

Praca domowa nr 3+4

Karol Degórski

11/01/2020

Badanie odczytywania danych z wykresu kołowego i słupkowego

Tematem mojej Pracy Domowej nr 3+4 jest zbadanie jak typ wykresu (kołowy, słupkowy) wpływa na odczytywanie prezentowanych wartości przez odbiorców oraz jak duży błąd mogą oni popełnić podczas ich odczytu. Do mojego badania wykorzystałem dane przedstawiające podział mandatów w Parlamencie Europejskim w 2019 roku. Pochodzą one ze strony: https://en.wikipedia.org/wiki/2019_European_Parliament_election. Wybrałem te dane, ponieważ zawierają one 8 różnych partii, a ponadto partie G/EFA i ID oraz ECR i NI (europosłowie bezpartyjni) mają taki sam procentowy udział. Pozwoliło mi to sprawdzić, czy osoby ankietowane (link do ankiety <https://docs.google.com/forms/d/e/1FAIpQLSfNSCzmCBWbpROnpNVt8Uahc18HP5UMfOmoJ4CXInjc1OfDLw/viewform>) są w stanie rozpoznać takie same kąty oraz takie same wysokości słupków. Na początku przygotowałem dane do sporządzenia wykresów.

```
library(ggplot2)
library(dplyr)

liczba_mandatow <- c(182, 154, 108, 74, 73, 62, 57, 41)
partia <- c("EPP", "S&D", "RE", "G/EFA", "ID", "ECR", "NI", "GUE/NGL")
dane <- data.frame(liczba_mandatow, partia)
dane <- dane %>% mutate(procent=round(liczba_mandatow/sum(liczba_mandatow)*100,0))

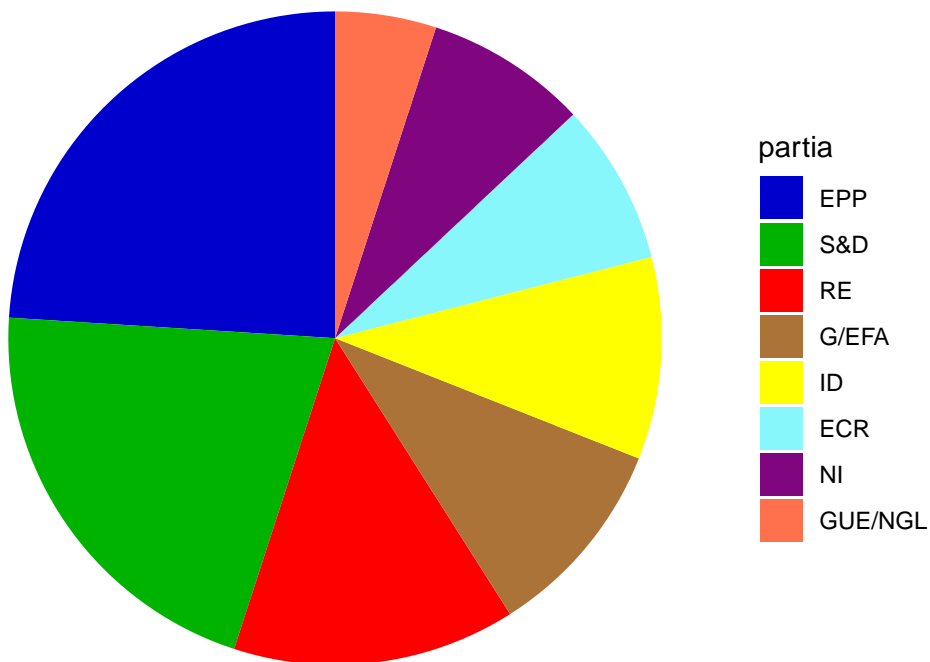
dane$partia <- factor(dane$partia, levels = c("EPP", "S&D", "RE", "G/EFA", "ID", "ECR",
                                             "NI", "GUE/NGL"))
```

Wykres kołowy

Przygotowany przeze mnie wykres kołowy przedstawiający powyższe dane:

```
ggplot(dane, aes(x="", y = procent, fill = partia)) +  
  geom_bar(stat="identity", width = 1) +  
  ggtitle("Procentowy udział partii w Parlamencie Europejskim w 2019 roku") +  
  coord_polar("y", start=0) + theme_void() +  
  scale_fill_manual(values = c("#0000CC", "#00B300", "#FF0000", "#AC7339",  
                                "#FFFF00", "#88F7FC", "#800680", "#FF704D"))
```

