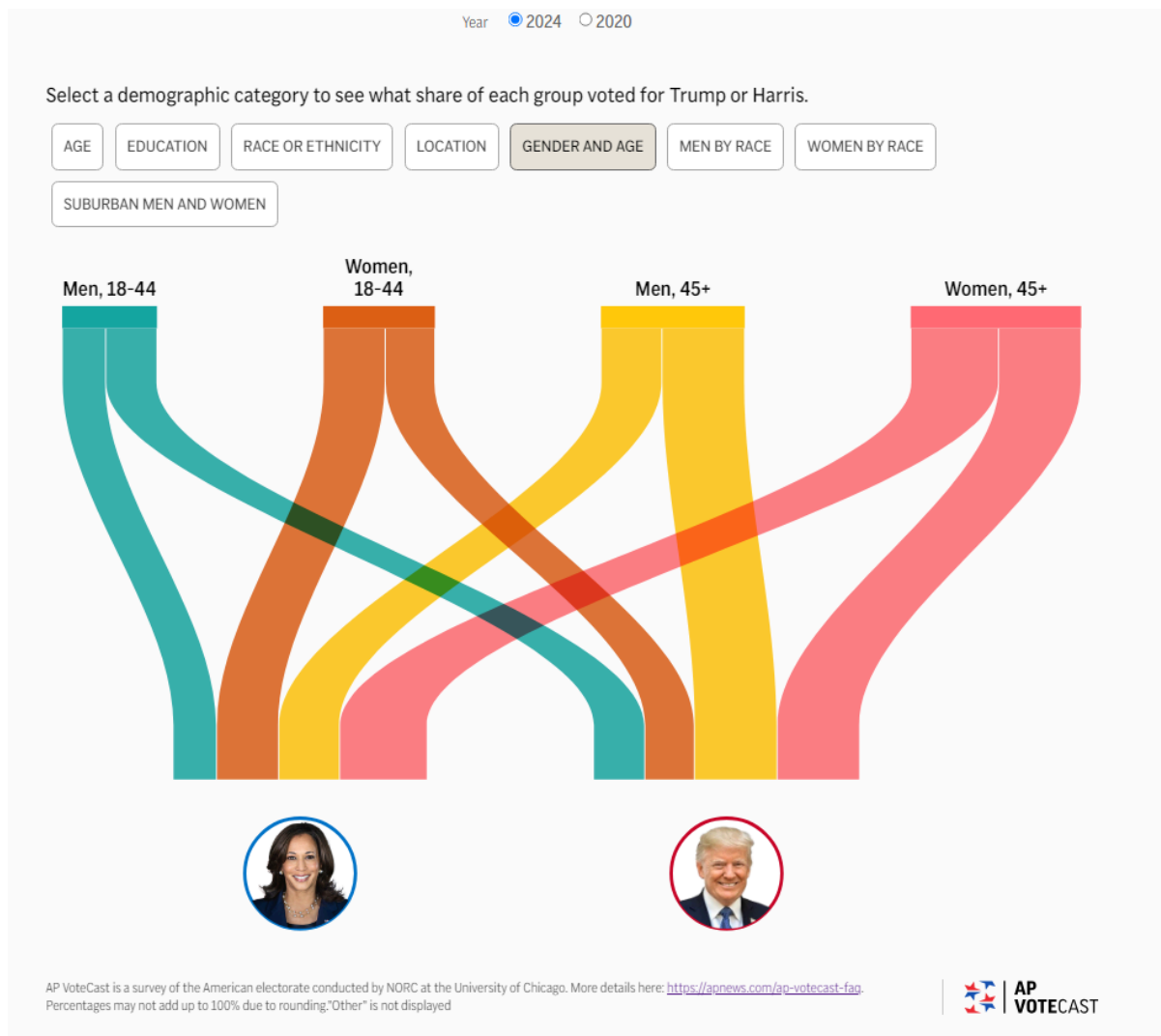


# Praca domowa 3

Zuzanna Zalewska

Wizualizacja, która zostanie poprawiona:



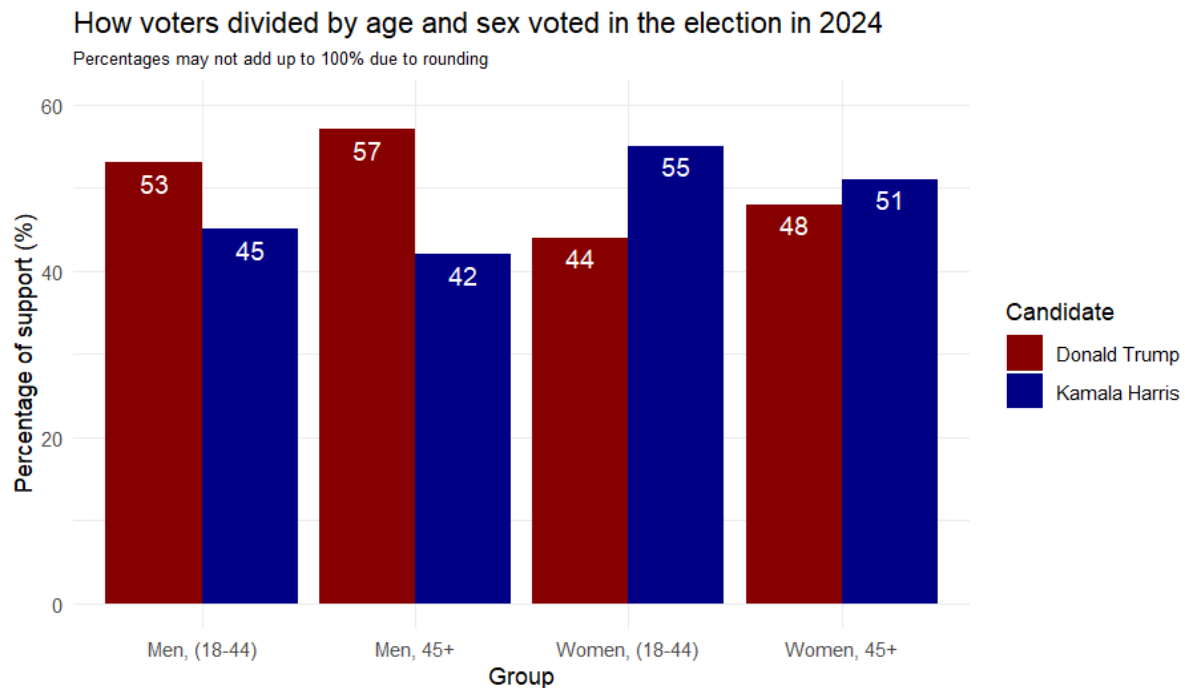
Źródło: <https://apnews.com/projects/election-results-2024/votecast/>

Data publikacji: 06.11.2024.

Elementy wizualizacji wymagające poprawy:

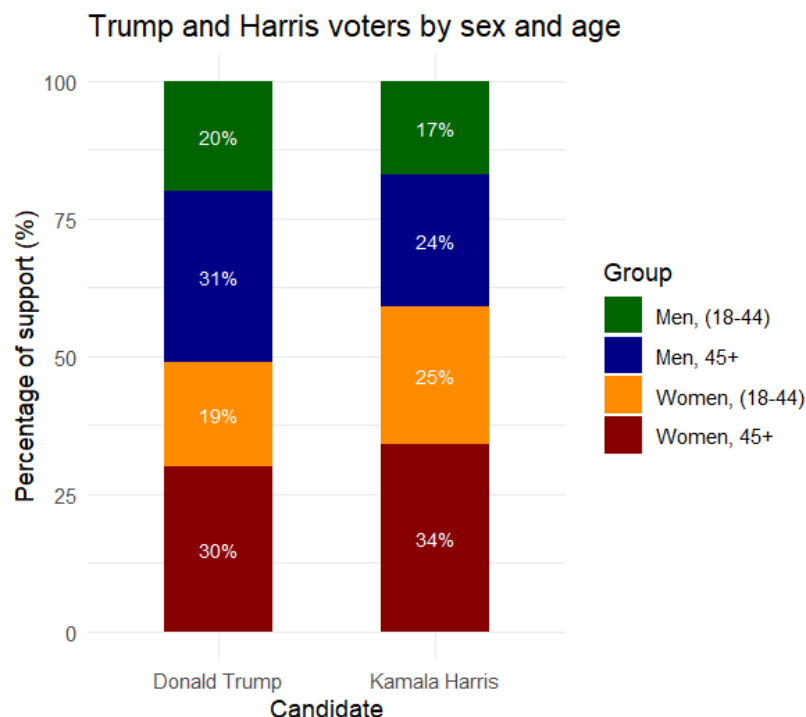
- brak tytułu oraz podpisanych osi na wykresie,
- dla każdej z grup wyborców ciężko jest określić, jaki procent danej grupy ludzi głosuje na danego kandydata,
- w sposób nieczytelny został przedstawiony podział na grupy ludzi popierające danego kandydata .

