IBM Data Science Certification Capstone Project:
"Segmenting and Clustering: Miami Neighborhoods."

Liliya Strong

July 10th, 2021¶

1. Introduction

1.1 Background

Miami is a beautiful coastal city in the USA. Its growing metropolitan area is blessed with a favorable climate and a low tax environment. Inspired by the famous real estate axiom "Location, location, and location," I will try to find optimal neighborhoods to live in in the Miami area.

1.2 Problem

The final decision on optimal property location is a complex process that entails many choices. I'm interested to see if walking proximity to schools, beaches, golf courses, many restaurants, or places you could go to exercise, do shopping, or do cultural activities influence property prices. I'll use data science tools of visualization, correlation and k-means clustering algorithm to group the neighborhoods into clusters.

1.3 Interest

The conclusion of differences/similarities between Miami neighborhoods will be useful for not only prospective buyers out of town but also local real estate brokers.

2. Data acquisition and cleaning

2.1 Data sources

I based my analysis on the following databases, APIs, libraries:

- Miami's neighborhood list and its typical home prices could be found in Zillow Database. Zillow website provides the Zillow Home Value Index (ZHVI): A smoothed, seasonally adjusted measure of the typical home value and market changes across a given region and housing type. It reflects the typical value for

homes in the 35th to 65th percentile range. I used the latest available dataset from the end of May, 2021. Source: https://www.zillow.com/research/data/
- Google API will help to get the latitude and longitude coordinates for Miami neighborhoods.
- Foursquare API will get me venue names, location and venue categories
- List of all schools and their GPS coordinates (latitude and longitude) could be found in Miami-Dade County's open data hub. Source: https://gis-mdc.opendata.arcgis.com/datasets/private-school?geometry=-81.005%2C25.493%2C-79.921%2C25.926
- Finally, for crime rate I'll use data of Area Vibes website, which compiles data from FBI and local law enforcement agency and when not available, also includes estimates based on demographic data. https://www.areavibes.com/miami-fl/most-dangerous-neighborhoods/. Here, the crime rate defined as the number of violent crimes (murder, rape, robbery, or assault) per 100,000 people living in each neighborhood.

2.2 Data cleaning

- Data downloaded or scraped from multiple sources were combined into one table. Zillow dataset had information for all of the US, but I only focused on Miami area (Miami and Miami Beach) and the latest property prices from May of 2021.
- Only large schools with over 100 enrollment rates were selected. Calculate distances between each school and neighborhood and only count those under 1 mile.
- Only venue names, their GPS coordinates and categories were selected from Foursquare API.

2.3 Feature selection

I created a detailed table describing location factors for each neighborhood in Miami and Miami Beach:

35 neighborhoods "Neighborhood Names" by 12 location factors:

- "City" -from Zillow
- "Typical Property Value, $" -from Zillow
- Restaurant Freq" – from Foursquare API select only columns from Venue Names with food related names (Restaurant | Bar | Café | Coffe) and add them up
- "Store Freq"- from Foursquare API select only columns from Venue Names with shopping related names (Store | Market | Shop), then add them up.
- "Sport Freq"- from Foursquare API select only columns from Venue Names with exercise related names (Gym | Yoga | Tennis | Stadium | Studio), then add them up.

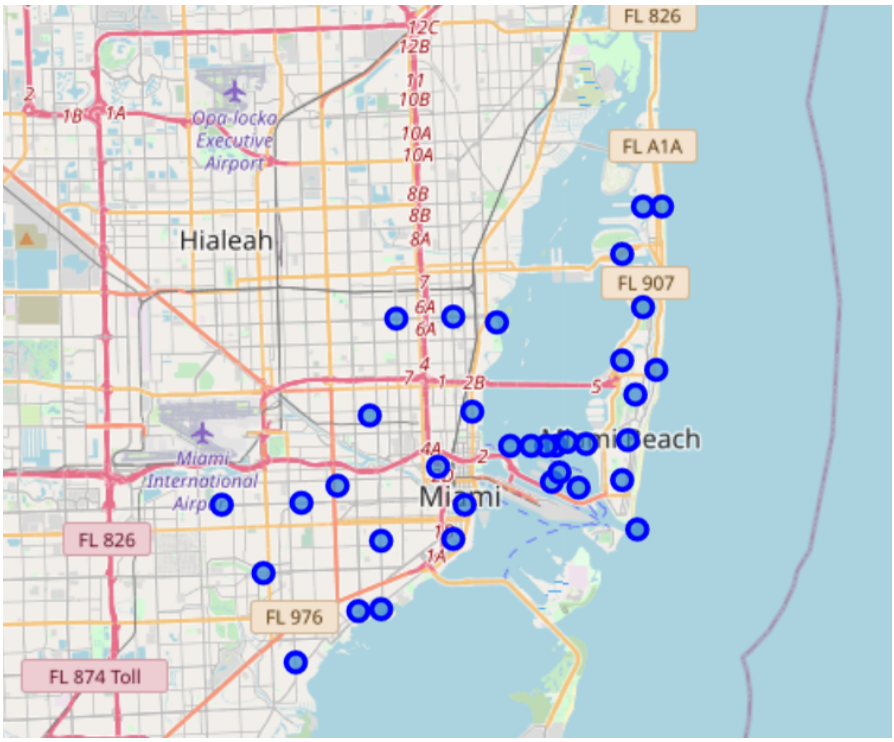
- "Art Freq" – from Foursquare API select only columns from Venue Names with cultural related names (Art | Museum | Event | Concert | Music | Theater), then add them up.
- "Golf | Park"- from Foursquare API select only columns from Venue Names with Golf or park and add their frequencies up.
- "Waterfront" – create Boolean of 1 or 0 if either Foursquare API has name related to "Beach" or name of the neighborhood has words "Island | Isle | Pointe | Point"
- "School Count" –I calculate the distance from each school to each neighborhood using Geopy library. Then add up only those which are less than 1 mile away
- "Crime per 100k" – use Area Vibes website dataset for each neighborhood.
- "Neighborhood Latitude" – Google API
- "Neighborhood Longitude" "– Google API

3. Exploratory Data Analysis

My analysis shows 35 neighborhoods in Miami and Miami Beach area with a massive difference in typical property prices. To understand the reason, I put together detailed descriptions of each neighborhood: 1) Map of Miami, 2) Statistical distribution of property prices in Miami 3) Top 10 most frequent venues, 4) Density of the essential categories such as restaurants, schools, gyms, cultural activities, shopping, and availability of golf course or beach in the vicinity.

