Choosing a Neighborhood in Miami

Hugo Castro

10/07/2020

Introduction

Is well known that Miami is an extremely popular U.S. city to live in, maybe because the tropical weather, the beautiful beaches, or the luxury nightlife in that area.

For those people who are moving into Miami because these “new popularity”, or for those who are already living there but want an upgrade in their life quality, choosing a neighborhood to live is not an easy task, there are many variables to consider into the equation.

The main idea of this project is to find the best neighborhoods to live in Miami, classified them into groups based on all the information we have, so people can choose correctly.

Data

The first data needed to complete this project was the list of all the neighborhoods of the city of Miami, including the coordinates of each one. (https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Miami)

The problem with these data set was that the coordinates where in a different format, there where lost data and not all the columns where useful. So before starting working with it I had to prepare it, creating to new columns for the latitude and the longitude, deleting the rows with lost data and choosing the columns that where needed.

In second place, all the venues from every neighborhood in Miami was extracted from Foursquare, where I was able to choose and search for six of the most desirable venues for a person when is choosing where to live:

1. Police stations (id: 4bf58dd8d48988d12e941735)
2. Fire stations (id: 4bf58dd8d48988d12c941735)
3. Schools (id: 4bf58dd8d48988d13d941735
4. Malls (id: 4bf58dd8d48988d1fd941735)
5. Hospitals (id: 4bf58dd8d48988d196941735)
6. Universities (id: 4bf58dd8d48988d1ae941735)

Process

In order to choose the correct neighborhood for someone, considering their preferences and what is considered important to live in a certain place, and to simplify the process, I want to divide this analysis into four major areas:

1. *Getting and preparing the data*

The first step was to find information about all the neighborhoods in Miami, the exact location of each one, and the number of different venues in the area. The first to where easy to find due to Wikipedia and Google, but the last one was a little bit more difficult.

	Neighborhood	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

To find the number and type (Police stations, Fire stations, Schools, Malls, Hospitals and Universities) of venues in each neighborhood, I used Foursquare API`s and a code to do it automatically.

Once I had all the data I needed in my reach, I start cleaning up the data frame, like dropping missing values, changing formats, creating columns, and deleting the ones that were not needed.

At the end of these process, the data frame ended like these:

	Neighborhood	Latitude	Longitude	Police Sations	Fire Stations	Schools	Malls	Hospitals	Universities
0	Allapattah	25.815	-80.224	0	1	1	0	1	0
1	Arts & Entertainment District	25.799	-80.190	0	1	1	2	1	1
2	Brickell	25.758	-80.193	0	1	0	2	1	5
3	Buena Vista	25.813	-80.192	2	0	1	4	0	0
4	Coconut Grove	25.712	-80.257	0	0	0	0	0	0

2. *Creating clusters*

The idea was to group all the neighborhoods according to the number and type of venues they have. Each group must be different from the other ones but the neighborhoods in each one must be similar, so, in order to accomplish that task, I decided to use clusters, more specifically the k-means method.

After using the “sklearn” library to create the clusters, I created a new column in the data frame to denotate in which of the tree cluster each of the neighborhoods belongs.

	Neighborhood	Latitude	Longitude	Police Sations	Fire Stations	Schools	Malls	Hospitals	Universities	Cluster
0	Allapattah	25.815	-80.224	0	1	1	0	1	0	0
1	Arts & Entertainment District	25.799	-80.190	0	1	1	2	1	1	0
2	Brickell	25.758	-80.193	0	1	0	2	1	5	2
3	Buena Vista	25.813	-80.192	2	0	1	4	0	0	0
4	Coconut Grove	25.712	-80.257	0	0	0	0	0	0	1

At the end, the three cluster where quite different from each other, showing the big differences there are between the neighborhoods in Miami.

