1.0-exploratory-analysis-globally

October 21, 2020

1 Imports

2 Goal

My goal is to visualize various aspect of the `COVID-19` pandemic. In this notebook we focus on a high level overview as well as the early days of the pandemic.

3 Data sources

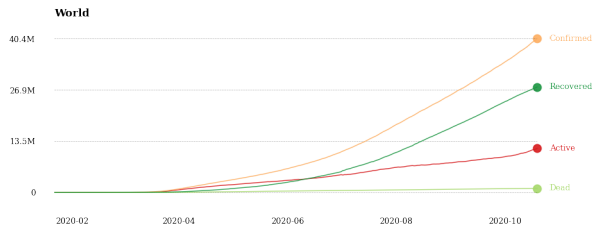
In this notebook I use data from the following sources: - https://github.com/CSSEGISandData/COVID-19 - JHU CSSE COVID-19 Data. - https://datahub.io/JohnSnowLabs/country-and-continent-codes-list - country codes and continents.

4 Data loading

See `src/visualizations/covid_data_viz.py` for code to load the data and reproduce the plots.

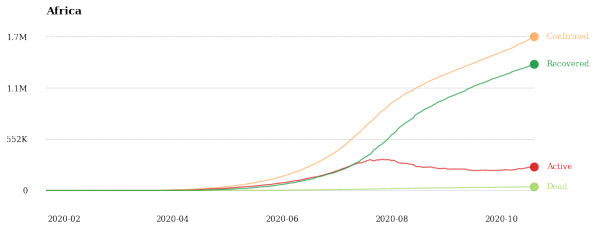
5 COVID19 - Globally

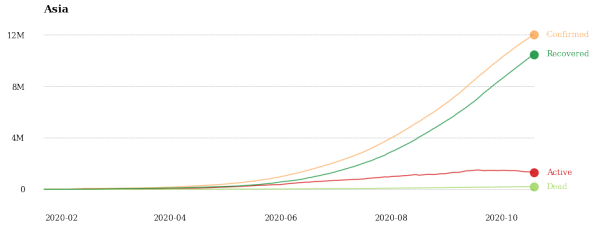
The `SARS-CoV-2` disease was first discovered in `Wuhan, Hubei, China` in December 2019. The disease spread rapidly throught the world. In this notebook we see how rapidly and to what extent. Below we plot types of cases worlwide. At the time of writing `(2020-10-05)` the `confirmed` cases are nearing 35 milion.

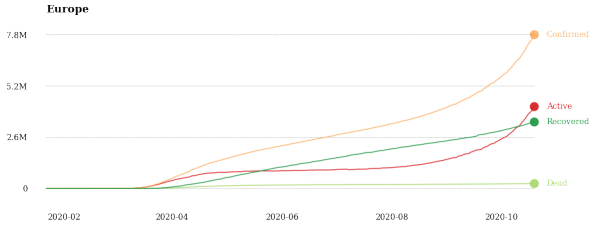


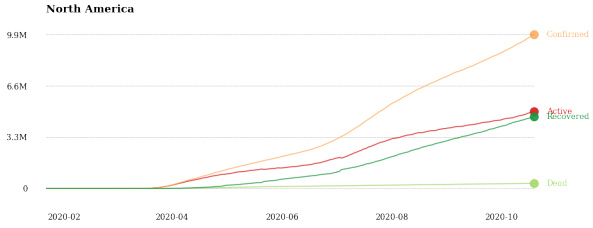
6 COVID19 - By continent

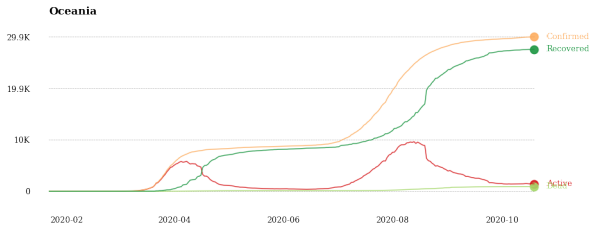
We drill down further into the data visualizing cases data at the continent level. Note that `Asia` is still exhibiting exponential growth and that the continent with least amount of `confirmed` cases is `Australia`.

