University of Toronto

Final Assignment

By:

Xinxuan Lin, Peggy Pu

Faculty of Information
INF2178
Group 15
April 16, 2023

Table of Contents

1. Introduction **3**
1.1 Literature review 3
1.2 Research Questions and Objectives 4
2. EDA **5**
2.1 Descriptive Statistics 5
2.2 T-tests 11
2.2.1 Assumption Validation 12
2.2.1.1 Perceived Race and Percentage of Strip Searches 12
2.2.1.2 Sex and Percentage of Strip Searches 13
2.2.1.3 Youth status and Percentage of Strip Searches 13
2.2.1.4 Negative action and Percentage of Strip Searches 14
2.2.1.5 Age group and Percentage of Strip Searches 14
2.2.1.6 T-test Conclusion 15
3. Method **15**
3.1 Data Description 15
3.2 Research Methods and Assumptions 16
3.2.1 Power Analysis 17
3.3 Research Question 1 18
3.4 Research Question 2 19
4. Result **19**
4.1 Research Question 1 - Logistics Model 19
4.1.1 Assumption Validation 24
4.2 Research Question 2 - ANCOVA 25
4.2.1 Assumption check 26
5. Discussion **27**
6. Conclusion **29**
References **30**

1. Introduction

Strip searches are a controversial practice in law enforcement that involves the complete removal of clothing from an individual in order to search for contraband or other items (BSB Solicitors, 2022). These searches can be invasive and humiliating and have been the subject of much debate and criticism in recent years (Lemke, 2022). In particular, concerns have been raised about the disproportionate use of strip searches against certain demographics, including people of color, and those who are arrested at a young age (Lemke, 2022). Previous research, as seen in our assignment 1, suggests that demographic factors such as race, sex, and age group significantly affect the percentage of individuals subjected to strip searches. However, other factors such as negative actions at the time of arrest were not considered. Therefore, this report aims to explore the relationship between race, age group, negative action at arrest, and the average number of times individuals are subjected to strip searches at the time of the arrest. To achieve this goal, we will analyze the same dataset used in the previous study, a comprehensive dataset of arrest records to determine the incidence of strip searches across different race, age groups, and negative actions at the time of the arrest. Our analysis will uncover any discernible patterns or trends, allowing us to better understand the relationship between these demographic factors, actions at the time of arrest, and the probability of being strip searched and how these factors affect the average number of strip searches. We hope to shed light on the potential biases and disparities that exist within the criminal justice system with regard to strip searches and to identify areas where reform may be needed.

1.1 Literature review

Several studies have examined the relationship between sex, race, age group at the time of arrest, and the probability of being strip searched. Research indicates that females are less likely to be strip searched than males. A study by Prison Policy Initiative (2019) found that although females make up a growing portion of arrests, male arrestees were still more likely to be strip searched than female arrestees. However, it should be noted that the reasons for the difference in strip search rates between males and females are not clear.

Another finding of the study is that women of color are more likely to be stopped and searched by the police than white women. For example, Black women are 1.5 times more likely to be stopped and searched than white women, even though they are less likely to be carrying contraband (Prison Policy Initiative, 2019). This disparity in police stops and searches can have significant consequences for women of color, such as increased exposure to violence and

trauma. Similarly, A study by Susan Nembhard and Lily Robin (2021) found that Black and Hispanic individuals were more likely to be strip searched than White individuals.

Additionally, research indicates that the age group at the time of arrest may also play a role in the probability of being strip searched. Daniel M. Blonigen's article explores the relationship between personality traits and age-related crime trends. The research has consistently shown that crime rates decline as people age, and the decline is steeper for violent crimes than for nonviolent crimes (Blonigen, 2021). Weale's article notes that strip searches are a controversial and potentially traumatizing practice that should only be used as a last resort (Weale and Dodd, 2022). Critics argue that the disproportionate use of strip searches on certain demographics, such as children is indicative of systemic issues within law enforcement (Weale and Dodd, 2022).

Finally, for arrests related to weapons and homicide-related actions, white people were less likely to be strip-searched but had a higher rate of finding items when searched, while Black people were more likely to be strip-searched but had a lower rate of finding items. This suggests that strip searches involving Black people arrested for weapons offenses were less effective, and further exploration is needed to minimize potentially unnecessary searches while protecting safety (Toronto Police Service, 2022).

1.2 Research Questions and Objectives

Our study aims to examine the relationship between individuals' demographics, actions at the time of arrest, and the average number of being subjected to strip searches. Our research questions are based on a thorough literature review and a preliminary analysis of the dataset, which is presented in the Descriptive Statistics and T-tests section below. Specifically, we will investigate the following research questions:

> RQ1: How does individuals' demographics and action at arrest affect their likelihood of being strip searched?
> RQ2: How does the proportions of being strip searched differ by individual's demographics, such as youth status, and the number of negative action at arrest

By answering these research questions, we hope to provide a deeper understanding of the factors that influence strip searches in our dataset. We believe that exploring the factors, such as sex, race, age, and action at arrest will shed greater light on the underlying patterns and disparities in the use of strip searches in law enforcement.

2. EDA

2.1 Descriptive Statistics

The dataset comprises 34,043 observations, with 7,114 females and 26,929 males. A breakdown of the number of people being arrested by gender and race during 2020 and 2021 is presented in Figure 1. Across all racial categories, males have a higher number of arrests than females, which could be attributed to their larger representation in the dataset. In terms of race, the majority of individuals arrested are white and black, followed by unknown or legacy race, East/Southeast Asian, South Asian, Middle-Eastern, Latino /indigenous people. Notably, Latino/indigenous males have the lowest number of arrests among all the racial groups, which could be due to the distribution of the population across different racial groups in the Greater Toronto Area.

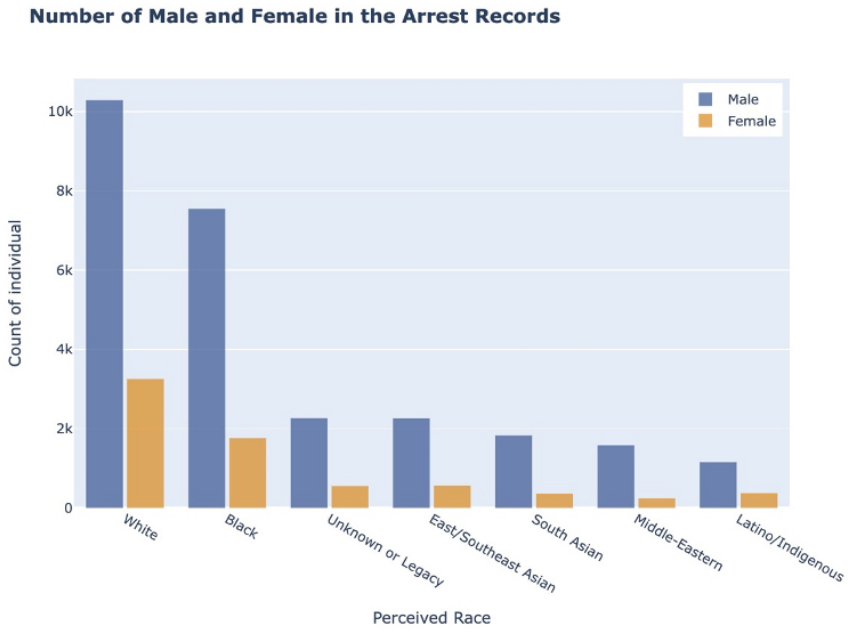

Figure 1: Distribution of males and females being arrested in each race group.

In Figure 2, the overall number of arrests for various racial and gender groups is presented. As with Figure 1, white and black individuals have the highest number of arrests, while the arrests for Latino/indigenous people are comparably lower.

