Data Science Capstone Project - The Battle of the Neighborhoods

RELOCATING TO MIAMI AFTER RETIREMENT

Introduction

▶ Miami is one of the most popular places to retire.

▶ Use data science skills to help customers who would consider relocating to Miami after retirement to get familiar with the city, as well as to find some areas they might be interested in.

▶ Focus on the analysis and comparison of different zip code areas for four specific category groups:

- hospital facilities
- grocery shopping markets
- indoor entertainment places
- outdoor opportunities

Data

▶ List of zips in Miami and related information, such as latitude and longitude coordinates and population of those zips. This information is required to plot the map, get the venue data, and do all the analysis.

▶ Use geopy library to get the geographical coordinates of Miami, FL.

▶ Use Foursquare API to get venues for each zip in Miami. The venue data in category groups like hospital facilities, grocery facilities, indoor facilities and outdoor facilities is particularly useful for this project. The venue data is also used to perform clustering on different zips.

Libraries

- pandas for data manipulation and analysis

- geopy library to get the geographical coordinates of a location

- Python Folium library to visualize geospatial data

- Foursquare API to explore venues for different zips

- Matplotlib library and associated modules for plotting

- Scikit-learn k-means as the method of unsupervised learning to cluster the zips
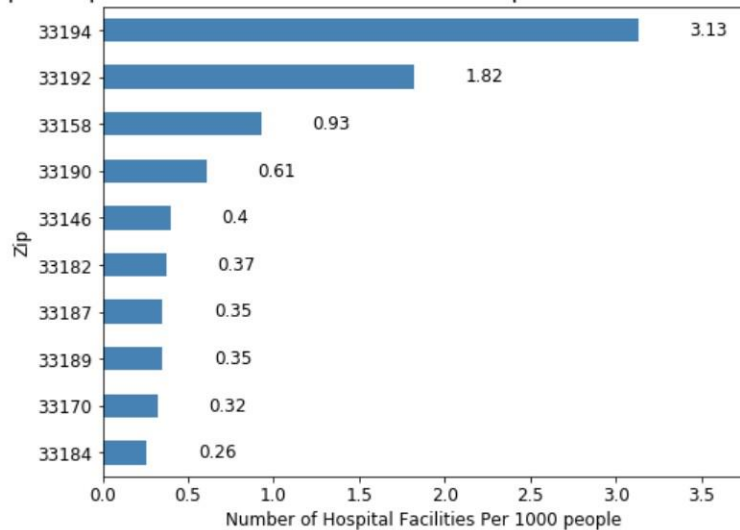
Map of Miami with its zips
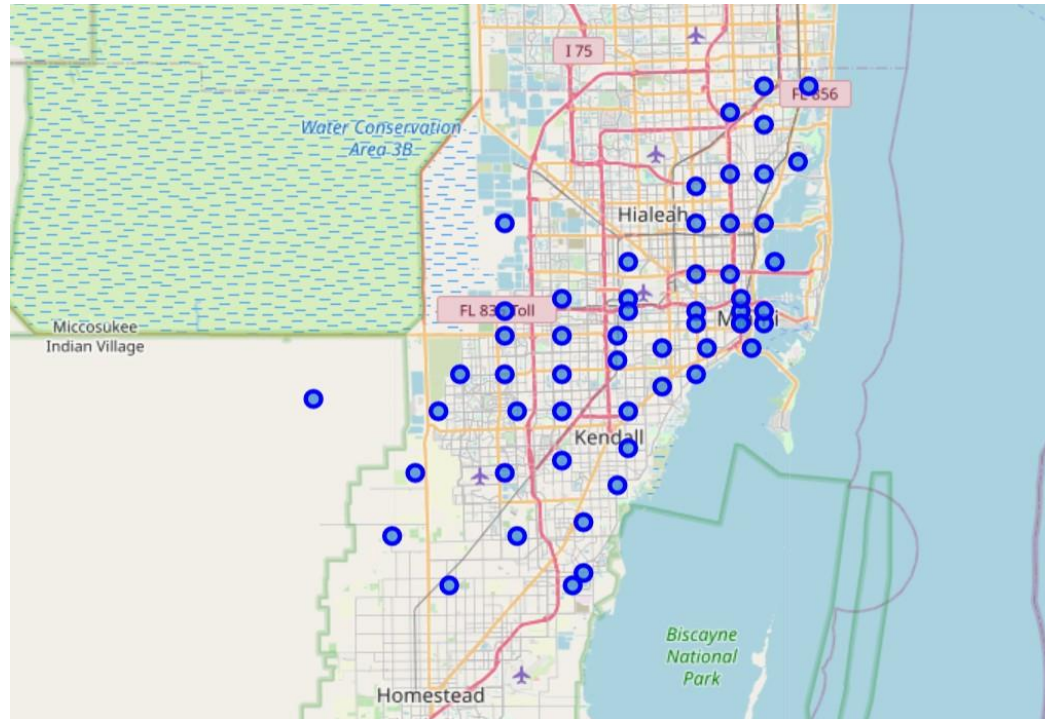
Analyze hospital facilities

▶ Hospital facilities include: hospitals, medical centers, urgent care centers, and emergency rooms.

