



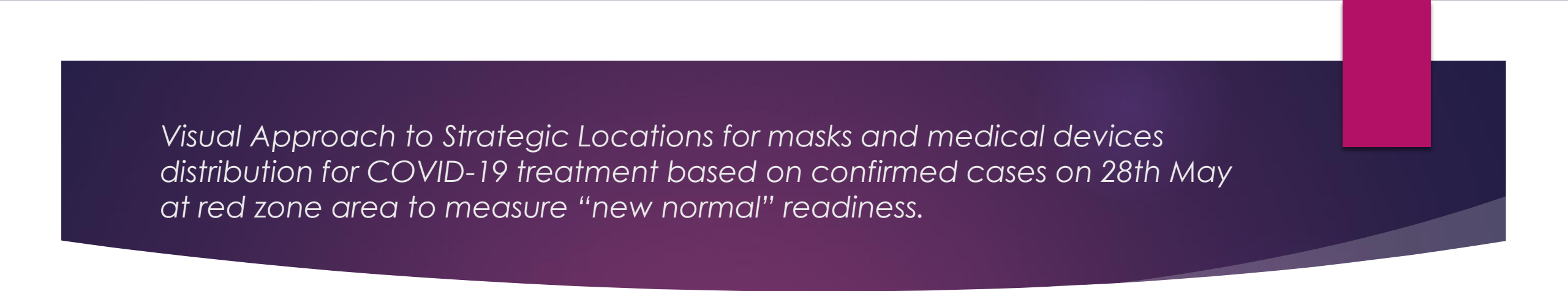
Capstone Project - The Battle of Neighborhoods Report

Applied Data Science Capstone by IBM/Coursera

Cahyati Supriyati Sangaji, 1 June 2020.

Table of contents

- ▶ Introduction: Business Problem
- ▶ Data
- ▶ Methodology
- ▶ Results and Discussion
- ▶ Conclusion



Visual Approach to Strategic Locations for masks and medical devices distribution for COVID-19 treatment based on confirmed cases on 28th May at red zone area to measure “new normal” readiness.

► Introduction: Business Problem

Since the beginning of 2020, Jakarta and many other cities around the world are currently under attack by an invisible army called ‘Novel Corona Virus’, also known as ‘Covid-19’. While every effort has been focusing on solving or minimizing problems, just like other professionals, Data Scientist takes a position to assess the current situations in the city, such as the availability, the number, and the geographical distribution (i.e. location) of health infrastructures, such as virus testing centers or authorized hospitals to treat affected patients. I would like to describe a simple analysis for determining a strategic location for the distribution of **masks and medical devices** for treatment COVID-19 based on confirmed cases on 28th May and the red zone area for “**new normal**” condition analysis.

Data

► Based on definition of our problem, factors that will influence our decision are:

1. The data Covid-19 cases per district are sourced from "[Riwayat File Covid-19 DKI Jakarta](#)"
2. The data total population in DKI Jakarta 2020 is sourced from statistik.jakarta.go.id
3. The data 10 most pupulation in DKI Jakarta 2020 per district is sourced from statistik.jakarta.go.id
4. The data Hospital for treatment covid-19 is sourced from megapolitan.kompas.com

► Following data sources will be needed to extract/generate the required information:

1. Process covid-19 positive case data collection on 28 May 2020 At 09.00.
2. The distribution of mask sales is based on the population in the DKI Jakarta area.
3. The distribution of mask sales is at most based on 5 districts with the most densely populated populations.
4. Creating new datasets from Hospital table that contains city, district, along with their latitudes and longitudes.

Methodology

- In the following METHODOLOGY section, we will describe the process of how to do a 'Visual' approach to better understand our data using data science and data analytics tool kits.

