Identify new optimal location for a Bowling Alley business.

By Shaq – The Battle of the Neighborhoods.

Introduction

Background & Problem

Californians by nature are fun loving people it does not matter what race, culture or region they belong to. They love to spend quality time with family and friends. Los Angeles (LA) is a densely populated metro city. To avoid the daily rat race, people live in the suburbs of LA and still like to be called as Angelenos. Lots of people live in the suburbs of LA and most of the suburbs in outskirt of LA fall under the vicinity of the largest valley in the world called San Fernando Valley (SFV). This is where the city of west hills is located.

A well-known Bowling chain is our audience and stakeholder. The sponsors/stakeholders are looking forward to open up a new location for their bowling alley in the SFV's West Hills city. *The problem is where should they open their Bowling Alley in the city in order to get minimum competition, maximum customer turnout in short find an optimal location?* This project specifically targets the stakeholders who are interested in opening a new location for **Bowling Alley** in the city of **West Hills of California,** USA.

Problem:
Finding an optimal location for Bowling Alley

- Bowlers Eye is a well known Bowling Alley chain.
- Looking for opening a new location in San Fernando Valley, California.
- Target areas with less competitors and maximum turn out of customers.
- Bowlers Eye will invest in more research if needed.

Data Acquisition and Cleaning

Data Preprocessing

- Street Address with the geolocation coordinates Dataset acquired from Redfin.com.
- Demographics data acquired from Niche.com.
- **Foursquare API** – For most common venues and category along with geolocations.
- City's geocoding data was acquired through GeoPy using Nominatim.
- Data was cleaned and combined into one dataset.
- Duplicate observations and irrelevant features were removed.

Demographics of West Hills

Education Levels		National
Master's degree or higher	18%	12%
Bachelor's degree	29%	20%
Some college or associate's degree	28%	29%
High school diploma or equivalent	18%	27%
Less than high school diploma	7%	12%

Household Income Brackets	
<$25k	8%
$25-$44k	9%
$45-$74k	13%
$75-$149k	33%
$150k+	36%

Gender	
Female	52%
Male	48%

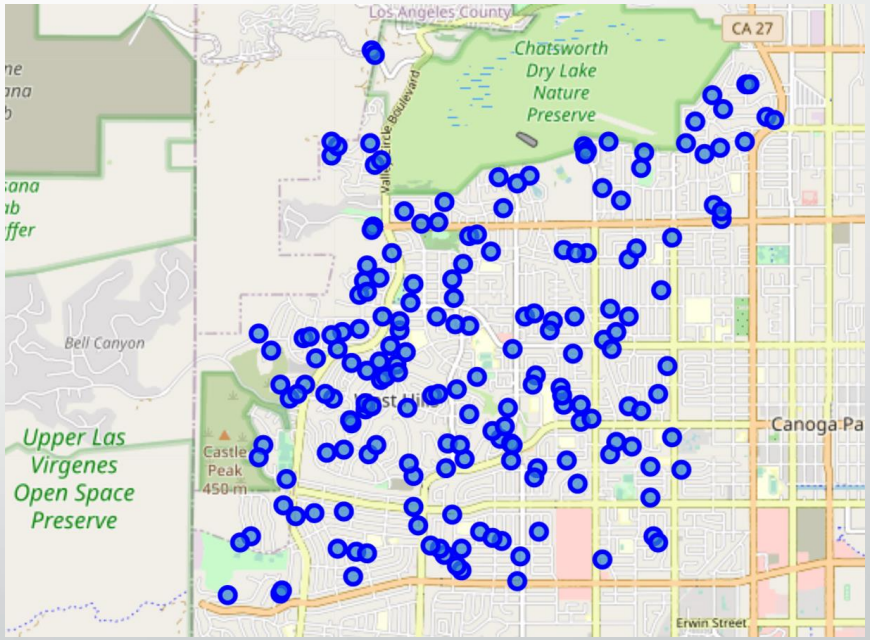
Age	
<10 years	11%
10-17 years	10%
18-24 years	7%
25-34 years	10%
35-44 years	12%
45-54 years	16%
55-64 years	16%
65+ years	18%

Methodology

- After acquiring and cleaning the data gathered through different sources, we got our master dataset reduced to *neighborhood, zip code, latitude and longitude* as main features.

