

Comparing Orange County, CA and Hennepin County, MN

Erick Campos

December 20, 2020

1. Introduction

Disclaimer – this business report is for the final project for Coursera Course “Applied Data Science Capstone.” While the data compiled is real, the proposed business case is entirely fictitious. Any real-world resemblance to this report is purely coincidental. Any thoughts and opinions presented are my own.

1.1 Background

Disneyland Resort is located in Orange County, CA and Walt Disney World is located in Orange County, FL. These two locations are on opposite coasts of the United States. These properties bring a large volume of economic value for its host city and the surrounding cities in its county. It could be advantageous to place a third Disney property somewhere in the middle of the United States. This project aims to compare Hennepin County, Minnesota to the location of Disneyland (the first Disney theme park) in Orange County, California to decide whether or not Hennepin County is a viable option for a third Disney property in the US. The comparison will be made using mapping with k-means clustering.

1.2 Problem

Data is restricted to the zip code data available for both counties, as well as the data recorded on the Foursquare API. Anything that exists in either county but is not available on the Foursquare database will be missed. The location of Hennepin County was picked based on an estimate of distance. The distance between Orange County, CA, Hennepin County, MN, and Orange County, FL would more or less create a triangle with sides of equal distance. The k-means factor selected is arbitrary in order to keep things consistent between the two counties.

1.3 Interest

The potential stakeholders of interest include The Walt Disney Company, as well as the residents of Hennepin. Things to consider are monetary benefits or costs as well as how adding a large theme park would affect the way of life in Hennepin County.

2. Data Acquisition and Cleaning

2.1 Data Sources

The two data sources used to extract the zip codes for Orange County, CA and Hennepin County, MN are found at <https://www.zipcodestogo.com/California/> and <https://www.zipcodestogo.com/Minnesota/> . The latitude and longitude data for the zip codes come from a function that calls upon the ArcGIS mapping API. The data source for the venues of interest for both counties will come from the Foursquare API.

2.2 Data Cleaning

The data from the two webpages listed in the previous section were transposed into two tables, one for California and one for Minnesota. The data tables created include columns for zip codes, cities, and counties. From there, the data tables were filtered on the counties of interest: Orange County and

Hennepin County. The data was then mapped to ensure that there were no data discrepancies. After that, the location data was matched to its corresponding venue data from the Foursquare API. Any rows containing NaN values were once again removed to avoid errors. The final step was clustering the location and venue data based on similar venues using a k-means clustering with a factor of 5.

