Assignment 2: Physical Properties of Lakes

Keith Bollt

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] "Z:/Hydrologic_Data_Analysis2"
```

```
#I am using a new computer, so I have to install again!
#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 ------------------------------------------------------------------------------ tidy
## 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
```

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

theme_set(theme_light())
```

Creating and analyzing lake temperature profiles

Single lake, multiple dates

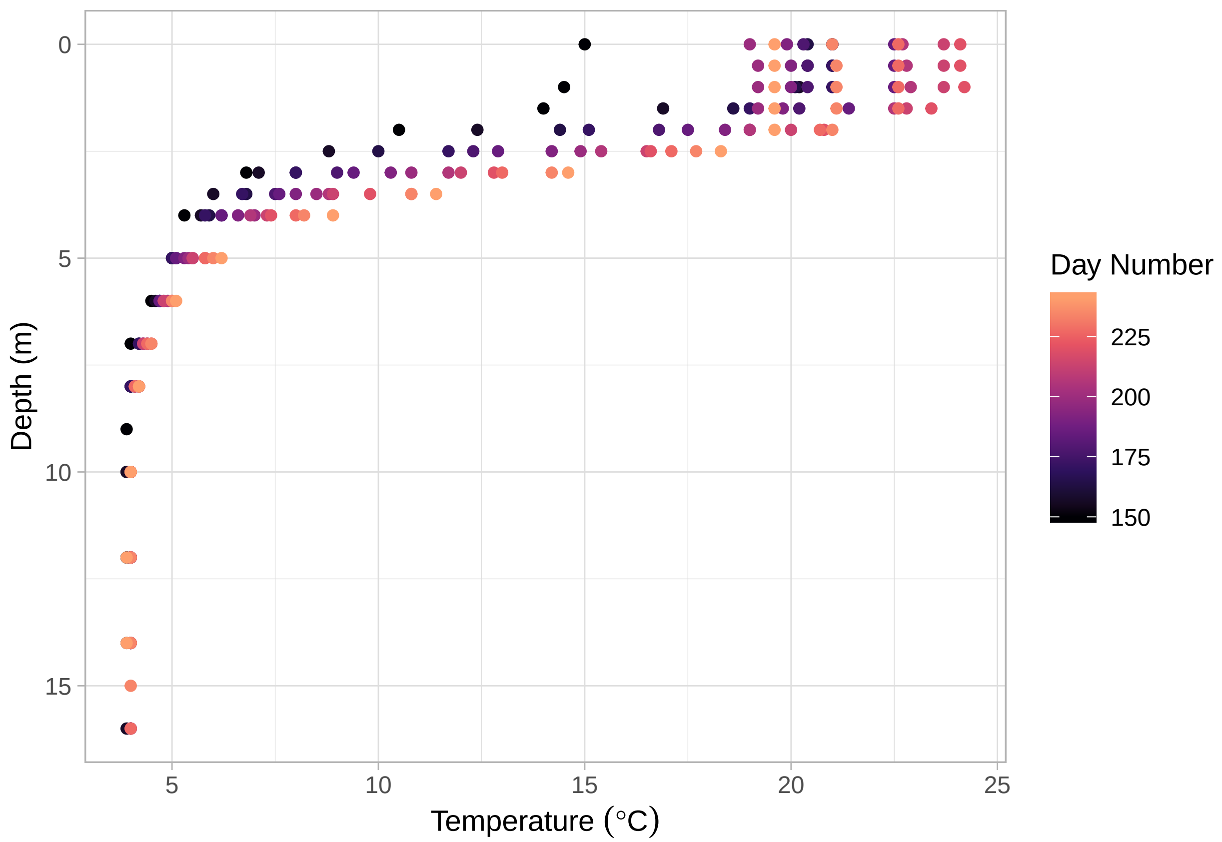
5. Choose either Peter or Tuesday Lake. Create a new data frame that wrangles the full data frame so that it only includes that lake during two different years (one year from the early part of the dataset and one year from the late part of the dataset).

```
Tuesday_NTL <- filter(NTLER_data, lakename == "Tuesday Lake")
Tuesday_filtered <- filter(Tuesday_NTL, year4 == 1984 | year4 == 2016)
```

