

Dubai Community Selection



Mohammad Sukhon

Introduction

My friend owns a restaurant in Abu Dhabi inspired from the famous Game Of Thrones show. The restaurant has received great coverage from the media [1].

The interior design resembles that of a typical restaurant/bar in the North of Westeros. It consists of around 6 counter stools for visitors to come and enjoy a fast meal. The restaurant is also registered with delivery companies. Outbound deliveries account for more than 80% of their business. The main reason behind having low physical customers is the location; as the restaurant is in a rundown section of the city. My friend initially selected this location because he was able to get a great deal on rent. Now my friend has accrued some cash from operations and is interested in expansion. His initial idea was to relocate within Abu Dhabi to a more prime location, but I convinced him not to change the setting as it became part of his restaurant's brand in Abu Dhabi and hence relocating might not yield the highest ROI. Instead, I convinced him to expand operations in Dubai which has a higher population, much higher influx of tourists and is generally a lot more western than Abu Dhabi and hence would appeal more to the local population.

This time he wants to select an optimal location from the beginning and has requested my help. My recommendation to him was to go for a community that satisfies the following criteria:

- Large population density
- Large number of venues which indicates strong business sentiment
- Demographics:
 - Young, single or small family
 - Predominantly European, North American and Western mindset
- Large influx of business visitors and leisure tourists
- Surrounded by other similar communities to facilitate delivery / pick up
- Consists of popular destination such as a Zoo Safari, Beach, entertainment park..etc

We decided to resort to Data science and Machine Learning to help us segment the different communities in Dubai which'll narrow down our choices.

Data

To get census data about the different communities in Dubai, we used the following Wikipedia link:

https://en.wikipedia.org/wiki/List_of_communities_in_Dubai

The main table in this link consists of the following structure:

Community #	Community Name in English	Community Name in Arabic	Area	Population	Population Density

There are 130 diverse communities in Dubai. All the data types are of type object or string and hence further processing is needed especially on the population density column.

We plan on using K-means unsupervised Machine Learning method to segment our communities based on the prevalent venue types. Hence, we'll need to ultimately have a data frame consisting of all communities along with their corresponding top venues. To get the venue information we will be using the Foursquare API. This requires us to get the latitude/longitude data per community; we will use the geopy library to get that information.

Methodology

Using pandas function `read_html`, we imported the table into a data frame consisting of the structure highlighted in the previous section. We conducted a prescreening of the data to eliminate obvious communities from our selection such as Al Qusais Industrial which consists mainly of auto garages. This reduced our community set from 130 to 95. We also noticed that certain key neighborhoods were missing and hence had to manually add those to the data frame.

After the prescreening, we systematically cleansed the data following the approach below:

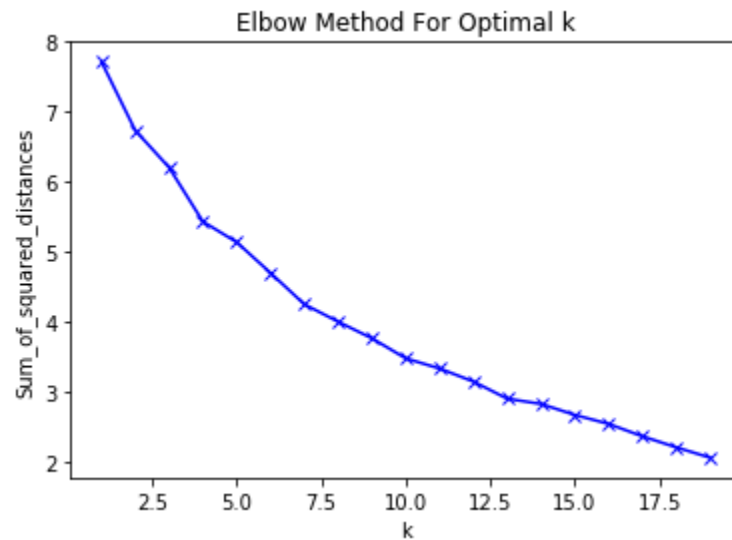
- Removed the `'/km2'` character from the population density column and converted the resulting number into a float
- Passed the data frame into the `geopy` function and identified a long list of “unrecognized” communities. We then adjusted the names of these communities to ones that the `geopy` function can accept. One rejected group category were the communities that had the numbers expressed in words such as `'Al Qouz First'`. Therefore, we had to create a function that extracts numerical words and converts them into numbers. The function is called `“wordcut”`. After all these iterations, a small number of communities remained unrecognized by `geopy` such as Al Bada. We did a small research on these communities and concluded they were not strategic to us and hence decided to discard them all. The new community count is now 88.

