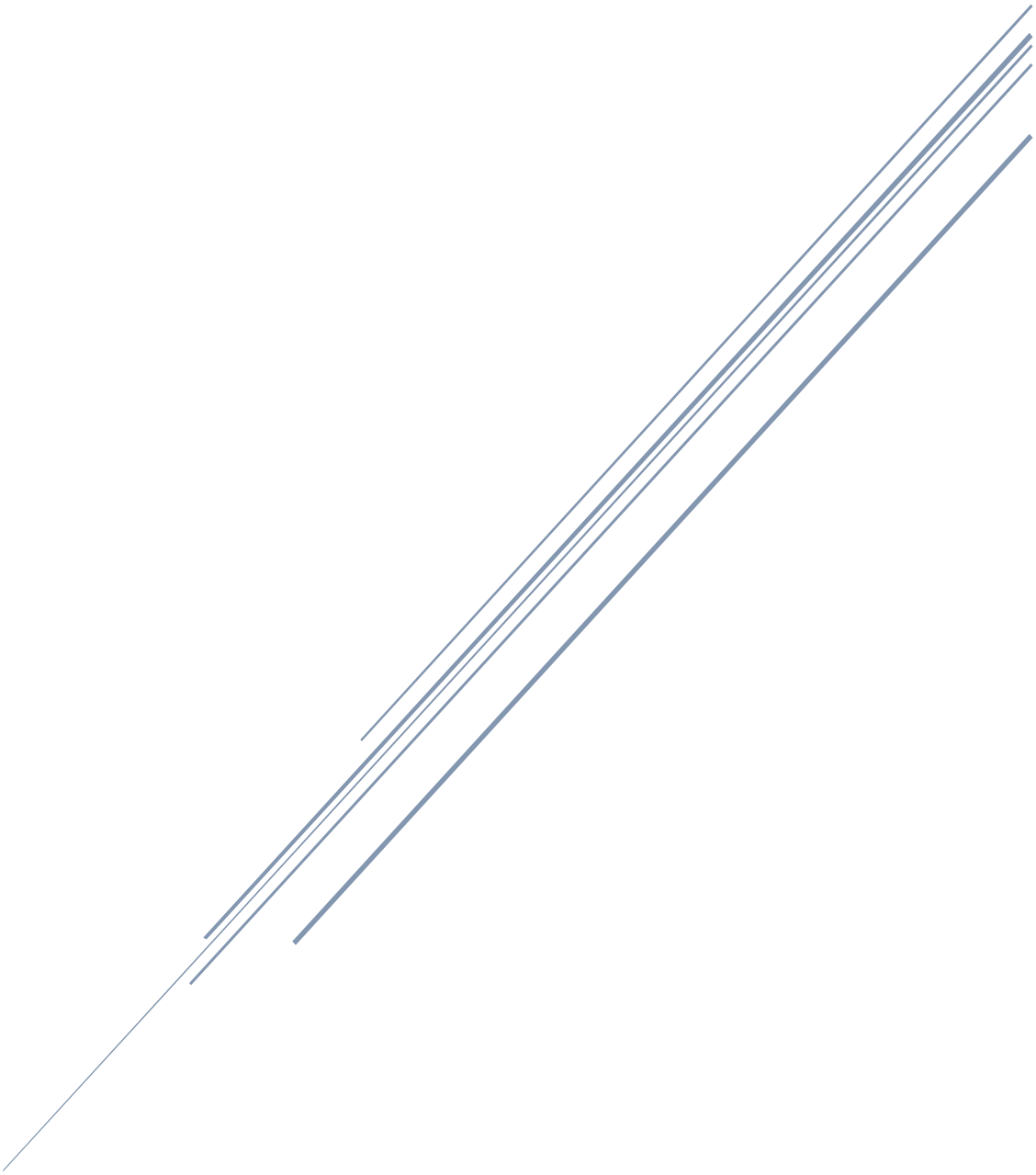
COURSERA CAPSTONE PROJECT

Opening an Italian Restaurant in Miami



March 2021

Business Problem

A friend of mine would like to open an Italian restaurant. He would like to open this restaurant in Miami. He knows there are already a lot of Italian restaurants in Miami, but he doesn't know where they are located and how well they are doing. He asked me to investigate the boroughs of Miami to see where other Italian restaurants are located and how well they are doing. This way he can determine what the best borough will be to open his own Italian restaurant.

