Battle of The Neighborhoods: Where in Los Angeles is the Best Location to Open a Pet Store?

Muhammad Reza Ardian

February 2021

1 Introduction

1.1 Background

With a population of 4 millions, city of Los Angeles is the second most populous city in the United States. The "City of Angels" is a city of sharp contrasts, home to people who hail from all parts of the globe and an important center of culture, business, media, and international trade. Being the world's hub of television, motion picture, and music industry, Los Angeles has a diverse economy and hosts businesses in a broad range of professional and cultural fields.

The city is divided into many different districts and neighborhoods, some of which were incorporated cities that merged with Los Angeles. These neighborhoods are well-defined enough that the city has signage marking nearly all of them.

Being a demographically and economically segregated city, it would be interesting to know in which part of LA is the best place to open a business. This work focuses on finding out which zip codes of LA are the best locations to open a business dedicated to pet such as pet store which sells pet foods and accessories or even pet clinic.

Pet foods and treat market is a fast-growing market. According to Euromonitor International, global pet food sales totaled $ 93.9 billion in 2019, an increase of 6.5 % over 2018. US pet food and treat consumers accounted for 36.5% of that by purchasing $34.3 billion of pet food and treats, of which nearly $6 billion was spent on dog and cat treats and mix-ins. US pet food sales grew in 2019 by 5.8% over 2018. Despite the pandemic, this number was projected to increase by 4% in 2020.

1.2 Problem

A good location is important to the success of a retail business. For small pet-related businesses, the advantages to opening a retail store in a busy area can include improved revenue and increased marketing exposure. In addition, the presence of similar businesses is a very important factor. Hence, deciding where to open a new business can be tricky sometimes so some important factors such as population, pet ownership, rent price, and number of existing similar businesses should be taken into account when looking for locations or zip codes to establish a new business dedicated to pet-related products.

1.3 Interest

This work would provide a recommendation for those who are planning to open a new pet store in LA, or anyone who is interested to know about the market of pet-related products in the region.

2 Data Acquisition and Cleaning

2.1 Data source

Data will be obtained from publicly available datasets, primarily those on http://laalmanac.com/ which has been a unique online reference devoted to providing statistical and narrative information about the people, places and stories of Los Angeles County. Some important datasets could be obtained from aforementioned website are:

- Pet ownership by zip codes
- Population by zip codes
- Retail rent price by zip codes
- Average income by zip codes

In addition, geography-related data such as coordinates, information about number of existing venues on each zip codes, as well as some details about the venues are obtained from Foursquare API call. When it comes to Foursquare, there are a huge number of venues in Los Angeles, much more than cities in some countries where Foursquare isn't so popular.

2.2 Data cleaning

Data provided by laalmanac.com websites are in the form of tables which are no so straightforward, for example on this link. Therefore, after fetching the data with either BeautifulSoup or pandas, it is necessary to obtain a proper dataframe by cleaning the unnecessary rows.

On the other hand, after getting venues from API call, it is necessary to filter the returned venues so that we only get pet-related venues from Foursquare, for example: pet store, pet shop, pet clinic, pet saloon.

3 Methodology

Python libraries used in this work are:

- pandas for loading the data into dataframes
- folium and geopy for map visualization
- requests and json for making API call
- seaborn for data visualization
- scikit-learn for clustering algorithm (KMeans clustering)

The first step was to load the data with read_html feature of pandas library from laalmanac.com for statistics about pet ownership and rent price as well as from opendatasoft.com for latitude and longitude.

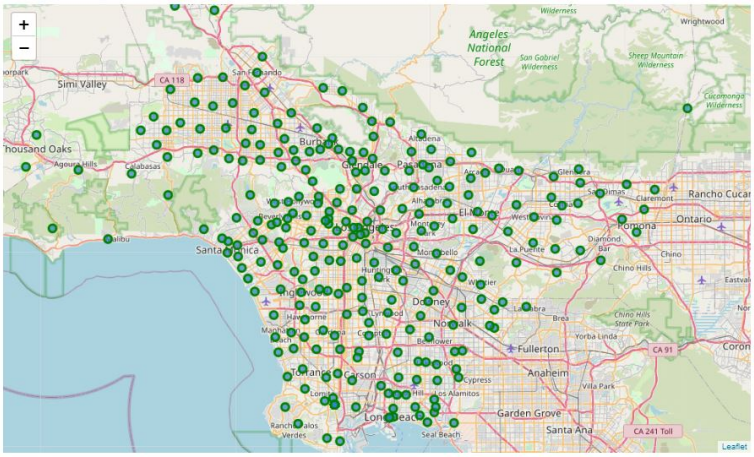
After that, API call was performed to get the venues with pet store and pet service category to find the number of pet-related establishments around a particular neighborhood. Some important parameters are radius=500 meters and limit=100.

