

Michał Syrkiewicz HW3

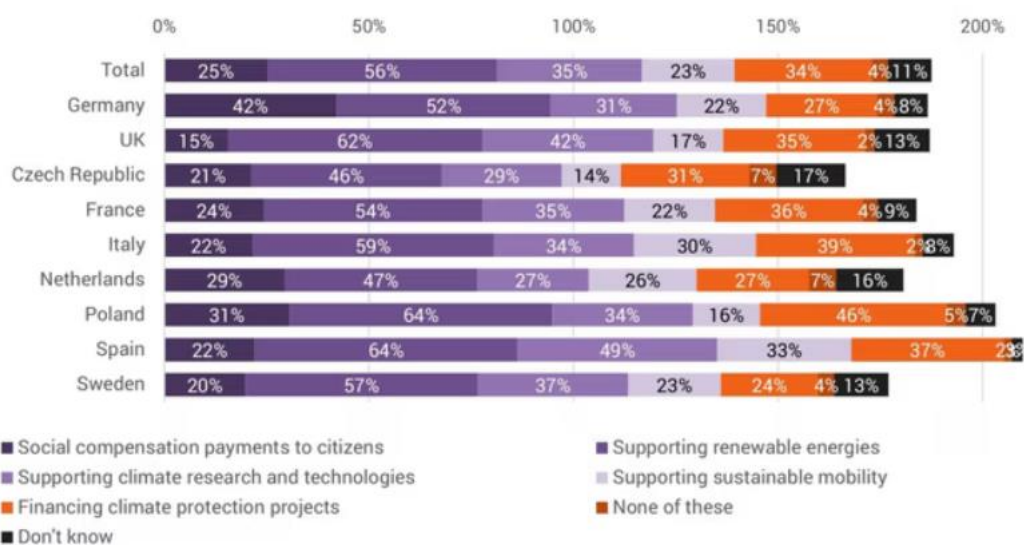
TWD 2024/2025

Błędna wizualizacja danych

Źródło: <https://wiadomosci.onet.pl/jak-europejczycy-walczą-o-zieloną-przyszłość-wyniki-nowego-badania/lxzdegi>

Data publikacji: 06.11.2024

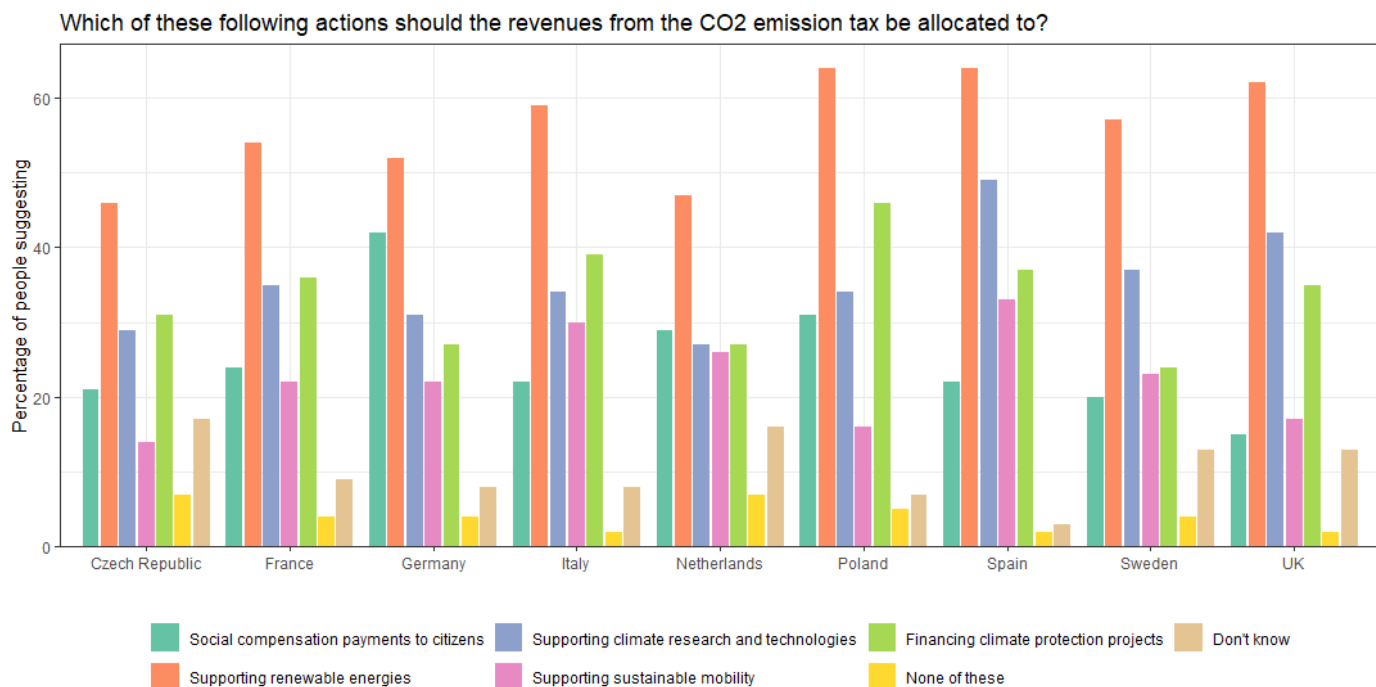
Na które z poniższych działań powinny być przeznaczone dochody z podatku od emisji CO₂?