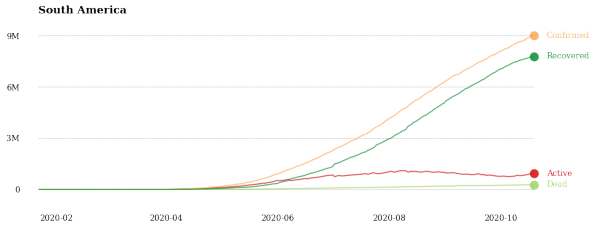






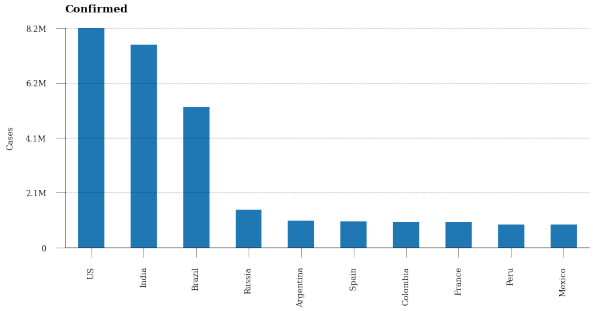


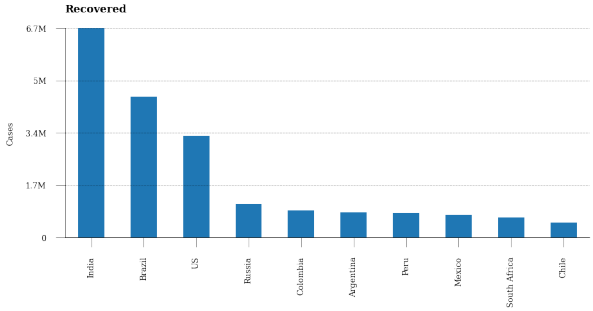


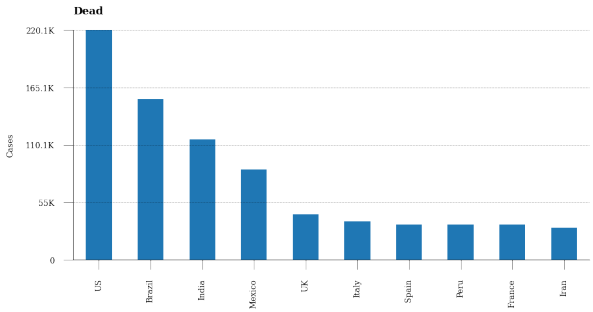


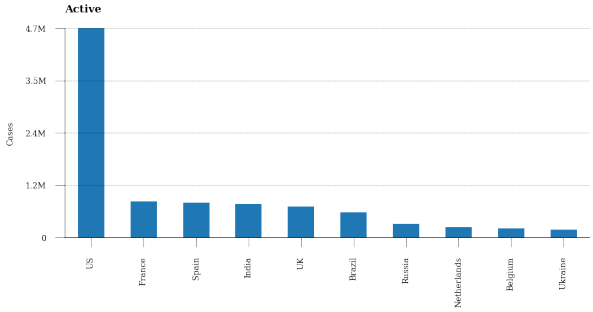
7 COVID - By country

Finally we can take a look at the data on a country level. First we show the countries with the most `cases`.

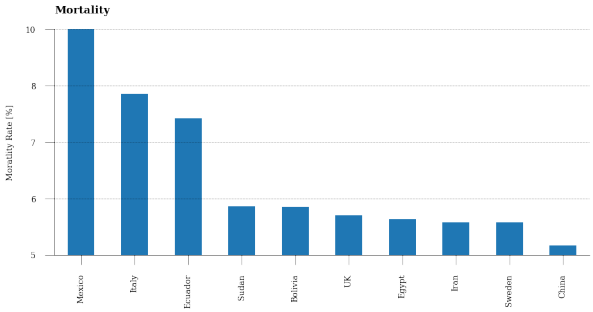








Then we show countries with `confirmed cases >= 5000` and the highest `mortality rate`.

