Capstone project - Recommending a location for a new hotel in Toronto

Freddy Adiv

1. Introduction
   1.1 Background
   A worldwide hotels firm is wishing to build their first hotel in Toronto which already have 518 hotels (according to tripadvisor.com, https://www.tripadvisor.com/Hotels-g155019-Toronto_Ontario-Hotels.html).

   1.2 The problem
   There are several key factors needed to be considered in order to decide the proper location to build the new hotel on:
      1.2.1 The proximity from the sea – a view to the sea is highly crucial for the firm's brand.
      1.2.2 People's traffic – The hotel must be in a central location, close to venues that are commonly used by tourists.
      1.2.3 The area to be selected should not be crowded – in the factor of ration of venues in the area to the quantity of hotels, the hotel should not precede the 5th most common venue.

2. The data
   The data to be used in this research will include the following sources:
   - Toronto's postal codes from https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
   - Toronto's geo-spatial data from https://cocl.us/Geospatial_data
   - Venues in Toronto from foursquare.com

   The data will be used to identify the boroughs of Toronto's, then clustered by venues commonly found on each borough to identify the possible locations for the new hotel.
   Once identified, find potential locations that answers the key factors mentioned in section 1.1.

3. Methodology
   3.1 Data preparation
   The postal codes of Toronto were scrapped from Wikipedia's page, any Not Assigned Borough records were removed from the data frame.
   The geo-spatial data was added to the postal codes data frame, enabling map representation of Toronto's neighborhoods.
   A new data frame of venues in Toronto was queried from the foursquare database, and connected by proximity to each other and to boroughs.
   3.2 Exploratory Data Analysis
   Using the geospatial data we were able to create a map of Toronto and its neighborhoods:

Figure 1 - Map of Toronto

After acquiring the venues list, we were able to cluster the neighborhoods based on their 10 most common venues.
All 36 neighborhoods were clustered into 10 different groups, according to the top 10 venues each have, out of 1706 available venues.
The following table shows the 5 most common venues in each cluster:

Table 1 - Common venues for clusters

Cluster ID	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Cluster 0	Supermarket	Pharmacy	Bakery Gym	Fitness Center	Music Venue
Cluster 1	Park	Hotel	Fast Food Restaurant	Pizza Place	Convenience Store
Cluster 2	Coffee Shop	Bakery	Clothing Store	Cosmetics Shop	Italian Restaurant
Cluster 3	Garden	Wings Joint	Farmers Market	Falafel Restaurant	Event Space
Cluster 4	Gym	Playground	Summer Camp	Wings Joint	Dim Sum Restaurant
Cluster 5	Park	Trail	Jewelry Store	Sushi Restaurant	Wings Joint
Cluster 6	Park	Lawyer	Bus Line	Swim School	Wings Joint
Cluster 7	Park	Playground	Trail	Building	Wings Joint
Cluster 8	Breakfast Spot	Gift Shop	Bookstore	Dog Run	Restaurant
Cluster 9	Health Food Store	Other Great Outdoors	Trail	Pub	Wings Joint

The segmentation of the clusters on the map can be seen in Figure 2:

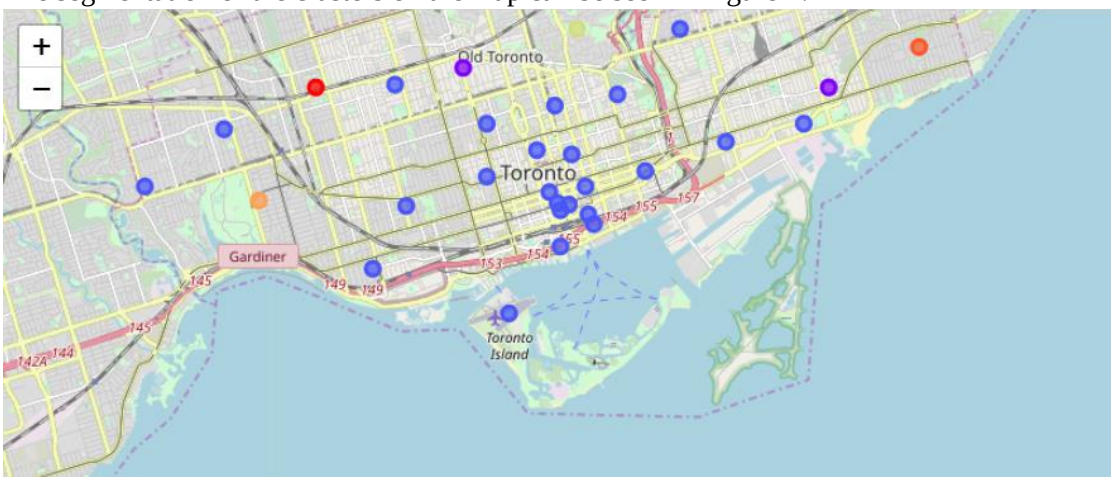

Figure 2 - Segmentation of Toronto to clusters

At this point we identified that cluster 2 have the most hotels in its top 5 venues, making it as a candidate collection of possible locations of a new hotel.

