D4J: Data for Justice to Advance Transparency and Fairness

Alex Dong, Anhthy Ngo, Christine Shen, Andrew Yeh
Center for Data Science, New York University
{awd275,an3056,zs1534,ay1626}@nyu.edu

Abstract

The Fort Bend County, Texas, District Attorney's Office has recently upgraded their case-management infrastructure and is looking to take the next step in analyzing the data they've collected. We investigate prosecutorial decision-making and provide the District Attorney's Office in Fort Bend County, Texas, with the data infrastructure to examine their own practices. The final products are a data dashboard built in R Shiny and a statistical analysis of disparity in outcomes for based on race and and gender for the top seven most-common crimes. We find evidence for moderate to large disparities at specific decision-making points for certain crimes.

1 Introduction

Prosecutorial decision-making is broadly considered the "black box" of the criminal justice system. Prosecuting attorneys have comprehensive discretion in formulating decisions that influence criminal case outcomes, more than any other actor in the American justice system(Frederick and Stemen, 2012). As key decision-makers, prosecutors operate in complex circumstances that constantly affect their actions and decisions. The prosecutorial process has been shown to be affected by three major drivers: (1) legal factors, (2) quasi-legal factors, and (3) extra-legal factors. Legal factors include strength of evidence and seriousness of offense; quasi-legal factors include defendant-victim relationship and age; and, lastly, extra-legal factors include legally impermissible factors such as defendant gender, ethnicity, and race.

Defendant demographic characteristics such as race, ethnicity, and gender have received significant research attention. For example, Albonetti (1986) found that prosecutors were generally more likely to file charges against men than women. Bias or prejudice based on extra-legal factors should be investigated and addressed. For example, prosecutors should not consider personal or political considerations in excercising prosecutorial discretion. A prosecutor should strive to eliminate implicit biases and act to mitigate any improper bias or prejudice when credibly informed that it exists within the scope of the prosecutor's authority (ABA, 2017).

With improved record-keeping and data analysis, district attorney (DA) offices can more easily identify approaches that can improve public safety, reduce recidivism, enhance diversion and treatment programs, and allow for greater transparency and fairness in the criminal justice system. We partner with the Litmus Group, at the NYU Marron Institute, to advance the Data for Justice (D4J) initiative and focus our efforts on the District Attorney's Office of Fort Bend County, Texas.[1]

We propose the development of an interactive visualization tool aimed to facilitate the DA Office's exploration of their own data. The provided data infrastructure will give prosecutors the

[1] https://marroninstitute.nyu.edu/programs/litmus

ability to highlight disparities and injustices where they exist and the tools to assess their responses to revealed problems.

2 Related Work

With the growing concerns about bias and disparities in the criminal justice system, advocacy groups are increasingly scrutinizing prosecutor's offices. As a result, prosecutor's offices around the country are developing tools such as interactive analytics dashboards to monitor their progress towards data driven decision-making. The NYU Marron Institute, for example, has worked with the Winnebago County, Wisconsin, District Attorney's Office to create the PROTECT dashboard with similar transparency goals. Publicly available dashboards from Philadelphia [2], Hennepin [3], and San Francisco [4] are also online. These dashboards include analytics on incidents, arrests, charges, bail, etc. and help to identify and reduce disparities within the system.

3 Problem and Approach

3.1 Task

The overarching mission of D4J is to promote fairness and equity within the criminal justice system. Our team aims to help prosecutors analyze their data to investigate potential sources of bias and the processes which they can then use to take corrective action if necessary. We provide two major deliverables that will expedite these goals for the District Attorney's Office of Fort Bend County, Texas: an interactive analytics dashboard and a bias report.

Note that our proposed solutions do not focus on a machine learning or deep learning approach for two reasons:

1. **Interpretability**: Our solution needs to be easily interpretable to people without a data science background.
2. **Scope:** Machine learning and deep learning approaches are well suited to problems of prediction. However, the scope of our problem in examining bias is more strongly related to statistical inference to address the question, "Is demographic information related to disparate outcomes?"

3.2 Methodology

The interactive analytics dashboard will serve as a tool to visualize and investigate disparities in the criminal justice system. The secondary purpose of the dashboard is to satisfy DA office reporting needs. The dashboard will act as a one-stop shop for case metrics and graphical representations that the DA office needs to report on a monthly basis.

The bias report is a formal document that will outline the detailed exploratory data analysis we perform on the data provided. We focus on the three major points of decision-making within the prosecutorial timeline and investigate any presence of bias during each point separately. Our approach to the bias report can be divided into three major sections:

1. **Deep Dive Data Analysis:** Through data visualizations and exploratory data analysis on various variables of interest, we are able to formulate reasonable data pre-processing choices and set up the framework to investigate biases within the prosecutorial timeline. We isolate three major points of decision-making vulnerable to biases: charge decision, case outcome, and sentencing.
2. **Logistic Regression:** We ran logistic regressions for each point of decision-making and provide an overview of any potential biases. This helps us formulate hypothesis and identify areas that need further investigation.

[2] https://data.philadao.com/
[3] https://www.hennepinattorney.org/about/dashboard/data-dashboard
[4] https://www.sfdistrictattorney.org/policy/justice-dashboard/

3. **difference-in-proportions Hypothesis Testing:** To dive deeper into the results of the logistic regression, we then ran hypothesis tests for each point of decision-making. The results of hypothesis testing help explain and enrich observations from logistic regression tests.

