Post Offices Analysis of Abu Dhabi City

Ziad Samhan

October 20, 2019

Table of Contents

1. Introduction

1.1. Background

Abu Dhabi is the largest emirate in the United Arab Emirates, occupying 84 percent of the national landmass territory. It has 200 islands and a long coastline stretching 700km. Its total area is 67,340 sq. km. Abu Dhabi city in the emirate is the federal capital of the UAE.,The emirate of Abu Dhabi lies on the coast of the Arabian Gulf and is bordered by Sultanate of Oman to the east, the Kingdom of Saudi Arabia to the south and the west and the emirate of Dubai to the northeast. The three main regions of the emirate are: The city of Abu Dhabi, Al Ain in the east, Al Dhafrah (earlier known as Al Gharbia) in the west, For this analysis I will concentrate in "The city of Abu Dhabi"[1]

1.2. Problem

Determining the best location for new post office requires analysis of current post offices distribution and population distribution within the "city of Abu Dhabi". However I was not able to gather data for population distribution within The "city of Abu Dhabi" Neighborhood wise, so I will use current offices with 2.5 Km radius as a coverage area for the post office.

1.3. Interest

Having more post offices means more coverage and less load for the currently operating offices, I will examine neighborhoods without post offices to try to determine best locations with minimum offices for the uncovered neighborhoods.

2. Data acquisition and cleaning

2.1. Data sources

For the analysis, I was able to gather data from post offices and neighborhoods using:
- Bayanat.ae, the official data portal of the UAE Government[2]
- Forsquare API, to get post offices of city of Abu Dhabi.
-BeautifulSoup4, for pulling neighborhoods from Wikipedia page containing list of neighborhoods of the city of Abu Dhabi[3]
-Nominatim, to convert neighborhoods' addressess into latitude and longitude.
-I used Google Map, to get the coordinates for some of neighborhoods.

2.2. Data Cleaning

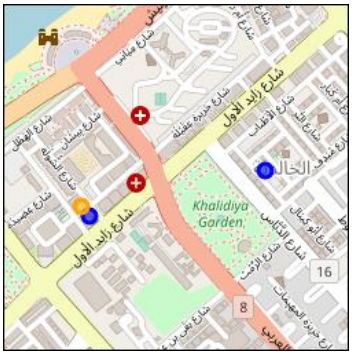
I have obtained 25 post office locations from "bayanat.ae"[4] covering Abu Dhabi emirate, where our interest is the city of Abu Dhabi. And since I was not able to obtain a Second-level Administrative Divisions of Abu Dhabi, I have dropped locations which located more than 50Km from Abu Dhabi City center, and International Airport Post Office was dropped since it is limited for Airport usage, so I ended up with 15 post Offices from "Bayanat.ae" and 16 from Forsquare.

Plotted Post Office locations, Blue circles represent locations obtain using Forsquare, where Orange circles are locations obtain from "Bayant.ae", as I found out some duplicates and some missing locations from Forsquare.

Al Bateen Post Office

Abu Dhabi Central Post Office

Hamdan Post Office



Al Khaldiyah Post Office

Chamber of Commerce Post Office

Shabiyat Musafah Post Office

Since Bayanat.ae data was more accurate than Forsquare, I have decided to use “Bayanat.ae” data for the rest of my analysis.

To obtain and clean up neighborhoods of the City of Abu Dhabi, I have Used BeautifulSoup4 to pull data out of Wikipedia page, containing a list of neighborhoods of the City of Abu Dhabi, I ended up with 43 neighborhoods.

“ (page does not exist)” was found in some neighborhoods’ names, and it was cleaned up.
I dropped “Al Qubesat” and “Al Rehhan” as I was unable to obtain accurate or even clear location for them.

