USING FOURSQAURE APP TO CHOOSE A CITY NEIGHBORHOOD TO RENT IN.

CAPSTONE ASSIGNMENT (C9_WK4,5_ASSIGNMENT_RD)

Author: ryan desousa

1. INTRODUCTION

- 1.1 Background and Business Problem
- As people relocate between cities for work or otherwise, they often start out in the new location by renting.
- As you move to a new city, how would one choose the area of the new city in which to live?
- This would depend a lot on the individuals' personal choices and lifestyle.
- Using rental data that is available online, I would like to explore the use of the Four Square API app to see if the combination can be used to get a better understanding of neighborhoods in a target city and if, based on user input and preference, any neighborhoods in a city are of greater interest to rent in as compared to others.
- This target audience are people who are moving to new cities and towns and who want to get a better idea how a different neighborhoods in the city will suit their needs.
- The city that I will be exploring will be Los Angeles, CA.

1. INTRODUCTION

1.2) Data

- The average rental rates for different neighborhoods in the city of Los Angeles are available at: https://www.rentcafe.com/average-rent-market-trends/us/ca/los-angeles/
- Extract data to csv and then into a dataframe and cleaned.
- Pair this data, with information on the different LA neighborhoods, obtained through the Four Square API app.
- Use this rental data and the API app to explore different venues in the LA neighborhoods to try get an idea of what are the high traffic offerings in the various neighborhoods.
- Use one hot encoding to try to find the ten most frequently occurring venues in the different neighborhoods.
- Apply K-Means Clustering Machine Learning algorithm on the acquired neighborhood-venue data to cluster the data into five clusters.
- Filter the Neighborhood-Venue frequency data set, using venues of my choice and look for neighborhoods where these venues are among the top three most frequent venues.
- Use the FourSquare API, to search any two of these neighborhoods further, with five venues of my choice and try to determine the frequency of occurrence of these venues, within a distance of 2000m of their neighborhood coordinates.
- Once I have all this information, I will try to put all this information together to look for similarities/differences in the two neighborhoods and see if any one neighborhood is more preferred to rent in than the other.
- The final goal is to see whether this data can be used to differentiate between the neighborhoods and be used by people to make a better decisions on choice of city neighborhood to rent in.

