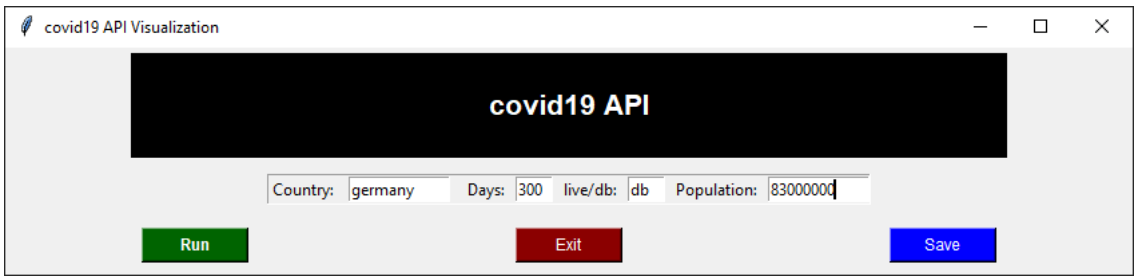
Web Scraping on covid-19 API using Python

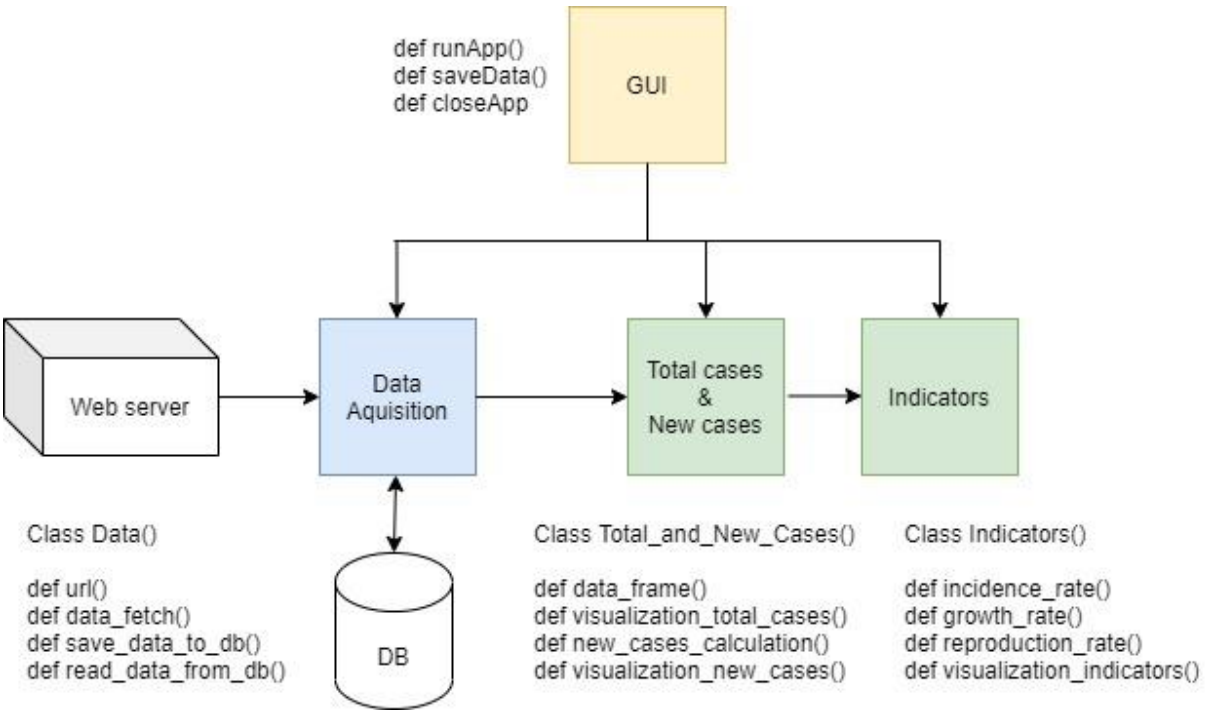


This application allows us to visualize covid-19 data from different countries. The user chooses the country and the range of days from today backwards. Furthermore, the user can choose to read the data “live” from the web server or from a “db” database, as well as save the data to the database. The data is obtained from a free API (https://covid19api.com/), which takes data from John Hopkins CSSE.

For this project, I technically do web scraping, data processing, simple analysis and data storage.

The application shows the total number of cases (confirmed, deaths, recovery and active) in **figure_1**, subsequently, the new cases are from the aforementioned parameters calculated and shown in **figure_2**, and at the end, indicators such as the incidence of 7 days (per 100.000 habitants) are also calculated and shown in **figure_3**.

The application consists of four modules, where each one performs a specific task as indicated in the following diagram.



The diagram shows the modules and the classes with their respective functions.

Figure_1. Shows raw data of accumulative cases. The active channel shows apparently outliers coming from the API server.

