**Using Geospatial APIs to Scout for Optimal Restaurant Locations**

Bryce Goodsite

August 10, 2020

**1. Introduction**

**1.1 Background**

Expanding a family run, brick-and-mortar store can be a major decision. One of the biggest of which is the location of the second venue. Much thought needs to go into selecting a new location, such as proximity to customers, rent, and the space available. In recent years, there have been many trends that have existed that help signify an area is “up-and-coming”.

**1.2 Problem**

I have been approached by the owner of a well-known, up-and-coming restaurant about 30 mins outside of Atlanta, way out in the suburbs. The restaurant is a brewpub that features house beers and high quality food, yet not high enough quality to scare off the average customer from going for a casual dinner/lunch. The restaurant has been tremendously successful locally and the owner wants to open a second location further into Atlanta proper, as he believes that there is a great opportunity for his business to do well in the metro area.

**2. Data acquisition and cleaning**

**2.1 Data sources**

The data for the neighborhoods of Atlanta was pulled from a list of the neighborhoods on a Wikipedia [page](https://en.wikipedia.org/wiki/Table_of_Atlanta_neighborhoods_by_population), while the rest of the information was ascertained using the Foursquare API.

**2.2 Data cleaning**

After importing a series of necessary libraries, the table was scraped from the Wikipedia page and converted to a data frame for use with the python code and the Jupyter Notebook. The table contained 161 neighborhoods and their respective populations, but lacked coordinates that would be usable for Foursquare data. To remedy this, the Nominatim function from the geopy library was used to convert the address name into geographic coordinates. First the neighborhood names were appended with the phrase “, Atlanta, GA” for the function to work correctly and not misidentify wrong coordinates not in the area. Then the neighborhoods that were not able to be located with the function were then cut from the data frame, as they were often much smaller in population. A map of the results was then made to check that the data was formatted correctly.

**3. Exploratory Data Analysis**

**3.1 Trimming the available geospatial data**

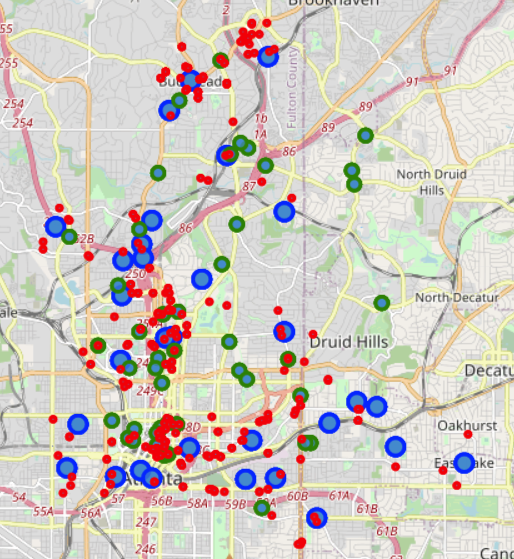
After successfully pulling all of the neighborhoods, the decision was made to focus the search of locations to North of Downtown, and within a certain corridor off of the major North-South highway. This was done as the founder of the restaurant wanted to not drive through all of the Atlanta traffic in order to move between his two locations. Being only a certain distance from the major highways would also help to decrease on travel time.

Another requirement was exposure to local residents, and having a certain population in a neighborhood was added as a requirement. From this, neighborhoods with less than 1200 people were cut from the list, yielding 33 locations. The information was then plotted on a Folium map to check for accuracy and to identify the preliminary candidates.

Now with the neighborhoods narrowed down, other factors need to be considered. The first is something called the “Starbucks Effect”, which describes the phenomena of how a Starbucks store opening increases home and property values. The data shows that between 1997 and 2014, properties closer to the coffee shop increased in value by 96%, compared to 65% for all U.S. residential properties. In short, the locations of Starbucks can be used as a proverbial “canary in the coal mine” for an up-and-coming area. The locations of these were identified and then plotted. A further analysis will be discussed in later sections. For scouting wealthy areas, it’s like heading into an exam with the answer key. Assuming that significant market research has gone into every Starbucks store opening, we can piggyback on those positive conclusions.

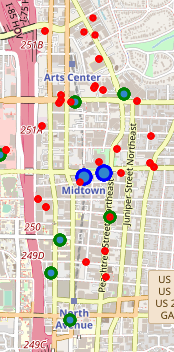
Along with Starbucks, independently owned cafes are also a sign of a neighborhood’s affluence and local economic health. This is due to a rise in city café culture, which is partly attributed, ironically, to the success of Starbucks. The retail chain got customers used to overpriced coffee, but the local café gave a sense of community and “trendiness” to that overpriced coffee. These also tend to have smaller margins than a Starbucks, so typically they concentrate in more affluent areas. These were identified and also mapped onto the combined Folium map, this time in red.

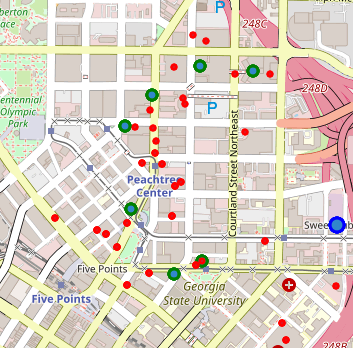
The Folium map now displays the neighborhood locations (blue), the Starbucks locations (green), and the independently owned cafes (red).

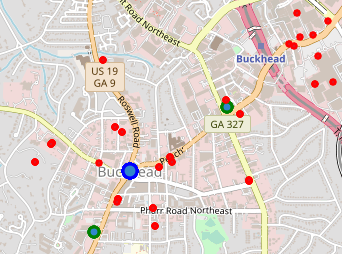


**5. Conclusions**

In this study, I used different qualitative indicators to help identify affluent neighborhoods for a secondary restaurant location. After narrowing down the location based off of geography alone, different factors such as the location of coffee shops provided an idea of what areas potentially supported the most discretionary consumer spending (i.e. certain retail locations were markers of wealth, and likewise pointed toward the local economy being able to support a new restaurant of similar background). After mapping all of the locations onto a Folium map with the aid of the Foursquare API, it was seen that several areas would be good candidates. Based off of these findings, the second restaurant should be located in **Downtown, Midtown,** or **Buckhead**.







**Downtown**

**Buckhead**

**Midtown**

Another good option at first glance appears to be Lenox. However, it can be seen that all of the respective venues are inside of the mall. This will not be counted for now, as the restaurant should not be opened inside of a mall for its second venue ever.

**6. Future directions**

After narrowing the location for the restaurant down to specific neighborhoods, it would be good to identify better economic indicators such as average home value, average household income, cost of leases for restaurants, crime rates, etc. All of these indicators (and more) are regularly used to identify good locations for chains such as Starbucks, but we want to narrow down the search based specifically on just a single restaurant. Furthermore, it would also behoove us to identify similar restaurants by the ratings, and identify the best mix of ratings of restaurants and the cost of doing business in an area. There are certain tools out there for this, however they typically cost money and would require an additional learning curve to utilize effectively.