Determining The Best ATM Location

Introduction/Business Problem

Centenary Bank is a commercial bank in Uganda licensed by the Bank of Uganda, the central bank and national banking regulator

Centenary Bank is of the second largest financial services provider in Uganda with over 1.8m consumers, employing over 2754 staff and has an asset base of UShs. 3.567 trillion as of December 31, 2019

The bank has a network of 71 bank branches together with 189 linked automated teller machines at 123 locations in the Central, Western, Northern, and Eastern Regions.

Being the fastest growing bank in Uganda with a high number of customers, the bank through ATMs extends its services to the customers in different regions. Determining the best ATM allocation is a crucial thing in extending the services to customers but a great challenge.

In this case an ATM cannot be placed in a region where there are no customers, or less populated. Since the bank has many ATMs all over the country, choosing good location where is no ATM or distant from another ATM or branch is not easy

Therefore, the goal of this exercise is to give a recommendation to the bank of the best location of putting an ATM.

Description of the data

In the execution of this project, I will use the following:

I will collect the offsite ATMs and branch data concerning their location and address from the bank website https://www.centenarybank.co.ug/index.php/branches/index. ATM and Branch data will be merged into one data frame because where there is a branch there is an ATM.

I will also extract data concerning of the Top most populated towns in Uganda from Wikipedia: https://en.wikipedia.org/wiki/List_of_cities_and_towns_in_Uganda.

I will use Forsqure API to get the most common venues of different towns in Uganda.

Methodology

In this section, I will describe the data analysis and how I used the data to yield the results.

Starting out, I scraped data from Centenary Bank website to create a dataframe with the offsite and branch ATMs of the bank: https://www.centenarybank.co.ug/index.php/branches/index

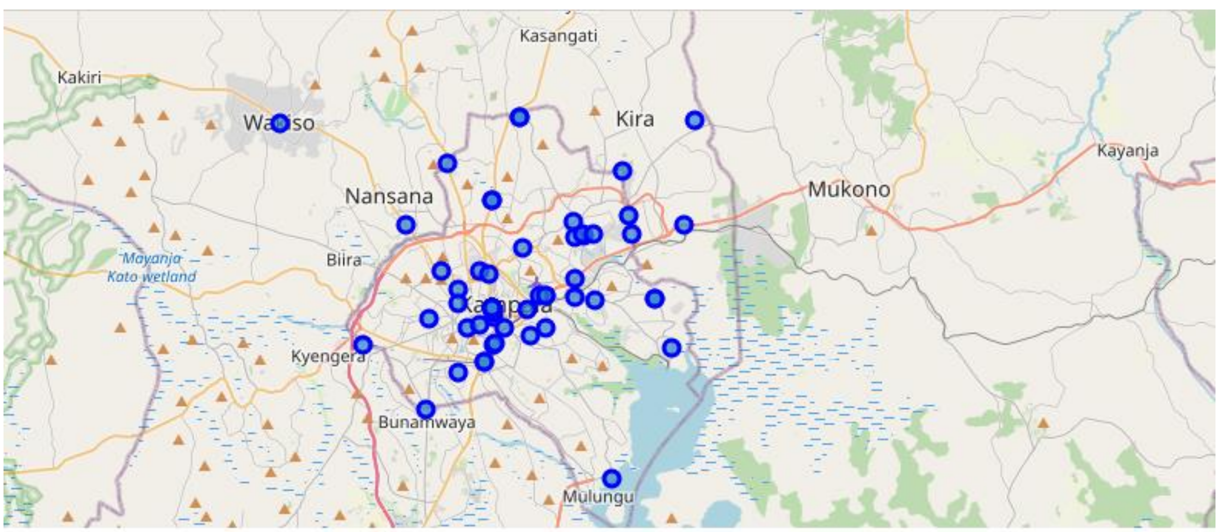
For this, I used the pandas read function. I had to clean the resulting data frame in terms of unnecessary information or data that could not be handled in a data frame, such as ATM location and Neighborhood(Address). The result is a nice data frame:

	Location	Neighborhood
0	Arua	Catholic Center Building, Near Christ the King...
1	Bugolobi	Spring Road, Middle East Hospital & shopping c...
2	Bugembe	Bugembe Town Council, Jinja-Iganga highway.
3	Busia	Busia Customs Road, Busia town
4	Iganga	Main Street, Iganga town

Then, I enabled geopy functions by installing the conda-forge geopy package. I used the nominatim function to add geospatial data to the data frame, that is the latitude and the longitude seen on the right side of the following table.

