

# NEIGHBOURHOOD ANALYSIS CHENNAI, SOUTH INDIA

A report on finding few optimal locations in Chennai to open a new shopping  
mall using Data science skills

Author – Poongodi P

June 11, 2021

## Table of Contents

<b>1. Introduction &amp; Business Problem:</b>	2
1.1 Background:	2
1.2 Problem Statement:	2
1.3 Who will be interested? :	2
<b>2. Data:</b>	2
2.1 Data Source:	2
2.2 Data cleaning:	2
<b>3. Methodology:</b>	3
3.1 Exploratory Data Analysis:	3
<b>4. Results:</b>	6
<b>5. Discussion:</b>	9
<b>6. Conclusion:</b>	9
<b>7. Future Direction:</b>	9
<b>8. Code:</b>	9
<b>9. References:</b>	9

## 1. Introduction & Business Problem:

### 1.1 Background:

This report deals with analysis of neighbourhoods of **Chennai, South India employing Datascience skills**. Chennai also known as Madras is the capital city of Indian State of **TamilNadu**. Located on the Coromandel Coast of the Bay of Bengal, it is one of the largest cultural, economic and educational centres of south India. According to the 2011 Indian census, it is the sixth-most populous city and fourth-most populous urban agglomeration in India. The city together with the adjoining regions constitutes the Chennai Metropolitan Area, which is the **36th-largest urban area by population** in the world.

Chennai is divided into four broad regions: North, Central, South, and West. North Chennai is primarily an industrial area. South Chennai and West Chennai, previously mostly residential, are fast becoming commercial, home to a growing number of information technology firms, financial companies and call centres. The city is expanding quickly along the Old Mahabalipuram Road and the Grand Southern Trunk Road (GST Road) in the south and towards Ambattur, Koyambedu and Sriperumpudur in the west.

### 1.2 Problem Statement:

With this background, the **purpose of this report is to leverage data science skills to find out few prime neighbourhoods in Chennai that will be appealing for an entrepreneur to open a new shopping mall**. Chennai hosts a large number of shopping malls. We will also have a brief look into it.

### 1.3 Who will be interested? :

Deciding where to put your business is as important as the business you decide to go. Intended audience of this report would be **large scale entrepreneurs** who wish to know about promising locations in Chennai city to open their **new** business specifically a shopping mall. This report also gives the stakeholders basic understanding about the existing malls.

**Note:** This project can be extended for other businesses such as restaurants or electronic shops with little modifications.

## 2. Data:

Success of a mall is directly proportional to the customer traffic. Our logic behind this project is to find out the places in the city **where the spending population traffic is high with less number of malls existing nearby**.

We will need the following data:

- Locations of **existing shopping malls** in the city
- Population** distribution of the city
- Locations of **spending population** in the city

### 2.1 Data Source:

- Locations of existing shopping malls** are obtained using **FoursquareAPI**.
- Premium details such as rating and reviews for each shopping mall are gathered using **FoursquareAPI premium services**.
- Population distribution of Chennai :  
Latest census data for Chennai is available for the year 2011. Ward wise population (**A city is divided into wards for administrative purposes – similar to boroughs**) for Chennai is obtained from District Census Hand Book (DCHB) from [census website of India](#)
- Choropleth map will be a good choice to visualize the population distribution. To plot choropleth we need to get shape/geojson file for Chennai. Thanks to <https://github.com/mickeykedia/India-Maps> for the geojson, shape and allied files for Chennai city!
- Details of prime localities and popular IT parks are collected from Wikipedia. Localities include major IT parks and large residential localities.

### 2.2 Data cleaning:

- Foursquare API returns lot of details. Extract only the name, geographical coordinates and id of the shopping malls.
- From Foursquare premium call, extract only rating and reviews for the malls.
- Population Data:  
Challenge here was, during 2011 when census was taken, Chennai city had 155 wards. Soon afterwards it was extended to have 200 wards. Population data for the newly included wards were collected from [DCHB of Kancheepuram](#). Population data along with ward numbers are extracted to 'Chennai\_wardwise\_population\_2011.csv'.
- Geojson file of Chennai had ward names missing for certain ward numbers. Update zone name as ward name in such cases

### 3. Methodology:

We first get the locations of existing shopping malls using FoursquareAPI. Here we search for shopping malls in Chennai city as a whole.

