Race and Police Violence: An Analysis on Whether a Civilian's Race Impacts Whether a Police Officer Will Use Physical Force During Stop and Frisks in New York

Abstract

Recently, police interaction with minorities has received more and more scrutiny, and tensions between the two groups have steadily risen. Many other studies have sought to examine aspects of this trend in attempts to gain more insight and possibly relieve some of the tensions. This paper uses NYPD stop-and-frisk data from 2005-2018 to analyze the relationship between civilian race and whether physical force is used by police during stops. The results demonstrate that minorities are stopped more frequently, and they experience physical force from a police officer more often than their white counterparts. This result is hard to generalize to other areas and further research is necessary to more closely examine this relationship and help control minority-police tensions.

Introduction

In the last decade, police violence against minorities has become an issue at the forefront of the media and the minds of American citizens. This violence became especially problematic in places like New York where the massive amounts of minorities being stopped skyrocketed to hundreds of thousands, only for the majority of those stopped to be found innocent of any crimes. Smith and Holmes (2014) demonstrate that black people are more likely than white people to report incidents of experiencing police brutality, and this has generally been accepted. Thus, the popular narrative has become that racial minorities, especially black people, are in greater danger of experiencing police brutality than white people are. However, some of the statistical analysis on the topic has generated doubts over the accuracy of this narrative. For example, a Michigan State study by Cesario and Brooks (2019) found no evidence for this trend, but the strength of conclusions was limited due to the infrequency of their measured response. In their study they examined fatal shootings of black people by police officers of different races, however, the analysis discussed in this paper is centered on the use of physical force by officers of the New York Police Department (NYPD). Both are uncommon responses which presents a unique problem for making conclusions.

This paper seeks to answer the question: does the race of a civilian influence the likelihood that an NYPD police officer will use physical force on the civilian? We hypothesized that, compared to white people, racial minorities would experience disproportionately more physical force when stopped by NYPD officers. While this study focuses solely on data for NYPD stops, the issue extends beyond simply New York. There is much opportunity for further research and analysis to examine and compare police brutality on a broader scale. This issue has become especially salient in the last decade as racial tensions have increased with the recent political climate and increased media. Through social media especially, incidents of police misconduct such as brutality or shootings have become readily available to huge amounts of people. Because of this, tensions between police officers and citizens, especially minority citizens, are extremely high. This research is important because analysis that can provide concrete quantitative information on police brutality against minorities could possibly begin to ease the tension between police officers and racial minorities.

Data

We obtained stop and frisk data from the New York Civil Liberties Union (NYCLU), a non-profit civil rights organization. The online database contains annual data on stop and frisks in New York City from 2003 to 2018. Each row in the dataset corresponds to an instance that an officer stopped a civilian. Police officers are supposed to report details of every stop. Each annual dataset includes up to 100 variables for each observation– such as the age of the person stopped, if the person was frisked, if a weapon was recovered, if physical force was used and the exact times of the stops. The data comes directly from police officers' reports on stops. We obtained a cleaner version of the dataset from our professor. This dataset combined all the annual datasets from 2005 to 2018 and contained a few additional variables.

Data Cleaning

For our analysis, we wanted to create the most accurate model to predict the probability that physical force was used by the police during a stop then conduct a log-likelihood test to assess if civilian race affected this probability. We began our data cleaning by selecting 18 explanatory variables that we believed would contribute to physical force being used on a civilian. Next, we separated date into month and year columns and changed binary variables that indicated Y/N to 1/0. We then verified that quantitative variables were recognized as numeric and that categorical variables were recognized as characters in R. We converted them if necessary.

For our response variable, we created a general physical force measure to contain any instance an officer used physical force during a stop. We merged 5 physical force variables: whether an officer hit the civilian with a weapon, whether the officer handcuffed the civilian, whether the officer used spray on a civilian, whether the officer drew and pointed a firearm at the civilian and whether the officer used other types of physical force into a single column. We coded a 1 if any of these types of physical force were used during a stop and coded a 0 if none were. Further, we cleaned the subject race column, so it included black, white and other race categories. Lastly, we deleted any of the rows containing any NA's - leaving our final sample size at 3,733,491 NYPD stops.

