

Praca Domowa 2

Yevhenii Vinichenko

Wykres z internetu

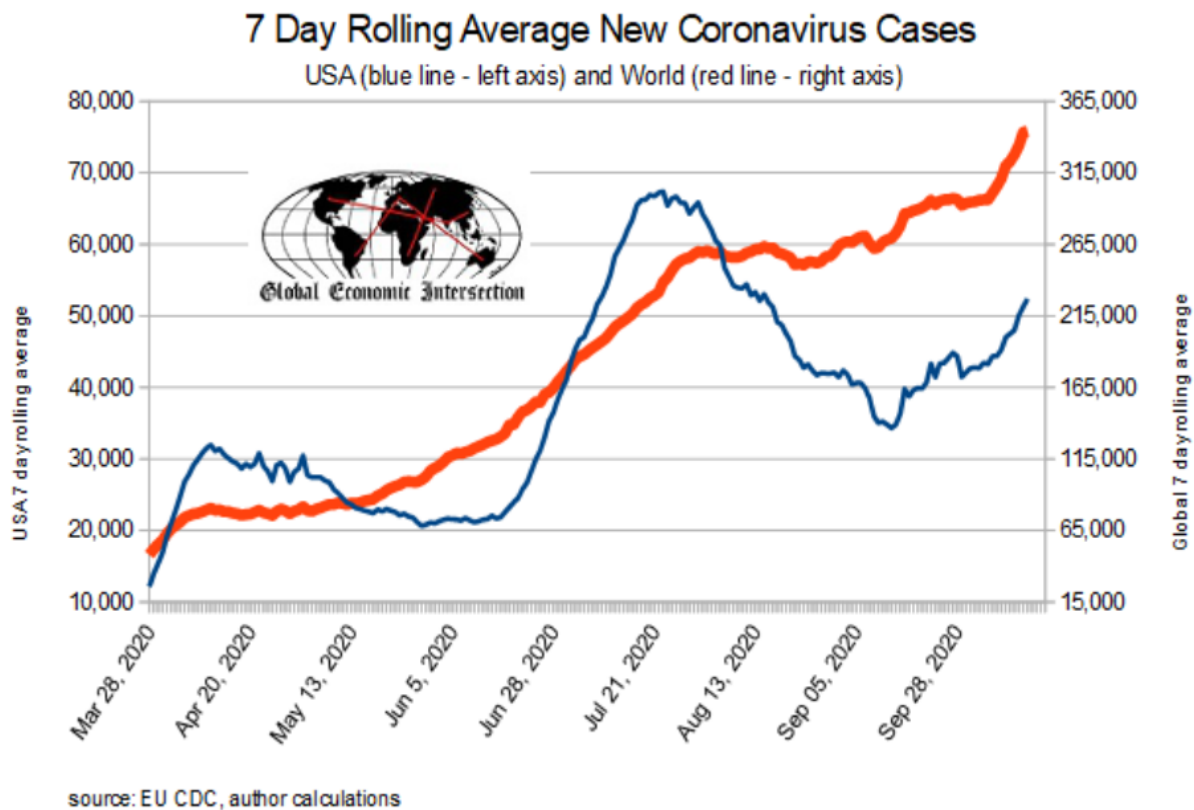


Figure 1: Wykres

Dany wykres powinien reprezentować średnią ruchomą zakażeń koronawirusem w Stanach Zjednoczonych i reszcie świata. Jego wadą, moim zdaniem, jest próba przedstawić dane o dużo różnych skalach na jednym wykresie. Co więcej, wykres nie posiada wyraźnej legendy, przez co nie od razu daje się zrozumieć która linia dotyczy których danych. Ponadto dane ze Stanów nie korelują (czego się można było spodziewać) z danymi z reszty świata, jako że w różnych krajach z koronawirusem walczą po różnemu, a połączenia międzynarodowe raczej nie działają.

Poprawienie wykresu

Kod z pliku źródłowego

```
dates = seq(as.Date("2020-03-28"),as.Date("2020-10-11"), by = "7 days")

x = seq(10000, 30000, by=4000)
a = seq(1000, 4000, by=1)
x = x + sample(a, length(x), replace = TRUE)

y = seq(30000, 20000, by=-2500)
a = seq(1000, 15000, by=1)
y = y + sample(a, length(y), replace = TRUE)

z = seq(20000, 80000, by=5000)
a = seq(5000, 10000, by=1)
z = z + sample(a, length(z), replace = TRUE)

v = seq(80000, 35000, by=-7500)
a = seq(1000, 4500, by=1)
v = v + sample(a, length(v), replace = TRUE)
v = v[-1]
v = v[-length(v)]

usa <- c(x, y, z, v)

tmp <- as.data.frame(list(usa, dates))
colnames(tmp) <- c("Cases", "Date")

world <- seq(65000, 380500, by=11000)

a <- seq(5000, 35000, by=1)
world = world + sample(a, length(world), replace = TRUE)

world <- as.data.frame(list(world, dates))
colnames(world) <- c("Cases", "Date")
```

