Selecting the Best Forward Sortation Area (FSA) for a New Restaurant in Toronto

Nassim Al-Abed

July 12, 2019

1. Introduction

Normally site selection is a task that carried out by Geographic Information Systems (GIS) software after specifying several criteria. GIS is very efficient in carrying out, such as task if the needed spatial data and metadata are available. For this project, the use of specialized GIS package is not advisable because we need to apply the knowledge gained in the courses of data science taken during the previous months in solving a site selection problem. Normally starting a new business and opening a new restaurant involves investing time in studying the market and calculating the risk associated with the restaurant site selection. Utilizing the free tools in data science in analysing the freely available data either from the internet or from governmental sources is advantageous.

1.1 Description of City of Toronto:

Toronto is the provincial capital of Ontario and the most populous city in Canada, with a population of 2,731,571 in 2016. Toronto census metropolitan area (CMA) has a population of 5,928,040, making it Canada's most populous metropolitan area. Toronto is the fastest growing city in North America and is the anchor of an urban agglomeration, known as the Golden Horseshoe in Southern Ontario, located on the northwestern shore of Lake Ontario. Toronto is an international center of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world. The total area of Toronto is 630.2 km^2 (https://en.wikipedia.org/wiki/Toronto). Figure 1 shows the city's location and its boundaries. The diverse population of Toronto reflects its current and historical role as an important destination for immigrants to Canada, where above 50 percent of residents belong to a visible minority population group, and over 200 distinct ethnic origins are represented among its inhabitants. While the majority of Toronto residents speak English as their primary language, over 160 languages are spoken in the city. So, we can easily say that Toronto is a global city. In 2016, the three most commonly reported ethnic origins overall were Chinese (332,830 or 12.5%), English (331,890 or 12.3%) and Canadian (323,175 or 12.0%), (https://en.wikipedia.org/wiki/Toronto).

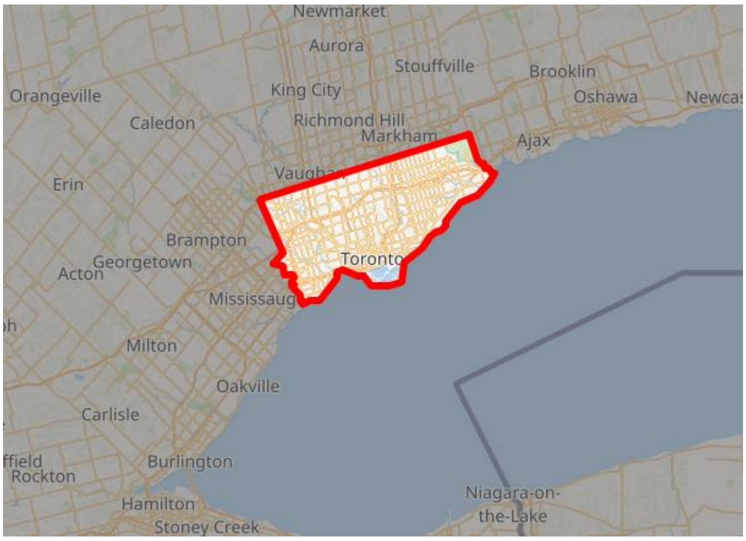


Figure 1. City of Toronto location and boundaries

2.2 Identifying the problem

First, not all of the needed data for this project are readily available, so I had to obtain the raw data and work on cleaning the data and converting the data to the format that I could use in this project. After cleaning and preparing the data Folium could be used in creating a choropleth map that can help us in finding a suitable Forward Sortation Area (FSA) for the new restaurant. The major task in this project was to select the best FSA for a new Chines restaurant based on two criterions, which will be outlined in the coming section.

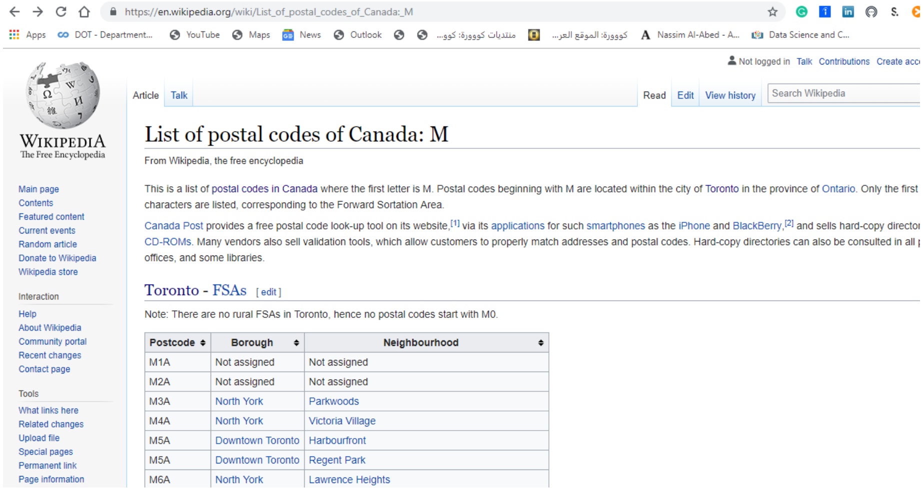
2.3. Identifying the site selection criteria

There are several factors that we could consider in carrying out an exercise similar to this project, but again selecting the approach and identifying the factors that will be considered in the selection criteria depends on the spatial data availability. What I found in this project that spatial data is not freely available for Toronto and utilizing some of the open data involved working on converting the format to the needed format for Folium using a GIS package as will be shown later. In this project, I selected two criteria to be considered for the selected FSA of the new Chines restaurant first one is to have at least 10% of the residents either Chines immigrants or from a Chines decedents, this was considered by assuming that they will favour the Chines cousin. The second criterion considered was the neighbourhood should have the highest population density in Toronto. This criterion was made to make sure that we will have enough potential customers from the same area.

2. Data acquisition and cleaning

I utilized the data from two sources for this project. The first source was the data coming from Wikipedia, where I used BeautifulSoup 4 library to scrape the information from the following site (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M).

The scraped information related to Toronto's neighbourhoods', boroughs, and postal codes. Figure 2 shows an image of the site with the available data.



List of postal codes of Canada: M

From Wikipedia, the free encyclopedia

This is a list of postal codes in Canada where the first letter is M. Postal codes beginning with M are located within the city of Toronto in the province of Ontario. Only the first characters are listed, corresponding to the Forward Sortation Area.

Canada Post provides a free postal code look-up tool on its website,[1] via its applications for such smartphones as the iPhone and BlackBerry,[2] and sells hard-copy directo CD-ROMs. Many vendors also sell validation tools, which allow customers to properly match addresses and postal codes. Hard-copy directories can also be consulted in all offices, and some libraries.

Toronto - FSAs [edit]

Note: There are no rural FSAs in Toronto, hence no postal codes start with M0.

Postcode	Borough	Neighbourhood
M1A	Not assigned	Not assigned
M2A	Not assigned	Not assigned
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Harbourfront
M5A	Downtown Toronto	Regent Park
M6A	North York	Lawrence Heights

Figure 2. The Wikipedia website for Toronto's districts, neighbourhoods and postal codes

During the process of cleaning the data scraped from Wikipedia website, any cell has a borough but a not assigned neighbourhood, and then the neighbourhood will be the same as the borough. I used the Geopy library to get the Latitude and Longitude values of Toronto. I used folium library to build a map for Toronto with the neighbourhoods superimposed on top of the map. Folium library builds on the data wrangling strengths of the Python ecosystem and the mapping strengths of the leaflet.js library. Figure three shows the neighbourhoods superimposed on top of Toronto's map.

