Assignment 9: Mapping

Jack Carpenter

OVERVIEW

This exercise accompanies the lessons in Water Data Analytics on mapping

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/NDWEUu73LooFJPVM8

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: 2022-04-12

Setup

1. Load the tidyverse, LAGOSNE, maps, dataRetrieval, sf, and nhdplusTools packages. Set your ggplot theme (can be theme_classic or something else).

```
#check wd
getwd()
```

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

```
#load packages
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(LAGOSNE)
library(maps)
```

```
##
## Attaching package: 'maps'
```

```
## The following object is masked from 'package:purrr':
```

```
##
##     map
```

```
library(dataRetrieval)
library(sf)
```

```
## Linking to GEOS 3.9.1, GDAL 3.4.0, PROJ 8.1.1; sf_use_s2() is TRUE
```

```
library(nhdplusTools)
```

```
## USGS Support Package: https://owi.usgs.gov/R/packages.html#support
```

```
#set ggplot theme
theme_set(theme_classic()+
            theme(axis.text = element_text(color = "black", size = 10),
                  legend.position = "right"))
options(scipen = 4)
```

LAGOS-NE

2. Choose five lakes in the LAGOS-NE database that are located within the same state and have chlorophyll data. Subset your data accordingly, and create two plots:

- A map of the five lakes within the state boundary, with each lake point as its own color.
- A boxplot with jittered points layered on top of chlorophyll concentrations in each lake (chlorophyll on y axis, lake on x axis), with each lake having a fill and/or color that matches the map.

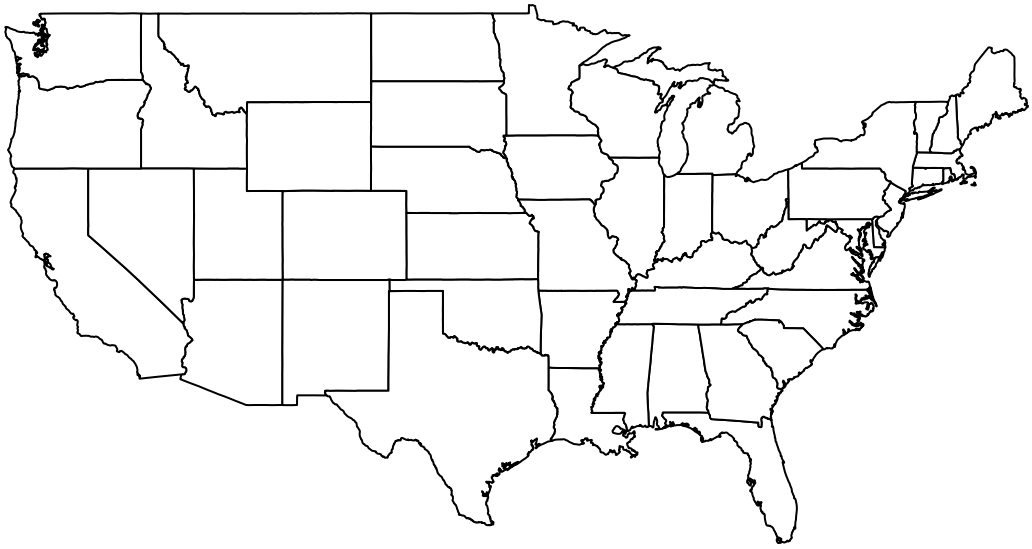
```
#first load LAGOS data
LAGOSdata <- lagosne_load()
```

```
## Warning in (function (version = NULL, fpath = NA) : LAGOSNE version unspecified,
## loading version: 1.087.3
```

