

**Semester 2, AY2020-21**

**BC2407 Analytics II: Advanced Predictive Techniques**

**Project Work (Group) Report**

**Analysing Police Misconduct in the United States of America (US)**

|  |  |
| --- | --- |
| **Member Name** | **Matriculation Number** |
| CHEW WEN QIANG, JEREMY | U1911921H |
| KENNETH CHIA | U1910780H |
| LIM ZHENG YI, DENNIS | U1910514C |
| TAN JUN HONG | U1920048B |
| WANG WAYNE | U1920270G |

**Seminar Tutor:** Prof. Neumann Chew

**Seminar:** 3

**Team:** 7

**Date:**  3rd April 2021

**Table of Content**

[**Executive Summary**](#_2et8ii924n2) **3**

[**1 Introduction**](#_nlo4stnxedq2) **4**

[1.1 Social Good Issue](#_f2518wtxsm15) 4

[1.2 Justification and Consequences](#_3fi1zl8wvqiw) 4

[1.3 Key Objectives](#_rvd87j2hsik5) 5

[**2 Data Exploration**](#_8nd7ghfmvr1d) **6**

[2.1 Data Acquisition](#_6e4yqx5wav3) 6

[2.2 Data Cleaning](#_xdpnb48tq90) 6

[2.3 Data Visualization](#_vmzxybsbyudb) 7

[**3 Root Cause Analysis of Police Misconduct**](#_46ndvw29a0tf) **9**

[3.1 Data Modelling](#_inltzyv47xu3) 9

[3.1.1 Logistic Regression](#_2jxsxqh) 9

[3.1.2 Multi-Adaptive Regression Splines (MARS)](#_jab3e1hdayme) 9

[3.2 Evaluation of Models](#_sgarzrymvxv6) 10

[3.3 Root Cause Analysis](#_cbii373b15db) 10

[3.3.1 Police Age](#_vjqkzbayu3mu) 10

[3.3.2 Police Ethnicity](#_dsu9gyedzvje) 10

[3.3.3 Victim Age](#_bkfyp886ktpq) 11

[3.3.4 Victim Gender](#_n4i6c67ww1co) 11

[3.3.5 Victim Ethnicity](#_pyspkqon6u5v) 11

[3.3.6 Repeated Complaints](#_pl692mwr793d) 12

[3.4 Limitations](#_967qltuletoc) 12

[3.5 Recommendations](#_ieuwjmrc1qxa) 12

[3.5.1 Scenario-based Hiring and Trainings](#_xc99nhdbixf0) 13

[3.5.2 Psychological Training](#_v19e27bjmmdc) 13

[3.5.3 Familiarisation](#_bgxhgeok7b2u) 14

[3.5.4 Community Campaigns](#_cql6patx8135) 14

[**4 Prediction Model for High-risk Police Officers**](#_py6dvtbd86x0) **14**

[4.1 Data Modelling](#_d0zwbpq0agxt) 14

[4.1.1 Logistic Regression](#_6ssa1b8yw6tv) 15

[4.1.2 MARS](#_f1v2nuxzum0h) 16

[4.1.3 CART](#_bvmpc2yr7651) 17

[4.1.4 Random Forest](#_kwpq4ieiaw5h) 18

[4.2 Evaluation of Models](#_99v9uhd4mun3) 20

[4.3 Limitations](#_a70mtfk4aoq1) 21

[4.4 Recommendations](#_sv5yfxinkjfz) 21

[4.4.1 Creating transparent disciplinary matrices](#_yf11wqesb7e9) 21

[4.4.2 Education-based discipline](#_hfngar472qap) 22

[4.4.3 Problem-solving approach](#_4e2fy1c1m2ib) 22

[4.4.4 Innovations of excessive force training](#_jsydr9716jp) 22

[**5 Conclusion**](#_go0306b24xek) **22**

[**References**](#_5456a364qg9w) **23**

[**Appendices**](#_xaug4wtetv5a) **26**

# Executive Summary

Police misconduct has gradually become one of the most severe and talked about issues in the United States (US) in recent years. Some of the reasons for police misconduct that are discussed include demographic factors, psychological and social-environmental factors.

This report depicts the process of the team crafting the best solution to tackle the issue of police misconduct by considering both the characteristics of the police officers and victims. The team utilized the dataset provided by New York City’s Civilian Complaint Review Board (CCRB), an official independent agency that investigates the complaints against NYC police officers in order to tackle the issue.

Our main approach to the issue is two-pronged: first, we used models with high explainability including Logistic Regression and MARS to identify the root causes of police misconduct. The team derived crucial insights as to why and under what circumstances do police misconduct happen. The models also identified predictors that were never mentioned in today’s society and assisted the team in discovering new variables that may prove to be influential on the issue of police misconduct. By identifying the main factors for police misconduct, we would then be capable of optimizing solutions to tackle the root causes and reduce the number of cases of police misconduct. Key factors identified include the number of repeated complaints against police officers, the victim’s ethnicity, the victim’s gender, the victim’s age, the police’s ethnicity and the police’s age.

The second segment of our approach requires the use of a highly accurate predictive model - CART and Random Forest to predict whether a police officer is at high risk of performing misconduct on the job. This model can be utilized in different cities in order to identify the key variables that are associated with police misconduct and will aid the cities in implementing optimal solutions to tackle the problem. The Random Forest managed to acquire a high predictive accuracy of 89.06%, while performing well for the other performance metrics including precision, recall and F1 score, which makes it very fitting for the purpose of curbing police misconduct.

The team acknowledges the limitations of the project and recommends future actions to refine the solution and model in order to better tackle the issue of police misconduct. Within our dataset, there are factors such as the number of repeated complaints recorded against a particular police officer which is too complicated to simply be taken at face value and should be investigated further to identify why that particular police officer has received repeated complaints. Considering the fact that the dataset contains only a limited number of variables, it should be recognized there may be other psychological and socio-environmental variables that can be used to explain the issue of police misconduct such as the working hours of police officers. Ultimately, our models and proofs of concept aim to offer insights on the root causes of police misconduct and provide curated recommendations to reduce the overall high rates of police misconduct prevalent in the US.

# 1 Introduction

## 1.1 Social Good Issue

With incidences of police misconduct in the US growing exponentially, the issue of police misconduct is undeniably becoming an increasingly pressing societal issue. This problem stems from various underlying reasons and causes that are either left unidentified or unaddressed, which led to the issue worsening. Essentially, identifying proper risk profiles with regards to police misconduct will lead to an effective strategy and implementation against police misconduct.

With regards to the current situation in the US, police misconduct is a persisting issue and has not fully been resolved effectively. As such, to better recommend strategies to the issue at hand, our team will be diving into the social issue of police misconduct to discover the causes of police misconduct and create a model that will be able to predict and identify police officers at higher risk of committing misconduct.

## 1.2 Justification and Consequences

To set the tone, the issue of police misconduct is becoming the centre of attraction in the US over the last few years due to the many incidents which involved systemic racism and police brutality. One notable incident is regarding the victim named George Floyd (BBC, 2020). On 25 May 2020, George Floyd, a black man, was arrested after a convenience store employee called 911 and informed the police that Mr. Floyd had purchased cigarettes using a counterfeit $20 bill. During the process of the arrest, George Floyd was pinned to the ground and one of the police officers knelt on his neck. This led to suffocation and subsequently, his death. This incident became one of the most-talked-about topics in the world and the protest and movement of “Black Lives Matter”. The case of George Floyd was not an isolated incident with many precedence cases such as Michael Brown in 2014 and Alton Sterling in 2016.

The ramifications of police misconduct results in a **predictive need** to resolve the issue. The consequences reach far deep socially and economically. Socially, an example of a ramification would be mistrust. The abuse of power to commit misconduct will ultimately lead to distrust between the citizens and the police officers, where citizens will no longer have trust in the police officers, who are supposed to ensure peace in the society. With regards to the African-American community, police mistrust is prevalent due to the many decades of tensed relations and hence are hesitant to call upon the police when help is needed (Rick Jervis, 2020). Additionally, they will be less cooperative towards a police officer’s orders. Economically, it results in resources used for social instability such as protest and strikes. Individuals may choose to be absent from work to join such protests. Subsequently, the government will have to assign more police resources towards controlling protests and the impending violence.

Next, the presence of **imperfect knowledge** is prevalent due to the controversial nature of this issue and even with numerous studies done, there has yet to be any conclusive findings that can entirely support the stands. According to a study conducted on 990 police fatal shootings using data compiled by The Washington Post in 2015 (Nix et al., 2017) suggested that black people who were fatally shot by police, as compared to white people, seemed to be twice as likely to be unarmed. In another study investigating the risk of being killed in the United States by use of force based on age, race–ethnicity, and sex (Lynne Peeples, 2020), it found that black men were 2.5 times more likely to be killed by police during their lifetime than white men. As a result, this has made the team curious to want to find out more, in hopes to identify whether systemic racism was a key factor of police misconduct.

Our team acknowledges that in pursuit of a solution to implement appropriate strategies and to hit the key objectives stated in Section 1.3, large amounts of data will be required. This considerable demand for data can be resolved with the **availability of training data** from New York City’s past complaint cases of police misconduct over 35 years. By using this dataset and our predictive model, the issue of police misconduct can be further mitigated.

## 1.3 Key Objectives

Our team proposes a **2-pronged approach** involving an analysis of significant variables (root causes) of high rates of police misconduct, followed by the development of a prediction model that predicts whether a police officer has high risk of committing misconduct (Figure 1).

Prior to our 2-step approach, in Section 2.2, adequate data cleaning was conducted to ensure that the accuracy of the models developed will not be compromised. Similarly, variables which were irrelevant for analysis or may have high multicollinearity with other variables were eradicated. Data anomalies were further rectified, and missing values were replaced accordingly. As more insights were discovered during the analysis, data cleaning was conducted iteratively. In Section 2.3, data exploration was performed to acquire preliminary understanding of the cleaned dataset. Our team utilized a variety of charts and plots, including stacked bar charts, box plots, scatterplots and correlation matrices.

