ECOL 620 Lab 1

Natalie Schmer

1/21/2021

# Load in packages

{  
 library(tidyverse)  
 library(rgdal)  
 library(viridis)  
 library(tidylog)  
 library(ggthemes)  
 library(ggrepel)  
 library(here)  
 library(sp)  
 library(cowplot)  
}  
  
#need to download old usmap   
install.packages("/Users/natalieschmer/Desktop/GitHub/ECOL\_620/data/usmap\_0.5.1.tar", repos = NULL, type="source")

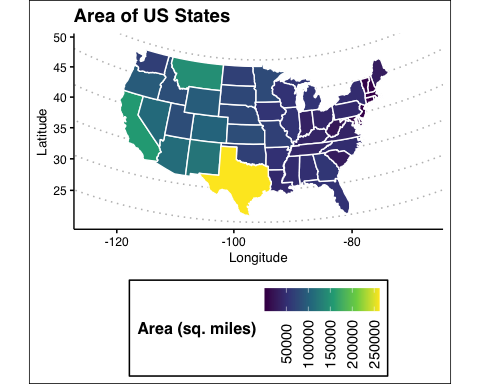
# 1. Briefly describe (1-2 sentences) what each of the following functions achieve.

* c() is for concatenate, and combines the arguments into a list or vector.
* geom\_smooth() is in {ggplot2} and draws a line through data to help show the pattern of the data, helps a lot when there is a lot on the plot. Need to specify the formula and fit arguments.
* rep() replicates values as specified in the arguments.
* filter() in {dplyr} subsets dataframe rows using values in columns that satisfy the given arguments.
* %>%() is a pipe! From {magrittr}, moves whatever object is on the left to the operation on the right.
* png() writes an object (like a plot) to a .png file. Need to specify file path as an argument and can also set height, width, other aesthetics.
* coord\_map() is for map projection in {ggplot2}.
* plot\_grid() from {cowplot} allows for arranging multiple plots in a grid. Can specify number of rows, columns, adjust alignments, labels with a variety of arguments.

# US Maps

## 2. Ploting the contiguous U.S. states colored by area

# Load the data we used earlier, join states to df with areas  
states <- as.data.frame(state.x77)  
states$region <- tolower(rownames(states))  
states\_map <- map\_data("state")  
fact\_join <- left\_join(states\_map, states, by = "region")  
  
(state\_areas <- ggplot(fact\_join, aes(long, lat, group = group))+  
 geom\_polygon(aes(fill = Area), colour = "white")+  
 scale\_fill\_viridis\_c(option = "D")+  
 ggthemes::theme\_clean() +  
 coord\_map("bonne", lat0 = 40)+  
 labs(y = "Latitude",   
 x = "Longitude",   
 fill = "Area (sq. miles)",  
 title = "Area of US States")+  
 theme(legend.position = "bottom",  
 legend.text = element\_text(angle = 90, vjust = 0.5))  
)



## 3 Three additional plots of the same data, each with different color scale and geographic projection, put together with plot\_grid function to make a three-panel horizontal plot.

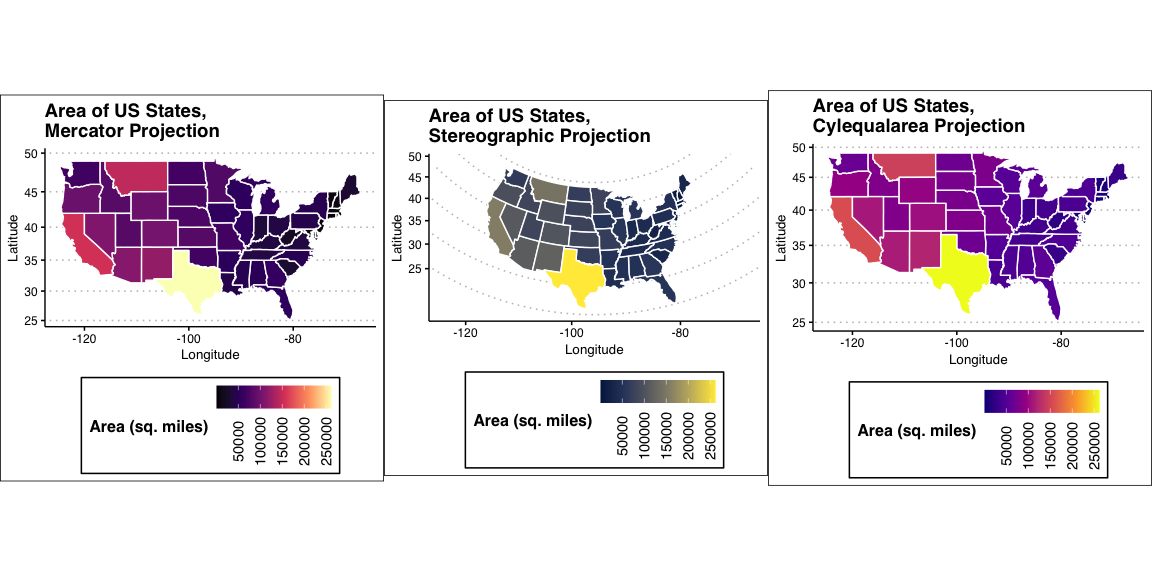
#1   
(state\_areas\_1 <- ggplot(fact\_join, aes(long, lat, group = group))+  
 geom\_polygon(aes(fill = Area), colour = "white")+  
 scale\_fill\_viridis\_c(option = "A")+  
 ggthemes::theme\_clean() +  
 coord\_map("mercator")+  
 labs(y = "Latitude",   
 x = "Longitude",   
 fill = "Area (sq. miles)",  
 title = "Area of US States, \nMercator Projection")+  
 theme(legend.position = "bottom",  
 legend.text = element\_text(angle = 90, vjust = 0.5))  
)

