

Praca domowa 2

Jan Wojtas

Temat Pracy

Poniższa praca ma na celu poprawienie oraz udoskonalenie istniejącego wykresu, który z wizualnego i informatycznego punktu widzenia zawiera błędy i niedociągnięcia.

Poprawiana wizualizacja - informacje ogólne.

Źródło: <https://twitter.com/PremierRP/status/1363220255338868743>

Oryginalna wizualizacja:

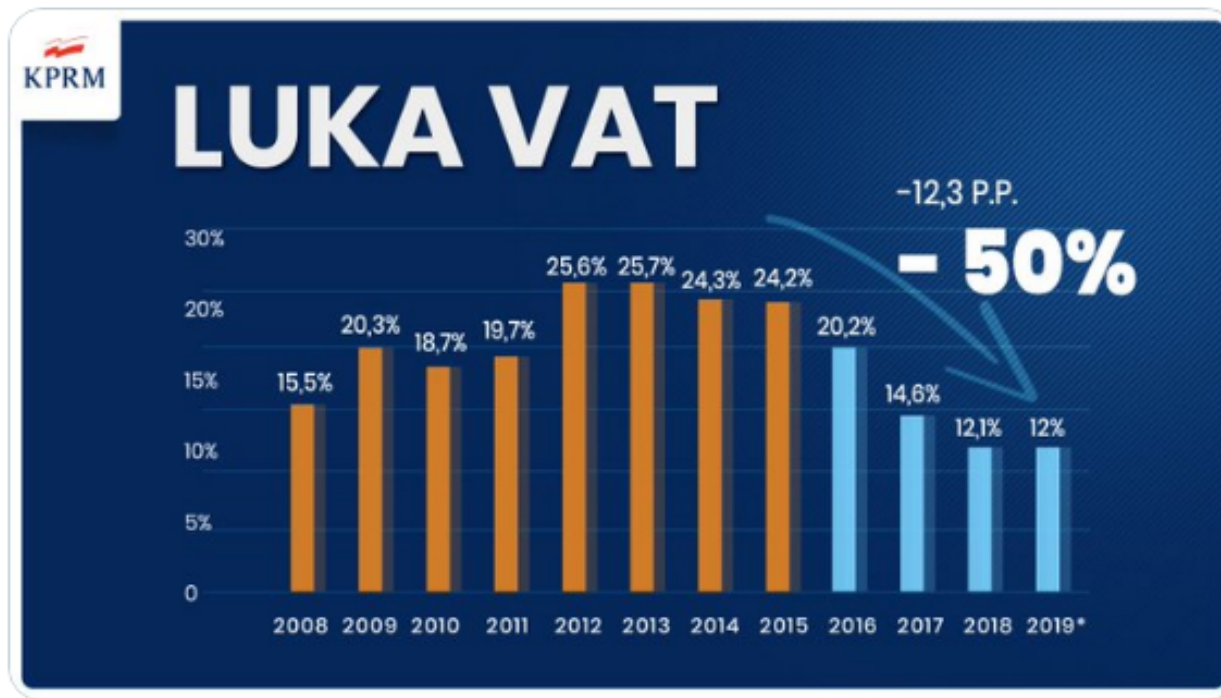


Figure 1: Wykres przedstawiający stan luki VAT w Polsce - lata 2008-2019

Powyższa wizualizacja pochodzi z Twittera Kancelarii Premiera Rzeczypospolitej Polskiej. Opisuje ona stan i zmianę rozmiaru luki VAT na przestrzeni lat 2008-2019. Wykres posiada jednak kilka niedoskonałości i błędów, które warto wymienić:

1. Informacja o zmniejszeniu się luki VAT o 12.3 punktów procentowych jest niezgodna z treścią wykresu: różnica między stanem z 2013 roku, w którym jej wartość osiągnęła 25.7%, a stanem z 2019 roku, gdzie wartość owej luki stanowiła 12% podatku wynosi 13.7 p.p..
2. Słupki wykresu są w 3D, co istotnie zaburza wizualny odbiór i prawdziwość wykresu. przykład: można odnieść wrażenie, że słupek z 2019 roku przewyższa słupek z 2018 roku, pomimo, że w istocie powinien być od niego niższy.
3. Miarka procentowa, znajdująca się po lewej stronie wizualizacji, jest nieprawidłowo skonstruowana - brakuje na niej wartości 25%, a pozostałe, obecne na niej wartości, są nieprawidłowo przesunięte względem poziomych linii, wyznaczających wartości co 5%.

Próba poprawy wizualizacji, uzasadnienie.

Poniżej znajduje się moja wizualizacja, w której podejmuję próbę udoskonalenia omawianego wykresu.

