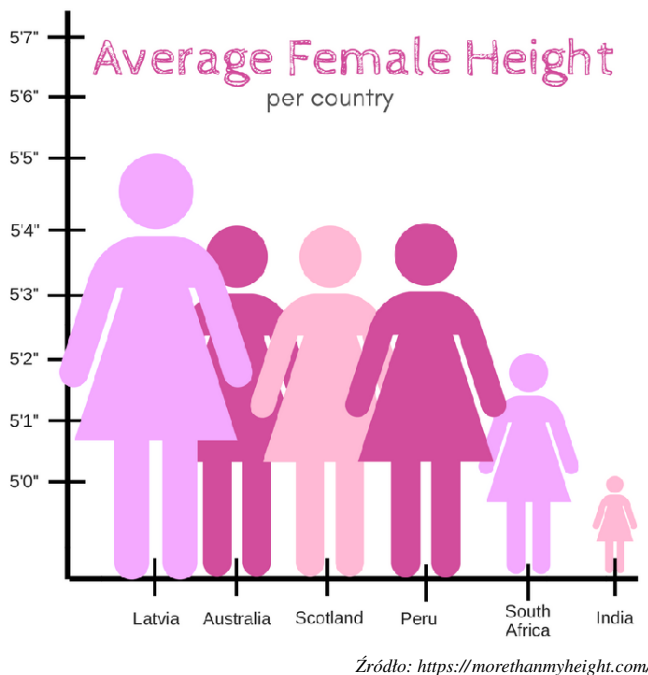


Praca domowa nr 2

Celem pracy domowej jest znalezienie źle zaprojektowanej wizualizacji danych (opublikowanej w ciągu ostatnich dwóch tygodni), opisanie jakie błędy zawiera, oraz poprawienie jej. Przeszukując odmęty Internetu natknąłem się na wiele grafik, którym swoim kiepskim dizajnem zaczęły rozprzestrzeniać się w formie memów (przykład rysunek 1a). Były też takie, które w sposób interesujący pokazywały dość dosłowny sposób wizualizacji danych (rysunek 1b)



(a) Średni wzrost kobiet w wybranym kraju

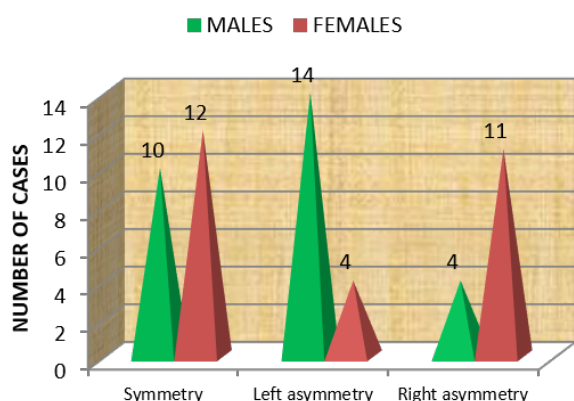


(b) Wykres słupkowy przedstawiający jaką część drzwi pomalowałem

Rysunek 1: Przykłady wykresów, które przykuły moją uwagę na tyle, że postanowiłem się nimi podzielić

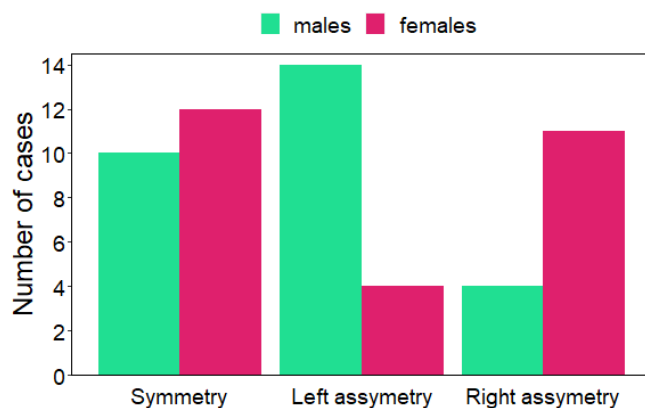
Żaden wykres jednak nie spełniał wymogów pracy domowej. Wtedy natknąłem się na Indyjskie czasopismo *International Journal of Research in Pharmaceutical Sciences*, którego *impact factor* wynosi zaledwie 0,24¹. Zawiera ono istną kopalnię źle zaprojektowanych wizualizacji. Trójwymiarowe wykresy słupkowe były modne na początku XXI wieku, kiedy to głównym narzędziem wizualizacji danych był Excel 2003. Czas na zmianę. Przedstawiam wam wykres piramidowy (rysunek 2a).