▶ The top 10 (or more, if there are ties) zips that have the most total numbers of hospital facilities:

Top 20 Zips In Miami That Have The Most Hospital Facilities

Hospital Facilities In The Top 20 Zips That Have The Most Hospital Facilities
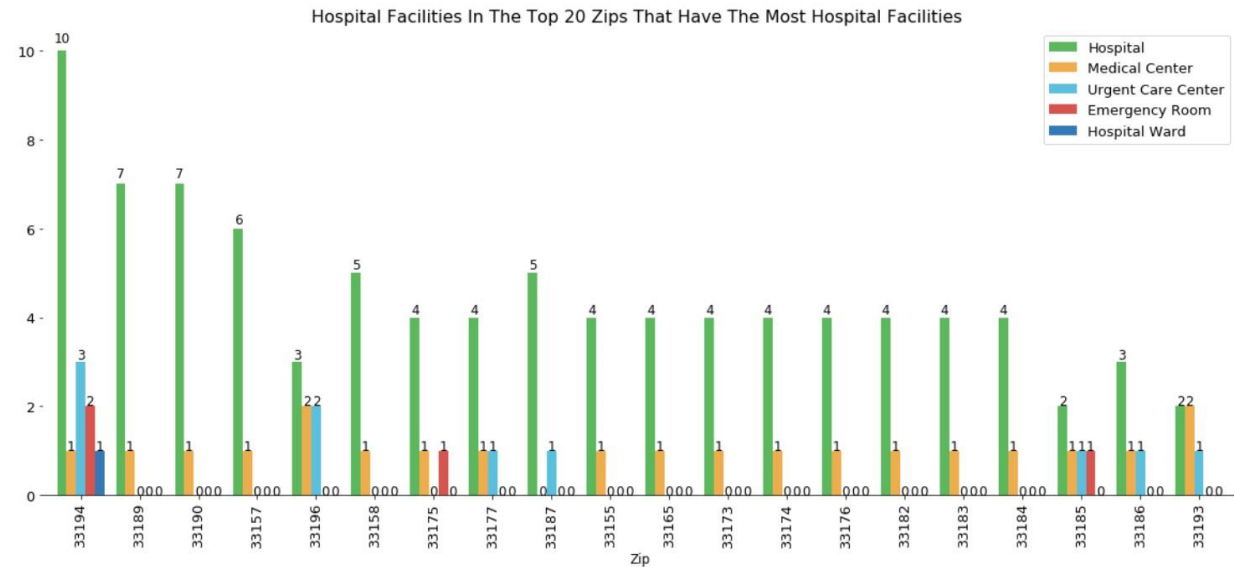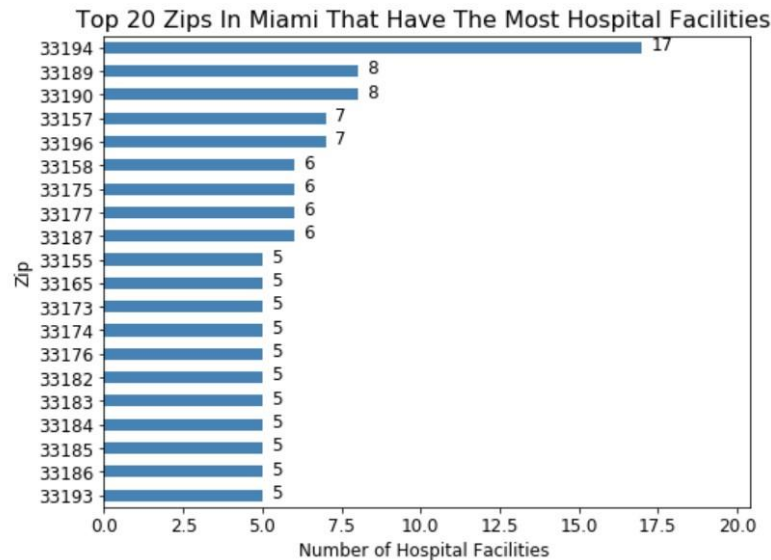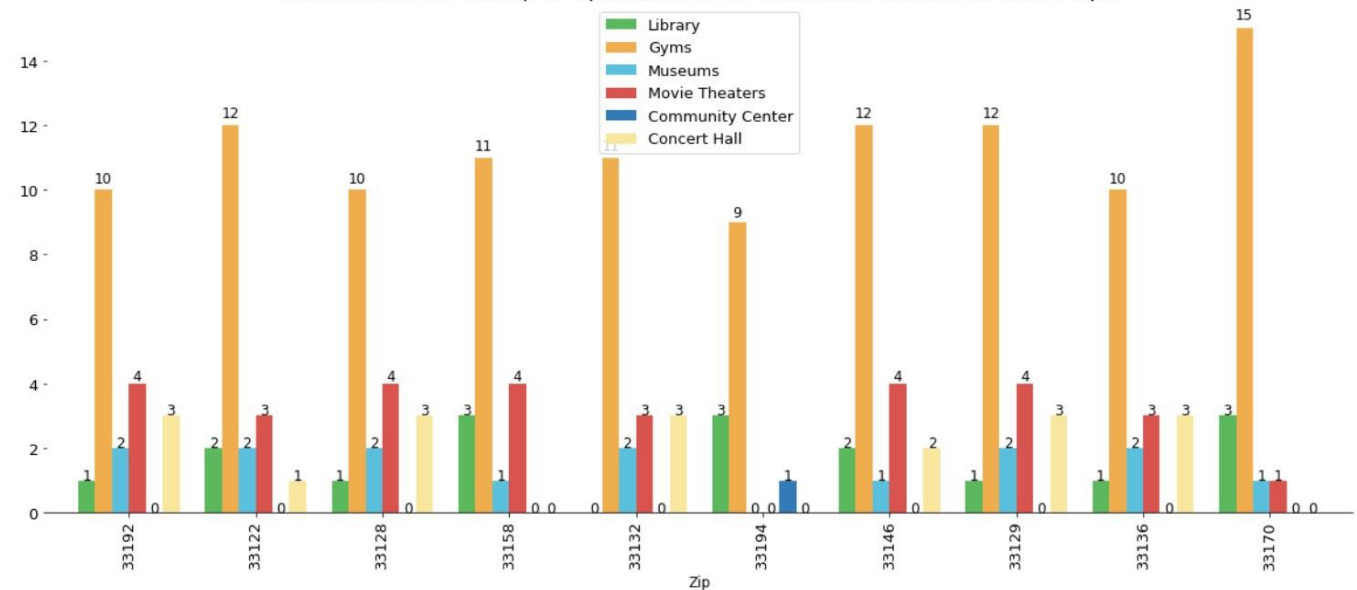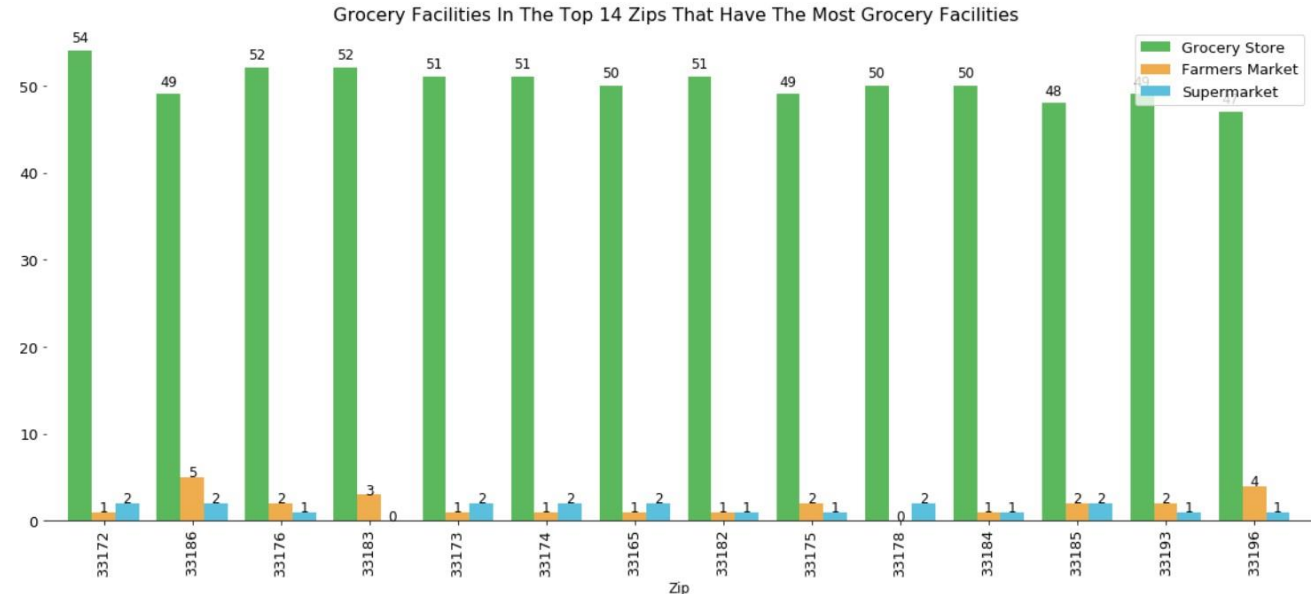
Analyze hospital facilities (cont.)

