

# COVID-19 outbreak

Eduard Bakstein, [eduard.bakstein@nudz.cz](mailto:eduard.bakstein@nudz.cz), <http://bakstein.net>, @edbakstein

12.3.2020

## Intro

The aim is to do exploratory analyses of the COVID-19 cases in selected countries and see how the models evolve as new cases are reported. Can be used as a quick&dirty basis for your own analyses. Special focus given to Czech Rep., as this is my country of origin.

Source code (R markdown): <https://github.com/ebakstein/covid19.git>

## Resources:

- <https://rviews.rstudio.com/2020/03/05/covid-19-epidemiology-with-r/> (basis of this rmd and analysis)
- <https://timchurches.github.io/blog/posts/2020-02-18-analysing-covid-19-2019-ncov-outbreak-data-with-r-part-1/>

## Interesting models and remarks

- <https://blog.ephorie.de/epidemiology-how-contagious-is-novel-coronavirus-2019-ncov>
- <https://arxiv.org/pdf/2002.00418v1.pdf>

## Data sources:

- [https://en.wikipedia.org/w/index.php?title=2020\\_coronavirus\\_outbreak\\_in\\_the\\_United\\_States&oldid=944107102](https://en.wikipedia.org/w/index.php?title=2020_coronavirus_outbreak_in_the_United_States&oldid=944107102)
- [https://github.com/CSSEGISandData/COVID-19/tree/master/csse\\_covid\\_19\\_data/csse\\_covid\\_19\\_time\\_series](https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series)

## TO DO

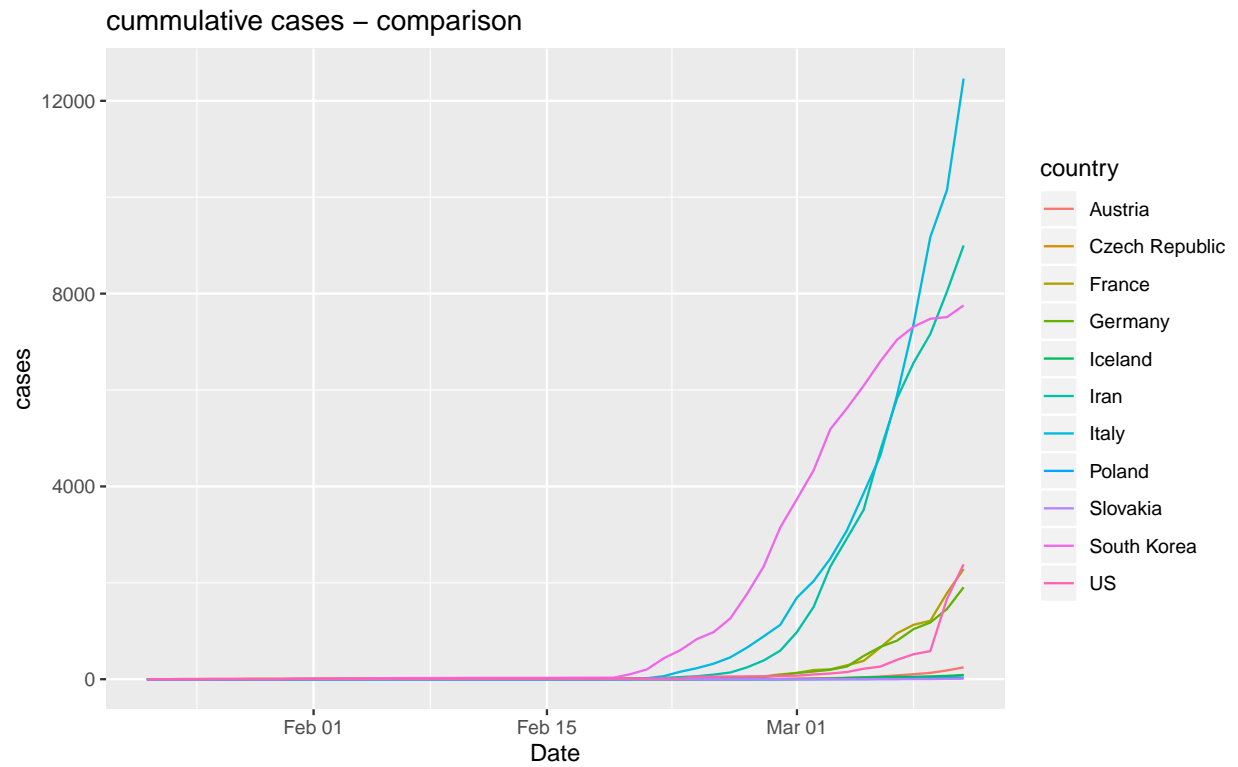
- modelling

## Load data

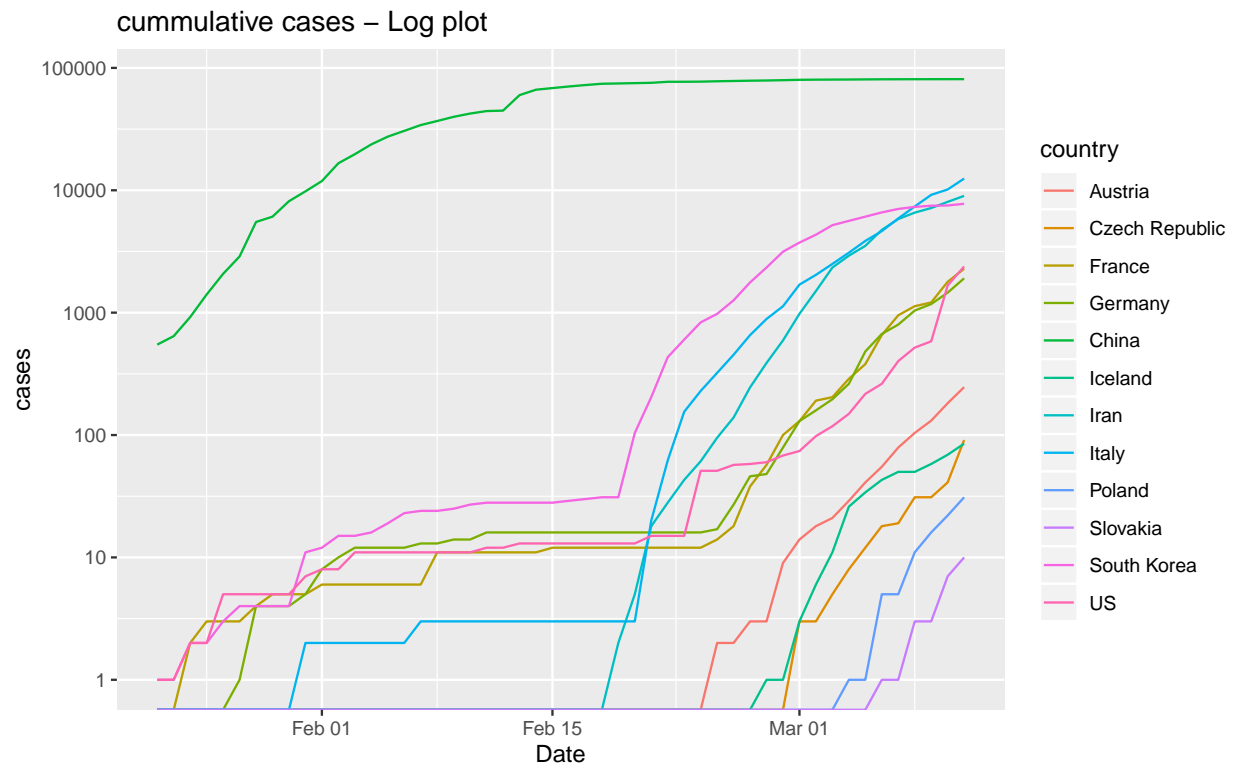
Using the Jons-Hopkins github repository to obtain current data, see [https://github.com/CSSEGISandData/COVID-19/tree/master/csse\\_covid\\_19\\_data/csse\\_covid\\_19\\_time\\_series](https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series) (updated daily around 23:59 UTC)