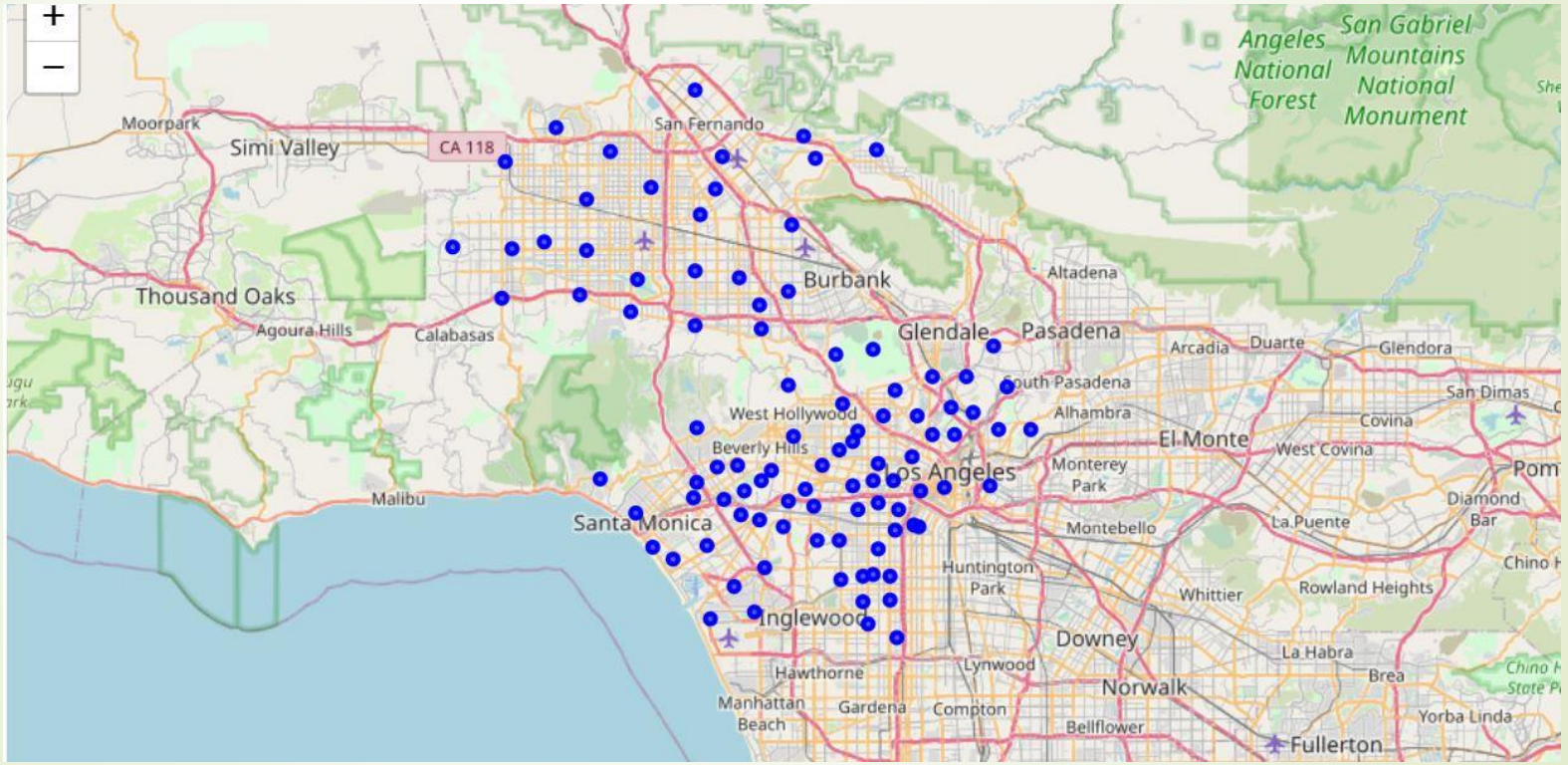
2 . DATA HANDLING

- Screen shot of the final treated dataset of LA Neighbohood, average monthly rent, latitude and longitude as a dataframe. (Total 97 neighborhoods)

	LA_Neighborhood	AverageMonthlyRent_USD	Latitude	Longitude
0	Jefferson Park, CA, USA	1355.0	34.027234	-118.317576
1	El Sereno, CA, USA	1396.0	34.081121	-118.177849
2	Vermont Vista, CA, USA	1445.0	33.941947	-118.285814
3	Vermont Knolls, CA, USA	1445.0	33.966819	-118.291670
4	Hyde Park, CA, USA	1484.0	33.980569	-118.330631