With the data about statistics as well as the presence of existing pet-related establishments (be it pet store or pet service) machine learning algorithm with KMeans clustering can be performed to find out which zip codes represent the best neighborhood in LA to open a new pet-related business.

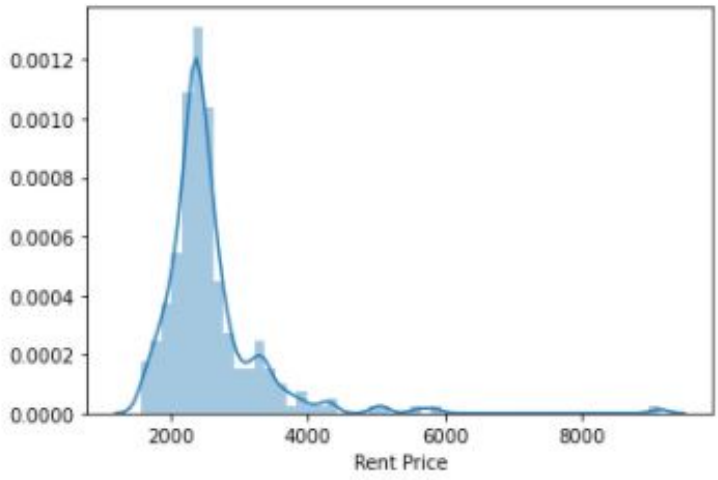
3.1 Exploratory Data Analysis

From the preliminary dataframe containing geographical coordinate and statistics, there are 267 zip codes which have complete data about the statistics (rent

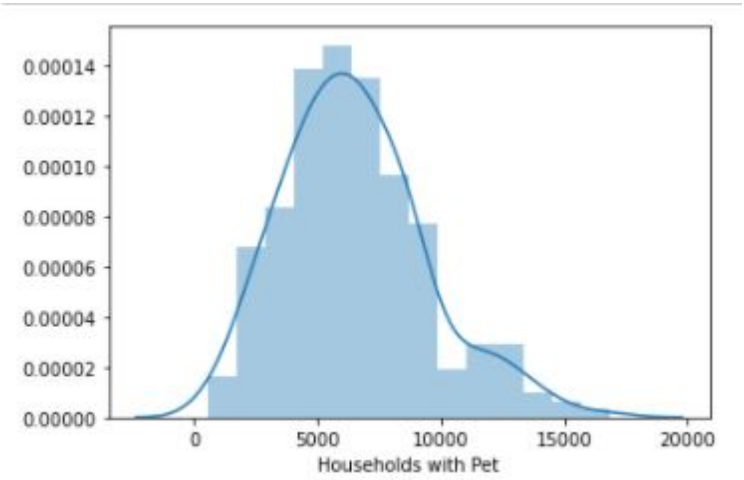
price and number of households with pets). Those 267 zip codes are distributed as follows:



The average rent price is USD 2569.7 with following distribution curve. It can be observed that there are many outliers which correspond to wealthy or business-oriented districts



Meanwhile, the shape of pet population's distribution curve is more symmetrical than rent price.

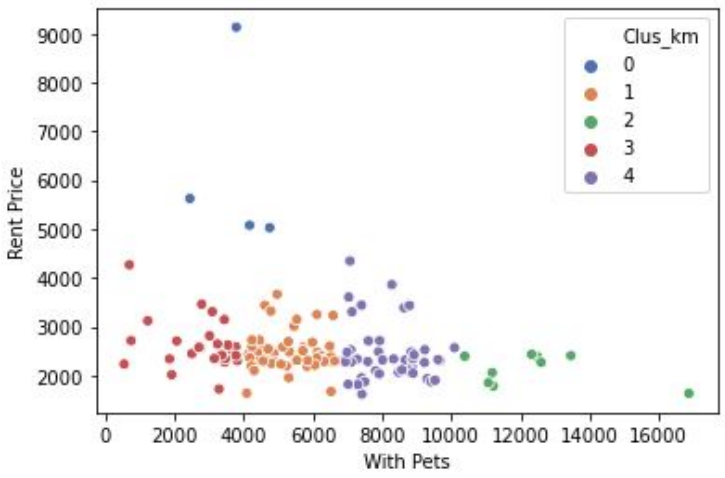


After making API call for each zip code, 175 venues with Pet Service category and 138 venues with Pet store category were obtained. Those are definitely a tiny amount considering the size of the city. Although this also depends on radius set on API call and the availability of the business within Foursquare data (we know that there are less places in foursquare than let's say googlemaps), we can consider this method as the right approach as it is the closest estimation to the number of (at least well-established) pet-related businesses.

Those 313 establishments are distributed in 130 zip codes. So, out of 267 zip codes we had, we have around 137 zip codes without any pet-related venues (at least according to Foursquare API call I made). This actually rendered the analysis much easier as we can already label the zip codes with pet venues and further cluster the rest of the dataset. Besides, the presence of similar business is an extremely important factor (if not the most) in small businesses' success. The clustering algorithm used in this work is KMeans clustering which is commonly used for aggregating data points together because of certain similarities.