Obtain data about population in each country by scraping the wikipedia page: '[https://en.wikipedia.org/wiki/List\\_of\\_countries\\_by\\_population\\_\(United\\_Nations\)](https://en.wikipedia.org/wiki/List_of_countries_by_population_(United_Nations))'

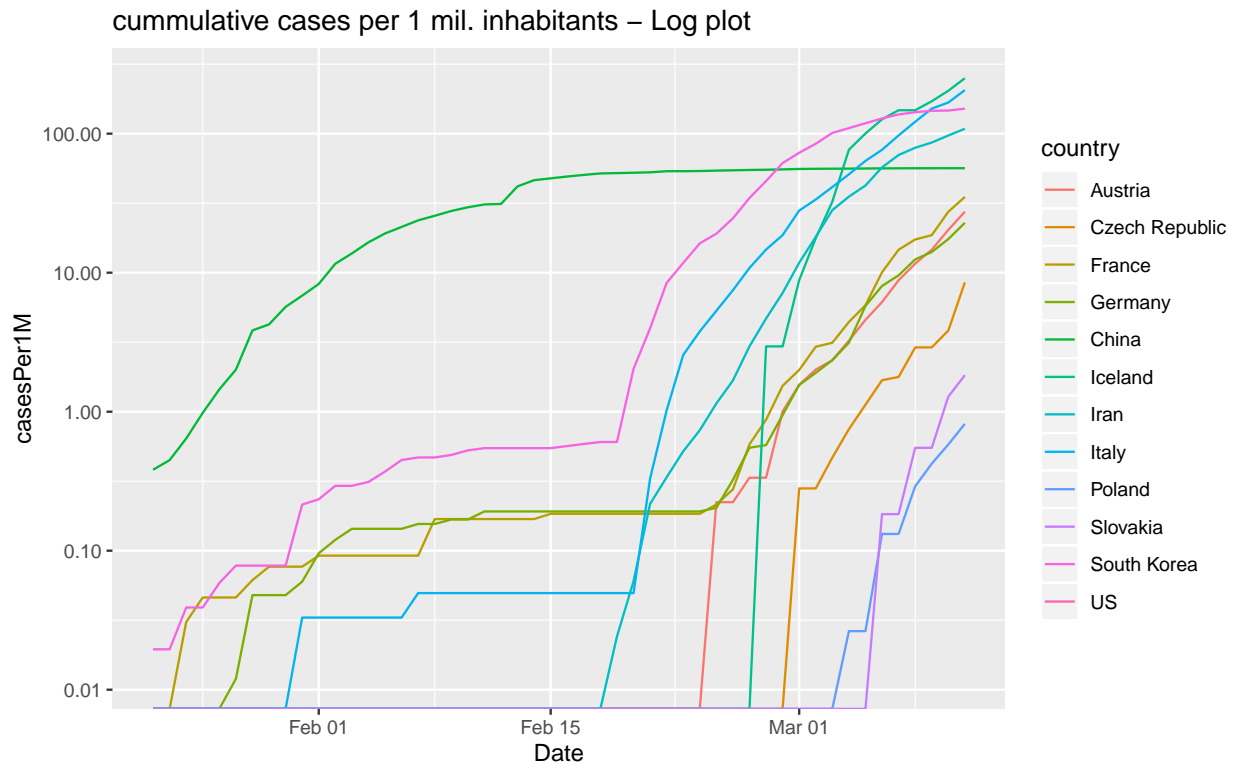
## Plot selected countries



Note: China number of cases too high to be plot on linear y axis with other countries

