**The Restaurant Battle of Neighborhoods in Mexico City**

Francisco J. Argüelles Vivas

October, 2021

**1.Introduction/Business Problem**

Mexico City is one of the top places to visit in the world [1]. It has many touristic places such as the historic center (designated a UNESCO World Heritage Site), colonial-era churches, El Templo Mayor (which preserves vestiges of the great Aztec empire), incredible museums, the Chapultepec Park, and many others [2].

Mexico City also possesses an extraordinary gastronomy, but for tourists locating the right place to eat can be complicated and challenging. Even for Mexicans in Mexico City, finding the right restaurant is a difficult task. Once, a friend of mine from India visited Mexico City for a conference. He asked me to show him the city, which I gladly did. Everything was going well until he told me he was a vegan, because I did not know where to find a restaurant for vegans! Mexican food is spicy, and the most delicious meals include beef or pork, which is a problem for many people around the world because of cultural, religious or health reasons.

Then, the purpose of this data science project is to help tourists in Mexico City to find restaurants of their preference. That is, in which neighborhoods of Mexico City they will find different options of restaurants (Fast food, Vegetarian food, Italian, Japanese and so on). Foreign tourists and also Mexicans are the target audience of this project.

**2.Source and description of the data**

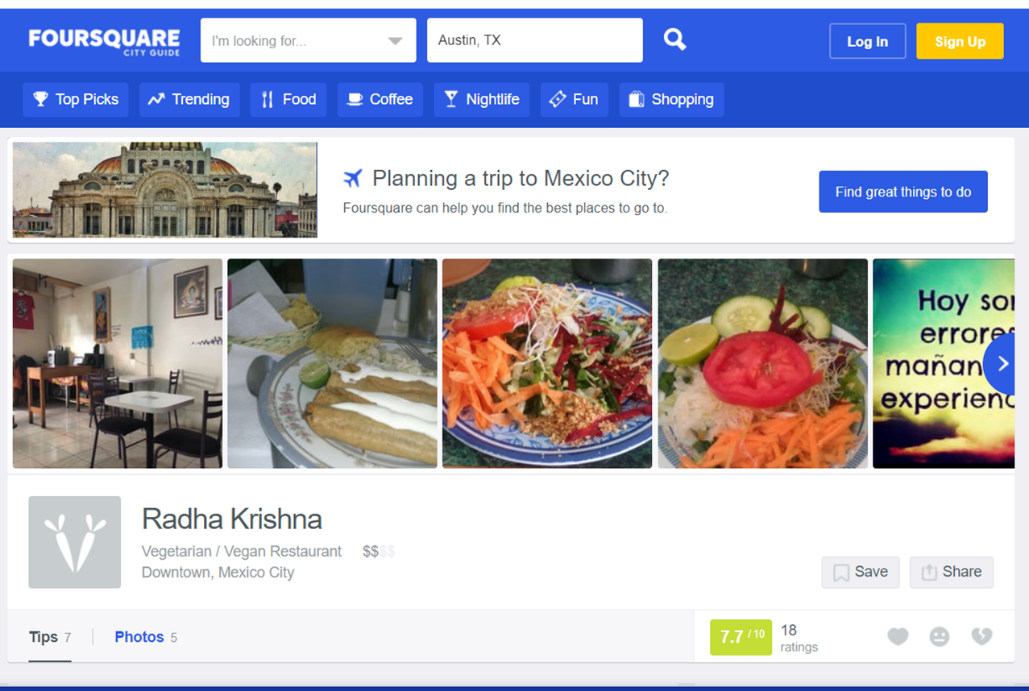
Postal codes from all boroughs and neighborhoods in Mexico City were taken from <https://micodigopostal.org/>. For example, Figure 1 shows some of the postal codes from neighborhoods from the borough “Cuauhtémoc” in Mexico City.



