

Introduction

A firm of medical supplies wants to open a new pharmacy in Toronto, and is interested in looking for new opportunities Scarborough.

Taking into account the density of this kind of commerces in the area, we will offer a set most likely boroughs to open their store.

This kind of study would be suitable for profiling any kind of store or venue in a sets of address, by means of geospatial tools such as Foursquare or Google Maps.

Data Section

We will focus on thos nighborhoods which are address in **North York, Toronto**. This locations will be provided by the postal code information settled in Wikipedia (source: [List of postal codes of Canada: M](#))

Geocoder was used to determinate latitude and longitude co-ordinates

The **Foursquare API** will be use to find locations and look for those area which lack of pharmacies in the nearbies.

The data retrieved contains information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Name of the venue e.g. the name of a store or restaurant
6. Venue Latitude
7. Venue Longitude
8. Venue Category

▼ Methodology

► Data Preparation - Wikipedia

The **postal information** from Wikipedia was downloaded and transform into a pandas dataframe, which had three columns:

- Postal Code
- Borough
- Neighborhood

Once done this, common extraction and data cleaning were done:

- Only the cells that have an assigned borough were processed; cells with a borough that is Not assigned were ignored.
- More than one neighborhood can exist in one postal code area.
- If a cell has a borough but a Not assigned neighborhood, then the neighborhood will be the same as the borough.

Using **Geocoder**, we retrieve the latitude and longitude for each neighborhood, and add them as columns. Our data frame will have five columns.

- Postal Code
- Borough
- Neighborhood
- Latitude
- Longitude

↳ 7 celdas ocultas

► Data Preparation - Foursquare

Using the dataset with the information of co-ordinates of the Neighbours, we will call the **Foursquare API** to get all the venues in a radius of 700m for each neighborhood.

The information is stored in a new dataframe with the columns

- Neighborhood
- Neighborhood Latitude
- Neighborhood Longitude
- Venue
- Venue Category

↳ 4 celdas ocultas

► Clustering over number of Pharmacies

With the results of our Foursquare API, we will get the number of Pharmacies that are in each neighborhood.

With this, we will clusterize the neighborhoods depending in the number of Pharmacies they have in ther nearbyes.

In addition, this information will be join to the main data set in order to make graphical analysis.

↳ 5 celdas ocultas

Results

Having analyzed **104 Neighborhoods**, **27 of them had at least a pharmacy** identified as venue in Foursquare.

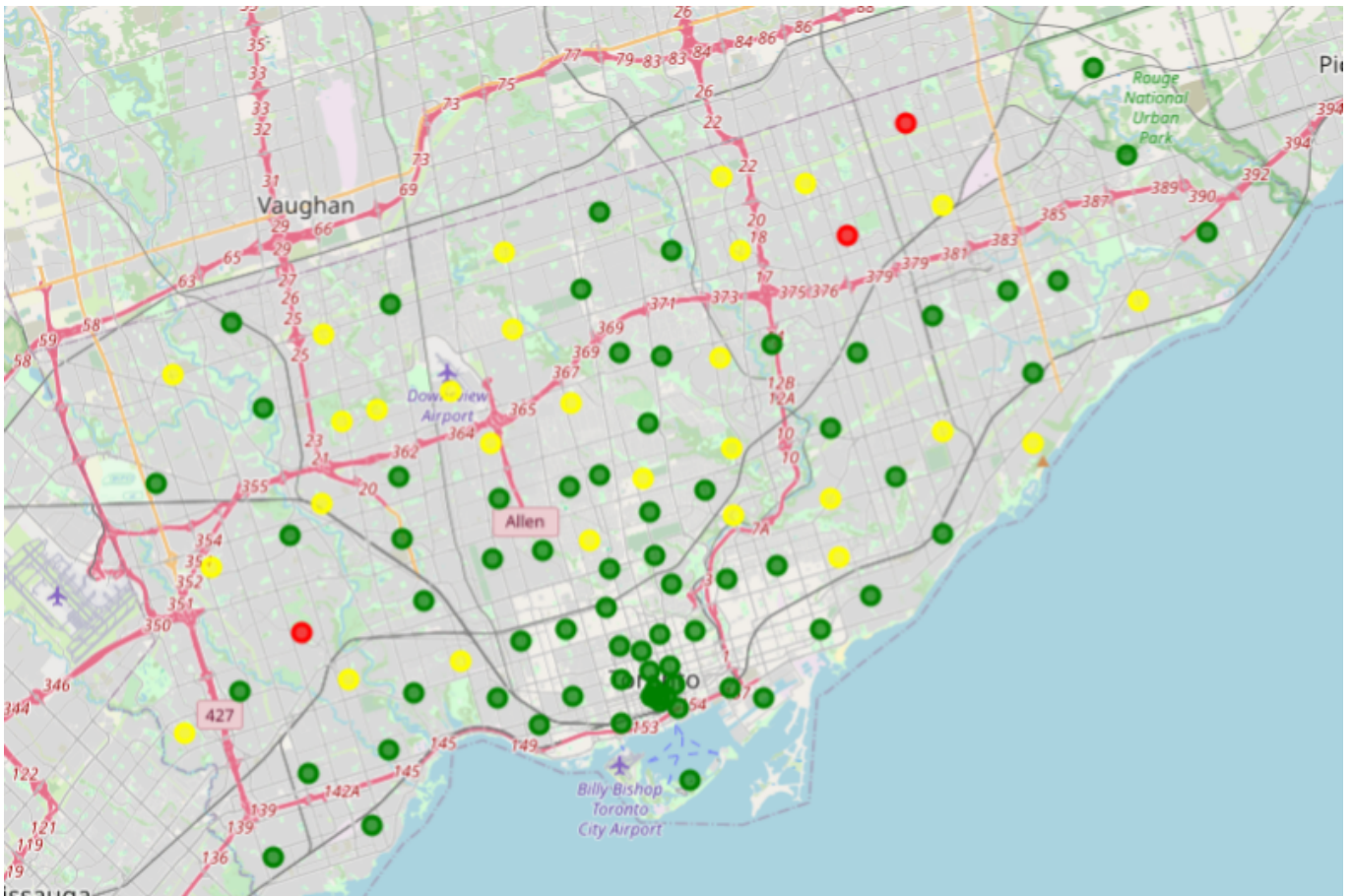
In addition, there are only three locations which have more than one Pharmacy:

- Clarks Corners, Tam O'Shanter, Sullivan
- Milliken, Agincourt North, Steeles East, L'Amoreaux East
- Islington Avenue, Humber Valley Village

Considering this, in order to get a clear view of the oportunities, we define three clusters of neighborhoods:

1. **Those with more than one pharmacy**
2. **Those with one pharmacy**
3. **Those with no pharmacies**

Representing the locations in Toronto map throws this results:



Having:

- in red, those with more than one pharmacy,
- in yellow, those with one pharmacy, and
- in green, those with no pharmacies.

Taking into account possible competitors, these are the number of stores of each firm in Scarborough:

Shoppers Drug Mart 23

Rexall 3

Cliffwood I.D.A. Pharmacy, Thorncrest Drug Store, Brimley Road Medical Center I.D.A., IDA Pharmacy, IDA High Park 1

Conclusion

According to the information from Foursquare, Scarborough offers a great opportunity to open pharmacies as there are very few of this stores.

Shoppers Drug Mart has over 23 stores, so should be taken into account when deciding the location to open.

