

# Analysing terrible graphs

Jan Taran

2024-11-19

Wykres został opublikowany 31.10.2024 przez GUS.

Wykres 3. Pracujący i wskaźnik zatrudnienia osób w wieku 15-89 lat i w wieku produkcyjnym

Kopiowanie danych z wykresu

Poprawianie wykresu

Dlaczego nowy wykres jest lepszy?

**Wykres został opublikowany 31.10.2024 przez GUS.** Strona z którego pochodzi pdf:

<https://stat.gov.pl/obszary-tematyczne/rynek-pracy/pracujacy-bezrobotni-bierni-zawodowo-wg-bael/aktywnosc-ekonomiczna-ludnosci-polski-2-kwartal-2024-roku,4,55.html>

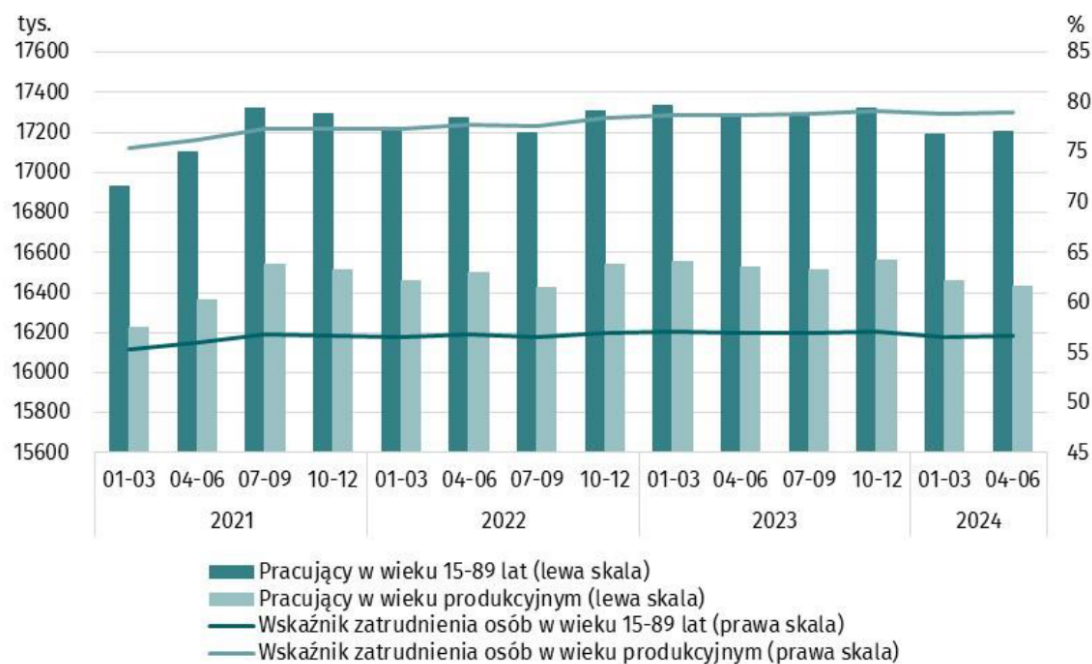
oraz sam plik pdf z którego pochodzi wykres:

[https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5475/4/55/1/aktywnosc\\_ekonomiczna\\_polski\\_za\\_2\\_kw\\_2024.pdf](https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5475/4/55/1/aktywnosc_ekonomiczna_polski_za_2_kw_2024.pdf)

## **Wykres 3. Pracujący i wskaźnik zatrudnienia osób w wieku 15-89 lat i w wieku produkcyjnym**

Na przedstawionym wykresie osie y nie zaczynają się od zera, co sprawia, że wydaje się, że są lekkie zmiany w wartościach na przestrzeni 3 lat, czy też lekki spadek na początku 2024 roku. Kolorystyka również jest fatalna, trzeba się chwilę przyjrzeć, aby móc cokolwiek odczytać z tego wykresu.