After cleansing the data set, we passed it through `geopy` and extracted per community the latitude / longitude information. We faced some timeout problems and had to increase the timeout parameter to 10 after which the `geopy` function worked smoothly.

To get the venue information, we used the `explore` API call; setting the radius to 1 km which results in an area of 3.14 Km² which is lower than the average community size of 8.9 km² minimizing the possibility of overlap. We extracted 4048 venues of which 304 were unique.

We used one-hot coding to convert the unique venue types into numerical variables and then per community computed the normalized frequency of occurrence of each venue type. This effectively added weight values for each unique venue category against each community. We will apply k-means to the venue category weights for the segmentation exercise.

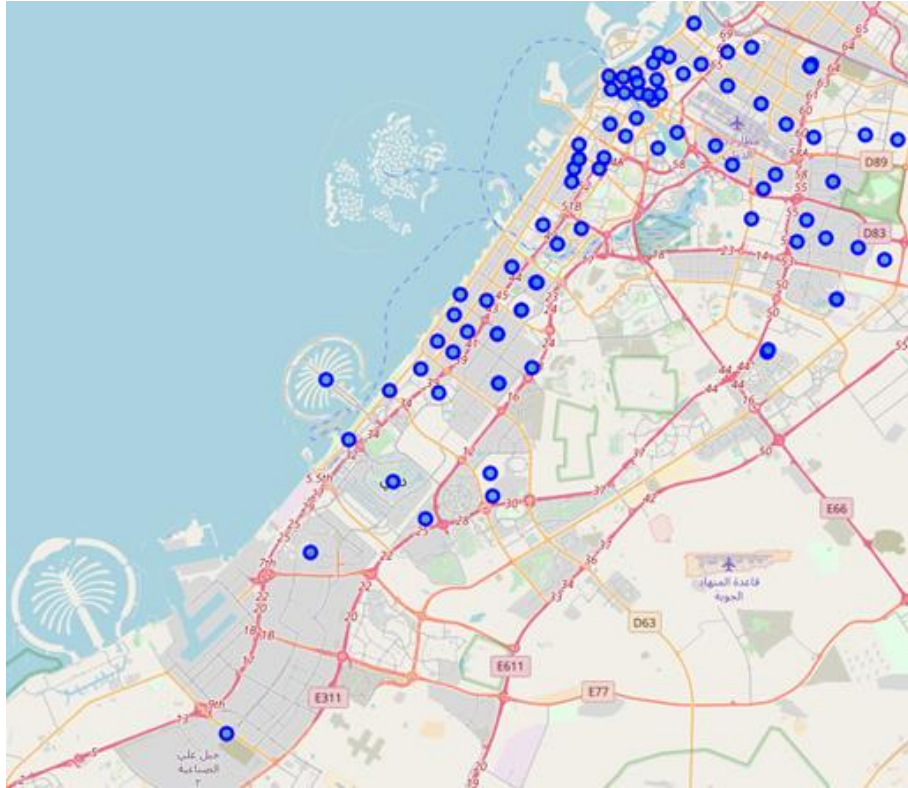
We started with a low K value of 5, this resulted in clusters which were hard to uniquely define. To arrive at an optimal K-value we decided to use the elbow method to plot the sum of distances square versus K. We got the following chart:



There are multiple elbow points, but we decided to go for K =13 given that the gradient precipitously drops afterwards and it has a low error value as well. At K = 13, the clusters made a lot more sense and we could uniquely describe them.

