Who Stays Single? Exploring the Factors Behind Non-Marriage in the United States in 2023*

Low Income and Education Levels Correlate with Higher Propensity to Remain Unmarried

Xizi Sun

December 3, 2024

This study examines the socio-demographic determinants of non-marriage in the United States in 2023 using logistic regression analysis on individual-level data. The findings highlight that individuals with lower income and education levels, particularly below a bachelor's degree, are significantly more likely to remain unmarried. Additionally, gender and race also play important roles, with men and certain racial groups, such as African Americans, showing higher unmarried rates. These results underscore structural inequities and socio-economic barriers shaping marital outcomes, offering insights into broader societal patterns and potential areas for intervention.

Table of contents

*Code and data are available at: https://github.com/LilianS77/US_Marriage.

1 Introduction

Marriage is widely perceived as a fundamental element of both social cohesion and economic stability. It profoundly influences individual lives and the broader social fabric. In recent decades, however, marital patterns in the United States have shifted noticeably, with a growing proportion of individuals choosing to remain unmarried. This trend prompts important questions about the factors underlying these decisions and the broader societal impacts of such changes. Understanding these issues is essential for addressing wider socio-economic inequalities and cultural transformations.

This study investigates the determinants of non-marriage by utilizing a dataset from IPUMS USA (IPUMS 2024), which provides survey data encompassing demographic, social, and economic variables. While previous research has primarily examined individual factors like education or income in isolation, few studies have explored the interplay between multiple socio-demographic factors in shaping marital decisions. This research addresses this shortfall by analyzing these factors in combination.

The estimand of this study is the reason an individual remains unmarried, analyzed in the context of their socio-demographic characteristics, including education level, income, race, gender, and age. This approach enables a deeper understanding of how these attributes collectively influence marital behavior.

The findings, derived through logistic regression analysis, reveal significant patterns. Higher education levels are strongly linked to higher marriage rates, while economic stability, as indicated by higher income, positively correlates with marital status. Marked disparities are evident among racial and gender groups. African Americans, for instance, demonstrate higher rates of non-marriage, whereas males are slightly more likely to remain unmarried than females. These findings underscore persistent structural inequities and cultural complexities influencing marriage in contemporary American society. Recognizing these inequities equips policymakers and social institutions to design targeted interventions that address inequality, foster opportunities for stable relationships, and adapt to evolving family dynamics in modern society.

The structure of this paper is as follows: Section 2 explains the dataset, variable selection, and measurement strategies employed to study marital status. Section 3 discusses the logistic regression methodology, including model rationale, prior specifications, and validation processes. Section 4 presents the key findings with supporting visualizations, emphasizing the roles of education, income, race, gender, and age in predicting non-marriage. Finally, Section 5 delves into the implications of the findings, acknowledges study limitations, and outlines potential areas for future research.

2 Data

2.1 Data Tool

The dataset was analyzed using R (R Core Team 2023) and utilized several R packages for data manipulation, visualization, and analysis, including ggplot2 (Wickham 2016) for creating elegant graphics, dplyr (Wickham et al. 2023) for data wrangling, and here (Müller and Bryan 2023) for simplifying file management. The data was processed efficiently using Apache Arrow (Richardson, Dunnington, and Developers 2023), and visualizations were further enhanced with the scales package (Wickham, Seidel, et al. 2023). logistic regression modeling was implemented using rstanarm (Cepeda, Gabry, et al. 2023), while statistical results were tidied with broom (Robinson 2023) and broom.mixed (Bolker et al. 2023). Tables were constructed using kableExtra (Zhu 2023) for a polished presentation. Reproducibility was ensured with knitr (Xie 2014). The bayesplot package (Gabry et al. 2023) was utilized for visualizing Bayesian model diagnostics and posterior predictive checks. Data was extracted from IPUMS USA (IPUMS 2024), and guidance on storytelling with data was drawn from Telling Stories with Data (Alexander 2023).

2.2 Measurement

For this study, data was specifically drawn from the American Community Survey (ACS) (Bureau 2024), a subset of IPUMS. The ACS conducts ongoing data collection throughout the year, selecting approximately 250,000 addresses each month to ensure the population's most current representation. The survey employs several collection methods to maximize participation. Initially, respondents receive a request to complete the survey online or return a paper questionnaire via mail. If no response is received, follow-up contact is made through telephone, utilizing a computer-assisted interview system (CATI). For those who remain unresponsive, a portion—approximately one-third—is selected for an in-person, computer-assisted personal interview (CAPI). In addition, the survey covers a wide range of social, economic and demographic characteristics, such as the variables of marital status, age and education analyzed in this study.

2.3 Outcome Variables

The primary outcome variable for this study is **Marital Status**, which categorizes individuals based on their marital state. The proportion of marital status categories is displayed in Figure 1. This variable allows for a comparison between individuals who have never married (`Not_Married`) and those who have (`Married`). For this study, **Not_Married**: Includes individuals who have never been married. **Married**: Includes individuals who are married, as well as those who are divorced, widowed, or separated.

2.4 Predictor Variables

The distribution of predictor variables is displayed in Figure 2. These variables capture demographic, socioeconomic, and personal characteristics, providing a comprehensive framework for analyzing factors associated with marital status. Below are the key predictor variables:

1. **Age**: A continuous variable representing the respondent's age in years.
2. **Gender**: A categorical variable indicating whether the respondent is male or female. Gender differences often play a role in marital patterns.
3. **Race**: A categorical variable categorized into White, Black, Asian, American Indian, and Other racial groups. This variable examines potential racial disparities in marital behavior.
4. **Education Level**: An ordinal variable indicating the highest level of education attained by the respondent. It is grouped into five categories: Below High School, High School, Some College, Bachelor's Degree, and Above Bachelor.
5. **Income**: A continuous variable measuring the respondent's annual income in dollars. Income reflects economic resources and may be associated with marital stability and decisions.

