

Assignment 6: GLMs (Linear Regressios, ANOVA, & t-tests)

Ruiqing Li

Spring 2023

OVERVIEW

This exercise accompanies the lessons in Environmental Data Analytics on generalized linear models.

Directions

1. Rename this file <RuiqingLi>_A06_GLMs.Rmd (replacing <FirstLast> with your first and last name).
2. Change “Student Name” on line 3 (above) with your name.
3. Work through the steps, **creating code and output** that fulfill each instruction.
4. Be sure to **answer the questions** in this assignment document.
5. When you have completed the assignment, **Knit** the text and code into a single PDF file.

Set up your session

1. Set up your session. Check your working directory. Load the tidyverse, agricolae and other needed packages. Import the *raw* NTL-LTER raw data file for chemistry/physics (NTL-LTER_Lake_ChemistryPhysics_Raw.csv). Set date columns to date objects.
2. Build a ggplot theme and set it as your default theme.

```
#1
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr  0.3.4
## v tibble  3.1.8      v dplyr  1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.2      v forcats 0.5.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(lubridate)

##
## Attaching package: 'lubridate'
##
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
library(here)
```

```
## here() starts at /Users/ruiqingli/Desktop/DataAnalytics/RWORK/EDA-Spring2023
```

```
library(agricolae)
getwd()
```

```
## [1] "/Users/ruiqingli/Desktop/DataAnalytics/RWORK/EDA-Spring2023"
```

```
Raw.NTL.LTER <- read.csv(here("Data/Raw/NTL-LTER_Lake_ChemistryPhysics_Raw.csv"), stringsAsFactors = TRUE)
```

```
Raw.NTL.LTER$sampleddate <- as.Date(Raw.NTL.LTER$sampleddate, format = "%m/%d/%y")
#2
```

```
Mytheme.A6 <- theme_classic(base_size = 14) +
  theme(axis.text = element_text(color = "pink"),
        legend.position = "top")
```

```
theme_set(Mytheme.A6)
```

Simple regression

Our first research question is: Does mean lake temperature recorded during July change with depth across all lakes?

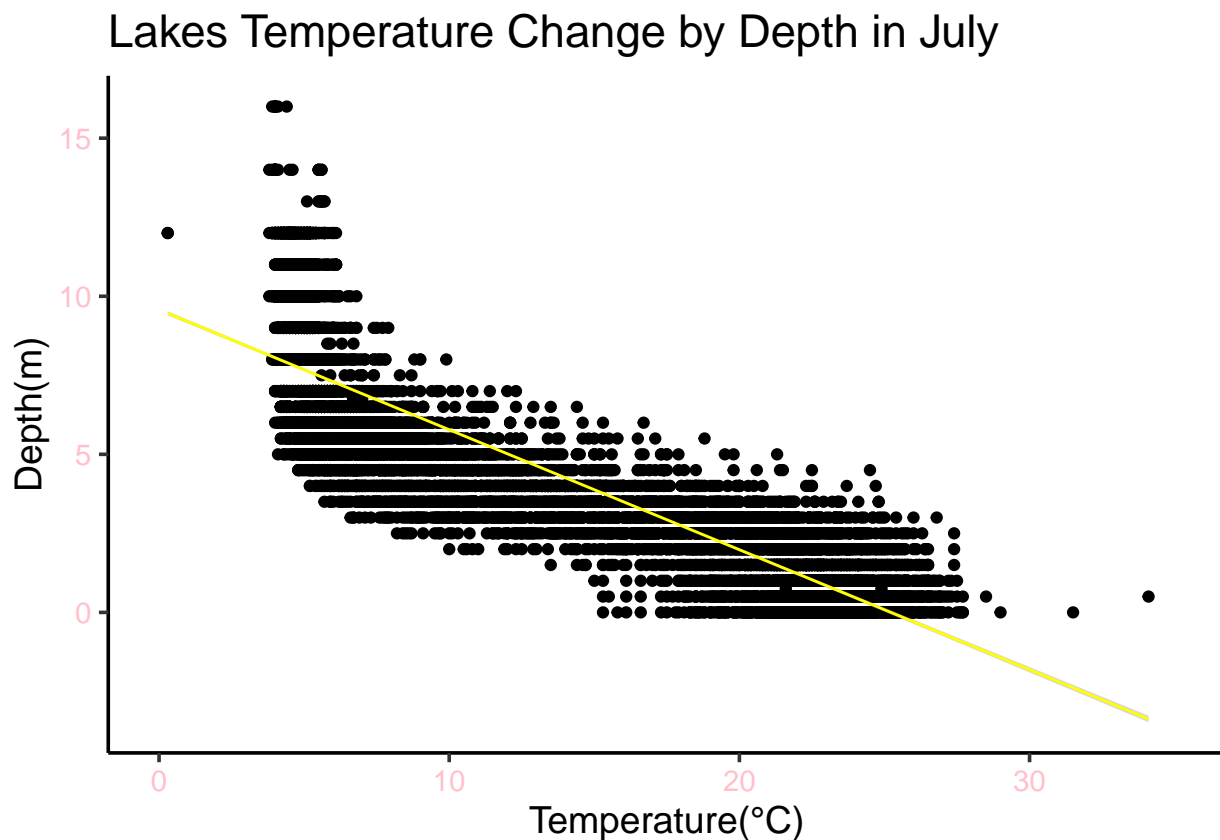
3. State the null and alternative hypotheses for this question: > Answer: H0: The mean lake temperature recorded during July does not change with depth across all lakes. Ha: The mean lake temperature recorded during July does change with depth across all lakes.
4. Wrangle your NTL-LTER dataset with a pipe function so that the records meet the following criteria:
 - Only dates in July.
 - Only the columns: lakename, year4, daynum, depth, temperature_C
 - Only complete cases (i.e., remove NAs)
5. Visualize the relationship among the two continuous variables with a scatter plot of temperature by depth. Add a smoothed line showing the linear model, and limit temperature values from 0 to 35 °C. Make this plot look pretty and easy to read.

```
#4
July.NTL.LTER <-
Raw.NTL.LTER %>%
  select(lakename, year4, daynum, depth, temperature_C) %>%
  filter(month(Raw.NTL.LTER$sampleddate) %in% 7) %>%
  na.omit()
```

```
#5
TemperatureDepth.NTL.LTER <-
ggplot(July.NTL.LTER, aes(x = temperature_C, y=depth)) +
  geom_point() +
  xlim(0, 35) +
```

```
geom_smooth(method=lm,color="yellow",size=0.5)+
  xlab("Temperature(°C)")+
  ylab("Depth(m)")+
  ggtitle("Lakes Temperature Change by Depth in July")
print(TemperatureDepth.NTL.LTER)
```

```
## 'geom_smooth()' using formula 'y ~ x'
```



6. Interpret the figure. What does it suggest with regards to the response of temperature to depth? Do the distribution of points suggest about anything about the linearity of this trend?

Answer: The temperature increase as the depth decrease. The distrubution of the trend suggests a negative linear relationship between temperature and depth.

7. Perform a linear regression to test the relationship and display the results

```
#7
depth.JulyLakes.Regression <- lm(July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth)
summary(depth.JulyLakes.Regression)
```

```
##
## Call:
## lm(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.5173 -3.0192  0.0633  2.9365 13.5834
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    21.95597    0.06792   323.3  <2e-16 ***
## July.NTL.LTER$depth -1.94621    0.01174  -165.8  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.835 on 9726 degrees of freedom
## Multiple R-squared:  0.7387, Adjusted R-squared:  0.7387
## F-statistic: 2.75e+04 on 1 and 9726 DF, p-value: < 2.2e-16
```

8. Interpret your model results in words. Include how much of the variability in temperature is explained by changes in depth, the degrees of freedom on which this finding is based, and the statistical significance of the result. Also mention how much temperature is predicted to change for every 1m change in depth.

Answer: There is a significant negative correlation between temperature and depth (higher temperature at lower depths), and that this model explains about 73.87 % of the total variance in temperature. The residual standard error is 3.835 on 9726 degrees of freedom. The F-statistic is 2.75e+04 on 1 and 9726 DF. The temperature is predicted to change 1.95°C every 1m change in depth.

Multiple regression

Let's tackle a similar question from a different approach. Here, we want to explore what might the best set of predictors for lake temperature in July across the monitoring period at the North Temperate Lakes LTER.