Results

The folium figure below displays the 88 communities in Dubai:



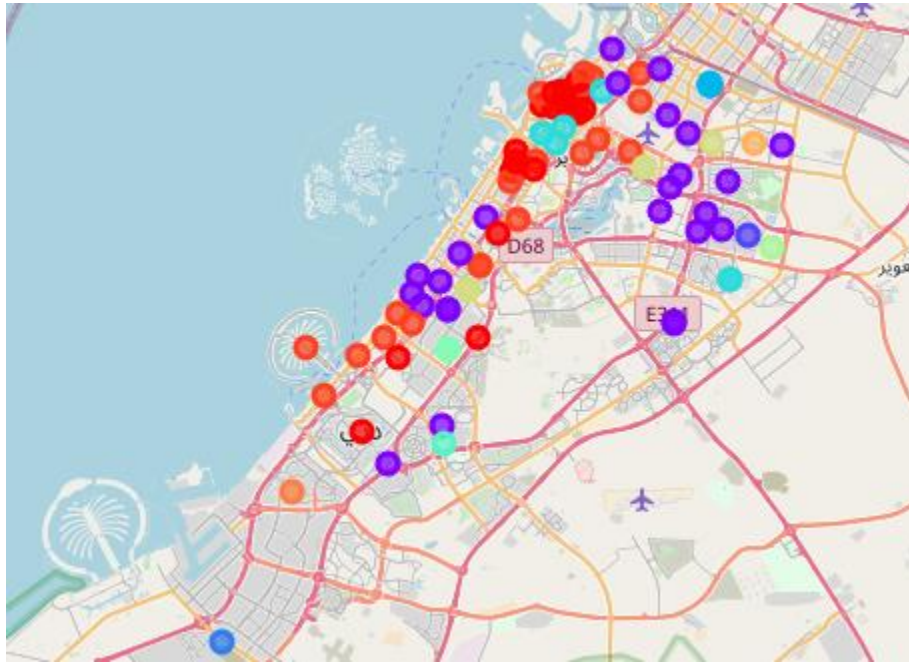
The communities below have the highest venue count in Dubai and will be given a priority within their clusters:

Rigga Al Buteen	Naif	Umm Hurair First	Port Saeed	Business Bay
Al Karama	Al Muraqqabat	Al Mina	Al Mankhool	Al Barsha First
Marsa Dubai	Al Wahda	Al Sabkha	Al Rigga	Downtown Dubai

The communities below have the highest population density in Dubai and again will be given priority with their clusters:

Ayal Nasir	Al Muraqqabat	Al Karama	Al Muteena	Hor Al Anz East
Naif	Al Ras	Hamriya	Al Satwa	Hor Al Anz
Al Murar	Al Buteen	Abu Hail	Al Hudaiba	Marsa Dubai

.After conducting the K-means clustering, we get 13 clusters, 5 of which are meaningful while the rest are single community clusters (such as Warqa 5 which is effectively a Zoo Safari) and can be treated as outliers. The clustered communities can be viewed in the diagram below.



Discussion

After carefully inspecting the clusters (details in the Appendix); we could identify a unique one which predominantly consists of Hotels and restaurants (cluster 6). Given that tourists and business visitors are a main customer segment that we want to attract, our community selection will fall in this cluster.

If we inspect the tables from the previous section; we get an intersection of four communities (densely populated and have a high venue count): **Naif**, **Al Karama**, **Al Muraqqabat** and **Marsa Dubai**. Out of this set, only **Marsa Dubai**, **Naif** and **Al Muraqqabat** fall in the cluster 6. After doing some external research, we decided to rule out **Naif** and **Al Muraqqabat** because these are non-touristic areas and are located on the ghetto side of Dubai. Most of the hotels in these two communities are old and non-rated and hence have relatively low occupancy rate.

