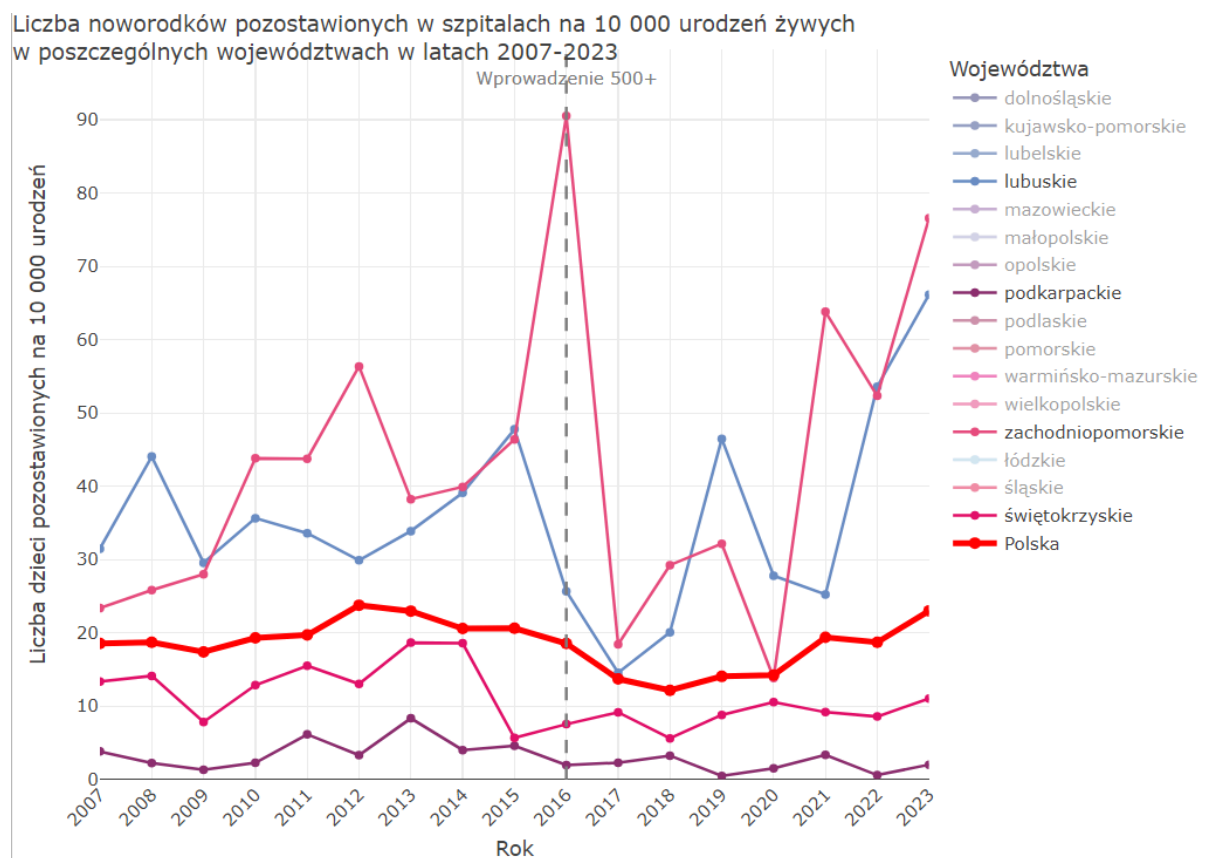


Zuzanna Zalewska

Praca domowa 4

W analizie danych zostały użyte dwie ramki danych udostępnione przez fundację Gajusz: Urodzenia żywe w Polsce 2007-2023 i Noworodki pozostawione w szpitalu 2007-2023. Wykorzystując te dane możliwe jest obliczenie liczby noworodków pozostawionych w szpitalach na 10 000 urodzeń żywych w poszczególnych województwach w podanych latach, co dostarcza nam konkretniejszych danych do analizy. Uwzględniony został także komentarz znajdujący się w jednej z ramek. Poniżej reprezentująca to wizualizacja wraz z potrzebnym kodem:



```

library(dplyr)
library(tidyrr)
library(plotly)

#przygotowanie danych do stworzenia wykresu
noworodki_w_szpitalu=read.csv("Noworodki_pozostawione_w_szpitalu.csv",comment.char = '#')
noworodki_w_szpitalu<-noworodki_w_szpitalu %>%
  rename_with(~substr(.,nchar(.)-3,nchar(.)),
    .cols=2:ncol()) %>%
  pivot_longer(cols=~Województwo,names_to="Rok",values_to="Pozostawione_noworodki")
urodzenia_zywe=read.csv("Urodzenia_zywe_w_Polsce.csv")
urodzenia_zywe=urodzenia_zywe %>%
  rename_with(~substr(.,2,nchar(.)),
    .cols=2:ncol()) %>%
  pivot_longer(cols=~Województwo,names_to="Rok",values_to="Urodzenia_zywe")
df=noworodki_w_szpitalu %>%
  inner_join(urodzenia_zywe,by=c("Województwo","Rok")) %>%
  mutate(Stosunek_pozostawionych_na_10000=round((Pozostawione_noworodki/Urodzenia_zywe)*10000,2))
df$Rok=as.numeric(df$Rok)

#kolory udostępnione przez fundację
colors <- c("#303174", "#315ca8", "#b5e0f3", "#884292", "#8c2a64", "#e62248", "#e4007e", "#ea4f7f")

#dla czytelności początkowego wykresu wybrane zostały niektóre województwa oraz Polska
wybrane_województwa = unique(df$Województwo)[c(4,8,13,16,17)]

fig <- plot_ly() %>%
  add_trace(
    data = df[df$Województwo != "Polska", ],
    x = ~Rok,
    y = ~Stosunek_pozostawionych_na_10000,
    type = 'scatter',
    mode = 'lines+markers',
    color = ~Województwo,
    colors = colors,
    hoverinfo = 'text',
    text = ~paste("Województwo:", Województwo,
      "<br>Rok:", Rok,
      "<br>Liczba dzieci pozostawionych na 10 000 urodzeń:", Stosunek_pozostawionych_na_10000),
    visible = ~ifelse(Województwo %in% wybrane_województwa, TRUE, "legendonly"),
    showlegend = TRUE
  ) %>%
  #wyszczególnienie danych dla Polski
  add_trace(
    data = df[df$Województwo == "Polska", ],
    x = ~Rok,
    y = ~Stosunek_pozostawionych_na_10000,