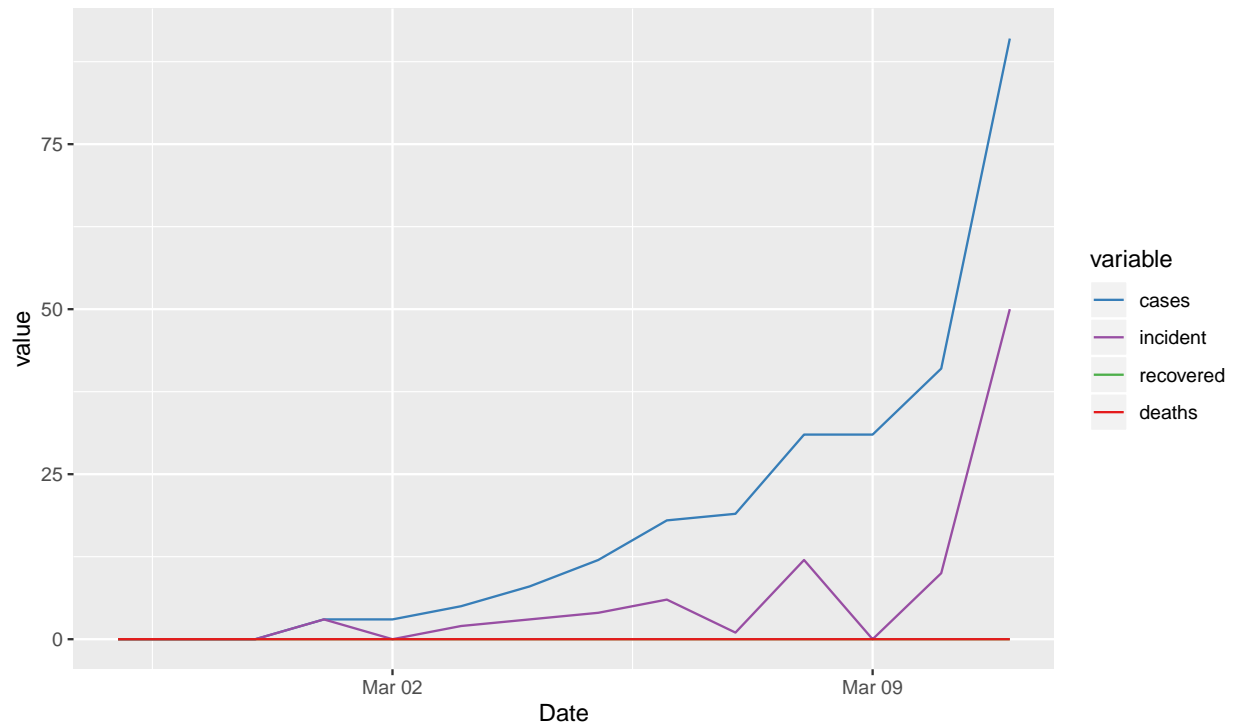


Note: China added to the log plot

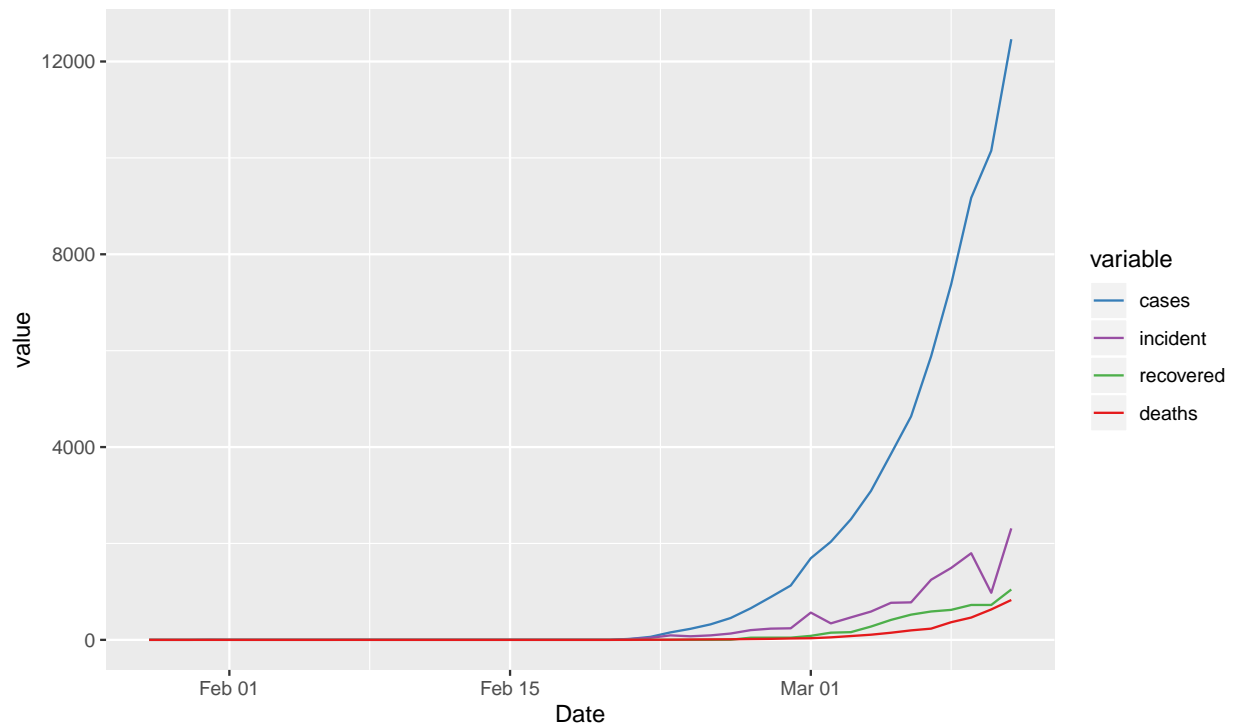