With the dashboard and bias report, we hope to provide the DA's office at Fort Bend County valuable insights and practical tools to promote fairness and transparency. In turn, we hope to help create safer communities and advance the national dialogue on best practices for local justice systems.

4 Data

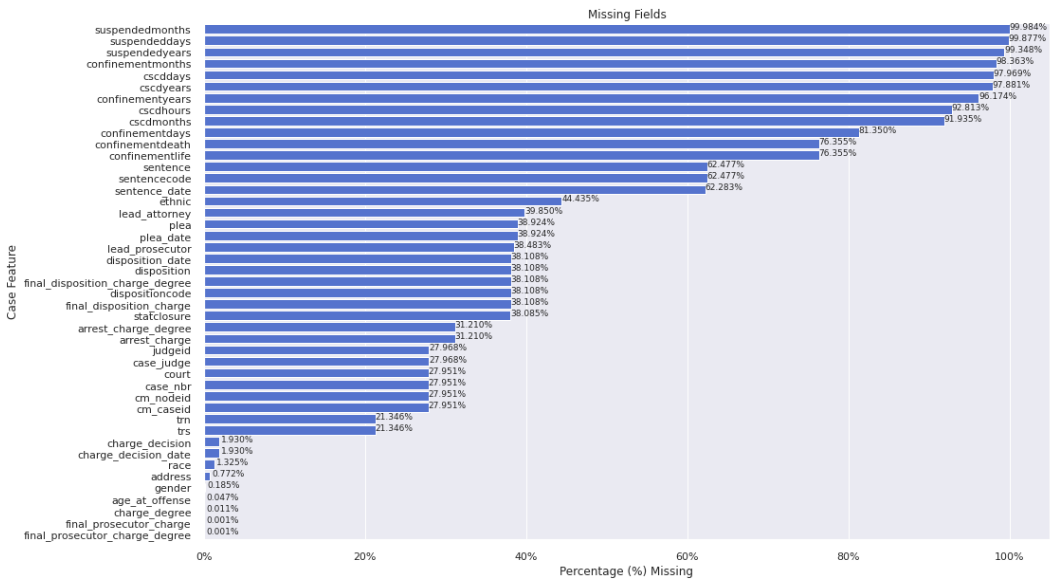


Figure 1: Case features with missing values. Approximately 80% of case features contain missing values.

We received two versions of criminal justice data sets from the Fort Bend County, Texas, District Attorney Office, one on October 15, 2020, and an updated version on November 19, 2020 with additional requested features. The new data set contains 134,791 observations and 56 features of historical prosecution data that dates back 10 years. The data set contains case features such as `race`, `gender`, `plea`, and `disposition` that record features of the defendant, crime description, charge severity, and case outcome.

In the data set provided, 45 out of 56 columns have missing feature values. Figure 2 displays the percentage of observations missing for each feature. Missing feature values provide challenges for investigating bias within the DA office. For example, around 40% of cases have ethnicity missing. This leads to problems with inference in relation to whether ethnicity being missing is random or whether its correlated with some other variable.

4.1 Data Pre-processing for the Dashboard

For the dashboard, we adopted the following process for data cleaning:

1. Binned continuous variables, e.g., age and confinement length.
2. Used a crime description to crime statute partial mapping to determine the statute based on crime description and therefore group crime descriptions together.
3. Created flags for prosecutional charge acceptance (whether to prosecute), conviction, and incarceration. These are used as dependent variables in the bias analysis.

