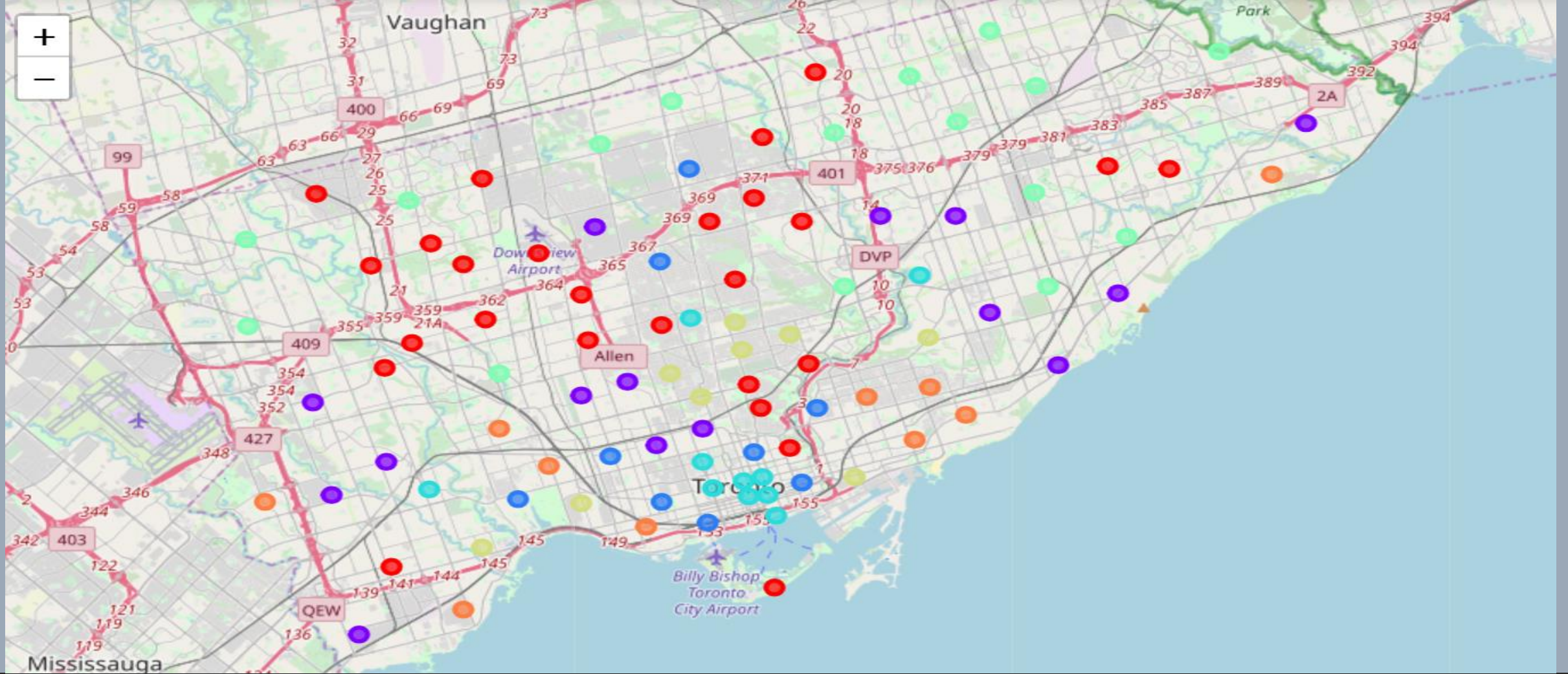# The Battle of the Neighborhoods

SEGMENTING NEIGHBORHOODS IN TORONTO FOR ANYWORDS BOOK VENTURES

INTRODUCTION

- Toronto City overview for AnyWords SAS

- AnyWords Strategic Plan and Overview

- Business problem:

1. Choice of initial neighborhoods to settle the first bookshop

2. The neighborhood should enable an easy replication

3. High demand for services

4. Low competition

- **Success Criteria: recommendation that meets the above points**

SEGMENTED NEIGHBORHOODS OF THE CITY OF TORONTO

DATA DESCRIPTION

Data requirements:

- Neighborhood information (names, coordinates, populations).

