

**University of Washington**

# **Predicting Racial Bias Tendencies in Police Stops**

Mahir Bathija, Taylor Merry, Hyewon Kim, Christopher Painton  
INFO 370 (Autumn 2018)  
12/7/2018

## Background

When police officer Darren Wilson fatally shot Michael Brown on August 9th, 2014, a nationwide debate was reborn centered around racial tensions between minority citizens and police officers. Waves of protest and outrage followed demanding justice, transparency, and change from our police forces. The aftermath of this incident and those similar has police departments expressing their eagerness to change perceptions of their racial bias and promises of removing discriminatory culture within their policing practices. However, we continuously see headline news focusing on the inequity of treatment towards minorities. As a result, there has been a steady rise of racial tensions between minorities and police officers.

The Stanford Open Policing Project, our primary data source, has conducted their own analysis into racial discrepancies within police traffic and pedestrian stops. Their team discovered that, while controlling for age and gender, officers stop blacks and Hispanics at higher rates than white drivers<sup>1</sup>. Additionally, officers were more likely to ticket, search, and arrest these black and Hispanic drivers<sup>2</sup>. In an effort to promote transparency, the study recognized that additional outside factors might remove the suspicion of racial bias in policing decisions. However, to test for discrimination in searches, their team built a threshold test (detailed below as our second source) that discovered police officers required less suspicion to search black and Hispanic drivers illustrating this “double standard that indicates discrimination”<sup>2</sup>. Along with their own research, the Stanford Open Policing Project has dedicated a section on their website to additional sources on this issue. Below we’ve listed some relevant research but note that these only make up a small scope of the publications investigating this issue.

### (1) Are Traffic Stops Prone to Racial Bias?

- a. This article details information about the strategies that states implement for collecting police stop data. While the article doesn’t focus on detailed research findings, it helps explain the origin of the Stanford Open Policing Project’s dataset. Only 31 states have formed some sort of protocol for collecting traffic stop information, however, the data collection across these states isn’t consistent. While many police departments do have stop information, they do not gather driver race information due to their policies against racial bias. In the context of our project, Washington State Patrol collects and interprets their data internally via random auditing twice a year, in efforts to fight racial discrepancies.

Ramachandran, Vignesh, et al. “Are Traffic Stops Prone to Racial Bias” (2016). The Marshall Project <https://www.themarshallproject.org/2016/06/21/are-traffic-stops-Prone-to-racial-bias>

---

<sup>1</sup> “Stop Rates.” Stanford Open Policing Project. <https://openpolicing.stanford.edu/findings/> (accessed October 19th, 2018)

<sup>2</sup> “After the Stop.” Stanford Open Policing Project. <https://openpolicing.stanford.edu/findings/> (accessed October 19th, 2018)

## **(2) Fast Threshold Tests for Detecting Discrimination**

- a. This study focused on a threshold test that was 75 times faster than existing models for detecting bias in police decisions. Their new modified test allows for quicker application and is much more capable of handling larger datasets. The data used for detecting discrimination in this specific study comes from New York City's Stop and Frisk Policy, described in the study below. This test was applied to over 700,000 traffic stops and "analyzed two specific decisions: the initial stop decision and the subsequent of whether or not to conduct a frisk"<sup>3</sup>. An important note was this assumption that the demographic breakdown of the communities reflected the breakdown of pedestrians that police officers interacted with while on the streets. As a result, they discovered that the stop thresholds for blacks and Hispanics are lower than whites, indicating a pattern of discrimination. These results match conclusions from previous research, while accounting for "potential infra-marginality".

Pierson, Emma, et al. "Fast Threshold Tests for Detecting Discrimination" (2017). Proceedings of Machine Learning. <https://5harad.com/papers/fasttt.pdf>

## **(3) Precinct or Prejudice? Understanding Racial Disparities in New York City's Stop-and-Frisk Policy**

- a. This study investigated New York City's Stop-and-Frisk Policy, a policing tactic to stop individuals suspected of criminal possession of a weapon. This study made note that police search tactics are typically focused in high crime neighborhoods, traditionally more black and Hispanic communities. While adjusting for the high crime rate and localized searches, the study found that blacks and Hispanics were still stopped at excessive rates within "low hit rate contexts". One of the factors they discovered was "in more than 40% of cases, the likelihood of finding a weapon (typically a knife) was less than 1%"<sup>4</sup>. Another factor is that there are lower thresholds to stop minorities compared to whites, despite simulating situations of whites within these high-crime rate areas. In conclusion, the study believes that a "statistically informed stopping strategy can be approximated by simple, easily implemented heuristics with little loss in efficiency"<sup>4</sup> that would lessen racial disparities and lead to more successful weapon confiscation.