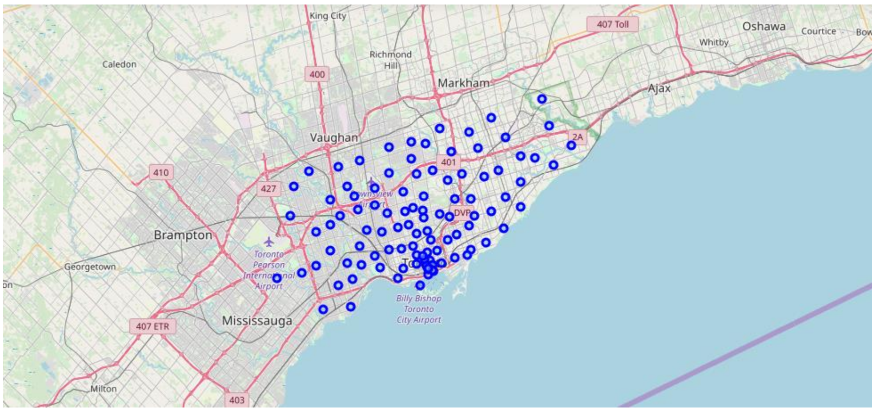
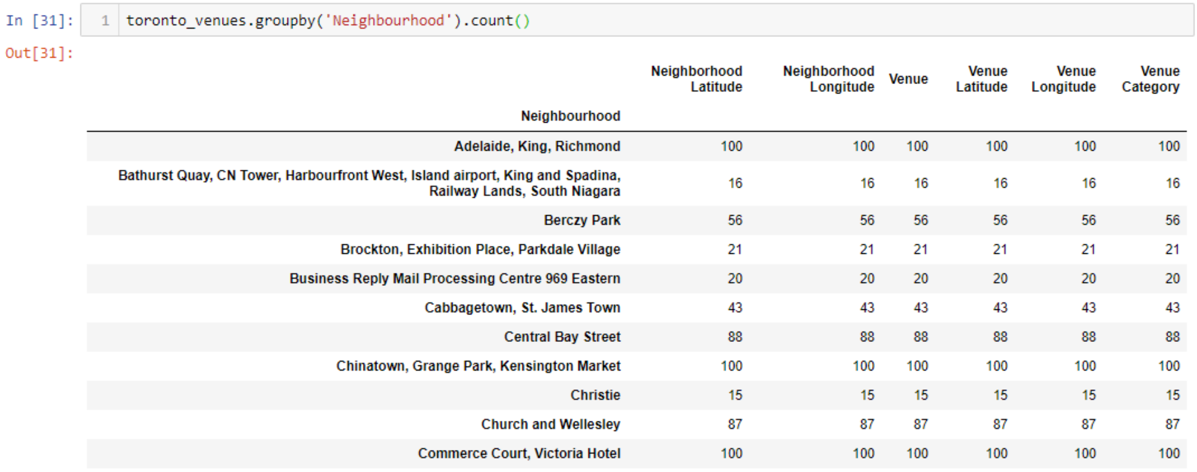


Figure 3. Toronto's neighbourhoods superimposed on top of the map

Methodology

Foursquare APIs were used to explore the neighbourhoods and to segment them based on the top ten venues. Foursquare is a social location service that allows users to explore the world around them. The venues were counted for each Borough as shown in figure 4.

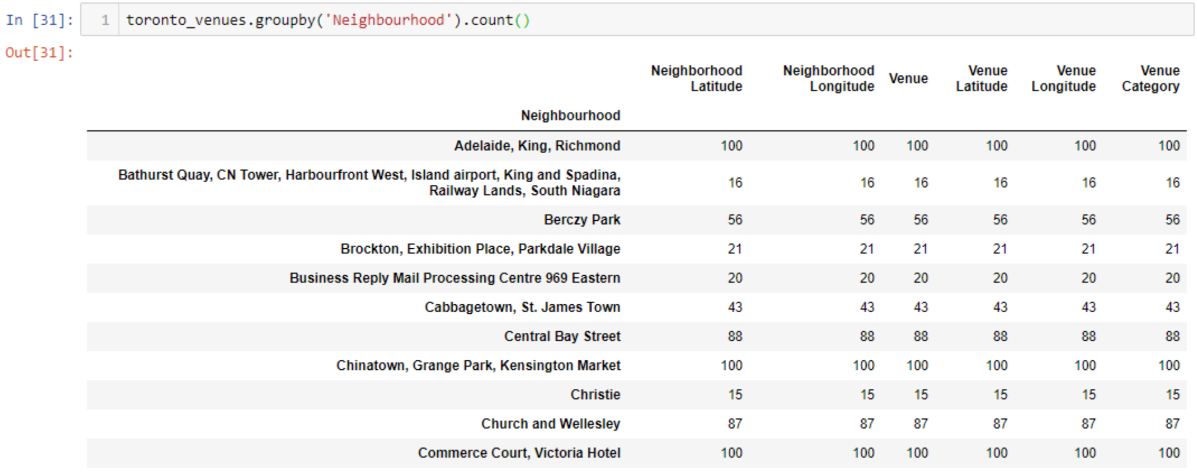


```
In [31]: toronto_venues.groupby('Neighbourhood').count()
Out[31]:
```

Neighbourhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Adelaide, King, Richmond	100	100	100	100	100	100
Bathurst Quay, CN Tower, Harbourfront West, Island airport, King and Spadina, Railway Lands, South Niagara	16	16	16	16	16	16
Berczy Park	56	56	56	56	56	56
Brockton, Exhibition Place, Parkdale Village	21	21	21	21	21	21
Business Reply Mail Processing Centre 969 Eastern	20	20	20	20	20	20
Cabbagetown, St. James Town	43	43	43	43	43	43
Central Bay Street	88	88	88	88	88	88
Chinatown, Grange Park, Kensington Market	100	100	100	100	100	100
Christie	15	15	15	15	15	15
Church and Wellesley	87	87	87	87	87	87
Commerce Court, Victoria Hotel	100	100	100	100	100	100

Figure 4. The venues counted for each Borough

Then the top five most common venues in each neighbourhood were found. k-means cluster was run to cluster the neighbourhood into 10 clusters as shown in figure 5.



```
In [31]: toronto_venues.groupby('Neighbourhood').count()
Out[31]:
```

Neighbourhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Adelaide, King, Richmond	100	100	100	100	100	100
Bathurst Quay, CN Tower, Harbourfront West, Island airport, King and Spadina, Railway Lands, South Niagara	16	16	16	16	16	16
Berczy Park	56	56	56	56	56	56
Brockton, Exhibition Place, Parkdale Village	21	21	21	21	21	21
Business Reply Mail Processing Centre 969 Eastern	20	20	20	20	20	20
Cabbagetown, St. James Town	43	43	43	43	43	43
Central Bay Street	88	88	88	88	88	88
Chinatown, Grange Park, Kensington Market	100	100	100	100	100	100
Christie	15	15	15	15	15	15
Church and Wellesley	87	87	87	87	87	87
Commerce Court, Victoria Hotel	100	100	100	100	100	100

Figure 5. k-means cluster for the venues

Figure 6 shows the visualization of the resulting clusters for Toronto.

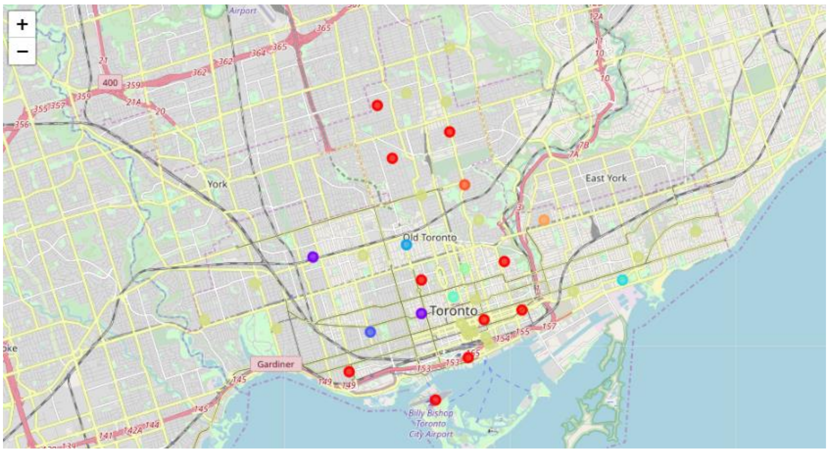
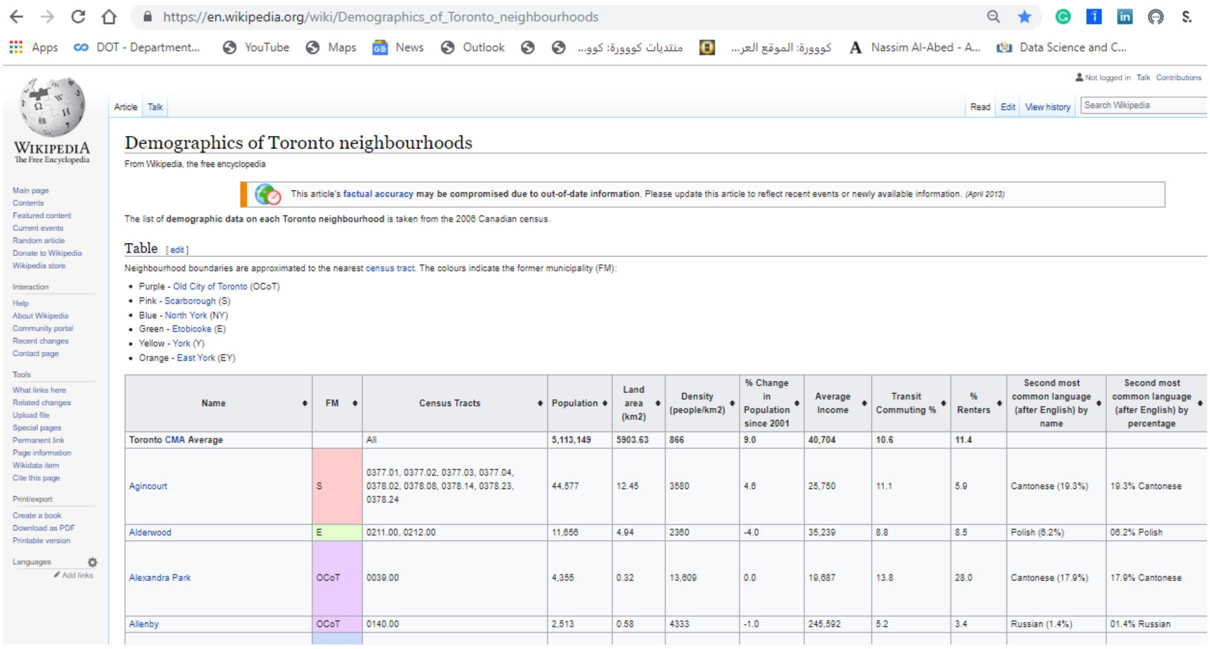


