Finding suitable locations to open a Cinema hall in Bengaluru, INDIA

Yashwanth S

1. Introduction / Business Problem

The aim of this project is to find suitable locations to open a cinema hall in the Bengaluru, India. The first requirement is that the new cinema hall should be easily accessible by its prospective customers and more specifically it should be located near a metro station. The number of cinema halls already existing in an area should also be considered so that fierce competition be avoided if possible. Apart from the obvious intended stakeholders, similar methodology could be used for other specific types of media-businesses. It can serve as an initial starting point of locations to consider to start their business. For the project objectives to be achieved, python geolocation libraries will used, along with the Foursquare API. Also, in order to create clusters of similar candidate locations, the K-Means machine learning clustering algorithm will be used.

2. Data

The necessary data for this project, based on the above stated requirements, are:

- The metro stations in the Bengaluru area
- Number of existing Cinema halls
- In addition, the distance to the nearest Cinema hall for every metro station will be used

In order to obtain the data, a combination of the **geopy** Python library and the **Foursquare**
API will be used:
1. *'Majestic'* will be considered as the center of Bengaluru. It is indeed one of the
most central location in the city. I will obtain its geospatial coordinates using the geopy
library

2. Having the coordinates of the ‘center’ of Bengaluru, the Foursquare API will be used to retrieve data for all the metro stations in radius of 30kms
3. To find the existing Cinema halls near the metro stations, the Foursquare API will again be used for every station. I will obtain data for all the Cinema halls located in a radius of 2000 meters of every metro station
Using the collected data, I will calculate the number of existing Cinema halls near each station. **I will also be able to determine the minimum distance to a Cinema hall for every metro station** from the 3rd step of the above process. This minimum distance to every bus station from a Cinema hall, along with the number of already existing Cinema halls near the station will be used as input to KMeans clustering algorithm to obtain the clusters of areas (metro stations).

3. Methodology

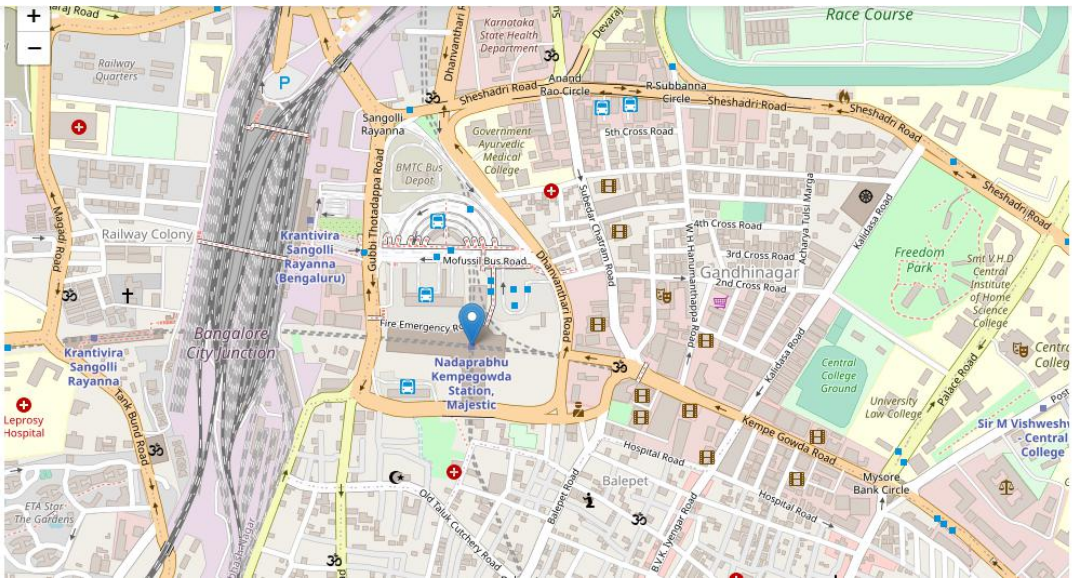
The objective of this project is to obtain information about metro stations in the greater metropolitan Bangalore area with potential for opening a theatre, and having as criteria:

Low number of already existing cinema halls

Minimum distance of each station to its nearby cinema theatres

The steps I followed to identify potential areas (metro stations) were: \

1. Considered Majestic as the ‘center’ of Bangalore (indeed probably the most central location of the city) and acquired its latitude and longitude geospatial
coordinates.