Data

The data I will be using will be coming from the Foursquare API. I will fetch the venue categories in the different boroughs and look at the most common venue in the boroughs. I will also be using the API to get venue search and venue likes, to see which boroughs have the most liked Italian restaurants. I will also be using the Wikipedia page of the Boroughs of Miami and their latitude and longitude information.

https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Miami

I will combine the list of boroughs (with their coordinates) with the Foursquare API data. This will give a good overview of the boroughs with their most common venues, but also with their most recommended venues.

Methodology

Before starting the data analysis, first the data needed to be cleaned. The table from the Wikipedia page had a few obsolete columns, a row without coordinates, as well as a totals row. After removing the boroughs without coordinates and the total row, I had to split the coordinates column into latitude and longitude (Figure 1). After the split I wanted to plot the data on a map (Figure 2). This way we get an idea of where the boroughs are located and how close they are from one another. We can see that most boroughs lay close to the water. The boroughs farter from the water are further apart. Seeing this, I would expect the boroughs close to the water to be similar to eachother.

	Borough	Latitude	Longitude
0	Allapattah	25.815	-80.224
1	Arts & Entertainment District	25.799	-80.190
2	Brickell	25.758	-80.193
3	Buena Vista	25.813	-80.192
4	Coconut Grove	25.712	-80.257
5	Coral Way	25.750	-80.283
6	Design District	25.813	-80.193
7	Downtown	25.774	-80.193
8	Edgewater	25.802	-80.190
9	Flagami	25.762	-80.316

Figure 1 First rows of Miami boroughs table

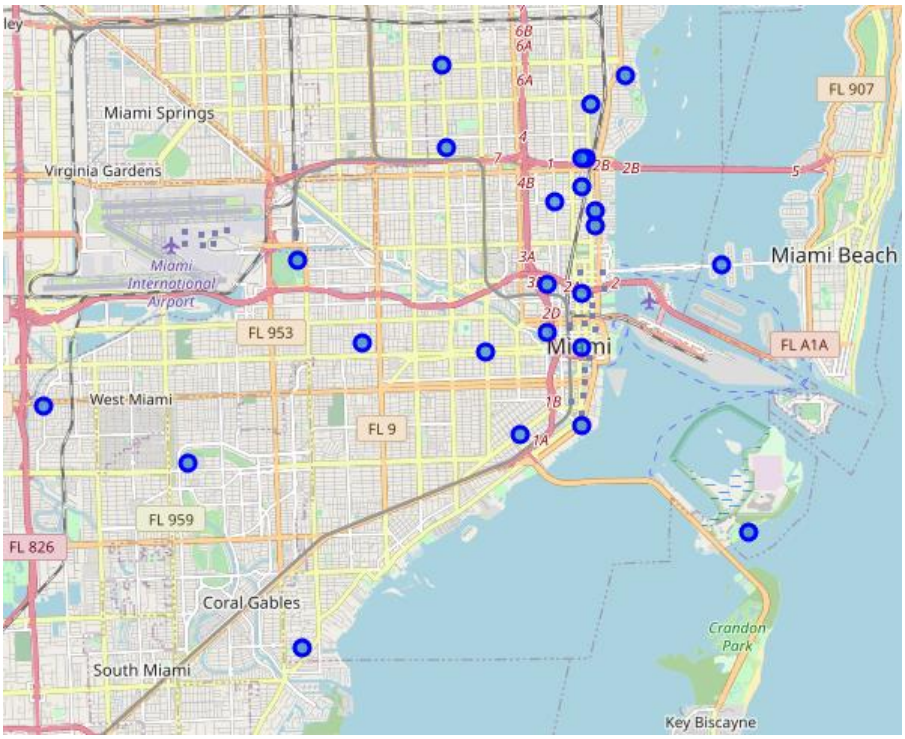


Figure 2 Boroughs of Miami on a map

After this first look, it was time to get the extra data from the Foursquare API. I used the API to get the top 100 venue information from the boroughs. I used one hot encoding to get a better look at the Italian restaurant category. Seeing the low frequencies of the different venues, I decided to only look at the boroughs that have the Italian restaurant category in their top 3. This brought the number of boroughs back from 24 to 8. I plotted these boroughs to get an idea of their locations (Figure 3). Purple locations are boroughs with Italian restaurants as one of the three most common venues, red locations are the other boroughs.

