Capstone Project - The Battle of Neighborhoods in Miami

Introduction:

This is the final assingment of Cousera Applied Data Science Capstones (part of Courseras Applied Data Science Specialization) course. This project is about learning how to use the knowleage given to my own project.

I am going to do mine Peer-graded Assingment of Miami, because I am interested of visiting there in the feture - more about my choice in the next chapter on this notebook.

This is my first time ever doing anything like this, actualla first time ever dealing with dataset, so I am going to keep this project so simply as possible.

Description of the problem and discussion of the backround:

I choosed to do this project of Miami neighborhoods and venues. This project is based on an idea, that trending travel agency wants to start selling "do it one your own" vagations and they hired me to analysize what they shuold recommentate to customers or book for them. Of course, best way on ideal way to sell "almous like you book it your self" and "live like locals" holidays, is to solve how does the locals spend their time.

Problem wich I´m trying to solve in this project is: If customer is about to travel or thingking of travelling to Miami - wich neighborhoods are the best to stay? People and customer want different things, so I will make on analysis from different perpectives: for beach holiday for cultural holiday for shopping holiday and for both beach and something else to do easily also.

I´m trying to solve this problem in this project by first gatering Miami Neighborhoods and gettin Forsquare data on venues and then analysing these datasets by gruoping them into venue gategories and finally clustering them to these fuor gategories I mentioned before: Beach Cultural Shopping Both Beach and something else to do also

Perhaps I will find some other gategorie on the way to, but this is the starting point of this project.

Description of the data and how it will be used to solve the problem:

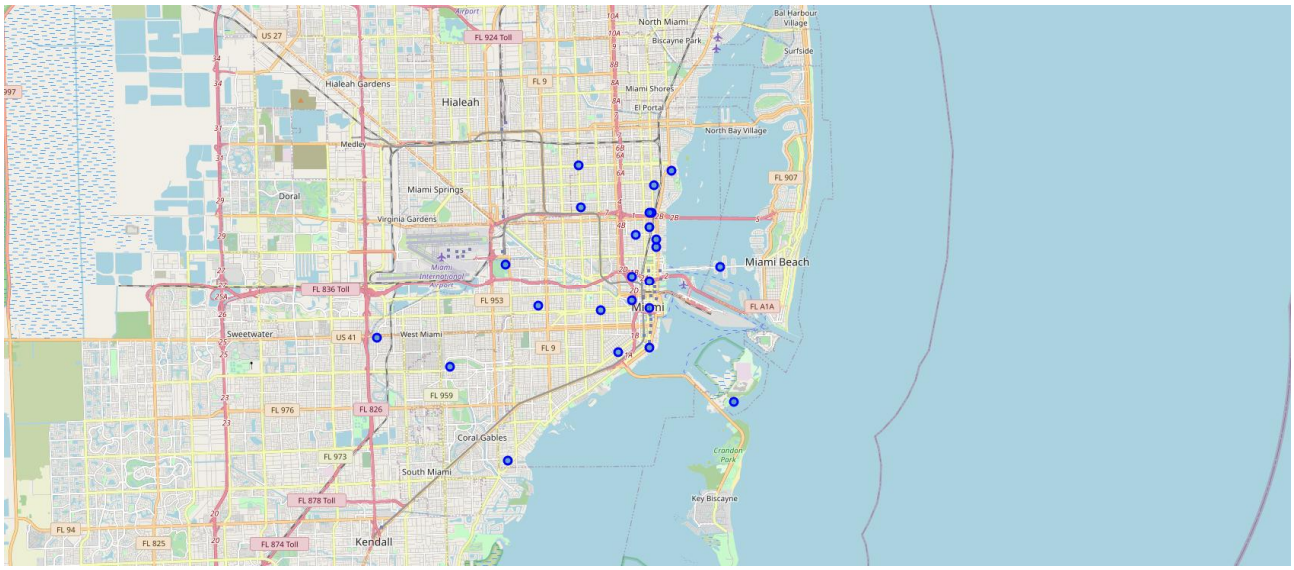
In my project I will use data from Wikipedia page https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Miami and Forsquare data off Miamis venues.

