

Capstone Project: The Battle of Neighborhoods — High-End Caribbean Restaurant

 Kenson Gilles Jan 7 · 6 min read

Coursera Capstone Project — IBM Data Science Professional

As a Data Scientist equipped with the skills and the tools to use location data to explore a geographical location, we will have the opportunity to be as creative as we want and come up with an idea to leverage the Foursquare location data to explore or compare neighborhoods or cities of our choice or to come up with a problem that we can use the Foursquare location data to solve.

Business Problem

Opening a new restaurant has always been every chef's dream, whether you are new or experienced in the culinary field you're both driven by the same charismatic law which is **passion**. Despite cultural differences between cuisines, customers still enjoy and appreciate each flavor set on their tables, nothing brings people together like a good dish. However, such a dream comes with a lot of planning, challenges, and let's not get started on the cost. It can be costly, time-consuming starting a business in the wrong neighborhood, or where restaurants are over-saturated.

This is where we come in as a Data Scientist, experts in data analysis. The objective of this project is to analyze and find the best location in Toronto Canada to open a new High-End **Caribbean Restaurant**. Using Data Science tools and methodology we'll be able to determine where in the city would be the safest and most profitable neighborhood to open a Caribbean Restaurant. We will take a close look at each neighborhood in the city by analyzing the most common **venues**, **population**, and **criminal data** rate to help us determine the safest borough.

Data

Requirements

1. The list of Toronto's neighborhoods.
2. Toronto's neighborhood coordinates (Latitude & Longitude).
3. Toronto's crime rate data.
4. Neighborhoods venues data.

Sources

1. We'll use Canada's postal codes list found on [Wikipedia](#).
2. Foursquare API for restaurant data.
3. Geocoder for the coordinates.
4. Due to limited data available, we'll use the [Toronto Public Safety Data Portal](#) dataset to analyze the criminal history.

Methodology

1. Collect, explore, and process the neighborhood data into the DataFrame.

```
# Let's import our python packages
import pandas as pd
import numpy as np
import requests

# We'll fetch and clean Canada's postal code data
---
postalDf.head(10)
```

	Postal Code	Borough	Neighbourhood
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
8	M9A	Etobicoke	Islington Avenue, Humber Valley Village
9	M1B	Scarborough	Malvern, Brimley

9	M1B	Scarborough	Malvern, Rouge
11	M3B	North York	Don Mills
12	M4B	East York	Parkview Hill, Woodbine Gardens
13	M5B	Downtown Toronto	Garden District, Ryerson

Image by Kenson Gilles

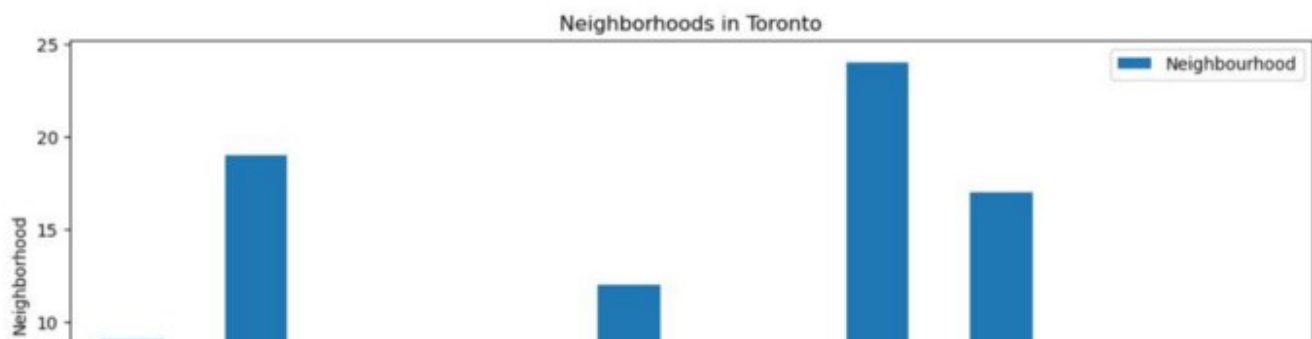
2. Assign the coordinates to each city accordingly using the geocoder.

```
# Using geocoder we'll assign the city coordinates
!pip install geocoder
---
postal_merge.head(10)
```

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
5	M9A	Etobicoke	Islington Avenue, Humber Valley Village	43.667856	-79.532242
6	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
7	M3B	North York	Don Mills	43.745906	-79.352188
8	M4B	East York	Parkview Hill, Woodbine Gardens	43.706397	-79.309937
9	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937

Image by Kenson Gilles

Very well, let's have a visual of how many neighborhoods are in each Borough.



