Tale of two cities: Business Venture Location Analysis for Salt Lake City and Provo

1. Business Problem: Introduction:

I am a data scientist at a consulting firm (based in Salt Lake City, Utah) that helps entrepreneurs navigate the process of opening a new business venture. My target audience is one of our clients who is interested in opening an eating establishment (either a coffee shop, a restaurant or bakery) in the state of Utah in the United States. She has heard that Utah is very welcoming to new businesses and needs help to figure out an optimal location.

Naturally, our client wants to look at the two most populous counties as potential business sites since these two counties will have the highest number of potential customers. Salt Lake City and Provo are two of the most prominent cities in Utah that would work well for a business location. The goal of this project is to analyze/compare the neighborhoods in the cities of Salt Lake City and Provo and find an optimal location for a business in each city. **The report generated from the analysis will be target specifically to our client, who is the stakeholder**.

Additionally, our firm would like to know how similar or different the neighborhoods within each county are and how the two counties compare with each other. This will enable us to help our client decide on where and what her first business venture should be.

Since I live in the suburbs of Salt Lake City, I am aware of the demographic differences between the two counties. I am personally interested in seeing how the demographic differences manifest in the location/venue data.

2. Data and Methodology:

The two populous counties in Utah are Salt Lake County and Utah County.

https://www.utah-demographics.com/counties_by_population

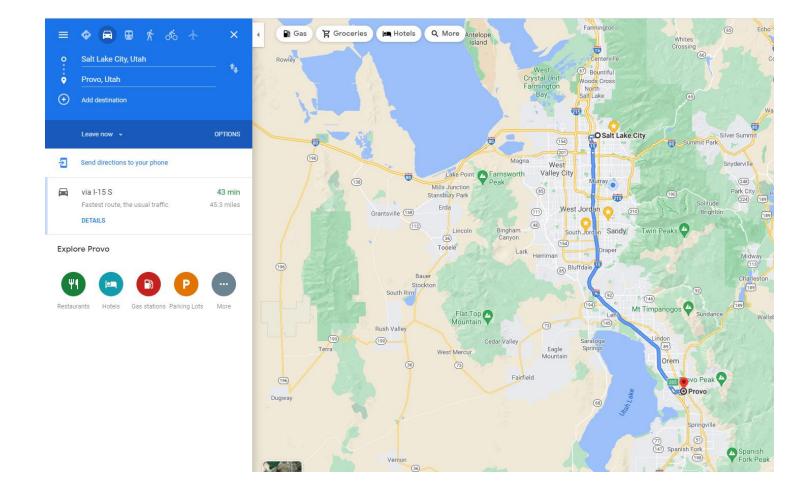
Here's a screenshot listing Utah counties by population

Rank	County	Population
1	Salt Lake County	1,133,646
2	<u>Utah County</u>	605,490
3	<u>Davis County</u>	345,767
1	Weber County	251,498
5	<u>Washington County</u>	165,811
5	Cache County	124,165
7	Tooele County	67,397
3	Box Elder County	53,946
7	Iron County	51,213
10	Summit County	41,103

The county seats (or the administrative centers) tend to be the most viable for starting a business. Salt Lake City and Provo are the county seats for Salt Lake and Utah County respectively. I will be choosing Salt Lake City and Provo neighborhoods as the focus of my analysis.

https://en.wikipedia.org/wiki/List_of_counties_in_Utah

The list of neighborhoods in each city is acquired by webscraping using BeautifulSoup. Using the neighborhood list, the coordinates for each neighborhood are acquired using the geolocater API and the coordinates superimposed on the map using Folium. Once we have the coordinates, FourSquare API is used to get the list of venues in the vicinity of each coordinate. After one-hot encoding of the venues, K-means algorithm is used for clustering the neighborhoods. The most frequently occurring venue categories are compiled within each city and then the two cities are compared. Since there are lots of eating establishments in both cities we will try to detect areas that are not already saturated. Based on the analysis, business venture recommendations can be made accordingly. The analysis is comprehensively summarized in the report and presentation which will be given to the stakeholder to make the final decision.



Salt Lake City and Provo are less than 50 miles apart and yet the demographics are quite different.

