Battle of Neighbourhoods

By Pedro Rocha 11/03/2021

1. Introduction

1.1 Background

Ciudad Autónoma de Buenos Aires (Autonomous City of Buenos Aires), is the capital and largest city of Argentina. Buenos Aires' quality of life was ranked 91st in the world in 2018, being one of the best in Latin America. In 2012, it was the most visited city in South America, and the second-most visited city of Latin America.

According to the World Travel & Tourism Council,tourism has been growing in the Argentine capital since 2002. In a survey by the travel and tourism publication Travel + Leisure Magazine in 2008, visitors voted Buenos Aires the second most desirable city to visit after Florence, Italy.

The city attracts visitors from all over the world, in part because of its rich gastronomic culture.

The gastronomic and hotel industry represented 3.4% of CABA's GDP in 2016. (Source: SSPMicro based on the Department of Statistics and Censuses of CABA. PBG participation in current prices and evolution based on PBG constant prices 2004)

1.2 Problem

This project analyzes the neighbourhoods in CABA to find the best ones to start a gastronomic business.

1.3 Interest

This project is of interest to people who plan to open a restaurant or people who already have one and are thinking of opening a new branch. This project is also interesting for people who plan to visit CABA to find out in which neighborhoods they can find a varied gastronomic offer.

2. Data

2.1 Data Sources

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

- Average income for each neighbourhood
- Is the neighbourhood touristic? Is it a Hotel zone?
- Other Touristic Gastronomy Offers in each neighbourhood
- Principal venues in each neighbourhood

You can find the data sets used in this project in the Buenos Aires Statistics Official Site:

Gastronomic Offer in CABA

Average total family income (ITF) according to commune. Buenos aires city. Years 2008/2019

The Average total family income is calculated by commune so I calculated a weighting based on the average price per M2 of apartments for sale in each neighborhood: Average price per m2 (dollars) of apartments for sale of 2 used rooms per neighborhood. Buenos aires city. 4th. 2006 / 4th quarter. quarter 2020

The principal venues in each neighbourhood were obtained from *Foursquare API*. You can check the documentation HERE

2.2 Data Cleaning

Data downloaded from multiple sources were combined into one table.
The only data about average income found was according to commune, so it was weighted based on the average price per M2 of apartments for sale in each neighborhood in order to calculate the average income per Neighbourhood.
The weighting was made as follows:
(Neighbourhood M2 AVG price / Commune M2 AVG price)

The geopy library was used to get each Neighbourhood coordinates and also was merged into the dataframe.

Villa General Mitre was renamed from the original list in order to use the geopy library.

2.3 Feature Selection

The features selected for the analysis are:

- Average Income per Family per Neighbourhood

Families with a higher income are more likely to eat out. So in the neighbourhoods with a higher average income a restaurant has more potential customers

- Hotels in the Neighbourhood

CABA is a touristic city and tourists are always good customers, so we think that Neighbourhoods with hotels can be a good target to start a restaurant.

- Restaurants in the Neighbourhood

People are likely to go to places where they have a variety of food options and then choose the restaurant they seem to like. So have other restaurant nearby is an opportunity to catch customers

- Principal venues categories in the Neighbourhood

This feature will help to know which kind of venues we can find in each Neighbourhood.

2.4 Exploratory Data Analysis.

CABA Neighbourhoods

CABA is composed of forty-eight neighborhoods (locally known as barrios). Since 2008, the city is also legally divided into communes, each one including one or more barrios.