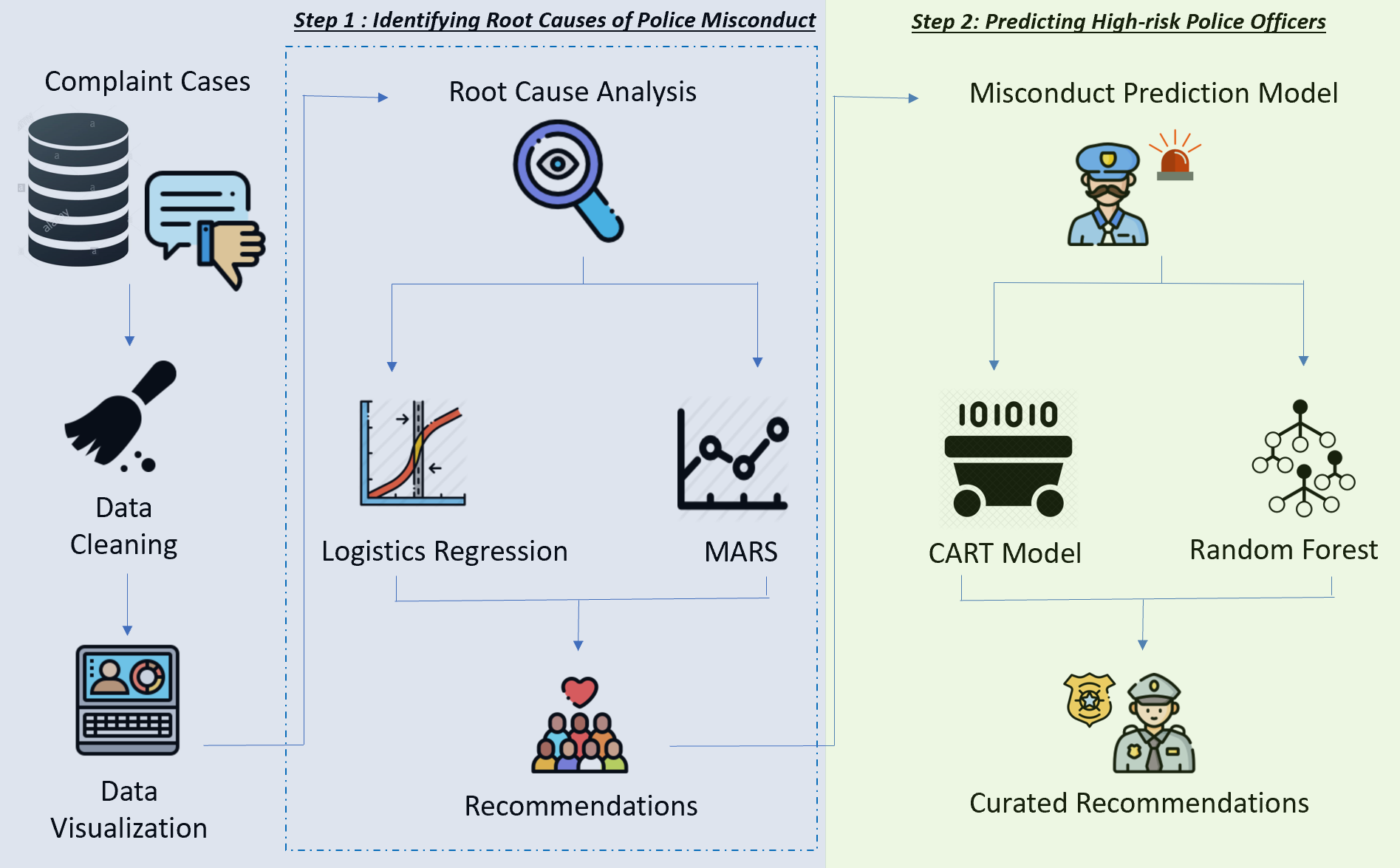
**

Figure 1: 2-pronged approach to mitigate high rates of police misconduct in US

In Section 3, to fully internalise the root causes of police misconduct, predictive models with high explainability, including Logistic Regression and Multivariate Adaptive Regression Spline (MARS), were utilized to identify key predictors for police misconduct. This offered critical insights to the key attributes that result in police misconduct, and emphasized on new variables or insights which were left unmentioned in current literature. With the key factors identified, our team proceeded to recommend solutions to specifically address the root causes or attributes identified, in order to help lower the high rate of police misconduct. This ultimately answers our key questions as to which variables are vital in concluding whether a police misconduct will occur, and whether police misconduct is truly a mere act of systemic racism. By performing root cause analysis, we provided holistic recommendations that address these key factors, and in effect, reduce the rates of police misconduct.

Next, in Section 4, our team has developed a prediction model to predict which risk level police officers will belong to in terms of misconduct. The outcome variable was based on the disposition of complaint investigation by CCRB - either misconduct is substantiated (1) or exonerated/unsubstantiated (0). In effect, the outcome variable is a binary categorical variable. Our team initially made use of multiple prediction models for analysis including logistic regression, MARS, Classification and Regression Tree (CART), as well as predictive models with higher predictive accuracy such as Random Forest and Neural Network models. Thereafter, according to the results acquired from each model, the models were compared based on a variety of metrics, and evaluated to determine the model which will serve the best purpose for addressing our issue on high police misconduct.

Finally, with our chosen model, our team will be able to predict whether a police officer is at risk of committing misconduct, and implement preventive and corrective measures curated to identified high-risk police officers early, in order to more effectively curb the high rates of police misconduct.

# 2 Data Exploration

## 2.1 Data Acquisition

The dataset is provided by New York City’s Civilian Complaint Review Board (CCRB), an official independent agency that investigates the complaints against NYC police officers before forwarding the findings to the police commissioner. Based on this dataset containing 35 years’ worth of civilian complaints about police officers in New York City, our analysis will revolve around critical information from each complaint case. These information include key details of the complainant (victim) and police officer, reasons for interaction and misconduct, outcome of the interaction as well as disposition (outcome) of each complaint case. The outcome as to whether police misconduct was indeed committed will be determined by CCRB.

## 2.2 Data Cleaning

Upon realization that missing data values existed in our dataset, iterative functions were created to replace missing categorical values with the mode and missing continuous values with the median. Additionally, two new critical columns were created, namely ‘Misconduct’ and ‘Repeated\_complaints’. The binary column ‘Misconduct’ was generated based on the original column ‘Board\_disposition’, where values of ‘substantiated’, and ‘unsubstantiated’ or ‘exonerated’ in *Board\_disposition* would be generated as ‘yes’ and ‘no’ in Misconduct respectively. This is supported by our key assumption that misconduct is only committed when CCRB’s disposition of the complaint is substantiated as the alleged conduct occurred has violated New York City Police Department’s (NYPD) rules. Extensive studies have provided insights that police officers with a history of repeated civilian complaints are more likely to commit serious misconduct in the future, regardless of the type of complaints received (Terrill & Ingram, 2015). As such, we have also created a binary ‘Repeated\_complaints’ column, where police officers were marked with the value ‘Yes’ in *Repeated\_complaints,* if their unique identification number under the column ‘Unique\_mos\_id’ appeared more than once.

Next, particularly in the ‘Complainant\_age\_incident’ column, erroneous values surfaced where there were complainants below the age of 3. To cope with these errors, these incorrect values were replaced with the median age of the complainants. In both ‘Complainant\_ethnicity’ and ‘Complainant\_gender’ columns, ambiguous values such as ‘refused’, ‘not described’ and ‘gender non-conforming’ were classified into a new category, ‘Unknown’. Finally, multi-categorical columns with numerous levels were re-categorised, such that the minority categories were collectively classified under a new factor, ‘Other’. These columns include ‘Command\_at\_incident’, ‘Command\_now’, ‘Contact\_reason’ and ‘Rank\_abbrev’.

Hence, the final dataset used for our analysis consists of 30 columns and 33359 rows, where each row represents the details of a complaint case sent in by a complainant.

## 2.3 Data Visualization

After cleaning the data, the next steps lie in conducting data visualizations to draw preliminary observations and highlight insightful findings from the dataset.

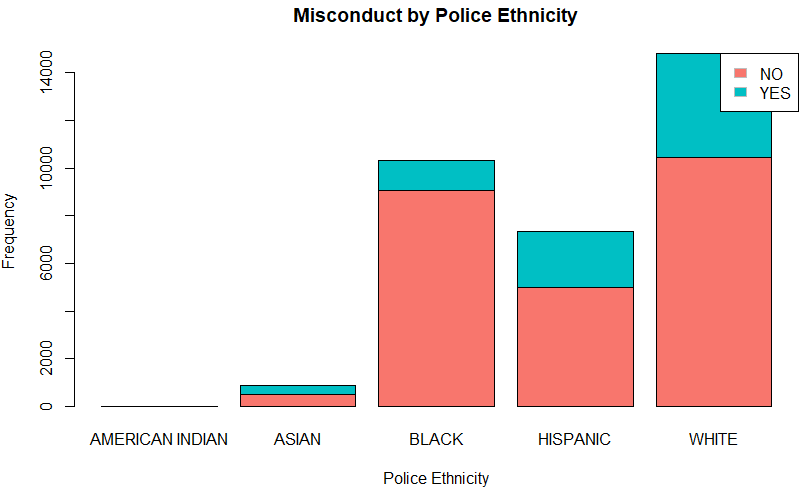
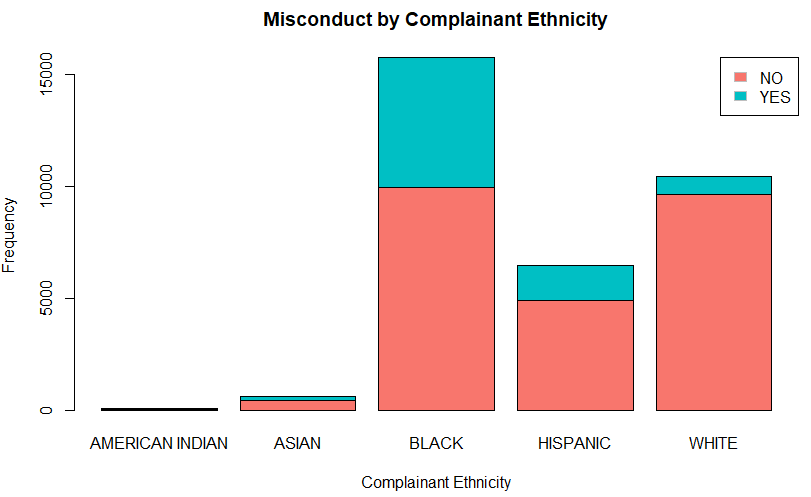


Figure 2: Misconduct by Complainant’s Ethnicity (left) and Police’s Ethnicity (right)

First, our team has utilized a stacked barchart to explore interesting observations regarding complainants’ ethnicity and misconduct. Based on the left side of Figure 2, it is observable that Black complainants make up the majority of the police complaints. As compared to other complainant ethnicities, a larger proportion of cases filed by Black complainants are found to have been substantiated by board disposition as misconduct. This further suggests that underlying racism could be present, since misconduct appears to be more apparent among Black victims. From this, our team hypothesizes that police misconduct is more likely to occur with Black victims.

When plotting a stacked barchart for police’s ethnicity and misconduct (right side of Figure 2), it becomes apparent that the majority of police complaints were directed towards White police officers. Additionally, a larger proportion of these cases were found to have been substantiated by board disposition as misconduct, when compared to other police ethnicities. As such, our team hypothesized that White cops are more likely to commit misconduct, more often than officers of other ethnicities.

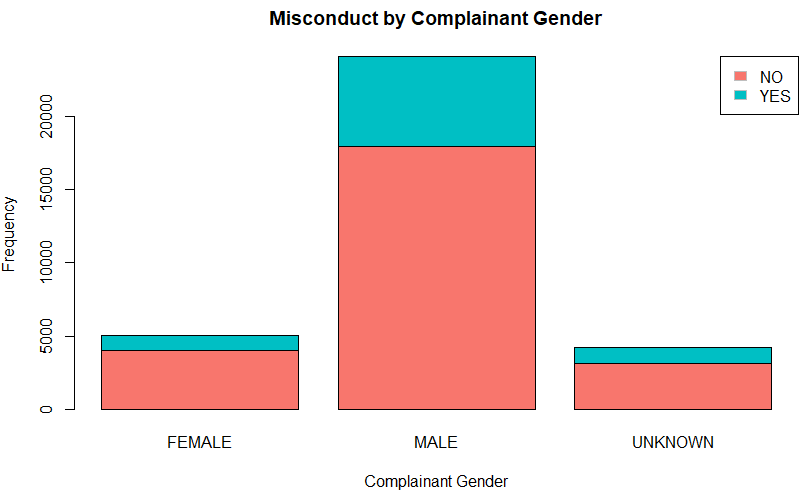


Figure 3: Misconduct by Complainant’s Gender (left) and Police’s Gender (right)

With reference to the stacked barchart for complainant’s gender and misconduct (left side of Figure 3), it is observable that most police complaints were submitted by male victims. In fact, as compared to females, a larger proportion of cases by male complainants are found to have been substantiated by board disposition as misconduct. Therefore, our team hypothesized that police misconduct is more likely to occur with male victims. Nonetheless, the higher occurrence of police misconduct experienced by male complainants could possibly be due to the higher crime rates by males as opposed to females.

Similarly, based on the stacked barchart for police gender and misconduct (right side of Figure 3), it is notable that the majority of police complaints are made against male cops. As a hypothesis, our team observed that male cops are more likely to perform police misconduct. However, it should be recognized that the disproportionate effect could be due to a naturally higher number of male officers being employed, as compared to female police officers.

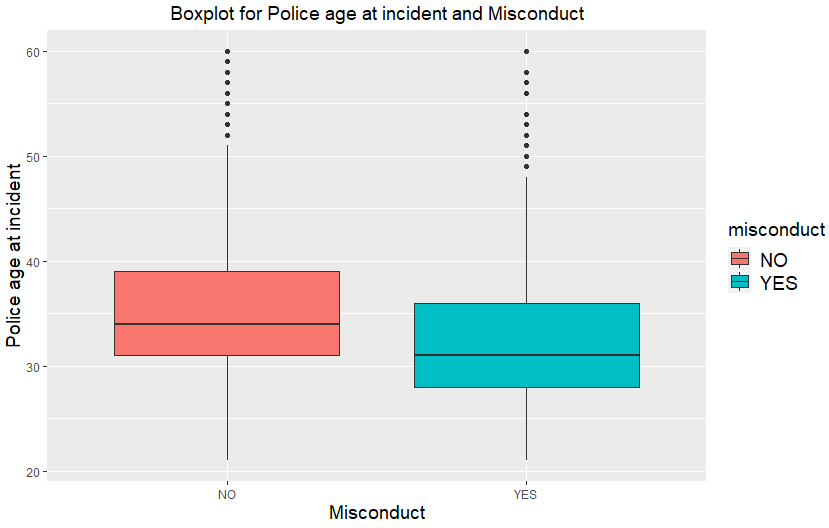


Figure 4: Misconduct by Police’s Age

According to Figure 4, our team used comparative boxplots to observe trends between police’s age during the incident and misconduct. It can be observed that for cases where misconduct was committed, the 25th percentile, median and 75th percentile of police’s age is substantially lower than for cases where misconduct was not conducted. This suggests that younger police officers might be more inclined to commit misconduct. This could inevitably be due to a lack of experience and appropriate training.

