Algorithmic Analysis: Law Schools & Bar Exams

Ethan Wong, Ria Mittal, & Nick DeLeon

About Us

Ethan Wong

Ria Mittal

Nick DeLeon

Table of Contents

01

Model Development

Domain Selection

Concentration on Education

Broader applicability for law school admissions

Data from the LSAC

7 years worth of data from 1991 - 1997

Approx. 22,400 Law Students

Captures information such as LSAT score, undergraduate GPA, gender, and race

Data Preparation

1. ## Dataset Overview

   Started with 22,407 Rows and 39 columns

2. ## Missing Value Strategy

   Retained all rows and used mode and algorithm imputation for nulls

3. ## Feature Selection Process

   Manual selection based on context, target correlation and outside EDA

4. ## Final Dataset for Modeling

   Reduced to 17 columns

Algorithm Training

Data Splitting

For gender, segmented by bar pass rates for men and women and group rations.

For race, maintained the distribution of the five different groups

Model Selection and Optimization

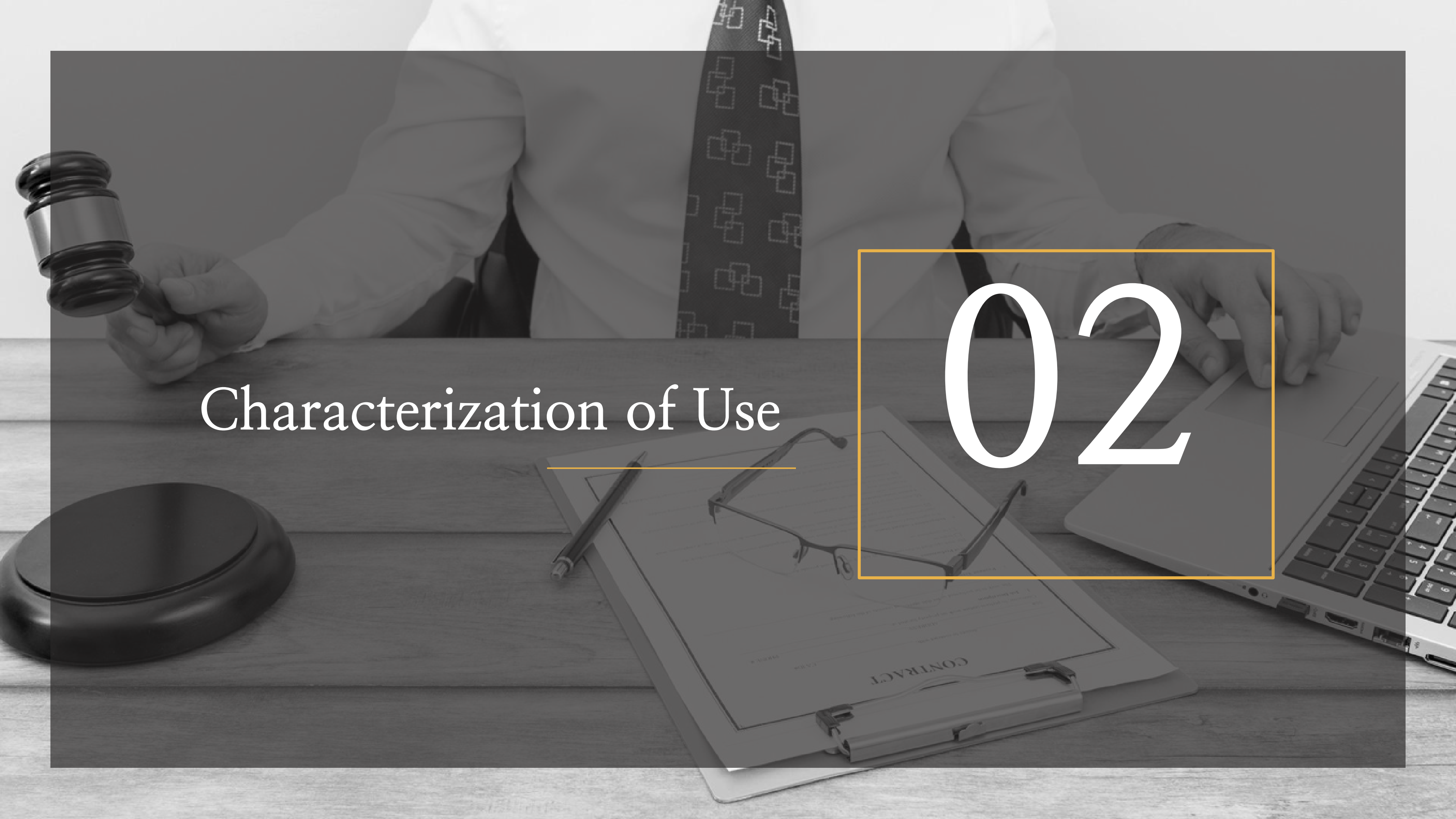
Utilized GridSearchCV for various models and hyperparameters