Goel, Sharad "Precinct or Prejudice? Understanding Racial Disparities in New York City's Stop-and-Frisk Policy" (2016). The Annals of Applied Statistics Vol.10 365-394. <https://5harad.com/papers/frisky.pdf>

---

<sup>3</sup> Pierson, Emma, et al. "Fast Threshold Tests for Detecting Discrimination" (2017). Proceedings of Machine Learning. <https://5harad.com/papers/fasttt.pdf> 6

<sup>4</sup> Goel, Sharad, et al. "Precinct or Prejudice? Understanding Racial Disparities in New York City's Stop-and-Frisk Policy" (2016). The Annals of Applied Statistics Vol.10 365 <https://5harad.com/papers/frisky.pdf>

## Motivation

While topics around racial bias within police policies are uncomfortable ones, we believe this discussion is an important one and feel that we have the perfect platform for sharing these ideas. This is a well-researched topic with publications across many different mediums that we wanted to help educate people on the gravity of this issue. These racial injustice claims are highly relevant in our current society that its crucial to closely examine if they hold any validity. Our project focuses primarily on data from the state of Washington, as we wanted to investigate these beliefs within our own home state. By combining police stop records with county demographics and voting history, we believe that we are conducting our own due diligence into Washington State police departments.

Although this is an evolving and complex issue, we felt this social responsibility to investigate and provide a new perspective into this long debate. This study is by no means a direct accusation of police being racist, as we understand the sensitivity behind each specific incident and situation. However, our team firmly believes in equal and consistent police practices and we want to contribute by carefully sharing our findings. The current political turmoil around this issue brings discomfort to many parties and we hope that by expanding on previous research, we will be able to provide some form of clarity.

## Research Question

Our primary research question is: *Can we predict the race of an individual from a traffic/pedestrian stop given the stop conditions, county demographics, and voting history?*

## Data Sources

### Washington Traffic Stop Records

The Stanford Open Policing Project provides traffic stop data for 31 states containing over 130 million traffic stop records. Their team is made up of journalists and academic researchers dedicated “to help researchers, journalists, and policymakers investigate and improve interactions between police and the public”<sup>5</sup>. Our team took their dataset for the state of Washington and removed unnecessary columns along with empty values leaving us with 5,343,498 observations spanning from 2009 to 2016. To help aid our model, we dummified important categorical values into binary features. These binary features include: if a certain violation was broken, was the offender a minority, is there a match in officer and offender race,

---

<sup>5</sup> The Stanford Open Policing Project. (2018, October 22). Retrieved from <https://openpolicing.stanford.edu>

and was an arrest made. We stored this information in the form of 0s (representing a no) and 1s (representing a yes). Specifically, for designating a minority, a one indicated a minority driver and a zero indicated a white driver.

### Washington Counties Voting Results

Since the traffic stop data contains the name of the county in which the stop occurred, we thought it would be interesting to join this information with county voting data. We believed that the political affiliation of each county could potentially be a factor for predicting driver race. The Secretary of State for Washington provided a website with csv files for the 2016, 2012, and 2008 presidential elections. For years outside of the election cycle, we assigned 2008 election data for stops within 2009-2010, 2012 election data for stops within 2011-2014, and 2016 election data for stops within 2015-2016. The voting results were broken down into three categories: Democratic, Republican, and other. We felt that presidential voting gave the best representation of a counties political affiliation as more people tend to vote and political preferences among communities don't drastically change.

### Washington County Race Demographics

We also wanted to include the demographic information to understand the representation of each county. The U.S. Census Quickfacts website provides this information on a national, state, and county-wide level. We scraped this information from the U.S Census website and stored all the counties' demographics into a single csv file. This file contains the county name, population of the county, racial group, and the percentage of the races made up from that county. The demographic breakdown percentage contained the following races: White, Asian, African American, Hispanic/Latino, and other. Since we only found the demographic breakdown for 2015, we will be making the assumption that the 2015 percentages is similar to that of 2009-2014 and 2016.

## Analysis

Our primary dataset is a combination of our three data sources: police stop information, demographic breakdown, and voting history. Based on a quick analysis of our dataset, our main columns of interest are: driver race, driver gender, police officer race, police officer gender, violations, search conducted, county demographic, county voting preference, and arrest/citation (Detailed in Appendix A). With the dataset prepared, we began some exploratory data analysis to best understand the information we were working with.

## Preliminary Data Analysis

Our initial findings discovered that King County had the highest numbers of police stops while Ferry County had the lowest number. When comparing the search rates (See Table 1), we found that searches were conducted on Hispanic and black offenders more often than white offenders. In addition, we discovered a trend of minority drivers being stopped during the “twilight hours” (10PM - 6AM) more often than white drivers (See Appendix B). During this period, black drivers were stopped almost twice as often than white drivers. The daytime hours (10 AM - 4PM) featured white drivers being stopped at the highest rates.

Race	Search Conducted	Search Not Conducted	Percentage of Stops with Search
Asian	6,838	333,620	2.01 %
Black	12,840	235,789	5.16 %
Hispanic	18,970	446,694	4.07 %
White	97,471	4,007,539	2.37 %
Other	4,346	72,799	5.63 %

**Table 1: Search Rate Counts and Percentages for Each Race.** Blacks (5.16 %) and Hispanics (4.07%) have the highest percentage of stops that ended with a search while Asians (2.01 %) had the lowest search percentages.

