

Segmenting and Clustering Socio-Economic Activities in Kenyan Counties

(Sammy Ongaya)

Business Statement

- Each County government is obligated to generate its own revenue and supplement the income disbursed from the national government. Therefore the each Counties stakeholders and administrator sought to analyse the most viable economical activities that will increase revenue and improve the livelihood of its people.
- Also for any organization and stakeholders wishing to expand their business operation to different Counties will need to understand the best County to invest in.
- With this analysis the stakeholders are well informed on the most appropriate economical activities to invest in each county and make informed decision that will yield high ROI.

Data

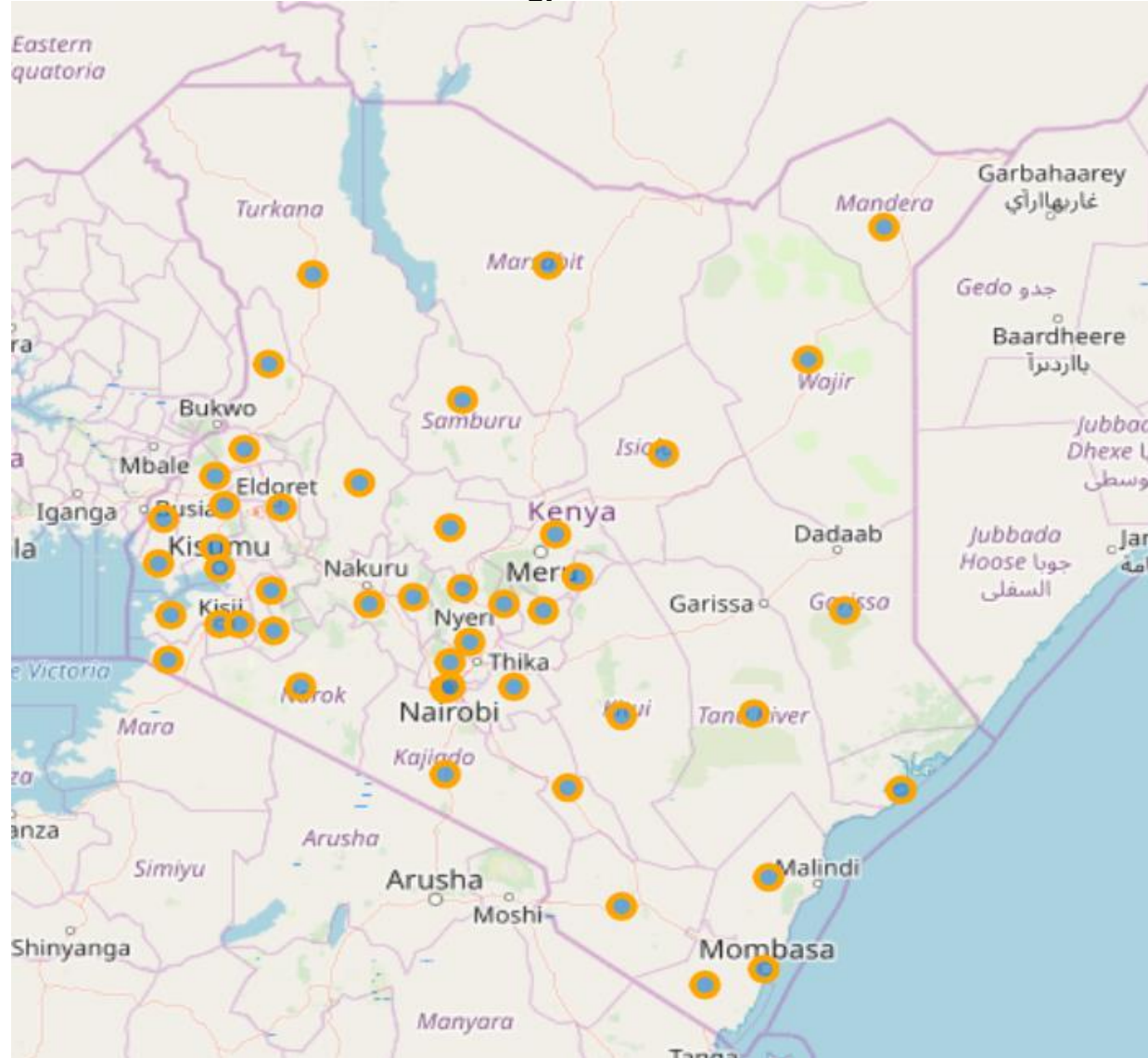
- We use data from 47 Counties in Kenya. The data is extracted from the [wikipedia](#) website through web scrapping using Python based library.
- The data has County name, Area in KM squared and Population Census as of 2009 among other details. We use the geopy library to get the coordinates of each county. We then create a map of Kenya and plot the coordinates of each county using the folium library.
- We use the Foursquare API to extract relevant information such as facilities within a radius of 25km for all Counties and get the venue categories for each facility.

