Behind Biden's 2020 Win: Analyzing Key Influencers through Regression Models and Visualizations*

Pengyu Sui   Lin Dai   Xuecheng Gao

March 16, 2024

This article replicates data to analyze the relationship between demographic variables and voting behavior in the 2020 presidential election in the United States. We study the correlation between demographic characteristics and voting behavior by analyzing voters' gender, race, age. By analyzing the 2020 U.S. election, it provides a new perspective on the future electoral pattern and provides new ideas for studying the world political electoral system.

Table of contents

1 Introduction 2

2 Data 3
2.1 Data Source . . . 3
2.2 Data clean . . . 3
2.3 Methodology . . . 4
2.4 Variables . . . 4
2.5 Measurement and visualization . . . 5

3 Model 7
3.1 Model set-up . . . 7
3.2 Model justification . . . 8

4 Results 9
4.1 Regression Model . . . 11

*Code and data are available at: https://github.com/simon0202sui/Political-support-in-the-US-2020.git

4.2 Race . . . 13

5 Discussion **14**
5.1 Bias . . . 14
5.2 Possible reason . . . 14
5.3 Limitations and Future Directions . . . 14

References **16**

1 Introduction

The 2020 US presidential election is a major historical event. This election has attracted not only the attention of the United States, but also the world. In that election, there was a very tight race between then-incumbent President Donald Trump and then-Democratic candidate Joe Biden. The 2020 U.S. election took place during the COVID-19 epidemic, which had a huge impact on American society, economy, and politics. At the same time, in 2020, many racial discrimination incidents occurred in the United States, which eventually triggered a large number of demonstrations, protests and social unrest. In such a social environment at the time, candidates and voters focused on issues such as government policies, how to respond to the U.S. economic recession, and how to ensure racial and gender equality. These background factors have led to a very special attitude of electors and voters in the 2020 U.S. presidential election. Against this background, Trump and Biden competed fiercely for votes, and ultimately Joe Biden won and became the President of the United States.

The presidential election in the United States uses the Electoral College. This system was developed to maximize the people's electoral rights and at the same time protect the stability of American politics. The Electoral College allocates 538 votes to each state in proportion to population. The votes are counted based on these 538 electoral votes. The person with more than 270 votes wins the election. When voters in the United States vote in the general election, they choose the electors who represent their states. There are 538 electors in total, representing the 50 states and Washington, DC.

This study hopes to further analyze the relationship between voters and electoral preferences, so we selected several variables to observe how the preferences of voters are affected in the 2020 U.S. presidential election.

2 Data

2.1 Data Source

In our research, we meticulously selected three pivotal factors for analysis: age group, gender, and race. This choice was informed by our aim to delve into the nuanced ways these variables influence political preferences and voter behavior in the context of the 2020 U.S. Presidential Elections. Utilizing the data from the 2020 Cooperative Election Study (CES) hosted in the Harvard Dataverse Repository, which comprises both pre-election and post-election inquiries with over 61,000 respondents, we sought to uncover insights into the electorate's dynamics (Schaffner, Ansolabehere, and Luks 2021). The CES dataset was particularly suitable for our study due to its comprehensive coverage of key topics relevant to our research questions, including detailed demographic information and voting patterns, thereby providing a rich basis for analysis (Schaffner, Ansolabehere, and Luks 2021).

The CES dataset's breadth, coupled with its public availability and the structured questionnaire that encompasses a wide range of topics — from demographic details to political leanings and issues of national concern — made it an optimal resource for our investigation. Our research leverages this dataset to explore the intricate relationships between voters' age groups, gender, race, and their electoral choices in the 2020 elections. However, it should be emphasized that due to the inherent flaws of the survey format, our statistics (support rate, number of supporters) will deviate from the real 2020 data as explained in detail in the discussion section of the postgraduate study Section 5.

2.2 Data clean

In this study, we conducted thorough cleaning and modification of the survey data to ensure its suitability for analysis. Initially, we loaded the data using the read_csv from readr package (Wickham, Hester, and François 2021) in R (R Core Team 2023) function and renamed variables for clarity and interpretability. Subsequently, comprehensive cleaning and modification steps were performed using the mutate and filter functions from the dplyr package (Wickham et al. 2021). We employed the case_when function to recode categorical variables such as "vote_intention," "support_biden," "party," "education," "race," "age_group," and "gender" into more meaningful categories. It's worth noting that age groups were delineated based on birth year ranges for analytical purposes. Finally, we filtered out observations containing "Unknown" values in any categorical variable, resulting in a dataset ready for further analysis. This rigorous data cleaning process ensured the integrity and reliability of the data, providing a solid foundation for subsequent statistical modeling and inference.

