

---

# Visual Exploration Assignment

COMP40610 Information Visualisation

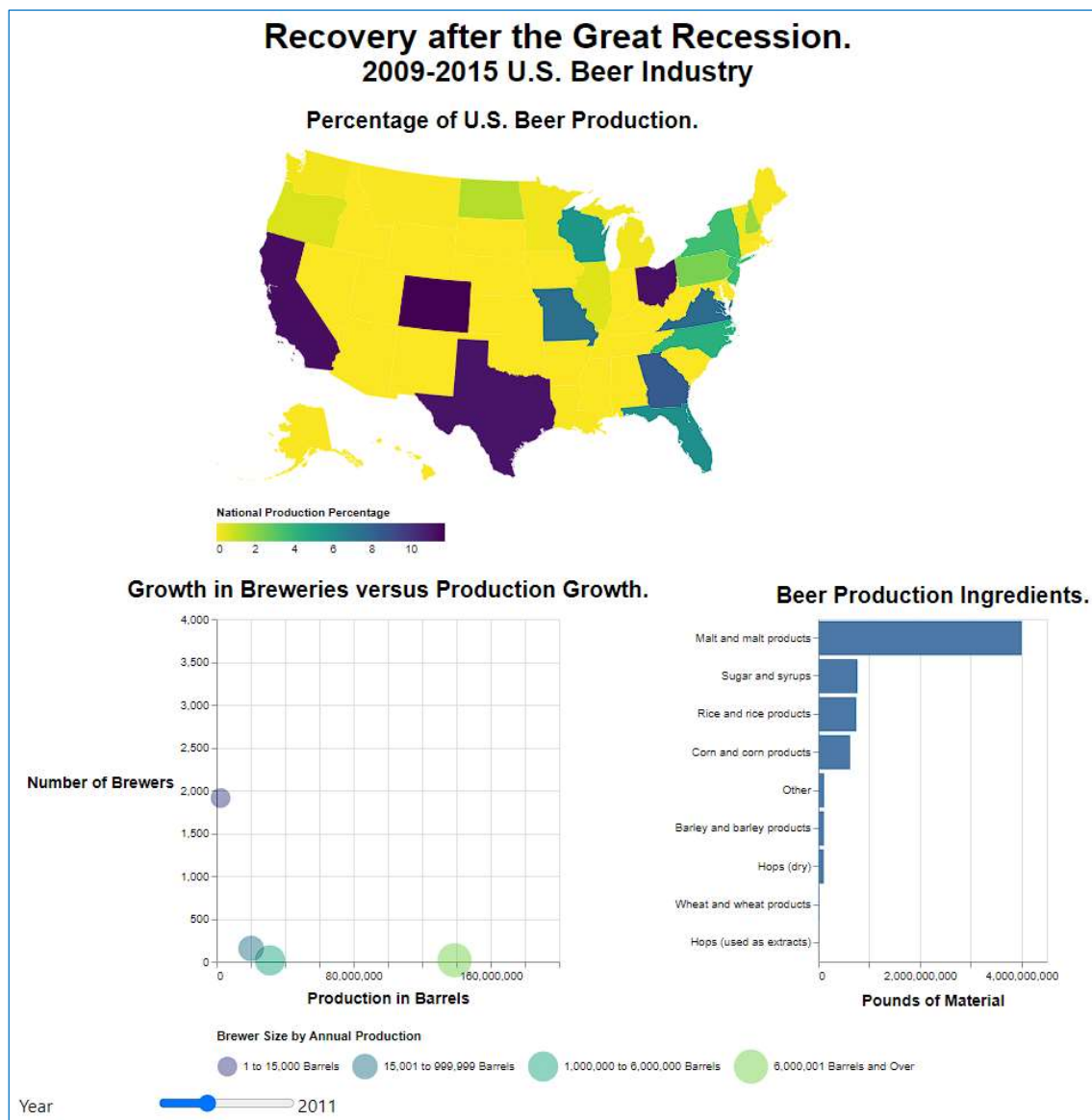


Pedro Antonio Mora Chacón  
Student# 22203184

UCD School of Computer Science  
1 DECEMBER 2023

---

## Screenshot.



## Beer Production Dataset overview.

These datasets stem from 2008 – 2019 reports filed with the U.S.A. Alcohol and Tobacco Tax and Trade Bureau (TTB) federal government agency [1] and encompasses various aspects of beer production. There are five datasets available for download which I used Jupyter Notebook and Python tools to download, transform and upload to my GitHub account the final datasets displayed in the visualisation.

Even though I started with five csv files from the website, 'brewing\_products.csv' and 'brewing\_materials.csv' are duplicates so I did not use 'brewing\_products.csv'. Nor did I use 'beer\_taxed.csv' file simply because it was not related to the topic I found researching the U.S.A. Beer Industry during that time.