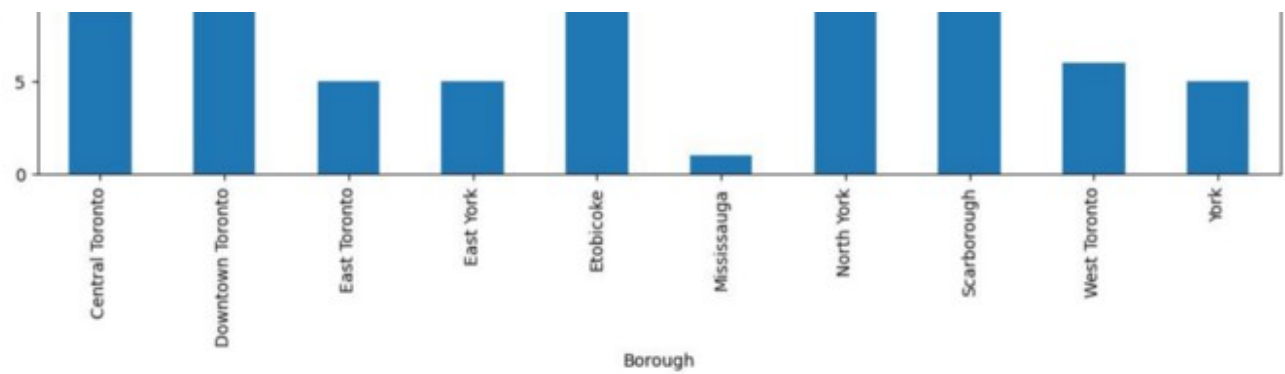


Image by Kenson Gilles

Based on our bar plot, North York seems to have the highest number of neighborhoods, following by Downtown Toronto in the second position. Let's continue by creating a map displaying our neighborhoods.

```
# Let's install and import folium to display our data on a map
!pip install folium
---
import folium
---

# create a map of Toronto using latitude and longitude values
map = folium.Map(location=[latitude, longitude], zoom_start=10)
```



3. We'll be using FourSquare API to retrieve all venues, and filter our data by looking for existing Caribbean restaurants.

```
# Using our FourSquare API credentials
CLIENT_ID = '<foursquare_client_id>'
CLIENT_SECRET = '<foursquare_client_secret>'
ACCESS_TOKEN = '<foursquare_access_token>'
VERSION = '20180604'
LIMIT = 30
---
```

```
# Search
search_query = 'Caribbean restaurant'
```

```
# Fetch venues from FourSquare
fsUrl = '<foursquare_api_url>'
fsReq = requests.get(fsUrl).json()
---
```

```
# From the returned JSON we'll retrieve our venues
venues = fsReq['response']['venues']
---
```

```
# Let's clean our data
# Filter by categories and remove anything not related to
# Caribbean Restaurant
filtered_columns = ['name', 'categories'] + [col for col in
fs_df.columns if col.startswith('location.')] + ['id']
---
```

```
fs_df_filtered = fs_df_filtered[(fs_df_filtered == 'Caribbean
Restaurant').any(axis=1)]
```

Let's display our clean data by cities, we'll be looking for the city with the most Caribbean Restaurant near Toronto.



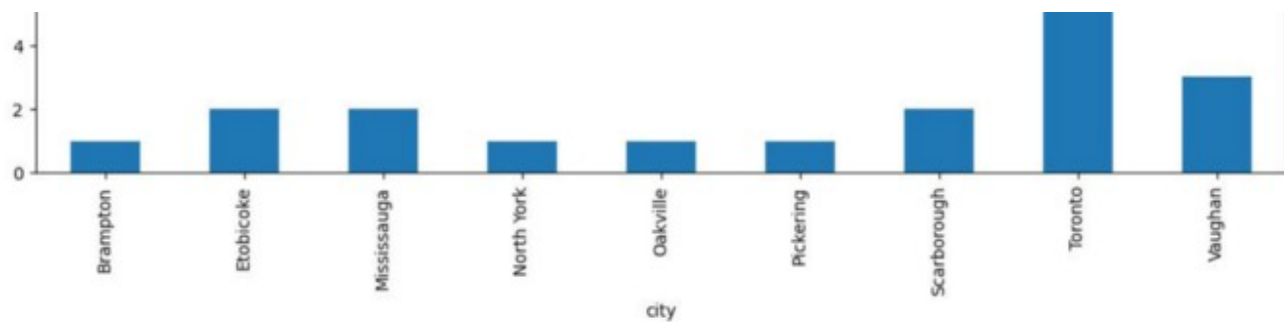


Image by Kenson Gilles

Clearly Toronto remains the city with the most Caribbean Restaurant, from this chart we can safely say Toronto is the most profitable city to open a Caribbean Restaurant. Let's dive deeper into our analysis.

Let's display a map with the Caribbean Restaurants in Toronto.