2.3 Methodology

To conduct our analysis, we employed the R programming language, complemented by an array of packages that facilitated our data handling, visualization, and statistical modeling processes (R Core Team 2023). These included tidyverse for comprehensive data manipulation (Wickham et al. 2019), dataverse for seamless access to the Dataverse datasets (Kuriwaki, Beasley, and Leeper 2023), arrow for efficient data integration (Richardson et al. 2024), and rstanarm for advanced Bayesian modeling (Goodrich et al. 2022). Additionally, we utilized knitr for dynamic reporting (Xie 2023), here for effective path management (Müller 2020), ggplot2 for sophisticated visualizations (Wickham 2016), scales for visual enhancements (Wickham, Pedersen, and Seidel 2023), modelsummary for concise model summaries (Arel-Bundock 2022), and marginaleffects for calculating the predictive impacts of our selected variables (Arel-Bundock 2023). This robust analytical framework enabled us to rigorously examine the impact of age group, gender, and race on voting behavior and political preferences.

2.4 Variables

There have 2 file each are clean data and simulated data, for each two file are both csv file. In the cleaned data has 7 variables, each are vote_intention, support_biden, party, education, race, age_group, gender. In the simulated data has 4 variables, each are race, age_group, gender, support biden. Because of the raw data is 800MB with 1000+ variable, it will not list each one here. You can find the variable statement in the /other/guide.

<u>Clean data</u>

- vote_intention, (1-Yes, 2-No, 3-other) (num)
- support_biden, (No, Yes, Other) (char)
- party, (Democratic, Republican, other) (char)
- education, (No HS, High school, College, Postgrad, other) (char)
- race, (White, Black, American Indian, Asian, Other)
- age_group(18~39, 31-61, 62-95) (char)
- gender (Male, Female) (char)

<u>Clean data</u>

- race, (White, Black, American Indian, Asian, Other) (char)
- age_group, ('18-30', '31-61', '62-95') (char)
- gender (Male, Female) (char)

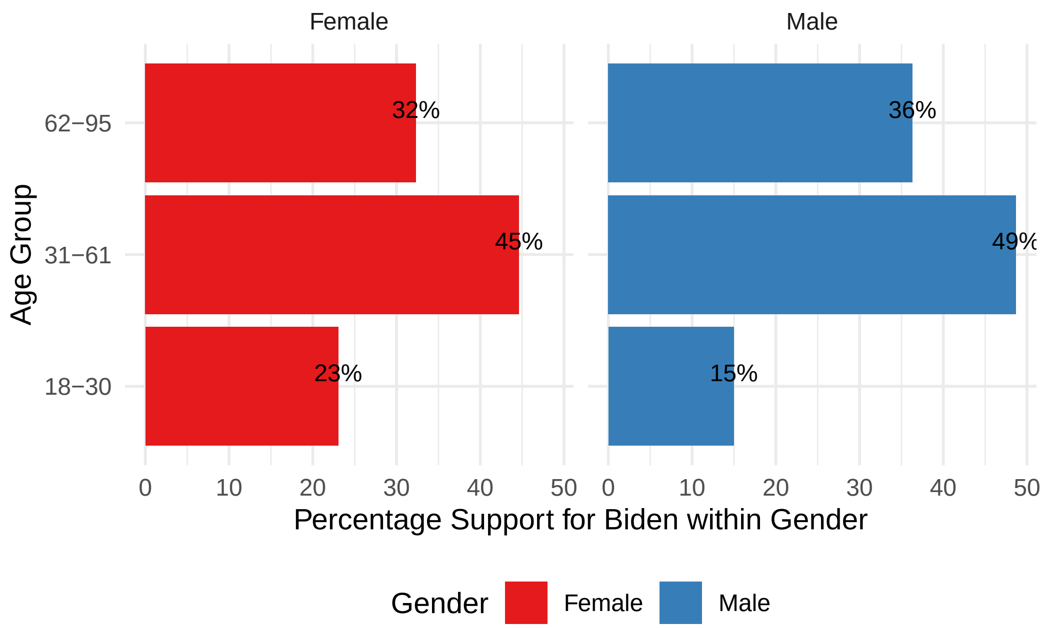

Figure 1: Percentage of Biden Support by Age Group and Gender

2.5 Measurement and visualization

In our analysis, we probed the relationship between age, gender, and support for Joe Biden in the 2020 Presidential Election. The dataset, after being meticulously cleaned and filtered to include only registered adult voters, was analyzed to display this dynamic. Our Figure 1, captioned "Percentage of Biden Support by Age Group and Gender," presents this relationship visually.

