

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

Jiawei Liang

## OVERVIEW

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

## Directions

1. Rename this file <FirstLast>\_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(tibbletime)

##
## Attaching package: 'tibbletime'
##
## The following object is masked from 'package:stats':
##
##   filter
```

```
library(lubridate)
```

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

```
library(dplyr)  
library(cowplot)
```

```
##  
## Attaching package: 'cowplot'  
##  
## The following object is masked from 'package:lubridate':  
##  
##    stamp
```

```
library(ggplot2)  
library(agricolae)  
library(formatR)  
getwd()
```

```
## [1] "C:/Users/Jiawei Liang/Documents/EDA-Fall2022/Assignments"
```

```
setwd("c:/Users/Jiawei Liang/Documents/EDA-Fall2022/Data/Raw/")  
NL_ChemistryPhysics <- read.csv("NTL-LTER_Lake_ChemistryPhysics_Raw.csv", stringsAsFactors = TRUE)  
NL_ChemistryPhysics$sampleddate <- as.Date(NL_ChemistryPhysics$sampleddate, format = "%m/%d/%Y")
```

```
# 2  
theme_default <- theme_set(theme_bw())  
theme_set(theme_default)
```

## 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: temperature changes with lake temperature Ha: temperature will not changes with lake temperature
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
class(NL_ChemistryPhysics$sampleddate)
```

```
## [1] "Date"
```

```
month(NL_ChemistryPhysics$sampleddate)
```

```
##      [1] 5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
##     [25] 5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
##    [49] 5  5  5  5  5  5  5  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
##   [73] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
##  [97] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [121] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [145] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [169] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [193] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [217] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [241] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [265] 6  6  6  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [289] 7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [313] 7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [337] 7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [361] 7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [385] 7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [409] 7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [433] 7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [457] 7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [481] 7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [505] 7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7
## [529] 7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  7  8  8  8  8  8
## [553] 8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8
## [577] 8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8
## [601] 8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8
## [625] 8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8
## [649] 8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8
## [673] 8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8
## [697] 8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8
## [721] 8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8
## [745] 8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  5  5
## [769] 5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## [793] 5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## [817] 5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## [841] 5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## [865] 5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5  5
## [889] 5  5  5  5  5  5  5  5  5  5  5  5  5  6  6  6  6  6  6  6  6  6
## [913] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [937] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [961] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [985] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [1009] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [1033] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
## [1057] 6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6  6
```

##	[1081]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[1105]	6	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	7	7	7	7	7
##	[1129]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[1153]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[1177]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[1201]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[1225]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[1249]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[1273]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[1297]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[1321]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[1345]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[1369]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[1393]	7	7	7	7	7	7	7	7	7	7	7	8	8	8	8	8	8	8	8	8	8
##	[1417]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[1441]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[1465]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[1489]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[1513]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[1537]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[1561]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[1585]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[1609]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	9	9
##	[1633]	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
##	[1657]	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
##	[1681]	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
##	[1705]	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
##	[1729]	9	9	9	9	9	9	9	9	9	9	9	9	9	5	5	5	5	5	5	5	5
##	[1753]	5	5	5																		

##	[2377]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[2401]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[2425]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[2449]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[2473]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[2497]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[2521]	8	8	8	8	8	8	8	8	8	9	9	9	9	9	9	9	9	9	9	9	9
##	[2545]	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
##	[2569]	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
##	[2593]	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
##	[2617]	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
##	[2641]	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
##	[2665]	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
##	[2689]	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
##	[2713]	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
##	[2737]	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
##	[2761]	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	6	6
##	[2785]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[2809]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[2833]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[2857]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[2881]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[2905]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[2929]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[2953]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[2977]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[3001]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[3025]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[3049]	6	6	6																		

##	[3673]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[3697]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[3721]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[3745]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[3769]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[3793]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[3817]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[3841]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[3865]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[3889]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[3913]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[3937]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[3961]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[3985]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4009]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4033]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4057]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4081]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4105]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4129]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4153]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4177]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4201]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4225]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4249]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4273]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4297]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4321]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[4345]	8	8	8																		

[illegible]

