

# **K-Means Clustering of Municipalities Based on Venue Categories in Miami-Dade County**

*(The Capstone Project of IBM Professional Certificate Courses at Coursera)*

Kangrui Lu

## **Table of Contents**

1. [Abstract](#)
2. [Introduction](#)
3. [Data and Methodology](#)
4. [Results](#)
5. [Conclusion and Discussions](#)
6. [Reference](#)

## Abstract

Miami-Dade County, known as Miami or Dade, is often recognized as one of the most culturally diverse and congregated area in the country. We used K-Means method to cluster the 35 municipalities of Miami-Dade County into 5 distinct groups based on the business venues located in each area. The findings would facilitate future city planning, distribution of resources, and site selection for future business initiative or development.

## Introduction

Miami-Dade County, located in south Florida, is the most populous county in the state, and one of the most populous county in the United States (United States Census Bureau, 2019). Besides its enormous population body, Miami-Dade County is also a hub for variety of businesses ranging from hospitality to financial institutions. There are 35 municipalities within the boundary of Miami-Dade, which are distinct in terms of population density, type of businesses, and lifestyle. In this study, we aimed to utilize K-Means method and the Foursquare location data to cluster these municipalities or cities into different groups based on the venues located in each area.

## Data Acquisition and Methodology

### 1) Data Acquisition and Exploratory Analysis

Firstly, I got a list of municipalities (or cities) in Miami-Dade County (Miami-Dade County GIS Open Data, 2019) from the Miami-Dade County Open Data Hub. There are 35 municipalities in Miami-Dade, including an 'Unincorporated Miami-Dade' area, which covers

mostly the Everglades National Park. This area was later removed from the dataset for analysis, given this area is uninhabited.

Then, considering a list of coordinates for each of the municipalities are difficult to find, I used the 'Nominatim' function and Foursquare API in a loop to generate the corresponding coordinates for each municipality according to their names, and insert the coordinates into the dataset.

To searching for venues, I defined a function 'getNearbyVenues' to loop through Foursquare API and get 100 venues for each municipality and insert back into the dataset. Within the Miami-Dade County, 6 municipalities have the number of venues under 100, with Sweetwater area having the least of 45 venues.

I collected the data of population by municipality of Miami-Dade County (United States Census Bureau, 2010) from the US Census Bureau and merged with the municipality dataset for later analysis.

## **2) Methodology**

To examine each municipality and conduct clustering, I used the K-means method with 5 clusters. Firstly, I calculated the frequency of occurrence of each venue category by municipality. Then, I selected the top 10 most common venue categories for each area by defining a 'most\_common\_venues' function.

Next, I applied the K-Means clustering method to cluster all the municipalities into 5 clusters based on the category of venues. The number of clusters was chosen based on the elbow methods, where 6 cluster would over split the municipalities, and 4 clusters would not properly group them into clusters with similar attributes.

To map the results, I used ‘folium’ function with geographic API municipality boundary GeoJSON file from Miami-Dade County Open Data Hub (Miami-Dade County GIS Open Data, 2020), and build the choropleth map of population by municipality. Then, I added markers according to the clusters onto the choropleth map.

## Results

The map below shows the clusters on top of population distribution within the Miami-Dade County. The different colors of markers resemble different clusters, and the colored area underneath shows the population of each municipality.

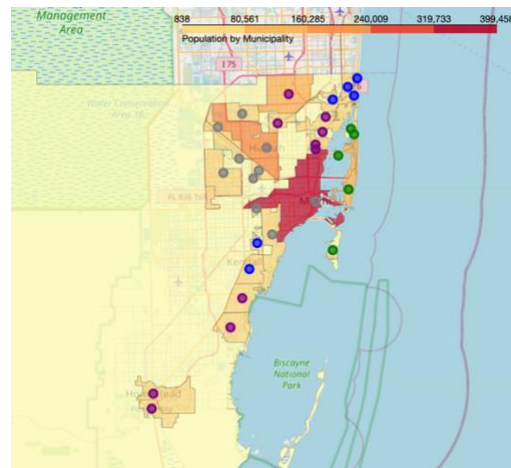


Figure – 1 (Map of Clusters with Population by Municipality)

