The Battle of Neighbourhoods-Coursera Capstone

Delhi vs Mumbai

The Business Problem

The business problem in this study assumes that people who would be interested in this study are those who would like to create a projection of potential life and activities in these metro city neighborhoods if the subject moves to live in one of them. The decision to choose one over the other would depend on popular venues in the neighborhoods in each of these metro cities.

Mumbai and Delhi are the two most important metro cities in India. There has been a war for supremacy in terms of quality of life, jobs, education, entertainment and recreational facilities that these cities have to offer to its residents. I attempt to analyze the neighborhoods in each of these two cities and try to understand what is popular in them and what they have to offer to someone who is contemplating to make a choice on seeking a life in either of the metro cities of India.

Delhi

Delhi, India's capital territory, is a massive metropolitan area in the country's north. In Old Delhi, a neighborhood dating to the 1600s, stands the imposing Mughal-era Red Fort, a symbol of India, and the sprawling Jama Masjid mosque, whose courtyard accommodates 25,000 people. Nearby is Chandni Chowk, a vibrant bazaar filled with food carts, sweets shops and spice stalls.



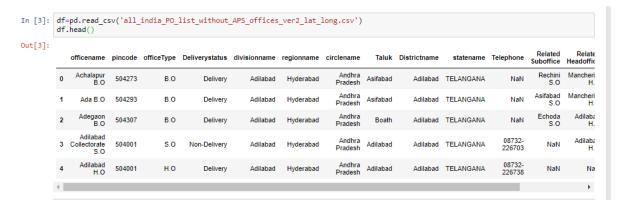
Mumbai

Mumbai (formerly called Bombay) is a densely populated city on India's west coast. A financial center, it's India's largest city. On the Mumbai Harbour waterfront stands the iconic Gateway of India stone arch, built by the British Raj in 1924. Offshore, nearby Elephanta Island holds ancient cave temples dedicated to the Hindu god Shiva. The city's also famous as the heart of the Bollywood film industry.



Data Used

For any "data science project" data is of paramount importance. For this study, we needed data about neighborhoods in each of these metro cities. The data published by the government on postal codes for all India would serve us well for this study. We will specifically download the CSV provided under https://data.gov.in/resources/all-india-pincode-directory-contact-details-along-latitude-and-longitude.



Methodology

In this study, we will download the CSV, read it into a pandas Dataframe and curate it to remove the data related to all other cities, towns, and places which are not Mumbai or Delhi, since we are only interested in comparing these two biggest metro cities in India.

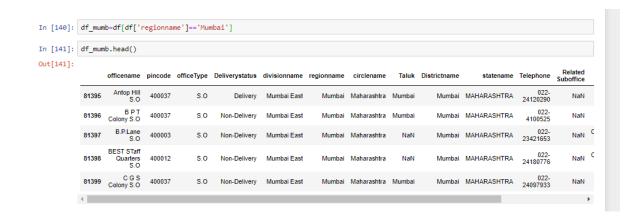
We shall then clean up the unnecessary columns in the CSV, which is not relevant or useful for our current study. Post office names (*office name*) will be used as the neighborhood names in each of the regions such as Mumbai or Delhi.

Neighborhood names with the same *Pincode* will be combined as a single row.

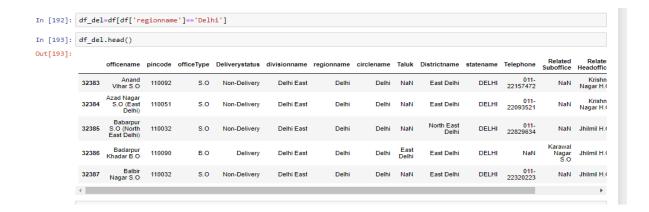
Foursquare API will be used to find the longitude and latitude of each of the neighborhoods in both Mumbai and Delhi. This will form the dataset we will use for this study.

The first few records of the dataset.

