Data Analysis on Arrests and Strip Searches

Chenyang Huan

Yueran Yang

INF2178H LEC 0101

Experimental Design for Data Science

Shion Guha

April 16, 2023

Introduction

In Canada, crime and violence are consistently considered as one of the top-of-mind concerns among Canadians, with only the unemployment rate being a greater concern (Government of Canada, 2022). Therefore, the government is devoted to addressing and preventing crime and violence problems as these are widely recognized as significant priorities. Nevertheless, when it comes to policing practices, some critics argue that police officers exhibit a higher propensity to arrest Black civilians than White civilians, especially for minor offences, when faced with evidence of illegal activity (Wortley & Jung, 2020). In light of the importance of equality in the criminal justice system, we will examine the relationships between different race groups and the number of arrests in our project.

Since arrests and searches are related to a wide range of criminal offences, analyzing the "Arrests and Strip Searches" dataset could provide valuable insights related to crime and violence in the city of Toronto. By analyzing the "Arrests and Strip Searches" dataset, we can identify potential biases or disparities in the criminal justice system, for example, whether a particular racial group is more likely to be arrested or whether certain racial groups are more likely to be targeted for strip searches. We hypothesize that the analysis of the "Arrests and Strip Searches" dataset will help policymakers and law enforcement agencies determine how to enhance fairness in law enforcement practices and equality in the criminal justice system. For our analysis, we will use the 2020-2021 arrests and strip searches dataset provided by Toronto Police Service.

The aim of this research project is to investigate the potential impact of racial and sex disparities on the number of arrests while controlling for the number of booked strip searches. Furthermore, the project aims to analyze the potential influence of the ArrestLocDiv and Occurance Category on the likelihood of individuals being cooperative during an arrest. Finally, the research project seeks to examine whether an individual's perceived race affects the likelihood of being booked for a strip search while taking into account their level of cooperation.

Through this project, we hope to assess the broader implications of these disparities on future interactions with the criminal justice system. By shedding light on these issues, we hope to contribute to a more equitable and just system.

Background Information

Tensions between the police and certain ethnic groups have been a longstanding issue in major Canadian cities, with numerous reports of discriminatory practices and excessive use of force against racialized communities. The inquiry initiated in Ontario in 1993 was a response to these issues and aimed to investigate allegations of systemic racism in the criminal justice system. The inquiry found evidence of discrimination against racialized communities, including disproportionate rates of arrest, harsher treatment in court, and over-representation in the prison system (Roberts et al, 1997).

During our midterm project, we conducted analyses that revealed significant differences in the mean number of arrests and the likelihood of being subjected to a strip search among different racial groups. These results underscore the importance of investigating potential inequalities within the criminal justice system to ensure that all individuals are treated fairly and justly. Further research in this area is needed to better understand and address these disparities.

Literature Review

For decades, there have been calls to acknowledge and address systemic racism in policing. According to a study, Canadian official crime statistics reflect the fact that Black people have a higher level of engagement in criminal activity compared to other racial groups. They are particularly over-represented in drug-related offences, gang activity and street-level violence. The study suggests that the disproportionate deployment of police in Black communities, results in Black people being subject to higher levels of police surveillance compared to other racial groups. Due to strict surveillance, Black offenders are more likely to be identified and arrested in comparison to other offenders (Wortley & Jung, 2020). Through an analysis of over 10,000 drug possession arrests conducted by the Toronto Police Service from 1996 to 2001, another study revealed that Black individuals (38%) were significantly more likely than White people (23%) to be taken to the police station for drug possession (Rankin et al, 2002).

In addition, booked arrests are formal records of arrests that are processed and fingerprinted at a local detention facility. Along with criminal convictions, they are the essential components of a criminal history record. These records are used by law enforcement to conduct future risk assessments, and in some cases, they may be accessed by non-criminal justice actors, such as employers (Raphael et al, 2019). Stolzenberg conducted a study to investigate the impact of a criminal suspect's prior record on their likelihood of being arrested. The study employed multivariate logistic regression analysis and found that suspects with a prior criminal record were significantly more likely to be arrested by police, with odds approximately 29 times higher than those without a criminal record. The study provides evidence supporting the argument that individuals with a criminal record may face increased scrutiny by law enforcement due to the negative social stigma attached to their labels. This labelling process can lead to the normalization and continuation of criminal behaviour. Individuals with a history of criminal behaviour are more likely to be detected and arrested by authorities, as law enforcement agencies maintain records of their fingerprints, DNA, photographs, personal background information, criminal associates, and modus operandi. These records allow authorities to identify and apprehend repeat offenders with greater accuracy, compared to individuals who have no criminal record (Stolzenberg et al., 2021).

Furthermore, Balfour's research indicates that women who commit crimes have distinct motivations compared to their male counterparts. Women may be more susceptible to committing crimes due to various social and economic factors that disproportionately affect them, including single parenthood, limited access to affordable childcare, living in poverty, inadequate employment opportunities, and unstable housing. These conditions can create significant challenges for women and may drive them to engage in criminal behaviour (Government of Canada, 2021). When accused and convicted, women are typically charged with less severe offences and are subject to more lenient sentencing. As a result, women are less likely to be incarcerated in comparison to men (Balfour, 2020).

Research Objective and Questions

Our study aims to investigate four research questions pertaining to the "Arrests and Strip Searches" dataset. To analyze the data effectively, we will employ t-tests to examine relevant variables and conduct a power analysis to determine the appropriate sample size. Furthermore, we plan to utilize ANCOVA and logistic regression models to gain further insight into the research questions at hand.

We will investigate the research questions that we have formulated based on our literature review and initial examination of the dataset. The section on descriptive Statistics provides an overview of our preliminary exploration of the data. The objective of our study is to address the following research question:

RQ1: Does perceived race have an effect on the number of arrests while controlling the number of booked strip searches?

RQ2: Does sex have an effect on the number of arrests while controlling the number of booked strip searches?

RQ3: Does the ArrestLocDiv and Occurance Category have an impact on the likelihood of individuals being cooperative during an arrest?

RQ4: Does an individual's perceived race have an impact on the likelihood of being booked for a strip search when taking into account their level of cooperation?

The first two research questions will be investigated through ANCOVA analysis, while the last two questions will be addressed using a logistic regression model.

Exploratory Data Analysis

Dataset Description

In our proposed project, we will use a dataset that includes information related to all arrests and strip searches that took place in the city of Toronto between 2020 and 2021. The dataset was collected by Toronto Police Service on November 2021, which contains 65,276 records and a total of 24 various attributes including arrest dates, perceived race, sex, age group, reasons for arrest, strip search, actions at arrest, search reasons, item found, etc. The dataset is available on Toronto Police Service Public Safety Data Portal and could be found through the following link: https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/about. Data is provided in either text format, in binaries 0 or 1 format or in numeric integer format. The full list of the variables in the dataset is shown in Table 1.

Table 1. The full list of the variables in the raw dataset.

Variable Name	Data Type	Variable Description
Arrest Year	int	The year the arrest took place in (2020 or 2021)
Arrest Month	str	The Quarter the arrest took place in (Jan-Mar, Apr-Jun, Jul-Sept, Oct-Dec)
EventID	int	The distinctive identifier is used to specify details of arrest incidents.
ArrestID	int	The distinctive identifier is used to specify details of arrest incidents.
PersonID	int	The distinctive identifier is used to specify details of arrest incidents.
Perceived Race	str	The race of the arrested individual
Sex	str	The sex of the arrested people
Age Group (at arrest)	str	The age group (under 17, 18-24, 25-34 35-44, 45-54, 55-64, over 65)
Youth at arrest (under 18 years)	str	Indicates whether the arrested individual is under 17 years old
ArrestLocDiv	str	The distinctive identifier is used to specify the location of arrest incidents.
StripSearch	int	Indicates whether an arrested person was strip-searched.
Booked	int	Indicates whether an arrested person was booked at a police facility within 24 hours of arrest.
Occurrence Category	str	The arrest reason.
Action at arrest - Concealed items	int	This data point indicates whether an arrested individual was uncooperative with the police during the arrest, which may be reflected in certain actions or behaviours exhibited during the incident.

Action at arrest - Resisted, defensive or escape risk	int	This data point indicates whether an arrested individual was uncooperative with the police during the arrest, which may be reflected in certain actions or behaviours exhibited during the incident.
Action at arrest - Mental instability or possibly suicidal	int	This data point indicates whether an arrested individual was uncooperative with the police during the arrest, which may be reflected in certain actions or behaviours exhibited during the incident.
Action at arrest - Assulted officer	int	This data point indicates whether an arrested individual was uncooperative with the police during the arrest, which may be reflected in certain actions or behaviours exhibited during the incident.
Action at arrest - Cooperative	int	This data point indicates whether an arrested individual was uncooperative with the police during the arrest, which may be reflected in certain actions or behaviours exhibited during the incident.
SearchReason - CauseInjury	int	Indicates the reason why an arrested individual was subjected to a strip search.
SearchReason - AssistEscape	int	Indicates the reason why an arrested individual was subjected to a strip search.
SearchReason - PossessWeapons	int	Indicates the reason why an arrested individual was subjected to a strip search.
SearchReason - PossessEvidence	int	Indicates the reason why an arrested individual was subjected to a strip search.
ItemsFound	int	Indicates whether or not any items or contraband were found on the arrested individual during the strip searches.