**Figure 1. Some postal codes from a borough in Mexico City called “Cuauhtémoc”. Asentamiento, Código Postal, Municipio, Ciudad mean neighborhood, postal code, borough and city, respectively; Ciudad de Mexico means Mexico City).**

Then, the postal codes from Mexico City were input in pgeocode, a python library that finds the GPS coordinates, region name, and municipality based on postal codes. After, the information from pgeocode was fed into Foursquare to search for different types of restaurants in the streets of Mexico-City neighborhoods. Figure 2 shows an example of a vegan/vegetarian restaurant in Mexico City: <https://foursquare.com/explore?mode=url&near=Mexico%20City%2C%20DF%2C%20Mexico&nearGeoId=72057594041458533&q=Radha%20Krishna>

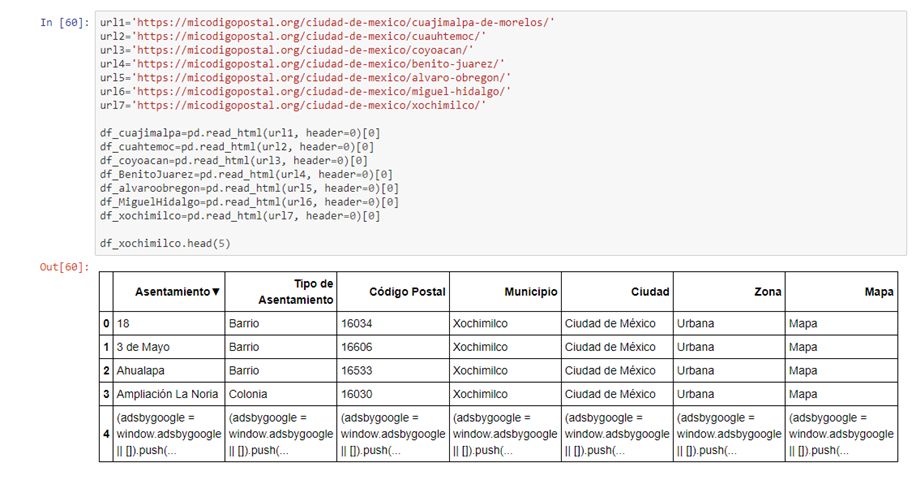
Finally, the data of restaurants in Mexico City were analyzed with the k-means clustering technique to give tourists suggestions of neighborhoods where to find the right restaurant for their taste.



