**Introduction to the problem**

The city council of the city of Chicago is interested in whether there is a correlation between the types of crime that frequently occur in a community area and the popular venues that the area has to offer to its citizens. For example, in an area with many banks and jewelry stores, robbery and burglary may be popular types of crime to be committed. If there is such a correlation, we can predict the potential types of crime that may arise in a certain community area and, therefore, will have better solutions to prevent those from happening.

**Data**

The Foursquare location data is very useful for this problem since we can retrieve the popular venues in a specific community area. In addition, we will make use of following data set

<https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2>

This data set is publicly available and can be downloaded as a csv file. It contains all reported crimes in the city of Chicago from 2001 and is updated on a regular basis. According to this data set, there are over 7 million reported cases. For simplicity, we will only consider the cases that occur during and after 2015. This results in a subset with over 1 million cases. Each observation in the data set contains useful information about the case, including the primary type of crime, the community area in which it happened and the precise location in terms of longitude and latitude. Using this data set, we will know the longitude and latitude of a community area and, therefore, easily retrieve the list of venues for that area by making an API call to Foursquare.