Predicting the Pattern of College Students in Auburn, AL

Emily Clark

December 4, 2020

1. Introduction
   1. Background

Auburn, AL is a college town with a sizeable non-student population. However, the small businesses in Auburn still rely heavily on student traffic for their income. Any new small business will need to consider its appeal to students in order to be successful. Auburn is known for its small-town feel despite its growing population. That atmosphere is very important to non-student residents, often being an attraction to incoming professors. If Auburn can maintain its appeal, it can continue to recruit top-notch professors and grow in prominence. Therefore, it is important from an economic development perspective to encourage growth among small businesses, and to do that successfully, planners must consider the behavior of the college students.

* 1. Problem

Prospective small business owners need to know which areas will be heavily trafficked by students. This project aims to point to areas within Auburn that are likely to have higher student volume based on the traffic at existing Auburn venues.

* 1. Interest

Pinpointing areas within Auburn that are likely to have heavy student traffic is an important consideration for anyone looking to open a business in the area.

1. Data
   1. Data Sources

Venue information was pulled from Foursquare’s API. I wanted to incorporate ratings into the analysis but was unable to do this at a reasonable scale due to the API’s data restrictions. I pulled the top venues along with their location and category (eg. Pizza restaurant) for clustering.

* 1. Data Cleaning

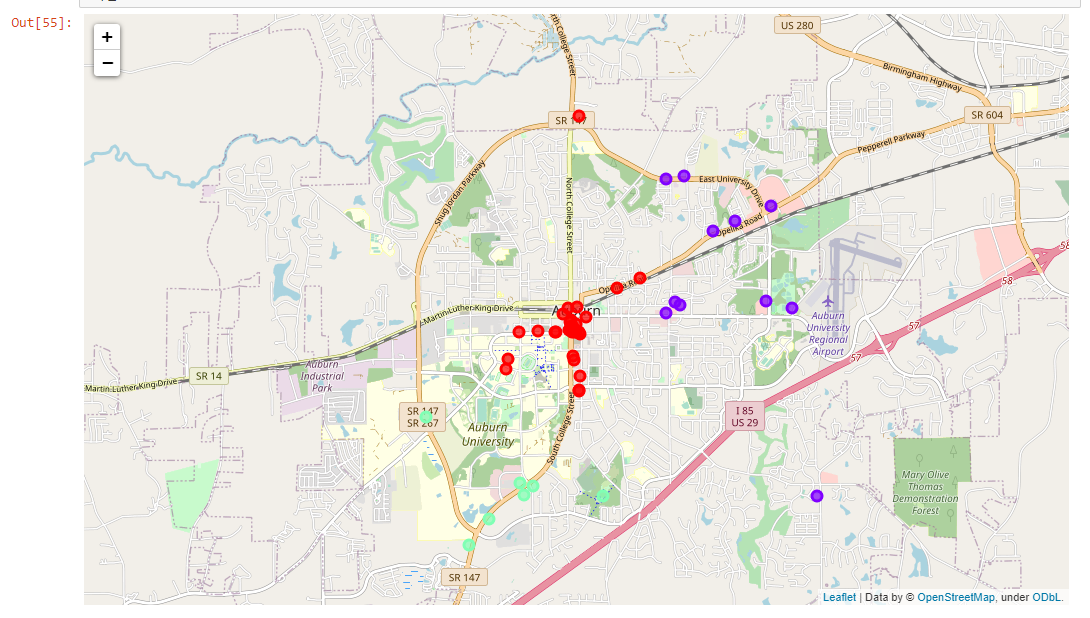
Data imported from the Foursquare API is stored in a JSON format. I used the json\_normalize function to flatten the JSON file so I could pull the relevant fields.

1. Methodology
   1. Data Import

I used the Foursqare API to download the top 50 Auburn, AL venues along with their coordinates and category. I used json\_normalize to flatten the imported data.

* 1. EDA

I did not have neighborhood information for Auburn, so the first step in my analysis was to determine how many distinct areas are present within the city. I used K-Means clustering based on the venue coordinates and visualized the clusters with a Folium map. There were three distinct clusters – one in downtown Auburn (campus), one in south Auburn, and one to the north in the Opelika area. Based on this information, I wanted to see which “neighborhood” – south Auburn or Opelika – is the most similar to the campus area in order to predict where students might spend more of their time.



* 1. Methodology

Once I determined the distinct “neighborhoods” within Auburn, the next step was to determine which area is most similar to the campus area. To do this, I pulled the top 10 most common venue types in each area and assigned dummy variables to them. Then I performed K-means clustering with k=2 to determine which of the two other neighborhoods – south Auburn or Opelika – is the most similar to campus.

1. Results

The k-means clustering algorithm revealed that the Opelika area is the most similar to the campus neighborhood based on the types of venues in each area.

1. Discussion

Auburn, AL has 3 distinct regions – South Auburn, campus, and Opelika. Of those, the campus and Opelika areas are the most similar in terms of venue types. For anyone interested in opening a small business, Opelika is a better location choice than south Auburn if the primary focus is to generate student traffic, since they will likely gravitate towards the area with similar venues as the ones that are already heavily trafficked on campus.

The most significant limitation to this project was the inability to pull venue ratings due to the API’s data constraints. Ideally, I would have liked to group venues by their ratings instead of their coordinates to get a better idea of student preferences.

1. Conclusion

After clustering venues in Auburn, AL to determine which areas students are most likely to visit, I concluded that the Opelika region is the most similar to campus and therefore is the most likely to be visited by the student population. This provides insight to potential small business owners looking for a location, since student traffic is essential for Auburn’s small businesses.