

Moving to Dubai

Ahmed Nady

28th April 2020

1. Introduction

1.1 Background

Movement from your home City to another Country/City is a very big challenge for you, you always need to search about a specific place that matches your requirements for satisfactions, Dubai is one of the cities that has different cultures and nationalities with different interests.

Every day many people go to Dubai for work or tourism and they always want to know where they should stay

1.2 Problem

When we move to a new place, we need to know that everything we need is available in the same neighborhood that we live in , so the question here , which neighborhoods the most interesting venues to us, which place have a good density of population not high neither low density

1.3 Interest

Any mover to Dubai for Work or tourism will be interested in this analysis, even Developers will be interested to know which neighborhoods can be targeted for their new developments to attract the buyers by the features of each neighborhood

2. Data acquisition and cleaning

2.1 Data sources

we need to start with Dubai Neighborhoods list with the area and population per each neighborhood [here](#) on Wikipedia, another important thing is the venues per each neighborhood so we can get that from Foursquare Places API ,so we extracted Longitude and Latitude for each Neighborhood and get the nearby venues that will help us in our Dataset

2.2 Data cleaning

Data Scrapped from Wikipedia that include Neighborhoods,Area,Population and Density and the columns as Name in Arabic and Neighborhood ID which is not needed so:

1- First remove the unneeded columns

2- we found sub-neighborhoods as Neighborhood first,Second ..etc that doesn't have any information or just a duplicate information from the main Neighborhood so removing those areas and remove duplication from the data frame we had used

3- unifying Numbers as float and remove any characters from it as removing "Km2" from Area and removing "/Km2" from Density then convert it to float to be able to analyze them as numbers

4- Removing missing values as we can't predict the Area or Density and depend only on complete values set

Final DataSet is

Table 1

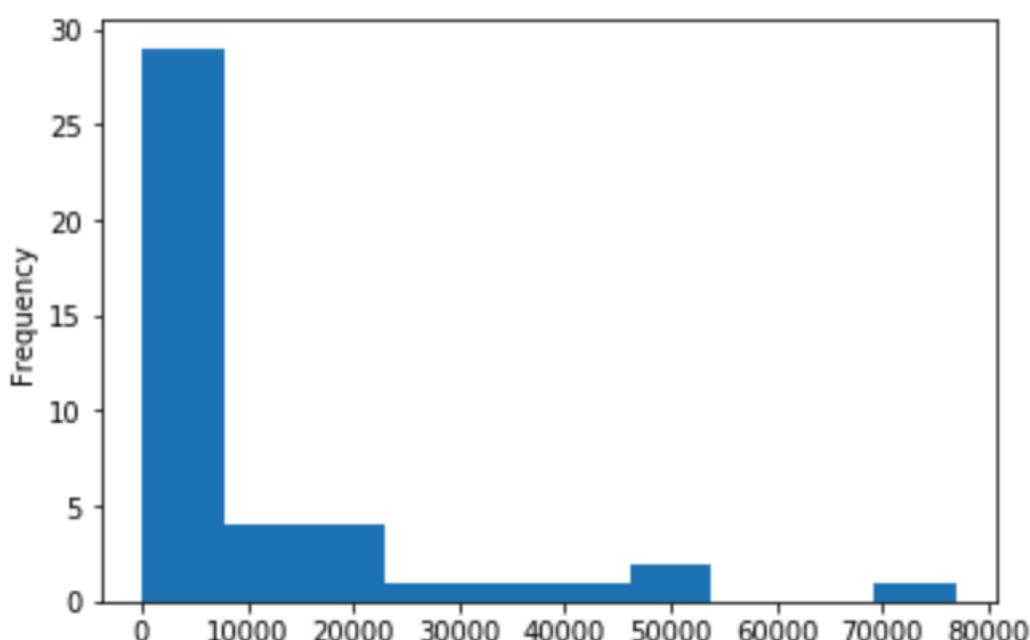
CommName	Area	Population	Density	long	lat
Abu Hail	1.27	1414	16861.4	55.329444	25.285942
Al Baraha	1.104	783	7086	55.320139	25.282576
Al Garhoud	4	4466	1116.5	55.355668	25.239831
Al Hudaiba	0.84	7699	9165	55.277459	25.240050
Al Jaddaf	7.3	990	409.5	55.171300	25.065700
Al Jafiliya	1.63	11619	7128	55.292050	25.233360
Al Karama	1.509	45674	30267	55.304755	25.244403
Al Kifaf	0.8	35	44	55.274809	25.227887

