The Battle of the Neighborhoods

IBM-Coursera Capstone Project

Table of Contents

Introduction: Business Problem

The purpose of this project is to find the optimal location of a restaurant in Northwest Mexico City, an area full of businesses, offices, hotels, museums and tourists. This project is aimed at entrepreneurs specialized in the food industry.

The project will provide some places ideal for a restaurant, taking into consideration that the place is not already crowded with restaurants and that it is close enough to an area easily accessible. It will also give some insight of what type of restaurants are already within an area to give options of what type of restaurants would be attractive for potential costumers. For example, if a place is not very crowded with restaurants and within the area there are no seafood restaurants, then a seafood restaurant would be a good option.

Data

An area of Northwest Mexico City will be divided into grids based on postal codes, neighborhoods and municipios (boroughs). This data will be used by the Foursquare API to gather data about restaurants within the zone, and the frequency of the type of restaurants within the area.

The postal codes of the neighborhoods of Mexico City divided by municipio (borough) is available on the following website:

https://micodigopostal.org/ciudad-de-mexico/

Data of the coordinates of the postal codes is found on the following website:

https://datos.cdmx.gob.mx/explore/dataset/codigos-postales-de-la-cdmx/table/

To obtain the data about restaurants, an area of Mexico City will be divided into a grid and that grid will be entered into the Foursquare API. Information about restaurants will be gathered and analyzed to obtain data about the type and number of restaurants within an area.

Methodology

This project will require postal codes and their coordinates to use them as inputs for the Foursquare API to retrieve data from venues around the coordinates. Only data from three

"municipios" (boroughs), located in Northwest Mexico City will be used. All data that falls into the categories of restaurants, comfort food and fast food places are taken into account.

The most popular types of restaurants are then obtained from the data. This will help to make choices about what kind of restaurants are preferred by customers but at the same time are not common around a particular area. Also, for each postal code, the number of restaurants and their proportion will be calculated. These will require basic statistical analysis (counting the frequency, obtaining the proportion of a given sample, etc.).

Once the most popular types of restaurants around the postal codes are obtained, by using the K-Means algorithm, wider areas gathered into clusters will give a more precise idea of how crowded of restaurants an area is and the type more common around that area. The areas not too crowded will be good candidates for a location of a new restaurant, especially if it is one of the popular types in Mexico City.

Results and Discussion

The analysis results show the *food preferences* in Northwest Mexico City, how crowded the *areas* of restaurants are, and the *top ten types of restaurants per neighborhood*.

Customers prefer **Mexican restaurants** and **taco places**, followed by the generic label of **"restaurant"**, in which probably are gathered restaurants serving home-made style and typical Mexican dishes, as well as **seafood restaurants** (Mexico City inhabitants love touristic destinations by the coast and seafood is not commonly cooked at home), **pizza** and **burger places**. *Fast food* and *Italian, Japanese, Argentinian, and Spanish* restaurants are less common but still popular. With around 60 different kinds of food venues, it can be concluded that customers in Mexico City love food diversity and that more exotic cuisines are welcome.

Venue Category	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude
Mexican Restaurant	646	646	646	646	646	646
Taco Place	486	486	486	486	486	486
Restaurant	217	217	217	217	217	217
Seafood Restaurant	137	137	137	137	137	137
Pizza Place	103	103	103	103	103	103
Burger Joint	101	101	101	101	101	101
Breakfast Spot	82	82	82	82	82	82
Sandwich Place	60	60	60	60	60	60
Sushi Restaurant	60	60	60	60	60	60
Italian Restaurant	58	58	58	58	58	58
Fast Food Restaurant	51	51	51	51	51	51
Snack Place	44	44	44	44	44	44
Japanese Restaurant	37	37	37	37	37	37
Argentinian Restaurant	33	33	33	33	33	33
Salad Place	29	29	29	29	29	29
Food Court	25	25	25	25	25	25
Comfort Food Restaurant	21	21	21	21	21	21
Spanish Restaurant	21	21	21	21	21	21
Vegetarian / Vegan Restaurant	19	19	19	19	19	19
Bistro	18	18	18	18	18	18

Figure 1. Top 20 food venues in Mexico City

The most crowded areas are *Polanco IV Sección*, with 47 restaurants around, followed by *Cuauhtémoc* (45), *Nueva Santa María* (44) and *Hipódromo* (42).

