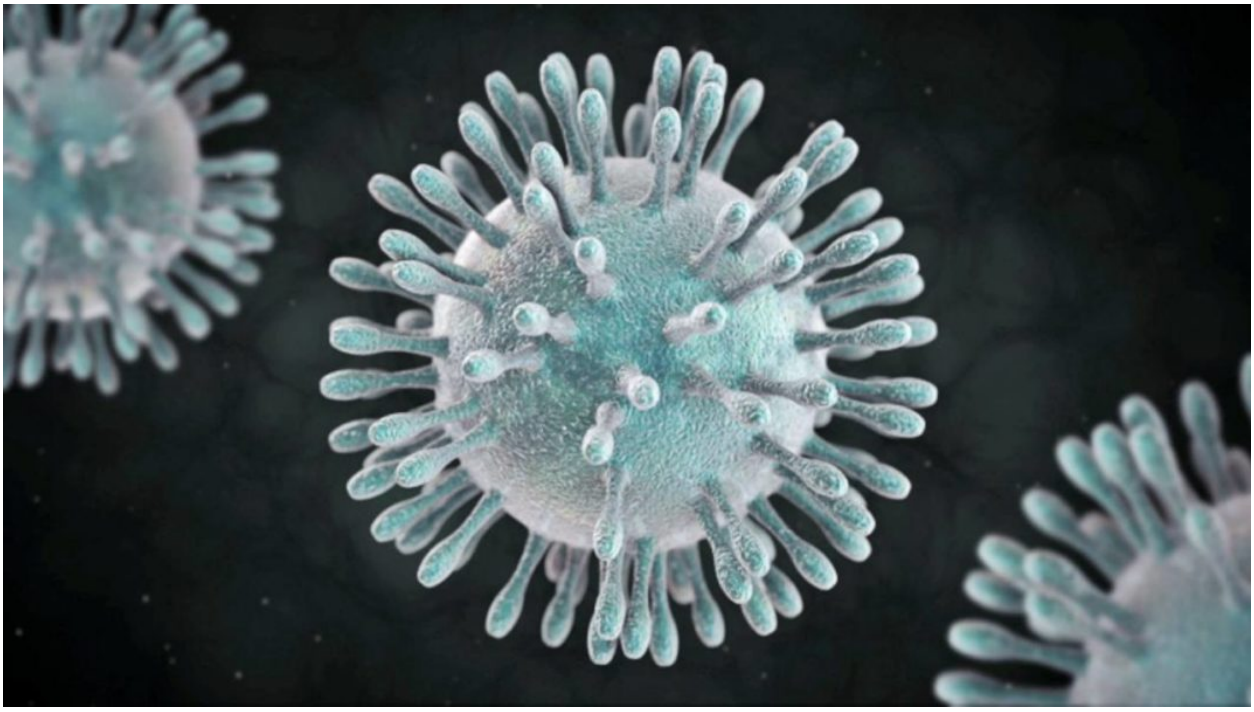
COVID-19 'Flattening The Curve' Graph

Hrvoje Krpan



Description

- In this workbook, I will show how to (re)create a graph that was presented by a Croatian health minister **Vili Beroš** by using ggplot package
- I saw graph originally at Slobodna Dalmacija portal. Link to original article.
- Each point on a red line is representing an average of 5 previous days. For instance, the red line over "ožu 30" (Mar 30 in Croatian) is the mean of cases that happened between 26.- 30. March.
- With a little adjustment of moving average, this plot can be used in salse purposes as well.

Install Packages

```
#devtools::install_github("covid19r/coronavirus")
#install.packages("tidyverse")
#install.packages("RcppRoll")
```

Load packages

```
library(coronavirus)
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------------
## v ggplot2 3.3.0     v purrr   0.3.3
## v tibble  3.0.0     v dplyr   0.8.5
## v tidyr   1.0.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
## -- Conflicts ------------------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(RcppRoll)
theme_set(theme_bw())
```

Getting the data

```
data("coronavirus")
tail(coronavirus)
```

```
##       Province.State Country.Region     Lat     Long       date cases      type
## 61220       Zhejiang          China 29.1832 120.0934 2020-04-04     1 recovered
## 61221       Zhejiang          China 29.1832 120.0934 2020-04-05     1 recovered
## 61222       Zhejiang          China 29.1832 120.0934 2020-04-06     0 recovered
## 61223       Zhejiang          China 29.1832 120.0934 2020-04-07     0 recovered
## 61224       Zhejiang          China 29.1832 120.0934 2020-04-08     2 recovered
## 61225       Zhejiang          China 29.1832 120.0934 2020-04-09     3 recovered
```

```
cro_covid19_basic <- coronavirus %>%
  filter(Country.Region == "Croatia")

tail(cro_covid19_basic)
```

```
##     Province.State Country.Region  Lat Long       date cases      type
## 232                       Croatia 45.1 15.2 2020-04-04    27 recovered
## 233                       Croatia 45.1 15.2 2020-04-05     6 recovered
## 234                       Croatia 45.1 15.2 2020-04-06     5 recovered
## 235                       Croatia 45.1 15.2 2020-04-07    37 recovered
## 236                       Croatia 45.1 15.2 2020-04-08    12 recovered
## 237                       Croatia 45.1 15.2 2020-04-09    40 recovered
```

```
cro_covid19 <- coronavirus %>%
    select(Country.Region, cases, type, date) %>%
    group_by(type) %>%
    filter(Country.Region=="Croatia") %>%
    pivot_wider(names_from = type, values_from = cases) %>%
    rename(country=Country.Region) %>%
    mutate(rollavg = roll_meanr(confirmed, n = 5)) %>%
    filter(confirmed != 0) %>%
    arrange(desc(date))
```

