NEW BREWERY IN TORONTO

DATA SCIENCE FINDINGS FOR MASTER BREWERS

SEAN MOORE | IBM APPLIED DATA SCIENCE CAPSTONE PROJECT

OUTLINE

1. Introduction
2. Business Problem
3. Data
4. Methodology
5. Results
6. Discussion
7. Conclusion

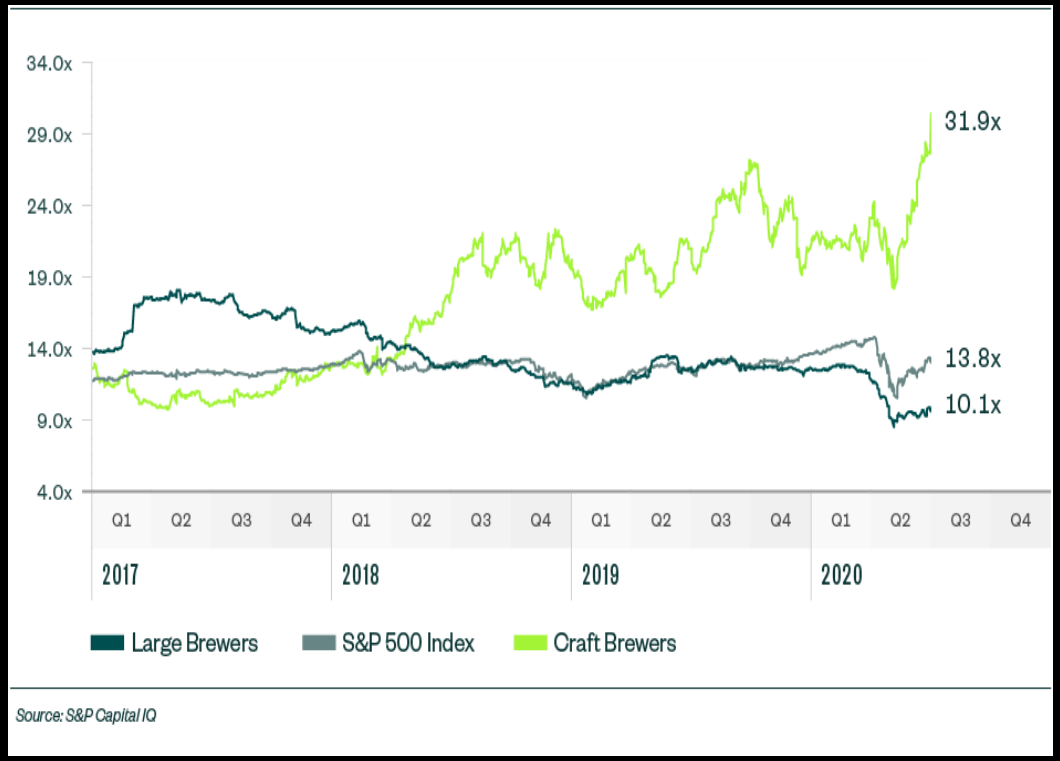
INTRODUCTION

Stakeholders for a team of Master Brewers are interested in opening a new Brewery.

The global pandemic impact of Covid has caused many disruptions across the globe.

There have been some positive results as it relates to Craft Brewers. As shown to the right, Craft Brewers have significantly outperformed Larger Brewers, and even the S&P 500 Index.

Though masters of their craft, the Stakeholders are unsure as to where to place the location of the Brewery, however they have decided the city of choice to be Toronto, CA.



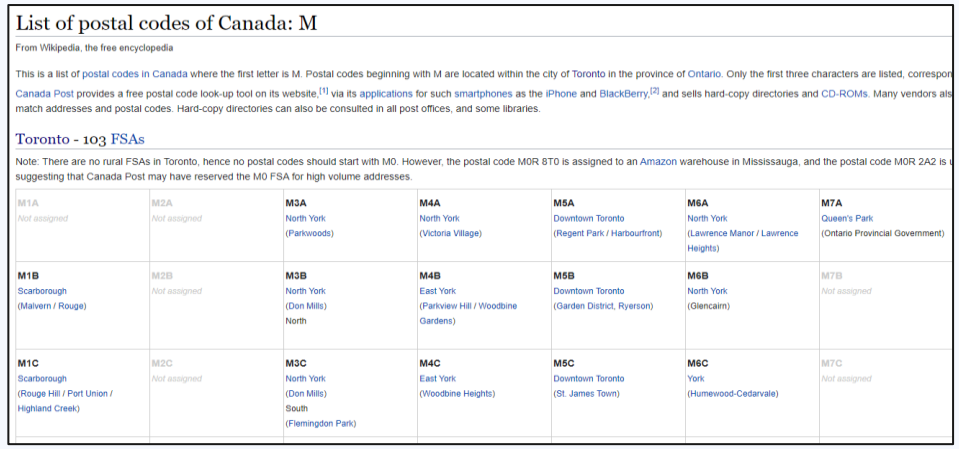
BUSINESS PROBLEM

The Stakeholders would like to identify a primary area for a new brewery in the city of Toronto. Additionally, the location should not be too close in proximity to another brewery as this may create unwanted competition.

