Identifying Medical Supply Zones in Delhi

Introduction:

As we all know that Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus. The outbreak was declared a Public Health Emergency of International Concern on 30 January 2020.

In this situation our aim is to provide suggestions for medical check-up booths and medicines supplies amid COVID-19 based on the areas, population, total number of cases in that area, hospitals in that area (to check the current medical facility status etc)

Target Audience: Government, medical agencies, NGOs and all related to COVID-19

Observed Problem: During the prime time of COVID-19, there was a requirement to identify the places where the medical and food supply is less, and cases are more. Due to panic situation, a lot number of small industries are shut down and due to which there is lack of medicine and food. So, I planned to identify the areas district wise in Delhi where proper supply of medicine and food is not there.

Description of data:

1. Call the Foursquare API to get the details of Hospitals in Delhi

	name	categories	address	lat	lng	labeledLatLngs	distance	cc	city	state	country	formatte
0	Dr. Ram Manohar Lohia Hospital	Hospital	Ram Manohar Lohia Hospital	28.624654	77.200609	[{'label': 'display', 'lat': 28.62465422213369...	1177	IN	New Delhi	Delhi	India	[Ram Ma Lohia Ho New Del
1	Sir Gangaram Hospital | सर गंगाराम अस्पताल	Hospital	Rajendra Nagar	28.638601	77.188884	[{'label': 'display', 'lat': 28.6386009566119,...	3016	IN	Delhi	Delhi	India	[Rajendr (Karol Ba Delhi, De
2	Sucheta Kriplani Hospital Cafeteria	Snack Place	NaN	28.634001	77.221702	[{'label': 'display', 'lat': 28.63400068219633...	2911	IN	NaN	NaN	India	[India]
3	army hospital research and referal	Hospital	near vasant vihar	28.586160	77.157681	[{'label': 'display', 'lat': 28.58616016121680...	5358	IN	Delhi	Delhi	India	[near va: (Dhaula Delhi, De
4	St. Stephen's Hospital	Hospital	Tis Hazari	28.666454	77.214489	[{'label': 'display', 'lat': 28.6664535219038,...	5940	IN	New Delhi	Delhi	India	[Tis Haza Delhi, De

2. Collect the information about the hospitals with proper address, Pin code, sub-districts etc

	Names	Sub Districts	Latidude	Longitude	State	Postal Code
Districts						
Central Delhi	Sir Gangaram Hospital	Rajender Nagar	28.638601	77.188884	Delhi	110060
New Delhi	Dr. Ram Manohar Lohia Hospital	Connaught Place	28.624654	77.200609	Delhi	110001
North Delhi	Tirath Ram Shah Hospital	Tis Hazari	28.670918	77.217385	Delhi	110054
North West Delhi	Karma Ayurveda Hospital	Pitam Pura	28.598992	77.189941	Delhi	110034
South Delhi	G.M. Modi Hospital	Saket	28.528471	77.215077	Delhi	110017
South West Delhi	army hospital research and referal	Dhaula Kuan	28.586160	77.157681	Delhi	110010
West Delhi	Orchid Hospital	Janakpuri	28.612374	77.086032	Delhi	110058

3. Collect the information about each district such as population, area, density etc.

	Code	District	Headquarters	Pincode	Area code	Latitude & Longitude	Population (2001)	Area (/km²)	Density (/km²)
0	CD	Central Delhi	Darya Ganj	110002	NaN	28° 51' N & 77° 25' E	644005	25	25759
1	ED	East Delhi	Preet Vihar	110003	NaN	28° 37' N & 77° 13' E	1448770	440	3293
2	ND	New Delhi	Connaught Place	110001	11.0	28° 36' N & 77° 12' E	302363	22	13744
3	NO	North Delhi	Sadar Bazar	110007	NaN	28° 63' N & 77° 22' E	779788	59	12996
4	NE	North East Delhi	Shahdara	110032	NaN	28° 4' N & 77° 2' E	1763712	52	33917