For some neighborhoods, I was unable to get the geolocation using Nominatim, for 5 of them it was due to spelling differences.
The list of neighborhoods below was changed to be able to obtain geolocation using Nominatim for them:

Name	New Name
Al Karama	Al Karamah
Al Maqtaa	Al Maqta
Al Meena	Al Mina
Al Moroor	Al Muroor
Al Khalidyah	al khalidiya

For the rest of the neighborhoods, I have used Google map to search and obtain the GeoLocations:

Neighborhood	Lat	Long
Al Manaseer [5]	24.453333	54.363889
Al Muzoon [6]	24.4081379	54.4338398
Al Ras Al Akhdar [7]	24.4596244	54.3190142
Bani Yas [8]	24.3063479	54.5995936
Al Zaab [9]	24.4667598	54.356339
Hideriyyat [10]	24.417739	54.3576465
Marina Village [11]	24.4772307	54.3179553
Qasr El Shatie [12]	24.4261292	54.380629
Al Samha [13]	24.695833	54.676111
Missing neighborhoods Added Manually		
Al Bahyah [14]	24.547706	54.6628323
Al Shahama [15]	24.524108	54.6770834
Al Rahbah [16]	24.605807	54.6944333

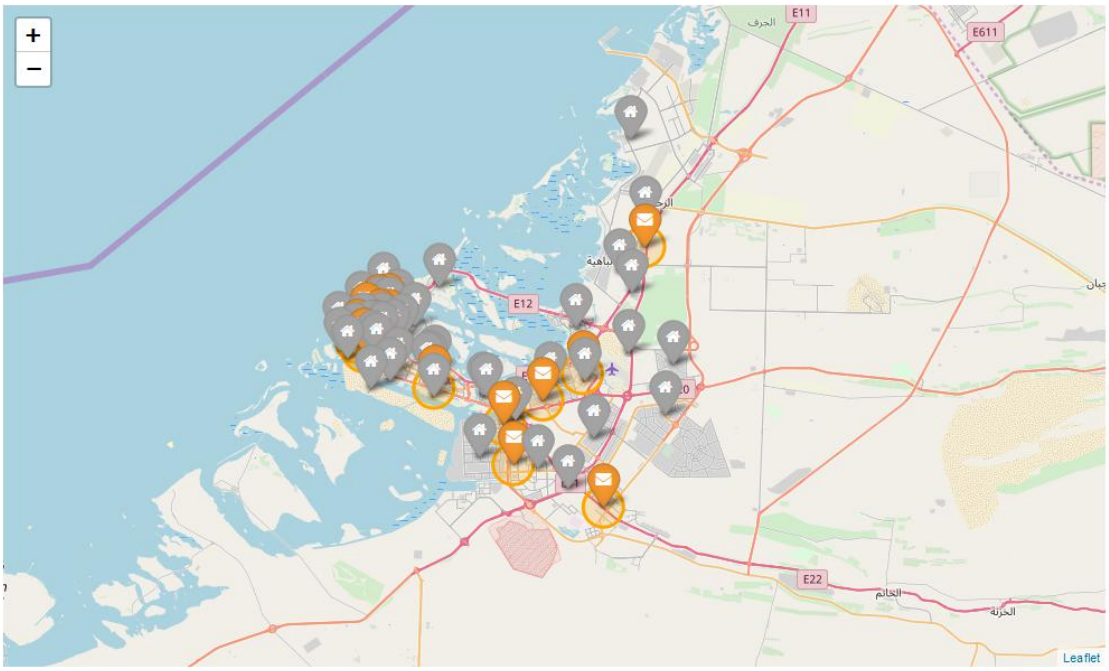
3. Methodology

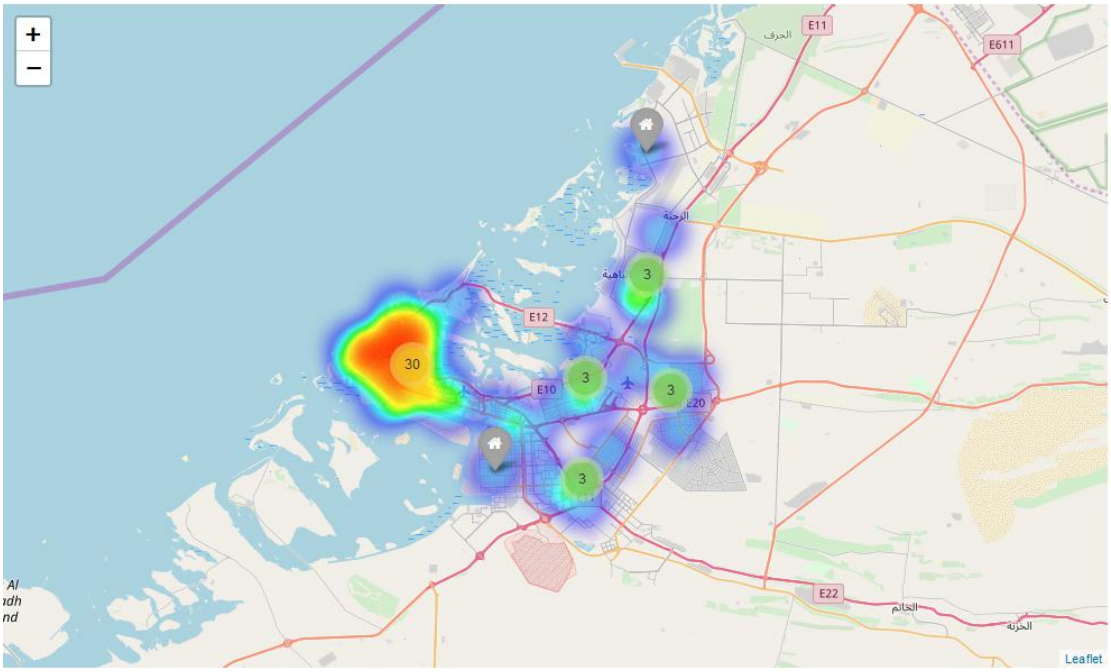
Before we can suggest a new post office location, we have to know what we currently have and how post offices cover Abu Dhabi city neighborhoods.