	Neighborhood	Restaurants
108	Polanco IV Sección	47
41	Cuauhtémoc	45
92	Nueva Santa María	44
59	Hipódromo	42
119	Roma Norte	41
48	Escandón I Sección, Escandón II Sección	38
34	Centro (Área 7)	37
60	Hipódromo Condesa	36
107	Polanco III Sección	35
109	Polanco V Sección	35
39	Condesa	34
49	Esperanza	34
153	Sector Naval, Clavería	33
56	Francisco I Madero, Popo	32
63	Ignacio Allende, Victoria de las Democracias	30
32	Centro (Área 5)	29
15	Ampliación Torre Blanca	29
149	Santa María la Ribera	28
58	Guerrero	28
68	Juárez	27
2	5 de Mayo, Deportivo Pensil	27
11	Ampliación Granada	27
154	Tabacalera	26
17	Anzures	26

Figure 2. Number of Restaurants per Neighborhood (fragment).

The list that contains the *top ten types of restaurants per neighborhood* gives a clear idea of the proportion of restaurants by type around a postal code. For example, neighborhood *Ampliación Granada* is very diverse, apart from the typical Mexican food and taco places, there are Italian, Mediterranean, Cantonese, and German restaurants, sushi and pizza places.

```
----Ampliación Granada----
                      venue  freq
0         Mexican Restaurant  0.19
1                 Restaurant  0.11
2         Italian Restaurant  0.07
3                 Taco Place  0.07
4           Sushi Restaurant  0.07
5                 Food Court  0.07
6  Mediterranean Restaurant  0.04
7      Cantonese Restaurant  0.04
8               Pizza Place  0.04
9         German Restaurant  0.04


----Ampliación Popo----
                    venue  freq
0     Mexican Restaurant  0.44
1             Taco Place  0.12
2           Burger Joint  0.06
3     Spanish Restaurant  0.06
4      German Restaurant  0.06
5             Pizza Place  0.06
6             Snack Place  0.06
7              Restaurant  0.06
8       Sushi Restaurant  0.06
9  Portuguese Restaurant  0.00


----Ampliación Torre Blanca----
                  venue  freq
0  Mexican Restaurant  0.31
1          Taco Place  0.31
2        Burger Joint  0.10
3          Restaurant  0.07
4  Italian Restaurant  0.03
5          Food Court  0.03
6         Pizza Place  0.03
7         Snack Place  0.03
8  Seafood Restaurant  0.03
9      Sandwich Place  0.03
```

Figure 3. List of the 10 Top Restaurant Types by Neighborhood (fragment).

Northwest Mexico City was divided into ten clusters, as illustrated in *Figure 4*.

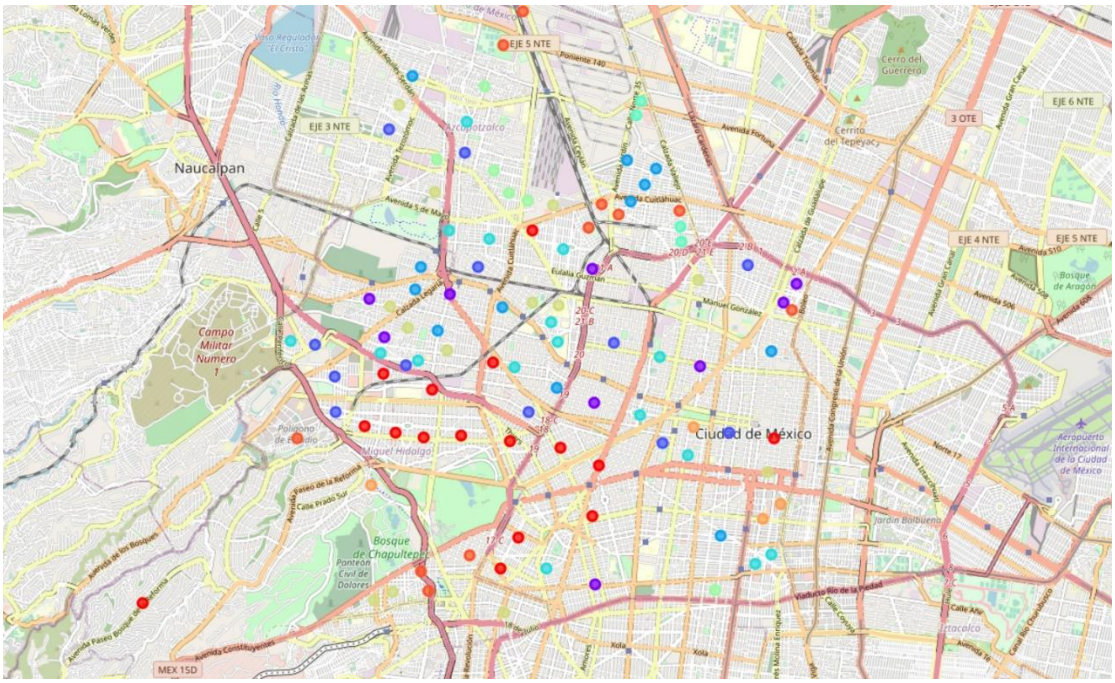

Figure 4. Postal Codes Gathered into Clusters.

