**Objective:**

The objective of the project is to find what type of food/drink business to open in the central part of Toronto .

The business can be a restaurant or a café.

To make that decision it is useful to explore the existing types of business and their concentration I the area of interest. Opening a café or restaurant in a neighborhood where there is already a concentration of similar businesses wouldn’t make sense. So we need to target the neighborhood where there is a need of our food/drinks services.

This could be useful for newcomers to Toronto to have an idea of where to establish their café/restaurant.

**Data:**

To solve the question we will need to :

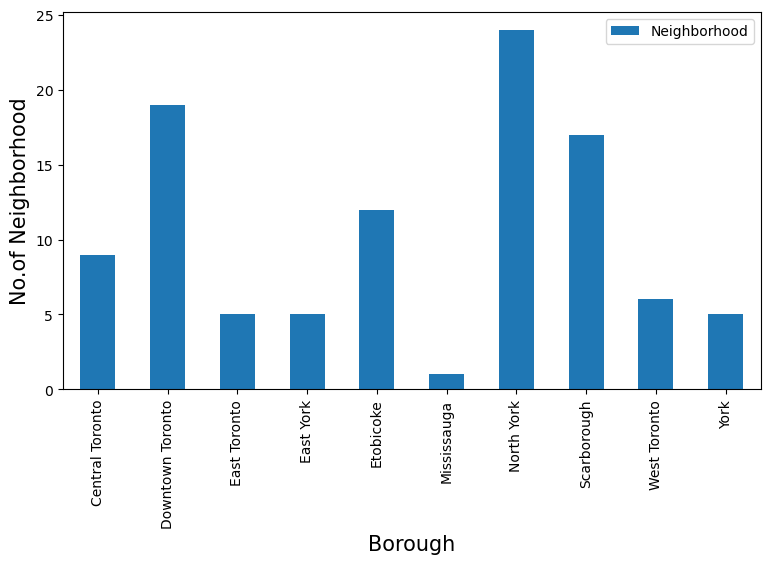
List of the neighborhood in Toronto and be able to place them geographically with coordinate. For that we will be using Wikipedia which has a compilation of neighborhood with coordinates.

We will also need a list of the existing businesses in each neighborhood . For that we will be using Foursquares API

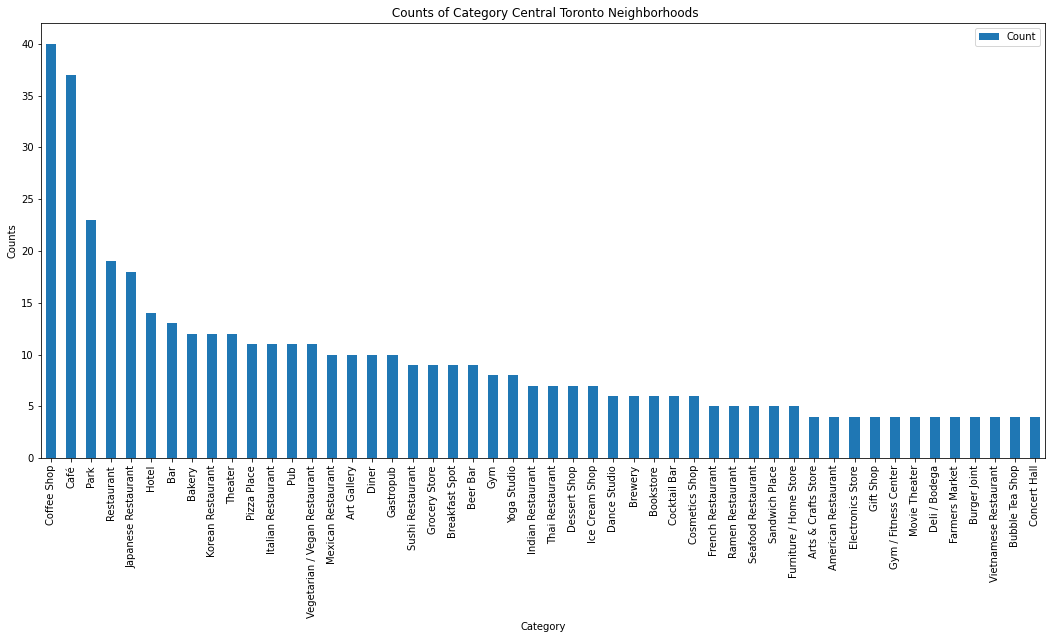
**Methodology**

First the list of neighborhood and borough with coordinates to place them relatively in Toronto is needed.

