The Battle of Neighborhoods - Report

1. Introduction/Business Problem

The city of Toronto is the most populous city in Canada, with a population of 2,731,571 in 2016. It is the centre of business, finance, arts and culture, being recognized as one of the most multicultural and cosmopolitan cities in the world. Its varied cultural institutions attract for over 25 million tourists each year, and is one of Canada's leading tourism destinations.

The booming tourism industry has propelled the increase in demand for hotels and hostels in the region. The demand for travel accommodations would also lead the increase in the appeal of laundry services. In a city with high tourist volume, tourists would often look for cheaper alternative than the hotel laundry services, and thus this would be a good business opportunity.

Thus, the problem statement is as follow:

What is the best place to set up a laundromat in the city of Toronto that is in close proximity to travel accommodations, yet at a distance from other competing laundromats?

2. Data

The most important data required to answer the problem statement includes the **locations and types of venues** around the Toronto city.

First, the **locations of travel accommodations** such as hotels, hostels and motels in the region will be extracted to estimate the flow of travellers and potential amount of customers that would frequent the laundromat as they are the primary target market. Next, the **locations of all existing laundromats or laundry services** will need to be obtained. This is to reduce the competition by setting up the laundromat a certain distance away from other existing laundromats.

An ideal location for the laundromat would be at the centre of a cluster of hotels and hostels as mentioned. A clustering algorithm can be performed on the locations of travel accomodations and laundromats separately to find the concentration of both types of the establishments. The ideal location can be inferred by comparing both clustering results.

3. Methodology

To search for the ideal locations for setting up the laundromats, the locations of the existing laundromats and hotels in Central Toronto (instead of the entire Toronto in order to limit the range of the data) are necessary. The coordinates of the neighborhoods in Toronto need to be found by querying the Foursquare API. The following steps were taken in the coding to achieve the purpose.

1. A *pandas* dataframe listing the boroughs and neighborhoods in Toronto is build by scraping and processing the Wikipedia page (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M_).

2. The geographical coordinates of each neighborhood was obtained from the .csv file downloaded from http://cocl.us/Geospatial_data.

3. The postal codes for each neighborhoods were matched to the information from the csv file. The original *pandas* dataframe was updated with the corresponding geographical coordinates.

4. The neighborhoods from borough of Central Toronto were extracted and listed in a separate dataframe.

5. Foursquare API was used to query each of the neighborhood for the intended venues (laundromats and hotels). The resulting locations were tabulated and stored in a new dataframe and visualized using Folium.

6. In order to observe the aggregation of the laundromats and hotels, hierarchical clustering was used to cluster the laundromats and hotels separately. Hierarchical clustering was used and the resulting clusters were shown on the Folium map.

7. The visualization on Folium map is further improved by obtaining the cluster centers for each group of laundromat and hotel with the size of the circle marker weighted in correspondence to the number of venues in the specific cluster.

4. Results

There is a total of 9 neighborhoods listed under Central Toronto. The laundromats and hotels around these neighborhoods were explored. Figure 1 shows the location of the venues (laundromats are represented by red; hotels by blue)

