Identifying neighborhoods for a new restaurant in Atlanta, GA

Elizabeth Niese

June 22, 2020

1 Background

For this project I am interested in opening a contemporary casual restaurant in Atlanta, GA with a focus on locally sourced food. With Atlanta's increasing population of educated professionals and the current interest in sustainable food, there should be at least one neighborhood in Atlanta that can support such a restaurant. To identify an optimal location for this restaurant I will need to look for areas of the city that has other mid to high end restaurants, indicating that local residents have the financial resources needed to eat out regularly, but is not oversaturated with such options. Another issue that would need to be addressed prior to opening such a restaurant is finding sufficient local farmers to purchase food from,but that is outside the scope of this project.

2 Data

The neighborhood data is available at

https://opendata.arcgis.com/datasets/d6298dee8938464294d3f49d473bcf15_196.geojson.

This data set includes neighborhood information such as name, statistical area, basic demographics, and neighborhood boundaries. To identify an appropriate neighborhood for the restaurant we will need to

1. identify the geographic location of each neighborhood,

2. use theFoursquare API to find venues in each neighborhood, and

3. use k-means clustering to identify which neighborhoods are likely to support the restaurant.

3 Methodology

After the initial download of the json file, certain features were of interest. These were the Statistical Area Code, neighborhood name, and latitude and longitude for each neighborhood. The geographic information was given as boundary coordinates, so to find a latitude and longitude for each neighborhood, the centroid of the polygon was computed. This is done by taking the average of each coordinate. A few neighborhoods had boundaries that were not amenable to this process. These were assigned a latitude and longitude of NaN and dropped later in the data-cleaning process.

	Neighborhood	African Restaurant	American Restaurant	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	...	Science Museum	Seafood Restaurant	Southern / Soul Food Restaurant	Sushi Restaurant	Tapas Restaurant	Tex-Mex Restaurant	R...
1	Peachtree Heights East, Peachtree Hills	0	1	0	0	0	0	0	0	0	...	0	0	0	0	0	0	
3	Peachtree Heights East, Peachtree Hills	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	
5	Peachtree Heights East, Peachtree Hills	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	
10	Peachtree Heights East, Peachtree Hills	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	
12	Peachtree Heights East, Peachtree Hills	0	1	0	0	0	0	0	0	0	...	0	0	0	0	0	0	

5 rows × 44 columns

One-hot encoding


Visualization of elbow method

The Foursquare API was used to find venue information for each neighborhood. Since the goal of this project is finding a good location for a restaurant, once the venue information was loaded into the dataframe, venues were restricted to those that contained the words *restaurant, museum, art*, and *studio*. These choices were made to help find a neighborhood which can support a higher-end restaurant and has other entertainment venues likely to draw patrons to the neighborhood.

One hot encoding was used to convert data into numeric form to prepare for k means clustering and can be seen in the table excerpt.

To determine the number of clusters that would be best, I used the elbow method available in the yellowbrick package. This method iterates k means clustering for a variety of values of k and determines the optimal number of clusters for the data. Using the elbow method I determined that 13 clusters would be used.

K-means clustering was used on the neighborhoods and the clusters were plotted on a map.

After mapping the clusters, I also viewed each cluster and the top venues for neighborhoods in that cluster. This allowed me to identify clusters containing venues that I want near my restaurant.
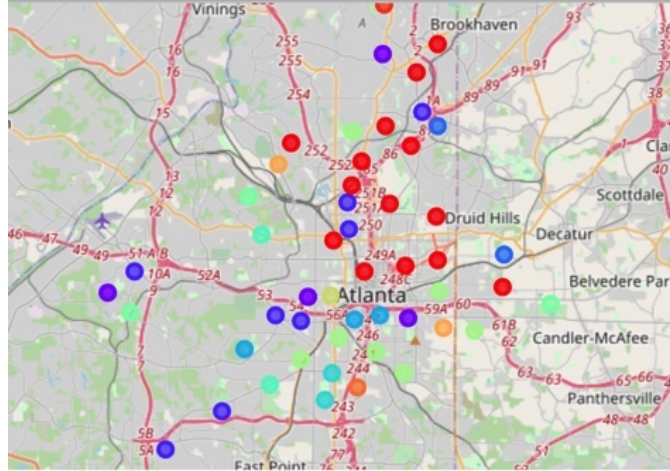
Map of neighborhoods color-coded by cluster

4 Results

Based on analysis of the clusters, I determined that cluster 4 and cluster 8 contained neighborhoods that would reasonably support the type of restaurant I intend to open. In particular, both clusters had cultural sites and a variety of restaurants. These two clusters also have neighborhoods close to each other, so there should be plenty of potential customers near the area in which a restaurant is opened.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
43	Mechanicsville	Restaurant	Japanese Restaurant	Indian Restaurant	History Museum	Greek Restaurant	German Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Design Studio
45	Capitol Gateway, Summerhill	Restaurant	History Museum	Japanese Restaurant	Indian Restaurant	Greek Restaurant	German Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Design Studio
67	Cascade Avenue/Road	Restaurant	Japanese Restaurant	Indian Restaurant	History Museum	Greek Restaurant	German Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Design Studio

