Mapping Toronto in the Context of Vancouver's Neighborhoods

By Kristian Correia

Introduction

Being two of the largest cities in Canada, Toronto and Vancouver serve as some of the largest hubs in the nation for business and tourism in the Country. It is not uncommon for residents of one of these cities to either travel to the other for vacation or to move to the other city. Given the large distance between these metropolises, it would be hard to scope out neighborhoods and getting a strong understanding of the other city before committing to spending a prolonged period of time in the other city.

In order to understand another city, it is beneficial to map it out in reference to a city that you understand much better. I am personally a resident of Vancouver and I would like to understand Toronto in greater depth. It would make sense to me to relate the neighborhoods in Vancouver that I am familiar with to the neighborhoods of Toronto to improve my understanding of the city's layout and which neighborhoods interest me.

I am intending with this project to leverage foursquare data and the lessons taught in the previous modules, to map out Toronto's neighborhoods in the context of how similar they are to the neighborhoods in Vancouver

I intend to create clusters of the different neighborhoods in Vancouver to build several general types of neighborhoods in the city. I will then feed the neighborhoods of Toronto through the same model to determine which cluster each borough aligns with.

Using this data, I will be able to determine what parts of Vancouver and Toronto are similar to one another and what areas of the Toronto will interest me to investigate further.

Data

I will need a substantial amount of data for each of the two cities in order to complete this project.

For each city I will need a list of the different neighborhoods, as we did in module three, I intend to use postal code FSAs in order to define my different neighborhoods. I will use web scrapping to devise tables of this data from the following sources:

Toronto: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Vancouver: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_V

I will also need to associate the different neighborhoods with their respective coordinates. I will use the geocoder package to obtain this data like I did in module 3.

Lastly I will need venue data to build profiles on the different neighborhoods in both Toronto and Vancouver. I intend on obtaining very granule data for each neighborhood and will accept the top 300 venues for each neighborhood using the Foursquare API. Specifically, I will be using venue categories which I will one hot encode, and sum up for each neighborhood and then normalize to account for

difference in total number of venues. This will allow for me to build a profile for each of the different boroughs. Which I can ultimately use to create clusters of the Vancouver neighborhoods and map Toronto to.

Methodology

In this project, we will be focusing on the 41 unique boroughs of Vancouver and determining which neighbourhoods are similar to one another. We will also be mapping Toronto in terms of its similarity to Vancouver and finding which neighborhoods are most similar. We will complete this objective through four steps:

We will retrieve the coordinate and venue data in all 41 postal code FSAs in Vancouver and 98 in Toronto. We will also be obtaining the venue category of the top 300 venues from each of get neighborhoods in both cities and build a normalized distribution of the venue types in each neighborhood. This first step was completed in the Data section of the report.

We will train a k-means cluster algorithm with the venue data from the Vancouver neighborhoods to cluster the neighborhoods into 8 categories. We will create a visual display of these clusters of neighborhoods on a map and we will create a profile of each cluster type noting what the average neighborhood in each cluster looks like.

We will use the k-means cluster to algorithm that we developed to categorize the neighborhoods in Toronto into the same clusters that we created for Vancouver. Similar to how we analyzed Vancouver, we will create a visual display of these clusters of neighborhoods on a map and we will create a profile of each cluster type noting what the average neighborhood in each cluster looks like.

We will lastly look to find which specific neighborhoods are closest to one another. We will do this by creating a distance matrix between all the neighborhoods in each Toronto against all the neighborhoods in Vancouver. We will use the Euclidean Distance of each pair neighborhood's compositions. We will then bring forward the 5 closest pairs of neighborhoods between each city. We will review if there are any insights comparing the details of each pair of neighborhoods.

