Analyzing Black Women's Opinions on Policing

STAT 499 Undergraduate Research (Directed Reading Program Winter 2025)

Authors: Lily Yao & Srinidhi Muralidharan

1. Introduction

The Innocence Project

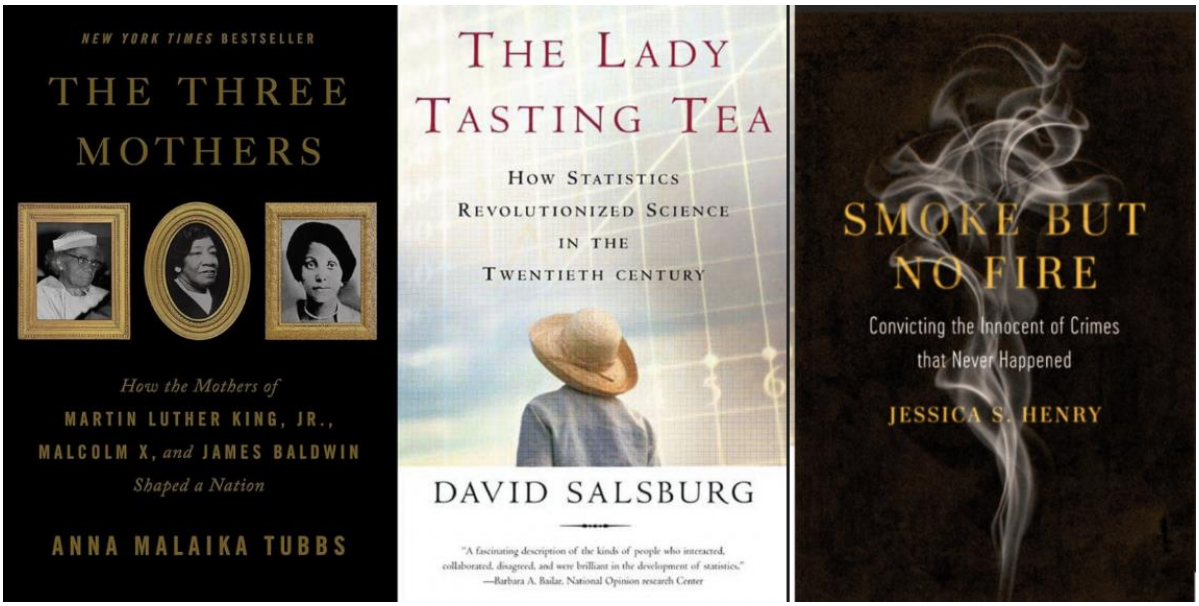
The main objective of our mentorship with Ms. Cindy Elder was to discuss The Innocence Project, the connection between statistics and prison, the history of statisticians, and use linear regression on our first data analysis project. Now in the beginning of the quarter, we read about The Innocence Project, an organization that works to exonerate people that are wrongfully convicted of a crime. The statistical findings about the Innocence Project were eye-opening. From an excerpt of *Smoke But No Fire: Convicting the Innocent of Crimes that Never Happened* by Jessica S. Henry, the reasons innocent people are convicted are police misconduct, forensic errors, witness misidentification, and false confessions. Specifically, police misconduct can happen due to pressure to meet arrest numbers, raise revenues from fines and court fees, as well as taking bribes. We discussed that many innocent people that The Innocence Project represent are people on death row. While Black men make up only 13.4% of the overall US population, they make up 41% on U.S death row (Morris and Schoeller 12). This statistic provides insights on the racial inequalities in the criminal justice system. Also, comparing the statistics on how wrongful convictions vary by race, exonerations of Black people convicted are later linked to police misconduct. For people of Hispanic descent, it is perjury, while White people are wrongfully convicted due to misleading or false forensic evidence (Morris and Schoeller 9). The numbers behind race and prison became a concern we wanted to explore.