In Wikipedia data I´ll found data of neigbourhoods and their geographical goordinates. Data also includes data of: Demomyn, Population 2020, Populalation/km2 and Sub-neighborhoods. I´m going to clean those out of my data, cause I´m going to keep this project simple. If this project wuold be about moving to Miami or bying a condo from there I think I wuold also use Population/km2 data and look closely to Sub-neighborhoods, but in this point the knowleage of those are not meaningfull.

More meaning to this project will come from Forsquare data of venues in Miami neighborhoods. I´m goint to use this data to locate places where local people often go. I believe that the best and cheepest way to get to know the city is to trust people who live there - we do they often go and what do they do? Forsquare data will tell me this and that is why I choose to use it in this project. Forsquare data is also really good for my purpose - lets not forget that I´m doing this project for my employer who wants to know what kind of travelling plans to make for their new trending market.

Methodology section

Data cleaning



Abowe you can see a map of Miami´s neighborhoods, each marked by blue cirkle. As we can see, neighbourhoods I´m handeling with are in centre of Miami and distances are not so far away from each other. Later on this means, that when I pick the venues nearby neighborhoods I will use 2 km radar to each Neighborhoods centre to locate the venues nearby.

Importin Forsquare data

Importin Forsquare data of Allapattah

To work with Forsquare data I will at first have to give my Forsquare client ID and SECRET and tell the version I want to use. After that I can go on adding the Forsquare data to my data of Miami

I will at first try does the Forsquare data work with my data with just one neighborhood. I pick up the first one which is Allapattah.

Because the information in Forsquare data is in the item keys i will make the get gategory type function and then clean the json and structure it into a pandas dataframe.

There are only 12 venues in Allapattah 1 km range from Allapattahs centre. It´s not much and perhaps by searching a little bit longer wuold give us little more result, but at this point is only relevant that function works.

Importing Forsquare data of Miamis neighborhoods

At first i will make the function work to many neighbourhoods at the same time. Like a told before, I´m going to use 3 km distance from neighbourhoods centre to locate the venues, cause areas are nearby, but still close to each other. And my goal is to think neighbourhoods for travelling - 3 km is somethong what almoust every one can walk. If I wuold run this with 2 km distance, I didn´t find any beaches.

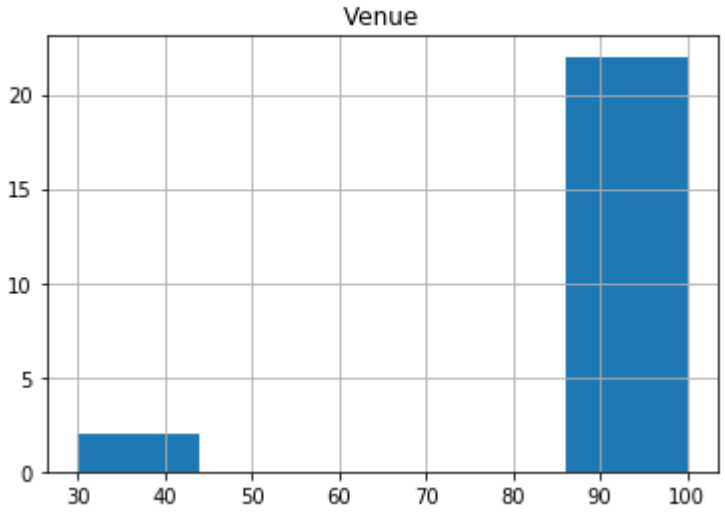
And then I will use function abowe to make a new dataframe of Miami´s neighbourhoods.

```
Allapattah
Arts & Entertainment District
Brickell
Buena Vista
Coconut Grove
Coral Way
Design District
Downtown
Edgewater
Flagami
Grapeland Heights
Liberty City
Little Haiti
Little Havana
Lummus Park
Midtown
Overtown
Park West
The Roads
Upper Eastside
Venetian Islands
Virginia Key
West Flagler
Wynwood
```

Cleaning the importet data of Miami´s venues.