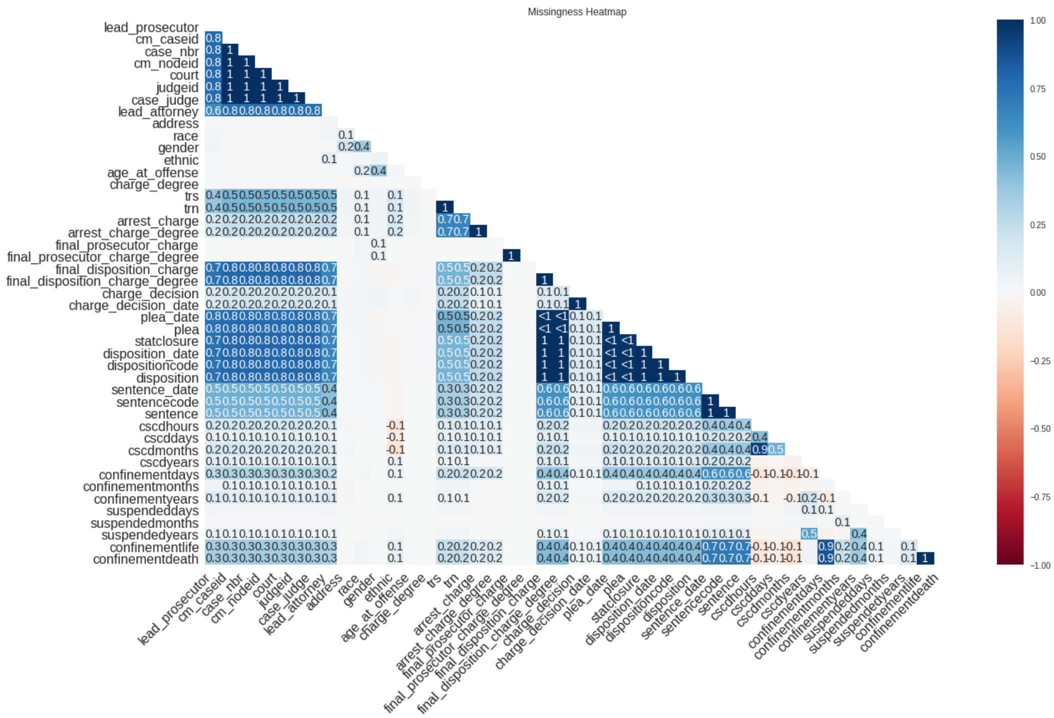


Figure 2: Heatmap of correlation of missingness between 2 columns. For example, the correlation between `case_nbr` and `cm_case_id` is 1, which means if one is present then the other one must be present. A value near -1 means if one variable appears then the other variable is very likely to be missing. A value near 0 means there is no dependence between the occurrence of missing values of two variables. A value near 1 means if one variable appears then the other variable is very likely to be present.

4. Collapsed low-frequency factors, thresholding at 1%, e.g., we collapsed lead prosecutors that appeared in fewer than 1% of the observations into an “Other Prosecutor" category.
5. Based on conversations with the District Attorney’s Office, grouped similar categories within features including charge degree, charge decision, and disposition in order to reduce the number of unique values per feature and facilitate comparisons.

4.2 Data Pre-processing for the bias report

For the bias report, we are specifically interested in six columns: `gender`, `race`, `ethnic`, `charge_decision`, `disposition` and `sentence`.

4.2.1 Independent variables

Gender Our dataset contains three options for `gender`: “Male," “Female," or “Missing."" There are 249 instances of missing values for `gender`. Since the proportion for missing gender data is not significant, we dropped all rows with missing gender.

Race and Ethnicity Our dataset contains eight options for race and two options for `ethnic`. There are 3,840 instances of missing values for `race` and more notably, 59,899 instances of missing values for ethnicity. Since only a small percentage of race data is missing, we dropped all rows with missing race data. Then, we combine the race and ethnicity columns to create one summary column: `race_and_ethnic`.

Since we have around 44% of missing ethnicity information, we replace missing data with a “Missing" tag in addition to the following imputation scheme: we combine “Black and Missing" with “Black and Non Hispanic," as well as “Asian and Missing" with “Asian and Non Hispanic." We justify this by noting that the amount of “Black and Non Hispanic" is much larger than “Black and Hispanic" (29299 vs 401), and the amount of “Asian and Non Hispanic" is much larger than “Asian and Hispanic" (2108 vs 24). Notably, we do not do this with “White and Missing" and “White and Non Hispanic". This is consistent with the demographics of Fort Bend County U.S Census Bureau (2019).

After cleaning, we end up with five main population groups in `race_and_ethnic`, which accounts for 94.4% of the total instances: “Black and Non Hispanic," “White and Missing," “White and Hispanic," “White and Non Hispanic," and “Asian and Non Hispanic."

Crime Type Cases vary drastically depending on the type of crime charged, so we include crime type as a dependent variable for our logistic regression, and control for it in the difference-in-proportions hypothesis tests. Here, we specifically examine the seven most-common crimes in our data set: “possession of marijuana $<$ 2oz" (10.06%), “assault causes bodily injury family violence" (7.61%), “driving while intoxicated" (4.42%), “driving without license with prev conviction/suspension/without insurance" (2.85%), “property theft between $50 and $500" (2.61%), “property theft between $100 < $750" (2.32%), and “assault causing bodily injury" (1.98%).

4.2.2 Dependent variables

We investigate the potential for gender, race, and ethnicity bias at three points of decision-making, charge decision, case outcome, and sentencing severity. For logistic regression we converted all three dependent variables to be binary: `charge_decision` is summarized as either accepted or not; `disposition` as either convicted or dismissed; and `sentence` as either jail or no_jail. For difference-in-proportions hypothesis tests, we treated the variables similarly except for `disposition`, where we define three outcomes: convicted, dismissed, and deferred adjudication.

5 Dashboard

The dashboard we implemented is built through Shiny, a package that builds interactive web apps directly from R. We chose Shiny because the platform allows highly interactive visualizations and allows us to incorporate statistical tools available through R. Additionally, the Marron Institute has previous experience with building Shiny dashboards for other District Attorney offices across the country, which includes institutional experience and code to modify the dashboard.

