

# COVID Exercise

## Data

[https://github.com/CSSEGISandData/COVID-19/tree/d8431834171d6aa66196fa71c6e4d8ff1de42cd4/csse\\_covid\\_19\\_data/csse\\_covid\\_19\\_daily\\_reports](https://github.com/CSSEGISandData/COVID-19/tree/d8431834171d6aa66196fa71c6e4d8ff1de42cd4/csse_covid_19_data/csse_covid_19_daily_reports)

## Goal

The goals of this exercise are

1. Understanding the quality of your coding skills. Therefore, the code should be production-oriented.
2. Understanding the way you think and deal with raw data that contains many flaws.

## Tasks

1. Load the first 6 months of the data.
2. Data cleansing
  - a. You're given raw data that contains many flaws. Please find missing values, misleading values, etc. in the data and fix them.
  - b. To find the flaws, you may need to analyze them first using statistics and graphs. Please keep the code and results for review.
3. Please describe in words only what will be the process in your opinion that is required for predicting the future number of active patients per country.

## Scope

This task should take up to 3 hours.

It is understandable not to find all flaws that exist, the goal is to find the most important ones in the given time.