# **NEIGHBORHOOD RESTAURANT CUSINE ANALYSIS OF LOS ANGELES**

**1. Introduction**

Los Angeles (LA) is the largest city in California. LA has a population of approximately four (4) million people and ranks third in having the largest population in North America. There are many characteristics that define LA including the city having ethnic diversity, a large metropolis and a wide-ranging economy. As such it has one of the largest Gross Domestic Product (GDP) in the world that has reached 1 trillion USD.

Development of any business within LA will require an in-depth analysis of several factors. This report however looks specifically into the restaurant business and how data analysis can point to several key indicators in determining specifics of the business which can suggest a possible location and a recommended type of cuisine to maximize profit. The following are the queries that the report will answer to determine these:

1) Determining neighborhood clusters according to the most popular type of restaurants cuisine e.g. Italian, Chinese.

2) Determining current rating of the different types of restaurant cuisine

3) Determining current trending types of restaurants cuisine

4) Analyze the tips for a restaurant cuisine to determine popularity

These queries will have the effect of showing the relationship between a location and a preferred cuisine as LA is ethnically diverse as mentioned before. In addition, even changes to colloquial preferences as per current trends in demand for certain foods can be shown by the queries.

Business personnel making the decisions can get an idea how the accumulated data points to via a simplified map showing the clusters. The use of graphs will also give insight into which location and cuisine can be most profitable based on indicators of restaurant ratings, trending and tips analysis for restaurants. These will then be compiled into a final table that can surmise the best location and cuisine a business owner starting a restaurant in LA can choose.

A businessperson’s interest in knowing this data will be due to him/her making an informed decision such that they gain the most amount of profit from their initial investment. By knowing the analysis of the existing data, they reduce the risk of a loss since location and cuisine of a restaurant are key decision factors that can determine how popular a restaurant is and the amount of revenue that can be drawn.

**2. Data Description**

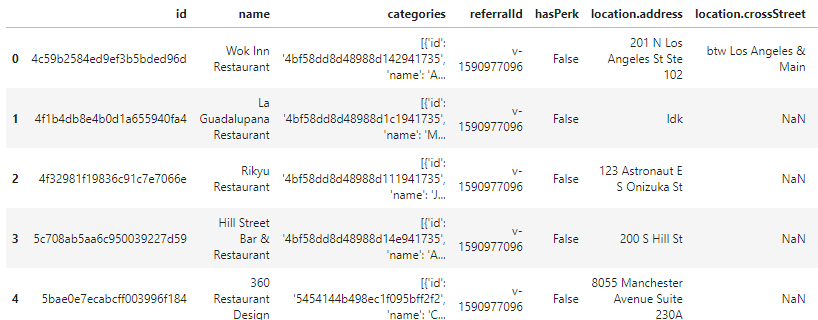
**2.1 Description of dataframe with Restaurant Name, Co-ordinates and Cuisines**

The data will be obtained from Foursquare Places API for Los Angeles, California. The co-ordinates for the latitude and longitude was obtained for LA, CA, a URL was created using a search query on ‘Restaurant’ which is used with a 1000 m radius in order to get the JSON file a snippet of which is shown in Figure 1.



**Figure 1: The figure shows a snippet of the JSON generated from the ‘Restaurant’ query**

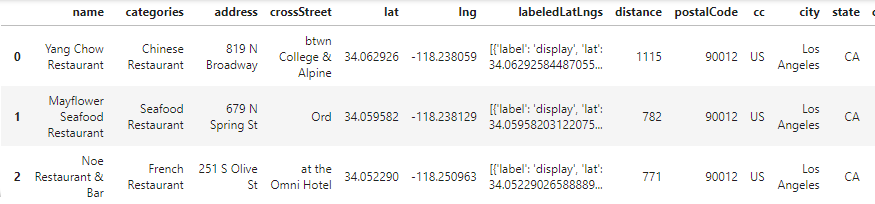
This will then be converted into a dataframe which is shown in Figure 2. As seen in this dataframe the categories are described via ID characters.



**Figure 2: The figure shows a snippet of the JSON data converted into a dataframe**

The important components of this dataframe from Figure 2 is the generated restaurant’s name and co-ordinates i.e. latitude and longitude that can be used after. Filtering will then be used in order to get the categories i.e. the cuisines required as mentioned above for the data analysis. This is show in Figure 3 where each restaurant’s cuisine is described in the dataframe under the ‘categories’ column. The data that shows cuisines can then be grouped to show the actual numbers of each for the respective neighborhoods this can give an idea of popularity of the cuisines and will be illustrated via graphs.

Figure 4 shows some of the restaurants that will be used in the analysis. This data will be used for clustering in the respective neighborhoods for the restaurants based on cuisine.



**Figure 3: The figure shows the filtered dataframe showing the cuisines for each restaurant**

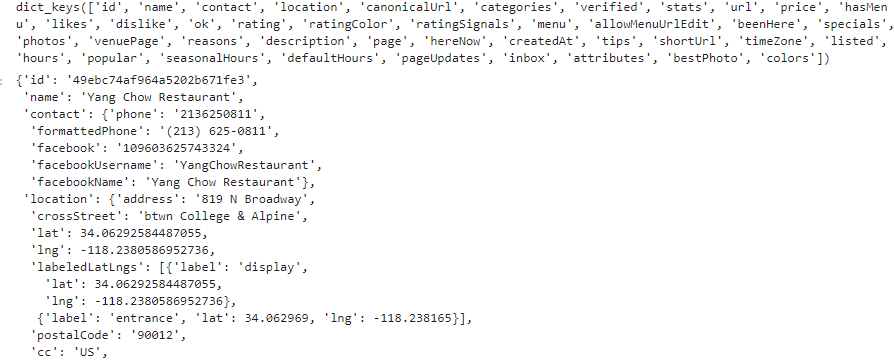


**Figure 4: The figure shows some of the restaurants whose data will be analyzed in the report**

**2.2 Description of JSON data to be used for ratings, trending and tips analysis**

The next JSON was generated to get specific venue information including the ratings, trending and tips for the venue, this is shown in Figure 5. This data will be used to determine how popular a restaurant is in each location based on cuisine. This data will be used to generate graphs for analysis.

This data will be placed into a dataframe and cleaned. The issues with this data are some venues are missing for example rating’s data and thus for analyses purposes will be removed from the dataset, this holds true for trending and tips analysis.



**Figure 5: This figure shows a snippet of the JSON that will be used to get specifics about the venue including ratings, trending and tips**