Female Voters without ID in Texas During 2016 United States Election*

Emily Su

2024 February 17

Abstract.

Table of contents

1 Introduction **1**

2 Data **1**
2.1 Data Source and Measurements 2
2.1.1 Reason Dataset 2
2.1.2 Cleaned Map Dataset 2
2.2 Methodology 3

3 Results **3**
3.1 Texas females who voted without a photo ID during 2016 United States election 3
3.2 Texas females who voted without a photo ID during 2016 United States election (Across Race) 6
3.3 Reasons of Texas females across different race without a photo ID during 2016 United States election 9

4 Discussion **10**

Appendix **11**
.1 Acknowledgments 11
.2 Datasets Used 11
.3 Tables from Results 11

*Data and code are available at: https://github.com/moonsdust/toronto-collisions. The replication on Social Science Reproduction Platform can be found here: https://doi.org/10.48152/ssrp-cbev-p353

1 Introduction

The estimand is the effect of photo ID requirements on female voter turnouts in Texas.

Reasonable Impediment Declaration (RID)

2 Data

This paper seeks to further investigate Fraga and Miller (2022)'s paper claim that the common reason registrants do not have photo IDs is not due to socioeconomic hardships. We will also extend their research into the demographics of who filled out a RID with a focus on female voters in Texas. Asians, who was excluded from a majority of the figures in the paper, were also included in this paper. In the analysis, the paper used two datasets, `cleaned_reason_data.csv` and `cleaned_map_data.csv`. `cleaned_reason_data.csv` is from the paper we replicated (Fraga and Miller 2022) where its original dataset "fraga_miller_reasons_replication" is a Stata file (StataCorp 2023). The other dataset is a hybrid dataset where female voter counts comes from "fraga_miller_reasons_replication" (Fraga and Miller 2022) and geometry data on Texas comes from Texas Open Data Portal's "County Map - StratMap_County_poly" dataset, which is publicly available (Texas Open Data Portal 2024).

2.1 Data Source and Measurements

2.1.1 Reason Dataset

`cleaned_reason_data.csv`

In 2016, individuals in Texas without a valid photo IDs could vote in the 2016 United States election as long as they filled out a RID (Fraga and Miller 2022). The RID form were available languages such as English, Vietnamese, Spanish, or Chinese. The form asked for the voter's name, signature, a form of non-photo id such as current utility bill and the reason they are not able to present a valid photo ID at the polling place. The form provided the following options for reasons: Family Obligations ("family" in the dataset); Photo ID applied for but not not received ("applied" in the dataset); Lack of Transportation ("transport" in the dataset); Lack of birth certificate or other documents needed to obtain acceptable photo ID ("birthcert" in the dataset); Disability or Illness ("disability" in the dataset); Work schedule ("work" in the dataset); Lost or stolen photo ID ("lost" in the dataset); Other reasonable impediment or difficulty ("other" in the dataset) (Fraga and Miller 2022). Due to the Texas Public Information Act, Fraga and Miller (2022) was able to request for 16,097 RID forms from the 2016 US

election. Fraga and Miller (2022) also obtained a copy of the Texas voter registration file in 2017 from the Texas Secretary of State. These two sources were then merged to get further information on voters from the RIDs. Since the RIDs do not contain information about a voter's race and ethnicity, they used a R package, wru (R Core Team 2023) to get estimates of every voter's race based on address, surname, sex, and birthday (Fraga and Miller 2022). The three sources were then recorded into a spreadsheet where each row represents one form containing information about the voter's county, gender, race, and reason they don't have a valid ID. For the RIDs where the voter has indicated "Other" and wrote down a reason, Fraga and Miller (2022) and their research assistants took note of the reason and categorized their reason into three categories based on their discretion: Relocation ("relocation" in dataset), Hardship ("hardship" in dataset), and ID-Capable ("id_capable" in dataset). Relocation are reasons that mention recently moving, being a student, or waiting for a new ID (Fraga and Miller 2022). Hardship are reasons where hardships prevented the voter from getting a valid photo ID (Fraga and Miller 2022). ID-Capable are given to voters who were able to get a photo ID in the past (Fraga and Miller 2022). This spreadsheet was then turned into a dataset, which is available on Dataverse (Kuriwaki, Beasley, and Leeper 2023).

