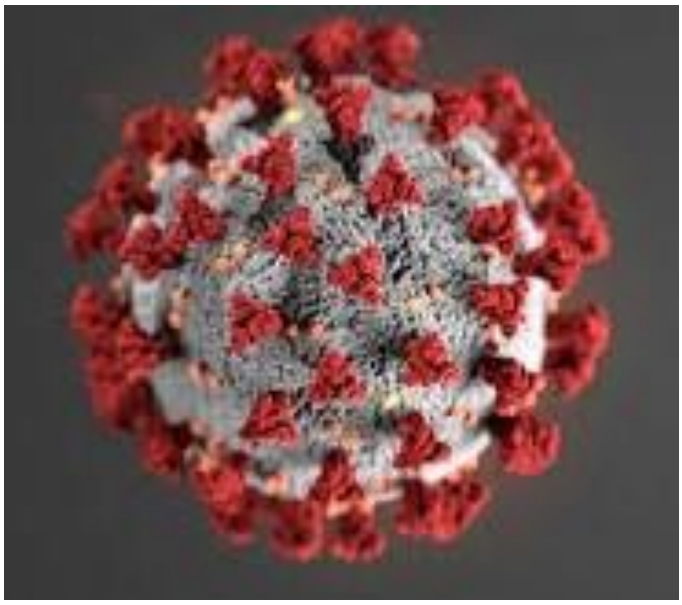
Coursera Capstone

IBM Applied Data Science Capstone

Understanding the impact of Pandemic & Recovery Status

Coronavirus Disease 2019 (COVID-19)



By:

Purnachand Kollapudi

May 2020

Contents:

Introduction:

The Chinese country office of the World Health Organization (WHO) on 31.12.2019 confirmed cases of unknown cause pneumonia found in Wuhan City, Hubei Province of China. The Chinese authorities have described a new form of coronavirus which was detected by laboratory experiments on 07.01.2020. This is a new strain that had not been previously found in humans until the epidemic in Wuhan, China was identified. Currently officially known as Coronavirus Disease 2019 (COVID-19), this "novel" coronavirus. It is from the virus family that causes illness ranging from common cold to more serious diseases such as Middle East Respiratory Syndrome (MERS-CoV) and Extreme Acute Respiratory Syndrome (SARS-CoV).

CoronaVirus:

Coronaviruses (CoV) derive their name from the fact that under electron microscopic examination, each virion is surrounded by the corona. Coronaviruses (CoV) are a large family of viruses that cause illness ranging from the common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS - CoV) and Severe Acute Respiratory Syndrome (SARS -CoV). So far, seven types of coronavirus are infecting people.

