The Battle of Neighborhoods

Introduction

Cairo is a sprawling city, and like you probably noticed in other big cities, not all neighborhoods were created equal.

The aim of this project is to use Foursquare location data to explore and compare neighborhoods of Cairo and use these data to solve a problem.

Problem Description

If a company is trying to start their own business in Cairo, where would you recommend that they setup their office? It would depend on a lot of factors like what is the type of business? and what are the target customers?

To help answering these questions we need to know what the most common venues in every neighborhood in Cairo are and what are the similarities and differences between these neighborhoods.

In this project we will cluster neighborhoods based on the most common venues like Hotels, Historic sites, restaurants, or parks. Which will help any company decide where to locate their office based on the pattern of location of these venues.

Target Audiences

Companies or individuals who are trying to find the best neighborhood in Cairo to locate their office based on the most common venues in each neighborhood.

Data description

I collected the all the neighborhoods in Cairo and their location data from Wikipedia in a CSV file. Then plugged these locations in Foursquare API to get all the venues in every neighborhood within 5 km.
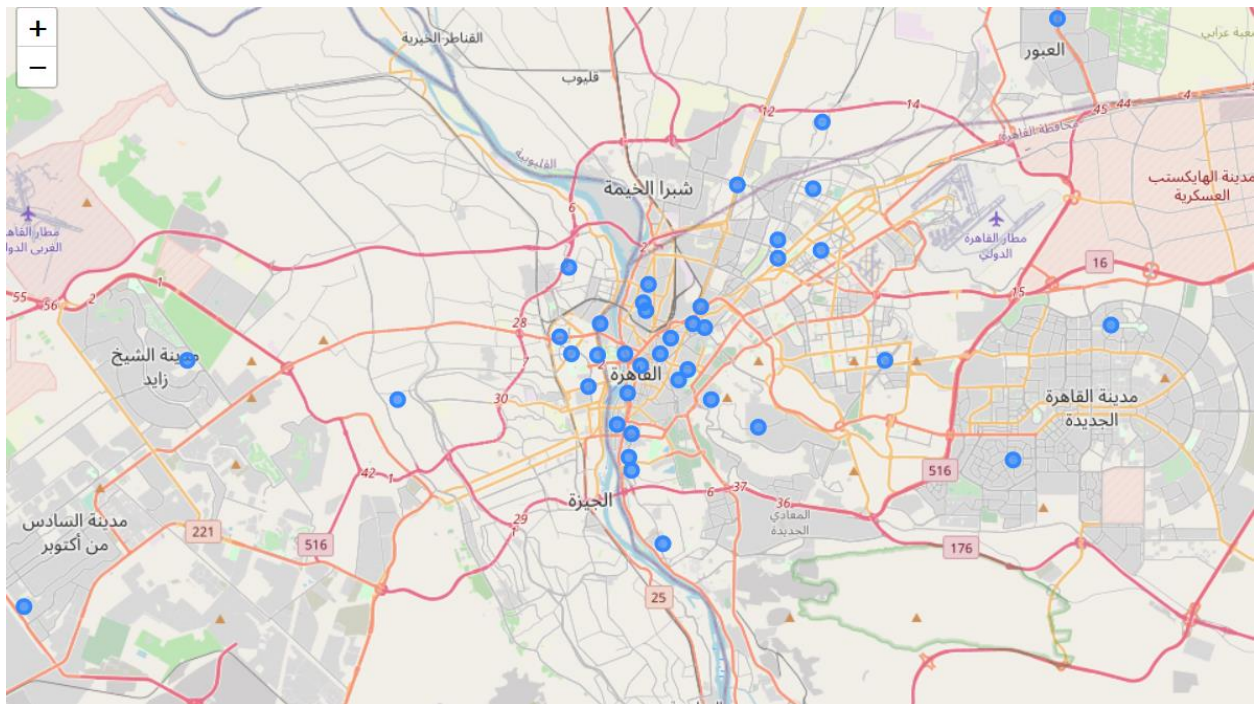
Foursquare API gives a lot of information about every venue, but I only used the Category description of the venue.

Methodology

Cairo's neighborhoods and their latitudes and longitudes was collected from Wikipedia and cleaned and saved into a CSV file.

