1. **INTRODUCTION**

**1.1. Description of the Problem**

One of the problems of small food businesses when they are just starting is the competition with established restaurants or with stores that offers the same menu but are already popular in that area. Another is the selection of locations for their new branches, since it is important to identity the nearby establishments and their market or customers so that they will have an idea of how their products will be received.

**1.2. Discussion of the Background**

Yummy Tokyo is a small Japanese minimalist casual restaurant concept that specializes in authentic, affordable, and yummy ramen. The business opened their first restaurant in Quezon City last February 2019 and since then has opened four more branches in various locations within Metro Manila in the last year. The business mainly targets professionals and families as their customers.

This project has two main objectives: (1) to identify the nearby venues that are ideal for food businesses and (2) to identify new locations in Metro Manila where Yummy Tokyo can expand their business. The project will focus on barangays and cities in Metro Manila, which is primarily the central business district in Luzon, Philippines. From the 2015 population census conducted by the Philippine Statistics Authority (PSA), there are about 13 million Filipinos residing in Metro Manila, which is more than 10% of the total national population which is about 106 million.

**1.3. Interest**

The study is primarily intended for the Yummy Tokyo restaurant but can be helpful for other small businesses. The study is not only limited to businesses that plans to expand to different locations but also to those who wants to start their business. The project is also not limited to food but also to other types of services.

1. **DATA**

**2.1. Data Sources**

The following data were collected for this project:

* The latest available census of Metro Manila which can be collected from the PSA website or from Wikipedia is of 2015 and was scraped using the python *BeautifulSoup* library. Collected with the census were the list of different barangays and cities in Metro Manila.
* The geocoder library was used to acquire the coordinates of the barangays and cities.
* The manila.json file was taken from github account of altcoder.
* The list of the Yummy Tokyo branches was prepared in an excel sheet. The locations of each branch were collected from google maps.
* The names and locations of other ramen restaurants were collected using FourSquare API, though the list that was generated were of the popular restaurants only.
* The venues near the five Yummy Tokyo branches were collected using the FourSquare API.
* The venues near the other stores were collected using the FourSquare API.
* The venues for each barangay were collected using the FourSquare API.
  1. **Data Cleaning**
* The census collected had missing population data, so these barangays were removed from the data set. There were originally 519 barangays from the original list but were cut down to 424 after removing the barangays with missing population data.
* There were originally 17 cities/municipalities listed, but due to unavailable population data, it was reduced to 14. Included in the cities/municipalities that were dropped are Caloocan, Pasay, and Pateros.
* Zip codes from the scraped data were removed from the data set since the coordinates of each place can be collected even with using only the barangay and city names.
* The population data were of type string and there were commas as thousand separators. The commas were removed, and the values were converted into type int.
* Some of the barangays of different cities have the same names, so using the *‘groupby’* function by ‘Barangay’ names merges these redundant names. In order to address this problem, after the coordinates of each locations were acquired, the barangay and city names were joined into one column using the *‘aggregate’* and ‘join’ functions.

**3. METHODOLOGY**

The *BeautifulSoup* library was used to scrape the data from Wikipedia. Unnecessary data were removed and barangays with empty population cells were dropped. After the data has been cleaned, the generated table had three (3) columns: ‘Barangay’, ‘City’, and ‘Population’. Using the python *geocoder* library, the coordinates of each barangay were collected and added to the table. A table of the 14 cities, their coordinates, and the population per city was generated as well.

A screenshot of a cell phone

Description automatically generated

Table 3.1 Head of Barangays, the City they belong to, their Coordinates, and Population Size

A picture containing table, black, wooden, display

Description automatically generated

Table 3.2 Head of Cities of Metro Manila, their Coordinates and Population

The locations and coordinates of the five Yummy Tokyo branches were taken from google map since the exact locations of these branches were not attainable when using the *geocoder* library and were not available in FourSquare.

A screenshot of a cell phone

Description automatically generated

Table 3.3 Branches, Locations, and Coordinates of the Five Yummy Tokyo Branches

With the use of the python *folium* library, the locations of the five branches were visualized on top of a choropleth map of the city populations. The blue circles mark each of the branches, and the labels were configured to pop-up only when clicked.