# 3 Root Cause Analysis of Police Misconduct

## 3.1 Data Modelling

With the aim to identify significant predictors of police misconduct, several models with high degree of explainability were used. These models included logistic regression and multi-adaptive regression splines (MARS).

### 3.1.1 Logistic Regression

Our team first identified that the outcome variable, *Misconduct*, is binary in nature. Following that, **logistic regression** was performed and based on the p-values, significant variables were identified. Generally, variables with p-values lower than 0.05 were identified to be significant. Additionally, this was verified by identifying the odds ratio confidence interval of each variable. If the odds ratio confidence interval excludes 1, this further shows that the variable is statistically significant. The results of the logistic regression is displayed in Appendix 6.

In addition, our team utilized a **backward elimination algorithm** to further verify which significant variables should be included in the model. Even though the selection algorithm can iteratively eliminate the least significant variables and pick out candidates of the model variables, it is noteworthy that the results of backward elimination should not be the sole representative of the best model. Instead, it should complement the p-values acquired from logistic regression of the full model and expert opinions and domain knowledge as well. The results of the backward elimination algorithm is shown below in Appendix 7.

### 3.1.2 Multi-Adaptive Regression Splines (MARS)

As we understand that utilizing one sole model may not be fully representative of the results, our team explored another analytical model, **MARS**, to harness further insights. The models for both 1st and 2nd degree MARS were performed. In particular, our team chose the 1st degree MARS model as the better model to identify significant predictors. This is because both the logistic regression model and the 1st degree MARS model showed similar statistical significance. Based on the 1st degree MARS model, important variables were identified according to three main metrics: the number of model subsets that include the variable (Nsubsets), the net decrease in Generalised Cross Validation (GCV), as well as the net decrease in Residual Sums of Squares (RSS). Utilizing the *‘evimp’* function, the more important variables were determined below in Figure 5.

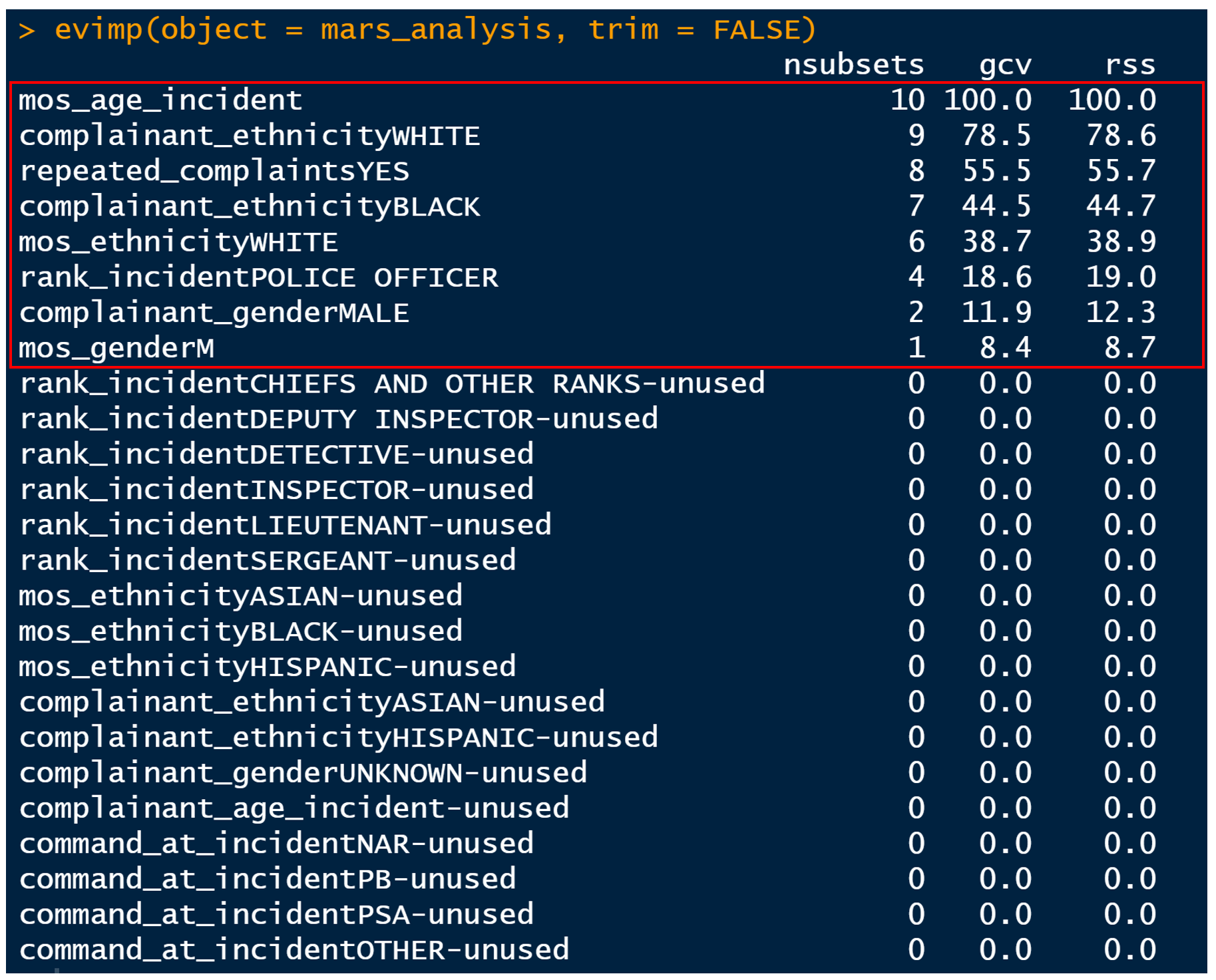


Figure 5: Identified significant variables using 1st degree MARS

## 3.2 Evaluation of Models

Based on both models, logistic regression and 1st degree MARS, our team identified **6 key predictor variables** that appeared to be statistically significant in causing police misconduct. Figure 6 illustrates the variables that are identified to be significant through the analysis of initial p-values and odds ratio confidence intervals through logistic regression, performance after backward elimination algorithms, analysis of important variables through MARS and further corroboration with research and expert opinions.

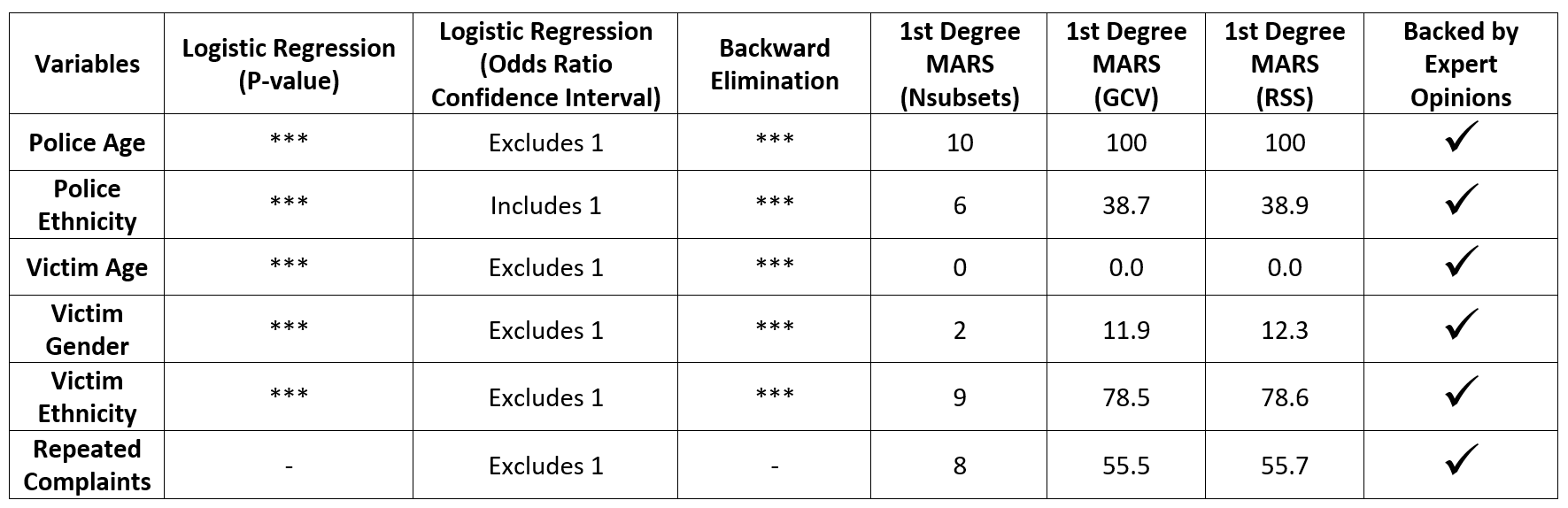


Figure 6: Identified significant variables across models and research

## 3.3 Root Cause Analysis

Based on our findings, we conducted thorough research and performed a root cause analysis on the 6 predictor variables that caused police misconduct.

### 3.3.1 Police Age

Police age, which refers to the age of the police officer during the time of the incident, was identified as one of the root causes for police misconduct. Through data exploration, it is evident that younger police officers are generally more likely to perform misconduct. Additionally, in a separate study done on the Network Structure of Police Misconduct in 2019, it showed that a white male officer appointed in 2005 who is 35 years old can expect to receive 0.70 civilian complaints in a year. Conversely, a similar officer of age 45 can expect to receive 0.39 civilian complaints in a year, which accounts for lesser complaints a year (Wood, Rithmayr and Papachristos, 2019). This aligns with our model towards the significance of Police Age being a root cause.

### 3.3.2 Police Ethnicity

Next, Police Ethnicity is identified as one of the root causes for police misconduct. Deriving insights from our visualisations, white police officers are at a higher risk of police misconduct, as compared to officers of other ethnicities. Referring to the same study done on Network Structure of Police Misconduct in 2019, it is conclusive that White officers are more likely to receive at least one complaint than Black or Hispanic police officers (Wood, Rithmayr and Papachristos, 2019). This can be attributed to the fact that more White police officers conceive and stereotype black people to be more violent than white people, as compared to Black officers and Americans overall. (Bacon, 2020). Another contributing factor is White police officers generally use more force as compared to Black and Hispanic officers, where the use of force per officer is 1.31, 1.24 and 1.28 for White, Hispanic and Black officers respectively. (Hauck and Yancey-Bragg, 2021). These factors concur with the significance of Police Ethnicity as a root cause to predict police misconduct.

### 3.3.3 Victim Age

Victim Age at the time of the incident is also identified as one of the root causes for police misconduct. In a study done on the demography of Police-Involved Homicides, it is shown that people aged between ages 20 to 29, experience the highest rates of police-involved killings as compared to all age groups. (Esposito, Edwards and Lee, 2020). Therefore, this is aligned to Victim Age being a significant root cause to predict police misconduct.

### 3.3.4 Victim Gender

Additionally, a root cause derived for police misconduct was Victim Gender. As observed from data exploration, male victims are more likely to be involved in misconduct cases. Cross referencing this to a research done by Proceedings of the National Academy of Sciences of the United States (PNAS), it corroborated that over the course a lifetime, with conditions observed over 2013 to 2018, about 52 in every 100,000 men and boys in the US are killed by the police, while 3 in every 100,000 women and girls are predicted to be killed by the police (Esposito, Edwards and Lee, 2020). Despite the fact that the gender ratio for the entire population in the United States stands at a ratio of around 1:1 (Ritchie & Roser, 2019), the number of males killed by police is substantially higher. As such, it is apparent that Victim Gender remains a significant root cause of police misconduct.

### 3.3.5 Victim Ethnicity

Victim Ethnicity is identified as one of the root causes for police misconduct. Black victims form a big bulk of the misconduct cases, as referenced to our data visualisations, which would imply that Black victims have a higher chance of getting involved in misconduct cases. A report by ProPublica stated that young black males are 21 times more likely to be killed by police as compared to young White males. (Wihbey and Kille, 2016). Another research study published by Rutgers School of Criminal Justice shows that Black men are 2.5 times more likely to be killed by police as compared to White men, while Black women are 1.4 times more likely to be killed by police in comparison to White women. (Edwards, Lee and Esposito, 2019). Moreover, another empirical research showed that Black people are at a 3 to 4 times higher risk of getting involved in less-than-lethal contact with police and they are also 3 times more likely to be involved in cases of police contact where force is considered as excessive (Motley and Joe, 2018).