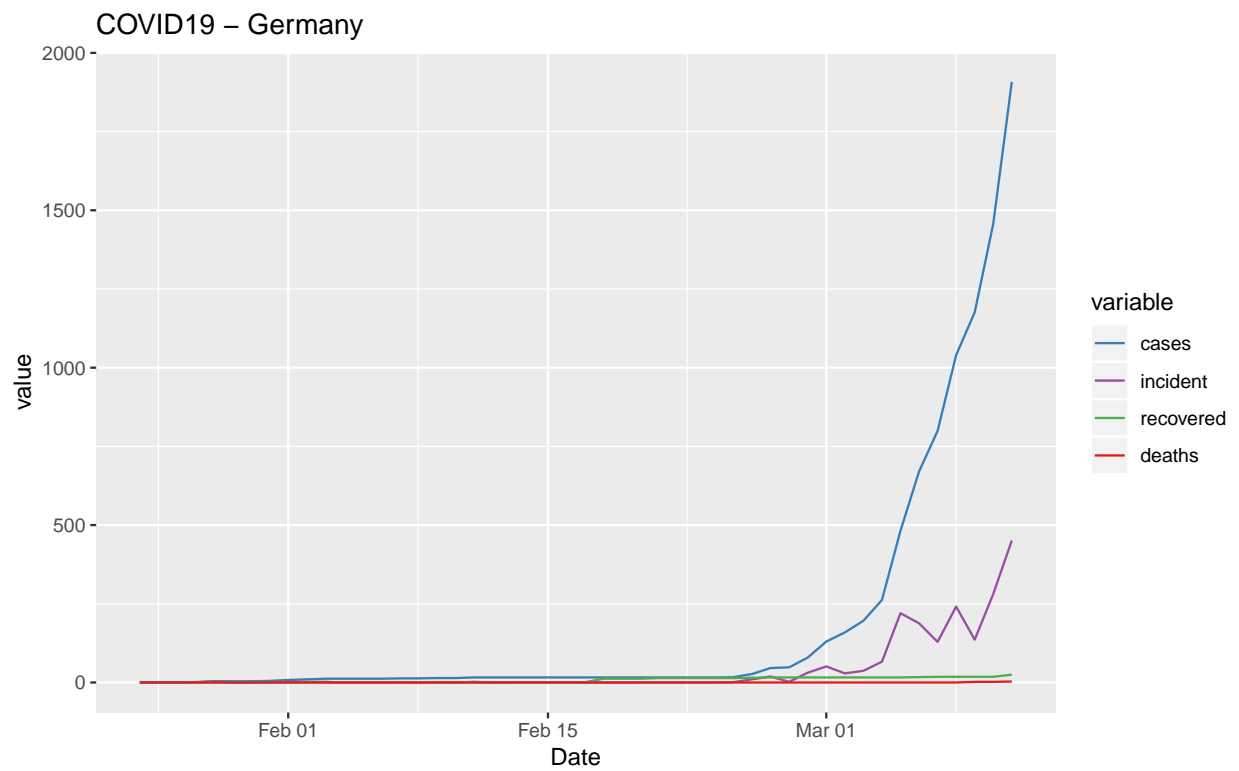
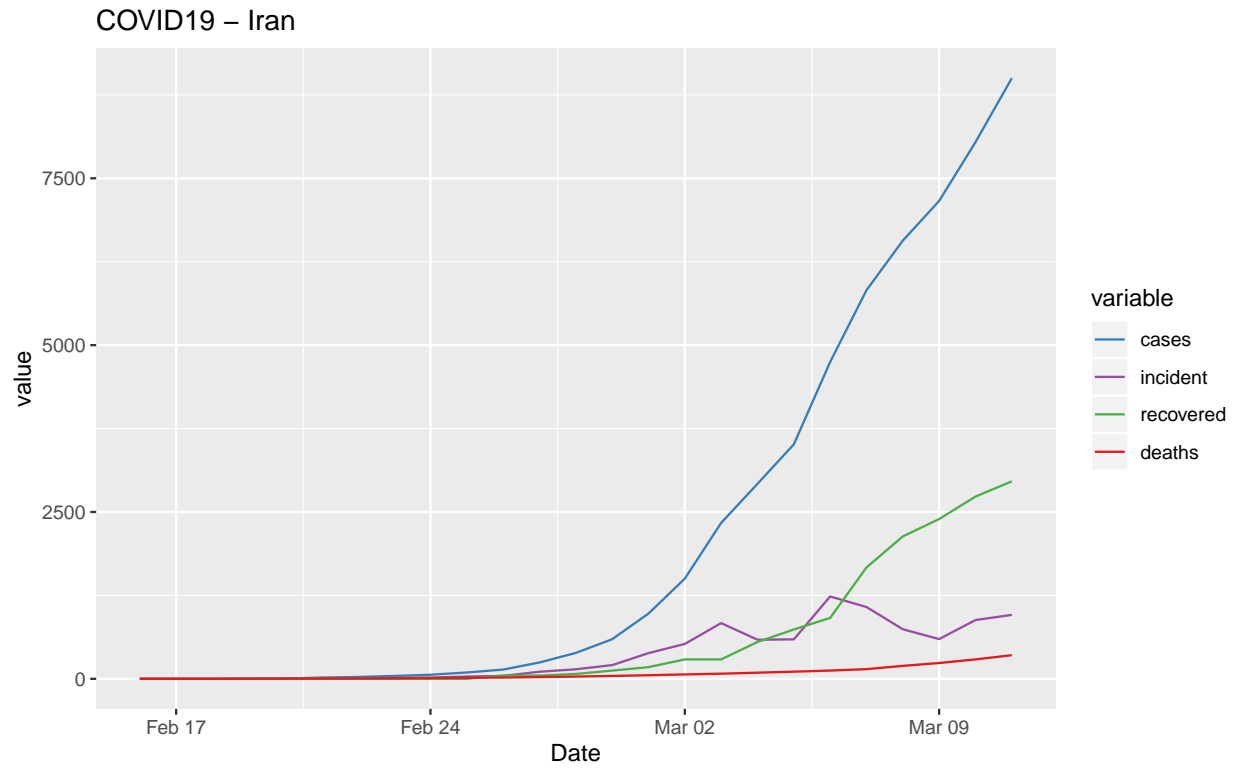
## Cummulative stats for Individual selected countries

