The Battle of the Neighborhoods - Location Decisions of Tim Horton and Second Cup in Toronto

Introduction

Background

Coffee is one of the most widely consumed beverages in the World and in Canada. According to Statista, Canadians drank on average 2.7 cups of caffeinated beverages per day in 2020. In the past decade, coffee shops have become a popular location to work, hang out with friends, or get a quick lunch. Tim Hortons, the biggest coffee shop chains in Canada, own 4,286 stores across the country.

Problem

In this report, I analyze the store location of Tim Hortons and Second Cup, the two biggest coffee franchises in Toronto.

Data acquisition, cleaning and analysis

Data description

I require three datasets in the research, namely the locations of Tim Hortons, the locations of Second Cups and a dataset of the coordinates of all Toronto neighborhoods.

Data sources

I retrieve the location of all Tim Hortons and Second Cup in Toronto using FourSquare API.

Data collection

I search the Tim Hortons and Second Cup stores within a radius of 5,000 meters of 103 neighborhoods of Greater Toronto, i.e. the 6 boroughs including Toronto, York, North York, East York, Scarborough and Etobicoke. Each request gets a maximum of 50 venues' information; hence I build two datasets, each having 5,150 Tim Hortons or Second Cup stores.

location.distance	location.formattedAddress	location.labeledLatLngs	location.lat	location.lng	location.neighborhood	location.postalCode	location.state	name
1254	['2050 Victoria Park Ave (btwn Cassandra Blvd....	[{'label': 'display', 'lat': 43.75281395751706...	43.752814	-79.314067	NaN	M1R 1V2	ON	Tim Hortons
866	['215 Brookbanks (York Mills Rd)', 'Toronto ON...	[{'label': 'display', 'lat': 43.76066827030228...	43.760668	-79.326368	NaN	M3A 1Z5	ON	Tim Hortons
1558	['1244 Lawrence Ave East (Curlew)', 'North Yor...	[{'label': 'display', 'lat': 43.74157865703848...	43.741579	-79.318966	NaN	M3A 1B9	ON	Tim Hortons
1492	['Lawrence', 'Toronto ON']	[{'label': 'display', 'lat': 43.74044697317023...	43.740447	-79.324169	NaN	NaN	ON	Tim Hortons / Esso
2170	['2501 Victoria Park Ave. (at Farmcrest Dr)', ...	[{'label': 'display', 'lat': 43.77156824953461...	43.771568	-79.320392	NaN	M1T 1A1	ON	Tim Hortons
ocation.distance	location.formattedAddress	location.labeledLatLngs	location.lat	location.lng	location.neighborhood	location.postalCode	location.state	name
1217	['Humbertown Plaza', 'Etobicoke ON']	[{'label': 'display', 'lat': 43.66227280880637...	43.662273	-79.519234	NaN	NaN	ON	Second Cup
2621	['3300 Bloor St W (Sun Life Centre)', 'Toronto...	[{'label': 'display', 'lat': 43.64533056349981...	43.645331	-79.522753	NaN	M8X 2W8	ON	Second Cup
1640	['265 Wincott Dr', 'Toronto ON M9R 2R7']	[{'label': 'display', 'lat': 43.67947508876911...	43.679475	-79.544773	NaN	M9R 2R7	ON	Second Cup
2829	['3008A Bloor Street West (Royal York)', 'Toro...	[{'label': 'display', 'lat': 43.64773783074008...	43.647738	-79.510773	NaN	M8X 1C2	ON	Second Cup
1641	['265 Wincott Drive', 'Toronto ON M9R 2R7']	[{'label': 'display', 'lat': 43.67951856654155...	43.679519	-79.544710	NaN	M9R 2R7	ON	Second Cup Coffee Co. featuring Pinkberry Froz...

