Assignment 2: Physical Properties of Lakes

Lindsay Roth

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/lindsayroth/Documents/MEM 2nd Year/HydroData/Hydrologic_Data_Analysis"
```

```
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
```

```
#update.packages(c("knitr", "stringr", "stringi"))

NTLdata <- read.csv("./Data/Raw/NTL-LTER_Lake_ChemistryPhysics_Raw.csv")

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).

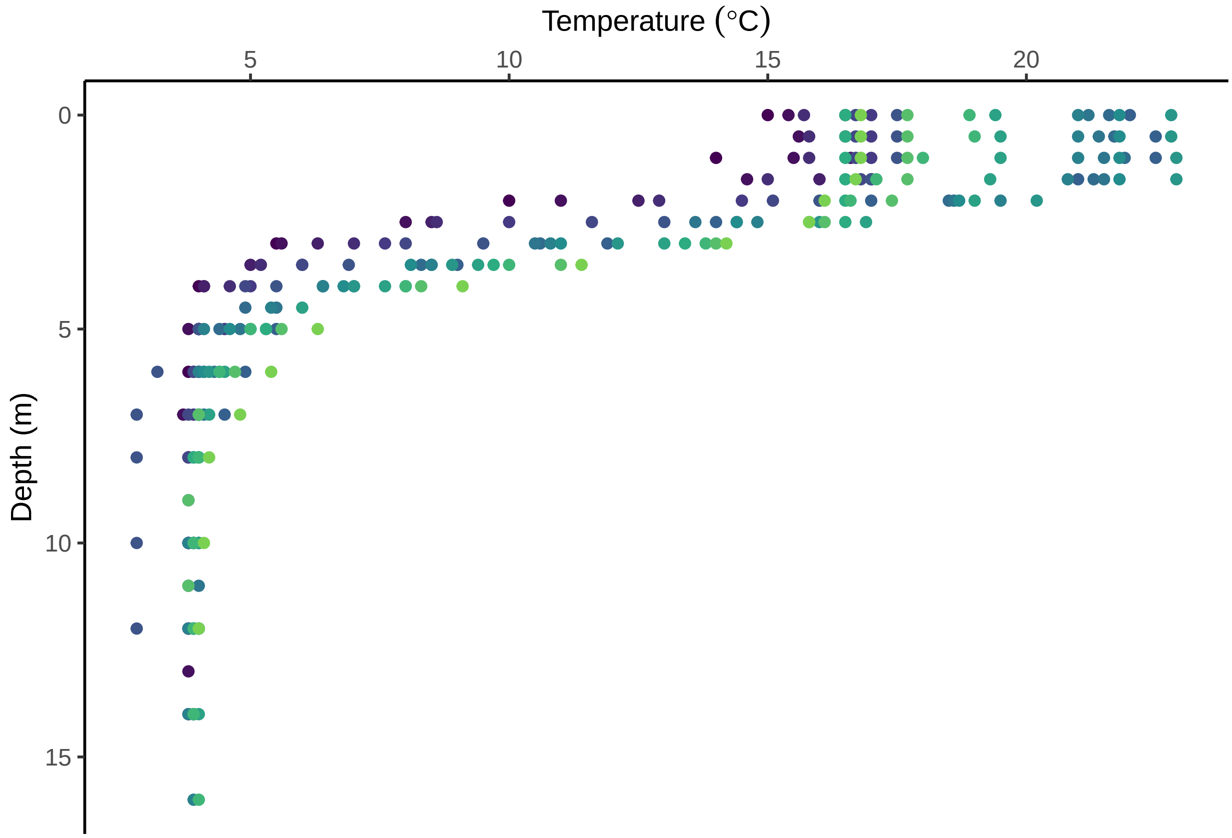
```
NTLdata$sampledate <- as.Date(NTLdata$sampledate, "%m/%d/%y")

TuesdayData <- NTLdata %>%
  filter(lakename == "Tuesday Lake")

