A Multifaceted Investigation of Strip Search Practices: Disentangling the Effects of Race, Age, Gender and Arrest History

Kaiyang Bian: 1003138464
Ruiyue Ma: 1009705240

A final paper submitted in conformity with the requirements for
the course of INF2178H
Faculty of Information
University of Toronto

Introduction

As an intrusive and compulsive practice for criminals at arrest, strip searches have faced controversial issues related to criminal justice, fairness and human rights. According to the Supreme Court of Canada, the strip search was identified as "one of the most extreme exercises of police power" and "inherently humiliating and degrading"(R.v. Golden, 2001). Due to the invasive nature of this practice, it leads to the high chance of trauma for the suspects and prisoners. A study found that strip searches impose harm to people with a history of trauma, and nearly half of prisoners incarcerated in Canada have experienced childhood abuse(Canadian Civil Liberties Association, 2022). The strip searches provokes the traumatized experience of those, which might cause more serious psychological damage. In addition, those who were arrested or prisoned might be innocent, waiting for trial or release on bail. Under this circumstance, the strip searches result in a secondary victimization for the innocent.

On the other hand, strip searches are believed to be useful practice for the police to ensure the person was not concealing any dangerous weapons that may cause possible injury to themselves, officers and other individuals whom they may be held in custody with(Toronto Police Service, 2022). With this reason, administrative officials in Ontario have unrestricted authority to approve strip searches at any time and in any situation according to the Ministry of Correctional Services Act and regulations(Canadian Civil Liberties Association, 2022). The law and regulation gives police officers in Ontario permission to determine the search target, place and reason for conducting strip searches. The report from the Ontario Independent Police Review Director indicated that there were an excessive amount of unjustified and illegal strip searches conducted by police in Ontario (McNeilly, 2019).

With the rising concern on the impact of strip searches, it's necessary to understand the rationale and patterns of strip searches, in order to maintain criminal justice and protect individuals from social prejudices, racial or sexual discrimination. How does race influence the frequency of being subjected to a strip search during an arrest? Are marginalized groups more likely to be subjected to this intrusive practice due to social bias? Does the strip search rate differ by gender and age groups? Are individuals with higher arrest frequency more likely to be strip searched?

Does the effects of strip search, race, age, and sex have an impact on the likelihood of items found at arrest? With deeper exploration into the research questions, this report tries to investigate the patterns of strip searches activities on the arrested individual in Ontario, providing support for the legislation authorities in Ontario on maintaining criminal and social justice. More details regarding the selected indicators are discussed further in the literature review.

The dataset of Arrests and Strip Searches was retrieved from the Public Safety Data Portal by the Toronto Police Service on a total of nearly 65000 arrested individuals in the GTA areas. The project applies ANOVA, ANCOVA and Logistic Regression models to study the relationship between predictors and outcomes.

Literature Review

The report from the Race and Identity-based Data Collection(2020) showed that the strip searches might enhance social bias as it targets racialized Canadians. The result found that one-third of people who were strip-searched at arrest are Black (Lemke, 2022), and Indigenous people were 1.3 times over-represented in strip searches(Race-Based Data Collection, 2020). The impact of racial profiling is not limited to countries of immigrants. Research conducted by the British Journal of Criminology (Newburn et al., 2004) has shown that the British police force frequently used the "stop and search' power on the African-Caribbean arrestees regardless of their age and offense in North London for the period of 1999 to 2000. In addition, there were raising concerns on whether there was sexual bias on the subject of strip searches. A study from the Boston Law University Review indicated that most strip searches at arrests targeted men of colour (Cooper, 2013). In comparison to the 43 female cases, 124 male cases alleging illegal strip searches were found in the analysis of criminal cases recorded in the Quicklaw and CanLil databases referencing the Golden decision(Psutka& Sheehy, 2016). This report highlighted that males have a higher possibility of experiencing an illegal strip search, compared to females. Moreover, a statistical analysis showed that the arrestees who were identified as young adults, typically aged between 17 to 23, were more likely to undergo a strip search than other age groups (Ha, 2011). The age pattern remained unchanged over time in different regions and continents as

the report from the Analysis and Policy Observatory (Grewock & Sentas, 2019) indicated that nearly 45% of the recorded strip searches were conducted on young adults aged 25 years and under in Australia during the year of 2017 and 2018.

Aside from the personal identities of the arrestees, studies showed that police suspicion contributed to the conduct of strip searches at arrest. According to the Columbia Journal of Law and Social Problems (MacGregor, 2003), the Nassau County Correctional Center, a primary detention facility for individuals who are arrested and held in custody in the United States, claimed that officials did not strip search individuals charged with misdemeanors and minor offenses unless they had reasonable suspicion that the arrestees were concealing weapons or contrabands. In addition, studies found that strip searches might be associated with types of offenses. The study indicated that drug related offenses accounted for 91% of all recorded reasons for police officers to conduct a strip search (Grewock & Sentas, 2019) . However, a report (Lemke, 2022) shows that a vast amount of items were discovered via a frisk search before contemplating a strip search, which leads to the doubt on whether conducting a strip search is indeed necessary for the searching procedures. Furthermore, the statistics reported to the Police Service Board showed that only 2% of the chance that objects were found during the strip and cavity searches. Among the objects found, only a small portion of items posed a risk at arrest (Lemke, 2022).

With the previous research and findings, it's worth investigating the impacts of personal traits and other relevant attributes on the strip search results due to the social bias and ethical concerns. Additionally, it helps to study whether the strip search was an effective conduct for the Toronto police officer to mitigate the risks of the arrestee at arrest.

Data Description

In our research project, we use the dataset that demonstrates information related to all arrests and strip searches conducted by the Toronto Police Service during the period of 2020 to 2010. The dataset contains a total of 65276 records of arrests and strip searches and 24 attributes related to information such as the arrest year and month, race, gender, age groups at arrest, arrest location division, strip search results, occurrence category, action at arrest and strip search reason. This dataset is accessible from the Public Safety Data Portal from Toronto Police Service collected in November 2022. The important attributes related to our study will be the strip search results, strip search reason, perceived race, sex and occurrence category. The original dataset consists of either ordinal data (such as Arrest ID, etc.), binary data in 0 or 1 format(strip search result, various search reasons etc.), or nominal data (perceived race, sex, age groups, occurrence category etc.). There are other missing values of some of the attributes, such as search reasons and items found. Therefore, data cleaning and preprocessing need to be conducted before our analysis.

Research Objective and Questions

Our study seeks to investigate if the strip searches rate differ by personal identities of the arrestees, which includes their race, gender and age at arrest. By conducting the analysis on relationship betweenstrip searched frequency and those attributes, we aimed to study if the perceived racial and sexual bias, as well as the age discrimination could impose obstruction of criminal justice. In addition, we investigated if the likelihood of items found at arrest differ due to the effect of factors including strip search outcome and other personal traits of the arrestee. By analyzing how the likelihood of items found at arrest varies by strip search outcome and other predictor variables, it can help the regulators and law enforcement agencies evaluate the effectiveness of their current search practices. Our research questions are constructed based on knowledge gained from our literature review and preliminary analysis of the data.

RQ1: *Are there any differences in strip search rate for different perceived races (such as white, black, indigenous, southeast asian, etc.)?*

RQ2: *Are there any significant differences in strip search rate for different age groups(such as aged 17 and below, aged 18 to 24, etc.) and sex?*

RQ3: *Are people with a higher frequency of arrests more prone to being strip searched, while controlling for perceived race and age group?*

RQ4: *Does the likelihood of items found at arrest vary by strip search, certain perceived race, sex and age group ?*

Those research questions will help us understand if there is any social bias that affects Ontario police officers on the decision to conduct search searches on arrestees, which might contribute to the further study of criminal justice.