Aimed for balanced accuracy with 10-fold Stratified Cross-Validation

Optimal Model Results

Logistic Regression with standardized data, tuned regularization, and balanced class weights

Achieved 78% accuracy and 88% AUC on testing data



02 Characterization of Use

Admissions Reform and Bias Reduction

Relying exclusively on LSAT scores to determine admissions has historically kept law schools predominantly white. Now, law schools use a culmination of factors to predict the likelihood of a student passing the bar. Does this eliminate racial bias?

X

- Transcripts,
- Test scores (e.g., LSAT),
- Life experience
- Work/Volunteer Experience
- Recommendations
- Resilience
- Goals

Y

Pass the bar

03

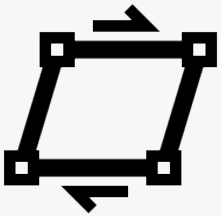
Ethical Considerations

Sampling Bias

Occurs when when the data used to train an algorithm fails to represent the target population accurately

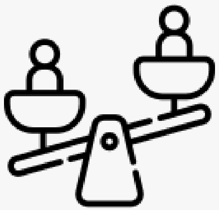
Immediate Effect

Skews Predicts about Current Applicants



Long Term Effect

Reinforces educational and professional inequities



Algorithmic Bias

Occurs when attributes are not applicable/unethically used to predict an outcome

For law school admissions, using demographic attributes perpetuates the history of exclusion among minorities.

Steps to take:

- Understand the context behind each attribute
- Make sure each attribute used in training the model is ethically justifiable
- If disparities between demographics occur, identify why and eliminate underlying bias

