Comparing two cities

Introduction

The following post is for my final project for the IBM Data Science Professional Certificate on Coursera. I am using a made up example to tell a story with a problem and a solution solved by data.

Business Problem

As a restaurant owner, finding a location in a new city can have a variety of problems and issues, especially when you do not know the city. There is a plethora of information out there to research and compare, so information overload can set in quickly. Let us suppose that the owner is doing quite well in the current location and wants to find a similar neighborhood in a different city. They like the blend and density of the population and other restaurants near the location, which provides a unique space in the landscape of food choice.

What the owner needs is to narrow down the choices. A map and a list can show the similar neighborhoods, which have the blends of the restaurants, population and income level for the known city so it can be compared to the new city. In this example, the two cities that I will be comparing are St. Louis, Missouri and Dallas, Texas.

Data Description

The important data that we need to find for both cities are the following:

- Population density
- Median household income
- Average number of restaurant categories

I will organize and group all the data by zip code and run a clustering algorithm to find similar areas. Then, color code the resulting clusters on a geographical map.

The data sources needed are the following:

- *uszipcode* library from Python Software Foundation (author: Sanhe Hu)
    - Will give us all the zip codes, income and population data for the given zip codes
- Foursquare API
    - Will give us all the restaurants within a 2 miles radius of each zip code

Methodology

I used a juypter notebook on the cognitivieclass.ai labs cloud platform for my analysis.

From the *uszipcode* library, I loaded the columns of interest for each zip code of each city in its own dataframe.

St. Louis data

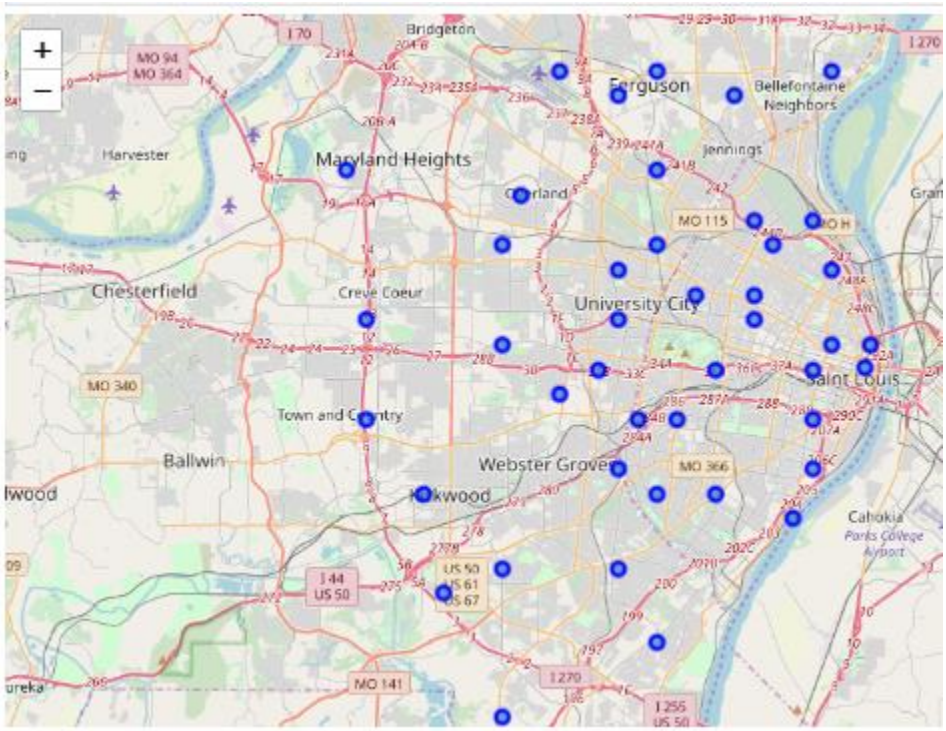
	zipcode	common city list	lat	lng	radius in miles	population	population density	median home value	median household income
0	63101	Saint Louis	38.631	-90.193	0.568182	2620	6947.0	237100.0	54417.0
1	63102	Saint Louis	38.640	-90.190	2.000000	2316	1552.0	194200.0	54018.0
2	63103	Saint Louis	38.630	-90.220	1.000000	6900	3196.0	178300.0	34719.0
3	63104	Saint Louis	38.610	-90.220	2.000000	18656	5399.0	192700.0	45498.0
4	63105	Clayton	38.650	-90.320	2.000000	17667	6647.0	559500.0	86031.0