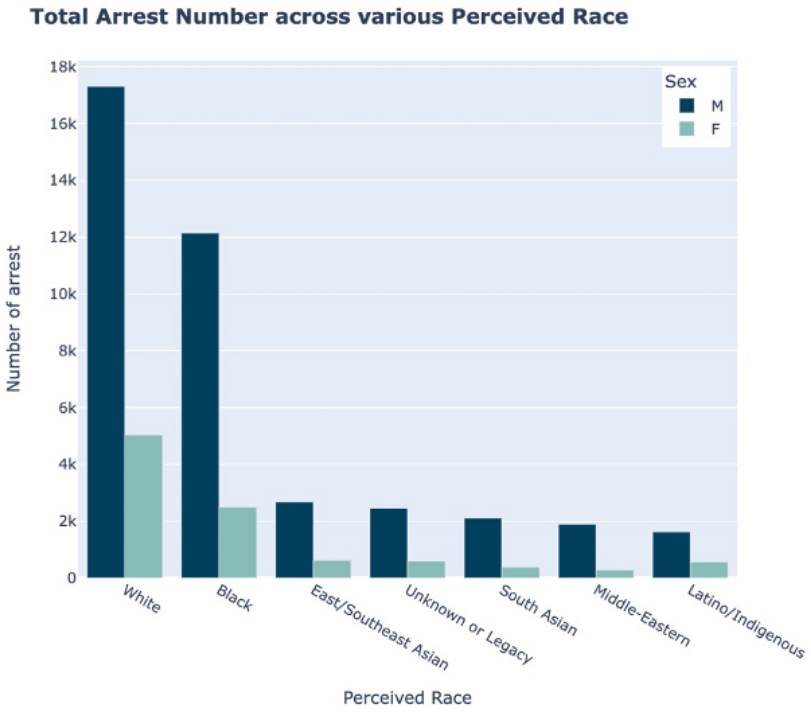

Figure 2: Number of total arrests for different sex groups across the races

There are 30,260 entries in the dataset for people who were not strip searched (with a strip search indication of 0) and 3,783 entries for those who were (with a strip search indication of 1), making the dataset significantly imbalanced. Figure 3 depicts the distribution of strip searched and not searched cases across racial groups, which follows the same trend as Figure 2. White individuals have the most non-strip search cases, while Latino/indigenous individuals have the least. However, this may be due to the uneven sample sizes between racial groups.

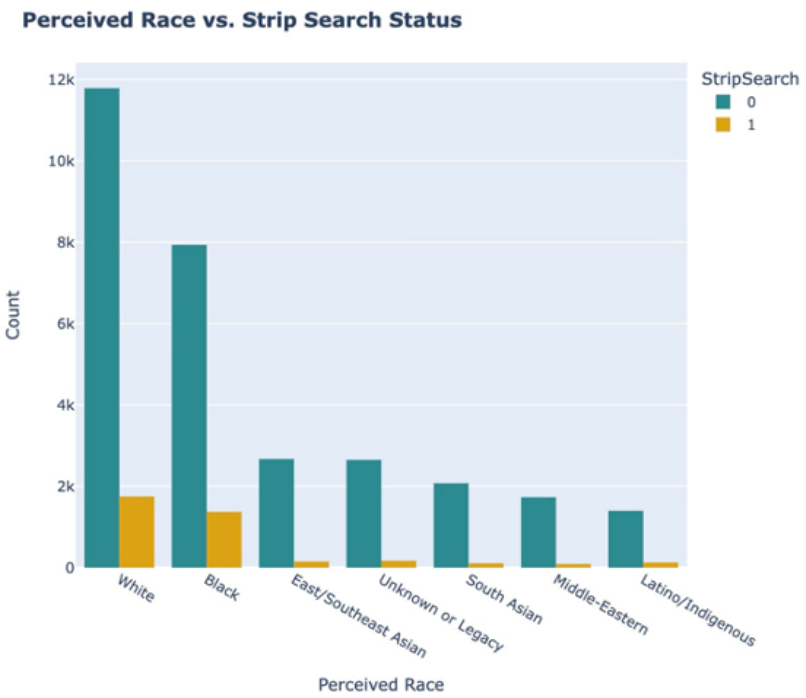

Figure 3: Count of cases for being strip searched and not being strip searched across the races.

Figure 4 displays the percentage of individuals who were strip searched and those who were not, for each racial group, while considering the sample size. This chart shows a different trend from Figure 3, with black individuals having the highest proportion of being searched, followed by white individuals, Latino/indigenous individuals, and the remaining four groups. Among the Middle Eastern, South Asian, East/Southeast Asian, and Unknown or legacy groups, the proportion of being searched is relatively lower and similar.

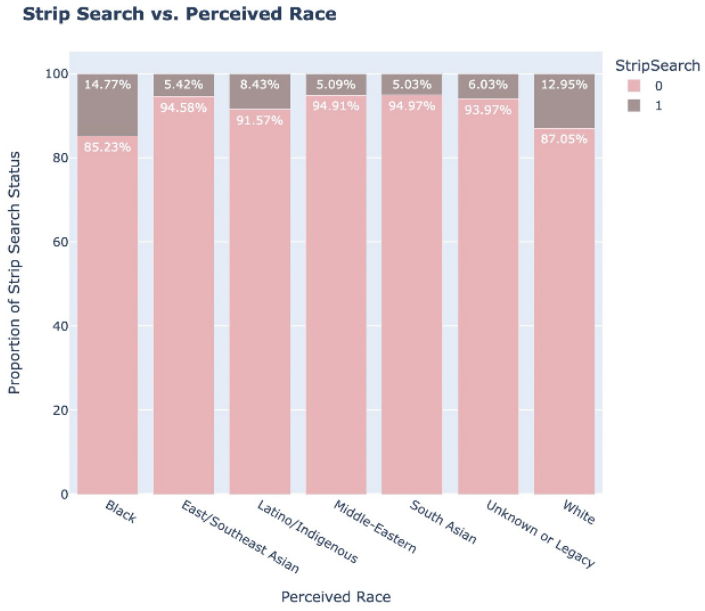

Figure 4: Distribution of strip searched and not searched cases across racial groups in percentage.

Figure 5 depicts the frequency of individuals who were strip searched and those who were not, across different age groups. The chart indicates a dramatic difference in the count of searched cases between groups. The age group of 25 to 34 years has the highest number of searched cases, followed by those aged 35 to 44, 18 to 24, and 45 to 54, and the other three groups. The senior group, aged 65 years and older, has the lowest number of searched cases.

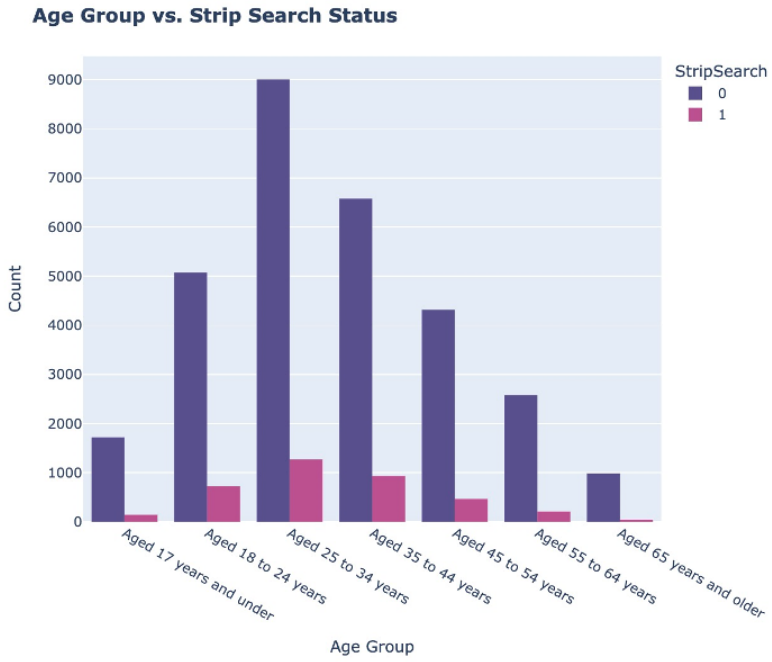

Figure 5：Count for being strip searched and not searched for different age groups

Figure 6 displays the percentage of individuals who were strip searched and those who were not, categorized by sex. The proportion of strip searches for males and females is relatively similar, suggesting that an individual's gender identity may not play a significant role in determining whether or not they are subjected to a search when arrested.

