##### The Battle of Neighborhoods

##### Toronto neighborhood - Vegan restaurants

##### Author: Hércules Sant'Ana da Silva José

##### July/2020

##### **Summary**

##### **Introduction Section:**

##### Introduction where we discuss the business problem about vegan restaurants in Toronto and who would be interested in this project.

##### **Data Section:**

##### Data where we describe the data that will be used to solve the problem and the source of the data.

##### **Methodology section:**

##### Methodology section which represents the main component of this report where we discuss and describe any exploratory data analysis that we did, any inferential statistical testing that we performed, if any, and what machine learnings were used and why.

##### **Results section:**

##### Results section where we discuss the results.

##### **Discussion section:**

##### Discussion section where we discuss any observations we noted and any recommendations to our audience based on the results.

##### **Conclusion section:**

##### Conclusion section is where we conclude the report.

##### **1. Introduction**

##### **1.1. Scenario**

##### Recent studies suggest that about 8% of the world's population is vegan. Few people around the world are committed to living a plant-based lifestyle. If you are interested in eating more cleanly, do not want to explore animals or simply enjoy how you feel as a vegan, it can sometimes be difficult to find great vegan restaurants in the world.

##### However, this is starting to change. You will find that the main cities in the world are slowly but surely developing strong vegan cultures, from coffee shops and cooperatives to some of the best vegan restaurants in the world. In short, you now have more options than simply ordering the only salad on the menu.

##### Toronto is one of the most ethnic and culturally diverse cities in the world. Located on the shores of Lake Ontario, it has people from all over the world from different cultures and ethnicities resulting in a lot of delicious delights. No matter what type of cuisine you like, you will surely find great restaurants that serve delicious food for you. Now, if you're a vegan, you may be wondering, all of this is good, but what if I'm looking for wonderful vegan cuisine to enjoy in Toronto? Will I be able to find something good? Will I as a student find good vegan restaurants near universities? Which neighborhoods can a company that specializes in vegan products promote its products?

##### **1.2. Business problem**

##### Identify existing vegan restaurants in Toronto, and identify the best neighborhoods where vegan companies can partner with restaurants to promote their products or open new physical stores.

##### **1.3. Audience**

##### Vegan audience that comes to Toronto to come to work, study or tourism, and also companies interested in selling and promoting their vegan products.

##### **2. Data**

##### **2.1. Data description**

##### For this analysis, we will be required to explore, segment, and cluster the neighborhoods in the city of Toronto. However, the neighborhood data is not readily available on the internet. So, we will use a Wikipedia page that contains all the information we need to explore and cluster the neighborhoods in Toronto. The wikipedia page is available in <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>.

##### We will be required to scrape the Wikipedia page and process the data, clean it, and then read it into a Pandas dataframe so that it is in a structured format that we can do our analysis.

##### We will use the Foursquare API to identify the existing vegan restaurants in Toronto and any existing restaurants in each neighborhood with a distance up to 500 meters. We will use a list of boroughs and neighborhoods of Toronto with their coordinates (latitude and longitude) to obtain the venues of each boroughs and neighborhoods of Toronto. The coordinates dataset is available in <http://cocl.us/Geospatial_data>.

##### **2.2. Data preparation**

##### The data preparation process consisted of obtaining postal codes for the city of Toronto from Wikipedia page. All borough and neighborhood with value "Not assigned" were removed from the dataframe.

##### Then we obtain the geographic data of the Toronto neighborhoods. We merged the two dataframe, creating a new one with all neighborhoods and their respective coordinates (latitude and longitude).

##### The dataframe has 10 boroughs and 103 neighborhoods for our analysis.

##### After this, we check the geographic data collected by plotting the neighborhoods on the map. We used the Toronto coordinates as start point.

##### The map below shows all of Toronto's neighborhoods.



Fig. - Toronto's neighborhoods

##### **2.3. Checking Foursquare API data**

##### We invoke the Foursquare API to check the returned data. In this test we verify how many venues exists around the University of Toronto at a distance of up to 1000 meters.