	Neighborhood	ZipCode	Latitude	Longitude
0	Cowper Ave	91304	34.214568	-118.647376
1	Asman Ave	91307	34.203795	-118.616474
2	Rudnick Ave	91304	34.226878	-118.613329
3	Community St	91304	34.221487	-118.640221
4	Mencken Ave	91304	34.214276	-118.649154
5	Hamlin St	91307	34.188693	-118.622719

- Used Folium to show the geographic location of the City and the Neighborhoods.



- Utilizing the Foursquare API to explore the areas and venues, a limit of 100 venues and 500 meter for each neighborhood was set. First the city data was merged with the Foursquare API. Below the table snapshot.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Cowper Ave	34.214568	-118.647376	Lazy J Park	34.212002	-118.644622	Park
1	Asman Ave	34.203795	-118.616474	Go's Mart	34.200696	-118.613450	Sushi Restaurant
2	Asman Ave	34.203795	-118.616474	Ginger Thai	34.200518	-118.613958	Thai Restaurant
3	Asman Ave	34.203795	-118.616474	Sze-chwan Inn	34.202187	-118.613009	Chinese Restaurant
4	Asman Ave	34.203795	-118.616474	Del Taco	34.199975	-118.614415	Fast Food Restaurant

- Then we checked how many venues were returned by each neighborhood.

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Wyandotte St	5	5	5	5	5	5
Angela Ave	4	4	4	4	4	4
Archwood St	31	31	31	31	31	31
Arminta St	1	1	1	1	1	1
Ashton Ct	2	2	2	2	2	2
...	...	...	...	...	...	...

- There were 117 unique categories return for all the venues.
- After Analyzing each neighborhood for different categories we got the below observations, with new dataframe size of (1179, 118).

	Neighborhood	ATM	Airport	American Restaurant	Arcade	Arts & Crafts Store	Astrologer	Baby Store	Bagel Shop	Bakery	...	Thai Restaurant	Theater	Theme Restaurant	Tourist Information Center	Trail	Video Game Store	V S
0	Cowper Ave	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	
1	Asman Ave	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	
2	Asman Ave	0	0	0	0	0	0	0	0	0	...	1	0	0	0	0	0	
3	Asman Ave	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	
4	Asman Ave	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	

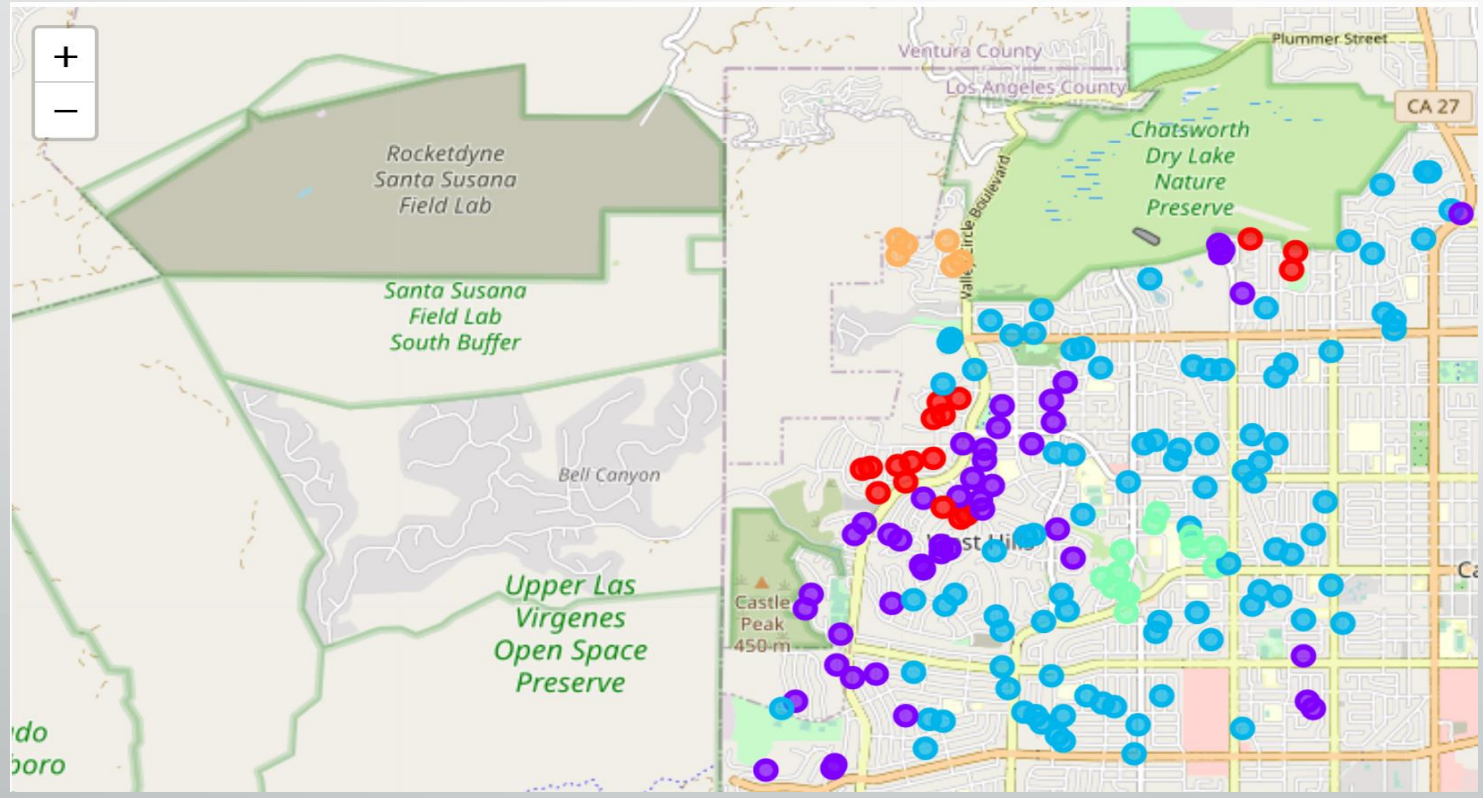
- We then grouped rows by neighborhood and by taking the mean of the frequency occurrence for each category. Snapshot is pasted below. After Analyzing each neighborhood for different categories we got the below observations, with new dataframe size of (1179, 118).

