COVID-19 outbreak

Eduard Bakstein, 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://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series

TO DO

- modelling

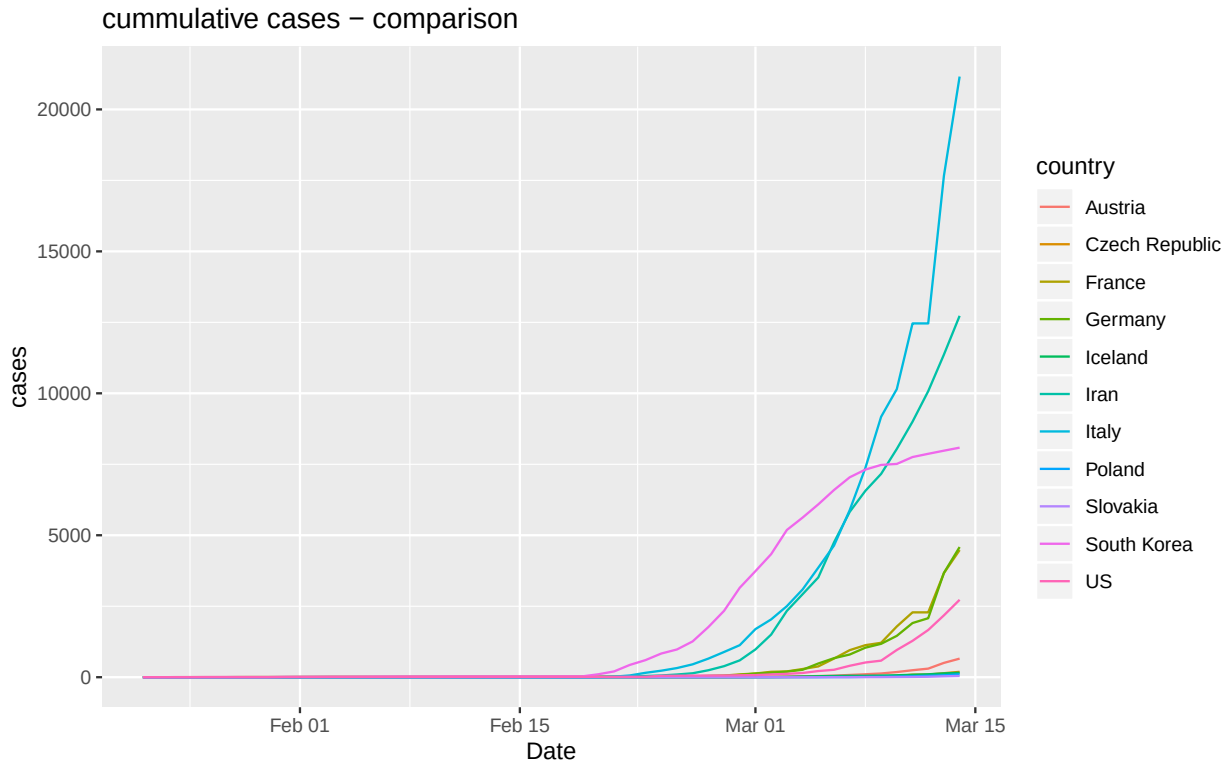
Load data

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 (updated daily around 23:59 UTC)

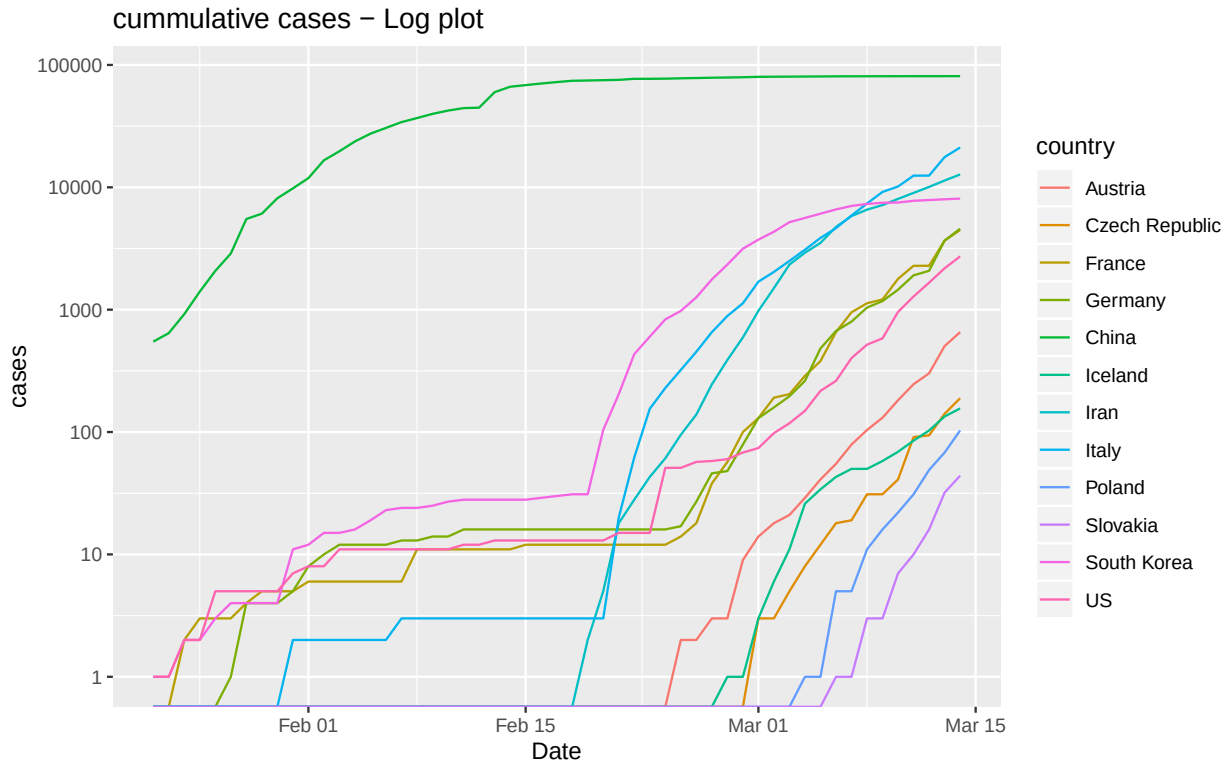
Obtain data about population in each country by scraping the wikipedia page: 'https://en.wikipedia.org/wiki/List_of_countries_by_population_(United_Nations)'

obtain latest czech data from MZ / apify https://api.apify.com/v2/key-value-stores/K373S4uCFR9W1K8ei/records/LATEST?disableRedirect=true (parsed from https://onemocneni-aktualne.mzcr.cz/covid-19)

