



Analysis of Mexico City's Venues

By: Marco Xoca

Introduction

- Mexico City is the capital, largest city, and most populous city in North America. This city is also one of the most important and cultural centers in the world.
- By 2015 the estimated population was about 9,000,000, also Mexico City, or CDMX has 16 municipalities
- In this project I'll find the information about landmarks, restaurants, shops, and more venues of each municipalities, in order to help people to find the best municipality to visit according at their interest



Data

In order to get all the information needed, we can list the data as below:

- List of each municipality in CDMX
- Coordinates (latitude and longitude) of each municipality in CDMX
- List of every venue by municipalities (obtained through Foursquare API)

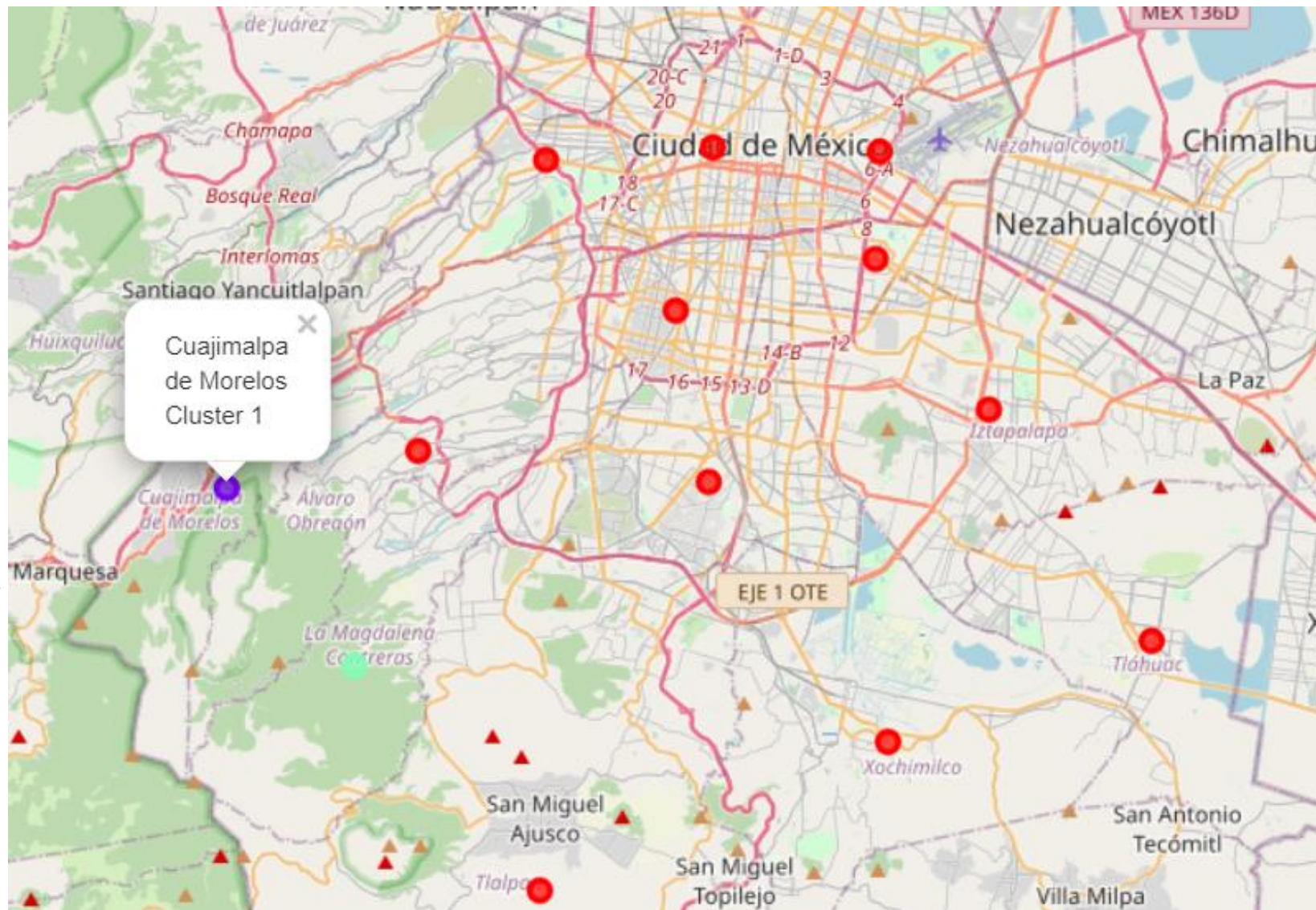
After the preprocessing and cleaning, the dataframe was the one in the right.

