SAFEST NEIGHBORHOOD IN TORONTO FOR OPENING A COMMERCIAL ESTABLISHMENT

Fer Vázquez

November 2020

Introduction

- Background:

Toronto is a great place to live, the shopping is great, thousands of restaurants and cafes to get a fantastic meal, and there are lots of things you can do at any hour from strolling through parks, catching a movie or concert, or watching some live sports. But opening a business in Toronto isn’t always so good, especially if you ask about crime. Fortunately, if you want to open your business in Toronto, with this project we will be looking to understand the crime, and which will be the best neighborhood to open your own business.

Introduction

- Business Problem:

The purpose of this project is to understand which neighborhood will be the best to open a commercial business in Toronto and which type of commercial business. The first task will be to understand which neighborhood is the safest by analyzing the crime data and the second task will be to analyze the 10 most common venue in these neighborhoods. We will use our knowledge of Data Science to do this analysis.

Data Acquisition

- To fetch the crime details of Toronto I used real world data set published on Kaggle. Though this dataset included type of crime, recorded time and coordinates of the criminal activity along with neighborhood.
- The second source of data is based on data from a Wikipedia, them I merged it with dataset that contains the location data for each neighborhood.
- The third data source is generated from Foursquare API.

Methodology

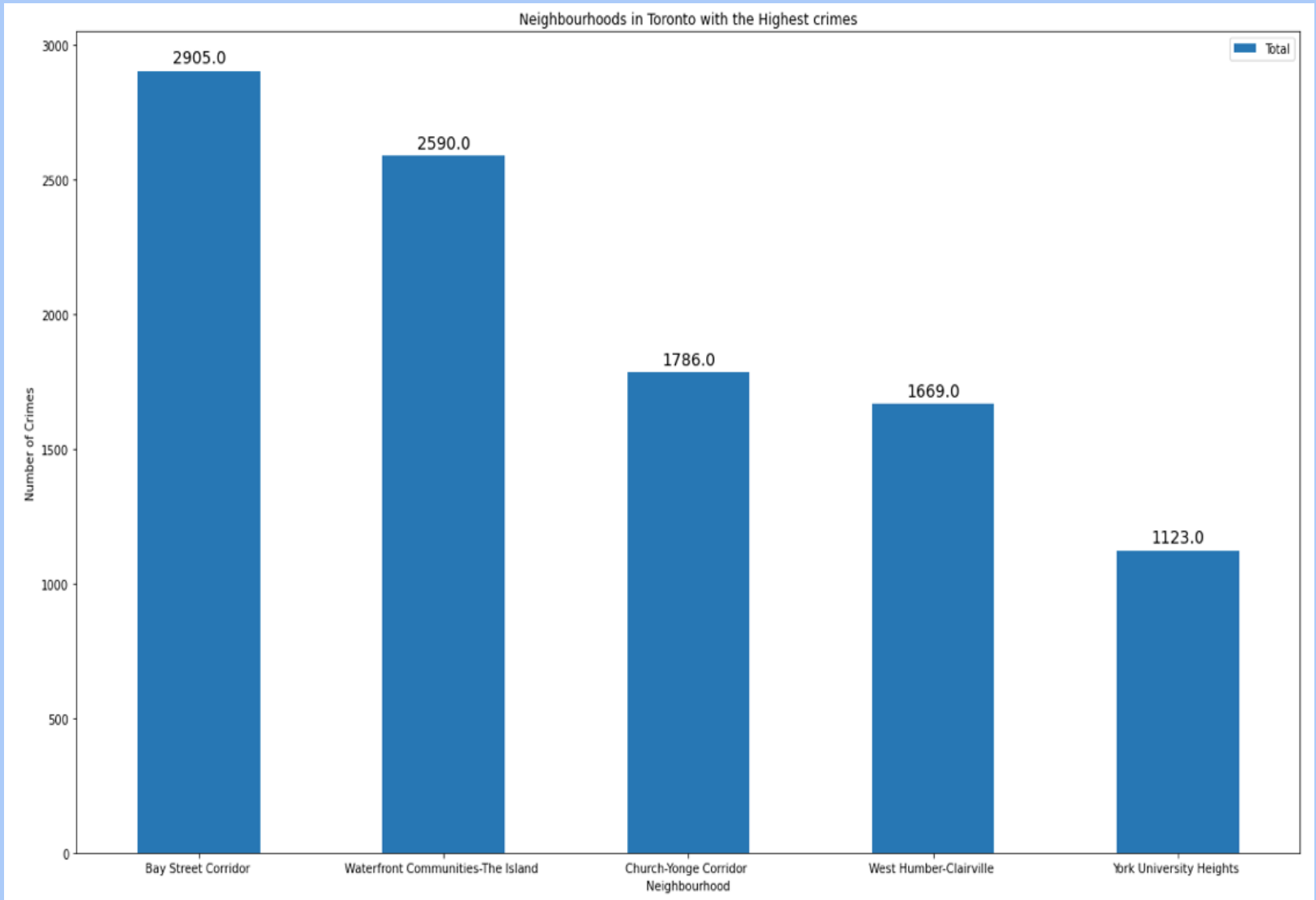
- Statistical summary of crimes:

```
[13]: df_crime_cat.describe()
```