	Neighborhood	ATM	Airport	American Restaurant	Arcade	Arts & Crafts Store	Astrologer	Baby Store	Bagel Shop	Bakery	...	Thai Restaurant	Theater	Tourist Information Center	Trail	Video Game Store
0	Wyandotte St	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.0	0.000000
1	Angela Ave	0.000000	0.0	0.0	0.0	0.0	0.333333	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.0	0.000000
2	Archwood St	0.035714	0.0	0.0	0.0	0.0	0.000000	0.0	0.035714	0.035714	...	0.0	0.0	0.0	0.0	0.035714
3	Arminta St	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.0	0.000000
4	Ashton Ct	0.000000	0.5	0.0	0.0	0.0	0.000000	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.0	0.000000

- Next Foursquare came handy in finding out the list of top 10 venue category for each area, below is the table snapshot.

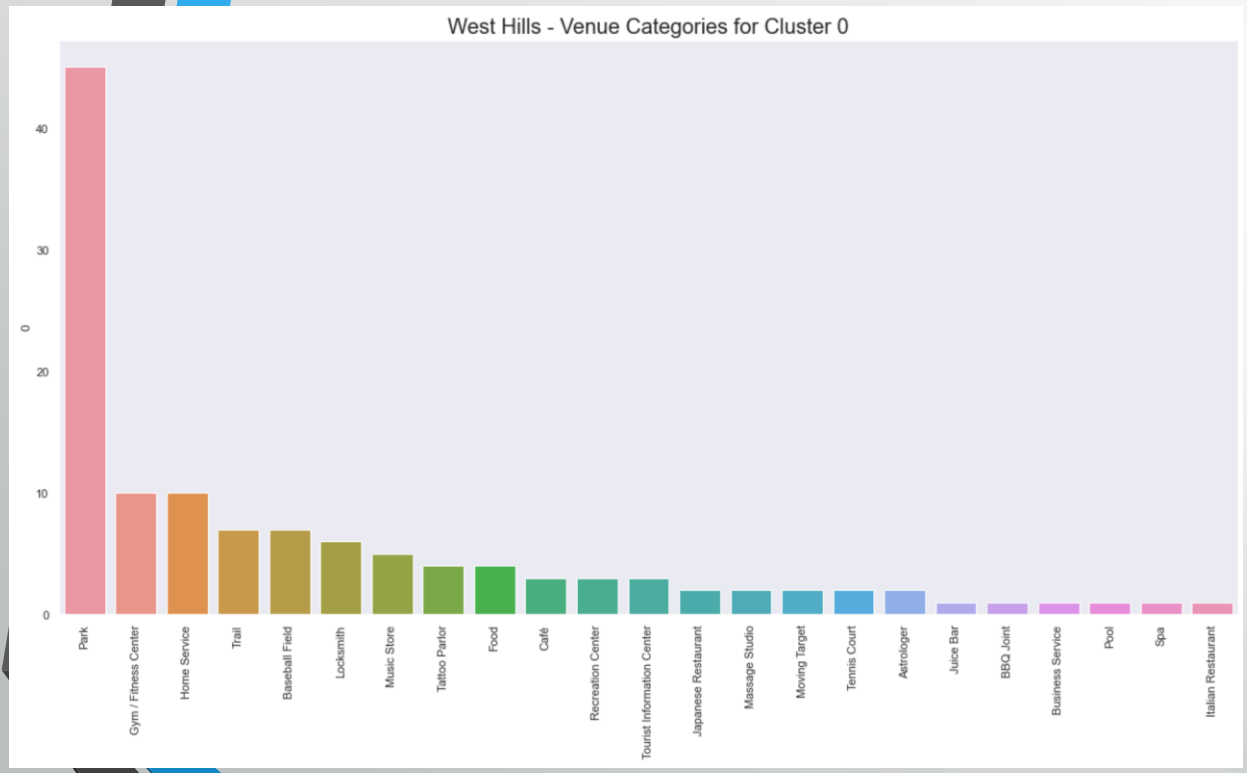
	Neighborhood	1st Most Common Venue	2dr Most Common Venue	3ct Most Common Venue	4plrd Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Wyandotte St	Pharmacy	Sandwich Place	Home Service	Gym	Wings Joint	Food Truck	Diner	Discount Store	Dive Bar	Donut Shop
1	Angela Ave	Astrologer	Home Service	Park	Food	Wings Joint	Furniture / Home Store	Discount Store	Dive Bar	Donut Shop	Dry Cleaner
2	Archwood St	Ice Cream Shop	Chinese Restaurant	Salon / Barbershop	Mexican Restaurant	Food & Drink Shop	Discount Store	Coffee Shop	Convenience Store	Cosmetics Shop	Pizza Place
3	Arminta St	Park	Wings Joint	Furniture / Home Store	Discount Store	Dive Bar	Donut Shop	Dry Cleaner	Fabric Shop	Fast Food Restaurant	Flower Shop
4	Ashton Ct	Airport	Park	Wings Joint	Furniture / Home Store	Discount Store	Dive Bar	Donut Shop	Dry Cleaner	Fabric Shop	Fast Food Restaurant