Tuesday1985 <- TuesdayData %>%
  filter(year4 == 1985)

Tuesday2015 <- TuesdayData %>%
  filter(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.

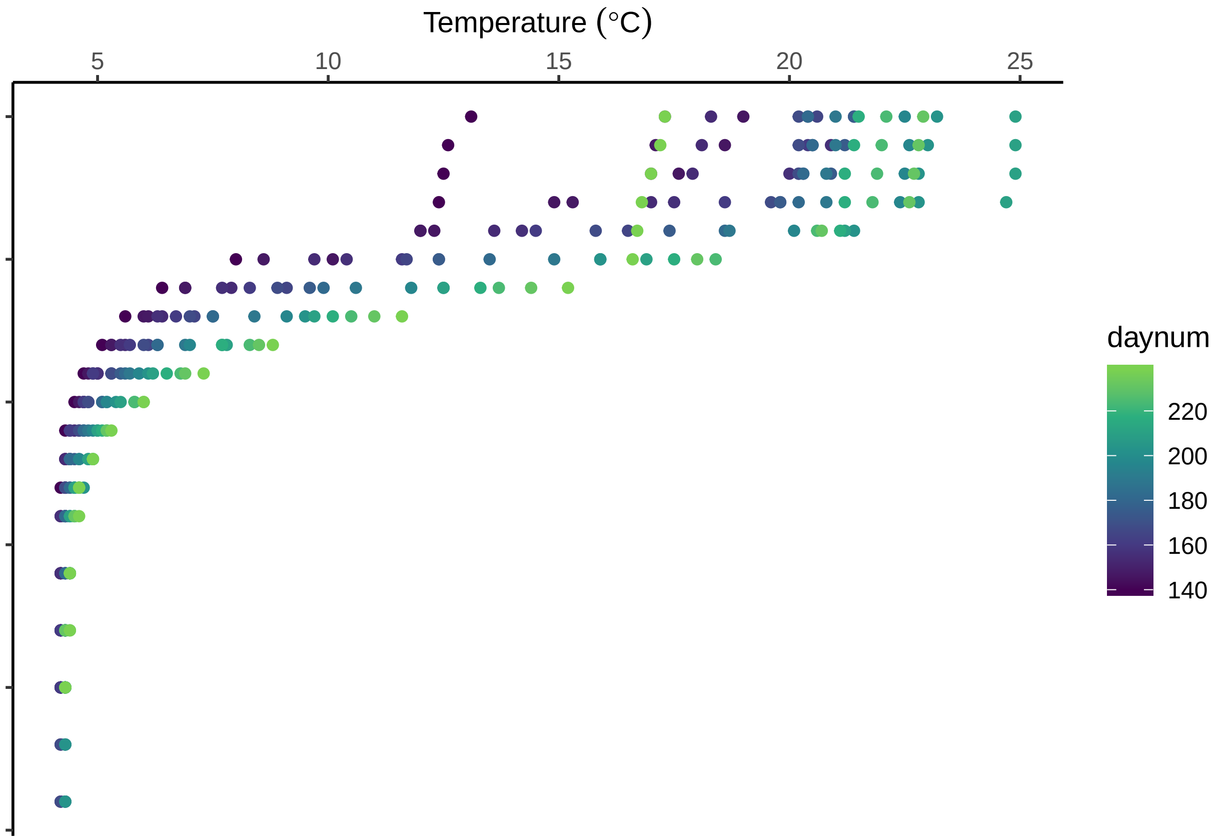
```
Tempprofiles1985 <-
  ggplot(Tuesday1985, aes(x = temperature_C, y = depth, color = daynum)) +
  geom_point() +
  scale_y_reverse() +
  scale_x_continuous(position = "top") +
  scale_color_viridis_c(end = 0.8) +
  #0.8 telling viridis not to use yellow
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)") +
  theme(legend.position = "none")
print(Tempprofiles1985)
```

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