## Statistical Analysis

As a continuation of our preliminary analysis, we wanted to get a sense of the general descriptive summary (See Appendix C) of our dataset, to present the nuances and potential issues that could occur with our machine learning model. The first thing we noticed was that the most stopped driver race was white by a significant margin, making up 78.39% of the total police stops in Washington. Another interesting observation from our descriptive table is that arrest/citations occur more frequently than warnings with an average of 1.4 violations per stop. This information gave us a sense of what to be cautious for when building our classification model.

## Chi Squared Independence Tests and Expected vs Observed Breakdown

For categorical variables, specifically violations, we used chi squared tests to see if violation type and driver’s race were independent. After each test returned low p-values, we were able to indicate that driver race and each of the violations are not independent of one another. The resulting contingency tables (observed vs expected) provided additional insight into stop tendencies (See Appendix D). Table 2 below demonstrates that each of the minority groups had

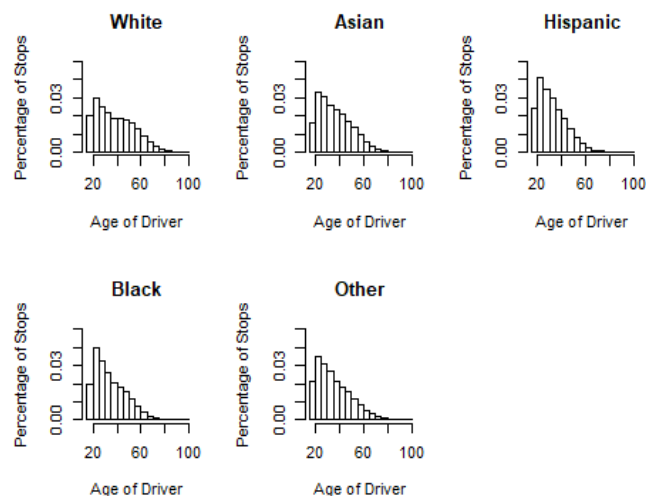
higher observed arrest/citation counts than their expected counts, while white remains as the only racial group with observed arrest/citation counts lower than their expected count.

	Arrest/Citation Observed	Arrest/Citation Expected	Percentage Change
Asian	198,652	176,229.54	+ 12.72 %
Black	139,903	128,696.56	+ 8.70 %
Hispanic	249,637	241,039.29	+ 3.57 %
Other	46,152	39,932.17	+ 15.58 %
White	2,076,409	2,124,855.43	- 2.33 %

**Table 2: Arrest/Citation Observed vs Expected Count.** Whites were the only racial group to have their counts decrease from expected to observed, as each minority group saw an increase of their expected to observed counts.

## ANOVA and Tukey Post Hoc Tests

We ran analysis of variance (ANOVA) tests on the numerical variables: driver age, time of day, number of violations, and Republican to Democrat voting ratio (See Appendix E). The p-values for each variables' ANOVA test were approximately zero (most likely due to the large sample size), meaning we rejected the null hypothesis that all of the means were equal. After the ANOVA tests, we performed Tukey Post Hoc tests to determine which races had different means. One of our most interesting findings (Figure 1 and Table 3) discovered that Hispanic drivers were 2 years younger than each of the other minorities and 6 years younger than white drivers. Black drivers were on average 4 years younger than white drivers. White drivers were 2 years older than drivers of any other race.



**Figure 1: Histogram of Age Grouped by Race.** Younger Hispanic and black drivers (< 25) have a higher percentage of stops compared to other races drivers of the same age.

Race 1 vs Race 2	Difference of Means
Hispanic vs Asian	-3.901
Hispanic vs Black	-1.989
White vs Black	4.185
White vs Hispanic	6.173
White vs Other	3.181

**Table 3: Driver Age Tukey Test Table.**

Hispanic drivers were almost 2 years younger than every other race on average along with 6 years younger than white drivers. Black drivers were on average 4 years younger than white drivers. White drivers were 2 years older than drivers of any other race.

## Primary Analysis

The final step in our research process was to use the scikit learn implementation of XGBoost for answering our research question. XGBoost is a gradient boosting framework designed to be highly efficient and gradient boosting is an ensemble method. Ensemble methods combine the predictions of several models to improve performance<sup>6</sup>. There are two types of ensemble methods: bagging methods and boosting methods. Boosting is when the model starts with weak learners (simple decision trees) and iteratively combines them while reducing bias/error to create a strong learner. We specifically chose to use XGBoost because it is known to be extremely accurate<sup>7</sup>. We wanted to prioritize accuracy over implementation time and explain ability because we didn't plan on utilizing our model for real time predictions and aren't trying to explain our algorithm to non-technical shareholders.

