

Exploring Real Estate Prices, Transportation, and Entertainment in Mexico City Boroughs with Python

Lauren Juskelis

June 17, 2019

1. Introduction

1.1 Background

Mexico City is one of the largest cities in the world, with a population of over 8 million people living within the city limits. The city is home to one of the most heavily used metro systems in the world, officially called the Sistema de Transporte Colectivo, that services more than 1.5 billion passengers per year. As the oldest capital city in the Americas, Mexico City has both world-class museums and archeological sites as well as all the amenities of a modern metropolitan center, from restaurants to city parks to night life. All of this leads the city to be quite overwhelming for new residents and businesses to make sense of.

1.2 Problem

As an expat living in Mexico City, I'm interested in developing a way for new businesses and residents to better understand the city in order to ultimately decide on where to establish themselves. New residents might be interested in finding an affordable location with plenty of transportation and things to do. New businesses might want to be close to transportation and other popular destinations within the city while minimizing their cost of operation. By using the available data as explained below, I develop an interactive map including information on transportation, real estate prices, and popular destinations.

1.3 Interest

Both businesses and individual residents may be interested in comparing real estate prices, access to low-cost transportation, and the availability of entertainment between different boroughs.

Businesses often have a target audience in mind, and could use this tool to identify where they should open up a new store or restaurant. For example, a high-end restaurant will want to set up in an area of town with similarly expensive options, ideally near other forms of entertainment like parks or theaters, and not too close to public transportation. On the other hand, a fast food chain might want to open a new branch near public transportation in a location where there are limited other food options.

Individuals looking to move cities or boroughs will likely have a housing budget in mind and want to have some entertainment options nearby. Depending on whether or not the person or family owns a car or uses public transportation, the household may want to be close to public transportation or be far from it to avoid congestion. This tool will help prospective residents identify the borough that might be best suited for their lifestyle.

2. Data Description and Source

GPS locations and names of metro stops were found through the Mexico City government's 'explore' feature [here](#). The data was already cleaned and contains the names, latitude, longitude, and metro line associated with each metro stop.

The geographic boundaries of each borough were found on the same Mexico City government website [here](#). The data was already cleaned and contained the name, municipal code, and shape as a series of GPS coordinates for each borough.

Raw data for median housing prices by borough were found on the Propiedades.com website [here](#). Data was searched for manually for each borough within Mexico City and then saved as a .csv file containing borough name and the corresponding housing value. Housing value was measured as the median cost of all single-family homes for sale by borough advertised on the Propiedades.com page on May 28, 2019.

Venue data was found using Foursquare's API accessed through the developer portal [here](#). By using the explore endpoint and the near and limit parameters, I was able to find the top 15 venues in each borough. This approach returned venues within a borough rather than venues near the geographical center of the borough, resulting in more appropriate findings.

3. Methodology

3.1 Set Up

Before all else, I began by installing and importing the relevant libraries:

```
import sys
import pandas as pd
import numpy as np

!{sys.executable} -m pip install requests
!{sys.executable} -m pip install geocoder
!{sys.executable} -m pip install geopy
!{sys.executable} -m pip install folium

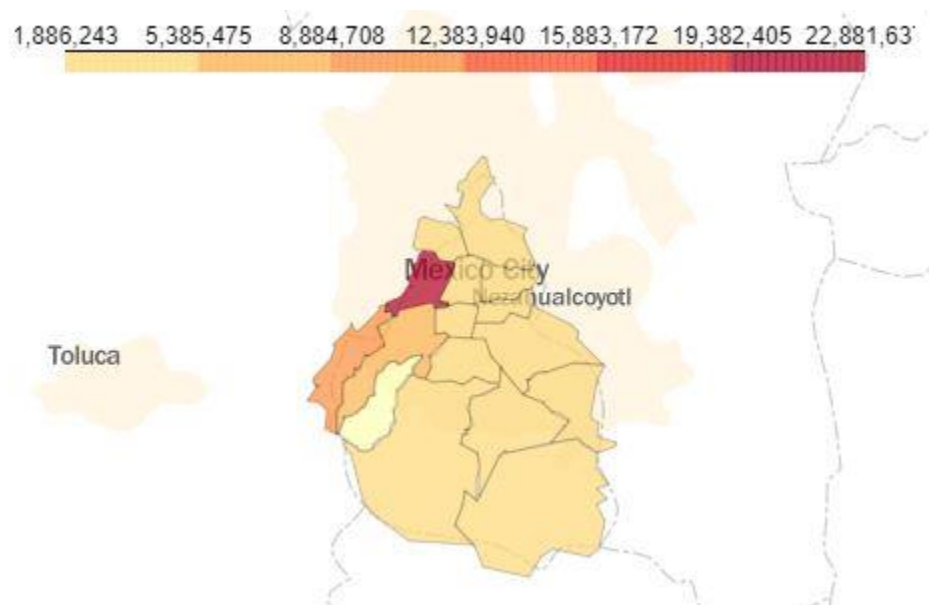
import requests
from geopy.geocoders import Nominatim
import folium
import matplotlib.colors as colors
import matplotlib.cm as cm
```