After a little while of working with data i realised that venue gatogories in the dataset include in one category also Neighbourhood. It´s not on venue, so I will clean it up and then count the venues again. Now data looks correctly to use later on.

```
There are 202 uniques categories.
```

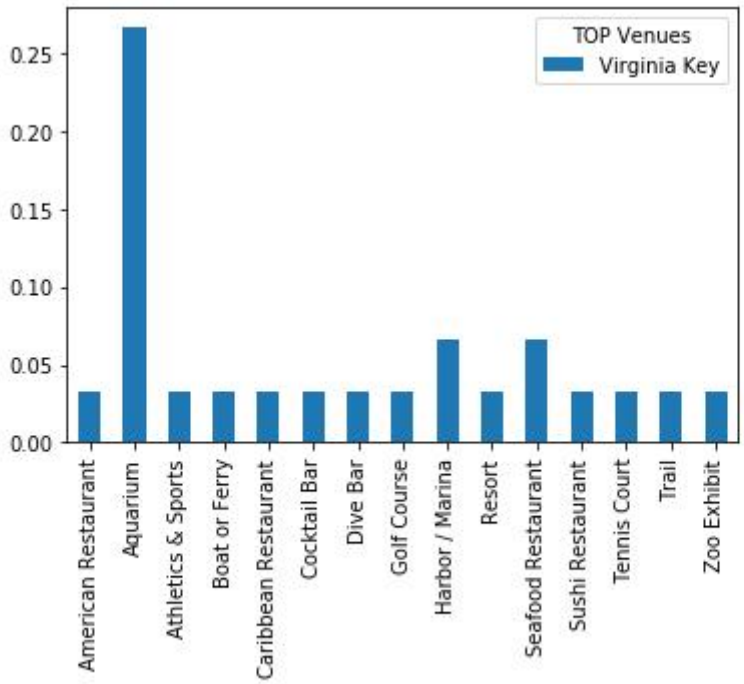


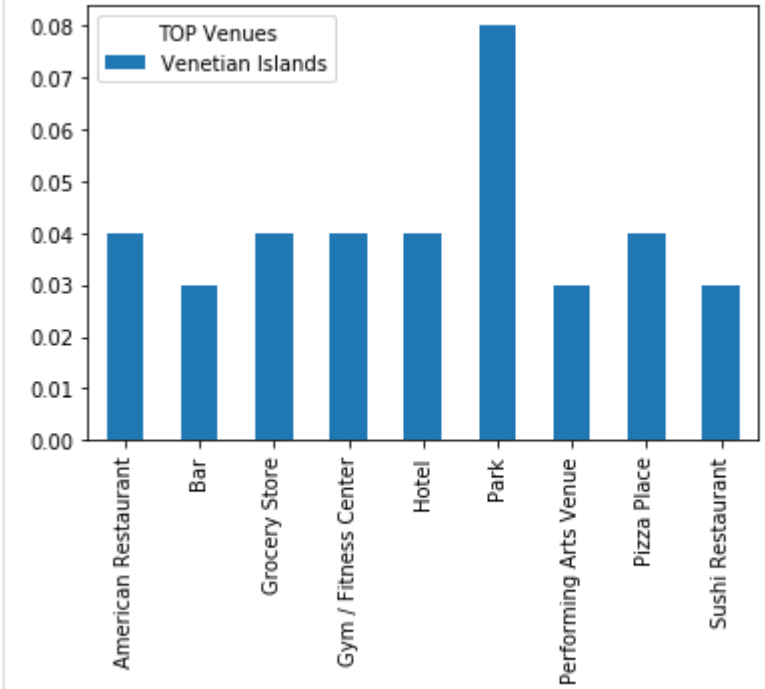
Most of the neighbourhoods have 88-100 unique venues - in Virginia Key and Liberty City are just few different venues.
Dataset telss me that I now have 2137 venues to analyze and unique categories of venues are 226. That means that local people (or visitors who use Forsquare) visit a lots of kinds different places. I could clean the amount a little bit less and visualize gategories, but to my goal it´s worthels, so I´m happy with this amount and go on.