Distribution of Counties by Population and Area

Top 10 Counties by Population						
	Code	County	Former Province	Area_km2	Population	Capital
46	47.0	Nairobi (County)	Nairobi (Province)	694.9	3138369	Nairobi (City)
36	37.0	Kakamega	Western	3033.8	1660651	Kakamega
21	22.0	Kiambu	Central	2449.2	1623282	Kiambu
31	32.0	Nakuru	Rift Valley	7509.5	1603325	Nakuru
38	39.0	Bungoma	Western	2206.9	1375063	Bungoma
11	12.0	Meru	Eastern	6930.1	1356301	Meru
44	45.0	Kisii	Nyanza	1317.9	1152282	Kisii
2	3.0	Kilifi	Coast	12245.9	1109735	Kilifi
15	16.0	Machakos	Eastern	5952.9	1098584	Machakos
8	9.0	Mandera	North Eastern	25797.7	1025756	Mandera
Least 10 Counties by Population						
	Code	County	Former Province	Area_km2	Population	Capital
23	24.0	West Pokot	Rift Valley	8418.2	512690	Kapenguria
30	31.0	Laikipia	Rift Valley	8696.1	399227	Nanyuki
27	28.0	Elgeyo-Marakwet	Rift Valley	3049.7	369998	Iten
12	13.0	Tharaka-Nithi	Eastern	2409.5	365330	Kathwana
9	10.0	Marsabit	Eastern	66923.1	291166	Marsabit
5	6.0	Taita–Taveta	Coast	17083.9	284657	Mwatate
3	4.0	Tana River	Coast	35375.8	240075	Hola
24	25.0	Samburu	Rift Valley	20182.5	223947	Maralal
4	5.0	Lamu	Coast	6497.7	191539	Lamu
10	11.0	Isiolo	Eastern	25336.1	143294	Isiolo

Top 10 Counties by Size						
	Code	County	Former Province	Area_km2	Population	Capital
22	23.0	Turkana	Rift Valley	71597.8	855399	Lodwar
9	10.0	Marsabit	Eastern	66923.1	291166	Marsabit
7	8.0	Wajir	North Eastern	55840.6	661941	Wajir
6	7.0	Garissa	North Eastern	45720.2	623060	Garissa
3	4.0	Tana River	Coast	35375.8	240075	Hola
8	9.0	Mandera	North Eastern	25797.7	1025756	Mandera
10	11.0	Isiolo	Eastern	25336.1	143294	Isiolo
14	15.0	Kitui	Eastern	24385.1	1012709	Kitui
33	34.0	Kajiado	Rift Valley	21292.7	687312	Kajiado
24	25.0	Samburu	Rift Valley	20182.5	223947	Maralal
Least 10 Counties by Size						
	Code	County	Former Province	Area_km2	Population	Capital
38	39.0	Bungoma	Western	2206.9	1375063	Bungoma
41	42.0	Kisumu	Nyanza	2009.5	968909	Kisumu(City)
35	36.0	Bomet	Rift Valley	1997.9	730129	Bomet
39	40.0	Busia	Western	1628.4	743946	Busia
44	45.0	Kisii	Nyanza	1317.9	1152282	Kisii
19	20.0	Kirinyaga	Central	1205.4	528054	Kerugoya / Kutus
45	46.0	Nyamira	Nyanza	912.5	598252	Nyamira
46	47.0	Nairobi (County)	Nairobi (Province)	694.9	3138369	Nairobi (City)
37	38.0	Vihiga	Western	531.3	554622	Vihiga
0	1.0	Mombasa (County)	Coast	212.5	939370	Mombasa (City)

