Map Activity

Libraries Used

```
library(colorspace)
library(broom)
library(tidyverse)
library(tmap)
library(maps)
library(sf)
```

The Data

First, we will read in a pre-downloaded set of data from the Census API.

```
census_data <- read_csv("census_data.csv")
```

Then, let's view the first 10 rows:

```
head(census_data, n = 10)
```

```
## # A tibble: 10 x 18
##    county_fips state_abbv county_name county state rShift popEstimate16
##          <dbl> <chr>      <chr>       <chr>  <chr>  <dbl>         <dbl>
##  1        1001 AL         Autauga Co~ Autau~ Alab~  0.802         55278
##  2        1001 AL         Autauga Co~ Autau~ Alab~  0.802         55278
##  3        1001 AL         Autauga Co~ Autau~ Alab~  0.802         55278
##  4        1001 AL         Autauga Co~ Autau~ Alab~  0.802         55278
##  5        1001 AL         Autauga Co~ Autau~ Alab~  0.802         55278
##  6        1001 AL         Autauga Co~ Autau~ Alab~  0.802         55278
##  7        1001 AL         Autauga Co~ Autau~ Alab~  0.802         55278
##  8        1001 AL         Autauga Co~ Autau~ Alab~  0.802         55278
##  9        1001 AL         Autauga Co~ Autau~ Alab~  0.802         55278
## 10        1001 AL         Autauga Co~ Autau~ Alab~  0.802         55278
## # ... with 11 more variables: unemployment <dbl>, medianIncome16 <dbl>,
## #   stateAbbrv <chr>, prcntGOP16 <dbl>, long <dbl>, lat <dbl>, group <dbl>,
## #   whiteNotLatino <dbl>, black <dbl>, domesticMig16 <dbl>,
## #   log_popEstimate16 <dbl>
```

This dataset contains information on county (by state) population estimates, median income, unemployment, and political shifts to the right of the spectrum.

Publications and research reports that use this data must cite it appropriately by including the following information:

- Ella Foster-Molina and Ben Warren. Partisan Voting, County Demographics, and Deaths of Despair Data. February 2019.

The codebook for the dataset is:

- county_fips: Five digit Federal Information Processing Standards code that uniquely identifies counties and county equivalents in the United States
- state_abbv: State postal abbreviation
- county_name: County name, may be identical to county variable
- county: County name
- state: State full name
- rshift: Percentage difference between the Republican presidential vote in that county in 2016 and 2012. For example, 46.7955% of Kent County in Delaware (FIPS 20001) voted for Romney in 2012. In 2016, 49.81482% of that county voted for Trump. Therefore, the county shifted towards the Republican presidential candidate by 3.01325%. Positive value mean leaning more Republican; negative values mean leaning less Republican.
- popEstimate16: Population in the county in 2016
- unemployment: Unemployment rate in 2016
- medianIncome16: Median household income in the county in 2016.
- prcntGOP16: Percent of the county that voted for the Republican presidential candidate, Donald Trump, in 2016.

