Visualisation of evolution of novel coronavirus

Marcelo Avila

17 März, 2020

Contents

load libraries

load data

```
cat_labels <- c("daily cases", "daily deaths",
                "cumulative cases", "cumulative deaths")
df <- df %>% mutate(category = factor(category,
                                labels = cat_labels),
                    continent=countrycode(location,
                                          origin = "country.name",
                                          destination = "continent"))
```

```
## Warning in countrycode(location, origin = "country.name", destination = "continent"): Some values we
```

There is data for 151 countries. The last update was on 16 of March

```
head(df)
```

```
## # A tibble: 6 x 8
##    date       location  category   value cca2  pop2020 value_per_100k_~ continent
##    <date>     <chr>     <fct>      <dbl> <chr>   <dbl>            <dbl> <chr>
## 1 2020-02-25 Afghanis~ daily cas~     0 AF    3.89e7                0 Asia
## 2 2020-02-25 Afghanis~ daily dea~     0 AF    3.89e7                0 Asia
## 3 2020-02-25 Afghanis~ cumulativ~     1 AF    3.89e7          0.00257 Asia
## 4 2020-02-25 Afghanis~ cumulativ~     0 AF    3.89e7                0 Asia
## 5 2020-02-26 Afghanis~ daily cas~     0 AF    3.89e7                0 Asia
## 6 2020-02-26 Afghanis~ daily dea~     0 AF    3.89e7                0 Asia
```

```r
# for thousand separator for plotting
fun_dot <- function(x) format(x, big.mark = " ",
                                 scientific = FALSE,
                                 decimal.mark = ".")

# negative in
`%!in%` <- Negate(`%in%`)

# y scale
log10_minor_break = function (...){
  function(x) {
    minx         = floor(min(log10(x), na.rm=T))-1;
    maxx         = ceiling(max(log10(x), na.rm=T))+1;
    n_major      = maxx-minx+1;
    major_breaks = seq(minx, maxx, by=1)
    minor_breaks =
      rep(log10(seq(1, 9, by=1)), times = n_major)+
      rep(major_breaks, each = 9)
    return(10^(minor_breaks))
  }
}
```

Cumulative Cases for Selected Countries

```r
sel_category <- "cumulative cases"  # new cases, new deaths,
                                    # cumulative cases, cumulative deaths
selected_countries <- c("Germany",
                        "Italy",
                        "United Kingdom",
                        "Brazil",
                        "China",
                        "Spain",
                        "South Korea",
                        "United States")

df %>%
  # wrangle
  filter(location %in% selected_countries, category==sel_category) %>%
  mutate(
    location=fct_reorder2(location, date, value),
    label = ifelse(date == "2020-02-29", #max(date),
                      as.character(location),
                      NA_character_)) %>%

  # plot
  ggplot(aes(x=date, y=value, color=location)) +
  geom_point(aes(shape="o")) +
  geom_line(aes(linetype=location)) +

  # labels
  ggrepel::geom_label_repel(aes(label = label),
```
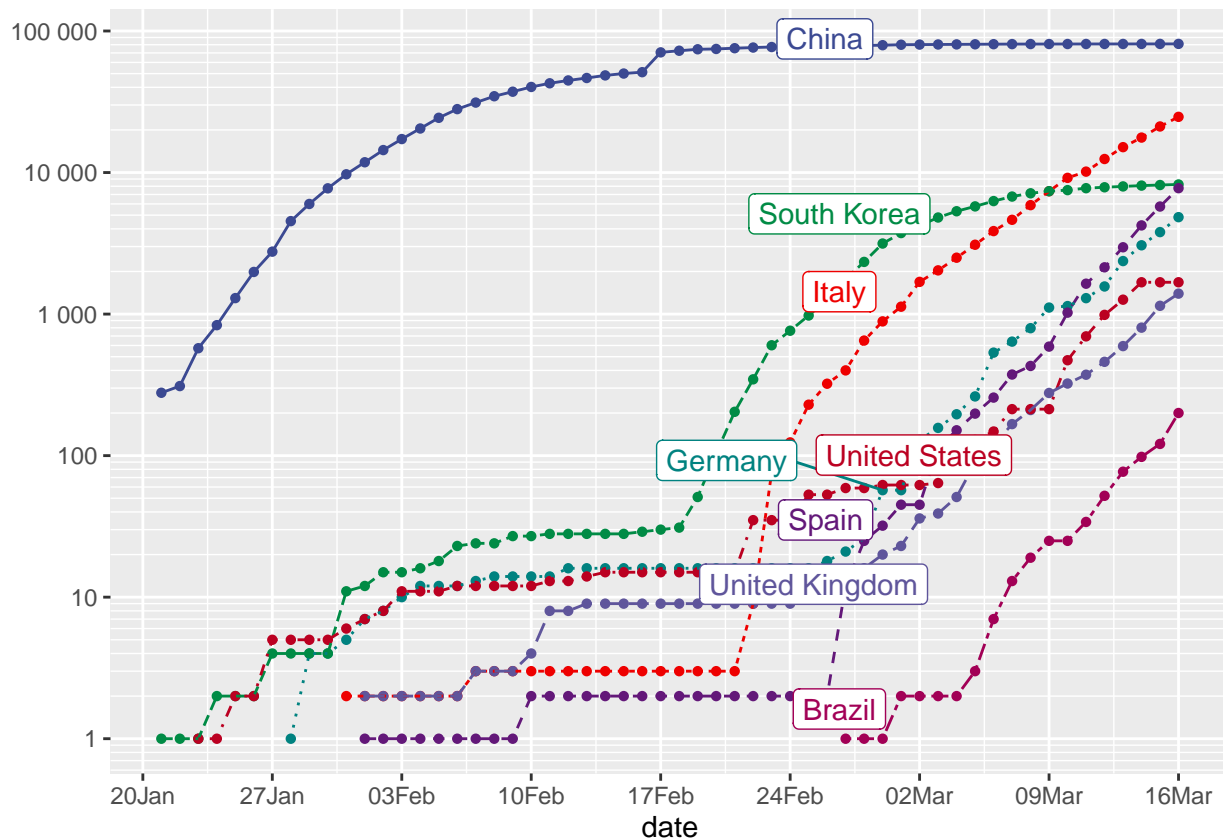
```
                              nudge_x = -1, nudge_y = 1/10,
                              na.rm = TRUE) +
# legend and scales
scale_y_log10(name="(log scale)", labels=fun_dot,
              breaks=10^(0:9),
              minor_breaks=log10_minor_break()) +
scale_x_date(date_breaks = "1 week",
             date_labels="%d%b") +
ggsci::scale_color_aaas() +
theme(legend.position = "none",
      axis.title.y = element_blank())
```


