INF2178 Final Project

“Examining Disparities in Strip Search Rates: A Comprehensive Analysis of Demographic Factors and Arrest Circumstances in Toronto”

Date: 16/04/23

Student 1: Moxuan Huang 1004814882

Student 2: Nicholas Leong 1005330447

Instructor: Shion Guha

Faculty of Information

Human-Centered Data Science

Table of Contents

Introduction

The use of strip searches by police officers has become a topic of controversy, as concerns have been raised about the impact of strip searches on individual rights, racial disparities, and gender equality. In this project, we aim to analyze the dataset “Arrests and Strip Searches” (RBDC-ARR-TBL-001), a large public dataset shared by the Toronto Police Services that can be found in the Public Safety Data Portal. A strip search is a type of search where a police officer removes some or all of a person's clothing and visually examines their body. We are interested in investigating how different factors such as race, gender, and age group, along with the type of action taken at arrest, are associated with the likelihood of a strip search being conducted. Our study seeks to determine whether there are statistically significant differences in strip search rates between various demographic groups and to understand the association between these factors and the likelihood of a strip search being conducted after controlling for relevant covariates. By examining these issues, we hope to identify whether strip searches are being conducted fairly and equitably across different demographic groups and situations. The results could contribute to broader discussions about policing practices, and the need for greater accountability and transparency in law enforcement. Ultimately, this study could provide insights into areas where policy changes or training may be required to improve practice, helping to promote fairness and justice in the policing of Toronto communities.

Literature review

Strip searches are an intrusive form of the police search in which an individual's clothing is removed for visual inspection. The use of strip searches by the police has been controversial, with many people and groups debating their role. Michael Grewcock and Vicki Sentas argue that police departments and local governments should address the impact of strip searches on the rights of individuals, particularly young people and individuals from marginalized communities [1]. They also emphasize the importance of a more balanced and proportionate approach to strip searches that emphasizes the rights and dignity of individuals while ensuring public safety.

According to David M Tanovich, gender and race-based violence has been perpetuated by the Toronto Police Service, citing multiple cases of strip searches of women resulting in sexual assault and abuse of prosecutorial power [2]. These acts can disproportionately punish individuals with a history of sexual assault or abuse who are already at high risk for post-traumatic symptoms. In addition, an independent expert evaluation report examining the first phase of the Toronto Police Service's race-based data collection strategy showed that Black and Indigenous women were disproportionately affected by strip searches, suggesting a pattern of racialized violence by Toronto police [3].

Michelle Psutka and Elizabeth Sheehy provide a thorough analysis of police strip searches, noting that they have become a pretext for sexual assault, abuse of prosecutorial power, and other violations of individual rights [4]. The authors argue that this practice has become normalized and routine, leading to a failure to recognize the violation of these rights and dignity inherent in these searches. They argue that the Toronto police have faced criticism for their

handling of strip search complaints, including a lack of transparency and accountability; these weaknesses have contributed to public mistrust of law enforcement.

Kirkup argued that the use of strip searches conflicts with principles of Canadian law, which recognizes and protects an individual's right to physical integrity and privacy [5]. Also, the use of strip searches is contrary to Canada's principles of law and violates an individual's right to physical integrity and privacy. Kirkup noted that the use of strip searches on transgender people has been criticized for failing to recognize the complexities of gender identity and for its impact on already marginalized communities. In addition, Kirkup pointed to an imperfect legal framework surrounding strip searches, leading to confusion about the legality of these searches and a lack of accountability for officials involved in these practices.

Toronto police research on the topic of arrests and strip searches highlights issues of gender, racial violence, sexual assault, prosecutorial power and the use of strip searches [3]. Evidence shows that these searches can lead to violations of individual rights and dignity and that an imperfect legal framework surrounding them confuses people about their legality. In addition, officials who carry out these practices are not held accountable for their actions. Toronto police must take steps to address these issues by increasing transparency about their policies and practices and providing training on the proper use of strip searches as well as recognizing their impacts on marginalized communities.

