

# Replication of ‘Sexism and the Far-Right Vote: The Individual Dynamics of Gender Backlash’ (Anduiza & Rico, 2022)

Reeka Estacio And Belynda Herrera

2025-02-11

For the final project, we (**Reeka Estacio** and **Belynda Herrera**) intend to reproduce the binary logit model for the year 2019, modeling the intention to vote for the Vox party. The regression table of this model is located in the first column of “Table 1. Predictors of Intention to Vote for Vox in 2019 and 2020” of the original paper.

In the paper, the model is described as:

$$vox_{it} = sexism_{it} + otherattitudes_{it} + controls_{it}$$

## Load data

```
# Read .dta file
dat <- read_dta("cleaned.dta")

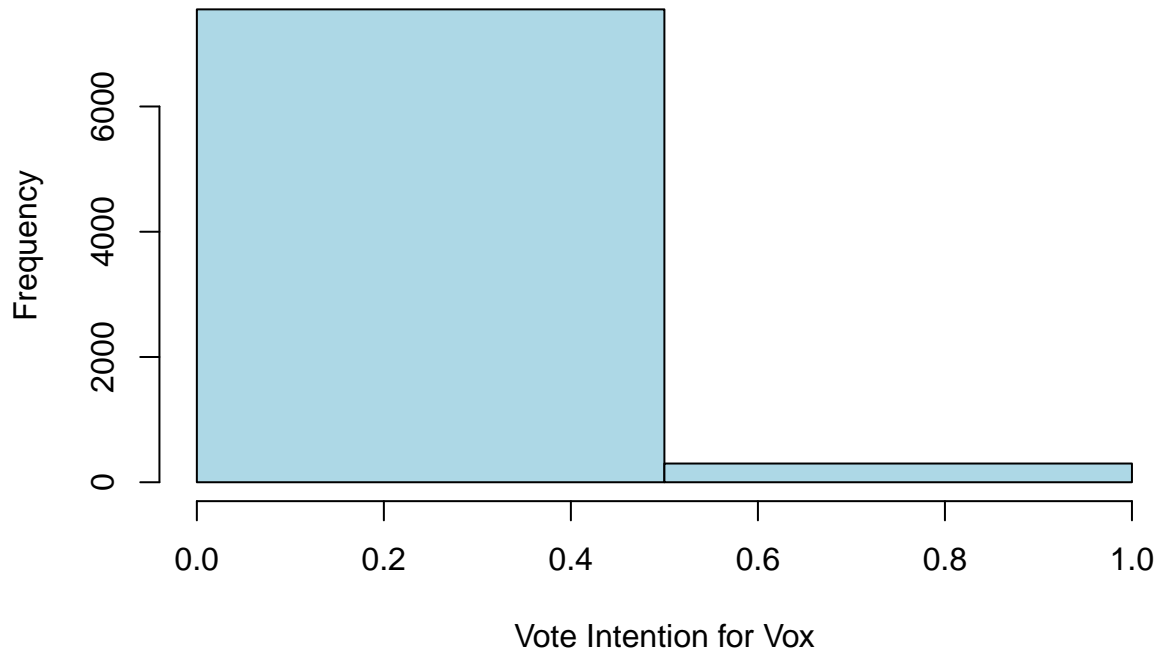
# Save as CSV
write.csv(dat, "cleaned.csv", row.names = FALSE)
```

## Histogram of the dependent variable

The dependent variable of this model is `vim_vox`. This is a binary variable indicating a person’s intention to vote for the Vox party.

```
hist(dat$vim_vox, breaks = 2, col = "lightblue", main = "Histogram of Intention to Vote for Vox",
      xlab = "Vote Intention for Vox", ylab = "Frequency")
```