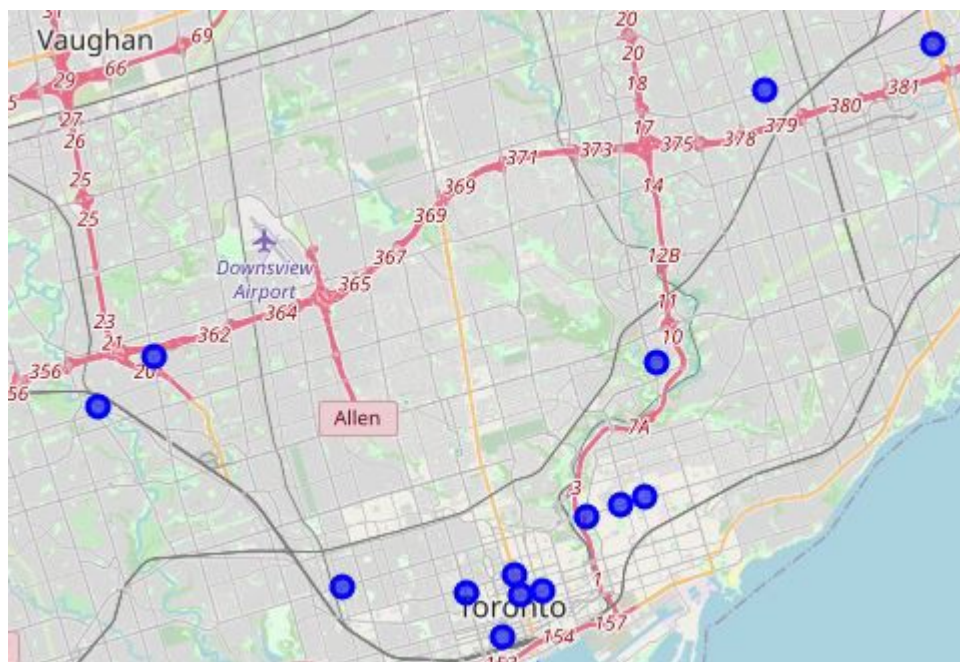


Image by Kenson Gilles

Most of them are located near Downtown Toronto, with a few scattered across. Perhaps the location, population, and neighborhood diversity have something to do with it.

As we all know, Downtowns from certain cities are not always the safest place to maintain a business, however, it can be the most profitable place to open one. They are well known for their criminal activities and overpopulated area.

4. We will now analyze each of Toronto's neighborhood crime rates from 2014 to 2018 using the **Toronto Public Safety Data Portal**, we'll then determine which neighborhood is the safest yet profitable to open and maintain a restaurant.

```
# Fetch Toronto's crime dataset
crimeUrl = '<toronto_public_data>'
crimeReq = requests.get(url=crimeUrl)

# Let's clean our data
# Rename and remove columns
# Retrieve only the average of all crime activities
# Add a new column summing the crimes average

toronto_crime.rename(columns = {
    ---
}, inplace = True)

toronto_crime.remove(---)
```

	Neighbourhood	Population	Assault	Auto Theft	Break Enter	Homicide	Robbery	Theft Over	Crime Rate
151	Milliken	26572	83.8	58.5	108.3	0.2	32.7	16.5	300.0
152	Pleasant View	15818	46.0	13.5	19.8	0.2	11.8	3.8	95.1
153	Wychwood	14349	70.2	13.2	34.0	0.3	13.8	2.3	133.8
154	Leaside-Bennington	16828	32.8	18.2	33.3	0.2	7.5	5.2	97.2
155	Briar Hill-Belgravia	14257	73.7	18.2	36.7	0.2	21.5	5.0	155.3

Image by Kenson Gilles

We currently have a record of 156 neighborhood crime to compare. Let us begin!

Assaults



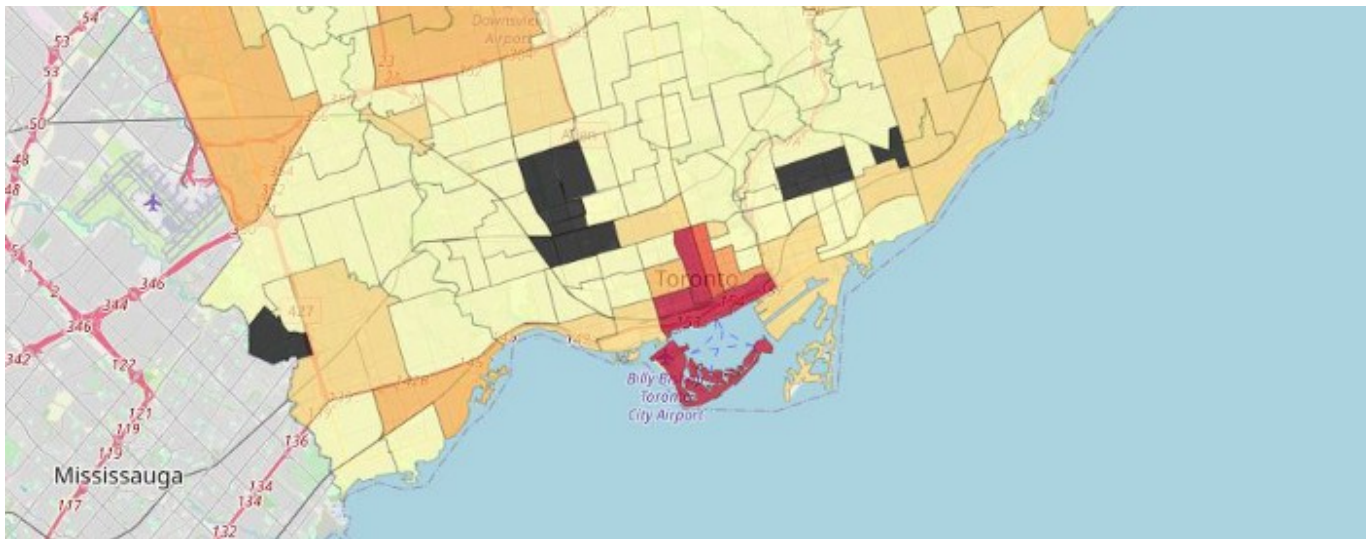


Image by Kenson Gilles (Toronto Assaults)