##	[6265]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6289]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6313]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6337]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6361]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6385]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6409]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6433]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6457]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6481]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6505]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6529]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6553]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[6577]	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6601]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6625]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6649]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6673]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6697]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6721]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6745]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6769]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6793]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6817]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6841]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6865]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6889]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[6913]	8	8	8	8	8	8	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
##	[6937]	9	9	9																		



##	[7561]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7585]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7609]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7633]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7657]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7681]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7705]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7729]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7753]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7777]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7801]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7825]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7849]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[7873]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	8	8	8	8	8
##	[7897]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[7921]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[7945]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[7969]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[7993]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[8017]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[8041]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[8065]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[8089]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[8113]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[8137]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[8161]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[8185]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[8209]	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
##	[8233]	8	8	8																		

##	[8857]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[8881]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[8905]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[8929]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[8953]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[8977]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[9001]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
##	[9025]	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	7	7	7	7
##	[9049]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9073]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9097]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9121]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9145]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9169]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9193]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9217]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9241]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9265]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9289]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9313]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9337]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9361]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9385]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9409]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9433]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9457]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9481]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9505]	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
##	[9529]	7	7	7																		

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

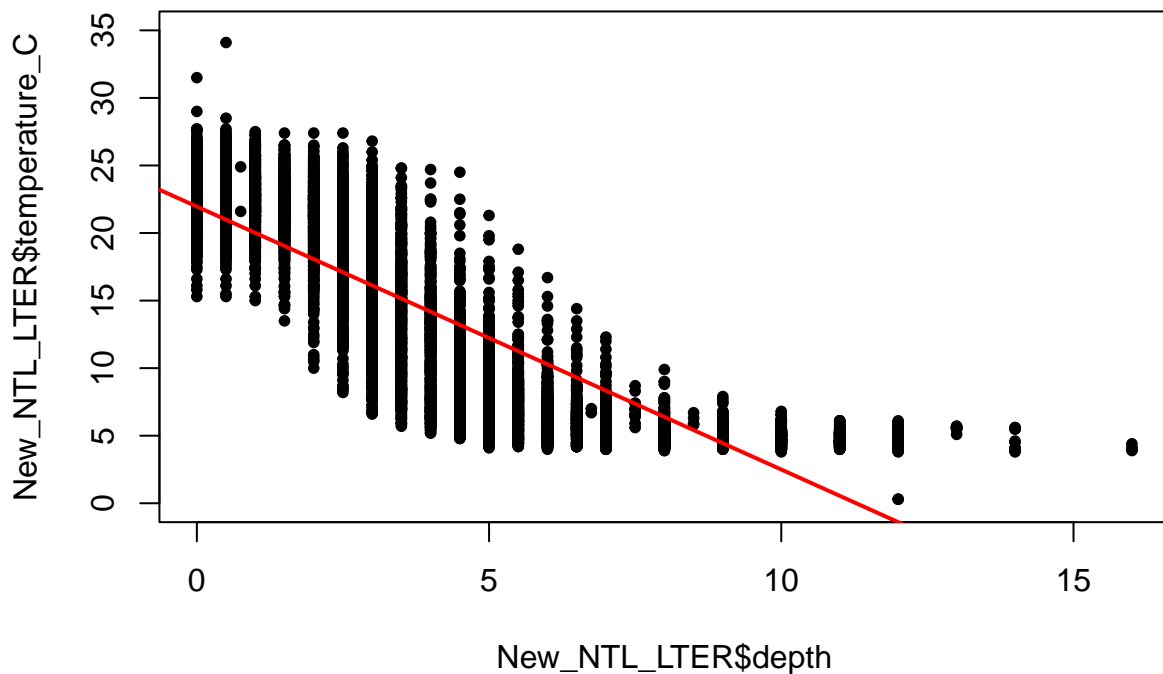
[illegible]

[illegible]



```
New_NTL_LTER <- filter(NL_ChemistryPhysics, month(NL_ChemistryPhysics$sampldate) ==
7) %>%
  select(lakename, year4, daynum, depth, temperature_C) %>%
  drop_na()

# 5
plot(New_NTL_LTER$depth, New_NTL_LTER$temperature_C, pch = 20, cex = 1, ylim = c(0,
35))
model = lm(temperature_C ~ depth, New_NTL_LTER)
abline(model, col = "red", lwd = 2)
```



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

Answer: It is obvious that as the depth increases, the temperature gradually decreases. The distribution of points suggests the linearity of this trend.