► The top 10 (or more, if there are ties) zips that have the most numbers of hospital facilities per 1000 people:

Analyze grocery facilities

▶ Grocery facilities include: grocery stores, farmers markets, and supermarkets.

▶ The top 10 (or more, if there are ties) zips that have the most total numbers of grocery facilities:

Analyze grocery facilities (cont.)

► The top 10 (or more, if there are ties) zips that have the most numbers of grocery facilities per 1000 people:

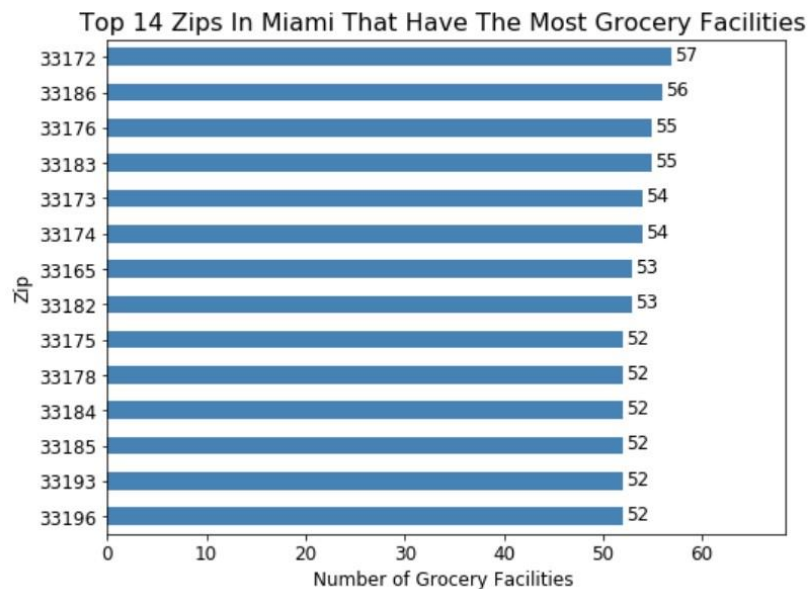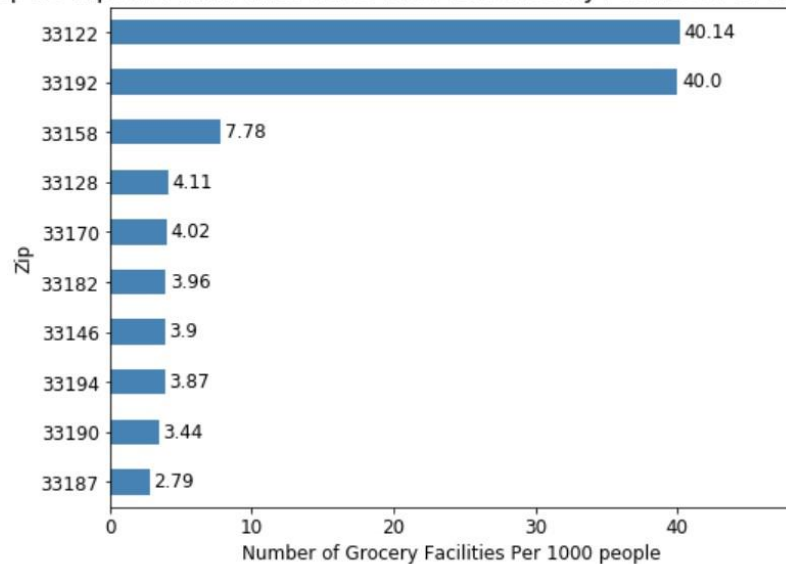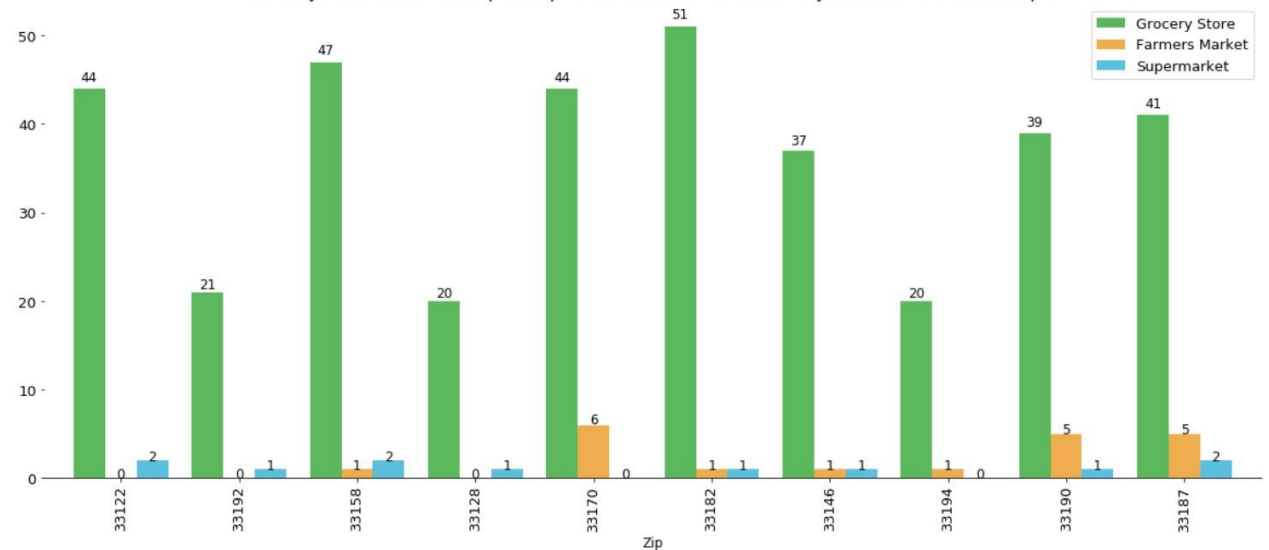Top 10 Zips In Miami That Have The Most Grocery Facilities Per 1000 people
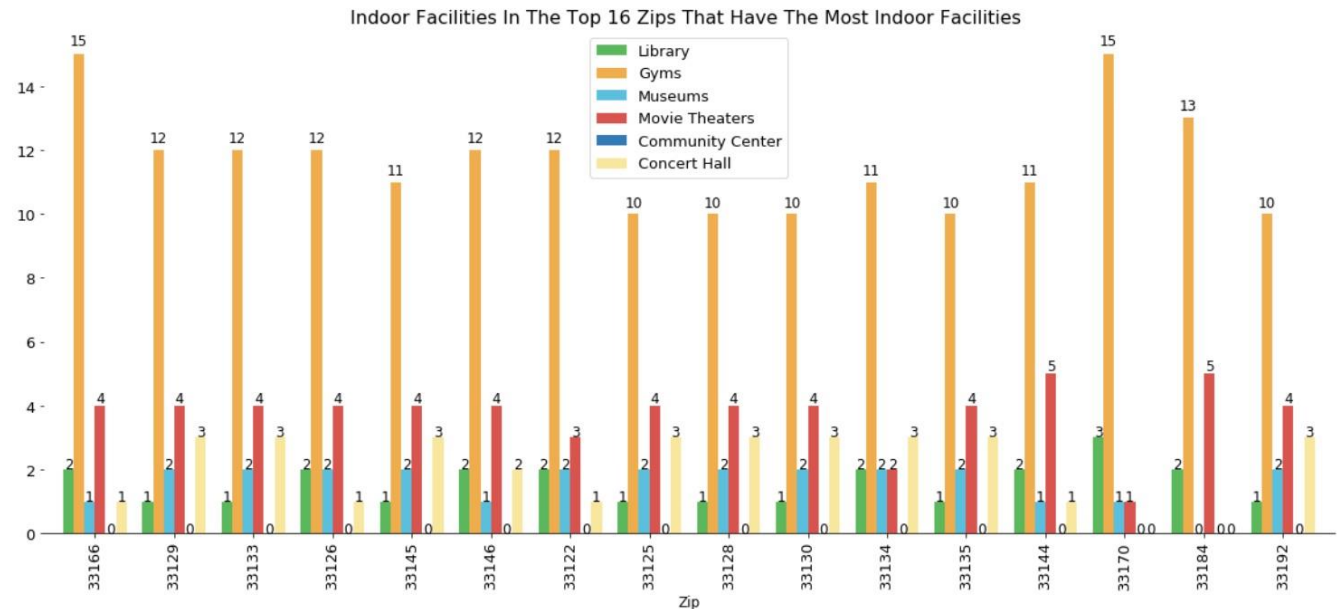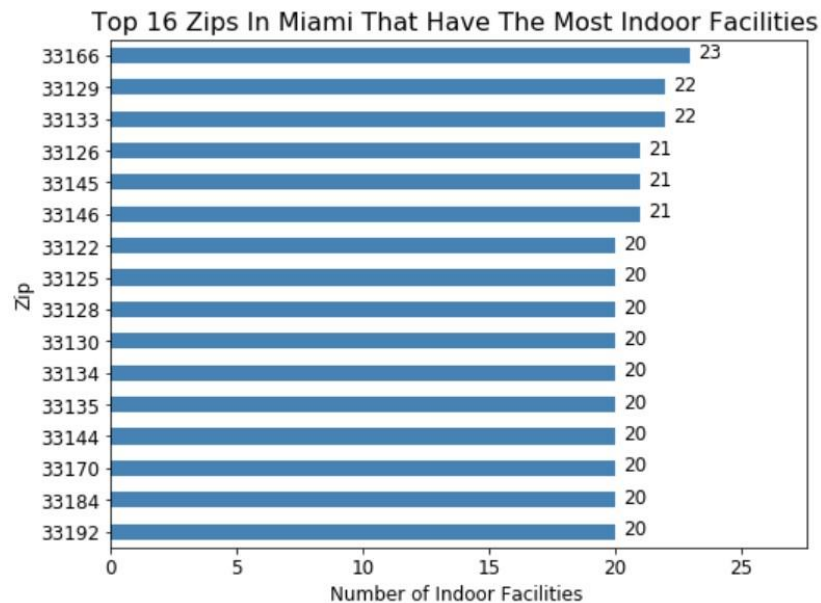