[https://api.foursquare.com/v2/venues/search?&client\\_id={}&client\\_secret={}&v={}&ll={}&categoryId={}&query={}&limit={}](https://api.foursquare.com/v2/venues/search?&client_id={}&client_secret={}&v={}&ll={}&categoryId={}&query={}&limit={})

From the resulting JSON extract name, venue id, latitude and longitude for all the returned venues.

To fetch ratings and reviews from Foursquare API use its premium call for venues with venue id being endpoint:

[https://api.foursquare.com/v2/venues/{id}?&client\\_id={}&client\\_secret={}&v={}&v={}](https://api.foursquare.com/v2/venues/{id}?&client_id={}&client_secret={}&v={}&v={})

Here venue id from previous result is passed to get the ratings for individual venues. Since this is a premium call, we will be able to execute this API only once per venue per day. From the resulting JSON extract ratings and tips.

**KMeans** clustering machine learning model is used to group the venues. Feature set includes ratings and location of the venues. Plotting the data using folium library we will get a picture of locations of existing malls and groupings of similar kind of malls. At this point stakeholders will get an idea about the existing business situation in the city. **High and medium rated malls are filtered to predict precise competition level.**

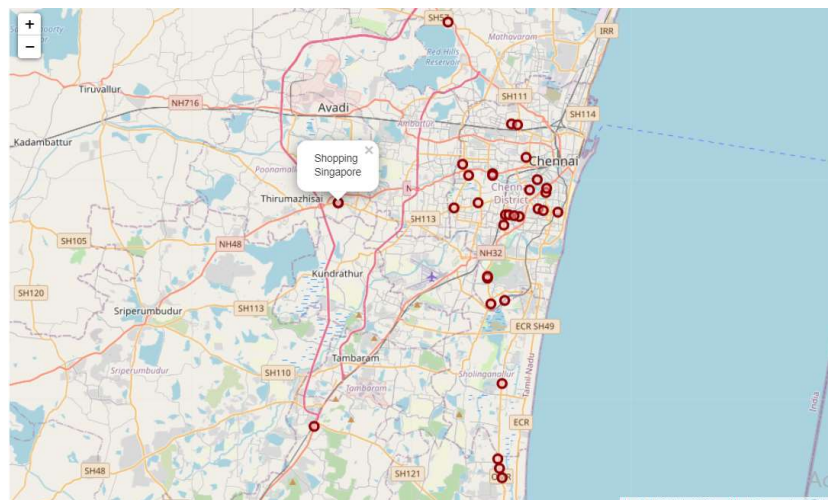
Success of a shopping mall highly depends on the customer traffic. The following criterions are taken into account to solve our problem. Fix areas:

- ✓ Based on the **population distribution**:  
Locations with high population are better places to get more customers
- ✓ Based on **purchasing power of the localities**:  
Localities having big residential properties, large IT parks, prominent companies attracts affluent customers.
- ✓ Based on the **availability of high rated malls nearby**:  
Competition level is low with less number of high rated malls nearby

#### 3.1 Exploratory Data Analysis:

##### 3.1.1 Locations of Shopping Malls in Chennai:

Using folium map plot the locations of existing malls in the city.

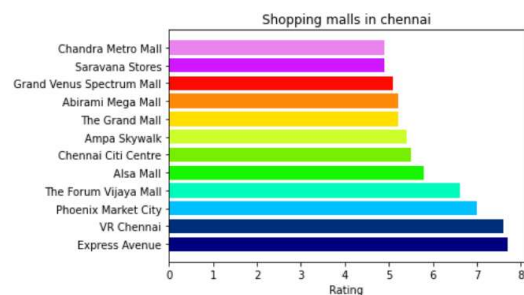


##### 3.1.2 Ratings and reviews:

This table lists the address, ratings and reviews of top 10 shopping malls in the city. This result is based on the views of foursquare website users.