First part of the data acquisition involves getting the names of all the neighborhoods in Salt Lake City and Provo

This information will need to webscraped using BeautifulSoup from the following webpages

https://en.wikipedia.org/wiki/List_of_Salt_Lake_City_neighborhoods

https://www.provo.org/government/city-council/neighborhood-program/neighborhoods

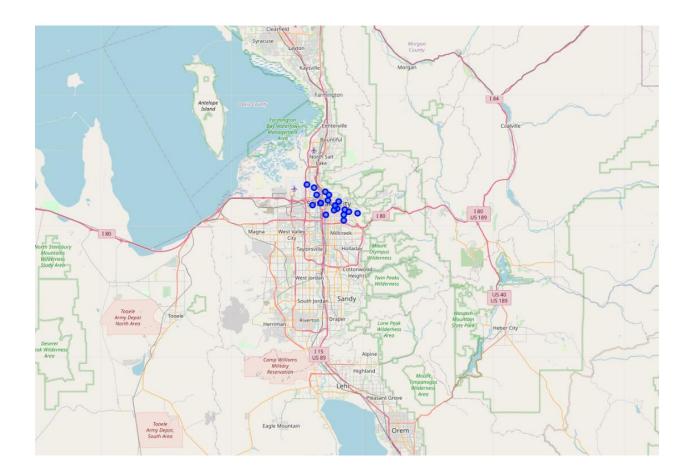
3. Analysis:

First we analyze, Salt Lake City Neighborhoods: Info was acquired through webscraping and Geolocator API and converted to a pandas DataFrame

[22]:		Neighborhood	Latitude	Longitude
	0	Sugarhouse	40.725304	-111.847947
	1	The Avenues	40.760827	-111.908884
	2	9th & 9th	40.749789	-111.865394
	3	15th and 15th	40.736873	-111.847987
	4	Downtown	40.760827	-111.908884
	5	Capitol Hill	40.783988	-111.897367
	6	Yalecrest	40.746822	-111.845507
	7	Liberty Park	40.745731	-111.874030
	8	Marmalade	40.777439	-111.888198
	9	Central City	40.755332	-111.871650
	10	University	40.766101	-111.890793
	11	East Bench	40.739892	-111.810322
	12	Foothill	40.743002	-111.834101
	13	Ballpark	40.736458	-111.896101
	14	Central City	40.755332	-111.871650
	15	East Central	40.764112	-111.861163
	16	Fairpark	40.777307	-111.920437
	17	Federal Heights	40.760827	-111.908884
	18	Poplar Grove	40.756123	-111.932037
	19	Rose Park	40.791787	-111.926857
	20	Westpointe	40.798557	-111.946605

There are 21 neighborhoods in total

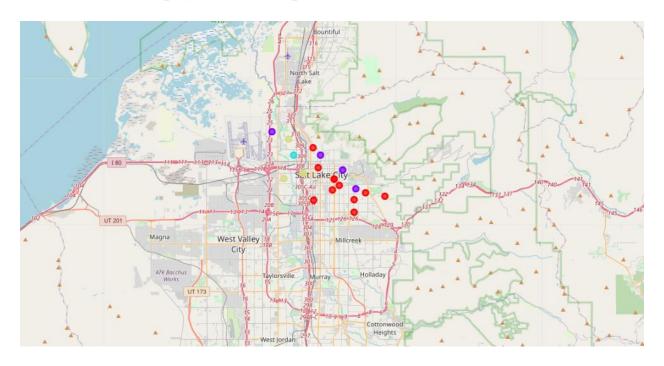
Using the Folium library the neighborhoods on superimposed on the Salt Lake City map



Using the latitudes and longitude data for each neighborhood, venue location data is obtained for each neighborhood and the most common venues are tabulated. K-means clustering algorithm is used to divide the neighborhoods into 4 clusters.

