

Capstone Project Report

The Battle of Neighborhoods

Finding the best cities in Botswana, Africa
based on Foursquare registered data



Map of Botswana, Africa

Introduction

In this project we will try to find the **most livable cities** in Botswana, Gaborone. Specifically this report will be targeted to individuals/ families interested in moving or buying a house in Botswana.

After having lived in Botswana for over a decade, with experience of having lived in different cities, I am often asked by new people moving to Botswana on which place is best suited to live! I hope to justify my answer by creating a comparative analysis between cities in Gaborone.

People prefer places that are in close proximity to schools, shopping malls, hospitals, grocery shops, theatre etc. Apart from proximity to a facility, the mean house price, school ratings, crime rates, road connectivity, weather conditions, water resources and recreation facilities in the city matter as well.

Botswana is a land locked country in Southern Africa. Botswana is topographically flat, with up to 70 percent of its territory being the Kalahari Desert. With just over 2 million people, it is one of the most sparsely populated countries in the world. From being one of the world's poorest countries, it has since transformed itself into an upper middle income country, with one of the world's fastest-growing economies.

We will use our data science powers to generate name of promising cities to live in Botswana, based on various conditions. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

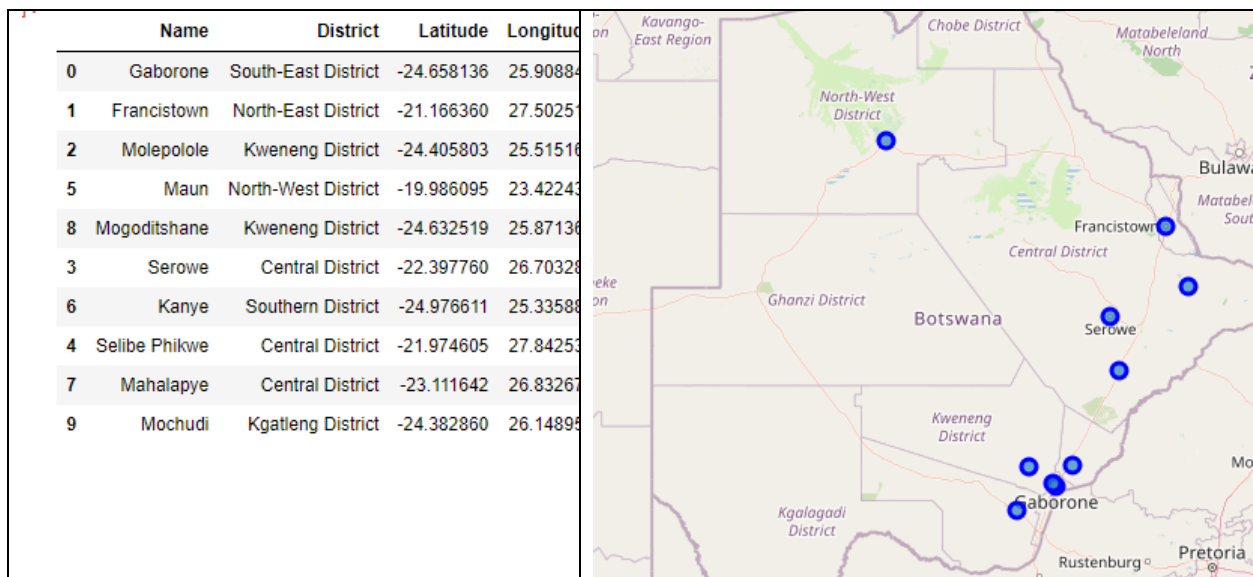
DATA

Based on the definition of our problem, factors that will influence our decision are:

- List of cities in Botswana from **Wikipedia** :
https://en.wikipedia.org/wiki/List_of_cities_in_Botswana
- Get coordinates for each City by running a geocoding web service (**Nominatim**, which is based on Open Street Map)
- Information about various venues in the city and their ratings using **Foursquare API**:
<https://foursquare.com/>

By combining the above data, further analysis will be performed. Correlations between density and all venue categories will be calculated. Charts and Cluster Algorithms will help to find solution to the Business Problem stated to 'Find the best livable cities in Botswana, Africa'.

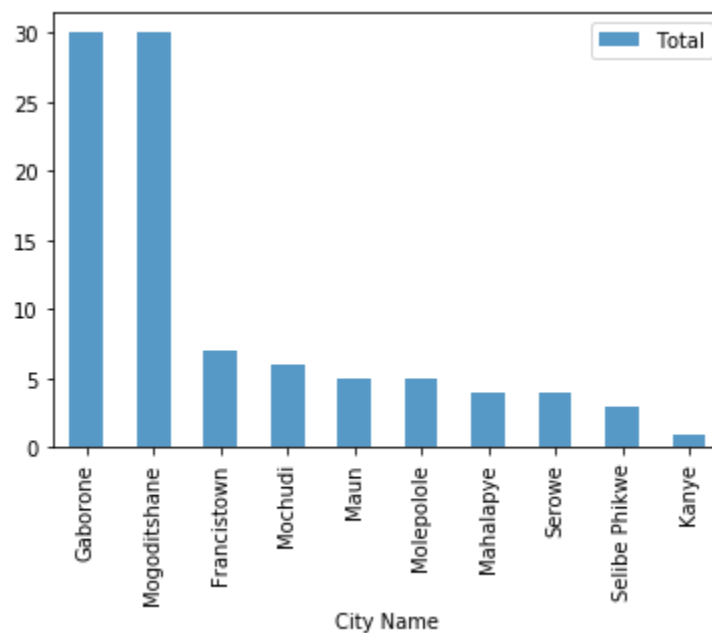
Our Analysis can only be as good as the data! Botswana is still far behind from other developed countries with tech-savvy people. I have noticed that there are very few businesses registered/ rated on Foursquare compared to the many operating in Botswana.