Figure 1. First 5 Rows of Arrests and Strip Searches Dataset

	Arrest_Year	Arrest_Month	EventID	ArrestID	PersonID	Perceived_Race	Sex	Age_group__at_arrest_	Youth_at_arrest__under_18_years
0	2020	July-Sept	1005907	6017884.0	326622	White	M	Aged 35 to 44 years	Not a youth
1	2020	July-Sept	1014562	6056669.0	326622	White	M	Aged 35 to 44 years	Not a youth
2	2020	Oct-Dec	1029922	6057065.0	326622	Unknown or Legacy	M	Aged 35 to 44 years	Not a youth
3	2021	Jan-Mar	1052190	6029059.0	327535	Black	M	Aged 25 to 34 years	Not a youth
4	2021	Jan-Mar	1015512	6040372.0	327535	South Asian	M	Aged 25 to 34 years	Not a youth

5 rows × 25 columns

Figure 1 displays a sample of the first five rows of the Arrests and Strip Searches dataset. In order to ensure the accuracy and reliability of the dataset, we performed several data-cleaning procedures to prepare the data for analysis.

To begin with, we conducted an initial check of the dataset to identify any missing values. We found that there are some missing values in "ArrestID", "Perceived_Race" "Age_group__at_arrest_", and "Occurrence_Category". Of particular concern were the 4 variables related to SearchReason, which contained over 57,000 missing values. In light of this significant amount of missing values, we would exclude those four variables from our analysis. Subsequently, we removed all the missing values in "ArrestID", "Perceived_Race" "Age_group__at_arrest_", and "Occurrence_Category" and filled all NaN values in the 4 RearchReason variables with 0 to eliminate the potential impact of these variables on our analysis.

In addition, we split the two-year timeframe (2020-2021) into eight quarters, each representing a distinct three-month interval. Although the time periods are discrete, the number of arrests in each quarter can be treated as a continuous variable for our subsequent analysis. After that, we created a dictionary to map the values in the "Occurrence_Category" to standardized categories and replace the original values in the column with their corresponding standardized categories. For instance, we replaced "Police Category - Administrative" and "Police Category - Incident" with "police". By combining those values into a single standardized category, we can simplify our exploratory data analysis and improve the consistency of the data.

Descriptive Statistics

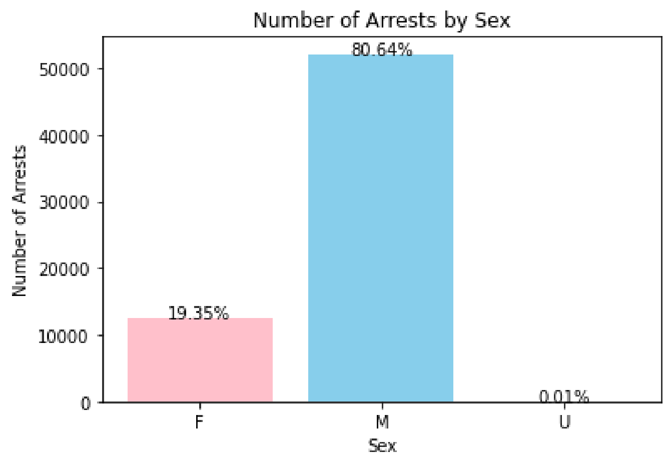
The exploratory data analysis (EDA) enabled us to identify several variables that are crucial for gaining insights into our research questions. These variables, which we carefully considered, include:

- Perceived Race
- Arrest Booked
- Sex
- StripSearch
- Action at Arrest (Cooperative)
- ArrestLocDiv
- Occurrence Category
- ObjectId

In this report, we will conduct a thorough exploratory data analysis by carefully examining multiple sections. Our approach will involve a comprehensive review of each section, with a focus on identifying unique insights and trends that emerge from our analysis.

Sex and Number of Arrests

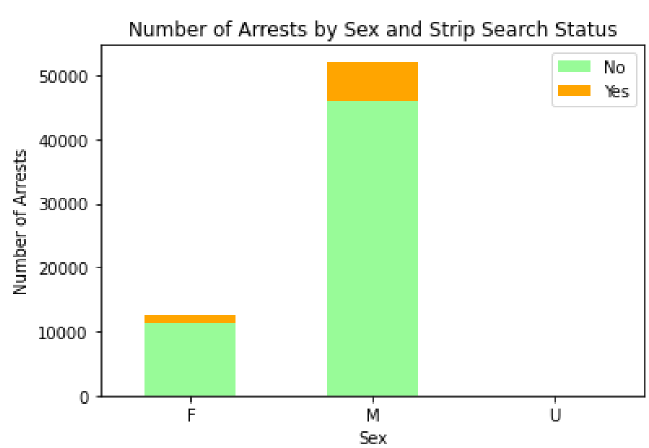
Figure 2. Number of Arrests by Sex



Since the y-axis of this figure was scaled in increments of 10,000, it is difficult to discern the number of arrests for unknown sex (9) from the chart (See Figure 2). The reason for the y-axis scale is that the number of male (52,106) and female (12,500) arrests are significantly higher than the number of arrests for individuals of unknown sex. Therefore, we added the percentage for each sex group. As shown, 80.64% of the arrested individuals were male, 19.35% were female, and the unknown sex group accounted for only 0.01%.

The chart below provides an overview of the number of arrests based on sex and strip search status (See Figure 3). While the male arrestee population may be five times larger than the female arrestee population, the data in this plot suggests that males are strip-searched more frequently than females, and by a larger margin than what might be expected based on population size alone.

Figure 3. Number of Arrests by Sex and Strip Search Status



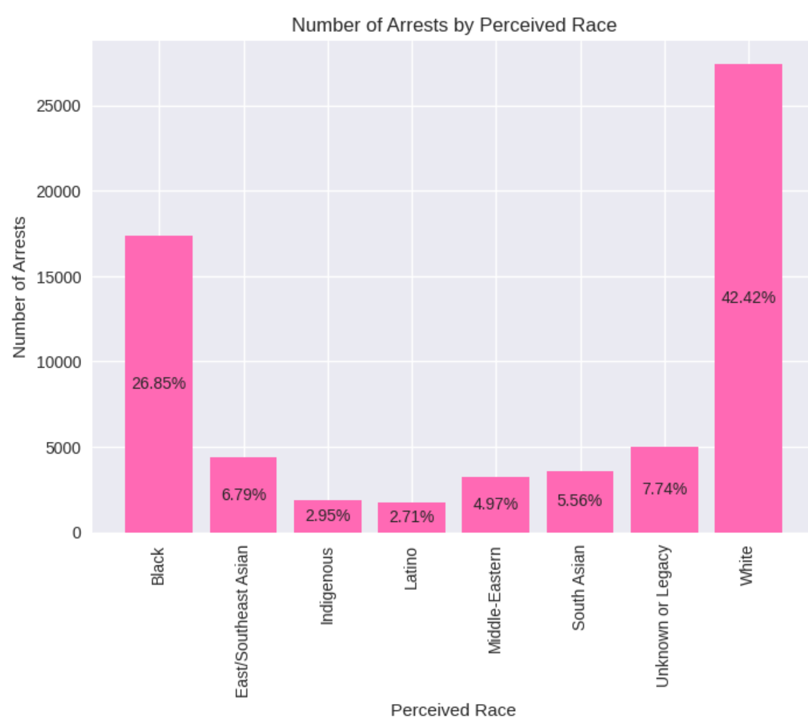
Perceived Races and Number of Arrests

Table 2. Number of Arrests by Perceived Race

Perceived Races	Number of Arrests
White	27,407
Black	17,352
Unknown or Legacy	5,002
East/Southeast Asian	4,388
South Asian	3,594
Middle-Eastern	3,213
Indigenous	1,907

Based on the grouping of individuals by their perceived race as presented in Table 2, our findings suggest that White individuals had the highest number of arrests (27,407), with Black individuals following closely behind (17,352), while Latino individuals had the lowest number of arrests (1,752).

