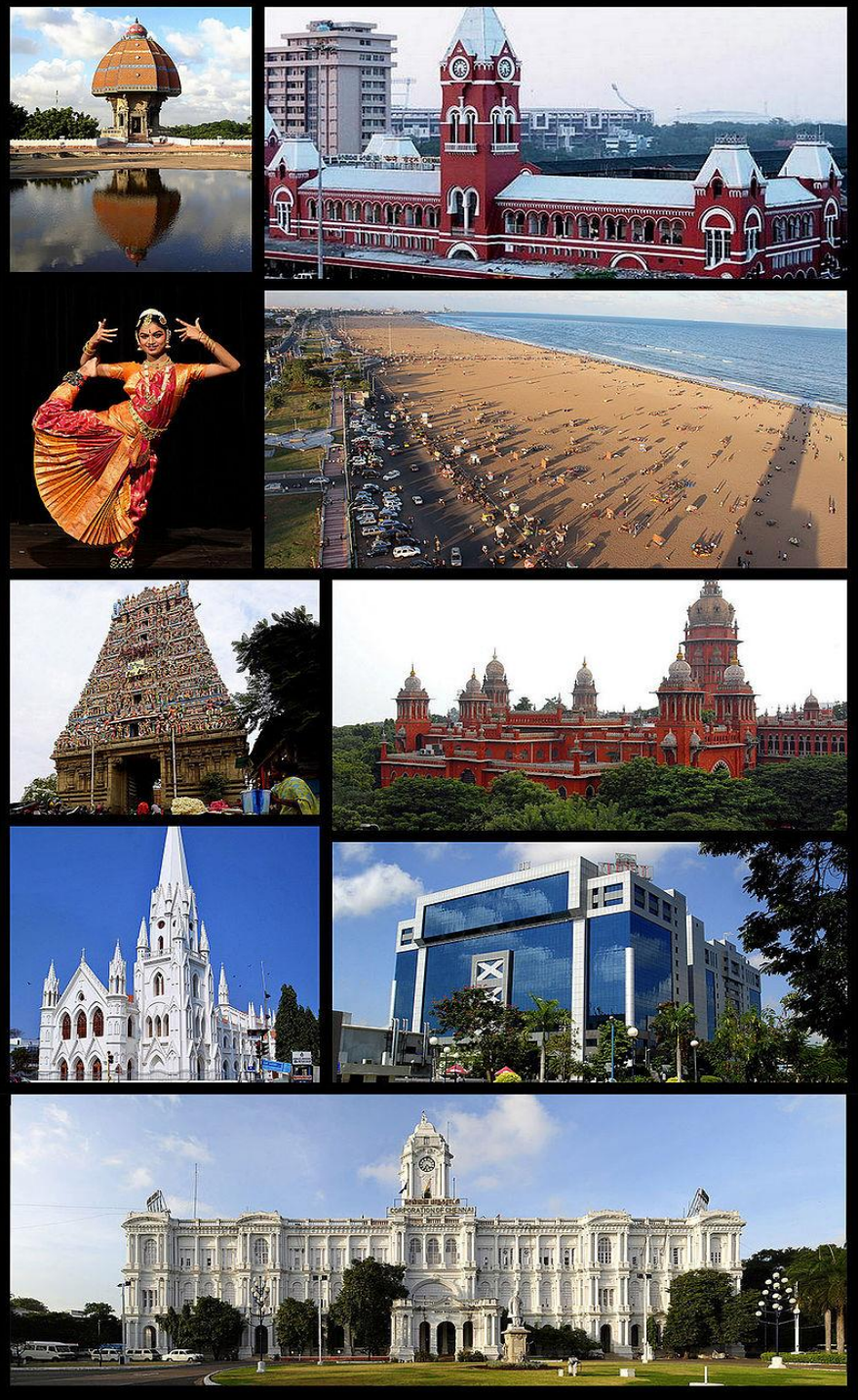
INTRODUCTION

For many shoppers visiting shopping malls is a great way to spend their leisure time during weekends. The different kinds are stupid gigantic, windowless urban, after-hours pedestrian, impressive urban, smooth, snow-covered, successful regional, massive mega, unfortunate modern, expansive pedestrian, massive regional, flashy corporate, big, faceless, dizzying underground and has to be high end pedestrian. It consists of restaurants ,spa ,kids-zones, theatres and arcades. It houses people of all age groups . It is a growing demand to build shopping malls in a developing country like India. Chennai , the capital of southern state Tamil Nadu is a fast growing economy and is the perfect time for builders to build shopping malls here. But the location of a mall is important in determining it's success.

