Numerous variables relating to 2016 election in 2021 US GSS*

Possible factors effecting 2016 election

SangWoong Lee   Young Suk

19 March 2022

Abstract

First sentence (what was done). Second sentence(what was found). Third sentence (why this matters). Fourth sentence.

1 Introduction

Donald Trump won the 2016 United States presidential election. From what we remembered, it was quite surprising considering his image outside of U.S or even in U.S. Also, from what we remembered, before the election, Hillary Clinton was winning in most surveys as well. There is a famous quote that we learn from the past. Therefore, we thought that even though we are living in 2022, it would be meaningful to analyze the election so that we get to know what U.S citizens thoughts, or in broad sense thoughts trends in nowadays and what factors they or we considered most important to decide our representative.

This paper is going to explore mainly possible variables that potentially had impacts on deciding factors in the election as well as their analysis. However, making a decision that this variable affects U.S citizens to lean on Trump side or Clinton side is not easy because there is nothing in this world absolute. Thus, as we mentioned above, we are going to focus on general flows of results instead of defining variables as absolute factors of the demarcation. Following that, this paper has structures: Data, Results, Discussion, Appendix and supplementary survey.

We have obtained dataset from U.S GSS, and sorted variables that we consider as potential factors affecting the election. Throughout analysis of relation between the variables and respondents' votes on the election, we could observe that people who have conservative thoughts on equality and welfare preferred Trump, and people who do not preferred Clinton. This may sound obvious considering their parties, however in result session, reader can observe that what specific areas or fields they become more conservative and progressive. We believe that this situation is not only happening in U.S, but it also happens all over the world. This claim is going to be conducted with supplementary survey session.

This report was created using `R` (R Core Team 2020) and its libraries `tidyverse` (Wickham et al. 2019), `dplyr` (Wickham et al. 2021), were used to process and clean the data, and `ggplot` (Wickham 2016) `bookdown` (Xie 2022) and `knitr` (Xie 2021) was used to format this report.

2 Data

Original dataset is collected through US GSS, which is aimed for the goal that enabling the cleanest comparison to show changes in public opinion in the United States. Usually, it is conducted in person, however

---

*Code and data are available at: https://github.com/SANGWOONG-LEE/2016electionanalysis

because of pandemic, this time it was conducted via online. Among numerous interesting findings from the dataset, we choose to analyze the 2016 United States presidential election. This slightly deviates from the US GSS goal, but we thought that it can clearly represent current situations of U.S. However, choosing specific variables related to the election was not easy since deciding who to vote is followed by countless and complex factors. After looking over the variables, we chose age,race,gender,degree,born US, family income, national spending on education, race, arms, welfare, the poor, the other countries, social security and immigration as potential variables.

We mainly divided the variables into two parts: personal information (Age, race, gender, degree, born US and family income) and thoughts on national policies (national spending on education, race, arms, welfare, the poor, the other countries, social security and immigration). As I mentioned above, deciding who to vote is followed by very complex system of thoughts. Thus, as we cut down the variables, they do not guarantee 100% of the vote. Also, we thought that excluding non-response completely would be a bias for our understandings in data, so even though it caused difficulties in interpretation, we decided to include it except "pres16", which is our main variable whether respondents voted for Clinton and Trump. This is because non-response does not have meanings with respect to our interests. There exists similar survey similar to ours, CNN exit polls, however it is not easy to obtain the datasets they used. Also, US GSS is more realistic than exit polls since people would think more rational in US GSS than exit polls.

Now, let's get into the variables in details. One thing that we focused at first was whether respondents were equivalently recruited. To do this, we have to look at respondents' personal information. If you take a look from figure 1 to 4. We can see some interesting observations as followed:

1) number of survey participants of age in the range of 20~40 is unexpectedly low compared to other ages.

2) There are more white race than others. However, considering US race demographics, it is very reasonable. For more details on US race demographics, please visit the link: https://www.census.gov/quickfacts/fact/table/US/PST045221.

3) For gender, there is slightly more female than male, and we will see later whether this has impacts on vote or not.

