Analysis

Prashan A. Welipitiya

This markdown is just an overview of the proccess that I used to get my information.

Data

I'm going to be pulling the covid numbers from the New York Times Github. They have total numbers of cases and deaths per county and per state.

```
#https://github.com/nytimes/covid-19-data
setwd("C:/Users/Prashan.Welipitiya/Desktop/covid-19-data")
covid_states = read.csv("us-states.csv")
covid_states$date <- ymd(covid_states$date)
covid_counties = read.csv("us-counties.csv")
covid_counties$date <- ymd(covid_counties$date)

head(covid_counties)
```

```
##         date    county      state  fips cases deaths
## 1 2020-01-21 Snohomish Washington 53061     1      0
## 2 2020-01-22 Snohomish Washington 53061     1      0
## 3 2020-01-23 Snohomish Washington 53061     1      0
## 4 2020-01-24      Cook   Illinois 17031     1      0
## 5 2020-01-24 Snohomish Washington 53061     1      0
## 6 2020-01-25    Orange California  6059     1      0
```

I'm going to be using a dataset that was part of a homework assignments in an old class. This dataset has a lot of important information that I am curious about on counties in the US. It includes percent populations of 2016 voting information, elderly, black, white, hispanic, asian, education and income.

```
county_votes16 <- readRDS(url("https://ericwfox.github.io/data/county_votes16.rds"))
head(county_votes16)
```

```
##   state         county clinton_pctvotes trump_pctvotes obama_pctvotes pct_pop65
## 1    AL Autauga County            23.96          73.44          26.58      13.8
## 2    AL Baldwin County            19.57          77.35          21.57      18.7
## 3    AL Barbour County            46.66          52.27          51.25      16.5
## 4    AL    Bibb County            21.42          76.97          26.22      14.8
## 5    AL  Blount County             8.47          89.85          12.35      17.0
## 6    AL Bullock County            75.09          24.23          76.31      14.9
##   pct_black pct_white pct_hispanic pct_asian highschool bachelors income
## 1      18.7      77.9          2.7       1.1       85.6      20.9 53.682
## 2       9.6      87.1          4.6       0.9       89.1      27.7 50.221
```

```
## 3      47.6      50.2          4.5       0.5       73.7      13.4 32.911
## 4      22.1      76.3          2.1       0.2       77.5      12.1 36.447
## 5       1.8      96.0          8.7       0.3       77.0      12.1 44.145
## 6      70.1      26.9          7.5       0.3       67.8      12.5 32.033
##   trump_win
## 1         1
## 2         1
## 3         1
## 4         1
## 5         1
## 6         0
```

```
#summary(county_votes16)
```

```
# To match the New York Times data, I'm going to add a collumn that changes the state abbreviation to t
county_votes16$state_Name <- state.name[match(county_votes16$state,state.abb)]

# And take the word county out of the county names.
county_votes16$county_Name <- as.character(county_votes16$county)
county_votes16$county_Name <- substr(county_votes16$county_Name,1,nchar(county_votes16$county_Name) - 7

head(county_votes16)
```

```
##   state         county clinton_pctvotes trump_pctvotes obama_pctvotes pct_pop65
## 1    AL Autauga County            23.96          73.44          26.58      13.8
## 2    AL Baldwin County            19.57          77.35          21.57      18.7
## 3    AL Barbour County            46.66          52.27          51.25      16.5
## 4    AL    Bibb County            21.42          76.97          26.22      14.8
## 5    AL  Blount County             8.47          89.85          12.35      17.0
## 6    AL Bullock County            75.09          24.23          76.31      14.9
##   pct_black pct_white pct_hispanic pct_asian highschool bachelors income
## 1      18.7      77.9          2.7       1.1       85.6      20.9 53.682
## 2       9.6      87.1          4.6       0.9       89.1      27.7 50.221
## 3      47.6      50.2          4.5       0.5       73.7      13.4 32.911
## 4      22.1      76.3          2.1       0.2       77.5      12.1 36.447
## 5       1.8      96.0          8.7       0.3       77.0      12.1 44.145
## 6      70.1      26.9          7.5       0.3       67.8      12.5 32.033
##   trump_win state_Name county_Name
## 1         1    Alabama     Autauga
## 2         1    Alabama     Baldwin
## 3         1    Alabama     Barbour
## 4         1    Alabama        Bibb
## 5         1    Alabama      Blount
## 6         0    Alabama     Bullock
```