

Business Venture Consultation

Business Problem Introduction:

I am a data scientist at a consulting firm that specializes in advising people on starting new business ventures. My target audience is one of the firm's clients who is interested in opening a business in the state of Utah in the United States. Our client has heard that Utah is a very welcoming to entrepreneurs and new businesses.

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. Our client needs advice on the best type of business to open in each of counties and decide on which neighborhood in each county would be the best site to open this business.

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.

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

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

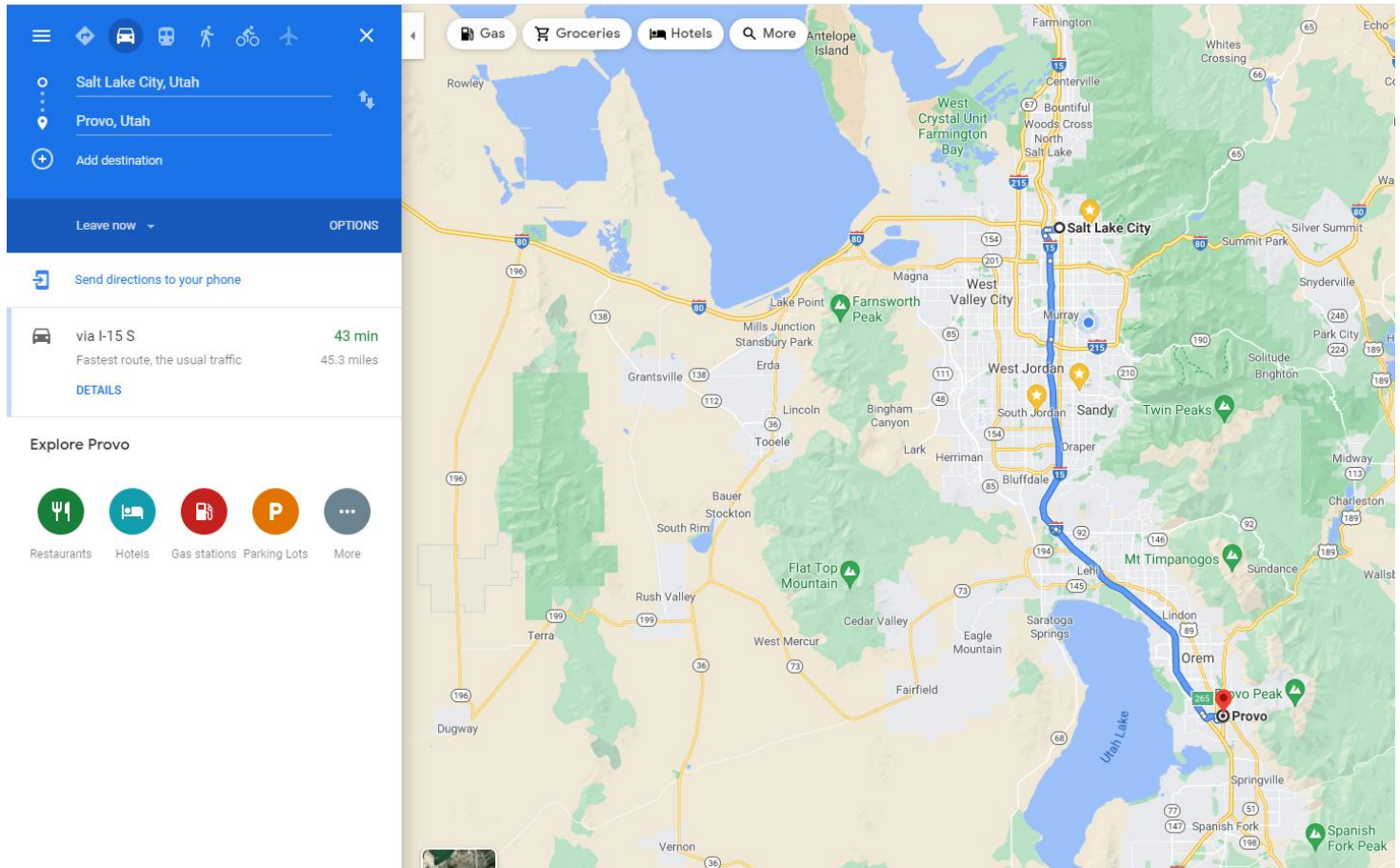
Here's a screenshot listing Utah counties by population

Rank	County	Population
1	Salt Lake County	1,133,646
2	Utah County	605,490
3	Davis County	345,767
4	Weber County	251,498
5	Washington County	165,811
6	Cache County	124,165
7	Tooele County	67,397
8	Box Elder County	53,946
9	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

My data science project entails analyzing the location data for all the neighborhoods in Salt Lake City and Provo. I will then cluster the neighborhoods within each city independently to figure out what type of business would work based on the existing businesses.



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

Questions to Address:

1. Which city and neighborhood within that city would be best suited for starting a business venture?
2. What type of business would do well once we decide on a location?

