Assignment 7: Water Quality in Rivers

Lydie Costes

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)

```
# 1. Verify working directory, load packages, set theme
getwd()
```

```
## [1] "/Users/lydiecostes/Documents/Duke/WaterDataAnalytics/Water_Data_Analytics_2022/Assignments"
```

```
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   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.1.1     v forcats 0.5.1

## -- 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(cowplot)
```

```
##
## Attaching package: 'cowplot'

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

```
library(dataRetrieval)

theme_set(theme_bw())
```

Hypoxia

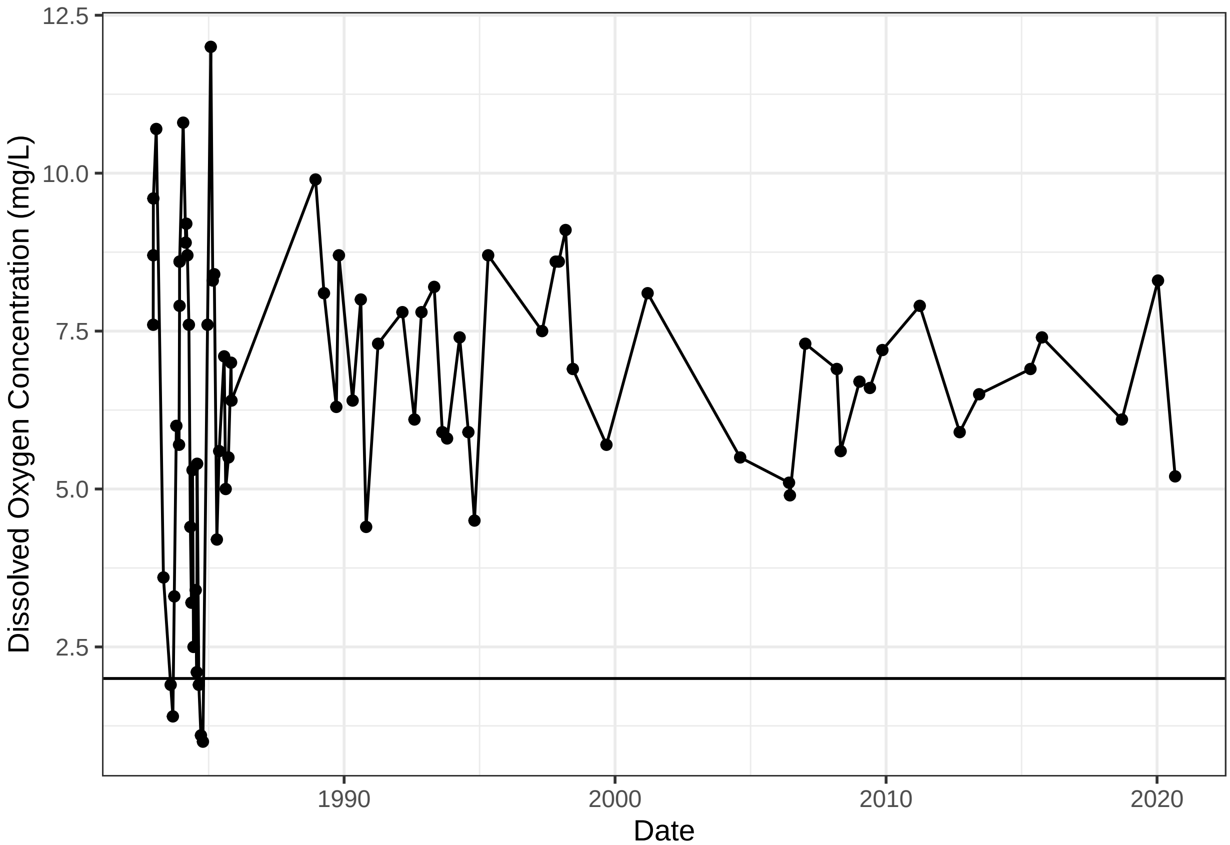
This assignment will look at another measure of water quality - oxygen concentration. Oxygen in the water column is very important for aquatic life, and so is considered a measure of water quality. Hypoxia (low oxygen) has many different definitions. For this assignment, we will use 2 mg/L O_2 as our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

```
# 2. Import data and select date and DO
NewHope <- readWQPqw(siteNumbers = "USGS-02097314",
                     parameterCd = "00300")

NewHopeDO <- NewHope %>%
  select(Date = ActivityStartDate,
         DO_mgL = ResultMeasureValue)
```

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

