The Battle of the Neighborhoods

Starbucks Coffee shops in Tunisia

Ashraf Ben Hammadi

December 2nd, 2020

A. Introduction

A.1. Background & description of the business problem

Tunisia is home of thousands of cafés. Well, a cafe is never more than a few feet away; and you are never more than a few hours from your next shot of Espresso or Cappuccino. Now adding to it a Starbucks, the long-awaited US coffee giant that has decided last year to finally open several branches in the country with an ambitious plan to conquer the spiritual home of espresso. This is a good decision as the Tunisian market should already have room for both traditional cafés and Starbucks, as they target different consumers. Unfortunately, for unknown reasons, Starbucks still did not open its doors in Tunisia.

The idea of this project is to imagine as if the decision of Starbucks business development is to go ahead an open several stores in Grand Tunis. [1]

One of the first challenges that could face Starbucks is to choose the coffee shops location. The location of a coffee shop, or just about any type of business for that matter, plays a huge role in the success of the shop. Choosing the right location is a key to any good business endeavor. Hence, the concept of this project is to put myself in Starbucks team 's shoes and try to find out for them the best locations.

A.2. Problem solving concepts

In order to find the best locations for the coffee shop stores, here is what we need to look for:

- **Demographics – the who**

We need to look into "who" is in the area I'm considering. There are certain types of people that likely will be looking for coffee more than others. So, because of this, we should look for locations that has higher populations of the kind of people we're talking about. Examples would be college students, shoppers, and lots of workforces.

- **Competition – the what**

Starbucks is one of the big names in the coffee industry. We all know who they are. However, Tunisian market is a bit different, Tunisia is still getting away from" the big chain restaurant and stores" and the coffee shops are locally owned. Therefore, the direction is to find locations that doesn't already have a bunch of locally owned coffee shops.

- **Neighboring businesses – the where**

We should investigate locations with businesses that if you're close to, it can actually help drive the sales up more. These businesses are in locations where you do want to be operating. Examples would be boutique's stores, bookstores/libraries and family fun venues.

- **Cost of living – the how (much)**

When we're choosing the locations, we have to think about the cost of living within each area. Some areas will be more expensive than others. Metropolitan and highly urban areas typically have higher costs of living. Although this means a higher cost for entry into the market, higher operating costs, and higher rent or mortgage costs, but in the other side there is major potential for high profits. So, the direction is to evaluate the cost living index of all neighborhoods of "Grand Tunis" and then to go for the areas with higher cost living.

When we consider studying all the above factors, we can think to leverage the Foursquare location data to explore and cluster the neighborhoods of Grand Tunis and then evaluate the result based on the above selective factors in order to answer the question : where would be recommended for Starbucks to open the coffee shop stores?

A.3. Data sources

The list of data set and sources used for the purpose of this project is listed as below:

1. The Second-level Administrative Divisions, Tunisia, 2015 – Published by the University of California, Berkeley. This file contains the administrative divisions of Tunisia which include all delegations and the geospatial data. The .json file has coordinates of all cities of Tunisia. I have converted it to csv file for easy processing in my Notebook. [2]

2. I used Foursquare API to get the most common venues of the given Borough of Grand Tunis.[3]

3. Obviously, the first three factors from section A.2 (demographics, Competition and Neighboring businesses) could be generated and assessed through the result given by the leverage of the Foursquare location data to explore or compare the neighborhoods of "Grand Tunis". What is pending is to link the result with the cost living factor. For this, we are going to use the data base of the INS [4] which contains the latest official survey of the population in Tunisia.

B. Methodology

As a data set, I have used the Second-level Administrative Divisions, Tunisia which was Published by the University of California Berkeley. The .json file has coordinates of all cities of Tunisia and I have converted it to csv file for easy processing in my Notebook.

My data set has all the main required features like Borough, Neighborhood, Latitude and Longitude information of Tunisia. I have cleaned the data and reduced it to the region of "Grand Tunis ": Grand Tunis is the name for the greatest metropolitan area in Tunisia, which assembles four of the following states: Tunis, Ariana, Manubah and Ben Arous. According to the latest population census in 2004, the area of Grand Tunis

is the home of 2,247,800 peoples. Therefore, the aim is to find the most convenient neighborhoods in Grand Tunis in order to open the first batch of Starbucks ‘s licensed stores.

	features__id	Borough	Neighborhood	Latitude	Longitude
0	zy373kj6079.1	Ariana	Ariana Médina	36.856560	10.193133
1	zy373kj6079.2	Ariana	Ettadhamen	36.842808	10.056228
2	zy373kj6079.3	Ariana	Kalaat El Andalous	37.151249	10.218749
3	zy373kj6079.4	Ariana	Mnihla	36.863815	10.131537
4	zy373kj6079.5	Ariana	Raoued	36.955280	10.230476