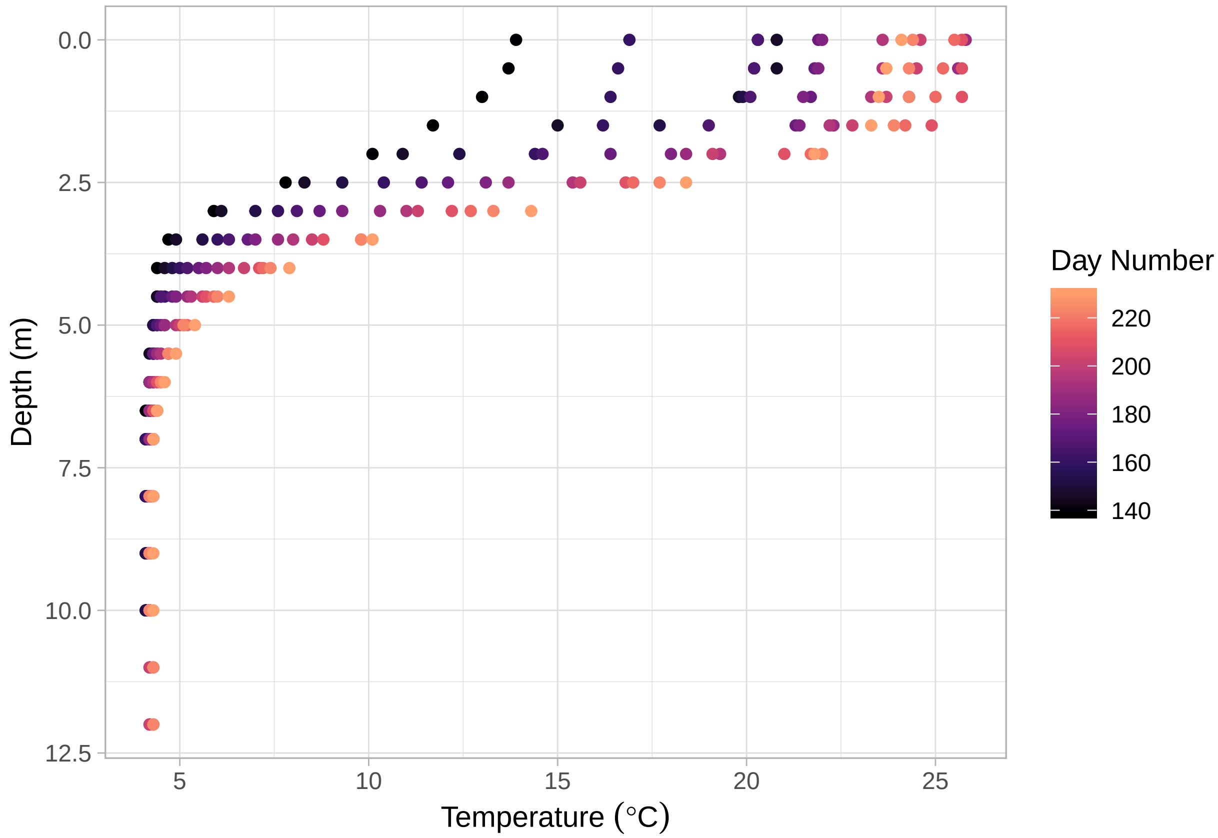
6. Create three graphs: (1) temperature profiles for the early year, (2) temperature profiles for the late year, and (3) a `plot_grid` of the two graphs together. Choose `geom_point` and color your points by date.