```
Tempprofiles2015 <-
  ggplot(Tuesday2015, 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) +
  #0.8 telling viridis not to use yellow
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)") +
  theme(axis.text.y = element_blank(), axis.title.y = element_blank())
print(Tempprofiles2015)
```

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

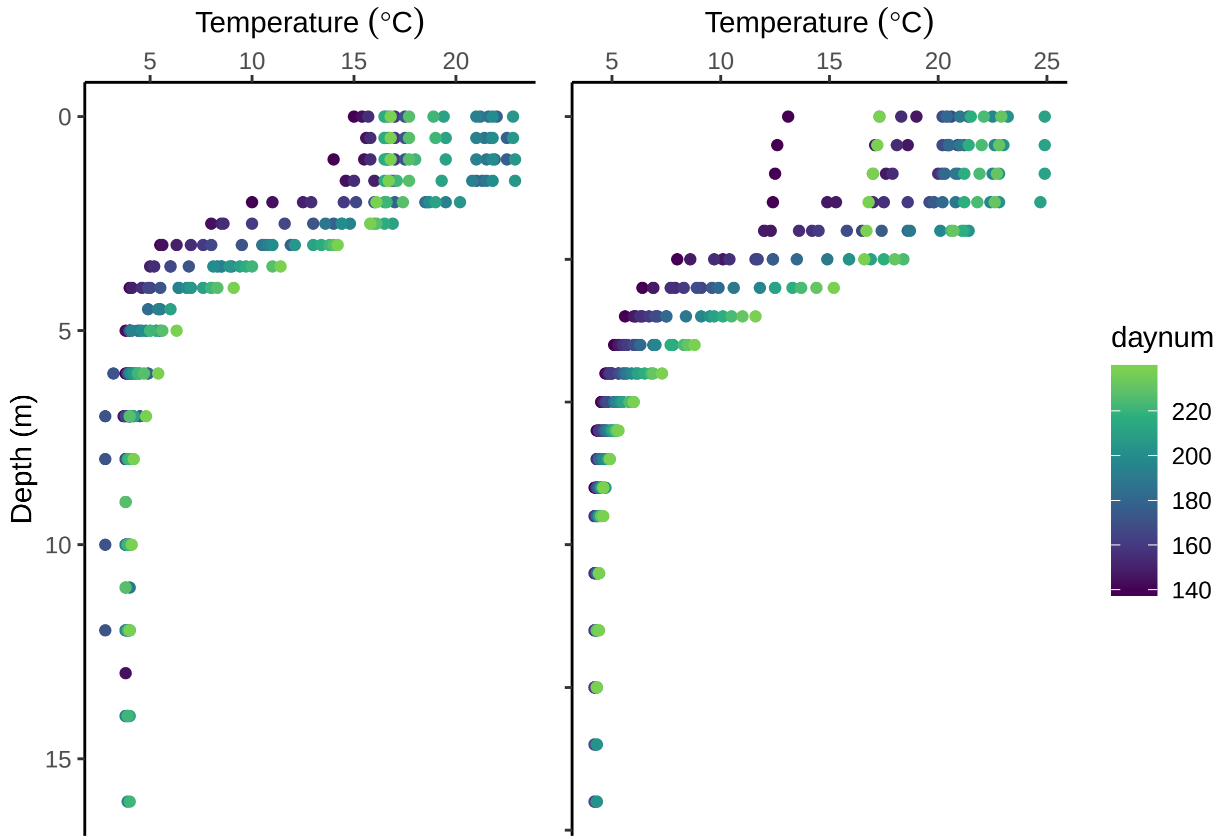