Unintended Consequences

Diversity & Inclusion

Outdated data

Favor People from Historically Successful Backgrounds

Education

Alter the curricula to cater to algorithm preferences

Compromising on quality

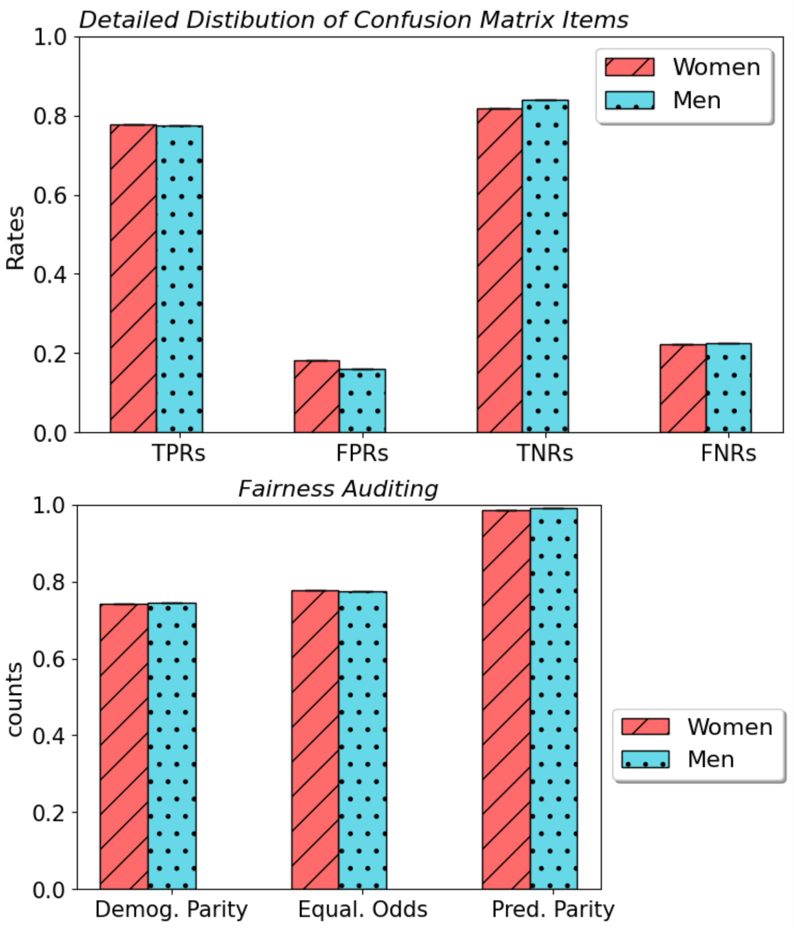
Ethical

Students alter demographic data

Compromising on data integrity & reliability

04

Empirical Bias Assessment



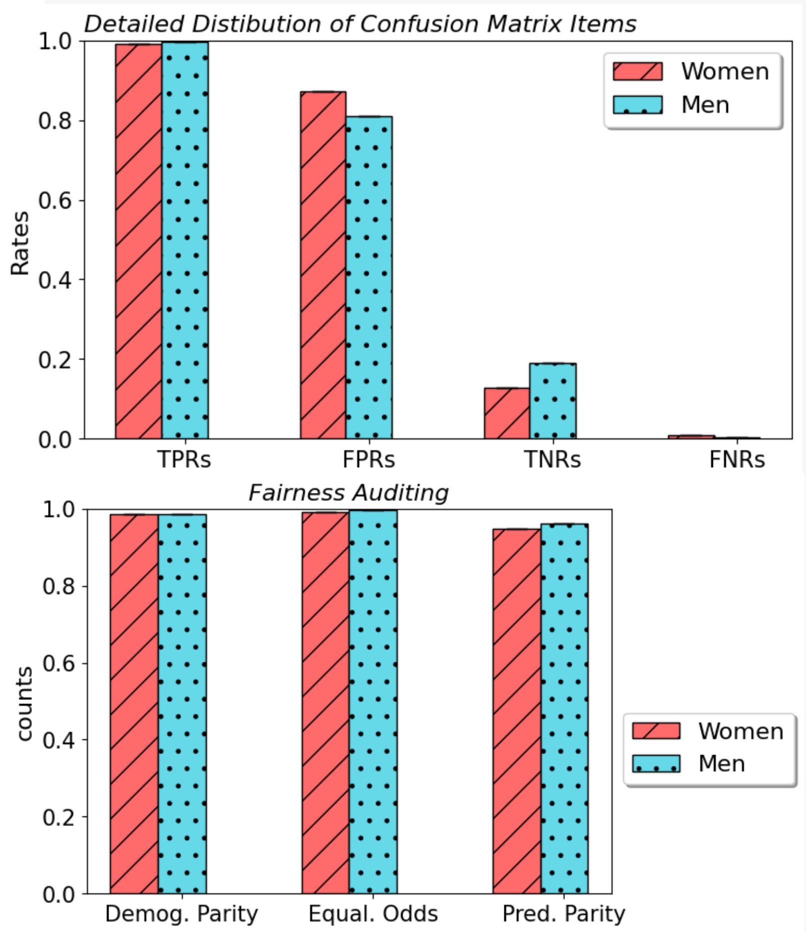
Analysis on Gender

Only a 0.02 difference between False Positive and True Negatives

No violations of demographic parity, equalized odds, or predictive parity

From a law school admissions stakeholder's perspective, minimizing false negatives is crucial: missing out on admitting capable candidates is detrimental.

