Assignment 7: Water Quality in Rivers

Jack Carpenter

OVERVIEW

This exercise accompanies the lessons in Water Data Analytics on water quality in rivers.

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, check your PDF against the key and then submit your assignment completion survey at https://forms.gle/AF6vXHWbeQGEnHpNA

Having trouble? See the assignment's answer key if you need a hint. Please try to complete the assignment without the key as much as possible - this is where the learning happens!

Target due date: 2022-03-22

Setup

1. Verify your working directory is set to the R project file. Load the tidyverse, lubridate, cowplot, and dataRetrieval packages. Set your ggplot theme (can be theme_classic or something else)

```
# setup
getwd()
```

```
## [1] "/Users/Jack/Documents/Duke/Spring 2022/Water Data Analytics/Water_Data_Analytics_2022"
```

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.8
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(cowplot)
library(dataRetrieval)

theme_set(theme_classic())
options(scipen = 4)
```

Hypoxia

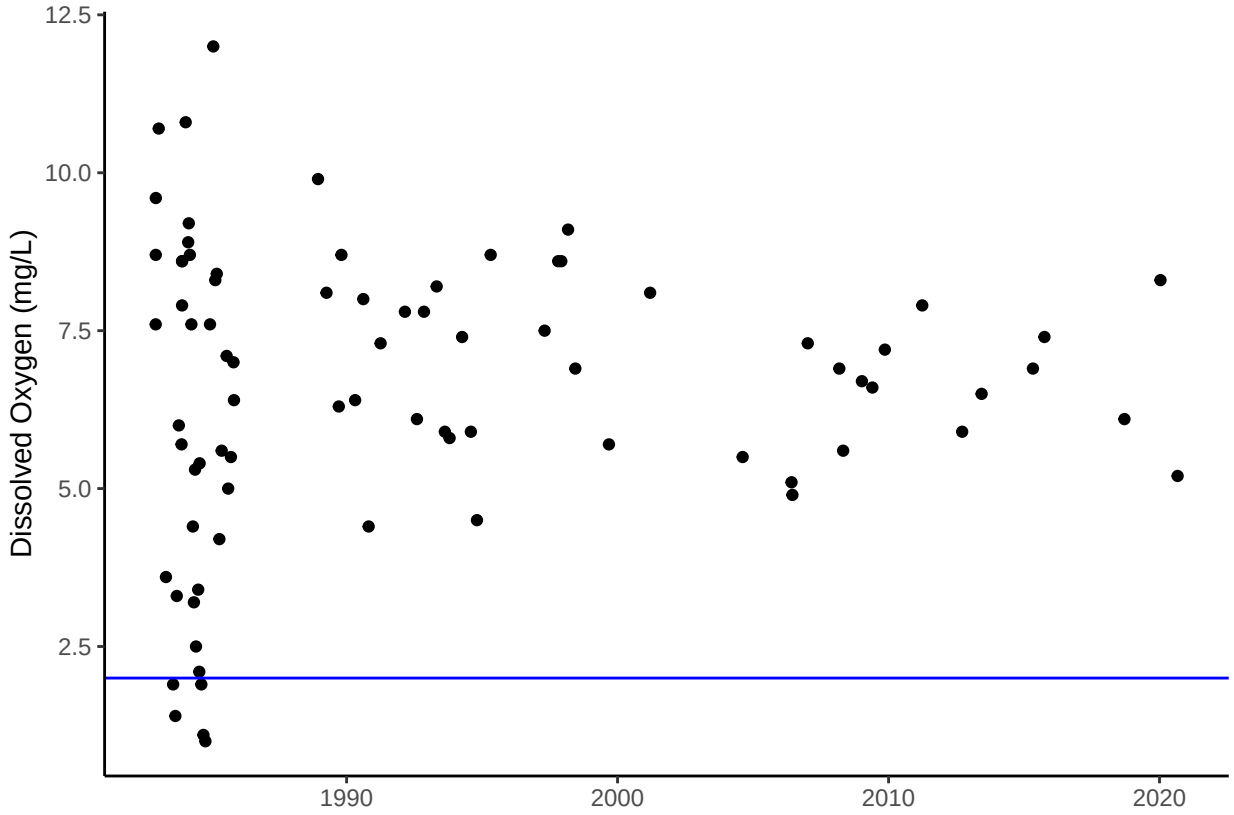
This assignment will look at another measure of water quality - oxygen concentration. Oxygen in the water column is very important for aquatic life, and so is considered a measure of water quality. Hypoxia (low oxygen) has many different definitions. For this assignment, we will use 2 mg/L O_2 as our cut-off.

2. Import the oxygen water quality data from New Hope Creek at Blands (using `readWQPqw()`, site code `USGS-02097314`, parameter code `00300`). Make a data frame called `NewHopeDO` that includes only the Date and dissolved oxygen concentration values. Rename the column names "Date" and "DO_mgL".