Model Creating and Analysis

We began by creating an initial logistic regression model with the 18 explanatory variables we selected to assess the probability that an officer used physical force during a stop (G = 756,637, df = 22). We then examined the output of the model (Table 1) and spent some time experimenting with other models to find the best model with some variation of the initial 18 variables. We ended up with an 11-variable model (G = 749,160, df = 15, Table 2) that reduced the G statistic by less than 1% (7,477) yet used 7 less variables. We then created a model with those 11 explanatory variables and each pairing as an interaction (G = 770,257, df = 109).

Because we had hundreds of terms in this model with the interactions, we obtained a data frame that only included terms that had significant p-values from the model output. Next, we examined this data frame, experimented with various models and ended up selecting 9 single variable and 20 interaction term to use as our final model. The single variables we used were whether a civilian was searched, whether the civilian had other contraband, whether the civilian had a firearm, whether the civilian had another weapon, whether the civilian was frisked, the suspected crime, stop duration, month and year. Each single variable and interaction term had large Wald statistics and statistically significant p-values (<.05), plus the model itself yielded a very large G statistic for that number of variables. We created the final model with (G = 761417, df = 40, Table 3) and without the race term (G = 759542, df = 38, Table 4). Finally, we conducted a log likelihood test to evaluate the impact of race.

Results

Race significantly affect the probability that a police officer used physical force during a stop and frisk (p = 2e^-16, G = 1875.1). A segmented bar chart is used to analyze the breakdown between race: the proportion of blacks, other races and whites that had physical force used on them was 22%, 23% and 16% respectively (Figure 1).Thus, police officers used physical force on blacks and other races at a similar rate, yet they used physical force on whites at a 38% lower rate than other races and a 44% lower rate than blacks. Further, by utilizing a stacked bar chart, we see that blacks, other races and whites were stopped at much different levels (Figure 2). Almost 2 million blacks were stopped, nearly 1.5 million civilians of other races were stopped yet just around .4 million whites were stopped (Table 5). Thus, blacks were stopped over 4 times as often as whites, and other races were stopped over 3 times as much as whites. Blacks make up over half of the stops, yet only around 25% of the population.

Further, we used a Confusion matrix (Table 6, Table 7) to validate the model. The accuracy rate of the model in predicting the outcome without race was 78.15%, and the accuracy rate of the model with race was 78.17% - for a difference of .02%. Thus, adding race just marginally improves the model's ability to predict whether physical force will be used on a civilian during a stop and frisk. A ROC curve (Figure 3) was used to determine the threshold (minimum predicted probability to determine a positive result) for the confusion matrix. The selected value was 0.6 because this yielded a true positive rate around 0.66 and a false positive rate under 0.2. A true positive means that the model predicted success and the actual outcome was a success. A false positive means the model predicted success when the actual outcome was failure.

Discussion

As this analysis has shown, there is evidence that minorities, specifically black people, experience physical force from NYPD officers more often than their white counterparts. Both figure 1 and 2 demonstrate that minorities are stopped more often, and a greater proportion of their stops result in physical force than when a white person is stopped. This provides support for our original hypothesis. However, concrete conclusions are hard to make because of the limited nature of these events, as well as the possibilities of confounding. This analysis did not control for things like location which could skew the results. If more problematic neighborhoods have denser minority populations, these results could be partially due to systematic problems in segregation rather than simply racial prejudice held by officers. Another potentially confounding factor is that racial minorities were stopped far more than white people. While this provides evidence for increased suspicion of racial minorities by NYPD officers, it potentially confounds this analysis on the use of physical force. Furthermore, as Cesario and Brooks (2019) discussed, it becomes difficult to make meaningful conclusions when measured responses are uncommon.

Some of the barriers to making generalizable conclusions from this research stem from the data itself. First, this data was only collected on NYPD officers and it is possible that New York is vastly different from many other areas in the country. Thus, it is hard to generalize these results to police officers outside of the NYPD and the civilians they serve. Second, this data was self-report. Officers were supposed to report the details on every single stop they made, but humans are bound to leave out information or make similar errors. This is evident in problems with officers incorrectly marking sex of the suspect, their race, etc. Furthermore, the variables for physical force that served as the response variable in this study could be influenced or lack credibility due to the nature of their collection.

