DETECTING NEARBY HOSPITAL VENUES FOR MEDICAL EMERGENCIES AND FINDING COVID-19 SEVERITY IN TAMILNADU DISTRICTS

1. INTRODUCTION

The swift spreading of the novel corona virus COVID-19 had brought the entire world to a stand-still. Researchers had suggested that its repercussions are still premature and yet to be realized completely. This pandemic has caught the entire world on tenterhooks. Scientists world over are running against time to develop an anti-body for this COVID-19 that would bring a cure and curtail its further advancements.

1.1 Problem Background

The state of interest, Tamilnadu (TN), located in southern India had been chosen for analyzing the Covid-19 impact. As we are in a covid-19 pandemic situation, citizens of populated cities often get infected by coming in close contact with an infected person whenever he or she coughs or sneezes. This covid-19 disease may lead to death for people who have chronic medical disorders and granny adults.

The urban cities Chennai, Madurai, Trichy and so on of TN, are considered as densely populated cities of TN. Most of the other district people reside in these urban cities, since they account for very high job market. In such populated urban cities, the spreading of Covid-19 virus is highly alarming. These cities are rich in having enough hospitals within its circumference to serve the needy. Besides, at the rural areas of TN, there are very minimum hospitals in close proximity.

Finding nearby hospitals based on current location (Latitude, Longitude) in case of medical emergencies for the patients is vital. To condense these odds of

happening, a health care category finder that would extract nearby locations of medical centers is important for providing medical emergencies on demand.

From the health ministry affairs, it is very important in finding out the severity of covid-19 cases over the districts of TN State. Many urban and rural districts of TN are affected by this disease and the count is highly alarming, it becomes necessary to group affected districts based on severity level which would go a long way in assisting health organizations to give importance for such places where the spreading of the disease is rapid.

1.2 Problem Description

When a patient is looking for a hospital in need of medical emergency, it is often better to locate existing hospitals and find the category of such hospitals. The purpose of this project is to find nearby hospital locations within a minimal radius from the current location in an optimal manner. The proposed project focuses on the districts with dense or less populated districts, finds the nearby hospitals based on Foursquare API. The distance between the current location and the hospital locations are computed. The project suggests most optimal neighborhood hospitals sorted by distance as a result.

Also, this project utilizes clustering algorithm to cluster districts of TN State's pandemic situation by considering fields such as number of active cases, recovered cases and deaths on day-to-day basis using web scrapping of covid-19 data from Wikipedia. From the clustering analysis, the impact of covid-19 spreading over the districts is analyzed.

The major Insights influenced in the problem are,

- Suggests the number of prevailing hospitals in the neighborhood areas
- Distance to the neighborhood hospitals from the current location
- Grouping of similar districts based on covid-19 cases where utmost care has to be taken

2. DATA DESCRIPTION

In this section, the various sources of data that will be used to analyze the problem of detecting nearby hospitals based on distance.

2.1 Data Collection

The following data sources are utilized to extract/generate the required information:

1. Location addresses of Tamil Nadu Districts will be obtained using Google Maps API reverse geocoding (geopy.geocoders)

2. Number of neighborhood hospitals and their locations will be obtained using Foursquare API

3. Latitude and Longitude Dataset of Tamil Nadu state will be obtained from the **URL:** https://www.mapsofindia.com/lat_long/tamilnadu/

   This dataset comprises of the following columns:

   Location: District Name(s) of TN

   Latitude: Latitude data of the District (DMS Format*)

   Longitude: Longitude data of the District (DMS Format*)

   *DMS Format: (° : degree, ' : minute, " : second)+Direction(N,E,W,S)

4. COVID-19 pandemic data of Tamil Nadu, India medical cases by district will be obtained from the **URL:**https://en.m.wikipedia.org/wiki/Template:COVID19_pandemic_data/India/Tamil_Nadu_medical_cases_by_district

   This dataset comprises of the following columns:

   District: TN District Name

   Diagnosed cases: Total No. of Diagnosed cases for COVID-19

   Deaths: Total No. of Deaths caused by COVID-19

   Recovered cases: Total No. of Recovered cases from COVID-19

   Active cases: Total No. of COVID-19 cases still active

Population: Total population of the District

Cases per M: Cases per Million

Last case reported on: Last reported date

3. DATA PRE-PROCESSING

In this section the core components required for the COVI-19 analysis of in Tamilnadu districts and detection of nearby hospitals for medical emergencies.

3.1 Dataset creation and Cleaning

The data for this project has been retrieved from multiple sources like TN latitude and longitude data and wiki page comprises of COVID-19 identified cases.

3.1.1 COVID-19 pandemic data

The COVID-19 pandemic data is extracted from Wikipedia webpages using Web Scraping with the help of BeautifulSoup library in Python. Taking a closer look at the dataset, the collected data contains the identifier comma (,) to represent thousands in the dataset.