These studies undoubtedly suggest that Black people are more likely to fall as victims of police misconduct. This may be attributed to the fact that Black people are already racially ostracized and perceived by White police officers as being violent (Bacon, 2020). With prejudice towards Black citizens, cases of misconduct will be more likely to happen. In a study done by Ravi Shroff, an assistant Professor at NYU Steinhardt and NYU CUSP, findings conveyed that upon getting stopped by police officers for a check, Black people are being searched 1.5 to 2 times more often than White people, while having lower probabilities of carrying illegal items (NYU, 2020). With Black people being searched more compared to people of other races, it sheds light on the underlying biases exhibited by police officers against Black people. Hence, with supported research, Victim Ethnicity is of significance for being a root cause.

### 3.3.6 Repeated Complaints

Lastly, Repeated Complaints appeared to be one of the root causes for police misconduct as well. Our data has shown that police officers who have histories of police misconduct have greater tendencies to commit misconduct again. This is reinforced by studies, showing that police officers who have a history of complaints, regardless of the type of complaints, are typically at a higher risk of performing serious misconduct in the future. (McCorkel, 2020). This ultimately aligns with our model which shows that Repeated Complaints is a significant root cause of police misconduct.

## 3.4 Limitations

Upon further analysis, our team realises that limitations do emerge. It is unquestionable that there may be other significant variables which are absent from our dataset. We acknowledge that behavioral and psychological variables could further explain or be identified as a root cause for police misconduct. For instance, the working hours of police officers could be a crucial variable that can be used in the case of police misconduct. In a study completed to observe the stability of implicit racial bias in police officers, it is found that when officers had less sleep before the testing was done in this study, they portrayed stronger association between Black people and weapons. (‌James, 2017). This shows that a police officer who had lesser sleep, possibly from late-night shifts, would have heightened implicit racial bias.

Another limitation lies in the lack of data on the psychological aspect of police misconduct, such as a police officer’s exposure to violence. A study has shown that prolonged exposure to violence can increase aggressive, violent practices and thinkings within police officers. This explains the use of excessive force as the number one reason for police misconduct cases. (Warren, 2015). Consequently, it brings insights that police officers working departments that deal with a higher number of violent cases are more likely to perform misconduct since they are exposed to relatively more violence.

## 3.5 Recommendations

Upon identifying the root causes, several recommendations can be ascertained to target the root causes of the occurrence of police misconduct. As aforementioned, the root causes are Police Age, Police Ethnicity, Victim Age, Victim Gender, Victim Ethnicity and Repeated Complaints.

### 3.5.1 Scenario-based Hiring and Trainings

Understanding that specific police attributes, namely age and ethnicity, are root causes for a higher tendency of police misconduct, hiring processes can pay special attention to these key attributes. Instead of completely eradicating the hiring of younger police officers or white police officers, more stringent hiring processes can be implemented to target the issue at its core. For instance, scenario-based hiring can be incorporated into their hiring processes. In these cases, candidates will be placed in various scenarios to evaluate their responses and actions. The responses and actions performed by the participants will be evaluated with deciding criteria, which helps to conclude whether the police officer should be hired. The assessor will be further entrusted with the duty of taking note of younger and white police officers who may have a tendency to make more mistakes and bad decisions in such tests. Consequently, potential candidates who commit unusually high mistakes should not be hired.

Scenario-based training can also be extended to original police officer training (Basham, 2019). These scenarios can utilise two mediums of role-playing and tabletop exercises. Typically, role-playing engages trainers and police trainees to actively participate by acting out their responses to several appropriate scenarios. On the other hand, the nature of tabletop exercises is more discussion-based. In such exercises, participants will verbalise their action, with respect to the scenario and context provided. Subsequently, the trainers will run through their decision-making and ethical aspects of the scenario.

As noted, these training scenarios should encompass a variety in the age, gender and ethnicity of the victim. This is necessary as it targets the identified root causes that results in victims being more likely involved in a police misconduct. Ultimately, more scenarios should be included as more root causes are identified in the future.

### 3.5.2 Psychological Training

Helping younger police officers and white police officers raise the awareness as to why they respond or behave in a particular way is essential. Psychological training undeniably fulfils this role. Psychological insights will help both police officers and their supervisors figure out behavioral aspects to focus on. Objectively, psychological insights should be geared towards justice and fairness. By tapping onto this approach, public trust in the police force can be readily fostered (Abrams, 2020).

Additionally, hired psychologists can emphasize on resolving implicit biases amongst police officers. Implicit bias refers to having certain assumptions or unconscious association towards any social group, and thereby, resulting in stereotyping. A real example can be referenced to which Oakland Police Officers treated Black people with less respect because of their background, upon reviewing the footage from their body cameras (Voigt et al., 2017). Eventually, with consistent sessions conducted by psychologists, police officers will be able to eliminate implicit biases, which raises the fairness and objectivity of police officers’ actions and decisions made towards a particular demographic group.

### 3.5.3 Familiarisation

Patrolling around local neighbourhoods assigned to police officers will help ease the root causes. According to statistics from Pew Research Center, 72% of respondents said it is important for a police officer to have good knowledge of the people, places, and culture. Conversely, only about 3% say that neighbourhood patrol is unnecessary (Mitchell, 2017).

In particular, patrolling around the neighbourhood appears to be more essential for white officers. Based on the above statistics, fully 84% of black officers and 78% of Hispanic officers believe that it is crucial to harness knowledge of their local neighbourhoods whereas only 60% of white officers feel the same way. This indicates a difference in mindsets and a probable lack of knowledge of their local neighbourhoods by white officers. Therefore, familiarity of local neighbourhoods should be an emphasis for white officers as well, since they are significantly more likely to commit police misconduct.

### 3.5.4 Community Campaigns

This recommendation entails fostering trust in the police force among the community. By having people from the public to actively participate in community meetings, police-sponsored youth activities and community safety fairs that allow interaction between the public and police officers would unquestionably develop trust.

More specifically, specially organised campaigns that encourage members of the public to role-play in scenarios that police officers could possibly face would help nurture deeper understanding and hone stronger ties between both parties. For example, a young, white individual from the public could participate in role-play sessions, involving actors emulating the role of a 30 year old, black, and male victim. Through such activities, the individual volunteer partaking in this activity will be able to better understand tough and quick decisions often made by police officers, as well as raise the awareness that the community can play a role in preventing police misconduct from happening.

# 4 Prediction Model for High-risk Police Officers

## 4.1 Data Modelling

With the aim to predict whether a police officer is at risk of commiting police misconduct, several models will be utilised. In particular, our team introduced two new models - **Classification and Regression Trees (CART) and Random Forest**, which have higher degree of predictability. Even though it should be acknowledged that **Logistic Regression and MARS** are not as high in predictability as the tree models, we continued to utilize them as predictive models to comprehensively exhibit the difference in predictive accuracies amongst all four models, before selecting the one that best meets our needs.

Notably, there are some differences in how the manner logistic regression and MARS were performed as compared to Section 3. In this section, our team selectively used the 6 variables that were deemed significant and important from Section 3. Subsequently, our team continued to perform a train-test split, with 75% of the data as train set and the rest as test set. Even though a train test split is not necessary for CART and random forest, we still fed only the train data to the two models, as we want to perform a fair comparison between the models. As such, since all the models are fed with the same data, proper evaluation of the model’s performance can be conducted.

Additionally, because of the overwhelmingly larger proportion of non-misconduct cases as compared to misconduct cases, the models have a natural tendency to favour the majority category. Hence, to rectify this, Synthetic Minority Oversampling Technique (SMOTE) was run on the train data to over-sample the minority data on the train set, such that the number of misconduct and non-misconduct cases are the same. In doing so, we are able to increase the classification accuracy. However, when measuring train set and test set accuracy, it should be recognised that the train and test set data used should remain untouched by SMOTE, so as to evaluate how well the model performs in real-world data, with the original proportions of misconduct kept intact.

### 4.1.1 Logistic Regression

Logistic Regression is a statistical model which uses a logistic function to model a categorical dependent variable. One advantage of the logistic regression model is that it is easy to implement and interpret the model coefficients to indicate variable importance.

For the logistic regression model, the following 6 variables, which have been found to show significance in predicting whether a case may result in police misconduct, were used.

1. mos\_age\_incident (Age of the involved police officer during the incident)
2. mos\_ethnicity (Ethnicity of the involved police officer)
3. complainant\_age\_incident (Age of the complainant during the incident)
4. complainant\_gender (Gender of the complainant)
5. complainant\_ethnicity (Ethnicity of the complainant)
6. repeated\_complaints (Whether the involved police officer has repeated complaints)

The summary of the resultant model containing the respective model coefficients are shown under Appendix 8.

Next, the odds ratio of each variable displayed in the left side of Figure 7 is obtained by computing the exponential of the model’s variable coefficients.

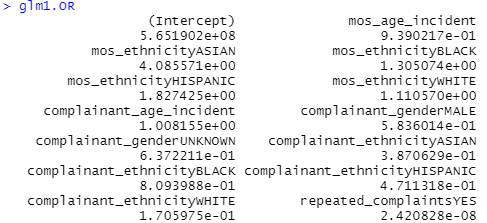
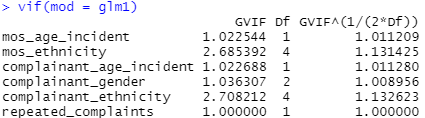
 

Figure 7: Odds ratios (left) and VIF (right) from logistic regression model

Based on the results, statistical interpretation can be derived. For instance, complainants who are Black are times more likely to be involved in a substantiated police misconduct case. Our team has further verified the potential of any multicollinearity problems arising within the model, by utilizing the VIF function. In our logistic regression model, since all of the VIF found is less than 5 (right side of Figure 7), it is conclusive that there is no multicollinearity issue in our model.

|  |  |
| --- | --- |
| Figure 8: Confusion matrix of  Logistic Regression Model results (Train set) | Figure 9: Confusion matrix of  Logistic Regression Model results (Test set) |

From the respective confusion matrices for the train set (Figure 8) and test set (Figure 9), the **model accuracy for the train set is 63.63%** with a true positive rate of 70.29% and a true negative rate of 61.43% whereas the **model accuracy for the test set is 63.75%** with a true positive rate of 70.60% and a true negative rate of 61.48%.

### 4.1.2 MARS

MARS has hinge functions, which is able split the data into sections so as to have a more localised prediction. We used the following parameters for MARS (Appendix 11):

* degree = 1
* nfold = 1, ncross = 1
* glm = list(family = binomial)), for categorical Y

Our team performed the model at higher degrees, as well as higher nfold and ncross values. However, the improvements in the results were insignificant. As such, lower values were utilized so that the model can be built faster and more efficiently.

All 12 terms were selected, along with 9 of the 13 predictors. The model concluded that the most important variable is the complainant’s ethnicity, followed by repeated complaints. Using the same train set and test set above, the results are as displayed below.

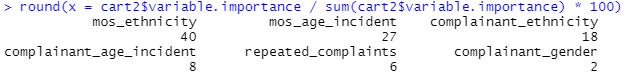
|  |  |
| --- | --- |
| Figure 10: Confusion matrix of  MARS Model results (Train set) | Figure 11: Confusion matrix of  MARS Model results (Test set) |

From the respective confusion matrices for the train set (Figure 10) and test sets Figure 11), the **model accuracy for the train set is 64.28%** with a true positive rate of 66.90% and a true negative rate of 63.42% whereas the **model accuracy for the test set is 64.58%** with a true positive rate of 67.90% and a true negative rate of 63.47%.

The accuracy for both train and test set is similar to that of logistic regression. This is likely due to the fact that the parameter, glm = list(family = binomial)) is used. As stated in package earth’s documentation, “if the response is binary or a factor, consider using the glm argument.”. As such, we are using the generalized linear models to do prediction, which is similar to logistic regression.

### 4.1.3 CART

CART is a model that uses a machine learning algorithm to derive a resulting decision tree where each fork is a split in a predictor variable and each end node is a prediction for the outcome variable. To generate the predictive model, our team used all the same variables earlier from the logistic regression model and first grew the tree to its maximum by setting minimum split = 2 and complexity parameter (cp) = 0. Subsequently, we found the optimal cp and pruned the cp by choosing the model with the optimal cp.