I have used python folium library to visualize geographic details of Grand Tunis and its boroughs. the center of map was set on through Nominatim search engine for OpenStreetMap data then I have used latitude and longitude values to get the visual as below:

Next step is to utilize the Foursquare API to explore the boroughs and segment them. The number of venues is limited to 100 venues and the radius to 500 meters for each borough from their given latitude and longitude information. Here is a head of the list Venues name, category, latitude and longitude information:

	name	categories	lat	lng
0	صحن تونسي عند لطفي	African Restaurant	36.859091	10.188879
1	Woodstock	Fast Food Restaurant	36.855739	10.187889
2	El Borj	Fast Food Restaurant	36.855375	10.195601
3	Baguette & Baguette Ariana	Burger Joint	36.857064	10.188687
4	La Friandise	Dessert Shop	36.858211	10.191946
5	Cafés BONDIN	Coffee Shop	36.858942	10.191740
6	Yasmine Coffee	Café	36.856276	10.194643
7	Kafteji Ellas	Fast Food Restaurant	36.858585	10.189685
8	cafe Belar	Café	36.857106	10.188519
9	Chill-out café	Café	36.857249	10.194428
10	Rendez-vous	Tea Room	36.856669	10.193113

The above was generated to all neighborhoods of Grand Tunis and the new data frame was as per the below shape:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Ariana Médina	36.85656	10.193133	صحن تونسي عند لطفي	36.859091	10.188879	African Restaurant
1	Ariana Médina	36.85656	10.193133	Woodstock	36.855739	10.187889	Fast Food Restaurant
2	Ariana Médina	36.85656	10.193133	El Borj	36.855375	10.195601	Fast Food Restaurant
3	Ariana Médina	36.85656	10.193133	Baguette & Baguette Ariana	36.857064	10.188687	Burger Joint
4	Ariana Médina	36.85656	10.193133	La Friandise	36.858211	10.191946	Dessert Shop

We can further analyze each neighborhood by providing how many venues were returned for each neighborhood.

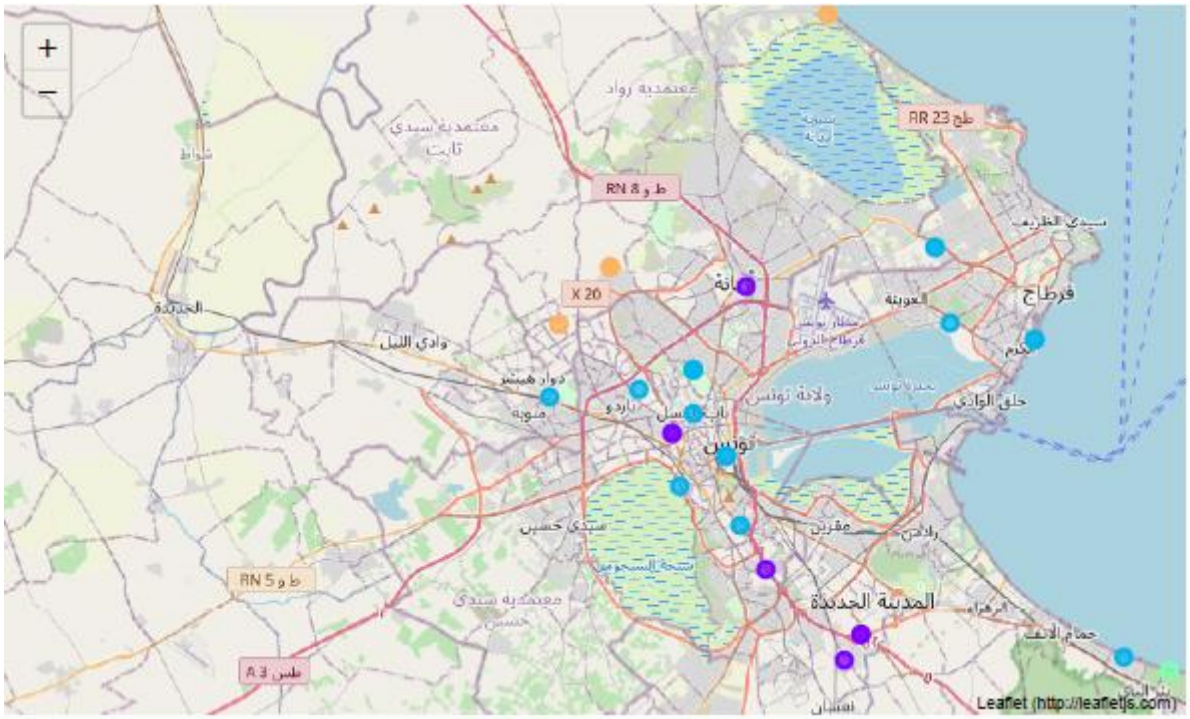
Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Ariana Médina	25	25	25	25	25	25
Bab Bhar	21	21	21	21	21	21
Bab Souika	15	15	15	15	15	15
Bardo	2	2	2	2	2	2
Bouhaira	13	13	13	13	13	13
Carthage	7	7	7	7	7	7
Cité El Khadra	3	3	3	3	3	3
El Menzah	3	3	3	3	3	3
El Mourouj	4	4	4	4	4	4
El Ouardia	3	3	3	3	3	3
El Tahrir	2	2	2	2	2	2