ABOUT CHENNAI

- **Chennai** , also known as **Madras** , is the capital of the Indian state of Tamil Nadu. Located on the Coromandel Coast off the Bay of Bengal, it is a primary cultural, economic and educational centre of south India.[*dubious* – *discuss*] 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.[14] Chennai is among the most-visited Indian cities by foreign tourists. It was ranked the 43rd-most visited city in the world for the year 2015.[15] The Quality of Living Survey rated Chennai as the safest city in India.[Chennai attracts 45 percent of health tourists visiting India, and 30 to 40 percent of domestic health tourists.[17] As such, it is termed "India's health capital".



SHOPPING MALLS IN CHENNAI

Notable shopping malls [edit]

Name	Location	Year	Size	Source
Spencer Plaza	Mount Road	1863	250,000 sq ft (23,000 m^2)	
Alsa Mall	Montieth Road, Egmore	1998		
Abhirami Mega Mall	Purasawalkam	2003		
Mayajaal	Kanathur, East Coast Road	2006	30,000 sq ft (2,800 m^2)	
Chennai Citi Centre	Dr.Radhakrishnan Salai, Mylapore	2006	117,600 sq ft (10,930 m^2)	
Ampa Skywalk	Poonamallee High Road, Aminjikarai	2009	650,000 sq ft (60,000 m^2)	
Express Avenue	Whites Road, Royapettah	2010	1,500,000 sq ft (140,000 m^2)	
Chandra Mall	Arcot road, Virugambakkam	2011	143,130 sq ft (13,297 m^2)	[4]
Coromandel Plaza	Navalur, OMR	2011	300,000 sq ft (28,000 m^2)	[5][6][7]
Spectrum Mall	Paper Mills Road, Perambur	2011	160,000 sq ft (15,000 m^2)	
Ramee Mall	Anna Salai, Teynampet	2012	225,000 sq ft (20,900 m^2)	[8]
Bergamo	Khader Nawaz Khan Road, Nungambakkam	2012	40,000 sq ft (3,700 m^2)	
Gold Souk Grandé Mall Chennai	GST Road, Vandalur	2015	800,000 sq ft (74,000 m^2)	[9][10]
Phoenix Market City (Chennai)	Velachery	2013	2,400,000 sq ft (220,000 m^2)	[11]
Forum Vijaya Mall	Arcot road, Vadapalani	2013	700,000 sq ft (65,000 m^2)	[12]
VR Chennai	Inner Ring Road, Anna Nagar West	2018	2,000,000 sq ft (190,000 m^2)	

BUSINESS PROBLEM

- The objective of this capstone project is to analyze and select the location for a new shopping mall in the city of Chennai using data science methodology. It provides the answer to the business question , which area is the best location to build a new shopping mall. If a property developer is looking to open a new shopping mall, where would you recommend to open it?

Target audience

- This project is targeted towards real estate builders and promoters in the city of Chennai. The project is timely as the city is in the process of westernization as the number of malls have also shown a significant increase. A new shopping mall can be opened by the real estate promoters based on the project findings (ie) the optimum location to open a new mall.

DATA REQUIRED

- To solve the problem , we need the following data,
- List of neighborhoods in the city of Chennai.
- Location of neighborhoods found.
- Venue data related to shopping malls. This can be later used for clustering.
-
- The Wikipedia page https://en.wikipedia.org/wiki/Category:Suburbs_of_Chennai contains a list of suburbs of Chennai. HTML Parsing is done using module such as Beautiful soup . Once scraping of web page is done to extract the required data, python module Geocoder is used to extract the latitude and longitude of the location
- The foursquares api is used to get the venue data of shopping malls in each location. Foursquares api will provide name, location and categories of venue data near the given location after we specify the radius and limits. We will use the Shopping Mall category. We will use KMeans clustering to cluster the locations with the shopping mall category and see the best cluster to build a new mall.

METHODOLOGY

- The Wikipedia page is scraped and downloaded into the notebook with the use of Beautiful Soap library. But then it is just a list of names with no meaningful insight. Then we use the python geocoder package to extract the latitude and longitude of the neighborhood extracted. We then make use of the folium plot to visualize the neighborhoods for sanity check. Next we use foursquares api to get the venue data. The calls can be made by using a client ID, Client key which can be obtained by simply registering as a developer with Foursquares.