A close up of a map

Description automatically generated

Figure 3.1 Choropleth Map of the City Populations and marked with the 5 Branches of Yummy Tokyo

Using the FourSquare API, the locations of other ramen/Japanese restaurants in the 14 cities were collected. A total of 14 venues were generated by FourSquare. The locations were visualized along with the Yummy Tokyo branches. The black circles mark the locations of the other restaurants.

A screenshot of a cell phone

Description automatically generated

Table 3.4 Popular Ramen Restaurants in Metro Manila

A close up of a map

Description automatically generated

Figure 3.2 Choropleth Map marked with the 5 Branches of Yummy Tokyo and 14 Other Ramen Restaurants

Using the FourSquare API, the venues near the branches were collected. The search was filtered to the following categories: (1) college and university, (2) professional and other places, (3) residence, (4) travel and transport, (5) shop and service, (6) outdoors and recreation, (7) nightlife spot, (8) event, and (9) arts and entertainment, which are all the main categories in the FourSquare category list, except for the food category. Food was not used as a filter since it was considered as a product of having the other venues. A total of 1206 venues were generated with 220 unique Venue Categories since the limit that was configured for each category was 100 with a radius of 500. By one hot encoding, the different Venue Categories were placed in columns and tabulated against the Yummy Tokyo branches. Then, the mean of the occurrence of each Venue Category was calculated and was used to determine the top 10 most common venue category for each branch. (In order to save space, Residential Building (Apartment / Condo) shall be referred to as Residential Building.)

A screenshot of a computer

Description automatically generated

Table 3.5 Top 10 Most Common Venue Categories near the Yummy Tokyo Branches

A screenshot of a cell phone

Description automatically generated

Figure 3.3 Summary of the Top 2 Most Common Venue Categories near the Yummy Tokyo Branches

For comparison, the nearby venues for the other restaurants were also collected using the FourSquare API using the same configuration and filter. A total of 2372 venues were generated with 195 unique Venue Categories. By one hot encoding, the different Venue Categories were placed in columns and tabulated against the restaurants. Then, the mean of the occurrence of each Venue Category was computed. The top 10 most common venue category for each branch are on Table 3.6. The summary of the Venue Categories most mentioned in the 1st and 2nd Most Common Venue are in Figure 3.4.

A close up of a piece of paper

Description automatically generated

Table 3.6 Top 10 Most Common Venue Categories near the Other Restaurants

A close up of a logo

Description automatically generated

Figure 3.4 Summary of the 10 Venue Categories most mentioned in the 1st and 2nd Most Common Venues near the Ramen/Japanese Restaurants

In order to identify the locations where Yummy Tokyo can expand their future branches, barangays that offer the same nearby venues were identified. The venues in each barangay were collected with the FourSquare API using the same filter, limit, and radius. The search resulted with a total of 26,325 venues with 465 unique Venue Categories. Like what was previously did, the 10 Most Common Venue Categories for each ‘Barangay and City’ were summarized in Table 3.7.

A screenshot of a computer

Description automatically generated

Table 3.7 Head of Top 10 Most Common Venue Categories in each Barangay

A screenshot of a cell phone

Description automatically generated

Figure 3.5 Summary of the 10 Venue Categories most mentioned in the 1st and 2nd Most Common Venues in each of the Barangay

The ‘Barangay and City’ column with the mean values were dropped so that the data can be used for the KMeans clustering. The optimum number of clusters of ‘2’ was derived with the Elbow Method using Distortion, but the Elbow Method using Inertia did not give a specific value. KMeans clustering was used because the data was not yet labeled, and an unsupervised machine learning method was necessary to look for the patterns common to the elements in the dataset. The KMeans clustering resulted in 2 clusters of sizes 252 for the 1st cluster and 169 for the 2nd cluster. A summary of the 10 most mentioned Venue Categories for each cluster were summarized in Figure 3.8 and Figure 3.9.

