Identification of a

suitable neighbourhood in Toronto

to start a retail store

to sell excellent Mediterranean olive oil

Capstone Project – The Battle of Neighborhoods –
 IBM / Coursera



The challenge is to find the most suitable neighbourhood for setting up a small sales outlet in Toronto to sell excellent Mediterranean olive oil

- A trading company specialized in trading with olive oil would like to expand internationally
- Because of
 - the large number of inhabitants with Mediterranean roots (who very much like to use olive oil)
 - the high average income (some olive oil varieties are expensive) and
 - the large number of Mediterranean restaurants (which are potential clients and can be considered multipliers)

Toronto was chosen to open a new retail store.

 However, Toronto is large. For this reason, the trading company has asked a consulting company to find the most suitable neighbourhood for setting up a small sales outlet in Toronto.



 Olive oil is a liquid fat obtained from olives, a traditional tree crop of the Mediterranean Basin. It is commonly used in cooking, whether for frying or as a salad dressing.

 Olive trees have been grown around the Mediterranean since the 8th millennium BC.

 Choosing a cold-pressed olive oil can be similar to selecting a wine. The flavor of these oils varies considerably and a particular oil may be more suited for a particular dish.



Sources:
Wikipedia,
International Olive Council,
https://atasteofolive.com/products/novello-di-gradassi

Spain is the largest producer of olive oil, followed by Greece and Italy

- Largest Producers of Olive Oil -

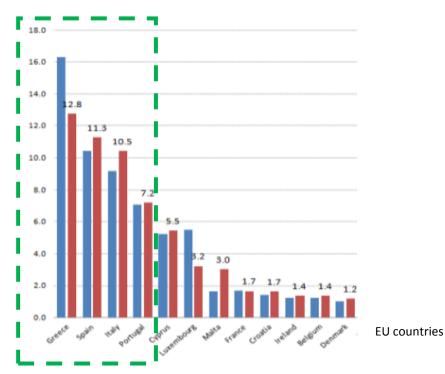
Country	Production [tonnes]	Production [% of global]
Spain	1,290,600	50
Greece	195,000	8
Italy	182,300	7
Turkey	178,000	7
World	2,586,500	100

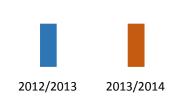
Source: International Olive Council

Per capita consumption is highest in Greece, followed by Spain, Italy, and Portugal

- Per capita olive oil consumption in EU countries [kg] -

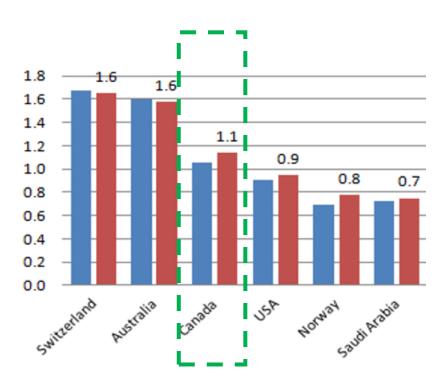
Per capita olive oil consumption [kg]

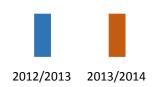




Consumption in Canada is far less, but rising

- Per capita olive oil consumption in some non-EU countries [kg] -





Because of its size, importance within Canada and diverse population Toronto is well suited to be part of the international expansion

- Toronto is the provincial capital of Ontario and the most populous city in Canada, with a population of 2,731,571 in 2016.
- Toronto is a centre of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world.
- The diverse population of Toronto reflects its current and historical role as an important destination for immigrants to Canada. Over 200 distinct ethnic origins are represented among its inhabitants.
- While the majority of Torontonians speak English as their primary language, over 160 languages are spoken in the city.



Business Problem

The trading company is looking for the most suitable neighbourhood for setting up a small sales outlet in Toronto taking into account three main parameters

- The trading company is looking for the most suitable neighbourhood for setting up a small sales outlet in Toronto.
- Important decision criteria are:
 - the **number of inhabitants with Mediterranean roots** as these inhabitants are the consumers who very much like to use olive oil
 - the **average income** in the neighbourhood as the olive oil varieties with the highest profit margins are expensive and
 - the **number of Mediterranean restaurants** as they are potential customers and can be considered multipliers
- The neighbourhood with the best combination of these parameters will be the first choice.