2 . DATA HANDLING

- The 97 neighborhoods plotted on a map of LA using Folium



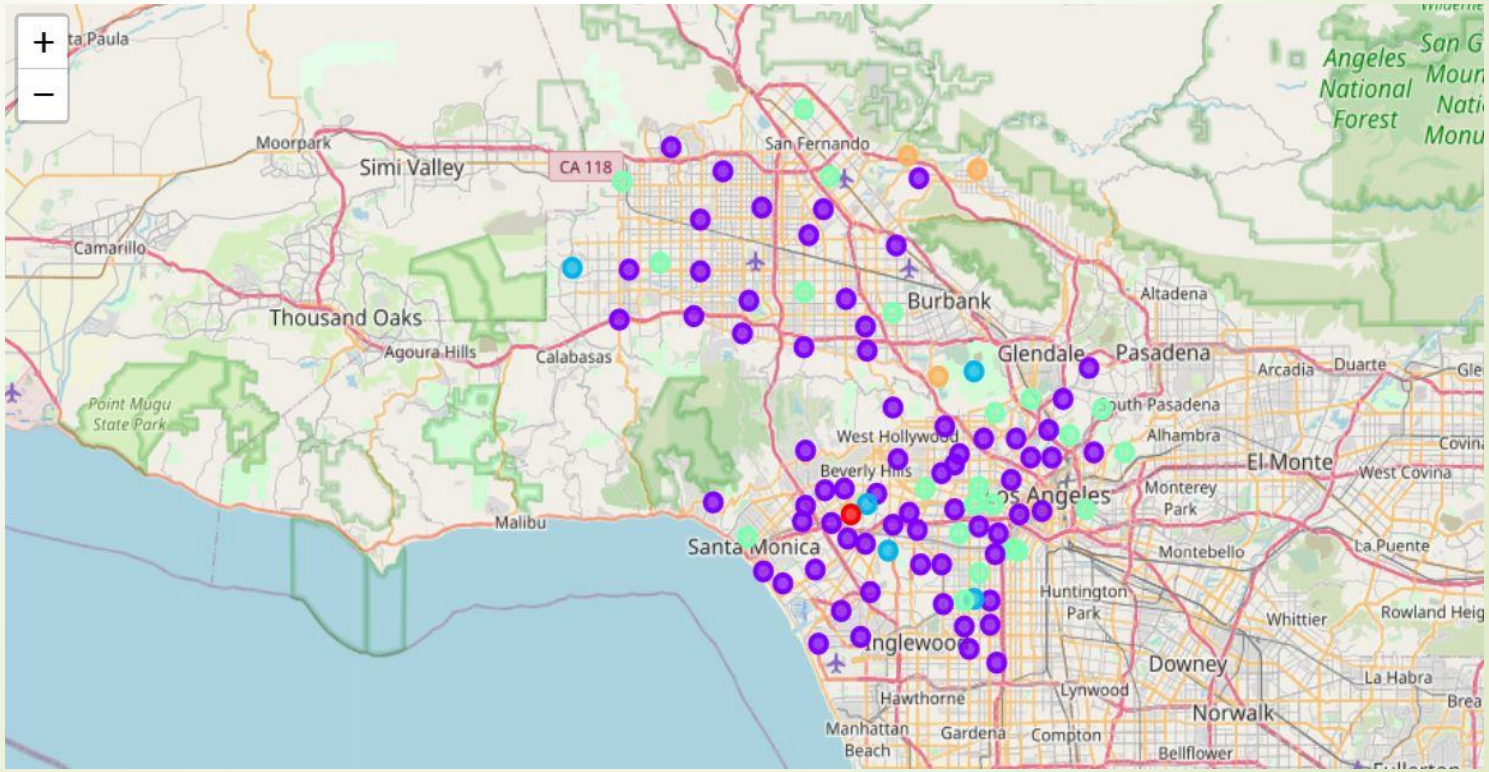
3) BATTLE OF LA NEIGHBORHOODS - METHODOLOGY: EXPLORE VENUES

- Using rental data and the FourSquare API app explored different venues in the LA neighborhoods to try get an idea of what are the high traffic offerings in the various neighborhoods.
- Used one hot encoding to find the ten most frequently occurring venues in the different neighborhoods as shown in the figure.

