INTRODUCTION

## Description of the data and objective to solve the problem

The data for this project is exclusively provided by Foursquare location data. A free Forequarter’s API key will be used to query the data needed from Foursquare database in JSON (JavaScript Object Notation) format and saved in CSV (Comma Separated Value) format for further manipulations. Due to the limited number of venues per results of a free API key, 100 venues will be queried for the 5 most populous cities/states for each of the 7 countries of the G7 countries.

To achieve the project objective, the following specifics objectives will have to be achieved:

1. Get and clean all data and prepare them for analysis
2. Cluster venues per cities/states and compare cities/states inside each country
3. Cluster venues by countries and compare the countries inside the G7
4. Cluster venues of the G7 and get a global insight about venues for the whole G7 countries.

Clustering algorithm that will used in this project is mainly k-means.