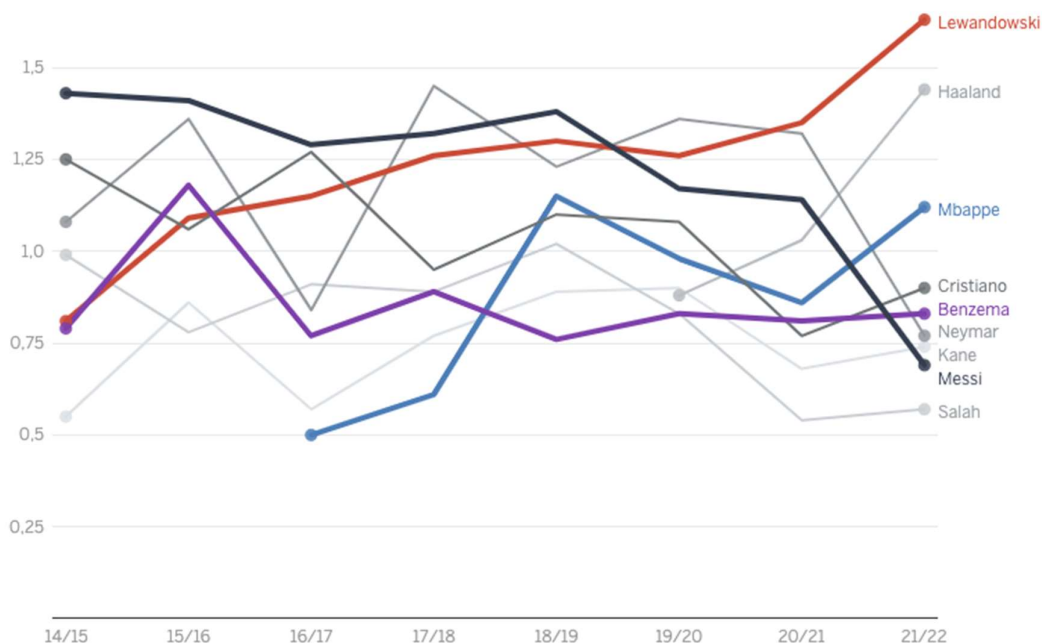


Poprawa wizualizacji związanej z przyznaniem Złotej Piłki



Powyższy wykres pochodzi z Twittera jednego z polskich dziennikarzy sportowych.

Link do jego posta:

https://twitter.com/mati_swiecicki/status/1452176519779045376/photo/1

O ile sam dziennikarz nie jest szczególnie znaną postacią, to wykres został zamieszczony na jednej z popularniejszych stron informacyjnych o tematyce sportowej.

<https://www.sport.pl/pilka/7,64946,27725311,lewandowski-deklasuje-rywali-konkurenci-do-zlotej-pilki-daleko.html>

Na celu miał pokazanie dominacji Roberta Lewandowskiego, lecz uważam, że można sporo poprawić zarówno od strony merytorycznej jak i wizualnej.

Merytoryka:

- Tytuł wykresu jest niejednoznaczny „Suma spodziewanych bramek i asyst...” nie wyjaśnia, co będziemy porównywać
- Pojawia się statystyka z obecnego sezonu, gdzie niektórzy zawodnicy rozegrali zaledwie 4 z 30 spotkań ligowych.
- Po przeczytaniu artykułu i komentarzy można zrozumieć, że wartości na osi Y liczymy według wzoru: $\frac{\text{liczba goli} + \text{liczba asyst}}{\text{liczba meczów}}$ w danym sezonie. Po sprawdzeniu statystyk poszczególnych graczy (google stats) okazuje się, że żaden z wyników nie jest nawet

bliski pokazanemu na wykresie. Przykład: Leo Messi w tym sezonie nie zdobył żadnej bramki ani asysty w spotkaniach ligowych, jednak jego wynik to prawie 0,75 punktu.

- Złota Piłka jest przyznawana co roku za osiągnięcia w poprzednim sezonie ligowym(20/21). Nie jest ważny progres osiągnięty przez ostatnie lata

Wizualizacja:

- Brak tytułu na wykresie, brak podpisów przy obu osiach
- Kolory linii 5 z 9 zawodników są prawie identyczne (szare), zlewają się z tłem
- Najbardziej widoczny kolor czerwony został przypisany zawodnikowi, który ma być uznany przez twórcę za najlepszego; delikatna manipulacja na co patrzeć
- Zbyt duża liczba sklasyfikowanych zawodników utrudnia czytelność wykresu

Mój pomysł na poprawę wykresu

Propozycja 1

Wykres słupkowy porównujący zawodników w poszczególnych sezonach.