METHODOLOGY

Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos.

After finding the list of cities, we extract the top 10 cities by population. We then use this list to connect to the Foursquare API to gather information about venues inside each and every city within 20km radius.



Bar Chart displaying the No. of nearby Foursquare Venues Vs. City

Clustering Approach:

To compare the similarities of two cities, we decided to explore the venues, segment them, and group them into clusters to find similar cities in Botswana. To be able to do that, we need to cluster data which is a form of unsupervised machine learning: **k-means clustering algorithm**.

City Name	Total	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
Gaborone	30	Hotel	Café	Shopping Mall	Restaurant	Coffee Shop	Steakhouse	Cocktail Bar	Convenience Store	Brazilian Restaurant	Bistro
Mogoditshane	30	Hotel	Shopping Mall	Café	Coffee Shop	Steakhouse	Food	Arts & Crafts Store	Asian Restaurant	Bistro	Brazilian Restaurant
Francistown	7	Shopping Mall	Airport	Pub	Hotel	Portuguese Restaurant	Bed & Breakfast	None	None	None	None
Mochudi	6	Convenience Store	Farm	Bar	Grocery Store	Hotel	Burmese Restaurant	None	None	None	None
Maun	5	Shopping Mall	Food	Portuguese Restaurant	Hotel	National Park	None	None	None	None	None

Various Common Venue Categories segmented and groped under the stated City

Step 7 - Cluster Cities using k-means Algorithm

Run *k*-means to cluster the neighborhood into 5 clusters.

```
In [49]: # set number of clusters
kclusters = 5

city_grouped_clustering = city_grouped_4_cluster.drop('City Name', 1)

#city_grouped_clustering

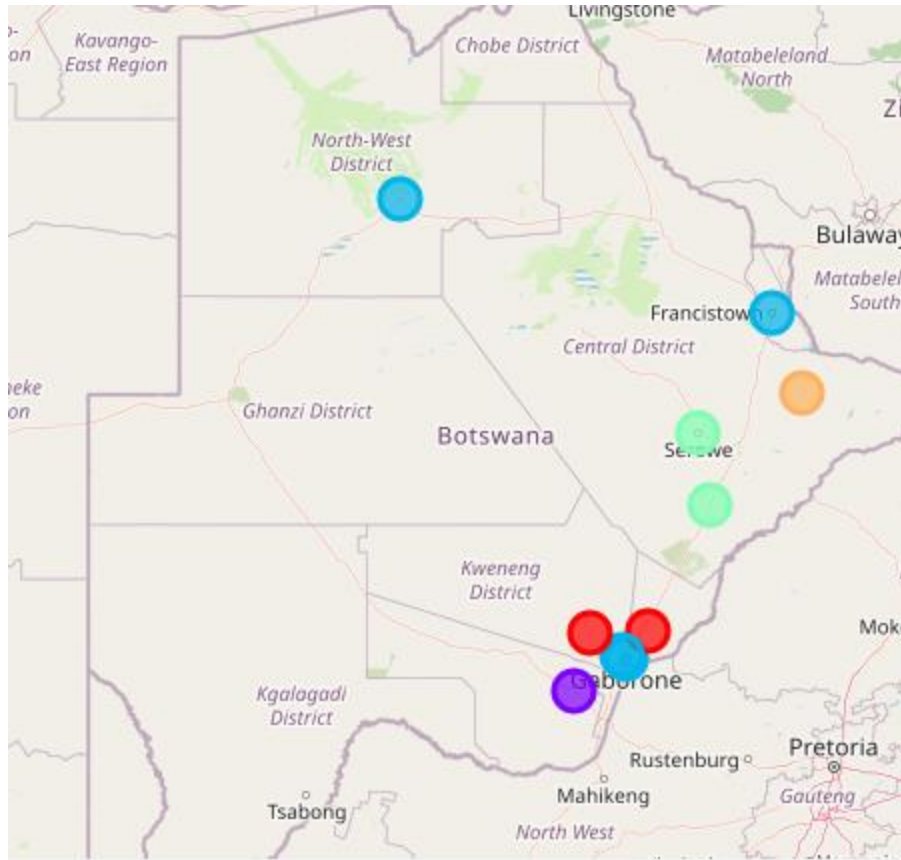
# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(city_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_

Out[49]: array([2, 2, 1, 3, 2, 0, 2, 0, 4, 3], dtype=int32)
```

Clustering Algorithm applied on the City, Venue details

RESULTS



Map of Clusters of top 10 Botswana Cities

The Clustering Algorithm divided the 10 Botswana Cities into 5 different Clusters based on similar venues.

