**Introduction: Business Problem****& Background**

In this project we will try to find an optimal location to open a new café in New York. This café is a multinational chain and is known to be loved by customers. Till now at all locations it has opened, the café has got consistently very high ratings

This café is targeted for tourists and hence would it would be optimum to locate it where there are more tourist attractions nearby but at the same time are not overly crowded by cafes.

Hence, it’s important to find locations that have enough tourist attractions as a first parameter. Once identified, these locations will have to be checked for existing cafes (the rating on each of the cafés might help also look at a qualitative outlook and hence offer a good café in an area where current cafes do not meet customer standards).

We will try and create location clusters which are most optimal for the required problem. We can explore further opportunities based on the data

**Data****– Approach and Requirements**

We are going to use the New York data (used in Week 3) + venues data from Foursquare to help us reach the desirable solution

Based on definition of our problem, factors that will influence our decision are:

* number of tourist attractions in the neighborhood
* existing number of cafés in these neighborhoods
* average rating of cafes in the neighborhood

Following data sources will be needed to extract/generate the required information:

* New York locations will be extracted from existing data shared during the previous courses
* Various venues will be explored and tourist places + cafes will be explored in every neighborhood using **Foursquare API**
* coordinate of New York for the initial map will be hardcoded from data available online
* Initially top tourist areas will be identified, post which café data will be analyzed and ideal locations will be found