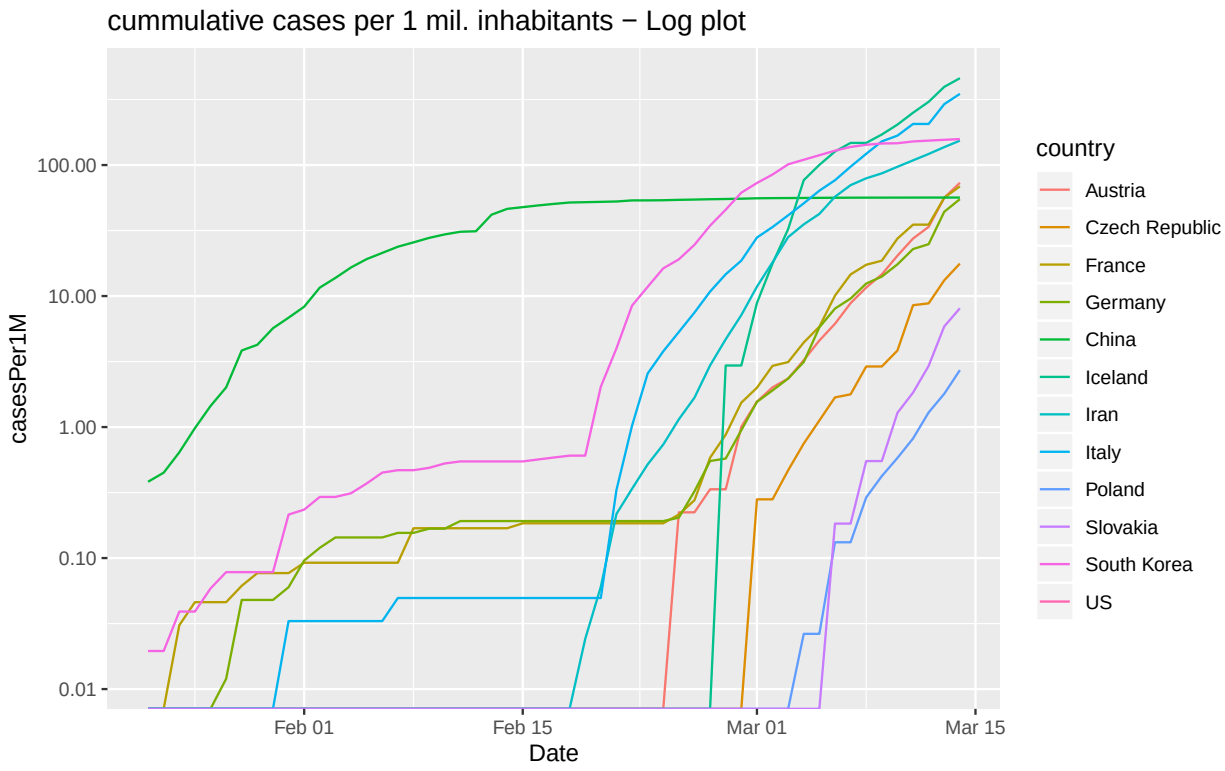
Plot selected countries



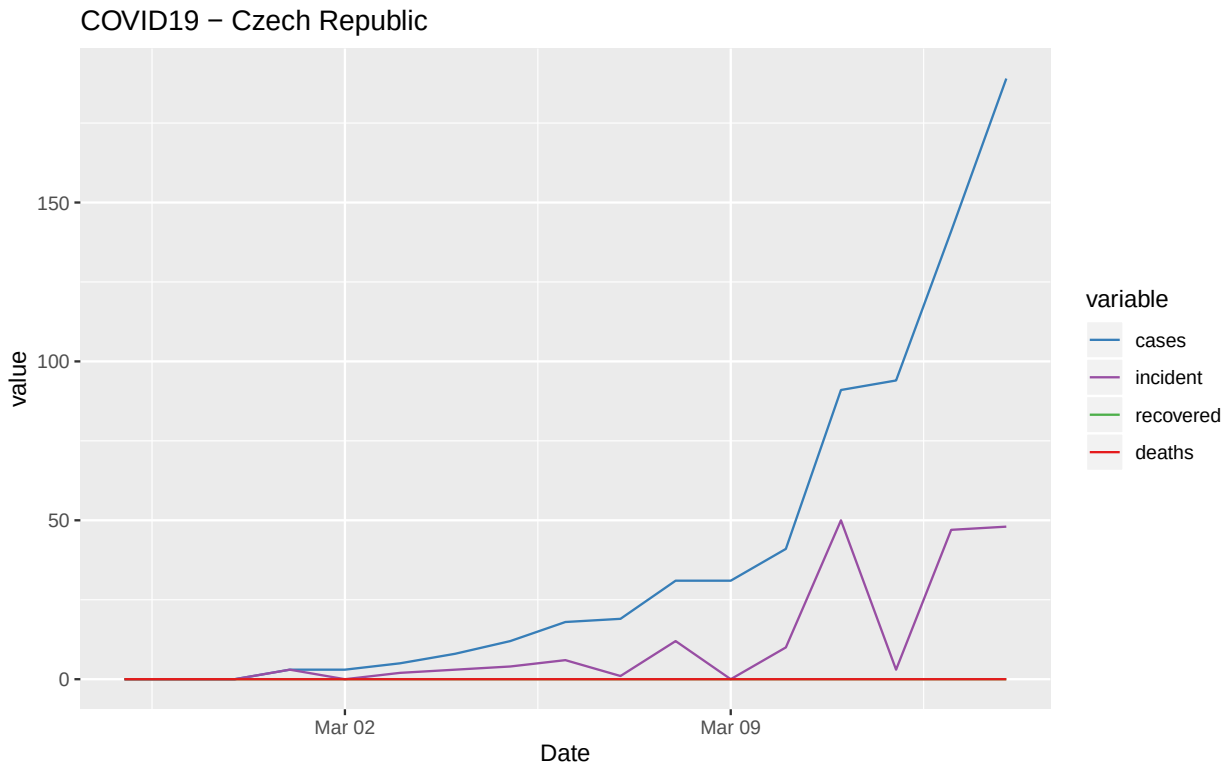
Note: China number of cases too high to be plot on linear y axis with other countries