Figure 6. k-means cluster visualization of the resulting Toronto’s clusters

Demographic data analysis

Also, I scraped Toronto’s demographic information from the following page utilizing BeautifulSoup: (https://en.wikipedia.org/wiki/Demographics_of_Toronto_neighbourhoods).

Figure 7 shows a snapshot of the site and its content.



Demographics of Toronto neighbourhoods

From Wikipedia, the free encyclopedia

This article's factual accuracy may be compromised due to out-of-date information. Please update this article to reflect recent events or newly available information. (April 2013)

The list of demographic data on each Toronto neighbourhood is taken from the 2006 Canadian census.

Table [edit]

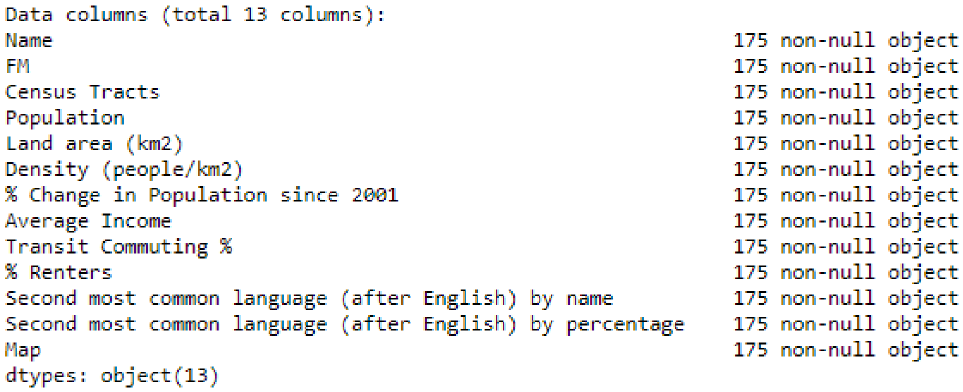
Neighbourhood boundaries are approximated to the nearest census tract. The colours indicate the former municipality (FM):

- Purple - Old City of Toronto (OCoT)
- Pink - Scarborough (S)
- Blue - North York (NY)
- Green - Etobicoke (E)
- Yellow - York (Y)
- Orange - East York (EY)

Name	FM	Census Tracts	Population	Land area (km2)	Density (people/km2)	% Change in Population since 2001	Average Income	Transit Commuting %	% Renters	Second most common language (after English) by name	Second most common language (after English) by percentage
Toronto CMA Average		All	5,113,149	5903.63	866	9.0	40,704	10.6	11.4		
Agincourt	S	0377.01, 0377.02, 0377.03, 0377.04, 0378.02, 0378.08, 0378.14, 0378.23, 0378.24	44,577	12.45	3580	4.6	25,750	11.1	5.9	Cantonese (19.3%)	19.3% Cantonese
Alderwood	E	0211.00, 0212.00	11,656	4.94	2360	-4.0	35,239	8.8	8.5	Polish (6.2%)	06.2% Polish
Alexandra Park	OCoT	0039.00	4,355	0.32	13,609	0.0	19,687	13.8	28.0	Cantonese (17.9%)	17.9% Cantonese
Allenby	OCoT	0140.00	2,513	0.58	4333	-1.0	245,592	5.2	3.4	Russian (1.4%)	01.4% Russian

Figure 7. The Wikipedia website for Toronto’s demographic information

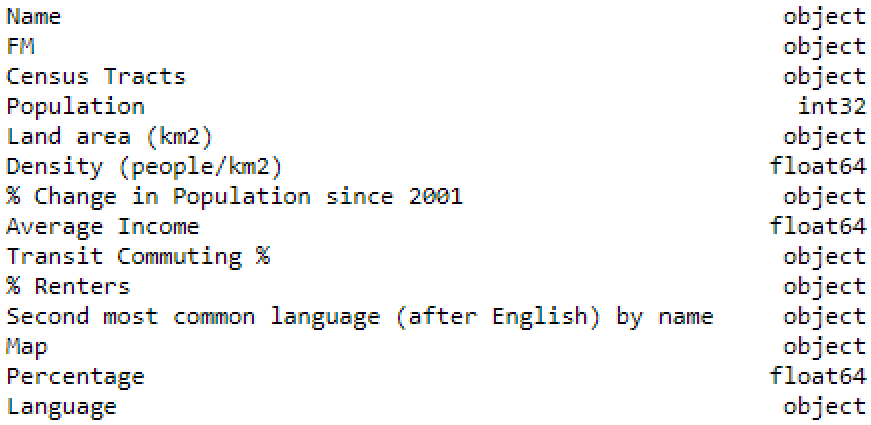
The demographic data for Toronto was scrapped from the Wikipedia site. There was no missing data points in the cells. All of the data types were object types as shown in figure 8, so we needed to convert some types to either float or integer types for further analysis of the data.



```
Data columns (total 13 columns):
Name                                                  175 non-null object
FM                                                    175 non-null object
Census Tracts                                         175 non-null object
Population                                            175 non-null object
Land area (km2)                                       175 non-null object
Density (people/km2)                                  175 non-null object
% Change in Population since 2001                     175 non-null object
Average Income                                        175 non-null object
Transit Commuting %                                   175 non-null object
% Renters                                             175 non-null object
Second most common language (after English) by name   175 non-null object
Second most common language (after English) by percentage   175 non-null object
Map                                                   175 non-null object
dtypes: object(13)
```

Figure 8. Raw Demographic data types

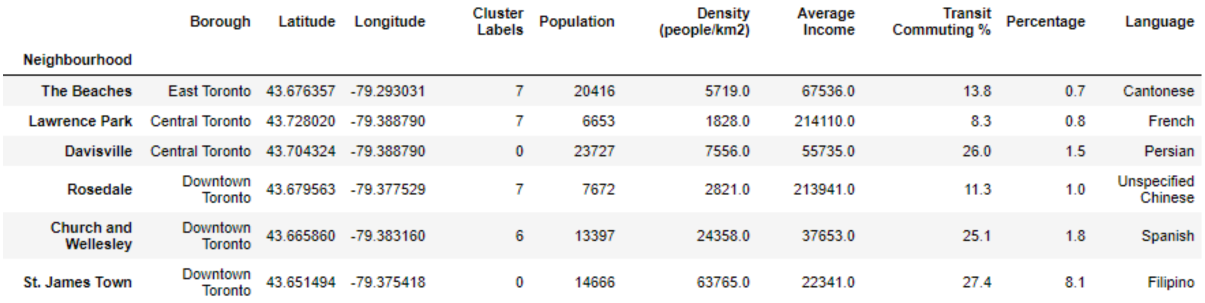
Some data columns in the data frame have commas in their cells. The commas will create issues for us when we convert the type from object to either float or integer. Therefore, we need to remove the commas and convert the data type. Figure 9 shows the data after the data type after conversion.



```
Name                                                   object
FM                                                     object
Census Tracts                                          object
Population                                              int32
Land area (km2)                                        object
Density (people/km2)                                  float64
% Change in Population since 2001                      object
Average Income                                        float64
Transit Commuting %                                    object
% Renters                                              object
Second most common language (after English) by name    object
Map                                                    object
Percentage                                            float64
Language                                               object
```