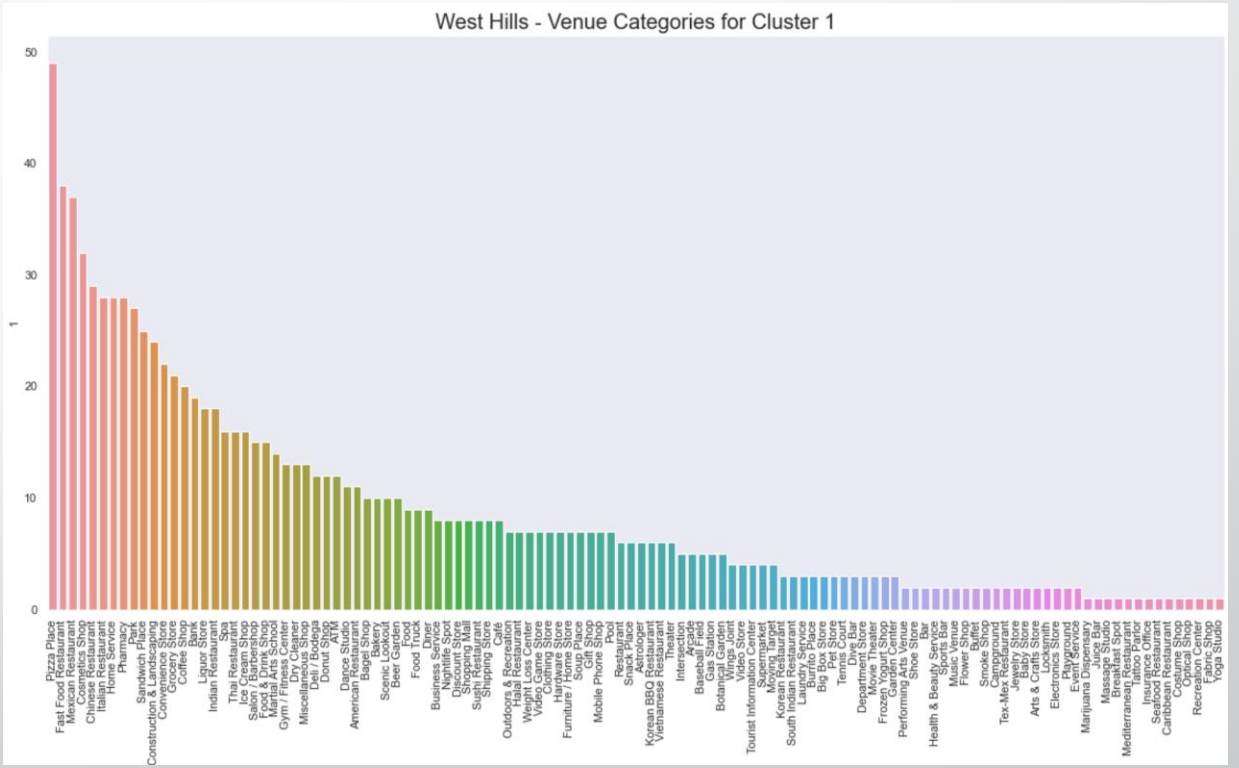
- There were some common venue categories among the neighborhoods. For which we used unsupervised learning K-Means algorithm to cluster the neighborhoods. This will help us identify the cluster with most venues of similar types and helps segmenting.
- Our clustering process gave us 5 Clusters, as below.
- We used Folium to plot the clusters on the map. Snapshot Below.

