IIT Bombay Venues

[Coursera Capstone Project]

Table of Contents

1. Introduction

Mumbai (formerly known as Bombay) is one of the top metro cities of India with rich heritage and culture. Indian Institute of Technology Bombay (IIT Bombay) situated in Mumbai, is a public engineering institution known for best education facilities in India.

1.1 Problem Statement

FIRST CONVOCATION

Lots of students stay away from their homes during their 4 year tenure of study at IIT Bombay. Being a student at IIT Bombay, I would like to reside in the neighbourhood of the institution. I will explore the venues of each neighbourhood of the institution and look for the neighbourhoods having ample hostels, malls, restaurants, departmental stores, book stores and tranportation etc. The best neighbourhood according to my preference would be my first choice for stay during my entire study term at IIT Bombay. :D This analysis may help students who wish to reside nearby IIT Bombay Campus.

2. Data

I downloaded the required data from a website published by government of india. Here is a link to the website... https://data.gov.in/resources/all-india-pincode-directory-contact-details-along-latitude-and-longitude.

This csv file, contains the list of pincodes of all cities of India, officenames (post office), the head office (HO) names of each officename etc. The Latitude and Longitude columns are incomplete for most areas. So, I will try to fill it with the help of Geopy. I am using Foursquare API to get the venues of each neighbourhood of the institution.

As I need only the details of 'Mumbai' city and 'Chembur' area in Mumbai of which IIT Bombay Campus is a sub-area, I filtered the dataframe accordingly.

A snippet of csv file is given below. Thank you!

officenam	pincode	officeType	Deliveryst	divisionna	regionnan	circlenam	Taluk	Districtna	statename	Telephone	Related Su	Related H	longitude	latitude
Achalapur	504273	B.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Asifabad	Adilabad	TELANGAN	NA	Rechini S.(	Mancheri	NA	NA
Ada B.O	504293	B.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Asifabad	Adilabad	TELANGAN	NA	Asifabad S	Mancheri	NA	NA
Adegaon E	504307	B.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Boath	Adilabad	TELANGAN	NA	Echoda S.(	Adilabad F	NA	NA
Adilabad (	504001	S.O	Non-Deliv	Adilabad	Hyderaba	Andhra Pr	Adilabad	Adilabad	TELANGAN	08732-226	NA	Adilabad F	NA	NA
Adilabad F	504001	H.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Adilabad	Adilabad	TELANGAN	08732-226	NA	NA	NA	NA
Agm Color	504251	S.O	Non-Deliv	Adilabad	Hyderaba	Andhra Pr	Bellampal	Adilabad	TELANGAN	08735-222	NA	Mancheri	NA	NA
Ainam B.C	504273	B.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Asifabad	Adilabad	TELANGAN	NA	Rechini S.(	Mancheri	NA	NA
Aknepalli	504251	B.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Luxettipet	Adilabad	TELANGAN	NA	Bellampal	Mancheri	NA	NA
Akoli B.O	504309	B.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Adilabad	Adilabad	TELANGAN	NA	Gumma S.	Adilabad F	NA	NA
Alegaon B	504201	B.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Chennur	Adilabad	TELANGAN	NA	Chinnoor	Mancheri	NA	NA
Aloor B.O	504110	B.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Nirmal	Adilabad	TELANGAN	NA	Chincholi	Adilabad F	NA	NA
Ambaripe	504202	B.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Khanapur	Adilabad	TELANGAN	NA	Peddur S.(	Adilabad F	NA	NA
Andavelli	504296	B.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Sirpur (t)	Adilabad	TELANGAN	NA	Sirpur Kha	Mancheri	NA	NA
Andugulpe	504231	B.O	Delivery	Adilabad	Hyderaba	Andhra Pr	Luxettipet	Adilabad	TELANGAN	NA	Kalyani Kh	Mancheri	NA	NA

3. Methodology

3.1 Formatting and cleaning data

1. We need the data pertaining to Districtname= 'Mumbai', so I will filter only the rows having Districtname= "Mumbai".
2. Since IIT Bombay's Related HeadOffice is "Chembur" , we will keep the part of dataframe .having Related HeadOffice= "Chembur".
3. Next we will, keep the columns officename, pincode, longitude, latitude.
4. We check for multiple occurences of each pincode.
5. We rename the column 'officename' as 'neighborhood' and combine the neighborhoods having same pincode with a ','.