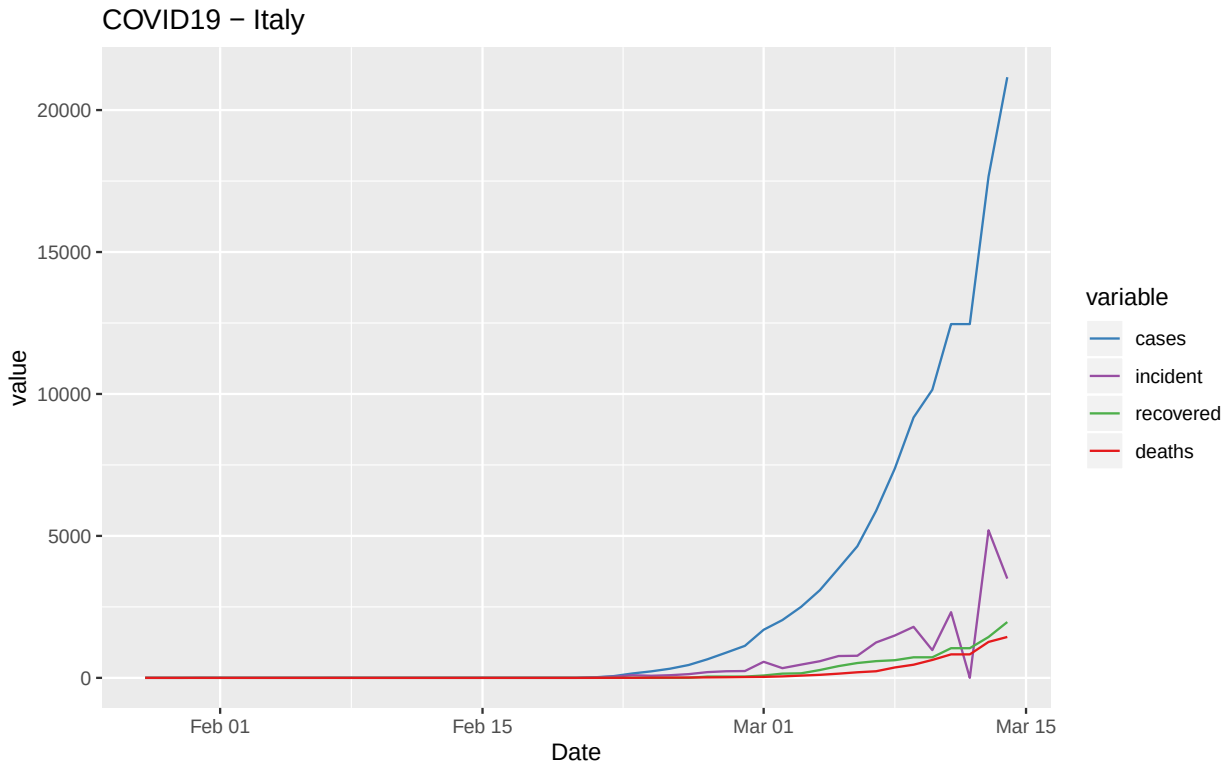


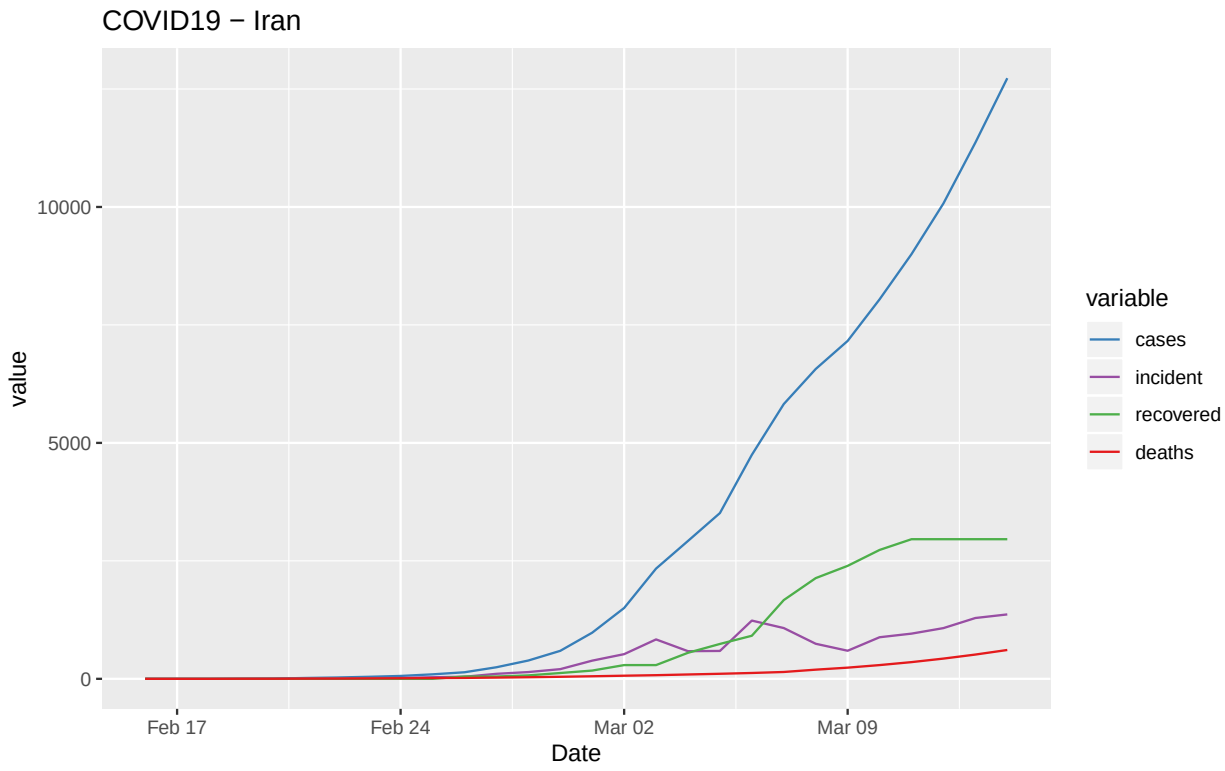
Note: China added to the log plot

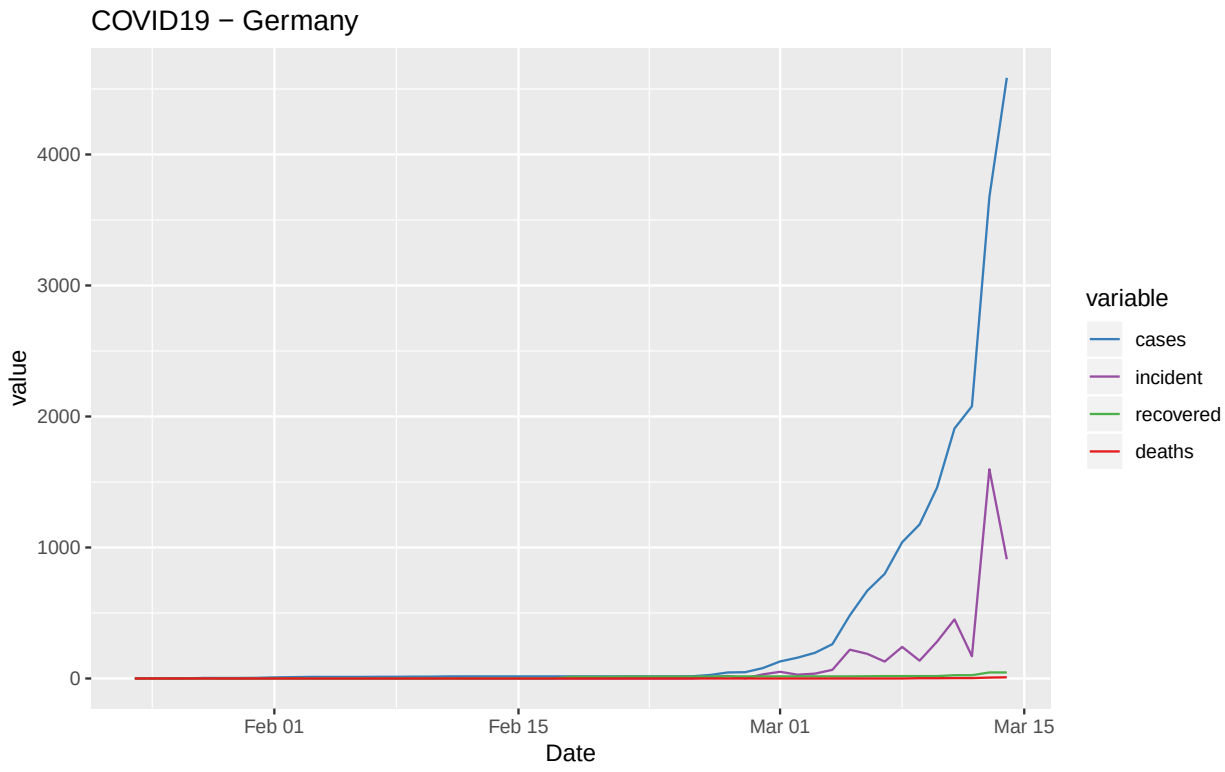

cummulative cases per 1 mil. inhabitants − Log plot
casesPer1M
100.00
10.00
1.00
0.10
0.01
Feb 01
Feb 15
Mar 01
Mar 15
Date
country
Austria
Czech Republic
France
Germany
China
Iceland
Iran
Italy
Poland
Slovakia
South Korea
US

