**Where to open an Indian Restaurant & Bar in USA**

Ekaansh Khosla

September 6, 2019

**1. Introduction**

As we many Indians migrate from India to USA, whether it is for higher studies, or getting a job so it is very nice idea to open up things for them so they hang out with their Indian as well as American friends.

As we know that an average person likes to eat in his or her cultural, like Italians like to eat in Italian style, Americans like to eat in American style, Chinese people like to eat in Chinese style so it is very important that something must be done so that Indian people could enjoy their culture outside of India so if it is possible that an Indian Restaurant could be open in USA then it will be great for the business men and as well as Indians.

As there are lots of Indians in California region, I researched that where in California an Indian restaurant can be set up. After doing some research I found that in Los Angles it will be very profitable to setup an Indian restaurant. So that university students as well as working could enjoy their weekends or so. By doing this we could also introduce Indian culture to Americans as well as people other than Indians and Americans who have migrated from their countries.

**2. Data**

I collected data for various places in California as well as Los Angeles from Kaggle [1] [2].

California data was in the following format Table 1:

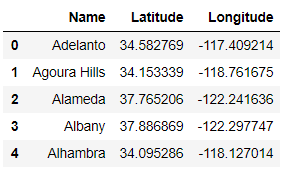


Table 1. Data format of California

So I got a readymade data which I could use for my analysis. After analyzing this data using foursquare API, I got to know that Los Angles would be the most suitable place to open up an Indian Restaurant.

So we got Los Angeles data from Kaggle in the following format Table 2:

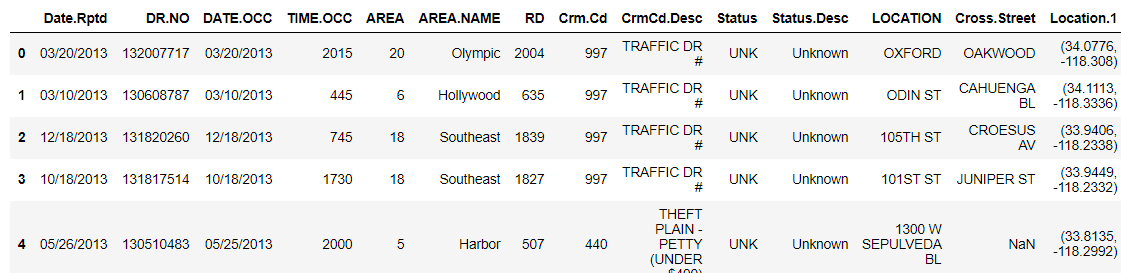


Table 2. Data format of Los Angles

In this only two column were useful that are AREA\_NAME and Location.1. so after deleting all the other columns and converting Location.1 into latitude and longitude values in type float we get the following data Table 3:

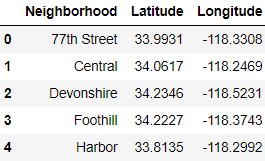


Table 3. Data format of Los Angles

Now the venues near these places can be explored using foursquare API.

**3. Methodology**

***Analyzing California***

First I explored places near California using the following code in which I used foursquare API Figure 1:

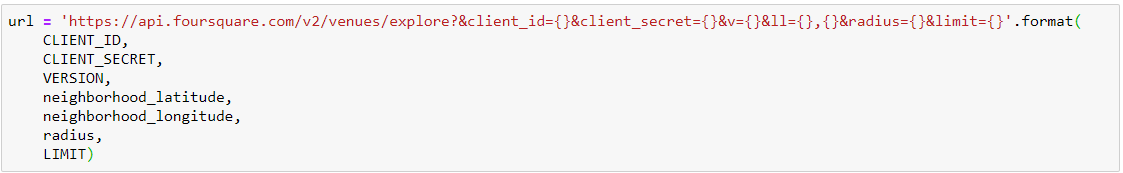


Figure 1. Code for exploring different places in California.