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

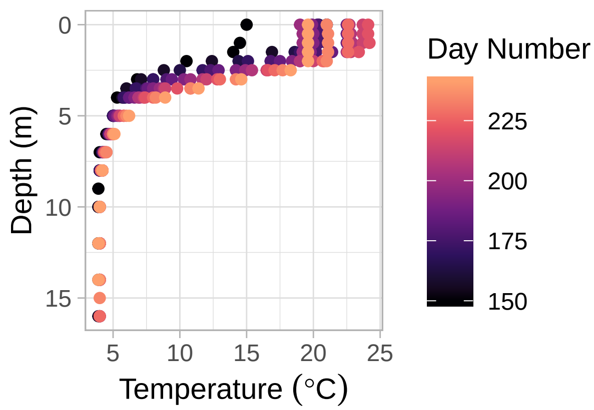
```
Tuesday_eightyfour <- filter(Tuesday_filtered, year4 == 1984)

Eightyfourplot <- ggplot(Tuesday_eightyfour, aes(x = temperature_C, y = depth, color = daynum))+
  geom_point()+
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)", color = "Day Number")+
  scale_y_reverse()+
  scale_color_viridis_c(end = 0.8, option = "magma")

print(Eightyfourplot)
```

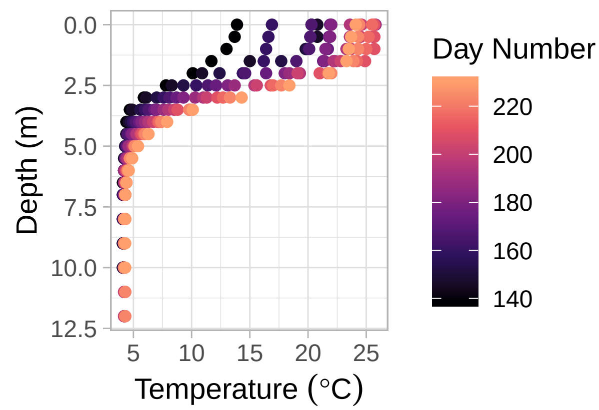



```
Tuesday_sixteen <- filter(Tuesday_filtered, year4 == 2016)

Sixteenplot <- ggplot(Tuesday_sixteen, aes(x = temperature_C, y = depth, color = daynum))+
  geom_point()+
  labs(x = expression("Temperature "(degree*C)), y = "Depth (m)", color = "Day Number")+
  scale_y_reverse()+
  scale_color_viridis_c(end = 0.8, option = "magma")

print(Sixteenplot)
```




```
Tuesday_comparison <-
  plot_grid(Eightyfourplot, Sixteenplot, ncol = 2, rel_widths = 1, 1.5)
print(Tuesday_comparison)
```





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

   I am seeing similar patterns in Tuesday Lake to those that we saw in class in Paul Lake. As the summer goes on, the surface warms, but the hypolimnion stays basically in the same place. By August, the surface is already starting to cool back down again for fall. The lake is warmer in 2016 than 1984 at all depths across the summer. From the naked eye, even the bottom of the lake looks like it is warmer in 2016.