Figure 9. Demographic data type after conversion

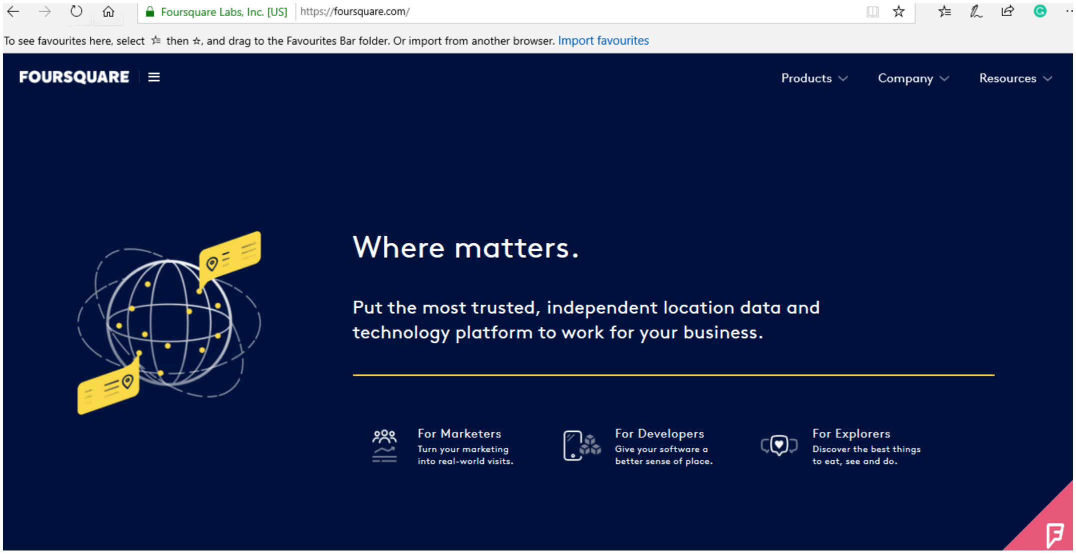
Some of the unnecessary data was dropped from the data frame and the Demographic data was appended with the coordinates as shown in figure 10.



	Borough	Latitude	Longitude	Cluster Labels	Population	Density (people/km2)	Average Income	Transit Commuting %	Percentage	Language
Neighbourhood										
The Beaches	East Toronto	43.676357	-79.293031	7	20416	5719.0	67536.0	13.8	0.7	Cantonese
Lawrence Park	Central Toronto	43.728020	-79.388790	7	6653	1828.0	214110.0	8.3	0.8	French
Davisville	Central Toronto	43.704324	-79.388790	0	23727	7556.0	55735.0	26.0	1.5	Persian
Rosedale	Downtown Toronto	43.679563	-79.377529	7	7672	2821.0	213941.0	11.3	1.0	Unspecified Chinese
Church and Wellesley	Downtown Toronto	43.665860	-79.383160	6	13397	24358.0	37653.0	25.1	1.8	Spanish
St. James Town	Downtown Toronto	43.651494	-79.375418	0	14666	63765.0	22341.0	27.4	8.1	Filipino

Figure 10. Demographic data appended to the coordinates

Then I used the Foursquare APIs to explore Toronto’s venues especially the restaurants to decide on the best location for a Chinese cousin restaurant based on selecting the neighbourhood with at least 10% of the residents to be of Chinese origin and with the highest population density. Figure 11 shows the Foursquare website on www.foursquare.com.


Foursquare Labs, Inc. [US] https://foursquare.com/

To see favourites here, select ☆ then ☆, and drag to the Favourites Bar folder. Or import from another browser. Import favourites

Where matters.

Put the most trusted, independent location data and technology platform to work for your business.

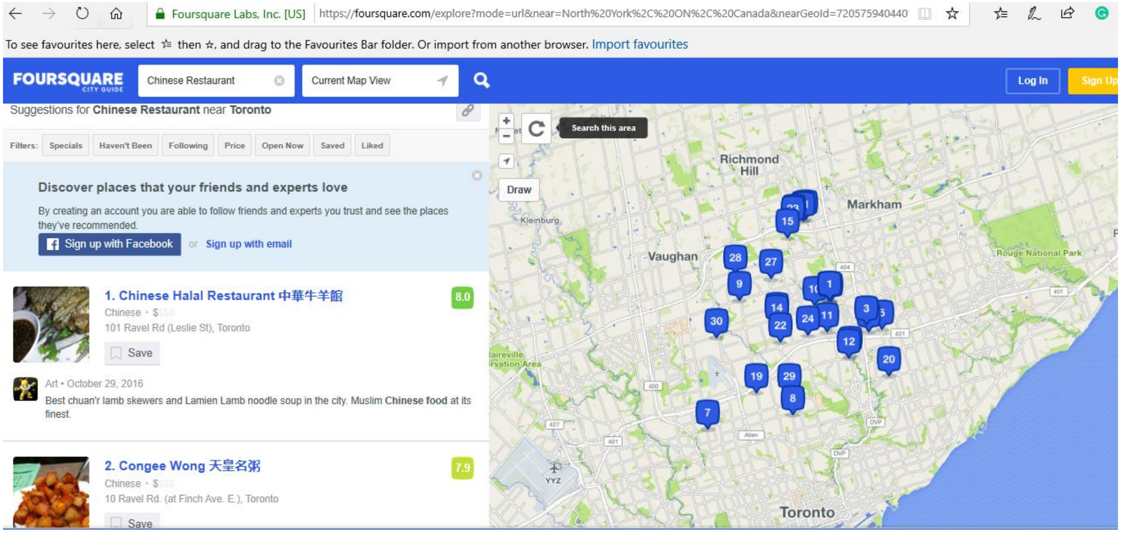
For Marketers
Turn your marketing into real-world visits.

For Developers
Give your software a better sense of place.

For Explorers
Discover the best things to eat, see and do.

Figure 11. The Foursquare homepage

Utilizing the Foursquare website, we can see that we have around 30 Chinese restaurants in the region. Figure 12 shows the locations of the restaurants in the selected FSA (Forward Sortation Area) of North York.


Foursquare Labs, Inc. [US] https://foursquare.com/explore?mode=url&near=North%20York%2C%20ON%2C%20Canada&nearGeoId=720575940440

To see favourites here, select ☆ then ☆, and drag to the Favourites Bar folder. Or import from another browser. Import favourites

Chinese Restaurant

Current Map View

Suggestions for **Chinese Restaurant** near **Toronto**

Filters: Specials Haven't Been Following Price Open Now Saved Liked

Discover places that your friends and experts love

By creating an account you are able to follow friends and experts you trust and see the places they've recommended.

Sign up with Facebook or Sign up with email

1. Chinese Halal Restaurant 中華牛羊館 8.0
Chinese · $
101 Ravel Rd (Leslie St), Toronto
Save

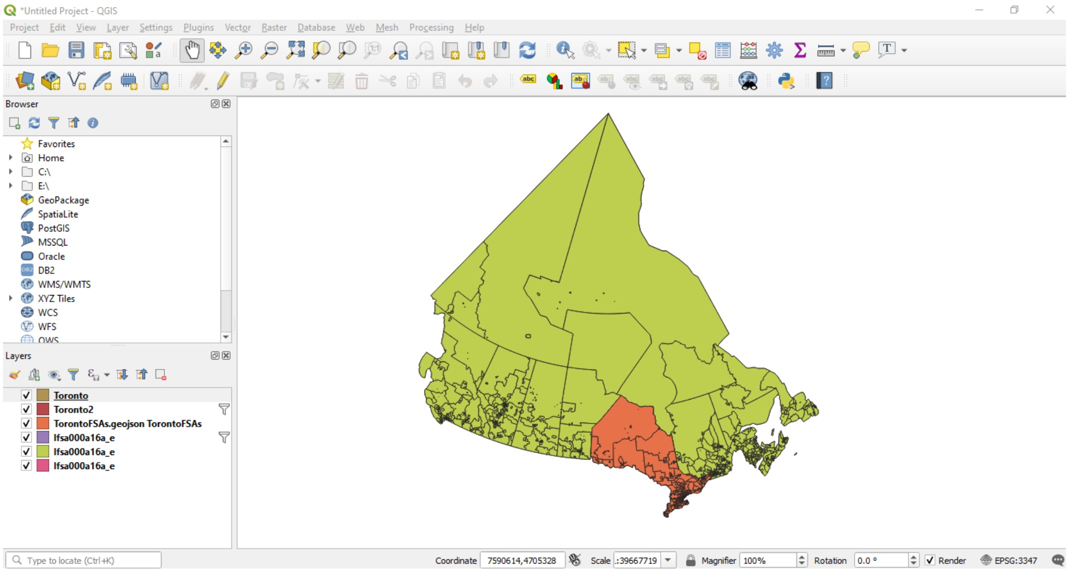
Art · October 29, 2016
Best chuan'r lamb skewers and Lamien Lamb noodle soup in the city. Muslim **Chinese food** at its finest.

