

# Capstone Project- The Battle of the Neighborhoods - Project Report

Applied Data Science Capstone by IBM/Coursera

By :Swapneel Prasanth Mishra

## Table of contents

- \* Introduction: Business Problem
- \* Data
- \* Methodology
- \* Analysis
- \* Results and Discussion
- \* Conclusion

## **Introduction: Business Problem**

In this project we will try to find an optimal location for a Yoga Centre. This report is essentially targeted for stake holders interested in starting a new **Yoga Centre or a multi-speciality fitness centre in Bangalore, India.**

Bangalore also known as the Silicon Valley of India is home to millions of Indians from all over India, working in various IT companies. A booming cosmopolitan population has meant huge growth potential in tertiary service industry. Naturally in recent times franchise model of businesses have taken hold in many parts of Bangalore. Same holds true for Health and Fitness sector as well.

With more and more people becoming fitness aware, has meant the cutthroat competition. Thus, Businesses must investigate niche offerings like Yoga, Zumba etc. A dedicated Yoga Centre is quite a lucrative offering since it addresses mind and body fitness. Also, the target audience larger, since Yoga can be taken up by people of all age groups compared to a regular Gym or Zumba centre which may be only target population between 20-40.

While a Gym may not be in direct competition of a Yoga centre will try to find locations, which does not have any fitness centre in the vicinity. Also, we will seek out locations which are in close vicinity of **residential, offices, colleges and commercial centres.**

We will utilize tools and techniques in data science to generate a few most favourable areas based on the above criteria

## DATA

Based on definition of our problem, factors that will influence our decision are

- Number of existing Gyms in the neighbourhood
- Number of and distance to Yoga Centres in the neighbourhood.
- Number of Homes, Offices, Colleges, Commercial activity in the neighbourhood.

Bangalore has Several suburbs and satellite locations which may not be ideal to start new business. So first we will determine the effective subset of Bangalore for our analysis.

Then we will define tightly packed grids of radius around 500 metres for our analysis

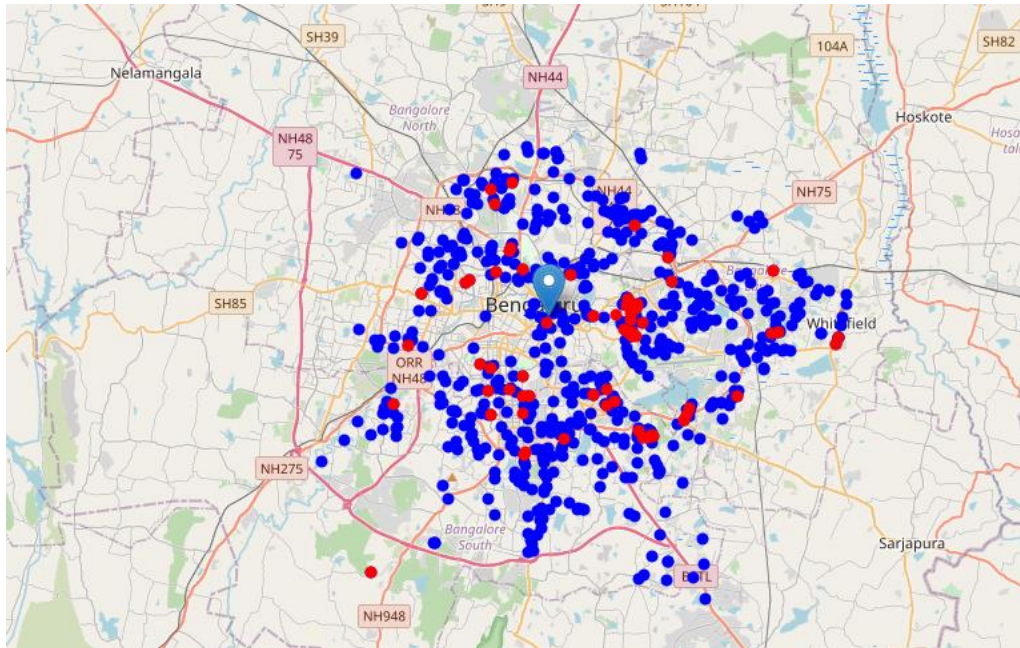
We will use the following data sources to generate the required information:

- The suburbs in Bangalore: We will use information in Wikipedia to get the list of suburbs in Bangalore.
- Then we will determine an effective radius which covers maximum number of suburbs.
- Then we generate the grid centres/locations and determine their approximate addresses using the OpenCage API
- In general, Spatial coordinates and reverse geocoding will be performed using **OpenCage API**.
- Number of Gyms, Yoga centres and amenities viz homes, offices, colleges, commercial centres in every location will be obtained using **Foursquare API**.
- Also, we will need several Python Libraries like, Pandas, Panda-SQL, NumPy, Folium to be used for our Analysis.

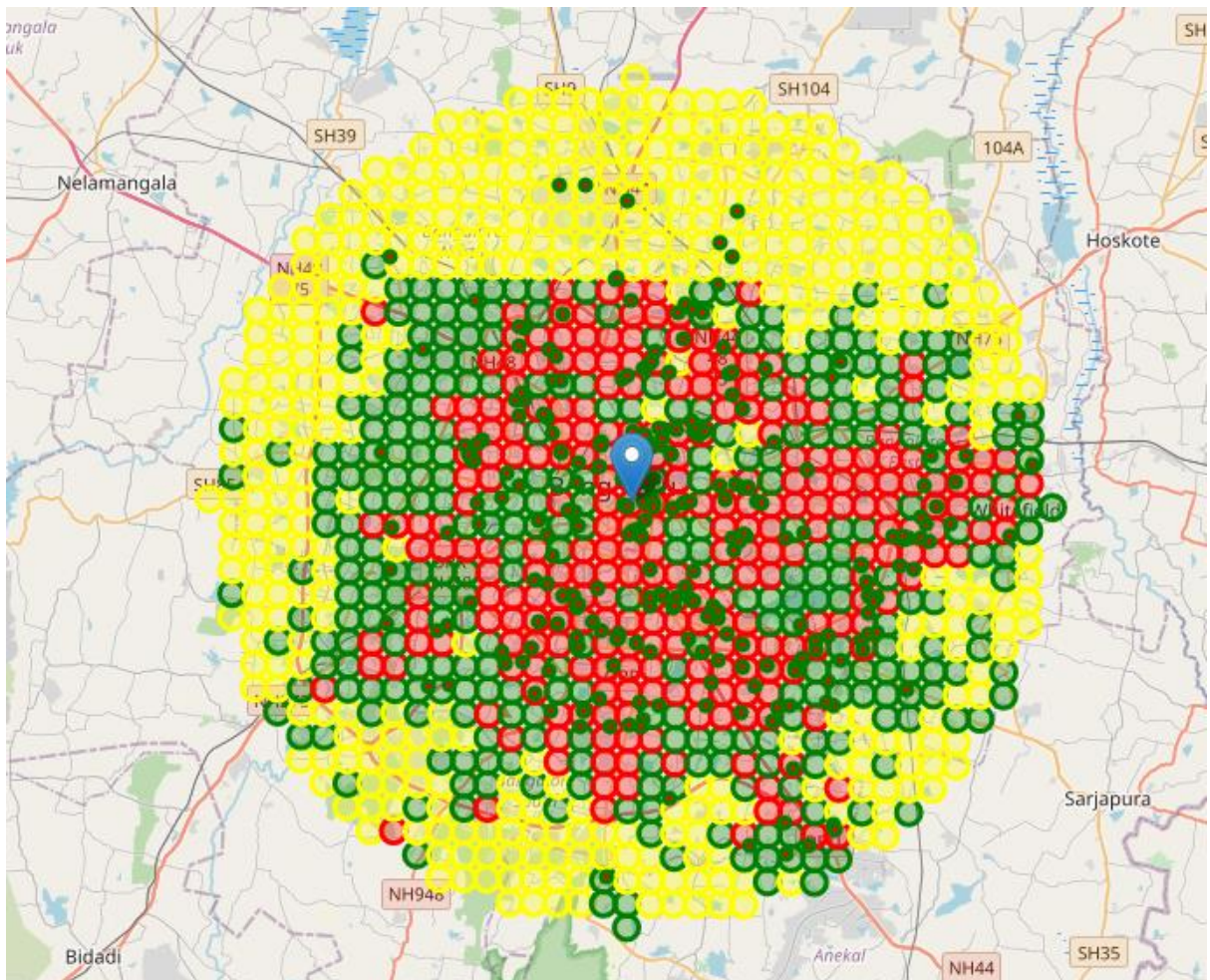
The raw data sourced from Foursquare and OpenCage APIs will be summarized in the below format and used for further refinement