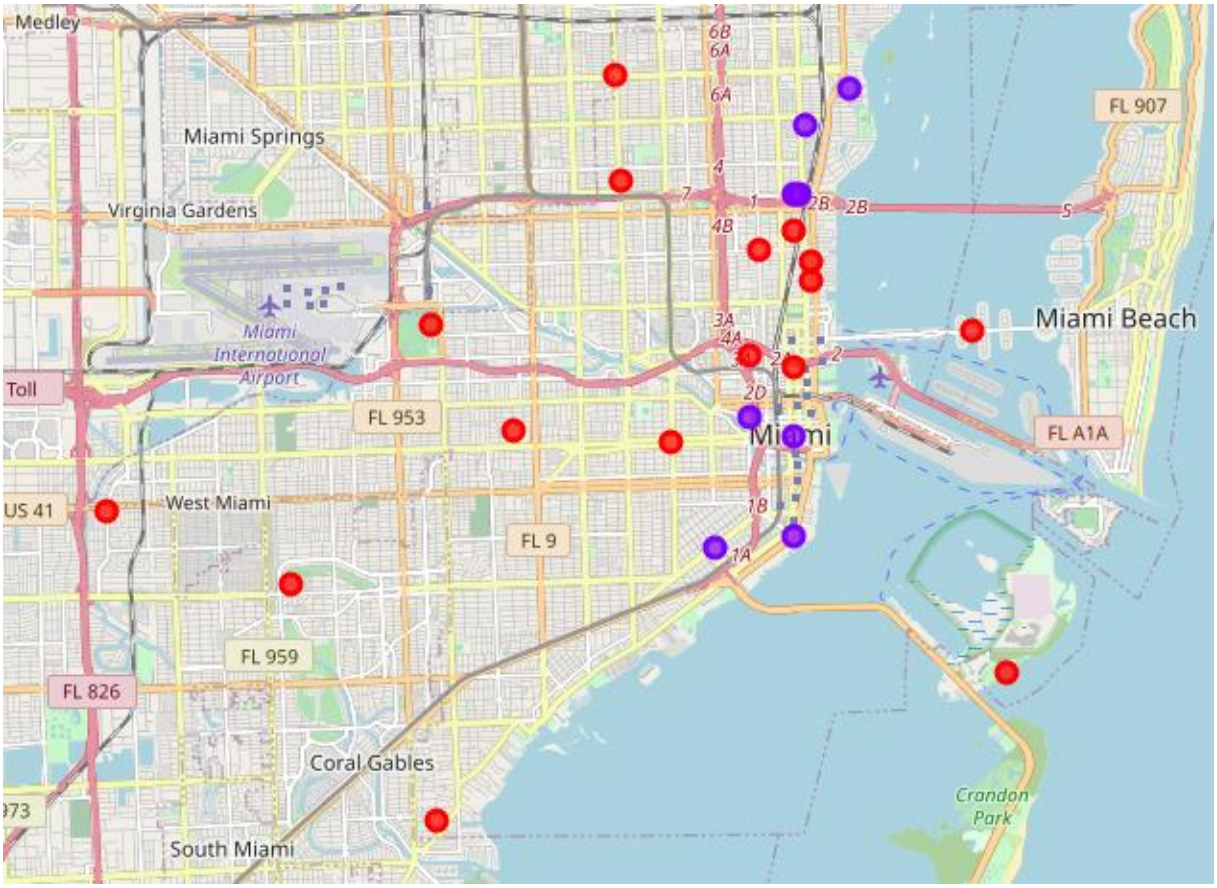


Figure 3 Boroughs with Italian restaurants as top 3 most common venue (purple) vs other boroughs (red)

For further investigation, I wanted to look at a few specific venues withing the purple boroughs. I want to see how many people liked the first 10 Italian restaurants for every borough the API gave back. I used the Foursquare API to get 10 venues within the category 'Italian Restaurant', using the venue ID's to get the amount of likes for all these venues. I grouped these likes by every borough (Figure 4). This table gives a good idea of the quality of the Italian restaurants within every borough.