| | borough | population | area | density | postal code | Latitude | Longitude |
|----|-----------------------|------------|----------|---------|---------------|-----------|------------|
| 0 | Azcapotzalco | 414711 | 33.66 | 12.635 | 02000 - 02999 | 19.485329 | -99.182107 |
| 1 | Benito Juárez | 385439 | 26.63 | 13.331 | 03000 - 03999 | 19.380642 | -99.161135 |
| 2 | Coyoacán | 620416 | 54.4 | 11.545 | 04000 - 04999 | 19.326667 | -99.150376 |
| 3 | Cuajimalpa de Morelos | 186391 | 74.58 | 2.328 | 05000 - 05999 | 19.324634 | -99.310729 |
| 4 | Cuauhtémoc | 531831 | 32.4 | 16.071 | 06000 - 06999 | 19.431373 | -99.149056 |
| 5 | Gustavo A. Madero | 1185772 | 94.07 | 12.683 | 07000 - 07999 | 19.504065 | -99.115864 |
| 6 | Iztacalco | 384326 | 23.3 | 16.953 | 08000 - 08999 | 19.396912 | -99.094330 |
| 7 | Iztapalapa | 1815786 | 117 | 15.563 | 09000 - 09999 | 19.349166 | -99.056799 |
| 8 | Magdalena Contreras | 239086 | 74.58 | 3.069 | 10000 - 10999 | 19.268977 | -99.268413 |
| 9 | Miguel Hidalgo | 372889 | 46.99 | 7.523 | 11000 - 11999 | 19.428062 | -99.204567 |
| 10 | Milpa Alta | 130582 | 228.41 | 507 | 12000 - 12999 | 19.139457 | -99.051095 |
| 11 | Tlalpan | 650567 | 312 | 2.085 | 14000 - 14999 | 19.198340 | -99.206221 |
| 12 | Tláhuac | 360265 | 85.34[9] | 4.032 | 13000 - 13999 | 19.276998 | -99.002822 |
| 13 | Venustiano Carranza | 430978 | 33.4 | 13.396 | 15000 - 15999 | 19.430495 | -99.093106 |
| 14 | Xochimilco | 415007 | 122 | 3.427 | 16000 - 16999 | 19.245145 | -99.090364 |
| 15 | Álvaro Obregón | 727034 | 96.17 | 7.347 | 01000 - 01999 | 19.336176 | -99.246820 |

Methodology

I got the geographical coordinates of Mexico City, which were:
longitude = 19.4326296 and
latitude = -99.1331785

With these coordinates, and using Folium's package, I made a map of the CDMX, with labels at each municipality.



