Introduction

Addressing racial inequality has been at the forefront of the Canadian government's program funding in recent years as the nation's governing body strives to emulate the diversity that is considered one of the "cornerstones of Canadian identity" (Government of Canada, 2021). However, there is still a lot of progress to be made in the pursuit of nationwide racial justice, as demonstrated by the 80% increase in race-targeted hate crimes reported by Statistics Canada, from 2019 to 2020 (2022). Although Canada's justice system is by definition expected to uphold fairness, racial discrimination and bias have been consistently present within the criminal justice and policing system. The issue of racial injustice was recently highlighted in a report issued by the Correctional Investigator of Canada in 2020 which divulged that although Indigenous people only account for 5% of the population of Canada, they make up 30% of the prison population federally (Office of the Correctional Investigator, 2020). In order to identify specific racial discrepancies present within policies and practices within the Canadian policing and criminal justice system, that can consequently lead to mitigation, it is integral that the effects of race on arrest-related police procedures are examined. There may be multiple indicators of over-policing/unjust police measures that can be analyzed in order to investigate racial discrepancies such as whether someone is strip searched or booked after an arrest. In addition, racial discrimination is often exacerbated by other demographic factors such as gender/sex. According to Goli Rezai-Rashti, connections between demographic attributes such as race, social class, and gender are often overlooked when discussing equity issues, resulting in the loss of insights into their comprehensive relationship with discrimination (Ontario Human Rights Commission). In addition, other factors outside of demographics may affect the policing procedures implemented such as the person's behavior at the time of the arrest. The person's behavior can range from cooperative to difficult or even violent during the arrest and intuitively these actions would influence the presence of any additional policing measures engaged by officers such as strip searches and police station bookings.

Considering the prevalent issues of racial inequality discussed above, we have decided to investigate the role of race in the likelihood of arrested individuals being booked at a police station or strip searched in order to uncover potential disparities. In addition to the prominent factor of perceived race, we have decided to include additional factors that may influence policing measures such as cooperative behavior at arrest and demographics like gender/sex. In our analysis, we will utilize a public dataset on arrests and strip searches recorded by the Toronto police service.

Literature Review

A debate which has been pervasive in the literature concerning criminal justice systems and their relationship to race is that of whether the differences within the impact of policing on different racial groups are "discrimination or good policing" (Wortley & Tanner, 2004). The claim that these disparities are due to racial discrimination is frequently supported by consistent overrepresentation observed in select data detailing criminal activity-related metrics such as offense categories. In addition, in a study published by the Toronto Star, an analysis of post-arrest treatment and race revealed that black offenders are on average subject to more harsh treatment when compared to white offenders. This is more concretely translated to white offenders being released at the crime scene in contrast to their black counterparts who are more likely to be brought back to the police station and detained until the bail hearing. To reinforce the prominence of racial issues within the policing sector, it was found in the 2017 OHRC report that 25.9% of black people, 24% of indigenous people, and 17.9% of other racial minorities, were checked on the street while only 8.6% of white people were street checked (Giwa et al. 2020). Furthermore, in another study conducted, it was reported that 44% of the black males included in the sample were stopped as well as questioned by police at least one time in the last two years while only 12% of the white males within the sample reported the same experience (Wortley and Tanner 2003). However, despite centering on specifically racial bias in police investigations, the authors claim that such discrimination cannot be fully explained without considering other demographic factors such as education, gender, and social class. The effect of gender/sex on policing measures and subsequent disparities has been investigated by many researchers as well in an effort to understand the intersectionality. Data from legal sentences in Pennsylvania was used to demonstrate that specifically for male defendants, there were overall harsher sentences for certain racial groups such as black and hispanic people when compared to white male defendants (Steffensmeier et al. 2017). These effects were not found to be as prominent in the female defendant population (Steffensmeier et al. 2017). Furthermore, additional studies have found that study participants were significantly more likely to believe in the validity of a call to police if the subject in question was male when compared to their female counterparts (Pica et al. 2019).

The intersection of these demographic attributes plays a significant role in highlighting the complex relationships with discrimination that the general population experiences. Overall, it is evident that racial bias and discrimination is extremely prevalent within the criminal justice and policing system and should be further investigated using relevant metrics that reflect the subtleties of profiling within said system alongside other demographic factors.

Research Objective and Questions

Within our project we intend to explore how race influences arrest-related policing procedures such as being detained at a police station or other post-arrest treatment. Similarly to the literature discussed above, we aim to identify potential disparities between racial groups

regarding criminal justice and policing behaviors. In addition, our analysis is set up to expand on the intersectional relationship between race and other demographic factors within the context of bias within the criminal justice system. The following research questions were selected based on insights gained from our preliminary literature review as well as our exploratory data analysis conducted which is elaborated on below.

RQ1: Is there a difference between the likelihood of getting **booked (taken into police custody)** amongst different **races and genders**?
RQ2: Is there a difference between the likelihood of getting **strip searched for a specific reason (ex: weapons)** amongst different **races**?
RQ3: What is the effect of **race** on the probability of being **strip searched**, given that the probability of being **cooperative at arrest** is fixed?

Exploratory Data Analysis (EDA)

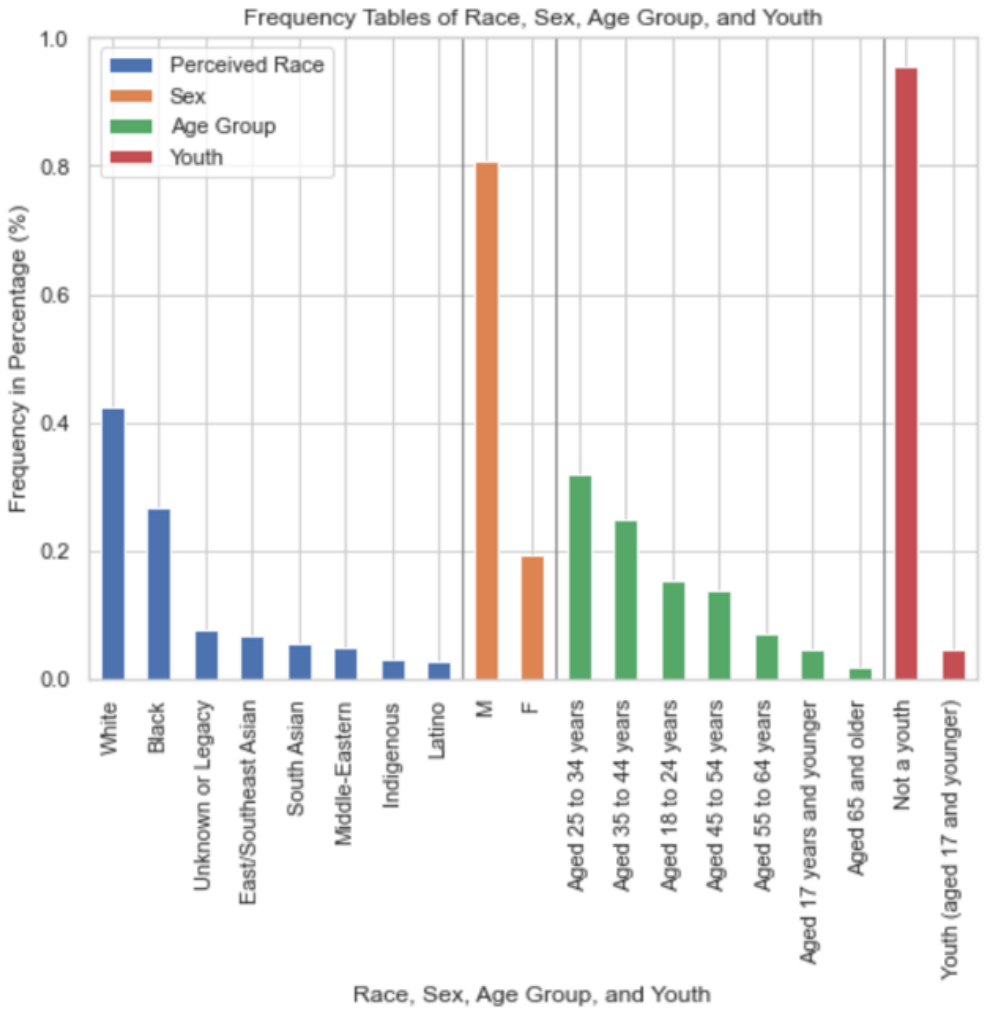
Data Cleaning

Before and during the process of conducting Exploratory Data Analysis (EDA), data cleaning was performed to ensure data accuracy, reliability, and consistency. According to the official Toronto Police Service website (*Arrests and strip searches (RBDC-arr-TBL-001)* 2022), due to a booking system error, there were records that did not indicate a booking when a strip search was taking place. Hence, the value of *Booked* was replaced with 1 when the strip search was 1. For the *Sex* column, there were nine times that the values were "U", which was treated as unknown sex, ending up being dropped. Both *Youth_at_arrest__under_18_years* and *Age_group__at_arrest_* had different values that meant the same, so changes were made so that the values were distinct and consistent. In addition, we found out that the *Perceived_Race* column had several null values, so we replaced them with "Unknown or Legacy".

