**Capstone Project**

**The Battle of Neighborhoods**

Version 1.0

By,

Sheikh Jawwad Hakeem

06 July 2019

Contents

[Introduction 3](#_Toc13270331)

[Data 3](#_Toc13270332)

[Methodology 3](#_Toc13270333)

[Brooklyn, New York 5](#_Toc13270334)

[North York, Toronto 5](#_Toc13270335)

[Machine Learning Technique 5](#_Toc13270336)

[Results 7](#_Toc13270337)

[Discussion 7](#_Toc13270338)

[Conclusion 7](#_Toc13270339)

# Introduction

Individuals and families migrate to different cities across the globe for multifarious reasons. Some migrate to join families, and some migrate to find better livelihood. Not much research goes into picking a place other than what is evident / on-the-table. Not to mention, family-members' and friends' experiences are relied upon the most.

With factual research in place, one can take decisions on quantifiable basis. I’d like to use quantifiable data to show how one can pick a place like their current habitat. This will be useful for all who are looking to change their existing address with a new one, and maintain a similar, if not the same, lifestyle.

I will be using the foursquare API to compare different localities – their neighborhoods and available amenities – to establish the best option for migration.

# Data

I will be comparing Brooklyn, New York with North York, Toronto. This comparison will help individuals and families evaluating to permanently or temporarily move between the two areas. The comparison will be based upon the amenities available in the neighborhoods of the respective cities. Using this comparison one can evaluate how similar or dissimilar their choice of destination will be, to their existing one.

The data would be based upon different cuisines, coffee shops, health facilities, parks, etc. Information based on these dimensions plays a vital role in deciding on a place to live.

# Methodology

Following are maps for Brooklyn and North York, with their respective neighborhoods highlighted on them. It is evident, from the illustration, that North York is not as populated as Brooklyn.

|  |  |
| --- | --- |
| **Brooklyn, New York** | **North York, Toronto** |
|  |  |

The same is reflected from the numbers below:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Neighborhood** | **Venue** | **Venue Categories** |
| **Brooklyn** | 70 | 2831 | 287 |
| **North York** | 24 | 251 | 111 |

This would be an attractive point for someone looking for a quieter neighborhood. Extracting the common areas between the two cities we come up with the following venue categories:

American RestaurantArts & Crafts StoreAsian RestaurantAthletics & SportsBakeryBankBarBaseball FieldBasketball CourtBeer StoreBike ShopBoutiqueBridal ShopBubble Tea ShopBurger JointBurrito PlaceBus LineBus StationButcherCaféCandy StoreCaribbean RestaurantChinese RestaurantClothing StoreCoffee ShopConvenience StoreCosmetics ShopDeli / BodegaDepartment StoreDessert ShopDim Sum RestaurantDinerDiscount StoreDog RunElectronics StoreEvent SpaceFalafel RestaurantFast Food RestaurantFood & Drink ShopFood CourtFood TruckFried Chicken JointFrozen Yogurt ShopFurniture / Home StoreGeneral EntertainmentGolf CourseGreek RestaurantGrocery StoreGymGym / Fitness CenterHome ServiceHotelIce Cream ShopIndian RestaurantIntersectionItalian RestaurantJapanese RestaurantJuice BarKids StoreKorean RestaurantLiquor StoreLoungeMassage StudioMediterranean RestaurantMetro StationMiddle Eastern RestaurantMiscellaneous ShopMovie TheaterOther Repair ShopParkPet StorePharmacyPizza PlacePlazaPoolPubRamen RestaurantRestaurantSalon / BarbershopSandwich PlaceShoe StoreShopping MallSmoke ShopSpaSporting Goods ShopSteakhouseSupermarketSupplement ShopSushi RestaurantTea RoomThai RestaurantTheaterToy / Game StoreVideo Game StoreVideo StoreVietnamese RestaurantWings JointWomen's Store

The following tables give a snapshot of the information that was extracted from various sources for the two cities:

## Brooklyn, New York

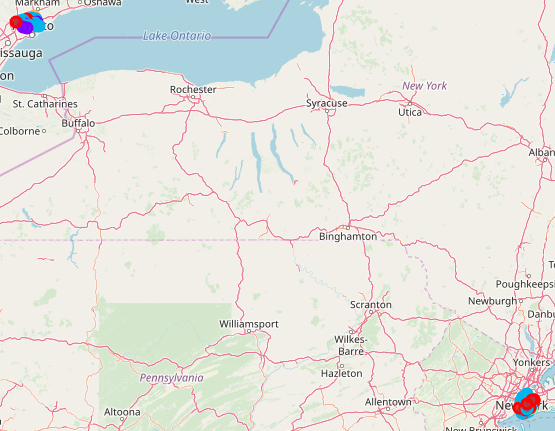


## North York, Toronto



## Machine Learning Technique

I’ve used the above information to extract similarities between the two cities through the means of clustering. K-means algorithm has been used to classify the neighborhoods of Brooklyn and North York. Following is the result of the clustering analysis:



The same when zoomed results in the following:

|  |  |
| --- | --- |
|  |  |

The color scheme shows the neighborhoods that are similar in characteristics.

The same can be conducted for a reduced number of clustering parameters, depending on the likes and dislikes of the migrant.

# Results

Based on the multi-dimensional analysis of the data we were able to help the user establish similar migration destination(s). The data is rich enough for the migrate to be able to pick and choose the basis for analysis as well.

# Discussion

As an option, the same strategy can be extended to determine movement within the city of New York as well. Also, based on the migrant’s interest, the selection of venue categories can be changed to suite any specific needs.

# Conclusion

The study provides data on the top venue categories for each city / neighborhood. The user/migrant may use this data to make an informed decision with respect to choosing a target city.