Data Acquisition & Cleaning:

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 be scraped 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>

The scraped data is cleaned and converted to a pandas data frame

Neighborhoods	
0	CARTERVILLE
1	DIXON
2	DOWNTOWN
3	EAST BAY
4	EDGEMONT
5	FOOTHILLS
6	FORT UTAH
7	FRANKLIN
8	FRANKLIN SOUTH
9	GRANDVIEW NORTH
10	GRANDVIEW SOUTH
11	INDIAN HILLS
12	JOAQUIN
13	LAKEVIEW NORTH
14	LAKEVIEW SOUTH
15	LAKEWOOD
16	MAESER
17	NORTH PARK
18	NORTH TIMPVIE
19	OAK HILLS
20	PLEASANT VIEW
21	PROVO BAY
22	PROVOST
23	PROVOST SOUTH
24	RIVERBOTTOMS
25	RIVERGROVE
26	RIVERSIDE
27	ROCK CANYON
28	SHERWOOD HILLS
29	SPRING CREEK
30	SUNSET
31	TIMP
32	UNIVERSITY
33	WASATCH

Then using the geolocator API the latitude and longitude of the neighborhoods are obtained

	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, Provo	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, Lakeview, Provo	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, Caryhurst	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, Provo	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, Provo	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

The neighborhood, latitude and longitude data will be used with the Foursquare API to give the most common venues in each neighborhood.

The neighborhoods will be clustered using the K-Means algorithm and based on the most common venue features in each neighborhood/cluster insights can be drawn as to where and what type of business venture should be started. We are looking for places where business will flourish-right location (for e.g. having restaurants in a non-residential neighborhood) and where there is not too much saturation for that particular type of business (e.g. too many coffee shops in a particular cluster)

[0]:	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	Cartersville	40.259880	-111.668621	2.0	Hotel	Pizza Place	Discount Store	Supermarket	Jewelry Store
1	Dixon	40.237168	-111.677765	2.0	Restaurant	Home Service	Campground	ATM	Mexican Restaurant
2	Downtown	40.233294	-111.661315	2.0	Mexican Restaurant	Sandwich Place	Asian Restaurant	Latin American Restaurant	Bar
3	East Bay	40.212666	-111.650474	2.0	Gym / Fitness Center	Mediterranean Restaurant	Performing Arts Venue	Park	New American Restaurant
4	Edgemont, Provo	40.286066	-111.651866	3.0	Massage Studio	ATM	Men's Store	Pizza Place	Performing Arts Venue
5	Foothills	40.238850	-111.634058	2.0	Water Park	Intersection	Theme Park	Taco Place	Mediterranean Restaurant
6	Fort Utah	40.237069	-111.702384	2.0	Farm	Airport Terminal	Construction & Landscaping	Park	Brewery
7	Franklin	40.228002	-111.670665	2.0	Train Station	Rock Club	Chinese Restaurant	ATM	Men's Store
8	Franklin South	40.222205	-111.666880	0.0	Construction & Landscaping	Train Station	Furniture / Home Store	ATM	Mexican Restaurant
9	Grandview North	40.257786	-111.684236	0.0	Construction & Landscaping	ATM	Plaza	Pizza Place	Performing Arts Venue
10	Grandview South	40.249176	-111.687033	1.0	Park	Clothing Store	ATM	Plaza	Pizza Place
11	Joaquin	40.240187	-111.650371	2.0	Dessert Shop	Snack Place	Gas Station	ATM	Pizza Place
12	Lakeview North, Lakeview, Provo	40.253732	-111.710958	1.0	Park	ATM	Men's Store	Pizza Place	Performing Arts Venue
13	Lakeview South	40.241539	-111.703080	0.0	Playground	Construction & Landscaping	Farm	Plaza	Pizza Place
14	Maeser	40.229487	-111.649537	0.0	Video Store	Construction & Landscaping	Gas Station	Discount Store	ATM
15	North Park	40.245514	-111.666089	2.0	Pizza Place	Park	Business Service	Gym / Fitness Center	Sandwich Place
16	North Timpview, Caryhurst	40.305259	-111.655147	2.0	Trail	Park	Financial or Legal Service	ATM	Men's Store
17	Oak Hills	40.254138	-111.629868	2.0	Scenic Lookout	Food	ATM	Playground	Performing Arts Venue
18	Provo Bay	40.211185	-111.711363	2.0	Airport Service	Mediterranean Restaurant	Performing Arts Venue	Park	New American Restaurant
19	Provost, Provo	40.225441	-111.630560	2.0	Massage Studio	Park	Home Service	Mediterranean Restaurant	Performing Arts Venue
20	Provost South	40.211571	-111.625508	2.0	Convenience Store	ATM	Plaza	Pizza Place	Performing Arts Venue
21	Riverbottoms	40.297065	-111.661324	2.0	Water Park	Sushi Restaurant	Kitchen Supply Store	Kids Store	Lingerie Store
22	Rivergrove	40.247876	-111.675328	1.0	Park	ATM	Men's Store	Pizza Place	Performing Arts Venue
23	Rock Canyon	40.269529	-111.639238	2.0	Playground	Park	Music Venue	Music Store	Men's Store
24	Sherwood Hills	40.289842	-111.638113	2.0	Food	Playground	ATM	Men's Store	Performing Arts Venue
25	Spring Creek	40.205466	-111.628930	0.0	Convenience Store	Construction & Landscaping	Furniture / Home Store	ATM	Movie Theater
26	Timp	40.238421	-111.664827	2.0	ATM	Snack Place	Laser Tag	Video Store	Latin American Restaurant
27	Brigham Young University	40.255011	-111.649657	2.0	Lake	Science Museum	ATM	Men's Store	Performing Arts Venue
28	Wasatch	40.252979	-111.638216	2.0	Sandwich Place	ATM	Men's Store	Performing Arts Venue	Park