Research objective and questions

We aim to investigate the potential differences in the number of strip searches conducted by the Toronto Police based on various demographic factors and the circumstances of arrest. Our research will primarily focus on the top 5 crime occurrence categories, including "Assault," "Assault & Other crimes against persons," "Robbery & Theft," "Warrant," and "FTA/FTC/Compliance Check/Parollee." With 31 types of occurrence categories in total, we believe that focusing on the top 5 categories will provide a higher representation, as smaller categories could introduce bias that impacts our results. Our study will analyze perceived race, gender, age, and action taken at arrest separately to identify their relationships with the number of strip searches. Through this study, we hope to outline any potential bias or discrimination in Toronto Police's strip search practices and try to identify any discrepancies and understand the underlying reasons behind them. We believe the following research questions can help us gain insights into the dataset from different perspectives:

Research Question 1

- What is the minimum sample size required to detect a statistically significant difference in strip search rates between different racial, gender, and age groups with a specified effect size, level of significance, and power?

Research Question 2

- Are there significant differences in strip search rates between different racial, gender, and age groups after controlling for covariates such as the type of action taken at arrest?

Research Question 3

- What is the association between race, gender, age group, and type of action taken at arrest on the likelihood of a strip search being conducted?

Dataset Description

The "Arrests and Strip Searches" dataset includes 24 attributes and a total of 65,276 records that provide information on all strip arrests and searches conducted by Toronto Police. This dataset contains information about each arrest, including the year between 2021 and 2022 and the month recorded quarterly in which it occurred, as well as the unique identifiers of the persons arrested and/or searched. The dataset also includes information about the person's race in eight different types such as White and Black; gender including male, female and unknown; age group at the time of arrest as well as whether they were under 18 years old at that time. The arrest locations are provided for each subdivision, indicating where arrests occurred within those boundaries. Some subdivisions could not be geo-coded or the arrest took place outside of the City of Toronto boundaries marked as XX.

Additionally, The dataset contains information on whether a strip search was performed, whether the person was booked to police within 24 hours of arrest, and 31 types of occurrence categories associated with the arrest including robbery & theft, assault and sexually related crime etc. The dataset also contains attributes related to actions taken at the time of arrest, including whether the person had concealed objects, was belligerent or violent, resisted arrest, was mentally unstable or potentially suicidal, assaulted police, or cooperated. Information on reasons

for searches including whether searches were to inflict harm, facilitate escape, find weapons or seek evidence are also contained in the dataset. Lastly, it contains information about any items found during searches.

Descriptive statistics/ EDA

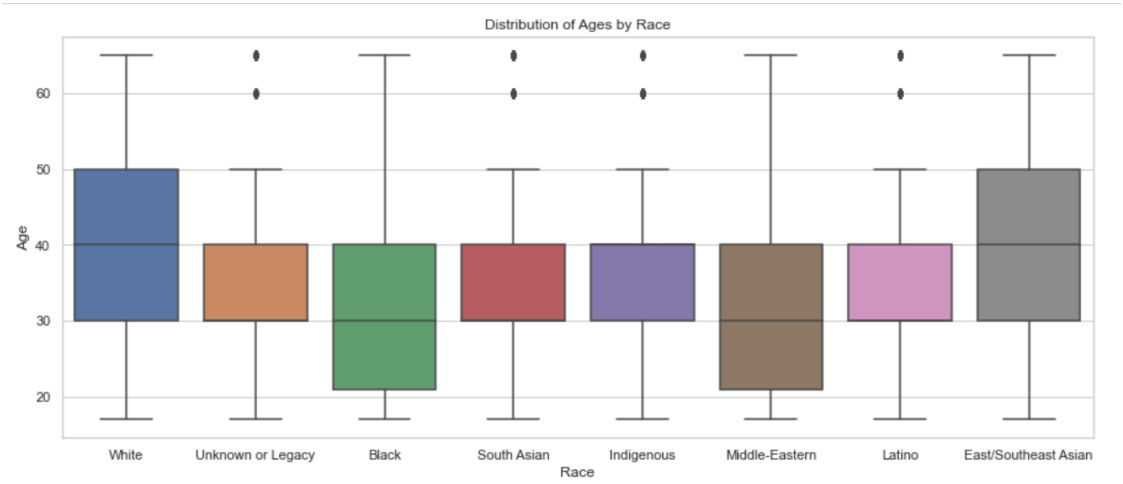

Figure 1, Distribution of ages by perceived race

The boxplot in figure 1 shows the distribution of ages at arrest for different races/ethnicities. We can see that the median age for each race/ethnicity falls between the 25th and 75th percentile range, indicating a relatively symmetrical distribution. However, we can see some variation in the age distributions across different races/ethnicities. For example, the median age for Black individuals at arrest appears to be slightly younger than that for Hispanic and

White individuals. Additionally, the interquartile range (IQR) for Black individuals appears to be slightly wider than that for other races/ethnicities. Overall, the plot suggests that there may be some differences in the ages at which individuals from different races/ethnicities are arrested.