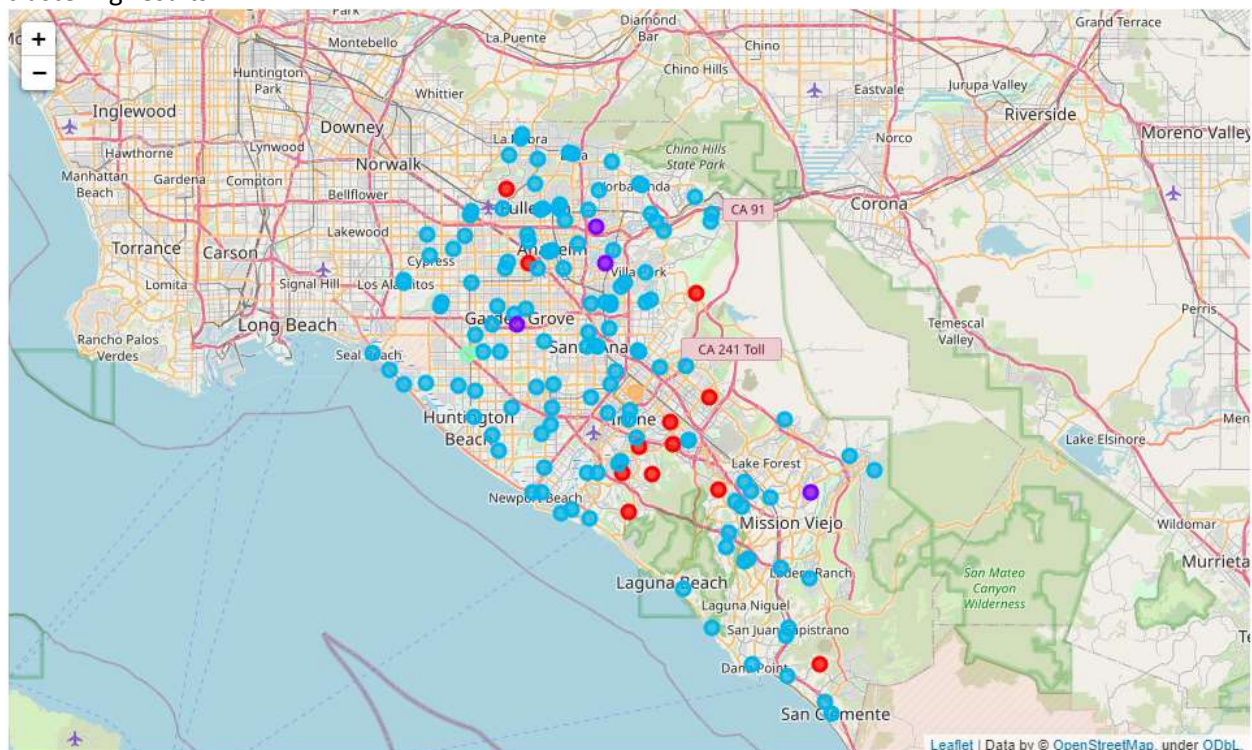
2.3 Feature Selection

The information to be examined is the way the venues for each county were clustered. The comparison between the two counties will be made using a k-means clustering with a factor of 5. The clustering will be determined based on the frequency of each venue transposed as a “one-hot” or “dummy” variable. The decision to place a Disney theme park in Hennepin County will be made based on the comparison between the two different sets of clusters. The clusters do not have to be an exact match – rather we must make assumptions as to why it would be justifiable to place a Disney property in Hennepin County based on the story that those clusters are telling us. We must keep in mind that the clusters in Orange County may have been influenced by Disneyland and not the other way around.

3. Exploratory Data Analysis

3.1 Visual mapping of Orange County, CA

The venue features for zip codes were extracted from the Foursquare API and then joined to the original table of zip code coordinates. The following map of Orange County was generated based on the k-means clustering results:



3.2 Orange County, Cluster 1 (Red)

The first cluster is as follows:

	City	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
16	Irvine	0.0	Park	Tennis Court	Restaurant	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop
17	Irvine	0.0	Park	Playground	Massage Studio	Gym	Food Court	Food & Drink Shop	Food	Food Service	Flower Shop	Fabric Shop
23	Irvine	0.0	Park	Gym	Food Truck	Playground	Fish & Chips Shop	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service
27	Irvine	0.0	Park	Dog Run	Pool	Scenic Lookout	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop
28	Irvine	0.0	Park	Pizza Place	Fish Market	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop	Flower Shop
30	Irvine	0.0	Park	Pool	Music Venue	Gym / Fitness Center	Fish & Chips Shop	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant
47	Laguna Hills	0.0	Park	Rental Car Location	Sports Bar	Convenience Store	Fish Market	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service
51	Newport Coast	0.0	Pool	Park	Flower Shop	Tennis Court	Indoor Play Area	Coffee Shop	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service
59	San Clemente	0.0	Park	Airport Terminal	Trail	Fish Market	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop
97	Anaheim	0.0	Park	Performing Arts Venue	Food Truck	Fish & Chips Shop	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish Market
117	Fullerton	0.0	Park	Pool Hall	Playground	Kitchen Supply Store	Fish Market	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service
135	Orange	0.0	Park	Train	Bike Rental / Bike Share	Yoga Studio	Fish Market	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop

Based on the common venues, this cluster can be labeled as recreational areas, since they mostly include parks, pools, playgrounds, and other outdoor activities. From my trivial knowledge, these zip codes and cities make up some of the more affluent areas in Orange County.