2.Based on the coordinates of Majestic, I obtained information about metro stations in a radius of 35 km using the Foursquare API.
At this stage I **removed from the above dataset 10rows of data** that although they are identified as 'Metro stations' by the Foursquare API, they are only used as depots or maintenance gathering for the metro carriages.

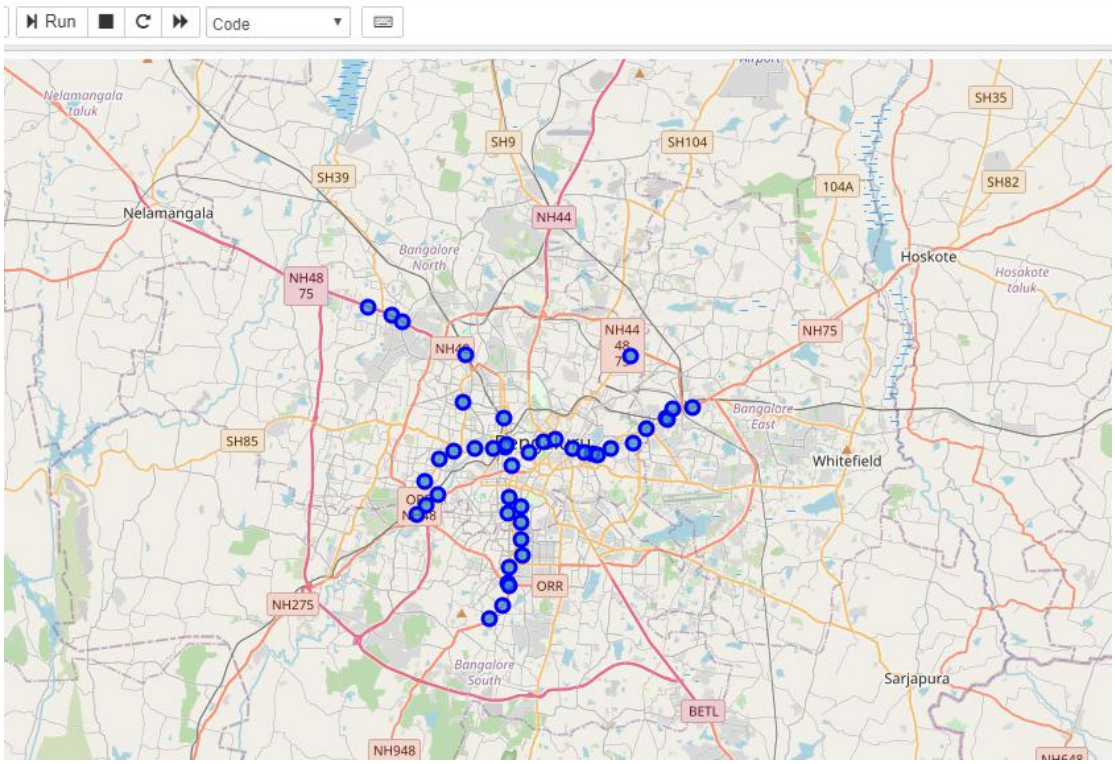
Out[14]:

	name	lat	lng	distance	postalCode	venue_type
0	MG Road Metro Station	12.975495	77.606777	3677	560001	Metro Station
1	Trinity Metro Station	12.972948	77.617340	4833	560001	Metro Station
2	Baiyappanahalli Metro Station	12.991302	77.655208	9097	NaN	Light Rail Station
3	Indiranagar Metro Station	12.978266	77.638634	7138	NaN	Metro Station
4	Mysore Road Metro Station	12.946738	77.530403	5623	NaN	Metro Station
5	Kempegowda Metro Station	12.976134	77.571482	158	NaN	Metro Station
6	Bangalore City Metro Station	12.975862	77.565700	778	560023	Metro Station
7	Chickpet Metro Station	12.966987	77.574840	993	560053	Metro Station
8	Shri M Visvesvaraya Metro Station	12.974097	77.584015	1221	NaN	Metro Station
9	Mantri Square Sampige Road Metro Station	12.991217	77.570917	1739	560003	Metro Station

