Capstone Project

The Battle of Neighborhoods

Murphy Chan Ngar Kay
October 22, 2020

Introduction and Background

The concept of neighborhoods

- help government and community organizations with their local planning by providing socio-economic data at a meaningful geographic area
- allowing researchers to examine changes over time

This project is

- illustrate how data analysis helps to be creative and come up with different ideas to solve problem
- Building a profile for each neighborhood for potential explanation of why a neighborhood is popular
- The cause of complaints in another neighborhood, or anything else related to neighborhoods

Problem Description

- Consider Hiroshima Company is going to startup a business of opening a sushi restaurant in Toronto which targeting middle class customers with location in :Etobicoke or Scarborough. On top of the food quality, service and any other interior elements of the restaurant, one of the most important thing is the location and neighborhoods of the restaurant location. This project is to illustrate how data impacts on the determination for problem solving. The following analysis mainly compares the 2 boroughs in Toronto :Etobicoke and Scarborough.

Description of Data

1. wiki, will be used for Postal code, Borough, and Neighborhood link : https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
2. The location data for analyzing further information for the neighborhoods with latitude and longitude coordinates are from the following link : http://cocl.us/Geospatial_data
3. Foursquare API will be called for the information of nearby venues and the most common venue to analysis the popularity of special location.
4. Data for the demographic statistic figures : https://open.toronto.ca/

Methodology

Getting the Segmenting and Clustering Neighborhoods, download the data from wiki and join the neighborhood that exist in one postal code area, and then map the Latitude and Longitude data.

Postalcode	Borough	Neighborhood
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Regent Park, Harbourfront
M6A	North York	Lawrence Manor, Lawrence Heights
M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government

Call the Foursquare API from the target location table for the nearby venues.

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Malvern, Rouge	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
Rouge Hill, Port Union, Highland Creek	43.784535	-79.160497	Royal Canadian Legion	43.782533	-79.163085	Bar
Rouge Hill, Port Union, Highland Creek	43.784535	-79.160497	SEBS Engineering Inc. (Sustainable Energy and ...	43.782371	-79.156820	Construction & Landscaping
ildwood, Morningside, West Hill	43.763573	-79.188711	RBC Royal Bank	43.766790	-79.191151	Bank
uildwood, Morningside, West Hill	43.763573	-79.188711	G & G Electronics	43.765309	-79.191537	Electronics Store

Methodology

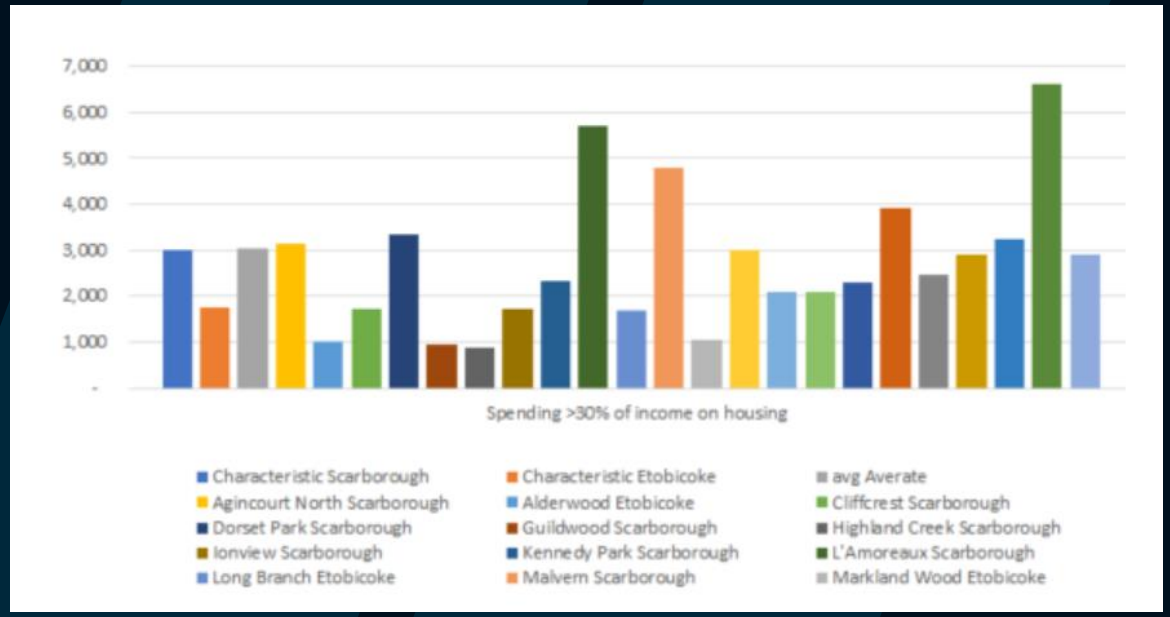
Massage the table and get the top 10 venues, and go through the k-mean clustering.

Postalcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353	2.0	Fast Food Restaurant	Vietnamese Restaurant	Clothing Store	Gym	Grocery Store
M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	43.784535	-79.160497	4.0	Bar	Construction & Landscaping	Vietnamese Restaurant	Clothing Store	Gym
M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711	1.0	Rental Car Location	Electronics Store	Medical Center	Intersection	Bank
M1G	Scarborough	Woburn	43.770992	-79.216917	1.0	Coffee Shop	Korean Restaurant	Mexican Restaurant	Vietnamese Restaurant	Hakka Restaurant
M1H	Scarborough	Cedarbrae	43.773136	-79.239476	1.0	Hakka Restaurant	Thai Restaurant	Athletics & Sports	Bakery	Bank

Plot the final result in map for comparision.
Details are in another notebook
: https://github.com/kayagi/Applied-Data-Science-Capstone/blob/main/CapstoneProject%20Data.ipynb

Methodology

Export the demographic statistic figures will be used for the comparison and scoring for each location. The analysis is based on the profile of different neighborhoods, including the Population, Income, and age distribution. The data source is from : https://open.toronto.ca/

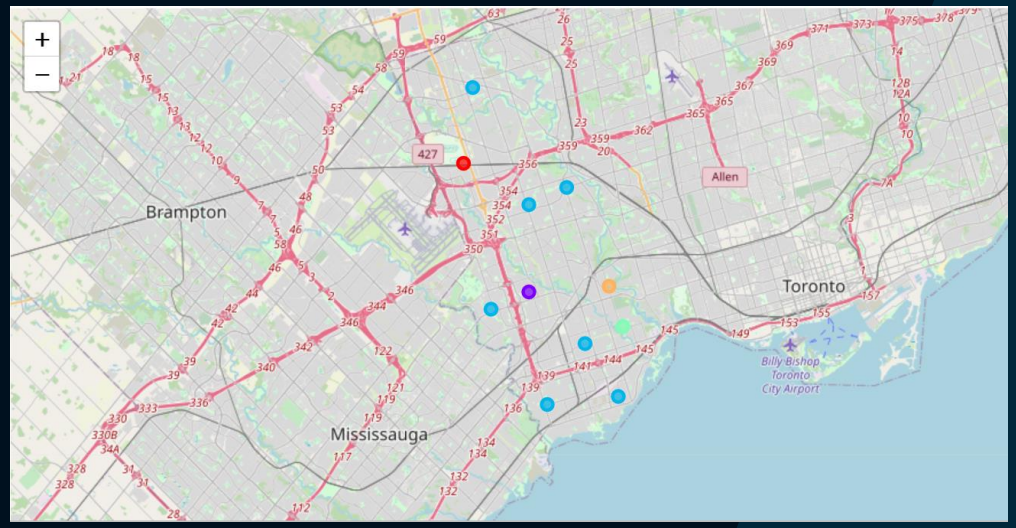


Results

Comparison of the Segmenting and Clustering Neighborhoods

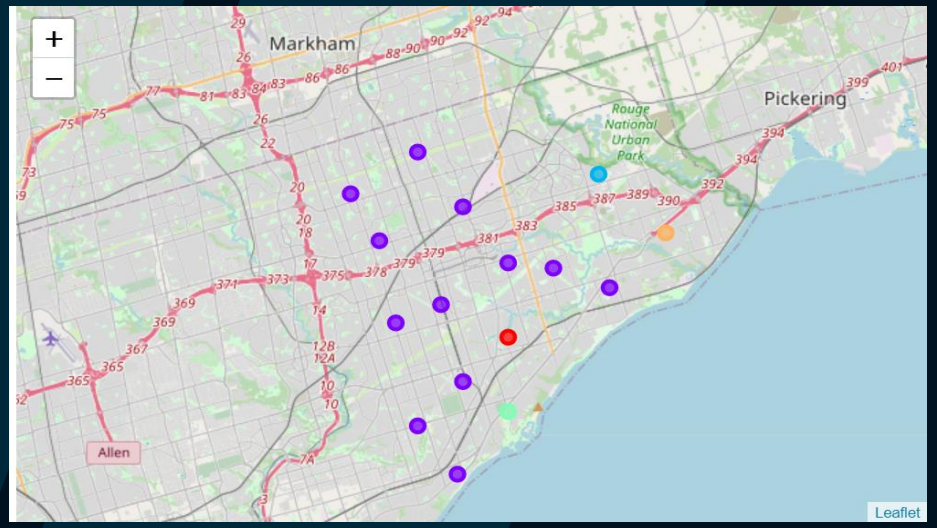
Etobicoke

- there are 41 distinct value of Venue Category and the cluster map is as follows :



