# BATTLE OF THE SISTER ISLANDS

Caroline Jean-Pierre



MADININA VERSUS KARUKERA
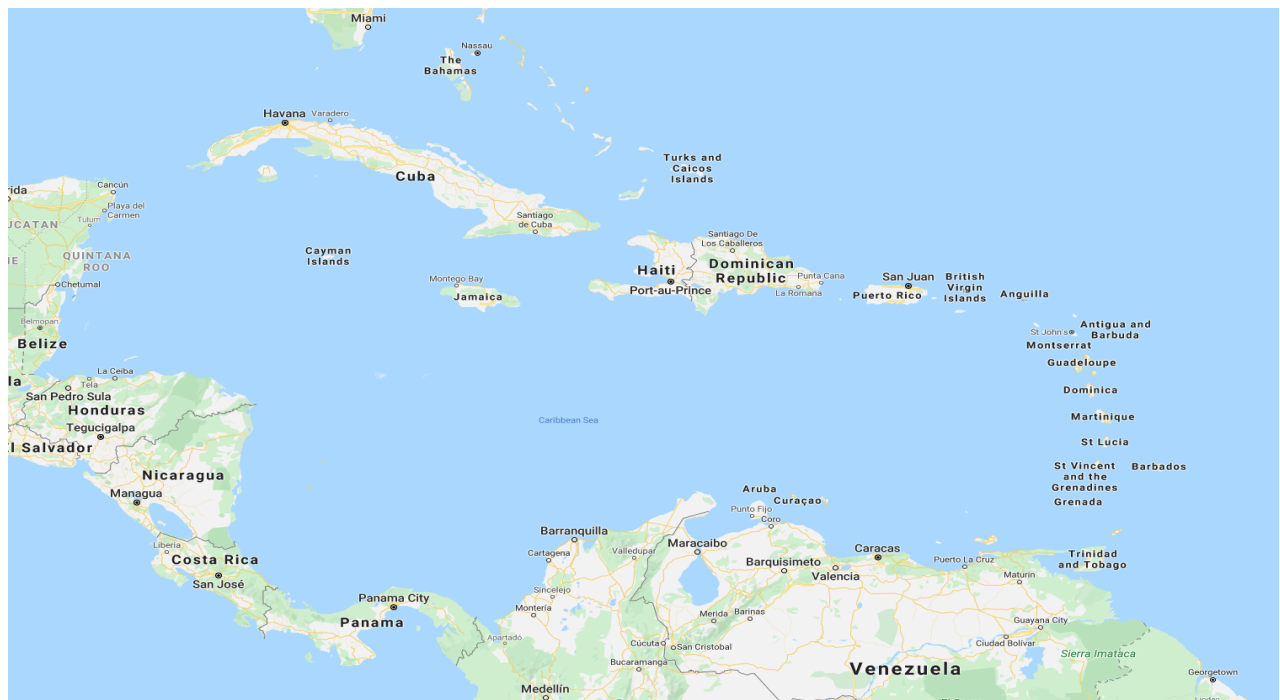
June 2020

# I  INTRODUCTION

## I.1. BACKGROUND

Martinique and Guadeloupe are two French Islands located in the heart of the Caribbean.

In PR Columbus era, the Arawak and the Caribs populated both Islands, the indigenous called Martinique "Madinina" which means "Islands of flowers" and Guadeloupe "Karukera, which means "Islands of beautiful waters".



The French settled in the Islands rapidly in the 17th century and occupied these territories for an almost uninterrupted period until their full integration as French Republic as oversea departments in 1946.

The history, culture and geography of the Martinique and Guadeloupe present many similarities, their inhabitants are used to call them the "sister islands". From the broader French population perspective, the two territories are seen as almost identical and often collectively referred as French west indies.

## I.2. BUSINESS PROBLEM AND INTEREST

Although Madinina and Karukera are often perceived as almost identical nowadays, they do present some dissimilarities.

This distinction that the first inhabitants made was based on the distinctive pattern of the landscape, and it may not be relevant in our modern economies. That being said, from an investment perspective the Sister Islands may not be interchangeable.

It could be interested to analyze what characteristics make the difference and can influence a choice of a business in one of the two territories.

The analysis aims to help potential entrepreneurs in the food industry to make a choice of opening a business in Martinique or Guadeloupe es considering the existing markets (for example the density of and the type of existing restaurants) and the key economical constraint in each of the island.

# II DATA COLLECTION AND PREPROCESSING

## II.1. DATA SOURCES

For market analysis, I used Foursquare to identify existing restaurants and their location in each of the Islands.

For key economical constraints, I used the French government data related to real estate transactions for last 3 years (i.e. 2017, 2018, 2019). These data are available for download in the website https://cadastre.data.gouv.fr/data/etalab-dvf/latest/csv/ and can be download per department (971 for Guadeloupe and 972 for Martinique)

## II.2. DATA PREPROCESSING AND CLEANING

### II.2.1. REAL ESTATE TRANSACTIONS

I imported the real estate transactions for each year and each island and used the pd.read_csv method to create data frames (on per year and island).

I merged the data frames with real estate yearly data to obtain data frames for each island.

I removed duplicated neighborhood in the data frame and checked if the data frames had NaN and null values.

As most of the NaN values were related to missing Latitude and Longitude of neighborhoods, I decided to drop the NaN values from the data frames.

As the source data was labeled in French, I changed the name of the columns from French to English.

### II.2.2. FOURSQUARE DATA

I imported the venues from foursquare with a limit of 100 and a radius of 500

I used the to_csv method to save the foursquare data into csv files for future use and circumvent the daily quotas of foursquare retrievals.

I used one-hot encoding and mean() method to create data frames with the frequency of venues categories per neighborhood.

# III ANALYSIS OF BOROUGHS AND REAL ESTATE MARKET

## III.1. METHODOLOGY

I used the data frames with real estate transactions to:

- Determine the number of boroughs and neighborhoods in each island:
- evaluate the number of real estate transactions in each island over the 3-year period
- evaluate the average price per meter square over the 3-year period
- identify boroughs with the highest number of transactions in both Guadeloupe and Martinique.

To analyze the weight of the most dynamic boroughs, I created bins to classify boroughs by number of real estate transactions. I used a pie chart to measure the weight of each bin in percentage and visualize the level of concentration in the real estate markets of Martinique and Guadeloupe.

I used the folium library to create maps of Martinique and Guadeloupe.

I added markers to indicate the location of the neighborhoods of the 4 most dynamic boroughs of Martinique and Guadeloupe

## III.2. RESULTS

### III.2.1. OVERVIEW OF BOROUGHS AND REAL ESTATE DATA

#### BOROUGHS AND NEIGHBORHOODS

- Guadeloupe has 34 boroughs and 2314 and a total of neighborhoods
- Martinique has 34 boroughs and a total of 2055 neighborhoods

#### NUMBER OF REAL ESTATE TRANSACTIONS

- There have been 8266 transactions in Guadeloupe over the last 3 years
- There have been 6551 transactions in Martinique over the last 3 years

#### PRICE PER METER SQUARE FOR A CONSTRUCTION

- The price for a construction in Guadeloupe is: 2493.91 per meter square
- The price for a construction in Martinique is: 2158.1 per meter square

#### MOST DYNAMIC BOROUGHS

##### GUADELOUPE

There were more than 600 hundred real estate transactions over the last 3 years in 5 boroughs:

- Saint-François
- Baie-Mahault
- Saint Martin
- Le Gosier
- Les Abymes

Guadeloupe - Number of real estate transaction per boroughs
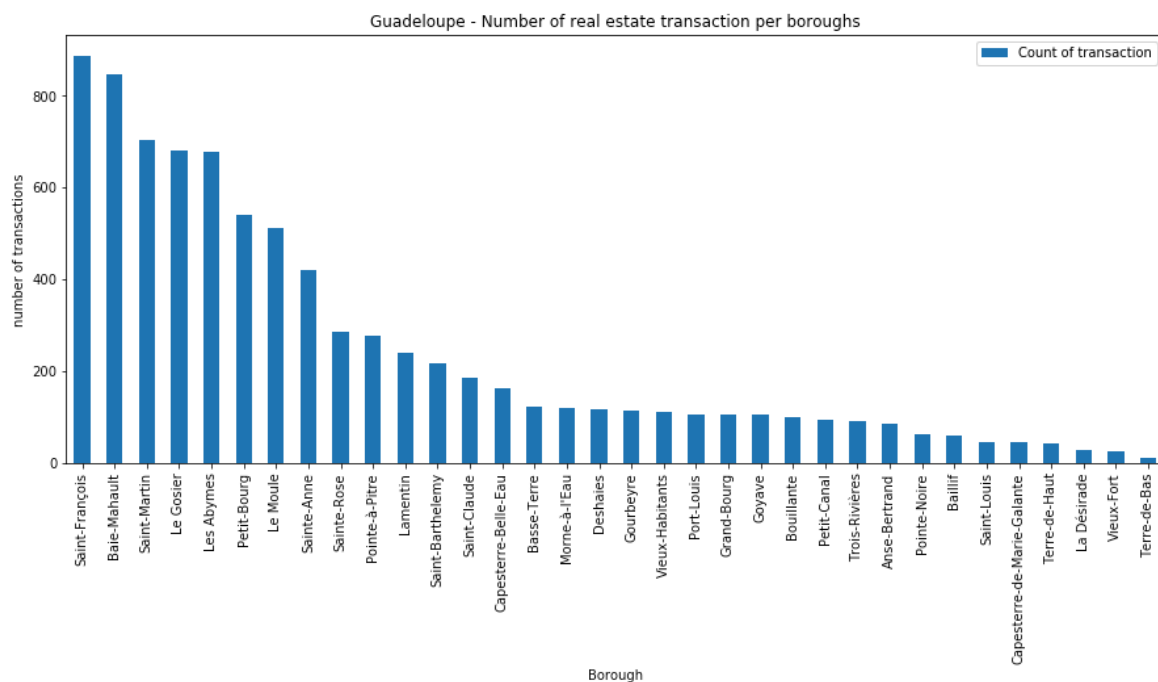
## MARTINIQUE

There were more than 500 hundred real estate transactions over the last 3 years in only 4 boroughs:

- – Fort de France
- – Le Lamentin
- – Schoelcher
- – Les Trois îlets

Fort de France is by far the most dynamic borough in Martinique with 1500 transactions over the period 2017-2019
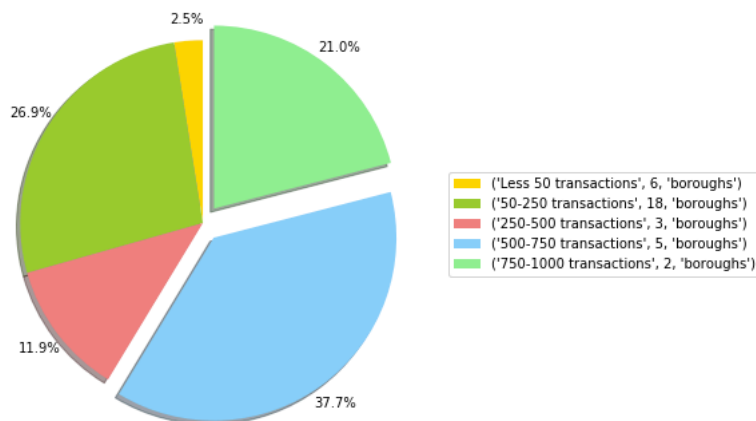


Martinique - Number of real estate transactions per boroughs
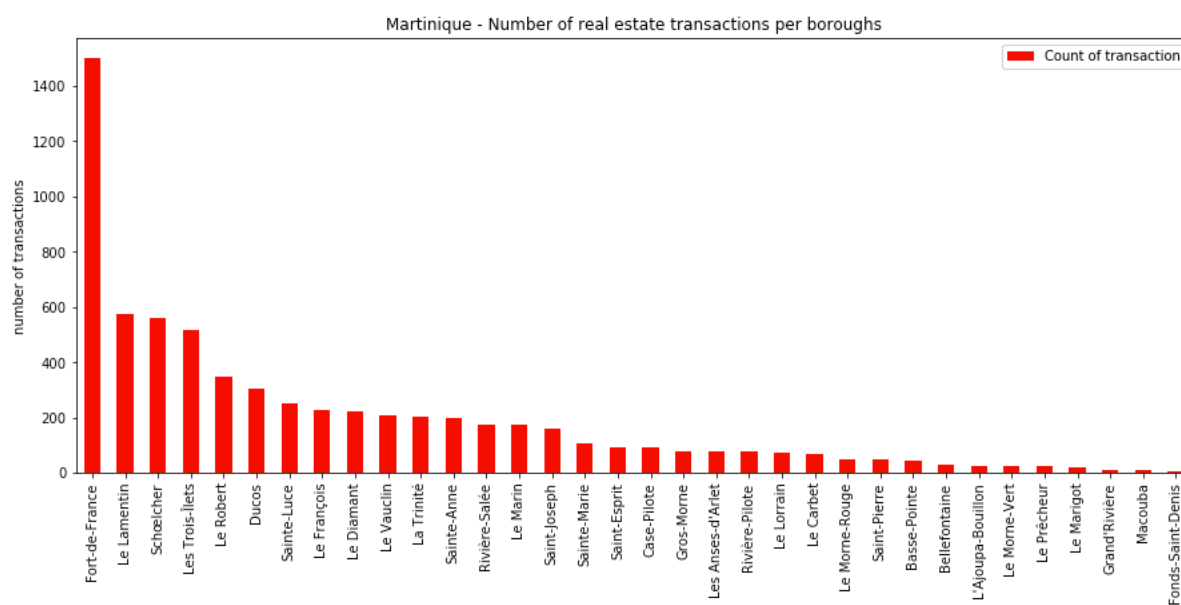
## GUADELOUPE

The real estate market is highly concentrated. The 2 most dynamic boroughs capture 21% of the transactions and the next 5 capture 37.7%:

In a nutshell, 20% of the boroughs gather 58.7 % of the market.

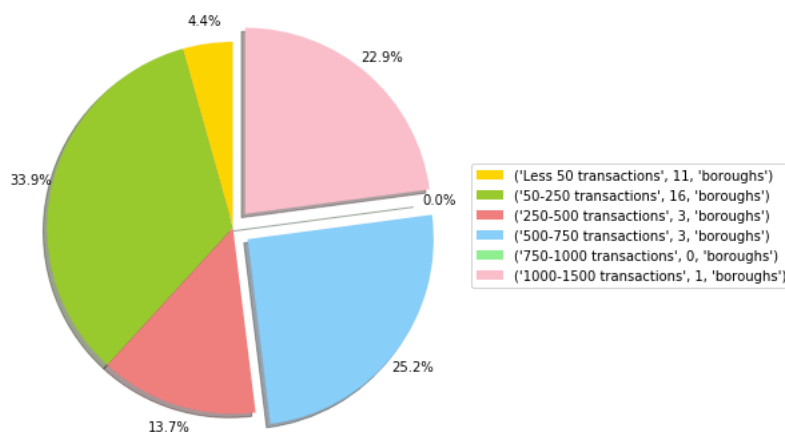Concentration of real estate market in Guadeloupe

2.5%

21.0%

26.9%

('Less 50 transactions', 6, 'boroughs')
('50-250 transactions', 18, 'boroughs')
('250-500 transactions', 3, 'boroughs')
('500-750 transactions', 5, 'boroughs')
('750-1000 transactions', 2, 'boroughs')

11.9%

37.7%

## MARTINIQUE

The real estate market is highly concentrated. The most dynamic borough captures 25% of the transactions and the next 6 capture 36.6%:

In a nutshell, 20% of the boroughs gather 61.8 % of the market.

