COURSERA CAPSTONE-Business Problem

1. Project description and discussion of the background:

In this project, we will use data science to evaluate how similar or dissimilar two boroughs/ cities are in terms of venues. We will work on Manhattan and Toronto boroughs to illustrate the concept.
The system will be helpful to solve several practical issues. For example:

- Business case 1: Company K which is already in Toronto wants to extend his business in Mahanttan, and want to know how similar or dissimilar Toronto and Mahnattan are in term of venues. This information is critical to company K, since it can decide to replicate the same distribution system or to change it.
- Business case2: Mr. Dupont is living now in the "Marble Hill" neighborhood in Manhattan. Mr. Dupont loves this neighborhood very much. He is leaving for Toronto and wants to find something quite similar to "Marble Hill" in Toronto.

Aside from those two business cases illustrated here, there are many other real-world problems that the tool developed here could solve.

2. Data :

We will use the data :

- For the Toronto borough
    - Web scrapping (url= https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) to get the list of postal code, Borough and Neighborhood in Toronto (
    - Geospatial data for Toronto: 'Geospatial_Coordinates.csv'
    - The data is saved in the file Toronto_Data.csv and is available in the same repository on Github.
- For Manhattan borough:
    - Download and process list of boroughs and neighborhoods in NewYork https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork_data.json
    - Save the data specific to Manhattan in a file called 'Manhattan_Data.csv' and available on Github

3. Foursquare API :

We use Foursquare API to obtain venues about each neighborhood.

4. Librairies

- Pandas
- Numpy

- Matplotlib
- Folium
- Plotly