During our initial model development phase, we actually were able to minimize false negatives… but at a cost.



Analysis on Gender Cont.

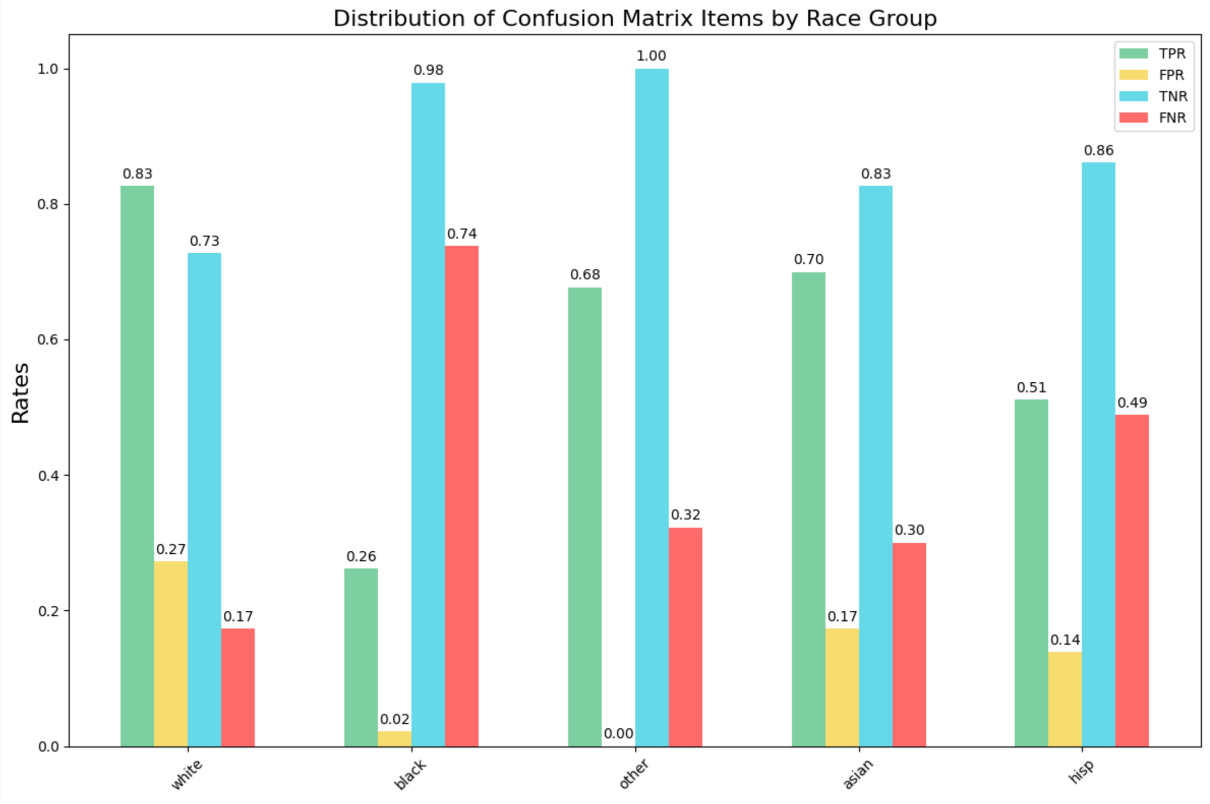
These are the results of a model optimized for ROC AUC and without adjusting for class imbalance