Results

We begin by creating a k-means cluster algorithm with the venue data from the Vancouver Neighborhoods. The results can be seen below in figure 1. We see that 27 of the 40 neighborhoods are placed in the same cluster while the other 13 neighborhoods are divided among the other 7 clusters. Nevertheless, we largely see that neighborhoods in specific clusters also tend to be geographically close to one another. In figure 2 we can see a summary of top venue types in each of the different clusters of Vancouver neighborhoods. We can see different focuses in different clusters where some are more residential oriented, while others have thriving nightlife or a large focus on nature and outdoor spaces.
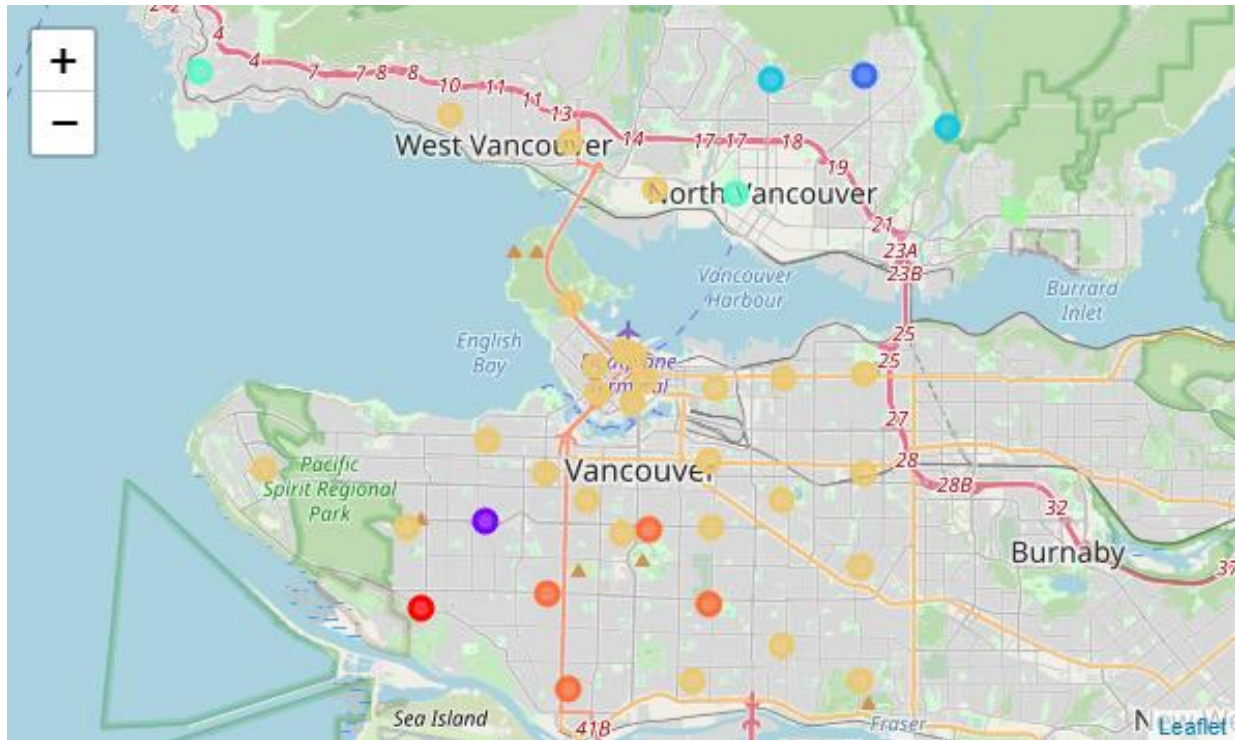
Figure 1 - Cluster map of Vancouver Neighborhoods

	Cluster Labels	Number of Neighborhoods	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	0	1	Grocery Store	Gym / Fitness Center	Japanese Restaurant	Event Space	Dry Cleaner	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	Elementary School	Ethiopian Restaurant
1	1	1	Caribbean Restaurant	Italian Restaurant	Bakery	Donut Shop	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	Elementary School	Ethiopian Restaurant	Event Space
2	2	1	Food & Drink Shop	Yoga Studio	Falafel Restaurant	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	Elementary School	Ethiopian Restaurant	Event Space	Fair
3	3	3	Trail	Pet Store	Park	Coffee Shop	Mountain	Bus Stop	Ski Chairlift	Falafel Restaurant	Farm	Fair
4	4	2	Playground	Boat or Ferry	Park	Falafel Restaurant	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	Elementary School	Ethiopian Restaurant	Event Space
5	5	1	Construction & Landscaping	Elementary School	Yoga Studio	Farm	Eastern European Restaurant	Electronics Store	Ethiopian Restaurant	Event Space	Fair	Falafel Restaurant
6	6	27	Coffee Shop	Park	Café	Bus Stop	Bank	Sushi Restaurant	Hotel	Bakery	Indian Restaurant	Sandwich Place
7	7	4	Chinese Restaurant	Coffee Shop	Dessert Shop	Bus Stop	Sushi Restaurant	Asian Restaurant	Pizza Place	Field	Bubble Tea Shop	Sandwich Place