Web Scrapping:

```
url = "https://en.m.wikipedia.org/wiki/Template:COVID-19_pandemic_data/India/Tamil_Nadu_medical_cases_by_district"
source = requests.get(url).text


soup = BeautifulSoup(source,"html.parser")
table=soup.find('table')
```

Retriving table Information from web

```
for row_cell in table.find_all('tr'):
    row=[]
```

```
    for data_cell in row_cell.find_all('th'):
      row.append(data_cell.text.strip())
      for data_cell in row_cell.find_all('td'):
        row.append(data_cell.text.strip())
    if len(row)==8:
      df.loc[len(df)]=row
```

df.head()

	District	Diagnosed cases	Deaths	Recovered cases	Active cases	Population	Cases per M	Last case reported on
0	District	Diagnosed cases[a]	Deaths	Recovered cases	Active cases[b]	Population	Cases per M	Last case reported on
1	Tamil Nadu	102,721	1,385	58,378	42,955	75,695,000	1,357	3 July 2020
2	Ariyalur	463	0	425	38	752,481	615	1 July 2020
3	Chengalpattu	6,139	106	3,113	2,920	2,556,423	2,401	3 July 2020
4	Chennai	64,689	996	40,111	23,582	7,088,000	9,127	3 July 2020

      The cleaning of the collected data from the web table is performed to create a uniform and usable dataset before they could be subjected to further analysis.

Convert required columns to number type (Removing Comma's)

```
import locale
from locale import atof
locale.setlocale(locale.LC_NUMERIC, '')
cases_dt[['Diagnosed cases', 'Deaths', 'Recovered cases','Active cases', 'Population', 'Cases per M']].applymap(atof)
```

Converting object to int type

```
cases_dt[['Diagnosed cases', 'Deaths', 'Recovered cases','Active cases', 'Population', 'Cases per M']] = cases_dt[['Diagnosed cases', 'Deaths', 'Recovered cases','Active cases', 'Population', 'Cases per M']].apply(pd.to_numeric)
```

```
print(cases_dt.dtypes)


District             object
Diagnosed cases      int64
Deaths               int64
Recovered cases      int64
Active cases         int64
Population           int64
Cases per M          int64
Last case reported on    object
```
dtype: object

```
cases_dt.head()
```

	District	Diagnosed cases	Deaths	Recovered cases	Active cases	Population	Cases per M	Last case reported on
0	Chengalpattu	8553	169	5695	2689	2556244	3346	14 July 2020
1	Chennai	79662	1295	62552	15815	4646732	17144	14 July 2020
2	Coimbatore	1480	10	338	1132	3458045	428	14 July 2020
3	Cuddalore	1565	6	1154	405	2605914	601	14 July 2020
4	Dharmapuri	260	1	103	156	1506843	173	14 July 2020

3.1.2 Latitude and Longitude Data of TN Districts

The Latitude and Longitude data had been extracted from the mapsofindia web page. The spherical coordinates (latitudes and longitudes) are in degrees-minutes-seconds (DMS) format. DMS data is recorded in the format: (° : degree, ' : minute, " : second) with direction (N,E,W,S). The range of minutes and seconds values is from 0 to 60.

Importing latitude and Longitude of TN Districts

```
import pandas as pd
```

```
Lat_Long_TN = pd.read_csv('TN Lat Lon12.csv', encoding= 'unicode_escape')
```

	Active	Location	Latitude	Longitude
0	1	Ariyalur	11° 8' N	79° 4' E
1	1	Chengalpattu	12° 42' N	80° 01' E
2	1	Chennai	13° 04' N	80° 17' E
3	1	Coimbatore	11° 00' N	77° 00' E
4	1	Cuddalore	11° 43' N	79° 49' E

Data Transformation (Convert DMS (degrees-minutes-seconds) to DD (decimal degrees) Coordinates)

The geographic coordinates can be converted to decimal degrees. It's just another way to represent the same location in a different format which can be used to plot markers over the map.

```python
import re
def dms2dd(degrees, minutes, direction):
    dd = float(degrees) + float(minutes)/60;
    if direction == 'S' or direction == 'W':
        dd *= -1
    return dd;
def parse_dms(lat,lon):
    lat_part = re.split('[^\d\w]+', lat)
    lon_part = re.split('[^\d\w]+', lon)

    lati = dms2dd(lat_part[0], lat_part[1], lat_part[2])
    lng = dms2dd(lon_part[0], lon_part[1], lon_part[2])
    return (lati,lng)
```

```
DD_Lat=[]
DD_Lng=[]
for i in range(len(Lat_Long_TN)):
  #print(Lat_Long_TN.loc[i, 'Latitude'], Lat_Long_TN.loc[i, 'Longitude'])
   latt=parse_dms(Lat_Long_TN.loc[i, 'Latitude'], Lat_Long_TN.loc[i, 'Longitude'])
  DD_Lat.append(latt[0])
  DD_Lng.append(latt[1])


Lat_Long_TN['DD_Lat']=DD_Lat
Lat_Long_TN['DD_Lng']=DD_Lng


Lat_Long_TN.head()
```

	Active	Location	Latitude	Longitude	DD_Lat	DD_Lng
0	1	Ariyalur	11° 8' N	79° 4' E	11.133333	79.066667
1	1	Chengalpattu	12° 42' N	80° 01' E	12.700000	80.016667
2	1	Chennai	13° 04' N	80° 17' E	13.066667	80.283333
3	1	Coimbatore	11° 00' N	77° 00' E	11.000000	77.000000
4	1	Cuddalore	11° 43' N	79° 49' E	11.716667	79.816667

4. METHODOLOGY

After data cleansing and data exploration on the two datasets, the problem finding nearby hospital location based on distance is implemented. Also, the impact of COVID-19 over TN districts is analyzed using K-means clustering technique.

4.1 Finding nearby Hospital Locations based on FourSquare API

The coordinates of a current location in which the nearby hospitals being searched, utilized using geocoding API and Nominatim.

