

Determination of Community Type

(based on demographic data for Miami-Dade County, FL)

Alexey D. Nechaev

Ph.D. in Law, L.L.M., M. Econ, M. Acc

April 8, 2020

Table of Contents

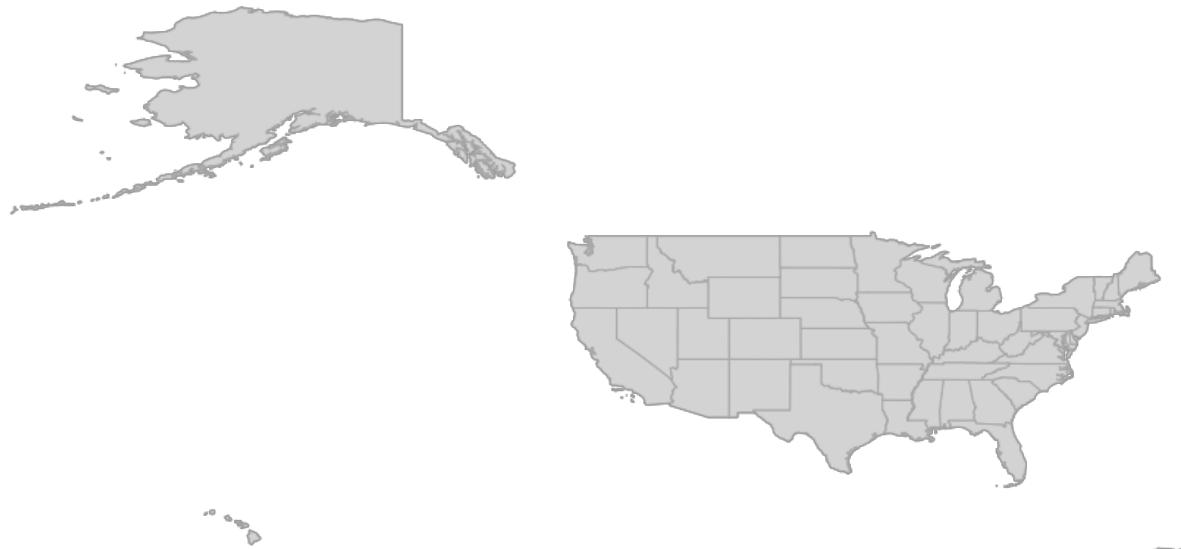
1. Introduction	3
2. Data.....	6
3. Methodology.....	8
4. Results	35
5. Discussion	61
6. Conclusion	61
7. Acknowledgement	61

1. Introduction

1.1. Background

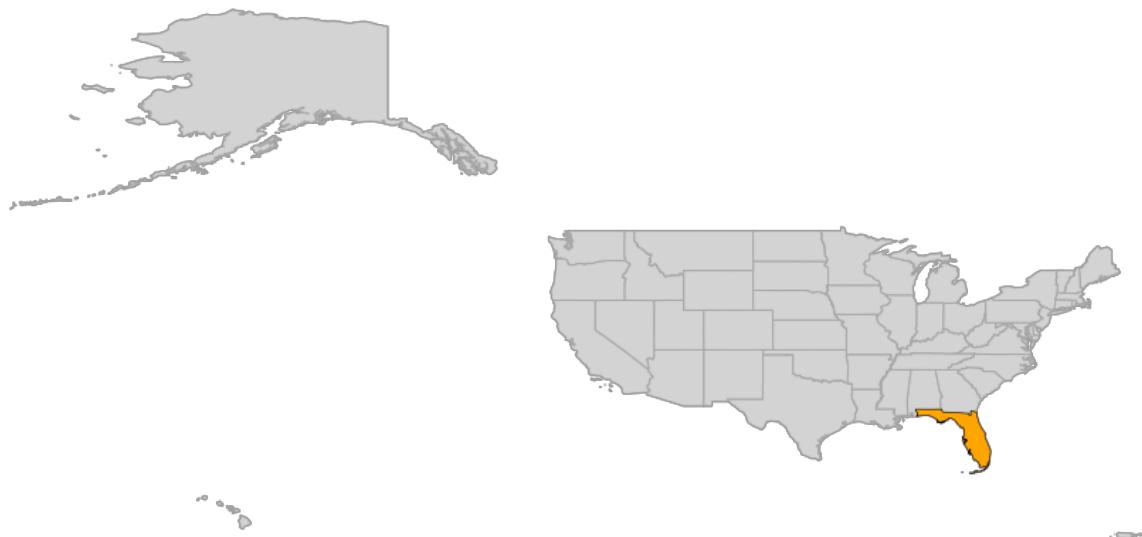
Local governments in the US have at least two tiers of local government (counties and municipalities) below the level of the state. The types of municipalities vary from state to state. The United States consists of 50 states, a federal district, five major territories, and various minor islands.

Map of the United States of America



Florida is the 27th state and the southernmost state in the US.

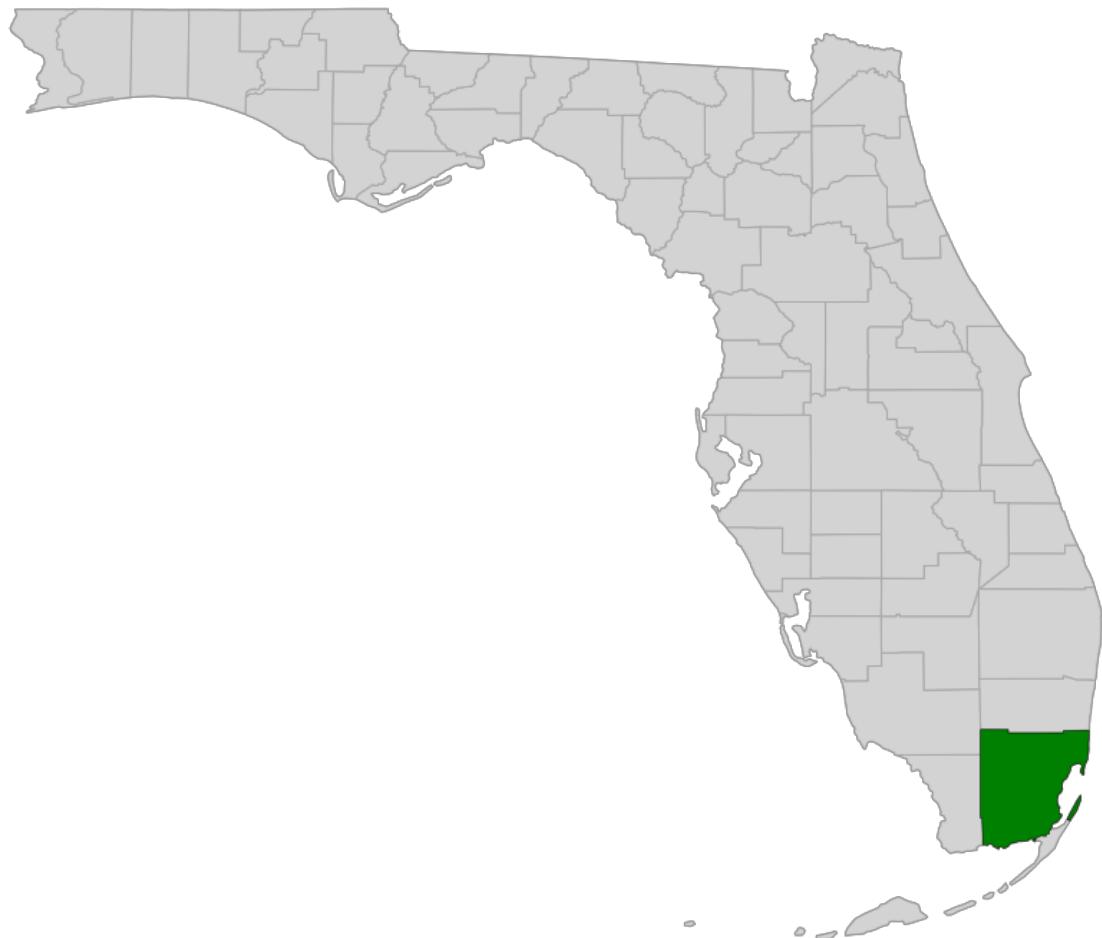
Map of the United States of America with Florida



According to Article VIII of the Constitution of the State Florida local governments in Florida are represented by the following types: counties and municipalities.

The State of Florida consists of 67 counties. Miami-Dade County is one of the counties of the State of Florida. It's situated in the southeastern part of Florida.

Map of Counties in Florida



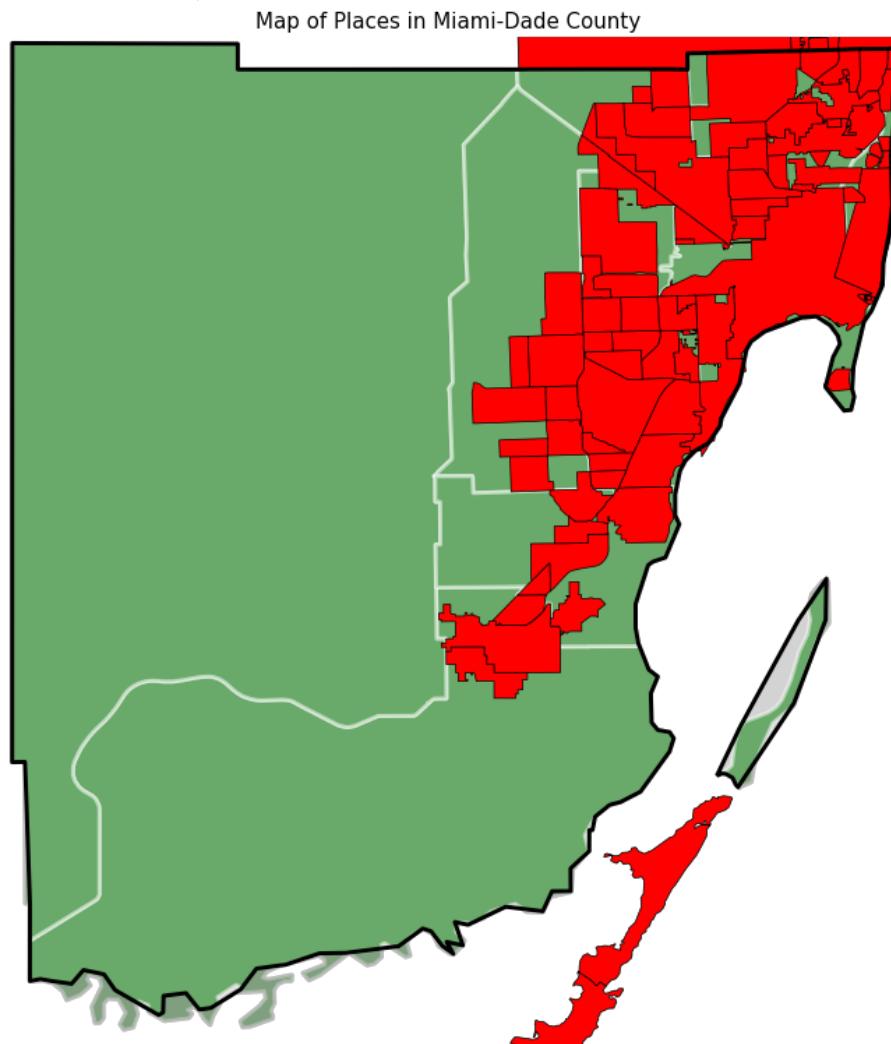
Municipalities in Florida are cities, towns and villages. However, there is no any legal distinction between these definitions.

There are 34 incorporated municipalities in Miami-Dade County, including cities, towns and villages. It's known that there are 19 cities, 6 towns and 9 villages in Miami-Dade County.

Moreover, there are a lot of Census-designed places (CDP). Census-designed places are delineated to provide data for population that are not legally incorporated under the laws of the State. According to the 2010 United States Census there are 37 CDPs in Miami-Dade County.

Municipalities and CDPs can be defined as communities (or places). The first ones are incorporated but the seconds are not corporate entities.

All communities in Miami-Dade County are presented on the map below in the black border of the County.



1.2. Problems

These observations in the background part of introduction lead us to the next problems:

- could the administrative division in Miami-Dade County be described as optimal?
- what type of municipality should be assigned to already existed incorporated communities and to CDPs (if they would be incorporated)?
- how could be changed the administrative division in Miami-Dade County?

The project aims to predict whether and how the types of communities should be distributed in Miami-Dade County.

1.3. Scope of Interest

On the one hand, the results of this project would be of interest to Miami-Dade County officials in prediction of further development of the county, including projecting of urbanization processes. On the other hand, it could be of interest to other specialists in the field of municipality and regional policy. It becomes more important due to the future results of CENSUS 2020.

2. Data

2.1. Literature Review on Demographic Data

As it was mentioned above according to the Florida law there is no certain legal criterion to distinguish cities, towns and villages. It's not the unique problem of Florida or even the US.

The executive director of the Wisconsin League of Municipalities D. Thompson pointed out that in Wisconsin “there are towns that operate like villages, villages that operate like towns, unincorporated villages that are actually towns and no real distinction between villages and cities”¹.

Population and housing size. The spokeswoman of the Colorado Municipal League S. Werner confirmed that “a town has a population of 2,000 or fewer; a city has a population of more than 2,000”². National Geographic consider that “a town is generally larger than a village, but smaller than a city. Some geographers further define a town as having 2,500 to 20,000 residents”. M.Rosenberg assumes that “villages are smaller than towns and towns are smaller than cities, though this is not always the case”³ He provides us with some examples of minimum population to determine community as a city: 200 residents (Sweden,

¹ https://lacrossetribune.com/town-village-city-what-s-the-difference/article_3ae0c206-d3e6-11e2-a6d7-001a4bcf887a.html

² <https://www.timescall.com/2019/10/20/johnnie-st-vrain-whats-the-difference-between-a-city-and-a-town/>

³ <https://www.thoughtco.com/difference-between-a-city-and-a-town-4069700#citation-2>

Denmark), 1,000 residents (Canada), 2,000 residents (France, Israel), 2,500 (Mexico), 50,000 (Japan). Some researchers are focusing on the number of **housing units** as a criterion.

Area size. E.E.Lampard conclude that a city, is “a relatively permanent and highly organized center of population, of greater size or importance than a town or village”⁴

Age distribution. Information that villages could be for older people and the young people go for cities⁵ gives us the criterion “age distribution”.

Racial diversity. The big cities could have more racial diversity than smaller towns and villages⁶.

2.2. Data limitations

Based on the literature review we choose 5 main factors (population, area, housing units, age and race) for further determination of type of municipality and CDP’s. Most of these factors have demographical nature. However, it would be inaccurate to consider that only demographic factors influence on the type of municipality. There are plenty of social factors as location of production, labor division, transportation system efficiency, cultural environment, medical service and so on. This project is concentrated only on demographical features of communities according to CENSUS 2010. It’s an important limitation that should be taken into account.

2.3. Other Data

In this project we also use a lot of data about location of communities, their borders and central points.

2.4. Data Sources

Most parts of data processed in this research are accumulated by CENSUS Bureau on their website <https://www.census.gov/>. Some of data (for instance, about location) can be received by in-built Python libraries. To find out the existence and the numbers of grocery and pharmacy stores we use Foursquare API.

⁴ <https://www.britannica.com/topic/city>

⁵ <https://news.ufl.edu/articles/2015/09/should-older-americans-live-in-places-segregated-from-the-young.html>
<https://www.thetimes.co.uk/article/villages-growing-older-as-young-head-for-the-cities-cvmdl3kk9>

⁶ <https://theconversation.com/diversity-is-on-the-rise-in-urban-and-rural-communities-and-its-here-to-stay-69095>

3. Methodology

In the research we use the following methods:

- data collection – public records, Foursquare queries;
- data preprocessing – methods of data cleansing, data editing, data reduction and data wrangling;
- data understanding – exploratory data analysis;
- modelling – clustering (K-means clustering and Hierarchical Clustering);
- evaluation – observation, comparison.

The problem that we are trying to solve in this research has the aim to find the structure of data, to group objects based on the similarity to one group of objects and dissimilarity to others. We will use K-means clustering for type of community's rearrangement and Hierarchical clustering to find out the optimal merging of communities.

In this research we use Python for achieving our goals.

After installing and importing the necessary libraries the steps of research are the following:

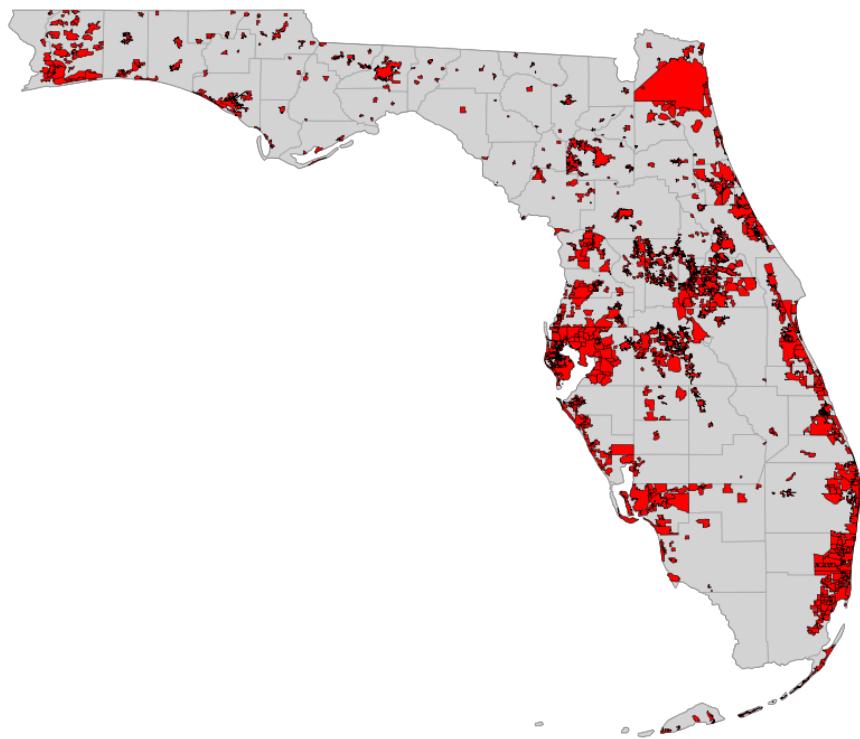
3.1. Visualization of geographical objects

Using the CENSUS data about shapes of the US, states of the US, Florida, counties and communities in Florida we created a range of schemes and maps applying Geopandas. It was necessary for representation of the research scope. Most number of schemes and maps are presented in the introduction of this project. Some more results of plotting are represented below.

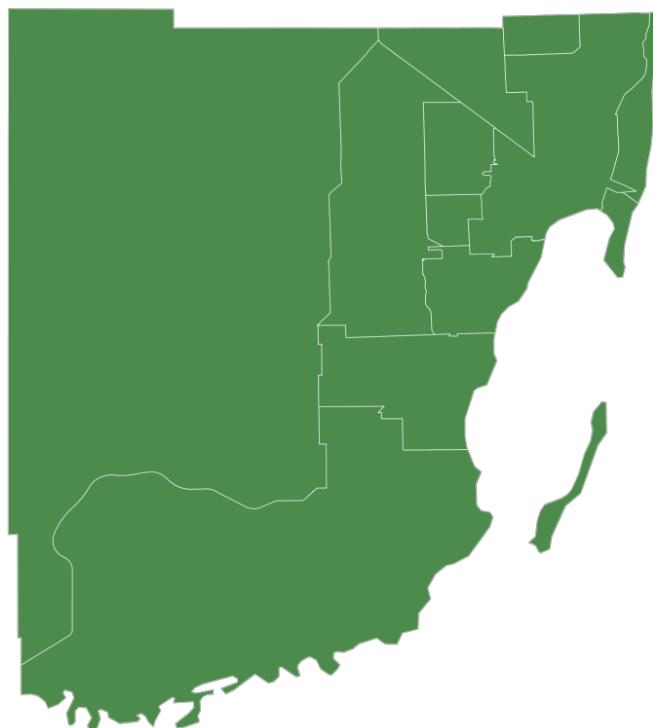
The process of plotting geodata consists of the next stages: downloading zipfile from CENSUS Bureau web-directory, unzipping it, applying GeoPandas functions to build GeoDataFrame and plotting schemes and maps.

This stage was basic because a lot of further representation of data is based on this scheme-maps.

Map of Places in Florida (with missing data)



Map of Places in Miami-Dade County



3.2. Data analysis and visualization

The next step is about collecting the data from CENSUS Bureau web-directory, reading it using Pandas ‘.read_csv’ function, data cleansing, data editing, data reduction and data wrangling.

It allows us to receive different DataFrames:

- 1) *df_PAH*, representing GEOID, Community, Type, Area, Population, Housing Units

	GEOID	Community	Type	Population	HousingUnits	Area
0	1590000US1208602681	Aventura	City	35762	26120	3.51
1	1590000US1208603275	Bal Harbour	Village	2513	2780	0.65
2	1590000US1208603975	Bay Harbor Islands	Town	5628	3199	0.53
3	1590000US1208606600	Biscayne Park	Village	3055	1324	0.63
4	1590000US1208609000	Brownsville	CDP	15313	5797	2.28

- 2) *df_AS*, representing GEOID, groups of age (0-17, 18-24, 25-44, 45-64, 65-..., Median Age, the proportion of number of males to 100 females)

	GEOID	<18%	18-24%	25-44%	45-64%	>=65%	MedianAge	M/100F
0	1590000US1208602681	15.4	5.5	27.8	24.8	26.6	46.1	84.6
1	1590000US1208603275	15.5	4.2	21.7	26.5	32.1	51.0	79.5
2	1590000US1208603975	21.7	5.0	30.0	26.1	17.1	41.0	87.0
3	1590000US1208606600	21.5	6.4	31.6	29.6	10.9	40.4	94.8
4	1590000US1208609000	29.6	11.6	24.8	23.1	10.8	31.1	90.8

This DataFrame is merged with df_PAH and becomes *df_PAHAS*, updated by real (not percentage) values.

Population	M	F	M/100F	MedianAge	<18	18-24	25-44	45-64	>=65	<18%
35762	16326	19436	84.6	46.1	5364	1788	9656	8583	10371	15.38
2513	1109	1404	79.5	51.0	377	101	528	653	854	15.50
5628	2618	3010	87.0	41.0	1182	281	1688	1463	1013	21.72
3055	1480	1575	94.8	40.4	642	183	947	886	397	21.50
15313	7254	8059	90.8	31.1	4441	1684	3675	3522	1991	29.62

3) df_R , representing GEOID, groups of race (White, Latino, Black, etc.)

	GEOID	Black%	Indian%	Asian%	Hawaiian%	OtherRace%	Latino%	White%
0	1590000US1208602681	3.9	0.1	1.8	0.0	2.0	35.8	57.9
1	1590000US1208603275	2.1	0.1	0.9	0.0	0.8	28.7	68.1
2	1590000US1208603975	2.5	0.2	1.2	0.0	2.2	46.3	49.1
3	1590000US1208606600	17.6	0.2	3.3	0.1	2.8	35.2	43.4
4	1590000US1208609000	74.7	0.3	0.0	0.0	3.3	25.7	1.2

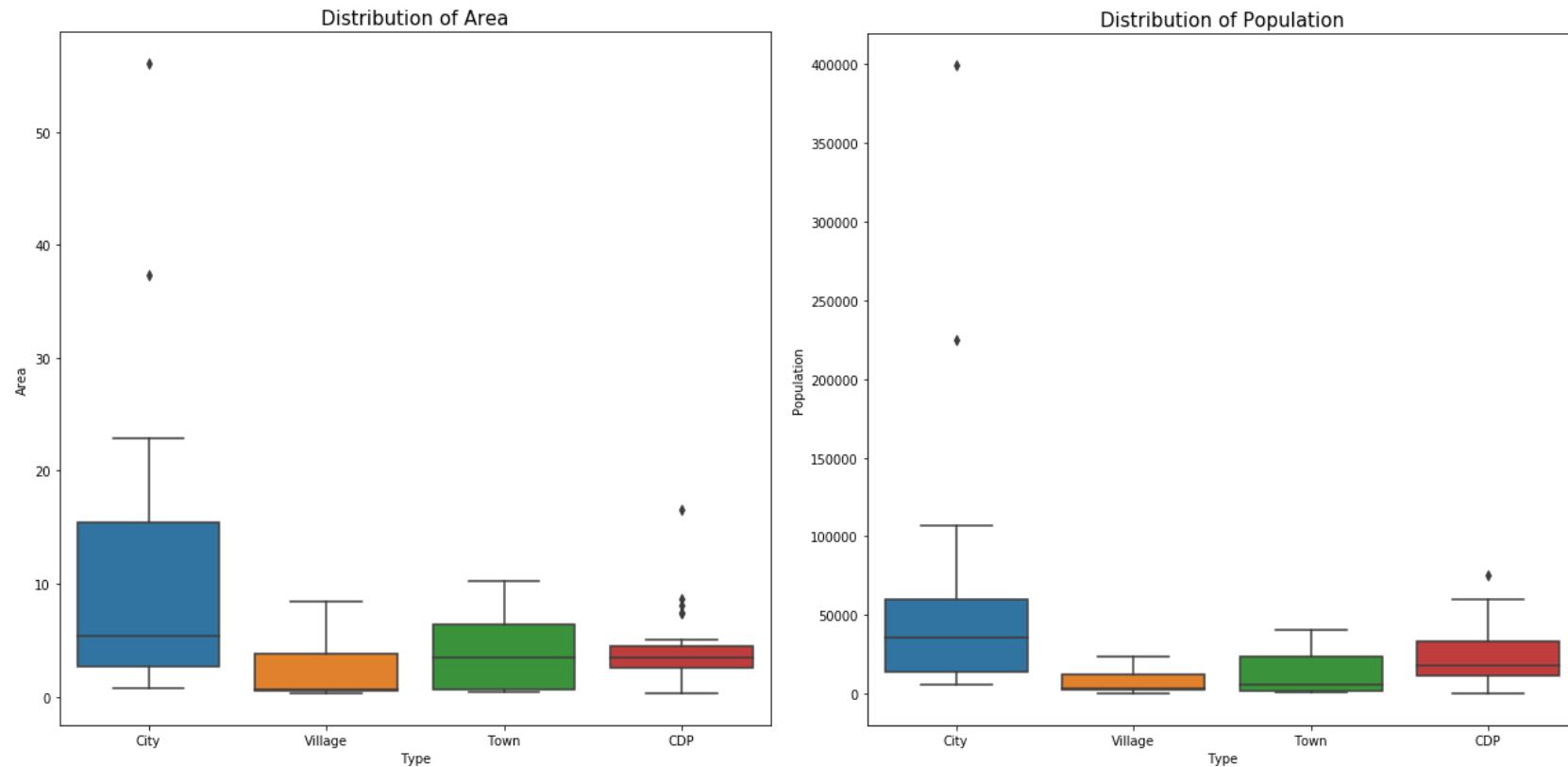
This DataFrame is merged with df_PAHAS and becomes df_PAHASR updated by real (not percentage) numbers for racial groups (White, Latino, Black, Others).

White	Latino	Black	Others
20384	12159	1073	2146
1709	704	50	50
2758	2476	113	281

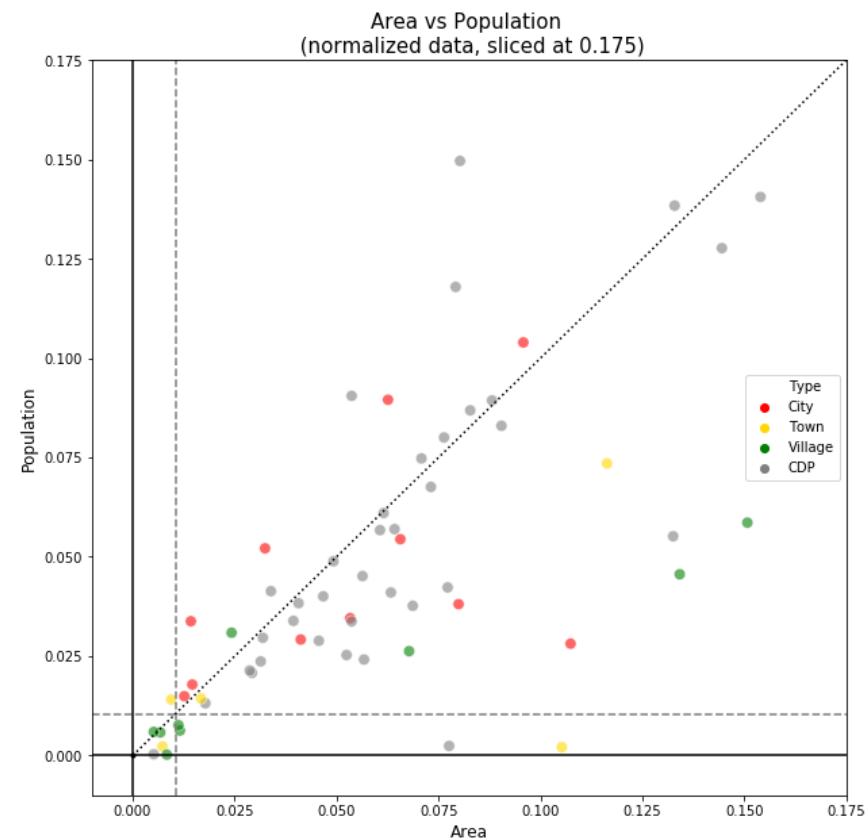
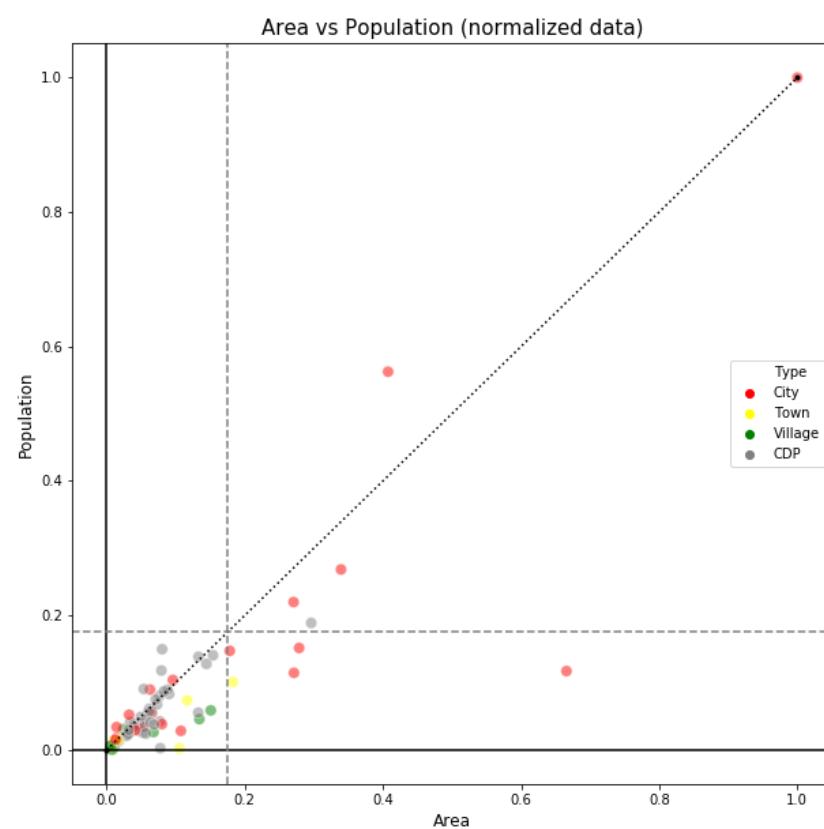
For each of discussed earlier DataFrames we receive descriptive statistics using functions ‘.value_counts()’ and ‘.describe()’. Descriptive statistical information for df_PAH is represented below.

