

# The Battle of Neighborhoods

## Introduction

The University of Southern California with its project named The Neighborhood Data for Social Change (NDSC) platform, provides different resources they said it is head to civic actors who want to learn about their neighborhood. The platform offers different areas of study such as demography, education, environment, health, among others. Taking advantage of all these data sets, especially of the demographic dataset, which provides information about the population distribution by age distribution, households, race and ethnicity, age distribution among others, it would be nice to understand the population distribution in Los Angeles by race and ethnicity to analyse where the different races are located and how the neighborhoods are clustered for the largest groups of people in Los Angeles. To do that, the idea is to find the two largest groups in Los Angeles by using the NDSC (White, Black, hispanic, Asian...) classification, clustering them by utilizing the Foursquare API, and comparing them to understand similarities and differences among races.

The importance of this analysis is that persons who want to invest or live in one of those neighborhoods can understand the influence of races in their neighborhoods, and they can make decisions based on this analysis.

## Table of contents

1. [Data Sources](#)
  - 1.1 [Understanding the NDSC dataset Race & Ethnicity](#)
  - 1.2 [Understanding the Foursquare API](#)
2. [Methodology](#)
3. [Analysis](#)
4. [Results and Discussion](#)
5. [Conclusion](#)

## 1- Data Sources

### 1.1 Understanding the NDSC dataset = Race & Ethnicity

Race & Ethnicity has 7 years records from 2010 to 2016, 8 race and ethnicity classification, and 262 neighborhoods which provides 2344 location points with the latitude and longitude coordinates of each location point.

Here is the link to the dataset: <https://usc.data.socrata.com/api/views/jxw5-xxv5/rows.csv>  
For this analysis, we will use only records from 2016, so let's prepare the data.

	Variable	Count	Neighborhood	latitude	longitude
138	Black Population	66.0	Mar Vista	34.016630	-118.437563
198	Black Population	19.0	Burbank	34.171425	-118.352775
317	Black Population	136.0	Burbank	34.173525	-118.342414
377	Black Population	51.0	Burbank	34.162038	-118.349580
455	Black Population	161.0	Burbank	34.164754	-118.338370

## 1.2 Understanding the Foursquare API

The Foursquare API is a source for developers who want to search for a specific type of venue around a given location as well as a specific venue with data such as the full address, working hours, menu among others. Furthermore, it can possible to explore popular spots in a given location and trending venues as well.

Here is the API link  
is: [https://api.foursquare.com/v2/venues/search?client\\_id=CLIENT\\_ID&client\\_secret=CLIENT\\_SECRET&ll=LATITUDE, LONGITUDE&v=VERSION&query=QUERY&radius=RADIUS&limit=LIMIT](https://api.foursquare.com/v2/venues/search?client_id=CLIENT_ID&client_secret=CLIENT_SECRET&ll=LATITUDE, LONGITUDE&v=VERSION&query=QUERY&radius=RADIUS&limit=LIMIT)

where:

1. CLIENT\_ID = your Foursquare ID
2. CLIENT\_SECRET = your Foursquare Secret
3. LATITUDE = latitude of the given place
4. LONGITUDE = longitude of the given place
5. VERSION = date of the database version 'yyyymmdd'
6. QUERY = the value you want to search
7. RADIUS = number of meter from the given place
8. LIMIT = number of records in the answer

Important information: With your free Foursquare account you can access:

1. 105M places
2. 2 Photos & 2 Tips per Venue
3. 2 Queries per Second (QPS)
4. 1 App per Account
5. Insight into API Usage
6. API Call Quota
7. 99,500 Regular Calls + 500 Premium Calls

Let's see how we can call for venues for the first five location points from the Race & Ethnicity dataset.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Mar Vista	34.016630	-118.437563	Ocean View Farms	34.013712	-118.440426	Garden
1	Mar Vista	34.016630	-118.437563	north venice little league	34.014600	-118.440083	Baseball Field
2	Mar Vista	34.016630	-118.437563	Mountain View Outlook	34.017705	-118.441549	Scenic Lookout
3	Mar Vista	34.016630	-118.437563	VENICE GARDENS	34.015355	-118.442468	Garden
4	Burbank	34.171425	-118.352775	Emerald Knights Comics and Games	34.172974	-118.354616	Toy / Game Store

## 2- Methodology

In this project, we will focus on the identification of similarities and differences among the two largest race and ethnicity populations in Los Angeles, especially in the neighborhoods in which each population group (2) has the most extensive participation.

