Market Segmentation for

Senses Coffee company

Hayfa Bu Hazzaa

1. Introduction

1.1Background

Senses is a coffee company that produce three brands of coffee, Ellite, Instance, and Chill. It is based on US. The owners and shareholders decided that it is the time to expand and penetrate another market. Canada is the first option to investigate since it is near to US and the market is somehow similar in taste to US, so no need to customize the product to suite the target country culture. Canada is very big country that is why it was initially decided to dedicate three offices to manage the operations in the country. Also it was decided to focus on specific market segments which are Hotels, Bakeries and Grocery Stores. The reasons of this selection are:

1. Senses Coffee has three main coffee brands that dedicated to Hotels, Bakeries and grocery stores.

	*Ellite* is fine coffee suitable for hotels markets,
	*Instance* is average coffee that is suitable from taste and price for Bakeries markets,
	***Chill*** is an iced coffee that is very popular with good price and suitable for Grocery stores markets.

2. Canada is very big country, and to penetrate it you need first to test the market and focus on the most venues that would generate good money.

The company investors are willing to invest in Canada market, but still need to see solid data about Canada market and specifically about the three market segments it will target. To be able to estimate the amount of the investment needed based on the initial market size. Data scientist team was hired to provide the investors with needed data. Data scientist team will need to collect the concerned data, clean it and then analyse it, this shall provide the needed insight for the investors.

1.2 Problem

Senses Coffee company hired data scientist to collect and analyse the data based on geo-demographic market segmentation. to provide insight that will drive the decision of entering the Canadian market and segment it across the three operation offices.

Essential data which will contribute to the market segmentation process includes: Canadian municipalities, Canadian municipalities' geographical coordinates (i.e. the longitudes and latitude data of the city), and the number and the category of potential customers in each municipality like hotels, Grocery Store and Bakery).

This project is a data clustering project and it is aimed to segment the Canadian municipalities into three marketing segments (i.e. 3 clusters). Although the above-mentioned data will contribute in the segmentation process, the segmentation itself will be done according to the number and the category of potential customers in each municipalities (i.e. the number of hotels, the number of Grocery stores and the number of Bakeries). In this approach, each operation office will operate a single of the new market.

Figure 1 below depicts an example of the potential market segmentation, and Figure 2 depicts an example of a geo-demographic market segmentation based on this descriptive analytics data clustering approach.