4.1.1 Geocoding API

Geocoders are used to resolve a location from a string. Geocoders accept credentials such as CATEGORY_ID, CLIENT_ID, CLIENT_SECRET, VERSION used to retrieve nearby location within a specified radius. It interacts with any API like foursquare to locate its neighborhoods.

Coordinates extraction using geocoder API

address = 'Alwarpet, Chennai, Tamilnadu'

```
geolocator1 = Nominatim(user_agent="ny_explorer")
location1 = geolocator1.geocode(address)
latitude1 = location1.latitude
longitude1 = location1.longitude
print('The geograpical coordinate of {} are {}, {}.'.format(address,latitude1, longitude1))
```

The geograpical coordinate of Alwarpet, Chennai, Tamilnadu are 13.0338602, 80.2545491.

4.1.2 FourSquare API Credentials including category

The Foursquare API returns a formatted JSON response that contains the credentials like CATEGORY_ID, CLIENT_ID, CLIENT_SECRET, and VERSION. It is a social location service that allows users to explore the world. The categories are supported globally by default.

Setting radius to find hospitals for Alwarpet, Chennai, Tamilnadu Location

```
LIMIT = 100
radius = 2000
url = 'https://api.foursquare.com/v2/venues/explore?&categoryId={}&client_id={
}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.
format(
    CATEGORY_ID,
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    latitude1,
    longitude1,
    radius,
    LIMIT)
url
```

Generated Hospital category data for Alwarpet, Chennai, Tamilnadu Location

A small sample of generated neighborhood hospitals along with geographical coordinate addresses returned by FourSquare API under hospital category has been shown in the table below.

Neighborhood hospital data returned by FourSquare API

```
nearby_venues.head(100)
```

	name	categories	lat	lng
0	kauvery hospital	Hospital	13.038214	80.257067
1	Balaji Dental & Craniofacial Hospital	Dentist's Office	13.035352	80.247048
2	Apollo Speciality Hospital	Hospital	13.033841	80.245289
3	ALL SMYLES COSMETIC DENTAL CENTRE CHENNAI INDIA	Dentist's Office	13.033604	80.250548
4	SKS Veterinary Hospital	Hospital	13.028984	80.258728
5	Madras ENT Research Foundation	Hospital	13.027266	80.254176
6	St. Isabel's Hospital	Hospital	13.039236	80.262297
7	Oliva Skin and Hair Clinic Alwarpet	Hospital	13.042096	80.259281
8	Sankara Nethralaya	Hospital	13.026432	80.261102
9	Dr. Agarwal's Eye Hospital	Hospital	13.046304	80.255363
10	Bharathiraja Hospital	Hospital	13.048036	80.245013
11	Seethapathy Clinic and Hospital	Hospital	13.048841	80.264537

4.2 Clustering severity of COVID-19 impact on TN Districts

The two datasets containing geographical coordinates information and COVID-19 cases are fused based on district column. Then, K-means clustering technique is applied to cluster similar districts based on the cases reported.

4.2.1 Data Integration

The fields of the above two datasets containing district-wise covid-19 cases and district-wise geographical coordinates are merged based on district column to visualize and gain further insights from the obtained data.

	Active	Location	Latitude	Longitude	DD_Lat	DD_Lng	District	Diagnosed cases	Deaths	Recovered cases	Active cases	Population	Cases per M	Last case reported on
0	1	Chengalpattu	12° 42' N	80° 01' E	12.700000	80.016667	Chengalpattu	6,139	106	3,113	2,920	2,556,423	2,401	3 July 2020
1	1	Chennai	13° 04' N	80° 17' E	13.066667	80.283333	Chennai	64,689	996	40,111	23,582	7,088,000	9,127	3 July 2020
2	1	Coimbatore	11° 00' N	77° 00' E	11.000000	77.000000	Coimbatore	645	1	252	392	3,172,578	203	3 July 2020
3	1	Cuddalore	11° 43' N	79° 49' E	11.716667	79.816667	Cuddalore	1,143	5	813	325	2,600,880	439	3 July 2020
4	1	Dharmapuri	12° 08' N	78° 13' E	12.133333	78.216667	Dharmapuri	107	0	41	66	1,502,900	71	3 July 2020
5	1	Dindigul	10° 22' N	78° 00' E	10.366667	78.000000	Dindigul	618	7	310	301	2,161,367	286	3 July 2020

4.2.2 K-means Clustering

For the clustering process, the K-means approach was used, which is an unsupervised machine learning algorithm. K-means will be computationally faster than other clustering algorithms.

This process also requires to set the parameter for the number of clusters. To be able to identify the optimal number for this parameter, the **silhouette score** was used.

The silhouette coefficient is calculated by,

Silhouette Coefficient = (x-y)/ max(x,y)

The district wise COVID-19 data is trained using K-means Clustering Algorithm to get the desired clusters. K-means Clustering is chosen to cluster TN districts based on COVID-19 Severity.

