Contents

Introduction/Business Problem:

Mumbai is a city known for its nightlife with loads of popular venues spread across the city which are frequented by youngsters, party goers and celebrities alike. An entrepreneur is interested in starting a venture in this lucrative space and wants to open a Lounge/Gastropub in the city. He would like to identify an appropriate location in the city to start this venture based on the venues in the locality and their popularity. This will form our business problem and we will use data from Foursquare API and other sources to explore the Nightlife venues in city to arrive at an appropriate location.

Data Used:

- We will use the dataset obtained from website http://creativecommons.org/licenses/by/3.0/ which contains the latitude longitude data for India at a pincode level. We will subset the above dataset in order to obtain latitude & longitude only for Mumbai at a pincode level which is our area of interest for this exercise.
  This is how the data will look after importing:

	postal code	Neighborhood	State	District	latitude	longitude
0	400001	Mumbai G.P.O.	Maharashtra	Mumbai	18.938536	72.836334
1	400001	Bazargate	Maharashtra	Mumbai	18.938536	72.836334
2	400001	Town Hall (Mumbai)	Maharashtra	Mumbai	18.938536	72.836334
3	400001	Tajmahal	Maharashtra	Mumbai	18.938536	72.836334
4	400001	Stock Exchange	Maharashtra	Mumbai	18.938536	72.836334
5	400001	M.P.T.	Maharashtra	Mumbai	18.938536	72.836334
6	400002	Kalbadevi	Maharashtra	Mumbai	18.948366	72.825935
7	400002	S. C. Court	Maharashtra	Mumbai	18.948366	72.825935
8	400002	Ramwadi	Maharashtra	Mumbai	18.948366	72.825935
9	400002	Thakurdwar	Maharashtra	Mumbai	18.948366	72.825935

  As seen above, multiple localities have a common pincode. We will combine this data at a pincode level and bring it in the below format:

|	postal code	Neighborhood	State	District	latitude	longitude
0	400001	Mumbai G.P.O.,Bazargate,Town Hall (Mumbai),Tajmahal,Stock Exchange,M.P.T.	Maharashtra	Mumbai	18.938536	72.836334
1	400002	Kalbadevi,S. C. Court,Ramwadi,Thakurdwar	Maharashtra	Mumbai	18.948366	72.825935
2	400003	Mandvi (Mumbai),Null Bazar,B.P.Lane,Masjid	Maharashtra	Mumbai	18.950000	72.833300
3	400004	Girgaon,Opera House,Ambewadi (Mumbai),Charni Road,Chaupati,Madhavbaug	Maharashtra	Mumbai	18.950000	72.816700
4	400005	Colaba,Holiday Camp,V.W.T.C.,Colaba Bazar,Asvini	Maharashtra	Mumbai	18.906900	72.810600
5	400006	Malabar Hill	Maharashtra	Mumbai	18.950000	72.783300
6	400007	Tardeo,Grant Road,S V Marg,N.S.Patkar Marg,Bharat Nagar (Mumbai)	Maharashtra	Mumbai	18.966700	72.816700
7	400008	Falkland Road,Mumbai Central,M A Marg,J.J.Hospital,Kamathipura	Maharashtra	Mumbai	18.967140	72.828659
8	400009	Noor Baug,Princess Dock,Chinchbunder	Maharashtra	Mumbai	18.958296	72.838943
9	400010	Mazgaon,Mazgaon Road,V K Bhavan,Mazgaon Dock,Dockyard Road	Maharashtra	Mumbai	18.970188	72.845963

- We will use Foursquare API to explore the venues falling in the 'Nightlife' category around various neighbourhoods in the city.
  This data will be exported in a pandas dataframe and will be brought in the below form for further analysis:

	postal code	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	id	Venue Latitude	Venue Longitude	Venue Category
0	400001	Mumbai G.P.O.,Bazargate,Town Hall (Mumbai),Tajmahal,Stock Exchange,M.P.T.	18.938536	72.836334	Town House Cafe	5263e1ba11d265711e8024bf	18.938550	72.833464	Bar
1	400001	Mumbai G.P.O.,Bazargate,Town Hall (Mumbai),Tajmahal,Stock Exchange,M.P.T.	18.938536	72.836334	The Clearing House	5810c12738faa0b87f99d404	18.935328	72.838263	Lounge
2	400001	Mumbai G.P.O.,Bazargate,Town Hall (Mumbai),Tajmahal,Stock Exchange,M.P.T.	18.938536	72.836334	5 Spice	4be18fe88815b713ef9e6406	18.933491	72.835955	Chinese Restaurant
3	400001	Mumbai G.P.O.,Bazargate,Town Hall (Mumbai),Tajmahal,Stock Exchange,M.P.T.	18.938536	72.836334	Press Club Mumbai	4dca583cae607b31c0bdf19c	18.940721	72.832469	Bar
4	400001	Mumbai G.P.O.,Bazargate,Town Hall (Mumbai),Tajmahal,Stock Exchange,M.P.T.	18.938536	72.836334	Sunlight Bar	4e19a3221f6eb9559885364c	18.944448	72.829234	Bar
5	400001	Mumbai G.P.O.,Bazargate,Town Hall (Mumbai),Tajmahal,Stock Exchange,M.P.T.	18.938536	72.836334	Bottle bar	517830f4498e1f75ba5bb195	18.930246	72.833423	Lounge
6	400001	Mumbai G.P.O.,Bazargate,Town Hall	18.938536	72.836334	The Bar	567ba37f498e1d6a14d30821	18.930338	72.833421	Cocktail

The 89 neighbourhoods in the city present in the above data are plotted on the map below
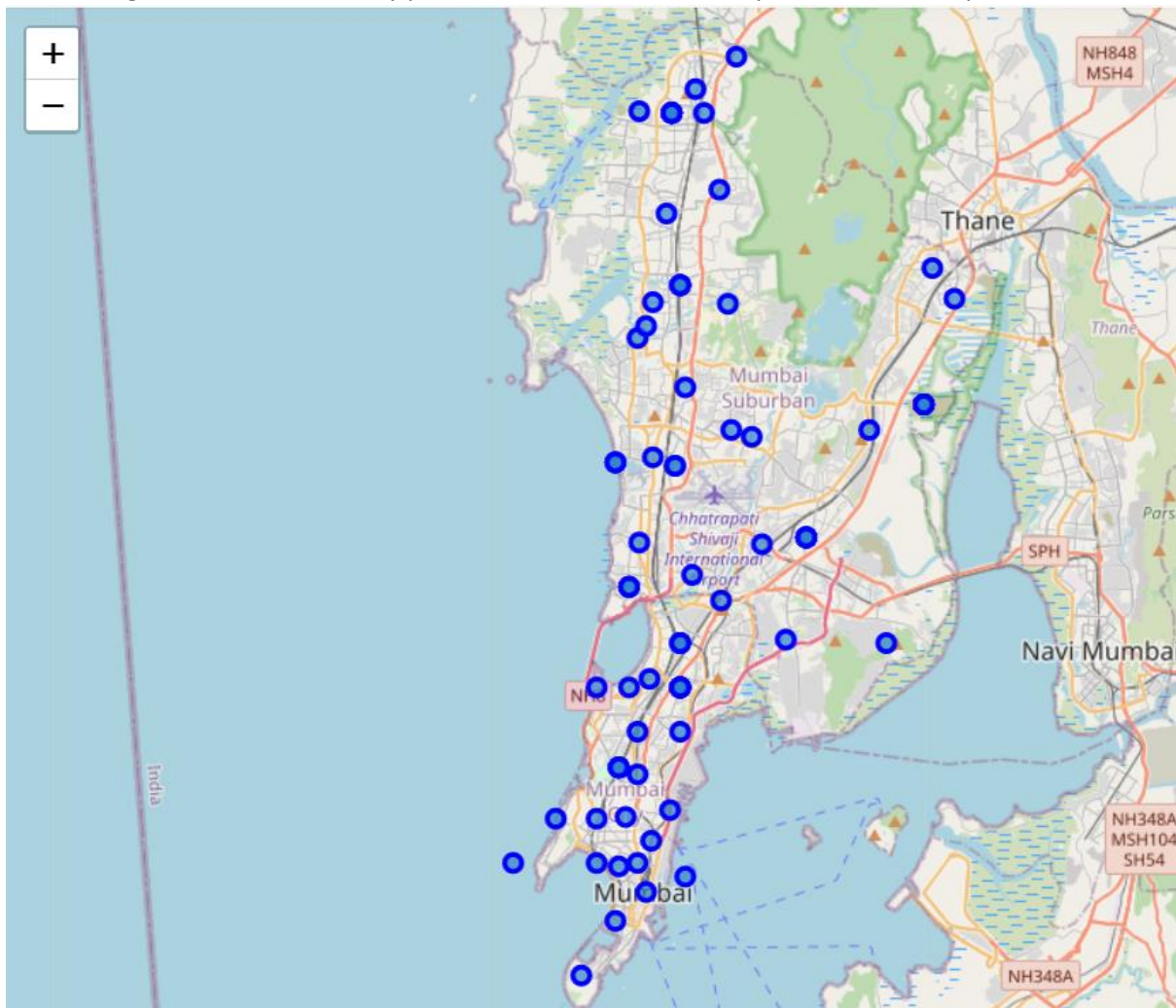
Methodology:

Now that the data collected and cleaned, we can discuss the methodology and the approach we are going to take to solve our problem i.e. recommend an optimum location for opening a Gastropub/Lounge.

We will look at density of all the venue categories in each neighbourhood and then run a k-means clustering algorithm to divide the neighbourhoods into clusters. Also, we will look at the top 10 popular categories in each neighbourhood to understand their preferences.
For deciding the optimum location of our Gastropub business, we will evaluate the clusters from 2 perspectives:
a) **Location:** The selected clusters must be spread out across the city so that the stakeholders have enough flexibility to choose from the various options.
b) **Popularity of our venue category:** We will also look at the popularity of various venue categories in each neighbourhood to understand their nightlife venue preferences. We need to ensure that the within the cluster that we select our venue category i.e. Gastropub is popular

Data analysis:

Let us also have a look at the count of various venue categories present in our data. We can see from the data below that Gastropubs fall in top 10 venue categories.
Also, other categories like bars and lounges might be saturated and we may face stiff competition if we enter into these categories. A category like Gastropub which is popular yet not saturated can give us a nice niche to operate in.

Venue Category	Count
Bar	189
Lounge	139
Pub	58
Nightclub	42
Hotel Bar	39
Hookah Bar	35
Cocktail Bar	21
Gastropub	20
Sports Bar	16
Restaurant	16
Café	16
Brewery	14
Hotel	14
Indian Restaurant	13
Wine Bar	10
Beer Garden	9
Italian Restaurant	9
Whisky Bar	8
Asian Restaurant	5
Mediterranean Restaurant	5

Tapas Restaurant	5
Karaoke Bar	3
Bistro	3
Gym	3
Coffee Shop	3
Chinese Restaurant	3
Music Venue	3
Tea Room	3
Food	3
Seafood Restaurant	3
Speakeasy	2
Dive Bar	2
Sake Bar	2
Campground	1
Dim Sum Restaurant	1
Mexican Restaurant	1
Molecular Gastronomy Restaurant	1
Gym / Fitness Center	1
Event Space	1
Athletics & Sports	1
Grocery Store	1
Beer Bar	1
Nightlife Spot	1

From the above analysis we also see that a total 726 venues belonging to 43 distinct nightlife categories were found in Mumbai neighbourhoods.
Out of the 89 neighbourhoods in our dataset we could find nightlife venues in only 78 neighbourhoods.
This might be due to lack of data in Foursquare about nightlife venues in these localities.

	postal code	Neighborhood	State	District	latitude	longitude
0	400006	Malabar Hill	Maharashtra	Mumbai	18.9500	72.7833
1	400013	Delisle Road	Maharashtra	Mumbai	18.9448	72.8524
2	400063	Sharma Estate,Goregaon East	Maharashtra	Mumbai	19.1624	72.8694
3	400082	Mulund Colony,Bhandup Complex	Maharashtra	Mumbai	19.1247	72.9488
4	400083	Kannamwar Nagar,Tagore Nagar	Maharashtra	Mumbai	19.1247	72.9488
5	400084	Barve Nagar	Maharashtra	Mumbai	19.1247	72.9488
6	400085	BARC	Maharashtra	Mumbai	19.1247	72.9488
7	400086	Ghatkopar West,Rifle Range,Sahakar Bhavan	Maharashtra	Mumbai	19.1247	72.9488
8	400087	Sandeepany Sadhanalya,NITIE	Maharashtra	Mumbai	19.1247	72.9488
9	400088	Trombay,T.F.Donar,Govandi	Maharashtra	Mumbai	19.0333	72.9333
10	400089	Tilak Nagar (Mumbai),Chembur Rs	Maharashtra	Mumbai	19.1247	72.9488

We will later cluster these 11 neighbourhoods into a separate cluster as we will not be able to make any decisions regarding these neighbourhoods due to lack of data. Let's see which neighbourhoods fall in this cluster.