2.1.2 Cleaned Map Dataset

`cleaned_map_data.csv`

One half of the dataset comes `cleaned_reason_data.csv` while the other half comes from the "County Map - StratMap_County_poly" dataset, which is available on Texas Open Data Portal (Texas Open Data Portal 2024). The dataset contains geometry data (also known as "the_geom" in the dataset) that reconstructs a map of Texas and it also reconstructs the boundaries of Texas' counties (also called "county" in the dataset). These boundaries coordinates were obtained by the Texas Open Data Portal team (Texas Open Data Portal 2024) using the United States Geological Survey's digital 7.5 minute Topology Maps, who obtain their data from satellites every 3 years (United States Geological Survey 2024).

2.2 Methodology

The datasets used in this paper were retrieved, simulated, cleaned, analyzed, and tested using the R programming language (R Core Team 2023), tidyverse (Wickham et al. 2019), haven (Wickham, Miller, and Smith 2023), knitr (Xie 2014), janitor (Firke 2023), dplyr (Wickham et al. 2023), ggplot2 (Wickham 2016), sf (Pebesma and Bivand 2023), readr (Wickham, Hester, and Bryan 2024), dataverse (Kuriwaki, Beasley, and Leeper 2023), and ggpubr (Kassambara 2023).

For the Reason dataset, a new column called "white" was constructed to indicate if the voter was white or not since it was not in the original dataset. For the cleaned map data, the columns, "female_count", "female_white", "female_black", "female_asian", and "female_latinx" were

created from `cleaned_reason_data.csv` by filtering for female voters and summing the number of voters of each race. The cleaned reason dataset has around 16104 observations while the cleaned map dataset has 254 observations.

3 Results

3.1 Texas females who voted without a photo ID during 2016 United States election

Table 1: Percentage of Texas Voters who are female versus not female who voted in 2016 United States election without a photo ID.

Female	Not Female
0.57	0.43

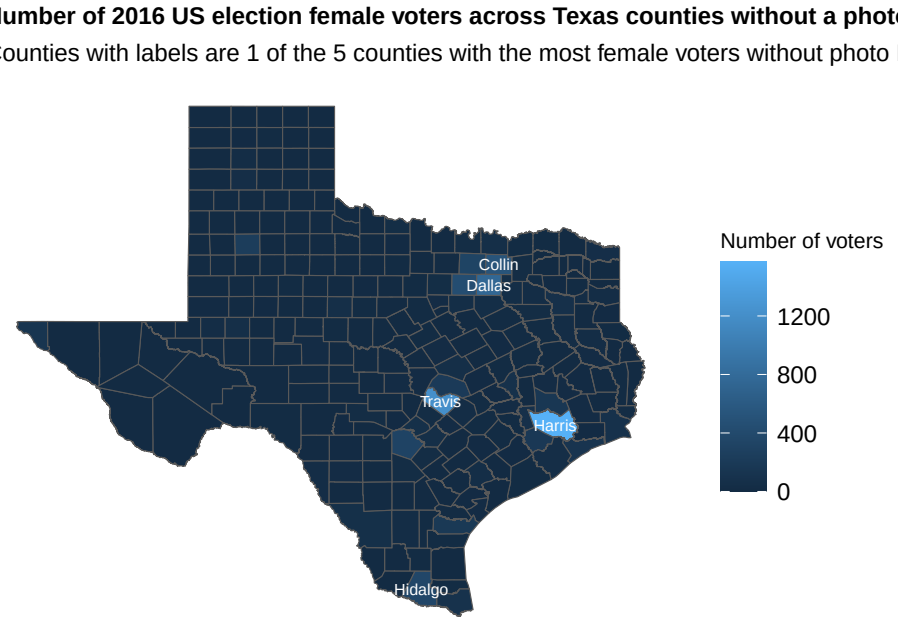

Figure 1: Number of female voters across Texas counties during 2016 United States election without a photo ID

Table 2: Number of female voters across Texas counties during 2016 United States election without a photo ID. The left column are the five county names with the most number of female voters without a photo ID. The right column is the number of female voters for each county.