Dallas data

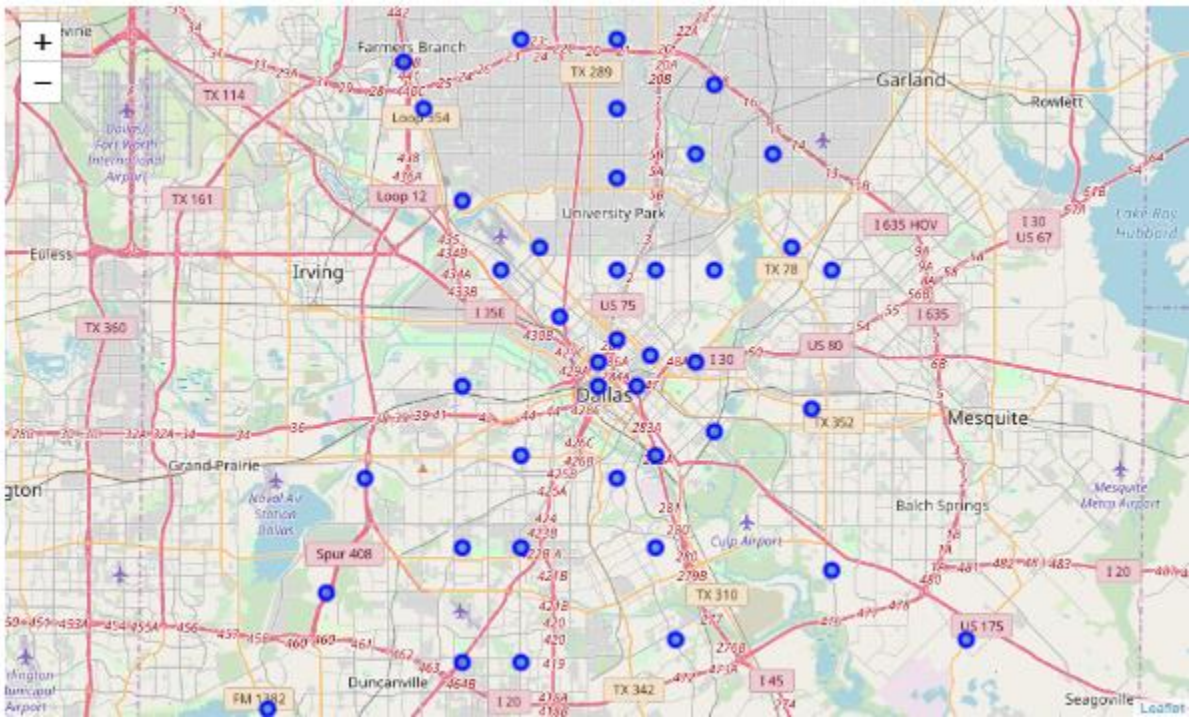
	zipcode	common city list	lat	lng	radius in miles	population	population density	median home value	median household income
44	75249	Dallas	32.64	-96.97	2.0	13373.0	2909.0	117700.0	57247.0
46	75252	Dallas	33.00	-96.79	2.0	24112.0	4778.0	280300.0	54650.0
47	75253	Dallas	32.67	-96.61	4.0	18450.0	1365.0	48200.0	33283.0
48	75254	Dallas	32.95	-96.80	2.0	23253.0	7006.0	294800.0	44283.0
52	75287	Dallas	33.00	-96.84	2.0	49004.0	8754.0	246400.0	46828.0

I used the folium library to visualize the layout of each city and overlaid the zip code locations on to the map.

St. Louis, MO Map



Dallas, TX map



The Foursquare API retrieved all of the restaurant locations for the given zip code. For each zip code, I limited the result set to 100 restaurants and a 2 mile (3219 meter) radius. Here is a small sample of the results:

	zipcode	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
3363	75287	33.0	-96.84	Dickey's Barbecue Pit	33.025397	-96.846218	BBQ Joint
3364	75287	33.0	-96.84	Papa John's Pizza	32.999467	-96.825826	Pizza Place
3365	75287	33.0	-96.84	Burger King	33.001843	-96.857385	Fast Food Restaurant
3366	75287	33.0	-96.84	AwShucks	33.027237	-96.841570	Seafood Restaurant
3367	75287	33.0	-96.84	Taco Cabana	33.025270	-96.845552	Taco Place