Mumbai



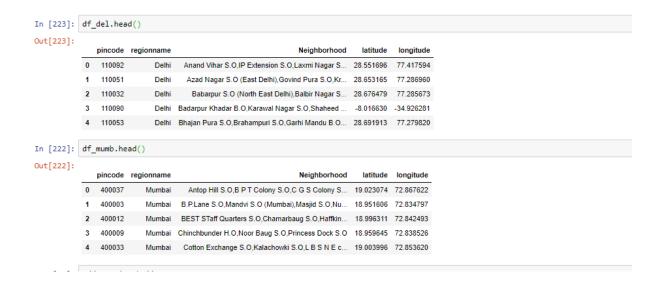
Delhi



We now see that there are the same *Pincode* values for different neighborhoods. The next step is to combine the rows having the same Pincode, we do this by changing the value of the neighborhood by building a comma-separated concatenation of neighborhood values for rows with the same Pincode.

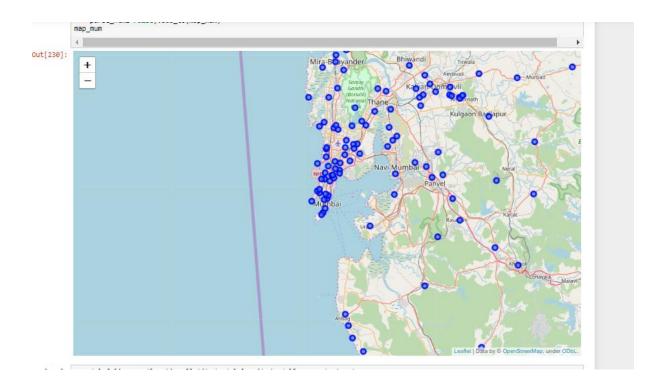
The next step is to enhance the dataset with the required information. We would need the longitude and latitude values for the neighborhoods. We will use the *Nominatim* library from *geocoders.geopy* package to find the longitude and latitude for each of the neighborhoods and would eventually create a dataset having all the necessary columns for our analysis.

Dataframe after cleanup

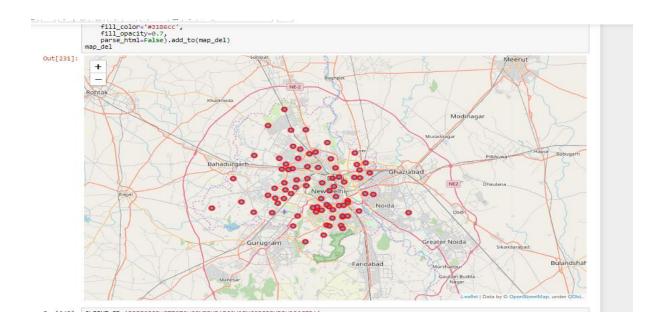


We now have the necessary information to visualize the neighborhoods for both the cities on a *folium* map.

<u>Mumbai</u>



<u>Delhi</u>



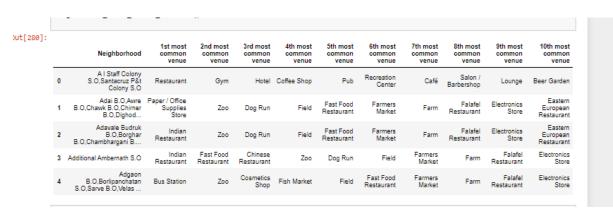
Analyzing the neighborhoods

Finding top venues near Mumbai neighborhoods

We will use the Foursquare API to find the top venues in the neighbourhoods of Mumbai. This will help us understand the nature of life Mumbai neighborhoods have to offer. We will iteratively make Foursquare API calls for each of the Mumbai neighborhoods in our dataset. Foursquare API returns the popular venues within 500m radius of this neighborhood.

Next, we will employ statistically and analytical methods to find the unique venues/venue categories in the Mumbai neighborhoods and we will build a Dataframe that calibrates each of the neighborhoods with the frequency of occurrence for each of the venue category.

We then create a dataset that lists the top 10 common venues against each of the neighbourhoods in Mumbai. We get a representation such as below for all the neighbourhoods in Mumbai.