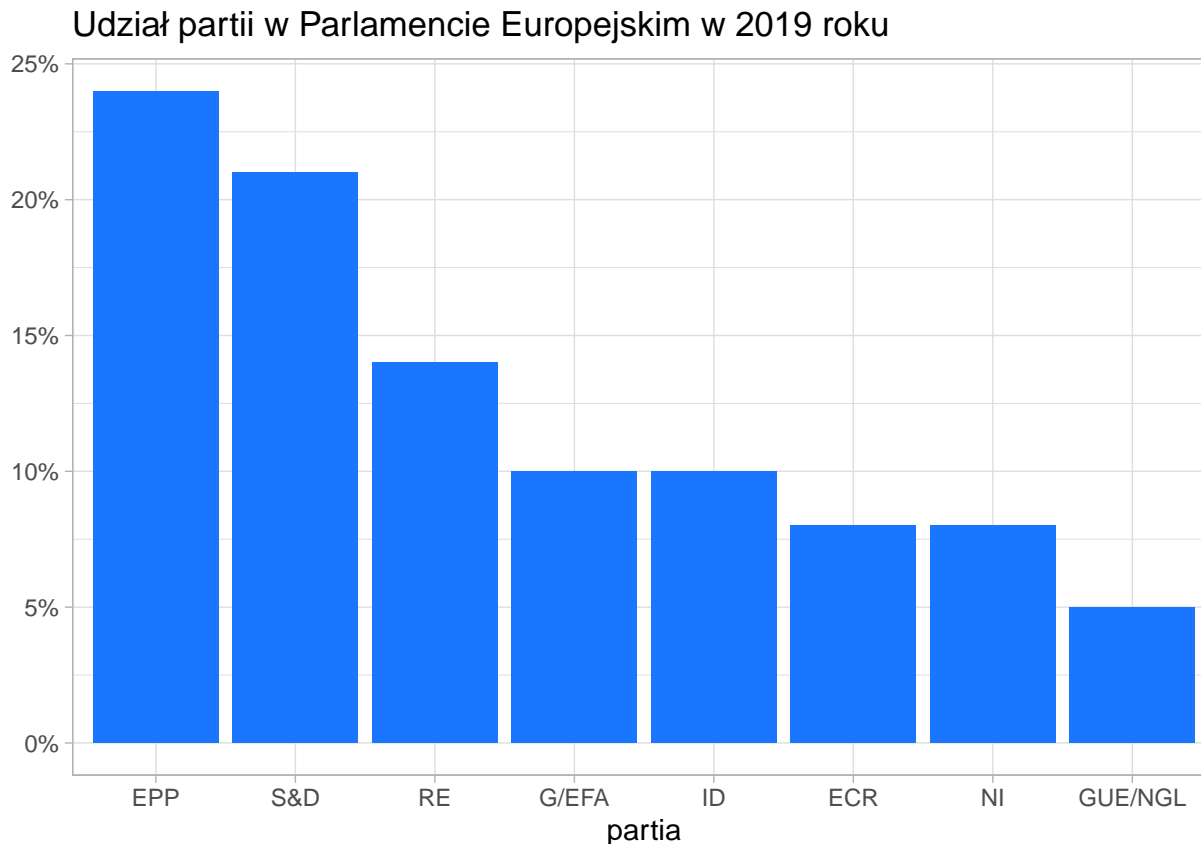
Procentowy udział partii w Parlamencie Europejskim w 2019 roku



Wykres słupkowy

Przygotowany przeze mnie wykres słupkowy przedstawiający powyższe dane:

```
ggplot(dane, aes(x =partia, y = procent, fill = partia)) +  
  geom_bar(stat="identity") + ylab(element_blank()) +  
  scale_y_continuous(labels = scales::percent_format(scale = 1, accuracy = 1)) +  
  ggtitle("Udział partii w Parlamencie Europejskim w 2019 roku") +  
  scale_fill_manual(values = rep("#1A75FF", length.out = 8)) + theme_light() +  
  guides(fill = FALSE)
```



