Dataset: https://www.kaggle.com/sudalairajkumar/covid19-in-india

1. Describe the Dataset

Coronaviruses are a large family of viruses which may cause illness in animals or humans. In humans, several coronaviruses are known to cause respiratory infections ranging from the common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). The most recently discovered coronavirus causes coronavirus disease COVID-19 - World Health Organization

The number of new cases are increasing day by day around the world. This dataset has information from the states and union territories of India at daily level.

India has the second-highest number of confirmed cases in the world (after the United States) with 29.3 million reported cases of COVID-19 infection and the third-highest number of COVID-19 deaths (after the United States and Brazil) at 367,081 deaths.A second wave beginning in March 2021 was much larger than the first, with shortages of vaccines, hospital beds, oxygen cylinders and other medicines in parts of the country.

2. Determine The Tasks of Dataset (excludes the mentioned ones too)

- Predict how well can the 21 day lockdown perform in containing the spread of the virus.
- Predicts in which states will increase cases in the near future.

3. What kinds of data provided by the dataset?

Structured Data : CSV

- Serial Number : Int
- Date of Observation : Date
- Time of Observation : Time
- State / Union Territory : String

4. Which Analytical Approach will be used?

- Descriptive

5. What are the use of the result of an analytical approach?

- Determine predictions of spikes in covid 19 cases in several regions of India.
- Estimate the supply of medical devices, isolation locations and the costs that must be allocated by the state to hospitals.

6. For whom the result will be beneficial?

- World Health Organizations
- Covid Survey Institute
- Hospital
- Government
- Ministry of Health

7. Can it be possible to improve the result with another approaches?

Yes, it can be use Predictive too.

8. Choose one notebook and analyse it in the framework of data science methodology.

- Notebook: https://www.kaggle.com/ashishkumarbehera/covid-19-analysis?select=StatewiseTestingDetails.csv
- This NoteBook is an effort to analyze the cumulative data of confirmed, deaths, and recovered cases over time. In this notebook, the main focus is to analyze the spread of this virus all over the india.
- Methodology:
    a. Problem Identification (Problem statement formation & Data source)
    b. Data collection
    c. Data Organization
    d. Build data profil tables and plots
    e. Explore data relationships.