From figure 5 to 14, they represent respondents' thoughts on national policies as well as relation between variables. Mostly they were collected from questions that whether you think that national spending on corresponding variables is about right, too much, or too little.For instance, let's take a look at figure 7. This graph represents respondents' thoughts on U.S national spending on improving black condition, and we sorted it with our interest, Clinton or Trump. For more example, let's take a look at figure 9.This graph represents relation between income and degree with respect to our interest. It generally shows that higher education you have more income you earn. We are going to talk about interesting results in more detailed later in result and discussion session. For more information regarding variables graph please check Appendix.
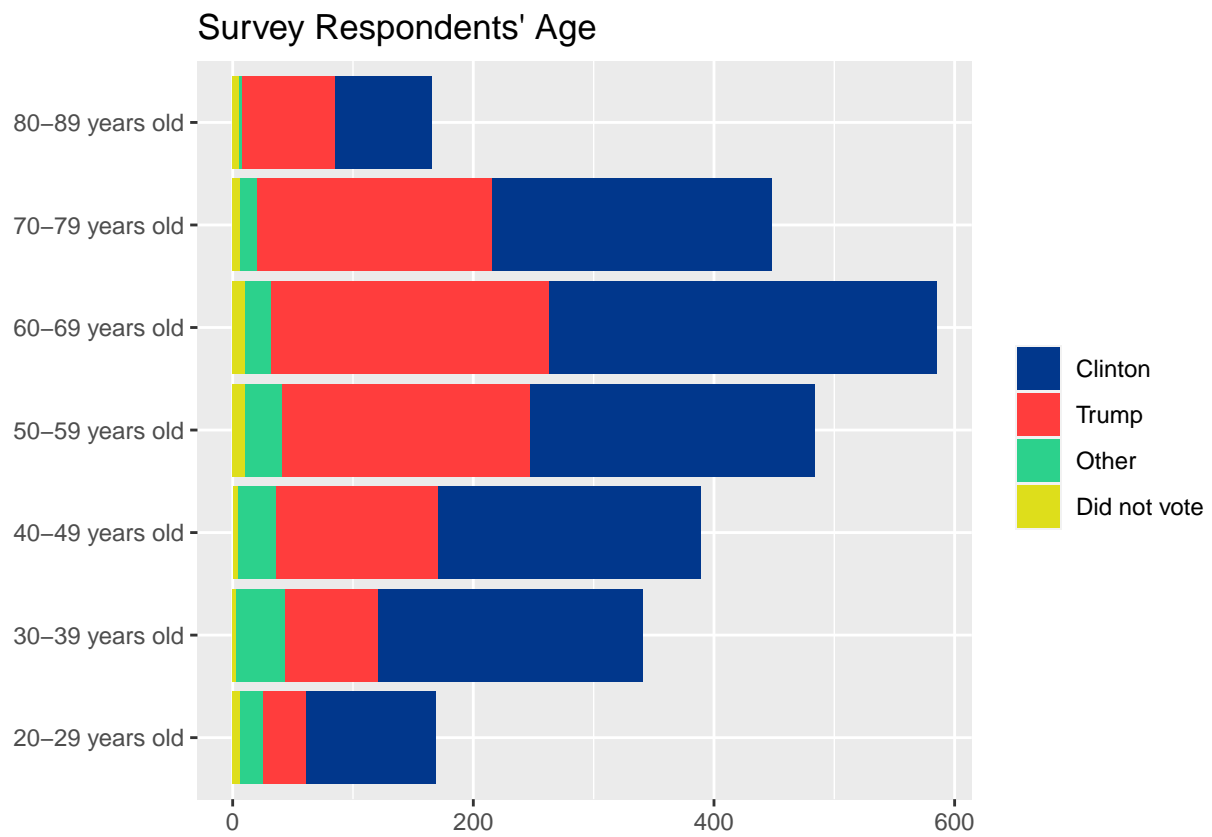
Figure 1: Distribution of survey respondents' age group.
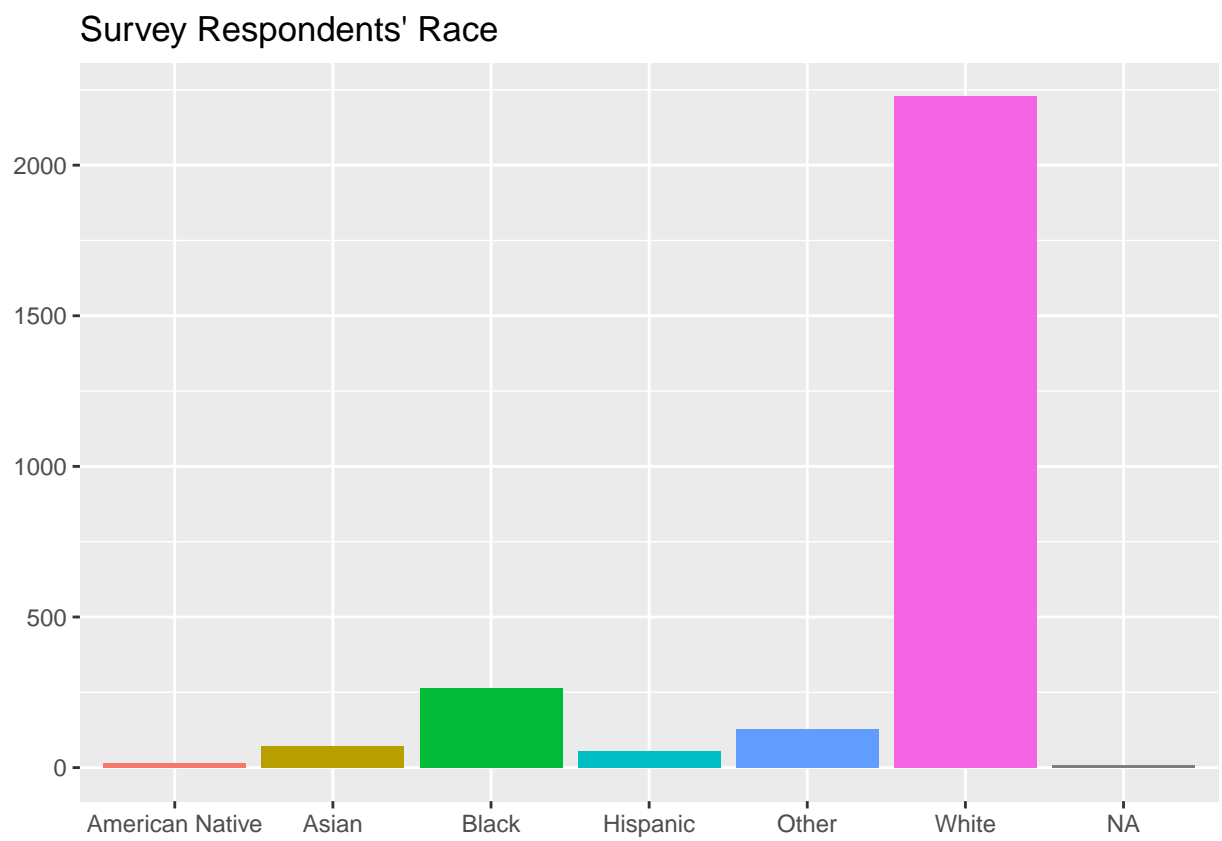
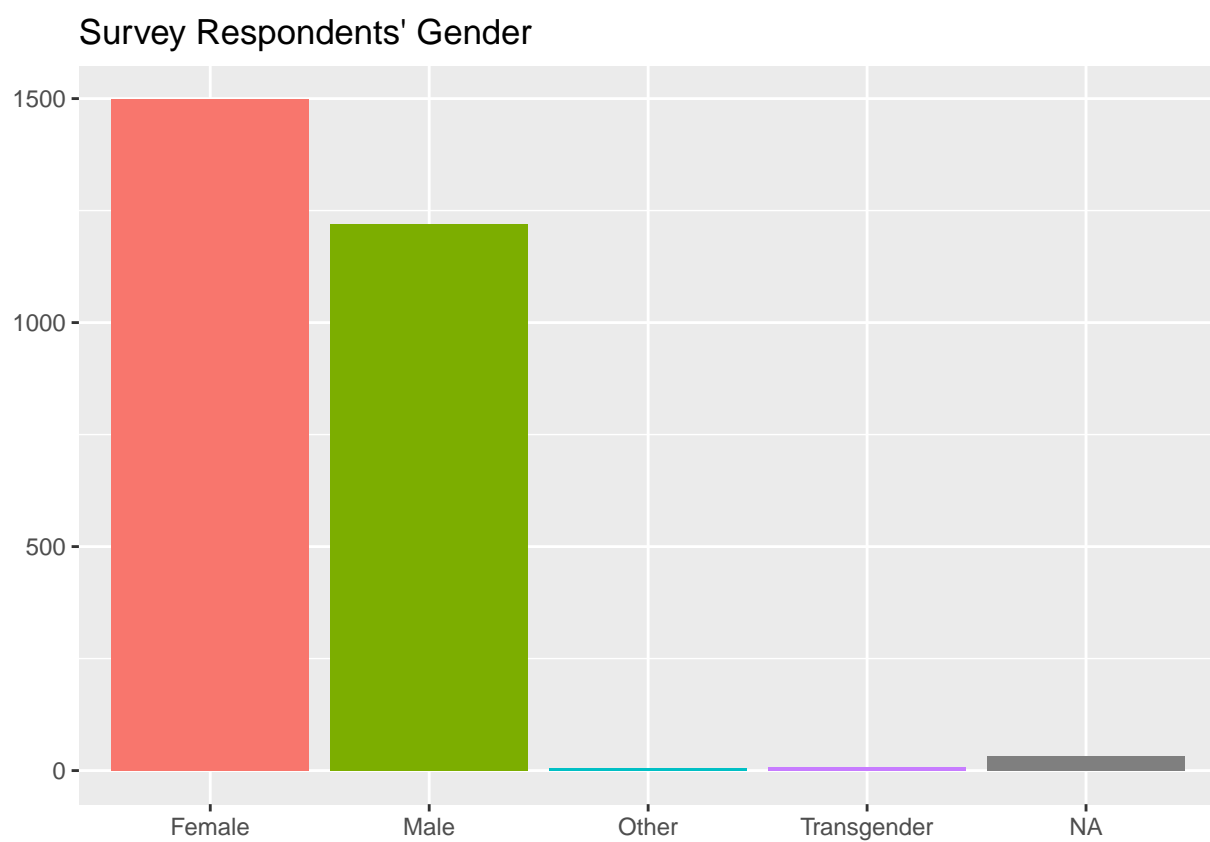Figure 2: Distribution of survey respondents' race.