The biggest challenge for implementing our classification model was with the disproportionate amount of whites (~78% of the stops) to the minority races. To adjust for the unbalanced dataset, we decided to randomly under sample the white training data to match the total count of minority drivers. The testing data was left unaltered, so that our results would be reported on something that more closely matched our original population sample. Another design decision was to use K-fold cross validation, but within the loop we under sampled only the training data as we didn't want our validation sets to present results that misrepresented our data. This K-fold cross validation ( $k = 5$ ) was used for tuning learning rate, number of estimators, and the scale position weight parameters. The final decision was to use binary classification as we felt this was easier to predict than multi classification and we already had a "Minority Driver" column prepared as the target variable.

For tuning the other hyperparameters, we used a grid search to iterate through different hyperparameter combinations. The best performing combinations (based on the best AUROC) was then implemented into the next grid search for finding the other hyperparameters. The following hyperparameters were tuned using this strategy: max depth (maximum depth of a tree), min\_child\_weight (minimum sum of weights of all observations), subsample (fraction of observations randomly sampled for each tree), colsample\_bytree (fraction of columns randomly sampled for each tree), gamma, and alpha.

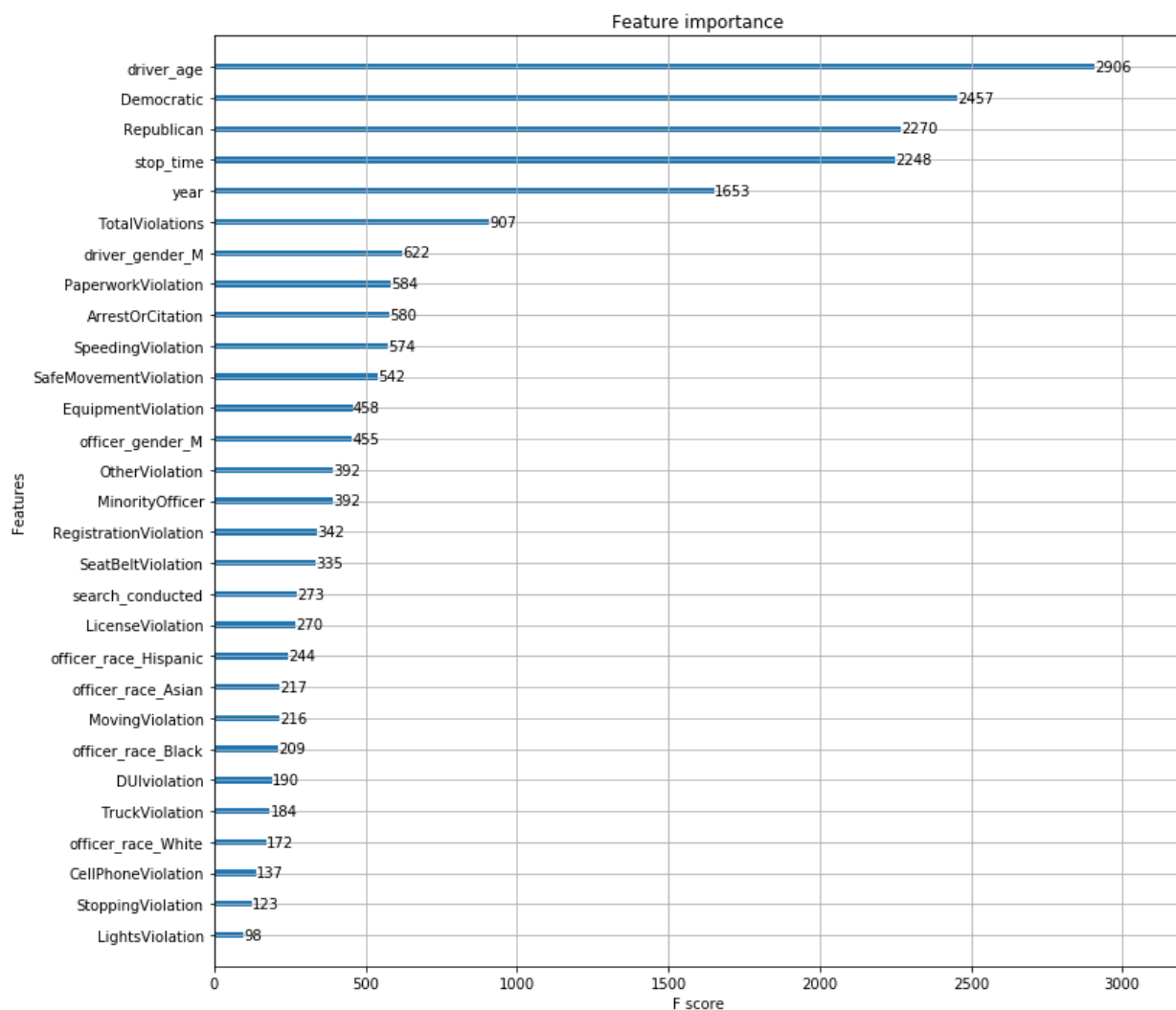
---

<sup>6</sup> "Ensemble Methods." *Scikit-Learn: Machine Learning in Python - Scikit-Learn 0.20.1 Documentation*, Scikit Learn, [scikit-learn.org/stable/modules/ensemble.html](https://scikit-learn.org/stable/modules/ensemble.html).

