IBM Data Science Capstone Project

Opening a new fast-food restaurant in Los Angeles, USA

Introduction

"Opening a new fast-food restaurant in Los Angeles"

- Los Angeles is very density populated city in west coast of USA. It is one of the major cultural hubs in the world with immigrants from many countries live in harmony. It is well known for Hollywood movie industry, night life and restaurants. Los Angeles has a very big market for restaurants and fast-food restaurants are very famous due to the tourism. Due to this, there are thousands of restaurants in the city and there is huge competition in the business. It is a very challenging task for any restaurant chain to plan their business expansion.
- In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening a fast-food restaurant in Los Angeles, USA.

Objective

- The objective is to use data science methodology and machine learning algorithms to answer the following critical questions in order recommend potential location for a new fast-food restaurant in Los Angeles
  - Which areas have potential Fast Food Restaurants Market?
  - Which areas are lacking Fast Food Restaurants?
  - What are the suitable neighborhoods in Los Angeles for opening a new Fast-Food Restaurant?

Target Audience

- This is very useful for some of the major fast food restaurant chains such as Burger King, Subway, McDonald's etc… It will help them in preparing a strategy for their business expansion also understand the areas with potential market for fast food restaurants.
- It will provide much more value add with further analysis on other influencing factors, measures and key external indicators. For example: impact of other similar restaurants, user ratings, market share, per capita income etc…

Data Sources

Data Requirements

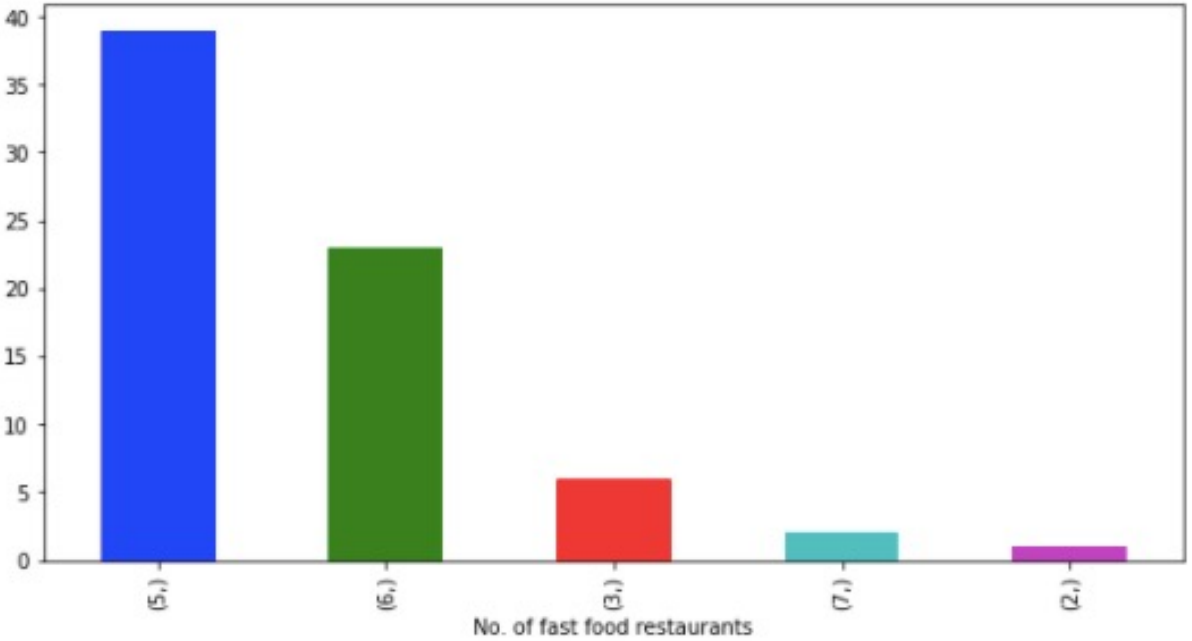
- We first need list of all neighborhoods in Los Angeles
- Population density of each neighborhood
- Latitude and longitude coordinates of each neighborhood
- Venue data, specifically, list of restaurants around each neighborhood

