CAPSTONE PROJECT

Neighborhoods in central Mexico City
By parks availability and quality perception

Spatial analysis- clustering

Isaac Perez

INTRODUCTION AND THE PROBLEM

How are neighborhoods in central Mexico City similar or different to each other based on the availability and quality perception of their public spaces and particularity parks?

Can we grouped neighborhoods according to their parks availability and how people rate them?

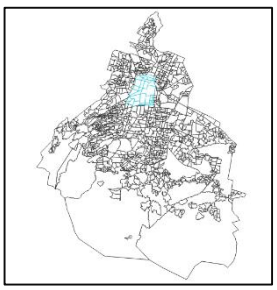
Are neighborhoods clustered with more and better parks associated with higher level of development and less poverty and the other way around?

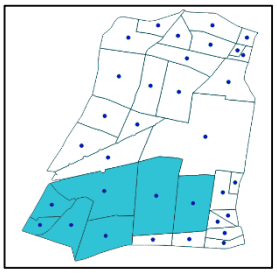
Public spaces such as parks are often unevenly distributed in cities. In some urban areas, usually those with high inequality, there are spatial patterns of concentrations and deficit of public spaces. Some areas might have not only more but also better public spaces than others. There are neighborhoods, on the other hand, that have comparatively few, less vibrant, even unsafe and poor quality public spaces or don't have any at all. This, in some cases, exacerbates, other forms of inequality and deprivation, prevents people to spend time outside and to meet with others. It might also mean they have less access to green spaces, impacting their health and the urban landscape.

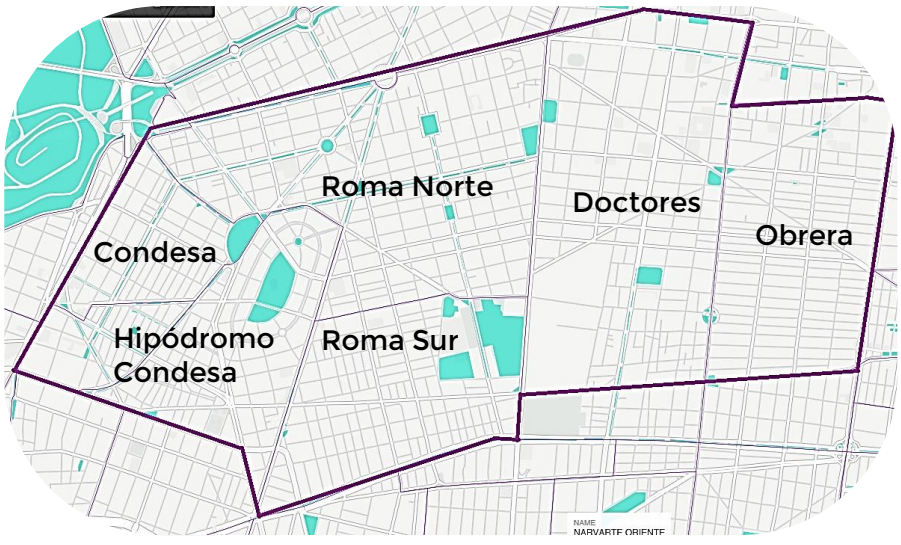
Mexico City stands as one of the largest metropolis in the world with about 22 million people and spanning over nearly 8 thousand square kilometers. The city is also full of contrast, with boroughs with a human development index comparable to Finland or France while also some comparable to countries in sub-Saharan Africa. About one third of the population lives in poverty and about 4 percent in extreme poverty, equivalent to 800 thousand people. The city also displays differences within and across boroughs and neighborhoods in terms of infrastructure, services and public spaces,

As shown in the map below, which shows neighborhoods in the central borough Cuauhtemoc and public spaces-parks in green, some neighborhoods in the west have more parks – and are more similar to each other in this regard, while the ones in the east are more similar to each other given their lack of parks. Cuauhtemoc houses about half a million people, areas with high, medium and low development.

Although this is something one could observed by looking a map like the one below and with this overlay, **there is not a tool, analysis or map that groups neighborhoods based on the availability and quality of parks. This prevents understanding this key dimension in the city, guide policy or interventions based on this knowledge and notion. This also would make more difficult for someone to take a decision to move to another neighborhood taking this aspect into consideration.**







THE GOAL

This project seeks to provide insights into the similarities and differences in neighborhood in central Mexico City (Cuauhtémoc Borough) based on the availability and ratings of parks. It intends to group – cluster the 33 neighborhoods in this borough according to these variables. It addition, it also aims at shedding light into how the resulted grouping is associated with the neighborhood social development.