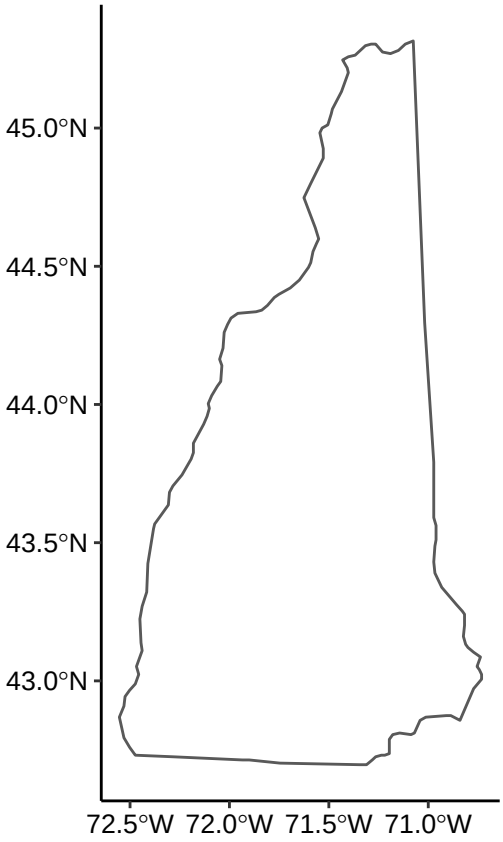
```
#load LAGOS dataframes
LAGOSlocus <- LAGOSdata$locus
LAGOSstate <- LAGOSdata$state
LAGOSnutrient <- LAGOSdata$epi_nutr

#joined at the hip
LAGOScombined <-
  left_join(LAGOSnutrient, LAGOSlocus) %>%
  left_join(., LAGOSstate) %>%
  select(lagoslakeid, sampledate, chla, nhd_lat, nhd_long, state)#choose data
```

```
## Joining, by = "lagoslakeid"
## Joining, by = "state_zoneid"
```

```
#state your name
states <- st_as_sf(map(database = "state", plot = TRUE, fill = TRUE,
                       col = "white"))
```

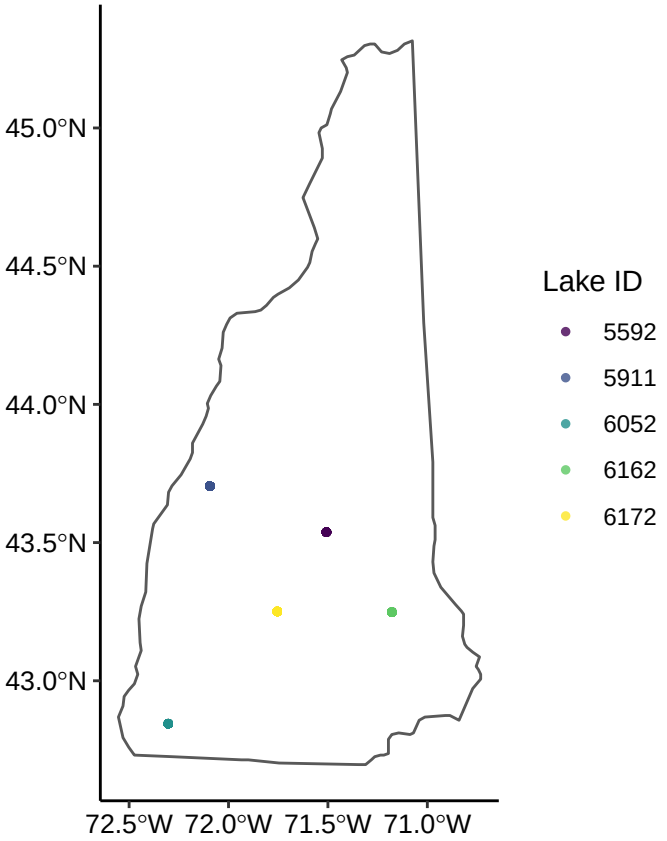


```
states.NH <- filter(states, ID == "new hampshire")

ggplot(states.NH) +
  geom_sf(fill = "white") #double check that it'll map
```