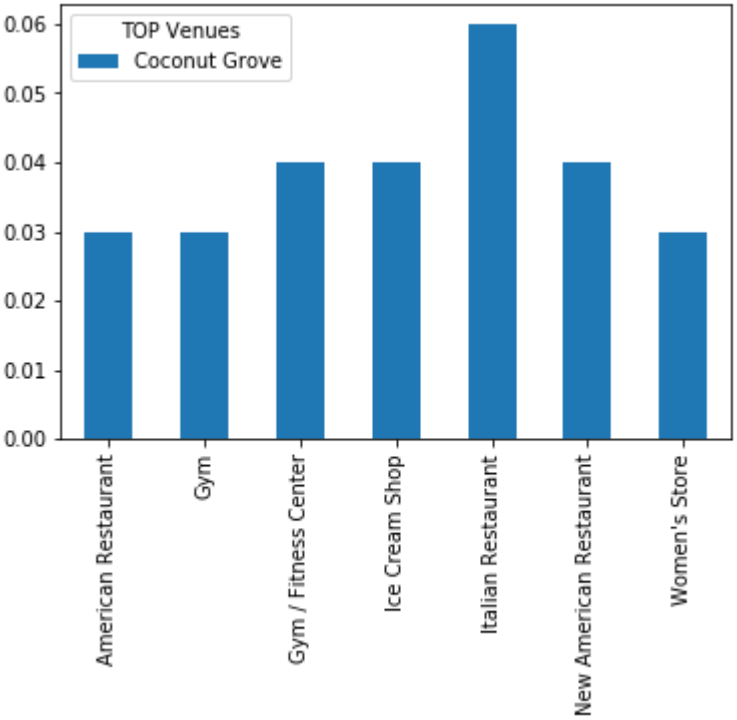
To use data abowe I need to index data so that it possible to count venues. That´s why I use one hot encoding to make a nwe datafreme to analyze more

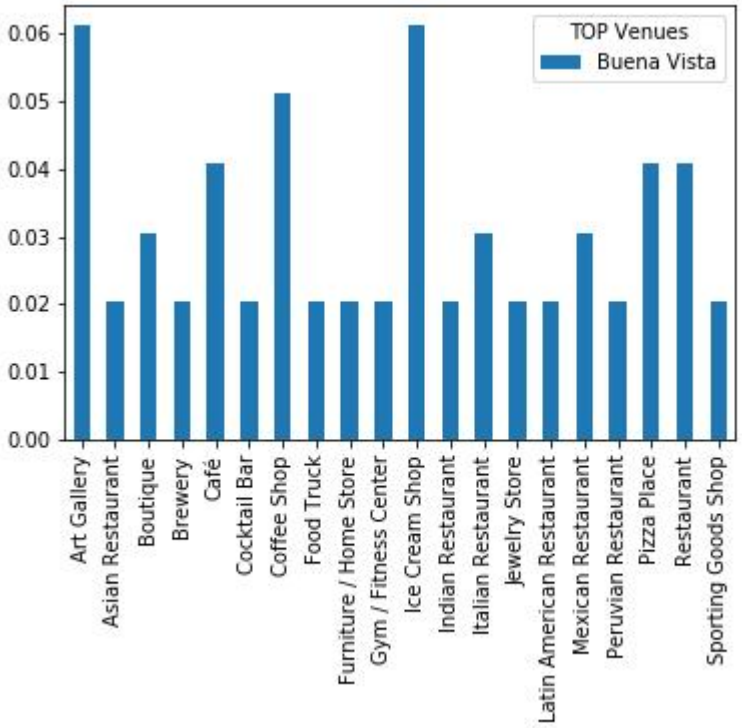
Analysis of the data¶

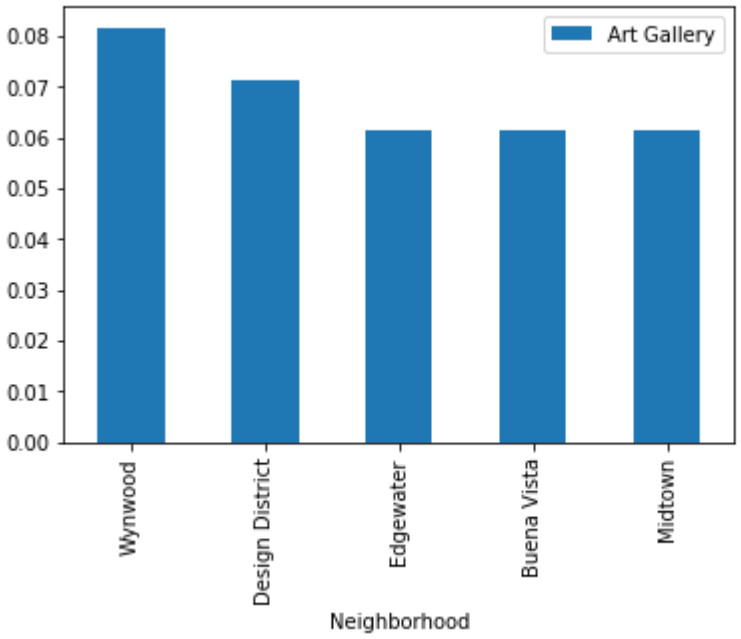
After encoding I can gruop neihgbourhoods and take the mean of frequency of occurence of each category. Result is shown below.