The dashboard was created to help prosecutors answer their own questions about the types of cases passing through the office and allows them to investigate implicit bias in decision-making. For example, the dashboard can be used to view conviction rates by race for the top-seven crime descriptions.[5]

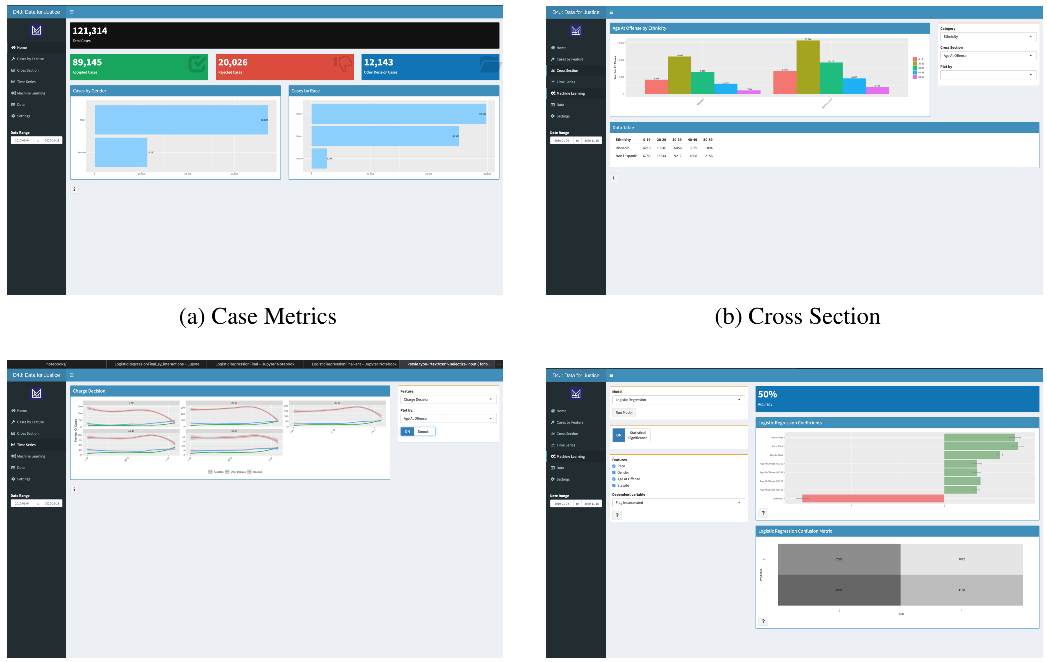
The dashboard will act as an integrated business intelligence tool that provides interactive data visualizations that display metrics, which will streamline the prosecutorial reporting process and allow for comprehensive analyses for Fort Bend County.

Interactivity Although the Fort Bend County office exhibits relatively advanced case-management systems compared to many other offices, it still employs static reporting wherein each data request must be manually processed by the IT department. The interactive dashboard gives decision-makers the flexibility to slice their data to the views they want to see and speed through Shiny’s interactivity. We allow end-users to restrict their analyses to specific time ranges and drill down by case features to different perspectives and more detail.

Case Metrics Based on conversations with the District Attorney’s Office, we report specific case metrics related to the DA office’s monthly reporting. We build these metrics directly into the dashboard. Additionally, we present cases by gender and cases by race on the home page to call attention to potential trends in cases by demographics. Figure 3a displays the home page of the dashboard that exhibits these metrics.

We allow the user to slice the data across up to three different features in the “Cross Section" tab to view interactions between features.

[5] In the cross section tab, filter “Feature" to `Race`, “Cross Section" to `Disposition`, and “Plot by" to `Crime Description`. In the Settings tab, set the threshold on max number of categories to display per feature to seven.



(a) Case Metrics

(b) Cross Section

(c) Time Series

(d) Machine Learning

Figure 3: In (3a) we display essential case metrics and visualizations necessary for monthly reporting. In (3b) we display cross section visualizations. In (3c) we display case counts over time for various features. In (3d) we display the dashboard's capabilities for interpretable machine learning. Images have been zoomed out to show the entirety of each tab's contents without scrolling. The included images are high resolution to allow reading specific text by zooming in.

Time-Series Analysis Visualizing case trends over time, for our purposes, are important to:

(**a**) Study trends in bias.

(**b**) Determine if mandated policies have been effective.

Historical data are necessary to explore factors influencing a prosecutor's decision to proceed with charges or not against a person arrested for an offense. Our dashboard visualizes trends in case counts over time broken down by a selection of case features. We also add an option to toggle locally weighted smoothing that creates an approximate function to capture important patterns in the data and ignore day-to-day noise.

Machine Learning We implement interpretable machine learning within our dashboard. Kim et al. (2016) defines interpretability as the degree to which a human can consistently predict the model's result. As discussed in Section 2, our dashboard solution does not focus on sophisticated deep learning or machine learning framework and instead we focus on interpretable results and inference over prediction.

Incorporating interpretable machine learning was not included within the scope of our project. However, we saw an opportunity to use machine learning for interactive inference. The machine learning tab allows the user flexibility to select:

1. **Model type:** Logistic regression and decision tree algorithms are currently implemented as they are common interpretable models, but the framework is there to include more models as needed. These models were chosen because of their ability to handle binary target variables as well as their interpretability.
2. **Model options:** For logistic regression, there is the capability to display or hide 95% confidence intervals for the feature coefficients. For decision tree, the option to tune max tree depth is available.

3. **Features:** Features are used as independent variables (predictors) in the machine learning predictive task. These are categorical features (e.g.,, `race`, `gender`, `age_at_offense`, and `statute`) that are one-hot-encoded. The features `race`, `gender`, and `age` are demographic characteristics, while `statute` is included to control for crime type.
4. **Dependent variable:** Choose to predict binary indicators for charge acceptance, conviction, or incarceration as discussed in Section 4.

