Exploring the Cuisine of Toronto

Paraic O'Gorman

May 03, 2021

1. Introduction

1.1 Background

In this project, we are going to look at the distribution of restaurants across the city of Toronto. Specifically this project will look at which neighbourhoods you can find a high concentration of certain types of cuisine.

Toronto has been named the most diverse city in the world by BBC Radio. Toronto has long been recognized as one of the most multicultural cities in the world, with over 230 different nationalities within Ontario's capital city. It has long been the entry point for many immigrants into the country, and these immigrants have brought with them a wide variety of cuisines. Today, these different cuisines can be found in the restaurants and diners that have sprung up in immigrant areas of the city, Italian restaurants in little Italy, Chinese restaurants in Chinatown etc.

1.2 Problem

Food tourism is a growing industry in Toronto. Over the years a number of companies have begun running food tours in the city, catering to the needs of these tourists. The goal of this project is to give tourists an idea of where a large number of restaurants catering to a given cuisine. This will be of particular use to foreign tourists hoping to get a taste of everything New York has to offer. It can also be used by food tourism companies to market specific city neighbourhoods to these tourists.

At the end of the project, I will suggest the path that food tours can take to most efficiently taste all Toronto has to offer.

1.3 Interest

This project would be of interest to restaurateurs in the greater Toronto area. It would also be of interest to tourist companies in Toronto interested in offering food tours through Toronto, and tourists visiting Toronto interested in sampling the cities cuisine.

2. Data acquisition and cleaning

2.1 Data sources

I will use Foursquare data to profile the restaurants that can be found in Toronto. Foursquare is an American technology company focusing on location data. Foursquare has a large repository of data on restaurants, but it does not have a complete picture of the restaurants that can be found in Toronto. On initial inspection, there are approximately 270 types of venues listed in the Toronto

area. Not all of these will be included in the scope of this project. Foursquare data can be found at https://api.foursquare.com

Postal codes and neighbourhood names in Toronto will be scraped from Wikipedia: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Spatial data will be taken from http://cocl.us/Geospatial_data

2.2 Data cleaning

Data downloaded or scraped from Wikipedia, Foursquare and the online data source http://cocl.us/Geospatial_data. The data is combined into one table.

Venues are identified in the Toronto area using Foursquare. 269 distinct venue types were identified. Not all these venues are in the scope of this project, as they are not restaurants. I remove any venues that are not one of the following types:

- Portuguese Restaurant
- Italian Restaurant
- Ramen Restaurant
- Thai Restaurant
- Modern European Restaurant
- Middle Eastern Restaurant
- Ethiopian Restaurant
- Asian Restaurant
- Korean Restaurant
- Colombian Restaurant
- Tibetan Restaurant
- Vietnamese Restaurant
- Mexican Restaurant
- Japanese Restaurant
- German Restaurant
- Irish Pub
- Belgian Restaurant
- Cuban Restaurant
- Filipino Restaurant
- Taiwanese Restaurant
- Greek Restaurant

3. Methodology

3.1 Exploration of Neighbourhood data

Having joined postal code, neighbourhood name, latitude and longitude, I explored how these neighbourhoods looked on a map of Toronto using the Folium mapping package. I set the latitude

and longitude as the average of the latitude and longitude of each of its neighbourhoods, and then plotted each of the neighbourhoods on a map.