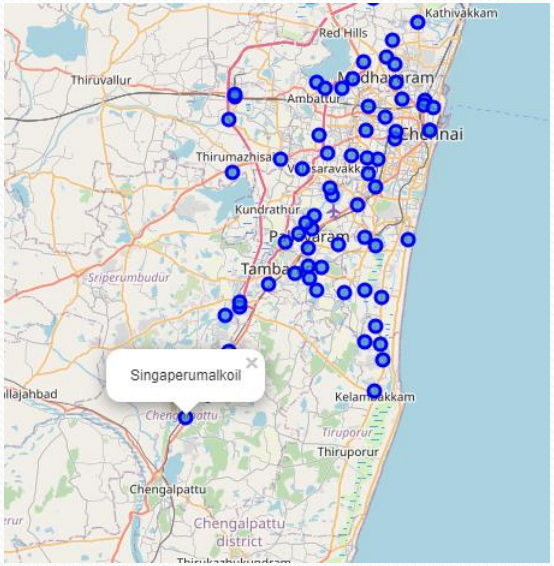
  Then we make calls to obtain various venue data near the specified neighborhoods from which we filter out the shopping mall category. The presence of shopping mall in each location is done by one hot encoding and we find if the given neighborhood as shopping mall or not. Then we cluster the different neighborhoods with the help of KMeans clustering . We visualize the clustered neighborhoods to analyze a pattern. Based on the clusters it would be useful for us to determine a place to open a new shopping mall.

EXPLORATORY DATA ANALYSIS

	Neighborhood
0	Alandur
1	Anna Nagar
2	Ashok Nagar, Chennai
3	Assisi Nagar
4	Ayanavaram

	Neighborhood	Latitude	Longitude
0	Alandur	13.000130	80.200600
1	Anna Nagar	13.083590	80.210200
2	Ashok Nagar, Chennai	13.035390	80.212200
3	Assisi Nagar	13.164570	80.232740
4	Ayanavaram	13.098830	80.232380
5	Chennai city	13.074622	80.242777
6	Chitlapakkam	12.932770	80.143870
7	Chromepet	12.952340	80.144110
8	Cowl Bazaar	12.988610	80.151000
9	Egattur (Kanchipuram District)	12.827250	80.228660
10	Guduvancheri	12.837900	80.053270
11	Iyyapanthangal	13.040920	80.136490
12	K. K. Nagar, Chennai	13.110350	80.213010
13	Kallikuppam, Ambattur	13.129720	80.183000
14	Kamarajapuram	13.120580	80.060470
15	Keelkattalai	12.956150	80.178850
16	Kelambakkam	12.793410	80.220100

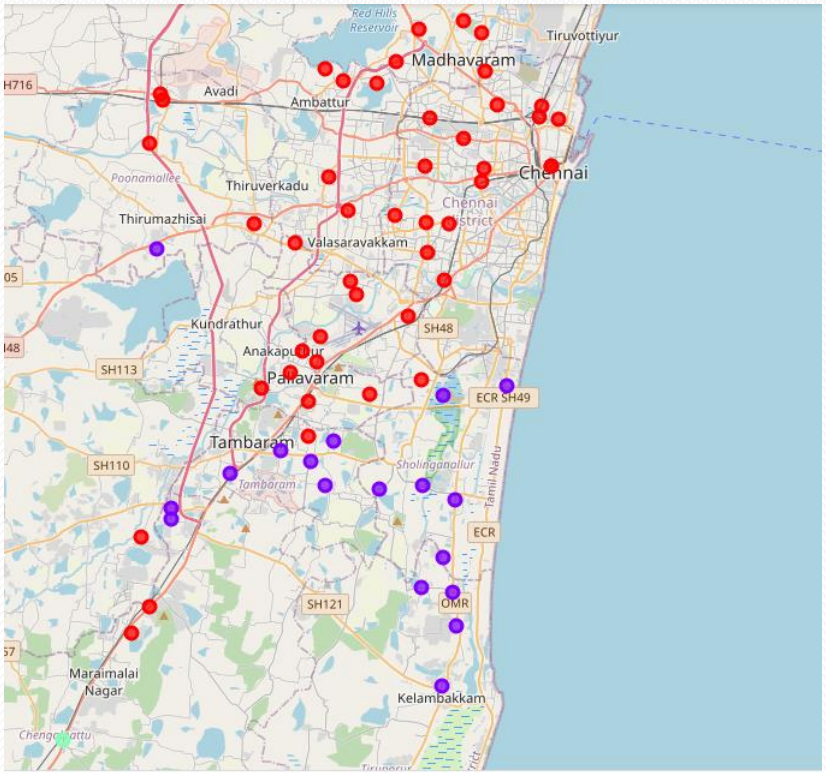
download.jpg



Classification

	Neighborhoods	Shopping Mall
0	Alandur	0.020000
1	Anna Nagar	0.020000
2	Ashok Nagar, Chennai	0.020000
3	Assisi Nagar	0.021277
4	Ayanavaram	0.020000

```
[ ]  kclusters = 3
```