```
#import data
NewHopeDO <- readWQPqw(siteNumbers = "USGS-02097314",
                       parameterCd = "00300", #DO concentrations
                       startDate = "",
                       endDate = "")
NewHopeDO <- NewHopeDO %>%
  select(ActivityStartDate, ResultMeasureValue) %>% #pick our two columns
  rename(Date = ActivityStartDate,
         DO_mgl = ResultMeasureValue) #change the names
```

3. Create a ggplot of oxygen concentrations over time. Include a horizonal line at 2 mg/l to show the hypoxia cutoff.

```
# Plot it to take a quick look-see
DOplot <- ggplot(NewHopeDO, aes(x = Date, y = DO_mgl)) +
  geom_point() +
  geom_hline(yintercept = 2, color = "blue") +
  labs(x = "", y = "Dissolved Oxygen (mg/L)")
DOplot
```



4. What do you notice about the frequency of hypoxia over time?

Hypoxia has actually decreased over time, with fewer low points and none below 5 mg/L after 2000. The only two hypoxic events occurred prior to 1990.

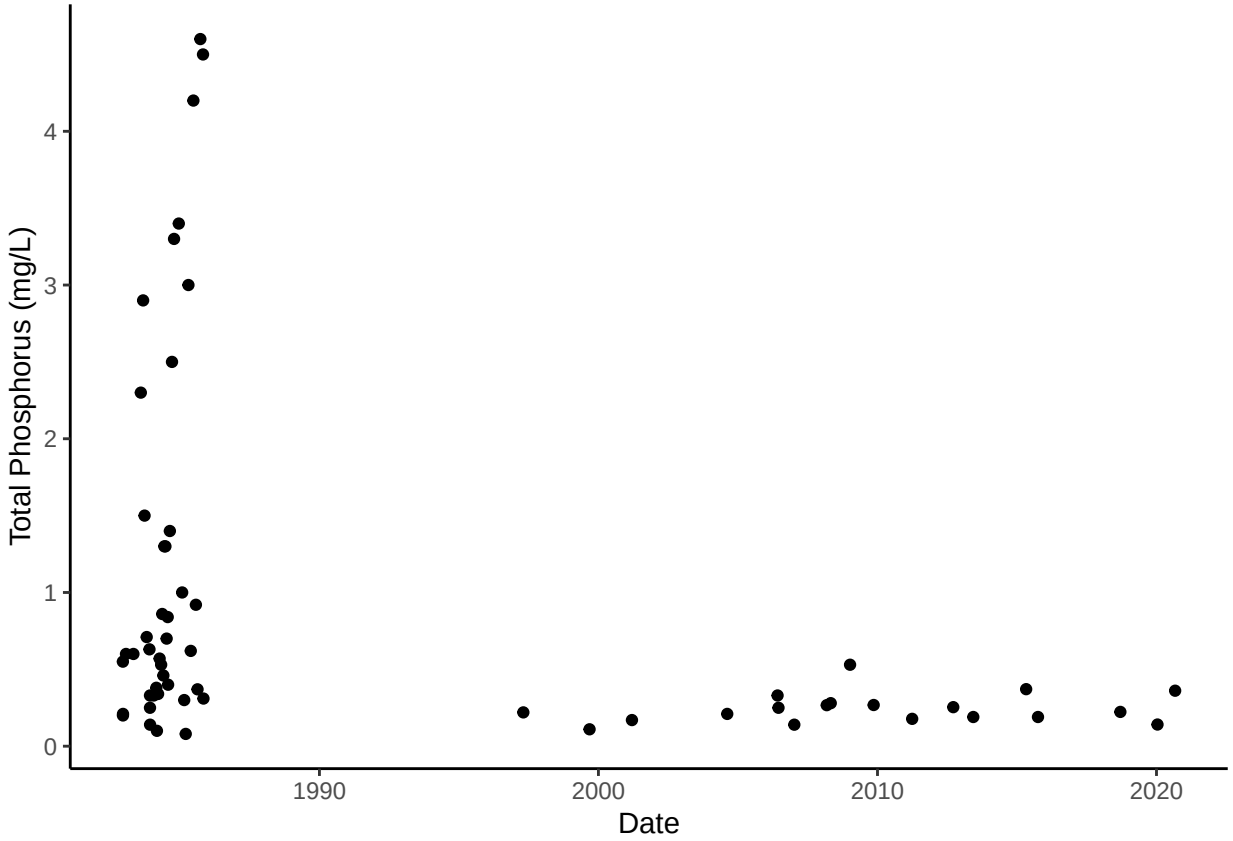
Nutrients

5. Often times hypoxia is associated with high nutrient concentrations, because abundant nutrients promote primary production which in turn increases respiration and depletes oxygen concentrations in the water (remember how oxygen concentrations were very low in the hypolimnion from the Physical Properties of Lakes week). Create a new data frame, called `NewHopeTP` with total phosphorus (parameter code `00665`) data from the same site Your data frame should have 2 columns: "Date" and "TP_mgL".