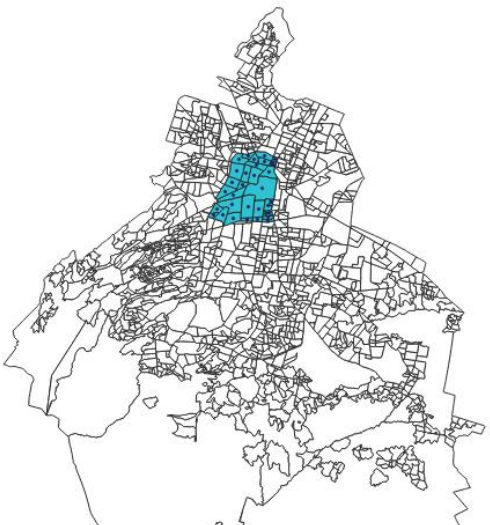
THE DATA

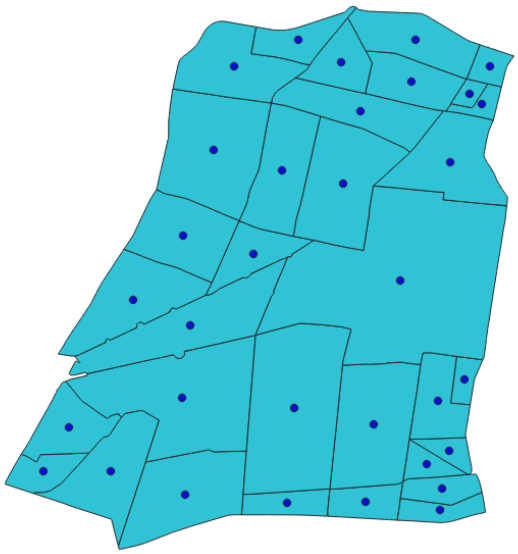
To develop this project data will be used from various sources including foursquare and leveraging the data available through the API.

1) The shape files of the Neighborhoods in Mexico City are available for from the city's Urban Development Department. The city has over a thousand neighborhoods distributed in 16 boroughs. For this particular project, only 33 that are part of the Cuauhtemoc borough will be filtered and used (highlighted in blue below). I have download the shape files, created new shape and json files with only Cuauhtemoc and also generated the centroids of each neighborhoods to have them both as points and polygons.
2) Data about the parks location and rating will be gathered using the foursquare API. There is data about 50 parks in Cuauhtemoc in a 2.5 kilometer radius from its centroid coordinates in their data set. Cuauhtemoc has about this radius, so this is an appropriate radius. The is also ratings for these venues in foursquare (see figure in the next page).
3) Data on the socio-economic development of neighborhoods is available from the city's social development department (See figure in the next page).

Using data on the parks locations and their ratings, two variables will be crated, number of parks and average parks ratings per neighborhood. Each neighborhood in a new data set-frame will have information on these two dimensions. K-means clustering then will be used to group neighborhoods based on these characteristics. The two variables will be normalized to facilitate the method execution. In addition, and once this step is completed, it will be explored what the most common development level of each group using bar plots and spatial overlays of these two dimensions.

1 Neighborhoods location data





2 Foursquare api data

Json file- request for parks

```
[35]: results = requests.get(url).json()
      results

[35]: {'meta': {'code': 200, 'requestId': '5ef946cc513bb60f5114b60b'},
       'response': {'venues': [{'id': '4cc8ae0b41e75481e20e5884',
          'name': 'Parque Alameda',
          'location': {'address': 'Av. Juárez 76',
           'crossStreet': 'Esq. con Azueta',
           'lat': 19.434515120135995,
           'lng': -99.14715262455131,
           'labeledLatLngs': [{'label': 'display',
             'lat': 19.434515120135995,
             'lng': -99.14715262455131}],
           'distance': 80,
           'postalCode': '06010',
```

Json to data frame

```
[39]: venues = results['response']['venues']

      # tranform venues into a dataframe
      dataframe = json_normalize(venues)
      dataframe.head(5)
```

	id	name	categories	referralId	hasPerk	location.address	location.crossStreet	location.la
0	4cc8ae0b41e75481e20e5884	Parque Alameda	[{'id': '4bf58dd8d48988d1fd941735', 'name': 'S...	v-1593395281	False	Av. Juárez 76	Esq. con Azueta	19.434515
1	4ec7df32f79041351d2d6868	Parque Ciudadela	[{'id': '4bf58dd8d48988d182941735', 'name': 'T...	v-1593395281	False	NaN	NaN	19.427322
2	53e8320c498ee437284d3f15	Parque Olof Palme	[{'id': '4bf58dd8d48988d163941735', 'name': 'P...	v-1593395281	False	Rio Grijalva	Rio Danubio	19.433327
3	4da21ca29aa4721e3473091a	Parque José Martí	[{'id': '4bf58dd8d48988d164941735', 'name': 'P...	v-1593395281	False	Reforma	Hidalgo	19.437105