Figure 4. Number of Arrests by Perceived Race with Corresponding Percentages

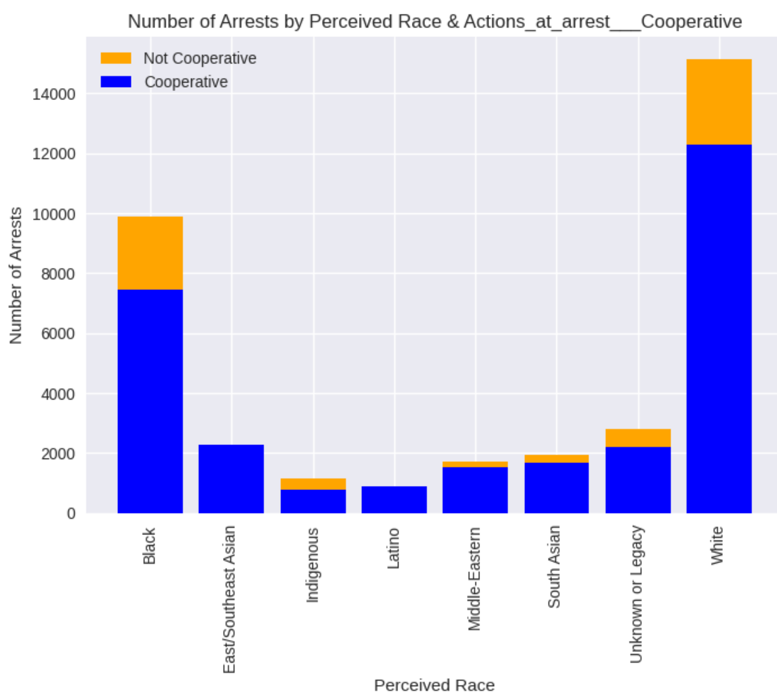


The bar chart illustrates the racial breakdown of arrests, with White individuals accounting for the largest percentage at 42.42%, followed by Black individuals at 26.85%. The percentage of arrests for Latino individuals is the lowest among the groups represented in the chart, at 2.71% (See Figure 4).

Perceived Races, Action at Arrest (Cooperative) and Number of Arrests

We grouped the data based on the arrested people's perceived race and whether they were cooperative during the arrest. After that, we produced a bar chart with the eight race groups as the x-axis, and the number of arrests as the y-axis (See Figure 5).

Figure 5. Number of Arrests by Perceived Race and Actions at Arrest_Cooperative



This plot displays the total number of arrests by perceived race, grouped by whether the individual cooperated or not. While White individuals account for approximately 10,000 more arrests than Black individuals, the proportion of non-cooperative individuals within each group is roughly similar (See Figure 5). To explore this further, we have created a table that shows the number of cooperative arrests and cooperative percentage for each racial group, which will provide additional insight into the patterns observed in the plot (See Table 3).

Table 3. Percentage of People Who Were Cooperative During the Arrests for Each Race Group

Perceived Race	Total Arrests	Actions_at_Arrest__Cooperative	Actions_at_Arrest__Cooperative %

Black	17,352	7,453	42.95%
East/Southeast Asian	4,388	2,267	51.66%
Indigenous	1,907	758	39.75%
Latino	1,752	892	50.91%
Middle-Eastern	3,213	1,509	46.97%
South Asian	3,594	1,662	46.24%
Unknown or Legacy	5,002	2,214	44.26%
White	27,407	12,283	44.82%

Table 3 shows the total arrests of each race group and their cooperation rate during the arrests. According to the table, there were 17,352 arrests of individuals perceived as Black, and of those arrests, 7,453 (42.95%) were cooperative. Similarly, there were 27,407 arrests of individuals perceived as White, and of those arrests, 12,283 (44.82%) were cooperative. East/Southeast Asian (51.66%) and Latino (50.91%) individuals appear to have a slightly higher proportion of cooperative arrests than other groups, while Indigenous individuals have a lower proportion (42.95%). Our analysis revealed that individuals of East or Southeast Asian descent exhibited the highest rate of cooperation during arrest, with a rate of 51.66%. In contrast, individuals of Black descent had the lowest rate of cooperation, with a rate of 42.95%.

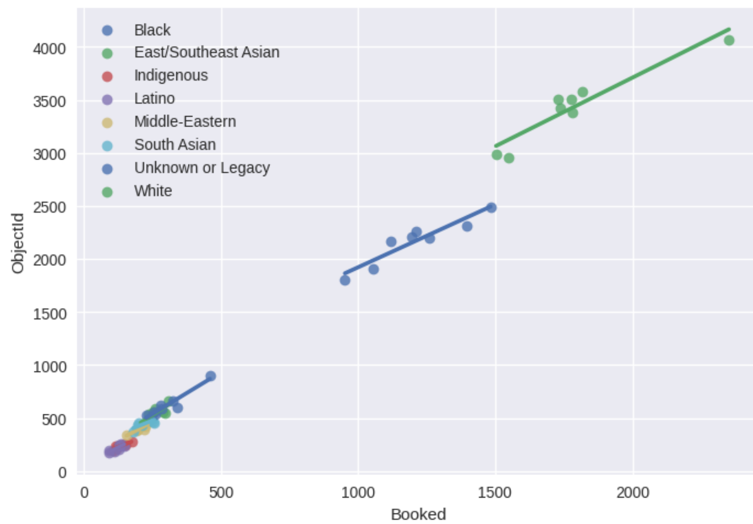
Booked, Perceived Races and Number of Arrests

Table 4. Percentage of People Who Were Booked for Each Race Group

Perceived Race	Total Arrests	Number of Booked	Booked %
Black	17,352	9,666	55.71%
East/Southeast Asian	4,388	2,094	47.72%
Indigenous	1,907	1,071	56.16%
Latino	1,752	964	55.02%
Middle-Eastern	3,213	1,630	50.73%
South Asian	3,594	1,776	49.42%
Unknown or Legacy	5,002	2,435	48.68%
White	27,407	14,247	51.98%

Table 4 shows the number of arrests and the proportion of arrests resulting in booking for each perceived race category. Overall, the table suggests that Indigenous (56.16%) and Black (55.71%) individuals are slightly more likely to be booked following an arrest than other groups, while East/Southeast Asian (47.72%) and South Asian (49.42%) individuals are less likely.

Figure 6. Regression Analysis of Booked vs Total Arrests by Perceived Race



The scatterplot with regression lines for each perceived race category shows a clear linear relationship between the number of booked strip searches and the total number of arrests (See Figure 6). Each race category is represented by a different regression line, with all lines showing a positive slope and a statistically significant relationship between the two variables.

Figure 7. Number of Arrests by ArrestLocDiv

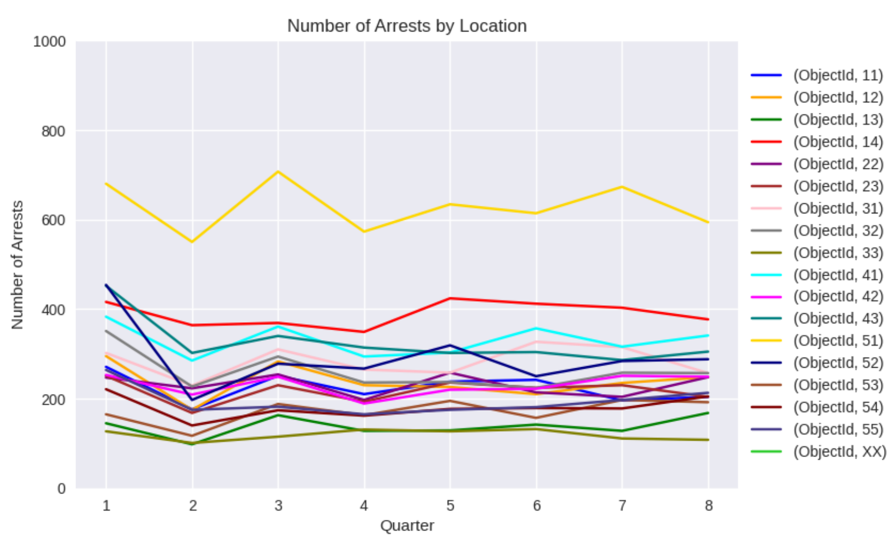


Figure 7 shows the number of arrests by arrest location for each quarter from 2020 to 2021. Each line represents the number of arrests at different locations, and the colour of each line is unique. The y-axis shows the number of arrests, and the x-axis represents the quarters of the two years. It is clear from the plot that ArrestLocDiv 51 has the highest number of arrests compared to other arrest locations. Additionally, we can see that the number of arrests fluctuates throughout the quarters, with some quarters having higher arrest rates than others. Our first logistic regression model will focus on analyzing ArrestLocDiv 51.

T-Test

Before performing a t-test, we need to check the normality of the data as well as the equality of variances between the two groups being compared. Once the data meets these assumptions, we can use a t-test to compare the means of the two groups based on the number of arrests in each quarter.

T-Test #1: Mean of Number of Arrests Between Male and Female

H_0 (Null Hypothesis): The mean number of arrests between males and females is equal.

H_a (Alternative Hypothesis): The mean number of arrests between males and females is not equal.

The mean number of arrests for male individuals is 6513.25 with a standard deviation of 490.29, while the mean number of arrests for female individuals is 1562.50 with a standard deviation of 244.09. The t-test comparing the means of the two groups yielded a t-statistic of 23.916 and a very small p-value of 9.407e-13. This indicates that there is a significant difference in the mean number of arrests between the two groups. Specifically, we have enough evidence to reject the null hypothesis that the means of arrests between male and female individuals are equal, in favour of the alternative hypothesis that they are not equal.

