IBM APPLIED DATA SCIENCE CAPSTONE

RELOCATING TO A NEIGHBOURHOOD IN TORONTO

INTRODUCTION

Toronto is the most populous city, the financial, Educational and technological hub of Canada. Both my friend and I are relocating to Toronto later this year, and will be staying in Toronto for 2 years.

Problem Description

Both my Friend and I will be staying about 2 years Toronto. We both really love Japanese food so I would like to stay in a neighbourhood which has some Japanese restaurants, among other things. Further, we do not want to spend a lot on rental and our budget for monthly rental expenses is approximately $1500 per month.

So, the problem description can be best summarised in the below table:

Problem description	Solution
➢ Choosing neighbourhood with low average real estate rental (Our budget is about $1500)	➢ Identify neighbourhood where average real estate rental is low (around $1500)
➢ Choose neighbourhood with some Japanese restaurants	➢ Identify neighbourhood with Japanese's restaurant as a common venue

Further, since will be in Toronto for 2 years, we will not have a car of our own so I would like the restaurants within 500 meters radius from the location.

Target Audience

Results of this Project can also be helpful for the following scenarios:

a. Specifically, anybody planning to relocate to Toronto for a short duration, do not want to miss tasty Japanese food and want to save cost on rental expenses
b. This analysis can be generalised for anybody/ any family planning to relocate to Toronto and wants good variety of restaurants around them

DATA SOURCE

Following data are chosen:

a. Foursquare API for identifying venues, venue categories such as cafes, restaurants
b. Wikipedia page titled 'List of Postal Codes of Canada_M'
c. Average real estate rental from Kaggle dataset on average rentals

THE FOURSQUARE API (https://developer.foursquare.com/places) allows application developers to interact with the Foursquare platform. The API allows querying places and users, exploring popular places, and checking out reviews and photographs for these places. In this project we will be using Foursquare explore API to get the list of venues for each neighbourhood based on the locational coordinates and venue categories. Venues are explored and selected from of 500 meters of locational coordinates (latitude and longitude) as we feel that is walking distance.

Toronto neighbourhoods data were scraped from a Wikipedia page titled 'List of Postal Codes of Canada: M' (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M). The dataset has list of postal codes along with name of borough and neighbourhoods. Geolocator was used the further enhance the dataset by latitudes and longitudes.

Real estate rentals in Toronto area is obtained from preformatted and precleaned csv file from Kaggle containing rentals and locational coordinates.

METHODOLOGY

Data Extraction and Cleaning

The web scrapped data on postal codes from Wikipedia page titled 'List of Postal Codes of Canada: M' was transformed into a pandas dataframe. A lot of missing values were identified where postal codes were not assigned to any borough or neighbourhood, and were dropped.

Geolocator Nominatim from Geopy library was used to include locational coordinates (latitude and longitude) of each neighbourhood, as make a dataset which has postal codes, borough name, neighbourhood name, latitude and longitude. In cases where Geopy could not provide the locational coordinates a precleaned csv file was used to fill the gaps.

It was observed that in many cases the neighbourhoods were grouped together for one single postal code as given below.

Postal code	Borough	Neighbourhood
M5A	Downtown Toronto	Regent Park, Harbourfront
M6A	North York	Lawrence Manor, Lawrence Heights