4. Lastly collect the information about the current COVID-19 cases in each district.

	SINo	State_Code	State	District_Key	District	Confirmed	Active	Recovered	Deceased	Delta_Confirmed	Delta
0	1	AN	Andaman and Nicobar Islands	AN_Nicobars	Nicobars	0	0	0	0	0	0
1	2	AN	Andaman and Nicobar Islands	AN_North and Middle Andaman	North and Middle Andaman	1	0	1	0	0	0
2	3	AN	Andaman and Nicobar Islands	AN_South Andaman	South Andaman	32	0	32	0	0	0
3	4	AP	Andhra Pradesh	AP_Anantapur	Anantapur	118	62	52	4	0	-3
4	5	AP	Andhra Pradesh	AP_Chittoor	Chittoor	151	74	77	0	9	6

Sources of Data:

1. Foursquare API
2. https://www.indianmirror.com/india-post/indianpincode.html to get area, density, population etc
3. https://api.covid19india.org/csv/latest/district_wise.csv to get latest COVID-19 cases
4. Wikipedia and other sites to get hospitals districts.

Assumption: Some data which I have collected is not accurate as it is not available online mainly the population data.

Exploratory data analysis:

Below are our target hospitals and areas of Delhi on which we are going to do the analysis:

Folium Map:
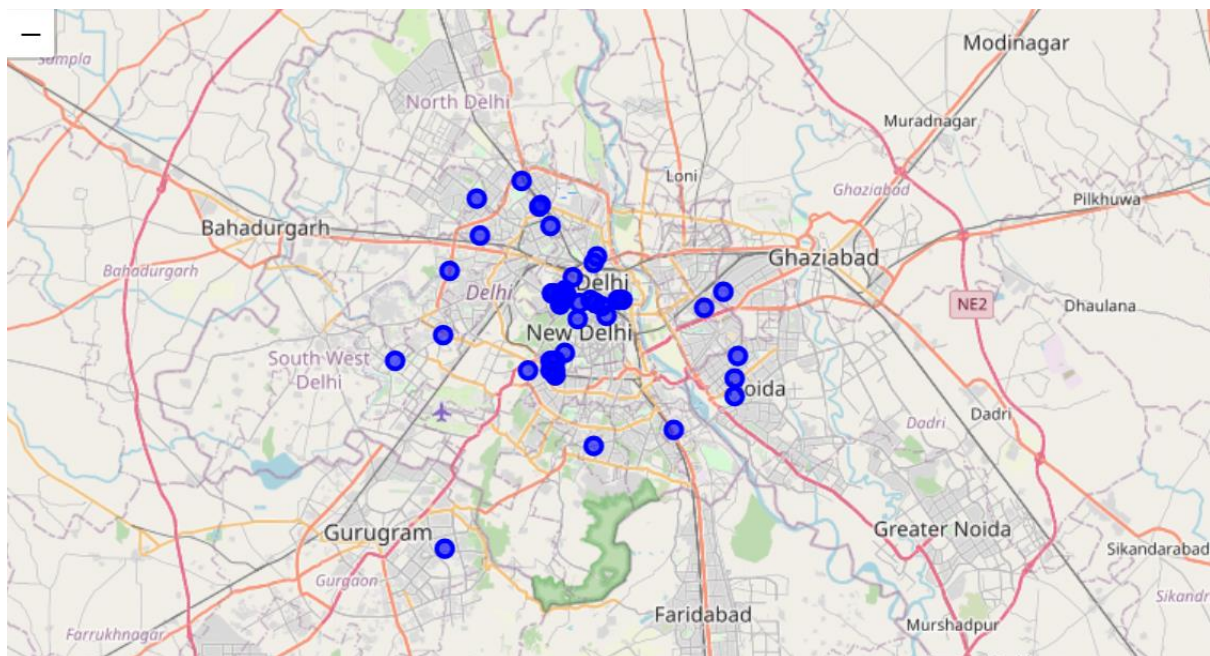
Below is the correlation matrix which shows the relation between the different features:

Few observations based on the below matrix:

1. Confirmed cases and Density have a high relation with value of 0.95

2. Area and Confirmed cases have a negative relationship. Which means inversely proportional to each other.
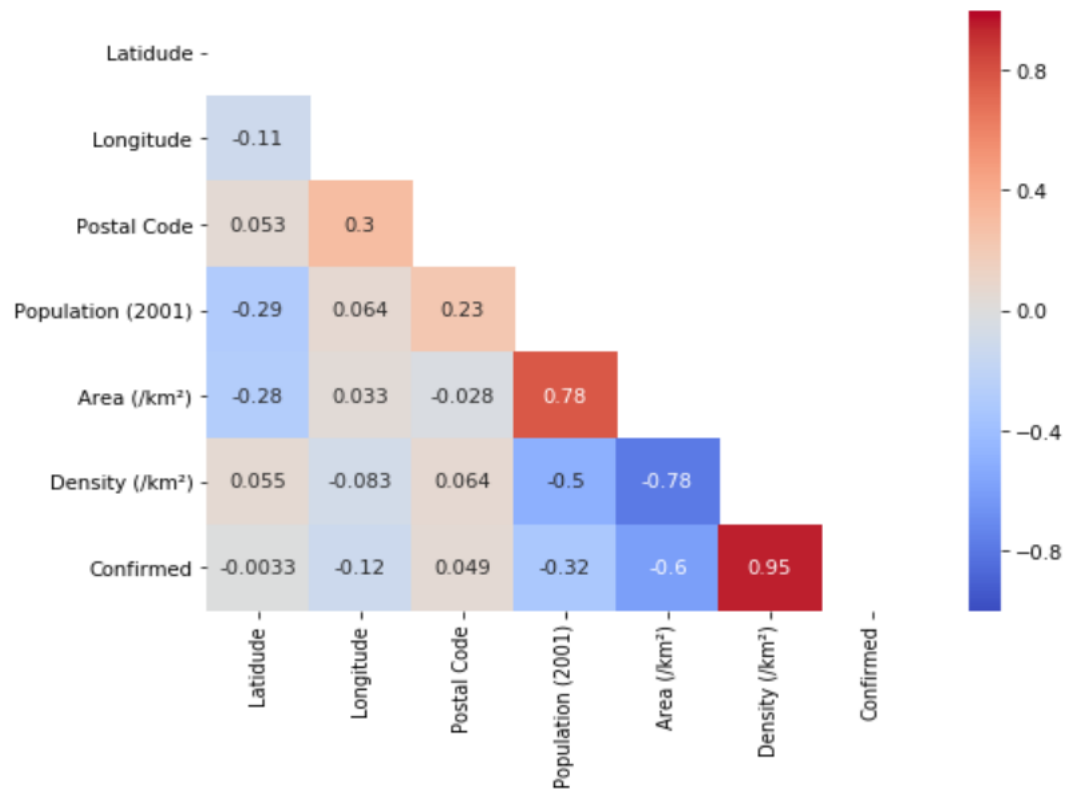
Now we will try to apply KMeans clustering algorithm, which will divide our regions into clusters based on similar parameters.