9. Run an AIC to determine what set of explanatory variables (year4, daynum, depth) is best suited to predict temperature.
10. Run a multiple regression on the recommended set of variables.

```
#9
depth.year4<-
  lm(July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth + July.NTL.LTER$year4 )
AIC(depth.year4)
```

```
## [1] 53756.97
```

```
summary(depth.year4)
```

```
##
## Call:
## lm(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth +
```

```
##      July.NTL.LTER$year4)
##
## Residuals:
##      Min        1Q    Median        3Q        Max
## -9.5543 -3.0227  0.0981  2.9492 13.7469
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   -1.104769    8.629545   -0.128  0.89813
## July.NTL.LTER$depth -1.946542    0.011733 -165.906 < 2e-16 ***
## July.NTL.LTER$year4  0.011538    0.004317   2.672  0.00754 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.834 on 9725 degrees of freedom
## Multiple R-squared:  0.7389, Adjusted R-squared:  0.7389
## F-statistic: 1.376e+04 on 2 and 9725 DF,  p-value: < 2.2e-16
```

```
depth.daynum<-
  lm(July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth + July.NTL.LTER$daynum )
AIC(depth.daynum)
```

```
## [1] 53679.36
```

```
summary(depth.daynum)
```

```
##
## Call:
## lm(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth +
##      July.NTL.LTER$daynum)
##
## Residuals:
##      Min        1Q    Median        3Q        Max
## -9.6174 -2.9809  0.0845  2.9681 13.4406
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   14.088588    0.855505   16.468 <2e-16 ***
## July.NTL.LTER$depth -1.946111    0.011685 -166.541 <2e-16 ***
## July.NTL.LTER$daynum  0.039836    0.004318   9.225 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.818 on 9725 degrees of freedom
## Multiple R-squared:  0.741, Adjusted R-squared:  0.741
## F-statistic: 1.391e+04 on 2 and 9725 DF,  p-value: < 2.2e-16
```

```
year4.daynum<-
  lm(July.NTL.LTER$temperature_C ~ July.NTL.LTER$daynum + July.NTL.LTER$year4 )
AIC(year4.daynum)
```

```
## [1] 66798.34
```

```
summary(year4.daynum)
```

```
##
## Call:
## lm(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$daynum +
##     July.NTL.LTER$year4)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.279  -7.158  -2.591   8.072  21.402
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -2.827705   16.944033  -0.167    0.867
## July.NTL.LTER$daynum  0.040484   0.008475   4.777 1.81e-06 ***
## July.NTL.LTER$year4   0.003779   0.008439   0.448   0.654
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.494 on 9725 degrees of freedom
## Multiple R-squared:  0.002363, Adjusted R-squared:  0.002158
## F-statistic: 11.52 on 2 and 9725 DF, p-value: 1.007e-05
```

```
year4<-
  lm(July.NTL.LTER$temperature_C ~ July.NTL.LTER$year4 )
AIC(year4)
```

```
## [1] 66819.14
```

```
summary(year4)
```

```
##
## Call:
## lm(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$year4)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.378  -7.221  -2.579   8.103  21.383
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.775522   16.888014   0.283   0.777
## July.NTL.LTER$year4  0.003975   0.008449   0.470   0.638
##
## Residual standard error: 7.502 on 9726 degrees of freedom
## Multiple R-squared:  2.276e-05, Adjusted R-squared: -8.006e-05
## F-statistic: 0.2213 on 1 and 9726 DF, p-value: 0.638
```

```
daynum<-
  lm(July.NTL.LTER$temperature_C ~ July.NTL.LTER$daynum )
AIC(daynum)
```

```
## [1] 66796.54
```

```
summary(daynum)
```

```
##
## Call:
## lm(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$daynum)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.320  -7.156  -2.594   8.077  21.399
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.722359    1.675347   2.819  0.00483 **
## July.NTL.LTER$daynum 0.040502    0.008475   4.779 1.79e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.494 on 9726 degrees of freedom
## Multiple R-squared:  0.002343, Adjusted R-squared:  0.00224
## F-statistic: 22.84 on 1 and 9726 DF, p-value: 1.786e-06
```

```
depth<-
  lm(July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth )
AIC(depth)
```

```
## [1] 53762.12
```

```
summary(depth)
```

```
##
## Call:
## lm(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.5173 -3.0192  0.0633  2.9365 13.5834
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    21.95597    0.06792   323.3  <2e-16 ***
## July.NTL.LTER$depth -1.94621    0.01174  -165.8  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.835 on 9726 degrees of freedom
## Multiple R-squared:  0.7387, Adjusted R-squared:  0.7387
## F-statistic: 2.75e+04 on 1 and 9726 DF, p-value: < 2.2e-16
```

```
year4.daynum.depth <- lm(data = July.NTL.LTER, July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth + July
AIC(year4.daynum.depth)
```

```
## [1] 53674.39
```

```
step(year4.daynum.depth)
```

```
## Start: AIC=26065.53
## July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth + July.NTL.LTER$year4 +
## July.NTL.LTER$daynum
##
##              Df Sum of Sq    RSS   AIC
## <none>                141687 26066
## - July.NTL.LTER$year4    1      101 141788 26070
## - July.NTL.LTER$daynum    1      1237 142924 26148
## - July.NTL.LTER$depth    1    404475 546161 39189

##
## Call:
## lm(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth +
## July.NTL.LTER$year4 + July.NTL.LTER$daynum, data = July.NTL.LTER)
##
## Coefficients:
## (Intercept) July.NTL.LTER$depth July.NTL.LTER$year4
## -8.57556 -1.94644 0.01134
## July.NTL.LTER$daynum
## 0.03978
```

```
#10
summary(year4.daynum.depth)
```

```
##
## Call:
## lm(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$depth +
## July.NTL.LTER$year4 + July.NTL.LTER$daynum, data = July.NTL.LTER)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.6536 -3.0000  0.0902  2.9658 13.6123
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -8.57556    8.630715  -0.994  0.32044
## July.NTL.LTER$depth -1.946437    0.011683 -166.611 < 2e-16 ***
## July.NTL.LTER$year4  0.011345    0.004299   2.639  0.00833 **
## July.NTL.LTER$daynum  0.039780    0.004317   9.215 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.817 on 9724 degrees of freedom
## Multiple R-squared:  0.7412, Adjusted R-squared:  0.7411
## F-statistic: 9283 on 3 and 9724 DF, p-value: < 2.2e-16
```


11. What is the final set of explanatory variables that the AIC method suggests we use to predict temperature in our multiple regression? How much of the observed variance does this model explain? Is this an improvement over the model using only depth as the explanatory variable?

Answer: All three explanatory variables (year4, daynum, depth) should be included. According to R-Squared Value, this model explained 74% of the total variance. The previous model using only depth explained about 73.87 % of the total variance in temperature. So, yes, it is an improvement.

Analysis of Variance

12. Now we want to see whether the different lakes have, on average, different temperatures in the month of July. Run an ANOVA test to complete this analysis. (No need to test assumptions of normality or similar variances.) Create two sets of models: one expressed as an ANOVA models and another expressed as a linear model (as done in our lessons).

#12

```
July.anova <- aov(data = July.NTL.LTER, July.NTL.LTER$temperature_C ~ July.NTL.LTER$lakename)
summary(July.anova)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## July.NTL.LTER$lakename      8  21642   2705.2      50 <2e-16 ***
## Residuals              9719 525813     54.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(July.anova)
```

```
## Analysis of Variance Table
##
## Response: July.NTL.LTER$temperature_C
##              Df Sum Sq Mean Sq F value    Pr(>F)
## July.NTL.LTER$lakename      8  21642   2705.2 50.003 < 2.2e-16 ***
## Residuals              9719 525813     54.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
July.anova.lm <- lm(data = July.NTL.LTER, July.NTL.LTER$temperature_C ~ July.NTL.LTER$lakename)
summary(July.anova.lm)
```

```
##
## Call:
## lm(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$lakename,
##     data = July.NTL.LTER)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -10.769 -6.614 -2.679 7.684 23.832
##
## Coefficients:
##
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 17.6664 0.6501 27.174 < 2e-16 ***
## July.NTL.LTER$lakenameCrampton Lake -2.3145 0.7699 -3.006 0.002653 **
## July.NTL.LTER$lakenameEast Long Lake -7.3987 0.6918 -10.695 < 2e-16 ***
## July.NTL.LTER$lakenameHummingbird Lake -6.8931 0.9429 -7.311 2.87e-13 ***
## July.NTL.LTER$lakenamePaul Lake -3.8522 0.6656 -5.788 7.36e-09 ***
## July.NTL.LTER$lakenamePeter Lake -4.3501 0.6645 -6.547 6.17e-11 ***
## July.NTL.LTER$lakenameTuesday Lake -6.5972 0.6769 -9.746 < 2e-16 ***
## July.NTL.LTER$lakenameWard Lake -3.2078 0.9429 -3.402 0.000672 ***
## July.NTL.LTER$lakenameWest Long Lake -6.0878 0.6895 -8.829 < 2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.355 on 9719 degrees of freedom
## Multiple R-squared: 0.03953, Adjusted R-squared: 0.03874
## F-statistic: 50 on 8 and 9719 DF, p-value: < 2.2e-16
```

```
anova(July.anova.lm)
```

```
## Analysis of Variance Table
##
## Response: July.NTL.LTER$temperature_C
## Df Sum Sq Mean Sq F value Pr(>F)
## July.NTL.LTER$lakename 8 21642 2705.2 50.003 < 2.2e-16 ***
## Residuals 9719 525813 54.1
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(July.anova)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$lakename, data = July.NTL.LTER)
##
## $'July.NTL.LTER$lakename'
## diff lwr upr p adj
## Crampton Lake-Central Long Lake -2.3145195 -4.7031913 0.0741524 0.0661566
## East Long Lake-Central Long Lake -7.3987410 -9.5449411 -5.2525408 0.0000000
## Hummingbird Lake-Central Long Lake -6.8931304 -9.8184178 -3.9678430 0.0000000
## Paul Lake-Central Long Lake -3.8521506 -5.9170942 -1.7872070 0.0000003
## Peter Lake-Central Long Lake -4.3501458 -6.4115874 -2.2887042 0.0000000
## Tuesday Lake-Central Long Lake -6.5971805 -8.6971605 -4.4972005 0.0000000
## Ward Lake-Central Long Lake -3.2077856 -6.1330730 -0.2824982 0.0193405
## West Long Lake-Central Long Lake -6.0877513 -8.2268550 -3.9486475 0.0000000
## East Long Lake-Crampton Lake -5.0842215 -6.5591700 -3.6092730 0.0000000
## Hummingbird Lake-Crampton Lake -4.5786109 -7.0538088 -2.1034131 0.0000004
## Paul Lake-Crampton Lake -1.5376312 -2.8916215 -0.1836408 0.0127491
## Peter Lake-Crampton Lake -2.0356263 -3.3842699 -0.6869828 0.0000999
## Tuesday Lake-Crampton Lake -4.2826611 -5.6895065 -2.8758157 0.0000000
```

