

# Finding Neighborhoods in Downtown Toronto similar to Parkwoods

Chibueze Louis Alagbogu

7<sup>th</sup> July, 2021

## 1. Introduction: Problem

In this project we will try to find locations in Downtown Toronto similar to Parkwoods. One of the most popular venue in Parkwoods is KFC, if KFC wishes to replicate their success from Parkwoods in Downtown Toronto, this project would be very much helpful. It can also be used by residents moving from Parkland and wish to be in a similar neighborhood in Downtown Toronto.

We will use our data science powers to generate similar and dissimilar neighborhoods to guide the stakeholders.

## 2. Data

Based on definition of our problem, the major factor that will influence our decision are the types of venues in each neighborhood

Following data sources will be needed to extract/generate the required information:

- List of Postal Codes in Canada gotten from the [Wikipedia page](#)
- CSV file containing Canadas [Geospatial coordinates](#)
- Types of venues in every neighborhood in Downtown Toronto will be obtained using *Foursquare API*
- Geographical coordinate of Downtown Toronto will be obtained using *Geocoder*

### 2.1 Scrapping and wrangling the data of Toronto Neighborhoods

Scrapping the Wikipedia webpage using *BeautifulSoup* to extract data such as:

- Postal Code
- Borough and
- Neighborhood of the various cities in Canada.

## 2.2 Reading Geospatial Coordinates

The geographical coordinates of the neighborhoods containing each postal code was gotten from a csv file downloaded through the link above.

## 2.3 Merging the data to add Latitude and Longitude to the Data Frame

The geospatial data containing the postal codes contains just three (3) columns:

- Postal Code
- Latitude
- Longitude

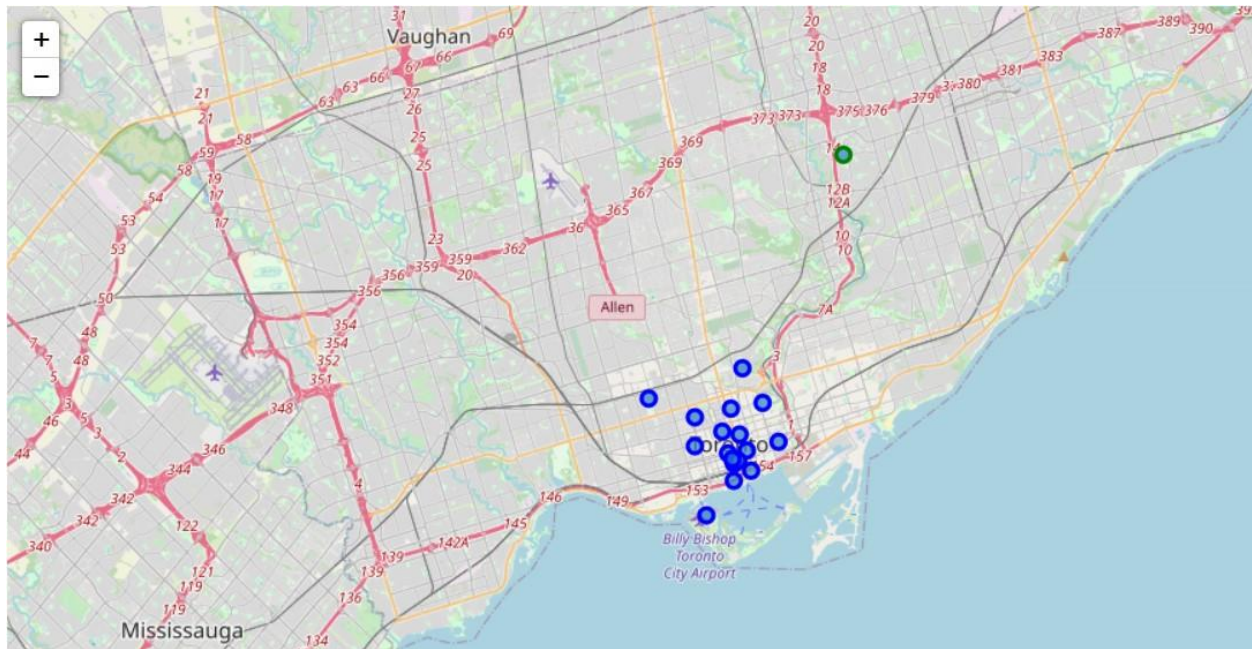
So, the data would be merged with the one gotten after scrapping the Wikipedia page, the criteria for merging both data would be the Postal Code. In line with our objective, the data from Parkwood neighborhood and Downtown Toronto borough would be separated and each would have their own individual data frames.