We follow a conventional 80-20 training-test split and display model coefficients, test accuracy, and a confusion matrix for predictions in the test set.

Unlike a typical machine learning problem where higher accuracy is the goal, in this case, we are more interested in the statistical significance of the coefficients. Highly significant coefficients could indicate that demographic information plays a role in determining outcomes, which may signify bias. In fact, a low test accuracy may be desirable because it signifies that demographic information cannot be used to accurately predict outcomes.

We made this tab interactive instead of reporting only our best models so that the District Attorney's Office gains experience and familiarity with machine learning tools through having access on the fly. The framework also allows them to create their own dependent variables, incorporate other demographic information, or run the models over specific time periods.

6 Bias Report

In our bias report, we focus on outcome differences at three major points of decision-making in the prosecutorial pipeline: decision on whether or not to charge, outcome of the case, and sentencing. We examine the impact of gender, race, and ethnicity at these three points in the pipeline. We examine the effect only on the seven most common crimes in our data set.

We approach the bias report as follows. We first provide an exploratory data analysis on some of the variables in question, and justify the data pre-processing choices. We then investigate any potential bias at the aforementioned three points of decision-making through two approaches: using hypothesis tests and odds ratios from logistic regression, and hypothesis tests from difference-in-proportions hypothesis tests.

6.1 Logistic Regression

We constructed three separate logistic regressions using `race_and_ethnic`, `gender`, and `crime_type`, attempting to predict `charge_decision`, `disposition`, and `sentence`. We investigated statistical significance for `race_and_ethnic` and `gender`, but not for `crime_type`. This is because we expect that the crime type will largely influence the outcome of `charge_decision`, `disposition`, and `sentence`, and it is not unreasonable for it do so.

In each of the regressions, we will analyze the statistical significance of the normalized coefficient. The normalized coefficient is calculated by subtracting the coefficient from the average of all the coefficients of one group of independent variables (e.g `race_and_ethnic`). We then report a z-score on the significance of the coefficient, which measures how confident we are of its deviation from the average coefficient in the group. We also report an odds ratio on the normalized coefficient, which can be interpreted as the chance of the positive outcome (being charged, convicted, or sentenced for jail).

Our results show evidence for large amounts of bias, or at the very least, differences in group outcome, especially for sentencing.

6.1.1 Charge Decision

Table 1: Logistic regression coefficients for predicting `charge_decision`

Feature Names	Coefficients	Coefficients Normalized	P-value	Odds Ratio
Asian and Non Hispanic	0.146	-0.017	7.14e-01	0.995
Black and Non Hispanic	0.107	-0.056	1.53e-02	0.983
White and Hispanic	0.257	0.094	2.43e-03	1.028
White and Missing	0.061	-0.102	3.22e-05	0.968
White and Non Hispanic	0.245	0.081	4.09e-02	1.025
Female	0.238	-0.170	2.92e-25	0.946
Male	0.578	0.170	1.42e-30	1.050
AGG ASSAULT W/DEADLY WEAPON	-1.244	nan	nan	nan
ASSAULT CAUSES BODILY INJURY FAMILY VIOLENCE	-1.529	nan	nan	nan
DRIVING W/LIC INV W/PREV CONV/SUSP/W/O FIN RES	1.331	nan	nan	nan
DRIVING WHILE INTOXICATED	0.836	nan	nan	nan
POSS MARIJ <2OZ	0.282	nan	nan	nan
THEFT PROP >=$100<$750	0.397	nan	nan	nan
THEFT PROP>=$50<$500	0.743	nan	nan	nan

We include statistical analysis for `crime_type` as it justifiably influences `charge_decision`.

Our results for the logistic regression predicting `charge_decision` are in Table 1. For `charge_decision`, a positive coefficient indicates that the charge decision is more likely to be accepted, while a negative coefficient indicates that the charge decision is more likely to be rejected. We see some statistical significance for the normalized coefficients, though the magnitude in the shifts of odds ratio is relatively minor. We decline to conclude any bias for this stage because the shifts in odds ratios are minor and are eclipsed by our results for `disposition` and `sentence`.

6.1.2 Disposition

Table 2: Logistic regression coefficients for predicting `disposition`.

Feature Names	Coefficients	Coefficients Normalized	P-value	Odds Ratio
Asian and Non Hispanic	-0.600	-0.581	3.07e-27	0.708
Black and Non Hispanic	0.200	0.219	2.77e-18	1.115
White and Hispanic	0.304	0.323	5.22e-22	1.169
White and Missing	-0.011	0.008	7.80e-01	1.004
White and Non Hispanic	0.012	0.031	4.63e-01	1.016
Female	-0.271	-0.223	4.99e-26	0.884
Male	0.175	0.223	1.53e-39	1.117
AGG ASSAULT W/DEADLY WEAPON	-0.789	nan	nan	nan
ASSAULT CAUSES BODILY INJURY FAMILY VIOLENCE	-0.579	nan	nan	nan
DRIVING W/LIC INV W/PREV CONV/SUSP/W/O FIN RES	0.564	nan	nan	nan
DRIVING WHILE INTOXICATED	1.219	nan	nan	nan
POSS MARIJ <2OZ	-0.548	nan	nan	nan
THEFT PROP >=$100<$750	-0.507	nan	nan	nan
THEFT PROP>=$50<$500	0.545	nan	nan	nan

We include statistical analysis for `crime_type` as it justifiably influences `disposition`.