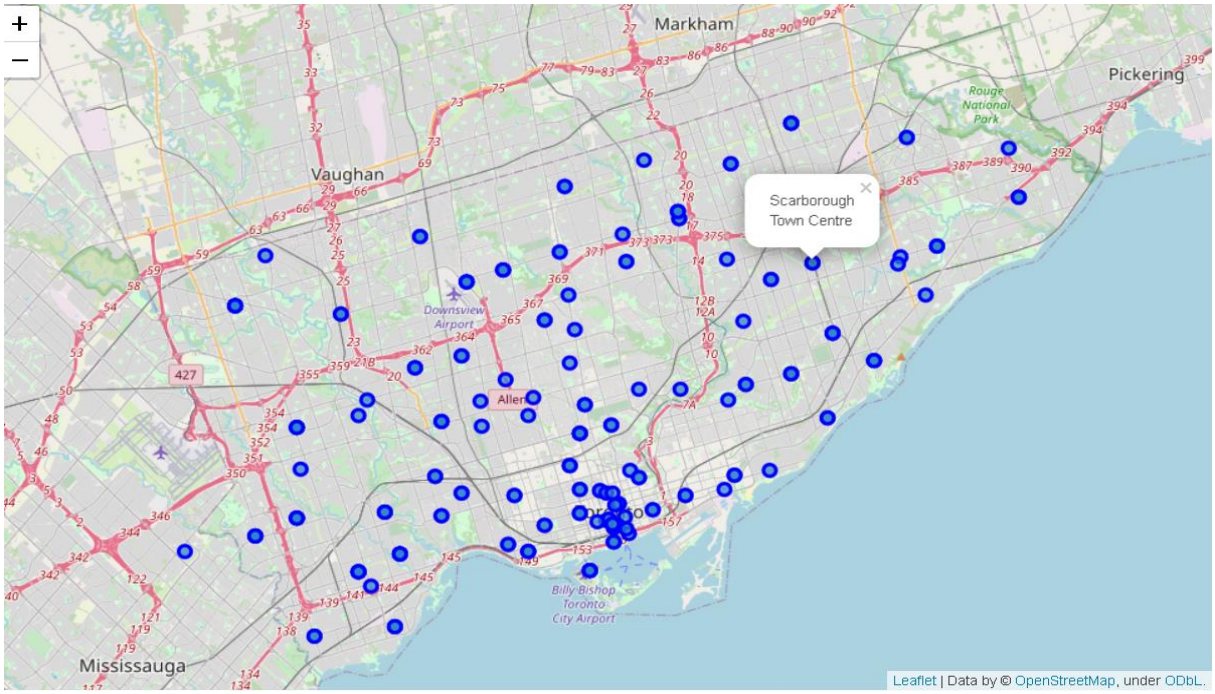
For all these cases strings in each of cells were split as given below, so that we have the complete list of all neighbourhoods in Toronto along with postal code, name of borough and locational coordinates.

Postal code	Borough	Neighbourhood
M5A	Downtown Toronto	Regent Park
M5A	Downtown Toronto	Harbourfront
M6A	North York	Lawrence Manor
M6A	North York	Lawrence Heights

After splitting, the list of 209 neighbourhoods across Toronto, each with Postal code, name of borough and locational coordinates (latitude and longitude) was obtained. A snapshot of 5 neighbourhoods are provided below (table I).

Postal code	Borough	Neighborhood	Latitude	Longitude
M5C	Downtown Toronto	St. James Town	43.669403	-79.372704
M4E	East Toronto	The Beaches	43.671024	-79.296712
M5E	Downtown Toronto	Berczy Park	43.647984	-79.375396
M5G	Downtown Toronto	Central Bay Street	43.660708	-79.385802
M6G	Downtown Toronto	Christie	43.664111	-79.418405
M6H	West Toronto	Dufferin	43.660202	-79.435719

These neighbourhoods were then put in the map using folio to visualise geospatially. The map of Toronto with the neighbourhoods are shown below:



Foursquare API's explore option was utilised to identify list of venues and its categories across each of the neighbourhoods. In total 5184 such venues were identified across 288 categories which are within 500 meters of locational coordinates (latitude and longitude of the neighbourhood).

The list of venue category is then joined with the neighbourhood name, locational coordinates of table I. A glimpse of the first 5 rows given below (table II). Latitude, longitude of neighborhoods and venues were used to join and develop the below data set.