**Figure 2. An example of a vegetarian/vegan restaurant in Mexico City shown in the website of Foursquare.**

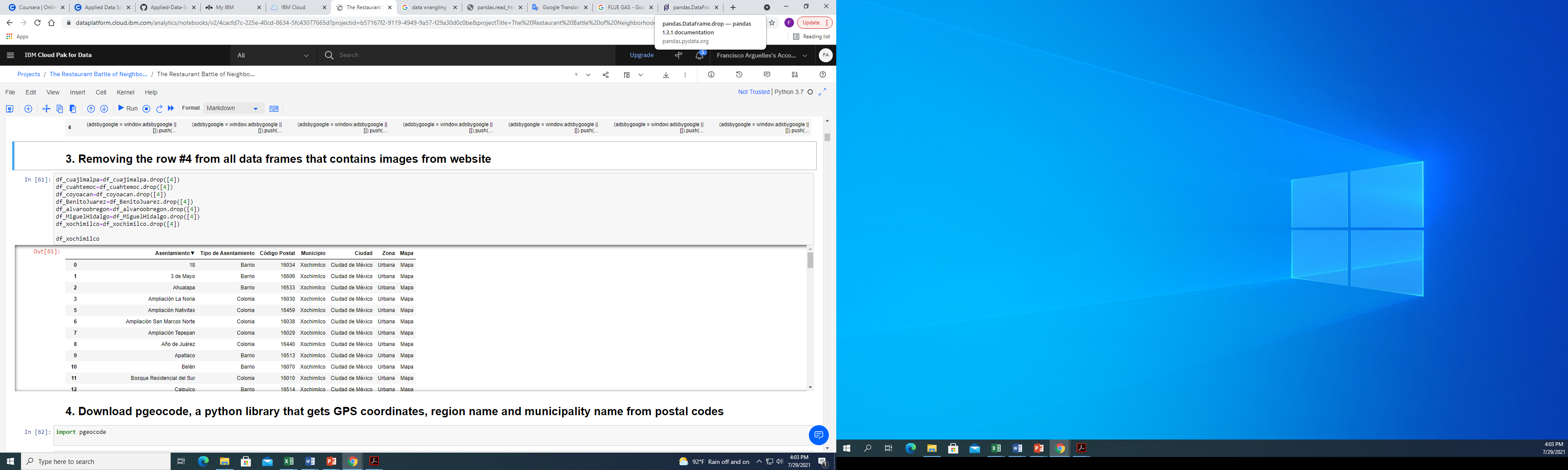
**3.Methodology**

Python was used in the IBM platform <https://myibm.ibm.com/dashboard/> to perform this project. First, Tables with the names and postal codes of boroughs and neighborhoods in Mexico City were obtained from <https://micodigopostal.org/>. Information of the seven most important boroughs in Mexico City was taken from this website: Cuajimalpa, Cuauhtémoc, Coyoacán, Benito Juárez, Alvaro Obregón, Miguel Hidalgo and Xochimilco. The pandas read function was used to read the data from <https://micodigopostal.org/>, as shown in **Figure 3.**

****

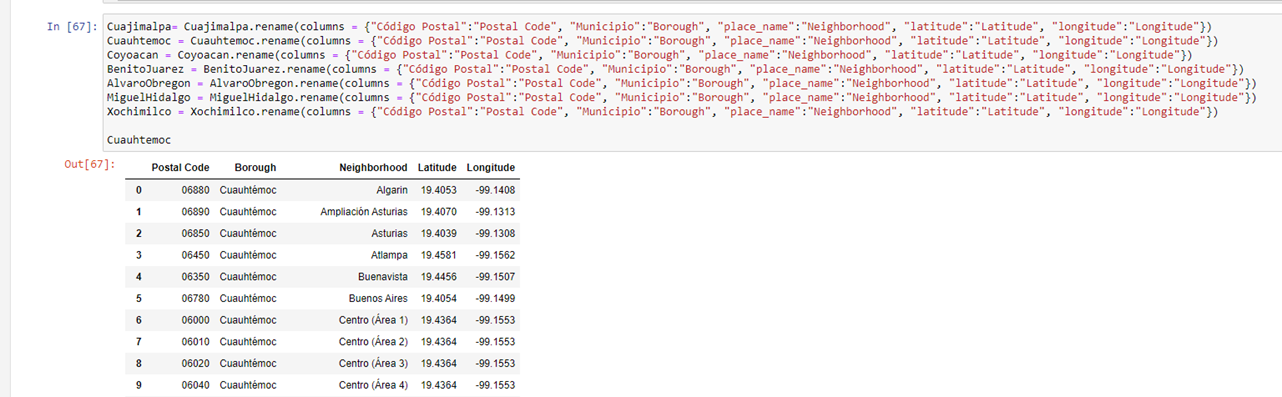
**Figure 3. Use of a pandas read function to get data from a website**

In Figure 3, the reader can see that the code of an image from the website was also store in in the row 4 of all data frames. To eliminate this row of all tables, the pandas drop function was used, as shown in **Figure 4**.



**Figure 4. Use of pandas drop function to eliminate an image code imported from a website from data frames.**

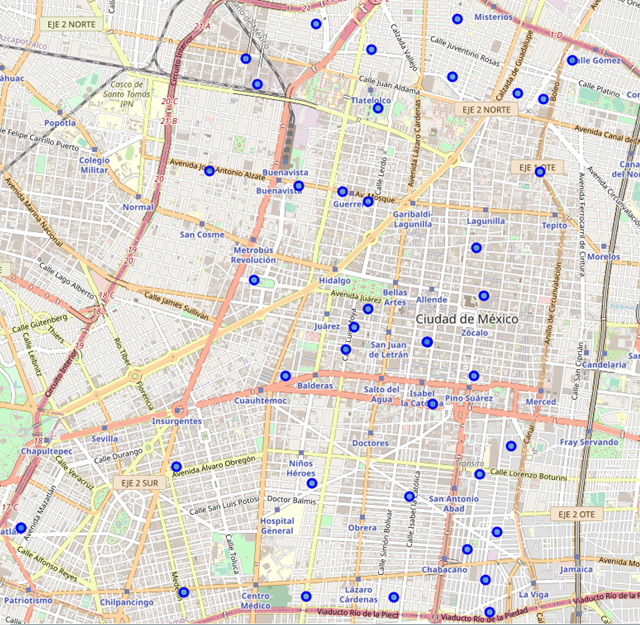
Then, the python library pgeocode was downloaded to get GPS coordinates, region name, and municipality through the postal codes of data frames shown in **Figure 3** (column called “código Postal”). After data cleaning and merging operations, new dataframes with relevant information of the boroughs and neighborhoods of Mexico City were created, as shown in **Figure 5** for Cuauhtémoc borough.