	Population	HousingUnits	Area
count	71.000000	71.000000	71.000000
mean	32230.788732	13017.450704	5.918592
std	54604.511066	24254.640352	8.519586
min	86.000000	33.000000	0.290000
25%	9876.500000	3650.500000	1.805000
50%	16898.000000	6629.000000	3.550000
75%	35729.000000	13468.500000	5.960000
max	399457.000000	183994.000000	56.060000

For the number of variables like Area, Population and Housing Units we create **Box Plots**. They allow us to define some outliers and analyze the distribution of this variables. Some Box Plots are presented below.



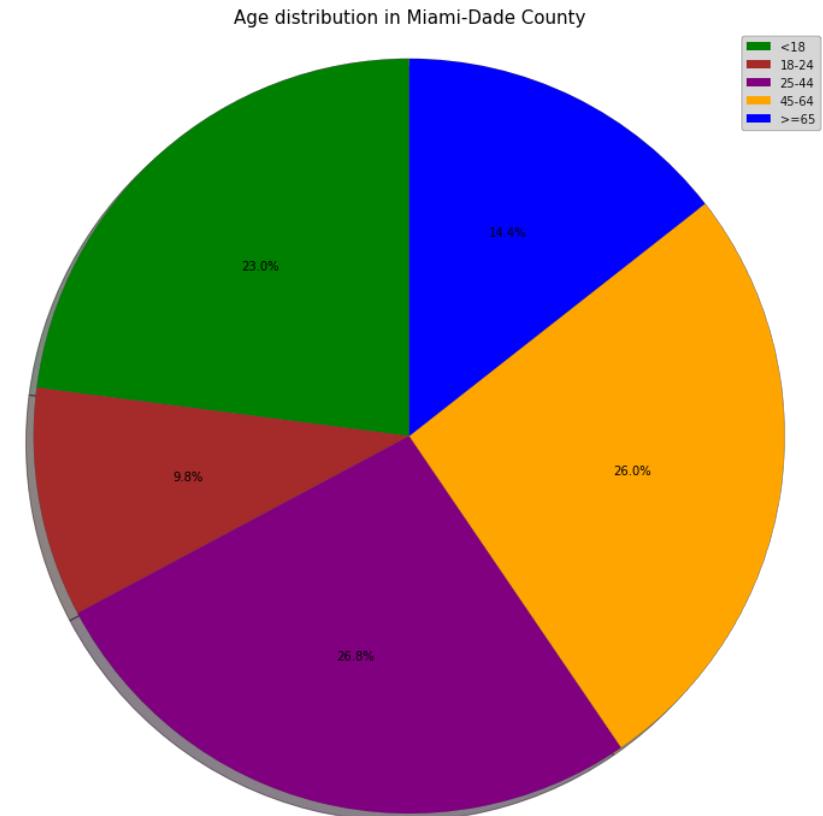
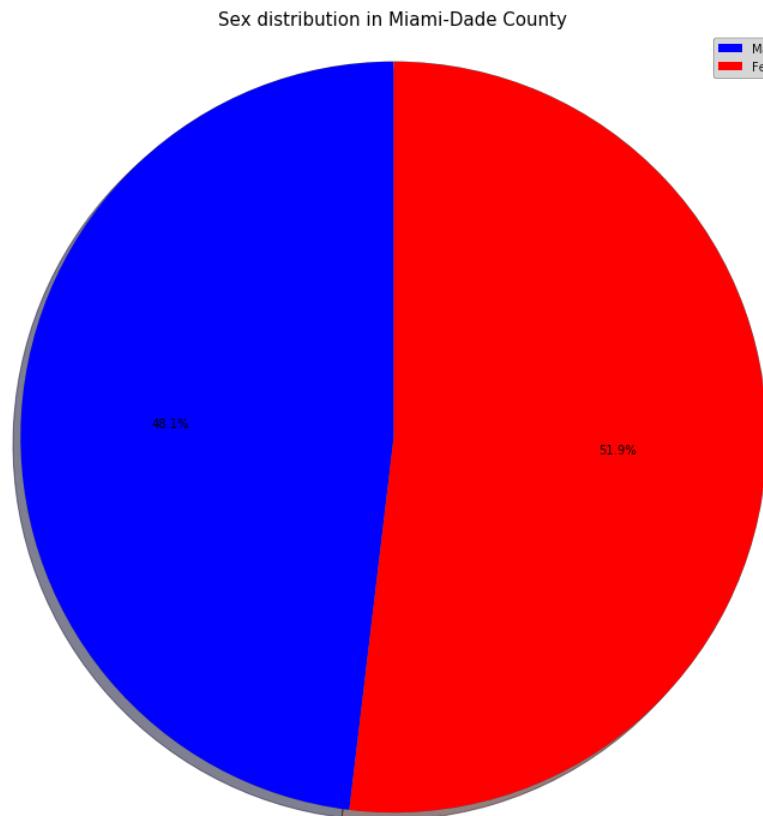
Based on Scatter Plots we conclude that it's impossible to determine the type of community just on the population, area or housing units and area. Some Scatter Plots are below.

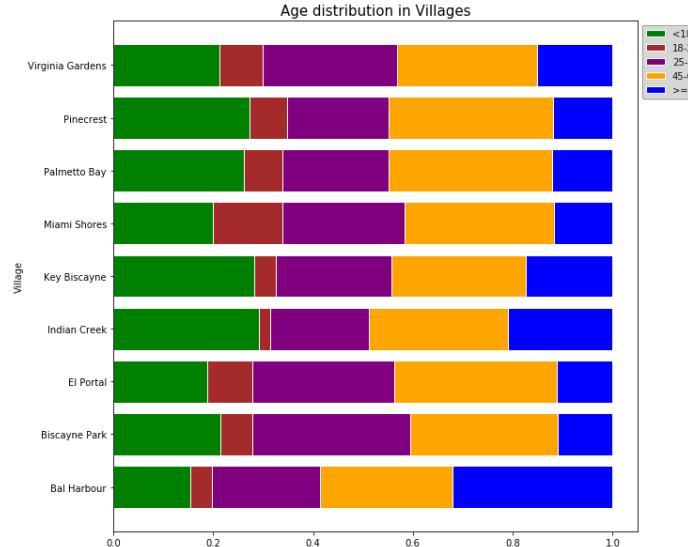
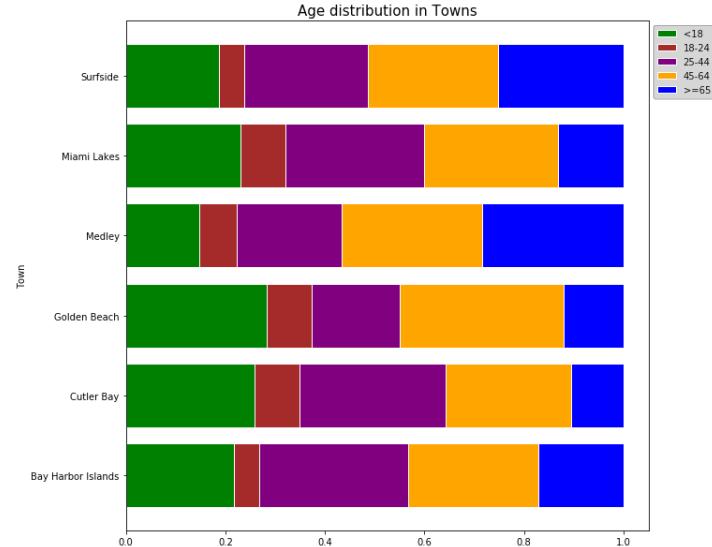
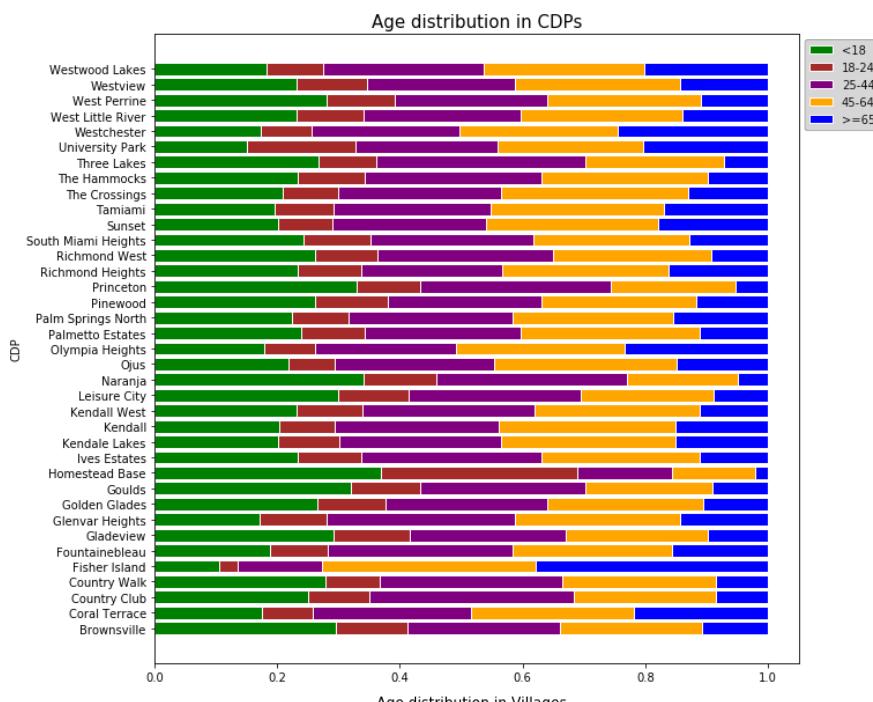
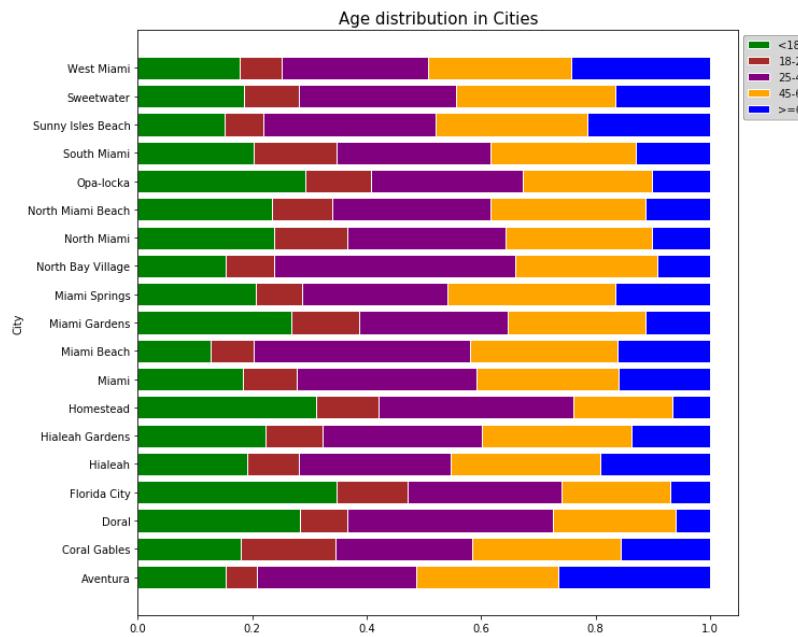


We use also **Pie Plots** and **Bar Charts** to represent the data about:

- 1) *sex distribution*
- 2) *age distribution*

We cannot find any specific and interesting data about sex differences in cities, towns and villages. But the age differences are crucial for some types of communities.

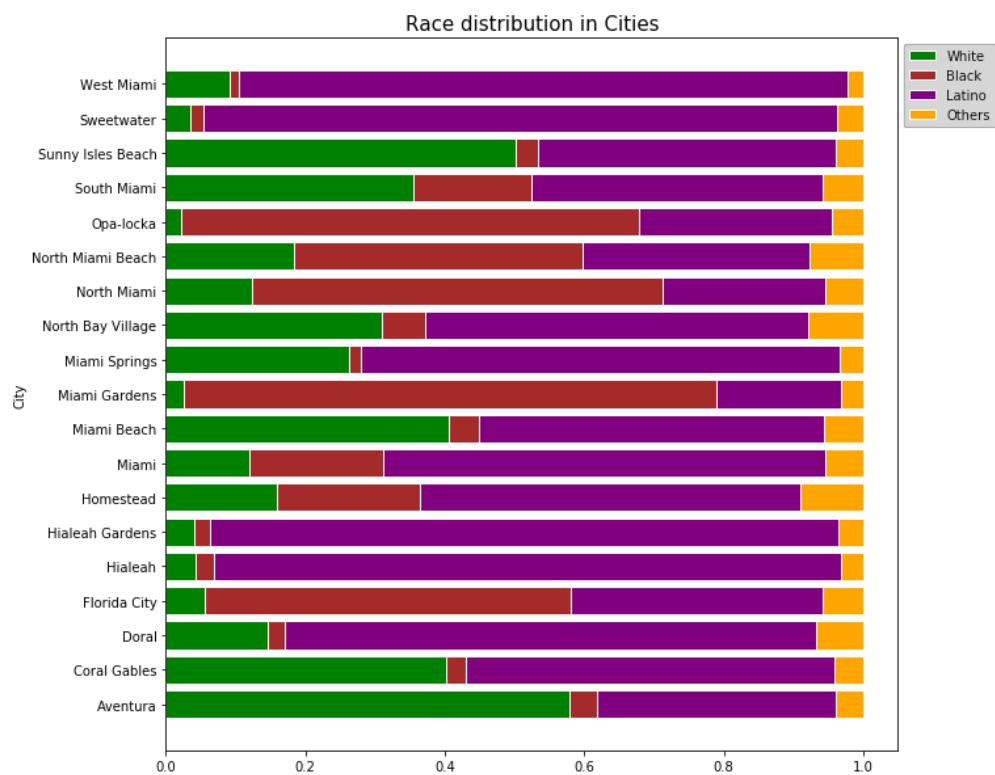
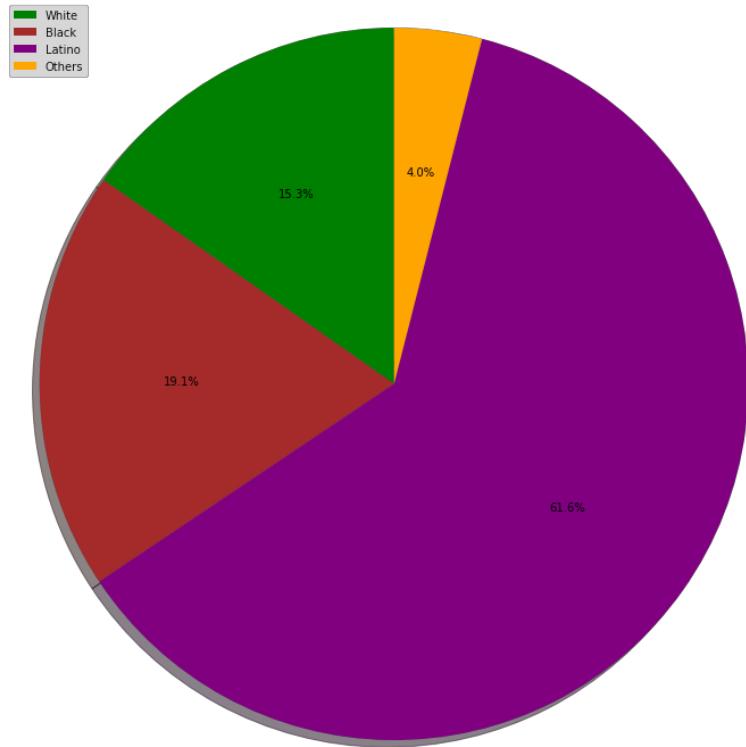


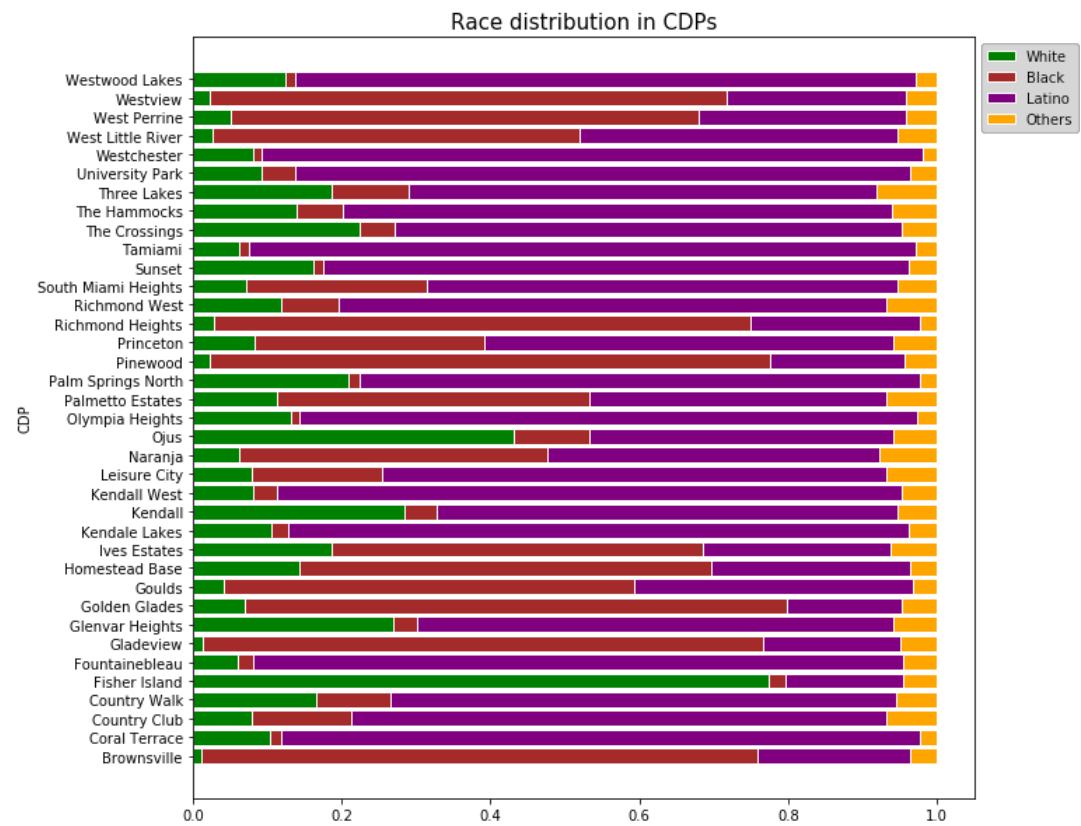
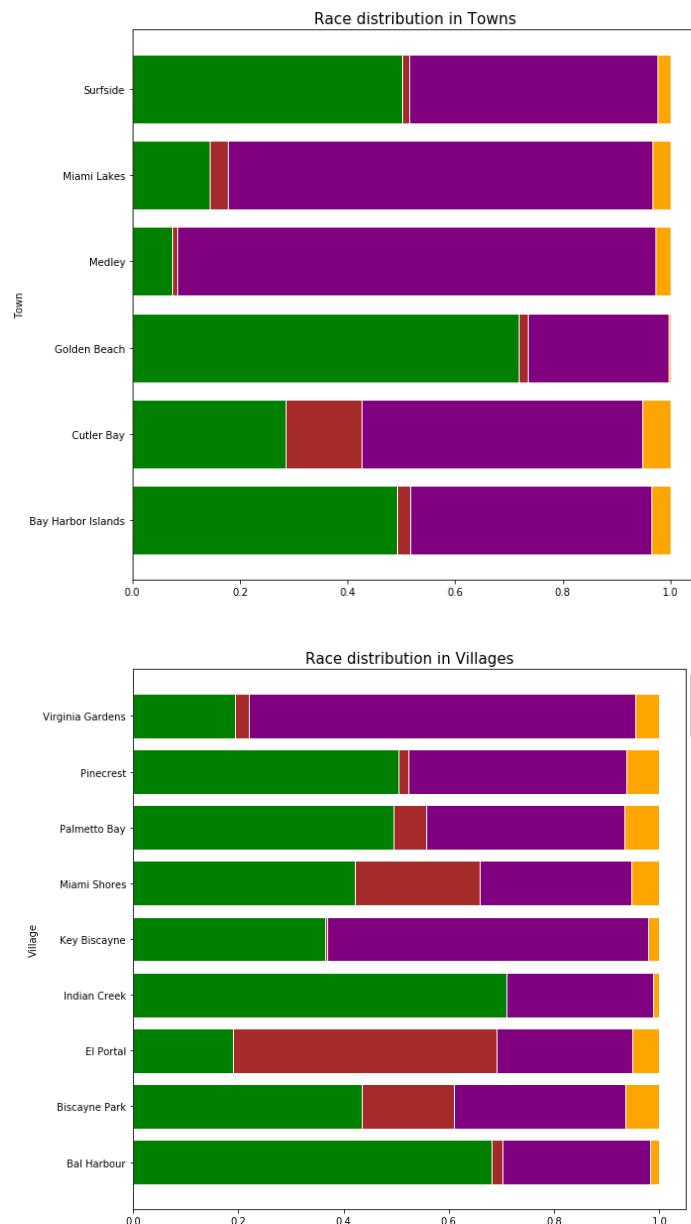


3) race distribution

Miami-Dade County is Latino oriented County. In most of cities Latino is the major race, but some east-north cities (Sunny Isles Beach, Aventura) has White Majority or some central-north has Black Majority (Hialeah).

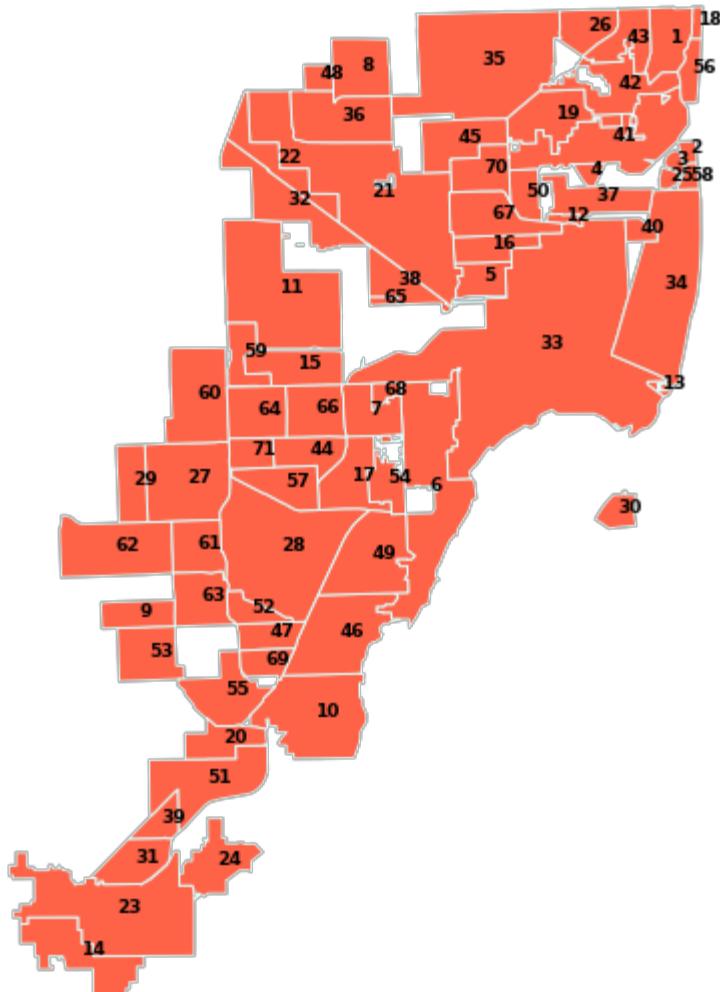
Race distribution (aggregated) in Miami-Dade County





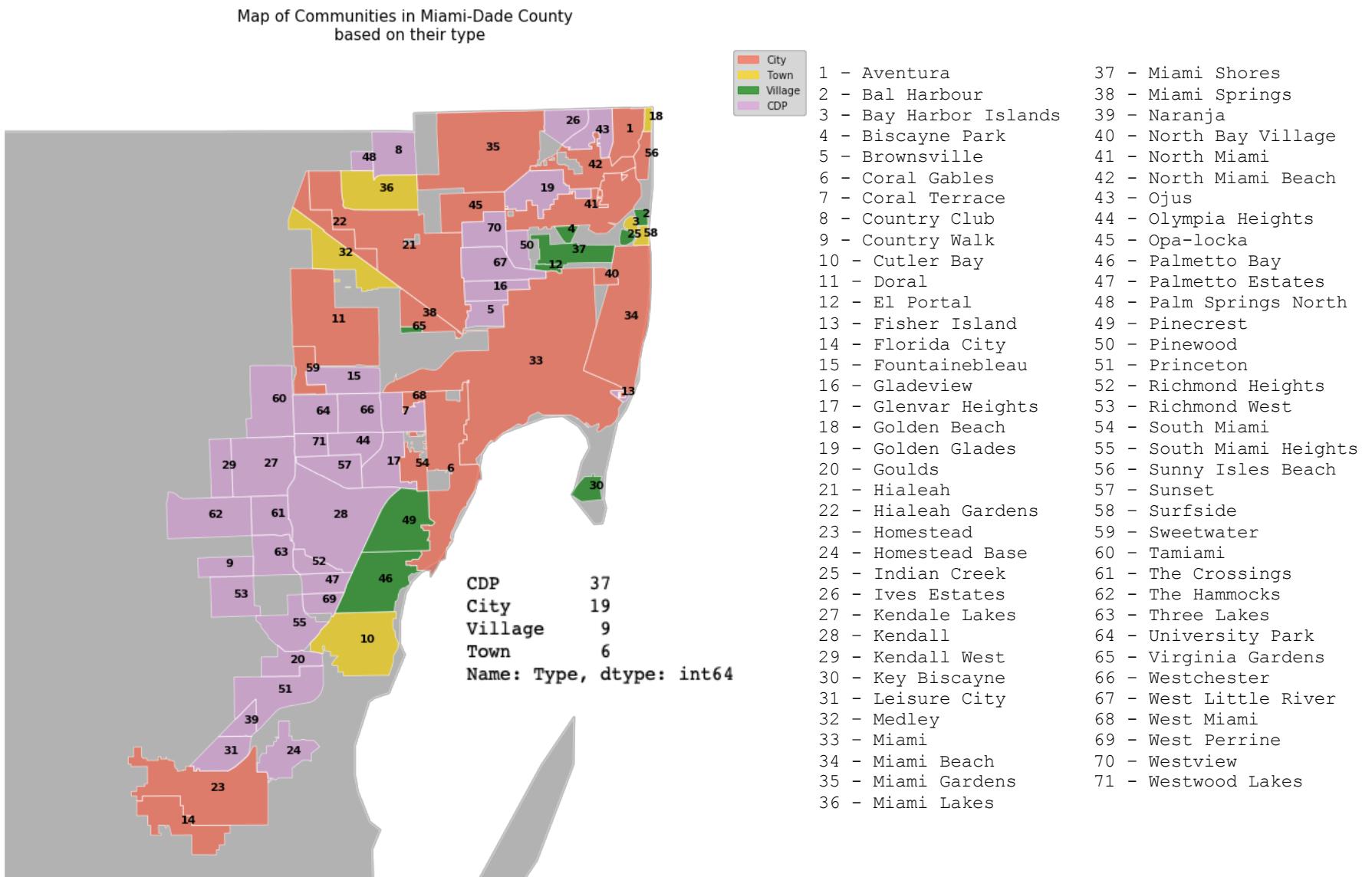
Using the list of Communities in Miami-Dade County and the GeoData polygons from CENSUS we plot all Miami-Dade County communities with their borders and numerate them in the alphabetic order. Unfortunately, it was impossible earlier because CENSUS place geometry data is not associated with certain Community) .