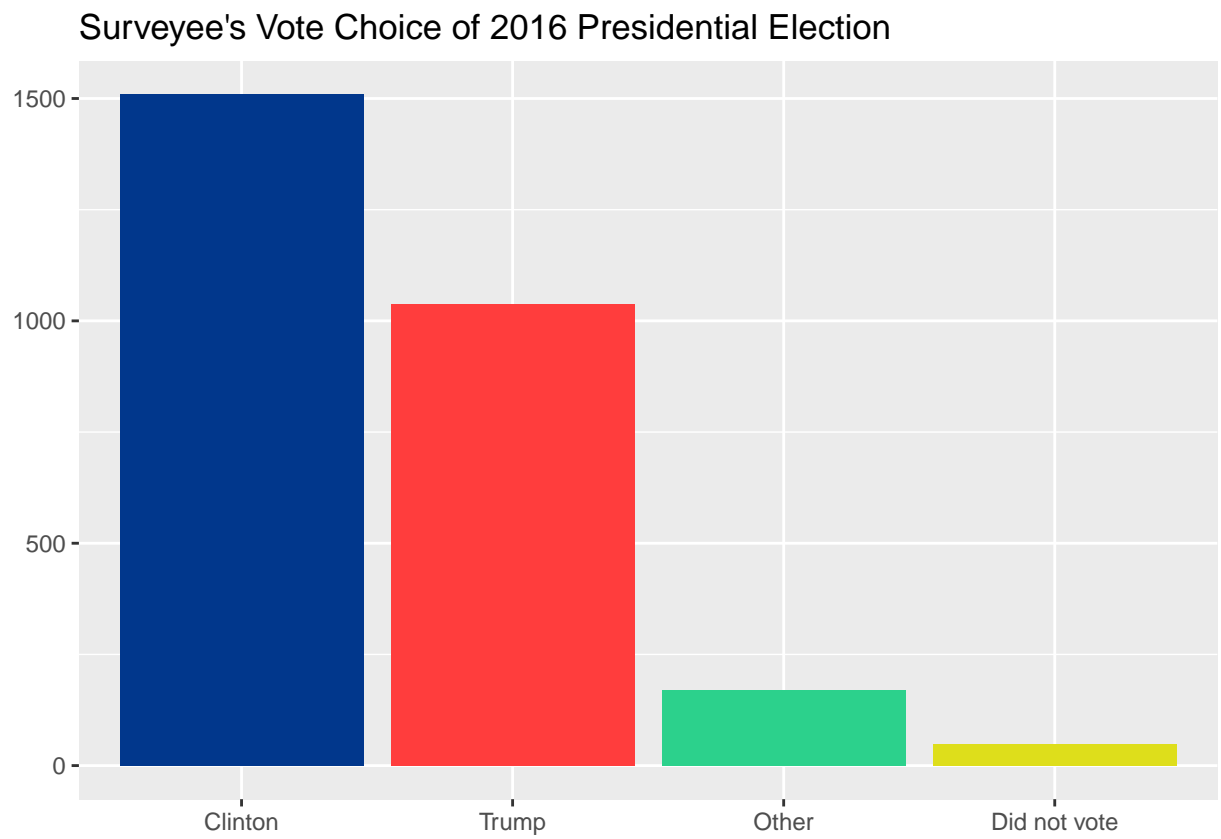Figure 3: Distribution of survey respondents' gender

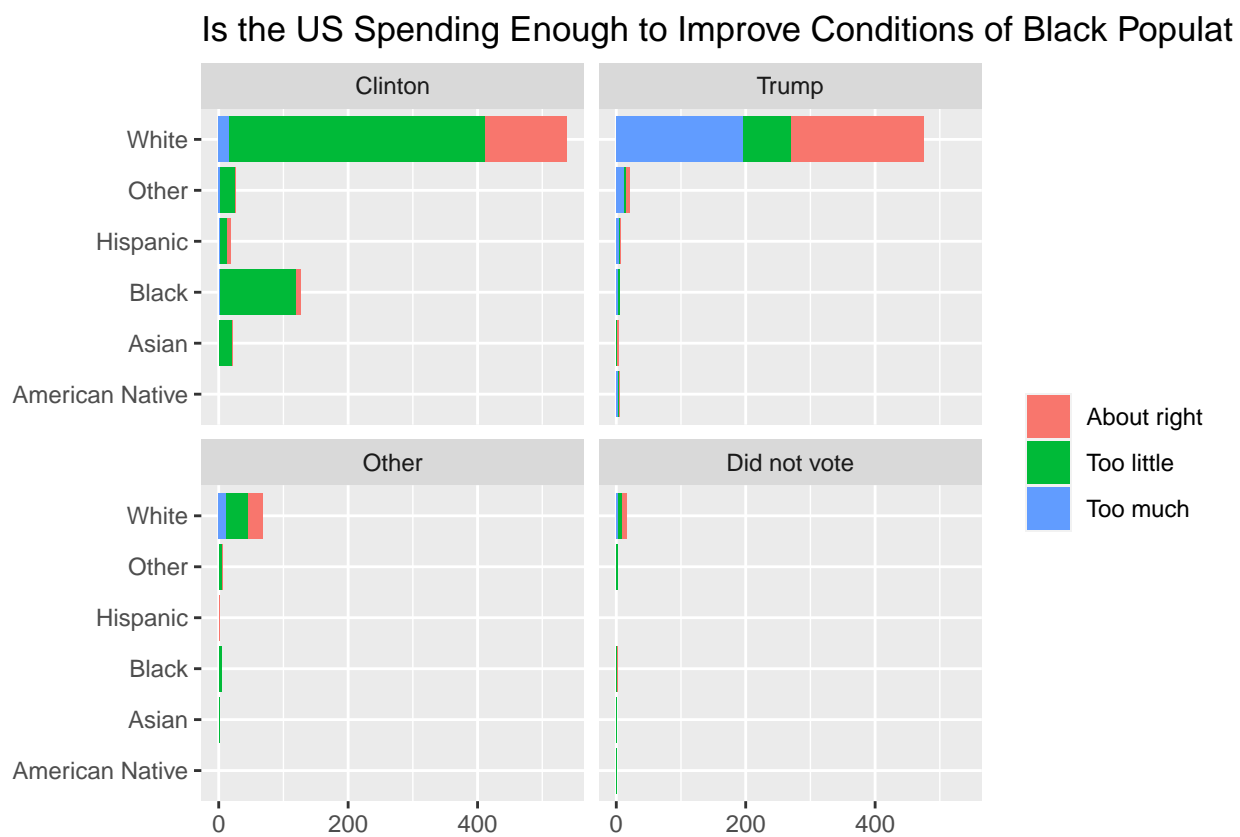Figure 4: Distribution of survey respondents