##### We created a function to extract the venue categories from json returned from API.

##### I've cleaned the returned json and I've structured it into a pandas dataframe.

##### 15 venues were returned by Foursquare near University of Toronto.

##### The table below shows the contents of the dataframe created from the json returned by the Foursquare API.

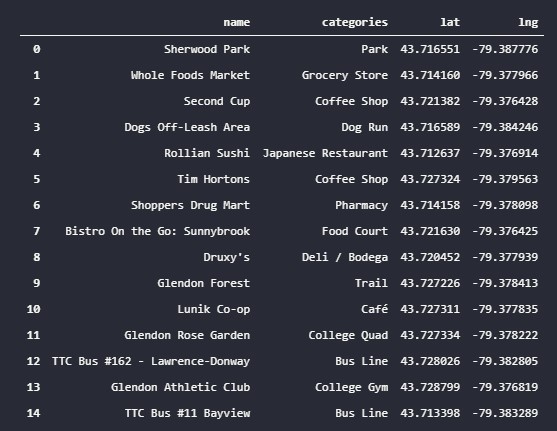


Fig. - Dataframe created from Foursquare API data

##### **3. Methodology**

##### We performed an exploratory analysis from the data returned by the Foursquare API. First we seek a specific category for vegan restaurants.

##### We then identified all other types of restaurants. Using a map of the city (provided by the Python library Folium), we plot all venues and visually compare the amount of vegan restaurants with the other types of restaurants.

##### Finally, we use the KMeans clustering technique to determine which neighborhoods are promising for companies to partner with local restaurants to sell vegan products.

##### We have created a function to facilitate the search for venues from the geographic data of the respective neighborhoods. We limited the search for venues to a distance of up to 500 meters. For performance reasons, the data collected was saved in a CSV file.

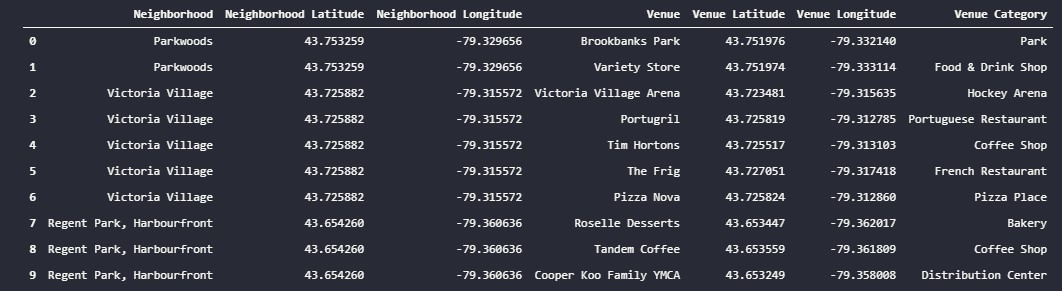


Fig. - Venues returned by Foursquare API

##### **4. Results**

##### Exploring our dataset, we identified a specific category for vegan restaurants called "Vegetarian / Vegan Restaurant".

##### When we counting all venues in the category, we found that there are only 18 venues registered.

##### In contrast, when we look for the word 'restaurant' in the description of the venues categories, we found 47 different categories.

##### Before applying the KMeans clustering technique, we grouped the neighborhoods together and created a dataframe containing only the venues within the 'Restaurant' category.

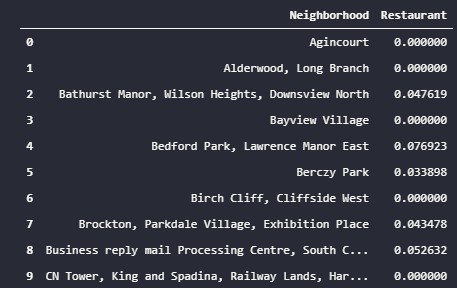


Fig. - Neighborhood grouped

##### We applied the K-Means clustering technique, where the neighborhoods are grouped into three clusters. The labels created were associated with each neighborhood contained in the dataframe.

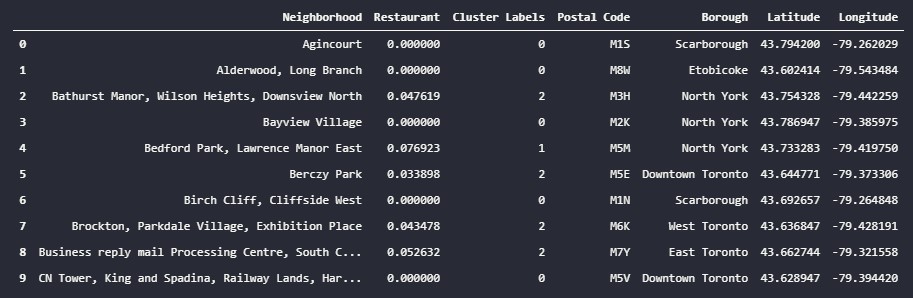


Fig. - Cluster labels associated with each neighborhood