

IBM Data Science Professional: Capstone Project Report
Optimal Locations to Open a Cafe with a Bookstore in Vancouver

Mike Hsiao

July 31st 2020

Introduction/Problem

The Metro Vancouver region is home to a great coffee culture - especially boutique micro roasters and small businesses. In this project, we will find the best city in the Metro Vancouver Region to open up a coffee shop.

Target Audience and Benefits of this Project

The target audience is for anyone/business owner who is looking to find a good location in Metro Vancouver to open up a coffee shop. The benefits of this project is that we will use Foursquare data and postal code data from Wikipedia and find out which regions have clusters of businesses that do not currently have a coffee shop. Thus, these locations are great for starting such a business.



Figure 1 - Picture of Vancouver from the City of Vancouver's Website

2.0 Approach With Data

2.1 Location Data

Using the data from Wikipedia of the different cities in Metro Vancouver, we can use it to find population data as well as geographical data.

https://en.wikipedia.org/wiki/Metro_Vancouver_Regional_District

2.2 Economic Data

Using property values as economic data, we can determine which city in the Metro Vancouver Region has the highest income. Higher income means that the residents of that region can afford more boutique style coffee.

<https://info.bcassessment.ca/news/Pages/Lower-Mainland-2020-Property-Assessments-in-the-Mail.aspx>

2.3 Foursquare Data

Using data from Foursquare, we can find the cities in Metro Vancouver that already have coffee shops and bookstores and which cities currently has a smaller number of coffee shops and bookstores. Using the API data, we can determine where would be the best areas to open up such a store.

3.0 Methodology

In this section we discuss the methodology to obtain data to help us determine which location in Metro Vancouver would be the best to open up a coffee shop with a bookstore. With a combination of data processing and the use of Foursquare API, we can determine which areas are the best.

3.1 Data Processing and Preparation

The main purpose of this project is to determine where in the Metro Vancouver Region would be the best place to have a coffee shop/bookstore. The first step after loading all necessary libraries is to find the data. As described in the sections above, Wikipedia was first used to locate a table with the names of different cities in Metro Vancouver. Then, irrelevant columns such as past population data was dropped. Because location data was not included on the Wikipedia page, additional location data was obtained from www.longlat.net. The data was obtained individually and then saved as a dataframe before being combined with the table taken from Wikipedia. The table is shown in Figure 2 below.

Out[7]:

	Name	Regional district[2]	Population (2016)[5]	Population density[5]	Latitude	Longitude
2	Burnaby	Metro Vancouver	232755	2568.7	49.267132	-122.968941
7	Coquitlam	Metro Vancouver	139284	1138.9	49.283764	-122.793205
11	Delta	Metro Vancouver	102238	567.4	49.084721	-123.058609
22	Langley	Metro Vancouver	25888	2533.6	49.074329	-122.559319
23	Maple Ridge	Metro Vancouver	82256	308.3	49.216667	-122.599998
27	New Westminster	Metro Vancouver	70996	4543.4	49.206944	-122.911110
28	North Vancouver	Metro Vancouver	52898	4465.1	49.316666	-123.066666
31	Pitt Meadows	Metro Vancouver	18573	214.7	49.233334	-122.683334

Figure 2 - A section of the table of cities in Metro Vancouver used as data.

3.2 Using Foursquare API to Obtain Nearby Venues

After the required location data is captured in the previous section, the locational information is passed to the Foursquare API to obtain a list of nearby populations. The first city, Burnaby is passed to the API as a to obtain values. The values are verified for accuracy before a function GetNearbyVenues is used that contains the data from the previous section. The function returns a list of popular locations in different cities in Metro Vancouver. The GetNearbyVenues function returns a total of 269 entries.

[27]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Burnaby	49.267132	-122.968941	Sarlo's Awesome Eatery	49.266429	-122.971001	Diner
1	Burnaby	49.267132	-122.968941	Sumas Park	49.265564	-122.969760	Park
2	Burnaby	49.267132	-122.968941	Halifax Park	49.268722	-122.962963	Playground
3	Coquitlam	49.283764	-122.793205	Evergreen Cultural Centre	49.285535	-122.790402	Art Gallery
4	Coquitlam	49.283764	-122.793205	City Centre Aquatic Complex	49.285570	-122.794140	Pool
5	Coquitlam	49.283764	-122.793205	Lafarge Lake	49.285767	-122.787858	Lake
6	Coquitlam	49.283764	-122.793205	Urban Gate Bar & Grill	49.281712	-122.796175	Bar
7	Coquitlam	49.283764	-122.793205	The Coffee Bun	49.279554	-122.793293	Café
8	Coquitlam	49.283764	-122.793205	Clever Cupcakes	49.280934	-122.795730	Cupcake Shop
9	Coquitlam	49.283764	-122.793205	Kulinarya	49.282102	-122.799525	Filipino Restaurant
10	Coquitlam	49.283764	-122.793205	Caffé Divano	49.283851	-122.793524	Coffee Shop
11	Coquitlam	49.283764	-122.793205	Legend House 天府	49.282209	-122.798716	Asian Restaurant
12	Coquitlam	49.283764	-122.793205	Snowy Village (Coquitlam)	49.281986	-122.793321	Dessert Shop
13	Coquitlam	49.283764	-122.793205	Ninja Bubble Tea	49.282695	-122.796946	Bubble Tea Shop
14	Coquitlam	49.283764	-122.793205	Browns Socialhouse Town Centre	49.282674	-122.798201	Gastropub
15	Coquitlam	49.283764	-122.793205	Grand Palace Seafood Restaurant	49.279897	-122.793852	Chinese Restaurant

Figure 3 - The data that the function GetNearbyVenues obtains from Foursquare API

3.3 Use of One Hot Encoding to Find Frequency

With the data from Figure 3, the technique of One Hot Encoding is used to categorize each business and convert them into 0's and 1's. Afterward, a function is used to group the venues by city in Metro Vancouver before being placed in a function to calculate frequency of each category per city.

```

----Coquitlam----
      venue  freq
0      Dessert Shop  0.06
1           Café  0.06
2      Art Gallery  0.06
3      Gastropub  0.06
4 Furniture / Home Store  0.06

```

Figure 4 - Once frequencies are determined per city, every city's top 5 categories will be displayed like this figure.