Data

Out[20]:

	PostalCode	Borough	Neighborhoods	BoroughLatitude	BoroughLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.65426	-79.360636	Roselle Desserts	43.653447	-79.362017	Bakery
1	M5A	Downtown Toronto	Regent Park, Harbourfront	43.65426	-79.360636	Tandem Coffee	43.653559	-79.361809	Coffee Shop
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.65426	-79.360636	The Distillery Historic District	43.650244	-79.359323	Historic Site
3	M5A	Downtown Toronto	Regent Park, Harbourfront	43.65426	-79.360636	Cooper Koo Family YMCA	43.653249	-79.358008	Distribution Center
4	M5A	Downtown Toronto	Regent Park, Harbourfront	43.65426	-79.360636	Distillery Sunday Market	43.650075	-79.361832	Farmers Market



List of postal codes of Canada: M

From Wikipedia, the free encyclopedia

This is a list of postal codes in Canada where the first letter is M. Postal codes beginning with M are located within the city of Toronto in the province of Ontario. Only the first three characters are listed, correspon

Canada Post provides a free postal code look-up tool on its website,[1] via its applications for such smartphones as the iPhone and BlackBerry,[2] and sells hard-copy directories and CD-ROMs. Many vendors als match addresses and postal codes. Hard-copy directories can also be consulted in all post offices, and some libraries.

Toronto - 103 FSAs

Note: There are no rural FSAs in Toronto, hence no postal codes should start with M0. However, the postal code M0R 8T0 is assigned to an Amazon warehouse in Mississauga, and the postal code M0R 2A2 is suggesting that Canada Post may have reserved the M0 FSA for high volume addresses.

M1A *Not assigned*	M2A *Not assigned*	M3A North York (Parkwoods)	M4A North York (Victoria Village)	M5A Downtown Toronto (Regent Park / Harbourfront)	M6A North York (Lawrence Manor / Lawrence Heights)	M7A Queen's Park (Ontario Provincial Government)
M1B Scarborough (Malvern / Rouge)	M2B *Not assigned*	M3B North York (Don Mills) North	M4B East York (Parkview Hill / Woodbine Gardens)	M5B Downtown Toronto (Garden District, Ryerson)	M6B North York (Glencairn)	M7B *Not assigned*
M1C Scarborough (Rouge Hill / Port Union / Highland Creek)	M2C *Not assigned*	M3C North York (Don Mills) South (Flemingdon Park)	M4C East York (Woodbine Heights)	M5C Downtown Toronto (St. James Town)	M6C York (Humewood-Cedarvale)	M7C *Not assigned*

Based on the problem the Stakeholders have presented, the following data will be used for the analysis:

Foursquare location data to identify how many breweries are in a given area.

Wikipedia to retrieve postal code data for Toronto.

Geospatial data for Toronto that will provide appropriate longitude and latitude coordinates to create necessary dataframes.

METHODOLOGY

BEAUTIFULSOUP

The initial process will be to collect the data and then extract the data for the city Toronto. This will provide necessary Postal Codes, Boroughs and Neighborhoods. This information will be scraped using BeautifulSoup.

```
In [5]: df.head()
```

Out[5]:

	PostalCode	Borough	Neighborhoods
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Queen's Park	Ontario Provincial Government

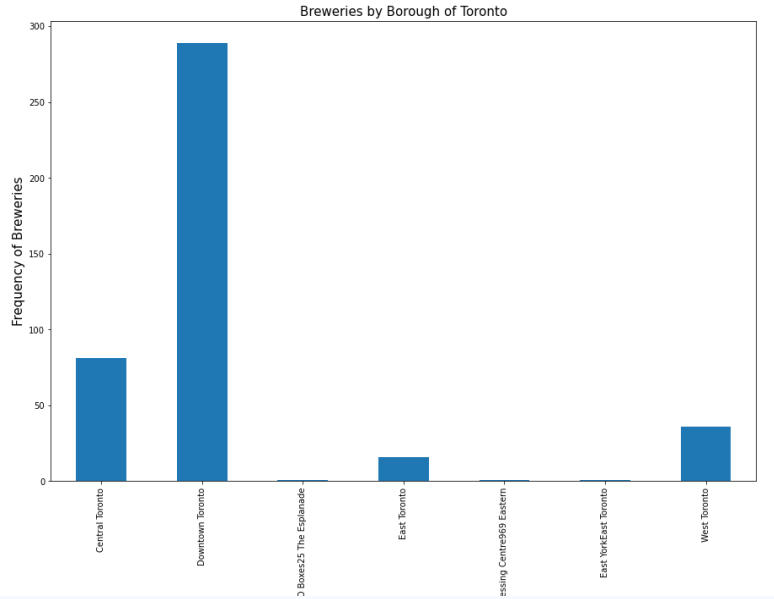
CREATE DATAFRAME AND MAP

Next, clean and merge the data with Geospatial data in order to create a complete DataFrame.