Approach

The analysis will be done using a six step approach

1	Start with the Toronto neighbourhoods from the assignment of week 3
2	Choose only the neighbourhoods with strong Mediterranean influence (a Mediterranean language as second most important language)
3	Add information on average income and number of inhabitants with Mediterranean roots using information from Wikipedia
4	Add information on the number of Mediterranean restaurants within each Neighbourhood using Foursquare data
5	Cluster the information from all these sources using k-means
6	Rank the neighbourhoods

The starting point of the analysis will be the table with the neighbourhoods that has already been used in the assignment of week 3 and has been combined with geospatial data

Postcode +	Borough +	Neighbourhood	
M1A	Not assigned	Not assigned	
M2A	Not assigned	Not assigned	
МЗА	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	
M1B	Scarborough	Malvern	
M2B	Not assigned	Not assigned	
МЗВ	North York	Don Mills North	
M4B	East York	Woodbine Gardens	
M4B	East York	Parkview Hill	
M5B	Downtown Toronto	Ryerson	
	Downtown		



Postal Code, Latitude, Longitude M1B,43.8066863,-79.1943534 M1C,43.7845351,-79.1604971 M1E,43.7635726,-79.1887115 M1G,43.7709921,-79.2169174 M1H,43.773136,-79.2394761 M1J,43.7447342,-79.2394761 M1K,43.7279292,-79.2620294 M1L,43.7111117,-79.2845772 M1M,43.716316,-79.2394761 M1N,43.692657,-79.2648481 M1P,43.7574096,-79.273304 M1R,43.7500715,-79.2958491 M1S,43.7942003,-79.2620294 M1T,43.7816375,-79.3043021 M1V,43.8152522,-79.2845772 M1W,43.7995252,-79.3183887 M1X,43.8361247,-79.2056361 M2H,43.8037622,-79.3634517 M2J,43.7785175,-79.3465557 M2K,43.7869473,-79.385975 M2L,43.7574902,-79.3747141 M2M,43.789053,-79.4084928 M2N,43.7701199,-79.4084928 M2P,43.7527583,-79.4000493 M2R,43.7827364,-79.4422593 M3A,43.7532586,-79.3296565 M3B,43.7459058,-79.352188 M3C 43 7258997 -79 340923



Po	stal Code	Borough	Neighbourhood	Latitude	Longitude
0	МЗА	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Harbourfront	43.654260	-79.360636
3	M5A	Downtown Toronto	Regent Park	43.654260	-79.360636
4	M6A	North York	Lawrence Heights	43.718518	-79.464763
5	M6A	North York	Lawrence Manor	43.718518	-79.464763
6	M7A	Queen's Park	Queen's Park	43.662301	-79.389494
7	M9A	Etobicoke	Islington Avenue	43.667856	-79.532242
8	M1B	Scarborough	Rouge	43.806686	-79.194353
9	M1B	Scarborough	Malvern	43.806686	-79.194353

Source:

https://en.wikipedia.org/wiki/List_of_postal codes of Canada: M

Source:

"http://cocl.us/Geospatial_data"

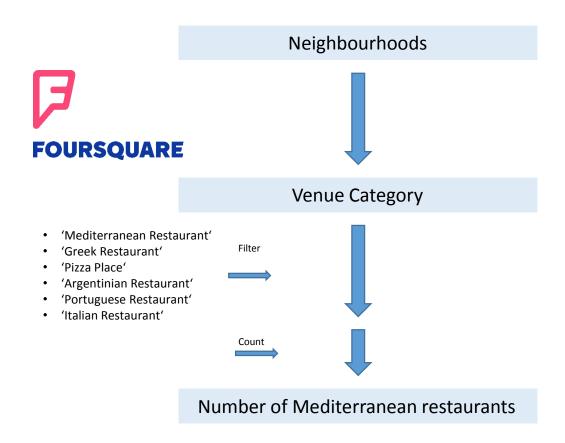
To filter the neighbourhoods on inhabitants with Mediterranean roots a Wikipedia showing the "second most common language" within a neighbourhood will be used