3. Methodology

3.1 Data Analysis:

applying data Analysis techniques using Pandas to get the density distribution and analyze it for all Neighborhoods

We had 43 Neighborhood that has identified by Foursquare API to get their data for this analysis we found that the Density Histogram as below



So we find that most of Neighborhoods has density below 10K per Kmsqr

And

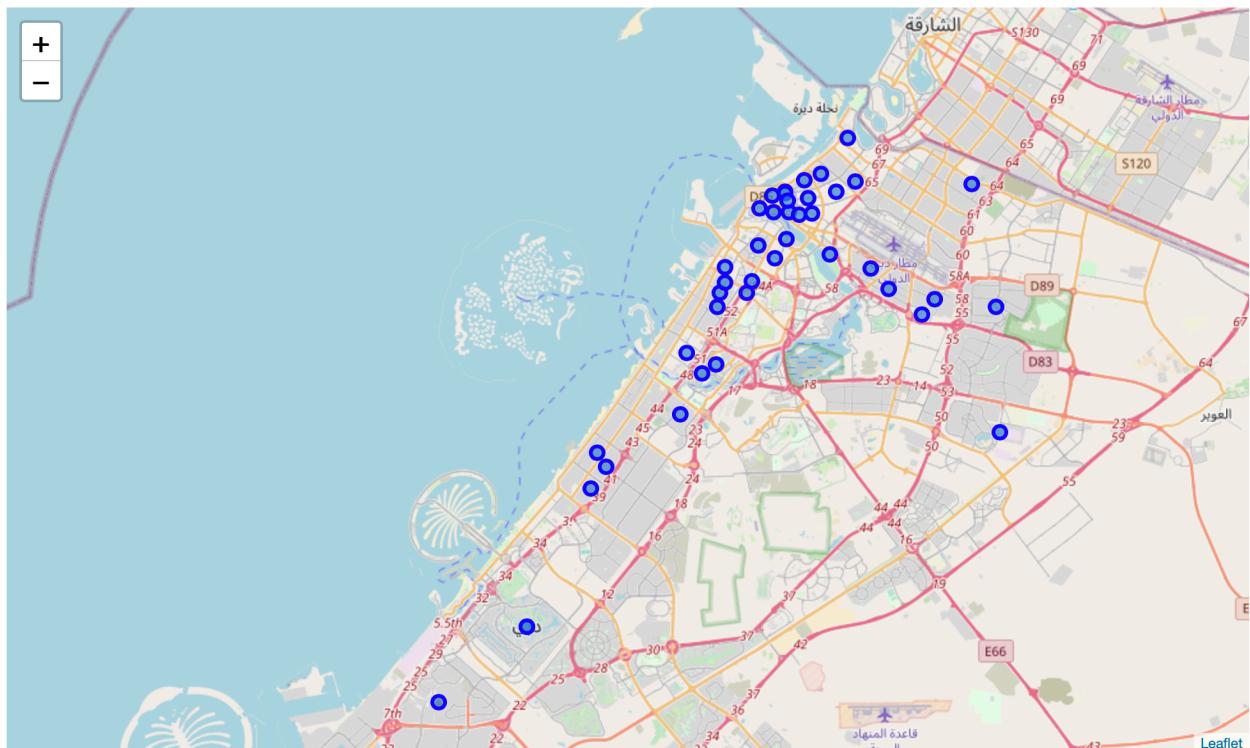
Maximum Neighborhood Density: 76923 per Kmsqr

Minimum Neighborhood Density: 44.0 per Kmsqr

Avg Density: 10790.9 per Kmsqr

And plotting the Neighborhoods over Dubai Maps:

The geographical coordinate of Dubai, AE are 25.0657, 55.1713.