5. RESULTS

The COVID-19 district wise cases contains fields such as District, Diagnosed cases, Deaths, Recovered cases, Active cases, Population, Cases per M, Last case reported. The dataset numeric columns are visualized in bar chart in having district in x-axis and number of cases in y-axis.
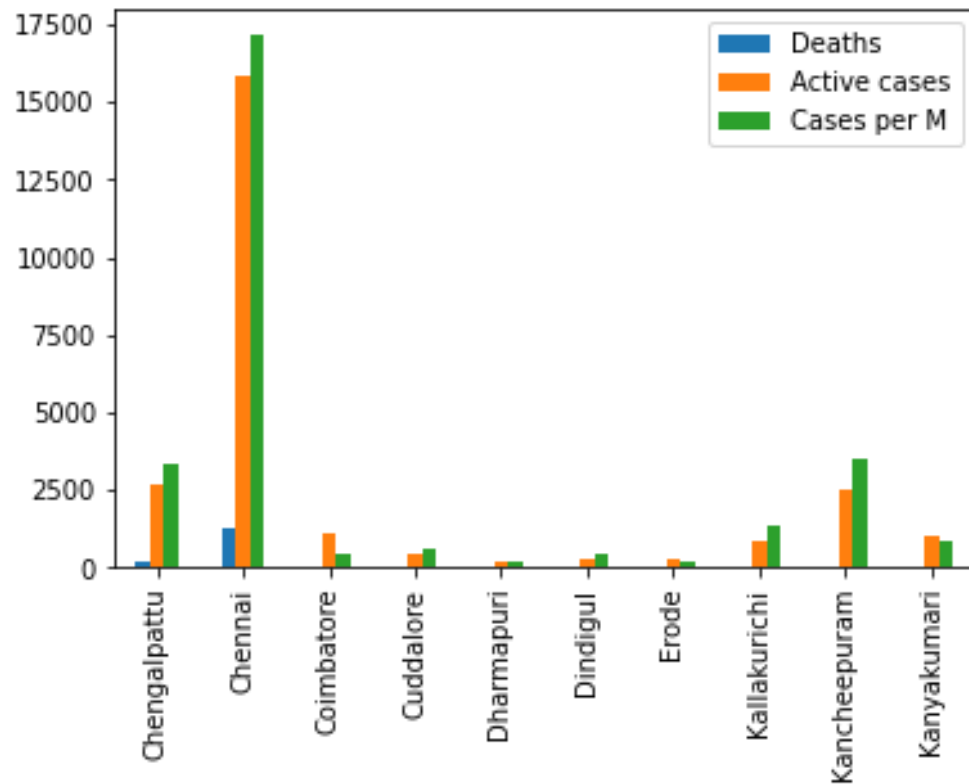
The heap map is used to visualize the cases impact over the numeric columns is shown
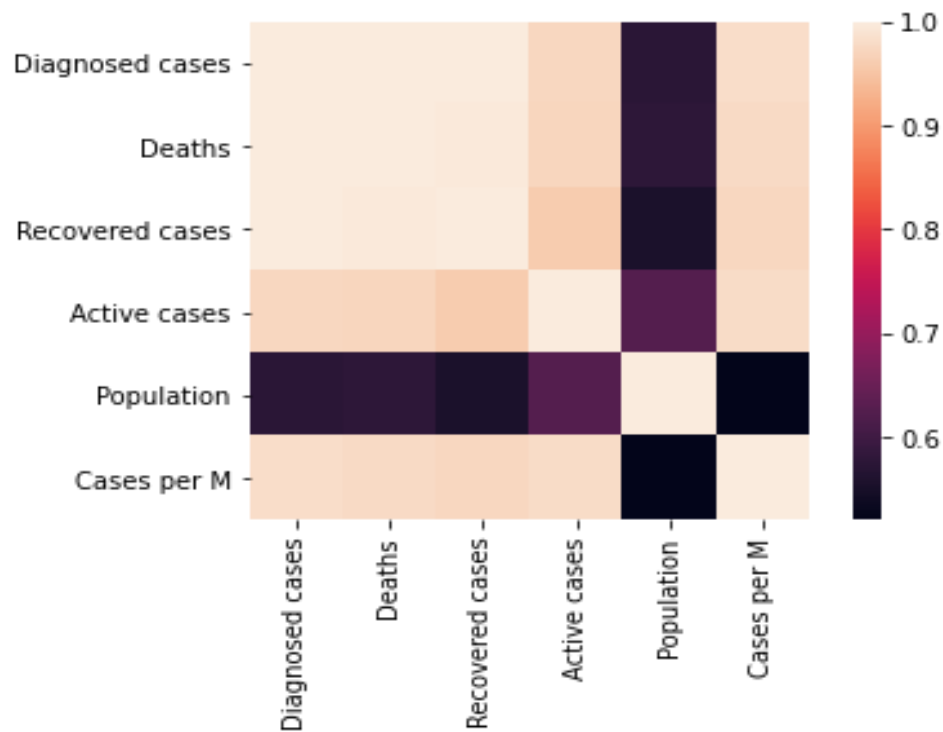
5.1 Calculating Distance for the hospitals extracted

The location coordinates of **'Alwarpet, Chennai, Tamilnadu'** is extracted using geocoder and Nominatim.

```
address = 'Alwarpet, Chennai, Tamilnadu'


geolocator1 = Nominatim(user_agent="ny_explorer")
location1 = geolocator1.geocode(address)
latitude1 = location1.latitude
longitude1 = location1.longitude
print('The geograpical coordinate of {} are {}, {}.'.format(address,latitude1, longitude1))
```

The geograpical coordinate of Alwarpet, Chennai, Tamilnadu are 13.0338602, 80.2545491.