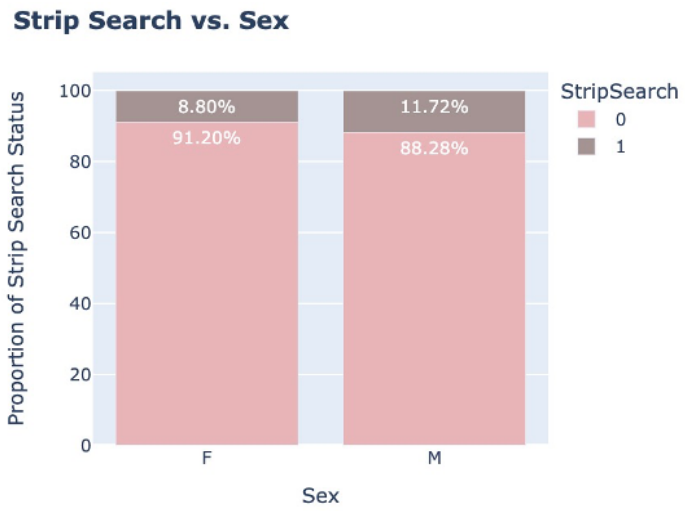

Figure 6：Proportion of each strip search status among different sex.

As previously noted, the distribution of data for strip searches and non-searches is highly unbalanced. In the case of the non-search group, the large number of observations with no negative actions at arrest results in the interquartile range and whiskers being compressed in Figure 7. The data for individuals who have undergone strip searches is right-skewed, with the number of negative actions greater than 3 being considered outliers. Both groups have a median of 0, but the mean number of negative actions for individuals who have been searched is higher than the other group. Additionally, the data for individuals who have been searched is more scattered and dispersed than the other group.

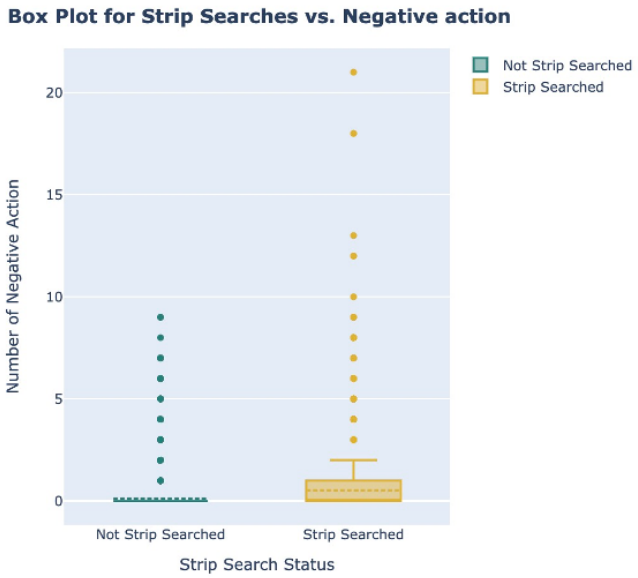

Figure 7: Boxplot for number of negative actions in different strip-searched categories.

Figure 8 shows that a large proportion of individuals in both youth categories had zero strip searches, causing the interquartile range and whiskers to be compressed. The median number of strip searches for both groups is zero. However, our analysis found that the mean percentage of individuals subjected to strip searches at the time of arrest was 6.06% for the youth category, which is lower than the other age group.

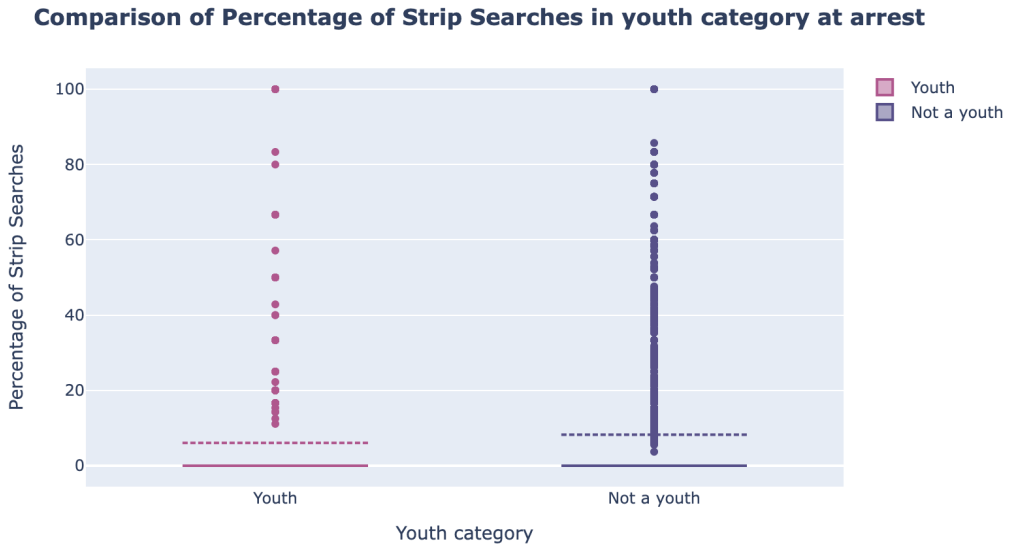

Figure 8: Boxplot to show the statistics of the percentage of strip searches in the youth category at arrest

Figure 9 illustrates the percentage of strip searches conducted for various age groups. Those aged 17 years and younger have the highest percentage of strip searches at 13.78%, while those aged 65 years and older have the lowest percentage of strip searches at 4.98%. This finding is consistent with previous research indicating that crime rates generally decrease as people age.

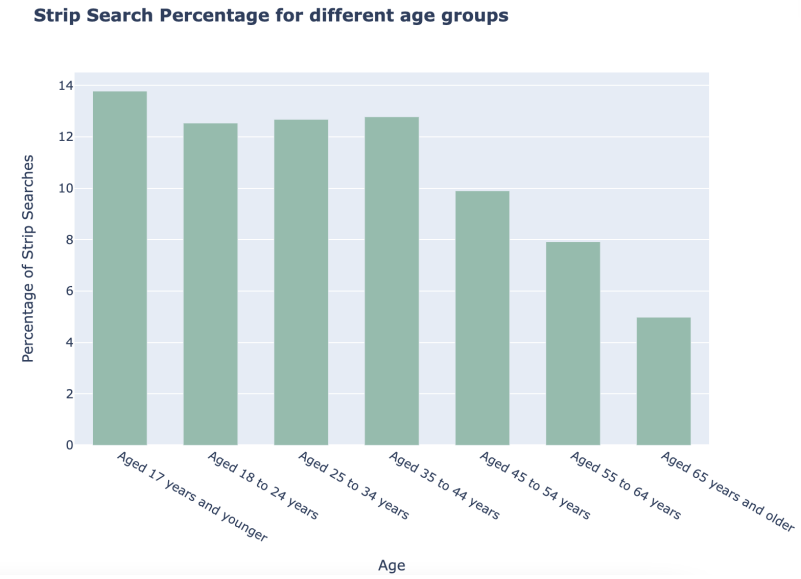

Figure 9: Barplot to show the statistics of the percentage of strip searches in different age groups at arrest

Figure 10 displays the correlation between the percentage of strip searches and the number of negative actions. The plot indicates a minor upward trend between the number of negative actions and the percentage of strip searches. In other words, as the number of negative actions increases, the percentage of strip searches also increases slightly. Furthermore, the percentage of strip searches for non-youth individuals shows a steeper increase than that of youth individuals. This implies that, for the same number of negative actions, non-youth individuals have a higher percentage of strip searches than youth individuals. However, this finding should be interpreted with caution due to the limited data points for the youth groups.

