

# Projekt - wykonanie mapy

Michał Celiński

2023-04-14

## Cel i temat projektu

Celem projektu było stworzenie dowolnej mapy obrazującej wybrane zagadnienie. Wybór padł na graficzne przedstawienie liczby goli zdobytych przez Roberta Lewandowskiego podczas meczów wyjazdowych w europejskich pucharach.

## Praca nad projektem

Na początku musiałem się zorientować czy w ogóle istnieją dane dotyczące tematu projektu, żebym nie musiał ręcznie tworzyć ramki danych. Na szczęście z pomocą przyszła strona sportowo - statystyczna *fbref.com*, z której pobrałem plik csv zawierający wszystkie bramki napastnika reprezentacji Polski zdobywane od czasu gry w Ekstraklasie. Następnie pozostała mi już jedynie praca z ramką danych uzyskaną z pliku csv, starałem się zautomatyzować wszystkie procedury, jednak w dwóch przypadkach musiałem “ręcznie” dodawać wiersze bądź kolumny. Raz też skorzystałem z *ChatGPT*, który utworzył mi plik csv przyporządkowując klubom, przeciwko którym Lewandowski strzelał gole współrzędne geograficzne miast, w których występują, abym mógł zaznaczyć te punkty na mapie.

## Skrypt

```
library(dplyr)
```

```
##
```

```
## Dołączanie pakietu: 'dplyr'
```

```
## Następujące obiekty zostały zakryte z 'package:stats':
```

```
##
```

```
##      filter, lag
```

```
## Następujące obiekty zostały zakryte z 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
library(maps)
```

```
## Warning: pakiet 'maps' został zbudowany w wersji R 4.2.3
```

```
library(mapdata)
```

```
## Warning: pakiet 'mapdata' został zbudowany w wersji R 4.2.3
```

```
library(ggplot2)
library(stringr)
library(ggrepel)
```

```
## Warning: pakiet 'ggrepel' został zbudowany w wersji R 4.2.3
```

```
library(patchwork)
```

```
## Warning: pakiet 'patchwork' został zbudowany w wersji R 4.2.3
```

```
bramki_lewego <- read.csv(file = 'bramki_lewego.csv')
```

```
df1 <- bramki_lewego %>%
  filter(Venue == "Away") %>%
  filter(Comp == "Europa Lg" | Comp == "Champions Lg" | Comp == "UEFA Cup") %>%
  select(Squad, Opponent) %>%
  group_by(Opponent) %>%
  summarise(Gole = length(Opponent))
```

```
Kraje <- c("Greece", "Netherlands", "Belgium", "UK", "Spain", "Portugal", "UK", "Croatia",
           "Ukraine", "Norway", "Azerbaijan", "Italy", "France", "UK", "UK", "France", "Greece", "Netherlands",
           "Austria", "Spain", "Serbia", "Italy", "Ukraine", "UK", "Russia")
```

```
df1 <- df1 %>%
  cbind(Kraje) %>%
  group_by(Kraje) %>%
  summarise(ile = sum(Gole))
colnames(df1) <- c("region", "ZdobyteBramki")
df1$ZdobyteBramki <- as.character(df1$ZdobyteBramki)
```

```
country <- map_data("world")
```

```
ramka_do_wykresu <- country %>%
  left_join(df1)
```

```
## Joining with 'by = join_by(region)'
```

```
ramka_do_wykresu$ZdobyteBramki <- replace(ramka_do_wykresu$ZdobyteBramki, is.na(ramka_do_wykresu$ZdobyteBramki), 0)
```

```
punkty2 <- bramki_lewego %>%
  filter(Venue == "Away") %>%
  filter(Comp == "Europa Lg" | Comp == "Champions Lg" | Comp == "UEFA Cup") %>%
  select(Opponent)
```

```
#listę drużyn z punkty2 przekazuję do ChatGPT, którego proszę o wygenerowanie ramki danych zawierających
```

```

miasta_goli <- read.csv("miasta_goli.csv")
miasta_goli <- miasta_goli %>%
  group_by(Miasto) %>%
  summarise(ile = length(Miasto))

miasta_goli$label = paste0(miasta_goli$Miasto, sep = " - ", miasta_goli$ile)
miasta_goli_pom <- read.csv("miasta_goli.csv")

miasta_goli_gotowe <- miasta_goli %>%
  left_join(miasta_goli_pom) %>%
  select(label, long, lat) %>%
  unique()

```

```
## Joining with 'by = join_by(Miasto)'
```

```

## Warning in left_join(., miasta_goli_pom): Each row in 'x' is expected to match at most 1 row in 'y'.
## i Row 1 of 'x' matches multiple rows.
## i If multiple matches are expected, set 'multiple = "all"' to silence this
## warning.

