Exploring Restaurants in Neighborhoods of Dubai

Applied Data Science Capstone Project

By,

Maryam Muhammad Zubair

Table of Contents

Introduction 4
Background 4
Research Questions 4
Data – Requirements, Sources and Preparation 5
Methodology 7
Descriptive Analysis 7
K-means Clustering 7
Results 9
Discussion 17
Conclusion 17
References 18

List of Figures

Figure 1 Dubai map with marked neighborhoods 5
Figure 2 Snapshot of the results from FOURSQUARE when food related venues are searched for Aweer, Dubai. 6
Figure 3 Most frequent restaurant categories in neighborhoods of Dubai 9
Figure 4 Frequency of restaurant categories in neighborhoods of Dubai 10
Figure 5 Most common venues in neighborhoods of Dubai 11
Figure 6 K-means clustered neighborhoods 12
Figure 7 Cluster 1 most frequent restaurant categories 13
Figure 8 Cluster 2 most frequent restaurant categories 14
Figure 9 Cluster3 most frequent restaurant categories 14
Figure 10 Cluster 4 most frequent restaurant categories 15
Figure 11 Cluster 5 most frequent restaurant categories 15
Figure 12 Most populated neighborhoods of cluster 3 16

Introduction

Background

The restaurant industry has high demand and is a sought-after market in Dubai, the most famous and populous [1] city of United Arab Emirates. Starting a restaurant is one of the many profitable business ideas in Dubai [2] but investors are faced with a location dilemma in addition to the other expected questions and prerequisites. When a city is already home to a variety of restaurants, investors want to avoid competition and yet attract a good number of the population. The multicultural community of Dubai provides investors with a range of options to explore before deciding what specific project and category of restaurant they want to invest in.

Research Questions

In this project, we will be exploring the different type and density of restaurants in neighborhoods of Dubai. We will try to answer the following questions:

1) What are the most frequent types of restaurant categories in neighborhoods of Dubai?
2) Can we segment neighborhoods based on their surrounding food venues?
3) Can we identify an underserved region for a breakfast spot in these neighborhoods?

The target audience of this project:

1. Investors interested in restaurant business – the project will give them a summary about the different categories and density of restaurants in neighborhoods of Dubai and potentially find a gap which they can target. It will also provide insights to the population in these neighborhoods to understand if a good customer base can be built.
2. Students of data science, who may want to study sample projects to understand how foursquare can be used to understand and help in solving venue related problems.

Data – Requirements, Sources and Preparation

To answer the above question, the data requirements are as follows:

1. List of the neighborhoods in Dubai
    a. Latitude and longitude
    b. Population

We download the data provided by Government of Dubai, Population and Statistics [3], to obtain a list of neighborhoods along with their population. This data is available for download on their website and provides Number of Estimated Population by Sector and Community in the year 2019. For the latitude and longitude for each of the communities/neighborhoods, we use Geolocator python package [4].

In the data preparation stage, we clean the data by excluding neighborhoods with a population less than 500. We also merge different divisions of the same neighborhood into one and sum their populations. For instance, Khawaneej One and Khawaneej Two are considered as Khawaneej. This is done for simplicity and after considering that the areas are not very large to require segregation. For this project, we only include the neighborhoods for which the coordinates are readily available through the geolocator. After all the data cleaning and wrangling, we have a total of 66 neighborhoods with their respective population and co-ordinates. Figure 1 shows a map of Dubai with all the 66 neighborhoods marked.

Figure 1 Dubai map with marked neighborhoods

2. Data on the different categories of restaurant around a neighborhood/community.

We find the data on nearby restaurants of each neighborhood using the Foursquare API. Foursquare is a data platform which allows one to access and search for a wide range of information on venues around a given location. There are various venue categories including shopping venues, food venues, recreation, art, entertainment etc. In the venue search, we use the categoryId of 'Food' to include all venues that fall under this main category. We use a radius on 5 km and limit of 100 restaurants. Figure 2 shows a snapshot of the different categories of restaurants in one of the neighborhoods, i.e. Aweer, Dubai. We repeat the process for all the neighborhoods and obtain a complete data set of all the food venues. We then use this data to study the density of different type of restaurants in each neighborhood and identify the most and least common categories across. This will help us identify gaps. We will also cluster the neighborhood based on food venues around to understand if certain neighborhoods have similar distribution of restaurants. For instance, if there are more breakfast spots in some neighborhoods than others.