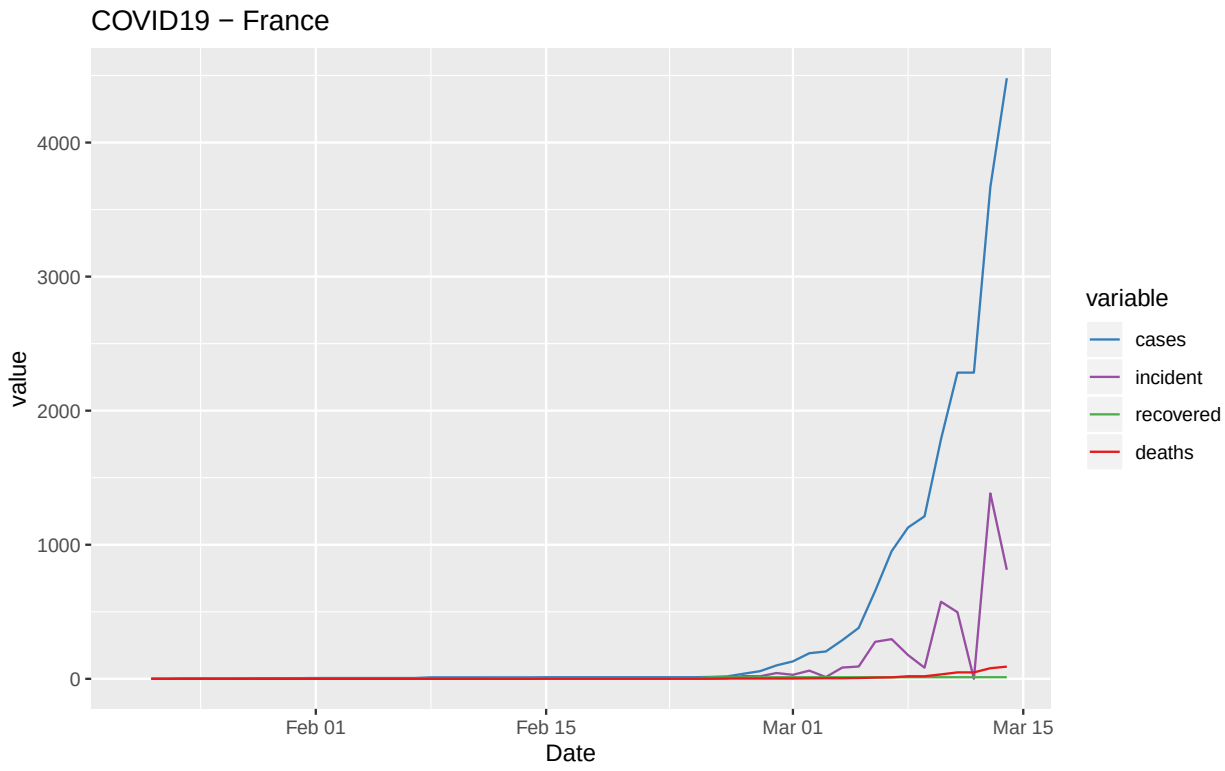
Cummulative stats for Individual selected countries