```

```

wg_klubow <- read.csv("miasta_goli.csv")
liczba_goli <- wg_klubow %>%
  group_by(Klub) %>%
  summarise(ile_goli = length(Klub))
wg_klubow_wspl <- wg_klubow %>%
  left_join(liczba_goli) %>%
  unique() %>%
  mutate(label = paste0("vs. ", Klub, " - ", ile_goli))

```

```
## Joining with 'by = join_by(Klub)'
```

```
#tworzę mapę
```

```

mapa_bramek <- ggplot() +
  geom_polygon(data = ramka_do_wykresu, aes(x = long, y = lat, group = group, fill = ZdobyteBramki), color = "black") +
  geom_point(data = miasta_goli_gotowe, aes(x = long, y = lat)) +
  geom_label_repel(data = wg_klubow_wspl, aes(x = long, y = lat, label = label), max.overlaps = 30, min.segment.length = 5) +
  theme_void() +
  coord_fixed(ratio=1.3, xlim = c(-13, 50), ylim = c(34,65)) +
  scale_fill_discrete(type = c('white', '#edf8e9', '#c7e9c0', '#a1d99b', '#74c476', '#31a354', '#006d2c'),
    name = "Liczba\bramek") +
  labs(subtitle = "Z podziałem na kraje, w których strzelał gole", caption = "") +
  scale_x_continuous(expand = c(0,0)) +
  scale_y_continuous(expand = c(0,0)) +
  theme(plot.background = element_rect(fill = "white"),
    legend.position = c(0.95, 0.75),
    legend.margin = margin(10,10,10,10),
    legend.background = element_rect(fill = "white", color = "black"),
    plot.margin = margin(t = -35),
    plot.subtitle = element_text(hjust = 0.5))

```

### ##wykres słupkowy

```
df3 <- bramki_lewego %>%
  filter(Venue == "Away") %>%
  filter(Comp == "Europa Lg" | Comp == "Champions Lg" | Comp == "UEFA Cup") %>%
  select(Date, Squad)

df3a <- str_split(df3$Date, pattern = "-", n = 3, simplify = TRUE)
df3 <- cbind(df3, df3a)
colnames(df3) = c("Data", "Klub", "Rok", "Miesiac", "Dzien")
df3$Rok <- as.numeric(df3$Rok)
df3$Miesiac <- as.numeric(df3$Miesiac)

df3 <- df3 %>%
  mutate(sezon1 = ifelse(Miesiac < 7, Rok - 1, Rok)) %>%
  mutate(sezon2 = sezon1 + 1)
df3$sezon2 <- substr(df3$sezon2, 3, 4)
df3$sezon <- paste(df3$sezon1, df3$sezon2, sep = "/")
df3 <- df3 %>%
  select(Klub, sezon)
df3pom <- df3 %>%
  group_by(sezon) %>%
  summarise(liczba_goli = length(sezon))

df3 <- df3 %>%
  left_join(df3pom) %>%
  unique()
```

## Joining with 'by = join\_by(sezon)'

*# ręcznie dodaję do ramki danych wiersz zawierający sezon, w którym RL9 nic nie strzelił*

```
pusty_sezon <- data.frame(Klub = "Dortmund", sezon = "2010/11", liczba_goli = 0)
df3 <- rbind(df3, pusty_sezon)
df3$Klub <- factor(df3$Klub, levels = c("Lech Poznań", "Dortmund", "Bayern Munich", "Barcelona"))
```

### *# tworzę wykres słupkowy*

```
wykres_bramek <- df3 %>%
ggplot(aes(x = sezon, y = liczba_goli, fill = Klub)) +
  geom_col(width = 0.6) +
  scale_y_continuous(breaks = 1:9, expand = c(0,0), limits = c(0, 10)) +
  scale_x_discrete(guide = guide_axis(angle = 45)) +
  scale_fill_manual(values = c("dodgerblue", "#FDE100", "#DC052D", "#004D98")) +
  theme_bw() +
  labs(x = "", y = "Liczba bramek", subtitle = "Z podziałem na sezony i kluby, które reprezentował") +
  theme(panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank(),
        panel.grid.minor.y = element_blank(),
        axis.text = element_text(size = 12),
        plot.subtitle = element_text(hjust = 0.5),
        axis.title = element_text(size = 11))
```