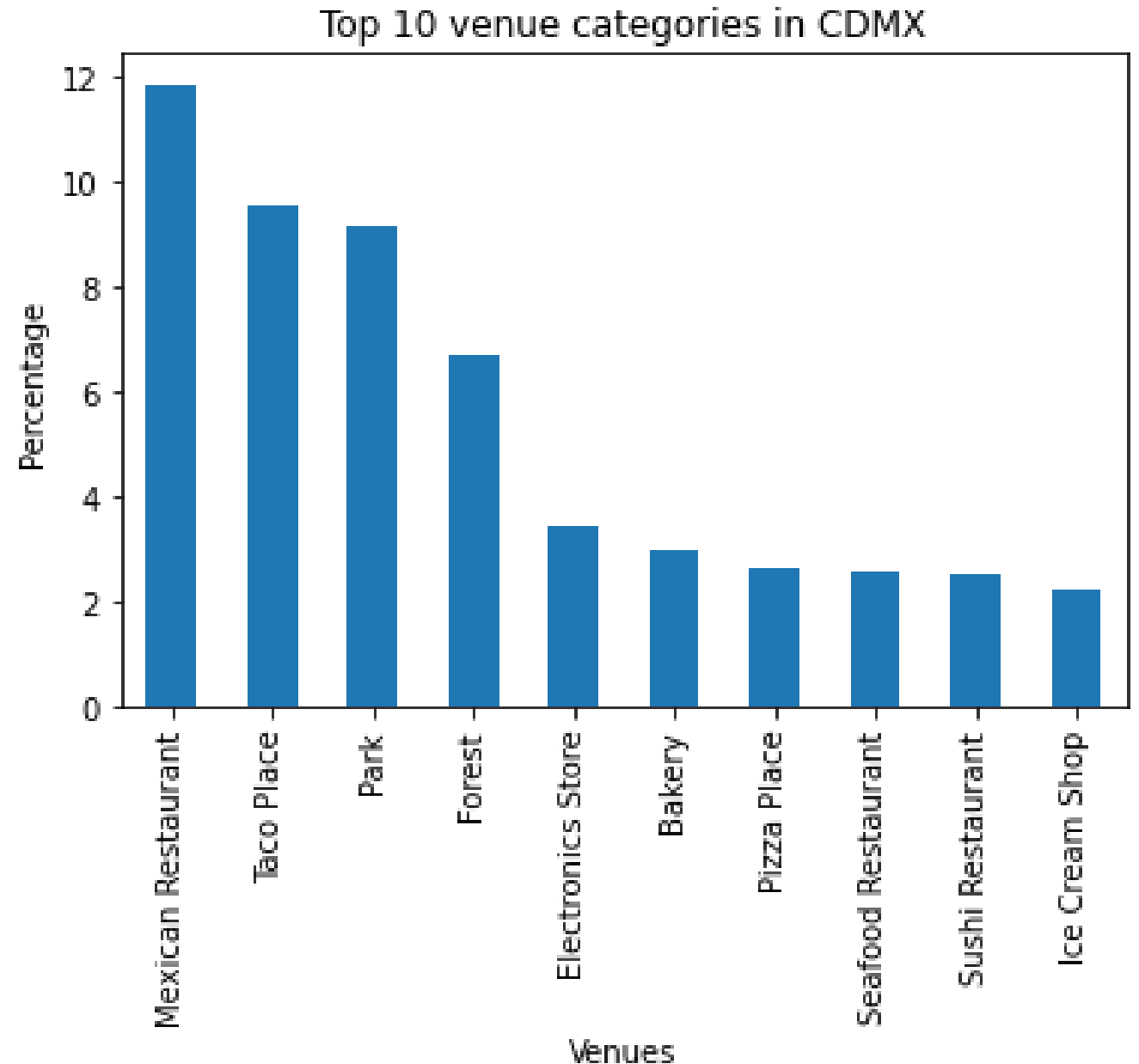
Methodology

Once having the CDMX's map, with the help of Foursquare API I got all the venues available by municipality (neighborhood), and the top 10 venue category by municipality

| | Neighborhood | 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 |
|---|-----------------------|-----------------------|-----------------------|-------------------------|-----------------------|---------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|
| 0 | Azcapotzalco | Mexican Restaurant | Ice Cream Shop | Taco Place | Seafood Restaurant | Breakfast Spot | Burrito Place | Bakery | Food Court | Japanese Restaurant | Garden |
| 1 | Benito Juárez | Pizza Place | Mexican Restaurant | Coffee Shop | Pet Store | Southern / Soul Food Restaurant | Greek Restaurant | IT Services | Ice Cream Shop | Dog Run | Deli / Bodega |
| 2 | Coyoacán | Taco Place | Seafood Restaurant | Ice Cream Shop | Gym / Fitness Center | Fast Food Restaurant | Park | Coffee Shop | Mexican Restaurant | Pizza Place | Pool |
| 3 | Cuajimalpa de Morelos | Park | Women's Store | Comfort Food Restaurant | Convenience Store | Cupcake Shop | Cycle Studio | Deli / Bodega | Design Studio | Dessert Shop | Diner |
| 4 | Cuauhtémoc | Mexican Restaurant | Taco Place | Deli / Bodega | Hotel | Coffee Shop | Art Gallery | Bar | Restaurant | Café | Bakery |