## 3. Methodology

In this project we will direct our efforts on detecting areas in Downtown Toronto that are similar to the Parkwoods Neighborhood using k-means clustering. The neighborhoods in Downtown Toronto that are similar to Parkwoods neighborhood would be clustered together.

### 3.1. Visualizing the Map of Downtown Toronto

*Geocoder* was used to get the coordinates of Downtown Toronto and we visualized the coordinates on a map using *folium*. Blue markers would be added to show the neighborhoods in Downtown Toronto and a green marker to show Parkland neighborhood.



### 3.2. Four Square API

Now that we have our location coordinates, Foursquare API was used to get info on venues in each neighborhood.

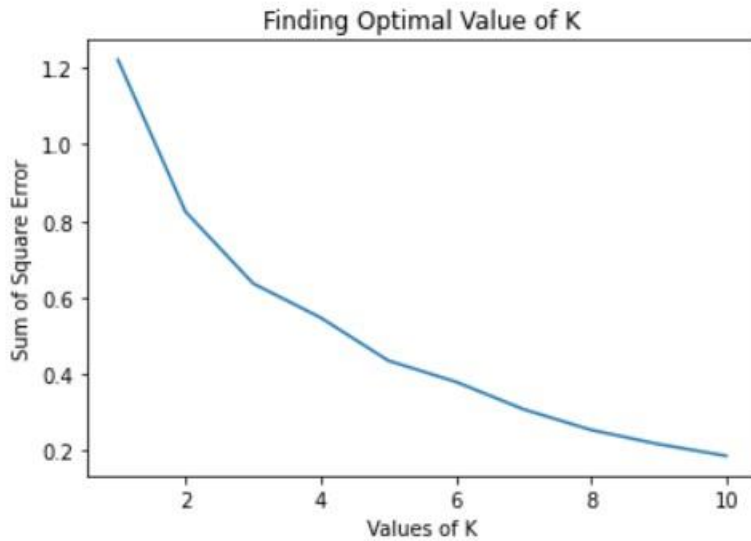
We have also grouped the data by venue category (according to Foursquare categorization) and applied One hot encoding. Now one hot encoding is used to convert all the different venue categories from strings to real numbers. One (1) and zero (0) is used to represent the availability of a venue category. If a venue category exists in a neighborhood, it is represented by one (1) and if it doesn't, it is represented by zero (0).

Converting the data to numerical values would make analysis easier along the line.

### 3.3. K-means Clustering

An unsupervised machine learning classification method known as k-means clustering is used. The neighborhoods in Downtown Toronto that are similar to Parkwoods neighborhood would be clustered together.

In order to choose an optimal value of K in the k-means clustering method, we used the Elbow method. In this method, the sum of square errors is plotted against the different values of K and the value that has the most critical bend is chosen as the optimal value.



From the plot, the optimal value of K is 3. So, we use K=3 for our k-means clustering. Then the labels from the k-means clustering are merged to the data frame. After merging the cluster label to the data frame, we check for the label of Parkwoods and then view all the neighborhoods containing similar labels and also, the neighborhoods containing dissimilar labels.

#### 4. Results

Our analysis shows that there are a great number of neighborhoods in Downtown Toronto which are similar to Parkwoods neighborhood. Only two (2) neighborhoods are different.