3.3 Orange County, Cluster 2 (Purple)

The second cluster is as follows:

	City	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
72	Mission Viejo	1.0	Home Service	Falafel Restaurant	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop	Fish Market	Flower Shop	Yoga Studio
101	Anaheim	1.0	Brewery	Music Venue	BBQ Joint	Home Service	Yoga Studio	Flower Shop	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop
126	Garden Grove	1.0	Men's Store	Shopping Plaza	Home Service	Bar	Fish Market	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop
138	Orange	1.0	Thrift / Vintage Store	Home Service	Flower Shop	Mexican Restaurant	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop	Fish Market

Although this data set is a lot smaller, it can be deduced that this area also comprises of recreational areas, but not outdoor recreation like the previous cluster. These zip codes have more indoor venues and restaurants, maybe more for fun nights out.

3.4 Orange County, Cluster 3 (Blue)

The third cluster is as follows:

	City	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
0	Buena Park	2.0	Baseball Stadium	Sushi Restaurant	Liquor Store	Playground	Yoga Studio	Fish Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop
1	Buena Park	2.0	Liquor Store	Korean Restaurant	Gastropub	Mexican Restaurant	Spa	Pizza Place	Fish Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service
2	Buena Park	2.0	Theme Park Ride / Attraction	Theme Park	Theme Restaurant	Theater	American Restaurant	Dessert Shop	Electronics Store	Pharmacy	Flower Shop	Park
3	La Palma	2.0	Korean Restaurant	Asian Restaurant	Bubble Tea Shop	Mexican Restaurant	Fast Food Restaurant	Bakery	Soup Place	Spa	Supermarket	Grocery Store
4	Buena Park	2.0	Korean Restaurant	Pizza Place	Gastropub	Fried Chicken Joint	Liquor Store	Rental Car Location	Spa	Mexican Restaurant	Shopping Mall	Fast Food Restaurant
...
144	Placentia	2.0	Coffee Shop	Mexican Restaurant	Pharmacy	Sandwich Place	Chinese Restaurant	Comfort Food Restaurant	Noodle House	Supermarket	Fried Chicken Joint	Breakfast Spot
145	Yorba Linda	2.0	Yoga Studio	Historic Site	Library	Basketball Court	Grocery Store	New American Restaurant	Boxing Gym	Pharmacy	Optical Shop	Cosmetics Shop
146	Yorba Linda	2.0	Mexican Restaurant	Yoga Studio	Historic Site	Cosmetics Shop	Pharmacy	Sandwich Place	New American Restaurant	Boxing Gym	Tex-Mex Restaurant	Library
147	Yorba Linda	2.0	Trail	Basketball Court	Park	Bed & Breakfast	Yoga Studio	Farm	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop
148	Anaheim	2.0	Dance Studio	Sushi Restaurant	Rental Car Location	Miscellaneous Shop	Flower Shop	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop

This cluster is the largest data set of all. Based on the common venues as well as my own trivial knowledge, I know that these zip codes are the surrounding areas of theme parks, including the Disneyland Resort and Knott's Berry Farm. This area is home to a large variety of cultures, which is why there are so many different restaurants, including Mexican, Korean, Filipino, and Vietnamese. This could be the most important cluster, as it looks like this is the cluster that has built itself around the theme park and entertainment industry.

3.5 Orange County, Cluster 4 (Green)

The fourth cluster is as follows:

	City	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
68	Westminster	3.0	Sporting Goods Shop	Yoga Studio	Fish Market	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop	Flower Shop

It only includes one zip code. I am going to consider this an outlier, as it includes a venue that the other clusters did not have: a farm.