The distance of the neighborhood hospitals is calculated based on the current location coordinates.

```
nearby_venues.head()
```

	name	categories	lat	lng	Distance
3	ALL SMYLES COSMETIC DENTAL CENTRE CHENNAI INDIA	Dentist's Office	13.033604	80.250548	0.434329
0	kauvery hospital	Hospital	13.038214	80.257067	0.555733
4	SKS Veterinary Hospital	Hospital	13.028984	80.258728	0.706312
5	Madras ENT Research Foundation	Hospital	13.027266	80.254176	0.734405
1	Balaji Dental & Craniofacial Hospital	Dentist's Office	13.035352	80.247048	0.829344

## 5.2 Clustering based on distance and adding cluster labels to the dataset

The neighborhood hospitals are clustered using K-means clustering and the cluster labels columns are added to the dataset.

```
nearby_venues.tail()
```

	Cluster Labels	name	categories	lat	lng	Distance
7	2	Oliva Skin and Hair Clinic Alwarpet	Hospital	13.042096	80.259281	1.049481
8	2	Sankara Nethralaya	Hospital	13.026432	80.261102	1.089126
9	2	Dr. Agarwal's Eye Hospital	Hospital	13.046304	80.255363	1.386506
10	1	Bharathiraja Hospital	Hospital	13.048036	80.245013	1.884593
11	1	Seethapathy Clinic and Hospital	Hospital	13.048841	80.264537	1.986330

5.3 Mapping Clustered Coordinates

The neighborhood hospitals are clustered based on distance. The nearby hospitals are plotted and visualized based on distance.
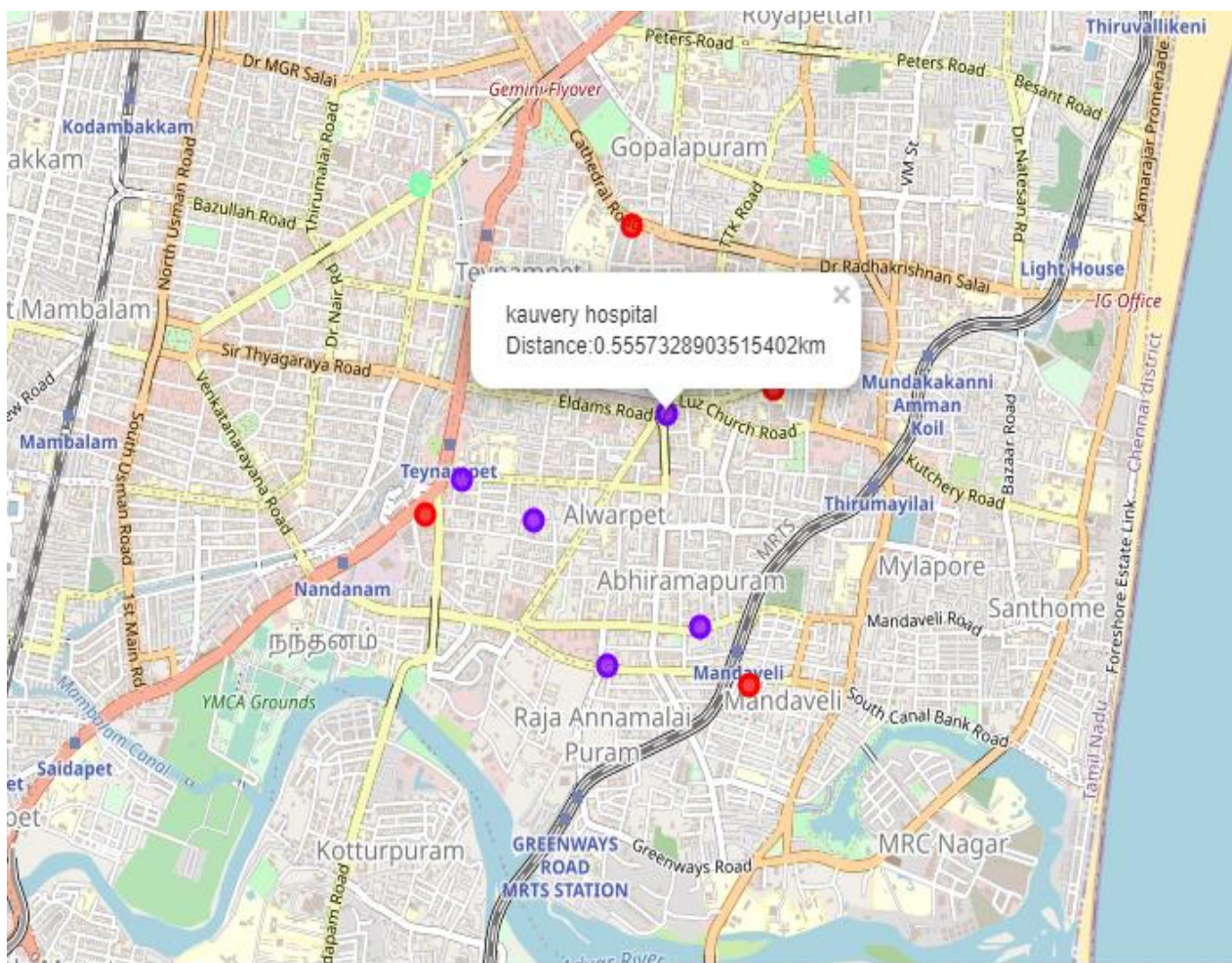
5.4 Clustering severity of COVID-19 impact on TN districts based on FourSquare API