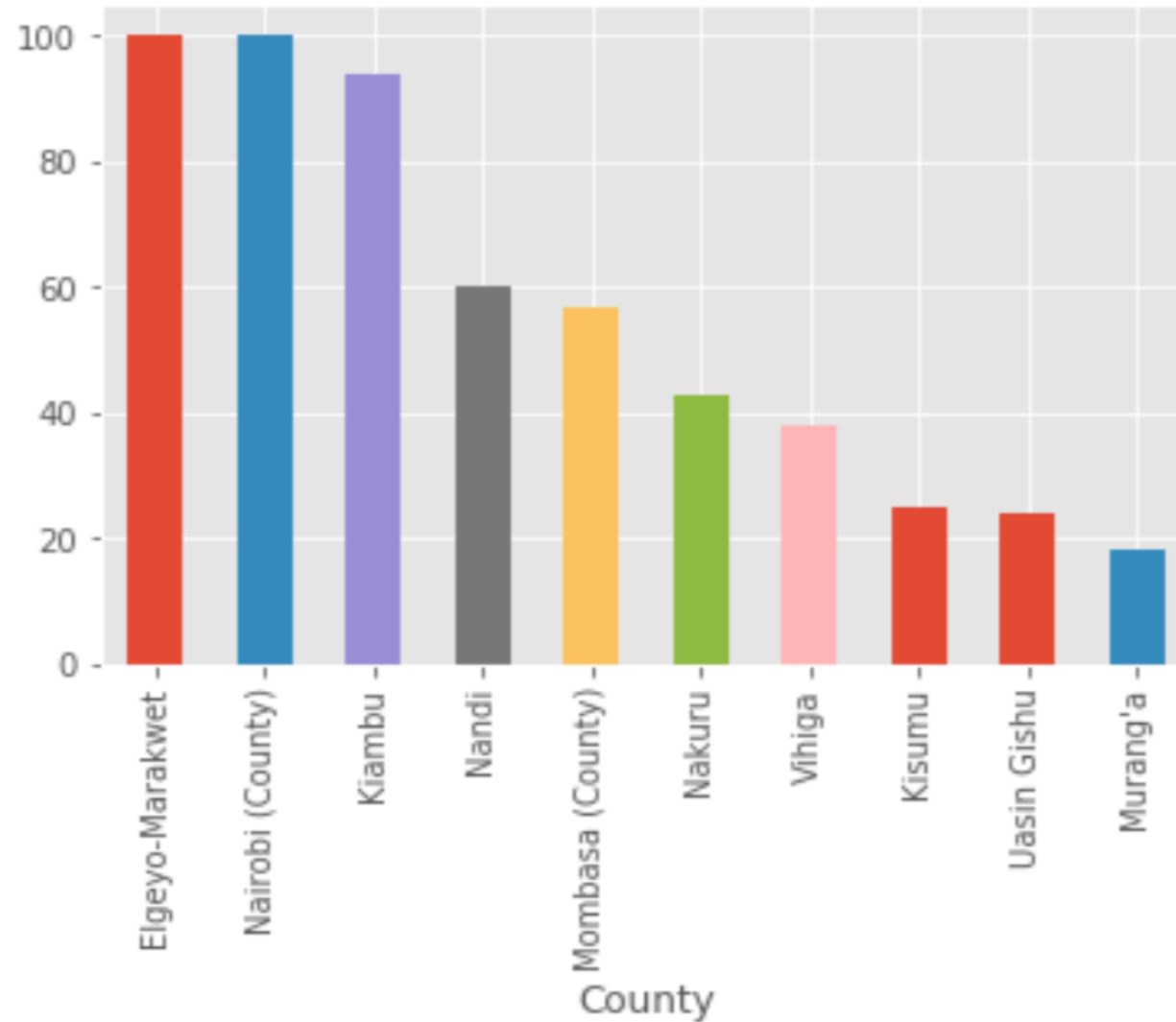
Kenya Counties



Counties and Venues

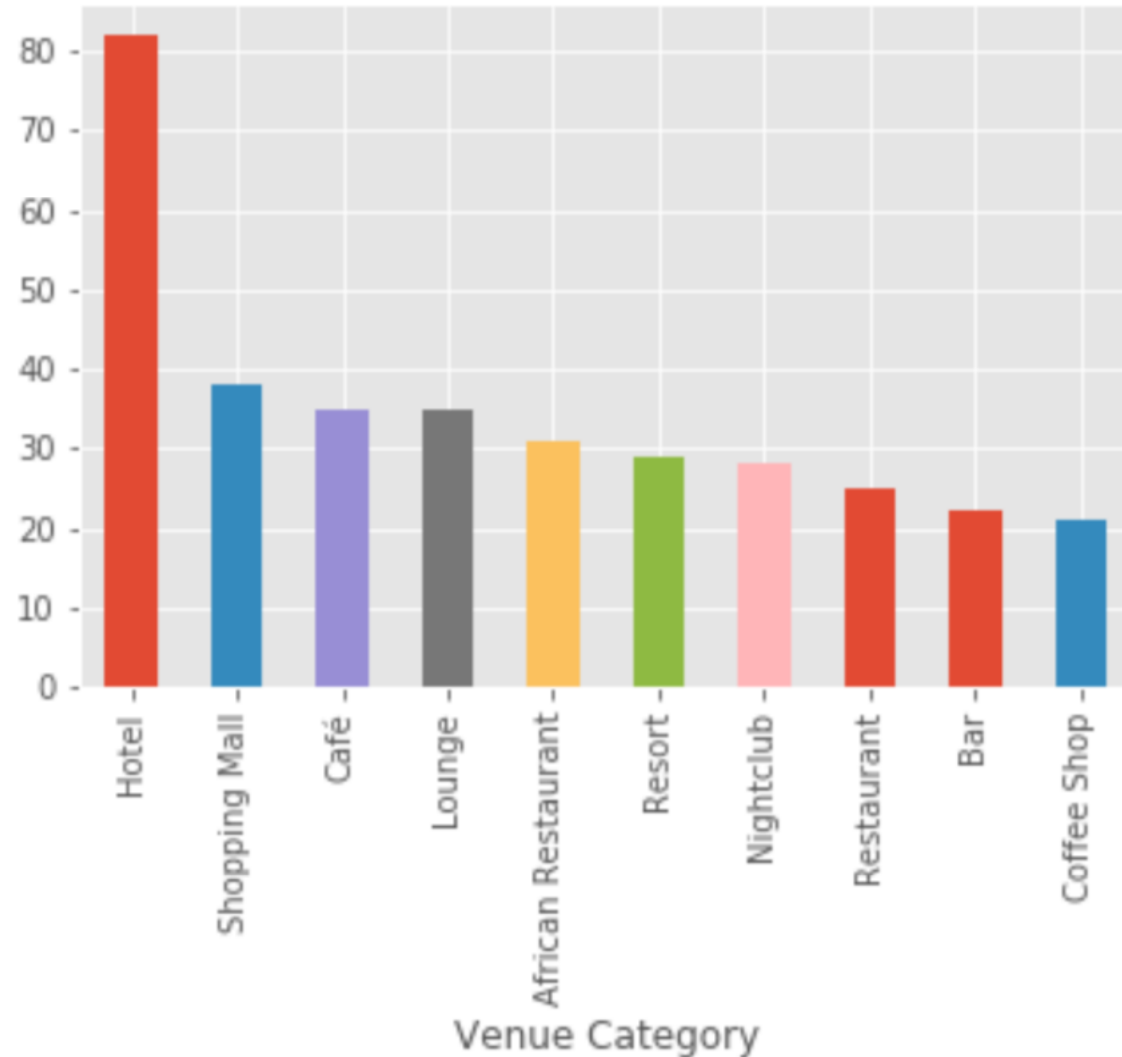
15 Counties with Most Venues		15 Counties with Least Venues	
Elgeyo-Marakwet	100	Lamu	5
Nairobi (County)	100	Trans-Nzoia	5
Kiambu	94	Nyamira	4
Nandi	60	Kericho	4
Mombasa (County)	57	Nyandarua	4
Nakuru	43	Homa Bay	4
Vihiga	38	Siaya	4
Kisumu	25	Busia	4
Uasin Gishu	24	Machakos	4
Murang'a	18	Kajiado	3
Nyeri	11	Tharaka-Nithi	3
Kirinyaga	10	Bungoma	3
Makueni	10	Kitui	1
Kakamega	9	Kwale	1
Bomet	8	Kilifi	1

