

# Intermediate Pandas Tasks for COVID-19 Dataset

## 1. Data Inspection & Preparation

- Load the dataset and display the first 5 rows.
- Show dataset shape and column info.
- Convert 'Date' column to datetime type.
- Fill missing values in numerical columns with 0.

## 2. Filtering & Selection

- Select data for a specific country (e.g., 'Pakistan') for the month of March 2020.
- Get all rows where Deaths is greater than 1000.
- Select only Country/Region and Confirmed columns for the first 10 rows.

## 3. Aggregations

- Group by Country/Region and get the total confirmed cases per country.
- Group by Country/Region and calculate average daily new cases per country.
- Find the top 5 countries with the highest average daily new cases.

## 4. Column Operations

- Create a new column Mortality Rate (%) =  $(\text{Deaths} / \text{Confirmed}) * 100$ .
- Create a new column Recovery Rate (%) =  $(\text{Recovered} / \text{Confirmed}) * 100$ .
- Round both new columns to 2 decimal places.

## 5. Sorting

- Sort the DataFrame by Mortality Rate (%) in descending order.
- Sort the DataFrame by Confirmed within each country.

## 6. Indexing & Resetting

- Set Country/Region as the index.
- Reset the index back to default.

## 7. Date-Based Analysis

- Find the first date each country reported more than 1000 confirmed cases.
- Find the last date each country had active cases greater than 10,000.