	Borough	Venue Likes
0	Brickell	184
1	Buena Vista	238
2	Design District	787
3	Downtown	127
4	Little Haiti	33
5	Lummus Park	91
6	The Roads	40
7	Upper Eastside	79

Figure 4 Amount of likes per borough (left)

After looking into these purple clusters and their venue likes, I wanted to see how a K-Means model would cluster the boroughs of Miami. Combining the venue likes of the boroughs with the clustering from the K-Means, I feel comfortable with recommending a borough to start an Italian restaurant.

Results

The K-Means clustering gave the result as shown in Figure 5. I added boundaries to the map to make it easier to see the different clusters. I also added the number of each cluster.

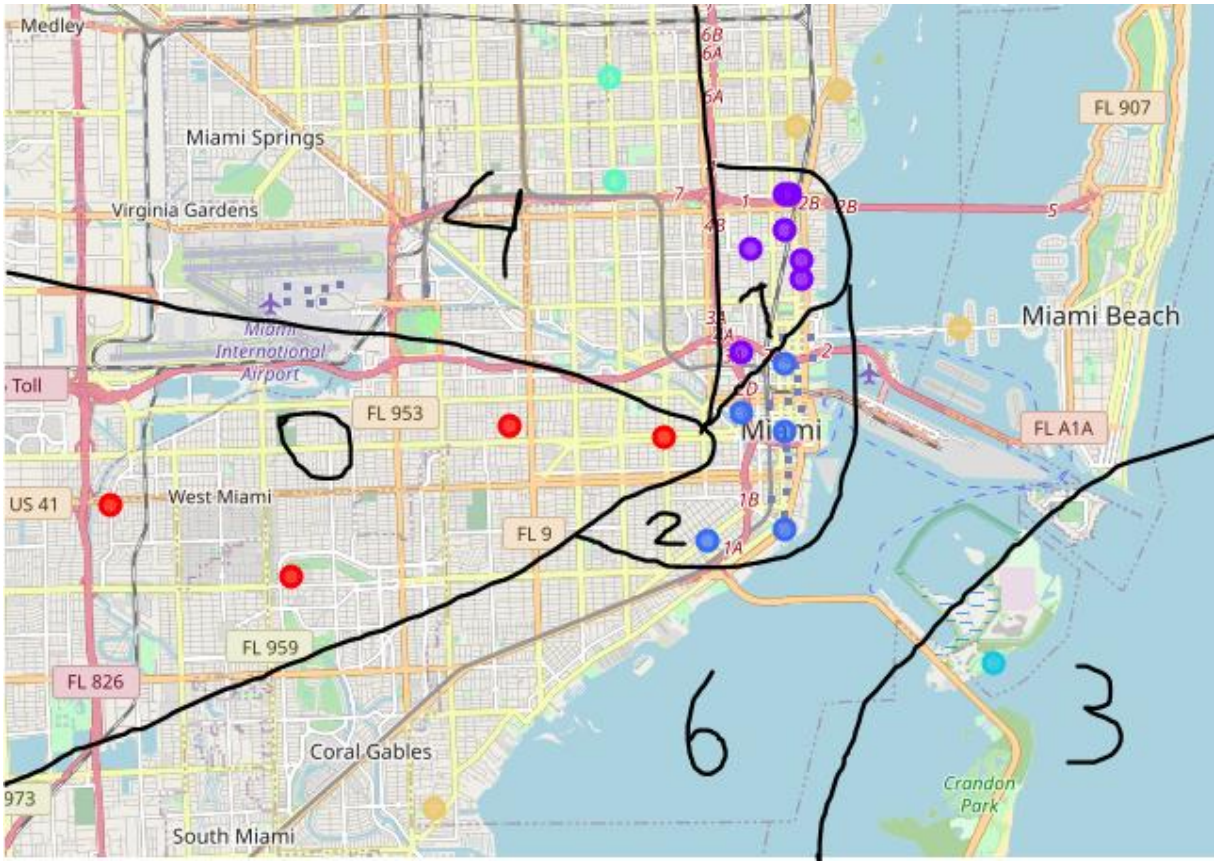


Figure 5 K-Means clusters of the Miami borough

I grouped the clusters with the venue likes data to get the table shown in Figure 6. The Italian Restaurants column is the number of boroughs where Italian restaurants were in the top 3 of most common venues. Venue Likes is the total likes for the Italian restaurants of the boroughs within the cluster. This table show a lot of likes in cluster 1 and cluster 2. An overview of all the boroughs with their venue likes, cluster group and the column that show if the borough had Italian restaurants as the most common venue can be seen in Figure 7.

	Cluster	Italian Restaurant	Venue Likes
0	0	0	0.0
1	1	2	1025.0
2	2	4	442.0
3	3	0	0.0
4	4	0	0.0
5	5	0	0.0
6	6	2	112.0

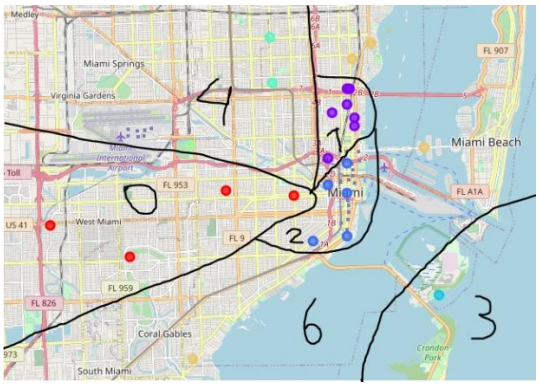