<sup>7</sup> "XGBoost: A Scalable Tree Boosting System." *Distributed Machine Learning Common*, University of Washington, [dmlc.cs.washington.edu/xgboost.html](https://dmlc.cs.washington.edu/xgboost.html).



Our final model had a prediction accuracy of 62.89%, F1 score of 0.44, and average precision score of 0.29. Despite the model's prediction accuracy, the low average precision score tells us that our output quality was very low with a high misclassification rate. The F1 score, the harmonic mean of the precision and recall, helped indicate that we had a poor model as it was below the expected 0.5 threshold for binary classification. Although our model had poor performance, we found the four most important predictive features to be driver age, Democratic, Republican and stop time (as shown in Figure 2). The next step for potential improvement would be to engineer more features (how stops were instigated, search types, search outcome) to hopefully increase the precision of our model. If provided more time and the necessary resources (computing power), we would further tune the hyper parameters by using 5-fold cross validation along with a wider range of hyperparameter values for the grid searches.



**Figure 2: Feature Importance for all the Predictors.** This figure shows us how much each of the variables matter when predicting race type. The higher the F score, the more important the feature is. Most important feature - driver age. Least important feature - Lights Violation.

## Work Division

We've broken this section into four separate paragraphs (organized by name) to detail specific contributions each member had on the project. All written sections of this paper and previous rough drafts were divided equally among each group member.

Taylor's primary role focused on leading the analysis sections of the project. Taylor engineered features (dummify categorical values) and prepared the dataset for the XGBoost model. Following the preparation, Taylor began working on the statistical tests for the numerical values, specifically ANOVA and Tukey Tests, and provided the tables and histograms (found in Appendix B). In addition, Taylor helped build the initial XGBoost model with Mahir and built baseline models (Random Forest and Decision Tree) to give context as to why our model had poor performance.

Mahir's primary contributions came from the initial statistical testing and building the XGBoost model. Majority of the EDA findings and descriptive stats table was completed by Mahir, along with the accompanying visuals. While Taylor was working on preparing our dataset for machine learning models, Mahir wrote the initial code for our XGBoost model and provided Kai with the under-sampling strategy for training our model.

Jessie's (Hyewon) primary job was to clean the dataset of "unusable" data, provide visualizations, and prepare the PowerPoint. Jessie prepared our dataset for analysis by removing all observations that were missing crucial information, cleaned the remaining observations, and assigned appropriate column names. The PowerPoint structure, visuals and information were all completed by Jessie.

Kai's (Christopher) primary task was with tuning the hyperparameters for our model and organizing/editing the writing sections of our paper(s). Kai started with the web scraping to gather the voting and demographic information. The chi squared independence tests were conducted by Kai along with the expected/observed tables. Once Mahir and Taylor finished the XGBoost model, Kai used K-Fold cross validation (using the under-sampling framework Mahir provided) and grid searches to find the optimal hyperparameters, presenting the final results to the group.

## Future Directions

Although our project showed a lot of initial promise, our model did not perform as we expected. We might attribute one limitation to our large white population and not being able to thoroughly research other sampling techniques. While we feel that our under-sampling strategy was sufficient, it would have been interesting to compare results across different sampling methods. When predicting something as complicated as race, a larger number of features may have resulted into a better performing model. If given the opportunity to “restart” our research, we would engineer more features into our model. One final limitation came from a lack of time and resources to thoroughly train and understand our model. Although we did run a 5 K-Fold cross validation on some of the hyperparameters, our hope was to run grid searches for all hyperparameters with 5-fold cross validation and wider ranges. In addition, we ran into challenges with implementing under sampling within the grid searches and simply ran out of time to optimize them properly.

Even with these limitations, it is important to note that even if our model was able to predict with high precision, the presence of these “trends” cannot be solely attributed to racial bias. Similarly, just because our model has a low precision score, we cannot simply state that no racial bias exists (as seen from our early analysis). Our hope is that our findings (and struggles) might spark other’s interest to continue investigation into this topic. The scope for our project only focused on one state and with this being an ongoing and evolving issue, data is continuing to be collected and will eventually be made available. While our model may not have displayed any particular trends, there are plenty of different directions for researching police practices that offer real potential for uncovering racial bias tendencies.