## Ward Lake-Crampton Lake	-0.8932661	-3.3684639	1.5819317	0.9714459
## West Long Lake-Crampton Lake	-3.7732318	-5.2378351	-2.3086285	0.0000000
## Hummingbird Lake-East Long Lake	0.5056106	-1.7364925	2.7477137	0.9988050
## Paul Lake-East Long Lake	3.5465903	2.6900206	4.4031601	0.0000000
## Peter Lake-East Long Lake	3.0485952	2.2005025	3.8966879	0.0000000
## Tuesday Lake-East Long Lake	0.8015604	-0.1363286	1.7394495	0.1657485
## Ward Lake-East Long Lake	4.1909554	1.9488523	6.4330585	0.0000002
## West Long Lake-East Long Lake	1.3109897	0.2885003	2.3334791	0.0022805
## Paul Lake-Hummingbird Lake	3.0409798	0.8765299	5.2054296	0.0004495
## Peter Lake-Hummingbird Lake	2.5429846	0.3818755	4.7040937	0.0080666
## Tuesday Lake-Hummingbird Lake	0.2959499	-1.9019508	2.4938505	0.9999752
## Ward Lake-Hummingbird Lake	3.6853448	0.6889874	6.6817022	0.0043297
## West Long Lake-Hummingbird Lake	0.8053791	-1.4299320	3.0406903	0.9717297
## Peter Lake-Paul Lake	-0.4979952	-1.1120620	0.1160717	0.2241586
## Tuesday Lake-Paul Lake	-2.7450299	-3.4781416	-2.0119182	0.0000000
## Ward Lake-Paul Lake	0.6443651	-1.5200848	2.8088149	0.9916978
## West Long Lake-Paul Lake	-2.2356007	-3.0742314	-1.3969699	0.0000000
## Tuesday Lake-Peter Lake	-2.2470347	-2.9702236	-1.5238458	0.0000000
## Ward Lake-Peter Lake	1.1423602	-1.0187489	3.3034693	0.7827037
## West Long Lake-Peter Lake	-1.7376055	-2.5675759	-0.9076350	0.0000000
## Ward Lake-Tuesday Lake	3.3893950	1.1914943	5.5872956	0.0000609
## West Long Lake-Tuesday Lake	0.5094292	-0.4121051	1.4309636	0.7374387
## West Long Lake-Ward Lake	-2.8799657	-5.1152769	-0.6446546	0.0021080

13. Is there a significant difference in mean temperature among the lakes? Report your findings.

Answer: Yes, there is. For every meter of difference in lake depth, the temperature among different lakes change between 2.3°C and 6.9°C.

14. Create a graph that depicts temperature by depth, with a separate color for each lake. Add a `geom_smooth` (method = "lm", se = FALSE) for each lake. Make your points 50 % transparent. Adjust your y axis limits to go from 0 to 35 degrees. Clean up your graph to make it pretty.

```
#14.

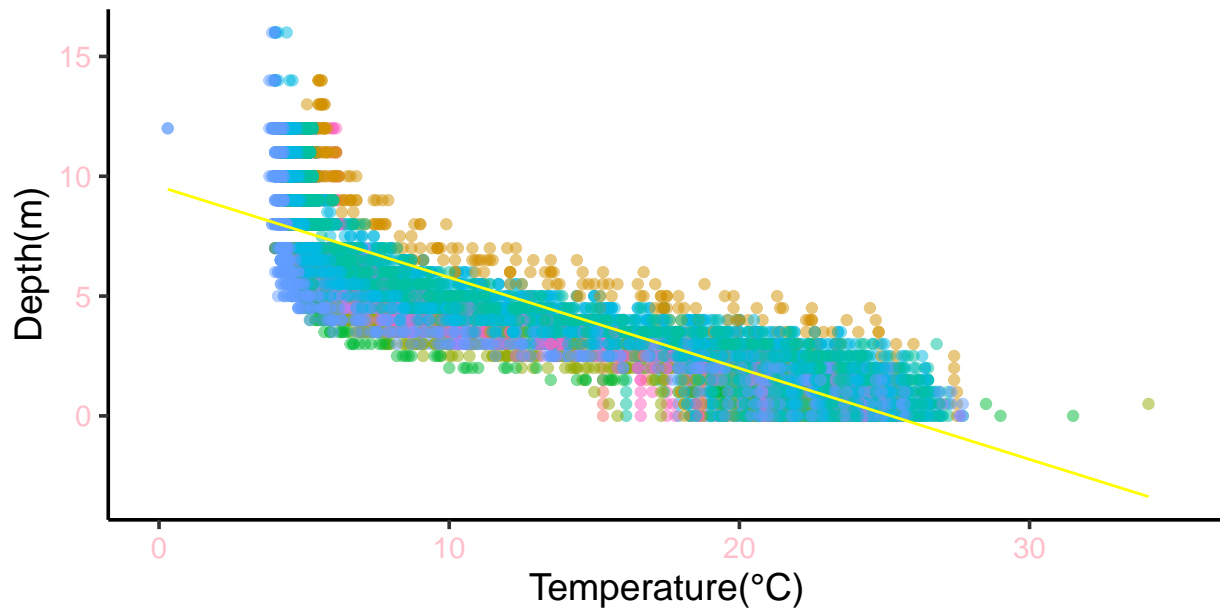
#For July
JulyLakeTemperatureDepth.NTL.LTER <-
  ggplot(July.NTL.LTER, aes(x = temperature_C, y=depth)) +
  geom_point(aes(color=lakename),alpha=0.5)+
  xlim(0, 35) +
  geom_smooth(method = "lm", se = FALSE,color="yellow",size=0.5)+
  xlab("Temperature(°C)")+
  ylab("Depth(m)")+
  ggtitle("Lakes Temperature Change by Depth")
print(JulyLakeTemperatureDepth.NTL.LTER)
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

Lakes Temperature Change by Depth

ne

- Central Long Lake
- East Long Lake
- Paul Lake
- Tuesday Lake
- Crampton Lake
- Hummingbird Lake
- Peter Lake
- Ward Lake



```
#For All Months
New.NTL.LTER <-
  Raw.NTL.LTER %>%
    select(lakename, depth, temperature_C) %>%
    na.omit()

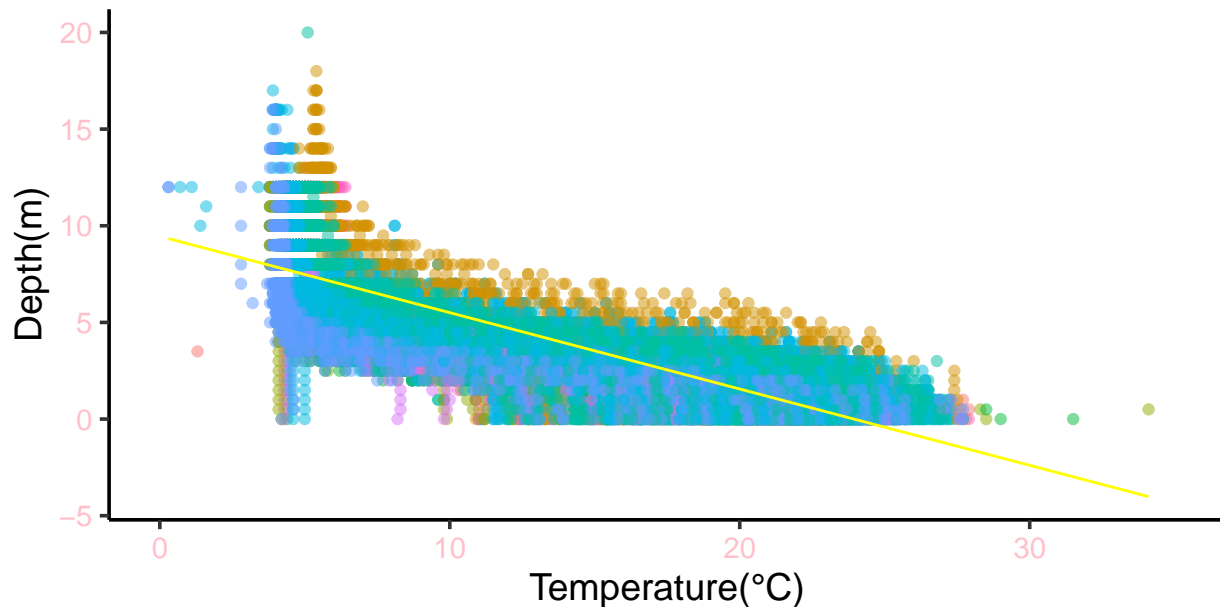
AllMonthsLakeTemperatureDepth.NTL.LTER <-
  ggplot(New.NTL.LTER, aes(x = temperature_C, y=depth)) +
    geom_point(aes(color=lakename),alpha=0.5)+
    xlim(0, 35) +
    geom_smooth(method = "lm", se = FALSE,color="yellow",size=0.5)+
    xlab("Temperature(°C)")+
    ylab("Depth(m)")+
    ggtitle("Lakes Temperature Change by Depth")
print(AllMonthsLakeTemperatureDepth.NTL.LTER)
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