Exploratory Data Analysis

According to the original dataset, we did some data cleaning to impute missing values and deleted the age group with the same meaning. We obtained a result of 52650 male, 12617 female and 9 arrestees with unisex. There are also 8 race groups (White, Black, East/Southeast Asian, South Asian, Middle-Eastern, Indigenous, Latino and unknown race) and 7 age groups for us to conduct the exploratory data analysis.

Table 1.1 Action_at_arrest_Copoperative by Perceived Race

Perceived_race	Count	Mean	Std
Black	17526.0	0.426	0.495
East/Southeast Asian	4415.0	0.514	0.500

Indigenous	1934.0	0.394	0.489
Latino	1768.0	0.507	0.500
Middle-Eastern	3237.0	0.468	0.499
South Asian	3613.0	0.461	0.499
Unknown or Legacy	5056.0	0.439	0.496
White	27723.0	0.444	0.497

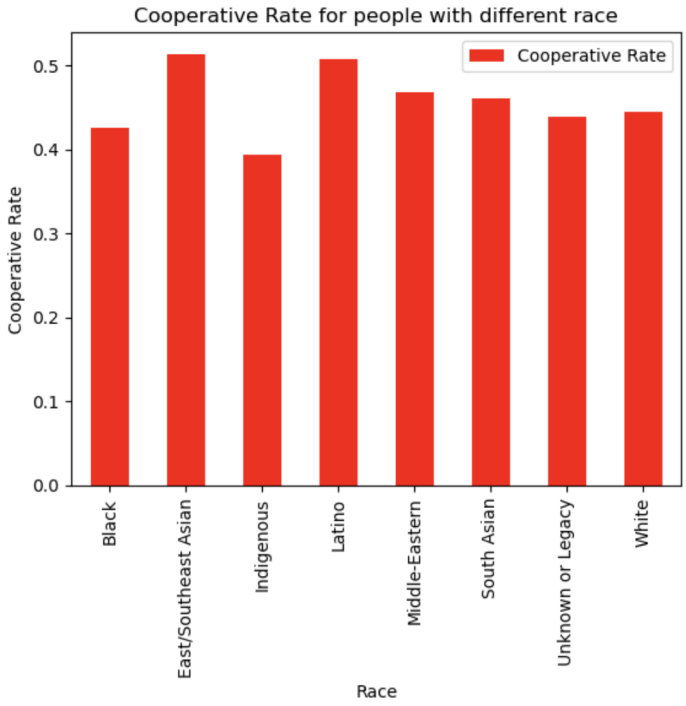
Table 1.2 Actions_at_arrest_Cooperative by Sex

Sex	Count	Mean	Std
Female	12617.0	0.430	0.495
Male	52650.0	0.450	0.497
Unisex	9.0	0.333	0.500

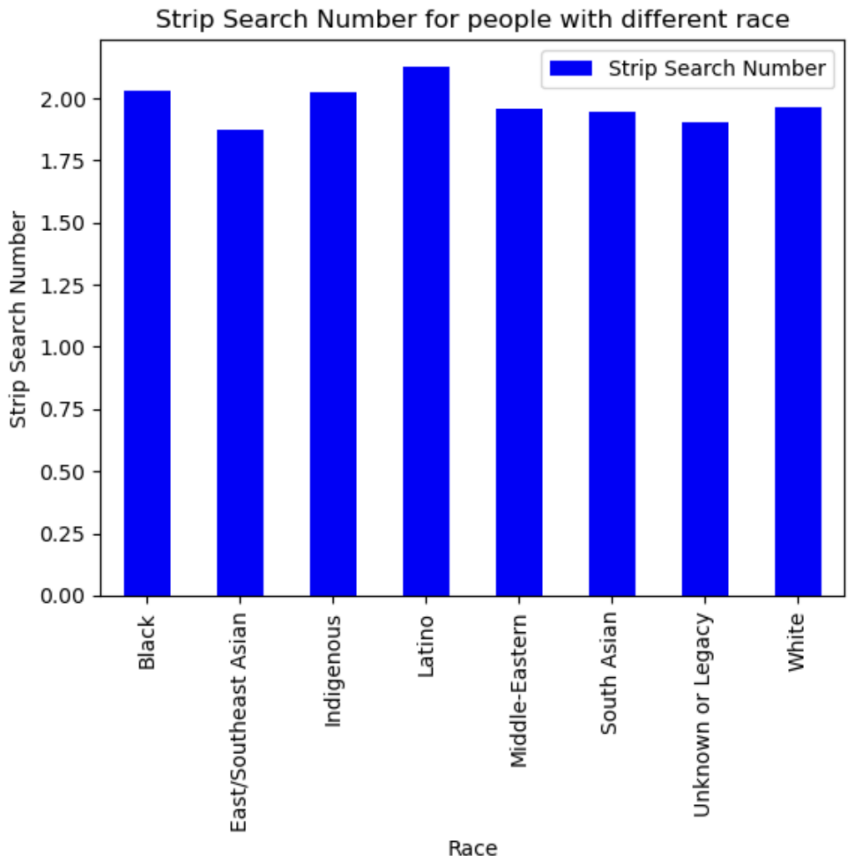
Table 1.3 Action_at_arrest_Cooperative by Age Group

Age_group_at_arrest	Count	Mean	Std
Aged 17 years and under	3042.0	0.464	0.499
Aged 18 to 24 years	10041.0	0.440	0.496
Aged 25 to 34 years	20949.0	0.436	0.496
Aged 35 to 44 years	16242.0	0.444	0.497
Aged 45 to 54 years	9066.0	0.4617	0.499
Aged 55 to 64 years	4590.0	0.455	0.498
Aged 65 and older	1322.0	0.488	0.500

Then we conducted a comparison on the cooperative action among individuals with different race, age and sex. We found the East/southeast Asian group has the highest cooperative rate in the race group, men have a slightly higher cooperative rate at arrest than women, and people who are 65 years older have higher cooperative rate at arrest.

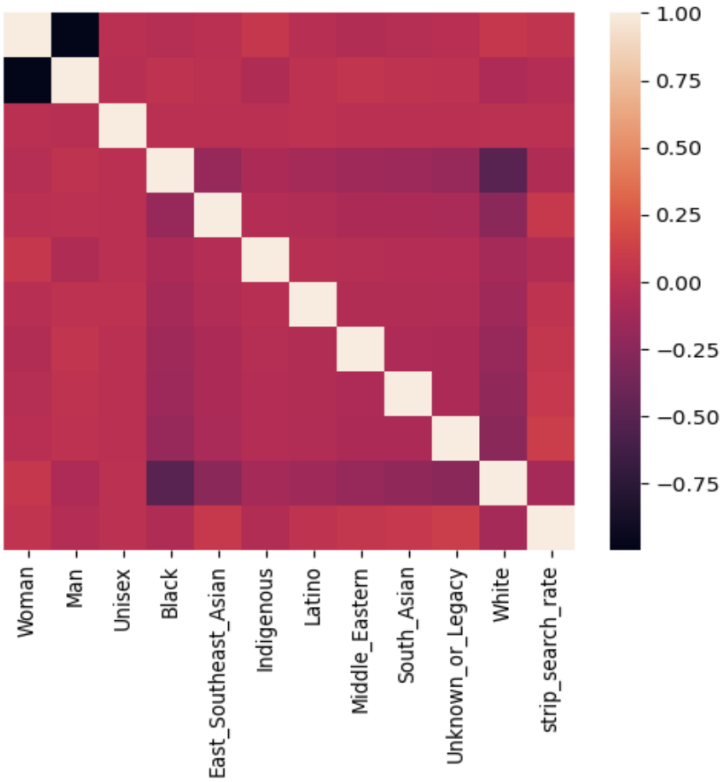


(Figure 1.0)



