Capstone Project

IBM Professional Certificate in Data Science

Opening an African Restaurant in Downtown Toronto, Canada

By Isaac Nyamekye

April 2020

Table of Contents

List of Tables

List of Figures

1 Introduction

Toronto is the most populous city in Canada and the provincial capital of Ontario. The city is very diverse and home to a large immigrant population. Because of the cultural diversity of its inhabitants, the city has a dynamic and diverse culinary scene

However, there are a few Ghanaian (West Africa) restaurants, especially, in downtown Toronto. An African restaurant chain would like to change this by opening a couple of Ghanaian restaurants in downtown Toronto.

To reduce competition with other African restaurants, the company would like to open their Ghanaian restaurants in neighbourhoods that have no or very few African restaurants. The purpose of this project was to identify such neighbourhoods in downtown Toronto.

The target audience of this project was mainly the senior executives of the African restaurant chain. This report would enable them to make an evidence-based decision on the best neighbourhoods to open their restaurants.

2 Data

This project used data from:

- o **Toronto Wikipedia page**: I scraped postal codes, boroughs and neighbourhoods' names from this page. Since the focus of my project is downtown Toronto, I only considered neighbourhoods in that borough.

- o **Geopy library**: This package was used to generate geolocation data (latitude and longitude) for the downtown Toronto neighbourhoods.

- o **Foursquare API**: From the location data obtained from web scraping and geocoding, I used the Foursquare API to generate African restaurant venue details for the respective neighbourhoods.

3 Methodology and Analysis

This section details the process I used to solve the business problem. The process comprised of:

- o **Data Acquisition and Cleaning:** This subsection outlines how I scraped and cleaned the neighbourhood data from Toronto's Wikipedia site. I also show how I used the Geopy library to generate the geolocation data needed for my analysis.

- o **Exploratory Data Analysis:** In this subsection, I explore how the neighbourhoods in downtown Toronto. I used the folium library to generate maps of the neighbourhoods and the Foursquare API to analyze the most common venues in each neighbourhood.

- o **Model Development:** This subsection shows the modelling process. I used K-mean clustering to determine neighbourhoods with no or few African restaurants.

3.1 Data Acquisition and Cleaning

3.1.1   Scrape Toronto Neighborhoods Data from Wikipedia

I used the BeautifulSoup Python library to scrap postal code, borough and name data from the Toronto Wikipedia page. Table 3.1 shows a subset of the data that was obtained from the page.

Table 3.1: Toronto Data from Wikipedia

	Postal code	Borough	Neighborhood
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park , Harbourfront
5	M6A	North York	Lawrence Manor , Lawrence Heights
6	M7A	Downtown Toronto	Queen's Park , Ontario Provincial Government

As the focus of my analysis was neighbourhoods in Downtown Toronto, I filtered out data for only the Downtown Toronto borough. Table 3.2 presents a subset of this data.

Table 3.2: Downtown Toronto Neighbourhoods Data

	Postal code	Borough	Neighborhood
4	M5A	Downtown Toronto	Regent Park , Harbourfront
6	M7A	Downtown Toronto	Queen's Park , Ontario Provincial Government
13	M5B	Downtown Toronto	Garden District, Ryerson
22	M5C	Downtown Toronto	St. James Town
31	M5E	Downtown Toronto	Berczy Park

3.1.2   Generate Latitude and Longitude for the Neighbourhoods

In this part, I generated latitude and longitude information for the downtown neighbourhoods using the pgeocode python library. This was to make the data usable in Foursquare API as the information is required to obtained venues common in these neighbourhoods. Table 3.3 presents a subset of the downtown neighbourhoods with their respective latitudes and longitudes.

Table 3.3: Downtown Toronto Neighbourhoods with Latitude and Longitude Data

	Postal code	Borough	Neighborhood	Latitude	Longitude
4	M5A	Downtown Toronto	Regent Park , Harbourfront	43.6555	-79.3626
6	M7A	Downtown Toronto	Queen's Park , Ontario Provincial Government	43.6641	-79.3889
13	M5B	Downtown Toronto	Garden District, Ryerson	43.6572	-79.3783
22	M5C	Downtown Toronto	St. James Town	43.6513	-79.3756
31	M5E	Downtown Toronto	Berczy Park	43.6456	-79.3754