County	Number of Female Voters
Harris	1570
Travis	1207
Dallas	752
Collin	518
Tarrant	442
Hidalgo	355

Table 3: Minimum, quartiles, median, mean, and maximum of female voter turnouts across Texas counties during 2016 United States election without a photo ID.

Number of female voters with no photo ID
Min. : 0.0
1st Qu.: 0.0
Median : 4.0
Mean : 35.7
3rd Qu.: 16.0
Max. :1570.0

Table 1 reveals 57% of Texas voters who filled out the RID from our dataset, which contains 16104 observations, were females while 43% were not females. From Figure 1 and Table 2, the five counties with the largest number of female voters who did not present a photo ID at the polling place in 2016 are Harris (1570 voters), Travis (1207 voters), Dallas (752 voters), Collin (518 voters), Tarrant (442 voters), and Hidalgo (355 voters).

We will first define the following: mean means the average value amongst the dataset, median means the data point found in the middle of a dataset, and the standard deviation is how far are values from the mean of a dataset. According to Table 3, the mean number of female voters without a photo ID across different counties was 36 voters, the median was 4 voters, and the standard deviation is 144 people.

In Figure 1, these counties are in a lighter colour, indicating a higher number of female voters who filled out RIDs. However, Travis and Harris are light blue compared to Dallas, which is a dark blue. When looking at the difference in female voters between Travis and Dallas, they have a gap of 455 voters. To understand the numbers, we will look at the counties' population numbers in 2016 and median income from 2012 to 2016 that are provided by the U.S. Census

Bureau. The counties' population on June 1, 2016 are the following: Harris (4,617,041), Travis (1,204,582), Dallas (2,587,462), Collin (942,453), Tarrant (2,021,746), and Hidalgo (850,187) (U.S. Census Bureau 2017). When looking at the population numbers, the populations of Harris, Travis, and Dallas are greater than one million in 2016. The percent (rounded to two decimal places) of each county's population who are female voters that filled out a RID are 0.03% for Harris, 0.10% for Travis, and Dallas, Collin, Tarrant and Hidalgo are under 0.01%. This reveals that Travis had the largest percent of female voters without a photo ID.

Now with median income, the counties' median incomes from 2012 to 2016 are $55,584 for Harris, $64,422 for Travis, $51,411 for Dallas, $86,188 for Collin, $60,373 for Tarrant, and $36,094 for Hidalgo with Texas' average median income being $55,508 (U.S. Census Bureau 2016). The average median income of the five counties are above the median income except for Dallas and Hidalgo.