It was decided to use radius of 2500m for coverage, so, for current post offices I have calculated the total neighborhoods within 2500m radius, and for the neighborhoods I have calculated total post offices within the same radius of 2500m.

	name	lng	lat	nbCount
4	Al Nadi Sayahi Post Office	54.379688	24.498306	3
5	Bani Yas Post Office	54.643620	24.285358	0
7	Hamdan Post Office	54.366886	24.496719	5
9	Al Falah Post Office	54.379513	24.481054	7
10	Al Musafah Post Office	54.521702	24.377821	1

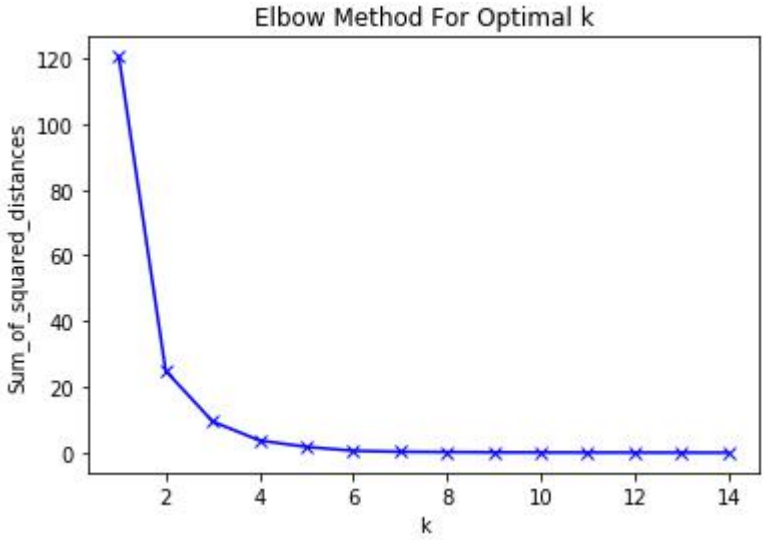
	Neighborhood	lat	lng	poCount
0	Al Aman	24.431990	54.426550	1
1	Al Bateen	24.451315	54.328144	2
2	Al Dhafrah	24.476147	54.369360	5
3	Al Falah	24.444696	54.728187	0
4	Al Karamah	24.465300	54.371509	4

I used python folium library to visualize geographic details of Abu dhabi City, its neighborhoods and post offices with coverage.





For analyzing current coverage, I have used unsupervised learning K-means algorithm to cluster the neighborhoods. K-Means algorithm is one of the most common cluster methods of unsupervised learning. First, I will run K-Means to cluster the neighborhoods into 3 clusters because when I analyzed the K-Means with Elbow method, it assured me the 3 degrees for optimum k of the K-Means.



