

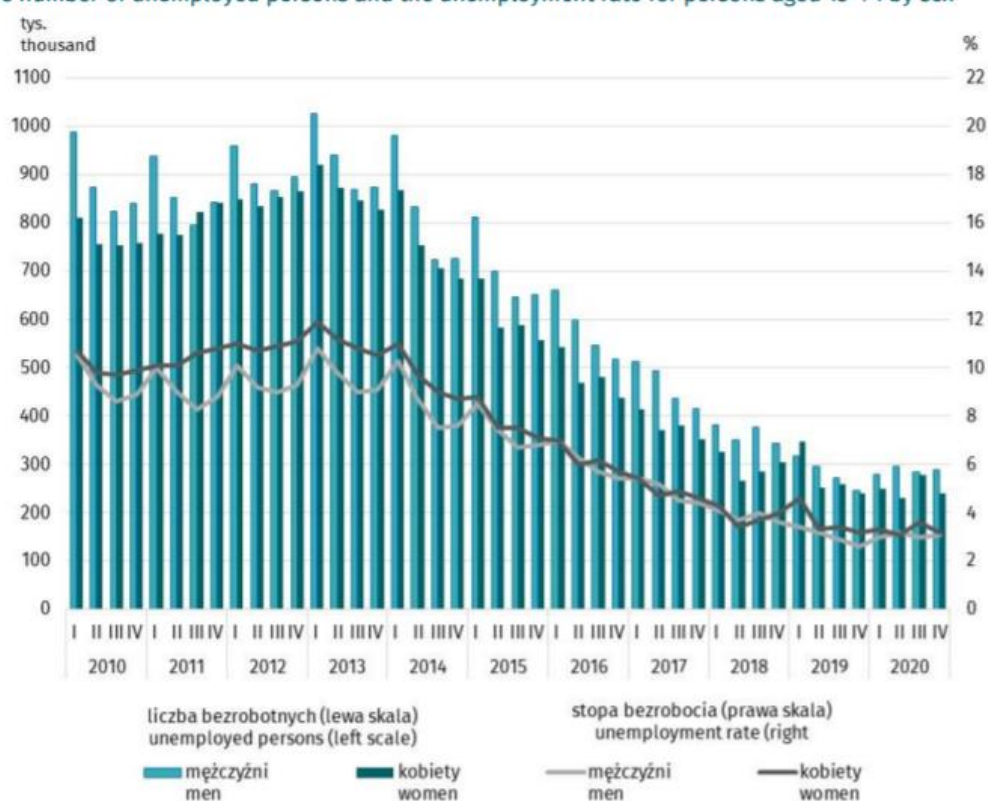
TWD – Praca domowa nr. 2

Mateusz Borowski

1. Wstęp

Poprawiana wizualizacja, to wykres znajdujący się na dole 42 strony raportu zatytułowanego „Wybrane aspekty rynku pracy w Polsce. Aktywność ekonomiczna ludności przed i w czasie pandemii COVID-19”.

Wykres 1.5. Liczba bezrobotnych i stopa bezrobocia ludności w wieku 15–74 lata według płci
Chart 1.5. The number of unemployed persons and the unemployment rate for persons aged 15–74 by sex