Lakes Temperature Change by Depth

ne

- Central Long Lake
- East Long Lake
- Paul Lake
- Tuesday Lake
-
- Crampton Lake
- Hummingbird Lake
- Peter Lake
- Ward Lake



15. Use the Tukey's HSD test to determine which lakes have different means.

```
#15
#For July
JulyDepth.anova <- aov(data = July.NTL.LTER, July.NTL.LTER$temperature_C ~ July.NTL.LTER$lakename)
summary(JulyDepth.anova)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## July.NTL.LTER$lakename      8  21642   2705.2    50 <2e-16 ***
## Residuals                9719  525813     54.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(JulyDepth.anova)
```

```
## Analysis of Variance Table
##
## Response: July.NTL.LTER$temperature_C
##              Df Sum Sq Mean Sq F value    Pr(>F)
## July.NTL.LTER$lakename      8  21642   2705.2  50.003 < 2.2e-16 ***
## Residuals                9719  525813     54.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
JulyDepth.anova.lm <- lm(data = July.NTL.LTER, July.NTL.LTER$temperature_C ~ July.NTL.LTER$lakename)
summary(JulyDepth.anova.lm)
```

```
##
## Call:
## lm(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$lakename,
##     data = July.NTL.LTER)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.769  -6.614  -2.679   7.684  23.832
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   17.6664     0.6501  27.174 < 2e-16 ***
## July.NTL.LTER$lakenameCrampton Lake    -2.3145     0.7699  -3.006 0.002653 **
## July.NTL.LTER$lakenameEast Long Lake   -7.3987     0.6918 -10.695 < 2e-16 ***
## July.NTL.LTER$lakenameHummingbird Lake -6.8931     0.9429  -7.311 2.87e-13 ***
## July.NTL.LTER$lakenamePaul Lake        -3.8522     0.6656  -5.788 7.36e-09 ***
## July.NTL.LTER$lakenamePeter Lake       -4.3501     0.6645  -6.547 6.17e-11 ***
## July.NTL.LTER$lakenameTuesday Lake    -6.5972     0.6769  -9.746 < 2e-16 ***
## July.NTL.LTER$lakenameWard Lake        -3.2078     0.9429  -3.402 0.000672 ***
## July.NTL.LTER$lakenameWest Long Lake  -6.0878     0.6895  -8.829 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.355 on 9719 degrees of freedom
## Multiple R-squared:  0.03953,    Adjusted R-squared:  0.03874
## F-statistic:    50 on 8 and 9719 DF,  p-value: < 2.2e-16
```

```
anova(JulyDepth.anova.lm)
```

```
## Analysis of Variance Table
##
## Response: July.NTL.LTER$temperature_C
##              Df Sum Sq Mean Sq F value    Pr(>F)
## July.NTL.LTER$lakename      8  21642   2705.2  50.003 < 2.2e-16 ***
## Residuals                 9719  525813    54.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(JulyDepth.anova)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = July.NTL.LTER$temperature_C ~ July.NTL.LTER$lakename, data = July.NTL.LTER)
##
## $'July.NTL.LTER$lakename'
##              diff            lwr            upr      p adj
## Crampton Lake-Central Long Lake -2.3145195 -4.7031913  0.0741524 0.0661566
```

```
## East Long Lake-Central Long Lake -7.3987410 -9.5449411 -5.2525408 0.0000000
## Hummingbird Lake-Central Long Lake -6.8931304 -9.8184178 -3.9678430 0.0000000
## Paul Lake-Central Long Lake -3.8521506 -5.9170942 -1.7872070 0.0000003
## Peter Lake-Central Long Lake -4.3501458 -6.4115874 -2.2887042 0.0000000
## Tuesday Lake-Central Long Lake -6.5971805 -8.6971605 -4.4972005 0.0000000
## Ward Lake-Central Long Lake -3.2077856 -6.1330730 -0.2824982 0.0193405
## West Long Lake-Central Long Lake -6.0877513 -8.2268550 -3.9486475 0.0000000
## East Long Lake-Crampton Lake -5.0842215 -6.5591700 -3.6092730 0.0000000
## Hummingbird Lake-Crampton Lake -4.5786109 -7.0538088 -2.1034131 0.0000004
## Paul Lake-Crampton Lake -1.5376312 -2.8916215 -0.1836408 0.0127491
## Peter Lake-Crampton Lake -2.0356263 -3.3842699 -0.6869828 0.0000999
## Tuesday Lake-Crampton Lake -4.2826611 -5.6895065 -2.8758157 0.0000000
## Ward Lake-Crampton Lake -0.8932661 -3.3684639 1.5819317 0.9714459
## West Long Lake-Crampton Lake -3.7732318 -5.2378351 -2.3086285 0.0000000
## Hummingbird Lake-East Long Lake 0.5056106 -1.7364925 2.7477137 0.9988050
## Paul Lake-East Long Lake 3.5465903 2.6900206 4.4031601 0.0000000
## Peter Lake-East Long Lake 3.0485952 2.2005025 3.8966879 0.0000000
## Tuesday Lake-East Long Lake 0.8015604 -0.1363286 1.7394495 0.1657485
## Ward Lake-East Long Lake 4.1909554 1.9488523 6.4330585 0.0000002
## West Long Lake-East Long Lake 1.3109897 0.2885003 2.3334791 0.0022805
## Paul Lake-Hummingbird Lake 3.0409798 0.8765299 5.2054296 0.0004495
## Peter Lake-Hummingbird Lake 2.5429846 0.3818755 4.7040937 0.0080666
## Tuesday Lake-Hummingbird Lake 0.2959499 -1.9019508 2.4938505 0.9999752
## Ward Lake-Hummingbird Lake 3.6853448 0.6889874 6.6817022 0.0043297
## West Long Lake-Hummingbird Lake 0.8053791 -1.4299320 3.0406903 0.9717297
## Peter Lake-Paul Lake -0.4979952 -1.1120620 0.1160717 0.2241586
## Tuesday Lake-Paul Lake -2.7450299 -3.4781416 -2.0119182 0.0000000
## Ward Lake-Paul Lake 0.6443651 -1.5200848 2.8088149 0.9916978
## West Long Lake-Paul Lake -2.2356007 -3.0742314 -1.3969699 0.0000000
## Tuesday Lake-Peter Lake -2.2470347 -2.9702236 -1.5238458 0.0000000
## Ward Lake-Peter Lake 1.1423602 -1.0187489 3.3034693 0.7827037
## West Long Lake-Peter Lake -1.7376055 -2.5675759 -0.9076350 0.0000000
## Ward Lake-Tuesday Lake 3.3893950 1.1914943 5.5872956 0.0000609
## West Long Lake-Tuesday Lake 0.5094292 -0.4121051 1.4309636 0.7374387
## West Long Lake-Ward Lake -2.8799657 -5.1152769 -0.6446546 0.0021080
```

```
#For All Months
```

```
All.anova <- aov(data = New.NTL.LTER, New.NTL.LTER$temperature_C ~ New.NTL.LTER$lakename)
summary(All.anova)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## New.NTL.LTER$lakename      8   57921    7240  155.7 <2e-16 ***
## Residuals          34747 1615571      46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(All.anova)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: New.NTL.LTER$temperature_C
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## New.NTL.LTER$lakename      8   57921    7240.1  155.72 < 2.2e-16 ***
```

```
## Residuals          34747 1615571    46.5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

All.anova.lm <- lm(data = New.NTL.LTER, New.NTL.LTER$temperature_C ~ New.NTL.LTER$lakename)
summary(All.anova.lm)

##
## Call:
## lm(formula = New.NTL.LTER$temperature_C ~ New.NTL.LTER$lakename,
##     data = New.NTL.LTER)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.436  -5.959  -2.559   6.549  24.321
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   16.7363     0.3240   51.660 < 2e-16 ***
## New.NTL.LTER$lakenameCrampton Lake    -2.5443     0.3833   -6.638 3.23e-11 ***
## New.NTL.LTER$lakenameEast Long Lake   -6.9570     0.3436  -20.248 < 2e-16 ***
## New.NTL.LTER$lakenameHummingbird Lake -6.6985     0.4775  -14.030 < 2e-16 ***
## New.NTL.LTER$lakenamePaul Lake        -3.9441     0.3316  -11.893 < 2e-16 ***
## New.NTL.LTER$lakenamePeter Lake       -4.4838     0.3309  -13.549 < 2e-16 ***
## New.NTL.LTER$lakenameTuesday Lake     -6.3896     0.3368  -18.974 < 2e-16 ***
## New.NTL.LTER$lakenameWard Lake        -4.3083     0.4395   -9.802 < 2e-16 ***
## New.NTL.LTER$lakenameWest Long Lake   -5.6778     0.3423  -16.587 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.819 on 34747 degrees of freedom
## Multiple R-squared:  0.03461,    Adjusted R-squared:  0.03439
## F-statistic: 155.7 on 8 and 34747 DF,  p-value: < 2.2e-16

anova(All.anova.lm)

## Analysis of Variance Table
##
## Response: New.NTL.LTER$temperature_C
##              Df Sum Sq Mean Sq F value    Pr(>F)
## New.NTL.LTER$lakename      8  57921  7240.1  155.72 < 2.2e-16 ***
## Residuals                 34747 1615571    46.5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

TukeyHSD(All.anova)

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = New.NTL.LTER$temperature_C ~ New.NTL.LTER$lakename, data = New.NTL.LTER)
##
```



```
## $'New.NTL.LTER$lakename'
##               diff          lwr          upr          p adj
## Crampton Lake-Central Long Lake -2.5442854 -3.7331780 -1.3553927 0.0000000
## East Long Lake-Central Long Lake -6.9570473 -8.0227648 -5.8913298 0.0000000
## Hummingbird Lake-Central Long Lake -6.6985124 -8.1794348 -5.2175900 0.0000000
## Paul Lake-Central Long Lake -3.9440682 -4.9727040 -2.9154324 0.0000000
## Peter Lake-Central Long Lake -4.4837864 -5.5102613 -3.4573116 0.0000000
## Tuesday Lake-Central Long Lake -6.3896413 -7.4341675 -5.3451152 0.0000000
## Ward Lake-Central Long Lake -4.3082596 -5.6715463 -2.9449730 0.0000000
## West Long Lake-Central Long Lake -5.6777623 -6.7395105 -4.6160141 0.0000000
## East Long Lake-Crampton Lake -4.4127620 -5.1405823 -3.6849417 0.0000000
## Hummingbird Lake-Crampton Lake -4.1542271 -5.4140285 -2.8944256 0.0000000
## Paul Lake-Crampton Lake -1.3997828 -2.0721371 -0.7274286 0.0000000
## Peter Lake-Crampton Lake -1.9395011 -2.6085446 -1.2704575 0.0000000
## Tuesday Lake-Crampton Lake -3.8453560 -4.5417779 -3.1489341 0.0000000
## Ward Lake-Crampton Lake -1.7639743 -2.8831342 -0.6448143 0.0000357
## West Long Lake-Crampton Lake -3.1334769 -3.8554727 -2.4114812 0.0000000
## Hummingbird Lake-East Long Lake 0.2585349 -0.8857499 1.4028198 0.9987916
## Paul Lake-East Long Lake 3.0129792 2.5954288 3.4305296 0.0000000
## Peter Lake-East Long Lake 2.4732609 2.0610627 2.8854591 0.0000000
## Tuesday Lake-East Long Lake 0.5674060 0.1121132 1.0226989 0.0035472
## Ward Lake-East Long Lake 2.6487877 1.6614645 3.6361109 0.0000000
## West Long Lake-East Long Lake 1.2792850 0.7857610 1.7728091 0.0000000
## Paul Lake-Hummingbird Lake 2.7544443 1.6446129 3.8642756 0.0000000
## Peter Lake-Hummingbird Lake 2.2147260 1.1068972 3.3225548 0.0000000
## Tuesday Lake-Hummingbird Lake 0.3088711 -0.8157039 1.4334461 0.9952041
## Ward Lake-Hummingbird Lake 2.3902528 0.9647057 3.8157999 0.0000071
## West Long Lake-Hummingbird Lake 1.0207501 -0.1198389 2.1613391 0.1224797
## Peter Lake-Paul Lake -0.5397183 -0.8434372 -0.2359993 0.0000013
## Tuesday Lake-Paul Lake -2.4455731 -2.8056140 -2.0855323 0.0000000
## Ward Lake-Paul Lake -0.3641914 -1.3113688 0.5829859 0.9582889
## West Long Lake-Paul Lake -1.7336941 -2.1410071 -1.3263812 0.0000000
## Tuesday Lake-Peter Lake -1.9058549 -2.2596746 -1.5520351 0.0000000
## Ward Lake-Peter Lake 0.1755268 -0.7693033 1.1203570 0.9997136
## West Long Lake-Peter Lake -1.1939759 -1.5958003 -0.7921515 0.0000000
## Ward Lake-Tuesday Lake 2.0813817 1.1169709 3.0457925 0.0000000
## West Long Lake-Tuesday Lake 0.7118790 0.2659563 1.1578017 0.0000259
## West Long Lake-Ward Lake -1.3695027 -2.3525401 -0.3864652 0.0005266
```

