NEW RESTAURANT SITE EVALUATION RECOMMENDATIONS

MICHAEL HALL

MJHALL@UCSD.EDU

AGENDA

Business Problem

Data

Methodology

Results

Discussion

Conclusion

BUSINESS PROBLEM

XYZ Restaurant Group has an existing restaurant in Solana Beach, CA . This restaurant has been very successful and XYZ Restaurant Group is looking to open second restaurant somewhere within San Diego County. They believe that the location is one of the key contributors to the success of any restaurant. They want to evaluate additional communities that are closely aligned to the characteristics of Solana Beach. Based on prior experience and research on other successful restaurants, XYZ Restaurant Group would like the have the locations evaluated on the demographics of the community. Some of the characteristics that they would like to be evaluated include:

- Population
- Income
- Nearby venues and attractions
- Age
- Housing ownership

The question that XYZ ultimately wants answered is –

“Which of the target communities most closely aligns with characteristics of our current, successful restaurant in Solana Beach?"

DATA

Dataset	Description	Use	Source	Format
Community Venues	Venue name, location, category and other firmographic information	Identify competitors and venues within a community	https://www.foursquare.com	JSON
Community Demographics	2010 US Census demographics aggregated and summarized by community	Profile of the population and other demographics of the community	http://datasurfer.sandag.org/	Excel
Community Location	Boundaries of the communities	Boundaries of the communities	http://rdw.sandag.org/	Geojson

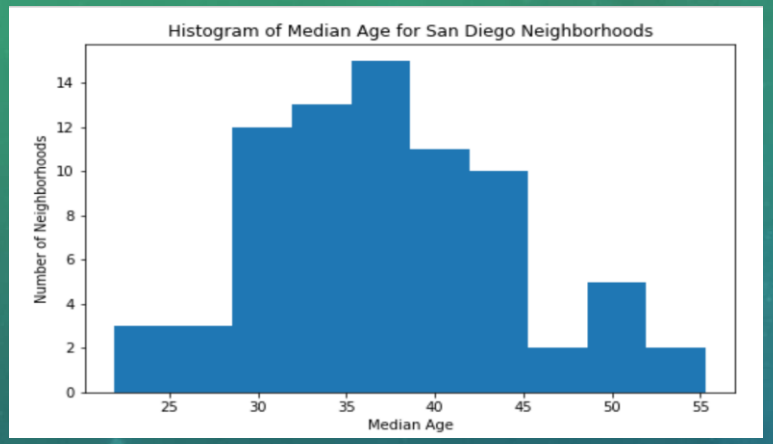
METHODOLOGY

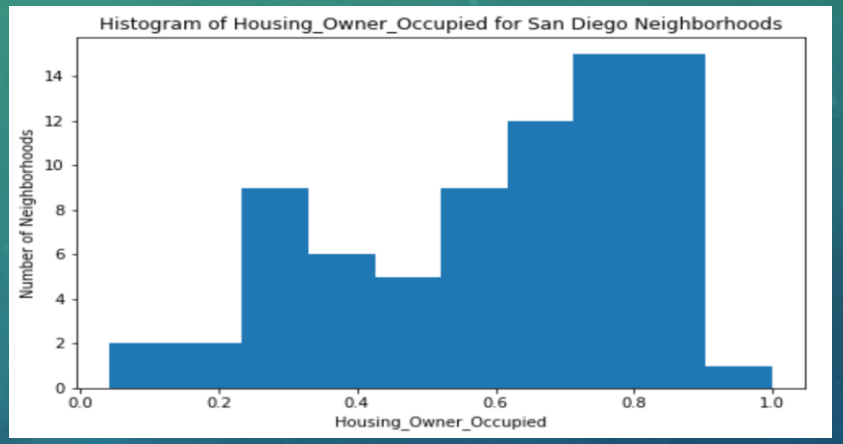
The 2010 census data obtained from San Diego Association of Governments (SANDAG) was used to identify the demographic characteristics of the neighborhoods. This data was already aggregated and summarized for each of the neighborhoods. Additional data preparation was done to filter the various statistics down to just the the few key attributes to be evaluated:

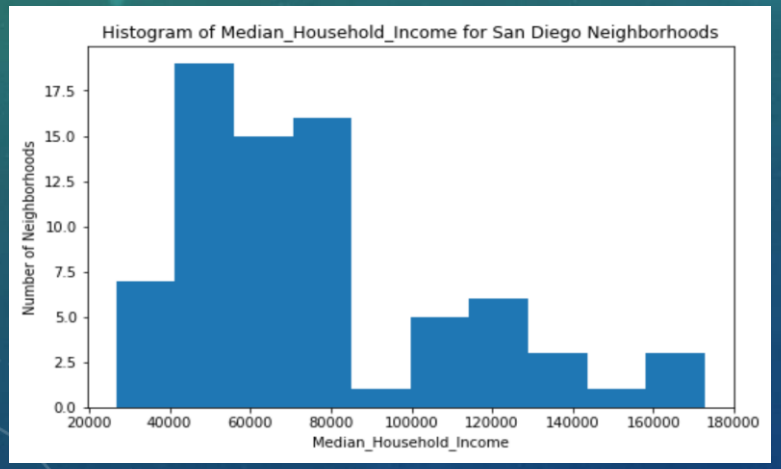
- Median household income
- Median age
- % of home ownership
- Number of households

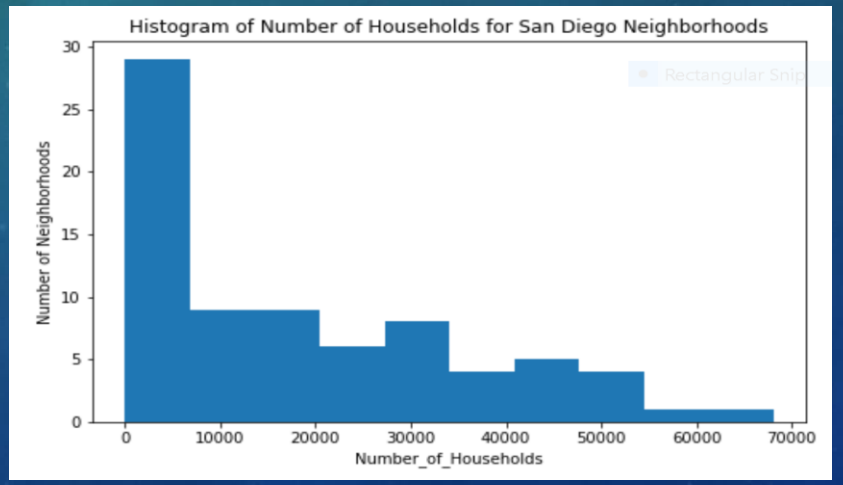
Each of the attributes were analyzed utilizing histograms and were segmented into three categories each based on the distributions