Cluster the neighbourhoods in Mumbai based on the similarity of top common venues

Given that we now have the required information regarding the top venues against each of the neighborhoods in Mumbai, let us now apply a clustering algorithm to group the neighborhoods based on the similarity in types of venues they have. By clustering, we also provide information to users on a common type of neighbourhoods in Mumbai. We will use the k-Means clustering approach to cluster the neighbourhoods. k will be selected as 5. This means that we will group the neighborhoods into 5 clusters. Each of the neighborhoods gets a Cluster Label assigned.

Clustering

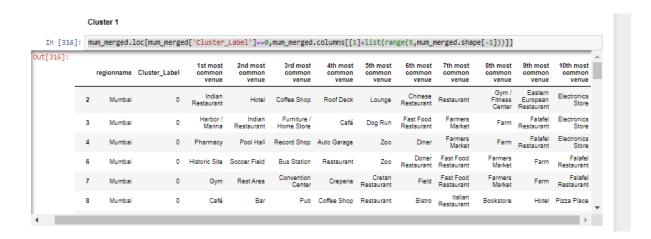
```
In [295]: mum_group_cluster=mum_grouped.drop('Neighborhood',1)
    model1=KMeans(n_clusters=5,random_state=5).fit(mum_group_cluster)
    model1.labels_
```

We will then use the dataset with cluster labels assigned to visualize the clusters in a *folium* map.

A piece of important information this map provides is that many neighborhoods in Mumbai are of similar nature concerning the venues they have around, indicated by the cluster marked in blue.

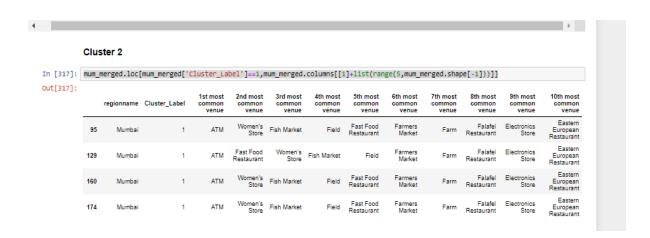
Let us now dig a little deeper into how the neighborhoods are clustered and what is the characteristic of the cluster that is very common across most neighborhoods in Mumbai.

Cluster 1



From this cluster we can conclude that this area has mix of all including eateries ,gym,historic site and many more. For people to have lots of variety of food café's and even for medical requirements pharmacy is also there. Also for people to exercise Gym facilty is also pretty common

Cluster 2



If someone has withdraw cash from here, he will have no problem as there are lots of ATMs and for women there are lots of women store.

-[246]-													
Out[319]:		regionname	Cluster_Label	1st most common venue	2nd most common venue	3rd most common venue	4th most common venue	5th most common venue	6th most common venue	7th most common venue	8th most common venue	9th most common venue	10th most common venue
		Mumbai	2	Indian Restaurant	Smoke Shop	Zoo	Doner Restaurant	Fish Market	Field	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
	1	Mumbai	2	Indian Restaurant	Middle Eastern Restaurant	Ice Cream Shop	Chinese Restaurant	Café	Restaurant	BBQ Joint	Indian Sweet Shop	Dessert Shop	Electronics Store
	5	Mumbai	2	Indian Restaurant	Movie Theater	Hockey Arena	Optical Shop	Zoo	Dog Run	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant
	31	Mumbai	2	Indian Restaurant	Mattress Store	Bus Station	loe Cream Shop	Café	Zoo	Donut Shop	Field	Fast Food Restaurant	Farmers Market
	35	Mumbai	2	Indian Restaurant	Snack Place	Fast Food Restaurant	Golf Course	Food Truck	loe Cream Shop	Dog Run	Farmers Market	Farm	Falafel Restaurant
	38	Mumbai	2	Vegetarian / Vegan Restaurant	Electronics Store	Dance Studio	Indian Restaurant	Zoo	Dog Run	Field	Fast Food Restaurant	Farmers Market	Farm
	40	Mumbai	2	Indian Restaurant	Juice Bar	Gym / Fitness Center	Coffee Shop	Donut Shop	Restaurant	Electronics Store	Bed & Breakfast	Bakery	Men's Store
	43	Mumbai	2	Indian Restaurant	Department Store	Multiplex	Chinese Restaurant	Lake	Hotel	Restaurant	Mughlai Restaurant	Eastern European Restaurant	Donut Shop

