

# Capstone Project – Locations for opening Chinese restaurant in Toronto

## 1. Introductions and the Problem

In this project we will try to find optimal location to open a Chinese restaurant. This report will be targeted to stakeholders interested in open such restaurant in Toronto, Canada. We will try to detect locations that are not already crowded with Chinese restaurant and locations that are highly populated. We would also prefer locations as close to downtown as possible, assuming that the first two conditions met.

## 2. Data acquisition and cleaning

2.1 Based on definition of our problem, factors that will influence our decision are:

- Number of existing restaurants
- Number of existing Chinese restaurants
- Population of neighborhood
- Distance of neighborhood from downtown

2.2 Following data sources will be needed to extract/generate the required information:

- Centers of initial candidate areas will be found from [Wiki page](#) and geographical information can be found from this [File](#)
- Population of Toronto can be gathered from this [webpage](#), this is population information for the year 2016. With the limitation of data online, we will use it as a reference even though it's not the most recent data
- Number of Chinese restaurants in every neighborhood will be obtained using Foursquare API
- Coordinate of Toronto downtown has been found through a simple Google search (43.6548, -79.3883)

2.3 Data cleaning

We will use postcode area as the candidate area in the initial analyzing level; the reason is our population data was generated according to postcode. We are using the population data as a factor to narrow down areas.

We will generate a table that contains Postcode, Latitude, Longitude, and Population. Then use the geographic information to find venues in Toronto, filter out all restaurants and Chinese restaurants. We can complete this step using Foursquare.

Here is the first 5 rows of the restaurants table:

	Area	Area Latitude	Area Longitude	Venue Name	Venue Latitude	Venue Longitude	Venue Category	Venue Id
0	M1B	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant	4bf58dd8d48988d16e941735
1	M1T	43.781638	-79.304302	KFC	43.780400	-79.300700	Fast Food Restaurant	4bf58dd8d48988d16e941735
2	M1W	43.799525	-79.318389	KFC	43.798938	-79.318854	Fast Food Restaurant	4bf58dd8d48988d16e941735
3	M1W	43.799525	-79.318389	McDonald's	43.798249	-79.318167	Fast Food Restaurant	4bf58dd8d48988d16e941735
4	M2J	43.778517	-79.346556	KFC	43.777600	-79.344200	Fast Food Restaurant	4bf58dd8d48988d16e941735

## 3. Methodology

In this project we will direct our efforts on detecting areas of Toronto that have low restaurant density, particularly those with low number of Chinese restaurants.

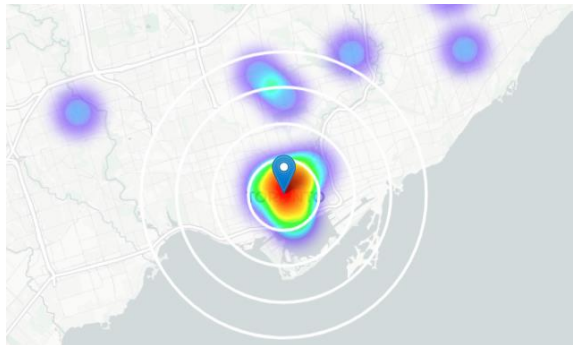
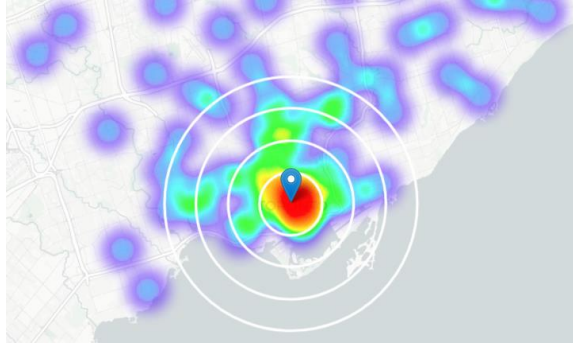
In first step we have collected the required data: location and type (category) of every restaurant in Toronto. We have also identified Chinese restaurants (according to Foursquare categorization).