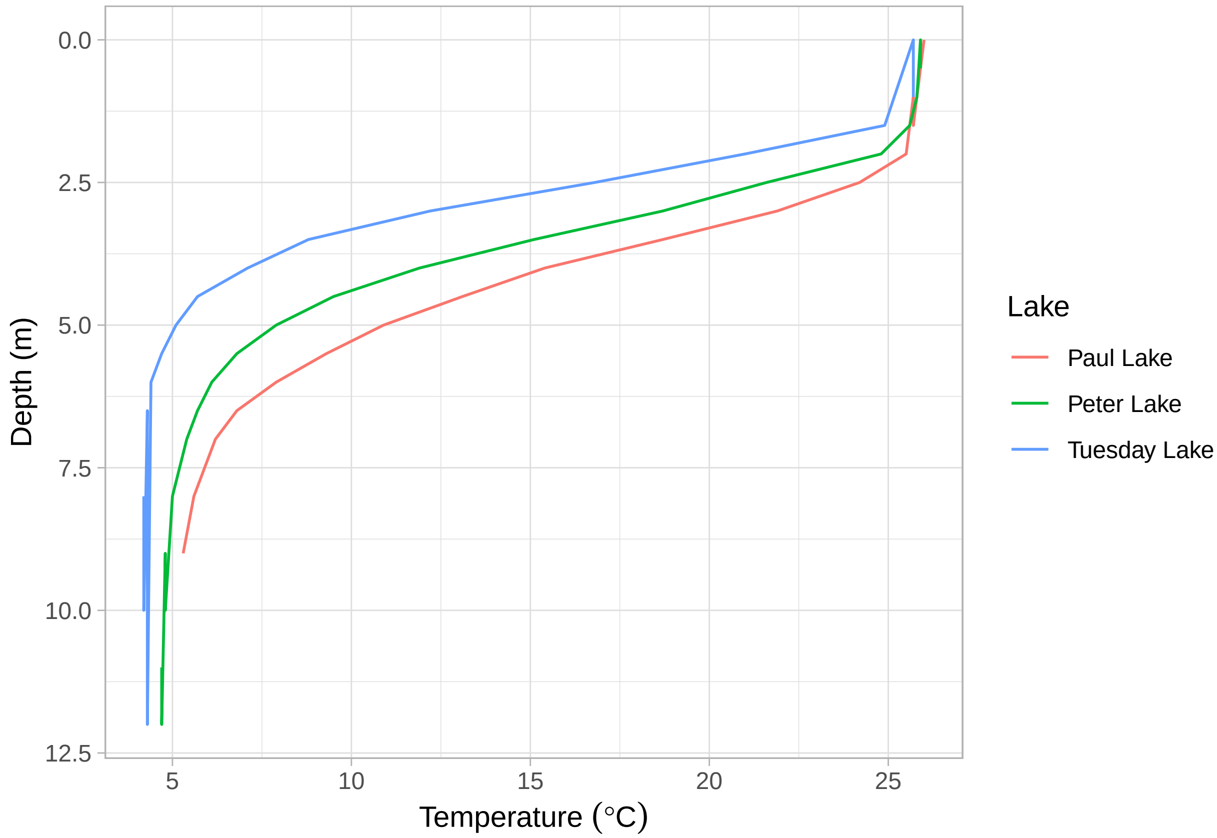
Multiple lakes, single date

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.

```
Julysixteen <- filter(NTLER_data,
                      lakename == "Peter Lake" | lakename == "Paul Lake" |lakename == "Tuesday Lake",
                      year4 == 2016, daynum == 207| daynum == 208 | daynum == 209)
```

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

```
threelakegraph <- ggplot(Julysixteen, aes(x = temperature_C, y = depth, color = lakename))+
  geom_line()+
  scale_y_reverse()+
   labs(x = expression("Temperature "(degree*C)), y = "Depth (m)", color = "Lake")

print(threelakegraph)
```



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

    (Paul Lake) Epilimnion: 0-2m, Thermocline: 2-7m, Hypolimnion: 7 meters and below (Peter Lake) Epilimnion: 0-1.7m, Thermocline: 1.7-6.5m, Hypolimnion: 6.5 meters and below (Tuesday Lake) Epilimnion: 0-1.5m, Thermocline: 2-6m, Hypolimnion: 6 meters and below

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.

```
Tuesdaysurface <- Tuesday_NTL %>%
  mutate(Month = month(sampledate)) %>%
  filter(depth == 0 & Month < 9)

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

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

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesdaysurface.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
#slope= 0.02 Celsius/year, p value = 0.56 (not statistically significant trend)

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

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesdaysurface.June)
##
## 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
#slope= -0.0002 Celsius/year, p value = 0.99 (not statistically significant trend)

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

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesdaysurface.July)
##
## 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
```

