

Strategic Clustering for Establishing a Restaurant in Mexico City

1. Introduction

1.1 Background

To have success in a restaurant, we need to consider multiple factors such as the type of the restaurant, ratings, prices, locality, etc. But one of the most important is understanding the customers: Which places do they prefer? What type of food is their favorite? What type of restaurant do they prefer? What do they expect about the menu? The answers to all these questions might be different for every neighborhood from the City. Knowing the neighborhood and what their customers expect or are used to might be the key to the fast growth of a restaurant. However, finding the best location in a big city such as Mexico City could be very difficult. So, implementing a Data Science analysis and using a Machine Learning algorithm to classify some of the neighborhoods from Mexico City, based on their most popular restaurant, will give some insight into the customers' preferences.

1.2 Problem

A seafood restaurant is going to be established in Mexico City. The owner wants to locate it in a strategic place, in a neighborhood where the customers like to go to big restaurants and where the restaurant has the opportunity to grow despite the competition of other restaurants already established.

The objective of this capstone project is to find the optimal neighborhood in Mexico City for the restaurant. The [Foursquare API](#) would be used to get the most popular restaurants of each neighborhood and then classify them to know the preferences of the people around the location

.

1.3 Interest

Restaurant analytics are really useful to restaurant owners. It delivers insights that help to understand what drives the profitability of the restaurant, saving time and winning money. The location of the restaurant could be one of the most important factors, as it could tell us a lot about what the customers like and what they are used to.

2. Data Acquisition and cleaning

2.1 Data source

The Mexican Postal Codes data are obtained from the [GreatData](#) Database. Which has the latitude and longitude information of each neighborhood of Mexico City. A sample data set can be freely downloaded in a CSV file. All these locations are used to extract information of venues, specifically of the restaurants, utilizing Foursquare API, specifically, its explore function.

2.2 Data Cleaning

The image below shows the data frame with the information of 1000 neighborhoods from Mexico City. The Type Code of each Neighborhood could be:

- ❖ 'BO': Barrio.
- ❖ 'CAMP': Campamento.
- ❖ 'COL': Colonia.
- ❖ 'EQUIP': Equipamiento.
- ❖ 'PBO': Poblado.

According to the database information previously downloaded.

	Postal Code	Sequence No.	State Code	Type Code	Place Name	Municipality Name	Major City Name	Latitude	Longitude
0	1000	1	DF	COL	San Angel	Álvaro Obregón	México	19.346257	-99.193169
1	1010	1	DF	COL	Los Alpes	Álvaro Obregón	México	19.360502	-99.193958
2	1020	1	DF	COL	Guadalupe Inn	Álvaro Obregón	México	19.349653	-99.197087
3	1030	1	DF	PBO	Axotla	Álvaro Obregón	México	19.359138	-99.176073
4	1030	2	DF	COL	Florida	Álvaro Obregón	México	19.356070	-99.181523
...
995	10840	1	DF	BO	Las Calles	La Magdalena Contreras	México	19.304080	-99.232064
996	10840	2	DF	BO	Plazuela del Pedregal	La Magdalena Contreras	México	19.304080	-99.232064
997	10900	1	DF	PBO	San Nicolás Totolapan	La Magdalena Contreras	México	19.297665	-99.242156
998	10910	1	DF	PBO	La Magdalena	La Magdalena Contreras	México	19.302101	-99.245791
999	10920	1	DF	COL	Las Huertas	La Magdalena Contreras	México	19.298948	-99.240467

The following table counts the number of neighborhoods with each type of code:

Type Code	Number of Neighborhoods
BO	55
CAMP	1
COL	890

EQUIP	2
PBO	52

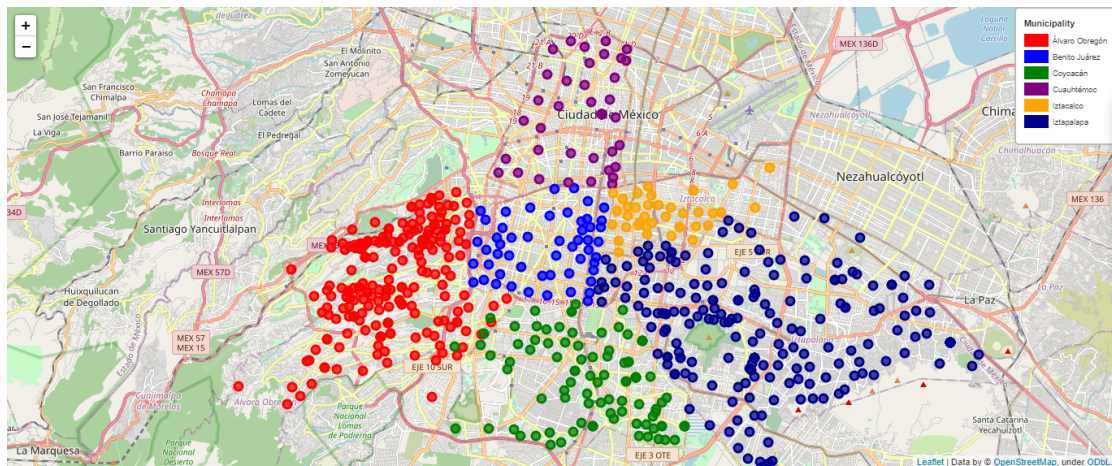
Only the neighborhoods with type codes 'BO', 'COI', and 'PBO' are selected. And only the municipalities below are selected, most of them represent the center of CDMX. Besides, due to there is a call quota of 950 API calls that can be made in a 24 hour period, we could not make a call for all the neighborhoods from the city.

Municipality names:

- ❖ Álvaro Obregón
- ❖ Benito Juárez
- ❖ Coyoacán
- ❖ Cuauhtémoc
- ❖ Iztacalco
- ❖ Iztapalapa

2.3 Feature selection

The map below, made with folium library, shows the neighborhoods of the 6 selected municipalities; they are grouped by the municipality.



We have 645 neighborhoods and the number of neighborhoods of each municipality is pointed out in the next data frame:

Number of Neighborhoods	
Municipality Name	
Benito Juarez	53
Coyoacan	95
Cuauhtémoc	41
Iztacalco	38
Iztapalapa	199
Alvaro Obregon	219

3. Exploratory Data Analysis

3.1 The 10 most common type of restaurants

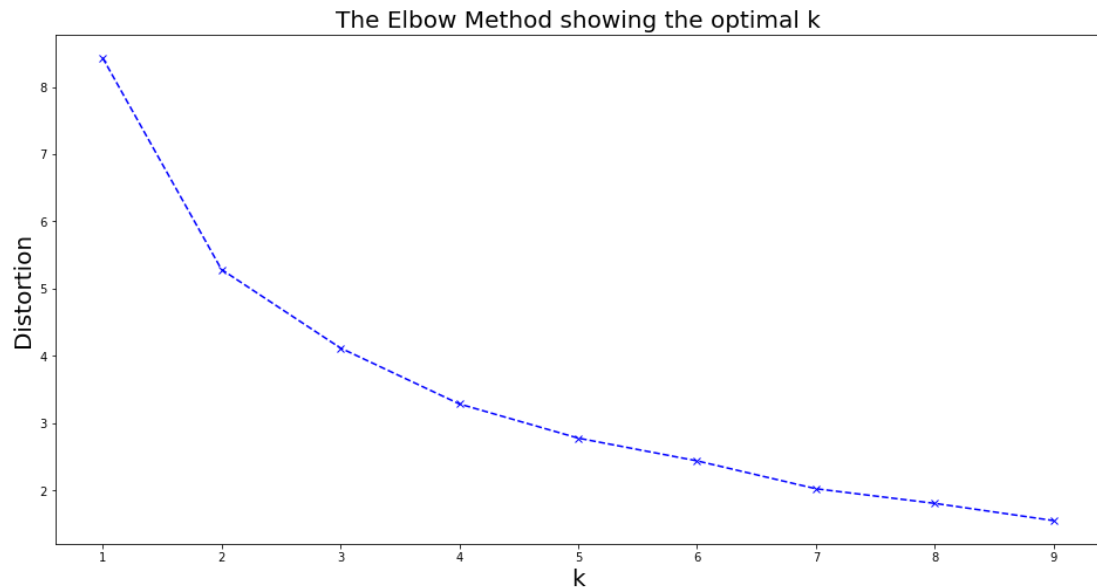
For each neighborhood, the 100 nearest venues in a radius of 500 meters were considered and just the venue categories related to food or restaurants were picked, then the 10 most common types or restaurants were calculated for each municipality, as it is shown in the data frame below.

