Toronto’s Best and Worst Neighborhoods for Italian Cuisine

Laura Desjarlais

August 8, 2020

# Introduction

Toronto is a major city in the province of Ontario in eastern Canada. The city has a population of over 2.7 million people. “Toronto is an international centre of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world”. Immigrants to Canada have most often chosen Toronto as home which is represented by the more than 200 distinct ethnic origins living in the city. Some members of this diverse population have segregated into pockets known as little Italy, Chinatown, Portugal village and Little India. If a neighborhood has an ethnic majority inhabiting it, then is this the best neighborhood to go for ethnic cuisine? (*Source:* [*https://en.wikipedia.org/wiki/Toronto*](https://en.wikipedia.org/wiki/Toronto)*)*

## 1.1 Business problem

The objective of this investigation will be to identify the neighbourhoods in Toronto that have the most Italian restaurants, which therefore offers a good selection of Italian cuisine, and determine if “Little Italy” is in that list. Secondly, the neighbourhood with the least Italian restaurants will be noted as a recommendation for the best neighbourhood to open an Italian restaurant because there would be little competition.

## 1.2 Stakeholders

Determining the neighbourhood with the most Italian restaurants will be useful information for the Toronto tourism board when advising tourists who are visiting Toronto. In fact, the number of Italian restaurants in each Toronto neighbourhood could be passed on to the tourism board. The neighbourhood with the least Italian restaurants will be used by the Toronto community agencies for small business development. So if a businessman is interested in opening an Italian restaurant then a Toronto small business agency could advise him based on the data from Foursquare.

# 2.0 Data acquisition and cleaning

## 2.1 Data sources

The Toronto postal codes which are available on the Wikipedia page, link is below, will be combined with latitude and longitude data from the Week 3 assignment for transformation into the location of Toronto neighborhoods. The latitude and longitude data will be used in mapping the neighborhoods.

Wikipedia page link: https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M

Source of latitude and longitude coordinates for Toronto neighborhoods:

'Geospatial\_Coordinates.csv'

The Foursquare location data will be used for making a list of all the venues in each Toronto neighborhood.

Link: Foursquare API

The Python Folium library will be used to map the neighbourhoods with the Italian restaurant venues, after the venues have been identified and filtered from Foursquare and combined with geospatial coordinate data.

## 2.2 Data cleaning and preparation

Once the Toronto postal codes were parsed and extracted from Wikipedia using Beautiful Soup into a data frame, the rows with no boroughs assigned to the postal code were dropped. The index values of the data frame were then reset.

The longitude and latitude coordinate data was brought into a data frame. The two data frames were combined so that the location of each postal code and therefore borough was assigned.

# 3.0 Methodology

A map was generated to understand the number of Toronto neighbourhoods and the vastness of the city. The Foursquare location data was used to obtain the venues in each neighbourhood. A function was used to pull the venues into a data frame that specified the pertinent data including the venue name, venue latitude, venue longitude and venue category. This data frame was filtered by using the venue category data column so that only the Italian restaurants remained.

The neighbourhoods were grouped to get an understanding of the total number of Italian restaurants in each neighbourhood. As well, the extra columns were dropped from the data frame so only the neighbourhoods and the number of Italian restaurants remained. The list of neighbourhoods was sorted from highest to lowest number of Italian restaurants to clarify which neighbourhood has the largest number.

The neighbourhood with the largest number of Italian restaurants, Davisville, was not as expected since “Little Italy” is the neighbourhood called Christie. A map of these neighbourhoods was created using the Python Folium library to visualize the proximity of Christie to Davisville to determine if perhaps people of Italian descent lived close to Davisville and opened restaurants there. Special markers were added to the map to highlight the locations of Davisville and Christie by using the Folium library tool called Featuregroups.