3.1 Visualization of Miami Map with all neighborhoods

There are 35 neighborhoods between Miami (aka Downtown) and Miami Beach (aka South Beach). The most expensive neighborhoods in the Miami area have a central location -- conveniently located btw touristy South Beach and arty/foodie Downtown Miami. The least expensive are more inland and are on Miami side.



3.2. Statistical Distribution of Typical Property Prices in Miami

From Zillow data, I discovered that the average property price is around $ 2 millions and that different neighborhoods have a vast price difference. For example, a typical home in Star Island ($28 mln) and Little Havanna ($200k) have almost 150 times price delta!

	Typical Property Value, $
count	35.000
mean	2,094,096.086
std	4,863,894.150
min	210,996.000
25%	346,186.000
50%	548,980.000
75%	995,299.500
max	28,298,329.000

3.3 Top 10 Most Frequent Venues for each Neighborhood

Using Foursquare API, I built a table with 10 most common venues. Looking at the table below, you can get a feel for the neighborhood. For example, we can see that the expensive neighborhoods have beach, spas, places to exercise, and nearby hotels.

	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	Alameda - West Flagler	Latin American Restaurant	Cuban Restaurant	Fast Food Restaurant	Pharmacy	Bank	Bakery	Mexican Restaurant	South American Restaurant	Italian Restaurant	Grocery Store
1	Allapattah	Latin American Restaurant	Cuban Restaurant	Clothing Store	Nightclub	Food & Drink Shop	Spanish Restaurant	Gas Station	Light Rail Station	Café	Snack Place
2	Bayshore	Hotel	Beach	Spa	Nightclub	Italian Restaurant	Cocktail Bar	Resort	Hostel	Hotel Pool	Coffee Shop
3	Belle Isle	Gym / Fitness Center	American Restaurant	Grocery Store	Clothing Store	Pizza Place	Bakery	Vegetarian / Vegan Restaurant	Park	Italian Restaurant	Coffee Shop
4	Biscayne Island	Boat or Ferry	Cruise Ship	Art Gallery	Seafood Restaurant	Park	Café	Harbor / Marina	Bus Station	Performing Arts Venue	Exhibit
5	Biscayne Point	Beach	Hotel	Italian Restaurant	Sandwich Place	Cuban Restaurant	Park	Bakery	Mediterranean Restaurant	Peruvian Restaurant	Breakfast Spot
6	Brickell	Hotel	Italian Restaurant	Seafood Restaurant	Argentinian Restaurant	Bar	Latin American Restaurant	Bakery	Pharmacy	Steakhouse	Japanese Restaurant
7	City Center	Hotel	Beach	Clothing Store	Bar	Italian Restaurant	Bakery	Vegetarian / Vegan Restaurant	Lounge	Pizza Place	Coffee Shop
8	Coral Way	Café	American Restaurant	Italian Restaurant	Japanese Restaurant	Restaurant	Bakery	Steakhouse	Ice Cream Shop	Argentinian Restaurant	Hotel
9	Di Lido Island	Bar	Park	Gym / Fitness Center	American Restaurant	Coffee Shop	Bakery	Italian Restaurant	Hotel	Gym Pool	Yoga Studio
10	Downtown	Hotel	Italian Restaurant	Seafood Restaurant	Peruvian Restaurant	Restaurant	Residential Building (Apartment /	Bakery	Bar	Grocery Store	Pizza Place