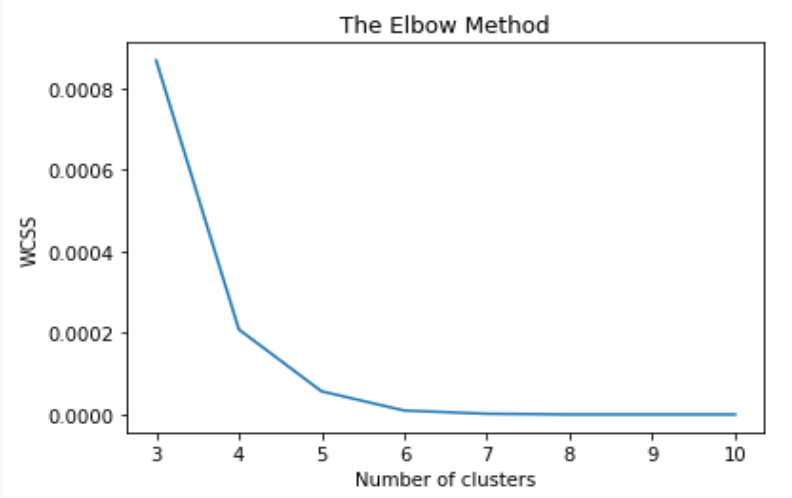
Once the DataFrame is created, A Folium Map will be generated. Then using Foursquare location data, K-Means Clustering will be used to identify specific clusters of Boroughs that will be desirable locations for a new Brewery.

METHODOLOGY

Downtown Toronto shows the highest frequency of Breweries; however, it was important to proceed with the cluster analysis to identify if there were opportunities within the identified clusters.



Before proceeding with K-Means, the Elbow Method was used to determine the number of clusters that will provide the best K value.



RESULTS

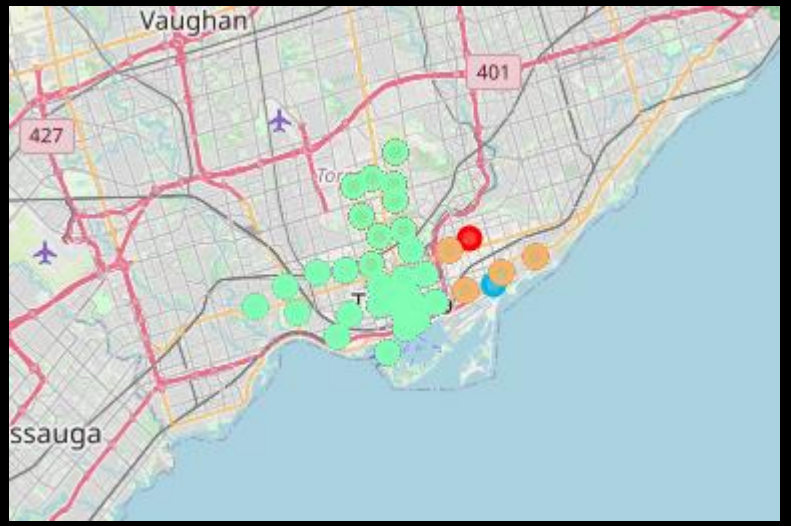
Cluster 0: This clusters consists primarily of Downtown, Central and West Boroughs, has no breweries.

Cluster 1: Mostly Downtown, though in different Neighborhoods, shows some of the highest frequency of Breweries.

Cluster 2: This cluster seems to be the least suitable due to the surrounding vicinity.

Cluster 3: This cluster, Downtown, Central, West and some East Boroughs has Breweries, though some of the lowest frequency.

Cluster 4: This cluster in East Toronto Shows the highest frequency, and would most likely be the most competitive.



DISCUSSION

Based on the above results, Clusters 2 and 4 should be eliminated as options due to the vicinity and high frequency respectively.

Cluster 1 also shows high frequency of breweries which still may cause increase competition, which is not suitable for the stakeholders.

Based on the analysis, Cluster 0 should be the primary cluster for a new Brewery.

Out[38]:

	PostalCode	Borough	Neighborhoods	Latitude	Longitude	Brewery	Cluster Labels
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	0.0	0
253	M4X	Downtown Toronto	St. James Town, Cabbagetown	43.667967	-79.367675	0.0	0

Cluster 3 would be the secondary Cluster which may also yield strong opportunity as there are some breweries which may prove to have an already established client base.

	PostalCode	Borough	Neighborhoods	Latitude	Longitude	Brewery	Cluster Labels
46	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	0.01	3
408	M4T	Central Toronto	Moore Park, Summerhill East	43.689574	-79.383160	0.01	3
13	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	0.01	3

CONCLUSION

This Project was to identify for the stakeholders a new location for opening a Brewery, while mitigating too close of proximity to other Breweries.

Overall, Downtown Toronto shows some of the highest frequency of Breweries.

However, by using k Means Clustering, it has been identified that in Clusters 0 and 3 there are opportunities for the stakeholders to review in the Downtown Areas which either have no Breweries or very little.

These locations would likely yield the best opportunity for success due to the concentration of activity in the Downtown Toronto area.