Auto Theft

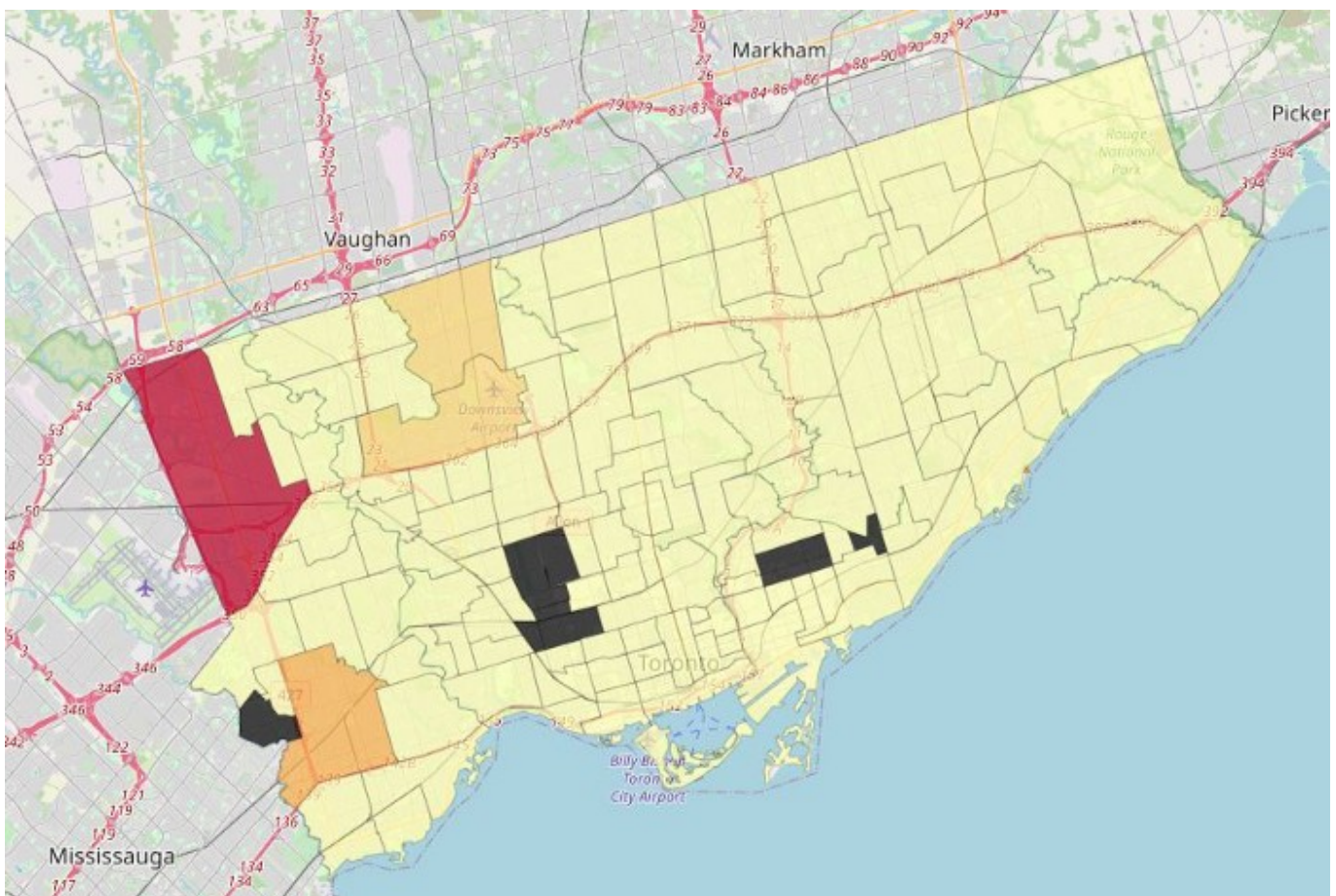


Image by Kenson Gilles (Toronto Auto Theft)

Break Enter



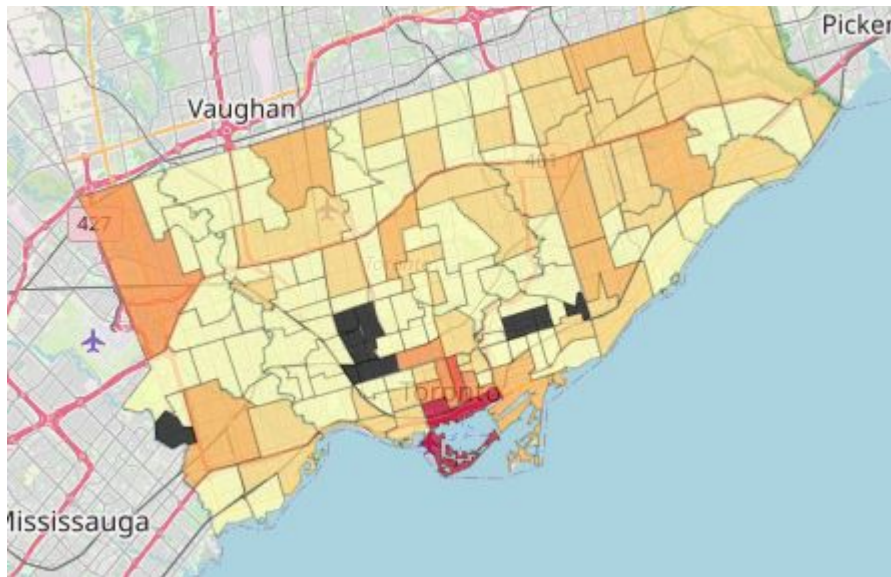


Image by Kenson Gilles (Toronto Break Enter)