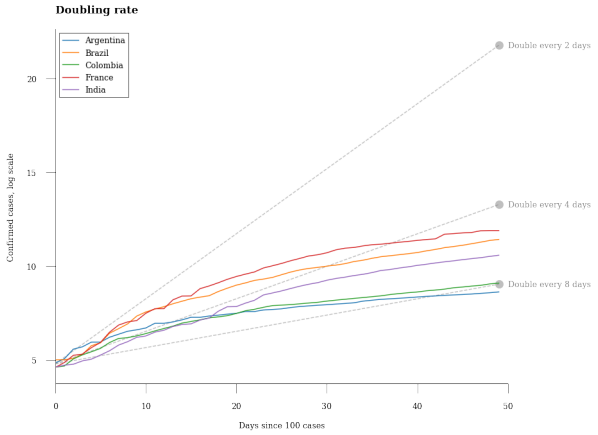


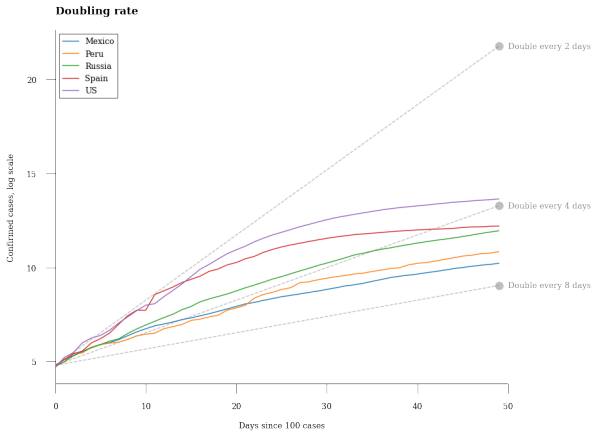
8 COVID19 - Growth rates

We proceed to visualize the early days of the pandemic. In most cases the growth was exponential as will be seen. We use the reindexed `confirmed_cases_t0.csv` file to plot growth curves starting at time `t0` ie the date when a country achieved `Confirmed >= 100`. The charts are logarithmically

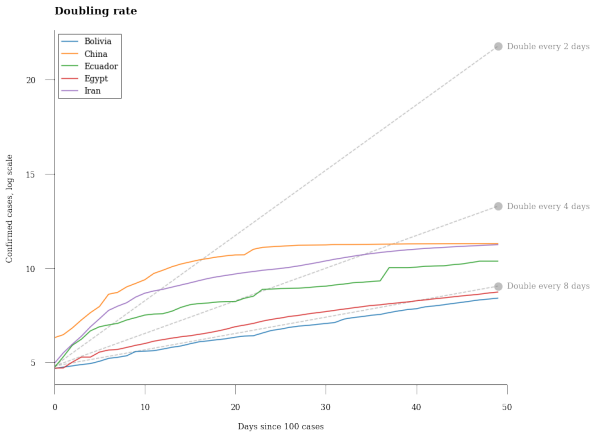
scaled. Note that exponential growth on a logarithmic chart is a line. This follows from simple math. Furthermore observer that most of the countries in the plot are able to combat high exponential growth within `30 - 40` days since the `100`th case.

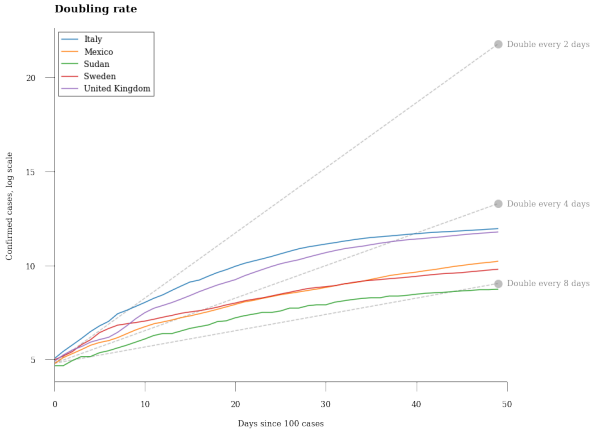
We plot exponential growth curves with different doubling rates against the countries with the most `confirmed cases`.





We plot the case growth curve against the countries with the highest `mortality rates`.





See the next notebook for more.