3.6 Orange County, Cluster 5 (Yellow)

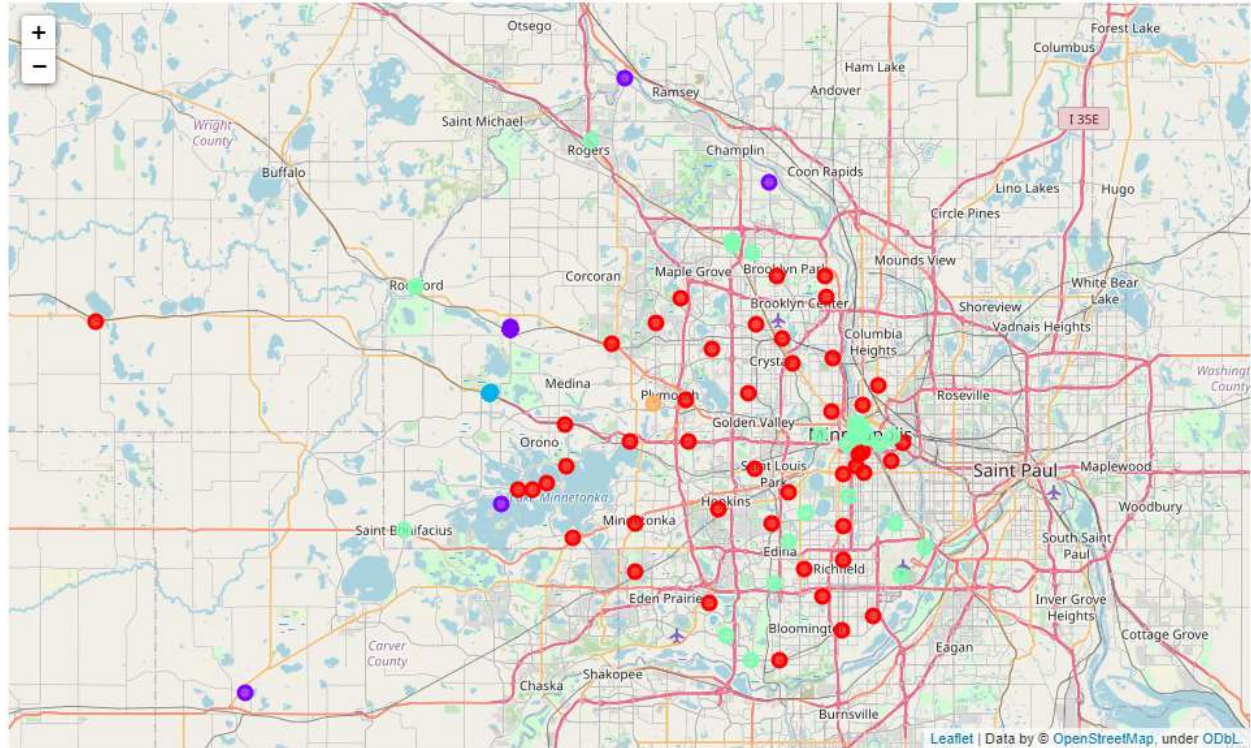
The fifth cluster is as follows:

	City	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
86	Irvine	4.0	Seafood Restaurant	Yoga Studio	Fabric Shop	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Financial or Legal Service	Fish & Chips Shop	Fish Market

This cluster is also only one data point. I will also consider it an outlier because of the "farm" data point.

3.7 Visual Mapping of Hennepin County, MN

The map for Hennepin County was also generated based on the Foursquare API and k-means clustering with a factor of 5. The following map is the result:



3.8 Hennepin County, Cluster 1 (Red)

The first cluster is as follows:

	City	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
1	Hopkins	0	Coffee Shop	Pet Store	Shoe Store	Electronics Store	Grocery Store	Furniture / Home Store	Bank	Italian Restaurant	Big Box Store	Esc
2	Osseo	0	Locksmith	Field	Roof Deck	Yoga Studio	Escape Room	Dutch Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Ev Spi
4	Crystal Bay	0	Food	Boat or Ferry	Post Office	Harbor / Marina	Art Gallery	Event Service	Electronics Store	Empanada Restaurant	English Restaurant	Esci Ro
6	Excelsior	0	Bakery	Coffee Shop	Clothing Store	Bridal Shop	Brewery	Boutique	Spa	Bank	Sushi Restaurant	Music St
7	Hamel	0	Coffee Shop	Salon / Barbershop	American Restaurant	Pizza Place	Automotive Shop	Café	Mobile Phone Shop	Martial Arts School	Food	Big E St
8	Hopkins	0	Park	Rental Car Location	Automotive Shop	Pizza Place	Baseball Field	Fast Food Restaurant	Brazilian Restaurant	Mexican Restaurant	Social Club	Mc Thea
9	Eden Prairie	0	Clothing Store	Sandwich Place	Ice Cream Shop	Bakery	Pizza Place	Chinese Restaurant	Fast Food Restaurant	Kids Store	Coffee Shop	Mol Phc St
10	Minnetonka	0	ATM	Locksmith	Rental Car Location	Furniture / Home Store	Arcade	Clothing Store	American Restaurant	Advertising Agency	Fast Food Restaurant	Farm Mar
11	Eden Prairie	0	Concert Hall	Skating Rink	Yoga Studio	Dry Cleaner	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Exhibit	Ev Spi
14	Long Lake	0	Pizza Place	Shopping Mall	Chinese Restaurant	Salon / Barbershop	Sandwich Place	Jewelry Store	Liquor Store	Spa	Gym / Fitness Center	Cof St
17	Minnetonka Beach	0	Campground	Golf Course	Harbor / Marina	Lake	Yoga Studio	English Restaurant	Dry Cleaner	Dutch Restaurant	Electronics Store	Empana Restaur
22	Spring Park	0	Harbor / Marina	BBQ Joint	Lake	Furniture / Home Store	Seafood Restaurant	Post Office	Trail	English Restaurant	Dutch Restaurant	Electron St
23	Wayzata	0	American Restaurant	Coffee Shop	Clothing Store	Irish Pub	Breakfast Spot	Restaurant	Frozen Yogurt Shop	Dry Cleaner	Bakery	Steakhou
24	Navarre	0	Coffee Shop	Jewelry Store	Pizza Place	Restaurant	Café	Fast Food Restaurant	Shop & Service	Bowling Alley	Liquor Store	Esc
28	Minneapolis	0	Sandwich Place	Flower Shop	Breakfast Spot	Park	Miscellaneous Shop	Other Repair Shop	Grocery Store	Escape Room	Dutch Restaurant	Electron St
30	Minneapolis	0	Bowling Alley	Liquor Store	Spiritual Center	Bakery	Playground	Mexican Restaurant	Latin American Restaurant	Bar	Art Gallery	Ev Serv
31	Minneapolis	0	Mexican Restaurant	Grocery Store	Taco Place	ATM	Non-Profit	Bookstore	Brewery	Shoe Store	Bus Station	C
32	Minneapolis	0	Pizza Place	Bar	Adult Boutique	Grocery Store	Brewery	Taco Place	Cosmetics Shop	Gym / Fitness Center	Salon / Barbershop	Rec St
35	Minneapolis	0	Business Service	Wine Bar	Speakeasy	Brewery	Miscellaneous Shop	Bus Station	Yoga Studio	Escape Room	Empanada Restaurant	Eng Restaur
36	Minneapolis	0	Grocery Store	Pharmacy	Bakery	Convenience Store	Park	Farmers Market	Yoga Studio	Electronics Store	Empanada Restaurant	Eng Restaur
37	Minneapolis	0	Dive Bar	American Restaurant	Yoga Studio	Art Gallery	Pizza Place	Speakeasy	Food Truck	Bookstore	Fish & Chips Shop	Brew
38	Minneapolis	0	Chinese Restaurant	Bus Station	Hotel	Pizza Place	Hockey Arena	Café	Sports Bar	Dessert Shop	Art Gallery	Bus L
40	Minneapolis	0	Grocery Store	Shoe Store	Coffee Shop	Video Store	Liquor Store	Yoga Studio	Kids Store	Shopping Mall	Sandwich Place	F
42	Minneapolis	0	Thai Restaurant	Bakery	Mexican Restaurant	ATM	Bus Station	BBQ Joint	Bank	South American Restaurant	Brewery	Pizza Pla
43	Minneapolis	0	Yoga Studio	Sandwich Place	Chiropractor	Fast Food Restaurant	Gym / Fitness Center	Laundry Service	Liquor Store	Mexican Restaurant	Other Great Outdoors	A

