COVID19 Final Project

Jared, Matt, Megan

Spring 2022

knitr::opts\_chunk$set(echo = TRUE,  
 fig.width=10,   
 fig.height=6,   
 fig.align = "center")  
  
# Load the needed package(s) here  
pacman::p\_load(tidyverse, COVID19,gridExtra, usmap)  
cov19 <- covid19(country = c("US"), level = 2, start = "2020-12-01", end = "2021-06-01")

## We have invested a lot of time and effort in creating COVID-19 Data Hub, please cite the following when using it:  
##   
## Guidotti, E., Ardia, D., (2020), "COVID-19 Data Hub", Journal of Open  
## Source Software 5(51):2376, doi: 10.21105/joss.02376.  
##   
## A BibTeX entry for LaTeX users is  
##   
## @Article{,  
## title = {COVID-19 Data Hub},  
## year = {2020},  
## doi = {10.21105/joss.02376},  
## author = {Emanuele Guidotti and David Ardia},  
## journal = {Journal of Open Source Software},  
## volume = {5},  
## number = {51},  
## pages = {2376},  
## }  
##   
## To hide this message use 'verbose = FALSE'.

#Renamed the administrative\_area\_level\_2 to state just for convenience and added three percentage variables  
cov19 <- rename(cov19, state = administrative\_area\_level\_2) %>%  
 mutate(deathRate = (deaths/confirmed)\*100,  
 infectionRate = (confirmed/population)\*100,  
 vaccineRate = (vaccines/population)\*100, na.rm = TRUE)  
  
# Change the default theme below:  
theme\_set(theme\_bw())

The next section is data cleaning

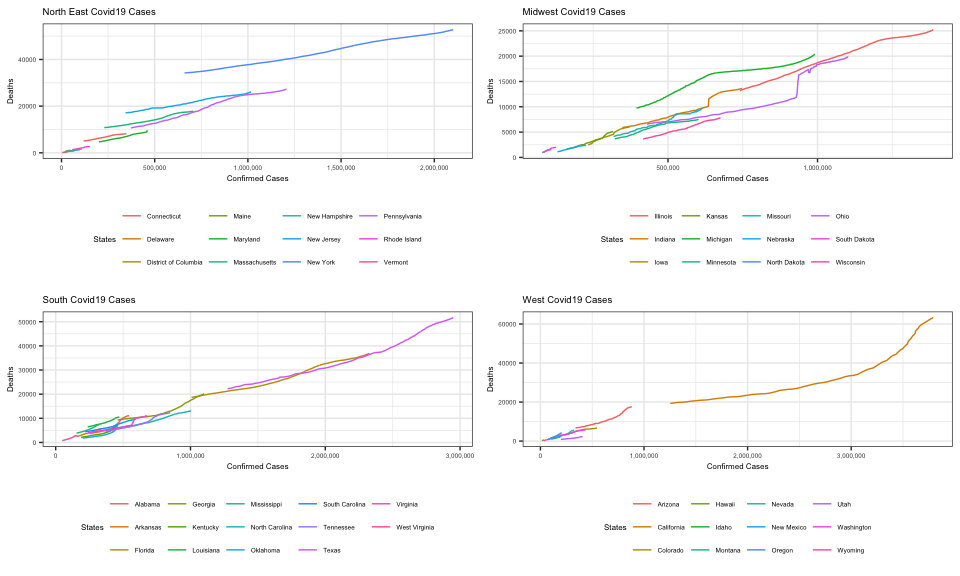
# total count of cov19 at last day of data  
cov19LastDay <- cov19 %>%  
 filter(date == "2021-06-01")  
# Cov19 by region  
cov19South <- cov19 %>%   
 filter(state %in%   
 c('Alabama', 'Florida', 'Georgia', 'Mississippi',   
 'South Carolina','Arkansas', 'Louisiana', 'Oklahoma',   
 'Texas', 'Kentucky', 'North Carolina', 'Tennessee', 'Virginia', 'West Virginia'))  
cov19NorthE <- cov19 %>%   
 filter(state %in%   
 c('Connecticut', 'Maine', 'Massachusetts', 'New Hampshire',   
 'Rhode Island', 'Vermont', 'New Jersey', 'New York', 'Pennsylvania',  
 'Delaware', 'District of Columbia','Maryland'))  
cov19MidW <- cov19 %>%  
 filter(state %in%  
 c('Iowa', 'Kansas', 'Missouri', 'Nebraska', 'North Dakota',   
 'South Dakota', 'Illinois', 'Indiana', 'Michigan', 'Minnesota', 'Ohio', 'Wisconsin'))  
cov19West <- cov19 %>%  
 filter(state %in%  
 c('California', 'Hawaii', 'Nevada','Oregon', 'Washington','Arizona',   
 'Colorado', 'Idaho', 'Montana', 'New Mexico', 'Utah', 'Wyoming'))

The next section is graphing

# North East plot  
northEastPlot <- ggplot(cov19NorthE, aes(x = confirmed, y = deaths, color = state))+  
 geom\_line()+  
 labs(title = "North East Covid19 Cases",  
 x = "Confirmed Cases",  
 y = "Deaths",  
 color = "States")+  
 theme(text=element\_text(size=6),  
 legend.position="bottom")+  
 scale\_x\_continuous(labels = scales::comma)  
  
