

DSC 680 – Applied Data Science

Craft Beer

Krystina Moses

Bellevue University

Abstract

There are many kinds of craft beer made by different breweries across the country. Some of them include lagers, pilsners, ales, and IPAs. While the beers are nice for some to enjoy, the presence of breweries help build the economy in places. Craft beers and their breweries give us something more to enjoy. It provides an experience, supports a smaller business, specialty brews are around for the seasons, and some can say it brings people together. There is a lot that goes into brewing a craft beer and history behind the brewery. We can analyze more of this information by being able to take the characteristics of the beers and identify its type from them. In addition, we can identify how breweries can provide more for the areas they are in including revenue and employment.

The datasets that I chose have information around craft beers as well as breweries across the United States. They were combined for analyzation. The focus for this project is to determine the style of beer based on its ABV as well as determine the next best place for a brewery. Gaussian Naïve Bayes and logistic regression models were created to determine what type a beer is based on it's ABV as well as where a good location for a brewery would be next. In review, it was determined that at this time Colorado has the highest number of breweries currently. Upon further analyzing the data and fitting the models and a Gaussian Naïve Bayes model is the best to determine the answers to our questions.

Craft Beer

Background

When people think about beer, they generally think of the main corporations such as Anheuser-Busch, Miller-Coors, Pabst Brewing Company, etc. and when we think about those companies, we often forget that craft beer is something different. Having a craft beer means that you are supporting a smaller business, most likely locally brewed, and have a different experience than you would from having a beer from a big named brand.

People from all over will go to breweries when they open just to see what they have to offer, something fresh. This helps local economies and while additionally providing more jobs, it is also known that breweries do contribute to the local tax base. Breweries do not exist for someone or a group of people to go out and get drunk. “They offer a more refined environment where professionals can socialize and appreciate a crafted beverage” (Upstream Brewing Team, 2020).

The business is not for everyone; however, it does benefit everyone. Upstream Brewing Team states that the “beer market keeps revenue circulating throughout the country.” The industry also provides employment, a place for events to take place, both professional and social, and they take pride in what they do. Having a brewery in your city and/or state supports you in one way or another.

Methods

Multiple methods were used during this project. Starting out with data cleaning and then moving on to exploratory data analysis, data visualization, and finally Gaussian Naïve Bayes Models and logistic regression models were fitted and tested. The datasets are from Kaggle, <https://www.kaggle.com/nickhould/craft-cans> and combined from two into one

complete dataset. Two target variables were focused on and they are ABV and state for two different tests. The additional variables within the dataset include ABV, IBU, ID, Beer, Style, Brewery ID, Ounces, Brewery, City, State, Style code and State code. The chart in Appendix A contains more information regarding each variable.

During the data cleaning phase, the two datasets were combined into one larger dataset. Additionally, column names were changed for clarification and distinction, and unneeded columns were removed. The columns that were changed consisted of the original name columns in the brewery dataset and the beer dataset individually. They were each changed to 'beer' and 'brewery' to better identify what type of name it was. I then went on to describe the dataset, view more information, and create a correlation (Appendix B). Before creating and fitting the models, additional columns were created as 'style code' and 'state code'. These columns consist of numbers that were assigned to each column of style and state for model prediction. Gaussian Naïve Bayes Models and logistic regression models were fitted for each target variable. When it was determined what would have the best outcome, it was used for each prediction of style and state.

Results

Based on the visuals (Appendix C and Appendix D), Colorado has the highest number of breweries at this current time and Grand Rapids is the city with the most. It is also known that the ABV and IBU have the closest relationship to each other within this dataset. Fitting each of the models for their predictions, it was determined that both models were effective for predicting the state outcome and the Naïve Bayes model was best for the style outcome. Each model showed the results of what was given to us in the dataset and we can use them to get more information and fill in the gaps.

Discussion/Conclusion

Breweries bring a lot to a community. Economies are being helped as well with having local breweries in town. With more knowledge of what they bring to the economy and to the states themselves, we can see that they increase employment as well as revenue for the city. Craft beer is popular among different age groups but is increasingly popular among millennials. Millennials are said to be adventurous and more are out drinking craft beer and enjoy the quality of it. This gives craft beer brewers an idea of who to market for. We can take this information and use it to identify the best place for the next brewery to open doors and find out what is important to reach out to new generations.

Acknowledgements

I would like to thank the breweries around the United States for creating a space of comfort for friends to hang out and enjoy good company and product.