	Address	Latitude	Longitude	X	Y	Distance	Gyms	YogaCent	Offices	Residenti	Commerci	Colleges
0	Bangalore	12.8098	77.59681	781902.5	1417520	18000	0	0	0	0	0	1
1	Byataraya	12.81929	77.55088	776902.5	1418520	17720.05	0	0	0	0	0	0
2	Byataraya	12.8192	77.56008	777902.5	1418520	17464.25	0	0	0	0	0	0
3	Gottigere	12.81911	77.56929	778902.5	1418520	17262.68	0	0	0	0	0	0
4	Bannerugl	12.81902	77.57849	779902.5	1418520	17117.24	0	0	0	0	0	0
5	Bannerugl	12.81893	77.5877	780902.5	1418520	17029.39	0	0	0	0	0	1
6	Bangalore	12.81884	77.59691	781902.5	1418520	17000	0	0	0	0	0	1
7	Bangalore	12.81875	77.60611	782902.5	1418520	17029.39	0	0	0	0	0	0
8	Bangalore	12.81866	77.61532	783902.5	1418520	17117.24	0	0	0	0	0	0
9	Yarandaha	12.81856	77.62452	784902.5	1418520	17262.68	0	0	0	0	0	0
10	Yarandaha	12.81847	77.63373	785902.5	1418520	17464.25	0	0	0	0	0	0

Below we can observe the Distribution of all **Gyms** and **Yoga Centres**



Below Plot roughly demarcates areas of potential interests. The regions marked in green are essentially the primary location of interest





## METHODOLOGY

In this project we will direct our efforts on detecting areas of Bangalore do not have a Gym or Yoga Centre. Also, we would try to look for locations which have offices, colleges, residential areas and commercial centres either in the location or in close vicinity.

This will give us an insight about the presence of a community in the given neighbourhood. Presence of a community and residents is necessary for any business.

From the data collected from the previous step we shall build a location profile, giving us the view of viability of each location. We do some exploratory analysis of the data to generate some additional parameters to aid us in our selection.

Below is the final form of the data post our exploratory analysis.

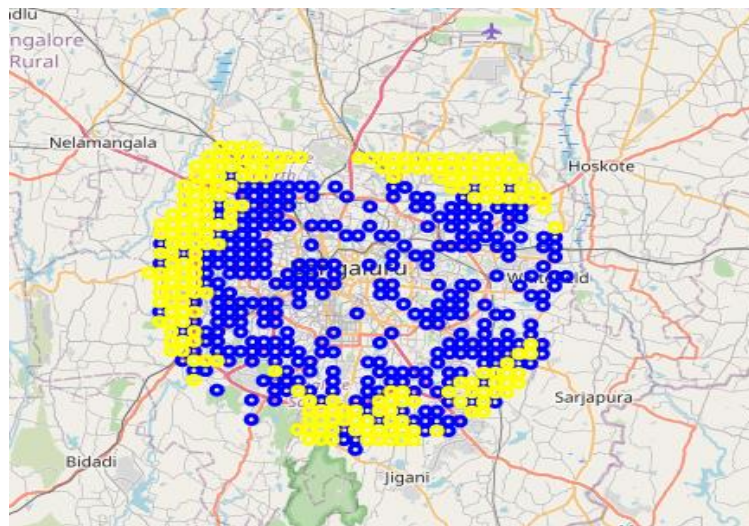
Address	Latitude	Longitud	X	Y	Distance	Gyms	YogaCent	Offices	Residenti	Commer	Colleges	Distance	Distance to Gym	Distance to Amenity	
Bangalore	12.8098	77.59681	781902.5	1417520	18000	0	0	0	0	0	0	1	10415.09	5062.130636	0
Byataraya	12.81929	77.55088	776902.5	1418520	17720.05	0	0	0	0	0	0	0	5496.633	4560.348385	4000
Byataraya	12.8192	77.56008	777902.5	1418520	17464.25	0	0	0	0	0	0	0	6381.579	4922.195531	3000
Gottigere	12.81911	77.56929	778902.5	1418520	17262.68	0	0	0	0	0	0	0	7296.309	4467.824516	2000
Bannerugl	12.81902	77.57849	779902.5	1418520	17117.24	0	0	0	0	0	0	0	8230.899	4107.108364	1000
Bannerugl	12.81893	77.5877	780902.5	1418520	17029.39	0	0	0	0	0	0	1	9179.285	3971.80341	0
Bangalore	12.81884	77.59691	781902.5	1418520	17000	0	0	0	0	0	0	1	9423.524	4084.373335	0
Bangalore	12.81875	77.60611	782902.5	1418520	17029.39	0	0	0	0	0	0	0	9610.985	4391.658311	1000
Bangalore	12.81866	77.61532	783902.5	1418520	17117.24	0	0	0	0	0	0	0	9896.427	4103.964013	2000

