



What's the best place for the next brewery?

Panagiota Kyriakou, June 2020 IBM Data Science capstone



Business Case

Business Goal: Identify candidate cities in the Twin Cities metro area (Minnesota) for a new brewery

Audience: This report can be used by an investor in the metro area or the brewery market to inform their decision and limit the number of cities they will investigate further as potential brewery location candidates.

Background

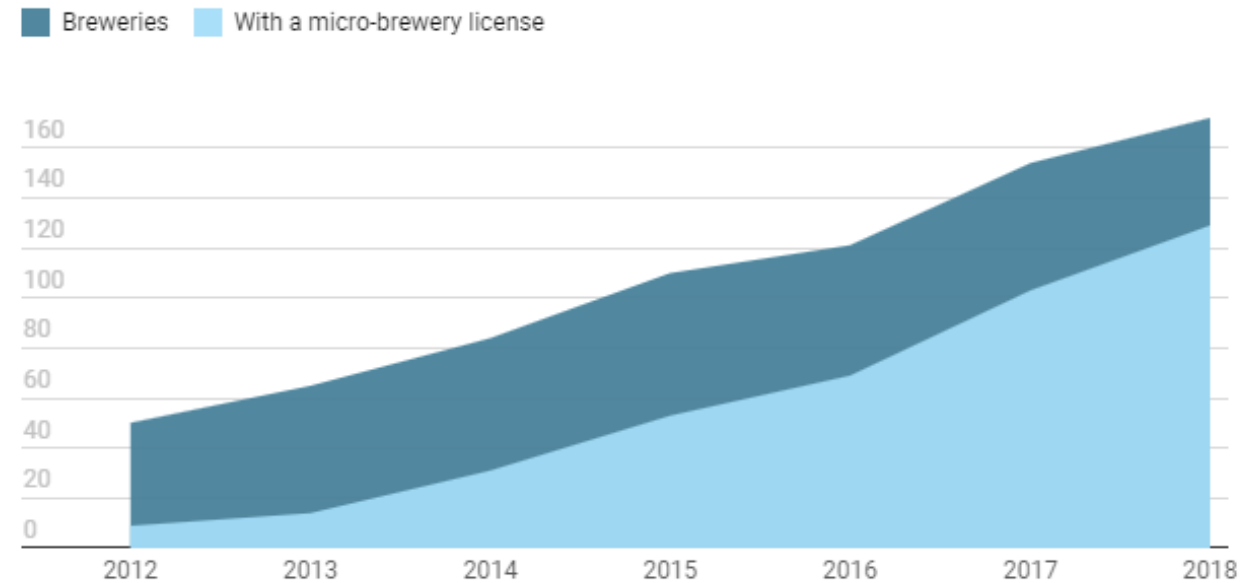
The brewery market in the Twin Cities has experienced a boost over the last decade.

Breweries are popping up across the metro area.

However there are still less breweries per capita than other cities in the US which implies that there is room for new breweries.



Microbreweries significant growth of breweries in MN



Data

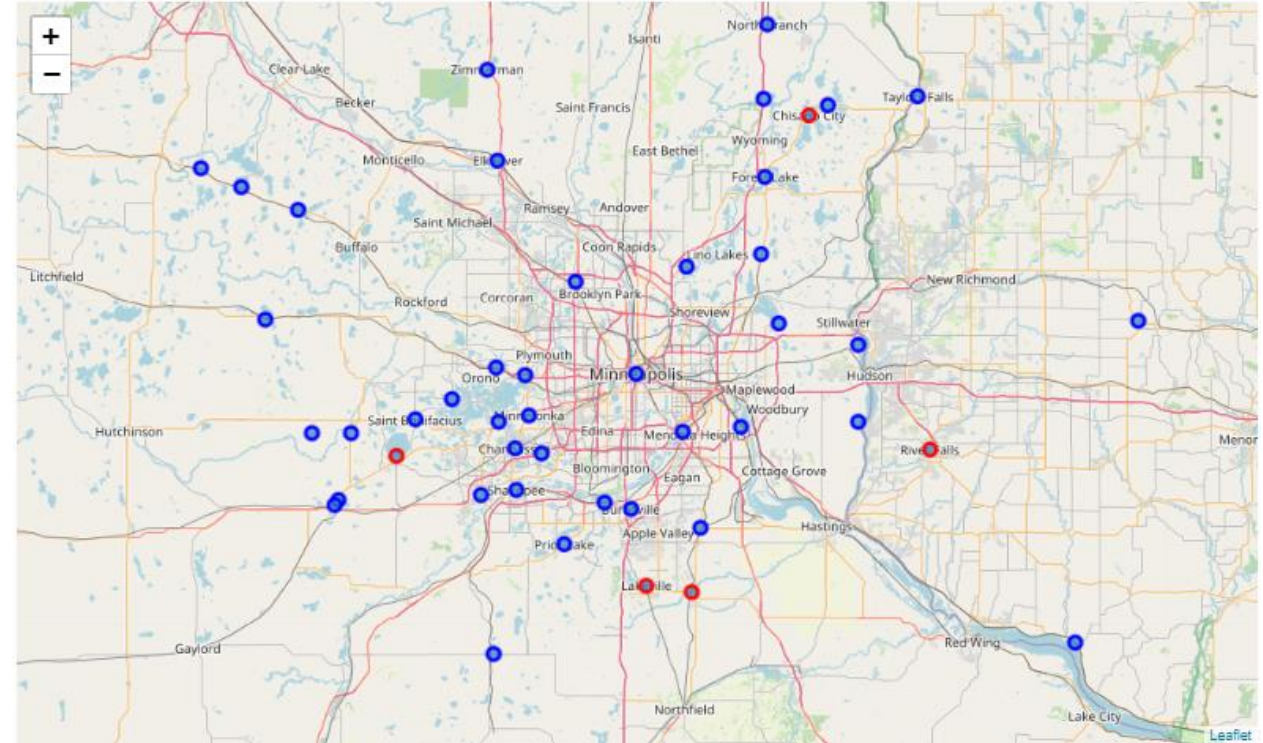
- A list of cities from the Twin Cities Metro area were retrieved manually from the website <https://www.zipcodestogo.com/Minnesota/>, saved to a file and loaded to a dataframe
- The geopy package was used to retrieve the latitude and longitude coordinates of each city
- The Foursquare API was used to retrieve the venue data of the cities

Methodology

- Imported the City data and retrieved their geospatial coordinates.
- Visualized the data using Folio
- Used the Foursquare API to retrieve the venues
- Grouped data by city and took the mean of the frequency of occurrence of each venue category
- Performed clustering and identifies the cluster with the most breweries
- Found the cities that were in the cluster but did not have a brewery and visualized all of them using Folio

Results

- The Twin Cities Metro Area has 142 cities
- The cities can be clustered in 6 clusters
- The cluster that has the most breweries has 94 includes cities (shown in the map)
- 79 of those cities do not have a brewery (shown in blue vs the ones that have, shown in red) and are good candidates for further investigation



Conclusions and Further Directions

- The business goal for this project was to identify cities in the Twin Cities Metro that could be good candidates to host the new brewery
- Through this analysis I successfully reduced the number of potential cities by **44%.**
- In order to get more specific recommendations we might need a more complex clustering algorithm or a different machine learning approach, such as supervised learning.
- Alternatively we can look at incorporating more data that describe the cities (population, income per capita, education, age) that might lead to a more predictive model.
- Finally I would recommend adding some data for the success of each brewery (sales, investment, beer production, twitter mentions etc) to inform the model on what a successful brewery looks like.