**Opening a New Hotel in Toronto, Canada**

Kaho Taniguchi

May 6, 2021

**1. Introduction**

In this project we are going to find a preferred location for a hotel. Especially this project is going to be targeted to stakeholders who are interested in opening a new hotel in Toronto, Canada.

Since when we think about tourism in Canada after the corona virus pandemic is over, Toronto must be the most attractive city. According to some tourism websites, Toronto is the most popular city in Canada and around 4,520,000 people visited there every year before the corona virus pandemic. In fact, 6 new luxury hotels opened in Toronto in 2020. For property developers, location of the hotel is one of the most important factors because it will determine whether the hotel will be success or not. So the business question is, 'In Toronto, if a property developer is going to open a new hotel, where would you recommend that they open it?' Today, there are lots of hotels in Toronto, so we will try to detect locations that are not already crowded with hotels. And, we are also particularly interested in areas with low hotel in vicinity.

We will use our data to generate a few most promising neighbourhoods based on this criteria. Advantages of each area will then be clearly expressed so that good location for property developers.

**2. Data**

According to definition of our problem, a factor that will influence our decision can be number of existing hotels in the neighbourhood, latitude and longitude coordinates of those neighbourhoods and venue data, particularly data related to hotels in Toronto.

So, this project is going to use following data sources to generate the required information. First, candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using Google Maps API reverse geocoding. Then, number of hotels and location in every neighbourhood will be obtained using Foursquare API (<https://foursquare.com/>). And we use a list of Postal Codes of Canada:M from Wikipedia (<https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>). Finally, we use Python Geocoder package for latitude and longitude coordinates for neighbourhoods.

**3. Methodology**

What this project is going to do on this paper is mainly 5 steps as follows:

1. Web scrape by using Python request and beautiful soup packages to get the list of neighbourhoods

2. Get the geographical coordinates of latitude and longitude using Python Geocoder package

3. Pull Hotel location data in Toronto via Foursquare API (https://foursquare.com)

4. Create the data frame of Toronto Neighbourhoods and their coordinates

5. Use K-means clustering to make groups to find a preferred place for a new hotel.

**4. Result**

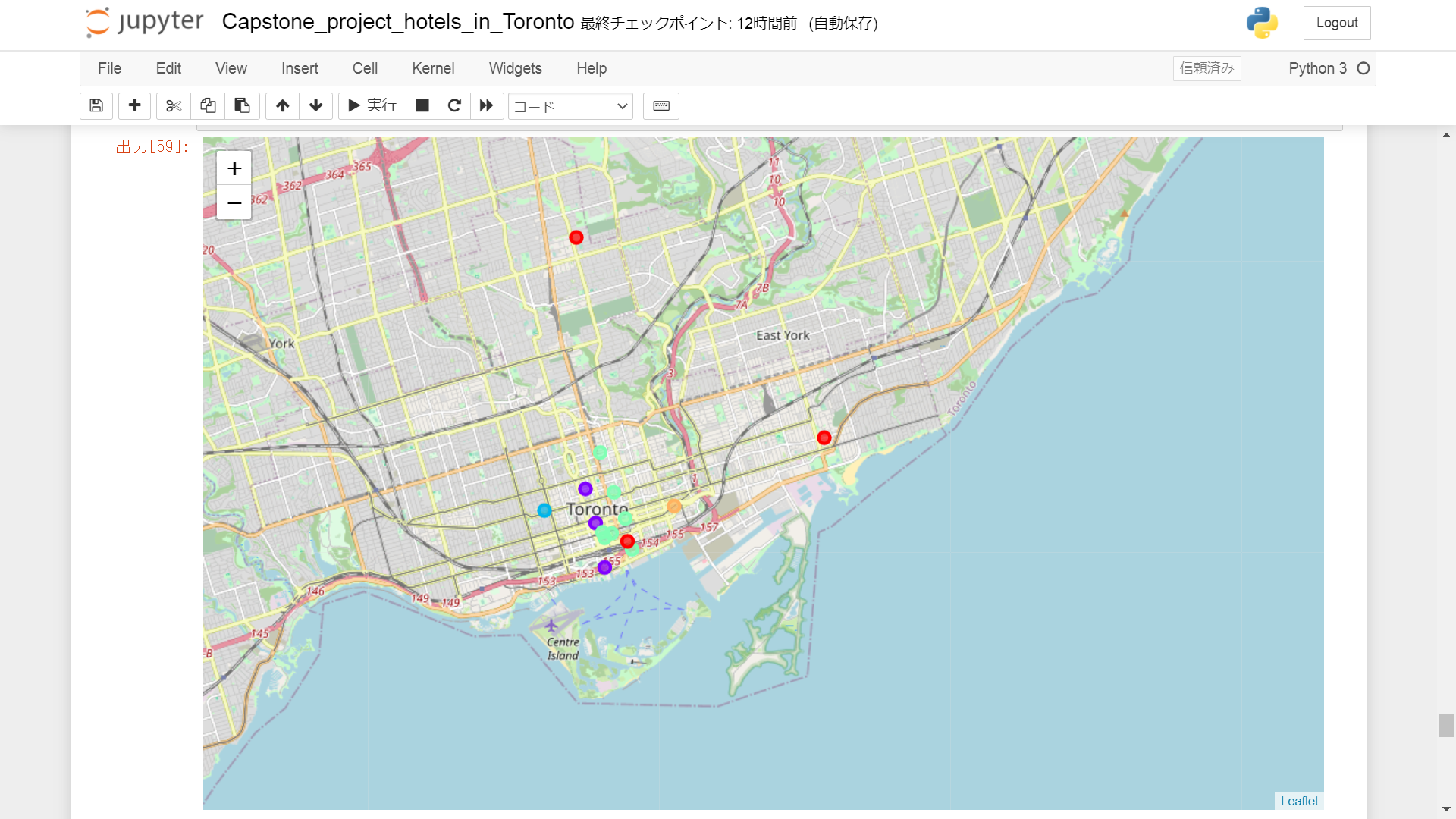
What our analysis found is as below,

Cluster 4 (green): the most crowded area

Cluster 1 (red) and Cluster 2 (purple): Neighbourhoods with moderate number of hotels

Cluster 3 and Cluster 5: Neighbourhoods with quite low number of hotels

Figure: clustering of existing hotels in Toronto



**5. Discussion**

Based on our analyse, the area grouped as cluster 3 and 5 can be recommended as a place for opening a new hotel with small number of existing hotels while the area grouped as cluster 1, 2 and 4 already have numbers of hotels.

This, of course, does not directly imply that those areas are actually the best locations for opening a new hotel. But the purpose of this analysis was to only provide information of areas where not crowded with existing hotels because it is entirely possible that there is a very good reason for small number of hotels in any of those areas, reasons which would make them unsuitable for a new hotel regardless of lack of competition in the area. Therefore recommended areas can be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

Our analysis shows that although there are many hotels in Toronto, there are still　some pockets of low hotel density. Highest concentration of hotels was detected east and central area in Toronto, so we focused our attention to downtown and west area, corresponding to boroughs Downtown Toronto.

**6. Conclusion**

Objective of this project was to identify Toronto areas with small number of hotels in order to make some help for property developer in narrowing down the search for preferred location for a new hotel. By calculating hotel density distribution with data　from Foursquare API, we identified general boroughs that justify further analysis.

Of course, final decision on optimal new hotel location will be made by property developer based on much more specific characteristics of neighbourhoods and locations in every recommended area, we can take into consideration additional factors by using data science power.