Part 1: Top 10 venue categories in each neighbourhood

- In the very first step we will encode our venue categories into numeric data using one hot encoding.
- Then we will find the average number of each venue categories present in each neighbourhood.
- Basis the average number of venue categories per neighbourhood we will arrive at top 10 categories in each neighbourhood

	Neighborhood	Asian Restaurant	Athletics & Sports	Bar	Beer Bar	Beer Garden	Bistro	Brewery	Café	Campground	...	Pub	Restaurant	Sake Bar
0	Agripada,Chinchpokli,Jacob Circle,Haines Road	0.0	0.0	0.500000	0.0	0.000000	0.00000	0.0	0.00000	0.25	...	0.000000	0.00000	0.0
1	Antop Hill,C G S Colony,B P T Colony,Wadala Truck Terminal	0.0	0.0	0.333333	0.0	0.000000	0.00000	0.0	0.00000	0.00	...	0.166667	0.00000	0.0
2	Anushakti Nagar	0.0	0.0	0.000000	0.0	0.000000	0.00000	0.0	0.00000	0.00	...	0.000000	0.25000	0.0
3	Audit Bhavan,Kherwadi,Bandra(East),B.N. Bhavan,Government Colony	0.0	0.0	0.400000	0.0	0.000000	0.00000	0.0	0.00000	0.00	...	0.100000	0.00000	0.0
4	Azad Nagar (Mumbai),Andheri	0.0	0.0	0.600000	0.0	0.000000	0.00000	0.0	0.00000	0.00	...	0.000000	0.00000	0.0
5	Bandra West	0.0	0.0	0.281250	0.0	0.031250	0.03125	0.0	0.03125	0.00	...	0.156250	0.03125	0.0
6	Bhandup East	0.0	0.0	0.275862	0.0	0.034483	0.00000	0.0	0.00000	0.00	...	0.034483	0.00000	0.0
7	Borivali	0.0	0.0	0.250000	0.0	0.000000	0.00000	0.0	0.00000	0.00	...	0.250000	0.00000	0.0
8	Borivali West	0.0	0.0	0.250000	0.0	0.000000	0.00000	0.0	0.00000	0.00	...	0.250000	0.25000	0.0
9	Chakala Midc	0.0	0.0	0.000000	0.0	0.000000	0.00000	0.0	0.00000	0.00	...	0.000000	0.25000	0.0

10 rows × 44 columns

Part 2: Clustering

We now begin the steps for clustering our data. We will divide the neighbourhoods in 4 clusters as follows:

- We use the average number of venue categories per location for our clustering exercise
- We will then merge clustered data & top 10 venue categories data in order to draw conclusions about popularity of venue categories in different clusters

The resulting dataset is shown below:

Neighborhood	latitude	longitude	Cluster Labels	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 D.,Bazargate,Town Hall umbai),Tajmahal,Stock Exchange,M.P.T.	18.938536	72.836334	2.0	Bar	Lounge	Chinese Restaurant	Indian Restaurant	Sake Bar	Cocktail Bar	Coffee Shop	Grocery Store	Gastropub	Food
Kalbadevi,S. C. t,Ramwadi,Thakurdwar	18.948366	72.825935	2.0	Lounge	Bar	Gastropub	Wine Bar	Coffee Shop	Gym	Grocery Store	Food	Event Space	Dive Bar
Mandvi (Mumbai),Null Bazar,B.P.Lane,Masjid	18.950000	72.833300	2.0	Bar	Hookah Bar	Lounge	Wine Bar	Coffee Shop	Gym	Grocery Store	Gastropub	Food	Event Space
Girgaon,Opera House,Ambewadi (Mumbai),Charni Chaupati,Madhavbaug	18.950000	72.816700	2.0	Hookah Bar	Indian Restaurant	Gastropub	Lounge	Coffee Shop	Gym	Grocery Store	Food	Event Space	Dive Bar
Colaba,Holiday Camp,V.W.T.C.,Colaba Bazar,Asvini	18.906900	72.810600	2.0	Bar	Lounge	Wine Bar	Coffee Shop	Gym	Grocery Store	Gastropub	Food	Event Space	Dive Bar