Descriptive statistics

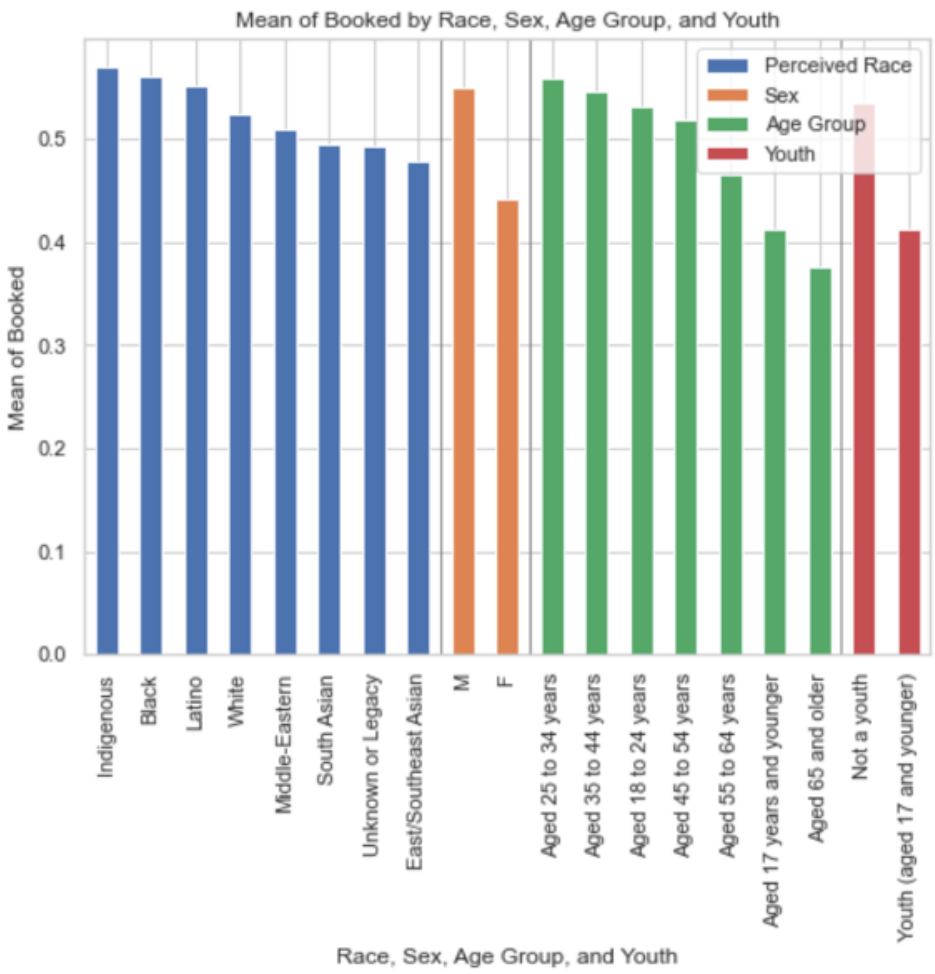
In order to gain more details about the dataset, the following combined frequency graph was made to visualize the selected variables. The graph illustrates the percentage frequency for each of the four variables, including the perceived race, sex, age group, and youth. For perceived race (*Perceived_Race*), the value "White" is the most accruing value, appearing around 40% of the time, followed by the value "Black". And there is a huge discrepancy in frequency between the top two races and the rest. For sex (*Sex*), the number of arrested males is four times the number of arrested females. Approximately 30% of the arrested people were aged from 25 to 34 years, while the age group of 65 and older is the least arrested. Moreover, the majority of people who were arrested are not youths.

Figure 1



In Figure 1, the y-axis is the percentage of frequency. However, we wanted to connect these variables with the mean of being booked as well, so Figure 2 was made. According to this graph, the two most obvious insights are that the mean of being booked for males is much higher than for females and the mean of being booked for non-youth is much higher for youth.

Figure 2



In order to gain insights on perceived race and strip searches, figure 3 was drawn, which is a histogram that shows the number of whether being strip searched or not for each perceived race. According to the graph, most people did not get strip searched, especially those who were perceived as white people. And for people who were strip searched, white people occupy the most in quantity, followed by black people.

Figure 3

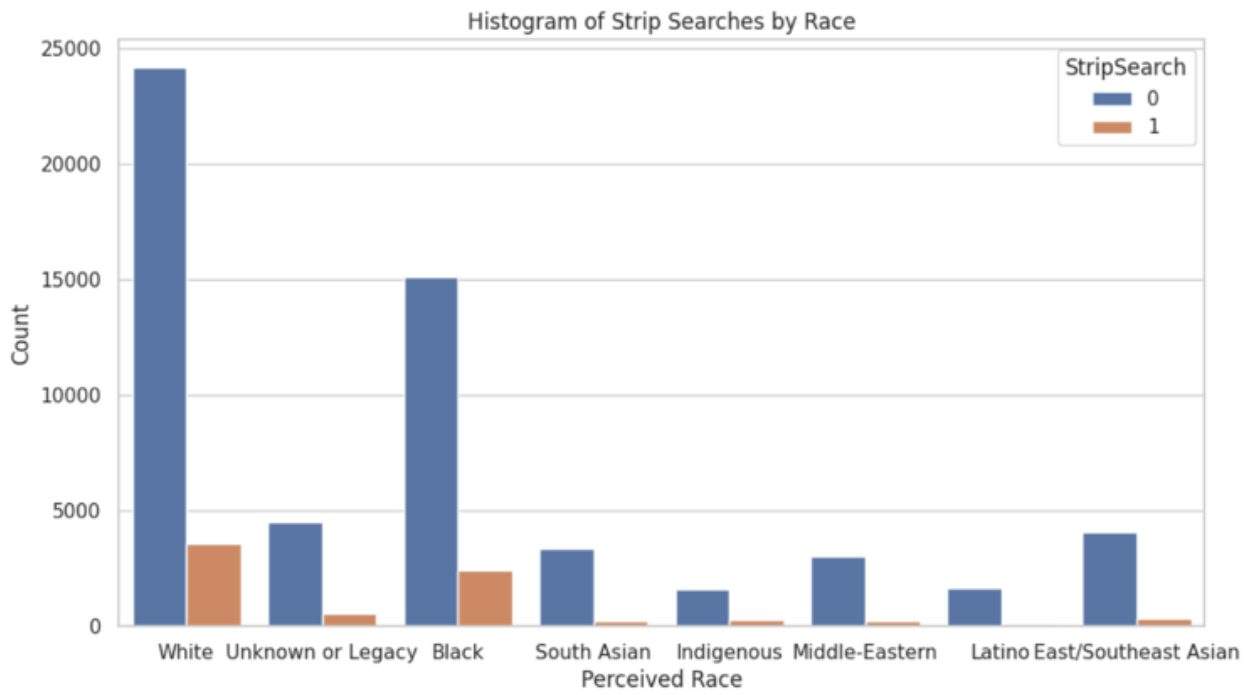


Figure 4 is a contingency table that shows the number of strip searches for each sex within each race. For those who were strip-searched, the number of black males rank first. And for those who were arrested but not strip searched, both black and white males take up great amounts.

Figure 4

	StripSearch	0	1
Perceived_Race	Sex		
Black	F	2757	255
	M	12335	2179
East/Southeast Asian	F	720	25
	M	3354	316
Indigenous	F	555	74
	M	1073	232
Latino	F	254	10
	M	1381	122
Middle-Eastern	F	346	18
	M	2663	210
South Asian	F	482	27
	M	2874	230
Unknown or Legacy	F	830	96
	M	3687	440
White	F	5389	778
	M	18763	2788

We were also curious about the relationship between perceived race and the search reason of possessing weapons. Hence, Figure 5 was drawn to illustrate the number of possessing weapons as the search reason for each race. As we can see from the graph, for all people that were searched because of weapon possession, black and white people take up great proportions.

Figure 5

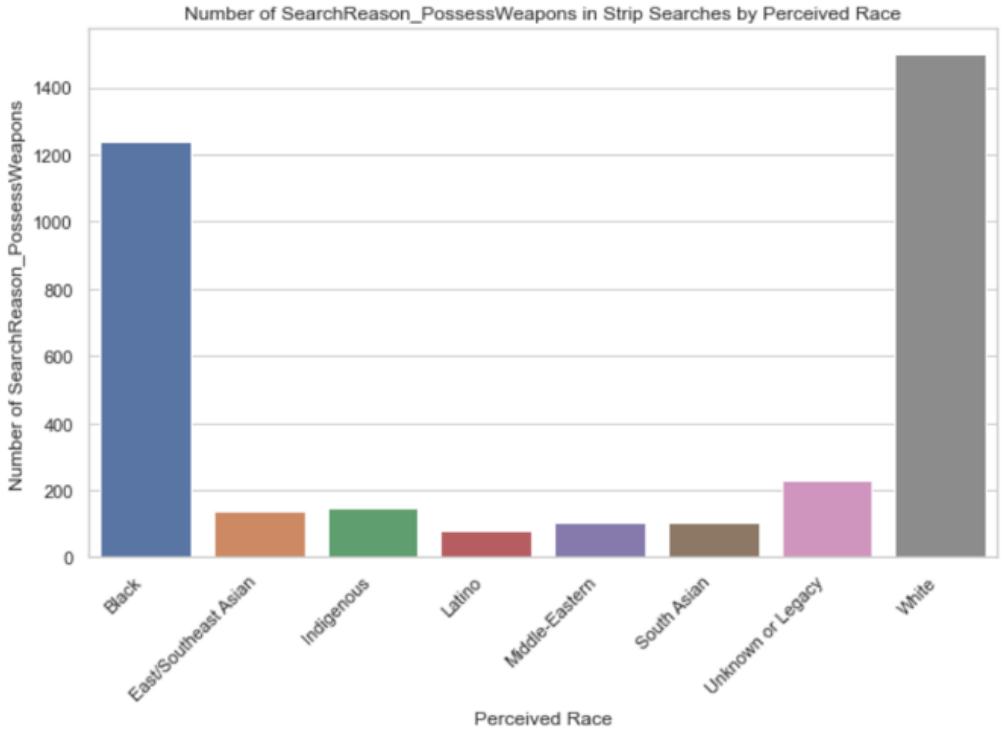
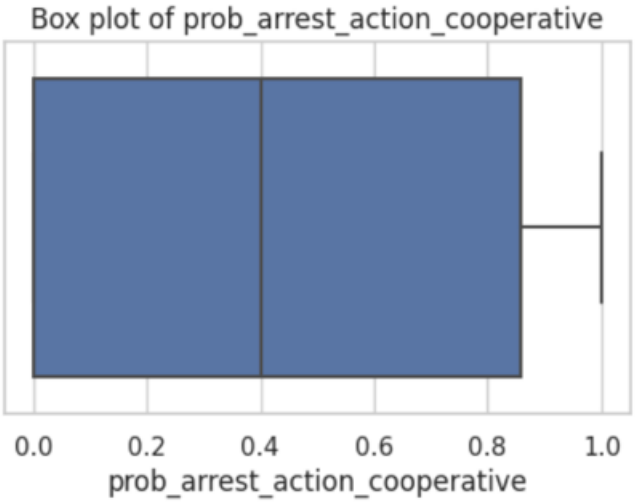