A snippet of resulting dataframe is given below.

	pincode	neighborhood
0	400042	Bhandup East S.O
1	400043	D.M. Colony S.O, Shivaji Nagar S.O (Mumbai)
2	400071	Sindhi Society S.O
3	400074	Chembur Extension S.O, FCI S.O, Mahul Road S.O
4	400075	Best Staff Colony S.O, Pant Nagar S.O, R.A.Nag...
5	400076	Powai Iit S.O
6	400077	Rajawadi S.O
7	400078	Bhandup Ind. Estate S.O, Bhandup West S.O, J.M...
8	400079	Psm Colony S.O, Vikhroli S.O
9	400080	Mulund Dd Road S.O, Mulund West S.O, Nahur S.O...
10	400081	Mhada Colony S.O, Mulund East S.O

3.2 Geopy for finding Latitudes and Longitudes

I have included Latitude and Longitude values for each neighbourhood with the help of Geopy.
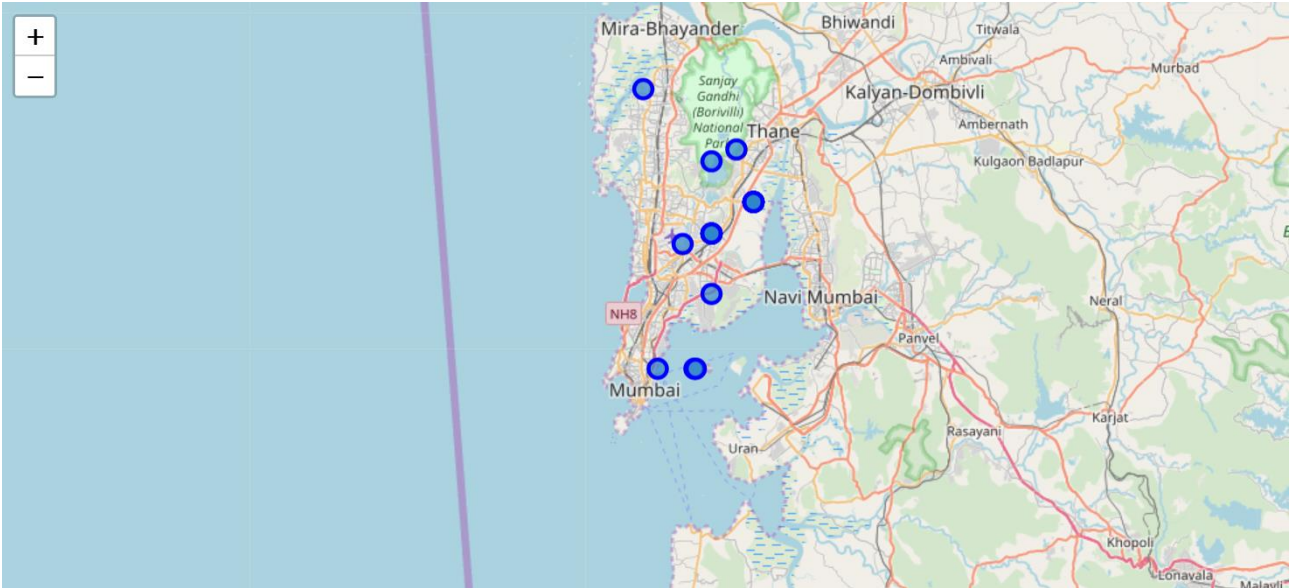A snippet of dataframe if given below.