Kod (wczytanie tabeli z Excela, stworzenie dodatkowej kolumny, wykres1)

```
library(readxl)
library(dplyr)
library(ggplot2)
library(tidyverse)
tabela <- read_excel("Zestawienie_zawodnicy.xlsx") # stworzona tabela

tabela <- tabela %>%
  mutate(index=(Gole+Asysty)/Mecze)

tabela %>%
  group_by(Sezon) %>%
  mutate(pozycja=rank(-index)) %>%
  ggplot() +
  geom_col(aes(Sezon,index,fill=Zawodnik,group=pozycja),position =
"dodge")+ # sortowanie słupków

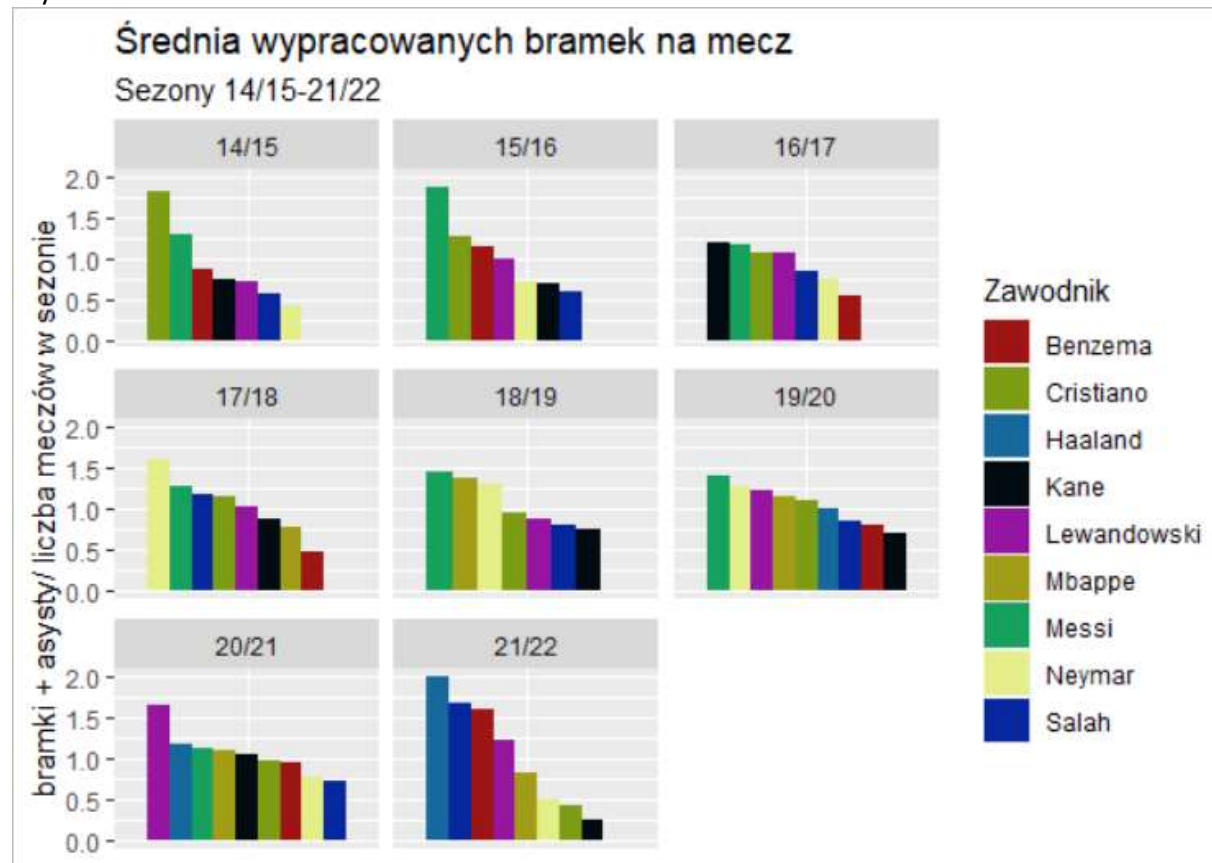
  scale_fill_manual(
    values =
c("#9d1515", "#7d9d15", "#15699d", "#020c12", "#9616a2", "#a29d16", "#16a25e", "#e
5ef8a", "#06279d")

  )+ # własne kolory
  facet_wrap(~Sezon,scales = "free_x")+

  labs(title = "Średnia wypracowanych bramek na mecz ",
        subtitle = "Sezony 14/15-21/22",
        x="",
        y="bramki + asysty/ liczba meczów w sezonie"
  )+

theme(axis.title.x=element_blank(),
      axis.text.x=element_blank(),
      axis.ticks.x=element_blank())
```