This multi-tiered selection process has helped us arrive at **Marsa Dubai** which is popularly called **Dubai Marina**. It satisfies all the criteria that we initially set; here is a quick recap:

Criterion	Rating	Commentary
Large population density		Dubai Marina is among the top 15 most densely populated areas in Dubai
Large number of venues to indicate strong business sentiment		Dubai Marina is among the top 15 communities with the highest venue count which can be a proxy of business sentiment
Demographics: Young, single or small family Predominantly European, North American and Western mindset		We did some offline research and there are numerous articles which point to the fact that marina has favorable demographics [2][3].
Large influx of business visitors and leisure tourists		Contained in cluster 6 and hotels are the most popular venue in Dubai Marina
Surrounded by other similar communities to facilitate delivery / pick up		Dubai Marina is surrounded by similar communities such as JLT and Palm Jumeirah which are highly touristic and have a decent population density
Consists of popular destination such as a Zoo Safari, Beach, entertainment park..etc		Dubai Marina is on the coast and has multiple hot beach destinations.

Conclusion

We have developed a multi-tiered community selection framework which incorporates machine learning and external research to arrive at an optimal restaurant location using several elements as proxy for potential demand or revenue. The data needed for this framework are readily available public data which can be applied to most major cities globally. We can further refine this framework by incorporating also the cost side of the equation in the form of rent, special permits, supply chain and others.

The framework developed can be part of a compelling business case to be presented to prospect investors.

Appendix

Cluster 1 - dominated by restaurants, some coffee shops and shopping

Neighborhood	Area	Population	Density	Lat	Lon	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Al Karama	1.509 km ²	45674	30267	25.244403	55.304755	Indian Restaurant	Café	Restaurant	Fast Food Restaurant	Asian Restaurant
Abu Hail	1.27 km ²	21414	16861.4	25.285942	55.329444	Park	Coffee Shop	Bakery	Restaurant	Dessert Shop
Al Mankhool	1 km ²	16013	16013	25.250827	55.295620	Indian Restaurant	Hotel	Fast Food Restaurant	Chinese Restaurant	Coffee Shop
Hor Al Anz	4.18 km ²	40342	9651	25.277042	55.337300	Indian Restaurant	Fast Food Restaurant	Market	Convenience Store	Bakery
Jumeira First	6.9 km ²	26499	3840	25.046345	55.189545	Fast Food Restaurant	Pizza Place	Ice Cream Shop	Shoe Store	Market
Umm Hurair First	2.3 km ²	4428	1926	25.254192	55.310743	Indian Restaurant	Coffee Shop	Café	Hotel	Middle Eastern Restaurant
Al Barsha South Second	5	7100	1411	25.069928	55.226624	Convenience Store	Coffee Shop	Bakery	Fast Food Restaurant	Department Store
Al Quoz Second	27.1 km ²	16719	617	25.124703	55.251207	Indian Restaurant	Hotel	Golf Course	Light Rail Station	Clothing Store
Nad Shiba 4	8.4 km ²	2563	6.16	25.154553	55.245027	Indian Restaurant	Fast Food Restaurant	Shopping Mall	Department Store	Food Court
Nad Shiba 2	8.4 km ²	2563	6.16	25.154553	55.245027	Indian Restaurant	Fast Food Restaurant	Shopping Mall	Department Store	Food Court

Cluster 2 - dominated by Cafe, some restaurants and misc venues such as jewelry stores and airport terminals

