

# Praca domowa 2 - raport

Autor: Michał Wielgosz

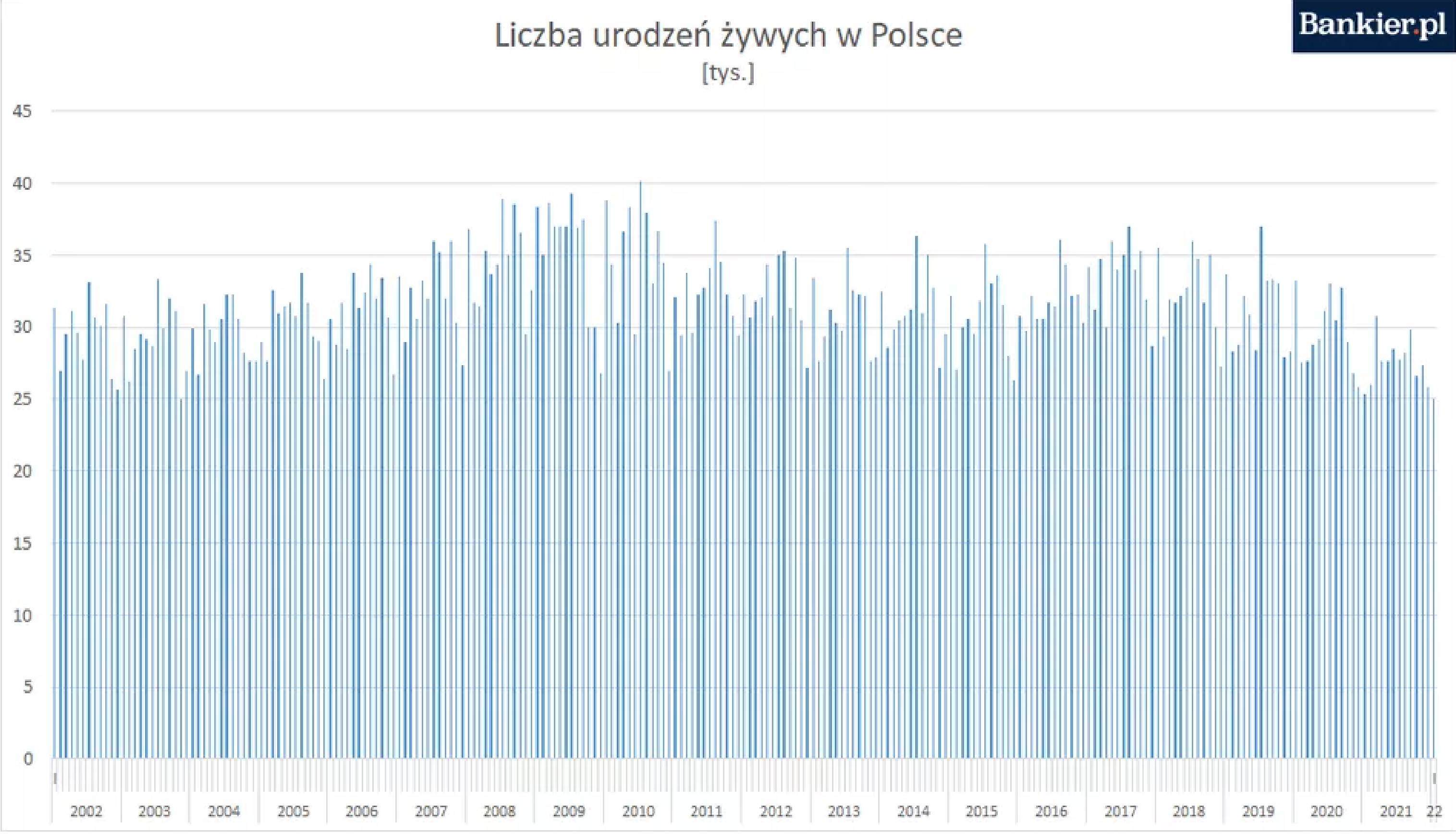
## Dane wykorzystane w pracy

Link do strony z poprawianym wykresem: <https://www.bankier.pl/wiadomosc/Ogromne-oczekiwania-inflacyjne-zalamanie-demografii-i-masakra-polskich-obligacji-Wykresy-tygodnia-8305741.html>