  
Figure 12: Variable Importance of Predictors in CART model

By checking the variable importance of each independent variable, it was found that police ethnicity has the greatest variable importance of 40%, followed by police age at incident at 27% and complainant ethnicity at 18% (Figure 12).

|  |  |
| --- | --- |
| Figure 13: Confusion matrix of CART Model results (Train set) | Figure 14: Confusion matrix of CART Model results (Test set) |

From the respective confusion matrices for the train set (Figure 13) and test set (Figure 14), the **model accuracy for the train set is 88.39%** with a true positive rate of 71.49% and a true negative rate of 93.99% whereas the **model accuracy for the test set is 86.40%** with a true positive rate of 68.53% and a true negative rate of 92.32%.

From these findings thus far, the accuracy using the CART model is far more superior as compared to the logistic regression model. Furthermore, it is observable that an improvement appears not only in the true positive rate, but in the true negative rate as well.

### 4.1.4 Random Forest

Random forest is a classification algorithm that consists of many decision trees that are combined to outperform a single decision tree in terms of predictive accuracy. Using bootstrap aggregating of both cases and the predictor variables, it is able to reduce biases and the influence of a strong predictor variable, resulting in a much better model.

However, one disadvantage is that it is a black box model, which insinuates that it is difficult to gain meaningful insights from the variables that the model finds significant. After all, a random selection of features and the massive number of trees make it even more difficult to understand and explain the model. However, since the focus lies in achieving higher predictive accuracy rather than higher explainability, random forest appears to be a good choice, as it is likely to outperform CART in terms of classification accuracy.

The following parameters were used for random forest:

* ntree = 500
* mtry left as default

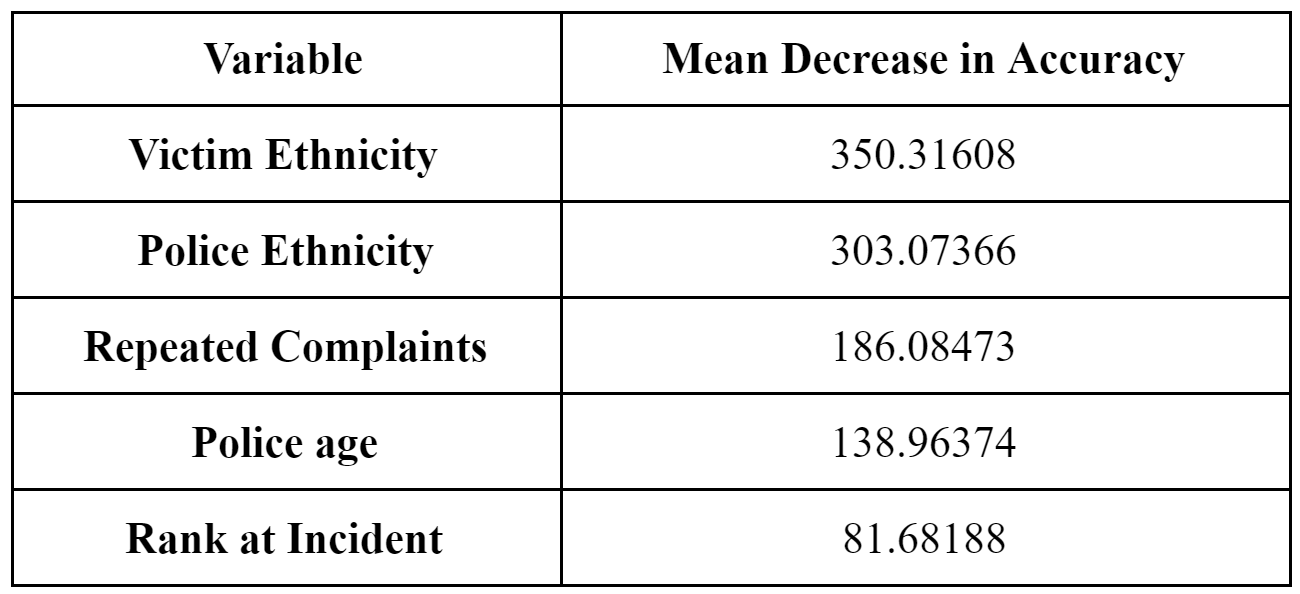
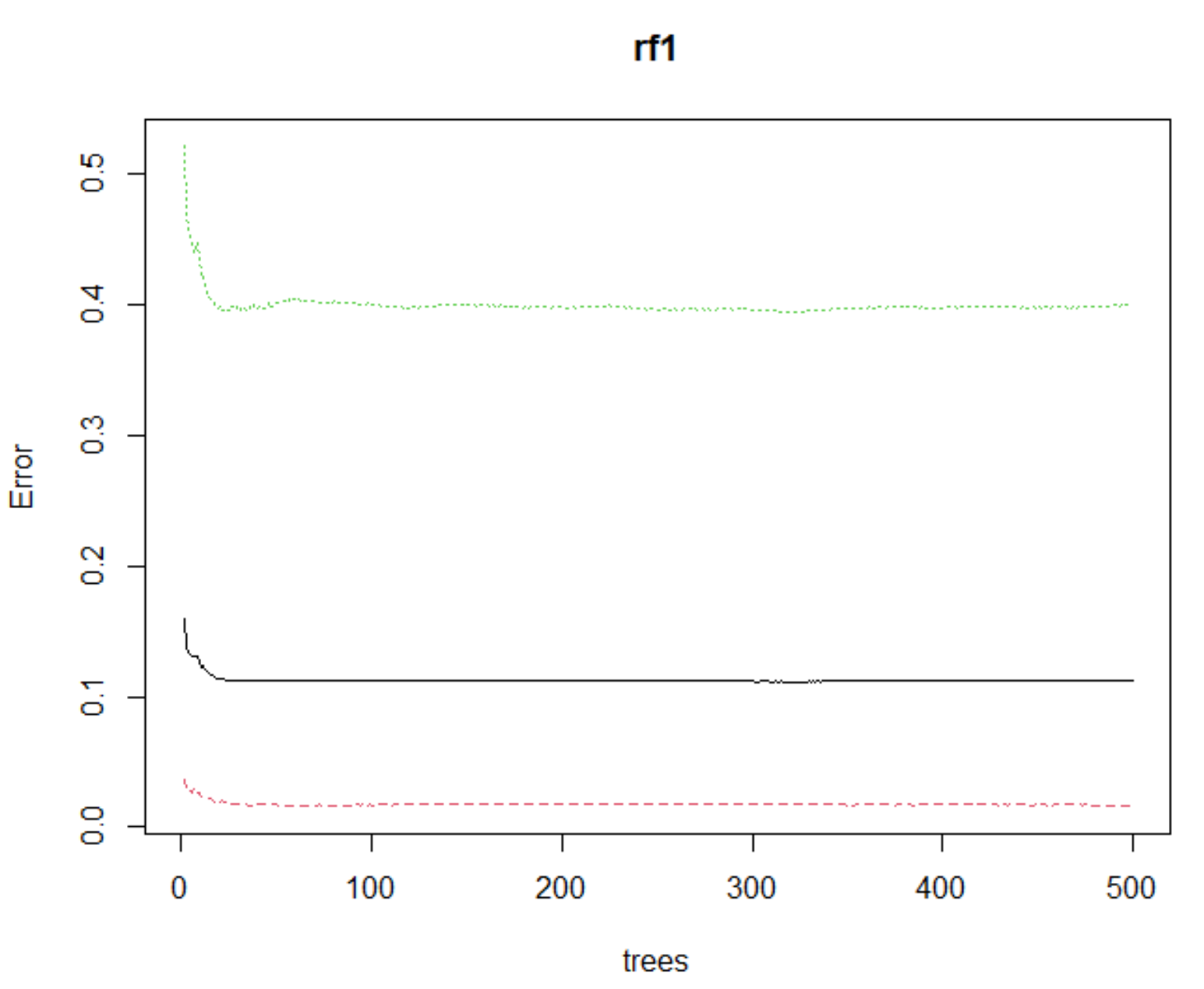


Figure 15: OOB error rates (left) and Variable Importance using Mean Decrease in Accuracy (right)

The OOB error rates were plotted using plot(x = rf1), and the results are shown in the left side of Figure 15. As observed, the error rate seems to stabilise in about 40 trees, showing that 500 trees is more than sufficient for the model. Before training the model, our team further ran random forest on 9 predictor variables to check on the variable importance, and the results are displayed on the right side of Figure 15.

Victim ethnicity is deemed most important, followed by police ethnicity. This clearly shows how ethnicity plays a very big part as determined by the random forest model. It uses the mean decrease in accuracy to determine the relationship between the predictors and response variable. As such, if the predictor variable is omitted, this will result in a large increase in error, hence deeming the predictor variable as important.

Using the same train set and test set above, the results are as displayed below.

|  |  |
| --- | --- |
| Figure 16: Confusion matrix of Random Forest Model results (Train set) | Figure 17: Confusion matrix of Random Forest Model results (Test set) |

From the respective confusion matrices for the train set (Figure 16) and test set (Figure 17), the **model accuracy for the train set is 89.54%** with a true positive rate of 61.60% and a true negative rate of 98.80% whereas the **model accuracy for the test set is 88.99%** with a true positive rate of 60.77% and a true negative rate of 98.34%.

## 4.2 Evaluation of Models

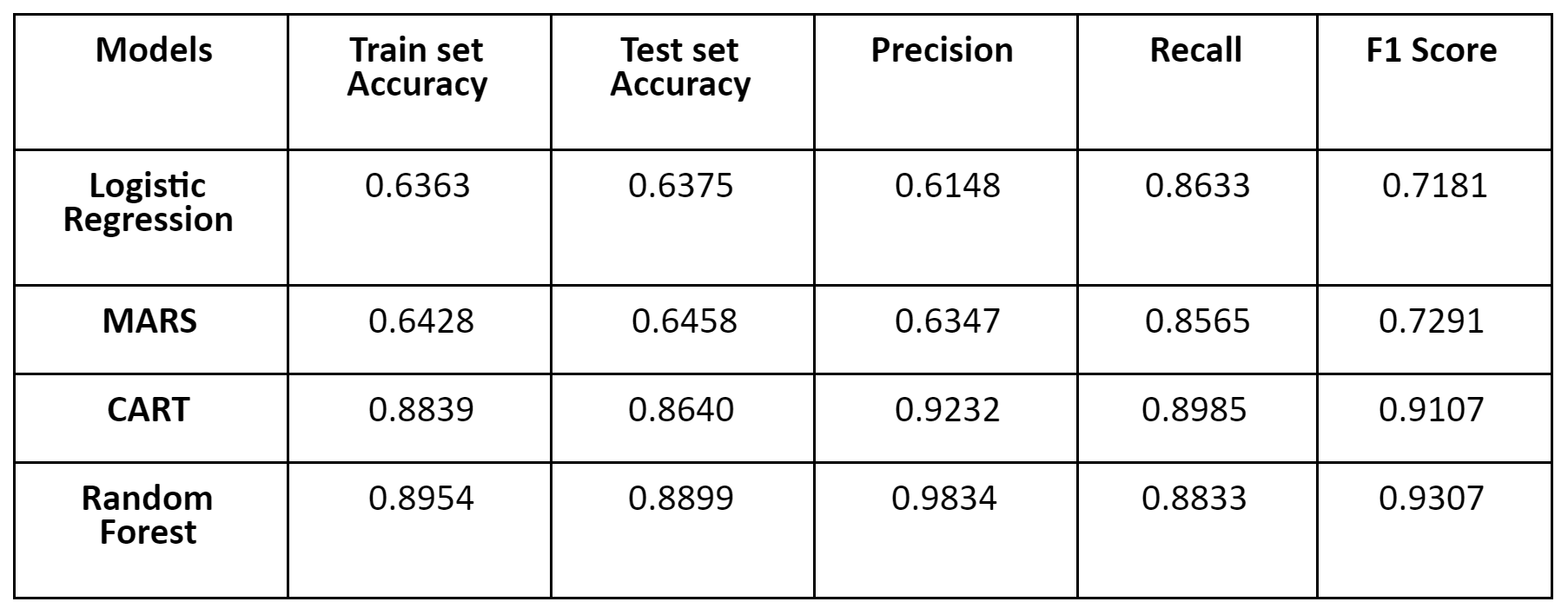
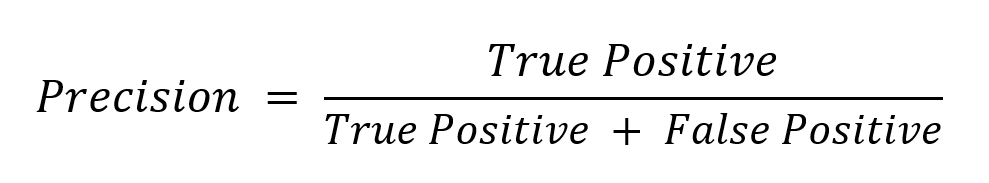
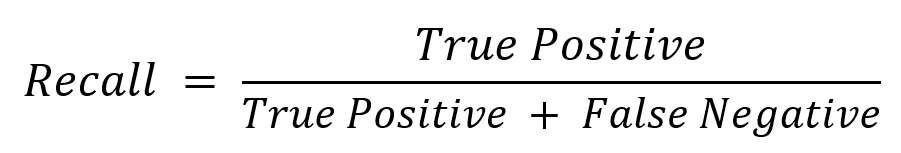


Figure 18: Performance Metrics of Models

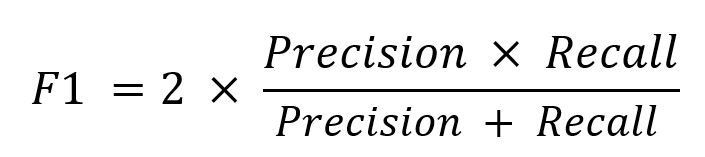
Referring to Figure 18, a comprehensive comparison of both train and test set accuracies for the 4 different predictive models has been performed. It is observable that the Random Forest has the highest train and test set accuracy. However, confusion matrix and accuracies should not be the sole accuracy to decide on the best model to use for our purpose. In our specific context, False Negatives appear to incur the highest societal cost because it involves a situation where the model predicts that a police officer is not at high risk of committing misconduct but in actual fact, is at high risk and requires curated measures. Therefore, it is critical for us to take into account other accuracy metrics such as Precision, Recall and the F1 Score that incorporates False Negatives.