Municipality Name	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
Alvaro Obregon	Mexican Restaurant	Seafood Restaurant	Restaurant	Food Truck	Sushi Restaurant	Fast Food Restaurant	Food	Italian Restaurant	Food Court	Japanese Restaurant
Benito Juarez	Mexican Restaurant	Seafood Restaurant	Restaurant	Food Truck	Italian Restaurant	Sushi Restaurant	Argentinian Restaurant	Japanese Restaurant	Fast Food Restaurant	Food
Coyoacan	Mexican Restaurant	Restaurant	Seafood Restaurant	Food Truck	Sushi Restaurant	Italian Restaurant	Argentinian Restaurant	Fast Food Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant
Cuauhtémoc	Mexican Restaurant	Restaurant	Seafood Restaurant	Argentinian Restaurant	Japanese Restaurant	Spanish Restaurant	Italian Restaurant	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Comfort Food Restaurant
Iztacalco	Mexican Restaurant	Restaurant	Food Truck	Seafood Restaurant	Fast Food Restaurant	Argentinian Restaurant	Sushi Restaurant	Chinese Restaurant	Food	Food & Drink Shop
Iztapalapa	Mexican Restaurant	Seafood Restaurant	Food Truck	Fast Food Restaurant	Restaurant	Food Court	Sushi Restaurant	Food	Asian Restaurant	Food & Drink Shop

In Iztaclco municipality, the seafood restaurants are less common than the restaurants from the other municipalities. So this is a good place to establish the restaurant to have as little competition as possible.

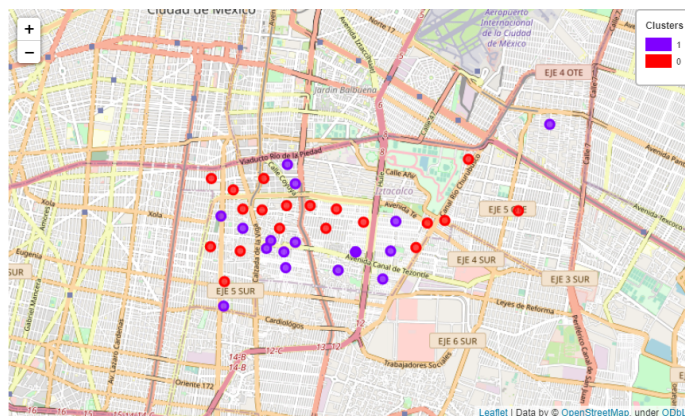
4. Predictive Modeling

The K-means algorithm is used to cluster the neighborhoods in Iztacalco municipality based on the category of the venue. The elbow method calculates the optimal number of clusters: $k=2$. The evaluation is shown below:



5. Results

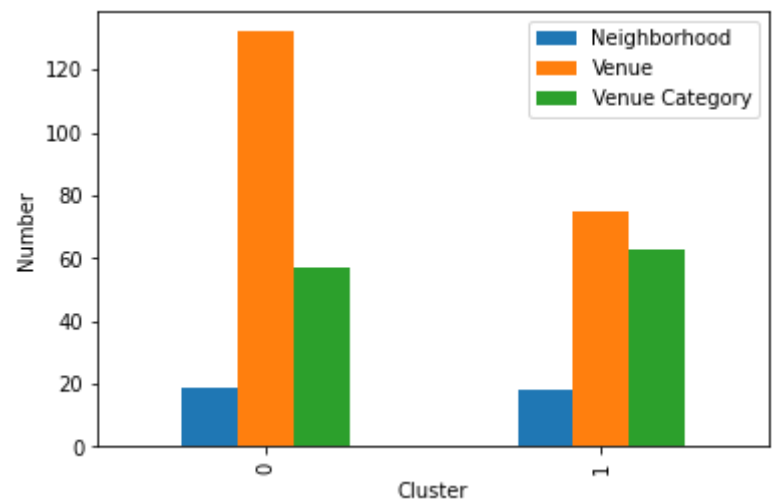
The location of each neighborhood classified into 2 clusters is shown in the next map. The neighborhoods from cluster 0 are marked with red circles and those from cluster 1 are marked with purple circles.



	Cluster 0	Cluster 1
0	Agrícola Oriental	Agrícola Pantitlán
1	El Rodeo	Ampliación Gabriel Ramos Millán
2	Ex-Ejido de La Magdalena Mixiuhca	Campamento 2 de Octubre
3	Gabriel Ramos Millán Sección Cuchilla	Carlos Zapata Vela
4	Gabriel Ramos Millán Sección Tlacotal	Cuchilla Agrícola Oriental
5	Granjas México	Fraccionamiento Coyuya
6	Juventino Rosas	Gabriel Ramos Millán
7	Los Picos de Iztacalco Sección 1a	Gabriel Ramos Millán Sección Bramadero
8	Los Reyes	INFONAVIT Iztacalco
9	Militar Marte	INPI Picos
10	Nueva Santa Anita	Jardines Tecma
11	Reforma Iztacchhuatl Norte	La Cruz
12	San Francisco Xicaltongo	Los Picos de Iztacalco Sección 1b
13	San Miguel	Los Picos de Iztacalco Sección 2a
14	Santa Anita	Reforma Iztacchhuatl Sur
15	Santiago Sur	San Pedro
16	Tlazintla	Santa Cruz
17	Viaducto Piedad	Santiago Norte
18	Zapotita	NaN

There are 19 neighborhoods in cluster 0 and 18 neighborhoods in cluster 1, 130 food venues in cluster 0 and 83 food venues in cluster 1, 53 different

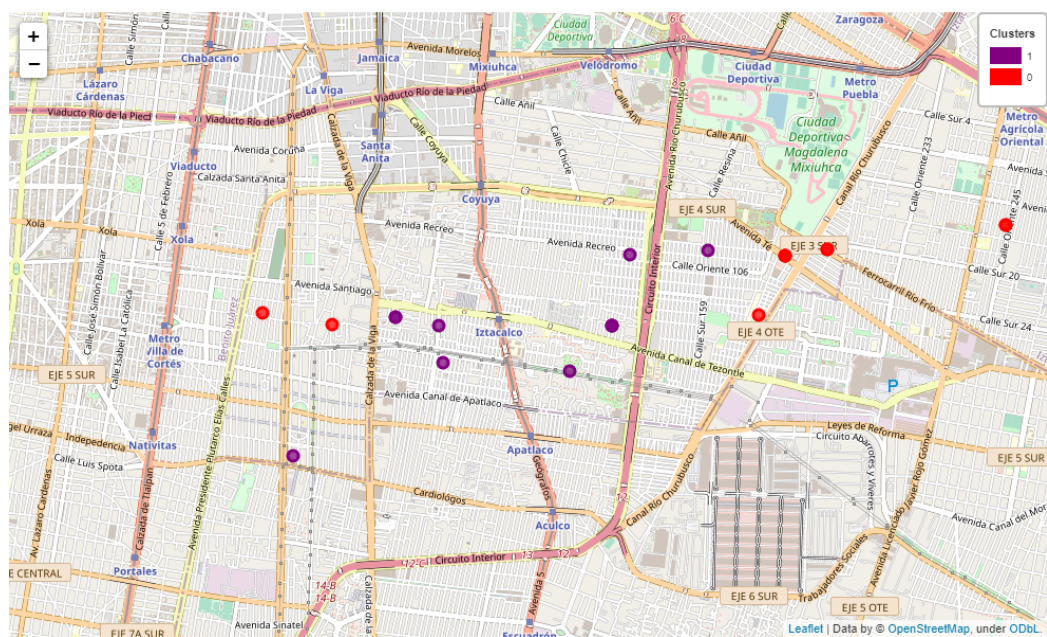
categories of venues in cluster 0, and 65 categories of venues in cluster 1.