Data Sources:

- Wikipedia pages

- Foursquare.com

	Neighbourhood	Population	Land Area	Density	Population %	Income	Commuting	2nd Language	2nd Language %	Latitude	Longitude
1	Agincourt	44577	12.45	3580	4.6	25,750	11.1	Cantonese (19.3%)	19.3% Cantonese	43.788	-79.2839
2	Alderwood	11656	4.94	2360	-4.0	35,239	8.8	Polish (6.2%)	06.2% Polish	43.6035	-79.5464
3	Alexandra Park	4355	0.32	13,609	0.0	19,687	13.8	Cantonese (17.9%)	17.9% Cantonese	43.6498	-79.4015
4	Allenby	2513	0.58	4333	-1.0	245,592	5.2	Russian (1.4%)	01.4% Russian	43.7077	-79.4127
5	Amesbury	17318	3.51	4,934	1.1	27,546	16.4	Spanish (6.1%)	06.1% Spanish	43.7011	-79.481
6	Armour Heights	4384	2.29	1914	2.0	116,651	10.8	Russian (9.4%)	09.4% Russian	43.7454	-79.4226
7	Banbury	6641	2.72	2442	5.0	92,319	6.1	Unspecified Chinese (5.1%)	05.1% Unspecified Chinese	43.7491	-79.3664
8	Bathurst Manor	14945	4.69	3187	12.3	34,169	13.4	Russian (9.5%)	09.5% Russian	43.7627	-79.4563
9	Bay Street Corridor	4787	0.11	43,518	3.0	40,598	17.1	Mandarin (9.6%)	09.6% Mandarin	43.6567	-79.3835
10	Bayview Village	12280	4.14	2,966	41.6	46,752	14.4	Cantonese (8.4%)	08.4% Cantonese	43.7782	-79.3828

METHODOLOGY

Data scraping

- Scraping data and structuring it
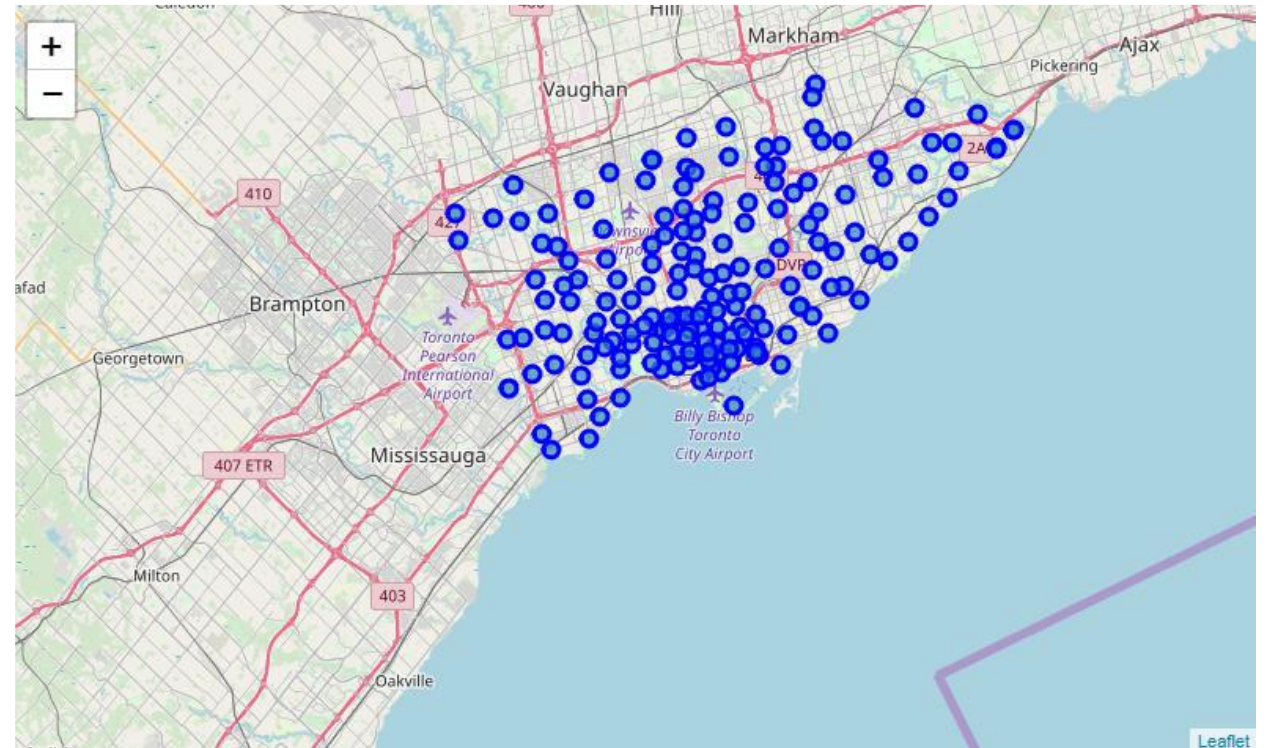
Foursquare.com

