

# STAT 443: Lab 1

Caden Hewlett (Student #40404246)

12 January, 2024

## Question 1

```
# read in data to an object called dat
dat <- read.csv("LakeLevels.csv")
head(dat, 10)
```

```
##           Date LakeLevel
## 1  1/1/2007   3732.65
## 2  1/2/2007   3732.65
## 3  1/3/2007   3732.65
## 4  1/4/2007   3732.64
## 5  1/5/2007   3732.64
## 6  1/6/2007   3732.64
## 7  1/7/2007   3732.64
## 8  1/8/2007   3732.64
## 9  1/9/2007   3732.64
## 10 1/10/2007  3732.64
```

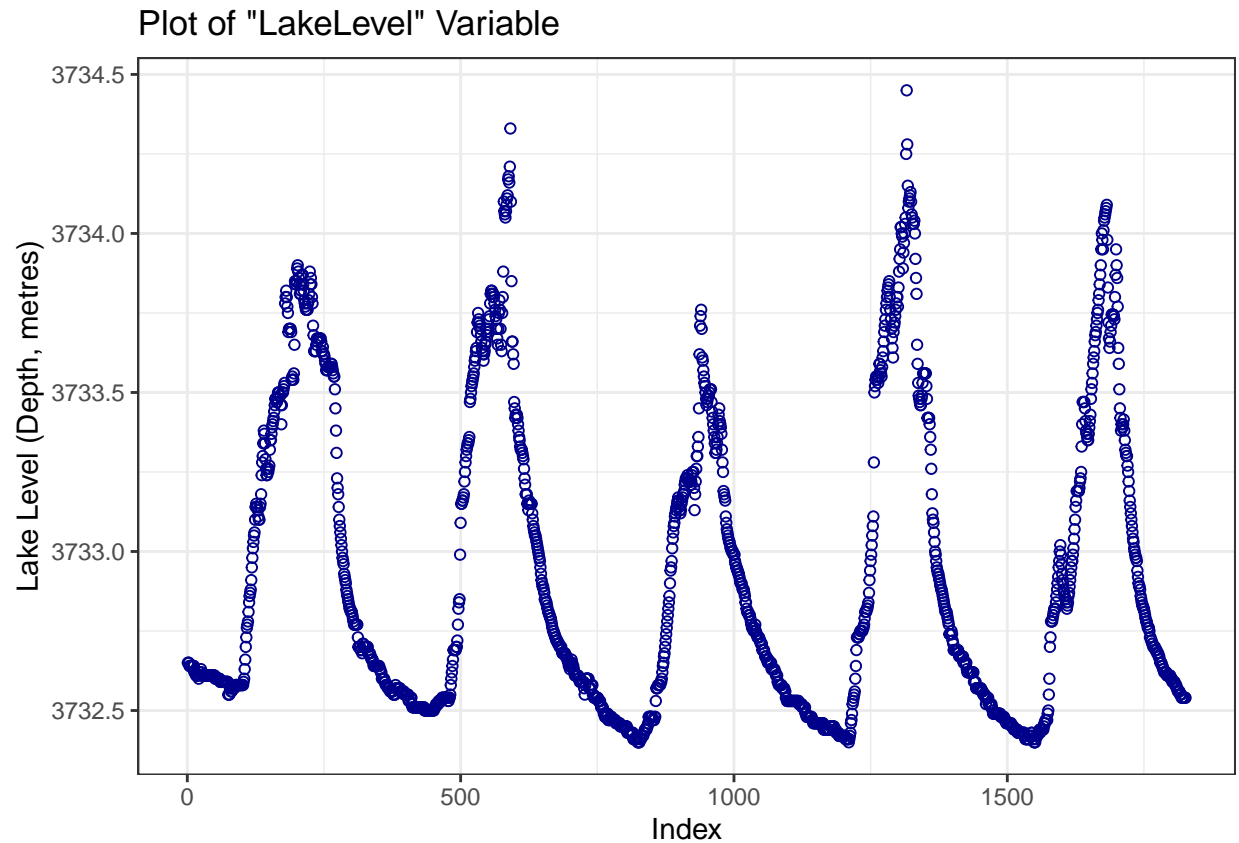
Then, we can check the `names` of this data frame.

```
names(dat)
```

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

Now, we create a plot of the `LakeLevel` variable.

```
ggplot(dat, aes(x = 1:nrow(dat), y = LakeLevel)) +
  geom_point(shape = 1, col = "darkblue") +
  xlab("Index") +
  ylab("Lake Level (Depth, metres)") +
  ggtitle("Plot of \"LakeLevel\" Variable") +
  theme_bw()
```



This plot is different from what we want for a time series because the value on the  $x$ -axis is the Index of the Lake Level data points in the `dat` dataframe rather than time. Hence, this is not a time series (yet!)