Figure 2 - Summary of clusters in Vancouver Neighborhoods

We then map the neighborhoods of Toronto through the same k-means cluster model that we fit using Vancouver neighborhoods data. This allows us to see what set of Neighborhoods in Vancouver, each neighborhoods in Toronto is similar to. We have displayed these results in a cluster map in Figure 3. We immediately see that most neighborhoods align with cluster 6 which is dense with coffees shops, parks banks and public transit stops. We also notice that the odd neighborhood in Toronto aligns with a

minority cluster indicating that these neighborhoods in Toronto are similar to very niche section in Vancouver.
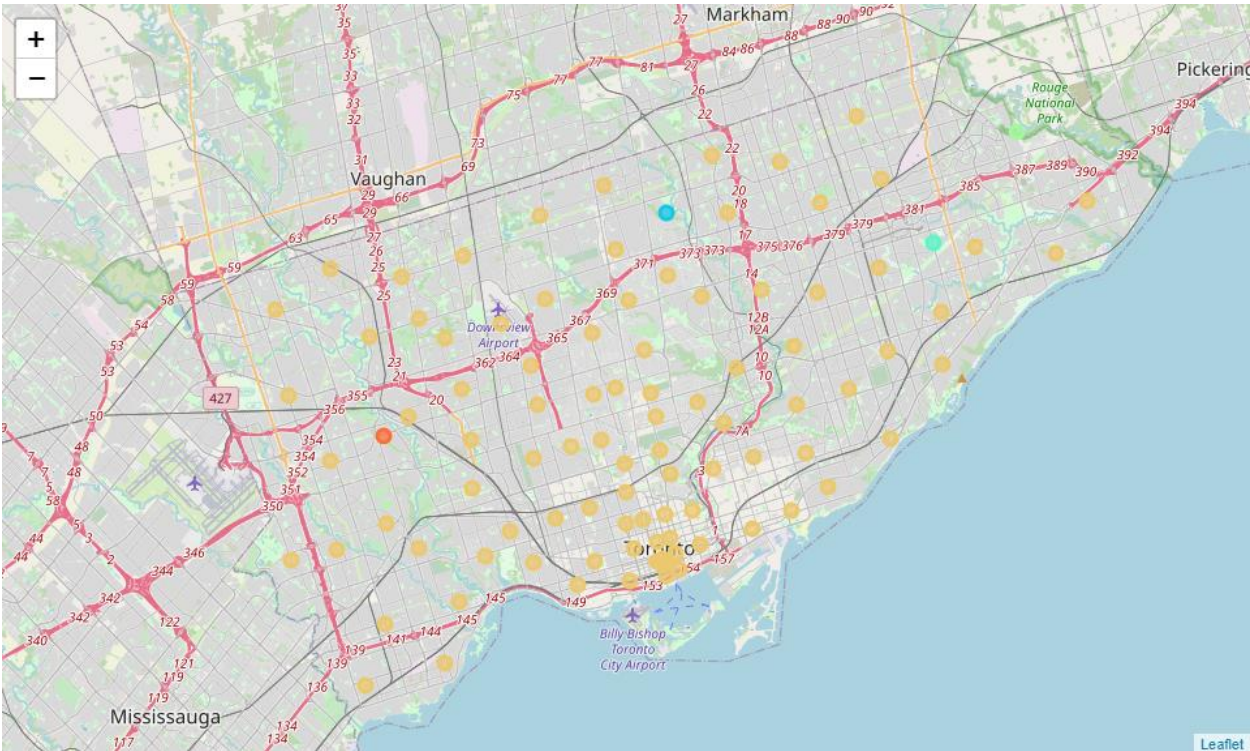

Figure 3 - Cluster Map Showing the Neighborhoods of Toronto as they align with the Neighborhoods of Vancouver

Lastly we create a distance matrix which takes the Euclidean distance of every pair of neighborhoods between the two cities. The idea is that the smaller the distance, the closer the composition of the two neighborhoods is, the more similar they are. In Figure 4, we review the 5 closest pairings of neighborhoods. We look into the details of these pairs neighborhoods in the project's notebook and we discuss these in further depth in the Discussion section of this report