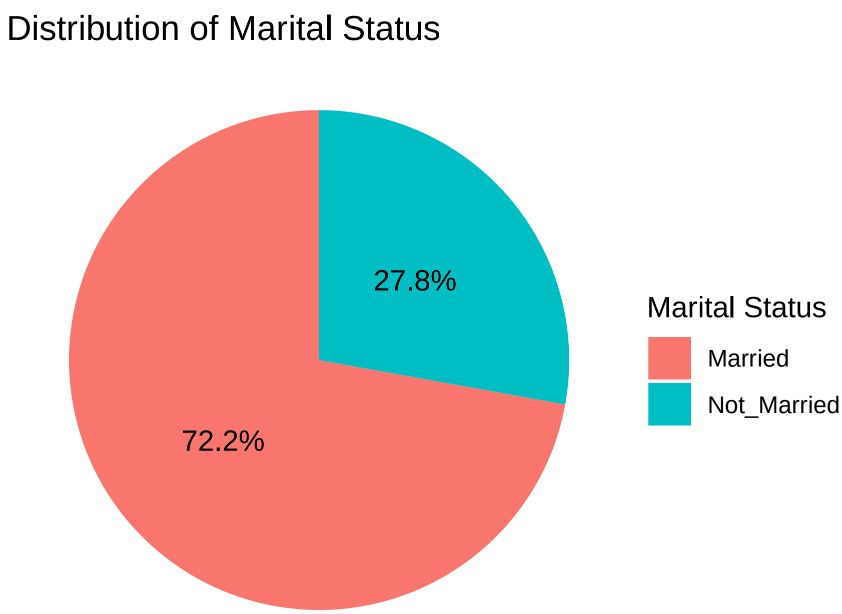


Figure 1: Proportion of Marital Status Categories

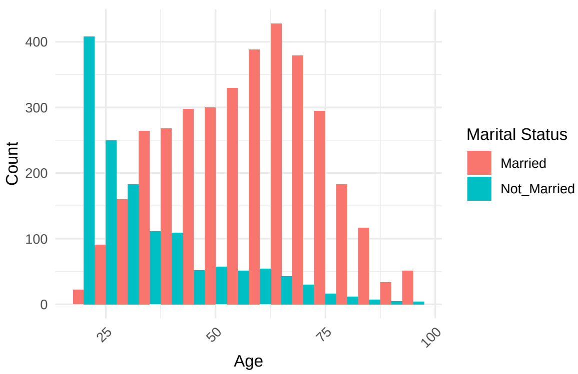
The unmarried results of the predictor variables are displayed in the Section 4.

2.5 Data Selection

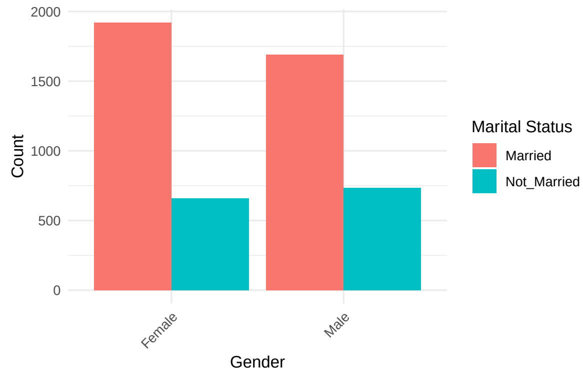
To examine the phenomenon of non-marriage in the United States, the IPUMS USA dataset was chosen as the primary data source, rather than IPUMS International. Although IPUMS International includes harmonized census data from 104 countries and encompasses over one billion individual records, its broad, global scope renders it less aligned with the objectives of this study. This research focuses specifically on societal patterns within the U.S. population. IPUMS USA, with its data derived from the American Community Surveys (ACS) and federal censuses, provides the precision and relevance needed to investigate non-marriage trends in the United States.

3 Model

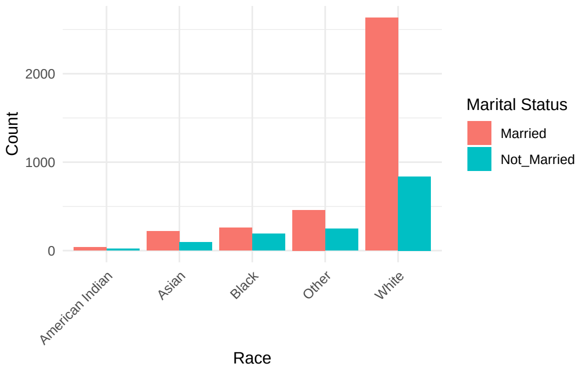
For this analysis, I apply a Logistic Regression Model to assess the likelihood of an individual not being married (outcome variable) based on several demographic and socioeconomic predictors (predictor variables). This model was chosen due to the binary nature of the outcome



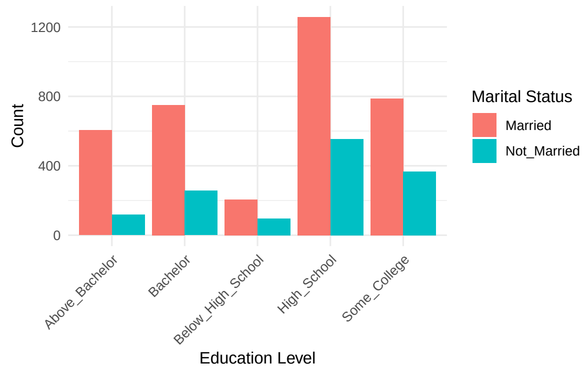
(a) Age Distribution by Marital Status



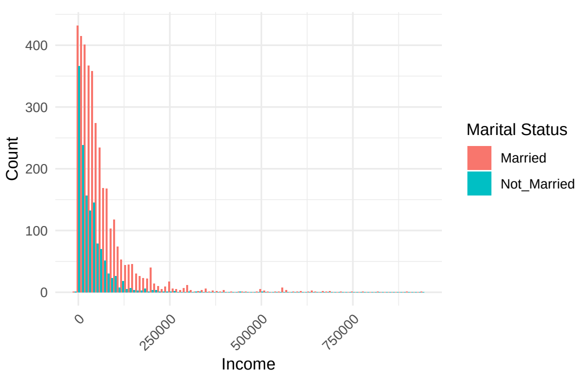
(b) Gender Distribution by Marital Status



(c) Race Distribution by Marital Status



(d) Education Level Distribution by Marital Status



(e) Income Distribution by Marital Status

Figure 2: Counts for Demographic Variables by Marital Status

variable, which distinguishes between individuals who are "Not Married" versus those who are married (including divorced, widowed, or separated).

3.1 Model Setup

3.1.1 Objective

The primary objective of the model is to analyze and predict the factors associated with individuals' marital status, focusing on identifying key predictors for individuals who have never been married (`Not_Married`).