T-Test #2: Mean of the number of arrests between White and Black people

H_0 (Null Hypothesis): The mean number of arrests between White and Black individuals is equal.

H_a (Alternative Hypothesis): The mean number of arrests between White and Black individuals is not equal.

The mean number of arrests for White individuals is 1712.94 with a standard deviation of 250.52, while the mean number of arrests for Black individuals is 1084.50 with a standard deviation of 188.09. The t-test comparing the means of the two groups yielded a t-statistic of 7.769 and a very small p-value of 1.142e-08.

The small p-value indicates that the probability of observing such a large difference in means by chance alone is very low, providing strong evidence against the null hypothesis that there is no difference in the mean number of arrests between White and Black individuals. Therefore, we can reject the null hypothesis in favour of the alternative hypothesis that the mean number of arrests between the two groups is not equal.

T-Test #3: Mean of Cooperative Arrests Among Different Perceived Race Groups

H_0 (Null Hypothesis): The mean of cooperative arrests among different race groups is equal.

H_a (Alternative Hypothesis): The mean of cooperative arrests among different race groups is not equal.

Among all races, only white and unknown races fail normality. A loop was applied to carry out T-Tests between the perceived races with equal sample sizes. Based on the p-values, it appears that in some cases (such as White vs. Unknown, or Black vs. East/Southeast Asian), the means are significantly different at a high level of statistical significance ($p < 0.001$). Note that there is not enough evidence to prove the mean of cooperative arrests is significantly different between unknown races and East/Southeast Asian, South Asian and Middle Eastern, Indigenous and Latino.

Research Design and Methods

Our study aims to explore the relationships among the variables of interest and draw meaningful conclusions from the results. To achieve this, we have adopted a comprehensive data analysis approach that includes power analysis, effect size calculation, power curve plotting, ANCOVA, and logistic regression.

Power analysis is a critical step in determining the sample size required for the study to detect significant effects. We calculate the required sample sizes for three different effect sizes (0.2, 0.5, and 0.8) with a statistical power of 0.8 and a significance level of 0.05. The effect size and power curve will be calculated and presented to ensure that the sample size is adequate to achieve the desired statistical power. The power curve is a graphical representation that illustrates how statistical power changes as the sample size increases or decreases for a given effect size and significance level. Statistical power increases as the sample size increases, meaning that there is a higher likelihood of detecting a significant effect if it truly exists in the population. Conversely, statistical power decreases as the sample size decreases, which means that there is a lower likelihood of detecting a significant effect.

In addition, ANCOVA was employed to examine the differences between the means of the dependent variables across the different groups while accounting for the influence of the identified covariate. ANCOVA allowed us to determine whether the observed differences were significant after controlling for potential confounding factors. Moreover, we applied logistic regression in this report, which was used to explore the relationship between a binary outcome variable and predictor variables.

To address our research questions, we will employ different statistical techniques. Specifically, we will use ANCOVA to investigate the effect of perceived race on the number of arrests while controlling for the number of booked strip searches (RQ1) and identify the impact of sex on the number of arrests while controlling for the number of booked strip searches (RQ2). Additionally, we will utilize a logistic regression model to examine how the ArrestLocDiv and Occurrence Category contributes to an individual's level of cooperation during an arrest (RQ3). This technique will enable us to identify how these two variables interact to affect an individual's level of cooperation during an arrest. Finally, we will also use logistic regression to investigate the relationship between an individual's perceived race, their level of cooperation, and the likelihood of being booked for a strip search (RQ4).

The ANCOVA analysis used in this study involved several steps to ensure the validity and reliability of the results. Firstly, a homogeneity test will be performed to check whether the variance of the dependent variable was equal across different groups, preventing incorrect conclusions about the statistical significance of the differences between the groups. Next, dummy variables will be created for the categorical variable and added to the data frame. An ANCOVA model will then be fitted using the dependent variable, independent variables and dummy variables. The summary statistics of the ANCOVA model will be printed, including the coefficient estimates, standard errors, t-values, p-values, and confidence intervals for each independent variable. The Breusch-Pagan test can be performed to check for homoscedasticity or equal variance of the residuals, while a Shapiro-Wilk test can be conducted to assess the normality of the residuals. These tests helped to ensure that the model assumptions were met and that the results of the ANCOVA were valid.

The logistic regression method used in this study involved several steps. Firstly, dummy variables will be created for the categorical variables in the dataset. Next, the independent variables and the dependent variable will be defined and the dataset was split into training and test sets using the train_test_split function. A heatmap will be plotted to visualize the correlations between the independent variables. Logistic regression can be performed using the logit function from the statsmodels library. The model summary will show the coefficient estimates, standard errors, z-values, and p-values for each independent variable. Odds ratios will also be calculated for each variable to measure the strength of the association between the independent and dependent variables. A prediction interval plot can be useful in evaluating the performance and uncertainty of the model's predictions while the confidence interval plot can help us to determine which variables are statistically significant and how strong their effects are in the logistic regression model.

Overall, these methods are specifically tailored to address our research questions, allowing us to derive valuable insights from the data. The integration of these data analysis techniques will enable us to thoroughly investigate the relationships among variables, pinpoint the critical factors affecting the outcome of interest, and establish well-founded conclusions based on our findings. The results of the analysis will be presented in the subsequent sections of the report, along with a discussion of the study's implications and limitations.

Power Analysis

By conducting a power analysis for different effect sizes (small: 0.2, medium: 0.5, large: 0.8) with a power of 0.8 and a significance level of 0.05, we have determined the necessary sample sizes for each effect size to achieve the desired statistical power. The results are as follows:

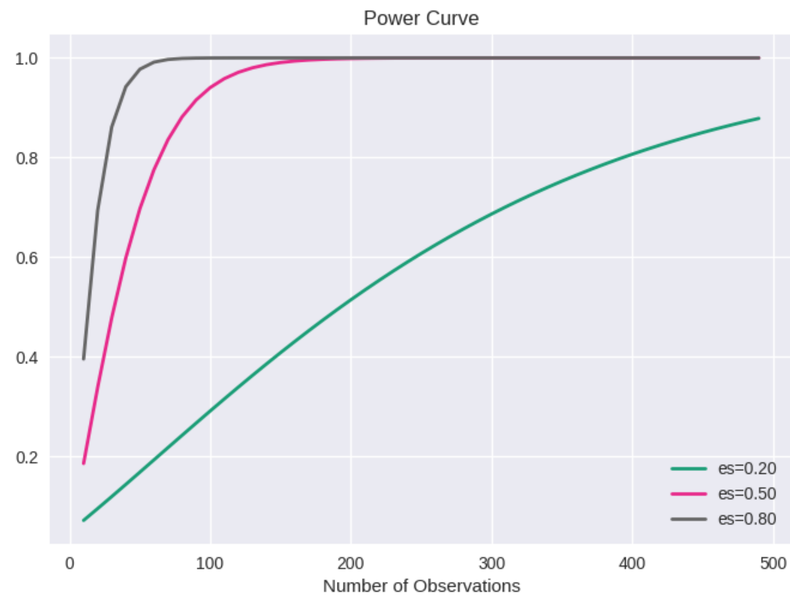
Table 5. Sample Sizes Needed to Observe a Statistically Significant Effect

For a small effect size (d = 0.2)	**For a medium effect size (d = 0.5)**	**For a large effect size (d = 0.8)**
393.405693	63.765611	25.524572

- For a small effect size (d = 0.2), approximately 393 participants are required in each group to achieve a power of 0.8. This means that if the true effect size in the population is small, we would need around 393 participants in each group to have an 80% chance of detecting the effect in our study at a 5% significance level.
- For a medium effect size (d = 0.5), approximately 64 participants are needed in each group to achieve a power of 0.8. In this case, if the true effect size is medium, having 64 participants in each group would provide an 80% chance of detecting the effect at a 5% significance level.
- For a large effect size (d = 0.8), about 26 participants are required in each group to attain a power of 0.8. If the true effect size is large, having 26 participants in each group would give us an 80% chance of detecting the effect with a 5% significance level.

To ensure the validity and reliability of our ANCOVA and Logistic Regression analyses, we will consider these sample size requirements when selecting participants for our study. By meeting the appropriate sample size thresholds, we can increase the likelihood of detecting true effects in our population and reduce the risk of Type II errors.

Figure 8. Power Curve



The three lines in the plot represent different effect sizes (0.2, 0.5, and 0.8) and the x-axis displays the sample sizes ranging from 10 to 500. Regarding the y-axis, it represents the statistical power of the test, which is a probability between 0 and 1. The plot shows that increasing the effect size or sample size results in higher statistical power, indicating that the test is more likely to detect a true effect (See Figure 8).

Results and Findings

ANCOVA

To address our first two research questions, we conducted two separate analyses of covariance (ANCOVAs).

RQ1: Does perceived race have an effect on the number of arrests while controlling the number of booked strip searches?