2. Congee Wong 天皇名粥 7.9
Chinese · $
10 Ravel Rd. (at Finch Ave. E.), Toronto
Save

Figure 12. The Foursquare.com homepage displaying the Chinese restaurants

In this project, I selected only the FSA with the highest population density utilizing Folium mapping capability. Choosing the exact location inside the FSA requires more analysis that can be done either by ArcGIS or QGIS software which has full spatial analysis capabilities, which are outside the scope of this project.

In order to create the choropleth map of Toronto's population density I had to use Folium capabilities utilizing the demographic data. The demographic data was downloaded from the StatsCan website: (https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2016-eng.cfm) based on the FSA (Forward Sortation Area—the first three digits of the Canadian Postal Code). A forward sortation area (FSA) is a way to designate a geographical unit based on the first three characters in a Canadian postal code. All postal codes that start with the same three characters—for example, K1A—are together considered an FSA. The use of Folium for creating a choropleth map requires a GeoJSON file for Toronto as an input. This file is not readily available online, or from other sources, so I had to create this file utilizing QGIS, which is free GIS software. I had to download QGIS and have it installed before I can start this process. Figure 13 shows the QGIS software.

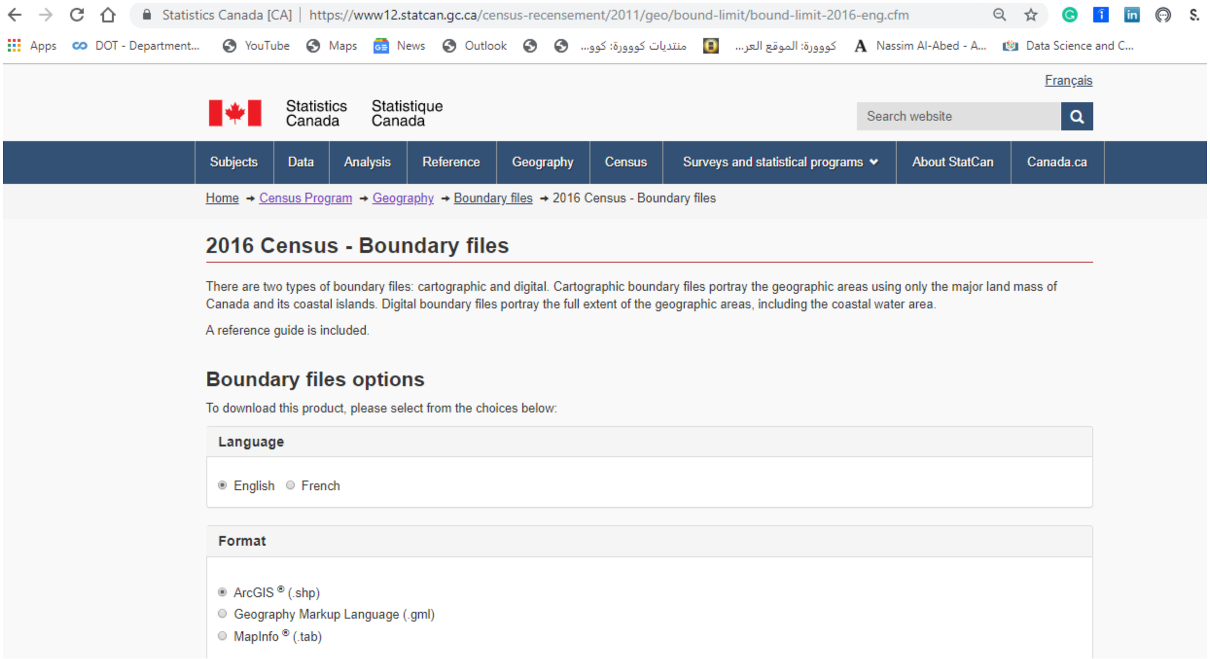


☑ Toronto
☑ Toronto2
☑ TorontoFSAs.geojson TorontoFSAs
☑ lfsa000a16a_e
☑ lfsa000a16a_e
☑ lfsa000a16a_e

Coordinate 7590614,4705328 Scale .:39667719 Magnifier 100% Rotation 0.0 ° ☑ Render EPSG:3347

Figure 13. QGIS Software used to open and create a GeoJSON file for Toronto

To download the data from StatsCan website, we need to select the format as ArcGIS shapefile format as shown in figure 14 and to choose the data type as forward sortation area (FSA) as shown in figure 15.



2016 Census - Boundary files

There are two types of boundary files: cartographic and digital. Cartographic boundary files portray the geographic areas using only the major land mass of Canada and its coastal islands. Digital boundary files portray the full extent of the geographic areas, including the coastal water area.

A reference guide is included.

Boundary files options

To download this product, please select from the choices below:

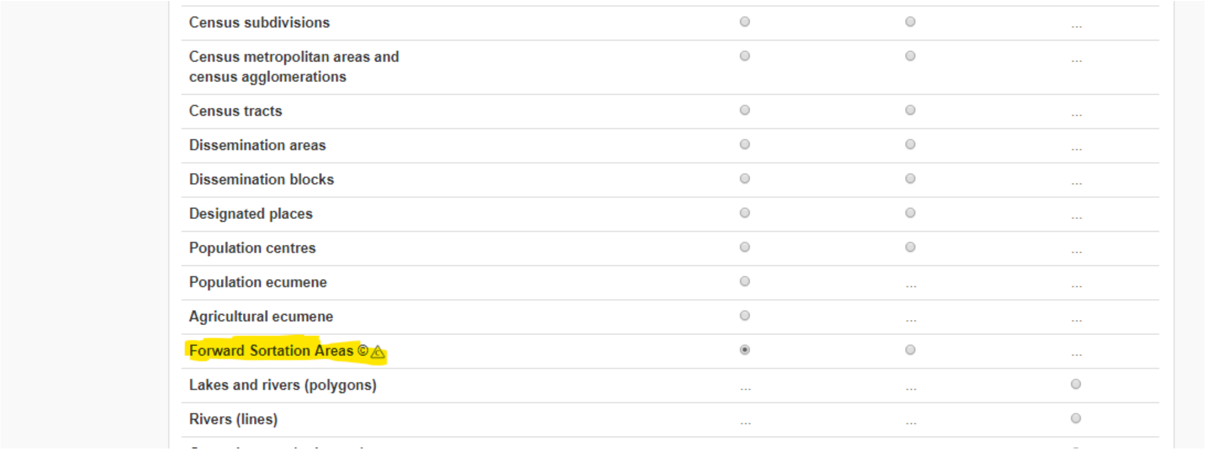
Language

● English ○ French

Format

● ArcGIS ® (.shp)
○ Geography Markup Language (.gml)
○ MapInfo ® (.tab)

Figure 14. Downloading the 2016 census data



Census subdivisions	○	○	...
Census metropolitan areas and census agglomerations	○	○	...
Census tracts	○	○	...
Dissemination areas	○	○	...
Dissemination blocks	○	○	...
Designated places	○	○	...
Population centres	○	○	...
Population ecumene	○	...	...
Agricultural ecumene	○	...	...
Forward Sortation Areas © ⚠	●	○	...
Lakes and rivers (polygons)	...	...	○
Rivers (lines)	...	...	○

Figure 15. Selecting the boundary file as a forward sortation area (FSA)

The downloaded boundary file will be as a zip file that contains several files, and one of them is the ArcGIS shapefile. ArcGIS software does not support the conversion of the shapefile to Geojson format, but QGIS software can convert the shapefile to Geojson format. In QGIS, I added the ArcGIS shapefile using the add layer option, and then add as vector layer as shown in figure 16.

