

Best Locations for a Gourmet Coffee Shop in Apex, NC

Jason Wise

IBM Data Science Professional Certificate Specialization
Capstone Project
January 2021

Location is critical to the success of a Coffee Shop

- The market for coffee drinkers is **STRONG**. Approximately 63% of Americans are coffee drinkers. Of those consumers, around 61% consume “gourmet” coffee¹. Additionally, around 36% of the population consumes coffee brewed outside of the home². In general, these statistics have an upward trend³.
- Building a strong business with repeat customers requires a great location, product and service. Product and Service alone will not sustain a retail business of this nature.
- Location is dependent on several factors – access to key roads, being adjacent to complimentary businesses, demographics and a lower concentration of direct competitors.
- For this project, I am working with a potential Gourmet Coffee Shop entrepreneur who desires to find the best locations to open an establishment in Apex, NC.

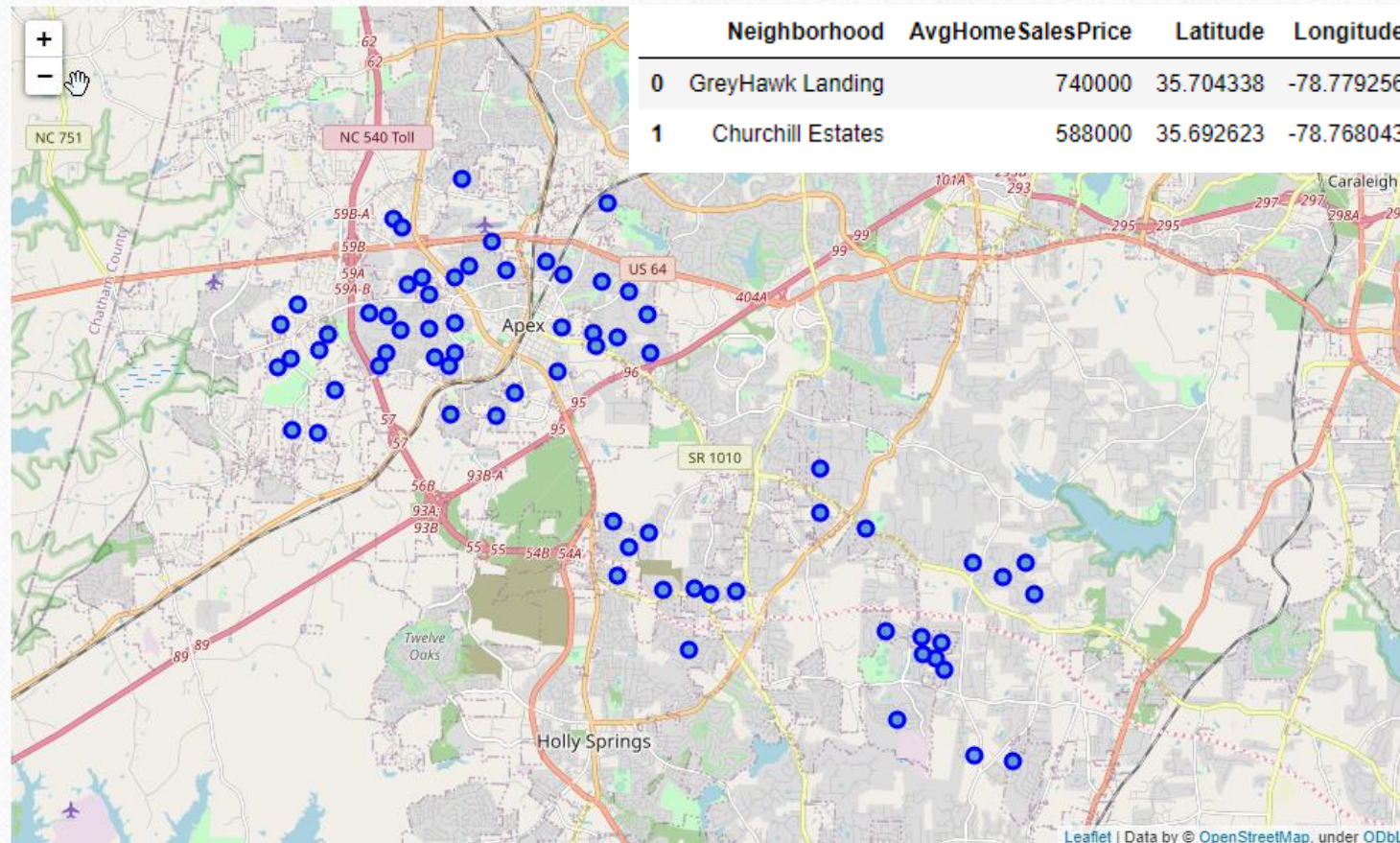
1. <https://www.washingtonpost.com/news/voraciously/wp/2019/03/28/americas-growing-affection-for-gourmet-coffee-and-other-takeaways-from-a-new-national-survey/>
2. <https://dailycoffeenews.com/2018/03/21/current-coffee-consumer-trends-inside-the-ncas-2018-report/>
3. <https://arctoscoffee.com/coffee-trends-show-increase-in-gourmet-and-youth-consumption/>

