Capstone Project

* **Part 1: Intro/Business problem**

**Finding the best location in GTA to open a Chinese food market**

Mr. Chen is a new immigrant who owned a chain supermarket company back in China.

He would like to apply his business talents to the Canadian market.

With everything else prepared, he now needs to know which community is the best choice for him.

His requirements for the ideal place are:

1. large Chinese population;

2. few competitors;

3. reasonable rent.

* **Part 2: Data**

To meet Mr. Chen’s requirements, we need both location data and demographic data for the GTA area. Hence the 2 main data sources will be **Statistics Canada,** **Foursquare** and **CMHC**.

1. To Find communities with large Chinese population:

This will be solved by leveraging the demographic data from Statistics Canada for GTA in the most recent years. First is to locate the promising communities where lots of Chinese people reside. We will look at two numbers, one is the total number of Chinese population, and the other is percentage of Chinese to the total population in that area. We will need these numbers for the most recent 3 years, as a reference to the demographic trend in communities.

Example: Scrape data from Statistics Canada for Chinese communities for most recent 3 years. Clean data and put it into a data frame, with community names, population for 3 years, and longitude/latitude. Then use folium to make a heat map showing the density of Chinese population in GTA area. We need to look for places with high density as well as watch the trend of population change in recent 3 years, which can be shown in a plot with normalized population data.

1. To find communities with few competitors:

Foursquare will be the main source of data for this step. In the previous step we get a list of communities with large and growing Chinese population. Using the geographical info of those communities, we put it into Foursquare and explore those areas.

Example: We got ABC community as a promising spot. By entering the geographical info of ABC community in Foursquare, we will explore some key words like “Chinese”, “Asian market”, etc. to see the number of existing competitors around the area. We then take note of the number for each community and divide the Chinese population by it, to see the potential market space for a new Chinese market in that area.

1. To find a place with reasonable rent

CMHC will be the data source for this step. With the shortlisted communities got from the previous steps, we need to find out which one has the best rent price.

Example: We will find corresponding CMHC rental price for business/warehouse for those communities, and put the data into a data frame. Then a simple plot will show which place has the best rent.