Map of Communities in Miami-Dade County

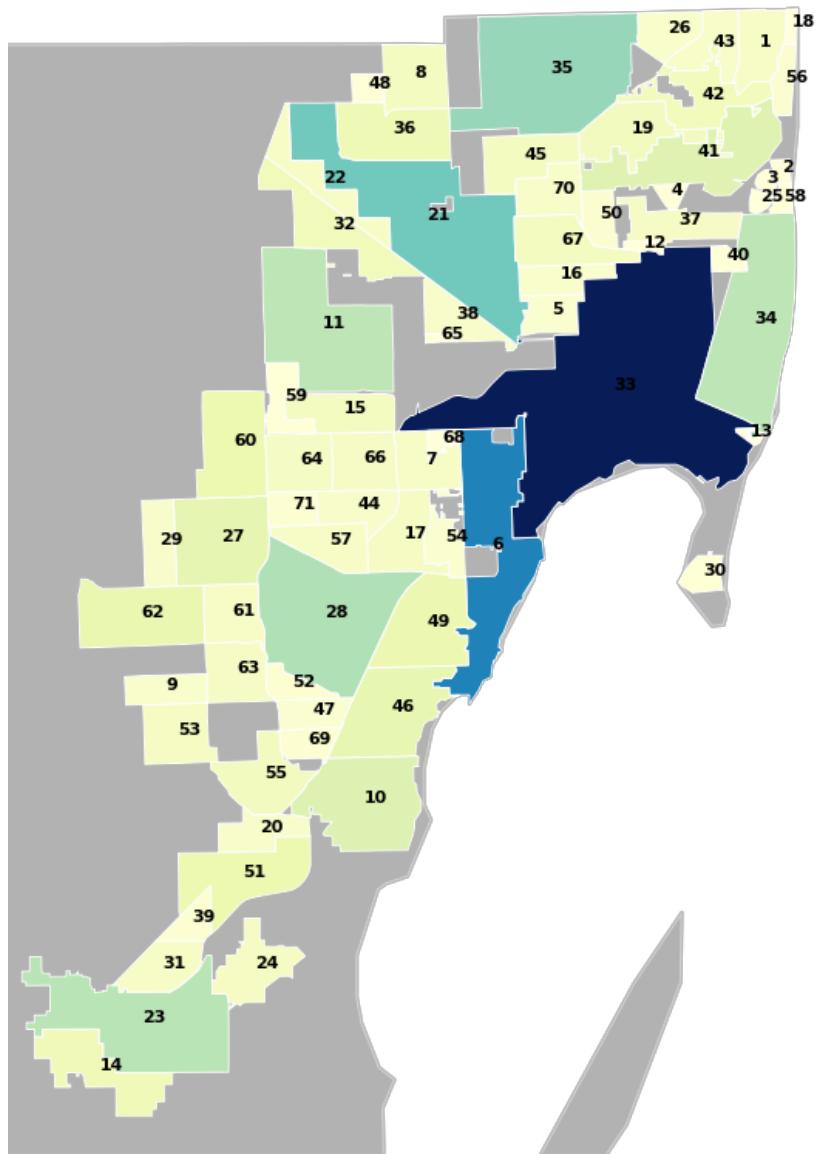


- 1 - Aventura
- 2 - Bal Harbour
- 3 - Bay Harbor Islands
- 4 - Biscayne Park
- 5 - Brownsville
- 6 - Coral Gables
- 7 - Coral Terrace
- 8 - Country Club
- 9 - Country Walk
- 10 - Cutler Bay
- 11 - Doral
- 12 - El Portal
- 13 - Fisher Island
- 14 - Florida City
- 15 - Fountainebleau
- 16 - Gladeview
- 17 - Glenvar Heights
- 18 - Golden Beach
- 19 - Golden Glades
- 20 - Goulds
- 21 - Hialeah
- 22 - Hialeah Gardens
- 23 - Homestead
- 24 - Homestead Base
- 25 - Indian Creek
- 26 - Ives Estates
- 27 - Kendall Lakes
- 28 - Kendall
- 29 - Kendall West
- 30 - Key Biscayne
- 31 - Leisure City
- 32 - Medley
- 33 - Miami
- 34 - Miami Beach
- 35 - Miami Gardens
- 36 - Miami Lakes
- 37 - Miami Shores
- 38 - Miami Springs
- 39 - Naranja
- 40 - North Bay Village
- 41 - North Miami
- 42 - North Miami Beach
- 43 - Ojus
- 44 - Olympia Heights
- 45 - Opa-locka
- 46 - Palmetto Bay
- 47 - Palmetto Estates
- 48 - Palm Springs North
- 49 - Pinecrest
- 50 - Pinewood
- 51 - Princeton
- 52 - Richmond Heights
- 53 - Richmond West
- 54 - South Miami
- 55 - South Miami Heights
- 56 - Sunny Isles Beach
- 57 - Sunset
- 58 - Surfside
- 59 - Sweetwater
- 60 - Tamiami
- 61 - The Crossings
- 62 - The Hammocks
- 63 - Three Lakes
- 64 - University Park
- 65 - Virginia Gardens
- 66 - Westchester
- 67 - West Little River
- 68 - West Miami
- 69 - West Perrine
- 70 - Westview
- 71 - Westwood Lakes

Further visualization is done for: **Type of Community, Area / Population / Housing size, Age / Race types**

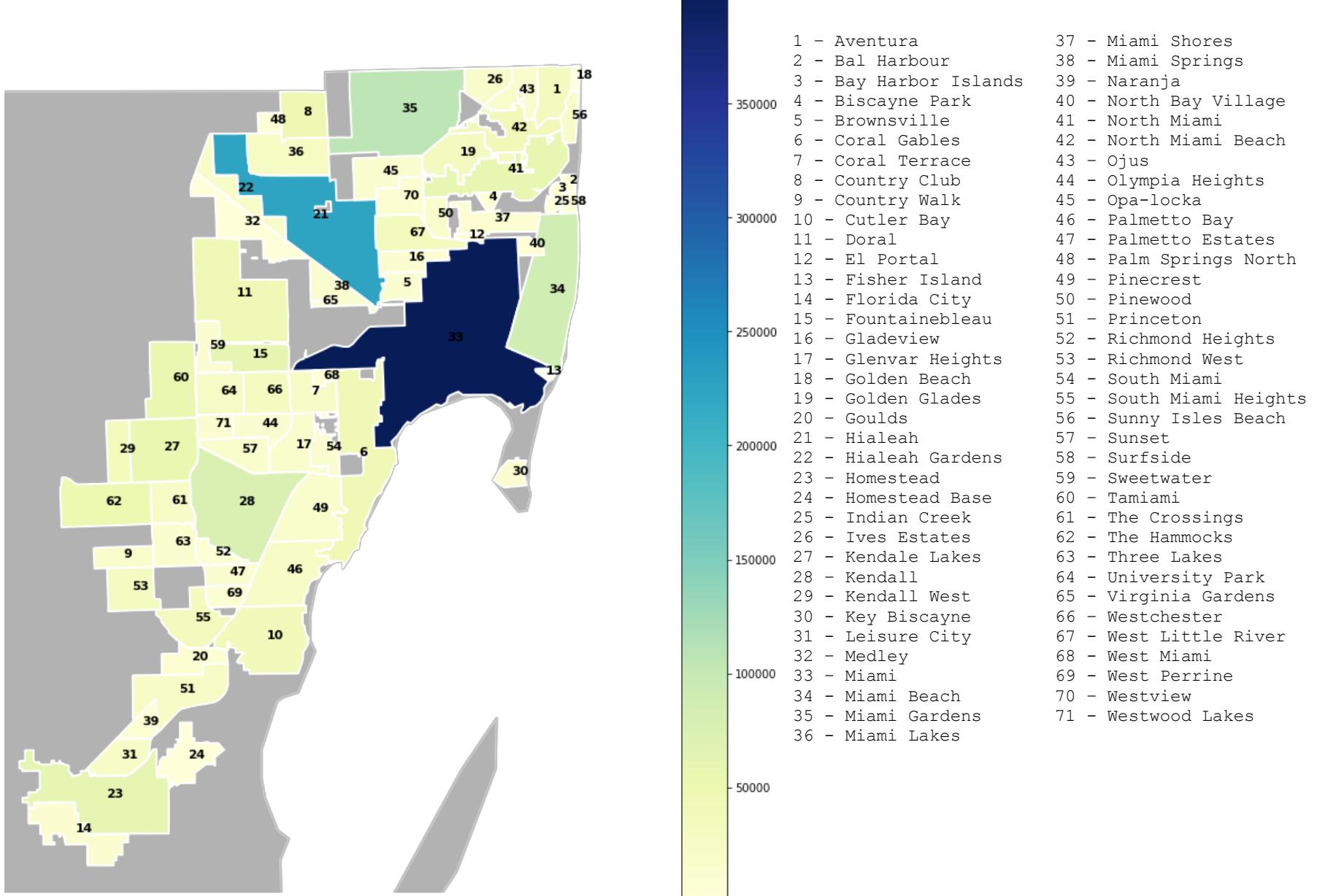


Map of Communities in Miami-Dade County
based on their area

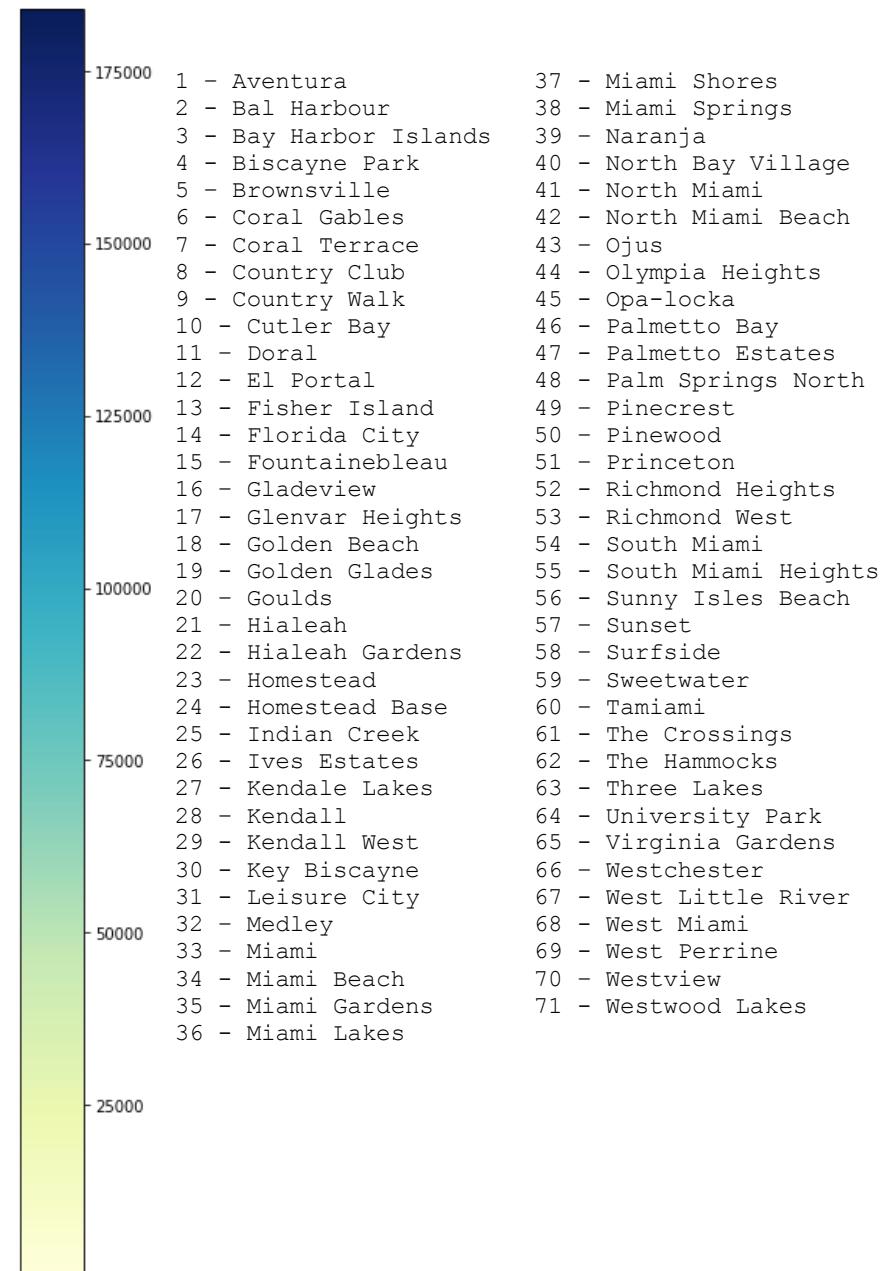
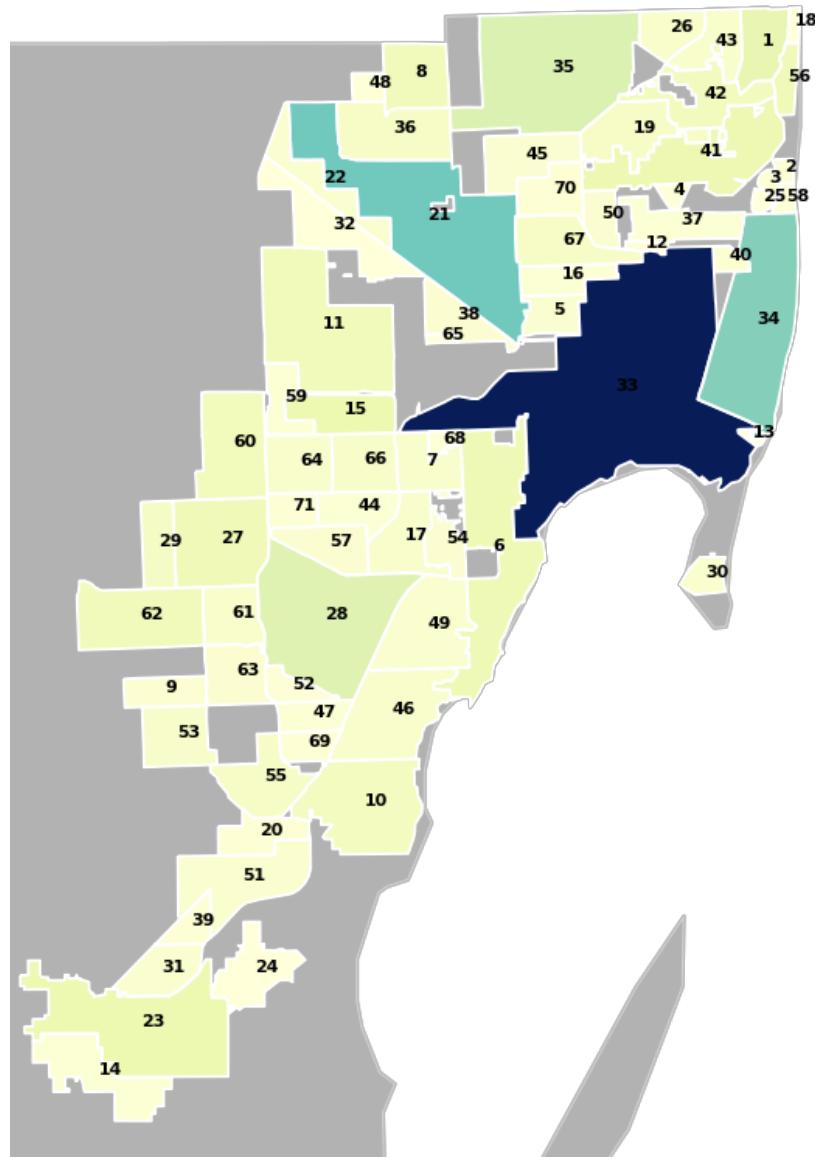


- 1 - Aventura
- 2 - Bal Harbour
- 3 - Bay Harbor Islands
- 4 - Biscayne Park
- 5 - Brownsville
- 6 - Coral Gables
- 7 - Coral Terrace
- 8 - Country Club
- 9 - Country Walk
- 10 - Cutler Bay
- 11 - Doral
- 12 - El Portal
- 13 - Fisher Island
- 14 - Florida City
- 15 - Fountainebleau
- 16 - Gladeview
- 17 - Glenvar Heights
- 18 - Golden Beach
- 19 - Golden Glades
- 20 - Goulds
- 21 - Hialeah
- 22 - Hialeah Gardens
- 23 - Homestead
- 24 - Homestead Base
- 25 - Indian Creek
- 26 - Ives Estates
- 27 - Kendall Lakes
- 28 - Kendall
- 29 - Kendall West
- 30 - Key Biscayne
- 31 - Leisure City
- 32 - Medley
- 33 - Miami
- 34 - Miami Beach
- 35 - Miami Gardens
- 36 - Miami Lakes
- 37 - Miami Shores
- 38 - Miami Springs
- 39 - Naranja
- 40 - North Bay Village
- 41 - North Miami
- 42 - North Miami Beach
- 43 - Ojus
- 44 - Olympia Heights
- 45 - Opa-locka
- 46 - Palmetto Bay
- 47 - Palmetto Estates
- 48 - Palm Springs North
- 49 - Pinecrest
- 50 - Pinewood
- 51 - Princeton
- 52 - Richmond Heights
- 53 - Richmond West
- 54 - South Miami
- 55 - South Miami Heights
- 56 - Sunny Isles Beach
- 57 - Sunset
- 58 - Surfside
- 59 - Sweetwater
- 60 - Tamiami
- 61 - The Crossings
- 62 - The Hammocks
- 63 - Three Lakes
- 64 - University Park
- 65 - Virginia Gardens
- 66 - Westchester
- 67 - West Little River
- 68 - West Miami
- 69 - West Perrine
- 70 - Westview
- 71 - Westwood Lakes

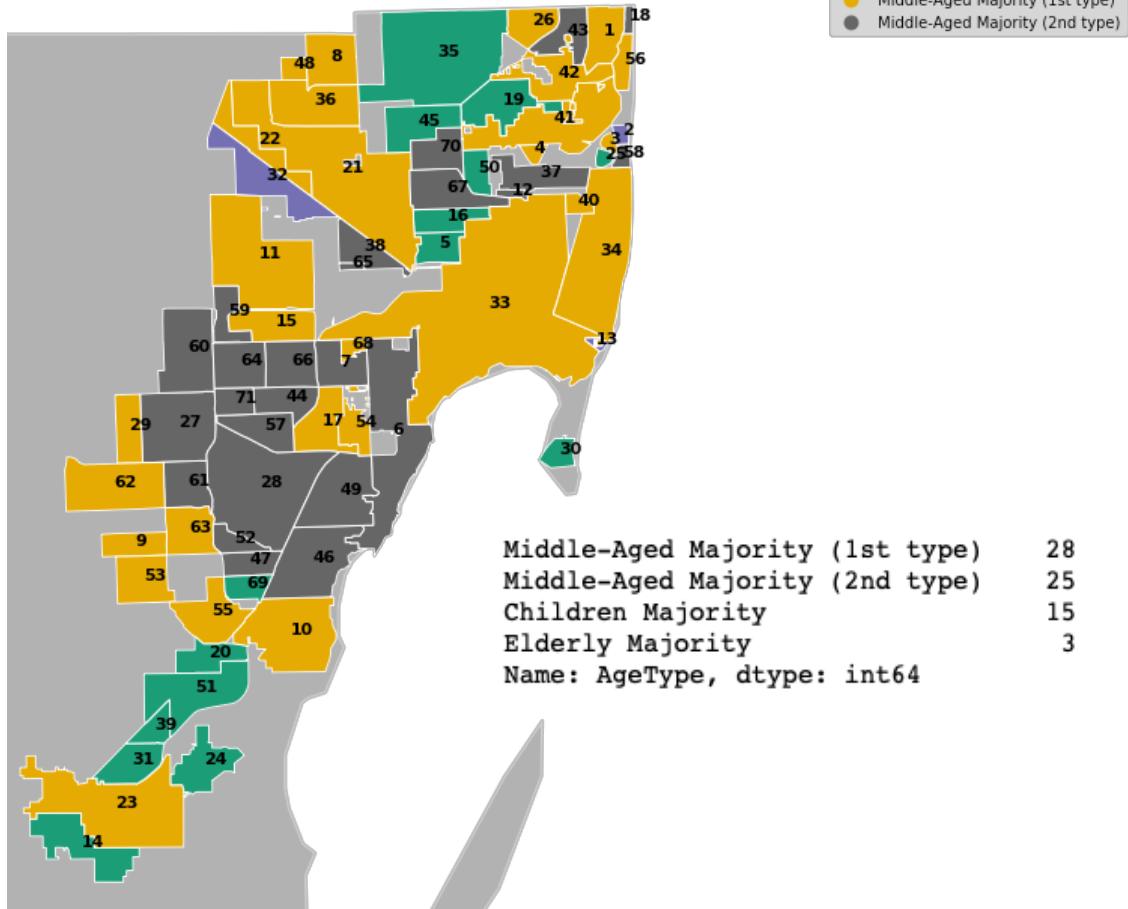
Map of Communities in Miami-Dade County
based on their population



Map of Communities in Miami-Dade County
based on their housing units

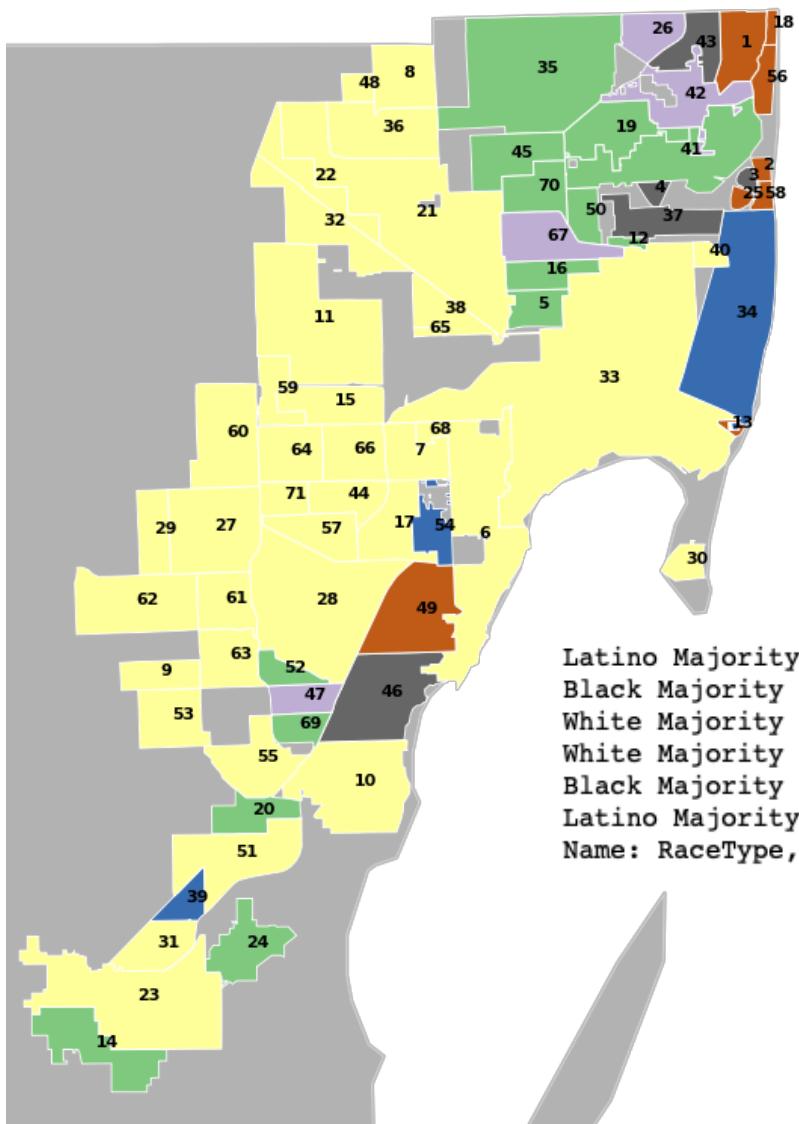


Map of Communities in Miami-Dade County
based on their age type



1 - Aventura	37 - Miami Shores
2 - Bal Harbour	38 - Miami Springs
3 - Bay Harbor Islands	39 - Naranja
4 - Biscayne Park	40 - North Bay Village
5 - Brownsville	41 - North Miami
6 - Coral Gables	42 - North Miami Beach
7 - Coral Terrace	43 - Ojus
8 - Country Club	44 - Olympia Heights
9 - Country Walk	45 - Opa-locka
10 - Cutler Bay	46 - Palmetto Bay
11 - Doral	47 - Palmetto Estates
12 - El Portal	48 - Palm Springs North
13 - Fisher Island	49 - Pinecrest
14 - Florida City	50 - Pinewood
15 - Fountainebleau	51 - Princeton
16 - Gladeview	52 - Richmond Heights
17 - Glenvar Heights	53 - Richmond West
18 - Golden Beach	54 - South Miami
19 - Golden Glades	55 - South Miami Heights
20 - Goulds	56 - Sunny Isles Beach
21 - Hialeah	57 - Sunset
22 - Hialeah Gardens	58 - Surfside
23 - Homestead	59 - Sweetwater
24 - Homestead Base	60 - Tamiami
25 - Indian Creek	61 - The Crossings
26 - Ives Estates	62 - The Hammocks
27 - Kendall Lakes	63 - Three Lakes
28 - Kendall	64 - University Park
29 - Kendall West	65 - Virginia Gardens
30 - Key Biscayne	66 - Westchester
31 - Leisure City	67 - West Little River
32 - Medley	68 - West Miami
33 - Miami	69 - West Perrine
34 - Miami Beach	70 - Westview
35 - Miami Gardens	71 - Westwood Lakes
36 - Miami Lakes	

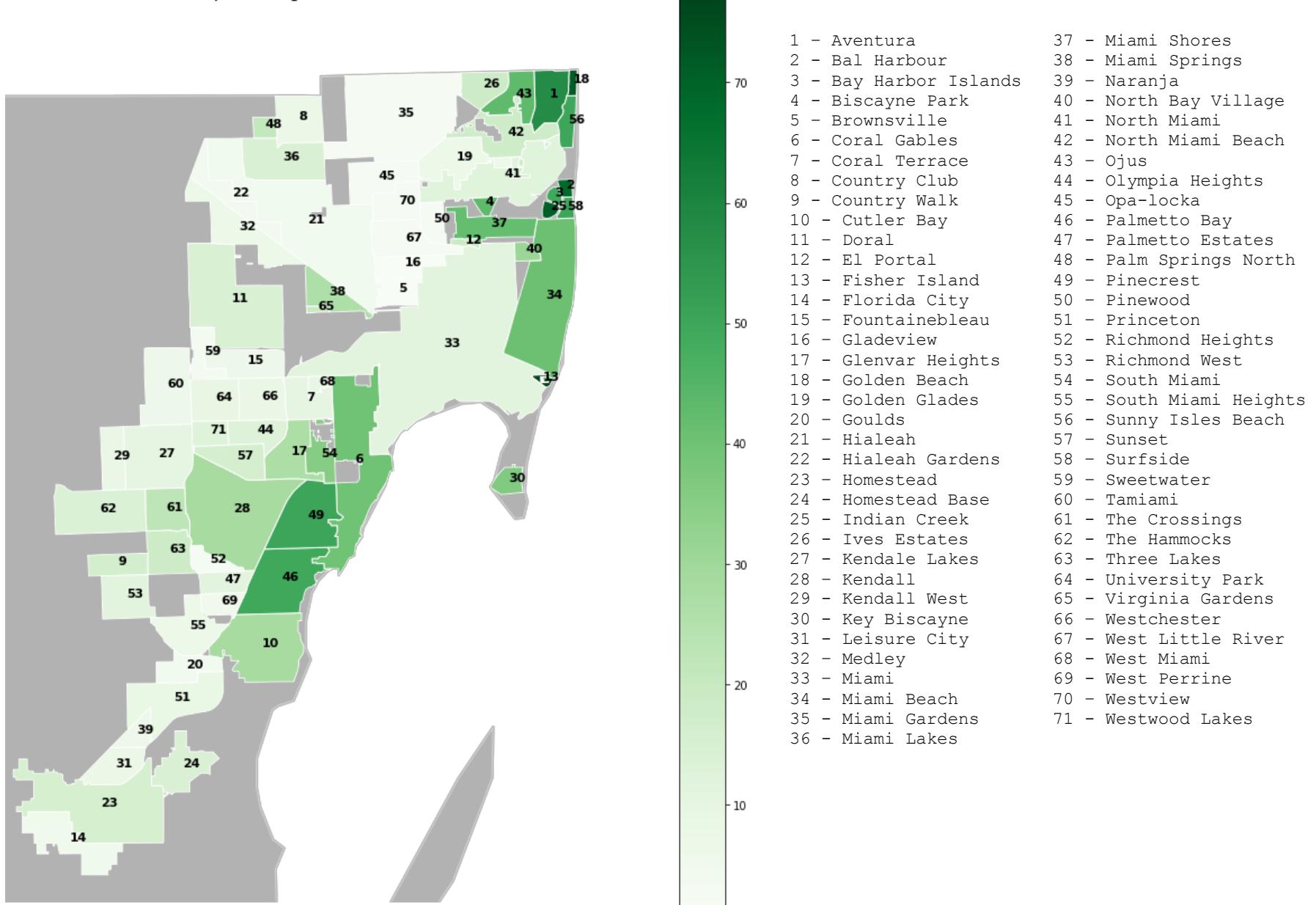
Map of Communities in Miami-Dade County
based on their major race