Second step in our analysis will be calculation and exploration of 'restaurant density' across different areas of Toronto - we will use heatmaps to identify a few promising areas close to center with low number of restaurants in general (and no Chinese restaurants in vicinity) and focus our attention on those areas.

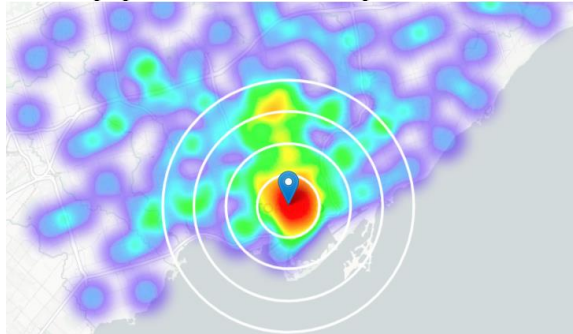
In third and final step we will focus on most promising areas and within those create clusters of locations that meet some basic requirements established in discussion with stakeholders: we will take into consideration locations with no more than two restaurants in radius of 500 meters, and we want locations without Chinese restaurants in radius of 800 meters. We will present map of all such locations but also create clusters (using k-means clustering) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

#### 4. Analysis

##### 4.1 check density of all restaurants (top) and Chinese restaurants (bottom):



##### 4.2 review population in a heatmap:



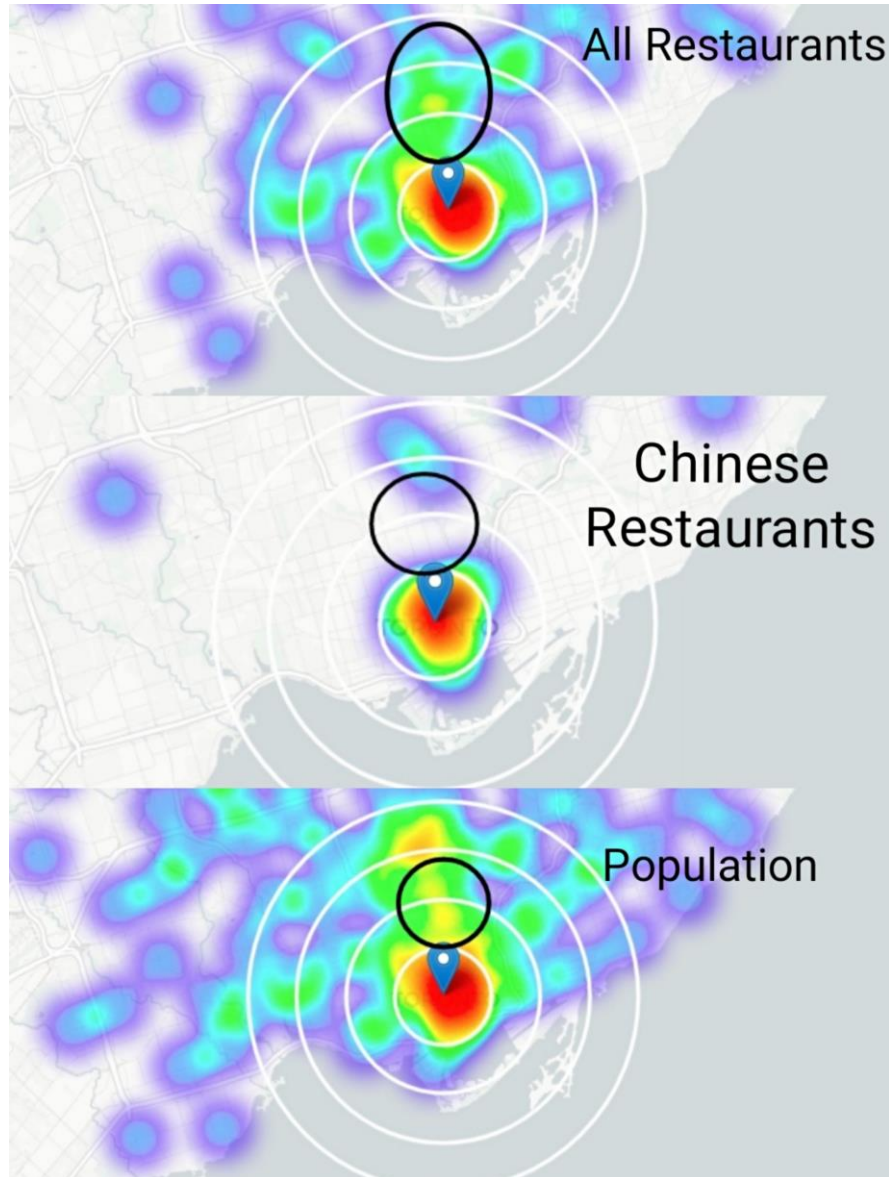
4.3 From the heatmaps above, we can tell that at about 2~6km north from downtown, there is an area (black circles) that:

4.3.1 the Chinese restaurants density is very low (no color on heatmap for Chinese Restaurants)

4.3.2 it's quite populated (yellow color in population heatmap)

4.3.3 the restaurants density is low (green color in All Restaurants heatmap)

So we will further review this area.



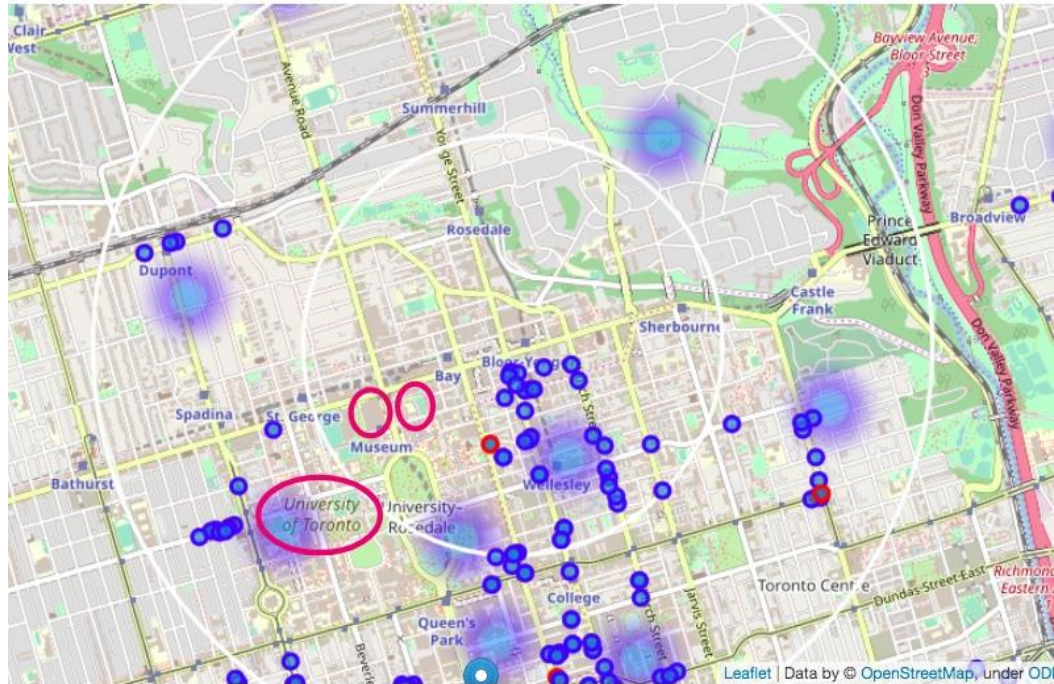
4.4 Zoom in to this area

We can find that the nearby area has fairly equal population (the light blue to purple color 'bubble' represents the population and are all in similar color). We could also find that University of Toronto, Royal Ontario Museum, and Gardiner Museum are in this area (red circles below). From Wikipedia, we could easily find that the total enrollment for U of T is 61,690(2019). Which means there are a lot of students who will come to campus and would potentially need food services. And the 2 museums would possibly bring tourists to the area.

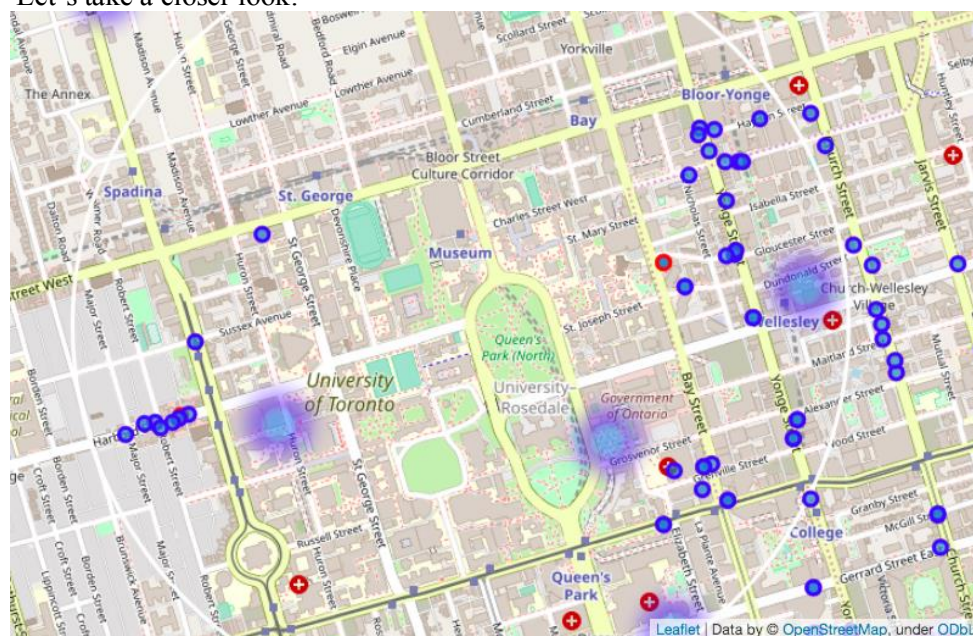


According to [timeout.com](https://www.timeout.com/toronto), Royal Ontario Museum is one of the best Museums in Toronto. "This museum's expansive collection of cultural and historic artifacts makes it a must-see. Stop by the Royal Ontario Museum to learn about everything from art of the First Peoples to modern fashion to the age of dinosaurs. The museum has an ever-revolving schedule of exhibitions and events, including the summer-long Friday Night Live, which transforms the galleries into a party with live DJs, food and drinks."

And just right across the street, there is the Gardiner Museum. "All things clay are on display at this museum dedicated to ceramics. Get your hands dirty at one of the Gardiner Museum's regular classes in hand building, wheel throwing and slip casting. If you'd prefer to leave the messy stuff to the experts, spend some time admiring the collection of some 4,000 pieces from the ancient Americas to today." - [timeout.com](https://www.timeout.com/toronto)



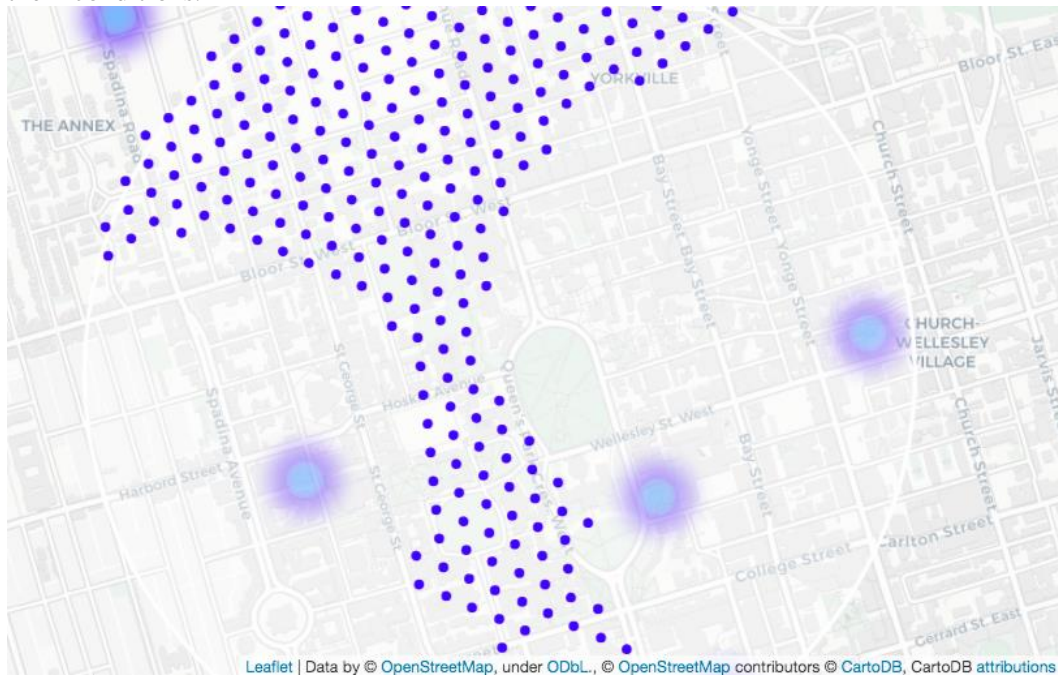
Let's take a closer look:



#### 4.5 Define new candidate areas

Now we'd like to create denser grid of location candidates restricted to our new region of interest -- the area in white circle above (let's make our location candidates 100m apart). There are 817 areas created in the area.

We calculate two most important things for each location candidate: number of restaurants in vicinity (we'll use radius of 500 meters) and distance to closest Chinese restaurant. Our goal is to find the location candidates that contain less than 2 restaurants and no Chinese restaurant within 800 meters. We found 256 location candidates that met the 2 conditions.



#### 4.6 Clustering

Cluster those locations to 15 cluster, and create centers of zones containing good locations, and the centers of the clusters will be the final result to present to stakeholders.