Output :

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 268 entries, 0 to 269
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype
---  -
 0   CITY        267 non-null   object
 1   DISTRICT    268 non-null   object
 2   POSITIVE    268 non-null   int64
dtypes: int64(1), object(2)
memory usage: 8.4+ KB
positive cases : 5061
```

```
CITY
JAKARTA BARAT      1122
JAKARTA PUSAT       922
JAKARTA SELATAN     888
JAKARTA TIMUR      1162
JAKARTA UTARA       953
KAB.ADM.KEP.SERIBU   14
Name: POSITIVE, dtype: int64
```

Methodology

- ▶ From all these process: data mining, preparation, exploration, the total number of Covid-19 confirmed positive cases in Jakarta is 5,061 per 28 May 2020, distributed across 6 main municipalities or cities in Jakarta, across 268 districts (or 'Kelurahan') out of just over 92.736 population of Jakarta. East Jakarta (Jakarta Timur) has the highest number of total POSITIVE cases with 1162 confirmed positives. Just like any other city, each city/municipality has many neighborhoods that can be used to pinpoint the location of the new proposed Covid-19 testing center along with further analysis of the neighborhood using FourSquare API and Folium map visualization technique.

Methodology

Need to get Latitude & Longitude of Jakarta city and the districts

	CITY	DISTRICT	POSITIVE	Latitude	Longitude
0	NaN	TOTAL	6929	-6.239089	106.818731
3	JAKARTA BARAT	PEGADUNGAN	51	-6.131373	106.701338