```
# 3. Create ggplot of O2 concentrations over time
ggplot(NewHopeDO, aes(x = Date, y = DO_mgL)) +
  geom_point() +
  geom_line() +
  geom_hline(yintercept = 2) +
  labs(y = "Dissolved Oxygen Concentration (mg/L)")
```



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

   It stopped occuring by the 90's, or perhaps even by the mid-1980's, but there is a gap in data. The leveling of oxygen concentrations beginning in the 90's is rather dramatic compared with the early 1980's.

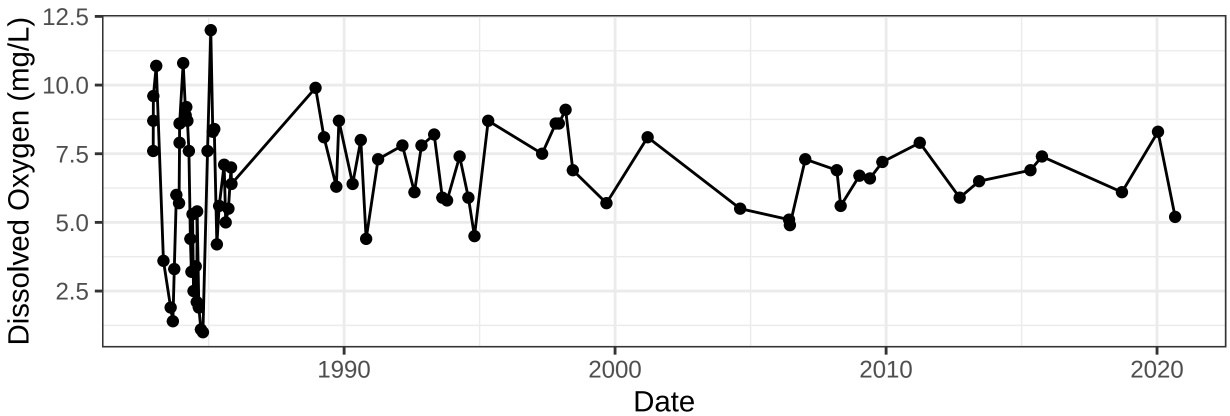
Nutrients

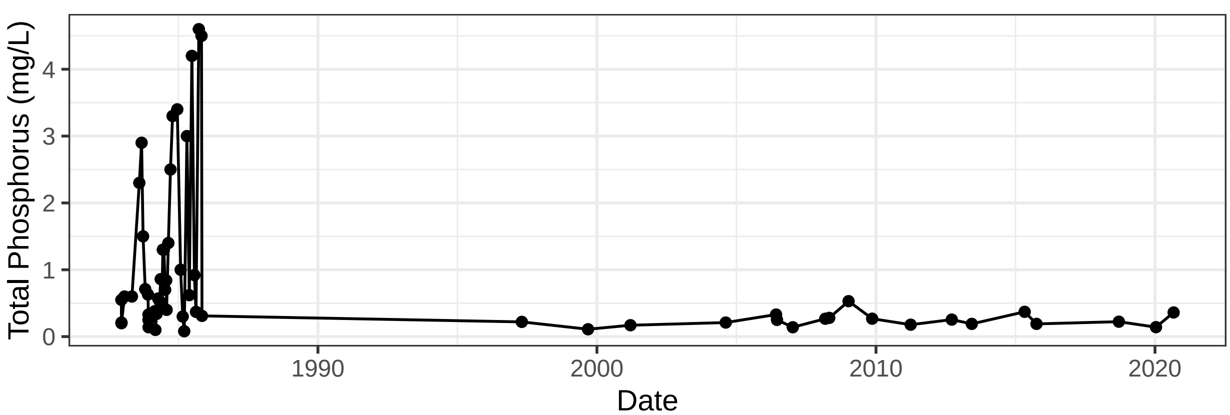
5. Often times hypoxia is associated with high nutrient concentrations, because abundant nutrients promote primary production which in turn increases respiration and depletes oxygen concentrations in the water (remember how oxygen concentrations were very low in the hypolimnion from the Physical Properties 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".

```
# 5. Import data and select date and TP
NewHope <- readWQPqw(siteNumbers = "USGS-02097314",
                     parameterCd = "00665")

NewHopeTP <- NewHope %>%
  select(Date = ActivityStartDate,
         TP_mgL = ResultMeasureValue)
```

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

```
# 6. Plot DO and TP
do_plot <- ggplot(NewHopeDO, aes(x=Date, y=DO_mgL)) +
  geom_line() +
  geom_point() +
  labs(y = "Dissolved Oxygen (mg/L)")

tp_plot <- ggplot(NewHopeTP, aes(x=Date, y=TP_mgL)) +
  geom_line() +
  geom_point() +
  labs(y = "Total Phosphorus (mg/L)")

