

# Similarity of the top 100 Cities in US based on their venues styles

Zimu Guo

March 2019

## 1 Induction

### 1.1 Background

The United States is a diverse country where you can find almost all possible kinds of entertainments. Therefore, traveller may find that exploring different categories of venues in US is fascinating. Someone may like to know the best destination where the venues would fit one's appetite. So, we can give them suggestions by providing the similarity score between cities based on their venues – if you love to stay in one city, then you should go to the other and have a try.

### 1.2 Problem

Data that might contribute to determining similarity between cities based on the categories of venues: How many venues are there in one city? What is the ratio of the number of venues in one style compare to the number of all venues? The project aim to divide the top 100 cities (based on population) in several groups based on data. If one loves feels good in a city, then he or she would also enjoy his or her life in the cities which are in the same group of the previous one.

## 2 Data

### 2.1 Necessary Data

1. The top 100 cities in US.
2. The venues in each city corresponding to their categories.

As we know the cities we are trying to explore, we can compare their features (venues' categories) and divide them in groups.

## 2.2 Data Source

1. Wikipedia: "List of United States cities by population"

[https://en.wikipedia.org/wiki/List\\_of\\_United\\_States\\_cities\\_by\\_population](https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population)

2. Data on FOURSQUARE, where we can find venues in each city and the information of each venue included it categories.

## 3 Data Analysis

### 3.1 Get the top 100 cities - Retrieve data from Wikipedia

Here is the table showed on the wikipedia website:

2017 rank	City	State <sup>[c]</sup>	2017 estimate	2010 Census	Change	2016 land area	2016 population density	Location
1	New York City <sup>[d]</sup>	New York	8,622,698	8,175,133	+5.47%	301.5 sq mi	780.9 km <sup>2</sup>	28,317/sq mi
2	Los Angeles	California	3,999,759	3,792,621	+5.46%	468.7 sq mi	1,213.9 km <sup>2</sup>	8,484/sq mi
3	Chicago	Illinois	2,716,450	2,695,598	+0.77%	227.3 sq mi	588.7 km <sup>2</sup>	11,900/sq mi
4	Houston <sup>[3]</sup>	Texas	2,312,717	2,100,263	+10.12%	637.5 sq mi	1,651.1 km <sup>2</sup>	3,613/sq mi
5	Phoenix	Arizona	1,626,078	1,445,632	+12.48%	517.6 sq mi	1,340.6 km <sup>2</sup>	3,120/sq mi
6	Philadelphia <sup>[e]</sup>	Pennsylvania	1,580,863	1,526,006	+3.59%	134.2 sq mi	347.6 km <sup>2</sup>	11,683/sq mi
7	San Antonio	Texas	1,511,946	1,327,407	+13.90%	461.0 sq mi	1,194.0 km <sup>2</sup>	3,238/sq mi
8	San Diego	California	1,419,516	1,307,402	+8.58%	325.2 sq mi	842.3 km <sup>2</sup>	4,325/sq mi
9	Dallas	Texas	1,341,075	1,197,816	+11.96%	340.9 sq mi	882.9 km <sup>2</sup>	3,866/sq mi
10	San Jose	California	1,035,317	945,942	+9.45%	177.5 sq mi	459.7 km <sup>2</sup>	5,777/sq mi
11	Austin	Texas	950,715	790,390	+20.28%	312.7 sq mi	809.9 km <sup>2</sup>	3,031/sq mi
12	Jacksonville <sup>[f]</sup>	Florida	892,062	821,784	+8.55%	747.4 sq mi	1,935.8 km <sup>2</sup>	1,170/km <sup>2</sup>
13	San Francisco <sup>[g]</sup>	California	884,363	805,235	+9.83%	46.9 sq mi	121.5 km <sup>2</sup>	18,569/sq mi
14	Columbus	Ohio	879,170	787,033	+11.71%	218.5 sq mi	565.9 km <sup>2</sup>	3,936/sq mi
15	Fort Worth	Texas	874,168	741,206	+17.94%	342.9 sq mi	888.1 km <sup>2</sup>	2,491/sq mi
16	Indianapolis <sup>[h]</sup>	Indiana	863,002	820,445	+5.19%	361.5 sq mi	936.3 km <sup>2</sup>	2,366/sq mi
17	Charlotte	North Carolina	859,035	731,424	+17.45%	305.4 sq mi	791.0 km <sup>2</sup>	2,757/sq mi
18	Seattle	Washington	724,745	608,660	+19.07%	83.8 sq mi	217.0 km <sup>2</sup>	8,405/sq mi
19	Denver <sup>[i]</sup>	Colorado	704,621	600,158	+17.41%	153.3 sq mi	397.0 km <sup>2</sup>	4,521/sq mi
20	Washington, D.C. <sup>[j]</sup>	District of Columbia	693,972	601,723	+15.33%	61.1 sq mi	158.2 km <sup>2</sup>	11,148/sq mi
21	Boston	Massachusetts	685,094	617,594	+10.93%	48.3 sq mi	125.1 km <sup>2</sup>	13,938/sq mi
22	El Paso	Texas	683,577	649,121	+5.31%	256.8 sq mi	665.1 km <sup>2</sup>	2,660/sq mi
23	Detroit	Michigan	673,104	713,777	-5.70%	138.8 sq mi	359.5 km <sup>2</sup>	4,847/sq mi
24	Nashville <sup>[k]</sup>	Tennessee	667,560	601,222	+11.03%	475.9 sq mi	1,232.6 km <sup>2</sup>	1,388/sq mi
25	Memphis	Tennessee	652,236	646,889	+0.83%	317.4 sq mi	822.1 km <sup>2</sup>	2,056/sq mi
26	Portland	Oregon	647,805	583,776	+10.97%	133.5 sq mi	345.8 km <sup>2</sup>	4,793/sq mi
27	Oklahoma City	Oklahoma	643,648	579,999	+10.97%	606.3 sq mi	1,570.3 km <sup>2</sup>	1,053/sq mi
								407/km <sup>2</sup>

Figure 1: List of United States cities by population (Wikipedia)

Use **BeautifulSoup** to retrieve data from the website. Parse the data from columns: 2017 rank, City and Location and put the results into the DataFrame `city_data` with columns: index column(2017 rank), City, Latitude, Longitude.