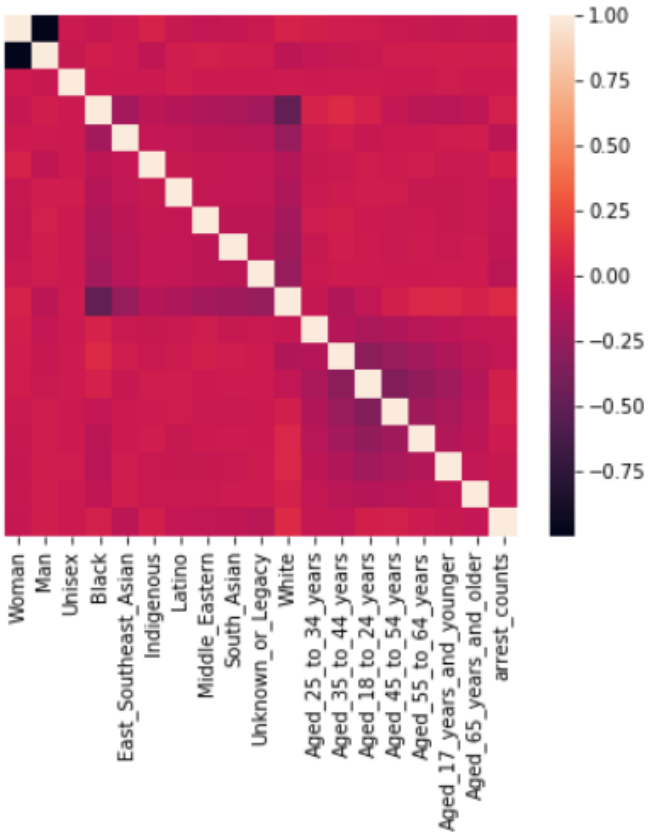
(Figure 1.1)

Then we found that some people were arrested several times, and we then decided to further explore the percentage of their strip search rate. We deducted the dataset and created a new dataframe containing each person only one time using a person ID. The strip search rate is conducted using times being arrested divided by times being strip searched. After data cleaning to remove the people with multiple sex or races, the final data frame has 34042 arrestees with the selected variables, including sex, race, arrest_counts, age_group_at_arrest, strip_search_count, strip_search_rate. We used the one-hot encoded method to set dummy variables for the sex and race category. Then used the concat method to combine these data frames together.



(Figure 1.3)

Figure 1.3 generated from the seaborn showed the correlation of different categorical attributes (individual variables from gender and race) and strip search rate. According to the result, we did not find strong correlation between these categorical attributes and the strip search rate.



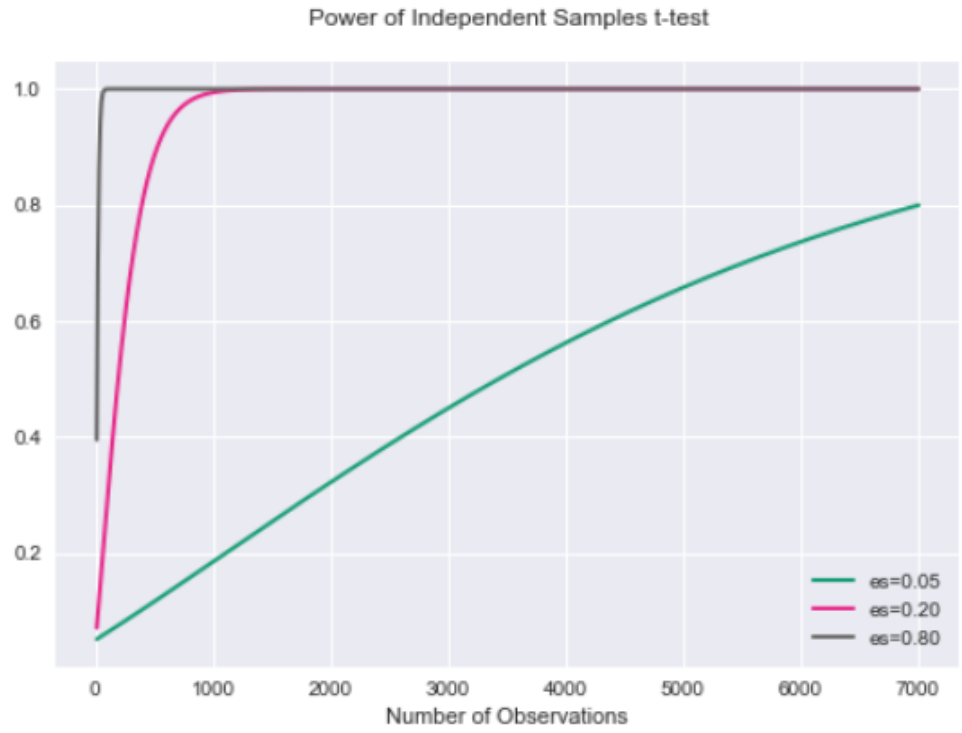
(Figure 1.4)

Figure 1.4 demonstrated the correlation between our predator variables (sex, race, age group and arrest counts). According to this heatmap, there is no strong correlation between attributes, which ensures our models avoid the issues with multicollinearity.

The data virtualization with bar charts and heat maps help us to understand the relationship between different categorical attributes and the strip search rate, as well as the correlation between different predators, which can be useful for us to develop our statistical models later on.

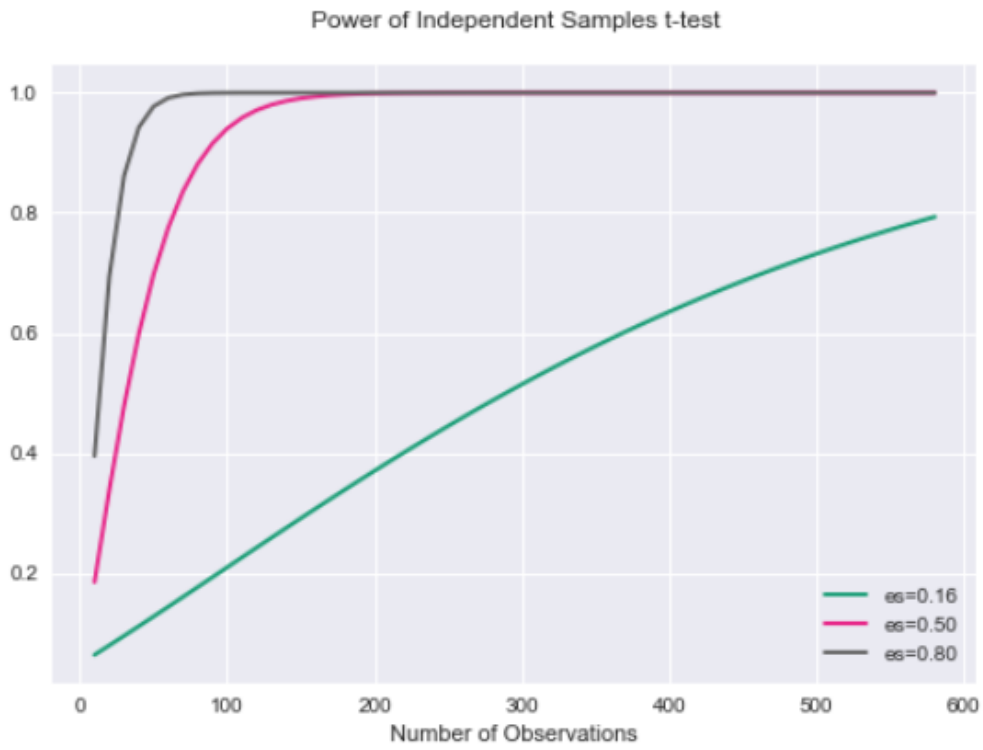
Power Analysis

Before we conduct t-test to analyze if there is significant difference in the strip search rate between different race groups, we calculated the effect size of the categorical variable using Cohen's D metric. As we are interested in finding whether the strip search rate differs between the Black and White race group, the effect size of the explanatory variable was 0.05. With this effect size, we calculated the sample size required for meeting the statistical power at 80% was 7016. With the selected power of 80%, the probability of Type II error is reduced to 20%, meaning that 80% chance that our t-test will correctly detect and reject the null hypothesis. The results indicated that a sample size of 4158 was needed for the White arrestee and size of 2859 was required for the Black arrestee. The following green curve on the power plot (Figure 1.5) illustrated the required sample size for our study. When the power is close to 0.8, the required sample size is 7016 arrestees at the significance level of 0.05.