	Location	Neighborhood	Latitude	Longitude
0	Arua	Catholic Center Building, Near Christ the King...	0.315820	32.575030
1	Bugolobi	Spring Road, Middle East Hospital & shopping c...	0.319457	32.620888
2	Bugembe	Bugembe Town Council, Jinja-Iganga highway.	0.470120	33.248820
3	Busia	Busia Customs Road, Busia town	0.463070	34.105660
4	Iganga	Main Street, Iganga town	0.315820	32.575030

I used python **folium** library to visualize geographic details of Kampala with ATM locations superimposed on top. I used latitude and longitude values to get the visual as below:



Now, foursquare data comes into play. I first did a view try-outs for the city district "Kampala", which is the capital city of Uganda and in central region, to see if the venues retrieved from foursquare seem reasonable and correct. That was the case.

Then, retrieved the foursquare data for all venues on foursquare with a distance of less than 500 meters from each ATM, as indicated as blue dots in the map above. The result was a list of 800 venues all over Uganda. Out of these 800 venues, 94 where unique categories among which were restaurants, coffee shops, gyms and ATMs.

	Neighborhood	ATM	Afghan Restaurant	African Restaurant	Airport	Art Gallery	Arts & Crafts Store	Auto Garage	BBQ Joint	Bakery	Bar	Basketball Court	Bed & Breakfast	Beer Garden	Bistro	Boarding House
0	Catholic Center Building, Near Christ the King...	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Catholic Center Building, Near Christ the King...	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Catholic Center Building, Near Christ the King...	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0

I used this information to create a data frame in which you can see the most common venue types for each neighborhood.

	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	Andrea Olal Road Opposite Shell petrol Station	Coffee Shop	Shopping Mall	Café	Bar	Italian Restaurant	Fast Food Restaurant	Park	Art Gallery	Grocery Store	Diner
1	Angwenchibange Parish Akaidebe Zone, Parcel Do...	Coffee Shop	Shopping Mall	Café	Bar	Italian Restaurant	Fast Food Restaurant	Park	Art Gallery	Grocery Store	Diner
2	Assessment Centre, Mulago Hospital	Café	Park	Hotel Bar	Nightclub	Lounge	Chinese Restaurant	Indian Restaurant	Flea Market	Gym	Garden
3	Ben Kiwanuka Street	Park	IT Services	Coffee Shop	Mobile Phone Shop	Burger Joint	Convenience Store	Ice Cream Shop	Fast Food Restaurant	Hotel	Video Store
4	Block 123, Plot 300\nMain Street, Kayunga Cent...	Post Office	Whisky Bar	Café	Chinese Restaurant	Clothing Store	Cocktail Bar	Coffee Shop	Construction & Landscaping	Convenience Store	Department Store

.

Now, with all this data, I could finally run an unsupervised machine learning algorithm, more specifically, a k-means clustering algorithm from the scikit-learn package. One could use the ellbow method to systematically define the k value, but I simply chose k to be 5.

To find clusters of different neighboring venues in the different Location, I first transformed the data frame with the ATM locations and venues, associated to location as seen in the picture below