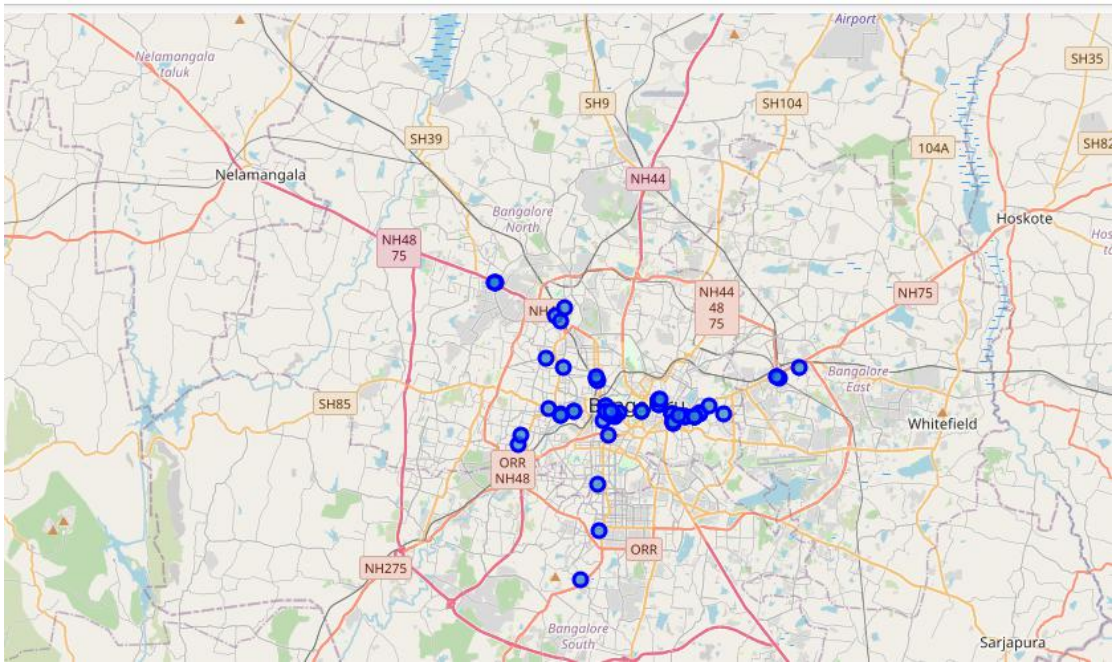
```
In [15]: stations_df.shape
```

```
Out[15]: (44, 6)
```

A visualization of the remaining metro stations on Bangalore map:



3.After the collection of metro stations information, I again utilized the Foursquare API to locate all the existing cinema halls in a radius of 2000 meters from each station. The resulting subcategories of businesses found were:



4. We look at the stations and match the corresponding cimeha halls and merge them to a single dataframe

Out[20]:

	station	lat	lng	Venue	Venue Latitude	Venue Longitude	Distance from Station	Venue Category
0	MG Road Metro Station	12.975495	77.606777	Swagath ONYX Theatre	12.974606	77.609505	311	Movie Theater
1	MG Road Metro Station	12.975495	77.606777	Fame Shankarnag	12.974602	77.609400	301	Movie Theater
2	MG Road Metro Station	12.975495	77.606777	BRV Theatre	12.979410	77.602287	653	Movie Theater
3	MG Road Metro Station	12.975495	77.606777	INOX Magrath Road	12.970127	77.609968	690	Multiplex
4	MG Road Metro Station	12.975495	77.606777	Screen 2	12.970259	77.609950	676	Multiplex
5	MG Road Metro Station	12.975495	77.606777	Screen 4 INOX Garuda	12.970080	77.609938	693	Multiplex
6	MG Road Metro Station	12.975495	77.606777	Garuda Mall	12.970140	77.609978	689	Shopping Mall
7	Trinity Metro Station	12.972948	77.617340	Fame Lido Mall	12.972786	77.616417	101	Movie Theater
8	Trinity Metro Station	12.972948	77.617340	Inox Lido Mall	12.973078	77.621315	431	Multiplex
9	Trinity Metro Station	12.972948	77.617340	Fame Lido	12.973042	77.621498	451	Multiplex

In [21]:
```
gyms_venues_df.shape
```

Out[21]: (79, 8)