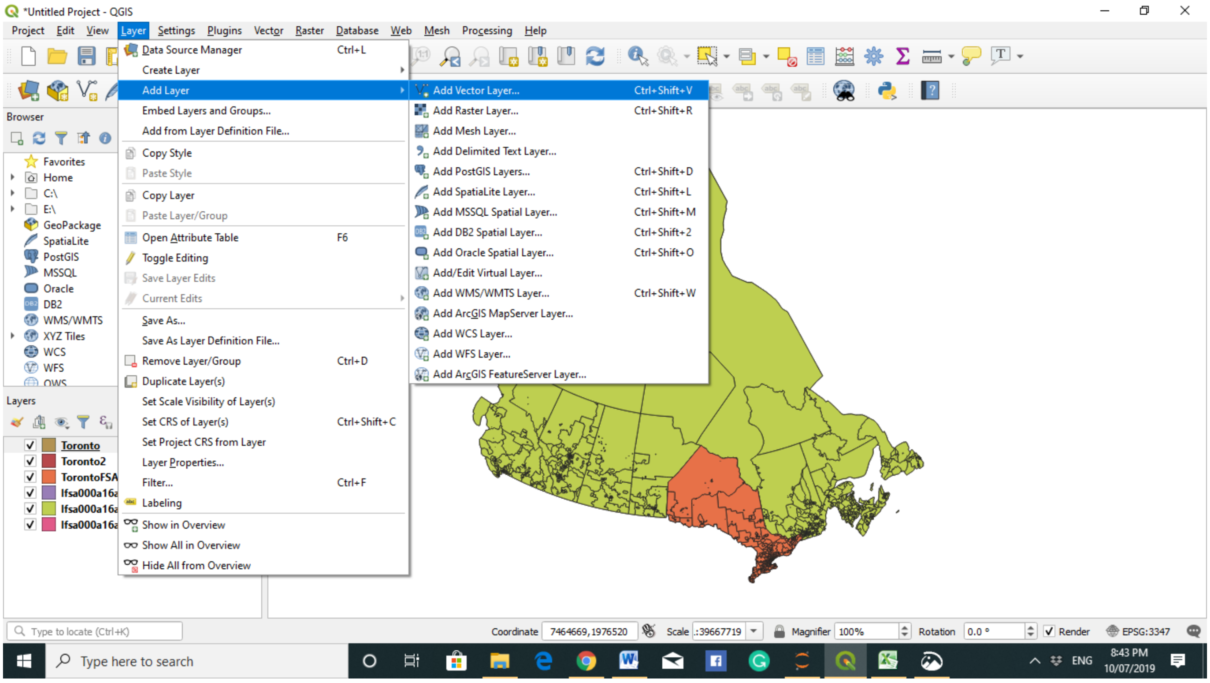
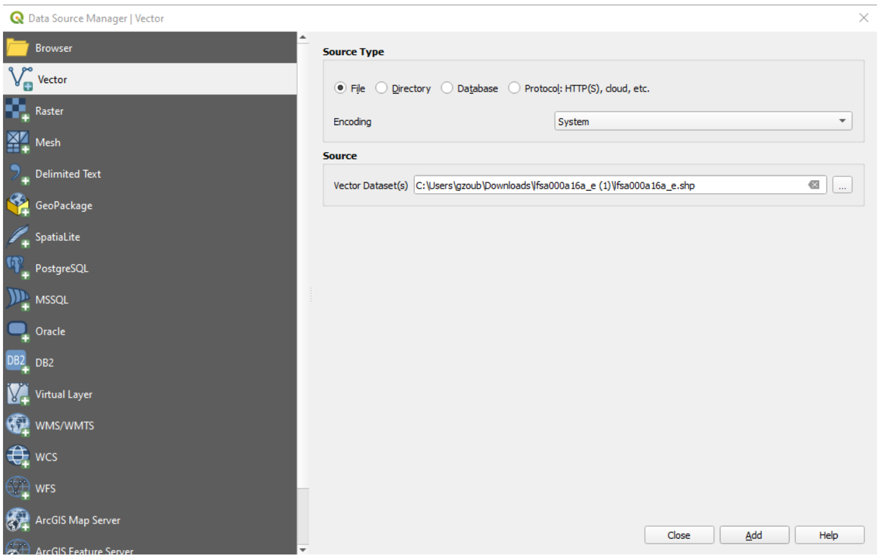

*Untitled Project - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help
Data Source Manager Ctrl+L
Create Layer
Add Layer
Embed Layers and Groups...
Add from Layer Definition File...
Copy Style
Paste Style
Copy Layer
Paste Layer/Group
Open Attribute Table F6
Toggle Editing
Save Layer Edits
Current Edits
Save As...
Save As Layer Definition File...
Remove Layer/Group Ctrl+D
Duplicate Layer(s)
Set Scale Visibility of Layer(s)
Set CRS of Layer(s) Ctrl+Shift+C
Set Project CRS from Layer
Layer Properties...
Filter... Ctrl+F
Labeling
Show in Overview
Show All in Overview
Hide All from Overview
Add Vector Layer... Ctrl+Shift+V
Add Raster Layer... Ctrl+Shift+R
Add Mesh Layer...
Add Delimited Text Layer...
Add PostGIS Layers... Ctrl+Shift+D
Add SpatiaLite Layer... Ctrl+Shift+L
Add MSSQL Spatial Layer... Ctrl+Shift+M
Add DB2 Spatial Layer... Ctrl+Shift+2
Add Oracle Spatial Layer... Ctrl+Shift+O
Add/Edit Virtual Layer...
Add WMS/WMTS Layer... Ctrl+Shift+W
Add ArcGIS MapServer Layer...
Add WCS Layer...
Add WFS Layer...
Add ArcGIS FeatureServer Layer...
Browser
Favorites
Home
C:\
E:\
GeoPackage
SpatiaLite
PostGIS
MSSQL
Oracle
DB2
WMS/WMTS
XYZ Tiles
WCS
WFS
OWS
Layers
Toronto
Toronto2
TorontoFSA
lfsa000a16a
lfsa000a16a
lfsa000a16a
Type to locate (Ctrl+K)
Coordinate 7464669,1976520
Scale :39667719
Magnifier 100%
Rotation 0.0 °
Render
EPSG:3347
Type here to search
ENG
8:43 PM
10/07/2019

Figure 16. Adding the shapefile to QGIS as a vector layer

Then we need to specify the shapefile name and the directory where it was saved as shown in Figure 17, and then we click add. The file is vast and covers the whole of Canada, so we need to filter only Ontario province and then Toronto city. The size reduction will be done utilizing the Filter option in QGIS as shown in Figure 18.



Source Type

● File ○ Directory ○ Database ○ Protocol: HTTP(S), cloud, etc.