	Mall Name	Address	Rating	Reviews
0	Express Avenue	2, Whites Road	7.7	avoid on Sundays!! way too many people- parking gets full, overdose of body odor & rowdy crowd!
1	VR Chennai	Jawaharlal Nehru Road (100 Feet Road), Ward 99	7.6	So many choices n unimaginable collections. Good Mall.
2	Phoenix Market City	Velachery Bypass Road, Gangai Nagar	7.0	Free WiFi by Tikona. You'll receive a PIN number through SMS, once you register in the Tikona's login page. Then you can access the Internet
3	The Forum Vijaya Mall	Ponammal St, Ward 130	6.6	IMAX will be ready by march !! Told by mall staff.
4	Alsa Mall	Montieth Road, Egmore	5.8	Bad place to come by car coz they don't have space for parking and even the so-called valet parking makes you wait for ages while they retrieve your car from the cluttered lot.
5	Chennai Citi Centre	10/11, Dr Radhakrishnan Salai	5.5	McD's roof top recommended for evening hangout!
6	Ampa Skywalk	Nelson Manickam Road, Aminjikarai	5.4	Be prepared for a good parking charge! And also parking slot is too confusing to spot our bikel
7	The Grand Mall	Dr Seetharaman Nagar Main Road, VGP Selva Nagar	5.2	Grand is in name only!! Nothing Grand
8	Abirami Mega Mall	Purusawalkam High Rd, Kellys	5.2	Go to Abirami Theatre and experience the newly renovated theatre.
9	Grand Venus Spectrum Mall	Paper Mills Rd, Periyar Nagar	5.1	Horrible experience to travel to this place. The Mall is like a shit congested building. The theatre is out of place and not suited for this location. Cannot travel by car to this mall. Will never go.

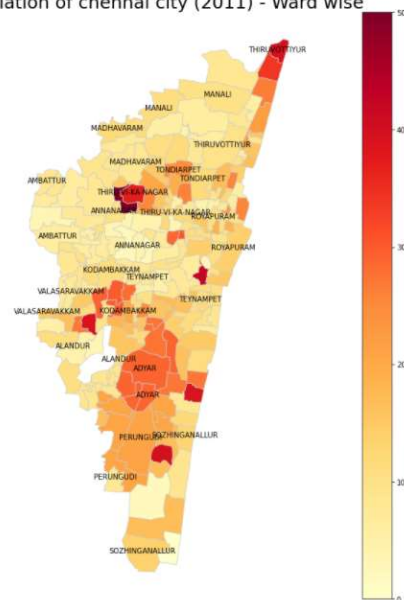
Bar chart showing high and medium rated malls



### 3.1.3 Population distribution:

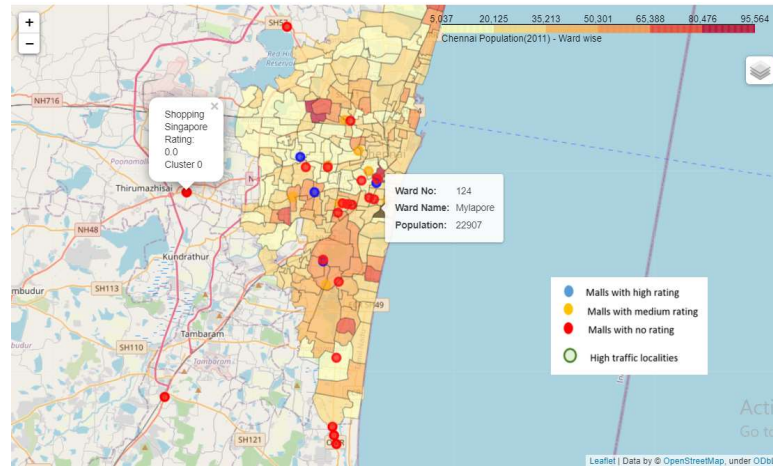
This map shows the population distribution across the city based on the census data of 2011.

