The liquor dataset from the state of Iowa’s Alcoholic Beverages Division contains over 12.5 million observations. All of the data work completed did not use this raw data and the processes used to clean the data are discussed in this section. To look at county variations in consumption, ensuring that all of the data entries had correctly specified counties was important. Upon further inspection, the original dataset stated that there were more counties in Iowa than existed. This was due to minor differences in spelling of county names. For example, O’brien county registered as a different county than Obrien county. After these adjustments were made, there were still several observations with null county entries. Luckily, all unique store identifiers had one county name associated with it. This meant that the null county entries could be entered using the same unique store identifier for an entry with county already entered.

One of the first operations we performed on the data was to convert the ‘Date’ column to datetime data type. This is important for all of the groupby functions that will be applied afterwards. The data was aggregated, in the beginning by date and county while sales in dollars and volume in liters were summed for each day. Two csv files were exported, one that contained daily sales and the other volume for each county in Iowa. Practically the same process was taken to create monthly sales and volume except that the aggregation was done at the monthly level. Two csv files were exported with the analogous monthly sales and volume for each Iowan county.

We merged the previously discussed cleaned liquor sale data using county level data on votes cast for each party in the 2016 Presidential election and population. To determine the partisan majority in each county, a column of Booleans was created in the voting data for if votes for Democrats was greater than Republicans. From this the data was further grouped by the majority party in the county to graph and make comparisons between liquor consumption of these counties.

One thing that was noted while looking at the data that caused suspicion was the fact that the dataset recorded no liquor sales on Sundays and very few on Saturdays. The authors conclude that this is likely due to the Iowa Alcoholic Beverages Division failing to collect data on these days. To the best of the knowledge of the authors no state liquor laws prohibit sales of these beverages on weekends. Given this complication and that some smaller counties often reported sales on only a few number of days in a week, the smallest temporal timeframe used in our analysis was one week.