Figure 6 Total venue likes and boroughs with Italian restaurants per cluster

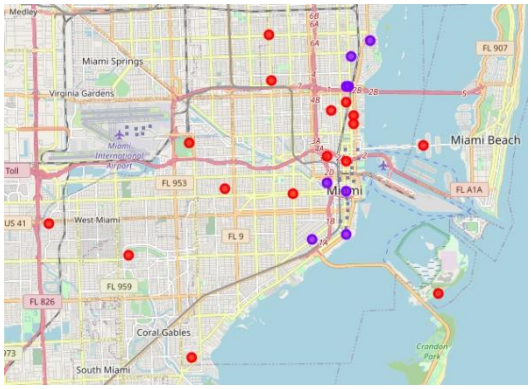
	Borough	Cluster	Latitude	Longitude	Italian Restaurant	Venue Likes
0	Allapattah	4	25.815	-80.224	0	0.0
1	Arts & Entertainment District	1	25.799	-80.190	0	0.0
2	Brickell	2	25.758	-80.193	1	184.0
3	Buena Vista	1	25.813	-80.192	1	238.0
4	Coconut Grove	6	25.712	-80.257	0	0.0
5	Coral Way	0	25.750	-80.283	0	0.0
6	Design District	1	25.813	-80.193	1	787.0
7	Downtown	2	25.774	-80.193	1	127.0
8	Edgewater	1	25.802	-80.190	0	0.0
9	Flagami	0	25.762	-80.316	0	0.0
10	Grapeland Heights	5	25.792	-80.258	0	0.0
11	Liberty City	4	25.832	-80.225	0	0.0
12	Little Haiti	6	25.824	-80.191	1	33.0
13	Little Havana	0	25.773	-80.215	0	0.0
14	Lummus Park	2	25.777	-80.201	1	91.0
15	Midtown	1	25.807	-80.193	0	0.0
16	Overtown	1	25.787	-80.201	0	0.0
17	Park West	2	25.785	-80.193	0	0.0
18	The Roads	2	25.756	-80.207	1	40.0
19	Upper Eastside	6	25.830	-80.183	1	79.0
20	Venetian Islands	6	25.791	-80.161	0	0.0
21	Virginia Key	3	25.736	-80.155	0	0.0
22	West Flagler	0	25.775	-80.243	0	0.0
23	Wynwood	1	25.804	-80.199	0	0.0

Figure 7 Table with all available borough information for this analysis

Discussion

Combing the K-Means clustering, total venue likes and the division on the Italian restaurant as most common venues (Figure 3, 5 and 6 below), it shows that cluster 1 and 4 have the most likes and cluster 4 has the most boroughs with Italian restaurants as most common venue.