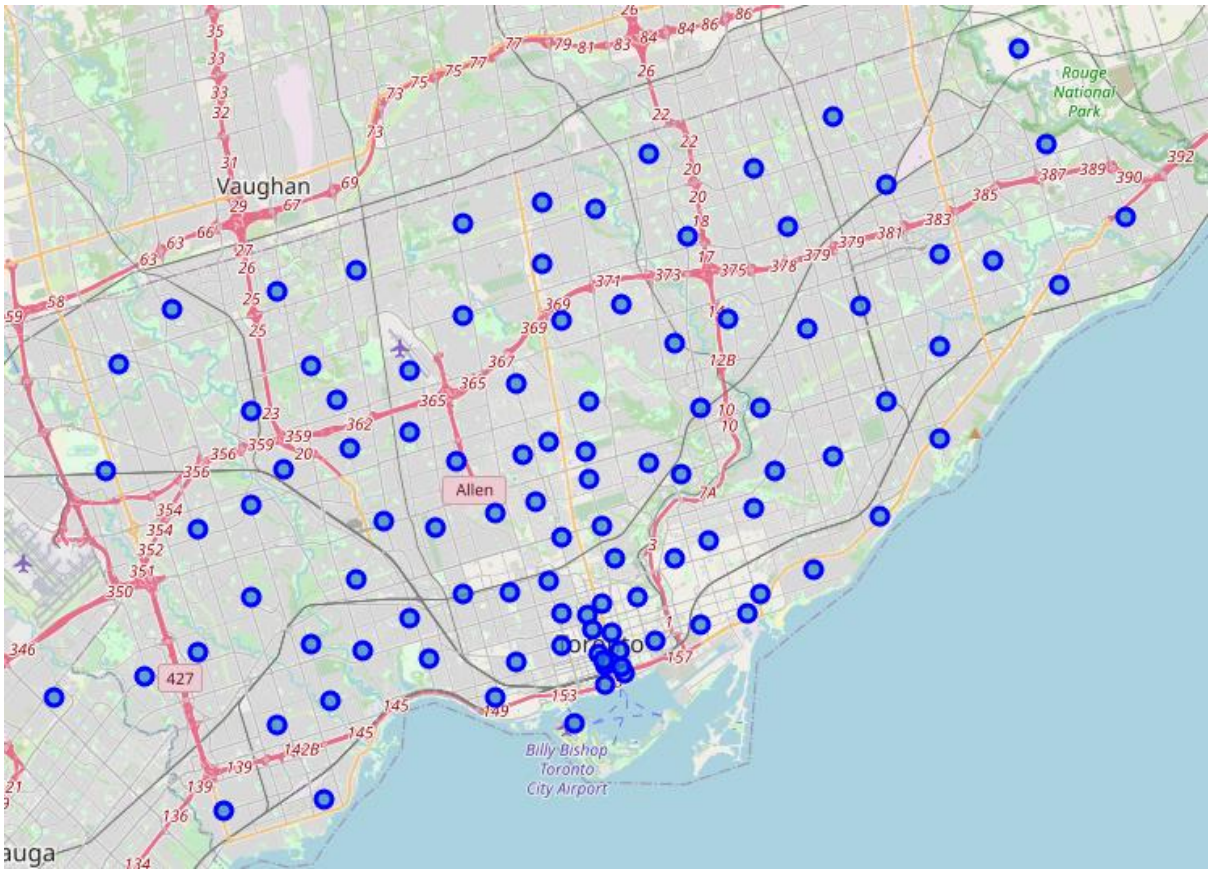

Figure 1. Plot of the 103 Toronto neighbourhoods

Having identified the 103 neighbourhoods in scope, I identified what venues are available within 500m of each neighbourhood. I then filtered this to get a view of what neighbourhoods contained a restaurant in the scope of this project. These neighbourhoods are mapped here.