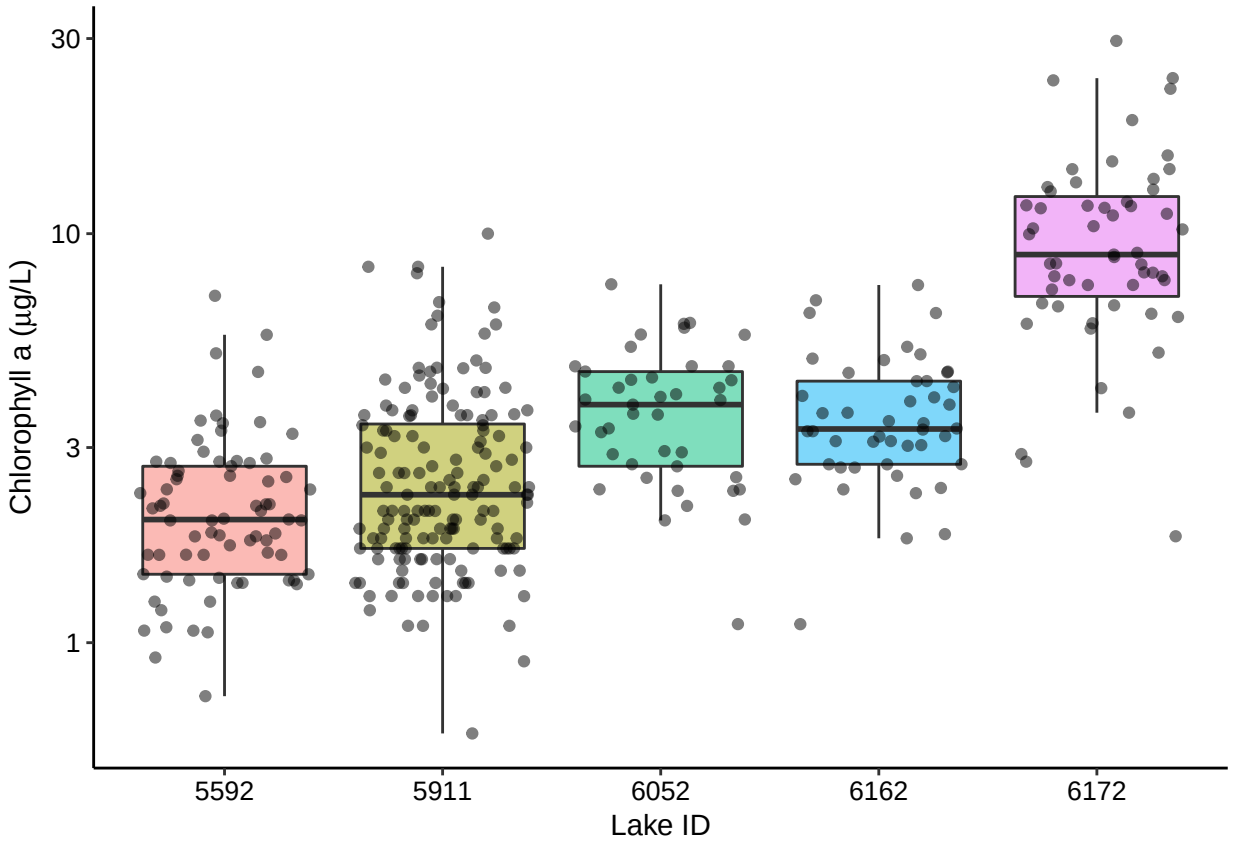



```
LAGOS.NH <- LAGOScombined %>%
  group_by(lagoslakeid) %>%
  filter(state == "NH") %>%
  drop_na(chla)
```

```
#choose 5 random lakes
samplelakes <- sample(x = LAGOS.NH$lagoslakeid, size = 5)

NH.Sample <- LAGOS.NH %>%
  filter(lagoslakeid %in% samplelakes)


#map it up, fuzzball
NH.sample.spatial <- st_as_sf(NH.Sample, coords= c("nhd_long", "nhd_lat"),
                              crs = 4326)

ggplot() +
  geom_sf(data = states.NH, fill = "white") +
  geom_sf(data = NH.sample.spatial, aes(color = as.factor(lagoslakeid)),
          alpha = 0.8, size = 1)+
  scale_color_viridis_d() +
  labs(color = "Lake ID")
```




```
ggplot(NH.Sample, aes(x = as.factor(lagoslakeid), y = chla,
                      fill = as.factor(lagoslakeid))) +
  geom_boxplot(outlier.shape = NA, alpha = 0.5) +
  geom_jitter(alpha = 0.5) +
  scale_y_log10() +
  scale_color_viridis_d() +
  labs(x = "Lake ID", y = expression("Chlorophyll a ("*mu*"g/L)", fill = ""))+
  theme(legend.position = "none")
```




```
#I can't get the colors to match, even though I'm using the same color scale
```