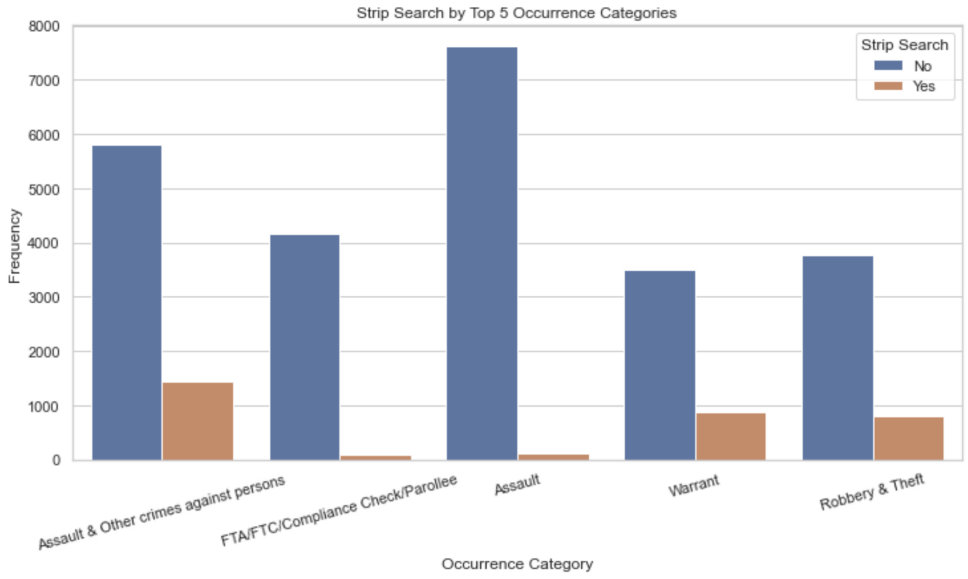

Figure 2, Frequency of conducting strip search based on top 5 crime occurrence categories.

The count plot presents the frequency of strip searches by the top 5 occurrence categories. Clearly, "Assault and Other Crimes Against a Person" has the highest number of strip searches, followed by "Weapons Offenses," "Drug and Narcotic Offenses," "Disorderly Conduct," and "Larceny/Theft Offenses." This suggests that individuals arrested for crimes related to violence, weapons, and drugs are more likely to be subjected to strip searches. In addition, we can see that the majority of occurrences in each category did not result in a strip search, as the bar for "No" is higher than the bar for "Yes" in each category. This could indicate

that strip searches are not performed in every arrest within these categories and are reserved for certain situations or individuals.