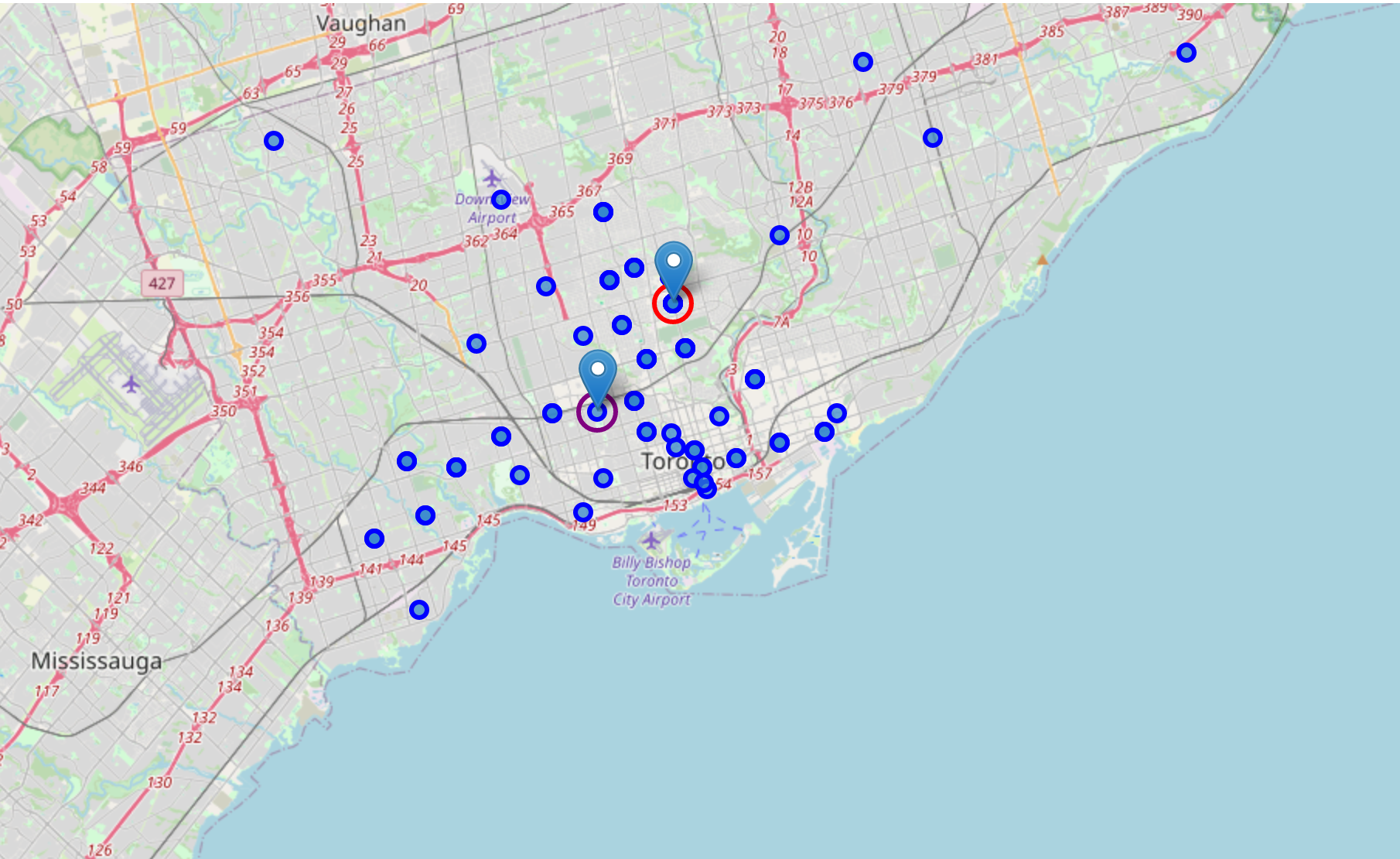
To determine the best neighbourhood(s) for opening an Italian restaurant, it was necessary to identify the neighbourhood(s) with no Italian restaurants and with other venues that would attract foot traffic. The assumption being that a good location for a new restaurant is one that is easily accessible and has high visibility. Using one hot coding in Python allowed the Toronto venue category column data to be transformed from categorical values into a data frame with integer values. The neighbourhoods were grouped, and the mean of the frequency was calculated for each group. Based on the mean, the top five venues in each neighbourhood group were sorted in ascending order and the information was transferred into a data frame by appending the five most common venues onto the neighbourhood group columns.