NHDPlus

3. Delineate and map the watershed and flowpaths in the network upstream of the Eno River gage in Durham (USGS-02085070). Your map should include all stream orders.

```
#what do we have?
ENOparams <- whatNWISdata(siteNumbers = "02085070")

#where is it
ENOcoords <- ENOparams %>%
  select(site_no, dec_lat_va, dec_long_va) %>%
  distinct()

#define a starting point
start_point <- st_sfc(st_point(c(ENOcoords$dec_long_va, ENOcoords$dec_lat_va)),
                      crs = 4326)
start_comid <- discover_nhdplus_id(start_point)

#navigate the NLDI
NLDI <- navigate_nldi(list(featureSource = "comid", featureID = start_comid),
                      mode = "upstreamTributaries",
                      distance_km = 1000)

#watershed and flowpath - this is just the syntax that nhdplus needs
subset_file <- tempfile(fileext = ".gpkg") #needs the "."
```

```
subset <- subset_nhdplus(comids = as.integer(NLDI$UT$nhdplus_comid),
                         output_file = subset_file,
                         nhdplus_data = "download",
                         flowline_only = FALSE,
                         return_data = TRUE,
                         overwrite = FALSE)
```

```
## All intersections performed in latitude/longitude.
```

```
## Reading NHDFlowline_Network
```

```
## Writing NHDFlowline_Network
```

```
## Reading CatchmentSP
```

```
## Writing CatchmentSP
```

```
## Warning: No nhdarea features found
```

```
#create data frames from the subset list
flowlines <- subset$NHDFlowline_Network
catchment <- subset$CatchmentSP
# double check they each are sf, tbl, and df classes
#class(flowlines)
#class(catchment)


#find gages near/in watershed
gages <- get_nwis(AOI = catchment) #AOI is area of interest
```

```
## Warning in if (AOI_type == "POINT") {: the condition has length > 1 and only the
## first element will be used
```

```
## Warning in if (AOI_type == "POINT") {: the condition has length > 1 and only the
## first element will be used
```

```
class(gages) #also as sf and df
```

```
## [1] "sf"         "data.frame"
```

```
gages <- st_intersection(gages, catchment) #now ONLY gages IN catchment
```

```
## Warning: attribute variables are assumed to be spatially constant throughout all
## geometries
```

```
#now map it out
plot_nhdplus("USGS-02085070") #first way to map - easiest and broadest
```

```
## Found invalid geometry, attempting to fix.
```

