Deciding Where to Open Shop (in Toronto)

By: Adarsh Ravikumar

The Business Problem

- In a large city, starting a new business can be challenging, especially with so many other competitors to deal with
- Choosing the right location is essential
- Stakeholders and shop owners would be the target audience
- Successful business = Profit = Everyone is happy \(\exists\)
- Things to consider
 - Surrounding venues
 - Foot traffic
 - Vehicle traffic
 - Target demographics

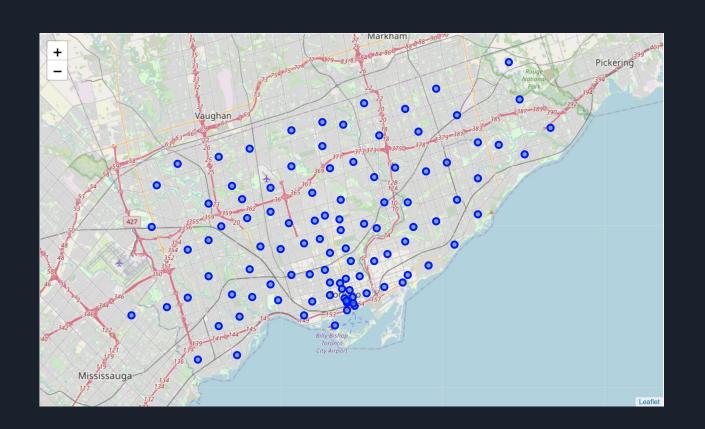
Datasets Being Used

- Dataset 1: Pedestrian and Vehicle Traffic Volumes
- Dataset 2: Neighbourhood Profiles
 - https://open.toronto.ca/dataset/neighbourhoods/
- Dataset 3: Postal Codes with Neighborhoods and Boroughs
 - https://en.wikipedia.org/wiki/List of postal codes of Canada: M
- Dataset 4: CSV with the Latitude and Longitude of Postal Codes
 - http://cocl.us/Geospatial data

Data Wrangling

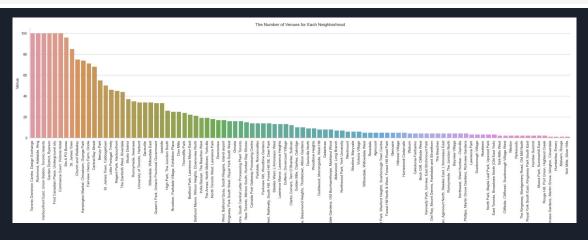
- Postal code information
 - Scraped Wikipedia page and read into data frame
 - Removing rows with unassigned boroughs
 - Grouping neighborhoods based on postal codes
- Traffic information data set
 - Selecting relevant rows
 - 8 Peak Hour Vehicle and Pedestrian Traffic information
- Neighborhood Profiles
 - Grouped multiple ethnicities into target demographic
 - Largest % target demographic neighborhoods were of interest

Mapping out the Neighborhoods with Foursquare



Getting Venue Information by Neighborhood

print(toronto venues.shape) toronto venues.head() (2167, 7)Neighborhood Neighborhood Neighborhood Venue Category Longitude Latitude Latitude Longitude Malvern, Rouge 43.806686 -79.194353 Wendy's 43.807448 -79.199056 Fast Food Restaurant Rouge Hill, Port Union, 43.784535 -79.160497 Royal Canadian Legion 43.782533 -79.163085 Bar Highland Creek Rouge Hill, Port Union. SEBS Engineering Inc. (Sustainable Construction & 43.784535 -79.160497 43.782371 -79.156820 Highland Creek Energy and ... Landscaping Guildwood, Morningside, West 43.763573 -79.188711 **RBC Royal Bank** 43.766790 -79.191151 Bank Guildwood, Morningside, West 43.763573 -79.188711 43.765309 -79.191537 G & G Electronics **Electronics Store**