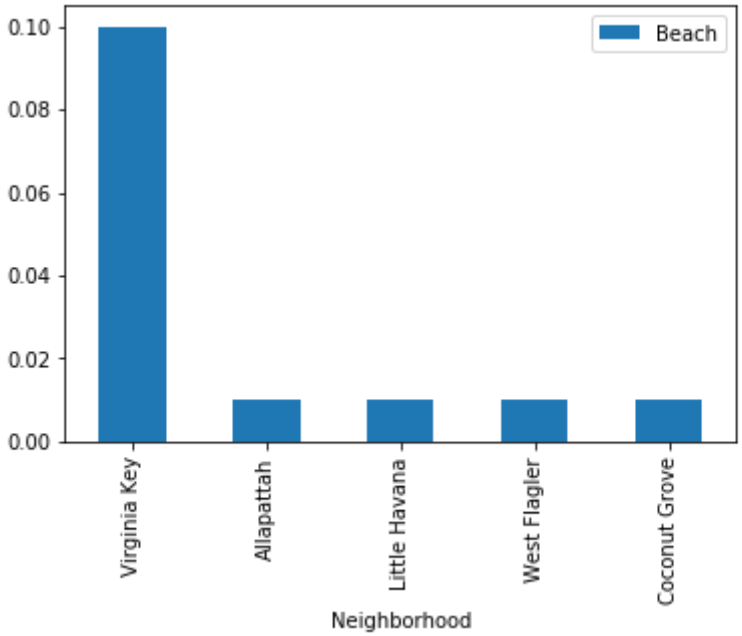

TOP Venues
Virginia Key
0.25
0.20
0.15
0.10
0.05
0.00
American Restaurant
Aquarium
Athletics & Sports
Boat or Ferry
Caribbean Restaurant
Cocktail Bar
Dive Bar
Golf Course
Harbor / Marina
Resort
Seafood Restaurant
Sushi Restaurant
Tennis Court
Trail
Zoo Exhibit

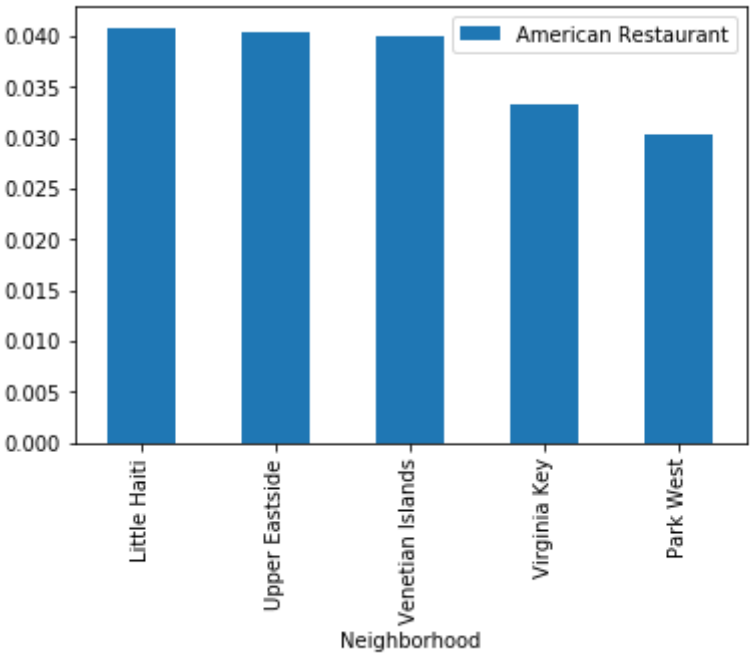

TOP Venues
Venetian Islands
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
American Restaurant
Bar
Grocery Store
Gym / Fitness Center
Hotel
Park
Performing Arts Venue
Pizza Place
Sushi Restaurant