Robbery

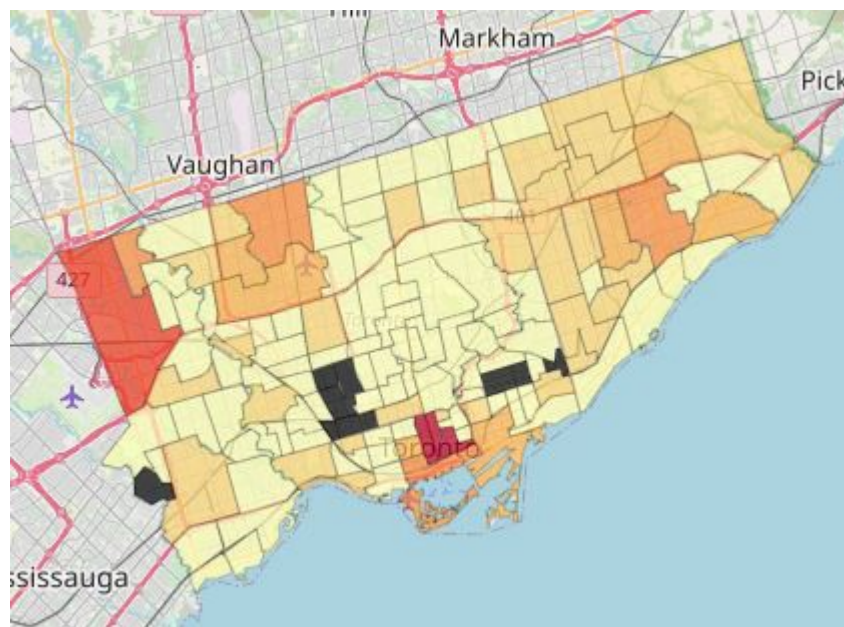
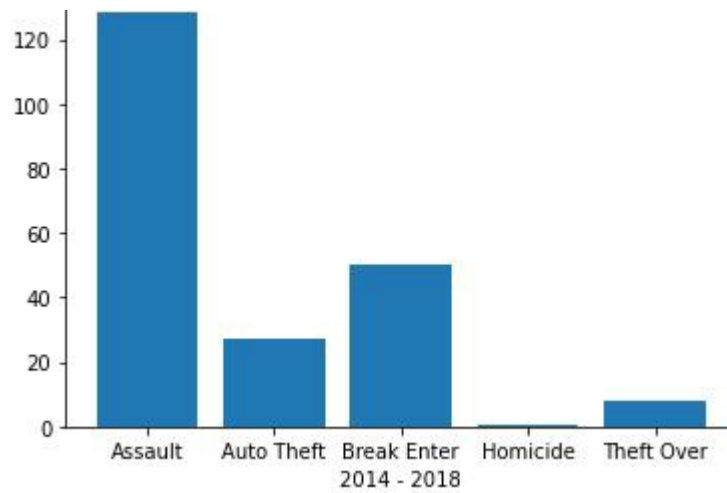


Image by Kenson Gilles (Toronto Robbery)

Which crime is the most common in Toronto? And will it prevent you from opening a business? According to the chart below 'Assault' is the most common crime in Toronto, we'll continue to investigate whether or not opening a business in Toronto is a high risk, and out of curiosity how many Caribbean Restaurants managed to maintain theres.

Average Crime by Category



5. We now have enough data to determine the most profitable and safest neighborhood in Canada. We will also use folium to display the data on a map.

Using the existing Caribbean Restaurant data retrieved from FourSquare, we'll create a map containing both the Restaurant and the neighborhood crime rate, starting from the safest neighborhood.