**Wykres 3. Pracujący i wskaźnik zatrudnienia osób w wieku 15-89 lat i w wieku produkcyjnym**



### Kopiowanie danych z wykresu

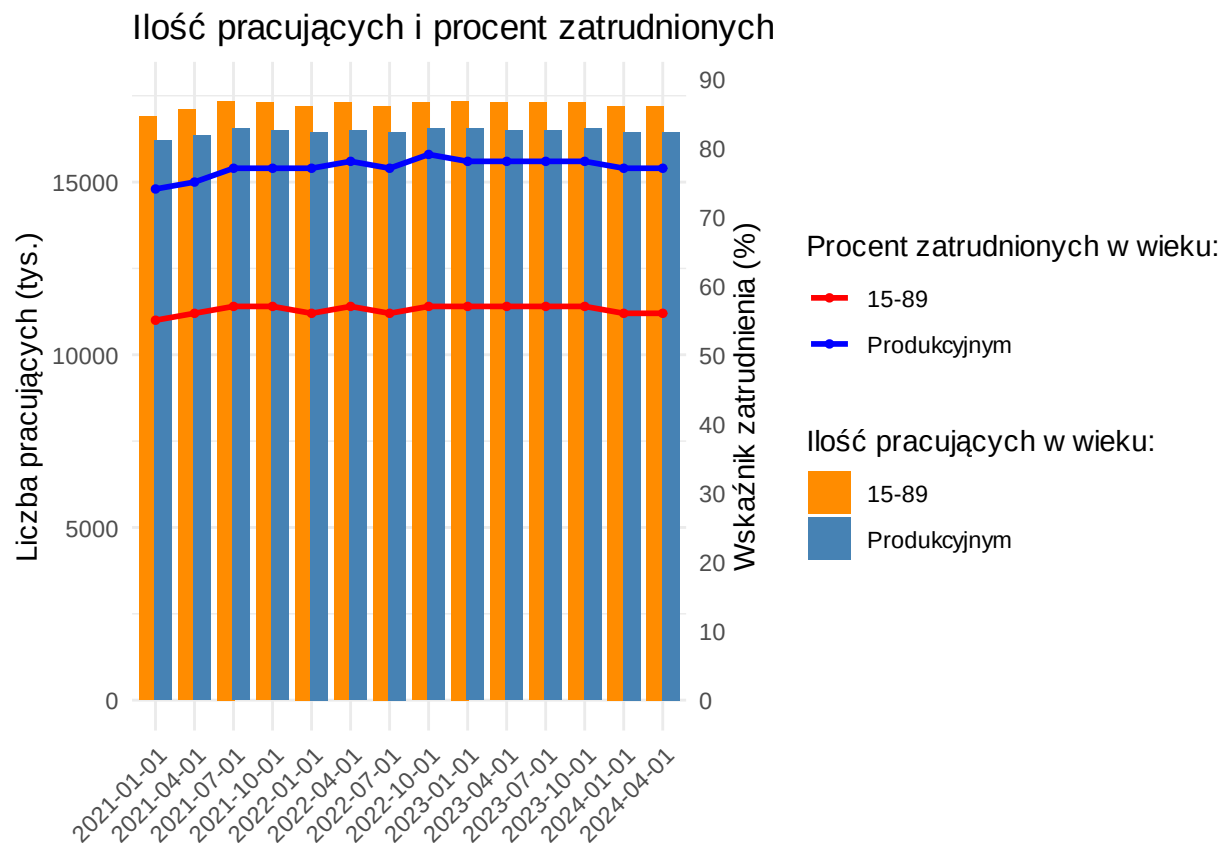
```
library(ggplot2)
library(tidyr)
data <- data.frame(
  Date = as.Date(c("2021-01-01", "2021-04-01", "2021-07-01", "2021-10-01",
                  "2022-01-01", "2022-04-01", "2022-07-01", "2022-10-01",
                  "2023-01-01", "2023-04-01", "2023-07-01", "2023-10-01",
                  "2024-01-01", "2024-04-01")),
  PracujacyWiek15_89 = c(16900, 17100, 17350, 17300,
                        17200, 17300, 17200, 17300,
                        17350, 17300, 17300, 17300,
                        17200, 17200),
  PracujacyWiekProd = c(16200, 16350, 16550, 16500,
                       16450, 16500, 16450, 16550,
                       16550, 16500, 16500, 16550,
                       16450, 16450),
  WskZatrudnieniaWiek15_89 = c(55, 56, 57, 57,
                               56, 57, 56, 57,
                               57, 57, 57, 57,
                               56, 56),
  WskZatrudnieniaWiekProd = c(74, 75, 77, 77,
                              77, 78, 77, 79,
                              78, 78, 78, 78,
                              77, 77)
)
```

## Poprawianie wykresu

```
data$Date <- factor(data$Date, levels = data$Date)

bar_data <-
  pivot_longer(data, cols = c(PracujacyWiek15_89, PracujacyWiekProd),
    names_to = "Category",
    values_to = "Value")

ggplot() +
  geom_bar(data = bar_data, aes(x = Date, y = Value, fill = Category),
    stat = "identity", position = position_dodge(width = 0.8)) +
  geom_line(data = data, aes(x = Date, y = WskZatrudnieniaWiek15_89 * 200,
    group = 1, color = "15-89"), linewidth = 1) +
  geom_line(data = data, aes(x = Date, y = WskZatrudnieniaWiekProd * 200,
    group = 1, color = "Produkcyjnym"), linewidth = 1) +
  geom_point(data = data, aes(x = Date, y = WskZatrudnieniaWiek15_89 * 200,
    color = "15-89"), size = 1) +
  geom_point(data = data, aes(x = Date, y = WskZatrudnieniaWiekProd * 200,
    color = "Produkcyjnym"), size = 1) +
  scale_y_continuous(
    name = "Liczba pracujących (tys.)",
    limits = c(0, 17600),
    sec.axis = sec_axis(~./200, name = "Wskaźnik zatrudnienia (%)",
      breaks = seq(0, 100, 10))
  ) +
  scale_fill_manual(
    labels = c("PracujacyWiek15_89" = "15-89",
      "PracujacyWiekProd" = "Produkcyjnym"),
    values = c("PracujacyWiek15_89" = "darkorange",
      "PracujacyWiekProd" = "steelblue")
  ) +
  scale_color_manual(
    values = c("15-89" = "red", "Produkcyjnym" = "blue")
  ) +
  labs(title = "Ilość pracujących i procent zatrudnionych",
    fill = "Ilość pracujących w wieku:",
    color = "Procent zatrudnionych w wieku:") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
    axis.title.y.right = element_text(angle = 90),
    axis.title.x = element_blank())
```



#### Dlaczego nowy wykres jest lepszy?

Na poprzednim wykresie odbiorca mógł odnieść wrażenie, że przedstawione dane jakoś się zmieniają na przestrzeni lat, gdzie w rzeczywistości te zmiany są znikome, co widać na moim nowym wykresie, gdzie oś Y zaczyna się od zera. Dodatkowo nowy dobór kolorów ułatwia rozpoznanie co oznaczają poszczególne słupki oraz linie. Według mnie legenda z boku wykresu również wpływa pozytywnie na jego czytelność.