Figure 1 - Potential market segmentation

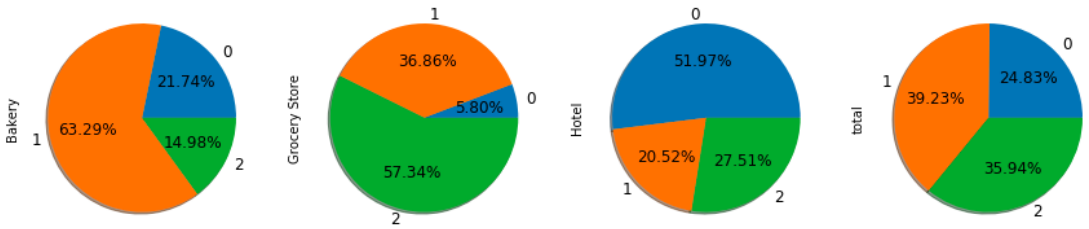
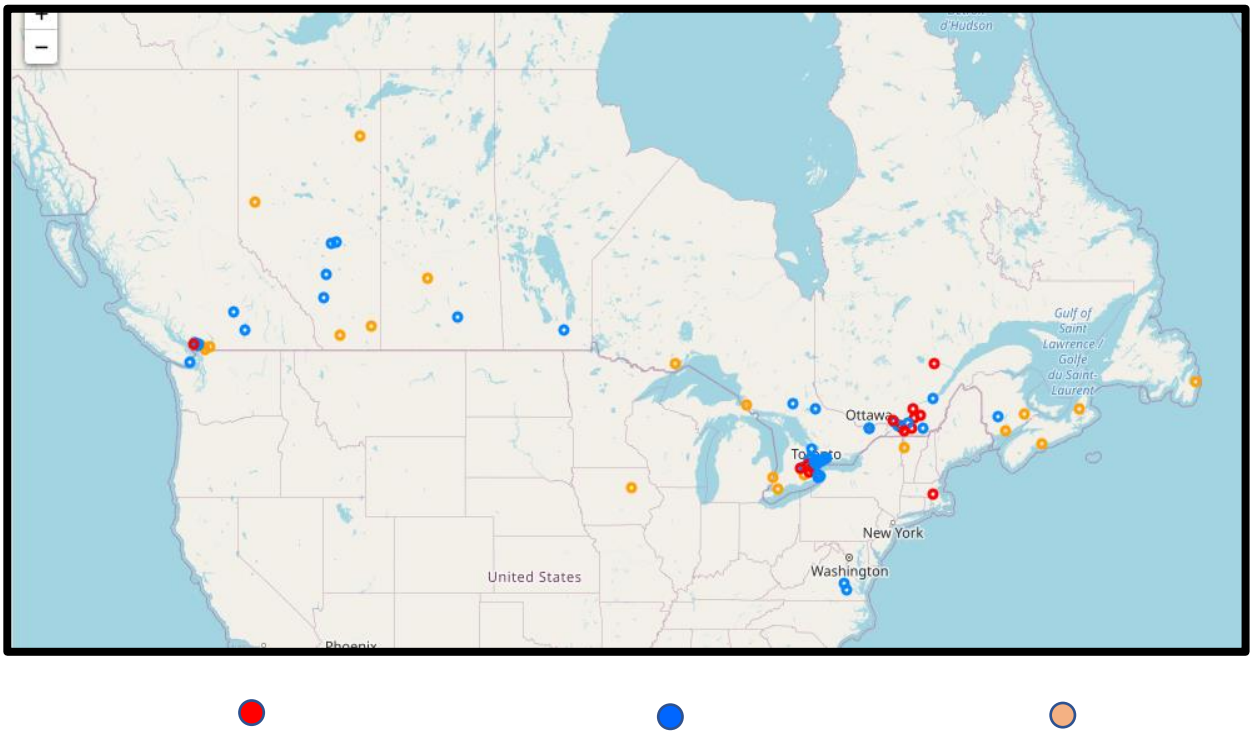


Figure 2 - Geo-demographic market segmentation



1.3 Interest

Senses Coffee company needs to select the municipalities that will be targeted by operation offices. Therefore, the company management and operation offices are interested to know the size of the new potential market in each segment. Other audiences who care about this problem include the company shareholders and Human capital department (see Table 1 below for an example).

Table 1 - An example of potential market segment size

(expressed in number of customers in each business line in each cluster)

Cluster Labels	Bakery	Grocery Store	Hotel
0	58	205	76
1	139	65	66
2	18	18	99

2. Data acquisition and cleansing

2.1Data sources

Table 2 below describes the datasets used to build the clusters and their corresponding data sources.

Table 2 - the datasets and their data sources

No	Dataset	Description	Data Source
1	List of the largest 100 municipalities in Canada	Data fields include Municipalities, Province, Growth rate and population. See Appendix I for an example of this dataset.	I scraped the following Wikipedia site to obtain this data https://en.wikipedia.org/wiki/List_of_the_100_largest_municipalities_in_Canada_by_population
2	Geo-Location data of each municipalities in Canada	Data fields include the longitude and latitude coordinates of each municipality. See Appendix II for an example of this dataset.	I obtained this data using the Python geocoding web services API.
3	Potential customers' data	Data fields include the venue name, category, longitude and latitude, See Appendix III for an example of this dataset.	I obtained this data by exploring the municipality's venues using the Foursquare API
4	Canada map GIS data	Data of Canada with the largest municipalities. See Appendix IV for an example if this dataset.	I obtained this data using the Folium API

2.2 Data Cleansing

I followed below steps to clean the data in order to be ready for further analysis.

1. I scraped the data of 100 largest municipalities in Canada from the Wikipedia page using Python. After investigation, I discovered that the column names in the Wikipedia page are not put in standard naming convention. a column name Population(2016) use special characters, and this jeopardize the Python program code. So, I modified the column name to Population2016 to include only the standard alphabetic character set.
2. I inserted the latitude and longitude coordinate columns structure to the data frame structure of the table read from the Wikipedia page.

3. I added the coordinates data from the geocoding web services and include it in the data frame.
4. Then I checked for any missing coordinates to drop nan values cells. Fortunately, all coordinates data were successfully retrieved by the API. Then I combined the venue data with the location data and the master data acquired from the Wikipedia (Table 3).
5. I then used Folium to create Canada map with all municipalities superimposed on top and used this map to visually verify the correctness of acquired data on the map (see Appendix IV).