plot_grid(do_plot, tp_plot, nrow = 2)
```





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 suggest that nutrient levels were higher in the 1980's, when hypoxia occurred. Phosphorus levels dropped dramatically since the 1990's, during which hypoxia events stopped occurring. Because phosphorus is often a limiting factor for algal growth, the lower P levels may be preventing surges in algal growth that could suck up all the oxygen.

Discharge and Dissolved Oxygen

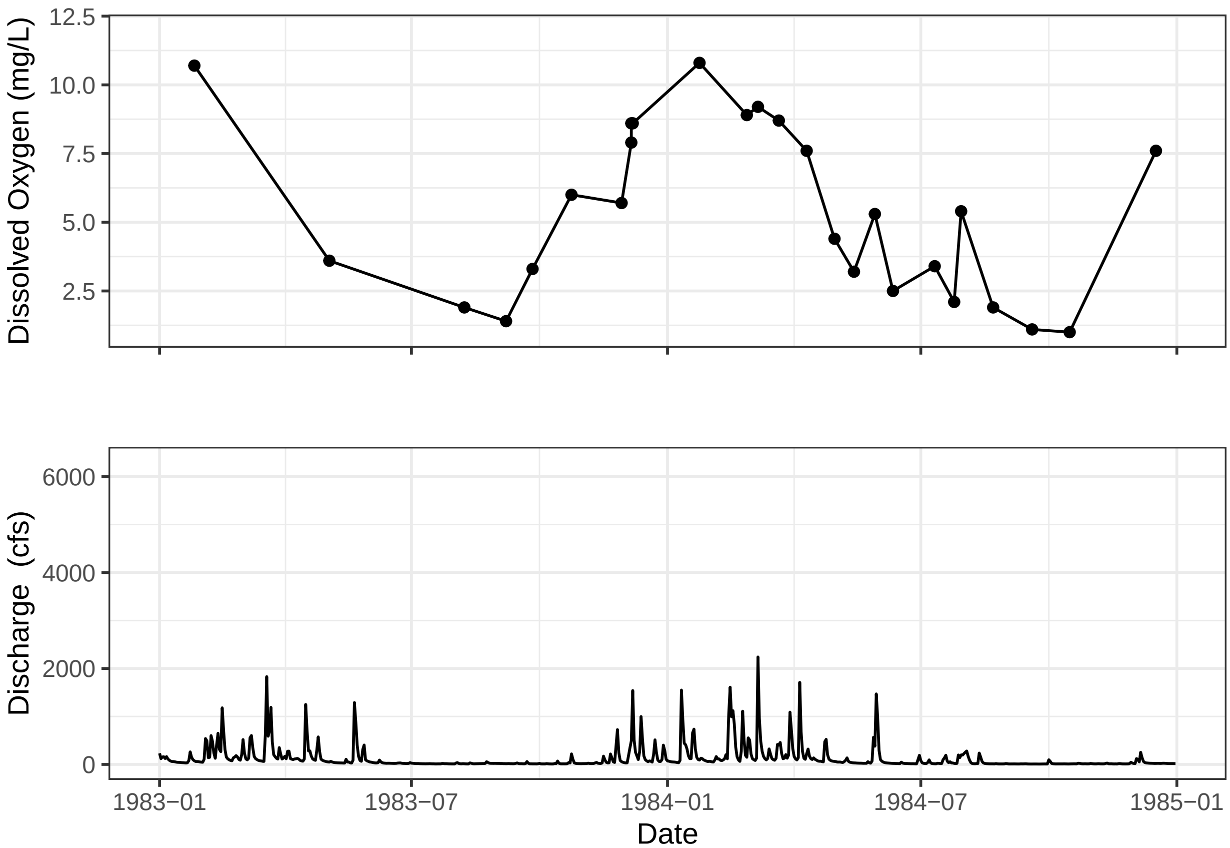
8. Turbulent flow in rivers mixes oxygen into the water column. As discharge decreases, water moves slower, and oxygen diffuses slower into the water from the atmosphere. Download and reformat the daily discharge data for New Hope Creek (function `readNWISdv()`, site `02097314`, parameter `00060`).

```r
# 8. Download and reformat discharge data
NewHopeFlow <- readNWISdv(site = "02097314",
                          parameterCd = "00060")

names(NewHopeFlow)[4:5] <- c("Discharge", "Approval.Code")
```

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).

```r
plot_do <- ggplot(NewHopeDO, aes(x = Date, y = DO_mgL)) +
  geom_line() +
  geom_point() +
  labs(x = "", y = "Dissolved Oxygen (mg/L)") +
  scale_x_date(limits = as.Date(c("1983-01-01", "1984-12-31"))) +
  theme(axis.text.x = element_blank())

plot_fl <- ggplot(NewHopeFlow, aes(x = Date, y = Discharge)) +
  geom_line() +
  labs(y = "Discharge  (cfs)") +
  scale_x_date(limits = as.Date(c("1983-01-01", "1984-12-31")))

plot_grid(plot_do, plot_fl, ncol = 1, align = "hv")
```

```
## Warning: Removed 55 row(s) containing missing values (geom_path).

## Warning: Removed 55 rows containing missing values (geom_point).

## Warning: Removed 13687 row(s) containing missing values (geom_path).
```



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

    Hypoxic events tend to happen in the dry season. Turbulent waters after a rainstorm could lead to more mixing of the water that could increase the amount of dissolved oxygen. If water levels and overall volume of water are lower, the water may be more turbid, allowing for more photosynthesis that uses up the oxygen present in the water.