(Figure 1.5)

Furthermore, we found the effect size of the other explanatory variable (Black vs Indigenous) was 0.16, which contributed to the analysis of whether the strip search rate differs between the Black and Indigenous group. The sample size needed for this study was 590. Among the total sample size, 559 Black arrestees and 32 Indigenous arrestees were required to conduct the test. As the following Figure 1.6 showed, the required sample size was 590 in order to obtain the power of 0.8 and significance level of 0.05.



(Figure 1.6)

Compared with the results from these two power analyses, our dataset has enough samples to conduct t-test which provided reliable results. With our current sample size, at least 80% of the statistical power was meeted.

T-test Result-Table 2.0

First IV	Second IV	DV	T statistics	P value
White (Race)	Black (Race)	Strip Search Rate	-1.27	0.2039
White (Race)	Indigenous (Race)		4.666	3.0981 e-6
White (Race)	East/southeast Asian (Race)		-17.44	1.468 e-67
Black (Race)	East/southeast Asian (Race)		-16.880	3.291 e-63
Black (Race)	Indigenous (Race)		5.177	2.301 e-07

To further explore the result of strip search rate on different race groups, we conducted the t-test by choosing different pairs of race groups to compare with. We tried to figure out if the strip search rates differ between the selected ethnic groups. The independent variables are different race groups shown on the table 1.0, which were randomly sampled and they were independent from each other. Since the size of each race group is large enough (>1000 for all chosen ethnic groups), we can be fairly tolerant with the data normality requirement as the sample mean will be approximately normal according to the Central Limit Theorem. The following results showed the relationship of strip search rate between two different races.

<u>White and Black</u>

H0 (Null Hypothesis): The strip search rate is the same for black people and white people.
H1 (Alternative Hypothesis): The strip search rate is different for black people and white people.

From the above result, we found that the T statistics is -1.27 . The p-value is 0.2039, which is larger than 0.05. Thus, we fail to reject the null hypothesis. We conclude there is no significant difference in strip search rate between black people and white people .

White and Indigenous

H0 (Null Hypothesis): The strip search rate is the same for Indigenous people and white people.
H1 (Alternative Hypothesis): The strip search rate is different for black people and white people.

From the above result, we found that the T statistics is 4.666 . The p-value is 3.0981 e-6, which is much smaller than 0.05. Thus, we can reject the null hypothesis. We conclude there is a significant difference in strip search rate between indigenous people and white people .

White and East/Southeast Asian

H0 (Null Hypothesis): The strip search rate is the same for east/southeast asian people and white people .
H1 (Alternative Hypothesis): The strip search rate is different for east/southeast asian people and white people .

From the above result, we found that the T statistics is -17.44 . The p-value is 1.468 e-67, which is much smaller than 0.05. Thus, we can reject the null hypothesis. We conclude there is a significant difference in strip search rate between east/southeast asian people and white people.

Black and East/Southeast Asian

H0 (Null Hypothesis): The strip search rate is the same for east/southeast asian people and black people .
H1 (Alternative Hypothesis): The strip search rate is different for east/southeast asian people and black people .

From the above result, we found that the T statistics is -16.880 . The p-value is 3.291 e-63, which is much smaller than 0.05. Thus, we can reject the null hypothesis. We conclude there is a significant difference in strip search rate between east/southeast asian people and black people.

Black and Indigenous

H0 (Null Hypothesis): The strip search rate is the same for Indigenous people and black people .

H1 (Alternative Hypothesis): The strip search rate is different for Indigenous asian people and black people .

From the above result, we found that the T statistics is 5.177 . The p-value is 2.301 e-07, which is much smaller than 0.05. Thus, we can reject the null hypothesis. We conclude there is a significant difference in strip search rate between Indigenous people and black people.

T-test Result-Table 2.1

First IV	Second IV	DV	T statistics	P value
Males	Females	Strip Search Rate	5.025	0.000
Aged 17 and younger	Aged 18 to 24		-0.238	0.812
Aged 18 to 24	Aged 25 to 34		-8.951	0.000
Aged 18 to 24	Aged 35 to 44		-10.676	0.000
Aged 18 to 24	Aged 45 to 54		-6.441	0.000
Aged 18 to 24	Aged 55 to 64		-1.989	0.046
Aged 18 to 24	Aged 65 and older		4.390	0.000

Similarly, we conducted the t-test to examine the difference of strip search rate on different gender and age groups. In this case, we selected the group aged 18 to 24 as the reference group to see if there are significant differences between age 18 to 24 and other age groups on the strip search rate, compared to the findings of literature review. According to table 2.1, there is a significant difference in strip search rates between males and females ($p<0.05$). In terms of the age, the p-value of 0.812 indicates that there is no significant difference in the strip search rate between arrestees aged 17 and younger and those aged 18 to 24. The rest of the t-test results indicated that there is significant difference between groups aged 18 to 24 and other age groups on the strip search rate.

The reason to perform t-tests is to compare the mean difference between the strip search rate between each race, sex and age group, so that it helps us to analyze the predator variables of the strip search models.

Research Design and Methods

This dataset was retrieved from the Public Safety Data Portal of Toronto Police Service, which consists of a total of 65276 samples of the arrest and strip searches records during the year of 2020 to 2021. As some of our research questions focused on analyzing the effect of arrestees' personal identity (race, gender and age groups) on the strip search result, we studied the relationship between strip searches rate and the attributes of perceived_race, sex, age_group_at_arrest and arrest count for each arrestee. Prior to the ANOVA and ANCOVA tests, we conducted the power analysis to determine the proper sample size for our studies using the desired power of 0.80. We also conducted several t-tests to analyze if there is significant difference between the means of two groups on the strip search rate. The t-tests were performed on the group comparisons including the difference between race groups, age groups and sex. To satisfy the requirement of ANOVA and ANCOVA test, we decided to create a continuous variable called strip search rate, which acted as the outcome variable in our study. Since the strip searches rate was not given in the original dataset, we calculated the strip searches rate for individual using the following formula:

Strip searches rate=Total count of strip search/Total count of arrest

As the dataset records all of the arrest events, which included the chance of one arrestee that might be arrested multiple times during the given time period. During the data preprocessing phase, we first calculated the total count of the arrest event of each arrestee via the unique personal ID. We found that the total number of arrestees is 37347. The sample was large enough for us to conduct the analysis. After the data preprocessing, we removed some of the misleading values which contained different labels for the same individual, then our sample size was reduced to 34042. Followed by calculating the total frequency of being strip searched, we used

the results of strip searches rate as the dependent variable for ANOVA and ANCOVA analysis. In addition, we calculated the total arrest number for each arrestee in this dataset, which is used as one of the predictors for our ANCOVA model.

RQ1: To conduct the analysis of this research question, we used a one-way ANOVA test. According to the purpose of the study, the independent variable is the perceived race, which is a categorical variable that consists of different levels. The relationships can be expressed as:

$$y(strip\ search\ rate) = \beta 0 + \beta 1(race) + e$$

RQ2: To analyze if the strip search rate is affected by age and gender, we conducted a two-way ANOVA test. In this case, the independent variables are sex (male and female) and different age groups with a total of 7 age levels, including aged 17 and below, aged 18 to24, aged 25 to 34, aged 35 to 44, aged 45 to 54, aged 55 to 64 and aged 65 and above. Both sex and age groups are categorical variables that consist of different levels, which are not dependent on one another.

$$y(strip\ search\ rate) = \beta 0 + \beta 1(sex) + \beta 2(age\ group) + e$$

We then set our hypothesis for each of the tests. Based on the result of one-way ANOVA and two-way ANOVA, interpretation can be conducted on how strip search rates differ depending on the age groups and sex of the arrestee.