I took radius of 5000 meters and limit of 100. then I sorted every places according to its most common venues. By analyzing this I decided to explore Los Angles more.

***Analyzing Los Angeles***

First I explored places near Los Angeles using the following code Figure 2 in which I used foursquare API:

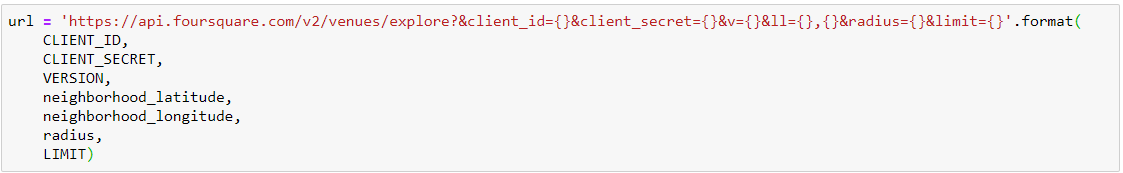


Figure 2. Code for exploring different places in Los Angeles.

I took radius of 5000 meters and limit of 100. Then I sorted every places according to its most common venues.

Then used K-means clustering in order to determine the major clusters corresponding to the data and using the clusters to determine where should we setup Indian Restaurant & Bar in Los Angeles.

**4. Results**

***Exploring California***

First I plotted the Map of California in which blue marks tell us the places of which the data is given, Plotted map is shown in figure 3.

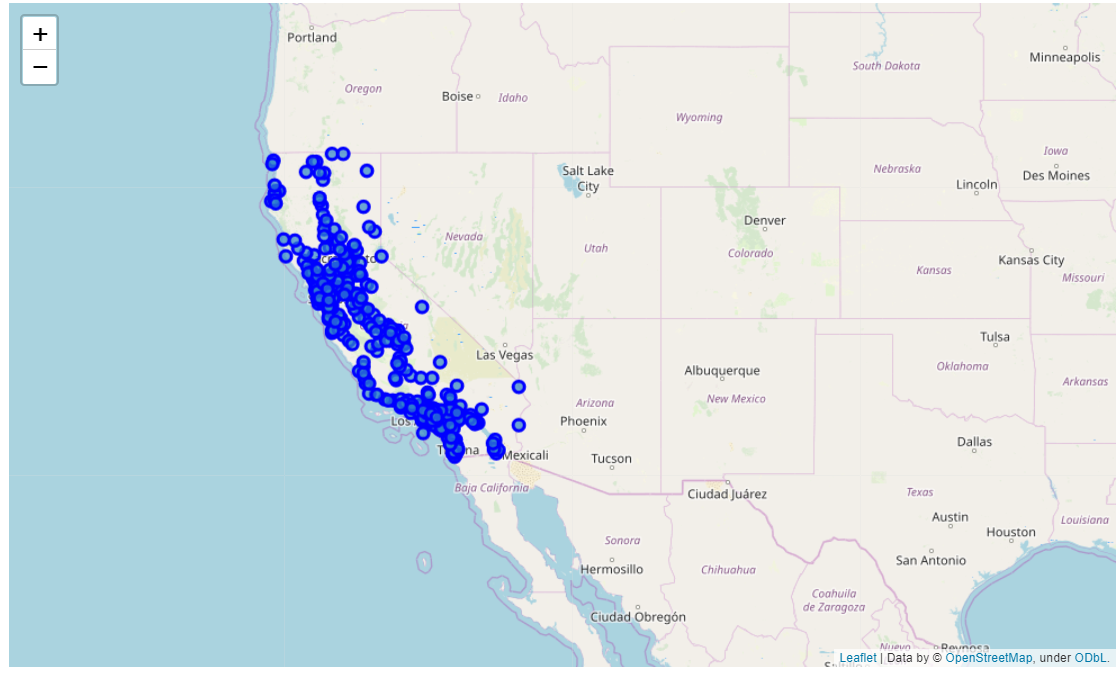


Figure 3. Places in California of which data was given on Kaggle data set.