	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, CA, USA	Sushi Restaurant	Taco Place	Gas Station	Wings Joint	Food Service	Flower Shop	Filipino Restaurant	Fast Food Restaurant	Farmers Market	Farm
1	Arleta, CA, USA	Movie Theater	Historic Site	Wings Joint	Donut Shop	Flower Shop	Filipino Restaurant	Fast Food Restaurant	Farmers Market	Farm	Dive Bar
2	Arlington Heights, CA, USA	Convenience Store	Art Gallery	Restaurant	Café	Latin American Restaurant	Donut Shop	Flower Shop	Filipino Restaurant	Fast Food Restaurant	Farmers Market
3	Atwater Village, CA, USA	Mexican Restaurant	Bakery	Pizza Place	Italian Restaurant	Farmers Market	Food Service	Food	Flower Shop	Filipino Restaurant	Fast Food Restaurant
4	Baldwin Hills - Crenshaw, CA, USA	Trail	Park	Music Venue	Wings Joint	Donut Shop	Flower Shop	Filipino Restaurant	Fast Food Restaurant	Farmers Market	Farm

4) BATTLE OF LA NEIGHBORHOODS - METHODOLOGY: CLUSTER VENUES

- Applied K-Means Clustering Machine Learning algorithm on the acquired neighborhood-venue data to cluster the data into five clusters and viewed using Folium Maps



5) BATTLE OF LA NEIGHBORHOODS - METHODOLOGY: LOOK FOR SPECIFIC VENUES OF INTEREST

- Filter the Neighborhood-Venue frequency data set, using 'Theater' as venue of my choice.
- Look for neighborhoods where this venue was among the top three most common venues.

	Neighborhood	AverageMonthlyRent_USD	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 M Comm Ve
0	Arleta, CA, USA	1634.0	34.241327	-118.432205	0	Movie Theater	Historic Site	Wings Joint	Donut Shop	Flower Shop	Filipino Restaurant	Fast F Restau
1	Larchmont, CA, USA	2140.0	34.079837	-118.317870	0	Indie Movie Theater	Korean Restaurant	Movie Theater	Park	American Restaurant	Cuban Restaurant	D St
2	Mid-City, CA, USA	2188.0	34.041527	-118.360370	0	Indie Theater	Gym / Fitness Center	Liquor Store	Theater	Food Truck	Dance Studio	D Boo
3	Hollywood, CA, USA	2485.0	34.098003	-118.329523	0	Coffee Shop	Movie Theater	Multiplex	Salon / Barbershop	Farmers Market	Wings Joint	D S
4	Exposition Park, CA, USA	3522.0	34.013654	-118.287211	0	Science Museum	College Football Field	Movie Theater	Park	Wings Joint	Filipino Restaurant	Fast F Restau

5) BATTLE OF LA NEIGHBORHOODS - METHODOLOGY: LOOK FOR SPECIFIC VENUES OF INTEREST

- Using the FourSquare API, search any two of these neighborhoods further (this will be based on user choice), with five venues of choice (again this will be based on user choice).
- The aim was to determine the frequency of occurrence of these five venues, within a distance of 2000m of their neighborhood coordinates.
- The five venues that I used here were Pharmacy, Market, Theater, Library and Restaurant.
- Use all this information together to look for similarities/differences in the two neighborhoods and see if any one neighborhood is more preferred to rent in than the other.
- This similarity or difference is based on the average monthly rent, number of venues found, the number of each type of venue found and the median distance from the neighborhood coordinates.
- This information is summarized in the Report Section.
- The final goal is to see whether this data can be used to differentiate between the neighborhoods and be used by people to make a better decisions on choice of city neighborhood to rent in.

6. FINAL REPORT - BATTLE OF THE NEIGHBORHOODS

- Neighborhood1: Exposition Park, LA, CA, USA.
- Neighborhood2: Arleta, LA, CA, USA.
- Both these neighborhoods were chosen from a list of neighborhoods that had 'Theatre' as the top three most common venues.
- These two neighborhoods were then searched using FourSquare API for presence of the venues Pharmacy, Market, Theater, Library and Restaurant within a radius of 2000 meters.

