**Objective:** Apply data science techniques to analyze a medical dataset combining time series and non-time series data, focusing on data sourcing, enrichment, and handling missing values.

Dataset: <https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_global.csv>

**Description**: Contains cumulative daily COVID-19 death counts by country/region, with static metadata (Province/State, Country/Region, Latitude, Longitude) and time series data

**Tasks**:

* Download the dataset from the provided URL.
* Load into a pandas DataFrame and explore the database using different techniques discussed in class.
* Save the raw data as raw\_covid\_deaths.csv for reproducibility.
* **Identify Data Types**: Separate time series (date columns) and non-time series (Province/State, Country/Region, Lat, Long) columns.
* **Transform to Long Format**: Use pd.melt() to create a long-format DataFrame with columns: Province/State, Country/Region, Lat, Long, Date, Deaths.
* **Handle Missing Values**:
  + For non-time series (e.g., Province/State): Fill with 'N/A'.
  + For time series (e.g., Deaths): Use forward-fill (ffill) to maintain cumulative trends.
* **Convert Dates**: Convert the 'Date' column to datetime format (pd.to\_datetime).
* Apply any model for time series (SARIMA, etc.) as well as one model (Logistic regression) for non-time series data (Lat, Long).
* Evaluate the model.