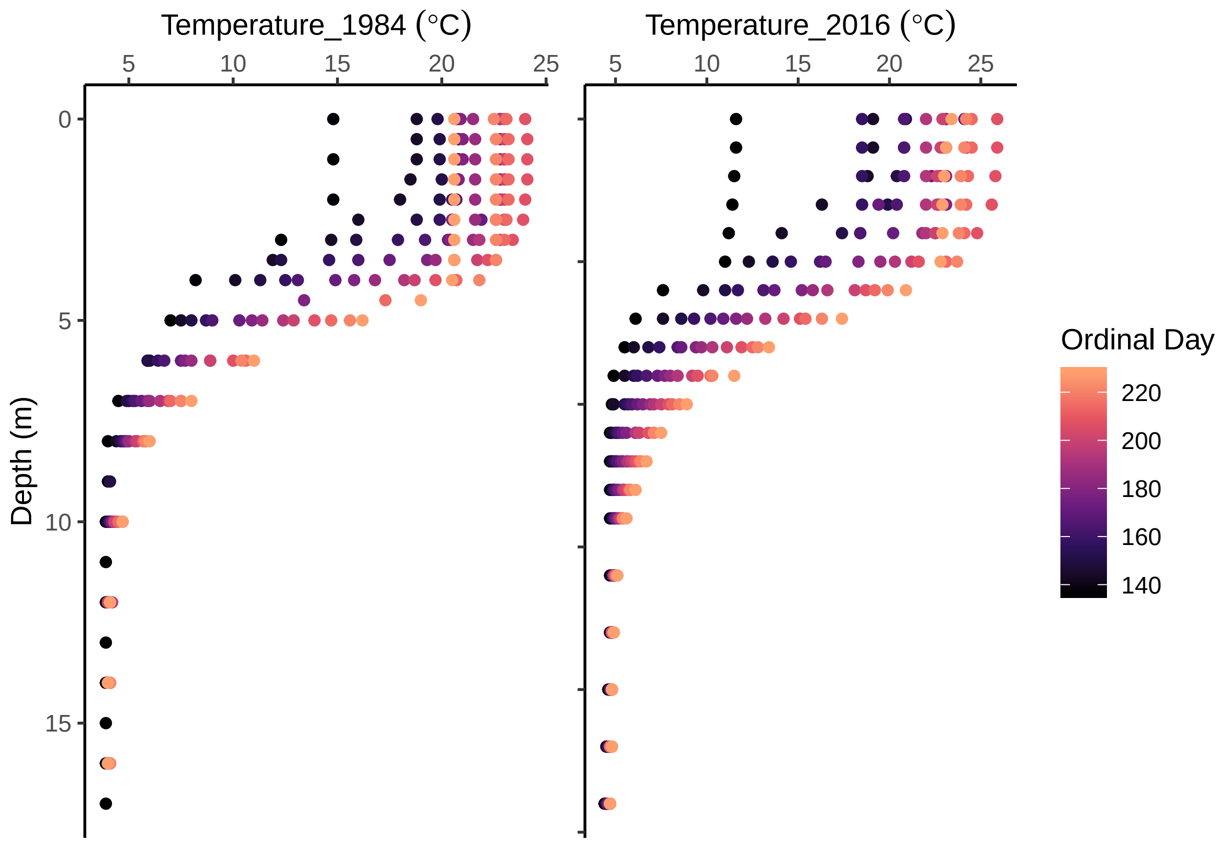
```
## Warning: Removed 28 rows containing missing values (geom_point).
```




```r
Tempprofiles19841 <-
  ggplot(Peterdata1984, 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_1984 "(degree*C)), y = "Depth (m)") +
  theme(legend.position = "none")

Tempprofiles20161 <-
  ggplot(Peterdata2016, 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_2016 "(degree*C)), y = "Depth (m)") +
  labs(color = "Ordinal Day") +
  theme(axis.text.y = element_blank(), axis.title.y = element_blank())

Physicalprofiles <-
  plot_grid(Tempprofiles19841, Tempprofiles20161,
            ncol = 2, rel_widths = c(1.25, 1.5))
```