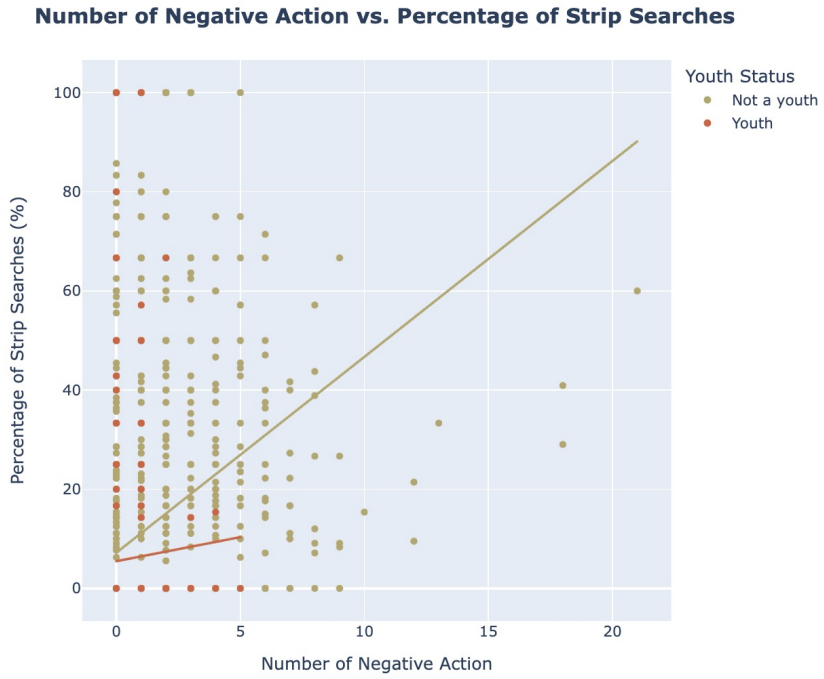

Figure 10: Scatterplot for number of negative action vs. percentage of strip searches brown down by youth status

2.2 T-tests

We conducted a Welch's t-test with independent variables to examine if there were any significant differences in the mean percentage of strip searches among different groups, considering the imbalance in sample sizes. The categorical variables tested were perceived race, sex, youth status, and age group. In the case of the continuous variable - negative action, it was divided into two groups: zero and non-zero.

-

The assumptions for Welch’s t-test are: (1) Independence errors; (2) continuous outcome variable; (3) categorical covariate with two levels; (4) outcome data for each group are approximately normality distributed. Since Welch’s t-test is conducted, there is no assumption for residuals to have constant variance.

2.2.1 Assumption Validation

The assumption of independence of errors is satisfied since each row in the dataset corresponds to a single individual, ensuring that there are no duplicate rows representing the same individual with different characteristics such as perceived race, sex, youth status, and age group.

The percentage of strip searches is a continuous variable representing the average number of strip searches among their total arrest. Perceived race is categorized as white or not white, while sex is classified as male or female. Youth status is determined as either youth or not a youth, and age group is divided into below 34 or above 35 years old. Negative action is classified as either no negative action or has negative action. All five covariates are two-level categorical variables.

To assess the normality assumption, we conducted a Shapiro-Wilk test for each variable, and the results are presented in Table 2.

Table 2: Shapiro-Wilk test result for all covariates

Variable	Level	P-value
Perceived Race	White	< 0.01
	Not white	< 0.01
Sex	Male	< 0.01
	Female	< 0.01
Negative Action	Has negative action	< 0.01
	No negative action	< 0.01
Youth Status	Youth	< 0.01
	Not a youth	< 0.01
Age Group	Below 34 years old	< 0.01
	Above 35 years old	< 0.01

The p-values for all groups are lower than the significance level (α=0.05), indicating strong evidence to reject the null hypothesis that the outcome data for each group follows a normal distribution. Therefore, we can conclude that the normality assumption is not met.

However, one of the reasons that may have caused this issue is that the Shapiro-Wilk test is more suitable for sample sizes smaller than 5,000, whereas the sample size in each group is greater than that, which may have resulted in inaccurate p-values.

2.2.1.1 Perceived Race and Percentage of Strip Searches

We observed a difference in the pattern of the percentage of strip searches across perceived race groups compared to the trend in total arrests. Therefore, we used Welch's t-test to

determine whether there is a significant difference in the percentage of strip searches between white and non-white people.

The hypothesis of the t-test are:

$Null\ Hypothesis\ (H_0)$: The population means of two independent levels: white people and non-white people, are equal.

$Alternative\ Hypothesis\ (H_a)$: The population means of two independent levels: white people and non-white people, are not equal.

The results of the Welch's t-test indicate that the mean percentage of strip searches for white people (M=8.28, SD=24.45) is higher than the mean percentage of strip searches for non-white people (M=7.34, SD=24.20). The p-value of 0.00035 is less than the significance level (α=0.05) with a 95% confidence interval of [0.41, 1.47]. Therefore, we have strong evidence to reject the null hypothesis and conclude that the mean percentage of strip searches differs between white and non-white people.

2.2.1.2 Sex and Percentage of Strip Searches

A Welch's t-test was performed to investigate whether there is a difference in the percentage of strip searches between male and female individuals. The hypotheses for the test are:

$Null\ Hypothesis\ (H_0)$: The population means of two independent levels: male and female, are equal.

$Alternative\ Hypothesis\ (H_a)$: The population means of two independent levels: male and female, are not equal.

The results of Welch's t-test indicate that the mean percentage of strip searches for male individuals (M=8.10, SD=24.80) is greater than that for female individuals (M=6.28, SD=22.30). The p-value (<0.00000) is less than the significant level ($\alpha = 0.05$), with 95% CI [1.22, 2.41]. Therefore, we have sufficient evidence to reject the null hypothesis that the mean percentage of strip searches is the same for male and female individuals.

2.2.1.3 Youth status and Percentage of Strip Searches

A Welch's t-test was performed to determine whether there is a difference in the percentage of strip searches between individuals who are classified as youth (below 17 years old) and those who are not. The hypothesis tested are:

$Null\ Hypothesis\ \left(H_0\right)$: The population means of two independent levels: youth and not a youth, are equal.

$Alternative\ Hypothesis\ (H_a)$: The population means of two independent levels: youth and not a youth, are not equal.

The results of Welch's t-test indicate that the mean percentage of strip searches for non-youth individuals (M=7.84 SD=24.48) is higher than the mean percentage of strip searches for youth (M=5.55, SD=21.07). The p-value (0.00001) is less than the significant level ($\alpha = 0.05$), with a 95% CI [-3.29, -1.30]. Therefore, there is strong evidence to reject the null hypothesis (H_0) that the mean percentage of strip searches is the same between youth and non-youth people.

2.2.1.4 Negative action and Percentage of Strip Searches

A Welch's t-test was carried out to determine if there is a difference in the percentage of strip searches between people who had a negative action at the time of arrest and people who did not have a negative action. The hypotheses are:

$Null\ Hypothesis\ \left(H_0\right)$: The population means of two independent levels: has negative action and no negative action, are equal.

$Alternative\ Hypothesis\ (H_a)$: The population means of two independent levels: has negative action and no negative action, are not equal.

The results of Welch's t-test show that the mean percentage of strip searches for people who had negative action (M=15.58, SD=23.47) is higher than the mean percentage of strip searches for people who did not have negative action (M=6.86, SD=29.80). The p-value (<0.00000) is less than the significant level ($\alpha = 0.05$), with a 95% CI [7.68, 9.76]. Thus, we have evidence to reject the null hypothesis that the mean percentage of strip searches is the same between people who had negative action and people who did not have negative action at the time of arrest.

2.2.1.5 Age group and Percentage of Strip Searches

A Welch's t-test was performed to investigate whether the proportion of strip searches varied between individuals who were below the age of 35 at the time of arrest and those who were older than 35. The hypothesis is:

H_0: The population means of two independent levels: under 35 and above 35 years old, are equal.

H_a: The population means of two independent levels: under 35 and above 35 years old, are not equal.