```
NewHopeTP <- readWQPqw(siteNumbers = "USGS-02097314",
                       parameterCd = "00665", #Phosphorus concentrations
                       startDate = "",
                       endDate = "")
NewHopeTP <- NewHopeTP %>%
  select(ActivityStartDate, ResultMeasureValue) %>% #pick out columns
  rename(Date = ActivityStartDate,
         TP_mgl = ResultMeasureValue) #changing names
```

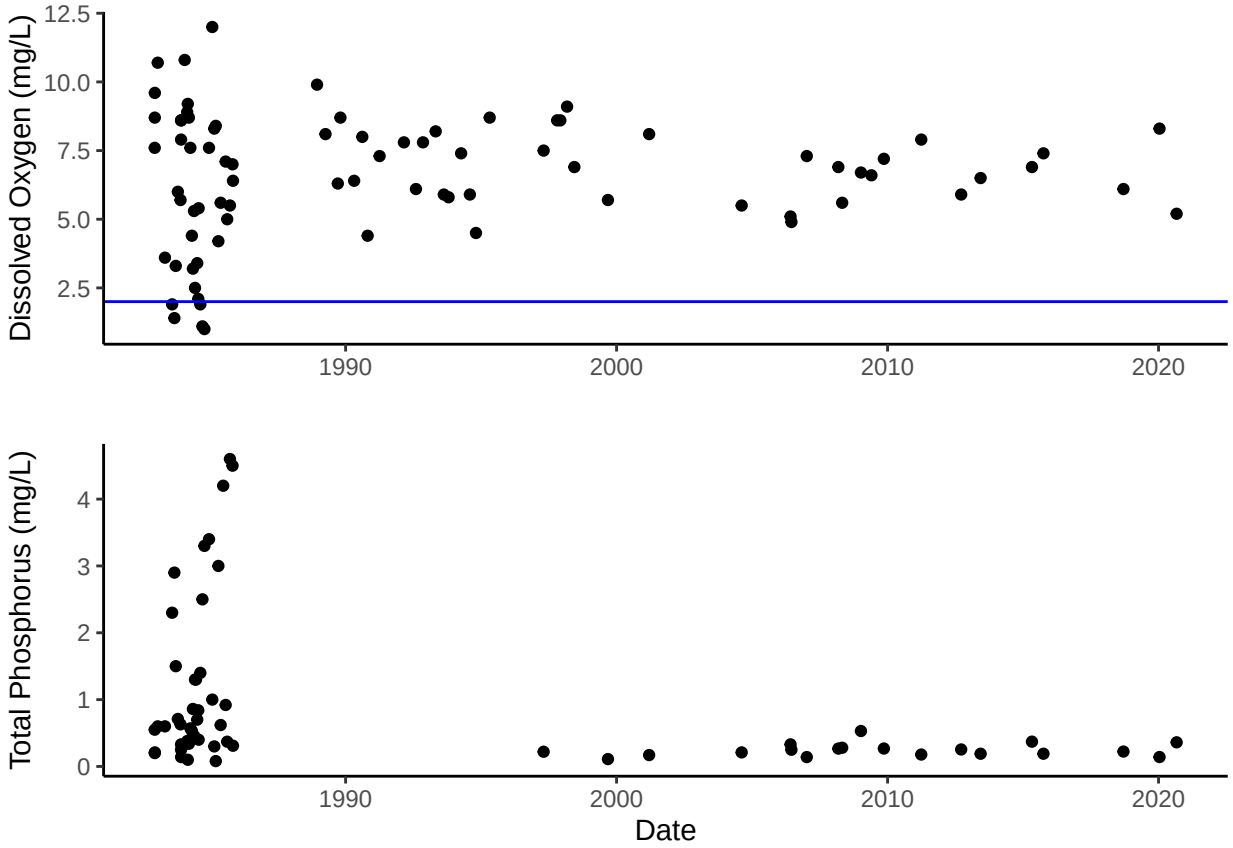
6. Create two ggplots stacked with `plot_grid` that show DO and TP concentrations over time.

