

Predicting a Viable Location for a Sushi Bar

Asterix Milind

Introduction

Background

Toronto is a cultural phenomenon, with residents from around the world and scores of languages, foods, customs, and celebrations. And it reflects food as well. One can find restaurants offering cuisines ranging from American to Indian, Chinese to Lebanese, and Korean to Mediterranean to name a few.

Problem Statement

Mr Kobayashi Maru wants to open a Sushi Bar. Opening a restaurant involves a lot of capital and the proprietor naturally expects the capital to generate a good return. For a restaurant to do well, a lot of factors come into play and location is an important factor. This project aims to predict a few locations that Mr Kobayashi Maru can choose from to open his Sushi Bar.

Data Acquisition and Cleaning

Data Sources

The boroughs, neighbourhoods data, and postal codes of Toronto were scraped from the [Wikipedia page](#).

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

Fig a. Wikipedia Page showing Neighbourhoods and Postal Codes of Toronto

The data was later merged with [Geospatial Data](#) over the postal codes. Restaurant data was fetched over the [Foursquare API](#).

	name	categories	lat	lng
0	Roselle Desserts	Bakery	43.653447	-79.362017
1	Tandem Coffee	Coffee Shop	43.653559	-79.361809
2	Impact Kitchen	Restaurant	43.656369	-79.356980
3	The Distillery Historic District	Historic Site	43.650244	-79.359323
4	Cooper Koo Family YMCA	Distribution Center	43.653249	-79.358008

Fig b. The output of Data via Foursquare API

Data Cleaning

The scraped data was first cleaned. Empty value cells and cells with the value “not assigned” were dropped. Then the data frame was merged with the data frame of geospatial data. With that, the co-ordinates were assigned to every neighbourhood. The

cells having borough names and neighbourhood names were allotted borough names under the neighbourhood column.

Feature Selection

Since our primary focus area is the restaurant. We filtered out the data returned by Foursquare API to include the Venue Category having the term “Restaurant”. Thus the data which were not required (Historical Monuments, Bakery, Cafe etc.) were dropped from the dataset.

Exploratory Data Analysis

Plotting of Neighbourhoods on a Map

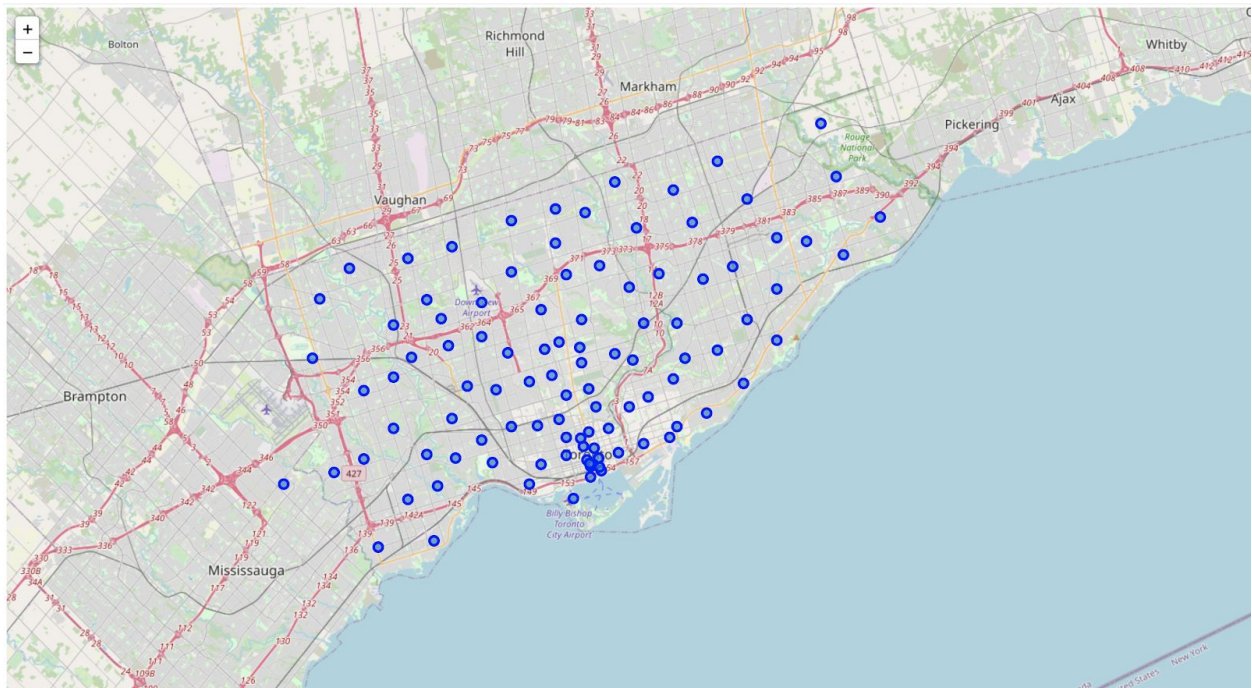


Fig c. Neighbourhood Plots

OneHotEncoding

OneHotEncoding was carried out on the data to create dummy variables for the Venue Category as it is a categorical variable.

