

# SARS-CoV-2 Outbreak Analysis

-Ashish Agarwal

## Introduction:

An ongoing outbreak of pneumonia caused by a novel coronavirus, currently designated as the **Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2)**, was reported recently. However, as SARS-CoV-2 is an emerging virus, we know little about it. In late December 2019, a number of local health authorities reported clusters of patients with pneumonia of unknown cause, which were epidemiologically linked to a seafood market in Wuhan, Hubei Province, China. On 30 January 2020, the World Health Organization (WHO) declared that CoVID-19 is a **“Public-Health Emergency of International Concern”**. The pathogen was identified by local hospitals using a surveillance mechanism for “Pneumonia of Unknown Etiology” that was established in the wake of the 2003 SARS outbreak with the aim of allowing timely identification of novel pathogens.

On 23 Feb 2020, the lock-down of Wuhan, a central city in China, has alarmed people all over the world of an emerging novel coronavirus that is posing a major public health and governance challenges and has now spread to 6 continents including 66 countries and still counting.

Despite the current mortality rate is less as of now, the emergence of large number of infected patients within short period of time could result in the collapse of health care system, and thus the mortality rate might be elevated. Every virus has **Basic Reproduction number (R0)** which implies how many people will get the disease from the infected person. As per initial research work R0 of SARS-CoV-2 is 2.7.

Every Pandemic has four stages:

Stage 1: Confirmed Cases come from other countries

Stage 2: Local Transmission Begins

Stage 3: Communities impacted with local transmission

Stage 4: Significant Transmission with no end in sight

Countries like Italy, USA, UK and France are currently in Stage 4, while India is in on the edge of Stage 3. Currently the goal of all scientists around the world is to "Flatten the Curve". SARS-CoV-2 currently has exponential growth rate around the world and flattening this growth typically implies even if the number of Confirmed Cases is increasing but the distribution of those cases should be over longer timestamp. To tackle the widespread, various countries are

imposing Travel Ban, Cross-Border shutdown, ban on immigrants are Testing, Contact Tracing and Quarantine.

## Objective:

The main objective of this project are:

- To study SARS-CoV-2 outbreak with the help of some basic visualizations' techniques. Comparison of China where it has originally originated from with the rest of the world.
- To Perform predictions and Time Series forecasting in order to study the impact and spread of the SARS-CoV-2 in coming days.
- To study some important factors that impact transmission rate of SARS-CoV-2 like Air Transmission, Relative humidity., Temperature.
- To study the responses of countries to SARS-CoV-2 i.e. to classify the changes in the number of infections in different countries, hoping to find out the factors that affect the country's ability to respond to it.
- To have a sensible study to classify the trend of the number of infected people under the influence of different factors, and then modify the policy according to the trend of the expected number of infected people.
- To analyze Growth Factor & Inflection.
- To have an Exploratory Data analysis, Forecasting and implementing the Machine Learning Models wherever needed.

## Value to Client(s):

Beneficial to concerned client(s) desiring to have a detailed analysis of the present outbreak to have essence of present situation in order to discover and analyze the trend, behavior, responses to various factors affecting the market and invent , strategize precautionary measures for any future pandemic like this.

## Data Sources\*:

- <https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset>
- <https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data>
- <https://www.kaggle.com/imdevskp/corona-virus-report>
- <https://www.kaggle.com/parulpandey/coronavirus-cases-in-india>
- <https://worldpopulationreview.com/countries/most-visited-countries/>

## Methodology:

- Importing required Python Packages and Libraries
- Data Preprocessing
- Feature Engineering
- Explanatory Data Analysis for having a general exploration of data and understanding the relationships between various attributes.
- Implementation of desired model for training and testing the dataset in order to forecast the future scenario.

## Expected Deliverable(s)\*:

- Date wise, Country wise statistics Analysis of the virus with time.
- Growth factor distribution of different types of cases worldwide
- Comparing infections in these countries with more than 500 infections.
- Statistical correlations with important factors like temperature, relative humidity and air travel, are performed to predict the transmission rates.
- Mortality and Recovery Rate analysis around the World
- Comparison of ICU beds within various Countries.
- World Bank Indicator Analysis regarding the comparison of the Patient's age with per capita GDP, average life span and number of ICU beds available.
- Linear Regression Model for Confirm Cases Prediction
- Time Series Forecasting

\*: Subject to change with advancement of the project