Final Project Proposal

Team Member: Mengran Xia, Karen Xu, He Liu

Title: Alcohol Consumption around the World

Abstract: We are interested in finding out how alcohol consumptions are difference from country to country, its relationship to students' health and whether if it is a factor of health conditions of citizens of each country.

- How are alcohol consumptions different from country to country? What is the average daily alcohol consumptions of a household in each country?
- How is alcohol consumptions influence students' school performance? What might cause students to consume alcohol?
- Does drinking alcohol increase the happiness level of people from different countries? What if we compare the happiness level to soda consumption?
- Do countries that consume the most beers would also likely to consume more wines?

We will use ggplot2, mapping, interactive graphs and text mining to complete our group project.

Data: For our analysis, we need data from the following sources

- National level data on Household drinking:
  - 2015 National Survey on Drug Use and Health (NSDUH): https://www.samhsa.gov/data/sites/default/files/NSDUH2015MRB/NSDUHmrbCAIquex2015.pdf
  - This dataset contains information collected from a national survey that provides each participant's demographic background and consumption of alcohol including frequencies and number of drinks each time.
- Global level alcohol consumption vs welfare:
  - https://www.kaggle.com/rjribeiro/alcohol-consumption
  - https://www.kaggle.com/marcospessotto/happiness-and-alcohol-consumption

- o The first dataset provides a more general introduction of alcohol consumption around the world broken down into three different drinking categories. This is especially helpful for the geospatial data visualization.

- o The second dataset adds two additional indicators of Happiness Score and HDI each provided by United Nations. This will help us observe the correlation between drinking habits and welfare by country.

- National level data on college student drinking:

  - o https://www.kaggle.com/uciml/student-alcohol-consumption

  - o This dataset focus on the drinking behavior of a smaller sample of students from different colleges. It also provides a comprehensive student background including parents' education, academic and extracurricular actives. This information can help us identify a pattern between social background and alcohol consumption.

Brainstorming: