

# **Open a New Chinese Restaurant in Toronto**

# Background and Problem

## Background

- Toronto is the most populous city in Canada
- Immigrant-friendly province
- More and more population

## Problem

- Where is the best place to open a new Chinese restaurant with the most profitable and least competitive?

# Data Acquisition and Cleaning

## Data Acquisition

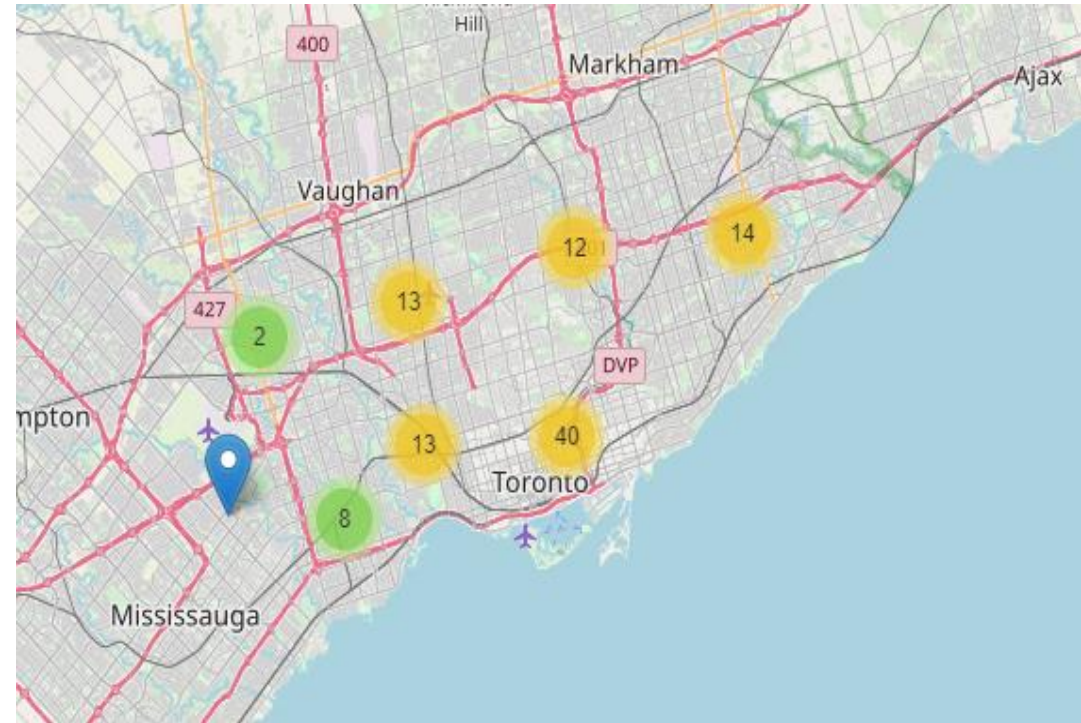
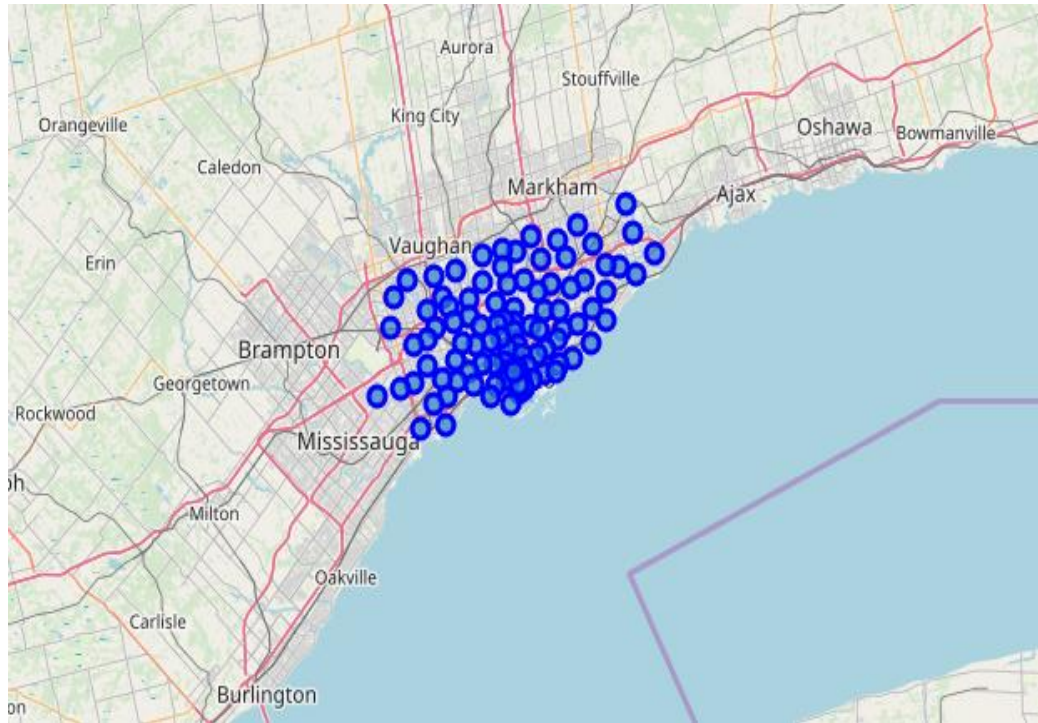
- “List of Postal code of Canada” table from Wikipedia
- Geographical coordinates from csv file
- Distribution of population by ethnicity in Toronto from Wikipedia
- Foursquare’s API to explore the venues

