

Homework 2

Mateusz Krzyżiński

18.10.2020

Wprowadzenie

Celem pracy domowej było przygotowanie wizualizacji utworzonej z wykorzystaniem `ggplot2` na podstawie wizualizacji dostępnej w prasie, telewizji lub internecie, opublikowanej w ciągu ostatnich 2 tygodni oraz ewentualna poprawa oryginalnej wizualizacji.

Wizualizacja źródłowa

Wizualizacją, którą wybrałem do poprawy jest wykres, który pojawił się na stronie internetowej Światowego Forum Ekonomicznego w artykule z 13 października zatytułowanym *Chart of the Day: This is how the \$88 trillion global economy is spread around the world* - [link](#).

Motywacja - czyli co jest nie tak z wizualizacją?

Przedstawiona wizualizacja jest niejako wariacją na temat wykresu kołowego, w którym koło zostało podzielone na wiele różnego rodzaju kształtów/wielokątów. Ich wielkość ma symbolizować wielkość PKB poszczególnych państw. Jednak jest to trudne w odbiorze - biorąc pod uwagę dwa podobnej wielkości obszary, na pierwszy rzut oka ciężko jest bowiem określić, który z nich jest większy. Takie przedstawienie zaciera też różnice pomiędzy wielkościami. Co więcej, z wykresu nie da się w łatwy sposób odczytać kolejności największych światowych gospodarek (biorąc pod uwagę PKB), a duże rozdrobnienie (duża ilość uwzględnionych państw) wpływa na to, że konieczne było zastosowanie linii pomocniczych.

Rezultat - czyli co udało się poprawić?

Stworzony wykres słupkowy jest znacznie łatwiejszy do odczytania. Wielkości PKB poszczególnych państw są symbolizowane przez długości słupków, dlatego łatwo jest odczytać, które z nich mają większy, a które mniejszy wskaźnik. Co więcej, państwa zostały ustawione w kolejności od największego do najmniejszego PKB. Dla czytelności wybrano także mniejszą liczbę państw (te, których wskaźnik stanowi ponad 0,5% w światowym PKB). Każdy słupek otrzymał także stosowną etykietę - wielkości (w bilionach (10^{12} - w krótkiej skali *trillion*) i procentowego udziału. Ponadto dobrano kolory, które intuicyjnie odpowiadają regionom świata (flaga olimpijska i popularne oznaczenia na mapach).

Gross Domestic Product (GDP) by Country 2019



2

Kod źródłowy

```
# Ładowanie pakietów
library(dplyr)
library(ggplot2)
library(formattable)

# Ładowanie i przygotowanie danych (link do nich znajduje się w artykule)
data <- read.csv("gdp_the_world_bank_data.csv", skip = 4)
metadata <- read.csv("gdp_the_world_bank_metadata.csv")

metadata <- metadata %>% select(Country.Code, Region)
data %>%
  select(Country.Name, Country.Code, X2019) %>%
  inner_join(metadata) %>%
  na.omit() %>%
  filter(Region != "") %>%
  arrange(-X2019) %>%
  mutate(X2019_pct = X2019/sum(X2019)) -> data

rest_of_the_world <- data %>% filter(X2019_pct < 0.005)

data %>%
  filter(X2019_pct > 0.005) %>%
  add_row(Country.Name = "Rest of the World", Region = "Rest of the World",
          X2019 = sum(rest_of_the_world$X2019),
          X2019_pct = sum(rest_of_the_world$X2019_pct)) %>%
  mutate(X2019 = X2019/(10^12)) -> final_data
final_data$Country.Name <- factor(final_data$Country.Name,
                                  levels = rev(final_data$Country.Name))

# Tworzenie wykresu
ggplot(final_data, aes(x = Country.Name, y = X2019, fill = Region,
  label = paste( sprintf("%0.2f", round(X2019, digits = 2)),
    " (", percent(X2019_pct), "%)", sep=""))) +
  geom_bar(stat = "identity") + #słupki
  geom_text(size = 3.5, hjust=-0.1) + #etykiety słupków
  xlab("Countries") + #tytuły osi i wykresu
  ylab("Gross Domestic Product (Trillions current US$)") +
  ggtitle("GDP by Country in 2019") +
  coord_flip(ylim = c(1, 25)) + #obrót
  theme_minimal() + #wygląd
  theme(legend.position="bottom",
        plot.title = element_text(hjust = 0.5, face = "bold", size=15),
        plot.margin = margin(0.5, 0.5, 0.5, 0.5, "cm"),
        axis.ticks = element_blank(),
        panel.grid.major.y = element_blank()) +
  scale_fill_manual(values = c("#fcd703", "#0a51c4", "#178f13", "#b80071",
    "#bf0000", "grey", "#d97400", "#9300b8"))

# Zapisanie wykresu
ggsave("good_chart.png", width = 22, height = 24, units = "cm")
```