```
TPplot <- ggplot(NewHopeTP, aes(x = Date, y = TP_mgl)) +
  geom_point() +
  labs(x = "Date", y = "Total Phosphorus (mg/L)")
TPplot
```




```
plot_grid(DOplot, TPplot,
          label_y = "AUTO", #keep our y labels
```

```
          label_x = "Date", #label our x-axis Date
          label_size = 10, #shrink from the automatic size 14
          ncol = 1, #stack them vertically
          align = "hv") #align them horizontally and vertically
```



7. What do these plots tell you about nutrient concentrations over time? How might nutrient concentrations relate to your previous plot of hypoxia events?

   These plots confer that nutruent concentrations over time have decreased to a mostly low level. The lack of nutruents correlates with the decrease in hypoxia events as well, indicating that the hypoxic events were potentially linked to nutrient availability like some sort of blooms.

Discharge and Dissolved Oxygen

8. Turbulent flow in rivers mixes oxygen into the water column. As discharge decreases, water moves slower, and oxygen diffuses slower into the water from the atmosphere. Download and reformat the daily discharge data for New Hope Creek (function `readNWISdv()`, site `02097314`, parameter `00060`).

```
NewHopeDischarge <- readNWISdv(site = "02097314",
                               parameterCd = "00060", #discharge in cfs
                               startDate = "",
                               endDate = "")
#check to see if this read in as a date
class(NewHopeDischarge$Date)
```

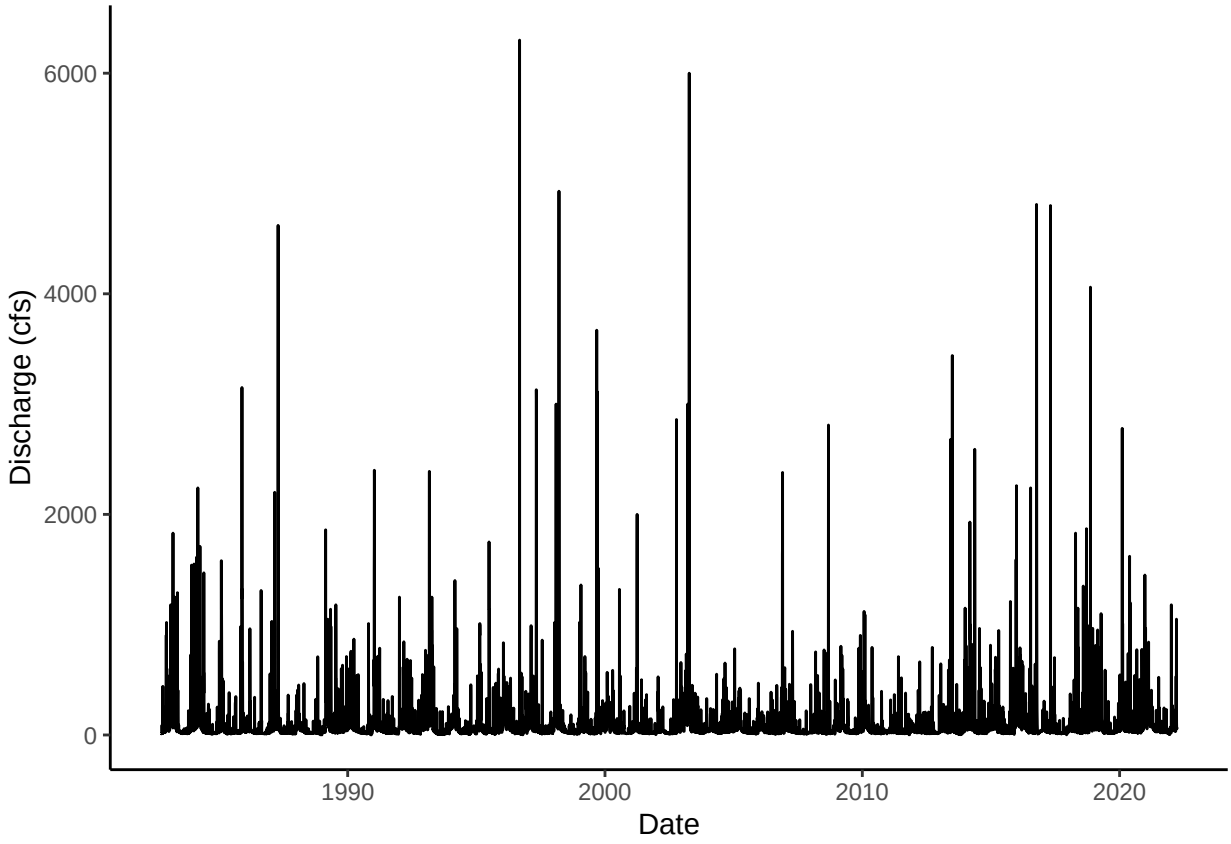
```
## [1] "Date"
```