DEMOGRAPHICS HISTOGRAMS

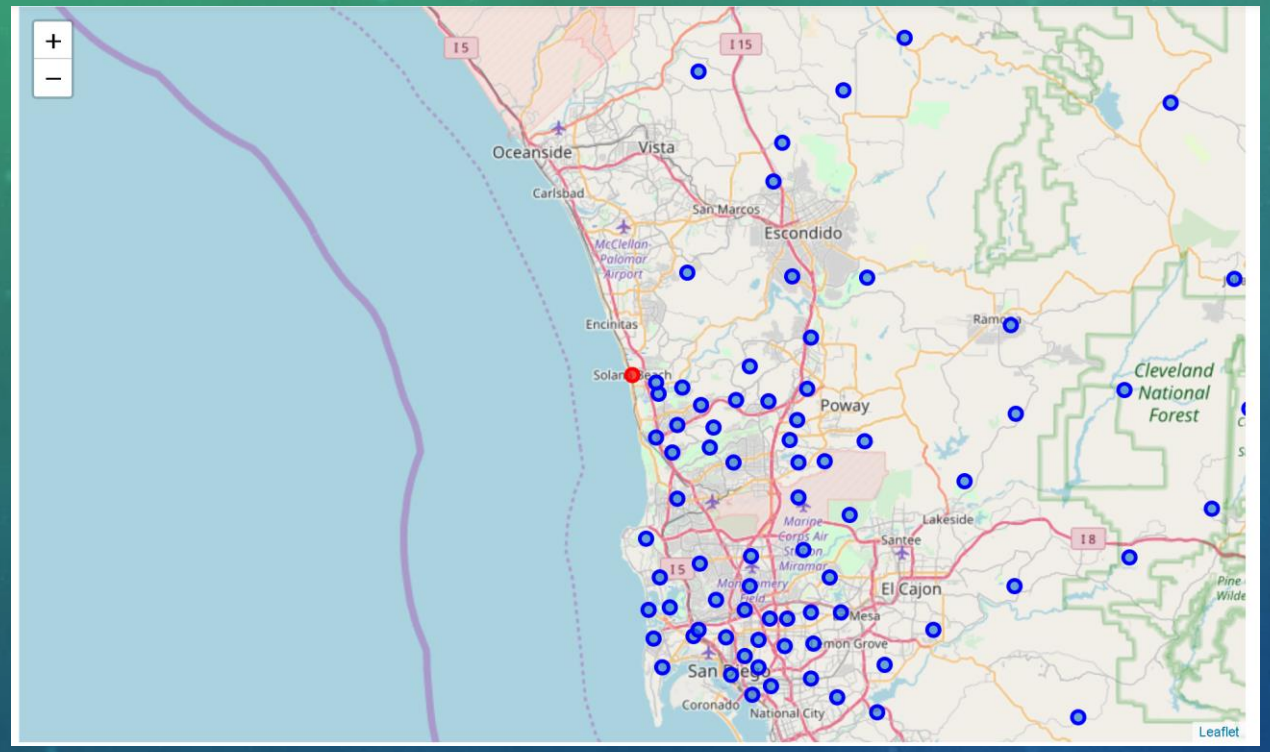








OBTAIN AND PLOT THE NEIGHBORHOODS



Solana Beach is represented by the red dot.

The other points are the centroids of each of the neighborhoods.

Solana Beach is in the “Via de la Valle” neighborhood.

OBTAIN UP TO THE 100 VENUES

	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 M Comn Ve
0	Alpine	Clothing Store	Shoe Store	Mexican Restaurant	Grocery Store	Sandwich Place	Accessories Store	Fast Food Restaurant	Kids Store	American Restaurant	Ca
1	Balboa Park	Zoo Exhibit	American Restaurant	Italian Restaurant	Pizza Place	Theater	Brewery	Farmers Market	Mexican Restaurant	Park	
2	Barona	Casino	American Restaurant	Café	Mexican Restaurant	Gym / Fitness Center	Athletics & Sports	Asian Restaurant	Steakhouse	Park	Ita Restau
3	Barrio Logan	Hotel	Park	Bar	Brewery	Mexican Restaurant	Breakfast Spot	Café	Steakhouse	Taco Place	Ita Restau
4	Black Mountain Ranch	Coffee Shop	Mexican Restaurant	Sandwich Place	Gym	Sushi Restaurant	Gym / Fitness Center	Golf Course	Grocery Store	Pizza Place	Video S
5	Bonsall	Golf Course	Farm	Mexican Restaurant	Garden Center	Fast Food Restaurant	Food & Drink Shop	Garden	Bed & Breakfast	Scenic Lookout	Liquor S
6	Borrego Springs	Scenic Lookout	Golf Course	Hotel	Home Service	Campground	New American Restaurant	Farm	Fast Food Restaurant	Eye Doctor	Fabric S
7	Boulevard	Restaurant	Food	RV Park	Mountain	Scenic Lookout	Resort	Zoo	Eye Doctor	Fabric Shop	
8	Carmel Mountain Ranch	Coffee Shop	Mexican Restaurant	Grocery Store	Sushi Restaurant	Chinese Restaurant	Italian Restaurant	Pizza Place	Sandwich Place	Donut Shop	Gr Restau
9	Carmel Valley	Coffee Shop	Seafood Restaurant	American Restaurant	Trail	Mexican Restaurant	Beach	Park	Restaurant	Golf Course	Ita Restau
10	Central Mountain	Trail	Waterfall	Zoo	Filipino Restaurant	Fabric Shop	Fair	Falafel Restaurant	Farm	Farmers Market	Fast F Restau

