Comparing Neighborhoods in Downtown Toronto and Manhattan: Are the Top trending venues similar?

# Introduction

The purpose of the analysis is to compare the Top Venues found in the Boroughs of Manhattan (NYC) and DownTown Toronto (TO), using the foursquare API and the data provided during the course.

Even though the two Boroughs differs in terms of population density and geographical areas, they are the most central Boroughs of the two cities with possible similarities in terms of Architectural structure and consumer preferences that could be identified using the top trending venues in the two Boroughs. Are Top Venues in the two cities showing similar patterns?

Here below the steps of the analysis:

**1.** Import and preprocess the data from the web in order to identify the location and the Neighborhoods belonging to the two Boroughs.

**2.** Identify top trending venues in each Neighborhood on the two Boroughs using Foursquare API and then run a cluster analysis to group similar Neighborhoods.

**3.** Analyse similarities and dissimilarities on the two Boroughs emerging from Cluster Analysis.

**4.**Conclusion and final thoughts.

**Target Audience**: Everyone who is interested in exploring the two Boroughs and would like to make a choice about which neighborhood to stay, based on individual preferences and purpose of the visit.

# Data

Data sources have been provided during the course:

**Canada Postal Codes**: <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>

**Geospatial Coordinates:** <http://cocl.us/Geospatial_data>

**New York City Data:** <https://cocl.us/new_york_dataset>

# Methodology

1. **Importing and Preprocessing**: The analysis has started by importing the necessary data from Web sources thanks to different python packages, such as Beautiful Soup and Pandas. Data have been preprocessed and visualized, in a way to make it easy to run the Cluster analysis and compare the results.
2. **Top Venues:** The analysis then moves on to find the current 10 Top Venues in each Neighborhood in the two Boroughs, using FourSquare API. Data have been organized by category (e.g. Restaurant, Coffee Shop, Theatre) and in tabular form in order to effectively run the cluster analysis. Code for the task has been provided in the Lab.
3. **Cluster Analysis:** Last Task in methodology section has been to run a Cluster Analysis on the Neighborhoods belonging to the two neighborhoods. Method used has been k-mean clustering and the maximum number of cluster has been set to 5. Data has been then visualized in order to get a sense on how the algorithm has grouped the different Neighborhoods.

A picture containing text, map

Description automatically generated

Clusters Downtown Toronto

A picture containing text, map

Description automatically generated

Clusters Manhattan

# Results

Result section will start with a brief description of the main characteristics of the clusters found in the Two neighborhoods:

**Downtown Toronto**

**Central Area 1 – Orange Cluster**: This Cluster includes neighborhoods that mostly belongs to the central area of the Borough. As shown on the map, this Clusters groups the majority of neighborhoods belonging to Downtown Toronto. Top venues in this Cluster are mostly Coffee Shops, Restaurants and Hotels. Neighborhoods belonging to this cluster are not only the most central in the city, but also the most active in terms of Business Activities. These areas are probably the most crowded in the city during the day and Night as well.

A screenshot of a cell phone

Description automatically generated

**Central Area 2 – Red Cluster**: Neighborhoods in this cluster are still in the central area of downtown Toronto. Coffee Shops still dominates the top Venues but seems that preferences in these neighborhoods goes more on Breakfast and sandwich places, compared to Central Area 1.

**Airport – Green Cluster**: Billy Bishop airport is the Downtown Toronto airport. The cluster is unique in terms of Top venues and is not a surprise that this includes mostly Airport related activities, such as Car Rentals, Airport Lounge and various bars and food places located in the terminal.

**Rosendale – Purple Cluster**: According to top venues and geographical location, Rosendale Neighborhood seems a quiet area surrounded by parks and Nature. Top Venues in this area include a Park, Playground, Dog Run, mostly related to open areas activities. Located away from the crowded city center, it is considered a residential area where the wealthiest individuals live.

**Christie – Blue Cluster**: Christie Neighborhood is characterized by a wider range of Top Trending venues such as Groceries, Christie Pits Park, Restaurants and Coffee Shops. It is considered an active multicultural neighborhood and much less residential compared to Rosendale.

**Manhattan**

**South – Red Cluster:** The two neighborhoods in this Cluster (Battery Park City and Civic Center) are characterized by two hotels in the first places and high quality restaurants and shops. There is also preference for Fitness activities with top venues identified in a Park, Yoga studio, Spa and Gyms.

**Central – Purple Cluster:** This cluster represent the most central neighborhoods in Manhattan. This is where most of the Touristic and business activities take place. We can see that top Venues are mostly represented by Coffee Shops and Restaurants (strong preference for Pizza Places and Italian Cousine). There is also a strong preference for Gyms, probably related to the great number of offices in the Area.

**A screenshot of a cell phone

Description automatically generated**

**North and Chinatown – Orange Cluster:** The orange cluster in Manhattan groups Chinatown, Harlem and, Inwood and Hamilton Heights. Preferences in this neighborhoods are strongly inclined toward ethnic restaurants, especially Mexican, African and Asian ones.

# Discussion

After the comments on the analysis of the clusters obtained, is it possible to return to the original question: Can Downtown Toronto and Manhattan be considered similar? Do they have a similar structure? Do the Top Trending venues show similar cultural and consumption patterns?

**Similarities**

Central locations of the two Boroughs shows similar patterns in terms of types of Activites. These are the most central areas in the two boroughs and the most active in terms of business, nightlife and Touristic Activities.

Top trending venues in **Central Area 1** in Downtown Toronto and **Central** in Manhattan are mostly related to Horeca sector with Coffee shops as the most top rated activities in the two Clusters.

**Dissimilarities**

1. Although Central areas are showing similar patterns, if we look closely at the two Clusters we can see that preferences **Central Area 1** are for a wider range of restaurants, while in Manhattan preferences are for Pizza Places and Italian Restaurants. There are also more Top rated Hotels compared to Central in Manhattan.

2. Downtown Toronto seems a more heterogeneous Borough. Clusters are more distinguished compared to Manhattan. Neighborhoods away from the center in Downtown Toronto are really different in terms of top venues and seem more residential. There is also an Airport area that is missing in Manhattan.

3. Although Clusters in Manhattan are dissimilar in terms of Top venues, we can see similar patterns between them, mostly dominated by Horeca sector, with the exception of Gyms that are top trending in Manhattan and not in downtown Toronto. Parks and outdoor activities are more Top rated in Downtown Toronto due to different geographical structure of the two Boroughs.

# Conclusion

This conclude the analysis of the two boroughs. Although the two are the most Central Boroughs in the cities of Toronto and New York, a closer analysis of top trending venues has discovered dissimilarities in terms of preferences mostly related to cultural and geographical reasons.