The DataFrame *city\_data* is shown below:

`city_data.head()`

	<b>City</b>	<b>Latitude</b>	<b>Longitude</b>
<b>0</b>	New York City	40.6635	-73.9387
<b>1</b>	Los Angeles	34.0194	-118.4108
<b>2</b>	Chicago	41.8376	-87.6818
<b>3</b>	Houston	29.7866	-95.3909
<b>4</b>	Phoenix	33.5722	-112.0901

Figure 2: DataFrame: *city\_data*

Find venues for each city and their names, latitudes, longitudes and their categories from FOURSQUARE. Parse the data and store them into the DataFrame *near\_venues* concatenated with the previous DataFrame *city\_data*

The DataFrame *nearby\_venues* is shown below:

`nearby_venues.head()`

	<b>City</b>	<b>City Latitude</b>	<b>City Longitude</b>	<b>Venue</b>	<b>Venue Latitude</b>	<b>Venue Longitude</b>	<b>Venue Category</b>
<b>0</b>	New York City	40.6635	-73.9387	Izzy's Brooklyn Smokehouse	40.664869	-73.937023	BBQ Joint
<b>1</b>	New York City	40.6635	-73.9387	Ali's Roti Shop	40.666436	-73.931346	Caribbean Restaurant
<b>2</b>	New York City	40.6635	-73.9387	Sweet Expressions	40.668382	-73.942337	Candy Store
<b>3</b>	New York City	40.6635	-73.9387	Bodega	40.668388	-73.932366	Deli / Bodega
<b>4</b>	New York City	40.6635	-73.9387	The Market Place	40.662153	-73.943010	Grocery Store

Figure 3: DataFrame: *nearby\_venues*

Find the 10 most common categories of venues in each city by first convert the categories as columns name and then use 0 or 1 to indicate the category is there or not. Use `get_dummies` from pandas to convert the *nearby\_venues* to DataFrame *city\_onehot*

The DataFrame `city_onehot` is shown below:

	city_onehot.head()														
	City	ATM	Accessories Store	Afghan Restaurant	African Restaurant	Airport	Airport Service	Airport Terminal	American Restaurant	Animal Shelter	Antique Shop	Arcade	Argen Resta	Armenian	Asian Restaurant
0	New York City	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	New York City	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	New York City	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	New York City	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	New York City	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 4: DataFrame: `city_onehot`

For each city, we first find the 10 most common categories of venues and use the information to do further clustering. The data was stored in DataFrame `city_venues_sorted`:

	city_venues_sorted.head()									
	City	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
0	Albuquerque	Mexican Restaurant	Brewery	Construction & Landscaping	Discount Store	Intersection	Event Service	Lawyer	Gas Station	Sandwich Place
1	Anaheim	Japanese Restaurant	Dance Studio	Home Service	Thai Restaurant	Burger Joint	Sushi Restaurant	Mediterranean Restaurant	Sandwich Place	Gym / Fitness Center
2	Arlington	Mexican Restaurant	Breakfast Spot	Rental Car Location	Donut Shop	Burger Joint	Gym	Smoke Shop	Martial Arts Dojo	Food Stand
3	Atlanta	Convenience Store	Park	Gay Bar	Pool	Wings Joint	Light Rail Station	Food	Tennis Court	Gas Station
4	Austin	Park	Coffee Shop	Sandwich Place	American Restaurant	Rental Car Location	Men's Store	Gift Shop	Italian Restaurant	Grocery Store

Figure 5: DataFrame: `city_venues_sorted`

Then, we are able to do the clustering using **K-mean**. As we successfully label all city, we put the latitude and longitude back. Create DataFrame `city_merge` contains city, latitude, longitude and it's 10 most common categories of venues. Then, we use `city_merge` to draw a map, with point on the position of each city – different colors of points represent different cluster.

The DataFrame `city_merge` is shown below:

	City	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
0	New York City	40.6635	-73.9387	0.0	Pizza Place	Grocery Store	Caribbean Restaurant	Deli / Bodega	Sandwich Place	Pharmacy	Donut Shop
1	Los Angeles	34.0194	-118.4108	0.0	Indian Restaurant	Convenience Store	Taco Place	Bakery	Thai Restaurant	Sushi Restaurant	Coffee Shop
2	Chicago	41.8376	-87.6818	0.0	Grocery Store	Diner	Coffee Shop	Fast Food Restaurant	Public Art	Bus Stop	Bakery
3	Houston	29.7866	-95.3909	0.0	Mexican Restaurant	Coffee Shop	Café	Ice Cream Shop	Food Truck	Gym / Fitness Center	American Restaurant
4	Phoenix	33.5722	-112.0901	0.0	Fast Food Restaurant	Food Truck	Sandwich Place	Nail Salon	Bar	Video Store	Pharmacy

Figure 6: DataFrame: `city_merge`

And here is the example result of clustering the cities into 10 clusters:

