1. **Introduction**

1.1 Background

The city of Ogden, UT has moved from being a slightly worn-down backwater region to a thrumming, thriving city in just a few decades. Perfectly situated a short distance from ski resorts, and considered by some to be a distant suburb of Salt Lake City (aided by the Frontrunner passenger train which runs through it), Ogden has become quite the destination location.

1.2 Problem

With greater movement, both from pedestrians and vehicles, both from natives and neighbors, the need to feed the horde becomes greater. Setting up shop in the wrong place, however, can mean huge customer acquisition costs for a restaurant, regardless of the city. Unless a new restaurant can quickly become popular on its own and stand apart from the crowd as a destination spot itself, the right location is absolutely critical.

1.3 Interest

Obviously, aspiring restaurant owners in Ogden, UT would be interested in knowing what the overall pattern is like for opening up a restaurant. Seeing such trends, and comparing against any outliers, would give a new restaurant owner a good idea of where they want to open up their establishment.

1. **Data**

2.1 Data Sources

All the data from this project comes from a geolocator app called Foursquare, whose API is pretty robust and easy to work with. Geolocations of all the sit-down restaurants in Ogden, UT were pulled from Foursquare, 43 in total, within a 5-mile radius of the city center. Said city center was pinpointed at 2401 Washington Boulevard. It should be stated that there are a couple of restaurants missing from this dataset, but the exceptions are few. It should also be stated that fast-food restaurants were not considered nor included in this project, and all projections, calculations, and conclusions refer to sit-down restaurants only.

2.2 Data Cleaning

Data pulled from Foursquare’s API was loaded into one table as a Python dataframe. Fortunately, all the restaurants queried from Foursquare had names, fully formatted addresses, and, most importantly, latitude/longitude (location.lat/location.lng) geolocations. There were, however, several unnecessary columns, which I removed from the dataframe. A full list of those columns and their meanings are as follows:

|  |  |
| --- | --- |
| **Column Dropped** | **Reason for Dropping** |
| id | This is the Foursquare ID, not important to this analysis. |
| categories | Contained redundant information already specified in the other columns |
| referralId | This was the same code (v-1595907793) for each and every result, meaning this was the code for this particular data pull. Removed due to perfect redundancy. |
| hasPerk | This was a dummy variable which was False for all results. Removed due to perfect redundancy. |
| location.labeledLatLngs | This contained information already specified in the location.lat and location.lng columns. Removed due to redundancy. |
| location.distance | A metric which showed how far away a particular data point was from the chosen city center. Since this analysis focuses on phenomena around individual datapoints and not distance from the city center, this was removed. |
| location.address/location.city/  location.state/location.country | These columns showed the components of the address of a particular restaurant, but since this was already included in the fully formatted address field, these columns were removed. |
| location.crossStreet | This field proved to be null for the majority of values, and provided a code irrelevant to geolocations. Removed for insignificance. |
| venuePage.id | This field proved to be null for the majority of values, and provided a code irrelevant to geolocations. Removed for insignificance. |
| delivery.id/ delivery.url/  delivery.provider.name/  delivery.provider.icon.prefix/  delivery.provider.icon.sizes/  delivery.provider.icon.name | These values all hinted at any delivery providers partnered with a given restaurant. This list may have indeed risen due to the pandemic to include more partners since last this data was compiled in Foursquare. For this analysis, however, since delivery partners can extend the reach of a given restaurant just the same as any other, this field was removed. |

1. **Methodology**

3.1 Initial Visualization

With the cleaned dataset, all the restaurants were placed onto a Folium map according to latitude/longitude geolocation to see what the overall spread looked like. As you can see in Figure 1, the spread is not entirely even, but there are several different clusters of restaurants.

Figure 1. Initial Visualization Result

A close up of a map

Description automatically generated

3.2 Clustering

The clusters needed to be isolated to study them more closely, so restaurants were placed on a graph by latitude and longitude. The DBSCAN method was then used to hone into these different clusters and separate them from their outliers, at an epsilon of 0.013795 (so small because the space between ticks on both the x [latitude] and y [longitude] axes were small) and a minimum grouping of 5. Each cluster was given a color, and any outliers were colored black. Clearly seen are the formation of three separate, distinct clusters. See Figure 2.