Data Description

- List of Apex, NC Subdivisions and Related Property Sales Figures – Triangle Realty
 - Neighborhood names and home sales price figures scraped from <http://trianglearearealty.com/apex-nc-subdivisions/>
- GPS Coordinates for Apex, NC Subdivisions - Google Maps Apex, NC
 - GPS data of the type needed, was not readily found and needed to be manually captured from [Google Maps](#)
- Subdivision Demographic Information – US Census via City-Data.com
 - United State Census data as tied to our neighborhoods was manually captured from several pages on city-data.com
- Foursquare Venues and Categories
 - The Foursquare Places API was used to capture venue descriptions and locations around the targeted radius from our neighborhoods foursquare.com

Methodology

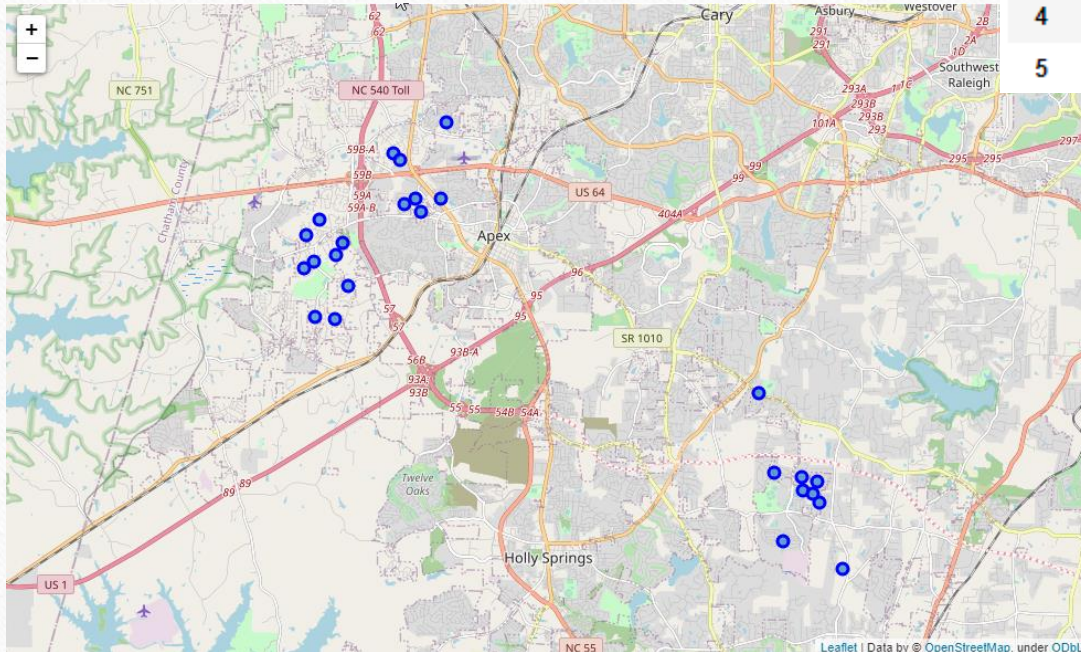
Map Neighborhoods to Visualize with Folium