The logistic regression model used in this study is:

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \beta_0 + \beta_1 \cdot \text{age} + \beta_2 \cdot \text{gender} + \beta_3 \cdot \text{Race} + \beta_4 \cdot \text{Income} + \beta_5 \cdot \text{education\_level}$$

The priors for the coefficients are defined as:

$$\begin{aligned}
\beta_0 &\sim \text{Normal}(0, 2.5),\\
\beta_1 &\sim \text{Normal}(0, 2.5),\\
\beta_2 &\sim \text{Normal}(0, 2.5),\\
\beta_3 &\sim \text{Normal}(0, 2.5),\\
\beta_4 &\sim \text{Normal}(0, 2.5),\\
\beta_5 &\sim \text{Normal}(0, 2.5).
\end{aligned}$$

Where:

- $\hat{p}$: Represents the probability that an individual is classified as `Not_Married`.
- β_0: The intercept term, representing the log-odds of being `Not_Married` when all predictors are zero.
- β_1: The coefficient for the individual's age.
- β_2: The coefficient for gender.
- β_3: The coefficient for the race category.
- β_4: The coefficient for annual income.
- β_5: The coefficient for the education level.

3.1.2 Priors Explanation

In this model, normal priors with a mean of 0 and a standard deviation of 2.5 are assigned to all coefficients and the intercept. This choice reflects neutrality, implying no prior expectation of direction or magnitude for the coefficients.

The standard deviation of 2.5 reflects moderate uncertainty in prior beliefs, allowing the model to incorporate reasonable variability in predictor effects while avoiding overly restrictive priors.

3.2 Model Justification

The logistic regression model was chosen for this study due to its suitability in predicting binary outcomes. Specifically, it was employed to classify individuals as either "Not Married" or otherwise. This method is widely used in social science research to analyze the relationship between a binary dependent variable and several independent variables. It is therefore well-suited for examining patterns in marital status.

The decisions regarding model design were guided by the dataset's structure and the variables under analysis. For instance, age was treated as a continuous variable to preserve its detail and avoid arbitrary groupings. Gender was modeled as a categorical variable with two levels—male and female—to reflect its binary nature in the data. Each racial category was included as a distinct level to capture the unique effects associated with different racial groups. Similarly, education and income were retained in their granular forms to maximize the use of the dataset's available information.

While the method depends on certain assumptions, such as linearity in the log-odds and independence of observations, it allows for the inclusion of interaction terms and accommodates a wide range of predictors. This balance of clarity, flexibility, and alignment with the study's objectives makes logistic regression an effective choice for this research.

Posterior checks for the model and MCMC convergence check can be found in Section A.2.

4 Results

This study integrates multiple visualizations to provide a comprehensive understanding of the demographic, economic, and social dynamics associated with marital status. Below, we summarize the key findings derived from each visualization.

Figure 3 and Table 1 highlight the coefficients for RaceWhite, RaceOther, and RaceAsian are negative, indicating that individuals from these racial groups are less likely to remain unmarried compared to others. The coefficient for RaceBlack is positive, and its confidence interval does not include 0, suggesting that RaceBlack group are more likely to remain unmarried compared to other racial groups. Individuals with lower levels of education (e.g., below high school) have a higher likelihood of remaining unmarried. The coefficient for males is positive, indicating that men are slightly more likely to remain unmarried compared to female.

Table 1: Summary of the model

term	estimate	std.error	conf.low	conf.high
(Intercept)	3.26	0.37	2.65	3.87
age	-0.08	0.00	-0.09	-0.08
genderMale	0.37	0.08	0.24	0.51
RaceAsian	-0.67	0.37	-1.26	-0.07
RaceBlack	0.20	0.34	-0.38	0.77
RaceOther	-0.67	0.35	-1.22	-0.10
RaceWhite	-0.67	0.33	-1.21	-0.11
Income	0.00	0.00	0.00	0.00
education_levelBachelor	0.30	0.14	0.07	0.54
education_levelBelow_High_School	0.48	0.21	0.15	0.83
education_levelHigh_School	0.21	0.14	-0.02	0.46
education_levelSome_College	0.18	0.15	-0.06	0.43

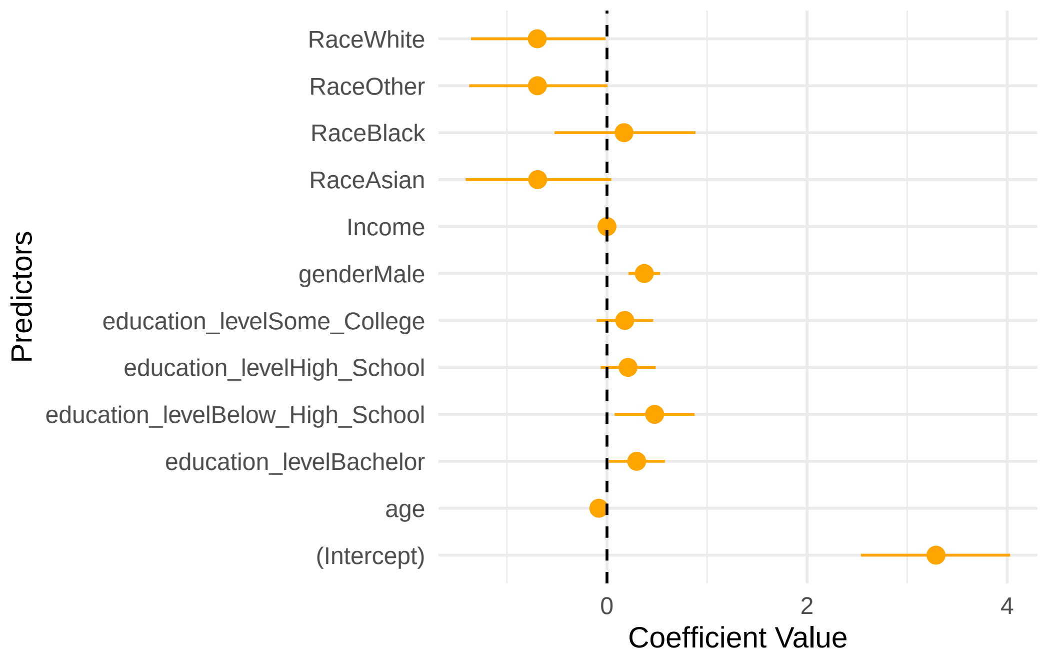


Figure 3: Coefficients of The Model

Figure 4 depicts the income distribution of unmarried individuals, revealing that the majority of unmarried individuals fall into lower income brackets. The distribution is heavily skewed, with a significant concentration of individuals earning below $50,000 annually. The density sharply declines for higher income levels, reflecting the economic constraints that may influence marital status. A small proportion of high-income individuals remain unmarried, as shown by the long right tail of the distribution.

