

# Open a Mexican Restaurant



IBM Applied Data Science Specialization

Capstone Project

Author: Norma Ruiz

# Definition of the problem

- I am going to open another Mexican Restaurant(Oaxacan Food) in Mexico City.
- I want to find neighbourhoods in specific districts of Mexico City that have the same places of interest (venues) as those in the neighbourhood where my first restaurant is located.

## Ciudad de México

División delegacional



Current  
neighbourhood:  
**Benito Juárez**  
Mexico City

# Candidate neighbourhoods



Cuauhtemoc



Coyoacan



Benito Juarez



Miguel Hidalgo

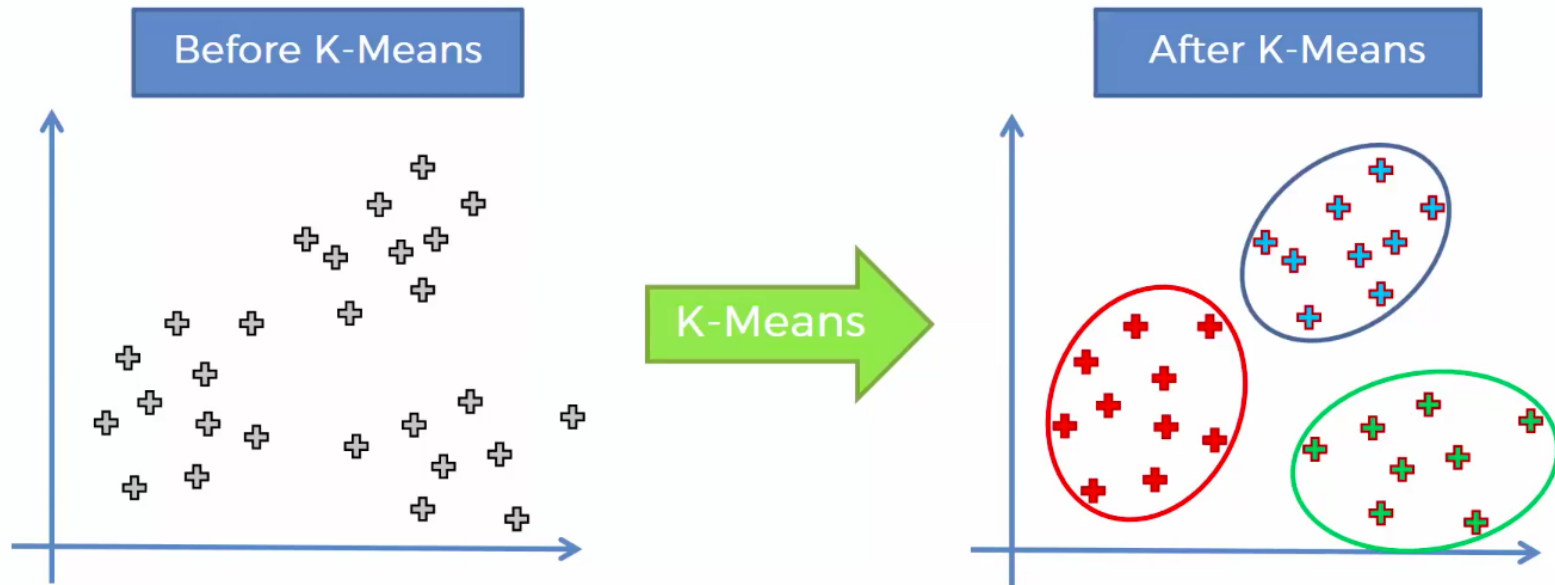
# Objective

- Find neighbourhoods similar to another one that serves as a point of comparison.
- In this case, the similarity is based on the venues in the neighbourhoods, according to the information provided by the Foursquare API.

