|  |
| --- |
|  |
|  |
| **Compare the neighborhoods of two cities of New York and Toronto.**  **August 10th 2019** |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

**Introduction**

**Background**

The City of New York, usually called either New York City (NYC) or simply New York (NY), is the [most populous city](https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population) in the [United States](https://en.wikipedia.org/wiki/United_States). With an estimated 2018 population of 8,398,748 distributed over a land area of about 302.6 square miles (784 km2), New York is also the [most densely populated](https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population_density) major city in the United States. A [global power city](https://en.wikipedia.org/wiki/Global_city),  New York City has been described as the [cultural](https://en.wikipedia.org/wiki/Culture_of_New_York_City), [financial](https://en.wikipedia.org/wiki/Wall_Street), and [media](https://en.wikipedia.org/wiki/Media_in_New_York_City) capital of the world, and exerts a significant impact upon commerce, entertainment, research, technology, education, politics, tourism, art, fashion, and sports.

Toronto is the provincial capital of [Ontario](https://en.wikipedia.org/wiki/Ontario) and the [most populous city in Canada](https://en.wikipedia.org/wiki/List_of_the_100_largest_municipalities_in_Canada_by_population), with a population of 2,731,571 in 2016. Current to 2016, the Toronto [census metropolitan area](https://en.wikipedia.org/wiki/Census_metropolitan_area) (CMA), of which the majority is within the [Greater Toronto Area](https://en.wikipedia.org/wiki/Greater_Toronto_Area) (GTA), held a population of 5,928,040, making it Canada's [most populous CMA](https://en.wikipedia.org/wiki/List_of_census_metropolitan_areas_and_agglomerations_in_Canada). Its [economy](https://en.wikipedia.org/wiki/Economy_of_Toronto) is highly diversified with strengths in technology, design, financial services, life sciences, education, arts, fashion, business services, environmental innovation, food services, and tourism.

**Problem**

New York City and the city of Toronto are financial capitals of their respective countries of USA and Canada. Both of the city host multi-cultural environments. New York and Toronto are huge cities. We are going to use data to find out how similar or dissimilar they are to each other. This information is going to be useful for anyone who plans on visiting in these cities.

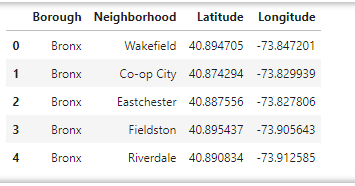
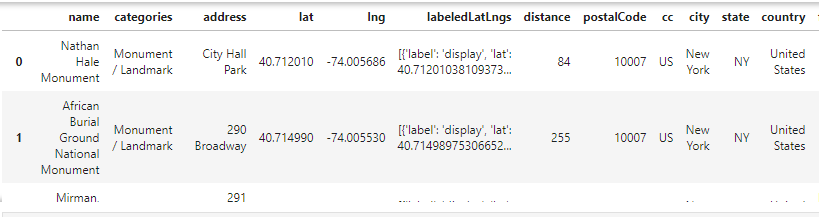
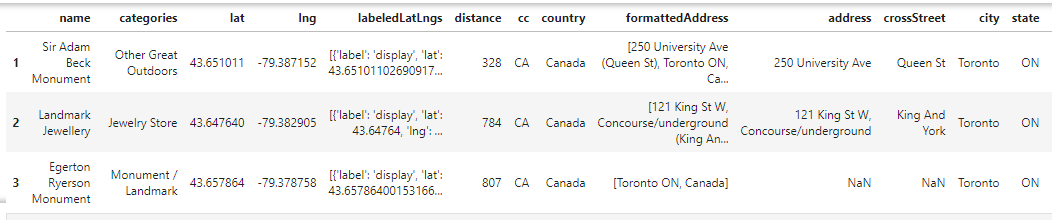
**Data acquisition and cleaning**

**Data Sources**

These two cities are huge and the number of factors/features that can be compared between them are too many. We are going to limit the number of data point that we are going to use. We will use sources like rent jungle to scrap of average rental information in these two cities.   
But we will be primarily using Foursquare for getting the location information of venues (features).  
We will use some or all of these features. The features we are shortlisting are

* Monument / Landmark
* Theaters / art galleries
* Coffee
* Food
* Parks / Trails
* Nightlife
* self defense
* Robbery/Assault

Here are the data and its sources that would be used in this solution.

1. **Neighborhood information:** For base map and location.  
   1. **Toronto**  
      **Source**: For the Toronto neighborhood data, a Wikipedia page exists that has all the information we need to explore and cluster the neighborhoods in Toronto. We will scrape the Wikipedia page and wrangle the data, clean it, and then read it into a pandas dataframe.  
      The below diagram shows Neighborhood data and coordinates information merged together.  
         