Cluster	Police Sations	Fire Stations	Schools	Malls	Hospitals	Universities
0	0.100000	0.000000	0.000000	0.300000	0.000000	0.000000
1	4.333333	2.666667	0.666667	3.666667	1.666667	3.333333
2	1.727273	0.636364	1.272727	1.727273	1.090909	0.181818

3. Ranking

Once the clusters where ready, I decided to create a ranking between the neighborhoods from Miami, so I gave a “weight” to each venue type like this:

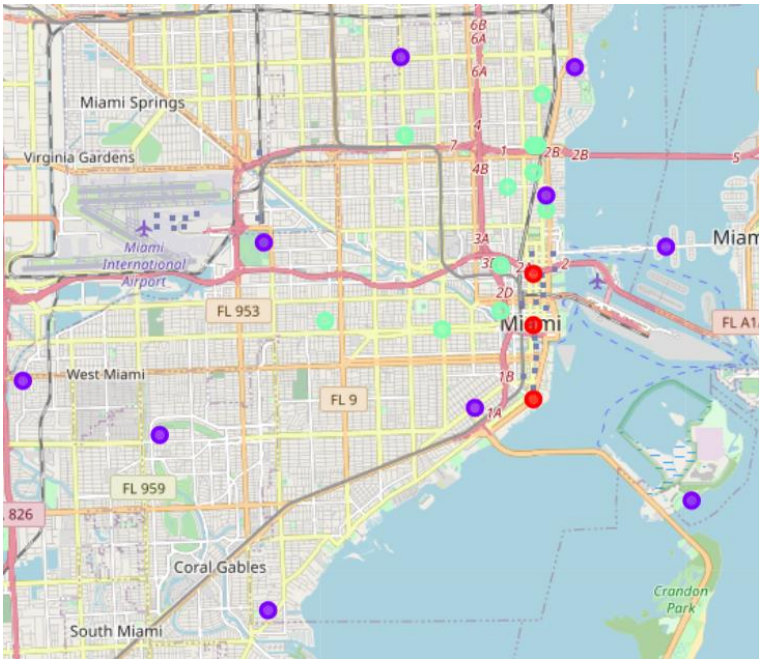
a. Police Station = 0.2
b. Fire Station = 0.2
c. School = 0.15
d. Mall = 0.05
e. Hospital = 0.3
f. University = 0.1

Multiplying these factors by the number of venues of each type and then adding the results, I was able to rank the neighborhoods according to the overall obtained value.

	Neighborhood	Cluster	Ranking
0	Downtown	1	3.85
1	Park West	1	2.60
2	Lummus Park	2	2.25
3	Overtown	2	2.10
4	Wynwood	2	1.25
5	Brickell	1	1.10
6	Little Haiti	2	1.05
7	Midtown	2	0.90
8	Arts & Entertainment District	2	0.85
9	West Flagler	2	0.80

4. *Map*

To finish the processing of information, I decided to create a map in order to represent the locations of the clusters and the neighborhoods to see if there where any relation between the clusters and the geographic position of each cluster.



Analysis

According the studio, Miami can be divided into 3 cluster, each one with different characteristics and an average number of venues.

1. Cluster 0: this cluster is the worst of all three because it has on average none of the listed venues. In the map they represented with a purple dot and they are usually at the borders of the city, meaning that probably those neighborhoods are suburbs with low density population. Additionally, all the neighborhoods in this cluster are on the lowest positions of the rankings, confirming that are the worst places to live in, considering the standard life quality.
2. Cluster 1: in here we have the top three neighborhoods of Miami, with the most police stations fire stations, malls, hospitals, and universities. They are at the top of the ranking and are located at the city center.
3. Cluster 2: these are neighborhood that are between cluster 0 and cluster 1, they have a good average of venues, a lot of schools, even more than cluster 1, they are usually in the middle of the ranking and are located between the suburbs and the city center.

As we can see, there is a clear relation between the number of venues, the cluster, the ranking, and the location of the neighborhood. The more venues or the higher the ranking, there is a bigger probability that the neighborhood is located in the city center.

Conclusion

So if someone wants to move into Miami and has no idea on where to live, he has to start choosing the neighborhoods in the city center and then start moving to the suburbs as he tries to match the price and the budget.

For a better life quality is better to chose one of this five neighborhoods:

1. Downtown
2. Park West
3. Lummus Park
4. Overtown
5. Wynwood