4	JAKARTA SELATAN	SENAYAN	17	-6.226911	106.809920
5	JAKARTA BARAT	KEBON JERUK	55	-6.192572	106.769726
6	JAKARTA UTARA	KELAPA GADING TIMUR	37	-6.166612	106.903732
...
265	KAB.ADM.KEP.SERIBU	PULAU TIDUNG	13	-5.803887	106.525306
266	JAKARTA BARAT	PINANGSIA	1	-6.135969	106.814264
267	JAKARTA SELATAN	SETIA BUDI	3	-6.207319	106.829340
268	JAKARTA BARAT	TAMBORA	2	-6.146614	106.801046
269	KAB.ADM.KEP.SERIBU	PULAU UNTUNG JAWA	0	-5.977613	106.707241

268 rows × 5 columns

Methodology

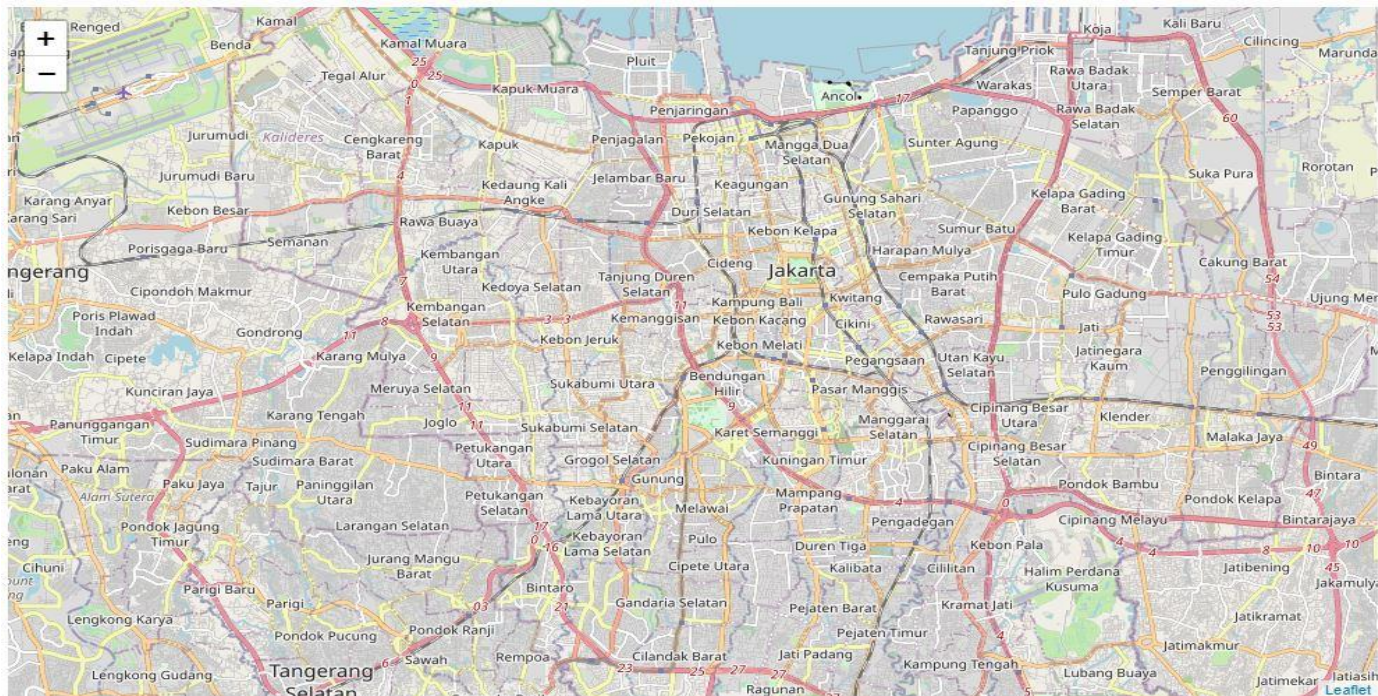
► Get the latitude and longitude Hospital

Besides, we also need to get the latitude and longitude of all Covid-19 testing centers in Jakarta that we have checked from the source www.kompas.com.

	Hospital	City	District	Latitude	Longitude
0	RSPI Sulianti Saroso	Jakarta Utara	Tanjung Priok	-6.129834	106.862590
1	RSUP Persahabatan	Jakarta Timur	Pulo Gadung	-6.204693	106.884848
2	RSPAD Gatot Soebroto	Jakarta Pusat	Senen	-5.434830	105.280931
3	RSUP Fatmawati	Jakarta Selatan	Cilandak	-6.311811	106.793485
4	RSU Bhayangkara Said Sukanto	Jakarta Timur	Kramat Jati	-6.269845	106.870943
5	RSAL Mintoahardjo	Jakarta Pusat	Tanah Abang	-6.210949	106.810810
6	RSUD Cengkareng	Jakarta Barat	Cengkareng	-6.142970	106.734888
7	RSUD Pasar Minggu	Jakarta Selatan	Pasar Minggu	-6.294225	106.820069
8	RSKD Duren Sawit	Jakarta Timur	Duren Sawit	-6.227008	106.937681
9	RS Peln	Jakarta Barat	Palmerah	-6.193061	106.803976