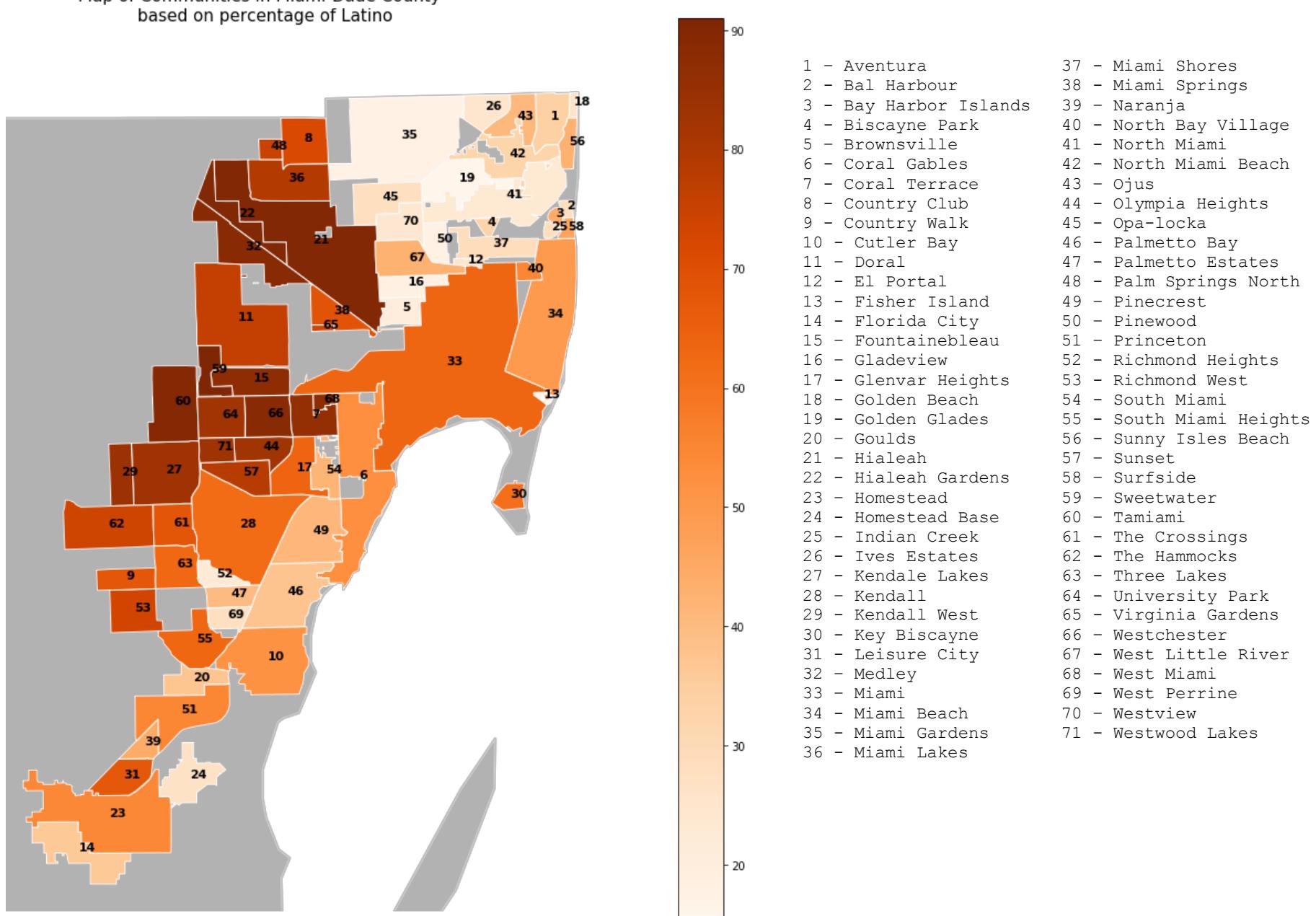
- Black Majority (strong)
- Black Majority (weak)
- Latino Majority (strong)
- Latino Majority (weak)
- White Majority (strong)
- White Majority (weak)

- 1 - Aventura
- 2 - Bal Harbour
- 3 - Bay Harbor Islands
- 4 - Biscayne Park
- 5 - Brownsburg
- 6 - Coral Gables
- 7 - Coral Terrace
- 8 - Country Club
- 9 - Country Walk
- 10 - Cutler Bay
- 11 - Doral
- 12 - El Portal
- 13 - Fisher Island
- 14 - Florida City
- 15 - Fountainebleau
- 16 - Gladeview
- 17 - Glenvar Heights
- 18 - Golden Beach
- 19 - Golden Glades
- 20 - Goulds
- 21 - Hialeah
- 22 - Hialeah Gardens
- 23 - Homestead
- 24 - Homestead Base
- 25 - Indian Creek
- 26 - Ives Estates
- 27 - Kendale Lakes
- 28 - Kendall
- 29 - Kendall West
- 30 - Key Biscayne
- 31 - Leisure City
- 32 - Medley
- 33 - Miami
- 34 - Miami Beach
- 35 - Miami Gardens
- 36 - Miami Lakes
- 37 - Miami Shores
- 38 - Miami Springs
- 39 - Naranja
- 40 - North Bay Village
- 41 - North Miami
- 42 - North Miami Beach
- 43 - Ojus
- 44 - Olympia Heights
- 45 - Opa-locka
- 46 - Palmetto Bay
- 47 - Palmetto Estates
- 48 - Palm Springs North
- 49 - Pinecrest
- 50 - Pinewood
- 51 - Princeton
- 52 - Richmond Heights
- 53 - Richmond West
- 54 - South Miami
- 55 - South Miami Heights
- 56 - Sunny Isles Beach
- 57 - Sunset
- 58 - Surfside
- 59 - Sweetwater
- 60 - Tamiami
- 61 - The Crossings
- 62 - The Hammocks
- 63 - Three Lakes
- 64 - University Park
- 65 - Virginia Gardens
- 66 - Westchester
- 67 - West Little River
- 68 - West Miami
- 69 - West Perrine
- 70 - Westview
- 71 - Westwood Lakes

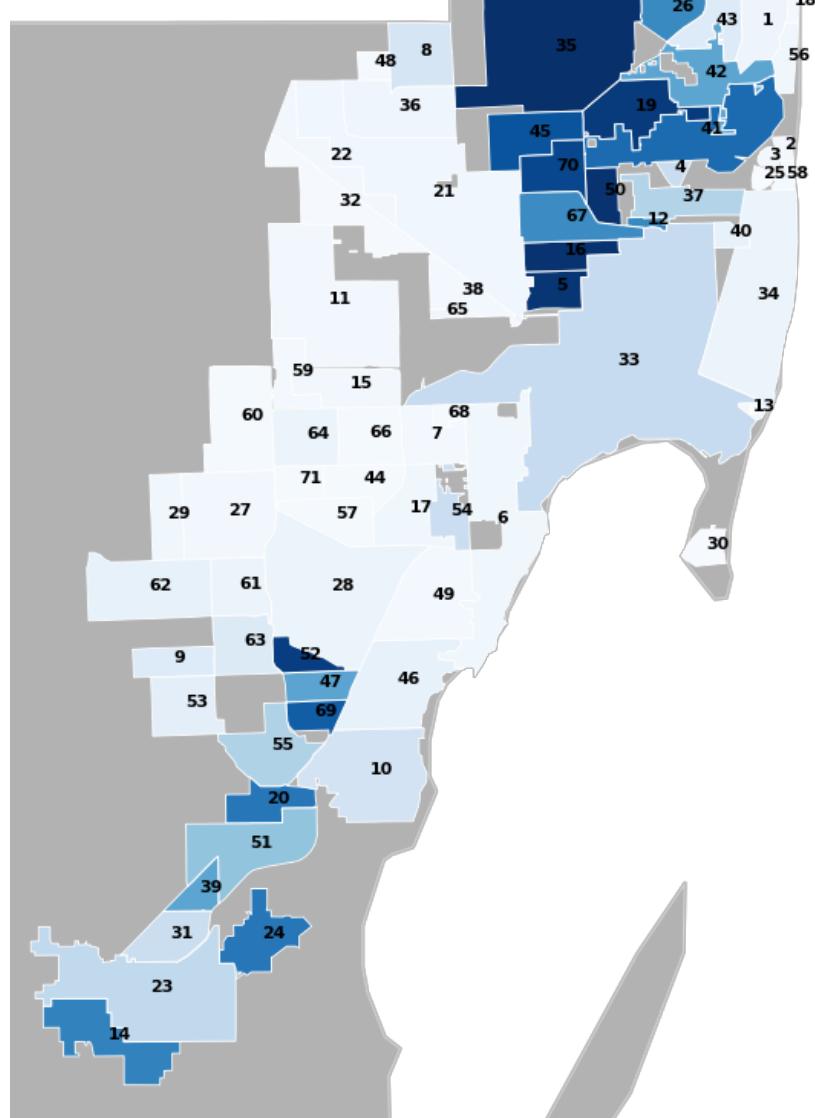
Map of Communities in Miami-Dade County
based on percentage of White



Map of Communities in Miami-Dade County
based on percentage of Latino



Map of Communities in Miami-Dade County
based on percentage of Black



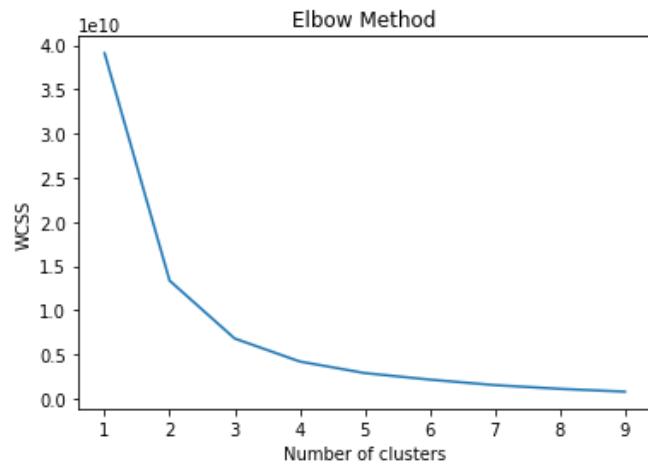
3.3.Determination of the real type of community (K-means Clustering)

Based on the literature review (section 2.1) and the data analysis and visualization (section 3.2) we choose the next variables: Area, Population, Housing Units, <18%, 18-24%, 25-44%, 45-64%, >=65%, White%, Black%, Latino%.

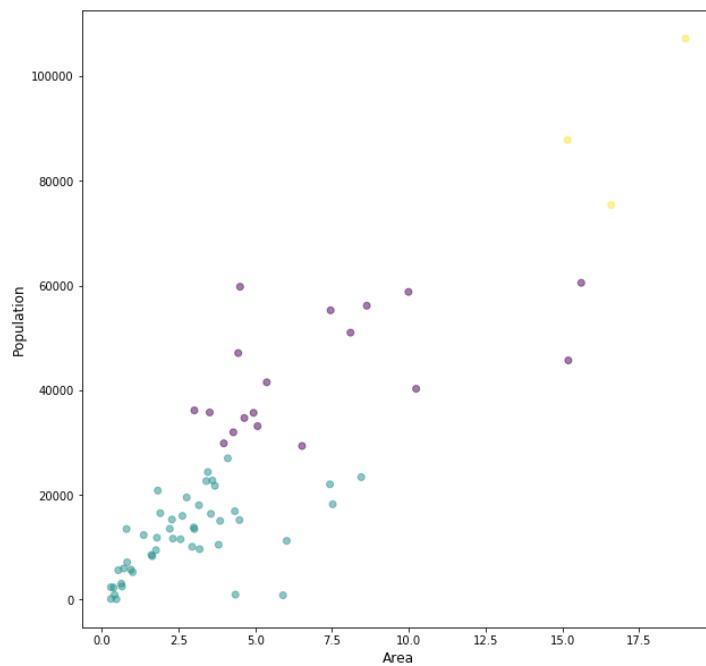
The choice of K-means Clustering is based on the next factors: 1) we have structured data; 2) we need to find similarity and dissimilarity with other elements based on demographic data; 3) we need unsupervised method; 4) we need non-overlapping clusters in the scope of demographic data.

Due to our determination of outliers we exclude Miami, Coral Gables and Hialeah from clustering and assign them ‘City’ type as fact..

Elbow Method shows us that 3 is the optimal number of clusters.



The distribution of communities on Area vs Population plot becomes clearer.



After clustering instead of the presumption that it's impossible to determine the label the clusters we create our code that can decode cluster numbers and assign to yellow point label "City", purple points - "Town" and blue points – "Village". We define DataFrame with the new Type of Communities in comparison with previous Type of Communities. It will be discussed deeper in the results section (section 4.1).

3.4. Determination of merged communities (Hierarchical Clustering)

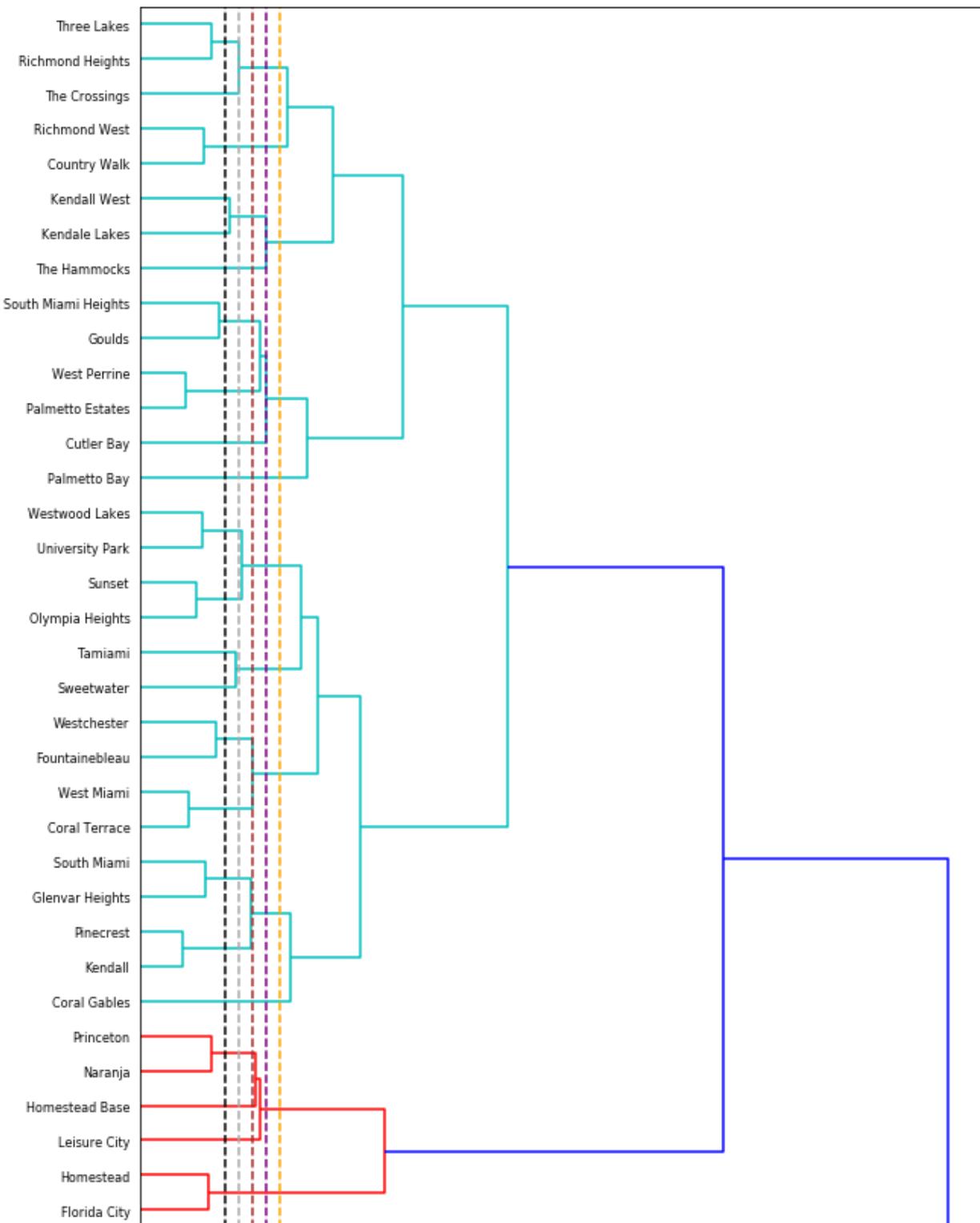
Using Geolocator we collect data about latitude and longitude of every community in Miami-Dade County and add them to our DataFrame.

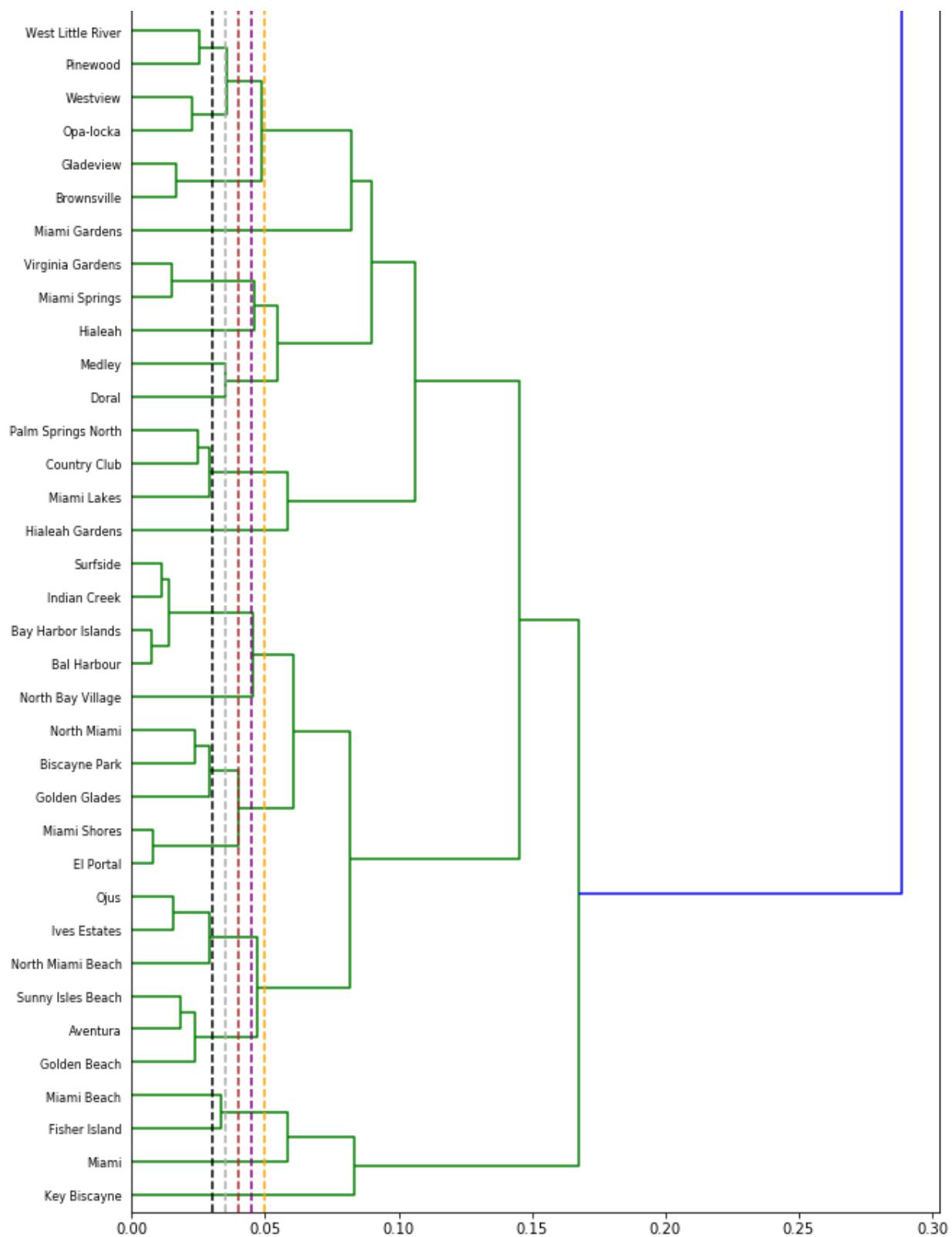
	GEOID	Community	geometry	Latitude	Longitude	Type	ClusterType	
0	1590000US1208602681	Aventura	POLYGON ((-80.15060 25.93773, 25.953257 -80.136646 -80.14915 25.942...))	25.93773	25.953257	-80.136646	City	Town
1	1590000US1208603275	Bal Harbour	POLYGON ((-80.13102 25.89913, 25.893650 -80.126442 -80.12457 25.899...))	25.89913	25.893650	-80.126442	Village	Village
2	1590000US1208603975	Bay Harbor Islands	POLYGON ((-80.13920 25.88877, 25.887595 -80.131156 -80.13764 25.891...))	25.88877	25.887595	-80.131156	Town	Village

The choice of Hierarchical clustering is based on the next factors: 1) we have structured data; 2) we need to find similarity and dissimilarity with other elements based on demographic data; 3) we need unsupervised method; 4) we need different possibilities to determine the number of merges clusters.

We use only one variable for this clustering – the distance between latitude and longitude of each community with others. Hierarchical Clustering gives us the following Dendrogram.

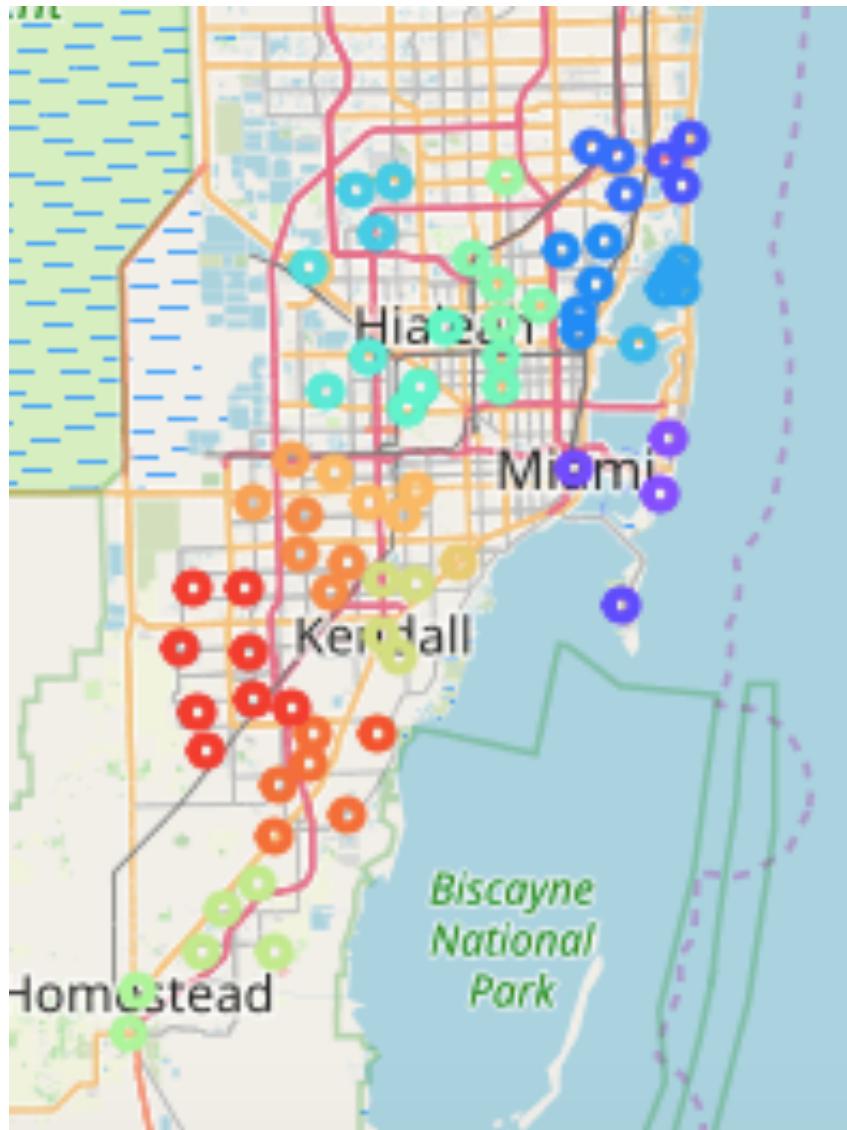
Dendrogram of communities merging
based on the distance





Due to our decision to not merge Miami, Hialeah and Coral Gables with any other communities (this communities are already outliers in the EDA) we choose the distance 0.045 It gives us 28 clusters – new merged communities.

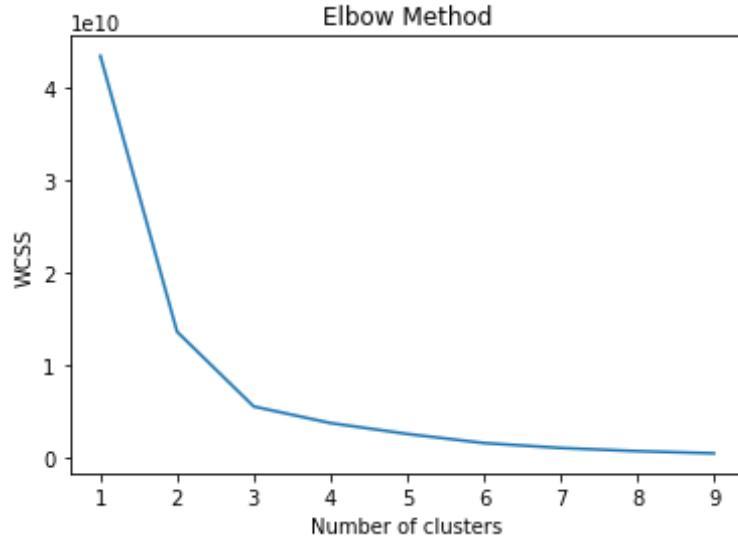
We have tried to add this label clusters to the map, but it doesn't give us a good representation of changes.



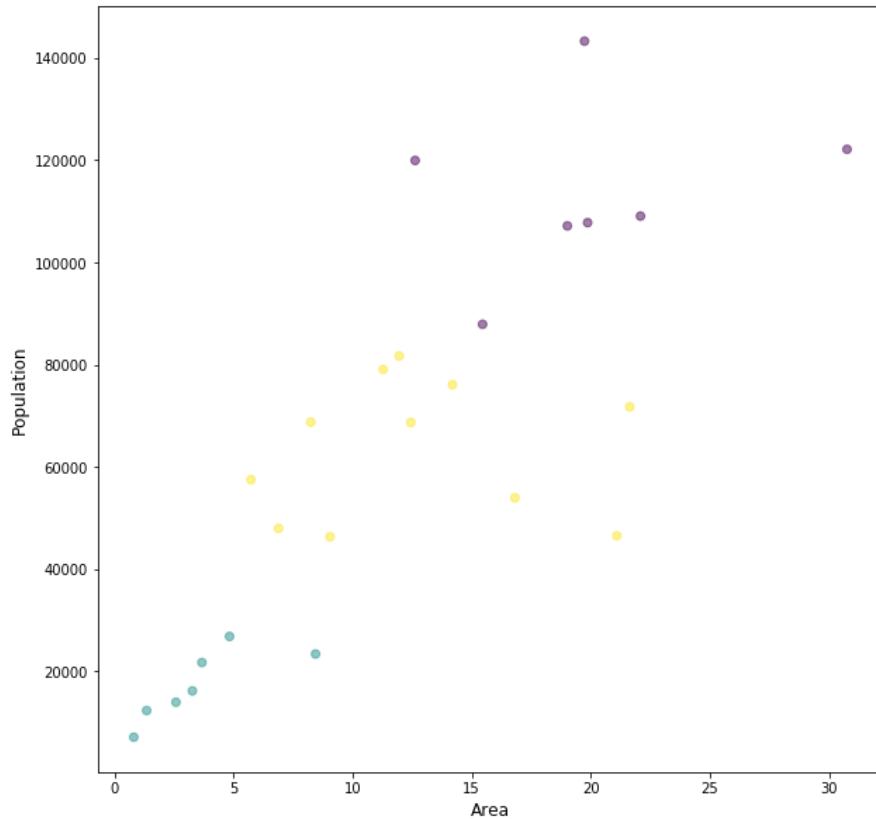
By using GeoPandas we join ('.dissolve' function) geometry of merged communities and plot new communities. We also aggregate their names, and main indicators. Moreover, we calculate percentage of races, ages and so on in merged communities. More results are in the section 4.2.

3.5. Determination of the merged community type (K-Means Clustering)

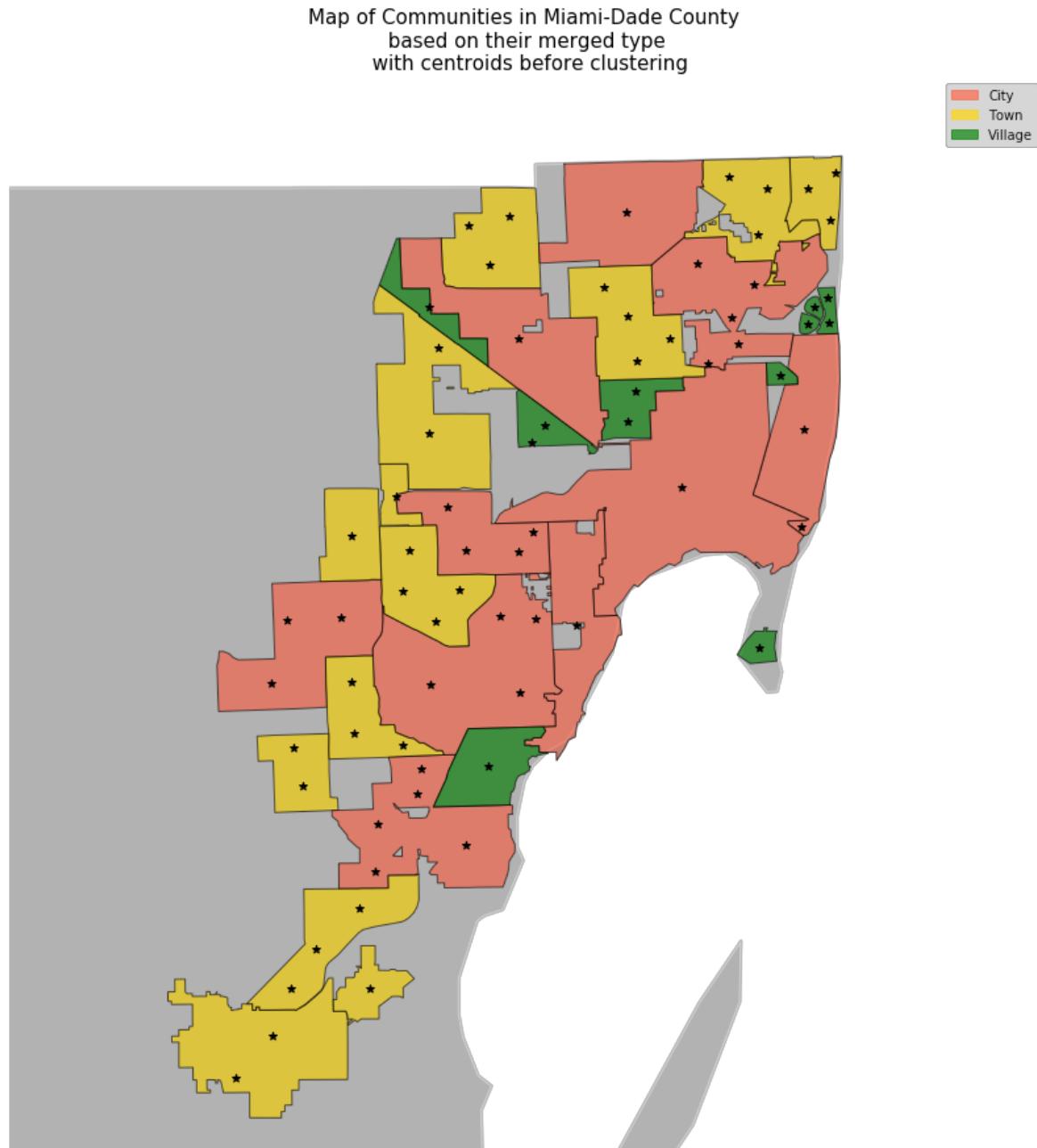
The type of merged community is undefined, so we decide to use K-clustering one more time. The reasons to choose this model are the same as in the section 3.3. We use the same variables for clustering and also exclude our 3 cities: Miami, Coral Gables and Hialeah. Elbow method again shows 3 clusters.



