# **Project – Analysing the trends of COVID-19 with Python**

**Problem Statement:** Given data about COVID 19 patients, write code to visualise the impact and analyse the trend of rate of infection and recovery as well as make predictions about the number of cases expected a week in future based on the current trends

# **COVID-19 Data Analysis and Prediction**

## **Introduction**

In this document, we’ll explore how to analyze COVID-19 data, visualize trends, and make predictions using Python. We’ll use the following libraries:

1. **pandas**: For data manipulation and aggregation.
2. **Plotly**: For interactive visualizations.
3. **Facebook Prophet**: For time series modeling and forecasting.

## **Steps**

### **1. Data Collection**

* Gather COVID-19 data from reliable sources (e.g., government health departments, WHO, or Kaggle datasets).
* Ensure that the data includes relevant information such as date, confirmed cases, recovered cases, and deaths.

### **2. Data Preprocessing**

* Load the data into a pandas DataFrame.
* Clean and transform the data (handle missing values, convert date strings to datetime objects, etc.).

### **3. Exploratory Data Analysis (EDA)**

* Calculate daily new cases, recovery rates, and mortality rates.
* Visualize the overall trend using line charts.
* Identify any anomalies or sudden spikes.

### **4. Interactive Visualizations with Plotly**

* Create interactive plots to visualize:
  + Total confirmed cases over time.
  + Daily new cases.
  + Recovery rates.
  + Mortality rates.
  + Geospatial distribution of cases (if available).

### **5. Time Series Modeling with Facebook Prophet**

* Install the prophet library (if not already installed).
* Prepare the data for Prophet:
  + Rename columns to “ds” (date) and “y” (total confirmed cases).
  + Optionally, include holidays or special events as additional regressors.
* Fit the Prophet model to the data.
* Make predictions for the next week.

### **6. Visualisation of Predictions**

* Plot the actual data along with the predicted values.
* Highlight the forecasted trend for the next week.