Neighborhood	Area	Population	Density	Lat	Lon	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Hamriya	0.72 km ²	15104	20890	25.287657	55.323997	Coffee Shop	Park	Market	Seafood Restaurant	Fishing Store
Hor Al Anz East	4.18 km ²	40342	9651	25.282082	55.347623	Café	Middle Eastern Restaurant	Fast Food Restaurant	Gym	Shopping Mall
Al Wahda	1.41 km ²	9856	6990	24.469070	54.373631	Café	Coffee Shop	Hotel	Italian Restaurant	Dessert Shop
Muhaisnah First	13 km ²	90870	6990	25.244086	55.412118	Movie Theater	Park	Hostel	Bowling Alley	Baby Store
Al Wasl	4.76 km ²	22153	4654	25.198898	55.257049	Café	Middle Eastern Restaurant	Ice Cream Shop	Burger Joint	Coffee Shop
Umm Al Sheif	1.87 km ²	6263	3350	25.132752	55.205857	Coffee Shop	Café	Jewelry Store	Restaurant	Grocery Store
Al Twar First	6.8 km ²	18457	2714	25.270966	55.362787	Airport Terminal	Coffee Shop	Airport Lounge	Indian Restaurant	Fast Food Restaurant
Al Twar Second	6.8 km ²	18457	2714	25.261550	55.381747	Coffee Shop	Café	Indian Restaurant	Mobile Phone Shop	Boutique
Al Twar Third	6.8 km ²	18457	2714	25.251084	55.396266	Fast Food Restaurant	Tunnel	Basketball Court	Indian Restaurant	Cafeteria
Umm Suqeim First	7.2 km ²	16459	2286	25.162427	55.209815	Coffee Shop	Café	Burger Joint	Bakery	Middle Eastern Restaurant
Umm Suqeim Third	7.2 km ²	16459	2286	25.138459	55.196458	Beach	Café	Restaurant	Fast Food Restaurant	Resort
Umm Suqeim Second	7.2 km ²	16459	2286	25.152020	55.206098	Café	Burger Joint	Coffee Shop	Snack Place	Middle Eastern Restaurant
Umm Hurair Second	2.3 km ²	4428	1926	25.238470	55.323143	Café	Coffee Shop	Sandwich Place	Gym	Lounge
Al Safa Second	4.5 km ²	6291	1398	25.159548	55.224873	Coffee Shop	Middle Eastern Restaurant	Furniture / Home Store	Cafeteria	Café
Al Safa First	4.5 km ²	6291	1398	25.176773	55.239138	Café	Middle Eastern Restaurant	Coffee Shop	Furniture / Home Store	Fast Food Restaurant
Umm Ramool	3.61 km ²	4629	1282	25.229996	55.365626	Auto Garage	Fast Food Restaurant	Breakfast Spot	Sporting Goods Shop	Pizza Place
Al Manara	2.1 km ²	2147	1022	25.143712	55.213574	Café	Coffee Shop	Health & Beauty Service	Supermarket	Salon / Barbershop
Mirdif	9.2 km ²	8115	882	25.221335	55.423499	Coffee Shop	Café	Grocery Store	Middle Eastern Restaurant	Gym
Quoz Industrial 3	27.1 km ²	16719	617	25.142060	55.231168	Art Gallery	Coffee Shop	Gym / Fitness Center	Café	Auto Garage
Quoz Industrial 1	27.1 km ²	16719	617	25.142060	55.231168	Art Gallery	Coffee Shop	Gym / Fitness Center	Café	Auto Garage
Za'abeel First	10 km ²	5283	528.3	25.132622	55.385739	Convenience Store	Italian Restaurant	Fast Food Restaurant	Tea Room	Donut Shop
Za'abeel Second	10 km ²	5283	528.3	25.133881	55.385947	Convenience Store	Pizza Place	Fast Food Restaurant	Tea Room	Gym / Fitness Center
Al Sufouh 1	7.6 km ²	3337	439	25.113261	55.169486	Beach	Coffee Shop	Clothing Store	Pharmacy	Indian Restaurant
Al Mizhar Second	11.2 km ²	4326	386.25	25.243065	55.460523	Café	Grocery Store	Ice Cream Shop	Trail	Pizza Place
Al Nahda Second	3.18 km ²	600	188.67	25.290592	55.376731	Fast Food Restaurant	Coffee Shop	Convenience Store	Indian Restaurant	Italian Restaurant
Al Shindagha	0.26 km ²	16	70	25.275464	55.294855	Pier	Boat or Ferry	Cruise	Market	Jewelry Store

