F

Finding the most common venues in Toronto

Saritha Aseen

7th July 2020

This report is for the capstone project of IBM Data Science Professional Certificate. In this project, the neighborhoods of Toronto are clustered into groups of similar neighborhoods. Foursquare API was used to retrieve data on the venues of the neighborhoods and their categories.

Business problem

ABC delivery service is an international delivery business. Their primary domain is courier and food delivery. They want to explore the possibility of striking partnership with health systems, grocery stores to expand their delivery business in these fields. Recently, the key business stakeholders decided to expand the delivery base to Toronto, the financial capital of Canada. ABC believes that convenience store and pharmacy delivery may leverage the drop in revenues faced from lesser restaurant food deliveries due to the COVID pandemic. They would like to analyze if convenience stores and pharmacies are among the major venue categories that exist in Toronto. This analysis will help them confirm if convenient store and pharmacies will be the right option to choose from to start their delivery diversifying project.

Data description

The Foursquare API is used to fetch and explore neighborhoods in Toronto. Then, I use the explore function to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters. Then, I use the *k*-means clustering algorithm to complete this task.

Finally, I use the Folium library to visualize the neighborhoods in Toronto and their emerging clusters.

The information I am focussing on are restaurants, convenience stores and pharmacies. We have neighborhood details about Downtown Toronto. We need to apply Neighborhood Segmentation and Clustering to analyze the neighborhood data and prioritize restaurants, convenience stores and pharmacies located in Toronto. Lastly, we can decide if Toronto is a good option to start with convenience store and drug store delivery.

Data Acquisition and Preparation

Firstly, we will take the processed Toronto data from week 3 but only include Downtown Toronto for this project, then explore the city by using Foursquare API and visualize the results separately.
Since clustering will be based on the categories of venues in the neighborhoods, we need data that specifies the venues in the neighborhoods and their categories. Of course we, in the first place, need a list of the neighborhoods of Toronto.

The figure below shows a map of Toronto with its neighborhoods represented as circles.

To acquire data on venues and their categories, Foursquare API is used. Foursquare is one of the world largest sources of location and venue data. To retrieve the venues and their categories in a given neighborhood, the coordinates—the latitude and the longitude—of the neighborhood are sent in the API request.

The API-request URL looks like the following:

```
https://api.foursquare.com/v2/venues/search?
&client_id=1234&client_secret=1234&v=20180605&
ll=40.89470517661,-73.84720052054902&radius=500&limit=100
```

Client_id and client_secret are credentials used to access the API service and are obtained when registering a Foursquare developer account, v indicates the API version to use, ll indicates the latitude and longitude of the desired location, radius is the maximum distance in meters between the specified location and the retrieved venues, and limit is used to limit the number of returned results if necessary.