Below is the neighborhoods table with Cluster:

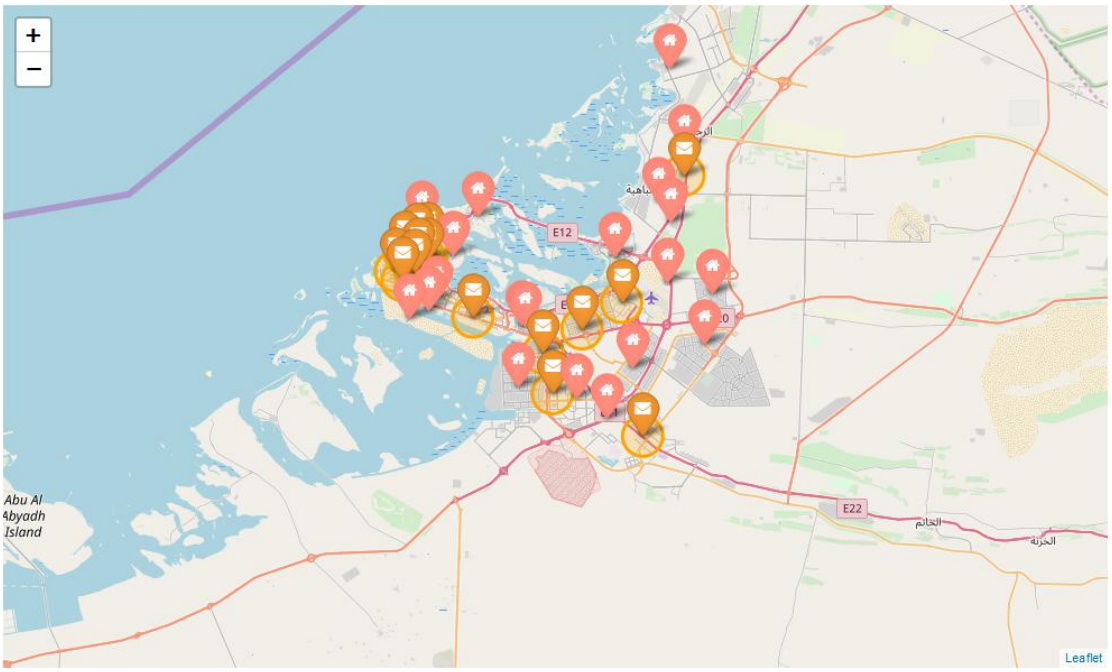
	Neighborhood	ClusterLabels	lat	lng	poCount
0	Al Aman	0	24.431990	54.426550	1
1	Al Bateen	0	24.451315	54.328144	2
2	Al Dhafrah	1	24.476147	54.369360	5
3	Al Falah	2	24.444696	54.728187	0
4	Al Karamah	1	24.465300	54.371509	4

Count of neighborhoods per cluster:

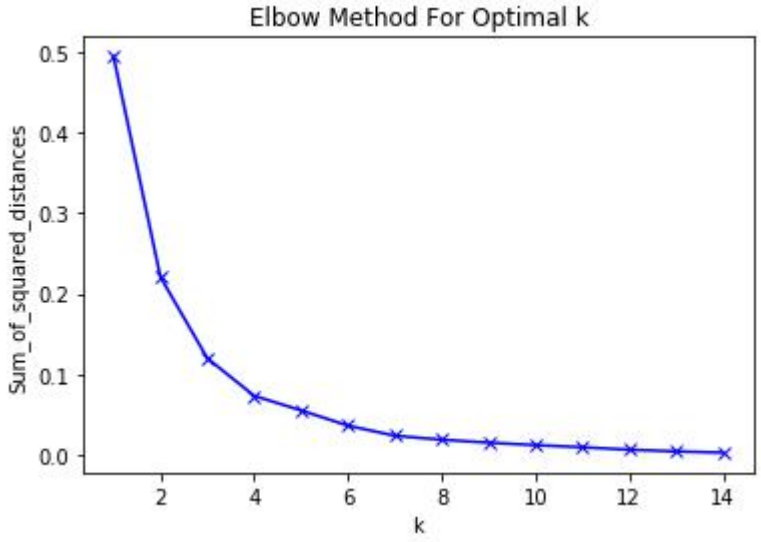
	poCount
	count
ClusterLabels	
0	15
1	9
2	20

I continued with the 20 neighborhoods within cluster 2 (with no or zero post offices covered), created a new data-frame for them, and visualized them using folium library:

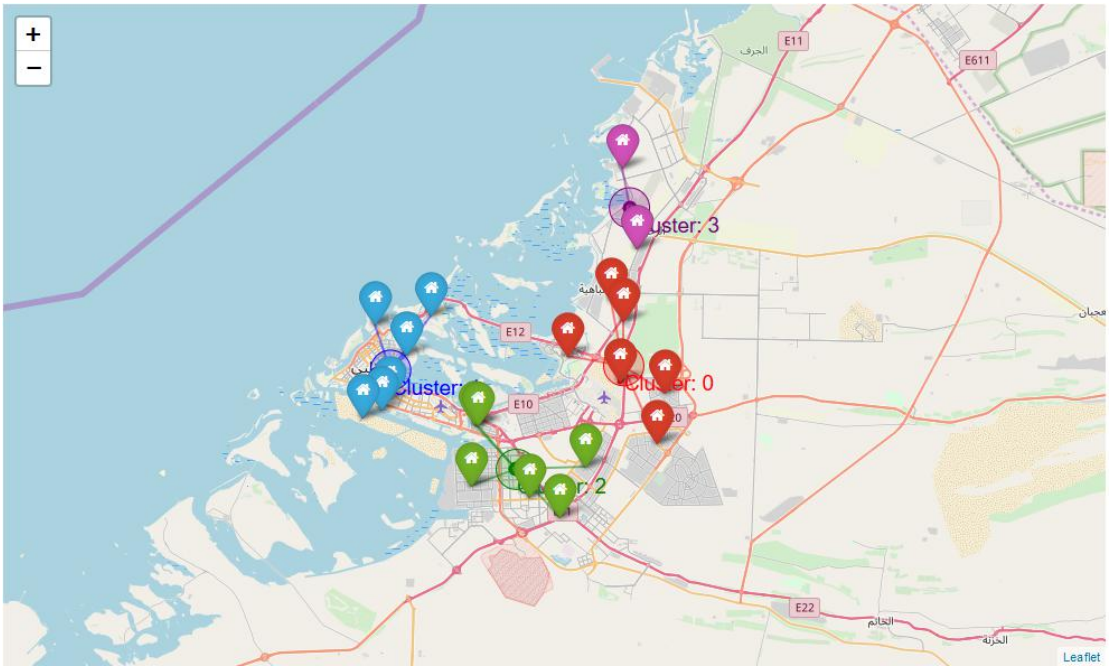
	Neighborhood	ClusterLabels	lat	lng	poCount
3	Al Falah	2	24.444696	54.728187	0
7	Al Madina	2	24.340878	54.490708	0
11	Al Maqta	2	24.407985	54.499372	0
13	Al Mina	2	24.520660	54.372529	0
15	Al Mushrif	2	24.436912	54.391006	0



I have used K-Mean algorithm to cluster the neighborhoods, aiming minimum post offices with maximum neighborhoods coverage, elbow method ensured that 4 degree for optimum k of the K-Means.

