



*Coursera: IBM Applied Data Science Capstone*

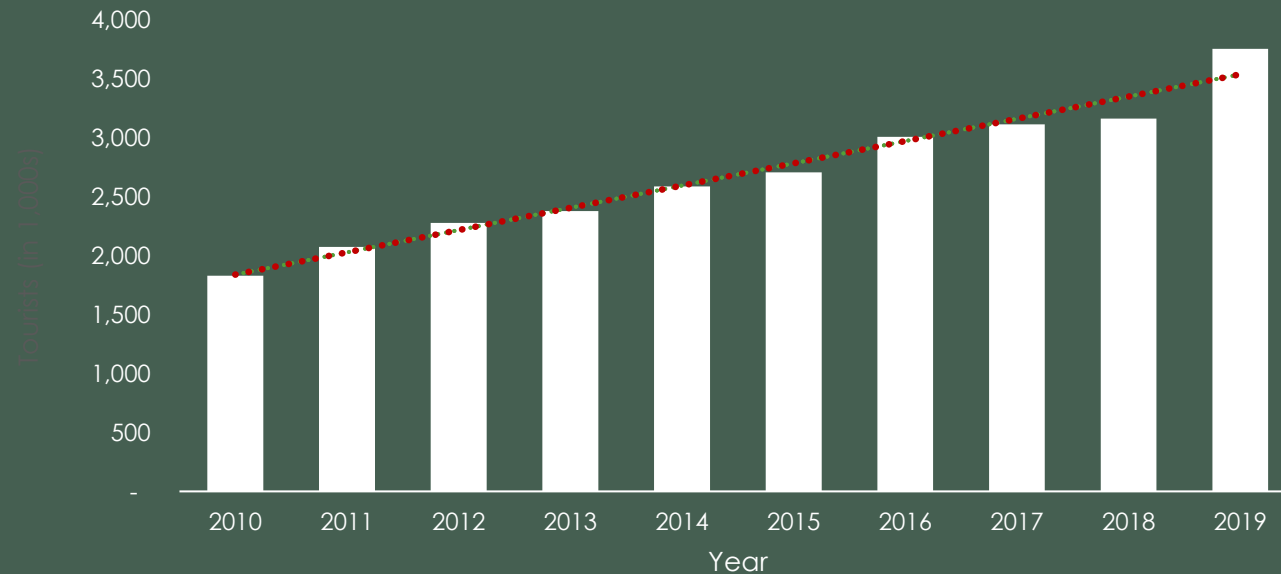
# Opening a New Hotel in Mexico City's Central Borough Cuauhtémoc