Limit the Dataset to Base Demographic Market

- Median Resident Age < 40
- Median Income > Mean of our Population
- Visualize with Folium

	Neighborhood	Median Resident Age	Median Household Income	AvgHomeSalesPrice
0	Covington	35.8	118780	378000
1	The Park At West Lake	35.8	118780	350000
2	Sawyers Mill	35.8	118780	228000
3	Villagio	36.1	128636	522000
4	Bella Casa	36.1	128636	490000
5	Holland Farm	36.1	128636	387000



Get Venue Information from Foursquare

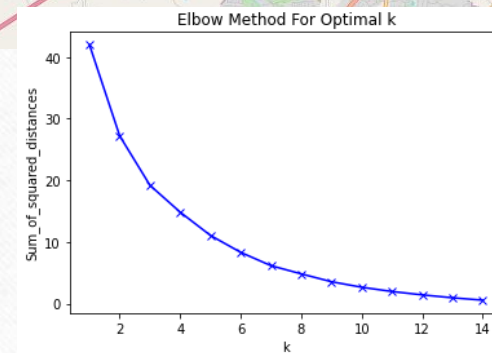
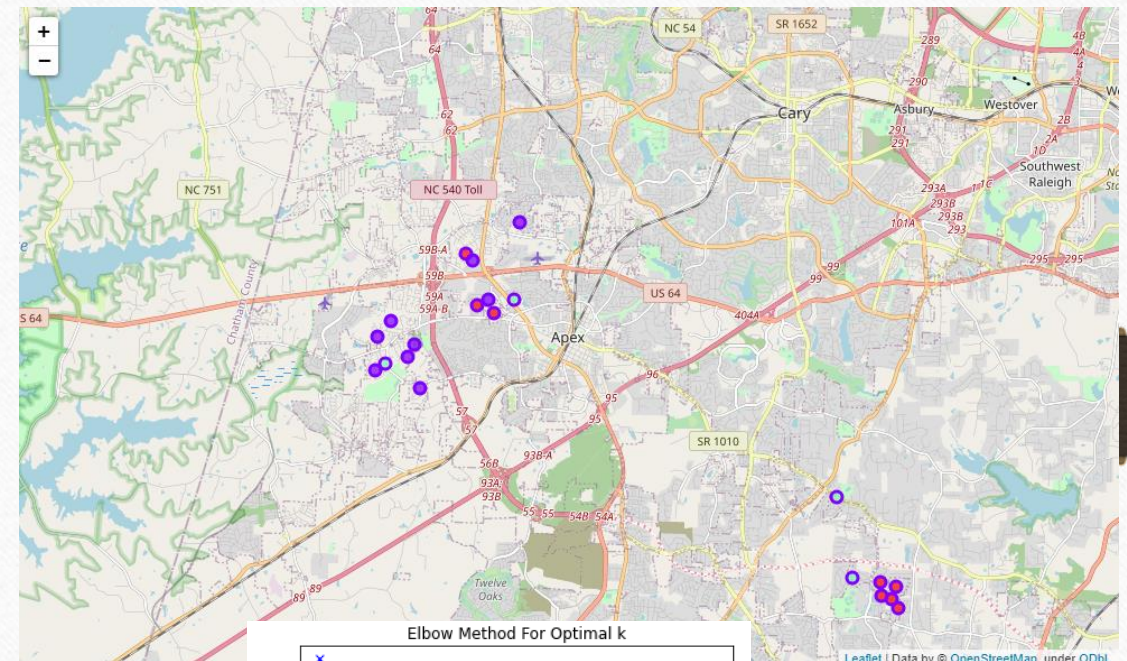
- Limit analysis to venues that are Complimentary OR Direct Competitors within a radius of 1.7 miles from each suburban neighborhood
- Complimentary businesses are those in categories that enhance our location's value when adjacent
- Direct Competitors help us to determine market saturation

