

# Maps

MSPA PREDICT 455-DL-SEC55

*Darryl Buswell*

# 1 Introduction

This assignment explores Census data for United States, with the aim of identifying trends or anomalies in the underlying data using both static and animated choropleth maps.

## 2 Data

Data for this assessment was obtained from the United States Census Bureau (Commerce 2015). An API key was obtained and used in order to pull relevant data from the the American Community Survey (ACS). This data includes population, age and sex data for the United States from 2009. A list of tables queried as part of this assessment can be found in Appendix A.

## 3 Data Exploration

A choropleth maps provides a useful method for visualizing spatial data. In R, perhaps the most simple method of generating a choropleth map is to use the ‘choroplethr’ package by Trulia (Innovation 2014). This package is able to pass ACS data and build a relevant choropleth map in a single R command.

We use the ‘state\_choropleth\_acs’ command to generate a choropleth map of the United States to show the population, median age and household income for each state over the years 2009 to 2014. These figures are shown in Appendix B. By default, ‘choroplethr’ divides the lower 48 states into nine equally sized buckets and colors the buckets using a sequential brewer scale. We see that states with the greatest population over this period include California, Texas, Illinois, Ohio, Pennsylvania, New York and Florida. Interestingly, these states also have recorded the greatest household income over this period. In terms of median age, Florida, West Virginia, Pennsylvania, Vermont and Maine have a relatively greater median age than other states.

We can also look at the data at the county level by using the ‘county\_choropleth\_acs’ command. This gives us greater detail, however there is a trade-off in terms of interpretability. That is, it can be more difficult to attribute individual states as having a relatively high or low characteristic, however we can see more broadly which general areas have a relatively high or low characteristic. For example, we generally see a greater population and higher level of income along the east and west coast of the United States, while a lower population and level of income within central United States.

The ‘choroplethr’ package does provide some flexibility in terms of how we categorize the data. We can for example, force the amount of data buckets to two. This allows us to more easily recognize counties which have a income level above or below a certain threshold. We can also zoom on particular states. For example, we have generated a map which zooms on states on the west coast of the United States (Washington, Oregon, California, Idaho, Nevada, Utah and Arizon). Figures for these data subsets are also shown in Appendix B.

Perhaps the most useful feature of the ‘choroplethr’ package is its ability to automatically generate an animated choropleth map. The package does this by appending a number of maps to a list of choropleth’s, generating a choropleth map image for each, and subsequently exporting these maps to a directory. The package will also create a html file with a JavaScript based player that allows the viewer to cycle through each choroplethr map. We have used this function to generate a set of choropleth map’s for population at the county level within the United States, for five year periods ending 2009 through to 2014. The generated images and JavaScript player can be found within the working directory of this markdown document.

## 4 Conclusion

We were able to leverage the ‘choroplethr’ package in order to generate a number of choropleth maps of ACS statistics for the United States. This package allowed us to easily identify states which have a relatively high or low population, median age and income. We were also able to investigate some of the flexibility of this package by zooming into particular states and creating an animated choropleth map of population within the United States at a county level.

## Appendix A Table Output

Table A1 ACS Table Query

Table	Description
B01002	MEDIAN AGE BY SEX
B01003	TOTAL POPULATION
B19001	HOUSEHOLD INCOME IN THE PAST 12 MONTHS (IN 2012 INFLATION-ADJUSTED DOLLARS)