On one hand, the 10 most common categories of venues for each cluster were calculated. As it is shown below:

Cluster		1st Most Common Venue Category	2nd Most Common Venue Category	3rd Most Common Venue Category	4th Most Common Venue Category	5th Most Common Venue Category	6th Most Common Venue Category	7th Most Common Venue Category	8th Most Common Venue Category	9th Most Common Venue Category	10th Most Common Venue Category
0	0	Seafood Restaurant	Mexican Restaurant	Restaurant	Argentinian Restaurant	Fast Food Restaurant	Food Truck	Sushi Restaurant	Food & Drink Shop	Chinese Restaurant	Japanese Restaurant
1	1	Mexican Restaurant	Seafood Restaurant	Restaurant	Food Truck	Fast Food Restaurant	Sushi Restaurant	Chinese Restaurant	Food & Drink Shop	Food Court	Food Stand

The seafood restaurants are marked in red for cluster 0 and purple for cluster 1. They are located on the next map:



6. Discussion

Seafood Restaurants are in the top four most common categories of restaurants in all the municipalities considered, which tells us that this kind of food is one of the favorites. Being Mexican Restaurants the most frequent category of the 6 municipalities analyzed. In Iztacalco, Seafood Restaurants are less common than in other considered municipalities. They are more common in Álvaro Obregón, Benito Juárez and Iztapalapa. For this reason, a seafood restaurant in Iztacalco might have the best opportunity to grow despite the competition of other restaurants of the same category.

Afterwards the neighborhoods in Iztacalco were clustered into 2 groups. Approximately, each cluster has the same amount of neighborhoods. The majority of the neighborhoods of cluster 0 are located in the west part of the municipality. Cluster 0 has a greater number of restaurants and venues of food. However, cluster 1 is more diverse, it has more categories of venues.

At the same time, to locate the restaurant in a place where the customers are expected or used to find a seafood restaurant. When we look at the most common venue categories of each cluster, the most common for cluster 0 is 'Seafood Restaurant'. So, establishing a seafood restaurant in a neighborhood of cluster 0 might be the best option

7. Conclusions

According to the category of the venue of the 6 municipalities analyzed, in order to find as little competition as possible for establishing a seafood restaurant: The Iztacalco municipality is the best option. The neighborhoods of the mentioned municipality were clustered in two groups based on their category of restaurants. And taking into consideration what kind of restaurants the customers might be expected to find around each neighborhood, establishing the restaurant in one neighborhood of cluster 0 might be the most strategic place to have a popular restaurant.

8. Future Directions

It would be interesting to analyze the 10 remaining municipalities that could not be considered due to the quota calls of our type of count of Foursquare API. And It is important to mention that this analysis depends on the current data and could quickly change, so, it needs constant updates.

9. References

- [1] [Jupyter Notebook:
Strategic-Clustering-for-Establishing-a-Restaurant-in-Mexico-City](#)
- [2] [GreatData](#)
- [3] [Foursquare API](#)