**The Battle of Neighborhoods Report**

1. **Introduction:**

1.1 Background

The City of New York, is the most populous city in the United States. It is diverse and is the financial capital of USA. It is multicultural. It provides lot of business oppourtunities and business friendly environment. It has attracted many different players into the market. It is a global hub of business and commerce. The city is a major center for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, advertising, legal services, accountancy, insurance, theater, fashion, and the arts in the United States. This also means that the market is highly competitive. As it is highly developed city so cost of doing business is also one of the highest. Thus, any new business venture or expansion needs to be analyzed carefully. The insights derived from analysis will give good understanding of the business environment which help in strategically targeting the market. This will help in reduction of risk. And the Return on Investment will be reasonable.

* 1. Business Problem:

The City of New York is famous for its excellent cuisine. It's food culture includes an array of international cuisines influenced by the city's immigrant history.

Thai food is quite diverse and mixed with various spices, It's make Thai food so popular in the United States, but only in Major cities. Therefore, the first thing that must be done before investing is to find the best location. It's a good opportunity to start investing in restaurants in a minor cities.

* 1. Target Audience

To recommend to the stakeholder or investor who was interested to invest Thai restaurant in Manhattan city at New York.

1. **Data acquisition and cleaning**

2.1 Data Source

New York Datasets has a total of 5 boroughs and 306 neighborhoods. In order to segement the neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the the latitude and logitude coordinates of each neighborhood. Luckily, this dataset exists for free on the web. Feel free to try to find this dataset on your own, but here is the link to the dataset: <https://geo.nyu.edu/catalog/nyu_2451_34572>

A screenshot of a cell phone

Description automatically generated

Foursquare API that will be leveraged to provision venues information for each neighborhood. We will use Forsquare API to explore neighborhoods in New York City with specific category for this case is Thai restaurant. The category id is “4bf58dd8d48988d149941735”

1. **Methodology**

3.1 Data Pre-processing and Understanding

Convert New York datasets into tabular form and then select the datasets only Manhattan city and visualize to ensure the data is selected correctly. Then I will use the Foursquare API to get Thai Restaurant in Manhattan city. The result from Foursquare API has 1095 records. Then visualize the neighborhood has few Thai restaurants.

A screenshot of a cell phone

Description automatically generated

A close up of a logo

Description automatically generated

A close up of a map

Description automatically generated

Thai Restaurant in Manhattan

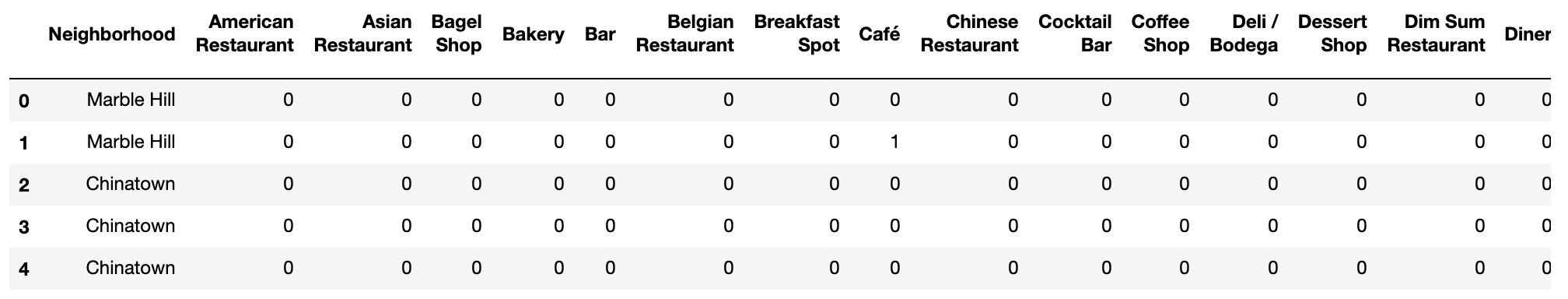
A screenshot of a cell phone

Description automatically generated

The neighborhood that has the least Thai restaurants

3.2 Feature Engineering

I use the one-hot encoding method to encode category data into a binary. then group the output from one-hot encoding to represents ratio Thai restaurant each neighborhood.



One-hot Encoding

A close up of a piece of paper

Description automatically generated

Ratio Thai Restaurant in Manhattan

3.3 Modeling

I use K-means clustering algorithms to cluster neighborhood at Manhattan. The evaluation of the model using the elbow method so the optimal K is 6 ks.

A screenshot of a cell phone

Description automatically generated

SSE using Elbow Method

1. **Results**

Cluster 0

**A screenshot of a cell phone

Description automatically generated**

Cluster 1

A screenshot of a cell phone

Description automatically generated

Cluster 2

A screenshot of a cell phone

Description automatically generated

Cluster 3

A screenshot of a cell phone

Description automatically generated

Cluster 4

A screenshot of a cell phone

Description automatically generated

Cluster 5

A screenshot of a cell phone

Description automatically generated

Visualize 6 Clusters

A close up of a map

Description automatically generated

Cluster 0 – Red point ()

Cluster 1 – Purple point ()

Cluster 2 – Blue point ()

Cluster 3 – Aquamarine point ()

Cluster 4 – Mint Green point ()

Cluster 5 – Orange point ()

1. **Discussion**
   1. This analysis perform on limited data. This may be correct or maybe not but if massive of data is available there will be better result.
   2. There is scope to explore restaurant in Manhattan only
2. **Conclusion**

The best place to investing the Thai restaurant is Central Harlem. Inferior are Roosevelt Island and Tribeca have a nice Asian Restaurant environment. Thus, the people who want to eat Asian food usually go to those places.