The process involved loading the cleaned data, , into R, where we singled out individuals who expressed support for Biden. We then grouped these individuals by age and gender to count the number of Biden supporters within each subgroup. Afterward, we aggregated these counts to find the total number of supporters by gender across all age groups.

By rejoining this data, we calculated the percentage of Biden supporters within each age group for each gender. This calculation was essential for our comparative analysis, and we visualized our findings in a horizontal bar chart. The chart distinctly illustrates the proportions with bars color-coded by gender and annotated with the exact percentages. This was achieved using the package, which allowed us to create a clear and aesthetically pleasing representation of the data, further facilitated by the division of the chart into separate panels for each gender for a straightforward comparison (Wickham 2016).

This visualization strategy was not only about presenting raw numbers but also about interpreting the data to gain insights into voter behavior patterns across different demographics. The chart serves as an empirical basis to discuss how age and gender intersect to affect political preferences, as indicated by support for Biden. It is through such an analysis that we seek to uncover underlying trends and factors that might have influenced the election outcome.

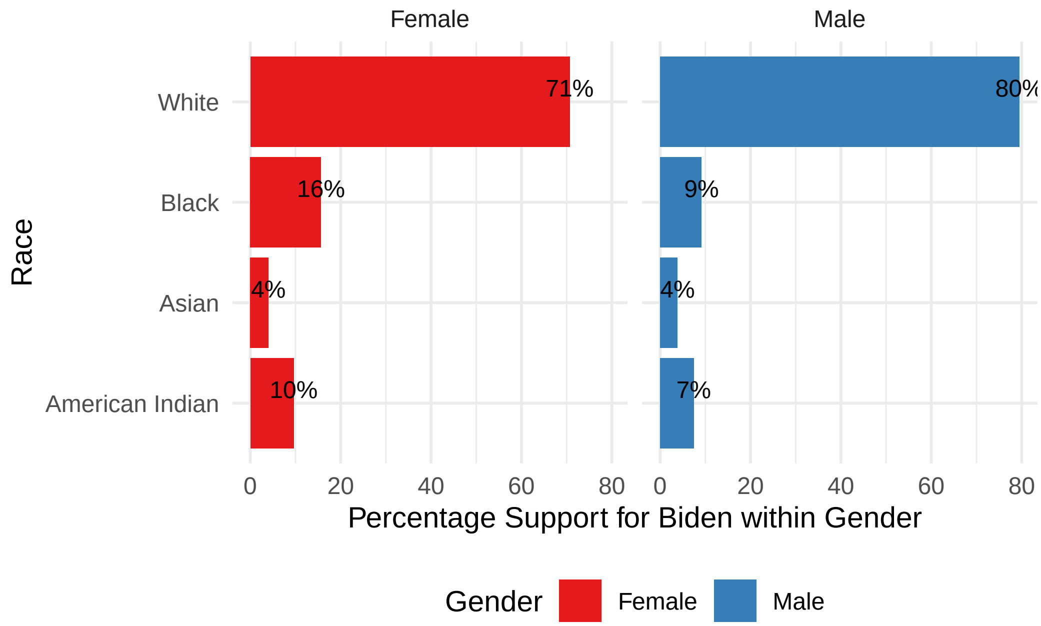

Figure 2: Percentage of Biden Support by Race and Gender

Additionally, we further dissected the electorate's support for Joe Biden in the 2020 Presidential Election by examining the intersection of race and gender. This approach allows us to understand the diversity of Biden's support base across different racial groups within each gender category. Our Figure 2, entitled "Percentage of Biden Support by Race and Gender," showcases these insights.

To construct this figure, we began by filtering the cleaned dataset to focus on respondents who supported Biden. We then proceeded to group these Biden supporters by race and gender to calculate the number of supporters within each intersectional category. Following this, we summed the total number of Biden supporters for each gender, which facilitated the calculation of the percentage of supporters by race within each gender.

Using the ggplot2 package (Wickham 2016), we visualized the calculated percentages in a horizontal bar chart. This chart is organized by race on the x-axis and displays the percentage of Biden's support on the y-axis, with bars filled according to gender. The bars are also annotated with the percentage values, providing a clear, numerical understanding of the data. To enhance comparison, the data is plotted with separate panels for each gender, enabling an at-a-glance assessment of support patterns across genders and races.

The minimalistic design of the chart, achieved with a theme that prioritizes data visibility, and the coordinated color palette allow for a straightforward interpretation of complex intersectional data. The placement of the legend at the bottom of the chart ensures that it complements the visual narrative without distracting from the primary data representation.

By interpreting this figure, we aim to draw conclusions about the distribution of Biden's support among racial groups and compare how this support varies between male and female voters. Such detailed visualization underscores the multifaceted nature of political support and is instrumental in our comprehensive analysis of voting behavior.