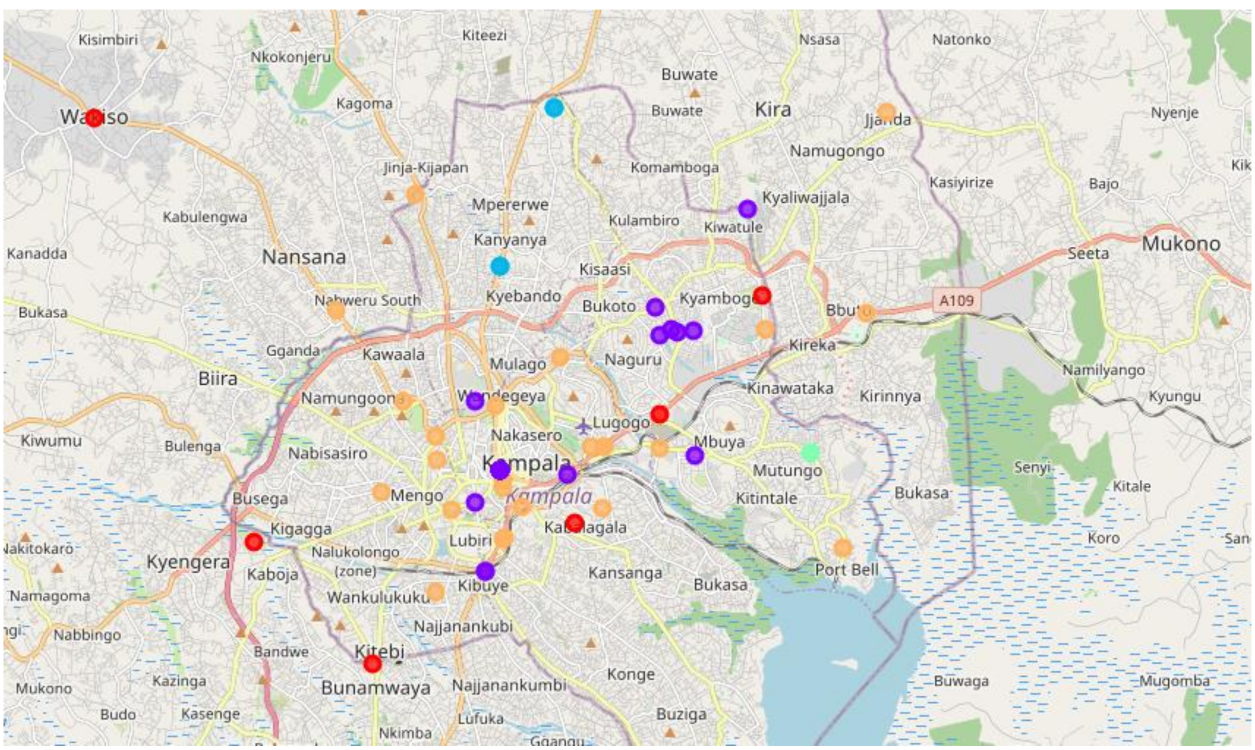
Results

And here already comes the result:

	Location	Neighborhood	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
0	Arua	Catholic Center Building, Near Christ the King...	0.315820	32.575030	1.0	Coffee Shop	Shopping Mall	Café	Bar	Italian Restaurant	Fast Food Restaurant	Park
1	Bugolobi	Spring Road, Middle East Hospital & shopping c...	0.319457	32.620888	1.0	Fast Food Restaurant	Bar	Shopping Mall	Whisky Bar	Bistro	Coffee Shop	Café
2	Bugembe	Bugembe Town Council, Jinja-Iganga highway.	0.470120	33.248820	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	Busia	Busia Customs Road, Busia town	0.463070	34.105660	4.0	Flea Market	Bus Station	Hotel	IT Services	Diner	Fast Food Restaurant	Cocktail Bar
4	Iganga	Main Street, Iganga town	0.315820	32.575030	1.0	Coffee Shop	Shopping Mall	Café	Bar	Italian Restaurant	Fast Food Restaurant	Park

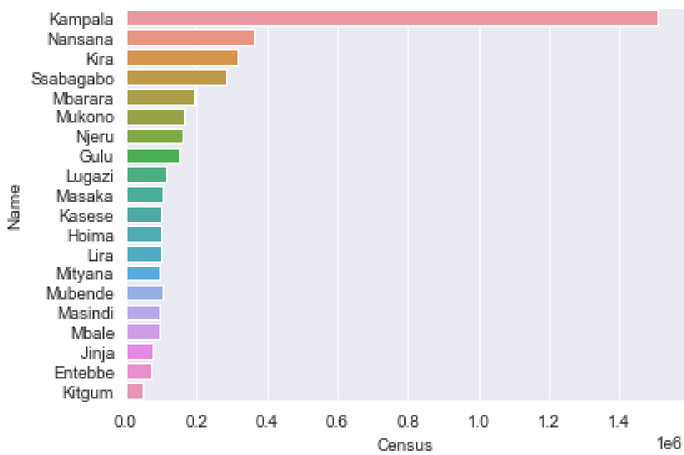
What we see in the table are the locations and their most common venues, and they now have been assigned five different cluster labels from 0 to 4.