Cluster 1

```
city_merged.loc[city_merged['Cluster Labels'] == 0, city_merged.columns[[1] + list(range(2, city_merged.shape[1]))]]
```

4]:

	City Name	Total	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	District	Latitude	Longitude
5	Mochudi	6	Convenience Store	Farm	Bar	Grocery Store	Hotel	Burmese Restaurant	None	None	None	None	Kgalega District	-24.382890	26.148950
7	Molepolole	5	Convenience Store	Stadium	Shopping Mall	Fast Food Restaurant	None	None	None	None	None	None	Kweneng District	-24.405803	25.515165

Cluster 2

```
city_merged.loc[city_merged['Cluster Labels'] == 1, city_merged.columns[[1] + list(range(2, city_merged.shape[1]))]]
```

5]:

	City Name	Total	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	District	Latitude	Longitude
2	Kanye	1	Shopping Mall	None	None	None	None	None	None	None	None	None	Southern District	-24.979611	25.335888

Cluster 3														
city_merged.loc[city_merged['Cluster Labels'] == 2, city_merged.columns[[1] + list(range(2, city_merged.shape[1]))]]														
56):	City Name	Total	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	District	Latitude Longitude
0	Francistown	7	Shopping Mall	Airport	Pub	Hotel	Portuguese Restaurant	Bed & Breakfast	None	None	None	None	North-East District	-21.166360 27.502515
1	Gaborone	30	Hotel	Café	Shopping Mall	Restaurant	Coffee Shop	Steakhouse	Cocktail Bar	Convenience Store	Brazilian Restaurant	Bistro	South-East District	-24.658136 25.908847
4	Maun	5	Shopping Mall	Food	Portuguese Restaurant	Hotel	National Park	None	None	None	None	None	North-West District	-19.986095 23.422435
6	Mogoditshane	30	Hotel	Shopping Mall	Café	Coffee Shop	Steakhouse	Food	Arts & Crafts Store	Asian Restaurant	Bistro	Brazilian Restaurant	Kweneng District	-24.632519 25.871364
Cluster 4														
city_merged.loc[city_merged['Cluster Labels'] == 3, city_merged.columns[[1] + list(range(2, city_merged.shape[1]))]]														
56):	City Name	Total	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	District	Latitude Longitude
3	Mahalapye	4	Eastern European Restaurant	Boarding House	Shopping Mall	Bus Station	None	None	None	None	None	None	Central District	-23.11642 26.832672
9	Serowe	4	Bus Station	Shopping Mall	Restaurant	Resort	None	None	None	None	None	None	Central District	-22.397760 26.703289
Cluster 5														
city_merged.loc[city_merged['Cluster Labels'] == 4, city_merged.columns[[1] + list(range(2, city_merged.shape[1]))]]														
56):	City Name	Total	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	District	Latitude Longitude
8	Selibe	3	Sports Club	BBQ Joint	Hotel	None	None	None	None	None	None	None	Central District	-21.974805 27.842536

DISCUSSION

We notice that there are many cities with very few nearby Venues as per Foursquare. But as a fact, that is not actually true. There are many businesses operating in these cities which have not been registered and rated on Foursquare by users. This supports the fact that Botswana users/ tourists are not tech-savvy yet. Thus the use of Foursquare to find the best livable cities in Botswana is not appropriate.

CONCLUSION

In this Capstone project, using k-means cluster algorithm we separated the 10 Botswana Cities into 5 (Five) different clusters, which have very-similar neighborhoods around them. Being someone who has lived in this country for over a decade, I confirm that project did cluster similar cities, though with limited data available in Foursquare.

Cluster 3 comprising of Gaborone, Francistown, Mogoditshane and Maun are the best cities with majority of venues tagged in the 20 km radius

Cluster 3														
city_merged.loc[city_merged['Cluster Labels'] == 2, city_merged.columns[[1] + list(range(2, city_merged.shape[1]))]]														
56):	City Name	Total	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	District	Latitude Longitude
0	Francistown	7	Shopping Mall	Airport	Pub	Hotel	Portuguese Restaurant	Bed & Breakfast	None	None	None	None	North-East District	-21.166360 27.502515
1	Gaborone	30	Hotel	Café	Shopping Mall	Restaurant	Coffee Shop	Steakhouse	Cocktail Bar	Convenience Store	Brazilian Restaurant	Bistro	South-East District	-24.658136 25.908847
4	Maun	5	Shopping Mall	Food	Portuguese Restaurant	Hotel	National Park	None	None	None	None	None	North-West District	-19.986095 23.422435
6	Mogoditshane	30	Hotel	Shopping Mall	Café	Coffee Shop	Steakhouse	Food	Arts & Crafts Store	Asian Restaurant	Bistro	Brazilian Restaurant	Kweneng District	-24.632519 25.871364