Neighborhoods with an average crime rate of less than 100

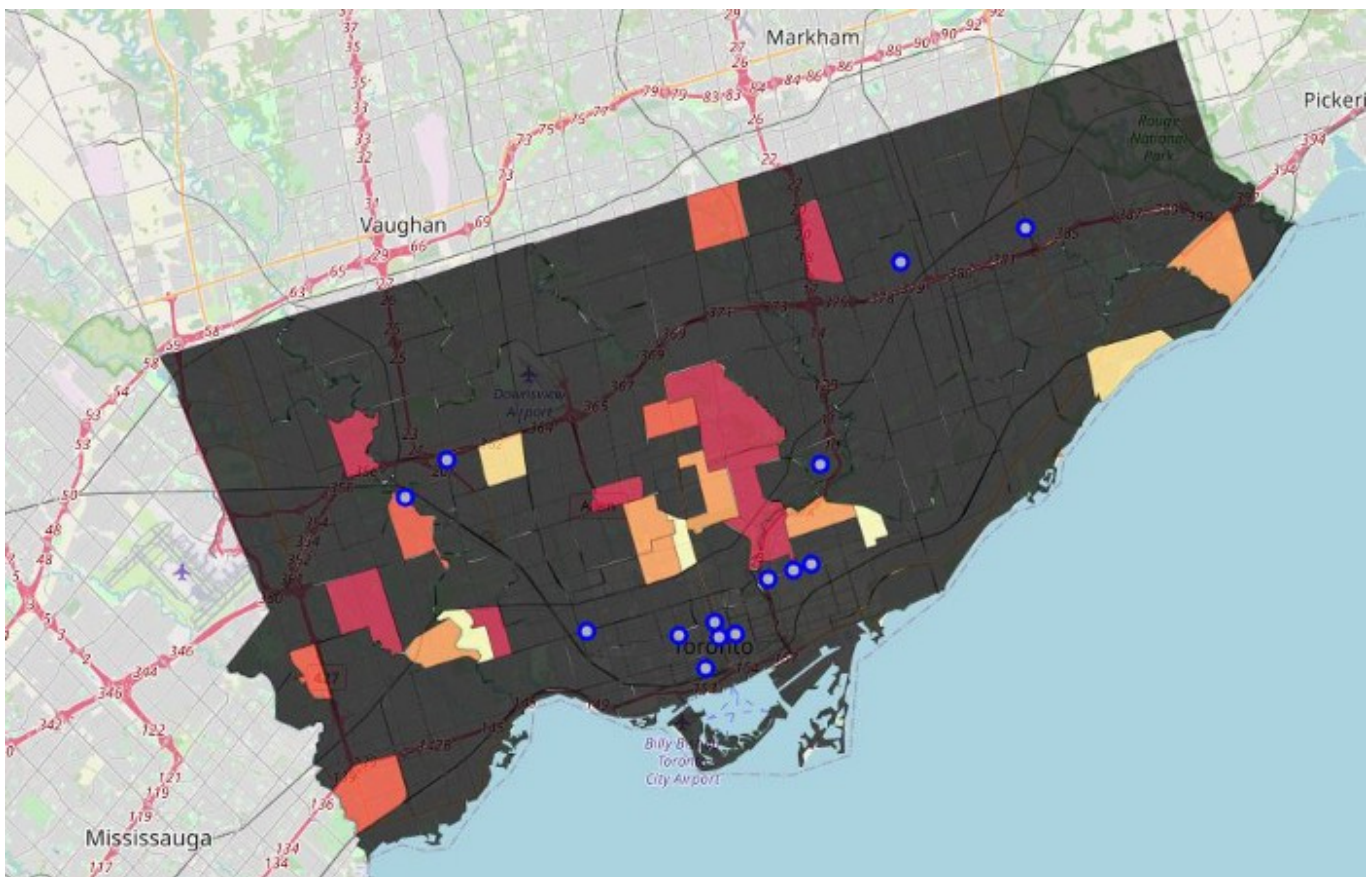


Image by Kenson Gilles

Neighborhoods with an average crime rate of above 100 and less than 500

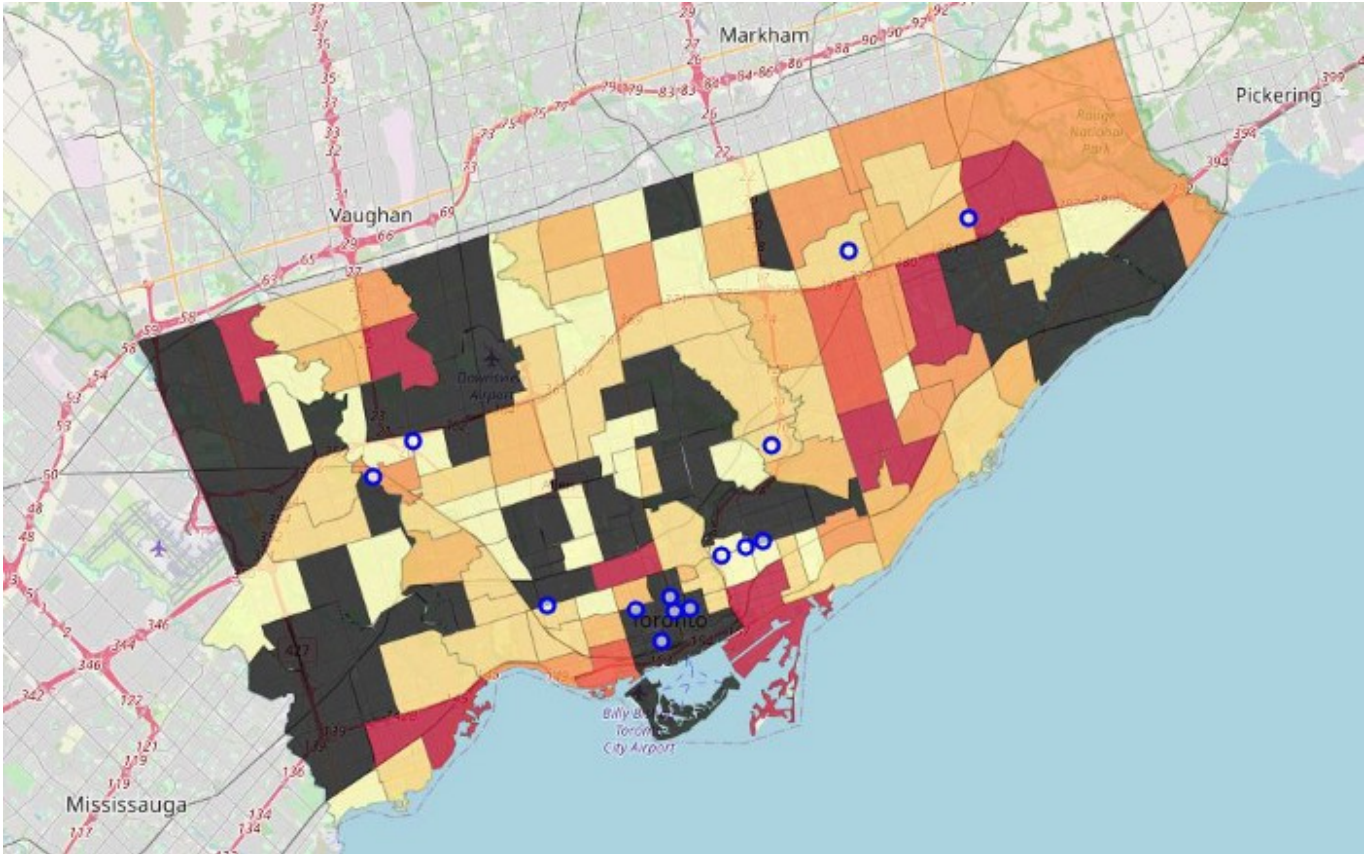


Image by Kenson Gilles