****

**Figure 5. Clean pandas dataframe of Borough Cuauhtémoc and its neighborhoods in Mexico City.**

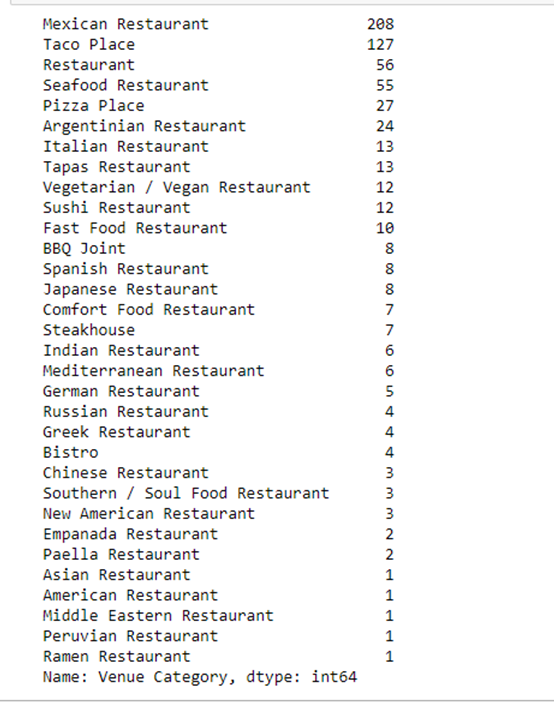
Actually, during the project execution it was found that the most important and relevant borough for tourists in Mexico City is Cuauhtémoc. Therefore, from here, we are focused only on the different types of restaurants on the neighborhoods of the Cuauhtémoc borough.

Then, **Figure 6** shows a map of the neighborhoods of Cuauhtémoc borough in Mexico City obtained with the folium package and the dataframe of Cuauhtémoc borough. It should be mentioned that, because pgeocode was unable to find all the neighborhoods of Cuauhtémoc borough, some of them were found manually using Google.



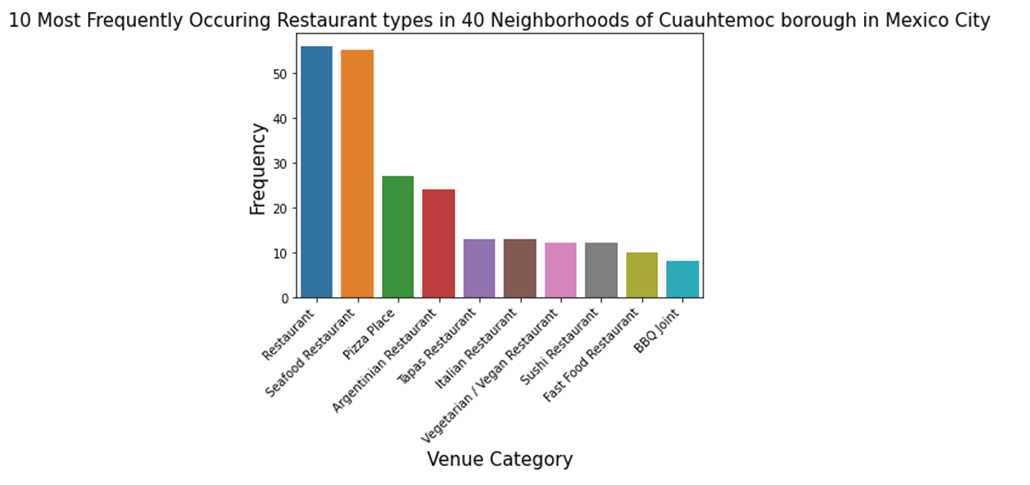
**Figure 6. Map of Cuauhtémoc** **borough created with Folium package. Blue dots are the neighborhoods.**

After, Foursquare was used to locate all the venues within a radius of less than 500 meters from the center of each neighborhood of Cuauhtémoc borough. 646 out of 2001 venues were restaurants. 31 types of restaurants were found, as shown in Figure 7. One of the first conclusions of this project was that Mexico City, and in particular the Cuauhtémoc borough, has too many Taco and Mexican restaurants, in comparison to other types of food. This is of course of common sense, but the proportion is disproportionate, in the opinion of the author of this project. 32% of restaurants are Mexican restaurants, and 19.6% are taco places, whereas, for example, Chinese restaurants only represent 0.46% of all restaurants. Therefore, Restaurants labeled as Mexican and Taco places were out of the data analysis since tourist can easily find them in the neighborhoods of Mexico City.