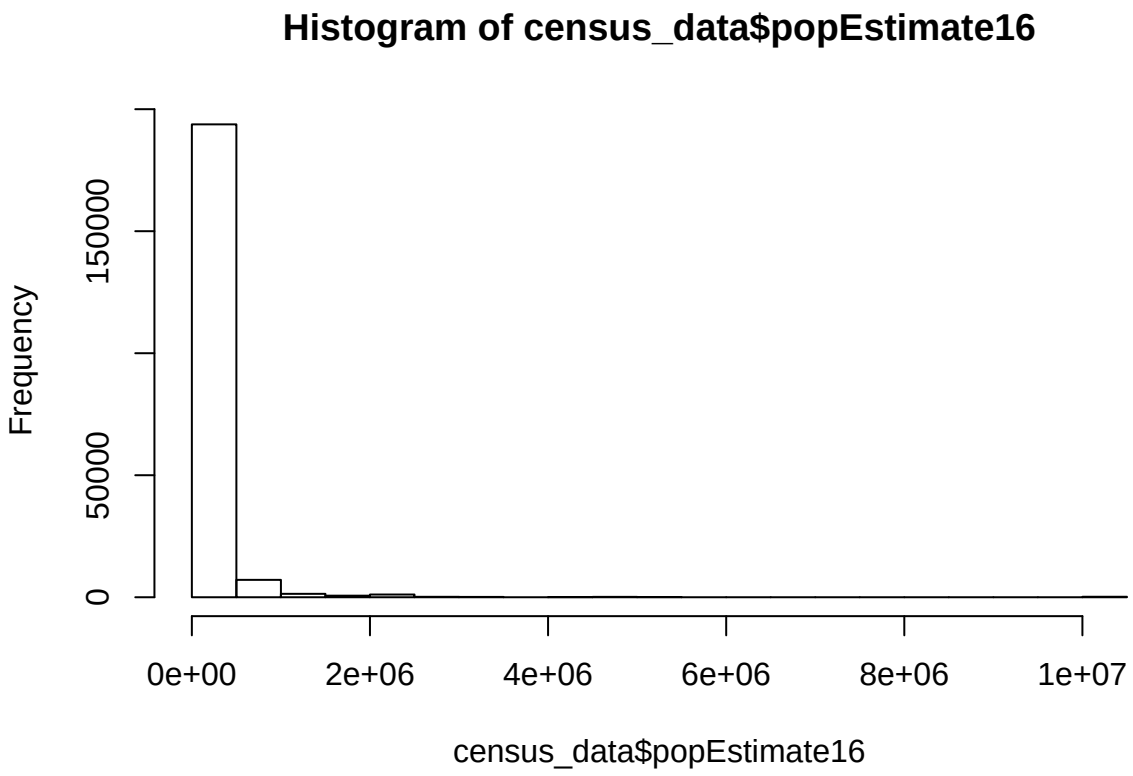
Graphics

Once we have the data set imported to R, we can begin to visualize some of the data.

Histogram

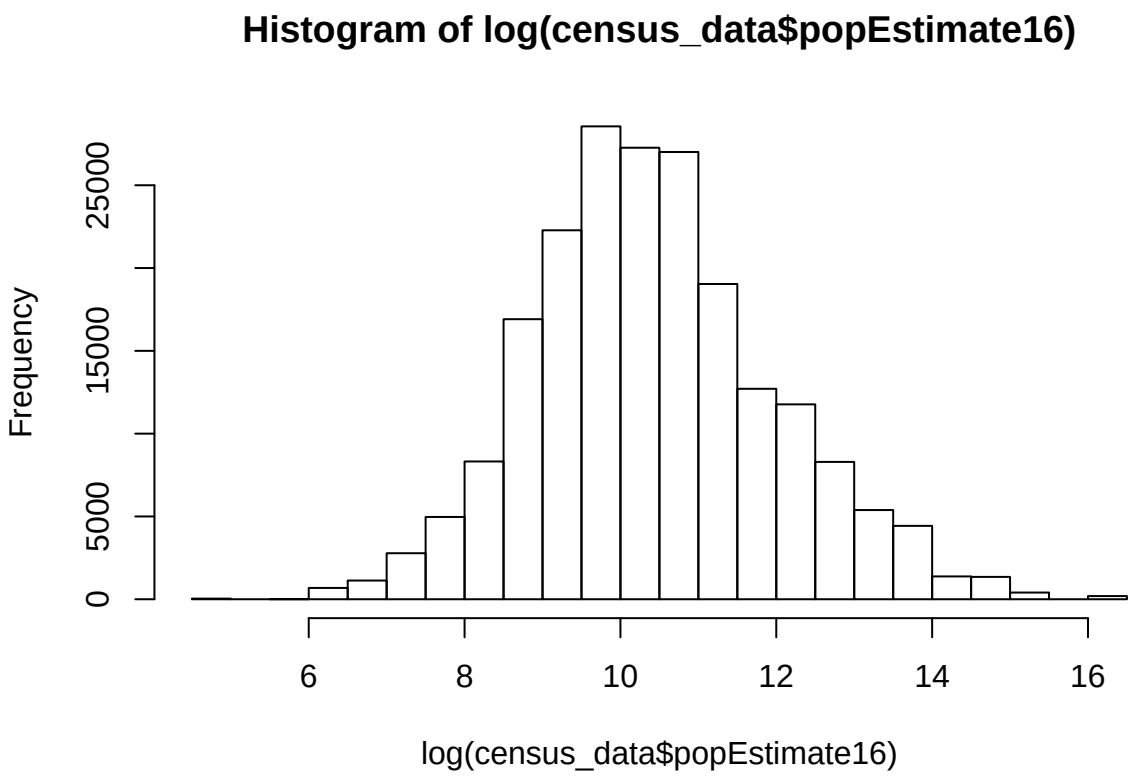
It is quite simple to make a simple histogram using the hist() function.

```
hist(census_data$popEstimate16)
```