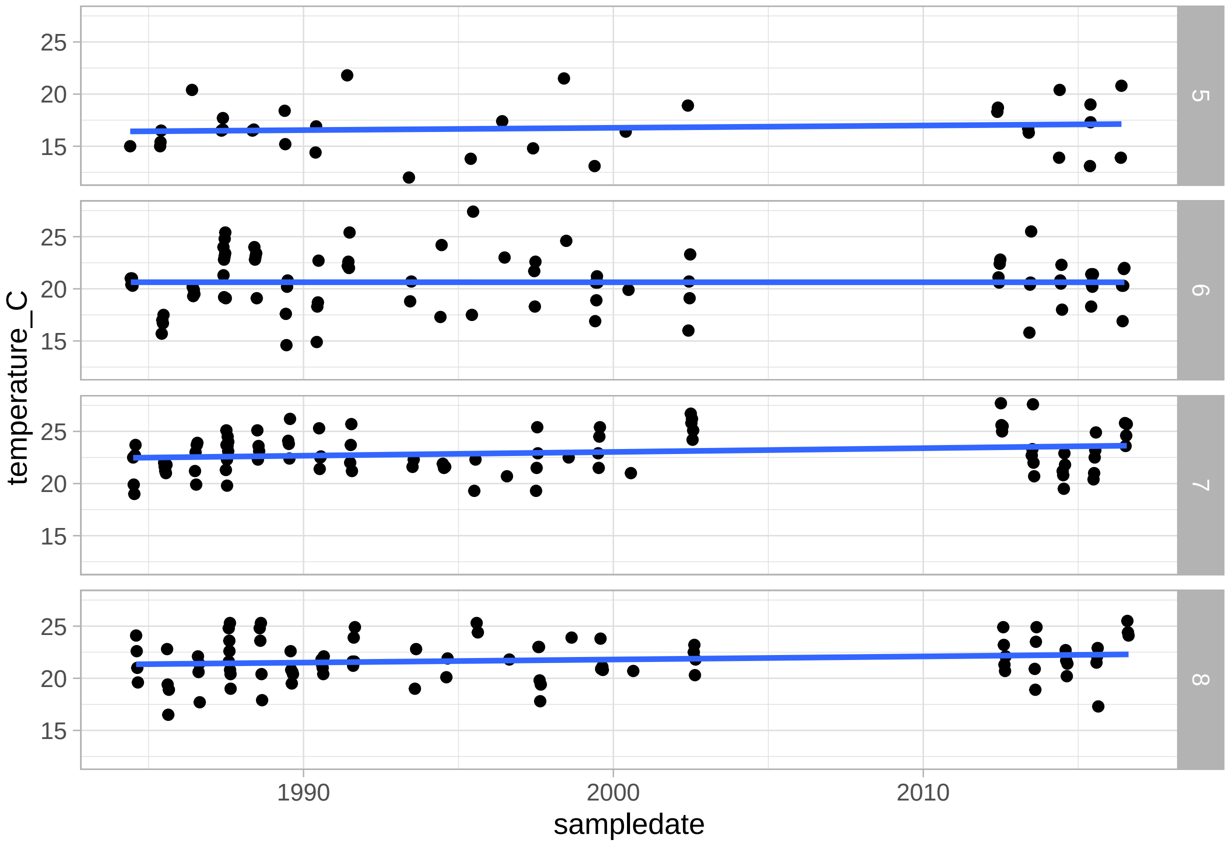
```r
#slope= 0.04 Celsius/year, p value = 0.06 (not statistically significant trend)

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

```
##
## Call:
## lm(formula = temperature_C ~ year4, data = Tuesdaysurface.August)
##
## 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
#slope= 0.03 Celsius/year, p value = 0.16 (not statistically significant trend)

ggplot(Tuesdaysurface, 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?

    I found for Tuesday Lake that for each of the four months sampled, there is not a significantly significant trend in surface temperature over the range of the dataset (altough July has a p-value of 0.057, which is pretty close). On the other hand, Paul Lake shows statistically significant, positive trends in surface temperature for the months of July (0.06 °C/year) and August (0.04 °C/year). Under the statistical test we conducted, Paul Lake is showing evidence of climate change, and Tuesday Lake is not showing evidence of climate change. Of course, this does not refute climate change, we would have to run other tests!