CWRU Data Analytics Boot Camp

Spring 2020

Project Proposal

Randy Dettmer, Evan Kamis, Ali Rizvi, Sumeet Dhawan

Impact of Brewery Trend on DUI

Background

It's a great time to be a craft beer fan in the U.S.! There are a ton of beer styles and brands to choose from and breweries have become very successful in the last several years. Breweries owe it all to beer lovers around the world! However, the other side of the story is that DUI’s are also increasing and we would like to see if there is a correlation of rise of breweries in US with the amount of DUI’s in the 50 states in the US.

Proposed Work

**Extract:** We plan to use the following Datasets of Breweries from Kaggle and Open Beer Database.

* US Breweries - (Kaggle) - link on brewery name
* Craft Beers Data Set -(Kaggle) - link on brewery name. This data was collected in January 2017 from [CraftCans.com](http://CraftCans.com).
* Open Beer Database: Last Updated 2016

We also plan to use a DUI database

* DUI Arrests, alcohol/vehicle deaths USA 2015 - (Kaggle) - link on the state in brewery
* Scrape for 2017 data

<https://www.responsibility.org/alcohol-statistics/state-map/state/ohio/>

**Transform**: We plan to clean up the data and apply filtering to narrow down breweries by state and link it to the DUI data by adding appropriate aggregation

**Load**: Link it to a SQL database so that it is in a clean readable format

**Who Does What**

1. Finding Data Source - Team - April 21
2. Cleaning up the data - Randy and Evan- April 23
3. Transforming the data - Ali and Sumeet
4. Loading the data - Team
5. Final Technical Report - Team - April 25

Daily Scrum to ensure project is on track : Zoom or Slack everyday to note any trouble