However, we notice a big gap between Fort de France with 1500 transactions and the rest of the boroughs. The second, third and fourth most dynamic boroughs only gathered between 500 and 750 transactions that less than a half of Fort-de France real estate transactions.

Concentration of real estate marketin Martinique

4.4%

22.9%

33.9%

0.0%

('Less 50 transactions', 11, 'boroughs')
('50-250 transactions', 16, 'boroughs')
('250-500 transactions', 3, 'boroughs')
('500-750 transactions', 3, 'boroughs')
('750-1000 transactions', 0, 'boroughs')
('1000-1500 transactions', 1, 'boroughs')

25.2%

13.7%

## III.2.2.  VISUALISATION OF THE NEIGHBORHOOD OF THE 4 MOST DYNAMIC BOROUGHS OF EACH ISLAND

The data for Guadeloupe included Saint Martin, which is administratively attached to Guadeloupe, but it is actually another Island with different population and economy.

A such we won't take into account Saint Martin data in the rest of our analysis as we want to focus on Guadeloupe and Martinique Islands.
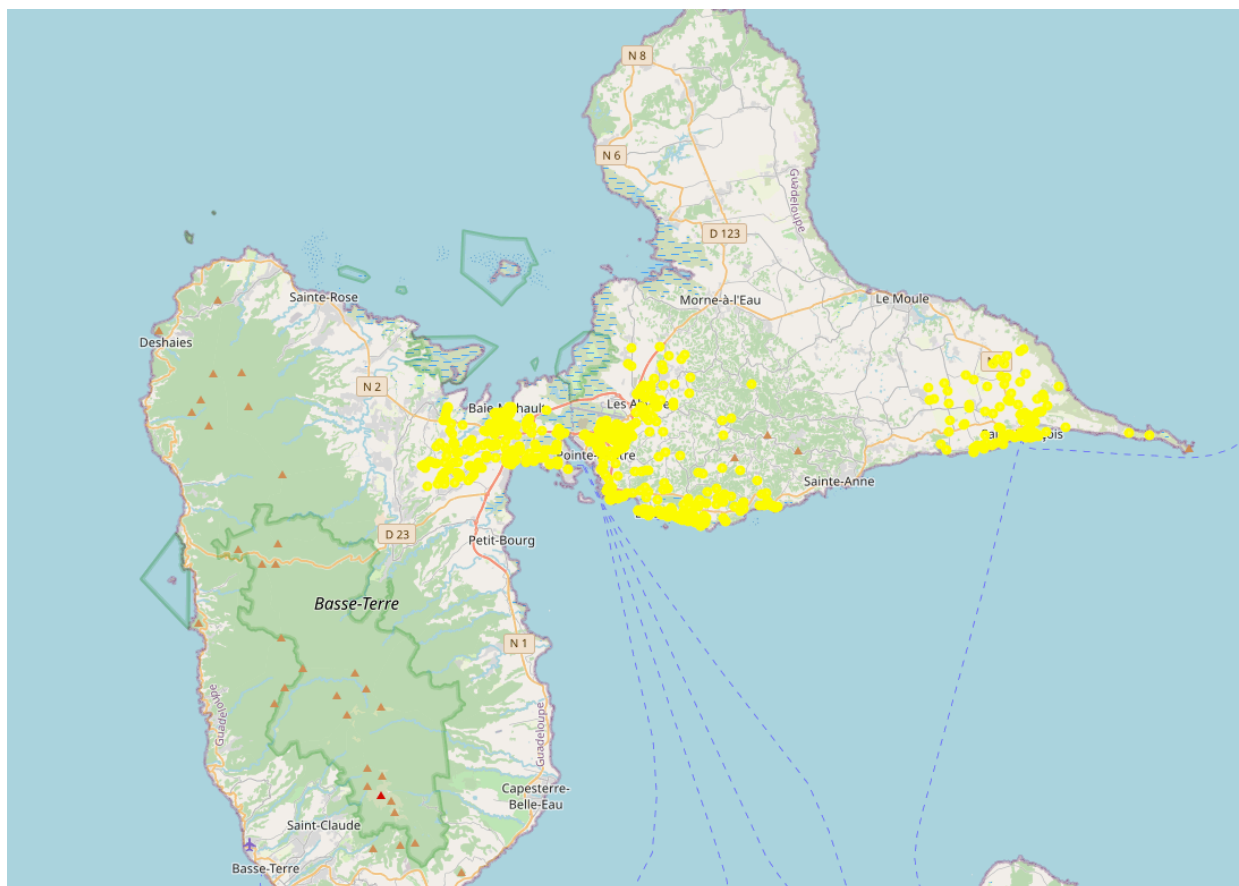
Based on the previous analysis we identified the four most dynamic boroughs of Guadeloupe as:
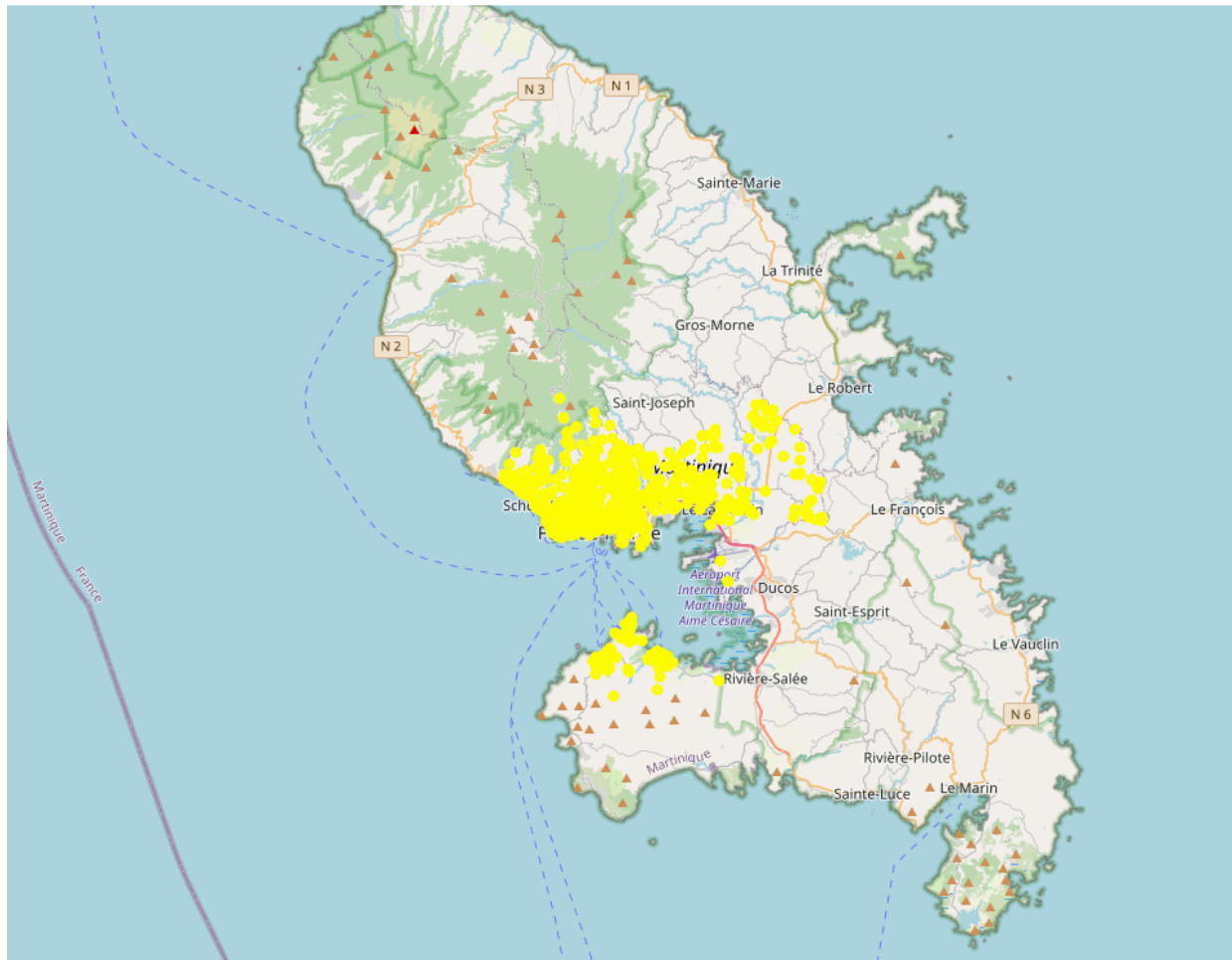
- Saint-François
- Baie Mahault
- Le Gosier
- Les Abymes

