Ice Cream in Las Palmas de Gran Canaria

Introduction & Problem Description

Gran Canaria is an island on the Northwest coast of Africa, belonging to Spain. It is a popular touristic destination with over 3.5 million tourists yearly. [1]

Las Palmas de Gran Canaria is the capital city of the Island. The amount of ice-cream shops is still somewhat scarce in comparison to the number of tourists and the many outdoors amenities that are present.

Finding a good location for a new ice-cream shop is essential, maybe even more so than other types of stores. As Las Palmas de Gran Canaria has several touristic areas with each their differences it can be quite challenging to find the best spot.

Data

To be able to analyze this problem I made use of the following data:

1. The information about the different neighborhoods of Las Palmas de Gran Canaria were gathered from the Spanish Wikipedia page. I scraped the website using the Python BeautifulSoup library. [3]
2. I gathered the coordinates of Las Palmas de Gran Canaria and the different neighborhoods with the Python GeoPy library. [4]
3. With the Foursquare API I gather the venues in the different neighborhoods of Las Palmas de G.C. As Foursquare provides a lot of diversity in the venue categories, I brought them together in bigger logical groups. (e.g. several types of restaurants will be grouped under the common name "*restaurant*") [5]
4. For the mapping visualizations, OpenStreetMap has been used with the Python Folium library. [6]

Methodology

Venue Categories

The amount of venue categories on foursquare is very diverse and many of them will not have an impact on our choice. When searching through the neighborhoods of Las Palmas de Gran Canaria there are not less than 114 unique categories, many of them showing great similarities.

I first merged categories to logical groups in order to have less diversity in groups, resulting in the below series of groups.

Venue Category	
Restaurant	347
Shops	78
Outdoors	77
Drinks	61
Fastfood	47
Culture	38
Sport	36
Coffee/Tea Room	36
Hotel	27
Beach	18
Supermarket	16
Ice Cream Shop	15
Breakfast Spot	9
Nightclub	8
Farmers Market	7
Concert Hall	7
Multiplex	7
Lounge	6

Restaurant is a venue that is overly present. In all the neighborhoods it is the first or second most occurring venue. We can therefor consider it an outlier in our data. Furthermore, as a venue it has little to no impact on the decision. It is not because one is a customer of a restaurant that he is inclined or not to buy ice cream before or after his visit to the restaurant. We therefor remove the venue entries *Restaurant* from our dataset

Sport venues are also removed as it is not having a decisive impact in our choice.

Therefor the remaining top groups and their occurrence in Las Palmas de Gran Canaria are as follow:

Top Venues in Las Palmas de Gran Canarias
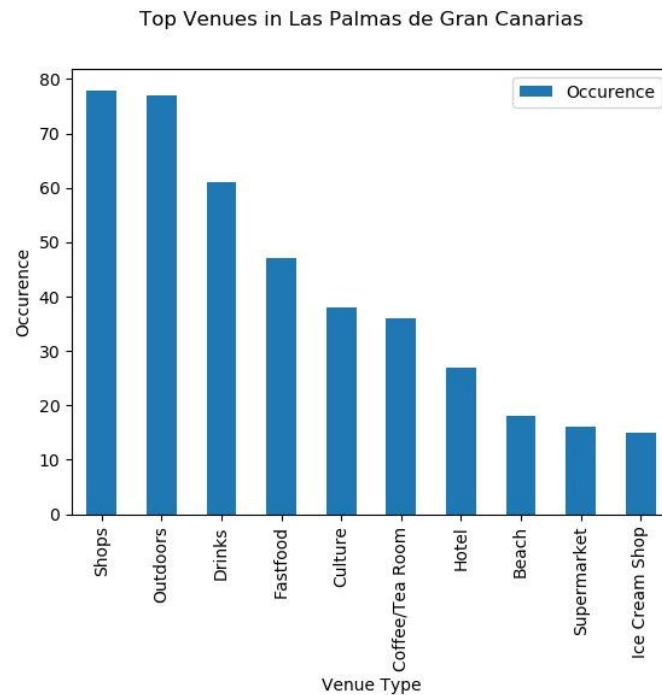
Neighborhood selection

I needed to select only the neighborhoods where there is still space for ice-cream shops. If a neighborhood has more than one ice-cream shop and less than four beaches, I considered that the marked in the neighborhood is saturated and removed it from further analysis.

I considered the number of beaches as they attract a lot of tourists and locals. If there are many beaches, there is more space for Ice Cream Shops.

In the below table you can see that the neighborhood *La Isleta* is not meeting our requirement, it will not be removed from the analysis.