5. We also locate the distance between the cinema hall and the nearest two metro stations and find distance between them

	station	lat	lng	Venue	Latitude	Longitude	Station_x	Category	Station_y
0	MG Road Metro Station	12.975495	77.606777	Swagath ONYX Theatre	12.974606	77.609505	311	Movie Theater	301
1	MG Road Metro Station	12.975495	77.606777	Fame Shankarnag	12.974602	77.609400	301	Movie Theater	301
2	MG Road Metro Station	12.975495	77.606777	BRV Theatre	12.979410	77.602287	653	Movie Theater	301
3	MG Road Metro Station	12.975495	77.606777	INOX Magrath Road	12.970127	77.609968	690	Multiplex	301
4	MG Road Metro Station	12.975495	77.606777	Screen 2	12.970259	77.609950	676	Multiplex	301
5	MG Road Metro Station	12.975495	77.606777	Screen 4 INOX Garuda	12.970080	77.609938	693	Multiplex	301
6	MG Road Metro Station	12.975495	77.606777	Garuda Mall	12.970140	77.609978	689	Shopping Mall	301
7	Trinity Metro Station	12.972948	77.617340	Fame Lido Mall	12.972786	77.616417	101	Movie Theater	101
8	Trinity Metro Station	12.972948	77.617340	Inox Lido Mall	12.973078	77.621315	431	Multiplex	101
9	Trinity Metro Station	12.972948	77.617340	Fame Lido	12.973042	77.621498	451	Multiplex	101
10	Indiranagar Metro Station	12.978266	77.638634	Fame Lido	12.974541	77.636075	499	Multiplex	499
11	Kempegowda Metro Station	12.976134	77.571482	Movieland Theater	12.978798	77.575473	524	Movie Theater	416
12	Kempegowda Metro Station	12.976134	77.571482	Santosh Theatre	12.974840	77.576178	529	Movie Theater	416
13	Kempegowda Metro Station	12.976134	77.571482	Abhinay theatre	12.973469	77.577840	750	Movie Theater	416
14	Kempegowda Metro Station	12.976134	77.571482	Triveni Theatre	12.975761	77.575299	416	Indie Movie Theater	416
15	Kempegowda Metro Station	12.976134	77.571482	Tribhuvan Theatre	12.975982	77.577605	664	Indie Movie Theater	416
16	Bangalore City Metro Station	12.975862	77.565700	Anjan Theater	12.975521	77.558984	729	Indie Movie Theater	729

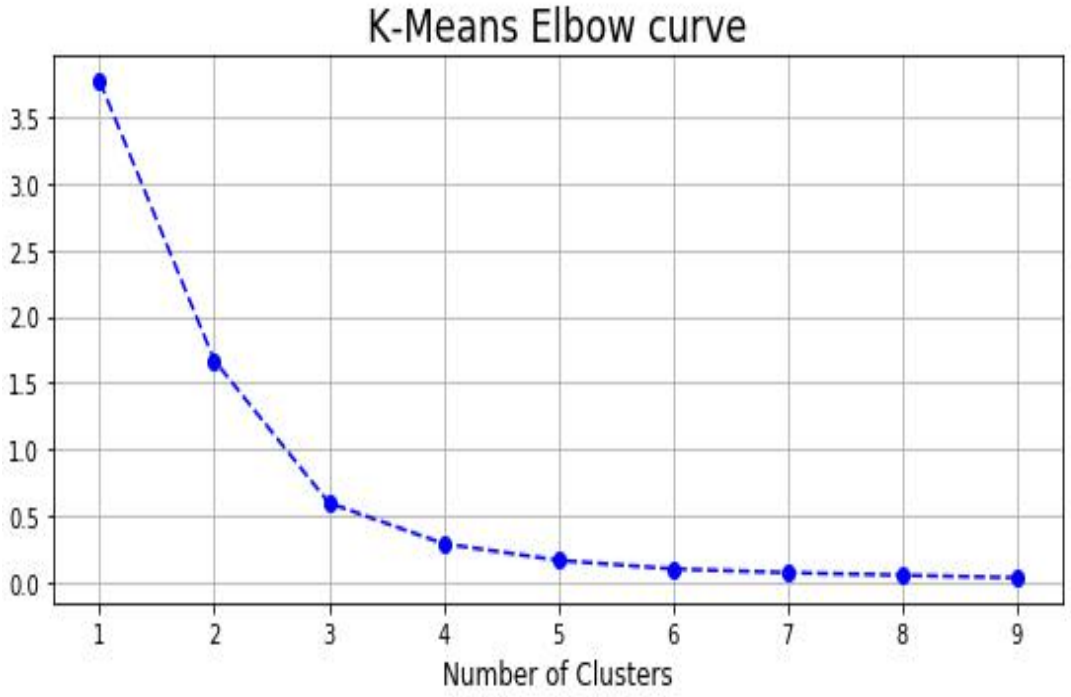
.