First, we incorporated precision as part of the few accuracy metrics because it evaluates the accuracy of predicting actual positives, amongst those that are positive predictions. With a high precision, the chosen model will have a higher accuracy in predicting police officers who will actually commit misconduct, over all the officers who are predicted to commit misconduct.



As for the recall metric, it was included in our analysis as False Negative cases are associated with higher cases of misconduct not being prevented. If the model were to predict that a police officer will not commit misconduct but in actual fact, will likely commit misconduct, detrimental societal consequences could arise. As such, it is pivotal to utilize this metric to assess which model has the highest percentage of results correctly classified.



With the precision and recall accuracy score computed, F1 score can be computed as well. The F1 score is the harmonic mean of both recall and mean (Koehrsen, 2018). This acts as an accuracy metric that balances precision and recall. Our model’s accuracy could be largely contributed by a large number of True Negatives which is not as important compared to False Negatives and False Positives which have higher societal costs. Therefore, by including the F1 score, our team was able to have an accuracy metric to evaluate our model accuracy, especially in a case where uneven class distribution (large number of True Negatives) is present.

All in all, our team has decided that the **Random Forest** model is deemed as the best model to predict whether a police officer will likely commit misconduct. Not only does the model have a high train and test set accuracy, it has the best results in terms of all the other performance metrics, including precision, recall and F1 score, as seen in Figure 18.

## 4.3 Limitations

Our team recognises that data collection mistakes, as well as sampling bias, could be present in our dataset. Since this dataset only contains the cases where there have been a complaint, our dataset could be skewed towards only police officers who have received a complaint before. However, it should be acknowledged that a more representative dataset that includes police officers who have yet to receive a complaint would better serve our predictive need.

Moreover, that key assumption that utilizing SMOTE will increase classification accuracy without having any potential consequences have been made for our model’s proof of concept. After conducting further research, our team recognises that this assumption may not hold true in the real world. As such, the reliance of this synthetic data manipulation technique poses a limitation to our analysis because synthetic data can never completely replace real data acquired, since there could be gaps within our analysis that we have failed to consider.

## 4.4 Recommendations

Dealing with officers with high risk of misconduct is a sensitive issue with programs mostly leading to corrective measures. In order to fully actualise the results provided by our predictive model, highly curated and relevant recommendations that encompasses both preventive and corrective measures have to be implemented and targeted to police officers, who are identified to be at higher risk of committing police misconduct. Most of these recommended measures revolve around discipline matrices taken from the Harvard Kennedy School and National Institute of Justice (Stephens, n.d.), by either making them more transparent or re-emphasizing their value towards the identified high-risk police officers.

### 4.4.1 Creating transparent disciplinary matrices

The purpose of creating disciplinary matrices as a preventive strategy is to allow every active duty police officer to be cognizant of the consequences of each offense and the degree of harm done when handling various situations. Allowing the disciplinary matrices to be transparent and known throughout will eliminate ambiguity when performing rules of engagement towards various parties. Each action will have a consequence associated with it and this will invoke police officers to think twice and abide strictly by the appropriate amount of force given in each situation. In fact, reemphasising these matrices from time to time will be a constant reminder for high-risk police officers to keep their level of force in check. An example of a disciplinary matrix is shown in Appendix 20.

### 4.4.2 Education-based discipline

As the name suggests, this corrective measure creates a change of police officers’ actions and behaviours through education rather than through a punitive method. Although punitive measures may lead to behavioural change, these measures may create resentment in police officers when they think it is duly unfair for them to face such discipline. This could eventually worsen if the discipline matrix is not properly conveyed or transparent enough. Essentially, when a police officer exhibits traits or nuances of behaviour which are signs that they may be at high risk of committing misconduct, they will be placed through an educational program which is less demeaning and more accepting, which leads to more ease towards behavioural change. Ultimately, this further prevents them from committing police misconduct when they are to be on active duty again.

### 4.4.3 Problem-solving approach

Problem-solving provides a more logical approach towards disciplining high-risk police officers in the perspective of higher-ups in the department. One of the more common methods of problem-oriented policing with regards to discipline is the SARA model created by researchers from the Police Executive Research Forum (PERF) and officers from Newport News Police Department (Braga, 2008). The SARA model (Appendix 21) consists of the following steps:

* Scanning – Identifying and Selecting problems for further study.
* Analysis – Breaking the issue down and tackling it with various perspectives.
* Response – Creating appropriate responses in relation to the analysis.
* Assessment – Assessing if the response has the desired impact.

### 4.4.4 Innovations of excessive force training

Identified high risk police officers can be assessed by psychologists by putting them into situations of adverse conditions or highly charged situations to see how they perform (‌Scrivner, 1994). Scenarios can involve the testing of their cultural sensitivity, if they would intervene to stop excessive force, interactions between citizens, decision-making under adverse conditions, defusing techniques, and conflict resolution. The list is non-exhaustive, and testing should be used when deemed fit.

Additionally, depending on the element that results in the police officers high risk of misconduct, various psychological treatments can be administered. For instance, the individual police officer who has a high tendency of hostility will be put through an anger management program which allows them to be aware of their own anger and how to handle and manage it. In effect, such psychological training could be incorporated as a complementary tool with excessive force training.

# 5 Conclusion

Recalling the 2-pronged approach, the use of logistic regression and MARS helped us to identify significant variables that can be used to predict police misconduct in the US while random forest was the selected model that predicted police officers at high risk of committing misconduct. These analytical solutions provide insights on measures that the police force, government and community can leverage on to tackle this pressing issue. On the whole, the government should adopt a stricter approach towards the issue by improving the quality of training for police officers and nurturing the trust between police officers and the general public. In doing so, this will lead to promising economic and social benefits for the US in the long run, while eradicating potential future detrimental societal repercussions.

A possible improvement to our analysis would be to obtain a dataset tailored to understand key predictors that lead to police misconduct in different geographical regions. By stepping across geographical boundaries and looking beyond New York City, our analysis could further examine possible different social and cultural influences that might exist in varying cities. Moreover, future analyses should also reduce the reliance on synthetic data manipulation techniques and leverage on newly found and updated resources that could lead to the betterment of research in this field.

Ultimately, it is imperative to further evaluate and improve the long-run feasibility of the proposed recommendations. As some changes are relatively large-scale, ease of implementation and receptivity plays a significant part for success. Along the way, analytics can be leveraged for the recording of KPIs for each recommendation. Additionally, there should be systems in place to accept new changes, should there be more root causes or findings identified in the future.

In essence, our models and proofs of concept aim to offer insights on the root causes of police misconduct and provide curated recommendations to reduce the overall high rates of police misconduct prevalent in the US.

# References

Abrams, Z. (2020). What works to reduce police brutality. Retrieved 27 March 2021, from <https://www.apa.org/monitor/2020/10/cover-police-brutality>

Bacon, P. Jr. (2020). How The Police See Issues Of Race And Policing. Retrieved 27 March 2021, from <https://fivethirtyeight.com/features/how-the-police-see-issues-of-race-and-policing/>

Basham, S. (2019). Education Offers Best Solution for Police Misconduct. Police Chief Magazine. Retrieved 27 March 2021, from <https://www.policechiefmagazine.org/education-offers-best-solution-for-police-misconduct/>

Braga, A. (2008). PROBLEM-ORIENTED POLICING AND CRIME PREVENTION 2nd edition. Retrieved 27 March 2021, from https://popcenter.asu.edu/sites/default/files/braga\_pop\_intro\_2.pdf

Edwards, F., Lee, H., & Esposito, M. (2019). Risk of being killed by police use of force in the United States by age, race–ethnicity, and sex. Retrieved 27 March 2021, from <https://www.pnas.org/content/116/34/16793/tab-figures-data>

Esposito, M., Edwards, F., & Lee, H. (2020). The Demography of Police-Involved Homicides. Retrieved 27 March 2021, from <https://www.asanet.org/news-events/footnotes/jul-aug-2020/features/demography-police-involved-homicides>

Hauck, G. & Yancey-Bragg, N. (2021). Police don't all act 'the same way': White officers use force more often, Chicago police study finds. Retrieved 27 March 2021, from <https://www.usatoday.com/story/news/nation/2021/02/11/chicago-police-study-white-cops-more-likely-arrest-use-force/4439259001/>

‌James, L. (2017). The Stability of Implicit Racial Bias in Police Officers. Police Quarterly, 21(1), 30–52. https://doi.org/10.1177/1098611117732974

Koehrsen, W. (2018). Beyond Accuracy: Precision and Recall. Retrieved 27 March 2021, from https://towardsdatascience.com/beyond-accuracy-precision-and-recall-3da06bea9f6c

McCorkel, J. (2020). Police officers accused of brutal violence often have a history of complaints by citizens. The Conversation. Retrieved 27 March 2021, from <https://theconversation.com/police-officers-accused-of-brutal-violence-often-have-a-history-of-complaints-by-citizens-139709>

‌Mitchell, T. (2017). 3. Police and the community. Pew Research Center’s Social & Demographic Trends Project; Pew Research Center’s Social & Demographic Trends Project. Retrieved 27 March 2021, from <https://www.pewresearch.org/social-trends/2017/01/11/police-and-the-community/>

Motley, R. O. & Joe, S. Jr. (2018). Police Use of Force by Ethnicity, Sex, and Socioeconomic Class. Retrieved 27 March 2021, from <https://cpb-us-w2.wpmucdn.com/sites.wustl.edu/dist/7/497/files/2018/07/Police-Use-of-Force-by-Ethnicity-Sex-and-Socioeconomic-Class-1h3fybv.pdf>

NYU (2020). Research Shows Black Drivers More Likely to Be Stopped by Police. Retrieved 27 March 2021, from https://www.nyu.edu/about/news-publications/news/2020/may/black-drivers-more-likely-to-be-stopped-by-police.html

Ritchie, H., & Roser, M. (2019). Gender Ratio. Retrieved 27 March 2021, from <https://ourworldindata.org/gender-ratio>

Stephens, D. (2011). New Perspectives in Policing Police Discipline: A Case for Change Executive Session on Policing and Public Safety. Retrieved 27 March 2021, from <https://www.ojp.gov/pdffiles1/nij/234052.pdf>

‌Scrivner, E. M. (1994). Controlling Police Use of Excessive Force: The Role of the Police Psychologist. Retrieved 27 March 2021, from<https://www.ojp.gov/pdffiles1/Digitization/150063NCJRS.pdf>

Terrill, W., & Ingram, J. R. (2015). Citizen Complaints Against the Police. Police Quarterly, 19(2), 150-179. doi:10.1177/1098611115613320

‌Voigt, R., Camp, N. P., Prabhakaran, V., Hamilton, W. L., Hetey, R. C., Griffiths, C. M., Jurgens, D., Jurafsky, D., & Eberhardt, J. L. (2017). Language from police body camera footage shows racial disparities in officer respect. Proceedings of the National Academy of Sciences, 114(25), 6521–6526. https://doi.org/10.1073/pnas.1702413114