The distribution of communities on Area vs Population plot becomes clearer too.



By using the code to label clusters we determine their types and plot them with the centroids of previous community's and CDP's areas.



It seems that this kind of illustration of changes is better than the map on the page 32. So, on the scheme-map you can see the number of communities (points) merged in the new ones. The results are presented in the section 4.2.

3.6. Determination of grocery and pharmacy store assignment to the new merged community

By using Foursquare API we find the representation of the most important grocery (Publix, WholeFoods, Aldi, Trader Joe's, Winn-Dixie, Presidente Supermarket, Bravo Supermarket) and pharmacy (CVS Pharmacy, Walgreens, Navarro Discount Pharmacy) chains.

Longitude and latitude of each store is preprocessed in GeoPandas to create Geo-object Point and then by applying the search function (.within() function) we determine that the store is in the borders of a certain merged community. The results are presented in section 4.3.

4. Results

The results of machine learning are presented in this section of our report. This section consists of 3 parts. In the section 4.1 we show the results of K-means clustering to determine the type of community in the pre-defined borders with comparison to the previous results.

The section 4.2 deals with Hierarchical Clustering to change the borders of communities by creating the new ones – merged communities. It gives the comparison with the existing distribution of some parameters.

Section 4.3 shows which stores should be assigned to each merged community.

4.1. Changing the type of Community in the pre-defined borders

The using of K-means clustering is changing a look of community type distribution. The total number of communities doesn't change (71 community) but CDPs disappear, the number of cities decreases from 19 to 6, and the number of towns and villages increases from 6 to 18 and from 9 to 47 respectively.

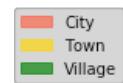
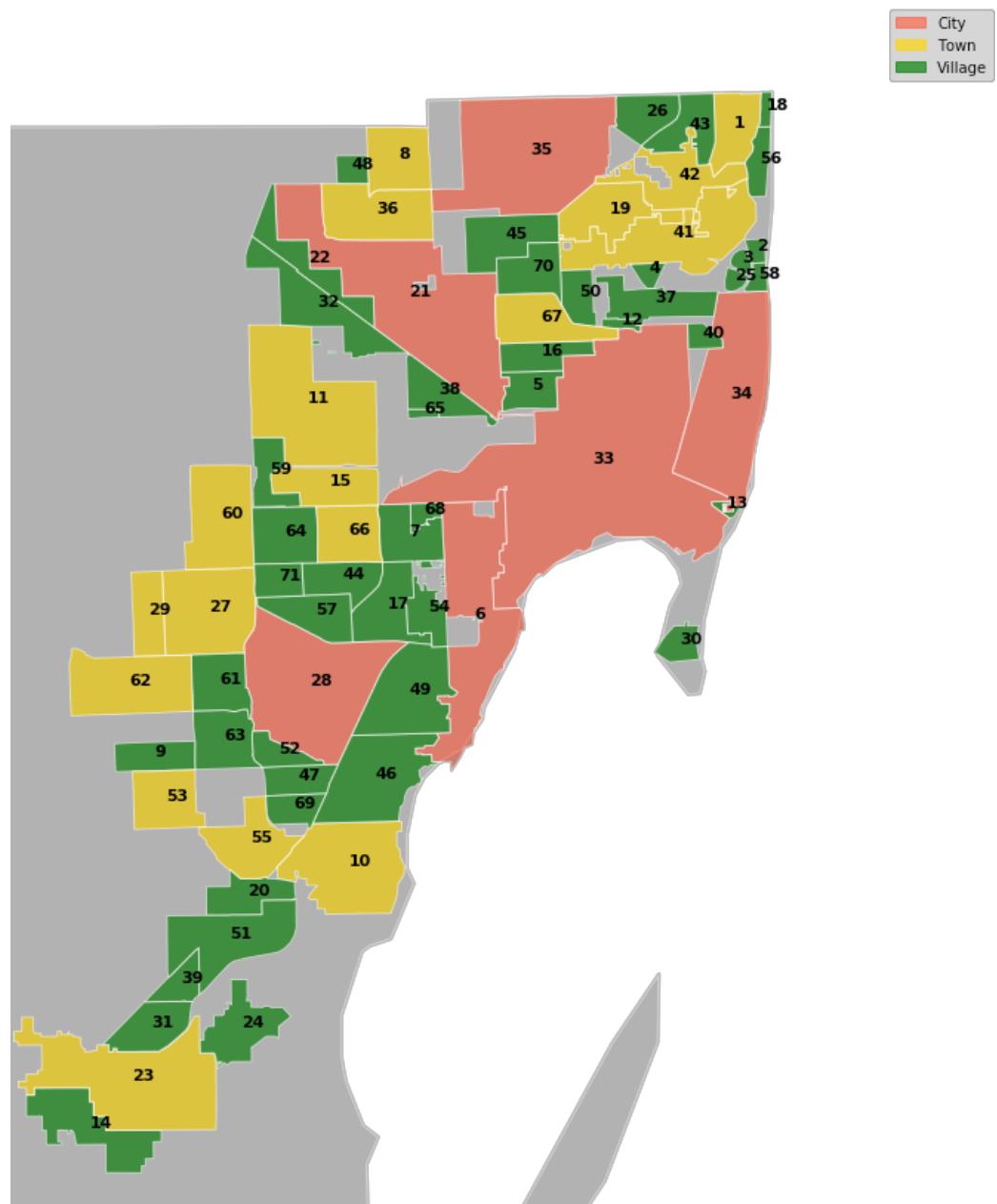
Some communities keep their previous status. Coral Gables (6), Hialeah (21), Miami (33), Miami Beach (34) and Miami Gardens (35) remain cities. It's 5 of 6 cities, including 3 predefined as cities during K-means Clustering. No one previous town or village receives the status of city. However, Kendall (28), the previous CDP, becomes the 6th city in Miami-Dade County.

Cutler Bay (10) and Miami Lakes (36) preserve the status of town (2 of 18 now). From city to the level of town the following transfers: Aventura (1), Doral (11), Homestead (23), North Miami (41), North Miami Beach (42). It's 5 new towns of 18 and 5 old cities of 19. No one previously defined village has received the status of town, but the rest (11 of 18 new towns) are formed from CDPs: Country Club (8), Fountainebleau (15), Golden Glades (19), Kendall Lakes (27), Kendall West (29), Richmond West (53), South Miami Heights (55), Tamiami (60), The Hammocks (62), Westchester (66), West Little River (67).

Another 4 of 18 previous towns become villages are Bay Harbor Islands (3), Golden Beach (18), Medley (32), Surfside (58). Another 9 of 19 previous cities have also become villages: Florida City (14), Hialeah Gardens, Miami Springs, North Bay Village (40), Opa-locka (45), South Miami (54), Sunny Isles Beach (56), Sweetwater (59), West Miami (68). All of 9 previously existing villages keep their status: Bal Harbour (2), Biscayne Park (4), El Portal (12), Indian Creek (25), Key Biscayne (30), Miami Shores (37), Palmetto Bay (46), Pinecrest (49), Virginia Gardens (65). The rest of CDPs (25 of 37 previous) excluding Kendall (28) becoming city and 11 CDPs forming towns is labeled as village.

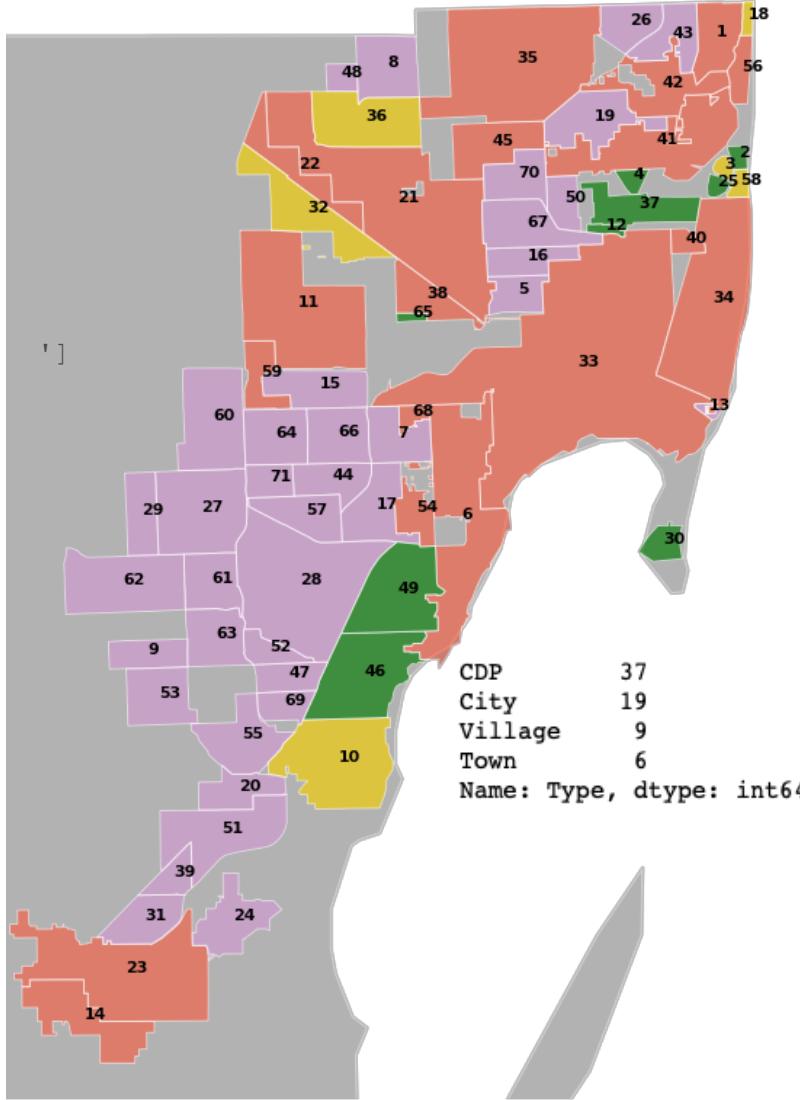
In the section 3.2 (p. 19) you can find the present distribution of communities based on their type. Now we are presenting the new types of communities and the comparison with the present distribution. Moreover, we can compare the distribution (age and race) before and after the changes.

Map of Communities in Miami-Dade County
based on their new type

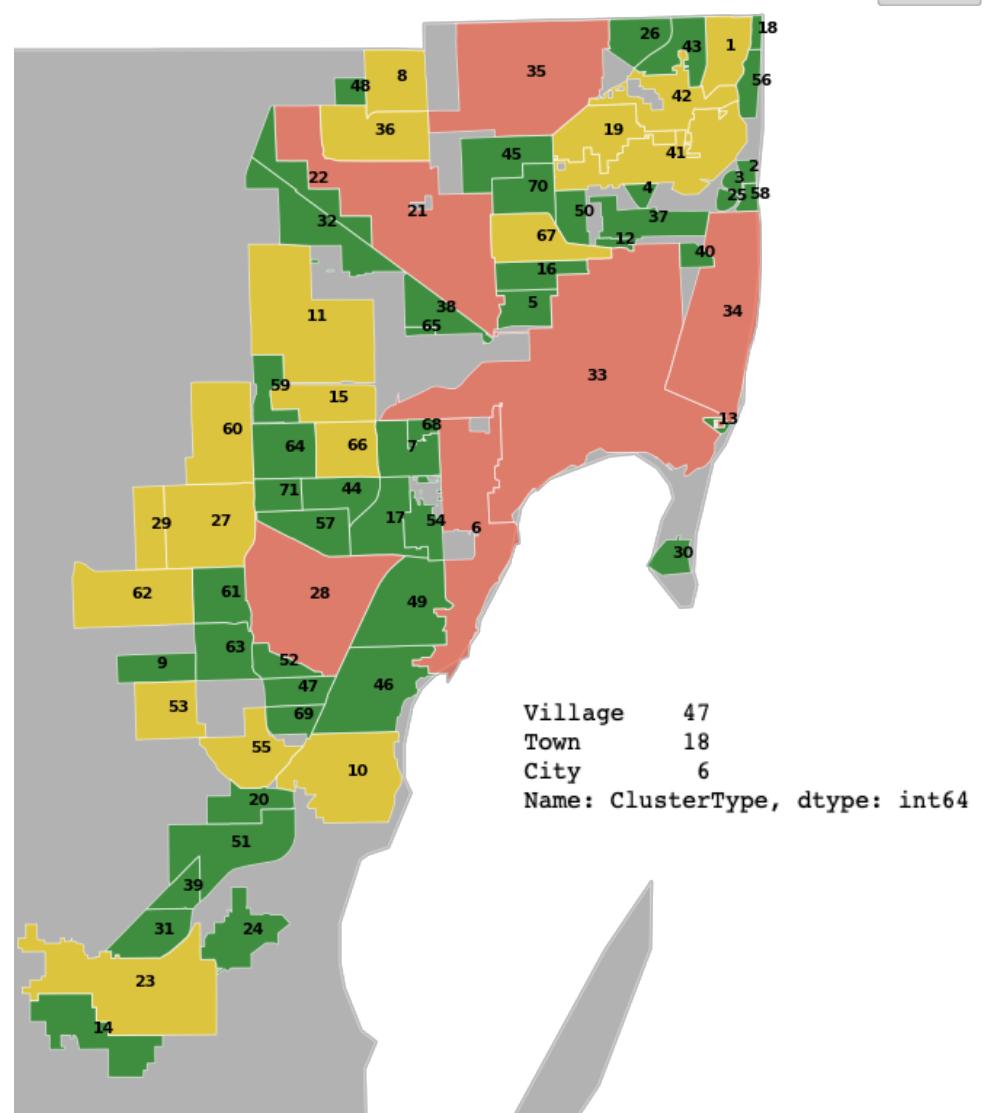


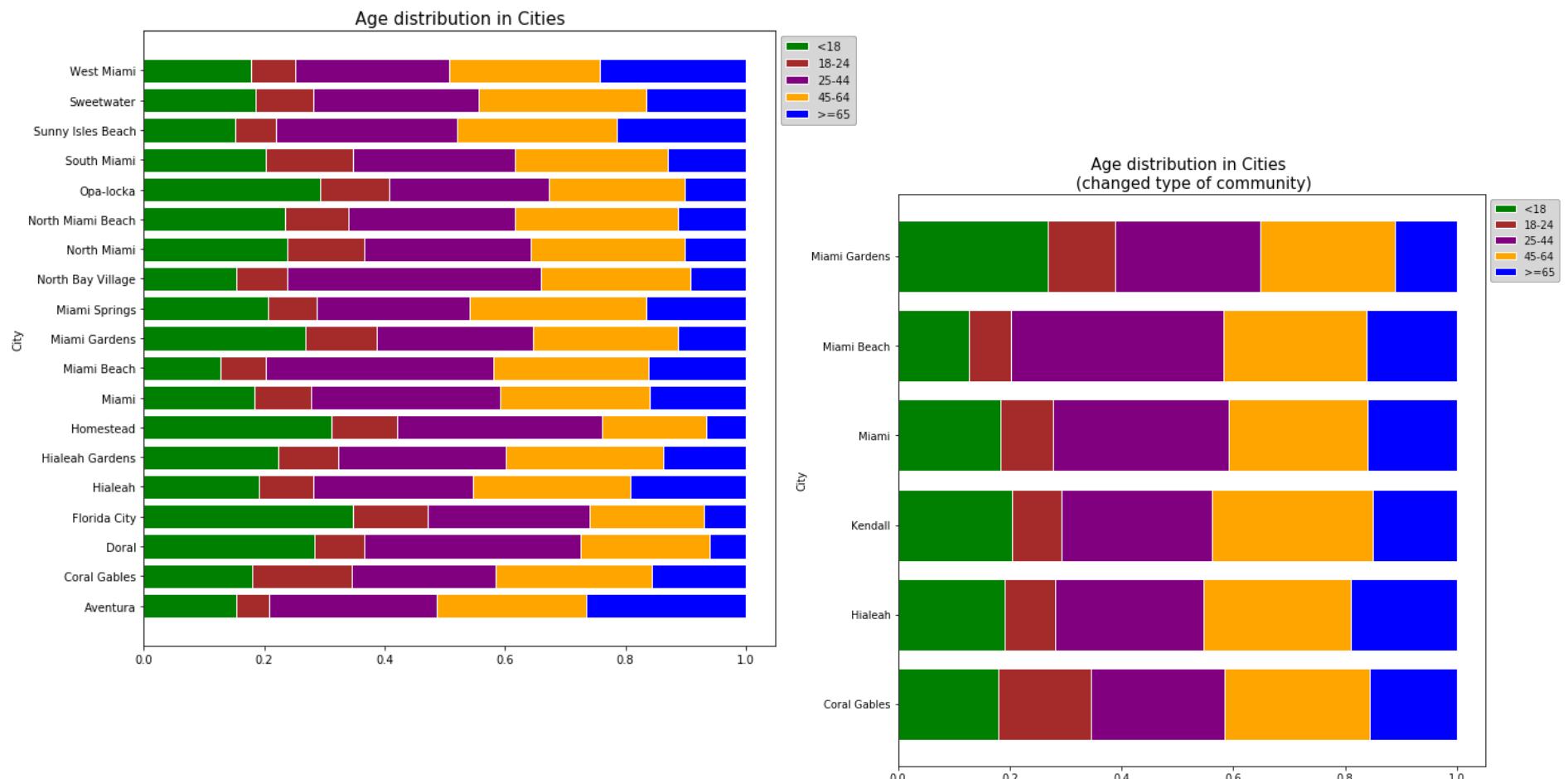
- | | |
|------------------------|--------------------------|
| 1 - Aventura | 37 - Miami Shores |
| 2 - Bal Harbour | 38 - Miami Springs |
| 3 - Bay Harbor Islands | 39 - Naranja |
| 4 - Biscayne Park | 40 - North Bay Village |
| 5 - Brownsville | 41 - North Miami |
| 6 - Coral Gables | 42 - North Miami Beach |
| 7 - Coral Terrace | 43 - Ojus |
| 8 - Country Club | 44 - Olympia Heights |
| 9 - Country Walk | 45 - Opa-locka |
| 10 - Cutler Bay | 46 - Palmetto Bay |
| 11 - Doral | 47 - Palmetto Estates |
| 12 - El Portal | 48 - Palm Springs North |
| 13 - Fisher Island | 49 - Pinecrest |
| 14 - Florida City | 50 - Pinewood |
| 15 - Fountainebleau | 51 - Princeton |
| 16 - Gladeview | 52 - Richmond Heights |
| 17 - Glenvar Heights | 53 - Richmond West |
| 18 - Golden Beach | 54 - South Miami |
| 19 - Golden Glades | 55 - South Miami Heights |
| 20 - Goulds | 56 - Sunny Isles Beach |
| 21 - Hialeah | 57 - Sunset |
| 22 - Hialeah Gardens | 58 - Surfside |
| 23 - Homestead | 59 - Sweetwater |
| 24 - Homestead Base | 60 - Tamiami |
| 25 - Indian Creek | 61 - The Crossings |
| 26 - Ives Estates | 62 - The Hammocks |
| 27 - Kendale Lakes | 63 - Three Lakes |
| 28 - Kendall | 64 - University Park |
| 29 - Kendall West | 65 - Virginia Gardens |
| 30 - Key Biscayne | 66 - Westchester |
| 31 - Leisure City | 67 - West Little River |
| 32 - Medley | 68 - West Miami |
| 33 - Miami | 69 - West Perrine |
| 34 - Miami Beach | 70 - Westview |
| 35 - Miami Gardens | - Westwood Lakes |
| 36 - Miami Lakes | |

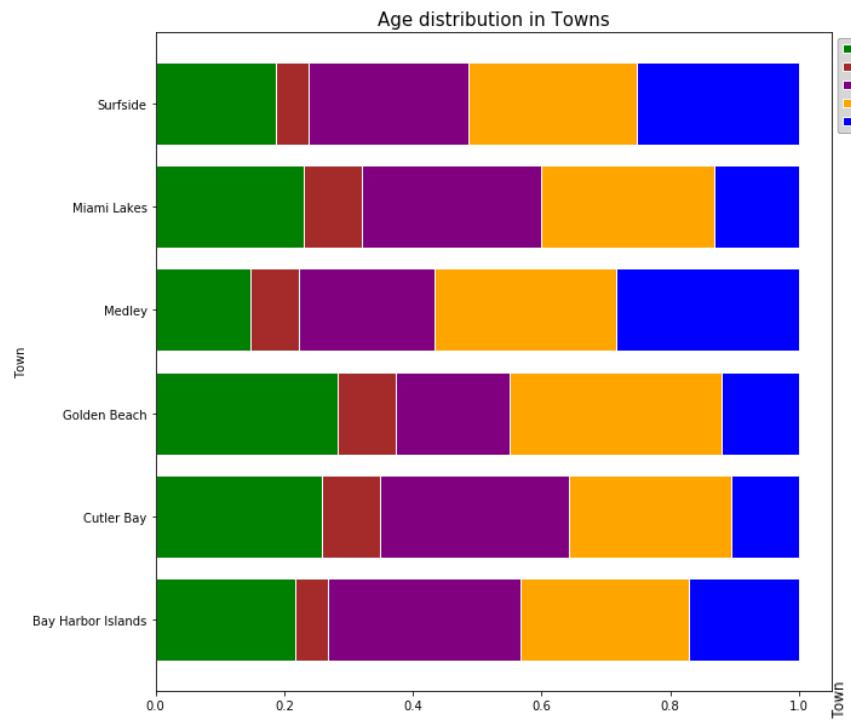
Map of Communities in Miami-Dade County
based on their type



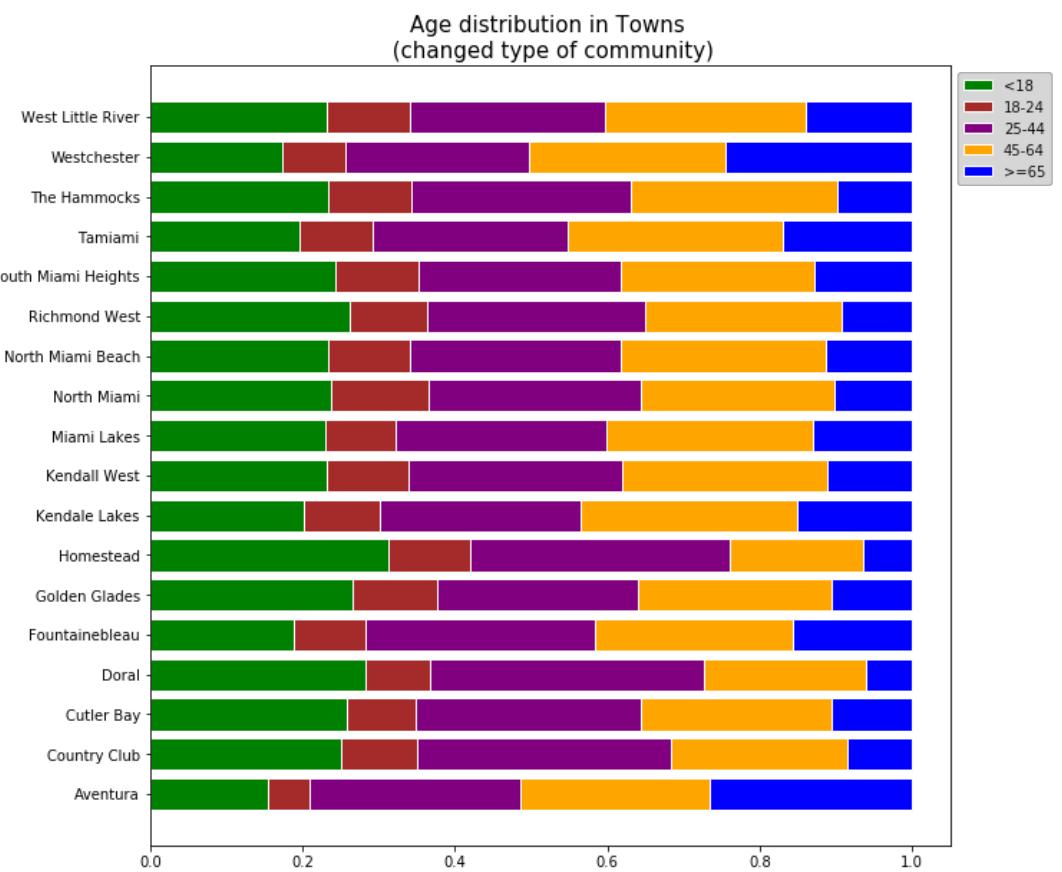
Map of Communities in Miami-Dade County
based on their new type