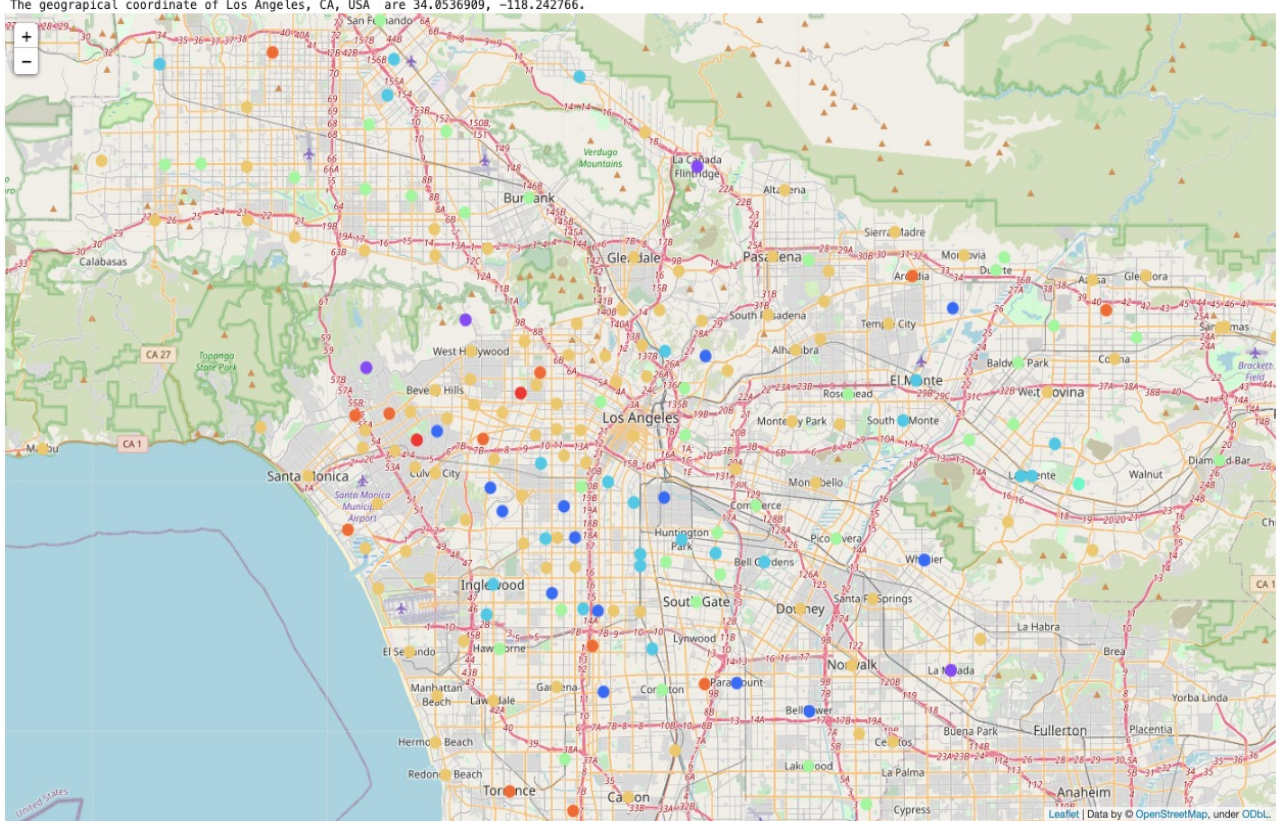
Data collection

- LA Times for neighborhood details and population density
  - ✓http://maps.latimes.com/neighborhoods/neighborhood/list/
  - ✓http://maps.latimes.com/neighborhoods/population/density/neighborhood/list/
- Geocoder library in python for latitude and longitude coordinates of neighborhoods
- Foursquare API for restaurants around each neighborhood

Clustering

- We created 8 clusters of the neighborhoods
- You can see the clusters with color coding in the map
- The clusters are distributed based on the restaurant categories