3 Model

3.1 Model set-up

Our analytical strategy is predicated on predicting the probability of an individual voting for Biden, utilizing predictors such as gender, race, and age group.

The Bayesian logistic regression model employed in our study is articulated as follows:

$$\begin{aligned}
y_i|\pi_i &\sim \text{Bernoulli}(\pi_i) && (1)\\
\text{logit}(\pi_i) &= \alpha + \beta \times \text{gender}_i && (2)\\
&+ \gamma \times \text{race}_i && (3)\\
&+ \delta \times \text{age group}_i && (4)\\
\alpha &\sim \text{Normal}(0, 2.5) && (5)\\
\beta &\sim \text{Normal}(0, 2.5) && (6)\\
\gamma &\sim \text{Normal}(0, 2.5) && (7)\\
\delta &\sim \text{Normal}(0, 2.5) && (8)
\end{aligned}$$

In this formulation:

y_i denotes the binary outcome for each respondent, coded as 1 if they voted for Biden and 0 for they didn't vote for Biden, π_i represents the probability that respondent i cast their vote for Biden, $gender_i$ is the gender of respondent i, $race_i$ delineates the race of respondent i, $agegroup_i$ categorizes the age group of respondent i. For the analysis, we leveraged the rstanarm package (Goodrich et al. 2022) for its Bayesian logistic regression capabilities, alongside the tidyverse package (Wickham et al. 2019) for data manipulation, and ggplot2 (Wickham 2016) for data visualization, all within the R statistical environment (R Core Team 2023). Logistic regression models are well suited for binary outcome data and, by employing the logit function, allow for the linear combination of predictor variables.

The Bernoulli distribution serves as the foundation of the logistic model, handling the binary nature of our response variable. The parameters of the model—α, the intercept, and β, γ,

δ, the slopes for each predictor—are assigned Normal priors with a mean of 0 and standard deviation of 2.5, following the default prior settings in rstanarm. These priors are weakly informative, letting the data largely influence the posterior distributions.

This model structure is executed to discern the impact of demographic factors on voter preference in the 2020 Presidential Election, intending to quantify these effects and understand their robustness within a probabilistic framework.

3.2 Model justification

In our study, we formulated a Bayesian logistic regression model to analyze the determinants of voter preference in the 2020 Presidential Election. The model was designed to predict the likelihood of an individual voting for Biden, using demographic factors as predictors. Here's a detailed explanation of the model components and the rationale behind our methodological choices:

Model Components and Their Interpretation: Binary Outcome Variable (y_i): This variable captures the voting decision of each respondent, with a value of 1 indicating a vote for Biden and 0 for not. The binary nature of y_i aligns with the logistic regression framework, which is optimized for analyzing dichotomous outcomes.

Probability of Voting for Biden (π_i): This represents the model's estimated probability that respondent i voted for Biden, based on their demographic characteristics. The logistic regression model calculates π_i through the logit link function, transforming the linear combination of predictors into a probability between 0 and 1.

Predictor Variables: The model includes three key predictors: Gender (gender_i): Reflects the gender identity of respondent i, recognizing gender as a significant factor in political preferences. Race (race_i): Captures the racial identity of respondent i, acknowledging the role of race in shaping electoral behavior. Age Group (age group_i): Categorizes respondents into age groups, allowing us to examine the influence of generational factors on voting patterns. Analytical Tools and Statistical Framework: Bayesian Logistic Regression (rstanarm): We utilized the rstanarm package (Goodrich et al. 2022) for its advanced Bayesian logistic regression capabilities, facilitating the incorporation of prior knowledge and the estimation of posterior distributions for model parameters. We chose to utilize a logistic regression model to analyze voter preferences because of its significant advantages in handling binary classification problems. Voter preferences can typically be regarded as a binary problem, i.e., choosing to support a particular candidate (such as Biden) or not. The logistic regression model is specifically designed to address such types of problems, transforming the linear combination of input features into probabilities between 0 and 1 to predict whether an event occurs. Moreover, the probabilities outputted by the logistic regression model are interpretable, allowing us to understand clearly the extent of influence each predictor variable has on the prediction outcome through the logit function. We assume that voters' gender, race, and age group have linear relationships with their voting preferences, and we estimate the coefficients of each predictor variable to assess their impact on voting preferences. Logistic regression is a classic statistical method widely used in analyzing social science data and predicting behavior, and its robustness and reliability have been validated in practice. Therefore, we believe that the logistic regression model is a suitable tool for analyzing voter preferences, providing us with insightful and easily understandable results.