To do that, we will first analyze each population group (White, Hispanic...) by using graphical descriptions, grouping data and cleaning data.

Second, we will select five neighborhoods for the two largest group. There are two approaches to do that. On one hand, we can select the neighborhoods in which the groups have the greatest number of citizens. On the other hand, we can select the neighborhoods in which the groups have the highest participation rate.

Third, We will cluster the location points given by the population dataset using **K-means clustering** the nearby venues in a **radius of 500 meters**, in order to understand how those neighborhoods are clustered and how each population group can influence those neighborhoods.

Finally, we will discuss the findings and make conclusions around them.

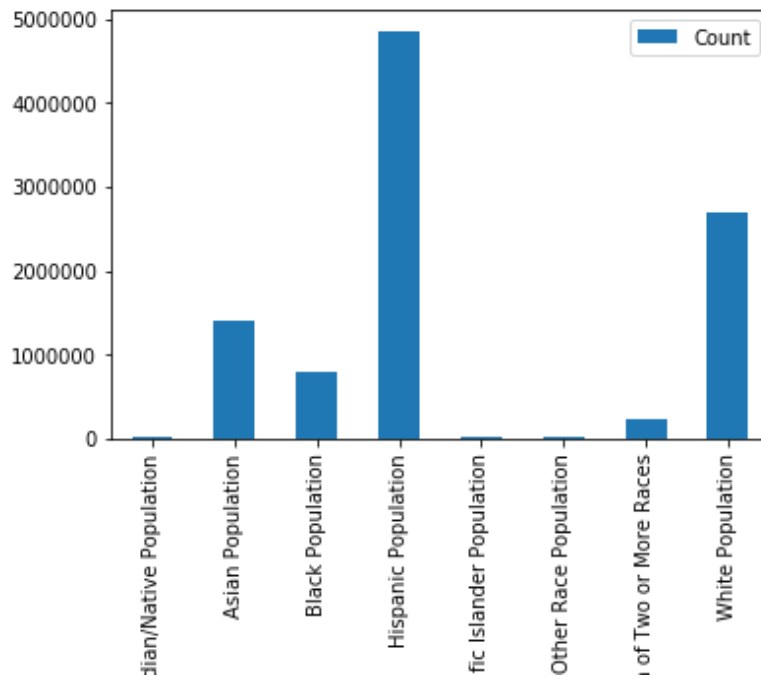
## 3- Analysis

Identify which population groups are the two largest.

```
In [12]: #Creating Dataframe
df2016 = pd.DataFrame(data=dataframe2016, columns=['Variable', 'Count'])
df2016 = df2016.groupby(['Variable']).sum()
df2016 = df2016.reset_index()
df2016.sort_values(by=['Count'], ascending=False)

Out[12]:
```

	Variable	Count
3	Hispanic Population	4861648.0
7	White Population	2687787.0
1	Asian Population	1413105.0
2	Black Population	801182.0
6	Population of Two or More Races	220878.0
5	Other Race Population	29351.0
4	Native Hawaiian/Other Pacific Islander Population	24439.0
0	American Indian/Native Population	18765.0



Let's create a data frame only for Hispanic Population which is the first largest in Los Angeles.

	Variable	Count	Neighborhood	latitude	longitude
120953	Hispanic Population	8710.0	Palmdale	34.568801	-118.018343
120266	Hispanic Population	6940.0	East Los Angeles	34.015861	-118.184074
120365	Hispanic Population	6838.0	South Gate	33.939464	-118.195780
119773	Hispanic Population	6795.0	San Fernando	34.294876	-118.427307
120371	Hispanic Population	6665.0	Lynwood	33.921784	-118.182427

Group all neighborhoods and identify the first ten.