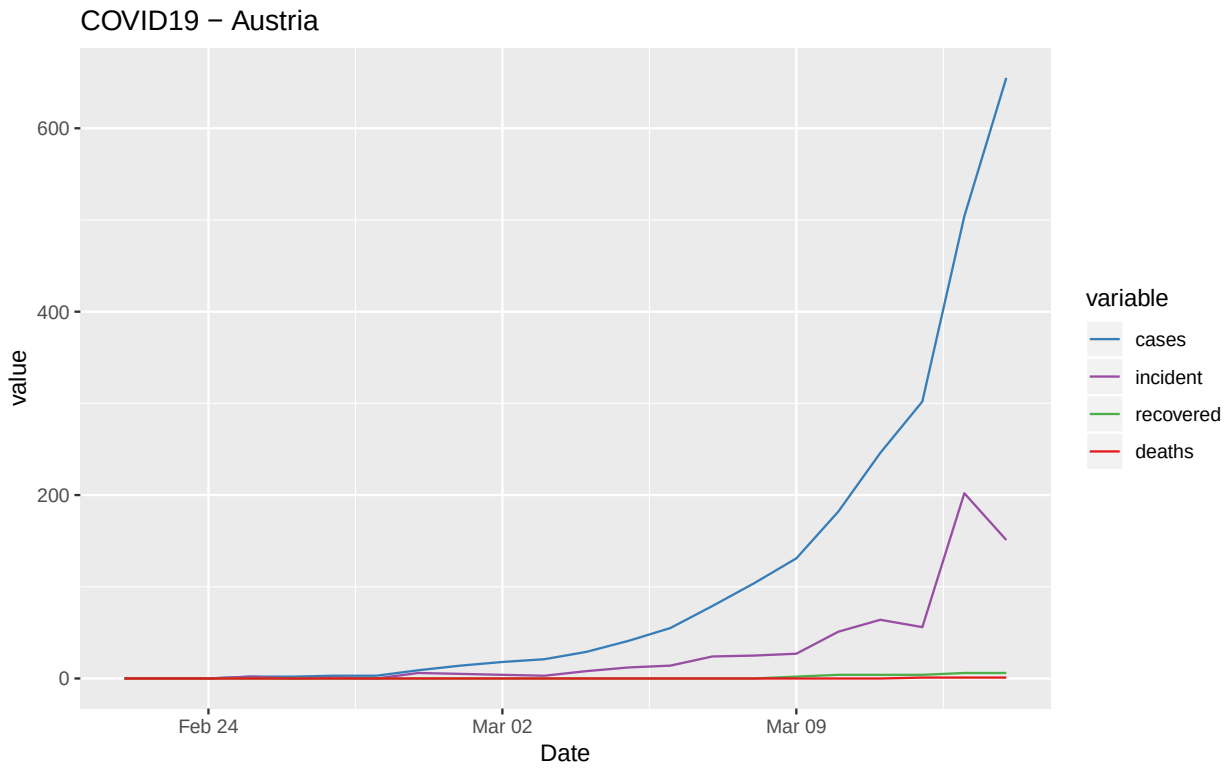


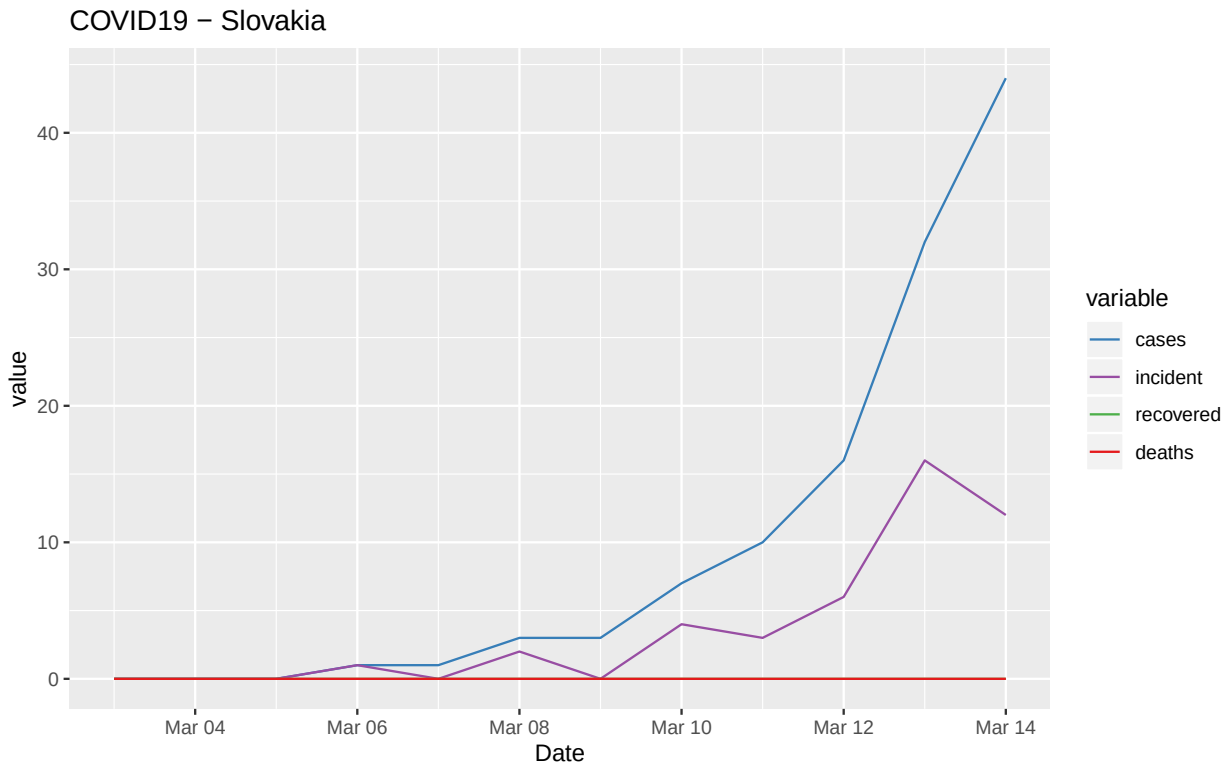


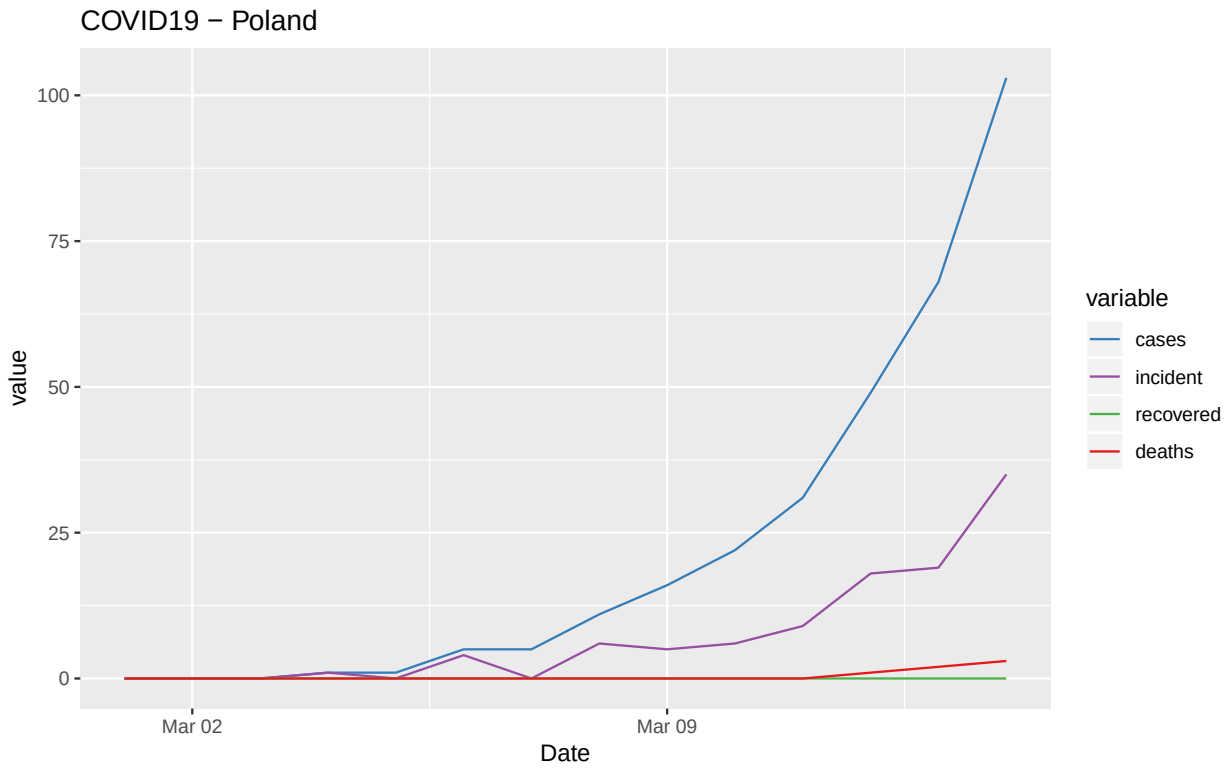

COVID19 − Iran
value
10000
5000
0
Feb 17
Feb 24
Mar 02
Mar 09
Date
variable
cases
incident
recovered
deaths