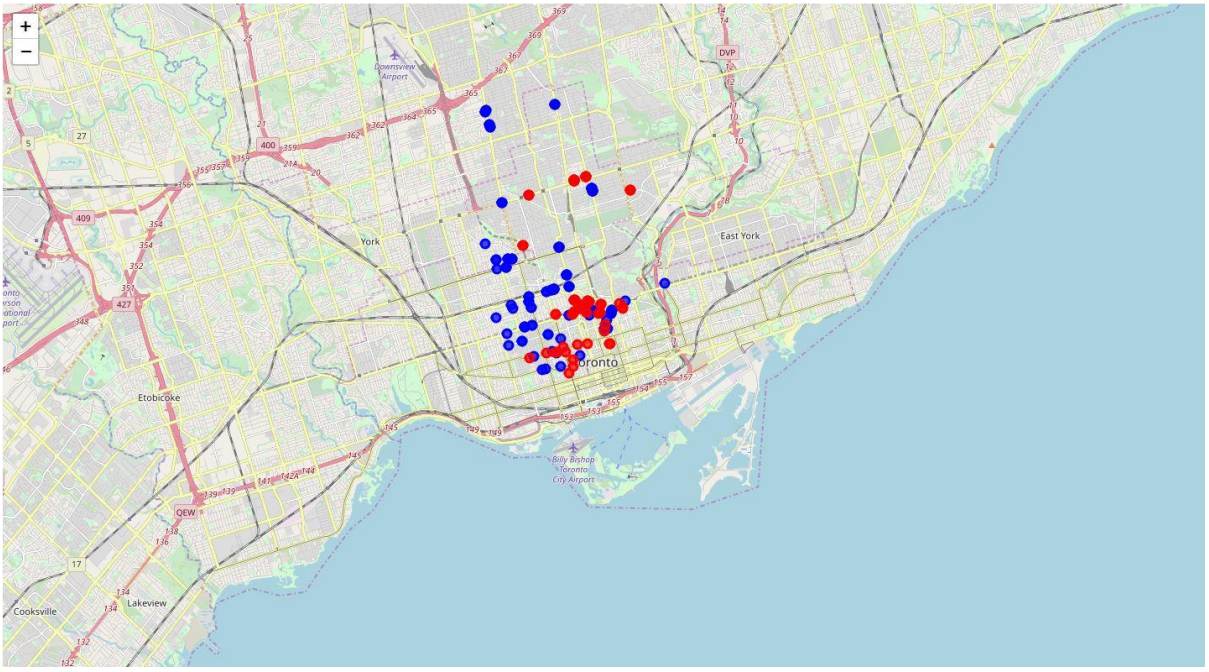

Figure 1: Folium map showing the location of laundromats (red) and hotels (blue)

From the Foursquare API, there are a total of 106 laundromats and 87 hotels in these neighborhoods.

A total of 5 clusters of laundromats and 4 clusters of hotels were produced from executing the Hierarchical clustering and were shown in Figure 2.

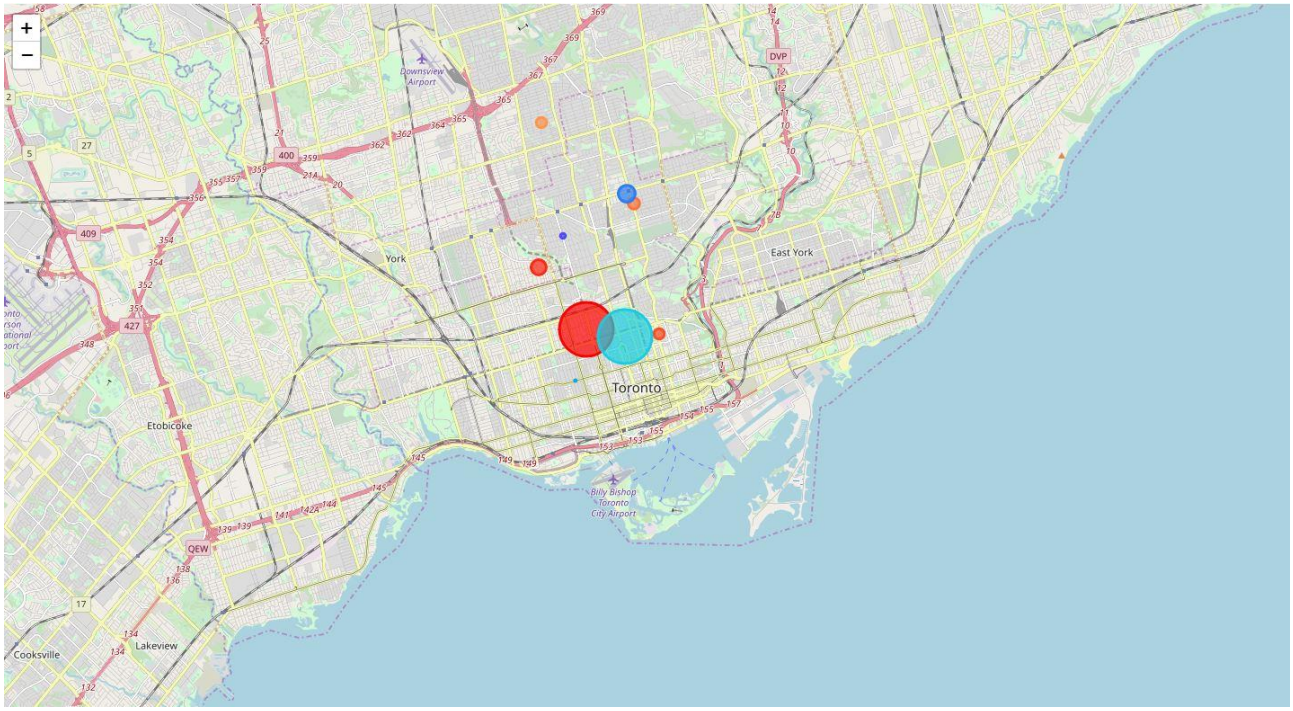

Figure 2: Folium map showing clusters of laundromats (red) and hotels (blue)

5. Discussions

1. From Figure 3, the laundromat and hotel clusters are found to be generally overlapping or close to each other. This confirms the expectation laundromats' business is dependent on the tourist population.