Sector	Ice Cream Shop	Beach
Arenales	3	4
Guanarteme	3	4
Santa Catalina	3	4
La Isleta	2	3
Las Rehoyas	1	0
Las Torres	1	2
Schamann	1	0
Vegueta-Triana	1	0
Escaleritas	0	0
La Paterna	0	0
San Cristóbal	0	0
Tafira	0	0
Tamaraceite	0	0
Marzagán	0	1
San Lorenzo	0	0

Neighborhood clustering

I used k-means clustering on the remaining neighborhoods for classification.

As k-means requires the number of clusters to be defined beforehand, I used the elbow method to help determine the optimal number of clusters.
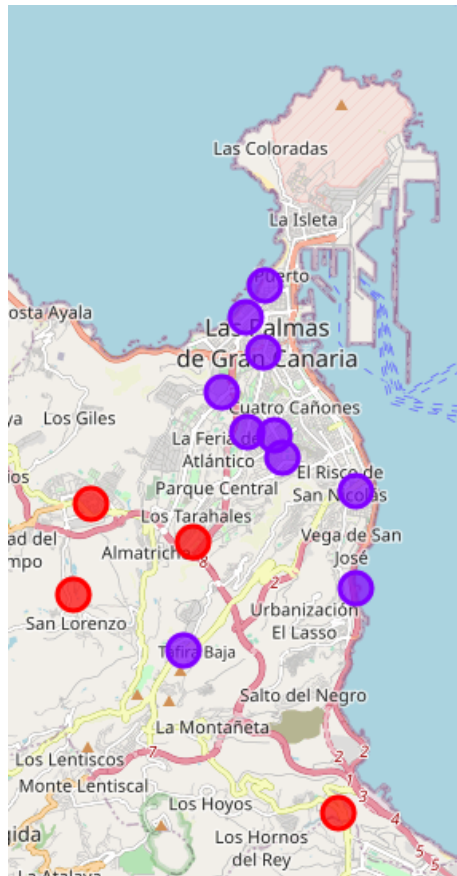
Elbow method

The Elbow Method showing the optimal k
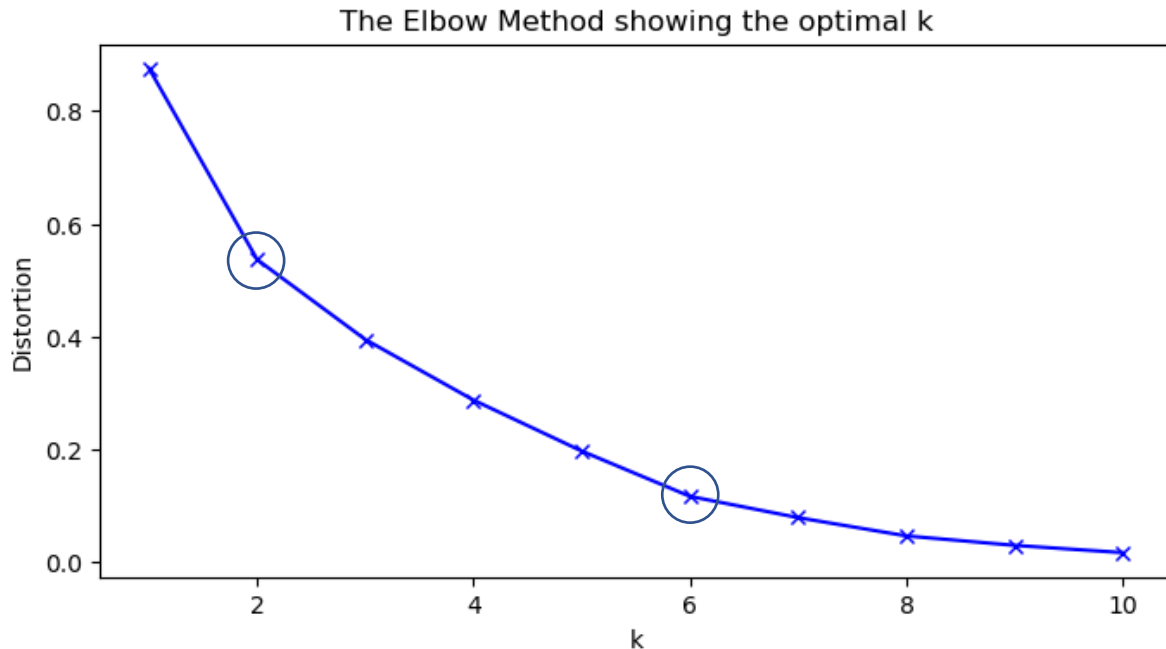
According to the elbow, the optimal number of clusters is two.

Although this is the optimal number of clusters, it showed only little information and did not help us much in choosing the appropriate neighborhood for the Ice Cream Shop.