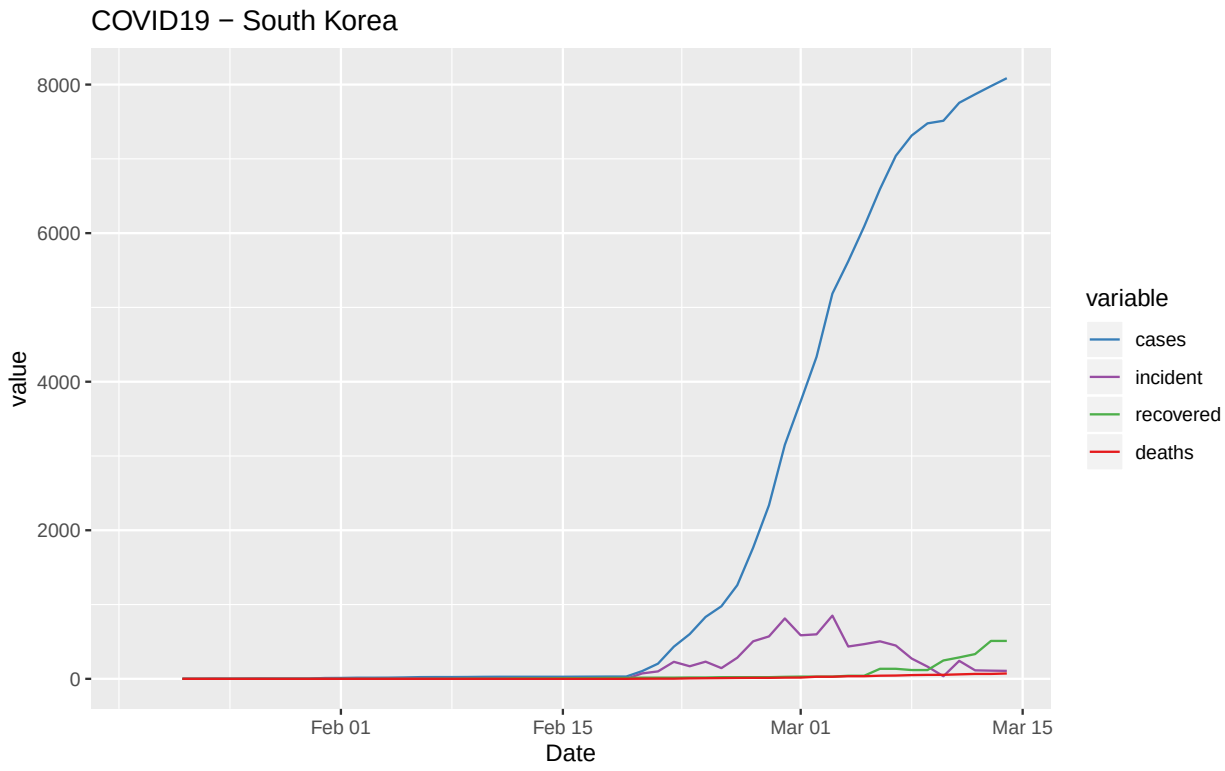

COVID19 − Germany
value
4000
3000
2000
1000
0
Feb 01
Feb 15
Mar 01
Mar 15
Date
variable
cases
incident
recovered
deaths

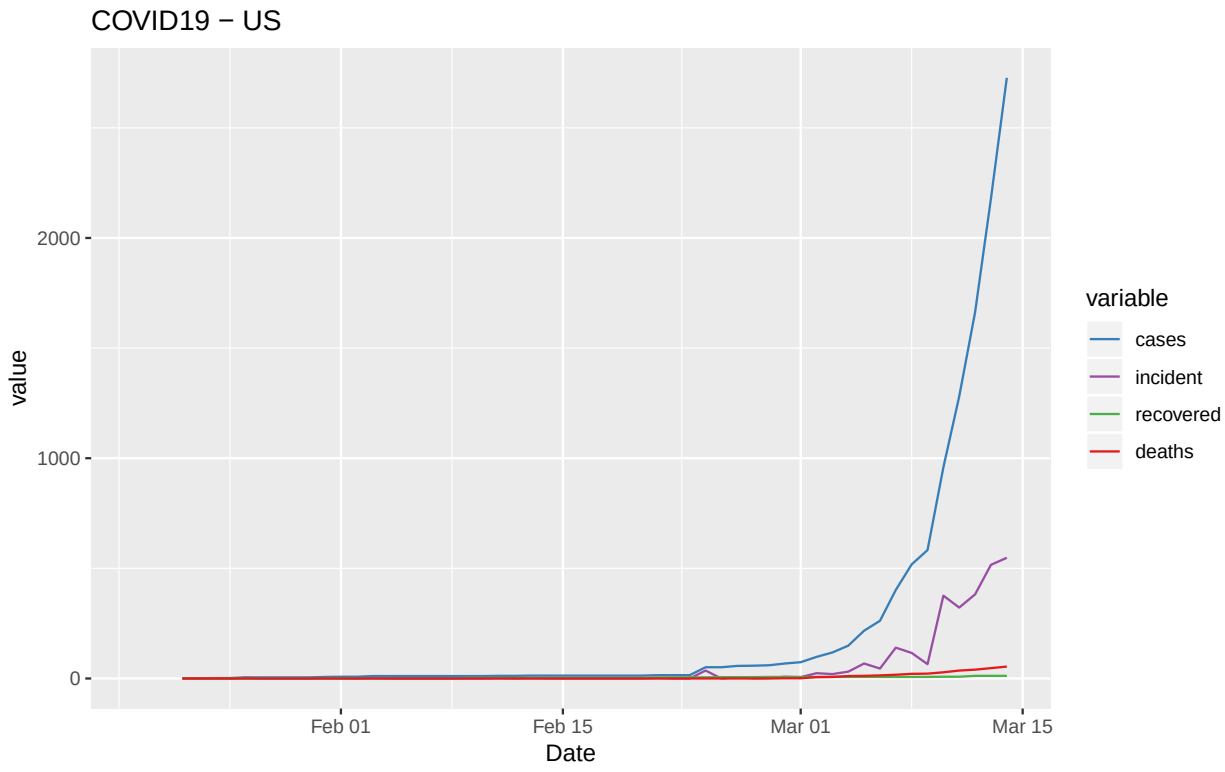

COVID19 − France
4000
3000
2000
1000
0
value
Feb 01
Feb 15
Mar 01
Mar 15
Date
variable
cases
incident
recovered
deaths

