**Analyzing The Community Areas of Chicago to start A New Restaurant**

Kerry Goode Jr

June 6, 2021

**Introduction**

The City ofChicago is the most populated city in the State of Illinois and the third most populated city in the United States. Chicago had an estimated population of 2.6 million people in 2019 which makes it the most populated city in the Midwestern United States and the fifth most populated city in North America.

The population of Chicago consist of people with various ethnicities from all over the world. The city has numerous restaurants which serve thousands of customers per day. Given the diversity in the population, there are a vast number of restaurants with different cuisines for everyone to enjoy. In this project we will examine the community areas in Chicago and make a recommendation on the location where a restaurant should be started.

**Business Problem**

A group of investors are looking to invest and start a franchise restaurant in Chicago. They have approached us to study the market and provide them with a recommendation on which community area in Chicago where they should invest and start their franchise. The main objective of this project is to obtain and analyze the right data about the various community areas in Chicago and recommend to our client a suitable location for their restaurant.

**Data**

The following data is needed to complete our analysis and to provide a recommendation:

1. Community Areas of Chicago
2. Geographical coordinates of the Community areas
3. Venue Data from FourSquare

**Community Area Data**

This data was extracted from the Community areas of Chicago page ([Category:Community areas of Chicago - Wikipedia](https://en.wikipedia.org/wiki/Category:Community_areas_of_Chicago)) using web scraping with BeautifulSoup library in Python. This will give us a specified list of community areas in Chicago.

**Geographical Coordinates of the Community Areas**

Geographical coordinates are needed for plotting maps to visualize the data. The geographical coordinates of the community areas were extracted using GeoPy library in python. Once we use GeoPy we will add two columns to our data frame with latitude and longitude information for each community area.

**Venue Data from FourSquare**

Using FourSquare data we extract Venue data. The venue data is used to study the venues in various community areas in Chicago. This data provides much needed details of the different resturants in those areas to understand how competitive it is in each area. The addition of this data will helpsus make a more informed decision at the conclusion of the project.