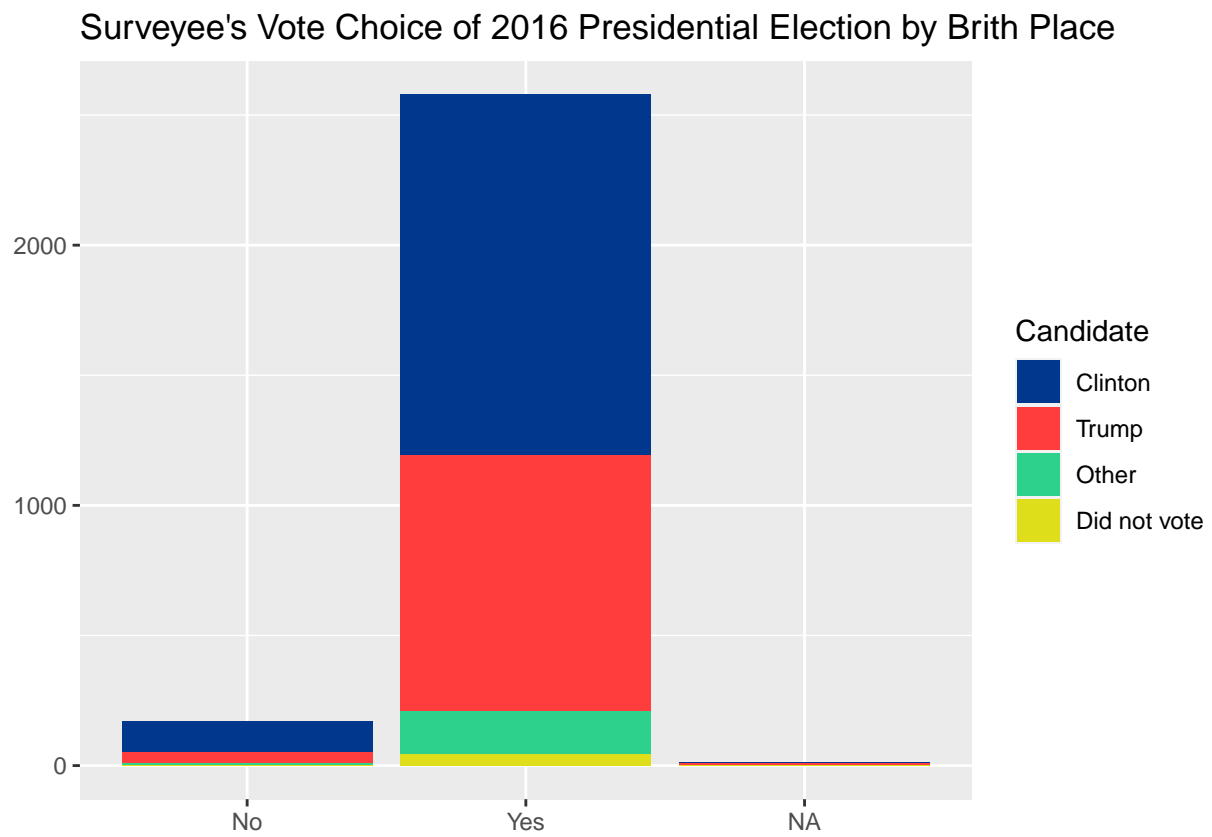Figure 5: Opinion on national spending on improving conditions of black population

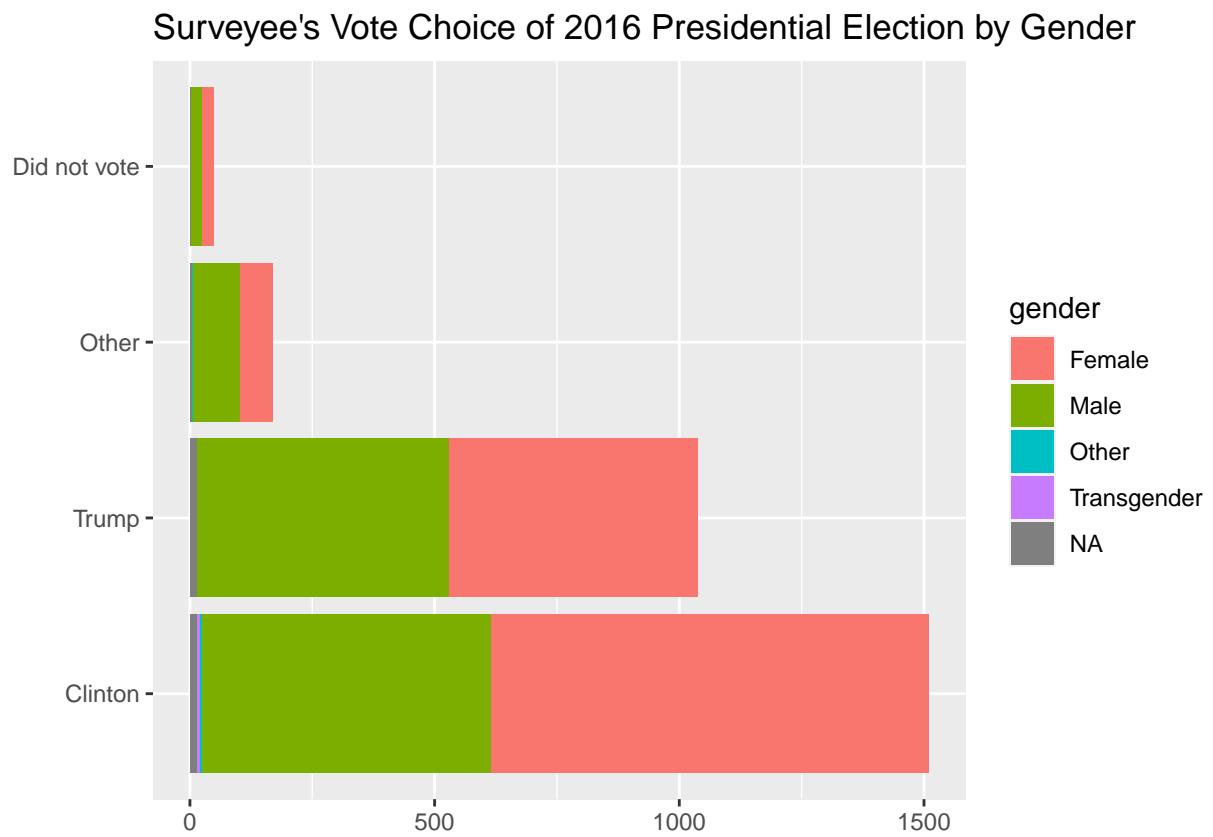Figure 6: Distribution of survey respondents' vote choice by birth place