Most Existing Venues

Table 1-1

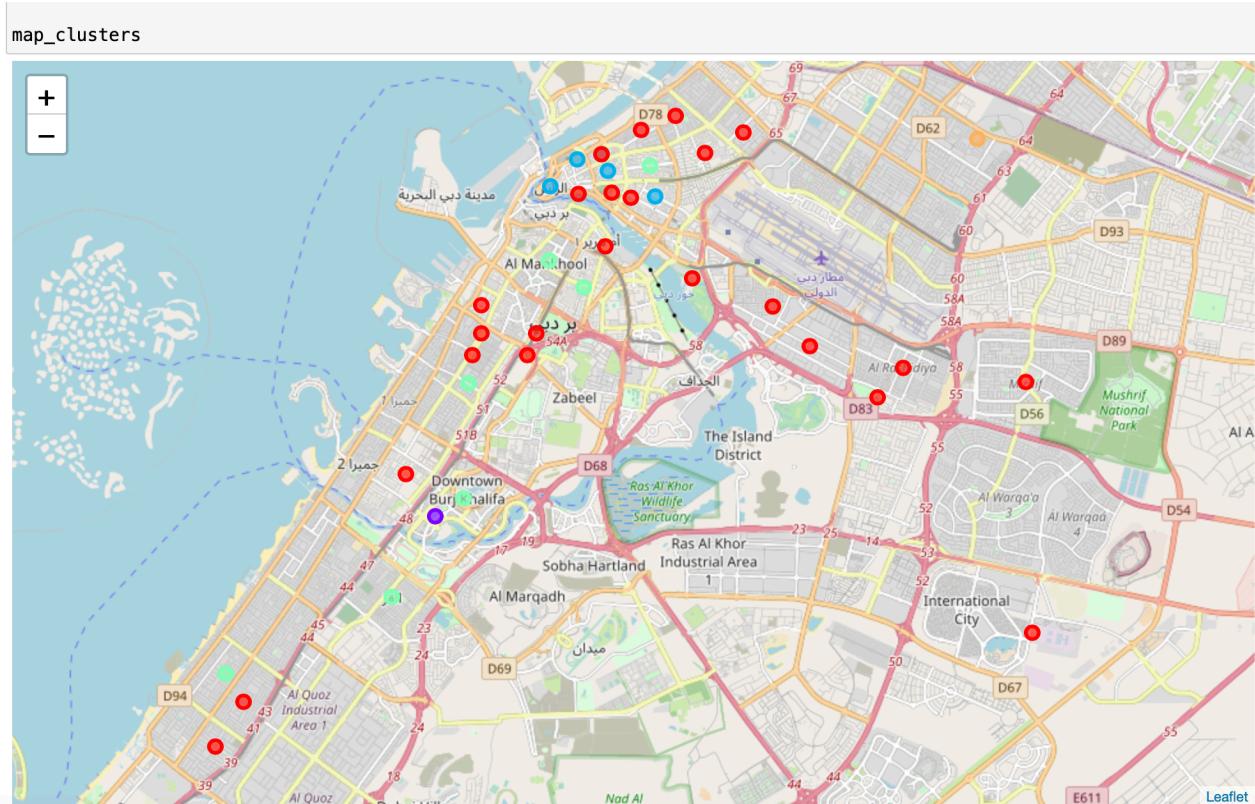
Venue Category	Counts
Hotel	57
Middle Eastern Restaurant	45
Café	41
Asian Restaurant	34
Coffee Shop	30
Indian Restaurant	29
Restaurant	19
Fast Food Restaurant	15
Ice Cream Shop	10
Gym	10

then Applying K-Means Clustering, to segment the Neighborhoods based on the Venues categories and Neighborhood attributes [Area,Population,Density]

We need 5 clusters to segment our Neighborhoods so that we can define each segment extracted

4. Results:

We got 5 Clusters represented on the Map



Cluster 1

it is in Red and representing the most common residential areas , including Parks, Cafes and Restaurants with different Density and Areas

Cluster 2:

it is in Purple and representing Ares with Night life so the theme is hotels, lounges and Nightclubs, and with below avg density

Cluster 3 :

it is in Blue and showing high Density Residential Areas because it is including Restaurants and markets

Cluster 4:

it is in Green and Representing Area for day activities theme , like beach, pool, gym, and restaurants and with different density

Cluster 5:

it is in Orange, and representing Business Area as it has business services, currency exchange centers and an avg density

5 Discussions:

so now based on this Results, we can find that:

- 1- Cluster 1 is good for Families
- 2- Cluster 2 is good for Singles
- 3- Cluster 5 is good for Business centers
- 4-Cluster 3 is not recommended if you like quite places as it will be very crowded
- 3- if you like day activities so cluster 4 is your target

Conclusion:

Finally, we can find that Dubai Neighborhoods can be segmented based on the style and density of population , and this analysis can be used to find a good neighborhood that matches your requirements or to find a good place to establish your new business , but also we need to refine this clusters getting new data sources as price per unit, transportation available per each neighborhood and it will help to define your requirements for the new place matching your requirements