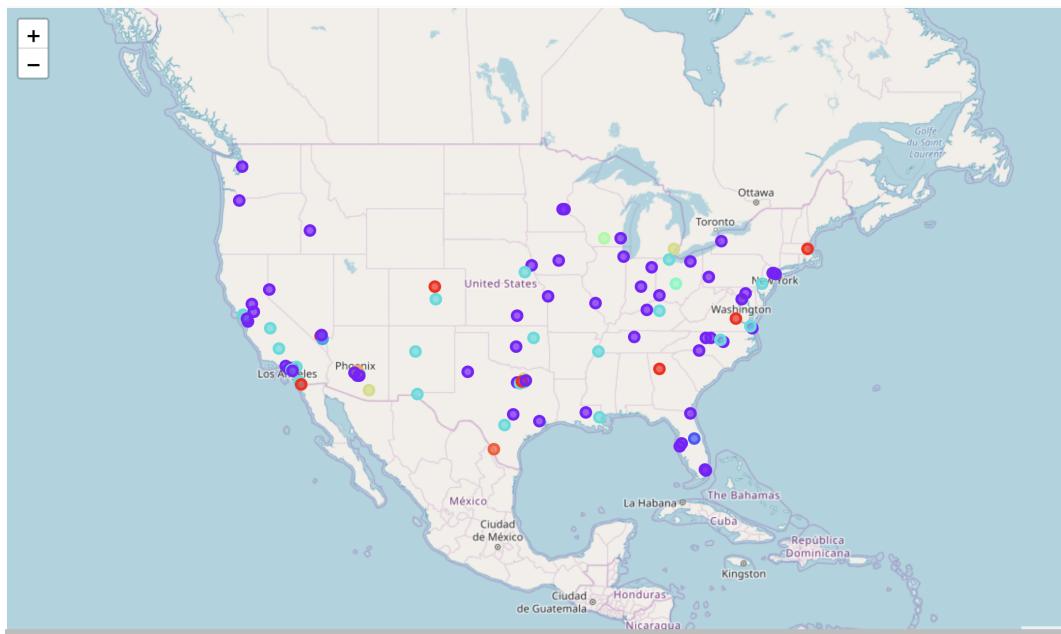


Figure 7: Map of 10 clusters

## 4 Result

### 4.1 Map for 2 Clusters - 10 Clusters

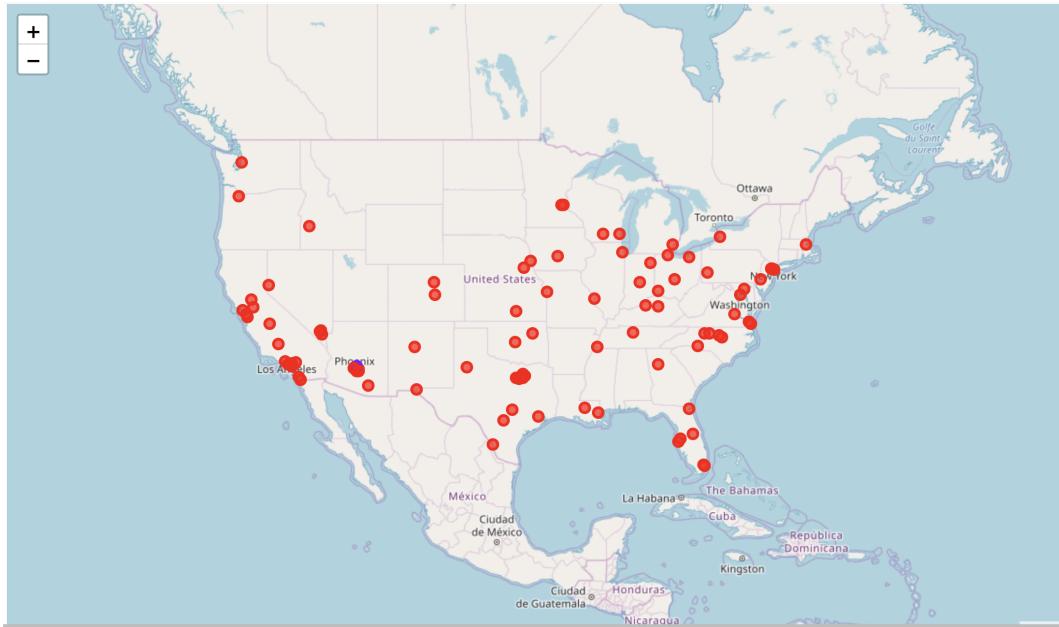


Figure 8: Map of 2 clusters

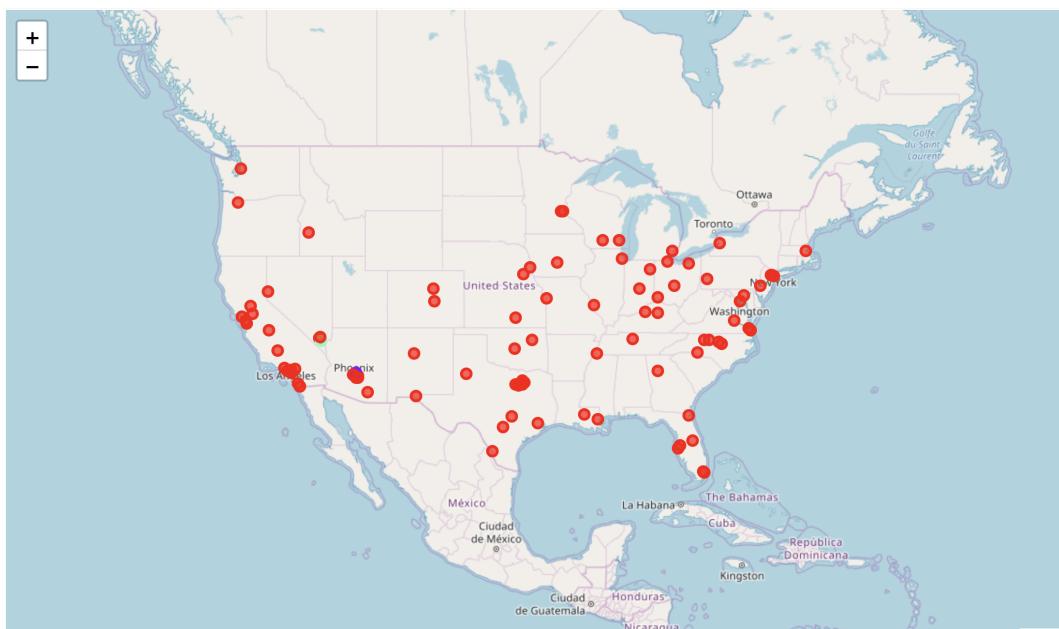


Figure 9: Map of 3 clusters

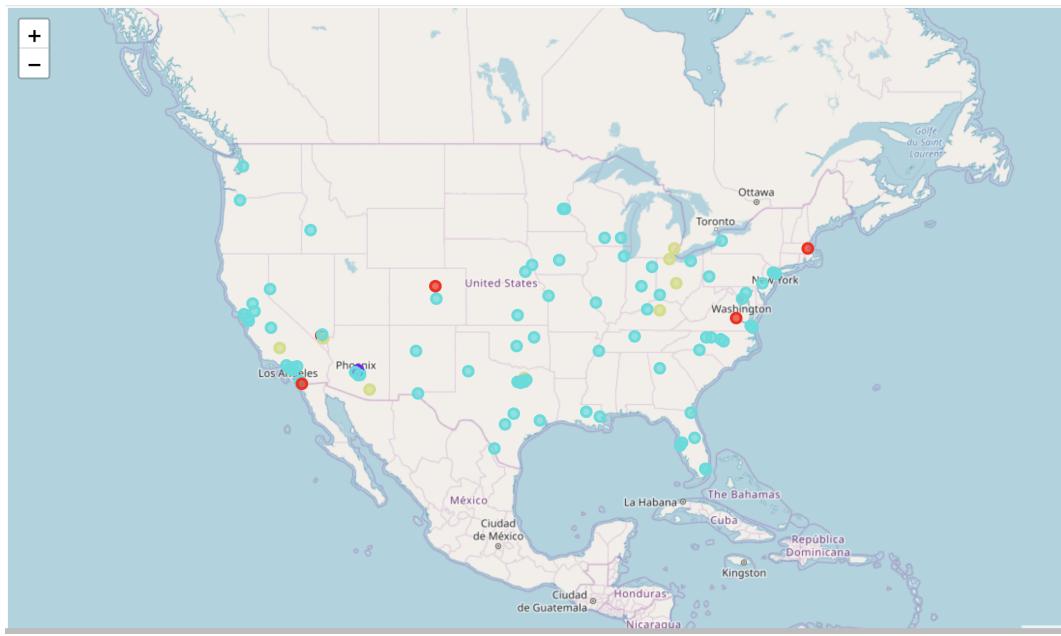


Figure 10: Map of 4 clusters

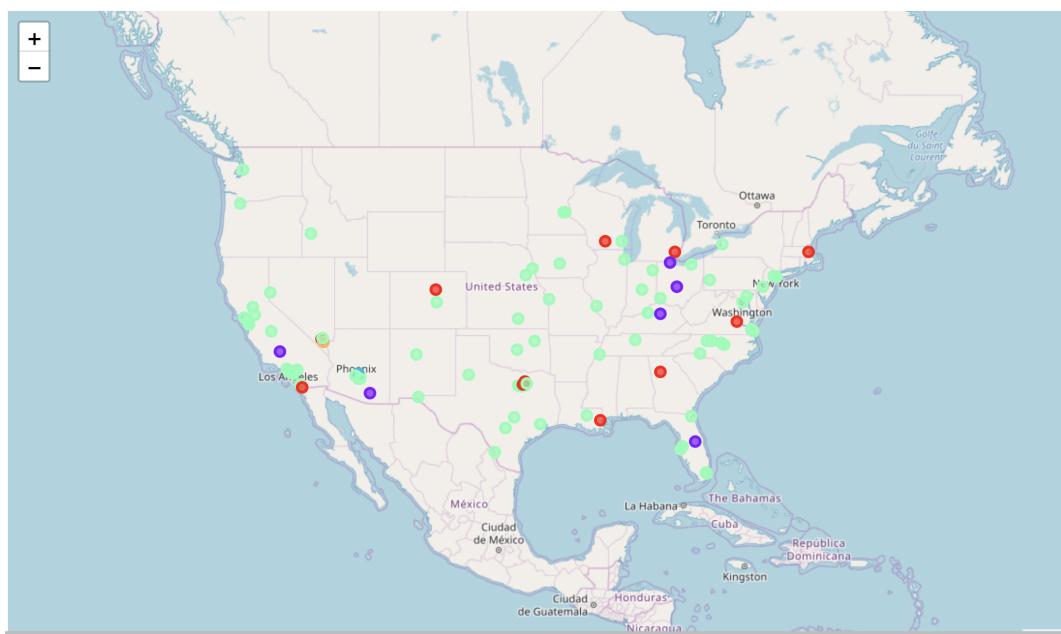


Figure 11: Map of 5 clusters