16. From the findings above, which lakes have the same mean temperature, statistically speaking, as Peter Lake? Does any lake have a mean temperature that is statistically distinct from all the other lakes?

Answer:

#For July Peter Lake-Paul Lake -0.4979952°C difference Ward Lake-Peter Lake 0.1755268°C difference

Maybe Ward Lake-East Long Lake 4.1909554°C difference

#For All Months Peter Lake-Paul Lake -0.5397183°C difference Ward Lake-Peter Lake 0.1755268°C difference

Maybe Ward Lake-Central Long Lake -4.3082596°C difference

17. If we were just looking at Peter Lake and Paul Lake. What's another test we might explore to see whether they have distinct mean temperatures?

Answer:

Two Sample T-Test

18. Wrangle the July data to include only records for Crampton Lake and Ward Lake. Run the two-sample T-test on these data to determine whether their July temperature are same or different. What does the test say? Are the mean temperatures for the lakes equal? Does that match you answer for part 16?

```
CramptonWard.NTL.LTER <-  
Raw.NTL.LTER %>%  
  select(lakename, year4, daynum, depth, temperature_C) %>%  
  filter(month(Raw.NTL.LTER$sampleddate) %in% 7, lakename == "Ward Lake"|lakename == "Crampton Lake") %>%  
  na.omit()  
print(CramptonWard.NTL.LTER)
```

##		lakename	year4	daynum	depth	temperature_C
## 1		Crampton Lake	1999	196	0.00	22.8
## 3		Crampton Lake	1999	196	0.50	22.6
## 5		Crampton Lake	1999	196	1.00	22.4
## 6		Crampton Lake	1999	196	1.50	22.2
## 7		Crampton Lake	1999	196	2.00	22.0
## 8		Crampton Lake	1999	196	2.50	21.9
## 9		Crampton Lake	1999	196	3.00	21.7
## 10		Crampton Lake	1999	196	3.50	21.3
## 11		Crampton Lake	1999	196	4.00	20.0
## 12		Crampton Lake	1999	196	4.50	17.1
## 13		Crampton Lake	1999	196	5.00	14.4
## 14		Crampton Lake	1999	196	5.50	12.1
## 15		Crampton Lake	1999	196	6.00	10.2
## 16		Crampton Lake	1999	196	6.50	9.1
## 17		Crampton Lake	1999	196	7.00	8.4
## 18		Crampton Lake	1999	196	8.00	7.1
## 19		Crampton Lake	1999	196	9.00	6.3
## 20		Crampton Lake	1999	196	10.00	5.8
## 21		Crampton Lake	1999	196	11.00	5.5
## 22		Crampton Lake	1999	196	12.00	5.4
## 23		Crampton Lake	2003	183	0.00	21.8
## 25		Crampton Lake	2003	183	0.50	21.9
## 27		Crampton Lake	2003	183	1.00	21.9
## 28		Crampton Lake	2003	183	1.50	21.8
## 29		Crampton Lake	2003	183	2.00	21.8
## 30		Crampton Lake	2003	183	2.50	21.3
## 31		Crampton Lake	2003	183	3.00	20.8
## 32		Crampton Lake	2003	183	3.50	19.9
## 33		Crampton Lake	2003	183	4.00	17.2
## 34		Crampton Lake	2003	183	4.50	15.2
## 35		Crampton Lake	2003	183	5.00	12.7
## 36		Crampton Lake	2003	183	5.50	11.2
## 37		Crampton Lake	2003	183	6.00	10.2
## 38		Crampton Lake	2003	183	6.50	8.8
## 39		Crampton Lake	2003	183	7.00	8.2
## 41		Crampton Lake	2003	183	8.00	6.8

