

[Final Assignment] The Battle of Neighborhoods

**Project to Find the Optimal Number of the
Toronto Financial Security Centers (TFSCs)**

21 July 2020

Junyup Kim

1. Introduction

According to the Major Crime Indicators (MCI) data of Toronto, robberies in financial institutions occurred 644 times in total from 2014 to 2019. This indicates that there are around 107 robberies on average in financial firms in Toronto per year. As one of the efforts to enhance the position of Toronto as a leading global business hub, let's assume that the Canadian Bankers Association (CBA) decides to lower the number of bank robberies as 50% of the current average level until 2025. As the first step to achieve this goal, CBA will support to establish a Toronto Financial Security Center (TFSC). This report will try to find **the optimal number of TFSCs by the clustering analysis**.

2. Data

The bank robbery case data from the Major Crime Indicators (MCI) database¹ will be used to perform the location clustering. Because this data includes the location information (i.e., longitude, latitude), it is possible to cluster the locations of crimes that occurred.

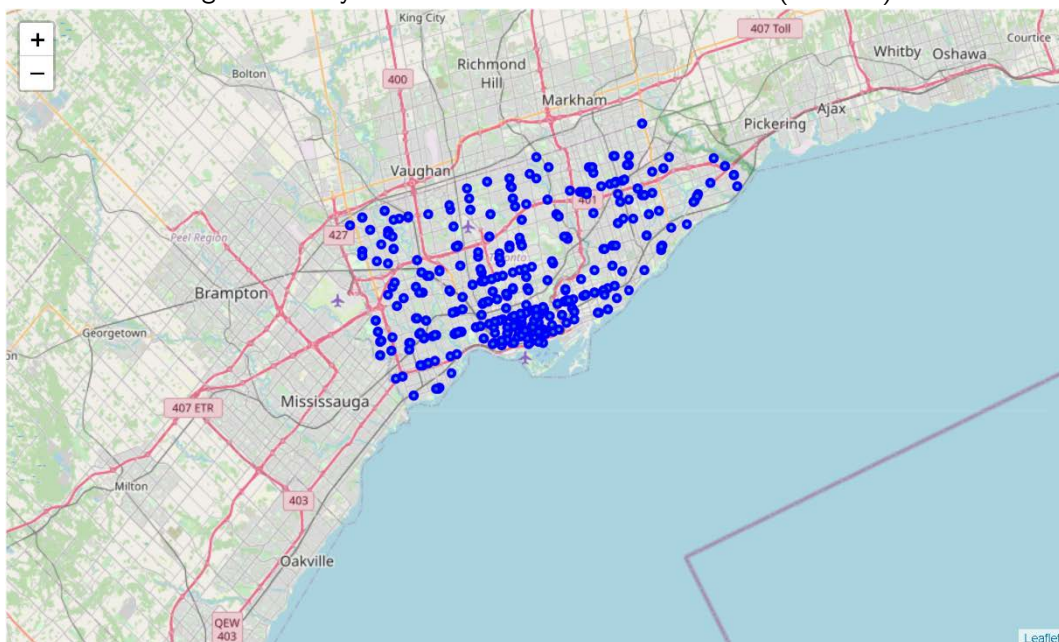
Meanwhile, the location information of banks in Toronto will be collected from the Foursquare location data. This will help to find bank clusters in Toronto and compare the location clusters of robbery in financial companies.

3. Methodology

[1] Location Clustering of the Robbery in Financial Institutions in Toronto

Before implementing K-means cluster analyses, it is necessary to visualize the locations that bank robberies occurred during the period between 2014 and 2019 using the longitude and latitude information and the Folium tool of Python. The result is as follows:

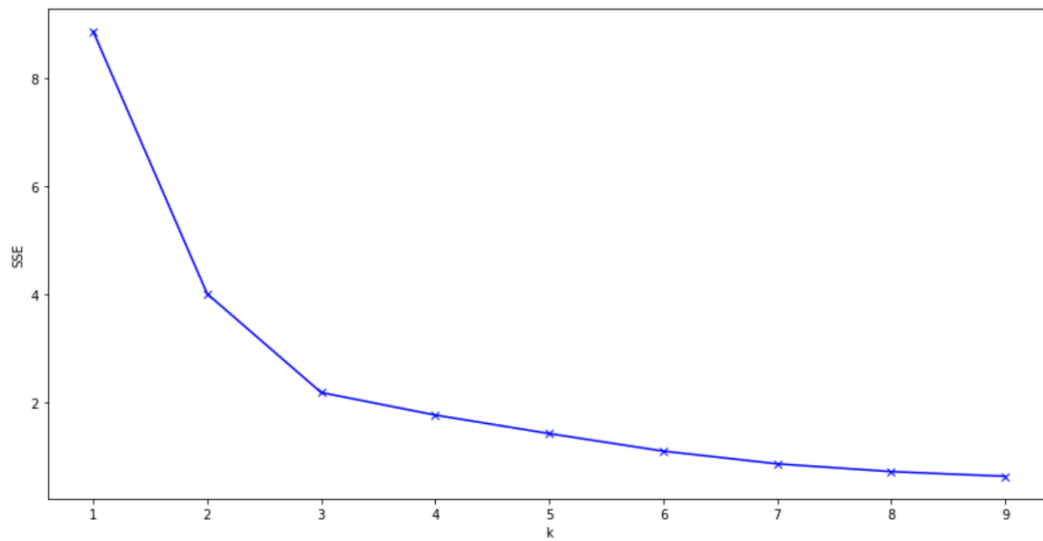
Fig 1. Robbery in Financial Institutions in Toronto (2014-19)



1. Toronto Police Service. Public Safety Data Portal (<https://data.torontopolice.on.ca/datasets/mci-2014-to-2019?geometry=-80.424%2C43.545%2C-78.338%2C43.893>, accessed: 18 July 2020).

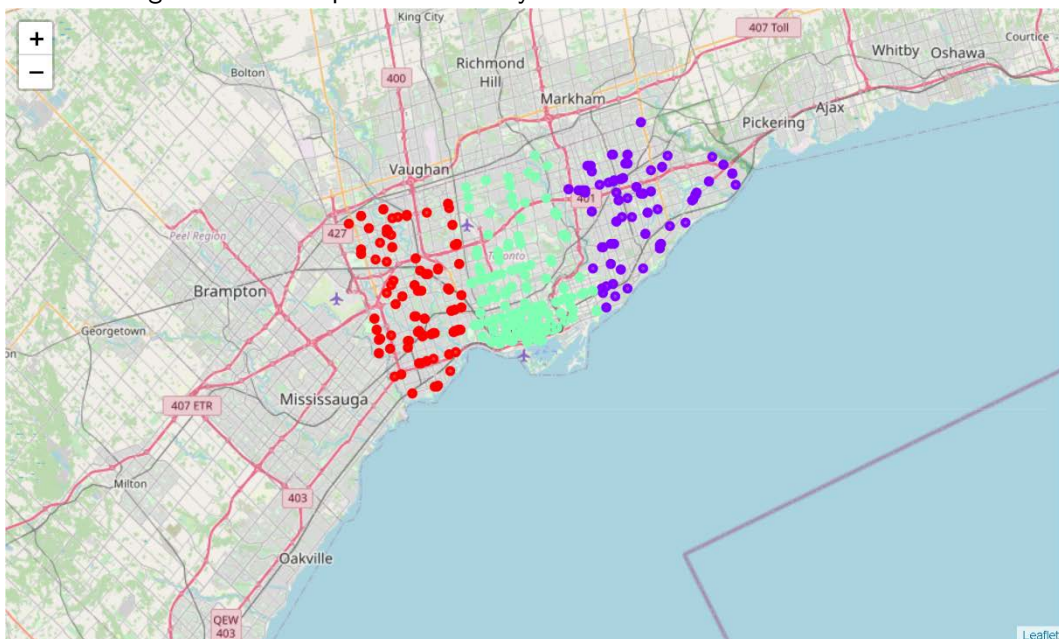
It is necessary to find the optimal number of clusters to perform clustering, According to the elbow method, the optimal number of clusters is three (i.e., $K = 3$) as shown in Fig 2.

Fig 2. The Elbow Method to Determine the Number of Optimal Clusters (K) for the Robbery in Financial Institutions in Toronto



The clustering result when $K = 3$ is described using the Folium tool in Fig. 3.

Fig 3. Cluster Map of the Robbery in Financial Institutions in Toronto



[2] Location Clustering of the Banks in Toronto

Ahead of building a bank location dataframe, it is necessary to make a dataframe having latitude and longitude coordinates based on the postal codes, including each borough and neighborhood in Toronto (Fig. 4).