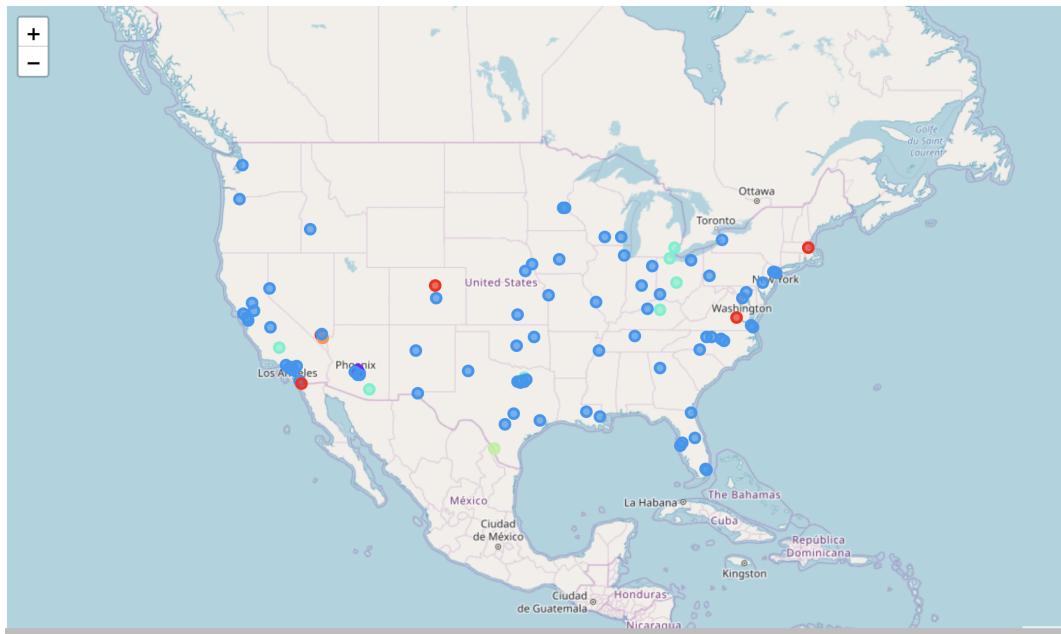


Figure 12: Map of 6 clusters

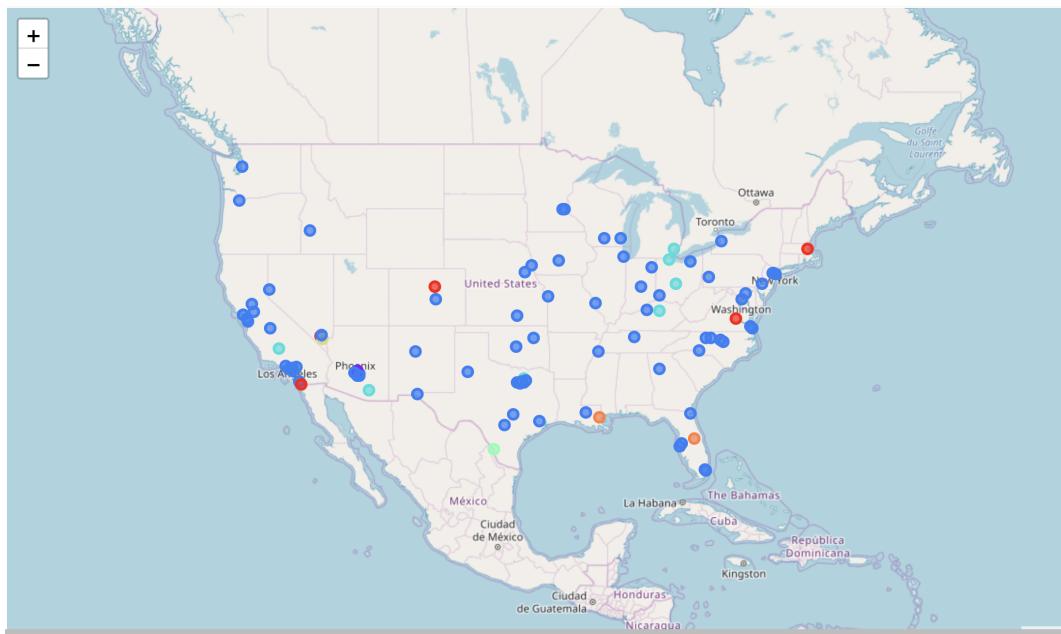


Figure 13: Map of 7 clusters

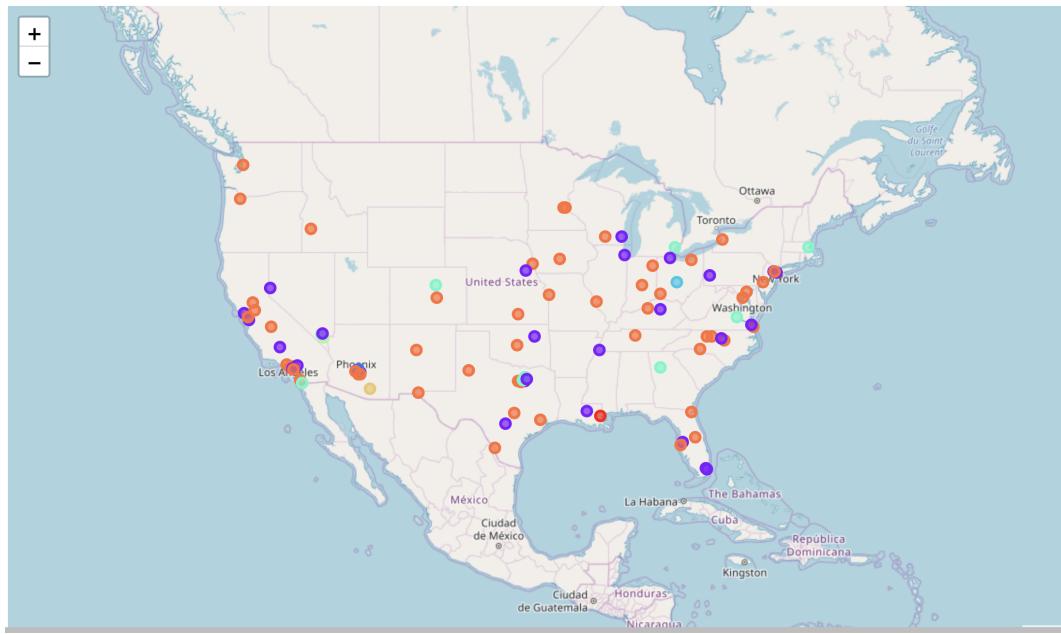


Figure 14: Map of 8 clusters

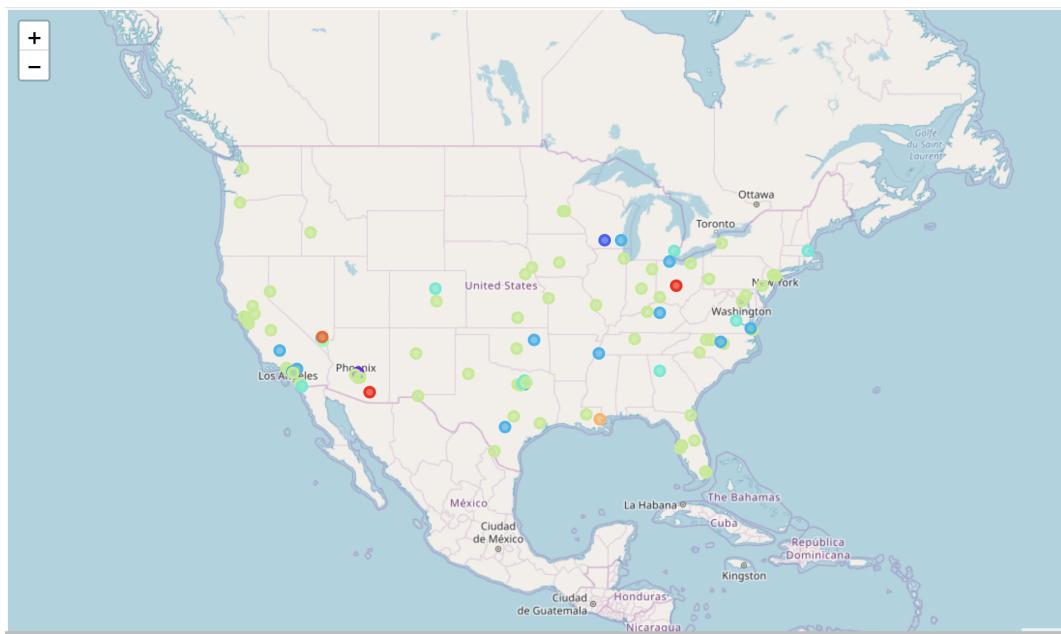


Figure 15: Map of 9 clusters

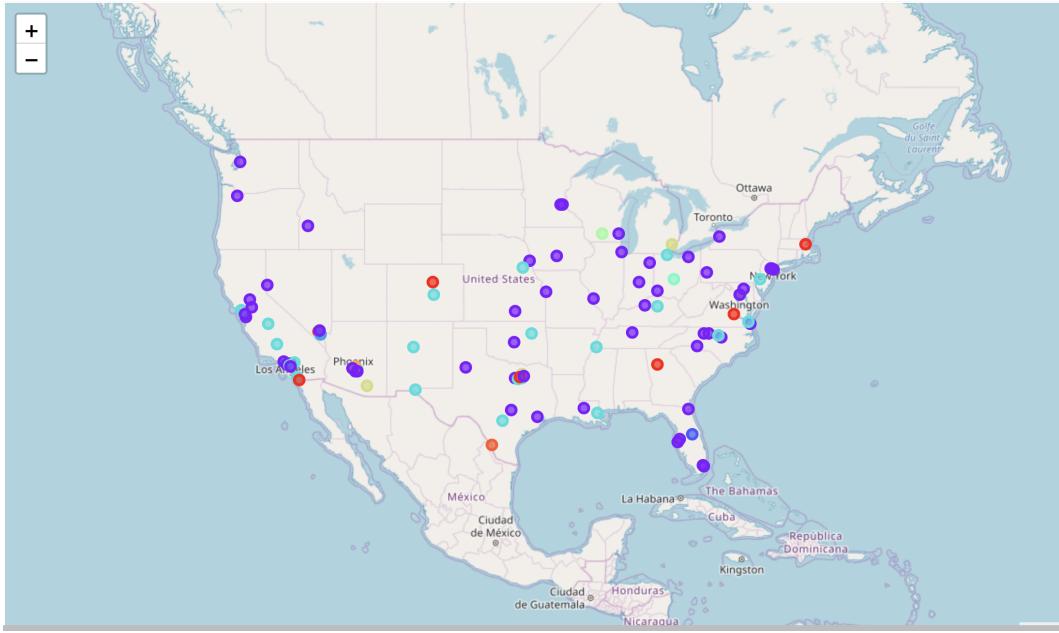


Figure 16: Map of 10 clusters

## 5 Discussion

### 5.1 Observation on the maps

First of all, we didn't see any boundary such as splitting north and south, or east and west whatever the value of  $k$  is. Secondly, we didn't find same color of points are intended to be closer to each other. Therefore, the points are distributed more like randomly, which means we can neither say that you will like the city nearby your favourite region, nor the west coast or the east area would be your best choice.