## Conclusion

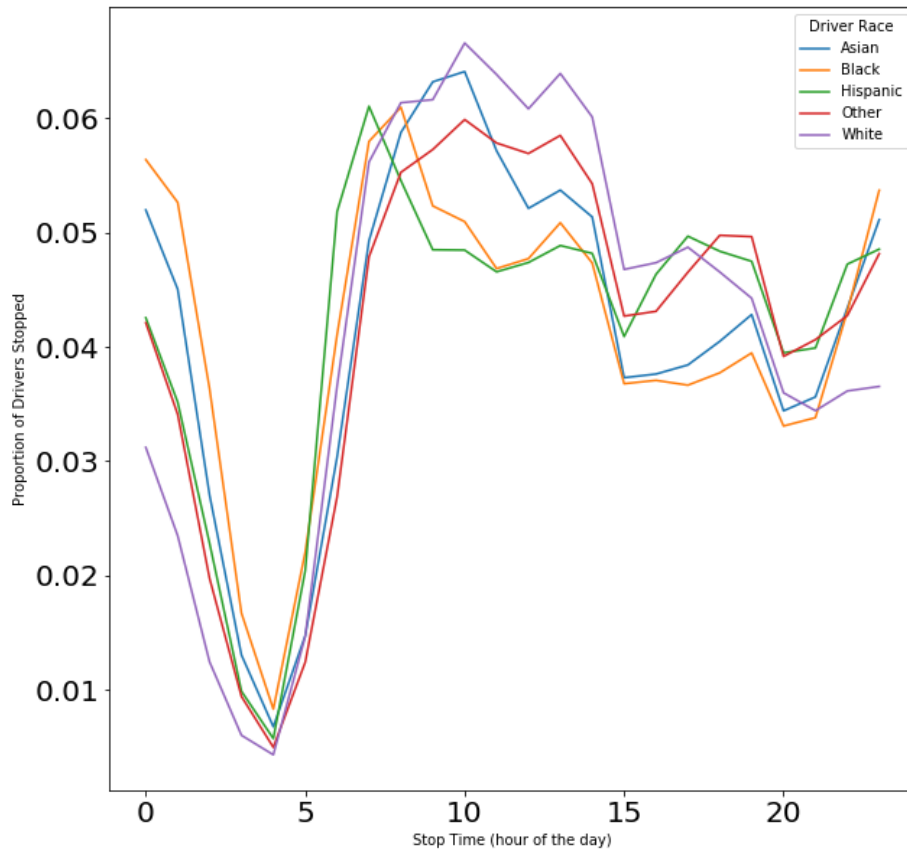
Despite our model not being able to accurately predict minority drivers, perhaps this is a good indication of fair police policies in Washington State. Our original motivation was simply to conduct our own due diligence into Washington State police departments and in a sense, we accomplished that goal. We would hope that many of the discriminatory practices that are too often highlighted in headline news aren’t common place in our police forces so close to home. With that being said, it is still important, and some may argue a duty, to continue to closely examine racial bias in police stops. Our statistical tests still indicated some potential of existing bias and we encourage further investigation into this very much real and active issue. Our research doesn’t mark the end, but rather a continuation of this ongoing inquiry into racial discrimination to discover, report, and share findings that will hopefully spark the social change of the future.

# Appendix

## Appendix A: Detailed Description of Variables Within Data Set

Data Source	Source URL	Variables Description
Stanford Open Policing Project	<a href="https://openpolicing.stanford.edu/data/">https://openpolicing.stanford.edu/data/</a>	<b>Stop Time:</b> Time of stop (by hours) <b>County Name:</b> Name of County where police stop occurred <b>Driver Gender:</b> Gender of driver stored as M or F <b>Driver Age:</b> Age of the offender <b>Driver Race:</b> Race of the offender grouped into 5 categories: Asian, Black, Hispanic, White, Other <b>Officer Race:</b> Race of the police officer grouped into 5 categories: Asian, Black, Hispanic, White, Other <b>Search Conducted:</b> True/False statement if search was conducted during test <b>Violations:</b> Stored as binary value, 1 indicating it was a violation, 0 indicating it was not <b>Minority Driver:</b> Binary column with 1 indicating it was a minority driver and 0 indicating it was not <b>Minority Officer:</b> Binary column with 1 indicating it was a minority officer and 0 indicating it was not <b>Arrest/Citation:</b> Binary column with 1 indicating offender was arrested or citation and 0 indicating they were not
U.S Census QuickFacts	<a href="https://www.census.gov/quickfacts/fact/table/wa/PST045217">https://www.census.gov/quickfacts/fact/table/wa/PST045217</a>	<b>Race.alone</b> (ex Asian.alone): Percentage breakdown of population that makes up that particular race <b>Two.or. more.race:</b> People who indicated they belonged to two race groups or more
Secretary of State Washington.gov	<a href="https://www.sos.wa.gov/elections/">https://www.sos.wa.gov/elections/</a>	<b>Democratic:</b> Percentage of county that voted for democratic presidential nominee <b>Republican:</b> Percentage of county that voted for republican presidential nominee <b>Other:</b> Percentage of county that voted for other party presidential nominees

## Appendix B: Exploratory Data Analysis Table



**Figure 3: Proportion of Drivers Stopped Against the Stop Time, grouped together by driver race.** Minorities witnessed a higher percentage of stops within the twilight hours (between 10 PM to 6 AM) while whites had a higher amount of stops during the midday time (10 AM to 4 PM).