Population of chennai city (2011) - Ward wise



### 3.1.4 Group the malls:

We will group the malls using kmeans clustering machine learning model based on location and rating. This map shows the details of grouped malls plotted over choropleth map showing population distribution. There are three groups based on ratings and locations - high rated malls, medium rated malls and malls that have no rating. Foursquare website lack ratings for many malls but since the project requirements stated to use Foursquare I have gone for it. I have extracted the **high and medium rated malls to a dataset to use it in final decision. This will give a good picture of precise competition level around recommended locations** Popup shows ward number, name and population count for every ward.

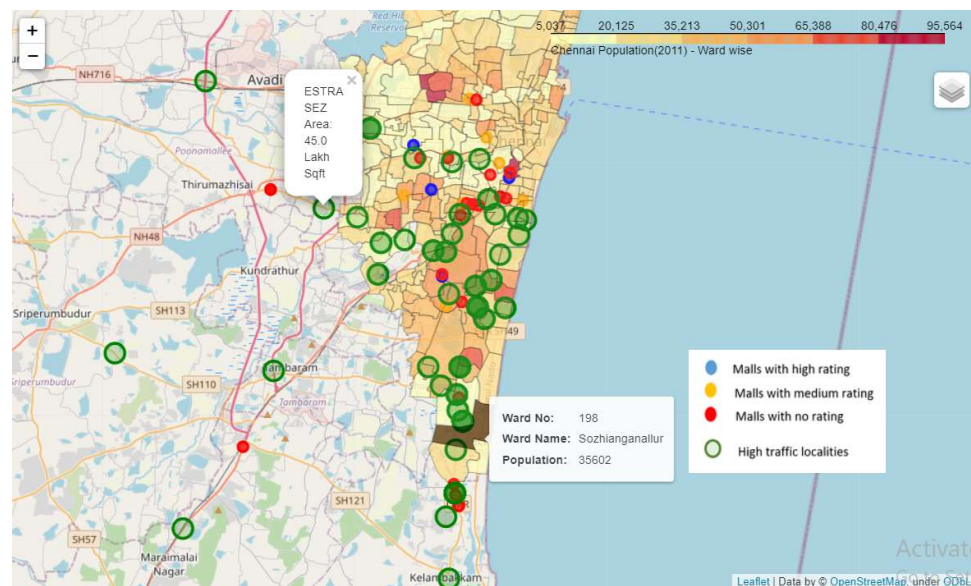


### 3.1.5 Locate spending population:

Success of a mall highly depends on the spending population around the locality. Chennai has a diverse array of economic sectors and is known for its **automobile industry and rich IT sector**. There are various **IT parks** that hosts numerous multinational IT companies from across the globe. Rich residential flats and villas started to emerge around these IT parks in the past decade to leverage the need. We can comfortably assume these areas as good source of spending population. Especially a shopping mall will flourish in these spots as there is huge need for shopping and entertainment. I collected the locations of IT parks and commercial places of other sectors from Wikipedia. These commercial localities are the places that serves large number employees in the city. The below table lists top 5 localities with respect to area. Hereafter we will refer them as **hotspots** in our report.

	Locality	Area	Address	Latitude	Longitude	Ward_no
0	DLF -IT park	70.0	Manapakkam,Chennai	13.014753	80.174493	157
1	Shriram The Gateway SEZ	50.0	Perungalathur,Chennai	12.927605	80.099965	
2	ESTRA SEZ	45.0	Ayyapanthangal, Chennai	13.037793	80.134942	150
3	Embassy Splendid Tech Zone	40.0	Pallavaram, Chennai	12.993374	80.172587	
4	ETL Infrastructure - Chennai One	37.8	Thoraipakkam	12.930069	80.230022	195

Plot the hotspots on the existing map. Green circles are the hotspots. We could see that these spots spread across central and southern part of Chennai.



From the above map we could see there are few areas which hosts dense spending population with no or few malls around. We will direct our research to narrow down to these spots.