Tekst ze strony opisujący wykres dla kontekstu:

“Rekordowo niska liczba urodzeń w Polsce”

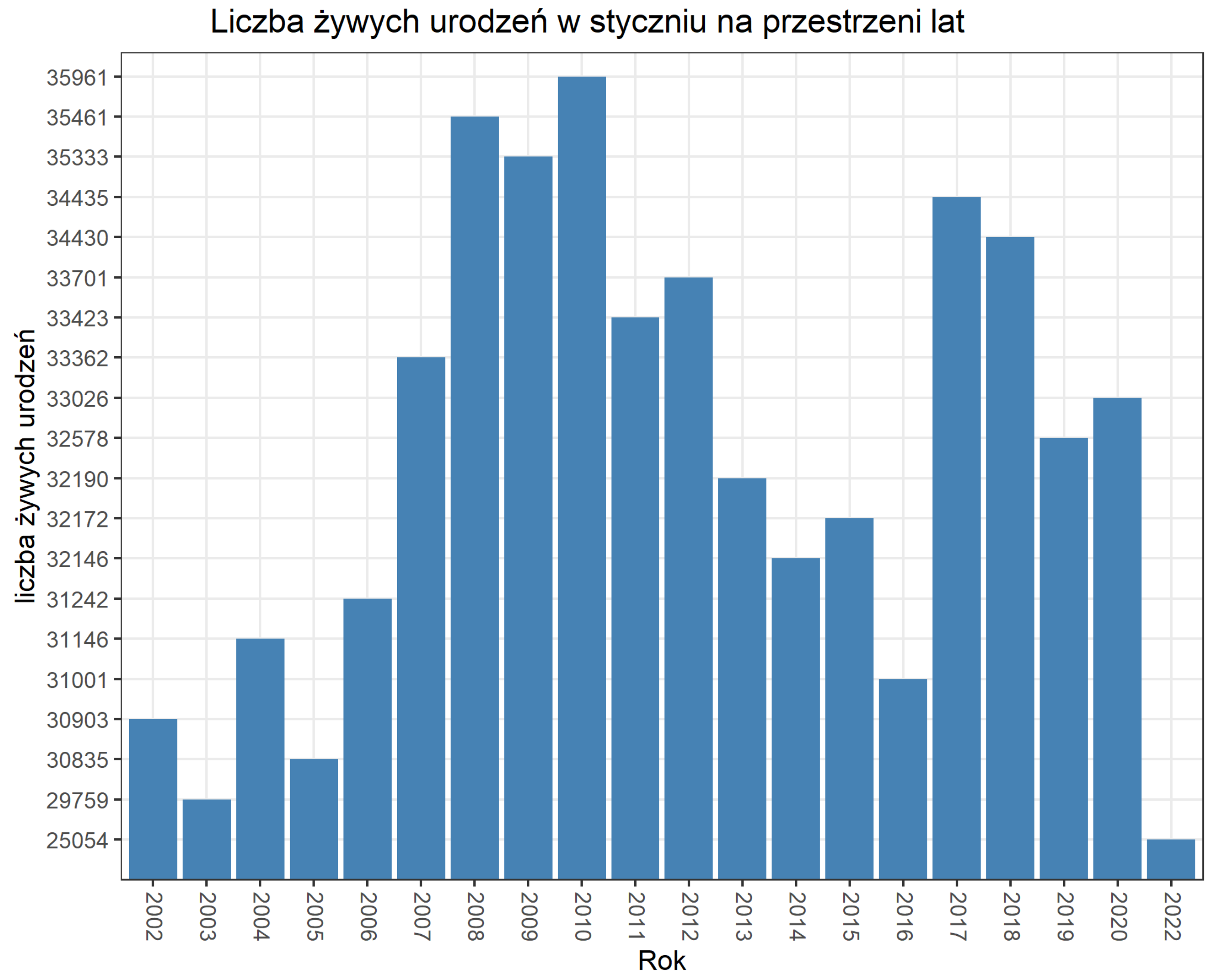
“W styczniu w Polsce urodziło się rekordowo mało dzieci - wynika z szacunków GUS. Liczba zgonów pozostawała wyraźnie wyższa niż przed pandemią.”