Data Manipulation and Visualization: The tidyverse package (Wickham et al. 2019) was employed for efficient data handling and preparation, while ggplot2 (Wickham 2016) was used for creating insightful visualizations of our data and model results. These tools enhance our ability to explore, model, and communicate the intricacies of voter behavior.

Bernoulli Distribution and Normal Priors: The choice of the Bernoulli distribution for modeling the binary response variable and Normal priors for the intercept and slopes (α, β, γ, δ) reflects a balance between model flexibility and informative constraints. The Normal priors, with a mean of 0 and standard deviation of 2.5, are weakly informative, allowing the observed data to primarily shape the posterior distributions.

This comprehensive modeling approach aims to illuminate the impact of key demographic variables on voter preference in the context of the 2020 Presidential Election. By leveraging a Bayesian logistic regression framework, we can not only estimate the effects of gender, race, and age group on the likelihood of voting for Biden but also assess the certainty of these effects, providing a nuanced understanding of electoral dynamics within a probabilistic setting.

4 Results

We will integrate the tables and charts derived from the regression model's summary with targeted comparative visualizations to elucidate the influence of race, age group, and gender on the likelihood of voting for Biden. These visual aids will help us dissect the nuances of how each factor contributes to voting preferences.

```
Warning:
`modelsummary` uses the `performance` package to extract goodness-of-fit
statistics from models of this class. You can specify the statistics you wish
to compute by supplying a `metrics` argument to `modelsummary`, which will then
push it forward to `performance`. Acceptable values are: "all", "common",
"none", or a character vector of metrics names. For example: `modelsummary(mod,
metrics = c("RMSE", "R2")` Note that some metrics are computationally
expensive. See `?performance::performance` for details.
 This warning appears once per session.
```

Table 1: **?(caption)**

(a)

	Support Biden
(Intercept)	2.475
genderMale	−0.432
raceAsian	0.831
raceBlack	0.855
raceWhite	−0.726
age_group31-61	−0.164
age_group62-95	−0.668
Num.Obs.	1000
R2	0.053
Log.Lik.	−484.513
ELPD	−491.9
ELPD s.e.	18.0
LOOIC	983.9
LOOIC s.e.	36.0
WAIC	983.7
RMSE	0.40

Explanatory models of political preferences based on gender, race, and Age group

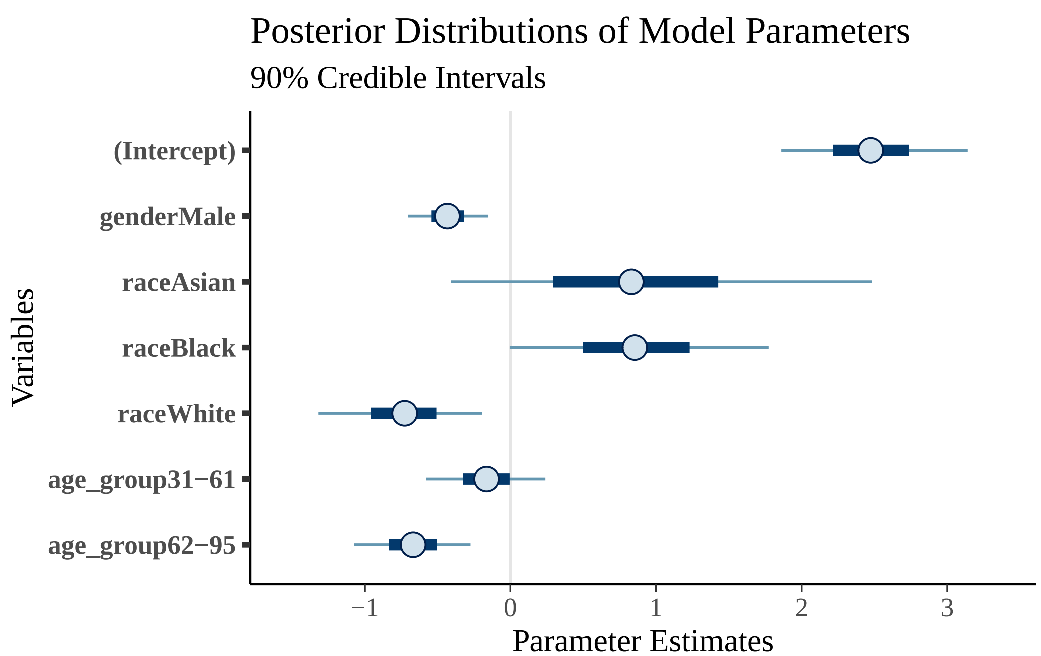

Figure 3: Posterior Distributions of Model Parameters

4.1 Regression Model

Intercept (2.475): The positive intercept suggests that when all the predictor variables are at their reference levels, there is a baseline inclination towards supporting Biden.