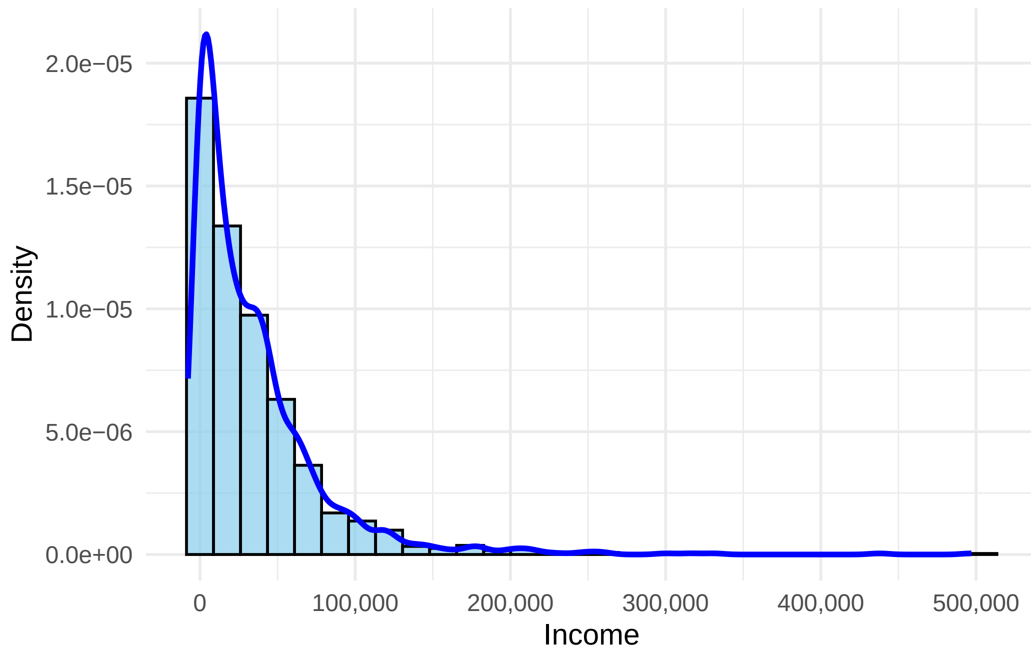


Figure 4: Income Distribution for Unmarried Individuals

Figure 5 shows the age distribution for unmarried individuals, which is dominated by younger age groups, particularly those in their 20s and early 30s. The density decreases substantially as age increases, illustrating that as individuals age, the likelihood of remaining unmarried diminishes.

The proportions of marital status by education level are presented in Figure 6. Individuals with higher education levels show significantly lower proportions of unmarried individuals, with 25% for those holding a bachelor's degree and 16% for those with education above a bachelor's degree. In contrast, individuals with lower educational attainment exhibit higher proportions of unmarried individuals, with 32%, 31%, and 32% for those with below high school, high school, and some college education, respectively.

Figure 7 and Figure 8 highlight the proportions of marital status by race and gender. Race disparities are evident, with "American Indian" and "Black" populations showing higher proportions of unmarried individuals, at 37% and 43%, respectively. In contrast, "White" and "Asian" groups exhibit the lowest proportions of unmarried individuals, at 24% and 30%,