In summary, there is 58 unique venue categories were returned by Foursquare, then I created a table which shows list of top 10 venue category for each borough as shown in the below table:

	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	10t Co
0	Ariana Médina	Café	Fast Food Restaurant	Coffee Shop	Food & Drink Shop	Burger Joint	Farmers Market	Dessert Shop	Metro Station	Pharmacy	Res
1	Bab Bhar	Electronics Store	Theater	Movie Theater	Bus Station	Dessert Shop	Garden	Gastropub	Mac & Cheese Joint	Café	C
2	Bab Souika	Plaza	Café	Metro Station	Market	Restaurant	Food Truck	Bakery	Bookstore	Diner	[
3	Bardo	Dessert Shop	Video Game Store	Café	Food & Drink Shop	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Electronics Store	Diner	Dep
4	Bouhaira	Café	Diner	Coffee Shop	Pizza Place	Cupcake Shop	Restaurant	Fast Food Restaurant	Snack Place	General Entertainment	Res

We have some common venue categories in the boroughs. In this reason I have used unsupervised learning K-means algorithm to cluster the boroughs. We are going to cluster the boroughs into 5 clusters (5-degree optimum k of the K-Means). Here is the new data frame with clusters and the correspondent map showing the clusters distribution on Grand Tunis map.

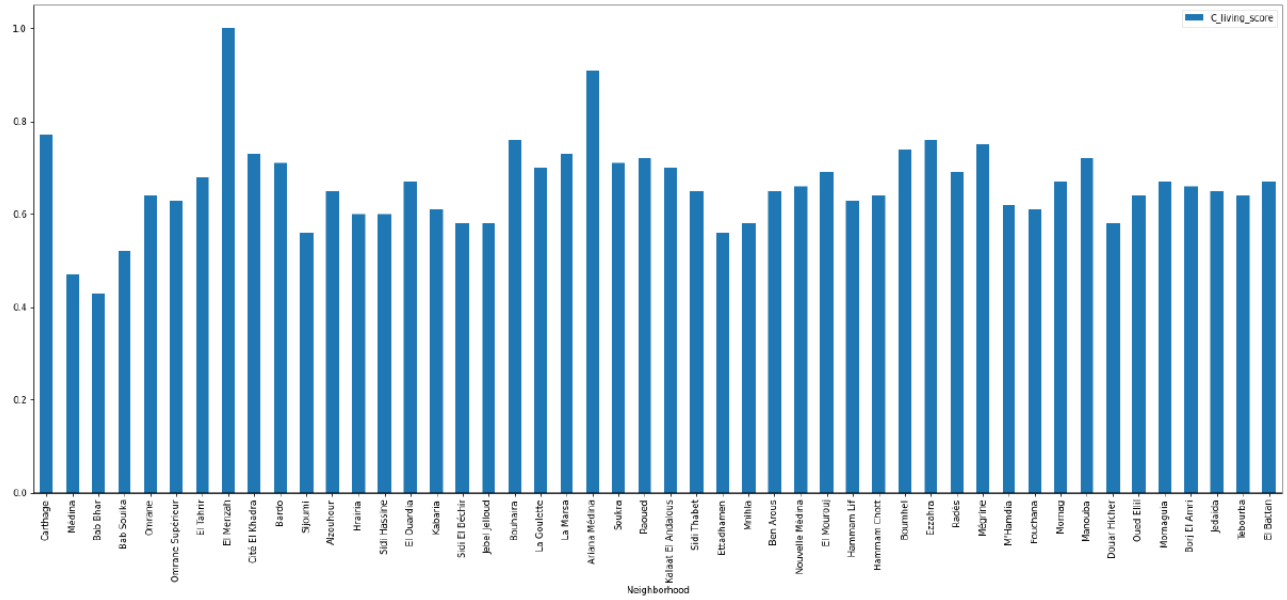
	features__id	Borough	Neighborhood	Latitude	Longitude	Cluster label	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	zy373kj6079.1	Ariana	Ariana Médina	36.856560	10.193133	1	Café	Fast Food Restaurant	Coffee Shop	Food & Drink Shop	Burger Joint	Farmers Market	Dessert Shop	Metro Station	Pharmacy	African Restaurant
1	zy373kj6079.4	Ariana	Mnihla	36.863815	10.131537	4	Coffee Shop	Bakery	Pizza Place	Video Game Store	Café	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Electronics Store	Diner
2	zy373kj6079.5	Ariana	Raoued	36.955280	10.230476	4	Coffee Shop	Video Game Store	Café	Food & Drink Shop	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Electronics Store	Diner	Dessert Shop
3	zy373kj6079.8	Ariana	Soukra	36.870617	10.279025	2	Furniture / Home Store	Health & Beauty Service	Outdoor Supply Store	Food Truck	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Electronics Store	Diner	Dessert Shop
4	zy373kj6079.23	Ben Arous	Hammam Chott	36.717083	10.385272	3	Video Game Store	Café	Food & Drink Shop	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Electronics Store	Diner	Dessert Shop	Department Store



