**Ideal Location for New Filipino Restaurant in Los Angeles, California**

Paris Ring

June 21, 2020

# 

## 1. Introduction

1.1 Background

Los Angeles, California is the home to nearly four million people and is the second most populous city in the United States. Los Angeles is known for the Hollywood Entertainment Industry, ethinc diversity and its extended metropolis. The Los Angeles metropolitan area is home to 13.1 million people. Los Angeles has a very diverse economy and has a wide variety of different businesses that make up its diversity.

Los Angeles is also home of the second-largest concentrated population of Filipinos in the world, only exceeded by Manila. The current Filipino population of the Greater Los Angeles area is ~606,000. As this population continues to grow, many other Filipinos in Los Angeles are looking to find their “little slice of home” within their hometown. But not only do Filipinos look for this tasteful cuisine, but many other Americans also are enjoying Filipino food, as it has grown in popularity within the last few years. With the population continuing to grow, there is a greater demand for more restaurants/bakeries. With the current COVID-19 situation, and the inability to travel to the Philippines, this would be a great time to open a new restaurant in a neighborhood of Los Angeles that does not currently have this cuisine available.

1.2 Problem

This project aims to identify an ideal location for a new Filipino chain restaurant opportunity for an aspiring entrepreneur.

1.3 Interest

An entrepreneur is looking for the opportunity to create a new Filipino chain restaurant in the Greater Los Angeles area. With the large population of Filipino-Americans as well as the growing popularity of Filipino food, and the lack of Filipino Chain Restaurants, the entrepreneur is interested in opening two restaurants in neighborhoods that do not have a huge Filipino-restaurant presence.

## 2. Data Acquisition and Cleaning

2.1 Data Sources

Los Angeles Neighborhood data was found [here](http://boundaries.latimes.com/sets/). The data was pulled from the LA Times website, from a GeoJSON. The Filipino Restaurant data was pulled by using the Foursquare API. The data was pulled for Filipino Restaurant data nearby Los Angeles.

2.2 Data Cleaning

Data that was pulled from Los Angeles Times from a GeoJSON was extracted to pull the Neighborhoods within Los Angeles county and the coordinates that make up the geospatial data. The data was transformed into a GeoDataframe. This GeoDataframe left us with 318 neighborhoods used for analysis.

The Filipino Restaurant/Venue data was pulled by connecting to the Foursquare API, and the data was transformed from a JSON to a dataframe to ease data cleansing. To prepare the data for analysis, a defined function was created to pull the category for the venue. The function created pulled the category, filters the category for each row, cleans the column names by keeping only the last term, then finally removes the categories that were not to be included in the analysis. These categories included: Church, Martial Arts, Car Wash, School, Non-Profit, Neighborhood, and any that had a null or missing category. This leaves us with 28 venues for our analysis.

## 3. Methodology

3.1 Data Analysis and Methods used for Analysis

28 Filipino venues were used to analyze the neighborhoods via clustering. Some of the categories included in the analysis were Food Trucks, Filipino Restaurants, Filipino Bakeries, Filipino Markets, and a Filipino Event Center. These were included in the analysis to not only show Filipino Restaurants, but also to show a Filipino presence in a given neighborhood or cluster to identify if the cluster or location would be ideal for a new Filipino chain-restaurant.

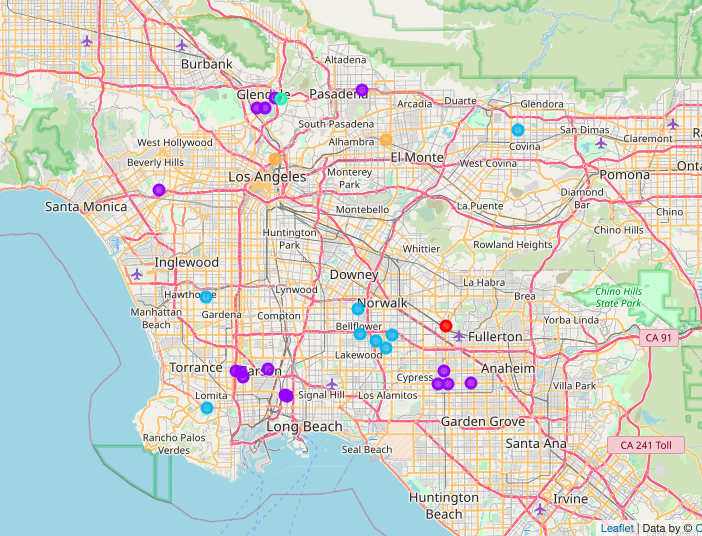
Once the data was acquired and prepared for analysis, both datasets were plotted onto a map using the python package, Folium. First, the geospatial data was plotted to show the Los Angeles data, then the Filipino Venue data was overlaid on the Los Angeles geospatial data, to show the venues around Los Angeles. Then Geopandas was used to join the geospatial data for Los Angeles data and the location of venues, to create one dataset for analysis. This would show the venues and their corresponding neighborhood based on the Los Angeles data.

After the data was joined together, analysis for each neighborhood was done using one hot encoding, to show each neighborhood and the count of each category within a neighborhood. Then the neighborhood data was grouped and the mean of the frequency of occurrence of each category was completed. There are 16 distinct neighborhoods that have the 28 venues.

For this analysis, clustering was used as the method of testing where the best locations would be for a new Filipino chain-restaurant. K-means was run to cluster the neighborhoods, using 5 as K, or the number of clusters.

3.2 Results

Cluster 0 has 1 venue for an asian restaurant. Cluster 1 has 16 venues, including several restaurants, markets and bakeries. Cluster 2 included all Filipino Restaurants. Cluster 3 included a Filipino outdoor event space. Finally, Cluster 4 included Filipino food trucks.



## 4. Discussion and Conclusion

With Clusters 0, 1 and 2 all having Filipino food, restaurant, bakery or market presence, it is not the ideal location(s) for the new filipino chain-restaurant. Cluster 3 (mint green color in above map), includes a Filipino Event Center. This means that there is a Filipino presence in this neighborhood already. With it being close to other neighborhoods that also have Filipino venues, but none within the Eagle Rock neighborhood, this is an ideal location for a new restaurant. In Cluster 4 (orange color in above map) includes two Food Trucks. This shows that there is no dine-in Filipino food presence in either of these neighborhoods. One of the food trucks is located in Elysian Park, which is located very close to downtown Los Angeles, making this also an ideal location to open a new dine-in restaurant. The other Food Truck is located in East San Gabriel, which is not too far from the El Monte airport, which shows there might be some more traffic in this neighborhood from the airport. This could also be an ideal location for a new restaurant.

In conclusion, Python was used to analyze Los Angeles geospatial data and Filipino Venue data from Foursquare’s API. After identifying the venues and the neighborhoods, k-means clustering was used to cluster the neighborhoods to identify areas that could be an ideal location for new restaurants. Clusters 3 and 4 both have ideal locations for a new dine-in chain restaurant for a new entrepreneur that is wanting to open new businesses. These locations both have Filipino venue presence, but lack the dine-in restaurant experience that would be ideal for a city like Los Angeles.