**Figure 7. Categories of restaurants found by Foursquare in the neighborhoods of Cuauhtémoc borough.**

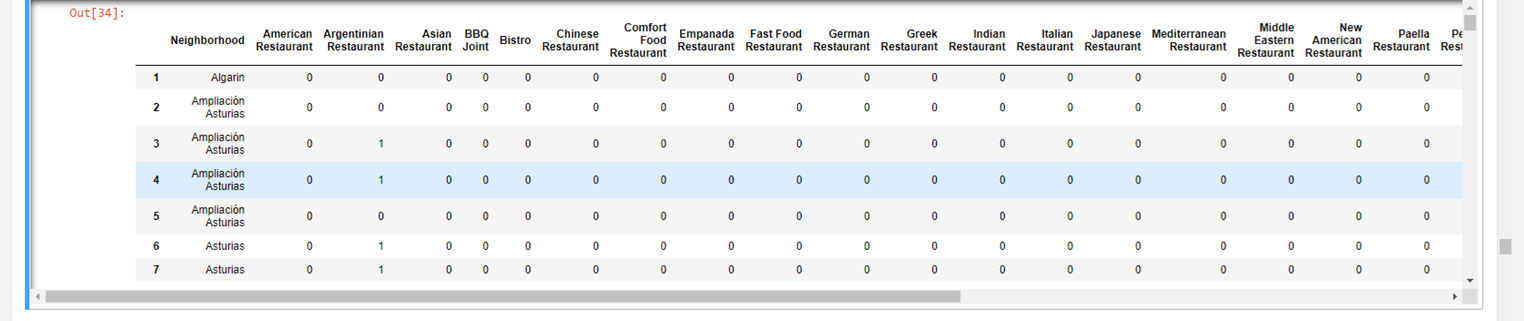
Then, a bar chart with the ten most common types of restaurants in the Cuauhtémoc borough of Mexico City was created with seaborn/matplotlib packages, as shown in the **Figure 8**.

****

**Figure 8. A plot chart with the ten most frequently occurring types of restaurants in the 40 neighborhoods of Cuauhtémoc borough in Mexico City.**

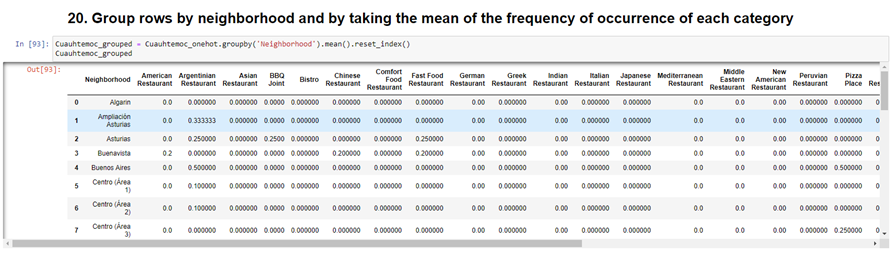
It must be clarified here that first bar labeled as “Restaurant” in Figure 8 referred to small and cheap restaurants run by families in the street of the neighborhoods (commonly known as “Fondas”). They are family businesses. In this type of restaurant, you find a variety of delicious and cheap Mexican dishes. After these restaurants, Seafood and then Pizza places and Argentinian restaurants are the most frequently occurring restaurants in the Cuauhtémoc borough of Mexico City.

Next step is to find clusters of restaurants by using k-means technique. To do so, first each neighborhood of Cuauhtémoc borough is analyzed by one-hot encoding, as shown in the **Figure 9**.