When comparing the number of neighborhood in the central part of Toronto( Central and Downtown) the downtown part seem to have more neighborhood and therefore would give more potential to places for business. It was decided to concentrate only on the borough Downtown part of Toronto.

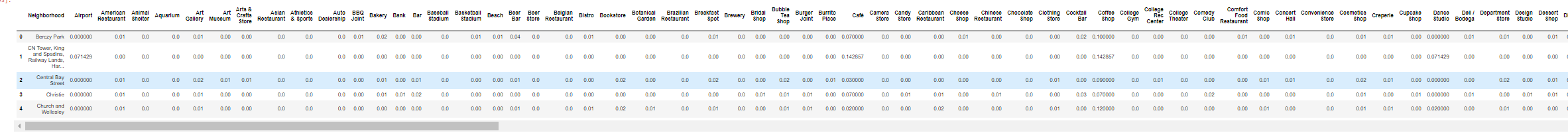


The totality of the existing venues was collected using the FourSquare API services. However this returns a wide range of venues.



The venues data is ranked per neighborhood . We also apply the one hot encoding to list each venue type presence/absence per neighborhood

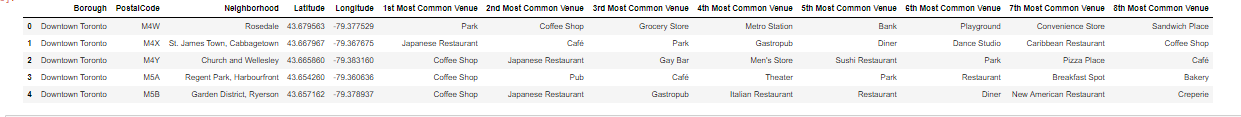
The frequency of each venue per neighborhood was calculated to “normalize” the presence of each type of venue





To be able to relate those venues to the place in Downtown Toronto we also need to merge the venues dataset with the neighborhood coordinates taken from Wikipedia.

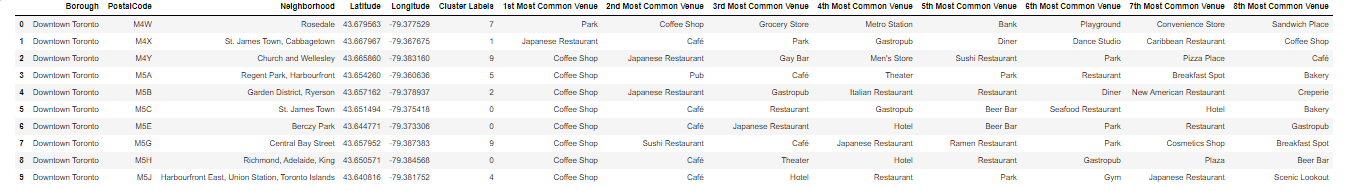
This will give a us a complete dataset with the venues geographically mappable per neighborhood.



We use machine learning techniques to find the most common existing business(venue) in each neighborhood) and also clustering on the basis of the venue occurrences geographic position



Results and Conclusion



Cluster 0 seem to have a lot of drink/food businesses (StJamesTown BercyPAk,Richmond/Adelaide/King) . This would not be a place to open such a business.

Cluster 1 has a restaurant and café as most 2 common venues: not the optimum place.

Cluster 3 and 4 are also café and restaurants rich.

**Cluster 5** has café in 1st and 3rd most present venues but no restaurant in the top 4 most present venues. So this could be a good place to open a restaurant.

The **cluster 7** shows primary venues being not orientated towards eating/drinking businesses so this would be the primary place of choice for opening a café or restaurant.