Clusters 1 (red) and **5** (aqua) are the most crowded with restaurants, each one gathers 16 postal codes. **Cluster 1** is the most diverse , and the most popular are the Mexican restaurants, generic restaurants and Italian. **Cluster 5** is the second of the most crowded areas. It has mainly Mexican restaurants and taco places, some areas with Japanese food and sushi, and also Italian food. It is not diverse and probably has more informal places to eat.

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
60	2830	0.0	Restaurant	Mexican Restaurant	Taco Place
80	6060	0.0	Restaurant	Taco Place	Mexican Restaurant
85	6140	0.0	Taco Place	Restaurant	Italian Restaurant
86	6170	0.0	Restaurant	Burger Joint	Taco Place
99	6500	0.0	Japanese Restaurant	Taco Place	Mexican Restaurant
100	6600	0.0	Mexican Restaurant	Italian Restaurant	Breakfast Spot
101	6700	0.0	Italian Restaurant	Pizza Place	Restaurant
130	11310	0.0	Restaurant	Mexican Restaurant	Sandwich Place
150	11520	0.0	Restaurant	Mexican Restaurant	Burger Joint
151	11529	0.0	Mexican Restaurant	Restaurant	Italian Restaurant
152	11530	0.0	Restaurant	Mexican Restaurant	Italian Restaurant
153	11540	0.0	French Restaurant	Mexican Restaurant	Italian Restaurant
154	11550	0.0	Mexican Restaurant	Restaurant	Seafood Restaurant
155	11560	0.0	Seafood Restaurant	Italian Restaurant	Taco Place
157	11590	0.0	Restaurant	Japanese Restaurant	Mexican Restaurant
162	11700	0.0	Sushi Restaurant	Restaurant	Seafood Restaurant

Figure 5. Cluster 1.

Clusters 2 (purple), **3** (purplish blue), **4** (blue), **8** (olive green) and **10** (solid orange) follow, each one of them gathers 9,12, 12, 13 and 11 postal codes, respectively. **Cluster 2** is full of Mexican restaurants, taco places, fast food and comfort food. **Cluster 3** is full of Mexican restaurants and

taco places. It also has areas of Japanese and Spanish restaurants. **Cluster 4** has mainly taco places, Mexican restaurants and in third place seafood restaurants. It looks like a place full of informal, comfort, fast types of food. **Cluster 8** is an area of taco places and Mexican restaurants, and in third place food courts and hamburgers. **Cluster 10** is an area of taco places and Mexican restaurants.

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
62	2860	1.0	Mexican Restaurant	Taco Place	Restaurant
89	6240	1.0	Mexican Restaurant	Taco Place	Fast Food Restaurant
91	6270	1.0	Mexican Restaurant	Restaurant	Taco Place
93	6300	1.0	Mexican Restaurant	Taco Place	Breakfast Spot
98	6470	1.0	Mexican Restaurant	Taco Place	Restaurant
103	6760	1.0	Mexican Restaurant	Taco Place	Restaurant
125	11270	1.0	Mexican Restaurant	Taco Place	Seafood Restaurant
138	11410	1.0	Mexican Restaurant	Taco Place	Bistro
144	11470	1.0	Mexican Restaurant	Taco Place	Restaurant

Figure 6. Cluster 2.

Clusters 6 (green dots near Calzada Vallejo), **7** (green) and **9** (light orange) are the least crowded, each one gathers 4, 6 and 4 postal codes, respectively. **Cluster 6** has a mix of fast food and Mexican restaurants. **Cluster 7** has a mix of taco places and Mexican restaurants. It also has Venezuelan and Latin American food. Both **clusters 6** and **7** are areas around the trendiest places where more affordable food is available. **Cluster 9** is an area of seafood and Mexican restaurants.