Fig.4 Location Information Based on the Postal Code of Toronto

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
5	M9A	Etobicoke	Islington Avenue, Humber Valley Village	43.667856	-79.532242
6	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
7	M3B	North York	Don Mills	43.745906	-79.352188
8	M4B	East York	Parkview Hill, Woodbine Gardens	43.706397	-79.309937
9	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
10	M6B	North York	Glencairn	43.709577	-79.445073
...					
95	M1X	Scarborough	Upper Rouge	43.836125	-79.205636
96	M4X	Downtown Toronto	St. James Town, Cabbagetown	43.667967	-79.367675
97	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280
98	M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North	43.653654	-79.506944
99	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
100	M7Y	East Toronto	Business reply mail Processing Centre, South C...	43.662744	-79.321558
101	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...	43.636258	-79.498509
102	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999

To find the bank location in Toronto, the exploration of venues in Toronto was performed first using the Foursquare API. The result includes every venue information in Toronto, including a bank. That is why it is necessary to extract the bank location information and make a new dataframe. The first five rows of the new dataframe is shown in Fig. 5.

Fig 5. Location Information of Banks in Toronto

	Postal Code	Borough	Neighbourhood	Latitude	Longitude	Bank
47	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	1
84	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	1
154	M4B	East York	Parkview Hill, Woodbine Gardens	43.706397	-79.309937	1
225	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	1
375	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711	1

As similar to the case of the bank crimes, locations of the banks in Toronto are shown in Fig 6. Also, the bank clustering result is visualized in Fig 7. The optimal number of clusters was set as three (i.e., $K = 3$) by applying the elbow method (Fig 8).

