Determining Business Prospects in Toronto

Introduction

Background

A new investor that just moved to Toronto, Ontario wants to invest in a new business but is not sure which is most profitable to him. Ontario is the second-largest province by total area in Canada. It is home to the nation's capital city, Ottawa, and the nation's most populous city, Toronto.

Business Problem

To help solve the problem, we will need to use machine learning algorithms to determine the most explored venues in a particular neighborhood. The business questions we are posed with are: How do we determine the venues in Ontario and what business do we recommend a business to the investor?

To solve this problem, we will make use of Foursquare Api to get the most explored venues in each neighborhood and depending on the neighborhood of choice of the investor, we will recommend a profitable business.

Data Section

For this project, we make use of the following data

1. Neighborhood in Toronto
   a. Source:
      https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
      - This data set contains required information to get the neighborhoods existing in Toronto
2. Longitude and Latitude of Each Neighborhood:
   a. Source:
      https://cocl.us/Geospatial_data/Geospatial_Coordinates.csv - This data set contains the longitude and latitude of each neighborhood in Ontario
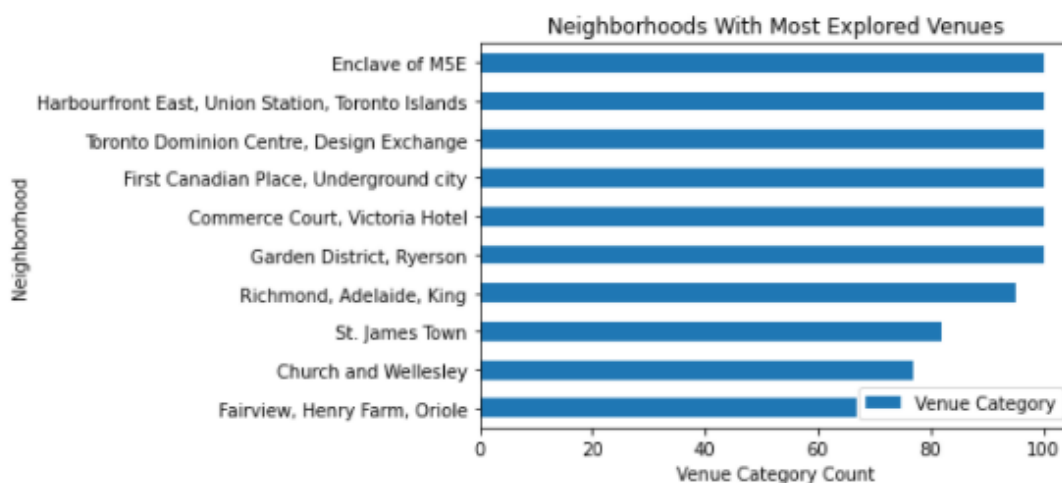3. Venues In Each Neighborhood

Methodology

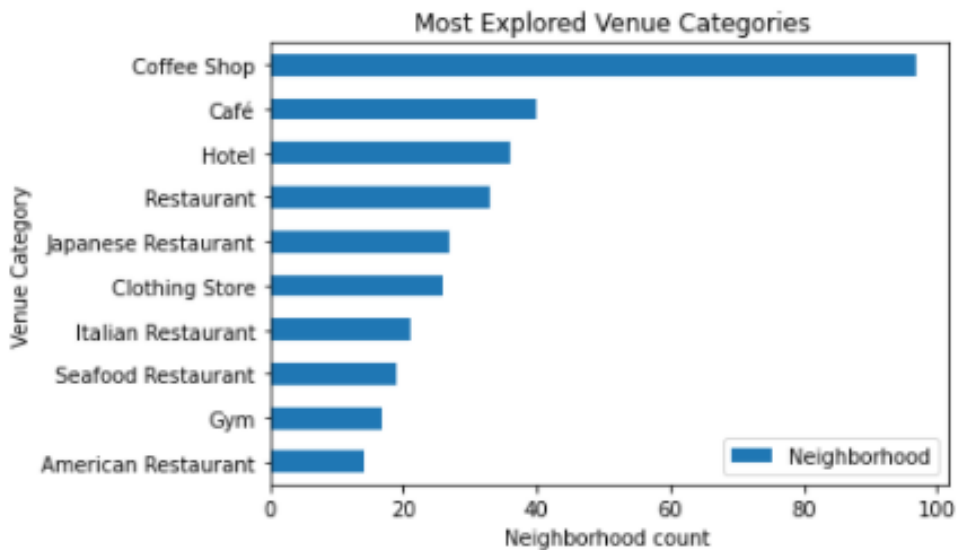The Methodology section includes:

1. 1.Collection of Inspection Data
2. 2.Exploration and Understanding Data
3. 3.Data preparation and pre-processing
4. 4.Modeling

Data Analysis and Visualization

The top 10 most explored Neighborhoods

The top 10 most explored venues in the top neighborhoods

Most Explored Venue Categories

Clustering Modelling

Using the KMeans Clustering algorithm and number of clusters set to 5, a clustering model was built to help determine the patterns in each neighbourhood.

Results and Discussions

We will examine our results in two aspects.

Our first exploration was to determine the most explored neighborhoods and the most explored venues in those neighbourhoods. From this, if our investor is looking to go into business in busy neighbourhoods, then we will recommend areas like: Church and Wellesley, First Canadian Place, Underground city, Enclave of M5E, Commerce Court, Victoria Hotel, Toronto Dominion Centre, Design Exchange and businesses like: coffee shop, cafe, hotel or restaurant.

Secondly, using the analysis of our clusters, the business prospect for our investor will depend on the cluster the neighborhood belongs to. For cluster 0, the most profitable business choice will be Cafeteria, Yoga Studio or Eastern European Restaurant. For cluster 1, the most profitable business choice will be baseball field, home service and yoga studio. For cluster 2, the most profitable business choice will be Park or Convenience Store. For cluster 3, the most profitable business choice will be Coffee Shop, Cafe or Pizza Place. For cluster 4, the most profitable business choice will be Cafeteria, Bar Yoga Studio.

Conclusion

In this study, we analysed the relationship between neighborhoods in Toronto and Exploration of Venues in those neighborhoods. We identified the most explored neighborhoods and the most explored venues. We developed Clustering models to determine the most profitable business prospects based on the neighborhood. This can be very useful in helping a business investor hoping to invest in or start a business in Ontario. For example, it could help identify what neighborhood to invest in and what type of business to invest in. With the help of the model, one can also recognize unexplored neighborhoods.