11	Fair Isle	Caribbean Restaurant	Grocery Store	Mediterranean Restaurant	Candy Store	Fast Food Restaurant	Tapas Restaurant	Bistro	Boat or Ferry	Coffee Shop	Harbor / Marina
12	Flagami	Cuban Restaurant	Fast Food Restaurant	Latin American Restaurant	Pharmacy	Hotel	Sandwich Place	Grocery Store	Spanish Restaurant	Coffee Shop	Mobile Phone Shop
13	Flamingo Lummus	Hotel	Beach	Pizza Place	Italian Restaurant	Mediterranean Restaurant	Art Gallery	Seafood Restaurant	Juice Bar	Yoga Studio	Park
14	Hibiscus Island	Boat or Ferry	Cruise Ship	Harbor / Marina	Pool	Park	American Restaurant	Gym	Pier	Hotel	Island
15	Isle of Normandy	Pizza Place	Hotel	Italian Restaurant	Coffee Shop	Resort	Beach	Grocery Store	Park	Café	Bistro
16	La Gorce	Hotel	Beach	Food Truck	Resort	Bar	American Restaurant	Playground	Shipping Store	Boat or Ferry	Sports Club
17	Liberty City	Fried Chicken Joint	Fast Food Restaurant	Discount Store	Cosmetics Shop	Park	Sandwich Place	Food & Drink Shop	Soccer Stadium	Storage Facility	Gas Station
18	Little Haiti	Italian Restaurant	Caribbean Restaurant	Art Gallery	Fast Food Restaurant	Pizza Place	Gas Station	Discount Store	Sandwich Place	Gym	Argentinian Restaurant
19	Little Havana	Latin American Restaurant	Bakery	Pharmacy	Pizza Place	Grocery Store	Asian Restaurant	Kids Store	Seafood Restaurant	Sandwich Place	Theater
20	Nautilus	Beach	Hotel	Resort	Italian Restaurant	Hotel Bar	Spa	Food Truck	Dim Sum Restaurant	Bistro	Mediterranean Restaurant
21	North Shore	Beach	Hotel	Italian Restaurant	Breakfast Spot	Sandwich Place	Park	Bakery	Brazilian Restaurant	Coffee Shop	Restaurant
22	North-East Coconut Grove	Pizza Place	Park	Bar	New American Restaurant	Fast Food Restaurant	Gym	Hotel	Sandwich Place	Liquor Store	Lingerie Store
23	Oceanfront	Hotel	Beach	Bar	Spa	Resort	Hotel Bar	Italian Restaurant	Cocktail Bar	Pool	American Restaurant
24	Overtown	Seafood Restaurant	American Restaurant	Nightclub	Restaurant	Gym	Theater	Café	Art Gallery	Italian Restaurant	Coffee Shop
25	Palm Island	Boat or Ferry	Harbor / Marina	Cruise Ship	Park	Exhibit	Zoo	Lawyer	Pool	Cuban Restaurant	Restaurant
26	Rivo Alto Island	American Restaurant	Clothing Store	Gym / Fitness Center	Italian Restaurant	Bakery	Grocery Store	Bar	Coffee Shop	Peruvian Restaurant	Vegetarian / Vegan Restaurant
27	San Marco Island	Boat or Ferry	Cruise Ship	Island	Harbor / Marina	Exhibit	Restaurant	Zoo	Resort	Bus Station	Seafood Restaurant
28	San Marino Island	Exhibit	Museum	Park	Zoo	Restaurant	Café	Mediterranean Restaurant	Gym	Farm	Travel Agency
29	Shenandoah	Cuban Restaurant	Smoke Shop	Latin American Restaurant	Mexican Restaurant	Pizza Place	Spanish Restaurant	Park	Bakery	Seafood Restaurant	South American Restaurant
30	South Pointe	Hotel	Seafood Restaurant	Italian Restaurant	Beach	Pizza Place	Restaurant	Cocktail Bar	Gym	Pharmacy	Sandwich Place
31	South-West Coconut Grove	Women's Store	Clothing Store	Lingerie Store	New American Restaurant	Italian Restaurant	Gym / Fitness Center	Coffee Shop	Park	Thai Restaurant	French Restaurant
32	Star Island	Boat or Ferry	Cruise Ship	Gym	Park	Italian Restaurant	Bakery	American Restaurant	Pool	Island	Harbor / Marina
33	Upper Eastside	Italian Restaurant	Park	Gym	Restaurant	Pizza Place	American Restaurant	Shopping Mall	Gas Station	Sandwich Place	Sushi Restaurant
34	Wynwood - Edgewater	Art Gallery	Ice Cream Shop	Coffee Shop	Restaurant	Pizza Place	Asian Restaurant	Mexican Restaurant	Gym / Fitness Center	Bar	Brewery

3.4 Density of the Essential Location Categories

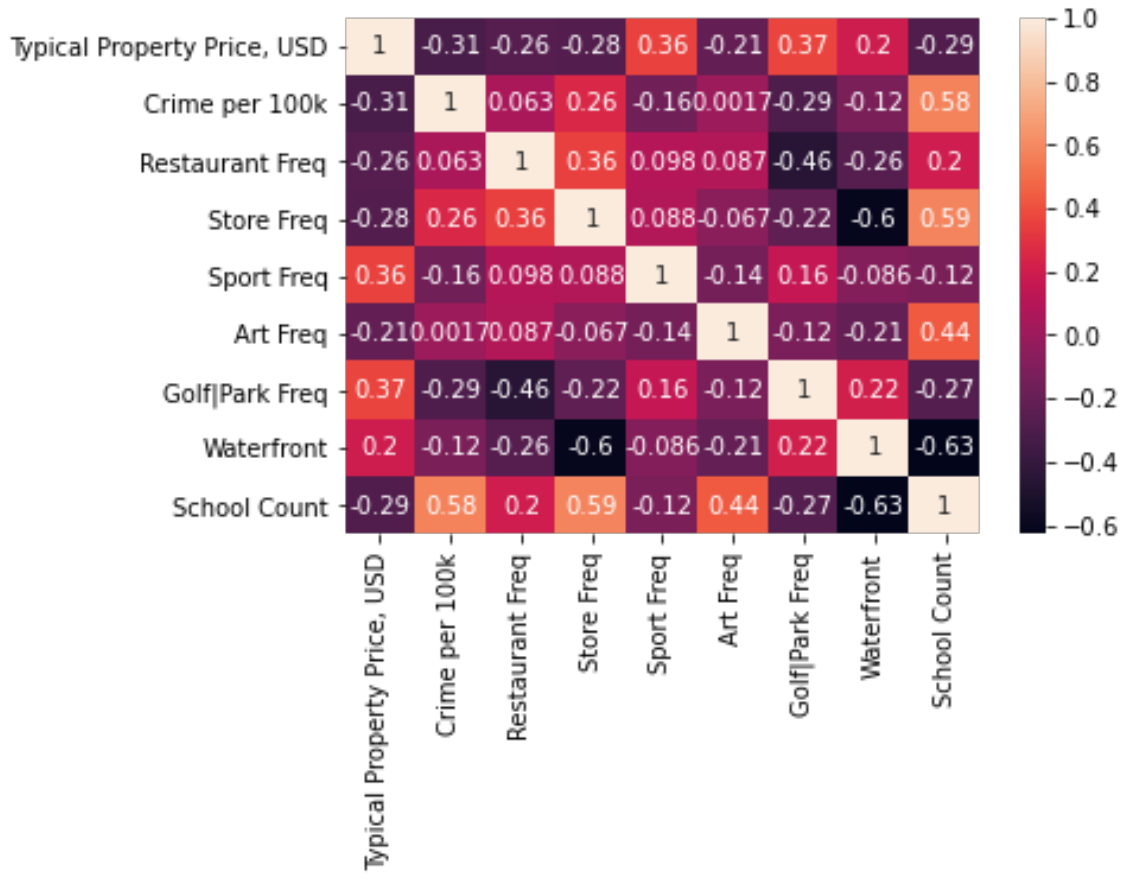
I calculated the density of the essential categories such as restaurants, schools, gyms, cultural activities, shopping, and availability of golf course or beach in the vicinity.

