

# RaportPD2

Wojciech Kosiuk

4 11 2021

## Praca Domowa 2

Wojciech Kosiuk

### Wstęp

Niniejszy raport został stworzony na potrzeby pracy domowej 2 z przedmiotu Techniki Wizualizacji Danych. Celem pracy domowej było znalezienie błędnie sporządzonego wykresu i poprawienie go.

### Wykres niepoprawny

Wykres, który zdecydowałem się poprawiać pochodzi z rocznej analizy GUS o tytule “Rynek wewnętrzny w 2020 roku”, która ukazała się na stronie stat.gov.pl 2 listopada 2021. Publikacja obejmuje min. branżę sprzedażową, gastronomiczną oraz podstawowe wskaźniki ekonomiczne. Zdecydowałem się wybrać Wykres 16. z tejże publikacji przedstawiający strukturę placówek gastronomicznych.

**Wykres 16. Struktura placówek gastronomicznych**

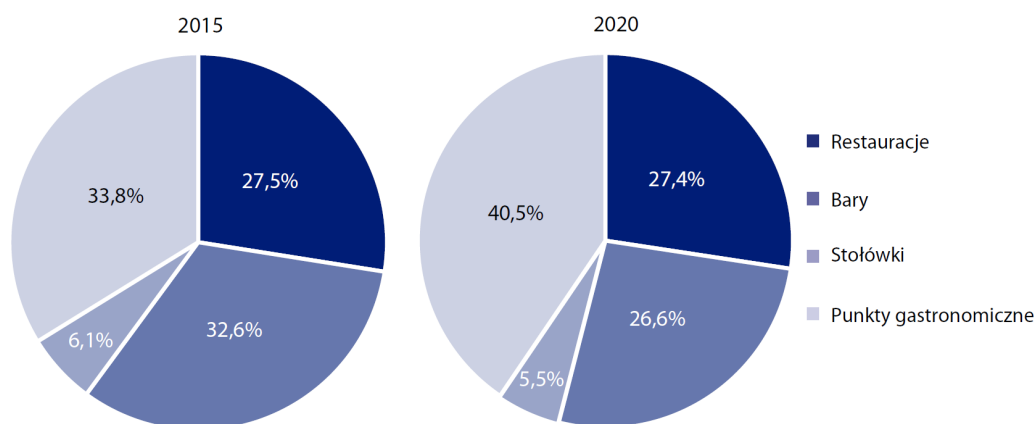


Figure 1: Struktura placówek gastronomicznych (GUS)

Źródło: <https://stat.gov.pl/obszary-tematyczne/ceny-handel/handel/rynek-wewnetrzny-w-2020-roku,7,27.html>

Strona 32, wykres 16

## Wykres poprawiony

### Opis

Do stworzenia wykresu użyłem bibliotek dostępnych w R:

- dplyr
- forcats
- ggplot2
- tidyr

oraz funkcji takich jak `fct_reorder()`, `pivot_longer()`, czy `geom_col()`.

### Kod do wykresu

```
library(dplyr)
library(forcats)
library(ggplot2)
library(tidyr)

if ( options()$stringsAsFactors )
  options(stringsAsFactors=FALSE)
my_data <- read.csv("C:/Users/wojtek/Desktop/pw/Techniki wizualizacji danych/Pd2/gusdane2.csv")

colnames(my_data) <- c("Struktura.placówek.gospodarczych","2015","2020")

my_data %>%
  mutate(Struktura.placówek.gospodarczych =
    forcats::fct_reorder(Struktura.placówek.gospodarczych,`2020`))%>%
  pivot_longer(!Struktura.placówek.gospodarczych,names_to = "Rok",values_to = "Procent")%>%
  ggplot(aes(x=Rok,y=Procent,fill=Rok)) +
  geom_col()+
  facet_grid(~Struktura.placówek.gospodarczych)+
  labs(title= "Struktura placówek gastronomicznych",
       x= "Porównanie w latach 2015 i 2020",
       y="Wartość procentowa")
```

## Wykres

Wykres słupkowy porównujący zmianę struktury placówek gastronomicznych pomiędzy latami 2015 i 2020 w zależności od rodzaju struktury.