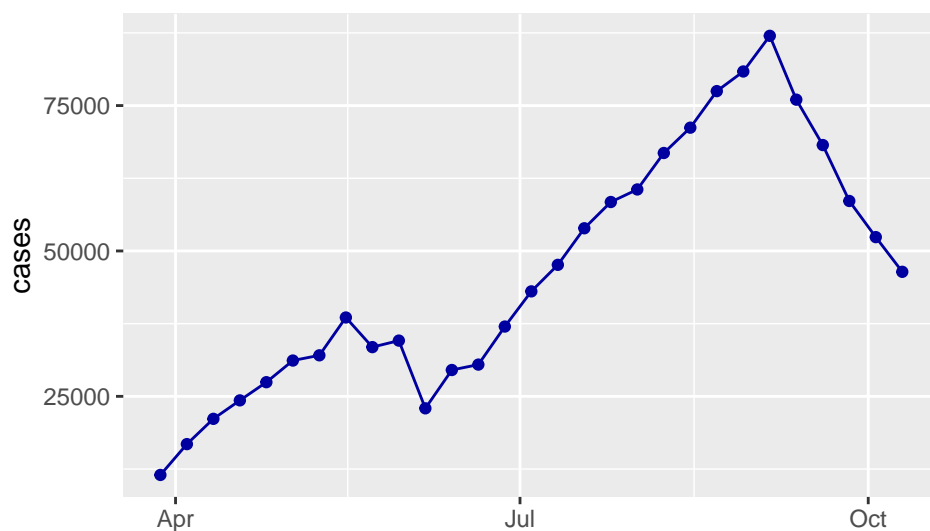
Kod do wykresu nowego

```
library(ggplot2)
library(cowplot)
source("lab2.R")

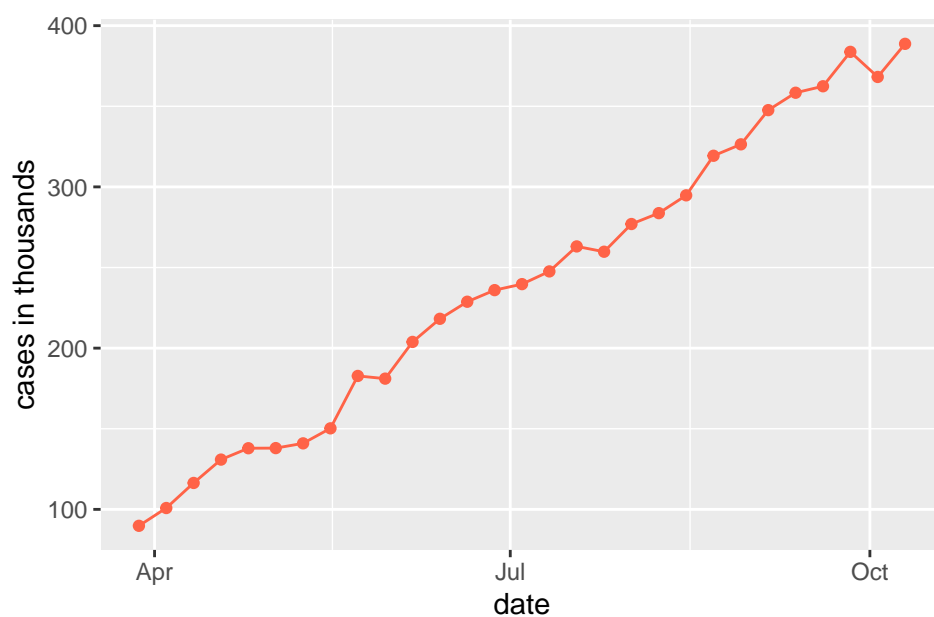
options(scipen = 10000)
p1 <- ggplot(tmp, aes(Date, Cases)) +
  geom_line(color = "#0000A0") +
  geom_point(color = "#0000A0") +
  labs(title = "7 Day Rolling Average New Coronavirus Cases", subtitle = "USA", x="",
        y = "cases")

world["Cases"] <- world["Cases"]/1000
p2 <- ggplot(world, aes(Date, Cases)) +
  geom_line(color = "#FF6347") +
  geom_point(color = "#FF6347") +
  labs(subtitle = "World", x = "date", y = "cases in thousands")
plot_grid(p1, p2, nrow = 2)
```

7 Day Rolling Average New Coronavirus Cases
USA



World



Uznałem, że lepiej podzielić pierwotny wykres na dwa, w celu uproszczenia rozumienia danych na nim przedstawionych.

Źródło

Wykres do stworzenia niniejszego raportu wzięto z [tej strony](#).