

DATA ANALYSIS FINAL ASSESSMENT INDIVIDUAL PART

Download the data set "Daily new confirmed COVID-19 cases per million people" from <https://ourworldindata.org/covid-cases> (as a csv file).

Using Anaconda, do the following tasks (you will need to clean the data first).

- (1) Bar plot the number of new cases for the countries (Canada, France, Germany, India, UK, USA) on 1 January 2021.
- (2) Fit an exponential regression model $y = ab^x$ to the number of daily new cases in the UK, between dates 25 May – 17 Jun 2021. What are a and b ? For the values of a and b obtained, plot $y = ab^x$ in the same plot with the number of new cases in the UK, between dates 25 May – 17 Jun 2021.
- (3) Use machine learning regression (`sklearn`) for the number of cases per day between 26 June–10 August in the USA, to predict the number of cases on 1 September.

Upload your Python file, *with your name and with comments to explain what you are doing*, to Github:

<https://github.com/jonitera/Data-Science-Final-projects>,
by 25 August (midnight BST).