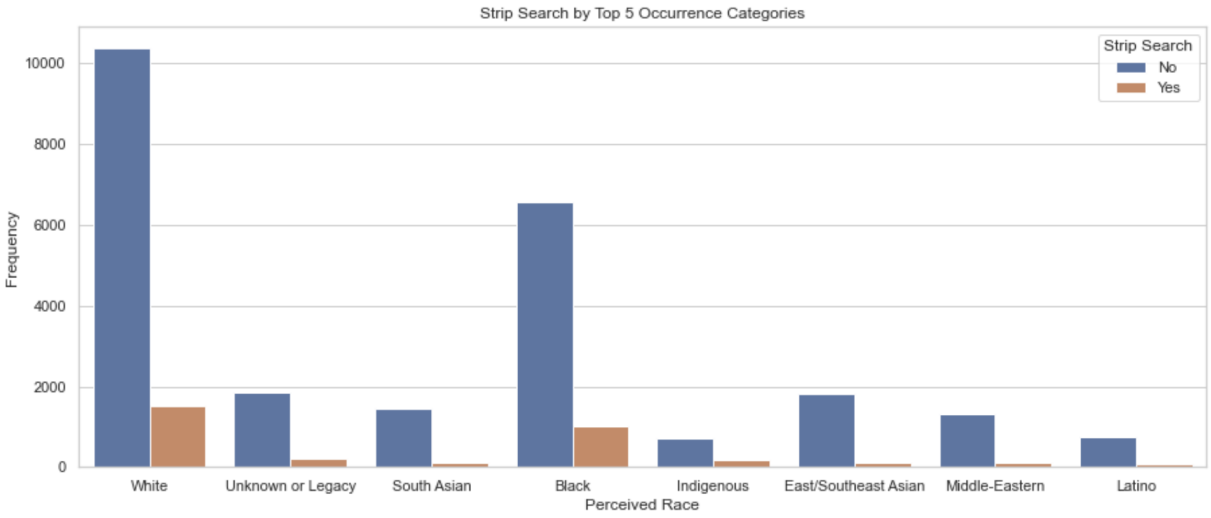

Figure 3, Frequency of conducting strip search based on perceived race

Based on the frequency of conducting strip searches based on perceived race, it appears that white individuals have the highest frequency of being searched, followed by black individuals. However, when we consider the proportion of individuals who are searched within each racial group, it is clear that black individuals have a higher proportion of being strip searched. This is definitely a concerning finding as it suggests that race may be playing a role in the decision to conduct strip searches.

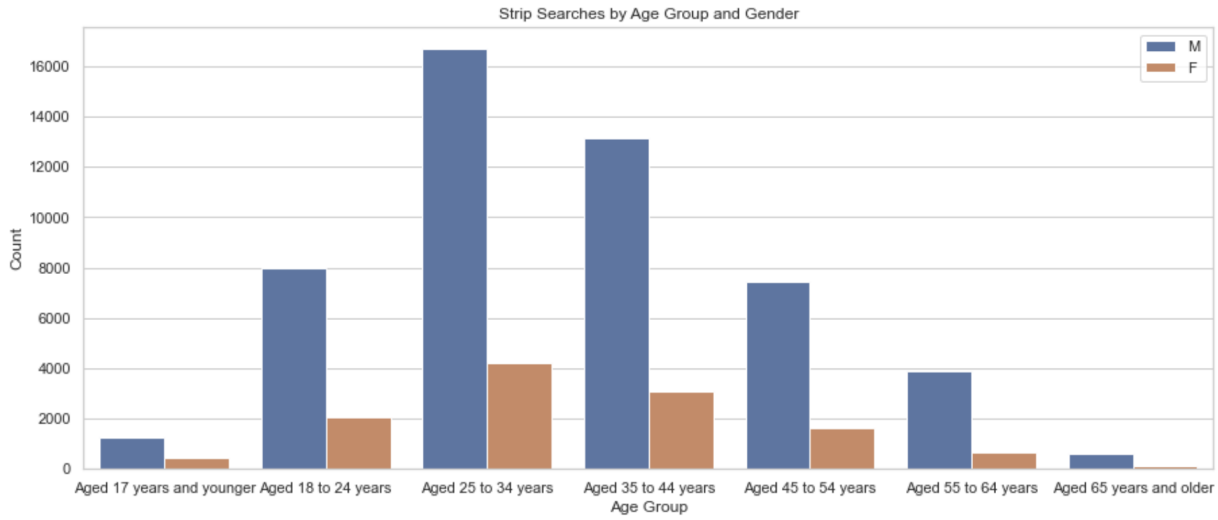

Figure 4, Frequency of conducting strip search based on age group at arrest and gender.

Figure 4 shows the distribution of strip searches by age group and gender. The plot indicates that males tend to have a higher frequency of strip searches than females in every age group from 17 years and younger to 65 years and older. The highest frequency of strip searches occurs in the 25-34 age group for both males and females, followed by the 35-44 age group. Additionally, the plot shows that the frequency of strip searches decreases as the age group increases for both males and females.