TOP Venues
Coconut Grove
0.06
0.05
0.04
0.03
0.02
0.01
0.00
American Restaurant
Gym
Gym / Fitness Center
Ice Cream Shop
Italian Restaurant
New American Restaurant
Women's Store


TOP Venues
Buena Vista
0.06
0.05
0.04
0.03
0.02
0.01
0.00
Art Gallery
Asian Restaurant
Boutique
Brewery
Café
Cocktail Bar
Coffee Shop
Food Truck
Furniture / Home Store
Gym / Fitness Center
Ice Cream Shop
Indian Restaurant
Italian Restaurant
Jewelry Store
Latin American Restaurant
Mexican Restaurant
Peruvian Restaurant
Pizza Place
Restaurant
Sporting Goods Shop


Art Gallery
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
Wynwood
Design District
Edgewater
Buena Vista
Midtown
Neighborhood





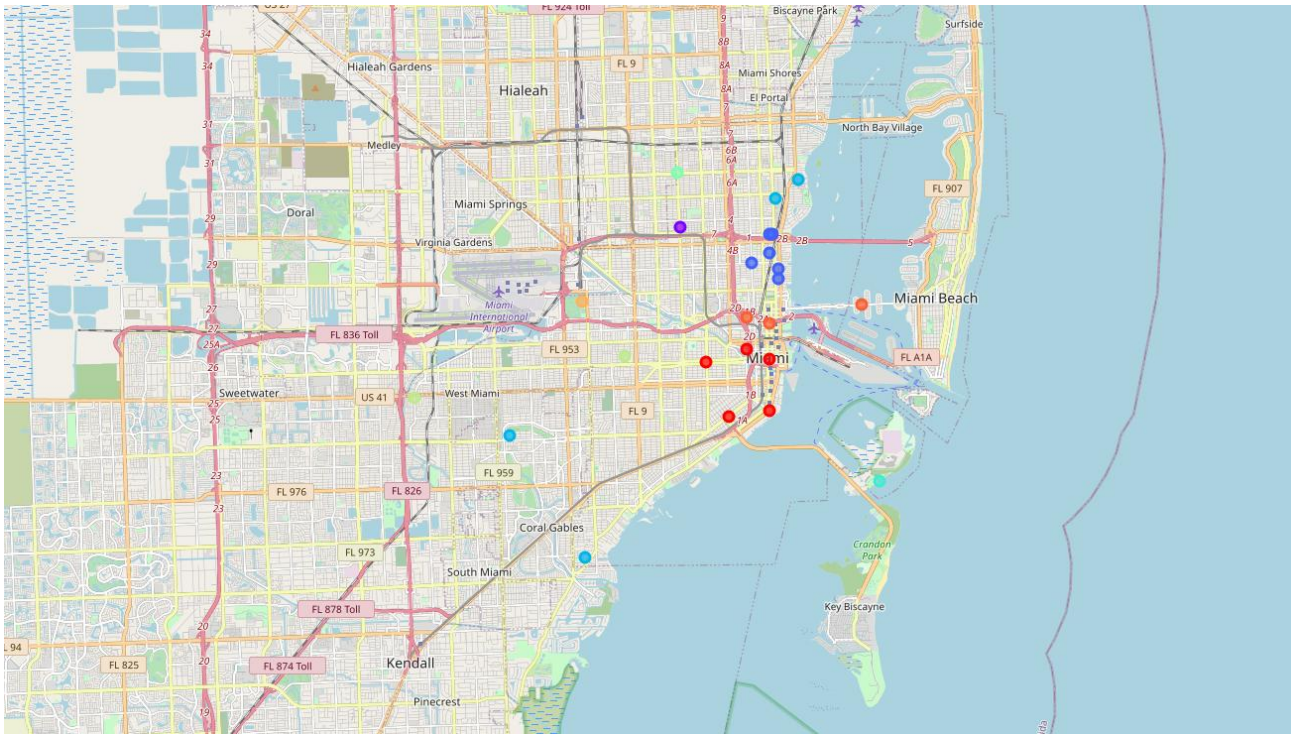
American restaurants you will find most likely from Little Haiti, Upper East Side and Venetian Islands
Now I make a top 5 venues list to every neigbourhoods to look at. By looking at top 5 venues I can say that neighbourhoods are not similar, some of them looklike the same, but some of them are unique

After examinating the list it´s hard to say directly how does the neihgbourhoods diverge to each others. Some of then most common venue is restaurant or hotel, some of art callery. Tho from this data I can make a default that every neihgborhood is a city neigbourhood - everyvere in Miami is something to do and something to eat. So basecly it doesn´t matter where to go there if u just want to hang around. Crapeland Heighs is doubtless an Airport area with airport services and in Brickell is many venues to be on a vagation. From Venetian Islands can doubtless get a ferry and in Virginia Key is also Harbor /Marina and beach.

