BATTLE OF THE NEIGHBORHOODS

Toronto

\*\*\*\*\*\* CAPSTONE PROJECT \*\*\*\*\*\*

**Kiah Wenger July 2020**

Capstone Project

Battle of the Neighborhoods

Table of Contents

Intro

* Business Problem
* Target Audience

Data Requirements

Methodology

Results

Discussion

Conclusion

**Introduction**

Business Problem

Moving is a big deal - you want to make sure everything you want in a neighborhood is close by. To solve this problem my project will be running a search engine to set criteria and search cities and neighborhood that fits their lifestyle. To do this we'll review foursquare data then perform segmentation and clustering. Finally taking the data from Foursquare we'll then segment areas Toronto based on the most common places returned.

For this project, I will be developing a recommendation system using Toronto as my search criteria.

Target Audience

People moving to a new neighborhood. Additionally entrepreneurs could also use the data for deciding what restaurants opportunities are in each neighborhood.

**Data**

First, the location data consists of the data used for this project will be acquired from Source - (<https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>). The datasets consists of the postal codes and suburb names.

Second, the Foursquare API will use Foursquare search feature [FOURSQUARE]. We will explore several neighborhood's venue data & details using longitude and latitude details.

**Methodology**

The project will use web APIs like FOURSQUARE as well as python packages to create maps, extract insights and provide helpful results as we examine the neighborhoods.

Packages includes:

Pandas - Library for Data Analysis

NumPy – Library to handle data in a vectorized manner

JSON – Library to handle JSON files

Geopy – To retrieve Location Data

Geocorder - For geolocation of neighborhoods

Requests – Library to handle http requests

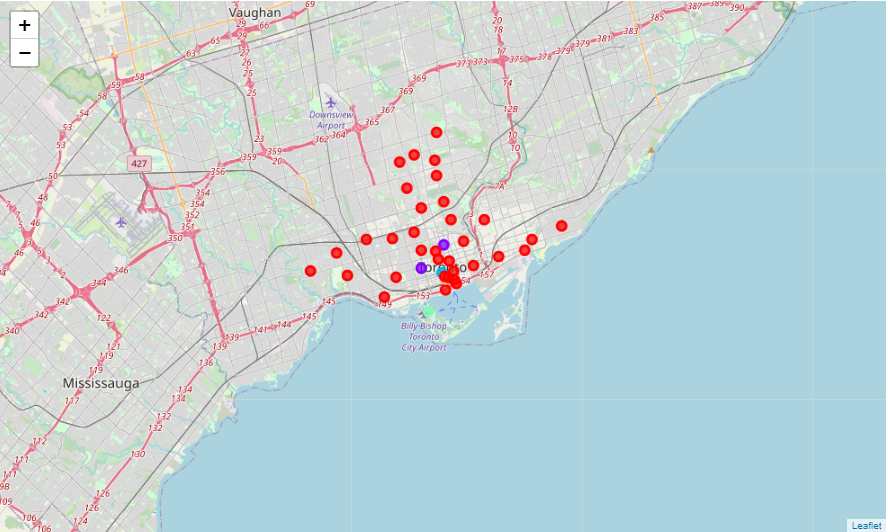
Matplotlib – Python Plotting Module

Sklearn – Python machine learning Library

Folium – Map rendering Library

**Results & Discussion**

As we review the clusters we can see clusters 2, 3 & 5 have coffee shops as the most common venue while cluster 4 has mostly airport related venues. Having this cluster option layout will allow anyone moving to a new neighborhood to decide which area has the most venue's they are looking for - such as cluster 3 having gyms and cafes which would appeal more to a younger crowd.



**Conclusion**

In this report we've reviewed the neighborhoods in Toronto and constructed a layout of the areas by the most common venues. Assuming the researcher is a younger person seeking gyms and coffee shops the neighborhood in cluster 3 would be the best pick. Meanwhile a commuter that flies for work would consider the neighborhood in cluster 4 with its proximity to the airport. Overall, this data will be very useful for anyone moving to a new area with a specific preference for what sort of venues they want readily available.