10	RSUD Tarakan	Jakarta Pusat	Cideng	-6.171638	106.810346
11	RSUD Koja	Jakarta Utara	Koja	-6.108922	106.900060
12	RSU Pertamina Jaya	Jakarta Pusat	Cempaka Putih Timur	-6.156662	106.886449

Methodology

- We then need to know how to get a map of the city that we are interested in (i.e. Jakarta) to present our data to the stakeholders using a ‘Visualization’ approach. We have downloaded all the required dependencies earlier in the report, and now we are ready to use the FOLIUM API service as described in the following section.



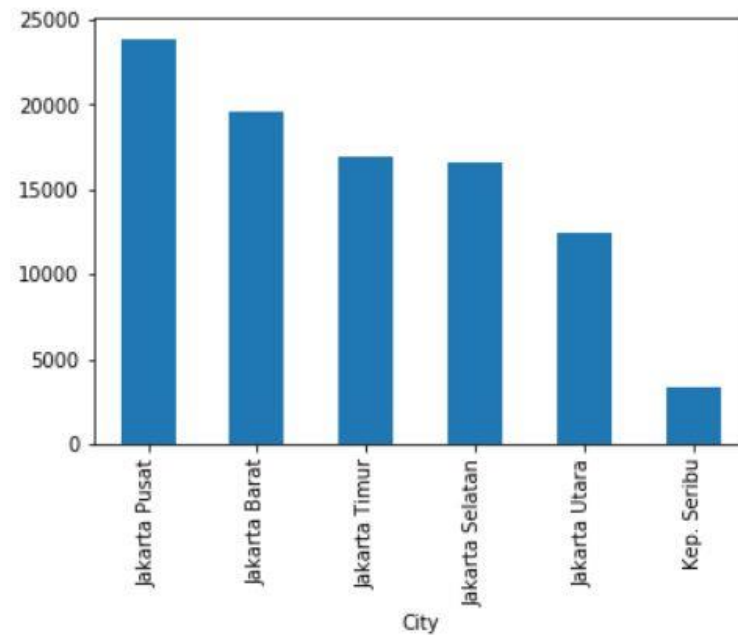
Methodology

- ▶ The map shows the main outer ring roads surrounding the city of Jakarta. It does NOT, however, show the official territorial boundary of the city concerning other administrative regions in the east, west, and south of Jakarta. However, because the author is from Indonesia, I know roughly which neighborhood belongs to Jakarta and which does not. In this scenario, we want to propose a strategic location (i.e. neighborhood) for the investing group within the Jakarta governmental area.

Results

The chart below is a chart of population density in Jakarta.

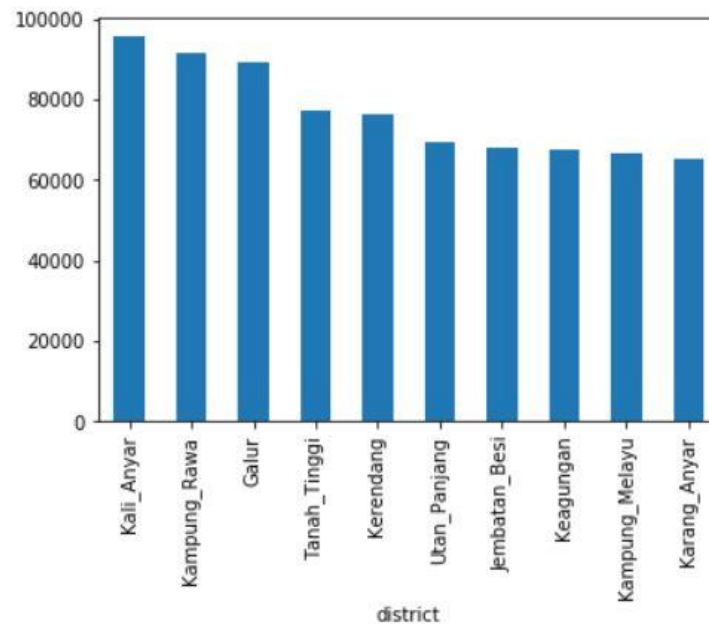
<matplotlib.axes._subplots.AxesSubplot at 0x23cdfa53688>



Results

The chart below is a chart of population density that is focused per district in Jakarta.

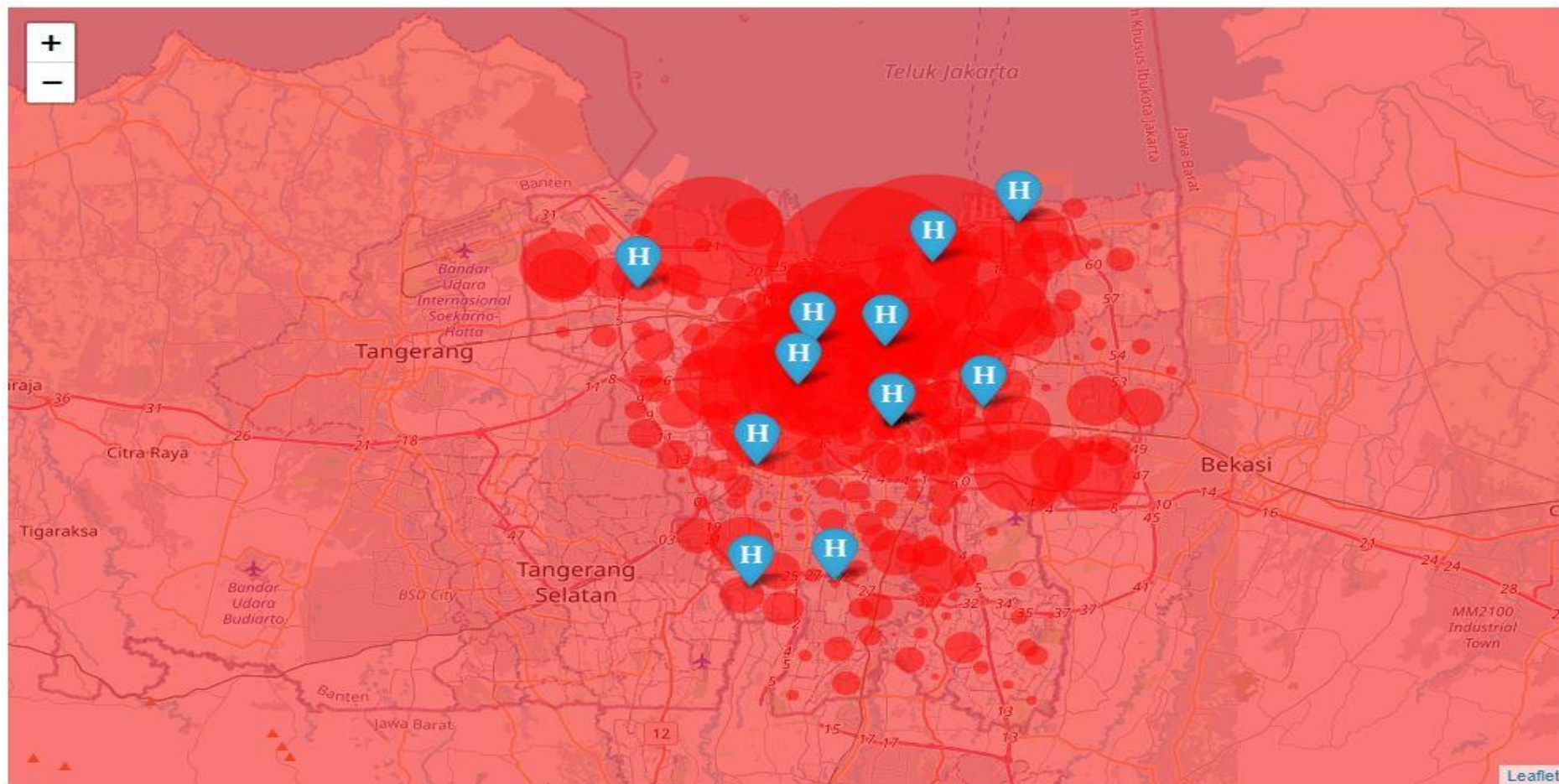
<matplotlib.axes._subplots.AxesSubplot at 0x23cdf9e7308>



Results

