BATTLE OF THE NEIGHBOURHOODS — REPORT BY RUSHAB SHAH

1. Introduction

1.1 Background

Kenya, often dubbed the gateway to East Africa, is one of the fastest developing nations in Africa. In recent years, Nairobi (the capital city of Kenya) has seen an influx of immigration partly due to more people moving to Africa as international companies expand their operations in Africa. Nairobi also features the headquartes of the United Nations Environmental Program, another international organization that has a lot of international visitors.

The growing expatriate community means that there is an increase in demand for accommodation. As Nairobi is still a developing country, information sources about the availability of facilities are limited in number and reliability. This project is therefore, aiming to create a guide for expatriates that are either visiting or looking to move to Nairobi.

1.2 Problem Statement

The primary objective of this project is to provide a range of options, based on personal preferences, for the ideal neighbourhood to live when in Nairobi. The project will take into consideration the diversity in demographic of the visitors to Nairobi and try to provide suitable options for each demographic. It will achieve this by identifying 100 places within 500 metres of each Neighbourhood in Nairobi and, categorizing and grouping them, to indicate what services and facilities are nearby.

2. Data Acquisition

Three data sources will be used for this project:

- i) <u>Wikipedia Data Table:</u> The table featuring Nairobi's constituencies and their wards is present on https://en.wikipedia.org/wiki/Nairobi_County. However, this table will need to be cleaned and processed so that it becomes a dataframe with a list of wards and each of their respective constituencies.
- ii) <u>Geocoder Python Package:</u> The Wikipedia data table featuring the list of constituencies and wards does not have any location data available. No other source for this data was found so the Geocoder python package will have to be used to find the latitude and longitude for each ward.
- iii) Foursquare API: The foursquare API will also be used to get data about the venues near each ward. We will look at 100 of the venues that are within 500 metres of each of the respective wards. This data will then be processed to understand the frequency of each of the categories.

3. Methodology

i) The first step of the cleaning process was to transform the Wikipedia table, shown in figure A below into a table with a list of wards with their constituency, shown in figure B below.



	Ward	Constituency
0	Kitisuru	Westlands
1	Parklands/Highridge	Westlands
2	Kangemi	Westlands
3	Karura	Westlands
4	Mountain View	Westlands
5	Kilimani	Dagoretti North
6	Kawangware	Dagoretti North
7	Gatina	Dagoretti North
8	Kileleshwa	Dagoretti North
9	Kabiro	Dagoretti North
10	Mutu-ini, Ngand'o	Dagoretti South
11	Riruta	Dagoretti South
12	Uthiru/Ruthimitu	Dagoretti South
13	Waithaka	Dagoretti South
4.4	Varan	Langata

Figure A Figure B

The Beautiful Soup and Panda libraries were used to import the Wikipedia table into a data frame in the notebook and the 'split' and 'tolist' functions were used to separate the Wards into a list.

ii) The geocoder package was used to get the latitude and longitude for each ward. Due to the unreliable nature of this package, the Ward names were run in several different variations (by adding the city and country names), if they did not work the first time.

After a preliminary visualization on a Folium map, several of the ward locations (latitudes and longitudes) were found to be outside the city of Nairobi and other wards featured duplicate latitude and longitudes. Since no accurate location data was found for these wards, it was decided to exclude these wards from the dataset. An example of a datapoint (ward location) that is outside the relevant range is shown in figure C below.

	Ward	Constituency	Latitude	Longitude
0	Kitisuru	Westlands	-1.286762	36.761933
1	Parklands	Westlands	-1.273904	36.814902
2	Kangemi	Westlands	-1.269551	36.757060
3	Karura	Westlands	-1.250684	36.814611
4	Mountain View	Westlands	54.997122	-1.591959
5	Kilimani	Dagoretti North	-1.287442	36.784523
6	Kawangware	Dagoretti North	-1.299296	36.760191
7	Gatina	Dagoretti North	-1.277659	36.752845

Figure C

iii) To create a visualization and confirm that the relevant Wards were within Nairobi, the datapoints in the table were used to create a Folium map.