Elementy wymagające poprawy w powyższej wizualizacji:

1. Wykres jest w 2 językach. Tytuł jest po polsku, reszta elementów w języku angielskim
2. Użycie w ten sposób procentów jest mylące i niepotrzebne. Sam musiałem się zastanawiać skąd się bierze fakt, że sumują się do tylu różnych wartości. Chodzi o to, że odpowiadający mógł wybrać kilka odpowiedzi.
3. Kategoria Total nie dość, że zbędna to nie wiadomo co oznacza. Domyślam się, że jest to średnia z wszystkich krajów, nie jestem jednak pewny.
4. Napisy na słupkach na siebie nachodzą i tworzą nieczytelne, zbite grupy np. końcowe wartości w Hiszpanii.

Poprawiona wizualizacja



Dlaczego jest lepsza od poprzedniej?

1. Wykres w jednym języku. Ja użyłem angielskiego.
2. Ułożenie słupków oraz nazwy skali sugerują, że każda odpowiadająca osoba mogła wybrać kilka odpowiedzi. Rozwiązuje to problem dziwnego sumowania się procentów na wykresach.
3. Łatwiej porównywać odpowiedzi pomiędzy poszczególnymi kategoriami w państwie oraz pomiędzy jedną kategorią w różnych państwach.
4. Usunięcie mało zrozumianej kategorii Total, zwiększa czytelność wykresu.
5. Mała stylistyczna zmiana – kraje ułożone alfabetycznie

Kod:

```
1 library(ggplot2)
2 library(dplyr)
3 library(tidyr)
4 countries <- c("Total","Germany","UK","Czech Republic","France","Italy","Netherlands","Poland","Spain","Sweden")
5 social_compasation <- c(25,42,15,21,24,22,29,31,22,20)
6 support_renew <-c(56,52,62,46,54,59,47,64,64,57)
7 support_research <- c(35,31,42,29,35,34,27,34,49,37)
8 supper_sus <- c(23,22,17,14,22,30,26,16,33,23)
9 finance<- c(34,27,35,31,36,39,27,46,37,24)
10 none <- c(4,4,2,7,4,2,7,5,2,4)
11 dont_know<- c(11,8,13,17,9,8,16,7,3,13)
12 df <- data.frame(
13   Country = countries,
14   `Social compensation payments to citizens` = social_compasation,
15   `Supporting renewable energies` = support_renew,
16   `Supporting climate research and technologies` = support_research,
17   `Supporting sustainable mobility` = supper_sus,
18   `Financing climate protection projects` = finance,
19   `None of these` = none,
20   `Don't know` = dont_know,
21   check.names=FALSE
22 )
23 df_long <- df %>%
24   pivot_longer(
25     cols = -Country,
26     names_to = "Category",
27     values_to = "Value")
28 category_order <- c( "Social compensation payments to citizens", "Supporting renewable energies",
29   "Supporting climate research and technologies","Supporting sustainable mobility",
30   "Financing climate protection projects","None of these", "Don't know")
31 max_v = max(df_long$Value)
32 df_long$Category <- factor(df_long$Category, levels = category_order)
33 df_long %>% filter(Country != "Total") %>%
34   ggplot(aes(x = Country, y = Value, fill = Category)) +
35   geom_col(position = position_dodge2(preserve = "single")) +
36   scale_fill_brewer(palette = "Set2") +
37   labs(
38     y = "Percentage of people suggesting",
39     x="",
40     title = "Which of these following actions should the revenues from the CO2 emission tax be allocated to?"
41   ) +
42   theme_bw() +
43   theme(
44     legend.position = "bottom",
45     legend.title = element_blank() ) + scale_y_continuous(expand = c(0,0),limits = c(0, max_v * 1.05))
```