Briefly, in regards to data cleaning and or manipulation, I used the scikit-learn *KNNImputer* to fill in values, removed constant columns, added calculated fields like percentage changes or State's percentage based on national output, and I added the State name and State id in regards to using the Vega geoshape

map [2]. I should note that because of the interactive tool incorporated into the visualisation that lets the user choose the year to display, there is a difference between the annual range of all the data and the one used. The reason the earliest date in the visualisation is 2009 is that there are different annual ranges between the dataset csv files and one has data starting from 2009. The reason for setting 2015 as the final year is that in 2015 the TTB filings used (Form 5130.9) faces out the beer production material reporting requirement [3]. The data goes to nearly zero by year 2016.

## Design considerations.

### Overall goal.

For the overall goal I wanted to show the U.S.A. Beer Industry recovery after the 2008 Great Recession. To show the changes in the data that might have lead to the Beer Industry recovery.

### Choropleth chart.

This is a U.S.A. map with a heatmap that displays each U.S. State's contribution to the national recovery by transforming their production into their percentage of the national production. The mark area channel represents the nominal U.S. States. The colour channel represents the State's ratio/percentage of the national production. I used the 'viridis' colour scheme to make it colour blind friendly. As the audience might not be from U.S.A. For Interaction, I also added U.S. State's name as a feature and included a tooltip for the map that gives the name, the State's actual production, and their actual percentage of the national production. Also, this chart has a filter that makes it bound to the slider at the bottom to let the user choose the year to display.

### Scatter plot.

This scatter plot chart displays the rapid increase in the smaller brewer size categories for the years following the recession. The circle marks which represent the nominal brewer size categories, the y axis represents the quantitative number of brewers for a category, and the x axis represents the quantitative number of barrels the brewer categories produced. The colour and size channels are both redundant encodings and together with the 50% colour opacity, they are meant to enhance the user perception of the marks even when the circles cluster together in the scatter plot. This chart is also interactive as it has a filter that makes it bound to the slider at the bottom to let the user choose the year to display. It includes a tooltip to display the specific x and y axis amounts. This also uses 'viridis' colour scheme to make it colour blind friendly.

### Bar chart.

The bar chart displays the gradual decrease in the expensive and dominant malt ingredient category. Significant because in 2009 there were studies with a new enzyme allowed brewing directly with less barley while saving energy in the process [4]. A sensory analysis study of the process with enzymes (instead of the malt ingredients) found that there were no significant differences between the two methods [5]. The chart bars length *mark* represents the beer production, the y axis represents the nominal Beer ingredient categories, and the x axis represents the quantitative weight in pounds of the beer production ingredients. The colour and size channels are both redundant encodings and together with the 50% colour opacity, they are meant to enhance the user perception of the marks even when the circles cluster together in the scatter plot. This bar chart is also interactive as it has a filter that makes it bound to the slider at the bottom to let the user choose the year to display. And it

includes a tooltip that gives the actual specific weight, and the difference from the previous year ingredients' use weight as a percentage change.

### Interaction consideration.

As mentioned above there were multiple interaction considerations taken. The first being the tooltips which were done with consideration given for providing additional information like transformational features without hindering the presentation by cluttering or making it more confusing. The second interaction consideration and just as significant is the annual slider tool that allows the user to compare the changes over the years by focusing on the years in question and seeing the movement resulting from their selection displaying a growth or decline in the metrics.

## Works Cited

- [1] R4DS Online Learning Community, "tidytuesday/data/2020/2020-03-31/readme.md at master · rfordatascience/tidytuesday," 31 March 2020. [Online]. Available: <https://github.com/rfordatascience/tidytuesday/blob/master/data/2020/2020-03-31/readme.md>. [Accessed 1 12 2023].
- [2] Vega-Lite, "Three Choropleths Representing Disjoint Data from the Same Table," [Online]. Available: [https://vega.github.io/vega-lite/examples/geo\\_repeat.html](https://vega.github.io/vega-lite/examples/geo_repeat.html). [Accessed 1 12 2023].
- [3] microbrewr.com, "Microbrewery taxes simplified; the quick guide to the complicated world of beer taxes," 23 March 2017. [Online]. Available: <https://microbrewr.com/microbrewery-taxes-simplified/>. [Accessed 1 12 2023].
- [4] Martin Heller State of Oregon, Dept of Environmental Quality, "Food Product Environmental Footprint Literature Summary: Beer," September 2017. [Online]. Available: <https://www.oregon.gov/deq/FilterDocs/PEF-Beer-FullReport.pdf>. [Accessed 1 12 2023].
- [5] N. E. P. H. N. A. M. N. O. R. a. A. K. Jesper Heddal Kløverpris, "Comparative Life Cycle Assessment of Malt-based Beer and 100 per cent Barley Beer.," *Brewing & Beverage Industry International magazine*, no. 2, pp. 34-36, 2010.