

# Toronto Neighborhoods & Venues



Suppose you have a friend who plans to move to Toronto, Canada but they want to live in an area similar to their neighborhood in downtown manhattan where there are many cafes, restaurants and shops.

## Problem

Suppose your friend enlists you to help them find a neighborhood that suits their needs. Which Toronto neighborhood should they live in?

The first approach to this problem is to collect all of the data and tools to make the best possible decision.

## Data/ Resources:

- 1) The following wikipedia page with Toronto neighborhoods categorized:  
[https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
- 2) Geospatial coordinates CSV file, with longitude and latitude coordinates of neighborhoods categorized by zip code.
- 3) Foursquare software will be utilized. local search-and-discovery mobile app which provides search results for its users. The app provides personalized recommendations of places to go near a user's current location based on users' previous browsing history and check-in history.

df\_GeoComplete.head(20)

Out[13]:

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Harbourfront, Regent Park	43.654260	-79.360636
3	M6A	North York	Lawrence Heights, Lawrence Manor	43.718518	-79.464763
4	M7A	Queen's Park	Queen's Park	43.662301	-79.389494
5	M9A	Etobicoke	Islington Avenue	43.667856	-79.532242
6	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353
7	M3B	North York	Don Mills North	43.745906	-79.352188
8	M4B	East York	Woodbine Gardens, Parkview Hill	43.706397	-79.309937
9	M5B	Downtown Toronto	Ryerson, Garden District	43.657162	-79.378937
10	M6B	North York	Glencairn	43.709577	-79.445073

In [ ]: # get venues

## 1. Dataset

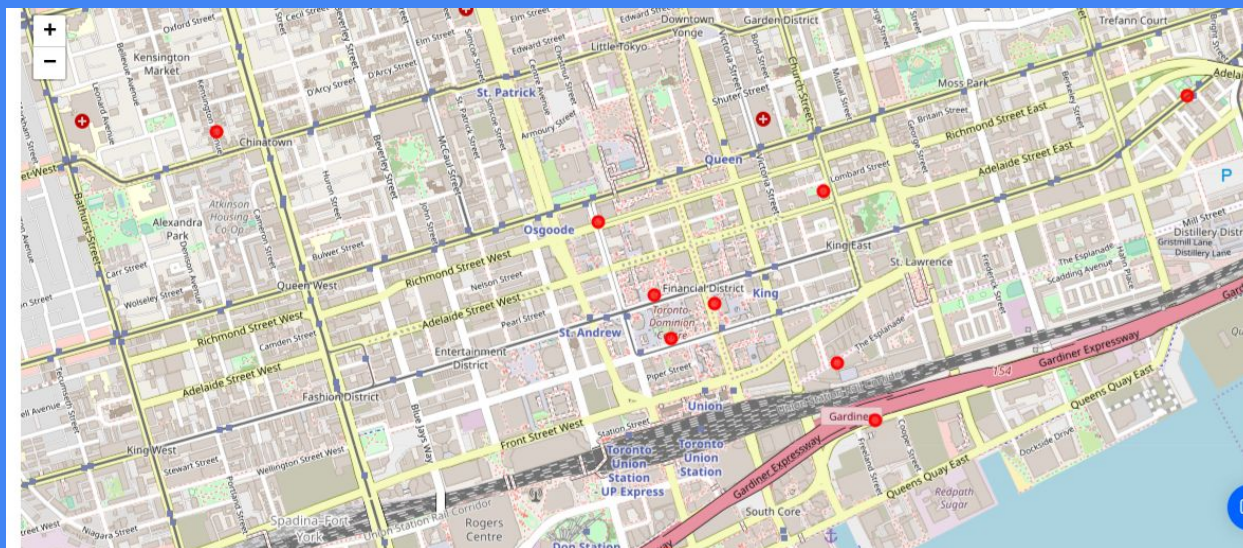
## Methodology

The data from the wikipedia page will be extracted using the software BeautifulSoup. BeautifulSoup is used to extract tables from websites. Once the table is extracted, it will be imported into our dataframe. Similarly, the geospatial coordinates CSV file along will also be added into our dataframe. The two tables have a common feature and that is they both have a column for zip codes. The two tables will be joined by zip code.

## Methodology Con't

Once the dataframe is set. We then utilize Foursquare to explore Toronto neighborhoods. We want to see all the venues the Toronto area provides for our consideration. Specifically, we are looking for neighborhoods with cafes, restaurants and shops.

The dataframe is then clustered. Four clusters are created. Each cluster has all the popular venues we need in our decision making. Each cluster represents a top neighborhood. Once clustered in the dataframe, the dataframe is then visualized.





## Conclusion

The cluster with that best fits our model will represent the best neighborhood for our friend. The neighborhood that we would recommend our friend to move to in Toronto, is Downtown Toronto.