## **Data Cleaning**

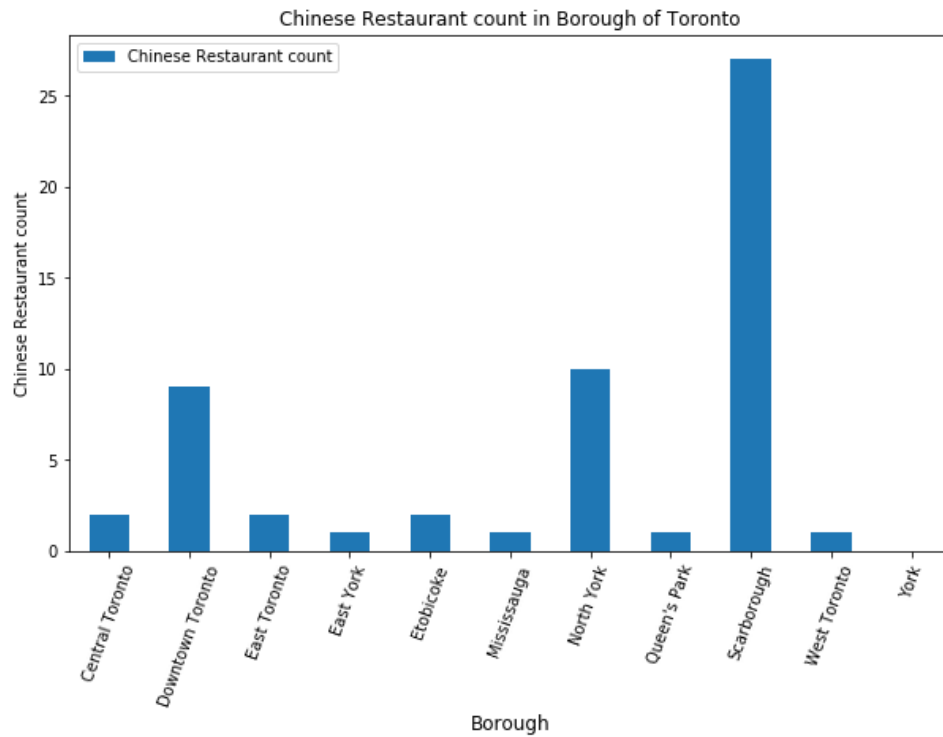
- Clean the postal code table with “not assigned” value
- Combine the postal code table with geographical coordinates table in one table
- Transform the distribution data into consolidation form
- Find the venues by Foursquare’s API