3.2 Housing data

Then I imported the .csv file containing median housing costs by borough and saved the result as a pandas dataframe:

	Borough	Median price
0	Tlalpan	4834600
1	Tláhuac	2092080
2	Venustiano Carranza	2662340
3	Cuauhtémoc	4208070
4	Iztapalapa	2226720
5	Azcapotzalco	2881370
6	Benito Juárez	5188090
7	Cuajimalpa de Morelos	11939500
8	Iztacalco	2735320
9	Gustavo A. Madero	2660400
10	Miguel Hidalgo	22675800
11	Magdalena Contreras	6029360
12	Coyoacán	4970550
13	Milpa Alta	2593840
14	Álvaro Obregón	8639000
15	Xochimilco	3349760

After opening the .txt file containing GPS coordinates defining the boundaries of each borough, I was able to make a choropleth map of the boroughs within Mexico City colored by housing cost:

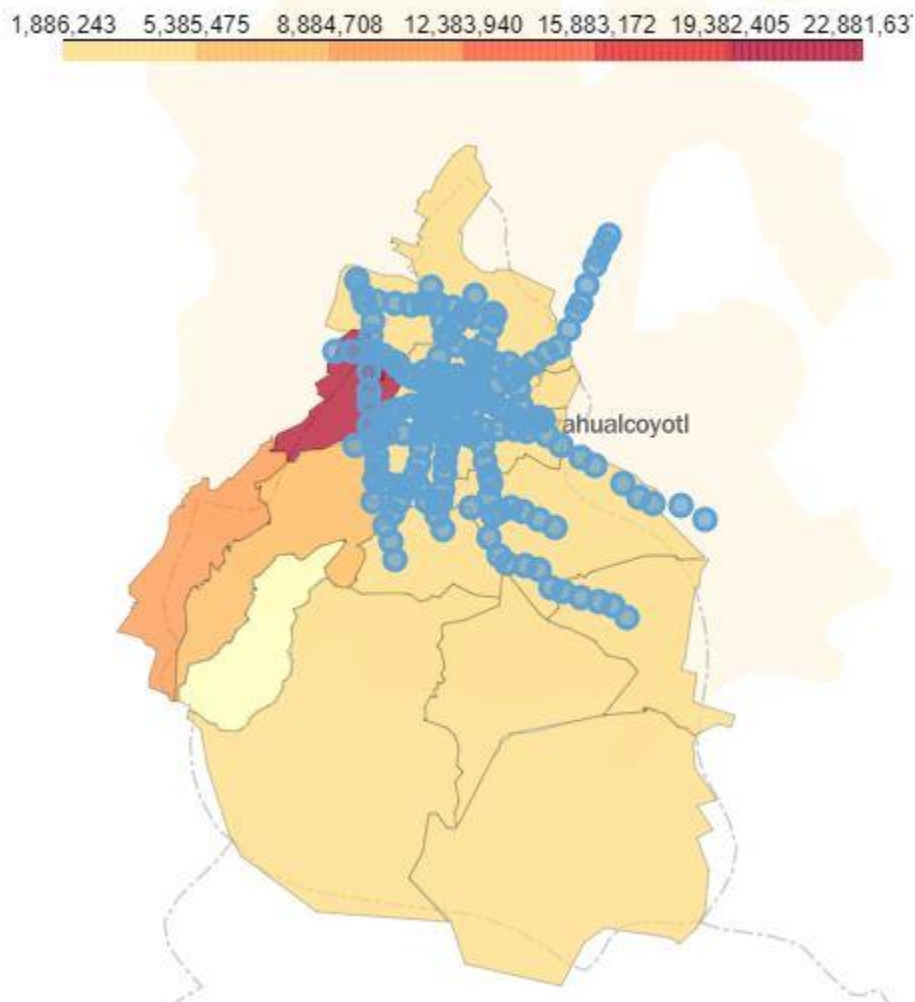


3.3 Metro Data

Next, I imported the .csv file containing information on metro stops, dropped irrelevant columns, and saved the result as a pandas dataframe:

	stop_name	stop_lat	stop_lon
0	Bellas Artes_1_2	19.435960	-99.141558
1	Canal del Norte_1	19.448667	-99.115895
2	Colegio Militar	19.448646	-99.171224
3	Buenavista_1_4	19.446158	-99.152899
4	Chapultepec	19.420256	-99.176588

Then I added the metro data to the existing choropleth map:



3.4 Foursquare Data

Finally, I entered my Foursquare developer credentials and defined a function to access the Foursquare API. The function was designed to iterate through a series of names, search each name in Foursquare, and return the top 15 venues (including their name, GPS coordinates, and category type). To do this, the

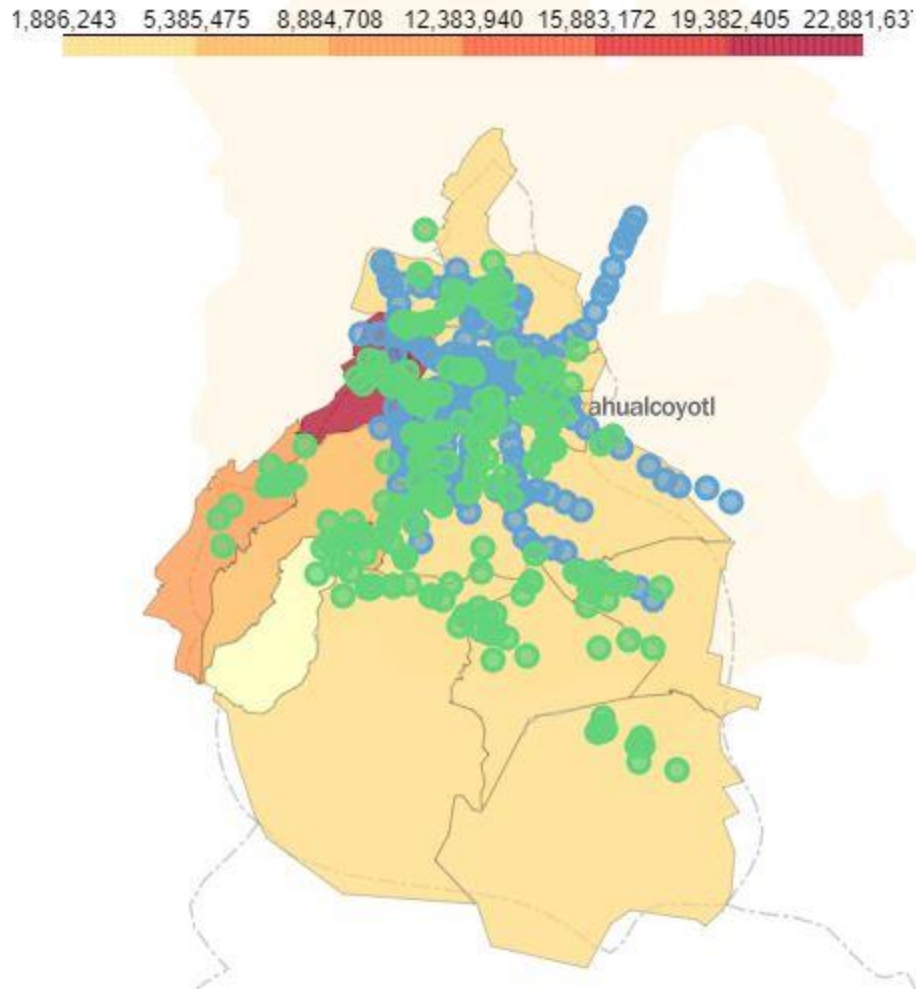
function used the explore endpoint and the near and limit parameters. I set 'limit' equal to 15, and 'near' equal to the name of a different borough with each iteration.