Municipality	SWEETWATER
1st Most Common Venue	Hotel
2nd Most Common Venue	Pizza Place
3rd Most Common Venue	Fast Food Restaurant
4th Most Common Venue	Convenience Store
5th Most Common Venue	American Restaurant
6th Most Common Venue	Gas Station
7th Most Common Venue	Chinese Restaurant
8th Most Common Venue	Sandwich Place
9th Most Common Venue	Discount Store
10th Most Common Venue	BBQ Joint

Figure – 2 (Cluster 1 overview)

The first cluster contains only the Sweetwater municipality, where several colleges, shopping malls, and the Miami International Airport are nearby. Therefore, the most common venue in this area is hotel along with restaurants and stores.

Municipality	NORTH MIAMI BEACH	AVENTURA	SOUTH MIAMI	PINECREST	INDIAN CREEK VILLAGE	SUNNY ISLES BEACH	GOLDEN BEACH
1st Most Common Venue	Coffee Shop	Clothing Store	American Restaurant	Coffee Shop	Coffee Shop	Beach	Clothing Store
2nd Most Common Venue	Mexican Restaurant	Ice Cream Shop	Bakery	Grocery Store	Pizza Place	Clothing Store	Furniture / Home Store
3rd Most Common Venue	Grocery Store	Coffee Shop	Coffee Shop	Park	Brewery	American Restaurant	Beach
4th Most Common Venue	Clothing Store	Cosmetics Shop	Shopping Mall	American Restaurant	Park	Coffee Shop	Cosmetics Shop
5th Most Common Venue	Sushi Restaurant	American Restaurant	Department Store	Bakery	Hotel	Department Store	American Restaurant
6th Most Common Venue	Chinese Restaurant	Grocery Store	Clothing Store	Department Store	American Restaurant	Ice Cream Shop	Grocery Store
7th Most Common Venue	Ice Cream Shop	Italian Restaurant	Asian Restaurant	Clothing Store	Mexican Restaurant	Gym	Hotel
8th Most Common Venue	Cosmetics Shop	Gym	Grocery Store	Furniture / Home Store	Café	Grocery Store	Burger Joint
9th Most Common Venue	Fast Food Restaurant	Furniture / Home Store	Pharmacy	Pharmacy	Restaurant	Toy / Game Store	Italian Restaurant
10th Most Common Venue	Middle Eastern Restaurant	Department Store	Cuban Restaurant	Electronics Store	State / Provincial Park	Sushi Restaurant	Japanese Restaurant

Figure – 3 (Cluster 2 overview)

The second cluster (blue markers on Figure - 1) contains a group of cities around North Miami Beach as well as the Kendall area. Several major shopping plaza and malls are located in these areas; therefore, clothing stores and coffee shops are relatively dense here.