Normalized comparison

```
df$continent %>% unique()
```

```
## [1] "Asia"     "Europe"   "Africa"   "Americas" "Oceania"  NA
```

```
df$category %>% unique() %>% as.character()
```

```
## [1] "daily cases"      "daily deaths"       "cumulative cases"
## [4] "cumulative deaths"
```

```r
make_plot <- function(dta, sel_category="cumulative cases",
                      sel_continent="Europe",
                      threshold=100) {

  # extract singular form of category
  singular_category <- scan(text = sel_category, what = "", quiet=TRUE)[2] %>%
    substr(1, nchar(.)-1)

  sel_continent_title <- ifelse(length(sel_continent)>1,
                                yes = "selected continents",
                                no = sel_continent)
  dta %>%
    filter(category == sel_category,
           continent %in% sel_continent) %>%
    group_by(location) %>%
    filter(value>=threshold) %>%
    mutate(date_shift=0:(n()-1)) %>% # generate date shifted after threshold
    ungroup(location) %>%
    arrange(date_shift) %>%
    # mutate for legend ordering and ggrepel postioning
    mutate(location=fct_reorder2(location, date_shift, value),
           # for ggrepel
           label = ifelse(date == max(date),
                          as.character(location),
                          NA_character_)) %>%
    # plot
    ggplot(aes(x=date_shift, y=value, colour=location)) +
    # geoms
    geom_point() +
    geom_line(aes(linetype=location)) +

    # ggrepel
    ggrepel::geom_label_repel(
      aes(label = label), nudge_x = 1, na.rm = TRUE) +

    # legend and scales
    scale_y_log10(
      name = paste(sel_category, "(log scale)"),
      labels=fun_dot, breaks = 10^(0:9), minor_breaks=log10_minor_break()) +
    scale_x_continuous(
      name = paste0("Days past since ", threshold, "th ", singular_category),
      breaks = seq(0, 1e5, 7), minor_breaks = 1:1e3) +
    scale_color_grey(start = 0, end = .3) +

    theme(legend.position = "none", axis.title.y = element_blank()) +
    #title
    ggtitle(paste("Evolution of novel coronavirus in", sel_continent_title),
            subtitle = paste(sel_category, "(log scale)"))
}


categories <- df$category %>% unique() %>% as.character() %>% purrr::set_names()
continents <- df$continent %>% unique() %>% as.character() %>% purrr::set_names()
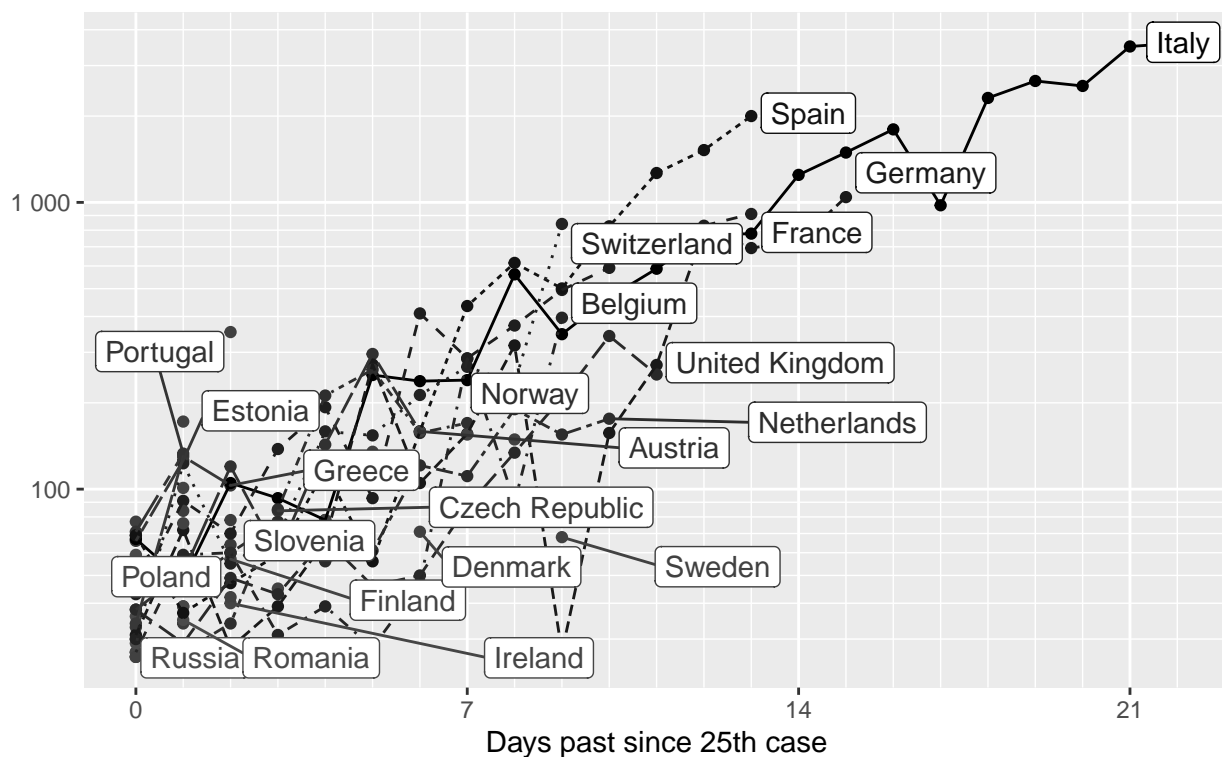```