Most Common Venue Category for Neighborhood

	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	Agincourt	Lounge	Breakfast Spot	Latin American Restaurant	Skating Rink	Clothing Store	Ethiopian Restaurant	Event Space	Escape Room	Electronics Store	Discount Store
1	Alderwood, Long Branch	Pizza Place	Sandwich Place	Coffee Shop	Pub	Pharmacy	Gym	Greek Restaurant	Discount Store	Department Store	Dessert Shop
2	Bathurst Manor, Wilson Heights, Downsview North	Coffee Shop	Bank	Pharmacy	Deli / Bodega	Shopping Mall	Bridal Shop	Sandwich Place	Diner	Restaurant	Middle Eastern Restaurant
3	Bayview Village	Japanese Restaurant	Café	Bank	Chinese Restaurant	Distribution Center	Dive Bar	Dog Run	Doner Restaurant	Donut Shop	Yoga Studio
4	Bedford Park, Lawrence Manor East	Sandwich Place	Italian Restaurant	Coffee Shop	Greek Restaurant	Thai Restaurant	Liquor Store	Comfort Food Restaurant	Juice Bar	Butcher	Café
5	Berczy Park	Coffee Shop	Restaurant	Bakery	Cocktail Bar	Beer Bar	Farmers Market	Cheese Shop	Seafood Restaurant	Sandwich Place	Beach

----Del Ray, Mount Dennis, Keelsdale and Silverthorn---frea 0 Turkish Restaurant Restaurant 0.25 Sandwich Place 0.25 Discount Store 0.25 3 Mediterranean Restaurant 0.00 ----Don Mills---freq venue 0.12 Beer Store 0.08 Coffee Shop Japanese Restaurant Asian Restaurant

	Neighborhood	Accessories Store	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Antique Shop	Aquarium
0	Malvern, Rouge	0	0	0	0	0	0	0	0	0	0	0
1	Rouge Hill, Port Union, Highland Creek	0	0	0	0	0	0	0	0	0	0	0
2	Rouge Hill, Port Union, Highland Creek	0	0	0	0	0	0	0	0	0	0	0
3	Guildwood, Morningside, West Hill	0	0	0	0	0	0	0	0	0	0	0
4	Guildwood, Morningside, West Hill	0	0	0	0	0	0	0	0	0	0	0

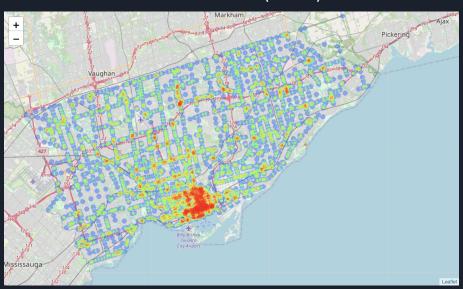
Target - "Middle Terranean"

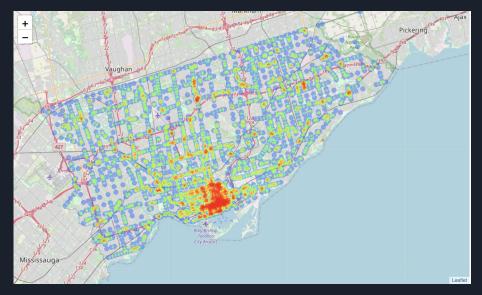
- For target population, ethnicities were combined into one larger target demographic group
- To see venue frequency, the following were combined into one
 - Doner Restaurant,
 Mediterranean Restaurant,
 Middle Eastern Restaurant,
 Falafel Restaurant, Greek
 Restaurant, Turkish Restaurant
- Target demographic population was divided by total population for each neighborhood to provide insight into the % density (on right)



8 Peak Hour Pedestrian and Vehicle Traffic

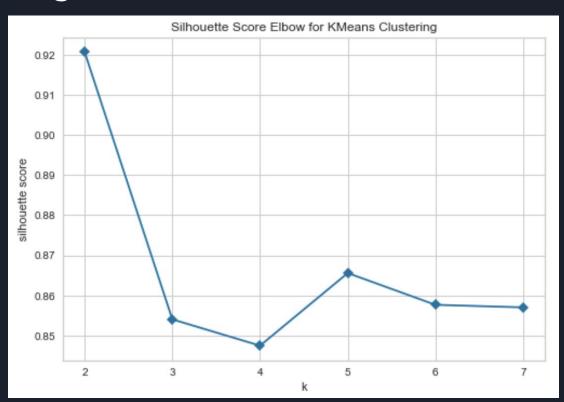
For Vehicles (Below)





K-means Clustering

- KElbowVisualizer helps figure out optimal cluster #
- Silhouette score is how similar values are to the clusters themselves
- Higher silhouette score = better
- Here, highest score at k = 2
- K-means clustering to be done with 2 clusters