After the frequencies are determined, the top most common venues of each city is the output. From that output, the information can be put in a dataframe to create a table that visualizes the most common venues in a city in Metro Vancouver.

Results

The results showed the frequencies of the top 10 most common venues in each city in the Metro Vancouver region. The table can be seen in figure 5 below.

out[39]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Burnaby	Playground	Park	Diner	Gym	Cosmetics Shop	Cupcake Shop	Department Store	Dessert Shop	Electronics Store	Falafel Restaurant
1	Coquitlam	Dessert Shop	Filipino Restaurant	Chinese Restaurant	Coffee Shop	Café	Pool	Lake	Bubble Tea Shop	Korean Restaurant	Clothing Store
2	Delta	Gym	Coffee Shop	Athletics & Sports	Bus Station	Fish & Chips Shop	Department Store	Dessert Shop	Diner	Electronics Store	Falafel Restaurant
3	Langley	Diner	Yoga Studio	Fish & Chips Shop	Cupcake Shop	Department Store	Dessert Shop	Electronics Store	Falafel Restaurant	Farmers Market	Fast Food Restaurant
4	Maple Ridge	Chinese Restaurant	Restaurant	Coffee Shop	Grocery Store	Sushi Restaurant	Pharmacy	Burger Joint	Nightlife Spot	Diner	Furniture / Home Store
5	New Westminster	Coffee Shop	Comedy Club	Pub	Historic Site	Sandwich Place	Gym	Japanese Restaurant	Mexican Restaurant	Other Great Outdoors	Cocktail Bar
6	North Vancouver	Gym / Fitness Center	Café	Chinese Restaurant	Vietnamese Restaurant	Greek Restaurant	Fast Food Restaurant	Department Store	Dessert Shop	Diner	Electronics Store
7	Pitt Meadows	Hotel	Liquor Store	Pub	Fish & Chips Shop	Cupcake Shop	Department Store	Dessert Shop	Diner	Electronics Store	Falafel Restaurant
8	Port Coquitlam	Coffee Shop	Sandwich Place	Ice Cream Shop	Bakery	Pizza Place	American Restaurant	Gym / Fitness Center	Falafel Restaurant	Japanese Restaurant	Convenience Store
9	Port Moody	Coffee Shop	Burger Joint	Grocery Store	Gym	Bank	Frozen Yogurt Shop	Japanese Restaurant	Farmers Market	Park	Pharmacy
10	Richmond	Chinese Restaurant	Clothing Store	Coffee Shop	Vietnamese Restaurant	Sushi Restaurant	Korean Restaurant	Bubble Tea Shop	Bank	Bakery	Fast Food Restaurant
11	Surrey	Gym	Fast Food Restaurant	Coffee Shop	Bank	Pharmacy	Burger Joint	Sandwich Place	Japanese Restaurant	Shipping Store	Gastropub
12	Vancouver[a]	Chinese Restaurant	Bubble Tea Shop	Dessert Shop	Coffee Shop	Outdoor Sculpture	Light Rail Station	Bus Stop	Filipino Restaurant	Diner	Electronics Store
13	White Rock	Japanese Restaurant	Tapas Restaurant	American Restaurant	Seafood Restaurant	Food & Drink Shop	Ice Cream Shop	Indian Restaurant	Fish & Chips Shop	Dessert Shop	Museum

Figure 5 - The resulting table that shows the most common venues in Metro Vancouver.

Discussion

From the table in figure 5, it is evident that coffee shops are prevalent across many cities in the Metro Vancouver region. For the sake of simplicity, a few assumptions will be made in this model to determine which city is the best to have a coffee shop / bookstore. Some of these assumptions are:

- Disregarding economic data and shopping trends amongst consumers - As economic data for each consumer in each city is hard to find, the correlation between if a consumer is more likely to drink coffee or not is furthermore difficult to find. Because of this, this project disregarded economic data and shopping trends amongst consumers.
- Disregarding consumer preference - As coffee shops with bookstores tend to be a niche, for the sake of discussion, bookstores data will be omitted. It is assumed that consumers who likes coffee shops will also enjoy books.
- Business competition - assuming that if there are cities with a high frequency of coffee shops, it can be assumed that the market is saturated and that opportunity to open a coffee shop is not present.
- Disregarding population - population is usually a big factor in decision making when it comes to local markets. However, for a more straightforward analysis, this is also omitted.

After considering all of these points of discussion. We can refer back to the product - figure 6, which presents the top 10 category of venues in each city. Cities without coffee shops or cafes in its top 10 include: Burnaby, Langley, Pitt Meadows and White Rock. This means that these 4 cities in the Metro Vancouver Region would be most suitable to open up a coffee shop due to lack of coffee shops in that area.

Conclusion

In conclusion, the best cities in the Metro Vancouver Region to open up a coffee shop would be in the cities of Burnaby, Langley, Pitt Meadows and White Rock.

References

- Wikipedia https://en.wikipedia.org/wiki/Metro_Vancouver_Regional_District
- Longitude and Latitude Information: www.longlat.net
- Coursera Lab: Segmenting and Clustering Neighbourhoods in New York City