```
plots_categ <- purrr::map(categories, ~make_plot(df, sel_category = .x,
                                                 threshold = 25))
plots_conts <- purrr::map(continents, ~make_plot(df, sel_continent = .x,
                                                 sel_category = "cumulative deaths",
                                                 threshold = 1))

plots_categ
```

```
## $`daily cases`
```
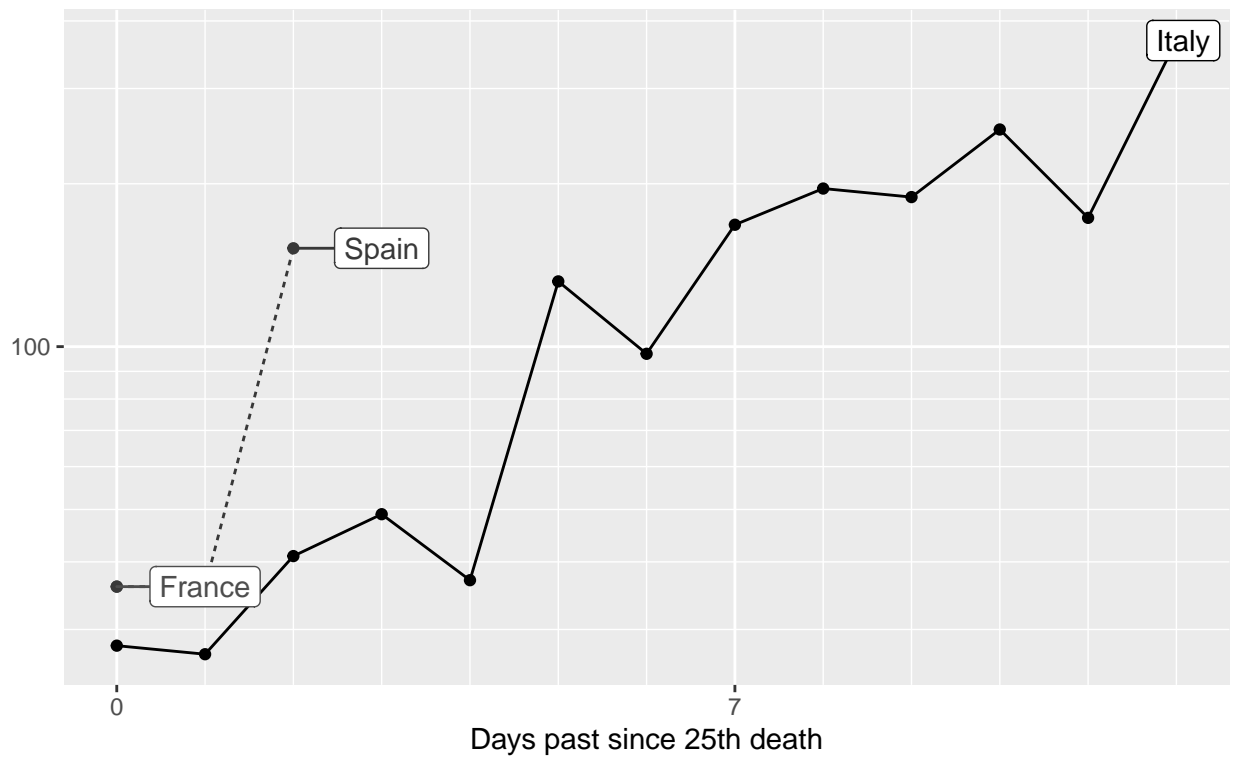
Evolution of novel coronavirus in Europe
daily cases (log scale)