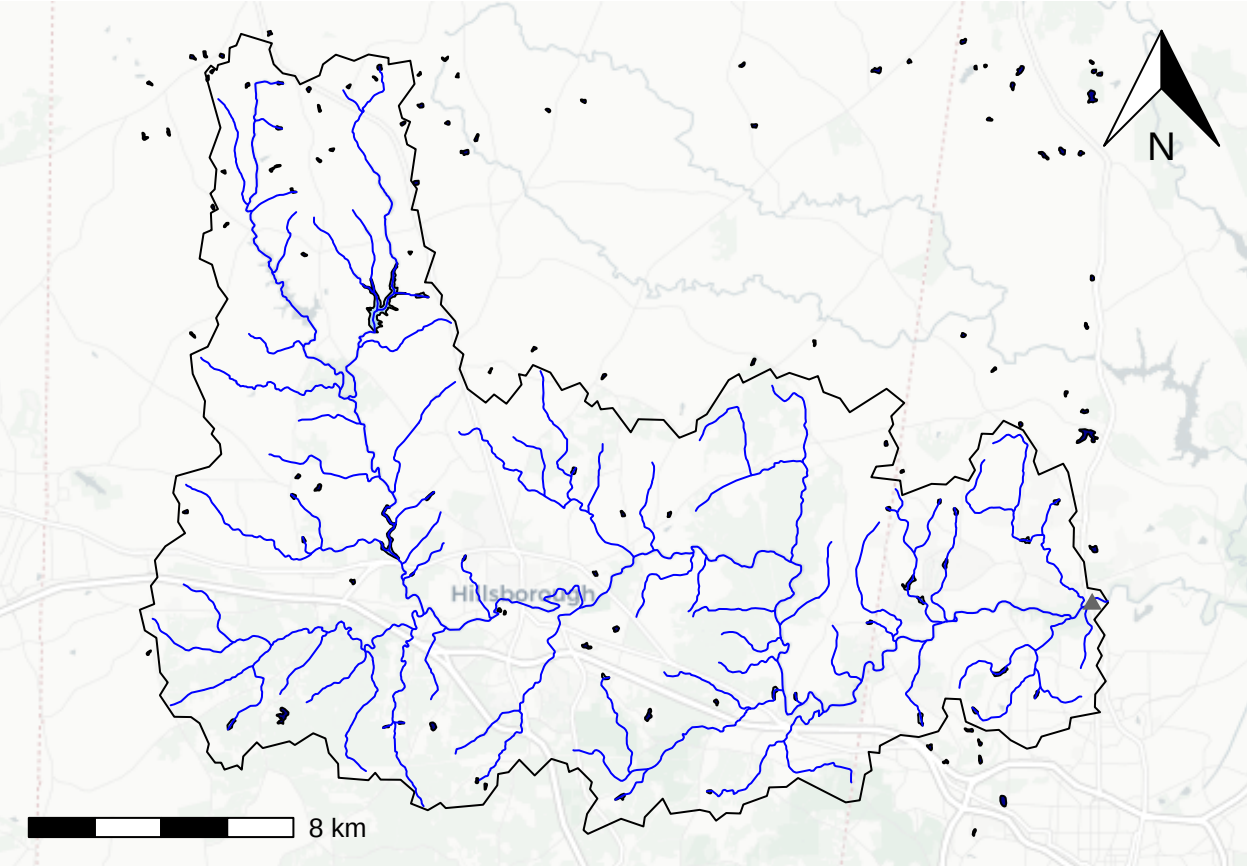
```
## Warning: No nhdarea features found
```

```
## Warning in showSRID(uprojargs, format = "PROJ", multiline = "NO", prefer_proj =
## prefer_proj): Discarded ellps WGS 84 in Proj4 definition: +proj=merc +a=6378137
## +b=6378137 +lat_ts=0 +lon_0=0 +x_0=0 +y_0=0 +k=1 +units=m +nadgrids=@null
## +wktext +no_defs
```

