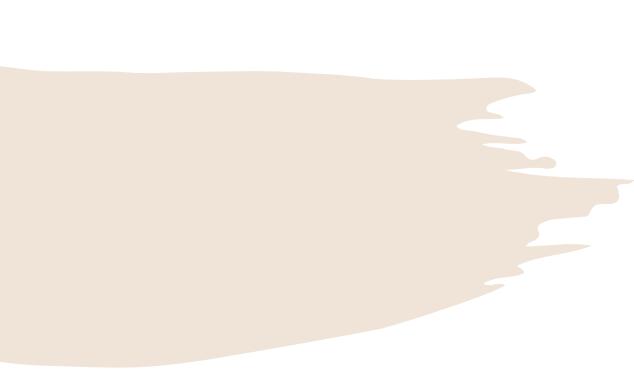




CLUSTERING RESULTS

VANCOUVER AND TORONTO
NEIGHBORHOOD COMPARISONS



I. Introduction & Business Problem

When an individual lives in a diverse and physically large country like Canada, one is bound to move from one city to another. Either at the personal level as one searches/move for jobs or at the business level where a restaurant might be interested in opening a branch in another city. These individuals or entities would be interested in neighborhoods that are similar to each other between the two cities.

Business Problem: How can we find a similar neighborhood in another city from our own?

Target Audience: Anyone interested in moving between cities including businesses interested in a strategy to identify new location across countries.



II. Data

To address this business problem and create a proof of concept solution, we will be interested in Toronto Neighborhoods and their similar neighborhoods in Vancouver. Geographically, they are quite far apart where Toronto can be considered to be on the East Coast of continental North America and Vancouver is on the West Coast.

List of Neighborhoods: Similar to the Toronto Lab, we will be obtaining a complete list of neighborhoods for the two cities using first [Wikipedia](#).

Geo Location: Using the geopy library, we would be able to join the list of neighborhoods to their respective coordinates.

Foursquare API: Then using the Foursquare API will we be able to obtain the characteristics and data on each neighborhood. Then the data will be joined with the list of neighborhood. This final dataset will be used for K-means clustering of neighborhoods to find clusters--groups of similar neighborhoods.



III. Methodology

In order to deliver on the business goal at hand of finding similar neighborhoods in a pool of Vancouver and Toronto Neighborhoods, the following was completed:

1. Getting a list of Neighborhoods from **Wikipedia**. A combination of **BeautifulSoup** and manual implementation was used to retrieve a list of Postal codes of the respective cities. Then using the **Google Maps API**, the postal code and province of where it came from was used to search for Geocoordinates.
2. After getting the coordinates for each postal code, the **Foursquare API** was used to retrieve a list of venues most popular in the neighborhood. Then, characteristics dummy variables were generated for each.
3. **K-Means Clustering** ($k=10$) was used to identify similar neighborhoods.



IV. Results

In general, the results showed well clustered groups with a few exceptions where groups are only one member with no other similar neighborhoods in it. This was unfortunate. Some good clusters deserving of attention below includes:

- **Cluster 2:** Includes a lot of Falafel Restaurants that were part of the most common venues.
- **Cluster 3:** Dominated by Zoos and parks in the top venues.
- **Cluster 4:** Construction & Landscaping venues with Ethiopian Restaurants
- **Cluster 5:** All of the neighborhood's top 10 venues matched exactly. They seem geographically close to each other too in Vancouver.
- **Cluster 10:** Trails and Parks