3.2 Exploratory Data Analysis

3.2.1   Explore the Neighbourhoods

In this section, I explored the various neighbourhoods in Downtown Toronto. This involved using the geopy library to generate the latitude and longitude of Downtown Toronto and creating a map of Toronto with downtown neighbourhoods superimposed using the folium library. This map is shown in Figure 3.1.
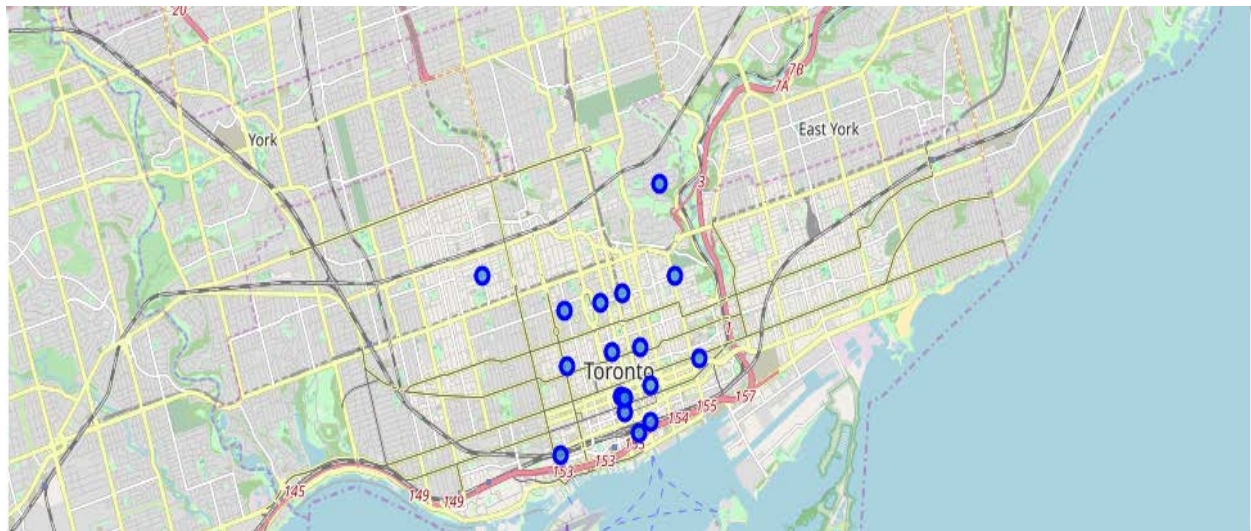

Figure 3.1: Map of Toronto with Downton Neighbourhoods Superimposed

I also used the Foursquare API to generate African restaurants within a 1km radius of each neighbourhood. Table 3.4 is a subset of this data and Table 3.5 shows the 3 most common venue categories.

Table 3.4: African Restaurants within 1km Radius of Downtown Toronto Neighbourhoods

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Regent Park , Harbourfront	43.6555	-79.3626	Ethiopian Keeffaa Forest Coffee	43.655958	-79.363977	Ethiopian Restaurant
1	Regent Park , Harbourfront	43.6555	-79.3626	King Solomon Queen Of Sheba	43.655731	-79.363986	Ethiopian Restaurant
2	Regent Park , Harbourfront	43.6555	-79.3626	Enat Buna	43.653920	-79.371570	Ethiopian Restaurant
3	Queen's Park , Ontario Provincial Government	43.6641	-79.3889	Ethiopian House	43.666599	-79.385669	Ethiopian Restaurant
4	Garden District, Ryerson	43.6572	-79.3783	Ethiopiques	43.656513	-79.377078	Ethiopian Restaurant

Table 3.5: Top 3 Common Venue Categories

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Berczy Park	Ethiopian Restaurant	Restaurant	Portuguese Restaurant
1	CN Tower , King and Spadina , Railway Lands , ...	Portuguese Restaurant	Restaurant	Ethiopian Restaurant
2	Central Bay Street	Ethiopian Restaurant	African Restaurant	Restaurant
3	Christie	Ethiopian Restaurant	African Restaurant	Restaurant
4	Church and Wellesley	Ethiopian Restaurant	Restaurant	Portuguese Restaurant

3.3 Model Development

I used *k*-means to group neighbourhoods into clusters. First, I used the elbow method to determine the optimal number of clusters (*k*) that would best represent the data. Figure 3.2 shows the elbow chart.
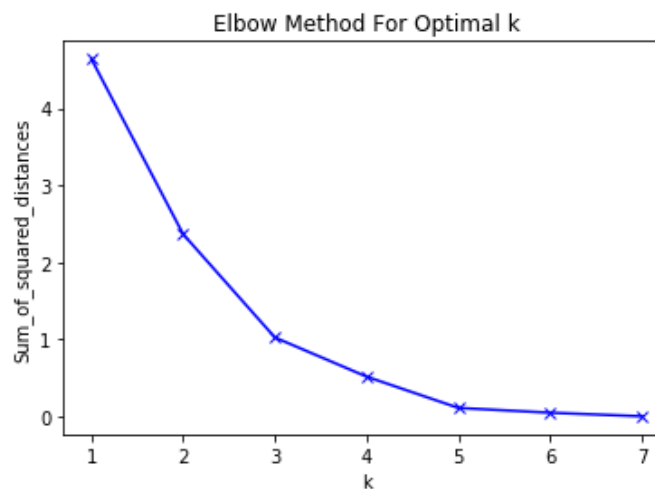