	name	categories	lat	lng
0	Al Raslan Cafe	Café	25.178707	55.545097
1	Al Raslaan Star Cafe الرسلان ستار كافيه	Café	25.178739	55.545027
2	Sports Restaurant مطعم الرياضة	Burger Joint	25.175851	55.542336
3	Zaffran AlAweer Cafe (مقهى زعفران العوير)	Café	25.160281	55.535183
4	Qalat al shay cafeteria (كافيتريا قلعة الشاي)	Cafeteria	25.176925	55.542434
5	Jernas Cafe	Café	25.178689	55.544648
6	كَفْتِيرْيَا الرَّسْلاَن (الْعُويرا)	Cafeteria	25.178810	55.545098

Figure 2 Snapshot of the results from FOURSQUARE when food related venues are searched for Aweer, Dubai.

Methodology

Descriptive Analysis

First, we prepare a list of unique categories of restaurants across the 5276 records of venues and neighborhoods and their respective frequencies. There were 97 unique categories of restaurants across the dataset. The results are shown in a bar plot and the 20 most frequent restaurant types in Dubai are listed in a table (Results section).

Second, the data from foursquare was wrangled to obtain the total number of restaurants per neighborhood. The total number of restaurants per neighborhood ranged between 11 and 100. Almost 70% neighborhoods returned the maximum number of restaurants, i.e. 100 while 17% neighborhoods returned less than 50 restaurants. The top 10 most frequent restaurants for each neighborhood were also recorded to understand the density of restaurants in each neighborhood.

K-means Clustering

To understand how neighborhoods cluster based on surrounding food venues, we performed a k means clustering analysis on the data. K-Means Clustering is a type of unsupervised learning which can be used to find groups within unlabeled data. K-mean divides the data into clusters where clusters are “a collection of data points aggregated together because of certain similarities” [5].This analysis was performed to understand how similar neighborhoods are in terms of food preferences. For instance, we expect that neighborhoods of the same cluster would have similar food venues and hence similar food behaviors and preferences. It will also help us understand and segregate our data. In short, k means clustering would allow 1) Neighborhood segmentation 2) understand the distribution of restaurants 3) identify underserved neighborhoods

The data was standardized before the clustering. The venue data was prepared by calculating mean frequencies for standardization. Then a k means clustering algorithm was run with a value of 5 clusters. Neighborhoods that returned a Nan value from the clustering were removed. The results were displayed on Dubai’s map with each neighborhood color coded according to the cluster it belongs to (Results section). The total number of restaurants for each cluster was also calculated.

Target cluster for breakfast spot

Clusters were analyzed for differences by calculating the most frequent categories in each cluster. The results were then used to identify a target cluster for opening a breakfast spot in Dubai. The population data was also used as a reference to extract the most populated neighborhoods of the target cluster.

Results

The top 20 restaurant categories 3 with their respective frequency are shown in Figure 3 and the frequency of the 97 different restaurant categories is shown in a bar graph in Figure 4. The neighborhoods are highly dominated by Cafes, Middle eastern restaurants, Indian and fast food restaurants. We can also observe that there are multiple cuisines and varieties on the top 20 list such as Italian, Japanese, Chinese, French, breakfast spots and steakhouses. A sliced snapshot of the most common venues in neighborhoods of Dubai are also displayed in Figure 5.

	Frequency
Café	611
Middle Eastern Restaurant	493
Restaurant	390
Indian Restaurant	303
Fast Food Restaurant	278
Bakery	198
Asian Restaurant	189
Italian Restaurant	174
Burger Joint	170
Seafood Restaurant	161
Chinese Restaurant	123
Sandwich Place	120
Japanese Restaurant	117
French Restaurant	117
Pizza Place	109
Breakfast Spot	105
American Restaurant	105
Steakhouse	79
Thai Restaurant	75
Fried Chicken Joint	74