	Cluster	Italian Restaurant	Venue Likes
0	0	0	0.0
1	1	2	1025.0
2	2	4	442.0
3	3	0	0.0
4	4	0	0.0
5	5	0	0.0
6	6	2	112.0

This gives multiple options as recommended borough for a new Italian restaurant. Looking only at the most liked venue, cluster 1 sticks out. This cluster contains 2 boroughs with highly liked Italian restaurants. Cluster 2 and 6 also stick out when we look at the boroughs with Italian restaurants as most common venue. Cluster 2 also takes the second place of clusters with the most venue likes. This leaves a few options, depending on what my friend wants for his business.

I will give my recommendations within a few scenarios. Depending on the needs of my friend, he has some options to choose from.

1) Go for the Borough with the most liked venues. This looks like the borough people go to for the very best Italian restaurants in Miami. This borough is Design District (Figure 4).
2) Go for a borough in the clusters with Italian restaurants being the most common venue. This is 2 out of 7 boroughs in cluster 1, 4 out of the 5 boroughs in cluster 2 and 2 out of 4 boroughs in cluster 6: Brickell, Buena Vista, Design District, Downtown, Little Haiti, Lummus Park, The Roads and Upper Eastside.
    a. You can take this a step further by also looking at the amount of likes for these boroughs. In that case Buena Vista will be on top with the most venue likes
3) Go for a borough with less or even no competition that is comparable to a borough in the same cluster that has Italian restaurants as common venue. This way you will be locating in an area that looks the same venue wise but has less competition. Because the boroughs were grouped together, we might expect the customer needs in these boroughs to be the same. This way you would fill a gab in these customer needs. This would be Arts & Entertainment District, Coconut Grove, Edgewater, Midtown, Overtown, Park West, Venetian Islands and Wynwood.
4) Go for a borough that doesn't fall in the other three recommendations. Even less competition than recommendation 3.

Conclusion

The boroughs in Miami are very interesting to analyze. This analysis focused on finding a good borough for a new Italian restaurant. Recommendations were made with the use of the Foursquare API, looking at a few Italian restaurants for interesting boroughs and using a model to group similar boroughs. After taking a closer look at the boroughs and some of the liked venues from the gather data and clusters, I came up with the following recommendations.

1) High risk, high reward: Design District is the borough with the most liked Italian restaurants in Miami. This seems to be the borough people go to for the best Italian restaurants. If you think you can handle the competition, this is the borough for you.

2) Go with the flow: A few boroughs seem to have popular Italian restaurants and some of these boroughs are also close together. The boroughs might be the place where the common man goes searching for Italian restaurants. Starting a business here means you are in the area where people will be looking for your type of restaurant. Boroughs in this group: Brickell, Buena Vista, Design District, Downtown, Little Haiti, Lummus Park, The Roads and Upper Eastside. If you are also looking for the borough with a lot of venue likes, Buena Vista is the one for you. You can probably expect a lot of competitions in these boroughs.

3) Not many competitors: By choosing a borough that is in the same cluster as a borough from recommendation 1 or 2, we find ourselves in an area that is comparable to the other boroughs, but doesn't have a lot of competition with the Italian restaurant scene. This way you could be filling a gab within the customer needs. This applies to Arts & Entertainment District, Coconut Grove, Edgewater, Midtown, Overtown, Park West, Venetian Islands and Wynwood.

4) Even less competitors: The remaining 10 boroughs are areas where there are even less Italian restaurants than the boroughs from recommendation 3. This might also be a good place to start your business. Starting here means you could be the only restaurant in close range, for people that don't want to eat far from their homes. On the other hand, it could also mean that people are not looking for a restaurant in your area, so you might need to advertise more.

There is no way of favoring one recommendation over the other. It depends on you needs as a business owner and on your business strategy.