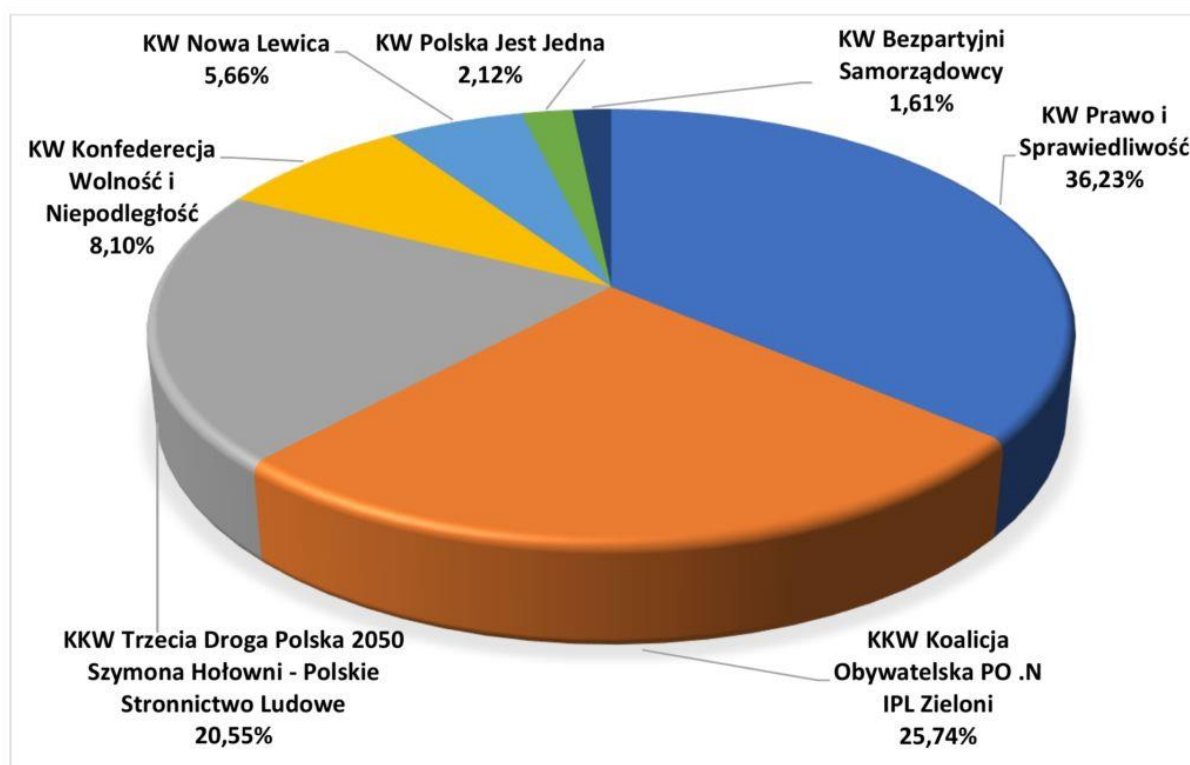


Praca domowa 2

Wizualizacja źródłowa

Zdecydowałem się dokonać analizy i poprawy wykresu reprezentującego ostateczne wyniki wyborów do Sejmu z dnia 15 października 2023 roku w gminie Niepołomice. Wykres ukazał się 24 października 2023 roku na portalu www.niepolomice.eu pod adresem <https://www.niepolomice.eu/inne/oficjalne-wyniki-glosowania-w-wyborach-2023-w-gminie-niepolomice/>. Oto on:



Co należy poprawić?

- 1) Głównym źródłem błędów, jakie zostały popełnione podczas tworzenia powyższego wykresu, jest zastosowanie wykresu kołowego. Dobrze wiadomo, że ludzki mózg ma problem z właściwym oszacowaniem kątów ostrych i rozwartych, co w naszym przypadku zaburza prawidłowe zrozumienie wykresu.
- 2) Co więcej, wykres jest trójwymiarowy. Wprowadzenie perspektywy zniekształca postrzeganie wielkości, co w połączeniu z faktem, że mamy wykres kołowy, przynosi tragiczne skutki. Wystarczy spojrzeć na obszar na wykresie zajmowany przez Koalicję Obywatelską i porównać go z polem przeznaczonym na Prawo i Sprawiedliwość. Obszary te wydają się do siebie zbliżone lub wręcz można stwierdzić, że więcej głosów otrzymała KO. Tymczasem dane wskazują coś zupełnie innego: 6658 głosów dla PiS-u i 4729 głosów dla KO.
- 3) Podpisy na wykresie składają się z pełnych, oficjalnych nazw zarejestrowanych komitetów wyborczych. Długie podpisy zaburzają czytelność wykresu i stanowią okazję do literówek, którą autor wykorzystuje. Nie stracimy nic na przekazie, gdy zastąpimy te wyrażenia potocznymi nazwami często wykorzystywanymi w mediach, na przykład zamiast "Komitet Wyborczy Nowa Lewica", napiszemy "Lewica".
- 4) Zastosowanie wykresu kołowego wprowadza konieczność rozróżnienia partii kolorami. Ich dobór zdaje się być zupełnie losowy, wskutek czego można nieświadomie przekazać jakiś komentarz na temat danej partii. Lepszym wyborem byłoby użycie kolorów, które jednoznacznie utożsamiane są z daną partią albo po prostu jednej barwy dla wszystkich komitetów.
- 5) Wartości na wykresie nie sumują się do 100, lecz do 100.01, co wynika z niepoprawnie przyjętych zaokrągleń.

