

# The Battle of Neighborhoods

Taejin Jung

March 04, 2020

## 1. Introduction

Background and business problem

In this project, I will define where is an optimal location for a restaurant. This report will be targeted to stakeholders who are interested in opening a Korean restaurant in Toronto, Canada. Various factors such as demand, type of restaurant would affect the success or failure of the restaurant. Therefore, descriptive and predictive capabilities of data science is necessary.

We need to find locations (and neighborhood) of unfulfilled demand for Korean cuisine. Also, we need to determine the location that have low competition and not overly crowded. We would also consider whether the location is close to popular city neighborhood.

By using some data science technique, this project will figure out which location is the best place to start a restaurant. The final location can be chosen by stakeholders.

## 2. Data Acquisition and cleaning

Based on definition of our problem, factors below will be considered.

- number of existing restaurants in the neighborhood (any type of restaurant)
- number of Korean restaurants and the distance in the neighborhood
- distance of neighborhood from popular neighborhoods

Therefore, in this project, I will:

- use the dataset that contains names and boroughs of the neighborhoods in Toronto that we used in the last assignment.
- After I got the names of all the neighborhoods of Toronto, I will merge latitude and longitude dataset that was provided in the last assignment.

- Next, I will use the foursquare API to find all types of restaurants within 250~500 meters radius for every neighborhood.

### 3. Methodology

In this project we will directly put our efforts on detecting areas of Toronto that have low Korean restaurant density but have high density of the other kind of restaurant.

In first glance, it seems that Korean restaurants are densely located in downtown and center of Toronto, and North York. Therefore, we are going to look for some spots in those three boroughs.

Second step in our analysis will be calculation and exploration of 'restaurant density' in three boroughs. We will use heatmaps to identify a few promising areas for Korean restaurants.

In third and final step we will focus on most promising areas by making clusters of locations that meet basic requirements: we will consider locations with no more than two Korean restaurants in radius of 250 meters, and we want locations which have the biggest number of restaurants. Since Toronto is already overly populated with restaurants, so I think that it would be useless to find locations without any restaurants. Low density of restaurants in Toronto would mean that there would be low population.

We will use K-Means cluster to identify general zones / neighborhoods / addresses which should be a starting point for final street level exploration and search for optimal venue location for stakeholders.

### 4. Result and Discussion

This analysis shows that there is a large number of restaurants in Toronto (over 2500 in Toronto). I think that there would be more candidate location because total number of venues are twice bigger than the number of restaurants. And the goal of this analysis is that whether there are some recommended places to start a Korean restaurant. Although there are more than 2500 restaurants in Toronto, this analysis shows that there are only 161 Korean restaurants. It is only 6 percent of overall restaurants.