genderMale (-0.432): Being male is associated with a decrease in the log-odds of supporting Biden by 0.432 units, suggesting that males are less likely to support Biden compared to the female.

raceAsian (0.831) and raceBlack (0.855): Both Asian and Black races are associated with an increase in the log-odds of supporting Biden by 0.831 and 0.855 units respectively, indicating that individuals of these races are more likely to support Biden compared to the Whites.

raceWhite (-0.726): Being White is associated with a decrease in the log-odds of supporting Biden by 0.726 units, suggesting that White individuals are less likely to support Biden, relative to the reference race group.

age_group31-61 (-0.164) and age_group62-95 (-0.668): Belonging to the age groups of 31-61 and 62-95 is associated with a decrease in the log-odds of supporting Biden by 0.164 and 0.668 units respectively, indicating that individuals in these age groups are less likely to support Biden.

R2 (0.053): The R-squared value is relatively low, indicating that only 5.3% of the variability in support for Biden is explained by the model. This implies there are other factors not included in the model that influence support for Biden. ## Gender

Combining the analysis from the jittered points on the violin plot with the distribution of the probability of voting for Biden across genders, we observe a clear difference in the central tendency of voting probabilities, with women displaying a higher median probability. The actual votes, represented by the jittered points, also align with this conclusion, supporting the notion that the likelihood of voting for Biden differs between genders. Moreover, referencing the regression model summary visualized in Table tbl-model-summary-table and Figure fig-model-summary-fig, the GenderMale coefficient of -0.432 further substantiates that men were less inclined to vote for Biden compared to women. Consequently, we deduce that in the 2020 election, Biden garnered more support from female voters. ## Age Group

According to the histogram Figure 1 in Measurement and visualization part, it can be concluded that among young men aged 18-30, the proportion of women supporting Biden is 23% higher than that of men, 15%. This reflects the lower support rate for Biden among young people, and may be due to young people's low identification with the policies and values promoted by the Democratic Party. Among middle-aged people aged 31-61, 49% of men support Biden, slightly higher than 45% of women. This age group is likely to be the main group supporting Biden. Among the elderly aged 62-95, 32% of women support Biden, slightly lower than 36% of men. A significant proportion of the elderly population has expressed support for Biden, which may be due to dissatisfaction with the Trump administration and the impact of Biden's policy stance on issues such as health protection for the elderly.