The Three Mothers

In one of our meetings, we read excerpts from *The Three Mothers: How the Mothers of Martin Luther King, Jr., Malcolm X, and James Baldwin Shaped a Nation* by Anna Malaika Tubbs. In the excerpts we read, the statistics were rather disturbing. Black pregnant women are 11 times more likely to face intimate partner violence than pregnant White women (Tubbs 5). However, Black women are more likely to not report the domestic violence they face to protect their partners that could potentially face police brutality (Tubbs 6). Furthermore, in 2017, Black women were imprisoned at twice the rate of white women as they couldn't pay for a lawyer or bail (Tubbs 6). Reading about the systemic inequalities faced by Black women, we finalized our plan to analyze one aspect of the complex relationship between Black women and the prison system.

Women in Statistics

With excerpts of *The Lady Tasting Tea* by David Salsburg, a fragment of our reading program was learning about the women that came before us as statisticians and scientists. For instance, during the World Wars, women worked as cryptographers, cryptologists, and human computers. While the field of science was heavily gender segregated, there were outliers like female theoretical physicist Maria Goeppert-Mayer. Along with her husband, she won a Nobel Prize for Physics in 1963 by discovering the nuclear shell model. Unfortunately, she was unacknowledged as a scientist and instead, the newspaper headline was “La Jolla Mother Wins Nobel Prize” (SupaduDev 93). Florence Nightingale, a self-taught statistician invented the pie chart to illustrate that wounds left unattended from war were causing deaths during the Crimean War (Salsburg 151). FN David, named after Florence Nightingale, worked in the Ministry of Home Security building statistical models in 1939. Her models helped estimate the number of casualties and the effects of bombs on electricity and sewage systems (Salsburg 153). We've taken inspiration from these influential women as future statisticians and kept it as a reason to pick a female-centric dataset to analyze.



2. Dataset

Overview

The *In Her Place Survey on Black Women and Policing* is a dataset from Harvard Dataverse on the experiences and opinions that Black women have had with Policing in the U.S. It was a national survey taken in October 2020 of women who are 18 or older. Some columns that were

of interest to us were age, region, political party, demographics, daily discrimination with police, and childhood vs adult interactions with police. The dataset intrigued us because of its unique variables such as “policetalk” (Ever received advice about how to interact with police?), “partyid”(Political party identification), “recent_attendedprotest”(Have you attended a rally or protest about police violence in the past 12 months?), and “pvch_ignoreda” (As a child, have you ever been dismissed or ignored by officer when reporting an incident involving self).

Authors

The principal investigators of the survey are Shannon Malone Gonzalez, a sociology professor at the University of North Carolina and Yasmin Irizarry, an African Diaspora Studies professor at University of Texas at Austin. Dr. Gonzalez has expertise in applied statistical modeling and gender studies with her research focusing on the relationship between marginality and policing. Dr. Irizarry research focuses on race in relation to schooling, social attitudes, and social control. She is also the founding director of the Numbers 4 Justice Lab at the University of Texas at Austin.