The full data set is not displayed here, but this cluster looks like it captured some of Minneapolis and its surrounding areas. There seems to be a large variety of venues in this cluster alone, mostly restaurants and the like. These can be considered recreational areas that do not include outdoor activities, as we saw in one of the Orange County clusters.

3.9 Hennepin County, Cluster 2 (Purple)

The second cluster is as follows:

	City	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
3	Champlin	1	Baseball Field	Bus Stop	Yoga Studio	Dutch Restaurant	Flower Shop	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Exhibit
5	Dayton	1	Baseball Field	Liquor Store	American Restaurant	Yoga Studio	Empanada Restaurant	English Restaurant	Escape Room	Event Service	Event Space	Exhibit
15	Loretto	1	American Restaurant	Bar	Baseball Field	Construction & Landscaping	Business Service	Caribbean Restaurant	Gym	English Restaurant	Escape Room	Event Service
18	Mound	1	Park	Business Service	Bar	Yoga Studio	Event Service	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Event Space
93	Young America	1	Art Gallery	Bar	Antique Shop	Baseball Field	Yoga Studio	Electronics Store	English Restaurant	Escape Room	Event Service	Event Space
101	Loretto	1	American Restaurant	Bar	Baseball Field	Business Service	Caribbean Restaurant	Furniture / Home Store	Gym	Yoga Studio	English Restaurant	Escape Room
102	Loretto	1	American Restaurant	Bar	Baseball Field	Business Service	Caribbean Restaurant	Furniture / Home Store	Gym	Yoga Studio	English Restaurant	Escape Room
103	Loretto	1	American Restaurant	Bar	Baseball Field	Business Service	Caribbean Restaurant	Furniture / Home Store	Gym	Yoga Studio	English Restaurant	Escape Room
104	Loretto	1	American Restaurant	Bar	Baseball Field	Business Service	Caribbean Restaurant	Furniture / Home Store	Gym	Yoga Studio	English Restaurant	Escape Room
105	Loretto	1	American Restaurant	Bar	Baseball Field	Business Service	Caribbean Restaurant	Furniture / Home Store	Gym	Yoga Studio	English Restaurant	Escape Room

This cluster is much smaller than the first one, and it actually looks like a lot of data points are repeated. This could be because many different zip codes are attributed to one city, and perhaps these particular cities are not very large. This can be the “baseball” cluster, as it looks like it revolves around the baseball field venue.

3.10 Hennepin County, Cluster 3 (Blue)

The third cluster is as follows:

	City	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
13	Maple Plain	2	Café	Construction & Landscaping	Garden	Post Office	Pub	Baseball Field	Escape Room	Electronics Store	Empanada Restaurant	English Restaurant
16	Maple Plain	2	Café	Construction & Landscaping	Garden	Post Office	Pub	Baseball Field	Escape Room	Electronics Store	Empanada Restaurant	English Restaurant
90	Maple Plain	2	Café	Construction & Landscaping	Garden	Post Office	Pub	Baseball Field	Escape Room	Electronics Store	Empanada Restaurant	English Restaurant
91	Maple Plain	2	Café	Construction & Landscaping	Garden	Post Office	Pub	Baseball Field	Escape Room	Electronics Store	Empanada Restaurant	English Restaurant
94	Maple Plain	2	Café	Construction & Landscaping	Garden	Post Office	Pub	Baseball Field	Escape Room	Electronics Store	Empanada Restaurant	English Restaurant
96	Maple Plain	2	Café	Construction & Landscaping	Garden	Post Office	Pub	Baseball Field	Escape Room	Electronics Store	Empanada Restaurant	English Restaurant
98	Maple Plain	2	Café	Construction & Landscaping	Garden	Post Office	Pub	Baseball Field	Escape Room	Electronics Store	Empanada Restaurant	English Restaurant
99	Maple Plain	2	Café	Construction & Landscaping	Garden	Post Office	Pub	Baseball Field	Escape Room	Electronics Store	Empanada Restaurant	English Restaurant
100	Maple Plain	2	Café	Construction & Landscaping	Garden	Post Office	Pub	Baseball Field	Escape Room	Electronics Store	Empanada Restaurant	English Restaurant

Although this cluster technically has many different data points, they all point to the same things. Again, this could be attributed to several zip codes corresponding to one city, and that city not being particularly large. I will have to label this cluster as an outlier, as it only really attributes to one “unique” row.

3.11 Hennepin County, Cluster 4 (Green)

The fourth cluster is as follows:

	City	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
0	Saint Paul	3	Café	Tennis Court	Campground	Light Rail Station	Museum	Athletics & Sports	Bar	English Restaurant	Dutch Restaurant	Electronics Store
12	Eden Prairie	3	Playground	Plaza	Dog Run	Field	Fast Food Restaurant	Farmers Market	Exhibit	Event Space	Event Service	Escape Room
19	Osseo	3	Bar	Coffee Shop	Greek Restaurant	Cajun / Creole Restaurant	Sandwich Place	Sporting Goods Shop	Thai Restaurant	Motorcycle Shop	Liquor Store	Gourmet Shop
20	Rogers	3	Fast Food Restaurant	Sporting Goods Shop	Coffee Shop	Gas Station	Dive Bar	Diner	Motel	Bar	Big Box Store	Restaurant
21	Saint Bonifacius	3	Coffee Shop	Video Store	Dive Bar	Bar	Baseball Field	Sandwich Place	Convenience Store	Park	Farmers Market	Exhibit
25	Minneapolis	3	Bar	Nightclub	Pizza Place	Café	American Restaurant	Yoga Studio	Scandinavian Restaurant	New American Restaurant	Sushi Restaurant	Food Truck
26	Minneapolis	3	Coffee Shop	Hotel	Theater	American Restaurant	Food Truck	Seafood Restaurant	Breakfast Spot	Steakhouse	Salad Place	Restaurant
27	Minneapolis	3	Coffee Shop	Hotel	Theater	New American Restaurant	Salon / Barbershop	Liquor Store	Seafood Restaurant	Gym	Bar	Plaza
29	Minneapolis	3	Coffee Shop	Pizza Place	Brewery	Men's Store	Park	Furniture / Home Store	Grocery Store	Antique Shop	Escape Room	Electronics Store
33	Minneapolis	3	Park	Market	American Restaurant	Butcher	Restaurant	Latin American Restaurant	Diner	Empanada Restaurant	Thrift / Vintage Store	Sushi Restaurant
34	Minneapolis	3	Athletics & Sports	Food Truck	Park	Event Service	Dutch Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Yoga Studio
39	Minneapolis	3	Coffee Shop	Bar	Sandwich Place	American Restaurant	Pub	Japanese Restaurant	Hotel	Italian Restaurant	Music Venue	New American Restaurant
41	Minneapolis	3	Yoga Studio	Park	Mexican Restaurant	Liquor Store	Gas Station	Gift Shop	Boutique	Coffee Shop	Credit Union	Electronics Store
47	Minneapolis	3	Pool	Football Stadium	Moving Target	Yoga Studio	Escape Room	Dutch Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Event Space
58	Minneapolis	3	Coffee Shop	Pizza Place	Golf Course	Automotive Shop	Event Space	Ice Cream Shop	Cafeteria	Burger Joint	Intersection	Gas Station
59	Minneapolis	3	Gym	Sandwich Place	Coffee Shop	Taco Place	Event Service	Music Venue	Bagel Shop	Athletics & Sports	Food Truck	Arts & Entertainment
60	Minneapolis	3	Coffee Shop	Hotel	Burger Joint	New American Restaurant	Salad Place	Bar	Food Truck	Bus Station	Pub	Chinese Restaurant
65	Minneapolis	3	Dry Cleaner	Food Truck	Vietnamese Restaurant	Paper / Office Supplies Store	Thrift / Vintage Store	Escape Room	Dutch Restaurant	Electronics Store	Empanada Restaurant	English Restaurant
68	Minneapolis	3	Rental Car Location	Airport Terminal	Coffee Shop	General Travel	Sandwich Place	Wine Bar	Convenience Store	Latin American Restaurant	Airport	Airport Service
69	Minneapolis	3	Theater	Bar	Sandwich Place	Bus Station	Coffee Shop	Hotel	Dive Bar	Rock Club	Pizza Place	Music Venue
70	Minneapolis	3	Coffee Shop	Pizza Place	College Rec Center	Bowling Alley	Fried Chicken Joint	Korean Restaurant	American Restaurant	Food Court	Bagel Shop	Mexican Restaurant
71	Minneapolis	3	Coffee Shop	Asian Restaurant	Bar	Park	New American Restaurant	Sushi Restaurant	River	Food Truck	Café	Spa
72	Minneapolis	3	Coffee Shop	Hotel	Burger Joint	New American Restaurant	Salad Place	Bar	Food Truck	Bus Station	Pub	Chinese Restaurant