:	Neighborh	od Latitu	ide	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
	0 Sugarho	use 40.725	304 -	111.847947	0	Fast Food Restaurant	Ice Cream Shop	Lingerie Store	American Restaurant	Park
	1 The Aver	ues 40.760	327 -	111.908884	3	Light Rail Station	Furniture / Home Store	Art Gallery	Dim Sum Restaurant	Music Venue
	2 9th 8	9th 40,749	789 -	111.865394	0	Coffee Shop	Gift Shop	Dessert Shop	New American Restaurant	Grocery Store
	3 15th and	5th 40.736	373 -	111.847987	0	Ice Cream Shop	Bakery	Middle Eastern Restaurant	Spanish Restaurant	Bookstore
	4 Downt	wn 40.760	327 -	111,908884	3	Light Rail Station	Furniture / Home Store	Art Gallery	Dim Sum Restaurant	Music Venue
	5 Capito	Hill 40.783	888 -	111.897367	0	Convenience Store	Bar	Salon / Barbershop	Nightclub	Theater
	6 Yaled	est 40.746	322 -	111.845507	1	Park	Trail	ATM	Middle Eastern Restaurant	Plaza
	7 Liberty	ark 40.745	31 -	111.874030	0	Zoo	Park	Coffee Shop	Diner	Lake
	8 Marma	de 40.777	139 -	111.888198	1	Park	Coffee Shop	Bed & Breakfast	American Restaurant	History Museum
	9 Central	ity 40.755	32 -	111.871650	0	Italian Restaurant	Boutique	Grocery Store	Bakery	Cosmetics Shop
1	0 Unive	sity 40.766	01 -	111.890793	0	Hotel	Sandwich Place	American Restaurant	Restaurant	Steakhouse
1	1 East Be	nch 40.739	392 -	111.810322	0	Scenic Lookout	Gym	Lake	ATM	Miscellaneous Shop
1	2 Foo	hill 40.743	002 -	111.834101	0	Convenience Store	Park	Bank	Chinese Restaurant	Flower Shop
1	3 Ball	ark 40.736	158 -	111.896101	0	ATM	Hardware Store	Deli / Bodega	Sandwich Place	Smoke Shop
1	4 Central	ity 40.755	32 -	111.871650	0	Italian Restaurant	Boutique	Grocery Store	Bakery	Cosmetics Shop
1	5 East Ce	tral 40.764	12 -	111.861163	1	Coffee Shop	Park	ATM	Restaurant	Pub
1	6 Fair	ark 40.777	307 -	111.920437	2	Arts & Entertainment	Basketball Court	ATM	Miscellaneous Shop	Pub
1	7 Federal Hei	hts 40.760	327 -	111.908884	3	Light Rail Station	Furniture / Home Store	Art Gallery	Dim Sum Restaurant	Music Venue
1	8 Poplar G	ove 40.756	23 -	111.932037	3	Construction & Landscaping	Moving Target	Mexican Restaurant	Miscellaneous Shop	Pub
1	9 Rose	ark 40.791	787 -	111.926857	3	Food Service	Golf Course	Mexican Restaurant	MTA	Plaza
2	0 Westpo	nte 40.798	57 -	111.946605	1	Park	Dive Bar	ATM	Miscellaneous Shop	Plaza

The clusters are displayed on the map



We then examine each cluster. Cluster 1 seems like a very vibrant neighborhood with a lot of venues

Cluster 1

: sa	ltlake_merged.	loc[saltlake_merged['(Cluster Labels'] == 0,	saltlake_merged.column	ns[[0] + list(range(4,	saltlake_merged.shape[1]))]]
:	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	
0	Sugarhouse	Fast Food Restaurant	Ice Cream Shop	Lingerie Store	American Restaurant	Park	
2	9th & 9th	Coffee Shop	Gift Shop	Dessert Shop	New American Restaurant	Grocery Store	
3	15th and 15th	Ice Cream Shop	Bakery	Middle Eastern Restaurant	Spanish Restaurant	Bookstore	
5	Capitol Hill	Convenience Store	Bar	Salon / Barbershop	Nightclub	Theater	
7	Liberty Park	Zoo	Park	Coffee Shop	Diner	Lake	
9	Central City	Italian Restaurant	Boutique	Grocery Store	Bakery	Cosmetics Shop	
10	University	Hotel	Sandwich Place	American Restaurant	Restaurant	Steakhouse	
11	East Bench	Scenic Lookout	Gym	Lake	MTA	Miscellaneous Shop	
12	Foothill	Convenience Store	Park	Bank	Chinese Restaurant	Flower Shop	
13	Ballpark	ATM	Hardware Store	Deli / Bodega	Sandwich Place	Smoke Shop	
14	Central City	Italian Restaurant	Boutique	Grocery Store	Bakery	Cosmetics Shop	