Methodology

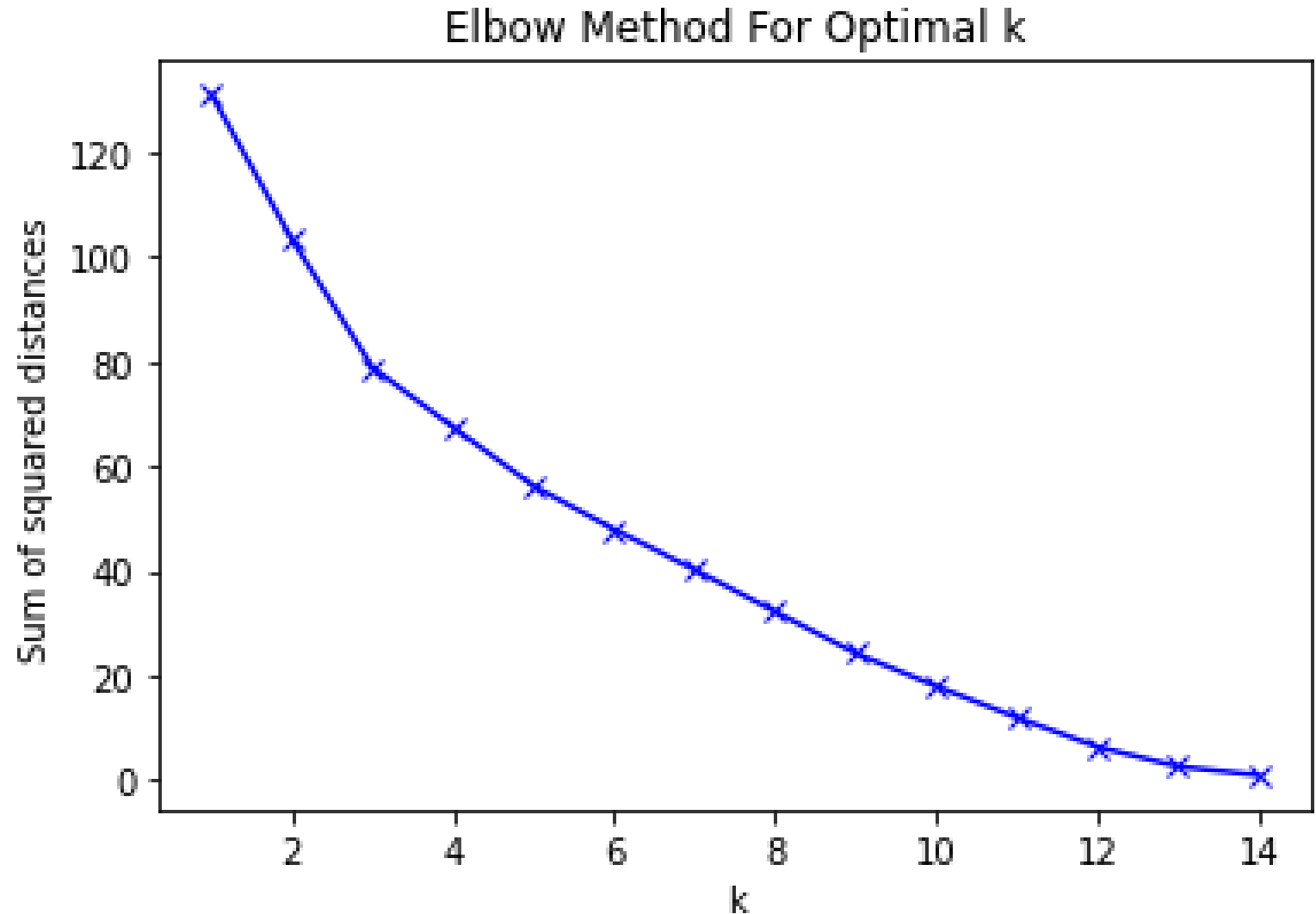
Using the one-hot encoding, and with the frequency dataframe, I made the follow dataframe with the top 10 most common venue in all the city