Fewer false negatives and 18% increase in test accuracy, but significant increase in false positives and decrease in true negatives

Consistent fairness metrics between genders

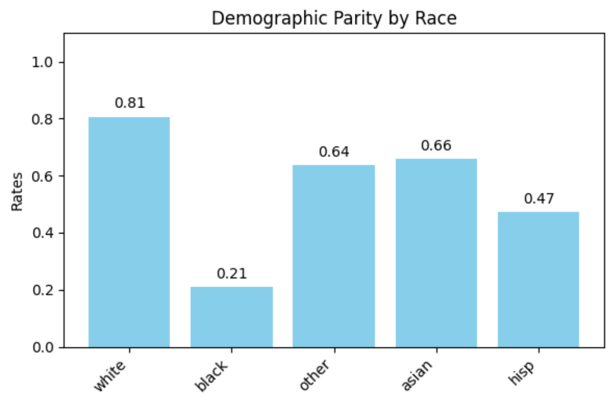
There is a trade-off between model utility and performance; optimizing for accuracy alone gives a skewed view of effectiveness

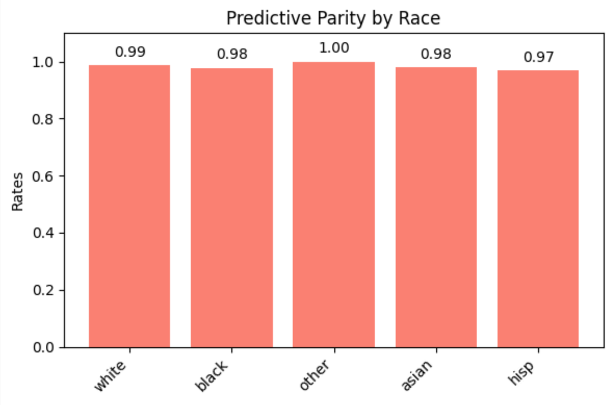
Analysis on Race

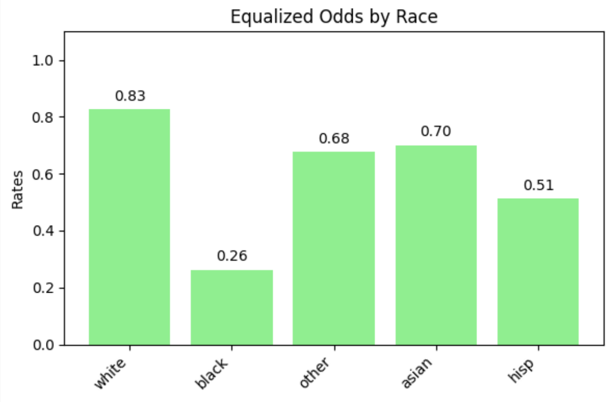


There is a significant amount of disparity in confusion matrix rates between White, Black, Asian, Hispanic, and “Other” racial groups.

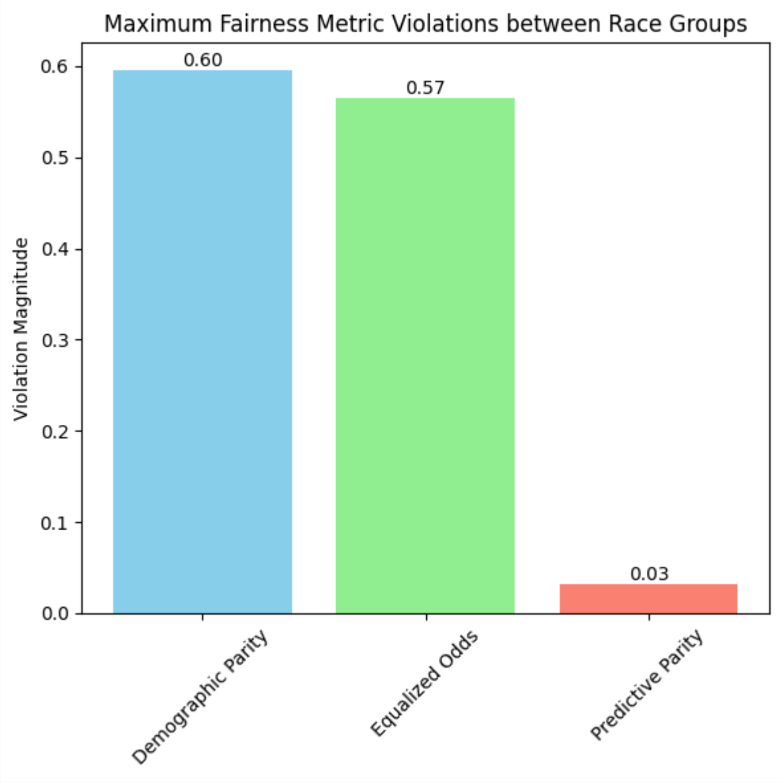
Analysis on Race Cont.