The results of Welch's t-test show that the mean percentage of strip searches for people under 35 years old (M=8.64, SD=25.85) is greater than the mean percentage of strip searches for people over 35 years old (M=6.69, SD=22.42). The p-value (<0.00000) is less than the significant level ($\alpha = 0.05$), with a 95% CI [1.44, 2.46]. Therefore, we have evidence to reject the null hypothesis that the mean percentage of strip searches is the same between people who are under 35 years old and those who are older than 35.

2.2.1.6 T-test Conclusion

According to Welch's t-test findings presented above, there were statistically significant differences in the mean percentage of strip searches for different categories of perceived race, sex, negative action, youth status, and age groups.

3. Method

3.1 Data Description

The dataset used for this project was provided by the Toronto Police Service (Toronto Police Service, 2022). The dataset has 34,043 observations that are grouped by person ID, which is assumed to be the identification number assigned to those who have been arrested. It includes demographic information on each individual, such as their perceived race, sex, age group at the time of arrest, and youth status at the time of arrest. These attributes are categorical variables with at least two levels.

Perceived race was treated as a 7-level categorical variable (East/Southeast Asian, White, Black, South Asian, Middle-Eastern, Unknown or Legacy, Latino/Indigenous) to examine the impact of each race on the outcome.

Age groups were categorized into 7-level variables: Aged 17 years and under, Aged 18 to 24 years, Aged 25 to 34 years, Aged 35 to 44 years, Aged 45 to 54 years, Aged 55 to 64 years, Aged 65 years and older.

Additionally, the dataset includes information about the arrest, including the indication of strip searches (0 for not being searched, and 1 for being searched), the number of strip searches, the strip search percentage (i.e., the average number of strip searches per person during 2020 and

2021), and number of any negative actions taken at the time of arrest, such as concealed items, violent behaviour, defensive or escape risk, and mental health issues. The variable *Strip Search Cont* counts the number of times each person was searched at the time of arrest without considering the total number of arrests, while strip search percentage represents the average number of being searched per person during the two-year period. Strip Search Status indicates whether or not the individual is being searched at the time of arrest. Negative action counts the number of inappropriate actions taken at the time of arrest for each person. To avoid any potential for misleading information, cases involving individuals with more than one perceived race and age group were removed from the dataset.

3.2 Research Methods and Assumptions

Based on the analysis of Figures 3 to 7 in the exploratory data analysis section, it was found that the proportion of being strip searched varies across different age groups, perceived race, and negative action. Figure 6 shows that there is no dramatic difference in the proportion of being strip searched between males and females, therefore sex would not be considered as a variable in the research. The research would instead focus on examining how other variables, such as an individual's demographics (e.g., race and age group) and number of negative actions at the time of arrest, impact the probability of being strip searched. To achieve this, a logistic model was fitted to determine the relationship between perceived race, age groups, and the number of negative actions with the probability of being strip searched. White was the reference group for perceived race, and aged 25 to 34 years was the reference group for the age groups.

The dataset was separated randomly into a training set and a testing set, with 80% of the data in the training set and 20% in the testing set. Due to the significant imbalance in the strip search data (with a much larger amount of data for cases where individuals were not strip searched), a stratified sampling technique was implemented to ensure that the proportions of being strip searched were similar between the training and test sets, as seen in the original dataset. Model accuracy and confusion matrix were calculated to evaluate the model.

To use logistic regression mode, the following assumptions were satisfied: (1) binary outcome; (2) independence of observations; (3) large sample size; (4) no multicollinearity between variables; (5) continuous independent variables are linearly related to the log odds of the outcome. The multicollinearity assumption and the linearity between continuous predictor and log odds of outcome were validated using correlation heatmap and regression plot, respectively.

The research would also explore how these variables impact the number of strip searches concerning the total number of arrests, which is the percentage of strip searches. Section 2.2.1.6 concluded that there were significant differences in the mean percentage of strip searches across various levels of youth status, negative action, age groups, race, and sex. However, since youth status and age groups both relate to age, only youth status will be considered in the study to determine variables that affect the percentage of strip searches to avoid collinearity between variables.

Previous research (midterm) has already explored the impact of perceived race and sex on the percentage of strip searches. Thus, the current study will focus on determining the relationship between youth status, number of negative actions, and the percentage of strip searches, which would be investigated by conducting a one-way ANCOVA.

To use ANCOVA, the following assumptions were satisfied: (1) continuous outcome; (2) categorical explanatory variables and continuous covariate; (3) independent errors; (4) normality assumption; and (5) constant variance; (6) linearity between the covariate and the dependent variable; (7) no interaction between the covariate and the independent variable. The same reason explained in the t-test section supported the first two assumptions. However, for performing a one-way ANCOVA, the number of negative actions an individual has at arrest is treated as a continuous variable, instead of a categorical variable. Thus, assumption (3) is also satisfied. The constant variance and normality of residuals assumption were validated using Levene's test and Shapiro-Wilk test, respectively. Assumption (6) was validated using scatterplot, and the last assumption was validated based on the results of the linear model with an interaction term.

A power analysis was performed to determine the sample size required to detect the statistically significant effect with a significance level of 0.05 and a power of 0.8.

3.2.1 Power Analysis

Before conducting a t-test to examine the difference in the percentage of strip searches (outcome variable) between two groups based on youth status (youth and non-youth), the effect size of the explanatory variable was measured usingCohen's D metric, which was 0.09.

After determining the effect size with Cohen's D metric, the required sample size was computed at 80% statistical power. The findings showed that the non-youth group would need a sample size of 16,081, while the youth group would require a sample size of 934. The dataset provided sample sizes of 32,176 and 18,867 for the non-youth and youth groups, respectively, both of

which exceeded the necessary sample sizes. Thus, the results obtained from the dataset were deemed reliable.

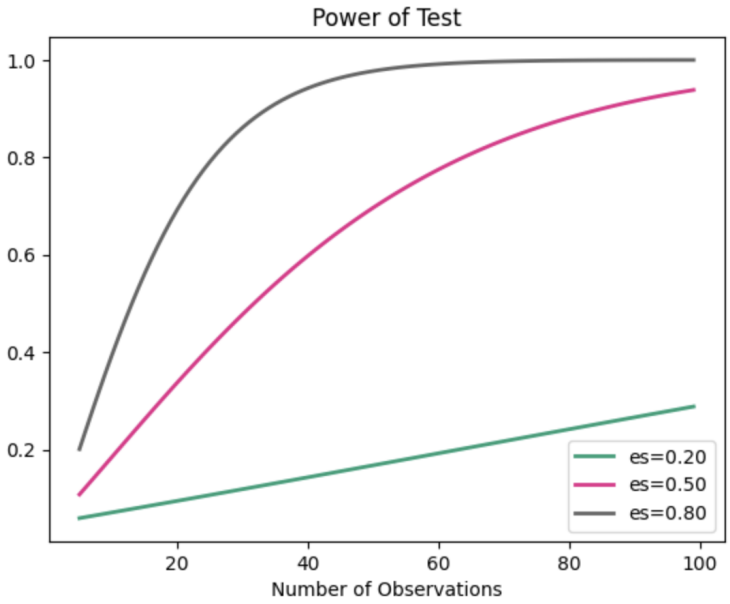

Figure 11: Power of Test

Figure 11 displays the correlation between the power and the sample size across various effect sizes. As depicted in the plot, as the number of observations increases, the power also increases for all three effect sizes. Additionally, at a constant level of sample size, a larger effect size corresponds to a higher power.

3.3 Research Question 1

The hypothesis is following:

Null hypothesis 1: There is no significant difference in the odds of being strip searched between other racial groups ($Race_j$) and the reference group ($Race_{black}$).

Alternative hypothesis 1: There is significant difference in the odds of being strip searched between other racial groups ($Race_j$) and the reference group ($Race_{black}$).

Null hypothesis 2: There is no significant difference in the odds of being strip searched between other age groups (Age_i) and the reference group (Age_{25-34}).