The post-hoc test was conducted in conjunction with the ANOVA to identify which specific group means are significantly different from each other. In this case, we run the Tukey's Honestly Significant Difference (HSD) test to determine which pairwise comparisons among the race, sex and age group are responsible for the difference.

RQ3: To analyze if the result of strip search rate is affected by the arrest frequency, perceived race and age, we conduct the ANCOVA test. In this case, the independent variables race and age groups consist of different levels. In this model, the arrest frequency as the covariate ranges from 1 to 24. The reference group is Black arrestees aged 17 and below. The results compare the

impacts of different predators on search rate between the reference group and other groups. More importantly, it examines the impact of the number of arrests on the strip search frequency.

$$y(strip\ search\ rate) = \beta 0 + \beta 1(arrest\ count) + \beta 2(race) + \beta 3(age\ group) + e$$

RQ4: To study the effects of strip search rate, race, gender and age group on the likelihood of the item found during arrest, we performed a Multiple Logistic Regression. The item found is a binary (0 or 1) outcome in the dataset. Differ from the method for ANOVA and ANCOVA, we use the arrest event in the original dataset without grouping data by Personal ID.

$$logit(P(item\ found)) = \beta 0 + \beta 1(strip\ search) + \beta 2(race) + \beta 3(selected\ age\ group) + \beta 4(selected\ sex) + e$$

In this Multiple Logistic Regression, the *logit* represents the natural logarithm and the *P(item found)* stands for the likelihood of an item discovered at arrest. The predictors in this model are strip search outcome (binary 0 or 1), selected race (categorical variable), selected age group (categorical variable), selected sex (categorical variable) and the *e* represents the random error. Due to the high number of groups for our study subjects (race, age group and sex), we select those variables that are important to our study, using the reference from the previous research. In the sex group, we select male as the main study object. As we want to investigate if there are differences between race groupon the likelihood of items found at arrest, all races (except for the unknown race) are selected in this model. Similarly, we choose the age groups that are meaningful for our study regarding the results of previous t-tests and ANCOVA. We turn the categorical variables into dummy variables with 1 and 0 values, which represents True and False. To interpret the results more intuitively, we used the odds ratios (e^{β}) instead of the log odds. β1 to β4 are the coefficients of the predictors, which represent the average change in the log of odd ratio when the predictor changes by one unit, holding other factors constant. After fitting the Multiple Logistic Regression, we used 20% of the test set to test the accuracy of the model. If the model provides a prediction greater than 0.5, then it is categorized as positive (1), otherwise it is negative (0). We compared the predicted result with the actual outcome of the item found to check the accuracy. We also assessed the performance of the model by computing the confusion

matrix, which shows the number of true positive, true negative, false positive and false negative predictors. Lastly, a prediction interval plot was used to visualize the uncertainty in our model's predictions..

Results and Interpretation

With the preliminary exploratory data analysis, we conducted the ANOVA, ANCOVA and Logistic Regression test to explore the result of our research questions. We found that there are strong relationships between strip search rate and different personal identity of the arrestee. Specifically, race and age of the arrestee showed a significant impact on the strip search rate. The detailed interpretation was shown in the following section.

Impact of Race on Strip Search Rate

We tried to determine whether there was a statistically-significant difference between the mean strip search rate (dependent variable) for different races including White, Black, East/Southeast Asian, Indigenous, Latino, Middle-Eastern, South Asian and Unknown or Legacy (8-level categorical variable). The following hypotheses are tested:

H0 (Null Hypothesis): The population means among all eight different races are equal.
H1 (Alternative Hypothesis): The population means among all eight different races are not equal. At least one mean is different from the others.

Table 3.0 : Result of One-way ANOVA on the Effect of Race on Strip Search Rate

	Sum of Squares	Degree of freedom	F-statistic	P-value
Race	66.523	7	179.055	1.43 e-261
Residual	1806.347	34034	NaN	NaN

From the above result, we found that the F-value is 179.06 . The p-value is 1.43 e-261, which is much smaller than 0.05. Thus, we can reject the null hypothesis. We conclude there is a significant difference in strip search rate between different races.

Figure 3.1 Tukey multiple comparison of means for different race groups

Group1	Group2	Mean Difference	P-value
Indigenous	Black	-0.058	0.001
	East/Southeast Asian	-0.1423	0.001
	Latino	0.124	0.001
	Middle Eastern	0.135	0.001
	South Asian	0.147	0.001
	White	0.054	0.001
Black	South Asian	0.089	0.001
Black	East/Southeast Asian	0.085	0.001
South Asian	White	-0.093	0.001
East/Southeast Asian	White	-0.089	0.001
Middle Eastern	White	-0.081	0.001

From the post-hoc test results from Figure 2.1, we selected the groups with the most significant differences. Excluding the Unknown_or_Legacy group, we found that the Indigenous people have overall high significant differences around 0.1 ($p<0.01$) from other race groups, based on the significance level of 0.05. In addition, there are significant differences between Black and South Asian for the mean difference of 0.089, Black and East/Southeast Asian whereas the mean difference is 0.085. Moreover, there are significant differences between South Asian and White for the difference of -0.09, East/Southeast Asian and White for the difference of -0.089, and Middle Eastern and White for the difference of -0.08.

On the other hand, Some groups showed no significant difference, including Black and White, East/Southeast Asian and Latino, East/Southeast Asian and Middle Eastern, East/Southeast Asian and South Asian , and Latino and Middle Eastern. These groups do not differ significantly from each other at the 0.05 level of significance.

Impact of Age and Sex on Strip Search Rate

For the second research question, we tried to analyze whether there was a statistically-significant difference in strip search rate(outcome variable) by age groups (total of 7-level categorical variables) and by sex (male, female and unisex). The following hypotheses are tested:

<u>The main effect of sex on strip search rate</u>

H0 (Null Hypothesis): The mean of strip search rate between different sex are equal.

H1 (Alternative Hypothesis): The mean of strip search rate between different sex are unequal.

<u>The main effect of age group on strip search rate</u>

H0 (Null Hypothesis): The mean of strip search rate among all seven age groups are equal.

H1 (Alternative Hypothesis): The mean of strip search rate among all seven age groups are not equal. At least one mean is different from the others.

<u>The interaction effect between sex and age group on strip search rate</u>

H0 (Null Hypothesis): The mean of strip search rate on the interaction effect between sex and age group are equal.

H1 (Alternative Hypothesis): The mean of strip search rate on the interaction effect between sex and age group are not equal.

Table 3.2 : Result of Two-way ANOVA on the Effect of Age group and Sex on Strip Search Rate

Model	Sum of Squares	Degree of freedom	F-statistic	P-value
Sex	0.000	2	0.000	9.971e-0.1
Age group at arrest	15.23	6	46.444	5.487e-39
Sex & Age group at arrest (Interaction)	0.617	12	0.941	4.875e-0.1

According to the result above(Table 3.2), we discovered that the sex does not have a statistically significant impact on the strip search rate, as evidenced by the F-value very close to 0 and it has a very high p-value of 9.97 e-0.1. As the p-value is greater than 0.05, we reject the null hypothesis and conclude that sex does not have significant effect on the strip search rates. On the other hand, the result showed that the F-value for age group is 46.44, with a p-value of 5.49 e-39 (p-value<0.05). Therefore, we reject the null hypothesis. This result showed that age group has a significant effect on the strip search rate. Finally, the result suggested that there is no significant effect between sex and age group on the strip search rate, as it has the F-value of 0.94 and p-value for 4.88 e-0.1 (>0.05). We fail to reject the null hypothesis in this case and conclude that the interaction between sex and age does not have an impact on the stirp search rate.