	Afghan Restaurant	American Restaurant	Asian Restaurant	Belgian Restaurant	Brazilian Restaurant
1110	1	0	0	0	0
1079	1	0	0	0	0
1495	0	1	0	0	0
3153	0	1	0	0	0
1466	0	1	0	0	0

Fig d. OneHotEncoding

Classification of Restaurants Using KMeans Algorithm

The restaurants were classified using the KMeans Algorithm. They were clustered based on their cuisines.

Cluster 0: American, Vietnamese, Greek, Mexican etc.

Cluster 1: Italian

Cluster 2: Sushi

Cluster 3: Thai

Cluster 4: Steakhouse, Fish & Chips, Sandwiches etc.

Plotting of all Restaurants along with their Neighbourhoods on a Map

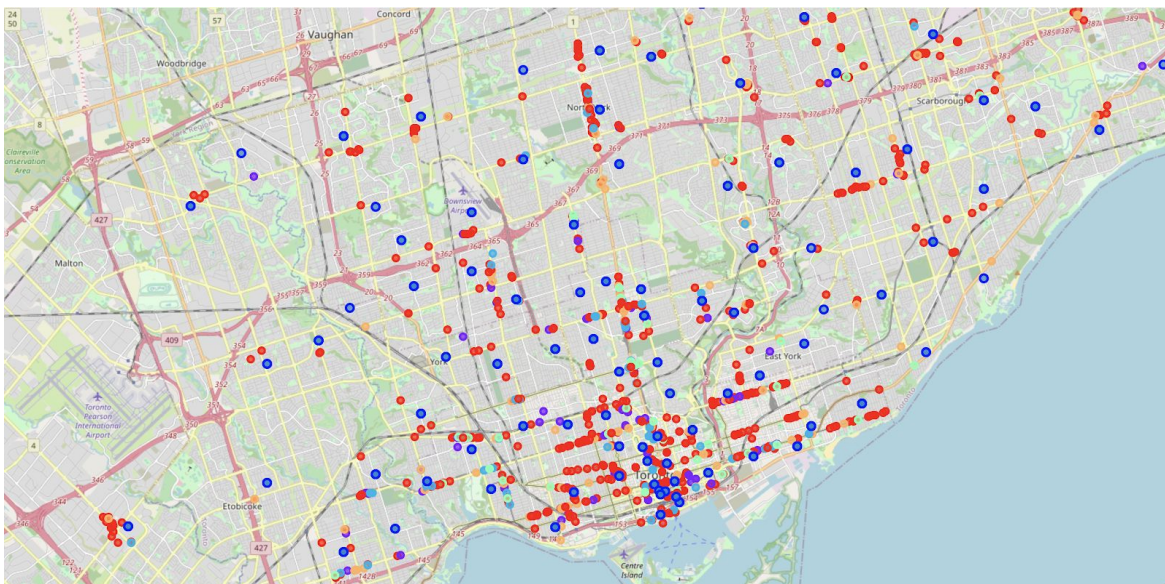


Fig e. All Cuisine Restaurants

Plotting of Cluster 2 (Sushi Restaurants) along with their Neighbourhoods on a Map