## 4. Results:

### 4.1.1 Filter based on population:

It's time now to filter out results from the data collected so far. First let us list out **top 10 densely populated wards** in the city.

WARD_NO	WARD_NAME
0	65 THIRU-VI-KA-NAGAR
1	63 Komaleeswaran pettai
2	1 Sathyavani muthu nagar
3	194 Okkiyam thuraiakkam
4	182 ADYAR
5	155 Ramapuram
6	64 Kolathur
7	2 Kathivakkam
8	75 THIRU-VI-KA-NAGAR
9	129 Saligramam

**Wards that have maximum number of spending population;**

[197, 180, 95, 184, 86, 170, 81, 185, 174]

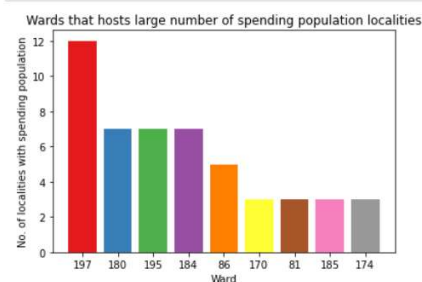
From the above two lists it is clear that there are **NO** matching ward numbers between the categories. Means there are no wards which are both densely populated and have dense spending population. This does not seem to make much sense. Reason behind this mismatch is the **census data is nearly 10 years old**. After the census, Chennai is converted to Greater Chennai enclosing 200 wards. The city has grown manifold since then. Spending population started growing in south Chennai. Also many of these population are floating population having their native somewhere else. So they may not get reflected in the city population census.

With this background, for now, till we get fresh census data, **we will exclude population density from our deciding factor**. However it is to be noted that when accurate census data is available, we can consider this as important influence factor to decide a business location.

### 4.1.2 Filter based on spending population:

#### 4.1.2.1 Wards having large number of hotspots:

Few wards have more number of spending population compared to the rest. This is the simple bar chart showing the ward numbers and number of spending localities ward wise:



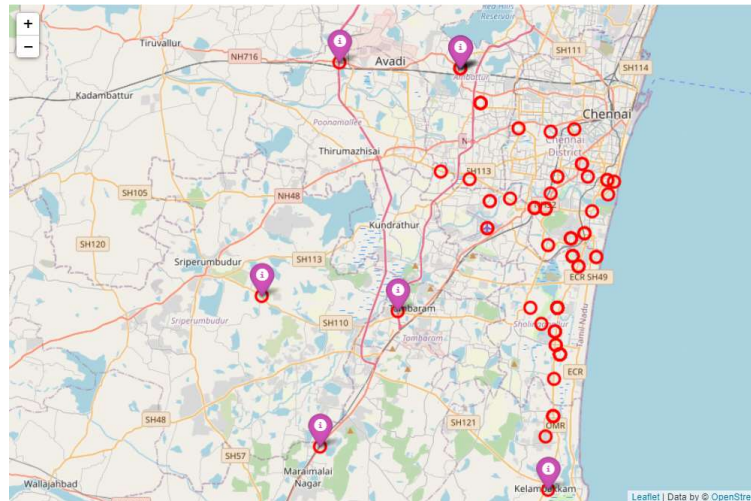
#### 4.1.2.2 Average distance of hotspots from malls:

Let us see if there are hotspots that has **no** malls within 5 km radius. Calculate the distance between each hotspot to all the malls and filter out hotspots that have no malls within 5 km radius. These are the places that are get caught:

Hot spots that has no malls within 5 kms:			
	Locality	Latitude	Longitude
0	Shriram The Gateway SEZ	12.927605	80.099965
1	SRM Tech Park	12.820757	80.036874
2	Ambattur Clothing Co	13.119375	80.150765
3	Tidel Park Avadi	13.124234	80.052364
4	TNPL	13.119375	80.150765
5	Alliance Orchid Tech Park	13.119375	80.150765
6	Arun Fabricators	12.939937	79.989277
7	Krishnan Sivagami Trust	12.786249	80.221994

Now we will get a clear idea of the locations by plotting on a map. Red circles are the locations with spending population. Purple icon are the localities that have no malls within 5 km radius.