Municipality	MIAMI SPRINGS	MIAMI LAKES	VIRGINIA GARDENS	HIALEAH	HIALEAH GARDENS	DORAL	MIAMI	MEDLEY	WEST MIAMI	CORAL GABLES
1st Most Common Venue	Cuban Restaurant	Coffee Shop	Hotel	Cuban Restaurant	Pizza Place	Grocery Store	Hotel	Cuban Restaurant	Cuban Restaurant	Italian Restaurant
2nd Most Common Venue	Hotel	Bakery	Cuban Restaurant	Bakery	Sandwich Place	Hotel	Italian Restaurant	Park	Bakery	American Restaurant
3rd Most Common Venue	Grocery Store	Cuban Restaurant	South American Restaurant	Sandwich Place	Cuban Restaurant	Latin American Restaurant	Seafood Restaurant	Hotel	Latin American Restaurant	Gym / Fitness Center
4th Most Common Venue	Coffee Shop	Pharmacy	Arepas Restaurant	Grocery Store	Fast Food Restaurant	Coffee Shop	Japanese Restaurant	South American Restaurant	Grocery Store	Ice Cream Shop
5th Most Common Venue	Mexican Restaurant	Ice Cream Shop	Mexican Restaurant	Latin American Restaurant	Pharmacy	Golf Course	American Restaurant	Latin American Restaurant	Café	Movie Theater
6th Most Common Venue	American Restaurant	Mexican Restaurant	Sushi Restaurant	Donut Shop	Grocery Store	Italian Restaurant	Park	Grocery Store	Italian Restaurant	Bakery
7th Most Common Venue	Food Truck	Grocery Store	Restaurant	Pharmacy	Latin American Restaurant	Japanese Restaurant	Peruvian Restaurant	Fast Food Restaurant	Restaurant	Cuban Restaurant
8th Most Common Venue	Italian Restaurant	Latin American Restaurant	Pizza Place	Coffee Shop	Hotel	Pharmacy	Grocery Store	Mexican Restaurant	Seafood Restaurant	Coffee Shop
9th Most Common Venue	Sandwich Place	Steakhouse	Coffee Shop	Italian Restaurant	Mexican Restaurant	Arepas Restaurant	Pizza Place	Coffee Shop	Pizza Place	Pizza Place
10th Most Common Venue	Breakfast Spot	Wings Joint	Pharmacy	Burger Joint	Bakery	Argentinian Restaurant	Lounge	Frozen Yogurt Shop	Bookstore	Burger Joint

Figure – 4 (Cluster 3 overview)

The third cluster (grey markers) locates at the central Miami, Coral Gables, and Hialeah areas, where most of the population reside. Given the local Cuban Americans comprise 35.4% of the county's population (United States Census Bureau, 2015), plus several Cuban and Latin American neighborhood, including Little Havana and Little Haiti locate here, the popularity of Cuban and Latin American restaurants is remarkably popular in this cluster of municipalities.

Besides, areas like downtown Miami, Coral Gables, and Virginia Gardens are popular travel destinations and entertainment sites, therefore boosting the hotel and other hospitality businesses in these places.

Municipality	NORTH BAY VILLAGE	KEY BISCAIYNE	MIAMI BEACH	SURFSIDE	BAY HARBOR ISLANDS	BAL HARBOUR
1st Most Common Venue	Beach	Aquarium	Hotel	Beach	Beach	Café
2nd Most Common Venue	Italian Restaurant	Beach	Beach	Italian Restaurant	Italian Restaurant	Coffee Shop
3rd Most Common Venue	Park	Park	Bakery	Hotel	Grocery Store	Hotel
4th Most Common Venue	Burger Joint	Argentinian Restaurant	Sushi Restaurant	Coffee Shop	Coffee Shop	Resort
5th Most Common Venue	American Restaurant	Italian Restaurant	Grocery Store	Resort	Park	Beach
6th Most Common Venue	Japanese Restaurant	Restaurant	Cocktail Bar	Burger Joint	Hotel	Supermarket
7th Most Common Venue	Pizza Place	Bakery	Art Gallery	Bakery	Burger Joint	Shopping Mall
8th Most Common Venue	Golf Course	Tennis Stadium	Gym / Fitness Center	Park	Resort	Golf Course
9th Most Common Venue	Bistro	Bar	Park	Grocery Store	Peruvian Restaurant	Flea Market
10th Most Common Venue	Brazilian Restaurant	Coffee Shop	Peruvian Restaurant	Peruvian Restaurant	Department Store	Mexican Restaurant

Figure – 5 (Cluster 4 overview)

The fourth cluster (green markers) congregated on the islands, such as Miami Beach, and Key Biscayne, which are among the most visited travel destinations in the world. Therefore, beaches, hotels, resorts, and a variety of restaurants with extremely diverse cuisines are found.