Lets first find the good value of k, which we can apply in our application.
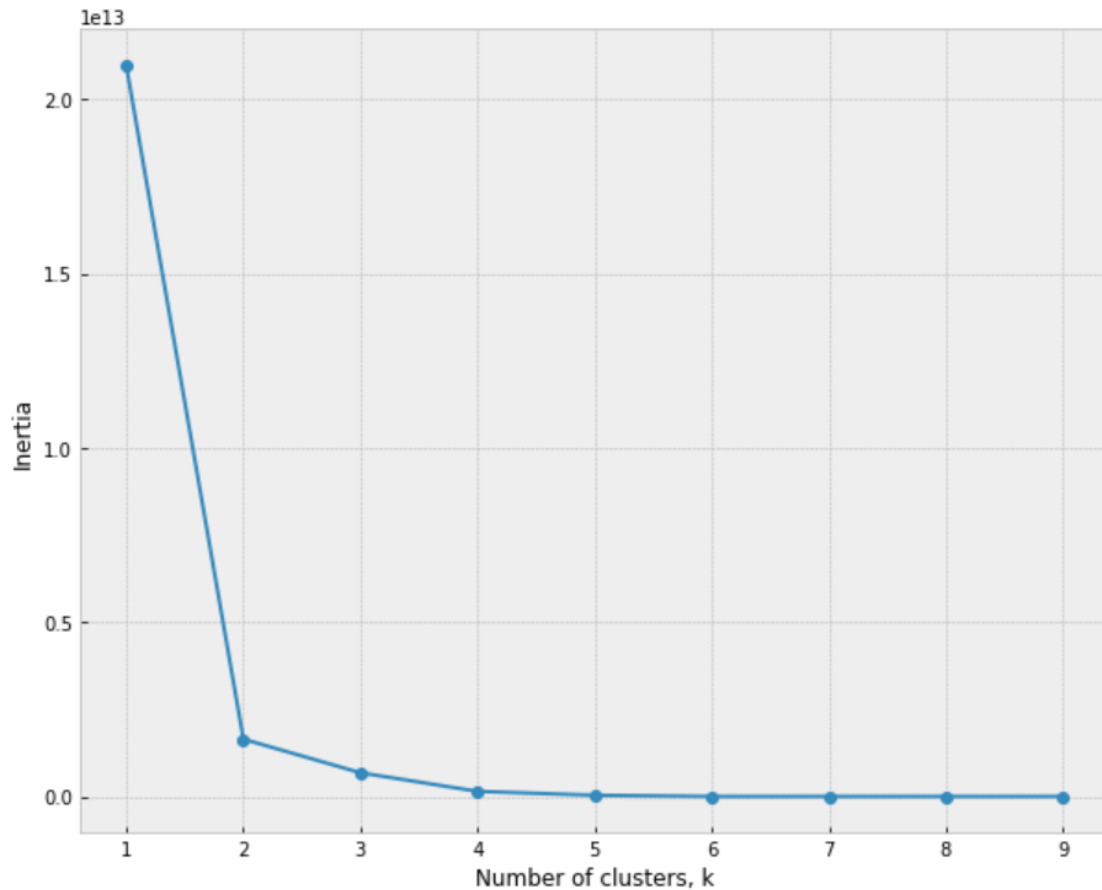
As COVID-19 is a sensitive matter, so here our decision is to take the value of K as 5, because after 5, inertia is not getting changed.

Now we will apply the KMeans clustering application using the below code:

```python
def k_means(n_clust, data_frame, true_labels):

    k_means = KMeans(n_clusters = n_clust, random_state=123, n_init=30)
    k_means.fit(data_frame)
    c_labels = k_means.labels_
    df = pd.DataFrame({'clust_label': c_labels, 'orig_label': true_labels.tolist()})
    ct = pd.crosstab(df['clust_label'], df['orig_label'])
    y_clust = k_means.predict(data_frame)
    display(ct)
    print('% 9s' % 'inertia  homo    compl   v-meas   ARI     AMI     silhouette')
    print('%i    %.3f  %.3f  %.3f  %.3f  %.3f   %.3f'
      %(k_means.inertia_,
      homogeneity_score(true_labels, y_clust),
      completeness_score(true_labels, y_clust),
      v_measure_score(true_labels, y_clust),
      adjusted_rand_score(true_labels, y_clust),
      adjusted_mutual_info_score(true_labels, y_clust),
      silhouette_score(data_frame, y_clust, metric='euclidean')))
```

```python
#privinding one of the optimal values of K
k_means(n_clust=5, data_frame=df_nonObj, true_labels=label)
```

Now we will discuss some observations:

First, we will discuss some relations between features:

Confirmed cases vs ratings of hospitals
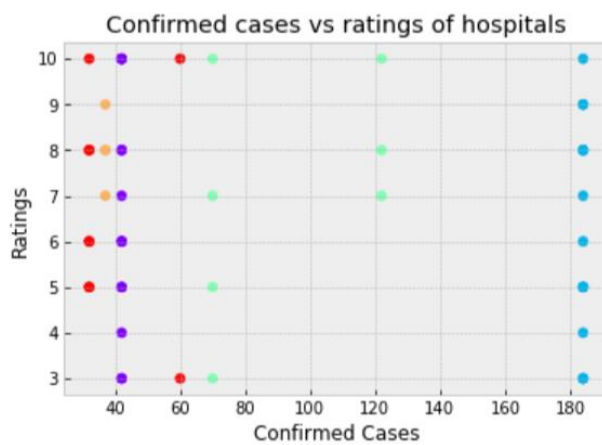
From the above graph we can say that there is no linear relation between the Ratings and Confirmed Cases


Confirmed cases vs Population
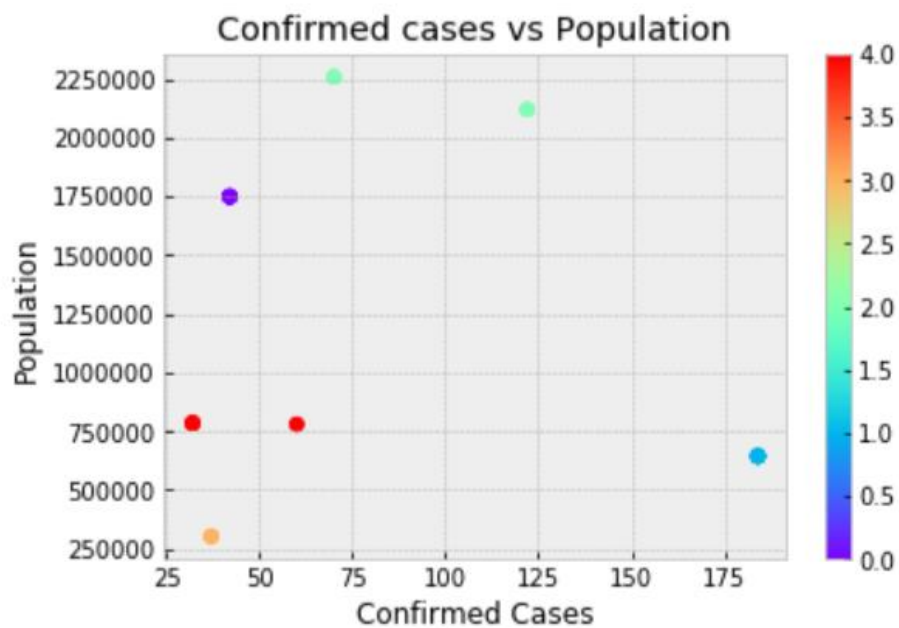
We have a linear relation between confirmed cases and population, which means up to some extent if population is increasing, then confirmed cases are also increasing.

Confirmed cases vs Density (/km²)
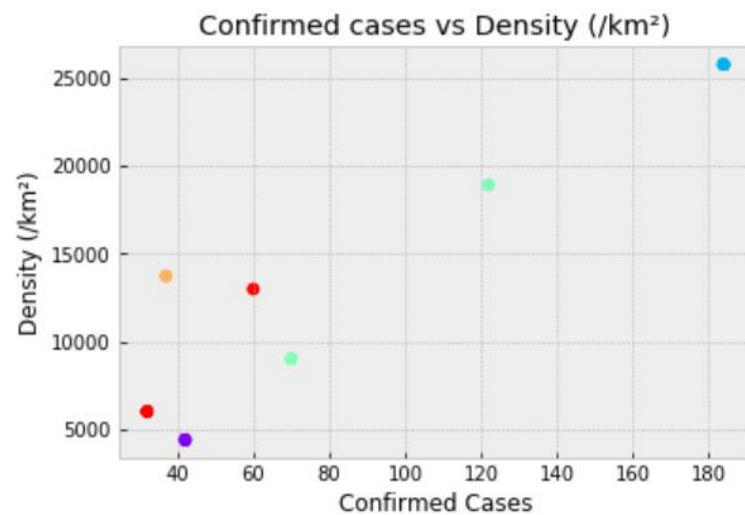
As we can see from the above graph that the density is linearly related to the Confirmed Cases.