Neighborhoods and venues in Cluster 4

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
18	Bush Mountain, Oakland City	Art Gallery	Vegetarian / Vegan Restaurant	Yoga Studio	Dance Studio	Indian Restaurant	History Museum	Greek Restaurant	German Restaurant	Fast Food Restaurant	Falafel Restaurant
42	Adair Park, Pittsburgh	Art Gallery	Yoga Studio	Dance Studio	Indian Restaurant	History Museum	Greek Restaurant	German Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant
44	Peoplestown	Art Gallery	Yoga Studio	Dance Studio	Indian Restaurant	History Museum	Greek Restaurant	German Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant
78	East Atlanta	Art Gallery	Yoga Studio	Dance Studio	Indian Restaurant	History Museum	Greek Restaurant	German Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant
81	Collier Hills, Collier Hills North, Colonial H...	Art Gallery	Art Museum	Yoga Studio	Design Studio	Indian Restaurant	History Museum	Greek Restaurant	German Restaurant	Fast Food Restaurant	Falafel Restaurant
96	Chosewood Park, Englewood Manor	Public Art	Art Gallery	Yoga Studio	Indian Restaurant	History Museum	Greek Restaurant	German Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant
101	Cabbagetown, Reynoldstown	Art Gallery	Restaurant	American Restaurant	Southern / Soul Food Restaurant	Dance Studio	History Museum	Greek Restaurant	German Restaurant	Fast Food Restaurant	Falafel Restaurant

Neighborhoods and venues in Cluster 8

Using the map, I decided that I would choose to build in either the Capitol Gateway or the Peoplestown neighborhood. These are on a main road and near each other. In fact there are multiple neighborhoods in clusters 4 and 8 near these two neighborhoods.
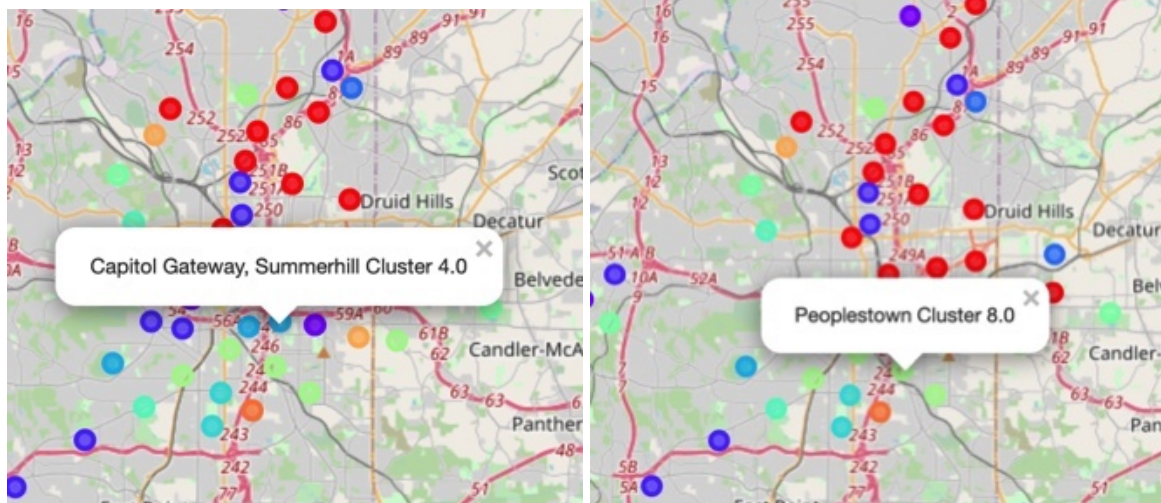
5   Discussion

For this project, I needed to access a data set for neighborhoods in a large city. I chose Atlanta since it is a major city in the southeast United States and because it had readily accessible neighborhood data. For this project, the scope was limited to looking at venues in neighborhoods and inferring neighborhoods that could support a contemporary casual style restaurant. Prior to opening a restaurant, other considerations such as average income of nearby neighborhoods, crime rates, and accessibility. The k-means clustering algorithm allowed me to group neighborhoods by prevalent venues. By mapping the neighborhoods by cluster, I was able to identify geographic regions of the city that had similar ability to support a restaurant.

Both the map and the cluster analysis were important for identifying neighborhoods for this restaurant. Using the cluster analysis, I was able to determine which clusters had top venues with characteristics indicating neighborhoods that could support a new restaurant. Using the map, I was able to identify neighborhoods that were geographically close to other neighborhoods with similar characteristics. Future work for determining location would require more detailed information within neighborhoods.