Figure 6 is the box plot for the probability of being cooperative as the action at arrest (*prob_arrest_action_cooperative*). This variable is created by first grouping the dataset based on *PersonID*, and for each *PersonID*, dividing the sum of the values for *Actions_at_arrest___Cooperative* by the number of rows for each person, which represents the probability of being cooperative as the action when the person was arrested. According to the box plot, the median of the probability of being cooperative as the action at arrest is around 0.4, with an interquartile range from 0 to approximately 0.85, as well as a maximum of 1.

Figure 6



T Tests

Prior to the t-tests, we created a continuous variable based on *PersonID*, *Perceived_Race*, and *Booked* columns called *prob_Booked*. The value of this variable is calculated by taking the average of the encoded values for *Booked* over the number of all entries for each unique *PersonID*, representing the probability of being booked for each perceived race within each *PersonID*. We also created a continuous variable called *prob_WeaponSearch* using the same method for creation as the previous variable, and the only difference between them is that we used the column *SearchReason_PossessWeapons* instead of using the column *Booked*. The value of *prob_WeaponSearch* indicates the probability of being strip searched due to the reason of possessing weapons for each perceived race within each *PersonID*. Moreover, another continuous variable called *prob_strip_search* was made, which stands for the probability of being strip searched *(prob_strip_search)*. It was created based on *PersonID* and *StripSearch* columns. The value of this variable is calculated by taking the sum of the encoded values for *StripSearch* over the number of all entries for each unique *PersonID*, representing the probability of being strip searched for each *PersonID*.

In addition, the assumptions for t-tests were checked before conduction. Two assumptions are fulfilled, which are (1) a nominal two-level explanatory variable and (2) a quantitative outcome variable. However, the assumptions of normality and independence of errors are not satisfied, which we will talk more about in the discussion and conclusion section at the end. Moreover, equal variance among the residuals was not assumed since we ran Welch's t-test.

Sex and Probability of Being Booked

We plotted a graph that contained the mean of being booked for males and females, and according to Figure 2, the number is higher for males and females. Instead of comparing the average of being booked for different genders, we wanted to see if the same result applied to the mean probabilities of being booked. Therefore, we conducted a t-test to analyze if there is a

difference in the mean probabilities of being booked (outcome variable) between males and females (two-level explanatory variable). The hypotheses for this test are the following:

H0 (Null Hypothesis): The population means of the probabilities of being booked for males and females are equal.

HA (Alternative Hypothesis): The population means of the probabilities of being booked for males and females are different.

The results indicate that the mean probability of being booked for males (M=0.50, SD=0.40) is higher than the mean probability of being booked for females (M=0.40, SD=0.40). With alpha established at 0.05, this is a statistically significant difference as the p-value (2.0478411815879689e-128) is less than 0.05, 95% CI [0.09, 0.10]. Therefore, we can reject the null hypothesis and conclude that there is a significant difference in the mean probabilities of being booked for males and females.

Race and Probability of Being Strip Searched

We were not only curious about the probability of being booked among races but also the probability of being strip searched among races. There are several distinct races, so we decided to group the races into two groups as done previously, which are white and non-white, in order to find out if they have the same average probabilities. Therefore, we conducted a t-test to see if there is a difference in the mean probabilities of being strip searched (outcome variable) between white and non-white (two-level explanatory variable) groups. The hypotheses for this test are the following:

H0 (Null Hypothesis): The population means of the probabilities of being strip searched for white and non-white people are equal.

HA (Alternative Hypothesis): The population means of the probabilities of being strip searched for white and non-white people are different.

The results indicate that the mean probability of being strip searched for white (M=0.13, SD=0.11) people is lower than the mean probability of being strip searched for non-white (M=0.23, SD=0.24) people. With alpha established at 0.05, this is a statistically significant difference as the p-value (1.1451132400160702e-17) is less than 0.05, 95% CI [0.01, 0.02]. Therefore, we can reject the null hypothesis and conclude that there is a significant difference in the mean probabilities of being strip searched for white and non-white people.

Youth and Probability of Being Booked

For the third t-test, we aimed to determine whether there is a statistically significant difference in the mean probabilities of being booked between youth and non-youth, similar to the previous test. Therefore, we implemented a t-test to discover if there is a difference in the mean probabilities of being booked (outcome variable) between youth and non-youth (two-level explanatory variable). The hypotheses for this test are the following:

H0 (Null Hypothesis): The population means of the probabilities of being booked for youth and non-youth are equal.

HA (Alternative Hypothesis): The population means of the probabilities of being booked for youth and non-youth are different.

As a result, the mean probability of being booked for non-youth (M=0.49, SD=0.40) is higher than the mean probability of being booked for youth (M=0.38, SD=0.38). With alpha established at 0.05, this is a statistically significant difference as the p-value (3.0818073676262326e-47) is less than 0.05, 95% CI [0.09, 0.12]. Therefore, we can reject the null hypothesis and conclude that there is a significant difference in the mean probabilities of being booked for youth and non-youth.

Race and Probability of Being Booked

We were also curious about the probability of being booked among races. However, there are multiple unique races, so we decided to group the races into two groups, which are white and non-white, and figure out if their mean probabilities are the same. Hence, we conducted a t-test to see if there was a difference in the mean probabilities of being booked (outcome variable) between white and non-white (two-level explanatory variable) groups. The hypotheses for this test are the following:

H0 (Null Hypothesis): The population means of the probabilities of being booked for white and non-white people are equal.

HA (Alternative Hypothesis): The population means of the probabilities of being booked for white and non-white people are different.

In the end, the results demonstrate that the mean probability of being booked for white people (M=0.50, SD=0.38) is higher than the mean probability of being booked for non-white people (M=0.47, SD=0.41). With alpha established at 0.05, this is a statistically significant difference as the p-value (1.528610007382848e-25) is less than 0.05, 95% CI [0.03, 0.04]. Therefore, we can reject the null hypothesis and conclude that there is a significant difference in the mean probabilities of being booked for white and non-white people.

Race and Probability of Being Searched for the Reason of Weapon Possession

We plotted the graph that illustrated the number of search reasons due to weapon possession for each race. And according to Figure 5, people whose skin color is white, rank at the top for being searched because of possessing weapons. Therefore, similar to the previous test, we separated race into two groups and conducted a t-test to analyze if there is a difference in the mean probabilities of being searched for the reason of weapon possession (outcome variable) between white and non-white groups (two-level explanatory variable). The hypotheses for this test are the following:

H0 (Null Hypothesis): The population means of the probabilities of being searched for the reason of possessing weapons for white and non-white groups are equal.

HA (Alternative Hypothesis): The population means of the probabilities of being searched for the reason of possessing weapons for white and non-white groups are different.

The results show that the mean probability of being searched for the reason of weapon possession for white (M=0.13, SD=0.30) people is higher than the mean probability of being searched for the reason of weapon possession for non-white (M=0.10, SD=0.28) people. With alpha established at 0.05, this is a statistically significant difference as the p-value (7.84315913031865e-44) is less than 0.05, 95% CI [0.03, 0.04]. Therefore, we can reject the null hypothesis and conclude that there is a significant difference in the mean probabilities of being searched for the reason of possessing weapons for white and non-white people.

Search Reason of Weapon Possession and Probability of Being Booked

We also conducted a t-test to analyze if there is a difference in the mean probabilities of being booked (outcome variable) for the reason of weapon possession and not for the reason of weapon possession (two-level explanatory variable). The hypotheses for this test are the following:

H0 (Null Hypothesis): The population means of the probabilities of being booked for the reason of possessing weapons and not for the reason of possessing weapons are equal.

HA (Alternative Hypothesis): The population means of the probabilities of being booked for the reason of possessing weapons and not for the reason of possessing weapons are different.

The results indicate that the mean probability of being booked for the reason of possessing weapons (M=0.75, SD=0.26) is higher than the mean probability of being booked not for the reason of possessing weapons (M=0.72, SD=0.26). With alpha established at 0.05, this is a statistically significant difference as the p-value (3.355770815754034e-05) is less than 0.05, 95% CI [0.01, 0.04]. Therefore, we can reject the null hypothesis and conclude that there is a significant difference in the mean probabilities of being booked for the reason of possessing weapons and not for the reason of possessing weapons.