Municipality	OPA-LOCKA	BISCAIYNE PARK	MIAMI GARDENS	HOMESTEAD	FLORIDA CITY	CUTLER BAY	MIAMI SHORES	PALMETTO BAY	NORTH MIAMI	EL PORTAL
1st Most Common Venue	Sandwich Place	Grocery Store	Caribbean Restaurant	Mexican Restaurant	Mexican Restaurant	Sandwich Place	Italian Restaurant	Grocery Store	Grocery Store	Italian Restaurant
2nd Most Common Venue	Seafood Restaurant	Italian Restaurant	Fast Food Restaurant	Grocery Store	Fast Food Restaurant	Grocery Store	Park	Burger Joint	Bakery	Pizza Place
3rd Most Common Venue	Fast Food Restaurant	Coffee Shop	Seafood Restaurant	American Restaurant	Pizza Place	Pharmacy	Pizza Place	Coffee Shop	Coffee Shop	Park
4th Most Common Venue	Discount Store	Pizza Place	Chinese Restaurant	Hotel	American Restaurant	Park	Burger Joint	Bakery	Cuban Restaurant	American Restaurant
5th Most Common Venue	Donut Shop	Burger Joint	Sandwich Place	Pizza Place	Gas Station	Seafood Restaurant	Gym	American Restaurant	Chinese Restaurant	Bistro
6th Most Common Venue	Pizza Place	Park	Grocery Store	Pharmacy	Sandwich Place	Cuban Restaurant	Grocery Store	Park	Mexican Restaurant	Gym
7th Most Common Venue	Fried Chicken Joint	Cuban Restaurant	Furniture / Home Store	Fast Food Restaurant	Donut Shop	Gym / Fitness Center	Bistro	Sandwich Place	Middle Eastern Restaurant	Grocery Store
8th Most Common Venue	Bakery	Spa	Pharmacy	Donut Shop	Hotel	Wings Joint	Caribbean Restaurant	Italian Restaurant	Pizza Place	Coffee Shop
9th Most Common Venue	Cuban Restaurant	Seafood Restaurant	Donut Shop	Steakhouse	Grocery Store	Ice Cream Shop	Cuban Restaurant	Gym / Fitness Center	Pet Store	Café
10th Most Common Venue	Grocery Store	Bistro	Park	Discount Store	Hardware Store	Pizza Place	Tapas Restaurant	Breakfast Spot	Restaurant	Caribbean Restaurant

Figure – 6 (Cluster 5 overview)

The fifth cluster (purple markers) locates both in the upper and in the very south side of the county. Those areas densely contain residential neighborhoods and small businesses. Therefore, grocery stores and restaurants are most common among these areas.

## Conclusion and Future Discussions

By examining the 5 clusters, we can see that, among the populated areas, restaurant, hotel, and grocery store, are three vital venue categories in Miami-Dade County. Moreover, hospitality industry plays a major role at the county, where diversity is greatly embraced. K-Means method created a clear clustering of venues categories for the municipalities within Miami-Dade County, which would be a reference for future businesses at Miami-Dade considering the popularity of type of business and population of each area. Possible future discussions on this topic could involve the tryout of Density-Based method for clustering the municipalities by examining the density of different venue categories or even customer feedback for specific venues.

## Reference

Miami-Dade County GIS Open Data. (2019, 2 21). Municipality. Retrieved from Open Data

Hub: <https://gis-mdc.opendata.arcgis.com/datasets/municipality>

Miami-Dade County GIS Open Data. (2020, 6 1). Municipal Coastal Boundary. Retrieved from

Open Data Hub: [https://gis-mdc.opendata.arcgis.com/datasets/municipal-coastal-](https://gis-mdc.opendata.arcgis.com/datasets/municipal-coastal-boundary/data)

boundary/data

United States Census Bureau. (2010). Population and Housing Occupancy Status: 2010 - County

- County Subdivision and Place for Florida. Retrieved from United States Census Bureau:

<https://www.census.gov/>

United States Census Bureau. (2015, 10 19). Miami-Dade County: Selected Social

Characteristics in the United States 2006–2010 American Community Survey 5-Year

Estimates. Retrieved from United States Census Bureau: <https://www.census.gov/>

United States Census Bureau. (2019, 7 1). QuickFacts Miami-Dade County, Florida; United

States. Retrieved from United States Census Bureau:

<https://www.census.gov/quickfacts/fact/table/miamidadecountyflorida,US/PST045219>