Applied Data Science Capstone

Determining restaurant location in Atlanta, GA

Elizabeth Niese, June 2020

Overview

- Background

- Data sources

- Methodology

- Results

- Conclusions and future work

Background

Opening a restaurant requires a knowledge of the optimal neighborhood and city surrounding it.

As a major business and travel hub, Atlanta, GA is an excellent location to consider opening a new restaurant.

In this project we use information about neighborhood venues to determine Atlanta neighborhoods that can support a contemporary casual restaurant.

Downtown Atlanta

Source: https://commons.wikimedia.org/wiki/File:Atlanta_Downtown_July_2010.JPG

Data

- Neighborhood data came from https://hub.arcgis.com/datasets/GARC::city-of-atlanta-neighborhood-statistical-areas and was downloadable as a .geojson file

- The data was given based on neighborhood statistical areas and instead of a single latitude/longitude for each neighborhood, boundary points were given.

- To effectively use this information and the Foursquare API to obtain venue information, these boundaries needed to be transformed into single data points. This was done by computing the centroid of each region to serve as the latitude/longitude.

- The Foursquare API was then used to find venues in each neighborhood.

Methodology

- Due to the wide variety of venues in the neighborhoods, it was difficult to identify neighborhoods that could support a restaurant using the full data set.

- To address this, venues were limited to types containing the words *restaurant, studio, museum,* and *art.*

- One hot encoding was used to convert the data into numeric form so that *k*-means clustering could be implemented.

- *K*-means clustering is an algorithm that groups together neighborhoods by common venues.

Methodology

- It is important to decide how many clusters to use when implementing *k*-means clustering. The elbow method was used to do this, implemented with the *yellowbrick* package.

Distortion Score Elbow for KMeans Clustering

Results

- Each cluster was analyzed based on top venues. The two clusters that looked most promising were those indexed by 4 and 8.

- Cluster 4 contains museums and a variety of restaurants.

	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

- Cluster 8 contains art galleries and museums, as well as dance and yoga studios.

	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

Results

- The neighborhoods, color coded by cluster, were plotted on a map.
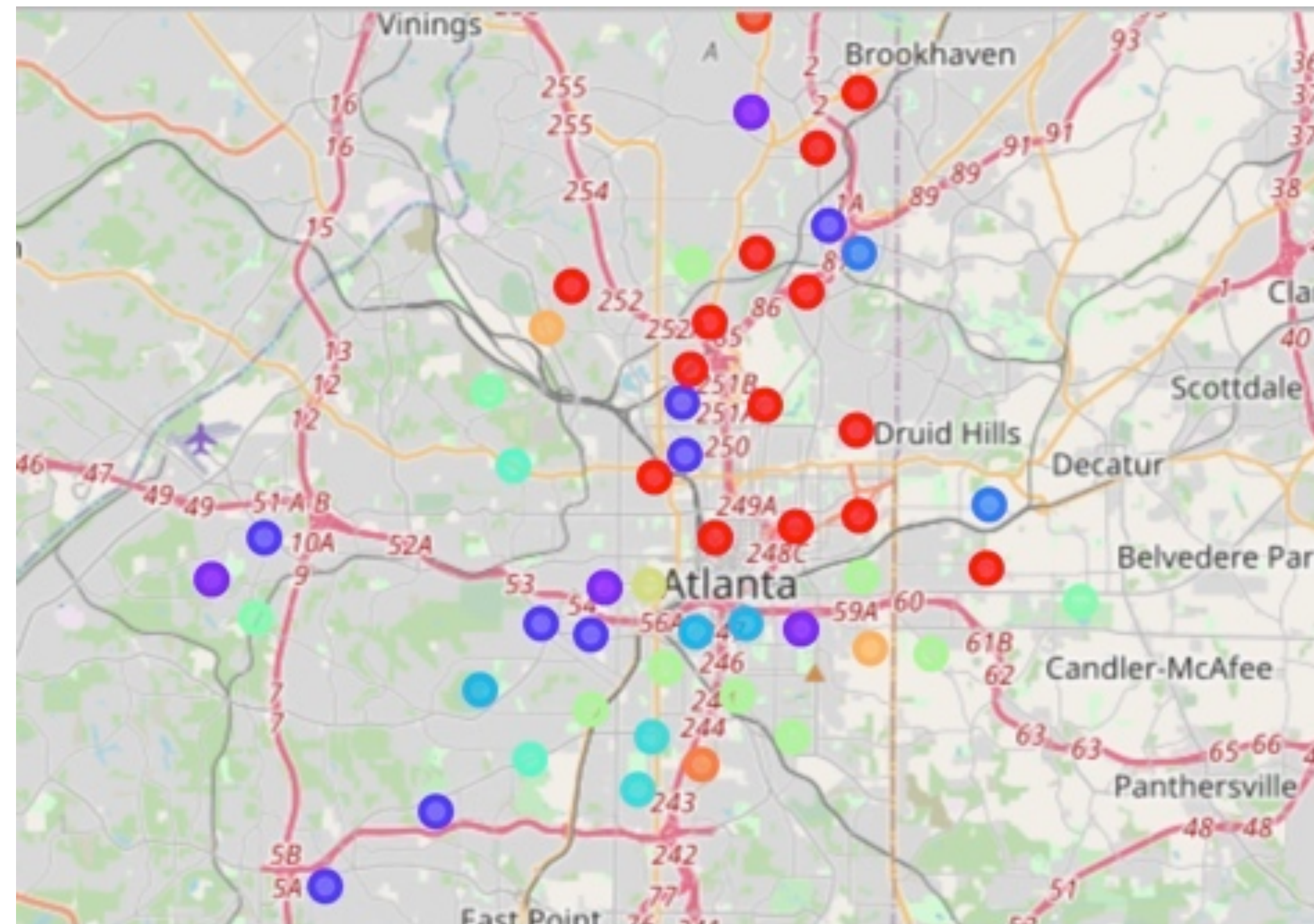
Discussion

Using the map and the cluster analysis, it was determined that the Capitol Gateway and Peoplestown neighborhoods would be good options. The two neighborhoods are close to each other and in proximity to other Cluster 4(teal) and 8 (lime green) neighborhoods.