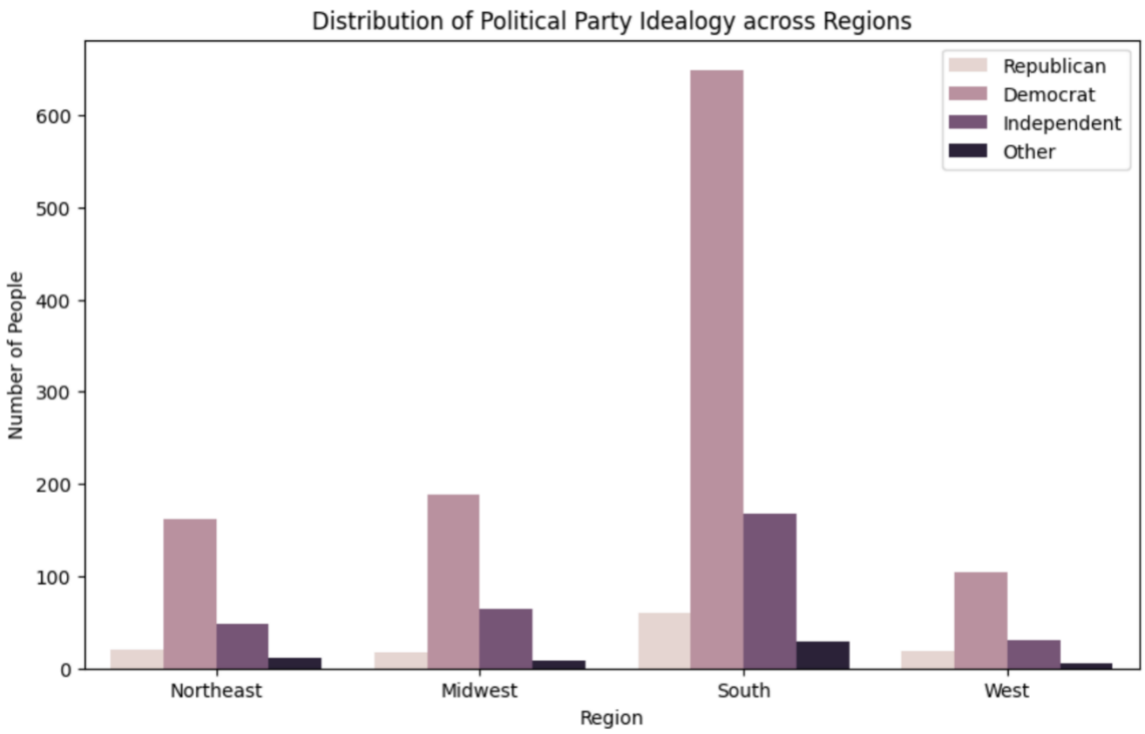
Challenges

A complication we encountered while selecting this dataset was the number of columns that are restricted to the public. While there are 1,603 observations in total, only 166 variables are public out of 1,700 variables. We are interested in columns about the opinions of Black women on defunding the police, abolishing the police department, and if they recommend talking to their kids about police. However, these are restricted access, and we’ve emailed the investigators for permission. As soon as we get access to these columns, we hope to expand our research questions.

Limitations

We’d figured that our data analysis should be considered with discretion when understanding the limitations of the dataset. While creating bar plots, we compared the region (Midwest, South, Northeast, and West) and the political party (Democrats, Republicans, Independent, and Other) that Black women support. We saw a huge spike in the number of women that are Democrats in the South compared to the other regions.

Connecting where the co-authors teach at, University of Texas at Austin and the University of North Carolina, we concluded that the dataset has convenient samples in the South since it would be easier for the researchers to survey people that they knew and are near their respective universities. The In Her Place dataset may be a “national” survey on the opinions of Black women, but we decided not to interpret this as a national survey but a survey of Black women that were willing to share their experience with a research study.



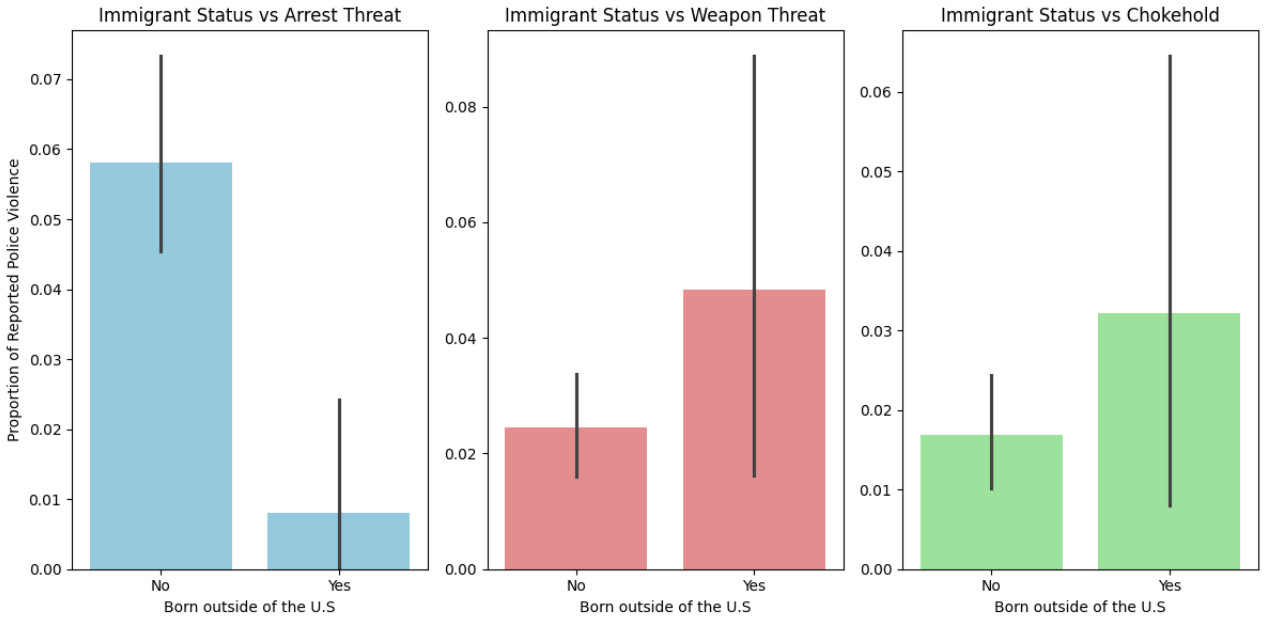
Graph I: Bar Graph of Distribution of Political Party Idealogy across regions of the U.S.