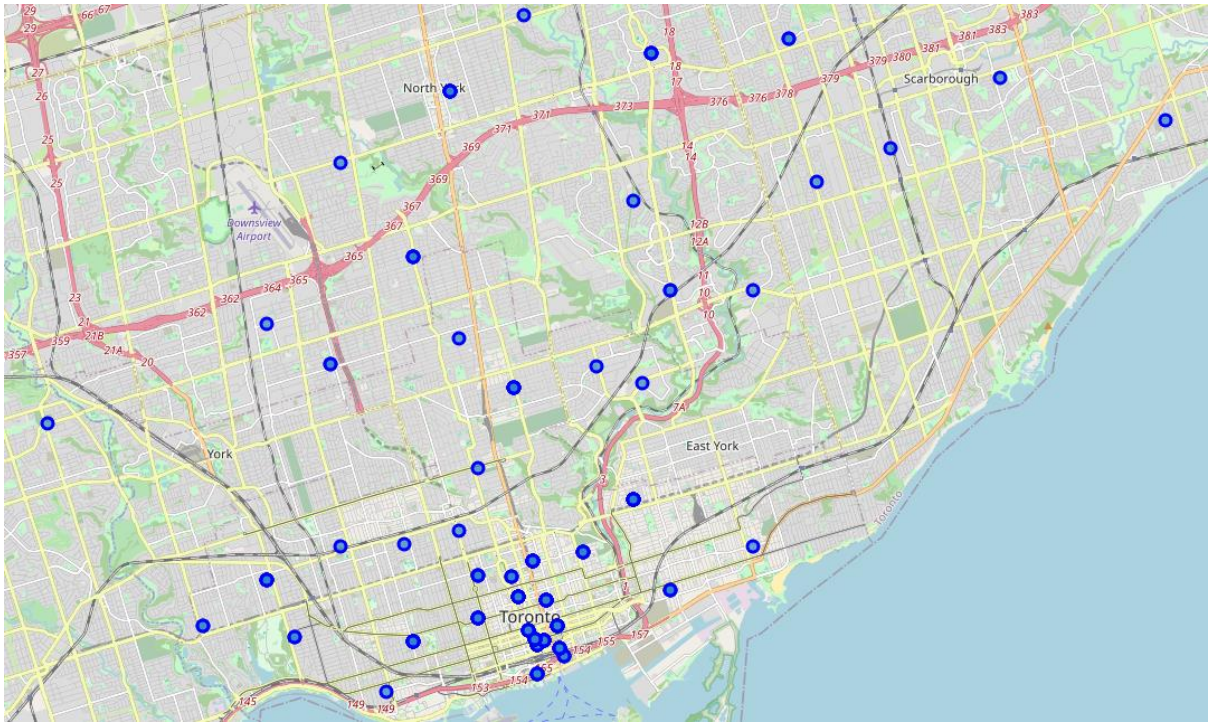

Figure 2. Plot of the 47 Toronto restaurant neighbourhoods

3.2 Clustering of Neighbourhood data

I reformatted the data so that each neighbourhood was analysed to identify the proportion in which each type of restaurant appears. For each neighbourhood I then identified the top 10 most common venues. This is the basis on which I chose to cluster the neighbourhoods.

To identify how best to cluster the neighbourhoods, I used the elbow and silhouette method. I wanted to keep the number of clusters small, as this is more realistic for a Toronto food tour, or for a short vacation to the city. In the figure below, you can find the output of the Elbow and Silhouette method.