Figure 3 Most frequent restaurant categories in neighborhoods of Dubai

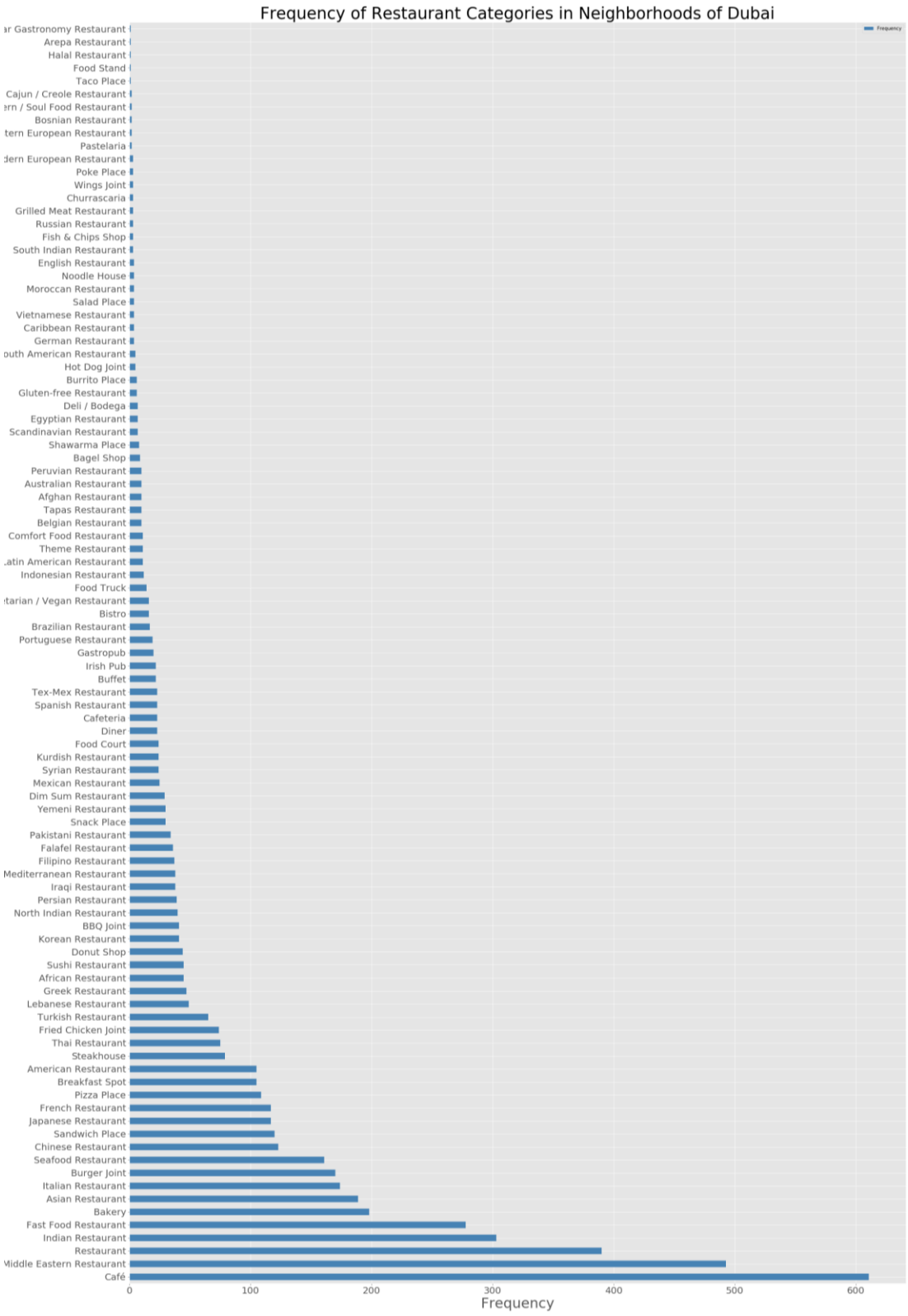

Figure 4 Frequency of restaurant categories in neighborhoods of Dubai

	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	AWEER , DUBAI	Café	Cafeteria	Indian Restaurant	Seafood Restaurant	Burger Joint	Fast Food Restaurant	Yemeni Restaurant	Filipino Restaurant	Donut Shop	Eastern European Restaurant
1	AYAS, DUBAI	Café	Restaurant	Burger Joint	French Restaurant	Middle Eastern Restaurant	Italian Restaurant	Breakfast Spot	Steakhouse	Japanese Restaurant	Turkish Restaurant
2	BARAHA, DUBAI	Middle Eastern Restaurant	Indian Restaurant	Café	Restaurant	Asian Restaurant	Fast Food Restaurant	Seafood Restaurant	Japanese Restaurant	Fried Chicken Joint	North Indian Restaurant
3	BARSHA SOUTH , DUBAI	Café	Restaurant	Middle Eastern Restaurant	Bakery	Seafood Restaurant	Burger Joint	Lebanese Restaurant	Breakfast Spot	Indian Restaurant	Turkish Restaurant
4	BARSHAA , DUBAI	Café	Restaurant	Seafood Restaurant	Middle Eastern Restaurant	Italian Restaurant	Bakery	African Restaurant	Thai Restaurant	French Restaurant	Asian Restaurant

Figure 5 Most common venues in neighborhoods of Dubai

K-means clustering

The k means clustering results are shown in Figure 5 with each neighborhood's color reflecting the cluster it belongs to. The total restaurants for each cluster is shown in the table below. Most neighborhoods clustered together according to their geographical locations with some exceptions, i.e. most neighborhoods geographically close to each other are clustered into one.

