

|  |
| --- |
| The Battle of Neighborhoods |
|  |
| May 15, 2021  *IBM Data Science Professional Certificate – Capstone Project*  Authored by: Omesha Virajini |

# Introduction

|  |
| --- |
| Business problem  Every city is unique in their own way and give something new. The city of New York is the most populous city in USA. It is diverse and is the financial capital of USA. It is multicultural and rich in heritage and developed enough from a foreign perspective.  Toronto, the capital of Ontario is a major Canadian city. It is the most populous city in Canada and fourth most populous city in North America. Both cities are very diverse and are the financial capitals of their respective countries.  Today Tourism is one of the most essential components of the economy. Most often people would like to visit countries which are having friendly environment. Tourists always eager to travel to different places on the basis of available information. The comparison between two cities always support to choose the specific places according to their choice.  Who would be interested?  Jane is a tourist, and she needs to explore how much these two cities, New York and Toronto are similar or dissimilar in aspects from a tourist point of view regarding food, accommodation, beautiful places etc. and to get more helpful information when making decisions about where to travel. Data Description of the data  For this project, Foursquare API will be used to gain the data of two cities, in terms of their neighborhoods. The information we want to focus on are restaurants, hotels, coffee shops, parks, theaters, art galleries, museums around each neighborhood. One Borough from each city will be chosen to analyze their neighborhoods. Therefore, Manhattan from New York and Downtown Toronto from Toronto will be considered.  How data will be used to solve the problem  According to the objectives of this project, Machine Learning technique called “Clustering” will be used to segment the neighborhoods with similar objects on the basis of each neighborhood data. Based on foot traffic in respective neighborhoods tourist’s areas and hubs will be located and that information will be used to decide the similarities or dissimilarities between two cities. MethodologyInitially we have selected Downtown Toronto and Manhattan to explore their neighborhoods. The data exploration, analysis and visualization for both boroughs are done in the same way but separately.Downtown TorontoFirst Downtown Toronto was considered.For Downtown Toronto case, we have extracted table of Toronto’s Borough from Wikipedia: http://zims-en.kiwix.campusafrica.gos.orange.com/wikipedia\_en\_all\_nopic/A/List\_of\_postal\_codes\_of\_Canada:\_MFirst the data was arranged according to our requirements.In the arrangement phase, which applied multiple steps including but not limited to, eliminating “Not assigned” values,Since I was not able to get the geographical coordinates of the neighborhoods using the Geocoder package, I have used the csv file that has the geographical coordinates of each postal code.Combine neighborhoods which have same geographical coordinates at each borough and sorted against the concerned borough.Set index for only Downtown Toronto and eliminate 'Postcode' column.For data verification and further exploration, Foursquare API was used to get the coordinates of Downtown Toronto and explore its neighborhoods.The neighborhoods are further characterized as venues and venue categories. Put blue dots on Bucharest map to see centers of neighborhoods.Toronto venuesThe top 10 venues for each neighborhood were obtained.**Clustering Neighborhoods - Downtown, Toronto**A new data frame was created including the cluster labels as well as the top 10 venues for each neighborhood.Manhattan, New YorkFor Manhattan, we used a saved data file which is already explored through foursquare API in which we have extracted all the boroughs of New York and then sorted against the concerned borough.Then we explored the Manhattan neighborhoods as venues and venue categories          The neighborhoods are further characterized as venues and venue categories. Put blue dots on Bucharest map to see centers of neighborhoods.    Manhattan venues      The top 10 venues for each neighborhood were obtained.  **Clustering Neighborhoods – Manhattan, New York**A new data frame was created including the cluster labels as well as the top 10 venues for each neighborhood.  ResultsDowntown TorontoCluster MapNow, we can examine each cluster and determine the discriminating venue categories that distinguish each cluster. Based on the defining categories, we can then assign a name to each cluster.**Cluster 1 (Restaurant, Café, Bar, Gym)****Cluster 2 (Park, Shops, Playground, Gym)****Cluster 3 (Airport, Harbor)****Cluster 4 (Restaurant, Spa, Pub, Museum, Night Club, Park)****Cluster 5 (Basketball Stadium, Market, Bar, Seafood)** Manhattan New York  Cluster map   **Cluster 1 (Restaurant, Café, Theater, Gym, Bar, Coffee Shop)**  **Cluster 2 (Restaurant, Seafood, Bar, Gym, Spa, Park, Museum)**    **Cluster 3 (Restaurant, Night Club)**  **Cluster 4 (Restaurant, Café, Bar, Gym, Spa)**    **Cluster 5 (Market, Park, Gym, Coffee)**    DiscussionAfter clustering the data of the respective neighborhoods, in Downtown and Manhattan, It can be seen that there are venues which can attract the Tourists.The neighborhoods are much similar in features like Theaters, Gyms, Parks, food places, Clubs, museums, Spas etc.As far as concern to dissimilarity, it differs in terms of some unique places like historical places, Airports and Harbors.When comparing the tourist places, it can be observed that the historical place is only situated in Downtown Toronto and the Monument or landmark venue is in Manhattan neighborhoods.Similarly, Airport facility, Harbor, Sculpture garden and Boat or ferry services are also available in Downtown Toronto while venues like Nightlife, climbing gym and Museums are present in Manhattan.As far as concern to recommendations, we recommend Downtown Toronto Neighborhoods will be considered first to visit. The tourists have an easily travelling access due to Airport facility, which not only saves time but also helps to save money. This saved money can be utilized to explore more, the attracting venues.ConclusionThe downtown Toronto and Manhattan neighborhoods have more like similar venues.As we know that every place is unique in its own way, so that’s argument is present in both neighborhoods.The dissimilarity exists in terms of some different venues and facilities but not on a larger extent. |
|  |
|  |
|  |