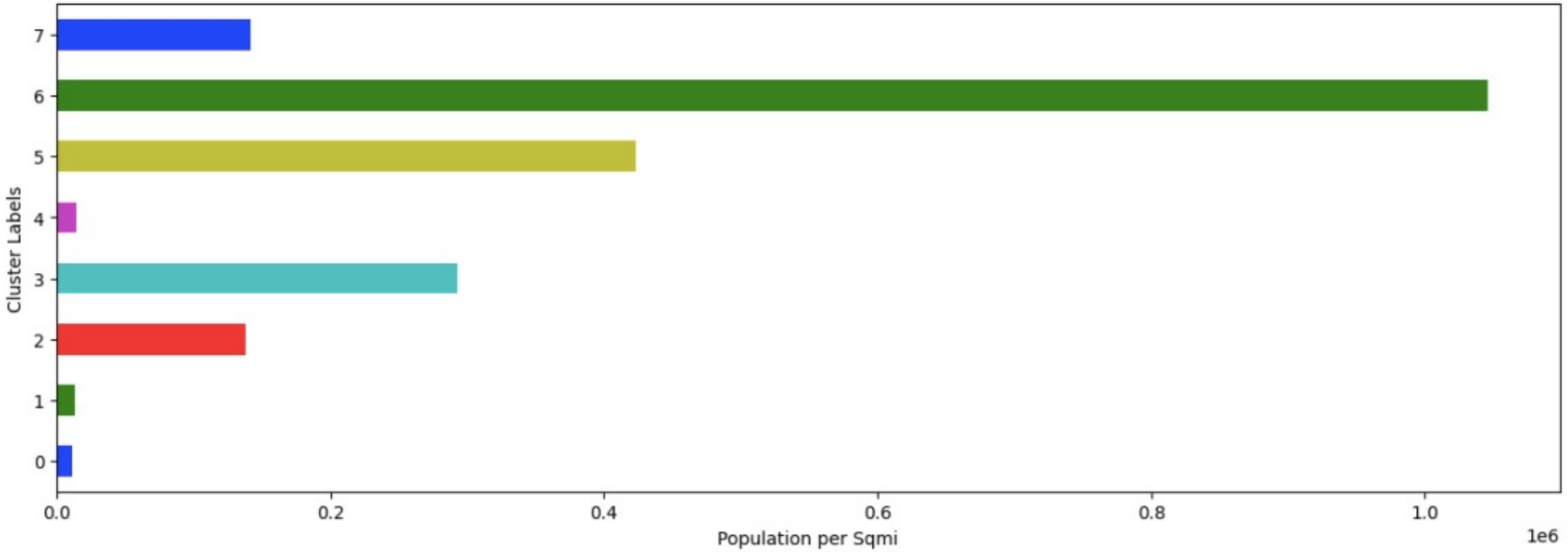




- Cluster 5 has highest number of fast-food restaurants
- Cluster 2 has lowest
- We can also notice that clusters 0, 1 and 4 do not have any fast-food restaurants