CLUSTER 2

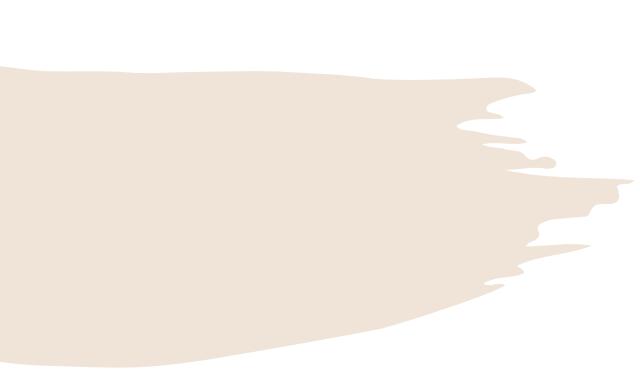
| PostalCode | Province | search | Latitude | Longitude | 1st Most Cor | 2nd Most Co | 3rd Most Cor | 4th Most Cor | 5th Most Cor | 6th Most Cor |
|------------|----------|---------|------------|------------|---------------|---------------|---------------|---------------|---------------|---------------|
| M5A | ON | M5A, ON | 43.6542599 | -79.360636 | Coffee Shop | Bakery | Park | Pub | Cafv© | Breakfast Sp |
| M1B | ON | M1B, ON | 43.8066863 | -79.194353 | Print Shop | Fast Food Re | Zoo | Falafel Resta | Electronics S | Ethiopian Re |
| M3B | ON | M3B, ON | 43.7459058 | -79.352188 | Japanese Re | Cafv© | Beer Store | Restaurant | Coffee Shop | Gym |
| M5B | ON | M5B, ON | 43.6560811 | -79.380171 | Clothing Stor | Coffee Shop | Italian Resta | Bubble Tea S | Restaurant | Cafv© |
| M6B | ON | M6B, ON | 43.709577 | -79.445073 | Pub | Japanese Re | Sushi Restau | Park | Zoo | Fair |
| M3C | ON | M3C, ON | 43.7258997 | -79.340923 | Japanese Re | Cafv© | Beer Store | Restaurant | Coffee Shop | Gym |
| M4C | ON | M4C, ON | 43.6953439 | -79.318389 | Cosmetics St | Spa | Curling Ice | Athletics & S | Beer Store | Pharmacy |
| M6C | ON | M6C, ON | 43.6937813 | -79.428191 | Hockey Arena | Tennis Court | Field | Trail | Fair | Electronics S |
| M9C | ON | M9C, ON | 43.6435152 | -79.577201 | Cosmetics St | Beer Store | Liquor Store | Park | Coffee Shop | Pet Store |
| M1E | ON | M1E, ON | 43.7635726 | -79.188712 | Mexican Res | Rental Car Lo | Bank | Moving Targ | Intersection | Electronics S |
| M4E | ON | M4E, ON | 43.6763574 | -79.293031 | Pub | Health Food | Trail | Zoo | Fair | Eastern Euro |
| M5E | ON | M5E, ON | 43.6447708 | -79.373306 | Coffee Shop | Cocktail Bar | Cafv© | Cheese Shop | Restaurant | Seafood Res |
| M1G | ON | M1G, ON | 43.7709921 | -79.216917 | Coffee Shop | Korean Resta | Indian Resta | Electronics S | Ethiopian Re | Event Space |
| M4G | ON | M4G, ON | 43.7090604 | -79.363452 | Sporting Goc | Coffee Shop | Burger Joint | Bank | Furniture / H | Pet Store |
| M5G | ON | M5G, ON | 43.6579524 | -79.387383 | Coffee Shop | Italian Resta | Sandwich Pla | Cafv© | Salad Place | Bubble Tea S |
| M1H | ON | M1H, ON | 43.773136 | -79.239476 | Caribbean Re | Fried Chicker | Bank | Athletics & S | Gas Station | Bakery |
| M3H | ON | M3H, ON | 43.7543283 | -79.442259 | Coffee Shop | Bank | Fried Chicker | Gas Station | Restaurant | Supermarket |
| M4H | ON | M4H, ON | 43.7053689 | -79.349372 | Sandwich Pla | Indian Resta | Pizza Place | Park | Burger Joint | Bus Line |
| M5H | ON | M5H, ON | 43.6505712 | -79.384568 | Coffee Shop | Cafv© | Restaurant | Deli / Bodeg | Gym | Clothing Stor |
| M6H | ON | M6H, ON | 43.6690051 | -79.442259 | Pharmacy | Bakery | Portuguese F | Art Gallery | Middle Easte | Supermarket |
| M2J | ON | M2J, ON | 43.7785175 | -79.346556 | Clothing Stor | Coffee Shop | Fast Food Re | Restaurant | Japanese Re | Bakery |
| M3J | ON | M3J, ON | 43.7679803 | -79.487262 | Caribbean Re | Coffee Shop | Miscellaneou | Massage Stu | Falafel Resta | Bar |
| M5J | ON | M5J, ON | 43.6408157 | -79.381752 | Coffee Shop | Aquarium | Cafv© | Hotel | Restaurant | Scenic Looko |
| M6J | ON | M6J, ON | 43.6479267 | -79.41975 | Bar | Asian Restau | Restaurant | Cafv© | Men's Store | Vegetarian / |
| M1K | ON | M1K, ON | 51.253775 | -85.323214 | Pizza Place | Shoe Store | Zoo | Fair | Eastern Euro | Electronics S |
| M4K | ON | M4K, ON | 43.6795571 | -79.352188 | Greek Resta | Coffee Shop | Italian Resta | Ice Cream St | Furniture / H | Restaurant |

CLUSTER 3

CLUSTER 4

CLUSTER 5

CLUSTER 10



V. Discussion and Conclusion

The goal of this project has value. Nevertheless, the results of the investigation showed that more refinement can be made as the choice for clusters. Known clusters could have been implemented to construct training and testing sets for the K-means clustering algorithm.

Having said that there were interesting clusters of Ethiopian restaurants and Trails were found to have commonality. At the same time, another clustering algorithm could have been used alongside as a point of comparison (i.e. hierarchical clustering).