ANCOVA 1: ObjectId ~ Booked + Perceived Race

To assess the homogeneity of the regression slopes across all racial groups, we conducted an examination of the interaction effect between the total number of booked strip searches and race. Specifically, we evaluated whether the slopes generated by the interaction terms were significantly different from each other. It is important to note that we used a benchmark of 0.05 to determine statistical significance. Upon examining the p-values for all interaction terms, we found that none of them were statistically significant, indicating that the slopes did not significantly differ across all races. Therefore, we conclude that the assumption of homogeneity is met. To sum up, the finding suggests that the relationship between the total number of arrests and the total number of booked strip searches is not significantly different across all racial groups, which strengthens the validity of our results.

Table 6. Model Summary of ObjectId ~ Booked + Perceived Race (ANCOVA 1)

Variable	Coefficient	Std.Error	z	P > | z |	[0.025	0.975]
Intercept	624.001	79.8545	7.81422	1.752e-10	463.969	784.033
Booked	1.27871	0.0641021	19.948	7.519e-27	1.15024	1.40717
East_Southeast_Asian	-410.203	66.6127	-6.15803	8.928e-08	-543.697	-276.708
Indigenous	-556.813	74.156	-7.50867	5.537e-10	-705.425	-408.201
Latino	-559.085	74.9529	-7.45915	6.673e-10	-709.294	-408.876
Middle_Eastern	-482.913	70.0159	-6.89719	5.559e-9	623.228	-42.598
South_Asian	-458.624	68.9415	-6.65236	1.398e-8	-596.786	-320.462
Unknow_or_Legacy	-387.958	64.1339	-6.04918	1.340e-7	-516.485	-259.43
White	524.655	45.8633	11.4395	3.631e-16	432.743	616.567

Note that the adjusted R-square is 0.998 which means our model fits the data well and the model explains 99.8% of the variability of the predictive variable, total number of arrests. Also, the p-value for all explanatory variables is smaller than the threshold of 0.05, thus all variables are significant predictors.

Holding all other variables constant, for every 1 unit increase in the book, the total arrests increases by 1.28 (See Table 6). Black is a reference group of the model. The intercept of 624.001 corresponds to the predicted value for Black people when all other variables are equal to zero. Holding other variables constant, the East/Southeast Asian group is 410.20 units of total arrests lower than the Black group, the Indigenous group is 556.81 units of total arrests lower, the Latino group is 559.09 units of total arrests lower, the Middle Eastern group is 482.91 units of total arrests lower, South Asian group is 458.62 units of total arrests lower, Unknown or Legacy group is 387.96 units of total arrests lower, White group is 524.65 units of total arrests higher, than the Black group on average.

To assess the presence of heteroskedasticity, we conducted a Breusch-Pagan test after constructing the ANCOVA model to examine whether the variance of the error terms is constant across the observations. The null hypothesis states that there is no heteroskedasticity, while the alternative hypothesis suggests the presence of heteroskedasticity. We obtained a p-value smaller than 0.05, which indicates that we should reject the null hypothesis and accept the alternative hypothesis, implying the existence of heteroskedasticity. As a result, our model fails to meet the homoscedasticity assumption and may not be entirely reliable.

Furthermore, we performed the Shapiro-Wilk test to determine if the residuals in our dataset are normally distributed. The null hypothesis assumes that the residuals are normally distributed. We observed a p-value less than 0.05, which suggests that we should reject the null hypothesis. However, this finding could be influenced by the large sample size of the dataset. In such cases, even small deviations from normality might produce significant p-values, leading to the rejection of the null hypothesis.

RQ2: Does sex have an effect on the number of arrests while controlling the number of booked strip searches?

ANCOVA 2: ObjectId ~ Booked + Sex

To assess the homogeneity of the regression slopes across all genders, we examined the interaction effect between the total number of booked strip searches and gender. Specifically, we evaluated whether the slopes generated by the interaction terms were significantly different from each other. We used a benchmark of 0.05 to determine statistical significance and found that none of the interaction terms were statistically significant. This indicates that the slopes did not significantly differ across all genders, which meets the assumption of homogeneity.

In summary, our finding suggests that the relationship between the total number of arrests and the total number of booked strip searches is not significantly different across all genders, which strengthens the validity of our results.

Table 7. Model Summary of ObjectId ~ Booked + Sex (ANCOVA 2)

Variable	Coefficient	Std.Error	z	P > | z |	[0.025	0.975]
Intercept	742.956	102.997	7.21335	1.450e-6	525.651	960.262
Sex[T.M]	1531.18	374.476	4.08887	0.000765	741.108	2321.26
Sex[T.U]	-741.634	123.352	-6.01232	1.3996e-5	-1001.88	-481.383
Booked	1.19445	0.128084	9.3255	4.26588e-8	0.924218	1.46469

Our regression model appears to fit the data well, as evidenced by an adjusted R-square of 0.997. This means that our model explains 99.7% of the variability in the dependent variable, the total number of arrests. In addition, all explanatory variables have p-values smaller than 0.05, indicating that they are significant predictors of the total number of arrests (See Table 7).

Specifically, gender was found to be a significant predictor. Holding all other variables constant, we found that every 1 unit increase in the number of booked strip searches was associated with a 1.19 unit increase in total arrests. The female gender group served as the reference group in our model. The intercept of 742.956 corresponds to the predicted value for females when all other variables are equal to zero. We found that on average and holding other variables constant, the male gender group had 1531.18 more total arrests than the female group, while the unknown gender group had 741.63 fewer total arrests than the female group.

Similarly, a Breusch-Pagan test and a Shapiro-Wilk test were performed to check the assumptions of the ANCOVA model. The p-value is 0.019 and 0.57 respectively, thus the model fails homoscedasticity but the residuals are normally distributed.

Logistic Regression

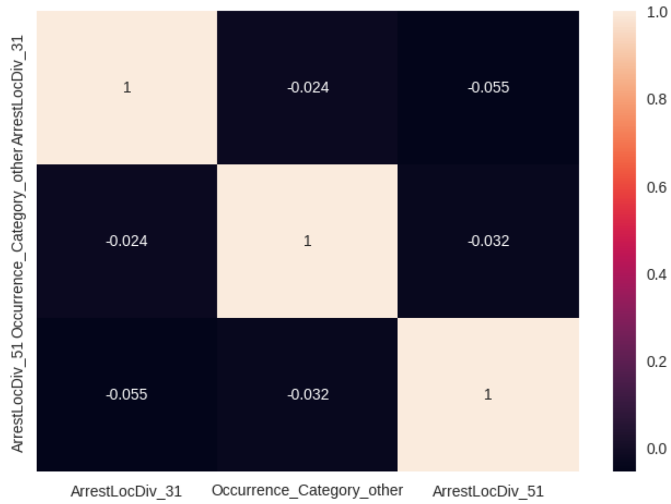
To investigate our last two research questions, we used logistic regression models. The models are presented below. Based on the results of our power analysis, we determined that the sample sizes used in our logistic regression models were appropriate to address our research questions.

RQ3: Does the ArrestLocDiv and Occurance Category have an impact on the likelihood of individuals being cooperative during an arrest?

Logistic 1: Actions_at_arrest_Cooperative ~ ArrestLocDiv_31 + Occurrence _ Category _ other + ArrestLocDiv_51

The first logistic regression model examines the relationship between the target variable (Actions_at_arrest___Cooperative) and three predictor variables: ArrestLocDiv_31, Occurrence_Category_other, and ArrestLocDiv_51. During the exploratory data analysis, we discovered that ArrestLocDiv_51 had the highest number of arrests compared to other locations, prompting us to investigate whether this location had an impact on the likelihood of individuals being cooperative during an arrest. To provide a comparison, we randomly selected ArrestLocDiv_31 as another predictor variable. Additionally, in the data pre-processing section, we categorized all arrest reasons into different groups, with "Other" representing "Other Offence," "Other Statute," and "Other Statute & Other Incident Type." Since arrest reasons falling under the "Other" category did not provide specific details like other categories, we were interested in determining whether this category would impact a person's level of cooperation during an arrest.

Figure 9. Correlation Heatmap for the variables in Logistic Regression 1



The heatmap data indicates that the correlation between the three variables (ArrestLocDiv_31, Occurrence_Category_other, and ArrestLocDiv_51) is very low, with correlation coefficients ranging from -0.055 to 0.024, suggesting a weak linear relationship between them. This suggests that these three variables can be used as separate independent variables in logistic regression analysis without causing multicollinearity issues (See Figure 9).