Table 3.3 Tukey multiple comparison of means for different age groups

Group1	Group2	Mean Difference	P-value
Aged 35 to 44 years	Aged 65 years and older	0.069	0.001
Aged 25 to 34 years	Aged 65 years and older	0.059	0.001
Aged 45 to 54 years	Aged 65 years and older	0.058	0.001
Aged 18 to 24 years	Aged 35 to 44 years	-0.043	0.001
Aged 55 to 64 years	Aged 65 years and	0.037	0.001

	older		

According to the result, we selected the top 5 age groups that are significantly different from others. As the result indicated, there are significant differences between the group aged 35 to 44 and group aged 65 and above for the significant difference of 0.069, followed by aged 24 to 34 and aged 65 and above for the significant difference of 0.060. The third one is between the aged 45 to 54 and age 65 and above for the significant difference of 0.058. In addition, There are significant differences between aged 18 to 24 and aged 35 to 44(mean difference=-0.043), and between aged 55 to 64 and 65 and above(mean difference=0.037).

Table 3.4 Tukey multiple comparison of means for different sex (FWER.0.5)

Group 1	Group 2	Meandiff	P	Lower	Upper	Reject
Female	Male	-0.022	0.001	-0.029	-0.014	True
Female	Unisex	0.096	0.740	-0.2217	0.413	False
Male	Unisex	0.117	0.648	-0.200	0.435	False

The Tukey HSD test shows that male and female do have significant differences in strip search rate as the p-value is less than 0.05, then we have enough evidence to reject the null hypothesis that male and female have equally the same strip search rate. However, although the absolute value of mean difference between male and unisex, and between female and unisex is larger than the difference between male and female, the sample size is too small so we fail to reject the null hypothesis that male and female has the same strip search rate as unisex since the p-value is 0.740 and 0.648 which are both larger than 0.05.

Impact of Arrest Count, Race and Age on Strip Search Rate

For the third research question, we tried to include the impact of time of arrest to analyze whether there was a statistically-significant difference in strip search rate(outcome variable) by arrest count (the covariate), by age groups (total of 7-level categorical variables) and by race (total of 8-level categorical variables). The following hypotheses are tested:

The main effect of sex on strip search rate

H0 (Null Hypothesis): There is no difference in the strip search rate between the Black and other racial groups.

H1 (Alternative Hypothesis): There is a significant difference in the strip search rate between the Black and other racial groups.

The main effect of age group on strip search rate

H0 (Null Hypothesis): There is no difference in the strip search rate between the individuals aged 17 and below and other age groups.

H1 (Alternative Hypothesis): There is a significant difference in the strip search rate between the individuals aged 17 and below and other age groups.

The main effect of arrest count on strip search rate

H0 (Null Hypothesis): There is no relationship between arrest counts and the strip search rate.

H1 (Alternative Hypothesis): There is a significant relationship between arrest counts and the strip search rate.

Table 4.0: ANCOVA Results for Strip Search Rate

Variable	B	Std.Error	t	Sig. (P)	Lower Bound	Upper Bound
Intercept	1.053	0.004	272.795	0.000	1.045	1.060
Race: East_Southeast_ Asian	0.035	0.003	10.438	0.000	0.028	0.041
Race: Indigenous	-0.026	0.007	-3.743	0.000	-0.039	-0.012
Race: Latino	0.022	0.005	4.281	0.000	0.012	0.032
Race: Middle_Eastern	0.030	0.004	7.484	0.000	0.022	0.037

Race: South_Asian	0.036	0.004	9.797	0.000	0.029	0.043
Race: Unknown_or_Legacy	0.050	0.003	14.794	0.000	0.043	0.056
Race: White	0.006	0.002	2.856	0.004	0.002	0.010
Age: 18 to 24 years	0.007	0.004	1.650	0.099	-0.001	0.015
Age: 25 to 34 years	0.000	0.004	0.093	0.926	-0.007	0.008
Age: 35 to 44 years	-0.004	0.004	-0.925	0.355	-0.012	0.004
Age: 45 to 54 years	-0.005	0.004	-1.272	0.203	-0.014	0.003
Age: 55 to 64 years	0.002	0.005	0.335	0.737	-0.008	0.011
Age: 65 years and older	0.013	0.006	2.197	0.028	0.001	0.025
Arrest Counts	-0.122	0.001	-202.881	0.000	-0.123	-0.121

Table 4.1: ANCOVA Model Summary for Strip Search Rate

R	R Square	Adjusted R Square	F	Sig. (P)	No. of observation
0.566	0.566	0.566	3166	0.000	34042

Interpretation of ANCOVA Result

The result of R-Squared 0.566 indicates that 56.6% of the variation in strip search rate can be explained by the chosen predators: arrest count, race and age group. According to the F-statistic of 3166 and P value close to 0.00 ($p<0.05$), the result indicates that the overall model is statistically significant.

In this model, the reference group are Black arrestees aged 17 and below, and the intercept indicates that the expected strip search rate is 1.05. The rest of race categories including East/Southeast Asian, Latino, Middle Eastern, South Asian, Unknown or Legacy and White have a positive and significant difference ($p<0.05$) on the strip search rate compared to the Black reference group. On average, the strip search rate for the White is 0.01 units higher than for the Black, holding other variable constant. Compared to the Black, the average strip search rate for the East/ Southeast Asian, the Latino and the Middle Eastern are 0.03 units, 0.02 units and 0.03 units higher respectively. The coefficient of 0.05 indicates the strip search rate for the Unknown or Legacy classification is 0.05 units higher than for the Black. The result also indicates that there is a statistically significant difference between the Black and the Indigenous people as the p value is close to 0, whereas the coefficient is -0.03, which indicates the strip search rate of Indigenous people is estimated to be 0.03 units lower than the strip search rate for the Black individuals, after accounting for race and arrest counts.

Except for the age group 65 years and older, there is no significant difference between the reference group (aged 17 and below) and other age groups on the strip search rate as their $p>0.05$. For those aged 65 years and older, the strip search rate is 0.01 units higher than for those aged 17 and below, after controlling the race and arrest counts. The results showed that there is a significant difference between arrestees aged 17 and below and those aged 65 years and older.

Moreover, the result shows the relationship between arrest counts and strip search rate is statistically significant as the p value is close to 0. With one unit increase in arrest counts, the coefficient of -0.12 shows that strip search rate decreases by 0.12 units after adjusting the effect for race and the age group.

Table 5.0 Result of the Logistic Regression Predicting the Likelihood of Item Found at Arrest

Model	Log Odds				Odds Ratio		
	Coef	Std err	Z val.	p	exp(Est.)	2.5%	97.5%
(Intercept)	-23.004	422.657	-0.054	0.957	0.000	-851.397	805.389
Male	-0.168	0.061	-2.763	0.006	0.845	-0.287	-0.049
Black	0.197	0.100	1.977	0.048	1.218	0.002	0.392
East Southeast Asian	-0.108	0.138	-0.784	0.433	0.898	-0.378	0.162
Indigenous	0.101	0.158	0.641	0.521	1.107	-0.208	0.411
Latino	0.226	0.179	1.261	0.207	1.253	-0.125	0.576
Middle Eastern	0.178	0.141	1.260	0.208	1.195	-0.099	0.454
South Asian	0.029	0.142	0.201	0.841	1.029	-0.250	0.307
White	0.092	0.096	0.961	0.337	1.096	-0.096	0.280
Aged 17 years and younger	0.477	0.094	5.086	0.000	1.612	0.293	0.661
Aged 18 to 24 years	-0.235	0.052	-4.508	0.000	0.790	-0.337	-0.133
Aged 65 years and older	0.143	0.182	0.789	0.430	1.154	-0.213	0.500
Strip Search	22.553	422.657	0.053	0.957	6.232*10 ^9	-805.840	850.946

