

The Battle of Neighborhoods Capstone Project

Matthieu Perrodin

EISTI, Cergy, France

Keywords: Candy Bar, Paris, Neighborhoods

1 Introduction and Business Problem

In 2001, Dylan Lauren pioneered the world's largest confectionery emporium and lifestyle brand, Dylan's Candy Bar. Its mission is to merge fashion, art and pop culture with candy to ignite the creative spirit and inner child in everyone that visits. This innovative concept has changed the way the world experiences candy today.

Dylan's Candy Bar houses over 7,000 confections, boasting an unparalleled selection of candies and candy-related gifts from around the world.



Figure 1. Dylan's Candy Bar 1011 Third Ave at East 60th Street New York, NY

The supposition is made that the Dylan's Candy Bar chain of stores wants to continue to expand internationally. For this reason, they want to open a store in Paris, emblematic city that would be a beautiful showcase for these fashion stores. Then the aim is to find the best location where this shop could be installed in Paris. For this, a study of the neighborhoods of Paris will be carried out. As these shops are very colorful like you can see in the photo, children are the target clientele by excellence. The study therefore concerns the ability to position the store according to the possible presence of children. Moreover,

several stores are already opened in the USA and could be an example of location where a fashion candy bar is likely to open.

Where is the best place for the candy bar to be seen by children ? Where is the best place to open a fashion candy bar like those already opened in the USA ? What neighborhoods have such criteria ?

2 Data

The first data used corresponds to the Paris neighborhoods. The data concerning Paris come from two different sources: on the one hand the wikipedia page of the list of administrative districts of Paris and on the other hand the site of opendata from the city of Paris. The opendata complete the information by giving the coordinates of the center of each neighborhood but also to obtain the geographical coordinates of these districts in a GeoJSON file:

1. [Wikipedia](#): Borough, Neighborhood, Number, Population and density for each neighborhood of Paris
2. [Opendata.Paris](#): Neighborhood, latitude, longitude and GeoJSON information about the neighborhoods.

Then data from Dylan's Candy Bar comes from the website of the chain.

- [Dylan's Candy Bar Website](#): list and of the stores and their address.

Finally, all the information about the neighborhoods in Paris and around the Dylan's bars comes from the Foursquare API. The request used is the search one.

3 Methodology

3.1 Paris Neighborhoods

The list of neighborhoods in Paris has been scrapped on the Wikipedia page cited earlier. The coordinates of each district were calculated from the Geocoder library. However, the coordinates returned were not precise and this is the reason why they were collected on opendata.paris. It is a csv file which was retrieved from this second site to have precise coordinates. The two files were then merged to obtain a dataframe containing all the information. In addition, the GeoJSON file was also recovered to obtain more precise visualizations of Paris with the Folium library.

	Borough(Arrondissement)	Neighborhood(Quarters)	Density	NoQuarters	Perimeter	Area	Latitude	Longitude
0	1er arrondissement	« du Louvre »	Saint-Germain-l'Auxerrois	1924.0	1	5057.549475	686900.6646	48.860650
1	1er arrondissement	« du Louvre »	Halles	21806.0	2	2066.411728	412458.4983	48.862289
2	1er arrondissement	« du Louvre »	Palais-Royal	11661.0	3	2168.839239	273696.7933	48.864660
3	1er arrondissement	« du Louvre »	Place-Vendôme	11316.0	4	2147.817602	269456.7806	48.867019
4	2e arrondissement	« de la Bourse »	Gaillon	7154.0	5	1866.962041	188012.2039	48.869307

Table 1. Dataframe merged of Paris neighborhoods

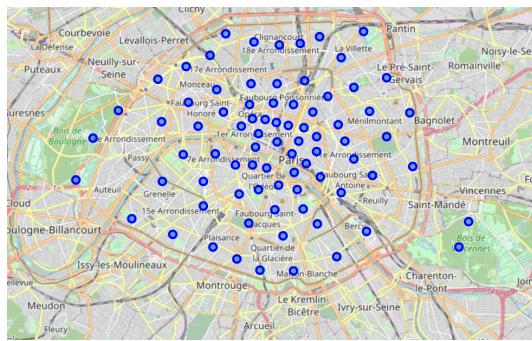


Figure 2. Paris neighborhoods

3.2 Specified category of venues in Paris Neighborhoods

First part of the analysis is about Paris neighborhoods. A list of values is selected as important criteria for the localisation of the shop. These values will be analysed in each neighborhood to choose the best ones thanks to Foursquare. The list of criteria selected to try to identify children presence and to see the competition from other stores is :