Research Design & Methods

Dataset Description

The dataset explored throughout this project contains demographic and arrest-related information on arrests and strip searches that took place in the Toronto police service's jurisdiction. The dataset consists of 65, 276 arrests recorded from 2020 to 2021 which may or may not have resulted in the individual arrested being booked at a police station or strip searched. This dataset has been made publicly available via the Toronto Police Service's Public Safety Data Portal and was last updated on November 10, 2022. Significant demographic characteristics within the dataset include perceived race, sex/gender, and age group. These characteristics are categorical variables displayed in text format. An important arrest-related attribute within the dataset indicates whether the individual was booked at a police station within 24 hours of the arrest which is numerically encoded as Yes = 1, and No = 0. Similarly, whether

the individual was strip searched by a police officer is also numerically encoded as Yes = 1, and No = 0. It is assumed that if an individual was strip searched then they must have been booked as well. Additional information includes indicator columns for the reason(s) attributed for each strip search such as injury caused, escape assisted, weapons possessed and evidence possessed. In each of these columns, the inclusion of the reason is numerically encoded as Yes = 1, and No = 0. Each of these numerically encoded variables is categorical in nature although their values are numerical.

Following the descriptive statistics, combination graphs, and t-tests conducted, we will fit our data into a statistical model in order to begin to answer our research questions.

Power Analysis

In order to assess whether the sample sizes of levels within significant categorical predictor variables are adequate for observing the effect on the outcome variable at a chosen statistical power, we can conduct power analysis. We have set our desired statistical power to a value of 0.8 or 80% as this is the current standard in experimental design. Similarly, we have set our significant level to a value of 0.05 as this is also a current standard within the field. In order to calculate the effect size, we have selected the Cohen's D method as this is an appropriate measure in the case that the standard deviations of the two groups are similar.

The most important predictor present within our research questions is undoubtedly that of perceived race. Therefore, we will be calculating the sample size required for the two most relevant groups within the race predictor variable in order to observe its effect on the probability of being strip searched, which is the outcome we have chosen as a measure of policing behavior. Since the race variable contains many categories, we have condensed the categories into two major groups of interest, white and non-white race for the power analysis.

Table 1: Power Analysis Summary & Results

Effect Size	0.067
Alpha (Significance Level)	0.05
Statistical Power	0.8
Sample Size Needed (White)	4037.011
Actual Sample Size (White)	27718
Sample Size Needed (Non-White)	2980.050
Actual Sample Size (Non-White)	37549

Looking at the results above, it is apparent that the required sample sizes needed for both the white race and non-white race samples are met within our dataset.

Research Question 1: Looking at the results of the t-test comparing the probability of being booked between different racial groups, we can see that with a p-value of 1.528e-25, the differences observed are statistically significant. Similarly, the results for the t-test comparing the probability of being booked between different sexes indicated a statistically significant difference between the mean probabilities amongst different sexes with a p-value of 2.047e-128. Due to our t-tests suggesting that there are statistically significant differences between the probability of getting booked and categorical demographic variables such as sex/gender and race, we will be using a 2-Way ANOVA model to further test for notable effects. In this model, we will have 2 categorical explanatory variables, gender/sex and race, as well as a continuous outcome variable, the probability of getting booked. Through the 2-Way ANOVA, we will also be exploring the potential interaction of gender/sex and race or how the two categorical demographic variables jointly affect changes in the probability of getting booked.

Hypotheses: The following hypotheses were made in the 2-Way ANOVA.

Null Hypothesis: H0: The means (of the probability of being booked) of all races are equal.

Alternative Hypothesis: H1: The mean (of probability of being booked) of at least one race is different.

Null Hypothesis: H0: The means (of probability of being booked) of the sexes/genders are equal.

Alternative Hypothesis: H1: The means (of probability of being booked) of the sexes/genders are different.

Null Hypothesis H0: There is no interaction between race and sex/gender.

Alternative Hypothesis H1: There is an interaction between race and sex/gender.

RQ1: Results & Findings:

Table 2: 2-Way ANOVA Results - Effects of Gender/Sex and Race on Probability of Being Booked

	Sum_squared	Degrees of freedom	F-Statistic	P_value
Perceived Race	223.488014	7.0	209.409039	6.808800e-309

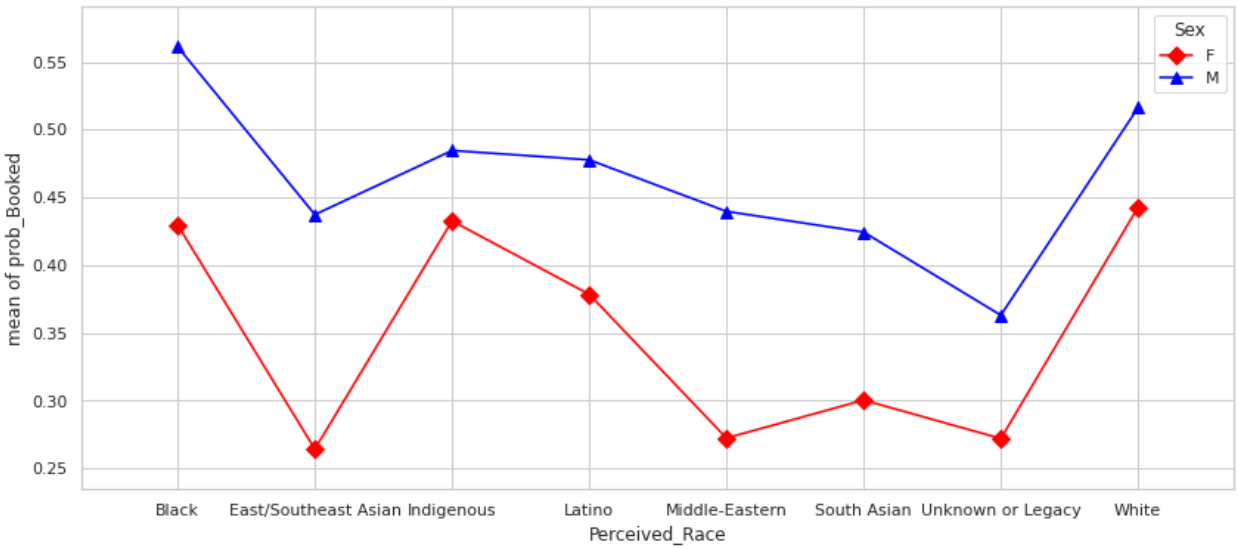
Sex/Gender	101.059944	1.0	662.854612	1.913825e-145
Interaction	11.912211	7.0	11.161783	3.377042e-14
Residual	9948.278712	65251.0	NaN	NaN

In the results obtained from the 2-Way ANOVA (Table 2), the interaction term has a p_value of 3.377e-14. Since the significance level is 0.05 and 3.377e-14 < 0.05, we can reject the null hypothesis and conclude that there is a statistically significant interaction between the sex/gender and race variables. Therefore we can state that the effect of gender/sex on the probability of being booked depends on the racial group being considered, as well as vice versa.

If we were to consider each main effect separate from any interaction effect, then we could make the following conclusions. For the perceived race variable since the p_value is 6.808e-309 which is less than 0.05, we can reject the null hypothesis and conclude that there is a statistically significant difference between the probability of being booked for at least one pair of races. Similarly, for the gender/sex variable, since the p_value is 1.913e-145 which is less than 0.05, we can reject the null hypothesis and conclude that there is a statistically significant difference between the probability of being booked for the two groups present (female and male).

However, due to there being an interaction between the two categorical explanatory variables, we cannot necessarily accept this interpretation as it may be misleading. In order to determine if the above conclusions are likely to be an accurate assessment of the main effects, we decided to create an interaction plot of the variables used in the 2-Way ANOVA.

Figure 7: Interaction Plot for Effects of Race and Gender/Sex on Probability of Being Booked



Looking at the interaction plot (Figure 7) we cannot see a trend where the probability of being booked can be mostly exclusively attributed to a change in either sex/gender or race while there appears to be hardly any change in the other explanatory variable. Therefore, we can conclude that only the interaction should be interpreted as significant and our most notable insight as a result of the 2-Way ANOVA is that both explanatory variables, gender/sex and race, affect the change in probability of being booked caused by the other. The presence of an interaction between the two explanatory variables is further confirmed by the deviation from parallelism observed for the change in probability of being booked for females and males between the following races: east/southeast Asian to unknown or legacy.

As discussed earlier, the main effects can be primarily disregarded as the interaction effect is significant and there is no indication of the significance of the main effects in the interaction plot. Therefore the results of the Tukey HSD test run for the main effect of perceived race (Table 2) are not particularly notable within the context of this project.

Table 3: Post-hoc Tukey HSD Test - Effects of Gender/Sex on Probability of Being Booked

group1	group2	meandiff	p-adj	lower	upper	reject
F	M	0.0956	0.001	0.0879	0.1033	True

Although not our primary focus due to the significant interaction, we conducted a Tukey HSD test for the effects of gender/sex on the probability of being booked (Table 3) as well, just to corroborate the results of the 2-Way ANOVA results since with only 2 groups within the gender/sex variable, we expect the significance of the difference discussed earlier to reappear

within these test results due to it referring to the only possible pairing within gender/sex, F (Female) and M (Male).

Since the interaction between race and gender/sex with regards to their effect on the probability of being booked, was found to be significant, we conducted a post-hoc Tukey HSD test to investigate which specific combinations of interacting pairs/groups contribute to the statistical significance observed. Each set of interacting pairs in Figure 2 with a p_value less than 0.05 was marked as "True" in the "reject" column indicating that the null hypothesis should be rejected and the difference observed is statistically significant. We can see that all interactions involving black males appear to have a significant difference. Similarly, most interactions involving indigenous males seem to have a significant difference as well. These observations may suggest a difference in the probability of certain racial groups and genders/sexes getting booked by a police station.