	Neighborhood	Count
130	Long Beach	201132.0
65	East Los Angeles	117396.0
170	Pomona	107027.0
159	Palmdale	100370.0
201	South Gate	92865.0
58	Downey	88842.0
27	Boyle Heights	86783.0
50	Compton	75440.0
251	Westlake	75434.0
156	Norwalk	74715.0

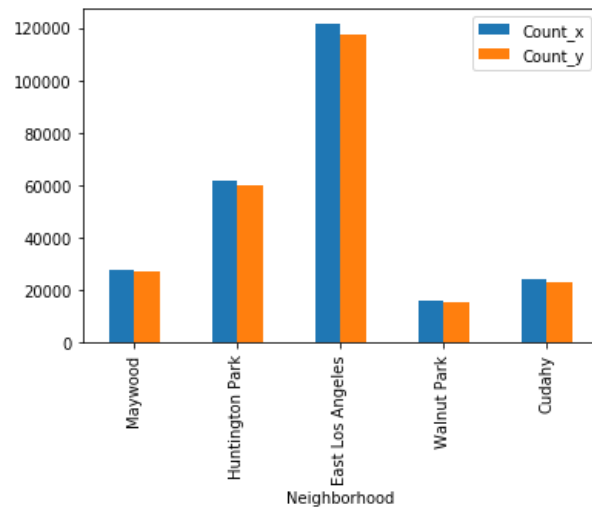
Calculate the total for each Neighborhoods without filtering

' J '.

	Neighborhood	Count
130	Long Beach	476623.0
190	Santa Clarita	195013.0
82	Glendale	191971.0
159	Palmdale	168089.0
170	Pomona	152353.0

Now, compare Hispanic participation with total participation for each neighborhood and identify the top 10.

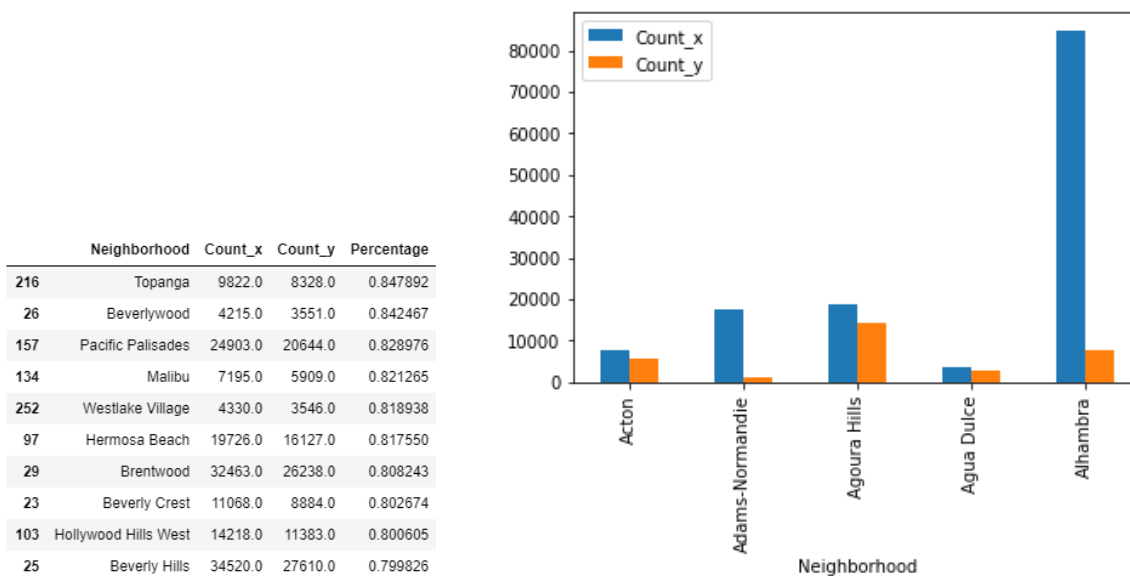
	Neighborhood	Count_x	Count_y	Percentage
140	Maywood	27672.0	27094.0	0.979112
104	Huntington Park	61630.0	59789.0	0.970128
65	East Los Angeles	121361.0	117396.0	0.967329
239	Walnut Park	15791.0	15197.0	0.962384
52	Cudahy	24106.0	23170.0	0.961171
21	Bell Gardens	42805.0	40867.0	0.954725
201	South Gate	97380.0	92865.0	0.953635
108	Inwindale	1319.0	1250.0	0.947688
20	Bell	35386.0	33091.0	0.935144
27	Boyle Heights	93164.0	86783.0	0.931508