Then I examine the name of venues by its category, there were 411 different categories. The result is shown as Table 4:



Table 4. Name of venues by its category

Then I sorted every venue category according to its most frequency at a place, results are shown below Figure 4:

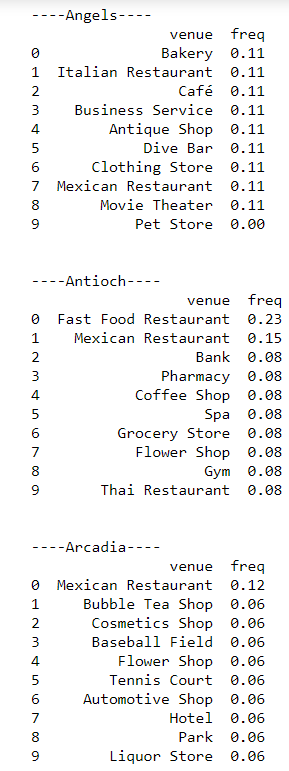


Figure 4. Sorted every venue category according to its frequency at a place

So exploring this I decided to open a restaurant and Bar in Los Angeles as there is no Indian Restaurant and Bar in its most frequent places. The most frequent venues data of Los Angeles is as follows Figure 5:



Figure 5. Most frequent places in Los Angeles

***Exploring Los Angeles***

First I plotted the Map of Los Angeles in which blue marks tell us the places of which the data is given, Plotted map is shown in figure 6.

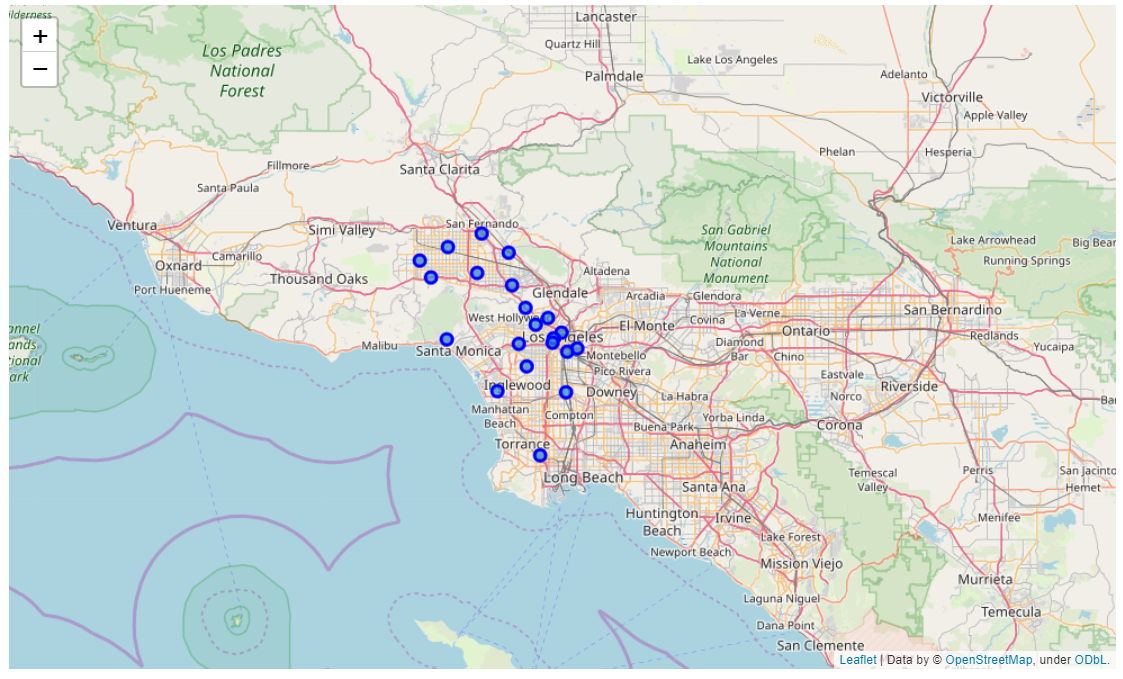


Figure 6. Places in Los Angeles of which data was given on Kaggle data set.

