

## INTRODUCTION

Whenever a restaurant starts to get famous and busy, it will eventually get into a major problem: lack of space. Since physical space is something which can be quite limiting in this kind of business, restaurant owners can always rise prices in order to meet the demand without increasing the offer. However, this price increase has also a ceiling and, whenever it's reach with the restaurant still quite busy, there's a problem where owner can feel he's losing money – which in fact he is, once, if he had more space, he could make more money. In this case, there are two solutions that can be met:

- a) Expand the current space;
- b) Open a branch in some other part of the city, rising, in this way, its target audience.

In this exercise, I will analyze a Mandau Bistro, a non-existent restaurant based on a real one, which is located at the Designers District, in Miami, FL. In the hypothetical situation, they have reached the ceiling and want to open a new branch in the city and now need to decide where to open it. The main idea is to find a neighborhood similar to its actual one as far as possible, in order to increase its influence radius in the city.

## DATA

In order to solve the problem, simple data will be used: a list of Miami neighborhoods, provided by Wikipedia, which includes Latitude and Longitude data, as well as population and density. The table can be found in the following URL: [https://en.wikipedia.org/wiki/List\\_of\\_neighborhoods\\_in\\_Miami](https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Miami). Also, in order to be able to classify the neighborhoods, we want to understand the most common venues in each neighborhood, as well as how many public venues there are. For this, the Foursquare API will be used.

## METHODOLOGY

To obtain this results, I will be dealing with the table mentioned before with pandas library on Python, preparing the tables I need in order to do such analyzes. We will need a table containing:

- Neighborhood name;
- Latitude;
- Longitude;
- Population;
- Density;
- Total number of venues;
- Type of venues per neighborhood.

Also, the four last venues will be normalized, so that the values magnitude won't interfere in the division I'm willing to do. Finally, with the data all prepared, we can apply the algorithm – K-Means – in order to cluster the data. The neighborhoods will be divided into 5 clusters, which is a good number as there are only 21 neighborhoods in the analysis

but I want to divide it into small clusters so that it will be possible to find really similar neighborhoods in this process.

After it's clustered, we can analyze only those neighborhoods within the same cluster as the Design District neighborhood, taking into account the distance from the current neighborhood as well as socioeconomic data.

## RESULTS

After applying the described methodology, the results were found as below: there are 8 neighborhoods (from the 21 analyzed) within the Cluster 0 (red one on the map) – the same cluster as the Design District. As it's possible to see, those are neighborhoods with similar density, in a wide population range, on which the most common venues include Cafes, Bars and Restaurants – different types of foods. This shows that in fact this is the ideal cluster to open a new bistro, since those are the neighborhoods which seems to have a more busy nightlife

Neighborhood	Population2010	Population/Km²	Lat	Long	Cluster Labels	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	Distance	Number of Categories
Buena Vista	9058	3540	25.813	-80.192	0	Jewelry Store	Café	Boutique	Italian Restaurant	Coffee Shop	Furniture / Home Store	Shopping Mall	Fast Food Restaurant	Big Box Store	Mediterranean Restaurant	0.312014	26
Design District	3573	3623	25.813	-80.193	0	Jewelry Store	Café	Boutique	Italian Restaurant	Coffee Shop	Furniture / Home Store	Shopping Mall	Fast Food Restaurant	Big Box Store	Mediterranean Restaurant	0.340687	26
Upper Eastside	12525	2513	25.830	-80.183	0	Gas Station	Pizza Place	Motel	Intersection	Tennis Court	Italian Restaurant	Furniture / Home Store	Sushi Restaurant	Yoga Studio	Farmers Market	1.799492	8
Overtown	6736	3405	25.787	-80.201	0	Wings Joint	Southern / Soul Food Restaurant	Baseball Stadium	Farmers Market	Athletics & Sports	Event Space	Food	Flea Market	Fish Market	Fast Food Restaurant	3.337442	5
Park West	4655	3635	25.785	-80.193	0	Nightclub	Café	Bar	Lounge	Theater	Hotel	Gastropub	Japanese Restaurant	Flea Market	Event Space	3.426426	28
Lummas Park	3027	3680	25.777	-80.201	0	Seafood Restaurant	Hotel	American Restaurant	Fish Market	BBQ Joint	Historic Site	Restaurant	Spanish Restaurant	Soccer Field	Golf Course	4.415501	14
The Roads	7327	4899	25.756	-80.207	0	Lawyer	Health & Beauty Service	Moving Target	Business Service	Frozen Yogurt Shop	Gas Station	Italian Restaurant	Latin American Restaurant	Farmers Market	Fast Food Restaurant	6.826385	9
Grapeland Heights	14004	4130	25.792	-80.258	0	Hotel	Gym / Fitness Center	Restaurant	Bar	Gym	Golf Course	Hotel Pool	Dessert Shop	Diner	Discount Store	7.159026	7