3.2 Texas females who voted without a photo ID during 2016 United States election (Across Race)

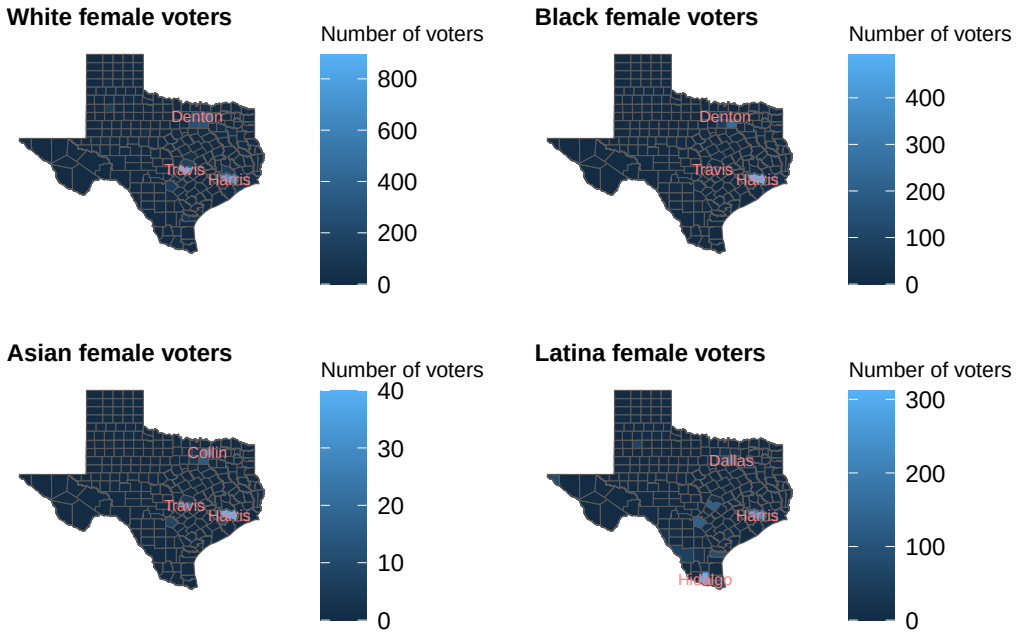

Figure 2: Number of female voters of different race across Texas counties during 2016 United States election without a photo ID

Table 4: Number of White female voters of different race across Texas counties during 2016 United States election without a photo ID. The left column are the five county names with the most number of White female voters without a photo ID. The right column is the number of White female voters for each county.

County	Number of Female Voters
Travis	895
Harris	706
Collin	432
Dallas	298
Tarrant	280
Denton	255

Table 5: Number of Asian female voters of different race across Texas counties during 2016 United States election without a photo ID. The left column are the five county names with the most number of Asian female voters without a photo ID. The right column is the number of Asian female voters for each county.

County	Number of Female Voters
Harris	40
Travis	27
Collin	24
Dallas	15
Fort Bend	12
Bexar	9

Table 6: Number of Black female voters of different race across Texas counties during 2016 United States election without a photo ID. The left column are the five county names with the most number of Black female voters without a photo ID. The right column is the number of Black female voters for each county.

County	Number of Female Voters
Harris	493
Dallas	248
Tarrant	81
Travis	64
Fort Bend	56
Denton	32

Table 7: Number of Latina female voters of different race across Texas counties during 2016 United States election without a photo ID. The left column are the five county names with the most number of Latina female voters without a photo ID. The right column is the number of Latina female voters for each county.

County	Number of Female Voters
Hidalgo	312
Harris	238
Dallas	137
Travis	136
Bexar	132
Nueces	89

Table 8: Minimum, quartiles, median, mean, and maximum of female voter turnouts across Texas counties during 2016 United States election without a photo ID categorized by race.

Number of White female voters with no photo ID	Number of Asian female voters with no photo ID	Number of Black female voters with no photo ID	Number of Latina female voters with no photo ID
Min. : 0.00	Min. : 0.0000	Min. : 0.000	Min. : 0.000
1st Qu.: 0.00	1st Qu.: 0.0000	1st Qu.: 0.000	1st Qu.: 0.000
Median : 2.00	Median : 0.0000	Median : 0.000	Median : 0.000
Mean : 21.47	Mean : 0.5787	Mean : 4.827	Mean : 7.185
3rd Qu.: 11.00	3rd Qu.: 0.0000	3rd Qu.: 0.000	3rd Qu.: 2.000
Max. :895.00	Max. :40.0000	Max. :493.000	Max. :312.000

Figure 2 shows the number of female voters who did not have a photo ID in Texas of the following groups: White, Black, Asian, and Latina. Figure 2, Table 6, and Table 4 reveals that for White and Black females, most female voters who filled out a RID were mostly in Denton, Travis, and Harris. Table 7, Table 5, and Figure 2 shows Hidalgo, Harris, and Dallas had the largest number of Latina voters fill out a RID while Harris, Travis, and Collin had the largest number of Asian female voters who filled out a RID.