We can plot these clusters on a map in order to visualize their locations and analysing each cluster in order to make our final decision regarding the location of the Gastropub.

Cluster label	Colour	Count of Neighbourhoods
0	Red	7
1	Green	13
2	Blue	58
3	Yellow	11
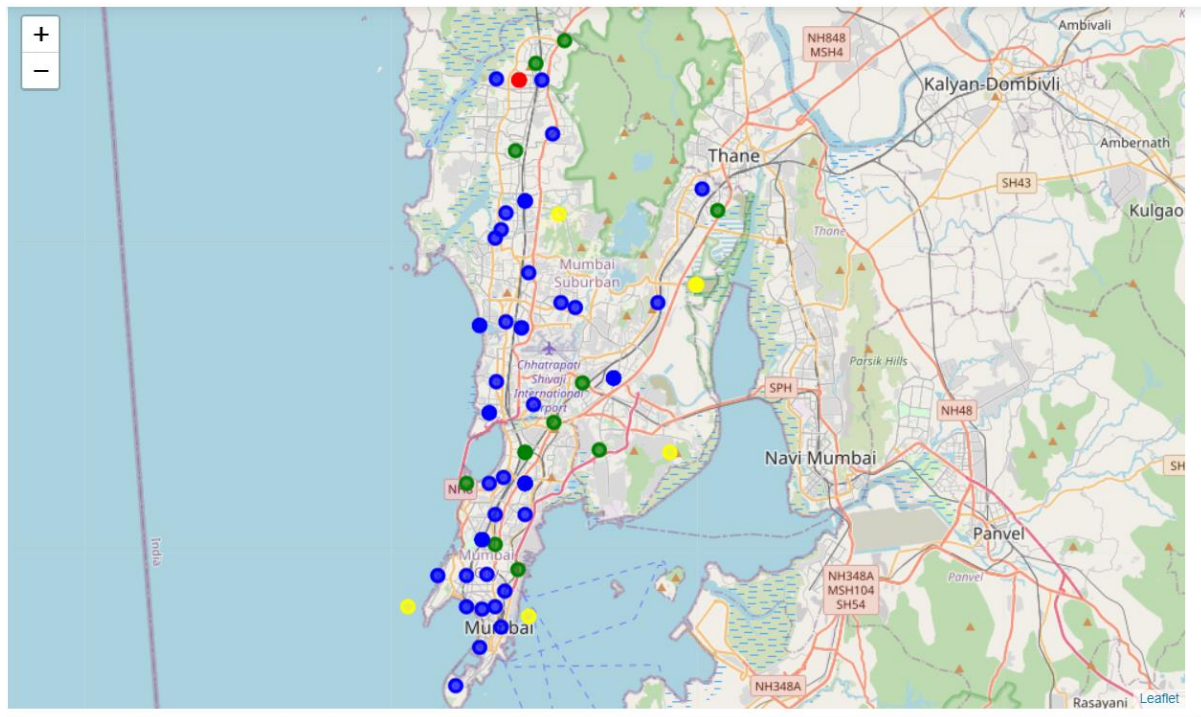
Results & Discussion:

From the above analysis we see that 58 out of 89 neighbourhoods fall in cluster 2. Cluster 1 & 0 has 13 & 7 neighbourhoods respectively. As pointed out earlier, cluster 3 is a cluster created to put together all the neighbourhoods where we could not find any venue information and hence this will not be used for making any decisions.