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

## Warning: Removed 28 rows containing missing values (geom_point).
```

```r
print(Physicalprofiles)
```



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

   From May to August, the temperature of the lake increases with the increase of atmospheric temperature, and then decreases with the decrease of atmospheric temperature. And also the temperarure decrease with increasing depth. There is significant temperature gradient in thermocline of the lake, which is larger when atmosperic temperature is higher. the temperature of hypolimnion of the lake is more stable. Diferences between the two years: 1. Highest temperature of the lake in 2016 is higher than that in 1984. 2. Lowest temperature of epilimnion of the lake in 2016 is lower than that in 1984. 3. Temperature variance of the lake in 2016 is much larger than that in 1984.

Multiple lakes, single date

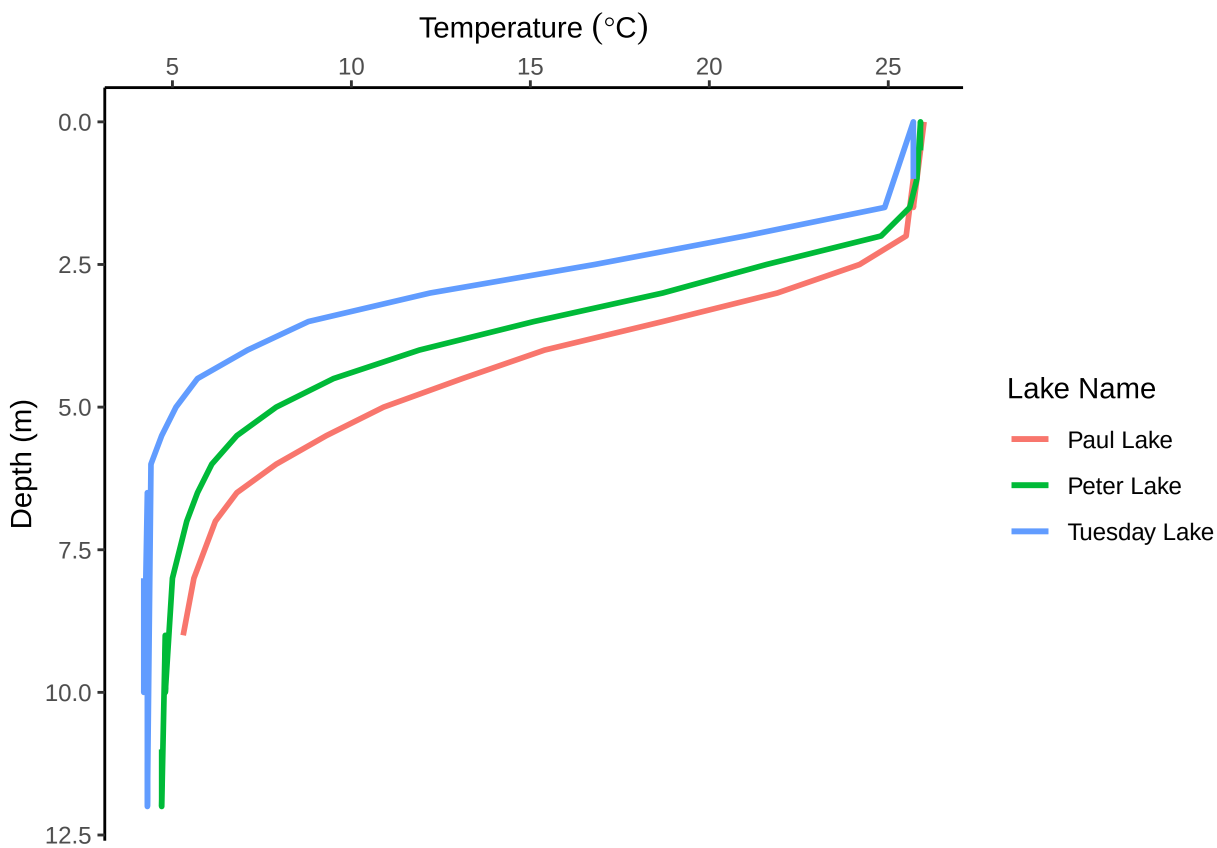
8. On July 25, 26, and 27 in 2016, all three lakes (Peter, Paul, and Tuesday) were sampled. Wrangle your data frame to include just these three dates.