NORMALIZE THE DATA FOR THE MODEL

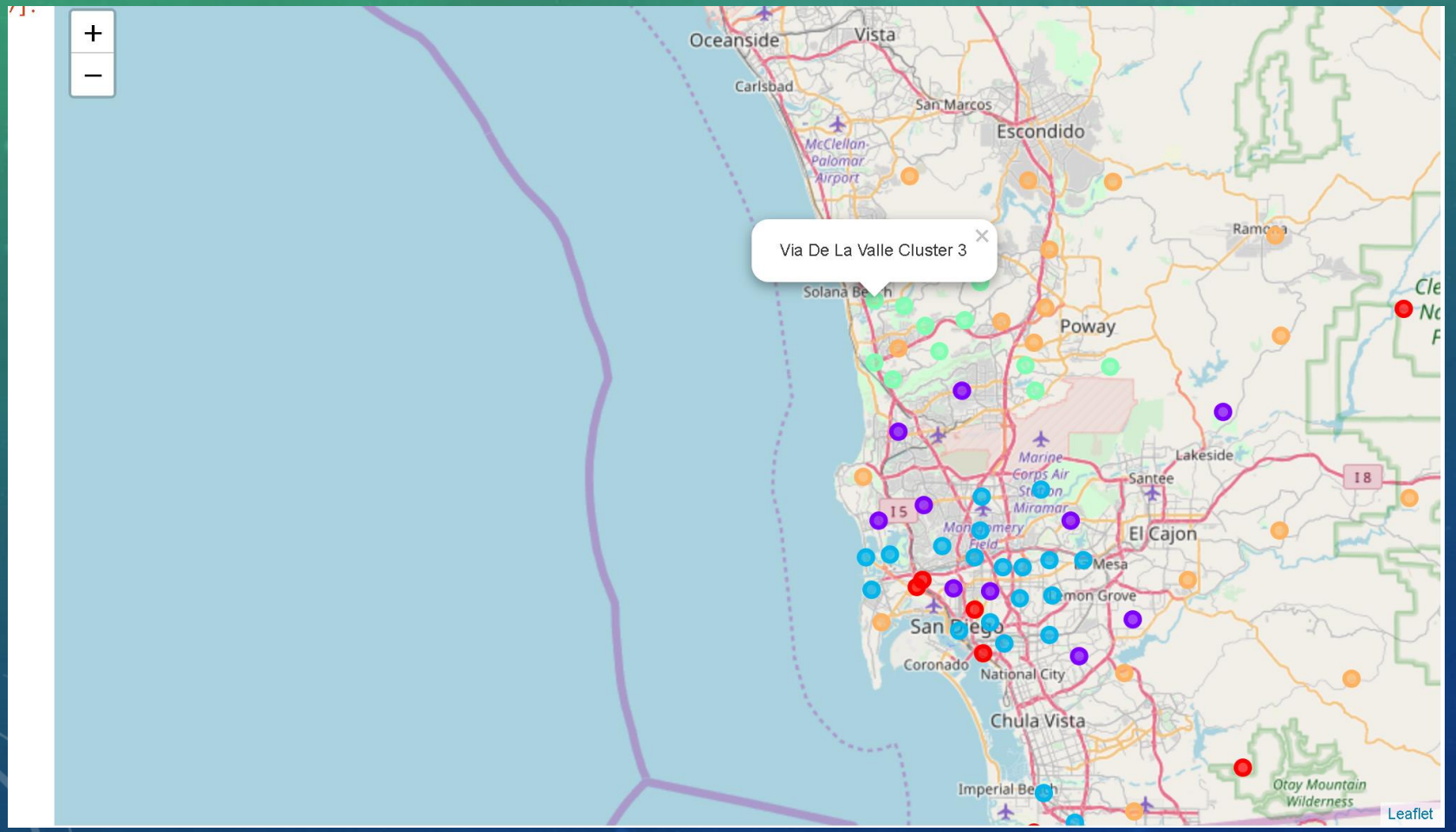
Lets merge the venue and demographic attributes for the neighborhoods

```
sandiego_demos = pd.merge(sandiego_grouped, census2010, on='Neighborhood', how='inner')
```

```
sandiego_demos.head()
```