Wnioski

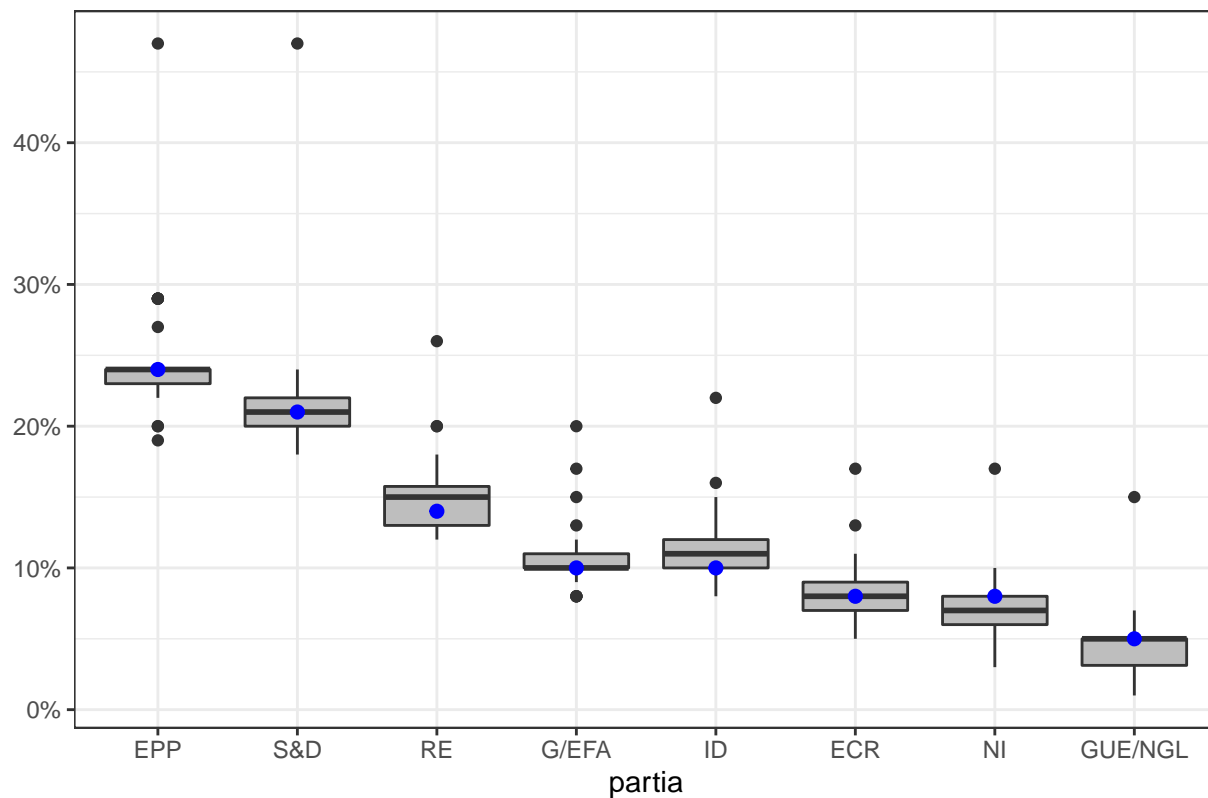
Przeprowadziłem ankietę z wykorzystaniem Google Forms na grupie studentów różnych kierunków. W ankiecie poprosiłem o podanie wartości odpowiadających procentowemu udziałowi danej partii najpierw na wykresie kołowym, a następnie na słupkowym. Na koniec zapytałem, który wykres był czytelniejszy. Do dnia 30 października 2020 roku otrzymałem 54 odpowiedzi i doszedłem do następujących wniosków:

- 98% respondentów uznało wykres słupkowy za czytelniejszy
- tylko jedna osoba potrafiła odczytać dokładne wartości z wykresu kołowego, a z wykresu słupkowego uczyniło to 37 osób, co odpowiada 69%.
- 17 osób (31%) podało inne wartości dla wycinków koła takiej samej wielkości, podczas gdy wysokości słupków nikt nie pomylił
- średni błąd odczytywania danych z wykresu kołowego wyniósł 6,5%, a z wykresu słupkowego 1%
- w przypadku wykresu kołowego największa wartość podana przez ankietowanych różniła się o 26 punktów procentowych od poprawnej wartości, natomiast w przypadku wykresu słupkowego wynosiła 4 punkty procentowe

Podsumowując, osoby odczytujące wartości z wykresów mają większe problemy z porównywaniem kątów niż wysokości słupków, dlatego też wykres słupkowy przedstawia dane w sposób zdecydowanie czytelniejszy niż wykres kołowy.

Na poniższych wykresach pudełkowych przedstawiłem w jaki sposób osoby ankietowane odczytywały wykres kołowy i słupkowy. Jako niebieskie kropki zazaczyłem wartości prawidłowe.

Odczytywanie warto ci z wykresu kołowego



Odczytywanie warto ci z wykresu słupkowego