```
#now rename the columns without useful names
NewHopeDischarge <- NewHopeDischarge %>%
```

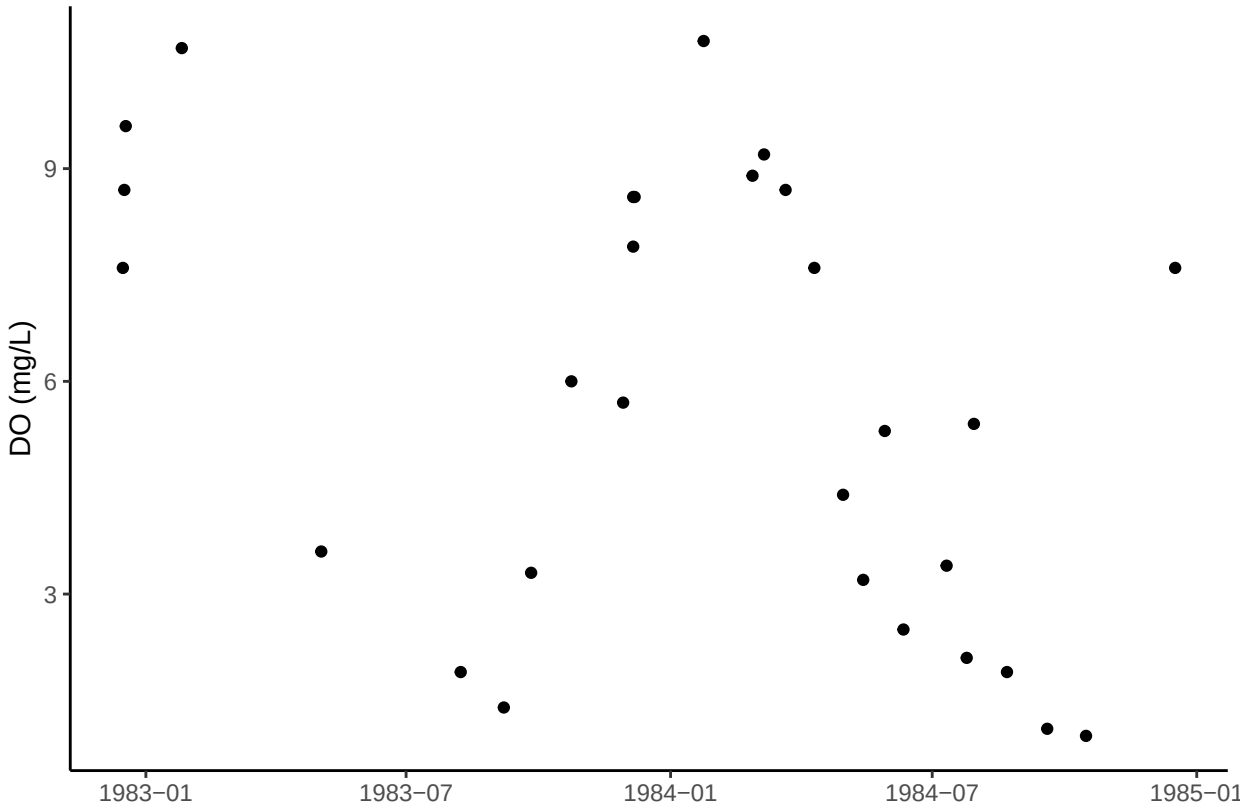
```
  select(Date = Date, Discharge = X_00060_00003)
```

9. Create two ggplots stacked with `plot_grid` that show DO concentrations and discharge over time, for the two years we observed hypoxia (1983 and 1984).

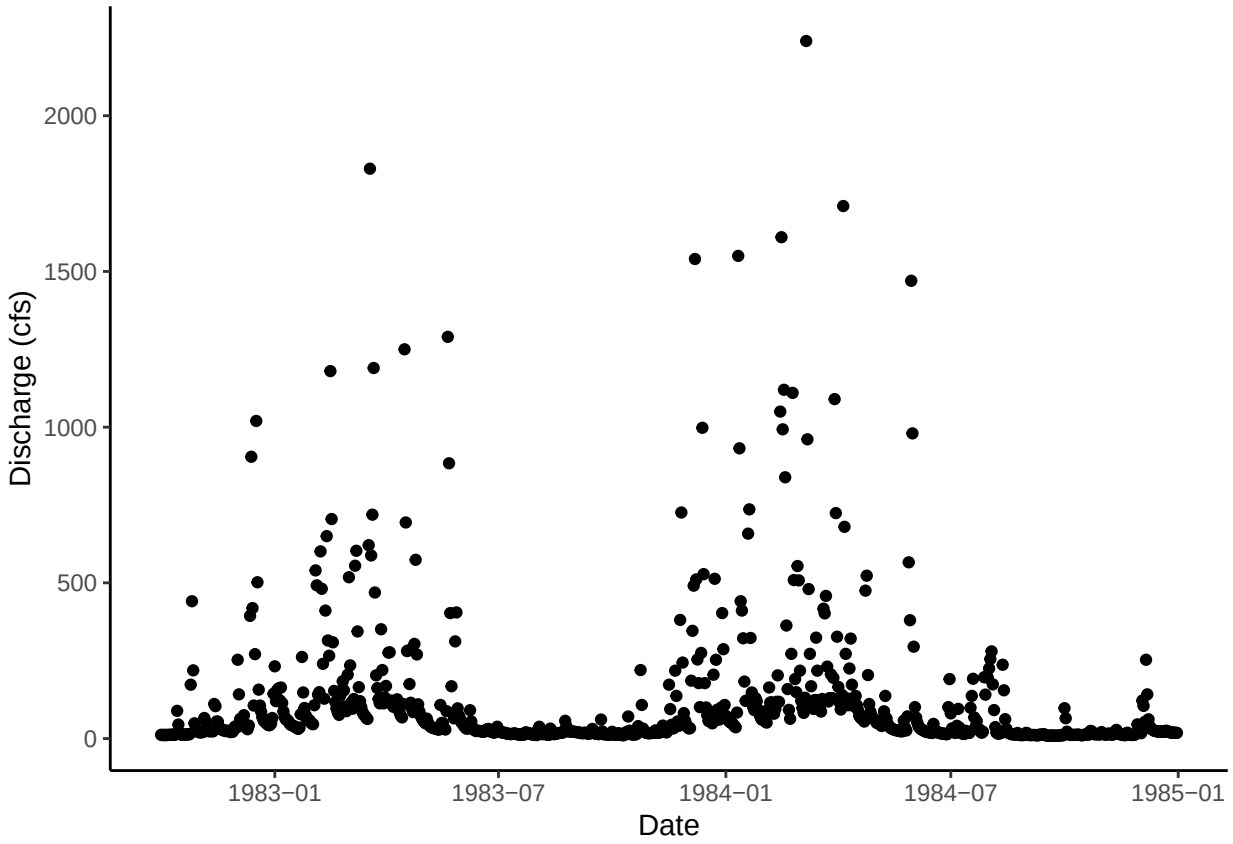
```
DischargePlot <- ggplot(NewHopeDischarge, aes(x = Date, y = Discharge)) +
  geom_line() +
  labs(x = "Date", y = "Discharge (cfs)")
