

Introduction

Utah is one of the nation's great outdoor states with incredible national parks, such as Arches National Park, Zion National Park, etc, and is also well known of its world famous ski area, such as Snowbird, Park City Mountain Resort, etc. Since there are millions of travelers coming to Utah each year for its incredible national parks and ski resorts, opening a business targeting travelers and residents will be a good idea. However, since these beautiful nature attractions are all over the place in Utah, researches on existing common venues near each attraction are necessary. Therefore, this study aims to utilize Foursquare website to find out and cluster the most common and popular existing venues near the Utah national and state parks. Additionally, the population of Utah Counties are collected to study how the local population affects the choices of venues near the attractions.

Data description

1. A list of names of Utah National and State Parks are collected. Nine Utah national attractions are collected from National Park Service Website. A list of Utah State Parks are web scrapped from Wikipedia.
2. A list of Geographical coordinates of Utah national and state parks are generated using geopy.
3. Foursquare API is utilized to generated the most common venues near each attractions.
4. A list of Utah County names and population are extracted from United States Census Bureau website.
5. A geojson file contains geographic information of Utah counties is collected from opendatasoft website.

Methodology

The information of Utah national and state parks are collected from National Park Service Website and Wikipedia to create a combined list. The list named `UT_parks` has a total of 48 attractions. Each attraction is passed to geopy to compute its latitude and longitude coordinate. A brief example of the data is shown in Table 1.

Table 1: An example of UT national parks with latitude and longitude coordinates.

	UT_parks	latitude	longitude
0	Arches National Park	38.731557	-109.563028
1	Bryce Canyon	37.618409	-112.142234
2	Canyonlands National Parks	38.347889	-109.858351
3	Capitol Reef National Park	38.067029	-111.155256
4	Cedar Breaks National Monument	37.642639	-112.844723

A python library called Folium is then utilized to visualize the geographic details of Utah and the blue markers on the map show the location of attractions as shown in Figure 1.

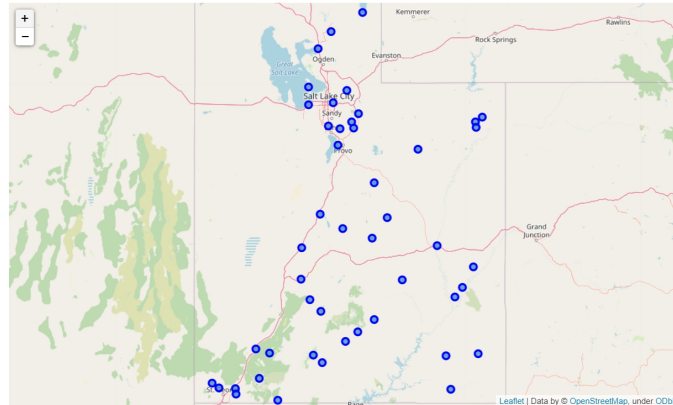


Figure 2: A map of Utah with markers indicating nature attractions.

A geoJson file is collected from Opendatasoft website containing details of Utah counties boundary. The boundary data is used to create a choropleth map of Utah counties with population collected from US Census Bureau. The choropleth map is created by using Folium library as shown in Figure 2.

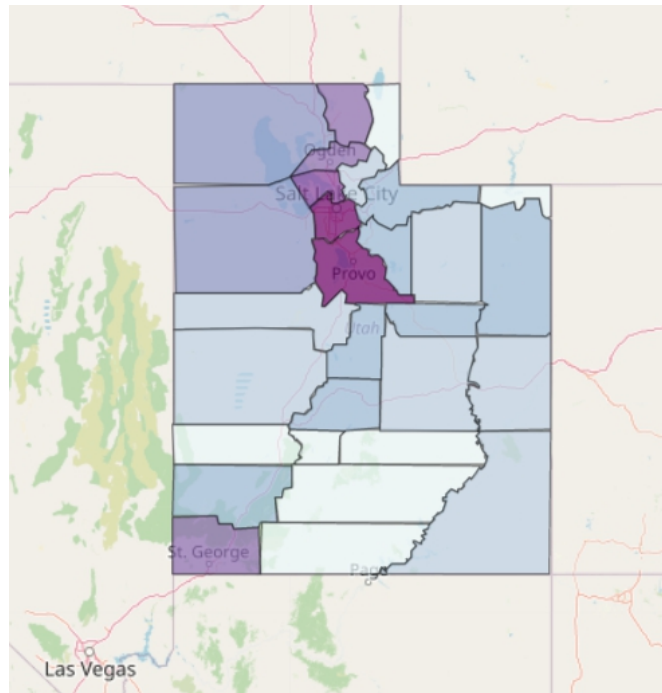


Figure 3: A choropleth map showing Utah counties with local population density.