Figure 7: Distribution of survey respondents' vote by gender
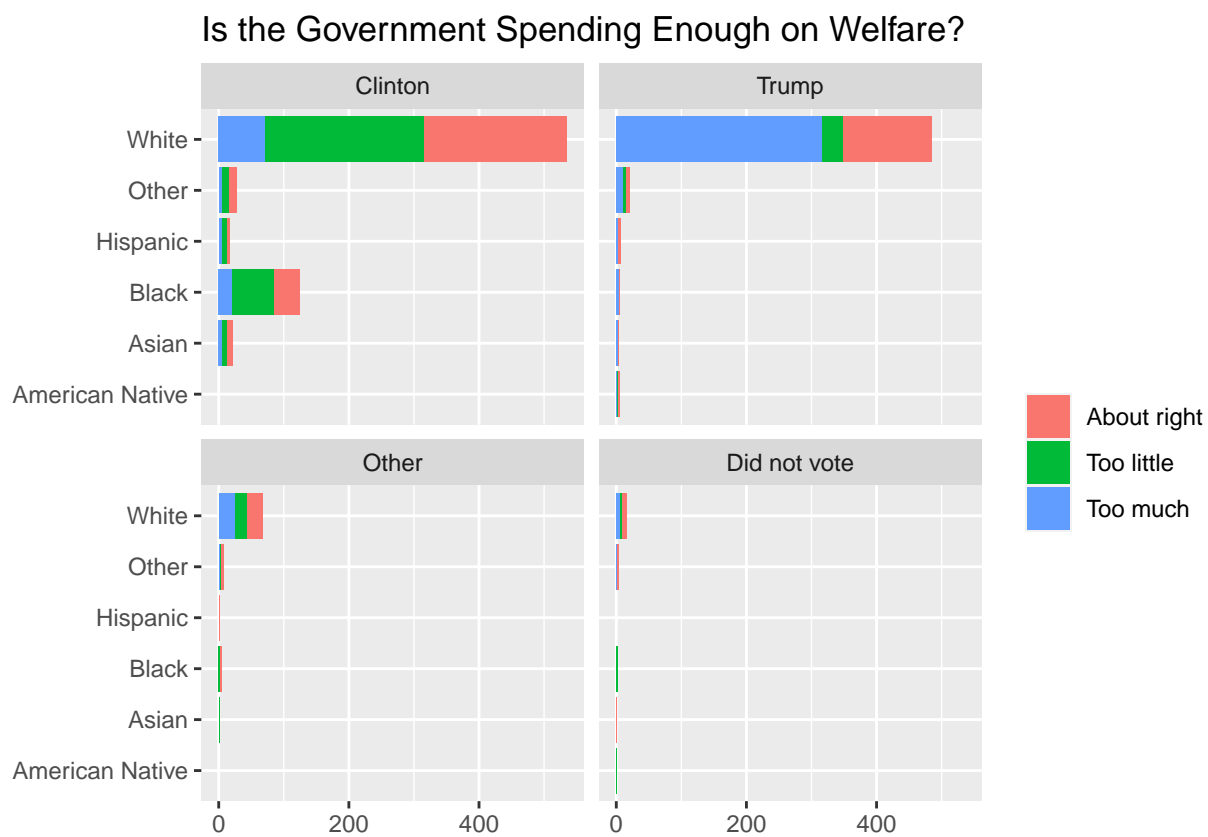
Figure 8: Opinion on national spending on welfare
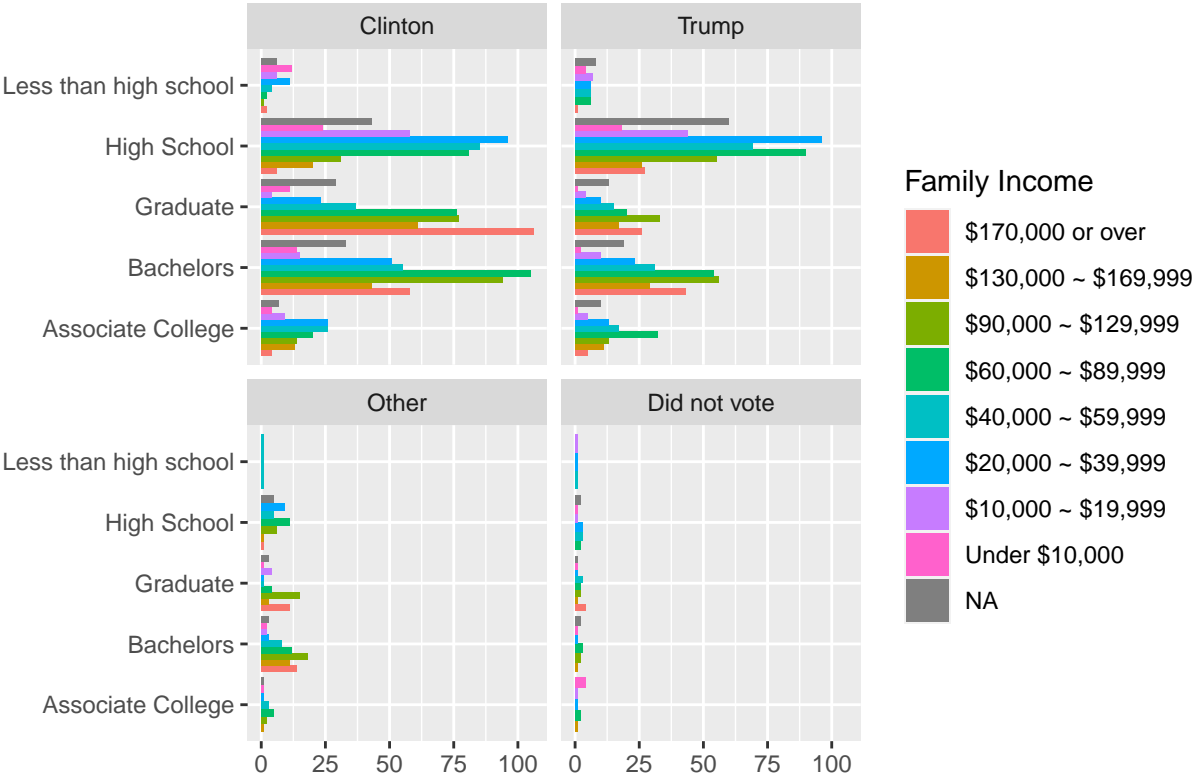
Education, Income and its Effect to the Vote