The other clusters were filled as below:

Neighborhood	Population2010	Population/Km²	Lat	Long	Cluster Labels	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	Distance	Number of Categories
Allapattah	54289	4401	25.815	-80.224	1	Boxing Gym	Department Store	Lounge	Event Space	Food	Flea Market	Fish Market	Fast Food Restaurant	Farmers Market	Yoga Studio	3.248677	3
Coconut Grove	20076	3091	25.712	-80.257	1	American Restaurant	Boat or Ferry	Fried Chicken Joint	Park	Yoga Studio	Empanada Restaurant	Fish Market	Fast Food Restaurant	Farmers Market	Event Space	13.274360	4
Coral Way	35062	4496	25.750	-80.283	1	Liquor Store	Pharmacy	Dive Bar	Seafood Restaurant	Burger Joint	Café	Golf Course	Mobile Phone Shop	Historic Site	Discount Store	11.721493	10
Flagami	50834	5665	25.762	-80.316	1	Seafood Restaurant	Bakery	Pet Store	Department Store	Fast Food Restaurant	Spanish Restaurant	Latin American Restaurant	Fish Market	Farmers Market	Event Space	13.823434	8
Liberty City	19725	3733	25.832	-80.225	1	Southern / Soul Food Restaurant	Grocery Store	Donut Shop	Flea Market	Fish Market	Fast Food Restaurant	Farmers Market	Event Space	Empanada Restaurant	Dog Run	3.802759	2
Little Haiti	29760	3840	25.824	-80.191	1	Yoga Studio	Shopping Mall	Gas Station	Ice Cream Shop	Italian Restaurant	Fast Food Restaurant	Liquor Store	Donut Shop	Dance Studio	Coffee Shop	0.916304	23
West Flagler	31407	4428	25.775	-80.243	1	Cuban Restaurant	Bakery	Pharmacy	Concert Hall	Comfort Food Restaurant	Gas Station	Latin American Restaurant	Fast Food Restaurant	Farmers Market	Event Space	6.863234	7

**Cluster 1 (purple):** it's also a cluster based on restaurants, but also with a small number of venues and a high value of population – it shows us it's a more residential cluster.

	Neighborhood	Population2010	Population/Km²	Lat	Long	Cluster Labels	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	Distance	Number of Venues
13	Little Havana	76163	8423	25.773	-80.215	2	Latin American Restaurant	Shoe Store	Pharmacy	Grocery Store	Discount Store	Mexican Restaurant	Restaurant	Fast Food Restaurant	Farmers Market	Event Space	5.305431	9