## 5.2 Detailed discussion on 10 clusters ( $k = 10$ )

Let's recall the map of 10 clusters:

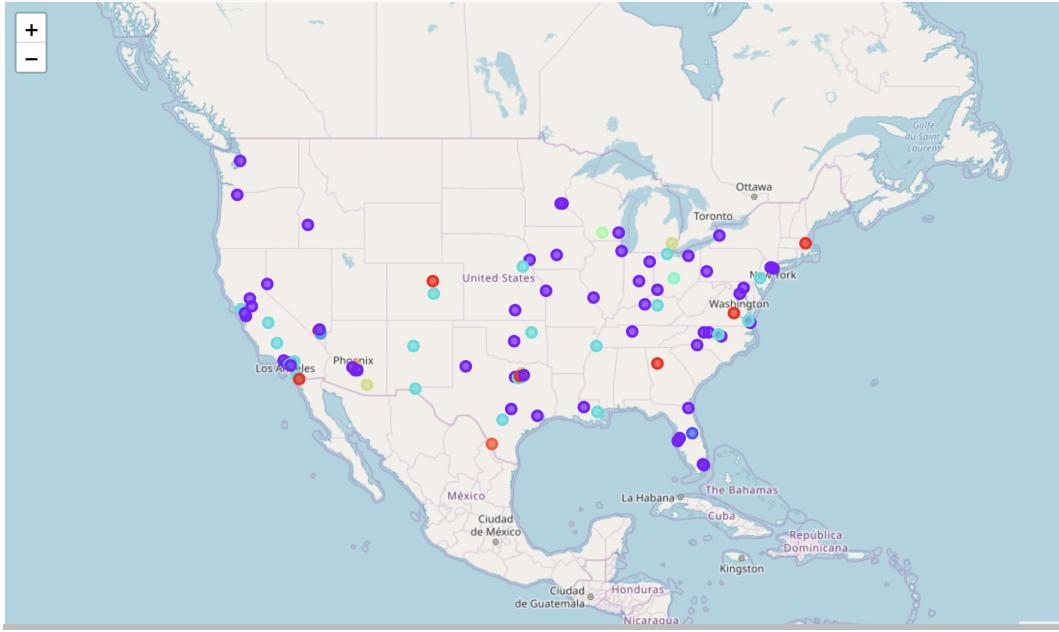


Figure 17: Map of 10 clusters

It contains points of following 10 colors: red, purple, dark blue, blue, light blue, blue-green, green, yellow, yellow-orange, orange-red

1. The first cluster with **RED** points contains **8** cities:

Cluster 1												
	City	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	
18	Denver	Recreation Center	Park	Pool	Restaurant	Zoo Exhibit	Farm	Event Service	Event Space	Exhibit	Ey	
20	Boston	Park	Beach	Baseball Field	Hot Dog Joint	Lighthouse	History Museum	Historic Site	American Restaurant	Trail	Pl	
27	Las Vegas	Park	Trail	Fast Food Restaurant	Skate Park	Convenience Store	Taco Place	Farm	Event Space	Exhibit	Ey	
37	Atlanta	Convenience Store	Park	Gay Bar	Pool	Wings Joint	Light Rail Station	Food	Tennis Court	Gas Station	Ni	
55	Honolulu	Park	Event Service	Sushi Restaurant	Martial Arts Dojo	Bus Stop	Cafeteria	Theater	Chinese Restaurant	Astrologer	Ve / v Re	
75	Chula Vista	Park	Outdoor Gym	Dance Studio	Intersection	Fast Food Restaurant	Event Space	Exhibit	Eye Doctor	Fabric Shop	Fa	
92	Irving	Baseball Field	Ice Cream Shop	Home Service	Convenience Store	Park	Donut Shop	Athletics & Sports	Sandwich Place	Farmers Market	Ex	
96	Richmond	Park	Beach	Trail	Garden	Forest	Bridge	Farm	Museum	Lake	Ge Ce	

