## **Project Proposal and Data selection/Milestone 1**

**What domain is this data going to come from?**

I'm going to concentrate on craft beers and breweries for this endeavor. The dataset, which focuses on artisan beers and brewery locations, was discovered on Kaggle.

Reference:

<https://justbeerapp.com/article/10-reasons-you-should-drink-craft-beer>

<https://www.nytimes.com/2018/02/27/business/craft-breweries-local-economy.html>

<https://www.tasteofhome.com/article/types-of-craft-beer/>

<https://www.craftbeer.com/beer/beer-styles-guide>

<https://eh.net/encyclopedia/a-concise-history-of-americas-brewing-industry/>

<https://www.upstreambrewing.com/blog/how-the-brewing-industry-helps-local-economy/>

<https://www.theatlantic.com/business/archive/2018/01/craft-beer-industry/550850/>

**Which Data?**

The Craft Beer Dataset from Kaggle is the dataset I'll be looking at. It can be found here: https://www.kaggle.com/nickhould/craft-cans. One dataset is for breweries, while the other is for craft brews. The brewery dataset comprises the name of the brewery as well as its location in terms of city and state. The alcohol by volume, international bitterness units, ID, beer name, beer style, brewery ID, and number of ounces are all included in the craft beers dataset.

**Research Questions? Benefits? Why analyzes these data?**

Both datasets will be examined in order to answer various queries.

• What are the features of each variety of beer?

• Where are the majority of breweries located? What is the smallest amount?

• Where would it make the most sense to open a new brewery if one were to open?

• What are the similarities and differences between the beers?

**Method**

To find the optimum outcome, I plan to examine various models, beginning with classification and clustering models. Each model will be able to help split the data into categories based on features and provide the most accurate analysis for addressing the research questions. Methods for exploratory data analysis and visualization will also be investigated.

**Potential Issues?**

I can see me dealing with a few challenges. Prior commitments, everyday activities, and focus are all possible factors that could put me off schedule. I can also see some possible difficulties in locating relevant studies. There is a lot out there, but all of the information collected so far is quite similar.

**Conclusion**

Various breweries across the country produce a variety of craft beers. Lagers, pilsners, ales, and IPAs are among them. While the beers themselves are enjoyable for some, the presence of breweries aids in the development of local economies. Craft beers and the breweries that make them provide us with more to enjoy. It provides an experience, supports a small business, specialty beers are available throughout the year, and some may argue that it draws people together. There's a lot that goes into brewing a craft beer, as well as the brewery's history. By being able to extract the properties of the beers and identify their type from them, we can study more of this data. Furthermore, we can determine how breweries may contribute more to the communities in which they operate, such as revenue and jobs.