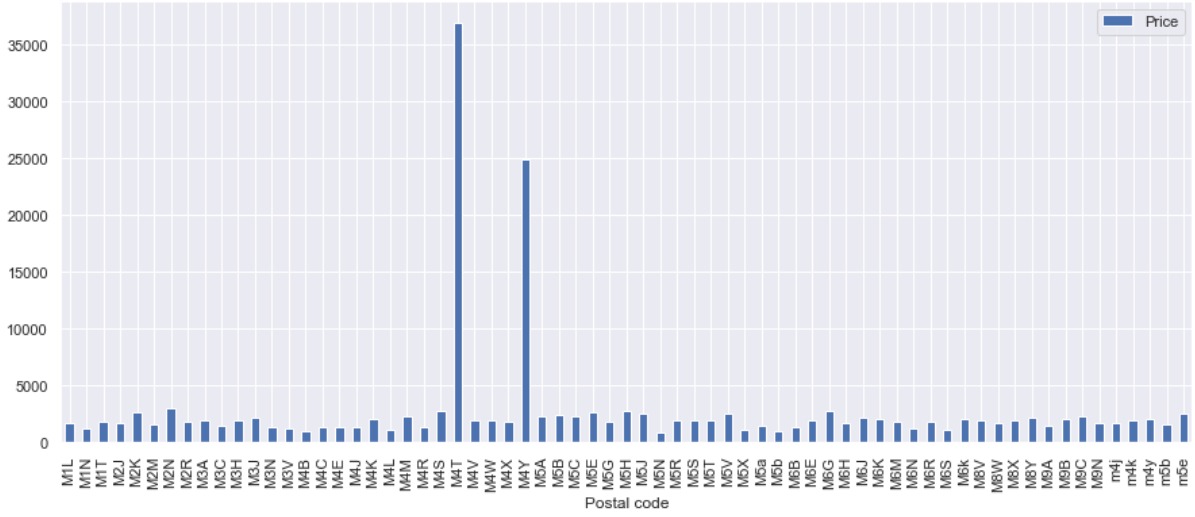
Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Parkwoods	43.758800	-79.320197	Allwyn's Bakery	43.759840	-79.324719	Caribbean Restaurant
Parkwoods	43.758800	-79.320197	LCBO	43.757774	-79.314257	Liquor Store
Parkwoods	43.758800	-79.320197	Petro-Canada	43.757950	-79.315187	Gas Station
Parkwoods	43.758800	-79.320197	Shoppers Drug Mart	43.760857	-79.324961	Pharmacy
Parkwoods	43.758800	-79.320197	Pizza Place	43.760231	-79.325666	Pizza Place

OneHotEncoder was used to encode the dataset of venues and was grouped based on neighbourhoods and then top venues are explored with a bar chart.

Dataset on average rentals in Toronto area is obtained from preformatted csv file from Kaggle containing rentals and locational coordinates was pre-cleaned.

Exploratory Analysis

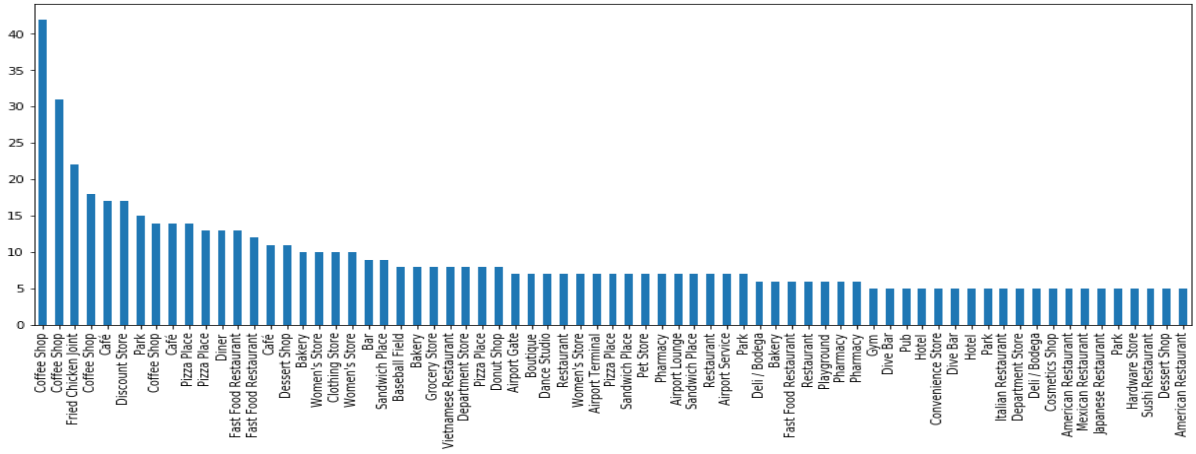
The dataset on rental was loaded. The average (mean) rental based on postal codes was computed using groupby and mean function. The below graph shows postal code wise average rentals in dollars.