- Retrieving venue information from API

K-Means ML algorithm

- Population count as 50% weightage

- Venues have another 50%

Scoring Weightage

- Population count as 50% weightage

- Venues have another 50%

RESULT

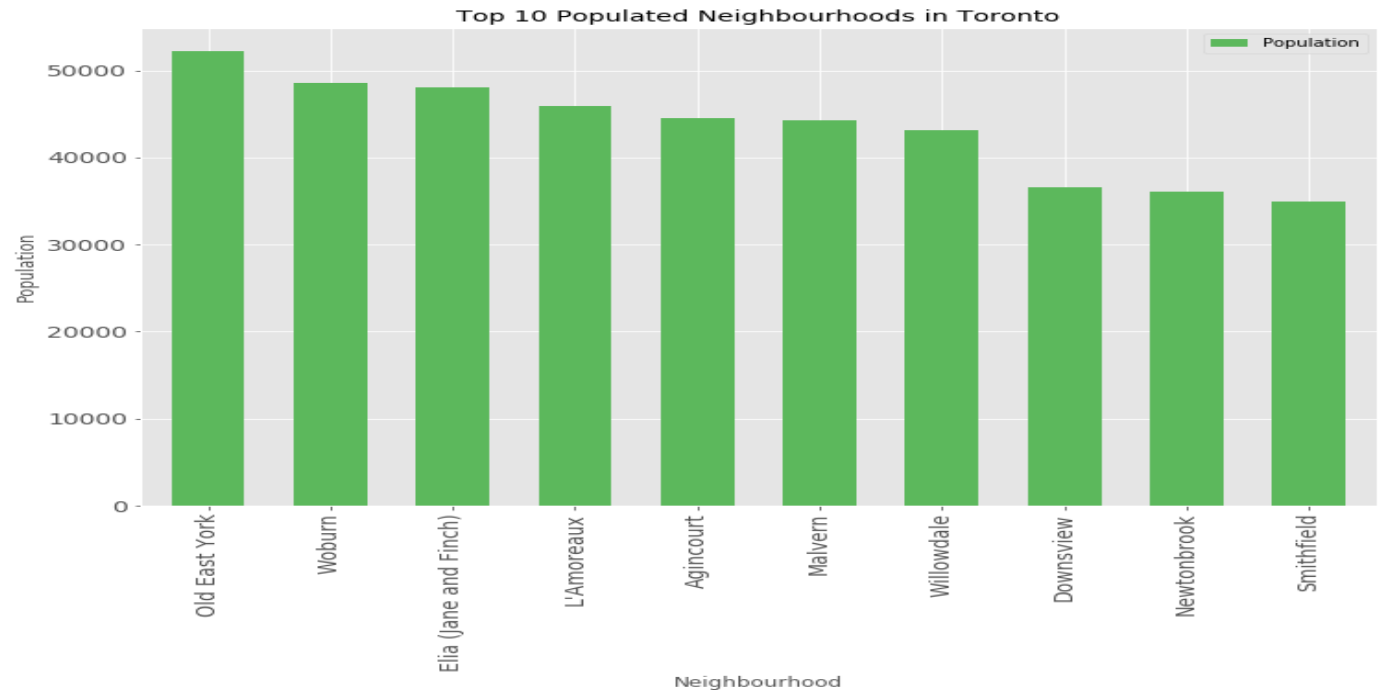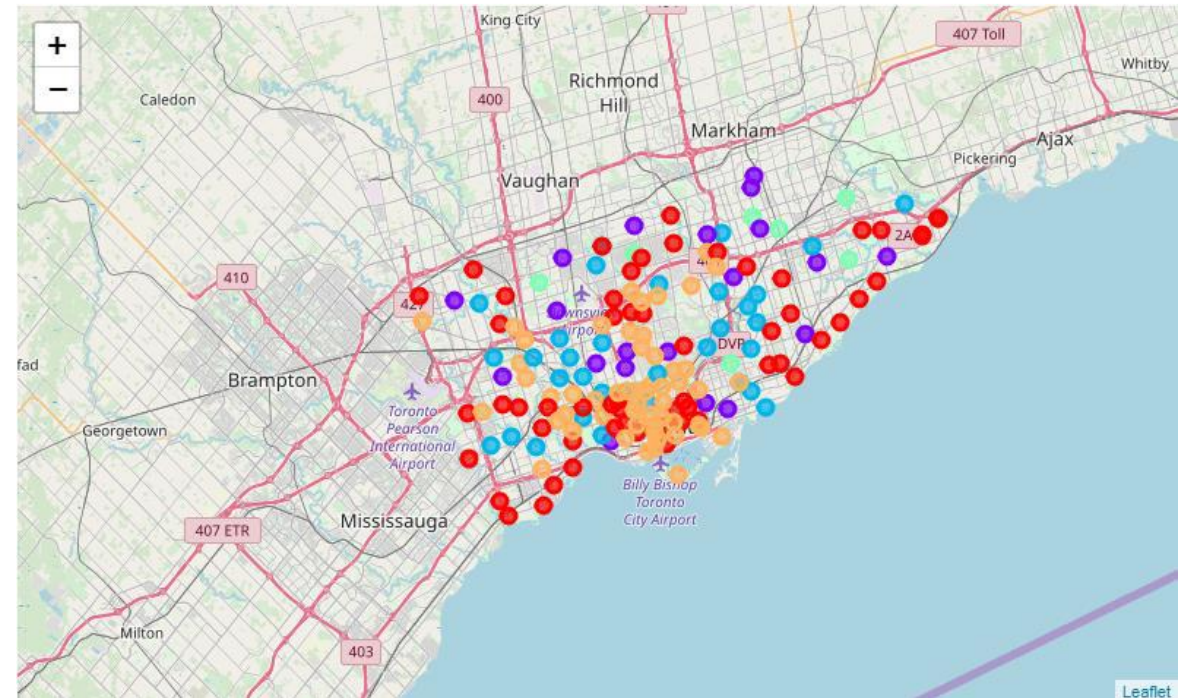Clustering Similar Neighborhoods