Scarborough

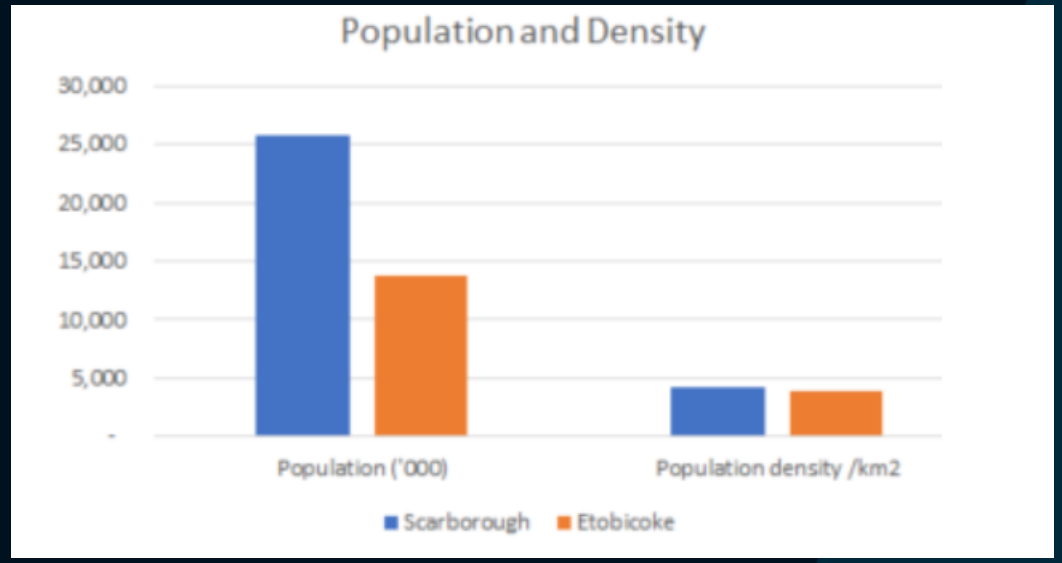
- there are 55 distinct value of Venue Category and the cluster map is as follows :