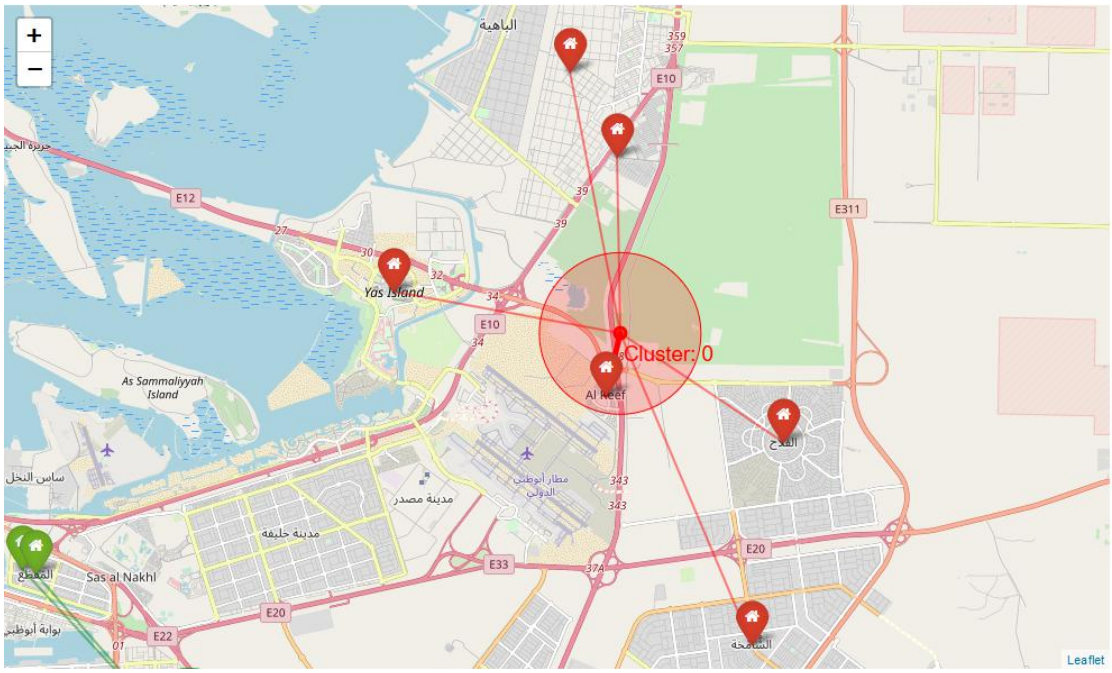


Visualizing clustered neighborhoods on map:

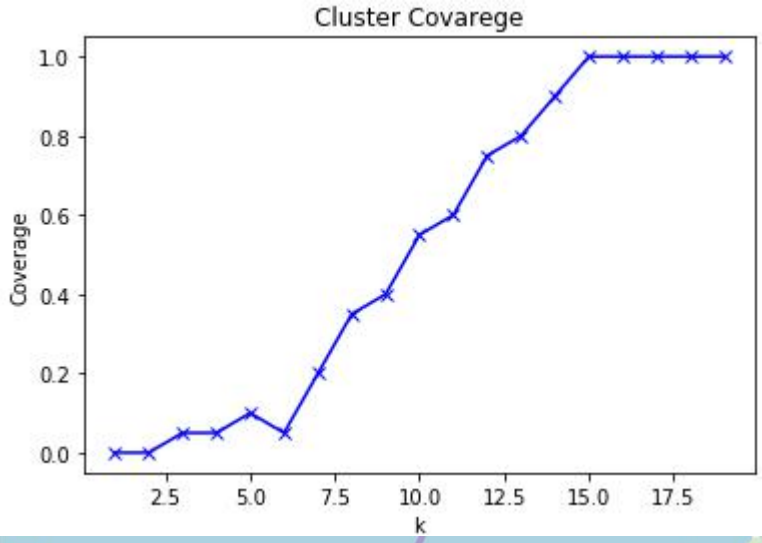


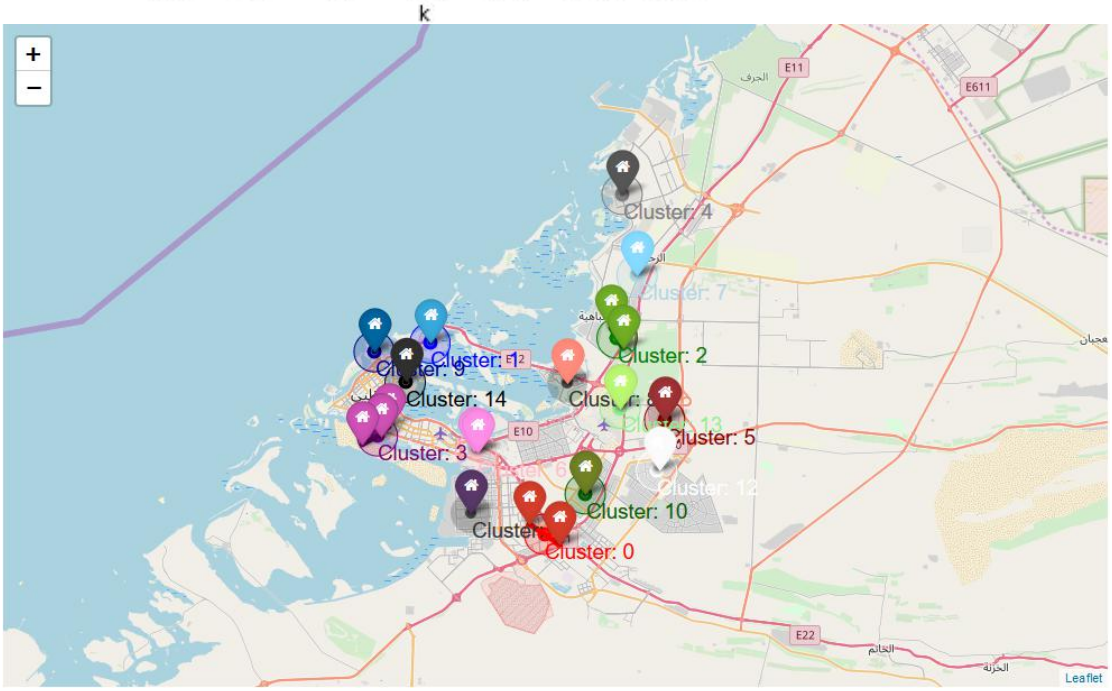
Using “cluster_centers_” as suggested post office location, I created a function to include covered neighborhoods within 2500m radius of suggested post office locations and ended up having only 1 neighborhood covered.