station	lat	lng	Venue	Venue Latitude	Venue Longitude	Distance from Station_x	Venue Category	Distance from Station_y
Banashankari Metro Station	1	1	1	1	1	1	1	1
Bangalore City Metro Station	1	1	1	1	1	1	1	1
Byappanahalli Metro Station	1	1	1	1	1	1	1	1
Chickpet Metro Station	2	2	2	2	2	2	2	2
Cubbon Park Metro Station	4	4	4	4	4	4	4	4
Dasarahalli metro station	2	2	2	2	2	2	2	2
Hosahalli Metro	2	2	2	2	2	2	2	2
Indiranagar Metro Station	1	1	1	1	1	1	1	1
Kempegowda Metro Station	5	5	5	5	5	5	5	5
Konanakunte Cross	1	1	1	1	1	1	1	1
M G Road	7	7	7	7	7	7	7	7
MG Road Metro Station	7	7	7	7	7	7	7	7
Magadi Road Metro Station	3	3	3	3	3	3	3	3
Majestic / Kempegowda Bus Stand	4	4	4	4	4	4	4	4

6. We later normalize the distance between the theatres and the metro station and include it in the dataframewhich will be used for clustering

	Venue	Venue Latitude	Venue Longitude	Distance from Station_x	Venue Category	Norm Min Distance from Station
0	Swagath ONYX Theatre	12.974606	77.609505	311	Movie Theater	-0.029455
1	Fame Shankarnag	12.974602	77.609400	301	Movie Theater	-0.029455
2	BRV Theatre	12.979410	77.602287	653	Movie Theater	-0.029455
3	INOX Magrath Road	12.970127	77.609968	690	Multiplex	-0.029455
4	Screen 2	12.970259	77.609950	676	Multiplex	-0.029455
5	Screen 4 INOX Garuda	12.970080	77.609938	693	Multiplex	-0.029455
6	Garuda Mall	12.970140	77.609978	689	Shopping Mall	-0.029455
7	Fame Lido Mall	12.972786	77.616417	101	Movie Theater	-0.304180
8	Inox Lido Mall	12.973078	77.621315	431	Multiplex	-0.304180
9	Fame Lido	12.973042	77.621498	451	Multiplex	-0.304180
10	Fame Lido	12.974541	77.636075	499	Multiplex	0.242523
11	Movieland Theater	12.978798	77.575473	524	Movie Theater	0.128512
12	Santosh Theatre	12.974840	77.576178	529	Movie Theater	0.128512
13	Abhinay theatre	12.973469	77.577840	750	Movie Theater	0.128512
14	Triveni Theatre	12.975761	77.575299	416	Indie Movie Theater	0.128512
15	Tribhuvan Theatre	12.975982	77.577605	664	Indie Movie Theater	0.128512
16	Anjan Theater	12.975521	77.558984	729	Indie Movie Theater	0.558457
17	Q Cinemas, Park Square, ITPL	12.971427	77.573794	507	Multiplex	0.128512

7.The **K-Means Machine Learning clustering algorithm** will be used to divide the stations and cinema hall data set into clusters of similar locations.The elbow method will be used to find the most suitable number of clusters.