Counties and Venues



Venue Categories

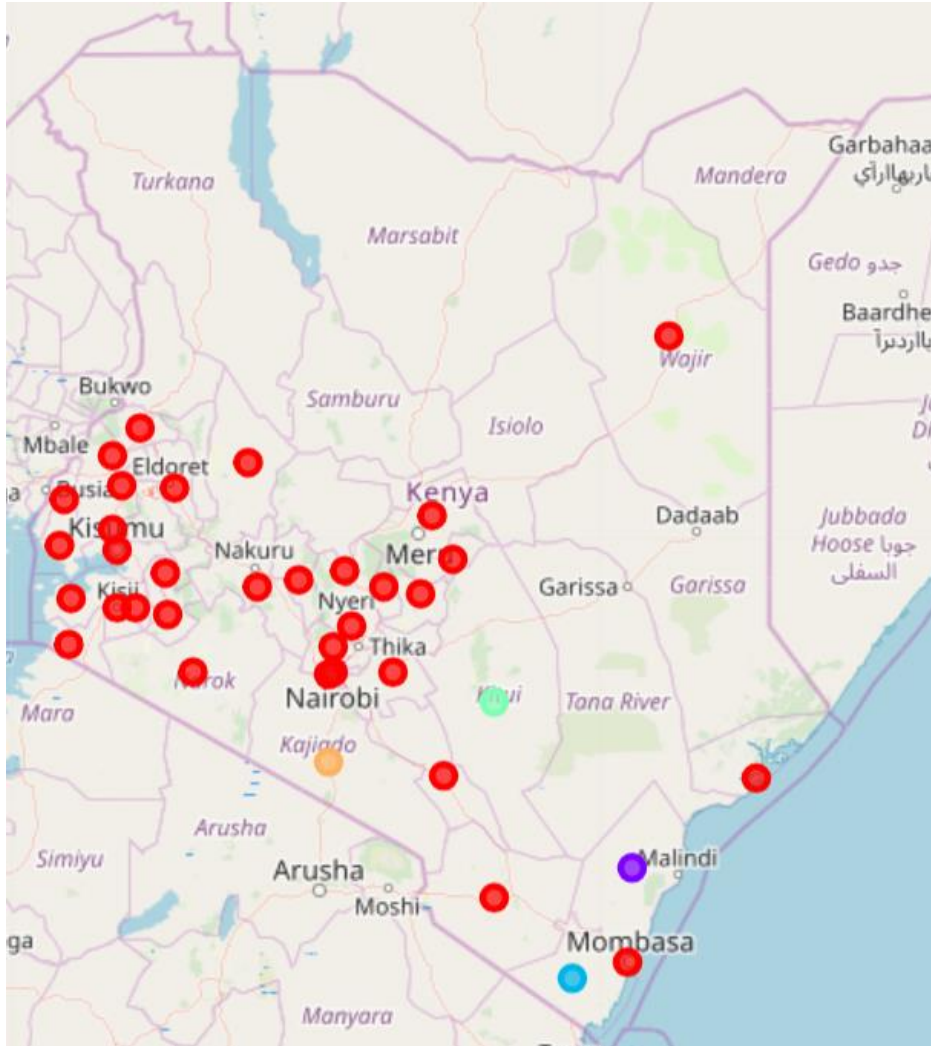
Top Venue Categories	
Hotel	82
Shopping Mall	38
Café	35
Lounge	35
African Restaurant	31
Resort	29
Nightclub	28
Restaurant	25
Bar	22
Coffee Shop	21
Least Venue Categories	
Gift Shop	1
General Travel	1
Gaming Cafe	1
French Restaurant	1
Farm	1
English Restaurant	1
Duty-free Shop	1
Dumpling Restaurant	1
Dessert Shop	1
Harbor / Marina	1