Cluster 2 has less number of neighborhoods but does have a some eating places

Cluster 2

Neighborhood 1st Most Common Venue 2nd Most Common Venue 3rd Most Common Venue 4th Most Common Venue 5th Most Common Venue 6 Yalecrest Park Trail ATM Middle Eastern Restaurant Plaza 8 Marmalade Park Coffee Shop Bed & Breakfast American Restaurant History Museum 15 East Central Coffee Shop Park ATM Restaurant Pub 20 Westpointe Park Dive Bar ATM Miscellaneous Shop Plaza	2]: saltlake_merged.loc[saltlake_merged['Cluster Labels'] == 1, saltlake_merged.columns[[0]							saltlake_merged.shape
8 Marmalade Park Coffee Shop Bed & Breakfast American Restaurant History Museum 15 East Central Coffee Shop Park ATM Restaurant Pub]:	Ne	ighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
15 East Central Coffee Shop Park ATM Restaurant Pub		6	Yalecrest	Park	Trail	ATM	Middle Eastern Restaurant	Plaza
	1	8	Marmalade	Park	Coffee Shop	Bed & Breakfast	American Restaurant	History Museum
20 Westpointe Park Dive Bar ATM Miscellaneous Shop Plaza	1	5	East Central	Coffee Shop	Park	ATM	Restaurant	Pub
	2	0	Westpointe	Park	Dive Bar	ATM	Miscellaneous Shop	Plaza

Cluster 3 seems more industrial

Cluster 3

[43]: s	altlake_merged	.loc[saltlake_merged['	Cluster Labels'] == 2,	Labels'] == 2, saltlake_merged.columns[[0] + list(range(4, saltlak		
[43]:	Neighborhood	1 st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	16 Fairparl	: Arts & Entertainment	Basketball Court	ATM	Miscellaneous Shop	Pub

Cluster 4 seems more like an office district

Cluster 4

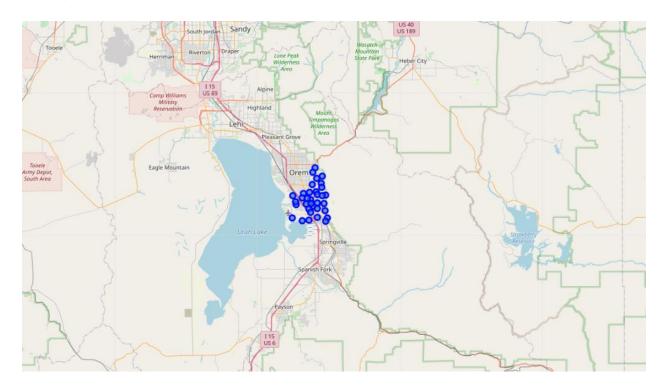


Next, we do a similar analysis on Provo neighborhoods: Info was acquired through webscraping and Geolocator API and converted to a pandas DataFrame

[58]:		Neighborhood	Latitude	Longitude
	0	Carterville	40.259880	-111.668621
	1	Dixon	40.237168	-111.677765
	2	Downtown	40.233294	-111.661315
	3	East Bay	40.212666	-111.650474
	4	Edgemont	40.286066	-111.651866
	5	Foothills	40.238850	-111.634058
	6	Fort Utah	40.237069	-111.702384
	7	Franklin	40.228002	-111.670665
	8	Franklin South	40.222205	-111.666880
	9	Grandview North	40.257786	-111.684236
	10	Grandview South	40.249176	-111.687033
	11	Indian Hills	40.276493	-111.639747
	12	Joaquin	40.240187	-111.650371
	13	Lakeview North	40.253732	-111.710958
	14	Lakeview South	40.241539	-111.703080
	15	Lakewood	40.207350	-111.670245
	16	Maeser	40.229487	-111.649537
	17	North Park	40.245514	-111.666089
	18	North Timpview	40.305259	-111.655147
	19	Oak Hills	40.254138	-111.629868
	20	Pleasant View	40.260597	-111.652561
	21	Provo Bay	40.211185	-111.711363
	22	Provost	40.225441	-111.630560
	23	Provost South	40.211571	-111.625508
	24	Riverbottoms	40.297065	-111.661324
	25	Rivergrove	40.247876	-111.675328
	26	Riverside	40.273823	-111.662062
	27	Rock Canyon	40.269529	-111.639238
	28	Sherwood Hills	40.289842	-111.638113
	29	Spring Creek	40.205466	-111.628930
	30	Sunset	40.205704	-111.686547
	31	Timp	40.238421	-111.664827
	32	Brigham Young University	40.255011	-111.649657
	33	Wasatch	40.252979	-111.638216