Alternative hypothesis 2: There is significant difference in the odds of being strip searched between other age groups (Age_i) and the reference group (Age_{25-34}).

Null hypothesis 3: There is no relationship between negative action and the probability of being strip searched.

Alternative hypothesis 3: There is relationship between negative action and the probability of being strip searched.

3.4 Research Question 2

The hypothesis is following:

Null hypothesis 1: The population means of two independent levels: youth and not a youth, are equal, controlling the number of negative actions at arrest.

Alternative hypothesis 1: The population means of two independent levels: youth and not a youth, are not equal, controlling the number of negative actions at arrest.

4. Result

4.1 Research Question 1 - Logistics Model

Table 3: Model summary for logistic model, with significance code: "***" for p-value < 0.0001, "**": 0.001, "*": 0.01, " ": 0.1.

Term	Estimate	Standard Error	Lower CI (Exponential)	Upper CI (Exponential)	OR
Intercept ***	-1.9195	0.045	0.1343	0.1601	0.1467
Negative Action ***	0.6365	0.025	1.7982	1.9860	1.8898
Ease/Southeast Asian ***	-0.9142	0.099	0.3304	0.4863	0.4008
Latino/Indigenous ***	-0.6139	0.110	0.4362	0.6716	0.5412

Middle Eastern ***	-1.0764	0.126	0.2663	0.4362	0.3408
South Asian ***	-0.9708	0.114	0.3032	0.4731	0.3788
Unknown or Legacy ***	-0.8564	0.097	0.3511	0.5137	0.4247
White	-0.0479	0.046	0.8714	1.0427	0.9532
Aged 17 and under ***	-0.4600	0.105	0.5134	0.7762	0.6313
Aged 18 to 24 *	0.1207	0.058	1.0076	1.2634	1.1283
Aged 35 to 44	0.0327	0.053	0.9306	1.1473	1.0333
Aged 45 to 54 **	-0.1845	0.066	0.7107	0.9462	0.8315
Aged 55 to 64 ***	-0.4292	0.088	0.5481	0.7732	0.6510
Aged 65 and older ***	-1.4771	0.218	0.1490	0.3497	0.2283

A logistic regression model was fitted to investigate how the number of negative actions, levels of perceived race, and categories of age group affect the probability of individuals being subjected to strip searches. Most of the features were found to be statistically significant as indicated by the asterisks (*) in Table 3.

The intercept represents the average log odds of strip searches for individuals who are black and aged between 25 to 34 years. The odds ratio was approximately 0.15, which implies that black individuals in the age group of 25 to 34 years have a probability of around 12.79% of being strip searched.

The increasing number of negative actions at the time of arrest was linked to an increased likelihood of an individual being subjected to a strip search, with a 95% confidence interval of [1.7983, 1.9680]. The odds ratio was approximately 1.89, indicating that for each additional unit increase in the number of negative actions, the odds of the individual being strip searched increased by around 1.89 times, keeping other variables constant.

Regarding perceived race, the odds of being strip searched for most races are significantly different from black individuals, except for white individuals. Specifically, East/Southeast Asians have statistically significantly lower odds of being strip searched compared to black individuals, with a 95% CI of [0.3304, 0.4863]. At a fixed level of age group and number of negative actions,

the odds of being strip searched are about 60% lower for East/Southeast Asians compared to black individuals, given an odds ratio of 0.4008.

Similarly, Latino/Indigenous individuals have statistically significantly lower odds of being strip searched compared to black individuals, with a 95% CI of [0.4362, 0.6716]. The odds of being strip searched are about 45.88% lower for Latino/Indigenous individuals compared to black individuals, given the odds ratio of 0.5412, while controlling for other variables.

Middle Eastern individuals also have statistically significantly lower odds of being strip searched compared to black individuals, with a 95% CI of [0.2663, 0.4362]. At a fixed number of negative actions and age groups, the odds of being strip searched are about 65.92% lower for Middle Eastern individuals compared to black individuals, given the odds ratio of 0.3408.

South Asians have statistically significantly lower odds of being strip searched compared to black individuals, with a 95% CI of [0.3032, 0.4731]. Controlling for other variables, the odds of being strip searched are about 62.12% lower for South Asians compared to black individuals, given the odds ratio of 0.3788.

The odds of being strip searched for people in unknown or legacy racial groups are statistically significantly different from black individuals, with a 95% CI of [0.3511, 0.5137]. Controlling for other variables, the odds of being strip searched are about 57.53% lower if the individual's race is unknown or legacy compared to black individuals, given the odds ratio of 0.4247.

In contrast, the odds of being searched for white individuals is not statistically significantly different from black individuals with 95% CI [0.8714, 1.0427]. The odds ratio is 0.9532, meaning that controlling other variables unchanged, the odds of being strip searched were only about 5% lower for white individuals compared to black individuals, and this difference was not statistically significant.

Regarding age groups, the odds of being strip searched for most age groups are significantly different from aged 25 to 34, except for the group of aged 35 to 44. Specifically, individuals aged 17 and under have statistically significantly lower odds of being strip searched compared to those aged 25 to 34, with a 95% CI of [0.5134, 0.7762]. The odds ratio for this group is 0.6313, indicating 40% lower odds of being strip searched than those aged 25 to 34, with other variables held constant.

On the other hand, individuals aged 18 to 24 have a statistically significantly higher odds of being strip searched compared to those aged 25 to 34, with a 95% CI [1.0076, 1.2634]. The odds

ratio is 1.1283, which means 12.83% higher odds of being strip searched for individuals aged 18 to 24 than those aged 25 to 34, keeping other variables constant.

The odds of being strip searched are significantly lower for individuals aged 45 to 54, 55 to 64, and 65 and older compared to those aged 25 to 34, with 95 % CI [0.7107, 0.9462], [0.5481, 0.7732], and [0.1490, 0.3497], respectively. Keeping other variables unchanged, the odds of being strip searched are about 16.85%, 34.90%, 77.17% lower for individuals aged 45 to 54, 55 to 64, and 65 and older, respectively, with corresponding odds ratios of 0.8315, 0.6510, and 0.2283.

Lastly, the odds of being strip searched for individuals aged 35 to 44 is not statistically significantly different from those aged 25 to 34, with a 95% CI [0.9306, 1.1473]. The odds ratio is 1.0333, which means the odds of being strip searched barely changes between individuals aged 35 to 44 and those aged 25 to 34, keeping other variables unchanged.

The fitted model was applied to the test set (containing features of age groups, perceived race, and number of negative actions) to generate predictions. These predictions were then compared to the actual outcomes in the test set, resulting in a test accuracy of approximately 89%. This means the model correctly predicted the outcome of 89% of the test dataset.

To estimate the likely range of values for a new observation, prediction intervals are typically computed with a 95% confidence level. After comparing predicted values with actual outcomes in the test set, it was found that 90.84% of target values fall within the prediction intervals, which is considered a favorable result. The table below displays several prediction intervals that were calculated, and it can be observed that all of the predicted values are encompassed by their corresponding intervals.

Table 4: Prediction Interval

Predicted Value	Prediction	Lower Bound	Upper Bound
0.1529	0	-0.5525	0.8582
0.1262	0	-0.5247	0.7772
0.0605	0	-0.4067	0.5276
0.0573	0	-0.3981	0.5127
0.0555	0	-0.3933	0.5044

Due to the fact that our model comprises three different features which are age groups, perceived race, and number of negative actions, it is difficult to represent the data points, prediction interval, and confidence interval in a two-dimensional plot. Therefore, generating a prediction interval plot is not feasible.

Furthermore, the dataset is highly imbalanced between strip-searched and not searched observations. Therefore, it is necessary to conduct a confusion matrix in order to analyze the model's prediction performance.

Table 5: Confusion Matrix for logistic model

	Predicted Negative	**Predicted Positive**
Actual Negative	6037	15
Actual Positive	735	22

Table 5 shows that the logistic model had a true negative of 6,037, which means that individuals who were not strip searched were correctly identified as not having been strip searched by the model. In other words, the model correctly predicted 6,037 cases of individuals who were not strip searched.