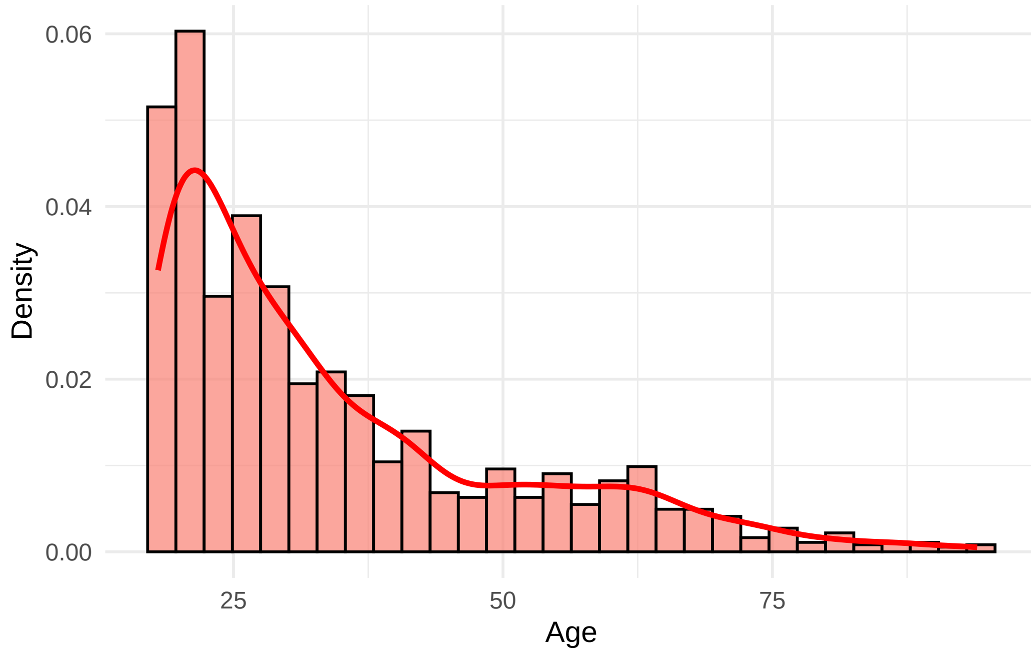


Figure 5: Age Distribution for Unmarried Individua

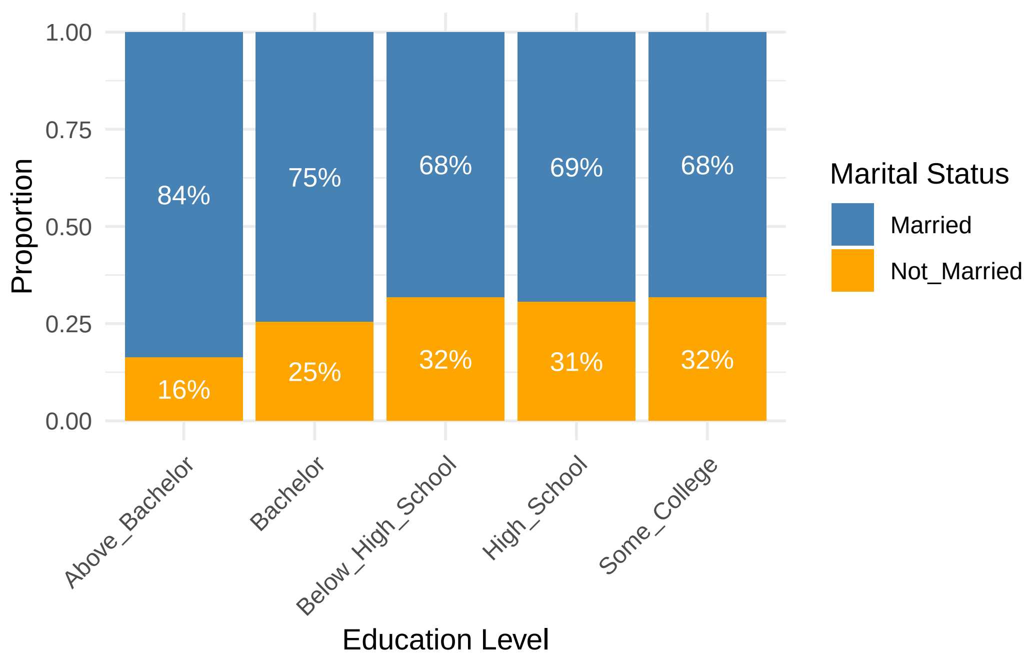


Figure 6: Proportions of Marital Status by Education Level

respectively. Gender analysis reveals that males are more likely to remain unmarried than females, with 30% of males unmarried compared to 26% of females.

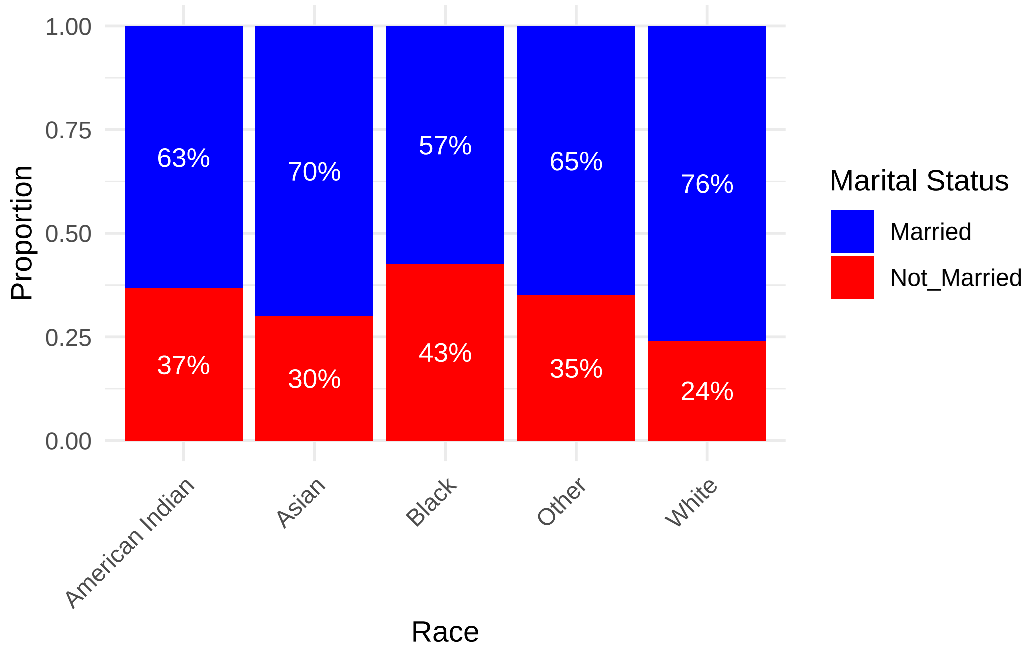


Figure 7: Proportions of Marital Status by Race

5 Discussion

5.1 Determinants of Non-Marriage and Societal Implications

This study investigates the socio-demographic factors influencing marital status, focusing on the probability of remaining unmarried in the United States. Through logistic regression analysis applied to individual-level microdata, significant predictors such as education, income, race, gender, and age were identified.

Higher education levels were strongly associated with lower non-marriage rates. For instance, Figure 6 shows that only 16% of individuals with education above a bachelor's degree were unmarried, compared to 32% for those with "Some College" education. Logistic regression coefficients further underscore this trend, with individuals holding higher degrees being substantially less likely to remain unmarried (e.g., the coefficient for "Below High School" = 0.48). Education thus acts as a stabilizing factor, equipping individuals with economic resources and social networks conducive to marriage.

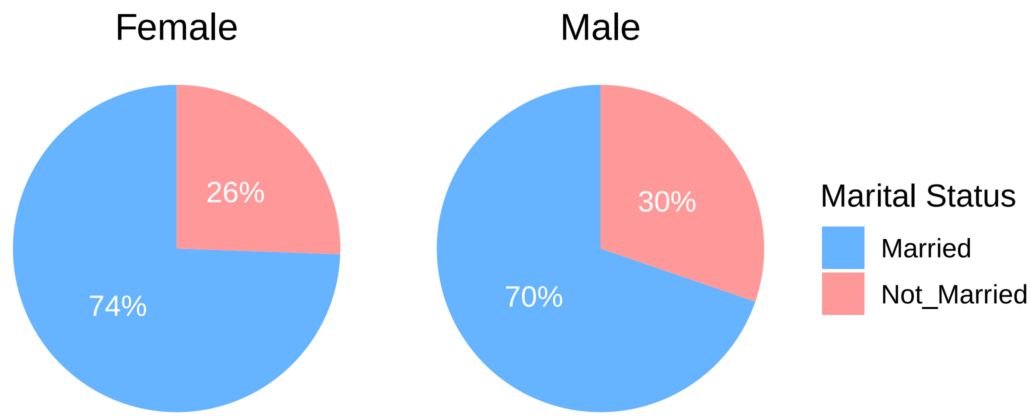


Figure 8: Marital Status Proportions by Gender

Similarly, income plays a complementary but less pronounced role. Figure 4 illustrates that unmarried individuals tend to cluster in lower income brackets, with the majority earning less than $40,000 annually. This highlights the persistent importance of financial stability in marital decisions, although its impact appears secondary to education.

The results also emphasize significant disparities across racial and gender lines. African Americans exhibited the highest non-marriage rates (43%), as shown in Figure 7, significantly surpassing those of Asian Americans (30%) and White individuals (24%). These patterns likely reflect structural inequalities, such as higher unemployment rates and barriers to wealth accumulation among African Americans. Gender differences were also evident, with men showing a slightly higher likelihood of remaining unmarried compared to women (30% versus 26%, respectively, Figure 8). This aligns with cultural norms where women are more closely tied to family roles and experience greater societal pressure to marry.