Looking at the results of the 2-Way ANOVA comparing the probability of being booked between different racial groups as well as between different sexes, we can see that there are statistically significant differences between certain groups such as black/indigenous males and most other possible combinations. With a p-value of $0.001 < 0.05$, black males appeared to have a significantly different probability of being booked when compared to all other groups. Similarly, with p-values of 0.001, 0.0023, $0.0161 < 0.05$, indigenous males appeared to also have a significantly different probability of being booked when compared to most other groups. Since our 2-Way ANOVA suggests that categorical demographic variables such as sex/gender and race have a notable effect on being booked, we will be building a logistic regression model to further investigate if these categorical variables can be used to predict whether someone will be booked or not. In this model, we will have 2 categorical features, gender/sex and race, as well as a categorical outcome variable, whether someone is booked or not (a value of 1 representing being booked and a value of 0 representing not being booked).

Through the logistic regression model, we will be exploring the strength and independence of the correlations between races and sexes/genders with being booked and the subsequent potential for prediction.

Logistic Regression Model:

Table 4: Logistic Regression Results - Effects of Gender/Sex and Race on Being Booked

	Coefficient	Standard Error	P_value	0.025	0.975
Intercept	-0.245	0.022	0.000	-0.288	-0.202

Black	0.126	0.022	0.000	0.083	0.169
Indigenous	0.184	0.053	0.001	0.079	0.288
Latino	0.054	0.056	0.332	-0.055	0.163
Middle Eastern	-0.090	0.042	0.030	-0.171	-0.008
Asian	-0.184	0.028	0.000	-0.240	-0.128
Unknown	-0.154	0.034	0.000	-0.221	-0.086
Male	0.447	0.022	0.000	0.403	0.491

Statistical Significance: The black, asian, indigenous, middle eastern, and male categories have p-values of 0.000/0.001/0.030. Since all these p-values are less than 0.05, we can conclude that there is a statistically significant relationship between being perceived as one of these races/sexes and whether or not a person is booked. In contrast, the latino category has a p-value of 0.332 which is greater than 0.05 and therefore tells us that there is not a statistically significant relationship between being perceived as latino and whether or not a person is booked.

Coefficients: With a coefficient value of 0.126 we can say that being perceived as black is associated with an average increase of 0.126 in the log odds of being booked in comparison to the baseline of having a perceived race of white. Similarly, with a coefficient value of 0.184 we can say that being perceived as indigenous is associated with an average increase of 0.184 in the log odds of being booked in comparison to the baseline of having a perceived race of white. With a coefficient value of -0.09 we can say that being perceived as middle eastern is associated with an average decrease of 0.09 in the log odds of being booked in comparison to the baseline of having a perceived race of white. Similarly, with a coefficient value of -0.184 we can say that being perceived as asian is associated with an average decrease of 0.184 in the log odds of being booked in comparison to the baseline of having a perceived race of white. Lastly, with a coefficient value of 0.447 we can say that the male sex is associated with an average increase of 0.447 in the log odds of being booked in comparison to the baseline of the female sex.

Table 5: Logistic Regression Results - Odds Ratios & Confidence Intervals - Effects of Gender/Sex and Race on Being Booked

	Lower CI	Upper CI	Odds Ratio
Intercept	0.749	0.817	0.783
Black	1.087	1.184	1.134

Indigenous	1.083	1.334	1.202
Latino	0.946	1.177	1.055
Middle Eastern	0.842	0.991	0.914
Asian	0.787	0.880	0.832
Unknown	0.801	0.917	0.857
Male	1.496	1.634	1.564

Odds Ratio/Confidence Intervals: The odds ratio for the black category is 1.134, which indicates that the odds of being booked with a perceived race of black is approximately 1.1 times higher than the odds that someone with a baseline perceived race of white is booked. Similarly, the odds ratio for the indigenous category is 1.202, which indicates that the odds of being booked with a perceived race of indigenous is approximately 1.2 times higher than the odds that someone with a baseline perceived race of white is booked. The odds ratio for the middle eastern category is 0.914, which indicates that the odds of being booked with a perceived race of middle eastern is approximately (1-0.914 = 0.086) times lower than the odds that someone with a baseline perceived race of white is booked. Similarly, The odds ratio for the middle asian category is 0.832, which indicates that the odds of being booked with a perceived race of asian is approximately (1-0.832 = 0.168) times lower than the odds that someone with a baseline perceived race of white is booked. The odds ratio for the male category is 1.564, which indicates that the odds of being booked for the male sex is approximately 1.6 times higher than the odds that someone with a baseline sex of female. The confidence intervals tell us the range within which the odds ratio would be 95% likely to fall if the analysis were to be repeated. For example, if the experiment was repeated, the odds ratio for the indigenous race would be 95% likely to fall between 1.083 and 1.334.

Table 6: Logistic Regression Model Evaluation - Effects of Gender/Sex and Race on Being Booked

<table>
<tr><td>Accuracy Score</td><td colspan="4">0.545</td></tr>
<tr><td rowspan="3">Confusion Matrix</td><td colspan="2" rowspan="2"></td><td colspan="2">Actual</td></tr>
<tr><td>Positive</td><td>Negative</td></tr>
<tr><td rowspan="2">Predicted</td><td>Positive</td><td>1384</td><td>4818</td></tr>
<tr><td>Negative</td><td>1119</td><td>5733</td></tr>
</table>

Based on the accuracy score of 0.545 generated, we can conclude that the logistic regression model built is not very effective at predicting whether someone will be booked or not based on the features used (perceived races, sex/gender) since only approximately 54% of the values predicted are accurate. The confusion matrix generated displays the true positives (1384), false positives (4818), true negatives (5733), and false negatives (1119). From these results we can see that the majority of the error can be attributed to false positives or in other words, the model mistakenly finding instances where someone is booked. This conclusion tells us that there might not be a strong enough correlation between the features chosen and the outcome selected (booked, yes/no) and perhaps the addition of more features with stronger effects on whether someone is booked could improve the model. In addition, there may be too many features (race: latino, middle eastern, asian) that do not have a strong enough correlation that are obscuring any effects previously noted (ex: differences for black/indigenous/male).

Research Question 2: Looking at the results of the t-test comparing the probability of being strip searched with the reasoning of potential weapon possession between different racial groups, we can see that with a p-value of 7.843e-44, the differences observed are statistically significant. Due to our t-test suggesting that there is a statistically significant difference between the probability of being strip searched with the reasoning of potential weapon possession and the categorical demographic variable of race, we will be using a 1-Way ANOVA model to further test for the presence of effects. In this model, we will have 1 categorical explanatory variable, race, as well as a continuous outcome variable, the probability of being strip searched with the reasoning of potential weapon possession.

Hypotheses: The following hypotheses were made in the 1-Way ANOVA.

Null Hypothesis: H0: The means (of the probability of being strip searched with the reasoning of potential weapon possession) of all races are equal.

Alternative Hypothesis: H1: The mean (of the probability of being strip searched with the reasoning of potential weapon possession) of at least one race is different.

RQ2: Results & Findings:

Table 7: 1-Way ANOVA Results - Effects of Race on Probability of Being Strip Searched Due to Suspected Weapon Possession

	Sum_squared	Degrees of freedom	F-Statistic	P_value
Perceived Race	110.815393	7.0	194.066788	3.790043e-286
Residual	5323.426328	65259.0	NaN	NaN

In the results obtained from the 1-Way ANOVA (Table 7), we can evaluate the significance of the effects of race on the probability of being strip searched due to suspected weapon possession. Since the p_value of the race variable is 3.79e-286 and therefore less than the significance level of 0.05 used, we can reject the null hypothesis. Therefore we can conclude that there is a statistically significant difference between the probability of being strip searched due to suspected weapon possession for different racial groups. Since there are more than two racial groups present within the explanatory race variable we conducted a post-hoc Tukey HSD test to determine which pairs of racial groups contributed to the statistically significant difference observed.

By observing the results of the Tukey HSD test for the 1-Way ANOVA conducted (Table 6), we can note that for each racial group pair with a p_value of less than 0.05, we can reject the null hypothesis, which indicates that the difference observed is statistically significant. We can see that most pairs were found to be statistically significant. Furthermore, it appears that all the pairs including either the black or indigenous race are statistically significant regarding their difference in their likelihood of being strip searched due to suspected weapon possession.

Assumption Checking:

The following underlying assumptions associated with the use of ANOVA statistical models were investigated in order to assess the validity of the ANOVA model results.

Assumption 1: Normality - Each sample was taken from a normal population distribution

The normality of the samples or groups was assessed using the Shapiro-Wilk test.

Shapiro-Wilk Test Results:

Statistic = 0.843, p_value = 0.0

Since the p_value is 0.0 which is less than our significance level (alpha = 0.05), we can conclude that the sampled data is not normally distributed.

Assumption 2: Independence of All Groups

For both outcome variables, probability of being booked and probability of being strip searched due to suspected weapon possession, the data included multiple entries for certain person IDs, indicative of multiple arrests for one individual. An individual within the dataset may also be associated with more than one racial group as the race variable refers to the race the individual was perceived to be by the police officer involved. The inclusion of one person in more than one group may inhibit the independence of the dataset and therefore violate this assumption.

Assumption 3: Equal Variances for All Groups

The homogeneity of the variances of the groups was assessed using the Levene test.

