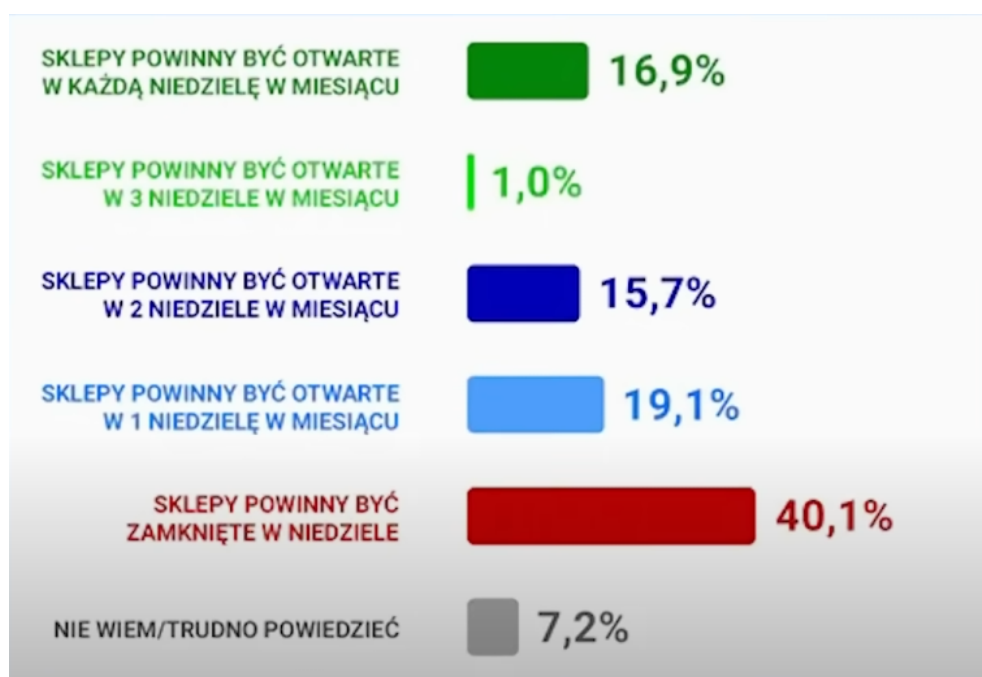


Praca domowa 3 z eksploracji danych

Maciej Turczyński

1 Wizualizacja źródłowa



Rysunek 1: Wizualizacja źródłowa

Źródło: <https://youtu.be/VI5kkAJmlqM?si=RnVukqoINGg9pFgK&t=939>

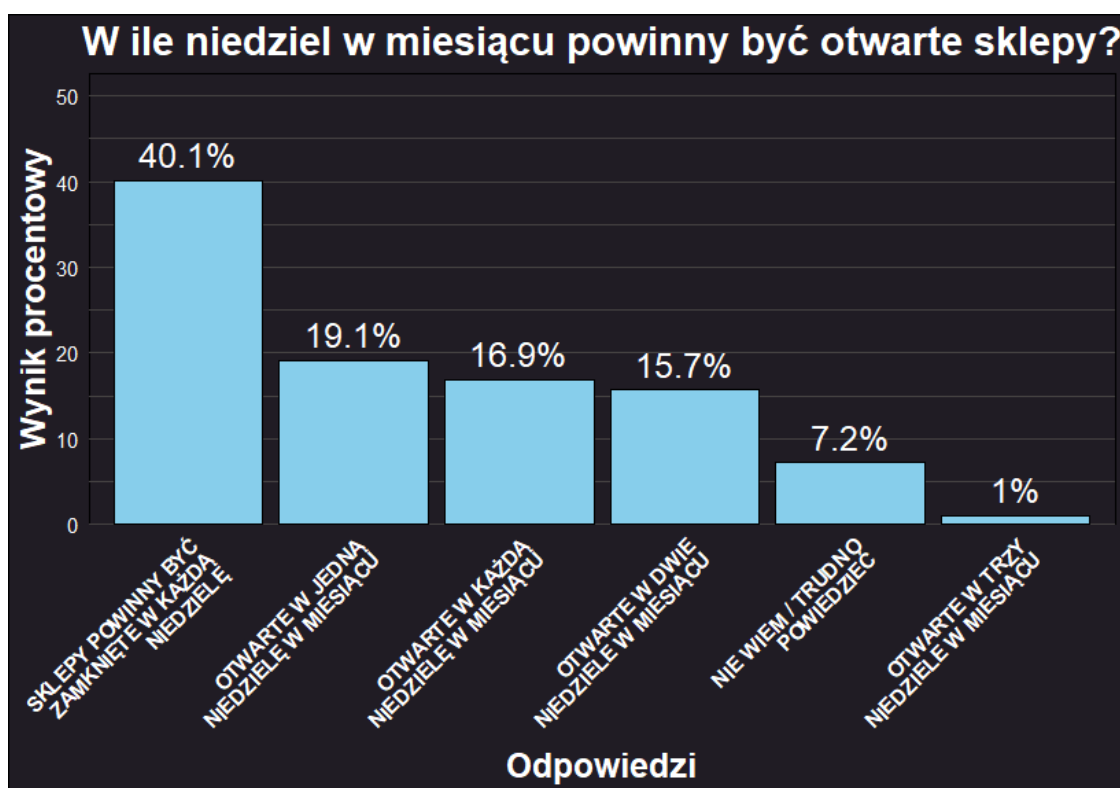
Data publikacji: 22 marca 2024

2 Elementy wymagające poprawy

Wykres jest mało czytelny oraz nie wygląda estetycznie. Słupki przedstawione bez siatki są trudne do porównywania, mimo zamieszczonych obok wartości liczbowych w procentach. Oprócz tego, każdy słupek jest w innym kolorze mimo tego, że wszystkie kolumny są podpisane.

3 Utworzona wizualizacja z kodem do jej wygenerowania

3.1 Wizualizacja



Rysunek 2: Utworzona wizualizacja

3.2 Kod do wygenerowania wykresu

```
1 library(ggplot2)
2 library(dplyr)
3 library(stringr)
4
5 data <-
6   data.frame(
7     c(
8       'OTWARTE W KAZDA NIEDZIELE W MIESIACU',
9       'OTWARTE W TRZY NIEDZIELE W MIESIACU',
10      'OTWARTE W DWIE NIEDZIELE W MIESIACU',
11      'OTWARTE W JEDNA NIEDZIELE W MIESIACU',
12      'SKLEPY POWINNY BYC ZAMKNIETE W KAZDA NIEDZIELE',
13      'NIE WIEM / TRUDNO POWIEDZIEC'
14    ),
15    c(0.169, 0.01, 0.157, 0.191, 0.401, 0.072)
16  )
17
18 colnames(data) <- c('Pytanie', 'Wynik_procentowy')
19 data$Pytanie <- str_wrap(data$Pytanie, width = 20)
20
21
22 data <- data %>%
23   arrange(-Wynik_procentowy) %>%
24   mutate(Pytanie = factor(Pytanie, levels = Pytanie))
25 ggplot(data, aes(x = Pytanie, y = Wynik_procentowy * 100)) +
26   geom_col(fill = "skyblue", color = "black") +
27   theme_minimal() +
28   theme(
29     axis.text.x = element_text(
30       angle = 45,
31       hjust = 1,
32       color = 'white',
33       size = 10,