10 most common venues in each County

	County	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	Baringo	Lake	Hotel	Zoo	Diner	English Restaurant	Electronics Store	Eastern European Restaurant	Duty-free Shop	Dumpling Restaurant	Dessert Shop
1	Bomet	African Restaurant	Café	Bar	River	IT Services	Vineyard	Travel & Transport	Lounge	Convenience Store	Convention Center
2	Bungoma	Historic Site	Shop & Service	Food	Zoo	Dessert Shop	Eastern European Restaurant	Duty-free Shop	Dumpling Restaurant	Diner	Department Store
3	Busia	Vegetarian / Vegan Restaurant	Cocktail Bar	Bus Station	River	Zoo	Dessert Shop	Eastern European Restaurant	Duty-free Shop	Dumpling Restaurant	Diner
4	Elgeyo-Marakwet	Hotel	Lounge	Café	African Restaurant	Italian Restaurant	Coffee Shop	Burger Joint	Bar	Shopping Mall	Chinese Restaurant
5	Embu	Beer Garden	Shopping Mall	Nightclub	Bus Station	Convenience Store	Diner	Electronics Store	Eastern European Restaurant	Duty-free Shop	Dumpling Restaurant
6	Homa Bay	Plaza	Beach	Hotel	Restaurant	Zoo	Diner	Electronics Store	Eastern European Restaurant	Duty-free Shop	Dumpling Restaurant
7	Kajiado	Campground	Mountain	Zoo	Ethiopian Restaurant	Coffee Shop	Concert Hall	Convenience Store	Convention Center	Department Store	Dessert Shop
8	Kakamega	Hotel	Lounge	Food	Department Store	Restaurant	Electronics Store	Convenience Store	Nightclub	English Restaurant	Concert Hall

Our Clustering Algorithm



- We have 5 clusters represented by colors red for cluster 1, blue for cluster 2, orange for cluster 3, green for cluster 4 and purple for cluster 5.
- Most data points fall under cluster 1 which comprises of food courts, hotels, resorts and restaurants.
- Some Counties such as Turkana, Marsabit, mandera and Samburu don't have venue data.

Conclusion

- This analysis provides valuable insights for investors who want to start business in Kenya in different Counties.
- This analysis is limited to the data collected from the foursquare API which is not sufficient.
- The analysis can be improved further by collecting more data.