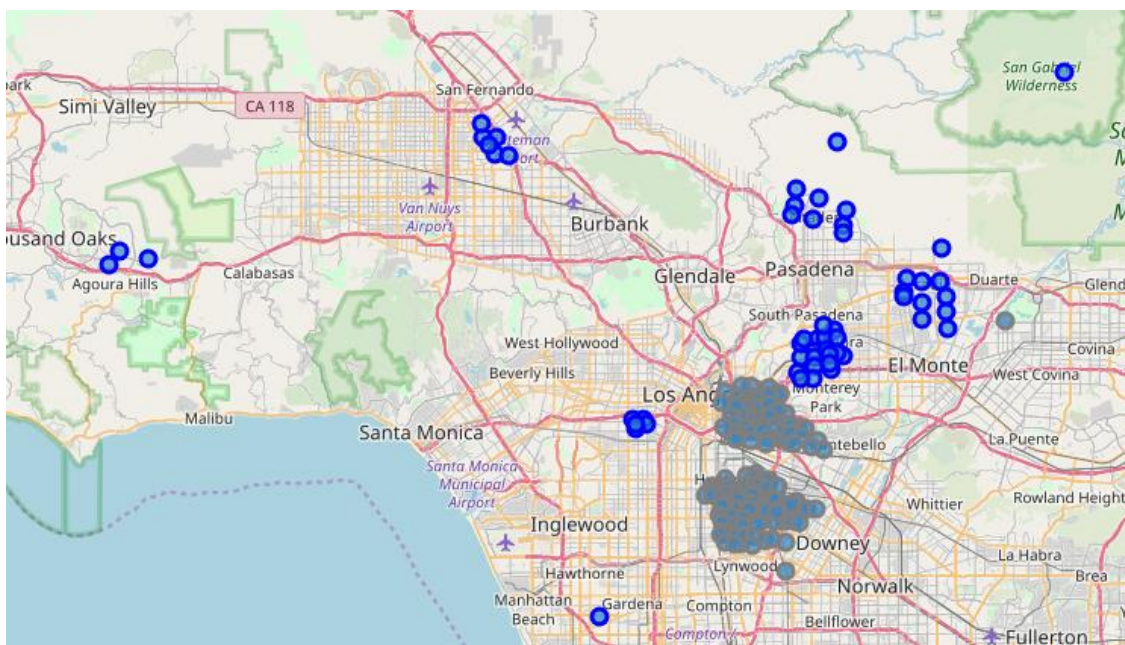
Repeat the process for the second largest group which is White Population

	Neighborhood	Count
130	Long Beach	132849.0
82	Glendale	116983.0
190	Santa Clarita	93790.0
192	Santa Monica	60297.0
217	Torrance	59390.0
31	Burbank	59352.0
164	Pasadena	52565.0
122	Lancaster	51839.0
196	Sherman Oaks	50390.0
177	Redondo Beach	41960.0

Compare with the total Neighborhoods



Plot the top ten neighborhoods for both groups



Clustering Neighborhoods

Merge both Top five dataframes which is necessary for clustering the neighborhoods

	Variable	Count	Neighborhood	latitude	longitude
4388	White Population	1407.0	Arcadia	34.162478	-118.024155
4408	White Population	2119.0	Altadena	34.200921	-118.136652
126733	White Population	348.0	Arleta	34.233766	-118.429943
128149	White Population	426.0	Arleta	34.232289	-118.418041
128150	White Population	404.0	Arleta	34.245902	-118.428032

Get nearby venues for all neighborhoods

»]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Altadena	34.200921	-118.136652	Farnsworth Park	34.199920	-118.131403	Playground
1	Altadena	34.200921	-118.136652	Charter Spectrum	34.197250	-118.135160	Business Service
2	Arleta	34.233766	-118.429943	Blackdog Yoga Studio	34.234398	-118.429553	Yoga Studio
3	Arleta	34.233766	-118.429943	Father's Donuts	34.235433	-118.433937	Donut Shop
4	Arleta	34.233766	-118.429943	Four Star Liquor And Market	34.235436	-118.434022	Liquor Store