	Neighborhood	lat	lng
0	Abbassia	30.065008	31.271710
1	Ain Shams	30.127778	31.329167
2	Azbakeya	30.052778	31.248333
3	Bulaq	30.053000	31.230000
4	Daher	30.074167	31.270000

This is a Folium map of the Cairo Neighborhoods before clustering them into groups.

Then used Foursquare API to explore areas around center of each neighborhood and find 10 most common venues in each neighborhood. I've set the radius of search to 5000 meters around the center of each neighborhood.

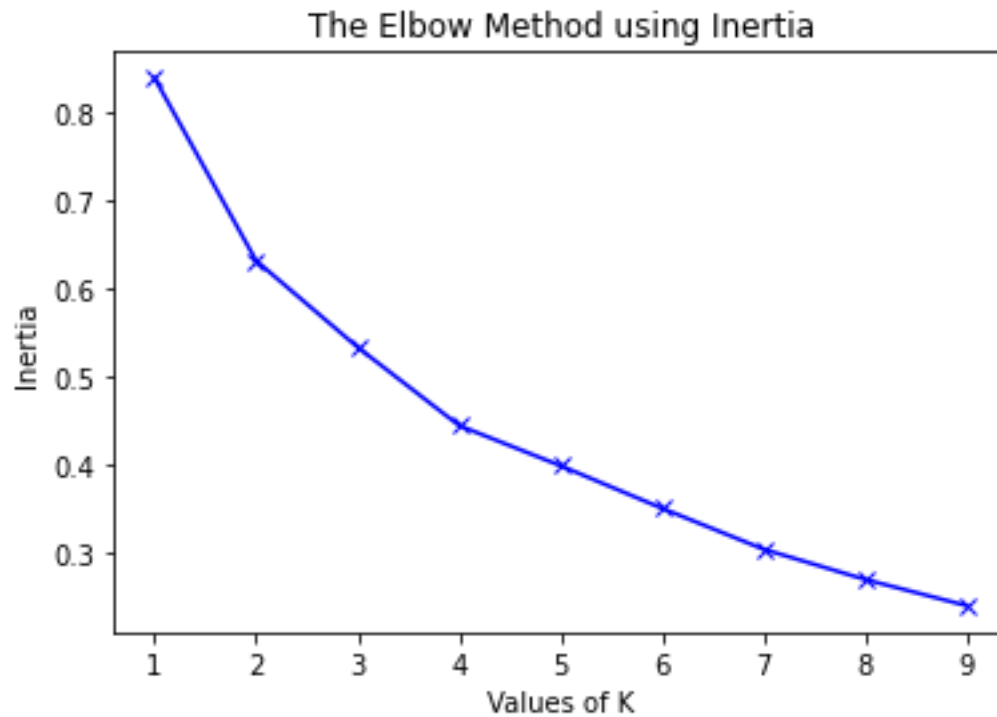This is an example of the search result merged with neighborhood data:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Abbassia	30.065008	31.27171	Al Haakem Mosque (جامع الحاكم بأمر الله)	30.054612	31.263587	Mosque
1	Abbassia	30.065008	31.27171	Bab Al-Fotouh (باب الفُتوح)	30.055631	31.263534	Historic Site
2	Abbassia	30.065008	31.27171	Sakakini Palace (قصر السكاكيني)	30.065388	31.266456	History Museum
3	Abbassia	30.065008	31.27171	Bayt El Suhaymi (بيت السحيمى)	30.052086	31.261915	Arts & Entertainment
4	Abbassia	30.065008	31.27171	El Moez St (شارع المعز لدين الله)	30.050482	31.261851	Historic Site

The result was 3639 venues returned by Foursquare in Cairo. Next step was to create a table, which shows a list of top 10 venues for each neighborhood:

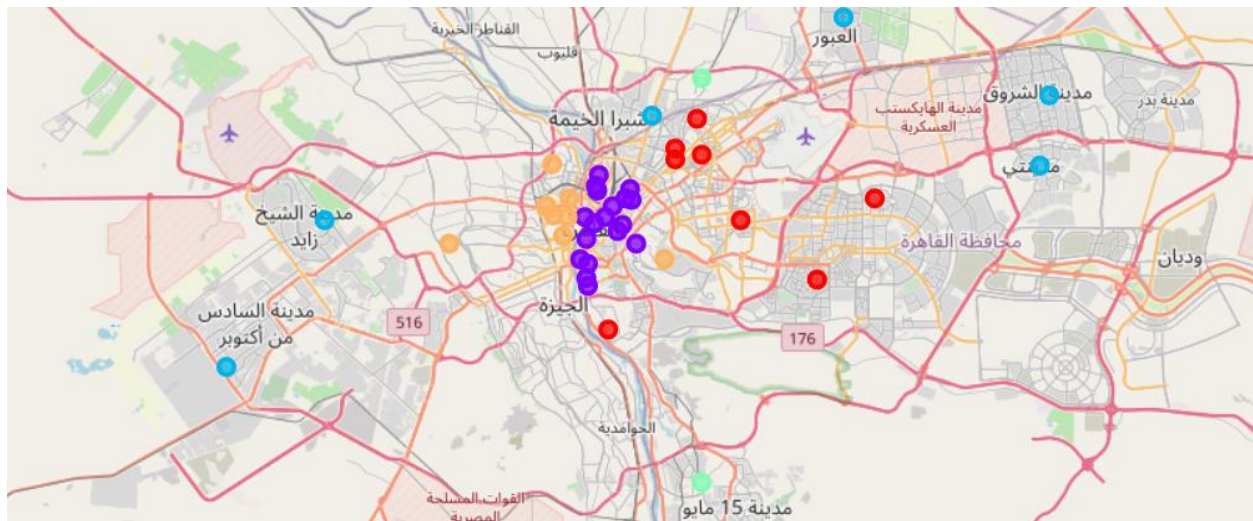|	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	6th of October	Fast Food Restaurant	Café	Kebab Restaurant	Syrian Restaurant	Plaza	Fried Chicken Joint	Coffee Shop	Pizza Place	Shopping Mall	Hookah Bar
1	Abbassia	Hotel	Café	Lounge	Egyptian Restaurant	Historic Site	Italian Restaurant	Mosque	Pastry Shop	Middle Eastern Restaurant	Bookstore
2	Agouza	Café	Hotel	Coffee Shop	Lounge	Sports Club	Sushi Restaurant	Middle Eastern Restaurant	Italian Restaurant	Restaurant	Seafood Restaurant
3	Ain Shams	Café	Convenience Store	Italian Restaurant	Dessert Shop	Gym / Fitness Center	Fast Food Restaurant	Coffee Shop	Sandwich Place	Lebanese Restaurant	Supermarket
4	Al-Darb al-Ahmar	Hotel	Café	Historic Site	Lounge	Egyptian Restaurant	Mosque	Theater	Bar	Hotel Bar	Pastry Shop

The next step is to cluster the neighborhood, so I used K-means clustering algorithm to group neighborhoods, based on the frequencies of venues in each neighborhood. These groups will be used to find similar neighborhoods in Cairo.

First we need to know the optimum number of clusters needed, so I used the Elbow method for optimal value of k in KMeans:

The Elbow Method using Inertia

From K-Means elbow method we can conclude that an optimal number of clusters would be 4.
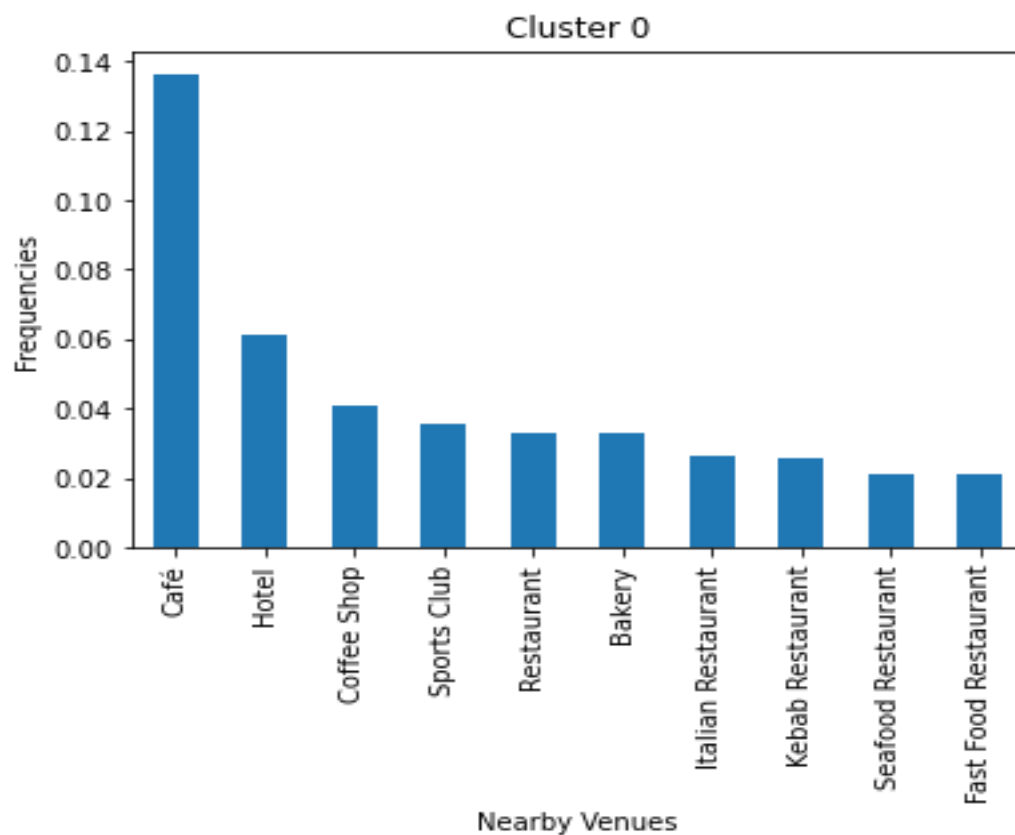
I ran the KMeans algorithm on the data and used Folium library to visualize the clustered neighborhoods on Cairo map.

Now we have a visual representation of clusters, and you can find the similar neighborhoods have the same colors.
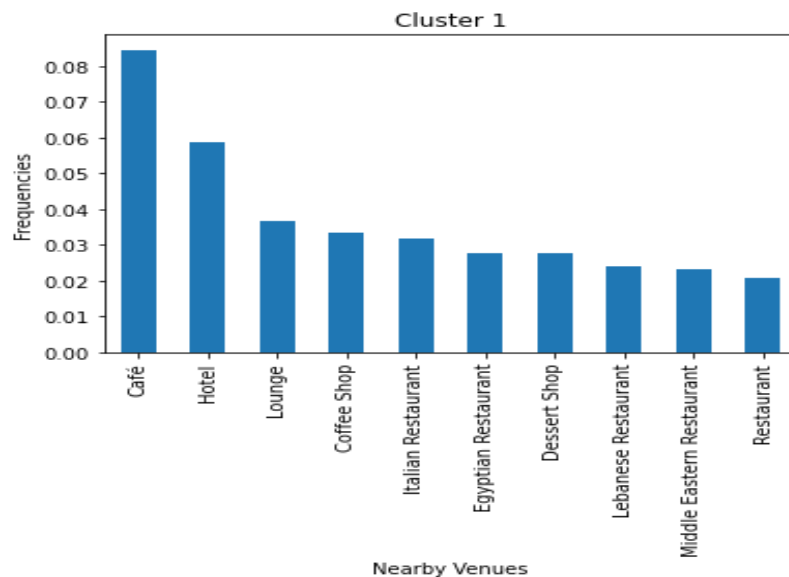
Cluster 0