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

```
# 7
liner7.regression <- lm(New_NTL_LTER$temperature_C ~ New_NTL_LTER$depth)
summary(liner7.regression)
```

```
##
```

```
## Call:
## lm(formula = New_NTL_LTER$temperature_C ~ New_NTL_LTER$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 ***
## New_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

cor.test(New_NTL_LTER$temperature_C, New_NTL_LTER$depth)

##
## Pearson's product-moment correlation
##
## data:  New_NTL_LTER$temperature_C and New_NTL_LTER$depth
## t = -165.83, df = 9726, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.8646036 -0.8542169
## sample estimates:
##          cor
## -0.8594989
```

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: Temperature and water depth are negatively correlated, and the negative correlation coefficient is -0.8594989, which is very close to -1. From the result we get that the p-value is smaller than 2.2e-16, indicating that the depth is very explanatory for the independent variable of temperature. Because if the p-value is less than 0.05, it means that the independent variable is very explanatory for the dependent variable. There is about 1.15 meters of temperature change for every meter of 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
TP_AIC <- lm(data = New_NTL_LTER, temperature_C ~ year4 + daynum + depth)
model_Tem_Predict <- lm(data = New_NTL_LTER, temperature_C ~ year4 + daynum + depth)
summary(model_Tem_Predict)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4 + daynum + depth, data = New_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) -8.575564   8.630715  -0.994  0.32044
## year4        0.011345   0.004299   2.639  0.00833 **
## daynum       0.039780   0.004317   9.215 < 2e-16 ***
## depth       -1.946437   0.011683 -166.611 < 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
```

```
# 10
TP10_regression <- lm(data = subset(New_NTL_LTER), temperature_C ~ year4 + daynum +
  depth)
summary(TP10_regression)
```

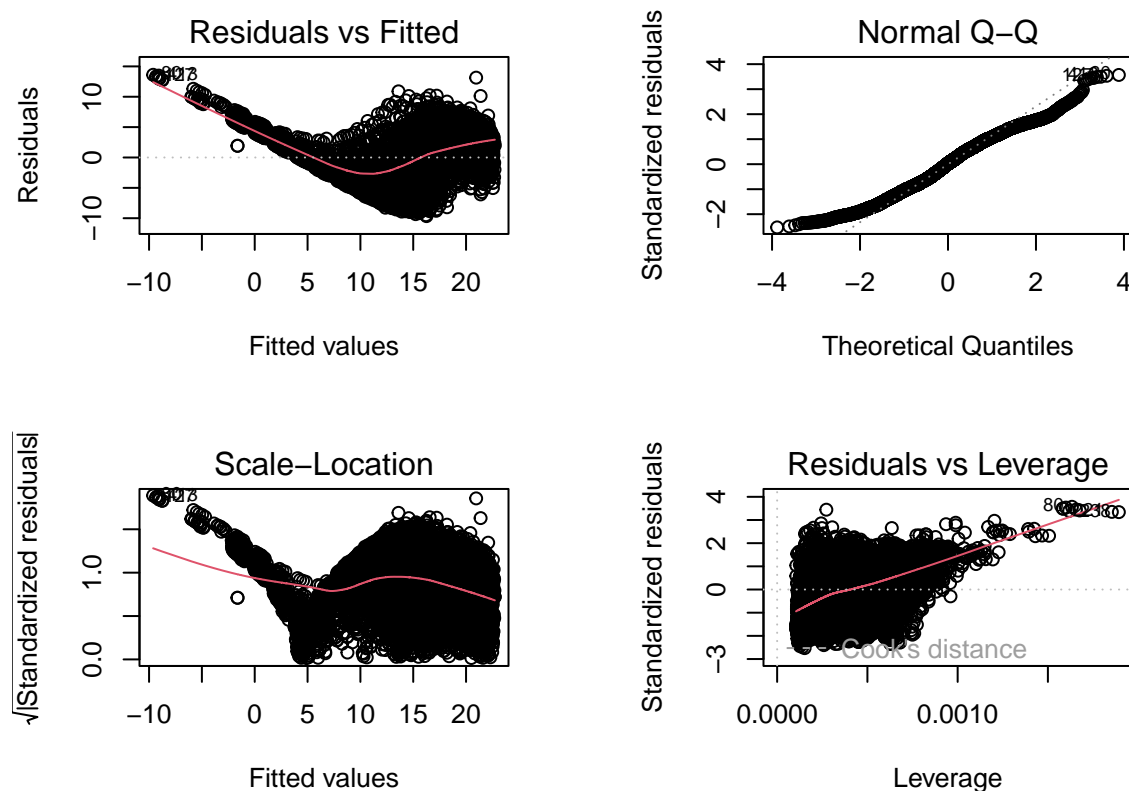
```
##
## Call:
## lm(formula = temperature_C ~ year4 + daynum + depth, data = subset(New_NTL_LTER))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.6536 -3.0000  0.0902  2.9658 13.6123
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept) -8.575564   8.630715  -0.994  0.32044
## year4        0.011345   0.004299   2.639  0.00833 **
## daynum       0.039780   0.004317   9.215 < 2e-16 ***
## depth       -1.946437   0.011683 -166.611 < 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
```