3. Data Visualizations

We have chosen to do exploratory data analysis and don't necessarily have a rigid set of research questions. Our objective was to find columns with adequate data and columns that were of interest to us. We wanted to explore columns like police talk, age, region, and if Black women are foreign-born as factors with how they've interacted with police. We have a few miscellaneous questions we wanted to answer after browsing through the dataset.

1. How does police interaction differ between immigrant statuses?
2. Do the political ideologies of Black women affect whether they have recently attended a police violence protest?
3. Does being ignored by police when reporting an issue as a child impact willingness to seek police help in adulthood?
4. Does experiencing a police talk at a young age impact arrest rates in adulthood?

Bar Graph of Immigrant Status vs various threats from Police.



Graph II: How does police interaction differ between immigrant status?

It appears that a larger portion of respondents born in the U.S. have faced the threat of arrest, which may attribute to more data collected from US-born respondents, whereas respondents born outside the U.S. experience a higher proportion of threats involving weapons (such as a baton, taser, or gun) and being placed in a chokehold by officers.

Noting that there might be biases in these visualizations because several factors could contribute to these differences. Selection bias may play a role, as individuals born outside of the country who are arrested may face deportation, making them less likely to appear in survey data. Also, response bias could influence the results, as individuals may be hesitant to report arrests or police encounters due to fear or social stigma, even in an anonymous survey. Differences in sample size also affect the reliability of estimates, as respondents born in the US make up a larger portion of the data, which may impact the perceived trends in police interactions.

Linear Regression

We decided to use linear regression to compare how different factors impact the experiences that Black women have had with police. Linear regression is used to determine how the average value of a continuous outcome or “y” varies with the value of a single predictor of “x”. The equation can be illustrated with Figure 1 or the equation $y = mx + b$. $b0$ is the average value of

the outcome at zero or the y-intercept. b_1 is the slope of the regression line while X_1 is the independent variable.