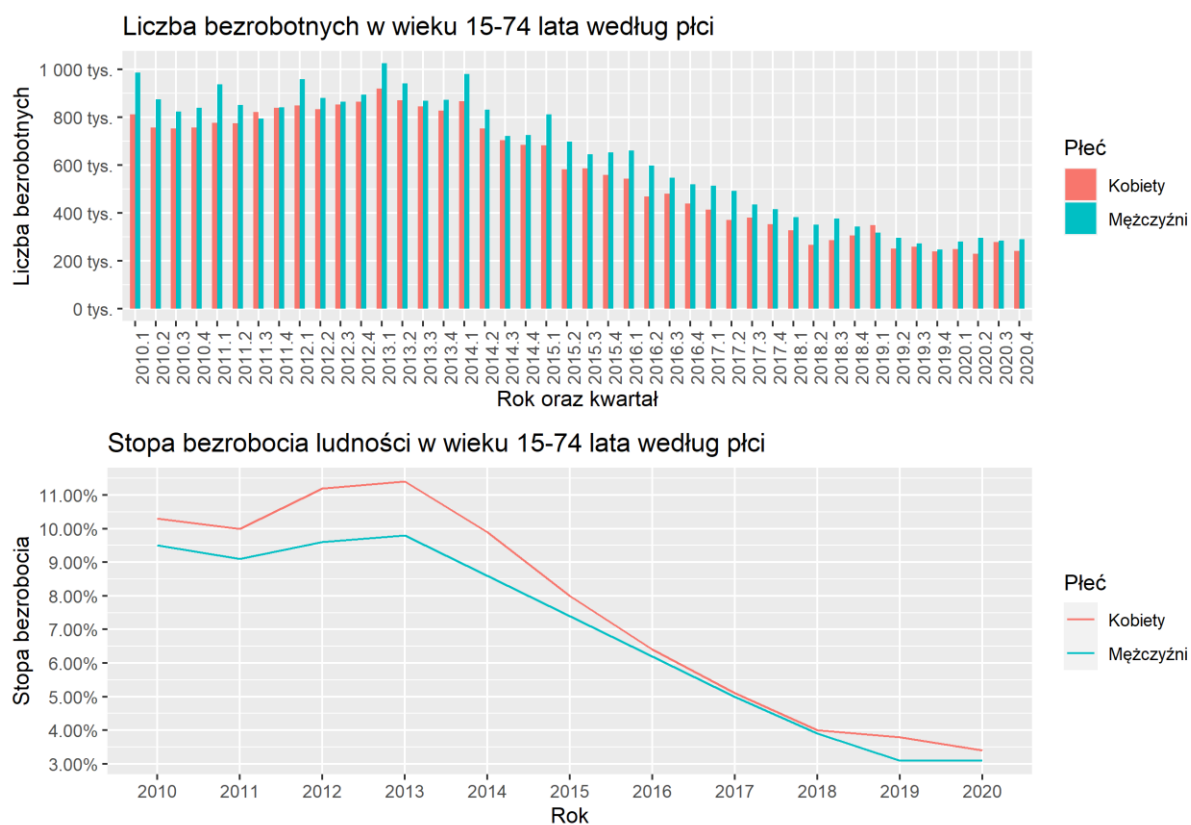


Rysunek 1 Wykres pochodzący z raportu „Wybrane aspekty rynku pracy w Polsce. Aktywność ekonomiczna ludności przed i w czasie pandemii COVID-19”

2. Dane

Raport nie zawiera danych za pomocą których wykres został wygenerowany. Dane, których użyłem do stworzenia mojej wizualizacji pochodzą ze strony bdl.stat.gov.pl. Niestety nie udało mi się dotrzeć do danych kwartalnych dotyczących stopy bezrobocia, dlatego w mojej wizualizacji używam danych rocznych. Przygotowane dane załączam w plikach „stopa_bezrobocia.csv” i „liczba_bezrobotnych.csv”.

3. Poprawiona wizualizacja



Rysunek 2 Poprawiony wykres

4. Kod

```
1 library(tidyverse)
2 library(dplyr)
3 library(ggplot2)
4 library(stringr)
5 library(cowplot)
6 library(scales)
7
8 liczba_bezrobotnych <- read.csv("liczba_bezrobotnych.csv", sep=";", encoding = "UTF-8")
9 stopa_bezrobocia <- read.csv("stopa_bezrobocia.csv", sep=";", encoding = "UTF-8")
10
11 plot_liczba_bezrobotnych <- liczba_bezrobotnych %>%
12   pivot_longer(-płeć, names_to = "rok", values_to = "liczba_bezrobotnych.w.tys") %>%
13   mutate(rok = str_sub(rok, 2, -1)) %>%
14   ggplot(aes(x = rok, y = liczba_bezrobotnych.w.tys, fill = płeć)) +
15   geom_bar(stat = "identity", position = "dodge", width = 0.5) +
16   scale_y_continuous(n.breaks = 8, labels = unit_format(unit = "tys.")) +
17   theme(axis.text.x = element_text(angle = 90, vjust = 1, hjust=1)) +
18   scale_fill_discrete(name = "Płeć",
19     labels = c("Kobiety", "Mężczyźni")) +
20   labs(title = "Liczba bezrobotnych w wieku 15-74 lata według płci",
21     x = "Rok oraz kwartał",
22     y = "Liczba bezrobotnych")
23
24 plot_stopa_bezrobocia <- stopa_bezrobocia %>%
25   pivot_longer(-płeć, names_to = "rok", values_to = "stopa_bezrobocia") %>%
26   mutate(rok = str_sub(rok, 2, -1),
27     stopa_bezrobocia = stopa_bezrobocia / 100) %>%
28   group_by(płeć) %>%
29   ggplot(aes(x = rok, y = stopa_bezrobocia, group = płeć, color = płeć)) +
30   geom_line() +
31   scale_y_continuous(n.breaks = 10,
32     labels = percent) +
33   scale_color_discrete(name = "Płeć",
34     labels = c("Kobiety", "Mężczyźni")) +
35   labs(title = "Stopa bezrobocia ludności w wieku 15-74 lata według płci",
36     x = "Rok",
37     y = "Stopa bezrobocia")
38
39 plot_final <- plot_grid(plot_liczba_bezrobotnych, plot_stopa_bezrobocia, nrow = 2)
40
41 ggsave(file = "stopa_i_liczba_bezrobotnych.png")
42
```

Rysunek 3 Kod źródłowy

5. Uzasadnienie

Rozbiecie wykresu na dwa osobne zwiększa jego czytelność, ponieważ znika problem z odczytaniem danych spowodowany nachodzącymi na siebie wykresami. Odczytanie wartości również jest teraz łatwiejsze, gdyż każdy wykres ma tylko jedną oś Y. Kolory na nowym wykresie są lepiej dobrane, dzięki czemu nie zlewają się tak bardzo jak na oryginalnej wizualizacji.