Interpretation of Logistic Regression

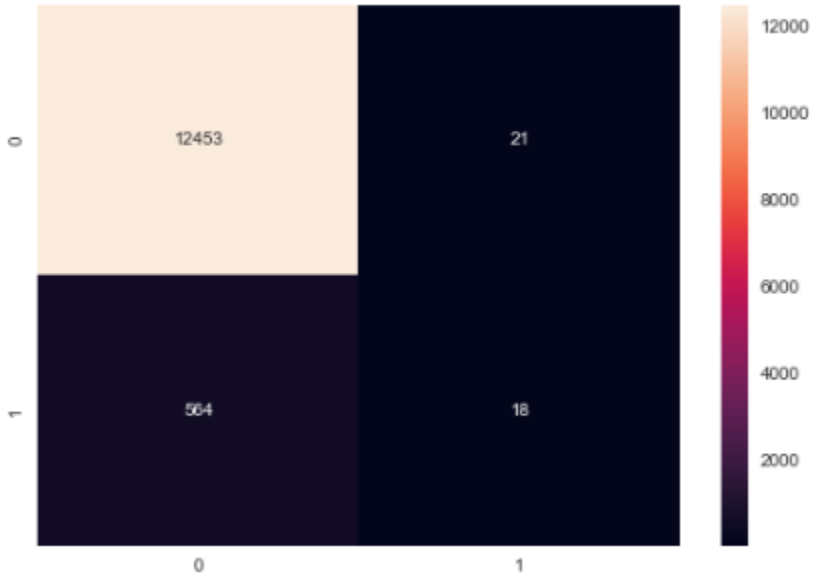
We perform a Logistic Regression to examine the effects of selected sex, race and age group on the likelihood of items found in the arrestees. In this model, the selected dependent variable is 'Male', 'Aged 17 and younger' , 'Aged 18 to 24 years', 'Aged 65 years and older', 'Strip search', and all the ethnicities ('East_Southeast_Asian', 'Indigenous', 'Latino', 'Middle_Eastern', 'Black', 'South_Asian', 'White'). The outcome variable is 'Items found'. According to the result of the R-squared for this Logistic Regression, this model shows a moderately good fit with a significant level of 0.05, which explains 57% of the variation in the strip search result. In this test, we have a total of 65276 observations, which represents the number of the arrest events. Among all the selected independent variables, four predictors are statistically significant with $p<0.05$, which includes male, the Black, the age groups of 17 years old and below and the age groups between 18 to 24 years. The remaining predictor variables in this model do not show a statistically significant likelihood of items found for p values greater than 0.05.

The intercept indicates that the likelihood of items found is close to 0 when all independent variables are 0. As the result indicated, an male arrestee has 0.85 times the odds of having an item found on them compared to other gender groups, holding all other variables constant. The coefficient for the male is statistically significant as its p-value is close to 0. Among all race groups, the Black arrestee has 1.22 times the odds of having items found on them compared to a non-Black arrestee, controlling for the other variables. In addition, the younger arrestee in the age groups below 27 demonstrates higher odds of having items found at arrest. An arrestee aged 17 years or younger has 1.61 times the odds of having an item found on them compared to those older than 17, holding all other variables constant. The arrestee aged 18 to 24 shows 0.79 times the odds of having an item found compared to those outside of this age range.

In terms of strip search conduct, the predicted odds ratio for reporting items found does not differ between arrestees being strip searched and those without. Other independent variables (East Southeast Asian, Indigenous, Latino, Middle Eastern, South Asian, White and Aged 65 years

and older) do not show statistically significant effects on the results of item founds with a p>0.05.

Confusion Matrix

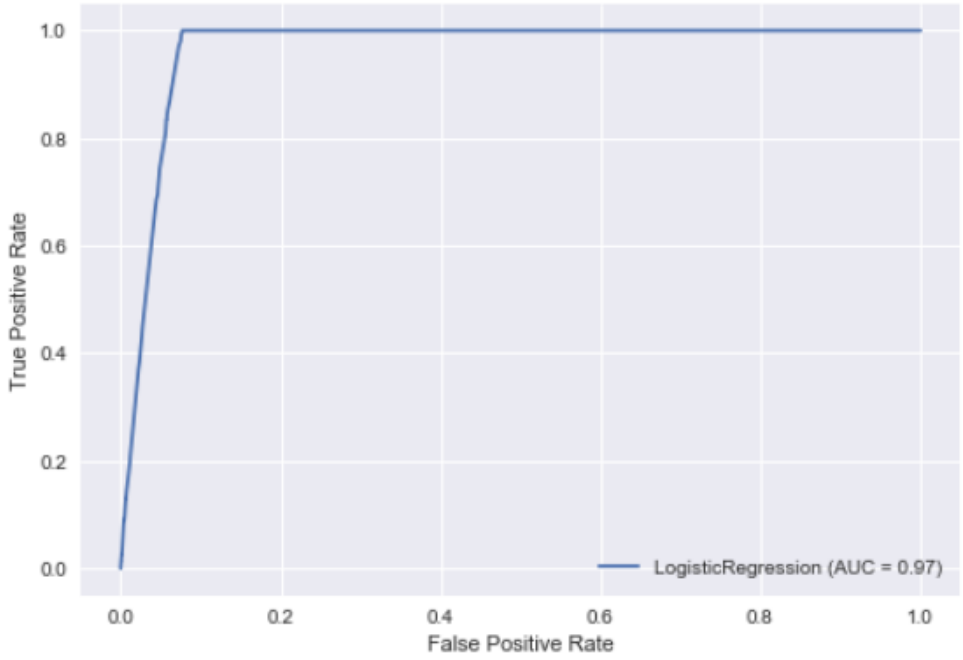


(Figure 2.0)

This confusion matrix (Figure 2.0) represents the performance of a binary classification model that predicted outcomes for a total of 13056 instances. According to Figure 2.0, there are 12453 instances indicating that the model correctly predicted that no item would be found on the arrestee. Regarding the true positive result, 18 times the model correctly predicted that an item would be found on the arrestee. On the other hand, there are 21 instances that the model indicates a Type I error, which predicts an item would be found on the arrestee when no item was found in reality. Lastly, 564 instances exhibit Type II error, which fails to predict the item found on arrestee.

With the calculation of the precision rate: True Positive/ (True Positive + False Positive)$= 18/(18+21) \approx 0.461$. The model correctly predicted the item would be found at 46.1% of the time. Also, the accuracy rate can be calculated with: (True Positive + True Negative) /total instances $= (18 + 12453)/(18+12453+21+564) \approx 0.956$. This result shows that the model is correct at

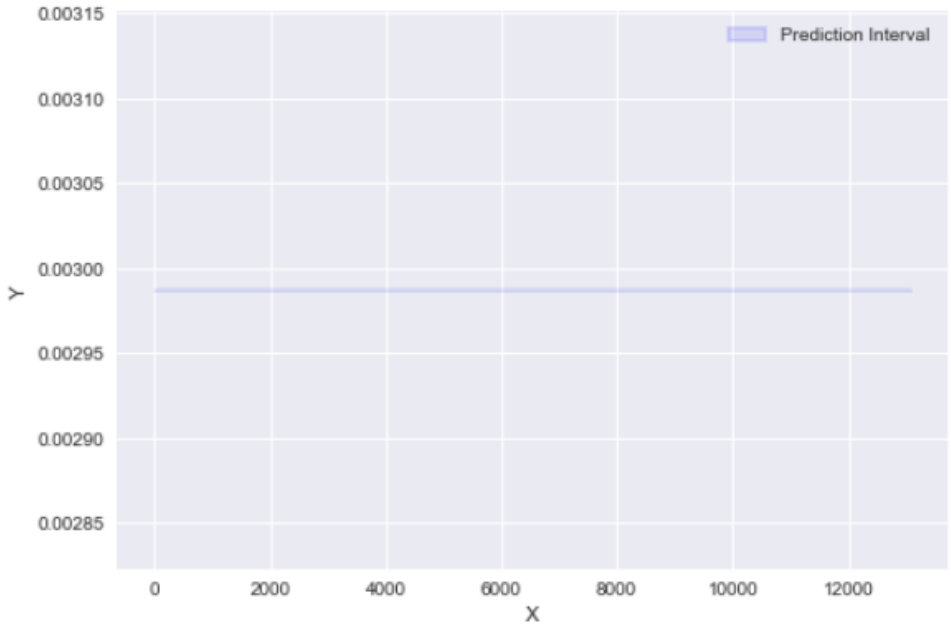
around 96% of the time. In addition, the sensitivity and specificity of the model is found at 3.1% and 99.8% , respectively. These results conclude that the model is good at identifying cases where no item is found on arrestee. However, it struggles to correctly identify cases where an item is found.