	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
0	Abbingdon	Coffee Shop	Furniture / Home Store	Sandwich Place	Salon / Barbershop	Fast Food Restaurant	Theater	Spa	Shipping Store	Gymnastics Gym	Gym
1	Beckett Crossing	Coffee Shop	Sandwich Place	Furniture / Home Store	Salon / Barbershop	Gas Station	Fast Food Restaurant	Theater	Spa	Shipping Store	Gymnastics Gym
2	Belmont	Convenience Store	Theater	Spa	Shipping Store	Sandwich Place	Salon / Barbershop	Gymnastics Gym	Gym / Fitness Center	Gym	Gas Station
3	Carriage Downs	Coffee Shop	Furniture / Home Store	Sandwich Place	Salon / Barbershop	Gas Station	Fast Food Restaurant	Theater	Spa	Shipping Store	Gymnastics Gym
4	Charleston Village	Sandwich Place	Salon / Barbershop	Coffee Shop	Furniture / Home Store	Spa	Gym	Fast Food Restaurant	Shipping Store	Gymnastics Gym	Convenience Store

Cluster Analysis using k-means via Scikit Learn

- Features – Average Venue Category Frequency & Demographics by Neighborhood
- Three Clusters Identified
 - **0** – Oldest Age Group, Lowest Income Group, Mid Home Values
 - **1** – Youngest Age Group, Mid Income Group, Lowest Home Values
 - **2** – Mid Age Group, Highest Income Group, Highest Home Values

	Cluster Labels	Median Resident Age	Median Household Income	AvgHomeSalesPrice
count	4.0	4.000000	4.000000	4.000000
mean	2.0	38.225000	136323.250000	509500.000000
std	0.0	1.968714	7080.491526	62745.517768
min	2.0	36.100000	128636.000000	438000.000000



Filter Cluster Data for Finely Targeted Demographics

- Dropped Cluster 0 - Oldest Age Group, Lowest Income Group, Mid Home Values. While a good market, this least represented our exact targeted demographic out of the three clusters.
- Dropped Neighborhoods with a relatively higher ratio of competitors to complimentary businesses (greater than 10%). This represents our measure of market saturation.
- Dropped the Jamison Park neighborhood due to it's outer suburban location – our client is only interested in more centrally-located neighborhoods.
- These changes resulted in reducing our list to the **Top 5 Neighborhoods**.

Cluster Labels		Neighborhood	Median Resident Age	Median Household Income	AvgHomeSalesPrice	Complimentary	Competitors	Ratio
0	1	Crocketts Ridge	36.1	128636	275000	2	0	0.0
1	1	Greenbrier	36.1	128636	223000	10	1	0.1
2	1	Hollands Crossing	36.1	128636	218000	2	0	0.0
3	1	Woodridge	37.0	145795	261000	6	0	0.0
4	2	Villagio	36.1	128636	522000	4	0	0.0