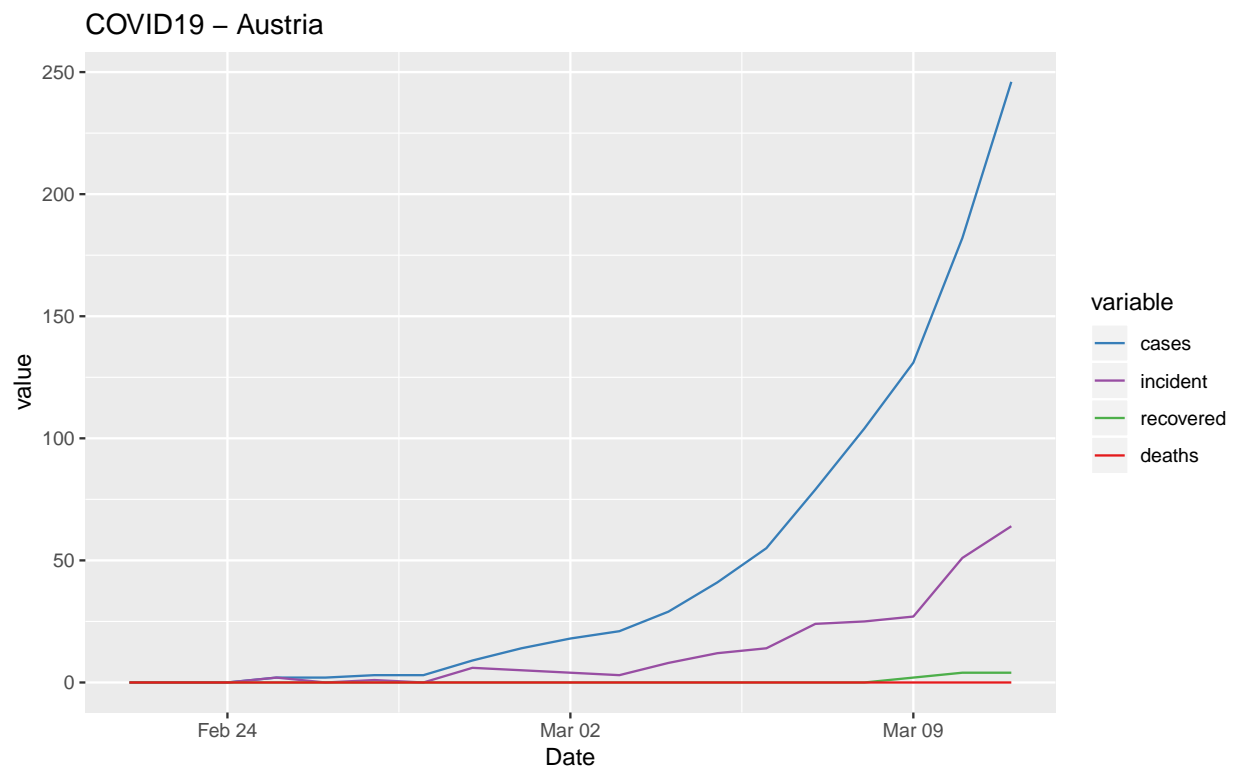
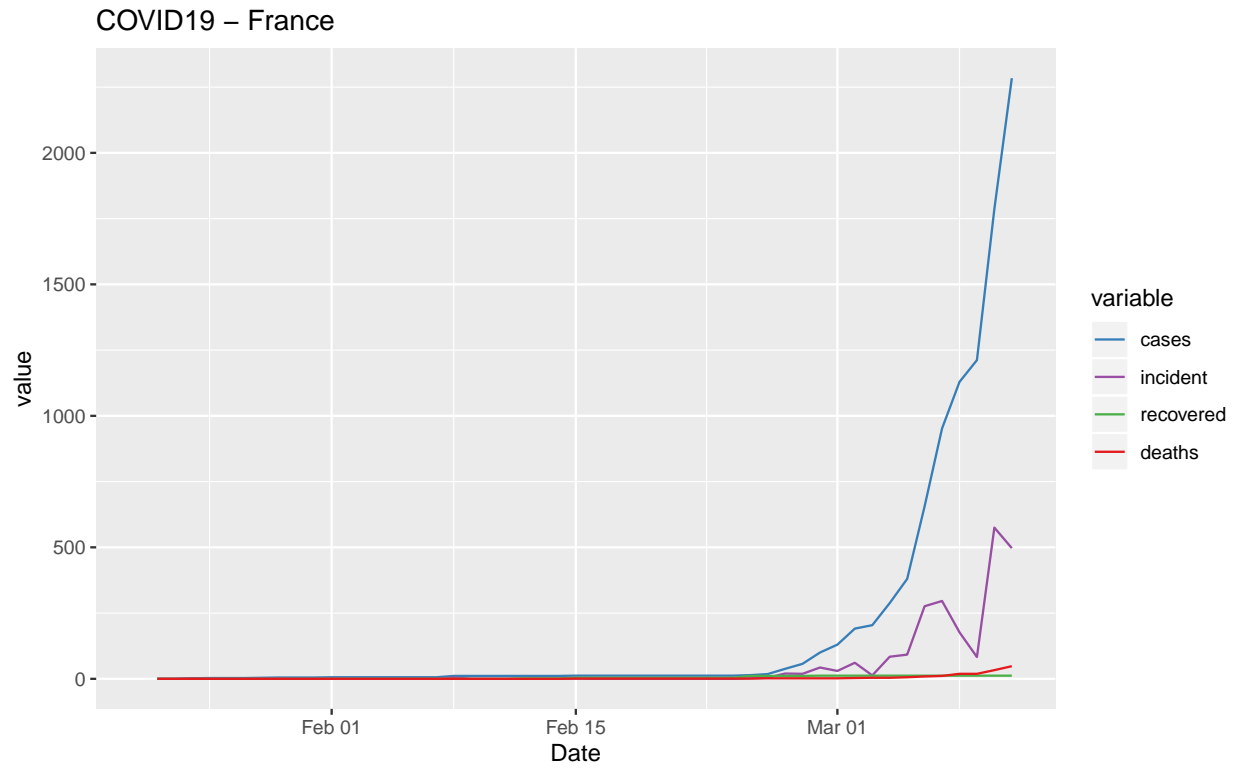
### COVID19 – Czech Republic