There are 34 neighborhoods in total

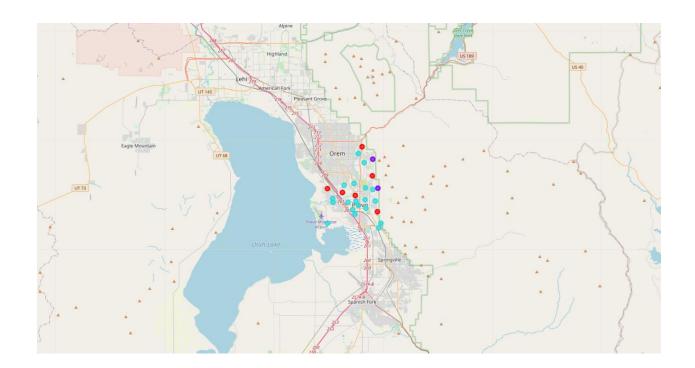
The Provo neighborhood information is displayed on a map using the Folium library



Using the latitudes and longitude data for each neighborhood, venue location data is obtained for each neighborhood and the most common venues are tabulated. K-means clustering algorithm is used to divide the neighborhoods into 4 clusters.

	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Carterville	40.259880	-111.668621	2.0	Hotel	Pizza Place	Cosmetics Shop	Gym / Fitness Center	Health & Beauty Service
1	Dixon	40.237168	-111.677765	2.0	Restaurant	Bakery	Home Service	Campground	ATM
2	Downtown	40.233294	-111.661315	2.0	Mexican Restaurant	Latin American Restaurant	Asian Restaurant	Sandwich Place	Bar
3	Edgemont	40.286066	-111.651866	2.0	Massage Studio	Health Food Store	ATM	Performing Arts Venue	Park
4	Foothills	40.238850	-111.634058	2.0	Intersection	Water Park	Taco Place	Massage Studio	Performing Arts Venue
5	Fort Utah	40.237069	-111.702384	2.0	Farm	Airport Terminal	Construction & Landscaping	Park	Brewery
6	Franklin	40.228002	-111.670665	2.0	Rock Club	Train Station	Chinese Restaurant	ATM	Mediterranean Restaurant
7	Franklin South	40.222205	-111.666880	2.0	Train Station	Gun Shop	Construction & Landscaping	ATM	Performing Arts Venue
8	Grandview North	40.257786	-111.684236	2.0	Construction & Landscaping	ATM	Plaza	Pizza Place	Performing Arts Venue
9	Grandview South	40.249176	-111.687033	0.0	Park	Clothing Store	ATM	Plaza	Pizza Place
10	Joaquin	40.240187	-111.650371	2.0	Dessert Shop	Gas Station	Mexican Restaurant	Pizza Place	Performing Arts Venue
11	Lakeview North	40.253732	-111.710958	0.0	Park	ATM	Mediterranean Restaurant	Pizza Place	Performing Arts Venue
12	Lakeview South	40.241539	-111.703080	2.0	Farm	Playground	ATM	Mediterranean Restaurant	Performing Arts Venue
13	Maeser	40.229487	-111.649537	2.0	Discount Store	Video Store	Construction & Landscaping	Gas Station	Chinese Restaurant
14	North Park	40.245514	-111.666089	0.0	Pizza Place	Park	Gym / Fitness Center	Sandwich Place	Café
15	North Timpview	40.305259	-111.655147	0.0	Trail	Park	ATM	Massage Studio	Performing Arts Venue
16	Oak Hills	40.254138	-111.629868	1.0	Food	Scenic Lookout	ATM	Playground	Performing Arts Venue
17	Pleasant View	40.260597	-111.652561	3.0	Electronics Store	Mediterranean Restaurant	Pizza Place	Performing Arts Venue	Park
18	Provo Bay	40.211185	-111.711363	2.0	Airport Service	Massage Studio	Performing Arts Venue	Park	New American Restaurant
19	Provost	40.225441	-111.630560	0.0	Entertainment Service	Park	Home Service	ATM	Mediterranean Restaurant
20	Provost South	40.211571	-111.625508	2.0	Convenience Store	Garden Center	ATM	Plaza	Pizza Place
21	Riverbottoms	40.297065	-111.661324	2.0	Brazilian Restaurant	Burger Joint	Yoga Studio	Clothing Store	Kids Store
22	Rock Canyon	40.269529	-111.639238	0.0	Park	Music Venue	ATM	Mediterranean Restaurant	Pizza Place
23	Sherwood Hills	40.289842	-111.638113	1.0	Food	ATM	Mediterranean Restaurant	Performing Arts Venue	Park
24	Spring Creek	40.205466	-111.628930	2.0	Construction & Landscaping	Business Service	Convenience Store	Lawyer	Furniture / Home Store
25	Timp	40.238421	-111.664827	2.0	MTA	Snack Place	Laser Tag	Video Store	Auto Garage
26	Brigham Young University	40.255011	-111.649657	2.0	Dessert Shop	Science Museum	Plaza	Pizza Place	Performing Arts Venue
27	Wasatch	40.252979	-111.638216	2.0	Home Service	Sandwich Place	ATM	Mediterranean Restaurant	Performing Arts Venue

