

# Praca domowa 2 - raport

Jakub Grunas

1.11.2021

## 1. Wstęp

Poniższy raport z pracy domowej nr. 2 zawiera przedstawienie wykresu danych znalezionej na stronie *koronawirusunas.pl*, który nie został wykonany w pełni prawidłowo, poprawioną wersję omawianego wykresu z kodem potrzebnym do jego utworzenia oraz uzasadnienie, co i dlaczego zostało skorygowane.

## 2. Oryginalny wykres

### Liczba zgonów z powodu COVID-19 w poszczególnych województwach

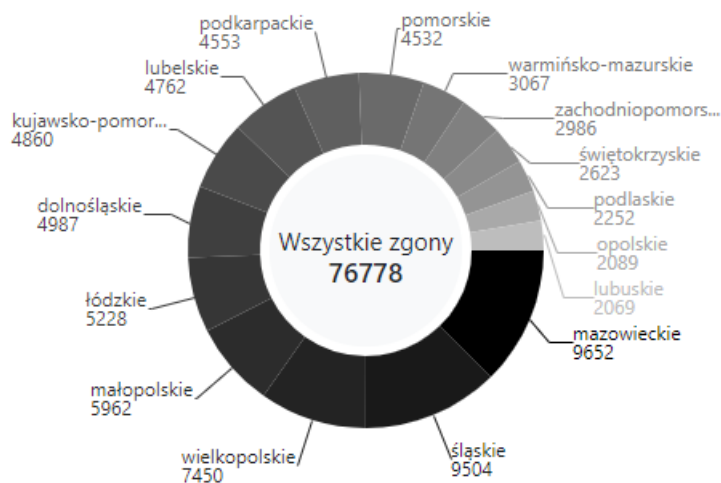


Figure 1: źródło: koronawirusunas.pl, dane z dnia 26.10.2021, godz. 22:00

### 3. Kod oraz poprawiony wykres

```
library(ggplot2)
library(dplyr)
library(forcats)

województwa = c('dolnośląskie', 'kujawsko-pomorskie', 'lubelskie',
                 'lubuskie', 'łódzkie', 'małopolskie',
                 'mazowieckie', 'opolskie', 'podkarpackie',
                 'podlaskie', 'pomorskie', 'śląskie',
                 'świętokrzyskie', 'warmińsko-mazurskie',
                 'wielkopolskie', 'zachodniopomorskie')

zgony = c(4987, 4860, 4762, 2069, 5228, 5962, 9652, 2089,
          4553, 2252, 4532, 9504, 2623, 3067, 7450, 2986)

df <- data.frame(województwa, zgony)
df <- df %>% mutate(województwa = fct_reorder(województwa, zgony))

ggplot(df, aes(x = zgony, y = wojewodztwa)) +
  geom_col(fill = 'red') + scale_x_continuous(limits = c(0,11000), expand = c(0.01,0)) +
  geom_text(aes(label = zgony), hjust = -0.1, size = 3.5) +
  labs(
    title = 'Liczba zgonów z powodu COVID-19 \nw poszczególnych województwach \nod początku pandemii',
    x = 'liczba zgonów',
    y = element_blank()
  ) +
  theme(
    panel.grid.major.y = element_blank(),
    panel.grid.major.x = element_line(colour = 'grey'),
    panel.grid.minor.x = element_line(colour = 'grey')
  ) +
  annotate(geom = 'label', x = 8750, y = 'świętokrzyskie', label = 'Wszystkie zgony:\n76778')
```