Levene Test Results:

Statistic = 86.934, p_value = 1.408e-126

The p_value obtained is less than our significance level of 0.05. Therefore we can conclude that the groups have statistically significant differences in variance, or in other words, the variances are unequal.

Research Question 3: From one of the t-tests above, we found out that there is a significant difference in the mean probabilities of being strip searched for white and non-white people. Hence, we are interested in using the probabilities of being strip searched as the outcome and perceived race as the categorical explanatory variable for the analysis for covariance (ANCOVA). However, instead of trying to discover if there is a statistically significant difference in the probabilities of being strip searched for white and non-white people, we want to investigate if there is a difference across all perceived racial groups. We chose the probability of being cooperative when arrested (*prob_arrest_action_cooperative*) as the continuous exploratory variable. Therefore, our research question for the analysis for covariance (ANCOVA) aims to figure out the effect of race on the probability of being strip searched, given that the probability of being cooperative at arrest is fixed.

Before conducting the analysis for covariance (ANCOVA), two continuous variables were created, which are *prob_strip_search* and *prob_arrest_action_cooperative* respectively. The first self-created variable stands for the probability of being strip searched (*prob_strip_search*), which is the outcome / dependent variable of our model. It was created based on *PersonID* and *StripSearch* columns. The value of this variable is calculated by taking the sum of the encoded values for *StripSearch* over the number of all entries for each unique *PersonID*, representing the probability of being strip searched for each *PersonID*. Similarly, *prob_arrest_action_cooperative* was created using the same method as *prob_strip_search*, but instead of using the column *StripSearch* for calculations, we used the column *Actions_at_arrest___Cooperative*. This newly

created variable is the continuous explanatory variable that represents the probability of being cooperative as the action at arrest.

In addition, we conducted the assumption check prior to running the model. Firstly, we checked for the independence of error. This assumption is not satisfied due to the method of our variable creation. We wanted to calculate the values in a way that could keep as much data as possible, otherwise, problems such as information loss and bias would arise. We ended up keeping almost all of the data in the dataset. However, this choice led to the situation that some rows contained the same values for each of the two self-created variables, causing the problem of dependence on error.

Secondly, we implemented two Shapiro-Wilk tests to check if the normality assumption was met. As a result, neither the outcome nor the continuous explanatory variable followed a normal distribution, meaning that the normality assumption was not satisfied.

Thirdly, we checked the linearity assumption to see if the relationship between the covariate and the outcome is linear, so we plotted a scatter plot with a best-fit line. According to Figure 8, there is a slight negative correlation between *prob_strip_search* and *prob_arrest_action_cooperative*, but no clear indication of linearity is presented. Therefore, the linearity assumption is unmet.

Figure 8

Moreover, we tested for the assumption of homogeneity of variance using Levene's test. The test is used to test if the variances of the probability of being strip searched (*prob_strip_search*) are equal across all the racial groups. The null hypothesis for this test is

equal variance. As a result, the p-value is around 1.60e-156, which is an extremely small value. Therefore, the null hypothesis is rejected, and we can conclude that there is a significant difference in the variance of the probability of being strip searched (*prob_strip_search*) across different racial groups. Thus, the homogeneity of variance assumption is not satisfied.

Lastly, the assumption of homogeneity of regression slopes is checked through making figure 9. This assumption aims to see if the relationship between the covariate (*prob_arrest_action_cooperative)* and the dependent variable (*prob_strip_search*) is the same across all racial groups. Based on Figure 8, the lines represent different racial groups, and they generally appear to be paralleled from each other. Hence, we concluded that this assumption is fulfilled.

Figure 9

After checking the assumptions, the analysis for covariance (ANCOVA) was conducted. The outcome / dependent variable is the probability of being strip searched (*prob_strip_search*), the continuous exploratory variable is the probability of being cooperative as the action at arrest (*prob_arrest_action_cooperative*), and the categorical exploratory variable is perceived race (*Perceived_Race*). The null hypothesis for our model is that there is no significant difference in the mean probability of being strip searched among different racial groups after controlling for the effect of the probability of being cooperative as the action at arrest.

RQ3: Results & Findings:

Table 5: ANCOVA Output

	Source	SS	DF	F	p-unc	np2
0	Perceived_Race	39.567086	7	103.719084	1.212401e-151	0.011003

1	prob_arrest_action_cooperative	2.890023	1	53.030280	3.321353e-13	0.000812
2	Residual	3556.404182	65258	NaN	NaN	NaN

The table above (Table 5) is the output of our analysis. It indicates that the p-value (p-unc) associated with *Perceived_Race* is 1.212401e-151, which is less than 0.05. Therefore, we can reject the null hypothesis and conclude that there is a statistically significant difference in the mean probability of being strip searched across the different races after controlling for the effect of the probability of being cooperative as the action at arrest. This finding suggests that race might be a factor in determining the likelihood of being strip searched. This may raise questions about the presence of discrimination in policing practices and could lead to further investigation or action to address any potential biases in the future.

Discussion & Conclusion

Discussion

Due to the lack of continuous variables in the original dataset, we needed to self-create variable(s) so that Welch's T-test and ANCOVA could be implemented. Before conducting Welch's t-test, we first created two continuous variables and checked if the assumptions of the t-test were met. As mentioned in the Exploratory Data Analysis (EDA) section, two assumptions are not being satisfied, which are normality and independence of error. To find out about the normality, we conducted a Shapiro-Wilk test and found out that the values of both variables we created are not normally distributed. The independence of error assumption is unmet due to the nature of our continuous variable creation method. We wanted to calculate the values in a way that could keep as much data as possible, otherwise, problems such as information loss and bias would arise. We ended up keeping almost all of the data in the dataset. However, this choice led to the situation that some rows contained the same values for each of the two self-created variables, causing the problem of dependence on error.

The assumption check before conducting the analysis for covariance (ANCOVA) indicated that the normality, independence of error, linearity, and homogeneity of variance assumptions were not met. For the normality assumption, we conducted Shapiro-Wilk tests on both the outcome and the continuous exploratory variables, discovering that they were not normally distributed. The independence of error assumption was unmet due to the same reason as the Welch's t-test, which is that we created the continuous variables in a way that could keep the most original data to reduce potential data loss and bias. Since the scatter plot of the outcome and the continuous exploratory variable did not illustrate a clear sign of linearity, it failed to satisfy this assumption. And for the homogeneity of variance, Levene's test was conducted, and we received a small p-value, which means that there is a significant difference in the variance of

the outcome across different racial groups. However, the assumption of homogeneity of regression slopes was fulfilled due to the generally paralleled lines in Figure 8.

In addition, there is an issue that would cause a discrepancy between the result of our analysis and the reality that the dataset records the perceived race of the arrested people instead of their true race. Perceived race is very subjective and different officers could perceive the same person in different races, which may or may not be their true race, leading to the inaccuracy of someone's race.

Conclusion

For the midterm project, we have two research questions in total. Our first research question aims to investigate if there is a difference between the likelihood of getting booked amongst different races and genders. According to our findings, there are interactions between gender and racial group. For example, there appears to be a notable difference in all interactions involving black males, as well as in most interactions involving indigenous males. In conclusion, the impact of gender/sex on the likelihood of being booked varies depending on the specific racial group under consideration. The second research question focuses on finding out if the probability of being strip searched with the reasoning of weapon possession of all races is equal. This is concluded that there is a significant difference in the probability of being strip searched due to weapon possession among different races. And similar to the findings for the first research question, the analysis indicates that there is a significant difference in the likelihood of being strip searched due to suspected weapon possession between all pairs of racial groups that include either Black or Indigenous individuals, which is interesting. In summary, the findings of both research questions indicate that there is a significant difference as long as the pairs involve black and indigenous people. This result potentially suggests and signals that black and indigenous people may be unfairly targeted and treated with regard to getting booked. Overall, our findings highlight the need for continued research and policy changes aimed at addressing disparities in the criminal justice system and ensuring that all individuals are treated fairly and equitably, regardless of their race or gender.

For the final project, we are focused on two research questions as well. Firstly, we want to see if there is a significant difference in the mean probability of being strip searched across the different races after controlling for the effect of the probability of being cooperative as the action at arrest. The objective of this research question is to figure out if there could be a potential racial bias in strip searches. As a result, the p-value associated with perceived race is very small, indicating that there is a significant difference in the probability of being strip searched across different racial groups. This suggests that there may be racial bias in the use of strip searches. However, this result should be taken with caution and consider other factors that could potentially lead to the difference, so further investigation is needed to come up with a more accurate conclusion in reality. Secondly, investigating the strength between races and sexes with the probability of being booked is another focus of ours. According to the results, being black or indigenous is associated with an average increase in the log odds of being booked in comparison

to the baseline of having a perceived race of white. In addition, the results also indicate that the odds ratio for the black and indigenous groups are both higher than 1, which means that the odds of being booked with a perceived race of black or indigenous is higher than the odds that someone with a baseline perceived race of white is booked. In summary, it is suggested that on average, individuals who are perceived as black or indigenous have a higher probability of being booked by law enforcement officers compared to individuals who are perceived as white. Therefore, there could be racial bias in the likelihood of being booked by law enforcement officers which is worth more investigation. Moreover, the results also show that on average, males are at higher risk of being booked than females. This may signal a gender bias but more investigations are needed to better understand the relationship between gender and booking. However, the logistic model only has an accuracy score of 54.5%, which indicates that it is not highly effective on the probability of being booked based on race and gender. The false positive contributes to most of the inaccuracy, meaning that the model incorrectly predicts that a person will be booked when they are not. Hence, actions such as considering other factors could be a way of improving the model.