# Steps followed

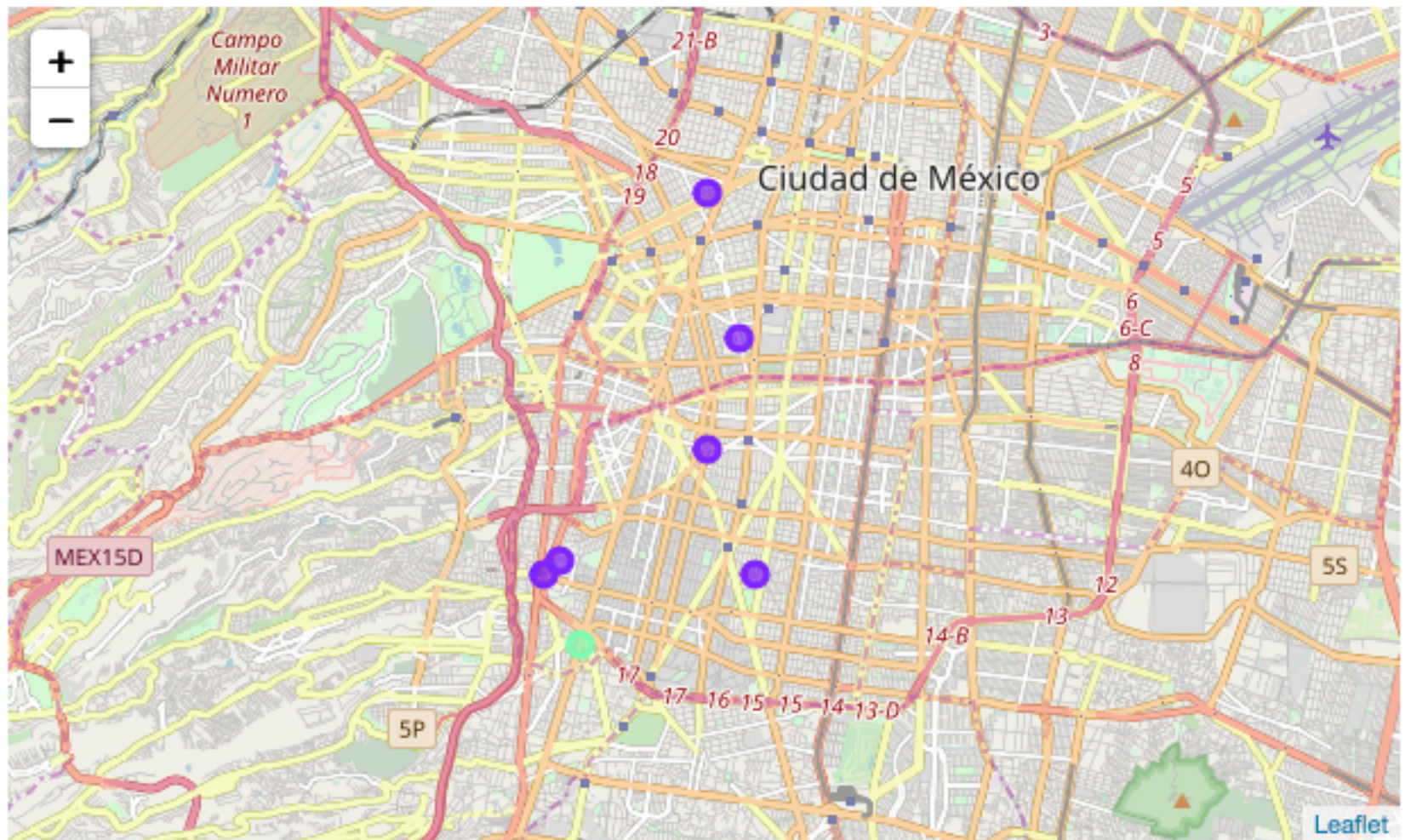
- Use Foursquare API to collect venues from current neighbourhood
  - Result: 100 rows for 1 neighbourhood
- Use Foursquare API to collect venues from 4 districts of Mexico City
  - Result: 11231 rows for 257 neighbourhoods
- Join both results and transform “venue category” using onehot-encoding
  - Result: 272 unique categories
- Group “venue categories” for each neighbourhood using the mean
- Use KMEANS to create clusters of similar neighbourhoods
- Find the closest neighbourhoods to the current neighbourhood using distance to centroid from KMEANS

# KMEANS Algorithm





# Similar neighbourhoods





# Conclusions

- It is possible to help people who are in a situation similar to the one described in this case, using public data available through the Foursquare API.
- The final decision cannot be based solely on the results of this analysis.
  - Rather it should be considered as a tool to narrow the options that must be investigated in greater detail.
- One way to enrich the results of the analysis would be by adding demographic and socioeconomic attributes to each of the neighbourhoods.