References

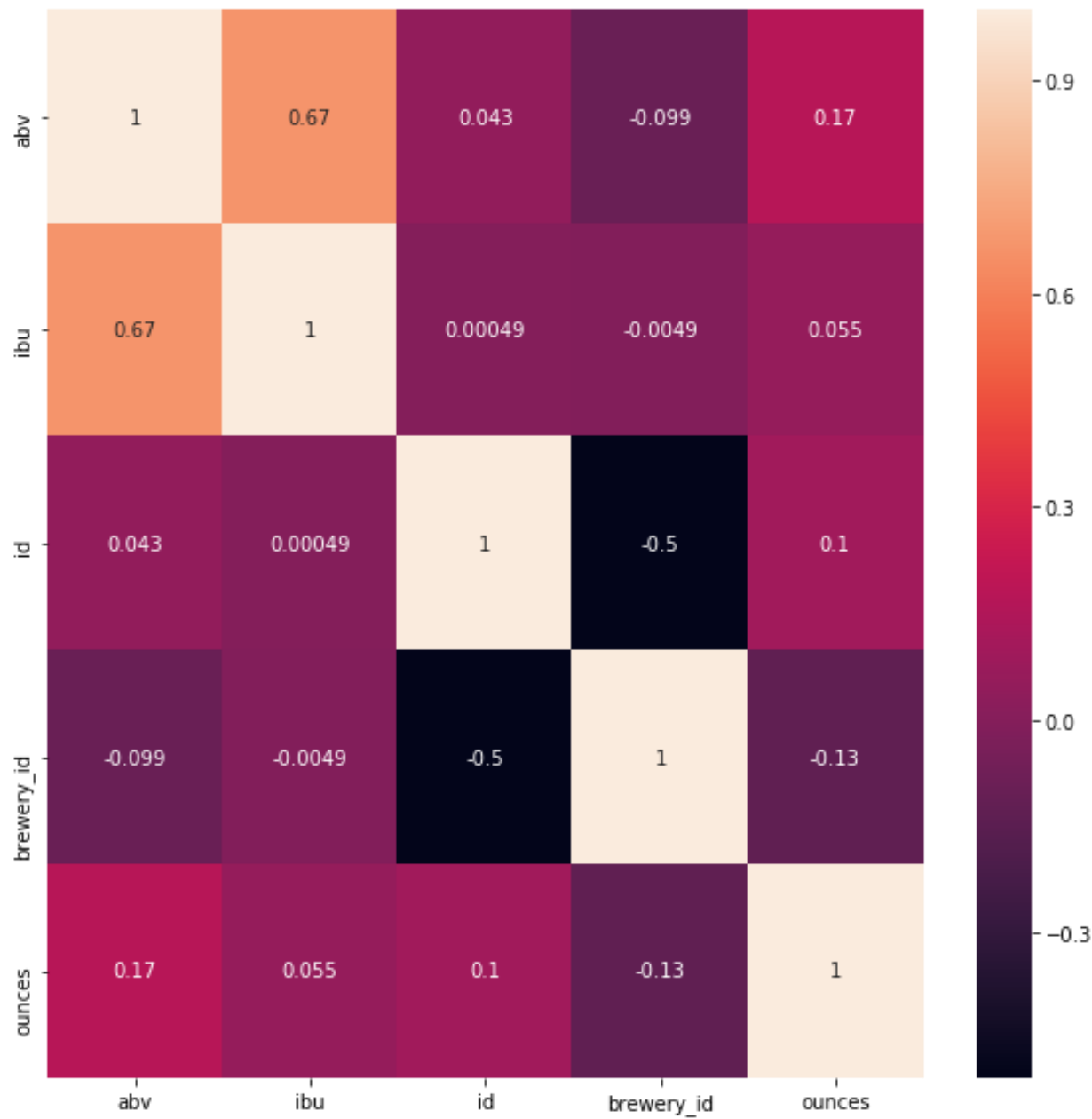
- The Beer Community. (2019, April 5). 10 Reasons You Should Drink Craft Beer. Retrieved from <https://justbeerapp.com/article/10-reasons-you-should-drink-craft-beer>
- Carneiro, J. (2020, March 02). What makes craft beer so popular with millennials?: DSM Food Specialties. Retrieved from https://www.dsm.com/food-specialties/en_US/insights/beverage/craft-beer-millennials-consumer-insights.html
- Celebrating A Year in Craft Beer. (2017, December 13). Retrieved from <https://www.brewersassociation.org/press-releases/2017-craft-beer-review/>
- Hughes, C. (2018, February 27). How Craft Breweries Are Helping to Revive Local Economies. Retrieved from <https://www.nytimes.com/2018/02/27/business/craft-breweries-local-economy.html>
- Mattison, L. D. (2019, August 09). 54 Types of Craft Beer You Really Need to Know. Retrieved from <https://www.tasteofhome.com/article/types-of-craft-beer/>
- Morris, C. (2016, June 08). 3 essential business growth lessons from the craft beer boom. Retrieved from <https://www.cnbc.com/2016/06/03/3-essential-business-growth-lessons-from-the-craft-beer-boom.html>
- Natewebman. (2018, May 08). Beer Styles Study Guide. Retrieved from <https://www.craftbeer.com/beer/beer-styles-guide>
- Nurin, T. (2020, July 14). Beer School: Is Getting A Degree In Brewing A Good Investment? Retrieved from <https://www.forbes.com/sites/taranurin/2016/09/22/what-do-students-at-beer-school-do-if-theyre-not-sitting-around-getting-drunk-all-day/>
- Snider, M. (2020, April 23). America's craft beer boom may go flat as coronavirus shutdown slows brewery taps. Retrieved from <https://www.usatoday.com/story/money/business/2020/04/21/coronavirus-pandemic-creates-brewing-crisis-craft-beer-industry/5151514002/>
- Stack, M. H. (n.d.). A Concise History of America's Brewing Industry. Retrieved from <https://eh.net/encyclopedia/a-concise-history-of-americas-brewing-industry/>
- Team, U. (2020, March 13). How the Brewing Industry Helps Local Economy. Retrieved from <https://www.upstreambrewing.com/blog/how-the-brewing-industry-helps-local-economy/>
- Thompson, D. (2018, January 23). Craft Beer Is the Strangest, Happiest Economic Story in America. Retrieved from <https://www.theatlantic.com/business/archive/2018/01/craft-beer-industry/550850/>