And the four most dynamic boroughs of Martinique are:

- Fort de France
- Le Lamentin
- Schoelcher
- Les Trois îlets

The map of the neighborhoods of the 4 most dynamic boroughs of Guadeloupe is as follows:



The map of the neighborhoods of the 4 most dynamic boroughs of Martinique is as follows:

# IV EXPLORE NEIGHBORHOODS OF THE MOST DYNAMIC BOROUGHS OF MARTINIQUE AND GUADELOUPE

## IV.1. METHODOLOGY

### IV.1.1. CREATE DATAFRAMES WITH GEOLOCATION OF THE NEIGHBORHOODS

I first created data frames with the borough, neighborhood, postal code, latitude and longitude for each island.

As almost half of the real estate transactions of each island happened in the top four dynamic boroughs, I decided to slice the data frames previously created to include only the 4 most dynamic boroughs (from a real estate market perspective) of each island.

I then obtained two data frames:

- Martinique top 4 boroughs with 857 neighborhoods
- Guadeloupe top 4 boroughs with 613 neighborhoods

Note: Guadeloupe initial data contained the borough "Saint Martin". Saint Martin is administratively attached to Guadeloupe, but it is actually another Island, so we won't take into account Saint Martin data in the rest of our analysis.

I used the geolocation of the neighborhoods of Martinique and Guadeloupe top 4 boroughs to obtain the venues of each neighborhoods for Martinique and Guadeloupe from foursquare and store them in a data frame.

I save the data frames in csv files for future use.

## IV.2.  RESULTS

Based on the real estate data, Martinique top 4 boroughs has 857 neighborhoods for which we have the latitude and longitude information and Guadeloupe top 4 boroughs has 613 neighborhoods for which we have the latitude and longitude information.

However, from foursquare data we imported:

- 7073 venues for 796 neighborhoods of Martinique, so 61 neighborhoods with real-estate transactions don't have any venue in foursquare
- 2790 venues for 520 neighborhoods of Guadeloupe, so 93 neighborhoods with real-estate transactions don't have any venue in foursquare

### IV.2.1.  IDENTIFY MOST COMMON VENUE CATEGORIES

#### GUADELOUPE

There are 148 venue categories in Guadeloupe data.

Almost 20% of the venues are Fast Food restaurants pizza places and Caribbean restaurants.
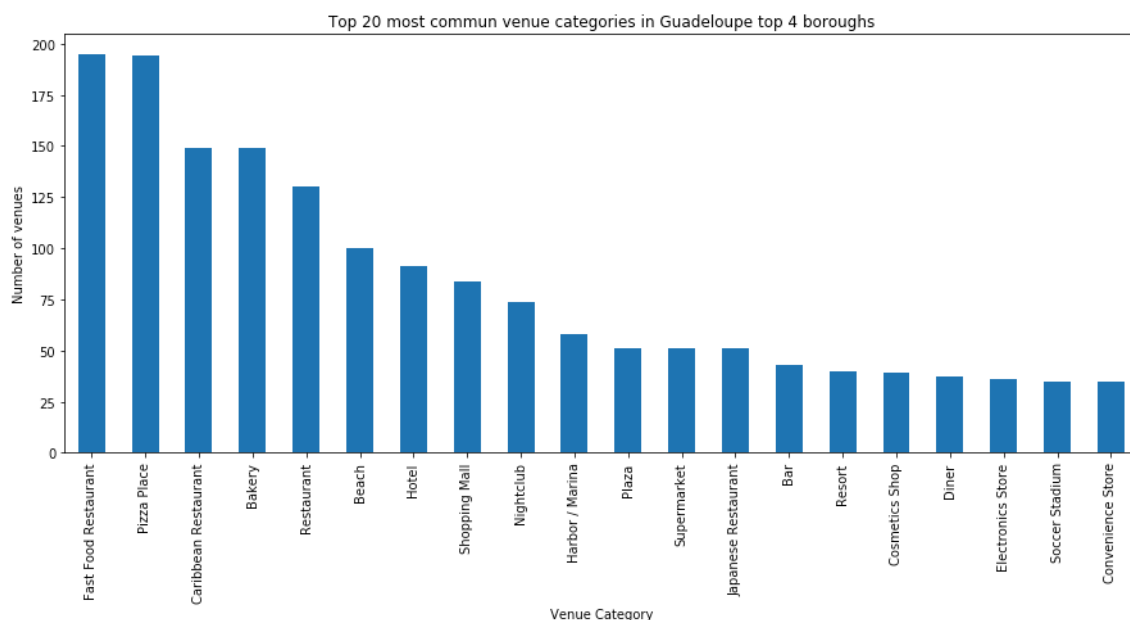


#### MARTINIQUE

There are 181 venue categories in Martinique data.

Like Guadeloupe, Pizza place, Fast Food and Caribbean restaurants are amongst the most common venues, however their relative weight is lower as these categories account for less than 13% of the venues.

There seems to be more diversity in Martinique venues as we note pharmacies as the 7th most common venues, spa and ice-cream shops are in the top 20 most common venues in Martinique.



