**Capstone Project - The Battle of Neighborhoods (Week 1)**

**1. Introduction**

**1.a Background**

In Module 3, we explored New York City and the city of Toronto and segmented and clustered their neighborhoods. Both cities are very diverse and are the financial capitals of their respective countries. One interesting idea would be to compare the neighborhoods of the two cities and determine how similar or dissimilar they are.

**1.b Problem**

Data required to determine similarity or dissimilarity between neighborhood might include official division of the city, data about arrangement of venues inside the neighborhoods, dominant venues, e.g. Hotels, Banks, and of course the location data for neighborhoods and venues. All of these for both cities, New York and Toronto. Base on that I could try to segment and cluster both city neighborhoods. There is no certainty, that these data will be enough.

**1.c Interest**

Similarities or dissimilarities between cities and they neighborhoods could be very useful for companies.

Let’s imagine that company ABC has headquarter or a branch in New York, in neighborhood BCD, and is satisfied with the location and surroundings. They are looking for location, to open a branch in Toronto. One of the requirements is to choose the neighborhood, which is very similar to one in New York.

Segmenting and clustering will be done for all neighborhoods in both cities together, based on popular venues in each neighborhood.

Let’s assume that company ABC is a SPA, located in neighborhood with many parks or gardens. I will try to find corresponding neighborhood in Toronto which is similar. Technically both neighborhoods should be placed in same cluster.

Here is a list of open questions:

* Is it possible to determine neighborhoods like the one in New York based on popular venues located in each neighborhood?
* Are there any patterns how neighborhoods are organized inside a city?
* How different or similar the cities are?