These findings reinforce how individual socio-demographic traits intersect with societal structures to shape marital outcomes. They also highlight systemic inequities that warrant targeted interventions.

5.2 Insights into Education, Income, and Marriage Patterns

Education emerged as the most significant predictor in this analysis. Higher education levels correlate strongly with lower non-marriage rates, as depicted in Figure 6 and Figure 3. These findings align with theories suggesting that education enhances economic stability, expands social capital, and provides greater agency in life choices, all of which promote marriage. However, as Figure 4 indicates, financial stability remains an essential but secondary factor in marital decisions.

The interplay between education and income underscores broader socio-economic disparities. For example, individuals with lower education levels not only face limited economic opportunities but are also more likely to remain unmarried, perpetuating cycles of inequality in family formation.

5.3 Gender and Race in Marital Trends

The analysis revealed gender and racial disparities. As shown in Figure 8, men exhibit higher non-marriage rates compared to women, likely reflecting cultural expectations and differing societal pressures. Racial disparities, however, are even more striking. African Americans, for instance, show the highest rates of non-marriage, suggesting systemic inequities such as income disparities and historical barriers to wealth accumulation. By contrast, Asian Americans, with the lowest non-marriage rates, may benefit from cultural norms emphasizing family cohesion and marriage stability.

These patterns illustrate how cultural norms and structural inequities converge to shape marital outcomes. Addressing these disparities requires policies that promote economic equity and social inclusivity.

5.4 Weaknesses and Next Steps

While this study provides valuable insights into the determinants of non-marriage, it is not without limitations. The cross-sectional nature of the data restricts causal inferences, making it difficult to ascertain how variables such as education and income evolve over time to influence marital decisions. Longitudinal data would allow for a deeper exploration of these dynamics.

Additionally, the dataset lacks qualitative dimensions, such as cultural attitudes or personal preferences, which are crucial for understanding the broader context of marital decisions. For instance, while financial stability and education are strong predictors, societal expectations and individual life goals may play equally important roles that remain unmeasured.

Future research should consider integrating mixed-method approaches to address these gaps. Furthermore, expanding the analysis to include international datasets could offer comparative insights, shedding light on how cultural and economic systems influence marriage patterns globally.

5.5 Conclusion

This study contributes to the understanding of non-marriage by emphasizing the roles of education, income, race, gender, and age. Education emerged as the most significant predictor, with higher levels strongly associated with lower non-marriage rates. Income, while secondary, remains a critical factor in promoting marital stability. Racial and gender disparities further highlight systemic inequities that shape marital trends.

By addressing these structural barriers, policymakers and institutions can promote equity in family formation. Future research should build on these findings, leveraging longitudinal and qualitative approaches to provide a more comprehensive understanding of marriage and its evolving role in contemporary society.

A Appendix

A.1 Data Details

A.1.1 Cleaned Data

The Table 2 shows the analysis data after data cleaning.

Table 2: Sample of cleaned data

Marital Status	Age	Gender	Race	Income	Education Level
Not_Married	73	Male	White	33900	Above_Bachelor
Married	43	Female	Asian	40000	High_School
Married	80	Female	White	13000	High_School
Married	66	Male	White	48000	Some_College
Married	52	Female	White	38100	Bachelor
Married	45	Female	American Indian	65000	High_School

A.2 Model Details

A.2.1 Posterior predictive check

Figure 9 provides an illustration of the posterior predictive check, showing the alignment between the observed outcome variable and the simulations derived from the posterior distribution. In a similar vein, Figure 10 depicts a comparison between the prior and posterior distributions, highlighting the degree to which the data informs and updates our estimates. Together, these figures offer strong evidence that the model achieves a reliable and accurate fit to the observed data.

A.2.2 Diagnostics

Figure 11 and Figure 12 provide insights into the MCMC diagnostics for the logistic regression model. Figure 11 show stable and well-mixed chains for all parameters, indicating convergence. fig-rhat shows $\hat{R}$ diagnostics are all close to 1, further confirming that the MCMC sampling has successfully converged to the posterior distribution.