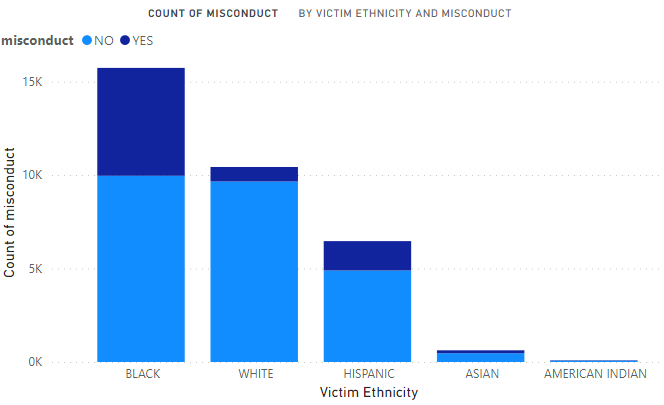
Warren, T. (2015). The Effects of Frequent Exposure to Violence and Trauma on Police Officers. Retrieved 27 March 2021, from <https://scholarworks.waldenu.edu/cgi/viewcontent.cgi?article=2328&context=dissertations>

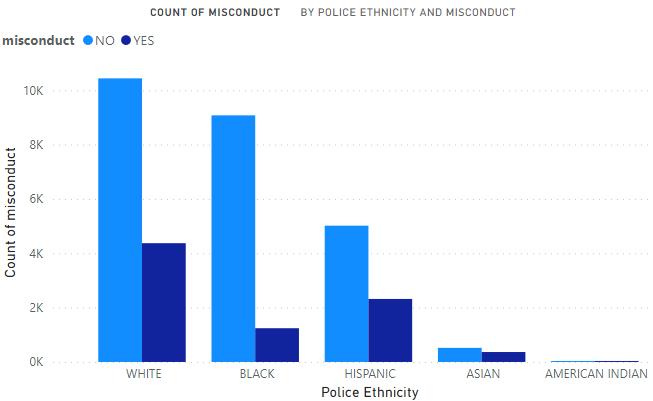
[Wihbey](https://journalistsresource.org/author/john-wihbey/), J. & [Kille](https://journalistsresource.org/author/leighton/), L. W. (2016), Excessive or reasonable force by police? Research on law enforcement and racial conflict. Retrieved 27 March 2021, from <https://journalistsresource.org/criminal-justice/police-reasonable-force-brutality-race-research-review-statistics/>

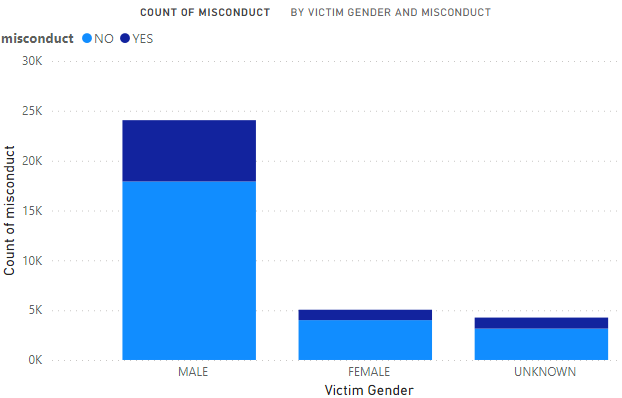
Wood, G., Rithmayr, D. & Papachristos, A. V. (2019), The Network Structure of Police misconduct. Retrieved 27 March 2021, from <https://www.asanet.org/sites/default/files/attach/journals/oct19sociusfeature.pdf>

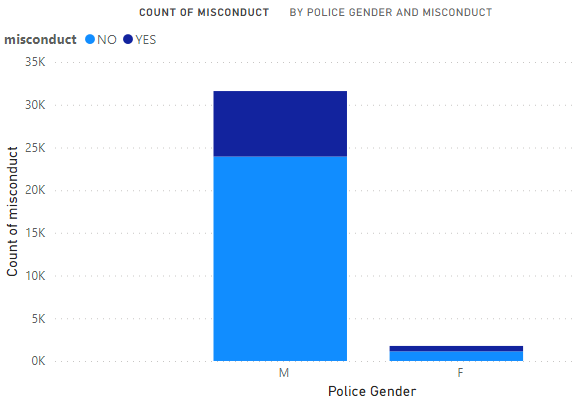
# Appendices

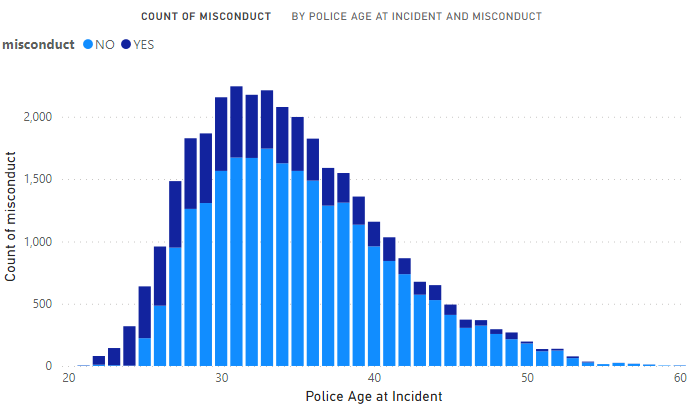
Appendix 1: Stacked barchart of Victim’s Ethnicity and Misconduct



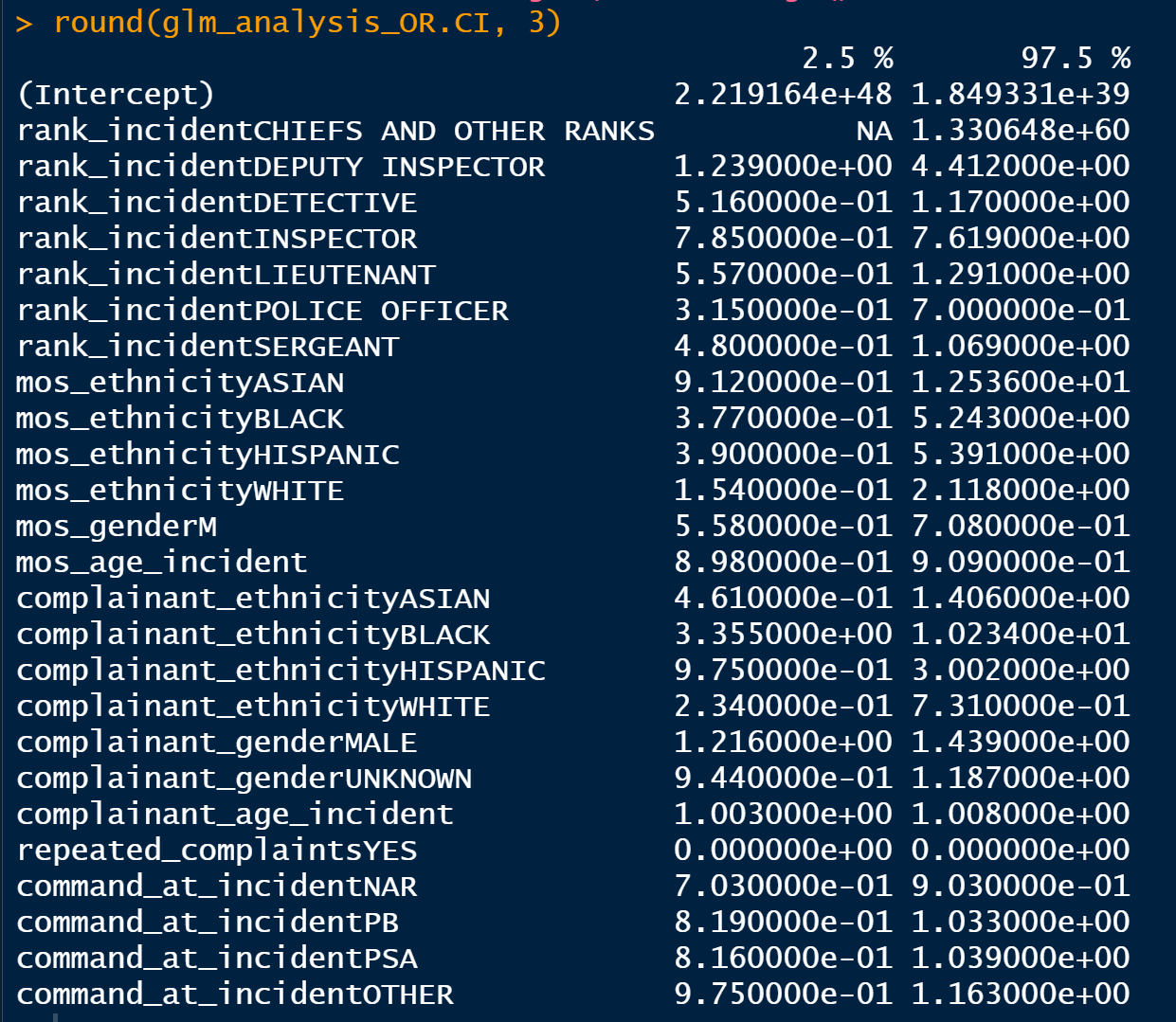
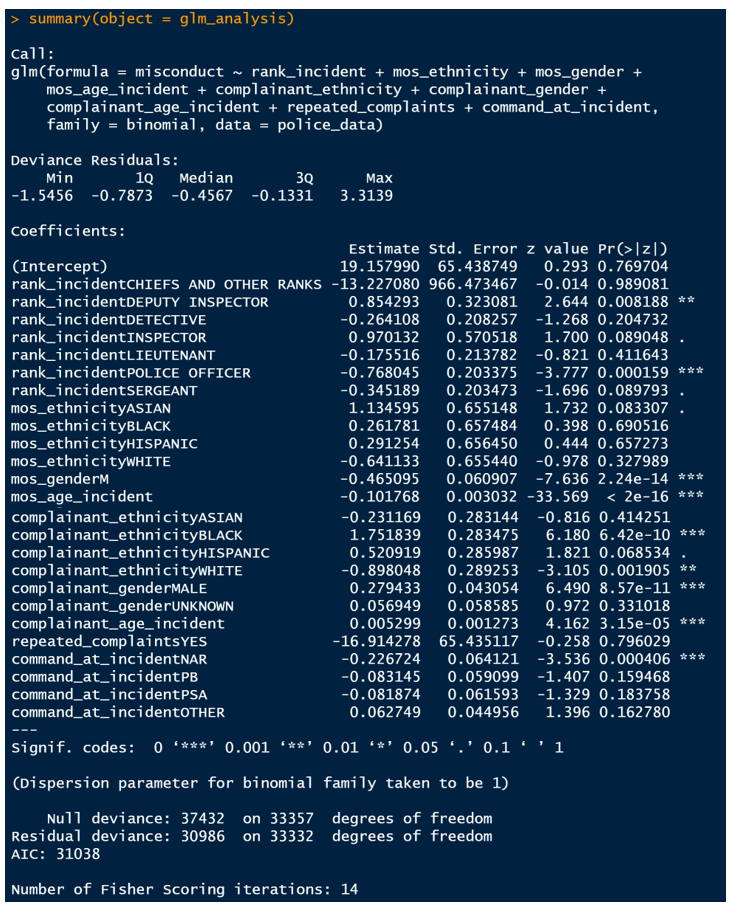
Appendix 2: Stacked barchart of Police’s Ethnicity and Misconduct

Appendix 3: Stacked barchart of Victim’s Gender and Misconduct

Appendix 4: Stacked barchart of Police’s Gender and Misconduct

Appendix 5: Density Distribution Chart of Police Age and Misconduct

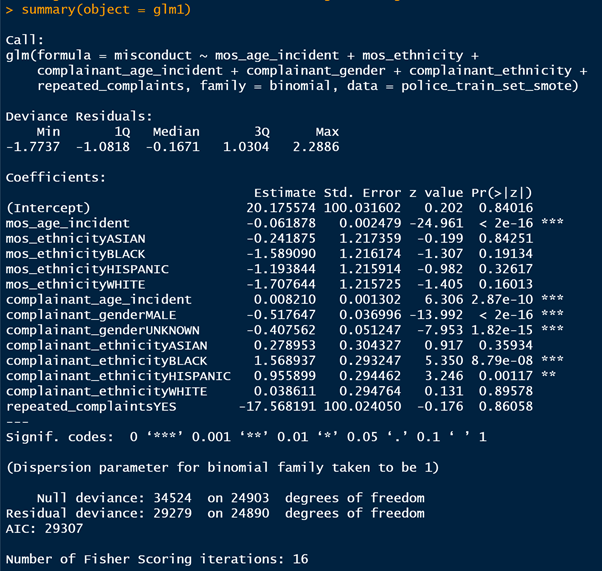
Appendix 6: Identified Significant Variables using Logistic Regression (P-values and Odds Ratio Confidence Interval