The neighbourhood groups were clustered into six distinct clusters based on the top five venues in each. Each cluster was classified based on the types of venues. The classification would help to determine which group of neighbourhoods would be good for opening an Italian restaurant. Out of the six clusters, cluster number three looked the most attractive for opening a new restaurant. The main borough in this cluster was North York but the specific neighbourhoods in the borough had yet to be identified. By using the data frame with the boroughs and neighbourhoods merged with the cluster data it was possible to pull out the neighbourhoods located in that specific North York location. The neighbourhoods of York Mills and Silver Hills were identified and mapped using the Folium library to observe the distance of these two neighbourhoods from other Toronto neighbourhoods that are not a good fit for a new Italian restaurant.

# 4.0 Results

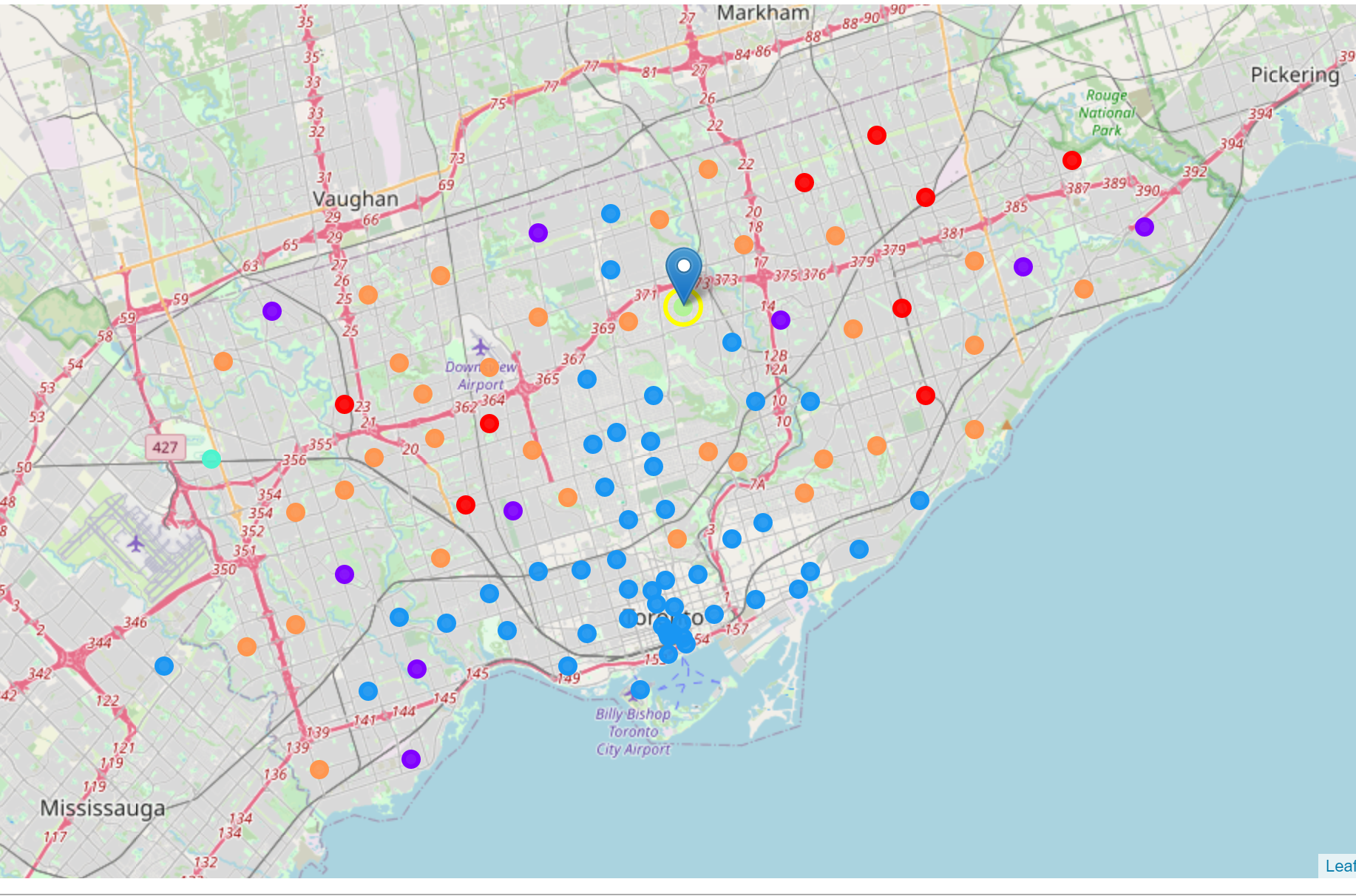
The results of this investigation were somewhat surprising. Firstly, visualizing all the Toronto neighbourhoods on a map gave a view of the expanse of the city. Determining the neighbourhood with the most Italian restaurants would be expected to be difficult with such a large selection of areas in a big city. After extracting and filtering the venues for Italian restaurants for the neighbourhoods using Foursquare, the neighbourhood called Christie, which is also known as “Little Italy”, only had one Italian restaurant. Whereas the neighbourhood with six and the most Italian restaurants was Davisville. After visualizing these two neighbourhoods on a map, figure 1, it was apparent that they were in relatively close proximity to each other. Christie is shown in purple and Davisville is shown in red.

