

# HW3

Mateusz Kiszka

1 kwietnia 2024

## 1 Wizualizacja źródłowa

### 1.1 Źródło

Wizualizacja źródłowa pochodzi z bloga FourWeekMBA, zajmującego się analizą modeli biznesowych różnych firm oraz oferującego kursy w tej tematyce. Wybrana przeze mnie wizualizacja pochodzi z opublikowanej 20 marca 2024 roku analizy modelu biznesowego firmy Microsoft: <https://fourweekmba.com/who-owns-microsoft/>

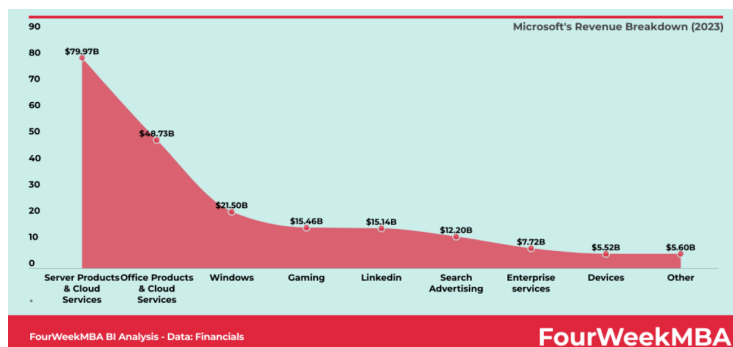


Figure 1: Wizualizacja źródłowa

### 1.2 Co jest nie tak?

Na pierwszy rzut oka wykres wydaje się być estetyczny i całkiem czytelny (może poza skalą na osi OY, która jest dość daleko od wykresu co utrudnia odczytywanie wartości). Jednak gdy przyjrzymy się na to co znajduje się na osi OX wykresu zrozumiemy jak bardzo bezsensowna jest ta wizualizacja. Autor chciał tutaj przedstawić rozkład przychodów spółki Microsoft, w zależności od oferowanej przez nich kategorii produktu. Jest to więc wykres mający przedstawić zależność zmiennej ilościowej od nominalnej. Zastosowany tutaj wykres liniowy sugeruje czytelnikowi ciągłość osi OX, co jest oczywiście bardzo dużym błędem.

## 2 Jak poprawić tę wizualizację?

Chcąc poprawić tę wizualizację musimy zacząć od zmiany typu wykresu. Postawiłem na wykres kolumnowy ustawiony poziomo. Zdecydowałem się na taki rodzaj wykresu z kilku powodów. Po pierwsze zmienne nominalne mają stosunkowo długie nazwy, które najłatwiej będzie odczytać gdy będą ustawione poziomo - w sposób naturalny dla czytelnika. Po drugie wykres kolumnowy dobrze oddaje skalę wielkości poszczególnych wartości. Łatwo zauważyć na nim relacje między przychodami z różnych źródeł. Oś OY jest tutaj ustawiona malejąco z wyjątkiem kategorii "Other" - mimo, że przychód z niej jest większy niż z "Device", to jako że jest to kategoria zbiorcza zawierająca wszystkie pozostałe źródła postanowiłem ją mimo wszystko umieścić na końcu. Pozostawiłem również szatę graficzną z wykresu źródłowego, ponieważ zauważyłem, że większość wykresów na tym blogu jej używa, więc chciałem zachować spójność poprawionego przeze mnie wykresu z pozostałymi wizualizacjami na blogu.

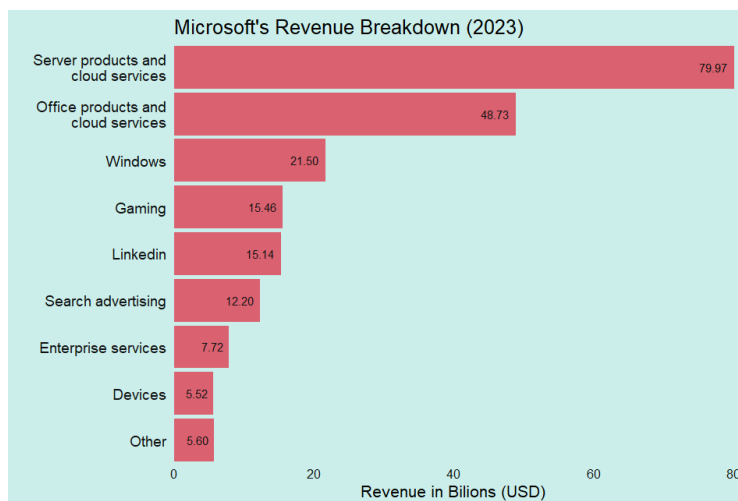


Figure 2: Poprawiona wizualizacja

### 3 Kod źródłowy

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

categories <- c("Server products and cloud services",
               "Office products and cloud services",
               "Windows", "Gaming", "Linkedin",
               "Search advertising", "Enterprise services",
               "Devices", "Other")

values <- c(79.97, 48.73, 21.50, 15.46, 15.14, 12.20, 7.72, 5.52, 5.60)

microsoft_data <- data.frame(categories, values)
new_labels <- c("Other", "Devices", "Enterprise services",
               "Search advertising", "Linkedin", "Gaming",
               "Windows", "Office products and\ncloud services",
               "Server products and\ncloud services")

microsoft_data %>%
  ggplot(aes(x = fct_rev(factor(categories, levels = categories)),
            y = values)) +
  geom_col(fill = "#d96170", color = "#d96170") +
  labs(title = "Microsoft's Revenue Breakdown (2023)",
       x = "",
       y = "Revenue in Bilions (USD)") +
  theme_minimal(base_size = 16) +
  theme(axis.text.x=element_text(color = "#0f1211",
                                size = 12)) +
  theme(axis.text.y=element_text(color = "#0f1211",
                                size = 14)) +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank()) +
  geom_text(aes(label = format(round(values, 2), nsmall = 2)),
            color = "#0f1211",
            position = position_dodge2(width = 0.5), hjust = 1.2, size 4)+
  scale_x_discrete(expand = c(0,0), labels = new_labels) +
  scale_y_continuous(expand= c(0,0)) +
  expand_limits(y = (c(0,80))) +
  coord_flip() +
  theme(panel.background = element_rect(fill = '#cbeeeb', colour = '#cbeeeb')) +
  theme(plot.background = element_rect(fill = '#cbeeeb', colour = '#cbeeeb'))
```