**ABC Hotel Group Toronto-Location Search**

**Section-1**

**Introduction/Business Problem**

ABC Hotel Group was seeking an opportunity to expand their hotel business and have decided to build a new luxury hotel in the city of Toronto. Now the top leadership team at ABC Hotel Group is facing the most of important thing for the business – picking an appropriate location.

**Background**

Toronto is not only the capital city of the province of Ontario, but also a center of business, finance, arts, and is recognized as one of the most multicultural and cosmopolitan cities in the world. Included luxury, cheap, 5-star and other type hotels, Toronto already has about 183 hotels with a total of almost 36,000 rooms (according to Wikipedia).

**The importance in a hotel industry**

In every business location plays a very important role in its success and hotels are no exception. Every hotel will have its own target groups or customers which include local national and international tourists, people coming for business purposes, conferences, wedding bookings and so on. Target groups may differ in each hotel but most of these groups prefer locations which are attractive and locations usually located near to different transport links. Therefore, a combination of reasonable venues nearby is easier to catch people’s attentions and thus would stay longer in their memories.

**Competition**

There is consumer “theory” –according to Greg Kahn, founder and CEO of Kahn Research Group in Huntersville, NC, and a behavioral research veteran who's done location research for Arby's, Buffets Inc., Home Depot, Subway and other major and minor players. He said: “Quite simply, the best place to be is as close to your biggest competitor as you can be."

Therefore, the senior leadership team at ABC Hotel Group needs a data-driven analytics report to help them make a decision.

**Section-2**

**Data**

To solve the business problem, we will use the following source data:

**Toronto Borouhs/Neighbourhoods**

(Wikipedia - Toronto Boroughs/Neighbourhoods: <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>)

It’s public data and its table contains a list of postal codes in Canada where the first letter is M. Postal codes beginning with M are located within the city of Toronto in the province of Ontario. Only the first three characters are listed, corresponding to the Forward Sortation Area.

**GeoCoder/Google Geolocation APIs**

Geocoding is the process of converting addresses (like "1600 Amphitheatre Parkway, Mountain View, CA") into geographic coordinates (like latitude 37.423021 and longitude -122.083739), which we can use to place markers or position the map.

The Maps JavaScript API provides a geocoder class for geocoding and reverse geocoding dynamically from user input.

**Foursquare APIs**

Foursquare is a local search-and-discovery service mobile app which provides search results for its users. The app provides personalized recommendations of places to go to near a user's current location based on users' "previous browsing history, purchases, or check-in history". Basically, the API service included:

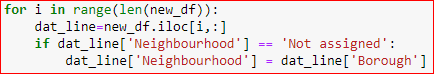
* Access to millions of fresh venue-related tips, tastes, photos & attributes from the Foursquare community
* Access places data - in real time - from any connected application
* Accurately assign a mobile app user to a specific location (Snap-to-Place)
* Enable users to search and discover venues via a mobile app or website
* Geo-tag content such as a photos, videos and more in your mobile app or website
* Know where your mobile app users go in the real world
* Build mobile audience segments for analysis, targeting and measurement

We first obtain the borough and its neighborhoods along with the postcodes from the public data. Based on the postcodes, we thus can obtain the unique latitude and longitude for each of the neighborhoods via (Google)GeoCoder. By using the Foursquare API, we can easily pull out the venues (included hotels) and its categories nearby each of the neighborhoods. After getting the location data for each of the neighborhoods in Toronto, we then manipulate the dataset and use K-mean algorithm to complete the data-driven analytics report eventually.

**Section-3**

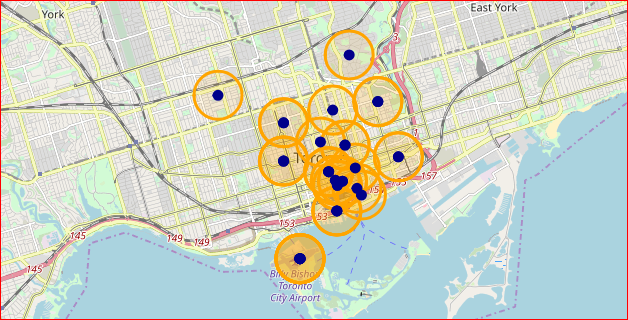
**Methodology**

We first scrape the Wikipedia webpage and extract the main information which included postal code, name of borough and its neighborhoods. We form a data frame and manipulate the data, such as to replace "Not assigned" in column 'Neighborhood' with the value in 'Borough':

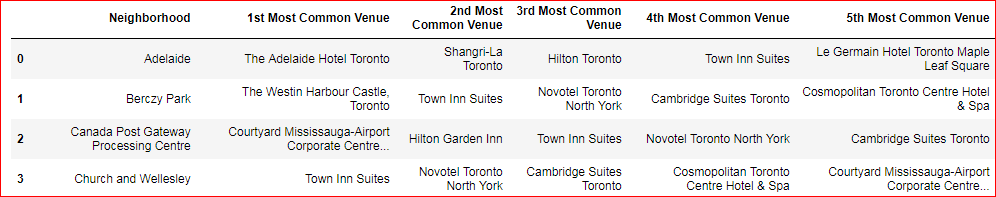


Secondary, we merge the latitude & longitude information onto the foregoing data frame using the given CSV file. Now all neighborhoods in the data frame are contained their latitude & longitude information base on their postal codes.