Then I examine the name of venues by its category, there were 177 different categories. The result is shown as Table 5:

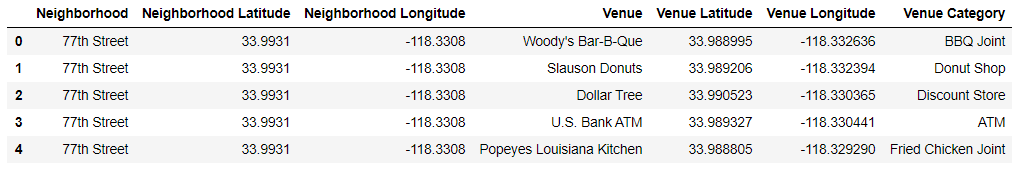


Table 4. Name of venues by its category

Then I hot encoded every category which are our features in order to fit the K-means Algorithm, the result is shown in Table 5.

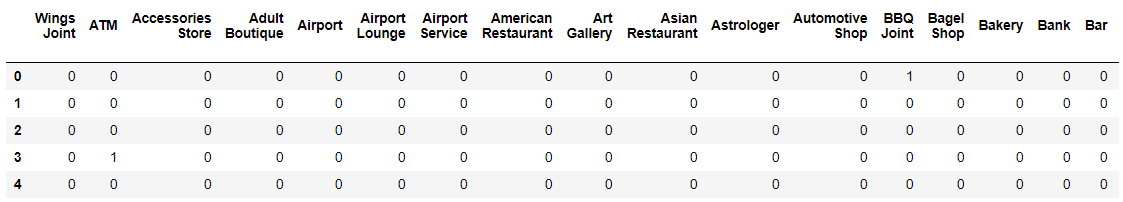


Table 5. Hot encoded category

Then I took the mean by using group by function and the sorted them according to their frequency and got the following results Table 6.

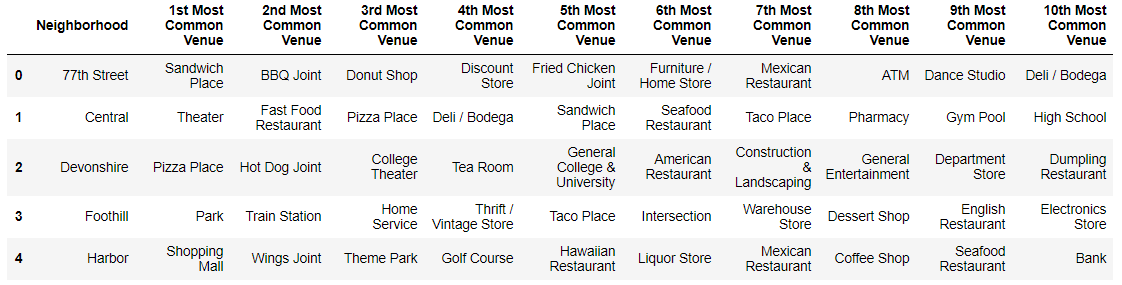


Table 6. Sorted every venue category according to its frequency at a place

Then I ran K-Means clustering on the hot encoded data after grouping them. The clusters I got is as follows Table 7:

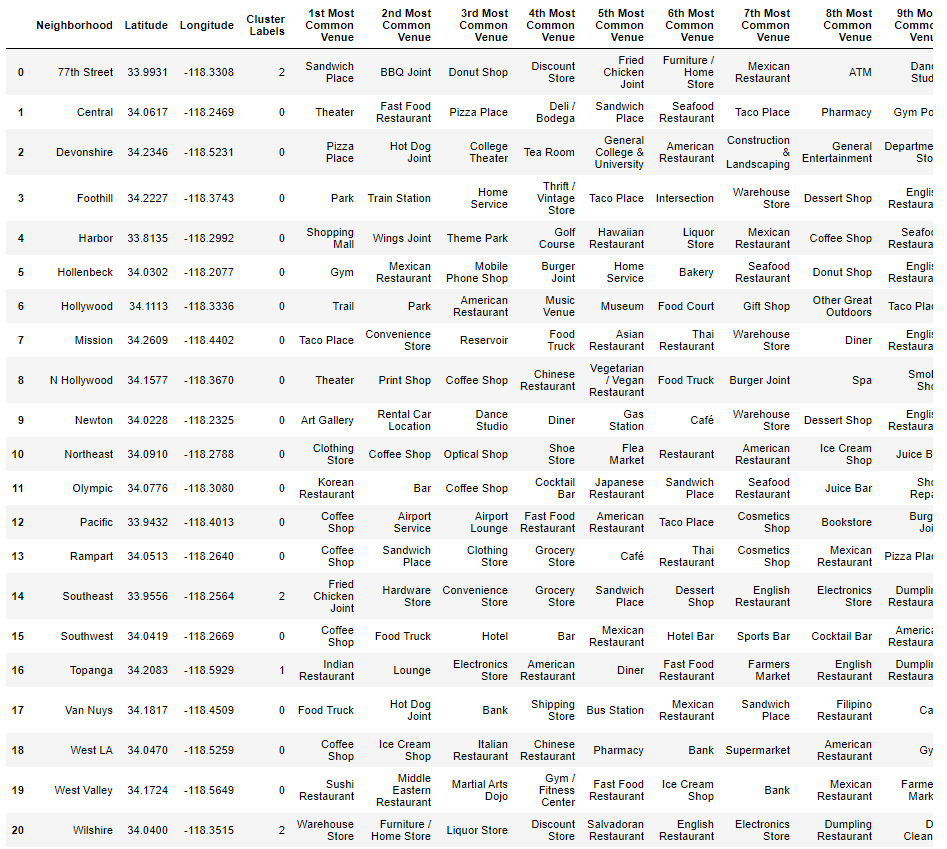


Table 7. Resultant clusters after running K-Means

From the above clustering we got to know that cluster with label=0 has mostly food items, restaurants near them and with cluster=1 we have Topanga which have the following top 10 venues:

**Topanga**

1. Lounge 0.25
2. American Restaurant 0.25
3. Electronics Store 0.25
4. Indian Restaurant 0.25
5. Wings Joint 0.00
6. New American Restaurant 0.00
7. Motel 0.00
8. Movie Theatre 0.00
9. Moving Target 0.00
10. Museum 0.00

Also, with cluster=2 we have 77th Street, Southeast and Wilshire and which have the following top 10 venues respectively:

**77th Street**

1. Mexican Restaurant 0.12
2. Sandwich Place 0.12
3. Fried Chicken Joint 0.12
4. Furniture / Home Store 0.12
5. Donut Shop 0.12
6. Discount Store 0.12
7. BBQ Joint 0.12
8. ATM 0.12
9. Other Great Outdoors 0.00
10. Optical Shop 0.00

**Southeast**

1. Fried Chicken Joint 0.33
2. Convenience Store 0.17
3. Hardware Store 0.17
4. Grocery Store 0.17
5. Sandwich Place 0.17
6. Nail Salon 0.00
7. Monument / Landmark 0.00
8. Motel 0.00
9. Movie Theatre 0.00
10. Moving Target 0.00

**Wilshire**

1. Warehouse Store 0.2
2. Discount Store 0.2
3. Liquor Store 0.2
4. Salvadoran Restaurant 0.2
5. Furniture / Home Store 0.2
6. Motel 0.0
7. Movie Theatre 0.0
8. Moving Target 0.0
9. Museum 0.0
10. Music Store 0.0

The above clustering is shown in graphical format in Figure 7.

