

Lab 2 - US States

Marcos Falcone

Loading Libraries

I will be using the US States spatial dataset that is included in the spData package.

```
library(sf)
```

```
## Linking to GEOS 3.5.1, GDAL 2.2.2, PROJ 4.9.2
```

```
library(spData)
```

```
## To access larger datasets in this package, install the spDataLarge  
## package with: `install.packages('spDataLarge',  
## repos='https://nowosad.github.io/drat/', type='source')`
```

Summary Statistics

Seven variables are included in the summary statistics below:

- GEOID: geographic identifiers
- NAME: state names
- REGION: region names
- AREA: area in square kilometers
- total_pop_10: total population in 2010
- total_pop_15: total population in 2015
- geometry: type of geometry (sfc_MULTIPOLYGON)

```
summary(us_states)
```

```
##      GEOID          NAME          REGION      AREA  
## Length:49      Length:49      Northeast: 9  Min.   :   178.2  
## Class :character Class :character Midwest :12  1st Qu.: 93648.4  
## Mode  :character Mode  :character South  :17  Median :144954.4  
##                                     West   :11  Mean   :159327.3  
##                                     3rd Qu.:213037.1  
##                                     Max.   :687714.3  
##      total_pop_10      total_pop_15      geometry  
## Min.   : 545579      Min.   : 579679      MULTIPOLYGON :49  
## 1st Qu.: 1840802      1st Qu.: 1869365      epsg:4269    : 0  
## Median : 4429940      Median : 4625253      +proj=long...: 0  
## Mean   : 6162051      Mean   : 6415823  
## 3rd Qu.: 6561297      3rd Qu.: 6985464  
## Max.   :36637290      Max.   :38421464
```

```
sapply(us_states,class)
```

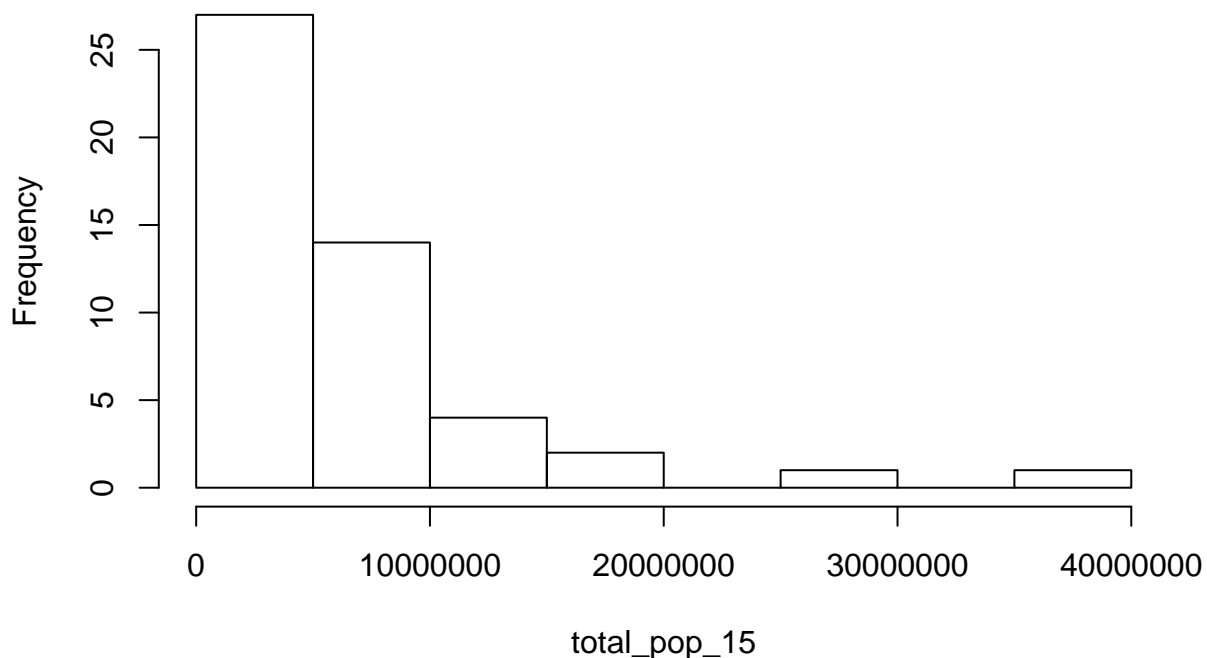
```
## $GEOID  
## [1] "character"  
##  
## $NAME  
## [1] "character"
```

```
##
## $REGION
## [1] "factor"
##
## $AREA
## [1] "units"
##
## $total_pop_10
## [1] "numeric"
##
## $total_pop_15
## [1] "numeric"
##
## $geometry
## [1] "sfc_MULTIPOLYGON" "sfc"
```

Plots

A histogram that shows population values for US states in 2015 in squared kilometres is provided below. I had an issue when making this plot, which was that the standard way to produce a histogram (without the ‘with’ at the beginning) didn’t work. What happens is that, when I run it, it keeps saying that the variable needs to be numeric, which according to the chunk code that contains the summary of the variables it actually is. To troubleshoot, I tried to then coerce the variable into being treated as numeric using the `as.numeric` function, but that doesn’t work as it keeps returning “(list) object cannot be coerced to type ‘double’”. For some reason, the way that I ended up producing the plot did not make it possible for me to make any of the modifications that I intended to produce (such as modifying the title, changing the number of bins, or adding color to the columns).

Histogram of total_pop_15



A map of US states ranked in terms of total 2015 population is shown below.

2015 Population by US State