6. FINAL REPORT - BATTLE OF THE NEIGHBORHOODS

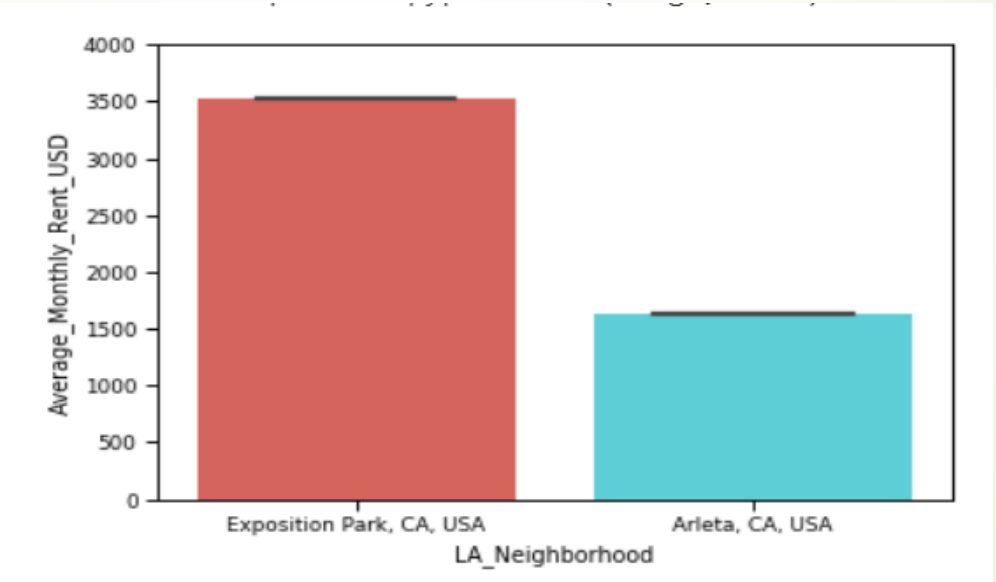
- The data for the two neighborhoods is summarized in Figure 6.1

FINAL NEIGHBORHOODS COMPARISON, SUMMARY by RD :

	LA_Neighborhood	Average_Monthly_Rent_USD	Radius (m)	Location	Found	Median Distance (km)
0	Exposition Park, CA, USA	3522.0	2000	Pharmacy	5	1.10
1	Exposition Park, CA, USA	3522.0	2000	Market	5	0.90
2	Exposition Park, CA, USA	3522.0	2000	Library	5	0.61
3	Exposition Park, CA, USA	3522.0	2000	Theater	4	0.73
4	Exposition Park, CA, USA	3522.0	2000	Restaurant	6	0.50
0	Arleta, CA, USA	1634.0	2000	Pharmacy	5	1.03
1	Arleta, CA, USA	1634.0	2000	Market	4	0.92
2	Arleta, CA, USA	1634.0	2000	Library	3	2.34
3	Arleta, CA, USA	1634.0	2000	Theater	1	1.74
4	Arleta, CA, USA	1634.0	2000	Restaurant	5	1.07