	Cluster	TotalNeighborhoods	TotalCovered
0	0	6	1
1	1	6	0
2	2	6	0
3	3	2	0



As I was aiming to have minimum post offices with maximum neighborhoods coverage, 4 clusters did not achieve the needed result, so, based on K-Mean Algorithm, I have created a function to calculate the maximum coverage with minimum clusters (post offices) and ended up with 15 clusters (post offices) to cover 20 uncovered neighborhoods.





	Cluster	TotalNeighborhoods	TotalCovered
0	0	2	2
1	1	1	1
2	2	2	2
3	3	3	3
4	4	1	1
5	5	1	1
6	6	2	2
7	7	1	1
8	8	1	1
9	9	1	1
10	10	1	1
11	11	1	1
12	12	1	1
13	13	1	1
14	14	1	1

0/15 Suggested Post Office(s) with no neighborhoods Covered.
15/15 Suggested Post Office(s) fully cover their clustered neighborhoods.
11/15 Suggested Post Office(s) cover 1 Neighborhood.


Saadiyat Beach
Cluster: 1


359
357
E10
39


Cluster: 3


Cluster: 4

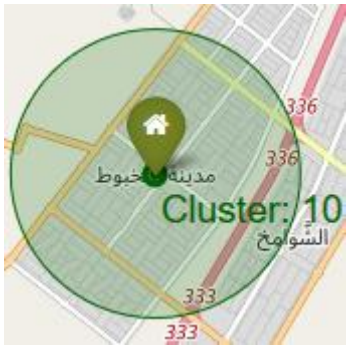

Cluster: 5


ساس النخل
بوابة أبوظبي


الرحبة
366
Cluster: 7


30
32
Cluster: 8
34


Cluster: 9
Lulu


336
336
Cluster: 10
333
333


Cluster: 11


E311


343

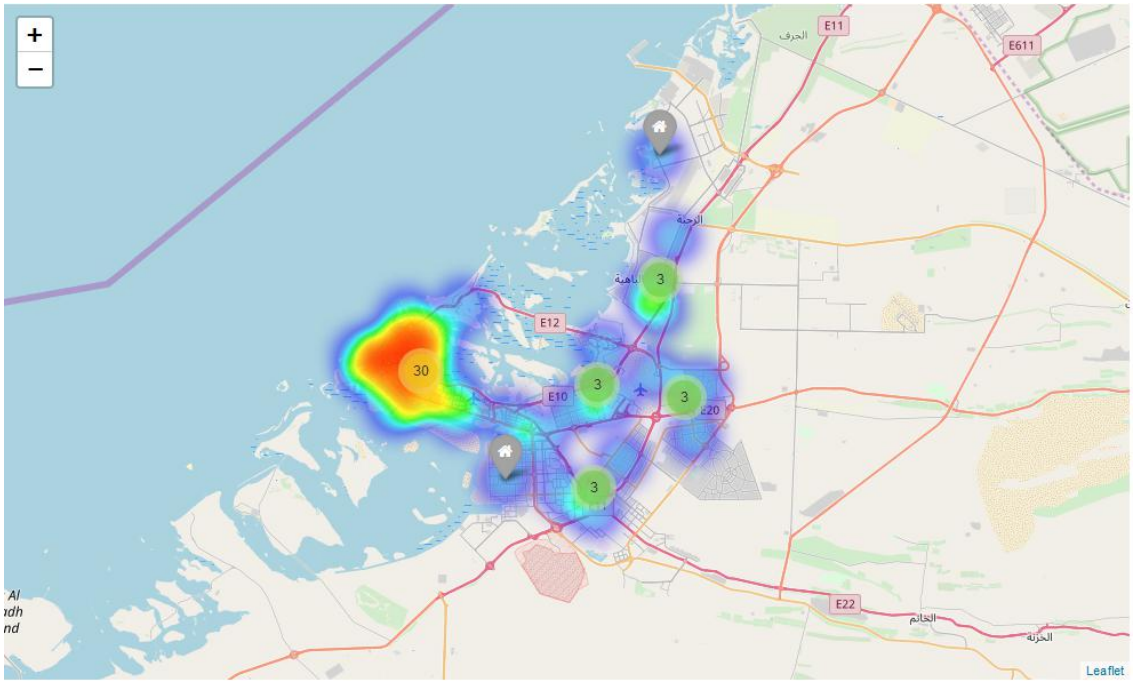

11
Cluster: 14
قصر البحر

4. Results

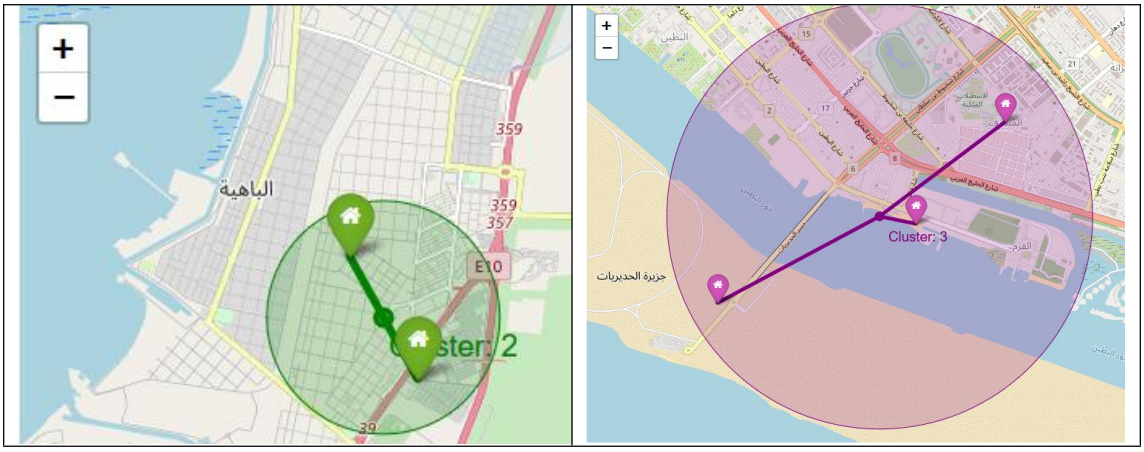
Started with the radius of 2.5 Km for coverage, I have used unsupervised learning K-means algorithm to cluster to the 20 uncovered neighborhoods into 4 clusters because when I analyzed the K-Means with elbow method it assured me the 4 degree for optimum k of the K-Means. However 4 k cluster resulted in only 1 neighborhood being covered, further analysis indicated that 15 clusters will cover all of the 20 uncovered neighborhoods.

