Capstone Project - The Battle of Neighborhoods (Week 1)

# Background:

A successful supermarket chain (Imtiaz Super Market) of grocery, utensils, toys and clothes and other household stuff wants to expand its business by opening new branches in USA. But before pouring such a huge investment and time into the mentioned idea the owners want to know that if store opening in USA would really provide considerable elevation in their present profit and that they can really convince the people to be their buyers despite of other stores already present for shopping. Stakeholders also interested to know the hobbies and likes dislikes of people in terms of food and ambiance to get an idea to design their supermarkets accordingly so they can target maximum number of people.

## Problem for Data Scientist:

* Provide a statistics using Foursquare location data base which cities/places in USA have shortage of Supermarkets and where there is a gap to enter for Imtiaz supermarket.
* Also provide the comments assessment of people about various supermarkets in the town and what is gap in them using rating system
* What gentry visits supermarkets and most likely to be our buyers so that we can plan the stuff quality and budget accordingly using Foursquare location data
* Using Foursquare database, provide clusters of people in diff cities who have similar type of hobbies and likes/dislikes in the context mentioned above to analyze which theme and entertainment/food section should be open in the supermarket to pull the crowd

# **Data** section

### We will use following set of data for the execution of this project:

* Get the latitudes and longitudes of different cities USA using geocoder to feed into Foursquare Location service
* Foursquare location data to access different locations of USA cities
* Foursquare data set of supermarkets, ratings, Comments and members
* Use Folium to represent clusters of data set on a map in given location
* Use K-Mean clustering to cluster similar groups