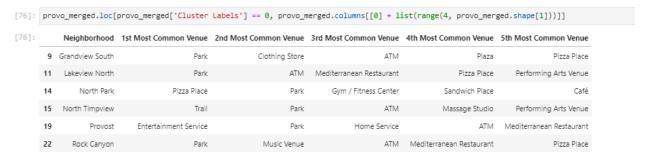
The clusters are displayed on the map



We examine each individual Provo neighborhoods cluster

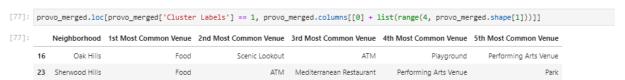
Cluster 1 seems more residential and office space

Cluster 1



Cluster 2 seems like it is more industrial

Cluster 2



Cluster 3 seems very vibrant and looks to be good place to start a business

Cluster 3



Cluster 4 seems more isolated

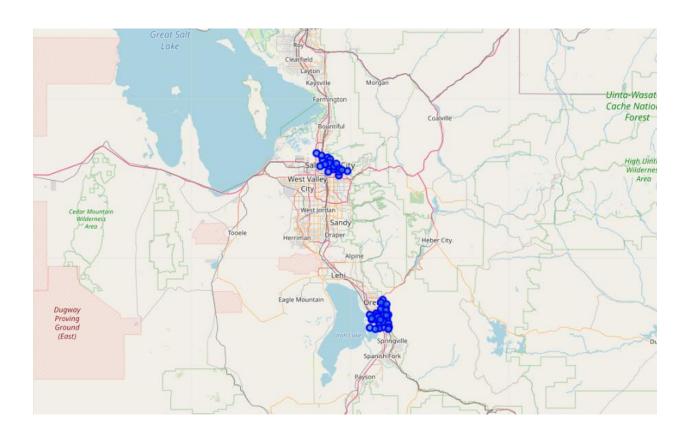
Cluster 4

[79]: provo_merged.loc[provo_merged['Cluster Labels'] ==3, provo_merged.columns[[0] + list(range(4, provo_merged.shape[1]))]]

[79]: Neighborhood 1st Most Common Venue 2nd Most Common Venue 3rd Most Common Venue 4th Most Common Venue 5th Most Common Venue 17 Pleasant View Electronics Store Mediterranean Restaurant Pizza Place Performing Arts Venue Park

4. Results and Discussion:

Salt Lake City and Provo neighborhoods are displayed on a single map to get a bird's eye view



Using the value_counts function the most popular venue categories are tabulated for each city. Coffee shops are very popular in Salt Lake City.