In order to cluster the severity of COVID-19 impact over TN districts, the latitude and longitude data of TN districts is merged with COVID-19 dataset by inner join on district column and location column in the datasets. The merged dataframe is shown below,

```
Merge_df.head()
```

|	Cluster Labels	Active	Location	Latitude	Longitude	DD_Lat	DD_Lng	District	Diagnosed cases	Deaths	Recovered cases	Active cases	Population	Cases per M	Last case reported on
0	3	1	Chengalpattu	12° 42' N	80° 01' E	12.700000	80.016667	Chengalpattu	8553	169	5695	2689	2556244	3346	14 July 2020
1	4	1	Chennai	13° 04' N	80° 17' E	13.066667	80.283333	Chennai	79662	1295	62552	15815	4646732	17144	14 July 2020
2	1	1	Coimbatore	11° 00' N	77° 00' E	11.000000	77.000000	Coimbatore	1480	10	338	1132	3458045	428	14 July 2020
3	3	1	Cuddalore	11° 43' N	79° 49' E	11.716667	79.816667	Cuddalore	1565	6	1154	405	2605914	601	14 July 2020
4	0	1	Dharmapuri	12° 08' N	78° 13' E	12.133333	78.216667	Dharmapuri	260	1	103	156	1506843	173	14 July 2020

5.5 Optimal number of clusters

The districts are divided into n clusters based on the severity of COVID-19 impact, where n is the number of clusters found using the optimal approach.
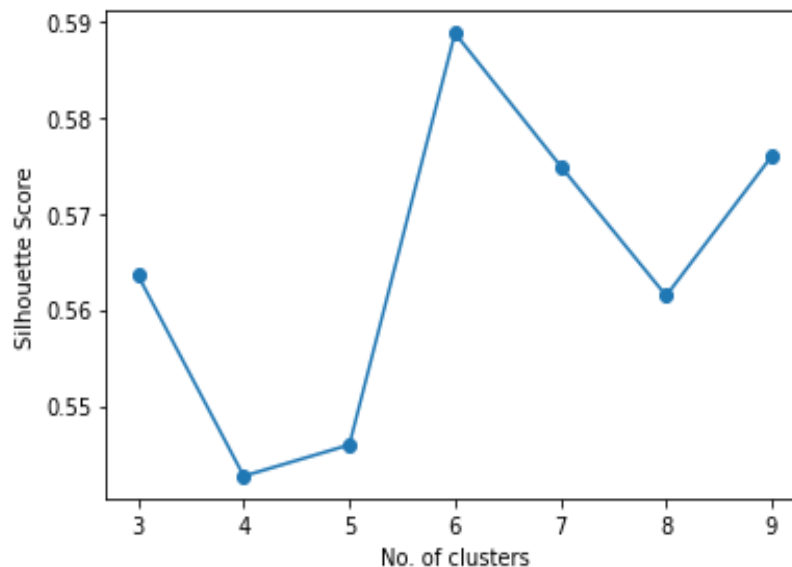
As per silhouette score, n=6 is opted as the best number for cluster separation

5.6 Clustering TN Districts based on COVID-19 Severity

The TN districts are visualized using different colors so as to make them distinguishable in the generated map. From the silhouette analysis, six classes are used to distinguish TN districts based on severity.
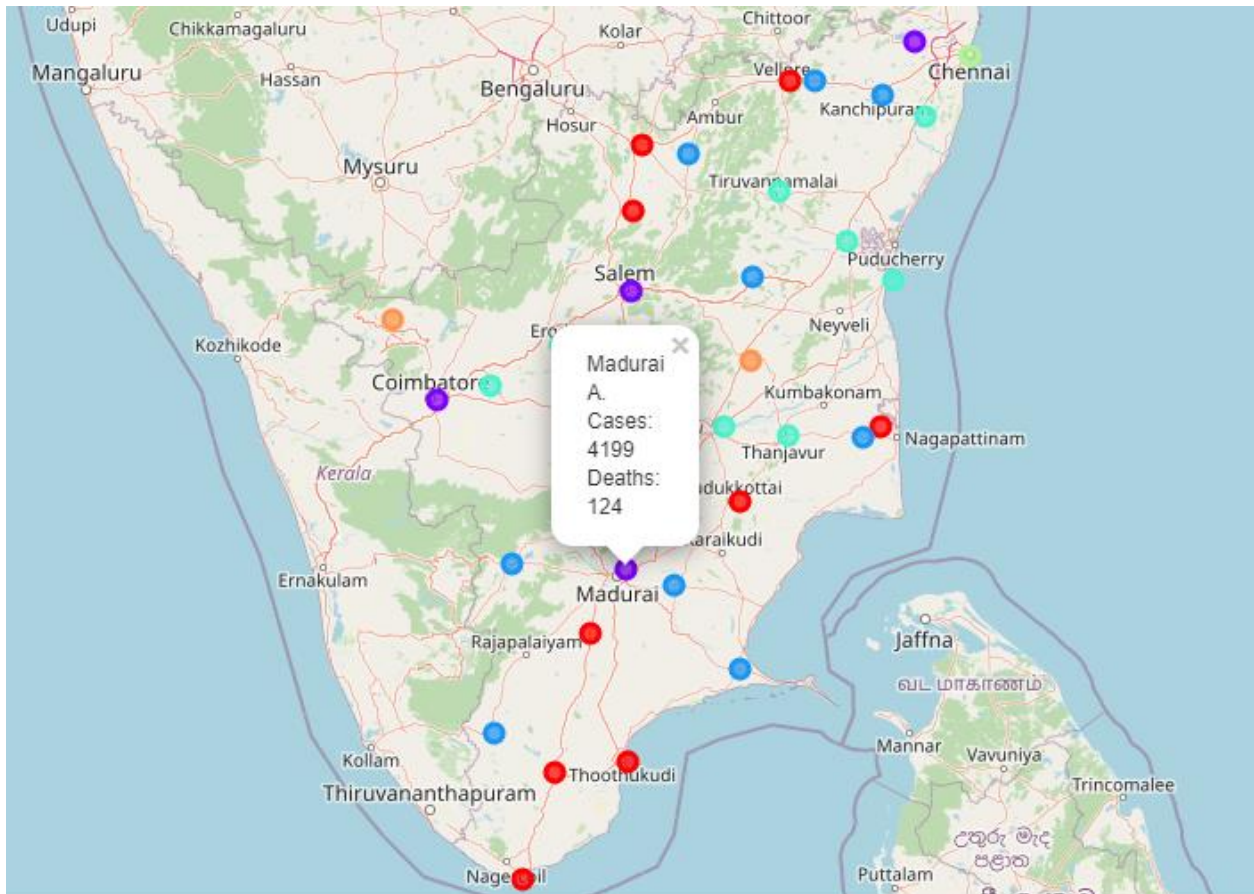
6. DISCUSSIONS

