

Introduction:

I like beer - I like drinking it, talking about it, and making it. In an alternate universe, I would start a business as a brewery somewhere near where I live, but in this universe, I will attempt to identify good locations for a new brewery in the Philadelphia suburbs, centered on Montgomery county so my evil-universe twin can follow his dream.

I will identify existing breweries and data on their rating as an indication of public perception of the quality of the brewery but not necessarily of the beer. Using this quality rating, I will then segment the breweries into groups based on their ratings and gather venues in the locality so I can do a clustering analysis to determine if there are commonalities among the rating categories.

Finally, If there are common traits for the better breweries, I will attempt to rate other locations in the target area that fit the positive while avoiding the negative criteria. The target areas will be based on geographic distance from other breweries. These locations will be considered as opportunities for a new brew-pub that can be ranked based on the data found.

Data:

- Location data from Foursquare
 - Brewery venues to be identified based on search
 - Venue ratings and number of ratings. I plan to use the ratings and the number of votes to segment the breweries into tiers. I suspect I may need to normalize the ratings based on the number of votes.
 - Trending venues near the target breweries which will be used to cluster the different ratings tiers.
 - Venues near the potential new sites for evaluation against the clustered results above.
- Pennsylvania spatial data - <https://www.pasda.psu.edu/uci/DataSummary.aspx?dataset=41> Geojson data to segment townships in the target county of Pennsylvania for choropleth plotting.
- The Penn State Data Center Municipal level data [here](#), to determine the population of Montgomery county and townships. This data can be used to identify potential customer base relative to population.
- Nominatim and Geolocator for translation of addresses into latitude and longitude coordinates.
- Kaggle.com - older data sets on beers and breweries can be used to help rate breweries found via Foursquare. Initial assessment of this data has not given me much confidence of use.
- www.brewersassociation.org offers limited free data on breweries and beer consumption.
- Untappd and BeerAdvocate databases of beers and breweries if it can be acquired for free. So far, I have not been able to get this data without paying.