34       face = 'bold'
35     ),
36     axis.text.y = element_text(color = 'white', size = 10),
37     axis.title.x = element_text(color = 'white', size = 17, face = '
38       bold'),
39     axis.title.y = element_text(color = 'white', size = 17, face = '
40       bold'),
41     plot.title = element_text(
42       color = 'white',
43       hjust = 0.5,
44       size = 20,
```

```

43     face = 'bold'
44   ),
45   panel.background = element_rect(fill = '#201c24'),
46   plot.background = element_rect(fill = '#201c24'),
47   panel.grid.major.x = element_blank(),
48   panel.grid.major.y = element_line(color = '#454242'),
49   panel.grid.minor.y = element_line(color = '#454242')
50 ) +
51 labs(x = "Odpowiedzi", y = "Wynik procentowy", title = "W ile
      niedziel w miesiacu powinny byc otwarte sklepy?") +
52 geom_text(
53   aes(label = paste0(round(
54     Wynik_procentowy * 100, 1
55   ), '%')),
56   vjust = -0.5,
57   size = 6,
58   color = 'white'
59 ) +
60 scale_y_continuous(limits = c(0, 50), expand = expansion(mult = c
      (0, 0.05)))

```

4 Uzasadnienie

Utworzony przeze mnie wykres jest lepszy, ponieważ:

- nie występuje podwójne oznaczenie słupków - wystarcza, że są one podpisane;
- brak kolorów słupków sprawia, że nie jest on sugestywny, co do tego czy dana odpowiedź powinna być odbierana w sposób pozytywny czy nie;
- na wykresie jest umieszczona siatka, co znacząco ułatwia odczytywanie wykresu;
- słupki umieszczone są malejąco, co poprawia czytelność wizualizacji;
- jest obiektywnie ładniejszy graficznie, domownicy pytani o to, który wykres jest lepszy (bez dodawania kontekstu, że jest to moja praca domowa) jednomyślnie stwierdzili, że ten stworzony przeze mnie.