```
##
## $`daily deaths`
```

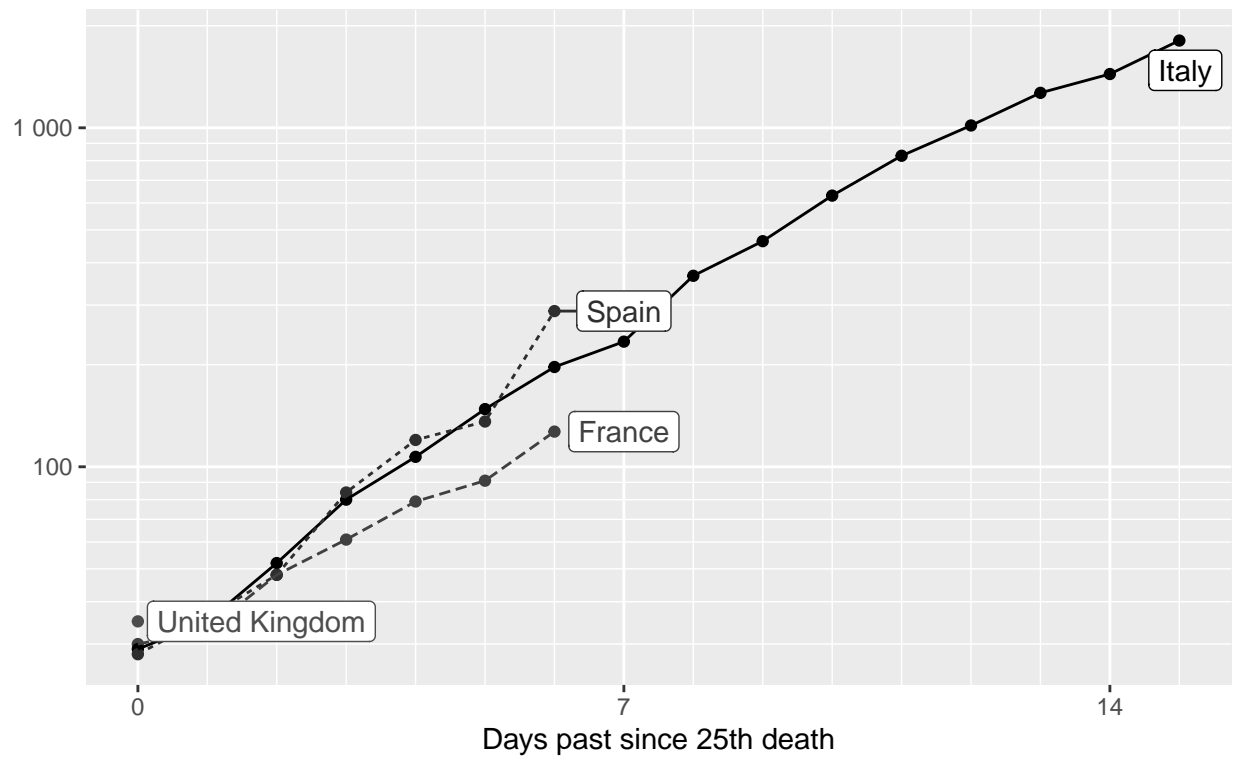## Evolution of novel coronavirus in Europe
daily deaths (log scale)


