New York City and Toronto City Neighborhood and Venues Analysis

1. Introduction

Nowadays, people, workers and industries are moving constantly from one location to another depending on their business, financial and family needs that are continuously changing and evolving.
Due to the Corona pandemic, employees, also supported by companies, are increasingly working from home, within the same country or in a country other than their workplace. This allowed the worker to be given more freedom than in the past, that is to travel to different places in the workplace. This puts the worker in front of a choice on which place, or rather neighborhood of a city, it might be more congenial to move to. Anyone who has moved has asked themselves the following questions: which neighborhood should I move to? What facilities would I like to have near my residence? I am happy in my neighborhood in New York City, what could be the most similar in the city of Toronto?
When we consider all these problems, we can create a map to compare the different cities with neighborhood superimposed on top and clustered based upon similarities, facilities and venues that each neighborhood offer.

For this project I have decided to apply the Data Science methodology to analyze the similarities of neighborhood of two big multinational and multicultural city, New York City and the Toronto City which are the financial capitals of their respective countries, US and Canada.

2. Data Description

To consider the problem we can list the datas as below:

- I have taken Toronto geographic data from: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M.
  Afterwards transformed it into a pandas data frame using BeautifulSoup. The file contains Postal Code Id, Borough and Neighborood for the city of Toronto.

	PostalCode	Borough	Neighborhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Queen's Park	Ontario Provincial Government

- I have decided to work on the csv file (https://cocl.us/Geospatial_data) have first imported and then red the csv file and after I have performed a left join between the two tables to add both Latitude and Longitude to Toronto dataset.

	PostalCode	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

- I have taken the NY data from: https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701ENSkillsNetwork/labs/newyork_data.json.
  The .json file contains Borough, Neighborhood, Latitude and Longitude for each Borough of New York.

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

- At latest, I used **Forsquare API** to get the most common venues of given Borough of New York and Toronto.