**Coursera Capstone Project**

**IBM Applied Data Science Capstone**

**The Battle of Neighborhoods**

**Need of Temporary Isolation Wards in Mumbai, India**

**May, 2020**



**Objective**

The objective of this capstone project is to analyze and select the best locations in the city of Mumbai, India to open temporary isolation wards for covid patients. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question: In the city of Mumbai, India, if the government is looking to open temporary isolation wards, where would you recommend that they open it?

**Introduction**

As India enters the final phase of lockdown 2.0, the preparedness of its healthcare infrastructure to combat covid-19 will come into increasing focus. Five weeks of lockdown have managed to slow the rate of growth of infection, with cases doubling in roughly 10 days. But the battle against the virus is far from over. As the lockdown is relaxed in a phased manner, and with expatriate Indians set to return, there could be a surge in cases.

What is the readiness of healthcare infrastructure of individual states? If infections continue to rise at the same rate in May as they have done so far in April, India could be facing a deficit in isolation beds by the end of May, and in intensive care unit (ICU) beds and ventilators by the first week of June, a Mint analysis suggests.

The infrastructure stress will be especially acute in eight high-burden states, led by Gujarat, Maharashtra and Delhi. In April, cases in all these eight states grew at a compounded daily rate of above 10%—or, a doubling in about seven days or less. At the other end are states whose case growth is slower and which have greater reserves of covid-critical infrastructure relative to their case count.

Healthcare capacity is also a moving target. Till the night of 27 April, the policy directive was to mandatorily isolate everyone who tested corona-positive in a standalone facility. This facility could be an entire hospital or a separate block where other patients are not allowed. This necessitated the establishment of isolation wards.

However, considering the scenario that these facilities might be overwhelmed, the Union health ministry issued a directive on 27 April allowing “very mild" and pre-symptomatic patients to isolate at home, provided they fulfil eight-point eligibility criteria. The key requirement was space for separate quarantining of family members, a challenge for most Indian households.

For this analysis, we have considered that all corona-positive patients need an isolation bed. Further, we have considered state-wise counts of isolation beds from a 5 April health ministry document that outlined capacity of various covid-critical infrastructure.

As of 4 April, India had 152,403 isolation beds. Rajasthan had the most (20,835), followed by Tamil Nadu (19,860) and Kerala (13,028). Maharashtra, which had the maximum number of coronavirus cases (8,100) as of 26 April, had 11,861 isolation beds. The ministry document projected a compounded daily growth rate (CDGR) of 2% in isolation beds. We have taken a 4% CDGR to factor in intensification of efforts.

Finally, to account for further augmentation by the private sector, a multiplier of 0.5 of the shares of the private sector in hospital beds in the state is applied to the count of isolation beds available.

To explain, Maharashtra has 78% hospitalization in private facilities, according to a 2017-18 National Sample Survey Office (NSSO) survey. So, we assumed an increase in the number of isolation beds in the state (11,861) by 39%, or a total of 16,487 on 4 April (with the private sector providing the additional beds).

For ICU beds and ventilators, we have considered a 2% CDGR since these facilities are more expensive and time-consuming to provide. In line with projections from various global studies, we have assumed that 5% of cases require ICU care and 3% of cases would be on the ventilator. Lastly, we have projected cases in each state increasing at the CDGR recorded by it between 1 April and 26 April. This ranges from around 14% in West Bengal to 1.5% in Goa. At their respective rates, Maharashtra and Gujarat will be the worst-affected states, with over 700,000 and 530,000 cases, respectively, by 31 May.

This would put immense strain on the healthcare infrastructure, with both states running a deficit in isolation beds by 17 May. Going by ministry numbers, Delhi is already running a deficit in isolation beds. By 31 May, India will fall short of isolation beds at an aggregate level, with West Bengal, Madhya Pradesh, Jharkhand and Uttar Pradesh joining them.

In terms of ICU beds and ventilators, India will show a shortage at an aggregate level by 7 June, with Delhi, Gujarat and Maharashtra again the worst-affected.

Governments at the state and Centre must double up efforts to contain case growth and augment health infrastructure.