## Appendix B Figure Output

Figure B1 United States Population by State

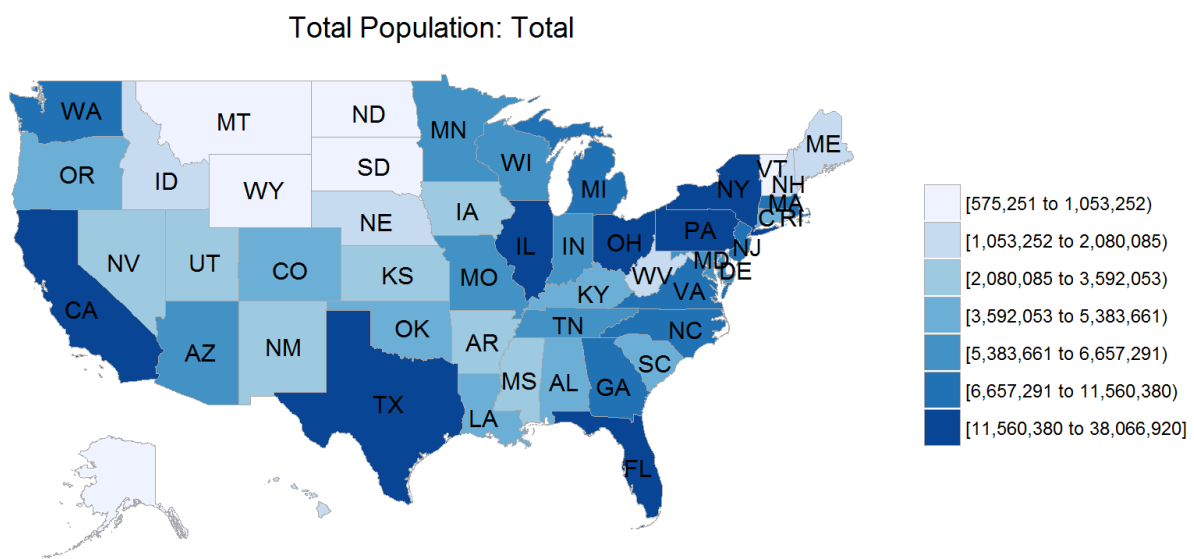


Figure B2 United States Median Age by State

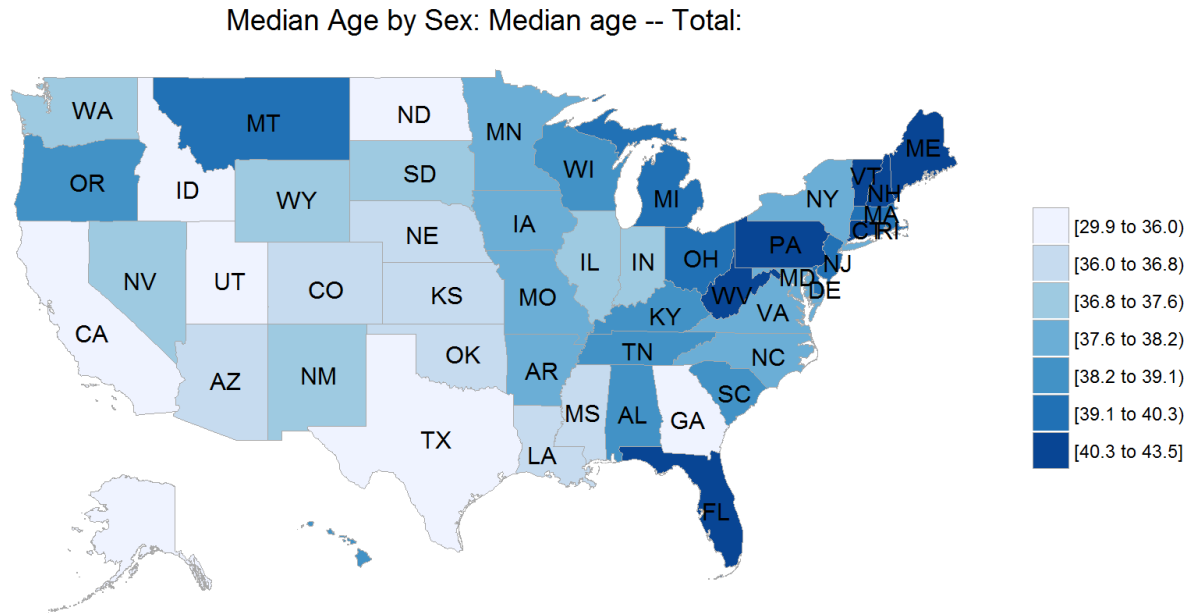




Figure B4 United States Population by County

