

PREDICTING THE MOST PROFITABLE NEIGHBORHOOD TO OPEN A FAST-FOOD RESTAURANT IN ATLANTA



Business Problem

- ▶ Data Analysis is a valuable tool in providing solutions that make human life easier in many areas.
- ▶ With data analysis you will be able to make decisions on customer trends and behavior prediction, increasing business profitability, and drive effective decision-making.
- ▶ Today, many sectors such as healthcare, manufacturing, finance services actively use data analysis.
- ▶ In this project, I analyzed data that I have and found the most effective solution by using the machine learning clustering algorithm to solve the problem of our customer who wants to open the second branch of his fast-food restaurant in Atlanta.

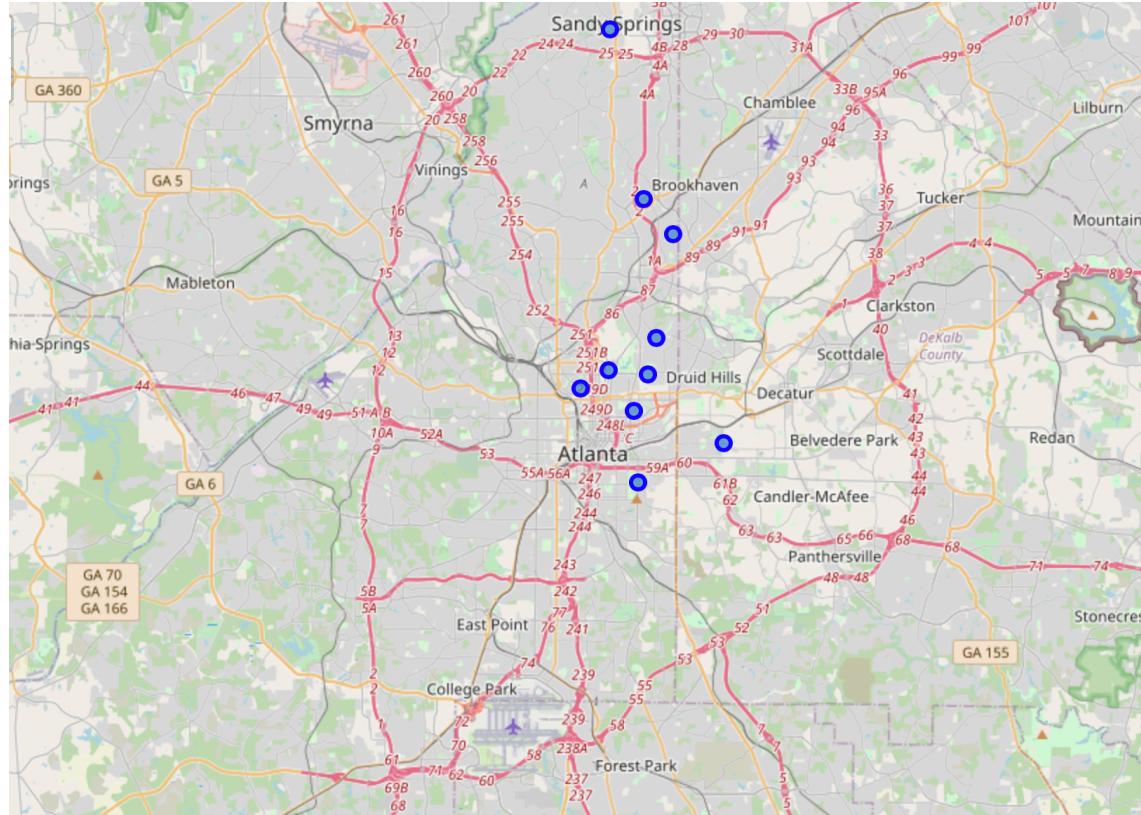
Data acquisition and Cleaning

- ▶ First, I gathered in a table the population data of the neighborhoods in the city of Atlanta that went to Wikipedia records in 2010 by building the code to scrape the Wikipedia page.
- ▶ In total, 161 rows and 3 features in the raw dataset.
- ▶ Eliminated some missing and duplicate values.
- ▶ Sorted the data by population since the first 10 crowded neighborhoods in Atlanta was sufficient for analysis
- ▶ Created the coordinate data of these neighborhoods in csv format.
- ▶ The required data was collected in a table and made ready for visualization and analysis.