Find the top 10 most common venues for all neighborhoods

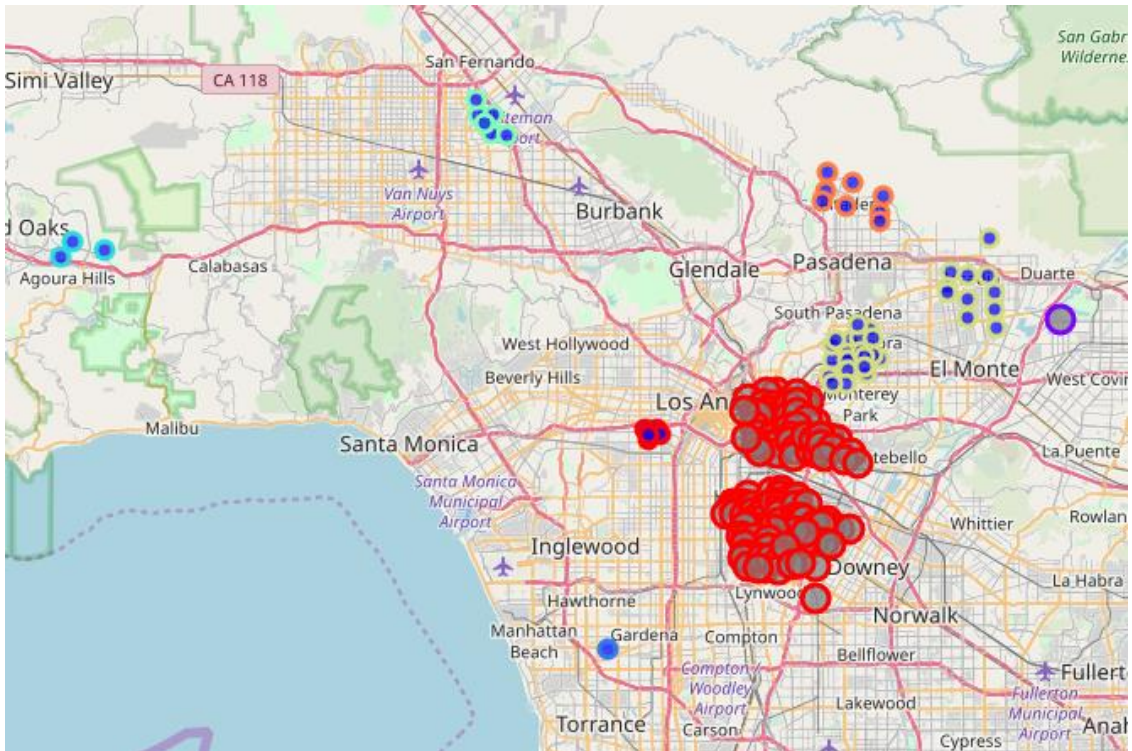
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Adams-Normandie	Sushi Restaurant	Mexican Restaurant	Taco Place	Playground	Seafood Restaurant	Ice Cream Shop	Café	Fried Chicken Joint	Restaurant	Fast Food Restaurant
1	Agoura Hills	Park	Yoga Studio	Laundry Service	Pizza Place	Photography Studio	Car Wash	Mediterranean Restaurant	Bowling Alley	Pub	Art Gallery
2	Alhambra	Fast Food Restaurant	Chinese Restaurant	Mexican Restaurant	Asian Restaurant	Vietnamese Restaurant	Convenience Store	Pizza Place	Liquor Store	Burger Joint	Bakery
3	Alondra Park	Park	Football Stadium	Spa	Food Stand	Food Court	Food & Drink Shop	Food	Flower Shop	Financial or Legal Service	Field
4	Altadena	Cosmetics Shop	Mexican Restaurant	Middle Eastern Restaurant	Coffee Shop	Grocery Store	Liquor Store	American Restaurant	Sandwich Place	Chinese Restaurant	Garden Center
5	Arcadia	Chinese Restaurant	Tea Room	Japanese Restaurant	Bakery	Noodle House	Dessert Shop	American Restaurant	Mobile Phone Shop	Asian Restaurant	Bank
6	Arleta	Mexican Restaurant	Liquor Store	Yoga Studio	Park	Bus Station	Food	Playground	Locksmith	Movie Theater	Beach Bar
7	Bell	Mexican Restaurant	Seafood Restaurant	Pizza Place	Burger Joint	Donut Shop	Market	Martial Arts Dojo	Grocery Store	Park	Gift Shop
8	Bell Gardens	Mexican Restaurant	Fast Food Restaurant	Coffee Shop	Fried Chicken Joint	American Restaurant	Food	Sporting Goods Shop	Ice Cream Shop	Thai Restaurant	Burger Joint