	Neighbourhood	Comuna		Neighbourhood	Comuna
0	Agronomía	15	24	Parque Chas	15
1	Almagro	5	25	Parque Patricios	4
2	Balvanera	3	26	Puerto Madero	1
3	Barracas	4	27	Recoleta	2
4	Belgrano	13	28	Retiro	1
5	Boedo	5	29	Saavedra	12
6	Caballito	6	30	San Cristóbal	3
7	Chacarita	15	31	San Nicolás	1
8	Coghlan	12	32	San Telmo	1
9	Colegiales	13	33	Vélez Sarsfield	10
10	Constitución	1	34	Versalles	10
11	Flores	7	35	Villa Crespo	15
12	Floresta	10	36	Villa del Parque	11
13	La Boca	4	37	Villa Devoto	11
14	La Paternal	15	38	Villa General Mitre	11
15	Liniers	9	39	Villa Lugano	8
16	Mataderos	9	40	Villa Luro	10
17	Montserrat	1	41	Villa Ortúzar	15
18	Monte Castro	10	42	Villa Pueyrredón	12
19	Nueva Pompeya	4	43	Villa Real	10
20	Núñez	13	44	Villa Riachuelo	8
21	Palermo	14	45	Villa Santa Rita	11
22	Parque Avellaneda	9	46	Villa Soldati	8
23	Parque Chacabuco	7	47	Villa Urquiza	12

Figure 1. CABA's Neighbourhoods List

Average Total Family Income per Neighbourhood

Top 15 Neighbourhoods per Average Total Family Income.

	Neighbourhood	Comuna	Weighted Income
24	Puerto Madero	1	133733.04
14	Belgrano	13	89658.24
39	Palermo	14	85061.00
18	Coghlan	12	80565.30
21	Villa Urquiza	12	79045.20
19	Saavedra	12	77525.10
34	Villa Luro	10	77204.40
40	Recoleta	2	75944.00
15	Colegiales	13	75248.88
16	Núñez	13	75248.88
25	Retiro	1	74134.62
5	Villa Ortúzar	15	73570.68
4	Villa Crespo	15	73570.68
1	Chacarita	15	71033.76
17	Caballito	6	69510.00

Figure 2. Top 15 Neighbourhoods per Average Total Family Income (2020) list

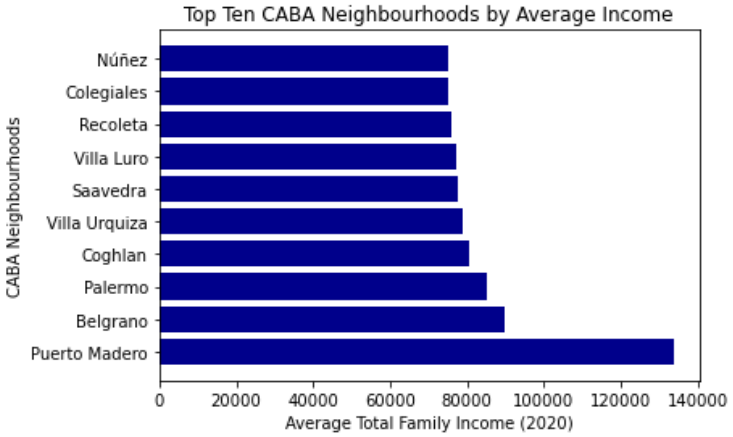

Figure 3. Top 10 Neighbourhoods by Average Total Family Income (2020)

Hotels in the Neighbourhood

To obtain the number of hotels in each neighborhood, the Foursquare API is used. The limit of results obtained is 100, so some neighborhoods show that value. This does not influence our analysis because then these values were used to create a binary variable that categorizes each neighborhood as a hotel zone or not.

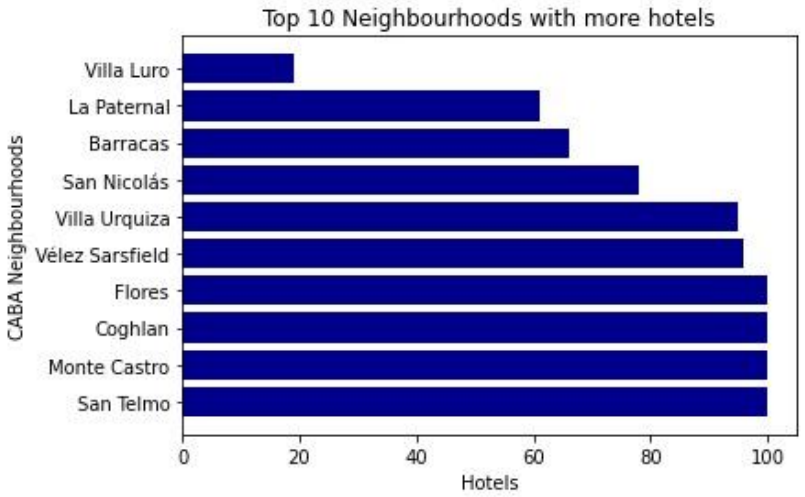

Figure 4. Top 10 Neighbourhoods with more hotels

Restaurants in the Neighbourhood

These values were used to create a binary variable that categorizes each neighborhood as a Gastronomical area or not.