Clustering with K=2

Cluster 1:

District	Sector	Latitude	Longitude	Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Ciudad Alta	La Paterna	28.091035	-15.450816	0	Shops	Fastfood	Outdoors	Coffee/Tea Room	Supermarket
Vegueta, Cono Sur y Tafira	Marzagán	28.039574	-15.419519	0	Shops	Fastfood	Gift Shop	Beach	Farmers Market
Tamaraceite-San Lorenzo	San Lorenzo	28.081055	-15.476745	0	Shops	Supermarket	Culture	Coffee/Tea Room	Bakery
Tamaraceite-San Lorenzo	Tamaraceite	28.098268	-15.472874	0	Shops	Coffee/Tea Room	Fastfood	Bus Station	Snack Place

Cluster 2:

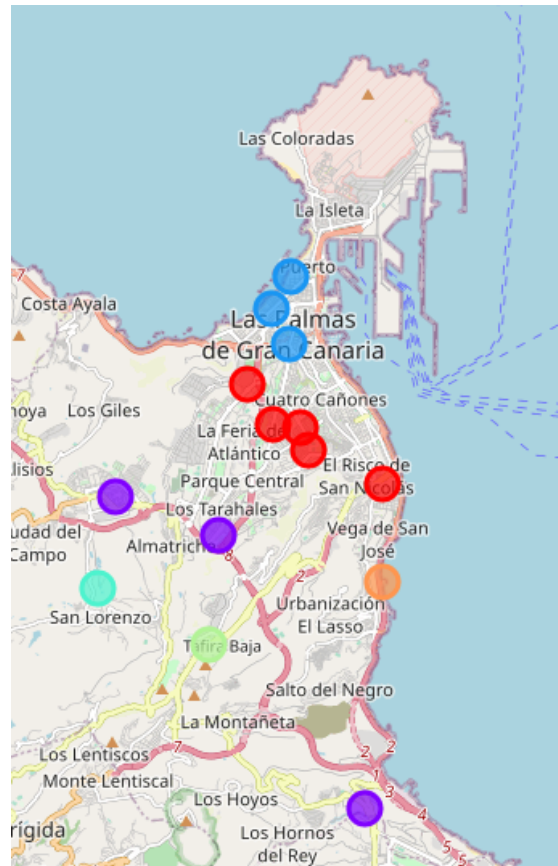| District	Sector	Latitude	Longitude	Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Centro	Arenales	28.127046	-15.435841	1	Drinks	Outdoors	Coffee/Tea Room	Hotel	Shops
Ciudad Alta	Escaleritas	28.112111	-15.439356	1	Shops	Outdoors	Drinks	Fastfood	Supermarket
Puerto-Canteras	Guanarteme	28.133622	-15.439566	1	Drinks	Outdoors	Coffee/Tea Room	Hotel	Culture
Ciudad Alta	Las Rehoyas	28.107005	-15.431840	1	Shops	Outdoors	Culture	Fastfood	Supermarket
Ciudad Alta	Las Torres	28.119548	-15.444735	1	Outdoors	Drinks	Shops	Fastfood	Supermarket
Vegueta, Cono Sur y Tafira	San Cristóbal	28.082157	-15.415947	1	Outdoors	Supermarket	Culture	College Auditorium	Shops
Puerto-Canteras	Santa Catalina	28.139699	-15.435295	1	Drinks	Outdoors	Coffee/Tea Room	Hotel	Culture
Ciudad Alta	Schamann	28.111440	-15.433229	1	Shops	Outdoors	Fastfood	Drinks	Supermarket
Vegueta, Cono Sur y Tafira	Tafira	28.070505	-15.452841	1	Shops	Snack Place	Pharmacy	Outdoors	Drinks
Vegueta, Cono Sur y Tafira	Vegueta-Triana	28.100697	-15.415786	1	Outdoors	Culture	Shops	Fastfood	Drinks

Clustering with K=6

As there is another small elbow at K = 6 I used amount for K to categorize the neighborhoods.

This gave better results on the clusters and made it easier to evaluate the different neighborhoods.
(see results section)

Results

After using K-means with K=6 we get the following clustered map of Las Palmas de Gran Canaria.

An observation that can be made directly is that most of the neighborhoods in the same clusters are geographically close to one another as well.