```
##
## $`cumulative cases`
```

Evolution of novel coronavirus in Europe
cumulative cases (log scale)
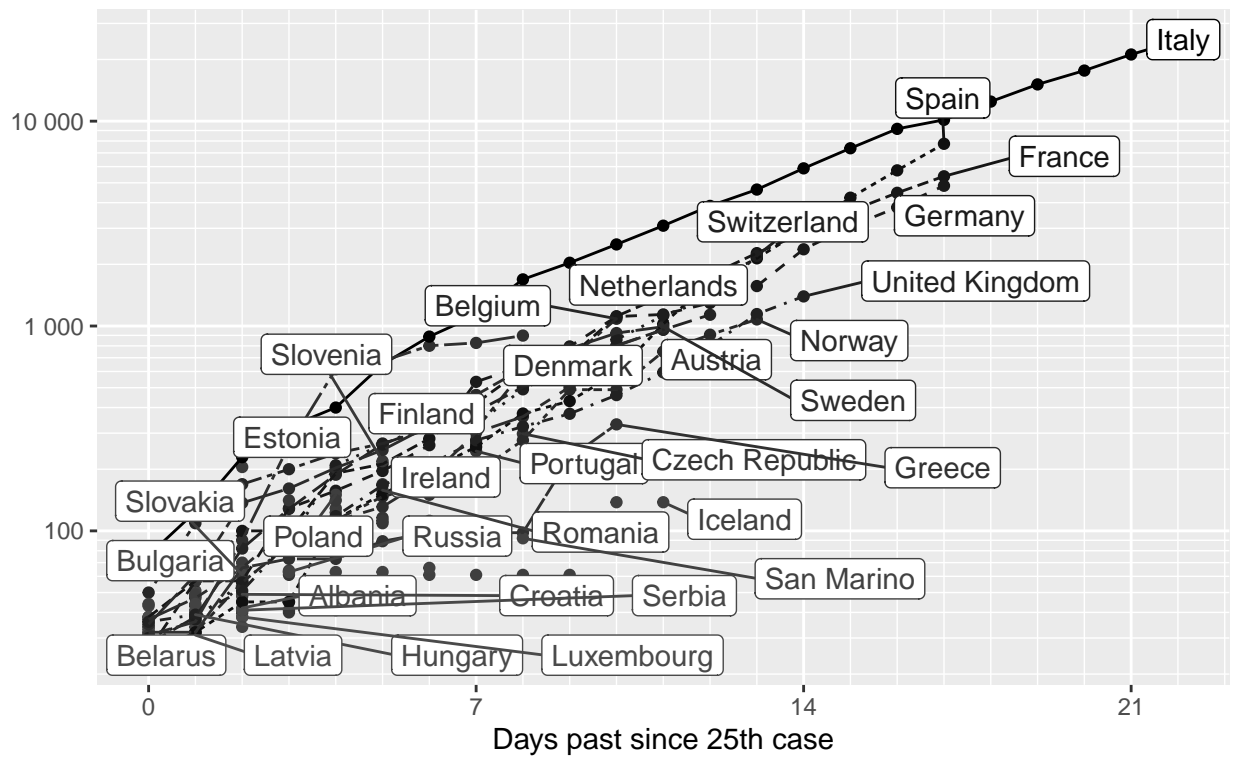

```
##
## $`cumulative deaths`
```

Evolution of novel coronavirus in Europe
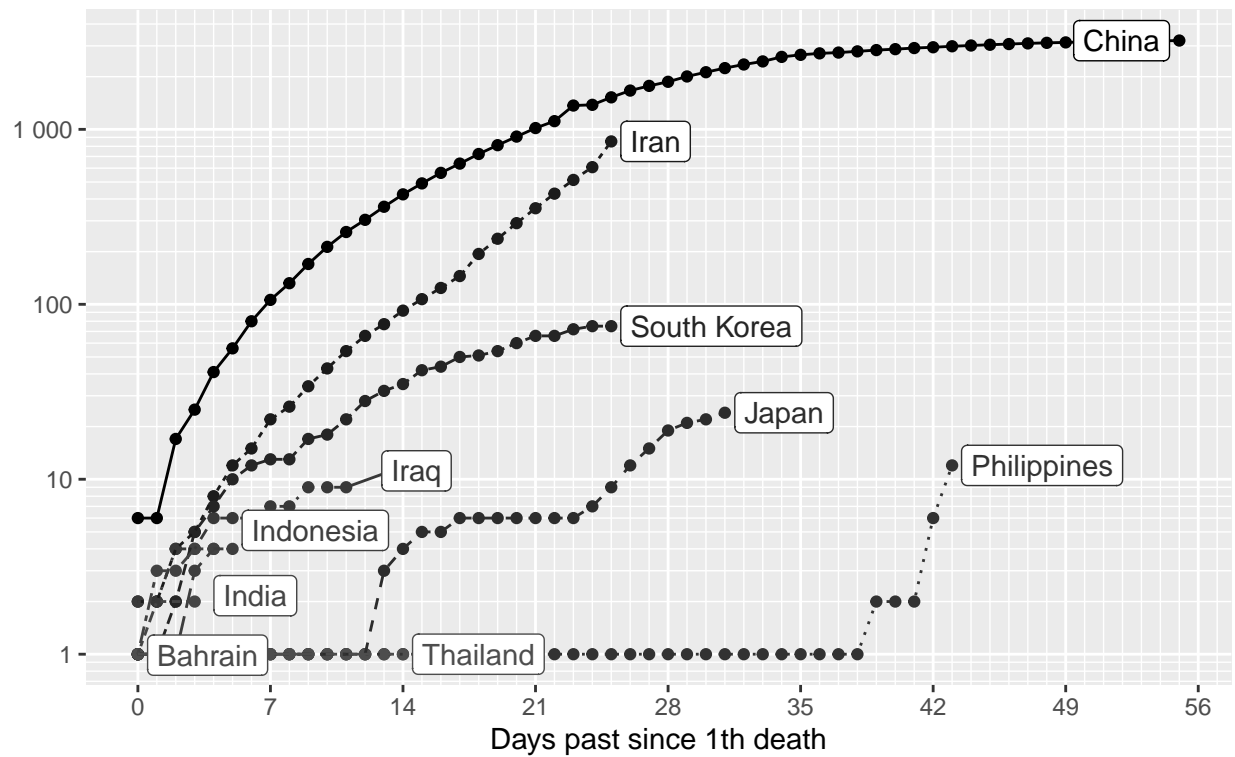cumulative deaths (log scale)

Days past since 25th death