When comparing the mean and median from Table 3 to Table 4 and Table 7, the five counties with the most number of White female and Latina who filled out a RID are above the mean and median number of female voters who filed out a RID. When observing Table 6 and Table 3, the number of Black female who filled out a RID was above the mean (36 voters) and median number of female voters (4 voters) who filled out a RID in the counties Harris, Dallas, Tarrant, Travis, and Fort Bend. However, with Denton, the number of Black females who filled out a RID (32 voters) were below the mean (36 voters) and above the median (4 voters) number of female voters who completed a RID. Table 3 and Table 5 reveals that Harris, which had the most number of Asian female voters without a photo ID at 40 voters, is above the mean and median of 36 voters and 4 voters respectively. However for the other four counties like Travis, the number of Asian female voters without a photo ID was below the mean but above the median.

Taking a look at Table 8, we see that the mean and median for White, Asian, Black and Latina females are below the mean and median of the number of female voters across Texas who filled out an RID. However, White females had a median of 2 voters while other groups of females had a median of 0. White females also has the highest mean number of voters who filled out an RID at 21 voters, followed by Latina with 7 voters and Black females with 5 voters, and Asian females had the smallest mean at 1 voter.

3.3 Reasons of Texas females across different race without a photo ID during 2016 United States election

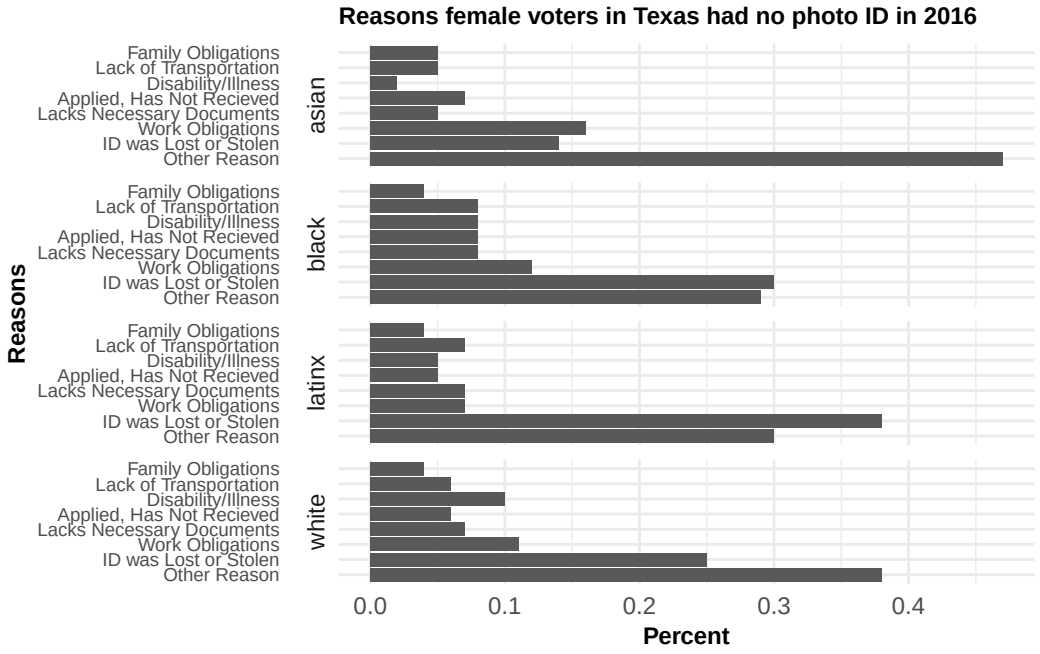

Figure 3: Impediments of female voters of different race across Texas counties during 2016 United States election without a photo ID

Figure 3 and Table 10 reveals the RID options that was selected by female voters who were unable to present a photo ID at the polling place. Figure 3 and Table 10 shows that for Asian females and White females, the option that was most commonly selected was "Other Reason", which made up 47% and 38% of their RIDs, respectively. For Black female and Latina voters, "ID was lost or stolen" was the most common reason at 30% and 38% of their RIDs.

