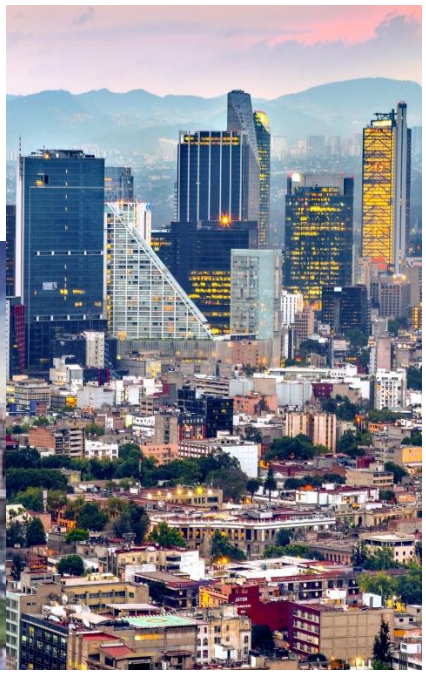
IBM Applied Data Science Project

Yhael Jacinto

Table of Contents

1. Introduction

Mexico City is the capital city of Mexico. It has a population of around 20 million people, which means your visit is going to be crowded. The weather is subtropical with temperatures ranging around 15 degrees and, at times, rainy. It is a good idea to visit the webpage for National Meteorological Service for better weather predictions. With an ALTITUDE of 2,240 metres people may suffer dizziness or headaches. Importantly, remember to enjoy the experience of visiting one of the biggest cities in the world and open your hearts and minds to the Mexican people. Moreover, that even though the country may suffer problems, Mexicans will try to help you in any way they can. Mexico City is big, crowded and suffers traffic congestion. The congress will be taking place is one of the most beautiful places in the city, yet it is occupied by tourists from every corner of the world. The Angel de la Independencia roundabout is also a hotbed for political and social protests, demonstrations, parades and several different types of events – be careful. In addition, Av. Paseo de la Reforma is the financial centre of the country, therefore their hours of operation are good to keep in mind.



2. Business Problem

The aim is to help tourists choose their destinations depending on the experiences that the neighbourhoods have to offer and what they would want to have. This also helps people make decisions if they are thinking about migrating to Mexico City, especifically the downtown, or even if they want to relocate neighbourhoods within the city. Our findings will help stakeholders make informed decisions and address any concerns they have including the different types of food, such as tacos, mole, gorditas, tlacoyos, sopes, tortas, tamales; museums, provision stores and what the city has to offer.

3. Data

In order to solve such a business problem, the following data will be required:

- List of neighbourhoods in Mexico City
- Description: Contains a list of all neighbourhoods in Mexico with their associated postal codes and boroughs
- Collection method: Web scraping using BeautifulSoup package
- Values of latitude and longitude of the associated neighbourhoods
- Description: Needed to specify the neighbourhoods' locations in order to further interact with Foursquare API
- ARCGIS APPI AND FOURSQUARE API DATA

Location data regarding venues that are present within those neighbourhoods

- Description: Consists of all venues in the neighbourhoods of Mexico
- Source: Foursquare database
- ARCGIS APPI AND FOURSQUARE API DATA
- Folium is a library that will allow me to build maps to better visualize these neighbourhoods. Lastly, I used an article from
    - StreetEasy that listed the top 6 most affordable neighbourhoods in Mexico.

4. Methodology

Data collection

First and foremost in order to start the project, the data mentioned above have to be collected accordingly. By using the BeautifulSoup package, it allows us to web scrape the necessary table that contains the data of Mexico's neighbourhoods from Wikipedia in the form of HTML data. Combining it with Pandas function to read HTML data, we can then turn it into a dataframe to be further prepared in Python.

To obtain the values of latitude and longitude of all the neighbourhoods in Mexico (which is needed to interact with the Foursquare API), the GeoSpace data of London is required.

Last but not least, location data regarding venues present within the Mexico and London neighbourhoods are to be collected using the Foursquare API. By defining our credentials (client ID and client secret), we can interact with the Foursquare API to gain access to their location data – in which we can select the location data of Mexico's neighbourhoods by including the previously mentioned latitudes and longitudes as parameters of the call request to the API.

Data wrangling/preparation

The dataframe of Toronto's neighbourhoods had missing values, whereby some postal codes had boroughs and neighbourhoods that were not assigned to them. As a result, rows with missing borough values are dropped, and neighbourhoods with missing values are replaced with the same values as the borough in their respective rows. The latitude and longitude values of all the neighbourhoods were also initially imported as a separate dataframe, so both dataframes were also merged with postal codes being the column to join based on. Afterwards, the location data obtained from using the

Foursquare API were also merged by grouping them via the neighbourhoods and by taking the mean of the frequency of occurrence of each venue category. To prepare for further data analysis, the data set was also further filtered to only contain the neighbourhoods and their associated mean frequencies of Chinese restaurants.

Analysis

It is important to first explain why clustering is an appropriate machine learning model to be used in formulating a solution for the aforementioned business problem. Clustering is an unsupervised machine learning model that groups a set of data points in such a way that data points within the same groups (also called as clusters) are more similar to each other compared to objects present in other clusters. This way, with the context of the business problem at hand, clustering can be used to group NeyYork neighbourhoods which have similar attributes.

5. Results

The neighbourhoods of Mexico are very multicultural. There are a lot of different cusines including the Mexican one, Indian, Brazilian, Italian, Turkish, Chineses and of course, from other parts of the country such as Puebla, Yucatan, Guadalajara. Mexico seems to take a step further in this direction by having a lot of Restaurants, bars, juice bars, coffee shops, Fish and Chips shop and Breakfast spots. It has a lot of shopping options too with that of the Flea markets, flower shops, fish markets, Fishing stores, clothing stores. The main modes of transport seem to be Buses and trains. For leisure, the neighbourhoods are set up to have lots of parks, golf courses, zoo, gyms and Historic sites. Overall, the city of Mexico offers a multicultural, diverse and certainly an entertaining experience.

England/London is relatively small in size geographically. There are a lot of hangout spots including many Restaurants and Bars. Mexico has a lot of Bistro's. Different means of public transport in London which includes buses, bikes, boats or ferries. For leisure and sight seeing, there are a lot of Plazas, Trails, Parks, Historic sites, clothing shops, Art galleries and Museums. Overall, London seems like the relaxing vacation spot with a mix of lakes, historic spots and a wide variety of cusines to try out.

6. Conclusion

The purpose of this project was to explore the cities of Meixco and London and see how attractive it is to potential tourists and migrants. We explored both the cities based on their postal codes and then extrapolated the common venues present in each of the neighbourhoods finally concluding with clustering similar neighbourhoods together.

Mexico City is the capital city of Mexico and one of the most important political, cultural, educational and financial centres in North America. Mexico City is the country's largest city as well as its most important political, cultural, educational and financial center.

Also Mexico City, Mexico's largest city and the most populous metropolitan area in the Western Hemisphere and can be compared to some of China big cities

We could see that each of the neighbourhoods in both the cities have a wide variety of experiences to offer which is unique in it's own way. The cultural diversity is quite evident which also gives the feeling of a sense of inclusion.

Both Mexico and London seem to offer a vacation stay or a romantic getaway with a lot of places to explore, beautiful landscapes, amazing food and a wide variety of culture. Overall, it's upto the stakeholders to decide which experience they would prefer more and which would more to their liking.