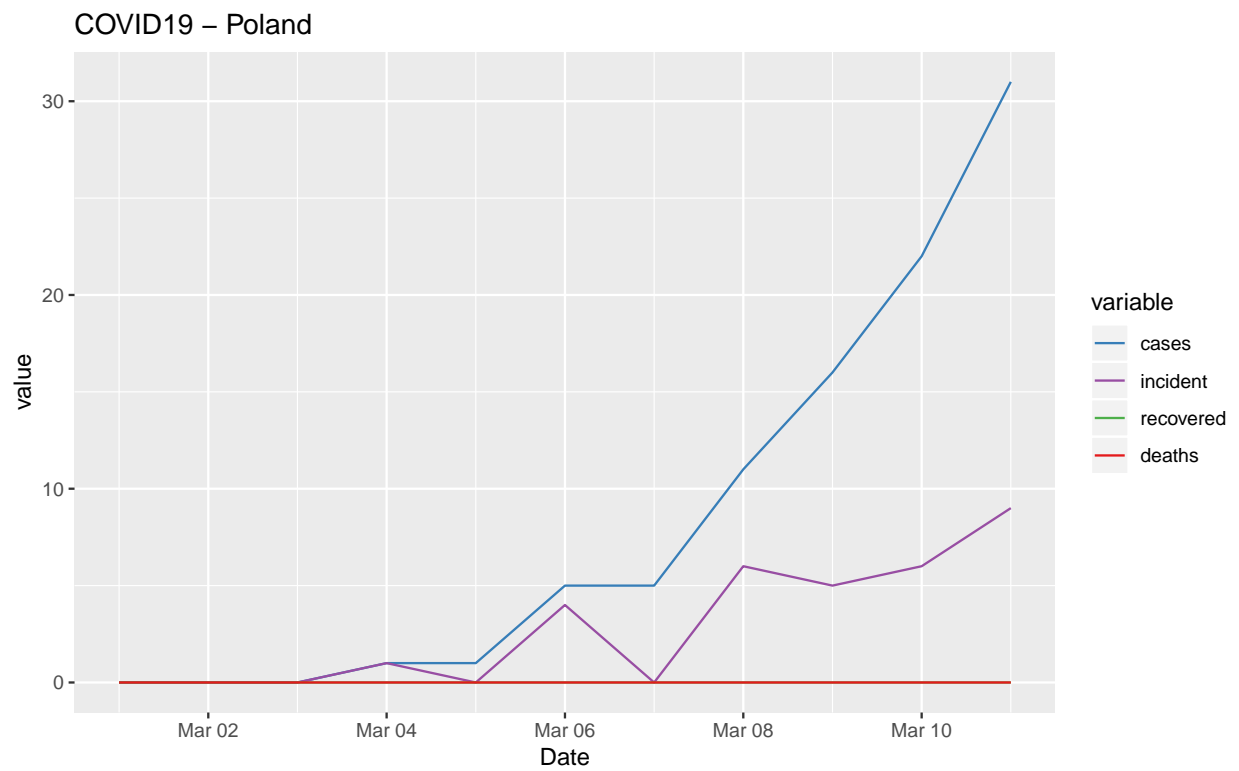
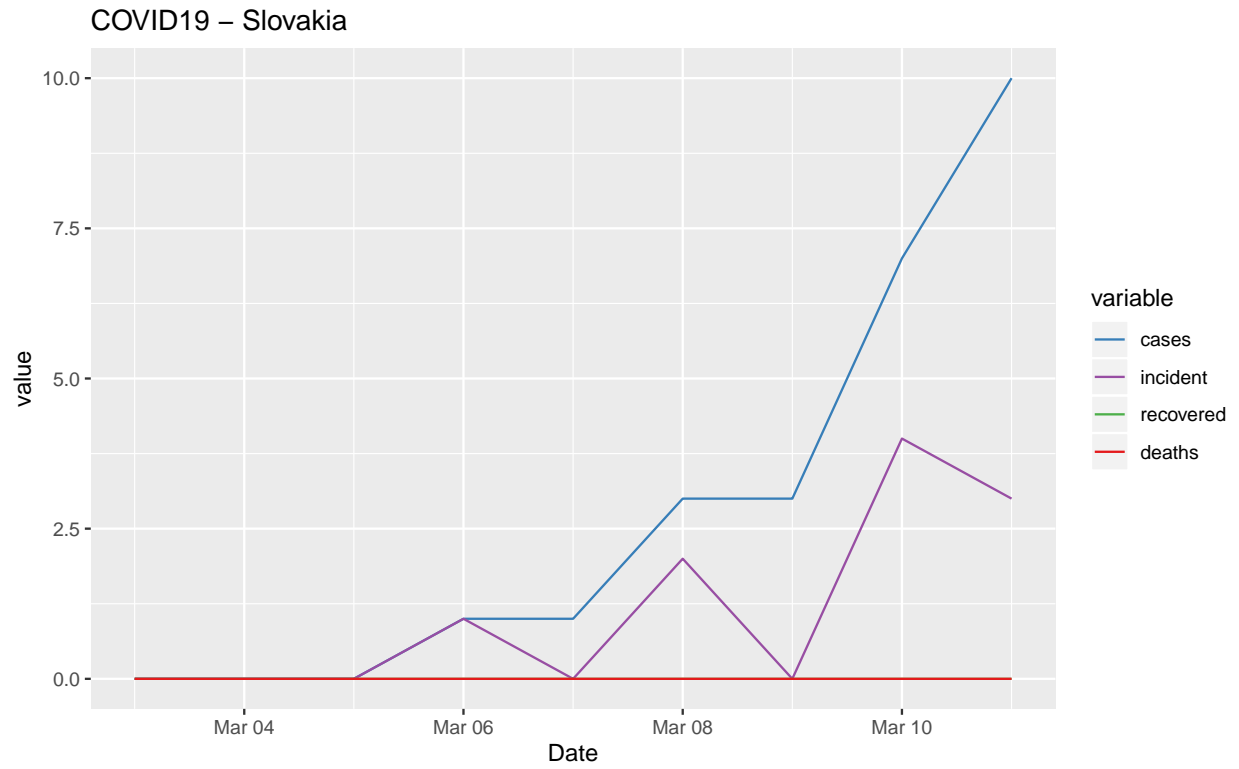