Neighborhood	City	Typical Property Price, USD	Restaurant Freq	Store Freq	Sport Freq	Art Freq	Golf|Park Freq	Waterfront	School Count	Crime per 100k	Latitude	Longitude
Alameda - West Flagler	Miami	397,024.000	0.441	0.215	0.000	0.032	0.032	0	4	269	25.772	-80.251
Allapattah	Miami	288,712.000	0.250	0.304	0.018	0.036	0.018	0	4	944	25.800	-80.227
Bayshore	Miami Beach	560,268.000	0.300	0.060	0.010	0.040	0.020	1	3	852	25.808	-80.130
Belle Isle	Miami Beach	685,340.000	0.360	0.220	0.100	0.040	0.030	1	0	237	25.791	-80.148
Biscayne Island	Miami	971,736.000	0.208	0.052	0.042	0.146	0.042	1	2	237	25.790	-80.175
Biscayne Point	Miami Beach	977,991.000	0.361	0.103	0.010	0.021	0.041	2	1	852	25.868	-80.128
Brickell	Miami	392,743.000	0.480	0.140	0.060	0.000	0.000	0	2	1195	25.760	-80.196
City Center	Miami Beach	327,291.000	0.350	0.170	0.000	0.050	0.010	1	2	852	25.793	-80.133
Coral Way	Miami	482,235.000	0.480	0.180	0.040	0.020	0.010	0	2	187	25.749	-80.265
Di Lido Island	Miami Beach	4,730,846.000	0.409	0.091	0.136	0.000	0.068	1	0	237	25.791	-80.159
Downtown	Miami	351,345.000	0.480	0.120	0.030	0.030	0.020	0	5	1195	25.771	-80.192
Fair Isle	Miami	761,856.000	0.484	0.161	0.000	0.032	0.032	1	1	187	25.737	-80.222
Flagami	Miami	341,027.000	0.378	0.257	0.014	0.000	0.027	0	4	227	25.772	-80.280
Flamingo Lummus	Miami Beach	257,367.000	0.320	0.110	0.050	0.040	0.020	1	2	852	25.779	-80.135
Hibiscus Island	Miami Beach	3,798,735.000	0.110	0.012	0.098	0.000	0.049	1	0	237	25.782	-80.158
Isle of Normandy	Miami Beach	289,870.000	0.400	0.120	0.010	0.020	0.040	2	2	852	25.853	-80.135
La Gorce	Miami Beach	2,028,854.000	0.203	0.068	0.017	0.000	0.034	1	0	237	25.836	-80.128
Liberty City	Miami	282,969.000	0.194	0.387	0.032	0.000	0.065	0	10	1919	25.832	-80.217
Little Haiti	Miami	370,677.000	0.400	0.200	0.074	0.053	0.011	0	6	1138	25.833	-80.196
Little Havana	Miami	210,996.000	0.303	0.224	0.000	0.092	0.039	0	5	585	25.778	-80.238
Nautilus	Miami Beach	1,012,608.000	0.263	0.070	0.018	0.035	0.000	1	2	852	25.818	-80.135
North Shore	Miami Beach	288,333.000	0.333	0.065	0.011	0.022	0.032	1	1	852	25.869	-80.121
North-East Coconut Grove	Miami	714,243.000	0.380	0.230	0.080	0.000	0.040	0	2	233	25.737	-80.230
Oceanfront	Miami Beach	449,860.000	0.290	0.050	0.010	0.040	0.010	1	2	852	25.815	-80.122
Overtown	Miami	212,994.000	0.400	0.110	0.060	0.170	0.010	0	10	851	25.784	-80.201
Palm Island	Miami Beach	5,274,093.000	0.083	0.000	0.000	0.000	0.083	1	0	237	25.779	-80.160
Rivo Alto Island	Miami Beach	4,977,389.000	0.420	0.200	0.100	0.010	0.030	1	0	237	25.792	-80.155
San Marco Island	Miami	4,928,254.000	0.167	0.000	0.000	0.033	0.033	1	0	237	25.790	-80.168
San Marino Island	Miami Beach	5,536,186.000	0.227	0.000	0.045	0.091	0.091	1	0	237	25.791	-80.163
Shenandoah	Miami	548,980.000	0.520	0.190	0.000	0.100	0.030	0	4	187	25.760	-80.222
South Pointe	Miami Beach	744,609.000	0.432	0.084	0.042	0.011	0.021	2	1	852	25.764	-80.130
South-West Coconut Grove	Miami	891,841.000	0.350	0.360	0.080	0.010	0.020	0	4	368	25.720	-80.253
Star Island	Miami Beach	28,298,329.000	0.273	0.078	0.104	0.000	0.052	1	1	237	25.777	-80.151
Upper Eastside	Miami	506,184.000	0.422	0.157	0.084	0.048	0.048	0	2	751	25.830	-80.180
Wynwood - Edgewater	Miami	401,578.000	0.410	0.260	0.050	0.100	0.020	0	8	782	25.802	-80.189

4. Analyzing and Clustering of Neighborhoods

4.1 Visualize Correlation btw Different Neighborhood Factors

To see if there is any relationship between 12 location factors, I built a correlation matrix. We can see that Waterfront, Golf course or Park and a lot of places to go for exercise ("Sport Freq") in the vicinity have a positive correlation with a property price. On the other hand, I had a surprising finding that the frequency of restaurants/stores/art/schools negatively correlates with property prices in Miami. But I could see how it makes sense. First, we don't know the ranking of those venues. Just to have density doesn't imply quality. Second, Miami's nature is tourism. Given that the tourist influx brings a lot of foot traffic to specific neighborhoods, this creates a high probability of crime. Indeed, crime rates negatively correlate with property prices and have a positive correlation with restaurants/stores/art factors. The number of schools also doesn't affect high-end real estate market. This is probably because of the demographics of high-net-worth individuals. They are usually not full-time residents of Miami.