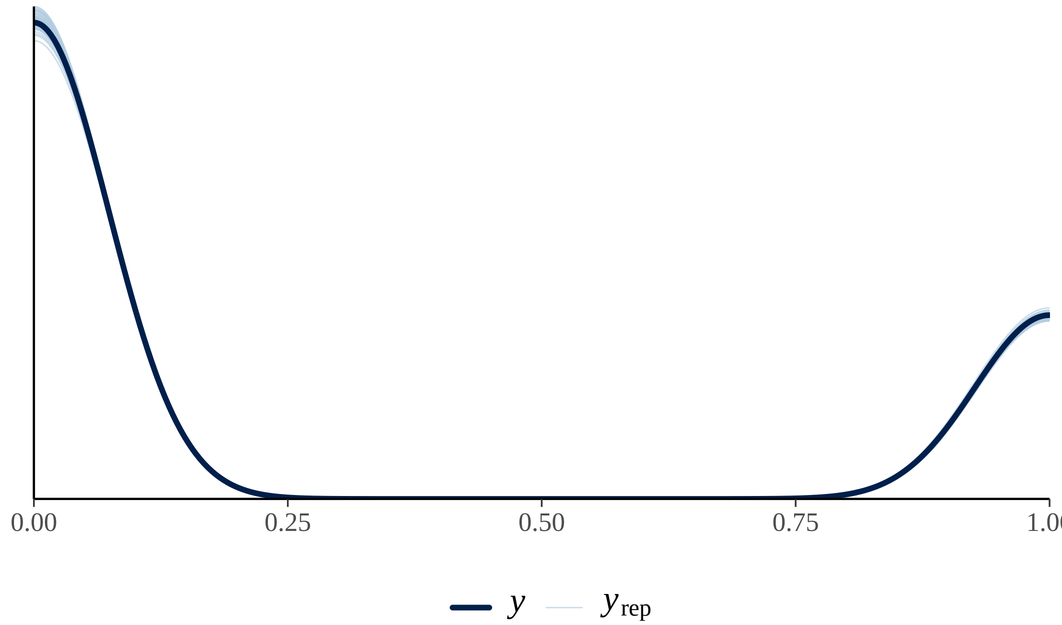


Figure 9: Posterior distribution for logistic regression model

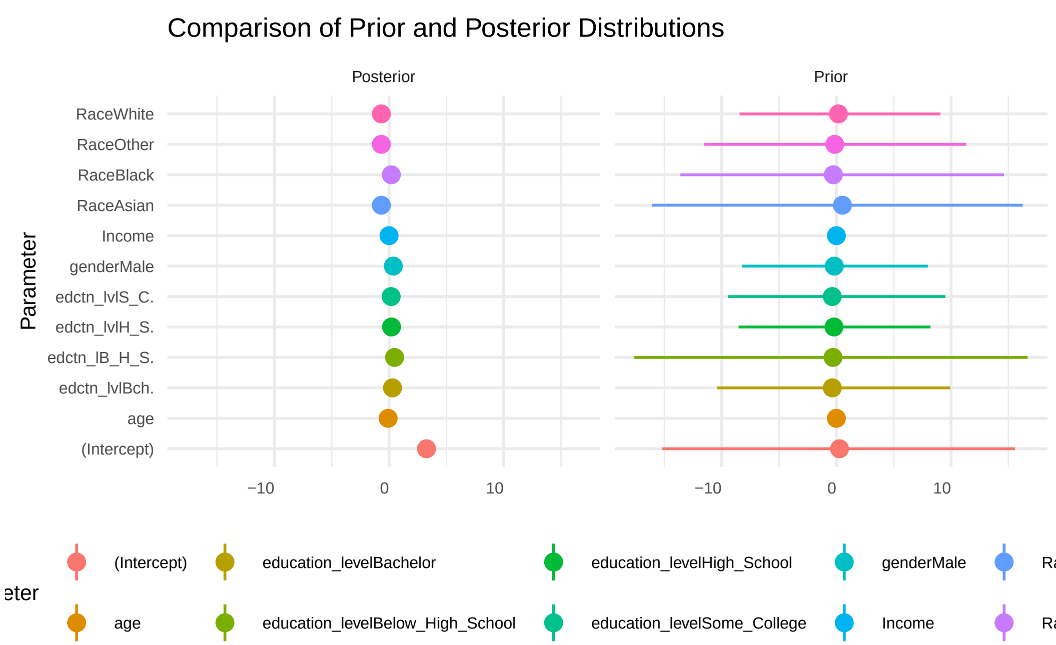
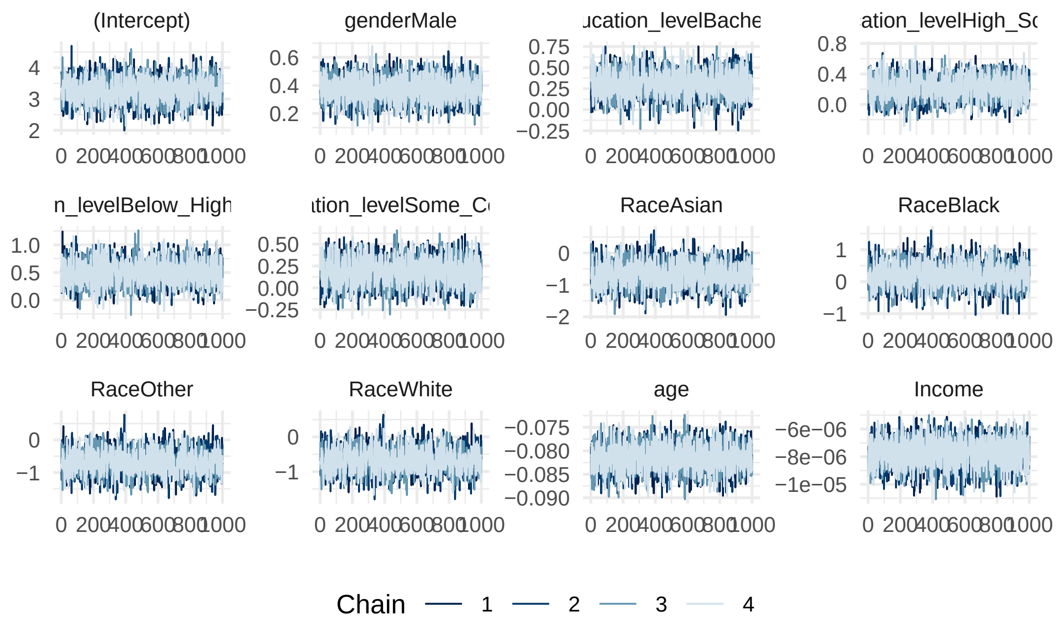


Figure 10: Comparing prior distribution with posterior distribution



(a) Intercept

Figure 11: MCMC Convergence Check: Trace Plots for Key Parameters

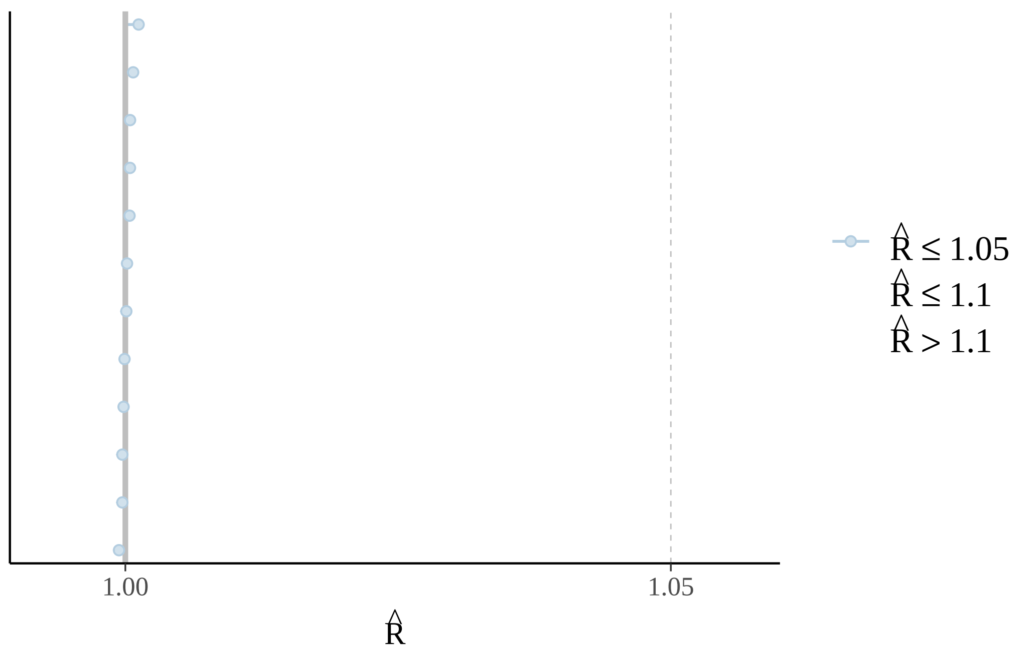


Figure 12: Rhat plot

A.3 Surveys, sampling, and observational data

A.3.1 Idealized Methodology

A.3.1.1 Overview

The objective of this survey is to explore the socio-demographic determinants of non-marriage in the United States. The study aims to understand factors influencing decisions to remain unmarried, including education, income, race, age, and cultural attitudes. A budget of $50,000 is allocated to collect diverse and representative data from individuals aged 18 and above. The methodology is designed to maximize the accuracy.

