

Beer Recommendation System

1. The problem I want to solve

It's common to see someone staring at a wall of beers at local supermarket, contemplating for over 10 minutes before grabbing one. They are not alone. As a beer enthusiast, I'm always looking for something new to try but I also dread being disappointed so I often spend too much time looking up a particular beer over several websites to find some kind of reassurance that I'm making a good choice. So it's necessary to build a recommendation system to leverage users historical ratings to create predictions for un-tasted beers and create a personalized recommendation for which beer to sample next.

2. Potential client and why do they care about this problem

My target audience will be the beer enthusiast, believe that drinker will find the results interesting. There are more than 2,748 different breweries operating in the US as of June, 2018 and more than 5,000 unique brands in the U.S. which may cause newfound confusion for the average consumer when shopping for beer. Given the huge rise in the brewery scene in the United States, it can be overwhelming trying to decide which new brewery or beer to try. My recommendation system will help them decide which beer to try next given their unique tastes and historical ratings.

3. Data source information

This dataset contains around 1.5 million reviews of Beer from BeerAdvocates. The dataset can be downloaded at <https://www.kaggle.com/rdoume/beerreviews>. The resulting data encapsulated 56 subcategories of beers; 4,964 unique beers; approximately 88 thousand unique user(reviewers); approximately 1.4 million user-review pairings.

4. Brief outline I'll follow to solve this problem

- Clean data and find average beer & user ratings.
- Choose only those beers that have at least n number of reviews, find n using EDA.
- Analysis and factor engineering: understand key drivers to beer ratings.
- Recommend Top 5 beers to users using Simple Recommender, Content Based Recommender, Collaborative Filtering and Hybrid Engine.