Cluster	Color	Restaurants
Cluster 1	Red	22
Cluster 2	Purple	2164
Cluster 3	Blue	2504
Cluster 4	Green	267
Cluster 5	Orange	319

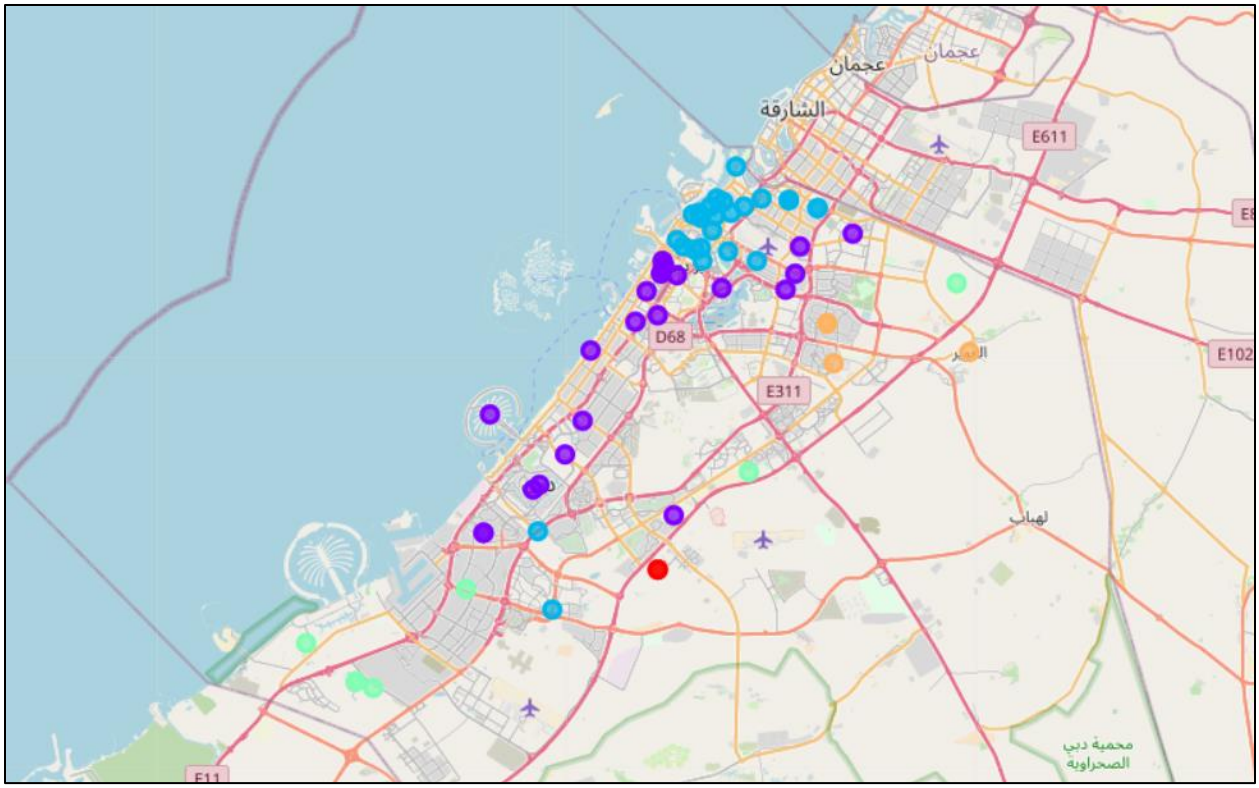

Figure 6 K-means clustered neighborhoods

Figures 7 to 11 show the top ten most frequent restaurants in each cluster with their respective frequencies indicating the differences among them. Cluster 1 has very few restaurants (22), given only two neighborhoods fall into that category, with mostly cafes and Lebanese restaurants. Cluster 2 neighborhoods are spread out across the city highly dominated by Cafes followed by restaurants, middle eastern, burger joints, Italian, bakeries and Indian restaurants. Breakfast spots and sea food restaurants also form a good portion of this cluster. While Cluster 3 includes neighborhoods closely located, it is dominated by middle eastern restaurants, Indian restaurants, and cafes. Fast food, Bakeries, Japanese and sea food also are found in good numbers. Cluster 4 is situated far from the city, includes industrial areas and is dominated by fast food, cafes and restaurants while cluster 5 mostly is house to cafes. A summarized list with top 10 restaurant categories for each cluster is found in table 1 below.