```r
Threelakesdata <- filter(NTLdata, (lakename == "Peter Lake"| lakename == "Paul Lake"|
                                      lakename == "Tuesday Lake")
                                & (sampledate == "2016-07-25" | sampledate == "2016-07-26" |
                                   sampledate == "2016-07-27"))
```

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

```r
Threelakesdata_paul <- filter(Threelakesdata, lakename == "Paul Lake")
Threelakesdata_peter <- filter(Threelakesdata, lakename == "Peter Lake")
Threelakesdata_tuesday <- filter(Threelakesdata, lakename == "Tuesday Lake")
Tempprofilesthreelakes <-
  ggplot(Threelakesdata, aes(y = depth)) +
  geom_line(aes(x = temperature_C, color = "1"), data = Threelakesdata_paul, size = 1) +
  geom_line(aes(x = temperature_C, color = "2"), data = Threelakesdata_peter, size = 1) +
  geom_line(aes(x = temperature_C, color = "3"), data = Threelakesdata_tuesday, size = 1) +
  scale_color_discrete(name = "Lake Name", labels = c("Paul Lake", "Peter Lake","Tuesday Lake")) +
  scale_y_reverse() +
  scale_x_continuous(position = "top") +
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)")
print(Tempprofilesthreelakes)
```

```
## Warning: Removed 2 rows containing missing values (geom_path).

## Warning: Removed 2 rows containing missing values (geom_path).

## Warning: Removed 2 rows containing missing values (geom_path).
```



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

    Epilimnion(m): Paul Lake(0.0-2.0), Peter Lake(0.0-1.5), Tuesday Lake(0.0-1.5) Thermocline(m): Paul Lake(2.0-7.5), Peter Lake(1.5-7.5), Tuesday Lake(1.5-6.0) Hypolimnion(m): Paul Lake(Below 7.5), Peter Lake(Below 7.5), Tuesday Lake(Below 6.0)

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.

```r
Peterdata <- filter(NTLdata, lakename == "Peter Lake")
Peterdata <- mutate(Peterdata,month(sampledate))
names(Peterdata)[12] <- c("Month")
Peterdata_Summer_surface <- filter(Peterdata, (Month  == "5" | Month  =="6" |
                                                 Month  =="7" | Month  =="8")
                                  & (depth == "0"))
Peterdata_surface_5 <- filter(Peterdata_Summer_surface, Month  == "5" )
Peterdata_surface_6 <- filter(Peterdata_Summer_surface, Month  == "6" )
Peterdata_surface_7 <- filter(Peterdata_Summer_surface, Month  == "7" )
Peterdata_surface_8 <- filter(Peterdata_Summer_surface, Month  == "8" )

Maylinear <- lm(data = Peterdata_surface_5, temperature_C ~ year4)
summary(Maylinear)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Peterdata_surface_5)
##
## 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
```

```r
Junelinear <- lm(data = Peterdata_surface_6, temperature_C ~ year4)
summary(Junelinear)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Peterdata_surface_6)
##
## 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
```