Next, I was able to feed the function the 'Boroughs' column we created earlier as part of the housing dataframe and create a new pandas dataframe with the resulting information:

	Venue	Venue Latitude	Venue Longitude	Venue Type
0	Bosque de Tlalpan	19.295363	-99.193256	Park
1	Superman - El Último Escape	19.295267	-99.209841	Theme Park Ride / Attraction
2	Medusa Steel Coaster	19.294227	-99.210888	Theme Park Ride / Attraction
3	La Castellana	19.286187	-99.157800	Liquor Store
4	Pista del Bosque De Tlalpan	19.295567	-99.193259	Trail

3.5 Final product

Using the geographical boundaries of each borough combined with data on housing, metro stops, and popular venues, I was able to create an interactive map that allows the user to easily identify housing cost, accessible public transportation, and popular destinations quickly:



4. Results

To anyone familiar with Mexico City, the resulting choropleth map is not very surprising. The most expensive parts of the city are indeed on the west side, and the borough of Miguel Hidalgo is infamous for astronomical housing prices and very wealthy residents. However, we also see that by only looking at data on the level of boroughs, we miss some data at the neighborhood level; certainly, not all areas within the same borough are the same. Much like New York City, the larger metropolitan area – or, in other words, the area that encompasses people who commute to the city daily – covers a much larger area than the city limits. This could also be taken into account in further analysis, as businesses wanting to capture the Mexico City market should also consider the nearby Estado de México market.

The metro and venue data are also not surprising, as most of the urban part of the city is centrally located, with city parks and forests in the southern part of the city. The expansion of the metro system beyond city limits once again shows us that the larger metropolitan area is not confined by city boundaries, and these areas should be studied in the future.

Lastly, we see that there are quite a few things missing from the map: where are the parks and roads? Some areas of the city are not able to be developed due to ecological reserves, and users with cars

would certainly want to see which parts of the city are accessible to them. Future analysis should take this into account.

5. Discussion

Despite its many oversights, this is a good start at a tool that might be very useful for businesses and individuals alike to start to get to know Mexico City. The fact that many of the findings of this exercise were not very surprising speaks to the accuracy and usefulness that the tool might offer: someone who knows nothing about Mexico City – in other words, the end user – wants to get key, “obvious” insights from which to build upon. With information on the metro system, the metropolitan layout of the city, and the reality that the economy (or at least transportation) of the city reaches beyond the city limits, the end user will walk away with at least a basic understanding of the city’s geography.

6. Conclusion

Mexico City, with a metropolitan population of more than 20 million, can be overwhelming for new coming businesses and individuals alike. This tool allows the user to easily identify the general price of each borough by using median housing prices as a proxy, and from there identify possible areas of interest by providing location data on metro stops and popular destinations. This might be useful for high-end businesses wanting to identify expensive areas far from public transit but close to other similar amenities, or for more affordable restaurants wanting to capture customers along their commute routes. Likewise, individuals with a car might prefer an affordable location with restaurants close by without considering metro access; individuals without a car might place high value on having a metro stop nearby. This is meant to illustrate that the tool presented is flexible and might be useful for a variety of users and applications.

Finally, the tool could be improved in the future to include data at a neighborhood level. Although this is a good initial overview for true newcomers, residents of the city know that both housing values and venue options vary greatly within each borough. Additionally, incorporating street maps would allow the user to identify usable and non-usable areas – for example, the entirety of Chapultepec Park is not labeled on this version of the choropleth map. This would allow the user to easily see which areas are real options for potential housing or business ventures.