Analysis on Race Cont.



The discrepancy between white and black individuals drives the highest violation in demographic parity and equalized odds. A small discrepancy between hispanic and “other” persons is the highest violation of predictive parity.

Next Steps

Intersectional Analysis

Investigate potential disparities between combinations of race and gender

New Data

Data needs to be representative, and not just in terms of recency

Other Bias Mitigation Approaches

Pre-processing, training time constraints, and post processing

Thanks

We are open to any questions!

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik.

Work Cited

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities In Commercial Gender Classification Proceedings of Machine Learning Research https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

De-Arteaga, M. (2024). S9-Class Ethics of AI: Measures of Algorithmic Bias. The University of Texas at Austin McCombs School of Business. *Canvas.* https://utexas.instructure.com/courses/1379214/files/76820760?module_item_id=13683842

Engler, A. (2021). Enrollment algorithms are contributing to the crises of higher education. *Research*. Brookings. https://www.brookings.edu/articles/enrollment-algorithms-are-contributing-to-the-crises-of-higher-education/

Koons, S. (2018). Researchers examine impact of law school admissions reform on universities. *Education*. https://www.psu.edu/news/education/story/researchers-examine-impact-law-school-admissions-reform-diversity/

Kuris, G. (2024). 7 Deciding Factors in Law School Admissions. Law Admissions Lowdown. U.S. News & World Report. https://www.usnews.com/education/blogs/law-admissions-lowdown/articles/7-deciding-factors-in-law-school-admissions#

Lowry, T. (2023). The Good, the Bad, and the Ugly of AI and Admissions. *Law:Fully*. LSAC. https://www.lsac.org/blog/good-bad-and-ugly-ai-and-admissions

Merritt, D. (2016). The Gender bias in Law School Admissions. Business & Practice. Bloomberg Law. https://news.bloomberglaw.com/business-and-practice/the-gender-bias-in-law-school-admissions#

Nussbaumer, J. (2006). Misuse of the Law School Admissions Test, Racial Discrimination, and the De Facto Quota System for Restricting African-American Access to the Legal Profession. *St. John's Law Review*, *80* (Winter 2006), 162–182. https://scholarship.law.stjohns.edu/cgi/viewcontent.cgi?article=1180&context=lawreview

Spencer, Christian. (2021) "More than a Third of White Students Lie about Their Race on College Applications, Survey Finds." The Hill, The Hill, 1 Nov. 2021, thehill.com/changing-america/enrichment/education/577722-more-than-a-third-of-white-students-lie-about-their/.

Usylesses. (2022). LSAC dataset - EDA and predictions. Kaggle. https://www.kaggle.com/code/eds8531/lsac-dataset-eda-and-predictions

Wightman, L., Ofer, D (2022). Law School Admissions Bar Passage. Kaggle. https://www.kaggle.com/datasets/danofer/law-school-admissions-bar-passage/data.