Wykres:

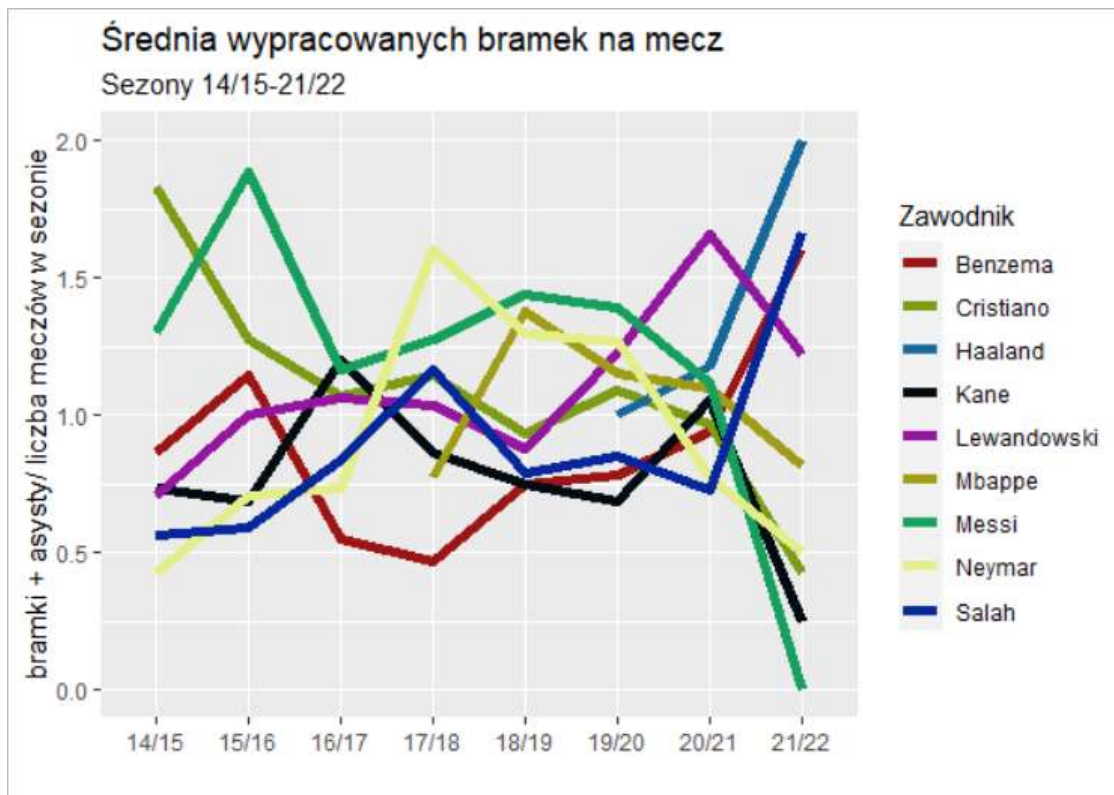


Propozycja 2

Wykres liniowy z poprawionymi kolorami, podpisami, grubością linii i odpowiednimi danymi

Kod (wykres 2):

```
tabela %>%
  ggplot(aes(x=Sezon,y=index,color=Zawodnik,group=Zawodnik))+
  scale_color_manual(
    values =
c("#9d1515", "#7d9d15", "#15699d", "#020c12", "#9616a2", "#a29d16", "#16a25e", "#e5ef8a", "#06279d")
  )+
  geom_line(size=2)+
  labs(title = "Średnia wypracowanych bramek na mecz ",
        subtitle = "Sezony 14/15-21/22",
        x="",
        y="bramki + asysty/ liczba meczów w sezonie"
  )
```



Dlaczego moje wykresy są lepsze?

Oba wykresy:

- Poprawa merytoryki; dokładny opis osi Y oraz tytuł i podtytuł wykresu
- Dane zgadzające się z rzeczywistością
- Różne kolory, nie zlewające się z tłem
- Wyraźna legenda
- Wykresy nie są stroniczne

Wykres1:

- Łatwiejsze porównanie zawodników w ostatnim, najbardziej istotnym sezonie(20/21)
- Bardziej czytelny wykres słupkowy

Wykres 2:

- Pogrubione linie pozwalają łatwiej odczytać wynik