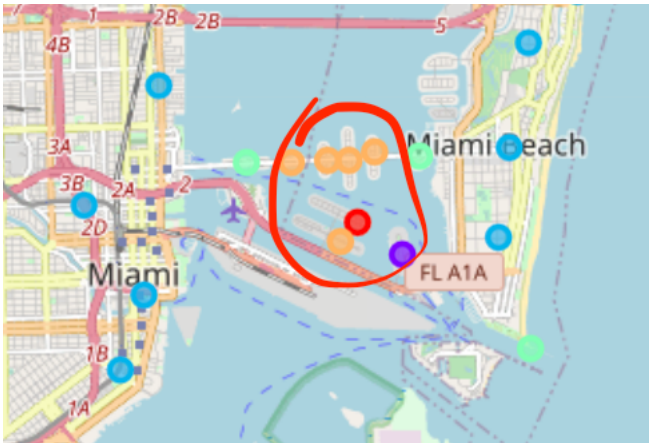


4.2 K-means Clustering Algorithm

Using 10 location factors (excluding latitude and longitude), I used k-mean clustering algorithm to discover neighborhood clusters of Miami. This clustering revealed the following: the most expensive neighborhoods in the Miami area have a central location -- conveniently located btw touristy South Beach and arty/foodie Downtown Miami. They all have a waterfront (which I define as either to have water view or beach) and golf course or park in the vicinity. They also have a lot of places to go to exercise, in that sense, the famous real estate axiom "Location, location, and location" proved to be true. Let me describe each neighborhood cluster in more detail:

"Upscale Islands":

Star, Hibiscus, Palm and other Venetian Islands



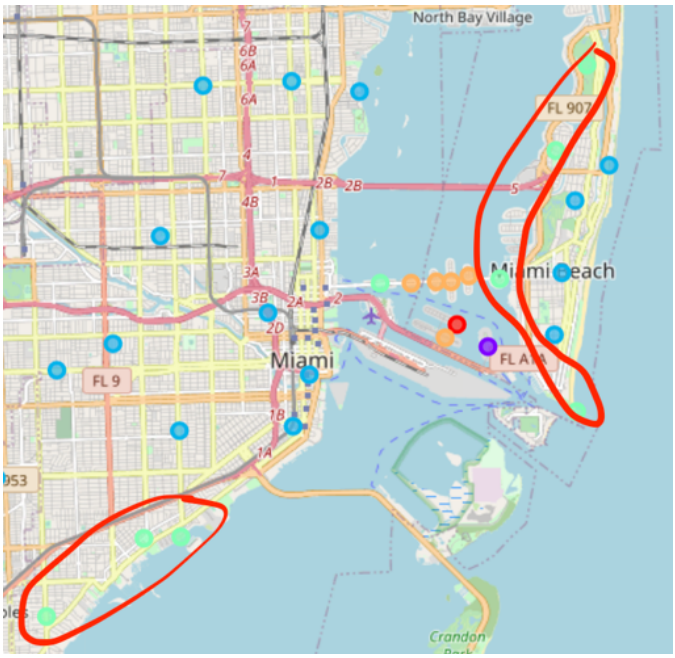
Looking at the map, we can see that all of these neighborhoods are small islands, which are right next to each other. They are conveniently located at the center between touristy Miami Beach (aka South Beach) and arty/foodie Miami (aka Downtown Miami). Looking at the top 10 venues in these islands, we can see that these upscale areas filled with water views, spas, places to exercise, and nearby hotels. They all have a golf course or park in the vicinity. There are no schools or stores. The violent crime rate is 50% less than the average of Miami. These locations obviously come with the price. The k-cluster algorithm rightly put Star Island in its own cluster (1) as it is especially expensive (typical property

value here is USD 28mln!). The houses on the Star Islands are much bigger than on other islands and this island has some celebrity owners (hence the name of the neighborhood).

Cluster Labels	Neighborhood	City	Typical Property Price, USD	Restaurant Freq	Store Freq	Sport Freq	Art Freq	Golf|Park Freq	Waterfront	School Count	Crime per 100k
0	Hibiscus Island	Miami Beach	3,798,735.000	0.110	0.012	0.098	0.000	0.049	1	0	237
1	Star Island	Miami Beach	28,298,329.000	0.273	0.078	0.104	0.000	0.052	1	1	237
4	Di Lido Island	Miami Beach	4,730,846.000	0.409	0.091	0.136	0.000	0.068	1	0	237
4	San Marco Island	Miami	4,928,254.000	0.167	0.000	0.000	0.033	0.033	1	0	237
4	Rivo Alto Island	Miami Beach	4,977,389.000	0.420	0.200	0.100	0.010	0.030	1	0	237
4	Palm Island	Miami Beach	5,274,093.000	0.083	0.000	0.000	0.000	0.083	1	0	237
4	San Marino Island	Miami Beach	5,536,186.000	0.227	0.000	0.045	0.091	0.091	1	0	237

"High-End Residential Neighborhoods":

Belle Isle, South Pointe, Nautilus, Biscayne Point, and La Gorce in Miami Beach; NE and SW Coconut Grove, and Fair Isle in Coconut Grove.

