

---

### PROBLEM DEFINITION

#### Introduction

Toronto is a financial capital of Canada, its downtown core has large concentration of high-rise office buildings for Canadian and International companies. For several decades Toronto downtown experienced rapid growth in residential condominiums, driven by professionals working in the financial service industry and proximity of their offices.

2020 global pandemic had significant impact on the operations of financial companies. Offices were shut down, and accent was made on the remote work from home. Toronto downtown core became a ghost town, with people looking for more suitable Toronto neighborhoods to relocate.

#### Business Problem

This research will focus on the people currently living in crowded Toronto downtown core, and considering relocating to other Toronto neighborhoods, which will better satisfy their interests and social needs. We will explore Toronto neighborhoods and build a data model, which will help people to find an optimal place to live depending on their specific preferences. Special consideration will be made for the 2020 COVID pandemic, in selecting new neighborhoods.

---

### DATA

#### Data source

This research uses two main data sources. The first one is Toronto neighborhood listing from Wikipedia. This data set is important to establish main framework, define neighborhood locations and names.

The second data set is a listing of Toronto amenities and facilities from Foursquare. This set of data provides important insights into each neighborhood and will allow compare neighborhoods to each other.

#### Data cleaning and wrangling

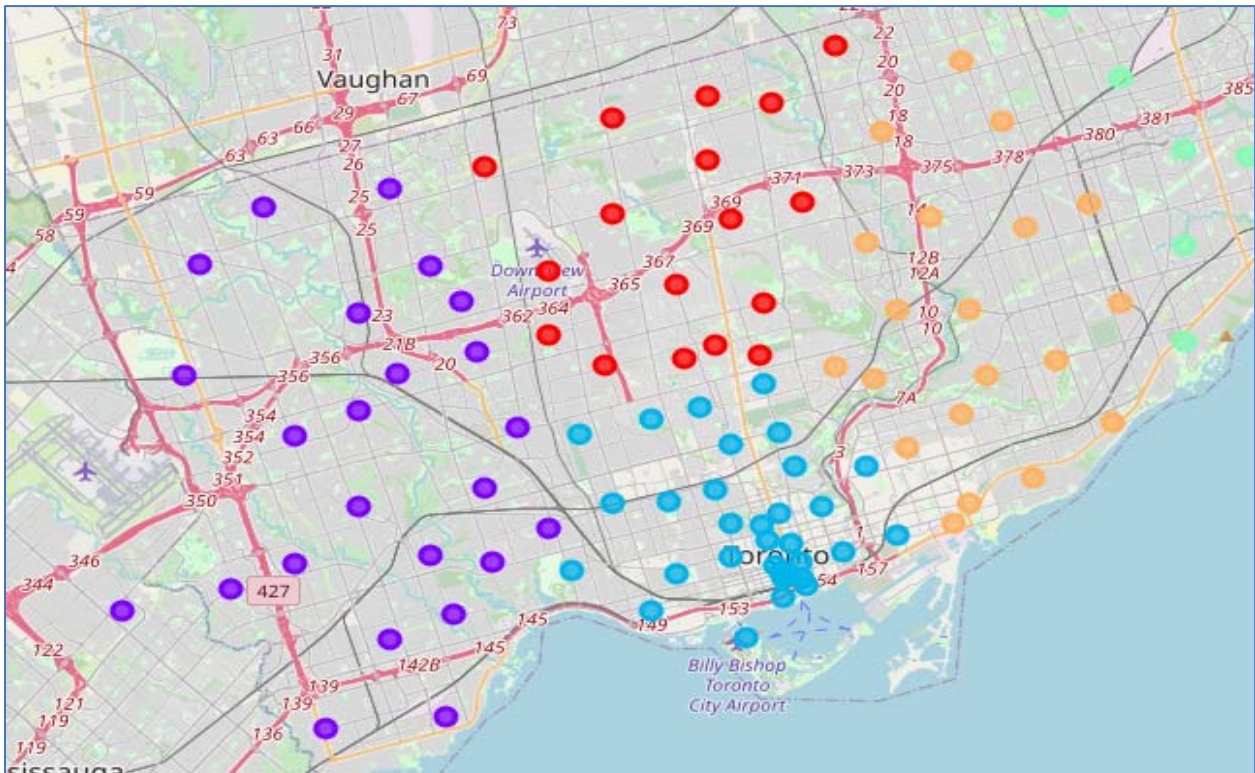
First, we will clean Toronto neighborhoods data set. The set contains postal codes “Not assigned” to specific boroughs, these lines have to be dropped, as they don’t contain any relevant data for our analysis. Next, we check that there is no “not assigned” neighborhoods, duplicate lines or missing values. Once it is done, we add geolocation longitude and latitude to each neighborhood.

Table 1. Example of neighborhood data.

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494

As a result, we have 103 unique neighborhoods and 10 boroughs. In order to make data easier to understand and visualize, we will cluster neighborhoods in several buckets and create a folium map of Toronto neighborhoods.

Table 2. Toronto neighborhood clusters.



The next data set contains Toronto facilities and venues from Foursquare. Considering that it is COVID pandemic, a lot of amenities are on lockdown. For this reason, we will disregard restaurants, and entertainment facilities, which are not essential services and have issues with social distancing rules. We will focus mainly on proximity of grocery chains, outdoor facilities, parks, public transit, and medical centers.

Original venue data from Foursquare was extracted do data frame and mapped based on the geolocation to existing neighborhoods from the previous step. Next, we drop the venues related to restaurants and entertainment facilities, as it is not relevant for our research, and focus on venues with outdoor activities, parks, dog playgrounds, grocery stores and medical offices. As a result, we have a solid data frame for further data analysis and modeling.

Based on the current data, we can check what are the most frequent facilities for each of the Toronto neighborhoods. Further we can rank those venues or focus on the specific ones.

Table 3. Most abundant activity for each neighborhood.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Agincourt	Skating Rink	Train Station	Discount Store	Grocery Store	Golf Course
1	Alderwood, Long Branch	Skating Rink	Train Station	Discount Store	Grocery Store	Golf Course
2	Bathurst Manor, Wilson Heights, Downsview North	Supermarket	Pharmacy	Train Station	Discount Store	Grocery Store
3	Bayview Village	Train Station	Dog Run	Harbor / Marina	Grocery Store	Golf Course
4	Bedford Park, Lawrence Manor East	Pharmacy	Grocery Store	Discount Store	Golf Course	Garden

### Data usage in future the models

Further this data will be used to customize neighborhood search based on the specific customer needs. For example, if the person has a dog, we can find neighborhoods with dog playgrounds, or if the person has special medical needs – neighborhood with a proximity of a medical Centre.