After having all the clusters and the common venues around the neighborhoods of all boroughs of Grand Tunis, we can say that we have covered the first three factors of our methodology which are demographics, Competition and Neighboring businesses.

For the remaining factor which is the cost living factor our data set was found in the data base of the INS [4]. I have used the “Census 2014 results per governorates “:it is the official survey of the population. The documents present statistical indicators from the general census of the population and the household 2014 by governorates. This was the latest census of population that we could found. The data available doesn’t give the required cost index factor but the survey provide several indicators related to the population per governorate like population density of the area, mode of housing (owning the property or renting) and the possession of some goods like cars ,electrical appliances …etc. I have used that information as weighing parameters to calculate an indicative factor, I called it C_living_score (cost living score) which give an idea about the areas that have higher costs of living and provide potential high profit. A head of the new generated data set that is going to be used in our notebook is displayed as follow (data frame and bar chart):

	Borough	Neighborhood	Neighborhoods Score	C_living_score	Rank
0	Tunis	Carthage	583.49	0.77	3
1	Tunis	Médina	353.92	0.47	47
2	Tunis	Bab Bhar	323.15	0.43	48
3	Tunis	Bab Souika	390.03	0.52	46
4	Tunis	Omrane	483.66	0.64	31



I have merged both data frames in order to have a compiled data frame which will be the basis of our decision making. I have modified it by dropping some columns and I kept only the ones that I need for the following section (Results).

C. Results

The below new data frame contains 22 neighborhoods distributed along the borough of Grand Tunis.

	Neighborhood	Cluster label	1st Most Common Venue	2nd Most Common Venue	C_living_score
0	Carthage	2	Park	Beach	0.77
1	Médina	1	Cafeteria	Café	0.47
2	Bab Bhar	2	Electronics Store	Theater	0.43
3	Bab Souika	2	Plaza	Café	0.52
4	Omrane	2	Soccer Stadium	Historic Site	0.64

This data frame will be our data base for the selection of the most suitable locations for Starbucks to set up their first stores in Grand Tunis. Our strategy is to find locations that doesn't already have a bunch of locally owned coffee shops and has higher cost living index. For that I have defined a new data frame which select locations based on two filtering conditions as follows:

1- Coffee shops not as 1st common venue.

2- Cost living index higher than 0.7.

This is the final data frame after conducting the conditional filtering:

	Neighborhood	Cluster label	1st Most Common Venue	2nd Most Common Venue	C_living_score
0	Carthage	2	Park	Beach	0.77
7	El Menzah	2	Soccer Stadium	Historic Site	1.00
8	Cité El Khadra	2	Soccer Stadium	Historic Site	0.73
9	Bardo	2	Dessert Shop	Video Game Store	0.71
16	La Goulette	2	Park	Beach	0.70
17	La Marsa	2	Furniture / Home Store	Health & Beauty Service	0.73
19	Soukra	2	Furniture / Home Store	Health & Beauty Service	0.71

D. Discussion

The result provides us with the most convenient neighborhoods (7 Neighborhoods) in Grand Tunis where Starbucks or the retails business partner could open the first batches of stores. Choosing among the above list will guarantee for the stores to be competitively positioned in high-demand areas.
Nevertheless, the result could be different if we change the strategy and the direction in which we want to take the business. For instance, we could go for the areas where the first common venue is coffeeshop, knowing that big names in the coffee industry account for about 70% of the profits made across the whole coffee industry, so competition won't be a factor that affect our choice of location but could be a great way to boost motivation and productivity.

E. Conclusion

The idea of this project was based on the following question: if Starbucks is looking to open stores in Grand Tunis, where would be recommended that they open it?
To solve the problem, I have been using location data to explore the geographical location of Grand Tunis and to compare the neighborhoods based on the Foursquare location data (the venues categories) in addition to other datasets that comprise social parameters.
Similar data analysis could be a good asset for investors and stakeholders in order to conclude optimal decisions.

References:

[1] Grand Tunis— Wikipedia

[2] The Second-level Administrative Divisions, Tunisia, 2015

[3] Foursquare API

[4] The National Institute of statistics (TUNISIA)