The venues near each attractions are found by utilizing Foursquare API. There are a total of 164 unique categories found by Foursquare. Then 10 most common venues are found for each attraction. K-mean clustering is used to cluster the attractions into groups. 5 clusters are found to be the optimum choice. The cluster result is shown in Figure 3.

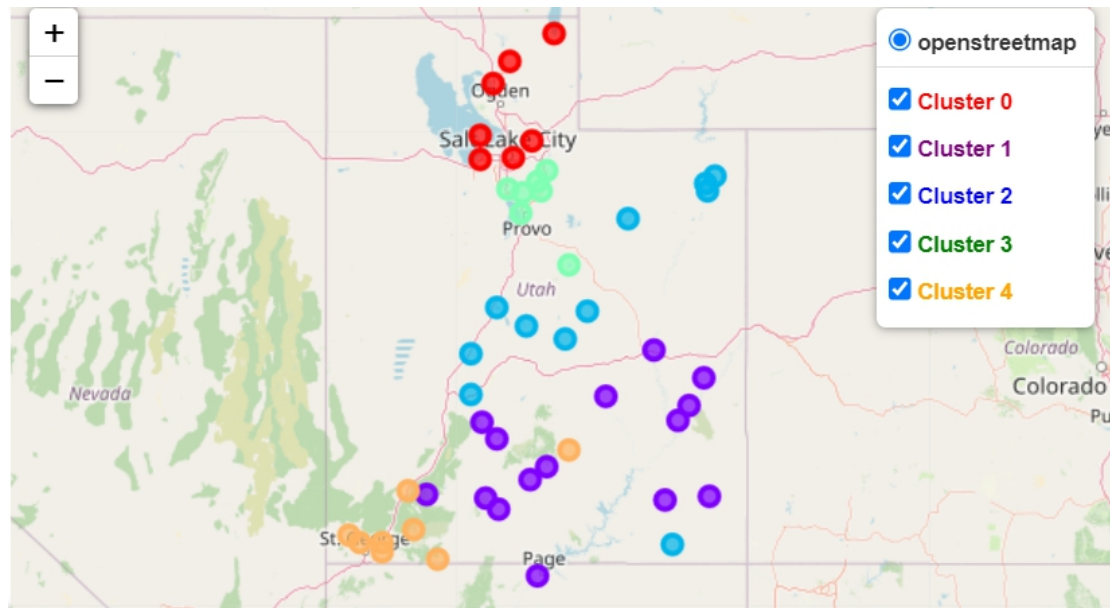


Figure 4: Attractions clustering results.

The cluster result is also plotted on the choropleth map to better visualize the relationship between the local population and the venue cluster results as shown in Figure 4.