Most popular eating venue in Salt Lake is a Coffee shop



Pizza and sandwich places are very popular in Provo

Most popular eating venue in Provo is either a Pizza or Sandwich Place

```
[84]: provo_venues['Venue Category'].value_counts().head(20)
[84]: Park
                               10
     Construction & Landscaping
     Pizza Place 6
Sandwich Place 6
     Mexican Restaurant
                               5
                               5
     Burger Joint
     Chinese Restaurant
     Gym / Fitness Center
                              3
     Home Service
     Train Station
                               3
     Café
     Video Store
                               3
     Rock Club
                               3
     Latin American Restaurant
                              3
     Convenience Store
     Hotel
     Gas Station
     Bakery
     Bookstore
     American Restaurant
```

The best location to open a new business would be in a less saturated market and in a vibrant neighborhood. I would choose either Central City or the 9th and 9th neighborhood for a business location in Salt Lake City.

saltlake_venues['	Neighbo	orhood'].	value_c	ounts()
		_		
University	97			
Central City	62			
9th & 9th	24			
Ballpark	14			
15th and 15th	10			
Downtown	8			
Sugarhouse	8			
The Avenues	8			
Capitol Hill	8			
Federal Heights	8			
Marmalade	7			
Liberty Park	6			
Foothill	5			
East Bench	3			
Rose Park	3			
Westpointe	3			
Poplar Grove	3			
Yalecrest	3			
Fairpark	3			
East Central	2			

For the Provo area I would choose either the Riverbottoms or Carterville neighborhood.

Downtown	45
Riverbottoms	25
Carterville	14
Timp	12
North Park	8
Spring Creek	6
Maeser	5
Fort Utah	5
Franklin South	4
Dixon	4
Provost	3
Franklin	3
North Timpview	3
Foothills	3
Rock Canyon	2
Brigham Young University	2
Edgemont	2
Joaquin	2
Oak Hills	2
Provost South	2
Wasatch	2
Grandview South	2
Lakeview South	2
Pleasant View	1
Provo Bay	1
Lakeview North	1
Grandview North	1
Sherwood Hills	1

Before choosing a final location, I see that the 9th & 9th neighborhood is already saturated with coffee shops so I would suggest the Central City neighborhood as the final choice for the business location in Salt Lake City

		counts
Venue Category	Neighborhood	
Hotel	University	6
Sandwich Place	University	6
American Restaurant	University	5
Bakery	Central City	4
Italian Restaurant	Central City	4
Grocery Store	Central City	4
Cosmetics Shop	Central City	4
Boutique	Central City	4
Restaurant	University	3
New American Restaurant	University	3
Coffee Shop	University	3
	9th & 9th	3
Steakhouse	University	3
Bar	University	3
Mediterranean Restaurant	University	3

Going through a similar analysis, I would choose the Riverbottoms neighborhood for the Pizza place since Carterville is already saturated.

			counts
V	enue Category	Neighborhood	
Mexic	an Restaurant	Downtown	4
	Pizza Place	Carterville	2
	Bar	Downtown	2
S	andwich Place	Downtown	2
	Rock Club	Downtown	2
Construction 8	& Landscaping	Spring Creek	2
	Burger Joint	Riverbottoms	2
	Park	North Park	2
Latin Americ	an Restaurant	Downtown	2
	Trail	North Timpview	2
	Bookstore	Downtown	2
Brazil	ian Restaurant	Riverbottoms	2
	Train Station	Franklin South	2
	Hotel	Carterville	2
As	ian Restaurant	Downtown	2

5. Conclusion:

Results show that coffee shops are most popular venue category in Salt Lake and Pizza and Sandwich Places are most popular venue category in Provo. Based on our analysis we recommend opening a **coffee shop** in the **Central City neighborhood of Salt Lake** and a **pizza or sandwich place** in the **Riverbottoms neighborhood of Provo**. To further refine the results additional research and analysis can be done on the demographics of the selected locations