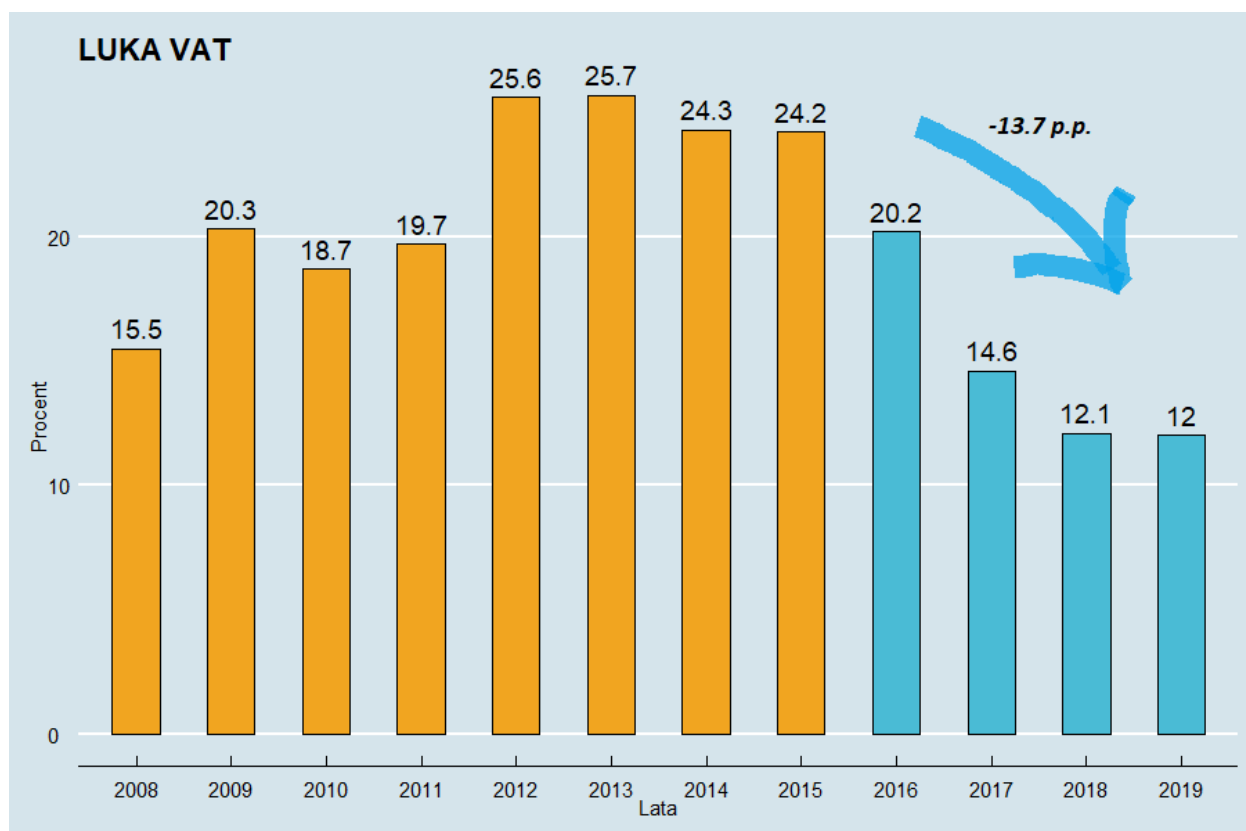


Figure 2: Poprawiony wykres

Wykres pozbywa się błędów, obecnych w pierwotnej wizualizacji. Nieprawdziwa informacja o spadku punktów procentowych zostaje zastąpiona prawidłową wartością, zgodną z treścią wykresu. Wykres jest 2D, co

zwiększa czytelność wykresu, słupki mają prawidłowe wysokości. Miarka procentowa, znajdująca się po lewej stronie wykresu, jest prawidłowo skorygowana.

Kod generujący wykres

```
1 library(ggplot2)
2 install.packages("ggthemes")
3 library(ggthemes)
4 lata <- c("2008", "2009", "2010", "2011", "2012", "2013", "2014", "2015", "2016", "2017", "2018", "2019")
5 wsp <- c("1", "1", "1", "1", "1", "1", "1", "2", "2", "2", "2")
6 wartosci_pr <- c(15.5, 20.3, 18.7, 19.7, 25.6, 25.7, 24.3, 24.2, 20.2, 14.6, 12.1, 12)
7
8 wykres <- data.frame(lata, wartosci_pr, wsp)
9
10 ggplot(wykres, aes(x = lata, y = wartosci_pr, fill = wsp)) + theme_economist() +
11   geom_col(width = 0.5, col = "black", show.legend = FALSE) + labs(title = "LUKA VAT", x = "Lata", y = "Procent ") +
12   scale_fill_manual(values = c("#F1A520", "#4ABBD5")) +
13   geom_text(aes(label = wartosci_pr), size = 5, vjust = -0.5)
14
15
```

Figure 3: Kod