## Kod i ulepszona wizualizacja

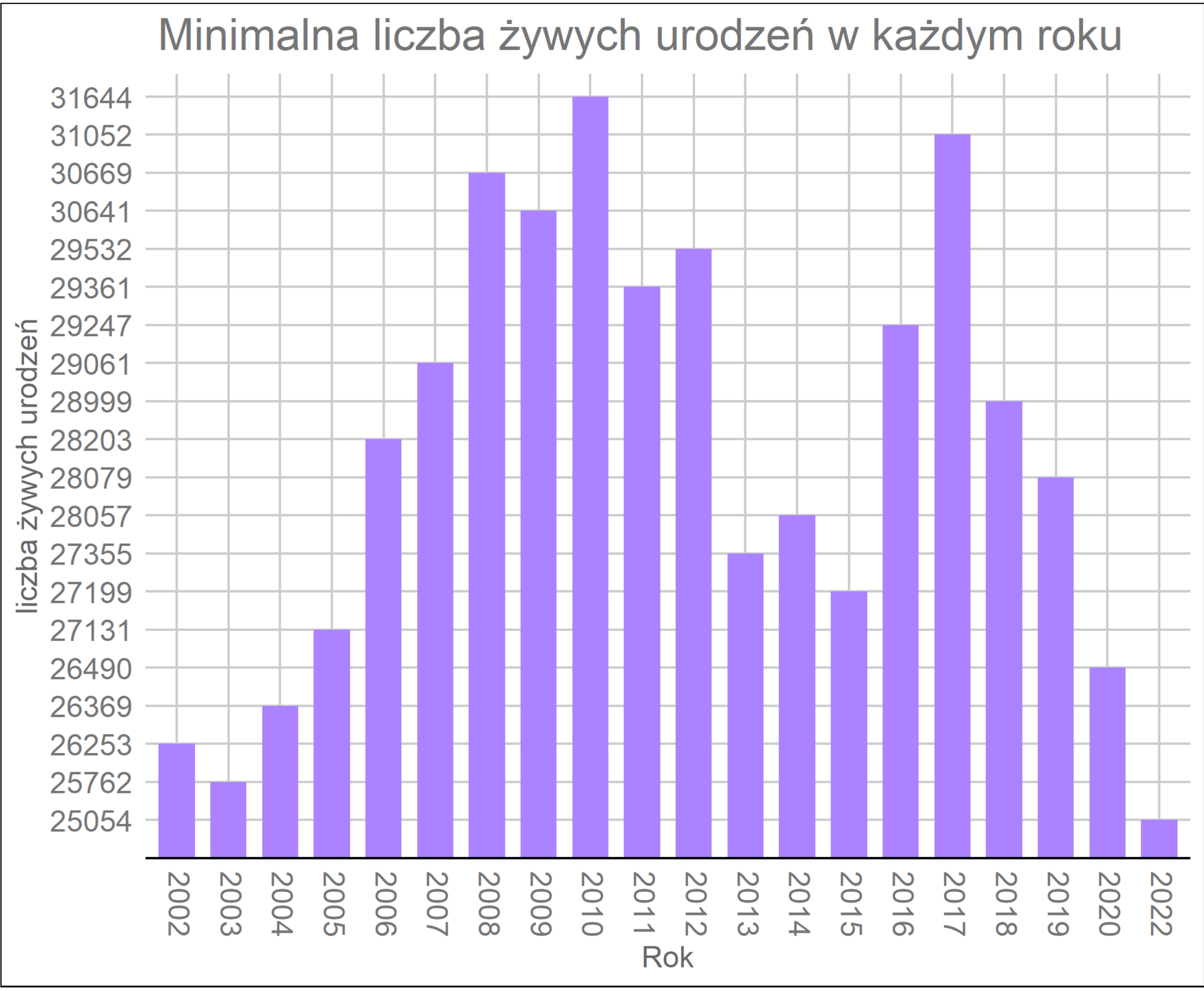
```
#interpretacja 1

ggplot(dff)+
  geom_col(aes(x=rok ,y=styczen,fill=styczen),fill="steelblue")+
  labs(y = "liczba żywych urodzeń", x = "Rok")+
  theme_bw()+
  theme(axis.text.x=element_text(angle = -90, vjust = 0.35,hjust = 0))+
  ggtitle("          Liczba żywych urodzeń w styczniu na przestrzeni lat")
```



```
#interpretacja 2

ggplot(dff,aes(x=rok ,y=Min,fill=Min))+
  geom_col(fill="mediumpurple1",width = 0.7)+
  theme_gdocs()+
  labs(y = "Liczba żywych urodzeń", x = "Rok")+
  theme(axis.text.x=element_text(angle = -90, vjust = 0.35,hjust = 0),plot.title=element_text(size=18))+
  ggtitle(" Minimalna liczba żywych urodzeń w każdym roku")
```