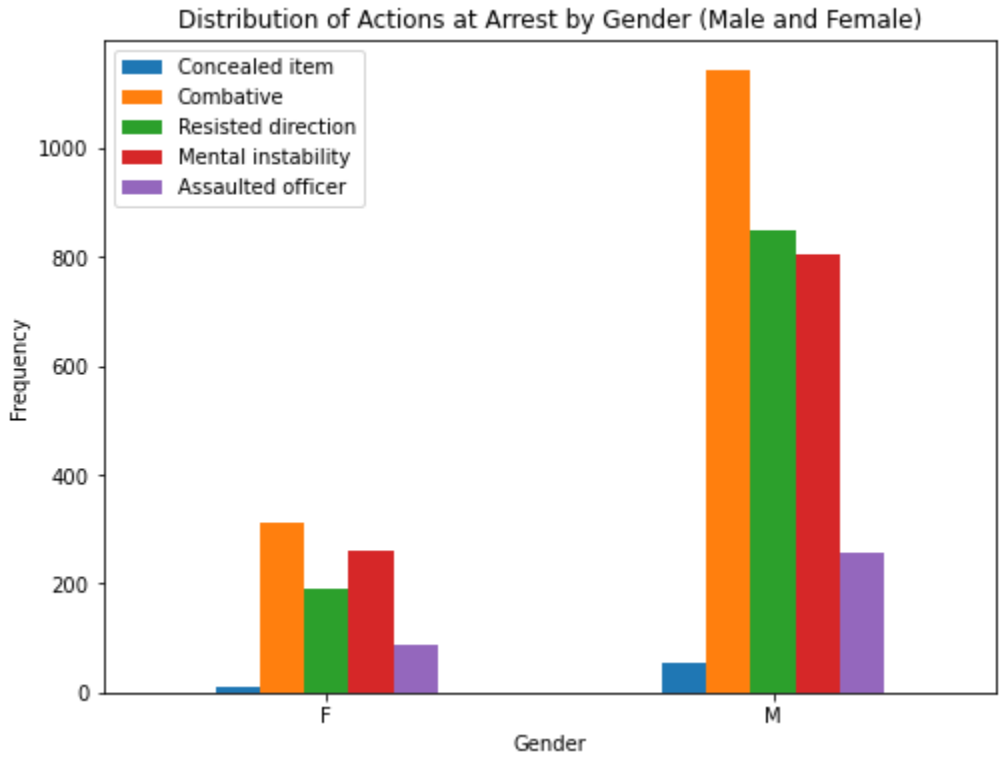

Figure 5, Distribution of actions at arrest by gender.

The bar plot in figure 5 displays the distribution of actions at arrest by gender, specifically for males and females. The actions at arrest considered in this plot include concealed item, combative, resisted direction, mental instability, and assaulted officer. The chart reveals that for both genders, the most common action leading to strip searches is "Combative," followed by "Resisted direction" and "Mental instability." The frequency of these actions is notably higher among males than females, indicating that males are more likely to experience strip searches due to these actions at arrest.