Encoding System

Source

Vector Dataset(s) C:\Users\gzoub\Downloads\lfsa000a16a_e (1)\lfsa000a16a_e.shp

Figure 17. Specifying the shapefile name and its location in QGIS

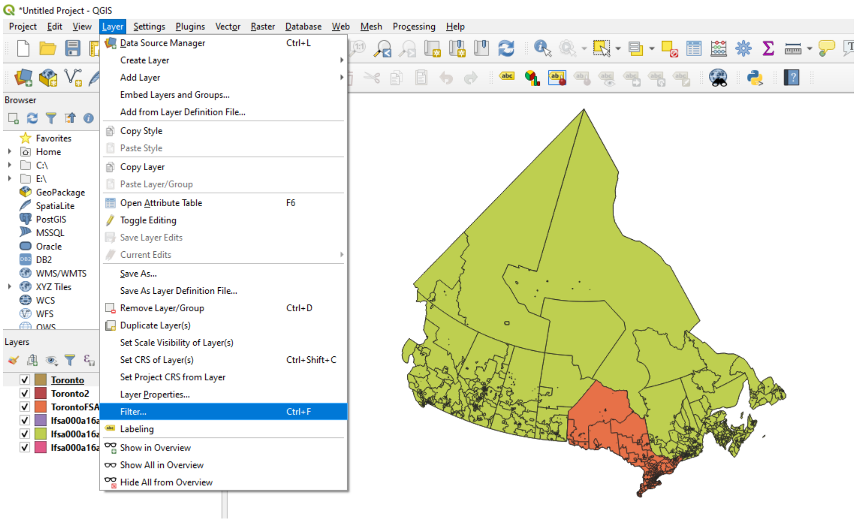
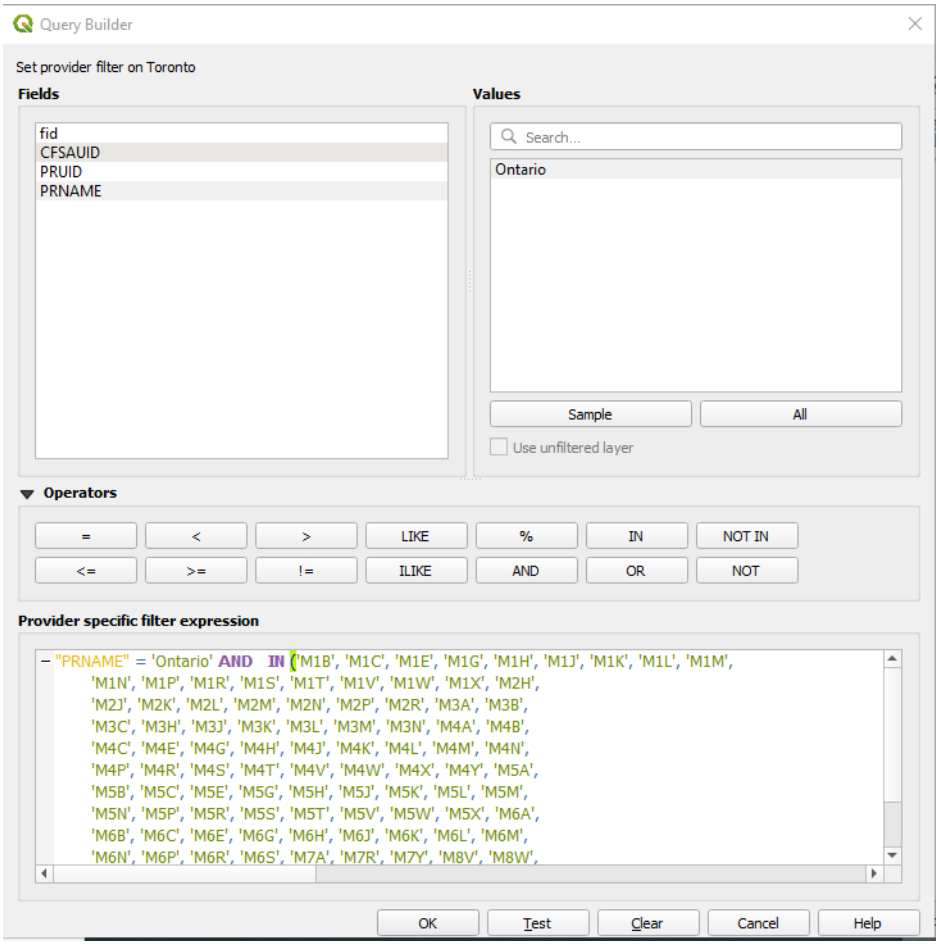


Figure 18. Using the Filter option in QGIS to select Ontario and then Toronto

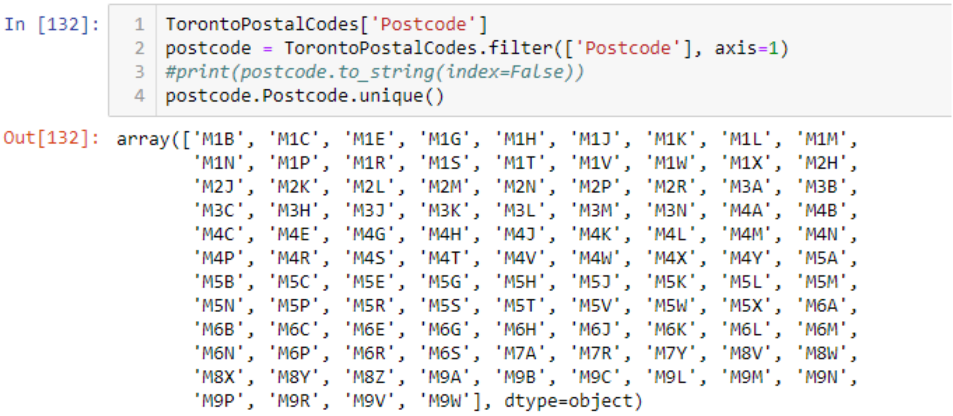
As shown in Figure 19, we filtered based on province name Ontario, and based on CFSAUID of Toronto.



```
"PRNAME" = 'Ontario' AND  IN ('M1B', 'M1C', 'M1E', 'M1G', 'M1H', 'M1J', 'M1K', 'M1L', 'M1M',
     'M1N', 'M1P', 'M1R', 'M1S', 'M1T', 'M1V', 'M1W', 'M1X', 'M2H',
     'M2J', 'M2K', 'M2L', 'M2M', 'M2N', 'M2P', 'M2R', 'M3A', 'M3B',
     'M3C', 'M3H', 'M3J', 'M3K', 'M3L', 'M3M', 'M3N', 'M4A', 'M4B',
     'M4C', 'M4E', 'M4G', 'M4H', 'M4J', 'M4K', 'M4L', 'M4M', 'M4N',
     'M4P', 'M4R', 'M4S', 'M4T', 'M4V', 'M4W', 'M4X', 'M4Y', 'M5A',
     'M5B', 'M5C', 'M5E', 'M5G', 'M5H', 'M5J', 'M5K', 'M5L', 'M5M',
     'M5N', 'M5P', 'M5R', 'M5S', 'M5T', 'M5V', 'M5W', 'M5X', 'M6A',
     'M6B', 'M6C', 'M6E', 'M6G', 'M6H', 'M6J', 'M6K', 'M6L', 'M6M',
     'M6N', 'M6P', 'M6R', 'M6S', 'M7A', 'M7R', 'M7Y', 'M8V', 'M8W',
```

Figure 19. Filtering the data of Toronto city in QGIS

To filter Toronto city data, I had to write a python code to find the unique postal codes of Toronto, as shown in figure 20.



```
In [132]:  TorontoPostalCodes['Postcode']
           postcode = TorontoPostalCodes.filter(['Postcode'], axis=1)
           #print(postcode.to_string(index=False))
           postcode.Postcode.unique()

Out[132]: array(['M1B', 'M1C', 'M1E', 'M1G', 'M1H', 'M1J', 'M1K', 'M1L', 'M1M',
                 'M1N', 'M1P', 'M1R', 'M1S', 'M1T', 'M1V', 'M1W', 'M1X', 'M2H',
                 'M2J', 'M2K', 'M2L', 'M2M', 'M2N', 'M2P', 'M2R', 'M3A', 'M3B',
                 'M3C', 'M3H', 'M3J', 'M3K', 'M3L', 'M3M', 'M3N', 'M4A', 'M4B',
                 'M4C', 'M4E', 'M4G', 'M4H', 'M4J', 'M4K', 'M4L', 'M4M', 'M4N',
                 'M4P', 'M4R', 'M4S', 'M4T', 'M4V', 'M4W', 'M4X', 'M4Y', 'M5A',
                 'M5B', 'M5C', 'M5E', 'M5G', 'M5H', 'M5J', 'M5K', 'M5L', 'M5M',
                 'M5N', 'M5P', 'M5R', 'M5S', 'M5T', 'M5V', 'M5W', 'M5X', 'M6A',
                 'M6B', 'M6C', 'M6E', 'M6G', 'M6H', 'M6J', 'M6K', 'M6L', 'M6M',
                 'M6N', 'M6P', 'M6R', 'M6S', 'M7A', 'M7R', 'M7Y', 'M8V', 'M8W',
                 'M8X', 'M8Y', 'M8Z', 'M9A', 'M9B', 'M9C', 'M9L', 'M9M', 'M9N',
                 'M9P', 'M9R', 'M9V', 'M9W'], dtype=object)
```

Figure 20. Finding the exclusive postal codes for Toronto

After clicking on OK in QGIS Query builder, the data for Toronto will be selected as shown in figure 21.