Taking the raw data at hand we will use the following scheme to determine a list of potential locations

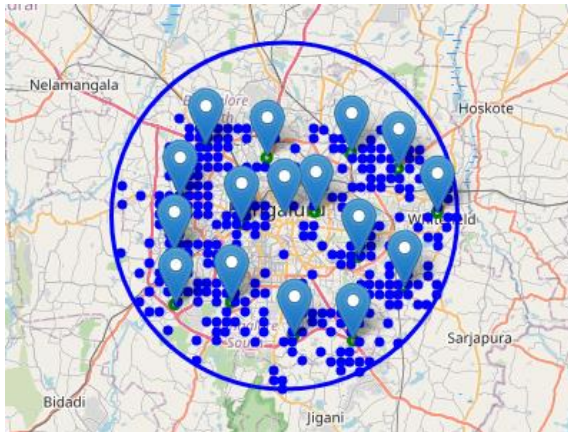
First, we will identify areas lacking a Gym or Yoga Centres but having the at-least one of the aforesaid amenities. These would be our Primary areas of focus

Next are the location which do not have a Gym or Yoga Centre and do not have the aforesaid amenities in their scope. But such neighbourhoods could have an amenity in the adjacent neighbourhood. Thus, all such locations which lack the aforesaid having some amenity within 4KM of their individual centres and from where the nearest Gym or Yoga Centre is farther than the average would be considered as potential secondary sites.

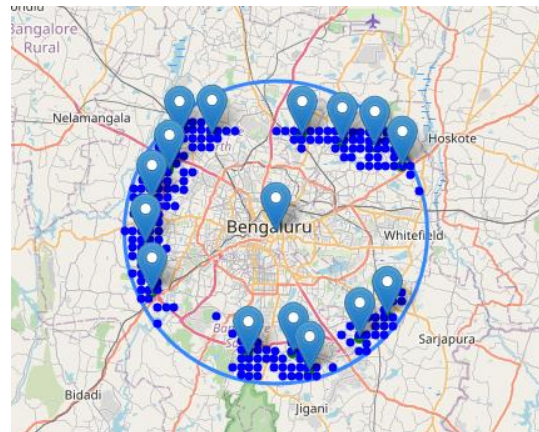
The plot below shows the potential primary sites in Blue and Secondary locations in yellow.



Once list of such sites is identified, we would use the K-Means clustering algorithm to identify 15 clusters of zones of primary and secondary interest.



Primary Cluster



Secondary Cluster.

The addresses of these locations would be the result of the project and starting point for further exploration.

## Results and Discussion

Our analysis shows that although for a city as populous as Bangalore is catered by few hundred Gyms and even fewer Yoga Centres. Thus, there is a great potential for new Yoga Centre or even a multi-speciality fitness centre.

Initially we divided the entire area of Bangalore into 500m radius circular cells and analysed each location. This helped us to arrive with data which showed which locations have fitness centres and which absolutely lack any

We tried assessing the feasibility of each location understanding the location profile. If a location has offices, homes, commercial centres or college, it indicates primarily that the cell is inhabited some economic activity can be inferred. Thus, such locations without a fitness centre are our primary area of interest. Now there could be areas which are remote but still within manageable distance from community centre. Such locations are the secondary set of locations that could be considered.

