

Kukla_Michal_HW2

Michał Kukla

2023-11-07

Źródło

Wykres pochodzi z posta na Facebooku:

<https://www.facebook.com/ZSPnr1Gorzyce/posts/pfbid0n2uEvFJhRm2iH7F4tLbU2AVqBNbA6Je3bRTjk9zhgKRmc2Biy4i>
z dnia 16.10.2023 r.

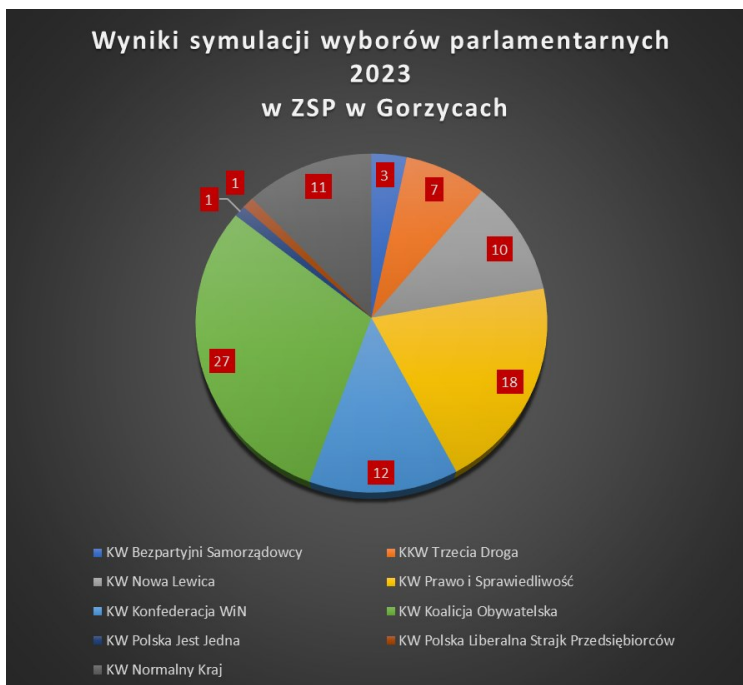


Figure 1: Oryginalna wizualizacja

Błędy w wizualizacji

1. W wykresie kołowym sporządzone wartości zazwyczaj sumują się do 100%. Przy takim wyborze wykresu należałoby oznaczyć wycinki liczbą procentów, a nie ilością oddanych głosów - wtedy wykres byłby czytelniejszy.
2. W wykresie poparcia partii politycznych wybranie wykresu kołowego jest raczej nieodpowiednie. Ludzki mózg ma skłonność do zawyżania wielkości kątów ostrych i zaniżania kątów rozwartych. Lepszy byłby tutaj wykres słupkowy.
3. Wykres jest nieczytelny: można by poszeregować wycinki koła od największego do najmniejszego, podobnie w kolejności od najbardziej do najmniej popularnego ugrupowania uszeregować legendę. Oprócz tego można by zastosować inne kolory (są 2 odcienie niebieskiego i 2 szarego, ciężkie w rozpoznaniu, który jest który).

Poprawiona wizualizacja

Poniżej znajduje się kod implementujący wizualizację oraz wykresy:

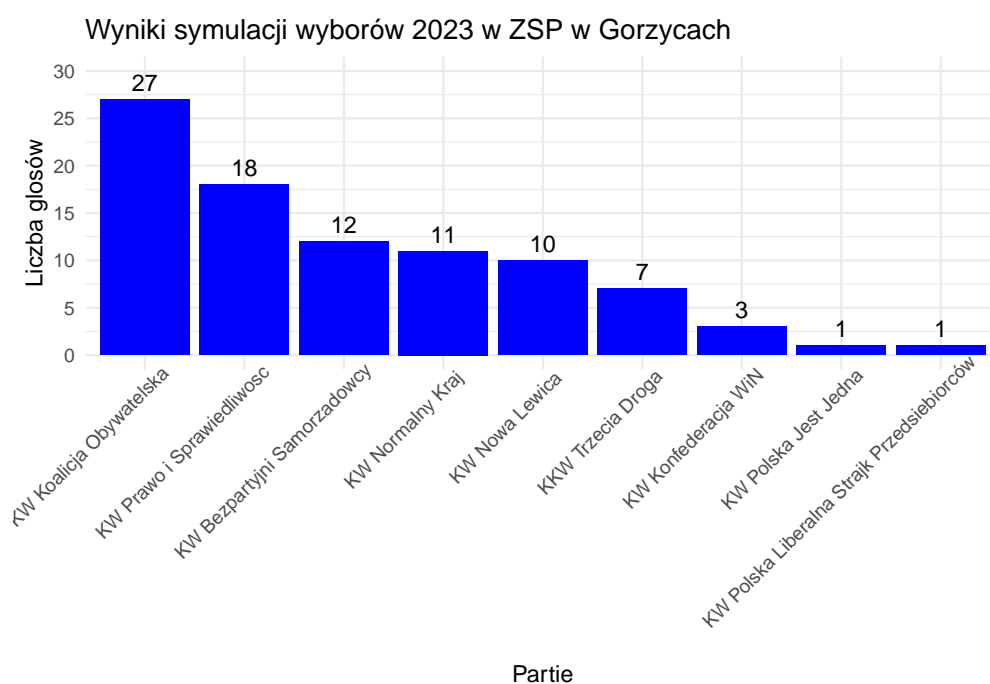
```
library(ggplot2)
library(dplyr)
```

```
df <- data.frame(Partie = c("KW Koalicja Obywatelska", "KW Prawo i Sprawiedliwość",
                           "KW Bezparyjni Samorządowcy", "KW Normalny Kraj",
                           "KW Nowa Lewica", "KKW Trzecia Droga", "KW Konfederacja WiN",
                           "KW Polska Jest Jedna", "KW Polska Liberalna Strajk Przedsiębiorców"),
                 liczba = c(27, 18, 12, 11, 10, 7, 3, 1, 1))
```

```
x = sum(c(27, 18, 12, 11, 10, 7, 3, 1, 1))
```