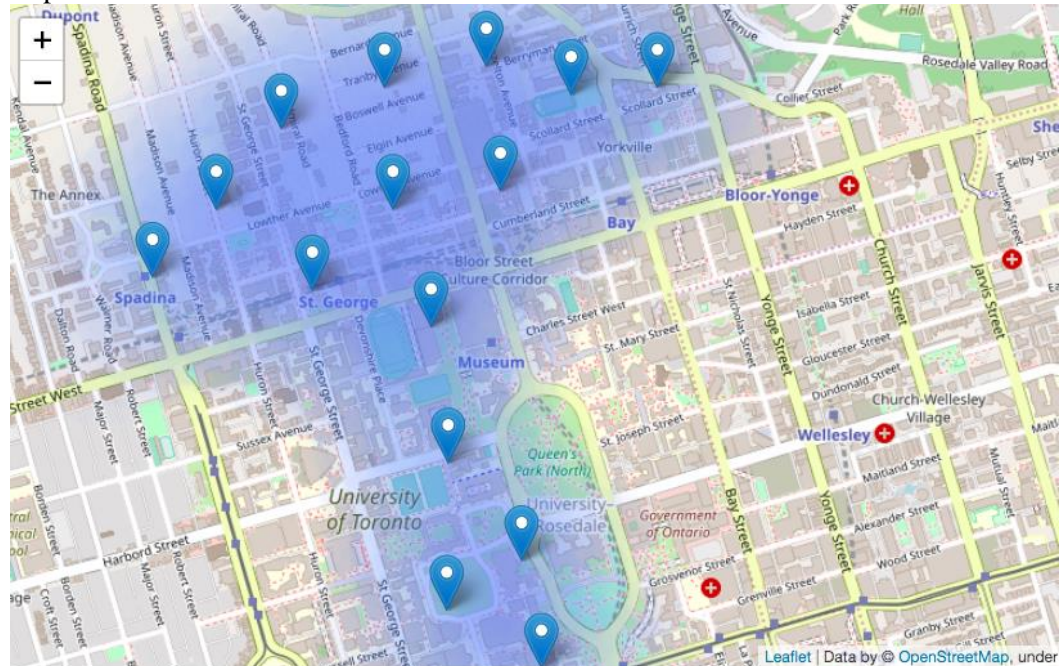




Addresses of centers of areas recommended for further analysis:

565 Huron St, Toronto, Ontario, M5R 2R6	=> 2.9km from Downtown, Toronto
13 Hoskin Ave, Toronto, Ontario, M5S	=> 1.8km from Downtown, Toronto
76 Hazelton Ave, Toronto, Ontario, M5R 2E2	=> 3.0km from Downtown, Toronto
160 College St, Toronto, Ontario, M5S 3E1	=> 1.0km from Downtown, Toronto
139 St George St, Toronto, Ontario, M5R 2L8	=> 2.5km from Downtown, Toronto
Amaya Express	=> 2.9km from Downtown, Toronto
34 Boswell Ave, Toronto, Ontario, M5R 1M4	=> 3.0km from Downtown, Toronto
23 Bedford Rd, Toronto, Ontario, M5R 2J9	=> 2.6km from Downtown, Toronto
58-98 King's College Cir, Toronto, Ontario, M5S	=> 1.3km from Downtown, Toronto
273 Bloor St W, Toronto, Ontario, M5S 1W2	=> 2.2km from Downtown, Toronto
135 Yorkville Ave, Toronto, Ontario, M5R 0C7	=> 2.6km from Downtown, Toronto
Queen's Park Cres W, Toronto, Ontario, M5S	=> 1.3km from Downtown, Toronto
38 Admiral Rd, Toronto, Ontario, M5R 2L5	=> 3.1km from Downtown, Toronto
39	
Spadina Rd, Toronto, Ontario, M5R 2S9	=> 2.9km from Downtown, Toronto
88 Scollard St, Toronto, Ontario, M5R 1G2	=> 2.9km from Downtown, Toronto

Map of the recommended addresses:



## 5. Results & Discussion

Our analysis shows that although there are a great number of restaurants in Toronto (more than 2000), there are pockets of low restaurant density fairly close to Downtown. Highest concentration of restaurants was detected at the Downtown area and about 6 to 8km north from Downtown, the population spread looks pretty much the same as the restaurants density, but it shows that the population in the area 2 to 6km north from Downtown is kind high. We notice that area (2~6km north from Downtown) is populated but not crowded with restaurants, so we focused our attention to that area. There are parks, university, Museums in that area, our attention was focused on area close to U of T, Royal Ontario Museum, and Gardiner Museum which offer a combination of popularity among students and tourists, strong socio-economic dynamics and a number of pockets of low restaurant density.

After directing our attention to this more narrow area of interest, we first created a dense grid of location candidates (spaced 100m apart); those locations were then filtered so that those with more than two restaurants in radius of 500m and those with a Chinese restaurant closer than 800m were removed.

Those location candidates were then clustered to create zones of interest, which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Result of all this is 15 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Chinese restaurants particularly. This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to Downtown, Toronto but not crowded with existing restaurants (particularly Chinese) - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore 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.

## **6. Conclusion**

Purpose of this project was to identify Toronto areas close to Downtown with low number of restaurants (particularly Chinese restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Chinese restaurant. By calculating restaurant density distribution from Foursquare data we have first identified general boroughs that justify further analysis (an area north from Downtown), and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.