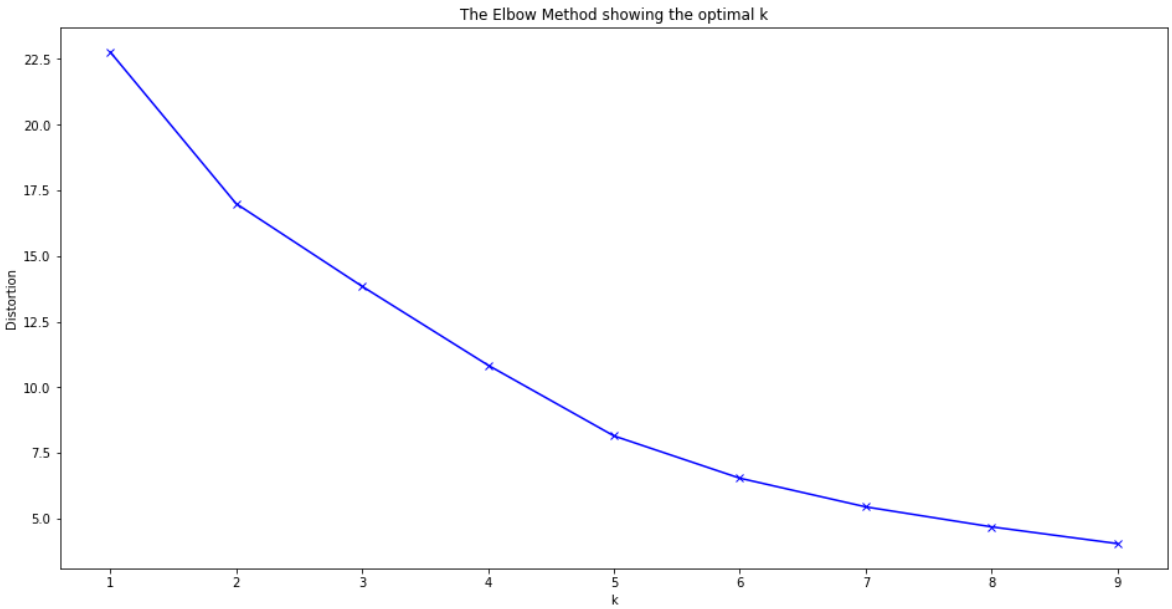

Figure 3. Output of the Elbow method

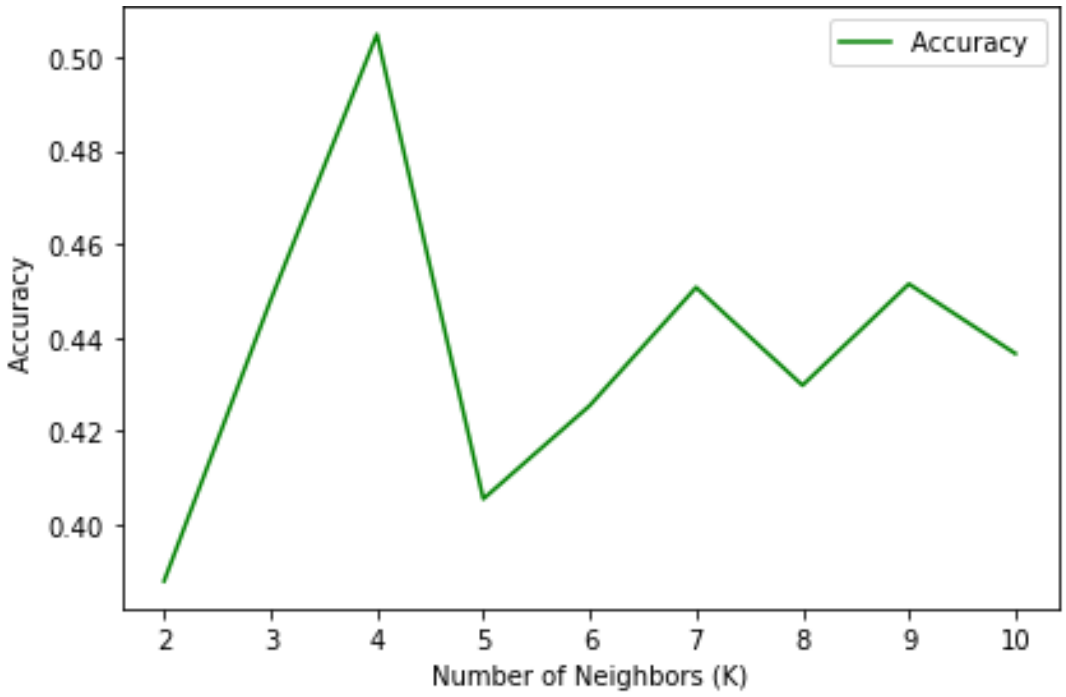

Figure 4. Output of the Silhouette method

On the basis of these results, I chose to cluster the neighbourhoods into 4 clusters using k-means clustering. I ran the k-means to cluster the neighbourhoods and plotted them on a map. Each cluster is assigned a different colour.