[13]:

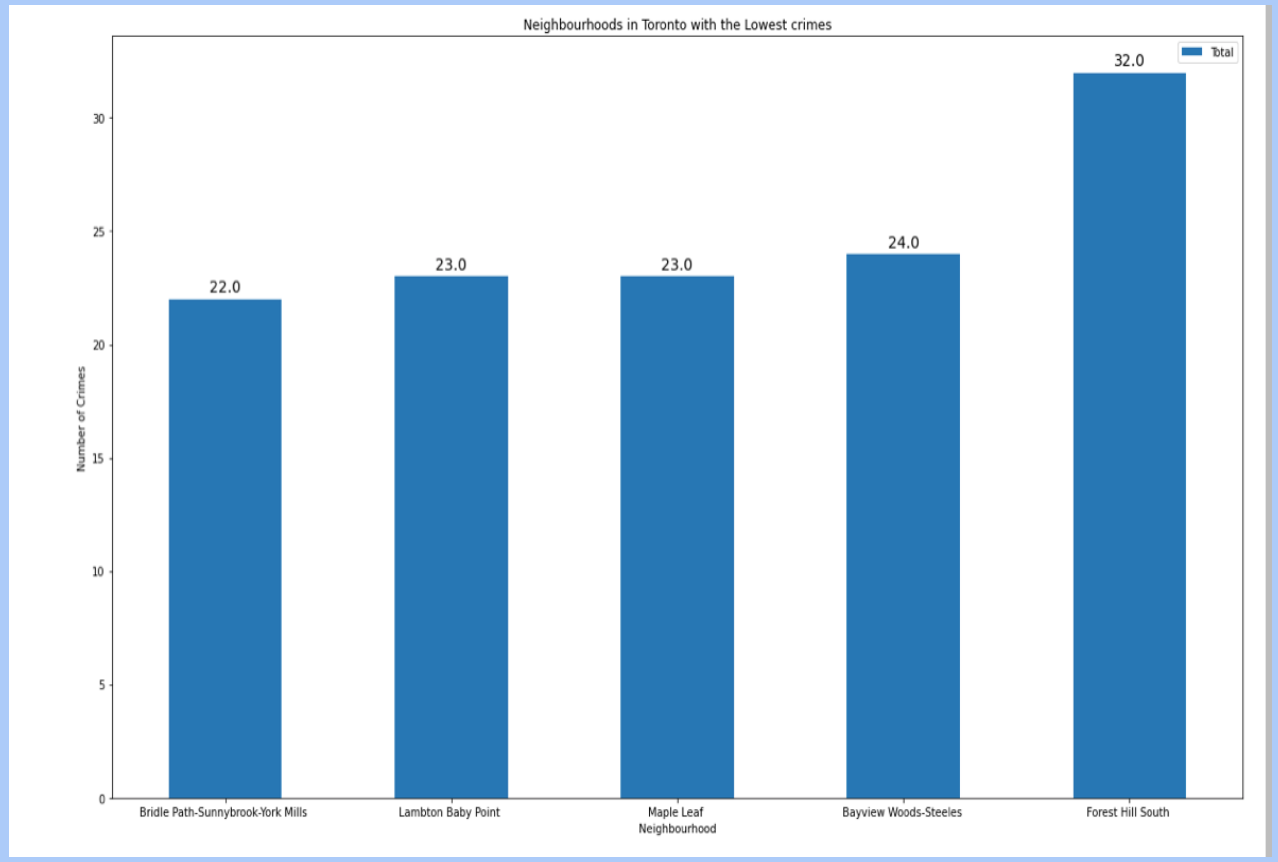
	occurrenceyearAssault	occurrenceyearAuto Theft	occurrenceyearBreak and Enter	occurrenceyearRobbery	occurrenceyearTheft Over	Total
count	141.000000	141.000000	141.000000	141.000000	141.000000	141.000000
mean	249.191489	33.673759	194.063830	71.985816	33.773050	582.666667
std	1485.352055	202.090379	1150.696201	425.861463	201.227496	3459.193320
min	3.000000	0.000000	5.000000	0.000000	0.000000	20.000000
25%	32.000000	2.000000	29.000000	14.000000	4.000000	86.000000
50%	66.000000	5.000000	61.000000	25.000000	7.000000	171.000000
75%	144.000000	16.000000	107.000000	43.000000	17.000000	291.000000
max	17568.000000	2374.000000	13681.000000	5074.000000	2381.000000	41078.000000