Mapping the Clusters

- Cluster_0
 - Red
 - Larger
- Cluster_1
 - o Purple
 - Smaller



Results

Cluster_1 has a solid amount of 'Middle Terranean' venues frequenting their most common venues

	Cluster Labels	Postal Code	Latitude	Longitude	Borough	Neighbourhood	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
11	1.0	M1R	43.750072	-79.295849	Scarborough	Wexford, Maryvale	Accessories Store	Sandwich Place	Bakery	Middle Eastern Restaurant	Auto Garage	Smoke Shop	Electronics Store
17	1.0	М2Н	43.803762	-79.363452	North York	Hillcrest Village	Mediterranean Restaurant	Athletics & Sports	Pool	Dog Run	Golf Course	Drugstore	Distribution Center
41	1.0	M4K	43.679557	-79.352188	East Toronto	The Danforth West, Riverdale	Greek Restaurant	Coffee Shop	Italian Restaurant	Bookstore	Restaurant	Cosmetics Shop	Ice Cream Shop
80	1.0	М6М	43.691116	-79.476013	York	Del Ray, Mount Dennis, Keelsdale and Silverthorn	Restaurant	Discount Store	Sandwich Place	Turkish Restaurant	Yoga Studio	Doner Restaurant	Diner
86	1.0	M7R	43.636966	-79.615819	Mississauga	Canada Post Gateway Processing Centre	Intersection	Coffee Shop	Hotel	Fried Chicken Joint	Sandwich Place	Gym	Mediterranean Restaurant

Comparing Target Population with Target Venues

- When comparing top values for highest "Middle Terranean" frequency (right) with top values for highest % demographic population (left) few are in common
- Wexford/Maryvale, Thorncliffe
 Park, York
- Suggests a possible slight correlation although probably very weak

Newtonbrook East

Old East York

Willowdale East

Broadview North

Flemingdon Park

Bayview Village

Victoria Village

Henry Farm

Thorncliffe Park

Wexford/Maryvale

Del Ray, Mount Dennis, Keelsdale and Silverthorn

Hillcrest Village

Wexford, Marvvale

The Danforth West, Riverdale

Canada Post Gateway Processing Centre

> Dufferin, Dovercourt Village

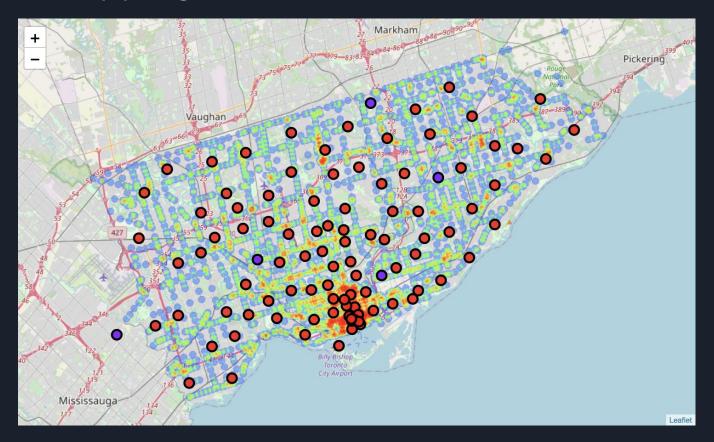
The Annex, North Midtown, Yorkville

Bathurst Manor, Wilson Heights, Downsview North

Bedford Park, Lawrence Manor East

> Thorncliffe Park

Overlapping Pedestrian Traffic with Clusters



Conclusion

- In conclusion, a k-means clustering model was used to cluster neighborhoods to help figure out where to open Mediterranean Restaurant
- Based on the results, Cluster_0 would be a better location to pursue
 - More neighborhoods around places with higher foot/pedestrian traffic
 - Cluster_1 has a relatively large amount of existing 'Middle Terranean' venues
 - Pre-existing competitors could make it hard to become popular
 - Since possible correlation is very weak, more neighborhoods could match both the target population % and target venue category which could bring more fruitful business results

Future Improvements

- Many things can be added for future models and decisions
- Other important factors to be accounted for
 - Accessibility
 - Parking
 - Competitor reviews
 - Customer dwelling/queue
 - Retail support relatively nearby
- Data about consumer demographics
- Specific store location information
 - Strip mall?
 - Shopping mall?
- Surrounding neighborhood socioeconomic information