Capstone Project

The Battle of Neighborhoods - Presentation

Applied Data Science Capstone by IBM / Coursera

Stephenas Ali

2019

Introduction: Business Problem

Target Audience:

- Stakeholders interested in opening an **coffee shop** near to **Downtown Miami**, in Florida.

Objectives:

- To detect **the most suitable location, in or around the surrounding neighborhoods**
- Select areas with **no existing coffee shops**
- Select areas with **popular venues** or businesses to attract potential customers

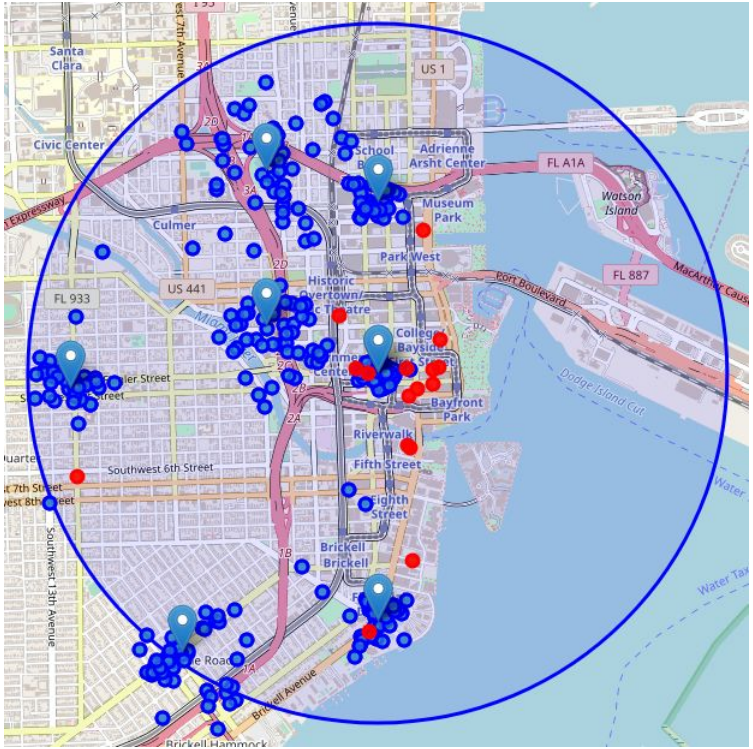
Data

Information will be gathered from the following data sources:

- A **Wikipedia** entry with a list of neighborhoods in Miami.
- **Foursquare API** - To identify existing coffee shops and other venues of interest.

Summary of Data

	Neighborhood	Population2010	Population/Km²	Latitude	Longitude
0	Brickell	31759	14541	25.758	-80.193
1	Downtown	71,000 (13,635 CBD only)	10613	25.774	-80.193
2	Little Havana	76163	8423	25.773	-80.215
3	Lummus Park	3027	3680	25.777	-80.201
4	Overtown	6736	3405	25.787	-80.201
5	Park West	4655	3635	25.785	-80.193
6	The Roads	7327	4899	25.756	-80.207



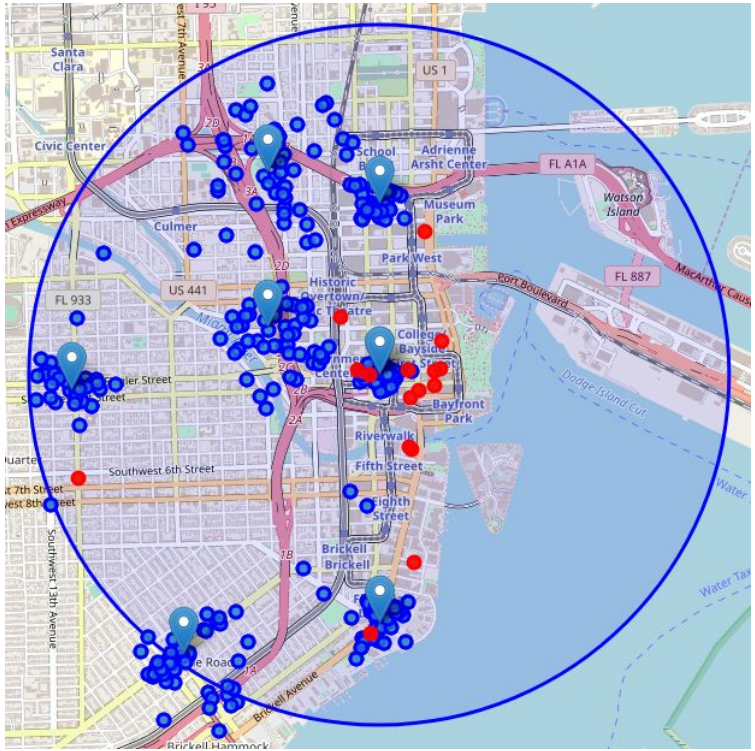
- Total number of places of interests: 350
- Total number of coffee shops: 16
- Percentage of coffee shops to general places of interests: 4.5
- Average number of coffee shops per neighborhood: 2.286