Csv file with parks in Cuauhtemoc

	id	name	categories	referralId	hasPerk	location.add	location.cros	location.lat	location.lng	location.labe	location.dist	lo
0	4cc8ae0b41e	Parque Alameda	[{'id': '4bf58cv-159339528:		False	Av. JuÃ¡rez 7	Esq. con Azu	19.4345151	-99.14715262	[{'label': 'dis	80	
1	4ec7df32f79	Parque Ciudadela	[{'id': '4bf58cv-159339528:		False			19.4273216	-99.14975289	[{'label': 'dis	918	
2	53e8320c498	Parque Olof Palme	[{'id': '4bf58cv-159339528:		False	Rio Grijalva	Rio Danubio	19.4333273	-99.17075567	[{'label': 'dis	2475	
3	4da21ca29aa	Parque JosÃ© MartÃ-	[{'id': '4bf58cv-159339528:		False	Reforma	Hidalgo	19.4371054	-99.14597898	[{'label': 'dis	248	
5	5e694bde14l	Dorothy Gaynor Parque Alm	[{'id': '4bf58cv-159339528:		False			19.435038	-99.146905	[{'label': 'dis	43	
6	5488464e498	parque para perros Pushkin	[{'id': '4bf58cv-159339528:		False	Morelia	Colima	19.4209741	-99.15446091	[{'label': 'dis	1757	
7	55fd8f8949d	Parque Galeria	[{'id': '4bf58cv-159339528:		False	Puebla 170	Roma Norta	19.4223491	-99.16258604	[{'label': 'dis	2154	
8	4d037f229d3	Parque San Fernando	[{'id': '4bf58cv-159339528:		False	Puente de A	San Fernand	19.4389003	-99.14609471	[{'label': 'dis	426	
9	4fdfe1bbe4l	Parque	[{'id': '4bf58cv-159339528:		False			19.4274756	-99.1428993	[{'label': 'dis	977	
10	527f9b111d	Parque Via 198	[{'id': '5032f cv-159339528:		False			19.4353684	-99.16816103	[{'label': 'dis	2193	
11	4ebf01607ee	Parque de Los Cuatro Vient	[{'id': '4bf58cv-159339528:		False			19.4626279	-99.13855394	[{'label': 'dis	3183	
12	50282730e4l	Parque Recreativo	[{'id': '4bf58cv-159339528:		False			19.4301949	-99.14349365	[{'label': 'dis	686	
13	5249ecb2498	Parque De Los CaÃ±tones	[{'id': '4bf58cv-159339528:		False			19.4301762	-99.15020693	[{'label': 'dis	641	

2 Socio-economic development level data

Cuauhtemoc Borough

33 Neighborhoods

Development Level (High, medium and low

Nombre de la Delegación	Nombre de la Colonia o Barrio	Habitantes	Indice de Desarrollo Social		
			Valor	Estrato	Grado
Cuauhtémoc	Algarín	5,556	0.90156	4	Alto
Cuauhtémoc	Ampliación Asturias	5,708	0.86853	3	Medio
Cuauhtémoc	Asturias	4,364	0.86687	3	Medio
Cuauhtémoc	Atlampa	14,433	0.86445	3	Medio
Cuauhtémoc	Buenavista	15,605	0.86896	3	Medio
Cuauhtémoc	Buenos Aires	5,772	0.79160	2	Bajo
Cuauhtémoc	Centro	61,229	0.75839	2	Bajo
Cuauhtémoc	Condesa	8,453	0.96018	4	Alto
Cuauhtémoc	Cuauhtémoc	11,399	0.95240	4	Alto
Cuauhtémoc	Doctores	44,703	0.85298	3	Medio
Cuauhtémoc	Esperanza	4,072	0.84225	3	Medio
Cuauhtémoc	Ex-Hipódromo de Peralvillo	11,711	0.84403	3	Medio
Cuauhtémoc	Felipe Pescador	1,988	0.88338	3	Medio
Cuauhtémoc	Guerrero	42,339	0.83420	3	Medio
Cuauhtémoc	Hipódromo	13,572	0.96310	4	Alto
Cuauhtémoc	Hipódromo Condesa	3,204	0.96037	4	Alto
Cuauhtémoc	Juárez	10,184	0.91452	4	Alto
Cuauhtémoc	Maza	2,503	0.82626	3	Medio
Cuauhtémoc	Morelos	36,590	0.74957	2	Bajo
Cuauhtémoc	Nonoalco Tlatelolco	27,843	0.96114	4	Alto
Cuauhtémoc	Obrera	35,224	0.83980	3	Medio
Cuauhtémoc	Paulino Navarro	5,307	0.82021	3	Medio
Cuauhtémoc	Peralvillo	20,213	0.84848	3	Medio
Cuauhtémoc	Roma Norte	27,770	0.91764	4	Alto
Cuauhtémoc	Roma Sur	17,435	0.93359	4	Alto
Cuauhtémoc	San Rafael	19,684	0.92592	4	Alto
Cuauhtémoc	San Simón Tolnáhuac	9,885	0.86051	3	Medio
Cuauhtémoc	Santa María Insurgentes	1,480	0.89881	3	Medio
Cuauhtémoc	Santa María La Ribera	40,960	0.90479	4	Alto
Cuauhtémoc	Tabacalera	3,267	0.89747	3	Medio
Cuauhtémoc	Tránsito	9,720	0.81415	3	Medio
Cuauhtémoc	Valle Gómez	6,281	0.78673	2	Bajo
Cuauhtémoc	Vista Alegre	3,377	0.88345	3	Medio