Grocery Facilities In The Top 10 Zips That Have The Most Grocery Facilities Per 1000 People
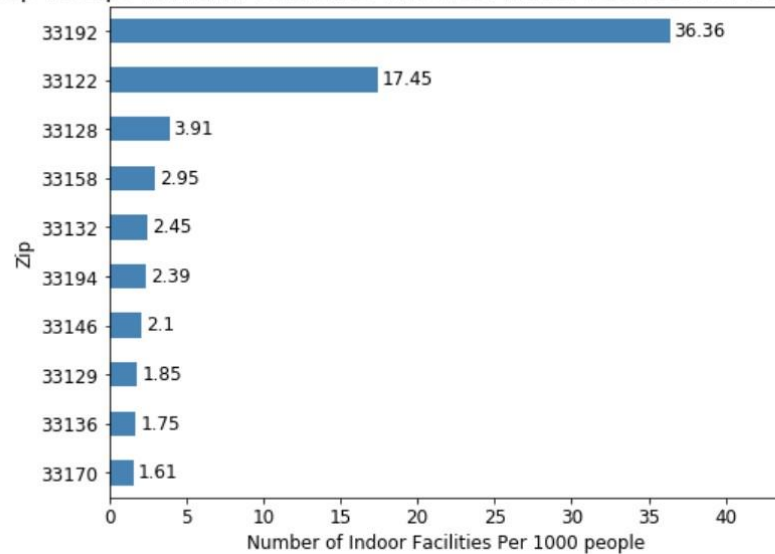
Analyze indoor facilities

▶ Indoor facilities include: libraries, gyms, museums, movie theaters, community centers, and concert halls.

▶ The top 10 (or more, if there are ties) zips that have the most total numbers of indoor facilities:

Analyze indoor facilities (cont.)

▶ The top 10 (or more, if there are ties) zips that have the most numbers of indoor facilities per 1000 people:

Analyze outdoor facilities

▶ Outdoor facilities include: parks, trails, gardens, beaches, and lakes.