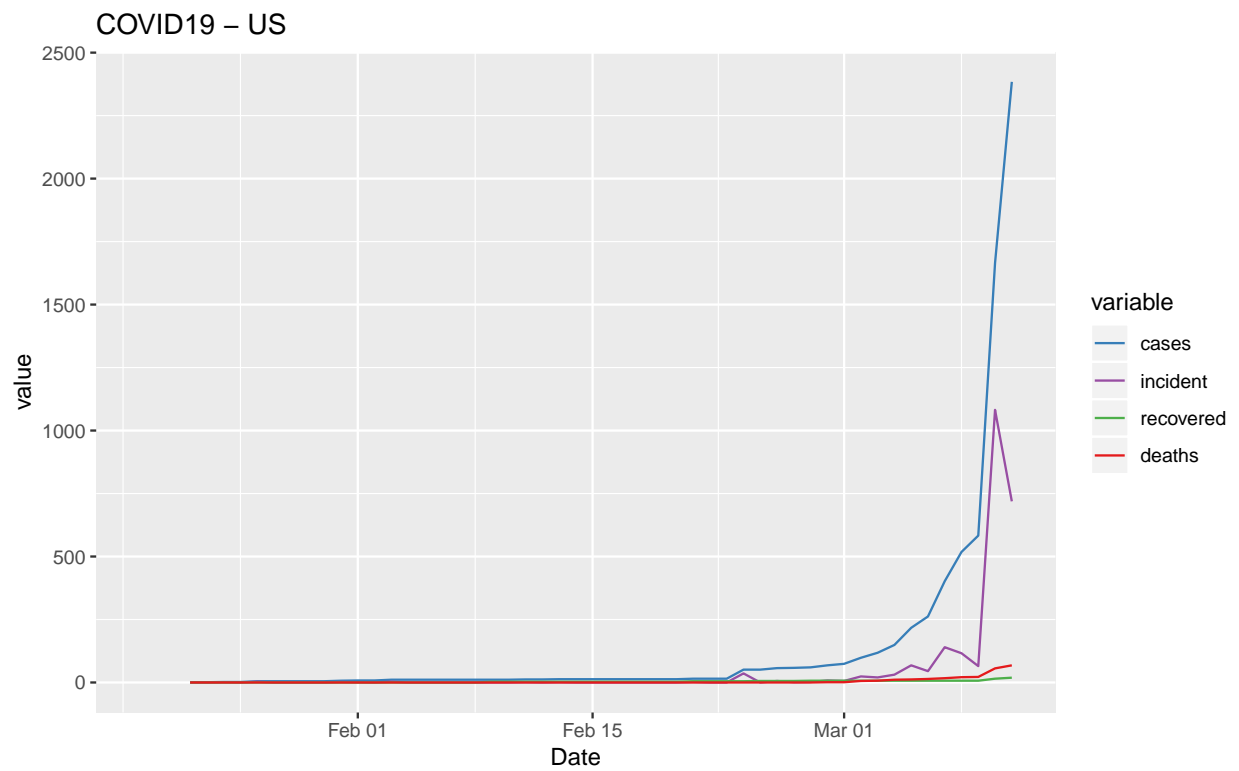
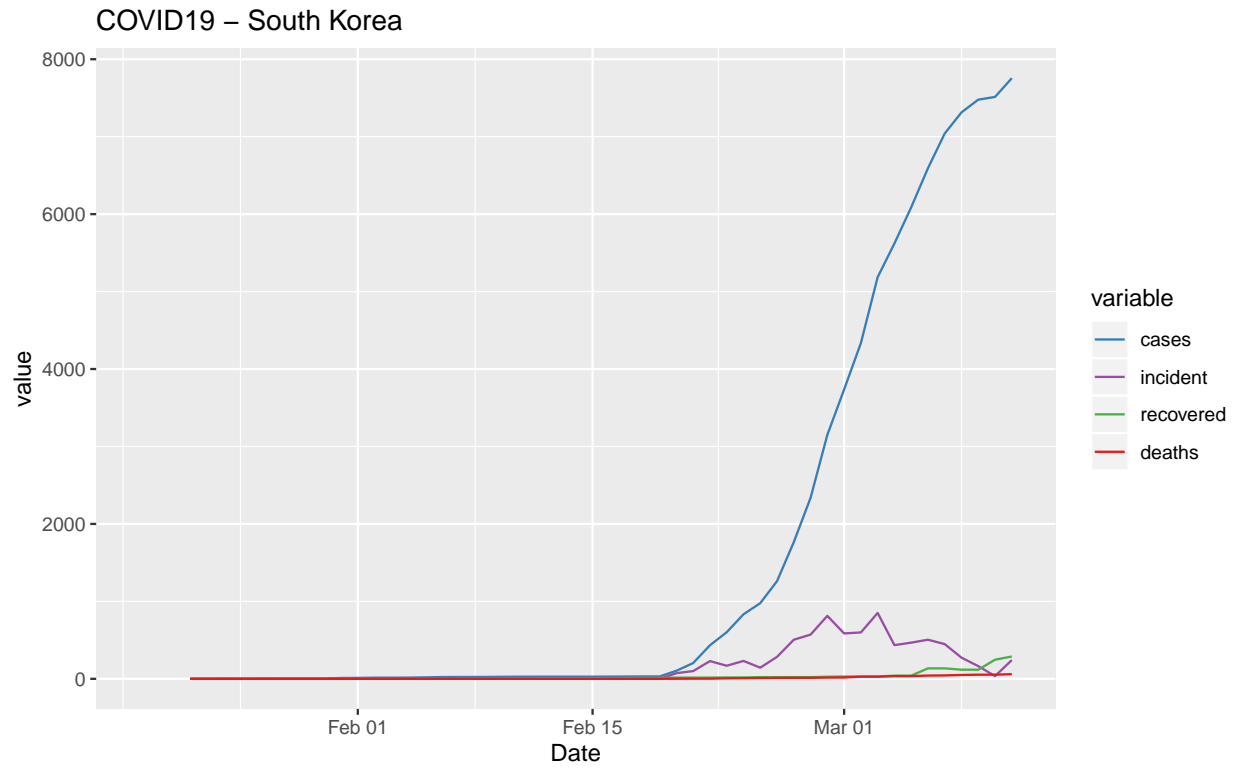
### COVID19 – Italy





