

Exploring the Cuisines of Los Angeles City

Ashish Kumar Vemula

June 13, 2020

1. Introduction and a discussion of the background

Los Angeles City (also simply known as L.A., and nicknamed the "City of Angels") is California's most populous city. The metropolitan area is the second most populated area in the U.S. and home to more than 17 million residents from all over the globe.

Los Angeles is known for its locations and its food. There are many food trucks, hotels, and restaurants that are renowned for their different delicacies. The cuisines of many cultures have taken root and it is filled with restaurants and delicacies that are made only in LA and also served with various dishes. With more restaurants opening every week, L.A. is topping the to-do lists of every food obsessive in America right now.

The aim of this project is to segment L.A's neighbourhoods categorically. In major clusters and test their cuisines. A good purpose is to look at the food preferences and taste of the neighborhood cluster. Further examination could reveal whether food has any relation with the neighborhood's diversity.

This project will help to understand the diversity of a neighborhood by leveraging venue data from Foursquare's 'Places API' and 'k-means clustering' machine learning algorithm. Exploratory Data Analysis (EDA) will help to discover further about the culture and diversity of the neighborhood.

Target Audience

- Stakeholders will be the one involved in using this quantifiable research to consider how various cultures and cuisines are spread over LA. Also, a new food vendor willing to open his or her restaurant can make use of this project. Or by a government authority to better examine and study the diversity of their city's culture.
- Business Analyst or Data Scientists who want to analyze LA's neighborhoods using Exploratory Data Analysis and other statistical & machine learning techniques to obtain all the data they need, perform some operations on it and finally be able to tell a story from it.

2. Source of the Data

- Zillow

Link - <https://www.zillow.com/research/data/>

A complete list of neighborhoods of Los Angeles was downloaded in CSV format from Zillow Housing Data.

- Foursquare API –

Link - <https://developer.foursquare.com/docs>

To get location and other information about various Food venues in LA I'm using Foursquare's explore API. Using the Foursquare's venue API with mentioned category (which gives venues recommendations), I'm fetching details about the food venues up present in LA and collected their names, categories and locations (latitude and longitude).

From Foursquare API (), I retrieved the following for each venue:

Filed	Details
Name	The name of the venue.
Category	The category type as defined by the API.
Latitude	The latitude value of the venue.
Longitude	The longitude value of the venue

3. Data Preparation & Data Cleaning

Once the neighborhood data of Los Angeles is collected from the Zillow House Value, geopy's nominatum package was used to fetch latitude and longitude details for every neighborhood. Total no.of neighborhoods - 95

Later Foursquare API was used to fetch FOOD venues for all the neighborhoods of LA within a radius of 5 miles and we found total 4750 venues.

4. EDA

Performed EDA on the cleaned data and found top food venues of each neighborhood.

5. Machine Learning Model

For clustering similar food categories, k-means clustering is used with optimized K value which was found using Silhouette method. We got 4 clusters and performed clustering on LA neighborhoods food venues.

6. Conclusion

From the results it is identified that 'Fast Food Restaurant' is the top and common food category in all the clusters. And also

- In cluster 0 'Japanese Restaurant's are the second most common venue
- In cluster 1 'Fried Chicken Joint' are the second most common venue
- In cluster 2 'Mexican Restaurant's are the second most common venue
- In cluster 3 'Seafood Restaurant's are the second most common venue