## Appendix C: Descriptive Stats of the Dataset

County with most stops	King County
Time between which most stops occurred	10:00am - 11:00am
Most stopped driver gender	Male
Median driver age	35
Most stopped driver race	White
Most frequent stop outcome	Arrest or citation
Most frequent search outcome	Not searched

Most frequent officer gender	Male
Most frequent officer race	White
Year with most stops	2015
Most frequent violation	Speeding violation
Mean number of violations per stop	1.4

**Table 4: Descriptive statistics about our dataset.** This breakdown gives a general description of our dataset as a whole. Although whites had the highest percentage of stops, this percentage is to be expected due to the high population of whites in Washington State. While majority of stops didn't involve a search, most stops ended with either an arrest or citation.

## Appendix D: Expected vs Observed Frequency Table

For the chi squared tests, we tested to see if the violation type was independent of driver's race by running separate tests for each violation. The resulting test statistics were rather large, and the p-values were all  $< 2.2e^{-16}$ , most likely due to the size of our dataset.

To delve deeper into the discrepancy between the observed and expected values, we wanted to see which violations have higher effects on particular races. The resulting table below (Table 5) displays the violations for each race that had higher observations values ( $> 25\%$  increase) than the expected values and the percentage increase from the expected values to the observed value.

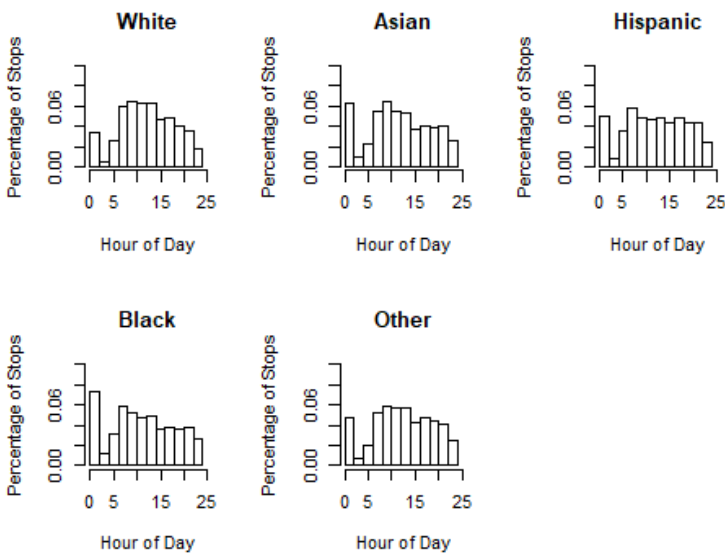
Race	Violation (Percentage Increase from Expected to Observed)
Asian	Moving (45.63%)
Black	DUI (63.43%), License (179.87%), Lights (35.7%), Moving (116.84%)
Hispanic	DUI (51.61%), License (59.26%), Paperwork (34.63%), Stopping (26.12%), Truck (27.90%)
Other	DUI (81.65%), License (137.63%)
White	

**Table 5: Violations with at least a 25% difference between expected and observed values for each race (when observed values are higher).** Here we can see that each of the minority races have some violations that had large increases of their observations over expected values. DUI's seem to be more common in minority violations and license violations saw the largest differences between expected to observed values. In an effort to promote transparency, it is worth noting that with the high population count for whites, it is difficult to see "drastic" differences between their expected and observed values.

## Appendix E: Histograms and Tukey Test Tables for ANOVA Tested Variables

ANOVA tests whether the means of multiple groups are different, so here we are testing if the mean of the variables (stop time, driver age, total violations, and Republican to Democratic ratio) for each race were statistically different.

### Stop Time

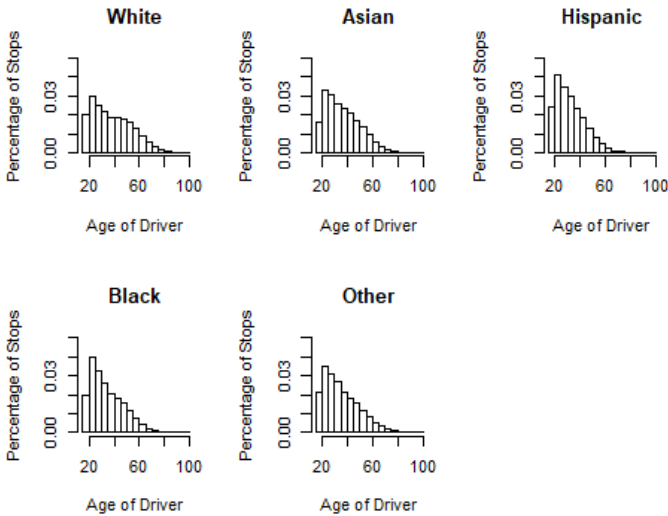


**Figure 4: Histogram of Stop Times Grouped by Race.** Stop times seem to be consistent across all races, except that minorities see a larger spike (higher percentage of stops) around 12 AM.