Before any generalizations toward police officers and minorities across the nation can be made, further research outside the NYPD is required. Though it would complicate the process, a different method of collecting data on incidents could improve the credibility of future studies as well. Furthermore, an interesting aspect that was explored in some previous studies was the race of the police officer that was guilty of the observed misconduct. This would lead to more meaningful conclusions on minority-police officer relations as it could provide support for structural issues in police departments, fundamental racial tensions, a combination of both or an entirely different issue. Further research and quantitative analysis must be completed in these areas to provide insight into police-minority tensions. Without concrete evidence the need for action has not been thoroughly demonstrated, and through inaction, tensions are only worsening.

References

Cesario, Joseph. “The Truth behind Racial Disparities in Fatal Police Shootings.” *MSUToday*, Michigan State University, 22 July 2019, https://msutoday.msu.edu/news/2019/the-truth-behind-racial-disparities-in-fatal-police-shootings/.

Smith, B. W., & Holmes, M. D. (2014). Police Use of Excessive Force in Minority Communities. *Social Problems*, *61*(1), 83–104.

“Stop-and-Frisk Data.” *New York Civil Liberties Union*, 3 Dec. 2019, https://www.nyclu.org/en/stop-and-frisk-data.

“The Changing Racial and Ethnic Makeup of New York City Neighborhoods.” *Furman Center for Real Estate and Urban Policy*, 2011.

Appendix

```
Coefficients:
                                                       Estimate Std. Error  z value Pr(>|z|)    
(Intercept)                                           1.493e+02  1.434e+00  104.145  < 2e-16 ***
nypd_data$OBSERVED_DURATION_MINUTES                   1.970e-03  2.374e-04    8.295  < 2e-16 ***
nypd_data$OTHER_PERSON_STOPPED_FLAG                   2.510e-01  3.311e-03   75.800  < 2e-16 ***
nypd_data$SEARCHED_FLAG                               7.781e-01  4.120e-03  188.838  < 2e-16 ***
nypd_data$BACKGROUND_CIRCUMSTANCES_VIOLENT_CRIME_FLAGY  1.603e-01  4.338e-03   36.951  < 2e-16 ***
nypd_data$SUSPECT_WEIGHT                              2.581e-04  3.162e-05    8.163 3.28e-16 ***
nypd_data$FIREARM_FLAG                                8.032e-01  3.141e-02   25.571  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALOther             -2.897e-01  9.674e-03  -29.944  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALSubstance         -6.304e-01  8.344e-03  -75.556  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTheft             -5.011e-01  6.808e-03  -73.603  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTresspass         -7.402e-01  8.932e-03  -82.864  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALWeapon            -2.740e-01  6.915e-03  -39.627  < 2e-16 ***
nypd_data$STOP_DURATION_MINUTES                       1.201e-02  2.520e-04   47.648  < 2e-16 ***
nypd_data$Month                                      -9.914e-03  3.993e-04  -24.827  < 2e-16 ***
nypd_data$OFFICER_EXPLAINED_STOP_FLAG                -5.189e-01  4.644e-02  -11.173  < 2e-16 ***
nypd_data$FRISKED_FLAG                                2.412e+00  4.477e-03  538.722  < 2e-16 ***
nypd_data$OTHER_CONTRABAND_FLAG                       5.117e-01  8.936e-03   57.258  < 2e-16 ***
nypd_data$OTHER_WEAPON_FLAG                           2.202e-01  2.152e-02   10.234  < 2e-16 ***
nypd_data$SUSPECT_SEXM                                1.178e-02  6.870e-03    1.715 0.086391 .  
nypd_data$SUSPECT_SEXOTHER                            6.678e-02  1.254e-02    5.324 1.02e-07 ***
nypd_data$SUSPECT_REPORTED_AGE                       -1.929e-04  5.751e-05   -3.354 0.000797 ***
nypd_data$SUSPECT_HEIGHT                              1.945e-03  4.655e-04    4.178 2.94e-05 ***
nypd_data$YEAR2                                      -7.552e-02  7.130e-04 -105.926  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3945356  on 3733490  degrees of freedom
Residual deviance: 3188719  on 3733468  degrees of freedom
AIC: 3188765
```