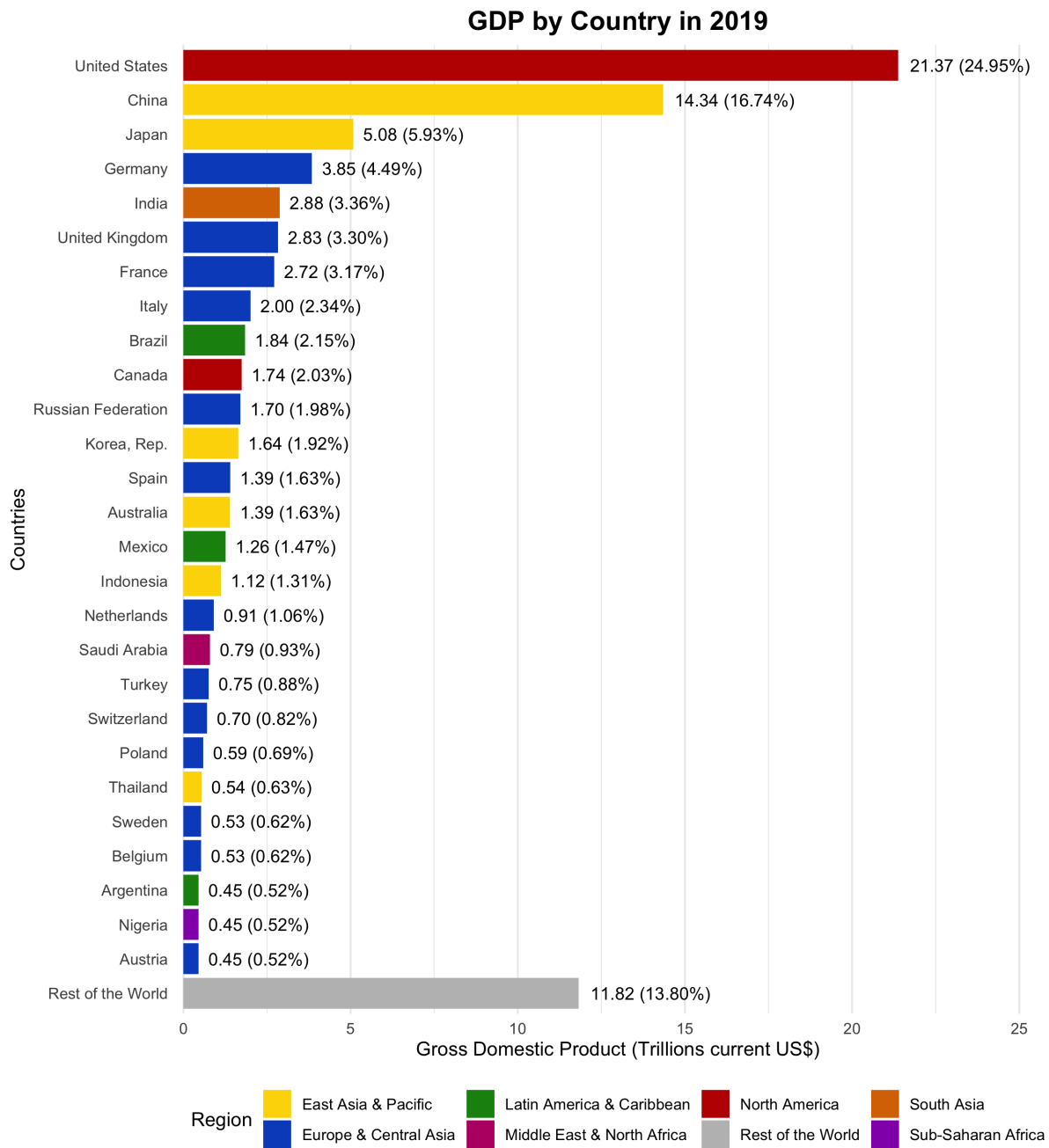


Figure 2: Poprawiona wizualizacja