

PRACA DOMOWA 2 – BASIA SEWERYN

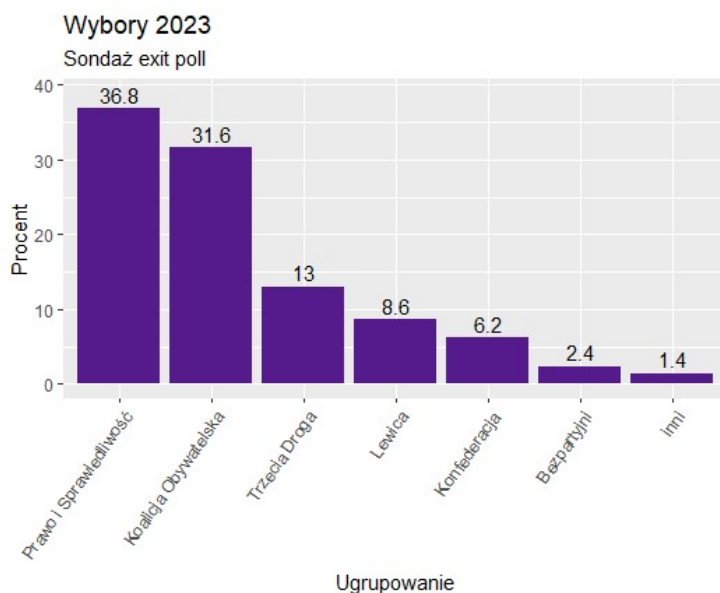
1. Wykres 1 – sondaż Exit Poll z tegorocznych wyborów parlamentarnych

Na portalu istotne.pl natknęłam się na prosty wykres słupkowy przedstawiający wyniki sondażu exit poll. Widziałam te dane już kilkakrotnie dlatego nie przyglądałam się wykresowi bardzo dokładnie, ale i tak od razu zauważyłam nieprawidłowość w wysokości słupków.



Źródło: [Są sondażowe wyniki wyborów - Bolesławiec - istotne.pl](https://istotne.pl),
data publikacji : 15.10.2023

POPRAWIONA WIZUALIZACJA:



Łatwo zauważyć, że słupki dla danych różniących się o 2,4 pkt. proc. są tej samej wysokości. Może być to bardzo mylące dla czytelnika, zwłaszcza dlatego, że wartość następnego słupka (dla Bezpartyjnych) jest dokładnie taka sama jaka powinna być różnica między wskazanymi słupkami.

Na poprawionej wizualizacji tych danych widać ewidentną różnicę między słupkami dla Lewicy i Konfederacji oraz lepiej oddana jest wartość słupka dla Bezpartyjnych, ponieważ na źródłowym wykresie wydawała się ona zdecydowanie mniejsza. Taki wykres nie tylko poprawnie pokazuje dane sondażu, ale też jest bardziej czytelny (choćby ponieważ ma oznaczoną oś pionową).