▶ The top 10 (or more, if there are ties) zips that have the most total numbers of outdoor facilities:
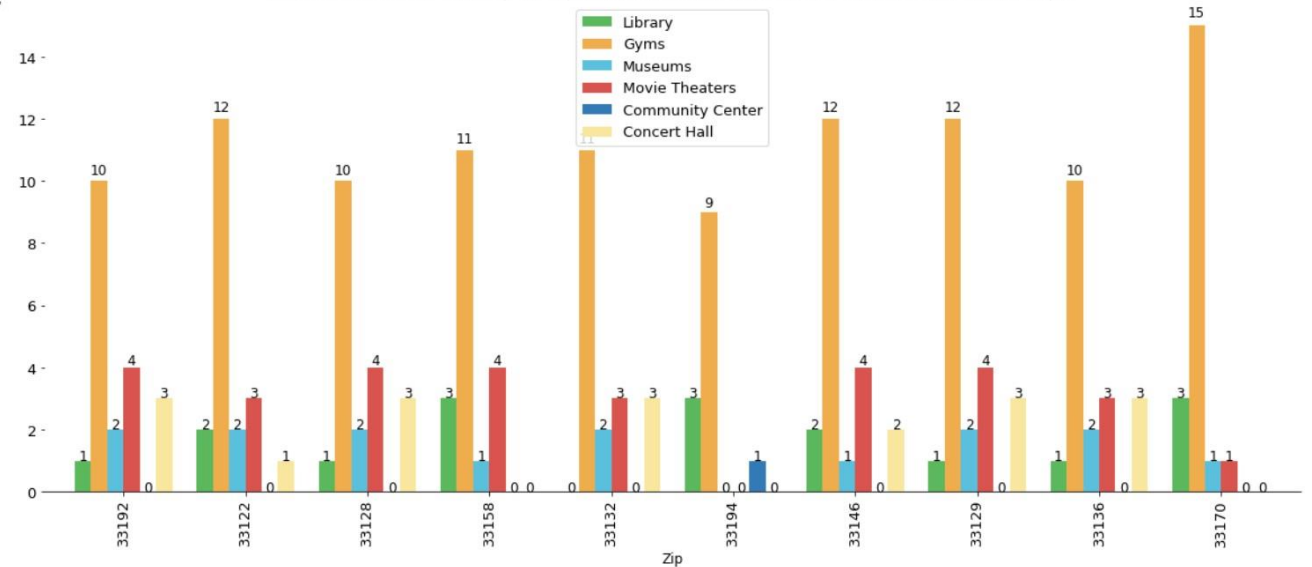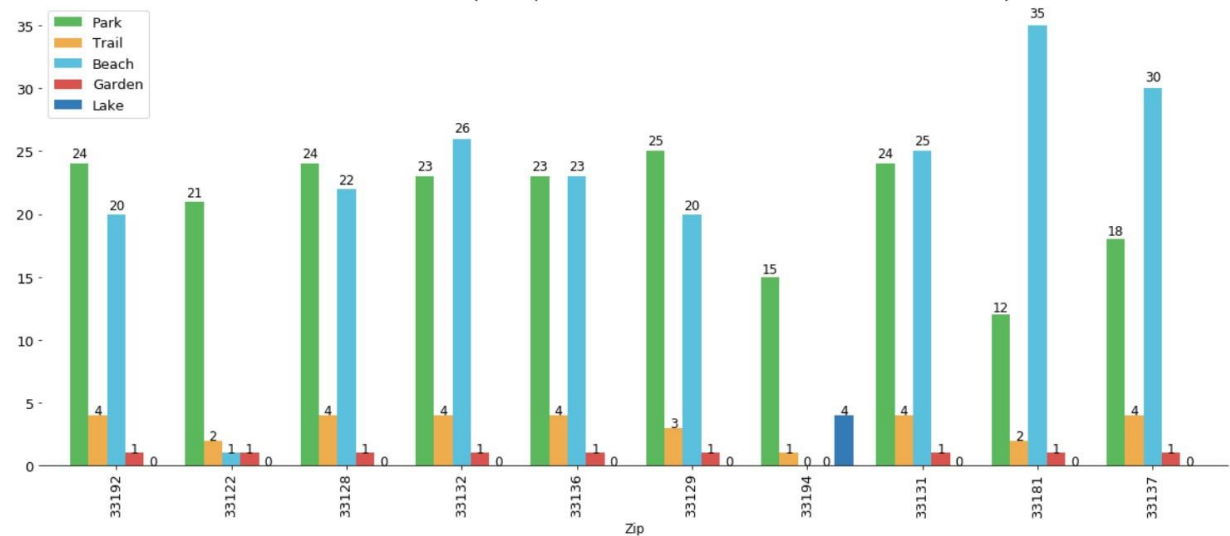
Analyze outdoor facilities (cont.)

▶ The top 10 (or more, if there are ties) zips that have the most numbers of outdoor facilities per 1000 people:

Clustering for hospital facilities
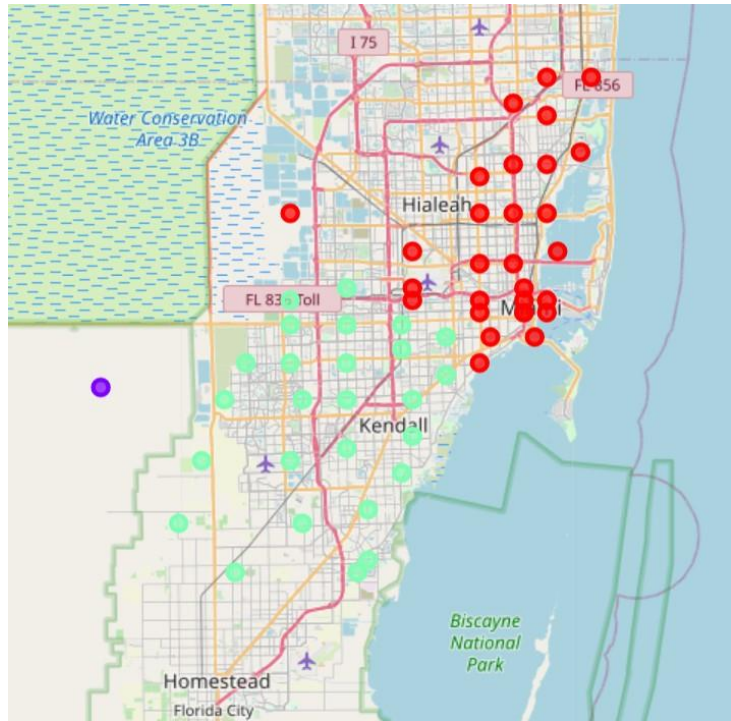
- Cluster 0: red, relatively fewer hospital facilities;
- Cluster 1: purple, relatively more hospital facilities;
- Cluster 2: green, medium number of hospital facilities.