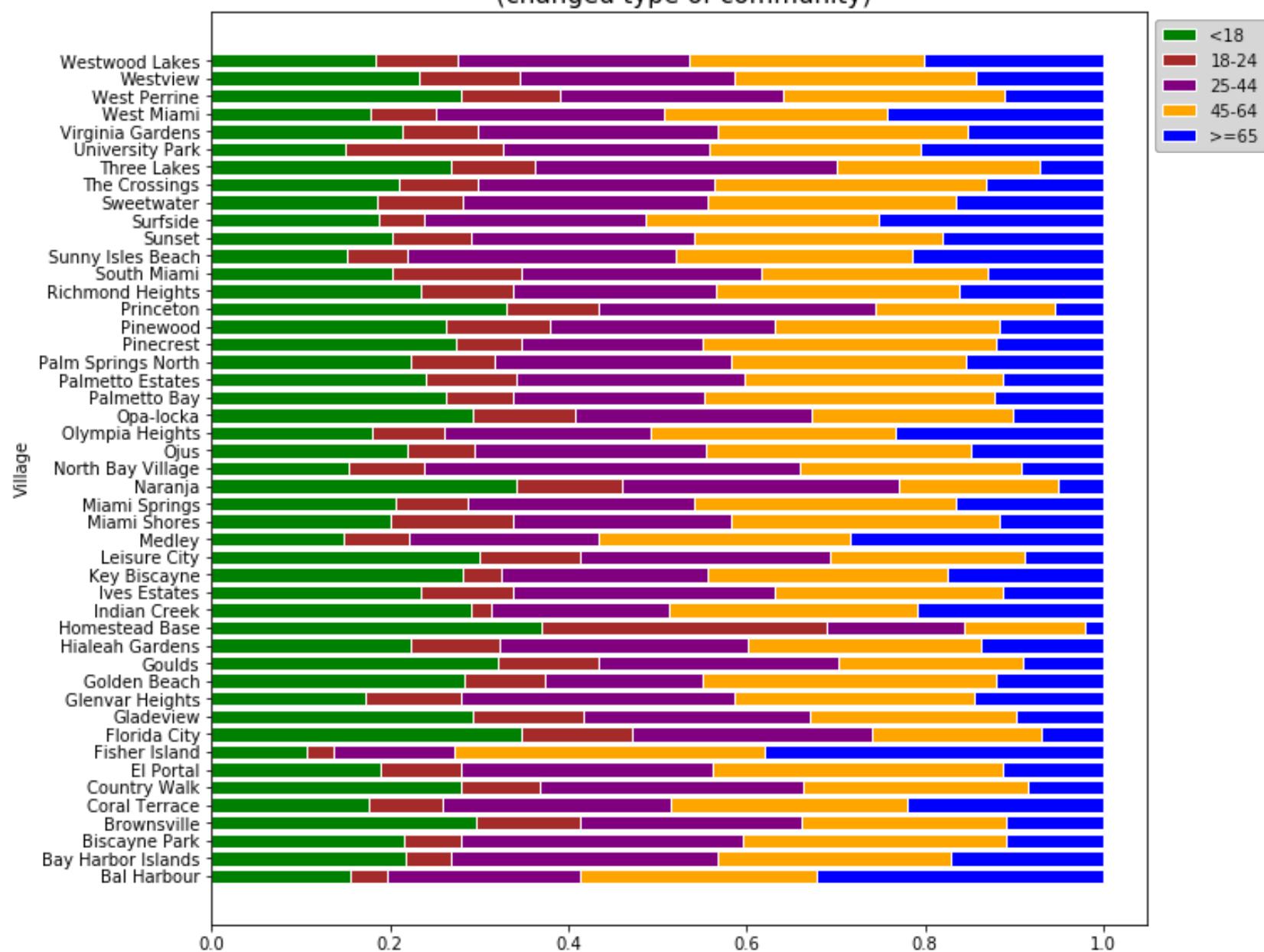


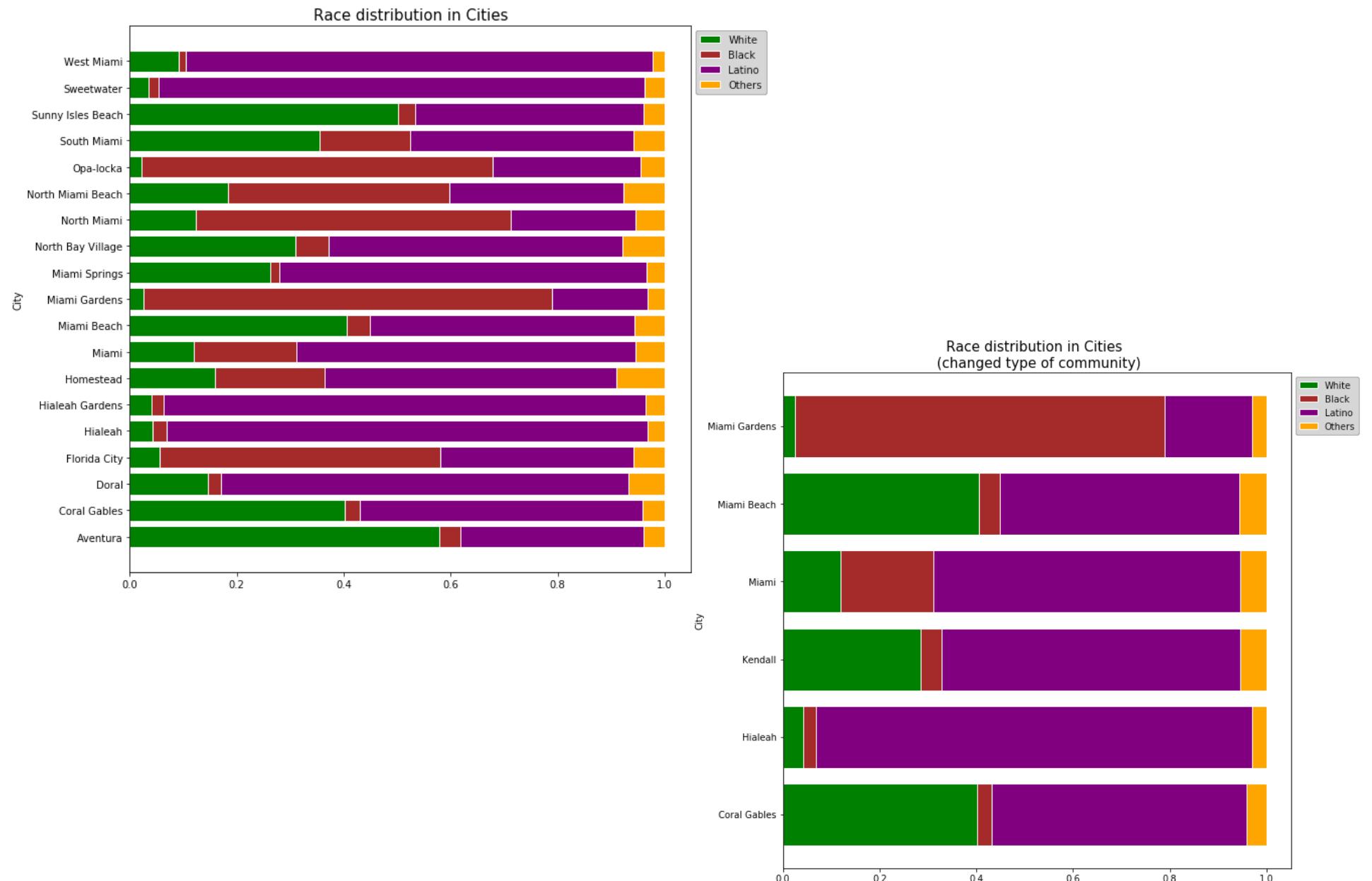


<18
18-24
25-44
45-64
≥65



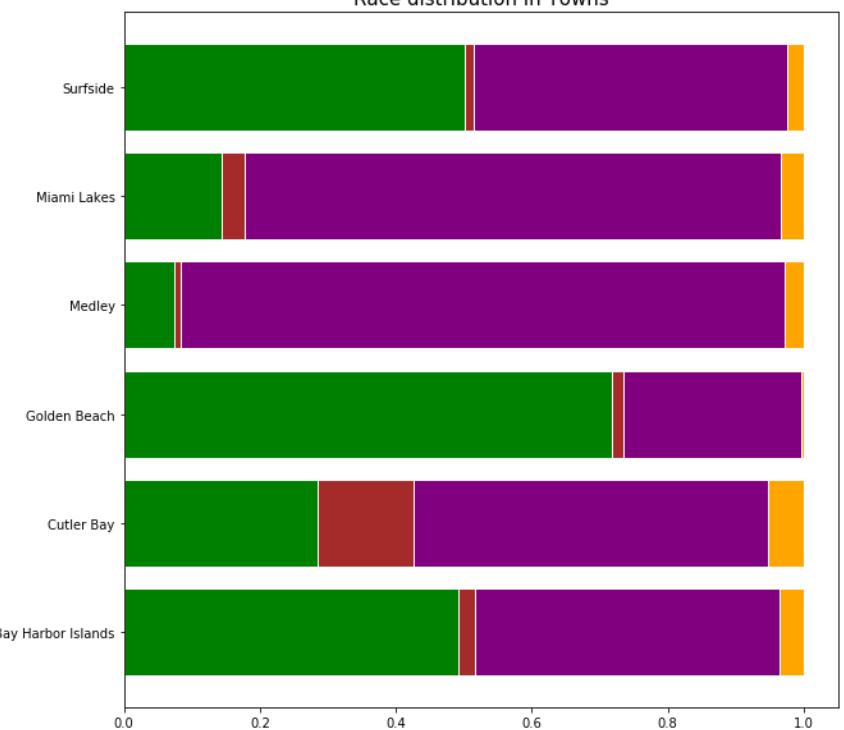
Age distribution in Villages (changed type of community)



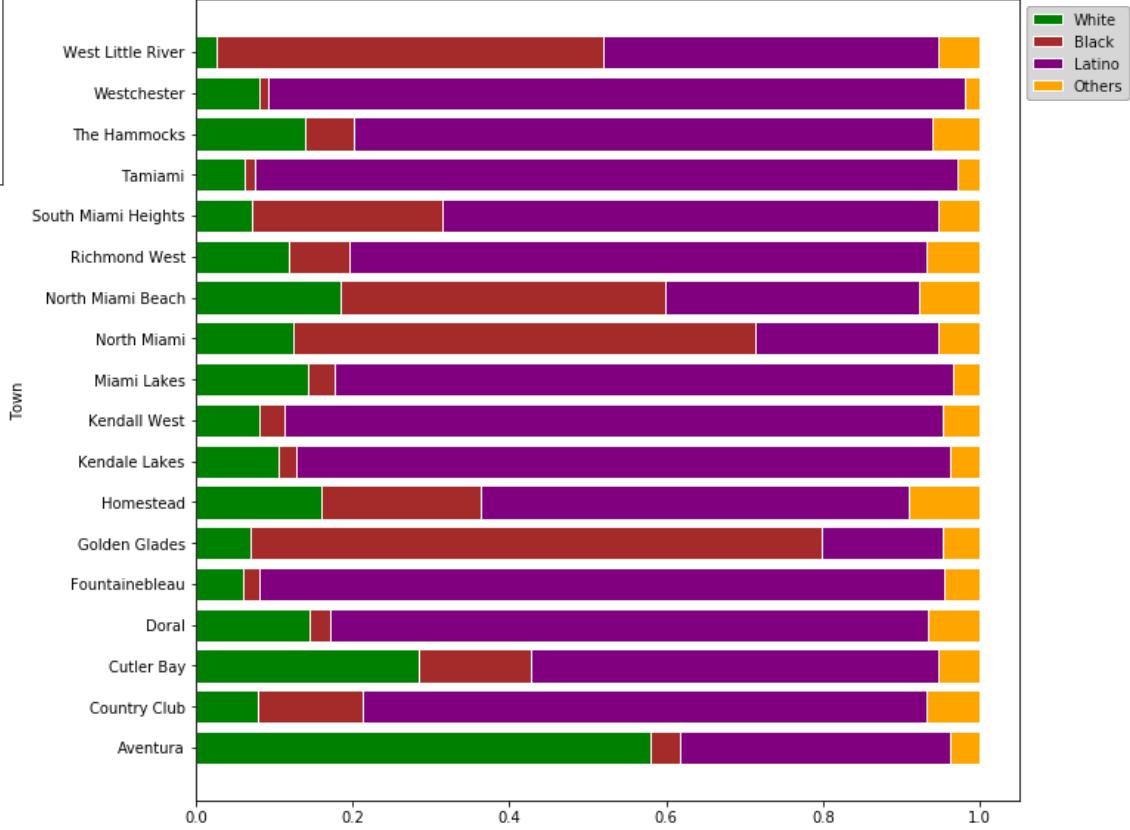


Race distribution in Towns

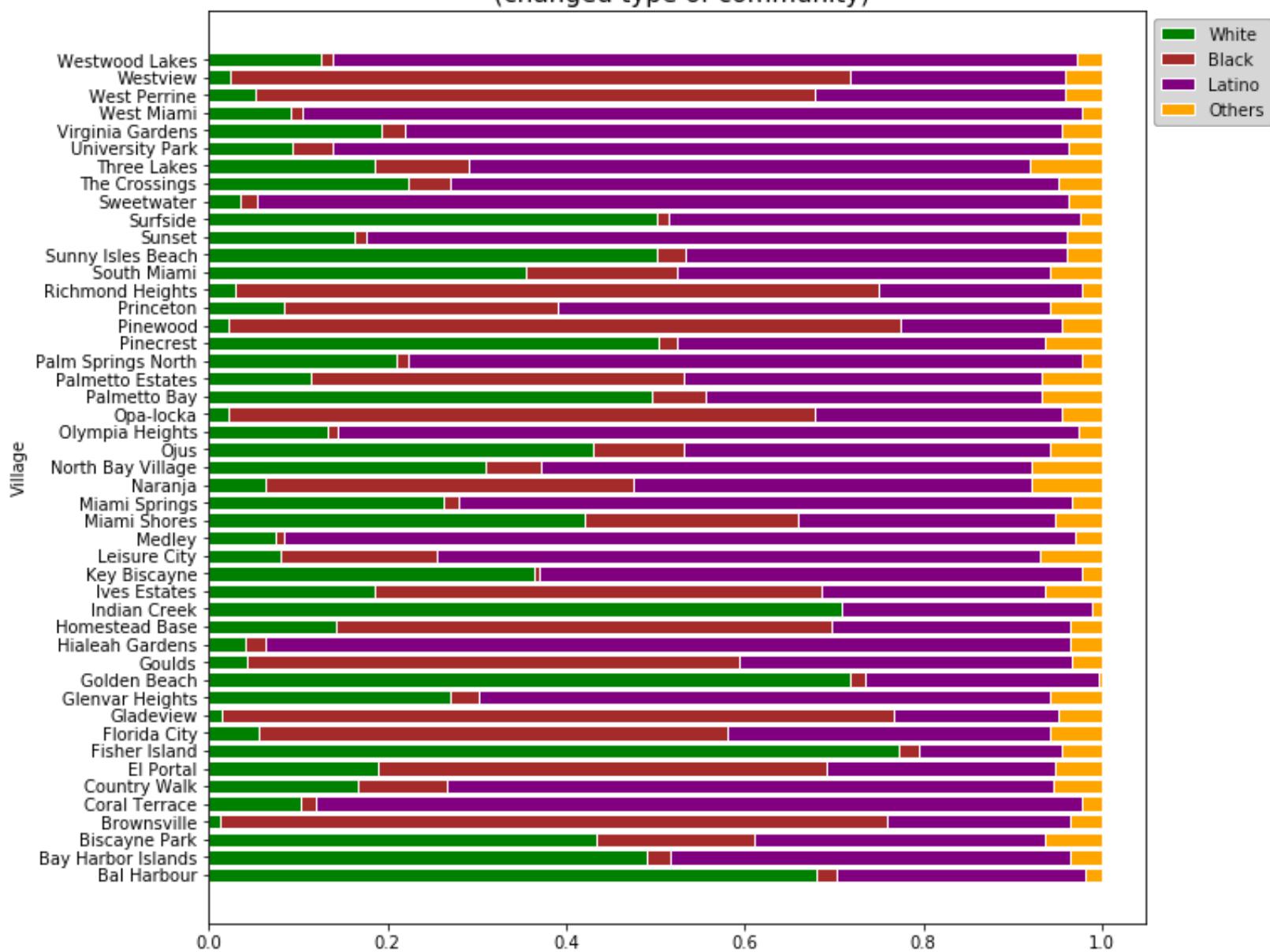
Town



Race distribution in Towns
(changed type of community)



Race distribution in Villages (changed type of community)



4.2. Changing the borders of Communities and determining the type of merged Community

Using Hierarchical Clustering we change the distribution of communities leaving the question about their types. Applying K-means Clustering we define the type of each merged community. The total number of communities in Miami-Dade County decreases from 71 to 28, including 10 cities, 11 towns and 7 villages.

Among 10 new formed **cities** there are 3 predefined during Hierarchical Clustering: Coral Gables (20), Hialeah (13) and Miami (2). Moreover, Miami Gardens (16) takes his title of city without any merge. Miami Beach which is a city at the present time and could save its status during type changes (section 4.1) will be joined by Fisher Island and they will form city “Fisher Island - Miami Beach” (1).

North Miami (previously city, risked as town) being joined by 3 villages and 1 CDP (Miami Shores, risked town) will remain city as “Biscayne Park - El Portal - Golden Glades - Miami Shores - North Miami” (6). The same pattern will have place with:

- 1) West Miami (previously city, risked as village) being joined by 3 CDPs (risked as 2 villages and 1 town) can save his city status as “Coral Terrace - Fountainebleau - Westchester - West Miami” (21);
- 2) South Miami (previously city, risked as village) being joined by Kendal (CDP, risked as city), 1 village and 1 more CDP (risked as villages) can keep city status as “Glenvar Heights - Kendall - Pinecrest - South Miami” (19).

Two new cities will be created from CDPs: “Kendale Lakes - Kendall West - The Hammocks” (26) and “Cutler Bay - Goulds - Palmetto Estates - South Miami Heights - West Perrine” (24).

Among **towns**: Aventura (previously city, risked as town), Sunny Isles (previously city, risked as village) being joined by Golden Beach (previously town, risked as village) will create town “Aventura - Golden Beach - Sunny Isles Beach” (4).

Miami Lakes (previously town, risked as town) save his positions being joined by CDPs and will create town “Country Club - Miami Lakes - Palm Springs North” (9).

Some CDPs will be merged and create new towns: “Country Walk - Richmond West” (27), “Homestead Base - Leisure City - Naranja – Princeton” (18), “Richmond Heights - The Crossings - Three Lakes” (28), “Olympia Heights - Sunset - University Park - Westwood Lakes” (23).

Medley (previously town, risked as village) save his town status being joined by Doral (previously city, risked as town) and they will form town “Doral – Medley” (11).

Homestead (previously city, risked as town) helps Florida City (previously city, risked as village) will create town “Florida City – Homestead” (17).

North Miami Beach (previously city, risked as town) will be joined by CDPs (risked as villages) – “Ives Estates - North Miami Beach – Ojus” (5).

Tamiami (previously CDP, risked as town) joins Sweetwater (previously city, risked as village) and they will form town “Sweetwater – Tamiami” (22).

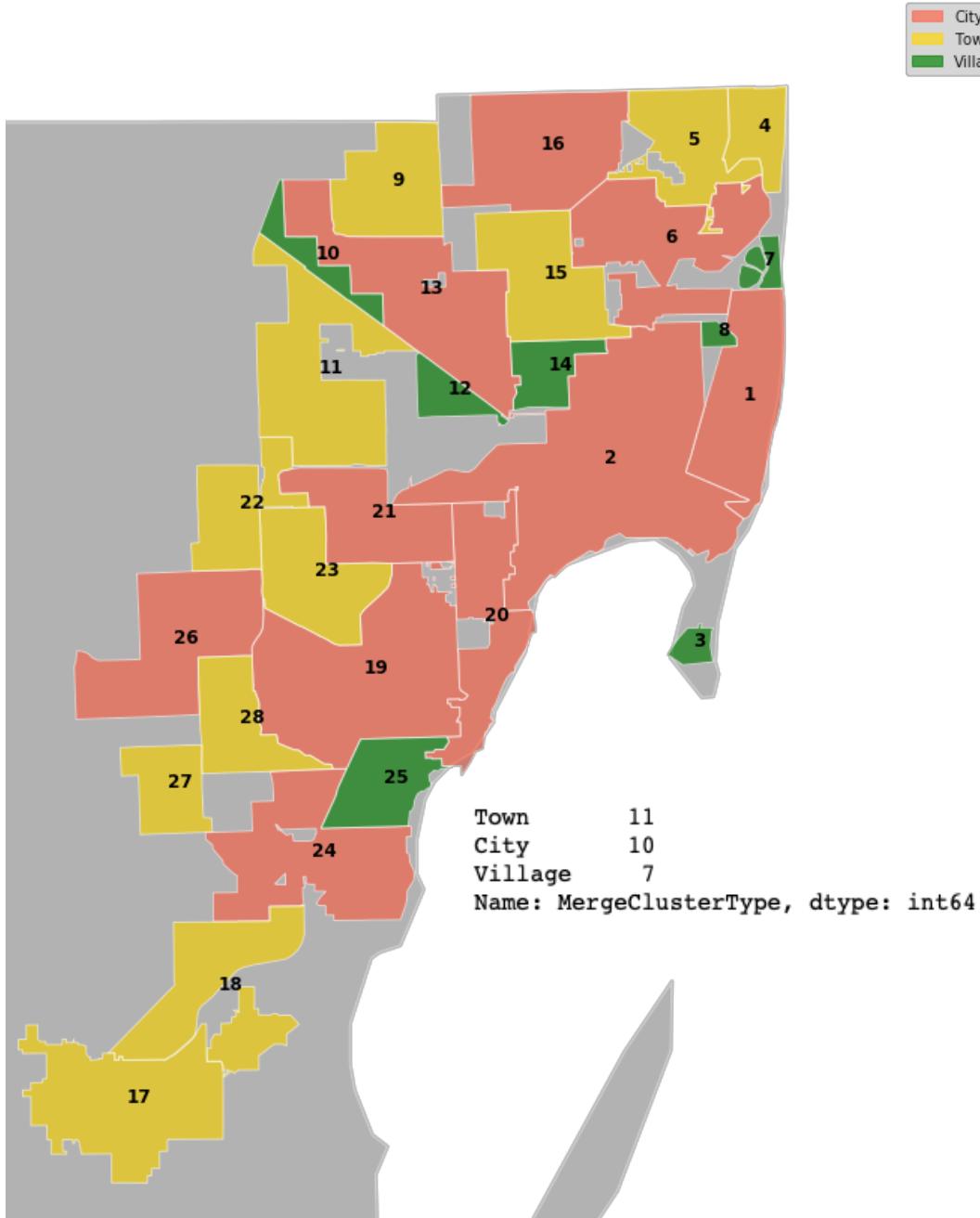
Opa-locka (previously city, risked as village) will be joined by CDPs (risked as 2 villages and 1 town) and they will create town “Opa-locka - Pinewood - West Little River – Westview” (15).

Key Biscayne (3), Palmeto Bay (25) keeps their **village** status without merging with any other communities. Hialeah Gardens (10), North Bay Village (8) are staying unmerged and as it was previously defined in K-Means Clustering in section 4.1 will be labeled as villages. “Bal Harbour - Bay Harbor Islands - Indian Creek – Surfside” (7) unless of merging 2 towns and 2 villages will be defined as village. It’s in accordance with K-Means Clustering which labeled each of them as villages.

Miami Springs (previously city, risked as village) being joined by Virginia Gardens (previously village, risked as village) will form village “Miami Springs - Virginia Gardens” (12). Some CDPs will form villages by their merging: “Brownsville – Gladeview” (14).

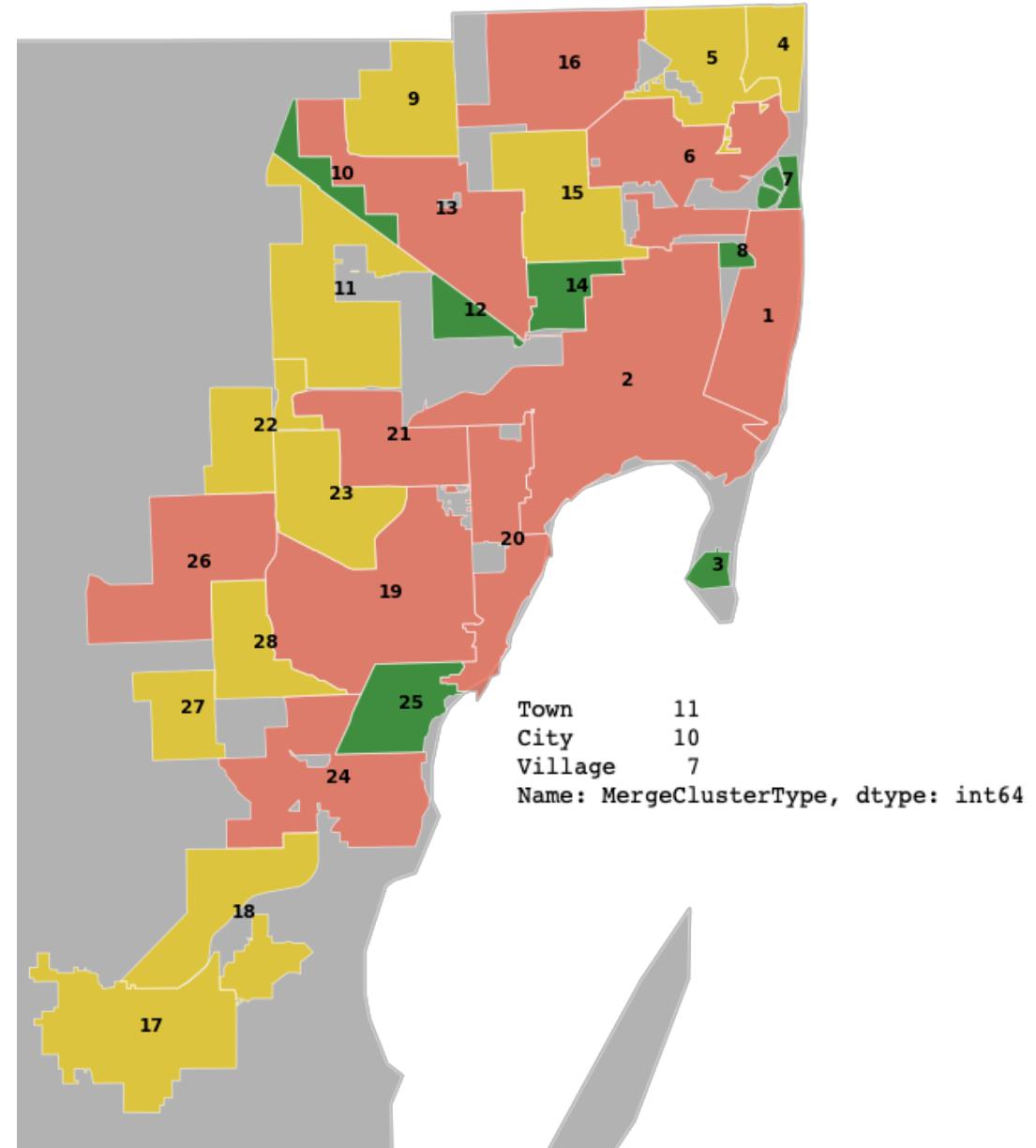
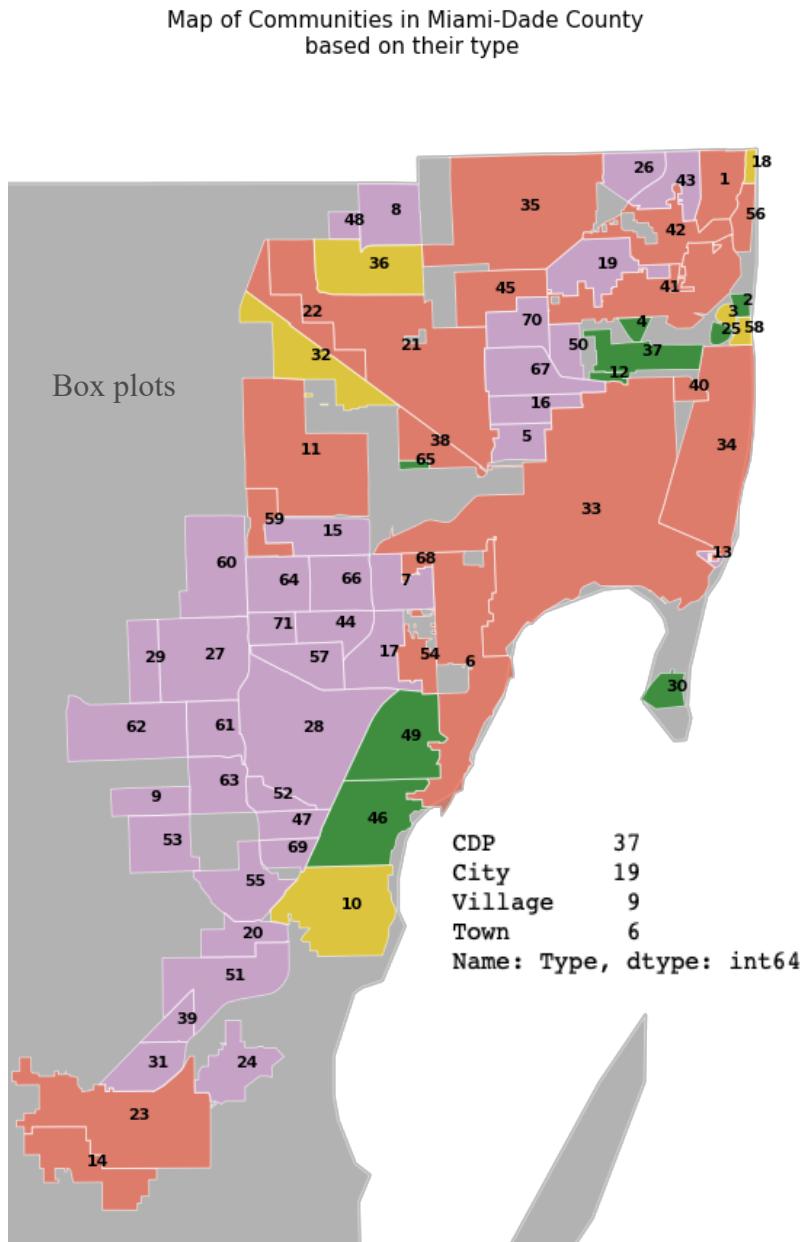
In the section 3.2 (p. 19) you can find the present distribution of communities based on their type. Now we are presenting the new communities, heir types of communities. Moreover, we can compare the distribution (age and race) before and after the changes.