DISCUSSION

- The results show the division of neighborhood into three clusters based on the frequency of shopping malls. The green clustered shows that the neighborhoods are outside Chennai and frequency is high.
- The red colored neighborhood shows that they are centered in the city and have a decent frequency. The malls are already spaced out in the central part of the city.
- The violet colored neighborhoods are suburbs and they contain very low frequency to malls. In some cases there are no malls present.
- AS we see the above depicted diagram, the violet clustered neighborhoods give a bright chance to open a new mall in the city of Chennai as the frequency of the malls is relatively low. The red clustered are city-centric and there will be a heavy competition to build a brand new shopping mall there. The green colored clustered neighborhoods already have shopping malls and also are out of scope for building new ones as it is outside the suburbs of the city.

```
[ ]
kl_merged.loc[kl_merged['Cluster Labels'] == 0]
```

51	Puzhal	0.020202	0	13.159460	80.207180
52	Saidapet	0.020000	0	13.020260	80.221310
56	Senji, Chennai	0.020000	0	13.083620	80.282520
61	Thiruneermalai	0.020000	0	12.959420	80.117180
62	Thiruperumbudur	0.020000	0	13.083620	80.282520
64	Vadapalani	0.020000	0	13.052260	80.211200
65	Virugambakkam	0.020000	0	13.055900	80.193490
46	Polichalur	0.020000	0	12.980710	80.140690
26	Madipakkam	0.020000	0	12.964480	80.208700
67	Washermanpet	0.020000	0	13.109500	80.287010
11	Iyyapanthangal	0.020000	0	13.040920	80.136490
7	Chromepet	0.020000	0	12.952340	80.144110
8	Cowl Bazaar	0.020000	0	12.988610	80.151000
5	Chennai city	0.020000	0	13.074622	80.242777
10	Guduvancheri	0.019231	0	12.837900	80.053270
25	Madhavaram Milk Colony	0.022222	0	13.157520	80.242830
12	K. K. Nagar, Chennai	0.020000	0	13.110350	80.213010
13	Kallikuppam, Ambattur	0.020000	0	13.129720	80.183000
14	Kamarajapuram	0.025000	0	13.120580	80.060470

	Neighborhood	Shopping Mall	Cluster Labels	Latitude	Longitude
63	Thuraipakkam	0.01000	1	12.90318	80.18455
60	Thalambur	0.00000	1	12.84851	80.20841
59	Tambaram	0.01000	1	12.92489	80.12818
57	Sholinganallur	0.01000	1	12.89760	80.22795
55	Semmencherry	0.00000	1	12.86557	80.22051
54	Sembakkam	0.01000	1	12.93055	80.15864
24	Madambakkam	0.00000	1	12.90529	80.15352
9	Egattur (Kanchipuram District)	0.01000	1	12.82725	80.22866
45	Perungalathur	0.01000	1	12.89279	80.06593
44	Perambakkam	0.01000	1	12.90563	80.20907
43	Peerkankaranai	0.00000	1	12.91224	80.09895
39	Pallikaranai	0.01000	1	12.95567	80.22080
37	Palavakkam	0.01000	1	12.96139	80.25743
16	Kelambakkam	0.00000	1	12.79341	80.22010
34	Nazarethpettai	0.01087	1	13.03710	80.05755
30	Mannivakkam	0.00000	1	12.88652	80.06590

```
[ ]  kl_merged.loc[kl_merged['Cluster Labels'] == 2]
```

	Neighborhood	Shopping Mall	Cluster Labels	Latitude	Longitude
49	Pudupakkam	0.044444	2	13.23039	80.21794
58	Singaperumalkoil	0.058824	2	12.76333	80.00350

The different clusters

LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

This project concentrates only on the frequency of the malls to cluster them, but various other factors like population, average salary can be taken into account for taking decisions.

The unavailability of those data for the time being compromises various factors that might determine the new location. I have used regular calls for the foursquare api. We can use premium account to make use of various other objects that is returned by the api.

Do find my original code in the repository.

CONCLUSION

I have scraped data from the webpage, added longitude and latitude locations for each neighborhood using the geocoder python package. Then we use regular calls to get location venues near the referenced location using foursquare api. We only choose shopping malls from categories. We then cluster the neighborhoods based on the frequency of shopping malls. We analyze the clusters and it is finally decided that Violet colored neighborhoods are the best place to build a new shopping mall in the city of Chennai. The best place would be MADAMBAKKAM or MANIVAKKAM as it is in the violet zone and also has no malls.