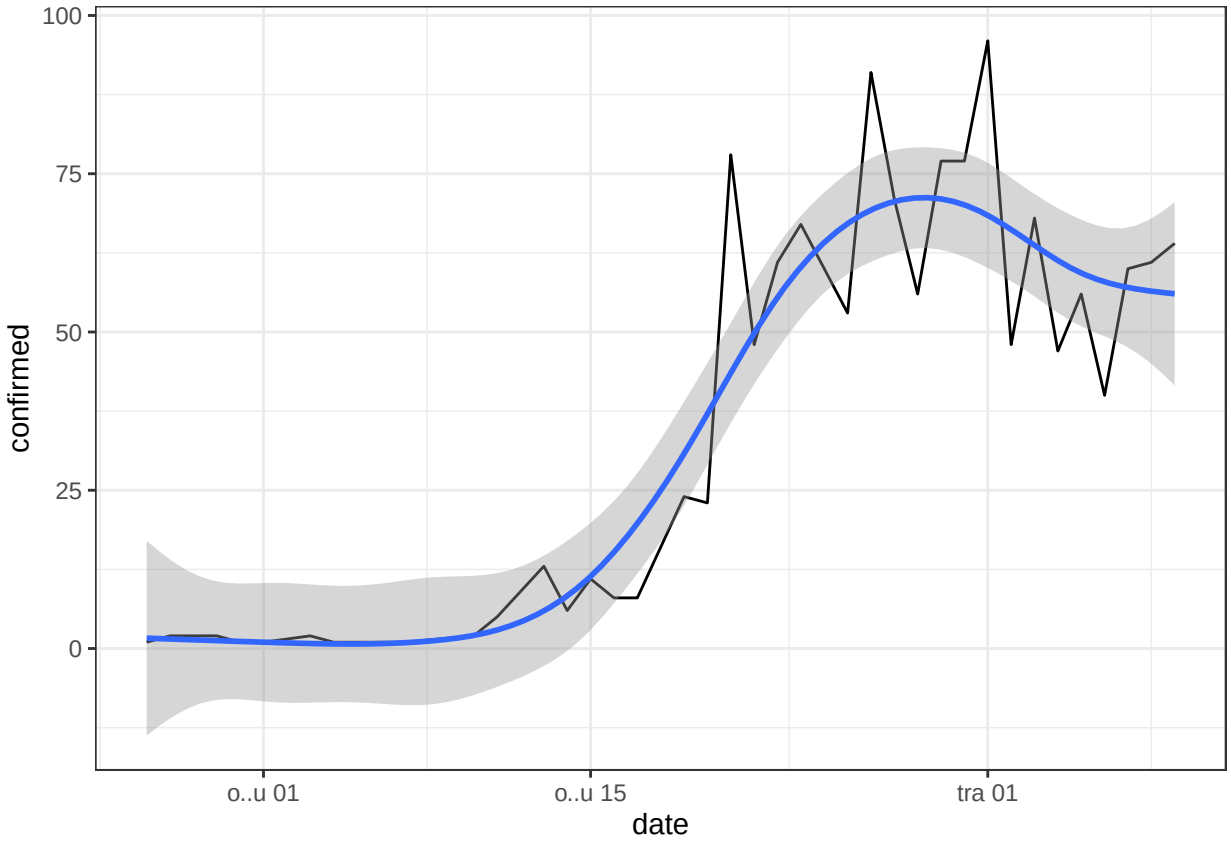
```
head(cro_covid19)
```

```
## # A tibble: 6 x 6
##   country date       confirmed death recovered rollavg
##   <chr>   <date>         <int> <int>     <int>   <dbl>
## 1 Croatia 2020-04-09        64     1        40    56.2
## 2 Croatia 2020-04-08        61     1        12    52.8
## 3 Croatia 2020-04-07        60     2        37    54.2
## 4 Croatia 2020-04-06        40     1         5    51.8
## 5 Croatia 2020-04-05        56     3         6    63
## 6 Croatia 2020-04-04        47     4        27    67.2
```

Data Visualization

```
cro_covid19 %>%
  ggplot(aes(x = date, y = confirmed)) +
  geom_line() +
  geom_smooth(method = "gam")
```

```
## `geom_smooth()` using formula 'y ~ s(x, bs = "cs")'
```




```
cro_covid19 %>%
  ggplot(aes(x = date, y = confirmed)) +
  geom_bar(stat = "identity",fill = "cornflowerblue", color = "white")+
  geom_line(aes(x = date, y = rollavg),color = "red", size = 1.1) +
  labs(subtitle="Number of confirmed cases per day and moving average of last 5 days",
```

```
      y="Confirmed cases",
      x="Date",
      title="COVID-19 - Confirmed Cases in Croatia",
      caption="Source: Johns Hopkins University Center") +
scale_x_date(date_breaks = '1 week', date_labels = "%b %d")
```