Figure 3.2: Elbow Method to Determine Optimal k

Based on the chart, I decided to run *k*-means to group the neighbourhoods into 3 clusters. Table 3.6 shows a subset of neighbourhoods and their cluster. Figure 3.3 presents this data on a map. Different coloured circles on the map represent the various clusters.

Table 3.6: African Restaurants within 1km Radius of Downtown Toronto Neighbourhoods

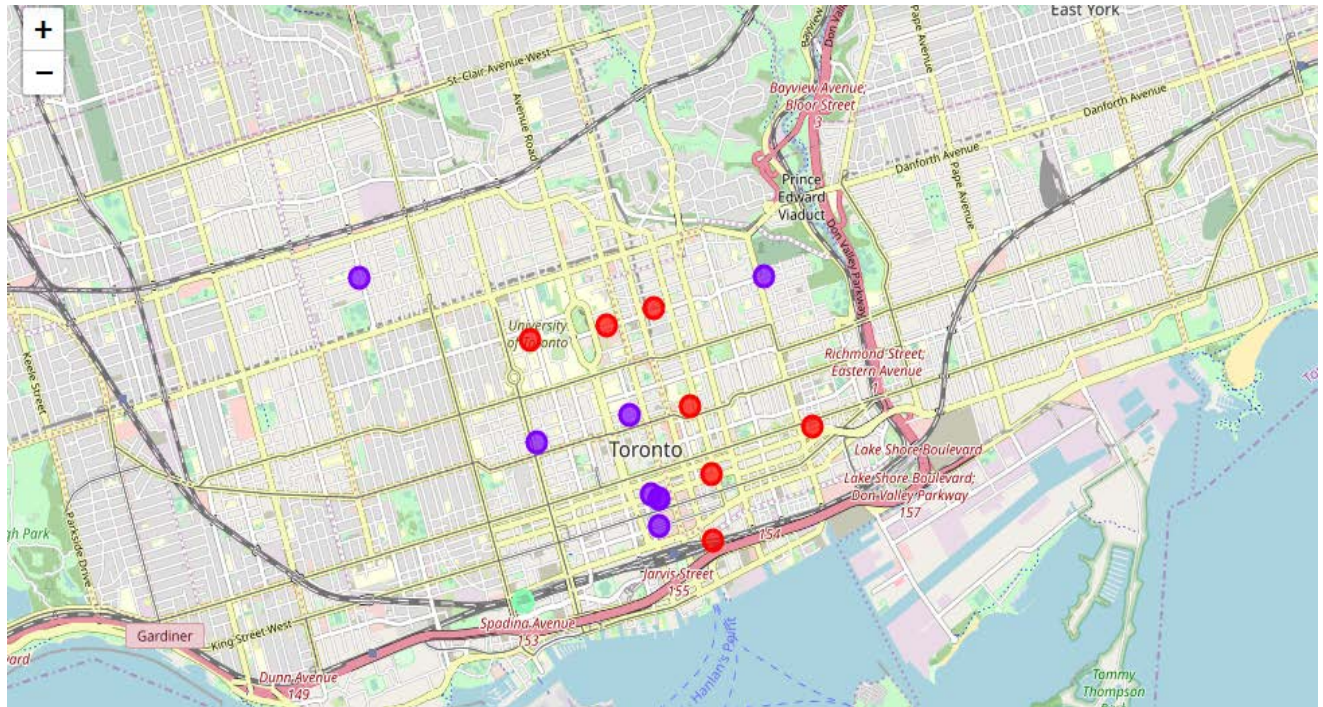|	Postal code	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
4	M5A	Downtown Toronto	Regent Park , Harbourfront	43.6555	-79.3626	0	Ethiopian Restaurant	Restaurant	Portuguese Restaurant
6	M7A	Downtown Toronto	Queen's Park , Ontario Provincial Government	43.6641	-79.3889	0	Ethiopian Restaurant	Restaurant	Portuguese Restaurant
13	M5B	Downtown Toronto	Garden District, Ryerson	43.6572	-79.3783	0	Ethiopian Restaurant	Restaurant	Portuguese Restaurant

Figure 3.3: Map of Downtown Toronto Neighbourhoods with Clusters

4 Results

Table 4.1 to Table 4.3 summarizes the characteristics of the various clusters. Cluster 0 has mainly Ethiopian Restaurants as the most common venues. Cluster 1 has mainly Ethiopian and other African Restaurants. Cluster 2 shows neighbourhoods with no or uncommon African Restaurant.

