Exploring Baton Rouge with Foursquare

Applied Data Science Capstone – Coursera by IBM – July 2021

Introduction

Since March 2020, the COVID-19 restrictions have shifted the way we work with employees working remotely from their homes. However, as these restrictions are lifted, employees are once again called back into the central offices; and those who accepted remote positions may now have to relocate. For the batch of new hires at my company in Baton Rouge, relocation will be a necessity; but many of us have never been to Baton Rouge and are unsure of the ideal location to look for housing based on our interests. I will use Foursquare's location-based data to explore the neighborhoods in Louisiana's capital city to offer my colleagues a comprehensive look at ideal areas to look for housing based on the venues available and overall neighborhood dynamic.

Data Exploration

The names of each neighborhood in Baton Rouge were obtained from Wikipedia[1] but I could not get the Geocoder to pull up the latitude and longitude for each neighborhood, so this data was collected through a general online search and manually added to a CSV file before being imported and merged with the original data frame. The resulting data frame has three columns and thirty rows, displaying the neighborhood's name, its latitude, and its longitude. I will use this data to retrieve the venue details from Foursquare and cluster the neighborhoods based on its top venue types.

Methodology

In this section, I will outline the steps taken to collect the data and create the clusters based on the most popular venue categories in the Baton Rouge neighborhoods.

Building the Data

I used Python's Beautiful Soup package to scrape the names of each neighborhood from Wikipedia[1] and created a new data frame with this data. There were 30 neighborhoods in total listed on the Wikipedia page and they are listed below for reference.

Downtown	Mid-City	Goodwood	Scotlandville	Park Forest
Spanish Town	McDonald Land	Hundred Oaks Addition	Shenandoah	Glen Oaks
Beauregard Town	Brookstown	Southdowns	Sherwood Forest	University Club
Garden District	Melrose Place	Gardere	Brownfields	Centurion Place
Old South Baton Rouge	Melrose East	Westminster	Zion City	Northdale
LSU/Lakeshore	Inniswold	Oak Hills Place	Monticello	Broadmoor

[1] https://en.wikipedia.org/wiki/Neighborhoods_in_Baton_Rouge,_Louisiana

The next step was to gather each neighborhood's latitude and longitude. Unfortunately, the Geocoder package was not working so I had to manually find the latitude and longitude through an online Google search and created an Excel file with the details. The only neighborhood where I could not find the latitude and longitude was 'Hundred Oaks Addition', as this turned out to be a retirement community; it was dropped it from the final data frame.

The image below shows the map of Baton Rouge with the neighborhoods marked based on their respective latitude and longitude.
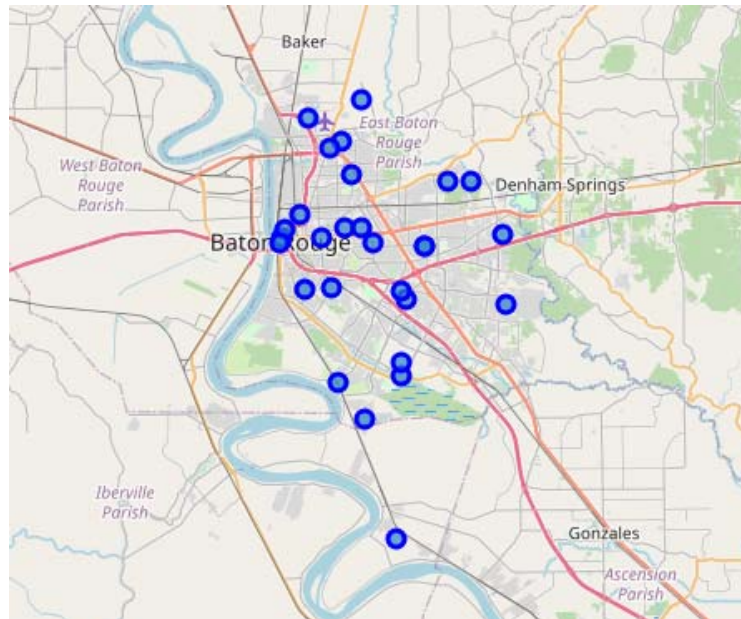

Figure 1: Map of Baton Rouge with Neighborhoods

Venue Details from Foursquare

Now that the neighborhood names and their respective locations are determined, it is time to get the top venues in each neighborhood. The request to Foursquare's API was for a maximum of 100 venues in each neighborhood but the most obtained was 42 (Downtown). Below is a table listing the four neighborhoods with the highest number of venues.

Neighborhood	Count
Downtown	42
Garden District	41
Beauregard Town	32
Spanish Town	17

The remaining neighborhoods had less than 10 venue types, four of which only had a single venue type. Overall, there were 93 unique venue categories among all neighborhoods. There were four neighborhoods that did not produce a result from Foursquare for a venue type, indicating that they are

purely residential with no venues or services within a 500-meter radius. These rows were dropped from the data frame.

The final step in preparing the data was to one-hot-encode the Venue Category column to calculate the frequency of venue type for each neighborhood and determine the top 10 venue categories among all neighborhoods. Below is a list of the top venue categories – many are repeated so I have only included them once.