Image showing the dissimilar neighborhoods:

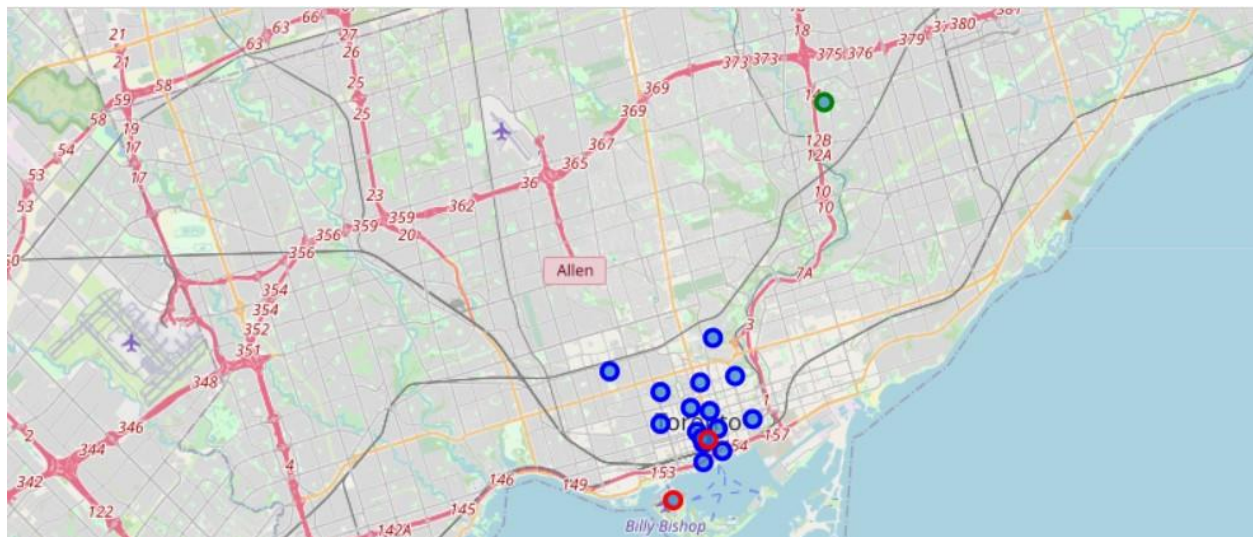
	PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels
10	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817	1
13	M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...	43.628947	-79.394420	2

It is also important to note that from our result, even though these neighborhoods stated above are dissimilar to Parkwoods neighborhood, they are also dissimilar to themselves. Those neighborhoods are NOT similar.

Image showing the similar neighborhoods below:

	PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels
0	M3A	North York	Parkwoods	43.753259	-79.329656	0
1	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	0
2	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	0
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	0
4	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	0
5	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383	0
6	M6G	Downtown Toronto	Christie	43.669542	-79.422564	0
7	M5H	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568	0
8	M5J	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752	0
9	M5K	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576	0
11	M5S	Downtown Toronto	University of Toronto, Harbord	43.662696	-79.400049	0
12	M5T	Downtown Toronto	Kensington Market, Chinatown, Grange Park	43.653206	-79.400049	0
14	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529	0
15	M4X	Downtown Toronto	St. James Town, Cabbagetown	43.667967	-79.367675	0
16	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280	0
17	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160	0

*Folium* was once again used to visualize the information using a map. The Blue marker represents the similar neighborhoods in Downtown Toronto and the red marker represents dissimilar neighborhoods while the green marker represents Parkwoods neighborhood.



## 5. Discussion

Based on the result, if one is moving from Parkwoods to Downtown Toronto and is curious to stay in a similar or dissimilar neighborhood from Parkwoods using Nearby venues as the criteria.

Neighborhoods in Downtown Toronto are all similar to Parkwoods bar Commerce court, Victoria hotel, CN Tower, King and Spadina and Railway Lands.

## 6. Conclusion

Purpose of this project was to identify neighborhoods in Downtown Toronto that are similar to Parkwoods Neighborhood. So, if one is moving from Parkwoods to Downtown Toronto and is curious to stay in a similar or dissimilar neighborhood from Parkwoods using Nearby venues as the criteria. Neighborhoods in Downtown Toronto are all similar to Parkwoods bar Commerce court, Victoria hotel, CN Tower, King and Spadina and Railway Lands.

Final decision should be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location, levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.