# The Battle of Neighborhoods | Introduction

## Introduction:

New York City, the most populous city in the United States, one of the greatest metropolises over the world, is a dream place for gourmet to seek delicious cuisine. Its food culture includes an array of international cuisines influenced by the city's immigrant history. The city is home to "nearly one thousand of the finest and most diverse haute cuisine restaurants in the world," according to Michelin. As of 2019, there were 27,043 restaurants in the city, up from 24,865 in 2017.

The purpose of this Project is to help Chinese restaurant owners to find the best place to start their business around their neighborhood. It will help people make smart and efficient decisions on selecting great neighborhoods out of numbers of other neighborhoods in New York.

## Data Preparation

Data used in the analysis are listed below:

- Neighborhoods in New York City https://cocl.us/new_york_dataset
  I cleaned the data and reduced it to boroughs of NYC so that I can use it to find geological locations for further venue analysis.
- Using Geopy to get geological location by address name
- Using the Foursquare API to get the most common venues of a given Borough of New York City.
- Using Foursquare API to get the venues' record of given venues of New York City.
- Data will be sorted based on rankings.
- Finally, the data will be visually assessed using graphing from Python libraries.

## Problem Statement

1. What is / are the best location(s) for Chinese cuisine in New York City?
2. In what Neighborhood and/or borough should the investor open a Chinese restaurant to have the best chance of being successful?

## Libraries Which are Used to Develop the Project:

Pandas: For creating and manipulating dataframes.
Folium: Python visualization library would be used to visualize the neighborhood cluster distribution by using an interactive leaflet map.
JSON: Library to handle JSON files.
Geocoder: To retrieve Location Data.
Beautiful Soup and Requests: To scrap and library to handle http requests.
Matplotlib: Python Plotting Module.