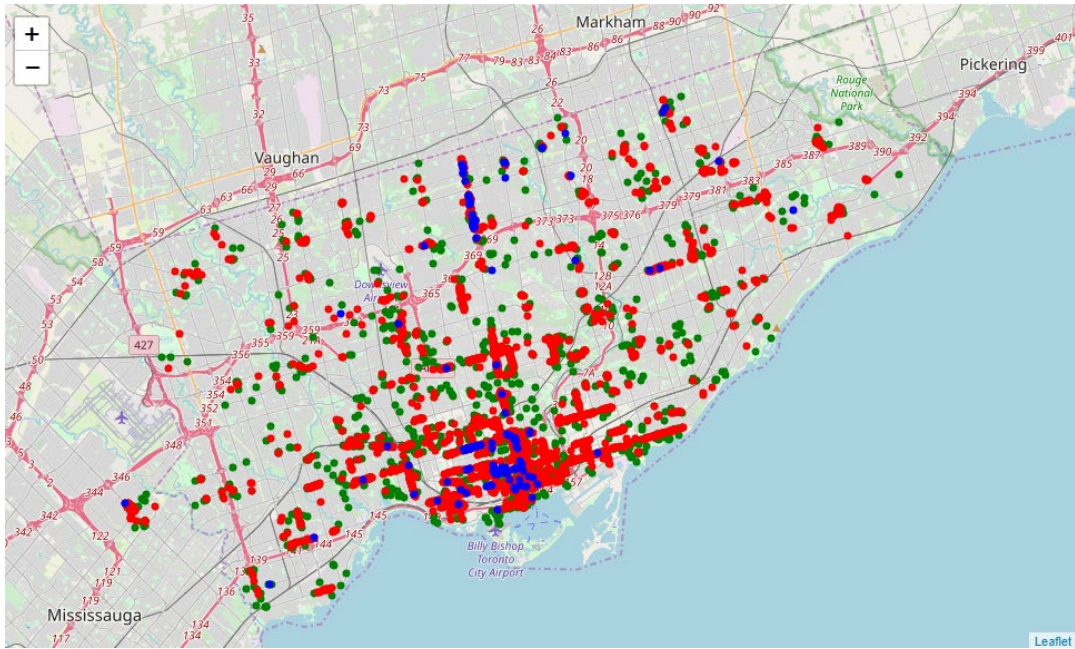


Figure 1. Map of locations of restaurants in Toronto

In this analysis, I chose recommended location as existing venues which are not restaurants because some venues are possibly in sale or try to change their business. And I chose the condition as more than 2 restaurants in radius of 250 meters, and no Korean restaurants in radius of 500 meters.

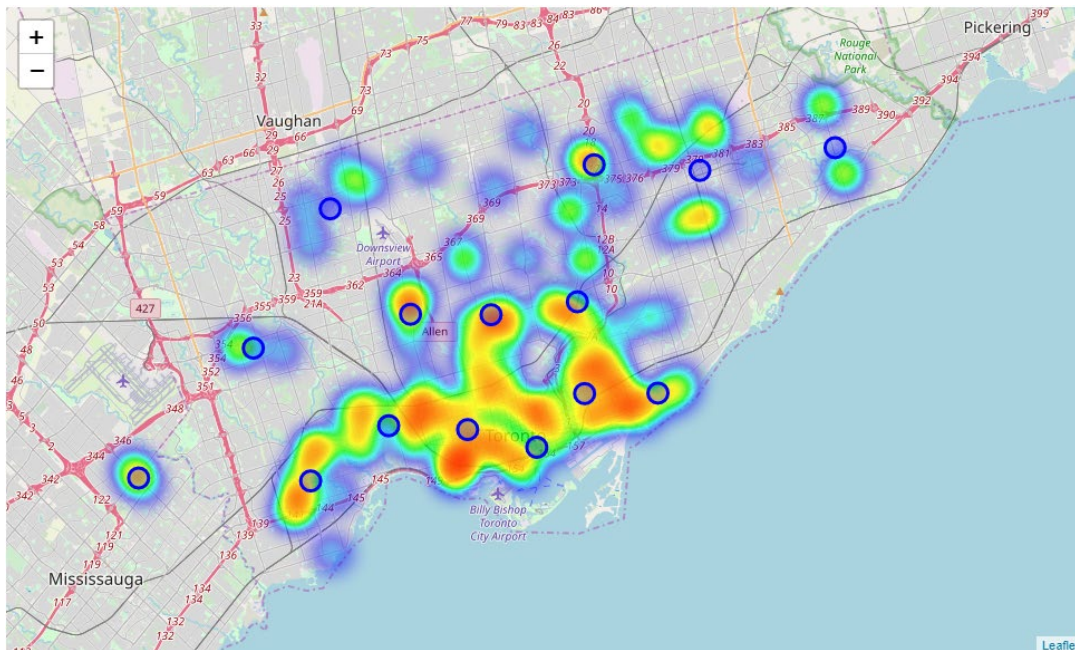


Figure 2. Map of centers of each cluster

Result of all this is 15 areas containing largest number of potential new Korean restaurant. I think that recommended zones would be considered only as a starting point for more detailed analysis which could result in location which has not only no nearby competitive place but also other factors should be taken into account. Also, the relevant conditions of stakeholders should be considered.

## **5. Conclusion**

This concludes our analysis. We have created 15 addresses which represent centers of candidate locations. I think Toronto is already overly populated place, so if we set the location that there is no restaurant nearby, it would be the place where less popular than others. Therefore, in my opinion, the candidate locations have more than 2 restaurants in radius of 250 meters which are not Korean cuisine, and no Korean restaurants in radius of 500 meters.

Final decision on optimal restaurant location will be made by stakeholders. It would be considered based on specific characteristics of neighborhoods and locations. Also, level of noise, real estate prices, and economic dynamics of neighborhoods.