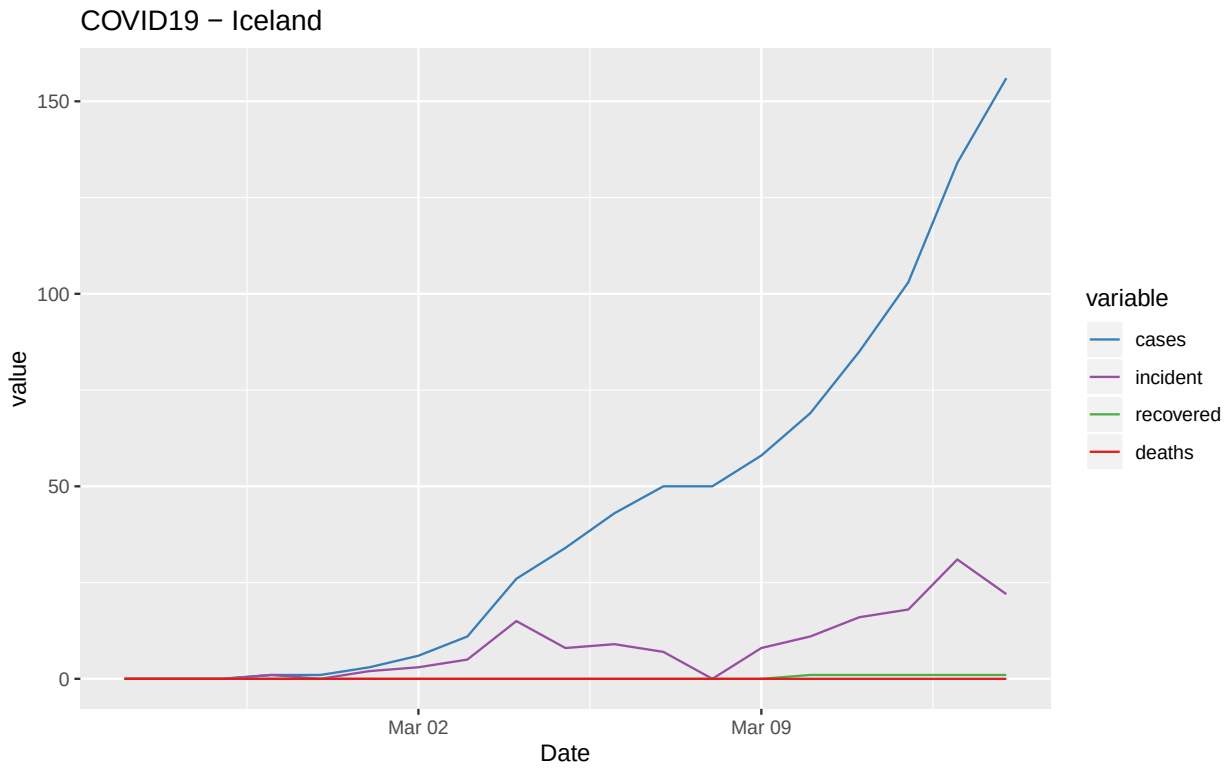

COVID19 − Austria
600
400
200
0
value
Feb 24
Mar 02
Mar 09
Date
variable
cases
incident
recovered
deaths

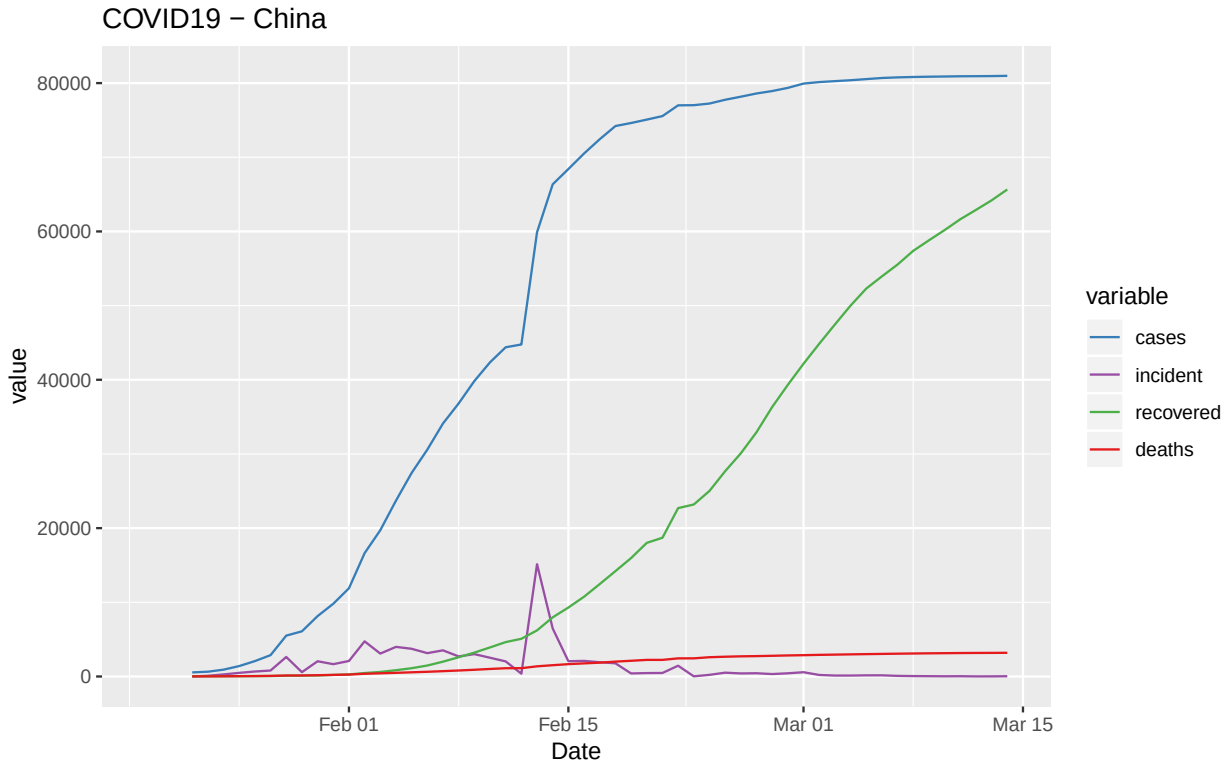

COVID19 − Slovakia
value
40
30
20
10
0
Mar 04
Mar 06
Mar 08
Mar 10
Mar 12
Mar 14
Date
variable
cases
incident
recovered
deaths


COVID19 − Poland
value
100
75
50
25
0
Mar 02
Mar 09
Date
variable
cases
incident
recovered
deaths


COVID19 − South Korea
8000
6000
4000
2000
0
value
Feb 01
Feb 15
Mar 01
Mar 15
Date
variable
cases
incident
recovered
deaths

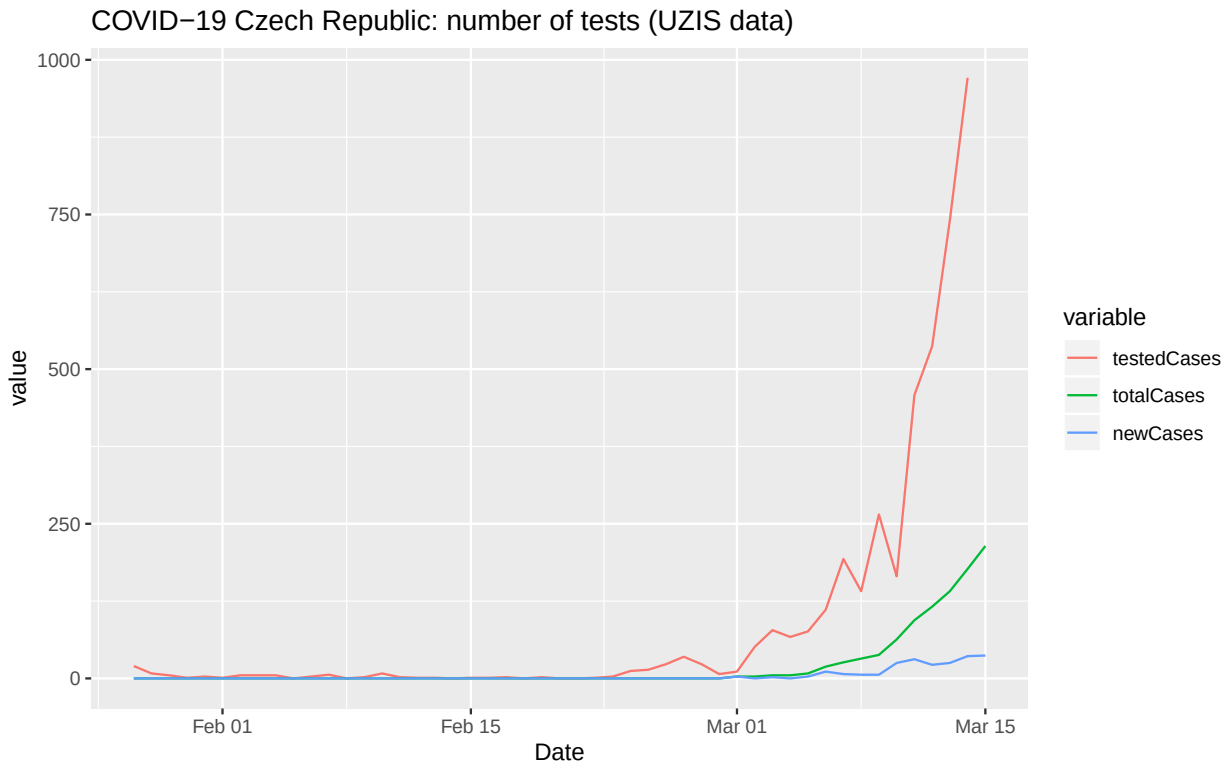

COVID19 − US
2000
1000
0
value
Feb 01
Feb 15
Mar 01
Mar 15
Date
variable
cases
incident
recovered
deaths