## Histogram of Intention to Vote for Vox



## Correlation matrix of the DV and IVs of the Model 1 (2019)

The correlation matrix produced below includes the dependent variable (`vim_vox`) and the exact independent variables used to produce the original model corresponding to the year 2019.

```
library(corrplot)
```

```
## corrplot 0.95 loaded
```

```
# Filter data for the year 2019
```

```
data_2019 <- dat %>% filter(year == 2019)
```

```
# Select relevant numeric variables from your model
```

```
subset_data <- data_2019 %>%  
  select(female, age, edu3_2, edu3_3, dhincome_all, livingpartner, intpol,  
         l2vim_vox, l2authoritarian, l2ideol, l2nativism, l2orgterr, l2pop6amz,  
         l2swim_msex, lsvim_vox, lsauthoritarian, lsideol, lsnativism,  
         lsorgterr, lspop6amz, lsswim_msex)
```

```
# Remove missing values
```

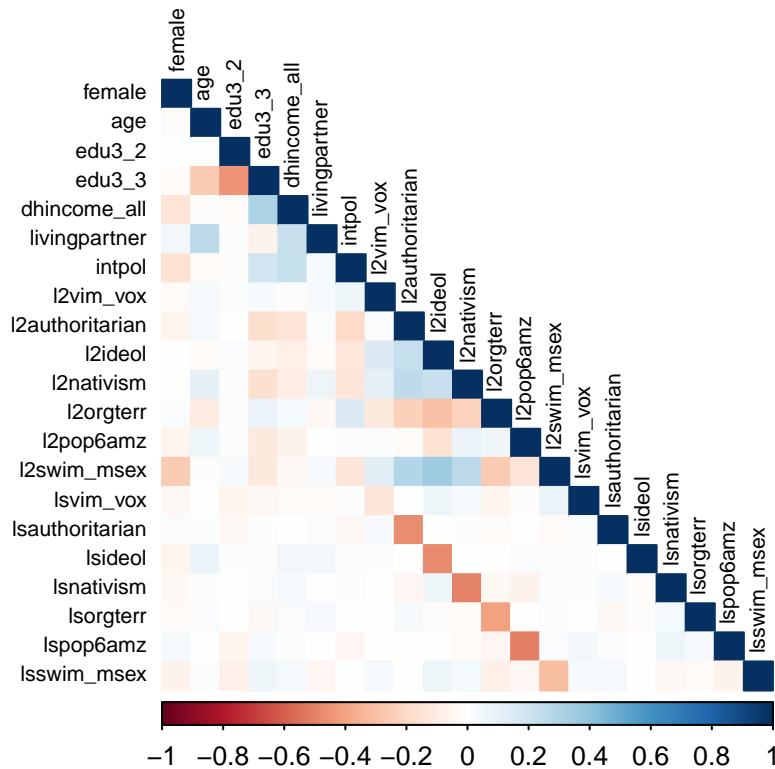
```
clean_data <- subset_data %>%  
  na.omit()
```

```
# Compute correlation matrix
```

```
cor_matrix <- cor(clean_data, use = "complete.obs")
```

```
# Plot correlation matrix
```

```
corrplot(cor_matrix, method = "color", type = "lower",
         tl.cex = 0.7, tl.col = "black", number.cex = 0.7)
```