```
plots_conts
```

```
## $Asia
```

Evolution of novel coronavirus in Asia
cumulative deaths (log scale)


```
## 
## $Europe
```

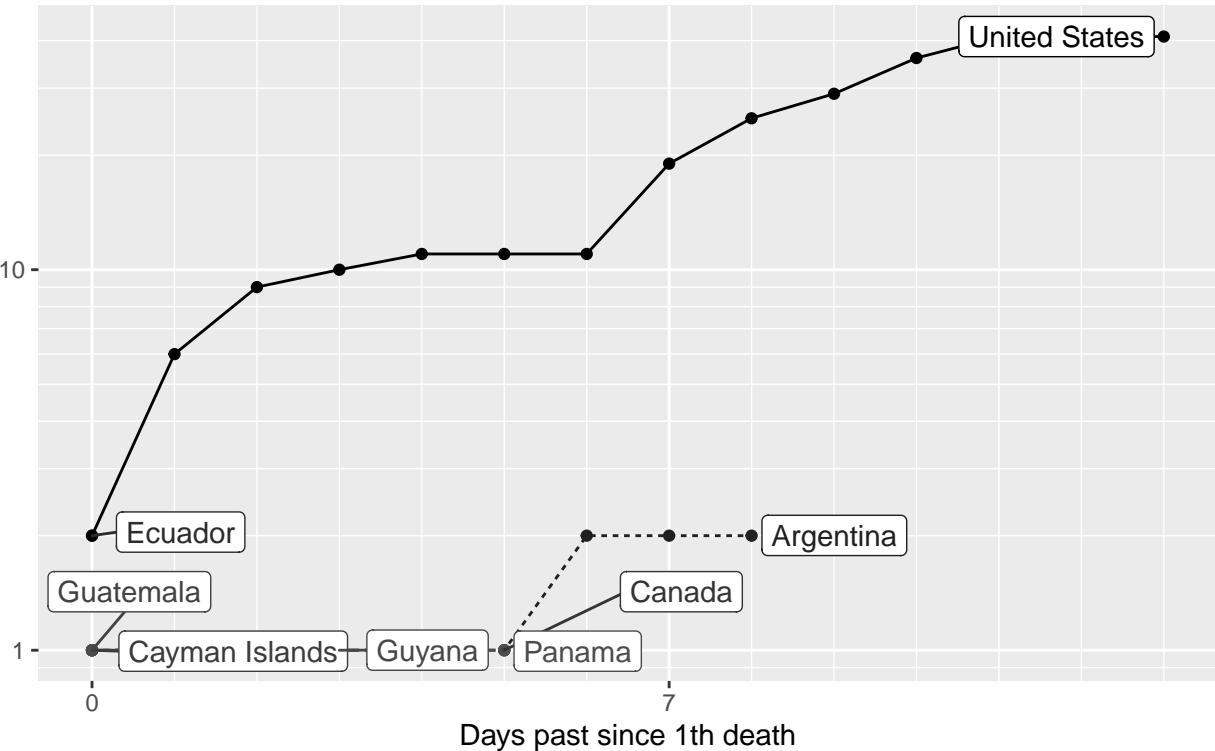## Evolution of novel coronavirus in Europe
cumulative deaths (log scale)

Days past since 1th death
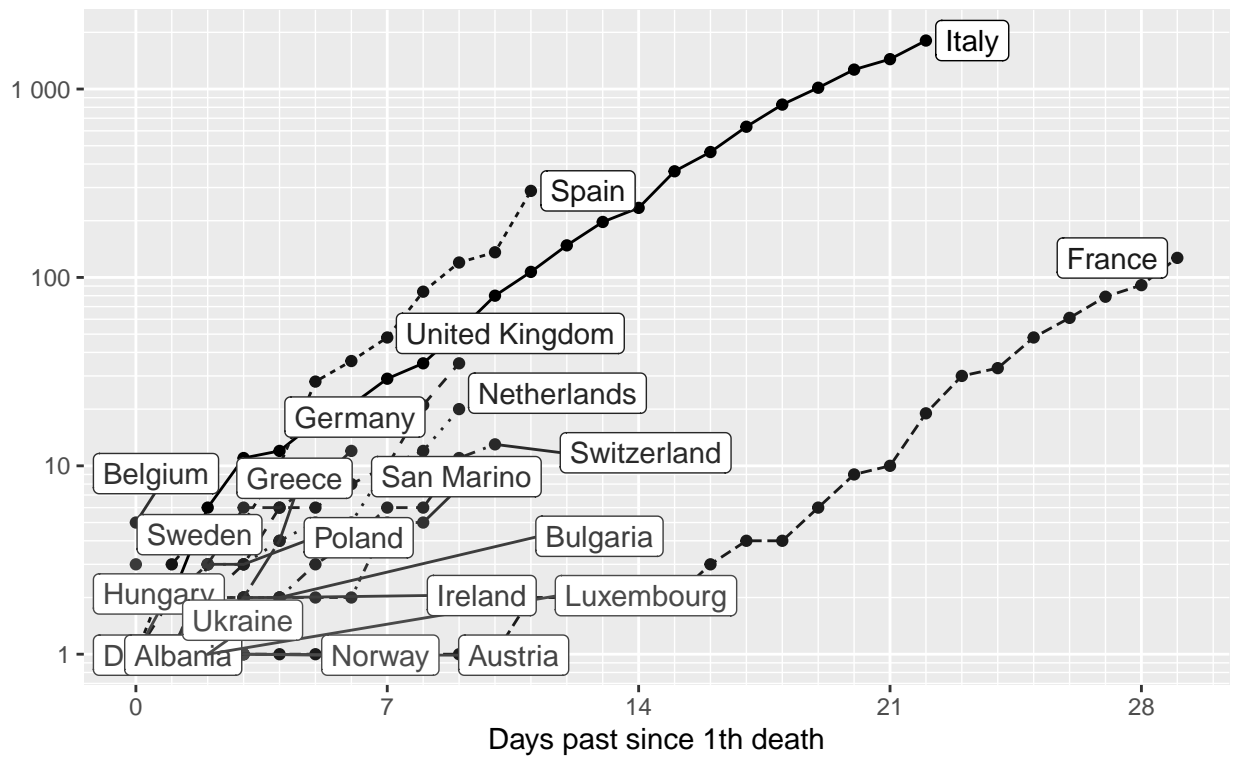
```
## 
## $Africa
```