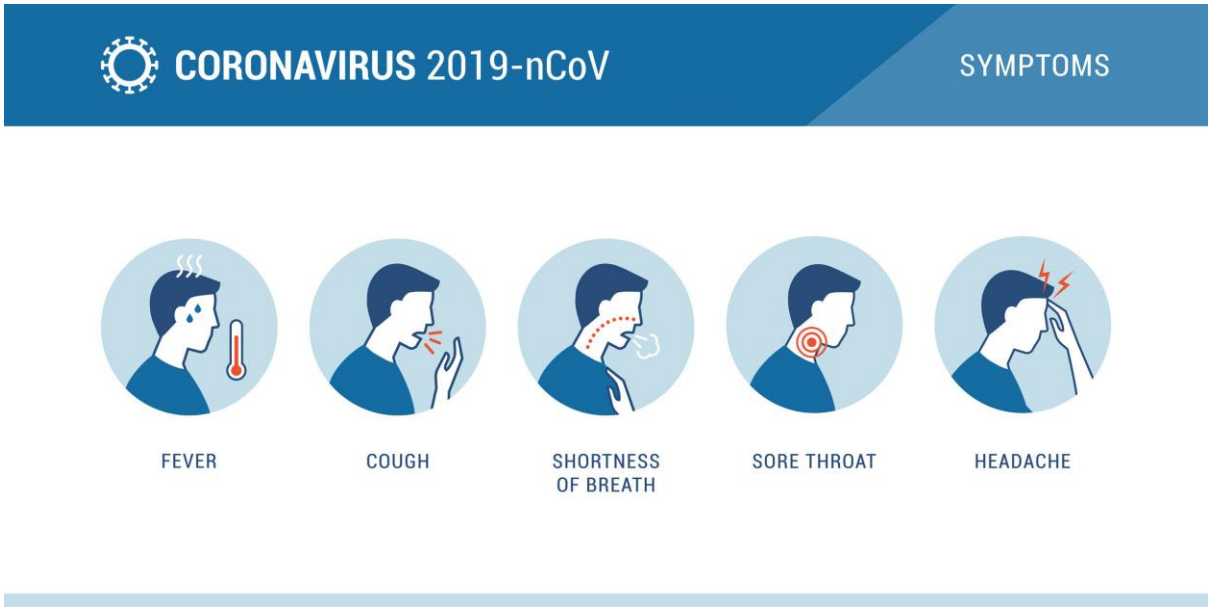


Fig.1.Corona Virus Symptoms

Direct Transmission: Person-to-Person

COVID-19 causes respiratory disease and is mainly transmitted in person-to-person. It can happen in the following circumstances:

- Between people who are in close contact with one another (within about 6 feet)
- Through respiratory droplets produced when an infected person coughs or sneezes
- These droplets can land in the mouths or noses of people who are nearby or possibly be inhaled into the lungs

Indirect Transmission: Other Causes

Contact with Infected Surfaces or Objects A person can possibly get COVID-19 by touching a surface or an object (e.g. doorknobs and table) that has the virus on it and then touching his own mouth, nose, or eyes.

Influential Survey:

	COVID-19	SARS	Influenza	Common Cough
Clinical Manifestations	Excessive fatigue; coughs; shortness of breaths; coughing up yellow or green mucus; chest X-ray shows scattered opacities in the lung	Coughs; breathing difficulties; fatigue; headache and diarrhea; fever	Running nose; sneezing; coughs; high temperature; muscle pain; diarrhea; vomiting	Nasal congestion; coughs; sore throat; throat discomfort; sneezing
Incubation Period	7-14 days	2-7 days	1-4 days	1 day
Ways of Transmission	Short distance droplets spread; close contact; contacts with animals	Short distance droplets spread; close contact	Coughs; sneezing and droplets spread; contact with secretions of an infected person	Droplets spread; contact with infected nasal secretions
Preventive Measures	Regular and frequent hand washing; check body temperature; use alcohol-based disinfectant; wear a surgical mask; enhance airflow; avoid contacts with animals or eat game meat	Cover mouth and nose when sneezing and coughing; regular and frequent hand washing; do not touch nose and mouth; wear a surgical mask; enhance airflow	Vaccination (flu shot); keep hands clean; wear a surgical mask; improve airflow	Regular hand wash, wear a surgical mask, boost your immune system

Statistics-as-on-date:

As per the statistics given by **Center for Systems Science and Engineering (CSSE)** is a research collective housed within the Department of Civil and Systems Engineering (CaSE) at **Johns Hopkins University (JHU)**. The team of CSSE works on a range of complex and interdisciplinary problems, united by the goal to better understand and improve societal, health, and technological systems for everyone**.**