Below we have our 5 clusters which we have mapped using Folium:
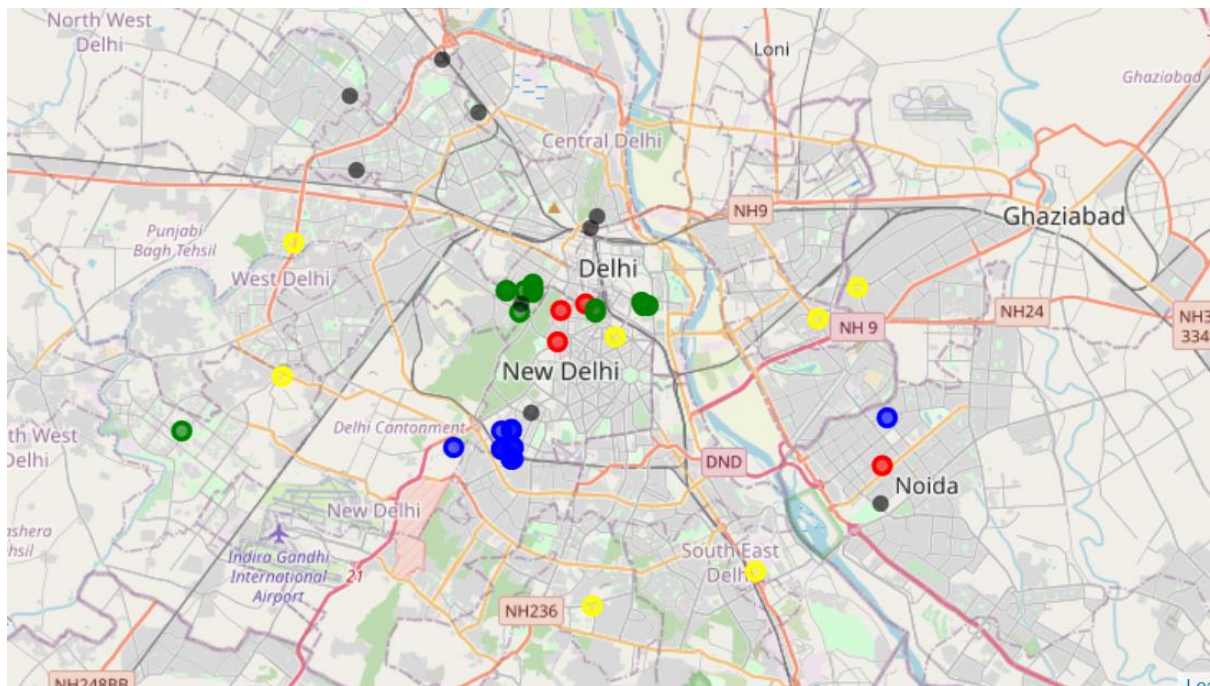
Cluster colours:

0 - Blue

1 - Green

2 - Yellow

3 - Red

4 – Violet

Below are the results which we have observed:

1. Blue cluster which are near New Delhi have low cases (at around 42 per district) and having low density of 4430 per square kms. Also, this cluster have higher number of hospitals.

2. Green cluster have highest cases of 184 and high density of around 25759 per square kms. This cluster have comparatively medium number of hospitals.

3. Yellow cluster have moderate cases of 70 and density of 9033 per square kms. This cluster have low number of hospitals.

4. Now red cluster have low cases of 37 per district and density of 13477 per square kms. This cluster have low number of hospitals.

5. Now Violet cluster have moderate cases of around 60 per district and density of 13000 square kms. Also having moderate number of hospitals.

Few recommendations:

1. Green cluster (mostly areas nearby Karol Bagh) needs to get more medical supplies in comparison to other clusters.

2. After green cluster, yellow cluster needs to be taken care because of low hospitals and high cases.

3. Then violet cluster (areas in north Delhi) needs supply of medicines and other amenities based on all the parameters.

Conclusion:

So, this concludes our observation. I hope these observations can help the required agencies. If we have more data about Delhi and COVID-19, then we can provide more accurate predictions.