Our results for the logistic regression predicting `disposition` are in Table 2. For `disposition`, a positive coefficient indicates that the disposition is more likely to be a conviction, while a negative coefficient indicates that the disposition is less likely to be a conviction. Here, we see a high statistical significance for the coefficients and moderate difference in the odds ratios. Importantly, we see that for `race_and_ethnic`, odds ratios range from .708 to 1.169, and for `gender`, males are 1.26x more likely to be convicted than females.

6.1.3 Sentencing

Table 3: Logistic regression coefficients for predicting `sentence`

Feature Names	Coefficients	Coefficients Normalized	P-value	Odds Ratio
Asian and Non Hispanic	-0.936	-0.872	3.08e-37	0.554
Black and Non Hispanic	0.228	0.292	7.20e-25	1.172
White and Hispanic	0.622	0.687	1.18e-80	1.405
White and Missing	-0.640	-0.575	1.79e-61	0.689
White and Non Hispanic	0.404	0.468	3.10e-23	1.277
Female	-0.557	-0.395	4.14e-57	0.780
Male	0.234	0.395	3.10e-96	1.234
AGG ASSAULT W/DEADLY WEAPON	0.050	nan	nan	nan
ASSAULT CAUSES BODILY INJURY FAMILY VIOLENCE	0.042	nan	nan	nan
DRIVING W/LIC INV W/PREV CONV/SUSP/W/O FIN RES	-0.134	nan	nan	nan
DRIVING WHILE INTOXICATED	-0.314	nan	nan	nan
POSS MARIJ <2OZ	-0.183	nan	nan	nan
THEFT PROP >=$100<$750	0.009	nan	nan	nan
THEFT PROP>=$50<$500	0.206	nan	nan	nan

We include statistical analysis for `crime_type` as it justifiably influences `sentence`.

Our results for the logistic regression predicting `sentence` are in Table 3. For `sentence` a positive coefficient indicates that the sentence is more likely to include jail time, while a negative coefficient indicates that the sentence is less likely to be include jail time. Here, we see high statistical significance for the coefficients, and a high difference in the odds ratios. Importantly, we see that for `race_and_ethnic` the odds ratios range from .55 to 1.4, implying that the defendant's race and ethnicity has a large impact on whether their sentence will include jail time. In addition, for `gender`, we see that males are 1.57x more likely to be sentenced to jail than females.

6.1.4 Remarks

In our analysis for the logistic regression, we showed large group differences that are also statistically significant. One might be tempted to conclude that "small" amounts of disparity exists at every point in the prosecutorial timeline. *However, as we show in the next section, large amounts of disparity seem to appear only at certain points of the prosecutorial timeline, dependent on the crime type.*

6.2 P-value from Difference-in-Proportions Hypothesis Test

We constructed three difference-in-proportion hypothesis tests for each of seven crimes. Our difference-in-proportion tests are the following: filed for crime vs charged for the crime, charged for the crime vs convicted for the crime, convicted for the crime vs sentenced to jail. These three tests provide a reasonable view of how group differences evolve during the prosecutorial process. Furthermore, we show that for different crimes, potential bias is evident at different points in the prosecutorial process.

We found evidence of bias for three different crimes: "possession of marijuana $<$ 2 oz," "assault causes bodily injury family violence," and "driving while intoxicated." We provide nine tables (three tables for each of the aforementioned crimes) for the difference-in-proportions hypothesis test in this report. *We conducted similar analyses for the other four crimes, but did not find evidence of bias in any of the tests, so we omit including those tables.*

6.2.1 Possession of Marijuana < 2 oz

Table 4: difference-in-proportions Hypothesis Test: Crime Filed vs Crime Charged

Race and Ethnicity	Crime Proportions	Charged Crime Proportions	P-value
Black and Non Hispanic	48.9%	49.6%	3.17e-01
White and Missing	21.0%	22.8%	9.46e-04
White and Hispanic	17.3%	15.4%	1.40e-04
White and Non Hispanic	8.2%	7.5%	7.96e-02
Asian and Non Hispanic	4.6%	4.6%	9.21e-01

Crime: Possession of Marijuana < 2 oz. Total Filed w/Crime: 12811. Total Charged w/Crime 9787.

Table 5: difference-in-proportions Hypothesis Test: Crime Charged vs Dispositions

Race and Ethnicity	Charged Crime Prop	Dismissed Prop	P-value Dismissed	Convicted Prop	P-value Convicted	Deferred Prop	P-value Deferred
Black and Non Hispanic	49.6%	49.6%	9.68e-01	54.9%	8.36e-07	41.8%	4.21e-11
White and Missing	22.8%	21.3%	4.16e-02	17.5%	1.59e-09	33.0%	1.89e-23
White and Hispanic	15.4%	15.5%	8.35e-01	17.7%	2.95e-03	12.1%	9.61e-05
White and Non Hispanic	7.5%	8.2%	1.89e-01	7.6%	8.86e-01	6.6%	1.50e-01
Asian and Non Hispanic	4.6%	5.4%	6.17e-02	2.3%	2.79e-08	6.5%	4.99e-04

Crime: Possession of Marijuana < 2 oz. Total Charged w/Crime 9787. Total Charged and Dismissed 4484. Total Charged and Convicted 2772. Total Charged and Deferred 2155.