## 42	Crampton Lake	2003	183	9.00	6.0
## 43	Crampton Lake	2003	183	10.00	5.4
## 44	Crampton Lake	2003	183	11.00	5.1
## 45	Crampton Lake	2003	183	12.00	5.0
## 46	Crampton Lake	2003	197	0.00	22.2
## 48	Crampton Lake	2003	197	0.50	21.9
## 50	Crampton Lake	2003	197	1.00	21.8
## 51	Crampton Lake	2003	197	1.50	21.8
## 52	Crampton Lake	2003	197	2.00	21.7
## 53	Crampton Lake	2003	197	2.50	21.7
## 54	Crampton Lake	2003	197	3.00	21.7
## 55	Crampton Lake	2003	197	3.50	21.6
## 56	Crampton Lake	2003	197	4.00	19.6
## 57	Crampton Lake	2003	197	4.50	17.3
## 58	Crampton Lake	2003	197	5.00	14.5
## 59	Crampton Lake	2003	197	5.50	12.5
## 60	Crampton Lake	2003	197	6.00	10.9
## 61	Crampton Lake	2003	197	6.50	9.8
## 62	Crampton Lake	2003	197	7.00	8.7
## 63	Crampton Lake	2003	197	8.00	7.0
## 64	Crampton Lake	2003	197	9.00	6.0
## 65	Crampton Lake	2003	197	10.00	5.5
## 66	Crampton Lake	2003	197	11.00	5.2
## 67	Crampton Lake	2003	197	12.00	5.1
## 68	Crampton Lake	2003	211	0.00	23.2
## 70	Crampton Lake	2003	211	0.50	23.3
## 72	Crampton Lake	2003	211	1.00	23.2
## 73	Crampton Lake	2003	211	1.50	23.3
## 74	Crampton Lake	2003	211	2.00	23.3
## 75	Crampton Lake	2003	211	2.50	23.3
## 76	Crampton Lake	2003	211	3.00	23.3
## 77	Crampton Lake	2003	211	3.50	23.2
## 78	Crampton Lake	2003	211	4.00	22.4
## 79	Crampton Lake	2003	211	4.50	20.6
## 80	Crampton Lake	2003	211	5.00	17.3
## 81	Crampton Lake	2003	211	5.50	14.6
## 82	Crampton Lake	2003	211	6.00	12.1
## 83	Crampton Lake	2003	211	6.50	11.0
## 84	Crampton Lake	2003	211	7.00	9.6
## 85	Crampton Lake	2003	211	8.00	7.4
## 86	Crampton Lake	2003	211	9.00	6.4
## 87	Crampton Lake	2003	211	10.00	5.7
## 88	Crampton Lake	2003	211	11.00	5.4
## 89	Crampton Lake	2003	211	12.00	5.2
## 90	Crampton Lake	2004	183	0.00	19.8
## 92	Crampton Lake	2004	183	0.50	19.8
## 94	Crampton Lake	2004	183	1.00	19.8
## 95	Crampton Lake	2004	183	1.50	19.8
## 96	Crampton Lake	2004	183	2.00	19.7
## 97	Crampton Lake	2004	183	2.50	19.5
## 98	Crampton Lake	2004	183	3.00	18.9
## 99	Crampton Lake	2004	183	3.50	18.4
## 100	Crampton Lake	2004	183	4.00	18.0
## 101	Crampton Lake	2004	183	4.50	17.5

##	102	Crampton Lake	2004	183	5.00	15.6
##	103	Crampton Lake	2004	183	5.50	13.4
##	104	Crampton Lake	2004	183	6.00	11.2
##	105	Crampton Lake	2004	183	6.50	10.2
##	106	Crampton Lake	2004	183	7.00	9.0
##	108	Crampton Lake	2004	183	8.00	7.6
##	109	Crampton Lake	2004	183	9.00	6.6
##	110	Crampton Lake	2004	183	10.00	6.0
##	111	Crampton Lake	2004	183	11.00	5.6
##	112	Crampton Lake	2004	183	12.00	5.5
##	113	Crampton Lake	2004	183	13.00	5.5
##	114	Crampton Lake	2004	183	14.00	5.5
##	115	Crampton Lake	2004	191	0.00	18.1
##	117	Crampton Lake	2004	191	0.50	18.1
##	119	Crampton Lake	2004	191	1.00	18.0
##	120	Crampton Lake	2004	191	1.50	18.0
##	121	Crampton Lake	2004	191	2.00	17.9
##	122	Crampton Lake	2004	191	2.50	17.9
##	123	Crampton Lake	2004	191	3.00	17.9
##	124	Crampton Lake	2004	191	3.50	17.8
##	125	Crampton Lake	2004	191	4.00	17.6
##	126	Crampton Lake	2004	191	4.50	17.4
##	127	Crampton Lake	2004	191	5.00	16.6
##	128	Crampton Lake	2004	191	5.50	13.8
##	129	Crampton Lake	2004	191	6.00	12.1
##	130	Crampton Lake	2004	191	6.50	10.9
##	131	Crampton Lake	2004	191	7.00	9.4
##	132	Crampton Lake	2004	191	8.00	7.6
##	133	Crampton Lake	2004	191	9.00	6.6
##	134	Crampton Lake	2004	191	10.00	6.0
##	135	Crampton Lake	2004	191	11.00	5.7
##	136	Crampton Lake	2004	191	12.00	5.6
##	137	Crampton Lake	2004	191	13.00	5.5
##	138	Crampton Lake	2004	191	14.00	5.5
##	139	Crampton Lake	2004	197	0.00	21.8
##	141	Crampton Lake	2004	197	0.50	21.8
##	143	Crampton Lake	2004	197	1.00	21.6
##	144	Crampton Lake	2004	197	1.50	21.6
##	145	Crampton Lake	2004	197	2.00	21.5
##	146	Crampton Lake	2004	197	2.50	20.9
##	147	Crampton Lake	2004	197	3.00	19.5
##	148	Crampton Lake	2004	197	3.50	19.1
##	149	Crampton Lake	2004	197	4.00	18.6
##	150	Crampton Lake	2004	197	4.50	17.9
##	151	Crampton Lake	2004	197	5.00	16.7
##	152	Crampton Lake	2004	197	5.50	14.8
##	153	Crampton Lake	2004	197	6.00	12.8
##	154	Crampton Lake	2004	197	6.50	11.1
##	155	Crampton Lake	2004	197	7.00	9.7
##	156	Crampton Lake	2004	197	8.00	7.7
##	157	Crampton Lake	2004	197	9.00	6.6
##	158	Crampton Lake	2004	197	10.00	6.1
##	159	Crampton Lake	2004	197	11.00	5.7
##	160	Crampton Lake	2004	197	12.00	5.6

##	161	Crampton Lake	2004	197	13.00	5.6
##	162	Crampton Lake	2004	197	14.00	5.5
##	163	Crampton Lake	2004	204	0.00	23.8
##	165	Crampton Lake	2004	204	0.50	23.8
##	167	Crampton Lake	2004	204	1.00	23.8
##	168	Crampton Lake	2004	204	1.50	23.8
##	169	Crampton Lake	2004	204	2.00	23.8
##	170	Crampton Lake	2004	204	2.50	23.2
##	171	Crampton Lake	2004	204	3.00	22.0
##	172	Crampton Lake	2004	204	3.50	20.5
##	173	Crampton Lake	2004	204	4.00	19.9
##	174	Crampton Lake	2004	204	4.50	18.5
##	175	Crampton Lake	2004	204	5.00	17.3
##	176	Crampton Lake	2004	204	5.50	15.5
##	177	Crampton Lake	2004	204	6.00	13.5
##	178	Crampton Lake	2004	204	6.50	11.3
##	179	Crampton Lake	2004	204	7.00	10.1
##	180	Crampton Lake	2004	204	8.00	7.8
##	181	Crampton Lake	2004	204	9.00	6.8
##	182	Crampton Lake	2004	204	10.00	6.2
##	183	Crampton Lake	2004	204	11.00	5.8
##	184	Crampton Lake	2004	204	12.00	5.7
##	185	Crampton Lake	2004	204	13.00	5.6
##	186	Crampton Lake	2004	204	14.00	5.6
##	187	Crampton Lake	2004	211	0.00	22.7
##	189	Crampton Lake	2004	211	0.50	22.7
##	191	Crampton Lake	2004	211	1.00	22.7
##	192	Crampton Lake	2004	211	1.50	22.7
##	193	Crampton Lake	2004	211	2.00	22.7
##	194	Crampton Lake	2004	211	2.50	22.7
##	195	Crampton Lake	2004	211	3.00	22.7
##	196	Crampton Lake	2004	211	3.50	22.6
##	197	Crampton Lake	2004	211	4.00	20.8
##	198	Crampton Lake	2004	211	4.50	18.1
##	199	Crampton Lake	2004	211	5.00	17.5
##	200	Crampton Lake	2004	211	5.50	15.8
##	201	Crampton Lake	2004	211	6.00	13.6
##	202	Crampton Lake	2004	211	6.50	11.5
##	203	Crampton Lake	2004	211	7.00	10.3
##	204	Crampton Lake	2004	211	8.00	7.8
##	205	Crampton Lake	2004	211	9.00	6.8
##	206	Crampton Lake	2004	211	10.00	6.2
##	207	Crampton Lake	2004	211	11.00	5.8
##	208	Crampton Lake	2004	211	12.00	5.7
##	209	Crampton Lake	2004	211	13.00	5.6
##	210	Crampton Lake	2004	211	14.00	5.6
##	211	Crampton Lake	2005	186	0.00	22.3
##	213	Crampton Lake	2005	186	0.50	22.3
##	215	Crampton Lake	2005	186	1.00	22.4
##	216	Crampton Lake	2005	186	1.50	22.4
##	217	Crampton Lake	2005	186	2.00	22.4
##	218	Crampton Lake	2005	186	2.50	22.4
##	219	Crampton Lake	2005	186	3.00	22.4
##	220	Crampton Lake	2005	186	3.50	22.3