# Exploratory Data Analysis

## Distribution of Neighborhood in Toronto



# The Distribution of Chinese Restaurant

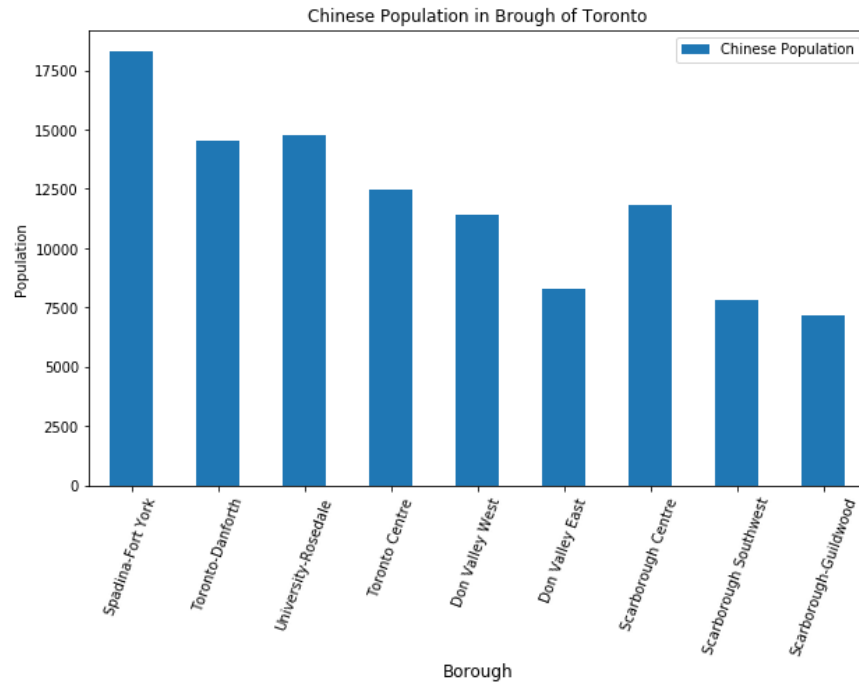


## 3 Boroughs with most Chinese restaurant

- Downtown Toronto
- North York
- Scarborough

# Chinese Population and Borough

## Top 3 population of Chinese

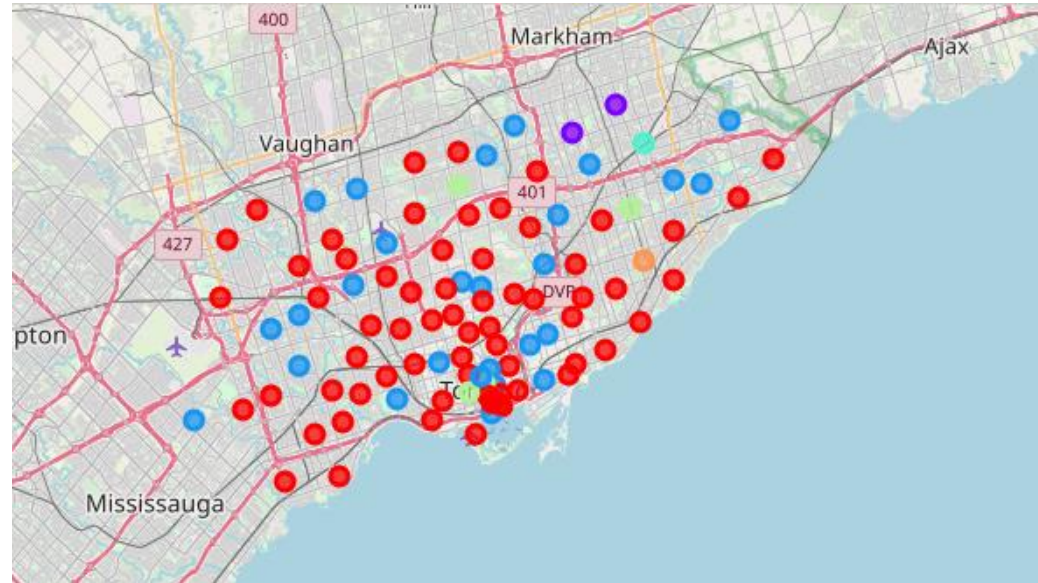
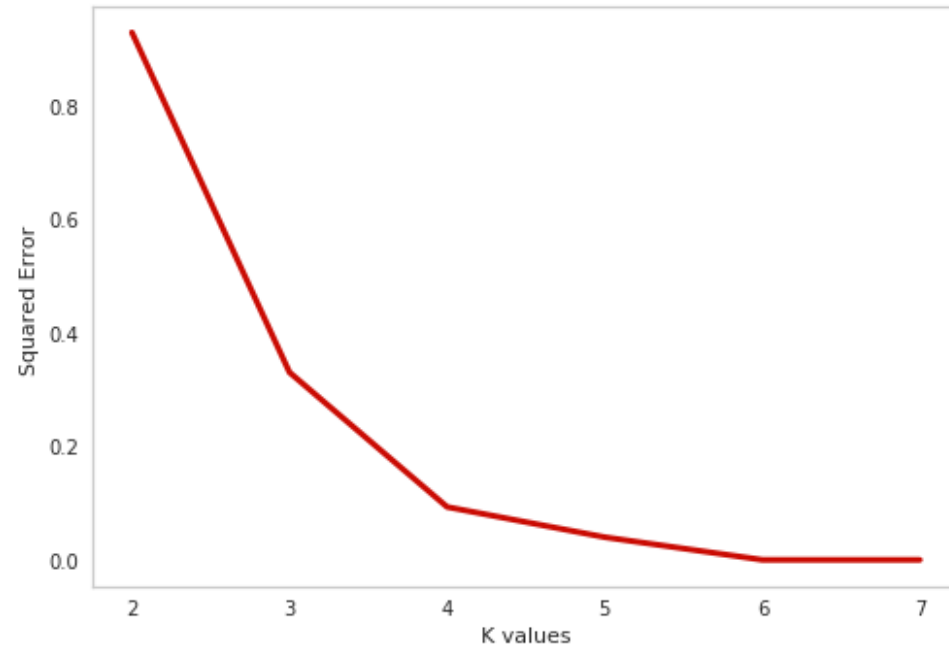