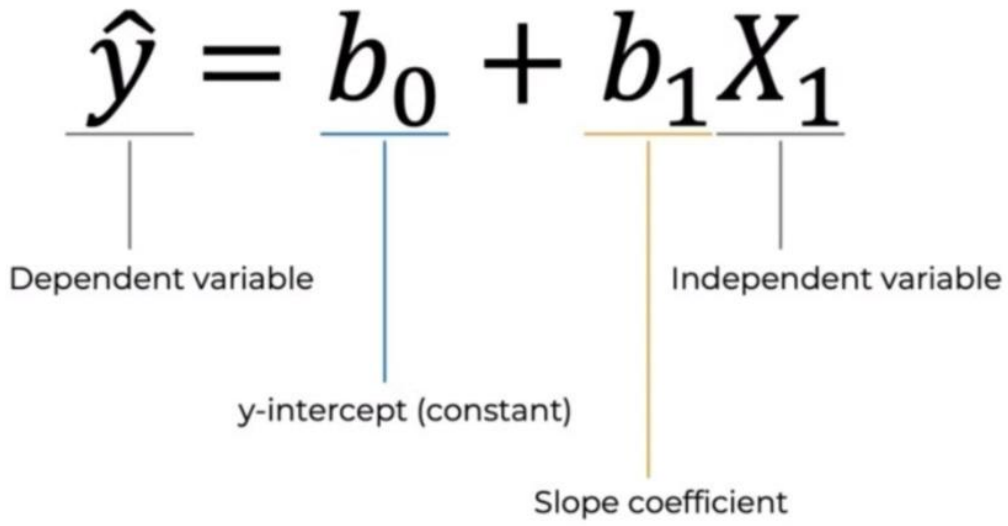


Figure I: Equation for Linear Regression

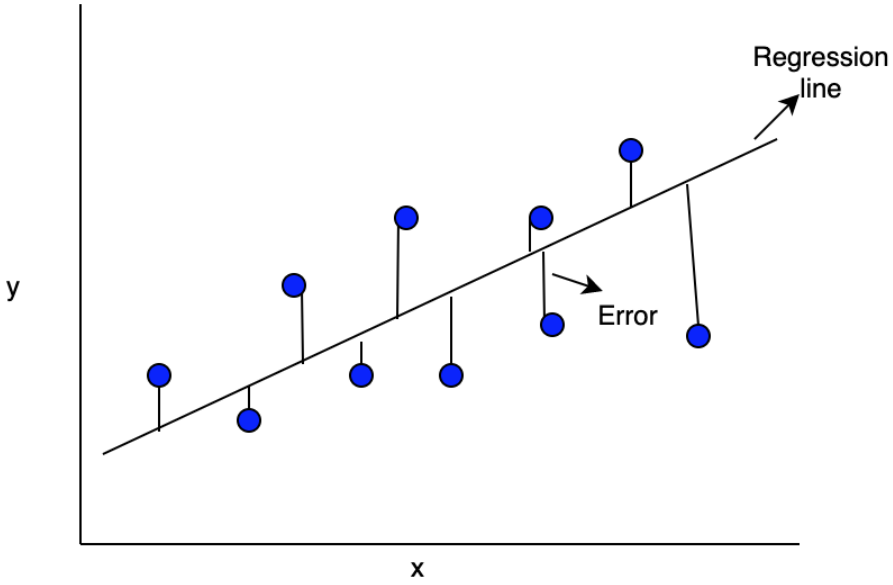
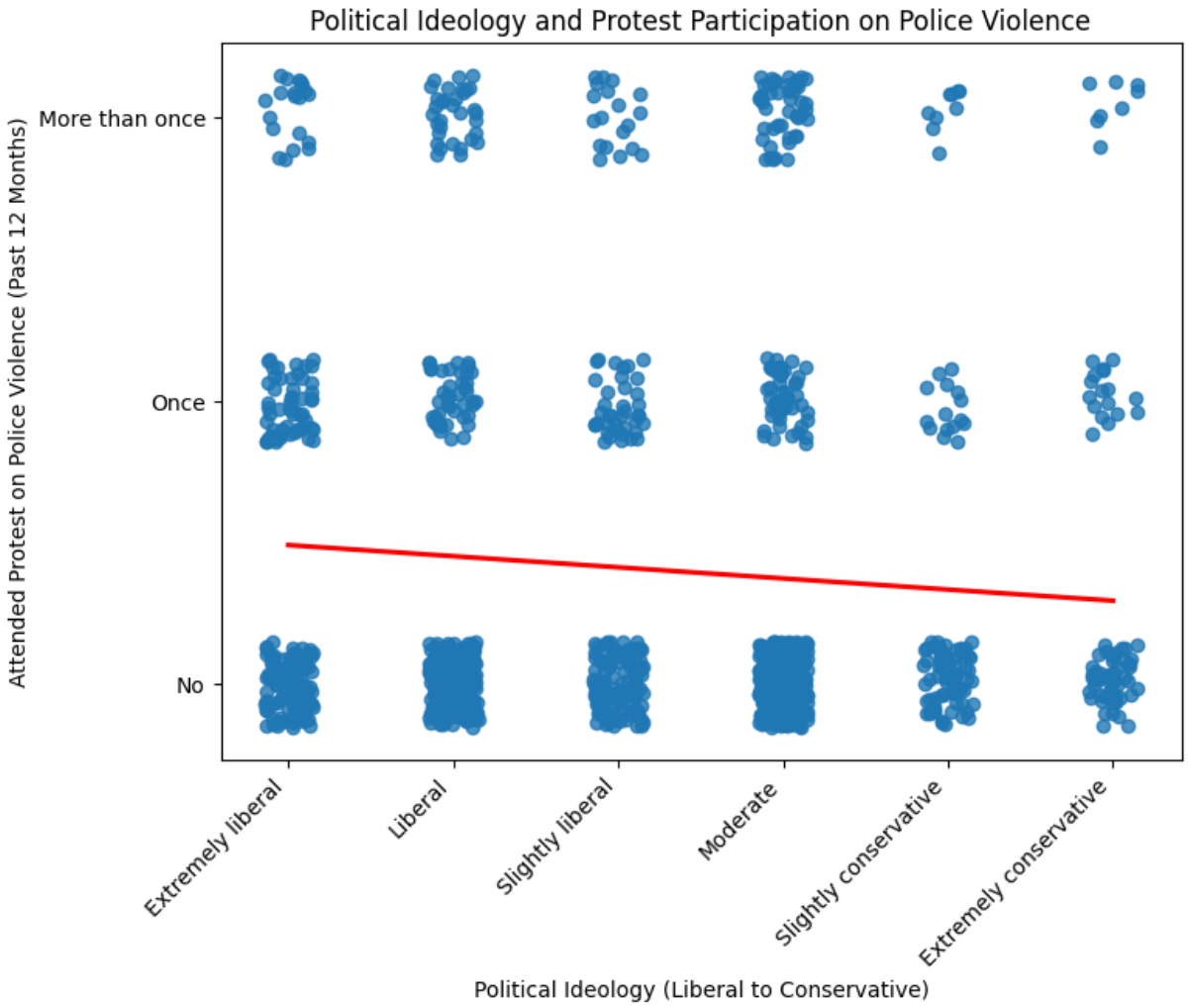