# South Plot  
southPlot <- ggplot(cov19South, aes(x = confirmed, y = deaths, color = state))+  
 geom\_line()+  
 labs(title = "South Covid19 Cases",  
 x = "Confirmed Cases",  
 y = "Deaths",  
 color = "States")+  
 theme(text=element\_text(size=6),  
 legend.position="bottom")+  
 scale\_x\_continuous(labels = scales::comma)  
# Midwest plot  
midWestPlot <- ggplot(cov19MidW, aes(x = confirmed, y = deaths, color = state))+  
 geom\_line()+  
 labs(title = "Midwest Covid19 Cases",  
 x = "Confirmed Cases",  
 y = "Deaths",  
 color = "States")+  
 theme(text=element\_text(size=6),  
 legend.position="bottom")+  
 scale\_x\_continuous(labels = scales::comma)  
# West plot  
westPlot <- ggplot(cov19West, aes(x = confirmed, y = deaths, color = state))+  
 geom\_line()+  
 labs(title = "West Covid19 Cases",  
 x = "Confirmed Cases",  
 y = "Deaths",  
 color = "States")+  
 theme(text=element\_text(size=6),  
 legend.position="bottom")+  
 scale\_x\_continuous(labels = scales::comma)

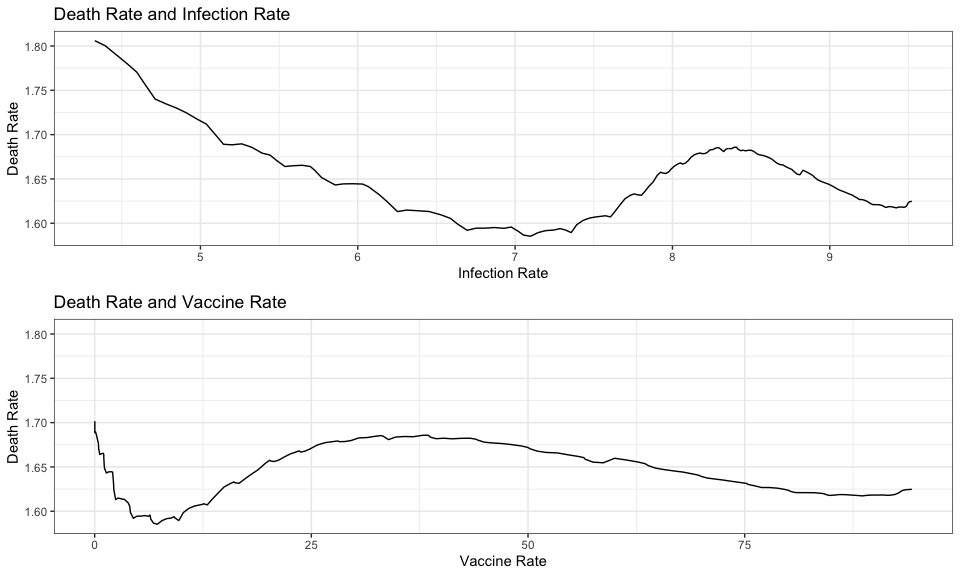
Grid arrange of the 4 plots above

grid.arrange(northEastPlot,midWestPlot, southPlot,westPlot, ncol = 2)



Plots for vaccine rate vs death rate and infection rate vs death rate with a grid arrange

# plot for death rate vs vaccine rate for each day  
cov19Average <- cov19 %>% group\_by(date) %>%  
 summarise(mean(infectionRate, na.rm = TRUE), mean(deathRate, na.rm = TRUE), mean(vaccineRate, na.rm = TRUE))  
  
deathVaccineRatesPlot <- ggplot(cov19Average, aes(x = cov19Average$`mean(vaccineRate, na.rm = TRUE)`, y = cov19Average$`mean(deathRate, na.rm = TRUE)`))+  
 geom\_line()+  
 labs(x = "Vaccine Rate",  
 y = "Death Rate",  
 title = "Death Rate and Vaccine Rate")  
  
deathInfectionRatesPlot <- ggplot(cov19Average, aes(x = cov19Average$`mean(infectionRate, na.rm = TRUE)`, y = cov19Average$`mean(deathRate, na.rm = TRUE)`))+  
 geom\_line()+  
 labs(x = "Infection Rate",  
 y = "Death Rate",  
 title = "Death Rate and Infection Rate")  
  
  
# Grid Arrange of two graphs  
grid.arrange(deathInfectionRatesPlot, deathVaccineRatesPlot)



Map of US with infection as the value

#filter out non states  
  
cov19StatesOnly <- cov19 %>% group\_by(state) %>% filter(state != "American Samoa", state != "Puerto Rico", state != "Northern Mariana Islands", state != "District of Columbia", state != "Guam", state != "Virgin Islands")  
  
## Map of covid cases  
  
plot\_usmap(data = cov19StatesOnly,  
 values = 'infectionRate',  
 regions = 'state',  
 color = "black")+  
 scale\_fill\_continuous(low = 'light blue', high = 'red', limits = c(0,15))+  
 labs(fill = "Infection Rate",  
 title = "Covid 19 Infection Rate")+  
 theme(plot.title = element\_text(hjust = .5, size = 16))