Table 8. Classification Report of Logistic Regression 1

	Precision	Recall	F1-Score	Support
0	0.98	0.56	0.71	62,352
1	0.05	0.66	0.10	2,263
Accuracy			0.56	64,615
Macro Avg	0.52	0.61	0.40	64,615
Weighted Avg	0.95	0.56	0.69	64,615

Table 8 provides an overview of the performance metrics of a binary classification model that predicts two classes, 0 and 1. Class 0 represents the cooperative actions at arrest, and Class 1 represents the non-cooperative actions at arrest. Precision measures the proportion of true positives to the total number of positive predictions made by the model. Recall, on the other hand, measures the proportion of true positives to the total number of actual positives in the dataset. F1-score provides a balance between precision and recall, and it is the harmonic mean of these two metrics. For Class 0, the precision is 0.98, meaning that 98% of the instances predicted as cooperative were indeed cooperative. However, the recall is 0.56, which means that the model identified only 56% of the cooperative instances correctly. The f1-score is 0.71, which is a balanced measure of precision and recall. For Class 1, the precision is 0.05, meaning that only 5% of the instances predicted as non-cooperative were actually non-cooperative. The recall is 0.66, indicating that the model identified 66% of the non-cooperative instances correctly. The f1-score is 0.10, which is quite low, indicating that the model's performance is not very good for this class. For Class 0 and 1, there are 62,352 and 2,263 instances in the dataset, respectively.

The model's accuracy is 0.56, indicating that 56% of the predictions made by the model are correct. In summary, this logistic regression model has moderate overall accuracy and performs well in predicting cooperative instances. However, its performance in identifying non-cooperative instances is poor, which might affect the reliability of the model in practical applications. This dataset is imbalanced, which could be one of the reasons why the model performs well in predicting cooperative instances but struggles with non-cooperative instances.

Table 9. Actions_at_arrest_Cooperative ~ ArrestLocDiv_31 + Occurrence_Category_other + ArrestLocDiv_51 (Logistic Regression 1)

Variable	Coefficient	Std.Error	z	P > | z |	[0.025	0.975]
Intercept	-0.1753	0.009	-20.007	0	-0.192	-0.158
ArrestLocDiv_31	0.8740	0.045	19.285	0	0.785	0.963
Occurrence_Category_other	-0.4021	0.029	-13.665	0	-0.46	-0.344
ArrestLocDiv_51	-0.3256	0.03	-10.708	0	-0.385	-0.266

In logistic regression, a positive coefficient indicates a positive relationship between the predictor variable and the target variable, while a negative coefficient indicates a negative relationship. Based on the logistic regression results, the coefficient for ArrestLocDiv_31 is 0.8740, suggesting a positive association with the target variable Actions_at_arrest_Cooperative. Compared with arrested people who were not arrested at ArrestLocDiv_31, the log odds of getting the arrested people to be cooperative during the arrest at ArrestLocDiv_31 are 0.874 higher, controlling for other features. Note this feature is statistically significant. The coefficient for ArrestLocDiv_51 is -0.3256, indicating that compared with arrested people who were not arrested at ArrestLocDiv_51, the log odds of getting the arrested people to be cooperative during the arrest at ArrestLocDiv_51 are 0.3256 lower, controlling for other features. Again, the feature is statistically significant. The coefficient for Occurrence_Category_other is -0.4021, indicating that when compared to other reasons for arrest that are not categorized under Occurrence_Category_other, the log odds of arrested individuals being cooperative during the arrest for reasons falling under Occurrence_Category_other are 0.4021 lower. It is important to note that this feature is statistically significant.

The logistic regression results suggest a positive association between ArrestLocDiv_31 and the likelihood of an individual being cooperative during an arrest, while there are negative associations between ArrestLocDiv_51 and Occurrence_Category_other with the likelihood of an individual being cooperative during an arrest. These findings indicate that individuals arrested in location 51 or for reasons categorized under 'Other' may be less likely to cooperate during the arrest. Therefore, it can be concluded that the location of arrest and arrest reasons do have an impact on an individual's level of cooperation during an arrest, which supports the initial hypothesis.

Table 10. Confusion Matrix of Logistic Regression 1

	Predicted Negative	Predicted Positive
Actual Negative	34,815	27,537
Actual Positive	762	1,501

A confusion matrix is a useful tool for evaluating the performance of logistic regression models. In our case, we used it to assess the performance of a model in predicting whether a person would cooperate during an arrest, based on the X variables: ArrestLocDiv31, Occurrence_Category_other, and ArrestLocDiv_51.

Table 10 shows the resulting confusion matrix, with the numbers of true positives, true negatives, false positives, and false negatives. Specifically, there were 34,815 true negatives (predicted not cooperative and actually not cooperative), indicating that the model correctly predicted that the person would not cooperate in these cases. Additionally, there were 1,501 true positives (predicted cooperative and actually cooperative), indicating that the model correctly predicted that the person would cooperate in these cases. However, there were also 27,537 false positives (predicted cooperative but actually not cooperative), which means that the model incorrectly predicted that the person would cooperate in these cases. Finally, there were 762 false positives (predicted not cooperative but actually cooperative), which means that the model incorrectly predicted that the person would not cooperate in these cases.

Table 11. Odds Ratio of Logistic Regression 1

	OR	Lower CI	Upper CI
Intercept	0.839239	0.824953	0.853772
ArrestLocDiv_31	2.396433	2.192756	2.619029
Occurrence_Category_other	0.668885	0.631396	0.708600
ArrestLocDiv_51	0.722074	0.680296	0.766417

We perform a logistic regression to examine the effects of ArrestLocDiv_31, Occurrence_Category_other, ArrestLocDiv_51 on the likelihood that the arrested individuals are cooperative during the arrest. The values in the second column represent the odds ratios for each predictor variable. The odds ratio for ArrestLocDiv_31 is 2.396433, indicating that individuals arrested at ArrestLocDiv_31 are more likely to exhibit cooperative behaviour during the arrest compared to those who were not arrested there. On the other hand, the odds ratio for Occurrence_Category_other is 0.668885, suggesting that individuals arrested for reasons falling under Occurrence_Category_other are less likely to exhibit cooperative behaviour during the arrest compared to those arrested for other reasons. Similarly, the odds ratio for ArrestLocDiv_51 is 0.722074, which means that individuals arrested at ArrestLocDiv_51 are also less likely to exhibit cooperative behaviour during the arrest compared to those who were not arrested there. The confidence intervals for all three odds ratios do not include 1, indicating that these associations are statistically significant.

Figure 10. Prediction Interval of Logistic Regression 1

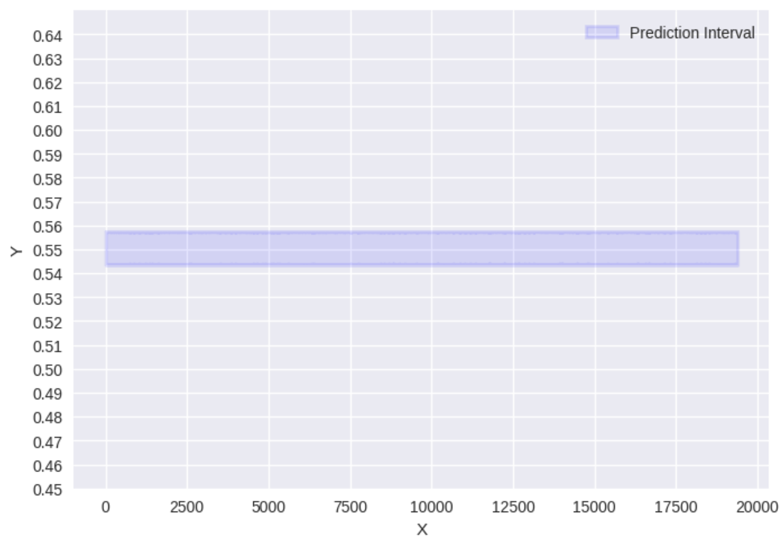
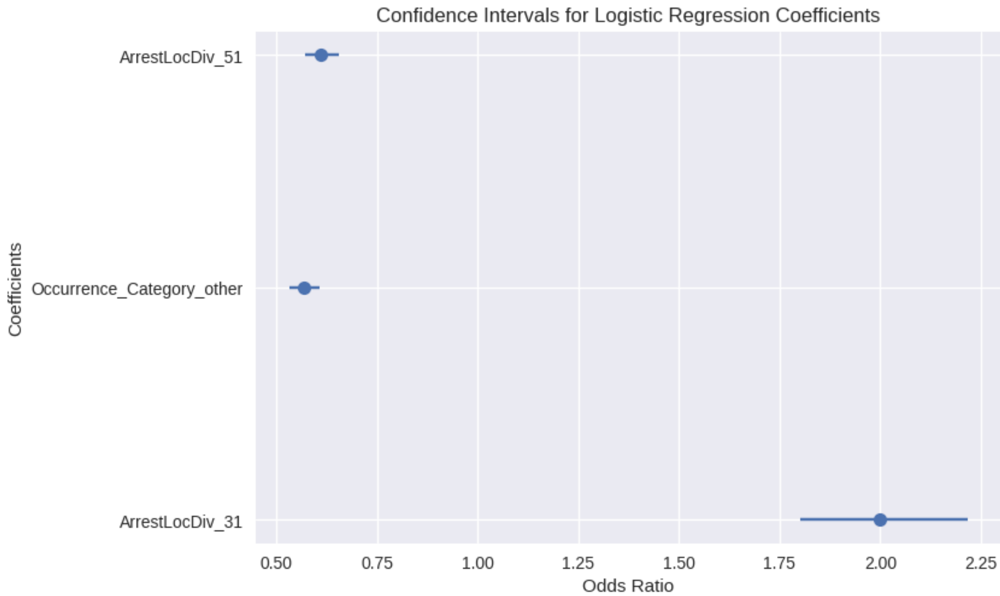


Figure 10 shows the prediction interval for our first logistic regression model. The shaded region represents the range of values where we expect the true probability of the outcome variable to lie with 95% confidence. The x-axis of the plot represents the index number of the data points in the test dataset, while the y-axis represents the probability of the model's predictions. Based on the plot, we can see that the model's predicted results mostly fall around 0.55, and the confidence interval is relatively stable, indicating that the model's predicted results are relatively accurate.