- Based on the graph results shown that areas need the distribution of masks the most is Central Jakarta (Jakarta Pusat) with the most populous population. Then 5 districts that most need a distribution of mask sales are Kali Anyar, Kampung Rawa, Galur, Tanah Tinggi, and Kerendang.

Results



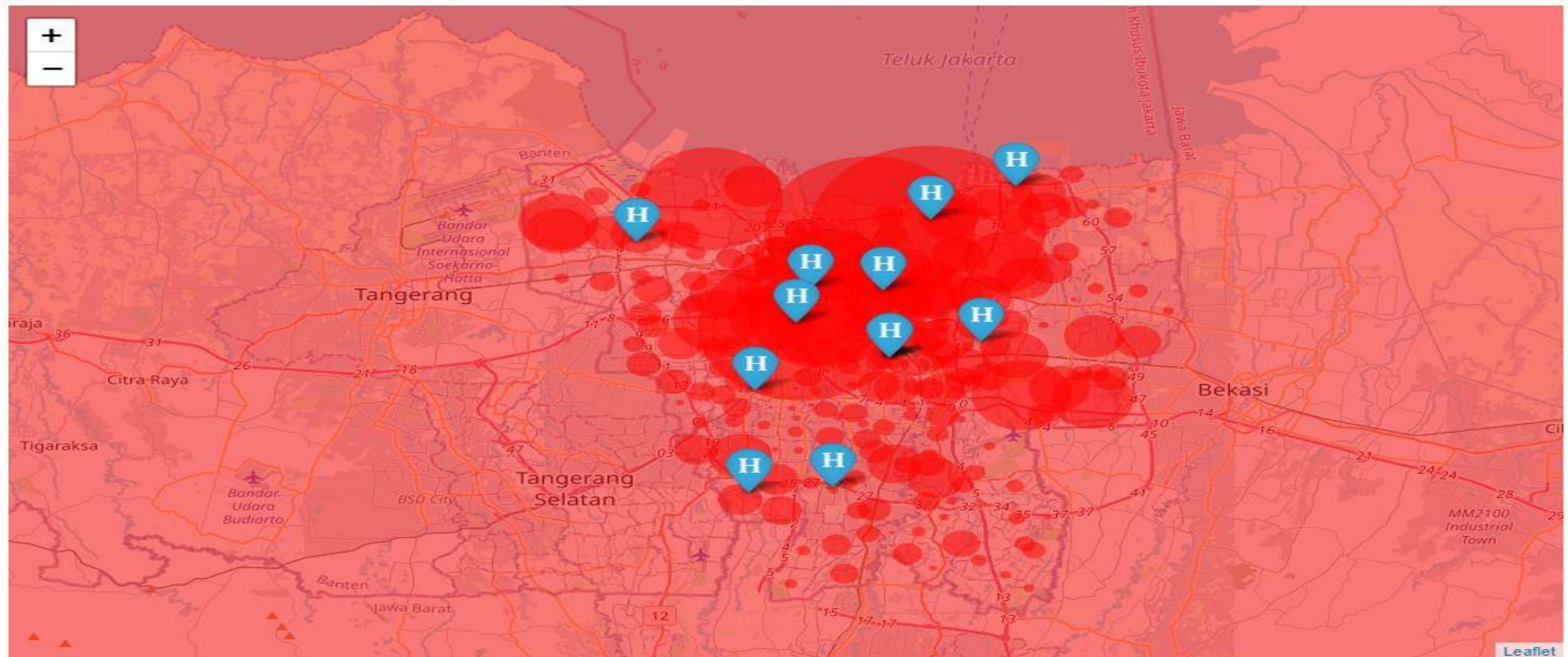
Results

- ▶ You can see from the results of the distribution of COVID-19 cases and the location of hospitals, almost all hospitals require a lot of medical equipment for COVID-19 treatment. In addition to Fatmawati hospital and the Pasar Minggu hospital, the distribution of the COVID-19 case is not as extensive as other hospitals.

Discussion

- ▶ We will try to analyze locations in the red zone based on the location of the hospital in the middle of the red zone. We determine based on the location of the Tarakan Hospital, Central Jakarta.
- ▶ Let's begin by trying to get the top 100 venues that are within Tarakan Hospital neighborhood and are within a radius of 500 meters of our candidate Covid-19 testing center using FOURSQUARE API.

Discussion



Discussion

- ▶ The analysis obtained is the location of the business which is in the Tarakan hospital neighborhood and is within a radius of 500 meters. then we also get the most congested cluster if businesses apply normal conditions in the red zone, potentially increasing cases of contracting the COVID-19 virus in the area.

Discussion

- ▶ We will try to analyze locations in the red zone based on the location of the hospital in the middle of the red zone. We determine based on the location of the Tarakan Hospital, Central Jakarta.
- ▶ Let's begin by trying to get the top 100 venues that are within Tarakan Hospital neighborhood and are within a radius of 500 meters of our candidate Covid-19 testing center using FOURSQUARE API.

Results and Discussion

- ▶ The project aims to provide information to local people who must be alerted to go out of the house from the distribution of the COVID-19 case in Jakarta. Provide information on which areas most need a lot of mask distribution according to population density in an area. Then provide information on which hospitals need the most medical equipment for COVID-19 treatment, maybe even additional medical personnel (doctors and nurses). Provide information on the business neighborhood which must implement health protocols with discipline when “new normal” conditions.

Conclusion

- ▶ This project helps mask sellers to distribute according to population density in Jakarta. Then it also helps the distribution of medical devices for corona care to hospitals that are estimated to have a large number of patients or even helps analyze which hospitals need additional medical personnel (doctors and nurses). Help business owners who are in one adjacent cluster to be more careful with the density of people in the business neighborhood.