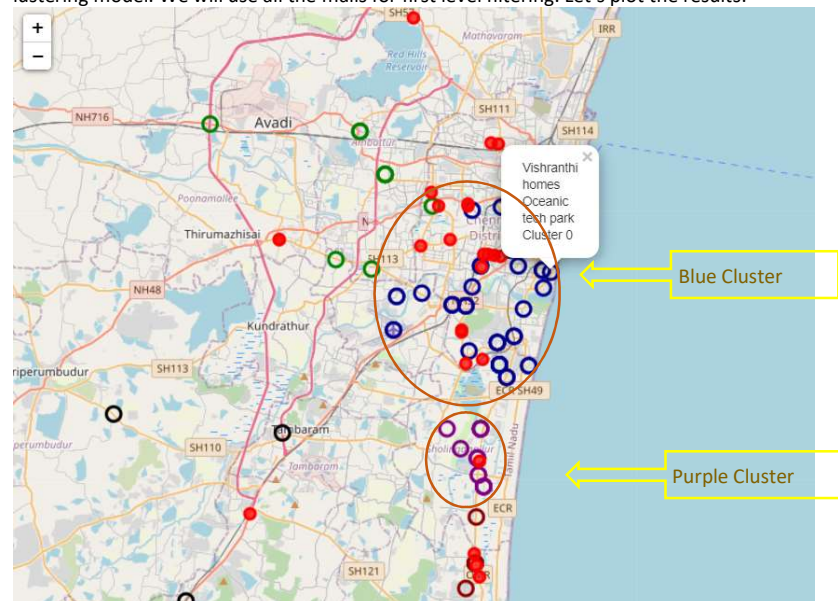


Oops! These locations look like outliers! Most of them are situated outside city limit. There are no other prominent spending population around these locations. We cannot recommend these places.

Also our calculations show that on average there are 4 to 5 malls exist within radius of 5 km for every hotspot.

#### 4.1.2.3 Narrow down the spot by clustering:

We will cluster the spending location into five groups based on the latitude and longitude using Kmeans clustering model. We will use all the malls for first level filtering. Let's plot the results:



Red circles are the locations of shopping malls. Coloured rings are hot spots. From the map above, we derive the following insights:

We can "exclude" the following localities from our 'target' for the following reasons:

1. North part of Chennai: It is primarily an industrial area. Does not have spending population. (includes Tiruvotriyur, Washermanpet, VOC Nagar etc)
2. Central part of Chennai: Quite large number of malls already exist in this part of Chennai (around Thousand Lights, Royapettah, Teynampet, West Mambalam)
3. West Chennai: Few malls exist already in this part. Does not seem to have large number of spending population as there are only few big commercial sites and IT parks
4. South most Chennai: (Navalur, Siruseri etc.,) Hosts three malls already. Although these are not high rated malls, we cannot recommend this location since there is not enough spending population around

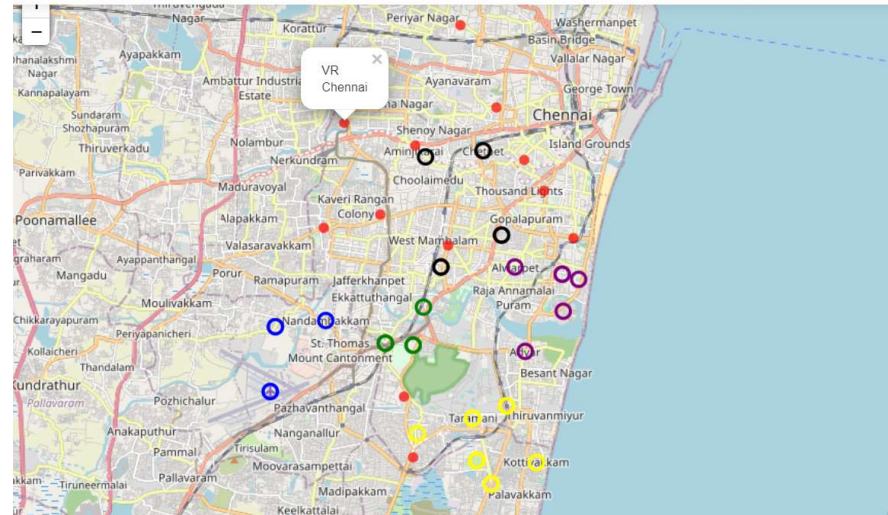
Following areas can be considered as promising localities:

1. **South East part of Chennai (Blue Cluster):** Areas such as MRC Nagar, Raja Annamalai Puram, Kotturpuram, Adyar, Besant Nagar and Thiruvannamiyur prove to be promising localities for opening a new mall as these are considered posh areas with large number of spending residents. It also encloses large number of big commercial firms.



2. **Purple Cluster:** Areas on OMR like Kottivakkam, Palavakkam, Kandhanchavadi, Thuraipakkam and Sholinganallur also proves to be promising localities as these encloses the fast growing IT parks with large number of spending residents around. Important point to be noted here is there are no big prominent malls exists in these localities. It hosts one mall with 6 prime IT parks around it. This mall is “Saravana Stores” which cannot be considered as a full-fledged shopping mall but a big retail showroom owned by a single company. We can safely discard the competition from this mall. So we will fix **centre of this cluster as one of our target**.

Let us dig further to see if we can precisely name few localities from Blue cluster. Group the localities in this cluster into five regions using kmeans. For this final level filtering we will **consider only high and medium rated malls**.

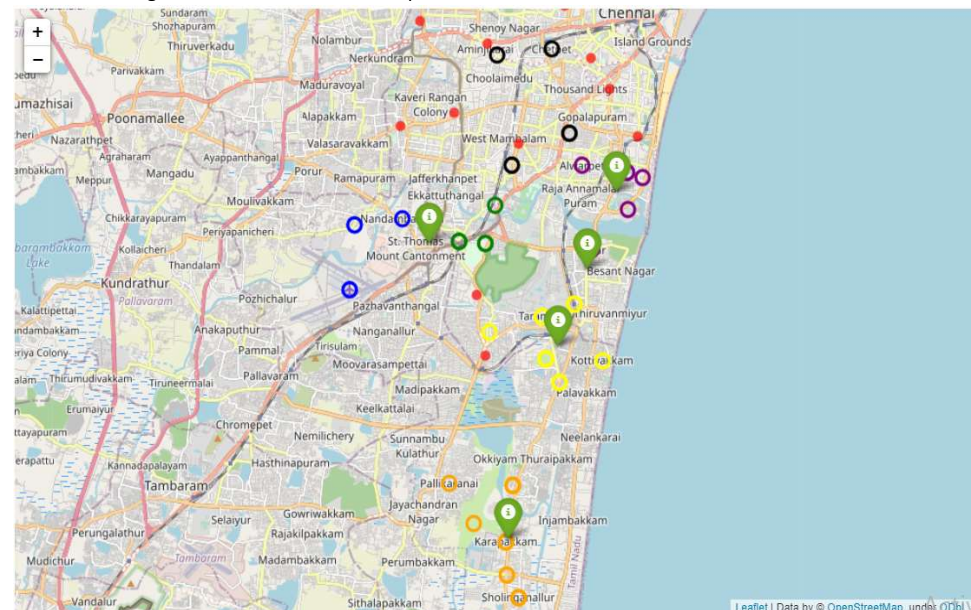


Red circles are high or medium rated shopping malls. We could see there is no hope around black cluster as there are bunch of malls with few hotspots. We can consider the other clusters.

- Get the **centre of Purple cluster as one of our target**.
- In the yellow cluster, exclude the outlier on western side and find the centre. As there are malls on western part of this cluster, we will **fix eastern part of the centre as one of our target**.
- One important promising locality that we could trace is the area in between purple and yellow cluster. Fix **the mean of above two target as one of our target**.
- We will consider blue and green clusters combined as there are no malls close by. It has sufficient hotspots. **Fix centre of these two clusters as a target**.

