Distribution Network of Craft Beer for a Start-Up

Eden Cheng

July 30, 2019

1. Introduction

1.1 Background

Our client is a startup craft beer brewer, they look for a distribution network of their craft beer in one of the area in Toronto. Since the supply of craft beer is limited, high target selling price, and the special flavor of the craft beer, they would like to find out the most suitable area to maximize their profit.

1.2 Challenge

There are some challenge that client would like to study, i) Craft beer with limited supply and "best tasting period", ii) High target selling price (i.e. 60% more expensive than branded beers e.g. Heineken, Budweiser) and iii) Special flavour like herbs, sours, salty lemon, etc. (Asian flavour)

1.3 Requirement from client

Client looks for an area with lots of bar/pubs/restaurants, to ensure that people in that area are willing to spend money on foods and drinks. Besides, area with Asian restaurants is preferred, as client thinks that it would be a selling point of craft beers with special flavor from Asia. Lastly, people are willing and affordable to spend money on the beers.

2. Data acquisition and cleaning

2.1 Data sources

Data could be found from below URL / API

- FourSquare developer API
- <List of postal codes of Canada> https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- <Latitude and Longitude data> http://cocl.us/Geospatial_data

2.2 Data cleaning

Data downloaded or scraped from multiple sources were combined into one table. There

were some of missing values from the source “List of postal codes of Canada”. In this regard, I simply discarded the record without value for both the fields “Borough” and “Neighbourhood”, or copy the value of “Borough” to “Neighbourhood”. (see below)

Postcode	Borough	Neighbourhood
M1A	Not assigned	Not assigned
M2A	Not assigned	Not assigned
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Harbourfront
M5A	Downtown Toronto	Regent Park
M6A	North York	Lawrence Heights
M6A	North York	Lawrence Manor
M7A	Queen's Park	Not assigned
M8A	Not assigned	Not assigned
M9A	Etobicoke	Islington Avenue
M1B	Scarborough	Rouge

3. Data analysis

To form a set of data for analysis, I scraped the data from the URL <List of postal codes of Canada>, which contains Postal code, Borough and Neighbourhood. Besides, I downloaded the Latitude and Longitude data and merged the 2 datasets together, so that we could find out the Venues based on the location. (see below)

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917

After that, based on the Latitude and Longitude, search nearby (i.e. radius = 1500m) Venue details of each “Neighbourhood” via FourSquare API, such as Venue name, Venue location, Category for further analysis. (see below)

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue ID	Category ID
0	Rouge, Malvern	43.806686	-79.194353	Images Salon & Spa	43.802283	-79.190565	Spa	4c669cbe63665481c948fa33	4bf58dd8d48988d144941735
1	Rouge, Malvern	43.806686	-79.194353	Canadians exhibit	43.817962	-79.193374	Zoo Exhibit	4c97975582b56dcb8320ebea	58daa1558bbb0b01f18ec1fd
2	Rouge, Malvern	43.806686	-79.194353	Caribbean Wave	43.798558	-79.195777	Caribbean Restaurant	4b914562f964a520d49ae33e3	4bf58dd8d48988d144941735
3	Rouge, Malvern	43.806686	-79.194353	Staples Morningside	43.800285	-79.196607	Paper / Office Supplies Store	4ccb612d3740b713f0606285	4bf58dd8d48988d121951735

Client looked for an area with lots of bar/pubs/restaurants with pricey menu, to ensure that people in that area are willing to spend money on foods and drinks, and preferable

to have Asian restaurants as client thinks that it would be a selling point of craft beers with special flavor from Asia. Therefore, I filtered out the records with key words like Restaurant, Asian, Japanese, Bar, Lounge, etc to indicate the type of venue. (see below)

Neighbourhood	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue ID	Category ID	Target_Restaurant
Guildwood, Morningside, West Hill	House Of Wong	43.773841	-79.183125	Asian Restaurant	4c02f2380d0e0f4764ea019a	4bf58dd8d48988d142941735	Yes
Woburn	Govardhan Thal	43.767496	-79.228476	Indian Restaurant	4bd48e734e32d13e4254b180	4bf58dd8d48988d10f941735	Yes
Woburn	Booster Juice	43.766159	-79.205821	Juice Bar	4de679298877e085f8ca16fe	4bf58dd8d48988d112941735	Bar
Woburn	Pho Dau Bo	43.761286	-79.225170	Vietnamese Restaurant	4b9ba5d8f964a5200a8336e3	4bf58dd8d48988d14a941735	Yes
Woburn	A&W	43.761066	-79.228290	Fast Food Restaurant	5b0d5ad9f8e9d7002c0db116	4bf58dd8d48988d16e941735	No