Race 1 vs Race 2	Difference of Means
Black vs Asian	-0.438
Hispanic vs Asian	0.436
Other vs Asian	0.679
White vs Asian	0.588
Hispanic vs Black	0.874
Other vs Black	1.117
White vs Black	1.026
Other vs Hispanic	0.243
White vs Hispanic	0.152
White vs Other	-0.091

**Table 6: Stop Time Tukey Test Table.** As seen from the histograms, the mean averages for times remained rather consistent. Blacks appeared to be stopped an hour later (on average) than whites.

## Driver Age



**Figure 5: Histogram of Age Grouped By Race.** Younger Hispanic and black drivers (25 and under) tend to have a higher percentage of stops compared to other races of the same age.

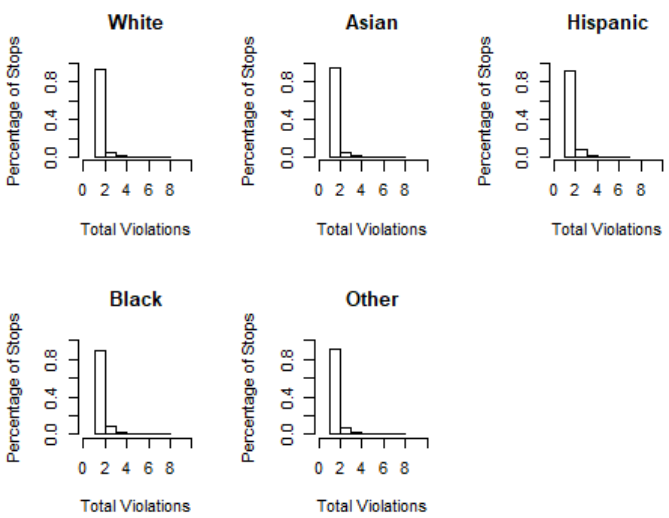
Race 1 vs Race 2	Difference of Means
Black vs Asian	-1.913
Hispanic vs Asian	-3.901
Other vs Asian	-0.910
White vs Asian	2.272
Hispanic vs Black	-1.989
Other vs Black	1.003
White vs Black	4.185
Other vs Hispanic	2.992
White vs Hispanic	6.173
White vs Other	3.181

**Table 7: Driver Age Tukey Test Table.**

Hispanic drivers were almost 2 years younger than every other race on average along with 6 years younger than white drivers. Black drivers were on average 4 years younger than white drivers. White drivers were 2 years older than drivers of any other race. This table displays the full results from Table 3 earlier in the report.



# Total Violations



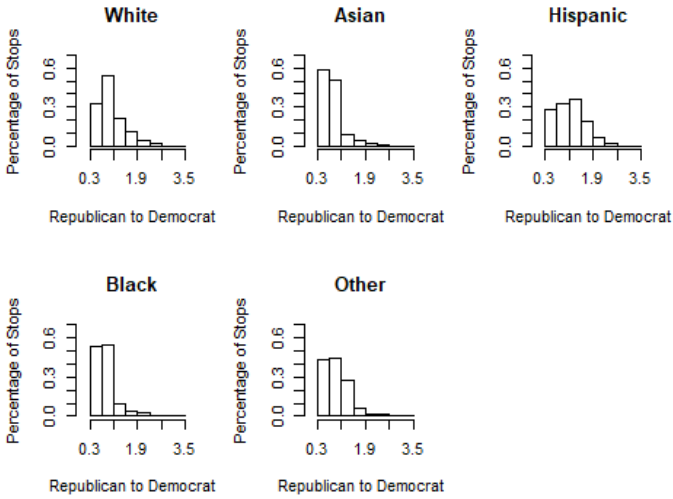
**Figure 6: Histogram of Age Grouped By Race.** Total violation count remains consistent across all of the race groups.

Race 1 vs Race 2	Difference of Means
Black vs Asian	0.175
Hispanic vs Asian	0.148
Other vs Asian	0.129
White vs Asian	0.030
Hispanic vs Black	-0.028
Other vs Black	-0.047
White vs Black	-0.146
Other vs Hispanic	-0.019
White vs Hispanic	-0.118
White vs Other	-0.099

**Table 8: Total Violations Tukey Test Table.**

No mean number of total violations was more than 0.2 violations greater than any other race.

## Ratio of Republican to Democratic Votes



**Figure 7: Histogram of Political Affiliation Grouped By Race.** Hispanics were on average pulled over more often in republican countries (based on the republican to democrat ratio).

Race 1 vs Race 2	Difference of Means
Black vs Asian	0.014
Hispanic vs Asian	0.375
Other vs Asian	0.149
White vs Asian	0.225
Hispanic vs Black	0.362
Other vs Black	0.135
White vs Black	0.211
Other vs Hispanic	-0.227
White vs Hispanic	-0.151
White vs Other	0.076

**Table 9: Republican to Democrat Votes Ratio Tukey Test Table.** A mean ratio .3 higher than the other races' ratios, demonstrates that Hispanics were pulled over more often in Republican counties than the other races.