covi19 ggplot2

Norberto Asensio

24/03/2020

Get COVI19 world data

Script detail to download data in https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-geographic-disbtribution-worldwide- (https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-geographic-disbtribution-worldwide-).

```
#these libraries are necessary
library(readxl)
library(httr)
#create the URL where the dataset is stored with automatic updates every day
url <- paste("https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-ge
ographic-disbtribution-worldwide-",format(Sys.time(), "%Y-%m-%d"), ".xlsx", sep =
"")
#download the dataset from the website to a local temporary file
GET(url, authenticate(":", ":", type="ntlm"), write_disk(tf <- tempfile(fileext =
".xlsx")))
```

```
## Response [https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-geo
graphic-disbtribution-worldwide-2020-03-24.xlsx]
##   Date: 2020-03-24 13:33
##   Status: 200
##   Content-Type: application/vnd.openxmlformats-officedocument.spreadsheetml.she
et
##   Size: 254 kB
## <ON DISK>  /var/folders/08/4_zvl9fd5259_lw8c12d7vzm0000gn/T//RtmpV3xgmZ/file178
de32bbd52d.xlsx
```

```
#read the Dataset sheet into “R”
data <- read_excel(tf)
#read the Dataset sheet into “R”
data <- read_excel(tf)
str(data)
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame':    6551 obs. of  8 variables:
##  $ DateRep                  : POSIXct, format: "2020-03-24" "2020-03-23" ...
##  $ Day                      : num  24 23 22 21 20 19 18 17 16 15 ...
##  $ Month                    : num  3 3 3 3 3 3 3 3 3 3 ...
##  $ Year                     : num  2020 2020 2020 2020 2020 2020 2020 2020 2020
2020 ...
##  $ Cases                    : num  6 10 0 2 0 0 1 5 6 3 ...
##  $ Deaths                   : num  1 0 0 0 0 0 0 0 0 0 ...
##  $ Countries and territories: chr  "Afghanistan" "Afghanistan" "Afghanistan" "A
fghanistan" ...
##  $ GeoId                    : chr  "AF" "AF" "AF" "AF" ...
```

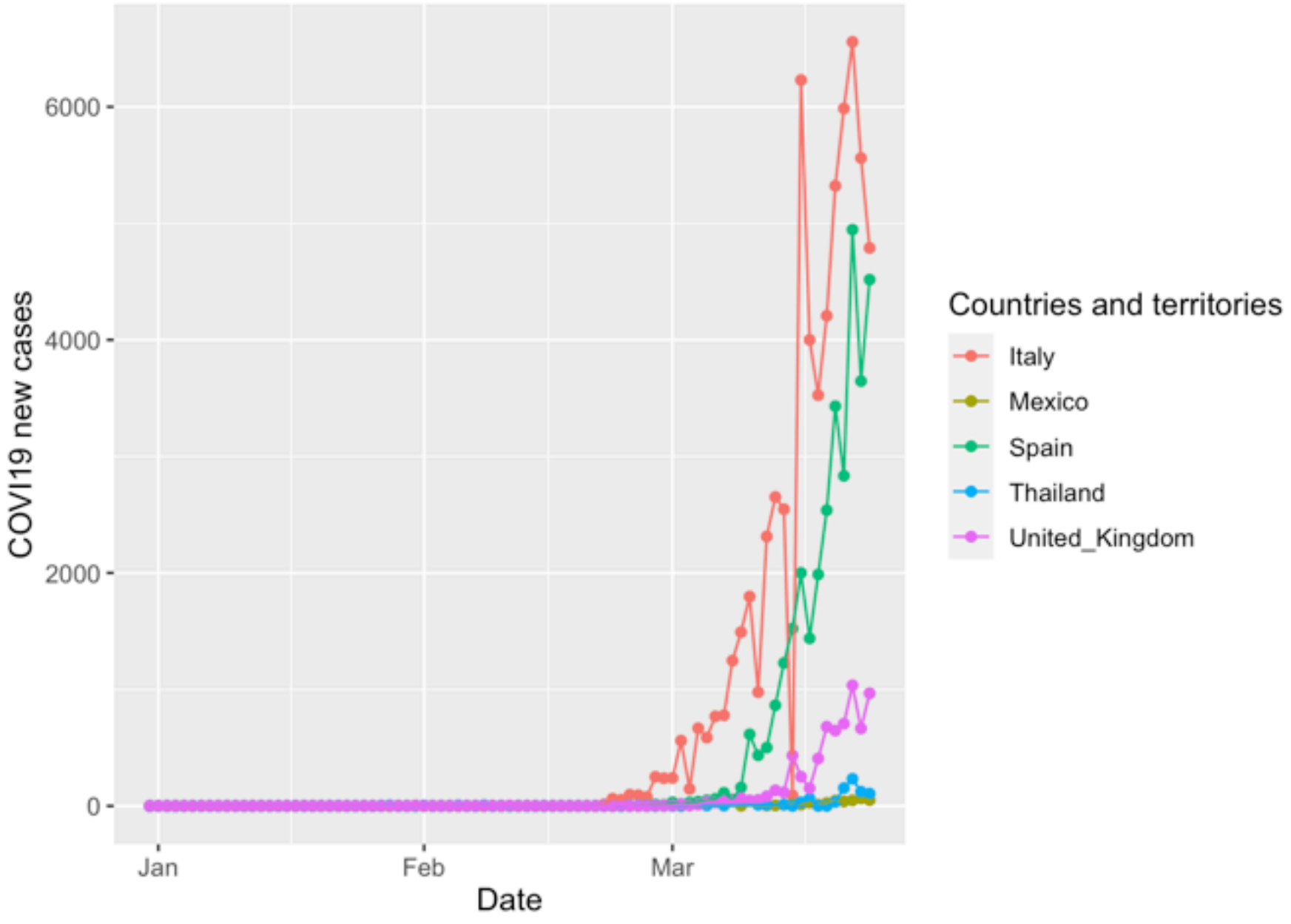
Subsample a database with countries of interest

```
severaldata <- data[which(data$`Countries and territories`=='Spain' |
                            data$`Countries and territories`=='Italy'|
                            data$`Countries and territories`=='Mexico'|
                            data$`Countries and territories`=='Thailand'|
                            data$`Countries and territories`=='United_Kingdom'), ]
table(severaldata$`Countries and territories`)#available days with data
```

```
##
##          Italy         Mexico          Spain       Thailand United_Kingdom
##             85             77             85             78             85
```

Scatterplots with daily new cases

```
library(ggplot2)
ggplot(severaldata, aes(x=DateRep, y=Cases, color=`Countries and territories`)) +
  geom_point()+geom_line()+
  xlab("Date") + ylab("COVI19 new cases") +
  theme(plot.title = element_text(lineheight=.8, face="bold", size = 20)) +theme(t
ext = element_text(size=12))
```


COVI19 new cases
6000
4000
2000
0
Jan
Feb
Mar
Date
Countries and territories
Italy
Mexico
Spain
Thailand
United_Kingdom