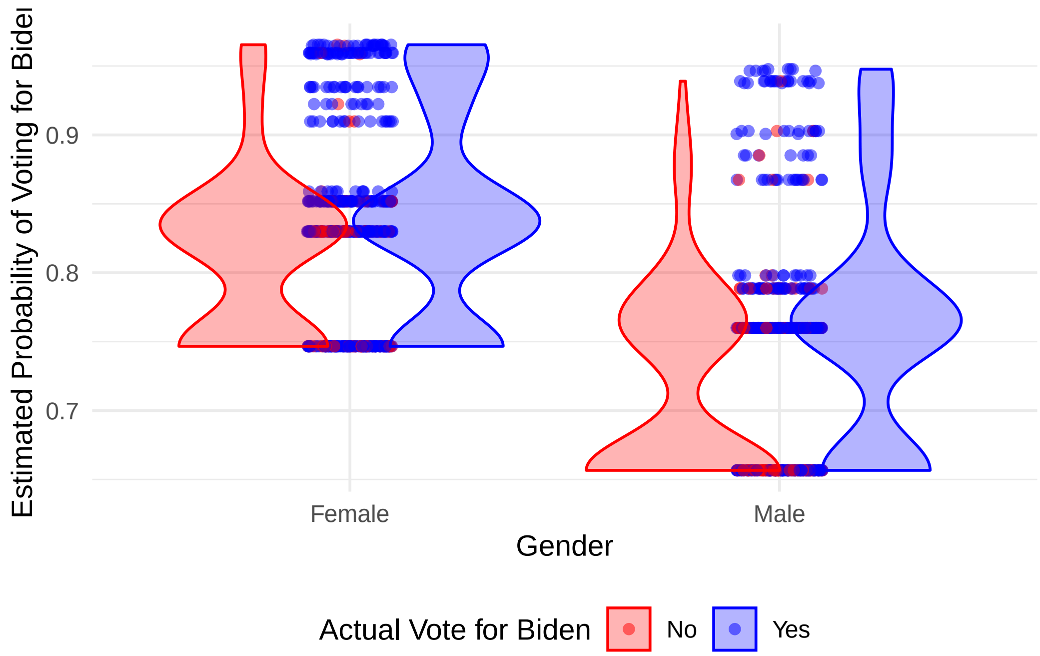

Figure 4: Predicted support of Biden given gender vs. actual support

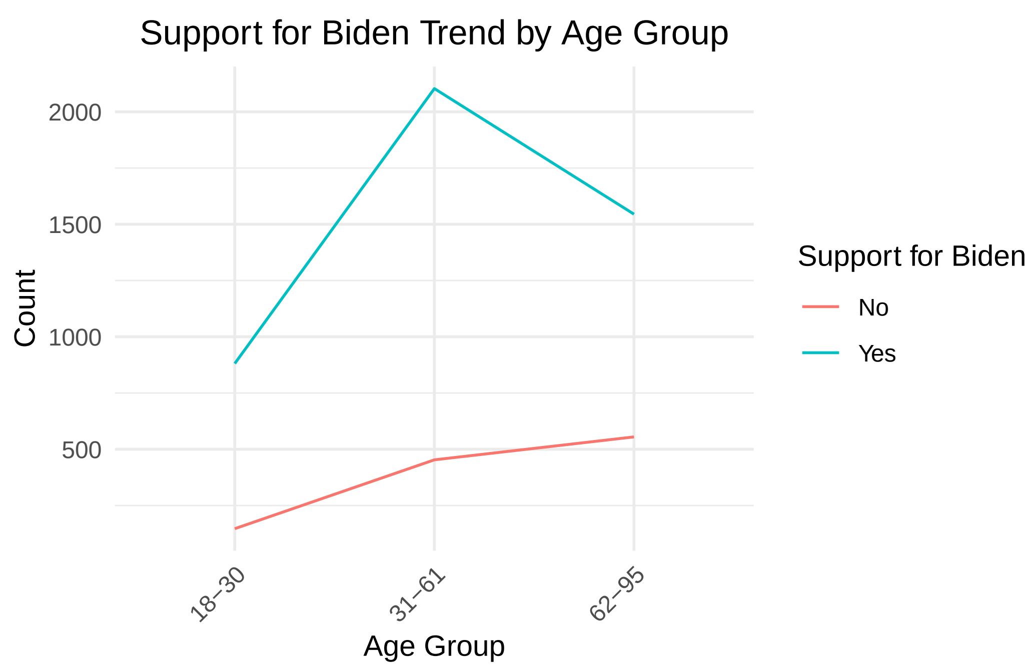

Figure 5: age group and support_Biden

Additionally, combining the insights from this line graph Figure 5 and the results shown by the regression model, it can be observed that the majority of Biden's supporters come from the 31-61 age group, indicating a predominantly middle-aged demographic. In comparison, younger adults who have just reached voting age and the older population above 61 years are less inclined to vote for Biden.# Discussion.