We then used the K means Clustering to arrive at 15 potential primary locations and 15 potential secondary locations

Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centres of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

### Addresses of Primary Locations.

Rayapuram Ward, Bengaluru - 560026, Karnataka	=> 4.6km from Bangalore Center
K R Puram, Bengaluru - 560 036, Karnataka	=> 12.8km from Bangalore Center
Begur, Bengaluru - 560076, Karnataka	=> 12.4km from Bangalore Center
4th Cross Road, Ganga Nagar Ward, Bengaluru - 560094, Karnataka	=> 6.3km from Bangalore Center
BWSSB Pipeline Track, Herohalli, Bengaluru - 560091, Karnataka	=> 11.1km from Bangalore Center
Kengeri Main Road, Jnana Bharathi, Bengaluru - 560074, Karnataka	=> 12.1km from Bangalore Center
Jogupalya, Bengaluru - 560007, Karnataka	=> 3.2km from Bangalore Center
India International School, Gunjur Varthur Main Road, Varthur, Bengaluru - 560087, Karnataka	=> 14.4km from Bangalore Center
Krishna Leela Park, Pipeline Road, Vasanthapura, Bengaluru - 560082, Karnataka	=> 10.5km from Bangalore Center
1st Cross Road, Horamavu, Bengaluru - 560 077, Karnataka	=> 9.5km from Bangalore Center
Bellandur Lake, Perl Paradise Community, Bellanduru, Bengaluru - 560103, Karnataka	=> 8.8km from Bangalore Center
Bangalore Elevated Tollways Limited, Begur, Bengaluru - 560 100, Karnataka	=> 14.7km from Bangalore Center
Umiya Woods, Ecumenical Christian Center Road, Hagadur, Bengaluru - 560066, Karnataka	=> 15.9km from Bangalore Center
HMT Ward, Bengaluru - 560057, Karnataka	=> 11.0km from Bangalore Center
Hemmigepura, Bengaluru - 560074, Karnataka	=> 14.7km from Bangalore Center

### Addresses of Secondary Locations

Thanisandra, Bengaluru - 56077, Karnataka	=> 11.8km from Bangalore Center
Gottigere, - 560083, Karnataka	=> 15.0km from Bangalore Center
Sunkadakatte, Machohalli - 560018, Karnataka	=> 15.3km from Bangalore Center
Bangalore Urban, Kodati - 560035, Karnataka	=> 16.5km from Bangalore Center
Shree Vinayaka Layout, Avalahalli - 560049, Karnataka	=> 16.9km from Bangalore Center
Anchepalya, Chikka Banavara - 560090, Karnataka	=> 16.2km from Bangalore Center
Kengeri, Bengaluru - 560074, Karnataka	=> 16.5km from Bangalore Center
Hudi, Bengaluru - 56077, Karnataka	=> 13.3km from Bangalore Center
Yarandahalli, Hulimangala - 560100, Karnataka	=> 16.8km from Bangalore Center
Begur-Koppa Road, Begur, Bengaluru - 560105, Karnataka	=> 14.4km from Bangalore Center
Dodda Bidarakallu, Bengaluru - 560073, Karnataka	=> 14.8km from Bangalore Center
Ullalu Upanagara, Kannalli - 560110, Karnataka	=> 15.6km from Bangalore Center
Begur, Bengaluru - 560 100, Karnataka	=> 16.0km from Bangalore Center
K. Channasandra, Bidarahalli - 560049, Karnataka	=> 15.5km from Bangalore Center
Shettihalli, Bengaluru - 560090, Karnataka	=> 13.8km from Bangalore Center

## **Conclusion**

Purpose of the project was to identify potential locations for starting a new Yoga Centre or Fitness Centre in Bangalore. Using the Foursquare data, we could gain insight about every location, the exact count of gyms or yoga centres, besides, we could get an overview of the location profile which helped us to narrow down the potential locations

We then clustered those locations to identify general zones which could use as starting point for further explorations

Final decision on optimal fitness centre/ Yoga Centre will be made by stakeholders based on specific characteristics of neighbourhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location real estate availability, prices, social and economic dynamics of every neighbourhood etc.