Methodology



Creating a map of Atlanta with neighborhoods



- ▶ Created a map that shows the ten most populous neighborhoods of Atlanta using Folium Library to gain a visually general knowledge of the neighborhoods.

Defining Foursquare Credentials and Version

- ▶ Utilized the Foursquare API to explore the most crowded neighborhoods and segment them.
- ▶ Reached the data by setting a 100-venue limit within 500 meters of these neighborhoods.
- ▶ As a result, 72 venues were returned by Foursquare.

	name	categories	lat	lng
0	Exhale	Spa	33.783294	-84.383368
1	Loews Atlanta Hotel	Hotel	33.783366	-84.383188
2	Café Intermezzo	Café	33.783136	-84.383470
3	Street Food Thursdays (& Mondays)	Food Truck	33.784558	-84.382534
4	Einstein's New American Restaurant		33.784143	-84.382086

Exploring Neighborhoods in Atlanta

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Downtown	35	35		35	35	35
Georgia Tech	12	12	12	12	12	12
Grant Park	9	9	9	9	9	9
Kirkwood	15	15	15	15	15	15
Midtown	72	72	72	72	72	72
Morningside/Lenox Park	3	3	3	3	3	3
North Buckhead	56	56	56	56	56	56
Old Fourth Ward	2	2	2	2	2	2
Pine Hills	3	3	3	3	3	3
Virginia-Highland	33	33	33	33	33	33

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0 Midtown	33.78302	-84.382332	Exhale	33.783294	-84.383368	Spa
1 Midtown	33.78302	-84.382332	Loews Atlanta Hotel	33.783366	-84.383188	Hotel
2 Midtown	33.78302	-84.382332	Café Intermezzo	33.783136	-84.383470	Café
3 Midtown	33.78302	-84.382332	Street Food Thursdays (& Mondays)	33.784558	-84.382534	Food Truck
4 Midtown	33.78302	-84.382332	Einstein's	33.784143	-84.382086	New American Restaurant

- ▶ Created a function to repeat the same process to all the neighborhoods have high population in Atlanta.
- ▶ 240 venues in different sectors in these ten crowded neighborhoods of Atlanta were returned by Foursquare.
- ▶ Obtained venues data in several different categories for each neighborhood.
- ▶ Totally, 106 unique categories were founded from all the returned venues.

Analyzing Each Neighborhood

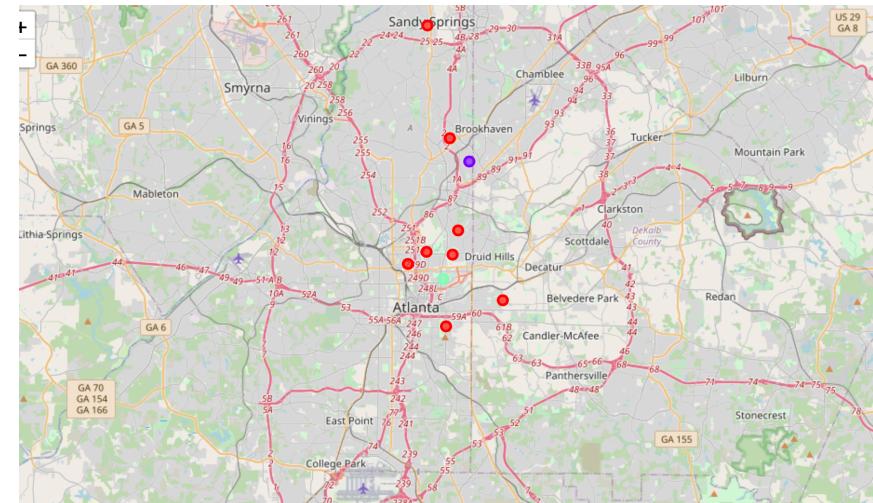
- ▶ In this part of the project, each neighborhood was analyzed separately by writing some functions to get more accurate results.
- ▶ Created a table which shows list of the top 10 venue categories for each neighborhood.