To make easier to anlysize neighbourhoods I will cluster neighbourhoods to 9 clusters to see wich of the neighbourhoods wuold be the best for my client customers.

I tested clustering in different amount of labels and there´s always few neighbourhoods clearly different than others. Airport area comes out allways and Virginia Key is the second one to come out from others. After while clustering I ended up to 9 clusters - my goal is to provide data on different kinds of neighbourhoods to spen time on, so I dont´t want clusters to bee too big and more clusters brought out just good things.

Visualising Miami´s neighborhoods venues



In the map shown abouw we can see the labels 0- 8 in different colours. Label 0 = Red Label 1 = Violet Label 2 = Blue Label 3 = Light Blue Label 4 = Turquoise Label 5 = Light Green Label 6 = Green Label 7 = Light Orange Label 8 = Orange

Cluster 0 contains like cluster 2 lots of restaurants and coffee shops, but not so many hotels. Restaurants also seem to be more like "normal" food than in cluster 2. Here visitors can eat most likely also cheeper than in cluster 2. Mostly neigbourhoods offer Art and relaxing tho. Desing and Arts and entertainmenst are most likely to been found from here

Cluster 1 is basicly Virginia Key. From here visitors can find Park, Beach, Harbor/Marina, Golf and Tennis and Athletics & Sports. Suonds like a great place for visitors who wants beach and sports and are willing to walk a little bit to go shopping and eating and so on.

Cluster 2 contains hotels and restaurant and bars. These neighbourhoods wuold be a great place to stay for those travellers who wants to eat well and drink enough. For my purpose = to find a neighbourhood where locals also go often and hang around and live in Airbnb this is not so good neighbourhoods.

Restaurants in this area´s are specialist very much - Italian, Seafood, Japanese, Latin American, Cuban, Asian and even Spanish Restaurangs are here.

Cluster 3 is also a city area. Not so many arts galleries than in cluster 0 or hotels than in cluster two - more shops and coffee places.

Cluster 3 seems to be neighbourhoods where people live. Here you can find in most common venues list Grocery store, Bakery, Bank, Pharmacy, Gym and so on. This sounds like neighbourhoods to stay like "normal" miami people. From the first most common venues I assume that Cuban and Latin American people might live here.

In cluster 6 I will find most common venue, Italian restaurant. Might be, that Italian people live here. Here´s also shops, sporting and parks ans so on. This looks like a great place to stay if you love good and exspecially Italian food.

Cluster 4 also seems like a place where people live (Nightclub, Gas Station, Park and food places). I assume that this neihgbourhood is not so rich -in most common venues are Discount Store and Sandwich Place.

Cluster 4 is basicly Venetian Islands. Most common venue is Parks. Here is Hotels, Grocery stores, restaurants and bars and also a gym. In neighborhood also is Performing Arts venue, so this seems to be a good place for visitors who wants a little bit of everything. If you also look the map where this neighbourhood is - you will find that also a beach is nearby. Based on Crocery store I wuold say that people also live here - so I think that Airbnb wuold also find here good places to stay.

Cluster 8 is Grapeland Heights - Airport is nearby here. I wuold not recommend to stay here.

Discussion section

Problem wich I was trying to solve in this project was:

If customer is about to travel or thingking of travelling to Miami - wich neighborhoods are the best to stay? People and customer want different things, so I will make on analysis from different perpectives: for beach holiday for cultural holiday for shopping holiday and for both beach and something else to do easily also.

In my analysis I have compared neihgbourhoods to each other and clustered them into 9 different types of categories besed on venues nearby neighbourhoods. Analysis gies a good picture where wuold be a great place to stay for different types of travellers.