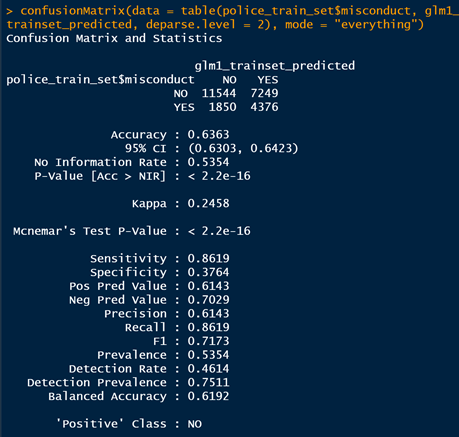
Appendix 7: Identified Significant Variables using Backward Elimination Algorithm



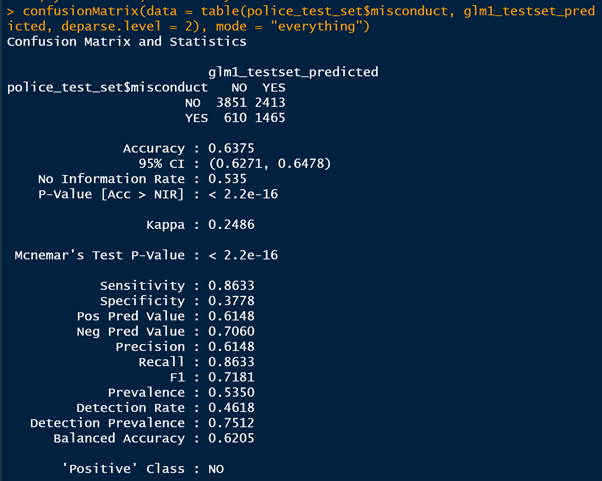
Appendix 8: Logistic Regression Output



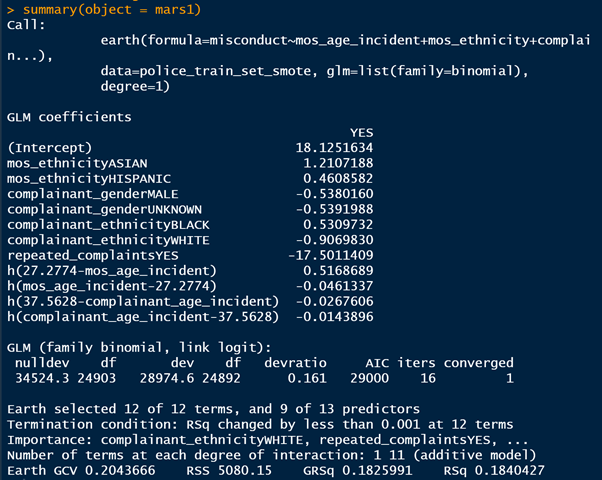
Appendix 9: Logistic Regression Train Set Accuracy



Appendix 10: Logistic Regression Test Set Accuracy and Other Performance Metrics



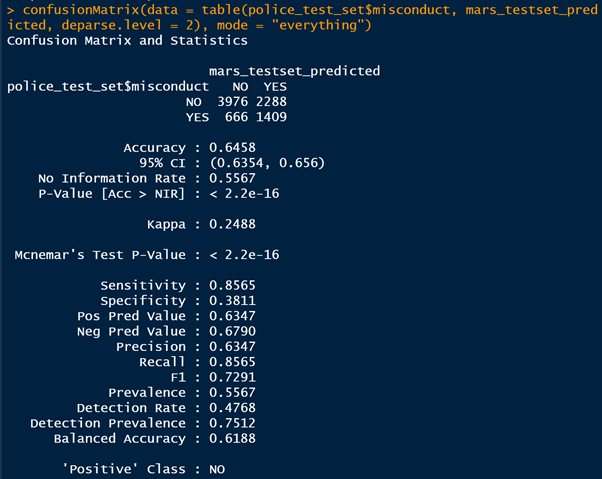
Appendix 11: MARS Output



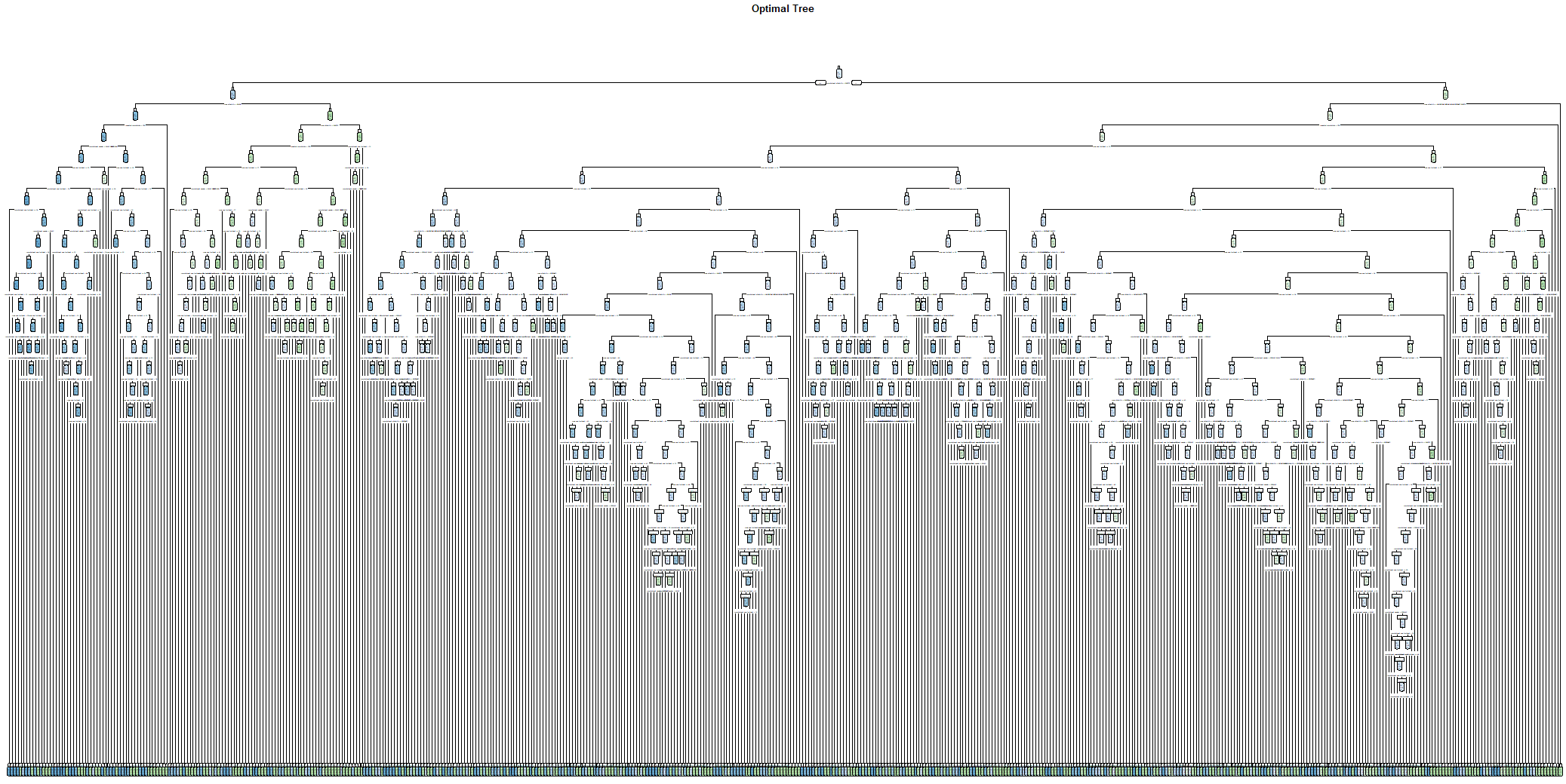
Appendix 12: MARS Train Set Accuracy



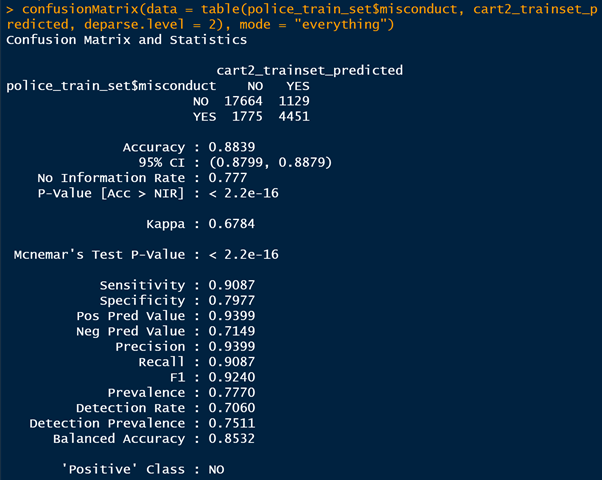
Appendix 13: MARS Test Set Accuracy and Other Performance Metrics



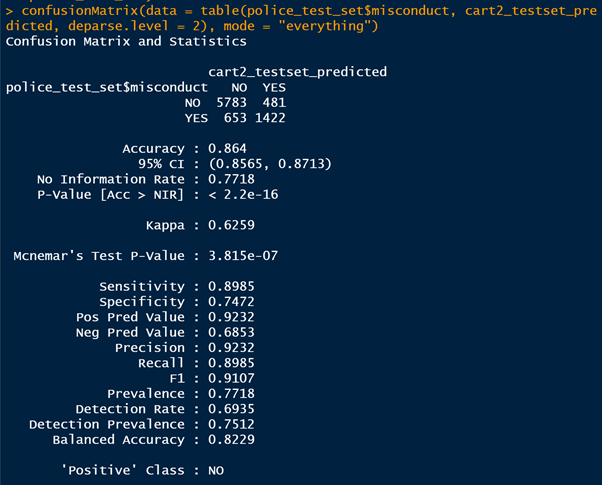
Appendix 14: Optimal Tree for CART Model



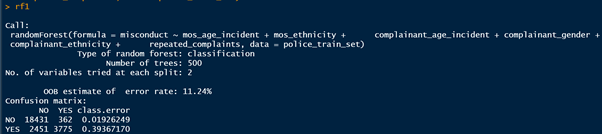
Appendix 15: CART Train Set Accuracy

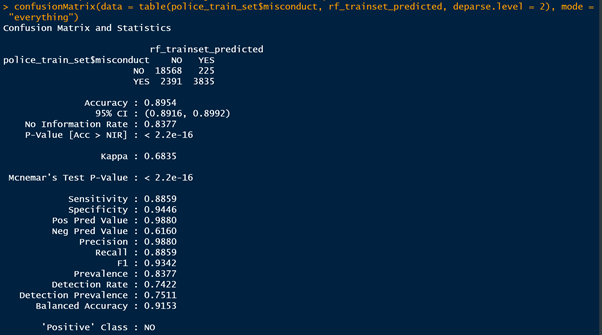


Appendix 16: CART Test Set Accuracy and Other Performance Metrics

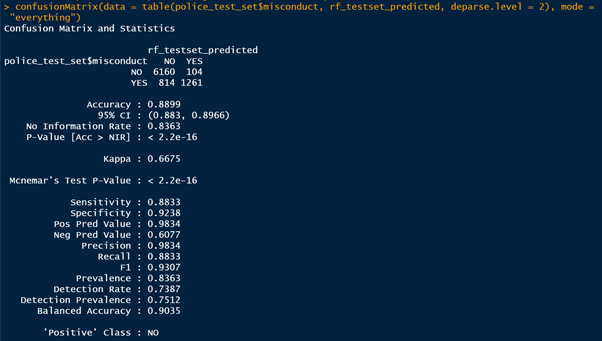


Appendix 17: Random Forest Output

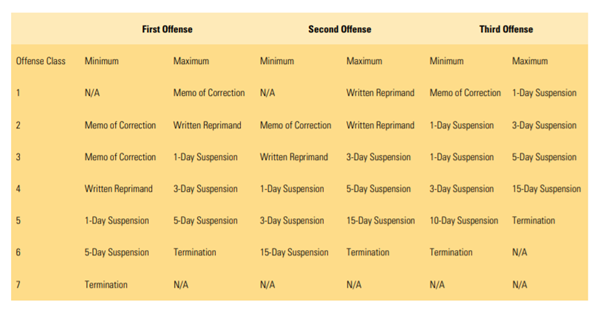


Appendix 18: Random Forest Train Set Accuracy

Appendix 19: Random Forest Test Set Accuracy and Other Performance Metrics



Appendix 20: Example of Discipline Matrix



Appendix 21: SARA Model