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	Downtown	Spa	Restaurant	Bakery	Vegetarian / Vegan Restaurant	Pizza Place	Bar	Middle Eastern Restaurant	Accessories Store	Park	Chinese Restaurant
1	Georgia Tech	Fast Food Restaurant	Sandwich Place	Chinese Restaurant	Food Court	College Theater	Coffee Shop	Music Venue	Restaurant	Bank	Athletics & Sports
2	Grant Park	Zoo Exhibit	Music Venue	Playground	Pharmacy	Park	Pool	Wine Bar	Video Store	Historic Site	Fast Food Restaurant
3	Kirkwood	Pet Store	Pizza Place	Bar	Coffee Shop	Mexican Restaurant	Breakfast Spot	Sandwich Place	Historic Site	Sports Bar	Juice Bar
4	Midtown	American Restaurant	Hotel	Seafood Restaurant	Spa	New American Restaurant	Italian Restaurant	Coffee Shop	Southern / Soul Food Restaurant	Gay Bar	Indian Restaurant

Clustering Neighborhoods

- ▶ There are the same kinds of venues in every neighborhood, it was easier to gather them in a group according to their similarities.
- ▶ Used the ‘k-Means’ which is the most popular unsupervised machine learning algorithm to cluster the neighborhoods.
- ▶ Created a new data frame that includes the cluster as well as the top 10 venues for each neighborhood.
- ▶ Finally, created the map below by visualizing the resulting clusters. Each color represents different cluster.

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
American Restaurant	Hotel	Seafood Restaurant	Spa	New American Restaurant	Italian Restaurant	Coffee Shop	Southern / Soul Food Restaurant
Spa	Restaurant	Bakery	Vegetarian / Vegan Restaurant	Pizza Place	Bar	Middle Eastern Restaurant	Accessories Store
Italian Restaurant	Playground	Furniture / Home Store	Dive Bar	Doctor's Office	Electronics Store	Exhibit	Farmers Market
Women's Store	Steakhouse	Boutique	Hotel	Italian Restaurant	Coffee Shop	Furniture / Home Store	Kids Store
Pool	Scenic Lookout	Furniture / Home Store	Dive Bar	Doctor's Office	Electronics Store	Exhibit	Farmers Market



Results

Cluster 1

	Population (2010)	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	16569	American Restaurant	Hotel	Seafood Restaurant	Spa	New American Restaurant	Italian Restaurant	Coffee Shop	Southern / Soul Food Restaurant	Gay Bar	Indian Restaurant
1	13411	Spa	Restaurant	Bakery	Vegetarian / Vegan Restaurant	Pizza Place	Bar	Middle Eastern Restaurant	Accessories Store	Park	Chinese Restaurant
3	8270	Women's Store	Steakhouse	Boutique	Hotel	Italian Restaurant	Coffee Shop	Furniture / Home Store	Kids Store	Accessories Store	Toy / Game Store
5	8030	Playground	Trail	Park	Zoo Exhibit	Furniture / Home Store	Doctor's Office	Electronics Store	Exhibit	Farmers Market	Fast Food Restaurant

- ▶ Examined each cluster and determined the discriminating venue categories that distinguish each cluster.
- ▶ Assigned a name to each cluster.
- ▶ The number of restaurants belonging to different cultures such as Italian and Chinese in the downtown of Atlanta is quite high.
- ▶ Although fast food restaurants are in the top ten on the most common venues list, they are not among the top three most common places.

Discussion and Conclusion

- ▶ Atlanta is a city developed in terms of art, culture and food and beverage tourism. Opening a restaurant in such a city can be advantageous as well as risky.
- ▶ In this project, I analyzed the relationship between the population and restaurant selection in the most crowded neighborhoods of Atlanta.
- ▶ I did visualization using the folium library to support my predictions.
- ▶ If I must interpret the concrete data, I worked on in my project, fast food restaurants in the downtown of Atlanta are less than other restaurants.
- ▶ For this reason, opening a new fast-food restaurant in the neighborhoods close to downtown may be logical in terms of investment, but it should be kept in mind that the rent of the venue here is high.
- ▶ Although opening a new place or restaurant in a big city such as Atlanta depends on many sociological and economic factors, I believe that I have reached the most accurate result with the available data.
- ▶ By using the analysis methods, I have used, this project can be studied in more detail and extensively if more features are achieved regarding these neighborhoods.