## Visual Representation of Missing Data

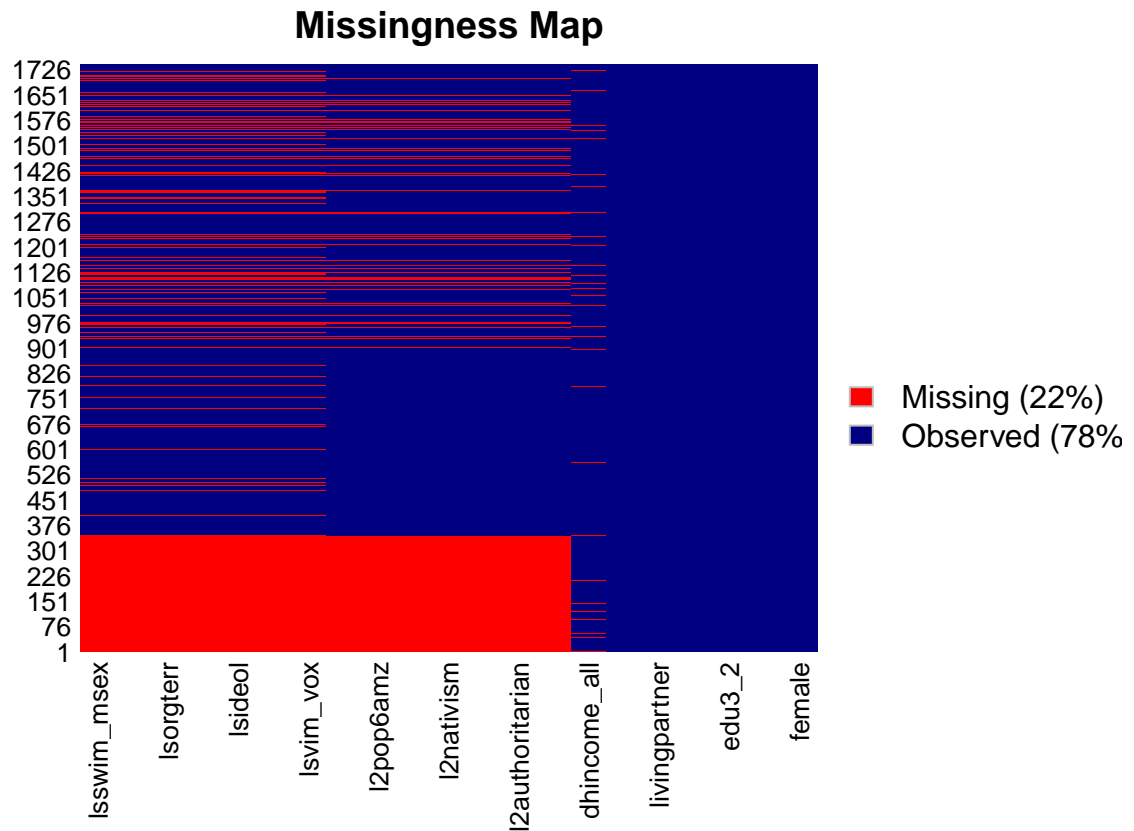
The missingness map highlights missing values in the variables used in this model.

```
# Load the Amelia package
library(Amelia)
```

```
## Loading required package: Rcpp
## ##
## ## Amelia II: Multiple Imputation
## ## (Version 1.8.3, built: 2024-11-07)
## ## Copyright (C) 2005-2025 James Honaker, Gary King and Matthew Blackwell
## ## Refer to http://gking.harvard.edu/amelia/ for more information
## ##
```

```
# Generate a missingness map
missmap(subset_data, col = c("red", "navyblue"), legend = TRUE)
```

```
## Warning: Unknown or uninitialised column: `arguments`.
## Unknown or uninitialised column: `arguments`.
## Warning: Unknown or uninitialised column: `imputations`.
```

