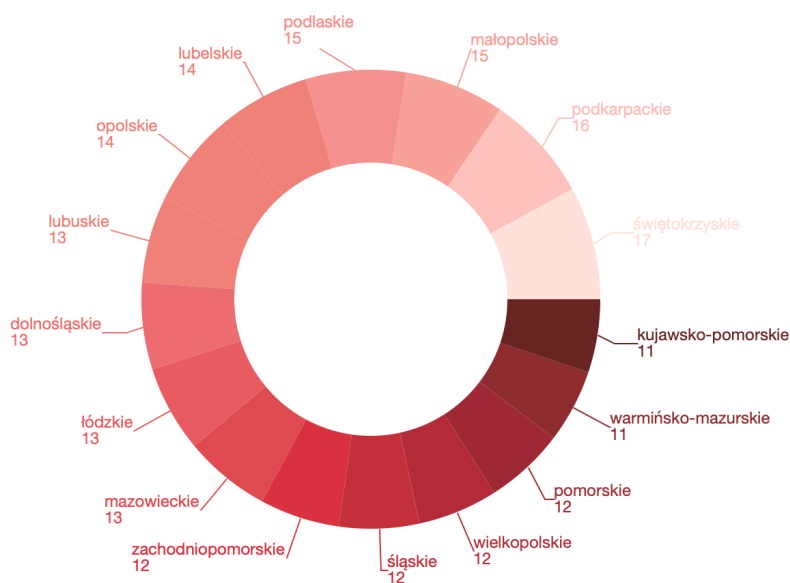


Wizualizacja danych o zakażeniach COVID-19 z wykorzystaniem pakietu ggplot2

Zuzanna Glinka
313347
31 października 2021

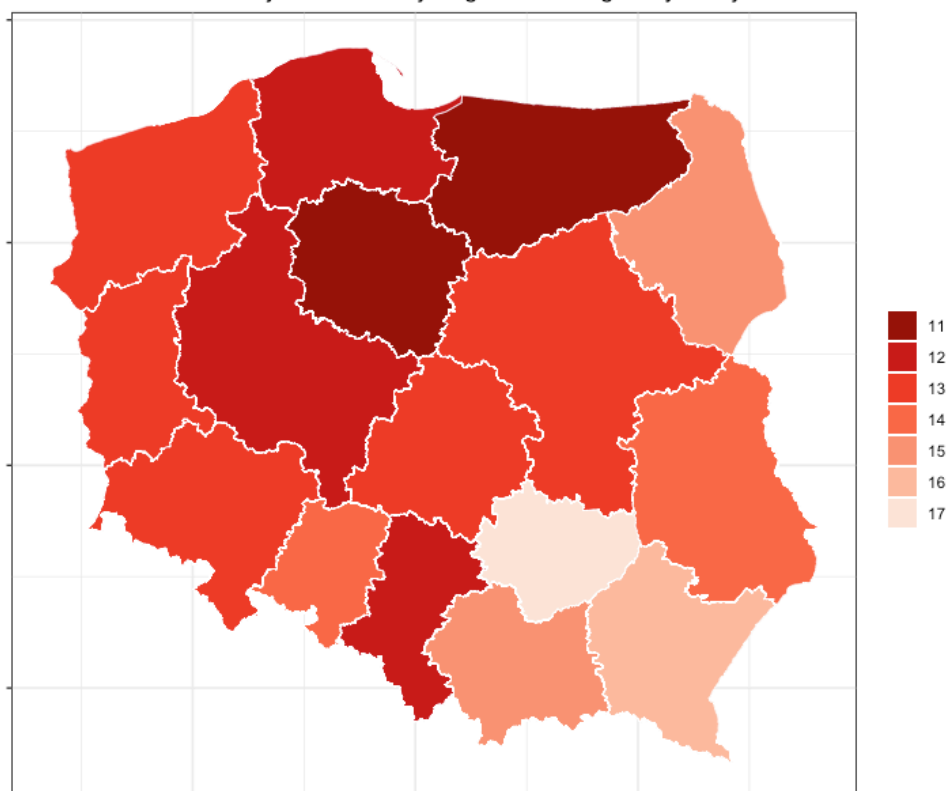
Poprawiany wykres pochodzi ze strony koronawirusunas.pl. Przedstawia dane związane z SARS-CoV-2 w Polsce.

Liczba mieszkańców województwa na jednego zakażonego w tym województwie



Poprawiona wizualizacja:

Liczba mieszkańców województwa na jednego zakażonego w tym województwie



Kod poprawionej wizualizacji:

```
1 library(ggplot2)
2 library(RColorBrewer)
3 library(rgdal)
4 regiony <- readOGR(dsn="/Users/zuzannaglinka/Downloads/Wojewodztwa/Województwa.shp")
5 regiony.df <- read.csv("/Users/zuzannaglinka/Documents/dane_zly_wykres.csv", sep = ";")
6 regiony@data$seq_id <- (1:nrow(regiony@data))
7 regiony@data$covid <- regiony.df$liczba.mieszkanców.na.jedego.zakażonego
8 regiony@data$województwo <- regiony.df$województwo
9
10 regiony@data$id <- rownames(regiony@data)
11 regionydata <- fortify(regiony)
12 regionydf <- merge(regionydata, regiony@data, by = "id")
13
14 p1 <- ggplot(data = regionydf, aes(x=long, y=lat, group = group,
15                                   fill=factor(covid))) +
16   geom_polygon() +
17   geom_path(color = "white", size = 0.2) +
18   scale_fill_manual(
19     values=c("#99000D", "#CB181D", "#EF3B2C", "#FB6A4A", "#FC9272", "#FCBBA1",
20             "#FEE5D9"),
21     breaks= c("11", "12", "13", "14", "15", "16", "17"),
22     labels=c("11", "12", "13", "14", "15", "16", "17"))+
23   coord_equal() +
24   theme(panel.background=element_blank()+
25         theme(panel.background= element_rect(color="black"))) +
26   theme(axis.title = element_blank(),
27         axis.text = element_blank()) +
28   labs(title = "Liczba mieszkańców województwa na jednego zakażonego w tym województwie",
29        fill = "")
```

Poprawiony wykres jest lepszy ze względu na lepszą czytelność. Typ wykresu został źle dobrany. Pierwotny wykres jest pierścieniowy. W powyższym przykładzie suma danych nie ma żadnej interpretacji. Wybór wykresu w postaci mapy pozwala lepiej zobrazować lokalizację problemu. Źle zostały również dobrane kolory wykresu. Dla takich samych danych są różne kolory (np. dla województw: pomorskiego, wielkopolskiego, śląskiego i zachodniopomorskiego) , a przy różnych wartościach są wręcz nierozróżnialne (np. dla lubuskiego i opolskiego).