- The describe function in python is used to get statistics of the crime data, this returns the mean, standard deviation, minimum, maximum, 1st quartile (25%), 2nd quartile (50%), and the 3rd quartile (75%) for each of the major categories of crime



Data Visualizations

- Neighborhoods with the highest crime rates



DATA VISUALIZATIONS

Neighborhood's with the lowest crime rates.

Modeling

- With the safest neighborhoods dataset, we were able to find the venues within a 500 meter radius of each neighborhood by connecting to the FourSquare API. This returns a response in json format containing all the venues in each neighborhood which we convert to a pandas data frame. This data frame contains all the venues along with their coordinates and category will look as follows:

(58, 7)

[166]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	York Mills West	43.752758	-79.400049	Kitchen Food Fair	43.751298	-79.401393	Convenience Store
1	York Mills West	43.752758	-79.400049	Tournament Park	43.751257	-79.399717	Park
2	North Park, Maple Leaf Park, Upwood Park	43.713756	-79.490074	Rustic Bakery	43.715414	-79.490300	Bakery
3	North Park, Maple Leaf Park, Upwood Park	43.713756	-79.490074	Maple leaf park	43.716188	-79.493531	Park
4	North Park, Maple Leaf Park, Upwood Park	43.713756	-79.490074	Mika's Trim	43.714068	-79.496113	Construction & Landscaping
5	Runnymede, Swansea	43.651571	-79.484450	Coffee Tree Roastery	43.649647	-79.483436	Café
6	Runnymede, Swansea	43.651571	-79.484450	Bryden's Pub	43.649259	-79.484651	Pub

Results

- After running the K-means clustering the result is as follows:

[47]:

	Postcode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	M2P	North York	York Mills West	43.752758	-79.400049	0	Convenience Store	Park	Yoga Studio	Candy Store	Diner	Dessert Shop	Construction & Landscaping	Coffee Shop
1	M6L	North York	North Park, Maple Leaf Park, Upwood Park	43.713756	-79.490074	2	Bakery	Construction & Landscaping	Park	Yoga Studio	Candy Store	Diner	Dessert Shop	Convenience Store
2	M6S	West Toronto	Runnymede, Swansea	43.651571	-79.484450	1	Café	Coffee Shop	Sushi Restaurant	Pub	Pizza Place	Italian Restaurant	Yoga Studio	Dessert Shop
3	M2K	North York	Bayview Village	43.786947	-79.385975	1	Chinese Restaurant	Bank	Café	Japanese Restaurant	Yoga Studio	Diner	Dessert Shop	Convenience Store
4	M6G	Downtown Toronto	Christie	43.669542	-79.422564	1	Grocery Store	Café	Park	Athletics & Sports	Candy Store	Baby Store	Restaurant	Coffee Shop

Discussion

- The objective of the business problem was to help entrepreneurs identify the safest neighborhoods in Toronto, and an appropriate venue to set up a commercial establishment such as Grocery Store, Book Store, Gym, Fitness Yoga, etc. This has been achieved by first making use of Toronto crime data to identify safest neighborhoods for any business to be viable. After selecting the neighborhoods, it was imperative to analyze the venues in the neighborhoods. We achieved this by grouping the neighborhoods into clusters to assist the entrepreneurs by providing them with relevant data about venues and safety of a given neighborhood.

Conclusion

- We have explored the crime data to understand different types of crimes in all neighborhoods of Toronto, this helped us to know the safest neighborhoods first. Once we confirmed the safest neighborhoods, we analyzed neighborhoods based on the common venues, to choose a neighborhood and the type of commercial business which best suits the business problem.