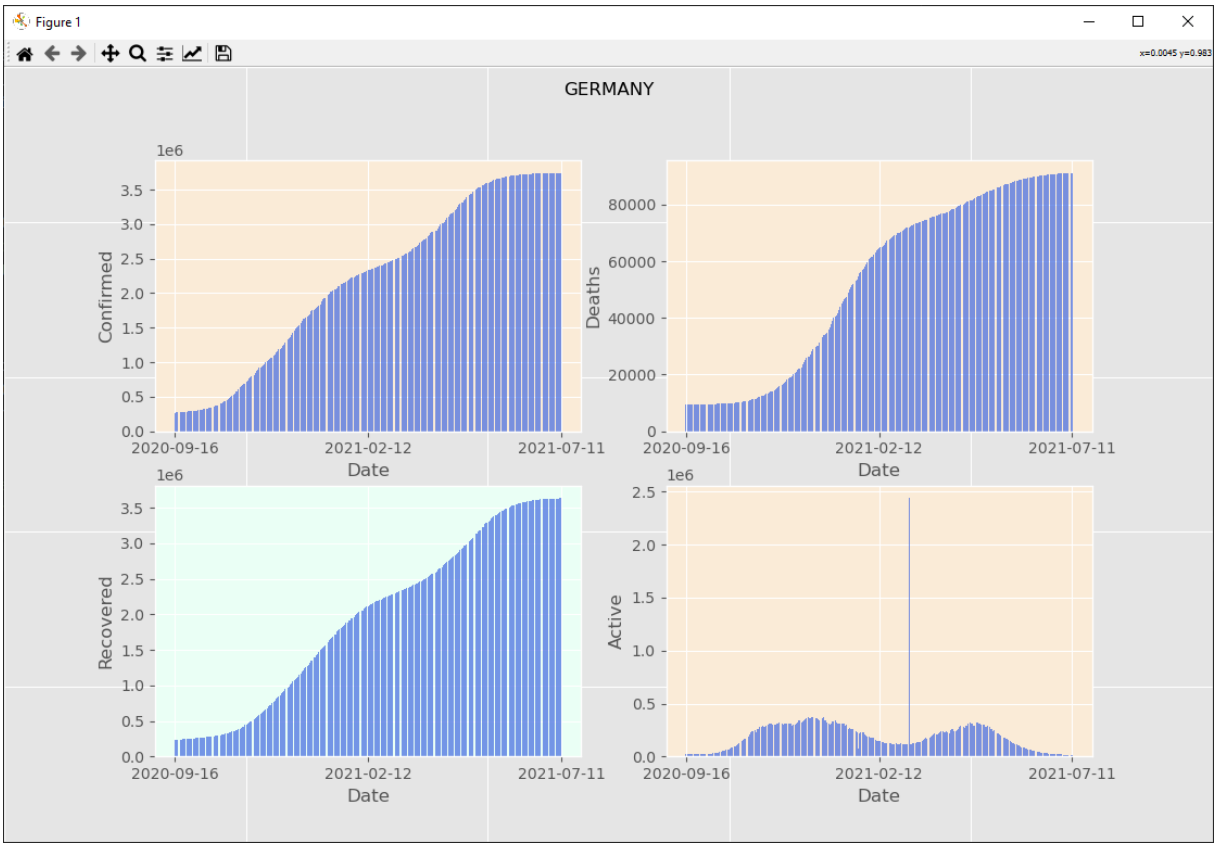


Figure_2. Shows the calculated “new_cases” in bars and the corresponding mean of 7 days in orange line plot.