Evolution of novel coronavirus in Africa

cumulative deaths (log scale)

Days past since 1th death
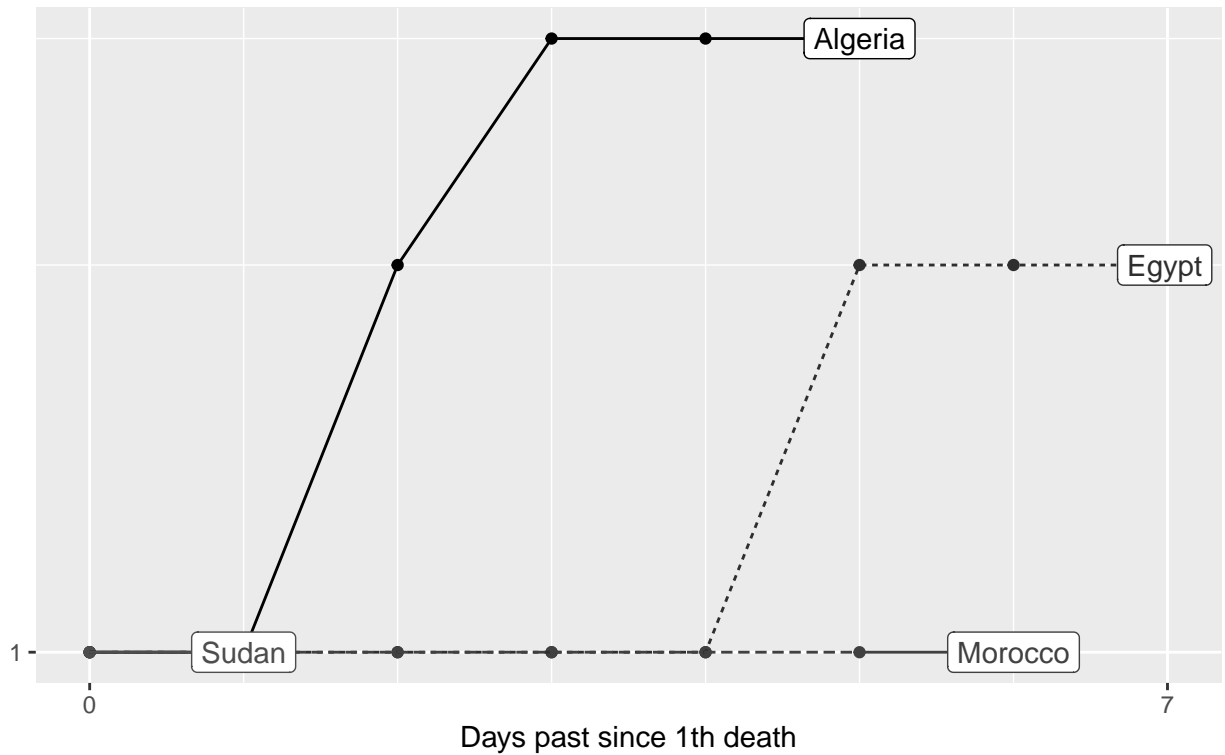
```
##
## $Americas
```

Evolution of novel coronavirus in Americas
cumulative deaths (log scale)

United States

10 -

Ecuador

Guatemala

Argentina

Canada

1 - Cayman Islands Guyana Panama

0   7

Days past since 1th death