I reorganized the category data with “one hot encoding” so I could get the average (mean) number of locations for a given zip code. From this data, I could sort all the restaurant categories by popularity to get a top 10 restaurant category list for each zip code.

	zipcode	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	63101	American Restaurant	Restaurant	Sandwich Place	Food Truck	Italian Restaurant	Pizza Place	Steakhouse	Mexican Restaurant	BBQ Joint	Bakery
1	63102	American Restaurant	Italian Restaurant	Food Truck	Steakhouse	Sandwich Place	New American Restaurant	Pizza Place	Restaurant	Breakfast Spot	BBQ Joint
2	63103	American Restaurant	Sandwich Place	New American Restaurant	Pizza Place	Food Truck	Italian Restaurant	Mexican Restaurant	Restaurant	Steakhouse	Café
3	63104	Pizza Place	Bakery	Restaurant	American Restaurant	Breakfast Spot	BBQ Joint	Sandwich Place	Food Truck	Mexican Restaurant	Café
4	63105	Italian Restaurant	American Restaurant	Pizza Place	Mexican Restaurant	Sandwich Place	Bakery	Thai Restaurant	Chinese Restaurant	Sushi Restaurant	Burger Joint

I added the *population density* and the *median household income* data to the "one hot encoding" data, which gives me all of the data of interest in one dataframe. As a preprocessing step, I scaled the data so the income and density columns would not skew the results.

I used the *K-means* machine learning algorithm from *Sci-kit Learn* library to cluster the data, because it is an easy and efficient way to find similarities from many dimensions of data. I used it to find 6 similar clusters of zip codes, which gives each zip code a label.

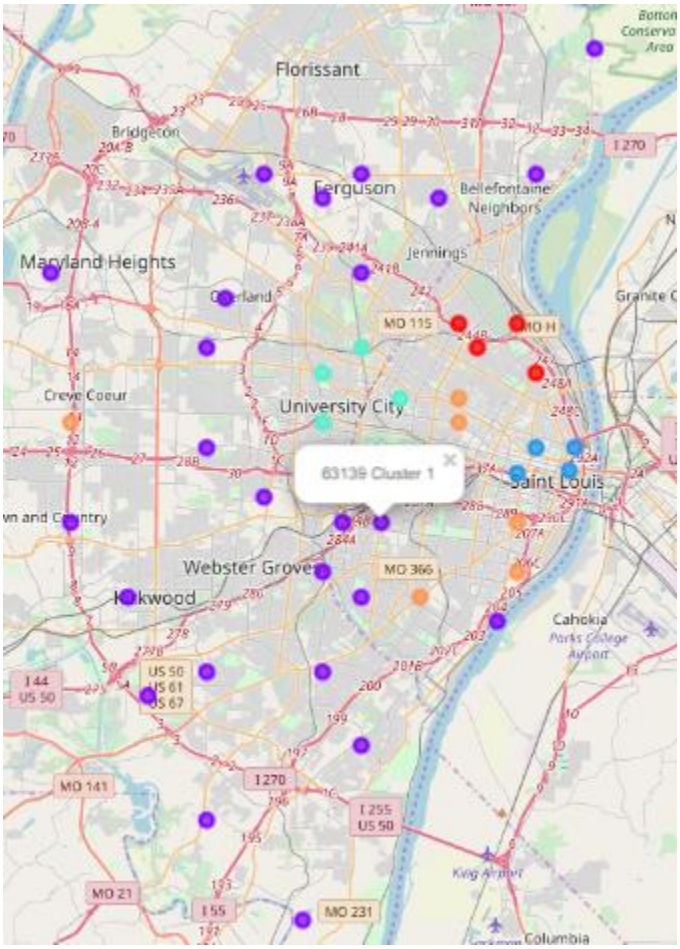
Results

As a final result, I combined all the interesting data, the labels and the top 10 results:

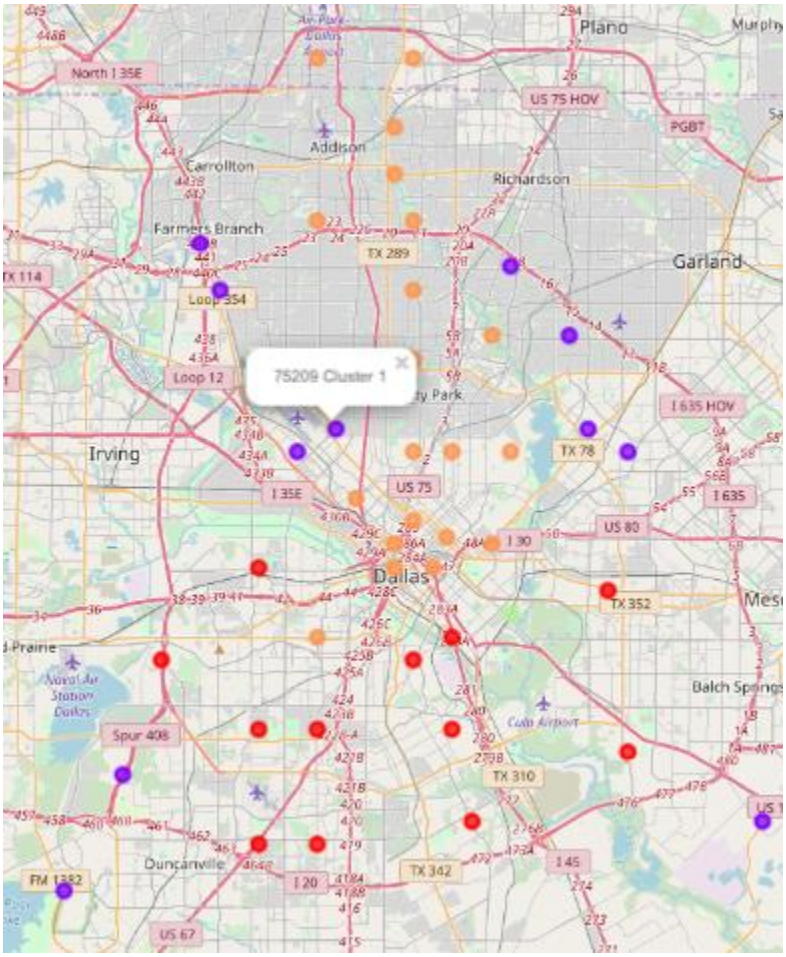
	zipcode	common city list	lat	lng	radius in miles	population	population density	median home value	median household income	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	63101	Saint Louis	38.631	-90.193	0.568182	2620.0	6947.0	237100.0	54417.0	2	American Restaurant	Restaurant	Sandwich Place	Food Truck	Italian Restaurant	Pizza Place	Steakhouse	Mexican Restaurant	BBQ Joint
1	63102	Saint Louis	38.640	-90.190	2.000000	2316.0	1552.0	194200.0	54018.0	2	American Restaurant	Italian Restaurant	Food Truck	Steakhouse	Sandwich Place	New American Restaurant	Pizza Place	Restaurant	Breakfast Spot
2	63103	Saint Louis	38.630	-90.220	1.000000	6900.0	3196.0	178300.0	34719.0	2	American Restaurant	Sandwich Place	New American Restaurant	Pizza Place	Food Truck	Italian Restaurant	Mexican Restaurant	Restaurant	Steakhouse
3	63104	Saint Louis	38.610	-90.220	2.000000	18656.0	5399.0	192700.0	45498.0	5	Pizza Place	Bakery	Restaurant	American Restaurant	Breakfast Spot	BBQ Joint	Sandwich Place	Food Truck	Mexican Restaurant
4	63105	Clayton	38.650	-90.320	2.000000	17667.0	6647.0	559500.0	86031.0	3	Italian Restaurant	American Restaurant	Pizza Place	Mexican Restaurant	Sandwich Place	Bakery	Thai Restaurant	Chinese Restaurant	Sushi Restaurant
5	63106	Saint Louis	38.640	-90.210	2.000000	11883.0	5267.0	89400.0	15126.0	2	American Restaurant	Sandwich Place	Pizza Place	Italian Restaurant	Fast Food Restaurant	Food Truck	Mexican Restaurant	New American Restaurant	Chinese Restaurant
6	63107	Saint Louis	38.670	-90.210	1.000000	11912.0	4962.0	68900.0	23276.0	0	Chinese Restaurant	Fast Food Restaurant	Diner	Sandwich Place	Fried Chicken Joint	American Restaurant	Café	Pizza Place	Food
7	63108	Saint Louis	38.650	-90.250	2.000000	21568.0	9627.0	265000.0	30157.0	5	Fast Food Restaurant	American Restaurant	Sandwich Place	New American Restaurant	Pizza Place	Chinese Restaurant	Café	Sushi Restaurant	Southern / Soul Food Restaurant
8	63109	Saint Louis	38.580	-90.300	2.000000	26946.0	7546.0	163400.0	51951.0	1	Pizza Place	Mexican Restaurant	Sandwich Place	Chinese Restaurant	BBQ Joint	Italian Restaurant	Greek Restaurant	Bakery	American Restaurant
9	63110	Saint Louis	38.630	-90.270	2.000000	17107.0	2732.0	158600.0	38036.0	5	Italian Restaurant	American Restaurant	Sandwich Place	Pizza Place	New American Restaurant	Bakery	Café	Burger Joint	Deli / Bodega

I mapped the zip codes and color coded the clusters so an owner can easily find their zip code color and compare it with the same color in the new city.