Name	♦ FM ♦	Census Tracts ≑	Population +	Land area ¢ (km2)	Density (people/km2)	% Change in Population since 2001	Average Income	† Transit † Commuting %	% Renters	Second most common language (after ¢ English) by name	Second most common language (after + English) by percentage	Мар
Toronto CMA Average		All	5,113,149	5903.63	866	9.0	40,704	10.6	11.4	_		
Agincourt	s	0377.01, 0377.02, 0377.03, 0377.04, 0378.02, 0378.08, 0378.14, 0378.23, 0378.24	44,577	12.45	3580	4.6	25,750	11.1	5.9	Cantonese (19.3%	19.3% Cantonese	Agincourt
Alderwood	Е	0211.00, 0212.00	11,656	4.94	2360	-4.0	35,239	8.8	8.5	Polish (6.2%)	06.2% Polish	
Alexandra Park	OCoT	0039.00	4,355	0.32	13,609	0.0	19,687	13.8	28.0	Cantonese (17.9%	17.9% Cantonese	Story 5 Alternative John St. Service Grant C.
Allenby	OCoT	0140.00	2,513	0.58	4333	-1.0	245,592	5.2	3.4	Russian (1.4%)	01.4% Russian	
Amesbury	NY	0280.00, 0281.01, 0281.02	17,318	3.51	4,934	1.1	27,546	16.4	19.7	Spanish (6.1%)	06.1% Spanish	Amesbury
Armour Heights	NY	0298.00	4,384	2.29	1914	2.0	116,651	10.8	16.1	Russian (9.4%)	09.4% Russian	Armont Helphts
Banbury	NY	0267.00	6,641	2.72	2442	5.0	92,319	6.1	4.8	Unspecified Chinese (5.1%)	05.1% Unspecified Chinese	
Bathurst Manor	NY	0297.01, 0310.01, 0310.02	14,945	4.69	3187	12.3	34,169	13.4	18.6	Russian (9.5%)	09.5% Russian	Authorit 1

The same table can be used to make a ranking on average income and the number of inhabitants with a certain origin

Name 6	FM ÷	Census Tracts	Population \$	Land area ‡ (km2)	Density (people/km2)	% Change in Population since 2001	Average Income	Transit \$	% Renters	Second most common language (after ¢ English) by name	Second most common language (after \$ English) by percentage	Map ≑
Toronto CMA Average		All	5,113,149	5903.63	866	9.0	40,704	10.6	11.4			
Agincourt	S	0377.01, 0377.02, 0377.03, 0377.0 0378.02, 0378.08, 0378.14, 0378.23, 0378.24		12.45	3580	4.6	25,750	1.1	5.9	Cantonese (19.3%	19.3% Cantonese	Agincourt
Alderwood	E	0211.00, 0212.00	11,656	.94	2360	-4.0	35,239	.8	8.5	Polish (6.2%)	06.2% Polish	
Alexandra Park	OC ₀ T	0039.00	4,355	.32	13,609	0.0	19,687	3.8	28.0	Cantonese (17.9%	17.9% Cantonese	Accounted a Service Se
Allenby	OCoT	0140.00	2,513	9.58	4333	-1.0	245,592	5.2	3.4	Russian (1.4%)	01.4% Russian	- A O.
Amesbury	NY	0280.00, 0281.01, 0281.02	17,318	3.51	4,934	1.1	27,546	16.4	19.7	Spanish (6.1%)	06:1% Spanish	Amodbury
Armour Heights	NY	0298.00	4,384	:29	1914	2.0	116,651	0.8	16.1	Russian (9.4%)	09.4% Russian	Armour Beight
Banbury	NY	0267.00	6,641	.72	2442	5.0	92,319	.1	4.8	Unspecified Chinese (5.1%)	05.1% Unspecified Chinese	
Bathurst Manor	NY	0297.01, 0310.01, 0310.02	14,945	4.69	3187	12.3	34,169	3.4	18.6	Russian (9.5%)	09.5% Russian	

To identify the number of Mediterranean restaurants within a neighbourhood Foursquare information will be used and filtered



As a last step a k-means clustering and a ranking will be carried out

Number of inhabitants with Number of Mediterranean Average income Mediterranean roots restaurants K-means-Clustering Ranking