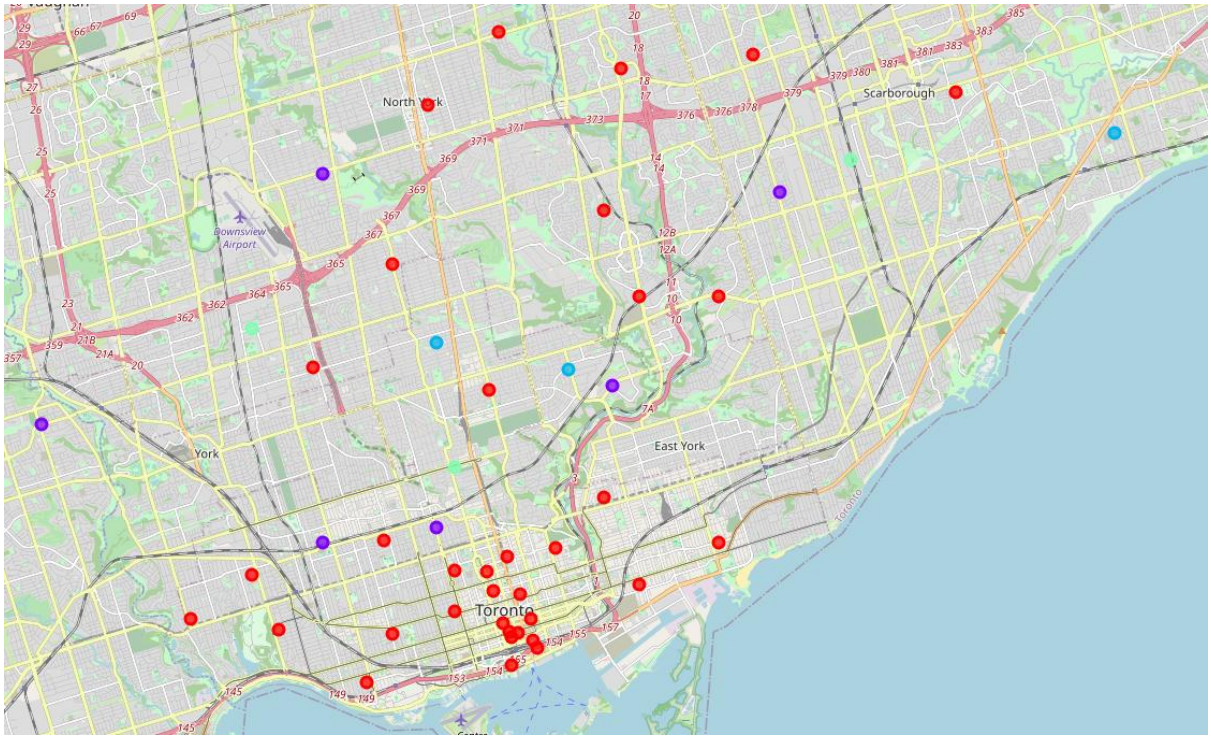

Figure 5. 4 clusters plotted over map of Toronto

I examined each cluster and determine the discriminating venue categories that distinguish each cluster. Based on the defining categories, I assigned a name to each cluster as below.

Cluster	Colour	Name
0	Red	Melting Pot Cluster
1	Purple	Middle Eastern Cluster
2	Blue	Mexican Cluster
3	Green	Asian Cluster

Having identified these 4 clusters, I set out defining an algorithm which would allow someone to visit a neighbourhood in each cluster in the most efficient manner. The distance between two neighbourhoods is defined using latitude and longitude. Iterating over the different clusters, I found the most efficient route is 'The Annex, North Midtown, Yorkville' to 'Summerhill West, Rathnelly, South Hill, Forest Hill SE, Deer Park' to 'Davisville' to 'North Toronto West'. These four locations are plotted here.