The model had a true positive of 22, which indicates that out of all the individuals who were actually strip searched, the model correctly predicted 22 of them. However, the model had a false negative of 735, meaning that the model incorrectly predicted that 735 cases were not strip searched when they actually were. Additionally, the model had a false positive of 15, which means that 15 cases were incorrectly predicted to be strip searched by the model when they were not actually searched.

Overall, there were 6,059 correct predictions and 750 incorrect predictions. The model only accurately predicted 22 out of the total of 757 actual strip-searched data. Thus, the model is not performing well in identifying the positive outcome (strip searched), with a recall of 0.029 and a precision of 0.595. However, for the total of 6,052 negative outcomes (not strip searched), the model can predict 99.7% of them correctly, indicating that the model is on average good at predicting the negative outcome.

4.1.1 Assumption Validation

In Section 3.1, we discussed that the outcome variable represents the strip search status, which is indicated by either 0 (not searched) or 1 (searched). Therefore, the assumption of a binary outcome is satisfied. Additionally, each row in the dataset corresponds to a unique individual, and there are no duplicated rows for any individuals, indicating that the assumption of independent observations is satisfied. As outlined in Section 3, the dataset consists of 34,043 observations, which is deemed sufficient for performing logistic regression. As a result, the assumption of a large sample size is satisfied.

The absence of multicollinearity assumption is satisfied as indicated by Figure 11, which shows very weak correlation between all the independent variables.

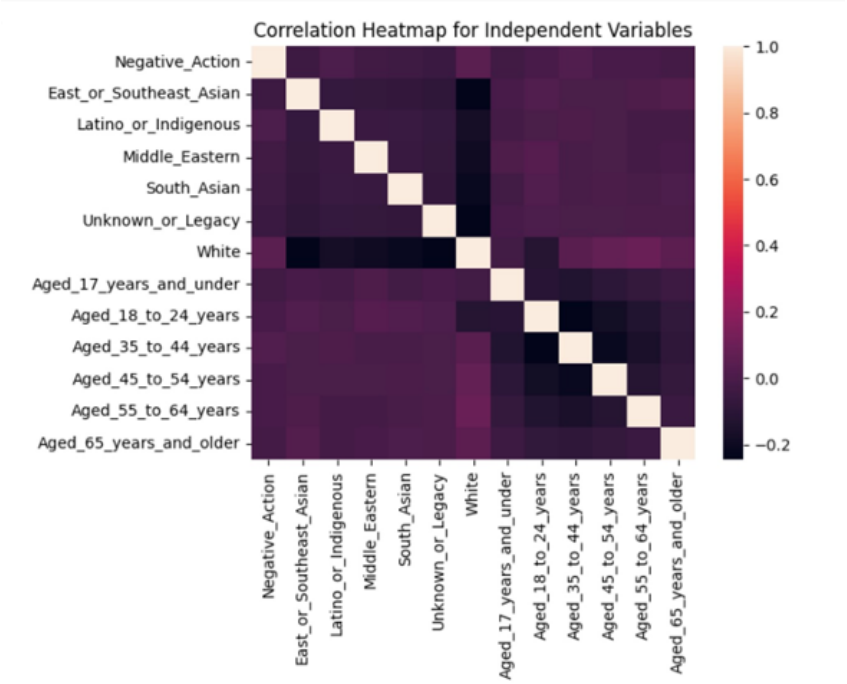

Figure 11 : Correlation heatmap for independent variables

A regression plot was created in Figure 12 to assess the linearity between the continuous variable, negative action, and log odds of being strip searched. The plot indicates an S-shaped curve, which suggests that the linearity assumption has been met.

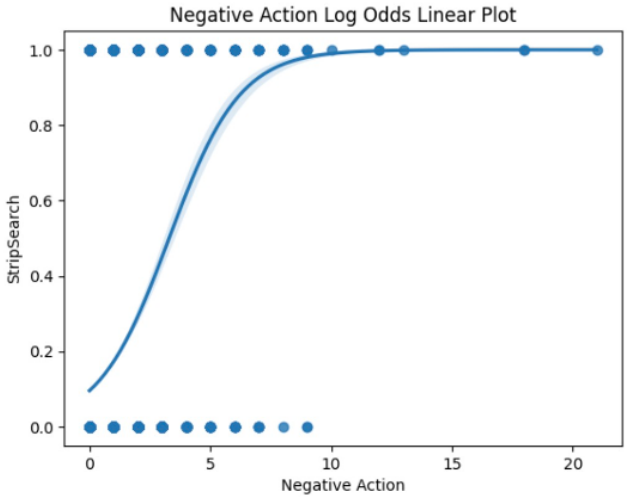

Figure 12: Regression plot for linearity assumption check.

4.2 Research Question 2 - ANCOVA

Table 5: ANCOVA Model Results

Source	SS	DF	F	p-unc
Youth Status	6612.711	1	11.3088	0.0008
Negative Action	201465.23	1	344.5375	<0.0001
Residual	19904588	34040	NaN	NaN

Table 5 displays that the p-values for youth status and negative action are both below the significance level of α=0.05, indicating that we have strong evidence to reject the null hypothesis that there is no difference in the mean percentage of strip searches for youth and non-youth individuals controlling for the influence of the number of negative actions at arrest. As a result, we can infer that youth status (F=11.3, P-value < 0.001) and number of negative actions (F = 344.54, p-value < 0.001) both have an impact on the percentage of strip searches conducted.

We hypothesized that the people's youth status would be able to predict the percentage of strip searches. From our result, we see that this is a statistically significant relationship between youth status and the percentage of strip search when controlling the negative actions at arrest.

4.2.1 Assumption check

To test the equal variance assumption, we conducted Levene's test. The results indicate that we have evidence to reject the null hypothesis that the variances of different groups are equal. Therefore, the assumption of constant variance is violated (statistic=59.08, $p<0.05$).

According to the results of the Shapiro-Wilk test, we have enough evidence to reject the null hypothesis and conclude that the residuals are not normally distributed ($W=0.39$, $p=0.0$), however, the result may not be accurate for a sample size greater than 5000.

Figure 13 illustrates a weak positive correlation between the percentage of strip searches and the number of negative actions at the time of arrest. This implies that as the number of negative actions increases, the percentage of strip searches also increases. The strength of the relationship is not significant when the number of negative actions is less than 10; however, when it exceeds 10, an upward trend between the outcome and negative action is noticeable. As a result, we can conclude that the linearity assumption is met.

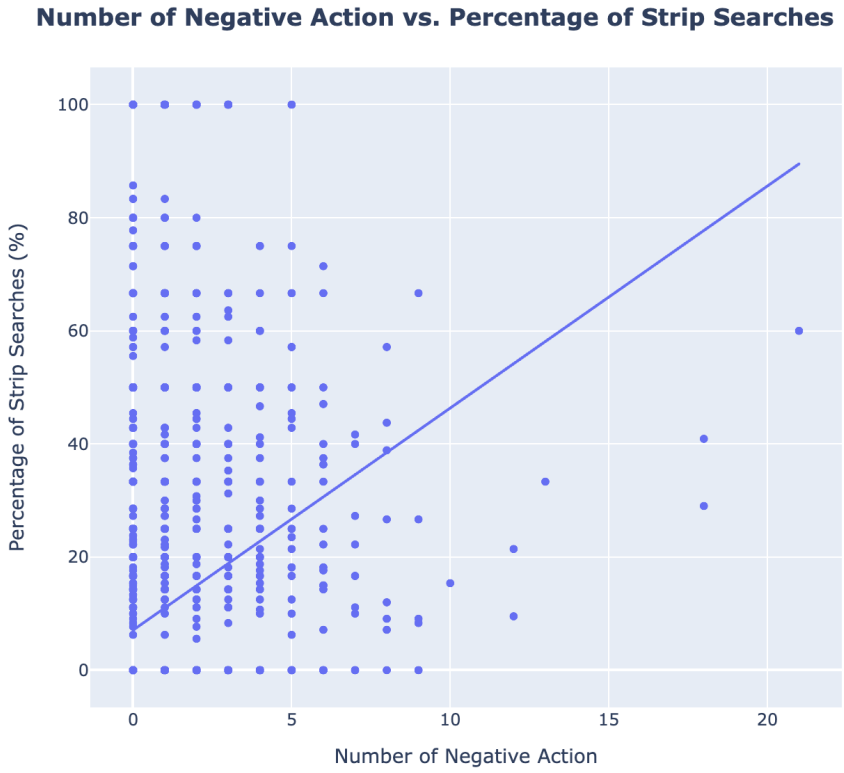

Figure 13: Plot for linearity assumption check

A multiple regression model is created with an interaction term (which is the multiplication of Youth and Negative Action) to investigate whether the impact of a one-unit increase in the number of negative actions on the outcome is consistent at all fixed levels of youth status, and vice versa.