- Home Service
- American Restaurant
- Intersection
- Trail
- Flower Shop
- Construction and Landscaping
- Convenience Store
- Dance Studio
- Discount Store
- Distillery
- Donut Shop

Clustering the Neighborhoods

The KMeans() function was used to separate the neighborhoods into clusters based on their top 10 most frequent venue categories. Originally, I had thought to separate the neighborhoods into 5 clusters, but it appeared that most neighborhoods were getting assigned to a single cluster, so I increased the number of clusters to 8. The table below shows each cluster number and the neighborhoods assigned to it; and most neighborhoods (12) are found in Cluster 3.

Cluster	Value Count
3	12
2	5
4	2
1	2
7	1
6	1
5	1
0	1

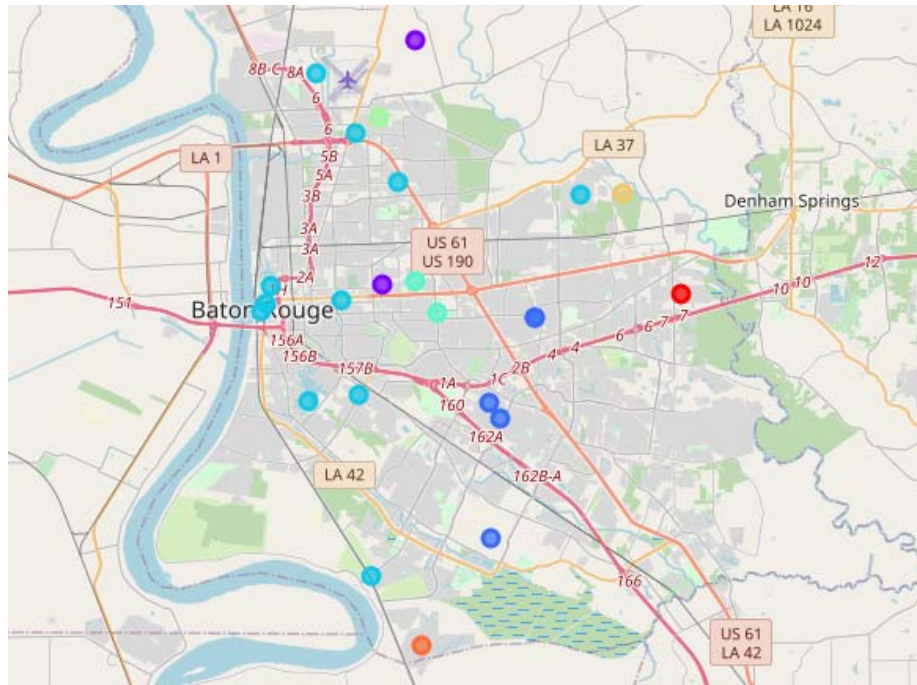Below is a map showing the clusters where each neighborhood was assigned.

Figure 2: Map of Baton Rouge with Clusters

Results

Cluster 3 is the largest with twelve total neighborhoods included, followed by Cluster 2 with five neighborhoods and Clusters 4 and 1 with two neighborhoods each. The remaining clusters are made up of a single neighborhood.

The twelve neighborhoods in Cluster 3 are: Downtown, Spanish Town, Beauregard Town, Garden District, LSU/Lakeshore, Mid-City, Brookstown, Southdowns, Gardere, Scotlandville, Monticello, and Glen Oaks. The top venues in this cluster are focused on food and active pursuits; and the table below shows the top 5 venue categories for this cluster:

Top Venues	
1	Sandwich Place
2	American Restaurant
3	Park, Trail, College Gym, Pool
4	Food
5	Fishing Store, Music Venue, College Gym

The second-largest cluster is Cluster 2 with five neighborhoods: Inniswold, Oak Hills Place, Broadmoor, Sherwood Forest, and Westminster. The top venues in this cluster do not seem to have a particular focus but looking at the map of Oak Hills Place, I can see that it is a large shopping area with the Louisiana Mall and a variety of large box stores. The table below shows the top 5 venue categories for this cluster:

Top Venues	
1	Home Service (Shops & Service)
2	Intersection (Travel & Transport)
3	Trail
4	Flower Shop
5	Construction and Landscaping

I did not break down the top venue categories for the other six clusters.

Discussion

For those looking for a neighborhood with a variety of activities and services, those in Cluster 3 offer the widest array of options with a mix of food venues and activity-based venues. The neighborhoods in Cluster 2, the second largest cluster, are best for those who prefer to be surrounded by the box stores; or those who need easy access to the interstate for a commute or frequent travel since 'Intersection' is the second most common venue.

Conclusion

Overall, Baton Rouge is a great city to live in; but before relocating, it is important to think about one's personal interests because some neighborhoods offer a variety of venue types (i.e. food and activities) while others are dominated by a select 1-2 venue types. Based on my preferences, the neighborhoods in Cluster 3 would be the best option for me since I want proximity to restaurants while still being able to keep active through multiple activities. However, additional factors such as housing availability and proximity to favorite venue types should also be considered before making the final choice on a neighborhood.

References

[1] https://en.wikipedia.org/wiki/Neighborhoods_in_Baton_Rouge,_Louisiana