```
lake_regression10 <- ggplot(NULL, aes(x = depth, y = daynum)) + geom_point(data = New_NTL_LTER,
  color = "red") + geom_smooth(data = New_NTL_LTER, color = "Blue") + xlim(0, 25) +
  ylim(0, 35)
print(TP10_regression)
```

```
##
## Call:
## lm(formula = temperature_C ~ year4 + daynum + depth, data = subset(New_NTL_LTER))
##
## Coefficients:
## (Intercept)      year4      daynum      depth
##   -8.57556      0.01134      0.03978     -1.94644
```

```
par(mfrow = c(2, 2), mar = c(4, 4, 4, 4))
plot(TP10_regression)
```



```
par(mfrow = c(1, 1))
```

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: The last set of explanatory variables is year4. A variance of 0.004299 was observed. This is an improvement over the model using only depth as the explanatory variable.

---

## 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

```
ChemistryPhysicslake_model.anova.2way <- aov(data = New_NTL_LTER, temperature_C ~  
  lakename)  
summary(New_NTL_LTER)
```

```
##          lakename      year4      daynum      depth  
## Peter Lake   :2872   Min.    :1984   Min.    :182.0   Min.    : 0.000  
## Paul Lake    :2660   1st Qu.:1992   1st Qu.:190.0   1st Qu.: 2.000  
## Tuesday Lake :1524   Median :1998   Median :198.0   Median : 4.500  
## West Long Lake:1026   Mean    :1999   Mean    :197.5   Mean    : 4.745  
## East Long Lake: 968   3rd Qu.:2006   3rd Qu.:205.0   3rd Qu.: 7.000  
## Crampton Lake : 318   Max.    :2016   Max.    :213.0   Max.    :16.000  
## (Other)      : 360  
## temperature_C  
## Min.       : 0.30  
## 1st Qu.    : 5.50  
## Median     :10.10  
## Mean       :12.72  
## 3rd Qu.    :20.80  
## Max.       :34.10  
##
```

```
ChemistryPhysicslake_model.anova.2way2 <- lm(data = New_NTL_LTER, temperature_C ~  
  lakename)  
summary(ChemistryPhysicslake_model.anova.2way2)
```

```
##  
## Call:  
## lm(formula = temperature_C ~ lakename, data = New_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 ***  
## lakenameCrampton Lake    -2.3145     0.7699  -3.006 0.002653 **  
## lakenameEast Long Lake   -7.3987     0.6918 -10.695 < 2e-16 ***
```

```
## lakenamHummingbird Lake -6.8931 0.9429 -7.311 2.87e-13 ***
## lakenamPaul Lake -3.8522 0.6656 -5.788 7.36e-09 ***
## lakenamPeter Lake -4.3501 0.6645 -6.547 6.17e-11 ***
## lakenamTuesday Lake -6.5972 0.6769 -9.746 < 2e-16 ***
## lakenamWard Lake -3.2078 0.9429 -3.402 0.000672 ***
## lakenamWest 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
```

```
TukeyHSD(ChemistryPhysicslake_model.anova.2way)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = temperature_C ~ lakenam, data = New_NTL_LTER)
##
## $lakenam
## 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: The most p value is 0 or near to 0. This means that there is no significant difference.