2. Wykres 2 – wykres bezrobocia w Polsce w latach 2013-2023 udostępniony przez Kancelarię Prezesa Rady Ministrów



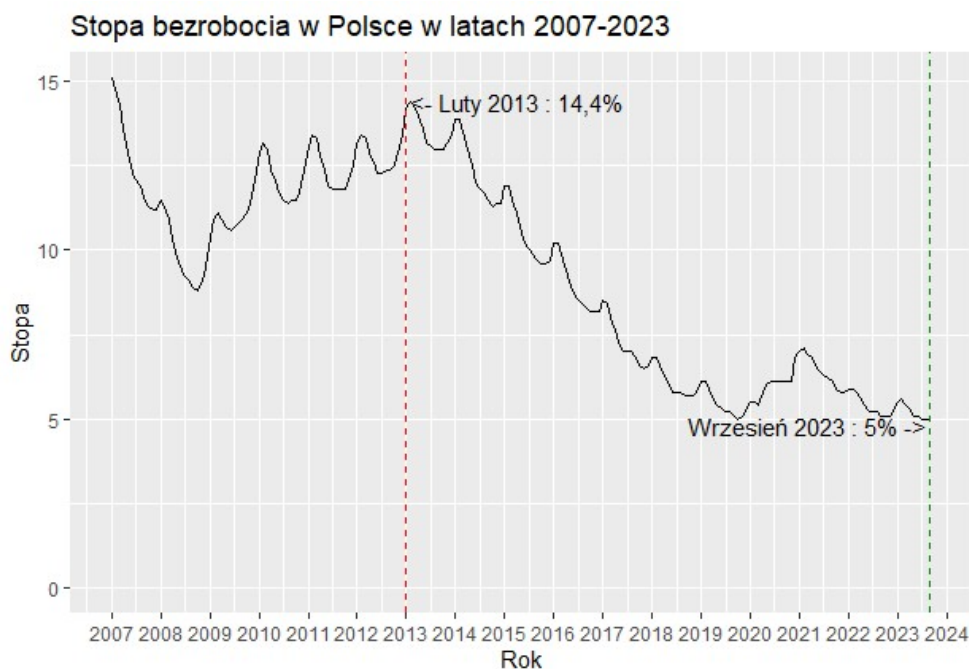
Na twitterze Premiera Mateusza Morawieckiego został udostępniony wizualizację stopy bezrobocia na podstawie danych z GUS-u, który bardziej przypomina szlaczek niż faktyczny wykres. Wykres nie ma ani osi poziomej, ani pionowej, zaznaczone są na niej tylko 2 skrajne wartości i nie ma on odpowiedniego tytułu. Została również zaznaczona jakaś pionowa prosta przerywana, która, jak można się tylko domyślać, oznacza zmianę rzędu (nie jest to początek tendencji spadkowej, gdyż ona ewidentnie zaczyna się wcześniej).

Źródło:

<https://x.com/PremierRP/status/1717130698371252711?s=20>,

Z <<https://twitter.com/MorawieckiM>> data publikacji: 25.10.2023

POPRAWIONA WIZUALIZACJA:



Poprawiony wykres, przedstawia kompletne dane od roku 2007 do 2023. Ma obydwie, dobrze opisane osie, a oś pionowa zaczyna się w zerze zapobiegając zaburzonemu odbiorowi danych. Wyróżnione daty zostały zaznaczone pionowymi prostymi, a ich wartości dodatkowo opisane na wykresie. Wykres jest bardziej czytelny i nie przedstawia faktycznych danych bez ich manipulacji.

3. Kod

Wykres 1:

```
gus %>% pivot_longer(!`rok/miesiac`, names_to = "miesiac", values_to = "procent") %>%
  na.omit() %>% mutate(rok_miesiac = paste(`rok/miesiac`, miesiac, sep = "_")) %>%
  filter(`rok/miesiac` > 2006) -> gus

gus %>% mutate(data = ymd(paste(`rok/miesiac`, miesiac, 1, sep = "-"))) %>%
  ggplot(aes(x = data, y = procent)) +
  geom_line(group=1) + scale_x_date(labels = date_format("%Y"), date_breaks = "1 year") +
  labs(title = "Stopa bezrobocia w Polsce w latach 2007-2023", x = "Rok", y = "Stopa") +
  geom_vline(aes(xintercept= ymd(paste(2013,1,1))), color="red", linetype="dashed") +
  geom_vline(aes(xintercept= ymd(paste(2023,9,1))), color="green4", linetype="dashed") +
  annotate("text", x = ymd(paste(2015,4,1)), y = 14.4, label = "<- Luty 2013 : 14,4%") +
  annotate("text", x = ymd(paste(2021,3,1)), y = 4.8, label = "Wrzesień 2023 : 5% ->") +
  expand_limits(x = ymd(paste(2007,1,1)), y = 0)
```

Wykres 2:

```
wybory <- data.frame(x = c("Prawo i Sprawiedliwość", "Koalicja Obywatelska",
  "Trzecia Droga", "Lewica", "Konfederacja", "Bezpartyjni", "inni"),
  y = c(36.8,31.6,13,8.6,6.2,2.4,1.4))
colnames(wybory) <- c("Ugrupowanie", "Procent")

wybory %>%
  mutate(Ugrupowanie = forcats::fct_reorder(wybory$Ugrupowanie,-wybory$Procent)) %>%
  ggplot(aes(x = Ugrupowanie, y = Procent)) + geom_col(fill = "purple4") +
  scale_x_discrete(guide = guide_axis(angle = 55)) +
  labs(title = "wybory 2023", subtitle = "Sondaż exit poll") +
  geom_text(aes(label = Procent), nudge_y = 2)
```