```
Julylinear <- lm(data = Peterdata_surface_7, temperature_C ~ year4)
summary(Julylinear)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Peterdata_surface_7)
##
## 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
```

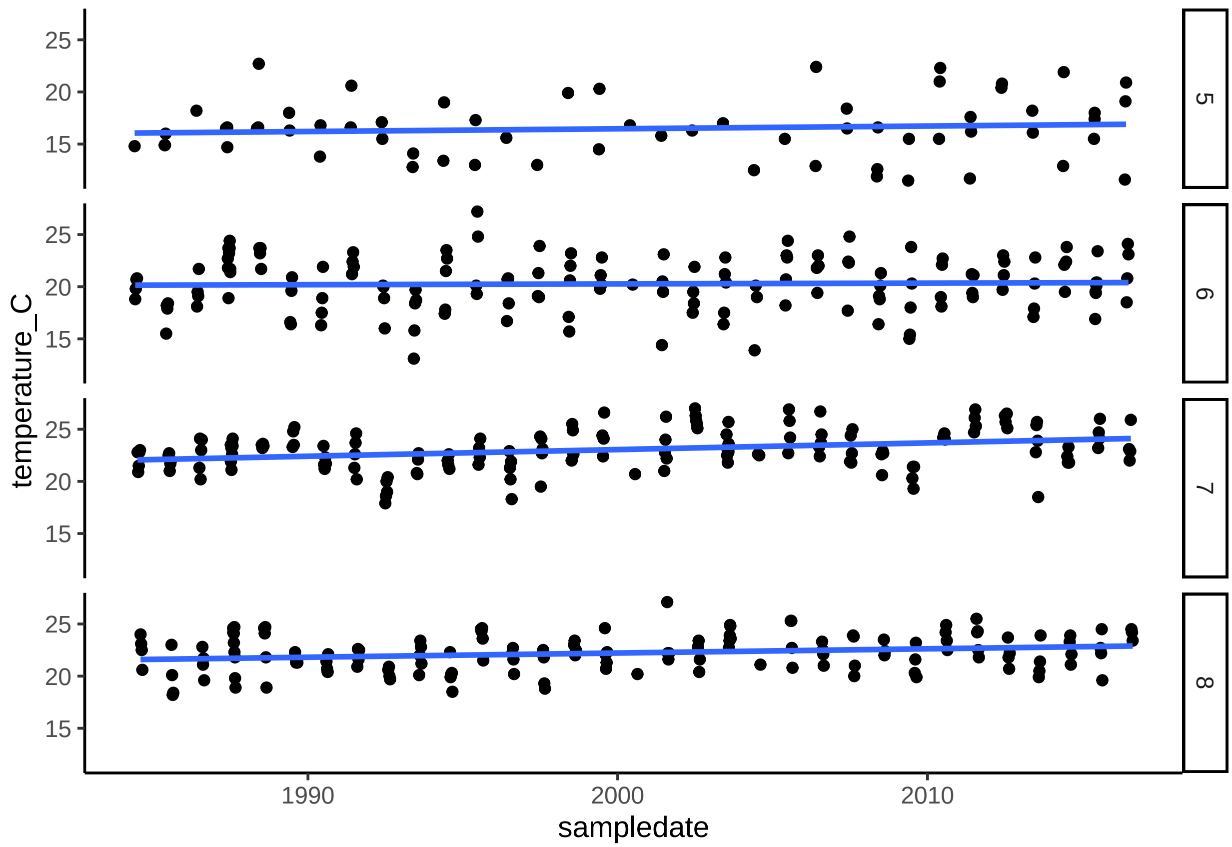
```
# Temperature increases 0.06 degree per year
# 0.06*31 = 1.86 degree increase over period of study for July
Auglinear <- lm(data = Peterdata_surface_8, temperature_C ~ year4)
summary(Auglinear)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Peterdata_surface_8)
##
## 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
```

```
# Temperature increases 0.04 degree per year
# 0.04*31 = 1.24 degree increase over period of study for August

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

```r
print(ggplotTchange)
```

```
## 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?

    They are quite similar. From May to August, the average temperature of surface water of both Paul Lake and Peter Lake per month is rising these years, especially the temperature in July and August, which has significant trends of rise. And the amount of rising temperarture of these two lakes warmed in July and August are same.