## Question 2

Now, we can test if this is a time series object.

```
is.ts(dat)
```

```
## [1] FALSE
```

Since it isn't, we'll convert it.

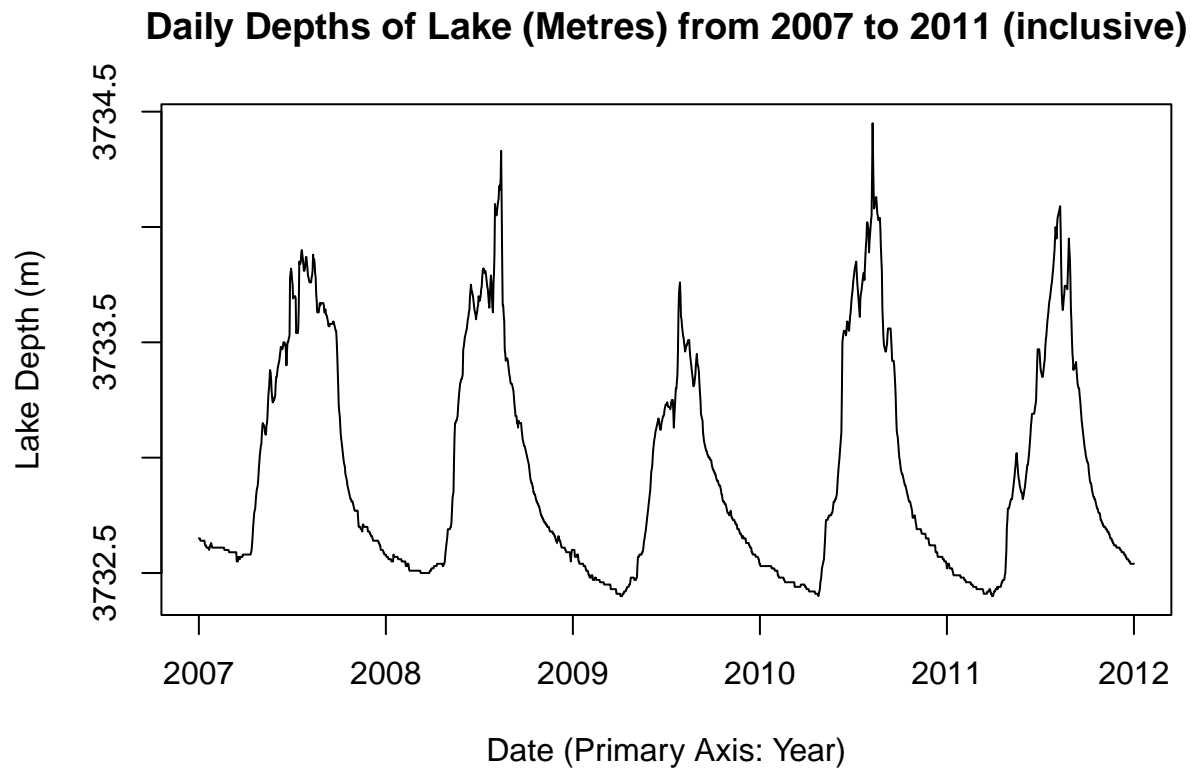
```
# convert to ts
x <- ts(dat$LakeLevel,
        start = c(2007, 1),
        end = c(2012, 1), # note: (2011, 12) omitted December
        frequency = 365)

# check that we haven't lost any data
length(x) == nrow(dat)
```

```
## [1] TRUE
```

### Question 3

```
plot(x,  
     main = "Daily Depths of Lake (Metres) from 2007 to 2011 (inclusive)",  
     ylab = "Lake Depth (m)",  
     xlab = "Date (Primary Axis: Year)"  
     )
```



### Question 4