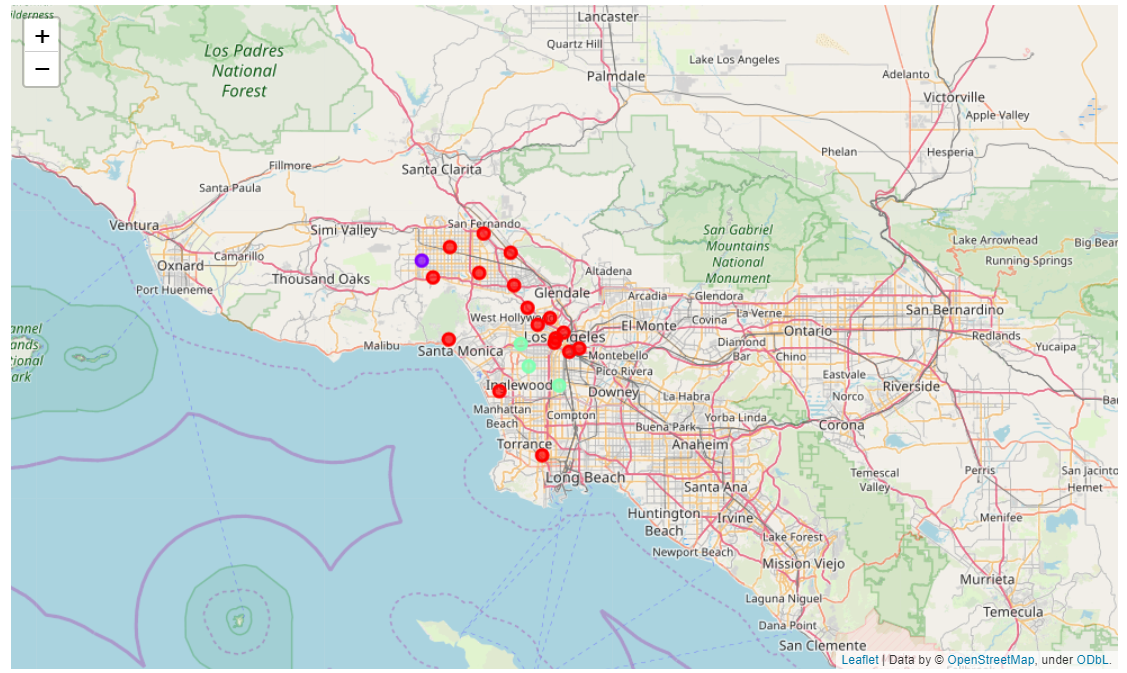


Figure 7. Clustering of venues in Los Angles.

Hence from the above data we got to know that there are not much Places to have food near Wilshire so we can setup an Indian restaurants there and as Liquor store is most common we can open a Restaurant & Bar over there where Indian people as well as American can have their drinks with Chicken tandoori, butter Chicken and all the awesome Indian food.

The following is the image taken from google maps and it also states that there are not much restaurants and bar near Wilshire Figure 8, Figure 9.