	index	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	1	Ain Shams	0	Café	Convenience Store	Dessert Shop	Italian Restaurant	Fast Food Restaurant	Gym / Fitness Center	Plaza	Coffee Shop	Lebanese Restaurant	Supermarket
1	10	Nasr City	0	Dessert Shop	Coffee Shop	Café	Supermarket	Hotel	Lebanese Restaurant	Multiplex	Ice Cream Shop	Food Court	French Restaurant
2	11	El Qobbah	0	Café	Convenience Store	Dessert Shop	Italian Restaurant	Lebanese Restaurant	Coffee Shop	Hotel	Supermarket	Bookstore	Egyptian Restaurant
3	17	Zeitoun	0	Café	Convenience Store	Italian Restaurant	Coffee Shop	Lebanese Restaurant	Dessert Shop	Hotel	Bookstore	Falafel Restaurant	Egyptian Restaurant
4	26	Heliopolis	0	Café	Convenience Store	Italian Restaurant	Hotel	Dessert Shop	Lebanese Restaurant	Egyptian Restaurant	Falafel Restaurant	Fast Food Restaurant	Gym / Fitness Center
5	27	Maadi	0	Italian Restaurant	Café	Middle Eastern Restaurant	Coffee Shop	Restaurant	Ice Cream Shop	Historic Site	Supermarket	Bakery	Gym / Fitness Center
6	30	El Rehab	0	Café	Coffee Shop	Dessert Shop	Shopping Mall	Supermarket	Seafood Restaurant	Lebanese Restaurant	Food Court	Furniture / Home Store	Burger Joint
7	31	Fifth Settlement	0	Dessert Shop	Coffee Shop	Clothing Store	Italian Restaurant	Lebanese Restaurant	Supermarket	Shopping Mall	Furniture / Home Store	French Restaurant	Convenience Store

This cluster represent the nighorhoods like Nasr City, Fifth Settlemnet and Rehab. If you're craving modern trappings, as it has huge malls, shopping avenues and well-regarded restaurants.
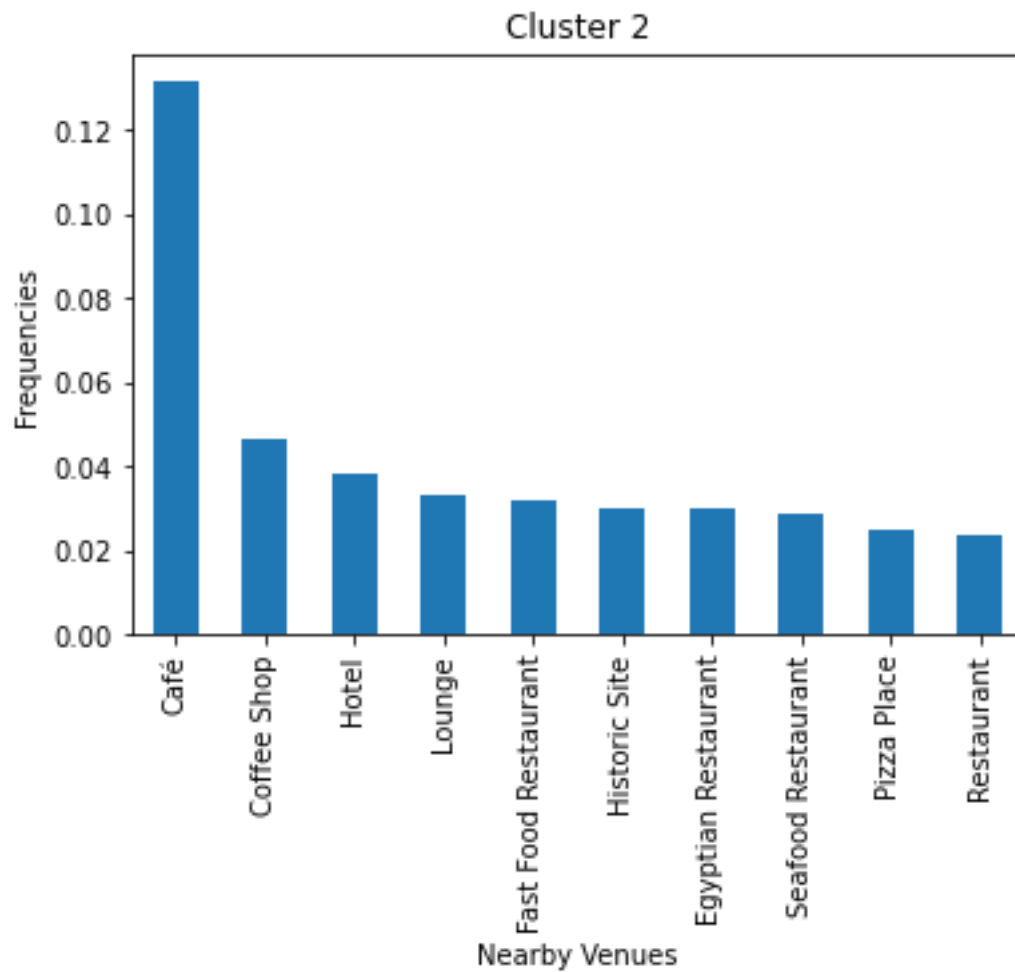
Cluster 1:

	index	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	0	Abbassia	1	Café	Hotel	Historic Site	Egyptian Restaurant	Lounge	Italian Restaurant	Middle Eastern Restaurant	Pastry Shop	Coffee Shop	Mosque
1	2	Azbakeya	1	Hotel	Café	Lounge	Egyptian Restaurant	Italian Restaurant	Middle Eastern Restaurant	Historic Site	Theater	Pastry Shop	Hotel Bar
2	3	Bulaq	1	Hotel	Lounge	Café	Italian Restaurant	Pastry Shop	Hotel Bar	Middle Eastern Restaurant	Sports Club	Seafood Restaurant	Pub
3	4	Daher	1	Café	Hotel	Egyptian Restaurant	Lounge	Restaurant	Coffee Shop	Historic Site	Pastry Shop	Italian Restaurant	Mosque
4	5	Al-Darb al-Ahmar	1	Hotel	Café	Historic Site	Lounge	Egyptian Restaurant	Hotel Bar	Theater	Pastry Shop	Mosque	Italian Restaurant
5	9	Manshiyat Naser	1	Café	Historic Site	Lounge	Hotel	Egyptian Restaurant	Kebab Restaurant	Juice Bar	Middle Eastern Restaurant	Italian Restaurant	Mosque
6	12	Rhoda Island	1	Lounge	Hotel	Historic Site	Lebanese Restaurant	Hotel Bar	Café	Nightclub	Dessert Shop	Seafood Restaurant	Gym / Fitness Center
7	13	El Sakkakini	1	Café	Hotel	Lounge	Egyptian Restaurant	Historic Site	Middle Eastern Restaurant	Italian Restaurant	Coffee Shop	Pastry Shop	Mosque
8	14	Shubra	1	Café	Lounge	Hotel	Egyptian Restaurant	Restaurant	Middle Eastern Restaurant	Pastry Shop	Sushi Restaurant	Coffee Shop	Italian Restaurant

Cluster 1 grouped most of the neighborhood east of the Nile river with Hotels, Lounges and Lebanese restaurants are the unique venues. Also this cluster represent old Cairo which is composed of the leftovers of cities that were capitals prior to Cairo, such as Fustat and al-Askar. This area also includes Coptic Cairo, which is known as the area that the Holy Family sought refuge in.
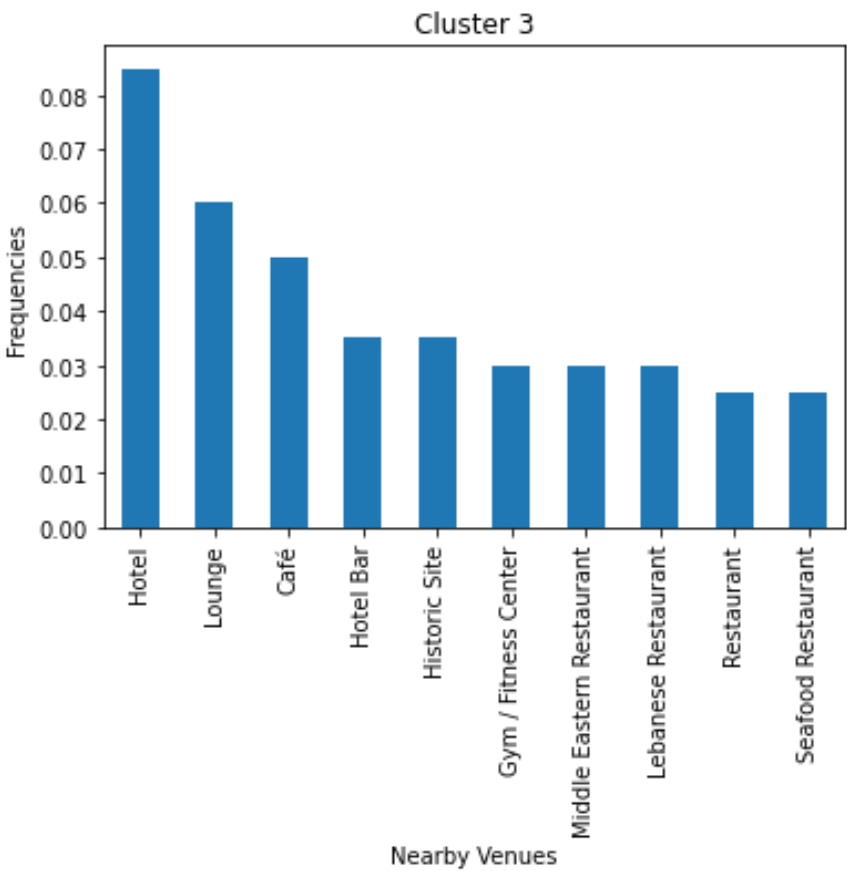
Cluster 2:

	index	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	23	El Matareya	2	Café	Supermarket	Bus Station	Pizza Place	Plaza	Convenience Store	Restaurant	Lighting Store	Gym	Fast Food Restaurant
1	29	Madinaty	2	Café	Coffee Shop	Kebab Restaurant	Supermarket	Plaza	Restaurant	Falafel Restaurant	Fast Food Restaurant	Middle Eastern Restaurant	Sports Club
2	32	El Shorouk	2	Shopping Mall	Café	Supermarket	Falafel Restaurant	Coffee Shop	Kebab Restaurant	Sports Club	Fast Food Restaurant	Campground	Pool
3	39	6th of October	2	Fast Food Restaurant	Café	Kebab Restaurant	Shopping Mall	Syrian Restaurant	Bakery	Coffee Shop	Fried Chicken Joint	Pizza Place	College Cafeteria
4	40	Sheikh Zayed City	2	Café	Coffee Shop	Shopping Mall	Sports Club	Supermarket	Fast Food Restaurant	Gym / Fitness Center	Food Court	Dessert Shop	Convenience Store
5	41	Obour (city)	2	Coffee Shop	Café	Kebab Restaurant	Supermarket	Shopping Mall	Italian Restaurant	Market	Club House	Movie Theater	Go Kart Track

Cluster 2 grouped most of the new neighborhoods and cities in New Cairo which share being away on the borders of Cairo.
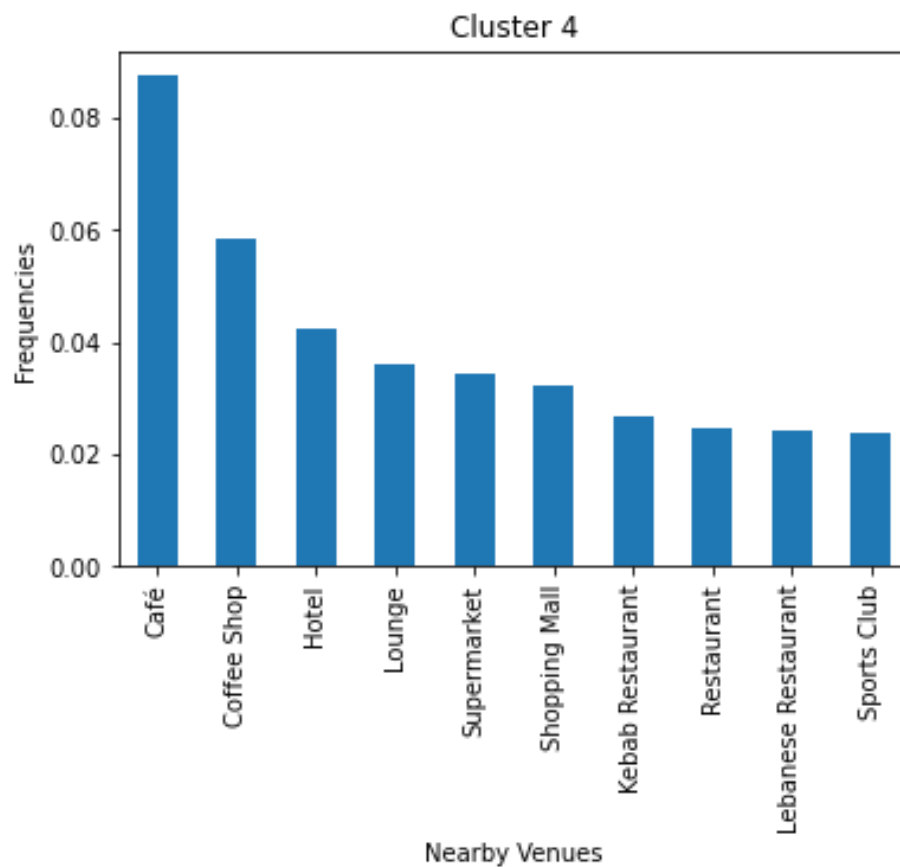
Cluster 3:

	index	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	6	El Marg	3	Café	Fast Food Restaurant	Bus Station	Sports Club	Falafel Restaurant	Coffee Shop	Pool	Gym	Metro Station	Fried Chicken Joint
1	7	Helwan	3	Café	Fried Chicken Joint	Park	Sports Club	Fast Food Restaurant	Seafood Restaurant	Coffee Shop	Pizza Place	Nightclub	Movie Theater

Cluster 4:

	index	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	8	Mokattam	4	Café	Historic Site	Dessert Shop	Restaurant	Fast Food Restaurant	Egyptian Restaurant	Coffee Shop	Juice Bar	Clothing Store	Malay Restaurant
1	28	Zamalek	4	Café	Hotel	Lounge	Coffee Shop	Sushi Restaurant	Middle Eastern Restaurant	Restaurant	Seafood Restaurant	Bar	Italian Restaurant
2	33	Agouza	4	Café	Hotel	Coffee Shop	Lounge	Seafood Restaurant	Sports Club	Sushi Restaurant	Middle Eastern Restaurant	Italian Restaurant	Restaurant
3	34	Mit Okba	4	Café	Coffee Shop	Seafood Restaurant	Hotel	Sushi Restaurant	Ice Cream Shop	Sports Club	Lebanese Restaurant	Middle Eastern Restaurant	Italian Restaurant
4	35	Imbaba	4	Café	Hotel	Middle Eastern Restaurant	Coffee Shop	Lounge	Egyptian Restaurant	Restaurant	Sushi Restaurant	Seafood Restaurant	Burger Joint
5	36	Kerdasa	4	Hotel	Café	Restaurant	Kebab Restaurant	Bakery	Falafel Restaurant	Coffee Shop	Flea Market	Sports Club	Mobile Phone Shop
6	37	Dokki	4	Café	Hotel	Sports Club	Lounge	Seafood Restaurant	Hotel Bar	Dessert Shop	Lebanese Restaurant	Sushi Restaurant	Coffee Shop
7	38	Mohandiseen	4	Café	Seafood Restaurant	Hotel	Coffee Shop	Italian Restaurant	Middle Eastern Restaurant	Sushi Restaurant	Ice Cream Shop	Sports Club	Bakery

Cluster 4 grouped most of the neighborhoods west of the Nile river which share the same venues as cluster 1 with Hotels, Lounges and Lebanese Restaurants are the unique venues.

Conclusion

Cairo boasts the largest city population in all of Africa and the Middle East, and navigating its vast districts can be a daunting prospect for any one. I've picked clustered the most similar neighborhoods, from ancient quarters to affluent modern districts, to help you make the most of any business decision of choosing a place to start your business or just planning a visit to one of these neighborhoods.

While the model didn't reach high level of accuracy in clustering these neighborhoods, its still a good start in the right direction.

The Foursquare API have a limited venues queries in the free version. By increasing the radius of search in every neighborhood you will be able to increase the number of venues queried from the API and increase the accuracy of the model in clustering the neighborhoods correctly.