Clustering: Elbow's method

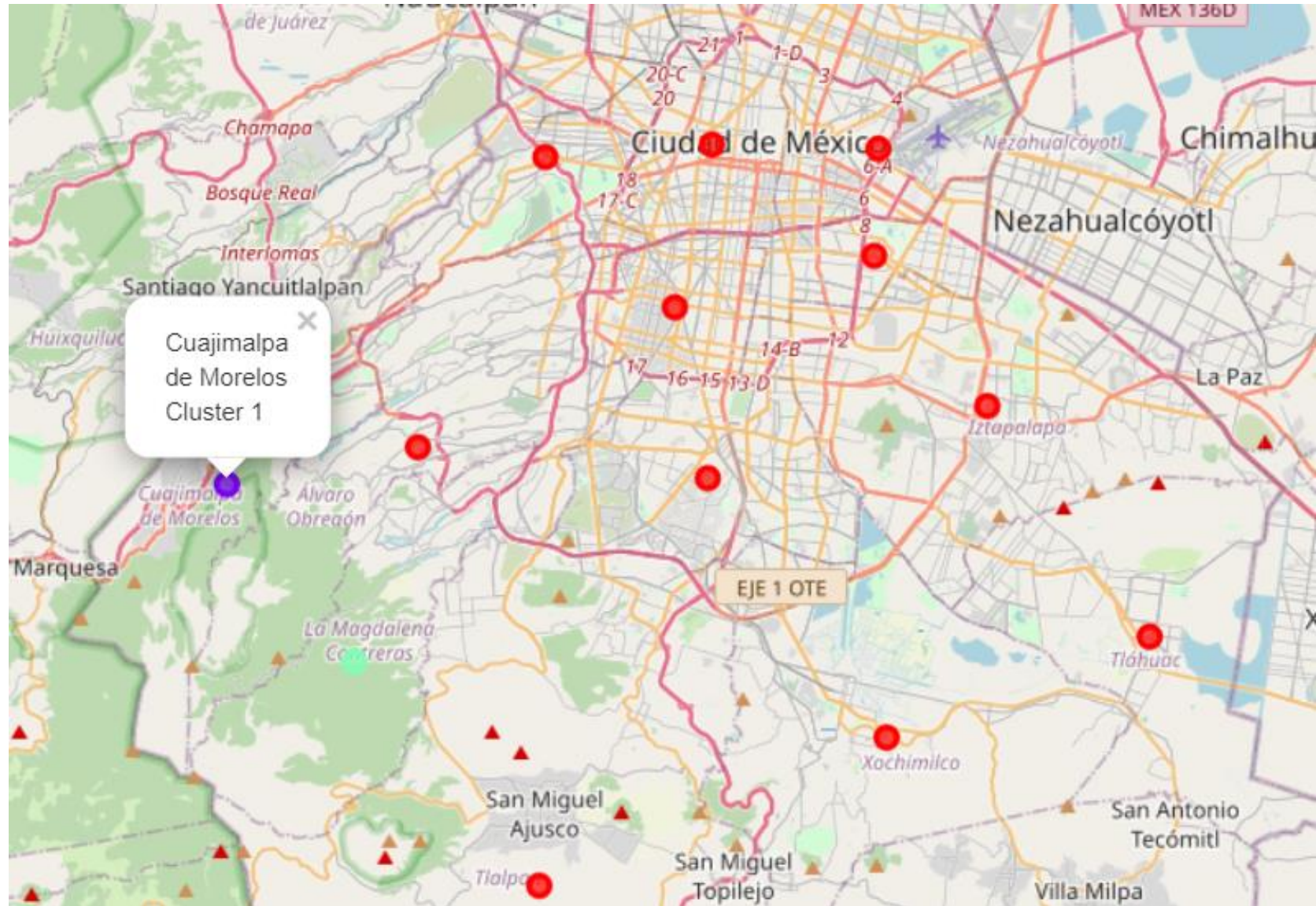
One step just before to start clustering, is to know how many clusters (or the k value) are the optimal for this problem.