6. FINAL REPORT - BATTLE OF THE NEIGHBORHOODS

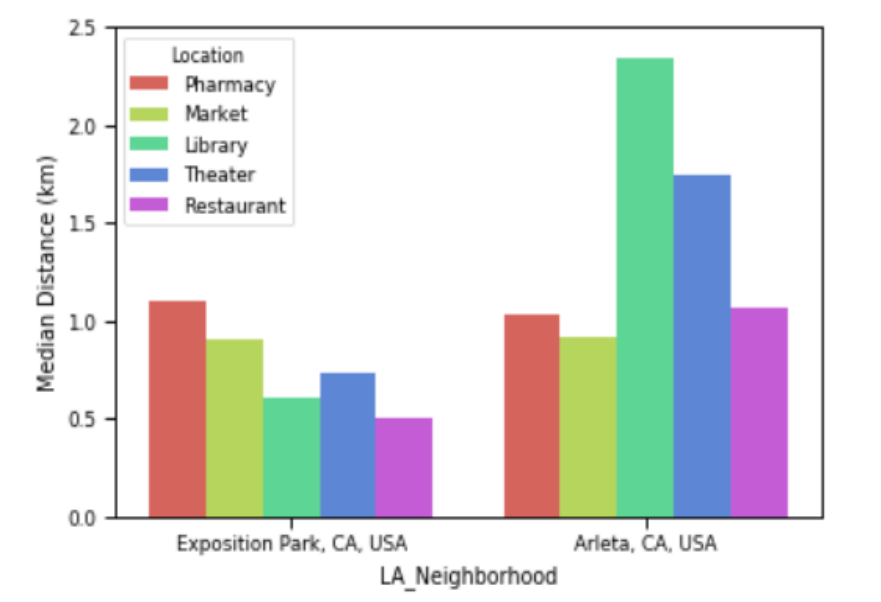
- 6.1 AVERAGE MONTHLY RENT



- The Average monthly rent is higher in Exposition Park than in Arleta neighborhood of Los Angeles.

6. FINAL REPORT - BATTLE OF THE NEIGHBORHOODS

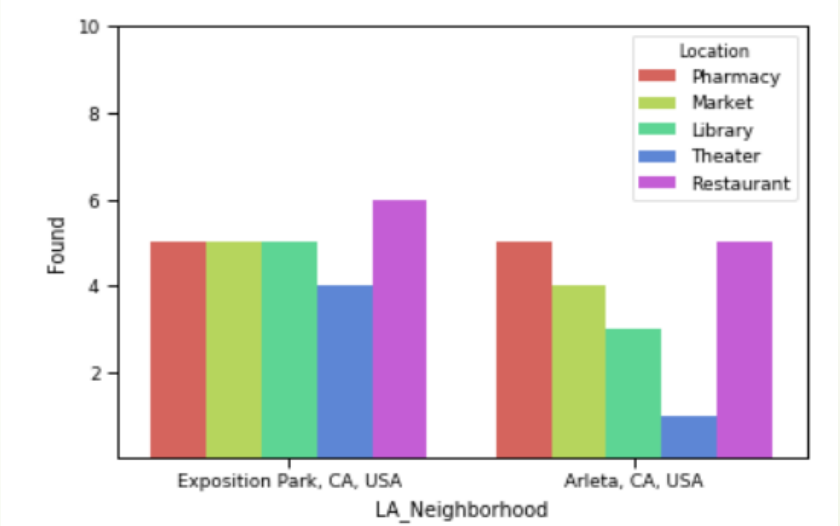
- 6.2 MEDIAN_DISTANCE



- Both Exposition Park and Arleta neighborhoods have all of the of the five venues of interest being represented.
- The median distance to the venues present appears to be higher in Arleta than in Exposition Park.