Figure II: Visual Representation of Linear Regression

The differences between the actual values and the predicted values in a linear regression graph are called errors. These errors have normal distributions, most errors are small, extreme outliers in errors are rare, and the errors should cancel out. The term “errors should cancel out” means that the differences between the actual and predicted values should have a mean of zero indicating that the model is unbiased and doesn’t under or overestimate the predicted outcome.

Also, with the In Her Place data, we have worked with binary and categorical data. When using binary or categorical data, if the predictor groups are too small, this can affect the reliability of the linear regression model. To combat this, we have selected columns with adequate amount of data and are cautious to not oversimplify our understanding of what the linear regression model tells us.

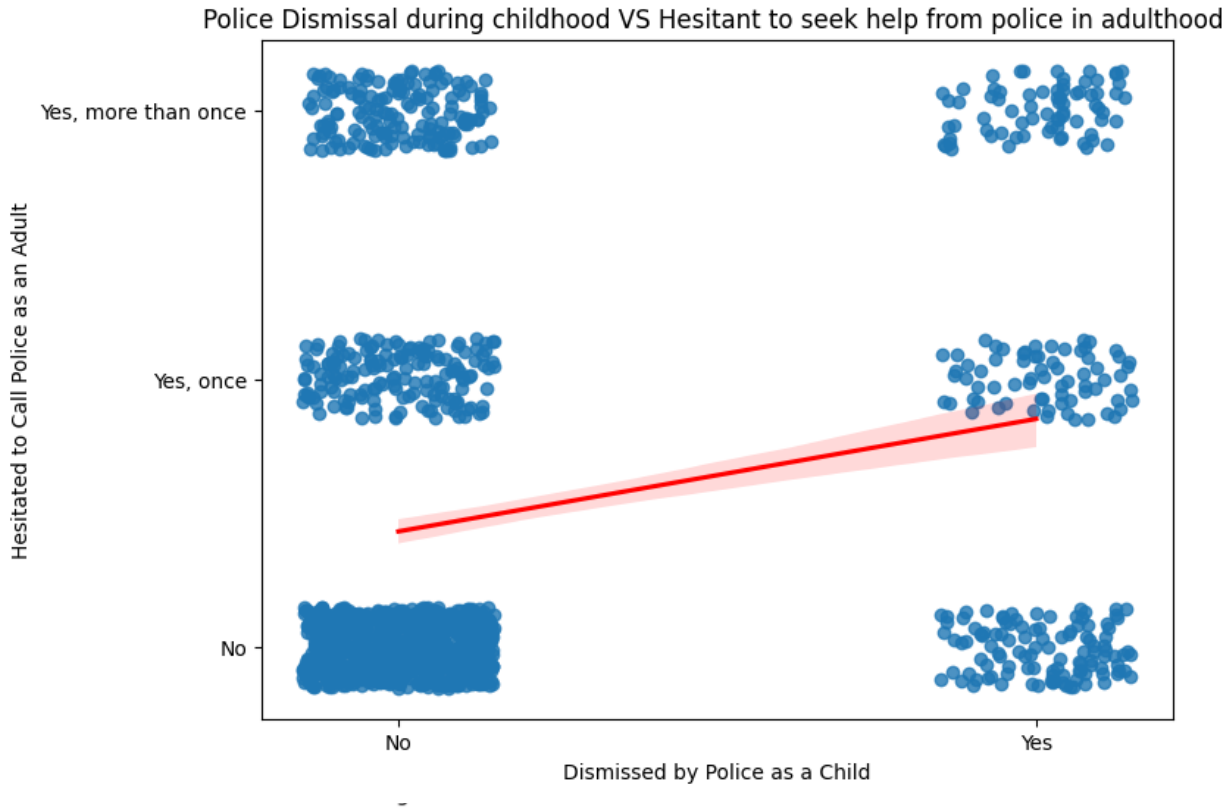



```
Intercept: 0.5096975332074305