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.

```
summary(New_NTL_LTER$lakename)
```

```
## Central Long Lake      Crampton Lake      East Long Lake      Hummingbird Lake
##              128              318              968              116
##      Paul Lake      Peter Lake      Tuesday Lake      Ward Lake
##              2660              2872              1524              116
##      West Long Lake
##              1026
```

```
Central_Long_Lake <- filter(New_NTL_LTER, lakename == "Central Long Lake")
Crampton_Lake <- filter(New_NTL_LTER, lakename == "Crampton Lake")
EastLong_Lake <- filter(New_NTL_LTER, lakename == "East Long Lake")
Hummingbird_Lake <- filter(New_NTL_LTER, lakename == "Hummingbird Lake")
Paul_Lake <- filter(New_NTL_LTER, lakename == "Paul Lake")
Peter_Lake <- filter(New_NTL_LTER, lakename == "Peter Lake")
Tuesday_Lake <- filter(New_NTL_LTER, lakename == "Tuesday Lake")
Ward_Lake <- filter(New_NTL_LTER, lakename == "Ward Lake")
WestLong_Lake <- filter(New_NTL_LTER, lakename == "West Long Lake")

Tem_depth <- ggplot(NULL, aes(x = depth, y = temperature_C)) + geom_point(data = Central_Long_Lake,
  color = "yellow", alpha = 0.5) + geom_point(data = Crampton_Lake, color = "red",
  alpha = 0.5) + geom_point(data = EastLong_Lake, color = "blue", alpha = 0.5) +
  geom_point(data = Hummingbird_Lake, color = "green", alpha = 0.5) + geom_point(data = Paul_Lake,
  color = "pink", alpha = 0.5) + geom_point(data = Peter_Lake, color = "black",
  alpha = 0.5) + geom_point(data = Tuesday_Lake, color = "purple", alpha = 0.5) +
  geom_point(data = Ward_Lake, color = "orange", alpha = 0.5) + geom_point(data = WestLong_Lake,
  color = "brown", alpha = 0.5) + geom_smooth(data = Crampton_Lake, color = "red",
  method = "lm", se = FALSE) + geom_smooth(data = Central_Long_Lake, color = "yellow",
  method = "lm", se = FALSE) + geom_smooth(data = EastLong_Lake, color = "blue",
  method = "lm", se = FALSE) + geom_smooth(data = Hummingbird_Lake, color = "green",
  method = "lm", se = FALSE) + geom_smooth(data = Paul_Lake, color = "pink", method = "lm",
  se = FALSE) + geom_smooth(data = Peter_Lake, color = "black", method = "lm",
  se = FALSE) + geom_smooth(data = Tuesday_Lake, color = "purple", method = "lm",
  se = FALSE) + geom_smooth(data = Ward_Lake, color = "orange", method = "lm",
  se = FALSE) + geom_smooth(data = WestLong_Lake, color = "brown", method = "lm",
  se = FALSE) + xlim(0, 18) + ylim(0, 35)
print(Tem_depth)
```