**Mumbai**

Mumbai is the capital city of the Indian state of Maharashtra. According to United Nations, as of 2018, Mumbai was the second most populous city in India after Delhi and the seventh most populous city in the world with a population of 19.98 million. As per Indian government population census of 2011, Mumbai was the most populous city in India with an estimated city proper population of 12.5 million living under Municipal Corporation of Greater Mumbai. Mumbai is the center of the Mumbai Metropolitan Region, the sixth most populous metropolitan area in the world with a population of over 23.64 million. In 2008, Mumbai was named an alpha world city. It has the highest number of millionaires and billionaires among all cities in India. Mumbai is home to three UNESCO World Heritage Sites: the Elephanta Caves, Chhatrapati Shivaji Maharaj Terminus, and the city's distinctive ensemble of Victorian and Art Deco buildings.

The seven islands that constitute Mumbai were originally home to communities of Marathi language speaking Koli people, who originated in Gujarat in prehistoric times. For centuries, the islands were under the control of successive indigenous empires before being ceded to the Portuguese Empire and subsequently to the East India Company when in 1661 Charles II of England married Catherine of Braganza and as part of her dowry Charles received the ports of Tangier and Seven Islands of Bombay. During the mid-18th century, Bombay was reshaped by the Hornby Vellard project, which undertook reclamation of the area between the seven islands from the sea. Along with construction of major roads and railways, the reclamation project, completed in 1845, transformed Bombay into a major seaport on the Arabian Sea. Bombay in the 19th century was characterized by economic and educational development. During the early 20th century it became a strong base for the Indian independence movement. Upon India's independence in 1947 the city was incorporated into Bombay State. In 1960, following the Samyukta Maharashtra Movement, a new state of Maharashtra was created with Bombay as the capital.

Mumbai is the financial, commercial and entertainment capital of India. It is also one of the world's top ten centers of commerce in terms of global financial flow, generating 6.16% of India's GDP and accounting for 25% of industrial output, 70% of maritime trade in India (Mumbai Port Trust and JNPT), and 70% of capital transactions to India's economy. Mumbai's billionaires had the highest average wealth of any city in the world in 2008. The city houses important financial institutions and the corporate headquarters of numerous Indian companies and multinational corporations.

**Healthcare Infrastructure in Mumbai**

Mumbai has a good public-health infrastructure. The health services for people in Mumbai are met through Hospitals and dispensaries run by the Municipal Corporation of Greater Mumbai (MCGM), Maharashtra state and the private sector. The MCGM has a network of four medical college-Hospitals, one dental college-Hospital, 16 municipal general Hospitals, six specialty Hospitals, 29 maternity homes, 175 municipal dispensaries and 183 health posts. In addition, the Maharashtra state government has one medical college Hospital, three general Hospitals and two health units located in Mumbai. The MCGM covers an area of 434 sq. km. The different health services are distributed MCGM runs several through Municipal Hospitals, clinics and dispensaries, as the chart below shows number of registered health care facilities in Mumbai.

With 18,555 cases and 696 deaths until May 17, Mumbai has the highest Covid-19 burden in the country.

Maharashtra government statistics show that about 30 per cent of Covid-19 cases who are admitted are symptomatic, including 3 per cent in need of critical care. While this is a small percentage, the city still faces an acute shortage of hospital beds for these critical patients.

**Data**

*To solve the problem, we will need the following data:*

• List of neighborhoods in Mumbai. This defines the scope of this project which is confined to the city of Mumbai, the city of the country of India in South Asia.

• Latitude and longitude coordinates of those neighborhoods. This is required in order to plot the map and also to get the location data.

• Location data, particularly data related to parks or recreational areas. We will use this data to perform clustering on the neighborhoods.

*Source of data and methods to extract:*

This Wikipedia page (https://en.wikipedia.org/Category:Suburbs\_in\_Mumbai) contains a list of neighborhoods in India, with around 50 neighborhoods. We will use web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and beautifulsoup packages. Then we will get the geographical coordinates of the neighborhoods using Python Geocoder package which will give us the latitude and longitude coordinates of the neighborhoods.

After that, we will use Foursquare API to get the venue data for those neighbourhoods. Foursquare has one of the largest databases of 105+ million places and is used by over 125,000 developers. Foursquare API will provide many categories of the venue data, we are particularly interested in the park or recreational areas category in order to help us to solve the problem put forward.

This is a project that will make use of many data science skills, from web scraping (Wikipedia), working with API (Foursquare), data cleaning, data wrangling, to machine learning (K-means clustering) and map visualization (Folium). In the next section, we will present the Methodology section where we will discuss the steps taken in this project, the data analysis that we did and the machine learning technique that was used.