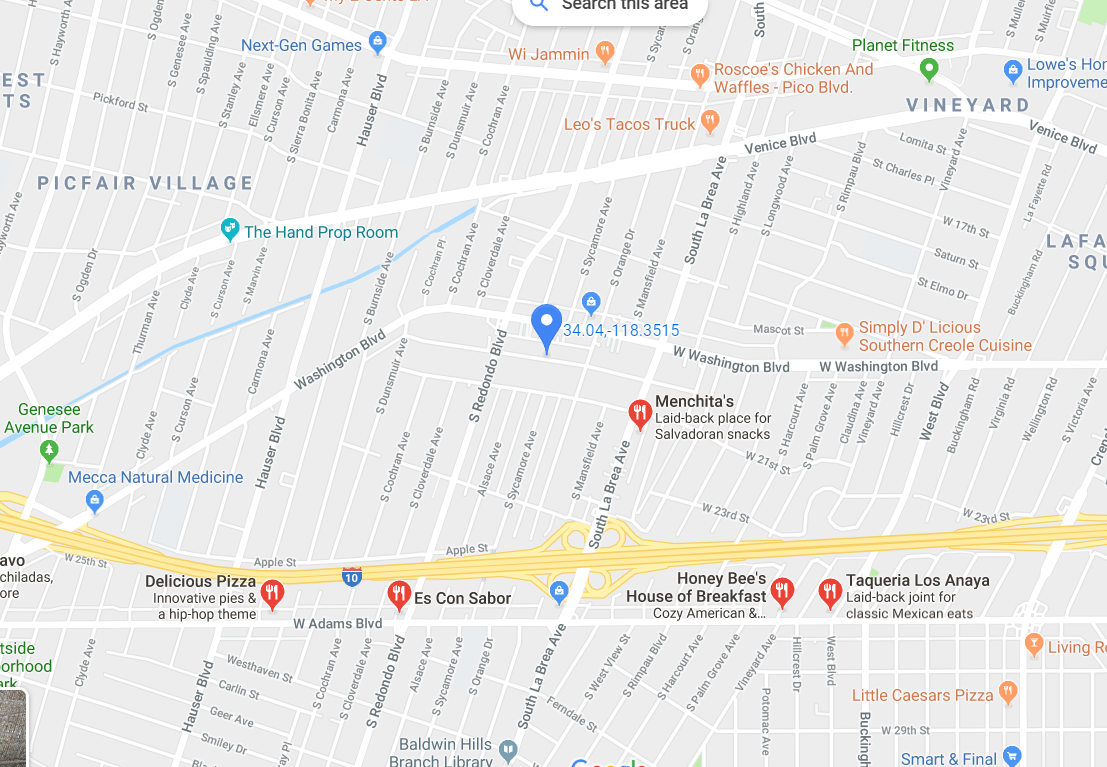


Figure 8. Google Maps, Restaurants in Wilshire

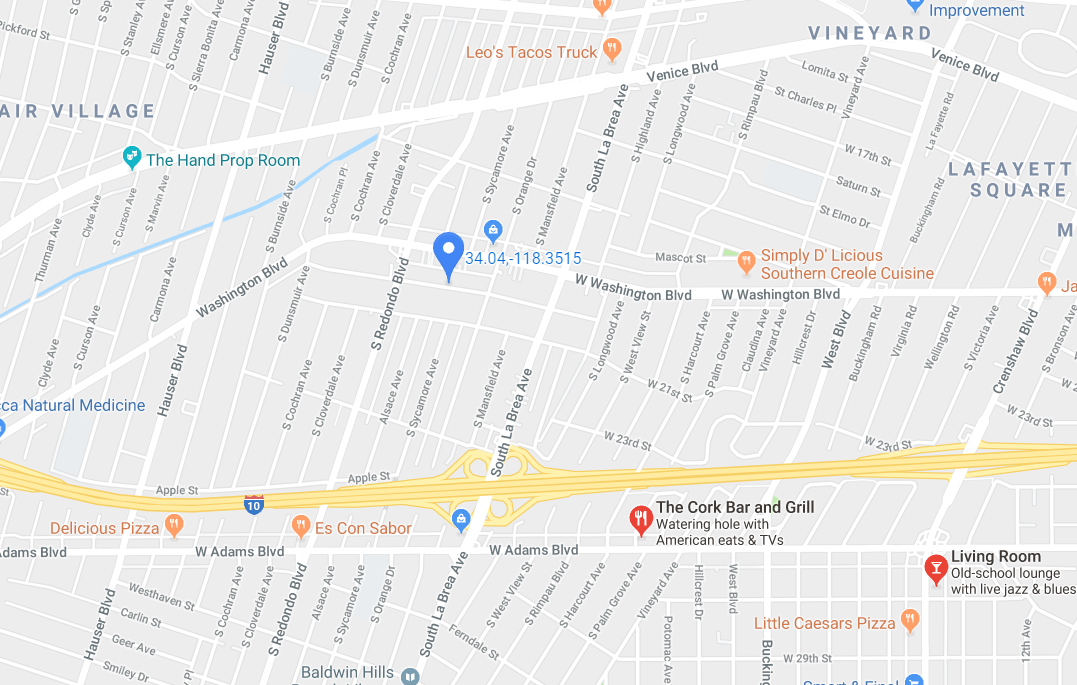


Figure 9. Google Maps, Bars in Wilshire

Hence by examine the google images we can set up and Indian Restaurants and Bar in the following area figure 10.