St. Louis, MO map



Dallas, TX map



The last bit for the owner is to now to find which cluster their restaurant is in and look at the details to find the blend that makes his restaurant a success.

Below are some of the details for each cluster:

Cluster 0 - Red

	common_city_list	population	population_density	median_home_value	median_household_income	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
6	Saint Louis	11912.0	4962.0	68900.0	23276.0	0	Chinese Restaurant	Fast Food Restaurant	Diner	Sandwich Place	Fried Chicken Joint	American Restaurant	Café	Pizza Place	Food	Bakery
14	Saint Louis	20775.0	4866.0	63300.0	26183.0	0	Fast Food Restaurant	Chinese Restaurant	Fried Chicken Joint	Sandwich Place	Pizza Place	American Restaurant	Diner	Snack Place	Food Court	Food
19	Saint Louis	10296.0	4260.0	52800.0	23996.0	0	Fast Food Restaurant	Chinese Restaurant	Fried Chicken Joint	Sandwich Place	Snack Place	American Restaurant	Diner	Pizza Place	BBQ Joint	Food
44	Saint Louis	11373.0	1937.0	72700.0	29816.0	0	Chinese Restaurant	Fast Food Restaurant	Food	Fried Chicken Joint	Sandwich Place	American Restaurant	Pizza Place	Truck Stop	Diner	Bakery
2	Dallas	15721.0	3324.0	65600.0	28602.0	0	Mexican Restaurant	BBQ Joint	Fried Chicken Joint	Fast Food Restaurant	Sandwich Place	Seafood Restaurant	Breakfast Spot	Deli / Bodega	Southern / Soul Food Restaurant	Chinese Restaurant
10	Dallas	73146.0	4151.0	87800.0	36704.0	0	Mexican Restaurant	Fast Food Restaurant	Pizza Place	Food	Wings Joint	Donut Shop	Fried Chicken Joint	Deli / Bodega	BBQ Joint	Taco Place
11	Dallas	24884.0	2418.0	73700.0	28802.0	0	Fast Food Restaurant	Mexican Restaurant	Pizza Place	Chinese Restaurant	Donut Shop	Fried Chicken Joint	Sandwich Place	BBQ Joint	American Restaurant	Taco Place
13	Dallas	14648.0	1741.0	56300.0	23433.0	0	Mexican Restaurant	Food	American Restaurant	Fried Chicken Joint	Pizza Place	Seafood Restaurant	Southern / Soul Food Restaurant	Fast Food Restaurant	BBQ Joint	Sandwich Place
14	Dallas	49416.0	3379.0	59700.0	23294.0	0	Fried Chicken Joint	Pizza Place	Fast Food Restaurant	BBQ Joint	Sandwich Place	Seafood Restaurant	Deli / Bodega	Breakfast Spot	Burger Joint	Food Truck
15	Dallas	80324.0	2940.0	73200.0	34480.0	0	Fast Food Restaurant	Fried Chicken Joint	Mexican Restaurant	American Restaurant	Diner	Fish & Chips Shop	Wings Joint	Deli / Bodega	Dim Sum Restaurant	Donut Shop
18	Dallas	41891.0	3701.0	161100.0	35827.0	0	Mexican Restaurant	Fast Food Restaurant	Sandwich Place	Pizza Place	Fried Chicken Joint	Taco Place	BBQ Joint	Deli / Bodega	American Restaurant	Breakfast Spot
20	Dallas	34034.0	5825.0	86900.0	33830.0	0	Fast Food Restaurant	Mexican Restaurant	Fried Chicken Joint	Pizza Place	Donut Shop	Food	Chinese Restaurant	Taco Place	Sandwich Place	Breakfast Spot
23	Dallas	55029.0	4828.0	86200.0	37720.0	0	Mexican Restaurant	Fast Food Restaurant	Pizza Place	Fried Chicken Joint	Chinese Restaurant	BBQ Joint	Food	American Restaurant	Wings Joint	Seafood Restaurant
28	Dallas	28682.0	3469.0	91200.0	36265.0	0	Fast Food Restaurant	Fried Chicken Joint	BBQ Joint	Sandwich Place	Mexican Restaurant	Chinese Restaurant	Pizza Place	Burger Joint	Restaurant	Snack Place
29	Dallas	14043.0	4057.0	115400.0	40151.0	0	Fast Food Restaurant	Mexican Restaurant	Fried Chicken Joint	Taco Place	Chinese Restaurant	Pizza Place	Donut Shop	Seafood Restaurant	Sandwich Place	Restaurant
33	Dallas	17101.0	2572.0	104500.0	29787.0	0	Fast Food Restaurant	Fried Chicken Joint	Chinese Restaurant	Seafood Restaurant	BBQ Joint	Pizza Place	Restaurant	Sandwich Place	Wings Joint	Mexican Restaurant
36	Dallas	27066.0	1008.0	78300.0	30870.0	0	BBQ Joint	Fried Chicken Joint	Fast Food Restaurant	Food	Seafood Restaurant	Pizza Place	Food Truck	Truck Stop	Restaurant	Diner