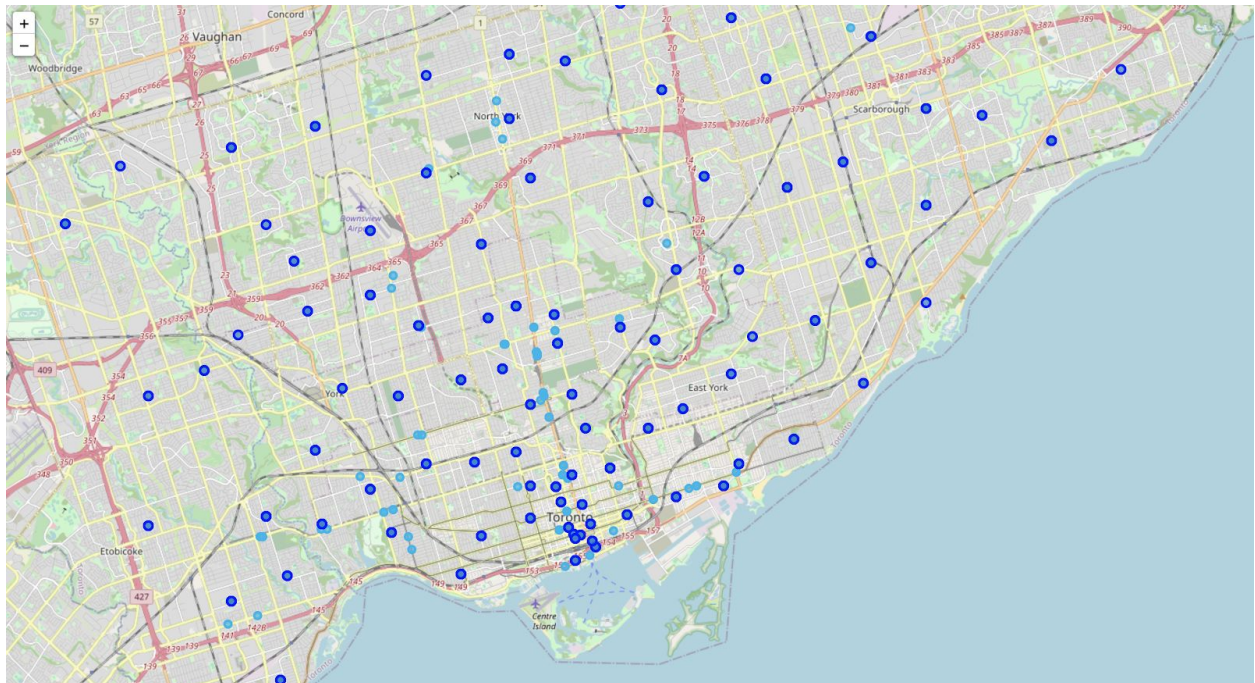
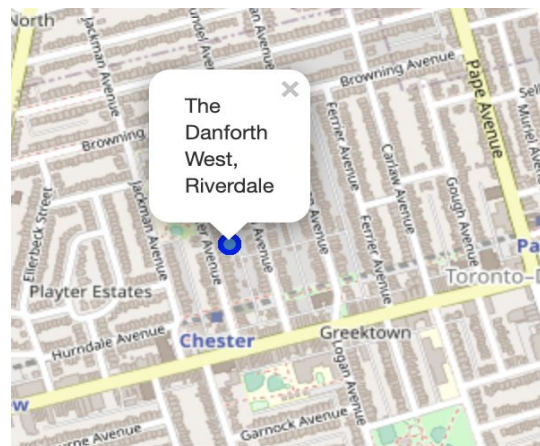
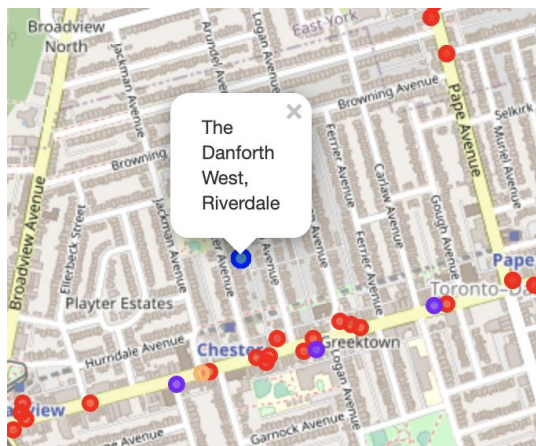


Fig f. Sushi Restaurants

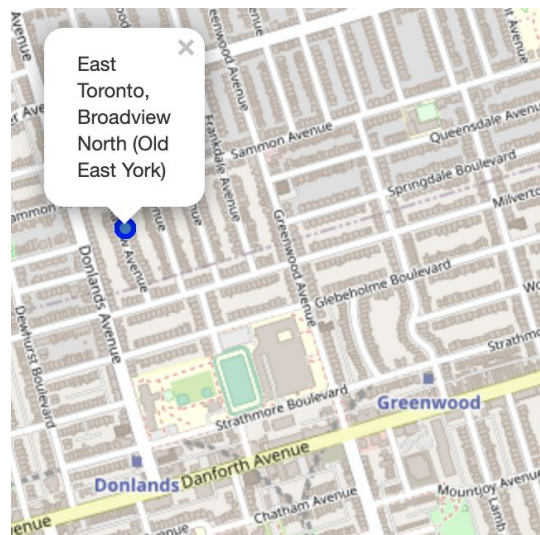
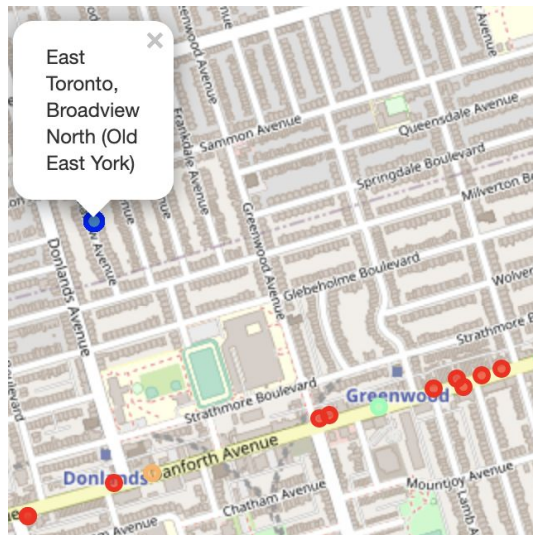
Results

Below are a few of the recommended neighbourhoods that seem to be a prime location for restaurants and we see that there is no competition to open a Sushi bar. On the left is a plot of all restaurants and on the right is a plot of the same area for Sushi restaurants.