Neighborhoods with an average crime rate of above 500

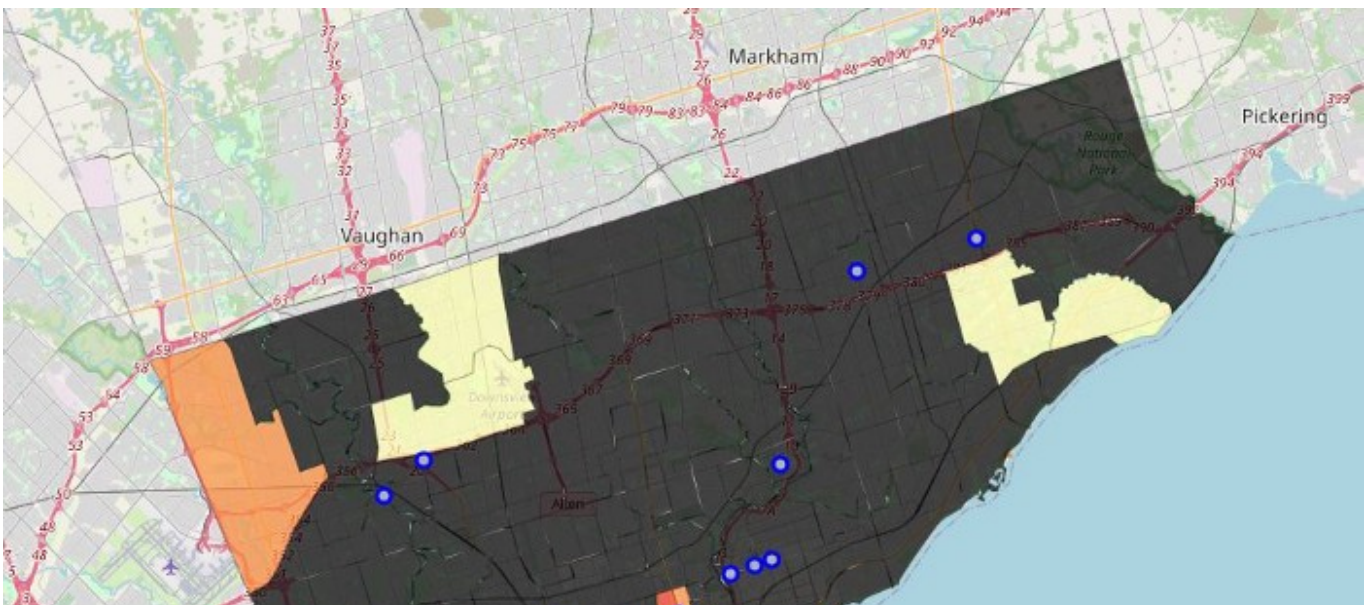




Image by Kenson Gilles

How about the population? Let's remind ourselves a restaurant needs people in order for it to thrive. :)

Let's take a look at the population.