Table 1: R output from initial model

```
Coefficients:
                                             Estimate Std. Error  z value Pr(>|z|)    
(Intercept)                                 1.480e+02  1.430e+00  103.510  < 2e-16 ***
nypd_data$SEARCHED_FLAG                     7.684e-01  4.111e-03  186.896  < 2e-16 ***
nypd_data$FRISKED_FLAG                      2.432e+00  4.462e-03  545.087  < 2e-16 ***
nypd_data$OTHER_CONTRABAND_FLAG             5.077e-01  8.923e-03   56.892  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALOther   -3.287e-01  9.617e-03  -34.181  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALSubstance -6.678e-01  8.259e-03  -80.854  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTheft   -5.277e-01  6.769e-03  -77.957  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTresspass -8.064e-01  8.860e-03  -91.010  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALWeapon  -3.262e-01  6.850e-03  -47.622  < 2e-16 ***
nypd_data$STOP_DURATION_MINUTES             1.436e-02  2.559e-04   56.124  < 2e-16 ***
nypd_data$FIREARM_FLAG                      8.523e-01  3.135e-02   27.184  < 2e-16 ***
nypd_data$Month                            -9.903e-03  3.987e-04  -24.841  < 2e-16 ***
nypd_data$YEAR2                            -7.500e-02  7.114e-04 -105.425  < 2e-16 ***
nypd_data$OTHER_WEAPON_FLAG                 2.208e-01  2.147e-02   10.286  < 2e-16 ***
nypd_data$SUSPECT_SEXM                      1.556e-02  6.662e-03    2.336   0.0195 *  
nypd_data$SUSPECT_SEXOTHER                  6.362e-02  1.244e-02    5.115 3.14e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3945356  on 3733490  degrees of freedom
Residual deviance: 3196196  on 3733475  degrees of freedom
AIC: 3196228
```