Figure 11. Confidence Interval of Logistic Regression 1



The plot shows the odds ratios and confidence intervals for the coefficients of our first logistic regression model. The points on the plot represent the estimated odds ratios for each predictor variable, and the error bars represent the 95% confidence intervals around these estimates. If the error bar does not cross the value of 1 on the horizontal axis, it suggests that the odds ratio estimate is statistically significant at the 95% confidence level. A value of 1 for the odds ratio indicates that the predictor variable has no effect on the outcome variable. In this plot, we can see that the odds ratio estimates for all predictor variables are statistically significant, as their error bars do not cross the value of 1. For example, the odds ratio estimate for ArrestLocDic_31 is 2, and its confidence interval [1.8, 2.2] does not include 1 (See Figure 11).

RQ4: Does an individual's perceived race have an impact on the likelihood of being booked for a strip search when taking into account their level of cooperation?

Logistic 2: Booked ~ Actions_at_arrest_Cooperative + Perceived_Race

The second logistic regression model examines the relationship between the target variable (Booked) and two predictor variables: Actions_at_arrest___Cooperative and Perceived_Race.

In the Literature Review section, we discussed how booked strip searches are part of an individual's criminal history record, and that those with a history of criminal behaviour are more likely to be arrested and detected by law enforcement. Some previous research has shown that individuals from certain racial and ethnic groups are more likely to be subjected to law enforcement scrutiny and surveillance, leading to disparities in arrest rates and the use of coercive measures such as strip searches. This raises questions about the potential for implicit bias and systemic discrimination in the criminal justice system, particularly with regard to the booking of strip searches. By examining the impact of race on the likelihood of being booked for a strip search, while controlling for the factor of action at arrest-cooperative, this research question aims to contribute to the understanding of potential racial disparities in law enforcement practices and to identify possible areas for policy and procedural improvements.

Table 12. Booked ~ Actions_at_arrest_Cooperative + Perceived_Race (Logistic Regression 2)

Variable	Coefficient	Std.Error	z	P > | z |	[0.025	0.975]
Intercept	-0.1312767	0.0153343	4.9922	5.970e-7	0.0465	0.106607
Actions_at_arrest_Cooperative	0.192421	0.00896714	21.45841787	<0.001	0.17484496	0.20999614
Perceived_Race_Black	0.268715	0.00987393	27.21455295	<0.001	0.24936169	0.2880675
Perceived_Race_Indigenous	0.301040	0.02775321	10.8470457	<0.001	0.24664403	0.3554366
Perceived_Race_Latino	0.285401	0.02916112	9.78703514	<0.001	0.22824508	0.34255666
Perceived_Race_South_Asian	0.023923	0.02024688	1.18203878	0.23719032	-0.01575129	0.06361649
Perceived_Race_Unknow_or_Legacy	-0.023671	0.01726812	-1.3707794	0.17044373	-0.05751629	0.01017473
Perceived_Race_White	0.122425	0.00824367	14.85084549	<0.001	0.10626782	0.13858299

The above is the summary of the second logistic regression model. The intercept term in the Logit Regression Results table represents the estimated log odds of the outcome variable when all predictor variables are equal to zero, which is -0.13. Holding all other variables constant, for each unit increase in cooperative action, the log odds of being booked for a strip search increased by 0.19. Holding all other variables constant, Black, Indigenous, Latino, South Asian and White dependents have higher log odds of being booked, while the Unknown group has a lower log odds of being booked, compared to the reference group, Middle-eastern. Among all the groups, the South Asian and Unknown groups have a p-value that is above the threshold of 0.05, indicating their coefficient estimate is not statistically significant (See Table 12).

Table 13. Confusion Matrix of Logistic Regression 2

	Predicted Negative	Predicted Positive
Actual Negative	3,595	5,589
Actual Positive	3,406	6,795

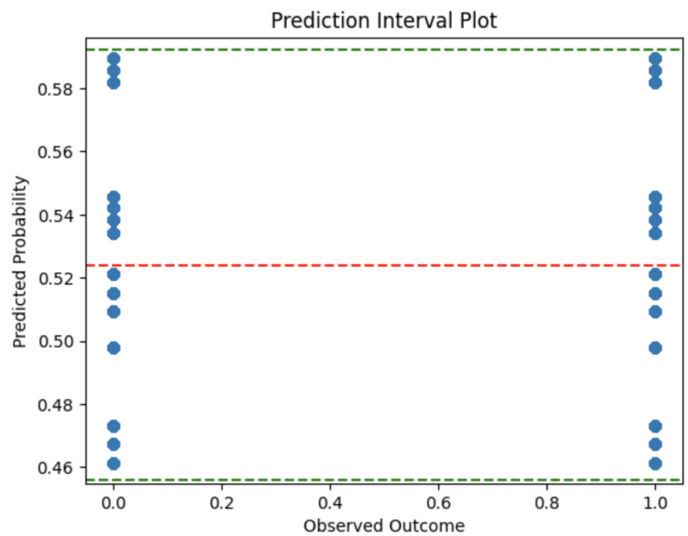
Noted that the accuracy of our second logistic regression model is 0.54. We used a confusion matrix to evaluate how well our model performed in predicting whether an individual would be booked for a strip search, based on their race and actions at arrest. The results of the confusion matrix are shown in Table 13, which shows the number of true positives, true negatives, false positives, and false negatives. In particular, the confusion matrix indicates that the model correctly predicted 3,595 cases as true negatives, where the person wasn't booked for a strip search and the model predicted that they indeed weren't booked. Additionally, there were 6,795 cases where the model correctly predicted that the person was booked for a strip search (true positives). However, there were also 5,589 cases where the model incorrectly predicted that the person was booked for a strip search (false positives). Finally, the model incorrectly predicted that the person wasn't booked for a strip search in 3,406 cases (false negatives).

Table 14. Odds Ratio of Logistic Regression 2

	OR	Lower CI	Upper CI
Actions_at_arrest_Cooperative	1.21218019	1.19106155	1.2336733
Perceived_Race_ Black	1.3082817	1.28320608	1.33384734
Perceived_Race_ Indigenous	1.35126382	1.27972348	1.42680346
Perceived_Race_ Latino	1.3302952	1.25639321	1.40854416
Perceived_Race_ South_Asian	1.02422128	0.98437211	1.06568362
Perceived_Race_ Unknow_or_Legacy	0.97660718	0.94410651	1.01022667
Perceived_Race_ White	1.13023481	1.11211968	1.148645

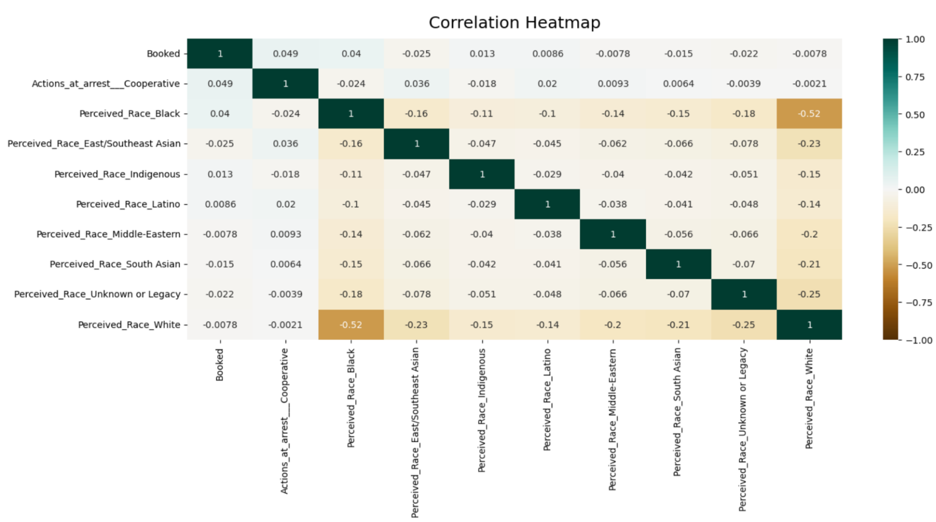
To investigate how likely an individual is booked for a strip search, we conducted a logistic regression analysis that considered the impact of perceived race and whether they are cooperative at arrest. The second column of results shows the odds ratio for each predictor variable. The results suggest that all attributes have a positive association with getting booked or not, except unknown race. All odds ratios except for unknown race and Southeast Asian are statistically significant, as their confidence intervals do not include 1.