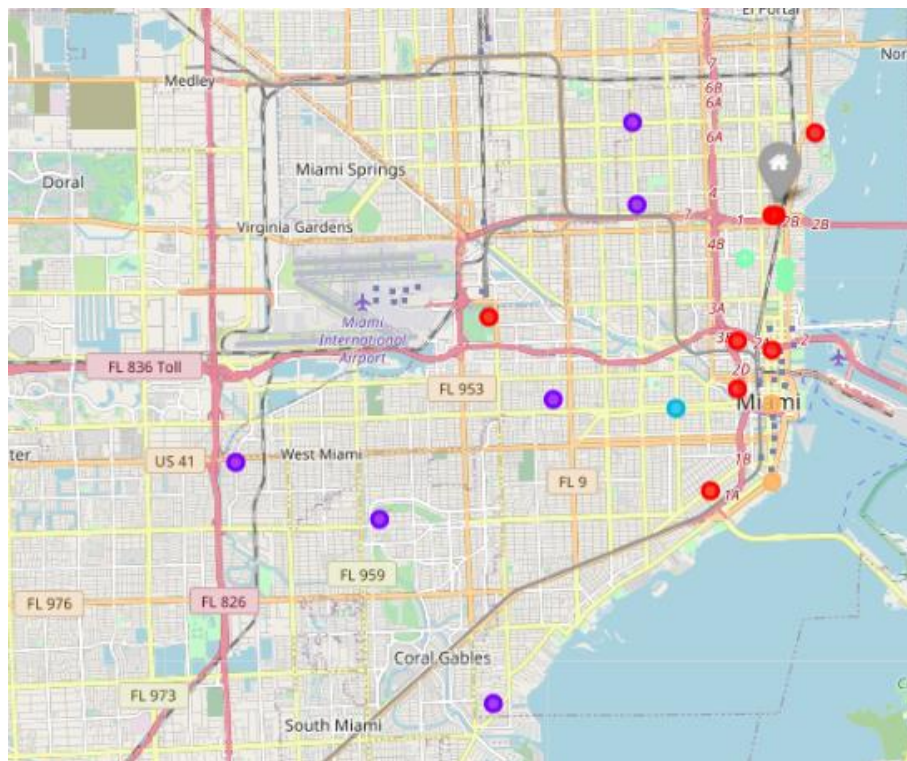
**Cluster 2 (blue):** Filled only with Little Havana, it's a quite big neighborhood with few venues: another example of a residential cluster.

	Neighborhood	Population2010	Population/Km²	Lat	Long	Cluster Labels	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	Distance	Number of Venues
1	Arts & Entertainment District	11033	7948	25.799	-80.190	3	Restaurant	Gym	Yoga Studio	Spa	Ice Cream Shop	Food Truck	Moving Target	Cupcake Shop	Paper / Office Supplies Store	Park	1.872783	19
8	Edgewater	15005	6675	25.802	-80.190	3	Restaurant	Gym	Sandwich Place	Art Gallery	Breakfast Spot	Pizza Place	Peruvian Restaurant	Cuban Restaurant	Coffee Shop	Ice Cream Shop	1.540520	48
20	Wynwood	7277	2983	25.804	-80.199	3	Art Gallery	Bar	Café	Food Truck	Theater	Peruvian Restaurant	Bakery	Caribbean Restaurant	Public Art	Coffee Shop	1.506737	80

**Cluster 3 (green):** It's more similar to cluster 0, with restaurants as the main type of venue. However, it's possible to see different venues types, such as gyms, yoga studios and art galleries.

	Neighborhood	Population2010	Population/Km²	Lat	Long	Cluster Labels	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	Distance	Number of Venues
2	Brickell	31759	14541	25.758	-80.193	4	Hotel	Italian Restaurant	Japanese Restaurant	Café	Bank	Juice Bar	Salon / Barbershop	Sandwich Place	Seafood Restaurant	Bar	6.428227	58
7	Downtown	71000	10613	25.774	-80.193	4	Hotel	Peruvian Restaurant	Café	Italian Restaurant	Brazilian Restaurant	Cosmetics Shop	Gym	Pharmacy	Coffee Shop	Lounge	4.649163	50

**Cluster 4 (orange):** It's a hotel cluster with some restaurants and cafes also.



## **DISCUSSION**

As we can see from the previous section, Cluster 0 is a quite busy cluster, with many neighborhoods on it. The most obvious choice for our Bistro to open is Grapeland heights: the most distant neighborhood from the Design District within the same cluster.

However, it's important to mention that further analysis should be done in order to guarantee that this is really the best option. For example, socioeconomical data should be analyzed, so that it could be understood whether the prices for the actual Bistro will be adequate to the new neighborhood. Also, housing price is something that should be taken into account, since, if it's too expensive to open this new place, it should not be worth it to open in the "most adequate" neighborhood and, then, new and more cheaper options can be analyzed.

## **CONCLUSION**

Although it's obvious that some subjective information should be taken into account to take a decision in that magnitude, it was also showed that Machine Learning can be a helpful hand to funnel the analysis into less data – which can be more comprehensive to the human brain. Also, it's important to note that the quality of the data utilized to make such analysis is the most important part of the whole process, and having the correct data is fundamental to obtain a consistent result.

In this case, we could set that our new bistro would be open on the Grapeland heights neighborhood through the numerical process applied. However, it's from that point that the human capabilities come into scene to validate this information and proceed with further analysis to obtain more detailed data.