For the continuation and future expansion of this project, the officials should ask the arrested people for their races and record it in the dataset instead of documenting the perceived race because the result findings could be biased or deviated if the dataset is analyzed, potentially leading to inaccurate decision and policy making and dissemination of information. Moreover, further investigations should be made to confirm if there are racial and gender biases. Lastly, officials should keep making efforts to improve and promote social justice, especially for black and indigenous people, creating a better society for all.

Appendix

Table 2: Post-hoc Tukey HSD Test - Effects of Race on Probability of Being Booked

group1	group2	meandiff	p-adj	lower	upper	reject
Black	East/Southeast Asian	-0.1304	0.001	-0.1505	-0.1104	True
Black	Indigenous	-0.0709	0.001	-0.0994	-0.0423	True
Black	Latino	-0.0757	0.001	-0.1054	-0.046	True
Black	Middle-Eastern	-0.1177	0.001	-0.1404	-0.0949	True
Black	South Asian	-0.1316	0.001	-0.1534	-0.1099	True
Black	Unknown or Legacy	-0.1922	0.001	-0.2112	-0.1732	True
Black	White	-0.0387	0.001	-0.0502	-0.0272	True
East/Southeast Asian	Indigenous	0.0596	0.001	0.0271	0.092	True
East/Southeast Asian	Latino	0.0548	0.001	0.0213	0.0883	True
East/Southeast Asian	Middle-Eastern	0.0128	0.8401	-0.0148	0.0403	False

East/Southeast Asian	South Asian	-0.0012	0.9	-0.0279	0.0255	False
East/Southeast Asian	Unknown or Legacy	-0.0617	0.001	-0.0862	-0.0372	True
East/Southeast Asian	White	0.0917	0.001	0.0724	0.111	True
Indigenous	Latino	-0.0048	0.9	-0.044	0.0343	False
Indigenous	Middle-Eastern	-0.0468	0.001	-0.081	-0.0126	True
Indigenous	South Asian	-0.0608	0.001	-0.0943	-0.0272	True
Indigenous	Unknown or Legacy	-0.1213	0.001	-0.1531	-0.0895	True
Indigenous	White	0.0321	0.0117	0.0042	0.0601	True
Latino	Middle-Eastern	-0.042	0.0072	-0.0772	-0.0068	True
Latino	South Asian	-0.0559	0.001	-0.0905	-0.0214	True
Latino	Unknown or Legacy	-0.1165	0.001	-0.1494	-0.0836	True
Latino	White	0.037	0.0031	0.0078	0.0662	True
Middle-Eastern	South Asian	-0.0139	0.8025	-0.0427	0.0149	False

Middle-Eastern	Unknown or Legacy	-0.0745	0.001	-0.1013	-0.0477	True
Middle-Eastern	White	0.079	0.001	0.0569	0.1011	True
South Asian	Unknown or Legacy	-0.0605	0.001	-0.0865	-0.0346	True
South Asian	White	0.0929	0.001	0.0719	0.114	True
Unknown or Legacy	White	0.1535	0.001	0.1353	0.1717	True

Table 3: Post-hoc Tukey HSD Test - Effects of Interaction of Gender/Sex and Race on Probability of Being Booked

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Black / F	Black / M	0.1315	0.001	0.1047	0.1583	True
Black / F	East/Southeast Asian / F	-0.1654	0.001	-0.2201	-0.1107	True
Black / F	East/Southeast Asian / M	0.0077	0.9	-0.0252	0.0406	False
Black / F	Indigenous / F	0.003	0.9	-0.0557	0.0616	False
Black / F	Indigenous / M	0.055	0.0023	0.0106	0.0993	True
Black / F	Latino / F	-0.0512	0.7728	-0.137	0.0347	False
Black / F	Latino / M	0.0481	0.0095	0.0058	0.0903	True
Black / F	Middle-Eastern / F	-0.1573	0.001	-0.2315	-0.083	True
Black / F	Middle-Eastern / M	0.01	0.9	-0.0249	0.0449	False
Black / F	South Asian / F	-0.1292	0.001	-0.1934	-0.0651	True
Black / F	South Asian / M	-0.0052	0.9	-0.0395	0.029	False
Black / F	Unknown or Legacy / F	-0.1576	0.001	-0.2079	-0.1074	True
Black / F	Unknown or Legacy / M	-0.0666	0.001	-0.0986	-0.0345	True
Black / F	White / F	0.0131	0.9	-0.0167	0.0428	False
Black / F	White / M	0.0865	0.001	0.0605	0.1126	True
Black / M	East/Southeast Asian / F	-0.2969	0.001	-0.3472	-0.2467	True
Black / M	East/Southeast Asian / M	-0.1238	0.001	-0.1485	-0.0991	True
Black / M	Indigenous / F	-0.1285	0.001	-0.183	-0.074	True
Black / M	Indigenous / M	-0.0766	0.001	-0.1152	-0.0379	True
Black / M	Latino / F	-0.1827	0.001	-0.2658	-0.0996	True
Black / M	Latino / M	-0.0834	0.001	-0.1197	-0.0472	True
Black / M	Middle-Eastern / F	-0.2888	0.001	-0.3598	-0.2178	True
Black / M	Middle-Eastern / M	-0.1215	0.001	-0.1488	-0.0941	True
Black / M	South Asian / F	-0.2607	0.001	-0.3211	-0.2004	True
Black / M	South Asian / M	-0.1368	0.001	-0.1632	-0.1103	True
Black / M	Unknown or Legacy / F	-0.2891	0.001	-0.3345	-0.2438	True
Black / M	Unknown or Legacy / M	-0.1981	0.001	-0.2217	-0.1745	True
Black / M	White / F	-0.1184	0.001	-0.1388	-0.0981	True
Black / M	White / M	-0.045	0.001	-0.0593	-0.0306	True
East/Southeast Asian / F	East/Southeast Asian / M	0.1731	0.001	0.1193	0.2268	True
East/Southeast Asian / F	Indigenous / F	0.1684	0.001	0.0959	0.2408	True
East/Southeast Asian / F	Indigenous / M	0.2204	0.001	0.1589	0.2818	True
East/Southeast Asian / F	Latino / F	0.1142	0.0045	0.0184	0.2101	True
East/Southeast Asian / F	Latino / M	0.2135	0.001	0.1535	0.2734	True
East/Southeast Asian / F	Middle-Eastern / F	0.0081	0.9	-0.0774	0.0937	False
East/Southeast Asian / F	Middle-Eastern / M	0.1754	0.001	0.1204	0.2304	True
East/Southeast Asian / F	South Asian / F	0.0362	0.9	-0.0408	0.1131	False
East/Southeast Asian / F	South Asian / M	0.1602	0.001	0.1056	0.2147	True
East/Southeast Asian / F	Unknown or Legacy / F	0.0078	0.9	-0.0581	0.0736	False
East/Southeast Asian / F	Unknown or Legacy / M	0.0988	0.001	0.0456	0.1521	True
East/Southeast Asian / F	White / F	0.1785	0.001	0.1266	0.2304	True
East/Southeast Asian / F	White / M	0.252	0.001	0.2021	0.3018	True
East/Southeast Asian / M	Indigenous / F	-0.0047	0.9	-0.0624	0.053	False
East/Southeast Asian / M	Indigenous / M	0.0473	0.0161	0.0042	0.0904	True

