

The city of coffee shops

Clustering Toronto Neighborhoods by Income

By Marcelo Ocampo

Objectives

- Guide potential business owners to open new location
- Group Toronto neighborhoods based on venues and income level
- Provide a picture of the current type of businesses in a location and residents' preferences

— — —

Data Description

— — —

1. Demographic data from Toronto segmented by neighborhood: Number of neighborhoods, Income level, Median income produced by Statistics Canada for the 2016 census.
Retrieved from: <https://open.toronto.ca/dataset/neighbourhood-profiles/>
2. Popular venue data from the Foursquare API: venue name, latitude, longitude and venue category. For more info please visit the Foursquare developer site
3. Data used in analysis:
 - a. 140 neighborhoods in the Toronto first data set
 - b. 10 neighborhoods with no latitude and longitude information
 - c. 15 neighborhoods with less than 4 venues in the Foursquare API

Data extraction and presentation

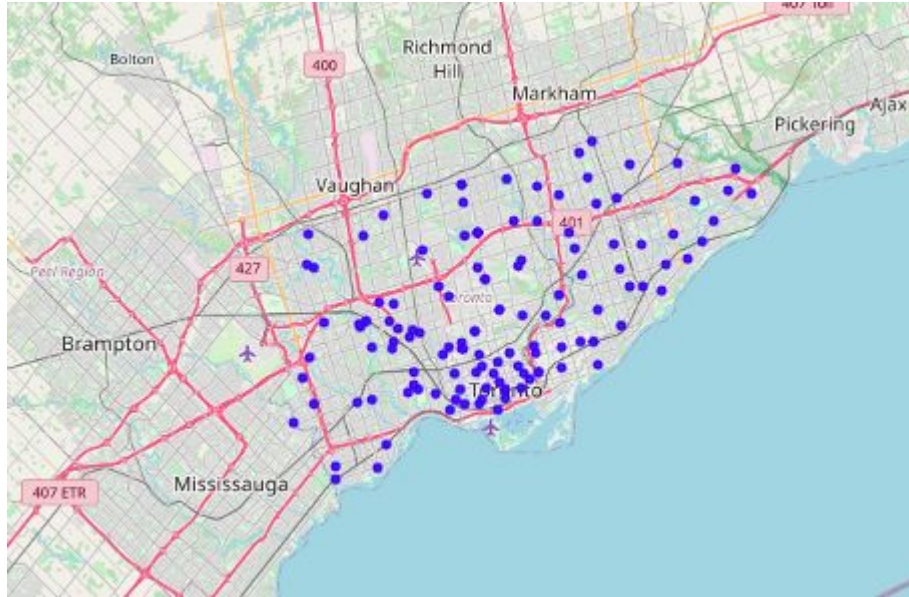
- Downloading data from Statistics Canada, formatting to columns
- Extracting income information and calculate income relative to median Toronto income for each neighborhood

	Neighborhood	Total income	Latitude	Longitude
0	Agincourt North	25005	43.808038	-79.266439
1	Agincourt South	20400	43.785353	-79.278549
2	Alderwood	10265	43.601717	-79.545232
3	Annex	26295	43.670338	-79.407117
4	Banbury	23410	43.742796	-79.369957

	Neighborhood	Percent income
0	Agincourt North	0.831035
1	Agincourt South	0.677989
2	Alderwood	0.341155
3	Annex	0.873907
4	Banbury	0.778025