By: Cynthia Valadez

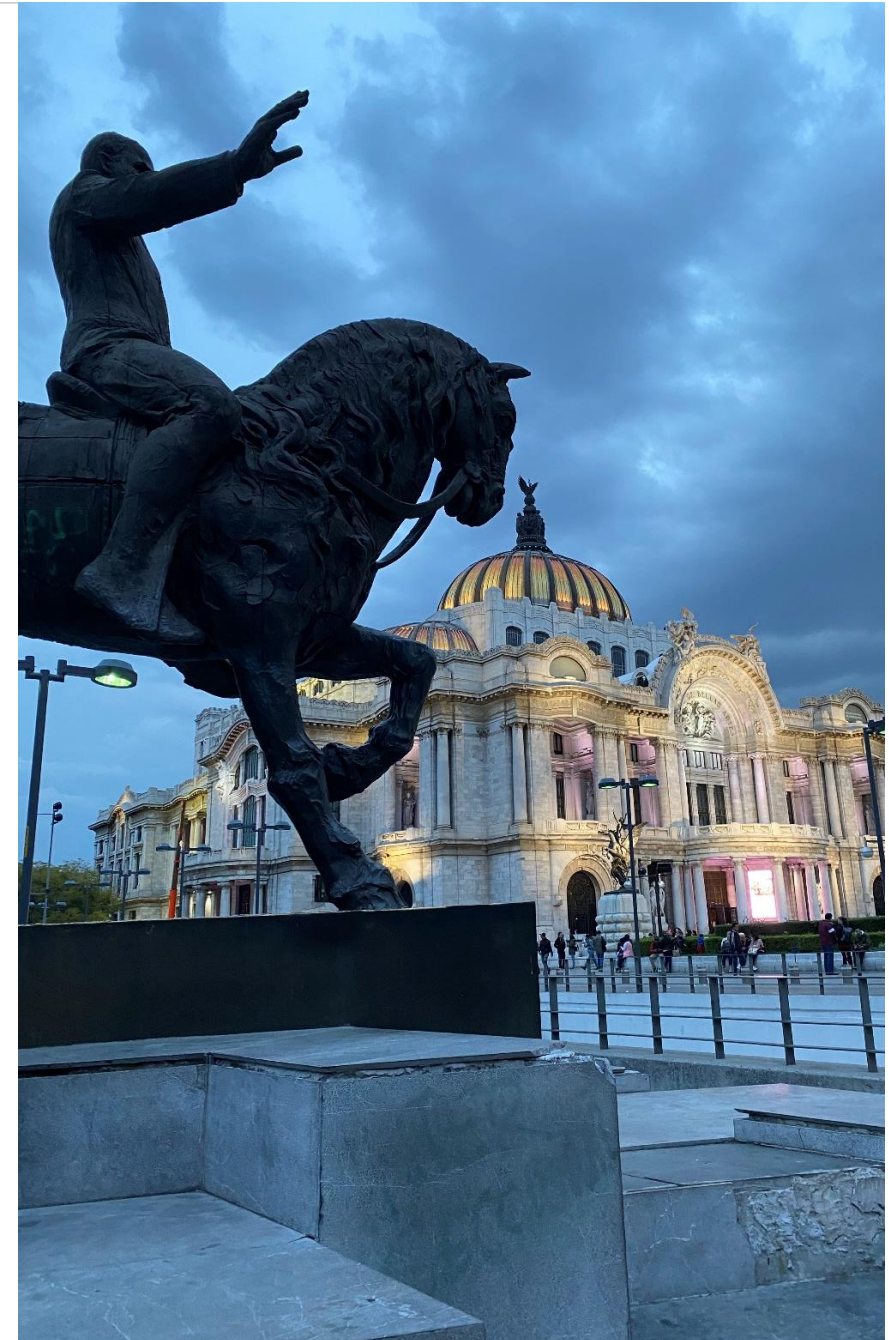
*February 2021*

# Introduction: *Mexico City Tourism*

Number of International Tourist Arrivals to Hotels in Mexico  
City from 2010 to 2019



Tourism in Mexico City has been booming in recent years. Foreign tourists have flocked to this bustling city in pursuit of more favorable weather, affordable experiences and cultural immersion. With this growth in visitors, there is also a growing need for more accommodation.





# Business Problem

- Identifying the ideal location to build a new hotel to provide additional accommodation to the growing number of tourists arriving to the city each year
- The main objective is to determine which area of Mexico City's central borough of Cuauhtémoc would be most favorable to develop a new hotel

# Data

Required Data	Sources
<ol style="list-style-type: none"><li>1. List of neighborhoods or “colonias” within the borough of Cuauhtémoc</li><li>2. Geographic coordinates (latitude/longitude) of each of the colonias</li><li>3. Nearby venue data for each of the colonias(specifically hotels)</li></ol>	<ol style="list-style-type: none"><li>1. Wikipedia webpage (<a href="https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Mexico_City#Cuahatemoc">https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Mexico_City#Cuahatemoc</a>)</li><li>2. Python's Geocoder library</li><li>3. Foursquare API (utilizing ID and Secret Key)</li></ol>

# Methodology

1. Scraping a Wikipedia webpage for list of neighborhoods by using Python's BeautifulSoup library
2. Use Python's Geocoder library to obtain the latitude and longitude coordinates of each of the neighborhoods
3. Use Foursquare API credentials to obtain venue data
4. Group data by each neighborhood and calculate the mean of the frequency of existing distinct venues
5. Filter venue category by Hotel
6. Utilize Machine Learning(k-means clustering) to cluster neighborhoods into three clusters based on hotel preponderance
7. Plot the clusters/neighborhoods into a map using the Folium package





# Results

- Cluster 0 has the highest frequency of hotels in each colonia followed by a more moderate frequency of hotels in colonias in Cluster 2.
- This leaves us with a lower preponderance or even nonexistence of hotels in colonias in Cluster 1 which encompasses neighborhoods located more on the outskirts of the Cuauhtémoc borough.

Cluster	Hotel Preponderance	Color on Map
Cluster 0	High	Red
Cluster 1	Low	Purple
Cluster 2	Moderate	Green



# Discussion/Conclusion

- The colonias in Cluster 1 present hotel developers an ideal opportunity to build a new hotel in areas that are not currently saddled with significant hotel competition. On the other hand, it is not advised to build new hotels where hotel competition is higher such as in colonias in Cluster 0 or even in Cluster 2.
- Notably, the colonias in the southern area of the Cuauhtémoc borough would be a more favorable option to build a hotel since those neighborhoods are at a closer proximity to some of the major tourist attractions—such as the Bosque de Chapultepec—than their northern cluster counterparts.
- Nevertheless, there are still other factors that must be considered before building a new hotel other than just competition. Some of these factors include the populations of each distinct colonia, the income distribution, and available land for development in each neighborhood.
- There would need to be even more research on local offerings and as well as hotel offerings to stand out. More intrinsic qualities such as aesthetic and theme would also have to be considered prior to development.



THE END.  
THANK YOU!



---