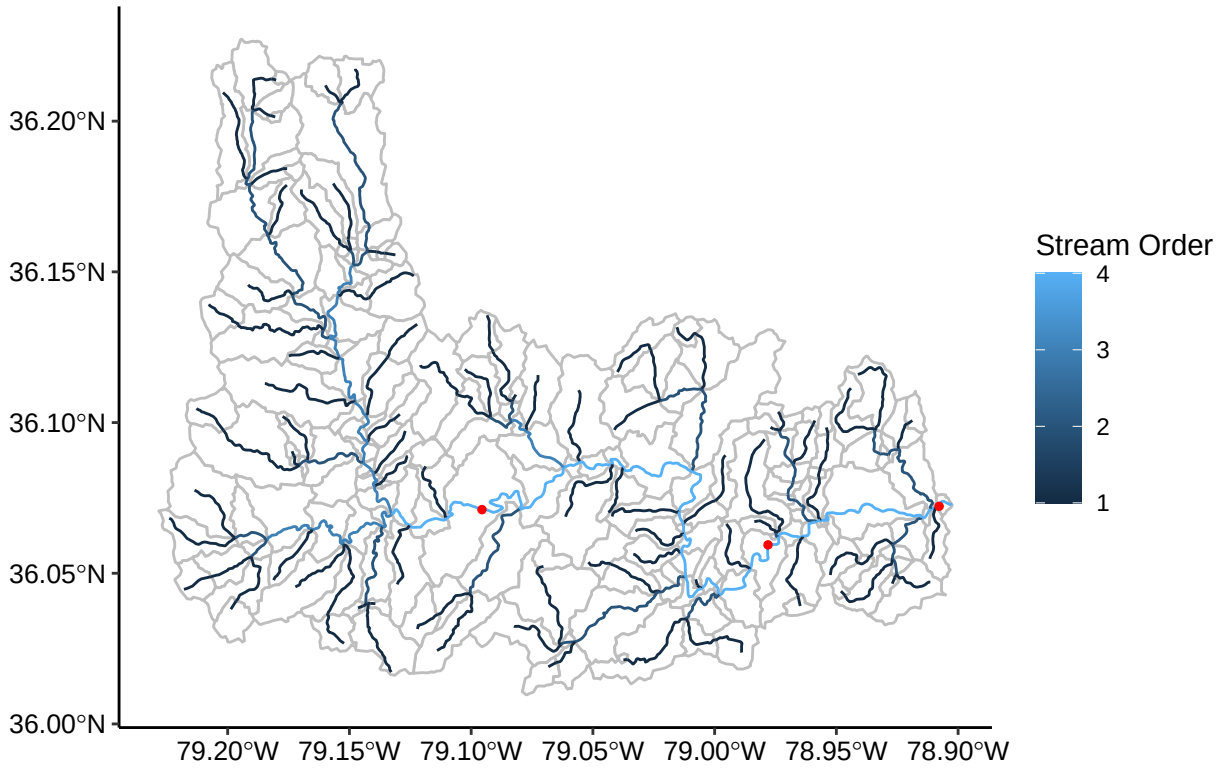
```
## Warning in showSRID(uprojargs, format = "PROJ", multiline = "NO", prefer_proj =
## prefer_proj): Discarded datum WGS_1984 in Proj4 definition
```

```
## Zoom: 11
```

```
## Map tiles by Carto, under CC BY 3.0. Data by OpenStreetMap, under ODbL.
## Audotdetect projection: assuming Google Mercator (epsg 3857)
```




```
ggplot() + #ggplot version, can control more aesthetics
  geom_sf(data = catchment, fill = "white", color = "grey", lwd = 0.5) +
  geom_sf(data = flowlines, aes(color = streamorde))+
  geom_sf(data = gages, color = "red", size = 1) +
  labs(color = "Stream Order")
```




```
#analysis
max(flowlines$totdasqkm) #total drainage area in km^2
```

```
## [1] 367.6815
```

```
summary(as.factor(flowlines$streamorde))
```

```
##   1   2   3   4
## 123  45  19  28
```

```
view(gages)
```

4. What is the upstream drainage area for this gage, and how are the flowpaths distributed with regard to stream order?

   The upstream drainage area is 367.68 square kilometers. Flowpaths are distributed with 123 first-order streams, 45 second-order streams, 19 thrid-order streams, and 28 fourth-order streams.

5. Are there any other gage sites in the upstream catchment? If so, where are they? > Yes there are two more, in Hillsborough and at Cole Mill Rd near Huckleberry Spring.