	neighborhood	pincode	latitude	longitude
0	Bhandup East	400042	19.0832	72.8748
1	Shivaji Nagar (Mumbai), D.M. Colony	400043	18.9596	72.8875
2	Sindhi Society, Chembur H.O	400071	19.0931	72.9049
3	FCI, Chembur Extension, Mahul Road	400074	19.0931	72.9049
4	R.A.Nagar, Best Staff Colony, Pant Nagar	400075	19.0931	72.9049
5	Powai Iit	400076	19.0931	72.9049
6	Rajawadi	400077	18.9596	72.8496
7	J.M. Road, Usha Nagar, Bhandup Ind. Estate, P....	400078	19.0931	72.9049
8	Vikhroli, Psm Colony	400079	18.9596	72.8881
9	Mulund Dd Road, Nahur, Mulund West, S.B. Road,...	400080	19.176	72.9317
10	Mhada Colony, Mulund East	400081	19.1649	72.9052
11	Bhandup Complex, Mulund Colony	400082	19.1247	72.9488
12	Kannamwar Nagar, Tagore Nagar	400083	19.1247	72.9488

Visualization of each neighborhood in Mumbai with Folium creates a map given below.

The blue circles represent the neighbourhoods of IIT Bombay in Mumbai City.
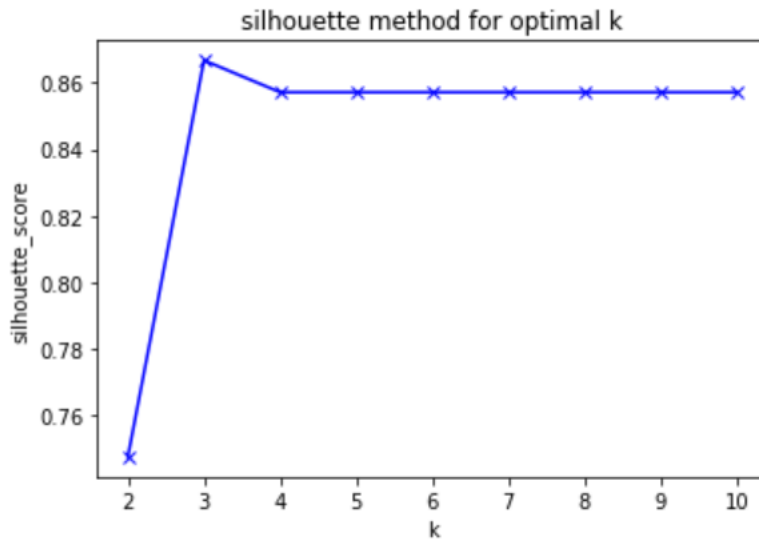

3.3 Foursquare API to find Venues

Foursquare API we get venues around each neighbourhood of IIT Bombay.

The snippet of dataframe containing 10 most popular venues is given below.

	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
0	BARC	Movie Theater	Clothing Store	Playground	Music Venue	Grocery Store	Dessert Shop	Cosmetics Shop	Bus Station	Beer Garden	Airport Lounge
1	Barve Nagar	Movie Theater	Clothing Store	Playground	Music Venue	Grocery Store	Dessert Shop	Cosmetics Shop	Bus Station	Beer Garden	Airport Lounge
2	Bhandup Complex, Mulund Colony	Movie Theater	Clothing Store	Playground	Music Venue	Grocery Store	Dessert Shop	Cosmetics Shop	Bus Station	Beer Garden	Airport Lounge
3	Bhandup East	Grocery Store	Cosmetics Shop	Beer Garden	Airport Lounge	Playground	Music Venue	Movie Theater	Dessert Shop	Clothing Store	Bus Station
4	FCI, Chembur Extension, Mahul Road	Playground	Bus Station	Music Venue	Movie Theater	Grocery Store	Dessert Shop	Cosmetics Shop	Clothing Store	Beer Garden	Airport Lounge

3.4 Clustering

To find the optimal 'K' to use K-means Clustering algorithm , I used the silhouette method.