4.2 Race

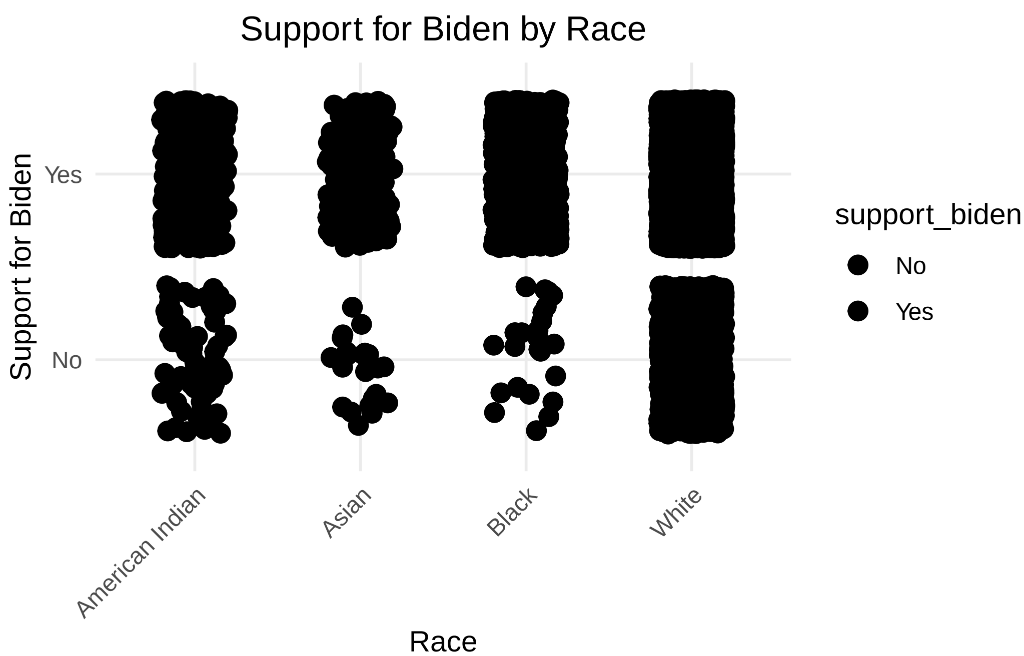

Figure 6: Race and support_Biden

The graph Figure 6 titled "Support for Biden by Race" is a jitter plot that illustrates the relationship between race and support for Biden. The plot segregates respondents into different racial categories: American Indian, Asian, Black, and White. Points are plotted to show individual responses, with the vertical axis categorizing the response into "Yes" for support and "No" for lack of support.

The dense clustering of points at the "Yes" level for races such as Asian and Black could suggest higher support for Biden among these groups. The White category shows a considerable number of points at both "Yes" and "No," indicating a more divided stance on support for Biden within this racial group. American Indian category appears to have fewer data points overall, which may suggest a smaller sample size or less representation in the dataset.

Combined with the bar chart Figure 2 in Measurement and visualization part, White voters have the highest support rate for Biden, especially white men, with an approval rate of 80%.

This may reflect the dissatisfaction of the American people with the Trump administration at the time and their recognition of Biden's policies and values on issues such as the economy and health care. Biden has higher support among black voters, ranking second. The support rate among women is 16%, and the support rate among men is 9%. It can be seen that although Biden has shown active stance on issues such as racial justice and social equity, most black voters still have reservations about his policies and promises. Indian voters rank second to last in support for Biden. Asian voters have the lowest support for Biden, with both women and men supporting 4%.

5 Discussion

5.1 Bias

Biden's support rate is not reasonable and the data are biased.

First of all, the researcher collected data through questionnaires, so selection bias may occur during the sample collection process. Second, we use Random sampling when selecting samples from the dataset, which causes the samples to not accurately represent the population. Third, there is an important reason for this result is the social background in 2020. During the epidemic period, many mass incidents occurred in the United States and caused many negative impacts. Most American citizens are disillusioned with the Trump administration. A majority of voters believe the Trump administration's policies have failed to deliver adequate returns. Therefore, they are more inclined to support any candidate who can replace Trump.

5.2 Possible reason

Biden's decisive advantage in the 2020 election.

First, the 2020 presidential election is very special. American citizens are generally dissatisfied with the Trump administration's handling of the COVID-19 pandemic. During the campaign, Biden proposed many response policies, and Biden promised to take more active measures to control the new crown epidemic. Second, Biden has proposed a series of economic recovery plans that will create many job opportunities. Third, Biden supports reforming health insurance to expand health insurance coverage and reduce medical costs.

5.3 Limitations and Future Directions

Our analysis, while comprehensive, presents certain limitations that stem from the nature of the dataset and the modeling approach employed. The survey, as noted by the Cooperative Election Study (CES), follows a methodology that seeks to balance precision and cost. However, large-scale survey data inherently come with caveats that must be considered. Instances of data tampering to align with specific political narratives have been documented in the past, and such risks must be acknowledged (**Rohan?**).

In advancing our research, incorporating a broader set of predictors would likely yield a more nuanced understanding of the American electorate. Variables such as family income, educational attainment, and geographic location could provide valuable insights into voter behavior and preferences. Additionally, including candidates from the Libertarian and Green parties would offer a more holistic view of the electoral landscape, necessitating an adjusted analytical model to accommodate multi-party choices.

Lastly, post-election survey data present a valuable opportunity for further study. Analyzing post-election responses could offer a retrospective understanding of voter sentiment and its alignment with pre-election attitudes. By reapplying our model to post-election data and juxtaposing it with pre-election findings, we could evaluate the predictive efficacy of our analysis and refine our methodology for future electoral studies.

References

Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

———. 2023. *Marginaleffects: Predictions, Comparisons, Slopes, Marginal Means, and Hypothesis Tests*. https://CRAN.R-project.org/package=marginaleffects.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Kuriwaki, Shiro, Will Beasley, and Thomas J. Leeper. 2023. *Dataverse: R Client for Dataverse 4+ Repositories*.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

Schaffner, Brian, Stephen Ansolabehere, and Sam Luks. 2021. "Cooperative Election Study Common Content, 2020." Harvard Dataverse. https://doi.org/10.7910/DVN/E9N6PH.

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.

Wickham, Hadley, Jim Hester, and Romain François. 2021. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.

Wickham, Hadley, Thomas Lin Pedersen, and Dana Seidel. 2023. *Scales: Scale Functions for Visualization*. https://CRAN.R-project.org/package=scales.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.