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	2000	2.0	Mexican Restaurant	Taco Place	Breakfast Spot
8	2090	2.0	Mexican Restaurant	Taco Place	Breakfast Spot
35	2480	2.0	Mexican Restaurant	Taco Place	Food Court
74	6000	2.0	Mexican Restaurant	Spanish Restaurant	Restaurant
78	6040	2.0	Mexican Restaurant	Spanish Restaurant	Japanese Restaurant
88	6220	2.0	Mexican Restaurant	Taco Place	Breakfast Spot
95	6400	2.0	Mexican Restaurant	Pizza Place	Italian Restaurant
118	11200	2.0	Mexican Restaurant	Taco Place	Restaurant
129	11300	2.0	Mexican Restaurant	Restaurant	Taco Place
146	11489	2.0	Mexican Restaurant	Taco Place	German Restaurant
148	11500	2.0	Mexican Restaurant	Restaurant	Breakfast Spot
149	11510	2.0	Mexican Restaurant	Spanish Restaurant	Restaurant

Figure 7. Cluster 3.

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
16	2150	3.0	Taco Place	Mexican Restaurant	Restaurant
41	2600	3.0	Taco Place	Mexican Restaurant	Restaurant
43	2640	3.0	Taco Place	Restaurant	Mexican Restaurant
44	2650	3.0	Taco Place	Mexican Restaurant	Restaurant
64	2900	3.0	Taco Place	Mexican Restaurant	Seafood Restaurant
87	6200	3.0	Taco Place	Mexican Restaurant	Seafood Restaurant
105	6800	3.0	Taco Place	Mexican Restaurant	Seafood Restaurant
126	11280	3.0	Taco Place	Mexican Restaurant	Italian Restaurant
128	11290	3.0	Taco Place	Mexican Restaurant	Pizza Place
136	11370	3.0	Taco Place	Mexican Restaurant	Restaurant
137	11400	3.0	Taco Place	Mexican Restaurant	Restaurant
141	11440	3.0	Taco Place	Mexican Restaurant	Sushi Restaurant

Figure 8. Cluster 4.

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
1	2010	4.0	Mexican Restaurant	Seafood Restaurant	Breakfast Spot
7	2080	4.0	Mexican Restaurant	Taco Place	Pizza Place
9	2099	4.0	Mexican Restaurant	Burger Joint	Restaurant
58	2800	4.0	Mexican Restaurant	Taco Place	Sushi Restaurant
77	6030	4.0	Mexican Restaurant	Taco Place	Japanese Restaurant
81	6070	4.0	Mexican Restaurant	Taco Place	Restaurant
84	6100	4.0	Mexican Restaurant	Taco Place	Italian Restaurant
94	6350	4.0	Mexican Restaurant	Taco Place	Restaurant
109	6860	4.0	Mexican Restaurant	Taco Place	Pizza Place
110	6870	4.0	Mexican Restaurant	Taco Place	Restaurant
120	11220	4.0	Mexican Restaurant	Restaurant	Taco Place
131	11320	4.0	Mexican Restaurant	Burger Joint	Restaurant
132	11330	4.0	Mexican Restaurant	Restaurant	Salad Place
133	11340	4.0	Mexican Restaurant	Taco Place	Restaurant
145	11480	4.0	Mexican Restaurant	Restaurant	Taco Place
147	11490	4.0	Mexican Restaurant	Taco Place	Restaurant

Figure 9. Cluster 5.

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
24	2320	5.0	Fast Food Restaurant	Breakfast Spot	Mexican Restaurant
26	2340	5.0	Breakfast Spot	Fast Food Restaurant	Mexican Restaurant
70	2960	5.0	Mexican Restaurant	Taco Place	Breakfast Spot
73	2990	5.0	Mexican Restaurant	Breakfast Spot	Italian Restaurant

Figure 10. Cluster 6.

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
4	2050	6.0	Mexican Restaurant	Taco Place	Burger Joint
5	2060	6.0	Mexican Restaurant	Restaurant	Taco Place
20	2240	6.0	Mexican Restaurant	Taco Place	Sushi Restaurant
40	2530	6.0	Mexican Restaurant	Latin American Restaurant	Taco Place
52	2730	6.0	Mexican Restaurant	Taco Place	Burger Joint
111	6880	6.0	Mexican Restaurant	Taco Place	Venezuelan Restaurant

Figure 11. Cluster 7.