However, it might be helpful to change the scale of the variable. We can do that with an easy call to the log() function.

```
hist(log(census_data$popEstimate16))
```

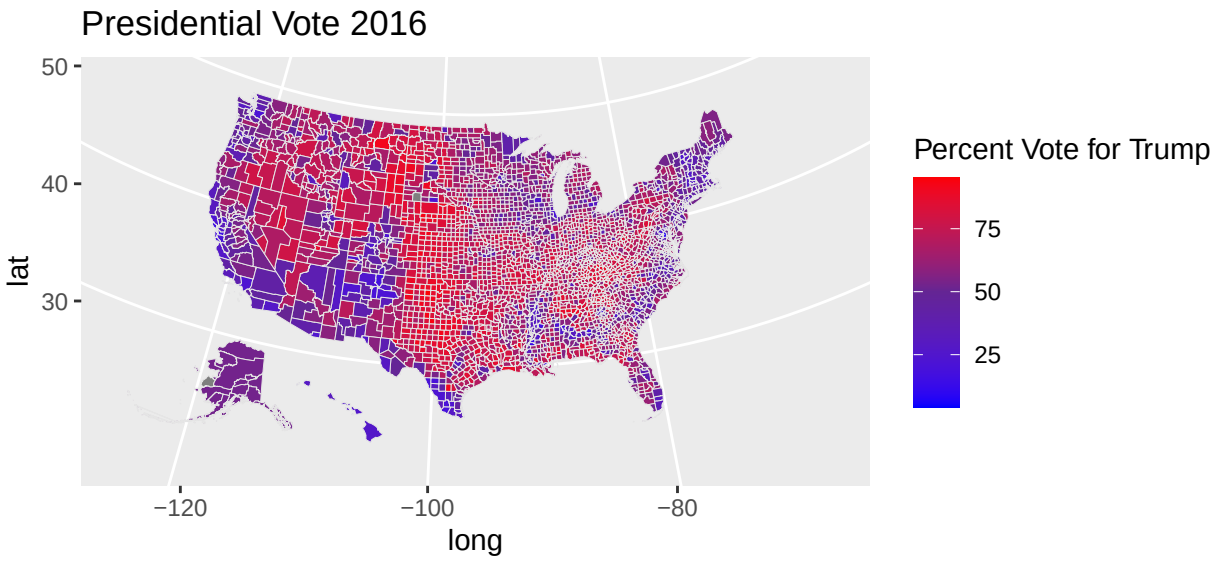


Choropleth Map

Using ggplot2, we can create an aesthetically pleasing and information filled map.

```
ggplot(
  data = census_data
  , aes(
    x = long
    , y = lat
    , group = group
    , fill = prcntGOP16
  )
) +
  geom_polygon(
    color = "gray90"
    , size = 0.05
  ) +
  coord_map(
    projection = "albers"
    , lat0 = 39
    , lat1 = 45
  ) +
  scale_fill_gradient2(
    low = "blue"
    , mid = scales::muted("purple")
    , high = "red"
```

```
  , midpoint = 50
) + labs(
  title = "Presidential Vote 2016"
  , fill = "Percent Vote for Trump"
)
```



Independent Activities

1. Use the provided data set (which can be downloaded from github) and create some graphics of your own!

- Use a histogram, scatter plot, and bar chart to explore different variables in the data set. Once you have found a variable that you are interested in, create a map of the entire US as well as a map of a particular state using it.

2. Review the tidyCensus package documentation to pull data of your own.

- Explore the possible variables (there are many!) and think of a question that you can answer using data visualization. Once you have a question to try to answer, pull the variables of interest. Remember, you may have to change the scale of some variables or mutate them in another way. Once your data is ready for graphing, have at it! Try out different graphs and see which ones lead you to an insightful conclusion.