A.3.1.2 Sampling Approach

A stratified sampling approach is adopted to ensure the representation of key demographic and socio-economic groups. Stratification variables include:

- Age: 18–29, 30–44, 45–64, 65+
- Gender: Male, Female, Non-binary
- Race/Ethnicity: White, Black, Hispanic, Asian, American Indian, Other
- Income: <$20,000, $20,000–$59,999, $60,000–$99,999, >$100,000
- Education: Under High School, High School, Some College, Bachelor's Degree, Above Bachelor's Degree *Geographic Region: Northeast, Midwest, South, West The target sample size is 5,000 respondents, yielding a margin of error of ±2.5% at a 95% confidence level.

A.3.1.3 Recruitment Strategy

- Online Surveys: Partner with established survey platforms like Prolific or Qualtrics, which offer access to diverse and validated panels.
- Community Engagement: Partnerships with community organizations for in-person recruitment.
- Mail Invitations: Surveys sent by mail, accompanied by pre-paid return envelopes or online access codes for respondents to complete the survey.
- Telephone Interviews: Random-digit dialing (RDD) to reach both landlines and mobile phones, ensuring coverage of participants without internet access.

A.3.1.4 Data Collection

Data collection will utilize platforms like Google Forms, with telephone and mail surveys structured to mirror the design and flow of the Google Form questionnaire. To reduce the likelihood of participants abandoning the survey, it has been tailored to take approximately five to ten minutes to complete.

A.3.1.5 Data Validation and Quality Control

- Pre-Survey Validation: Pilot testing with diverse focus groups to refine question phrasing and structure.
- Ongoing Quality Checks: Automated checks for inconsistent or incomplete responses during the survey.
- Post-Survey Weighting: Application of population weights to correct for sample imbalances across demographic strata.

A.3.1.6 Budget Allocation

- Survey Recruitment: $30,000
- Incentives: $10,000
- Data Cleaning and Analysis: $5,000
- Miscellaneous (software, outreach): $5,000

A.3.2 Idealized Survey Questions

Welcome Message
Welcome! Thank you for participating in this survey on marital status and the factors influencing people's decision to remain unmarried. Your responses are anonymous and will only be used for research purposes. This survey will take approximately 5–10 minutes to complete.

Screening Questions
Q1: Are you currently residing in the United States?
- Yes
- No *(If "No," terminate the survey with a thank-you message.)*

Q2: What is your age?
- Under 18 *(Terminate the survey with a thank-you message.)*
- 18 or older

Consent to Participate
Before proceeding, please read the following:
- Your participation in this survey is voluntary.
- You may skip any question or withdraw at any time.
- All responses are anonymous and will be aggregated for research purposes.

Do you agree to participate in this survey?
- Yes, I agree to participate.
- No, I do not agree to participate. *(If "No," terminate the survey with a thank-you message.)*

Survey Questions

Q3: What is your gender?
- Male
- Female
- Non-binary
- Prefer not to say

Q4: What is your race/ethnicity?
- White
- Black or African American
- Asian
- Hispanic or Latino
- Native American or Alaska Native
- Other

Q5: What is your highest level of education?
- Less than high school
- High school diploma or equivalent
- Some college
- Bachelor's degree
- Postgraduate degree

Q6: What is your total annual income?
- Less than $20,000
- $20,000–$49,999
- $50,000–$79,999
- $80,000–$99,999
- $100,000 or more

Q7: What is your age group?
- 18–29
- 30–44
- 45–59
- 60 or older

Q8: What is your current marital status?
- Never married
- Married
- Divorced
- Widowed
- Separated

Q9: If never married, have you ever been in a serious relationship (e.g., cohabitation, long-term dating)?
- Yes
- No

Q10: How important do you consider marriage in your life?
- Very important
- Somewhat important
- Not very important
- Not at all important

Q11: What are the main reasons for not being married? *(Select all that apply)*
- Financial concerns
- Lack of a suitable partner
- Personal choice
- Career priorities
- Family or cultural pressures
- Other (please specify)

Q12: Do you feel societal pressure to get married?
- Strongly agree
- Agree
- Neutral
- Disagree
- Strongly disagree

Q13: Do you plan to get married in the future?
- Yes
- No
- Unsure

Q14: If unsure or no, what factors might influence your decision?
- Improved financial stability

- Finding the right partner
- Change in personal values
- Other (please specify)

Closing Message
Thank you for completing this survey! Your responses have been recorded. If you have any questions about this survey or its purpose, please contact xizi.sun@mail.utoronto.ca. Have a great day!

References

Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. https://tellingstorieswithdata.com/.

Bolker, Ben et al. 2023. *broom.mixed: Tidying Methods for Mixed Models*. https://bbolker.github.io/broom.mixed/.

Bureau, U. S. Census. 2024. *American Community Survey (ACS): Comprehensive Socioeconomic Data Collection*. https://www.census.gov/programs-surveys/acs.

Cepeda, Gabriel A., Jonah Gabry, et al. 2023. *rstanarm: Bayesian Applied Regression Modeling via Stan*. https://mc-stan.org/rstanarm/.

Gabry, Jonah et al. 2023. *bayesplot: Plotting for Bayesian Models*. https://mc-stan.org/bayesplot/.

IPUMS. 2024. "IPUMS USA." https://usa.ipums.org/usa/#:~:text=IPUMS%20USA%20collects,%20preserves%20and%20harmonizes.

Müller, Kirill, and Jennifer Bryan. 2023. *here: A Simpler Way to Find Your Files*. https://here.r-lib.org/.

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, Dewey Dunnington, and Apache Arrow Developers. 2023. *arrow: Integration to Apache Arrow*. https://arrow.apache.org/.

Robinson, David. 2023. *broom: Convert Statistical Objects into Tidy Tibbles*. https://broom.tidyverse.org/.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. https://ggplot2.tidyverse.org/.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2023. *dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org/.

Wickham, Hadley, Dana Seidel, et al. 2023. *scales: Scale Functions for Visualization*. https://scales.r-lib.org/.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.

Zhu, Hao. 2023. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://haozhu233.github.io/kableExtra/.