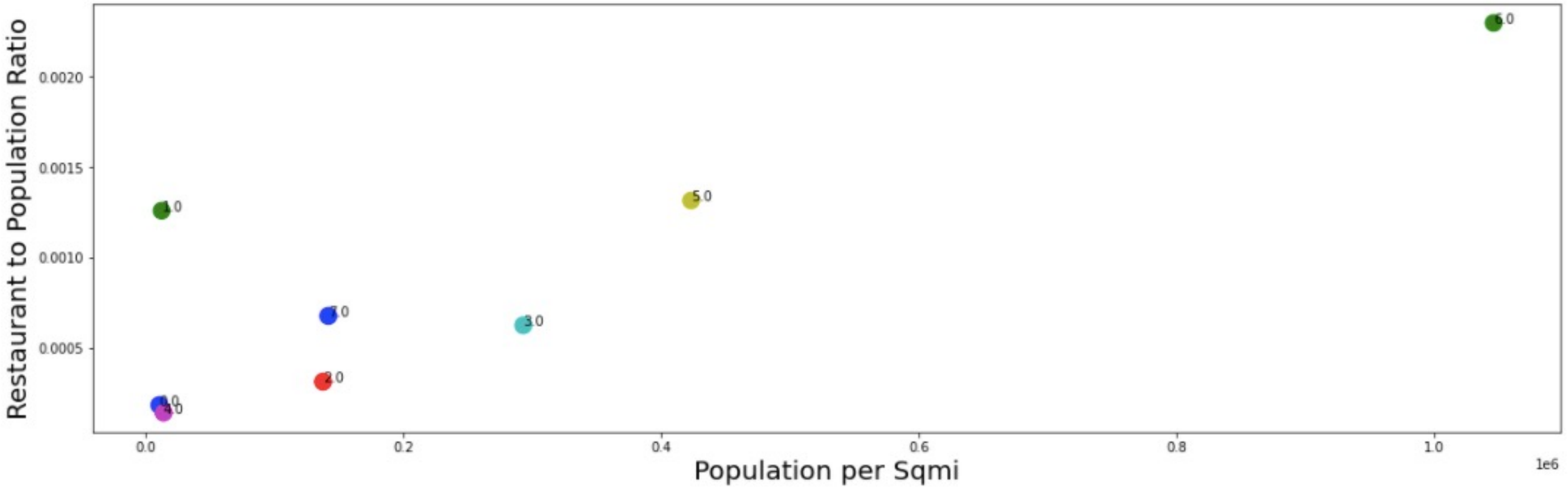
Population Density in each cluster

- Population density is very less in Clusters 0, 1 and 4 so these cannot be suitable places to open restaurants.
- Hence, we need to further review the clusters 2, 3, 6, 5, & 7



