**Asian Ventures in New York**

Jinyuan Ng

31 May 2020

1. **Introduction**

Background

New York (NY), is the most populous city in the United States. With an estimated 2019 population of 8,336,817 distributed over about 302.6 square miles (784 km2), New York is also the most densely populated major city in the United States. NY also boasts being the Top 7th Most Visited City in the world, receiving 13.6million international visitors in 2019. This fact alone is enough to entice investors from all around the world to invest in NY, being opening a business or in real estate etc. However, it may seem daunting for Asian investors, as it’s hard to predict which of the major 5 boroughs should they open their business in or which area for property investments. Therefore this dataset would be advantageous for them, to predict which borough and their respective neighbourhood would welcome their business as compared to the rest. Using Foursquare API, I’ll create a map based on each borough with chosen locations as map markers.

Aims

This project aims to predict the top common venues within each borough, as well as consumer’s expenditure preferences.

Stakeholders

Asian investors who’s keen to venture into a business/ invest in NYC, as well as Asians who are planning to visit / migrate / study in NYC.

1. **Data Acquisition and cleaning**

Data Acquisition

I made use of Foursquare API to obtain coordinates of the respective boroughs, their respective neighborhoods and their common venues. I also obtained data sets from the US Bureau of statistics and graphs from NYC Travel & Tourism Report 2019 to understand NYC consumers more.

Data Cleaning

It would be too much of a hassle to go through all the common venues per neighborhood, so I decided to find the top 5 common venues per neighborhood first, before consolidating only the neighborhoods which has a Asian restaurant / business in the top 5 common venues per borough. Therefore, it’d be easier for investors to see in a glance how welcoming of the neighborhoods would be with their businesses.

The data sets obtained has many categories, such as ‘Food’ being separated into ‘Food at home’, ‘Food away from home’, ‘Alcohol’ etc, and ‘Income’ being separated into consumers per unit. Therefore I only chose the rows ‘Income after taxes’, ‘Food away from home’ and ‘Housing’ from all consumer units as Food and Housing emerged as the Top 2 in terms of annual expenditure distribution in the year 2017-2018.