- Spadina-Fort York
- Toronto Danforth
- University-Rosedale

# Predictive Modelling

## Clustering

Find  $K = 6$ , then cluster the dataset from 0 to 5





# Examine the Clusters

## Cluster 0

```
print(df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 0].shape)
df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 0]
```

(67, 7)

i]:

	Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	Chinese Restaurant
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497	0.0	0.0
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711	0.0	0.0
5	M1J	Scarborough	Scarborough Village	43.744734	-79.239476	0.0	0.0
7	M1L	Scarborough	Clairlea, Golden Mile, Oakridge	43.711112	-79.284577	0.0	0.0
8	M1M	Scarborough	Cliffcrest, Cliffside, Scarborough Village West	43.716316	-79.239476	0.0	0.0
9	M1N	Scarborough	Birch Cliff, Cliffside West	43.692657	-79.264848	0.0	0.0
11	M1R	Scarborough	Maryvale, Wexford	43.750072	-79.295849	0.0	0.0
17	M2J	North York	Fairview, Henry Farm, Oriole	43.778517	-79.346556	0.0	0.0
19	M2L	North York	Silver Hills, York Mills	43.757490	-79.374714	0.0	0.0
20	M2M	North York	Newtonbrook, Willowdale	43.789053	-79.408493	0.0	0.0

# Cluster 1

```
print(df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 1].shape)
df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 1]
```

(2, 7)

3]:

	Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	Chinese Restaurant
14	M1V	Scarborough	Agincourt North, L'Amoreaux East, Milliken, St...	43.815252	-79.284577	1.0	5.0
15	M1W	Scarborough	L'Amoreaux West	43.799525	-79.318389	1.0	5.0

# Cluster 2

```
print(df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 2].shape)
df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 2]
```

(27, 7)

]:

	Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	Chinese Restaurant
0	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353	2.0	1.0
3	M1G	Scarborough	Woburn	43.770992	-79.216917	2.0	1.0
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476	2.0	1.0
13	M1T	Scarborough	Clarks Corners, Sullivan, Tam O'Shanter	43.781638	-79.304302	2.0	1.0
16	M2H	North York	Hillcrest Village	43.803762	-79.363452	2.0	1.0
18	M2K	North York	Bayview Village	43.786947	-79.385975	2.0	1.0
24	M3A	North York	Parkwoods	43.753259	-79.329656	2.0	1.0
26	M3C	North York	Flemington Park, Don Mills South	43.725900	-79.340923	2.0	1.0
28	M3J	North York	Northwood Park, York University	43.767980	-79.487262	2.0	1.0

## Cluster 3

```
print(df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 3].shape)
df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 3]
```

(1, 7)

]:

	Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	Chinese Restaurant
12	M1S	Scarborough	Agincourt	43.7942	-79.262029	3.0	7.0

## Cluster 4

```
: print(df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 4].shape)
df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 4]
```

(4, 7)

0]:

	Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	Chinese Restaurant
10	M1P	Scarborough	Dorset Park, Scarborough Town Centre, Wexford ...	43.757410	-79.273304	4.0	2.0
21	M2N	North York	Willowdale South	43.770120	-79.408493	4.0	2.0
56	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383	4.0	2.0
66	M5T	Downtown Toronto	Chinatown, Grange Park, Kensington Market	43.653206	-79.400049	4.0	2.0

## Cluster 5

```
print(df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 0].shape)
df_new_cn_rest.loc[df_new_cn_rest['Cluster Labels'] == 5]
```

(67, 7)

.]:

	Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	Chinese Restaurant
6	M1K	Scarborough	East Birchmount Park, Ionview, Kennedy Park	43.727929	-79.262029	5.0	4.0

# Conclusion and Future Direction

## **Conclusion**

- Harbourfront is the best place
- Located on the south of Downtown Toronto in few kms
- With high population of Chinese and least competition

## **Future Direction**

- Use data from recent years since the data use in this project is based on 2016 census
- Use multiple API to find more accuracy venues information