Figure 2.

A close up of a piece of paper

Description automatically generated

3.3 Final Visualization

In order to carry on further analysis, the data points needed to be returned to the map, and clearly marked by whether they made it into a cluster or not. This was done by creating two different dataframes, or tables, each with a new column, titled ‘core\_members\_mask’. This new column would contain a Boolean value, True meaning a given restaurant belonged to one of three clusters, and False meaning a given restaurant was not included in any cluster. Then the datapoints were returned to the map through Folium, and for simplicity’s sake, any data points housed in a cluster was placed on the map as a red pin, and any outliers were colored blue (as seen in Figure 3). Since these three clusters were far enough apart from one another to clearly define separation, no additional coloration was necessary.

Figure 3.

A close up of a map

Description automatically generated

1. **Results**

4.1 The Core Data

The three separate cluster groups were dubbed the Northern, Central, and Southern Clusters. Upon closer examination of the map, we note that the Southern Cluster exists near a shopping mall. The Middle Cluster exists near "the Junction," or the city center, which, coincidentally, also has many places for shopping, not to mention plenty of activities for kids and adults, and is also situated next to the Ogden Temple, widely considered the most important religious site in the city. The Northern Cluster exists at the crossroads of two high-movement streets, where also stands a popular gym, a Shopko, and a few convenience stores.

4.2 The Outliers

The DBSCAN returned 13 data points that did not fit into any of the three clusters. Further scrutiny of the area surrounding these Outliers reveals that they are much more similar to the clusters than one might think, situating themselves near smaller pockets of high-movement areas, whether that be near a museum, office space, apartment complex, etc.

1. **Discussion**

5.1 Go Where The People Are

Opening up shop near one of those three clusters seems to be a pretty safe bet for those who can afford the real estate prices, since those are high-movement areas. More importantly, all three of those high-movement areas are probably going to remain high-movement areas during and post-pandemic, because the attractions at their core are going to keep attracting people regardless. Now, more than ever before, it’s critical to open your restaurant near where the people go, or you’ll quickly struggle with customer acquisition costs.

5.2 Dealing with Competition

The question remains, however, whether moving your new restaurant into an area that’s already saturated with restaurants is wise. The competition for customers might be too fierce, one could argue, and for a new restaurant owner without an established name, said competition could kill your business. While that competition does exist on a certain level, one might find that building a restaurant next to an established culinary group might actually be a great way to market your business, because you benefit from the visibility. Think about it, your target market as a new restaurant owner is people who go to restaurants regularly. Where do those restaurant goers go? To the place where the restaurants are, of course. People like options. They want to be able to try new things, while at the same time staying close to what is familiar to them. What better way is there to advertise your new restaurant to restaurant goers than to include your edibles in the mix of available options? Instead of thinking that you’re trying to compete with their favorite place, customers will likely be grateful to you for providing them with a new experience.

5.2 Looking Elsewhere

That leaves the Outliers. One would think that for being Outliers, that they would follow a different set of rules than the data points of the core clusters. However, in the case of almost all of the Outliers, the vast majority are near high-movement areas too, albeit smaller pockets of movement. One Outlier restaurant is situated at the entrance of a dinosaur park for kids, which I personally visited not long ago. Another Outlier is placed in a small strip mall next to a grocery store. Another is right next to a hospital, and the list continues. Though these restaurants don’t benefit from being in one of the three restaurant hubs, they each have a specific reason for being built where they are.

1. **Conclusion**

If someone is considering opening up a restaurant in Ogden, UT, and can afford the real estate prices, said someone might do well to open up a restaurant in one of the city’s three restaurant clusters, because you will benefit from the movement of the main attractions as well as from being included in one of the restaurant hubs. If you can’t afford the real estate prices of city-center hosting, or if you want to open your restaurant in another location in the city of Ogden, then you might do well to build your establishment next to a smaller, yet still high-movement, pocket of the city to keep your customer acquisition costs low.