Appendix A

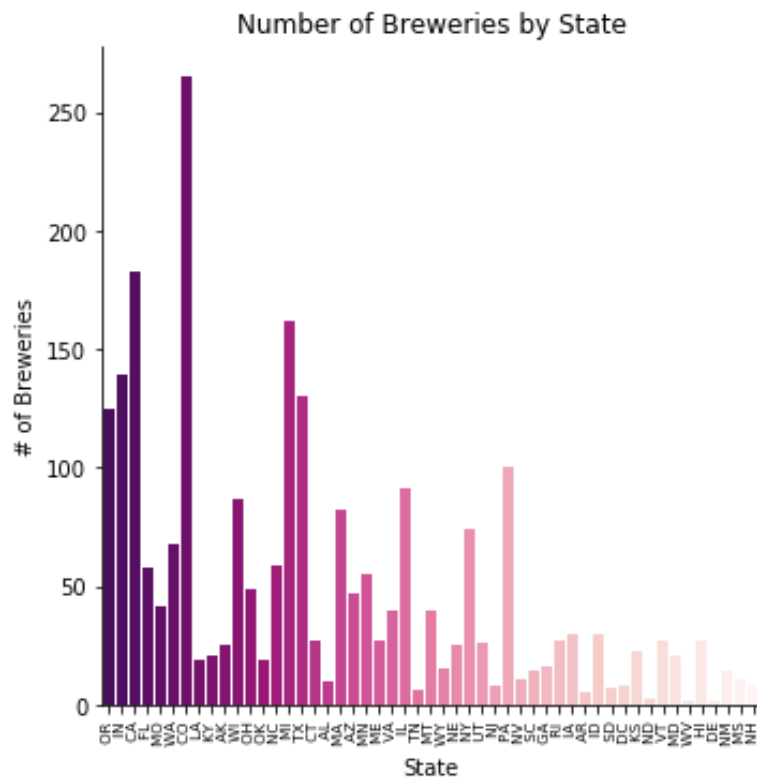
News dataset columns and descriptions

Column Name	Data Type	Description
ABV	Float	The alcoholic content by volume with 0 being no alcohol and 1 being pure alcohol
IBU	Float	International bittering units, which describe how bitter a drink is.
ID	Integer	Unique ID.
Beer	Object	Name of the beer.
Style	Object	Beer style (lager, ale, IPA, etc.)
Brewery ID	Integer	Unique identifier for brewery that produces this beer.
Ounces	Float	Size of beer in ounces
Brewery	Object	Name of the brewery
City	Object	Location of brewery (city)
State	Object	Location of brewery (state)
Style code	Integer	Code assigned to the style of beer
State code	Integer	Code assigned to the State location

Appendix B



Appendix C



Appendix D