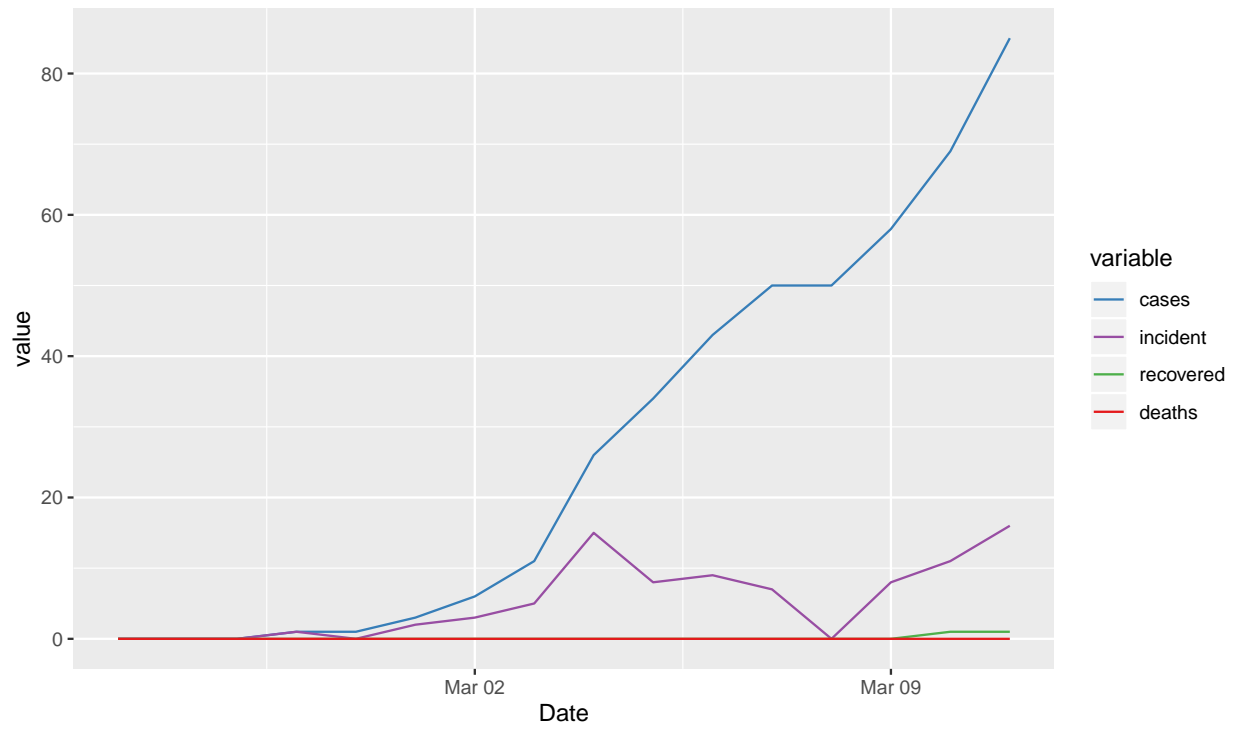




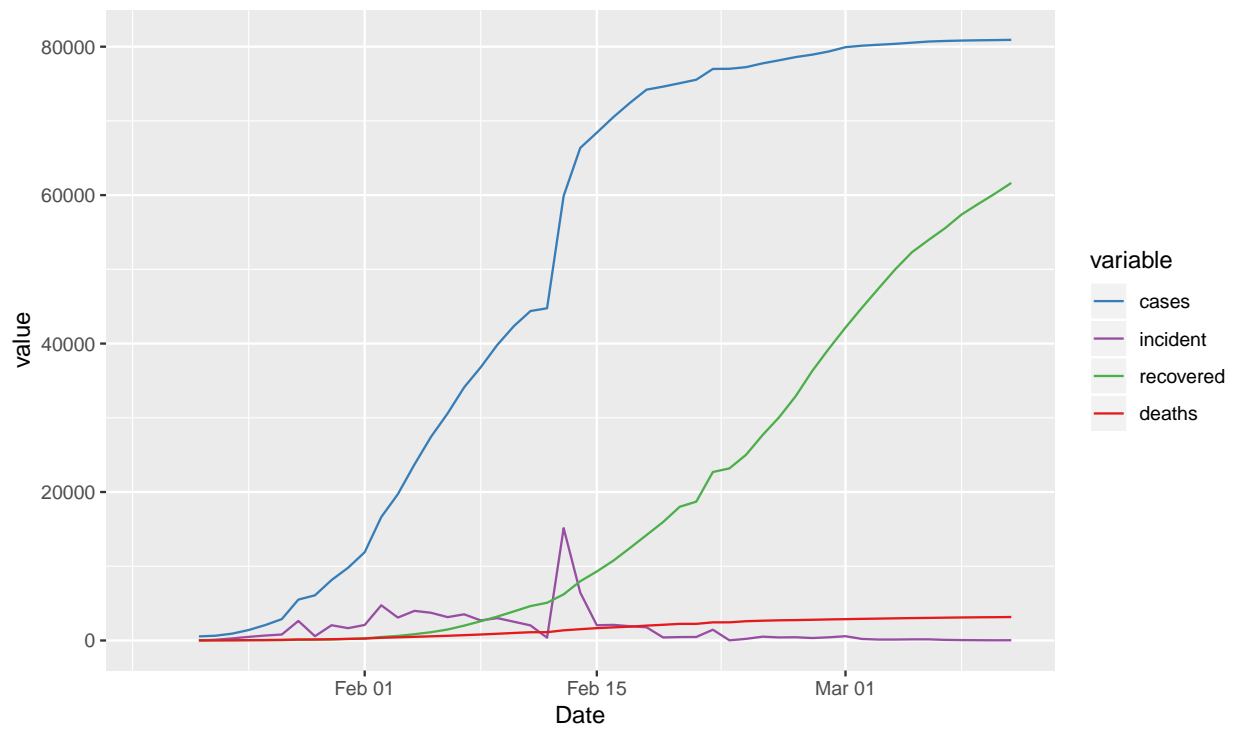




COVID19 – Iceland

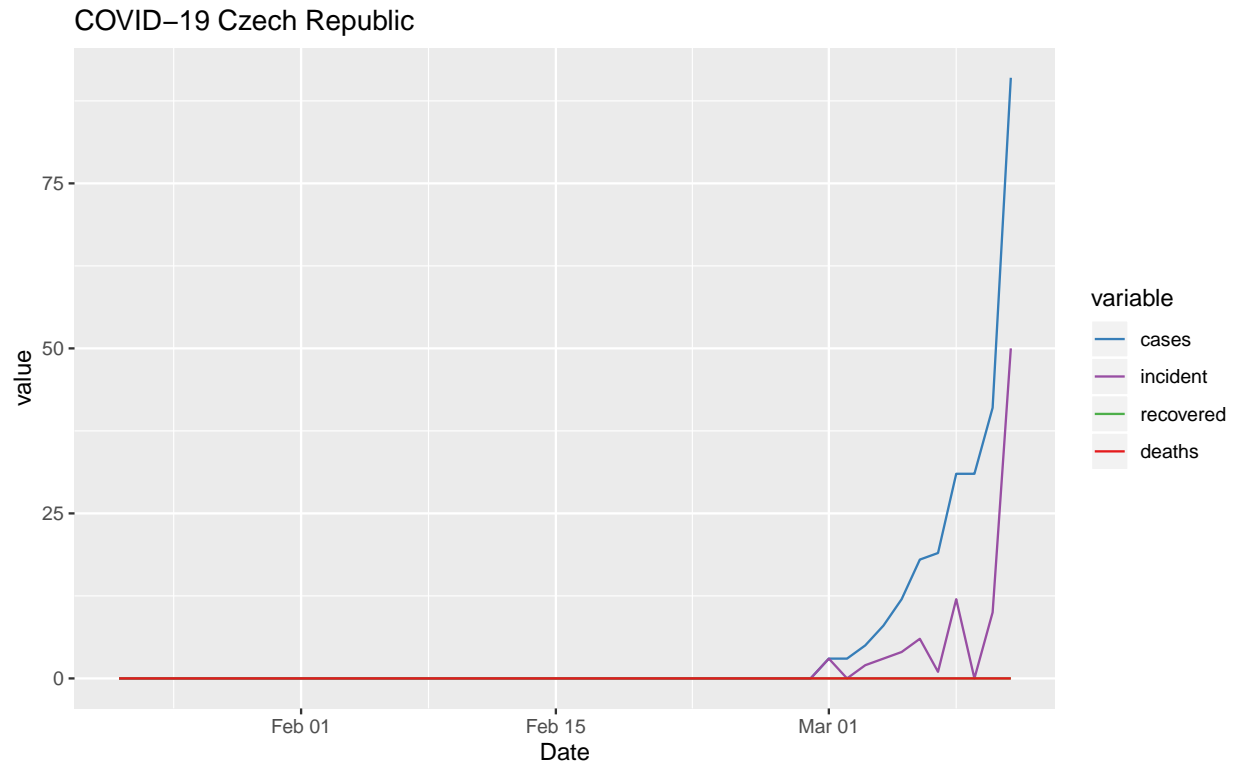


COVID19 – China



## Plot and analyze Czech Rep.

exploratory plots



first case in CZ: 2020-03-01

### Epidemiological modelling - TBD

TBD according to: <https://rviews.rstudio.com/2020/03/05/covid-19-epidemiology-with-r/> And <https://timchurches.github.io/blog/posts/2020-03-10-modelling-the-effects-of-public-health-interventions-on-covid-19-transmission-part-1/>

Store data locally for further reff