(Figure 2.1)

In addition, we also conducted the AUC curve to evaluate the performance of the model. The result of AUC of 0.97 indicates that the performance of our Logistic Regression model is fairly good.

Prediction Interval Plot



(Figure 2.2)

The plot of prediction interval (Figure 2.2) was generated to visualize the uncertainty in predictions of our Logistic Regression model. According to the plot, the observations fall at the interval between 0.00295 to 0.003. In this case, the prediction of item founds is close to 0.003 as shown in the plot, which indicates that the model's predicted probability of the items found is very low (0.3%). This might be due to the unproportional samples of the dataset as the items were only found at a very low instance at all arrest events.

Discussion

The findings of this project implicates some limitations. As we discover that there might be some discrepancy on the original dataset, as some arrestees who shared the same personal ID were recorded under different genders or different race groups for various arrest events. We suspect that there might be some errors for the record of this data. Some of the attributes contain contradictory meanings. For the action at arrest in the data, some arrestees are categorized as cooperative but also fall under other action lists, such as combative, violent or spitter, which causes confusions in our study. Also, there are large portions of unknown values, such as the unknown race, which might significantly affect the result of our study. Given the limitation of data types for the original dataset, there are not enough continuous variables to construct a stronger ANCOVA model. Most of the attributes are either categorical or binary variables, which imposed difficulty on conducting t-tests and ANCOVA tests. In this dataset, the predator variables that we studied did not have strong correlations with the strip search rate. With the current ANCOVA model, the predators represented 57% of the result of strip search rate. Additionally, the distribution of the arrest count is highly skewed, as most of the individuals had arrests fewer than five times which impacted the reliability of the model.

Furthermore, the findings on our Logistic Regression model indicates that there is no high correlation between independent variables and the predictor variables. When constructing a prediction interval, the binary data is either 0 or 1, which makes it hard to visualize any trend being predicted between an independent variable and dependent variable. The data overlaps itself on 0 and 1 which makes it hard to see how much data falls to a certain extent. The variables in the data also lack linearity of the relationship, which affects the predictions of the model. The unequal distribution of the number of instances for items found is significantly smaller than the result without, which leads to poor performance in predicting the positive class.

There are some improvements that could be made for our future studies. The accuracy of study on factors that contribute to the strip search result can be improved if more data is collected. For instance, data of whether the person is trying to hide anything before being arrested is useful to investigate if strip search is necessary compared to other forms of searching conduct. The job

classification of the arrestee could also be useful for this analysis. Instead of having age as a categorical value, age is better recorded as a continuous variable for further analysis. Having multiple continuous variables can help build linear relationships, to better utilize them as independent variables for logistic regression and covariate for ANCOVA test.

Conclusion

In general, our findings show that certain personal identity of the arrestee including race and age have significant impact on the strip search rate, whereas sex has little impact on the strip search rate for this dataset if we include the effect of unisex. The arrest count also has a significant impact on the strip search rate. However, there is a negative relationship between strip search rate and arrest count, which indicates whether the arrestee is a habitual offender does not have a significant impact on the police's decision to conduct a strip search. Lastly, the conduct of strip search does not have an impact on the likelihood of items found at arrest. Therefore, we conclude that strip search might not be an effective conduct for the Toronto Police Office to locate and remove the risk of items hidden at arrest.

We use quantitative methods to analyze the result of our research questions. The results of RQ1 showed that it's necessary to study the effect of race on strip search rate, as we discovered that there is a significant difference between the perceived race of the arrestee and the chance of being strip searched by the police officer.

In addition, RQ2 examined the impacts of other personal attributes on the effect of strip search rates. As the result shown, we discovered the age group also has a significant impact on the strip search rates. There is a significant difference between the strip search rate for male and females.These results are aligned with some of the findings from previous research and studies that race, age and sex has significant impacts on strip search rates.

The results of RQ3 indicated that even though the arrest count is associated with strip search rates, higher arrest frequency does not contribute to the higher probability of being strip

searched. This finding shows that there is very little association between the times of arrest and strip search rate.

Finally, the RQ4 proved that the items found at arrest are associated with the selected group. 1)Male has 0.845 times the odds of having an item found at arrest, holding all other variables unchanged. 2)The Black arrestee has 1.28 times the odds of having an item found on them compared with other race groups, holding other variables constant. 3)The arrestee aged 17 years or younger has 1.61 times the odds of having items found, controlling for other variables. 4) The arrestee aged 18 to 24 years has 0.79 times the odds of having items found on them, which presents a lower chance of items found compared to other groups. In addition, the strip search conduct did not contribute to the results of whether any objects were discovered at arrest.

These findings from our study can inform policy decisions and provide a deeper understanding of the factors influencing the outcomes of police search. More importantly, it will contribute to the social and criminal justices, which is crucial for the legislation authorities in Ontario to study the fairness and ethics on the appropriate use of strip search in the current criminal system.

References

Canadian Civil Liberties Association. (2022, June 20). *Strip searches in Ontario prisons*. CCLA. Retrieved February 24, 2023, from https://ccla.org/major-cases-and-reports/strip-searches-in-ontario-prisons/#:~:text=Ontario's%20law%20does%20the%20opposite,reasons%20to%20justify%20strip%20searches.

Cooper, R. F. (2013). We Are Always Already Imprisoned: Hyper-Incarceration and Black Male Identity Performance . In *Boston University Law Review* (Vol. 93, Ser. 1185, p. 1202). essay, Boston University School of Law.

Grewcock, M., & Sentas, V. (2022, September 16). *UNSW Report: Rethinking Strip Searches by NSW Police*. Analysis & Policy Observatory. Retrieved April 12, 2023, from https://rlc.org.au/resources/papers-and-submissions/unsw-report-rethinking-strip-searches-nsw-police

Ha, D. (2011). Blanket policies for strip searching pretrial detainees: an interdisciplinary argument for reasonableness. *Fordham Law Review, 79(6),* 2721-2760.

Lemke , M. (2022, June 6). *Strip searches are ineffective, unnecessary and target racialized Canadians*. The Conversation. Retrieved February 24, 2023, from https://theconversation.com/strip-searches-are-ineffective-unnecessary-and-target-racialized-canadians-185187

McNeilly, G. (2019, March). *Breaking the golden rule - a review of police strip searches in Ontario*. 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

MacGregor, D. L. (2003). Stripped of all reason the appropriate standard for evaluating strip searches of arrestees and pretrial detainees in correctional facilities. *Columbia Journal of Law and Social Problems, 36(2),* 163-208.

Newburn, T.,Shiner, M. & Hayman, S. (2004). Race, Crime and Injustice? *Strip Search and the Treatment of Suspects in Custody*, *The British Journal of Criminology*, 44(5), 677–694, Retrieved February 24,2021, from https://doi.org/10.1093/bjc/azh043

Psutka,M. & Sheehy. (2016). Strip-Searching of Women in Canada: Wrongs and Rights, 2016 94-2 *Canadian Bar Review* 241, 2016 CanLIIDocs 152, Retrieved February 24, 2023, from https://canlii.ca/t/74q,

R. v. Golden, 2001 SCC 83 (CanLII), [2001] 3 SCR 679, Retrieved February 24, 2023, from https://canlii.ca/t/51xm

Toronto Police Service. (2022, July 15). *2020 race and identity-based data collection findings*. 2020 RBDC Findings. Retrieved February 24, 2023, from https://www.tps.ca/race-identity-based-data-collection/2020-rbdc-findings/