Table 2: R output from reduced model

```
Coefficients:
                                                                          Estimate Std. Error  z value Pr(>|z|)    
(Intercept)                                                              1.511e+02  1.431e+00  105.606  < 2e-16 ***
nypd_data$SEARCHED_FLAG                                                  9.166e-01  1.784e-02   51.376  < 2e-16 ***
nypd_data$OTHER_CONTRABAND_FLAG                                          2.271e+00  2.494e-02   91.061  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALOther                                -5.721e-01  2.000e-02  -28.607  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALSubstance                            -1.062e+00  1.867e-02  -56.871  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTheft                                -9.998e-01  1.562e-02  -63.997  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTresspass                            -1.430e+00  1.943e-02  -73.611  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALWeapon                               -1.918e-01  1.831e-02  -10.477  < 2e-16 ***
nypd_data$STOP_DURATION_MINUTES                                          1.133e-02  7.579e-04   14.956  < 2e-16 ***
nypd_data$FIREARM_FLAG                                                   8.431e-01  7.256e-02   11.619  < 2e-16 ***
nypd_data$Month                                                         -1.012e-02  3.990e-04  -25.356  < 2e-16 ***
nypd_data$YEAR2                                                         -7.636e-02  7.116e-04 -107.307  < 2e-16 ***
nypd_data$OTHER_WEAPON_FLAG                                              4.308e-01  3.768e-02   11.432  < 2e-16 ***
nypd_data$FRISKED_FLAG                                                   2.003e+00  1.616e-02  123.941  < 2e-16 ***
nypd_data$SUSPECT_RACE_DESCRIPTIONOTHER                                  7.404e-02  2.927e-03   25.294  < 2e-16 ***
nypd_data$SUSPECT_RACE_DESCRIPTIONWHITE                                 -1.481e-01  5.366e-03  -27.594  < 2e-16 ***
nypd_data$SEARCHED_FLAG:nypd_data$OTHER_CONTRABAND_FLAG                 -2.045e-01  1.854e-02  -11.032  < 2e-16 ***
nypd_data$SEARCHED_FLAG:nypd_data$CRIME_DESCRIPTION_GENERALOther         1.022e-01  2.532e-02    4.037 5.42e-05 ***
nypd_data$SEARCHED_FLAG:nypd_data$CRIME_DESCRIPTION_GENERALSubstance     1.685e-01  2.174e-02    7.750 9.16e-15 ***
nypd_data$SEARCHED_FLAG:nypd_data$CRIME_DESCRIPTION_GENERALTheft        -3.252e-02  1.844e-02   -1.763 0.077902 .  
nypd_data$SEARCHED_FLAG:nypd_data$CRIME_DESCRIPTION_GENERALTresspass    -6.249e-02  2.298e-02   -2.719 0.006547 ** 
nypd_data$SEARCHED_FLAG:nypd_data$CRIME_DESCRIPTION_GENERALWeapon       -1.731e-01  1.836e-02   -9.430  < 2e-16 ***
nypd_data$SEARCHED_FLAG:nypd_data$STOP_DURATION_MINUTES                 -1.121e-02  5.687e-04  -19.714  < 2e-16 ***
nypd_data$SEARCHED_FLAG:nypd_data$FIREARM_FLAG                           2.409e-01  7.788e-02    3.094 0.001978 ** 
nypd_data$SEARCHED_FLAG:nypd_data$OTHER_WEAPON_FLAG                     -2.412e-01  4.541e-02   -5.313 1.08e-07 ***
nypd_data$OTHER_CONTRABAND_FLAG:nypd_data$FRISKED_FLAG                  -1.798e+00  2.540e-02  -70.778  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALOther:nypd_data$FRISKED_FLAG          3.096e-01  2.260e-02   13.700  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALSubstance:nypd_data$FRISKED_FLAG      4.511e-01  2.042e-02   22.091  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTheft:nypd_data$FRISKED_FLAG          5.483e-01  1.677e-02   32.689  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTresspass:nypd_data$FRISKED_FLAG      7.654e-01  2.139e-02   35.786  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALWeapon:nypd_data$FRISKED_FLAG        -6.698e-02  1.930e-02   -3.470 0.000521 ***
nypd_data$STOP_DURATION_MINUTES:nypd_data$FRISKED_FLAG                   5.357e-03  6.268e-04    8.547  < 2e-16 ***
nypd_data$OTHER_CONTRABAND_FLAG:nypd_data$STOP_DURATION_MINUTES         -5.124e-03  6.897e-04   -7.429 1.09e-13 ***
nypd_data$OTHER_CONTRABAND_FLAG:nypd_data$FIREARM_FLAG                  -2.915e-01  7.609e-02   -3.831 0.000128 ***
nypd_data$CRIME_DESCRIPTION_GENERALOther:nypd_data$STOP_DURATION_MINUTES     -2.522e-03  9.109e-04   -2.769 0.005629 ** 
nypd_data$CRIME_DESCRIPTION_GENERALSubstance:nypd_data$STOP_DURATION_MINUTES -7.995e-04  8.989e-04   -0.889 0.373745    
nypd_data$CRIME_DESCRIPTION_GENERALTheft:nypd_data$STOP_DURATION_MINUTES      6.306e-03  7.845e-04    8.038 9.13e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTresspass:nypd_data$STOP_DURATION_MINUTES  9.146e-03  1.309e-03    6.989 2.77e-12 ***
nypd_data$CRIME_DESCRIPTION_GENERALWeapon:nypd_data$STOP_DURATION_MINUTES     5.202e-03  8.471e-04    6.141 8.18e-10 ***
nypd_data$STOP_DURATION_MINUTES:nypd_data$FIREARM_FLAG                  -4.733e-03  1.742e-03   -2.716 0.006599 ** 
nypd_data$FIREARM_FLAG:nypd_data$OTHER_WEAPON_FLAG                      -4.515e-01  1.158e-01   -3.897 9.74e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3945356  on 3733490  degrees of freedom
Residual deviance: 3183939  on 3733450  degrees of freedom
AIC: 3184021
```