Figure 9: Distribution of votes by degree and family income
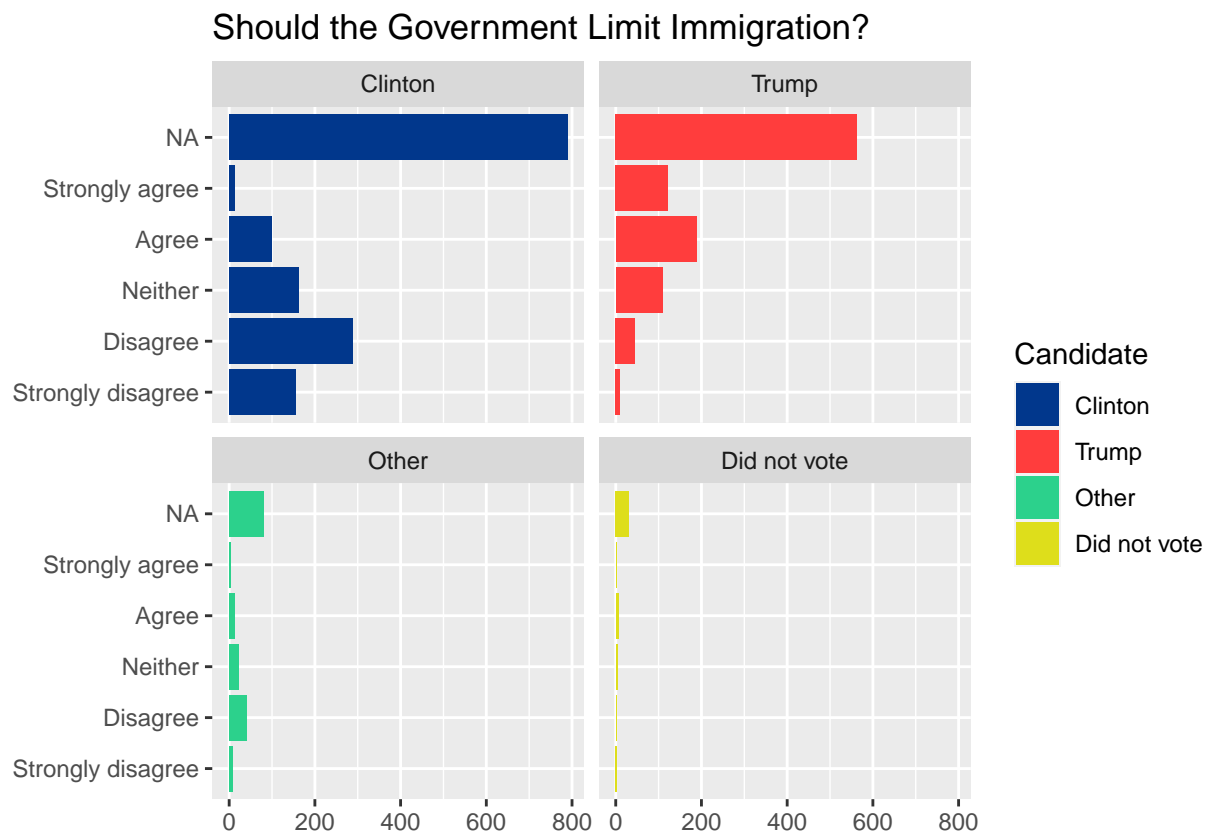
Figure 10: Opinion on wheter the US should limit immigration to protect their culture
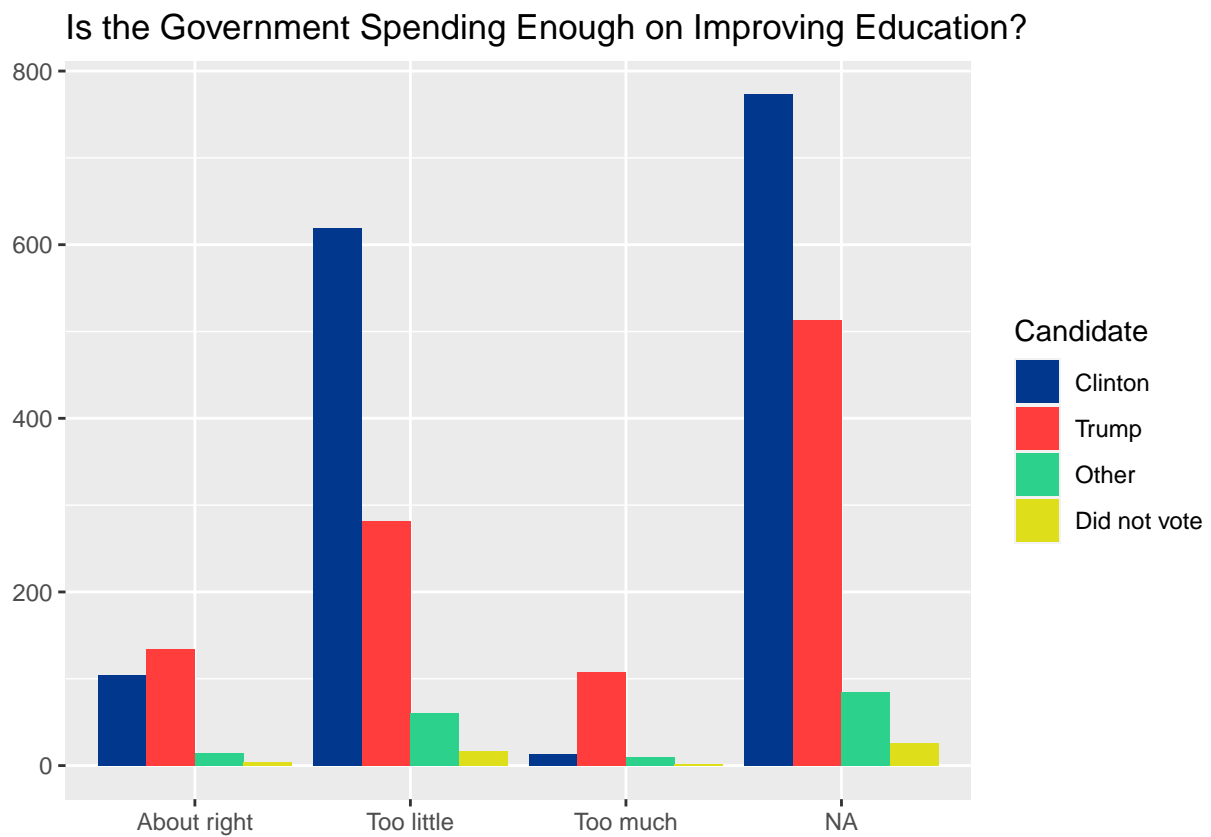
Figure 11: Surveyee's opinion on government's budget on education improvement