For those who love Indian cuisine, they lots of Indian restaurant including a veg restaurant, Middle east restaurant and many more. Movie goer can always find it easy going to a theatre here. There's departmental store to help cosnumers buy wide variety of goods

Cluster 4



If people want mobile ,they'll easily get here. For those who want to commute vie train wont find any difficulty here .There is a fishing store and zoo nearby



For those gym lovers and animal lovers this is place they'll easy access to their favourite places

Finding top venues near Delhi neighbourhoods

We will use the Foursquare API to find the top venues in the neighborhoods of Delhi. This will help us in understanding the nature of life Delhi neighborhoods have to offer. We will iteratively make Foursquare API call for each of the Delhi neighborhoods in our dataset.

Foursquare API returns the following response as the popular venues close to 500m radius of this neighborhood.

Next, we will employ statistically and analytical methods to find the unique venues/venue categories in the Delhi neighbourhoods and will build a Dataframe that calibrates each of the neighbourhoods with the frequency of occurrence of each of the venue category

From our analysis, we see that there are 14 unique venue categories in Delhi neighbourhoods. ATMs, Arts and Crafts stores, Burger Joints, Cafes, Gardens, Gyms, Multiplexes, Museums, Pizza places, Indian restaurants, Shopping malls, Water Parks, Gardens and Hotels being some of them.

We then create a dataset that lists the top 10 common venues against each of the neighborhoods in Delhi. We get a representation such as below for all the neighborhoods in Delhi.

90]:	Neighborhood	1st most common venue	2nd most common venue	3rd most common venue	4th most common venue	5th most common venue	6th most common venue	7th most common venue	8th most common venue	9th most common venue	10th most common venue
1 2	505 A B Workshop S.O,A F Palam S.O,Aps Colony 	Restaurant	Gay Bar	Bakery	Latin American Restaurant	Arepa Restaurant	Bar	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Steakhouse	Burger Joint
	A F Rajokari S.O,Rajokari B.O	ATM	Fast Food Restaurant	Fried Chicken Joint	French Restaurant	Food Truck	Food Stand	Food Court	Food & Drink Shop	Food	Fondue Restaurant
	A.G.C.R. S.O,Ajmeri Gate Extn. S.O,Darya Ganj	Indian Restaurant	Hotel	Restaurant	Hostel	History Museum	Cricket Ground	Stadium	Road	Historic Site	Asian Restaurant
	A.K.Market S.O,Multani Dhanda S.O,Pahar Ganj S	Hotel	Pizza Place	Motel	Fast Food Restaurant	Restaurant	Coffee Shop	Arts & Crafts Store	Miscellaneous Shop	Hostel	Food Truck
	Abul Fazal Enclave-I 4 S.O, Jamia Nagar	Indian Restaurant	Athletics & Sports	Coffee Shop	Gym / Fitness Center	Boat or Ferry	Hotel Pool	Farmers Market	Food Truck	Food Stand	Food Court

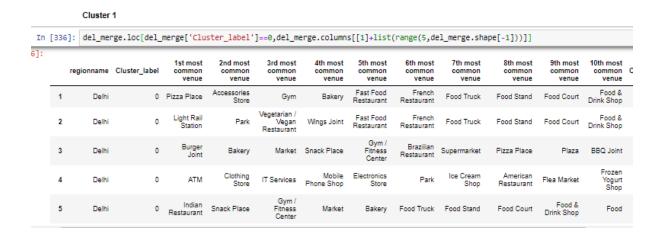
Cluster the neighborhoods in Delhi based on the similarity of top common venues

Given that we also have the required information regarding the top venues against each of the neighborhoods in Delhi, let us now apply a clustering algorithm to group the neighborhoods based on the similarity in types of venues they have. By clustering, we also provide information to users on a common type of neighborhood in Delhi. We will use the k-Means clustering approach to cluster the neighbourhoods. k will be selected as 5. This means that we will group the neighborhoods into 5 clusters. Each of the neighborhoods gets a Cluster Label assigned.

```
del_grouped_cluster=del_grouped.drop('Neighborhood',1)
model2=KMeans(n_clusters=5,random_state=5).fit(del_grouped_cluster)
model2.labels_
```

We will then use the dataset with cluster labels assigned to visualize the clusters in the *folium* map.