Figure 18: Cluster 1: cities with red points

2. The second cluster with **PURPLE** points, which is the largest cluster among the 10, contains **55** cities:

Cluster 2									
	City	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
0	New York City	Pizza Place	Grocery Store	Caribbean Restaurant	Deli / Bodega	Sandwich Place	Pharmacy	Donut Shop	Fast Food Restaurant
1	Los Angeles	Indian Restaurant	Convenience Store	Taco Place	Bakery	Thai Restaurant	Sushi Restaurant	Coffee Shop	Sandwich Place
2	Chicago	Grocery Store	Diner	Coffee Shop	Fast Food Restaurant	Public Art	Bus Stop	Bakery	Taco Place
3	Houston	Mexican Restaurant	Coffee Shop	Café	Ice Cream Shop	Food Truck	Gym / Fitness Center	American Restaurant	Bar
9	San Jose	Fast Food Restaurant	Playground	Fried Chicken Joint	Golf Course	Sandwich Place	Bubble Tea Shop	Caribbean Restaurant	Shipping Store
10	Austin	Park	Coffee Shop	Sandwich Place	American Restaurant	Rental Car Location	Men's Store	Gift Shop	Italian Restaurant
11	Jacksonville	Sandwich Place	Park	Deli / Bodega	Art Gallery	Coffee Shop	Convenience Store	Pub	Restaurant
14	Fort Worth	Bar	Mexican Restaurant	Steakhouse	Hotel	Historic Site	Burger Joint	Tourist Information Center	Theater
15	Indianapolis	Bar	Brewery	Pub	American Restaurant	Gastropub	Coffee Shop	Theater	New American Restaurant
16	Charlotte	Pizza Place	Coffee Shop	Ice Cream Shop	American Restaurant	Fast Food Restaurant	Bakery	Sandwich Place	Bed & Breakfast
17	Seattle	Exhibit	Coffee Shop	Theater	Sculpture Garden	Bar	Mexican Restaurant	Plaza	Vietnamese Restaurant
19	Washington, D.C.	Coffee Shop	American Restaurant	Bar	Pizza Place	New American Restaurant	Cocktail Bar	Bakery	Italian Restaurant
23	Nashville	Hotel	Restaurant	Coffee Shop	Bar	Sandwich Place	Pizza Place	Music Venue	Museum
25	Portland	Sandwich Place	Coffee Shop	Cocktail Bar	Mexican Restaurant	Sushi Restaurant	Bar	Thai Restaurant	Bed & Breakfast
26	Oklahoma City	Hotel	Sandwich Place	American Restaurant	Coffee Shop	Restaurant	Bar	Steakhouse	Burger Joint
28	Louisville	IT Services	Movie Theater	Café	Business Service	Playground	Video Store	Flower Shop	Flea Market
29	Baltimore	Coffee Shop	Indian Restaurant	Restaurant	Gay Bar	Bar	Italian Restaurant	Vegetarian / Vegan Restaurant	Theater
30	Milwaukee	Grocery Store	Bar	Video Store	Cosmetics Shop	Convenience Store	Sandwich Place	Deli / Bodega	Other Great Outdoors
34	Sacramento	Coffee Shop	Mexican Restaurant	Dive Bar	Gym	Salon / Barbershop	Vietnamese Restaurant	New American Restaurant	Bar
36	Kansas City, MO	Bar	Gym	Theater	Music Venue	Zoo Exhibit	Fast Food Restaurant	Eye Doctor	Fabric Shop
38	Long Beach	Airport	Deli / Bodega	Airport Terminal	Gym	Rental Car Location	Lawyer	Shipping Store	Furniture / Home Store
39	Omaha	Hotel	Park	Breakfast Spot	Arts & Crafts Store	Coffee Shop	Gym / Fitness Center	Intersection	Pizza Place
40	Raleigh	Bar	Trail	Mexican Restaurant	Park	Italian Restaurant	Steakhouse	Spa	Sushi Restaurant
42	Miami	Pharmacy	Pizza Place	Seafood Restaurant	Latin American Restaurant	Fish Market	Hotel	Spanish Restaurant	Convenience Store
43	Virginia Beach	Electronics Store	Gym	Athletics & Sports	Coffee Shop	General Entertainment	Event Service	Exhibit	Eye Doctor
45	Minneapolis	Coffee Shop	Theater	Hotel	Park	Diner	Grocery Store	Miscellaneous Shop	Thrift / Vintage Store
49	Wichita	Sandwich Place	Park	Mexican Restaurant	Pizza Place	Gym	American Restaurant	Hotel	Flower Shop
50	Cleveland	Bar	Pub	Coffee Shop	Park	Art Gallery	Food Truck	Pizza Place	Ice Cream Shop
51	Tampa	Cuban Restaurant	Italian Restaurant	Flower Shop	Pawn Shop	Shoe Store	Baseball Field	Chinese Restaurant	Seafood Restaurant
54	Anaheim	Japanese Restaurant	Dance Studio	Home Service	Thai Restaurant	Burger Joint	Sushi Restaurant	Mediterranean Restaurant	Sandwich Place
60	Stockton	Pool	Video Store	Noodle House	Chinese Restaurant	Smoke Shop	College Arts Building	Athletics & Sports	Grocery Store
61	St. Louis	Sandwich Place	Coffee Shop	Pizza Place	Fast Food Restaurant	Gym	Theater	Gay Bar	Furniture / Home Store
62	Saint Paul	Coffee Shop	Sandwich Place	American Restaurant	Pizza Place	Restaurant	Hotel	Italian Restaurant	Bar
64	Pittsburgh	Pizza Place	Park	Gym / Fitness Center	Cosmetics Shop	Grocery Store	Rental Car Location	Bank	College Basketball Court
65	Cincinnati	Zoo Exhibit	Coffee Shop	Pizza Place	Ice Cream Shop	American Restaurant	Bakery	Sandwich Place	Park
67	Greensboro	Spa	Cosmetics Shop	Coffee Shop	Steakhouse	Burger Joint	Kids Store	Mobile Phone Shop	Hotel
69	Newark	Donut Shop	Italian Restaurant	Pharmacy	Seafood Restaurant	Lounge	Fast Food Restaurant	BBQ Joint	Food
72	Irvine	Pizza Place	Locksmith	Chinese Restaurant	Japanese Restaurant	Dog Run	Donut Shop	Golf Driving Range	Golf Course
74	Jersey City	Science Museum	Fast Food Restaurant	Park	Coffee Shop	Skate Park	Garden Center	Latin American Restaurant	Bar
77	Fort Wayne	Park	Bar	Theater	Pizza Place	American Restaurant	Food Truck	Mexican Restaurant	Cajun / Creole Restaurant
78	St. Petersburg	Park	Bar	Coffee Shop	Restaurant	Fast Food Restaurant	Performing Arts Venue	Harbor / Marina	Café
80	Buffalo	Intersection	Pharmacy	Bar	Clothing Store	Jazz Club	Restaurant	Donut Shop	German Restaurant
82	Lubbock	Pharmacy	BBQ Joint	Grocery Store	Bakery	Construction & Landscaping	Burger Joint	Mobile Phone Shop	Bookstore
83	Chandler	Coffee Shop	Gift Shop	Pharmacy	Skate Park	Donut Shop	Grocery Store	Gym	Mediterranean Restaurant
85	Reno	Business Service	Bar	Trail	American Restaurant	Pizza Place	Sandwich Place	Fast Food Restaurant	Eye Doctor
86	Glendale	Home Service	Mexican Restaurant	Bar	Gas Station	Coffee Shop	American Restaurant	Bakery	German Restaurant
88	Winston-Salem	Bar	American Restaurant	Park	Salon / Barbershop	Coffee Shop	Italian Restaurant	Food Truck	Taco Place
89	North Las Vegas	Gym / Fitness Center	Pool	American Restaurant	Cafeteria	Zoo Exhibit	Field	Eye Doctor	Fabric Shop
90	Gilbert	Mexican Restaurant	Hotel	Burger Joint	Clothing Store	Electronics Store	Lingerie Store	Steakhouse	Snack Place
93	Hialeah	Pizza Place	Sandwich Place	Fast Food Restaurant	Bank	American Restaurant	Latin American Restaurant	Construction & Landscaping	Video Game Store
94	Garland	BBQ Joint	Bakery	Theater	Furniture / Home Store	Fast Food Restaurant	Mexican Restaurant	Thrift / Vintage Store	Light Rail Station
95	Fremont	Hotel	Chinese Restaurant	American Restaurant	Vietnamese Restaurant	Gym	Sandwich Place	Bubble Tea Shop	Mexican Restaurant
97	Boise	French Restaurant	Business Service	Bakery	Lawyer	Coffee Shop	Park	Arts & Crafts Store	Dog Run
98	Baton Rouge	Fast Food Restaurant	Greek Restaurant	Pizza Place	Italian Restaurant	Seafood Restaurant	Fried Chicken Joint	Gas Station	Bank
99	Des Moines	Italian Restaurant	Deli / Bodega	Taco Place	Baseball Stadium	Chinese Restaurant	Bar	Grocery Store	Austrian Restaurant