Table 6: Interaction model results

	Estimates	**Std err**	**P-value**	**Lower CI**	**Upper CI**
Intercept	7.19	0.139	< 0.001	6.917	7.463
Youth Status [Youth]	-1.71	0.589	0.004	-2.867	-0.557
Negative Action	3.95	0.212	< 0.001	3.536	4.366
Youth : Negative Action	-2.98	1.65	0.071	-6.217	0.259

According to the results presented in the table, the p-value associated with the interaction term is higher than the significance level (0.05). Therefore, there is no significant evidence to reject the null hypothesis that the interaction term does not have a meaningful impact on the model. This suggests that there is no interaction effect between youth status and negative action, and the assumption of homogeneity of the slope is met.

5. Discussion

The results of the logistic regression model indicate that the number of negative actions, perceived race, and age group all have a significant effect on the probability of individuals being subjected to strip searches. The intercept suggests that black individuals aged 25 to 34 years have the highest probability of being strip-searched, with a probability of around 12.79%. The number of negative actions during arrest was positively associated with an increased likelihood of being strip-searched. This finding is consistent with previous studies that have linked more severe criminal charges and violent behavior during the arrest process with an increased likelihood of strip searches.

Perceived race was also found to be a significant predictor of strip searches. The results show that individuals from East/Southeast Asian, Latino/Indigenous, Middle Eastern, South Asian, and unknown or legacy racial groups had significantly lower odds of being strip searched compared to black individuals. However, white individuals did not have statistically significant odds of being strip searched compared to black individuals. These findings suggest that race plays a role in the decision to conduct a strip search and that black individuals are disproportionately affected.

Age group was also found to have a significant effect on the probability of strip searches. Individuals aged 17 and under had statistically significantly lower odds of being strip searched compared to those aged 25 to 34, while individuals aged 18 to 24 had statistically significantly higher odds of being strip searched compared to those aged 25 to 34. Individuals aged 35 to 44 had similar odds of being strip-searched compared to those aged 25 to 34. However, individuals aged 45 to 54, 55 to 64, and 65 and older had statistically significantly lower odds of being strip-searched compared to those aged 25 to 34.

These findings have important implications for policymakers and law enforcement officials. Strip searches are invasive and can be traumatic for individuals, and their use should be limited to situations where there is reasonable suspicion that the individual is concealing contraband. The results of this study suggest that there may be racial bias in the use of strip searches and that black individuals are disproportionately affected. Law enforcement officials should be aware of these biases and work to eliminate them through training and policy changes. Additionally, the findings suggest that age should also be taken into account when making decisions about strip searches, as younger individuals may be more likely to be subjected to them.

Our study has several limitations, the t-test's normality assumption is not met because our sample size exceeds 5000, which is the threshold at which the Shapiro-Wilk test is reliable. Thus, the p-value may be unreliable. The smaller sample size of individuals who have been strip searched means that it may not be representative enough to properly study the relationship between the mean percentage of strip searches and other variables. The use of a single dataset, which may not be representative of all law enforcement agencies in Canada.

Our analysis did not consider other variables that may impact the likelihood of strip searches, such as the severity of the crime or the location of the arrest. And there is an imbalance in the number of people between the youth and non-youth groups, which may lead to inaccurate results in the ANCOVA analysis. Furthermore, the linearity assumption between the percentage of strip searches and the number of negative actions has been violated, which could also result in inaccurate results.

Despite these limitations, our study sheds light on the disproportionate use of strip searches in law enforcement, which may have implications for civil liberties and human rights. Moving forward, it is important for law enforcement agencies to reevaluate their policies and procedures regarding strip searches and ensure that they are not influenced by implicit biases or stereotypes. Future implications should investigate other variables that may impact the likelihood of strip searches such as location at arrest and the severity of the crime that has an

impact on the probability of being searched and explore potential solutions to reduce the use of strip searches in law enforcement while still maintaining public safety.

6. Conclusion

In conclusion, this study aimed to investigate the relationship between strip searches and the number of negative actions, perceived race, and age group. The logistic regression model found most of the features to be statistically significant. The number of negative actions during arrest was positively associated with the likelihood of an individual being subjected to a strip search, while perceived race and age group were also significant predictors.

The odds of being strip searched for black individuals aged 25 to 34 were the highest among all the age groups and races. However, individuals from East/Southeast Asian, Latino/Indigenous, Middle Eastern, South Asian, and unknown/legacy racial groups were less likely to be strip searched compared to black individuals, even after controlling for other variables. Similarly, individuals aged 18 to 24 were more likely to be strip searched, while those aged 45 to 54, 55 to 64, and 65 and older were less likely to be strip searched compared to those aged 25 to 34. To be more specific, when comparing youth and adults, youth had a lower proportion of strip searches.

Overall, these findings suggest that strip searches are not applied uniformly to all individuals but instead are influenced by their perceived race and age group. As such, law enforcement agencies must be vigilant in ensuring that their policies and procedures do not lead to discrimination in the application of strip searches. Further research is needed to investigate the reasons behind these disparities and to identify potential solutions to reduce them.

References

Lemke, M. (2022, June 17). *Strip searches are ineffective, unnecessary and target racialized Canadians*. Phys.org. Retrieved April 11, 2023, from https://phys.org/news/2022-06-ineffective-unnecessary-racialized-canadians.html

Lemke, M. (2022, July 18). *Policing toronto: Strip searching in a divided city - the bullet*. Socialist Project. Retrieved April 11, 2023, from https://socialistproject.ca/2022/07/policing-toronto-strip-searching-in-a-divided-city/

Police powers: Strip searching. BSB Solicitors. (2022, August 19). Retrieved April 11, 2023, from https://www.bsbsolicitors.co.uk/blog/police-powers-strip-searching/

Prison Policy Initiative. (2019, May 14). *Policing women: Race and gender disparities in police stops, searches, and use of Force*. Prison Policy Initiative. Retrieved April 11, 2023, from https://www.prisonpolicy.org/blog/2019/05/14/policingwomen/

Susan Nembhard and Lily Robin. (2021, August). Racial and Ethnic Disparities throughout the Criminal Legal System A Result of Racist Policies and Discretionary Practices. Urban Institute. Retrieved April 12, 2023, from https://www.urban.org/sites/default/files/publication/104687/racial-and-ethnic-disparities-throughout-the-criminal-legal-system.pdf

Daniel M. Blonigen. (2021). Explaining the relationship between age and crime: Contributions from the developmental literature on personality. Clinical Psychology Review, Volume 30, Issue 1, 89-100.

Arrests and strip searches (RBDC-arr-TBL-001). Toronto Police Service Public Safety Data Portal. (2022, November 10). Retrieved April 13, 2023, from https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/about

Sally Weale and Vikram Dodd. (2022, August 8). *Revealed: Met police strip-searched 650 children in two-year period*. The Guardian. Retrieved April 11, 2023, from https://www.theguardian.com/uk-news/2022/aug/08/police-data-raises-alarm-over-welfare-of-strip-searched-children