## Uzasadnienie

- W oryginalnym wykresie ciężko domyślić się co autor chciał zaprezentować, fakt, że styczeń 2022 r. cechował się najmniejszą liczbą żywych urodzeń w porównaniu do poprzednich styczniów? Czy może w porównaniu do wszystkich poprzednich lat? Stąd też dwie interpretacje, stosowne do obu przypadków
- Znaczne zwiększenie czytelności wykresu, dzięki zmniejszeniu ilości pokazanych danych przy zachowaniu przekazu
- Skale mocno precyzyjne, umożliwiają porównanie poszczególnych lat

## Reszta kodu plus źródło danych

Dane wykorzystane do obliczenia liczby żywych urodzeń w styczniu 2022 r.

<https://stat.gov.pl/podstawowe-dane/>

Dane wykorzystane do utworzenia tabeli

<https://demografia.stat.gov.pl/BazaDemografia/Tables.aspx>

Nie wiem czy ten kod też ma być, bo tematem był ggplot

```
library(ggplot2)
library(readxl)
library(dplyr)
library(scales)
library(ggthemes)
library(qlcMatrix)

#####

#Sprawdzana jest wiedza z ggplota, a tabelki możemy tworzyć nawet od ręki, więc uznałem, że zostawię tu tego potworka...

df2<-read_excel("pl_uro_2002_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df3<-read_excel("pl_uro_2003_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df4<-read_excel("pl_uro_2004_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df5<-read_excel("pl_uro_2005_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df6<-read_excel("pl_uro_2006_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df7<-read_excel("pl_uro_2007_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df8<-read_excel("pl_uro_2008_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df9<-read_excel("pl_uro_2009_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df10<-read_excel("pl_uro_2010_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df11<-read_excel("pl_uro_2011_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df12<-read_excel("pl_uro_2012_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df13<-read_excel("pl_uro_2013_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df14<-read_excel("pl_uro_2014_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df15<-read_excel("pl_uro_2015_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df16<-read_excel("pl_uro_2016_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df17<-read_excel("pl_uro_2017_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df18<-read_excel("pl_uro_2018_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df19<-read_excel("pl_uro_2019_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)
df20<-read_excel("pl_uro_2020_00_6.xls") %>%
  select(3:14) %>%
  slice(9:9)

#Tutaj kolumna odręcznie, bo naturalnie nie ma jeszcze tabelek na rok 2021(Nigdzie na gusie nie ma, nie wiem skąd wzięli do artykułu!), a co #dopiero 2022, więc, na 2022 musiałem wyliczyć na kartce...

df22<-c(25054,"Na","Na","Na","Na","Na","Na","Na","Na","Na","Na","Na")

#####

dff <- rbind(df2, df3,df4,df5,df6,df7,df8,df9,df10,df11,df12,df13,df14,df15,df16,df17,df18,df19,df20,df22) %>%
  mutate(rok = c("2002","2003","2003","2004","2004","2005","2005","2006","2006","2007","2007","2008","2008","2009","2009","2010","2010","2011","2011","2012","2012","2013","2013","2014","2014","2015","2015","2016","2016","2017","2017","2018","2018","2019","2019","2020","2020","2021","2021","2022","2022"),before = 1) %>%
  rowwise() %>%
  mutate(Min = min(...3,...4,...5,...6,...7,...8,...9,...10,...11,...12,...13))
colnames(dff) <- c("rok","styczen","luty","marzec","kwiecień","maj","czerwiec","lipiec","sierpień","wrzesień","październik","listopad","grudzień","Min")
```