Figure 19: Cluster 2: cities with purple points

3. The third cluster with **DARK BLUE** points contains only **1** city Orlando:

**Cluster 3**

```
city_merge.loc[city_merge['Cluster Labels'] == 2.0, city_merge.columns[0] + list(range(4, city_merge.shape[1]))]
```

	City	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
71	Orlando	Movie Theater	Gym	Lake	Furniture / Home Store	Zoo Exhibit	Fast Food Restaurant	Exhibit	Eye Doctor	Fabric Shop	Factory

Figure 20: Cluster 3: cities with dark blue points

4. The fourth cluster with **BLUE** points contains only **1** city Henderson:

**Cluster 4**

```
city_merge.loc[city_merge['Cluster Labels'] == 3.0, city_merge.columns[0] + list(range(4, city_merge.shape[1]))]
```

	City	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
63	Henderson	Auto Workshop	Golf Course	Event Service	Exhibit	Eye Doctor	Fabric Shop	Factory	Falafel Restaurant	Farm	Farmers Market

Figure 21: Cluster 4: cities with blue points

5. The fifth cluster with **LIGHT BLUE** points, which is the second largest cluster among the 10, contains **23** cities:

**Cluster 5**

```
city_merge.loc[city_merge['Cluster Labels'] == 4.0, city_merge.columns[0] + list(range(4, city_merge.shape[1]))]
```

41	Colorado Springs	Bridal Shop	Intersection	Gas Station	Fast Food Restaurant	Park	Athletics & Sports	Dog Run	Italian Restaurant	Bo
44	Oakland	Mexican Restaurant	Pizza Place	Grocery Store	Coffee Shop	Fast Food Restaurant	Food Truck	Clothing Store	Gym / Fitness Center	Gy
46	Tulsa	Convenience Store	Dive Bar	Shoe Store	Laser Tag	Liquor Store	Gas Station	Paper / Office Supplies Store	Big Box Store	Ph
47	Arlington	Mexican Restaurant	Breakfast Spot	Rental Car Location	Donut Shop	Burger Joint	Gym	Smoke Shop	Martial Arts Dojo	Fo
48	New Orleans	Theme Park	Moving Target	Intersection	Zoo Exhibit	Field	Eye Doctor	Fabric Shop	Factory	Fal Re
52	Bakersfield	Fast Food Restaurant	Pizza Place	Bank	Grocery Store	Mexican Restaurant	Buffet	Business Service	Bakery	Cic Stc
56	Santa Ana	Mexican Restaurant	Thai Restaurant	Playground	Donut Shop	Liquor Store	Intersection	Grocery Store	Pizza Place	Co & Lar
57	Riverside	Fast Food Restaurant	Liquor Store	Park	Mexican Restaurant	Convenience Store	Video Store	Comic Shop	Coffee Shop	Nal
59	Lexington	Fast Food Restaurant	Gun Shop	Rental Car Location	ATM	Pharmacy	Discount Store	Sandwich Place	Seafood Restaurant	Sh Ma
70	Lincoln	Mexican Restaurant	Discount Store	Park	Coffee Shop	Sandwich Place	Intersection	Social Club	Lingerie Store	Gai
73	Toledo	Fast Food Restaurant	Cosmetics Shop	Grocery Store	Nightclub	Pharmacy	Bank	Discount Store	Disc Golf	Ele Stc
76	Durham	Grocery Store	Trail	Fried Chicken Joint	Intersection	Seafood Restaurant	Sandwich Place	Café	Park	Dis Stc
87	Norfolk	Convenience Store	Mexican Restaurant	Intersection	Flea Market	Video Store	Furniture / Home Store	Fried Chicken Joint	Gym	Dis Stc

Figure 22: Cluster 5: cities with light blue points

6. The sixth cluster with **BLUE-GREEN** points contains only **1** city Columbus:

**Cluster 6**

```
city_merge.loc[city_merge['Cluster Labels'] == 5.0, city_merge.columns[0] + list(range(4, city_merge.shape[1]))]
```

City	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 Columbus	Fried Chicken Joint	Fast Food Restaurant	Intersection	Zoo Exhibit	Field	Exhibit	Eye Doctor	Fabric Shop	Factory	Falafel Restaurant

Figure 23: Cluster 6: cities with blue-green points

7. The seventh cluster with **GREEN** points contains also only **1** city Madison:

**Cluster 7**

```
city_merge.loc[city_merge['Cluster Labels'] == 6.0, city_merge.columns[0] + list(range(4, city_merge.shape[1]))]
```

City	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
81 Madison	Trail	Stadium	Baseball Field	Garden	Fast Food Restaurant	Exhibit	Eye Doctor	Fabric Shop	Factory	Falafel Restaurant

Figure 24: Cluster 7: cities with green points

8. The eighth cluster with **YELLOW** points contains **3** cities:

**Cluster 8** [1](#)

```
city_merge.loc[city_merge['Cluster Labels'] == 7.0, city_merge.columns[0] + list(range(4, city_merge.shape[1]))]
```

City	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
22 Detroit	Fast Food Restaurant	Neighborhood	Optical Shop	Park	Construction & Landscaping	Liquor Store	Eye Doctor	Fabric Shop	Factory	Falafel Restaurant
32 Tucson	Construction & Landscaping	Fast Food Restaurant	Business Service	Zoo Exhibit	Field	Exhibit	Eye Doctor	Fabric Shop	Factory	Falafel Restaurant
68 Plano	Fast Food Restaurant	Park	Convenience Store	American Restaurant	Pizza Place	Carpet Store	Business Service	Zoo Exhibit	Exhibit	Eye Doctor

Figure 25: Cluster 8: cities with yellow points

9. The ninth cluster with **YELLOW-ORANGE** points contains only **1** city Scottsdale:

**Cluster 9**

```
city_merge.loc[city_merge['Cluster Labels'] == 8.0, city_merge.columns[0] + list(range(4, city_merge.shape[1]))]
```

City	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
84 Scottsdale	Beach	Zoo Exhibit	Event Space	Eye Doctor	Fabric Shop	Factory	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant

Figure 26: Cluster 9: cities with yellow-orange points

10. The tenth cluster with **ORANGE-RED** points contains also only **1** city Laredo:

Cluster 10											
	City	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
79	Laredo	Hotel	Coffee Shop	Dessert Shop	Snack Place	Café	Zoo Exhibit	Fast Food Restaurant	Eye Doctor	Fabric Shop	Factory

Figure 27: Cluster 10: cities with orange-red points

## 6 Conclusion

The points in same color, i.e. the cities in same cluster, are similar to each other based on the categories of venues. Therefore, if you have a favourite city and given the table above, then you can simply look up the table and find the similar cities you would love.