Table 3 - Combined Wikipedia data, location data and venue data

	Municipality	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
8251	North Vancouver	49.320713	-123.073783	Indian Fusion	49.327991	-123.072434	Indian Restaurant
8252	North Vancouver	49.320713	-123.073783	Earnest Ice Cream	49.312330	-123.079756	Ice Cream Shop
8253	North Vancouver	49.320713	-123.073783	iLoveKickboxing.com Vancouver	49.283143	-123.100976	Boxing Gym
8254	North Vancouver	49.320713	-123.073783	Waterfront Park	49.311548	-123.085778	Park
8255	North Vancouver	49.320713	-123.073783	Coal Harbour Seawall	49.291304	-123.123276	Trail
8256	North Vancouver	49.320713	-123.073783	Fairmont Pacific Rim	49.288227	-123.116932	Hotel
8257	North Vancouver	49.320713	-123.073783	Revolver	49.283187	-123.109288	Coffee Shop
8258	North Vancouver	49.320713	-123.073783	Stanley Park	49.302488	-123.141718	Park
8259	North Vancouver	49.320713	-123.073783	Loden Hotel	49.287690	-123.123574	Hotel
8260	North Vancouver	49.320713	-123.073783	Rosewood Hotel Georgia	49.283429	-123.118911	Hotel
8261	North Vancouver	49.320713	-123.073783	Lynn Headwaters Regional Park	49.355070	-123.023958	Trail

2.3 Feature Selection

After data cleansing, there were 8719 venues listed under all categories. First, I took out the dup licated rows it was 100 rows and I took it out. To end up with total 8619 venues. Then I listed the unique categories (310 unique categories), this is important to select the right category name th at I need to keep and drop others.

Table 4- the first 10 rows in the venue categories

	Venue Category	Count_Category
0	Park	630
1	Coffee Shop	470
2	Café	341
3	Restaurant	298
4	Grocery Store	288
5	Brewery	246
6	Hotel	236
7	Bakery	210
8	Ice Cream Shop	204
9	Pizza Place	171

The above table gives the confidence that the Canadian market is good and there are many pot ential hotels, Grocery stores and Bakery in Canada. Then I dropped all venues except the venue s with categories (Hotels, Grocery Store and Bakery), the numbers of each venue category is sh own in table 4.

Table 4 – Venues counts of Grocery Store, Hotels, Bakery

	Venue Category	Count_Category
4	Grocery Store	288
6	Hotel	236
7	Bakery	210

Appendix I –Example of the Canadian municipalities in the Wikipedia page

Rank (2016)	Municipality	Province	Municipal status	Land area (km², 2011)	Growth Rate 2011–2016	Population (2016)	Population (2011)	Population (2006)	Population (2001)	Population (1996)
1	***Toronto***	Ontario	City	630.2	4.46%	2,731,571	2,615,060	2,503,281	2,481,494	2,385,421
2	*Montreal*	Quebec	Ville	365.1	3.34%	1,704,694	1,649,519	1,620,693	1,583,590	1,547,030
3	*Calgary*	Alberta	City	825.3	12.99%	1,239,220	1,096,833	988,193	879,003	768,082
4	**Ottawa**	Ontario	City	2,790.2	5.76%	934,243	883,391	812,129	774,072	721,136
5	**Edmonton**	Alberta	City	684.4	14.82%	932,546	812,201	730,372	666,104	616,306
6	Mississauga	Ontario	City	292.4	1.14%	721,599	713,443	668,549	612,925	544,382
7	***Winnipeg***	Manitoba	City	464.1	6.27%	705,224	663,617	633,451	619,544	618,477
8	*Vancouver*	British Columbia	City	115.0	4.64%	631,486	603,502	578,041	545,671	514,008
9	Brampton	Ontario	City	266.3	13.31%	593,638	523,911	433,806	325,428	268,251
10	Hamilton	Ontario	City	1,117.2	3.26%	536,917	519,949	504,559	490,268	467,799
11	**Quebec City**	Quebec	Ville	454.1	2.96%	531,902	516,622	491,142	476,330	473,569
12	Surrey	British Columbia	City	316.4	10.60%	517,887	468,251	394,976	347,820	304,477
13	Laval	Quebec	Ville	247.1	5.34%	422,993	401,553	368,709	343,005	330,393
14	***Halifax***	Nova Scotia	Regional municipality	5,490.3	3.34%	403,131	390,096	372,679	359,111	342,851
15	London	Ontario	City	420.6	4.83%	383,822	366,151	352,395	336,539	325,669
16	Markham	Ontario	City	212.6	9.03%	328,966	301,709	261,573	208,615	173,383
17	Vaughan	Ontario	City	273.5	6.22%	306,233	288,301	238,866	182,022	132,549
18	Gatineau	Quebec	Ville	343.0	4.11%	276,245	265,349	242,124	226,696	217,591
19	*Saskatoon*	Saskatchewan	City	209.6	10.89%	246,376	222,189	202,340	196,861	193,653
20	Longueuil	Quebec	Ville	115.6	3.58%	239,700	231,409	229,330	225,761	227,408
21	Kitchener	Ontario	City	136.8	6.42%	233,222	219,153	204,668	190,399	178,420
22	Burnaby	British Columbia	City	90.6	4.27%	232,755	223,218	202,799	193,954	179,209