      (sample data)
   2. **New York city  
      Source:** The 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood.  
      This dataset exists for free on the web. here is the link to the dataset: <https://geo.nyu.edu/catalog/nyu_2451_34572>  
      (sample data)
2. **Foursquare Data:**We make calls to the Foursquare API for this purpose of getting feature information. We will URL to send a request to the API to search for a specific type of venues like, to explore each particular venue (Monument / Landmark, Theaters / art galleries, Coffee, Food, Parks / Trails and Nightlife. Also, we will use the visualization library, Folium, to visualize the results.  
   For example:  
    Shown below is the sample dataset for venue of Monument / Landmark for both New York and Toronto respectively.
   1. **New York**  
        
      (sample data)
   2. **Toronto city**  
        
      (sample data)
3. **Average rental information:   
   Source:** The data for rents of 1bedroom, 2 bedroom and 3-bedroom averages are available on rentjungle.com. The data for Toronto is limited to few years when compared with New York city. So we will only use data for year that are available for both cities.<https://www.rentjungle.com/average-rent-in-toronto-rent-trends/>  
   <https://www.rentjungle.com/average-rent-in-new-york-rent-trends/>

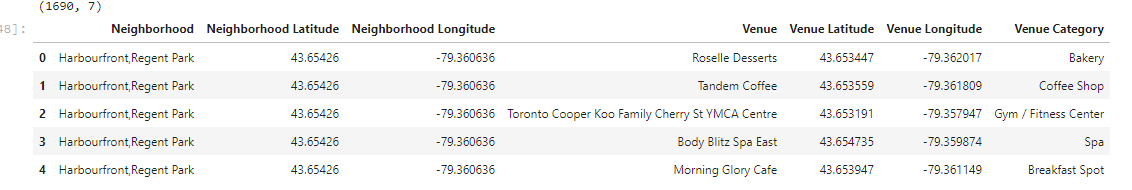
|  |  |
| --- | --- |
| **Toronto** | **New York City** |
|  |  |

(sample data)

**Data Analysis**

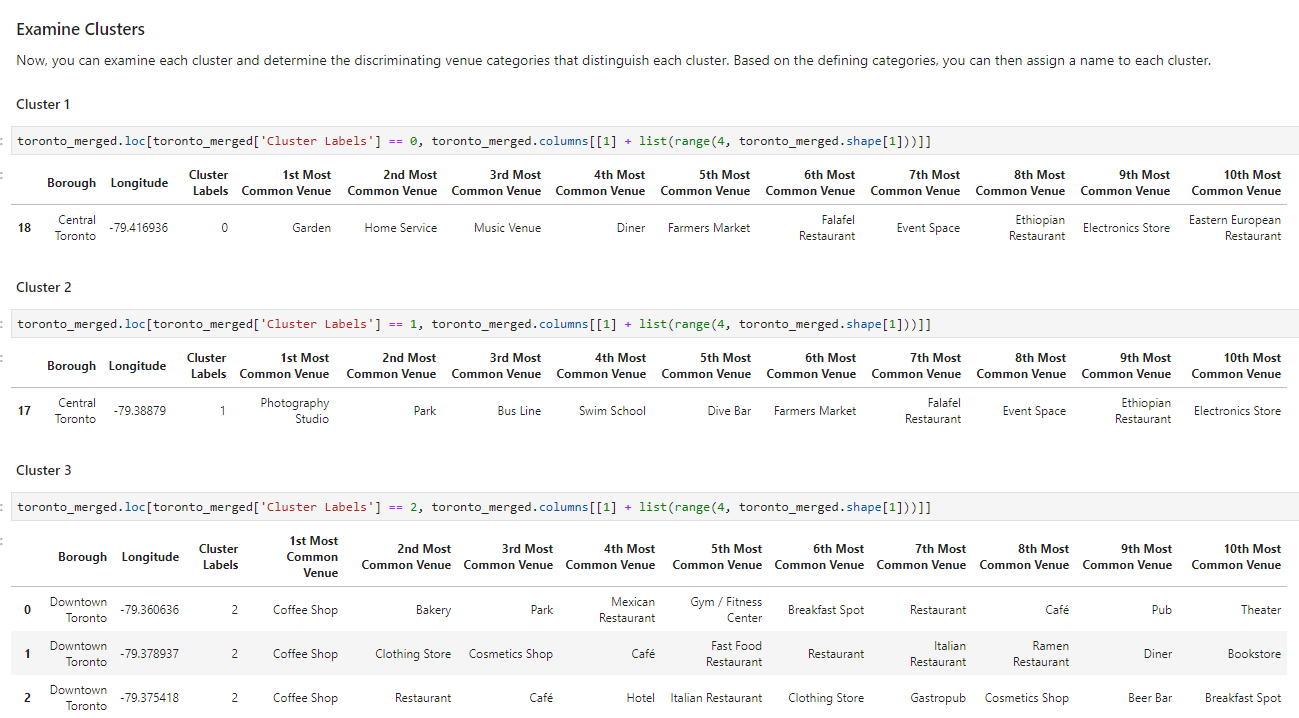
Due to limitation of how many calls Four square allows, the problem was converted to selecting top 10 venues in each of the cities.

These are not the same as we had set out in the beginning.



#### From the Foursquare lab in the previous module, we know that all the information is in the items key. Before we proceed, let's borrow the get\_category\_type function from the Foursquare lab.

The clusters reveal a very diverse set of venues that do not match the ones we had set out in the beginning.



**Solution to the problem**

These two cities are similar in weather and the kind of venues that are available to both the cities. New york being the larger of the two there are more options available. But having said that Toronto is not far lagging being and has higher percentage of tourism revenue compare the city that is half the size of New York city