

Praca domowa 2

Mikołaj Piórczyński

1. Wstęp

Celem pracy domowej nr. 2 było znalezienie w prasie, telewizji lub internecie wizualizacji opublikowanej w ciągu ostatnich 2 tygodni i poprawienie jej z wykorzystaniem pakietu `ggplot2`. Niniejszy raport przedstawia efekty tej pracy.

2. Oryginalny wykres

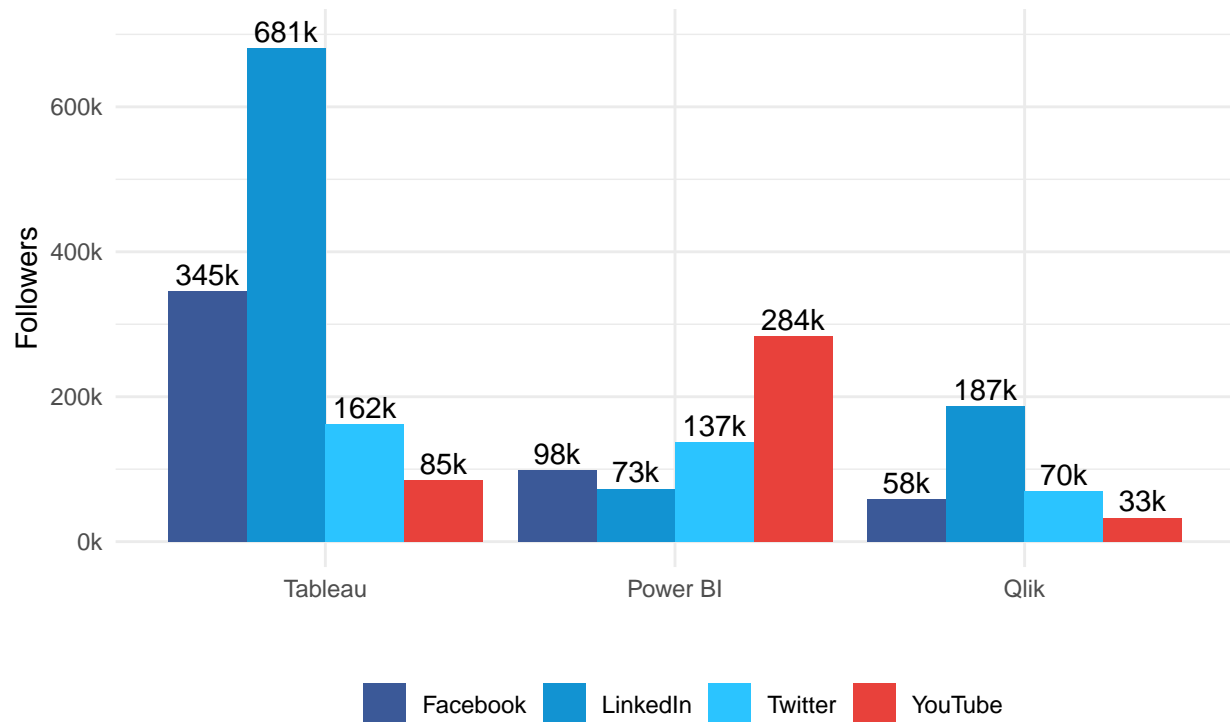


Figure 1: Znaleziony wykres, źródło: https://www.reddit.com/r/dataisbeautiful/comments/qjpb3/followers_of_the_top_3_business_intelligence

3. Poprawiony wykres

Where #BI hangs out

Followers of top 3 business intelligence tools in 4 chosen social medias in thousands



4. Kod źródłowy

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

df <- data.frame(tool = rep(c("Tableau", "Power BI", "Qlik"), each = 4),
                 media = rep(c("LinkedIn", "Facebook", "Twitter", "YouTube"), 3),
                 followers = c(681, 345, 162, 85, 73, 98, 137, 284, 187, 58, 70, 33))

df %>%
  mutate(tool = forcats::fct_reorder(tool, -followers)) %>%
  ggplot(aes(x = tool, y = followers, fill = media)) +
  geom_col(position = 'dodge') +
  geom_text(aes(
    label = paste0(followers, 'k'),
    position = position_dodge(width = 0.9),
    vjust = -0.3
  )) +
  labs(
```

```

title = "Where #BI hangs out",
subtitle = "Followers of top 3 business intelligence
           tools in 4 choosen social medias in thousands",
x = "",
y = "Followers"
) +
scale_y_continuous(limits = c(0, 700), labels = label_number(suffix = "k")) +
scale_fill_manual(values = c("#3b5998", "#1293d2", "#2bc4ff", "#e8413b")) +
theme_minimal() +
theme(
  plot.title = element_text(size = 18, face = 'bold'),
  legend.title = element_blank(),
  legend.position = "bottom"
)

```

5. Uzasadnienie

Dane na oryginalnym wykresie zostały zwizualizowane w formie wykresu **circular packing**, korzystającego z pól kół do przedstawienia danych o followersach, co utrudnia porównywanie przedstawionych wartości, ponieważ człowiek ma problem z porównywaniem pól figur. Dodatkowo umieszczenie ikonk portali wewnątrz kół zaburza postrzeganie ich pól. Warto również zauważyć, że wykresy typu **circular packing** zwykle służą przedstawieniu hierarchi pomiędzy grupami w danych, natomiast w tym przypadku takiej hierarchi nie ma, przez co jego zastosowanie wydaje się zbędne. Zmiana typu wykresu na wykres słupkowy pozwala na łatwe i szybkie porównywanie wartości pomiędzy poszczególnymi grupami poprzez porównywanie długości słupków. Dodatkowo dla ułatwienia nad słupkami zostały umieszczone dokładne wartości followersów, nieczytelne na oryginalnym wykresie w przypadku małych kół, a także podtytuł wykresu został zmieniony na bardziej informatywny.