##	221	Crampton Lake	2005	186	4.00	22.3
##	222	Crampton Lake	2005	186	4.50	21.5
##	223	Crampton Lake	2005	186	5.00	17.9
##	224	Crampton Lake	2005	186	5.50	15.3
##	225	Crampton Lake	2005	186	6.00	13.5
##	226	Crampton Lake	2005	186	6.50	12.3
##	227	Crampton Lake	2005	186	7.00	10.8
##	229	Crampton Lake	2005	186	8.00	8.8
##	231	Crampton Lake	2005	186	9.00	7.4
##	232	Crampton Lake	2005	186	10.00	6.0
##	233	Crampton Lake	2005	186	11.00	5.8
##	234	Crampton Lake	2005	186	12.00	5.7
##	235	Crampton Lake	2005	192	0.00	25.7
##	237	Crampton Lake	2005	192	0.50	25.7
##	239	Crampton Lake	2005	192	1.00	25.4
##	240	Crampton Lake	2005	192	1.50	25.4
##	241	Crampton Lake	2005	192	2.00	25.3
##	242	Crampton Lake	2005	192	2.50	24.2
##	243	Crampton Lake	2005	192	3.00	23.5
##	244	Crampton Lake	2005	192	3.50	22.9
##	245	Crampton Lake	2005	192	4.00	22.5
##	246	Crampton Lake	2005	192	4.50	21.4
##	247	Crampton Lake	2005	192	5.00	19.5
##	248	Crampton Lake	2005	192	5.50	16.5
##	249	Crampton Lake	2005	192	6.00	14.6
##	250	Crampton Lake	2005	192	6.50	12.9
##	251	Crampton Lake	2005	192	7.00	11.4
##	253	Crampton Lake	2005	192	8.00	9.0
##	255	Crampton Lake	2005	192	9.00	7.5
##	256	Crampton Lake	2005	192	10.00	6.5
##	257	Crampton Lake	2005	192	11.00	6.0
##	258	Crampton Lake	2005	192	12.00	5.7
##	259	Crampton Lake	2005	192	13.00	5.6
##	260	Crampton Lake	2005	199	0.00	27.5
##	262	Crampton Lake	2005	199	0.50	27.5
##	264	Crampton Lake	2005	199	1.00	27.5
##	265	Crampton Lake	2005	199	1.50	27.4
##	266	Crampton Lake	2005	199	2.00	27.4
##	267	Crampton Lake	2005	199	2.50	27.4
##	268	Crampton Lake	2005	199	3.00	26.0
##	269	Crampton Lake	2005	199	3.50	24.8
##	270	Crampton Lake	2005	199	4.00	23.7
##	271	Crampton Lake	2005	199	4.50	22.5
##	272	Crampton Lake	2005	199	5.00	19.8
##	273	Crampton Lake	2005	199	5.50	17.1
##	274	Crampton Lake	2005	199	6.00	15.3
##	275	Crampton Lake	2005	199	6.50	13.5
##	276	Crampton Lake	2005	199	7.00	12.0
##	278	Crampton Lake	2005	199	8.00	9.0
##	280	Crampton Lake	2005	199	9.00	7.7
##	281	Crampton Lake	2005	199	10.00	6.6
##	282	Crampton Lake	2005	199	11.00	6.1
##	283	Crampton Lake	2005	199	12.00	5.7
##	284	Crampton Lake	2005	199	13.00	5.7

##	285	Crampton Lake	2005	206	0.00	25.3
##	287	Crampton Lake	2005	206	0.50	25.1
##	289	Crampton Lake	2005	206	1.00	25.0
##	290	Crampton Lake	2005	206	1.50	24.9
##	291	Crampton Lake	2005	206	2.00	24.9
##	292	Crampton Lake	2005	206	2.50	24.9
##	293	Crampton Lake	2005	206	3.00	24.8
##	294	Crampton Lake	2005	206	3.50	24.8
##	295	Crampton Lake	2005	206	4.00	24.7
##	296	Crampton Lake	2005	206	4.50	24.5
##	297	Crampton Lake	2005	206	5.00	21.3
##	298	Crampton Lake	2005	206	5.50	18.8
##	299	Crampton Lake	2005	206	6.00	16.7
##	300	Crampton Lake	2005	206	6.50	14.4
##	301	Crampton Lake	2005	206	7.00	12.3
##	303	Crampton Lake	2005	206	8.00	9.9
##	305	Crampton Lake	2005	206	9.00	7.9
##	306	Crampton Lake	2005	206	10.00	6.8
##	307	Crampton Lake	2005	206	11.00	6.1
##	308	Crampton Lake	2005	206	12.00	5.8
##	309	Crampton Lake	2005	206	13.00	5.7
##	310	Crampton Lake	2006	186	0.00	22.7
##	312	Crampton Lake	2006	186	0.50	22.8
##	314	Crampton Lake	2006	186	1.00	22.8
##	315	Crampton Lake	2006	186	1.50	22.8
##	316	Crampton Lake	2006	186	2.00	22.8
##	317	Crampton Lake	2006	186	2.50	22.5
##	318	Crampton Lake	2006	186	3.00	22.6
##	319	Crampton Lake	2006	186	3.50	22.6
##	320	Crampton Lake	2006	186	4.00	20.3
##	321	Crampton Lake	2006	186	4.50	17.4
##	322	Crampton Lake	2006	186	5.00	15.2
##	323	Crampton Lake	2006	186	5.50	13.7
##	324	Crampton Lake	2006	186	6.00	12.1
##	325	Crampton Lake	2006	186	6.50	10.8
##	326	Crampton Lake	2006	186	7.00	9.6
##	327	Crampton Lake	2006	186	7.50	8.3
##	328	Crampton Lake	2006	186	8.00	7.4
##	329	Crampton Lake	2006	186	8.50	6.3
##	330	Crampton Lake	2006	186	9.00	6.0
##	331	Crampton Lake	2006	186	10.00	5.5
##	332	Crampton Lake	2006	186	11.00	5.2
##	333	Crampton Lake	2006	186	12.00	5.1
##	334	Crampton Lake	2006	200	0.00	25.4
##	336	Crampton Lake	2006	200	0.50	25.4
##	338	Crampton Lake	2006	200	1.00	25.4
##	339	Crampton Lake	2006	200	1.50	25.4
##	340	Crampton Lake	2006	200	2.00	25.4
##	341	Crampton Lake	2006	200	2.50	25.4
##	342	Crampton Lake	2006	200	3.00	25.4
##	343	Crampton Lake	2006	200	3.50	24.1
##	344	Crampton Lake	2006	200	4.00	22.5
##	345	Crampton Lake	2006	200	4.50	19.8
##	346	Crampton Lake	2006	200	5.00	17.5

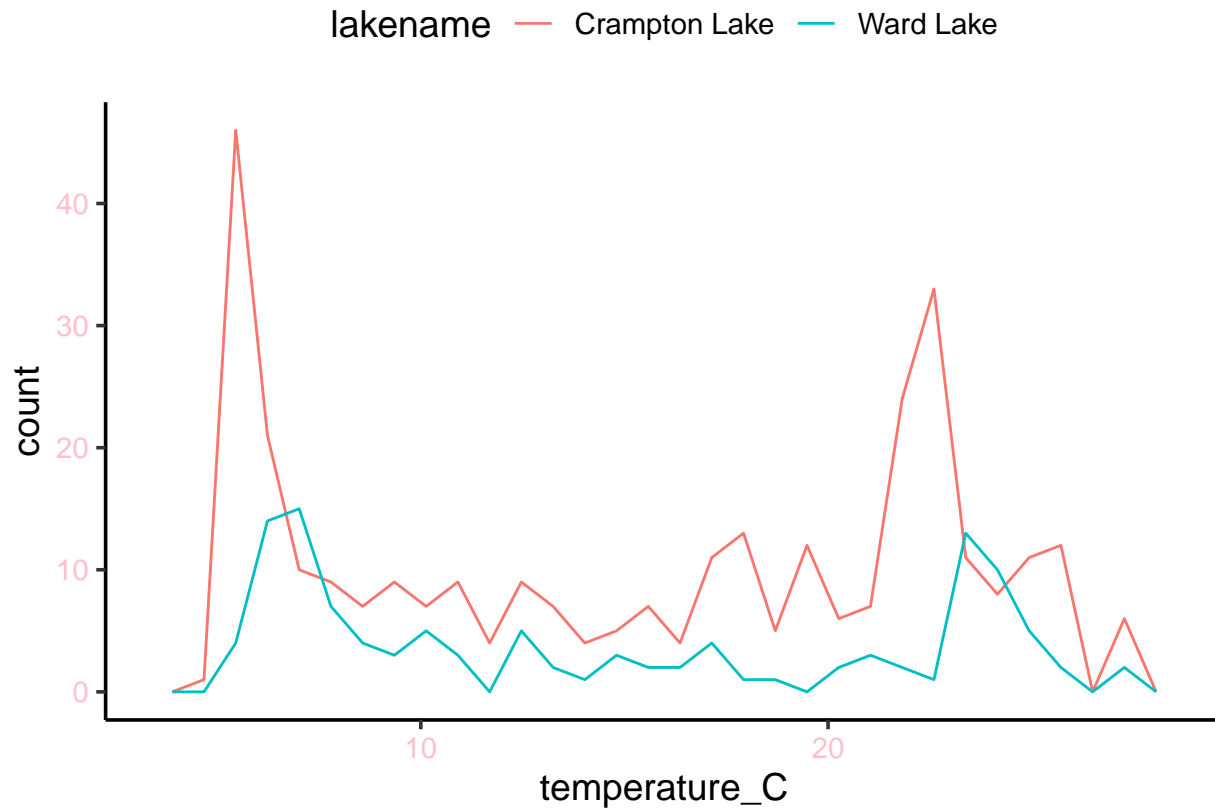
## 347	Crampton Lake	2006	200	5.50	15.0
## 348	Crampton Lake	2006	200	6.00	13.3
## 349	Crampton Lake	2006	200	6.50	11.4
## 350	Crampton Lake	2006	200	7.00	9.6
## 351	Crampton Lake	2006	200	7.50	8.7
## 352	Crampton Lake	2006	200	8.00	7.5
## 353	Crampton Lake	2006	200	8.50	6.7
## 354	Crampton Lake	2006	200	9.00	6.4
## 355	Crampton Lake	2006	200	10.00	5.7
## 356	Crampton Lake	2006	200	11.00	5.4
## 357	Crampton Lake	2006	200	12.00	5.1
## 358	Crampton Lake	2006	200	13.00	5.1
## 359	Ward Lake	2010	188	0.00	25.5
## 361	Ward Lake	2010	188	0.50	25.3
## 362	Ward Lake	2010	188	0.75	24.9
## 363	Ward Lake	2010	188	1.00	24.8
## 364	Ward Lake	2010	188	1.50	23.7
## 365	Ward Lake	2010	188	2.00	21.8
## 366	Ward Lake	2010	188	2.50	18.6
## 367	Ward Lake	2010	188	3.00	16.6
## 368	Ward Lake	2010	188	3.50	14.6
## 369	Ward Lake	2010	188	4.00	12.5
## 370	Ward Lake	2010	188	4.50	10.4
## 371	Ward Lake	2010	188	5.00	9.0
## 372	Ward Lake	2010	188	5.50	7.7
## 373	Ward Lake	2010	188	6.00	7.1
## 374	Ward Lake	2010	188	6.50	6.9
## 375	Ward Lake	2010	188	6.75	6.7
## 376	Ward Lake	2010	195	0.00	24.1
## 378	Ward Lake	2010	195	0.50	24.1
## 380	Ward Lake	2010	195	1.00	24.0
## 381	Ward Lake	2010	195	1.50	23.9
## 382	Ward Lake	2010	195	2.00	23.2
## 383	Ward Lake	2010	195	2.50	20.3
## 384	Ward Lake	2010	195	3.00	17.4
## 385	Ward Lake	2010	195	3.50	14.7
## 386	Ward Lake	2010	195	4.00	12.5
## 387	Ward Lake	2010	195	4.50	10.6
## 388	Ward Lake	2010	195	5.00	8.8
## 389	Ward Lake	2010	195	5.50	7.7
## 390	Ward Lake	2010	195	6.00	7.2
## 391	Ward Lake	2010	195	6.50	7.0
## 392	Ward Lake	2010	202	0.00	23.9
## 394	Ward Lake	2010	202	0.50	23.8
## 396	Ward Lake	2010	202	1.00	23.7
## 397	Ward Lake	2010	202	1.50	23.3
## 398	Ward Lake	2010	202	2.00	23.1
## 399	Ward Lake	2010	202	2.50	20.7
## 400	Ward Lake	2010	202	3.00	17.0
## 401	Ward Lake	2010	202	3.50	15.1
## 402	Ward Lake	2010	202	4.00	12.6
## 403	Ward Lake	2010	202	4.50	10.7
## 404	Ward Lake	2010	202	5.00	9.1
## 405	Ward Lake	2010	202	5.50	7.9

