

Praca domowa 3 z Eksploracji Danych

Patrycja Gromada

01.04.2024r

1 Niepoprawna wizualizacja

Poniżej znajduje się wybrana wizualizacja z dn. 26.03.2024r.
Wizualizacja przedstawia wyniki sondażu przeprowadzonego przez IRCenter na zlecenie TVP3 Warszawa odnośnie wyborów na prezydenta Warszawy.



Źródło: https://twitter.com/zofia_malisz/status/1772653404847735133

Oryginalnie wizualizacja ukazała się na stronie TVP3 Warszawa:

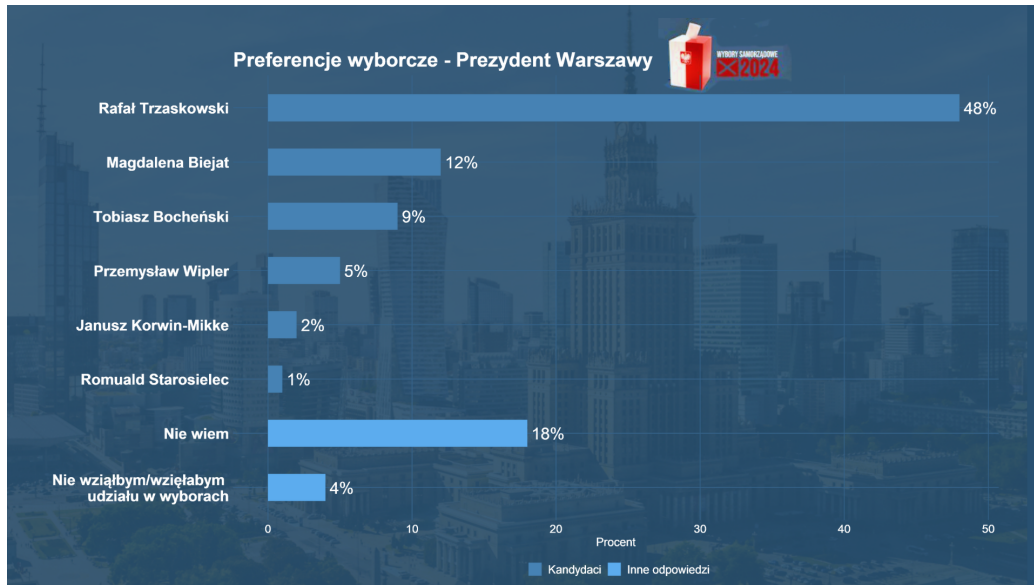
<https://warszawa.tvp.pl/76633707/>

kto-prowadzi-w-wyscigu-o-fotel-prezydenta-warszawy-mamy-nowy-sondaz

1.1 Poprawki

Słupki "Magdalena Biejał" 12% i słupki "Nie wiem" 18% mają bardzo zaburzone proporcje. Również słupki z 4% wyglądają na taki sam, jak powyższy 2%. Ogólnie mamy tutaj do czynienia z zaburzonymi proporcjami słupków oraz użyciem dwóch kolorów, mimo, że słupki są podpisane.

2 Poprawiona wizualizacja



2.1 Kod

```
library(ggplot2)
library(dplyr)
library(extrafont)

df <- data.frame(
  k = c(
    "Rafa Trzaskowski", "Magdalena Biejat", "Tobiasz
      Boche ski", "Przemys aw Wipler",
    "Janusz Korwin-Mikke", "Romuald Starosielec", "Nie
      wiem", "Nie wzi bym/wzi abym \nudzia u w
      wyborach"
  ),
  p = c(48, 12, 9, 5, 2, 1, 18, 4))

fill <- factor(df$k %in% c("Nie wiem", "Nie wzi bym/
  wzi abym \nudzia u w wyborach"),
```

```

        labels = c("Kandydaci", "Inne
                    odpowiedzi"))

df$k <- factor(df$k, levels = rev(unique(df$k)))

col_grid <- rgb(235, 235, 235, 100, maxColorValue = 255)

wykres <- ggplot(df, aes(x = k, y = p, fill = fill)) +

  geom_col(width = 0.5) +
  geom_text(aes(label = paste0(p, "%")), position =
    position_nudge(y = 0.3), size = 7, color = "white",
    hjust = 0) +
  coord_flip() +
  labs(
    title = "Preferencje wyborcze - Prezydent Warszawy",
    x = NULL, y = "Procent"
  ) +

  theme_minimal() +
  theme(
    legend.position = "bottom",
    text = element_text(size = 14, family = "Arial",
      color = "white"),
    plot.title = element_text(size = 24, family = "Arial"
      , color = "white", face="bold"),
    legend.text = element_text(size = 14, family = "Arial"
      , color = "white"),
    axis.text.y = element_text(size = 18, family = "Arial"
      , color = "white", face="bold"),
    axis.title.y = element_text(size = 14, family = "
      Arial", color = "white"),
    plot.caption = element_text(size = 12, family = "
      Arial", color = "white"),
    axis.text.x = element_text(size = 14, family = "Arial"
      , color = "white"),
    panel.grid.major.y = element_line(color = "steelblue4"
      ),

```

```

    panel.grid.major.x = element_line(color = "steelblue4",
    ),
    panel.grid.minor.x = element_blank()
) +

theme(plot.background = element_rect(fill = "white",
    colour = "white")) +
scale_fill_manual(values = c("Kandydaci" = "steelblue",
    "Inne odpowiedzi" = "steelblue2")) +
guides(fill = guide_legend(title = "")) +
theme(plot.background = element_rect(fill = "
    transparent", color = NA))

ggsave("wykres1.png", plot = wykres, width = 16, height =
    9, dpi = 300)

```

Tło i czerwony znaczek zostały dodane w Canva.

2.2 Uzasadnienie

Przygotowany wykres jest lepszy od oryginalnego, ponieważ poprawnie przedstawia proporcje pomiędzy długościami słupków. Została również podana legenda w celu uzasadnienia użycia dwóch kolorów. Uznałam, że kolory mogą zostać, ponieważ odróżniają one nazwiska kandydatów od innych odpowiedzi.