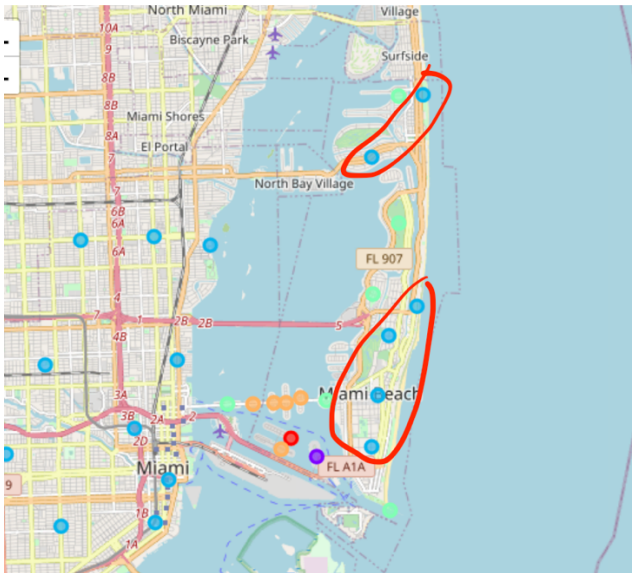


Miami Beach neighborhoods have condos and a higher crime rate given it is at the heart of tourism. Coconut Grove mostly houses, which is quieter with less crime and more schools nearby. Both areas are by the water and next to golf courses, but much more livable than upscale island cluster. They have higher restaurants density. The typical property prices here around USD 1mln; except La Gorce, which is $2mln (because it has its own golf club).

Cluster Labels	Neighborhood	City	Typical Property Price, USD	Restaurant Freq	Store Freq	Sport Freq	Art Freq	Golf|Park Freq	Waterfront	School Count	Crime per 100k
3	Belle Isle	Miami Beach	685,340.000	0.360	0.220	0.100	0.040	0.030	1	0	237
3	North-East Coconut Grove	Miami	714,243.000	0.380	0.230	0.080	0.000	0.040	0	2	233
3	South Pointe	Miami Beach	744,609.000	0.432	0.084	0.042	0.011	0.021	2	1	852
3	Fair Isle	Miami	761,856.000	0.484	0.161	0.000	0.032	0.032	1	1	187
3	South-West Coconut Grove	Miami	891,841.000	0.350	0.360	0.080	0.010	0.020	0	4	368
3	Biscayne Island	Miami	971,736.000	0.208	0.052	0.042	0.146	0.042	1	2	237
3	Biscayne Point	Miami Beach	977,991.000	0.361	0.103	0.010	0.021	0.041	2	1	852
3	Nautilus	Miami Beach	1,012,608.000	0.263	0.070	0.018	0.035	0.000	1	2	852
3	La Gorce	Miami Beach	2,028,854.000	0.203	0.068	0.017	0.000	0.034	1	0	237

"Mid-Tier Neighborhoods of Miami Beach":

North Shore, Isle of Normandy, Bayshore, Oceanfront, City Center, and Flamingo Lummus.

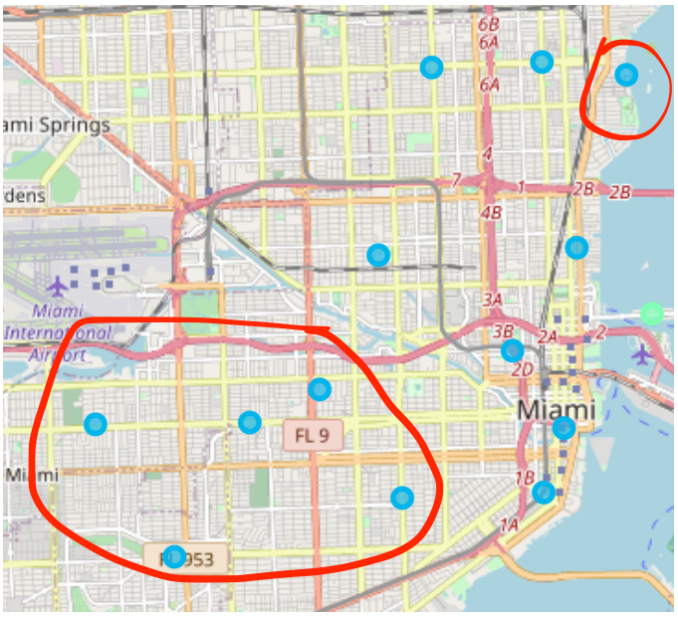


Typical prices of USD 300-500k. Given they are in the heart of touristy areas, they have so much to offer: many restaurants, bars, hotels in the vicinity. But this comes with a price of double the average Miami crime rate and smaller condo apartments.

Cluster Labels	Neighborhood	City	Typical Property Price, USD	Restaurant Freq	Store Freq	Sport Freq	Art Freq	Golf|Park Freq	Waterfront	School Count	Crime per 100k
2	Flamingo Lummus	Miami Beach	257,367.000	0.320	0.110	0.050	0.040	0.020	1	2	852
2	North Shore	Miami Beach	288,333.000	0.333	0.065	0.011	0.022	0.032	1	1	852
2	Isle of Normandy	Miami Beach	289,870.000	0.400	0.120	0.010	0.020	0.040	2	2	852
2	City Center	Miami Beach	327,291.000	0.350	0.170	0.000	0.050	0.010	1	2	852
2	Oceanfront	Miami Beach	449,860.000	0.290	0.050	0.010	0.040	0.010	1	2	852
2	Bayshore	Miami Beach	560,268.000	0.300	0.060	0.010	0.040	0.020	1	3	852

"Mid-Tier Neighborhoods of Miami":

Shenandoah, Coral Way, Alameda-West Flagler, Flagami, Upper Eastside, Little Havana.