Methodology

1. Select the neighborhoods closest to Downtown Miami
2. Categorize each neighborhood relative to the average number of Coffee Shops
3. Select neighborhood with the lowest number of Coffee Shops
4. Recommend the best location to setup a new Coffee Shop in the selected neighborhood

1. Select the Neighborhood

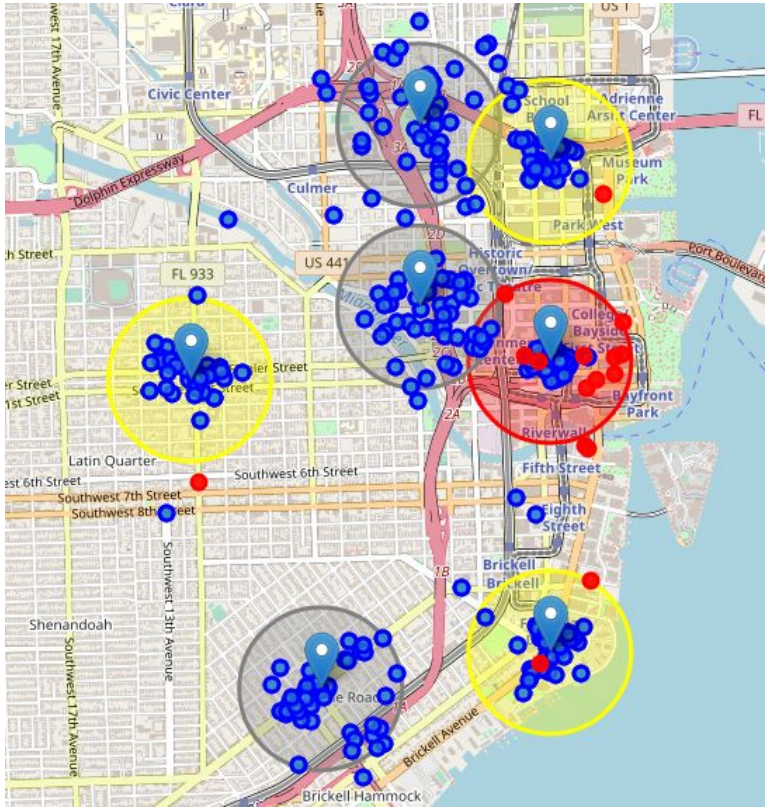
The haversine formula was used to calculate and select the most suitable neighborhoods closest to Downtown Miami



2. Neighborhood Categorization

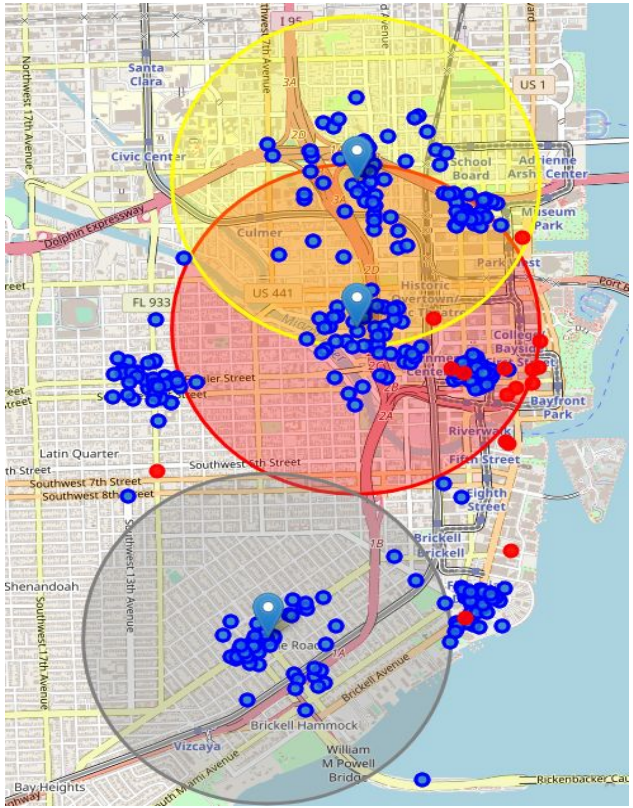
Each neighborhood were categorized according to the average number of coffee shops. The categories included:

1. Red - Above average
2. Yellow - Below average
3. Gray - No Coffee Shops



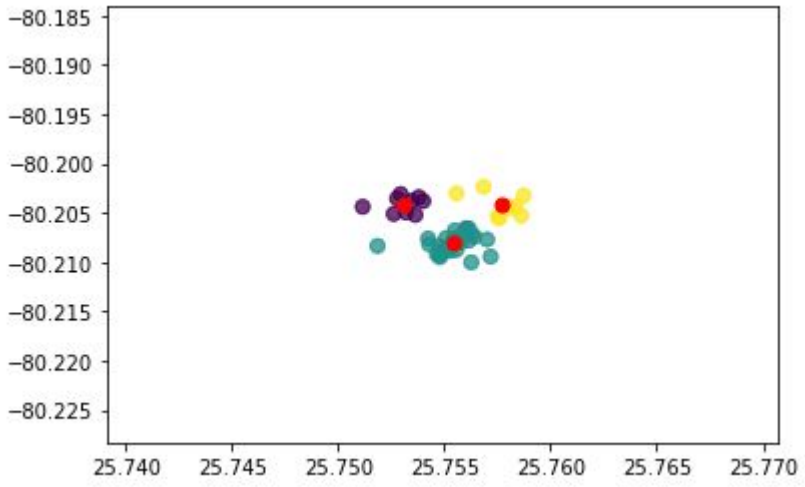
2. Neighborhood Categorization

- The neighborhoods with no coffee shops were selected and the procedure was repeated.
- The radial distance was increased from 500 meters to 1.25km to find coffee shops in a wider range.
- And the data was then re-categorized.



3. Neighborhood with Lowest No. of Coffee Shops

- Eventually neighborhood **The Roads** was selected
- K-means clustering was then used to find the largest cluster of venues of interest.
- A cluster of 3 was applied on the venues dataset and a simple scatter plot was created.

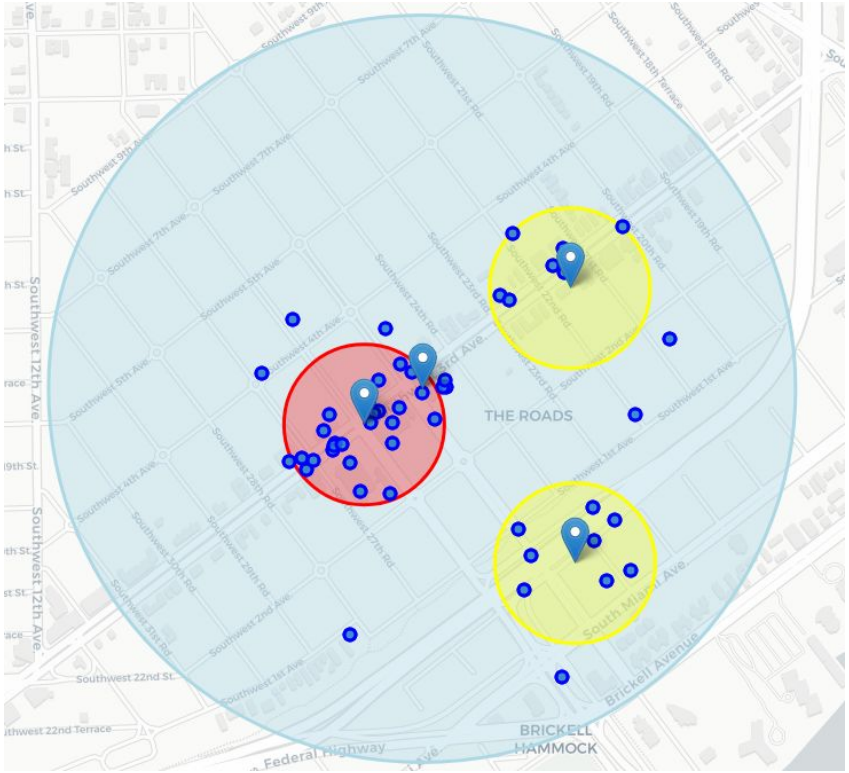


	Latitude	Longitude	Centroid	Venues	Category
0	25.753058	-80.204135	Centroid 1	16	yellow
1	25.755396	-80.208083	Centroid 2	48	red
2	25.757696	-80.204231	Centroid 3	12	yellow

4. Recommend Best Location

The centroids created by the k-means algorithm were then categorized and plotted onto the map.

Using the map we can tell that the best location to open a new coffee shop may be in the neighborhood of **The Roads**, at the intersection of **SW 3rd Avenue** and **SW 26th Road**.



Conclusion

Purpose of this project was to aid stakeholders in narrowing down the search for optimal location for a new coffee shop in the vicinity of Downtown Miami. To accomplish this we:

- Identified prospective neighborhoods, and then identified existing coffee shop distributions and supporting businesses using Foursquare data.
- Applied various predictive models to determine the most viable option for a new coffee shop.
- Then recommended solution satisfying low competition, supportive business establishments and whilst also being close to Downtown Miami.