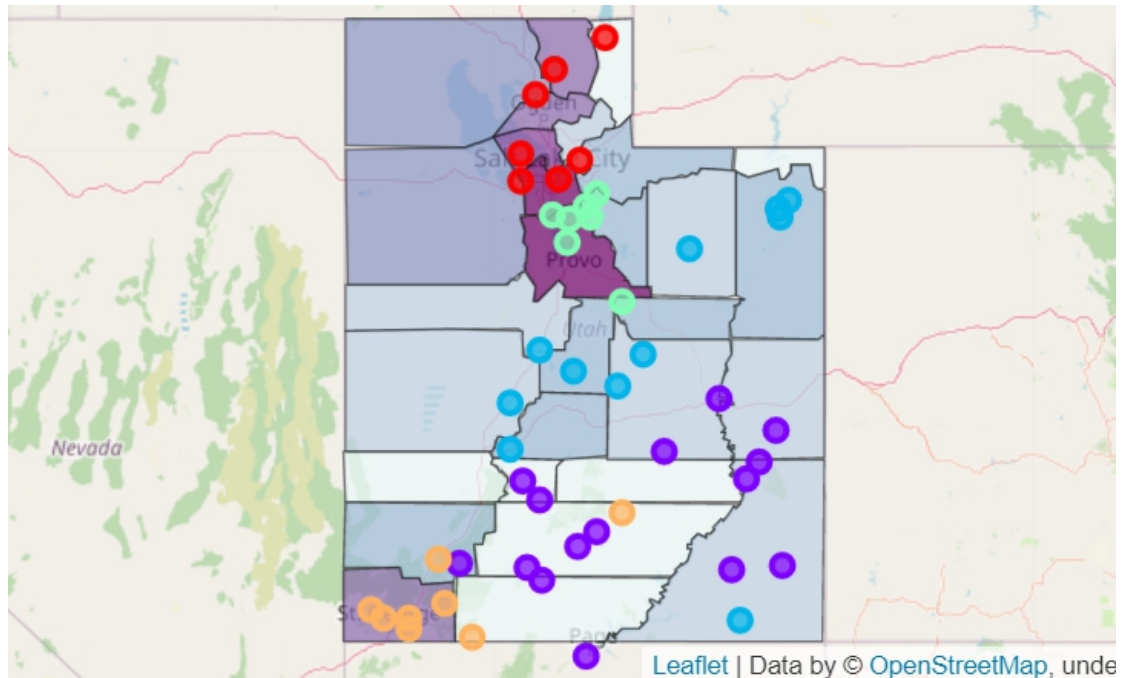


Figure 5: Attractions clustering results on Utah choropleth map.

Results

From Figure 4, it clearly shows a strong relationship between local population and existing common venues. After clustering the common venues near all famous Utah attractions, the types of common venues are significantly affected by the local population. Table 2 shows the most common venues for Cluster 3 (green markers). The 1st most common venue for Cluster 3 is Ski Area which is reasonable because all of Utah's top ski areas are located in Salt Lake County and Utah County which also has the largest local population density. Figure 4 clearly indicates that Cluster 3 is located at the largest population areas.

Table 2: The most common venues for the Cluster 3 group.

	Parks	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
7	Timpanogos Cave	Ski Area	Park	Grocery Store	Bakery	Ski Lodge
14	Deer Creek State Park	Ski Area	Park	American Restaurant	Café	Mexican Restaurant
18	Flight Park State Recreation Area	Park	Coffee Shop	Ski Area	Fast Food Restaurant	Grocery Store
28	Jordanella State Park	Ski Area	Coffee Shop	Park	Mexican Restaurant	Grocery Store
37	Scofield State Park	Ski Area	Grocery Store	Fast Food Restaurant	Bakery	Park
44	Utah Lake State Park	Ski Area	Fast Food Restaurant	Coffee Shop	Grocery Store	Ski Lodge
45	Wasatch Mountain State Park	Ski Area	Park	Mexican Restaurant	Fast Food Restaurant	American Restaurant

Table 3 shows the most common venues for Cluster 1 (purple markers). The most common venues of Cluster 1 are Scenic Lookouts and Trails which are also reasonable because Cluster 1 has nearly all of the national parks grouped such as Arches National Park, Bryce Canyon, and Canyonlands National Parks. Since the parks in Cluster 1 are located in the least local population, the most common business are Hotel and American Restaurant.

Table 3: The most common venues for the Cluster 1 group.

	Parks	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Arches National Park	Scenic Lookout	Trail	American Restaurant	National Park	Tourist Information Center
1	Bryce Canyon	Scenic Lookout	Trail	American Restaurant	Hotel	National Park
2	Canyonlands National Parks	Scenic Lookout	Trail	American Restaurant	Hotel	National Park
4	Cedar Breaks National Monument	Scenic Lookout	Trail	Hotel	Resort	National Park
5	Glen Canyon	Scenic Lookout	National Park	Hotel	Trail	American Restaurant
6	Natural Bridges National Monument	Scenic Lookout	National Park	Motel	Hotel	Gas Station
9	Anasazi State Park Museum	Scenic Lookout	Trail	Hotel	American Restaurant	National Park

Discussion

Even though the results suggest a strong relationship between the local population and the common venues near the attractions, it's difficult to visualize the optimum choice of venue types. Therefore, Figure 5 is created to better visualize the relationship. Since public venues cannot become choices of personal business, venues such as Scenic Lookout, Trail, and Tourist Information Center are excluded from the plot. Figure 5 shows the percentage of each common venue with size of marker indicating the population density.