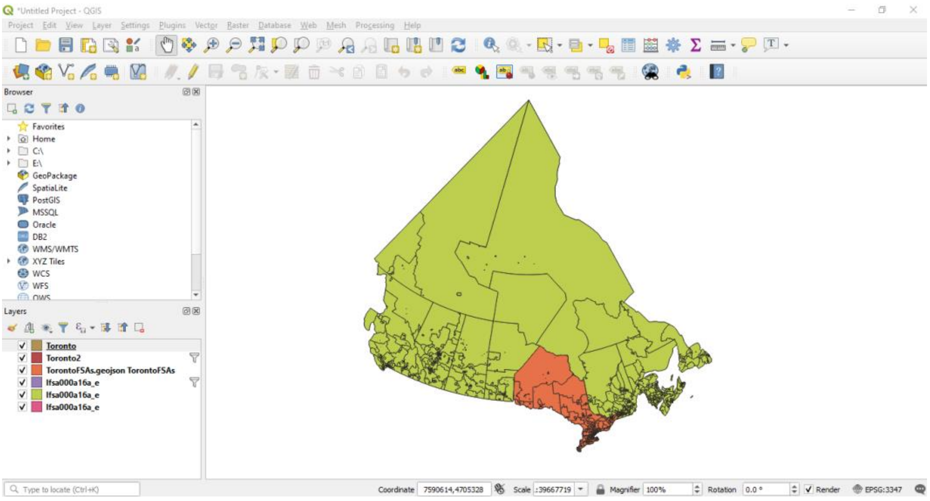


Figure 21. Saving the filtered layer in QGIS

Then after filtering the data, I selected save as option in QGIS as shown in figure 22. Then we need to specify in the save vector layer as the format that we need to save the layer into and the folder where to save it, and we need to select EPSG:4326 – WGS 84 which specify that the data will be in geographic projection as shown in figure 23.

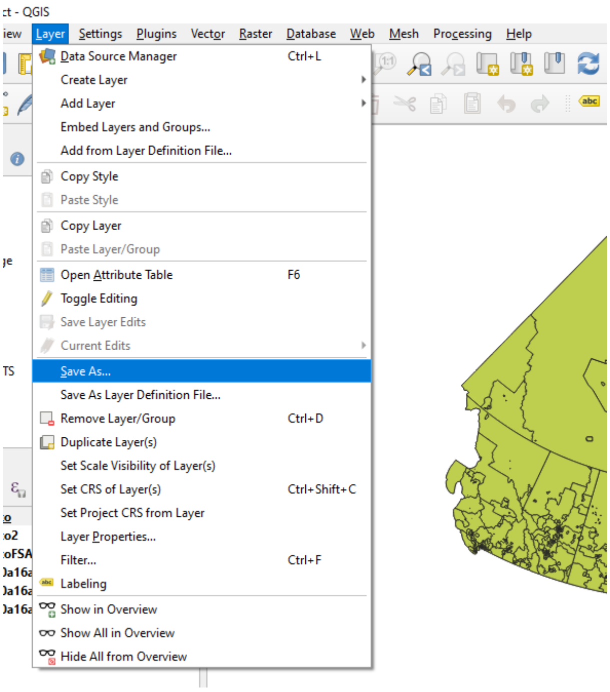
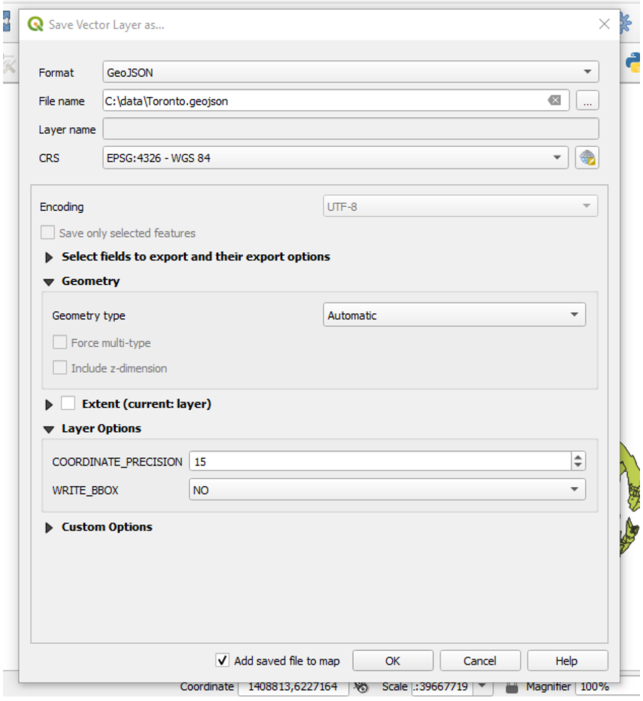


Figure 22. Saving the filtered layer in QGIS



Save Vector Layer as...

Format: GeoJSON

File name: C:\data\Toronto.geojson

Layer name:

CRS: EPSG:4326 - WGS 84

Encoding: UTF-8

□ Save only selected features

▶ Select fields to export and their export options

▼ Geometry

Geometry type: Automatic

□ Force multi-type

□ Include z-dimension

▶ □ Extent (current: layer)

▼ Layer Options

COORDINATE_PRECISION: 15

WRITE_BBOX: NO

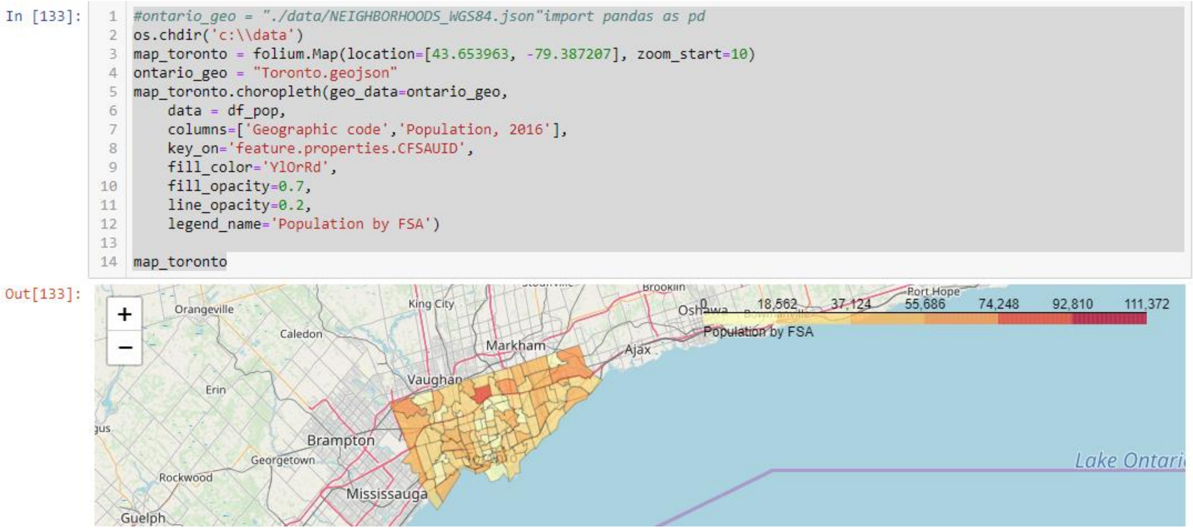
▶ Custom Options

☑ Add saved file to map

Figure 23. Selecting the Geojson file format in QGIS

Results

Then using Folium and specifying the Geojson file as shown in the code in figure 24, the choropleth map of Toronto’s population density drawn in the chart.



In [133]:

```
#ontario_geo = "./data/NEIGHBORHOODS_WGS84.json"import pandas as pd
os.chdir('c:\\data')
map_toronto = folium.Map(location=[43.653963, -79.387207], zoom_start=10)
ontario_geo = "Toronto.geojson"
map_toronto.choropleth(geo_data=ontario_geo,
    data = df_pop,
    columns=['Geographic code','Population, 2016'],
    key_on='feature.properties.CFSAUID',
    fill_color='YlOrRd',
    fill_opacity=0.7,
    line_opacity=0.2,
    legend_name='Population by FSA')

map_toronto
```

Out[133]:

Figure 24. Drawing the population density for Toronto

The demographic data analysis of Toronto showed that the Chines language was spoken by more than 10% in North York. Using the spatial analysis utilizing folium showed that North York had the highest population density in Toronto as shown in figure 24, even though there are around 30 Chinese restaurants in the area, North York fulfilled the conditions set at the start of the analysis so the new restaurant shall be located there.

Discussion and Conclusion

The demographic data analysis of Toronto showed that the Chines language was spoken by more than 10% in North York. Spatial analysis utilizing folium library showed that North York had the highest population density in Toronto, even though there are around 30 Chinese restaurants in the area, North York fulfilled the conditions set at the start of the analysis. This project showed that data science could be used to conduct spatial data analysis utilizing Python and folium library. Utilizing QGIS software was necessary because I could not find any available Geojson file for Toronto. Geojson files are necessary in order to create choropleth maps. North York satisfied the selected site selection criteria set at the start of the project for the new restaurant. The utilization of Foursquare API was necessary in order to find Toronto’s venues. Selecting the exact location of the new restaurant inside the Forward Sortation Area (FSA) was outside the scope of this project.