    type = 'scatter',
    mode = 'lines+markers',
    line = list(color = "red", width = 4),
    marker = list(color = "red", size = 8),
    name = "Polska",
    hoverinfo = 'text',
    text = ~paste("Kraj:", Województwo,
      "<br>Rok:", Rok,
      "<br>Liczba dzieci pozostawionych na 10 000 urodzeń:", Stosunek_pozostawionych_na_10000),
    visible = TRUE,
    showlegend = TRUE
  ) %>%

```

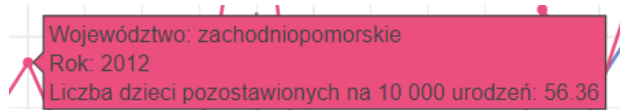
```

layout(
  title = list(text="Liczba noworodków pozostawionych w szpitalach na 10 000 urodzeń żywych
                w poszczególnych województwach w latach 2007-2023",font=list(size=15),xanchor='left',x=0),
  xaxis = list(title = "Rok",tickmode='linear',dtick=1,tickangle=-45,range = c(min(df$Rok), max(df$Rok))),
  yaxis = list(title = "Liczba dzieci pozostawionych na 10 000 urodzeń",
                range = c(0, max(df$Stosunek_pozostawionych_na_10000) * 1.1)),
  showlegend = TRUE,
  legend = list(title = "Województwa", font = list(size = 14)),
  #stworzenie przerywanej linii wskazującej na rok wprowadzenia 500+
  shapes = list(
    list(
      type = "line",
      x0 = 2016, x1 = 2016,
      y0 = 0, y1 = 1,
      xref = "x", yref = "paper",
      line = list(color = "gray", dash = "dash", width = 2)
    )
  ),
  annotations = list(
    list(
      x = 2016,
      y = 0.98,
      xref = "x",
      yref = "paper",
      text = "wprowadzenie 500+",
      showarrow = FALSE,
      font = list(size = 12, color = "gray")
    )
  )
)
fig

```

Komentarz do wizualizacji:

Aby zachować czytelność wykresu pozostawione zostały wybrane województwa wraz z Polską, ale użytkownik może wybrać samodzielnie interesujące go województwa poprzez kliknięcie ich w legendzie, zatem uzyskujemy możliwość porównywania danych pomiędzy województwami w określonym czasie. Dane dla Polski zostały bardziej wyszczególnione. Ponadto po najechnaniu na odpowiednią kropkę wyświetlone są szczegółowe informacje:



Została dodana linia przerywana reprezentująca wprowadzenie programu 500+ (komentarz do jednej z ramek), który mógł wypłynąć na dane (ogółem dla Polski i części województw jest widoczny spadek po tym roku, ale w ostatnich latach mamy do czynienia ze wzrostem). Co ciekawe są województwa mające znaczne różnice pomiędzy kolejnymi latami (jak na przykład zachodniopomorskie), a także pozostające na mniej więcej stałym poziomie (tak jak podkarpackie).