Cluster 1 - Purple (partial results)

	common_city_list	population	population_density	median_home_value	median_household_income	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
8	Saint Louis	26946.0	7546.0	163400.0	51951.0	1	Pizza Place	Mexican Restaurant	Sandwich Place	Chinese Restaurant	BBQ Joint	Italian Restaurant	Greek Restaurant	Bakery	American Restaurant	Deli / Bodega
10	Saint Louis	20313.0	6275.0	90900.0	30334.0	1	Mexican Restaurant	Chinese Restaurant	Bakery	Pizza Place	Fast Food Restaurant	Fried Chicken Joint	Vietnamese Restaurant	American Restaurant	Café	Sandwich Place
13	Breckenridge Hills	36201.0	4105.0	92800.0	41771.0	1	Fast Food Restaurant	Pizza Place	Mexican Restaurant	Chinese Restaurant	Sandwich Place	American Restaurant	Fried Chicken Joint	Café	Deli / Bodega	Indian Restaurant
18	Webster Groves	33969.0	4048.0	216200.0	68348.0	1	Pizza Place	Sandwich Place	Mexican Restaurant	New American Restaurant	American Restaurant	Fried Chicken Joint	Chinese Restaurant	Italian Restaurant	Deli / Bodega	Asian Restaurant
20	Normandy	26602.0	3705.0	80500.0	33704.0	1	Fast Food Restaurant	Chinese Restaurant	Sandwich Place	Pizza Place	Food	Diner	Café	Fried Chicken Joint	American Restaurant	BBQ Joint
21	Kirkwood	38495.0	2719.0	255400.0	80682.0	1	Fast Food Restaurant	Sandwich Place	Pizza Place	American Restaurant	Bakery	Restaurant	Mexican Restaurant	Steakhouse	Chinese Restaurant	Burger Joint
22	Affton	49308.0	3927.0	147400.0	52877.0	1	Pizza Place	Fast Food Restaurant	Sandwich Place	Bakery	Chinese Restaurant	Deli / Bodega	Restaurant	Mexican Restaurant	Food	Greek Restaurant
23	Clayton	10417.0	1190.0	656800.0	118816.0	1	American Restaurant	Italian Restaurant	Café	Chinese Restaurant	Pizza Place	Bakery	Snack Place	Mexican Restaurant	Sandwich Place	Breakfast Spot
24	Lemay	32201.0	3242.0	123800.0	44458.0	1	Pizza Place	Fast Food Restaurant	American Restaurant	Sandwich Place	Bakery	Chinese Restaurant	Fried Chicken Joint	Mexican Restaurant	Donut Shop	Breakfast Spot
25	Sappington	15112.0	3253.0	185800.0	64436.0	1	Fast Food Restaurant	Sandwich Place	Pizza Place	American Restaurant	Steakhouse	Chinese Restaurant	Breakfast Spot	Italian Restaurant	Restaurant	Mexican Restaurant
26	Sappington	4939.0	785.0	395100.0	88469.0	1	Fast Food Restaurant	Sandwich Place	Pizza Place	Chinese Restaurant	Mexican Restaurant	American Restaurant	Steakhouse	Breakfast Spot	Sushi Restaurant	Burger Joint
27	Sappington	29356.0	1955.0	211300.0	73440.0	1	Fast Food Restaurant	Pizza Place	Sandwich Place	Chinese Restaurant	Bakery	Fried Chicken Joint	Restaurant	Mexican Restaurant	Café	Deli / Bodega
28	Saint Louis	52718.0	2523.0	198000.0	65861.0	1	Pizza Place	Chinese Restaurant	Bakery	American Restaurant	Mexican Restaurant	Greek Restaurant	Breakfast Spot	Comfort Food Restaurant	Sandwich Place	Fast Food Restaurant
30	Des Peres	16769.0	1205.0	491000.0	132500.0	1	Sandwich Place	Fast Food Restaurant	Pizza Place	Bakery	Chinese Restaurant	Snack Place	Burger Joint	Mexican Restaurant	Breakfast Spot	Deli / Bodega
31	Olivette	13988.0	2658.0	251900.0	55687.0	1	Chinese Restaurant	Fast Food Restaurant	Pizza Place	Mexican Restaurant	Sandwich Place	Café	Indian Restaurant	Korean Restaurant	Fried Chicken Joint	Vietnamese Restaurant