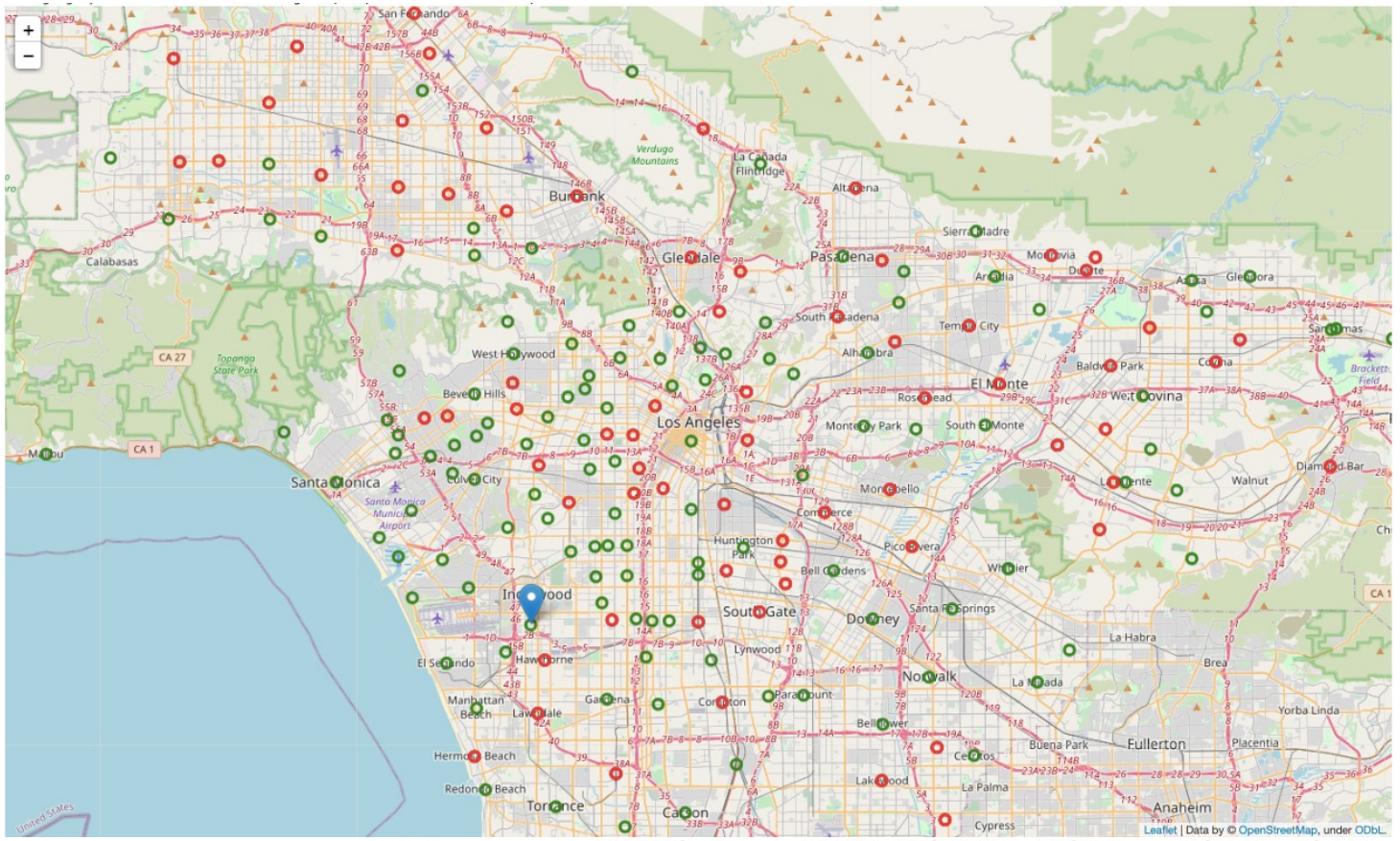
Restaurant count vs. Population Density

- Clusters 2 and 3 have very less number restaurants given the population density.
- Where as, cluster 5 and 6 already have good number of restaurants.



Most suitable neighborhood

- With in Clusters 2 and 3, 'Lennox' neighborhood has least number of restaurants, given the population density.
- In the map
  - Red circles has fast food restaurants in the neighborhoods
  - where as, Green circles do not have any fast-food restaurants but has other categories of restaurants
- The selected neighborhood i.e. 'Lennox' is highlighted, and you can see there are limited restaurants in the surrounding neighborhood and no fast-food restaurants near by.
- ‘Lennox’ has 9 restaurants in total however zero fast food restaurants.
- Hence, it is the most suitable place to open a new fast-food restaurant.



	Region	Neighborhood	Latitude	Longitude	Population per Sqmi	Restaurant Count	Fast Food Restaurant Count	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
112	South Bay	Lennox	33.939031	-118.357443	21557	9	0.0	3	Mexican Restaurant	Taco Place	Pizza Place	Seafood Restaurant	Food Truck
92	Southeast	Huntington Park	33.982704	-118.212034	20223	4	0.0	3	Fried Chicken Joint	Mexican Restaurant	Burger Joint	Food	African Restaurant
89	South L.A.	Historic South-Central	34.016230	-118.267308	19474	9	1.0	3	Mexican Restaurant	Restaurant	Taco Place	Fast Food Restaurant	Donut Shop
33	South L.A.	Central-Alameda	34.004015	-118.247784	18760	4	0.0	3	Donut Shop	Taco Place	Food	Mexican Restaurant	African Restaurant
198	South L.A.	Vermont-Slauson	33.983691	-118.291542	18577	5	0.0	2	Food	Sandwich Place	Burger Joint	African Restaurant	Mongolian Restaurant

Conclusion and Future Recommendations

- We were able to identify the suitable neighborhood for opening a new fast-food restaurant by using k-Mean clustering algorithm and exploratory analysis on the data.
  - As you can see in the final map, combining exploratory analysis with clustered data provided better insight and helped with decision making
  - Cluster 3 and cluster 2 are the viable options to open a fast food restaurant business, respectively
  - With 9 restaurants and zero fast food restaurants, ‘Lennox’ is the most suitable neighborhood to open a new restaurant
- We can improve our modeling by including more influencing factors, measures, and key external indicators and identifying the correlation
  - For example:
    - Impact of other similar restaurants
    - User ratings
    - Market share
    - Demographic data
    - Per capita income etc...

Thank You