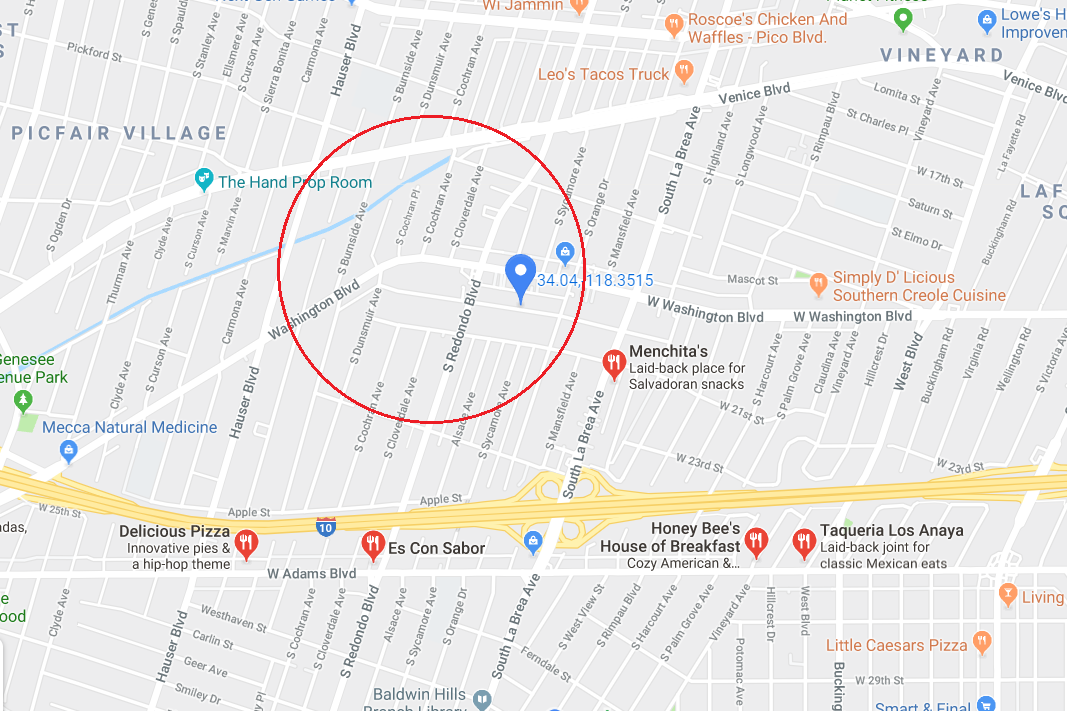


Figure 10. Place where and Indian Restaurant & Bar can be opened.

**5. Discussions**

From the study we got to know that foursquare API is a great tool to explore new places around the world. In the study I observed that we can apply many machine learning models in order to find a suitable spot for many activities for the people over that place.

I have little suggestion that is the foursquare API would have been more accurate then we could have performed the above study more accurately.

**6. Conclusions**

We successfully suggested the place where we can set up an Indian Restaurant and Bar in Los Angeles by using foursquare API and K-Means clustering.

But if the foursquare API would have been more accurate then we could have performed the above study more accurately, and we could have suggested business people about the place more accurately so that it would have been profitable for both business people as well as people living in that region.

**7. References**

1. <https://www.kaggle.com/camnugent/california-housing-feature-engineering>
2. <https://www.kaggle.com/kingburrito666/los-angeles-crime>