The graph show the optimal 'K' value as '3'. So, we cluster with K=3.

After applying the clustering algorithm, we get the following dataframe.

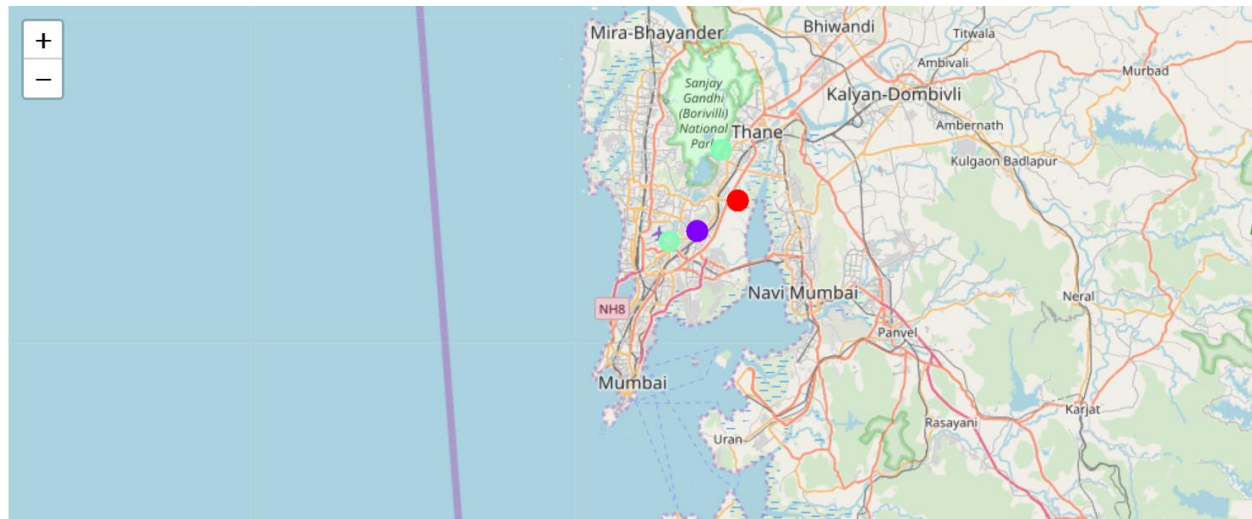|	Neighborhood	pincode	latitude	longitude	ClusterLabels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	Bhandup East	400042	19.0832	72.8748	2	Grocery Store	Cosmetics Shop	Beer Garden	Airport Lounge	Playground	Music Venue	Movie Theater	Dessert Shop	Clothing Store
1	Sindhi Society, Chembur H.O	400071	19.0931	72.9049	1	Playground	Bus Station	Music Venue	Movie Theater	Grocery Store	Dessert Shop	Cosmetics Shop	Clothing Store	Beer Garden
2	FCI, Chembur Extension, Mahul Road	400074	19.0931	72.9049	1	Playground	Bus Station	Music Venue	Movie Theater	Grocery Store	Dessert Shop	Cosmetics Shop	Clothing Store	Beer Garden
3	R.A.Nagar, Best Staff Colony, Pant Nagar	400075	19.0931	72.9049	1	Playground	Bus Station	Music Venue	Movie Theater	Grocery Store	Dessert Shop	Cosmetics Shop	Clothing Store	Beer Garden
4	Powai lit	400076	19.0931	72.9049	1	Playground	Bus Station	Music Venue	Movie Theater	Grocery Store	Dessert Shop	Cosmetics Shop	Clothing Store	Beer Garden

4. Results

4.1 Visualization

Given below is the map of clusters formed out of IIT Bombay's neighborhoods.

4.2 Examining Clusters

The dataframes of each cluster is given below.