4 Results

The KMean clustering was performed with k=5 on zip codes without pet venues and the parameters taken into account are only number of households with pet and rent price. The clustering and plotting resulted in below chart:

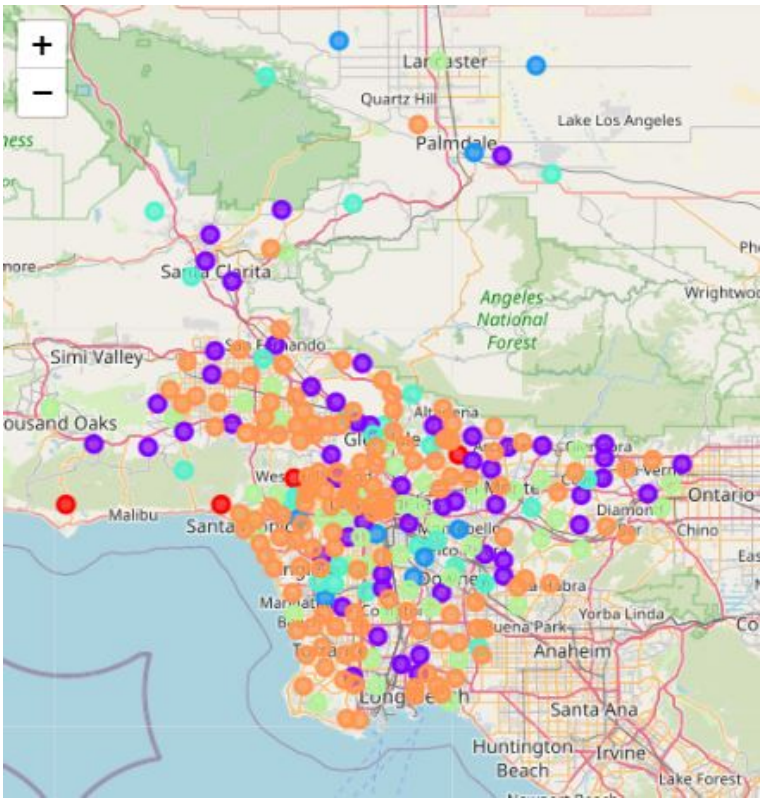


Those chart consists only the zip codes without pet-related business (which make up around half of original dataset) while the zip codes with pet-related business are labelled separately (labelled as '5', so 6 clusters in total). From the clustering chart, it can be observed that the zip codes belonging to green cluster (labelled as 2) would be the best places to open a new pet shop, as it has following favorable characteristics:

- Zero presence of venues with pet store / pet service category
- Highest number of households with pet
- Rent price is among the lowest in the city

5 Discussions

The geographic distribution of the clusters are as follow:



Different clusters here don't have a tendency to be geographically clustered together. In fact, the cluster which can be considered as the best one (cluster number 2, blue color on the map) tend to spread across the metropolitan area with some zip codes within that cluster are located quite far away from downtown LA (around Palmdale, Antelope Valley region). On the other hand, some zip codes around downtown also fall on this "best" cluster. The 9 zip codes belonging to this cluster are:

	Zip Code	Latitude	Longitude	Cities/Communities	With Pets	Rent Price	No Pet Establishment	Clus_km
18	90034	34.030560	-118.39804	Los Angeles (Palms)	13454	2411	True	2
63	90011	34.007063	-118.25868	Los Angeles (Southeast Los Angeles)	12491	2385	True	2
104	90250	33.914614	-118.35092	Hawthorne (Holly Park)	16863	1637	True	2
125	90280	33.944264	-118.19490	South Gate	12320	2437	True	2
143	93550	34.565480	-118.08549	Lake Los Angeles, Palmdale	11216	1788	True	2
150	93536	34.728857	-118.32683	Del Sur, Fairmont, Lancaster, Metler Valley, N...	11182	2062	True	2
204	90640	34.014212	-118.11358	Montebello	10388	2398	True	2
221	90201	33.972814	-118.17385	Bell, Bell Gardens, Cudahy	12598	2272	True	2
243	93535	34.691946	-117.97756	Hi Vista, Lake Los Angeles, Lancaster, Rooseve...	11072	1855	True	2

Therefore, it is recommended to open a new pet store in those zip codes. We can see an obvious winner here in this "battle of neighborhood" which is 90250 / Hawthorne(Holly Park) with the highest number of households with pet and lowest rent price.

6 Conclusions

The answer of the question "which part of LA is the best to open a new pet store?" is the 9 zip codes belonging to cluster 2, with 90250 / Hawthorne(Holly Park) as the absolute best zip code. The reason being that the cluster has pretty much a lot of things to make a new pet business successful: very few existing pet-related establishment, high pet population, and the rent prices are among the lowest in the city.