```
##
## $Oceania
```

Evolution of novel coronavirus in Oceania
cumulative deaths (log scale)


```
##
## $<NA>
```
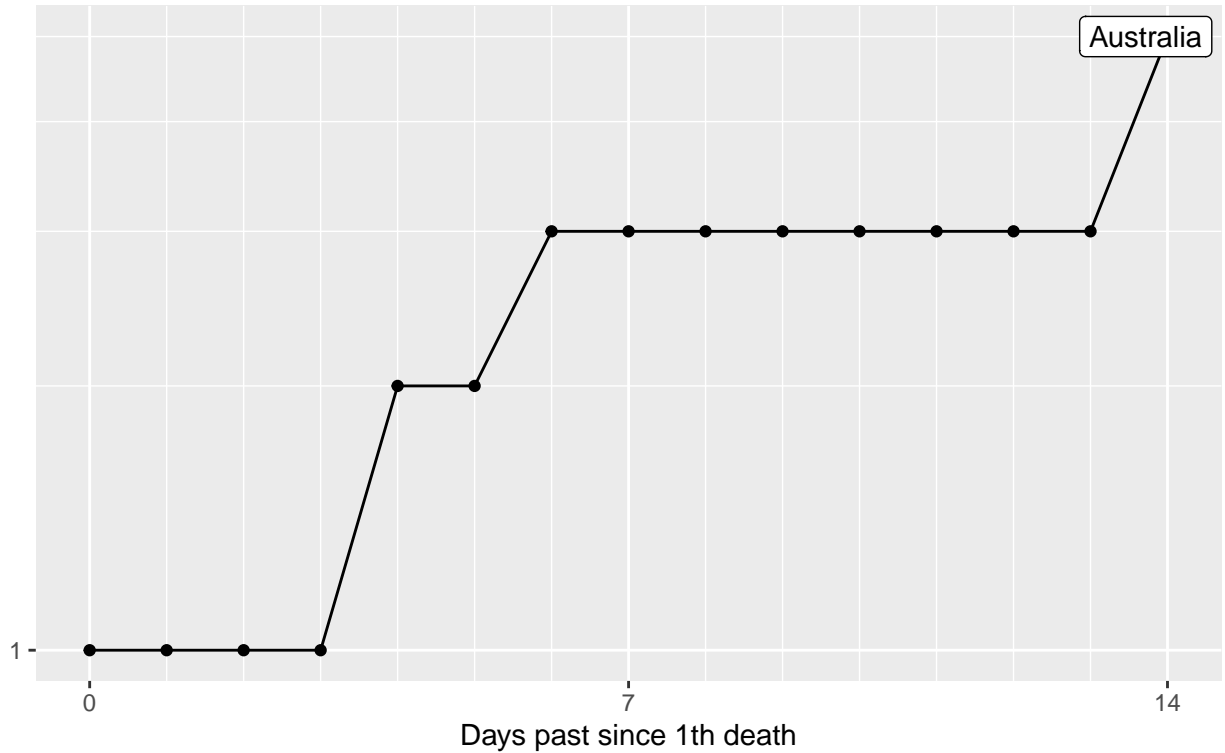
Evolution of novel coronavirus in NA
cumulative deaths (log scale)

Days past since 1th death