Cluster 4 - all communities here adhere to one structure more or less

Neighborhood	Area	Population	Density	Lat	Lon	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Muhaisanah Fourth	13 km²	90870	6990	25.280555	55.410502	Business Service	Gym	Fast Food Restaurant	Hotel	Pet Service
Muhaisanah Second	13 km²	90870	6990	25.280555	55.410502	Business Service	Gym	Fast Food Restaurant	Hotel	Pet Service
Muhaisanah Third	13 km²	90870	6990	25.280555	55.410502	Business Service	Gym	Fast Food Restaurant	Hotel	Pet Service
Nasr	1.5 km²	2469	1646	25.281865	55.411010	Business Service	Gym	Fast Food Restaurant	Restaurant	Cheese Shop

Cluster 6 - dominated by hotels, restaurants and cafes; main distinctive element is the prevalence of hotels

Neighborhood	Area	Population	Density	Lat	Lon	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Ayal Nasir	0.17 km²	13077	76923	25.275307	55.303120	Hotel	Middle Eastern Restaurant	Fast Food Restaurant	Shopping Mall	Café
Naif	0.561 km²	28571	50928	25.272775	55.311351	Hotel	Middle Eastern Restaurant	Coffee Shop	Café	Fast Food Restaurant
Al Murar	0.41 km²	19831	48368	25.276677	55.309763	Hotel	Middle Eastern Restaurant	Asian Restaurant	Fast Food Restaurant	Café
Al Muraqqabat	0.78 km²	32395	41532	25.266308	55.324222	Hotel	Middle Eastern Restaurant	Coffee Shop	Restaurant	Café
Al Ras	0.20 km²	6812	34060	25.268885	55.296075	Hotel	Indian Restaurant	Middle Eastern Restaurant	Shopping Mall	Café
Al Buteen	0.07 km²	2364	33771	25.263057	55.320584	Hotel	Middle Eastern Restaurant	Coffee Shop	Restaurant	Café
Al Sabkha	0.13 km²	2627	20207	25.267181	55.303487	Hotel	Middle Eastern Restaurant	Indian Restaurant	Café	Asian Restaurant
Al Muteena	1.12 km²	18094	16155	25.273892	55.322630	Hotel	Indian Restaurant	Coffee Shop	Fast Food Restaurant	Asian Restaurant
Al Satwa	2.87 km²	30147	10504	25.221161	55.273836	Coffee Shop	Café	Middle Eastern Restaurant	Hotel	Restaurant
Marsa Dubai	5	45500	9300	25.087754	55.146172	Hotel	Coffee Shop	Restaurant	Café	Beach
Al Hudaiba	0.84 km²	7699	9165	25.240050	55.277459	Hotel	Middle Eastern Restaurant	Café	Asian Restaurant	Coffee Shop
Al Barsha First	4	35500	8875	25.111504	55.197187	Hotel	Clothing Store	Coffee Shop	Café	Indian Restaurant
Al Jafiliya	1.63 km²	11619	7128	25.233360	55.292050	Hotel	Coffee Shop	Restaurant	Lounge	Spa
Al Baraha	1.104 km²	7823	7086	25.282576	55.320139	Hotel	Fast Food Restaurant	Park	American Restaurant	Café
Downtown Dubai	2.68 km²	18,698[2]	6975	25.197092	55.279036	Boutique	Hotel	Coffee Shop	Café	Lebanese Restaurant
Al Rigga	0.89 km²	5684	6386	25.267316	55.312268	Hotel	Middle Eastern Restaurant	Coffee Shop	Fast Food Restaurant	Restaurant
Business Bay	46 km²	191000	4200	25.188828	55.265022	Hotel	Lounge	Middle Eastern Restaurant	Indian Restaurant	Restaurant
Jumeira Second	6.9 km2	26499	3840	25.233175	55.277371	Café	Asian Restaurant	Middle Eastern Restaurant	Hotel	Coffee Shop
Jumeira Third	6.9 km2	26499	3840	25.233175	55.277371	Café	Asian Restaurant	Middle Eastern Restaurant	Hotel	Coffee Shop
Rigga Al Buteen	0.684 km²	1394	2038	25.265992	55.317428	Hotel	Middle Eastern Restaurant	Restaurant	Café	Coffee Shop
Palm Jumeirah	4	10500	1842	25.118501	55.133063	Beach	Gym	Hotel	Mediterranean Restaurant	Track
Trade Centre 2	1.91 km²	3353	1755.5	25.228012	55.289702	Hotel	Lounge	Restaurant	Nightclub	Theme Park
Al Mina	3.10 km²	4183	1350	17.066641	-96.719500	Hotel	Mexican Restaurant	Café	Italian Restaurant	Taco Place
Al Garhoud	4 km²	4466	1116.5	25.239831	55.355668	Hotel	Airport Terminal	Fast Food Restaurant	Coffee Shop	Mobile Phone Shop
Nad Shamma	1.36 km²	1208	888.34	25.217577	55.383785	Fast Food Restaurant	Middle Eastern Restaurant	Gym / Fitness Center	Furniture / Home Store	Bakery
Al Sufouh 2	7.6 km²	3337	439	25.124439	55.186900	Hotel	Resort	Beach	Harbor / Marina	Cocktail Bar
Al Jaddaf	7.3 km²	2990	409.5	25.065964	55.171340	Golf Course	Fast Food Restaurant	Steakhouse	Hotel	Spa
Port Saeed	1.84 km²	448	243.4	25.246647	55.333871	Coffee Shop	Middle Eastern Restaurant	Hotel	Clothing Store	Indian Restaurant
Al Nahda First	3.18 km²	600	188.67	25.288351	55.363138	Middle Eastern Restaurant	Café	Hotel	Seafood Restaurant	Bakery