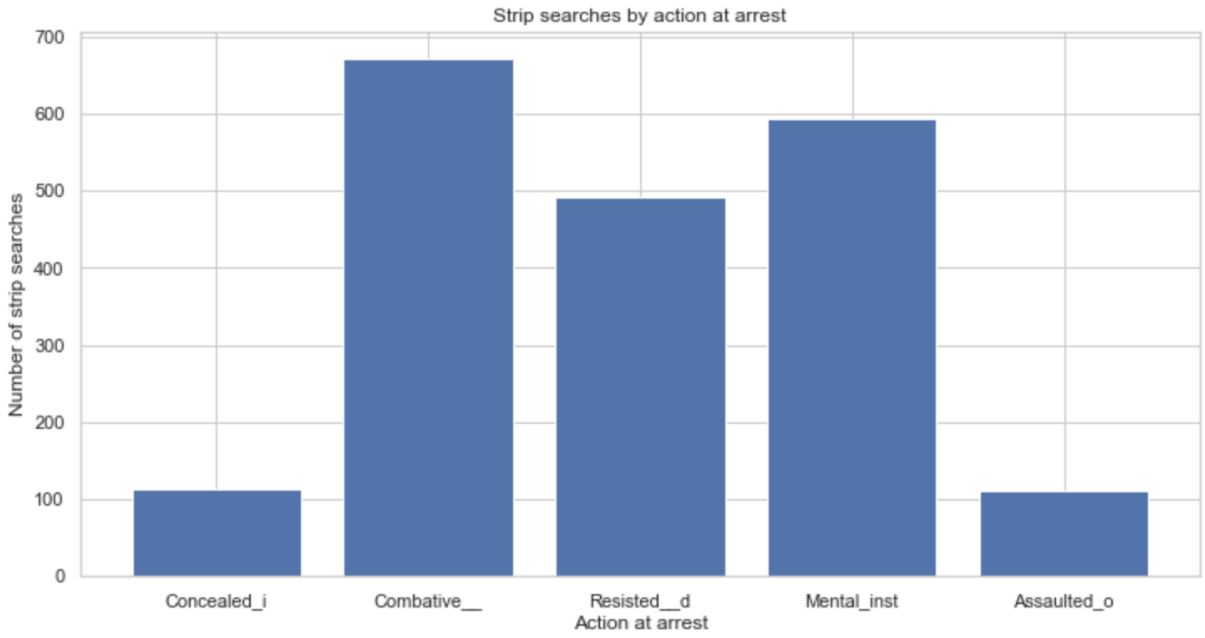

Figure 6, the number of strip searches conducted based on different actions of the arrest.

The bar plot shows the number of strip searches conducted for different actions at the time of the arrest. The plot indicates that the most common action that led to strip searches was being "Cooperative", occurring more than five times higher than the action "Combative" ranking second. This is an important observation and suggests that even individuals who are compliant and cooperative during the time of arrest are not exempt from being subjected to strip searches. It may indicate a broader issue with the use of strip searches as a routine procedure rather than a measure taken only when necessary. The fact that cooperative individuals are still subjected to strip searches at a higher frequency compared to those who conceal items or are combative raises questions about the appropriateness and necessity of these searches. It is important to note that the data presented in this plot does not provide information about the circumstances that led to the strip searches and whether they were conducted in a lawful and non-discriminatory manner.

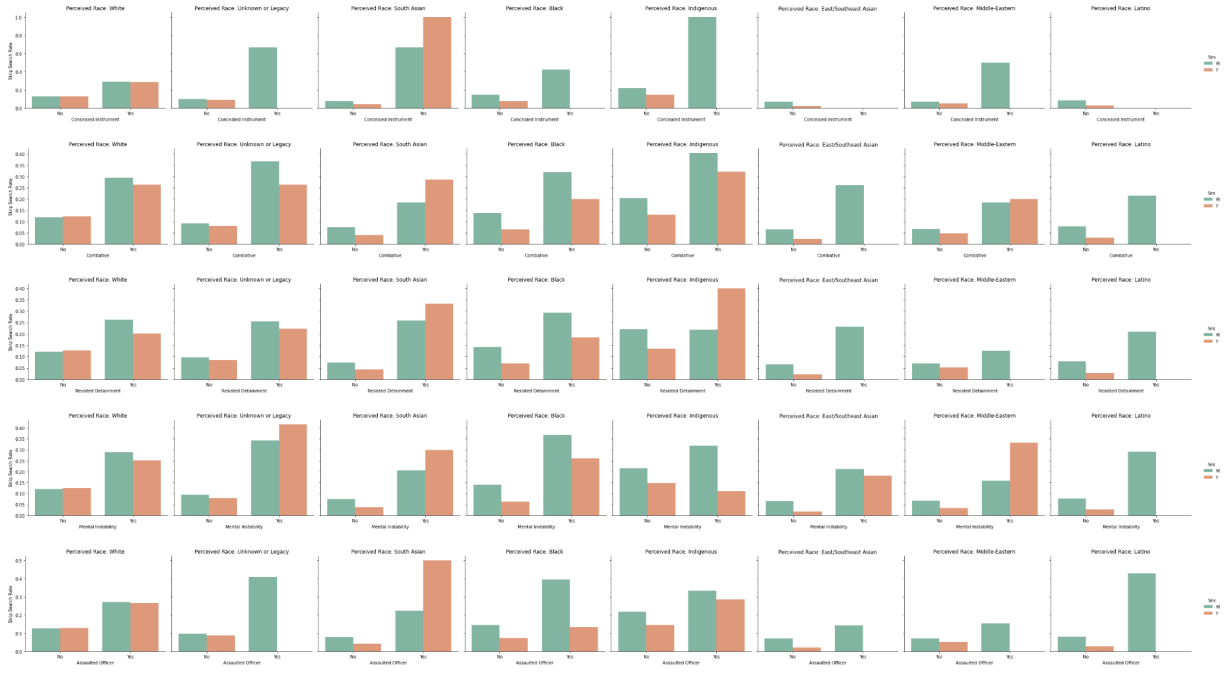

Figure 7, Strip search rates based on perceived race and actions at arrest.