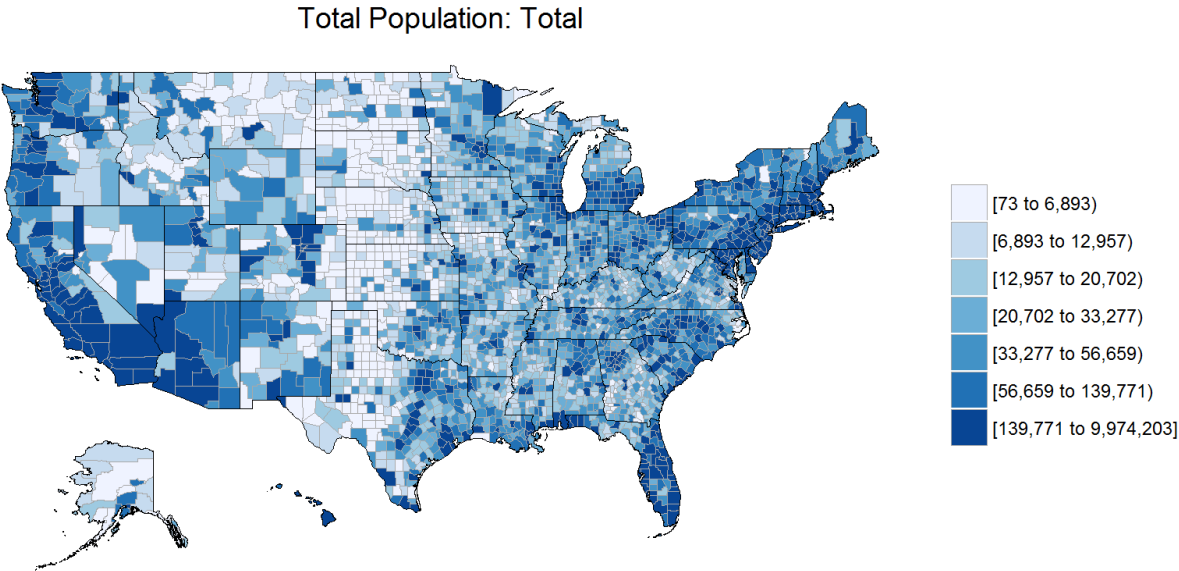


Figure B5 United States Median Age by County

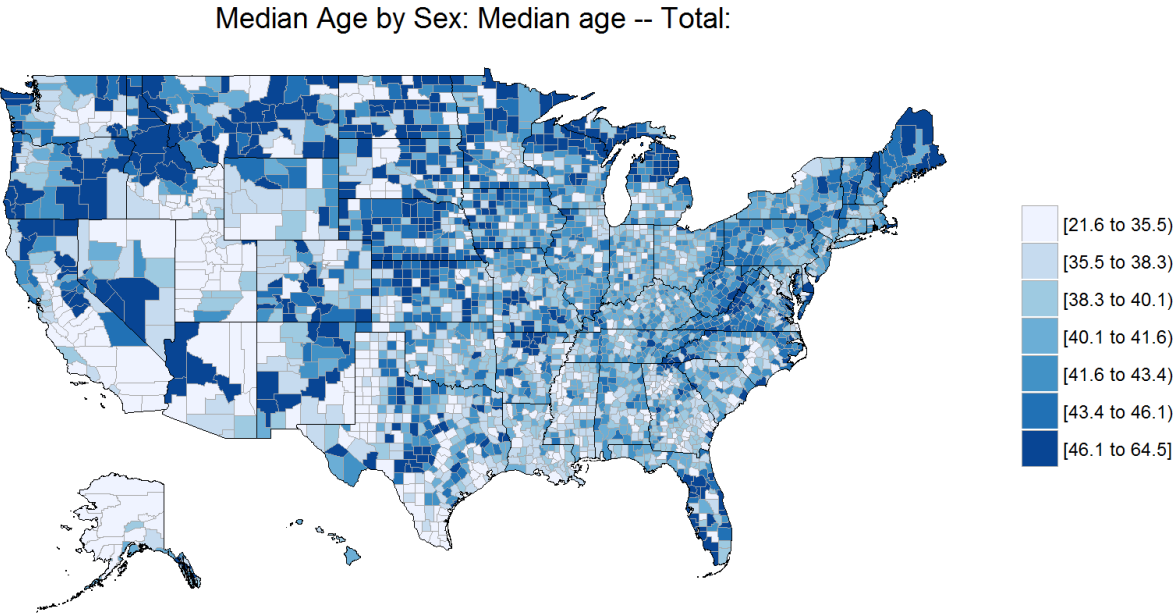




Figure B6 United States Average Income by County

. Household Income in the Past 12 Months (in 2014 Inflation-Adjusted Dollars): Total:

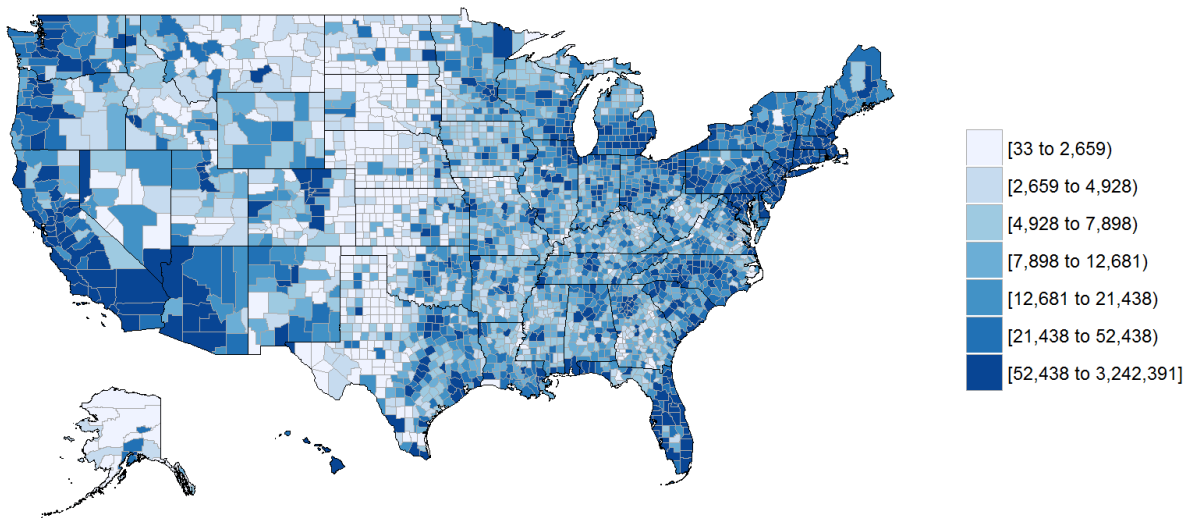


Figure B7 United States High/Low Income by County

I. Household Income in the Past 12 Months (in 2014 Inflation-Adjusted Dollars): Total:

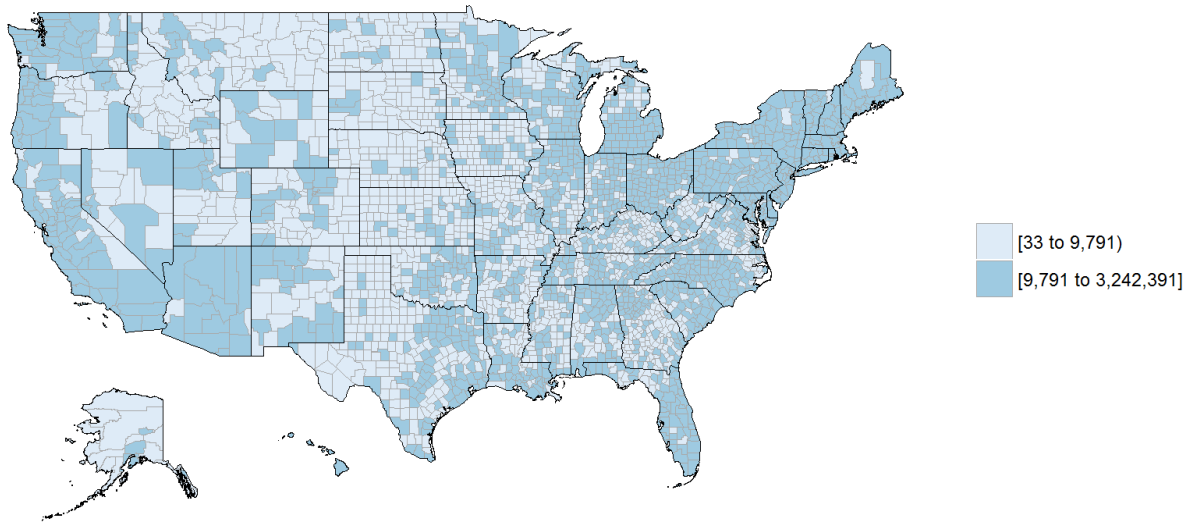
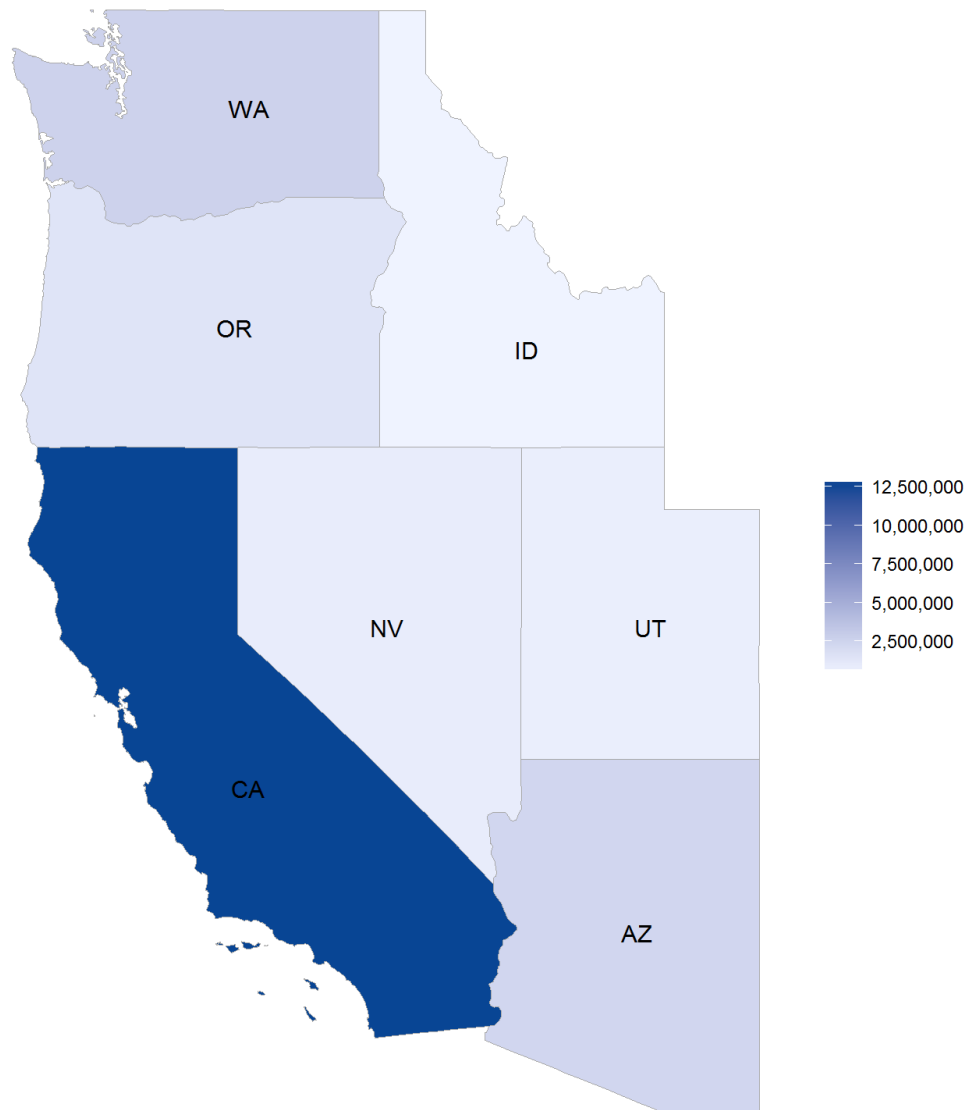


Figure B8 United States Average Income by County

001. Household Income in the Past 12 Months (in 2014 Inflation-Adjusted Dollars): Total:



## References

- Commerce, U.S. Department of. 2015. “American Community Survey.” <https://www.census.gov/acs/www/>.
- Innovation, Truila Tech &. 2014. “The Choroplethr Package for R.” <http://www.trulia.com/blog/tech/the-choroplethr-package-for-r/>.