Cluster 13 - mostly cafeterias and casual food restaurants

Neighborhood	Area	Population	Density	Lat	Lon	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Al Rashidiya	4.41 km²	22326	5062	25.224953	55.390507	Cafeteria	Coffee Shop	Middle Eastern Restaurant	Bakery	Light Rail Station
Al Mamzar	3.35 km²	2260	674.6	25.303247	55.343453	Beach	Cafeteria	Café	Grocery Store	Juice Bar
Al Quoz First	27.1 km²	16719	617	25.169145	55.252975	Cafeteria	Ice Cream Shop	Coffee Shop	BBQ Joint	Entertainment Service
Al Quoz Third	27.1 km²	16719	617	25.169070	55.253398	Cafeteria	Ice Cream Shop	Historic Site	Bowling Alley	Nightclub
Nad Hamar	8.4 km²	2563	6.16	25.201893	55.376374	Cafeteria	Sandwich Place	Clothing Store	Fast Food Restaurant	Basketball Court
Al Warqa'a First	21.6 km²	5	0.23	25.190247	55.402941	Cafeteria	Bookstore	Tea Room	Burger Joint	Market
Al Warqa'a Second	21.6 km²	5	0.23	25.200975	55.408048	Burger Joint	Middle Eastern Restaurant	Fast Food Restaurant	Cafeteria	Café
Al Warqa'a Third	21.6 km²	5	0.23	25.191657	55.419019	Shopping Mall	Ice Cream Shop	Tea Room	Bookstore	BBQ Joint

References

1. <https://www.thenational.ae/arts-culture/hidden-food-gems-fans-will-love-this-game-of-thrones-inspired-sandwich-shop-in-khalidiya-1.47219>
2. <https://www.thenational.ae/business/dubai-marina-s-take-on-european-style-living-with-the-wow-effect-in-pictures-1.197815>
3. <https://londonlovesbusiness.com/which-are-the-best-areas-for-young-people-to-live-in-the-uae/>