We can now use the cluster labels to show the location marked with a cluster-specific color on a map, again using folium:



Since customer numbers are important, I decided to extract the data of the most populated towns in Uganda and plotted on a seaborn bar char.

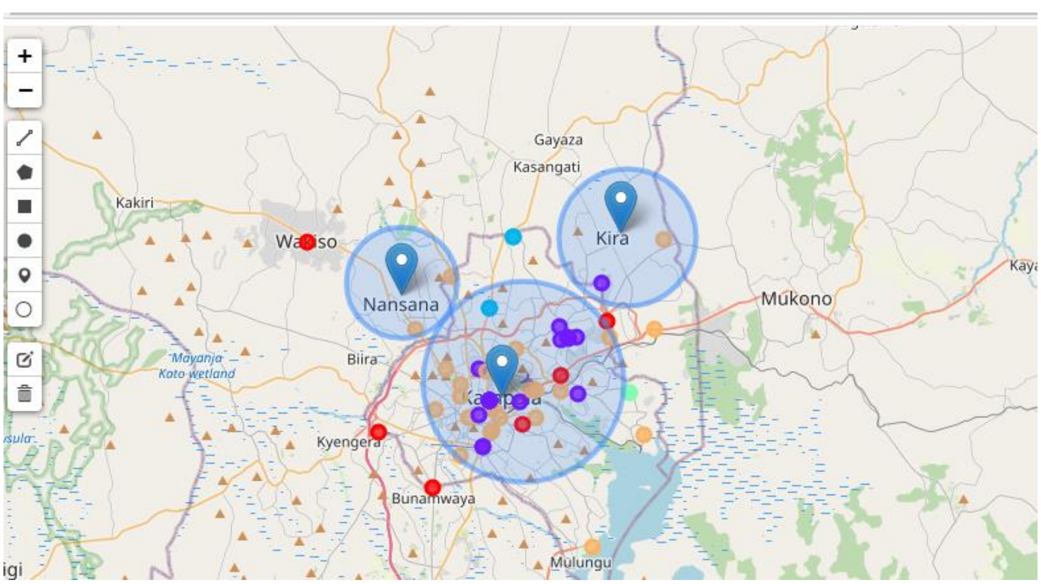
	Rank	Name	Census
0	1	Kampala	1,507,080
1	2	Nansana	365,124
2	3	Kira	317,157
3	4	Ssabagabo	283,272
4	5	Mbarara	195,531



With the help of the Foursquare we also got their latitudes and Longitudes.

	Rank	Name	Census	Latitude	Longitude
0	1	Kampala	1507080	0.315820	32.575030
1	2	Nansana	365124	0.366980	32.528900
2	3	Kira	317157	0.337188	32.580896
3	4	Ssabagabo	283272	0.289466	32.604420
4	5	Mbarara	195531	-0.616670	30.650000

Finally we plot the Location of the most populated towns on the map of our clusters to visualize how the clusters are distributed in the most populated towns.



Discussion

Data Science is an interesting field, will all free open source softwares that require less computer performance one can come up with such great import business projection. We just have to get to know the available open source packages and learn how to use them.

However I noticed that some of the coordinates provided by geopandas were not all acculate. Some ATMs were allocated out of Uganda as far as India but majority were in range.

In future if more data concerning customers is availed this can give us bigger picture and wide view of the analysis

Conclusion

In conclusion from my project, Kampala City/Town has the biggest population and more ATM distribution. The neighboring towns like Kira and Nansana have few ATMs but with promising growth which can be a suitable of ATM location considering the population they have.

I believe the Bank can use this for future prospection of business.