Top 20 most common venue categories in Martinique top 4 boroughs
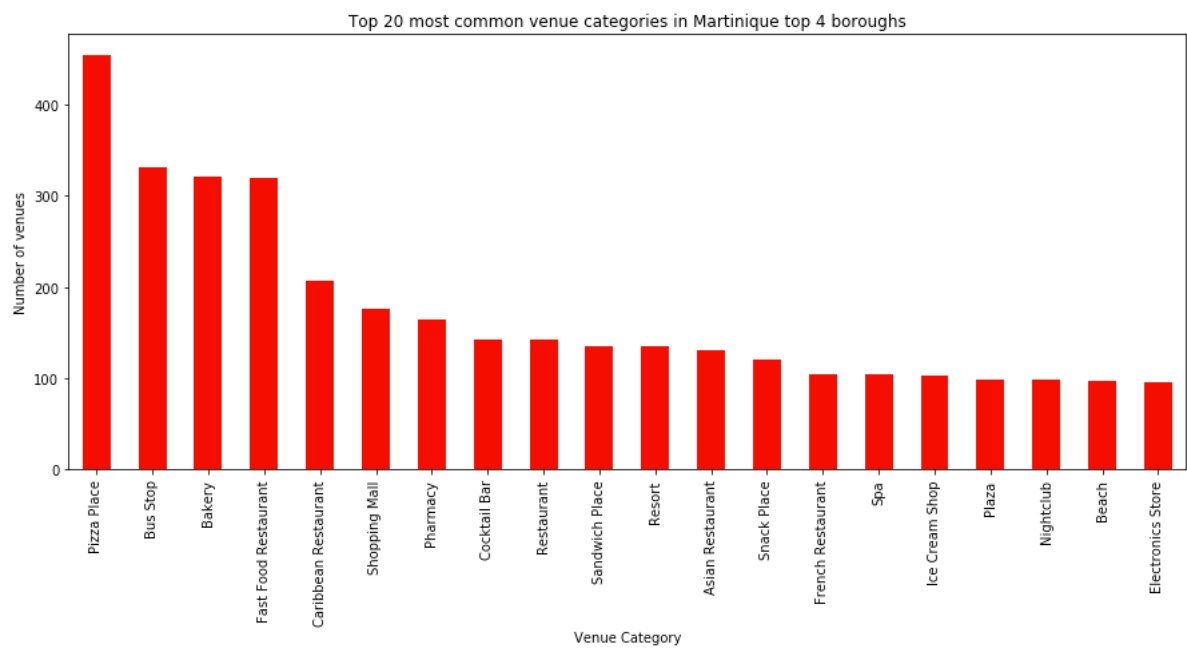
## IV.2.2. IDENTIFY NEIGHBORHOODS WITH HIGHEST CONCENTRATION OF VENUES

The area with the highest concentration of venues in Guadeloupe only count between 15 and 25 venues.



Top ten Guadeloupe Neighborhood per number of venues
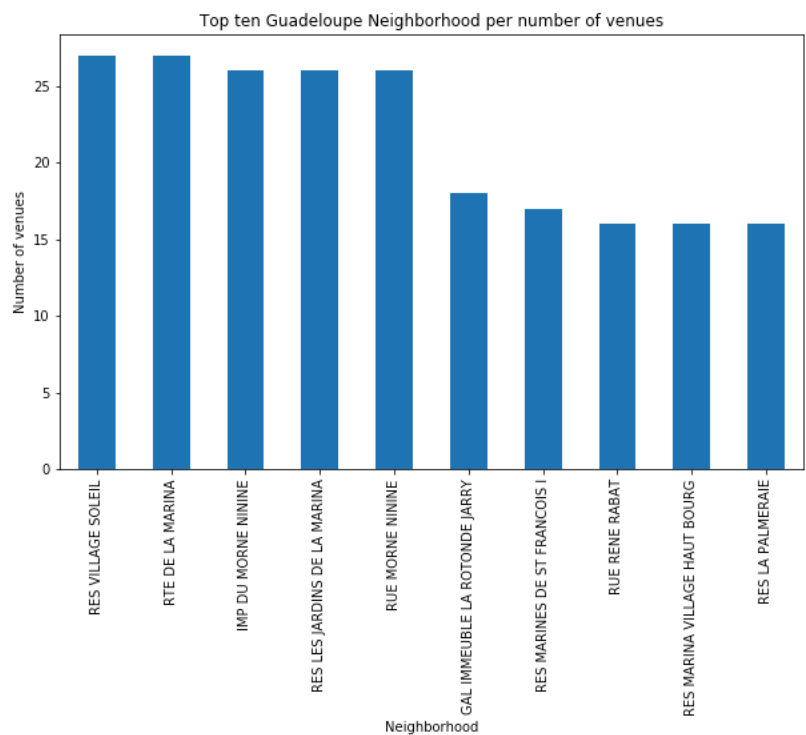
The area with the highest concentration of venues in Martinique count over 50 venues.

Top ten Martinique Neighborhoods per number of venues
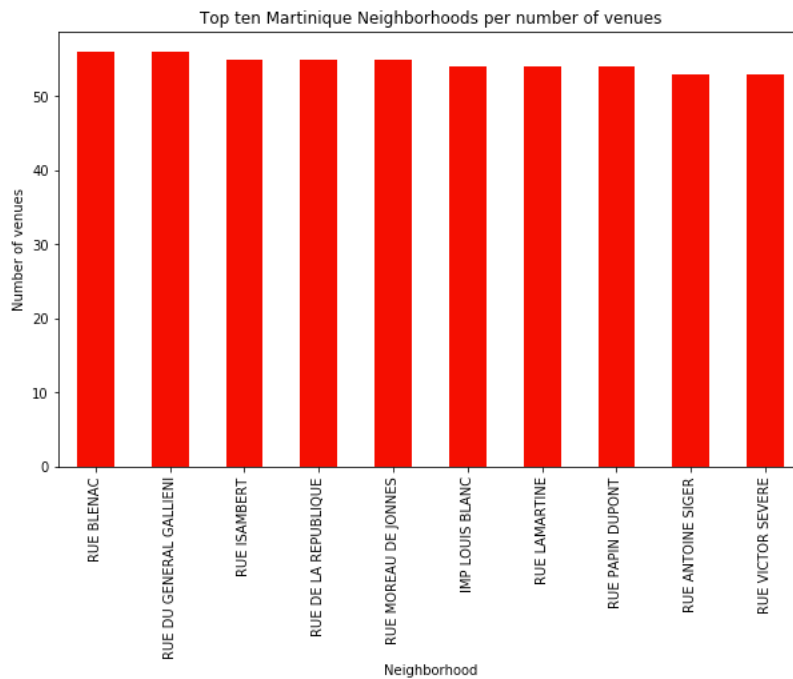
# V CLUSTER AND SEGMENT NEIGHBORHOODS

## V.1. METHODOLOGY

### V.1.1. COMBINED DATA FOR MARTINIQUE AND GUADELOUPE

I combined the Guadeloupe and Martinique data frames to have a unique data frame of Neighborhoods and venues.

I identified that some neighborhoods in Martinique and Guadeloupe had the same name so I modified the name of these neighborhoods to ensure they would be distinct in the analysis.

### V.1.2. FREQUEQUENCY OF VENUE CATEGORY PER NEIGHBORHOOD

In order to identify similarities of neighborhood, I used one-hot encoding to obtain a data frame identifying venue categories of each venue (i.e. if a venue is in a category, the value is 1 and if a venue is not in a category the value is 0.

I then grouped the data per neighborhood and applied the mean method to obtain the frequency of a venue category in each Neighborhood.

### V.1.3. RUN K-MEAN CLUSTERING

I run K-mean clustering to segment neighborhoods based on the frequency of venue categories in the neighborhood.

I set the number of clusters to 5 clusters

### V.1.4. DATA ANALYSIS AND VISUALISATION

In order to identify the specificities of each cluster, I created a data frame with the top 10 most common venues in each cluster.

I then sliced the data frame per cluster and visualized the first, second, third and fourth most common venue categories in the cluster.