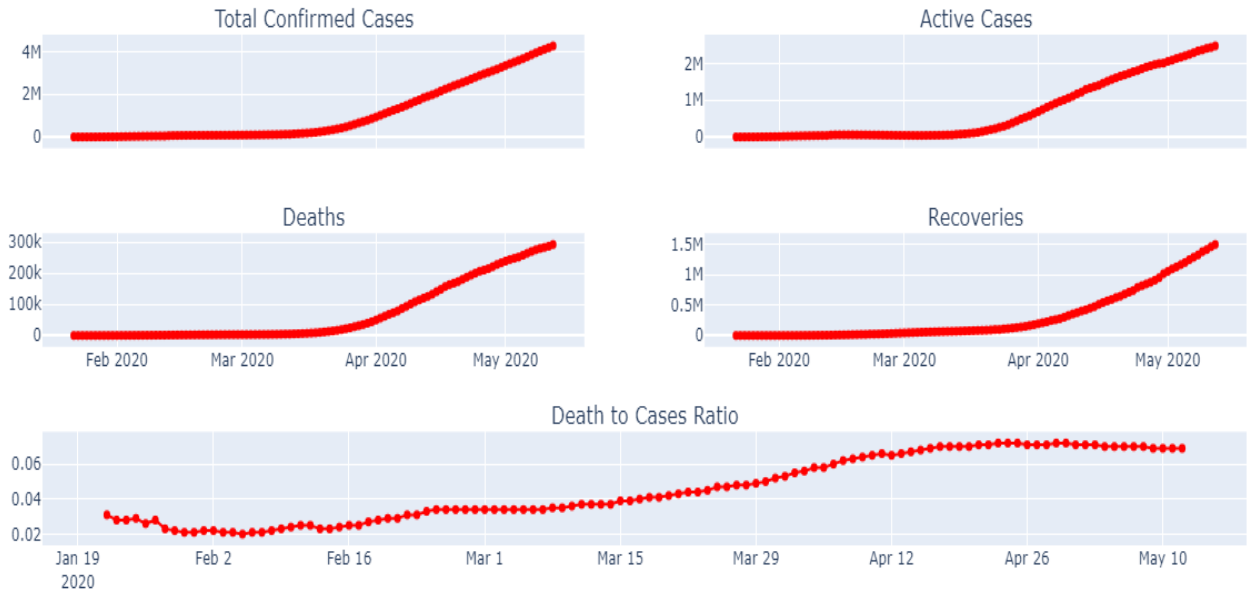


Fig.2. Statistics-As-On-Date by CSSE, JHU

Problem Statement:

With respect to the COVID-19 outbreak, the WHO Secretariat works with Taiwanese health experts and authorities, following established procedures, to facilitate a fast and effective response and ensure connection and information flow.

The innovators who are leveraging disruptive technologies to work on it and find unique and decisive solutions to improve the management of the pandemic and contain further outbreaks. The new ideas that emerge will help us and our countries to step back and observe the changes and figure out ways of taking advantage of a horizon of innovative opportunities that are emerging.

Research Objectives:

In this capstone project, the spread of the COVID-19 pandemic across large number of nations is in recent times is collected and analysed to identify the average recovery status of each country and visualized the same using geo maps to identify the clustered zones to predict the cause for recovery.

The focus areas for this project are as follows:

- Easy detection of infected persons in each country and recovery status
- Regular monitoring of the spread of the virus and predict outcomes
- Identifying the clustered zones to predict the reason for recovery
- Low cost and easy to implement

Dataset Description:

The dataset has been collected from an interactive web-based dashboard hosted by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University, to visualize and track reported cases in real-time. The dashboard, first shared publicly on January 22, illustrates the location and number of confirmed COVID-19 cases, deaths and recoveries for all affected countries. It was developed to provide researchers, public health authorities and the general public with a user-friendly tool to track the outbreak as it unfolds. Further, all the data collected and displayed is made freely available, initially as google sheets, now in a GitHub repository, along with the feature layers of the dashboard, which are now included in the ESRI Living Atlas.

Additional data sources are relied upon for reporting on regions outside China. These include U.S. county and state health departments, multiple national government health departments, as well as data aggregating websites including 1point3acres, Worldometers.info, BNO and the COVID Tracking Project (testing and hospitalizations), which rely on a combination of reporting from local health departments and local media reports. The full list of sources is maintained on our **CSSE COVID19 GitHub Repository**. All dashboard data curation and updates are coordinated by a team at JHU.

This capstone project has been collected "**time_series_covid19_recovered_global.csv**", it consists of **256 Countries** information represented as **rows** and **116 fields**(attributes) including **id, province/state, country/region, latitude and longitude** and also day wise recovery status from **22nd January,2020 to 12th May,2020** etc.

This project is aimed to address the following.

- Day-wise recovery status of each country and/or province or State.
- Latitude and longitude coordinates of those neighbourhoods. This is required in order to plot the map and also to get the venue data.
- Venue data, particularly data related to countries to find out the clustering on the neighbourhoods.
- Understanding the correlation between the attributes
- This analysis helps all others to predict the reasons for healthy recovery like lockdown, maintaining social distance, work from home etc.

References:

- https://www.mohfw.gov.in/pdf/DGSOrder04of2020.pdf
- https://www.mygov.in/hi/covid-19/
- https://www.codechef.com/COVDHACK?itm_campaign=contest_listing
- https://systems.jhu.edu/research/public-health/ncov/