For deciding the optimum location of our Gastropub business, we will evaluate the clusters from 2 perspectives: a.) Location b.) Popularity of our venue category

a) **Location:** From map we can see that cluster 2 is spread quite uniformly throughout the city making it a favourable cluster for opening our business. This gives us multiple options in the city and we can select a location that suits our budget. The places in South Mumbai will tend to be much more costly than those in the Western suburbs.

b) **Popularity:** Based on the information available for the top 10 venue categories we can see that Gastropubs feature predominantly in cluster 2 & cluster 0. However, cluster 0 has only 7

neighbourhoods and locations are very near to each other (they have same lat longs in our data and are represented by the single red dot in the cluster map). Choosing cluster 0 will limit our options and also we cannot rely on the data as the venue information for these locations is also very limited.

From the above analysis we can say that **cluster 2** neighbourhoods can be good potential candidates for opening a Gastropub given the location options and the popularity of our venue category. Specifically, we would recommend locations in cluster 2 where Gastropubs are present in top 10 locations.

Thus, we will end up with 46 locations to choose from in cluster 2 as shown below. The final decision will be made in discussion with the stakeholders and additional considerations like location cost, local permits, zoning details etc. will come into play which have not been considered for our analysis.

postal code	Neighbourhood	latitude	longitude
400001	Mumbai G.P.O.,Bazargate,Town Hall (Mumbai),Tajmahal,Stock Exchange,M.P.T.	18.93854	72.83633
400002	Kalbadevi,S. C. Court,Ramwadi,Thakurdwar	18.94837	72.82594
400003	Mandvi (Mumbai),Null Bazar,B.P.Lane,Masjid	18.95	72.8333
400004	Girgaon,Opera House,Ambewadi (Mumbai),Charni Road,Chaupati,Madhavbaug	18.95	72.8167
400005	Colaba,Holiday Camp,V.W.T.C.,Colaba Bazar,Asvini	18.9069	72.8106
400007	Tardeo,Grant Road,S V Marg,N.S.Patkar Marg,Bharat Nagar (Mumbai)	18.9667	72.8167
400009	Noor Baug,Princess Dock,Chinchbunder	18.9583	72.83894
400014	Naigaon (Mumbai),Dadar,Dadar Colony	19.0201	72.8381
400021	Nariman Point,New Yogakshema	18.9274	72.8241
400025	Prabhadevi,New Prabhadevi Road	19.0164	72.8294
400026	Cumballa Sea Face,Gowalia Tank,Dr Deshmukh Marg,Cumballa Hill	18.9667	72.8
400030	Worli Police Camp,Century Mill,Worli Colony,Worli Sea Face	19.0167	72.85
400031	Wadala,Kidwai Nagar (Mumbai),Wadala Rs	19.0167	72.85
400032	High Court Building (Mumbai),Mantralaya (Mumbai),Secretariate	19.0167	72.85
400033	Tank Road,Reay Road,Cotton Exchange,Kalachowki,L B S N E collage	19.0167	72.85
400034	Tulsiwadi,Haji Ali	19.0167	72.85
400035	Rajbhavan (Mumbai)	19.0167	72.85
400037	Antop Hill,C G S Colony,B P T Colony,Wadala Truck Terminal	19.0167	72.85
400050	Bandra West	19.05517	72.82995
400052	V.P. Road,Danda,Khar Colony,Khar Delivery	19.07171	72.83411

400053	Azad Nagar (Mumbai),Andheri	19.1309	72.8526
400054	Santacruz Central,Santacruz(West)	19.05517	72.82995
400055	Vakola,Santacruz(East)	19.05517	72.82995
400056	Vileparle(West),IRLA	19.1045	72.8395
400057	Vileparle Railway Station,Hanuman Road,Vileeparle (East)	19.10074	72.8485
400058	H.M.P. School,Andheri Railway Station	19.10074	72.8485
400060	Jogeshwari East	19.1699	72.8504
400061	Madh,Vesava	19.15	72.8333
400065	S R P F Camp,Nagari Niwara,Aareymilk Colony	19.1699	72.8504
400066	Rajendra Nagar (Mumbai),Borivali East,Magthane,Daulat Nagar (Mumbai),S. K.Nagar	19.1699	72.8504
400067	Kandivali West,Kandivali RS,Charkop	19.1699	72.8504
400072	Vihar Road,Sakinaka	19.0737	72.9009
400074	Chembur Extension,FCI,Mahul Road	19.0737	72.9009
400075	R.A.Nagar,Pant Nagar,Best Staff Colony	19.0737	72.9009
400076	Powai Iit	19.0737	72.9009
400077	Rajawadi	19.0737	72.9009
400078	P.H. Colony,Bhandup Ind. Estate,Usha Nagar,J.M. Road,Bhandup West	19.0737	72.9009
400079	Vikhroli,Psm Colony	19.1149	72.9267
400080	S.B. Road,Nahur,Mulund Dd Road,Nehru Road (Mumbai),Mulund West	19.176	72.9522
400091	Borivali	19.235	72.8598
400092	Borivali West	19.2361	72.8338
400101	Kandivali East	19.20586	72.86612
400102	Jogeshwari West,Oshiwara	19.15404	72.83663
400104	Motilal Nagar,Goregaon (Mumbai),Goregaon RS,Bangur Nagar	19.16305	72.8393