A difference between what Fraga and Miller (2022) found and our result is that in terms of the commonly cited reason is that Fraga and Miller (2022) found that Black voters cited "Other Reason" and ID was Lost or Stolen equally in 30% of their RIDs. In our results as seen in Table 10 and Figure 3, "Other Reason" was selected less, which composed 29% of RIDs from Black female voters.

Table 9: Other impediments of female voters of different race across Texas counties during 2016 United States election without a photo ID

percent	reasons	race
0.32	Relocation	white

Table 9: Other impediments of female voters of different race across Texas counties during 2016 United States election without a photo ID

percent	reasons	race
0.33	Hardship	white
0.29	ID-Capable	white
0.27	Relocation	black
0.34	Hardship	black
0.34	ID-Capable	black
0.16	Relocation	latinx
0.31	Hardship	latinx
0.46	ID-Capable	latinx
0.47	Relocation	asian
0.30	Hardship	asian
0.18	ID-Capable	asian

Figure 4 and Table 9 shows the percent of the type of reasons that was given by female voters who selected the "Other Reason" option in the RID. Table 9 and Figure 4 reveals that for Asian females who selected this option, their most common reasons was related to "Relocation" (at 47%). For Black females, their most common reasons was related to "ID-Capable" and "Hardship" at 34% for each. For Latinas, their most common reason was related to "ID-Capable) at 46% and for White females, it was"Hardship" at 33%. These results are also seen in Fraga and Miller (2022)'s study with Black, Latinx, and White voters in Texas for the barriers mentioned in the "Other Reason" option.

4 Discussion

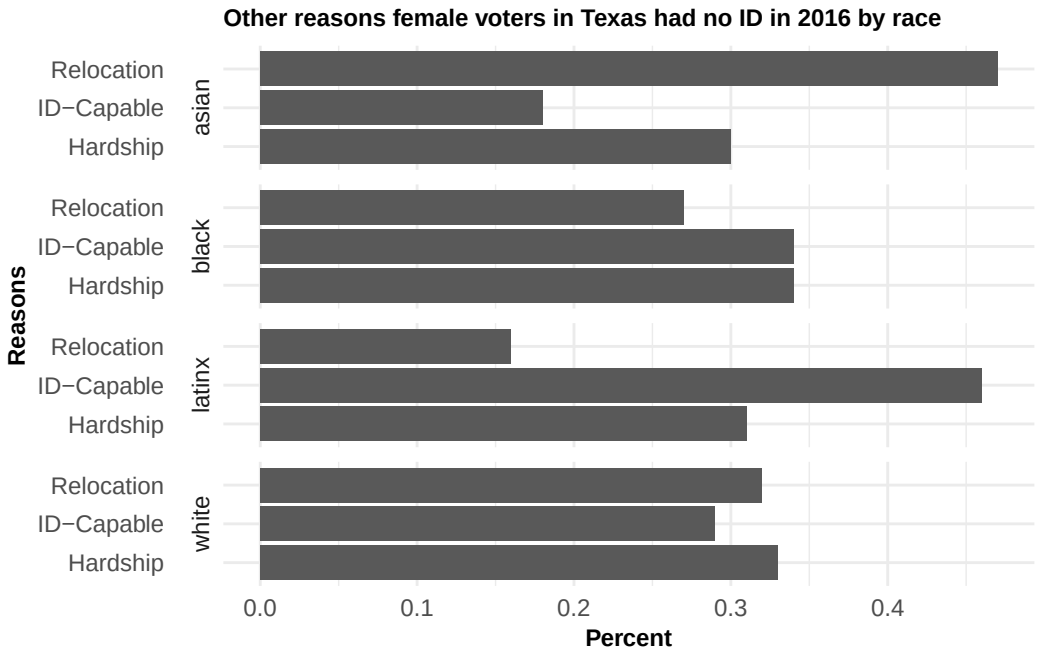

Figure 4: Other impediments of female voters of different race across Texas counties during 2016 United States election without an ID