Cluster 1(Red)

District	Sector	Latitude	Longitude	Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Ciudad Alta	Escaleritas	28.112111	-15.439356	0	Shops	Outdoors	Drinks	Fastfood	Supermarket
Ciudad Alta	Las Rehoyas	28.107005	-15.431840	0	Shops	Outdoors	Culture	Fastfood	Supermarket
Ciudad Alta	Las Torres	28.119548	-15.444735	0	Outdoors	Drinks	Shops	Fastfood	Supermarket
Ciudad Alta	Schamann	28.111440	-15.433229	0	Shops	Outdoors	Fastfood	Drinks	Supermarket
Vegueta, Cono Sur y Tafira	Vegueta-Triana	28.100697	-15.415786	0	Outdoors	Culture	Shops	Fastfood	Drinks

Cluster 2 (Purple)

District	Sector	Latitude	Longitude	Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Ciudad Alta	La Paterna	28.091035	-15.450816	1	Shops	Fastfood	Outdoors	Coffee/Tea Room	Supermarket
Vegueta, Cono Sur y Tafira	Marzagán	28.039574	-15.419519	1	Shops	Fastfood	Gift Shop	Beach	Farmers Market
Tamaraceite-San Lorenzo	Tamaraceite	28.098268	-15.472874	1	Shops	Coffee/Tea Room	Fastfood	Bus Station	Snack Place

Cluster 3 (Blue)

District	Sector	Latitude	Longitude	Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Centro	Arenales	28.127046	-15.435841	2	Drinks	Outdoors	Coffee/Tea Room	Hotel	Shops
Puerto-Canteras	Guanarteme	28.133622	-15.439566	2	Drinks	Outdoors	Coffee/Tea Room	Hotel	Culture
Puerto-Canteras	Santa Catalina	28.139699	-15.435295	2	Drinks	Outdoors	Coffee/Tea Room	Hotel	Culture

Cluster 4 (Cyan)

District	Sector	Latitude	Longitude	Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Tamaraceite-San Lorenzo	San Lorenzo	28.081055	-15.476745	3	Shops	Supermarket	Culture	Coffee/Tea Room	Bakery

Cluster 5 (Green)

District	Sector	Latitude	Longitude	Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Vegueta, Cono Sur y Tafira	Tafira	28.070505	-15.452841	4	Shops	Snack Place	Pharmacy	Outdoors	Drinks

Cluster 6 (Orange)

District	Sector	Latitude	Longitude	Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Vegueta, Cono Sur y Tafira	San Cristóbal	28.082157	-15.415947	5	Outdoors	Supermarket	Culture	College Auditorium	Shops

Discussion

For clustering the neighborhoods of Las Palmas de Gran Canaria many different algorithms can be used with different results. I only used k-means algorithm here.

By using other algorithms as well and expanding the data with pedestrian traffic in outdoor areas and traffic in the different amenities, this would go more in detail.

If we do not limit ourselves to Las Palmas, other regions of the Island could be added as well, as Telde (an area on the east coast) and the south of the island (another big touristic place). This would not only give a good comparison of the different regions, but help us to make a better choice.

Conclusion

As expected, Cluster 3 (Blue) came out quite good. There are several Ice Cream Shops present already in Cluster 3 but being the most touristic area seems it still seems to be a good choice for a new Ice Cream Shop.

Another, more surprising result is *San Cristobal* in Cluster 6 (Orange). Although it is not the typical tourist area and has no access to a beach, it has a lot of amenities that are good for Ice Cream Shops (outdoor and cultural activities and student campus')

The results give good indications which clusters and neighborhoods might be interesting for a new Ice Cream Shop but is not conclusive and more data needs to be gathered to have a better overview of it.

References

[1] Gran Canaria – Wikipedia

[2] Las Palmas de Gran Canaria - Wikipedia

[3] BeautifulSoup Python library

[4] Geopy

[5] Foursquare API

[6] Folium Python library

[7] OpenStreetMap

Appendix 1

List of neighborhoods of Las Palmas de Gran Canaria and their coordinates.

District	Neighborhood	Latitude	Longitude
Centro	Arenales	28.127046	-15.435841
Ciudad Alta	Escaleritas	28.112111	-15.439356
Ciudad Alta	La Paterna	28.091035	-15.450816
Ciudad Alta	Las Rehoyas	28.107005	-15.431840
Ciudad Alta	Las Torres	28.119548	-15.444735
Ciudad Alta	Schamann	28.111440	-15.433229
Puerto-Canteras	Guanarteme	28.133622	-15.439566
Puerto-Canteras	La Isleta	28.151998	-15.426672
Puerto-Canteras	Santa Catalina	28.139699	-15.435295
Tamaraceite-San Lorenzo	San Lorenzo	28.081055	-15.476745
Tamaraceite-San Lorenzo	Tamaraceite	28.098268	-15.472874
Vegueta, Cono Sur y Tafira	Marzagán	28.039574	-15.419519
Vegueta, Cono Sur y Tafira	San Cristóbal	28.082157	-15.415947
Vegueta, Cono Sur y Tafira	Tafira	28.070505	-15.452841
Vegueta, Cono Sur y Tafira	Vegueta-Triana	28.100697	-15.415786