

Applied Data Science Capstone Project

Determining the best neighbourhood in Toronto
to open a new coffee shop.

John Koplinae

January 20, 2021



Business Problem

Starting a small business has a host of challenges that accompany it. One such important decision is where to open? and a coffee shop is definitely no exception.

So what neighbourhood in Toronto, Ontario would be good to open a coffee shop?

This presentation looks to help answer that question.

Let's dive into the data!



Data

So what data is needed?

- Neighbourhood information
 - Name
 - Geographical location (Latitude and Longitude)
 - Population
- Coffee Shop information
 - Number of existing coffee shops per neighbourhood

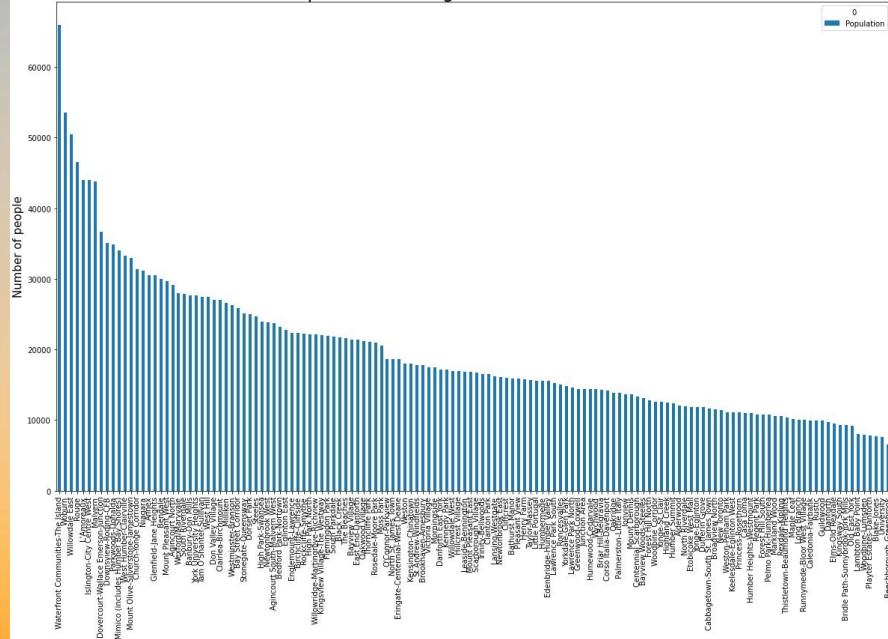
Sources:

- Open Data Portal for the City of Toronto - <https://open.toronto.ca/>
 - Using the data files (.csv) for Neighbourhood Boundaries and Profiles.
 - <https://open.toronto.ca/dataset/neighbourhood-profiles/>
- Foursquare API
 - To acquire the venue data for coffee shops in these neighbourhoods and the number of shops.
- Geocoder package for latitude and longitude coordinates

Toronto neighbourhood data

Cleaned population and location data for Toronto's 140 neighbourhoods.

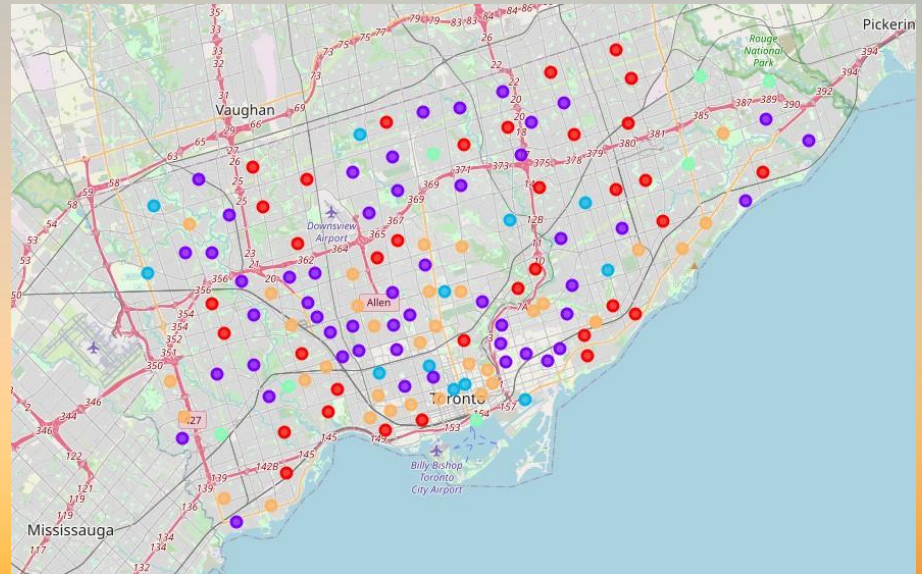
Population of neighbourhoods in Toronto



Normalizing the population and number of coffee shops per neighbourhood.

	Neighbourhood	Coffee Shop	Normalized Population
0	Agincourt North	0.038462	0.441688
1	Agincourt South-Malvern West	0.000000	0.360430
2	Alderwood	0.200000	0.182877
3	Annex	0.119048	0.463126
4	Banbury-Don Mills	0.200000	0.420175

$K = 5$ was used to cluster the neighbourhoods.



Results

- After the neighbourhoods were clustered and plotted, each cluster was reviewed. The chart below are the findings:

	Coffee Shop Frequency	Population
Cluster #1	Low to none	medium
Cluster #2	Low to none	Low
Cluster #3	Medium	Medium
Cluster #4	Low	High
Cluster #5	Medium	Medium - Low

- Cluster #4 has the ideal characteristics for opening a coffee shop. Low number of shops already, means less competition. High population which would lend itself to more customer traffic. As coffee shops are dependent on neighbourhood patronage and foot traffic.

	Neighbourhood	Coffee Shop	Normalized Population	Cluster Labels
58	Islington-City Centre West	0.086957	0.667016	3
66	Lambton Baby Point	0.000000	0.667440	3
73	Malvern	0.000000	0.664421	3
105	Rouge	0.000000	0.705415	3
122	Waterfront Communities-The Island	0.142857	1.000000	3
129	Willowdale East	0.000000	0.765160	3
132	Woburn	0.000000	0.811448	3

Cluster #4



Thank You

For joining and viewing this presentation