Cluster 2 – Dark Blue

	common_city_list	population	population_density	median_home_value	median_household_income	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 Louis	2620.0	6947.0	237100.0	54417.0	2	American Restaurant	Restaurant	Sandwich Place	Food Truck	Italian Restaurant	Pizza Place	Steakhouse	Mexican Restaurant	BBQ Joint	Bakery
1	Saint Louis	2316.0	1552.0	194200.0	54018.0	2	American Restaurant	Italian Restaurant	Food Truck	Steakhouse	Sandwich Place	New American Restaurant	Pizza Place	Restaurant	Breakfast Spot	BBQ Joint
2	Saint Louis	6900.0	3196.0	178300.0	34719.0	2	American Restaurant	Sandwich Place	New American Restaurant	Pizza Place	Food Truck	Italian Restaurant	Mexican Restaurant	Restaurant	Steakhouse	Café
5	Saint Louis	11883.0	5267.0	89400.0	15126.0	2	American Restaurant	Sandwich Place	Pizza Place	Italian Restaurant	Fast Food Restaurant	Food Truck	Mexican Restaurant	New American Restaurant	Chinese Restaurant	Steakhouse

Cluster 3 – Light Blue

	common_city_list	population	population_density	median_home_value	median_household_income	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
4	Clayton	17667.0	6647.0	559500.0	86031.0	3	Italian Restaurant	American Restaurant	Pizza Place	Mexican Restaurant	Sandwich Place	Bakery	Thai Restaurant	Chinese Restaurant	Sushi Restaurant	Burger Joint
11	Saint Louis	20368.0	6157.0	138200.0	27383.0	3	American Restaurant	Sandwich Place	Pizza Place	Fast Food Restaurant	Thai Restaurant	Bakery	Indian Restaurant	Café	Italian Restaurant	Chinese Restaurant
29	University City	30084.0	5973.0	209200.0	55976.0	3	Italian Restaurant	Pizza Place	American Restaurant	Chinese Restaurant	Thai Restaurant	Sandwich Place	Steakhouse	Mexican Restaurant	Indian Restaurant	Taco Place
32	Saint Louis	8161.0	2733.0	62100.0	25239.0	3	Chinese Restaurant	Fast Food Restaurant	Sandwich Place	American Restaurant	Pizza Place	Thai Restaurant	Wings Joint	Bakery	Fried Chicken Joint	Indian Restaurant

Cluster 4 - Yellow

	common_city_list	population	population_density	median_home_value	median_household_income	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
9	Dallas	7482.0	3316.0	54900.0	15258.0	4	American Restaurant	Fried Chicken Joint	Fast Food Restaurant	Southern / Soul Food Restaurant	German Restaurant	Food	Hot Dog Joint	Chinese Restaurant	Restaurant	Café