2. A large concentration of laundromats and hotels alike can be found surround the Annex neighborhood, which is shown in the large clusters shown in Folium map (Laundromat Cluster 4 and 5; Hotel Cluster 3).

3. Further observation shows that there is a relatively smaller hotel concentration at the Forest Hill neighborhood with low number of laundromats surrounding it (Hotel Cluster 2). The closest cluster of laundromats is located further away in the neighborhood of Humewood (Laundromat Cluster 3).

4. There is a small cluster of laundromats in the neighborhood of Caribou park (Laundromat Cluster 1) which does not have any neighboring hotel clusters. This is the location populated with boarding schools and hospitals.

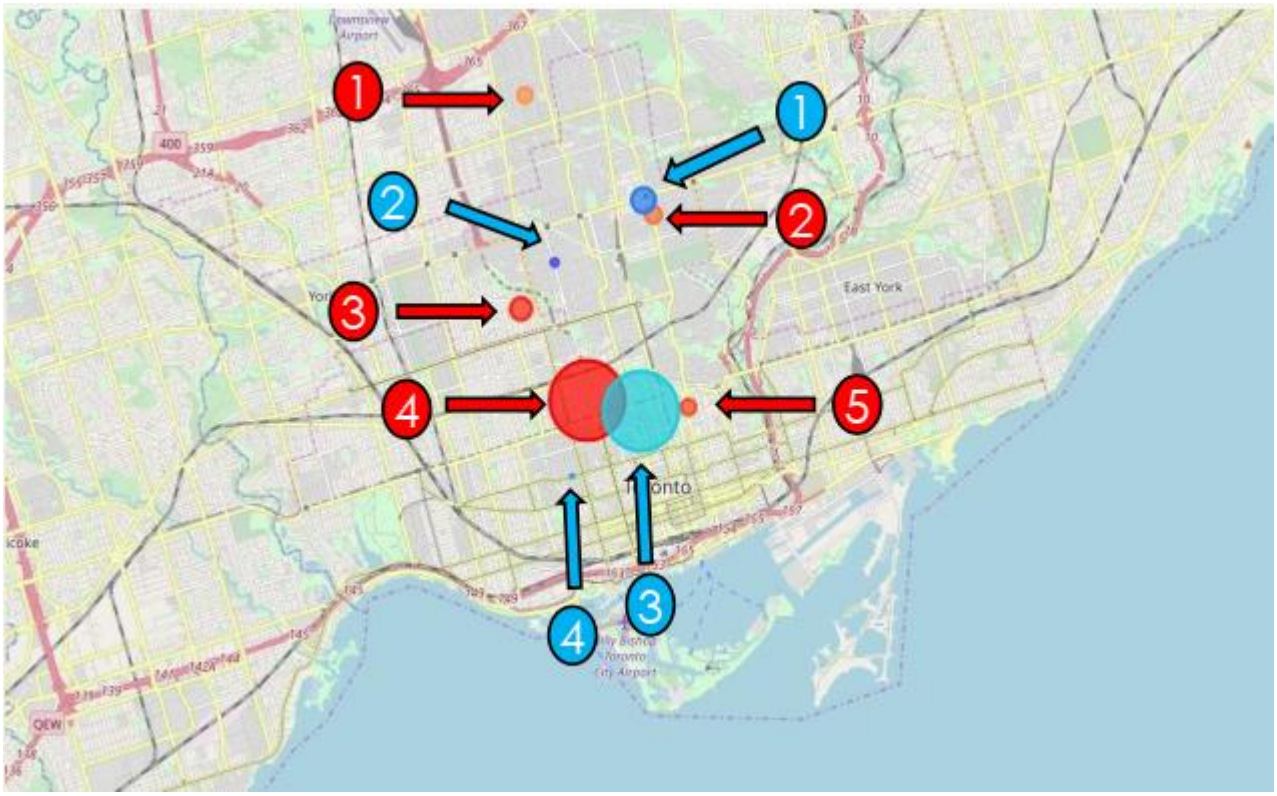

Figure 3: : Folium map showing labelled clusters of laundromats (red) and hotels (blue)

Conclusion

The most suitable place for setting up a new laundromat is in the neighborhood of Forest Hill (Hotel Cluster 2) in Central Toronto. This fulfil both the requirement of being in close proximity to travel accommodations and a distance away from other competing laundromats.