Cluster 0

	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
7	Bhandup Complex, Mulund Colony	Movie Theater	Clothing Store	Playground	Music Venue	Grocery Store	Dessert Shop	Cosmetics Shop	Bus Station	Beer Garden	Airport Lounge
8	Kannamwar Nagar, Tagore Nagar	Movie Theater	Clothing Store	Playground	Music Venue	Grocery Store	Dessert Shop	Cosmetics Shop	Bus Station	Beer Garden	Airport Lounge
9	Barve Nagar	Movie Theater	Clothing Store	Playground	Music Venue	Grocery Store	Dessert Shop	Cosmetics Shop	Bus Station	Beer Garden	Airport Lounge
10	BARC	Movie Theater	Clothing Store	Playground	Music Venue	Grocery Store	Dessert Shop	Cosmetics Shop	Bus Station	Beer Garden	Airport Lounge
11	Ghatkopar West, Rifle Range, Sahakar Bhavan	Movie Theater	Clothing Store	Playground	Music Venue	Grocery Store	Dessert Shop	Cosmetics Shop	Bus Station	Beer Garden	Airport Lounge

Cluster 1

	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	Sindhi Society, Chembur H.O	Playground	Bus Station	Music Venue	Movie Theater	Grocery Store	Dessert Shop	Cosmetics Shop	Clothing Store	Beer Garden	Airport Lounge
2	FCI, Chembur Extension, Mahul Road	Playground	Bus Station	Music Venue	Movie Theater	Grocery Store	Dessert Shop	Cosmetics Shop	Clothing Store	Beer Garden	Airport Lounge
3	R.A.Nagar, Best Staff Colony, Pant Nagar	Playground	Bus Station	Music Venue	Movie Theater	Grocery Store	Dessert Shop	Cosmetics Shop	Clothing Store	Beer Garden	Airport Lounge
4	Powai Iit	Playground	Bus Station	Music Venue	Movie Theater	Grocery Store	Dessert Shop	Cosmetics Shop	Clothing Store	Beer Garden	Airport Lounge
5	J.M. Road, Usha Nagar, Bhandup Ind. Estate, P....	Playground	Bus Station	Music Venue	Movie Theater	Grocery Store	Dessert Shop	Cosmetics Shop	Clothing Store	Beer Garden	Airport Lounge

Cluster 2

	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
0	Bhandup East	Grocery Store	Cosmetics Shop	Beer Garden	Airport Lounge	Playground	Music Venue	Movie Theater	Dessert Shop	Clothing Store	Bus Station
6	Mulund Dd Road, Nahur, Mulund West, S.B. Road,...	Music Venue	Dessert Shop	Playground	Movie Theater	Grocery Store	Cosmetics Shop	Clothing Store	Bus Station	Beer Garden	Airport Lounge

5. Discussions

After examining, the areas in Clusters 0, we see that they are more common in terms of movie theatres, clothing stores, dessert shops and departmental stores, whereas cluster 1 areas are more popular for bus stations i.e transport.

So students, who prefer departmental stores proximity, could stay in the Neighborhoods of cluster 0 i.e Bhandup Complex, Mulund Colony, Kannamwar Nagar, Tagore Nagar, Barve Nagar, BARC, Ghatkopar West, Rifle Range, Sahakar Bhavan .

If they prefer quick transportation they can choose Neighborhoods of cluster 1, i.e Sindhi Society, Chembur H.O, FCI, Chembur Extension, Mahul Road, R.A.Nagar, Best Staff Colony, Pant Nagar, Powai Iit, .M. Road, Usha Nagar, Bhandup Ind. Estate.

6. Conclusion

The results of this report provide a general overview of ideal neighbourhoods around a premier institution named IIT Bombay. This report is helpful for students or people who wish to reside nearby the institute.

The link to code of the analysis is given below.

https://github.com/sree-deep/Coursera_Capstone/blob/master/CourseraProject_IITBombay.ipynb