Some statistical analysis was done to understanding the rental prices in Toronto was done. It was observed that certain neighbourhoods where real estate rentals were steep is increasing the mean rental across neighbourhoods. The table below summaries the findings:

Statistical Analysis	Results
Mean	$ 4,243.57
Median	$ 2,200.00
Standard Deviation	$ 31,902.63
Skewness	+ 16.58
Kurtosis	274.21

The list of venue categories was also explored to check the availability of Japanese restaurants. It was observed that there are small number of Japanese restaurants.



A lot of duplicate items were observed which will be subsequently be merged.

The data set shown in table I is grouped based on neighbourhood, and 5 most common venue type was identified across each of 209 neighbourhoods. A snapshot of top 5 is given below (table III) showing most common venue category.

Postal code	Borough	Neighborhood	Latitude	Longitude	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
M5C	Downtown Toronto	St. James Town	43.669403	-79.372704	Coffee Shop	Pizza Place	Café	Grocery Store	Playground
M4E	East Toronto	The Beaches	43.671024	-79.296712	Beach	Japanese Restaurant	Breakfast Spot	Bar	Tea Room
M5E	Downtown Toronto	Berczy Park	43.647984	-79.375396	Coffee Shop	Italian Restaurant	Café	Restaurant	Japanese Restaurant
M5G	Downtown Toronto	Central Bay Street	43.660708	-79.385802	Coffee Shop	Sandwich Place	Japanese Restaurant	Middle Eastern Restaurant	Diner
M6G	Downtown Toronto	Christie	43.664111	-79.418405	Korean Restaurant	Coffee Shop	Indian Restaurant	Mexican Restaurant	Café

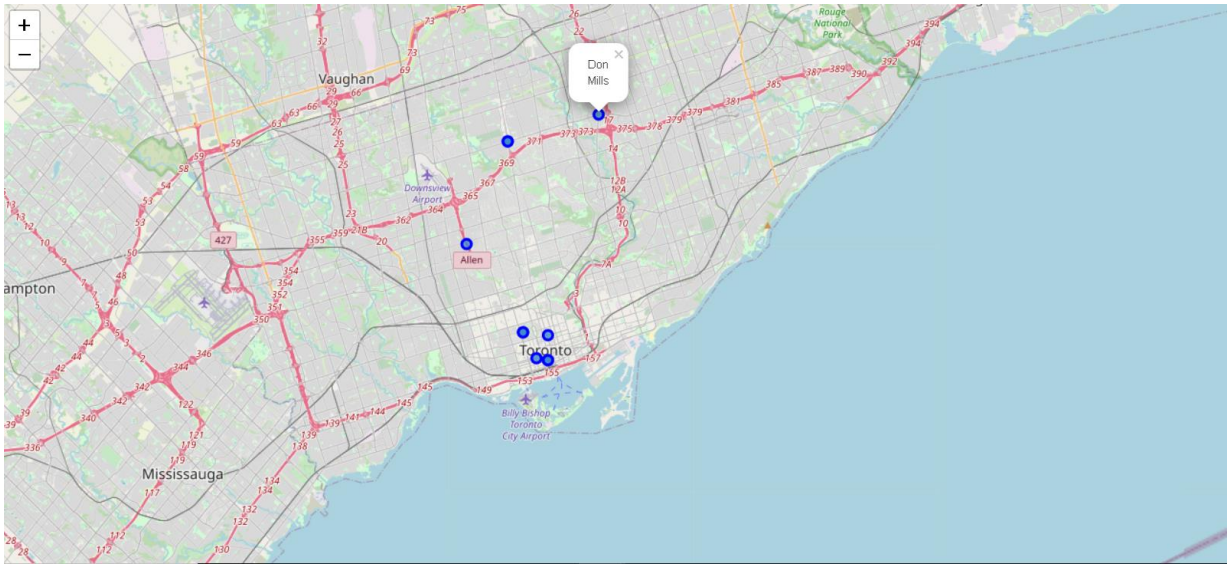
RESULTS

The final data set was developed by merging data set shown in table III with average rental prices of each neighbourhood. This has all the datapoints required to take a decision i.e. name of the neighbour, 5 most common venues, average rental and locational coordinates to map in folio.

