**Coursera Capstone Project – IBM Data Science Professional**

**Introduction/Business Problem**

The number of immigrants to the cities like Toronto, Paris, New York etc. has been on a rising note over the past few years. These cities have been the financial hubs and also capital cities of their respective countries. As a reason, lot of immigrants are moving to these cities to find jobs etc. Also, lot of tourists will be visiting these cities throughout the year. When people wish to move to these cities or make a trip to visit them, they are not quite enough resources for them to be aware of how similar or dissimilar the new city is going to be in terms of schools, weather, life style, food etc. from the city they are moving from. This makes it quite hard for them to decide and be prepared for the new city. So, I would like to come with a solution where given two cities, it can tell how similar/dissimilar the two cities are.

**Data**

Make a list of cities and use the foursquare api to get the data. Once we get the data, we need to come up with a one feature vector for each city by aggregating the data in some way so that some distance measure like Euclidean distance, Cosine similarity can be used to find the similarity measure.