- Density population
 - Candy stores
 - Elementary Schools
 - Playgrounds
 - Toy/Games/Video Games Stores
 - Athletics and Sports
 - Sport Clubs

For each of these criteria, a function is used to get a dataframe of all the venues of this specified category near a point in each neighborhood (except density which is already known). Then the dataframe is group by neighborhood to count the number of specific venues of each neighborhoods. Finally this criteria is plot in a Choropleth map of Paris according to the quantitative feature.

3.3 Dylan's Candy Bars

The second step concerns the Candy Stores. The list of stores in the United States is retrieved and the geographic coordinates are obtained through Geocoder. Then the Foursquare search query is applied to each of these addresses to get venues around the stores. These are all the categories of venues that are collected here. At the same time, it is the list of venues by district

of Paris which is collected in a manner similar to the work on New York and Toronto.

The two dataframes containing the frequencies of the categories of venues in each district are grouped together. The K-means algorithm is then applied to try to identify the districts of Paris similar to the districts where the Dylan's stores are installed. The Kmeans algorithm was executed with $K = 9$. This number of clusters was selected because for fewer clusters, the number of districts of Paris present in the cluster with the Dylan's bars was too large.

4 Results

4.1 Neighborhoods of Paris

The following graphs show the number of visitors to each district according to the categories:

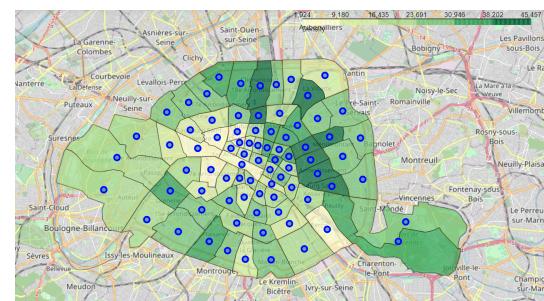


Figure 3. Density in Paris neighborhoods

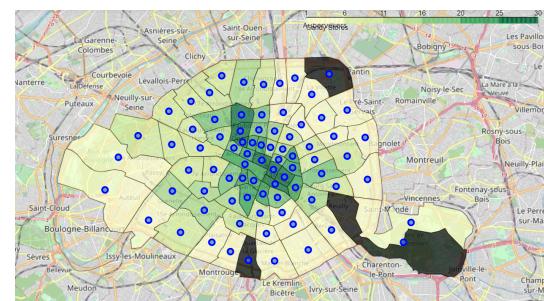


Figure 4. Candy Stores in Paris neighborhoods

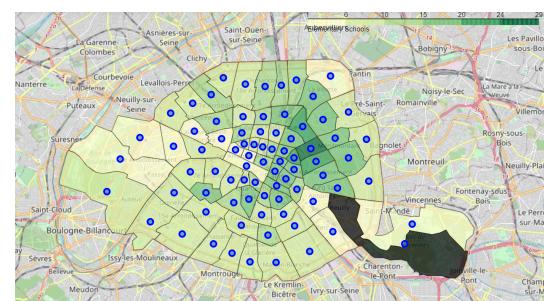


Figure 5. Elementary Schools in Paris neighborhoods

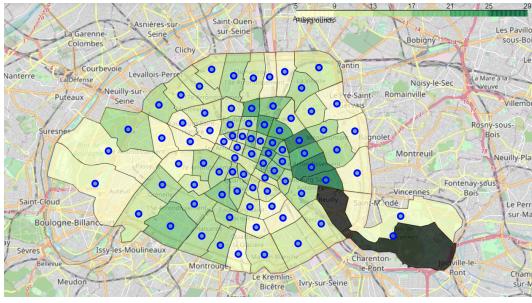


Figure 6. Playgrounds in Paris neighborhoods

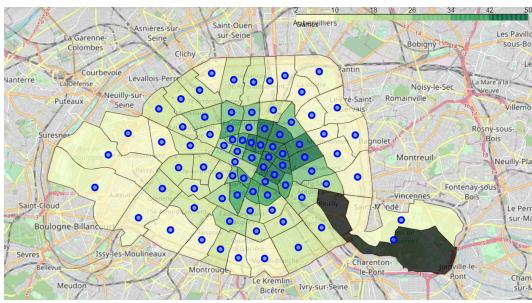


Figure 7. Toy/Games/Video Games Stores in Paris neighborhoods

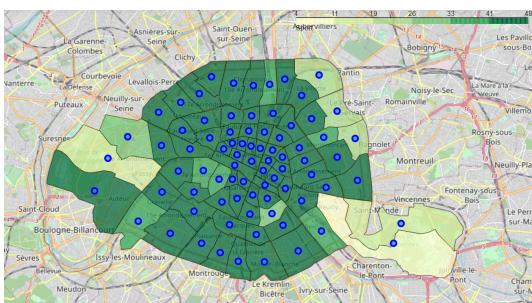


Figure 8. Athletics and Sports in Paris neighborhoods

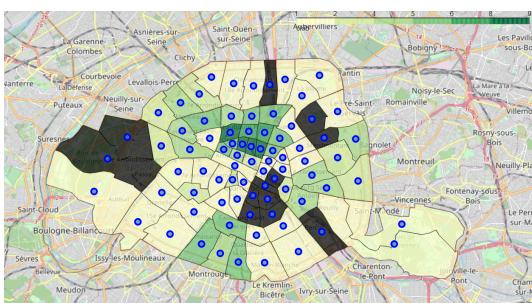


Figure 9. Sport Clubs in Paris neighborhoods

4.2 Dylan's Candy Bars and K-Means

The algorithm therefore created 9 clusters. All the Dylan's stores are in the same cluster except the one in the Bahamas and the one in Honolulu. These two stores are each in a cluster because of their specificity. The others are therefore in the same cluster with 4 districts of Paris:

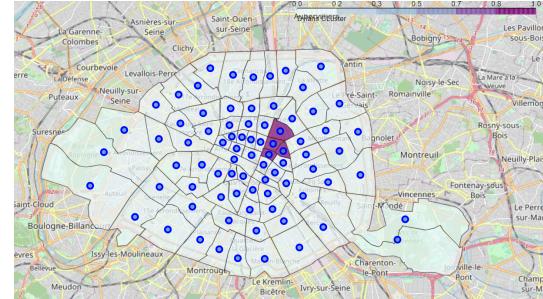


Figure 10. Paris neighborhoods in Dylan's Stores Cluster

5 Discussion

5.1 Analysis

The analysis of Paris neighborhoods give the fellow insights :

1. Density is high in the east of Paris.
2. The candy stores are strongly present in the center of Paris, in districts with fewer people.
3. Elementary schools and playgrounds are mostly located in the east of Paris and have a fairly similar density.
4. Toys and games stores are more widely present, still located more east of Paris.
5. Finally for sport, athletics and sports are present everywhere in the capital and are not a relevant indicator for the choice of the district. Conversely, sports clubs are present in the center of Paris.

It seems that the districts rather towards the east seem indicated to install the new store. Indeed, there are few existing candy stores in this area. In addition, many schools, playgroups and sellers of toys and games are present there and increase the number of children potentially present in this area.

The four neighborhoods similar to the neighborhoods of Dylan's bars are precisely in this area which emerges from the analysis of the neighborhoods. The clustering of the districts according to all the comers therefore confirms the ideal location for the opening of a new store. They are at the crossroads of concentrations of elementary schools and playgrounds on one side, sports clubs on the other side and right in the middle of the concentration of games and toys shops.

5.2 Limits

On the one hand, the list of criteria selected to analyze the districts of Paris relates mainly to the potential presence of children. However, the potential clientele could be increased and other criteria could then be added. In addition, other criteria concerning children could have been incorporated.

On the other hand concerning clustering, the districts that stand out are those most similar to American districts. However, certain Parisian districts could have specificities which separate them from districts of other countries and yet could be good locations to welcome a new store.

6 Conclusion

The recommendation and therefore the answer to the business problem is to open a candy bar in one of these four districts: Arts-et-Métiers, Enfants-Rouges, Porte-Saint-Martin and Sainte-Avoie. They are similar to the neighborhoods of Dylan's bars in the United States and correspond to the initial analysis looking for the potential presence of children.