The Battle of Neighborhoods

Toronto vs New York

(1) Introduction

Tourism is the centre of the economy. When people plan their trip, they will search many information and decide where to visit. Thus there are always some comparison in the mind of every tourist. Here we hope to use a data-mining approach to compare two city, and see which of them can face more customers.

Here we will compare the tourist spots between Toronto and New York. We aim to compare them because they are both multicultural but diverse in many ways. We will focus on extracting the location data about restaurant, accommodation and others.

(2) Data Description

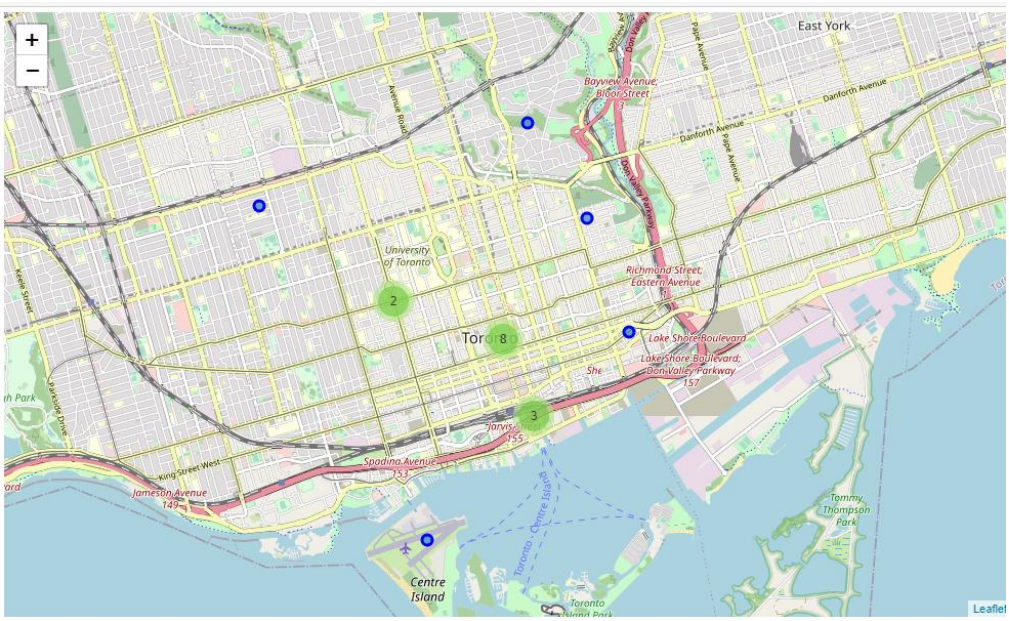
To solve this problem, we will access Foursquare API to explore the data of two cities. The databank is very useful because it is comprised of many information about the neighbourhood, like restaurants, hotels, parks, and others. Here we select Manhattan from New York and Downtown Toronto from Toronto. We will apply clustering to segment the neighbourhoods data. These objects aimed at locating tourist's areas and hubs, consequently to facilitate the comparison about dissimilarity for the choice of visit.

(3) Methodology

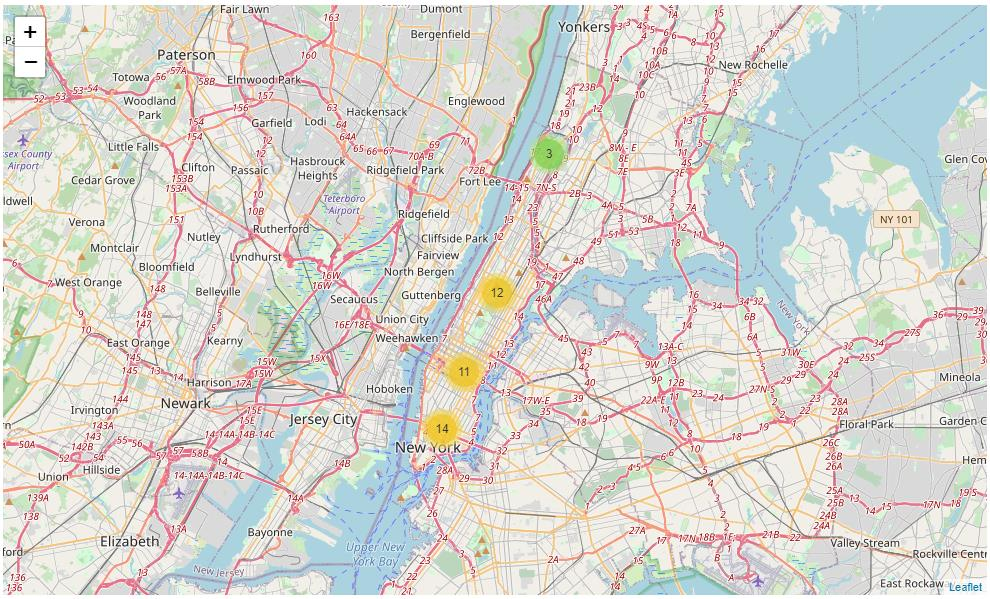
We have first extracted location information from the internet, then by using 'folium' package to visualise the map. Further we correlated with data from Foursquare to acquire more contents per location, like type of venue and etc. Last part is we using clustering to sort out the data of two cities.

(4) Results

For the city of Toronto (as figure below), we choose Downtown Toronto as representative place because I assume most of tourists will visit this zone. Thus, I have extracted table of Toronto's Borough from a Wikipedia page. Then I sorted out the data by applying multiple steps, for example eliminating "Not assigned" values, and combining neighborhoods that have similar geographical coordinates. Next I incorporated Foursquare API to get the information of Downtown Toronto and explore its neighbourhoods. The neighbourhoods are further characterized by venues.



For Manhattan (as figure below), I applied the location data from a saved data file in week3 which is already explored through foursquare API. In this case, we have extracted all the boroughs of Manhattan and then sorted against the concerned borough.



I noticed that the historical place is only situated in Downtown Toronto and the Monument or landmark venue is only in Manhattan neighbourhoods. Airport facility, harbor, sculpture garden and boat or ferry services are also available in Downtown Toronto. Venue for nightlife, climbing gym and museums are present in Manhattan.

After clustering the data of Downtown Toronto and Manhattan, I think both cities have venues which can be explored and attract the Tourists. I will think so because the neighbourhoods are much similar in features like theaters, opera houses, food places, clubs, museums, parks etc. But as mentioned, they are different in terms of some unique places like historical places and monuments.

(5) Discussion

For a long-term travel, I will recommend Manhattan because the density of venue in Manhattan is higher than Downtown Toronto. If that's a short-term travel, I will recommend Downtown Toronto, because the tourists will have a short access to airport.

(6) Conclusion

After visualising the venue on the map, we can see both the downtown Toronto and Manhattan have many similar venues, thus both can attract tourists. If to visit historical places is important for the tourist, Downtown Toronto will have more venue to visit.