	Hospital Facilities			
Cluster Labels	mean	min	max	count
0	0.006552	0.0000	0.0200	29
1	0.212500	0.2125	0.2125	1
2	0.052308	0.0300	0.0800	26

Clustering for grocery facilities
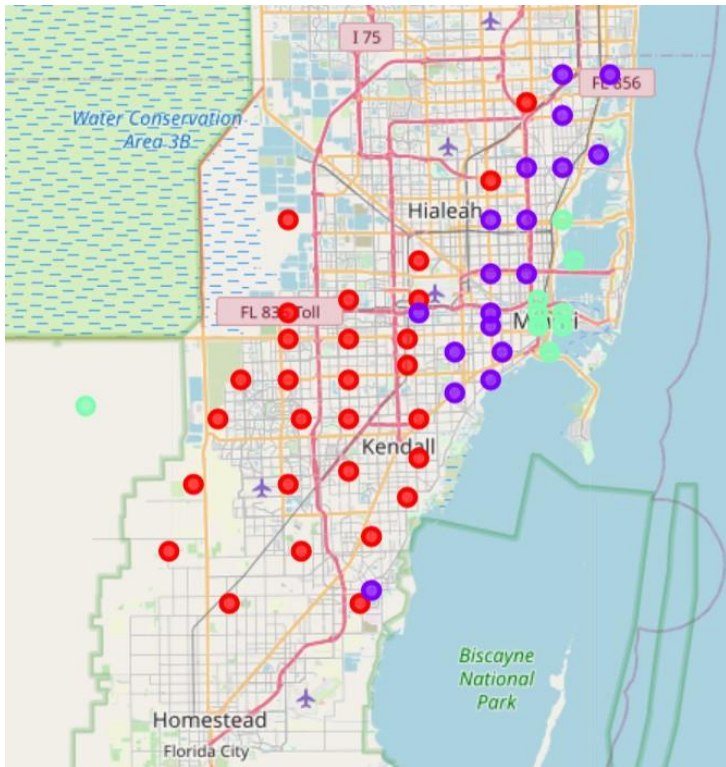
- Cluster 0: red, relatively more grocery facilities;
- Cluster 1: purple, medium number of grocery facilities;
- Cluster 2: green, relatively fewer grocery facilities.

Grocery Facilities

Cluster Labels	mean	min	max	count
0	0.503214	0.44	0.5700	28
1	0.350556	0.29	0.4200	18
2	0.218250	0.19	0.2625	10

Clustering for indoor facilities
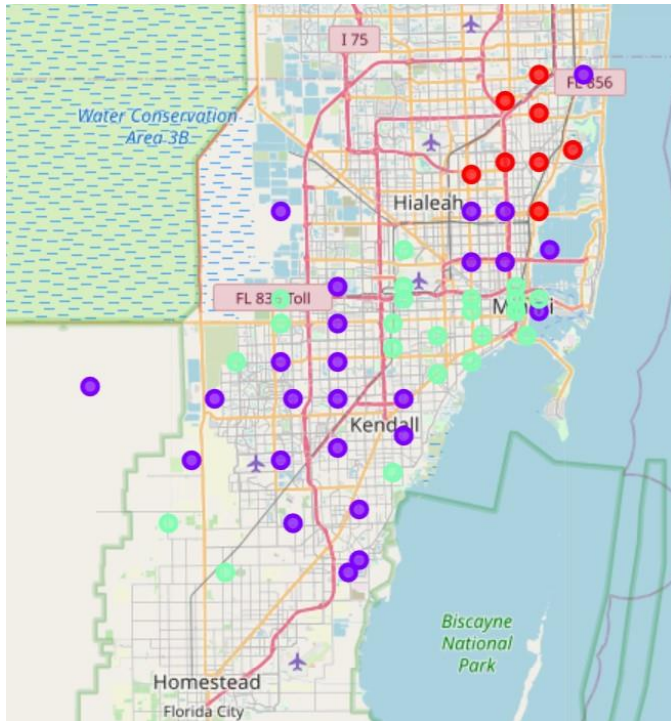
- Cluster 0: red, relatively fewer indoor facilities;
- Cluster 1: purple, medium number of indoor facilities;
- Cluster 2: green, relatively more indoor facilities.