Table 6: difference-in-proportions Hypothesis Test: Convicted vs Jail

Race and Ethnicity	Convicted Proportions	Jail Proportions	P-value
Black and Non Hispanic	54.9%	55.7%	5.65e-01
White and Missing	17.5%	15.7%	1.07e-01
White and Hispanic	17.7%	19.0%	2.78e-01
White and Non Hispanic	7.6%	7.5%	9.25e-01
Asian and Non Hispanic	2.3%	2.0%	5.88e-01

Crime: Possession of Marijuana < 2 oz. Total Charged and Convicted 2772. Total Convicted and Jailed 2056.

Inspecting tables 4, 5, and 6 reveals that there are not many group differences in the decision of whether to charge, large differences in the disposition (especially for convicted and deferred adjudication), and not much in the sentencing. The largest discrepencies are for “White and Missing" for receiving deferred adjudication (from 22.8% to 33.0%), and “White and Missing" for being convicted (from 22.8% to 17.5%).

6.2.2 Assault Causes Bodily Injury Family Violence

Table 7: difference-in-proportions Hypothesis Test: Crime Filed vs Crime Charged

Race and Ethnicity	Crime Proportions	Charged Crime Proportions	P-value
Black and Non Hispanic	40.5%	38.5%	2.87e-02
White and Missing	23.3%	19.2%	5.51e-08
White and Hispanic	18.1%	22.6%	6.21e-10
White and Non Hispanic	12.1%	13.7%	8.79e-03
Asian and Non Hispanic	6.0%	6.0%	9.76e-01

Crime: Assault Causes Bodily Injury Family Violence. Total Filed w/Crime: 9686. Total Charged w/Crime 412.

Table 8: difference-in-proportions Hypothesis Test: Crime Charged vs Dispositions

Race and Ethnicity	Charged Crime Prop	Dismissed Prop	P-value Dismissed	Convicted Prop	P-value Convicted	Deferred Prop	P-value Deferred
Black and Non Hispanic	38.5%	36.8%	2.27e-01	41.0%	1.67e-01	35.5%	1.66e-01
White and Missing	19.2%	21.7%	2.76e-02	14.7%	1.99e-03	23.8%	1.07e-02
White and Hispanic	22.6%	18.3%	2.52e-04	30.4%	6.91e-07	21.4%	5.25e-01
White and Non Hispanic	13.7%	15.7%	4.57e-02	11.4%	6.49e-02	13.5%	9.11e-01
Asian and Non Hispanic	6.0%	7.5%	3.85e-02	2.5%	1.83e-05	5.9%	8.74e-01

Crime: Assault Causes Bodily Injury Family Violence. Total Charged w/Crime 4218. Total Charged and Dismissed 1726. Total Charged and Convicted 895. Total Charged and Deferred 547.

Table 9: difference-in-proportions Hypothesis Test: Convicted vs Jail

Race and Ethnicity	Convicted Proportions	Jail Proportions	P-value
Black and Non Hispanic	41.0%	42.1%	6.42e-01
White and Missing	14.7%	13.3%	3.85e-01
White and Hispanic	30.4%	32.0%	4.83e-01
White and Non Hispanic	11.4%	10.3%	4.63e-01
Asian and Non Hispanic	2.5%	2.3%	8.76e-01

Crime: Assault Causes Bodily Injury Family Violence. Total Charged and Convicted 895. Total Convicted and Jailed 769.

Inspecting tables 7, 8, and 9 reveals that there is not much group differences in disposition or sentencing, but large differences in charge decision. The discrepancies are for the `race_and_ethnic` variable for “White and Missing" (from 23.3% to 19.2 %) and “White and Hispanic" (18.1% to 22.6 %).

6.2.3 Driving While Intoxicated

Table 10: difference-in-proportions Hypothesis Test: Crime Filed vs Crime Charged

Race and Ethnicity	Crime Proportions	Charged Crime Proportions	P-value
Black and Non Hispanic	20.2%	20.1%	8.93e-01
White and Missing	45.1%	45.6%	6.35e-01
White and Hispanic	18.0%	18.1%	8.41e-01
White and Non Hispanic	9.7%	9.5%	7.11e-01
Asian and Non Hispanic	7.0%	6.7%	5.52e-01

Crime: Driving While Intoxicated. Total Filed w/ Crime: 5624. Total Charged w/Crime 4887

Table 11: difference-in-proportions Hypothesis Test: Crime Charged vs Dispositions

Race and Ethnicity	Charged Crime Prop	Dismissed Prop	P-value Dismissed	Convicted Prop	P-value Convicted	Deferred Prop	P-value Deferred
Black and Non Hispanic	20.1%	19.3%	5.72e-01	19.7%	6.72e-01	28.6%	5.76e-01
White and Missing	45.6%	48.9%	6.41e-02	48.2%	2.27e-02	28.6%	3.66e-01
White and Hispanic	18.1%	13.4%	5.46e-04	17.9%	8.28e-01	14.3%	7.93e-01
White and Non Hispanic	9.5%	8.4%	2.90e-01	8.5%	1.31e-01	14.3%	6.66e-01
Asian and Non Hispanic	6.7%	10.0%	3.82e-04	5.7%	6.54e-02	14.3%	4.24e-01