From figure 5, it suggests that Coffee Shop is the best choice to start a business at the location with the largest population. The second best choice is the Mexican Restaurant. For the location with the least population, Hotel is the best choice for a new business and American Restaurant is the second best choice. Hotel is also the best choice at the location with moderate population. Bar, Plaza, and Golf Course are the three worst choices for a new business in Utah regardless of local population.

The take away messages found from Figure 5 are very reasonable. In the area with larger population, many venues are popular and share similar percentage. For example,

although Coffee shop has the highest percentage which is approximate 35%, Mexican Restaurant, Fast food Restaurant and Grocery Store are also very popular with around 20~30%, which means people more likely prefer various options of venues in these areas. On the other hand, in the area with less population, the options of venues are more distinct. Hotel and American Restaurant are the only two options in these areas with has 45% and 38%, respectively. Gas Stations, Resorts, and Plazas only have less than 10% of percentage in these areas.

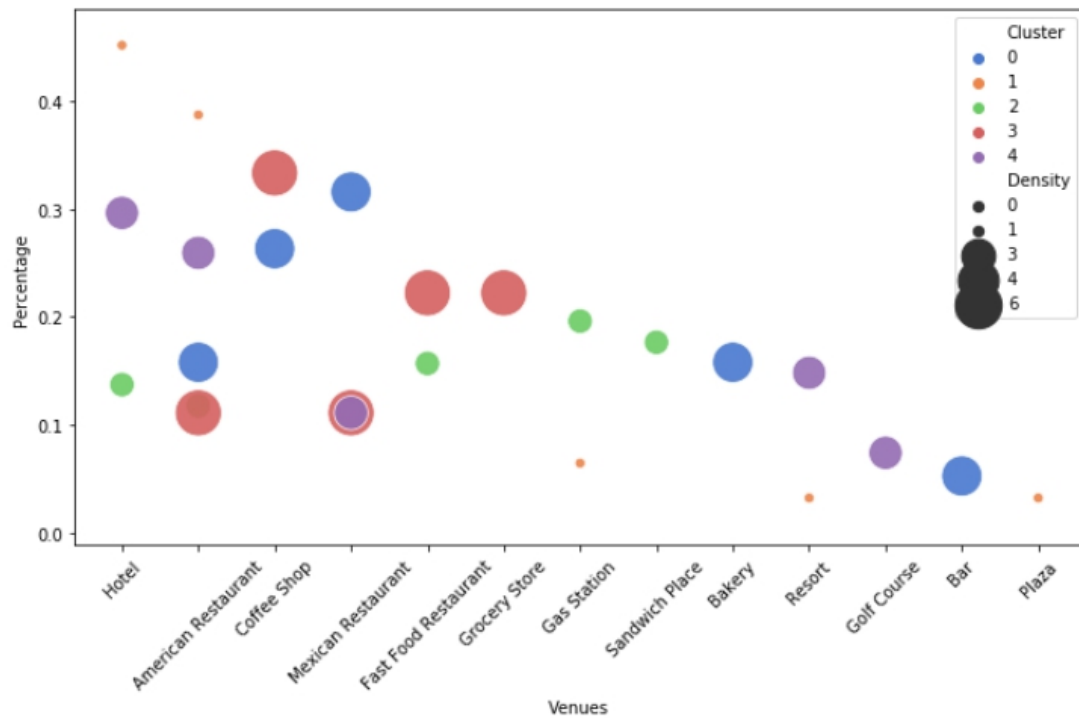


Figure 6: The relationship between the percentage of venues and venue types with marker size indicating population density and color indicating cluster number.

Conclusion

This study focuses on investigating the popular options for opening a new business mainly targeting travelers and local resident in Utah near nature attractions such as national and state parks. Existing common venues are collected by using Foursquare API and clustered to visualize the relationship between population and the types of venues. The results suggest that the success of opening a new business is significantly based on the local of the business due to population density. Coffee shop and Restaurants are the best opinions to open a business in the area with large population. Hotel and American Restaurant are the best opinions to open a business in the area with less population.