BATTLE OF THE NEIGHBORHOODS

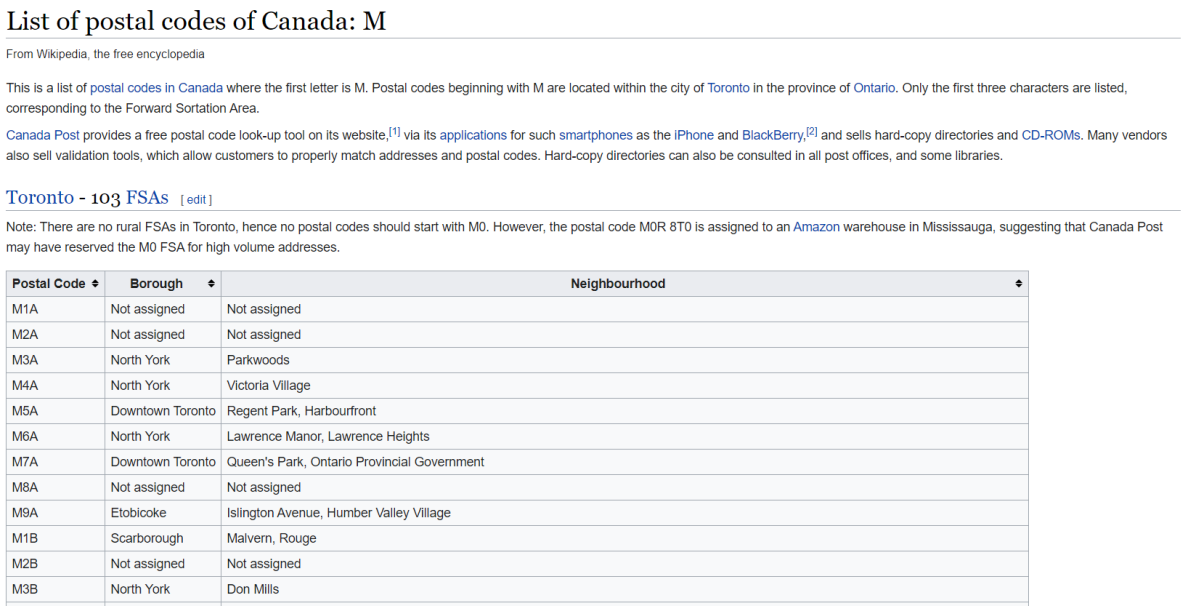
Exploring Suitable Locations for New Sushi Restaurants in Toronto, Canada using Data Science

Overview of Data:

The data required for the analysis will be obtained from multiple sources. The list of neighbourhoods in Toronto will be sourced from Wikipedia, the Geographical location of the neighbourhoods from a csv file and Venue data of Sushi restaurants from Foursquare. The Venue data will help find which neighbourhood is best suitable to open a Sushi restaurant in Toronto.

1) Toronto Neighborhood Data

The list of Toronto neighborhoods is sourced from Wikipedia (Fig.1). (“https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M”).



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 three 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 directories and 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 post offices, and some libraries.

Toronto - 103 FSAs [edit]

Note: There are no rural FSAs in Toronto, hence no postal codes should start with M0. However, the postal code M0R 8T0 is assigned to an Amazon warehouse in Mississauga, suggesting that Canada Post may have reserved the M0 FSA for high volume addresses.

Postal Code	Borough	Neighbourhood
M1A	Not assigned	Not assigned
M2A	Not assigned	Not assigned
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Regent Park, Harbourfront
M6A	North York	Lawrence Manor, Lawrence Heights
M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
M8A	Not assigned	Not assigned
M9A	Etobicoke	Islington Avenue, Humber Valley Village
M1B	Scarborough	Malvern, Rouge
M2B	Not assigned	Not assigned
M3B	North York	Don Mills

Fig.1 A screen grab of the Wikipedia page containing Toronto Neighborhood data

The list contains Postal Codes, Name of Boroughs and Neighbourhoods. The data is available in a format which is not suitable for the analysis. Therefore, the data

is scraped from the Wikipedia page. Data scraping is done from the website as it is suitable for the analysis. The scraped data is then wrangled, cleaned and read into Pandas data frame so that it is in a structured format (Fig.2).

	Postalcode	Borough	Neighborhood
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park, Harbourfront
5	M6A	North York	Lawrence Manor, Lawrence Heights
6	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government

Fig.2 Scraped Data in Pandas Data Frame

2) Geographical Location data

The Geographical coordinates of the Toronto neighbourhoods with the respective Postal Codes was sourced from the website https://cocl.us/Geospatial_data. The data is in csv format (Fig.3a). The data was converted to Pandas data frame (Fig 3b).

	A	B	C
1	Postal Code	Latitude	Longitude
2	M1B	43.80669	-79.1944
3	M1C	43.78454	-79.1605
4	M1E	43.76357	-79.1887
5	M1G	43.77099	-79.2169
6	M1H	43.77314	-79.2395

Fig.3a Geospatial data in csv format

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

Fig.3b Geospatial Data in Pandas Data Frame

3) Venue Data using Foursquare

The Neighborhood data frame and geospatial data frame were merged to get a new data frame (Fig.4).

	Postalcode	Borough	Neighborhood	Latitude	Longitude
37	M4E	East Toronto	The Beaches	43.676357	-79.293031
41	M4K	East Toronto	The Danforth West, Riverdale	43.679557	-79.352188
42	M4L	East Toronto	India Bazaar, The Beaches West	43.668999	-79.315572
43	M4M	East Toronto	Studio District	43.659526	-79.340923
44	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790

Fig.4 Merged Data Frame

Then using Foursquare credentials (client ID, client secret and version) and the data in the merged data frame, the venue data is extracted (Fig.5)

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	The Beaches	43.676357	-79.293031	Glen Manor Ravine	43.676821	-79.293942	Trail
1	The Beaches	43.676357	-79.293031	The Big Carrot Natural Food Market	43.678879	-79.297734	Health Food Store
2	The Beaches	43.676357	-79.293031	Grover Pub and Grub	43.679181	-79.297215	Pub
3	The Beaches	43.676357	-79.293031	Upper Beaches	43.680563	-79.292869	Neighborhood
4	The Danforth West, Riverdale	43.679557	-79.352188	MenEssentials	43.677820	-79.351265	Cosmetics Shop

Fig.5 Venue Data extracted from Foursquare API

This Venue data is used for further analysis.