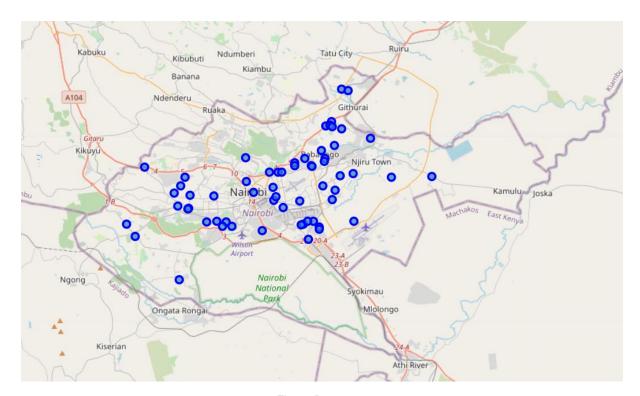


Figure D

iv) The Foursquare API was used to acquire data for 100 venues and the categories of these venues, that were within a 500 metre radius of each datapoint. This created the table shown by Figure E.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Kitisuru	-1.286762	36.761933	Toranj	-1.285340	36.762384	Persian Restaurant
1	Kitisuru	-1.286762	36.761933	Le Palanka	-1.286145	36.761896	African Restaurant
2	Kitisuru	-1.286762	36.761933	The Arbor	-1.287757	36.761574	Breakfast Spot
3	Kitisuru	-1.286762	36.761933	MumyDada - Nigerian & Kenyan Cuisine	-1.284435	36.764809	Bar
4	Kitisuru	-1.286762	36.761933	The Ambo Garden	-1.283841	36.761392	Lounge

Figure E

One hot coding was applied to the venue category and then the data was grouped by Neighborhood (Wards are now called Neighborhoods) to form another table showing the 10 most common venues for each location, as shown below.

	Neighbourhood	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	Dandora Area I	Campground	Wine Bar	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant	Fast Food Restaurant
1	Eastleigh North	African Restaurant	Bus Station	Flea Market	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant
2	Embakasi	Lounge	Wine Bar	French Restaurant	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant	Fast Food Restaurant
3	Harambee	African Restaurant	Steakhouse	Ethiopian Restaurant	Lounge	Shopping Mall	Fried Chicken Joint	Nightclub	Pizza Place	Pharmacy	Convenience Store
4	Huruma	Restaurant	Bar	Wine Bar	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant

Figure F

v) K Means Clustering was then applied to group the Neighbourhoods into several clusters. Different values of K were tested to identify the value to be used.

For each iteration, a map with different colours representing the clusters was plotted, and tables showing the clustered groupings were created.

4. Results

A K value of 6 was used to plot the following clustered map.

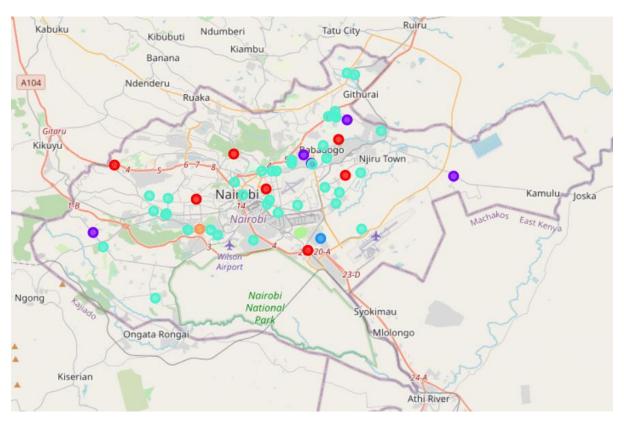


Figure G

The datapoints for each cluster are shown below:

Cluster 0:

	Ward	Constituency	Latitude	Longitude	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
3	Karura	Westlands	-1.250684	36.814611	0	Hotel	Boutique	North Indian Restaurant	Park	Gun Range	Gift Shop	Convenience Store	Cosmetics Shop	Cricket Ground	Department Store
5	Kilimani	Dagoretti North	-1.287442	36.784523	0	Hotel	Lounge	Seafood Restaurant	Gym	Gaming Cafe	Japanese Restaurant	Ethiopian Restaurant	Music Venue	Nightclub	Gym / Fitness Center
11	Uthiru	Dagoretti South	-1.260194	36.718377	0	Gym	Pub	Dry Cleaner	Bus Station	Wine Bar	Food & Drink Shop	Cosmetics Shop	Cricket Ground	Department Store	Diner
34	Lucky Summer	Ruaraka	-1.239308	36.898786	0	Gym	Soccer Field	Market	Wine Bar	Food & Drink Shop	Convenience Store	Cosmetics Shop	Cricket Ground	Department Store	Diner
36	Imara Daima	Embakasi South	-1.328590	36.874400	0	Train Station	Gym	Café	Wine Bar	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner
47	Komarock	Embakasi Central	-1.268377	36.904706	0	Gym / Fitness Center	Gym	Pizza Place	Shopping Mall	Wine Bar	Flea Market	Convenience Store	Cosmetics Shop	Cricket Ground	Department Store
65	Nairobi Central	Starehe	-1.279296	36.840374	0	African Restaurant	Cosmetics Shop	Soccer Field	Lounge	Moving Target	Food Court	Cricket Ground	Department Store	Diner	Dry Cleaner

Cluster 1:

	Ward	Constituency	Latitude	Longitude	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
12	Waithaka	Dagoretti South	-1.313882	36.700900	1	Bar	Antique Shop	Other Great Outdoors	Italian Restaurant	Wine Bar	Food Court	Cricket Ground	Department Store	Diner	Dry Cleaner
28	Kasarani	Kasarani	-1.223271	36.905904	1	Bar	Wine Bar	French Restaurant	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant	Fast Food Restaurant
30	Ruai	Kasarani	-1.268997	36.991548	1	Bar	Restaurant	Shopping Mall	Wine Bar	Food & Drink Shop	Convenience Store	Cosmetics Shop	Cricket Ground	Department Store	Diner
33	Mathare North	Ruaraka	-1.251788	36.870772	1	Bar	Fish & Chips Shop	Wine Bar	French Restaurant	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant
72	Huruma	Mathare	-1.258309	36.876976	1	Restaurant	Bar	Wine Bar	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant

Cluster 2:

Ward C	Constituency	Latitude	Longitude	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
51 Embakasi	Embakasi Fast	-1.318697	36.884473	2	Lounge	Wine Bar	French Restaurant	Cosmetics	Cricket Ground	Department	Diner	Dry Cleaner	Ethiopian Restaurant	Fast Food Restaurant

Cluster 3:

	Ward	Constituency	Latitude	Longitude	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	Kitisuru	Westlands	-1.286762	36.761933	3	African Restaurant	Bar	Korean Restaurant	Lounge	Diner	Chinese Restaurant	Breakfast Spot	Beer Garden	Persian Restaurant	Bakery
1	Parklands	Westlands	-1.273904	36.814902	3	Hotel	Department Store	Bookstore	Gastropub	Hotel Bar	Indian Restaurant	Museum	Nightclub	Coffee Shop	Pub
6	Kawangware	Dagoretti North	-1.299296	36.760191	3	Bar	Fast Food Restaurant	Wine Bar	Shopping Mall	Coffee Shop	Gift Shop	Lounge	Italian Restaurant	Ice Cream Shop	Gym / Fitness Center
8	Kileleshwa	Dagoretti North			3	Bar	Fast Food Restaurant	Wine Bar	Shopping Mall	Coffee Shop	Gift Shop	Lounge	Italian Restaurant	Ice Cream Shop	Gym / Fitness Center
9					3	Cocktail Bar	Movie Theater	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant	Fast Food Restaurant	Fish & Chips Shop
10					3	Italian Restaurant	Auto Garage	Wine Bar	French Restaurant	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant
13	Karen			36.709588	3	Shopping Mall	Café	Coffee Shop	Gym / Fitness Center	Indian Restaurant	Pub	Japanese Restaurant	Snack Place	Italian Restaurant	Supermarket
14	Nairobi West	Langata	-1.367127	36.751154	3	Art Gallery	Café	Garden	Resort	Wine Bar	Food & Drink Shop	Cosmetics Shop	Cricket Ground	Department Store	Diner
15	Nyayo Highrise	Langata	-1.316005	36.801733	3	Lounge	Athletics & Sports	Gun Range	BBQ Joint	Bar	Wine Bar	Fast Food Restaurant	Food & Drink Shop	Flea Market	Fish & Chips Shop
16	South C	Langata	-1.320401	36.830525	3	Coffee Shop	Beer Garden	Wine Bar	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant
17	Laini Saba	Kibra	-1.312348	36.796114	3	Pharmacy	Hotel Bar	Food Court	Convenience Store	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant
20	Sarang'ombe	Kibra	-1.312267	36.777459	3	African Restaurant	Flea Market	Campground	Pub	Convenience Store	Gun Range	Fish & Chips Shop	Cosmetics Shop	Health & Beauty Service	Cricket Ground
21	Githurai	Roysambu	-1.216305	36.895938	3	African Restaurant	Pub	Bar	Bus Station	Moving Target	Fast Food Restaurant	Food & Drink Shop	Sausage Shop	Lounge	Sculpture Garden
22	Kahawa West	Roysambu	-1.185685	36.905793	3	Bus Stop	Sports Bar	Bar	Flea Market	Wine Bar	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner
24	Roysambu	Roysambu	-1.220427	36.891057	3	Hookah Bar	Shopping Mall	Department Store	Café	Pizza Place	Burger Joint	Restaurant	Lounge	Food Court	Coffee Shop
25	Kahawa	Roysambu	-1.186963	36.912022	3	Bus Stop	Sports Bar	Wine Bar	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant
26	Clay City	Kasarani	-1.219800	36.894539	3	Pub	Hookah Bar	Food & Drink Shop	Hostel	Steakhouse	Fast Food Restaurant	Moving Target	Cosmetics Shop	Cricket Ground	Department Store
27	Mwiki	Kasarani	-1.232190	36.933101	3	Botanical Garden	Child Care Service	Bus Station	Convenience Store	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant	Fast Food Restaurant
29	Njiru Shopping	Kasarani	-1.221505	36.896905	3	African Restaurant	Sculpture Garden	Pub	Steakhouse	Fast Food Restaurant	Lounge	Food & Drink Shop	Convenience Store	Cosmetics Shop	Cricket Ground
31	Babadogo	Ruaraka	-1.244134	36.886464	3	Outdoor Sculpture	Business Service	Bus Stop	Bus Station	Bookstore	Food Court	Cricket Ground	Department Store	Diner	Dry Cleaner
32	Utalii	Ruaraka	-1.255719	36.861215	3	African Restaurant	Pool	Bus Stop	Stadium	Juice Bar	Food & Drink Shop	Cosmetics Shop	Cricket Ground	Department Store	Diner
42	Dandora Area I	Embakasi North	-1.254487	36.889509	3	Campground	Wine Bar	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant	Fast Food Restaurant
45	Kayole North	Embakasi Central	-1.265984	36.916945	3	Convenience Store	Bus Station	Fast Food Restaurant	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant
50	Upper Savanna	Embakasi East	-1.283888	36.821886	3	Coffee Shop	Café	African Restaurant	Bar	Fried Chicken Joint	Nightclub	Sandwich Place	Italian Restaurant	Fast Food Restaurant	Ice Cream Shop
52	Utawala	Embakasi East	-1.311575	36.917260	3	Café	Pub	Optical Shop	Wine Bar	Coffee Shop	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner
54	Umoja I	Embakasi West	-1.277867	36.888089	3	Shoe Store	Pizza Place	Bar	Lounge	Fried Chicken Joint	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner
55	Umoja II	Embakasi West	-1.281873	36.899340	3	Convenience Store	Nightlife Spot	Sports Bar	Fast Food Restaurant	Wine Bar	Food & Drink Shop	Cosmetics Shop	Cricket Ground	Department Store	Diner
57	Maringo	Makadara	-1.292071	36.865726	3	Arcade	Wine Bar	French Restaurant	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant	Fast Food Restaurant
59	Harambee	Makadara	-1.290941	36.897216	3	African Restaurant	Steakhouse	Ethiopian Restaurant	Lounge	Shopping Mall	Fried Chicken Joint	Nightclub	Pizza Place	Pharmacy	Convenience Store
60	Makongeni	Makadara	-1.297993	36.850010	3	Paper / Office Supplies Store	Steakhouse	Wine Bar	Food & Drink Shop	Convenience Store	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner
61	Pumwani	Kamukunji	-1.291626	36.841265	3	Bakery	Bus Station	Wine Bar	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant
62	Eastleigh North	Kamukunji	-1.288144	36.843018	3	African Restaurant	Bus Station	Flea Market	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant
67	Pangani	Starehe	-1.264899	36.837271	3	Shopping Mall	Gym	Beer Garden	Department Store	Fast Food Restaurant	Moving Target	Metro Station	Brewery	Restaurant	Food & Drink Shop
70	Hospital	Mathare	-1.264837	36.845300	3	Bus Stop	Bakery	Bar	Smoke Shop	Wine Bar	French Restaurant	Cricket Ground	Department Store	Diner	Dry Cleaner
73	Ngei	Mathare	-1.258833	36.861253	3	Pool	Stadium	Shopping Mall	Wine Bar	Food & Drink Shop	Convenience Store	Cosmetics Shop	Cricket Ground	Department Store	Diner
74	Mlango Kubwa	Mathare	-1.264950	36.849155	3	Pool	Bus Stop	Smoke Shop	Wine Bar	Food & Drink Shop	Convenience Store	Cosmetics Shop	Cricket Ground	Department Store	Diner
15	Kiamaiko	Mathare	-1 259506	36.877642	3	BBQ Joint	Restaurant	Wine Bar	Food Court	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant

Cluster 4:

	Ward	Constituency	Latitude	Longitude	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
38	Kwa Reuben	Embakasi South	-1.314298	36.869721	4	Recreation Center	Wine Bar	Cocktail Bar	Convenience Store	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant
41	Kariobangi North	Embakasi North	-1.315138	36.867340	4	Recreation Center	Wine Bar	Cocktail Bar	Convenience Store	Cosmetics Shop	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant

Cluster 5:

Ward Co	onstituency	Latitude	Longitude	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
19 Makina	Kibra -	1.311237	36.786862	5	Athletics & Sports	Wine Bar	French Restaurant	Cosmetics	Cricket Ground	Department Store	Diner	Dry Cleaner	Ethiopian Restaurant	Fast Food Restaurant

5. Discussion

The best value for K in the K means clustering algorithm was found to be 6 as it delivered the most even spread of clusters. The 6 clusters formed are of varying lengths with Cluster 3 being the largest with 37 Neighborhoods and, Cluster 2 and 5 the smallest with just 1 neighbourhood. It is easier to see the similarities in the smaller clusters, for example, cluster 4 has the same top 10 most common venues and most elements in Cluster 1 have Bar as their most common venue.

Not all the datapoints were used, as the location data (latitude and longitude) could not be found for all the datapoints, however, the data points used represented most neighbourhoods in the city of Nairobi. It is also worth noting that some of the wards do not have any other information available about them on Wikipedia (data source), so the accuracy of this information is limited to the availability of the data.

The similarities in the largest cluster, Cluster 3, are harder to see because the cluster has many more elements. Despite this, there are many venues that make frequent appearances for example, restaurants, cafes, bars, bus stops, shopping malls etc. The venues present display a wide range of facilities and services that potential expatriates may be interested in. This indicates that this cluster probably features the best neighbourhoods to stay in for those visiting Nairobi.

6. Conclusion

It can be concluded that the K means Clustering algorithm was successfully performed to group the neighbourhoods in Nairobi, according to the occurrence of different venues. This has helped in understanding the availability of facilities and services to different neighbourhoods in Nairobi. There were some limitations in acquiring the location data for several neighbourhoods, however, this is not thought to have a huge influence on the outcome.

The clusters developed by this model will provide visitors with a detailed guide about the availability of different services in Nairobi. However, to make a more informed decision, more variables will be needed concerning the demographic of the individual/group, price flexibility, proximity to work etc. and the importance of each of these factors. These are all suggestions for possible improvements to the model.