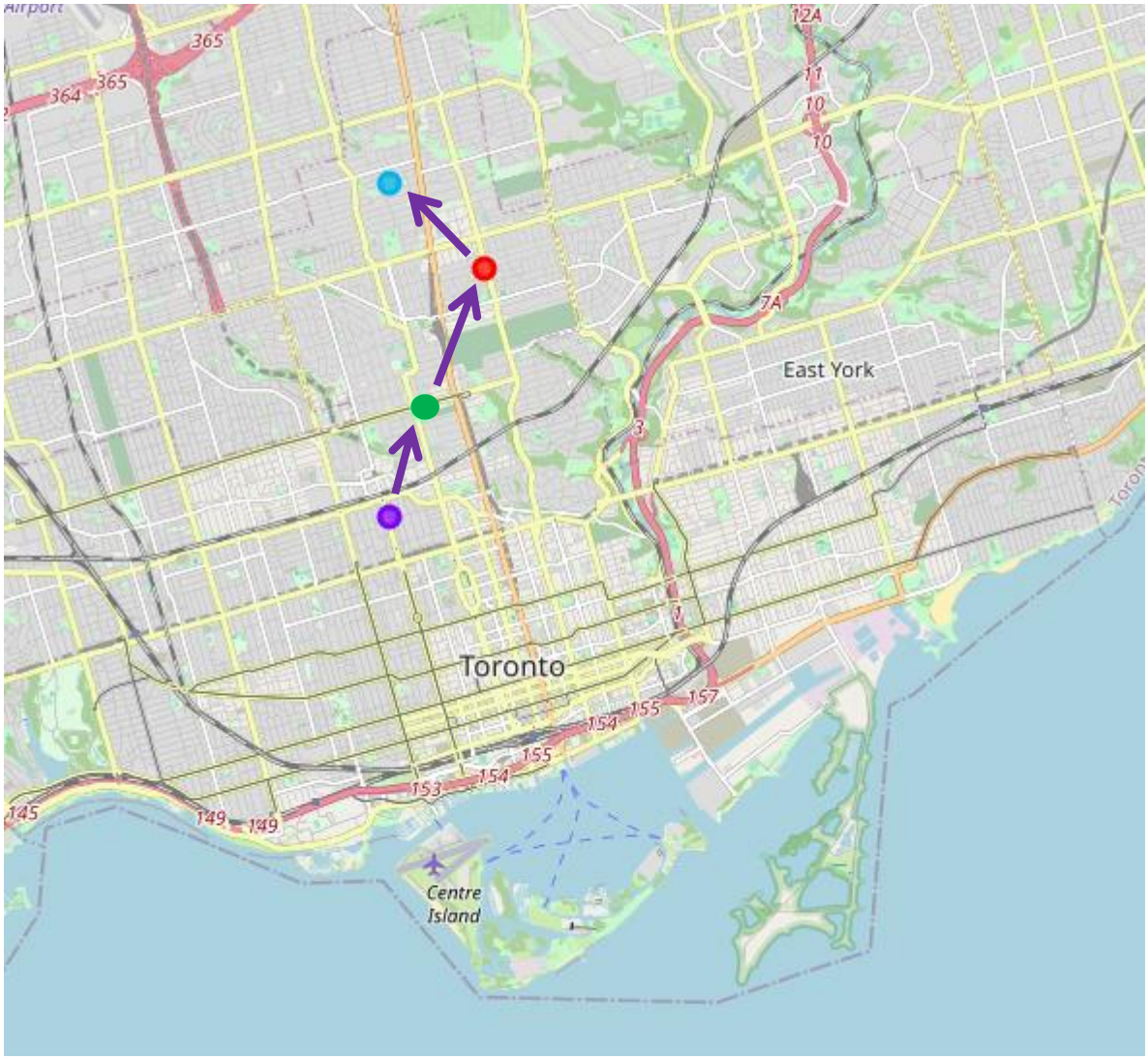

Figure 6. Most efficient path for visiting restaurants in all 4 clusters

4. Results

3.1 Analysis of Clusters

Having clustered the neighbourhoods into 4 clusters, I assigned a name to each of them based on their defining characteristics.

3.1.1 Cluster 0 - Melting Pot Cluster

PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
M4A	North York	Victoria Village	43.725882	-79.315572	0.0	Portuguese Restaurant	Asian Restaurant	Korean Restaurant
M7A	Queen's Park	Ontario Provincial Government	43.662301	-79.389494	0.0	Japanese Restaurant	Italian Restaurant	Mexican Restaurant
M3B	North York	Don Mills North	43.745906	-79.352188	0.0	Japanese Restaurant	Korean Restaurant	Tibetan Restaurant
M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	0.0	Japanese Restaurant	Italian Restaurant	Middle Eastern Restaurant
M6B	North York	Glencairn	43.709577	-79.445073	0.0	Asian Restaurant	Japanese Restaurant	Colombian Restaurant
M3C	North York	Don Mills South	43.725900	-79.340923	0.0	Asian Restaurant	Italian Restaurant	Korean Restaurant
M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	0.0	Italian Restaurant	Asian Restaurant	Irish Pub
M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	0.0	Japanese Restaurant	Thai Restaurant	Greek Restaurant

Figure 7. Sample of data from Cluster 0

The 'Melting Pot' Cluster is characterised by a broad variety of restaurant choices - Japanese, Italian, Asian, Thai etc. It is the most varied of the clusters.

