IBM Applied Data Science Capstone

Optimal Locations for Hospitals in Chennai city

ABILASH B

Introduction:

Chennai is the capital of the Indian state of Tamil Nadu. Located on the Coromandel Coast of the Bay of Bengal, it is one of the largest cultural, economic and educational center of south India. 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.

Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus. Most people infected with the COVID-19 virus will experience mild to moderate respiratory illness and recover without requiring special treatment.

The rising number of cases has resulted in a deepening healthcare crisis that has gripped several states. India is reeling under a severe second wave of Covid-19 and many states are struggling to cope with the rising numbers. The second wave has hit a country which was unprepared. This has led to increased no. of cases where the hospitals are overcrowded and there is also shortage of oxygen supply in some areas.

Business Problem:

This project aims to analyze and solve the following problems

1. List and visualize all the hospitals in and around the Chennai city
2. Identify the neighborhoods around the hospitals
3. What are all the best locations to open new public health centers
4. Identify areas that lack hospital centers compared to other areas
5. To alleviate over-crowding in hospitals and enhance the medical facilities to the citizens

Target Audience:

The purpose of this project is to help people through identification of optimal placement of new health centers by exploring the facilities around their neighborhood. It will help people navigate to the nearest neighborhoods with medical facilities out of numbers of other postal areas in Chennai, India.

Data Description:

Chennai city's demographics show that it is a large and ethnically diverse metropolis. With its diverse society, comes diverse infrastructure. There are many different kinds of infrastructure in Chennai city, each

belonging to different categories like hospitals, schools, colleges and apartments.

1. Chennai Postal codes
   a. Data source:

      https://www.mapsofindia.com/pincode/india/tamil-nadu/chennai/

   b. Description:

      This data contains list of major boroughs and their postal codes

2. Geocoding API
   a. Data source:

      https://developers.google.com/maps/documentation/geocoding/overview

   b. Description:

      The Geocoding API provides a direct way to access these services via an HTTP request. We leverage this API to obtain latitude and longitudes of hospitals

3. Hospital Location information
   a. Data source:

      https://nationalinsurance.nic.co.in/sites/default/files/Chennai%20PPN%20List%20of%20Hospitals.pdf

b. Description:

The data source contains list of hospitals and their addresses which are located in and around the Chennai city

4. Foursquare API

a. Data source:

https://developer.foursquare.com/

b. Description:

Foursquare helps us identify the various venues near our neighborhood through we'll able to identify the facilities around the hospitals
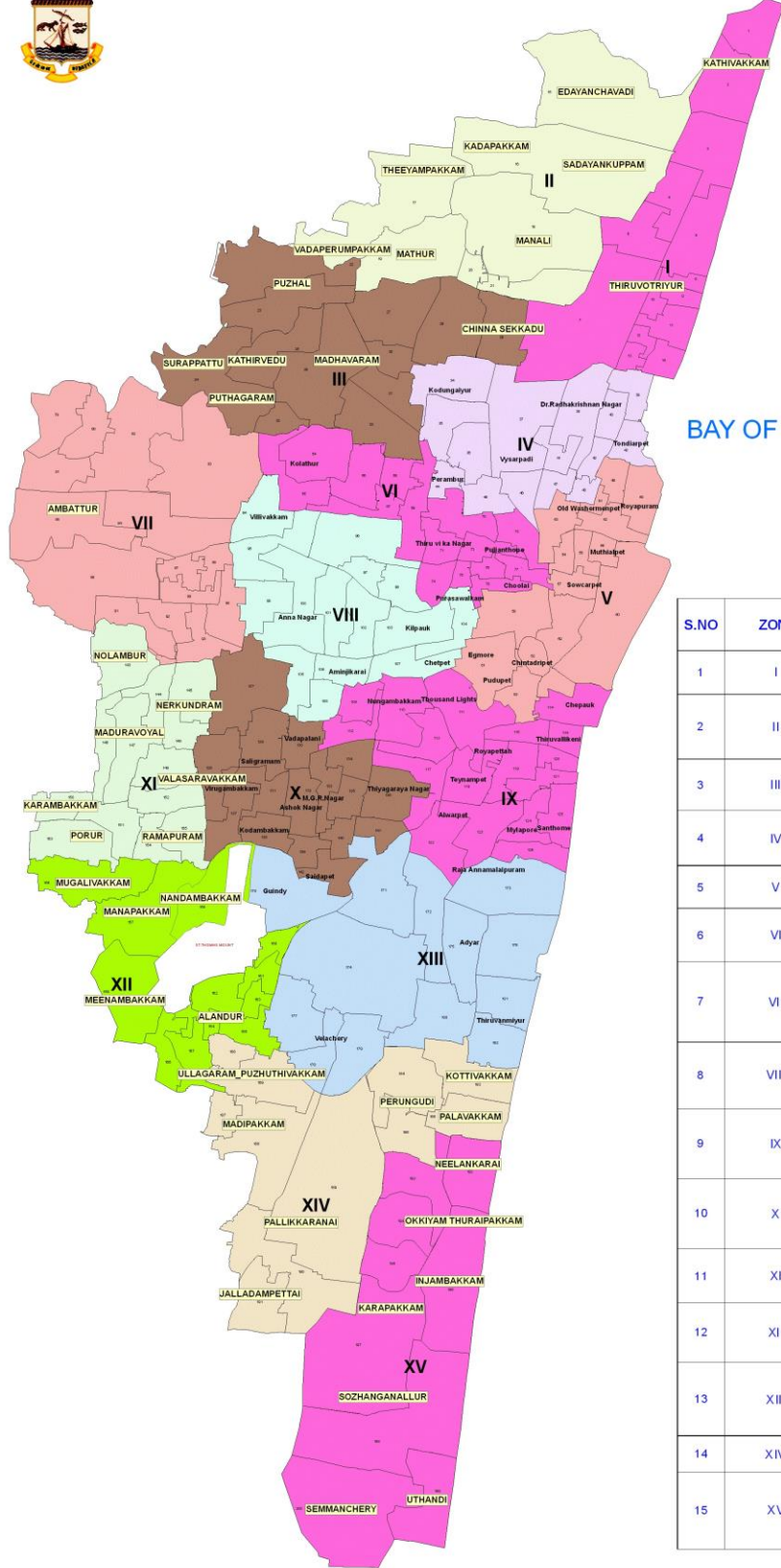
Work flow:

Our plan is to leverage the aforementioned data sources to explore and examine the medical centers throughout the city. This project would use various data science skills including web scraping, Geocoding API, Foursquare API, data cleaning, data wrangling, map visualization (Folium) and a clustering based machine learning approach (K-means clustering).

Methodology:

CORPORATION OF CHENNAI

BAY OF BENGAL

KATHIVAKKAM
EDAYANCHAVADI
KADAPAKKAM
SADAYANKUPPAM
THEEYAMPAKKAM
II
MANALI
VADAPERUMPAKKAM MATHUR
PUZHAL
THIRUVOTRIYUR
I
CHINNA SEKKADU
SURAPPATTU KATHIRVEDU MADHAVARAM
III
PUTHAGARAM
Kodungalyur
Dr.Radhakrishnan Nagar
IV
Tondiarpet
Vysarpadi
Perambur
Old Washermenpet Royapuram
Kolathur
VI
Villivakkam
AMBATTUR
VII
Thiru vi ka Nagar
Pujianthope
Muthialpet
Cheolai
Sowcarpet
V
Anna Nagar
VIII
Purasawalkam
Kilpauk
Chetpet
Egmore
Chintadripet
Aminjikarai
Pudupet
NOLAMBUR
NERKUNDRAM
Chepauk
Nungambakkam Thousand Lights
MADURAVOYAL
Vadapalani
Thiruvallikeni
Saligramam
Royapettah
XI VALASARAVAKKAM
Virugambakkam
X M.G.R.Nagar
Teynampet
IX
KARAMBAKKAM
Ashok Nagar
Thiyagaraya Nagar
Alwarpet
Mylapore Santhome
PORUR RAMAPURAM
Kodambakkam
Saidapet
Raja Annamalaipuram
MUGALIVAKKAM
NANDAMBAKKAM
Guindy
MANAPAKKAM
S.T.THOMAS MOUNT
Adyar
XIII
XII
MEENAMBAKKAM
ALANDUR
Velachery
Thiruvanmiyur
ULLAGARAM PUZHUTHIVAKKAM
KOTTIVAKKAM
PERUNGUDI
MADIPAKKAM
PALAVAKKAM
NEELANKARAI
XIV
PALLIKKARANAI
OKKIYAM THURAIPAKKAM
INJAMBAKKAM
JALLADAMPETTAI
KARAPAKKAM
XV
SOZHANGANALLUR
SEMMANCHERY
UTHANDI