```
TuesdayTempProfiles <-
  plot_grid(Tempprofiles1985, Tempprofiles2015,
            ncol = 2,
            rel_widths = c(1, 1.25))
```

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

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

```
print(TuesdayTempProfiles)
```



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

   For the most part, 1985 and 2015 are similar. The warmest months are around the month of July, and the coolest monts are May and August. However, there are a couple of noticable differences. The temperature range in the epilimnion is greater in 2015 than it is in 1985. There is deeper mixing of surface temperatures in 2015, resulting in a deeper epilimnion in 2015. The thermocline in 2015 also appears to extent to deeper depths than 1985.

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.

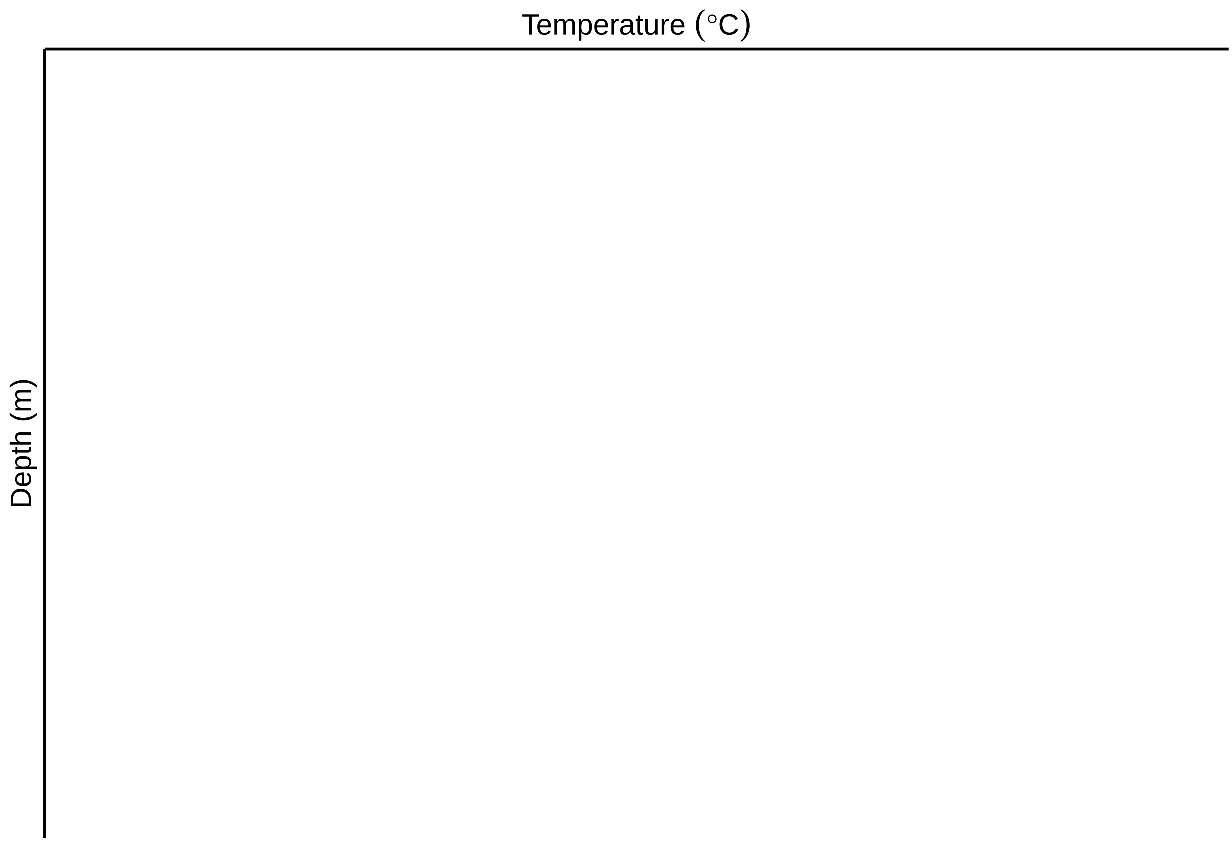
```
JulyData <- NTLdata %>%
  filter(lakename == "Peter Lake" | lakename == "Paul Lake" |
           lakename == "Tuesday Lake") %>%
  filter(sampledate == 2016-07-25 | sampledate == 2016-07-26 |
           sampledate == 2016-07-27)
```

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

```
as.numeric(JulyData$temperature_C)
```

```
## numeric(0)
```

```r
TempprofilesJuly<-
  ggplot(JulyData, aes(x = temperature_C, y = depth, color = lakename)) +
  geom_point() +
  scale_y_reverse() +
  scale_x_continuous(position = "top") +
  scale_color_viridis_d(end = 0.8) +
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)")
print(TempprofilesJuly)
```



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

    Paul Lake: The epilimnion is approximately 1.5-2 m deep, the thermocline from 2-6 m, and the hypolimnion is 6 m and lower. Peter Lake: The epilimnion is approximately 2-2.5 m deep, the thermocline from 2.5-7.5 m, and the hypolimnion 7.5 m and lower. Tuesday Lake: The epilimnion is about 2.5 m deep, the thermocline 2.5-8 m, and the hypolimnion 8 m and deeper.

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
Tuesdaydata.surface <- TuesdayData %>%
  mutate(Month = month(sampledate)) %>%
  filter(Month == 5 | Month == 6 | Month == 7 | Month == 8) %>%
  filter(depth == 0.00)
```

```r
Tuesdaydata.May <- Tuesdaydata.surface %>%
  filter(Month == 5)

Tuesdaydata.Jun <- Tuesdaydata.surface %>%
  filter(Month == 6)

Tuesdaydata.Jul <- Tuesdaydata.surface %>%
  filter(Month == 7)

Tuesdaydata.Aug <- Tuesdaydata.surface %>%
  filter(Month == 8)

#Step 4
May.Test <- lm(temperature_C ~ year4, Tuesdaydata.May)
summary(May.Test)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesdaydata.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
Jun.Test <- lm(temperature_C ~ year4, Tuesdaydata.Jun)
summary(Jun.Test)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesdaydata.Jun)
##
## 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
Jul.Test <- lm(temperature_C ~ year4, Tuesdaydata.Jul)
summary(Jul.Test)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesdaydata.Jul)
##
## 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
```