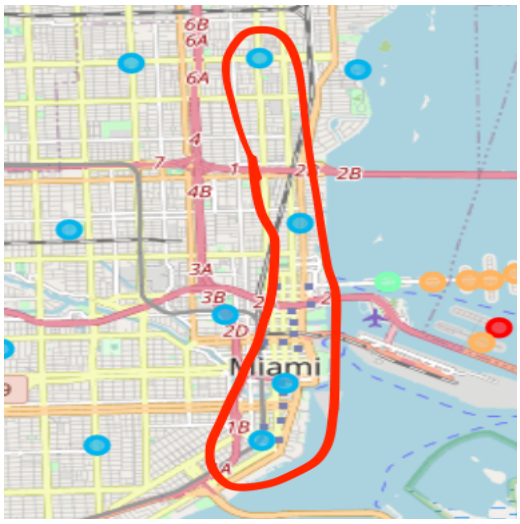
Typical prices of USD 300-500k. No beach, but some golf or park in the vicinity. They are mostly residential because have high density of stores and restaurants. A lot of public schools in the area. Good density of cultural activities as well. The crime rate is lower than

Miami Beach but higher than Upscale islands. These are excellent areas to live in for families.

Cluster Labels	Neighborhood	City	Typical Property Price, USD	Restaurant Freq	Store Freq	Sport Freq	Art Freq	Golf|Park Freq	Waterfront	School Count	Crime per 100k
2	Little Havana	Miami	210,996.000	0.303	0.224	0.000	0.092	0.039	0	5	585
2	Flagami	Miami	341,027.000	0.378	0.257	0.014	0.000	0.027	0	4	227
2	Alameda - West Flagler	Miami	397,024.000	0.441	0.215	0.000	0.032	0.032	0	4	269
2	Coral Way	Miami	482,235.000	0.480	0.180	0.040	0.020	0.010	0	2	187
2	Upper Eastside	Miami	506,184.000	0.422	0.157	0.084	0.048	0.048	0	2	751
2	Shenandoah	Miami	548,980.000	0.520	0.190	0.000	0.100	0.030	0	4	187

"Trendy Miami Neighborhoods":

Downtown, Brickell, and Wynwood-Edgewater.



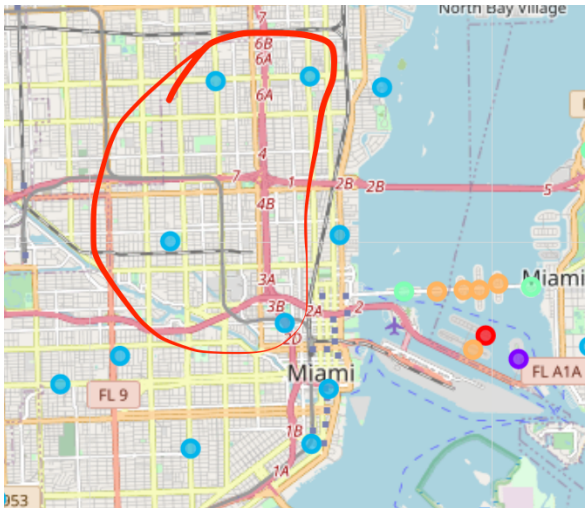
Historically, these neighborhoods have been dangerous, and still, the violent crime rate is almost 3 of Miami's average, but things are changing. Miami Police headquarters are between Overtown and Downtown. Despite the crime, these areas are considered trendy because each neighborhood has its unique purpose. Downtown has business/government/museums, Brickell is a business district, and Wynwood is an art district. These neighborhoods have high-rise condo residential developments along the water along Brickell Ave, Biscayne Blvd, and Edgewater. They have the highest density of restaurants and cultural activities in Miami. Yet the neighborhoods don't end there but go

much more inland, where unfortunately still poverty, homeless people, and crime. There is no beach but some parks in the vicinity.

Cluster Labels	Neighborhood	City	Typical Property Price, USD	Restaurant Freq	Store Freq	Sport Freq	Art Freq	Golf|Park Freq	Waterfront	School Count	Crime per 100k
2	Downtown	Miami	351,345.000	0.480	0.120	0.030	0.030	0.020	0	5	1195
2	Brickell	Miami	392,743.000	0.480	0.140	0.060	0.000	0.000	0	2	1195
2	Wynwood - Edgewater	Miami	401,578.000	0.410	0.260	0.050	0.100	0.020	0	8	782

"Dangerous Lower-Tier Neighborhoods":

Liberty City, Allapattah, Little Haiti, and Overtown.



These neighborhoods are poor and don't have any waterfront. And have more than double the average crime rate of Miami. Liberty City is an exceptional case; with 1919 violent crimes per 100k people, it is considered one of the most dangerous neighborhoods in the US. Typical property prices here are USD 200-300k. Not surprisingly, Miami policy headquarter is located in Overtown.

Cluster Labels	Neighborhood	City	Typical Property Price, USD	Restaurant Freq	Store Freq	Sport Freq	Art Freq	Golf|Park Freq	Waterfront	School Count	Crime per 100k
2	Overtown	Miami	212,994.000	0.400	0.110	0.060	0.170	0.010	0	10	851
2	Liberty City	Miami	282,969.000	0.194	0.387	0.032	0.000	0.065	0	10	1919
2	Allapattah	Miami	288,712.000	0.250	0.304	0.018	0.036	0.018	0	4	944
2	Little Haiti	Miami	370,677.000	0.400	0.200	0.074	0.053	0.011	0	6	1138

5. Conclusions

The purpose of this project was to identify the best neighborhoods in Miami to aid stakeholders in narrowing down the search for an optimal location to buy or rent a property. By calculating the density of schools, crime rates, restaurants, stores, sport, and cultural activities, I could give a general quantitative description for each neighborhood.

