IBM Applied Data Science Capstone Project

Introduction and Business Problem

Mexican cuisine was declared an intangible heritage of humanity by UNESCO in 2010. If you were to ask any Mexican what the best thing about Mexico is, chances are the vast majority will answer that it's the food. Mexican gastronomy is famous all over the world, yet it's almost impossible to get **real** Mexican food anywhere other than Mexico, a few restaurants in some southern USA states and a very limited number of Mexican-owned restaurants in other countries.

When talking to tourists or other visitors who just arrived to Mexic0 City, be it for work or leisure, it's universal: they always want to know what's the best place to get authentic Mexican Food. It shouldn't be a hard question. And depending on who you ask, you'll probably always get a different answer. And in a city of more than 20,000,000 habitants in 1,465 km², with thousands of authentic Mexican restaurants, the prospect can be daunting. That's why, for this project we will use the power of the data provided by Foursquare to process information about only the best options for authentic Mexican cuisine. We will create clusters of the best options for eating the best Mexican food.

Data

Most of the data will come form the Foursquare API, where we will be paying special attention to:

- Venue name and id
- Venue Category (Only Mexican restaurants, museums, monuments, historical places of interest and hotels will be considered).
- Venue rating
- Venue location
- Venue coordinates

Methodology

We processed these restaurants, and placed them in clusters. We will pay special attention to those of top quality (i.e.: most likes, since we have no access to ratings), close to city center. Why close to city center? Because if you are a tourist will limited time, that's where most historical sites are, and you can just walk to them.
We have already collected the info of the most liked Mexican restaurants and obtained their coordinates. Next, we will create clusters based on how great they are (top, middle bottom) Finally we will process and visualize these clusters to define the best options for eating authentic Mexican food, and see if we can find any other interesting conclusions.

Analysis

We will now arrange these restaurants in bins, from the very best, to the average to the bottom of the top 50.

This will require to get some stats and distribution information from the data frame.

We divided the data in 3 bins, using the percentiles in 33% and 66% to get the top, middle and bottom restaurants. We then proceeded to create our clusters for these bins using the K-Means algorithm. We then plotted them in the city map as can be seen below:
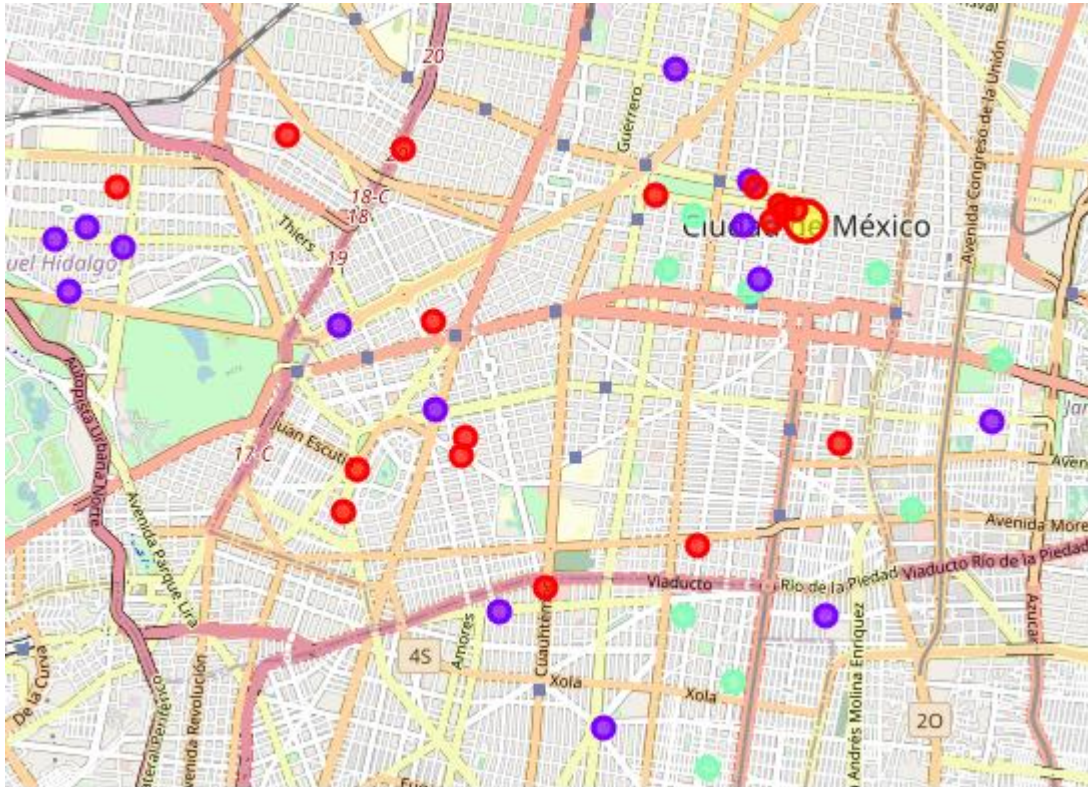
From the top cluster (id 0) we got the following restaurants as the top choices when visiting the city:

name	lat	lng	Likes Count	Bin	Cluster Label
La Casa de Toño	19.403356	-99.155597	6637	Top	0
La Casa de Toño	19.424880	-99.165188	5064	Top	0
La Polar	19.438798	-99.167759	3381	Top	0
Café de Tacuba	19.435682	-99.137591	3213	Top	0
Taquería El Abanico	19.414927	-99.130358	2818	Top	0
Azul Histórico	19.432880	-99.136197	2593	Top	0
El Cardenal	19.433746	-99.135216	2590	Top	0
La Casa de Toño	19.439875	-99.177742	2421	Top	0
El Parnita	19.414036	-99.162730	2043	Top	0
El Cardenal	19.434967	-99.146196	1872	Top	0
Azul Condesa	19.412896	-99.171659	1554	Top	0
Villa María	19.435656	-99.192375	1419	Top	0
Café El Popular	19.434142	-99.135460	1369	Top	0

Results and Discussion

From this data analysis, we can see that even though there are top quality authentic Mexican restaurants pretty much all over the city, there's quite a decent concentration of them **very near** the city center.

This provides a great reason why in this case it would be great to book a room in a hotel near the city center. Not only tourists are close to several great authentic Mexican restaurants, but by being there they are also a few blocks away from many of Mexico city historic landmarks. Also interesting, amongst the top restaurants, there are 3 from "La Casa de Toño" a restaurant very dear to Mexicans, since it started very small less than 20 years ago, but due to competitive prices, excellent service and great food as become extremely big and popular, without a drop in quality. Even if any visitors will wander away from the city center, they'll be able to easily find a Casa de Toño and get a great meal.

Conclusion

There is no shortage of great options for getting quality authentic Mexican food in Mexico City, as was to be expected before starting this project. However, we have found that for tourists the city center, or Zócalo, is a great option for eating, as it is for sight-seeing. Even though the stakeholders will be the ones to make the final decision, we can see from these clusters that wherever they choose to book a room, even if it's far from the city center, it's likely they'll have a quality options for dinner nearby, even though we only used 50 restaurants in this project.