**Final Project Proposal**

Leland Randles

DATA608, Spring 2019

For my final project, I will create and deploy a Shiny app using CDC (Centers for Disease Control and Prevention) data combined with 2010 census data.

Specifically, I am going to use the “Alcohol-Attributable Deaths due to Excessive Alcohol Use” dataset found here: <https://chronicdata.cdc.gov/Alcohol-Related-Disease-Impact/Alcohol-Attributable-Deaths-due-to-Excessive-Alcoh/yqqr-a7uy>. This dataset provides actual/estimated average yearly #’s of deaths due to excessive alcohol use for the period spanning 2006-2010. It separates the data by state, sex, condition type, condition ("Suicide", "Hypothermia", "Motor-vehicle traffic crashes", etc.) and age range. To incorporate per-capita figures into the analysis, I am going to pull state population numbers from the table on this web page: <https://en.wikipedia.org/wiki/2010_United_States_Census>.

The dataset requires considerable manipulation because of the format of the data. Some rows duplicate other rows. For example, in addition to showing the number of deaths by condition, there are rows where the “condition” is “subtotal all causes”, which is really the sum of the other rows. Therefore, you must filter out the “subtotal all causes” rows or you will double-count or include confusing data in your visualization. Similar issues exist for 4 or 5 of the other data elements.

This data is of interest because excessive alcohol use is a public health issue that costs lives, brings anguish to families, and exacts financial costs on society.

I will include multiple interactive visualizations: a) one which allows a user to compare states against each other based on sex, condition, and age range, b) another which allows the user to compare states against the national rate for these same criteria, c) one which focuses on the “Under 21” age group category, as this is of special interest given that the legal drinking age is 21 in the United States, and d) a map visualization. I expect these visualizations will be created using ggplot2 and some or all of the following packages: mapview, leaflet, and/or tmap.