Table 4.1: Cluster 0 – Neighbourhoods with Ethiopian Restaurants as the Most Common Venues.

	Postal code	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
4	M5A	Downtown Toronto	Regent Park , Harbourfront	43.6555	-79.3626	0	Ethiopian Restaurant	Restaurant	Portuguese Restaurant
6	M7A	Downtown Toronto	Queen's Park , Ontario Provincial Government	43.6641	-79.3889	0	Ethiopian Restaurant	Restaurant	Portuguese Restaurant
13	M5B	Downtown Toronto	Garden District, Ryerson	43.6572	-79.3783	0	Ethiopian Restaurant	Restaurant	Portuguese Restaurant
22	M5C	Downtown Toronto	St. James Town	43.6513	-79.3756	0	Ethiopian Restaurant	Restaurant	Portuguese Restaurant
31	M5E	Downtown Toronto	Berczy Park	43.6456	-79.3754	0	Ethiopian Restaurant	Restaurant	Portuguese Restaurant
121	M5S	Downtown Toronto	University of Toronto , Harbord	43.6629	-79.3987	0	Restaurant	Ethiopian Restaurant	Portuguese Restaurant
165	M4Y	Downtown Toronto	Church and Wellesley	43.6656	-79.3830	0	Ethiopian Restaurant	Restaurant	Portuguese Restaurant

Table 4.2: Cluster 1 – Neighbourhoods with Ethiopian and Other African Restaurants as the Most Common Venues.

	Postal code	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
40	M5G	Downtown Toronto	Central Bay Street	43.6564	-79.3860	1	Ethiopian Restaurant	African Restaurant	Restaurant
41	M6G	Downtown Toronto	Christie	43.6683	-79.4205	1	Ethiopian Restaurant	African Restaurant	Restaurant
49	M5H	Downtown Toronto	Richmond , Adelaide , King	43.6496	-79.3833	1	Ethiopian Restaurant	African Restaurant	Restaurant
67	M5K	Downtown Toronto	Toronto Dominion Centre , Design Exchange	43.6469	-79.3823	1	African Restaurant	Restaurant	Portuguese Restaurant
76	M5L	Downtown Toronto	Commerce Court , Victoria Hotel	43.6492	-79.3823	1	Ethiopian Restaurant	African Restaurant	Restaurant
130	M5T	Downtown Toronto	Kensington Market , Chinatown , Grange Park	43.6541	-79.3978	1	African Restaurant	Ethiopian Restaurant	Restaurant
156	M4X	Downtown Toronto	St. James Town , Cabbagetown	43.6684	-79.3689	1	Ethiopian Restaurant	African Restaurant	Restaurant
157	M5X	Downtown Toronto	First Canadian Place , Underground city	43.6492	-79.3823	1	Ethiopian Restaurant	African Restaurant	Restaurant

Table 4.3: Cluster 2 – Neighbourhoods with Ethiopian and Other African Restaurants as the Most Common Venues.

	Postal code	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
139	M5V	Downtown Toronto	CN Tower , King and Spadina , Railway Lands , ...	43.6404	-79.3995	2	Portuguese Restaurant	Restaurant	Ethiopian Restaurant

5 Discussions

My analysis shows that neighbourhoods in cluster 0 and cluster 1 have African restaurants as the most common venues. To open the restaurants in neighbourhoods with no or few African restaurants, the results show that neighbourhoods in cluster 2 would be the best place for the African Restaurant chain to open their proposed Ghanaian restaurants.

The best neighbourhoods to open the proposed Ghanaian restaurants are CN Tower, King and Spadina, Railways Lands, Harbourfront West, Bathurst Quay, South Niagara and Island airport

6 Conclusions

The objective of this project was to identify neighbourhoods in downtown Toronto that would be good candidate locations for opening Ghanaian restaurants. By clustering African restaurants in Downtown Toronto using Foursquare and Toronto neighbourhoods data, I have identified CN Tower, King and Spadina, Railways Lands, Harbourfront West, Bathurst Quay, South Niagara and Island airport as neighbourhoods that would be optimal for opening Ghanaian restaurants.