Table 1 10 Most Common Venues in Clusters

Cluster	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
1	Restaurant	Café	Lebanese Restaurant	French Restaurant	Sushi Restaurant	South American Restaurant	Diner	Fast food Restaurant		
2	Café	Restaurant	Middle Eastern Restaurant	Burger Joint	Italian Restaurant	Bakery	Indian Restaurant	French Restaurant	Seafood Restaurant	Breakfast spot
3	Middle Eastern Restaurant	Indian Restaurant	Café	Restaurant	Fast Food Restaurant	Asian Restaurant	Bakery	Seafood Restaurant	Japanese Restaurant	Sandwich Place
4	Fast Food Restaurant	Café	Restaurant	Italian Restaurant	Pizza Place	Indian Restaurant	Sandwich Place	Burger Joint	Middle Eastern Restaurant	Chinese Restaurant
5	Café	Middle Eastern Restaurant	Chinese Restaurant	Fast Food Restaurant	Burger Joint	Sandwich Place	Indian Restaurant	Pizza Place	Restaurant	Bakery

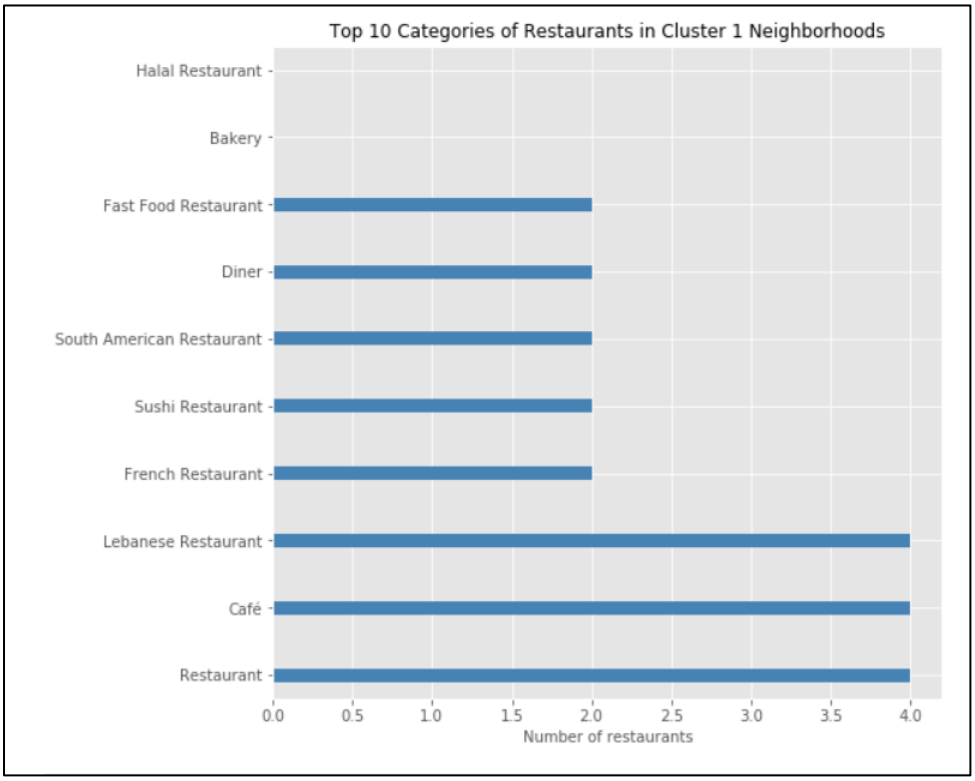

Figure 7 Cluster 1 most frequent restaurant categories

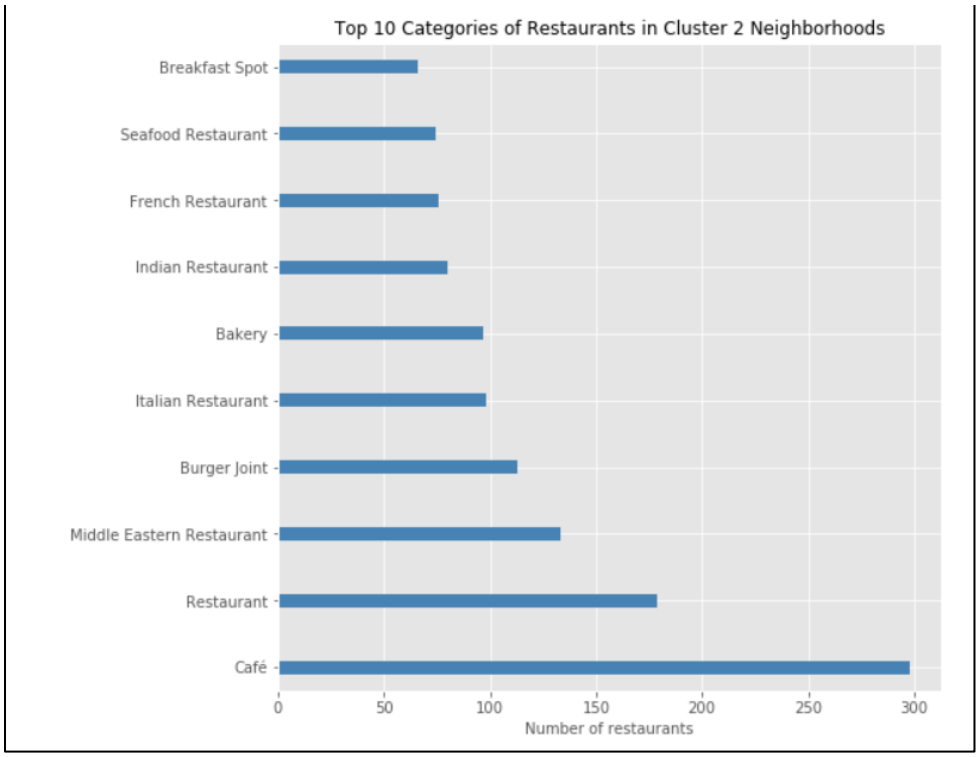

Figure 8 Cluster 2 most frequent restaurant categories

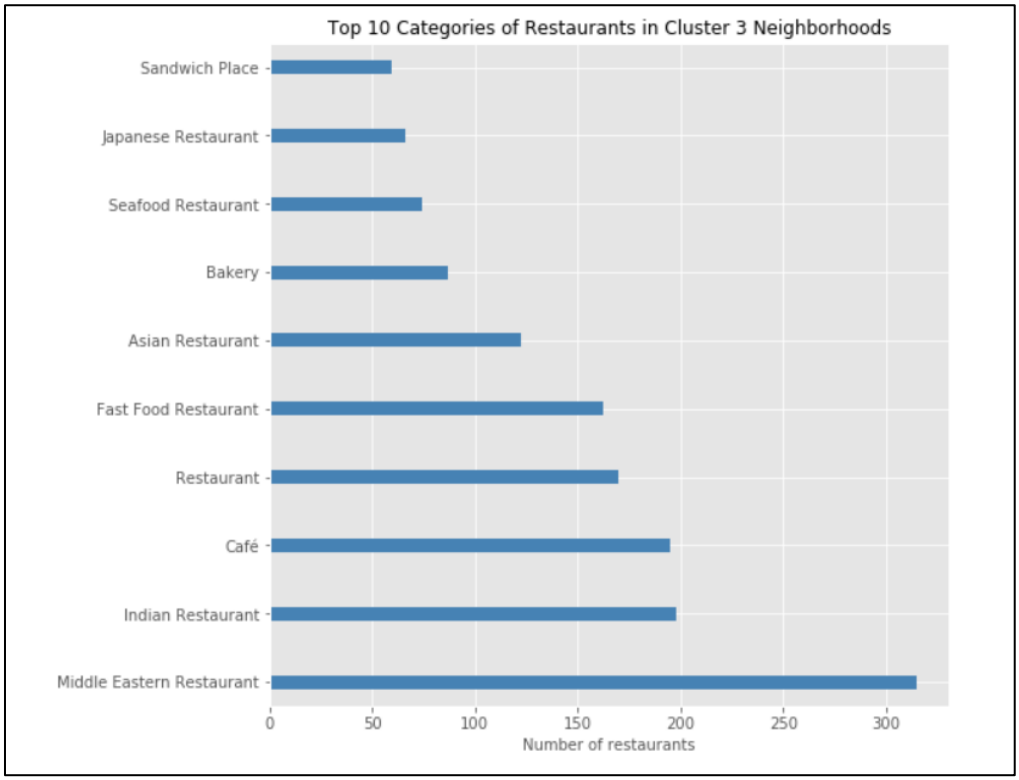

Figure 9 Cluster3 most frequent restaurant categories