Table 3: R output of final model with race variable

```
Coefficients:
                                                                          Estimate Std. Error  z value Pr(>|z|)    
(Intercept)                                                              1.507e+02  1.429e+00  105.442  < 2e-16 ***
nypd_data$SEARCHED_FLAG                                                  9.118e-01  1.783e-02   51.147  < 2e-16 ***
nypd_data$OTHER_CONTRABAND_FLAG                                          2.265e+00  2.492e-02   90.904  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALOther                                -5.881e-01  1.999e-02  -29.423  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALSubstance                            -1.064e+00  1.866e-02  -57.010  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTheft                                -1.005e+00  1.562e-02  -64.346  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTresspass                            -1.427e+00  1.942e-02  -73.464  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALWeapon                               -1.883e-01  1.830e-02  -10.290  < 2e-16 ***
nypd_data$STOP_DURATION_MINUTES                                          1.116e-02  7.539e-04   14.802  < 2e-16 ***
nypd_data$FIREARM_FLAG                                                   8.390e-01  7.251e-02   11.570  < 2e-16 ***
nypd_data$Month                                                         -1.017e-02  3.988e-04  -25.503  < 2e-16 ***
nypd_data$YEAR2                                                         -7.617e-02  7.110e-04 -107.131  < 2e-16 ***
nypd_data$OTHER_WEAPON_FLAG                                              4.200e-01  3.766e-02   11.153  < 2e-16 ***
nypd_data$FRISKED_FLAG                                                   2.013e+00  1.615e-02  124.630  < 2e-16 ***
nypd_data$SEARCHED_FLAG:nypd_data$OTHER_CONTRABAND_FLAG                 -2.063e-01  1.853e-02  -11.134  < 2e-16 ***
nypd_data$SEARCHED_FLAG:nypd_data$CRIME_DESCRIPTION_GENERALOther         1.053e-01  2.531e-02    4.160 3.19e-05 ***
nypd_data$SEARCHED_FLAG:nypd_data$CRIME_DESCRIPTION_GENERALSubstance     1.729e-01  2.173e-02    7.957 1.76e-15 ***
nypd_data$SEARCHED_FLAG:nypd_data$CRIME_DESCRIPTION_GENERALTheft        -3.150e-02  1.844e-02   -1.709 0.087472 .  
nypd_data$SEARCHED_FLAG:nypd_data$CRIME_DESCRIPTION_GENERALTresspass    -5.771e-02  2.298e-02   -2.512 0.012017 *  
nypd_data$SEARCHED_FLAG:nypd_data$CRIME_DESCRIPTION_GENERALWeapon       -1.710e-01  1.835e-02   -9.322  < 2e-16 ***
nypd_data$SEARCHED_FLAG:nypd_data$STOP_DURATION_MINUTES                 -1.104e-02  5.666e-04  -19.477  < 2e-16 ***
nypd_data$SEARCHED_FLAG:nypd_data$FIREARM_FLAG                           2.395e-01  7.786e-02    3.076 0.002098 ** 
nypd_data$SEARCHED_FLAG:nypd_data$OTHER_WEAPON_FLAG                     -2.329e-01  4.538e-02   -5.134 2.84e-07 ***
nypd_data$OTHER_CONTRABAND_FLAG:nypd_data$FRISKED_FLAG                  -1.798e+00  2.539e-02  -70.802  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALOther:nypd_data$FRISKED_FLAG          3.079e-01  2.259e-02   13.626  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALSubstance:nypd_data$FRISKED_FLAG      4.403e-01  2.041e-02   21.568  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTheft:nypd_data$FRISKED_FLAG          5.461e-01  1.677e-02   32.565  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALTresspass:nypd_data$FRISKED_FLAG      7.587e-01  2.138e-02   35.482  < 2e-16 ***
nypd_data$CRIME_DESCRIPTION_GENERALWeapon:nypd_data$FRISKED_FLAG        -7.496e-02  1.930e-02   -3.884 0.000103 ***
nypd_data$STOP_DURATION_MINUTES:nypd_data$FRISKED_FLAG                   5.325e-03  6.246e-04    8.525  < 2e-16 ***
nypd_data$OTHER_CONTRABAND_FLAG:nypd_data$STOP_DURATION_MINUTES         -5.057e-03  6.858e-04   -7.374 1.66e-13 ***
nypd_data$OTHER_CONTRABAND_FLAG:nypd_data$FIREARM_FLAG                  -2.866e-01  7.609e-02   -3.767 0.000165 ***
nypd_data$CRIME_DESCRIPTION_GENERALOther:nypd_data$STOP_DURATION_MINUTES     -2.538e-03  9.073e-04   -2.798 0.005144 ** 
nypd_data$CRIME_DESCRIPTION_GENERALSubstance:nypd_data$STOP_DURATION_MINUTES -8.545e-04  8.943e-04   -0.955 0.339351    
nypd_data$CRIME_DESCRIPTION_GENERALTheft:nypd_data$STOP_DURATION_MINUTES      6.159e-03  7.823e-04    7.873 3.45e-15 ***
nypd_data$CRIME_DESCRIPTION_GENERALTresspass:nypd_data$STOP_DURATION_MINUTES  9.044e-03  1.307e-03    6.921 4.48e-12 ***
nypd_data$CRIME_DESCRIPTION_GENERALWeapon:nypd_data$STOP_DURATION_MINUTES     5.155e-03  8.447e-04    6.103 1.04e-09 ***
nypd_data$STOP_DURATION_MINUTES:nypd_data$FIREARM_FLAG                  -4.683e-03  1.724e-03   -2.715 0.006622 ** 
nypd_data$FIREARM_FLAG:nypd_data$OTHER_WEAPON_FLAG                      -4.472e-01  1.158e-01   -3.862 0.000112 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3945356  on 3733490  degrees of freedom
Residual deviance: 3185814  on 3733452  degrees of freedom
AIC: 3185892
```