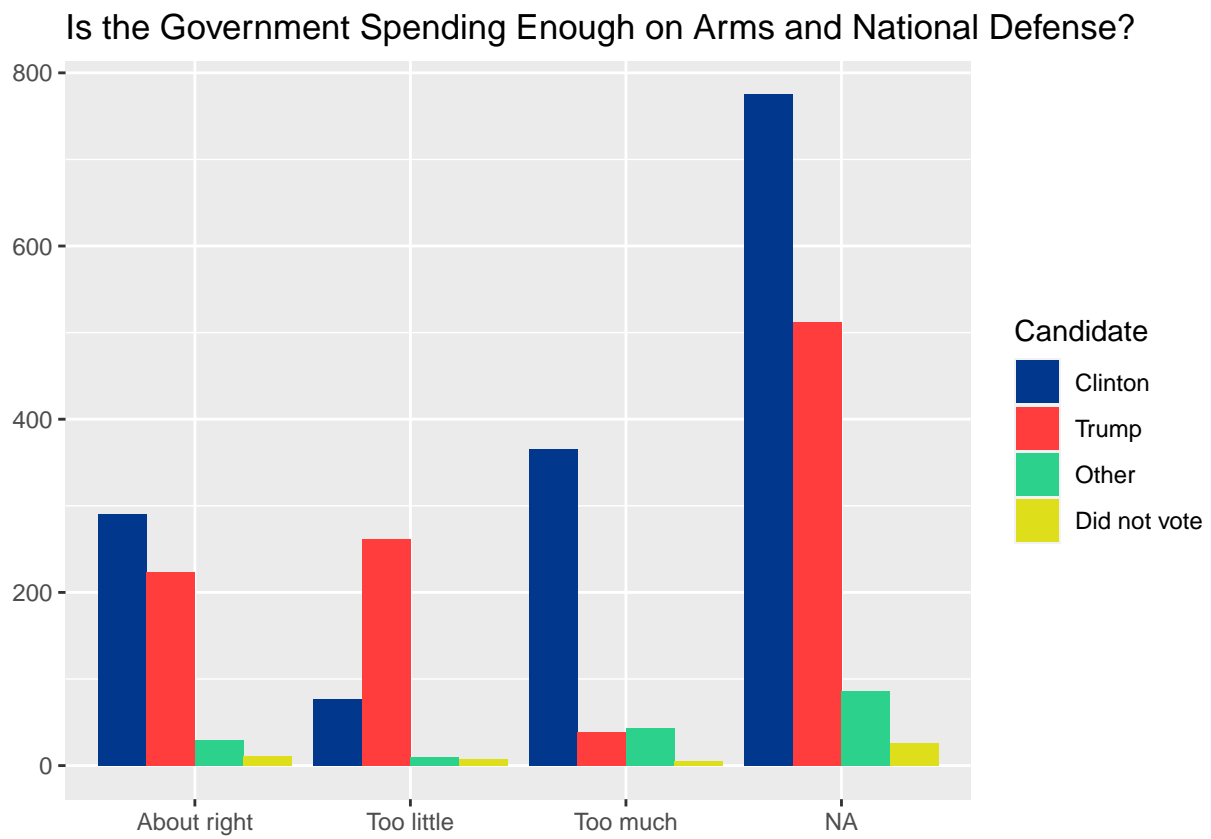
Figure 12: Surveyee's opinion on government's budget on arms
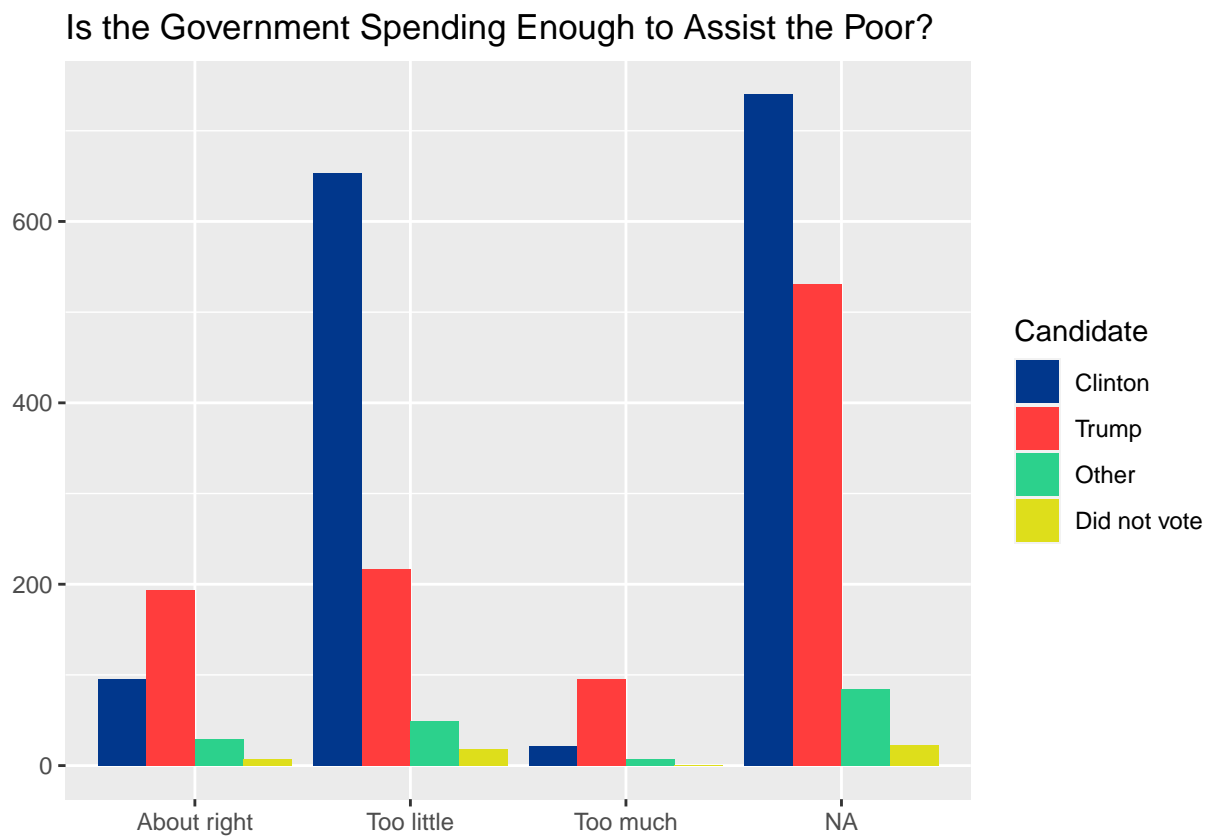
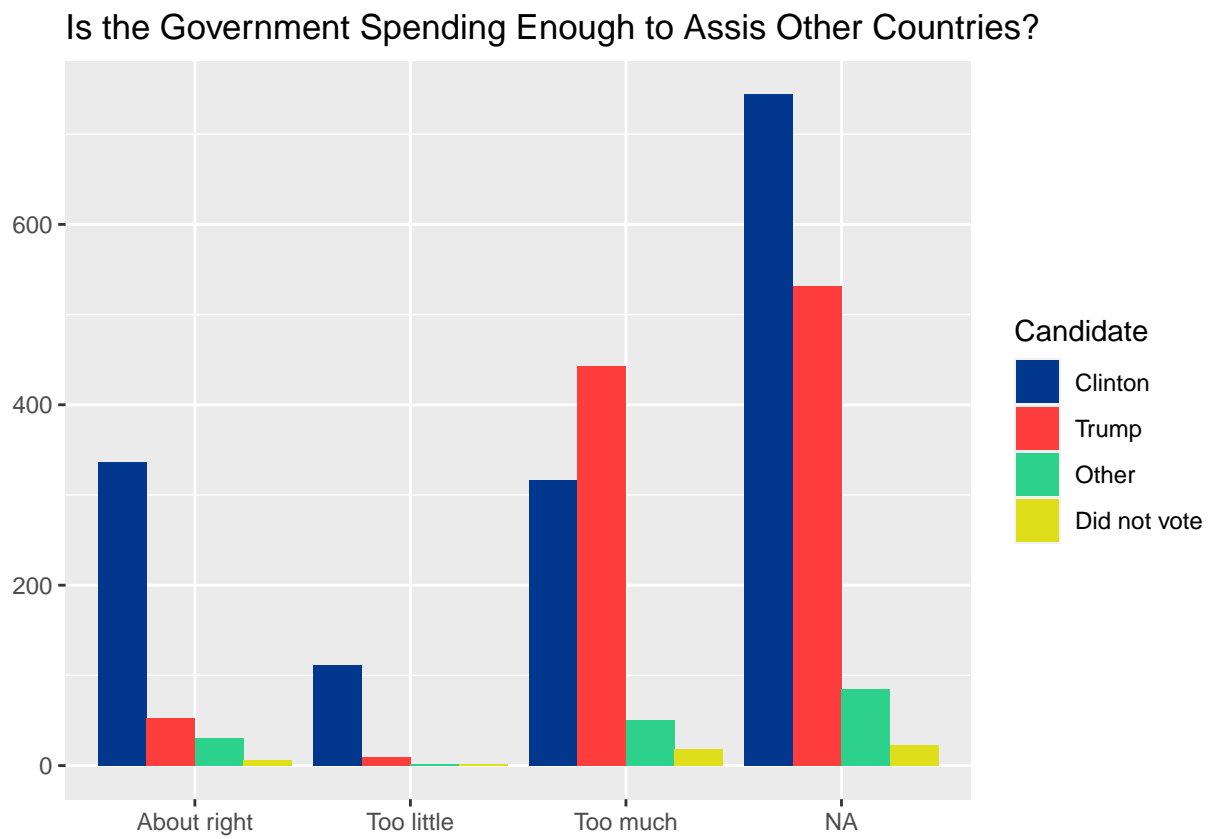Figure 13: Surveyee's opinion on government's budget on assisting the poor