I identified the following neighborhood clusters in Miami:

- "Upscale Islands": Star, Palm, and Venetian, Hibiscus islands in descending price order from USD 28 to 4mln (Houses only.) They are located conveniently between Miami Beach and Miami. They all have a waterfront and golf course or park in the vicinity. There are no stores or schools nearby but many places to go out for restaurants or exercise. The violent crime rate is 237 per 100k people, which is 50% less than the average of Miami.

- "High-End Residential Neighborhoods": South Beach and Coconut Grove by the water and near golf courses. (Many are houses) Neighborhoods include Belle Isle, South Pointe, Nautilus, Biscayne Point, and La Gorce in Miami Beach; NE and SW Coconut Grove, and Fair Isle in Coconut Grove. Typical property prices are around USD 1 million, except La Gorce, which is $2mln (because it has its own golf club). These neighborhoods are more residential because they have double the frequency of "Upscale Islands" in the number of stores, restaurants, and schools nearby. The crime rate is very low, similar to "upscale Islands," except some Miami Beach neighborhoods such as South Pointe, Biscayne Point, and Nautilus, which have a violent crime rate of 850 per 100k people or 100% more than the average of Miami.

These three neighborhoods are close to many touristy hotels and hence more foot traffic and crime in the vicinity.

- "Mid-Tier Neighborhoods of Miami Beach": North Shore, Isle of Normandy, Bayshore, Oceanfront, City Center, and Flamingo Lummus. Typical prices of USD 300-500k. Given they are in the heart of touristy areas, they have so much to offer: many restaurants, bars, hotels in the vicinity. But this comes with a price of double the average Miami crime rate and smaller condo apartments.

- "Mid-Tier Neighborhoods of Miami": Shenandoah, Coral Way, Alameda-West Flagler, Flagami, Upper Eastside, Little Havana. Typical prices of USD 400-550k. No beach but golf or park in the vicinity. They are more residential because they have a double frequency of "Upscale Islands" in stores, restaurants, and schools. The crime rate is lower than Miami Beach but higher than the Upscale islands. They are excellent areas to live in for families.

- "Trendy Neighborhoods": Downtown, Brickell, and Wynwood-Edgewater. Historically, these neighborhoods have been dangerous, and still, the violent crime rate is almost 3 of Miami's average, but things are changing. Miami Police headquarters are between Overtown and Downtown. Despite the crime, these areas are considered trendy because each neighborhood has its unique purpose. Downtown has business/government/museums, Brickell is a business district, and Wynwood is an art district. These neighborhoods have high-rise condo residential developments along the water along Brickell Ave, Biscayne Blvd, and Edgewater. They have the highest density of restaurants and cultural activities in Miami. Yet the neighborhoods don't end there but go much more inland, where unfortunately still poverty, homeless people, and crime. There is no beach but some parks in the vicinity.

- "Dangerous Lower-Tier Neighborhoods": Liberty City, Allapattah, Little Haiti, and Overtown. These neighborhoods are poor and don't have any waterfront. And have more than double the average crime rate of Miami. Liberty City is an exceptional case; with 1919 violent crimes per 100k people, it is considered one of the most dangerous neighborhoods

in the US. Typical property prices here are USD 200-300k. Not surprisingly, Miami policy headquarter is located in Overtown.

My project focused on location factors such as availability and density of restaurants, schools, stores, art, places one could go to exercise, and crime rates. I identified that the waterfront, the availability of a park, golf course, places one could go to exercise, and the low crime rate significantly affects property prices. However, the density of schools, restaurants, stores, and art negatively correlates with property prices. This is surprising, but I could see how it makes sense. First, we don't know the rankings of schools/art/stores nearby. To have density doesn't imply quality. Second, Miami's nature is tourism. Tourist influx brings a lot of foot traffic and a high probability of crime to specific neighborhoods (South Beach, Downtown, Wynwood, and Brickell). Finally, the density of public schools doesn't affect the high-end real estate market. This is probably because of the demographics of high-net-worth individuals. They are usually not full-time residents of Miami. That's why most of these buyers are after the convenience of the location and fantastic water views. (see Upscale Islands or High-end residential clusters)

My results might be surprising for some Miami locals (including me) because the violent crime rate is higher than it is perceived. For example, the trendy neighborhood cluster (Downtown, Wynwood, and Brickell) has double Miami's average crime rate. This makes sense if you look on the map to see how Google API defines a particular neighborhood area. For example, according to Google Maps Brickell, Downtown, Wynwood/Edgewater areas are much bigger than the perceived narrow strip along the water. The neighborhoods don't end with high-rise luxury residential condominiums along Brickell Ave, Biscayne Blvd, and Edgewater, but go further inland. And here, unfortunately, we still have a lot of poverty, homeless people, and crime.

In conclusion, waterfront, a healthy lifestyle, and tourism define the Miami area real estate. These factors might not work in other cities, but one could use my data science methodology to discover what works in the city of their choice.

6. Future directions

The final decision on optimal property location is a complex process that entails many choices. I only cover some factors; others include noise levels, proximity to major roads, real estate availability, prices per square foot, area of the property, year build, social and economic dynamics, etc.

I also see how I could improve on my data science research. First, I used almost two years old data from the FBI, which only published calendar 2020 data in September 2021. The situation might have improved since then. Second, it will be helpful to include the ranking of schools and restaurants and see if proximity to them affects property prices.

My project focused on location factors such as availability and density of restaurants, schools, stores, art, places one could go to exercise, and crime rates. I discovered that waterfront, a healthy lifestyle, and tourism define the Miami area real estate. These factors might not work in other cities, but one could use my data science methodology to discover what works in the city of their choice.