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

## Warning: Removed 5 rows containing missing values (geom_smooth).

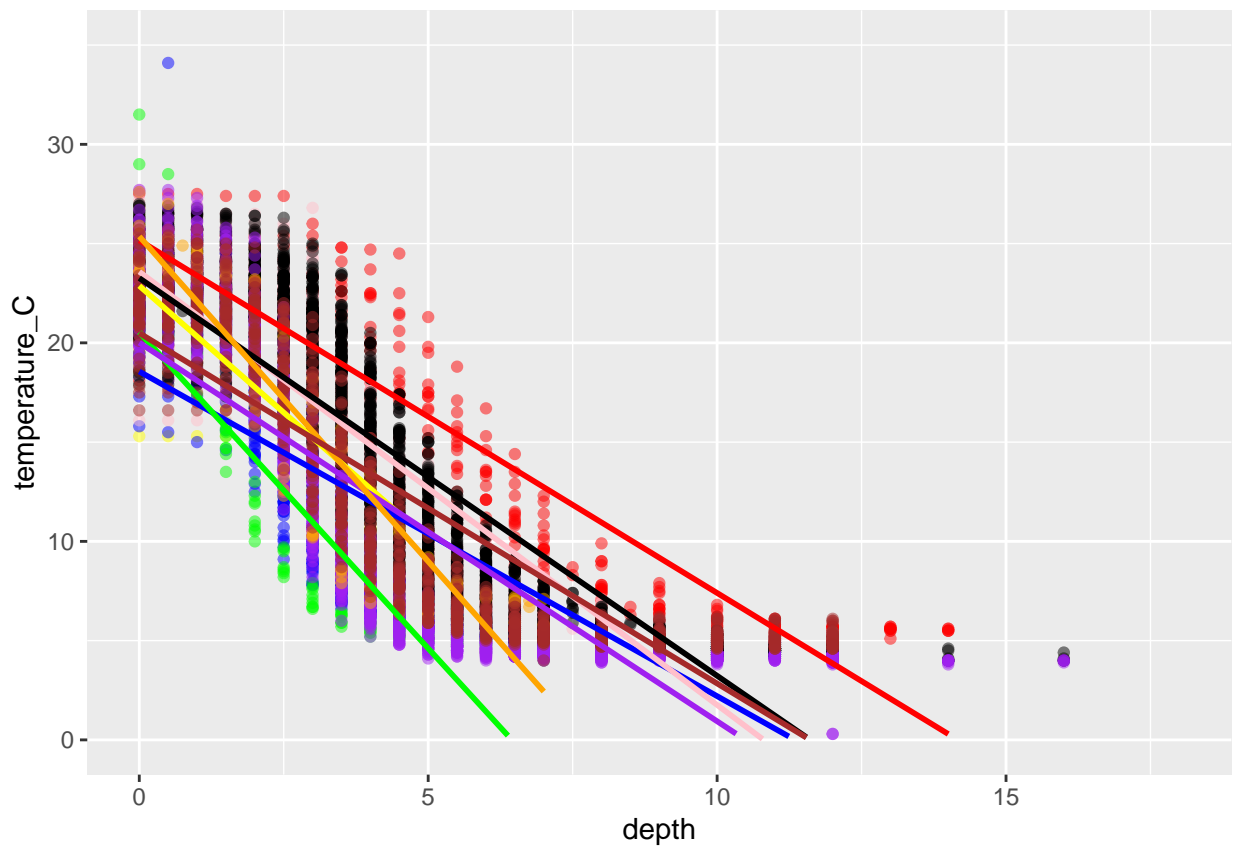
## Warning: Removed 7 rows containing missing values (geom_smooth).

## Warning: Removed 8 rows containing missing values (geom_smooth).

## Warning: Removed 22 rows containing missing values (geom_smooth).

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

## Warning: Removed 3 rows containing missing values (geom_smooth).
```



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



```
# 15
```

```
TukeyHSD(ChemistryPhysicslake_model.anova.2way)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = temperature_C ~ lakename, data = New_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

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:The temperatures of Paul Lake has the same mean temperature as Peter Lake. Crampton Lake is warmer than other average temperatures and with the increasing of depth, the mean temperature decreases quickly in Hummindbird Lake.

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: We can do the Comparison tests, because the result could accurately compare the means of two groups.

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?

```
summary(New_NTL_LTER$lakename)
```

```
## Central Long Lake      Crampton Lake      East Long Lake      Hummingbird Lake
##              128              318              968              116
##      Paul Lake      Peter Lake      Tuesday Lake      Ward Lake
##      2660              2872              1524              116
## West Long Lake
##      1026
```

```
CarmWardLake <- filter(New_NTL_LTER, lakename == "Crampton Lake" | lakename == "Ward Lake")
meancompare1 <- t.test(CarmWardLake$temperature_C ~ CarmWardLake$lakename)
meancompare2 <- lm(CarmWardLake$temperature_C ~ CarmWardLake$lakename)
view(meancompare2)
```

Answer: Their July temperature are same. Mean temperature are not equal. Match.