## V.2.    RESULTS

The combined foursquare data for Martinique and Guadeloupe resulted in 9863 venues spread in 222 venue categories and 1316 neighborhoods

### V.2.1.   IDENTIFY SIMILARITIES IN THE NEIGHBORHOODS OF THE TWO ISLANDS

As a result of the unsupervised clustering, 5 clusters of neighborhoods have been created.
The number of neighborhoods per clusters is as follows:

| Cluster Labels | Neighborhood |
|---|---|
| 0 | 148 |
| 1 | 706 |
| 2 | 260 |
| 3 | 65 |
| 4 | 137 |

The composition of cluster per neighborhood is as follows:

| Cluster Labels | % of Martinique | % of Guadeloupe |
|---|---|---|
| 0 | 39.19 | 60.81 |
| 1 | 59.77 | 40.23 |
| 2 | 63.85 | 36.15 |
| 3 | 98.46 | 1.54 |
| 4 | 62.77 | 37.23 |

Excepted the cluster 3 which is almost exclusively composed of neighborhoods of Martinique, all clusters are mostly balanced with a ration of 60-40 % of neighborhood per island.

### V.2.2.   VISUALIZE CLUSTERS

The cluster 0 in red with 148 neighborhoods is more present in Guadeloupe and is mostly located in the coastal areas of the islands

The cluster 1 in purple is by far the biggest cluster with 706 neighborhoods that is 54% of the neighborhoods.

The cluster 2 in blue is the second biggest cluster with 260 neighborhoods is more present in Martinique and mostly located toward the center of the islands

The cluster 3 in light green is almost exclusively located in Martinique and only count 65 neighborhoods. It is located mostly in the inner land, further from the coastal areas.

The cluster 4 in orange with 137 neighborhoods is more present in Martinique and mostly located in peripheries of cluster 1
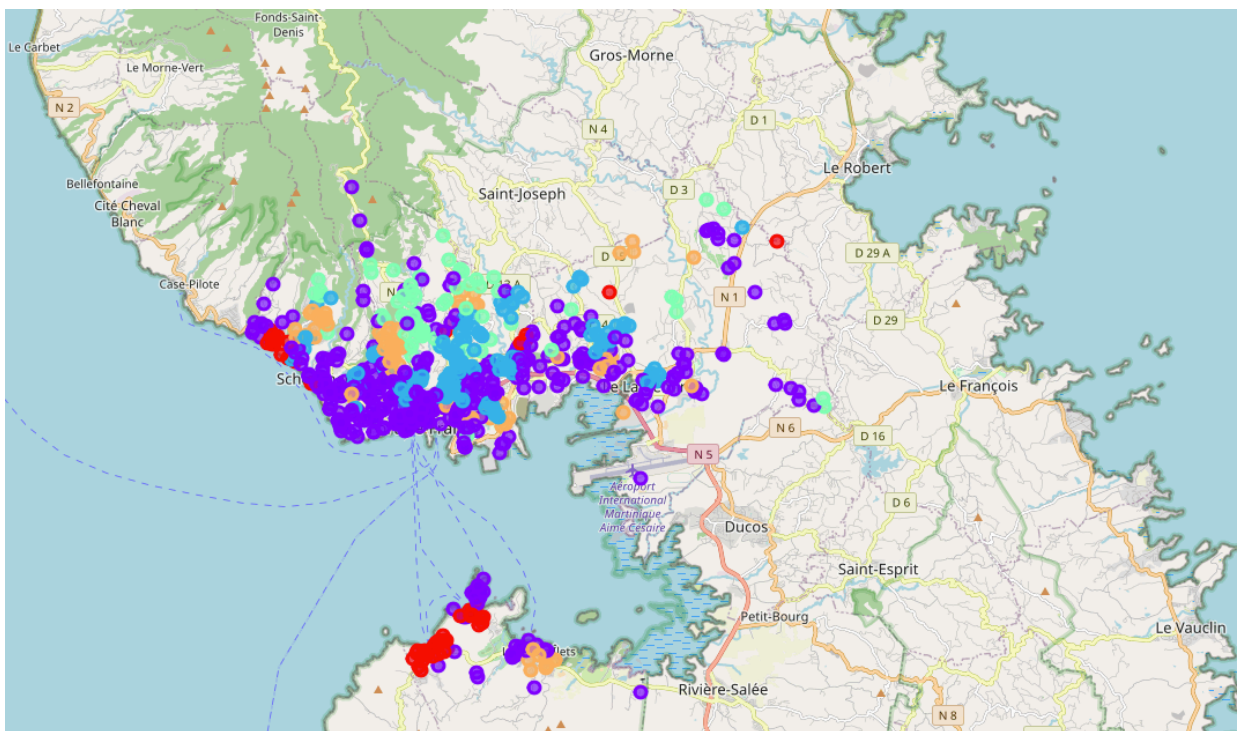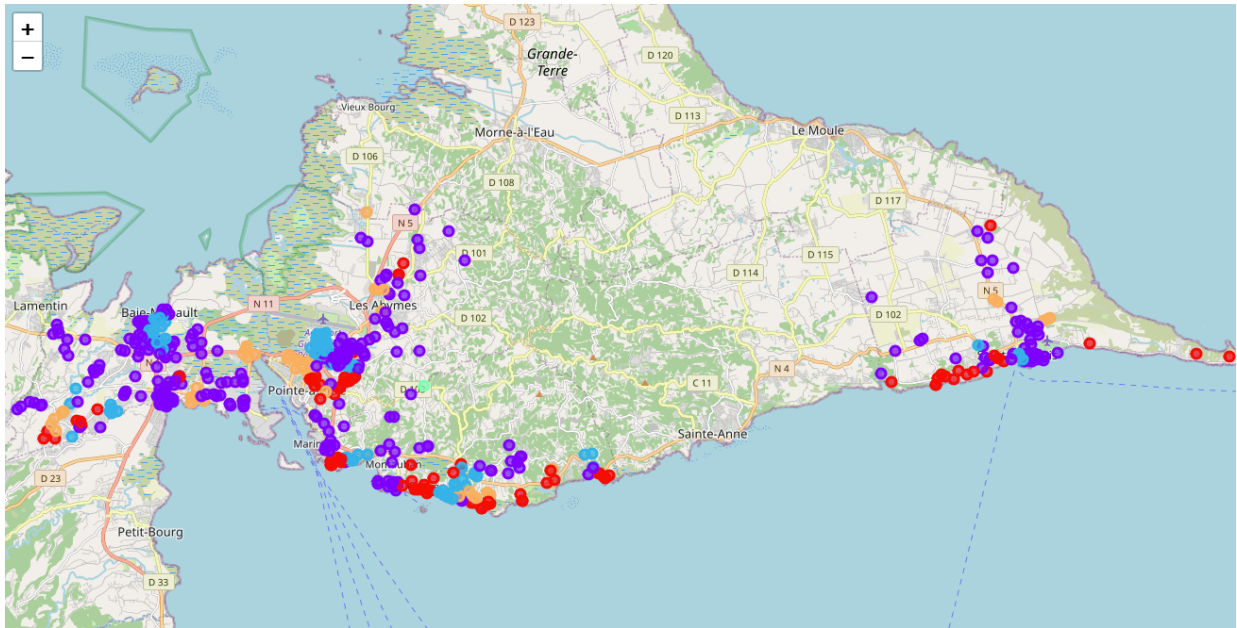
## V.2.3.   ANALYSE EACH CLUSTER BASED ON THEIR 4 MOST COMMON VENUES

### CLUSTER 0 - ANALYSIS

Cluster 0 neighborhoods seem to be characterized by coastal and leisure activities:

The neighborhoods of these clusters are surrounded by Caribbean restaurants, beaches and Yoga studios

Cluster 0, first most common venues

Cluster 0, second most common venues

Cluster 0, third most common venues

Cluster 0, fourth most common
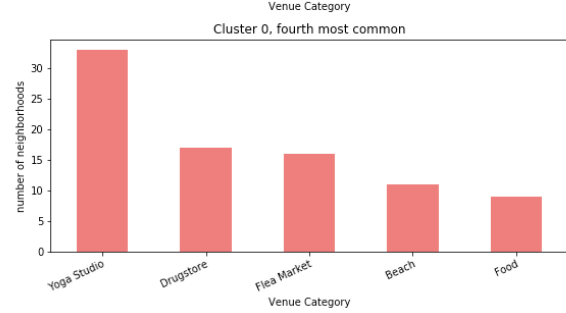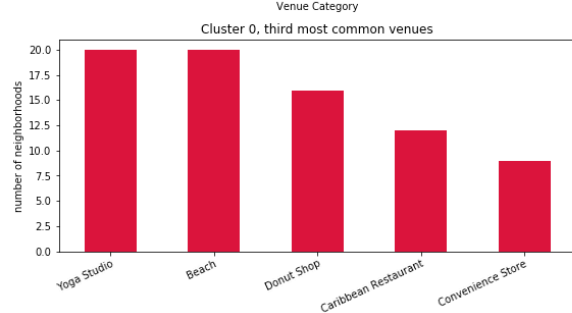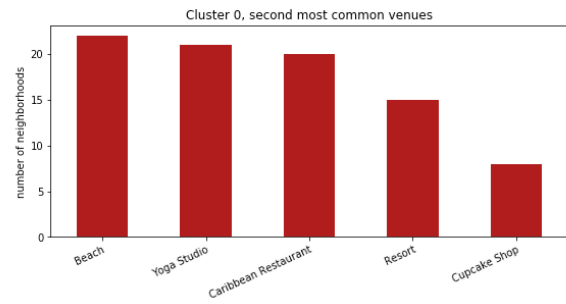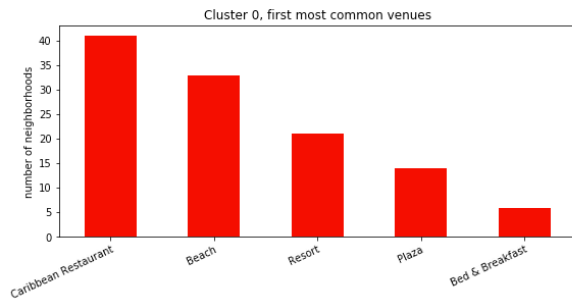
## CLUSTER 1 – ANALYSIS



Cluster 1, first most common venues

Cluster 1, second most common venues

Cluster 1, third most common venues

Cluster 1, fourth most common venues
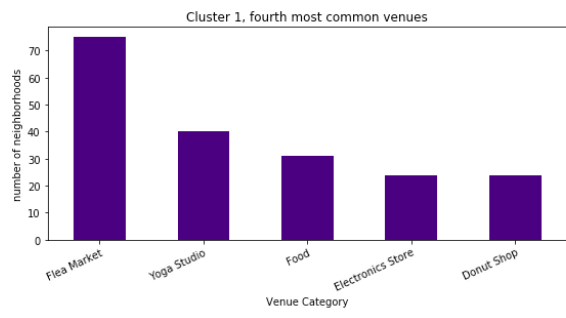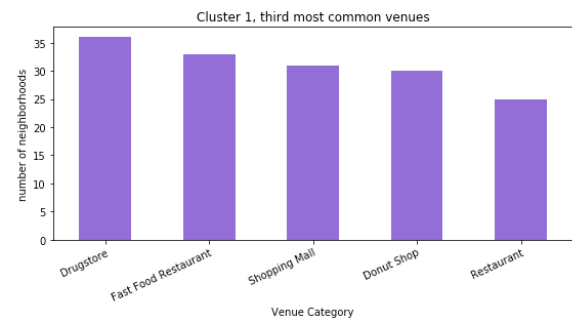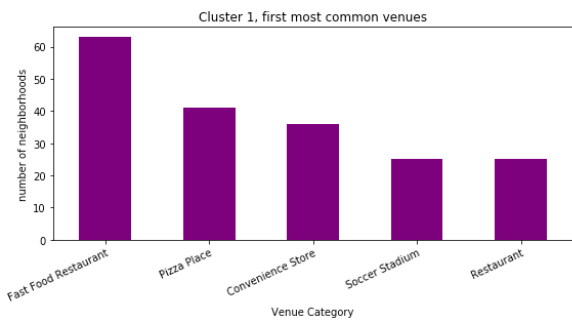
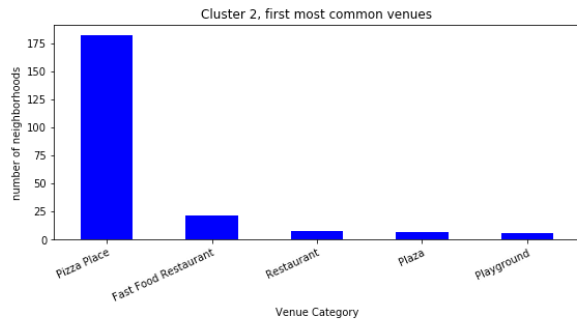Cluster 1 neighborhoods seems to be characterized by city and active lifestyle activities:

The neighborhoods of these clusters are surrounded mostly by Fast Food restaurants, Pizza places, Yoga studios, markets and shopping venues.

## CLUSTER 2 – ANALYSIS

Cluster 2 neighborhoods seems to be characterized by family or residential activities:

The neighborhoods of these clusters are invaded by Pizza place :175 out of 260 neighborhoods of the cluster has a pizza place as most common venue.
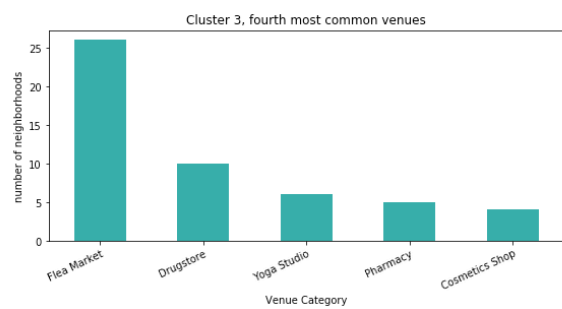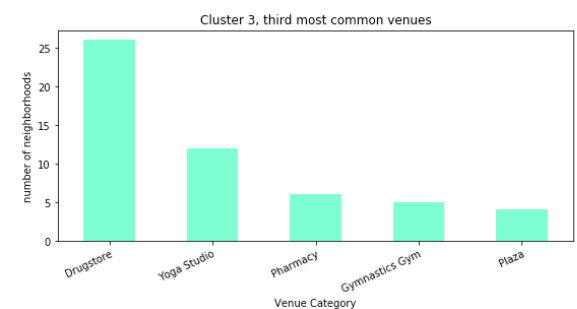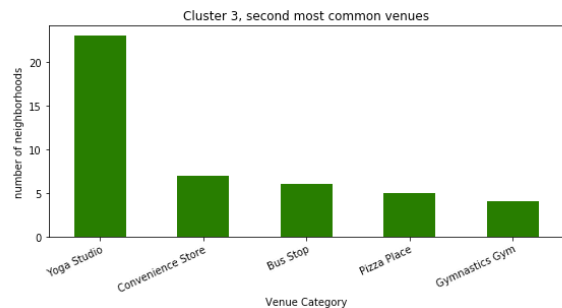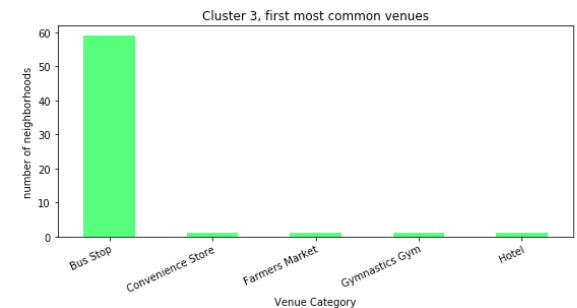
Pharmacies and Bakery are also very common in the neighborhoods of the second cluster.
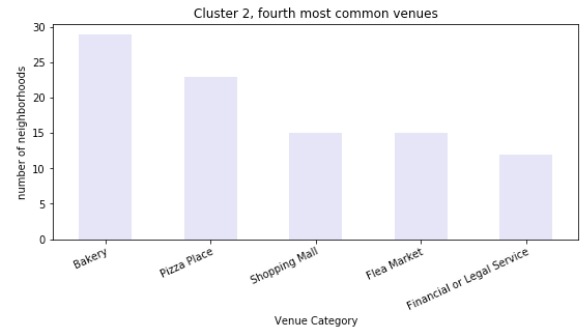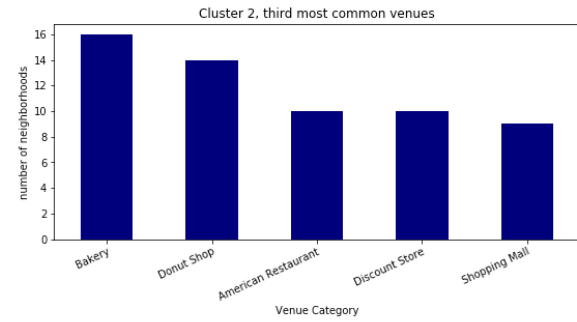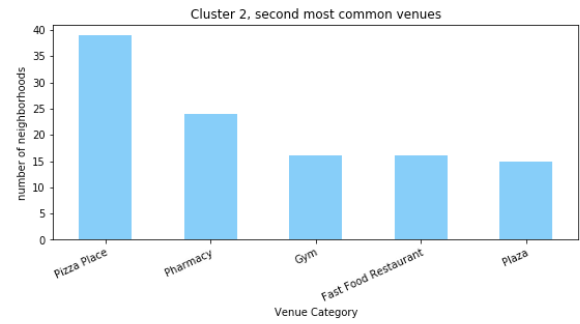
## CLUSTER 3 – ANALYSIS

Cluster 3 neighborhoods seems to be characterized by rural and very limited business activities:
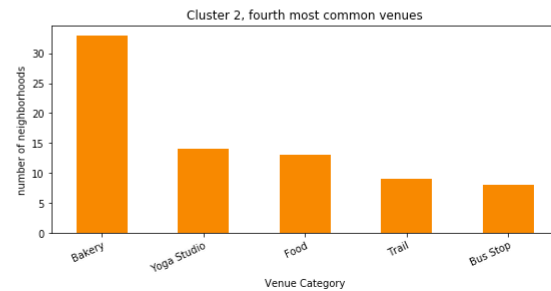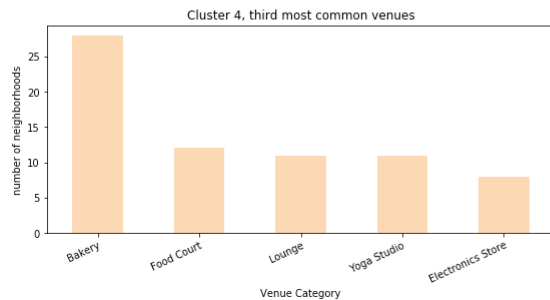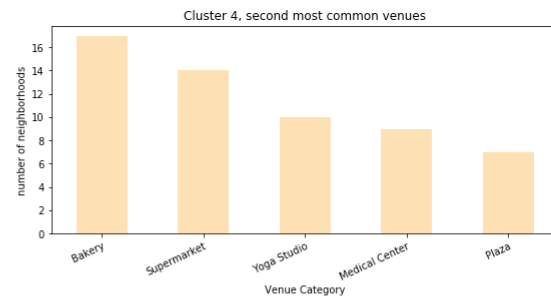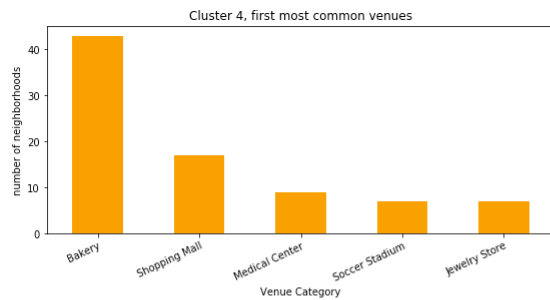
The neighborhoods of these clusters are characterized by almost no restaurants, business activities are limited to Yoga Studio, drugstores



## CLUSTER 4 – ANALYSIS

Cluster 4 neighborhoods seems to be characterized by convenience stores, supermarkets and shopping activities:

The neighborhoods of these clusters are surrounded by bakeries, supermarkets, medical centers.

Cluster 4, first most common venues — Cluster 4, second most common venues — Cluster 4, third most common venues — Cluster 2, fourth most common venues

# VI CONCLUSION

Our analysis identified both similarities and dissimilarities in the sister islands.

Over the period 2017-2019 Guadeloupe real estate market has been more dynamic and the housing sale price is more expensive than Martinique's.

Both Islands have a concentrated real estate market with a fifth of the boroughs gathering half of the transactions. However, Martinique market shows a huge gap between the capital Fort-de-France and the rest of the territory. In Guadeloupe, there is more balance and continuity between the different boroughs.

From a geography perspective, the most dynamic boroughs of Martinique are highly concentrated in the center of the island. In Guadeloupe, the most dynamic boroughs are spread in the center and the north est and cover a bigger part of the territory.

Despite the less expensive housing sale price, business implantations in Martinique are much higher and diverse than in Guadeloupe. In average the four most dynamic boroughs of Martinique count 60% more venues per neighborhood than Guadeloupe's four most dynamic boroughs.

Both islands food industries are largely dominated by Fast Food restaurants and pizza places, there seems to be room for opportunities in this area of business in both islands.

In the health sector, it is interesting to note that pharmacies are amongst the top 10 most common venues in Martinique whereas there are not even included in the top 20 most common venues of Guadeloupe.

The neighborhoods segmentation reveals that both Martinique and Guadeloupe present a very large number of neighborhoods mostly located in the center of the island in urban areas that are surrounded mainly by Fast Food restaurants, Pizza places, Yoga studios, and shopping venues.

Close to these urban areas are more residential neighborhood characterized by very high concentration of pizza places followed by bakeries and pharmacies.

The costal neighborhoods of the boroughs of Martinique and Guadeloupe are mainly caracterized by Caribbean restaurants and Yoga studios close to beaches although there are more neighborhoods of this type in Guadeloupe boroughs.

Finally, both islands also have neighborhoods dedicated to retail and shopping activities.

To conclude, both islands presents business opportunities. However, competition in Guadeloupe seem to be less fierce as the economic implantation is much lower than Martinique. That being said, the fixed cost of installation that is related to real-estate is higher in Guadeloupe.