Toronto neighborhood map

— — —



Getting venues with Foursquare

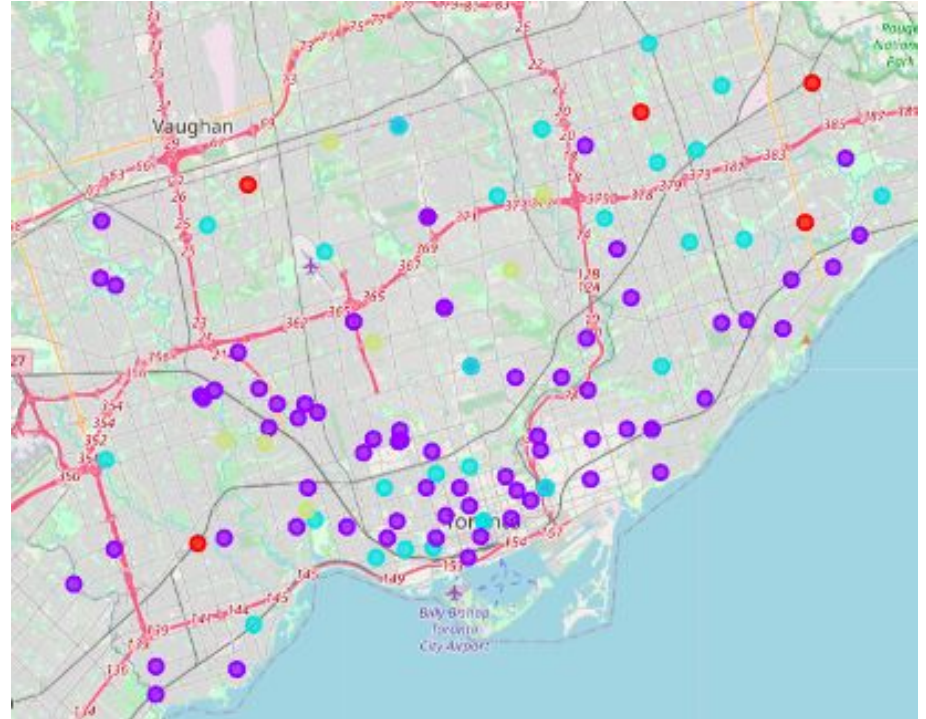
— — —

- Use Foursquare to get the different venues in each neighborhood (using a 500 mts radius)
- Excluding 15 neighborhoods with less than 4 venue results to provide more stability to our results
- Get the information on the top 5 most common venue categories per neighborhood

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Agincourt North	Bank	Bakery	Liquor Store	Pizza Place	Sporting Goods Shop
1	Agincourt South	Chinese Restaurant	Cantonese Restaurant	Hong Kong Restaurant	Coffee Shop	Asian Restaurant
2	Alderwood	Pizza Place	Pharmacy	Gym	Sandwich Place	Pub
3	Annex	Pizza Place	Bistro	Gym	Coffee Shop	Park
4	Banbury	Park	Tennis Court	Auto Garage	Electronics Store	Dog Run

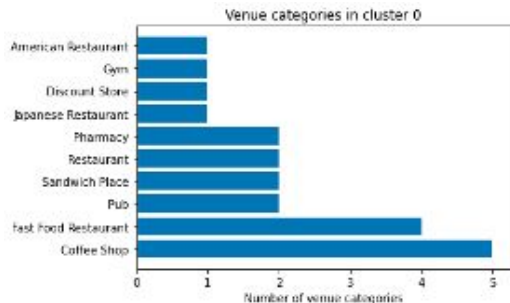
Clustering

Form cluster groups taking into account venue categories and relative income into 4 big clusters

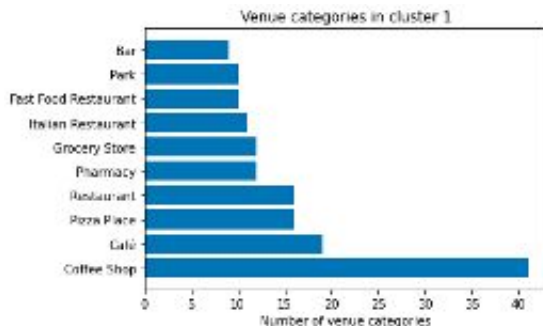


Results

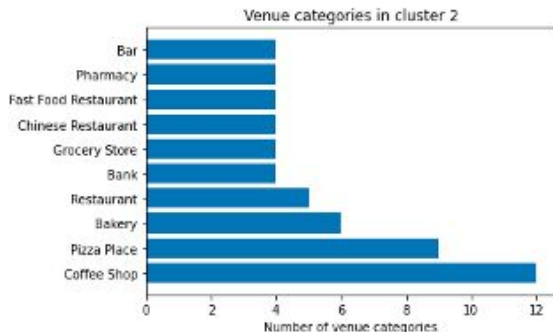
Cluster 1: High Income



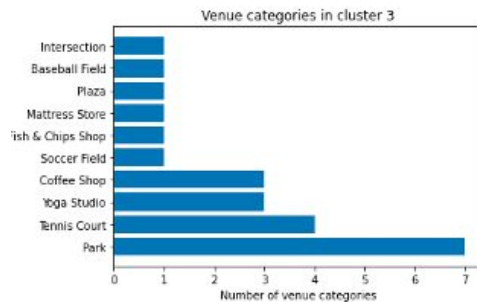
Cluster 2: Low Income



Cluster 3: High-medium Income



Cluster 4: Low-medium Income



Results

- **High income cluster:** tendency to quick eats and fast-food - Strong commercial sector
- **Low income cluster:** full-service restaurants - Residential areas - Further research for price elasticity
- **High-medium income cluster:** day-to-day consumption and dining - Residential
- **Low-medium income cluster:** outdoor entertainment and exercising - No eats

Next Steps

- Further research to provide a more valuable insight on price levels and elasticity
- More information needed to complement this data such as preferences and field research
- Complimentary demographic data depending on business type