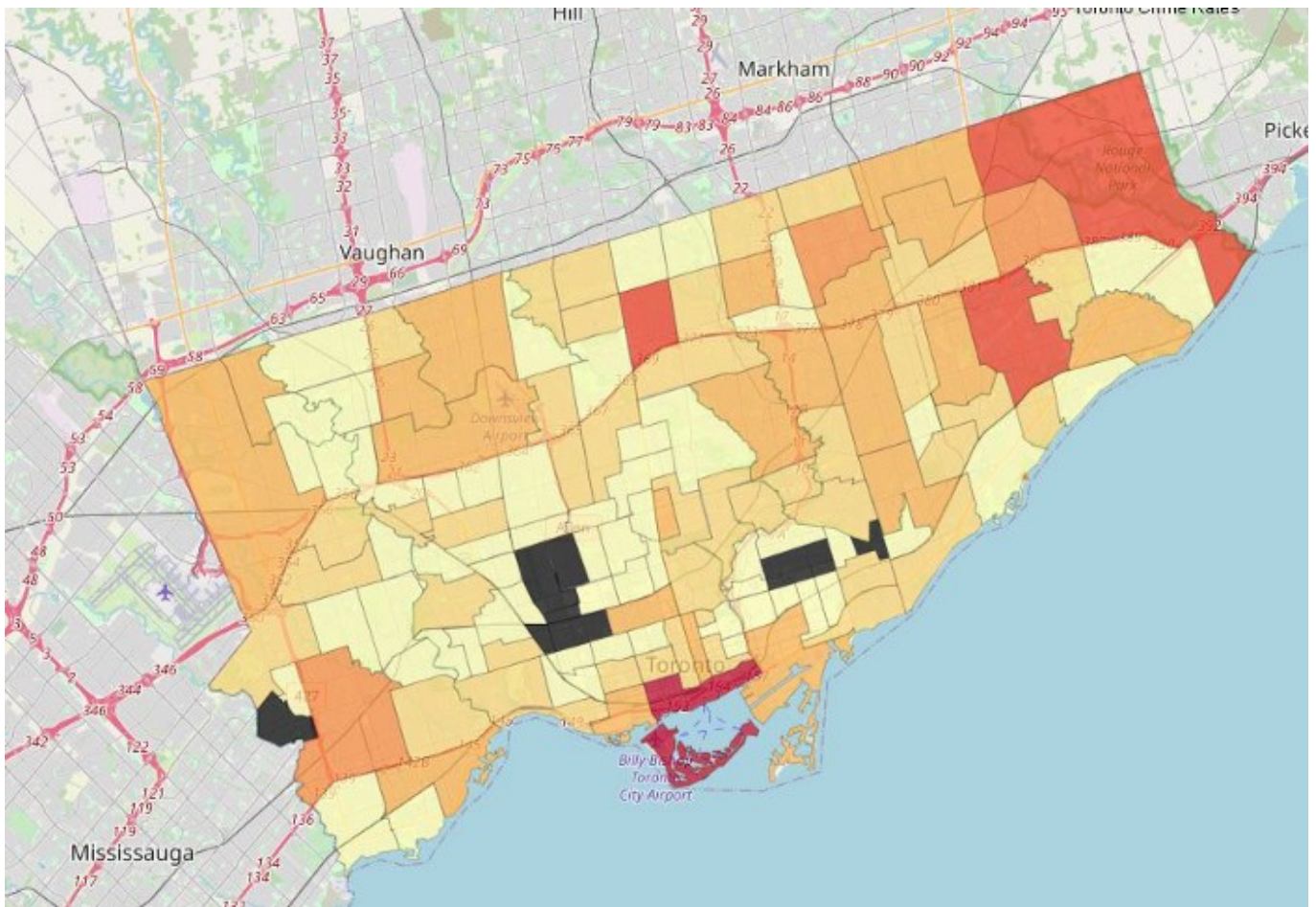


Image by Kenson Gilles

Results

From the above analysis, we clearly deduct that the safest neighborhoods to open a Caribbean Restaurant in Toronto are:

	Neighbourhood	Population	Assault	Auto Theft	Break Enter	Homicide	Robbery	Theft Over	Crime Rate
0	Yonge-St.Clair	12528	31.0	4.3	23.3	0.0	5.7	4.3	68.6
1	Yonge-St.Clair	12528	31.0	8.8	12.5	0.0	5.3	1.2	58.8
2	Lambton Baby Point	7985	31.0	4.3	23.3	0.0	5.7	4.3	68.6
3	Lambton Baby Point	7985	31.0	8.8	12.5	0.0	5.3	1.2	58.8
36	Centennial Scarborough	13362	46.5	9.5	19.2	0.3	7.5	2.0	85.0
44	Broadview North	11499	65.2	5.3	17.7	0.3	9.3	1.3	99.1
45	Princess-Rosethorn	11051	24.0	22.3	30.2	0.3	13.5	3.8	94.1
47	Etobicoke West Mall	11848	49.3	11.2	16.0	0.7	6.7	2.7	86.6
48	Forest Hill North	12806	30.8	15.0	36.3	0.0	9.7	2.3	94.1
54	Guildwood	9917	41.5	2.7	14.5	0.2	8.2	1.7	68.8

Image by Kenson Gilles

The safest and profitable borough to open a Caribbean Restaurant is in **East York**, because it is not over saturated with Caribbean Restaurants, it has one of the safest neighborhoods in Toronto with a decent amount of people.

However, the most profitable neighborhood to open a Caribbean Restaurant is in Downtown Toronto.

Discussion

It is very interesting to note they are more Chinese, Indian, and American restaurants in Toronto than Caribbean, although we did analyze Toronto's diversity group, but can safely say they are a lot of Caribbean descent currently leaving in the city.

We also noticed the lack of Caribbean Restaurant presence in other crowded areas with fewer crime rates, such as North York and Don Mills.

Conclusion

In this project, we analyzed over 100+ neighborhoods in Canada, the venues, crime rates, and population to determine where would be the ideal neighborhood to open a Caribbean Restaurant.