By looking at the cluster data, in finding the nearby hospitals we can see that the three clusters are grouped based on the most vital column *distance*. Cluster 0 is the one that we are the most interested in, due to the fact the group has the hospitals locations very near to the selected location.

The first cluster (Cluster label 0) has the hospitals situated less than 1 km from the chosen location *Alwarpet*.

Cluster 2 (Cluster label 2) is the cluster, where we see list of the list of hospitals situated (1 km approx).

Cluster 3 (Cluster label 1) contains hospitals situated more than 1.5 km from the given location.

The table shows that five hospitals (<1 km) which is closer to the given location, two hospitals are located (>1.5 km) and five hospitals are located (≈1 km) for the given radius (2000) in the foursquare API.

Cluster Labels	name	categories	lat	lng	Distance
0	ALL SMYLES COSMETIC DENTAL CENTRE CHENNAI INDIA	Dentist's Office	13.033604	80.250548	0.434329
0	kauvery hospital	Hospital	13.038214	80.257067	0.555733
0	SKS Veterinary Hospital	Hospital	13.028984	80.258728	0.706312
0	Madras ENT Research Foundation	Hospital	13.027266	80.254176	0.734405
0	Balaji Dental & Craniofacial Hospital	Dentist's Office	13.035352	80.247048	0.829344
2	Apollo Speciality Hospital	Hospital	13.033841	80.245289	1.003141
2	St. Isabel's Hospital	Hospital	13.039236	80.262297	1.030461
2	Oliva Skin and Hair Clinic Alwarpet	Hospital	13.042096	80.259281	1.049481
2	Sankara Nethralaya	Hospital	13.026432	80.261102	1.089126
2	Dr. Agarwal's Eye Hospital	Hospital	13.046304	80.255363	1.386506
1	Bharathiraja Hospital	Hospital	13.048036	80.245013	1.884593
1	Seethapathy Clinic and Hospital	Hospital	13.048841	80.264537	1.986330

By considering the severity of COVID-19 impact over TN districts extracted from web scrapping, the districts are classified as six groups based on severity. The optimal number of classes (k=6) is obtained through silhouette metric.

The first cluster (Cluster label 0) the biggest one has 10 districts and the second cluster has (Cluster label 1) has 4 districts grouped together based on the COVID-19 identified cases.

	Cluster Labels	Active	Location	Latitude	Longitude	DD_Lat	DD_Lng	District	Diagnosed cases	Deaths	Recovered cases	Active cases	Population	Cases per M	Last case reported on
17	0	1	Pudukkottai	10° 23' N	78° 52 ' E	10.383333	78.866667	Pudukkottai	729	9	414	306	1618345	450	14 July 2020
33	0	1	Vellore	12° 55' N	79° 11' E	12.916667	79.183333	Vellore	3098	10	1344	1744	1614242	1919	14 July 2020
27	0	1	Tirunelveli	8° 44' N	77° 44' E	8.733333	77.733333	Tirunelveli	1935	11	935	989	1665253	1162	14 July 2020
25	0	1	Thoothukudi	8° 48' N	78° 11' E	8.800000	78.183333	Thoothukudi	2497	17	1121	1359	1750176	1427	14 July 2020
14	0	1	Namakkal	11° 13 ' N	78° 13' E	11.216667	78.216667	Namakkal	203	1	96	106	1726601	118	14 July 2020
13	0	1	Nagapattinam[c]	10° 50' N	79° 44' E	10.833333	79.733333	Nagapattinam[c]	383	1	183	199	1616450	237	14 July 2020
11	0	1	Krishnagiri	12° 32' N	78° 16' E	12.533333	78.266667	Krishnagiri	274	7	165	102	1883731	145	14 July 2020
9	0	1	Kanyakumari	8° 5' N	77° 32' E	8.083333	77.533333	Kanyakumari	1613	10	553	1050	1870374	862	14 July 2020
35	0	1	Virudhunagar	9° 35' N	77° 57' E	9.583333	77.950000	Virudhunagar	2427	19	919	1489	1942288	1250	14 July 2020
4	0	1	Dharmapuri	12° 08' N	78° 13' E	12.133333	78.216667	Dharmapuri	260	1	103	156	1506843	173	14 July 2020
30	1	1	Tiruvallur	13° 09' N	79° 57' E	13.150000	79.950000	Tiruvallur	7292	133	4014	3145	3728104	1956	14 July 2020
12	1	1	Madurai	9° 58' N	78° 10' E	9.966667	78.166667	Madurai	6990	124	2667	4199	3038252	2301	14 July 2020
2	1	1	Coimbatore	11° 00' N	77° 00' E	11.000000	77.000000	Coimbatore	1480	10	338	1132	3458045	428	14 July 2020
20	1	1	Salem	11° 39' N	78° 12 ' E	11.650000	78.200000	Salem	2026	9	1024	993	3482056	582	14 July 2020

The third cluster (Cluster label 2) another biggest cluster has 10 districts and the fourth cluster has (Cluster label 3) has 9 districts grouped together based on the COVID-19 identified cases.