Danforth West



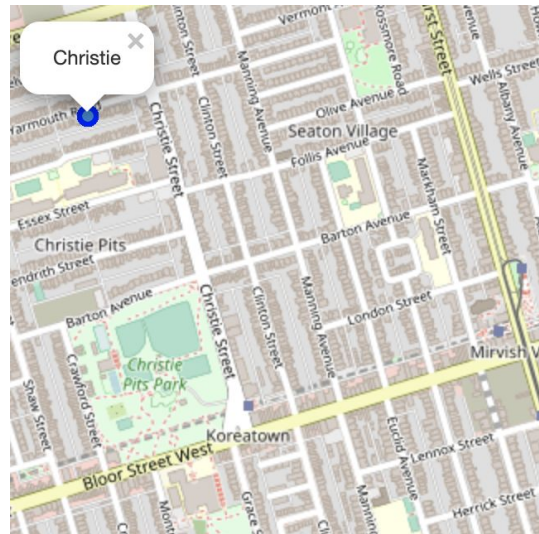
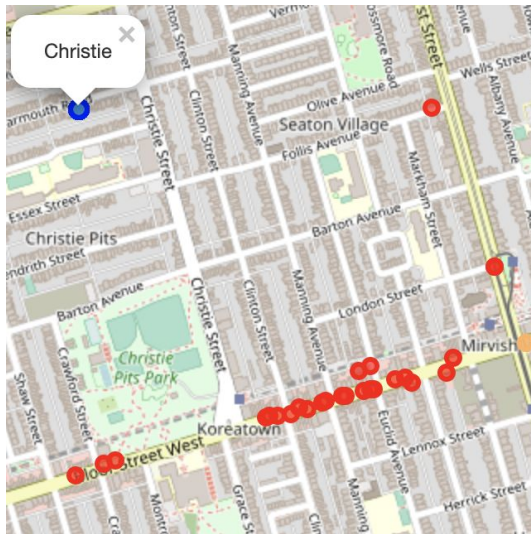
Broadview North



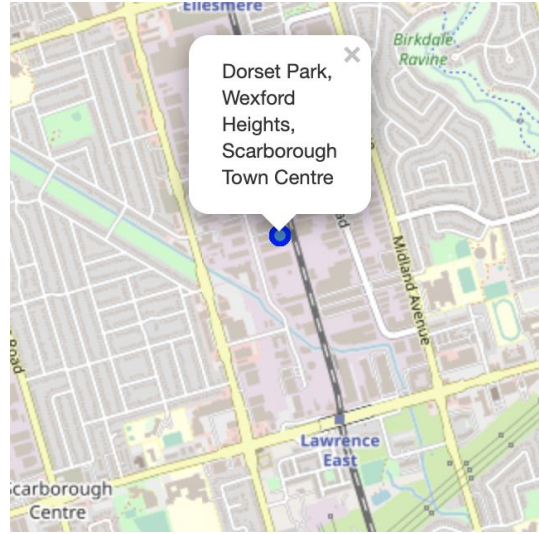
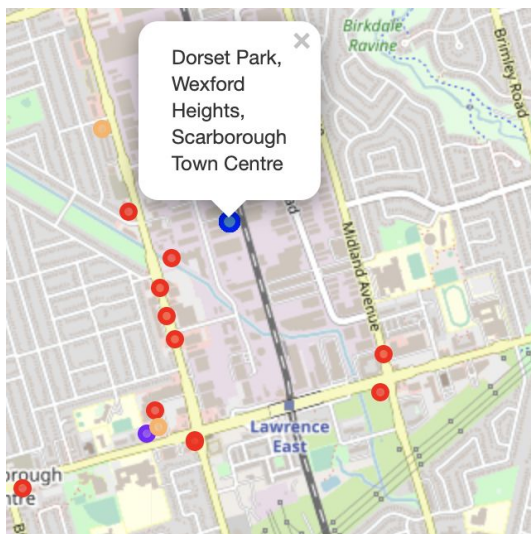
Commerce Court



Christie



Dorset Park



Conclusion

Based on the results and plots above; Danforth West, Broadview North, Commerce Court, Christie, and Dorset Park are the neighbourhoods suggested to Mr Kobayashi Maru to open his new Sushi Bar.