6. FINAL REPORT - BATTLE OF THE NEIGHBORHOODS

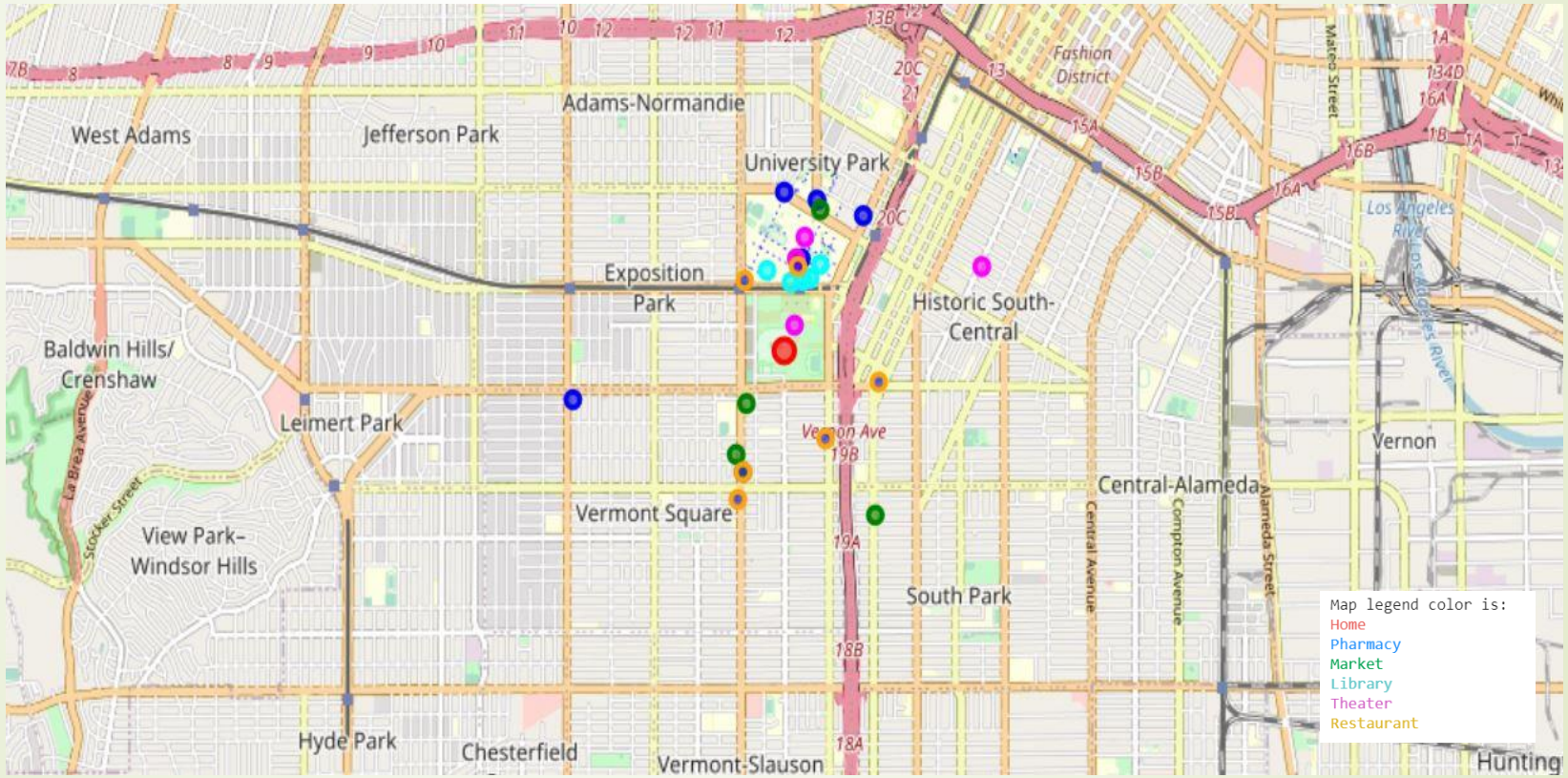
- 6.3 NUMBER OF VENUES FOUND



- Both Arleta and Exposition Park have all of the five venues present within 2000m of the target co-ordinates.
- Exposition Park has the same number or more of each venue type found, within 2000m of the target coordinates.