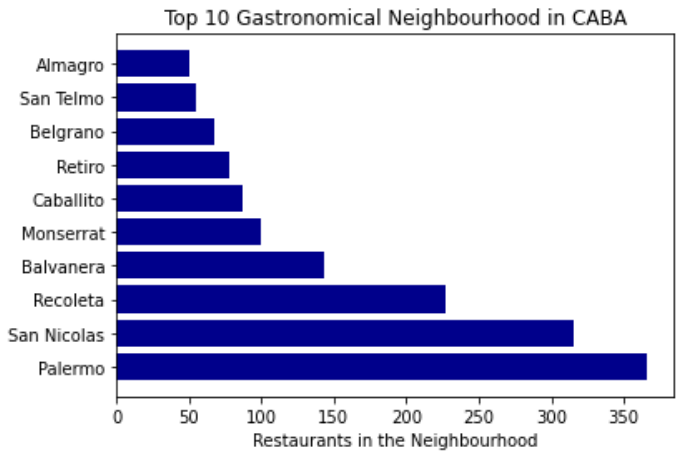

Figure 5. Top 10 Gastronomical Neighbourhood in CABA

3. Classification Model

In order to cluster the neighbourhoods I'll use the K-means algorithm.
K-means algorithm identifies *k* number of centroids, and then allocates every data point to the nearest cluster, while keeping the centroids as small as possible.
So, a key part of this model is to select the optimal k value.
For that I used the elbow method:

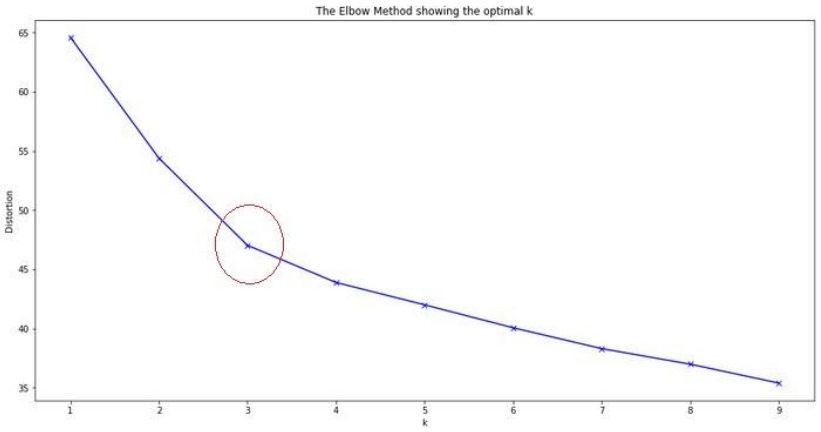

Figure 6. Elbow Method

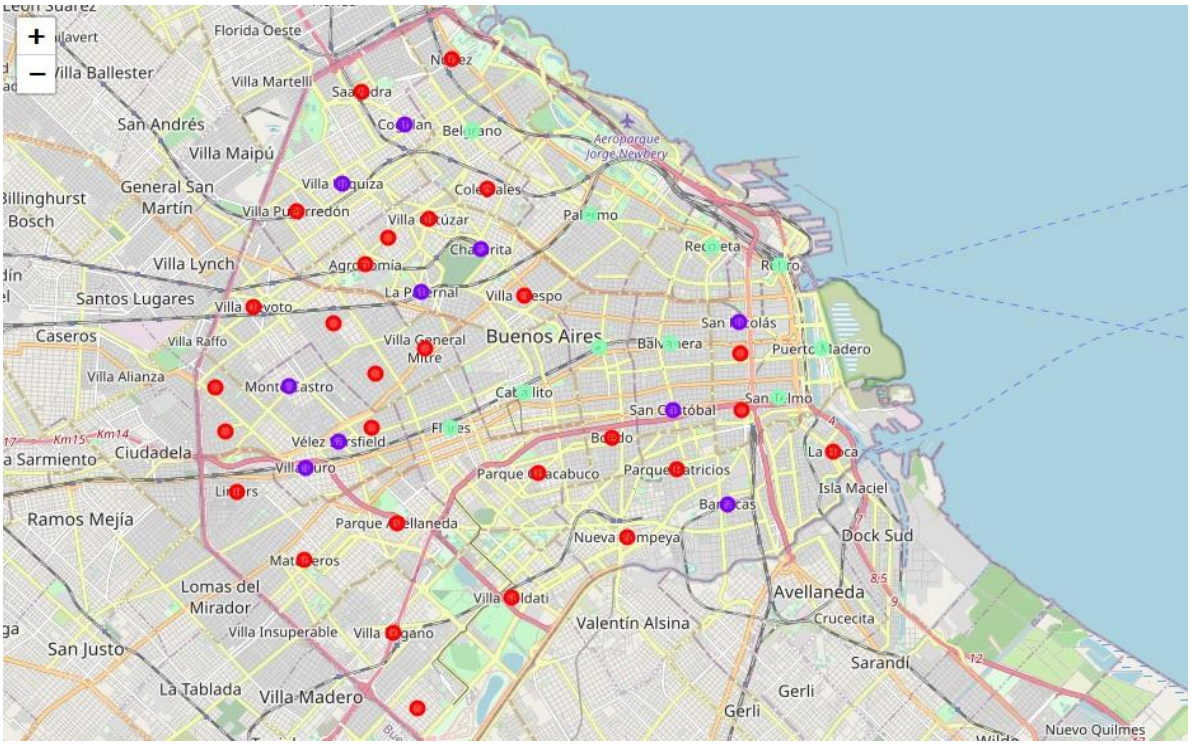

Figure 7. CABA - Clusters

4. Conclusions

We have segmented CABA in three clusters, now we can determine what are the characteristics of each.

Red Cluster: represents the neighborhoods that are not hotel or gastronomic areas.

Purple Cluster: represents the principal hotel neighbourhoods in CABA

Green Cluster: gastronomy neighbourhoods in CABA and we can see that most are high income too.

Our problem is where to start a gastronomical business in CABA, so our target neighbourhoods are the ones in the green cluster.

	Neighbourhood	Weighted Income	Hotel Area	Gastronomy Area	1st Most Common Venues	2nd Most Common Venues	3rd Most Common Venues	4th Most Common Venues	5th Most Common Venues
6	Almagro	59530.41	0	1	Restaurant	Gaming Cafe	Fried Chicken Joint	French Restaurant	Food Truck
8	Balvanera	46279.25	0	1	Burger Joint	Yoga Studio	Garden Center	Furniture / Home Store	Fried Chicken Joint
14	Belgrano	89658.24	0	1	Bakery	Yoga Studio	Flea Market	Gaming Cafe	Furniture / Home Store
17	Caballito	69510.00	0	1	Pizza Place	Yoga Studio	Fire Station	Fried Chicken Joint	French Restaurant
24	Puerto Madero	133733.04	0	1	Cultural Center	Yoga Studio	Fish Market	Furniture / Home Store	Fried Chicken Joint
25	Retiro	74134.62	0	1	Beer Bar	Yoga Studio	Flea Market	Gaming Cafe	Furniture / Home Store
27	San Telmo	56691.18	1	1	Sandwich Place	Yoga Studio	Fire Station	Fried Chicken Joint	French Restaurant
28	Flores	53624.28	1	1	Bar	Yoga Studio	Flea Market	Gaming Cafe	Furniture / Home Store
39	Palermo	85061.00	0	1	Office	Yoga Studio	Fish Market	Furniture / Home Store	Fried Chicken Joint
40	Recoleta	75944.00	0	1	Bagel Shop	Yoga Studio	Gaming Cafe	Furniture / Home Store	Fried Chicken Joint

Figure 8 - Gastronomical Neighbourhoods in CABA - Green Cluster