| S.NO	ZONE	NAME OF THE ZONE	WARD NUMBER
1	I	THIRUVOTRIYUR	1 TO 14
2	II	MANALI	15 TO 21
3	III	MADHAVARAM	22 TO 33
4	IV	TONDIARPET	34 TO 48
5	V	ROYAPURAM	49 TO 63
6	VI	THIRU-VI-KA NAGAR	64 TO 78
7	VII	AMBATTUR	79 TO 93
8	VIII	ANNA NAGAR	94 TO 108
9	IX	TEYNAMPET	109 TO 126
10	X	KODAMBAKKAM	127 TO 142
11	XI	VALASARAVAKKAM	143 TO 155
12	XII	ALANDUR	156 TO 167
13	XIII	ADYAR	170 TO 182
14	XIV	PERUNGUDI	168,169 183 TO 191
15	XV	SOZHANGANALLUR	192 TO 200

Neighborhoods Visualization:
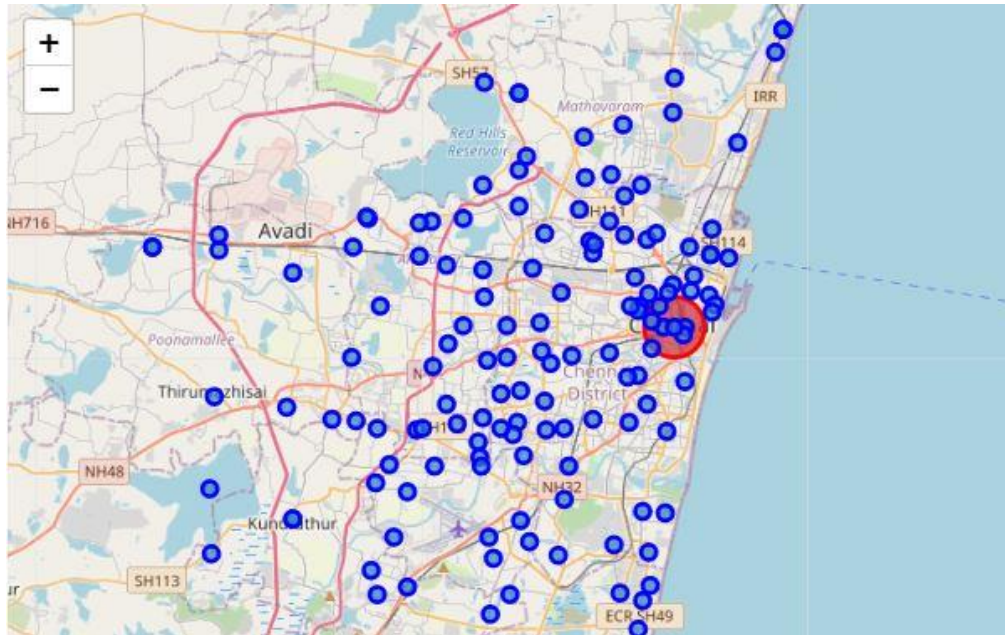
Chennai Hospital Visualizations:
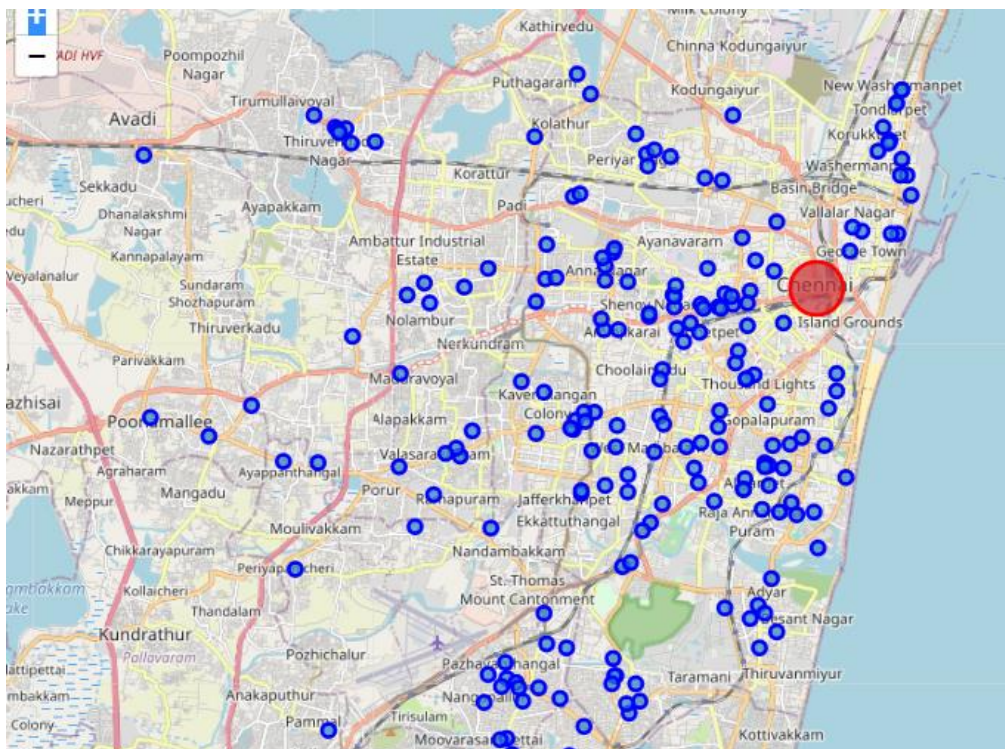
Hospital & Neighborhoods Visualizations:
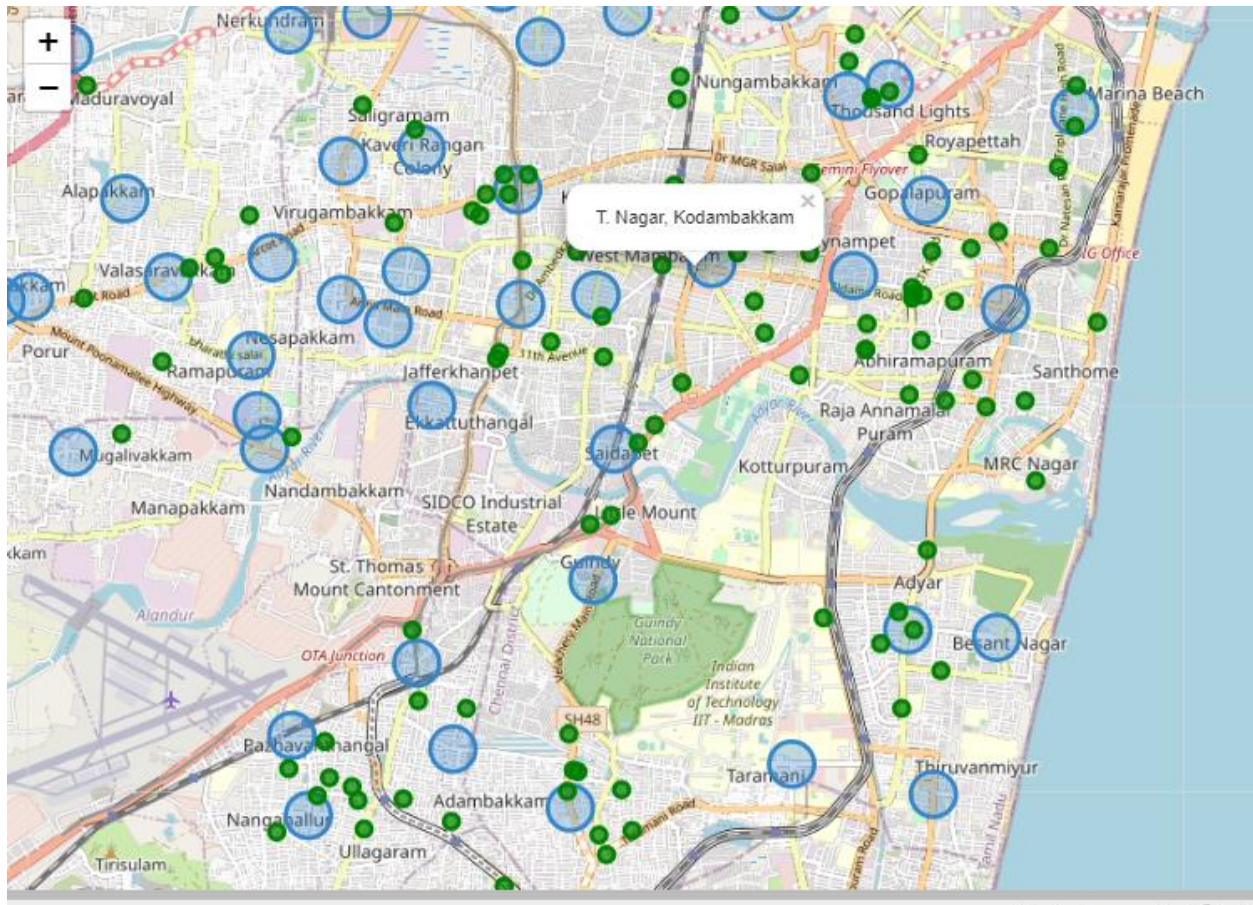
In this project we will direct our efforts on detecting areas of Chennai that have low hospital density, particularly those with those with a lot of venues. We will limit our analysis to area ~5km around each neighborhood.