Data cleaning

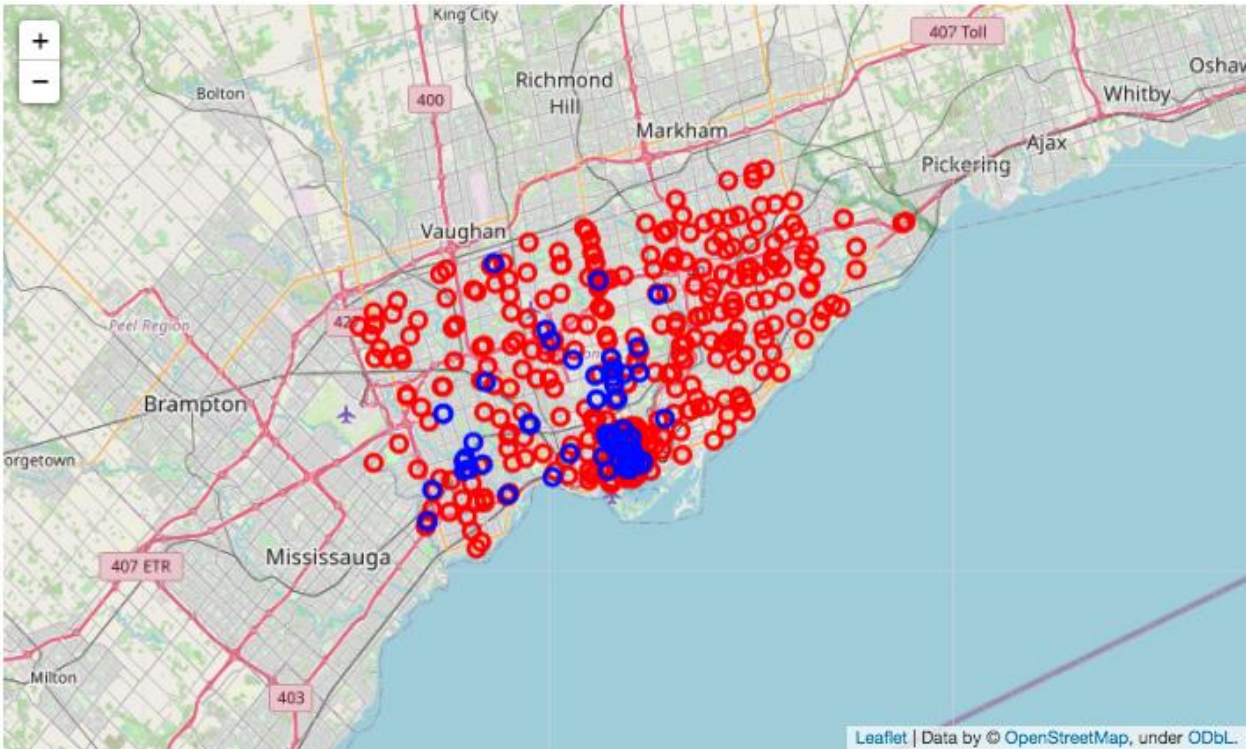
A large number of duplicates are observed as the 25,000-square-meter search area may overlap. I observed

I first select the stores located in Greater Toronto. Each store has a unique id, with which I detect and remove the duplicates in the dataframes. As a result, our data size reduces to 385 and 67 rows.

In the next step, I fill in the missing values in the columns of neighborhood and postal codes. I obtain the address of a given coordinate with Geopy package, then extract the postal codes with re package. I then find the corresponding neighborhood from the Toronto neighborhood dataset.

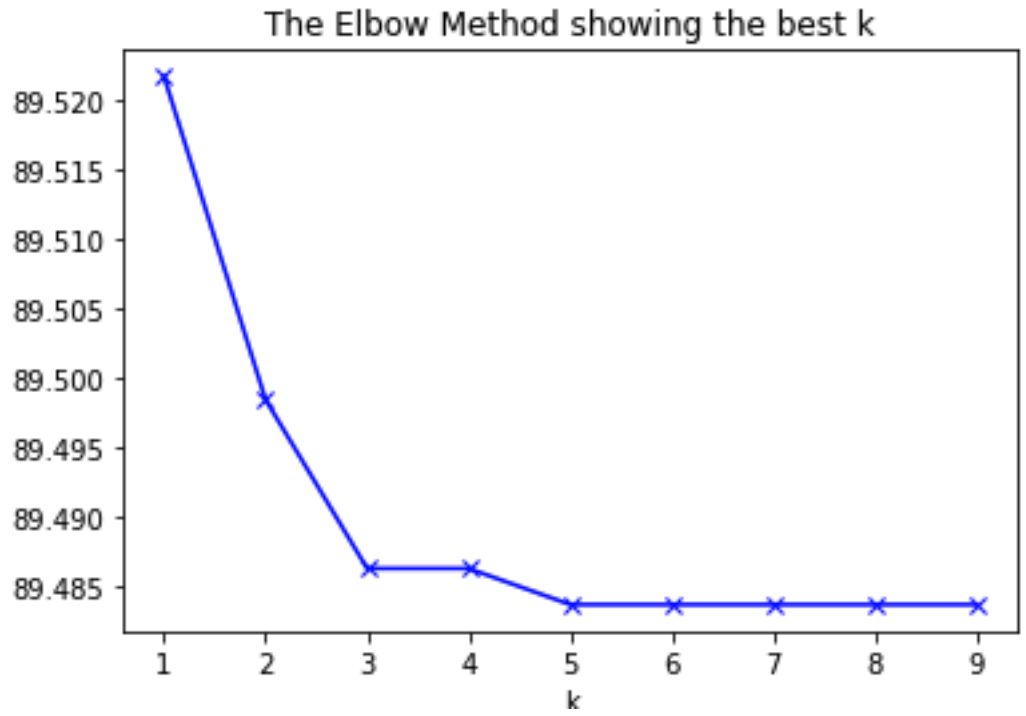
Data analysis

Tim Hortons owns 385 stores and Second Cup 67 stores in Toronto. As I draw the map, I observe that both chain focus on downtown area, hoIver, Tim Hortons covers a larger area in the city.