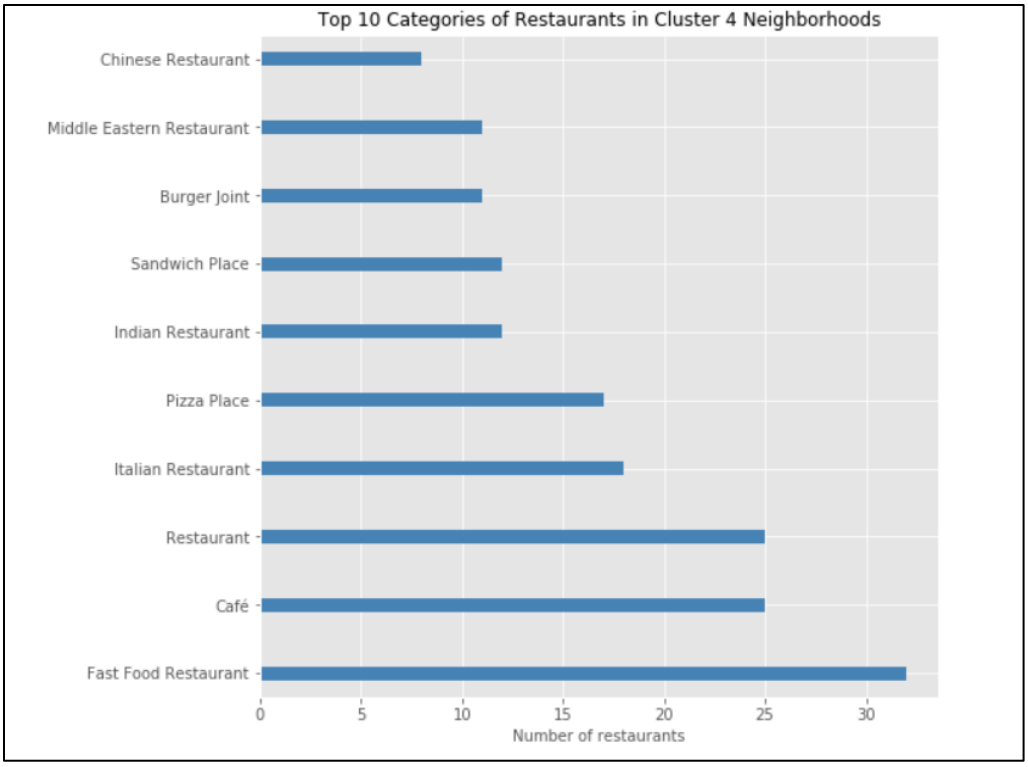

Figure 10 Cluster 4 most frequent restaurant categories

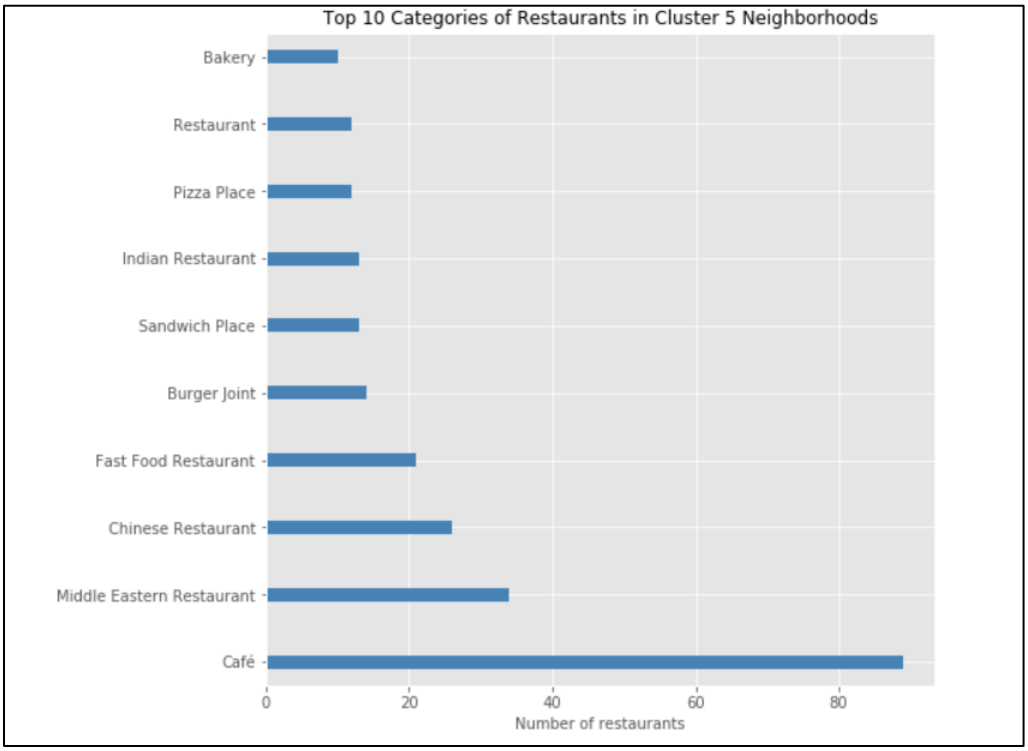

Figure 11 Cluster 5 most frequent restaurant categories

Breakfast spot

A target cluster is identified to open a breakfast spot in Dubai. We notice that only cluster 2 has breakfast spots in the top 10 categories. To attract a good percentage of the population, it will be a good idea to open a breakfast spot in empty niche like cluster 3. The top 5 most populated neighborhoods of cluster 3 are shown below in Figure 11. As we can see there are very few breakfast spots in these neighborhoods of central Dubai although the population is high. So as a starting point, we suggest considering these neighborhoods, (Muhaisanah, Nahda, Hor Al Anz, Karama and Qusais) as they are underserved for a breakfast spot and look further into other factors.

	Neighborhood	Breakfast Spot	Total Restaurants	Total population
20	MUHAISANAH , DUBAI	1	100	196316
9	NAHDA , DUBAI	1	100	93385
18	HOR AL ANZ, DUBAI	2	100	83187
4	KARAMA, DUBAI	2	100	75066
10	QUSAIS , DUBAI	1	100	65292

Figure 12 Most populated neighborhoods of cluster 3

Discussion