	Neighborhood	Zoo Exhibit	ATM	Accessories Store	Adult Boutique	Afghan Restaurant	Airport	Airport Lounge	Airport Terminal	American Restaurant	Amphitheater	Antique Shop	Aquarium	Arcade	Art Gallery
0	Alpine	0.00	0.0	0.035714	0.00	0.0	0.0	0.0	0.0	0.035714	0.00	0.0	0.0	0.0	0.00
1	Balboa Park	0.11	0.0	0.010000	0.01	0.0	0.0	0.0	0.0	0.050000	0.01	0.0	0.0	0.0	0.00
2	Barona	0.00	0.0	0.000000	0.00	0.0	0.0	0.0	0.0	0.142857	0.00	0.0	0.0	0.0	0.00
3	Barrio Logan	0.00	0.0	0.000000	0.00	0.0	0.0	0.0	0.0	0.020000	0.00	0.0	0.0	0.0	0.01
4	Black Mountain Ranch	0.00	0.0	0.000000	0.00	0.0	0.0	0.0	0.0	0.000000	0.00	0.0	0.0	0.0	0.00

RESULTS - USING K-MEANS CREATE 5 CLUSTERS



RESULTS FOR CLUSTER 3

	Neighborhood	Age_Young	Age_Middle	Age_Old	Rent	Rent_or_Own	Own	Low_Income	Middle_Income	High_Income	Small_Number_Households	Middl
4	Black Mountain Ranch	0	1	0	0	0	1	0	0	1	1	
13	Del Mar Mesa	0	1	0	0	0	1	0	0	1	1	
17	Fairbanks Ranch Country Club	0	1	0	0	0	1	0	0	1	1	
33	Miramar Ranch North	0	1	0	0	0	1	0	0	1	0	
47	Pacific Highlands Ranch	0	1	0	0	0	1	0	0	1	1	
54	Rancho Encantada	0	1	0	0	0	1	0	0	1	1	
60	Scripps Miramar Ranch	0	1	0	0	0	1	0	0	1	0	
68	Torrey Highlands	0	1	0	0	0	1	0	0	1	1	
69	Torrey Hills	0	1	0	0	1	0	0	0	1	0	
70	Torrey Pines	0	0	1	0	0	1	0	0	1	0	
75	Via De La Valle	0	0	1	0	0	1	0	0	1	1	

The current restaurant neighborhood

THE CHARACTERISTICS OF CLUSTER THREE

- Middle Age
- Own home
- High Income
- Small to medium number of households
- Coffee shops, Mexican restaurants and seafood restaurants are the most common venues

The other neigborhoods identified within the cluster include: Black Mountain Ranch, Del Mar Mesa, Fairbanks Ranch Country Club, Miramar Ranch North, Pacific Highlands Ranch, Rancho Encantanda, Scripps Miramar Ranch, Torrey Highlands, Torrey Hills and Torrey Pines.

DISCUSSION

- The resulting analysis was able to clearly identify a set of neighborhoods that were closely aligned to the neighborhood where the current restaurant is located. In addition, the identified attributes were consistent with the original hypothesis that income, age and home ownership would be contributing factors to a successful location. The one inconsistent result was that the lower population/number of households was identified. This was the result of the current restaurant being in a middle sized population center.
- The recommendation is to evaluate potential restaurant sites in the identified neighborhoods: Black Mountain Ranch, Del Mar Mesa, Fairbanks Ranch Country Club, Miramar Ranch North, Pacific Highlands Ranch, Rancho Encantanda, Scripps Miramar Ranch, Torrey Highlands, Torrey Hills or Torrey Pines.

CONCLUSION

The use of k-means clustering proved to be an effective tool for identifying neighborhoods similar to the current restaurant location. Additional factors that could be evaluated in the future include targeting specific locations rather than general neighborhood centroids. In this particular case, certain areas covered broad areas and the results may not have been as accurate versus looking at smaller, more targeted locations. In addition, additional demographic variables, such as number of children and commute times could be utilized.

Thank You

Michael Hall

mjhall@ucsd.edu