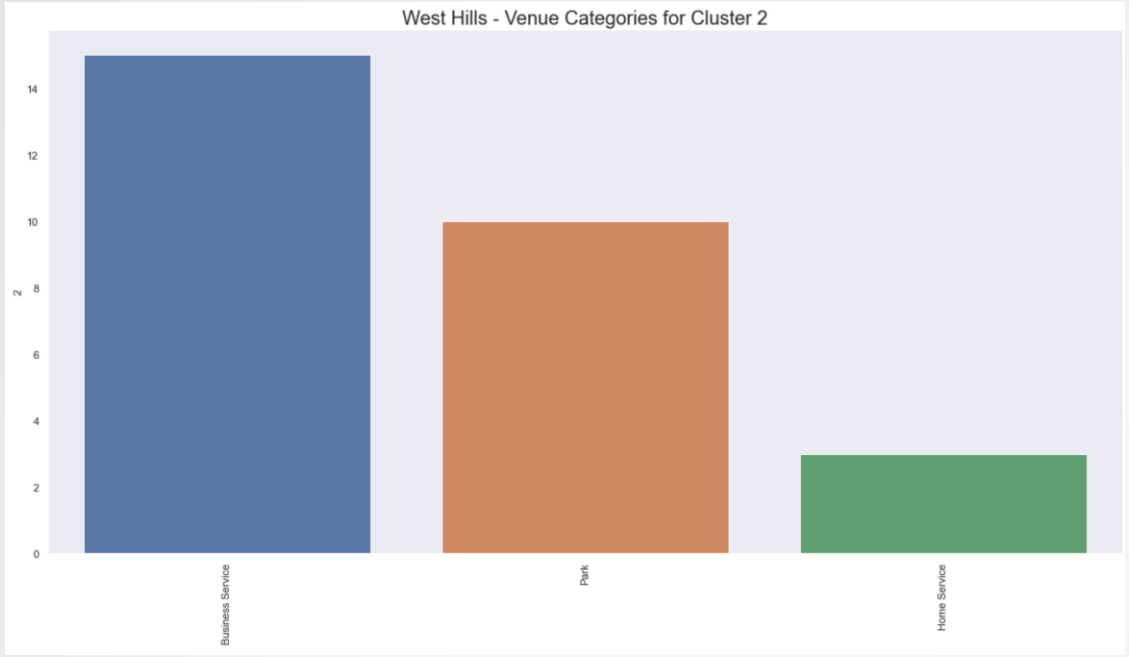


Results

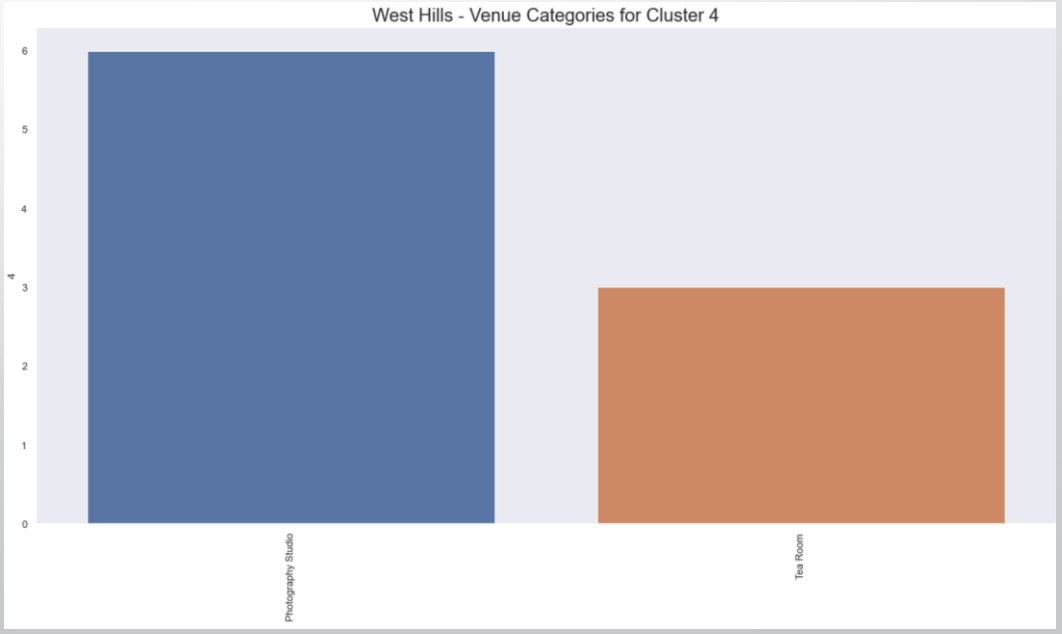
After clustering on venues and categories during the analysis we also discovered that a Cluster 0 and Cluster 1 have the greatest number of venues and activities. We discovered that there were many venues were common among the areas, however they all lacked the recreation activity precisely as Bowling. Please notice the clusters charts below.






West Hills - Venue Categories for Cluster 2
2
0
2
4
6
8
10
12
14
Business Service
Park
Home Service


West Hills - Venue Categories for Cluster 3
3
0
1
2
3
4
5
6
7
8
Park


West Hills - Venue Categories for Cluster 4
4
0
1
2
3
4
5
6
Photography Studio
Tea Room

Discussion

As our initial approach to find the most happening neighborhood and minimal competition in the bowling business, this report would be very beneficial to the client for decision making for a perfectly optimal location to open the new bowling alley. With all the given information it looks like locations in Cluster 1 would be more suitable. Lastly, there could be more considerations such as more detailed supportive data in order to come up with a more accurate finding.

Our research revealed that a great number of venues/businesses are in West Hills, however it appears to have a very small number of family recreational activity in the vicinity. Specially a Bowling Alley type of activity. Most of the business are in the north-west of the city, however we found that the businesses are all over in cluster 4 as we use the K-Means clustering. Whatever the case we did not find any bowling alley in any of the neighborhoods. Practical our client (A well-known Bowling Chain) could open a new location either in the center of the city or north-west or east (which is covered in cluster 1). For future direction we could acquire more data on the city population behavior, income, age percentage area wise and so on, then we could do some more analysis and provide better accuracy.

Conclusion

The purpose of this research was to identify the best optimal location for opening a new branch of a Bowling Alley, with minimal competitor presence and maximum venue activity. We are successful in empowering our client/stakeholder to make the decision on their own, as we presented the statistics of the similar businesses in a neighborhood and common happening places and so on. With the help of our research we found that there is no bowling alley in the city of West hills, which means zero or no competition in the similar type of business that our client is in.

Final decision on optimal location for the Bowling Alley will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended area.