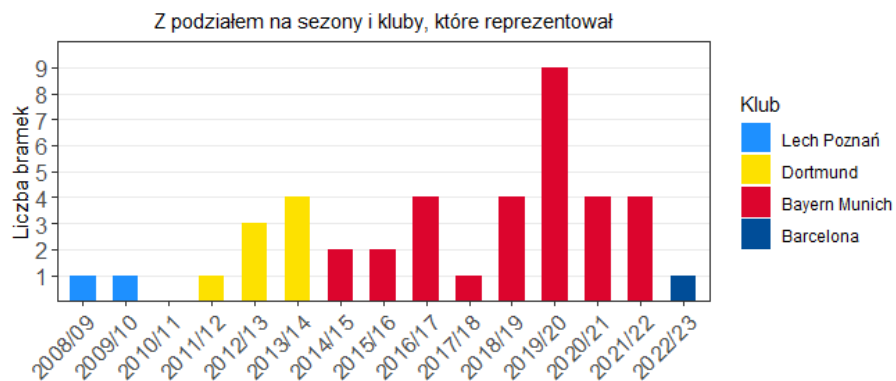
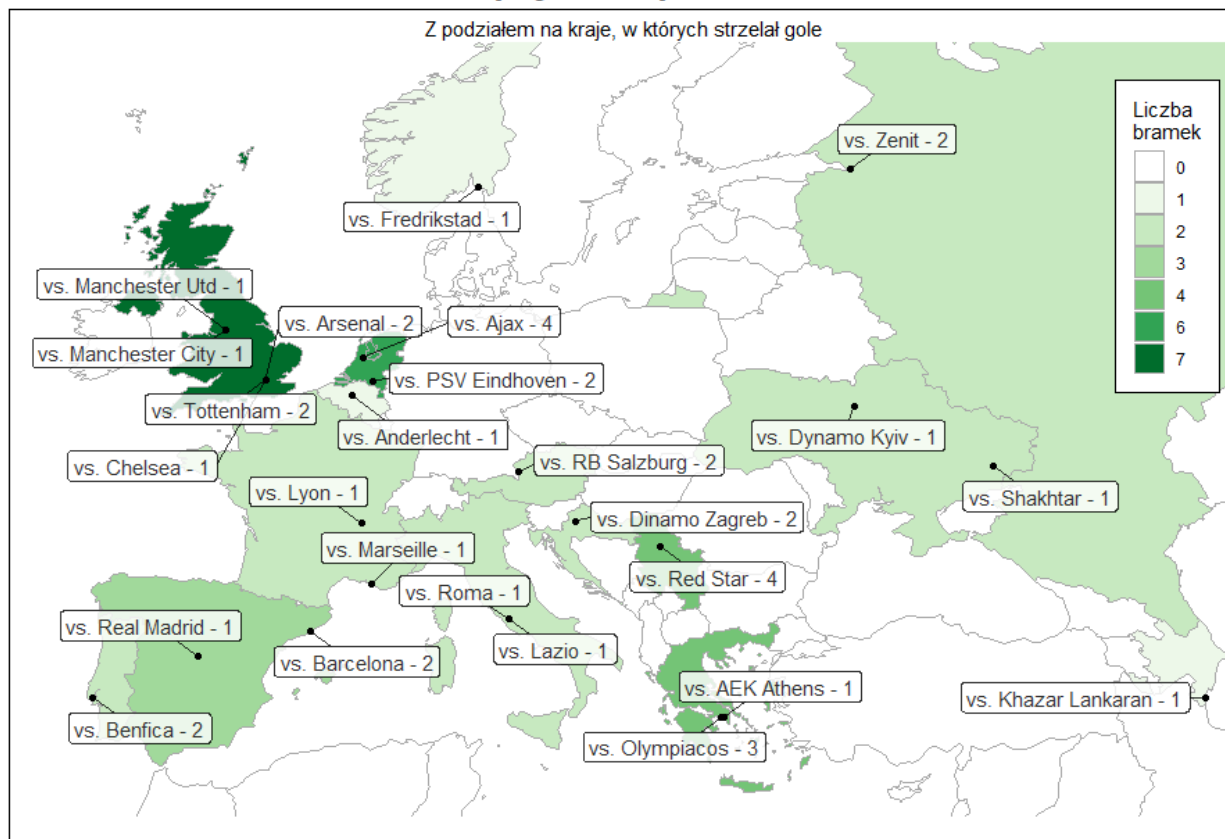
*# patchwork - łącznie mapy i wykresu słupkowego*

```
ostateczny_wynik <- mapa_bramek / (plot_spacer() + wykres_bramek + plot_spacer() + plot_layout(widths =  
  plot_layout(heights = c(0.75, 0.25)) +  
  plot_annotation(title = "Bramki Roberta Lewandowskiego w meczach wyjazdowych \neuropejskich pucharów"  
    caption = "Dane: fbref.com, Autor: Michał Celiński",  
    theme = theme(plot.title = element_text(size = 20, hjust = 0.5, face = "bold"))))
```

## Wynik pracy

Niestety, aby wykres dobrze wyglądał powinien być określonej wielkości i w odpowiednich proporcjach. Zapisalem go więc do pliku png, który wyświetlam poniżej:

## Bramki Roberta Lewandowskiego w meczach wyjazdowych europejskich pucharów



Dane: fbref.com, Autor: Michał Celiński