Cluster 2 areas are shown in the following map:

Figure3 - Cluster 2 on Toronto's map

We can now find Cluster 2 areas that doesn't have hotels in their top 5 most common venues, as shown on the following table:

Table2 - Cluster 2 areas without many hotels

	Postalcode	Borough	Neighborhood	Latitude	Longitude
2	M5A	Downtown Toronto	Harbourfront, Regent Park	43.654260	-79.360636
9	M5B	Downtown Toronto	Ryerson, Garden District	43.657162	-79.378937
20	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
24	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
25	M6G	Downtown Toronto	Christie	43.669542	-79.422564
30	M5H	Downtown Toronto	Adelaide, King, Richmond	43.650571	-79.384568
37	M6J	West Toronto	Little Portugal, Trinity	43.647927	-79.419750
41	M4K	East Toronto	The Danforth West, Riverdale	43.679557	-79.352188
43	M6K	West Toronto	Brockton, Exhibition Place, Parkdale Village	43.636847	-79.428191
54	M4M	East Toronto	Studio District	43.659526	-79.340923
69	M6P	West Toronto	High Park, The Junction South	43.661608	-79.464763
73	M4R	Central Toronto	North Toronto West	43.715383	-79.405678
80	M5S	Downtown Toronto	Harbord, University of Toronto	43.662696	-79.400049
81	M6S	West Toronto	Runnymede, Swansea	43.651571	-79.484450
84	M5T	Downtown Toronto	Chinatown, Grange Park, Kensington Market	43.653206	-79.400049
86	M4V	Central Toronto	Deer Park, Forest Hill SE, Rathnelly, South Hi...	43.686412	-79.400049
87	M5V	Downtown Toronto	CN Tower, Bathurst Quay, Island airport, Harbo...	43.628947	-79.394420
96	M4X	Downtown Toronto	Cabbagetown, St. James Town	43.667967	-79.367675
99	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
100	M7Y	East Toronto	Business Reply Mail Processing Centre 969 Eastern	43.662744	-79.321558

Placing these areas' coordinates on the map enables us to identify which locations are close to the sea:

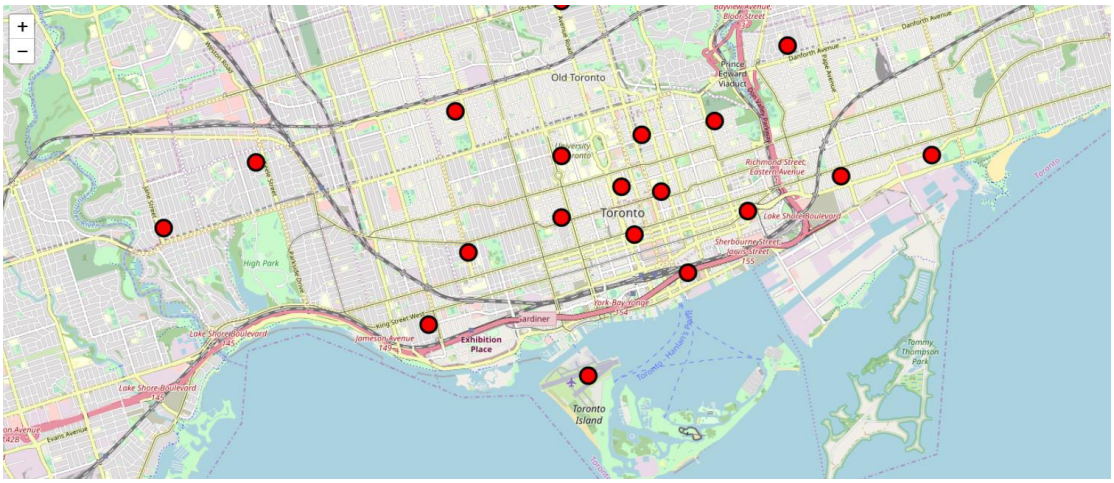

Figure4 - Potential locations

4. Results

Our findings are that the following locations are good candidates for a new hotel:

- Business Reply Mail Processing Centre 969 Eastern, East Toronto
- Studio District, East Toronto
- Harbourfront, Regent Park, Downtown Toronto
- Berczy Park, Downtown Toronto
- Brockton, Exhibition Place, Parkdale Village, West Toronto
- CN Tower, Bathurst Quay, Island airport, Harbourfront West, King and Spadina, Railway Lands, South Niagara, Downtown Toronto

5. Conclusion

   In this research I analyzed the location of venues around Toronto in order to identify optional locations for a new hotel. The findings show that there are six locations which were clustered as typical places for hotels.
   This approach my help recommending on places for other purpose such as stores, coffee shops etc...

6. Resources

   3.3 Notebook on Watson:
   https://eu-de.dataplatform.cloud.ibm.com/analytics/notebooks/v2/f39d67d0-bd28-4500-a00e-9b2a5213caee/view?access_token=71ef818db51dfd8e12620e22c8b4449bb78abcfb3753d9958cb16616846280a7

   3.4 Notebook on GitHub:
   https://github.com/FredAdiv/Coursera_Capstone/blob/master/Capstone%20Notebook.ipynb