Poprawiona wizualizacja



Kod niezbędny do wygenerowania wizualizacji:

```
library(ggplot2)

setwd("C:\\TWD\\Prace domowe\\Praca domowa 2\\RowickiMikołaj")

df <- data.frame(
  KomitetWyborczy = c("Prawo i\n Sprawiedliwość", "Koalicja\n Obywatelska",
    "Trzecia Droga", "Konfederacja", "Lewica",
    "Polska jest jedna", "Bezpartyjni\n Samorządowcy"),
  LiczbaGlosow = c(6658, 4729, 3776, 1488, 1040, 389, 295),
  ProcentGlosow = c(36.2, 25.7, 20.6, 8.1, 5.7, 2.1, 1.6)
)

p <- ggplot(df, aes(x = reorder(KomitetWyborczy, -ProcentGlosow),
  y = ProcentGlosow)) +
  geom_bar(stat = "identity", fill = alpha("blue", 0.8),
    position = position_dodge(width = 10)) +
  ggtitle("Oficjalne wyniki wyborów do sejmu w gminie Niepołomice") +
  labs(x = "Komitet wyborczy",
    y = "Procent głosów") +
  theme(plot.title = element_text(hjust = 0.5,
    margin = margin(b = 20)),
    axis.title.x = element_text(margin = margin(t = 10)),
    axis.title.y = element_text(margin = margin(r = 10)),
    panel.background = element_rect(fill = "white"),
    panel.grid = element_line(color = "gray", size = 0.1)) +
  geom_text(aes(label = LiczbaGlosow),
    vjust = -0.5, size = 3, color = "#4D4D4D") +
  scale_y_continuous(labels = scales::label_percent(scale = 1),
    limits = c(0, 40))

ggsave("plot.png", plot = p, width = 8 )
```

Co się zmieniło?

- 1) Przede wszystkim podjąłem decyzję o rezygnacji z wykresu kołowego, który zaburzał nam odczyt danych i postawiłem na dobrze znany wykres słupkowy. Dzięki temu w jasny sposób jesteśmy w stanie odczytać różnice pomiędzy liczbą głosów zdobytych przez poszczególne partie.
- 2) Zrezygnowałem również z wykresu trójwymiarowego, by nie wprowadzać zaburzenia wynikającego z perspektywy.
- 3) Uznałem, że pełne nazwy komitetów wyborczych można zastąpić nazwami głównych partii, które wchodzi w ich skład. Są to powszechnie znane nazwy, jednoznacznie utożsamiane przez wszystkich z poszczególnymi ugrupowaniami, więc nie tracimy żadnej dokładności, zaś mniej słów na wykresie poprawia poziom czytelności.
- 4) Zdecydowałem się na jednokolorowy, estetyczny kolor dla wszystkich słupków (niebieski), by nie wywołać przypadkiem negatywnych skojarzeń z któryś z ugrupowań.
- 5) Uznałem, że zamiast pisać przybliżone wartości procentowe nad słupkami i narażać się niepotrzebnie na błąd wynikający z zaokrągleń, warto dodać na wykresie dokładną liczbę głosów, jaką otrzymał dany komitet. Nasze dane dotyczą wyłącznie gminy Niepołomice, są to więc stosunkowo małe liczby, które nie zaburzają czytelności wykresu. Są też dokładniejsze niż jakkolwiek przybliżone wartości procentowe, a więc wnoszą nam dodatkową informację, która umożliwia nam realne porównanie poparcia komitetów Bezpartyjnych Samorządowców i Polska jest jedna. Wartości procentowe nie są w stanie poprawnie ukazać różnicy między nimi.