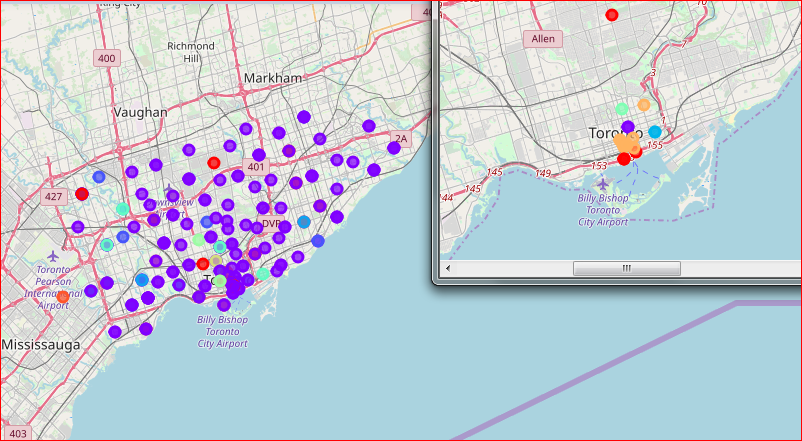
Third, we use google geocoder to map the neighborhood data from the data frame to a map so that we can visualize the data points in the area we target. We can also slice the data frame into a small portion – e.g. Downtown Toronto area:



Foursquare API is a good tool that we’ll use it to explore the neighborhoods and segment them. We define a function to help pull out the existing venues nearby (we set radius=500) each of the neighborhoods along with their latitude & longitude information and categories. Now we’re able to list the first 10 venues of most common through manipulating the data frame and the defined function of sorting the venues in descending order, as well as to list the first 5 of the most common hotels:



In order to reach our goal, we run K-mean algorithm to cluster the neighborhoods so that we’re easier to evaluate if a combination of the venues nearby is reasonable, if the density of hotel is too high.

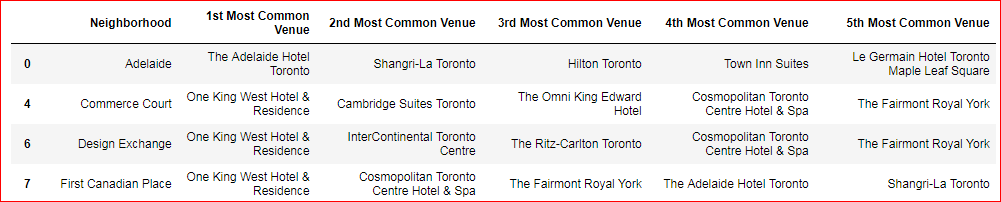


Finally we’ll optimize the neighborhood clusters and examine the clusters and make statistical inference – of which we’ll obtain a more reasonable location as our recommendation.

**Section-4**

**Results / Discussion**

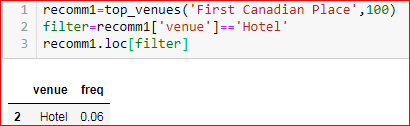
Shangri-La Hotel & the Fairmont Royal York hotel have been very successful in the hotel industry and they are ranked at 3rd and 8th among of the prestigious hotels. According to Greg Kahn’s economic theory, we find a neighborhood – First Canadian Place:



In this neighborhood, both Shangri-La Hotel & the Fairmont Royal York hotel are there and its combination of venues nearby is slightly better than the other close neighborhoods:

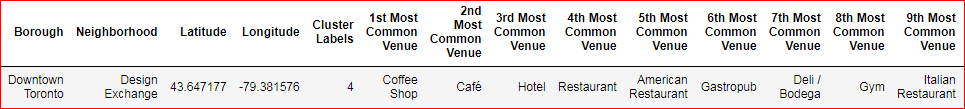


The combination in this neighborhood is reasonable for an luxury hotel and the density of hotel is also reasonable:

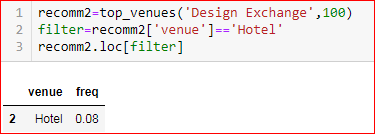


Therefore, the neighborhood of First Canadian Place is our recommendation for the project.

The other recommendations such as “Design Exchange” and “Toronto Dominion Centre”, they also have a good combination of venues nearby:



However, its density of hotel is a bit higher:



**Section-5**

**Conclusion**

Using google geocoder & Foursquare API, we can easily explore the neighborhoods and segment them and extract those venues nearby. Furthermore, we’re able to list the top N most common venues for each of the neighborhoods. With the location data, we run a machine learning algorithm – K-mean to cluster the neighborhoods, and therefore we’re easier to evaluate if a combination of the venues nearby is reasonable, if the density of hotel is too high for this project. This is to conclude the analytics report.