**Introduction**

In This project will analyse the New York City data. Along with it, we will find the most visited commercial shop according to the number of check-ins, then we will try to find the neighbourhood’s that are lacking the given type of shop which Could be potential business opportunity.

**Target Audience**

The target audience of this project is any one that is interested in opening a shop but also have no idea of what kind of shop they want to open and in which neighbourhood should the shop will be situated.

**Data Section**

The data comes from **Dingqi Yang** from the following ink [https://sites.google.com/site/yangdingqi/home/foursquare-dataset](#_top)

It contains 227,428 check-ins in New York City. The data contains a file in tsv format. Each file contains 8 columns, which are:

* User ID (anonymized)
* Venue ID (Foursquare)
* Venue category ID (Foursquare)
* Venue category name (Foursquare)
* Latitude
* Longitude
* Time zone offset in minutes (The offset in minutes between when this check-inoccurred and the same time in UTC)
* UTC time

**Application**

We will find the most visited type of commercial shops according to the number of check-ins given in the data which came from the internet, then we will try to find neighbourhood’s that has none of this type of shop. Examples are for 2000 venues, and the red dot is the centre neighbourhood which has the most number of Bars between selected coordinates. We did find two neighbourhood’s that are closest to it having none Bars within 4 kilometres.