#2  
(state\_areas\_2 <- ggplot(fact\_join, aes(long, lat, group = group))+  
 geom\_polygon(aes(fill = Area), colour = "white")+  
 scale\_fill\_viridis\_c(option = "E")+  
 ggthemes::theme\_clean() +  
 coord\_map("stereographic")+  
 labs(y = "Latitude",   
 x = "Longitude",   
 fill = "Area (sq. miles)",  
 title = "Area of US States, \nStereographic Projection")+  
 theme(legend.position = "bottom",  
 legend.text = element\_text(angle = 90, vjust = 0.5))  
)

#3  
(state\_areas\_3 <- ggplot(fact\_join, aes(long, lat, group = group))+  
 geom\_polygon(aes(fill = Area), colour = "white")+  
 scale\_fill\_viridis\_c(option = "C")+  
 ggthemes::theme\_clean() +  
 coord\_map("cylequalarea", lat0 = 40)+  
 labs(y = "Latitude",   
 x = "Longitude",   
 fill = "Area (sq. miles)",  
 title = "Area of US States,\nCylequalarea Projection")+  
 theme(legend.position = "bottom",  
 legend.text = element\_text(angle = 90, vjust = 0.5))  
)

### 3-panel of the different color scales and projections

plot\_grid(state\_areas\_1, state\_areas\_2, state\_areas\_3, nrow = 1)



# US University Data

## 4. Plot Colorado colleges and universities within the top 96th quantile of total student enrollment, add in geom\_text\_repel(data=NAME\_OF\_YOUR\_FILTERED\_DATA, aes(label=NAME, x=LON, y = LAT), force=20, size=2)

# Read in the csv of university data   
colo\_unis <- read.csv(here::here("data/data\_for\_lab1/Colleges\_and\_Universities/colorado\_universities.csv"))   
  
# Filter to only CO   
colo\_unis <- colo\_unis %>%   
 filter(LSTATE == "CO")  
# check   
unique(colo\_unis$LSTATE)

## [1] "CO"

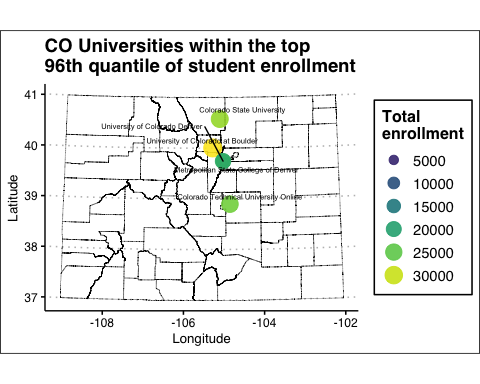
# Calculate and filter the 96th quantile universities  
(quan\_96 <- quantile(colo\_unis$TOT\_ENROLL, .96, na.rm = T))

## 96%   
## 16616.8

co\_96 <- colo\_unis %>%   
 filter(TOT\_ENROLL >= quan\_96)  
  
# Load in the CO shapefile   
co\_counties <- readOGR(here::here("data/data\_for\_lab1/counties/Colorado\_County\_Boundaries.shp"))

## OGR data source with driver: ESRI Shapefile   
## Source: "/Users/natalieschmer/Desktop/GitHub/ECOL\_620/data/data\_for\_lab1/counties/Colorado\_County\_Boundaries.shp", layer: "Colorado\_County\_Boundaries"  
## with 64 features  
## It has 9 fields  
## Integer64 fields read as strings: OBJECTID NUM\_FIPS