## 406	Ward Lake	2010	202	6.00	7.3
## 407	Ward Lake	2010	202	6.50	7.1
## 408	Ward Lake	2010	202	7.00	7.1
## 409	Ward Lake	2010	209	0.00	23.6
## 411	Ward Lake	2010	209	0.50	23.6
## 413	Ward Lake	2010	209	1.00	23.5
## 414	Ward Lake	2010	209	1.50	23.4
## 415	Ward Lake	2010	209	2.00	23.1
## 416	Ward Lake	2010	209	2.50	21.0
## 417	Ward Lake	2010	209	3.00	18.2
## 418	Ward Lake	2010	209	3.50	15.7
## 419	Ward Lake	2010	209	4.00	13.4
## 420	Ward Lake	2010	209	4.50	11.1
## 421	Ward Lake	2010	209	5.00	9.2
## 422	Ward Lake	2010	209	5.50	8.0
## 423	Ward Lake	2010	209	6.00	7.4
## 424	Ward Lake	2010	209	6.50	7.2
## 425	Ward Lake	2010	209	6.75	7.0
## 426	Ward Lake	2012	187	0.00	27.6
## 428	Ward Lake	2012	187	0.50	27.0
## 430	Ward Lake	2012	187	1.00	24.6
## 431	Ward Lake	2012	187	1.50	20.4
## 432	Ward Lake	2012	187	2.00	15.8
## 433	Ward Lake	2012	187	2.50	12.3
## 434	Ward Lake	2012	187	3.00	10.2
## 435	Ward Lake	2012	187	3.50	7.9
## 436	Ward Lake	2012	187	4.00	6.9
## 437	Ward Lake	2012	187	4.50	6.4
## 438	Ward Lake	2012	187	5.00	6.1
## 439	Ward Lake	2012	187	5.50	5.8
## 440	Ward Lake	2012	187	6.00	5.7
## 441	Ward Lake	2012	194	0.00	25.9
## 443	Ward Lake	2012	194	0.50	25.1
## 445	Ward Lake	2012	194	1.00	24.5
## 446	Ward Lake	2012	194	1.50	20.8
## 447	Ward Lake	2012	194	2.00	16.5
## 448	Ward Lake	2012	194	2.50	12.6
## 449	Ward Lake	2012	194	3.00	10.3
## 450	Ward Lake	2012	194	3.50	8.2
## 451	Ward Lake	2012	194	4.00	7.2
## 452	Ward Lake	2012	194	4.50	6.6
## 453	Ward Lake	2012	194	5.00	6.3
## 454	Ward Lake	2012	194	5.50	6.0
## 455	Ward Lake	2012	194	6.00	5.9
## 456	Ward Lake	2012	201	0.00	23.6
## 458	Ward Lake	2012	201	0.50	23.3
## 460	Ward Lake	2012	201	1.00	23.2
## 461	Ward Lake	2012	201	1.50	21.5
## 462	Ward Lake	2012	201	2.00	16.9
## 463	Ward Lake	2012	201	2.50	13.2
## 464	Ward Lake	2012	201	3.00	10.3
## 465	Ward Lake	2012	201	3.50	8.6
## 466	Ward Lake	2012	201	4.00	7.7
## 467	Ward Lake	2012	201	4.50	6.8

```
## 468 Ward Lake 2012 201 5.00 6.4
## 469 Ward Lake 2012 201 5.50 6.2
## 470 Ward Lake 2012 201 6.00 6.0
## 471 Ward Lake 2012 201 6.50 5.9
## 472 Ward Lake 2012 201 7.00 5.8
## 473 Ward Lake 2012 208 0.00 24.3
## 475 Ward Lake 2012 208 0.50 24.2
## 477 Ward Lake 2012 208 1.00 24.2
## 478 Ward Lake 2012 208 1.50 22.3
## 479 Ward Lake 2012 208 2.00 17.5
## 480 Ward Lake 2012 208 2.50 13.7
## 481 Ward Lake 2012 208 3.00 10.5
## 482 Ward Lake 2012 208 3.50 8.5
## 483 Ward Lake 2012 208 4.00 7.6
## 484 Ward Lake 2012 208 4.50 6.8
## 485 Ward Lake 2012 208 5.00 6.4
## 486 Ward Lake 2012 208 5.50 6.1
## 487 Ward Lake 2012 208 6.00 6.0
## 488 Ward Lake 2012 208 6.50 5.9
## 489 Ward Lake 2012 208 7.00 5.8
```

```
ggplot(CramptonWard.NTL.LTER, aes(x = temperature_C, color = lakename)) +
  geom_freqpoly()
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



```
twosample <- t.test(CramptonWard.NTL.LTER$temperature_C ~ CramptonWard.NTL.LTER$lakename)
twosample
```

```
##
## Welch Two Sample t-test
##
## data: CramptonWard.NTL.LTER$temperature_C by CramptonWard.NTL.LTER$lakename
## t = 1.1181, df = 200.37, p-value = 0.2649
## alternative hypothesis: true difference in means between group Crampton Lake and group Ward Lake is not equal to 0
## 95 percent confidence interval:
## -0.6821129 2.4686451
## sample estimates:
## mean in group Crampton Lake mean in group Ward Lake
## 15.35189 14.45862
```

```
twosample2 <- lm(CramptonWard.NTL.LTER$temperature_C ~ CramptonWard.NTL.LTER$lakename)
summary(twosample2)
```

```
##
## Call:
## lm(formula = CramptonWard.NTL.LTER$temperature_C ~ CramptonWard.NTL.LTER$lakename)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.3519  -7.5286   0.1947   7.0481  13.1414
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      15.3519    0.4087   37.56 <2e-16
## CramptonWard.NTL.LTER$lakenameWard Lake  -0.8933    0.7906   -1.13  0.259
##
## (Intercept)          ***
## CramptonWard.NTL.LTER$lakenameWard Lake
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.289 on 432 degrees of freedom
## Multiple R-squared:  0.002946, Adjusted R-squared:  0.0006383
## F-statistic: 1.277 on 1 and 432 DF, p-value: 0.2592
```

Answer: The mean in group Crampton Lake is 15.35189, and the mean in group Ward Lake is 14.45862, which are really close but not exact. The difference in TukeyHSD(JulyDepth.anova) for these two lakes is 0.8932661 which is the same as the result from the two-sample T-test we performed.