Crime: Driving While Intoxicated. Total Charged w/Crime 4887. Total Charged and Dismissed 918. Total Charged and Convicted 2914. Total Charged and Deferred 7

Table 12: difference-in-proportions Hypothesis Test: Convicted vs Jail

Race and Ethnicity	Convicted Proportions	Jail Proportions	P-value
Black and Non Hispanic	19.7%	22.4%	5.86e-02
White and Missing	48.2%	34.8%	1.32e-14
White and Hispanic	17.9%	27.4%	2.00e-11
White and Non Hispanic	8.5%	11.2%	6.68e-03
Asian and Non Hispanic	5.7%	4.2%	5.36e-02

Crime: Driving While Intoxicated. Total Charged and Convicted 2914. Total Convicted and Jailed 1131.

Inspecting tables 10, 11, and 12 reveals that there are not many group differences in charge decision or disposition, but large differences in sentencing. The largest discrepancies for the `race_and_ethnic` variable are: “White and Hispanic” make up 17.9% of those that were convicted for driving while intoxicated yet only 27.4% of those sentenced to jail. “White and Missing" makes up 48.2% of those that were convicted for driving while intoxicated yet only 34.8% of those sentenced to jail.

6.2.4 Remarks

Our analysis here shows that for most crimes and at most points in the prosecutorial decision-making process, we could not find statistically significant bias. However, we show evidence for large potential bias at a few decision-making points. For possession of marijuana < 2 oz, we found large and statistically significant discrepancies in disposition. For assault causes bodily injury family violence, we found large and statistically significant discrepancies in charge decision. Lastly, for driving while intoxicated, we found large and statistically significant discrepancies in sentencing.

6.3 Discussion

We acknowledge the complexity of bias in the criminal justice system and make note of the following points.

First, the scope of our analysis is limited to only the decisions made in the District Attorney's office. Despite observing discrepancies between the actual population distribution versus the one in our data set, we do not have access to any details related to arrests or the referring law enforcement agencies. In our analysis, we do not observe sufficient evidence to conclude bias across charge decisions. However, it is worth reflecting whether biases are propagated. Future research including data from law enforcement agencies will enrich the understanding of potential bias in the criminal justice system.

Second, we acknowledge the complex and intersectional nature of race and class. For example, we found evidence of possible racial bias in sentencing, but the disparity may be caused by an unobserved factor correlated with race such as difference in access to resources between different communities. In approaching the topic of bias in the criminal justice system, it is paramount to take varying factors into considerations.

Third, we note how our analysis is limited by our data. On the face of it, our results here support the possibility of bias in the criminal justice system. However, the large amount of missing ethnicity data (44%) affected the rigor and comprehensiveness of our analysis. We ran logistic regression and difference-in-proportion tests twice–before and after imputing missing ethnicity data. The results varied between these two sets of experiments. We observed slightly different patterns of potential bias in both crime types and in the three points of decision-making. To promote justice and transparency in the criminal justice system, change could start from something seemingly as small as improving record keeping.

Finally, we acknowledge the possibility of confounding variables that could explain some of the differences in group outcome. One variable that is highly relevant, but we did not have direct access to, is whether or not the defendant was a repeat offender[6]. Our data set spanned ten years, and the timeline for a repeat driving while intoxicated offense is also five years. Another possibility is the existence of county sponsored programs or certain laws that target specific crimes which could explain the large amount of observed bias.

7 Conclusion

We started this project with the goal of helping prosecutors' offices leverage their data. The first part of our solution was a dashboard that allows flexible, instant visualizations and an on-demand, deep dive into their own data. We also conducted a bias report which discovered large amounts of disparate outcomes exist at certain, limited points in the prosecutorial decision-making timeline. We hope that this leads to further investigation as to possible causes of these disparities. Finally, we hope that similar efforts are made for prosecutor's offices in other counties – criminal justice is a contentious topic in 2020, and efforts such as these will lead to greater transparency and trust in the criminal justice system.

8 Student Contributions

Ant: Developed dashboard, final paper writing
Andrew: Developed dashboard, developed statistical methodology, final paper writing
Christine: Wrote bias report, final paper writing
Alex: Developed statistical methodology, wrote bias report, final paper writing

[6] The data does include a unique ID per defendant which would allow this analysis in the future.

References

ABA. 2017. Prosecution Function. *CRIMINAL JUSTICE STANDARDS for the PROSECUTION FUNCTION*, 4.

Celesta A. Albonetti. 1986. Criminality, prosecutorial screening, and uncertainty: toward a theory of discretionary decision making in felony case processing.

Bruce Frederick and Don Stemen. 2012. The anatomy of discretion: An analysis of prosecutorial decision making – technical report.

Been Kim, Rajiv Khanna, and Oluwasanmi O Koyejo. 2016. Examples are not enough, learn to criticize! criticism for interpretability. In *Advances in Neural Information Processing Systems*, volume 29, pages 2280–2288. Curran Associates, Inc.

U.S Census Bureau. 2019. U.s. Census Bureau QuickFacts: Fort Bend County, Texas.