Appendix

.1 Acknowledgments

I would like to acknowledge Alexander (2023) for some of the R code used in this paper to produce the tables and graphs.

.2 Datasets Used

Due to the length of the datasets horizontally, presenting glimpses of the datasets was not possible in the paper. However, the cleaned datasets that was used can be found here: https://github.com/moonsdust/voting/tree/main/data/analysis_data.

.3 Tables from Results

Table 10: Impediments of female voters of different race across Texas counties during 2016 United States election without a photo ID

percent	reasons	race
0.04	Family Obligations	white
0.06	Applied, Has Not Recieved	white
0.06	Lack of Transportation	white
0.07	Lacks Necessary Documents	white
0.10	Disability/Illness	white
0.11	Work Obligations	white
0.25	ID was Lost or Stolen	white
0.38	Other Reason	white
0.04	Family Obligations	black
0.08	Applied, Has Not Recieved	black
0.08	Lack of Transportation	black
0.08	Lacks Necessary Documents	black
0.08	Disability/Illness	black
0.12	Work Obligations	black
0.30	ID was Lost or Stolen	black
0.29	Other Reason	black
0.04	Family Obligations	latinx
0.05	Applied, Has Not Recieved	latinx
0.07	Lack of Transportation	latinx
0.07	Lacks Necessary Documents	latinx
0.05	Disability/Illness	latinx
0.07	Work Obligations	latinx
0.38	ID was Lost or Stolen	latinx
0.30	Other Reason	latinx
0.05	Family Obligations	asian
0.07	Applied, Has Not Recieved	asian
0.05	Lack of Transportation	asian
0.05	Lacks Necessary Documents	asian
0.02	Disability/Illness	asian
0.16	Work Obligations	asian
0.14	ID was Lost or Stolen	asian
0.47	Other Reason	asian

References

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

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://github.com/sfirke/janitor.

Fraga, Bernard L., and Michael G. Miller. 2022. "Who Does Voter ID Keep from Voting?" *The Journal of Politics* 84 (2). https://doi.org/10.1086/716282.

Kassambara, Alboukadel. 2023. *Ggpubr: 'Ggplot2' Based Publication Ready Plots*. https://CRAN.R-project.org/package=ggpubr.

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

Pebesma, E, and R Bivand. 2023. *Spatial Data Science: With Applications in r*. Chapman; Hall/CRC. https://doi.org/10.1201/9780429459016.

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

StataCorp. 2023. *Stata Statistical Software: Release 18*. College Station, Texas: StataCorp LLC. www.stata.com/.

Texas Open Data Portal. 2024. "County Map - StratMap_county_poly." https://data.texas.gov/dataset/County-Map-StratMap_County_poly/m3yf-ffwm/about_data.

United States Geological Survey. 2024. "United States Geological Survey." www.usgs.gov.

U.S. Census Bureau. 2016. "Texas Counties: Median Household Income, 2012-2016." https://www.texascounties.net/statistics/householdincome2016.htm.

———. 2017. "Population Estimates for Texas Counties, 2010-2017: Arranged in Descending Order." https://www.tsl.texas.gov/ref/abouttx/popcnty201011.html.

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, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2024. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.

Wickham, Hadley, Evan Miller, and Danny Smith. 2023. *Haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*. https://CRAN.R-project.org/package=haven.

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. http://www.crcpress.com/product/isbn/9781466561595.