Search criteria was set on this data set to identify the neighbourhoods which has Japanese Restaurants in top 5 venue categories.

11 Neighbourhood of Toronto where Japanese Restaurants featured in top 5 venue category, with average real estate rental ranging from $1100 to $24853.16

These neighbourhoods were then put in the map using folio to visualise geospatially. The map of Toronto with the neighbourhoods where Japanese restaurants are among top 5 categories are shown below:



The full list of neighbourhoods will the relevant datapoints in provided below (table IV).

Postal code	Borough	Neighborhood	Latitude	Longitude	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	Price
M5X	Downtown Toronto	First Canadian Place	43.648429	-79.382280	Coffee Shop	Café	Hotel	Japanese Restaurant	Restaurant	1100.00
M5X	Downtown Toronto	Underground city	43.648429	-79.382280	Coffee Shop	Café	Hotel	Japanese Restaurant	Restaurant	1100.00
M6B	North York	Glencairn	43.708712	-79.440685	Japanese Restaurant	Pub	Tennis Court	Playground	Grocery Store	1300.00
M3C	**North York**	**Don Mills**	**43.775347**	**-79.345944**	**Clothing Store**	**Coffee Shop**	**Japanese Restaurant**	**Fast Food Restaurant**	**Tea Room**	**1516.66**
M2J	North York	Fairview	43.778517	-79.346556	Clothing Store	Coffee Shop	Fast Food Restaurant	Japanese Restaurant	Cosmetics Shop	1763.50
M2J	North York	Henry Farm	43.778517	-79.346556	Clothing Store	Coffee Shop	Fast Food Restaurant	Japanese Restaurant	Cosmetics Shop	1763.50
M2J	North York	Oriole	43.778517	-79.346556	Clothing Store	Coffee Shop	Fast Food Restaurant	Japanese Restaurant	Cosmetics Shop	1763.50
M5G	Downtown Toronto	Central Bay Street	43.661425	-79.386067	Coffee Shop	Sandwich Place	Gastropub	Sushi Restaurant	Japanese Restaurant	1813.46
M2R	North York	Willowdale	43.761510	-79.410923	Coffee Shop	Japanese Restaurant	Grocery Store	Pharmacy	Fried Chicken Joint	1867.00
M2N	North York	Willowdale	43.761510	-79.410923	Coffee Shop	Japanese Restaurant	Grocery Store	Pharmacy	Fried Chicken Joint	3000.00
M4Y	Downtown Toronto	Church and Wellesley	43.661195	-79.382114	Coffee Shop	Japanese Restaurant	Diner	Burger Joint	Sandwich Place	24853.16

As stated earlier, we had 2 criteria as provided in **Problem Description** section i.e.

Problem description	Solution
➢ Choosing neighbourhood with low average real estate rental (Our budget is about $1500)	➢ Identify neighbourhood where average real estate rental is low (around $1500)
➢ Choose neighbourhood with some Japanese restaurants	➢ Identify neighbourhood with Japanese's restaurant as a common venue

From the table above it can be seen Don Mills neighborhood in North York borough fits both the criteria:

- Japanese Restaurants as top 3 most common venue
- Average rental price of $ 1516.66

So, Don Mills is the ideal location to relocate.

About Don Mills

Don Mills is a mixed-use neighbourhood in the North York district of Toronto, Ontario, Canada. It was developed to be a self-supporting town and part of Toronto metropolitan area in 1998. It has a population of 25,435 (2006) and of mixed ethnicity.