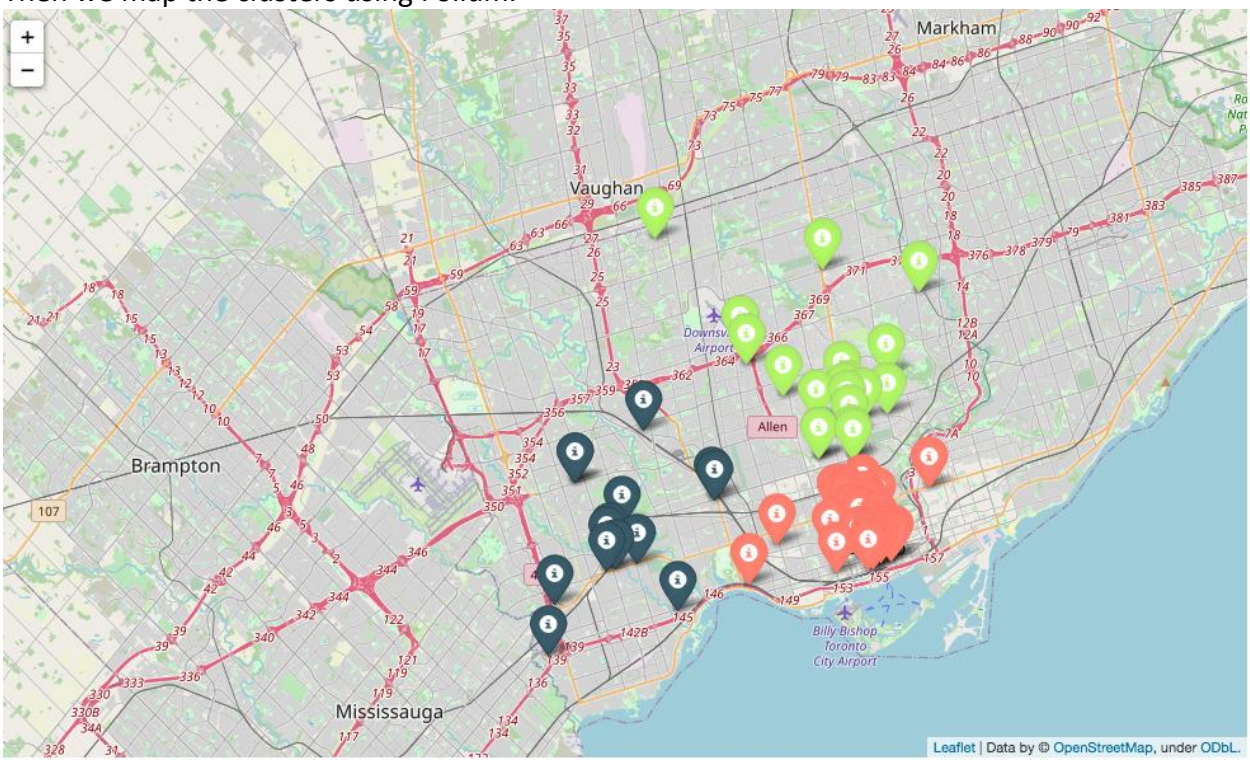


Methodology and Results

I practiced clustering algorithm on two datasets – Tim Hortons and Second Cup locations in Toronto. For each dataset, I calculate the best number of clusters, with which I label the stores of different clusters.



Then we map the clusters using Folium.



Discussion

We observe that Second Cup has no store in the north west of Toronto. It will be more interesting to combine the demographic data of the neighborhoods to understand the two chains' location decisions.

Conclusion

We collect data of Tim Hortons and Second Cup store locations in the municipality of Toronto. We observe the clusters of both chains' store locations, and find that Tim Hortons have a much larger network covering the city. The locations shall be mainly decided by the population including the visitors.

Thank you for your patience reading this. This submission is mainly for the purpose to practice the codes provided in the course materials. I literally got lost my objective but had to use clustering. So here it is ☺