DischargePlot
```



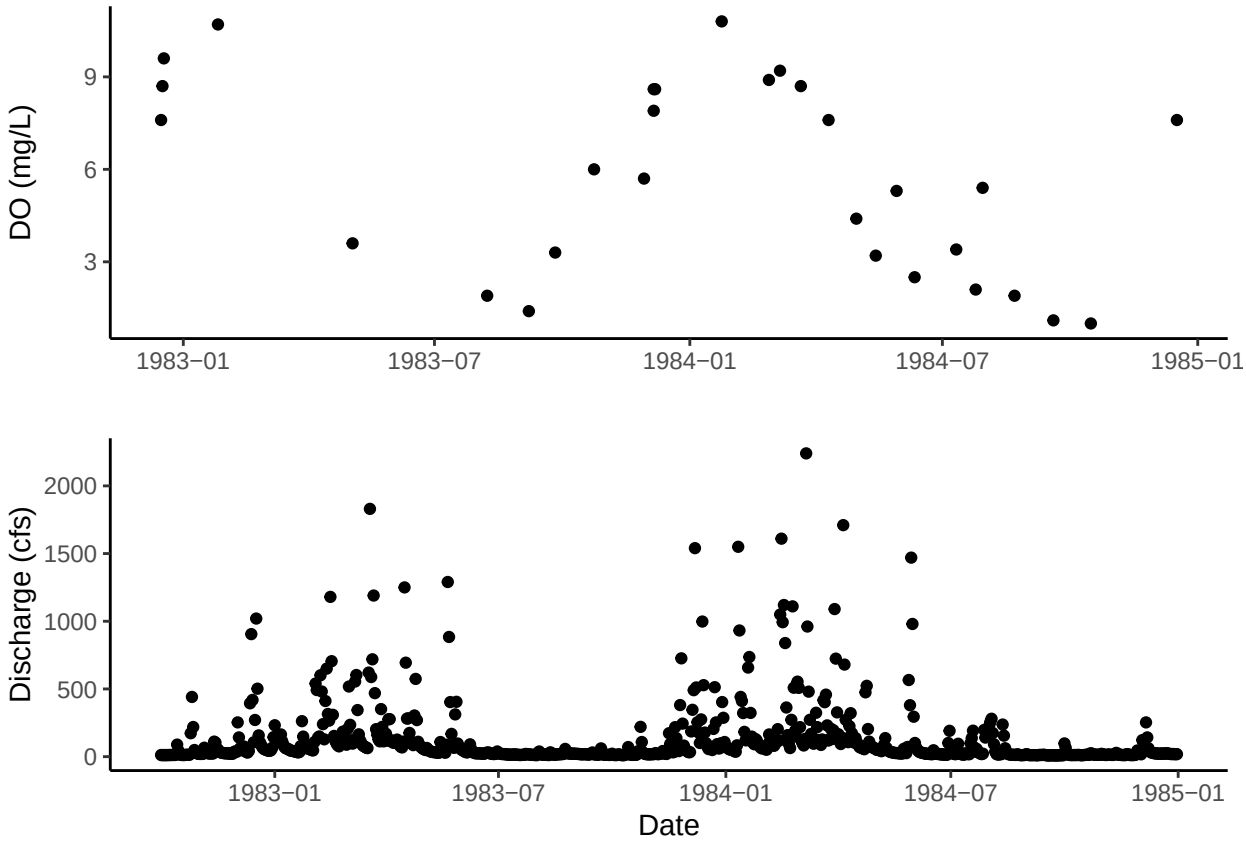

```
#now we have to zoom our plots into the desired dates (1983 and 1984)
DOplot2 <- ggplot(subset(NewHopeDO, Date < "1985-01-01"),
                  aes(x = Date, y = DO_mgl))+
  geom_point() +
  labs(x = "", y = "DO (mg/L)")
DOplot2
```




```
DischargePlot2 <- ggplot(subset(NewHopeDischarge, Date < "1985-01-01"),
                         aes(x = Date, y = Discharge)) +
  geom_point() +
  labs(x = "Date", y = "Discharge (cfs)")
DischargePlot2
```




```
plot_grid(DOplot2, DischargePlot2,
          ncol = 1,
          align = "hv")
```



10. Do hypoxic events happen in the wet or the dry season? Why might that be?

    Looks like hypoxic events occur in the dry season (low discharge), probably because less flow means less water movement and less oxygenation of the water column. Less water movement means both a smaller surface area for atmosphere-water gas exchange and also likely a smoother flow with fewer rapids/riffles to actively mix oxygen into the water.