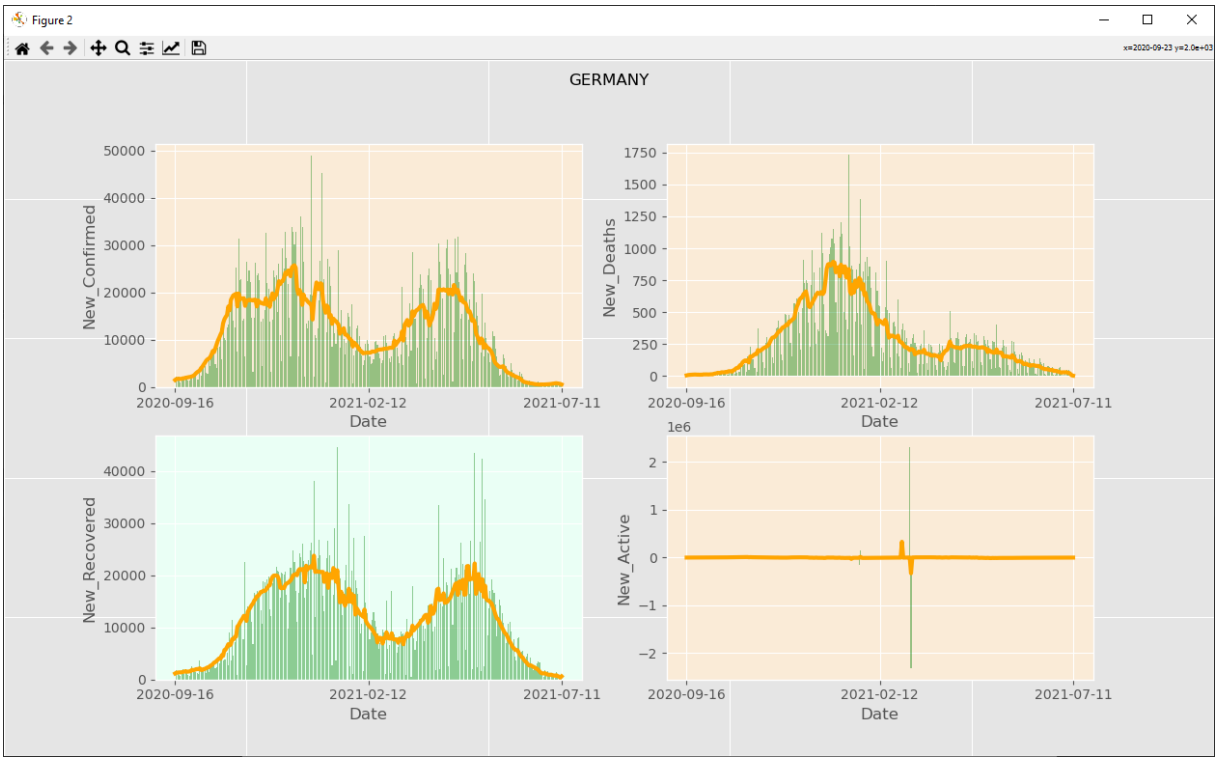
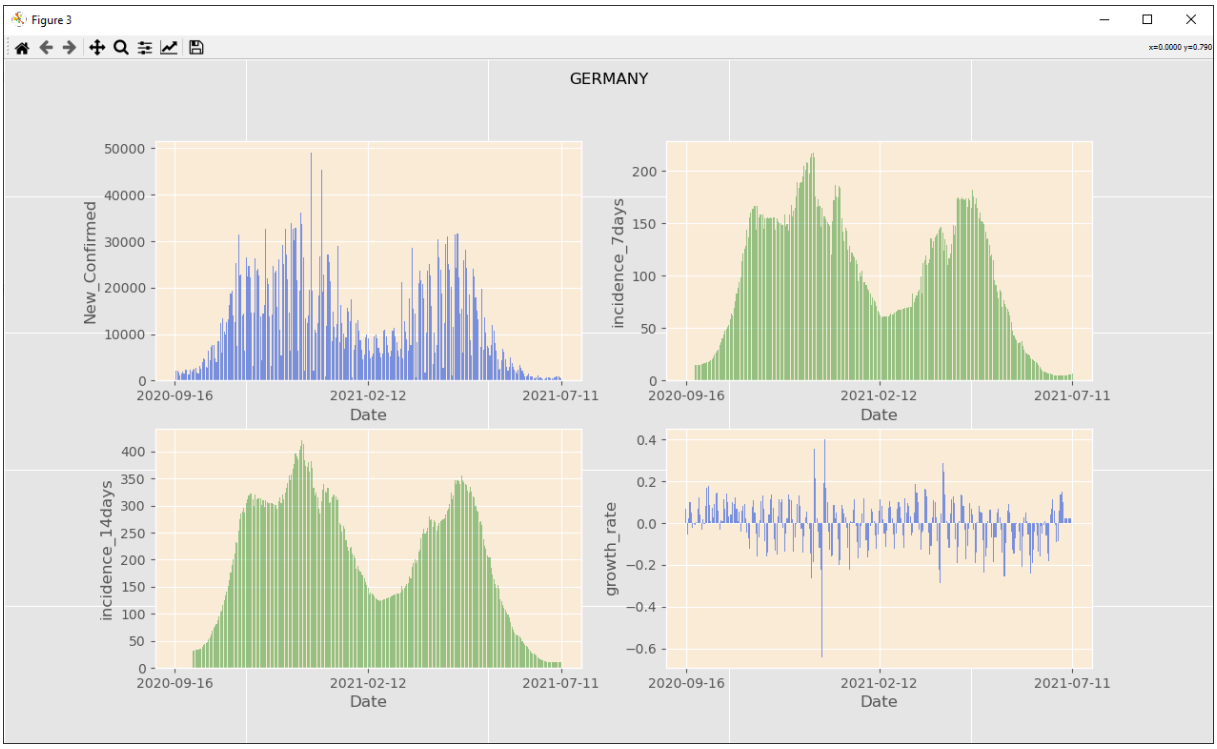


Figure_3. Shows the standard calculation of the incidence for 7 and 14 days respectively (per 100.000 habitants). Furthermore, “growth rate” is the coefficient of an exponential function fitted for 7 days in the “new_confirmed” channel.



The objective of this application was to show technical skills on web scraping and data visualization. Deep discussions about the calculation of indicators that characterize the pandemic are out of the scope of this small project and require domain knowledge.