

Context

Data processing lab practical with Python basics and Python libraries NumPy and Pandas

Coronavirus disease 2019 (**COVID-19**) is a contagious disease caused by severe acute respiratory syndrome **coronavirus** 2 (SARS-CoV-2). The first case was identified in Wuhan, China, in December 2019. The disease has since spread worldwide, leading to an ongoing pandemic.

Monitoring the spread of the coronavirus disease is expected to monitor epidemiological trends, rapidly detect new cases, and based on this information, provide epidemiological information to conduct risk assessment and guide disease preparedness.

In this practical you will have the chance to do initial exploratory about a COVID-19 dataset with learned skills of Python basics and Python data science libraries NumPy and Pandas.

For above dataset details following is implemented:

1. Good structure of Python Jupyter Notebook
 - a. Containing title cells, subtitle cells.
 - b. Python codes are reasonably separated into groups (code cells) with functionalities.
 - c. Containing meaningful comments and sensible variable and function names.
2. download csv data with pandas with below code:
 - `import pandas as pd`
 - `deaths_df = pd.read_csv('https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_global.csv')`
 - Identify and handle missing values in the dataset.
 - Remove duplicate entries if any.
 - Convert the date column to a consistent date format (e.g., YYYY-MM-DD)
 - Save the cleaned dataset to a new CSV file.
3. Display first 5 rows of the loaded data (1 mark) and do a short summary about the data (2marks)
4. Calculate the mean and median of the daily cases.
5. Get daily death cases worldwide (hint: summarizing daily death cases over all countries.)
6. Get daily increase of deaths cases via defining a function (hint: use the death cases of today minus the death cases of yesterday from the data obtained in task 5.)
7. Visualize the data obtained in task 5 with library matplotlib