I made a plot of the sum of squared distances with different k -values



Clustering: with $k=3$

Using the KMeans library, I divided all the venues data in 3 clusters.



Results

Seeing all the tables and maps, we can notice that almost all the venues are just in one cluster, which has a big relation in the urbanized area

Also, we can explore venues by cluster, like the venues in cluster 1 order by common

| | population | Longitude | Cluster Labels | 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 |
|---|------------|------------|----------------|-----------------------|-----------------------|-----------------------|-----------------------|---------------------------------|-----------------------|-----------------------|-----------------------|
| 0 | 414711 | -99.182107 | 0 | Mexican Restaurant | Ice Cream Shop | Taco Place | Seafood Restaurant | Breakfast Spot | Burrito Place | Bakery | Food Court |
| 1 | 385439 | -99.161135 | 0 | Pizza Place | Mexican Restaurant | Coffee Shop | Pet Store | Southern / Soul Food Restaurant | Greek Restaurant | IT Services | Ice Cream Shop |
| 2 | 620416 | -99.150376 | 0 | Taco Place | Seafood Restaurant | Ice Cream Shop | Gym / Fitness Center | Fast Food Restaurant | Park | Coffee Shop | Mexican Restaurant |
| 4 | 531831 | -99.149056 | 0 | Mexican Restaurant | Taco Place | Deli / Bodega | Hotel | Coffee Shop | Art Gallery | Bar | Restaurant |
| 5 | 1185772 | -99.115864 | 0 | Performing Arts Venue | Hot Dog Joint | Fast Food Restaurant | Scenic Lookout | Athletics & Sports | Movie Theater | Park | Shopping Mall |

Results

Cluster 2 and 3 are respectively:

| | population | Longitude | Cluster Labels | 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 |
|---|------------|------------|----------------|-----------------------|-----------------------|-------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|
| 3 | 186391 | -99.310729 | 1 | Park | Women's Store | Comfort Food Restaurant | Convenience Store | Cupcake Shop | Cycle Studio | Deli / Bodega | Design Studio | Dessert Shop | Diner |

| | population | Longitude | Cluster Labels | 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 |
|---|------------|------------|----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|
| 8 | 239086 | -99.268413 | 2 | Forest | Women's Store | Fast Food Restaurant | Convenience Store | Cupcake Shop | Cycle Studio | Deli / Bodega | Design Studio | Dessert Shop | Diner |

Discussions

As we can see in the last three figures, if we like to eat Mexican food, pizza or go to arts venue, we should go in the cluster 1 areas. Otherwise, if we like to visit parks, we should go to cluster 2 areas. Finally, if we like to go to the forest, we should go to cluster's 3 areas.

We can notice that in all CDMX, the most common venues are, obviously, Mexican restaurants, followed by Taco places, parks and forest, being such a good place to free activities and good food places.

Being Mexico City one of the bigger cities in the world, probably it is better to analyze venues by municipality. Also, foursquare has not a lot of data of CDMX's venues, so it would be a great idea using another venues API, like the Google's one, or a local government API to get all the necessary data to a better analysis



Conclusion

With all this analysis we can conclude that, according to our likes, we can visit different municipalities of CDMX.

If we like all kind of restaurants, we should go to any area in cluster 1, if we like to go to a park or shopping, we should go to any area in cluster 2, showed by the violet label.

Finally, if we like to go to a forest, shopping, or food restaurant, we should go to any area in cluster 3.

This kind of analysis is very useful if we like to visit a new city, but we have no idea what to expect, and with this we can cluster our favorite places and planning in a better way our travel.

Also, this analysis is very straightforward, and you don't need to be the grater programmer of all time in order to do a very well-made analysis and clustering of venues

