#VA’s 100 Best Beers

library(rvest)

#Define webpage information is coming from

url <- "https://www.beeradvocate.com/lists/state/va/"

webpage <- read\_html(url)

#Scrape Rank data

rank\_data\_html <- html\_nodes(webpage, '.hr\_bottom\_light:nth-child(3) b')

#Convert Rank data to text and define as number

rank\_data <- html\_text(rank\_data\_html)

rank\_data <- as.numeric(rank\_data)

#Scrape Ratings data

ratings\_data\_html <- html\_nodes(webpage, '.hr\_bottom\_light:nth-child(4) b')

#Convert Rating data to text and define as number

ratings\_data <- html\_text(ratings\_data\_html)

ratings\_data <- as.numeric(ratings\_data)

#Scrape Name of Beer

Name\_of\_beer\_html <- html\_nodes(webpage, '.hr\_bottom\_light a b')

#Convert Name data to text and change to factor

name\_of\_beer\_data <- html\_text(Name\_of\_beer\_html)

name\_of\_beer\_data <- as.factor(name\_of\_beer\_data)

#Scrape Producer

producer\_data\_html <- html\_nodes(webpage, '#extendedInfo a:nth-child(1)')

#Convert producer data to text and change to factor

producer\_data <- html\_text(producer\_data\_html)

producer\_data <- as.factor(producer\_data)

#Scrape Style of beer data

style\_of\_beer\_html <- html\_nodes(webpage, '#extendedInfo br+ a')

#Convert style of beer to text and change to factor

style\_of\_beer\_data <- html\_text(style\_of\_beer\_html)

style\_of\_beer\_data <- as.factor(style\_of\_beer\_data)

#Scrape ABV content

abv\_html <- html\_nodes(webpage, '#extendedInfo')

#Convert abv to text

abv\_data <- html\_text(abv\_html)

#Clean abv data

abv\_data <- gsub(".\*/","",abv\_data)

abv\_data <- gsub("%.\*","",abv\_data)

#Change abv data to numeric

abv\_data <- as.numeric(abv\_data)

#Remove html, url ,and webpage from workspace; I like a clean workspace

rm(abv\_html,Name\_of\_beer\_html,producer\_data\_html,rank\_data\_html,ratings\_data\_html,style\_of\_beer\_html,url,webpage)

#Convert variables into single data frame

Breweries.df <- data.frame(Beer = name\_of\_beer\_data, Style = style\_of\_beer\_data, ABV = abv\_data, Rank = rank\_data, Ratings = ratings\_data, Brewery = producer\_data)

#Import Brewery\_address\_and\_geocoords spreadsheet

#Merge spreadsheet with Breweries.df

Breweries <- merge(Breweries.df, Brewery\_address\_and\_geocoords, by = "Brewery")

#Map Breweries

library(ggplot2)

library(ggmap)

library(qmap)

map <- qmap('Virginia', zoom = 7, maptype = 'hybrid')

map + geom\_point(data = Breweries, aes(x=lon, y=lat), size = 2, alpha = 0.5, col = 'red')