A close up of a map

Description automatically generated

Figure 3.6 Elbow Method Using Distortion

A close up of a map

Description automatically generated

Figure 3.7 Elbow Method Using Inertia

A screenshot of a computer

Description automatically generated

Table 3.8 Head of Cluster 1

A screenshot of a cell phone

Description automatically generated

Figure 3.8 Summary of the 10 Venue Categories most mentioned in the 1st and 2nd Most Common Venue in Cluster 1

A screenshot of a computer

Description automatically generated

Table 3.9 Head of Cluster 2

A screenshot of a cell phone

Description automatically generated

Figure 3.9 Summary of the 10 Venue Categories most mentioned in the 1st and 2nd Most Common Venue in Cluster 2

1. **RESULTS**

As seen in Figure 3.1, two of the restaurant’s branches are in Quezon City which has the largest population. Followed by \*Taguig and Pasig which are 3rd and 4th in population size, and then by Makati which has the 8th largest population among the 14 listed cities. On the other hand, other restaurants are more concentrated in the City of Manila and in Makati, as seen in Figure 3.2. (\*In the map, json file included Fort Bonifacio in Makati, but it is supposed to be part of Taguig)

The Yummy Tokyo branches, whose main target markets are professionals and families, have located their branches near offices and residential buildings. This is supported by the data collected in FourSquare, which is shown in Table 3.5 and in Figure 3.3, which shows that the 1st Most Common Venue for all branches are ‘Offices’, and 4 of the branches have the ‘Residential Buildings’ as either the 2nd or 3rd Most Common Venue Category.

In Table 3.6 and Figure 3.4, ‘Offices’, ‘Clothing Stores’, and ‘Residential Building’ and ‘Hotels’ were the most common venues, which is almost the same with that of Yummy Tokyo’s.

With the use of the KMeans clustering method, the different barangays were clustered into 2. In Figure 3.8, the three most mentioned venue categories for the 1st and 2nd Most Common Venues in Cluster 1 are (1) Offices, (2) Residential Buildings, and (3) Convenience Stores. While for Cluster 2, they are (1) Churches, (2) Convenience Stores, and (3) Housing Developments. Since Cluster 1 fits the criteria for the ideal location of the branch, these barangays will be considered as the prospects.

To narrow down the choices further, the barangays in Cluster 1 were filtered to those that have the Offices and Residential Buildings in their 1st and 2nd Most Common Venue. A total of 27 barangays were identified which will be referred to as ‘Recommended Locations’. In Figure 4.1, These locations are added to the map of the branches and other stores and are marked with red circles.

A screenshot of a computer

Description automatically generated

Table 4.1 Head of the Barangays that have the Venue Categories Offices and Residential Buildings (Apartment / Condo) as their 1st or 2nd Most Common Venue

A close up of a map

Description automatically generated

Figure 4.1 Choropleth Map marked with Yummy Tokyo Branches (Blue), Other Restaurants (Black) and Prospect Recommended Locations (Red)

1. **DISCUSSION AND RECOMMENDATIONS**

The Recommended Locations in the results section must only be treated as a reference in choosing the location for the new branches. It was observed that a coordinate in a barangay different from which was provided by python *geocoder* library will yield a different list of nearby venues. Unlike in Table 3.5, where the 1st Most Common Venue for all branches were offices, the 1st Most Common Venue for the barangay where the branches are located gave-out mix results. And only the Tektite branch in San Antonio, Pasig gave the same 1st and 2nd Most Common venue.

A screenshot of a computer

Description automatically generated

Table 5.1 Top 10 Most Common Venues for the Barangays of the Yummy Tokyo Branches

The study also limited the category of the other stores to Ramen/Japanese restaurants. There could be a variety in the category ID for Type of Food/Restaurants so that the study can also explore the results for other categories.

Using the FourSquare API for looking for other Ramen/Japanese restaurants yield very few results. The results were only limited to the popular stores, even the Yummy Tokyo stores were not identified.

The study could be improved if the data of revenue of a store is available and used. By having the revenue of the store and the trend of its growth, a model can be constructed of the effects of the different venues to the success of a store.

1. **CONCLUSION**

In this study, I used the FourSquare API to find the different venues near the Yummy Tokyo branches, venues near other Ramen/Japanese restaurants, and venues in Barangays in order (1) to identify the venue categories that are ideal for food businesses and (2) to identify prospect locations for Yummy Tokyo’s expansion.

The data shows that Yummy Tokyo and other Ramen/Japanese restaurants have located their stores near (1) Offices, (2) Residential Buildings, and (3) Clothing Stores (or more commonly known as malls). It is logical since the likely customers are those who can afford the service.

On the 2nd part of the study, recommended locations were identified. These are the barangays in Cluster 1 that have offices and residential buildings in their top 10 most common venues. The result of Cluster 1 was further narrowed down by filtering the barangays to those that have offices and residential buildings on their 1st and 2nd most common venue.