Table 4: R output of final model without race variable

BLACK	OTHER	WHITE
1957081	1411315	365095

Table 5: Officer stops broken down by race

```
> table(nypd_data$GeneralPhysicalForce, PredictForcewR > 0.6)
```

	FALSE	TRUE
0	2891298	16490
1	798696	27007

Table 6: Confucian matrix for model with race

```
table(nypd_data$GeneralPhysicalForce, PredictForce > 0.6)
```

	FALSE	TRUE
0	2891788	16000
1	799644	26059

Table 7: Confucian matrix for model without race

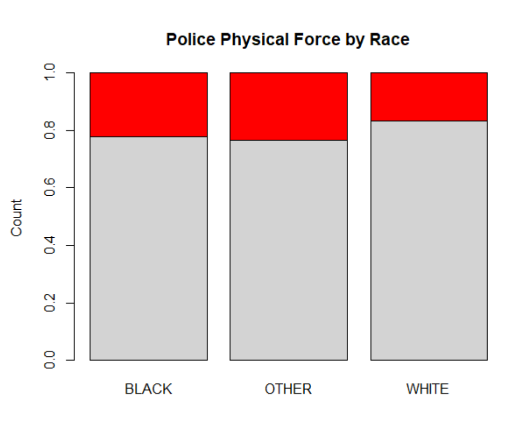

Figure 1: Police physical force by race (proportions)

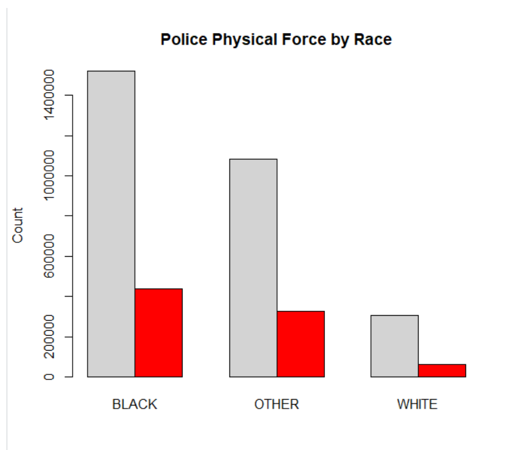

Figure 2: Police physical force by race (counts)

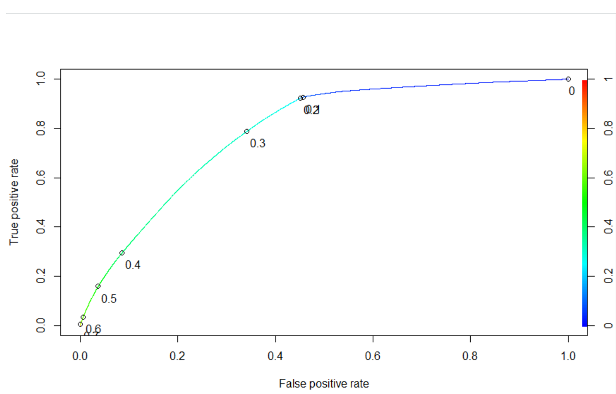

Figure 3: ROC curve for Confucian matrix