Further, I summarized the “Target Restaurant” and form a table which indicated the number of Bar, Asian Restaurant and Other Restaurant in each “Neighbourhood”, calculated their percentage, and Bar-to-Restaurant & Asian-to-Other Restaurant Ratio. After that, I use K-Means Clustering method to categorize the Neighbourhood and confirmed to select Category 3 as our targeted areas for further study, as they have good Bar-to-Restaurant & Asian-to-Other Restaurant Ratio.

Target_Restaurant Neighbourhood	Bar	No	Yes	Sum	Bar%	No%	Yes%	Yes_No_Ratio	Bar_Restaurant_Ratio	Cluster Labels
Cabbagetown, St. James Town	9	13	19	41	21.951220	31.707317	46.341463	1.461538	0.281250	0
Willowdale South	2	12	25	39	5.128205	30.769231	64.102564	2.083333	0.054054	0
Brockton, Exhibition Place, Parkdale Village	10	23	11	44	22.727273	52.272727	25.000000	0.478261	0.294118	1
Business Reply Mail Processing Centre 969 Eastern	14	18	10	42	33.333333	42.857143	23.809524	0.555556	0.500000	1
Chinatown, Grange Park, Kensington Market	11	23	8	42	26.190476	54.761905	19.047619	0.347826	0.354839	1
Christie	12	17	10	39	30.769231	43.589744	25.641026	0.588235	0.444444	1
Harbord, University of Toronto	13	22	5	40	32.500000	55.000000	12.500000	0.227273	0.481481	1
Studio District	11	16	13	40	27.500000	40.000000	32.500000	0.812500	0.379310	1
Dovercourt Village, Dufferin	19	18	7	44	43.181818	40.909091	15.909091	0.388889	0.760000	2
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

To prioritize the areas, we search extra venues information via FourSquare API, such as Rating and Price, to check if those Bar/Restaurant have good reputation and pricey menu. We got the average value of Rating and Price for each “Neighbourhood, and normalized them to come up a Scoring for our final judgement. (see below)

Target_Restaurant Neighbourhood	Bar	No	Yes	Sum	Bar%	No%	Yes%	Yes_No_Ratio	Bar_Restaurant_Ratio	Cluster Labels	Rating	Price	Norm_Rating	Norm_Price	Scoring
Little Portugal, Trinity	17	21	10	48	35.416667	43.750000	20.833333	0.476190	0.548387	0	8.606250	2.170213	1.000000	1.000000	2.000000
Christie	11	18	10	39	28.205128	46.153846	25.641026	0.555556	0.392857	0	8.253846	2.027027	0.667851	0.685946	1.353807
Studio District	12	19	9	40	30.000000	47.500000	22.500000	0.473684	0.428571	0	8.355000	1.975000	0.480449	0.571833	1.052282
Cabbagetown, St. James Town	9	17	15	41	21.951220	41.463415	36.585366	0.882353	0.281250	0	8.260976	1.714286	0.674630	0.000000	0.674630
Business Reply Mail Processing Centre 969 Eastern	9	23	10	42	21.428571	54.761905	23.809524	0.434783	0.272727	0	7.545238	2.000000	0.000000	0.620667	0.620667

4. Conclusion

The study is based on geographical data scarped from URL and venue details via FourSquare API. By forming a set of relevant data, I selected relevant type of venues,

performed basic statistically analysis, and clustered them by using K-Means Clustering. With the help of venue details from FourSquare (i.e. rating and Price index), I could prioritize the selected areas and finally identified “Little Portugal, Trinity” is the best place for client to set up their distribution network of their Craft Beer. It is because there are many good rating bars and restaurants with pricey menu in this area, which implied a larger market. High ratio of Asian restaurant, which implied a higher chance that people accept the special flavour of Craft beer.

In case deeper study is required, we could extract extra data from FourSquare (e.g. menu, comment, etc) for a more sophisticated analysis.