Let´s start by excluding neighbourhoods where I wuoldn´t recommend to travel if what clients customers wants a great holiday in singns of beach, cultural, shopping or a little bit of them all. Exluded neighbourhoods are on my opinion based on this analyzis: cluster 6 Liberty City - presumably great places to live and stay - but so much like living area and far from the sea and city and actually not much to do. cluster 8 Grapeland Heights - who wants to stay at airport? Nothing else needen to stay

So from 9 clusters I take out 2 and now I have 7 clusters. From them I can say: City areas: clusters 0 and 2 and 3 Possibilities to have beach and parks nearby and still be in the middle of everythin, clusters 1 and 7 Want to live like locals? Cluster 4 and 5 is the best one for that

Hotels and city areas: There is two this kind of areas, how are they different? Cluster 0 seems to be a city area with art galleries, Coffee shops, restaurants and boutiques. Most common venues are not hotels - actually in the 10 most common venues are not any hotels. Boutiques are in the most common

venues only in Buena Vista - so if I wuold like to go shopping that wuold be a great place. Cluster 2 is lots of like cluster 0, but it most common venues are hotels. There´s also much more Bars and parks than in cluster 0. Cluster 3 is for live like locals in neighbourhoods ful with restaurants, art galleries, coffee shops and so on - this seems like a good place. There are also bars if you are thirsty. My conclusion is that cluster 0 is centre and 2 and 3 a little bit further.

If city is what customers wants - these are the neighbourhoods to go. For customers in my assingment wuold be the best either cluster 8 or 3 - three if customer is thirsty often.

Beach and parks sections are found in clusters 1 Virginia Key and 7 Venetian Islands. Difference between these clusters is mostly that in cluster 1 beach is nearby and there´s lots of opportunities do to sports aswell. Cluster 7 provides more cultural experiments and happenings than sports, but also have more restaurants in common venues. So wich one to choose debends most of customer. Perhaps there will be more of apartments tho in cluster 7 to stay than in cluster 1.

Want to live like lokals i found to clusters, 4 and 5. Compairing these to each other it seems, that in cluster 5 is mostly Italian food and gyms and coffee shops and very much to do. In cluster 4 is not so many choises to eat different foods and the most common venues is Cuban and Latin American restaurant. I wuold recommend to visit cluster 5 of these clusters.

So now i have information from every cluster.

My conclusion is: For Beach holiday - both cluster 1 and 7 are good, but beach is clearly in cluster 1. For cultural holiday - cluster 0 is the best for this For shopping holiday - also cluster 0, but good choises are also clusters 2 and 3. If customer is in my project wanting to stay at Airbnb cluster 3 is the best. For both beach and something else to do - cluster 7 is the best choice.

For something else? As I did the analysis I found out one more. For live like locals in the city holiday - cluster 5 is my choice for that!

At this point it shuold be said, that my analysis is of Miami neighbourhoods. Miami Beach is an independed city, so Miami Bech neighbourhoods are not in my project. In the goal is to gi Miami, not Beach this results are relevant. But if thinking to go beach holiday, shuold really think about Miami Beach also. In Miami the nearest place of Miami Beaches is Venetian Islands.

Conclusion section

In this project I have been studing and analysing neigbourhoods of Miami.

Firs I have to say - that like a mentionde before - Miami Beach is an independed city, so that´s why beaches don´t come out so well as they can. If wanting a beach holiday, shuold so the research to Miami Beaches Venues also. At this project I concentrate only for Miami.

My goal was to find out where in Miami to go if customers want beach holiday, cultural holiday, or shopping holiday or something between them.

My recommendation will be:

For Beach holiday - both cluster 1 and 7 are good, but beach is clearly in cluster 1. (Virginia Key,Venetian Islands,Park West, Overtown)

For cultural holiday - cluster 0 is the best for this (Arts & Entertainment District, Buena Vista, Design District, Edgewater, Midtown, Wynwood)

For shopping holiday - also cluster 0, but good choises are also clusters 2 and 3. If customer is in my project wanting to stay at Airbnb cluster 3 is the best. (Arts & Entertainment District, Buena Vista,

Design District, Edgewater, Midtown, Wynwood / Brickell, Downtown, Little Havana, Lummus Park, The Roads / Allapattah)

For both beach and something else to do - cluster 7 is the best choice. (Venetian Islands,Park West, Overtown)