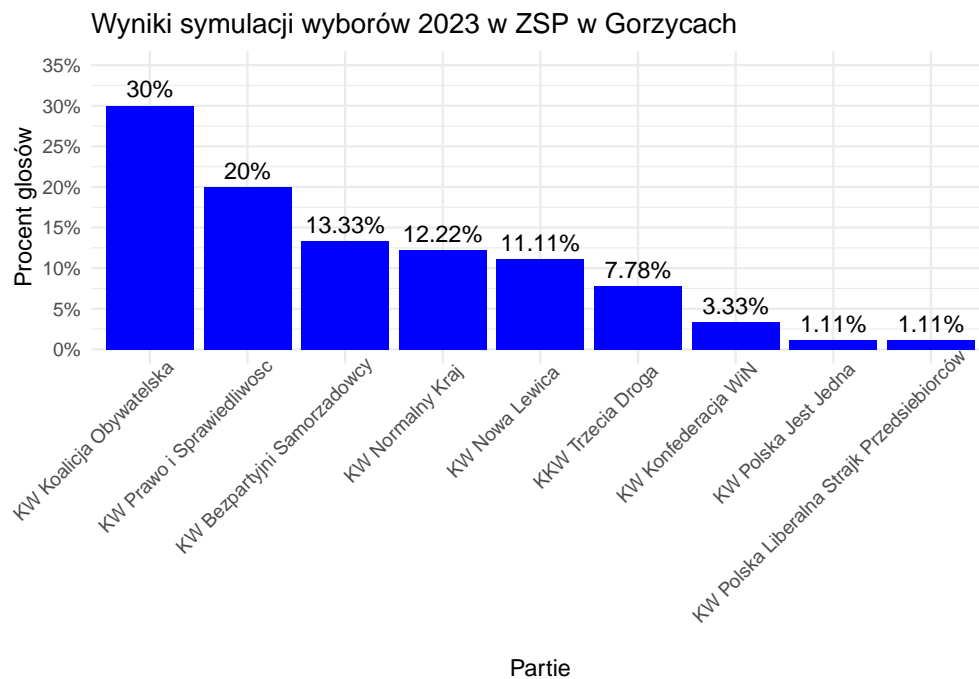
```
df %>%
  mutate(procenty = round((liczba/x*100),2)) -> df1
```

```
dobry_wykres1 <-
  ggplot(df, aes(x = reorder(Partie, -liczba), y = liczba)) +
  geom_bar(stat = "identity", fill = "blue") +
  geom_text(aes(label = liczba), vjust = -0.5) +
  labs(title = "Wyniki symulacji wyborów 2023 w ZSP w Gorzycach", x = "Partie", y = "Liczba głosów") +
  scale_y_continuous(limits = c(0,30),
                    breaks = seq(0, 30, by = 5)) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 0.9))
```



Uważam, że wykres byłby jeszcze czytelniejszy, gdyby zamiast liczby oddanych głosów rozważać procent wszystkich głosów. Oto ostateczny wykres z implementacją:

```
dobry_wykres2 <-  
  ggplot(df1, aes(x = reorder(Partie, -procenty), y = procenty)) +  
  geom_bar(stat = "identity", fill = "blue") +  
  geom_text(aes(label = paste0(procenty, "%")), vjust = -0.5) +  
  labs(title = "Wyniki symulacji wyborów 2023 w ZSP w Gorzycach", x = "Partie", y = "Procent głosów") +  
  scale_y_continuous(labels = scales::percent_format(scale = 1), limits = c(0,35),  
    breaks = seq(0, 35, by = 5)) +  
  theme_minimal() +  
  theme(axis.text.x = element_text(angle = 45, hjust = 0.9))
```



Zalety nad oryginalną wizualizacją:

1. Wykres jest dużo czytelniejszy - od razu widać kolejność popularności partii.
2. Czytelne osie sprawiają, że od razu można odczytać dokładny wynik partii.
3. Ilość głosów została przeformatowana na procent wszystkich głosów.