**INFO 3300 Project 1 Visualization Description**

**Names/netids:** Wyatt Queirolo (wfq2), Rashaad Ahmad (ra459), Kimberly Lee (kwl38)

The data we used came from a Kaggle “[Craft Beers Dataset](https://www.kaggle.com/nickhould/craft-cans)” containing two csv files, beers.csv and breweries.csv. The two tables are linked by a brewery\_id column which identifies the brewery that produces each beer. For the beers.csv table, we focused mostly on the abv (Alcohol by Volume), style, and brewery columns. To filter the dataset, we deleted all the beer entries that lacked an abv value. We also deleted unnecessary columns such as ibu which did not contribute to our dataset.

Our variables included location and average abv (within each state). Location was displayed by states on a map of the United States of America. The map used hexagons to represent each American state. This map, using hexagons as states, was chosen so that the viewer would not confuse the size of the state to be a variable. We found the map online and further developed and modified it to represent the abv variable through color. The average abv per state is what determined the color of the state’s hexagon. We calculated the average abv per state based on all the breweries each state had. Since the state data was in the breweries.csv dataset and the abv data was in the beers.csv dataset, we had to match each beer to a state using the primary key of the dataset, add up the abv per beer, and divide by the total number of beers used for the state to end up with average abv per state. Once we had all the corresponding averages for each state, we filled each with color. The darkness of the color corresponded with how high the average abv for a single state was (as shown by legend).

We also included a pie chart to display the top 5 beer types in the country. This information was found by using the “style” column on the beers.csv table. As a pie chart suggests, each sector corresponds to a certain overall percentage.

The visualization of the USA map showed one especially interesting thing. Arkansas had a distinctively bright yellow color to it which corresponds to a low average abv. Out of all the states, Arkansas actually had the lowest average abv. Upon further research, according to arktimes.com, "Arkansas is a relatively small market (for beer), so we get passed over a lot. Oftentimes we're one of the last states to get different new products." This could help explain the anomaly Arkansas demonstrated in the map. Since Arkansas is a small market, beer producers there might not produce as much of a variety resulting in lower abv beer rather than a mix.

We were also surprised to see Utah in the middle of the pack of abv, as establishments with “beer only” type licenses as well as grocery stores and gas stations are only able to sell beer with up to 4% abv.

The variation of the chart also can show which states have more craft breweries. Craft beers tend to have a higher abv than mass-produced beers, thus states with higher average abvs (Colorado, California) are more likely to have more craft breweries.

The pie chart visualization made it clear what the top 5 beer types were: Ales, India Pale Ales, Lagers, Porters and Stouts. The pie chart allowed to see how these top 5 beer types stacked up against all the other beer types produced.