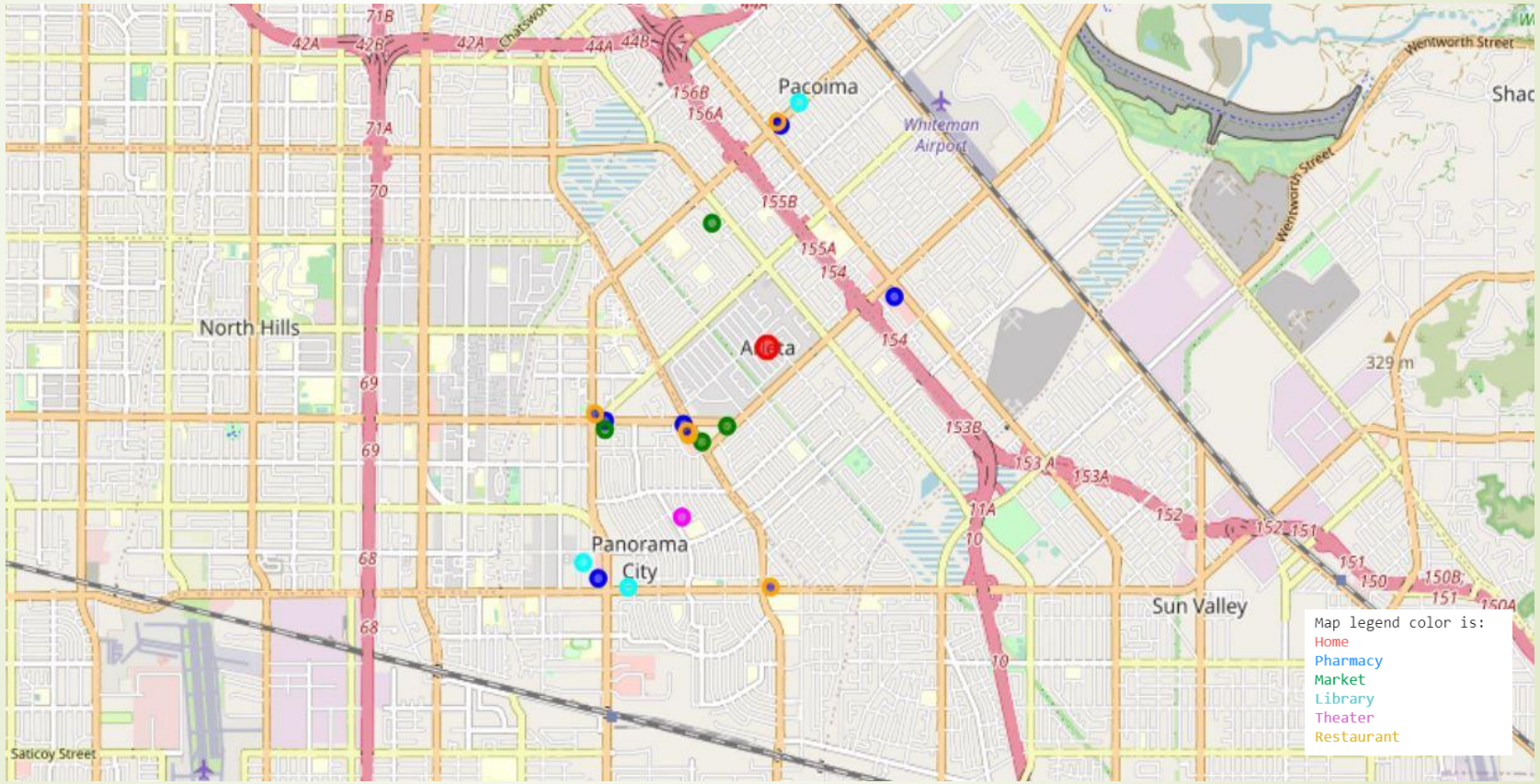
6. FINAL REPORT - BATTLE OF THE NEIGHBORHOODS

- Map of found venues in Exposition Park



6. FINAL REPORT - BATTLE OF THE NEIGHBORHOODS

- Map of found venues in Arleta



7. CONCLUSION

In Conclusion:

- I think that it is possible to use the neighborhood data, rental data and the FourSquare API to explore and search neighborhood for venues of ones choice and come to an informed decision on which neighborhood is better to rent in.
- This choice will likely be very personal.

Future Improvements:

- More venues can be searched per neighborhood using a bigger radius.
- More than two neighborhoods can be compared.
- I think that adding crime data for various neighborhoods would add a new dimension to the data, that would help make a more informed decision. (I was unable to locate suitably paired neighborhood - rent - crime data.)