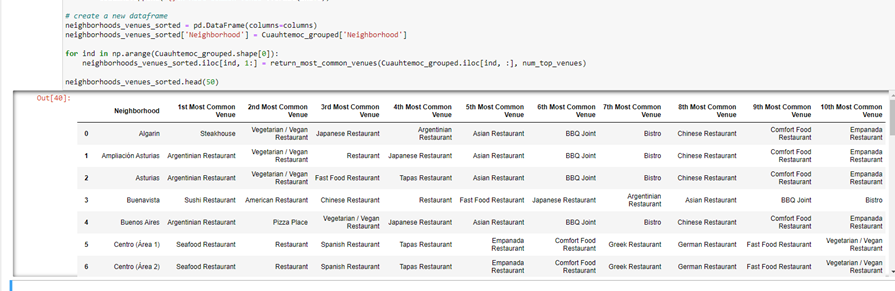
**Figure 9. Analysis of the occurrence of each type of restaurants in the neighborhoods of Cuauhtémoc borough by one-hot encoding (0/1).**

Then, the mean of the frequency of each category was determined for each neighborhood, as shown in the **Figure 10** for seven neighborhoods in the Cuauhtémoc borough.



**Figure 10. Frequency of each category of restaurants in each neighborhood of the Cuauhtémoc borough.**

After, a new pandas dataframe was created with the ten most common types of restaurants in each neighborhoods of Cuauhtémoc borough. As an example, **Figure 11** shows the 10 most common types of restaurants for six neighborhoods of the Cuauhtémoc borough.

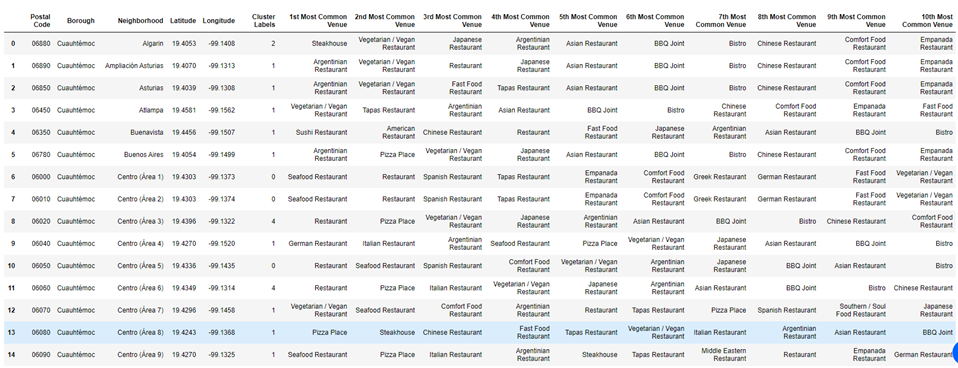
****

**Figure 11. The 10 most common types of restaurants in six neighborhoods of the Cuauhtémoc borough in Mexico City.**

Finally, the k-means clustering algorithm was applied to the dataframe shown in Figure 11 to obtain the restaurant clusters that will help tourists to find restaurants of their preference. A k value of 5 was chosen for the k-means technique.

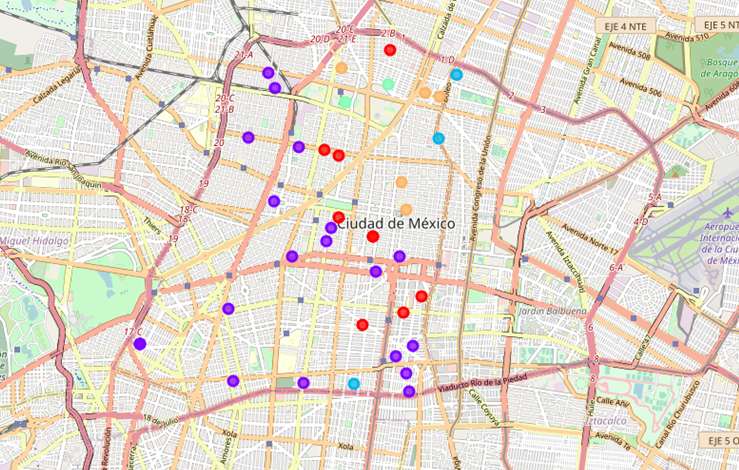
**4.Results**

**Figure 12** shows a pandas dataframe with the clusters of neighborhoods of the Cuauhtémoc borough and their most common types of restaurants, labeled from 0 to 4.