Setting the 7 clusters

	Variable	Count	Neighborhood	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
4388	White Population	1407.0	Arcadia	34.162478	-118.024155	5	Chinese Restaurant	Tea Room	Japanese Restaurant	Bakery	Noodle House	Dessert Shop	American Restaurant	Mobile Phone Shop	Asian Restaurant	Bank
4408	White Population	2119.0	Altadena	34.200921	-118.136652	6	Cosmetics Shop	Mexican Restaurant	Middle Eastern Restaurant	Coffee Shop	Grocery Store	Liquor Store	American Restaurant	Sandwich Place	Chinese Restaurant	Garden Center
126733	White Population	348.0	Arleta	34.233766	-118.429943	4	Mexican Restaurant	Liquor Store	Yoga Studio	Park	Bus Station	Food	Playground	Locksmith	Movie Theater	Beach Bar
128149	White Population	426.0	Arleta	34.232289	-118.418041	4	Mexican Restaurant	Liquor Store	Yoga Studio	Park	Bus Station	Food	Playground	Locksmith	Movie Theater	Beach Bar
128150	White Population	404.0	Arleta	34.245902	-118.428032	4	Mexican Restaurant	Liquor Store	Yoga Studio	Park	Bus Station	Food	Playground	Locksmith	Movie Theater	Beach Bar
Whites							Hispanics							Mexicans		

Split the dataset into Hispanic and White. Then, plot them differentiating population groups





## Showing each cluster

### Cluster No 1

```
In [40]: LAX_merged.loc[LAX_merged['Cluster Labels'] == 0, LAX_merged.columns[[1] + list(range(5, LAX_merged.shape[1]))]]
```

Out[40]:

	Count	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
128695	57.0	0	Sushi Restaurant	Mexican Restaurant	Taco Place	Playground	Seafood Restaurant	Ice Cream Shop	Café	Fried Chicken Joint	Restaurant	Fast Food Restaurant
128696	179.0	0	Sushi Restaurant	Mexican Restaurant	Taco Place	Playground	Seafood Restaurant	Ice Cream Shop	Café	Fried Chicken Joint	Restaurant	Fast Food Restaurant
128699	518.0	0	Sushi Restaurant	Mexican Restaurant	Taco Place	Playground	Seafood Restaurant	Ice Cream Shop	Café	Fried Chicken Joint	Restaurant	Fast Food Restaurant
128704	212.0	0	Sushi Restaurant	Mexican Restaurant	Taco Place	Playground	Seafood Restaurant	Ice Cream Shop	Café	Fried Chicken Joint	Restaurant	Fast Food Restaurant
128705	177.0	0	Sushi Restaurant	Mexican Restaurant	Taco Place	Playground	Seafood Restaurant	Ice Cream Shop	Café	Fried Chicken Joint	Restaurant	Fast Food Restaurant
119226	4528.0	0	Mexican Restaurant	Fast Food Restaurant	Food Truck	Pizza Place	Convenience Store	Pharmacy	Restaurant	Seafood Restaurant	Grocery Store	Taco Place
119227	4719.0	0	Mexican Restaurant	Fast Food Restaurant	Food Truck	Pizza Place	Convenience Store	Pharmacy	Restaurant	Seafood Restaurant	Grocery Store	Taco Place

### Cluster No 2

```
In [41]: LAX_merged.loc[LAX_merged['Cluster Labels'] == 1, LAX_merged.columns[[1] + list(range(5, LAX_merged.shape[1]))]]
```

Out[41]:

	Count	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
119884	1250.0	1	Auto Garage	Automotive Shop	Women's Store	Farmers Market	Football Stadium	Food Truck	Food Stand	Food Court	Food & Drink Shop	Food