#### 4.1.2.3 Recommended optimal spots:

Plot the five target discussed above in the map.



**Green location icons** are the proposed spots in Chennai city that can be explored further to open a new shopping mall.

The following table lists the address of the recommended spots. This also shows number of high rated malls within 3 km radius of the proposed spot.

	Latitude	Longitude	Address	Postalcode	High rated malls within 3 km
0	13.000513	80.256191	Vasanthan Bhavan, Dr Muthulakshmi Salai, Bakthavatchalam Nagar, Ward 175	600020	0
1	12.912769	80.228272	Prestige Cyber Towers, Rajiv Gandhi Salai, Navalur, Zone 15 Sholinganallur	603103	0
2	13.025520	80.266486	Mandaveli, Ward 126, Zone 9 Teynampet, Chennai	600028	1
3	12.975507	80.245895	Nehru Nagar Industrial Estate, Ward 183, Zone 14 Perungudi, Chennai	600041	1
4	13.008968	80.200370	M Auto, Paul Wells Road, St. Thomas Mount Cantonment, Alandur	600016	1

## 5. Discussion:

Chennai's rate of growth has seen a deep rise for past couple of decades. While north Chennai remains almost unchanged, the southern part has expanded drastically, IT firms being a major reason. Although a bunch of shopping malls already exists in the city, considering its rate of growth, it will not be a bad idea deciding to open a new mall. This report analyses the Chennai's neighbourhood to find an optimal location to open a new shopping mall.

Starting a new retail business such as malls works out well in densely populated areas. Directing our analysis based on this criteria faced a critical problem. Latest census data available is for the year 2011. Current year (2021) is the next census year. Owing to COVID and other reasons, latest data is not available yet. This made our analysis difficult in this direction.

Other important factor that decides success of a retail business is spending population. I tried to locate the places where there is huge spending population. To locate this, I searched for firms where huge number of peoples are employed. Obviously most of these places turned out be big IT parks. Few of them were big companies from other sectors like finance and automobiles. I have collected their addresses in a file. Totally there were 90 such hotspots. Apart from the commercial spots, these localities turned out to be a rich residential spots owing to the growth of real estate aiming the employees of these firms. Obviously these are the best localities we want to aim for a new retail business. Most of these falls on southern part of Chennai. I could find there is by average 4 or 5 malls within the radius of 5 kms surrounding these hotspots. On carefully narrowing down, I filtered few spots with dense spending population that has less number of high rated malls within radius of 3 kilometres.

## 6. Conclusion:

Purpose of this project was to use Data Science skills to identify areas in Chennai that has dense spending population with less number of malls nearby to aid the stakeholders to find an optimal location to open a new shopping mall. Malls density distribution across the city was fetched using Foursquare. Location with spending population is collected from Wikipedia. Based on the locations of spending population, the area is clustered using kmeans model. The spots are narrowed down using dense spending population with few high rated malls nearby. Results were displayed using maps and tables. These spots are recommended to the stakeholders for further exploration.

## 7. Future Direction:

**'Rating' of the malls could be a valuable feature to get precise competition level.** Foursquare did not provide ratings for all malls. Most of the malls has no reviews or ratings. Since this project demands the usage of Foursquare API, I did not use other APIs such as Google places. Google places gives more precise data as far as Indian neighbourhoods are considered. This project can be further fine-tuned by using accurate ratings of existing malls.

Population density can also be a good deciding factor once latest census data is available.

So this project can further be fine-tuned taking into account the couple of fore said features to achieve better results.

Apart from shopping malls, this project can also be used to fix localities for other retail businesses such as a restaurant or an electronic shop by using the respective data.

## 8. Code:

[Github Repository](#)

## 9. References:

1. [Foursquare API documentation](#)
2. [Geopy Documentation](#)
3. [District Census Handbook – Chennai](#)
4. [District Census Handbook – Kancheepuram](#)
5. [Chennai's IT parks – Wikipedia](#)
6. [Geojson file for Chennai](#)

Thank you!