Figure 14: Surveyee's opinion on government's budget on assisting other countries

3 Results

The questions asking personal information shows some obvious but also interesting results. Figure 1 shows that younger generation did not participate in the survey as much as older generation. It can also be seen that while almost all age groups preferred Clinton over Trump, compared to younger generation, older age groups shows more interest to Trump. Females participants outnumbers male participants by significant amount as seen on Figure 3, and many of female participants' vote went to Clinton, while male votes divided almost evenly(Figure 6). Figure 8 shows some very interesting responses. Considering the distribution of votes, votes of white population are fairly split between Clinton and Trump, with more votes to Clinton, but all other races heavily favors Clinton. Some aspects that are not so obvious are ones involving education, family income and its correlation with the candidate preference. AS seen on figure 9, People with higher education and higher income voted for Clinton while lower income and education votes are quite evenly split.

Figures 9-14 focuses on the questions that involves asking their opinions of various subjects of government policies and how supporters of different candidate , figure 8 shows that white population's general opinion is that the government's budget on welfare is not lacking, with many of them thinking government is spending too much, whereas other races opinions are evenly split with black community leaning more towards the opinion that it is lacking. We can also observe from the same figure that Clinton supporters generally think that the budget on welfare is lacking and Trump's supporters think budget is too much. Figure 10 shows answers to question on immigration and whether the government should limit immigration to protect american culture, and it shows that out of the people who answered the question, majority of Clinton voters disagrees and majority of Trump voters agrees. We start to see a trend here that Trump supporters have more conservative opinions and Clinton supporters are more open to ideas on improvement in social equality or welfare. We see this trend continue on other figures. On Figure 12, many Clinton voters think government is spending too much of national defense while very few Trump voters think the same, while many trump voters think they spend too little when very few Clinton voters have the same opinion. For the question on if they think the government is spending enough to assist other countries (Figure 14), while Clinton's supporters opinions are divided quite evenly between "about right", and "too much", majority of Trump supporters think that the government is spending too much. Figure 11 and Figure 13 shows where Clinton voters and Trump voters share the same opinion on government spending on education and to assist the poor. While majority of people who thinks that the government is spending too much on both occasions are Trump supporters, popular opinions from both parties is that the government is spending too little on education and to assist the poor.

4 Discussion

In this session, we are going to explore important questions that can be raised from our findings. At this point, you may end up with question marks in meanings this analysis possess. Although this paper is a research of the past, this definitely has something to tell us about the U.S or even the world. Hopefully following questions helps you to understand our paper.

4.1 Summary

We have taken survey data to observe the trend in voters of 2016 presidential election of the United States. In general, just as we expected by the trend we saw during our observation of the data, Hillary Clinton supporters were moved by improvement in social equality or welfare such as "Assistance to the Poor", "nation welfare", "nation education" and "Assistance to the other countries". On the other hand, Donald Trump supporters were interested by more conservative ideas that US is spending too much on the sessions above, and shows their support to Trump's slogan "Make America Great Again".

4.2 Disagreement Between Races, Gender, Ages, and Wealth

Even though the United States and many other countries around the world have been trying hard to improve social welfare and equality, throughout this paper we have observed that there still exists disagreement within gender, race, wealth and age, and this had massive impact on politics and was crucial factor on people's decision on who to vote for as the president. Despite the popular movements voicing for equality between genders, and races and such in recent years, it seems that the opinions of people in different groups, are very divided.

4.3 Weaknesses and how we should proceed in the future

The data from U.S GSS contained large amount of data and information and the distribution of survey respondents were quite accurate in terms of race. However, there were significantly more females than males and significantly less people of ages in the 20s, 30s and 40s. It is important, especially for the topic of this paper, to have the data on the younger generation and how they view government policies and how it effects their vote for the presidential election. To gain more knowledge in this trend of disagreement between people of different race, gender, age and wealth when it comes to the election, an additional survey could be made to answer WHY this disagreement is happening, by not just asking for their opinion but also asking about what they experience and witness. In the future, survey should be including questions that can allow us to better capture and understand the current trend in agreement and disagreement between different groups of people, it can allow us to study how it could effect the election in the future.

Appendix

1) Figure 6 represents people who were born in US and their votes on the candidates. Most respondents were born in US.

2) Figure 8 represents national spending on welfare. People who advocated more funds in improvement of welfare shows preference in Clinton and people who did not shows preference in Trump.

3) Figure 10 represents whether U.S should limit immigration in order to protect their own cultures or not. Even though many answered "Not Applicable", we can observe two different patters in Clinton and Trump.

4) Figure 11 represents national spending on education with votes. People who want more funds in education improvement shows preference in Clinton and vice versa.

5) Figure 12 represents national spending on the arms with respect to votes for the candidates. People who thought too much clealry shows the preference toward Clinton.

6) Figure 13 represents national spending on the poor with respect to respondents' votes on the candidates. From "About right" to "Too much", there exits more Trump supporters than Clinton's and vice versa.

7) Figure 14 represents national spending on aids to foreign countries with respect to the votes for candidates . There exists more Trump supporters in "Too much" comparing to Clinton's supporters.

A Supplementary survey

Our supplementary survey is available here: LINK

A.1 Preamble:

The purpose of this survey is ~~

By proceeding with this survey, you understand that we will use your responses to better understand the 2016 U.S presidential election in the GSS. The survey is voluntary, and if you decide to participate you can skip questions and withdraw at any time.

A.2 Questions

1.
2.
3.

A.3 End Page

Thank you ~~

B   References

U.S. Census Bureau quickfacts: United States. (n.d.), from https://www.census.gov/quickfacts/fact/table/US/PST045221

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation.* https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r.* https://yihui.org/knitr/.

———. 2022. *Bookdown: Authoring Books and Technical Documents with r Markdown.* https://github.com/rstudio/bookdown.