In this project, we were to find and summarize the different categories of restaurants in neighborhoods of Dubai using foursquare. The different ethnic cuisines are a reflection of the multicultural community the place has. It also represents that the population in interested in a variety of tastes. It is not a surprise that cafés are a hotspot in a place like Dubai. We were also able to cluster different neighborhoods based on surrounding food venues. Interestingly, neighborhoods that are geographically close are also homogenous in their surrounding food venues and tend to have the same food environment. As expected, we found that the neighborhoods farther away from the city do not have as many restaurants as central Dubai. We also able to identify an underserved region for a breakfast spot in Dubai.

Notwithstanding the insights, this study has several limitations. First, we were not able to find location data for some neighborhoods. This data may have provided deeper insights. Population income would also be an interesting addition, the community income may be related to the type of food environment built around them. For instance, high end and expensive restaurants may be situated in high income neighborhoods and so on.

Conclusion

Dubai is house to many different cuisines and has something to suit all tastes. Foursquare is a useful platform to discover venue information. Future studies should leverage platforms like foursquare for a deeper understanding of how venues are related to other factors in the society, for instance relationship of obesity and density of restaurants in a place.

References

[1] "Population of Cities in United Arab Emirates (2020)," [Online]. Available: Population of Cities in United Arab Emirates (2020).

[2] "What are profitable business ideas in Dubai?," September 2019. [Online]. Available: https://quikrgulf.com/blog/market-trends/what-are-profitable-business-ideas-in-dubai/.

[3] "Population and Vital Statistics," 2019. [Online]. Available: https://www.dsc.gov.ae/en-us/Themes/Pages/Population-and-Vital-Statistics.aspx?Theme=42&year=2019#DSC_Tab1.

[4] "GEOCODER PYTHON," [Online]. Available: https://geocoder.readthedocs.io/index.html..

[5] "List of Communities in Dubai," [Online]. Available: https://en.wikipedia.org/wiki/List_of_communities_in_Dubai.