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
32	2459	7.0	Taco Place	Mexican Restaurant	Food Court
37	2500	7.0	Burger Joint	Mexican Restaurant	Taco Place
57	2790	7.0	Mexican Restaurant	Taco Place	Burger Joint
83	6090	7.0	Taco Place	Mexican Restaurant	Argentinian Restaurant
114	6920	7.0	Taco Place	Mexican Restaurant	Burger Joint
127	11289	7.0	Taco Place	Mexican Restaurant	Burger Joint
134	11350	7.0	Taco Place	Mexican Restaurant	Breakfast Spot
139	11420	7.0	Taco Place	Breakfast Spot	Mexican Restaurant
140	11430	7.0	Burger Joint	Taco Place	Restaurant
142	11450	7.0	Taco Place	Mexican Restaurant	Burger Joint
163	11800	7.0	Taco Place	Mexican Restaurant	Burger Joint
165	11820	7.0	Taco Place	Mexican Restaurant	Restaurant
166	11830	7.0	Taco Place	Mexican Restaurant	Sushi Restaurant

Figure 12. Cluster 8.

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
79	6050	8.0	Mexican Restaurant	Seafood Restaurant	Restaurant
106	6820	8.0	Seafood Restaurant	Mexican Restaurant	Taco Place
107	6840	8.0	Seafood Restaurant	Mexican Restaurant	Fast Food Restaurant
116	11040	8.0	Restaurant	Seafood Restaurant	Mexican Restaurant

Figure 13. Cluster 9.

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
19	2230	9.0	Taco Place	Salad Place	Mexican Restaurant
23	2310	9.0	Taco Place	Salad Place	Argentinian Restaurant
59	2810	9.0	Taco Place	Mexican Restaurant	Sushi Restaurant
66	2920	9.0	Taco Place	Mexican Restaurant	Seafood Restaurant
68	2940	9.0	Taco Place	Restaurant	Mexican Restaurant
69	2950	9.0	Taco Place	Mexican Restaurant	Pizza Place
92	6280	9.0	Restaurant	Taco Place	Mexican Restaurant
161	11650	9.0	Restaurant	Taco Place	Mexican Restaurant
167	11840	9.0	Taco Place	Mexican Restaurant	Burger Joint
168	11850	9.0	Taco Place	Mexican Restaurant	Comfort Food Restaurant
169	11860	9.0	Taco Place	Mexican Restaurant	Seafood Restaurant

Figure 14. Cluster 10.

Conclusion

The area around clusters *2, 3, 4, 8 and 10* is a good candidate for a new restaurant given the initial conditions of finding an area which is not very crowded. *Cluster 3* looks like a good place for an International Cuisine restaurant, such as Italian, Argentinian or a Bistro, which are popular venues in Mexico City. The area around clusters *2, 4, 8 and 10*, already full of comfort food and informal restaurants, is a good candidate for a seafood, pizzas, burgers or a sandwich place.

Another way to look for the ideal area for a restaurant is to find a neighborhood and see its trends in the list of the top ten types of restaurants per neighborhood. For example, neighborhood *Anzures* looks already diverse with its mix of Mexican, Japanese, Italian, Spanish, Peruvian and Pakistani food. An International cuisine restaurant would be ideal in that area.

```
----Anzures----
                 venue  freq
0          Restaurant  0.19
1   Japanese Restaurant  0.12
2    Mexican Restaurant  0.12
3    Italian Restaurant  0.08
4          Taco Place  0.08
5    Spanish Restaurant  0.04
6         Pizza Place  0.04
7    Seafood Restaurant  0.04
8   Peruvian Restaurant  0.04
9  Pakistani Restaurant  0.04


----Anáhuac II Sección, Anáhuac I Sección----
                   venue  freq
0      Mexican Restaurant  0.33
1           Burger Joint  0.11
2             Restaurant  0.11
3             Taco Place  0.06
4     Japanese Restaurant  0.06
5             Food Court  0.06
6            Pizza Place  0.06
7      Seafood Restaurant  0.06
8  Comfort Food Restaurant  0.06
9          Sandwich Place  0.06


----Argentina Antigua----
                 venue  freq
0   Mexican Restaurant  0.44
1          Taco Place  0.25
2      Sandwich Place  0.06
3   Seafood Restaurant  0.06
4         Snack Place  0.06
5         Pizza Place  0.06
6          Restaurant  0.06
7    Paella Restaurant  0.00
8     Ramen Restaurant  0.00
9  Peruvian Restaurant  0.00
```

Figure 15. List of the 10 Most Common Venues in a Neighborhood (fragment).