This cluster can be labeled as the heart of Minneapolis. Many of these data points are placed in the center of Minneapolis and include a large variety of restaurants, eateries, and other fun things to do.

3.12 Hennepin County, Cluster 5 (Yellow)

The fifth cluster is as follows:

	City	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
67	Minneapolis	4	Music Venue	Yoga Studio	Hospital	Fish & Chips Shop	Field	Fast Food Restaurant	Farmers Market	Exhibit	Event Space	Event Service

Again, this cluster only contains one unique row so I will have to label it as an outlier.

4. Conclusions

Orange County and Hennepin County share more similarities than I had first assumed, based on the two large clusters for each of the corresponding counties. The city of Minneapolis seems to share similar venues as the city of Anaheim, which would make it seem like a good candidate for a Disney theme park. However, one must keep in mind that only 2 out of the 5 clusters for each county seemed to share similar attributes. Each county had 2 outliers that were not very similar. One must also consider the fact that Anaheim and its surrounding areas were in a way built around Disneyland. Minneapolis and its surrounding areas have all of those fun venues, restaurants, and eateries without the presence of a large theme park.

Ultimately my recommendation to proceed with Hennepin County as a host for a Disney theme park would be... to investigate Hennepin County further. There are many other factors that should be considered that are not captured in this initial set of data, which are discussed in the following section.

5. Future Directions

In the future, the first think I would recommend is to cluster the data differently. I aimed to keep things consistent by using a k-means factor of 5 for both counties, but it might be useful to find the optimal k-means factor for both counties, even if it means sacrificing some of the consistency (for example, if Orange County would be better clustered by a factor of 7 and Hennepin County by a factor of 3).

Something to keep in mind about this data is that it is coming from the Foursquare API, which for the most part is "crowdsourced," meaning that it depends heavily on user input. The data could be biased because people would be more inclined to report or "check-in" at restaurants and other fun things to do versus other buildings and venues such as hospital, churches, and government buildings.

I think that the people living in those areas are an important factor in making this decision, and this initial dataset does not capture very much of who is living there. Assumptions can be made, but it would be helpful to know things such as income and age of the individuals there. It would paint a much clearer picture as to who is living in those cities and who is going to those venues.