

# **Capstone Project - The Battle of the Neighborhoods: Toronto vs Manhattan**

Alfredo Aranguren Spirow

April 11, 2021

## **1. Introduction**

### **1.1 Background**

Despite New York and Toronto are not the capital of their respective countries (Washington is the USA capital, and Ottawa that of Canada), these cities are the financial centers of USA and Canada respectively. In this Capstone Project for Coursera, neighborhoods of Toronto and Manhattan will be compared to determine how similar or dissimilar they are.

### **1.2 Problem**

Neighborhoods of Toronto and Manhattan will be compared to determine how similar or dissimilar they are in basis of venues in Foursquare data; and choosing the City of Toronto, will be determined a recommendation to someone to open (or not) a restaurant in this city, in basis of Foursquare venues data, and where would be recommended to open it.

### **1.3 Interest**

New York and Toronto are the financial centers of USA and Canada respectively, and are interesting cities to compare, in this case for a Capstone Project for IBM/Coursera Applied Data Sciences Course. In this Project also will be determined a recommendation to someone to open (or not) a restaurant in the City of Toronto, this could be a real case for open a new restaurant for an independent investor or a restaurant franchise, and where would be recommended to open it.

## **2. Data acquisition and cleaning**

### **2.1 Data sources**

Comparison of Neighborhoods of Toronto and Manhattan is performed in basis of venues in Foursquare data; also for the City of Toronto, the recommendation to someone to open (or not) a restaurant in this city, is carry out in basis of

Foursquare venues data. Data for Toronto neighborhoods was acquired from scrapping through Wikipedia site for Canada Postal Codes ([https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)), and Geospatial Data for Toronto and New York is retrieved from sites on course server, as a .csv file for Toronto (Geospatial\_Coordinates.csv) and a JSON file for New York and Manhattan (newyork\_data.json).

Data for venues in both analysis is requested through Foursquare Places API, with a developer account in this platform for accessing this Foursquare API with credentials associated to account. After data is acquired as JSON files from Foursquare API requests, this data is converted through Pandas libraries to dataframes.

## **2.2 Data cleaning**

Data downloaded or scraped from sources described above were converted into dataframes, and then processed for acquire the necessary information required to perform analysis. Geospatial Data was retrieved for New York City, but in this project was required only data for Manhattan Borough of New York, then a new dataframe with Geospatial Data for Manhattan was acquired from the original dataframe for New York.

## **2.3. Feature selection**

Analysis was carried out for Manhattan from Foursquare Places API, resulting in a Manhattan venues dataframe (manhattan\_venues) with 3235 rows, one for each venue in Manhattan, and 7 columns with Geospatial information (latitude and longitude) for venues and correspondent neighborhoods, including also venue category column. Performed analysis for Toronto from the same Places API resulted in a Toronto venues dataframe (toronto\_venues), with 2111 rows, one for each venue in Toronto, and the same 7 columns with Geospatial information and a category column same as before.

### 3. Exploratory Data Analysis

#### 3.1. Toronto Top Venues

Toronto and Manhattan neighborhoods venues were reviewed and then compared to determine how similar or dissimilar they are analyzing their particular venues frequency on each city. In the case of the City of Toronto Top venues were analyzed and the results are viewed on the next graph:

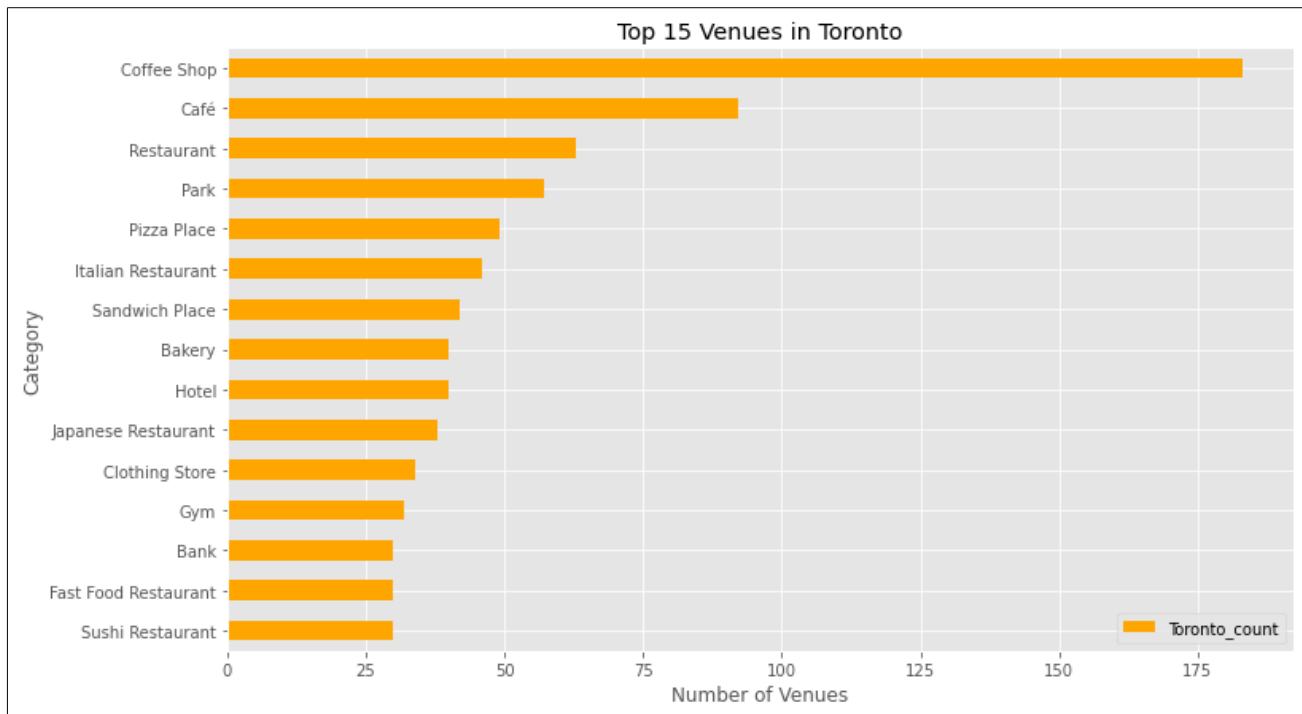


Fig.1. Top 15 Venues in Toronto, Venues data from Foursquare Places API.

It could be observed for Figure 1, than 8 of 10 top venues for Toronto are food places (Coffee Shop, Café, Pizza Place, Restaurants and a Bakery), only a Park is on the first places (4<sup>th</sup> place) and Hotel is the 10<sup>th</sup> place on top venues for Toronto. Next top venues are Japanese Restaurant, Clothing Store, Gym, Bank, Fast food Restaurant, and a Sushi Restaurant. In basis of this top venues for the City of Toronto, it could be concluded for food tastes than Torontians likes the most Coffee Shops and Cafés, followed by Italian food (Pizza Place and Italian Restaurant), then by Sandwich Place and Bakery, and last in this top venues for food in this list, are the Asian restaurants (Japanese and Sushi Restaurant).

### 3.2. Manhattan Top Venues

Top venues were analyzed for Manhattan and the results are displayed on the next graph:

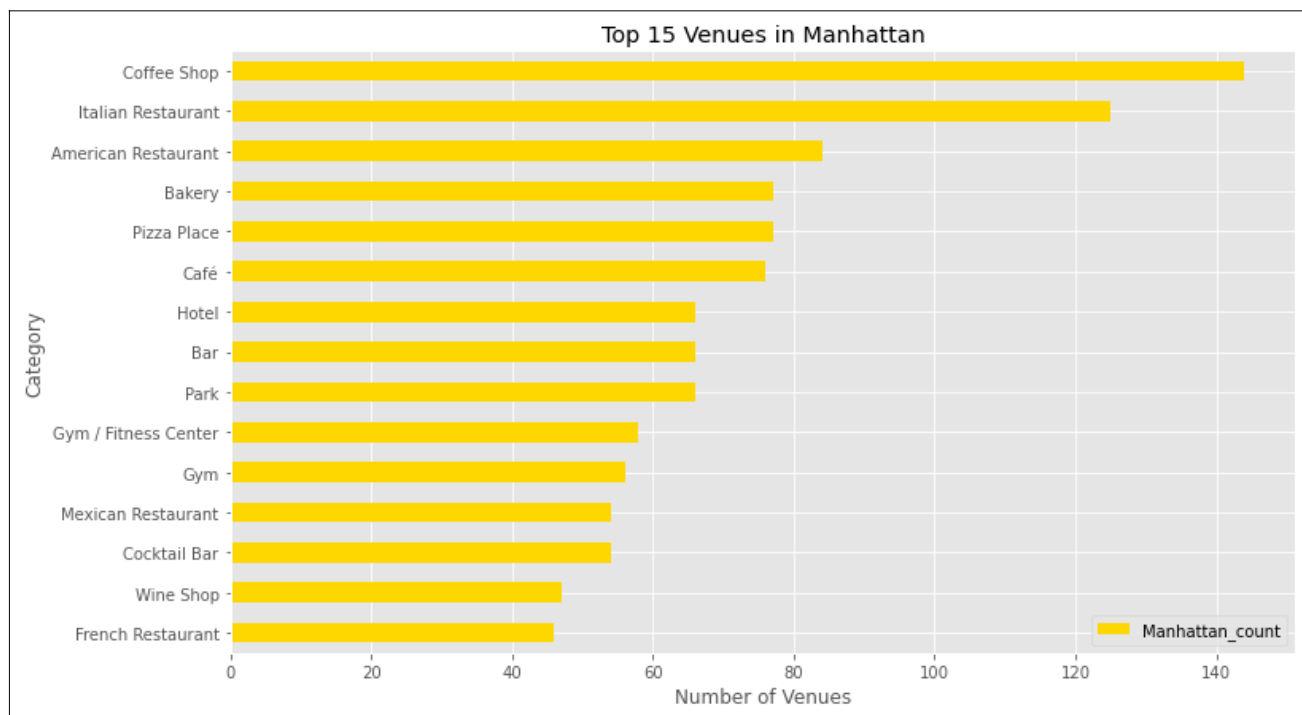


Fig.2. Top 15 Venues in Manhattan, Venues data from Foursquare Places API.

It could be viewed for Figure 2, than 6 of 10 top venues for Toronto are food places (Coffee Shop, Pizza Place, Restaurants, Bakery and Café). Next top venues for Manhattan are Hotel, Bar, Park, Gym, followed next by Mexican restaurants, Cocktail bar, Wine Shop and French Restaurant. In basis of this top venues for Manhattan, it could be concluded for food tastes than New Yorkers likes the most Coffee Shops, followed by Italian food (Pizza Place and Italian Restaurant), then by American Restaurant and Bakery, and last in this top venues for food in this list, are the Mexican and French restaurants.

### 3.3. Manhattan and Toronto Top Venues comparison

Top venues were analyzed for both Manhattan and Toronto, and the results are viewed on the next graph:

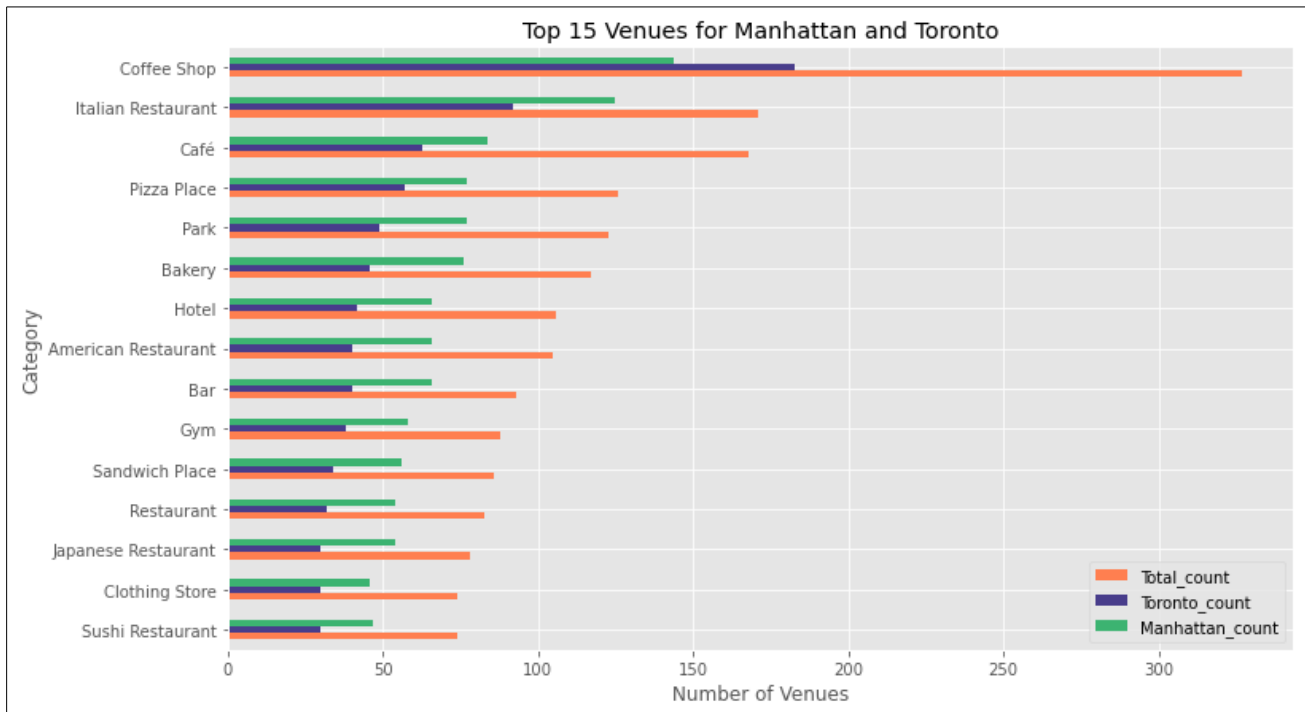


Fig.3. Top 15 Venues for Manhattan and Toronto, Venues data from Foursquare Places API.

It could be observed for Figure 3, that 4 of the first top venues for Toronto and Manhattan are food places (Coffee Shop, Pizza Place, Italian Restaurant, and Café). In this comparison graph, both New Yorkers and Torontonians like the most Coffee Shops, and New Yorkers like Italian restaurants more than Torontonians. In number of venues could be concluded in general that comparing Manhattan (and New York) and Toronto, these are very similar in tastes of their people, New Yorkers and Torontonians, but have some differences as could be noticed better on Figures 1 and 2, some of them is related to their habits, from analysis of venues on these figures, New Yorkers seem to like gym and park training more than Torontonians.

### 3.4. Analysis for Restaurant Opening in Toronto

Analysis of venues on figures 2 and 3, allows to conclude that could be recommended opening of a Restaurant in Toronto, despite current situation of covid-19; off course, most of these restaurants and coffee shops on top of venues does food delivery, and it is recommended also (IMHO) opening of restaurant with reduced dinning room and oriented to prepare most of meals for delivery on the currents circumstances. Anyway, an analysis of venues and restaurants distribution on maps, also allows to recommend some places for opening a restaurant in Toronto, as is shown on next figure:

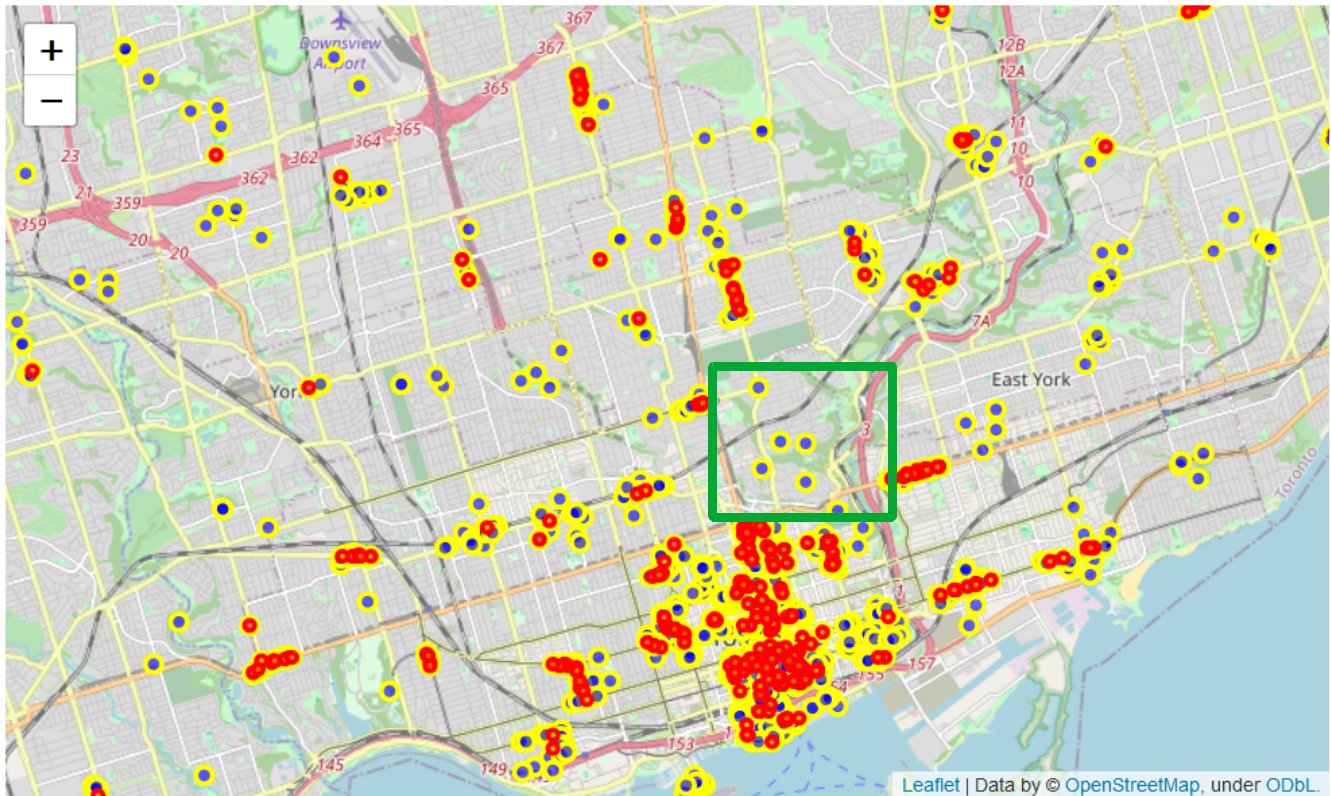


Fig.4. Map of Toronto with Venues (circles in blue), and Restaurant (in red). In green frame area with recommended places to opening a restaurant. Data sources: Foursquare venue data and Leaflet base map data.

Area in green frame on figure 4, comprises at least three neighborhoods in which are several venues different to restaurant (circles in blue), and without restaurants (circles in red). The geospatial analysis of this venue data from Foursquare on map, allows to recommend opening of a restaurant on the neighborhoods of Sherbourne (in first place), Rosedale and Summerhill (in third place, is the further of Toronto Center).

## 4. Conclusions

Foursquare is a very useful data source for data sciences, lets analysis of venues with geospatial data in almost all the world and updated frequently. Comparing Manhattan to Toronto, could be concluded than both New Yorkers and Torontians likes the most Coffee Shops, and New Yorkers likes Italian restaurant more than Torontians. In number of venues could be concluded in general than comparing Manhattan (and New York) and Toronto, these are very similar on tastes of their people, New Yorkers and Torontians, but have some differences, some of them is related to their habits, from analysis of venues, New Yorkers seem to like gym and park training more than Torontians. Analysis of venues allows to conclude that could be recommended opening of a Restaurant in Toronto. The geospatial analysis of the venue data from Foursquare on map, allows to recommend opening of a restaurant on the neighborhoods of Sherbourne (in first place), Rosedale and Summerhill (in third place, is the further of Toronto Center).