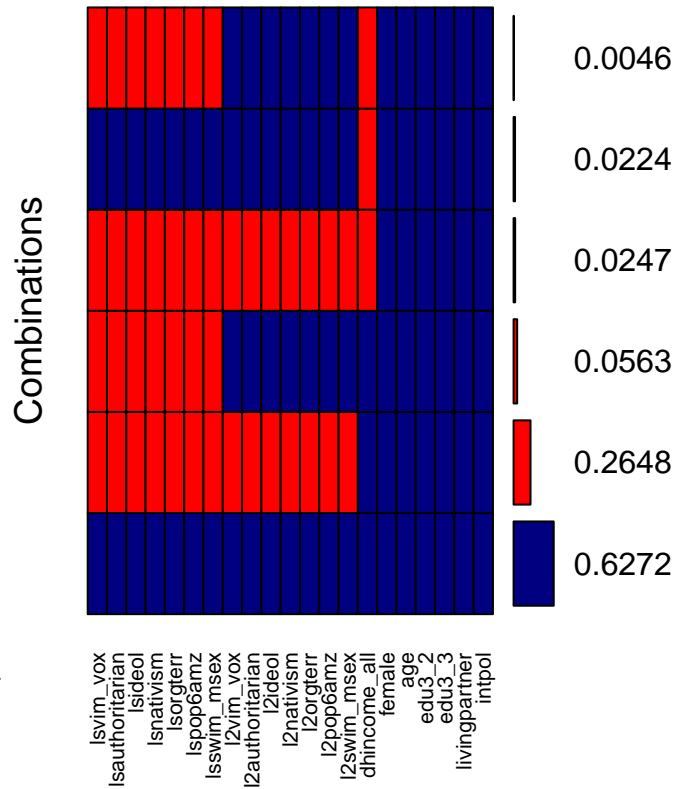
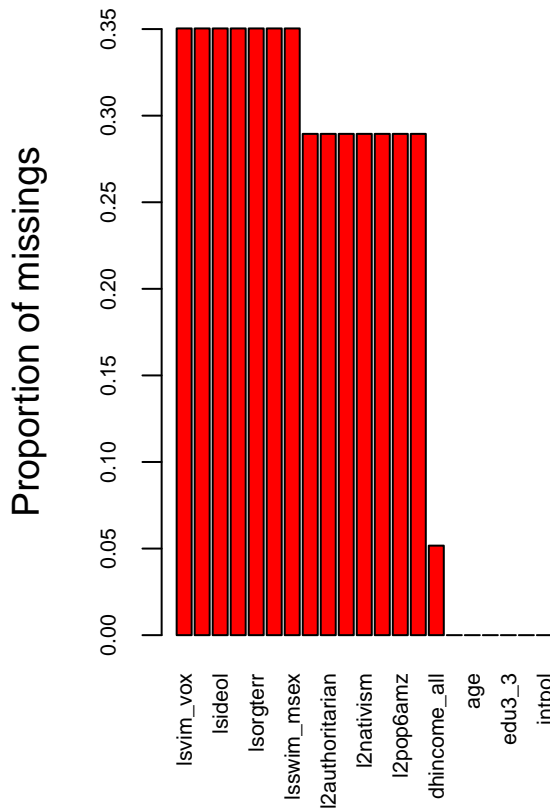


The bar plot shows the proportion of missing values per variable (missing data in red, observed data in blue). The aggregation plot shows frequency of missingness across different combinations of the variables.

```
library(VIM)

## Loading required package: colorspace
## Loading required package: grid
## VIM is ready to use.
## Suggestions and bug-reports can be submitted at: https://github.com/statistikat/VIM/issues
##
## Attaching package: 'VIM'
## The following object is masked from 'package:datasets':
##
##     sleep

# Bar plot of missing data proportions
aggr(subset_data, col = c("navyblue", "red"), numbers = TRUE, sortVars = TRUE,
     labels = names(subset_data), cex.axis = 0.7, gap = 2,
     ylab = c("Proportion of missings", "Combinations"))
```



```
##
## Variables sorted by number of missings:
## Variable Count
## lsvim_vox 0.35037335
## l2authoritarian 0.35037335
## lsideol 0.35037335
## lsnativism 0.35037335
## lsorgterr 0.35037335
## lspop6amz 0.35037335
## lsswim_msex 0.35037335
## l2vim_vox 0.28948880
## l2authoritarian 0.28948880
## l2ideol 0.28948880
## l2nativism 0.28948880
## l2orgterr 0.28948880
## l2pop6amz 0.28948880
## l2swim_msex 0.28948880
## dhincome_all 0.05169443
## female 0.00000000
## age 0.00000000
## edu3_2 0.00000000
## edu3_3 0.00000000
## livingpartner 0.00000000
## intpol 0.00000000
```

## Appendix

We used ChatGPT in this assignment to help make the correlation matrix. It was helpful for understanding how to create the correlation plot for the specific variables used in the model we intend to replicate. It also highlighted key factors we missed when we initially generated the plot (e.g. filtering to only include data for the year 2019).

[Link to conversation](#)