A piece of important information this map provides is that the neighborhoods in Delhi are of diverse nature concerning the venues they have around, indicated by the clusters marked in different colors. Also, we did see earlier that we did not have too many venue categories for the neighbourhoods that were returned for the neighbourhoods in Delhi.



This is area has wide variety of places that people find common from eateries to ATM, clothing store etc

Cluster 2

	Cluste	Cluster 2												
	del_merge.loc[del_merge['Cluster_label']==1,del_merge.columns[[1]+list(range(5,del_merge.shape[-1]))]]													
Out[337]:	regionname		Cluster_label	1st most common venue	2nd most common venue	3rd most common venue	4th most common venue	5th most common venue	6th most common venue	7th most common venue	8th most common venue	9th most common venue	10th most common venue	
	20	Delhi	1	Electronics Store	Bank	Wings Joint	Fish & Chips Shop	Fried Chicken Joint	French Restaurant	Food Truck	Food Stand	Food Court	Food & Drink Shop	
	31	Delhi	1	Electronics Store	Pizza Place	Bank	Fish & Chips Shop	Fried Chicken Joint	French Restaurant	Food Truck	Food Stand	Food Court	Food & Drink Shop	
	CI4.	2												

For those who are looking to buy electronic Appliances this is the place to be. Along with that people who want to visit the bank or wants have a bite of the pizza place wont have a problem here

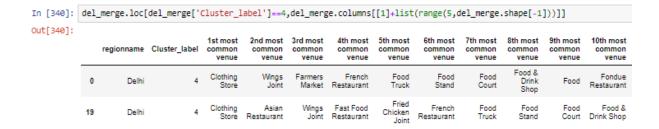


If someone has withdraw cash from here, he will have no problem as there are lots of ATMs and there are a lot of fast food restaurant, Motorcycle shop etc.

Cluster 4

Cluster 4 In [339]: del_merge.loc[del_merge['Cluster_label']==3,del_merge.columns[[1]+list(range(5,del_merge.shape[-1]))]] out[339]: 10th most 1st most 2nd most 3rd most 4th most 5th most 6th most 7th most regionname Cluster_label common common common common common common common common common venue Food & Drink Shop Fast Food Restaurant Chinese Restaurant Seafood Restaurant Wings Joint 49 Delhi Fish & Fast Food French Food Food Stand Food Court Indian Dessert 75 Delhi Café Chips Restaurant Restaurant Shop Restaurant Shop Truck Shop Vegetarian Food & Coffee Shop Fast Food Vegan taurant Drink Shop 79 Delhi Café Food Station Market Fish & Chips Shop Food & Drink Shop Fast Food Restaurant Chinese Restaurant Food Court 86 Preserve Restaurant

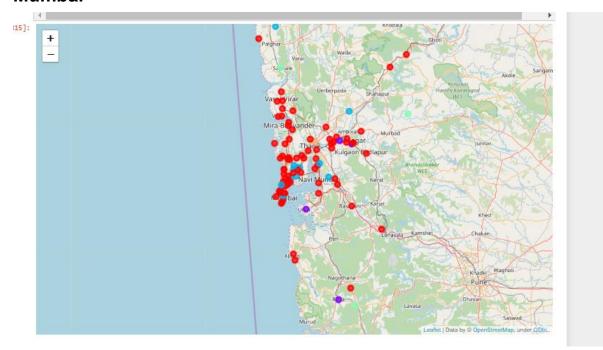
This is place for the foodies as there are a lot of places to eat here.



This is the place for those who want to buy clothing's or wish to have Asian Food.

Maps of Mumbai And Delhi After Clustering

Mumbai



Delhi

