Introduction

Los Angeles, officially the City of Los Angeles and often abbreviated as L.A., is the largest city in California. It has an estimated population of nearly 4 million and is the second-largest city in the United States, after New York City, and the third-largest city in North America, after Mexico City and New York City. Los Angeles is known for its Mediterranean climate, ethnic and cultural diversity, Hollywood entertainment industry, and its sprawling metropolitan area. According to the 2010 Census, the racial makeup of Los Angeles included: 49.8% Whites, 9.6% African Americans, 0.7% Native Americans, 11.3% Asians, 0.1% Pacific Islanders, 23.8% from other races, and 4.6% from two or more races. Los Angeles is home to people from more than 140 countries speaking 224 different identified languages. Ethnic enclaves like Chinatown, Historic Filipinotown, Koreatown, Little Armenia, Little Ethiopia, Tehrangeles, Little Tokyo, Little Bangladesh, and Thai Town provide examples of the polyglot character of Los Angeles.

This project explores the most popular restaurants and its type of cuisine in different neighborhoods in central Los Angeles. Central L.A. has 26 neighborhoods and diverse racial makeup: 46.1% Latino, 26.4% White, 16.2% Asian, 8.2% Black and 3.1% Other. This results in a diverse makeup of cuisines in central L.A. and the analysis data can be used to decide what type of restaurant should you open if you plan to open a restaurant in this area. This project answers the questions “Which type of cuisine is the most popular one in each neighborhood in central L.A.?” and “What are the most similar neighborhoods in terms of types of cuisine?”

Data

To answer the above questions, I first use data from Los Angeles Times to identify different neighborhoods in central LA. LA Times also provides neighborhoods information in other areas in LA.

Geopy will be used to convert an address into latitude and longitude values so that we can further use FourSquare API.

All data related to locations and quality of different restaurants will be obtained via the FourSquare API utilized via the Request library in Python.

I then perform a K-means clustering on these neighborhoods and pick the cluster the most similar neighborhoods together.