5. Discussion

As included in the introduction section, determining the best location for new post offices requires more than geoLocations. Demographic information will add a major factor for suggesting a location, and even the number of post offices within each neighborhood, which may be the reason that we have around 30 post offices within Abu Dhabi city central (downtown), as the population / companies are to found more inside it.



Second-level Administrative Divisions are another data needed for a better analysis and suggestion of a location, which can help us locate post office(s) within neighborhood boundaries, for example, some suggested locations are not accurate, as in Cluster 2; the suggested location was to cover “Al Bahyah” and “Al Shahama” neighborhoods, however, it resulted in covering partial parts of them. Another example can be found in cluster 3, where the suggested location covering 3 neighborhoods was located on coast.



6. Conclusion

It is clear that more post offices do not necessarily mean more or better coverage, determining the location of a new post office will indeed depend on the study and analysis of demographic data.

7. References

[1] The Official Portal of the UAE Government
https://www.government.ae/en/about-the-uae/the-seven-emirates/abu-dhabi
[2] The official data portal of the UAE Government, https://bayanat.ae/en
[3] Abu Dhabi - Wikipedia https://en.wikipedia.org/wiki/Abu_Dhabi
[4] Emirates post offices GIS location
http://data.bayanat.ae/en_GB/dataset/emirates-post-offices-gis-location
[5] Google Maps