	Indoor Facilities			
	mean	min	max	count
Cluster Labels				
0	0.102500	0.08	0.12	8
1	0.165700	0.14	0.18	25
2	0.201304	0.19	0.23	23

# Clustering for outdoor facilities
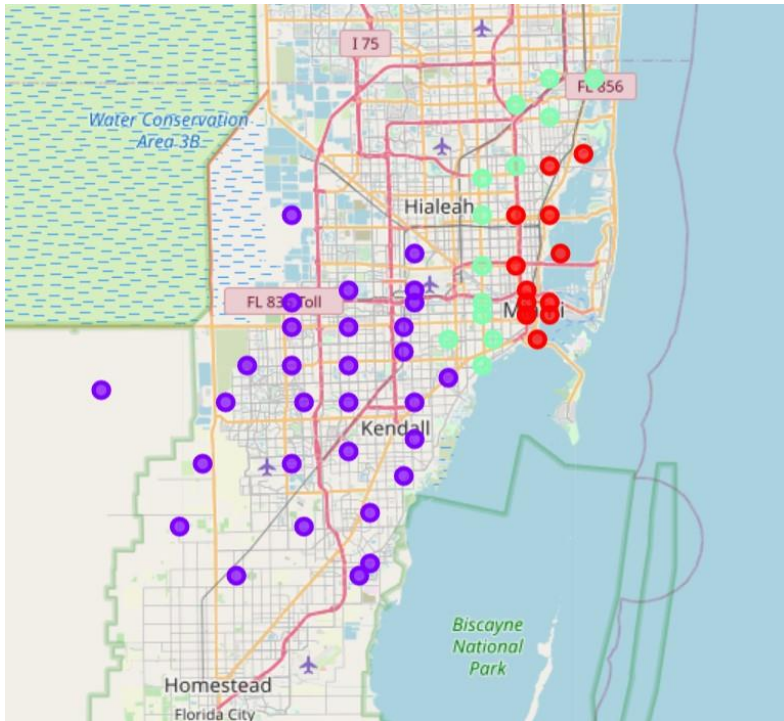
- Cluster 0: red, relatively more outdoor facilities;
- Cluster 1: purple, relatively few outdoor facilities;
- Cluster 2: green, medium number of outdoor facilities.

Outdoor Facilities

Cluster Labels	mean	min	max	count
0	0.503846	0.45	0.55	13
1	0.205333	0.15	0.28	30
2	0.377692	0.31	0.42	13

Result

The zips in most of the clusters with relatively more facilities:

	zip	Hospital Facilities	Grocery Facilities	Indoor Facilities	Outdoor Facilities	Total	primary_city	county	latitude	longitude	irs_estimated_population_2015
0	33144	1	1	1	0	3	Miami	Miami-Dade County	25.76	-80.31	23450
1	33155	1	1	1	0	3	Miami	Miami-Dade County	25.74	-80.31	41220
2	33158	1	1	1	0	3	Miami	Miami-Dade County	25.64	-80.31	6430
3	33170	1	1	1	0	3	Miami	Miami-Dade County	25.56	-80.46	12430
4	33182	1	1	1	0	3	Miami	Miami-Dade County	25.78	-80.41	13380
5	33184	1	1	1	0	3	Miami	Miami-Dade County	25.76	-80.41	19510
6	33185	1	1	1	0	3	Miami	Miami-Dade County	25.73	-80.45	26580
7	33187	1	1	1	0	3	Miami	Miami-Dade County	25.60	-80.51	17200
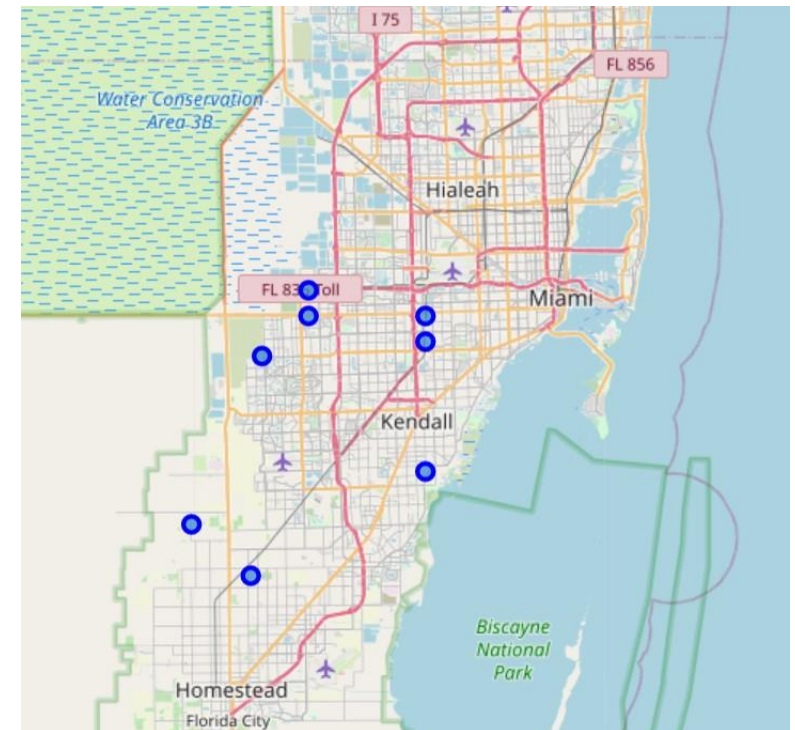
Discussion and Conclusion

- Found top zips with the highest numbers of total and average facilities for each category group, and plotted bar charts to demonstrate the numbers and to compare each venue in the category group for the top zips.

- With the k-means clustering method, for each category group, clustered the zips into 3 clusters based on the concentration, and marked them with different colors on the Miami map.

- Found the zip list appearing in most clusters having more facilities.

- The final decision on which zips of the city they would move to will be made by customers based on their individual or specific requirements/interests.

- With minor changes, this analysis can also be used by customers who would relocate to any city in the country with other preferences.

- Furthermore, can extend this project to analyze and compare the neighborhoods within each zip on the final list.