# Set the min and max enroll  
min\_enroll <- min(colo\_unis$TOT\_ENROLL)  
max\_enroll <- max(colo\_unis$TOT\_ENROLL)  
  
# Plot   
ggplot() +  
 geom\_polygon(data = co\_counties,   
 aes(x = long,   
 y = lat,   
 group = group),   
 fill = NA,   
 color ="black",   
 lwd =.1) +  
 geom\_point(data = co\_96,   
 aes(x = LON,   
 y = LAT,   
 size = TOT\_ENROLL,   
 colour = TOT\_ENROLL),   
 alpha =.9) +  
 coord\_map("bonne", lat0 = 40)+  
 theme\_bw()+  
 theme(panel.grid.minor = element\_blank(),  
 panel.grid.major = element\_blank())+  
 theme(panel.border=element\_blank())+  
 scale\_color\_viridis(limits = c(min\_enroll, max\_enroll),  
 breaks = seq(5000, 30000, by=5000),   
 name = "Total\nenrollment")+  
 guides(color= guide\_legend(), size=guide\_legend())+  
 scale\_size\_continuous(limits = c(min\_enroll,max\_enroll),   
 breaks = seq(5000, 30000, by=5000),  
 name = "Total\nenrollment") +  
 geom\_text\_repel(data = co\_96,   
 aes(label = NAME,   
 x = LON,   
 y = LAT),   
 force = 20, size = 2)+  
 theme\_clean()+  
 labs(y = "Latitude",   
 x = "Longitude",  
 title = "CO Universities within the top\n96th quantile of student enrollment")



## 5. Other universities: On, Wisconsin!

Wisconsin county boundary shapefile downloaded from <https://data-wi-dnr.opendata.arcgis.com/datasets/county-boundaries-24k>

# load in shapefiles for WI counties and universities  
wi\_counties <- readOGR(here::here("data/County\_Boundaries\_24K-shp/County\_Boundaries\_24K.shp"))

## OGR data source with driver: ESRI Shapefile   
## Source: "/Users/natalieschmer/Desktop/GitHub/ECOL\_620/data/County\_Boundaries\_24K-shp/County\_Boundaries\_24K.shp", layer: "County\_Boundaries\_24K"  
## with 72 features  
## It has 7 fields

us\_uni <- readOGR(here::here("data/data\_for\_lab1/Colleges\_and\_Universities/CollegesUniversities.shp"))

## OGR data source with driver: ESRI Shapefile   
## Source: "/Users/natalieschmer/Desktop/GitHub/ECOL\_620/data/data\_for\_lab1/Colleges\_and\_Universities/CollegesUniversities.shp", layer: "CollegesUniversities"  
## with 7126 features  
## It has 38 fields

# Filter out WI schools  
wi\_uni <- subset(us\_uni, LSTATE=="WI")  
  
# check crs   
proj4string(wi\_counties)

## [1] "+proj=tmerc +lat\_0=0 +lon\_0=-90 +k=0.9996 +x\_0=520000 +y\_0=-4480000 +ellps=GRS80 +towgs84=0,0,0,0,0,0,0 +units=m +no\_defs"

proj4string(wi\_uni)

## [1] "+proj=longlat +datum=WGS84 +no\_defs"

# reproject counties with the uni projection   
wi\_counties\_reprj <- spTransform(wi\_counties, "+proj=longlat +datum=WGS84 +no\_defs")  
  
# make wi unis a df   
wi\_unis\_df <- as.data.frame(wi\_uni)  
  
#remove ins\_size =<0 as those are not reported or not applicable  
wi\_unis\_df <- wi\_unis\_df %>%   
 dplyr::filter(INST\_SIZE > 0)

### WI Plot

ggplot() +  
 geom\_polygon(data = wi\_counties\_reprj,   
 aes(x = long,   
 y = lat,   
 group = group),  
 fill = NA, color ="black", lwd=.1) +  
 geom\_point(data = wi\_unis\_df %>% filter(NAME != "University of Wisconsin-Madison"),   
 aes(x = LON,   
 y = LAT,   
 size=INST\_SIZE,   
 colour=INST\_SIZE),   
 alpha=.7) +   
 geom\_point(data = wi\_unis\_df %>% filter(NAME == "University of Wisconsin-Madison"),   
 aes(x = LON,   
 y = LAT,   
 size=INST\_SIZE),   
 alpha=.6, color = "blue", show.legend = F) +  
 coord\_map("bonne", lat0 = 40)+  
 theme\_clean()+  
 theme(panel.grid.minor=element\_blank(),  
 panel.grid.major=element\_blank())+  
 theme(panel.border=element\_blank())+  
 scale\_color\_viridis(limits = c(1, 5),  
 labels = c("Under 1000",  
 "1000 - 4999",  
 "5000 - 9999",  
 "10000 - 19999",  
 "20000 and above"),  
 name = "Institution\nsize") +  
 guides(color= guide\_legend(), size=guide\_legend()) +  
 scale\_size\_continuous(limits=c(1, 5),  
 breaks=seq(1, 5, by= 1),  
 labels = c("Under 1000",  
 "1000 - 4999",  
 "5000 - 9999",  
 "10000 - 19999",  
 "20000 and above"),  
 name = "Institution\nsize")+  
 labs(y = "Latitude",   
 x = "Longitude",  
 title = "University of Wisconsin-Madison")