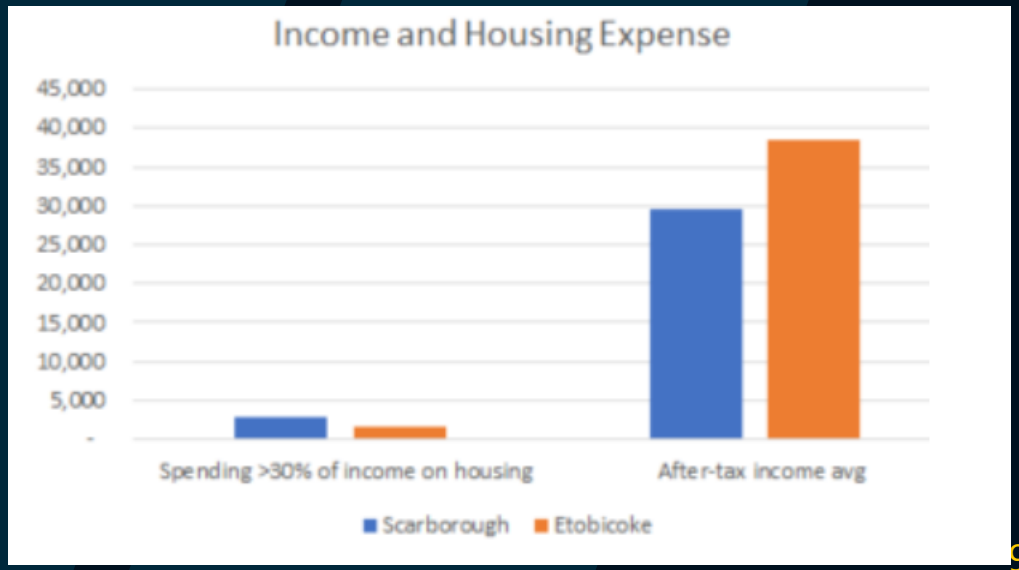
Results

Comparison in Etobicoke and Scarborough

Population and density

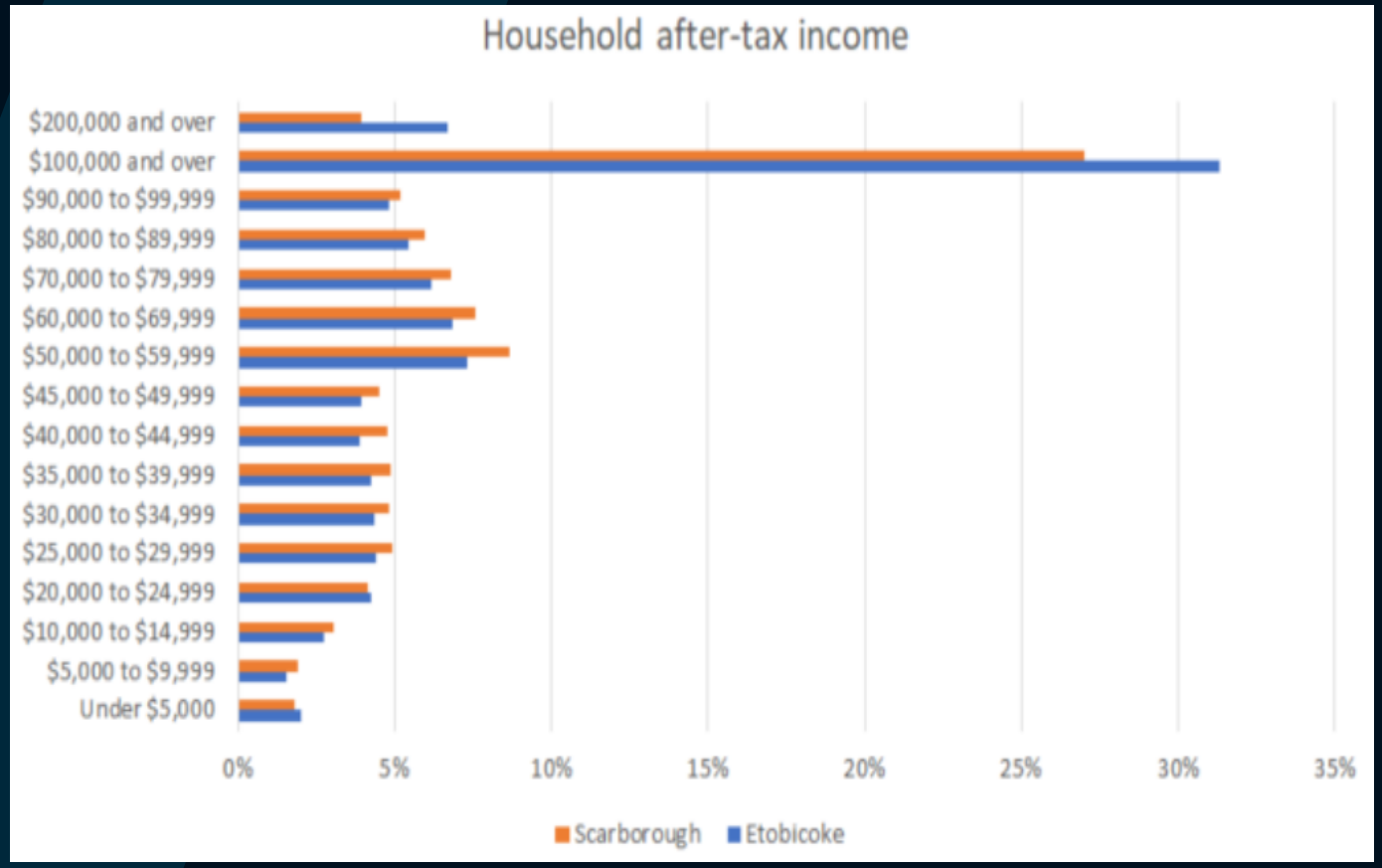


Income & Housing Expense



Results

Comparison of the household income by range in Etobicoke and Scarborough



Discussions

To make a more precise analysis.
1. For opening a foreign cuisine restaurant, we can also count the diversity of Immigrants in the location.
2. If the target customer group is family, we can further look into the figures on household with or without children and the age range in the location.

Conclusion

Using the data to analyze the segmenting and clustering Neighborhoods in Etobicoke and Scarborough, we can come up with a final decision on optimal location of the sushi restaurant location is recommended on Scarborough.

ncluding the nearby venue, population, population density, income and housing expense, and the household income range.

Additional factors like convenience of each location, environment condition, rental price and existing competitors are good to be included for decision making.

THANK YOU!