COVID19 − Iceland
150
100
50
0
value
Mar 02
Mar 09
Date
variable
cases
incident
recovered
deaths

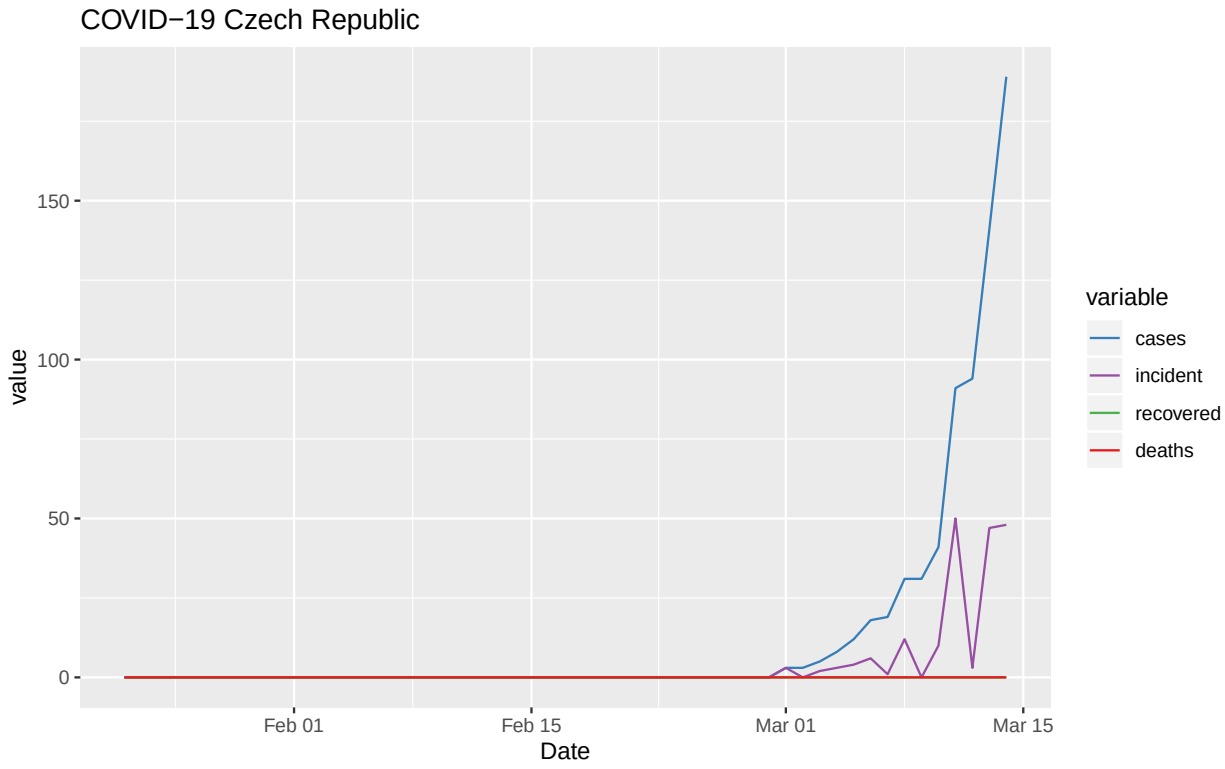

COVID19 − China
80000
60000
40000
20000
0
value
Feb 01
Feb 15
Mar 01
Mar 15
Date
variable
cases
incident
recovered
deaths

Plot and analyze Czech Rep.

exploratory plots: UZIS data

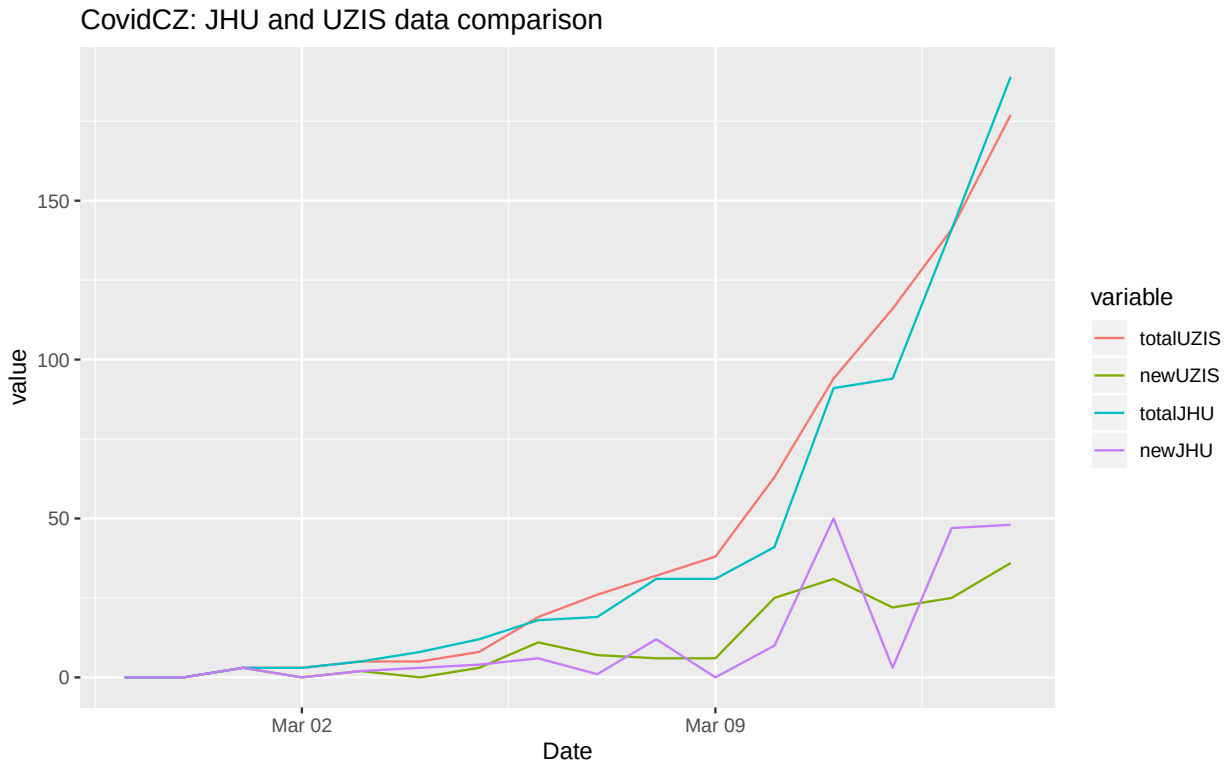


exploratory plots: JHU data



first case in CZ: NA

compare JHU and ÚZIS data



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