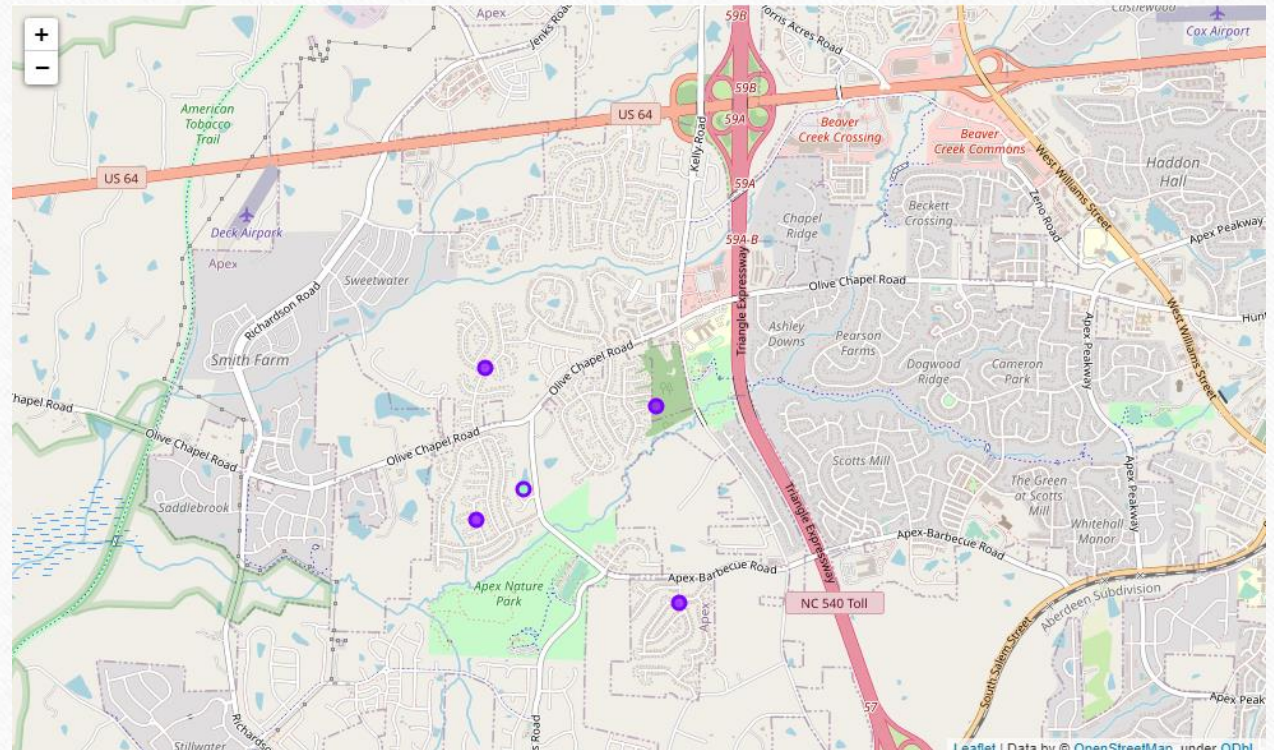
Results

Targeting the Best Locations

- Through my analysis I was able to gather, refine, segment and pinpoint our dataset to be highly targeted with results representative of the client's ideal customer and ideal business location:
 - Demographics - Young - Age Group – Under 40
 - Demographics - Relatively Affluent – Median Income Above the Average of our Targeted Population
 - Location - Complimentary Businesses within the targeted radius of 1.7 miles
 - Location - Low Concentration of Direct Competitors within the targeted radius of 1.7 miles

Top 5 Neighborhoods

- This represents the locations of the Top 5 Neighborhoods and their respective cluster associations.
- In addition to having favorable demographics, being near complimentary businesses, and having low competitor saturation; they are adjacent to major roads and interstates, making for customer ease of access.



Discussion

Observations

- Apex, NC offers a solid demographic based that aligns with the client's target market
- The analysis presented some limitations in the form of a limited data set with a reduced number of features. This meant that our dataset was more closely aligned and resulted in a more “rounded” optimal k analysis.
- Demographic data was roughly aligned to neighborhoods given differing shapes of the geographic borders of census vs subdivision data. This resulted in data results being averaged over adjacent areas.
- With additional resources, this study could be improved through the use of more finely targeted demographic data, as well as qualitative data about competitors and complimentary businesses.

Recommendations

- I recommend the client target a retail location that is located within 1.7 miles of one the Top 5 Neighborhoods
- Additionally, the target location should be adjacent to one or more Complimentary Businesses to promote increased and dedicated customer traffic.
- This analysis can be further improved through a qualitative study of adjacent complimentary businesses and a study of direct competitors in the area. This qualitative analysis will help to refine the business plan and further differentiate the business opportunity.

	Cluster Labels	Neighborhood	Median Resident Age	Median Household Income	AvgHomeSalesPrice	Complimentary	Competitors	Ratio
0	1	Crocketts Ridge	36.1	128636	275000	2	0	0.0
1	1	Greenbrier	36.1	128636	223000	10	1	0.1
2	1	Hollands Crossing	36.1	128636	218000	2	0	0.0
3	1	Woodridge	37.0	145795	261000	6	0	0.0
4	2	Villagio	36.1	128636	522000	4	0	0.0

Conclusion

- The town of Apex, NC represents a thriving and growing population of currently around 60,000 residents. As a suburb of tech savvy Raleigh, it is by nature, well suited for young, educated and growing families.
- The town demographics are well aligned to the client's target market, with many opportunities to start a business in a location with strong potential and an optimistic growth outlook.
- While the study methodology can be improved and expanded upon with a qualitative analysis, the results provide a useful tool for finalizing a business location with a Gourmet Coffee Shop retail location within the town of Apex, NC.