¹wg https://www.resurchify.com/all_ranking_details_2.php?id=25316



Źródło: <https://pharmascopie.org/ijrps/article/view/3267/8020#figures>

(a) Wykres przed liftingiem, grafika została opublikowana 01.10.2020



Źródło: dane pochodzą z [1], wizualizacja opracowana samodzielnie

(b) Wykres po liftingu

Rysunek 2: Oba wykresy przedstawiają porównanie wzorów zatoki czołowej

Poprawiony wykres przedstawiłem na rysunku 2b. Jak widać uległ on kilku przemianom, ponieważ oryginalna grafika posiada wiele błędów, a mianowicie:

1. trójwymiarowe przedstawienie słupków, które w zasadzie nie są słupkami a ostrosłupami czworokątnymi. Cel przedstawienia danych w ten sposób pozostaje dla mnie zagadką. Perspektywa ta sprawia, że trudno odczytać dokładne wartości danych. Autorzy jednakże umieścili wartości liczbowe na czubkach piramidy, co rozwiewa wszelkie wątpliwości dotyczące wysokości tych starożytnych grobowców. Wykres jednak zawiera liczby całkowite, które w sposób prosty można przedstawić na wykresie nie umieszczając jednocześnie zbędnych liczb
2. tło jest moim zdaniem zbędnym ozdobnikiem, dlatego się go pozbyłem. Podobnie usunąłem poziome linie pomocnicze

Tabela 1: Kod źródłowy generujący wykres

```
library(ggplot2)
library(colortools)

setwd(dirname(rstudioapi::getActiveDocumentContext()$path))

frontalSinusData <- data.frame(number = c(10, 12, 14, 4, 4, 11),
                                gender = factor(c("male", "female", "male", "female",
                                                    "male", "female")),
                                pattern = factor(c("Symmetry", "Symmetry",
                                                    "Left_assymetry", "Left_assymetry",
                                                    "Right_assymetry", "Right_assymetry")))

frontalSinusData$gender <- factor(frontalSinusData$gender,
                                  levels = c("male", "female"))

frontalSinusData$pattern <- factor(frontalSinusData$pattern,
                                   levels = c("Symmetry", "Left_assymetry",
                                              "Right_assymetry"))

plotColor <- complementary("#DF206D")

betterPlot <- ggplot(data = frontalSinusData, aes(x = pattern, y = number, fill = gender)) +
  geom_col(position = "dodge") +
  xlab("") +
  ylab("Number_of_cases") +
  theme(
    text = element_text(size = 24),
    axis.text.x = element_text(vjust = -1),
    axis.text = element_text(color = "black"),
    axis.ticks.length.x = unit(0, "cm"),
    panel.border = element_rect(fill = NA),
    panel.background = element_blank(),
    legend.title = element_blank(),
    legend.position = "top"
  ) +
  scale_fill_manual(values = c(plotColor[2], plotColor[1]),
                    breaks = c("male", "female"),
                    labels = c("males", "females")) +
  scale_y_continuous(expand = c(0, 0),
                     limits = c(0, max(frontalSinusData$number) + 0.5),
                     breaks = seq(0, 16, 2))

png(file = "better_plot.png", width = 574, height = 381)
print(betterPlot)
dev.off()
```

Bibliografia

- [1] Pinki Rai, Amit Kumar, Saxena Namita, Mehrotra, and Prachi Saffar Aneja. Morphometric analysis of frontal sinus dimensions using digital radiographs. *International Journal of Research in Pharmaceutical Sciences*, 11(4), 2020. 2