Coefficient: -0.039639634422109896
R-squared: 0.00695641531725355
```

Graph III: Does the political Idealogy of Black women affect whether they have recently attended a police violence protest?

Using simple linear regression, among the black female respondents to the In Her Place Study this year, we estimate that for each one-unit increase in ideological conservatism (from more

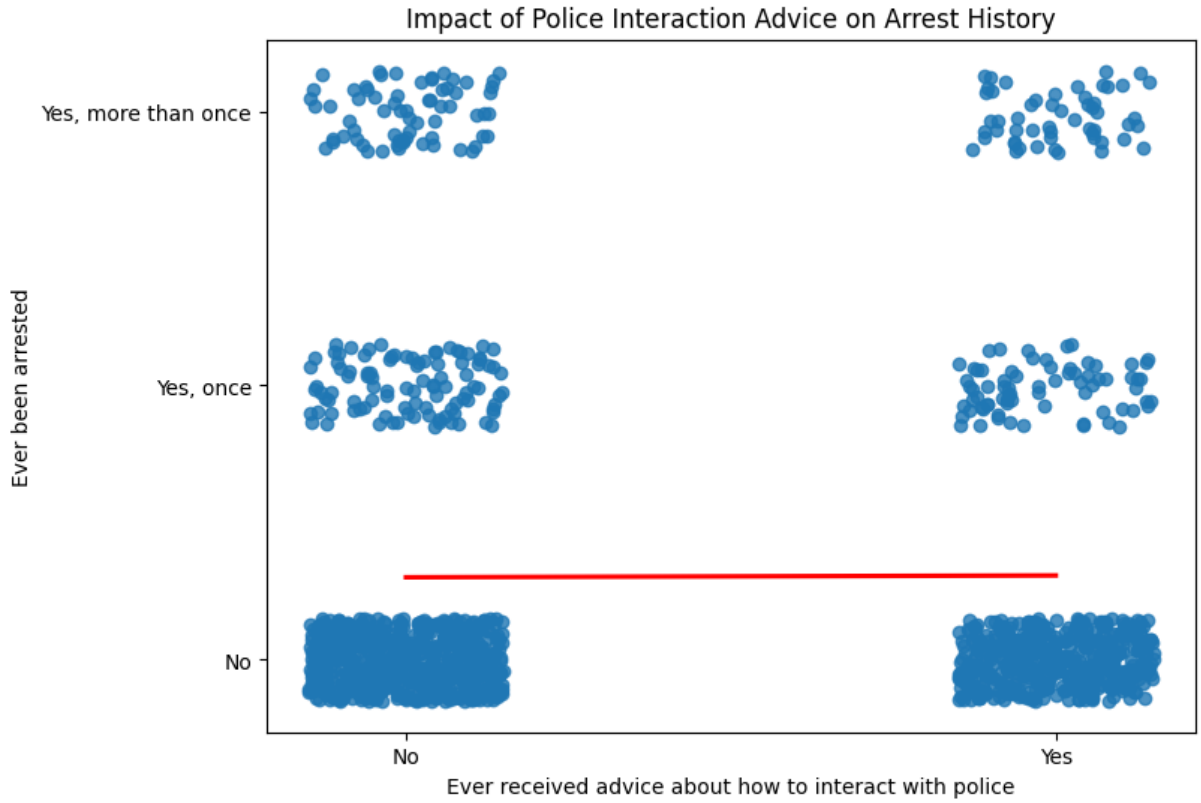
liberal to more conservative), the average protest attendance category decreases by 0.03963. This suggests that black women who identify as more conservative tend to report slightly lower protest attendance on average. However, this is an estimate from the model and does not imply causation. However, this is an estimate from the model and does not imply causation. Other factors, such as political party affiliation or broader social influences, may also contribute to this pattern.




```
Intercept: 0.4224598930481285
Coefficient: 0.4034660328777969
R-squared: 0.04001207732588763
```