Figure 1. Map of Toronto neighbourhoods with Italian restaurants.



The borough in Toronto that had some interesting venues to attract foot traffic but also had no Italian restaurants in the top five venues was North York. This was determined after clustering the neighbourhoods into six clusters with common features. The purpose of clustering was to be able to narrow down and select the cluster which had characteristics of a neighbourhood(s) that would be the most desirable to open an Italian restaurant. Out of the six clusters, cluster number three was selected as a good area for a new Italian restaurant. The cluster was labelled as “Foot traffic neighbourhoods” and had the characteristics of having parks, pools, and dog runs but without an Italian restaurant in the top five venues. Thus, the borough of North York, yellow circled marker in figure 2, stood out as a good candidate for opening a new Italian restaurant. Referring to the data frame that associated the specific borough in each cluster with the neighbourhoods, the neighbourhoods of York Mills and Silver Hills were identified.

Figure 2. Map of Toronto neighbourhood clusters.



# 5.0 Discussion

The primary observation made during this investigation was that assumptions can’t be made about the characteristics of neighbourhoods without some detailed analysis of the data for types of venues. A good example of this was “Little Italy” not having the greatest number of Italian restaurants in Toronto. Therefore, using Foursquare is highly recommend as a very useful tool for gaining an understanding of neighbourhoods.

# 6.0 Conclusion

Christie is known as “Little Italy” due to the number of residents with Italian heritage. The Toronto tourism board may have anticipated that Christie would be a good neighbourhood to send tourists for Italian cuisine but as it turns out the restaurant selection there would be limited. Davisville was the neighbourhood with the greatest number of Italian restaurants. After further investigation on the internet, Davisville is quite affluent and the proximity to Christie is close. According to Wikipedia, “approximately 33% of Davisville's households earn more than $60,000 per year, which is more than double Toronto, which has an approximately 16.2% of its households earning more than $60,000 per year” (source: <https://en.wikipedia.org/wiki/Davisville_Village>). The reason that Italians may have built restaurants in Davisville is therefore because their businesses are more profitable in Davisville.

The Toronto small business association will find it useful to know which neighbourhoods are recommended for development of a new Italian restaurant after data analysis. The basis for selecting York Mills and Silver Hills was on the existing venues that would attract foot traffic and not having any existing Italian restaurants within the top five types of venues.