I.   Data description

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

- The Toronto Neighborhood Coordinates are obtained from the List of Postal Codes of Canada (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) and the coordinates of each postal code (https://cocl.us/Geospatial_data), clean the data that don't have Borough or Neighborhood, group by Boroughs and Postal code. Then, the Postal Code is matched between the first data and the postal code coordinates, to finally get the coordinates of each neighborhood. And then finally takes only the Center of Toronto

In [68]: toronto_data.head()

Out[68]:

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M4E	East Toronto	The Beaches	43.676357	-79.293031
1	M4K	East Toronto	The Danforth West,Riverdale	43.679557	-79.352188
2	M4L	East Toronto	The Beaches West,India Bazaar	43.668999	-79.315572
3	M4M	East Toronto	Studio District	43.659526	-79.340923
4	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790

.

- The New York Neighborhood Coordinates was downloaded from the Json file (https://cocl.us/new_york_dataset), then capturing the Borough, Neighborhoods and the location (latitude and longitude), and finally takes the Manhattan's information. And then finally takes only the Center of Manhattan.

In [71]: manhattan_data.head()

Out[71]:

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

- The Paris Neighborhood Coordinates is loaded from the CSV file (https://opendata.paris.fr/explore/dataset/arrondissements/download) after the information is cleaned and structured in a data frame. And then finally takes only the Center of Paris.

```
In [69]: paris_data.head()
```

Out[69]:

	Neighborhood	Latitude	Longitude
0	Bourse	48.8682792225	2.34280254689
1	Temple	48.86287238	2.3600009859
2	Reuilly	48.8349743815	2.42132490078
3	Louvre	48.8625627018	2.33644336205
4	Hôtel-de-Ville	48.8543414263	2.35762962032

- The venues of each neighborhood are obtained by Foursquare with a limit of 100 venues and a radius of 500 of the center of each neighborhood. For those venues is saved the latitude, longitude and venue category

```
In [72]: venues_merged.head()
```

Out[72]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	The Beaches	43.6764	-79.293	The Big Carrot Natural Food Market	43.678879	-79.297734	Health Food Store
1	The Beaches	43.6764	-79.293	Grover Pub and Grub	43.679181	-79.297215	Pub
2	The Beaches	43.6764	-79.293	Guru Raghavendra Ji	43.680187	-79.292337	Astrologer
3	The Beaches	43.6764	-79.293	Upper Beaches	43.680563	-79.292869	Neighborhood
4	The Danforth West,Riverdale	43.6796	-79.3522	Pantheon	43.677621	-79.351434	Greek Restaurant