### Cluster No 3



```
In [42]: LAX_merged.loc[LAX_merged['Cluster Labels'] == 2, LAX_merged.columns[[1] + list(range(5, LAX_merged.shape[1]))]]
```

Out[42]:

	Count	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
129998	1512.0	2	Park	Football Stadium	Spa	Food Stand	Food Court	Food & Drink Shop	Food	Flower Shop	Financial or Legal Service	Field

## Cluster No 4

```
In [43]: LAX_merged.loc[LAX_merged['Cluster Labels'] == 3, LAX_merged.columns[[1] + list(range(5, LAX_merged.shape[1]))]]
```

Out[43]:

	Count	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
130140	5659.0	3	Park	Yoga Studio	Laundry Service	Pizza Place	Photography Studio	Car Wash	Mediterranean Restaurant	Bowling Alley	Pub	Art Gallery
130143	3280.0	3	Park	Yoga Studio	Laundry Service	Pizza Place	Photography Studio	Car Wash	Mediterranean Restaurant	Bowling Alley	Pub	Art Gallery
130148	5389.0	3	Park	Yoga Studio	Laundry Service	Pizza Place	Photography Studio	Car Wash	Mediterranean Restaurant	Bowling Alley	Pub	Art Gallery

## Cluster No 5

```
In [44]: LAX_merged.loc[LAX_merged['Cluster Labels'] == 4, LAX_merged.columns[[1] + list(range(5, LAX_merged.shape[1]))]]
```

Out[44]:

	Count	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
126733	348.0	4	Mexican Restaurant	Liquor Store	Yoga Studio	Park	Bus Station	Food	Playground	Locksmith	Movie Theater	Beach Bar
128149	426.0	4	Mexican Restaurant	Liquor Store	Yoga Studio	Park	Bus Station	Food	Playground	Locksmith	Movie Theater	Beach Bar
128150	404.0	4	Mexican Restaurant	Liquor Store	Yoga Studio	Park	Bus Station	Food	Playground	Locksmith	Movie Theater	Beach Bar
128151	218.0	4	Mexican Restaurant	Liquor Store	Yoga Studio	Park	Bus Station	Food	Playground	Locksmith	Movie Theater	Beach Bar
128152	340.0	4	Mexican Restaurant	Liquor Store	Yoga Studio	Park	Bus Station	Food	Playground	Locksmith	Movie Theater	Beach Bar

## Cluster No 6

```
In [45]: LAX_merged.loc[LAX_merged['Cluster Labels'] == 5, LAX_merged.columns[[1] + list(range(5, LAX_merged.shape[1]))]]
```

Out[45]:

	Count	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
4388	1407.0	5	Chinese Restaurant	Tea Room	Japanese Restaurant	Bakery	Noodle House	Dessert Shop	American Restaurant	Mobile Phone Shop	Asian Restaurant	Bank
129282	1595.0	5	Chinese Restaurant	Tea Room	Japanese Restaurant	Bakery	Noodle House	Dessert Shop	American Restaurant	Mobile Phone Shop	Asian Restaurant	Bank
129283	1192.0	5	Chinese Restaurant	Tea Room	Japanese Restaurant	Bakery	Noodle House	Dessert Shop	American Restaurant	Mobile Phone Shop	Asian Restaurant	Bank
129284	899.0	5	Chinese Restaurant	Tea Room	Japanese Restaurant	Bakery	Noodle House	Dessert Shop	American Restaurant	Mobile Phone Shop	Asian Restaurant	Bank
129285	852.0	5	Chinese Restaurant	Tea Room	Japanese Restaurant	Bakery	Noodle House	Dessert Shop	American Restaurant	Mobile Phone Shop	Asian Restaurant	Bank
129286	842.0	5	Chinese Restaurant	Tea Room	Japanese Restaurant	Bakery	Noodle House	Dessert Shop	American Restaurant	Mobile Phone Shop	Asian Restaurant	Bank

## Cluster No 7

```
In [46]: LAX_merged.loc[LAX_merged['Cluster Labels'] == 6, LAX_merged.columns[[1] + list(range(5, LAX_merged.shape[1]))]]
```

Out[46]:

	Count	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
4408	2119.0	6	Cosmetics Shop	Mexican Restaurant	Middle Eastern Restaurant	Coffee Shop	Grocery Store	Liquor Store	American Restaurant	Sandwich Place	Chinese Restaurant	Garden Center
129344	3422.0	6	Cosmetics Shop	Mexican Restaurant	Middle Eastern Restaurant	Coffee Shop	Grocery Store	Liquor Store	American Restaurant	Sandwich Place	Chinese Restaurant	Garden Center
129345	1480.0	6	Cosmetics Shop	Mexican Restaurant	Middle Eastern Restaurant	Coffee Shop	Grocery Store	Liquor Store	American Restaurant	Sandwich Place	Chinese Restaurant	Garden Center
129346	882.0	6	Cosmetics Shop	Mexican Restaurant	Middle Eastern Restaurant	Coffee Shop	Grocery Store	Liquor Store	American Restaurant	Sandwich Place	Chinese Restaurant	Garden Center
129354	476.0	6	Cosmetics Shop	Mexican Restaurant	Middle Eastern Restaurant	Coffee Shop	Grocery Store	Liquor Store	American Restaurant	Sandwich Place	Chinese Restaurant	Garden Center
129355	1557.0	6	Cosmetics Shop	Mexican Restaurant	Middle Eastern Restaurant	Coffee Shop	Grocery Store	Liquor Store	American Restaurant	Sandwich Place	Chinese Restaurant	Garden Center
129356	2963.0	6	Cosmetics Shop	Mexican Restaurant	Middle Eastern Restaurant	Coffee Shop	Grocery Store	Liquor Store	American Restaurant	Sandwich Place	Chinese Restaurant	Garden Center
129357	4192.0	6	Cosmetics Shop	Mexican Restaurant	Middle Eastern Restaurant	Coffee Shop	Grocery Store	Liquor Store	American Restaurant	Sandwich Place	Chinese Restaurant	Garden Center

## 4- Results and Discussion

Our analysis identified that Hispanic Population with 48,34% and White Population with 26,73% are the largest population groups in Los Angeles. Moreover, Race & Ethnicity is very differentiated among neighborhoods. For instance, the Hispanic group (gray points) shows that the top 10 neighborhoods with the highest participation rate are over 90% which means that in all those neighborhoods most households are Hispanic communities. On the other hand, the White group (blue points) has its top ten participation rate in neighborhoods between 80% - 85%, considerably less than the Hispanic group. Another important finding is the location of the top 10 White neighborhoods which is far away from the center of Los Angeles while the Hispanic group is closer.

After identifying our Top 10 neighborhoods for both population groups based on participation rate, we clustered those neighborhoods into 7 clusters. Once we plotted all clustered location points into a map, it is remarkable the similarities among Hispanic neighborhoods which are in only 2 clusters, one of them with more than 95% of the Hispanic neighborhood while White neighborhoods have different clusters depending on the location. Probably the similarity between Hispanic neighborhoods is based on the distance among each other.

One of the most important findings was that the biggest White group cluster is very influenced by Asian food such as Chinese, Japanese, and Sea food. On the other hand, the biggest Hispanic group is influenced by Mexican Restaurant, Fast Food Restaurant, Food Truck, Pizza Place, and Convenience Store. Another important finding is that only one White neighborhood appears in a Hispanic cluster which denotes how different those races are.

Given that Hispanic neighborhoods are highly influenced by Mexican Restaurant and Fast Food Restaurant, it seems an opportunity for investors who want to start business with different approaches such as Asian food, Italian food. Furthermore, for those who want to live in Los Angeles, must deal with segregation as it is clear after this analysis.

## 5- Conclusion

The purpose of this project was to identify how is composed the neighborhoods in Los Angeles (LA), finding similarities and differences among them. By using the Race & Ethnicity dataset, the analysis focused on the two largest population by races, in order to understand how those races influenced the LA neighborhoods. The results obtained showed an impressive neighborhood segregation among Hispanic population and White population, leading to a clear cluster classification and easy identification of potential investments such as Asian food and Italian food in Hispanic neighborhoods which according to the clusters, those are no common places there.

Of course, it is necessary further analysis to be sure about investments, however the resultant clusters serve as a starting analysis.