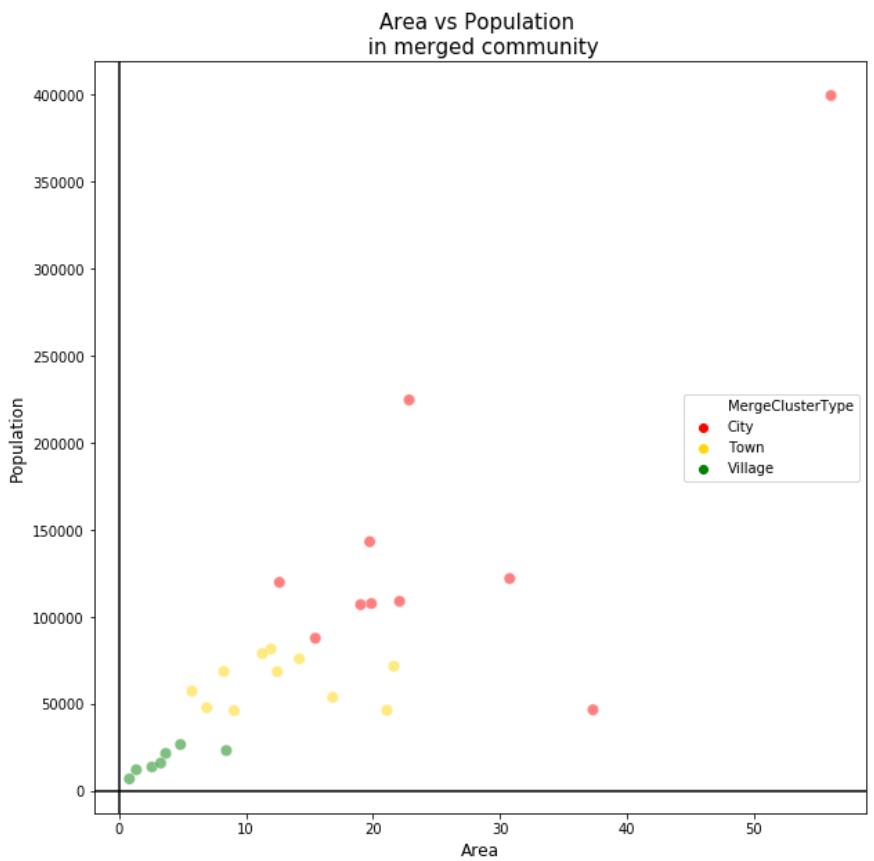
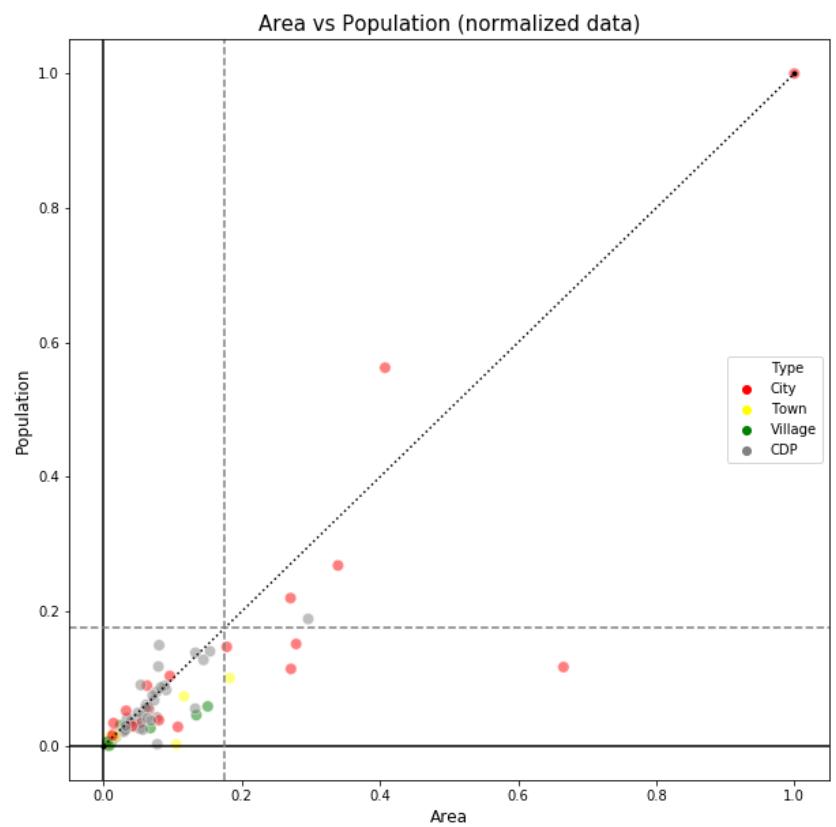
Map of Communities in Miami-Dade County
based on their merged type



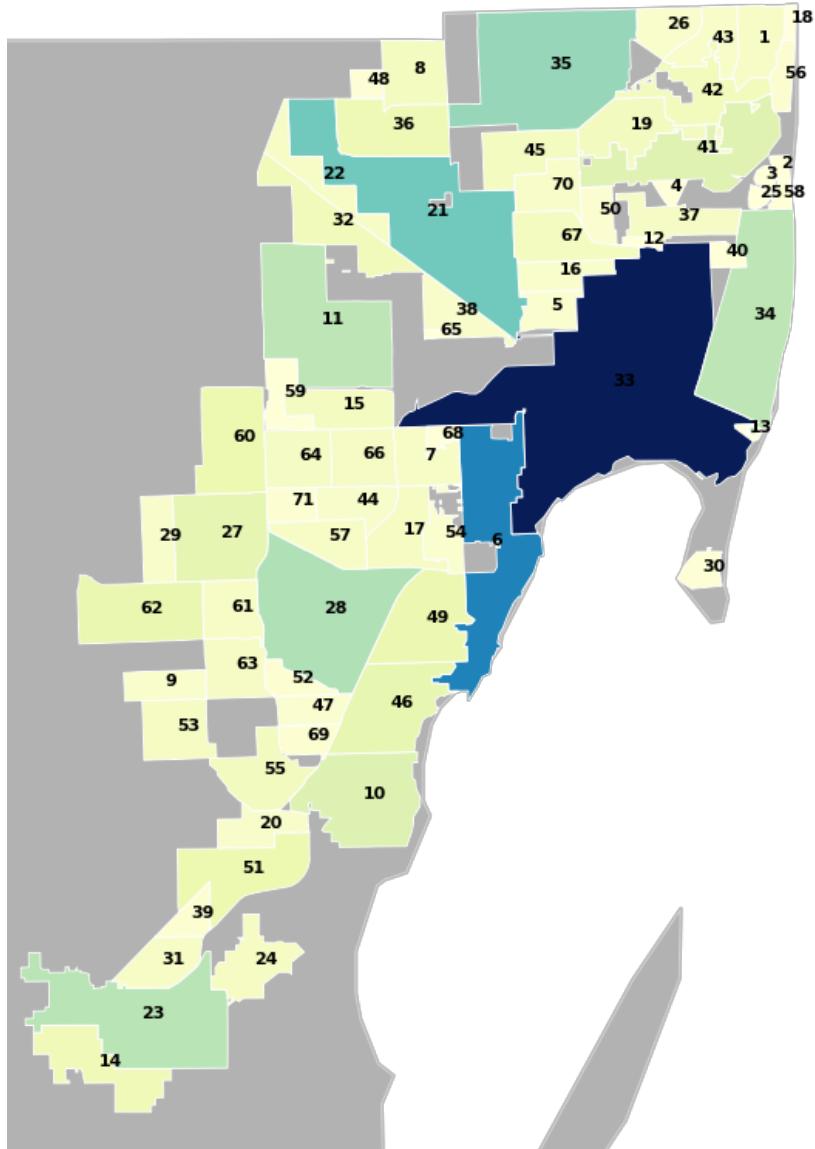
- 1 - Fisher Island - Miami Beach
- 2 - Miami
- 3 - Key Biscayne
- 4 - Aventura - Golden Beach - Sunny Isles Beach
- 5 - Ives Estates - North Miami Beach - Ojus
- 6 - Biscayne Park - El Portal - Golden Glades - Miami Shores - North Miami
- 7 - Bal Harbour - Bay Harbor Islands - Indian Creek – Surfside
- 8 - North Bay Village
- 9 - Country Club - Miami Lakes - Palm Springs North
- 10 - Hialeah Gardens
- 11 - Doral – Medley
- 12 - Miami Springs - Virginia Gardens
- 13 - Hialeah
- 14 - Brownsville – Gladview
- 15 - Opa-locka - Pinewood - West Little River – Westview
- 16 - Miami Gardens
- 17 - Florida City – Homestead
- 18 - Homestead Base - Leisure City - Naranja – Princeton
- 19 - Glenvar Heights - Kendall - Pinecrest - South Miami
- 20 - Coral Gables
- 21 - Coral Terrace - Fountainebleau - Westchester - West Miami
- 22 - Sweetwater – Tamiami
- 23 - Olympia Heights - Sunset - University Park - Westwood Lakes
- 24 - Cutler Bay - Goulds - Palmetto Estates - South Miami Heights - West Perrine
- 25 - Palmetto Bay
- 26 - Kendall Lakes - Kendall West - The Hammocks
- 27 - Country Walk - Richmond West
- 28 - Richmond Heights - The Crossings - Three Lakes

Map of Communities in Miami-Dade County
based on their merged type

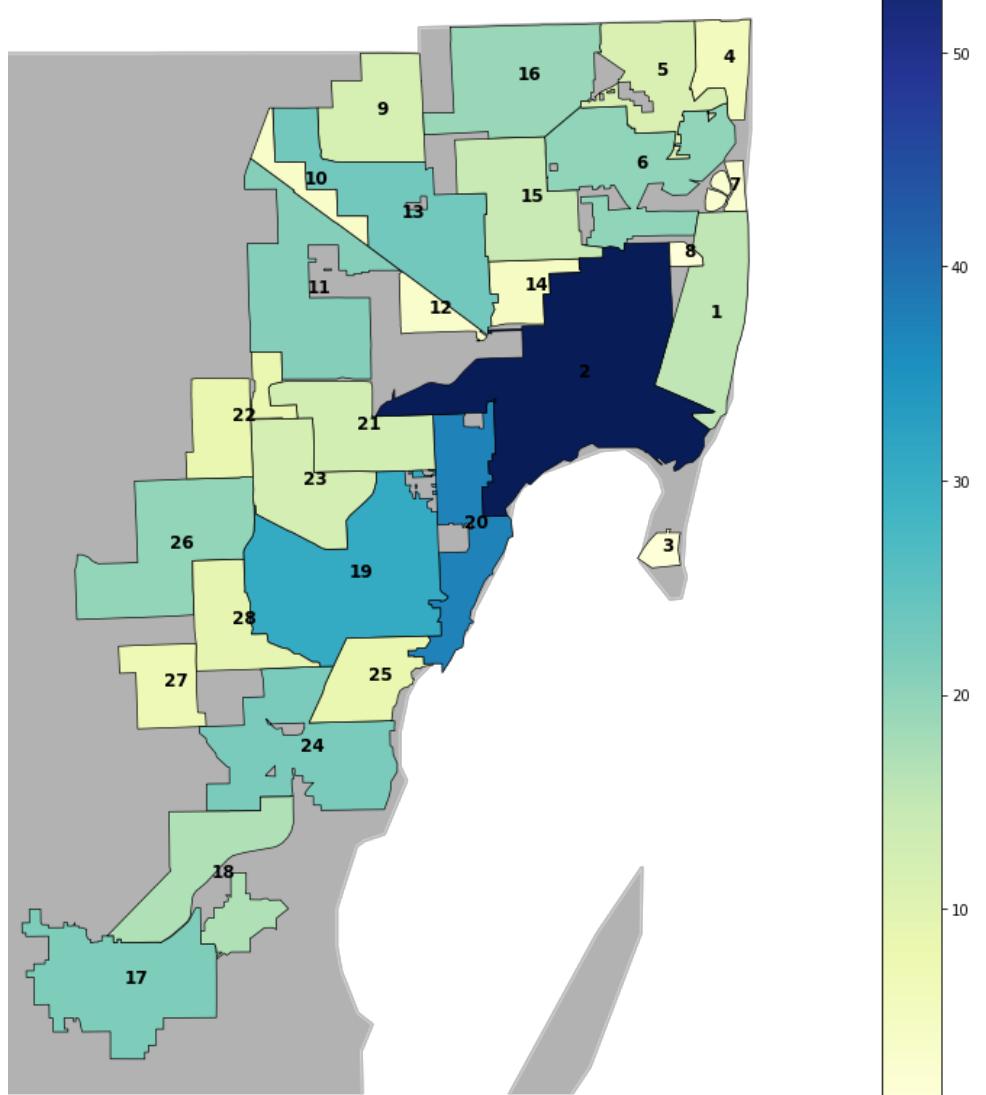




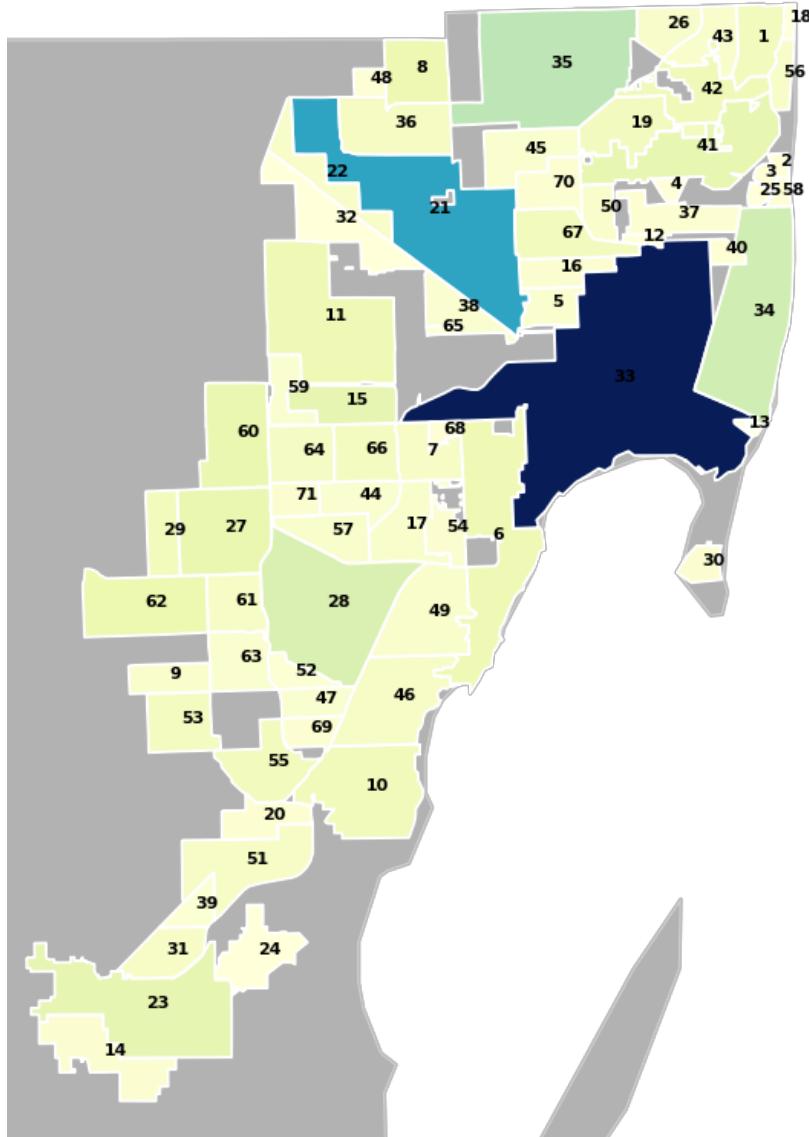
Map of Communities in Miami-Dade County
based on their area



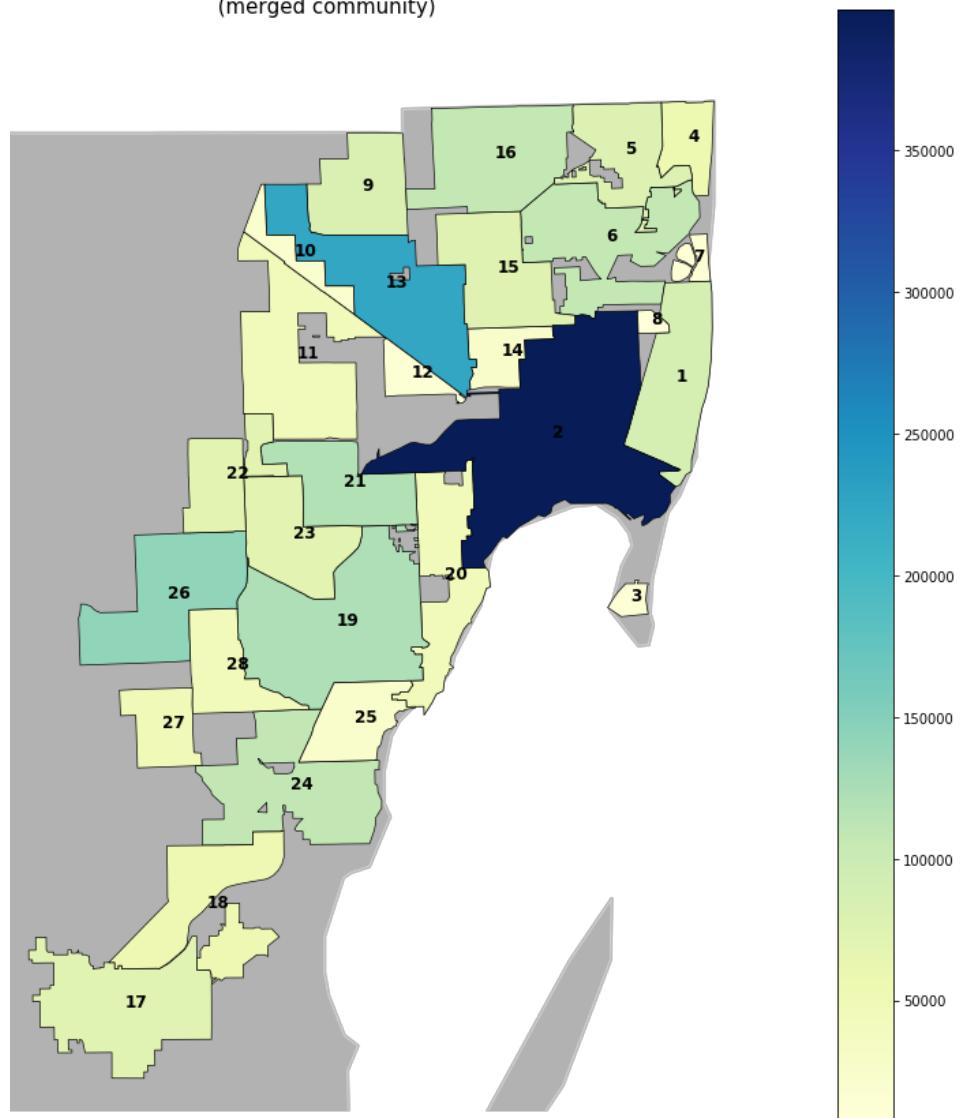
Map of Communities in Miami-Dade County
based on their area
(merged community)



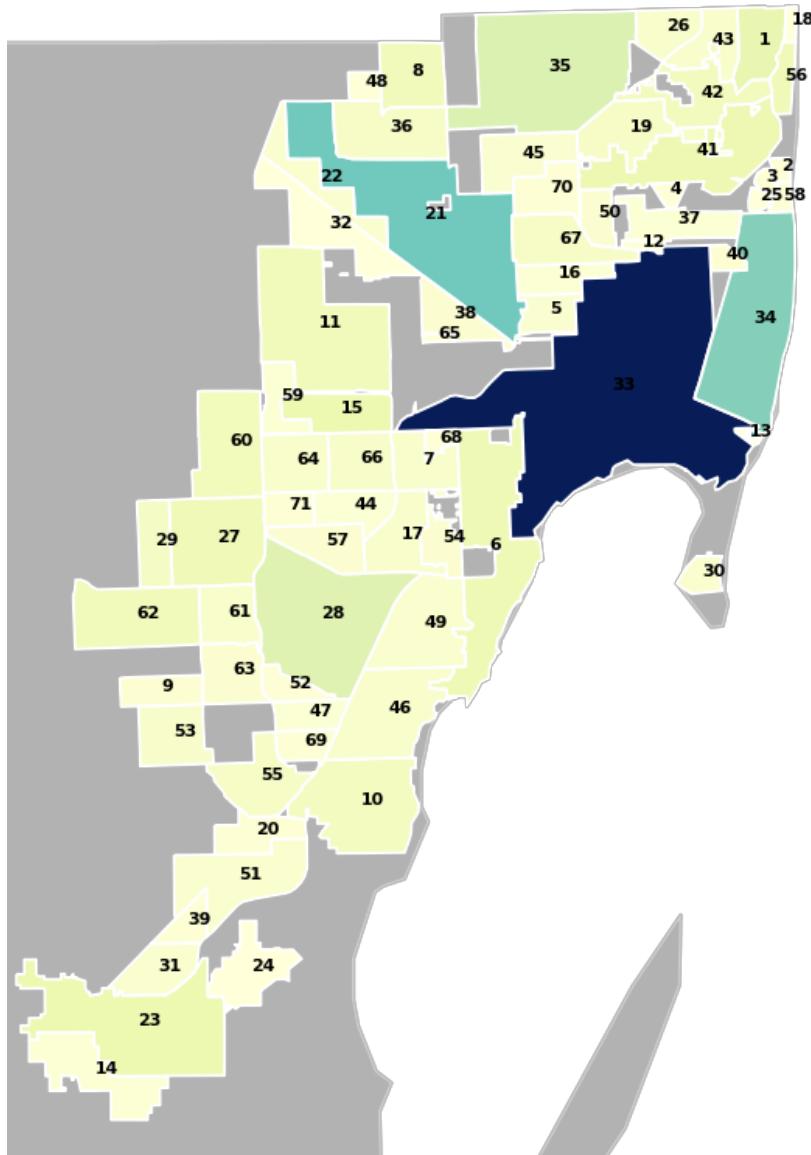
Map of Communities in Miami-Dade County
based on their population



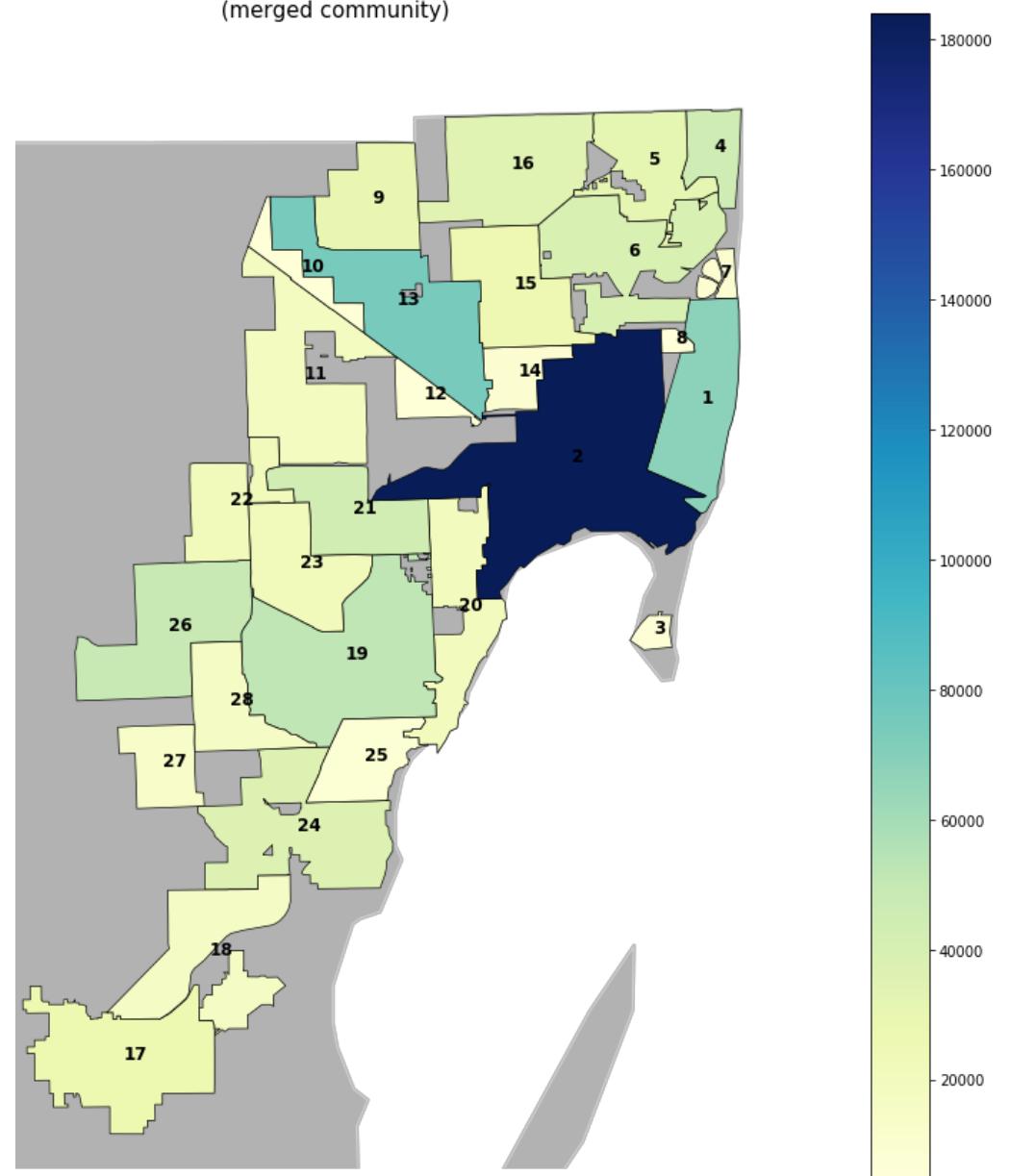
Map of Communities in Miami-Dade County
based on their population
(merged community)



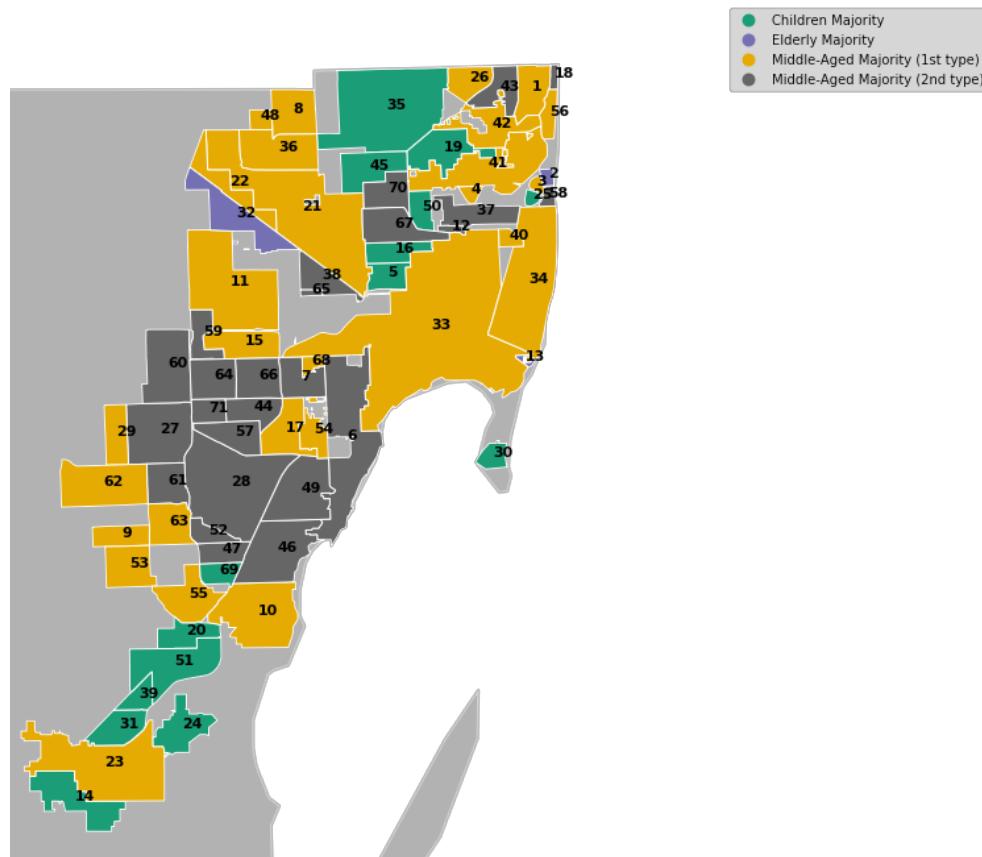
Map of Communities in Miami-Dade County
based on their housing units



Map of Communities in Miami-Dade County
based on their housing units
(merged community)

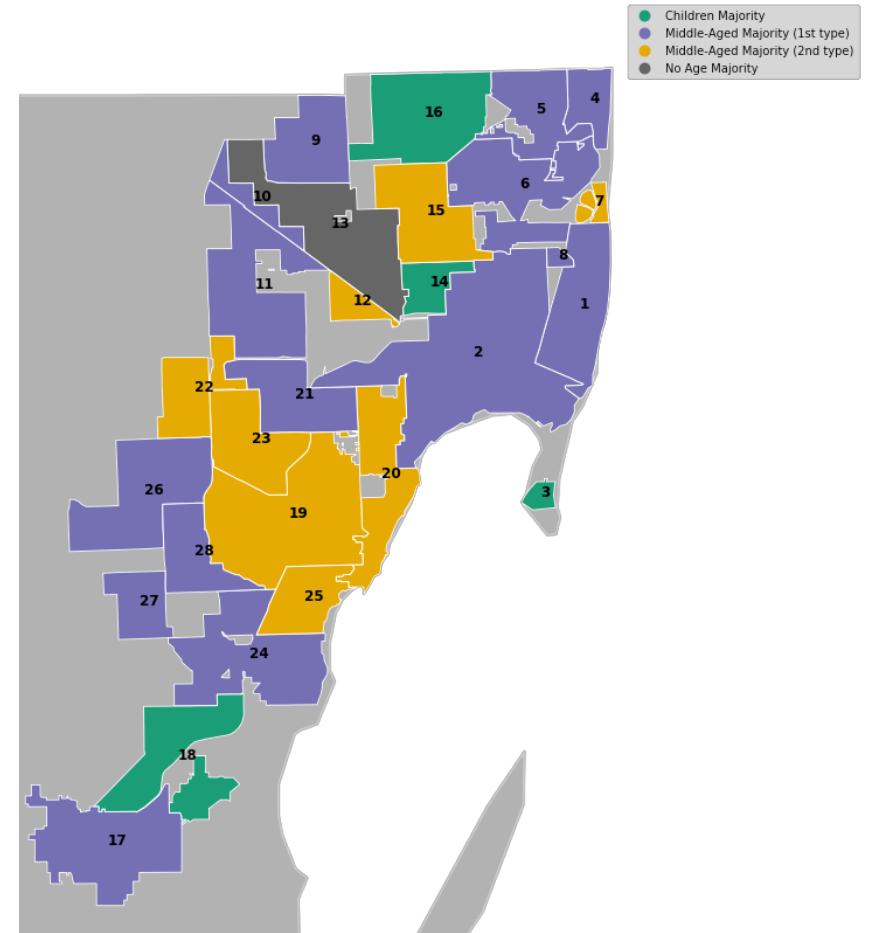


Map of Communities in Miami-Dade County
based on their age type



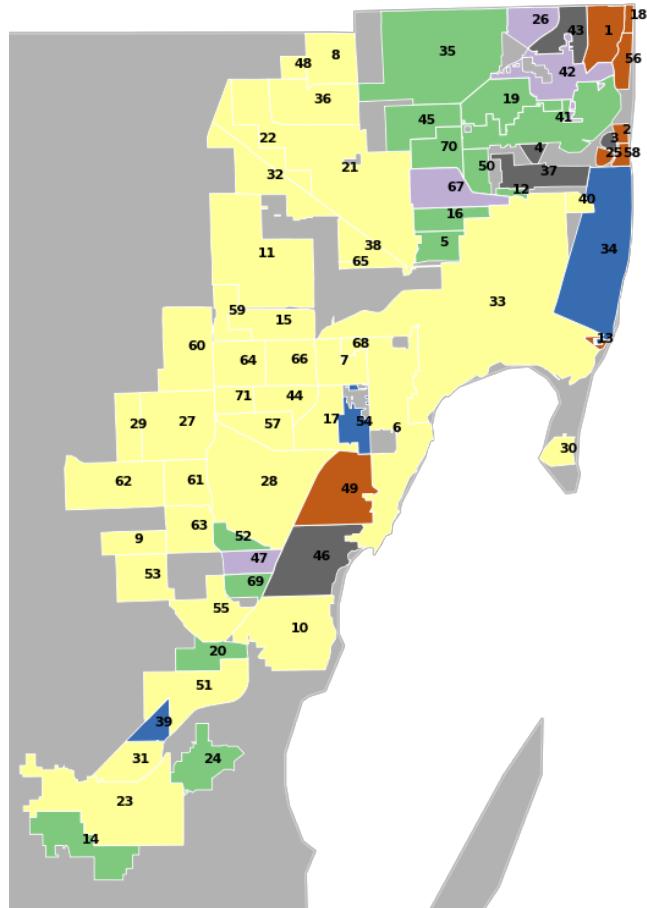
Middle-Aged Majority (1st type)	28
Middle-Aged Majority (2nd type)	25
Children Majority	15
Elderly Majority	3
Name: AgeType, dtype: int64	