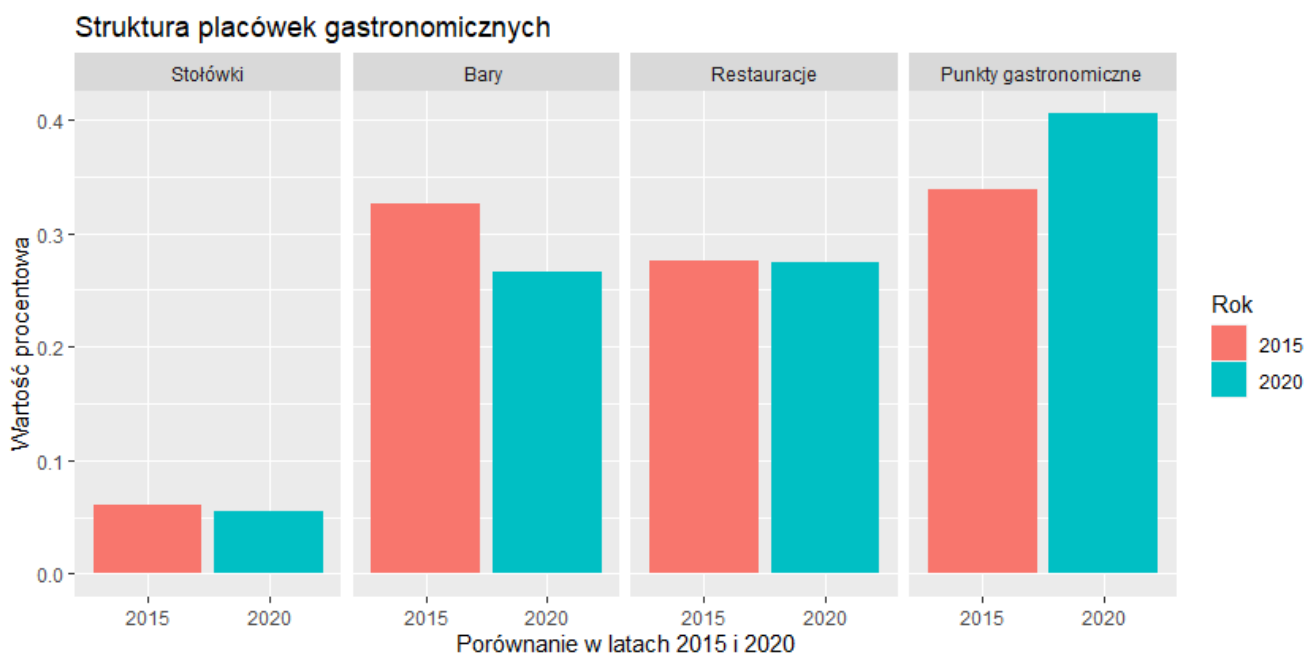


Figure 2: Struktura placówek gastronomicznych (Poprawiony wykres)

## Porównanie

Wykres sporządzony przez GUS jest wykresem kołowym, co sprawia, że dane przedstawione na nim są nieczytelne. Mój wykres jest wykresem słupkowym, w którym słupki są posortowane rosnąco względem udziału procentowego w 2020 roku. Dzięki temu dobrze widać relacje między różnymi strukturami placówek gastronomicznych.

Dodatkowo, wykres GUS przedstawia zmianę udziału procentowego na przestrzeni 5 lat (2015-20) jako dwa różne wykresy. Przez to ciężko jest odczytać zmianę wartości w czasie, gdyż trzeba spoglądać raz na jeden wykres, raz na drugi. W przypadku mojego wykresu, na jednym wykresie zostały przedstawione dwa rodzaje słupków. Jeden z danymi z 2015 roku, drugi z danymi z 2020 roku, gdzie dla każdej struktury dane z 2015 i 2020 znajdują się obok siebie.

Co więcej, oprócz odpowiednich podpisów na osi x, dane z różnych lat są u mnie zaznaczone różnymi kolorami słupków, co zwiększa czytelność wykresu.

Podsumowując, wykres kołowy nigdy nie jest dobrym pomysłem i w tym przypadku jest nie inaczej. Zastosowanie wykresu słupkowego, który na pozór wydaje się prostszy do zrobienia, pozwala znacznie zwiększyć przejrzystość i czytelność wykresu.