Graph IV: Does being ignored by police when reporting an issue as a child impact willingness to seek police help in adulthood?

Using simple linear regression, among the black female respondents to the In Her Place Study this year, we estimate that the average number of adult women who hesitated to call police is 0.4 higher per person for respondents who had been ignored by police during childhood, compared to respondents who are not.




```
Intercept: 0.3003851091142488
Coefficient: 0.006528887513070286
R-squared: 2.6984493695314526e-05
```

Graph V: Does experiencing a police talk at a young age impact arrest rates in adulthood?

Based on the simple linear regression among the black female respondents to the In Her Place Study this year, we estimate that there is no significant association between childhood exposure to a police talk and the likelihood of arrest in adulthood. This suggests that, on average, having a police talk during childhood does not appear to affect the likelihood of arrest as an adult, based on our model estimate.

3. Conclusion

Therefore, with our mentor Ms. Elder, we've learned about the Innocence Project and the statistics of racism and sexism in prison systems. With the In Her Place data, we've attempted exploratory data analysis to understand the patterns of Black women's experiences with police in relation to age, police talk, political idealogy and more. We hope to extend our analysis as soon as we get access to the restricted columns.

Sources

Morris, Phillip, and Martin Schoeller. “Sentenced to Death, but Innocent: These Are Stories of Justice Gone Wrong.” *History*, 18 Feb. 2021, www.nationalgeographic.com/history/article/sentenced-to-death-but-innocent-these-are-stories-of-justice-gone-wrong.

Henry, J. S. (2020). *Smoke but no fire: Convicting the innocent of crimes that never happened* (1st ed.). University of California Press.

Tubbs, A. M. (2021). *The three mothers: How the mothers of Martin Luther King, Jr., Malcolm X, and James Baldwin shaped a nation*. Flatiron Books.

Salsburg, D. (2001). *The lady tasting tea: How statistics revolutionized science in the twentieth century*. W. H. Freeman.

SupaduDev. “Women in Science: Struggle & Success, the Tale of Mileva Einstein-Marić, Einstein’s Wife.” *MIT Press*, 27 May 2022, mitpress.mit.edu/women-in-science-struggle-success-the-tale-of-mileva-einstein-maric-einsteins-wife/.

Malone Gonzalez, Shannon; Irizarry, Yasmiyn, 2024, "In Her Place Survey on Black Women and Policing", https://doi.org/10.7910/DVN/UXTOVN, Harvard Dataverse, V1

Figure III: https://www.youtube.com/watch?v=HYgHKxWs1b4 (Super Data Science, 2023, timestamp 0:34)

Figure IV: https://community.cloudera.com/t5/Community-Articles/Understanding-Linear-Regression/ta-p/281391 (Cloudera Community,2019, Figure 2)