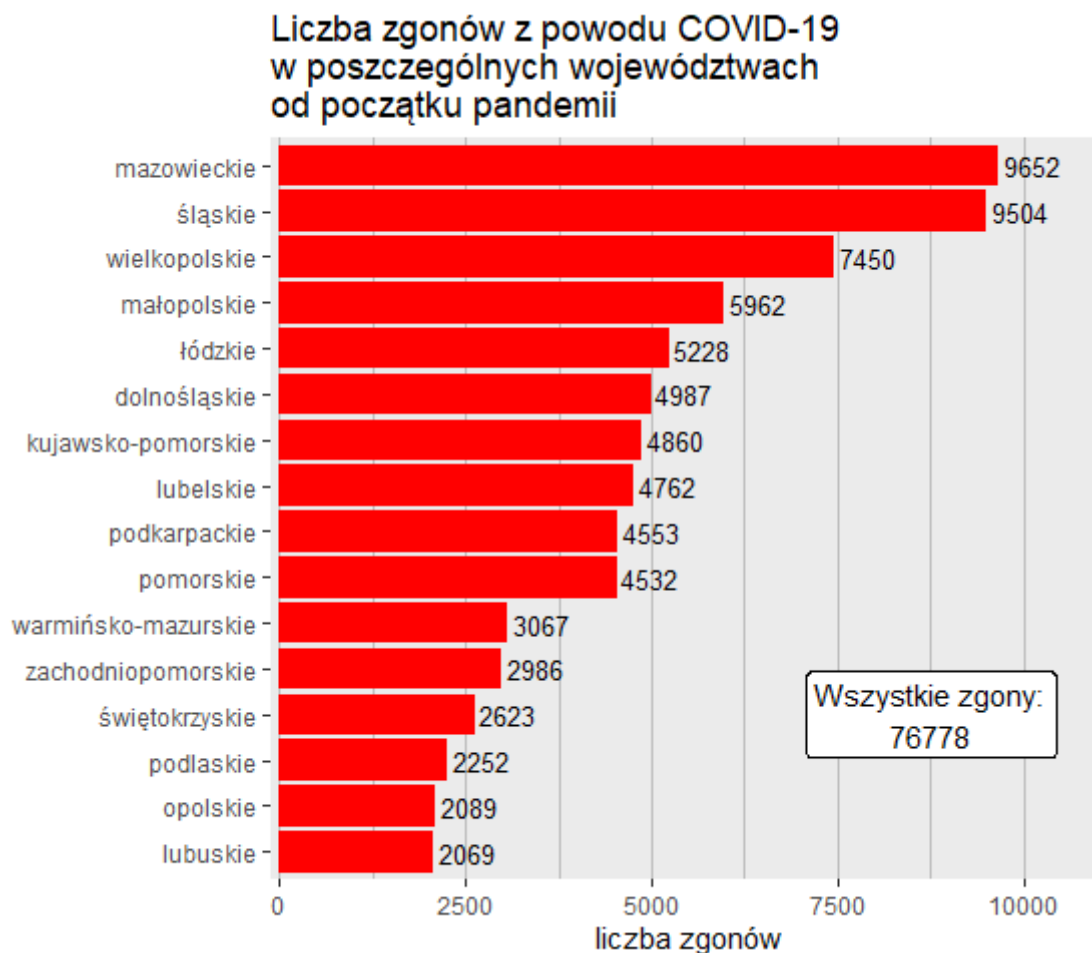


Figure 2: Poprawiony wykres

## 4. Uzasadnienie

- Wykres opublikowany na stronie jest wykresem pierścieniowym, który, podobnie jak wykresy kołowe, nie jest najlepszym wyborem do przedstawienia korelacji między poszczególnymi wartościami. Dlatego typ wykresu zmieniony został na kolumnowy, gdzie wyraźnie i szybko widać różnicę w przedstawianych wartościach pomiędzy województwami.
- Na oryginalnym wykresie obecna była zbędna skala kolorów, która nie wносиła do wykresu żadnych nowych informacji, a ponadto, moim zdaniem, utrudniała odczytywanie wartości, ze względu na podobieństwo sąsiadujących kolorów. Na poprawionym wykresie nie została ona uwzględniona. O wartościach informuje nas jedynie długość słupków oraz dodane, widoczne obok nich *labels*.
- Tytuł wykresu również został skorygowany - jednoznacznie wiemy teraz, od kiedy zbierane są przedstawione dane.