Cluster 5 – Orange (partial results)

	common_city_list	population	population_density	median_home_value	median_household_income	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	Saint Louis	18656.0	5399.0	192700.0	45498.0	5	Pizza Place	Bakery	Restaurant	American Restaurant	Breakfast Spot	BBQ Joint	Sandwich Place	Food Truck	Mexican Restaurant	Café
7	Saint Louis	21568.0	9627.0	265000.0	30157.0	5	Fast Food Restaurant	American Restaurant	Sandwich Place	New American Restaurant	Pizza Place	Chinese Restaurant	Café	Sushi Restaurant	Southern / Soul Food Restaurant	Fried Chicken Joint
9	Saint Louis	17107.0	2732.0	158600.0	38036.0	5	Italian Restaurant	American Restaurant	Sandwich Place	Pizza Place	New American Restaurant	Bakery	Café	Burger Joint	Deli / Bodega	Sushi Restaurant
12	Saint Louis	13167.0	5187.0	66600.0	24561.0	5	Fast Food Restaurant	Sandwich Place	Chinese Restaurant	Pizza Place	American Restaurant	New American Restaurant	Café	Fried Chicken Joint	Gastropub	Mexican Restaurant
15	Saint Louis	43540.0	7909.0	111500.0	40251.0	5	Pizza Place	Bakery	Mexican Restaurant	Chinese Restaurant	Sandwich Place	Vietnamese Restaurant	Fast Food Restaurant	American Restaurant	BBQ Joint	Café
16	Richmond Heights	9163.0	3889.0	230000.0	62593.0	5	Italian Restaurant	Bakery	Sandwich Place	Burger Joint	Mexican Restaurant	Pizza Place	American Restaurant	New American Restaurant	Chinese Restaurant	Breakfast Spot
17	Saint Louis	26704.0	7961.0	99800.0	28630.0	5	Pizza Place	Bakery	Sandwich Place	Restaurant	Mexican Restaurant	Vietnamese Restaurant	American Restaurant	BBQ Joint	Breakfast Spot	Fast Food Restaurant
40	Creve Coeur	20593.0	1457.0	373700.0	102708.0	5	Sandwich Place	Pizza Place	Café	Food Truck	American Restaurant	Donut Shop	Mexican Restaurant	Breakfast Spot	Indian Restaurant	Italian Restaurant
0	Dallas	9409.0	6514.0	520000.0	67979.0	5	American Restaurant	Burger Joint	Steakhouse	New American Restaurant	Mexican Restaurant	Restaurant	Seafood Restaurant	Pizza Place	Japanese Restaurant	Italian Restaurant
1	Dallas	1666.0	2358.0	211100.0	79681.0	5	American Restaurant	Mexican Restaurant	Steakhouse	Japanese Restaurant	New American Restaurant	Burger Joint	Seafood Restaurant	French Restaurant	Sandwich Place	Italian Restaurant
3	Dallas	26279.0	10277.0	254400.0	63551.0	5	American Restaurant	Burger Joint	Mexican Restaurant	Pizza Place	Steakhouse	New American Restaurant	Seafood Restaurant	Japanese Restaurant	Taco Place	Sushi Restaurant
4	Dallas	23061.0	5302.0	937100.0	108913.0	5	Mexican Restaurant	Pizza Place	French Restaurant	New American Restaurant	Restaurant	Seafood Restaurant	Burger Joint	Sandwich Place	Italian Restaurant	Taco Place
5	Dallas	36248.0	8516.0	268000.0	52192.0	5	Mexican Restaurant	New American Restaurant	Pizza Place	Burger Joint	Taco Place	American Restaurant	French Restaurant	Restaurant	Vietnamese Restaurant	Greek Restaurant
7	Dallas	30171.0	5080.0	120900.0	45170.0	5	Mexican Restaurant	Pizza Place	Fast Food Restaurant	Taco Place	American Restaurant	Italian Restaurant	Sandwich Place	Fried Chicken Joint	Gastropub	Burger Joint
12	Dallas	32950.0	4553.0	336900.0	69168.0	5	Mexican Restaurant	Taco Place	Burger Joint	American Restaurant	Italian Restaurant	BBQ Joint	Sandwich Place	Fast Food Restaurant	Pizza Place	Restaurant

Discussion

In my perspective, there are a lot of different metrics to look at when finding a new restaurant. The complexity can be overwhelming when there are too many dimensions to look at, all at once, over many different areas.

I chose to use a K-means machine learning algorithm in my analysis to input all the valuable dimensions of data and find similar environments in two different cities. I understand the analysis does not give you a single answer to choose from. The goal was to reduce the number of areas to search through to find a similar environment.

Conclusion

The resulting data may not be able to show you the perfect place to open a new restaurant, but this drastically reduces the pool of search results. Think of it as a starting point to see similar areas in a different city.

References

- *uszipcode* library from Python Software Foundation (author: Sanhe Hu)
- Foursquare API
- Coursera - IBM Data Science Professional Certificate
- Comparing two cities – Blogger post