- Grouping based on common venues

Neighborhood with high population

- Ranking the neighborhoods based on population count

Topmost common venues

- Ranking the neighborhoods

DISCUSSION

RESULT REVIEW

First cluster looks to offer the best business implementation opportunity. Hubermede neighborhood in the first cluster looks to meet the key business criteria whereby it will offer the highest demand (high population) and lowest competition (lower services supply).

	Neighbourhood	Population	Income	Commuting	2nd Language	2nd Language %	Latitude	Longitude	Population Score	Venue Score	Total Score
60	Humbermede	14778	24,297	11.8	Punjabi (9.7%)	09.7% Punjabi	43.7421	-79.5407	0.611729	0.0	0.305865
61	St. James Town	14666	22,341	27.4	Filipino (8.1%)	08.1% Filipino	43.6709	-79.3733	0.607093	0.0	0.303547
62	Cliffcrest	14531	38,182	12.5	Tamil (1.5%)	01.5% Tamil	43.7249	-79.2263	0.601505	0.0	0.300752
63	Humber Valley Village	14453	80,618	12.0	Ukrainian (3.9%)	03.9% Ukrainian	43.6671	-79.528	0.598276	1.0	0.799138
64	Harbourfront / CityPlace	14368	69,232	16.0	Unspecified Chinese (2.4%)	02.4% Unspecified Chinese	43.6416	-79.3902	0.594758	0.6	0.597379

CONCLUSION

❖ RECCOMENDATION SUMMARY

- Region: 1st cluster

- Neighborhood: Hubermede

Additional Note:

- Updated information is crucial to consider as part of the business implementation plan