In first step we have collected the required **data: location and type (category) of every hospital and venue within 5km from each neighborhood center** We have obtained the hospital data separately along with the latitudes and longitudes

The next step is we group the venues based on their categories: Services, Sports, entertainment, Hospital and others

We one hot encode the venue categories. We identify their total in each neighborhood. The venue categories are put in to the group/bins specified above. Then we compute, no.of venue groups per hospital in a given neighborhood.

We use this information to identify areas with least hospitals via clustering.

Discussion:
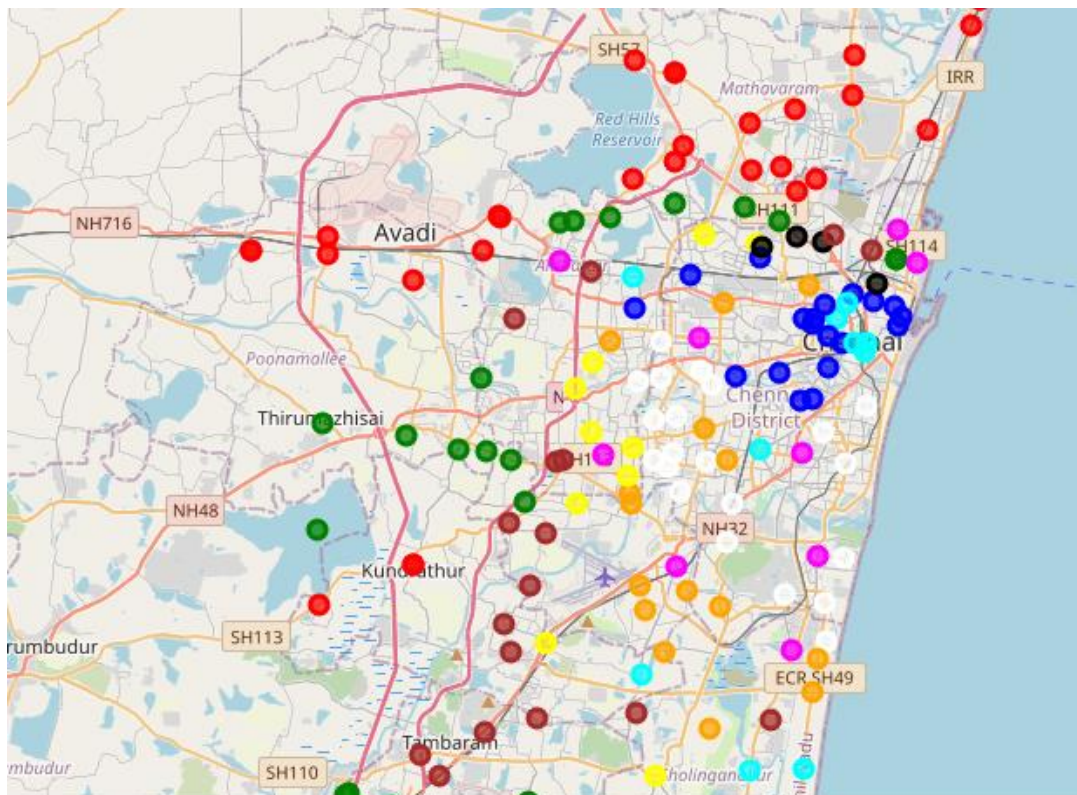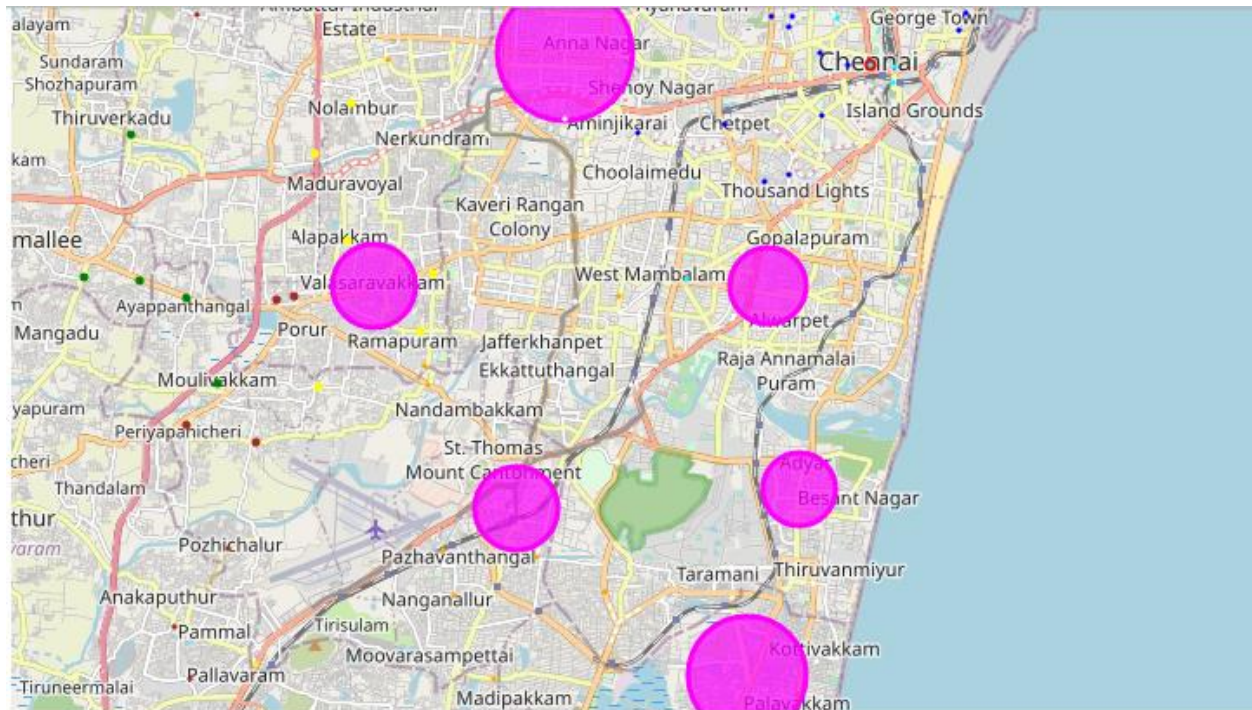
Results:

Our results have identified regions with lower hospital facilities.

We identify that the hospitals are placed in a nash equilibrium structure.

Few of the clusters indicate even thoughy they're not very densely populated, they're work/ commute related places. These regions need to have even more better medical services at their disposal to meet the emergency situations.

Further more, few of the hotels & restaurants near the hospitals can be used as temporary COVID care regions.

By this way, new camps can be opened in shorter time. Venues with amenities, restaurants and sporting camps can come forward to open COVID care camps.

Regions in Cluster 2 and Cluster 5 can be the places to start considering opening new facilities

Conclusion:

Purpose of this project was to identify Chennai areas close to center with low number of hospitals in order to aid stakeholders in narrowing down the search for optimal location for a new COVID care center. By calculating hospital density distribution from Foursquare data we have first identified general boroughs that justify further analysis. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decission on optimal COVID centers will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like housing areas, transport facilities, venue availability, social and economic dynamics of every neighborhood etc.