Results

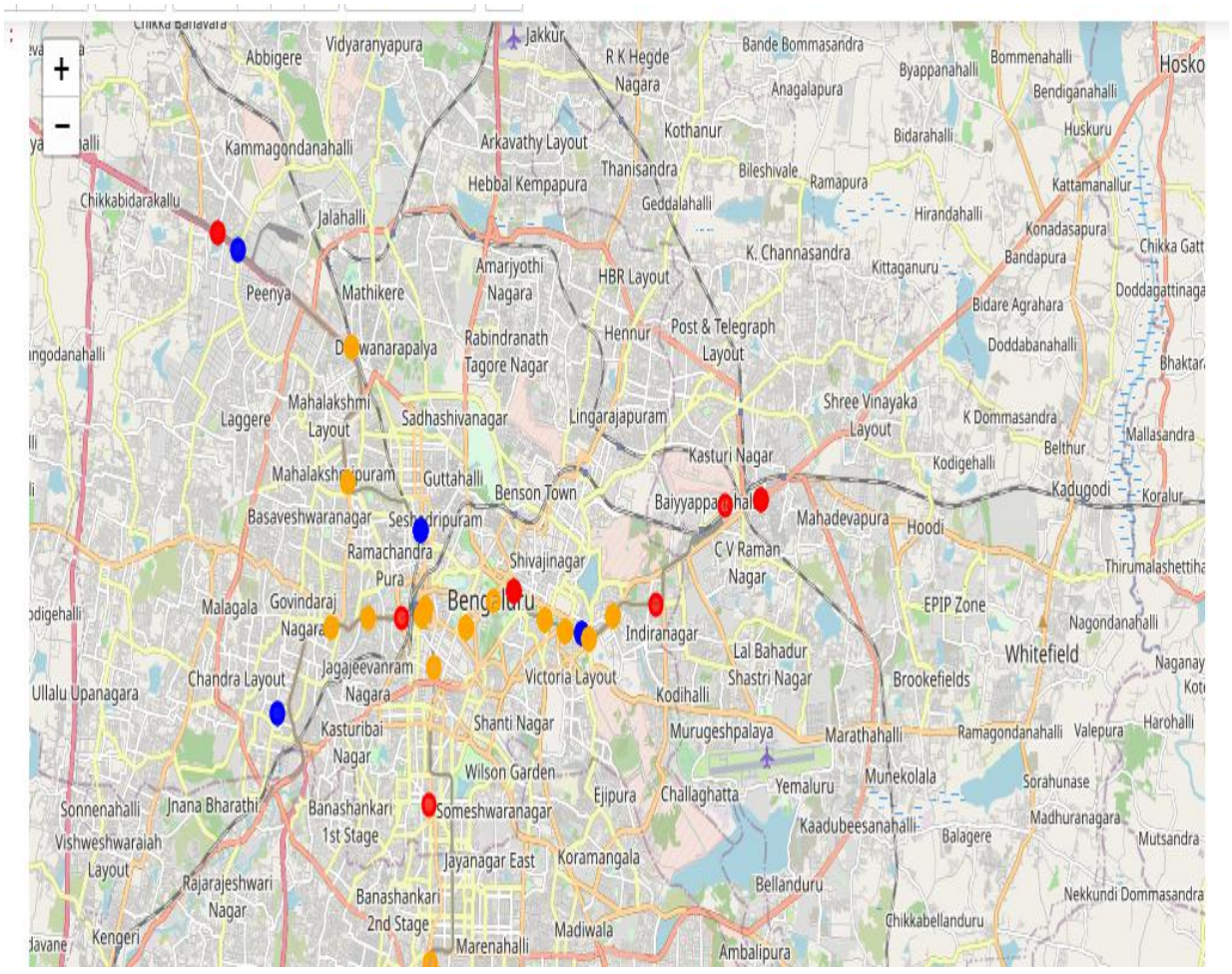
After executing the K-Means clustering algorithm three clusters of Metro stations were created, identified by their respective colors on the following map:

The colors indicate as follows:

Red: **Low potential**
There already exists enough number of cinema halls near metro stations

Orange: **Average potential**
There exists fair number of cinema halls but not all are very near to the metro station

Blue: **High Potential**
There exists very low no of cinema halls near metro stations



5. Discussion

Clusters of areas (in our case Metro stations) were identified as groups of similar in their potential locations for opening a gym.
Possible areas that were not in the Foursquare database should also be examined so that it can be determined if it is just lack of data about these stations or indeed there are no cinema hall in the vicinity of the stations.
A lot more factors can be considered when choosing an appropriate location. Some examples of extra factors can be:

- Population density in the area
- Number of businesses operating in the area
- Average age and household income in the area
- Property prices in the area

6. Conclusion

The above results can be **a good starting point** for a prospective businessman that is interested in opening a cinema hall. Similar methodology can be used for other types of businesses probably with customized criteria.
With the availability of a number of different tools and Machine Learning algorithms, it is possible to find solutions (or possible solutions) to an ever increasing number of problems and queries.