# Capstone Project

## Introduction/Business Problems

**Project Description:**

In a city of your choice, if someone is looking to open a restaurant, where would you recommend that they open it?

The city of my choice is Hong Kong, due to the fact I am from there and I would like to know more about the city through this project. Therefore, the aim of this Capstone Project will be to help people explore different possibilities and take a better decision when it comes to opening a restaurant. Similarly, if a contractor is trying to start their own business, where would you recommend that they setup their office?

**Selection Criteria:**

For the purpose of this project, the definition of a good location to open a restaurant is a place with parking and good visibility of the restaurant. On the other hand, the definition of a good location for an office is the low price with a good location meaning employees and clients will be able to assess to that office easily.

**The Location:**

Hong Kong is a city and special administrative region of China in the eastern Pearl River Delta by the South China Sea with over 7.4 million people with various nationalities in a 1,104-square-kilometer territory, making Hong Kong one of the most densely populated places in the world. Due to the fact that Hong Kong is a small place with a high population, parking places will not be available or will be too high costs for a restaurant, therefore this project will be locating the restaurant in the most visible place also locating the most convenient office.

**Foursquare API:**

This project would use Four-square API as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business.

**Work Flow:**

Using credentials of Foursquare API features of nearby places of the neighborhoods, however due to the HTTP request limitations the number of places per neighborhood parameter would be set to 100 and the radius parameter would be set to 500.

**Clustering Approach:**

For this study, K-mean clustering will be the main tool.

**Libraries:**

*Pandas*: For creating and manipulating data frames.

*Folium*: Python visualizing library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.

*Scikit Learn*: For importing k-means clustering.

*JSON*: Library to handle JSON files.

*Geopy*: To retrieve Location Data.

*Requests*: Library to handle http request.

*Matplotlib*: Python Plotting Module.