Fig 6. Banks in Toronto

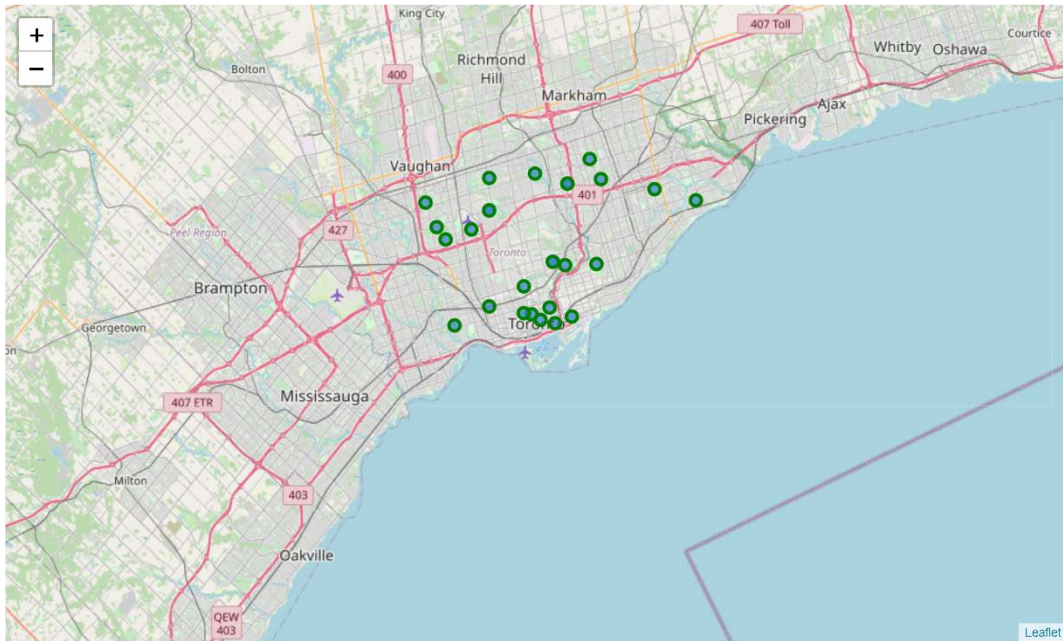


Fig 7. Cluster Map of the Banks in Toronto

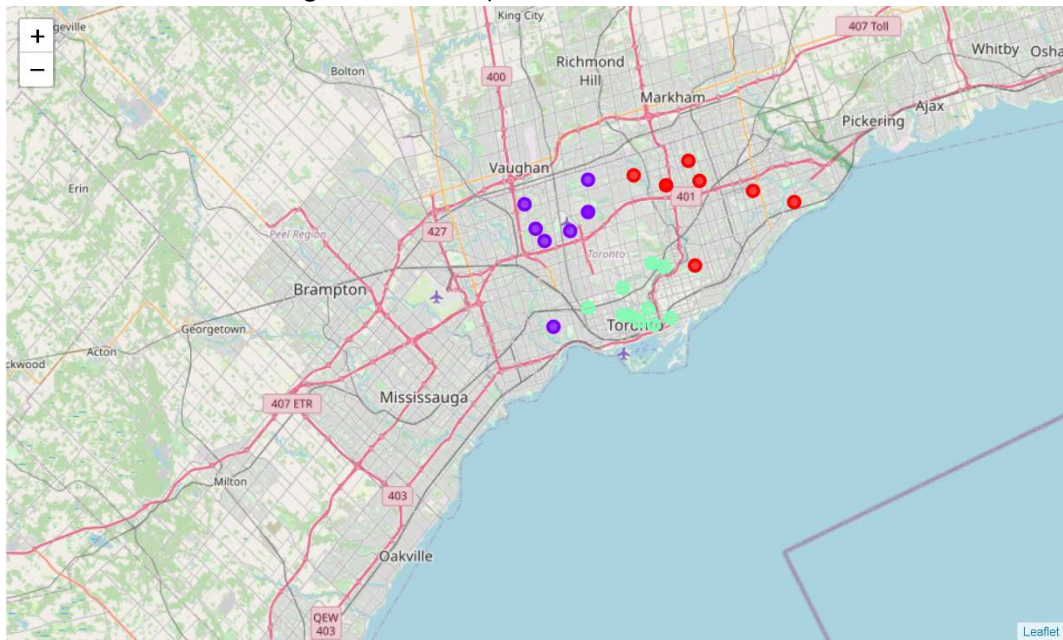
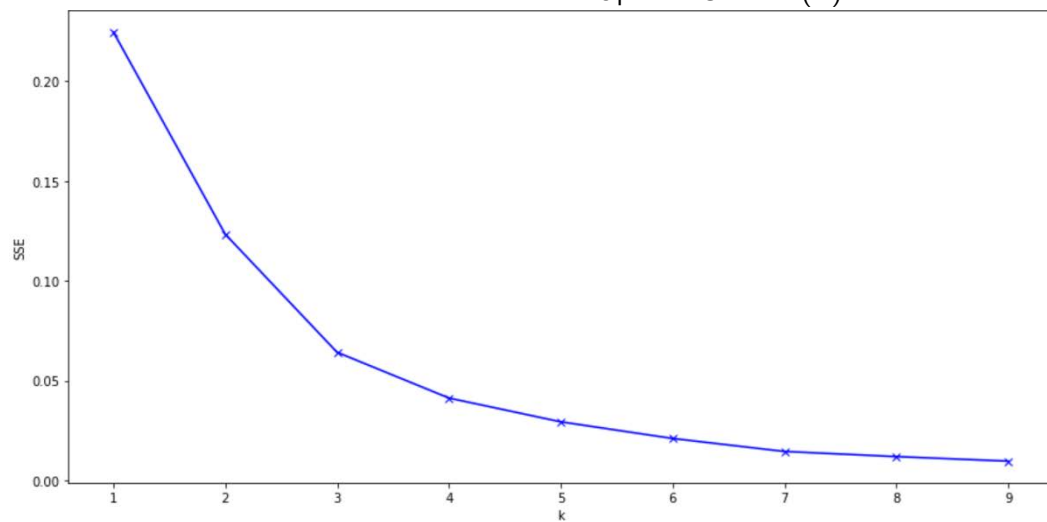


Fig 8. The Elbow Method to Determine the Number of Optimal Clusters (K) for the Banks in Toronto



4. Result and Discussion

According to the clustering analyses, there are three clusters in bank robberies and bank locations in Toronto. This implies that **at least three TFSCs in Toronto need to be established if there let be one TFSC in each cluster.**

The detailed location of each TFSC can simply be the centroid of each cluster. However, it is necessary to regard more factors such as transportation, population, and existing public security enforcement (e.g., police stations). Moreover, it seems that the Foursquare database does not keep the full location information for the complete list of financial institutions in Toronto. These should be supplemented through further research

5. Conclusion

To conclude, it can be necessary to have three TFSC in Toronto, given that there are three clusters both in bank locations and robbery cases in financial institutions.