# Location Recommender System for Opening a New Business in Your City

**Ogety Sai Vamsi**

**Date: September 15, 2019**

## 1. Introduction

### 1.1 Background

The city of Phoenix is the most populous city in the Arizona and in fact it is the 5th most populous cities in the United States. The region offers a skilled, diverse workforce fed by globally recognized universities and community colleges. The Cost of living is 42% less than Northern California and it offers a robust pro-business climate. The City is busy with entrepreneurial activity and resources to support it. Because of many resources and low cost of opening a business many people are actively seeking to start their businesses.

### 1.2 Problem

Suppose you are an entrepreneur who is planning to start a business and Let's say you want to open a Rental Car business in Phoenix. Choosing a location for your business is one of the key things in business. If you are suppliers or distributors are living nearby, your cost of operations will be less. It is also important to know whether it's a centre for products or services you are providing. The Ideal location to open your business would be in the most happening place in the city but It depends on the type of business and the budget you have. You can't open an Amusement park in the centre of a Tech- hub. You need to know about the buying or leasing costs and knowledge about growing business hubs with has potential for the future. This process is very painstaking, and it takes a lot of time and sometimes the information you have gathered may or may not be accurate or may be its just somebody's opinion.

This Project aims to suggest locations (Neighbourhoods) by running Clustering Algorithms on the Venues data provided by Foursquare API.

## 2. Data Source

We will be using New York city Neighbourhood Data Set. Luckily, this dataset exists for free on the web. Here is the link to the dataset: https://geo.nyu.edu/catalog/nyu_2451_34572
It has a total of 5 boroughs and 306 neighbourhoods. We extract Borough, Neighbourhood, Latitude and Longitude coordinates of each neighbourhood from the dataset to explore and segment the neighbourhoods.

## 3. Strategy to Solve the problem

Location (Neighbourhood) to the user is suggested considering the following steps:

1. Select the neighbourhoods where they provide services or products like yours.
   i.e. If you want to open a Car Rental business the neighbourhood suggested will be ones where they have Automotive businesses.
2. After selecting the relevant neighbourhoods, we filter them by checking whether any of these neighbourhoods have already the businesses that you want to establish.
   i.e. If the location has already a rental car business and is doing well opening your business there will be risky So we suggest the other remaining neighbourhoods to the user.

We will be using K-means Clustering algorithm to group neighbourhood. The grouping of the neighbourhood is done based on how similar each neighbourhood is and the metric used here is to compare the top 10 venues in each neighbourhood. We will be using Foursquare API to explore Venues around a location. The libraries that we are going to use are Pandas for Data-pre-processing and Exploratory Data Analysis. Matplotlib and Folium for Data Visualization. Geocode for converting Addresses to Latitude and Longitude.