	Vancouver Neighborhood	Toronto Neighborhood	Distance
1571	North Vancouver - Northwest Central	Bayview Village	0.666667
467	Killarney	Steeles West, L'Amoreaux West	0.800000
3220	Waterfront / Coal Harbour / Canada Place	Toronto Dominion Centre, Design Exchange	0.940000
3165	Waterfront / Coal Harbour / Canada Place	First Canadian Place, Underground city	0.980000
3201	Waterfront / Coal Harbour / Canada Place	Richmond, Adelaide, King	0.990000

Figure 4 - Closest Neighborhoods in the Distance Matrix between Vancouver and Toronto

Discussion

From this exercise, there are several notable observations that we can value moving forward. The first being that at its highest level, there is a common thread between most neighborhoods in both cities. Of the 40 neighborhoods in Vancouver, 27 fell into the same cluster and when categorizing the 98

neighborhoods in Toronto, 94 fell into this same cluster. We can conclude that many neighborhoods in both cities have a similar composition of venues where they tend to largely contain coffee shops, cafes, green spaces, public transit routes and a variety of restaurants.

While we did see a number of similar neighborhoods that did not stand out from one another, we are able to use this data to portion off a few unique neighborhoods in Toronto that align with unique areas of Vancouver. We can see 4 neighborhoods in Toronto that were placed in minority clusters from the Vancouver data. We will focus on 3 of these from our categorization analysis and one will be reviewed in the distance matrix review that we completed.

The neighborhood Westmount in Toronto was placed in cluster 7. We can conclude that this neighborhood is distinct from the others in Toronto and has a larger distribution of East Asian ethnicity based restaurant. This indicates that similar to the 4 neighborhoods in Vancouver that are in this cluster, these neighborhoods have a large cultural influence and are distinct from the others in this sense.

The Malvern, Rouge Neighborhood in Toronto was placed in cluster 5 which solely contained the North Vancouver – Inner East Neighborhood. We can draw to the similarity that both neighborhoods have a distinct focus on green space and active living along with a broad influence of cuisine from different cultures.

The Cedarbrae neighborhood in Toronto was placed in cluster 4. Similar the case of Malvern Rouge, it has similar to the neighborhoods in its clusters on the grounds of large focus on green space and active living along with a broad influence of cuisine from different cultures. We do see that the types of ethnic influence on the restaurants in these separate clusters are different from one another.

In addition to the categorization model that we completed we also created a distance matrix to individually measure how different each pair of neighborhoods between the different cities was. After reviewing the 5 most similar pair of neighbourhoods we can draw a number of conclusions.

The North Vancouver - Northwest Central neighborhood in Vancouver and the Bayview Village neighborhood in Toronto were found to be the most similar to one another. They were also both categorized in the same niche cluster. We can see the that they have a number of commonalities, specifically the large portion of Yoga studios, Schools, Eastern European restaurants, Trails, Electronic Shops and Parks. This is an excellent case of both the cluster model and the distance matrix finding two communities that have much in common.

The next closest pairing was the Killarney neighborhood in Vancouver and the Steeles West, L'Amoreaux West in Toronto. They both have a commonly high proportion of fast food restaurants, Chinese restaurants, and coffee shops.

The next three closest pairings were the Waterfront / Coal Harbour / Canada Place neighborhood in Vancouver to the Toronto Dominion Centre, Design Exchange neighborhood, the First Canadian Place, Underground city neighborhood and the Richmond, Adelaide, King neighborhood in Toronto. We can conclude that their common link between these four neighborhoods is that they serve as the central areas in their respective cities for commerce. We can see that in cases there is a large overlap in common venues including coffee shops, hotels and more North American style restaurants.

Conclusion

While this exercise did display a commonality that links many neighborhoods in both cities, it was also able to shine a light on certain niche neighborhoods that are distinct from others in their city and yet are similar to a neighborhood in the other city. There are several unique pockets of Toronto that we can expect to be similar in experience to the neighborhoods of Vancouver. These are neighborhoods have substantial range and are filled with nature, recreation, multicultural influence or serve as large hubs for the cities thriving business sector, that are useful. As a future extension of this project, I think it would be insightful to run a second iteration of this process, however, the clusters would be created from only the 27 neighborhoods in Vancouver that were assigned to cluster 6. This may allow for us to create an even more granular understanding of these neighborhoods.