Figure 12. Prediction Interval of Logistic Regression 1



The red dashed line serves as a reference line at the mean predicted probability while the two green dashed lines indicate the upper and lower bounds of a 95% prediction interval. The interval is calculated as the mean predicted probability plus or minus 1.96 times the standard deviation of the predicted probabilities. The above plot helps visualizes how well the model predictions match the observed outcomes. Noted that the points are widely spread out, which suggests the model's predictions are less reliable and the prediction interval might be widened (Figure 12).

Figure 13. Correlation Heatmap for the variables in Logistic Regression 1



The heatmap data indicates that the correlation between the perceived races and booked is very low, with correlation coefficients ranging from -0.525 to 0.049, suggesting a weak linear relationship between them. This suggests that these variables can be used as separate independent variables in logistic regression analysis without causing multicollinearity issues (See Figure 13).

Discussion

Regarding logistic regression 1, the model is statistically significant with a p-value of 1.207e-158. The pseudo-R-squared value is a measure of how well the logistic regression model fits the data. In our case, the value of 0.82% indicates that the model only explains a small portion of the variability in the target variable, Actions_at_arrest___Cooperative. We evaluated its performance using the evaluation metrics. For Class 0 and 1, there are 62,352 and 2,263 instances in the dataset, respectively. This dataset is imbalanced, which could be one of the reasons why the model performs well in predicting cooperative instances but struggles with non-cooperative instances. To improve the performance of the model, several approaches could be considered. First, it may be possible to identify additional relevant features to better capture the relationship between the predictor variables and the target variable. Additionally, addressing the class imbalance in the target variable could be useful, as this may improve the model's ability to predict cooperative individuals. We could use sampling techniques, such as oversampling or undersampling, to balance the classes. Another approach would be to try different algorithms, such as Random Forest or XGBoost, to see if they perform better than logistic regression on this particular dataset. Finally, hyperparameter tuning and collecting more data could also potentially improve the model's performance.

The second logistic regression model suggests that there is a positive association between the perceived race, whether or not being cooperative at arrest and the likelihood of being booked for a strip search. The dummy-coded race variables indicate that Black, Indigenous, Latino and South Asian races have a stronger influence on the outcome as the coefficients are larger than the Unknown and White race groups. The variable of being cooperative or not also has a positive association with the likelihood of being booked. The p-values for Unkown and South Asian groups are not statistically significant at the 0.05 level, but all other variables have p-values less than 0.05, indicating that they are statistically significant predictors of being booked for a strip search. Overall, the model suggests that race and cooperative behaviour are important factors in determining whether an individual will be booked for the strip search.

While the inclusion of indicators of whether a person was booked at a police station within 24 hours following an arrest event in the dataset is helpful, it is important to note that there are some limitations to this data. One major limitation is that due to issues with the booking template, there may be some records where a person was strip-searched, but the data does not indicate a booking, with a value of 0. In such cases, the users of this dataset must presume that a booking took place, which could introduce inaccuracies and biases into any analysis or research conducted using this data. Additionally, the dataset's coverage may be limited to specific geographic areas, time periods, or types of arrests, which could also limit its generalizability and applicability to other contexts.

Conclusion

RQ1: Does perceived race have an effect on the number of arrests while controlling the number of booked strip searches?

Our analysis has shown that perceived race does have a significant effect on the number of arrests while controlling for the number of booked strip searches. Compared to the Black group, the East/Southeast Asian, Indigenous, Latino, Middle Eastern, South Asian, and Unknown or Legacy groups have lower predicted values of total arrests, while the White group has a higher predicted value of total arrests.

RQ2: Does sex have an effect on the number of arrests while controlling the number of booked strip searches?

Our analysis indicates that sex does have a significant effect on the number of arrests while controlling for the number of booked strip searches. Our findings reveal that gender is a significant predictor of total arrests, with males having a higher predicted value of total arrests compared to females, while the unknown gender group had a lower predicted value of total arrests compared to females. Specifically, holding all other variables constant, we found that every 1 unit increase in the number of bookings was associated with a 1.19 unit increase in total arrests.

RQ3: Does the ArrestLocDiv and Occurance Category have an impact on the likelihood of individuals being cooperative during an arrest?

Our analysis suggests that the location of the arrest and the reason for the arrest have an impact on the likelihood of an individual being cooperative during an arrest. Our logistic regression model specifically revealed that when holding other factors constant, the likelihood of an individual being cooperative during an arrest increases if they were arrested at ArrestLocDiv_31. On the other hand, the model indicated that the likelihood of an individual being cooperative during an arrest decreases if they were arrested at ArrestLocDiv_51 or if the reason for the arrest falls under the Occurrence_Category_other. The confusion matrix showed that the model correctly predicted true negatives and true positives in most cases. However, the model also had a significant number of false positives and false negatives, indicating that there is still room for improvement in predicting the likelihood of an individual being cooperative during an arrest.

RQ4: Does an individual's perceived race have an impact on the likelihood of being booked for a strip search when taking into account their level of cooperation?

Our logistic regression analysis showed that perceived race and level of cooperation at arrest have a positive association with the likelihood of an individual being booked for a strip search, except for the Unknown and Southeast Asian race groups. Our logistic regression model found that holding all other variables constant, Black, Indigenous, Latino, South Asian, and White individuals had higher log odds of being booked for a strip search compared to the reference group of Middle Eastern individuals. However, the coefficients for the South Asian and Unknown groups were not statistically significant. The confusion matrix showed that our model correctly predicted true negatives and true positives in most cases, but had a significant number of false positives and false negatives. Therefore, there is still room for improvement in predicting whether an individual will be booked for a strip search based on their race and level of cooperation during an arrest.

In conclusion, the test results of our analyses revealed significant differences in the mean number of arrests and the likelihood of being booked for a strip search across different gender and race groups. Specially, we found that certain racial groups and gender groups were more likely to be arrested while controlling if they were booked for a strip search. Also, perceived race and action at arrest are important factors that impact if the defendant will be booked for a strip search. The finding highlights the need for further investigation into potential inequalities in the criminal justice system, particularly in regard to how race may impact the likelihood of being arrested and booked for the strip search. Further research is needed to better understand the underlying factors driving these disparities and to develop effective policies and interventions that promote fairness and equality for all individuals in the criminal justice system. In addition, efforts should also be made to improve training and protocols for officers to de-escalate situations and increase the likelihood of cooperation during an arrest. These steps can help promote safer and more effective interactions between law enforcement and individuals involved in arrests.

This project will provide insight into how perceived race and gender interact with the number of arrests. Knowledge gained from this project will contribute to existing literature and help police and policymakers improve equality in the criminal justice system.

Reference

Balfour, G. 2020. “Women in Custody.” In Adults Corrections in Canada, by J. Winterdyk and M. Winrath, (pp. 147-166). de Sitter Publications.

Government of Canada, D. of J. (2021, July 7). *State of the Criminal Justice System Annual Report 2020 report: Focus on Women*. State of the Criminal Justice System Annual Report. Retrieved April 15, 2023, from https://www.justice.gc.ca/eng/cj-jp/state-etat/2021rpt-rap2021/

Government of Canada, D. of J. (2022, August 17). *Public perception of crime and justice in Canada: A review of opinion polls*. Canadians' Concerns in Relation to Crime. Retrieved February 22, 2023, from https://www.justice.gc.ca/eng/rp-pr/csj-sjc/crime/rr01_1/p1.html

McNeilly, G. (2019, March). *Breaking the golden rule - office of the independent police review director*. Office of the Independent Police Review Director. Retrieved February 24, 2023, from https://www.oiprd.on.ca/wp-content/uploads/OIPRD_Breaking-the-Golden-Rule_Report_Accessible.pdf

Rankin, Jim, Quinn, Jennifer, Shephard, Michelle Simmie, Scott and Duncanson, John. 2002. “Singled Out: An Investigation into Race and Crime.” Toronto Star. October 19: A1.

Raphael, S., & Rozo, S. V. (2019). Racial disparities in the acquisition of Juvenile Arrest Records. *Journal of Labor Economics*, *37*(S1). https://doi.org/10.1086/701068

Roberts, J. V., & Doob, A. N. (1997). Race, ethnicity, and Criminal Justice in Canada. *Crime and Justice*, *21*, 469–522. https://doi.org/10.1086/449256

Stolzenberg, L., D’Alessio, S. J., & Flexon, J. L. (2021). *The usual suspects: Prior criminal record and the probability of arrest.* Police Quarterly, 24(1), 31-54.

Wortley , S., & Jung, M. (2020). (rep.). *Racial Disparity in Arrests and Charges: An analysis of arrest and charge data from the Toronto Police Service* (pp. 4–5). Toronto, Ontario: Government of Ontario.