Map of Communities in Miami-Dade County
based on their age type
(merged communities)



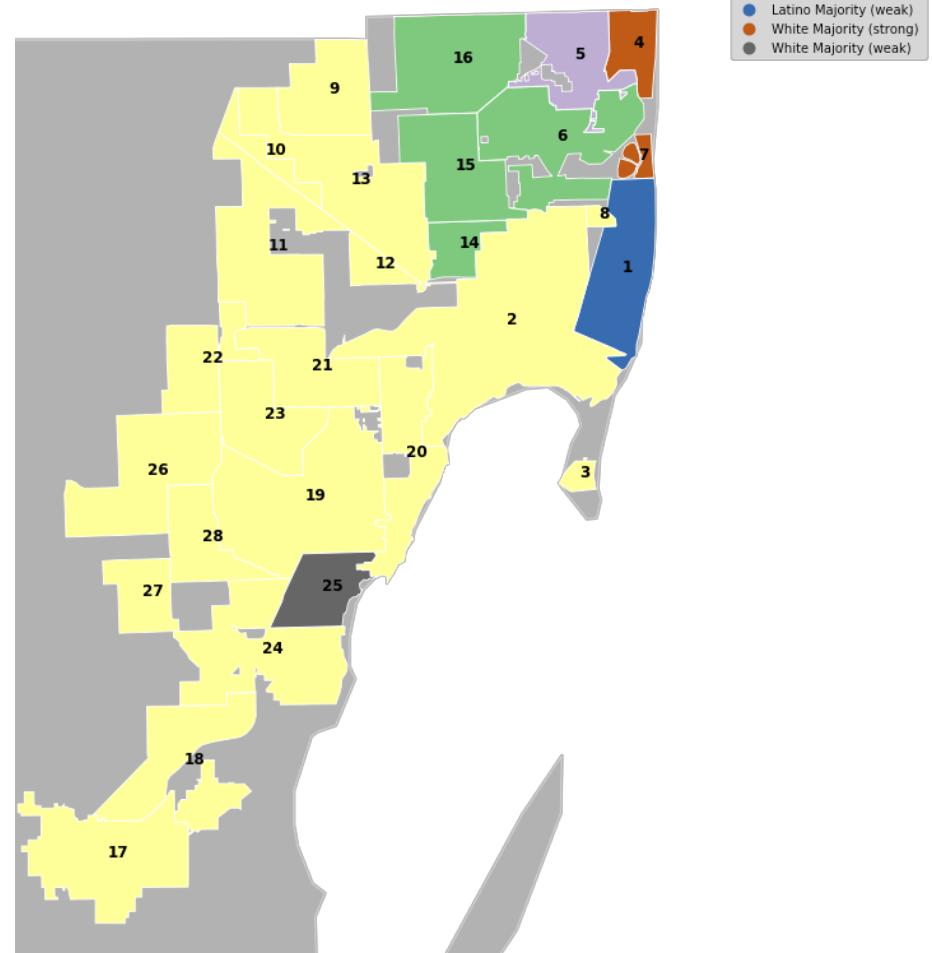
Middle-Aged Majority (1st type)	15
Middle-Aged Majority (2nd type)	8
Children Majority	4
No Age Majority	1
Name: AgeType, dtype: int64	

Map of Communities in Miami-Dade County
based on their major race



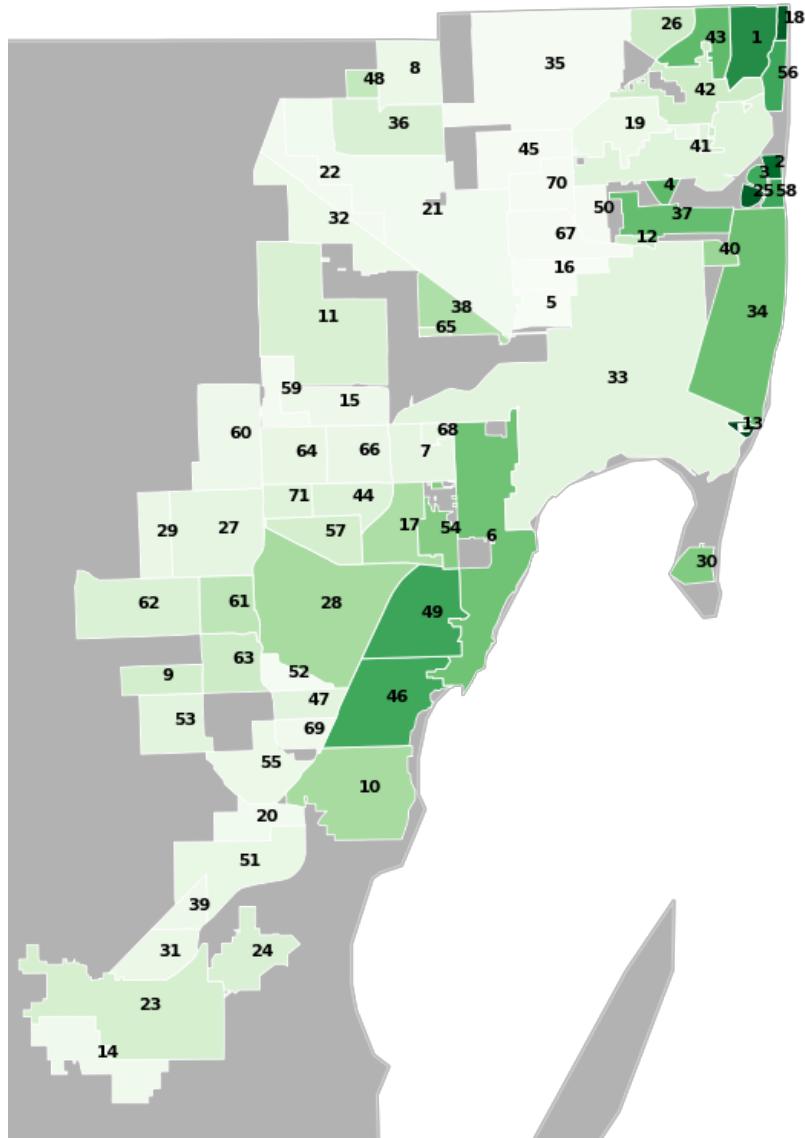
Latino Majority (strong)	37
Black Majority (strong)	14
White Majority (strong)	8
White Majority (weak)	5
Black Majority (weak)	4
Latino Majority (weak)	3
Name: RaceType, dtype: int64	

Map of Communities in Miami-Dade County
based on their major race
(merged communities)

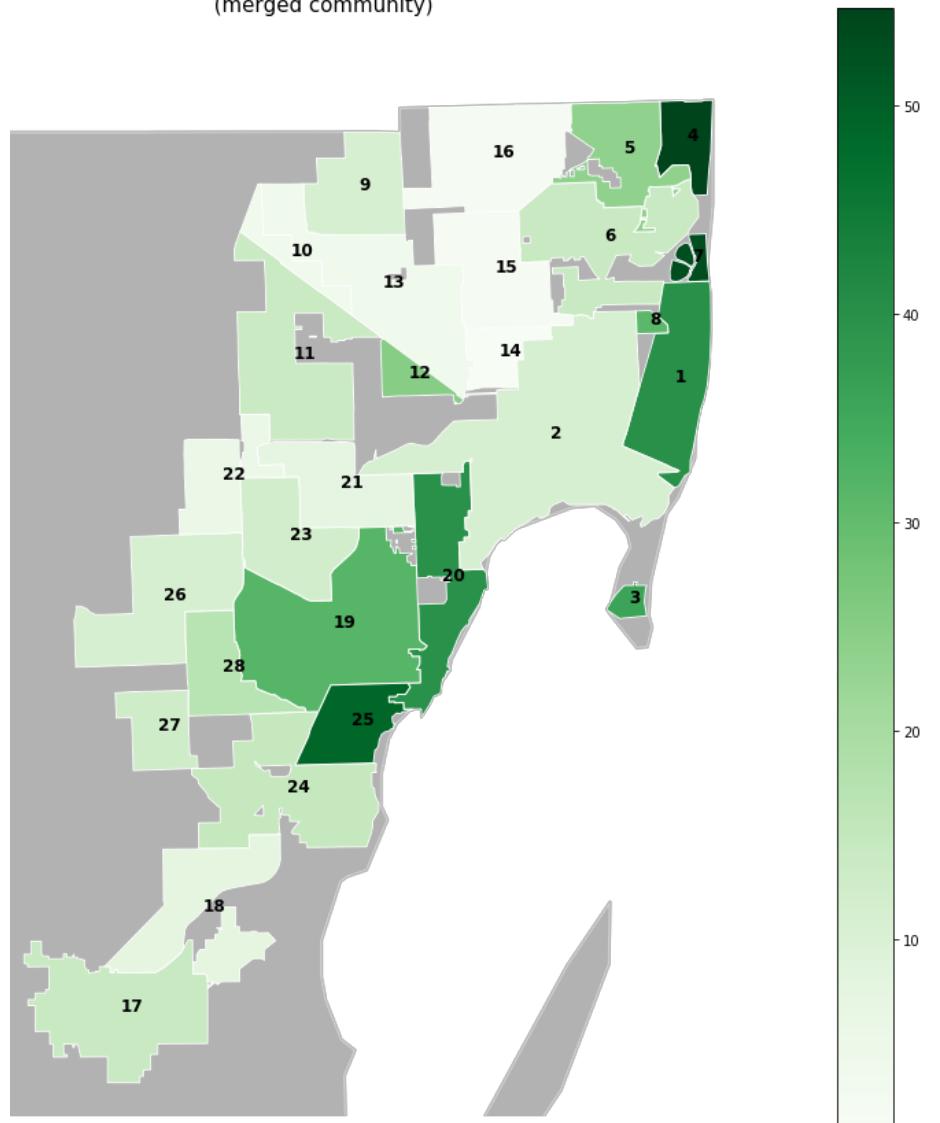


Latino Majority (strong)	19
Black Majority (strong)	4
White Majority (strong)	2
Black Majority (weak)	1
White Majority (weak)	1
Latino Majority (weak)	1
Name: RaceType, dtype: int64	

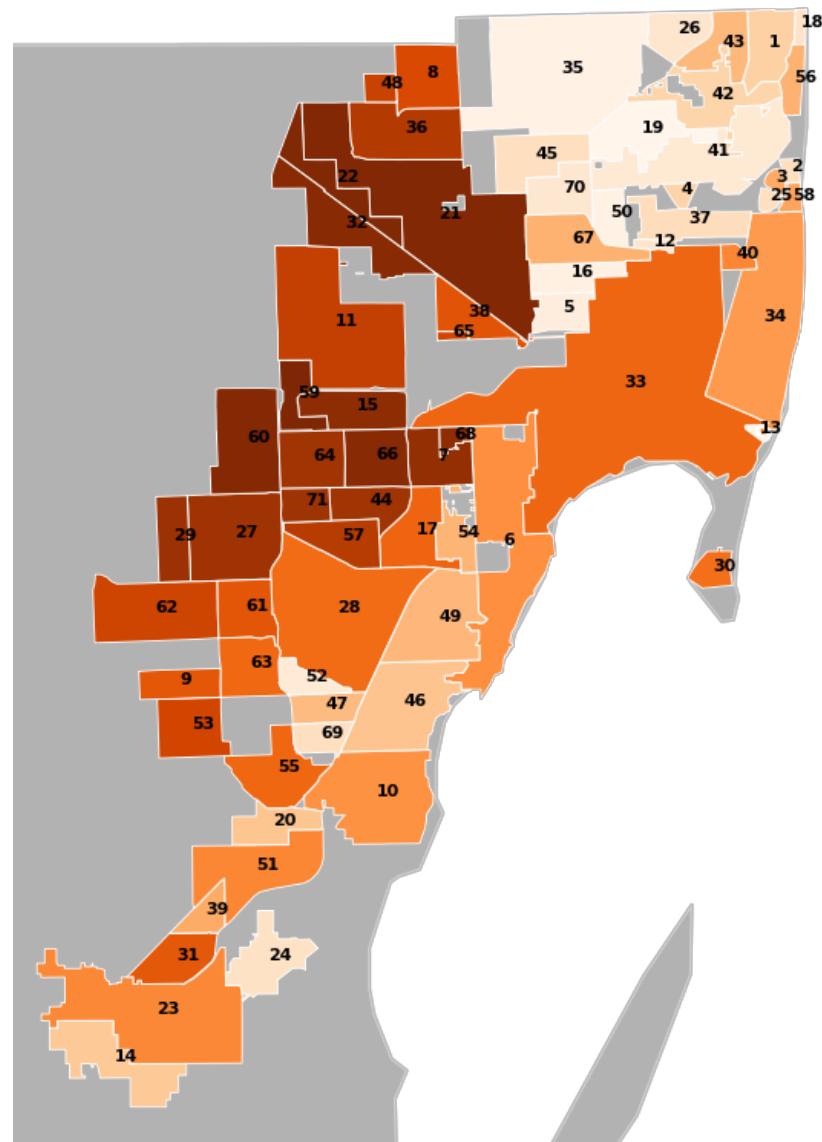
Map of Communities in Miami-Dade County
based on percentage of White



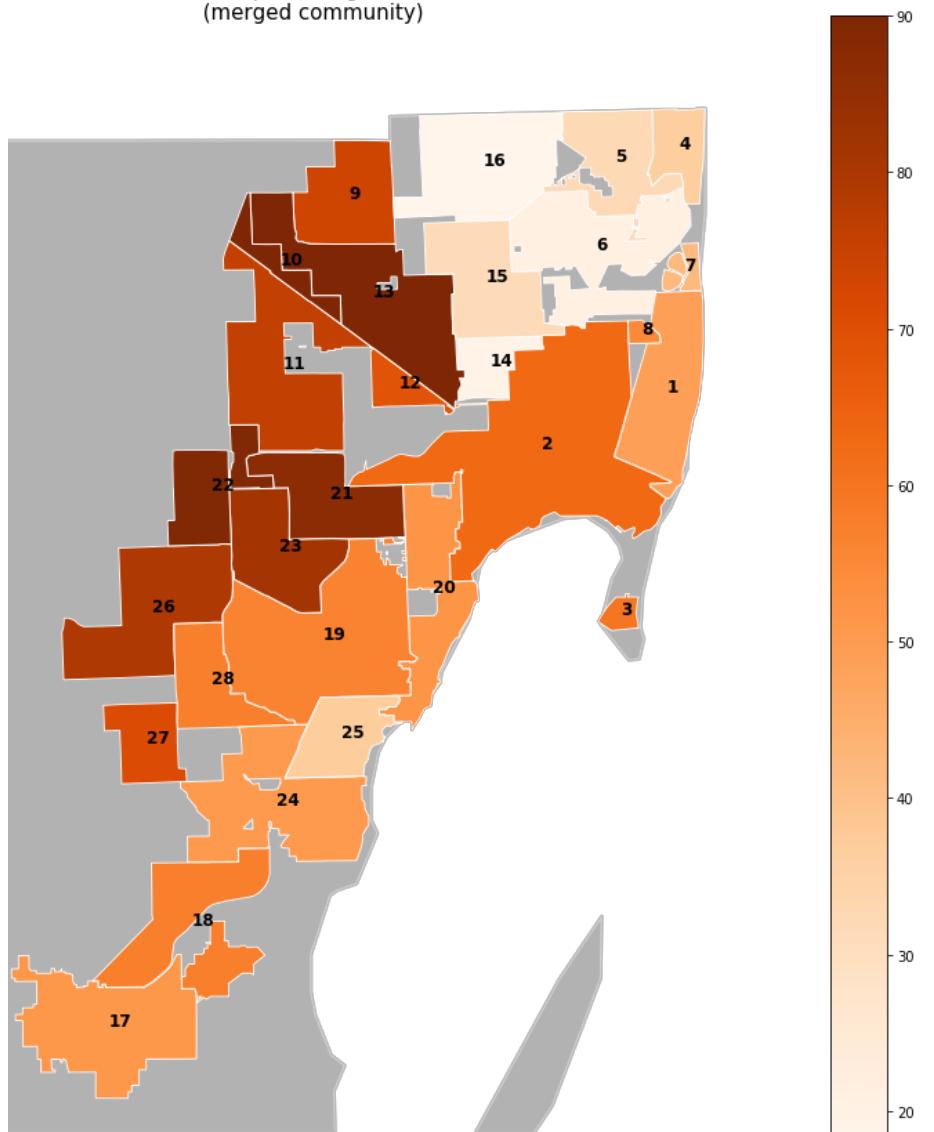
Map of Communities in Miami-Dade County
based on percentage of White
(merged community)



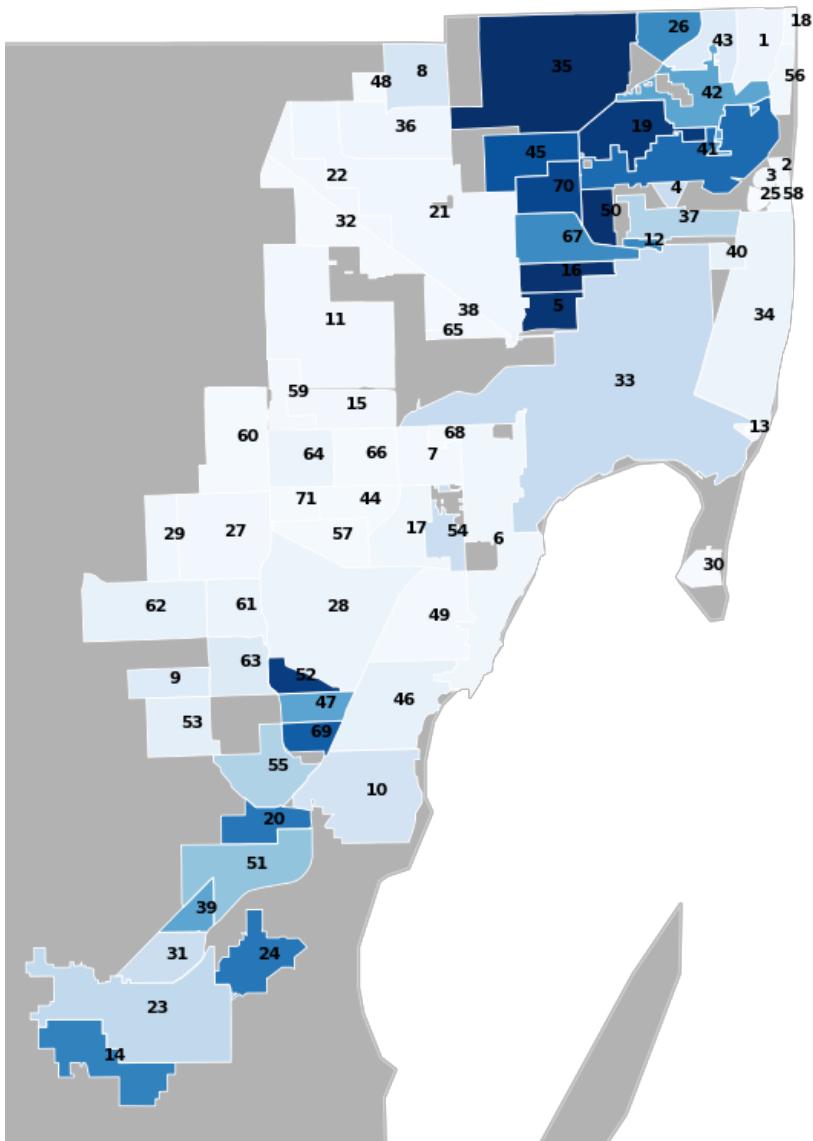
Map of Communities in Miami-Dade County
based on percentage of Latino



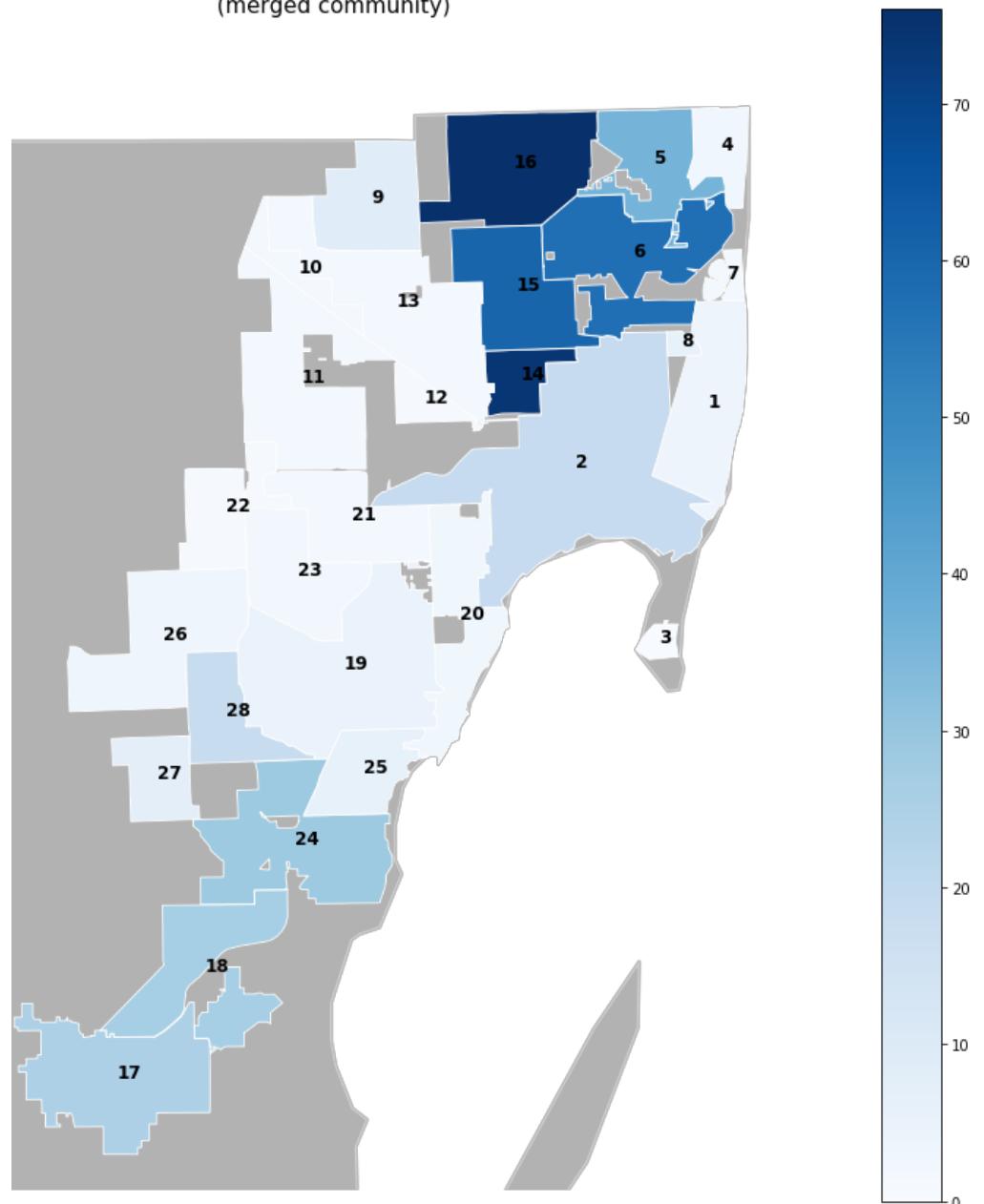
Map of Communities in Miami-Dade County
based on percentage of Latino
(merged community)



Map of Communities in Miami-Dade County
based on percentage of Black



Map of Communities in Miami-Dade County
based on percentage of Black
(merged community)



4.3. Finding and assigning places to the merged communities

Based on the merged communities and their coordinates (collected by Geolocator) with using Foursquare API we find out more than 400 locations of grocery and pharmacy stores. All of them were grouped according to the new merged communities. So, this data gives us the opportunity to define how well chains will be presented in each of merged community.

Cluster	Name	
1	CVS	8
	Navarro Discount Pharmacy	1
	Publix	4
	Trader Joe's	1
	Walgreens	11
	Whole Foods	1
2	Bravo Supermarkets	1
	CVS	16
	Navarro Discount Pharmacy	9
	Presidente	7
	Publix	11
	Walgreens	17
3	Whole Foods	1
	Winn-Dixie	6
	CVS	1
	Winn-Dixie	1
4	CVS	4
	Navarro Discount Pharmacy	2
	Publix	3
	Walgreens	3
	Whole Foods	2
5	Winn-Dixie	1
	ALDI	1
	CVS	3
	Publix	4
6	Walgreens	3
	Winn-Dixie	2
	Bravo Supermarkets	1
	CVS	6
	Navarro Discount Pharmacy	1
	Presidente	2
7	Publix	4
	Walgreens	4
	Whole Foods	1
	Winn-Dixie	1
	CVS	1
	Publix	1
8	Presidente	1
	CVS	5
9	Navarro Discount Pharmacy	2
	Publix	4
	Walgreens	1

	Winn-Dixie	2
10	Publix	1
11	ALDI	1
	CVS	3
	Navarro Discount Pharmacy	1
	Publix	4
	Walgreens	1
	Winn-Dixie	1
12	Publix	1
13	ALDI	2
	CVS	8
	Navarro Discount Pharmacy	11
	Presidente	2
	Publix	6
	Walgreens	10
	Winn-Dixie	5
14	Navarro Discount Pharmacy	1
	Walgreens	1
15	Bravo Supermarkets	3
	Presidente	1
	Walgreens	4
	Winn-Dixie	1
16	ALDI	2
	CVS	1
	Presidente	2
	Publix	1
	Walgreens	4
	Winn-Dixie	1
17	ALDI	1
	Bravo Supermarkets	1
	CVS	2
	Navarro Discount Pharmacy	1
	Publix	4
	Walgreens	5
	Winn-Dixie	1
18	CVS	2
	Publix	1
	Walgreens	2
19	CVS	16
	Navarro Discount Pharmacy	2
	Presidente	1
	Publix	6
	Trader Joe's	1
	Walgreens	6
	Whole Foods	1
	Winn-Dixie	4
20	CVS	1
	Navarro Discount Pharmacy	1
	Publix	3
	Walgreens	1
	Whole Foods	1
21	ALDI	1
	CVS	9
	Navarro Discount Pharmacy	2

		Publix	5
		Walgreens	4
		Winn-Dixie	2
22		ALDI	1
		Bravo Supermarkets	1
		CVS	3
		Navarro Discount Pharmacy	4
		Publix	1
		Walgreens	3
		Winn-Dixie	1
23		CVS	7
		Navarro Discount Pharmacy	1
		Presidente	2
		Publix	3
		Walgreens	4
		Winn-Dixie	2
24		ALDI	1
		CVS	4
		Navarro Discount Pharmacy	1
		Presidente	1
		Publix	4
		Walgreens	3
		Winn-Dixie	1
25		ALDI	1
		CVS	1
		Publix	1
		Walgreens	2
26		ALDI	2
		CVS	7
		Navarro Discount Pharmacy	4
		Presidente	1
		Publix	7
		Walgreens	2
		Whole Foods	1
		Winn-Dixie	2
27		CVS	1
		Navarro Discount Pharmacy	1
		Publix	4
		Walgreens	1
28		Bravo Supermarkets	1
		CVS	4
		Navarro Discount Pharmacy	1
		Publix	2
		Walgreens	2
		Winn-Dixie	1

Name: COUNTER, dtype: int64

5. Discussion

In both scenarios of improvement of the community type in Miami-Dade County we have received the communities divided by type and forming non-overlapping clusters (pp. 28 and 33). In both directions we determine the necessity of decreasing the number of cities (pp. 38, 48). The second approach due to the merge of communities allows us to reform community borders as well. Without any doubt merging of communities increased the area (p. 50), population (p. 51) and housing units (p. 52) in the merged communities. At the same time, it will lead to more equal distribution of age - no more elderly communities, more young communities (p.53).

6. Conclusion

Two scenarios of improvement of the community type in Miami-Dade County (inside the predefined borders and in merged communities) are presented in this research.

The first scenario is the only about determination of community type based on demographic data and using K-Means Clustering. It leads us to decreasing the number of cities in Miami-Dade County and increasing the number of towns and villages. In global aspect it gives us some criteria to determine the type of community based on demographic data.

The second scenario is two-step approach combining Hierarchical Clustering to determine the number of merging communities and K-Means Clustering solving the same problem as in the first scenario but based on aggregated data.

Moreover, project consists of the mechanism to find and assign any place to one or other merged community using Foursquare API. We take the chance to distribute most important grocery and pharmacy stores and assign them to one or other merged community.

Future development we can see in the following directions. First, CENSUS 2020 will provide us with new data for the same models and will allow us to evaluate the results of the modeling, maybe even to adjust the parameters. Second, we can use previous CENSUS data (2000, 1990 and even more) to look through the dynamics. Third, the modeling here has limitation on data – only demographics. So, we can use economic and social statistics and other data for deeper understanding the processes of community type determination.

7. Acknowledgement

I would like to say thank you for IBM and Coursera for this possibility to learn the new methods of data analysis remotely during the COVID-19 outbreak worldwide. Especially I would like to thank you our lectors: S. Aghabozorgi, R. Ahija, A. Akison, P. Lin, J. Santarcangelo.