# Capstone Project Proposal 01

### ***Covid-19 Global Forecasting and Analysis***

#### by: Reeti Bhagat

A close up of a flower

Description automatically generated

*Image Credits :*[*Scientific Animations*](http://www.scientificanimations.com/wiki-images/)*under a*[*CC BY-SA 4.0*](https://creativecommons.org/licenses/by-sa/4.0/)*license*

## **Background**

**Coronavirus** is a family of viruses that can cause illness, which can vary from *common cold* and *cough* to sometimes more severe disease. **Middle East Respiratory Syndrome (MERS-CoV)** and **Severe Acute Respiratory Syndrome (SARS-CoV)** were such severe cases with the world already has faced.  
**SARS-CoV-2 (n-coronavirus)** is the new virus of the coronavirus family, which first *discovered* in 2019, which has not been identified in humans before. It is a *contiguous* virus which started from **Wuhan** in **December 2019**. Which later declared as **Pandemic** by **WHO** due to high rate spreads throughout the world. Currently (on the date 10 June 2020), this leads to a total of *411K+ Deaths* across the globe, including *180K+ deaths* alone in *Europe*.

### **The following two curves shows why we need to flatten the curve and follow the protective measures:A close up of a device Description automatically generated**

A screenshot of a cell phone

Description automatically generated

# Goal:

Pandemic is spreading all over the world; it becomes more important to understand about this spread. It is an effort to analyze the cumulative data of confirmed, deaths, and recovered cases over time. In this study, the main focus is to analyze the spread trend of this virus all over the world.

1.To determine Cumulative Confirmed cases and Cumulative Recovery Vs Cumulative Death Analysis.

2. To perform Covid-19 Daily analysis

3.To perform Covid-19 data analysis

4.Covid-19 Spread Analysis in different countries

5.Comparision of mortality and death rates (Variation of Deaths and Confirmed cases of different countries over time)

6. Global prediction of Covid-19

7. Covid-19: USA, spread analysis

`

# Data Source:

I will be using data from different sources as[John Hopkins CSSE](https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series),

Data. World and world meter.

<https://github.com/CSSEGISandData/COVID-19>  
2019 Novel Coronavirus COVID-19 (2019-nCoV) Data Repository by Johns Hopkins CSSE  
This dataset is updated on daily basis by Johns Hopkins CSSE

* [WHO](https://www.who.int/emergencies/diseases/novel-coronavirus-2019)
* [COVID-19 Tracker by Johns Hopkins University](https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6)
* <https://www.worldometers.info/population/>
* <https://data.world/liz-friedman/us-covid-19-data-from-nytimes/workspace/file?filename=us-counties.csv>
* [https://query.data.world/s/n7hugkdu54wcxk4tbk5v5hytwhxrck'`](https://query.data.world/s/n7hugkdu54wcxk4tbk5v5hytwhxrck'%60)
* [https://services7.arcgis.com/LXCny1HyhQCUSueu/arcgis/rest/services/Definitive\_Healthcare\_USA\_Hospital\_Beds/ FeatureServer/0/query?where=1%3D1&outFields=\*&outSR=4326&f=json](https://services7.arcgis.com/LXCny1HyhQCUSueu/arcgis/rest/services/Definitive_Healthcare_USA_Hospital_Beds/%20FeatureServer/0/query?where=1%3D1&outFields=*&outSR=4326&f=json)
* <https://query.data.world/s/a77r7urucnw5vluyh6ubt477pkbyyp>
* <https://data.world/usafacts/coronavirus-in-the-us>

## **Approach**

1. Understand the variables in the data
2. Clean data to exclude redundant entries
3. Explore and construct models
4. Visualization and forecasting using best model

## **Deliverables**

* Report/Paper
* Jupyter Notebooks with intermediate data analysis

# 