Appendix II – Example of Canadian municipalities' Geo-Location data

	Municipality	Province	Land area(km2, 2011)	Growth Rate 2011–2016	Population2016	lat	lng
0	Toronto	Ontario	630.20	4.46%	2731571	43.6535	-79.3839
1	Montreal	Quebec	365.10	3.34%	1704694	45.4972	-73.6104
2	Calgary	Alberta	825.30	12.99%	1239220	51.0534	-114.063
3	Ottawa	Ontario	2790.20	5.76%	934243	45.4211	-75.6903
4	Edmonton	Alberta	684.40	14.82%	932546	53.5354	-113.508
5	Mississauga	Ontario	292.40	1.14%	721599	43.5903	-79.6457
6	Winnipeg	Manitoba	464.10	6.27%	705224	49.8955	-97.1385
7	Vancouver	British Columbia	115.00	4.64%	631486	49.2609	-123.114
8	Brampton	Ontario	266.30	13.31%	593638	43.6858	-79.7599
9	Hamilton	Ontario	1117.20	3.26%	536917	43.2561	-79.8729
10	Quebec City	Quebec	454.10	2.96%	531902	46.826	-71.2352
11	Surrey	British Columbia	316.40	10.60%	517887	51.2715	-0.341452
12	Laval	Quebec	247.10	5.34%	422993	48.071	-0.77235
13	Halifax	Nova Scotia	5490.30	3.34%	403131	44.6486	-63.5859
14	London	Ontario	420.60	4.83%	383822	51.5073	-0.127647
15	Markham	Ontario	212.60	9.03%	328966	43.8543	-79.3268
16	Vaughan	Ontario	273.50	6.22%	306233	43.7942	-79.5268
17	Gatineau	Quebec	343.00	4.11%	276245	45.4284	-75.7106
18	Saskatoon	Saskatchewan	209.60	10.89%	246376	52.1318	-106.661
19	Longueuil	Quebec	115.60	3.58%	239700	45.5172	-73.4467

Appendix III – Example of potential customers data

	Municipality	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Toronto	43.653482	-79.383935	Downtown Toronto	43.653232	-79.385296	Neighborhood
1	Toronto	43.653482	-79.383935	Byblos Toronto	43.647615	-79.388381	Mediterranean Restaurant
2	Toronto	43.653482	-79.383935	Elgin And Winter Garden Theatres	43.653394	-79.378507	Theater
3	Toronto	43.653482	-79.383935	Art Gallery of Ontario	43.654003	-79.392922	Art Gallery
4	Toronto	43.653482	-79.383935	St. Lawrence Market (South Building)	43.648743	-79.371597	Farmers Market
5	Toronto	43.653482	-79.383935	Hailed Coffee	43.658833	-79.383684	Coffee Shop
6	Toronto	43.653482	-79.383935	Alo	43.648574	-79.396243	French Restaurant
7	Toronto	43.653482	-79.383935	Delta Hotels by Marriott Toronto	43.642882	-79.383949	Hotel
8	Toronto	43.653482	-79.383935	Yeti Nails & Spa	43.647938	-79.396330	Cosmetics Shop
9	Toronto	43.653482	-79.383935	Pai	43.647923	-79.388579	Thai Restaurant

Appendix IV – Example of Canadian GIS data