Autor wizualizacji próbując przedstawić dwie różne informacje na jednym wykresie sprawia, że staje się on mało informatywny. Lepszym rozwiązaniem będzie stworzenie dwóch oddzielnych wykresów.



Kod generujący wykres:

```
#dane pochodzące ze źródła
df<-data_frame(Group=c("Men, (18-44)","Women, (18-44)","Men, 45+","Women, 45+"),
               Kamala_Harris=c(45,55,42,51),
               Donald_Trump=c(53,44,57,48))
df<-pivot_longer(df,cols=c("Kamala_Harris","Donald_Trump"),names_to="Candidate",values_to="perc")
ggplot(df,aes(x=Group,y=perc,fill=Candidate))+
  geom_bar(stat="identity",position = "dodge")+
  theme_minimal()+
  labs(title="How voters divided by age and sex voted in the election in 2024",
       subtitle = "Percentages may not add up to 100% due to rounding",
       x="Group",
       y="Percentage of support (%)")+
  ylim(0,60)+
  scale_fill_manual(labels=c("Kamala_Harris"="Kamala Harris","Donald_Trump"="Donald Trump"),
                   values=c("Kamala_Harris" = "darkblue", "Donald_Trump" = "darkred"))+
  geom_text(aes(label=perc),color="white",position = position_dodge(width = 0.9),vjust = 1.5,size=4)+
  theme(plot.subtitle = element_text(size = 8))
```



Kod generujący wykres:

```
#dane oszacowane na podstawie pierwotnego wykresu
df2<-data_frame(Group=c("Men, (18-44)", "Women, (18-44)", "Men, 45+", "Women, 45+"),
                  Kamala_Harris=c(17,25,24,34),
                  Donald_Trump=c(20,19,31,30))
df2<-pivot_longer(df2,cols=c("Kamala_Harris", "Donald_Trump"),names_to="Candidate",values_to="perc")
ggplot(df2,aes(x=Candidate,y=perc,fill=Group))+
  geom_bar(stat="identity",width = 0.5)+
  labs(title="Trump and Harris voters by sex and age",
        y="Percentage of support (%)")+
  theme_minimal()+
  geom_text(aes(label =paste0(perc, "%")),position = position_stack(vjust = 0.5), size = 3,color="white") +
  scale_fill_manual(values = c("darkgreen", "darkblue", "darkorange", "darkred"))+
  scale_x_discrete(labels=c("Kamala_Harris"="Kamala Harris", "Donald_Trump"="Donald Trump"))
```

Pierwszy z wykresów w sposób dokładny dostarcza informacji, jak podane grupy ludzi głosowały. Zastosowane zostały także odpowiednie kolory reprezentujące kandydatów. Wysokości słupków można porównywać nie tylko pomiędzy daną grupą, ale także pomiędzy różnymi grupami, co było utrudnione przy pierwotnym wykresie. Drugi wykres dostarcza nam w wygodny sposób możliwość porównywania danej wartości zarówno dla pojedynczego kandydata, jak i dla obydwu kandydatów.