****

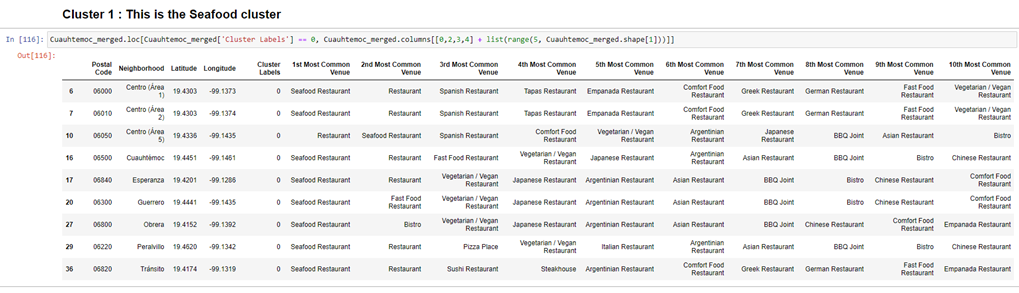
**Figure 12. Clusters of neighborhoods of the Cuauhtémoc borough and their most common types of restaurants, labeled from 0 to 4.**

**Figure 13** shows in colored circles the clusters of restaurants in the neighborhoods of the Cuauhtémoc borough in Mexico City. This map was built using Folium package.

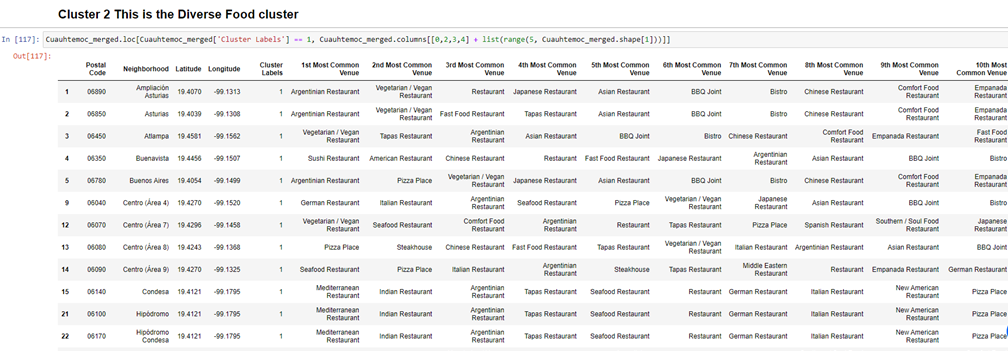
****

**Figure 13. Restaurant clusters in the neighborhoods of the Cuauhtémoc borough in Mexico City.**

Finally, **Figures 14 to 18** present the five clusters of restaurants located in the neighborhoods of the Cuauhtémoc borough in Mexico City. Cluster 2 is not fully visualized because of limited space in this report, but readers can find the whole cluster in the GitHub of the author: <https://github.com/PereiraSostiene?tab=repositories>



**Figure 14. The Seafood cluster.**



**Figure 15. The Diverse Food cluster.**



**Figure 16. The Fast Food/Vegetarian-Vegan cluster.**



**Figure 17. The Pizza cluster.**

****

**Figure 18. The Mexican Food cluster.**

**4.Discussion (Discussion section where you discuss any observations you noted and any recommendations you can make based on the results)**

As I mentioned above, the first observation I found was that in Mexico City is not easy to find diverse food because Mexican Food (Mexican dishes, Tacos and Fondas) is dominant in comparison to other types of Food. For example, Chinese restaurants represents only 0.46% of all restaurants in Mexico City, whereas Mexican Food represent 60%. In a way, this strengthens the development of this project, since we can tell tourists the specific street and neighborhoods where to find food other than Mexican one.

In this report I only proposed to form five food clusters, and one of them resulted as a “diverse food cluster”, which means that many types foods, that are not abundant, are together located in specifics neighborhoods of the Cuauhtémoc borough in Mexico City. It is probable that trying with more clusters we can have an “Argentinian cluster” or a “Mediterranean cluster.

I would suggest to entrepreneurs to invest in the restaurant industry in Mexico City. It seems a business opportunity to invest in restaurants offering foreign dishes.

**5.Conclusions**

A data science project was developed to help tourists in Mexico City to locate restaurant of their taste. Five clusters were found in the Cuauhtémoc borough (the most important one for tourism): the seafood, the diverse-food, the Fast Food/Vegetarian-Vegan, the pizza, and the Mexican-food clusters. The code can be easily modified to find restaurants in other boroughs of Mexico City.

**6. References**

1. <https://www.journeymexico.com/blog/mexico-city-rough-guides-2020>

2. <https://www.visitmexico.com/en/mexico-city>