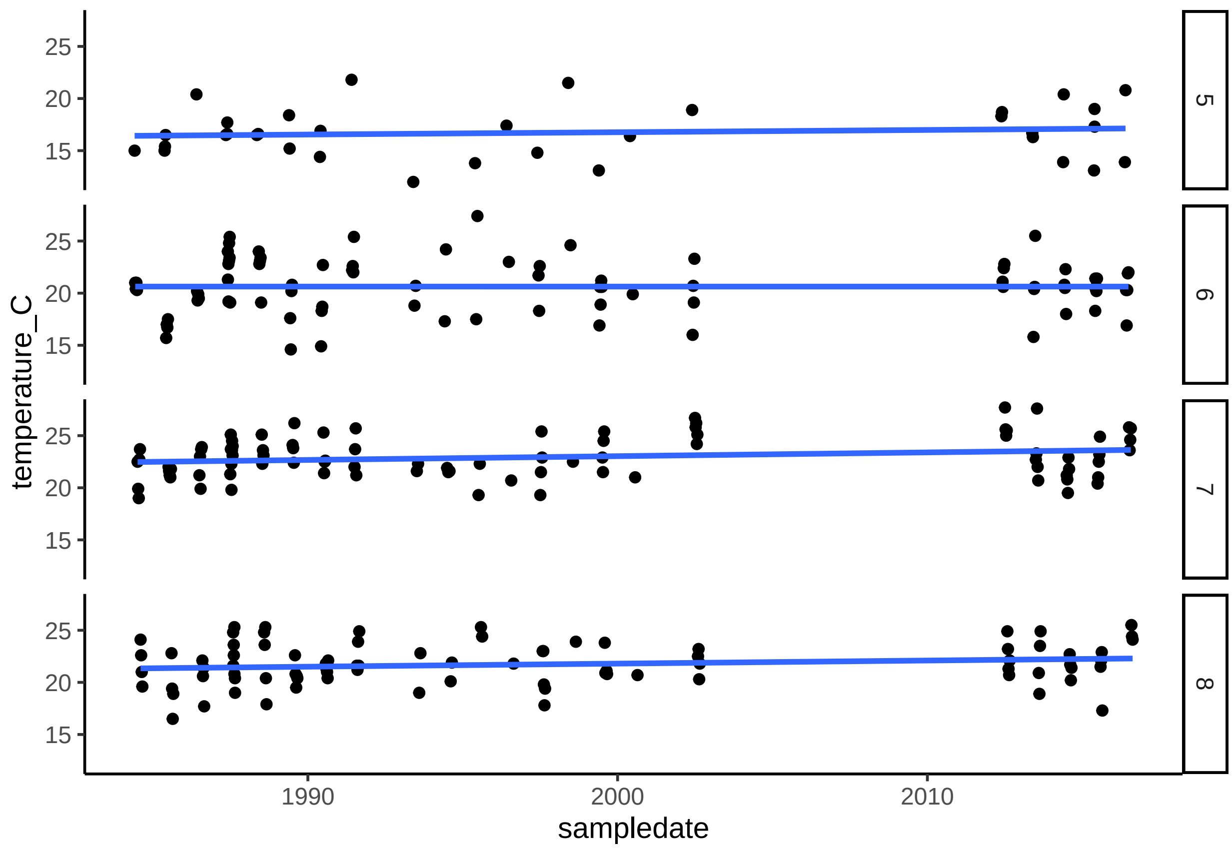
```r
Aug.Test <- lm(temperature_C ~ year4, Tuesdaydata.Aug)
summary(Aug.Test)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesdaydata.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
```

```r
TempChangePlot <-
  ggplot(Tuesdaydata.surface, aes(x = sampledate, y = temperature_C)) +
  geom_point() +
  geom_smooth(se = FALSE, method = lm) +
  facet_grid(rows = vars(Month))
print(TempChangePlot)
```

```
## 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?

    There are some similar trends between the Paul Lake and Tuesday Lake tests. The largest coefficients were for the months of July and August, and the smallest coefficients were for the months of May and June. However, for Paul Lake the tests for July and August were statistically significant, while none of the tests for Tuesday Lake were statistically significant.