East/Southeast Asian / M	Latino / F	-0.0588	0.5634	-0.1441	0.0264	False
East/Southeast Asian / M	Latino / M	0.0404	0.0586	-0.0006	0.0814	False
East/Southeast Asian / M	Middle-Eastern / F	-0.1649	0.001	-0.2385	-0.0914	True
East/Southeast Asian / M	Middle-Eastern / M	0.0024	0.9	-0.031	0.0357	False
East/Southeast Asian / M	South Asian / F	-0.1369	0.001	-0.2002	-0.0736	True
East/Southeast Asian / M	South Asian / M	-0.0129	0.9	-0.0455	0.0197	False
East/Southeast Asian / M	Unknown or Legacy / F	-0.1653	0.001	-0.2145	-0.1161	True
East/Southeast Asian / M	Unknown or Legacy / M	-0.0742	0.001	-0.1046	-0.0439	True
East/Southeast Asian / M	White / F	0.0054	0.9	-0.0225	0.0333	False
East/Southeast Asian / M	White / M	0.0789	0.001	0.055	0.1028	True
Indigenous / F	Indigenous / M	0.052	0.3042	-0.013	0.1169	False
Indigenous / F	Latino / F	-0.0541	0.8702	-0.1523	0.044	False
Indigenous / F	Latino / M	0.0451	0.5208	-0.0185	0.1086	False
Indigenous / F	Middle-Eastern / F	-0.1603	0.001	-0.2484	-0.0721	True
Indigenous / F	Middle-Eastern / M	0.0071	0.9	-0.0518	0.0659	False
Indigenous / F	South Asian / F	-0.1322	0.001	-0.212	-0.0525	True
Indigenous / F	South Asian / M	-0.0082	0.9	-0.0667	0.0503	False
Indigenous / F	Unknown or Legacy / F	-0.1606	0.001	-0.2297	-0.0915	True
Indigenous / F	Unknown or Legacy / M	-0.0696	0.0033	-0.1268	-0.0123	True
Indigenous / F	White / F	0.0101	0.9	-0.0459	0.0661	False
Indigenous / F	White / M	0.0836	0.001	0.0294	0.1377	True
Indigenous / M	Latino / F	-0.1061	0.0057	-0.1964	-0.0158	True
Indigenous / M	Latino / M	-0.0069	0.9	-0.0575	0.0437	False
Indigenous / M	Middle-Eastern / F	-0.2122	0.001	-0.2915	-0.1329	True
Indigenous / M	Middle-Eastern / M	-0.0449	0.047	-0.0896	-0.0003	True
Indigenous / M	South Asian / F	-0.1842	0.001	-0.2541	-0.1143	True
Indigenous / M	South Asian / M	-0.0602	0.001	-0.1043	-0.0161	True
Indigenous / M	Unknown or Legacy / F	-0.2126	0.001	-0.2701	-0.1551	True
Indigenous / M	Unknown or Legacy / M	-0.1215	0.001	-0.164	-0.079	True
Indigenous / M	White / F	-0.0419	0.037	-0.0826	-0.0011	True
Indigenous / M	White / M	0.0316	0.2471	-0.0065	0.0697	False
Latino / F	Latino / M	0.0992	0.0134	0.0099	0.1885	True
Latino / F	Middle-Eastern / F	-0.1061	0.0615	-0.2143	0.002	False
Latino / F	Middle-Eastern / M	0.0612	0.5166	-0.0248	0.1472	False
Latino / F	South Asian / F	-0.0781	0.3756	-0.1796	0.0234	False
Latino / F	South Asian / M	0.0459	0.9	-0.0399	0.1317	False
Latino / F	Unknown or Legacy / F	-0.1065	0.0091	-0.1998	-0.0131	True
Latino / F	Unknown or Legacy / M	-0.0154	0.9	-0.1003	0.0695	False
Latino / F	White / F	0.0642	0.3896	-0.0198	0.1483	False
Latino / F	White / M	0.1377	0.001	0.0549	0.2206	True
Latino / M	Middle-Eastern / F	-0.2053	0.001	-0.2835	-0.1272	True
Latino / M	Middle-Eastern / M	-0.038	0.1451	-0.0806	0.0046	False
Latino / M	South Asian / F	-0.1773	0.001	-0.2459	-0.1087	True
Latino / M	South Asian / M	-0.0533	0.0015	-0.0953	-0.0113	True
Latino / M	Unknown or Legacy / F	-0.2057	0.001	-0.2616	-0.1498	True
Latino / M	Unknown or Legacy / M	-0.1146	0.001	-0.1549	-0.0743	True
Latino / M	White / F	-0.035	0.125	-0.0735	0.0035	False
Latino / M	White / M	0.0385	0.0202	0.0028	0.0742	True
Middle-Eastern / F	Middle-Eastern / M	0.1673	0.001	0.0929	0.2417	True
Middle-Eastern / F	South Asian / F	0.028	0.9	-0.0638	0.1199	False
Middle-Eastern / F	South Asian / M	0.152	0.001	0.0779	0.2261	True
Middle-Eastern / F	Unknown or Legacy / F	-0.0004	0.9	-0.0831	0.0824	False
Middle-Eastern / F	Unknown or Legacy / M	0.0907	0.0023	0.0176	0.1638	True
Middle-Eastern / F	White / F	0.1703	0.001	0.0982	0.2425	True
Middle-Eastern / F	White / M	0.2438	0.001	0.1731	0.3145	True
Middle-Eastern / M	South Asian / F	-0.1393	0.001	-0.2036	-0.0749	True
Middle-Eastern / M	South Asian / M	-0.0153	0.9	-0.0499	0.0194	False
Middle-Eastern / M	Unknown or Legacy / F	-0.1677	0.001	-0.2182	-0.1171	True
Middle-Eastern / M	Unknown or Legacy / M	-0.0766	0.001	-0.1091	-0.0441	True
Middle-Eastern / M	White / F	0.003	0.9	-0.0272	0.0333	False
Middle-Eastern / M	White / M	0.0765	0.001	0.0499	0.1031	True
South Asian / F	South Asian / M	0.124	0.001	0.06	0.188	True
South Asian / F	Unknown or Legacy / F	-0.0284	0.9	-0.1022	0.0454	False
South Asian / F	Unknown or Legacy / M	0.0627	0.0515	-0.0002	0.1255	False
South Asian / F	White / F	0.1423	0.001	0.0806	0.204	True
South Asian / F	White / M	0.2158	0.001	0.1558	0.2758	True
South Asian / M	Unknown or Legacy / F	-0.1524	0.001	-0.2025	-0.1023	True
South Asian / M	Unknown or Legacy / M	-0.0613	0.001	-0.0931	-0.0295	True
South Asian / M	White / F	0.0183	0.714	-0.0111	0.0478	False
South Asian / M	White / M	0.0918	0.001	0.0661	0.1175	True
Unknown or Legacy / F	Unknown or Legacy / M	0.0911	0.001	0.0424	0.1397	True
Unknown or Legacy / F	White / F	0.1707	0.001	0.1236	0.2178	True
Unknown or Legacy / F	White / M	0.2442	0.001	0.1993	0.2891	True
Unknown or Legacy / M	White / F	0.0796	0.001	0.0527	0.1065	True
Unknown or Legacy / M	White / M	0.1531	0.001	0.1304	0.1758	True
White / F	White / M	0.0735	0.001	0.0542	0.0928	True

Table 6: Post-hoc Tukey HSD Test - Effects of Race on Probability of Being Strip Searched Due to Suspected Weapon Possession

group1	group2	meandiff	p-adj	lower	upper	reject
Black	East/Southeast Asian	-0.0959	0.001	-0.1104	-0.0813	True
Black	Indigenous	0.0316	0.001	0.0109	0.0524	True
Black	Latino	-0.0774	0.001	-0.099	-0.0558	True
Black	Middle-Eastern	-0.0913	0.001	-0.1079	-0.0747	True
Black	South Asian	-0.1005	0.001	-0.1163	-0.0846	True
Black	Unknown or Legacy	-0.1016	0.001	-0.1154	-0.0878	True
Black	White	-0.0124	0.001	-0.0207	-0.004	True
East/Southeast Asian	Indigenous	0.1275	0.001	0.1039	0.1511	True
East/Southeast Asian	Latino	0.0184	0.298	-0.0059	0.0428	False
East/Southeast Asian	Middle-Eastern	0.0046	0.9	-0.0155	0.0246	False
East/Southeast Asian	South Asian	-0.0046	0.9	-0.024	0.0148	False

East/Southeast Asian	Unknown or Legacy	-0.0057	0.9	-0.0235	0.0121	False
East/Southeast Asian	White	0.0835	0.001	0.0695	0.0975	True
Indigenous	Latino	-0.1091	0.001	-0.1375	-0.0806	True
Indigenous	Middle-Eastern	-0.1229	0.001	-0.1478	-0.098	True
Indigenous	South Asian	-0.1321	0.001	-0.1565	-0.1077	True
Indigenous	Unknown or Legacy	-0.1332	0.001	-0.1563	-0.11	True
Indigenous	White	-0.044	0.001	-0.0643	-0.0236	True
Latino	Middle-Eastern	-0.0138	0.6993	-0.0395	0.0118	False
Latino	South Asian	-0.023	0.1008	-0.0482	0.0021	False
Latino	Unknown or Legacy	-0.0241	0.0463	-0.0481	-0.0002	True
Latino	White	0.0651	0.001	0.0438	0.0863	True
Middle-Eastern	South Asian	-0.0092	0.8861	-0.0301	0.0118	False
Middle-Eastern	Unknown or Legacy	-0.0103	0.7233	-0.0298	0.0092	False

Middle-Eastern	White	0.0789	0.001	0.0628	0.095	True
South Asian	Unknown or Legacy	-0.0011	0.9	-0.02	0.0177	False
South Asian	White	0.0881	0.001	0.0728	0.1034	True
Unknown or Legacy	White	0.0892	0.001	0.076	0.1024	True

Reference

Arrests and strip searches (RBDC-arr-TBL-001). Toronto Police Service Public Safety Data Portal. (2022, November 10). Retrieved February 27, 2023, from https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/about

Gouvernement du Canada. (2021, June 23). Government of Canada. Canada.ca. Retrieved February 28, 2023, from https://www.canada.ca/en/canadian-heritage/campaigns/anti-racism-engagement/anti-racism-strategy.html

Iderd: Fighting racism means looking back as well as looking forward. Canadian Labour Congress. (2022, March 22). Retrieved February 28, 2023, from https://canadianlabour.ca/iderd-fighting-racism-means-looking-back-as-well-as-looking-forward/

The move towards an intersectional approach. The move towards an intersectional approach | Ontario Human Rights Commission. (n.d.). Retrieved February 28, 2023, from https://www.ohrc.on.ca/en/intersectional-approach-discrimination-addressing-multiple-grounds-human-rights-claims/move-towards-intersectional-approach

Wortley, S., & Tanner, J. (2004). Discrimination or 'good' policing? The racial profiling debate in Canada. Our Diverse Cities, 1(Spring), 197-201.

Intersectionality of race, ethnicity, gender, and age on criminal ... (n.d.). Retrieved April 17, 2023, from https://journals.sagepub.com/doi/abs/10.1177/0731121416679371

Pica, E., Thompson, L. E., Pozzulo, J., & Sheahan, C. L. (2019, September 12). *Perceptions of police conduct when race and gender are considered - journal of police and criminal psychology*. SpringerLink. Retrieved April 16, 2023, from https://link.springer.com/article/10.1007/s11896-019-09346-1

Giwa, S., Mullings, D. V., Adjei, P. B., & Karki, K. K. (2020, September 16). *Racial erasure: The silence of social work on police racial profiling in Canada - journal of human rights and social work*. SpringerLink. Retrieved April 16, 2023, from https://link.springer.com/article/10.1007/s41134-020-00136-y