Let us plot the above list of preferred locations on the map
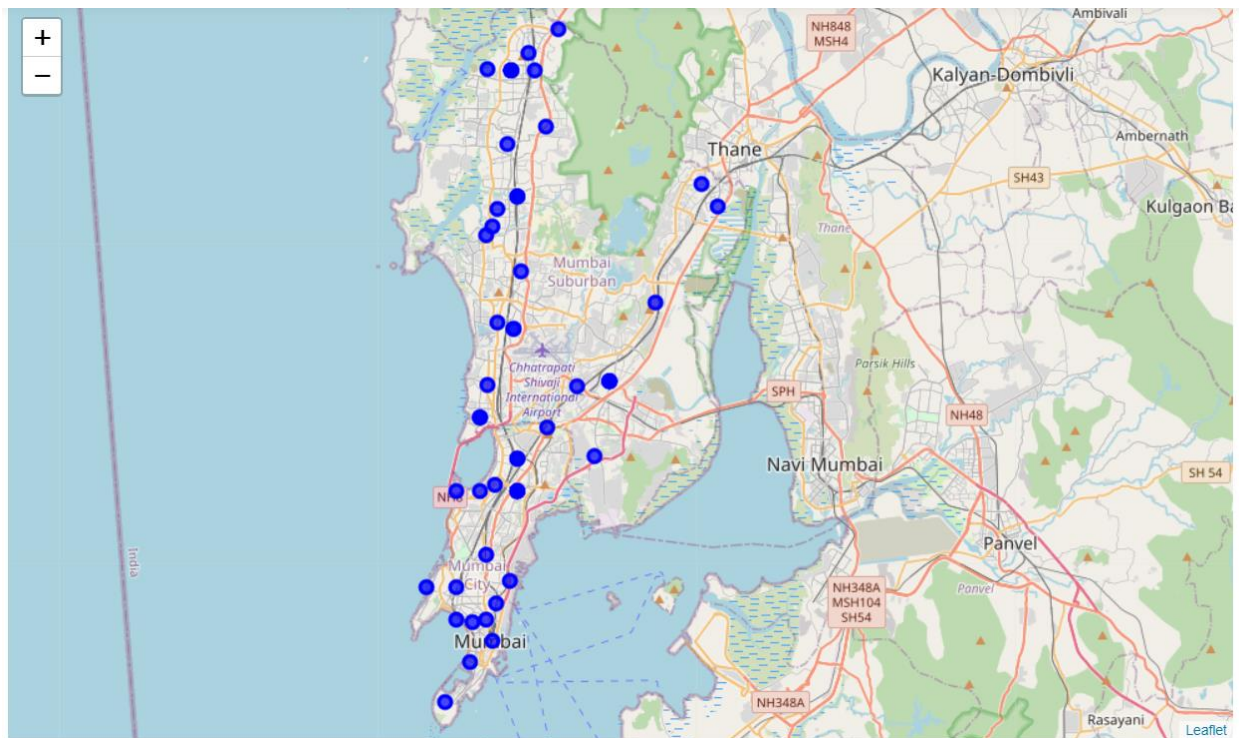
Conclusion:

We started with a problem statement of finding an optimum location for starting a business in **'Nightlife'** category in Mumbai. From the data we concluded that we will be entering **Gastropub** category which gives us a unique niche. Also, basis the clustering analysis and popular venues across locations we have concluded that we will open the Gastropub in **cluster 2**. This will give us a shortlist of **45 locations** across the city to choose from.