2	1	Kallakurichi	11° 44' N	78° 57' E	11.733333	78.950000	Kallakurichi	1904	6	1075	823	1370281	1389	14 July 2020
2	1	Tiruvarur	10° 46' N	79° 38' E	10.766667	79.633333	Tiruvarur	786	1	478	307	1264277	622	14 July 2020
2	1	Tirupattur	12° 29' N	78° 33' E	12.483333	78.550000	Tirupattur	461	1	272	188	1111812	415	14 July 2020
2	1	Theni	10° 0' N	77° 28' E	10.000000	77.466667	Theni	1916	20	684	1212	1245899	1538	14 July 2020
2	1	Kancheepuram	12° 50' N	79° 45' E	12.833333	79.750000	Kancheepuram	4091	53	1499	2539	1166401	3507	14 July 2020
2	1	Sivagangai	9° 52' N	78° 28' E	9.866667	78.466667	Sivagangai	1003	18	551	434	1339101	749	14 July 2020
2	1	Tenkasi	8° 58' N	77° 21' E	8.966667	77.350000	Tenkasi	824	3	325	496	1407627	585	14 July 2020
2	1	Ramanathapuram	9° 22' N	78° 52 ' E	9.366667	78.866667	Ramanathapuram	1956	38	1193	725	1353445	1445	14 July 2020
2	1	Karur	10° 58' N	78° 07' E	10.966667	78.116667	Karur	207	5	145	57	1064493	194	14 July 2020
2	1	Ranipet	12° 55' N	79° 20' E	12.916667	79.333333	Ranipet	1645	13	796	836	1210277	1359	14 July 2020
3	1	Villuppuram	11° 57' N	79° 32' E	11.950000	79.533333	Villuppuram	1723	20	1034	669	2093003	823	14 July 2020
3	1	Thanjavur	10° 47' N	79° 10' E	10.783333	79.166667	Thanjavur	738	11	422	1212	2405890	307	14 July 2020
3	1	Cuddalore	11° 43' N	79° 49' E	11.716667	79.816667	Cuddalore	1565	6	1154	405	2605914	601	14 July 2020
3	1	Tiruchirappalli	10° 50' N	78° 46' E	10.833333	78.766667	Tiruchirappalli	1715	27	964	724	2722290	630	14 July 2020
3	1	Dindigul	10° 22' N	78° 00' E	10.366667	78.000000	Dindigul	946	12	645	289	2159775	438	14 July 2020
3	1	Tiruppur	11° 05' N	77° 20' E	11.083333	77.333333	Tiruppur	319	4	185	130	2479052	129	14 July 2020
3	1	Erode	11° 20' N	77° 46' E	11.333333	77.766667	Erode	451	7	196	248	2251744	200	14 July 2020
3	1	Tiruvannamalai	12° 15' N	79° 07' E	12.250000	79.116667	Tiruvannamalai	3224	24	1848	1352	2494875	1292	14 July 2020
3	1	Chengalpattu	12° 42' N	80° 01' E	12.700000	80.016667	Chengalpattu	8553	169	5695	2689	2556244	3346	14 July 2020

The fifth cluster (Cluster label 4) being the smallest cluster, it has only 1 district (Capital of TN: Chennai). This district forms a distinct cluster, because the particular district is highly affected by COVID-19 virus because of its dense population and increasing cases day by day. The sixth cluster has (Cluster label 5) has 2 districts (Perambalur & Nilgiris) grouped together, these districts have the minimum number of COVID-19 identified cases.

Cluster Labels	Active	Location	Latitude	Longitude	DD_Lat	DD_Lng	District	Diagnosed cases	Deaths	Recovered cases	Active cases	Population	Cases per M	Last case reported on
5	1	Nilgiris	11° 29' N	76° 44' E	11.483333	76.733333	Nilgiris	258	1	107	150	735394	351	14 July 2020
5	1	Perambalur	11° 14' N	78° 56 ' E	11.233333	78.933333	Perambalur	178	1	161	16	565223	315	14 July 2020
4	1	Chennai	13° 04' N	80° 17' E	13.066667	80.283333	Chennai	79662	1295	62552	15815	4646732	17144	14 July 2020

7. CONCLUSION

Rural communities have extended struggled to maintain right to use to quality health care facilities. During the pandemic time, entire country is affected irrespective of rural and the urban areas because of rapid spreading of diseases. The urban class people gets immediate treatment as more hospitals are sited. A diversity of elements contribute to the problems in the rural areas such as shortages of physicians, casuality facilities, underinsured residents and so on. This project helps government and the health care organizations to identify areas where utmost care has to be taken and provide on-time medical facilities for the needy. The grouping of districts based on the severity of virus spreading helps health organizations in developing new approaches to support rural communities to deliver high-quality, cost-effective medical services on demand.