3.1.2 Cluster 1 - Middle Eastern Cluster

PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
M3H	North York	Bathurst Manor, Wilson Heights, Downsview North	43.754328	-79.442259	1.0	Middle Eastern Restaurant	Asian Restaurant	Korean Restaurant
M4H	East York	Thorncliffe Park	43.705369	-79.349372	1.0	Middle Eastern Restaurant	Asian Restaurant	Korean Restaurant
M6H	West Toronto	Dufferin, Dovercourt Village	43.669005	-79.442259	1.0	Middle Eastern Restaurant	Asian Restaurant	Korean Restaurant
M9P	Etobicoke	Westmount	43.696319	-79.532242	1.0	Middle Eastern Restaurant	Asian Restaurant	Korean Restaurant
M1R	Scarborough	Wexford, Maryvale	43.750071	-79.295849	1.0	Middle Eastern Restaurant	Asian Restaurant	Korean Restaurant
M5R	Central Toronto	The Annex, North Midtown, Yorkville	43.672710	-79.405678	1.0	Middle Eastern Restaurant	Asian Restaurant	Korean Restaurant

Figure 8. Sample of data from Cluster 1

The 'Middle Eastern' Cluster is characterised by its high concentration of Middle Eastern Restaurants, but it also has a high proportion of Asian and Korean restaurants. I assigned it this name because it is the only one of the clusters to boast such a high concentration of Middle Eastern restaurants.

3.1.2 Cluster 2 - Mexican Cluster

PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711	2.0	Mexican Restaurant	Asian Restaurant	Korean Restaurant
M4G	East York	Leaside	43.709060	-79.363452	2.0	Mexican Restaurant	Asian Restaurant	Korean Restaurant
M4R	Central Toronto	North Toronto West	43.715383	-79.405678	2.0	Mexican Restaurant	Asian Restaurant	Korean Restaurant

Figure 8. Sample of data from Cluster 2

The 'Mexican' Cluster is characterised by its high concentration of Mexican Restaurants, but it also has a high proportion of Asian and Korean restaurants. I assigned it this name because it is the only one of the clusters to boast such a high concentration of Mexican restaurants.

3.1.2 Cluster 3 - Asian Cluster

PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	3.0	Vietnamese Restaurant	Korean Restaurant	Tibetan Restaurant
M1P	Scarborough	Dorset Park, Wexford Heights, Scarborough Town...	43.757410	-79.273304	3.0	Vietnamese Restaurant	Korean Restaurant	Tibetan Restaurant
M4V	Central Toronto	Summerhill West, Rathnelly, South Hill, Forest...	43.686412	-79.400049	3.0	Vietnamese Restaurant	Korean Restaurant	Tibetan Restaurant

Figure 8. Sample of data from Cluster 3

In the Asian Cluster, the top 6 most common restaurants are all based around Asian cuisine, with Vietnamese, Korean and Tibetan restaurants featuring prominently.

3.2 Results of Shortest Path Algorithm

I found the most efficient route to visit all four clusters is, in order

1. 'The Annex, North Midtown, Yorkville'
2. 'Summerhill West, Rathnelly, South Hill, Forest Hill SE, and Deer Park'
3. 'Davisville'
4. 'North Toronto West'

This route can be used by food tourism companies to efficiently plan a tour of Toronto that encapsulates all the most prominent cuisines of the city.

5. Discussion and Recommendation

Based on the observations of my results, I have the following observations:

- This report can be used to experience as diverse a range of cuisines as Toronto has to offer.
- The most efficient way to visit all four clusters is outlined in the results section. This can be used by food tourism companies and tourists.

- The analysis of each of the clusters can also be used to know what restaurants are prominently available in each neighbourhood, allowing tourists to plan their meals which staying in Toronto.
- If a tourist is taking a day trip to Toronto and wants to experience the wider variety of cuisine with little hassle, then a ‘Melting Pot’ cluster allows them to experience this with minimal travel.

6. Conclusion

There is a wide and varied restaurant industry in Toronto. This report is an attempt to use data and the tools I learned in this course to identify how a food tourism company or tourist could navigate the cuisine of Toronto. I found the nighborhoods of Toronto could be clustered into 4 groups – the ‘Melting Pot’, ‘Middle Eastern’, ‘Mexican’ and ‘Asian’ clusters. The path that could be used to most efficiently visit one neighbourhood in each of the clusters is as follows:

1. 'The Annex, North Midtown, Yorkville'
2. 'Summerhill West, Rathnelly, South Hill, Forest Hill SE, and Deer Park‘
3. ’Davisville‘
4. ’North Toronto West'