The series of bar plots display the strip search rates based on perceived race and actions at arrest for both males and females. Six actions at arrest are considered, including concealed instrument, combative, resisted detainment, mental instability, assaulted officer, and others. The data is further broken down by perceived race, providing insight into potential racial disparities in strip search rates. For each action at arrest, the strip search rates are compared between males and females across different perceived race categories. The charts reveal varying patterns of strip search rates for different actions and races, indicating that race and gender may play a role in the likelihood of a strip search being conducted. Notably, some racial groups exhibit higher strip search rates for specific actions, suggesting that further investigation is needed to understand the potential impact of racial bias on strip search practices.

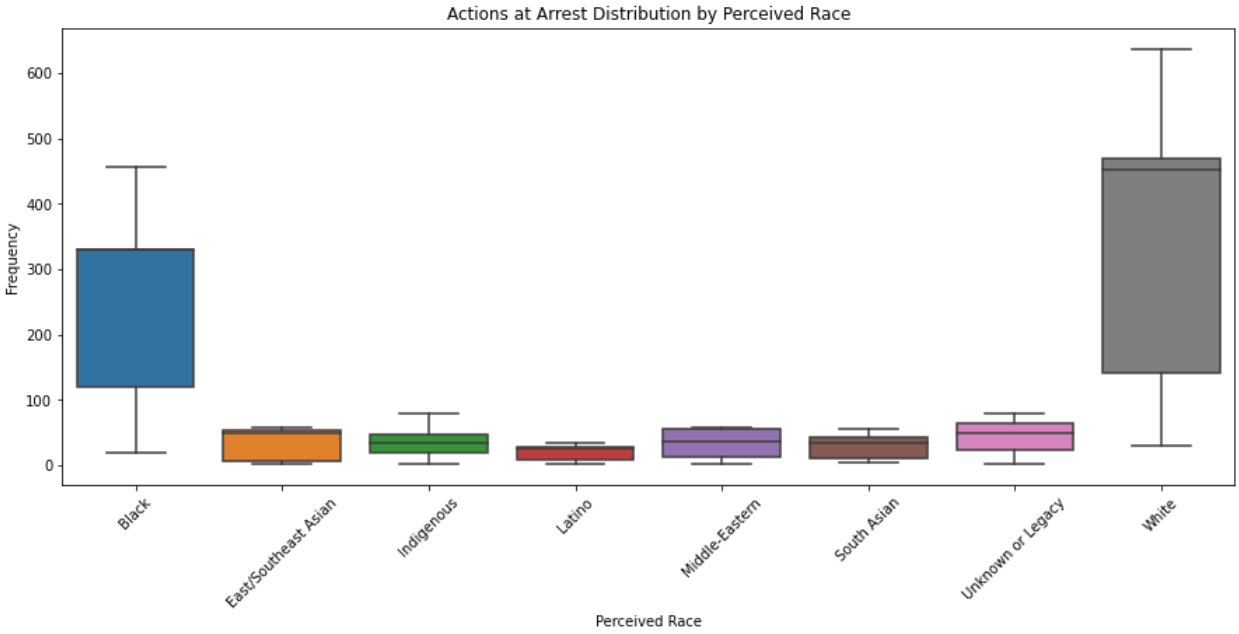

Figure 8, Actions at arrest distribution by perceived race.

The box plot illustrates the distribution of actions at arrest across different perceived racial groups. It is evident that White individuals have the highest frequency of actions at arrest, followed by Black individuals. The whiskers for these two racial groups are much longer than those for the other races, indicating a wider range of actions at arrest frequencies. In contrast, other racial groups, such as Asian, Hispanic, and Native American, have lower frequencies and are nearly on the same level. These groups also exhibit a smaller range of actions at arrest frequencies, as indicated by their shorter whiskers. This visualization highlights the need to consider potential racial disparities in actions at arrest and emphasizes the importance of addressing any underlying biases or systemic issues that may contribute to these disparities, particularly between White and Black individuals and the other racial groups.

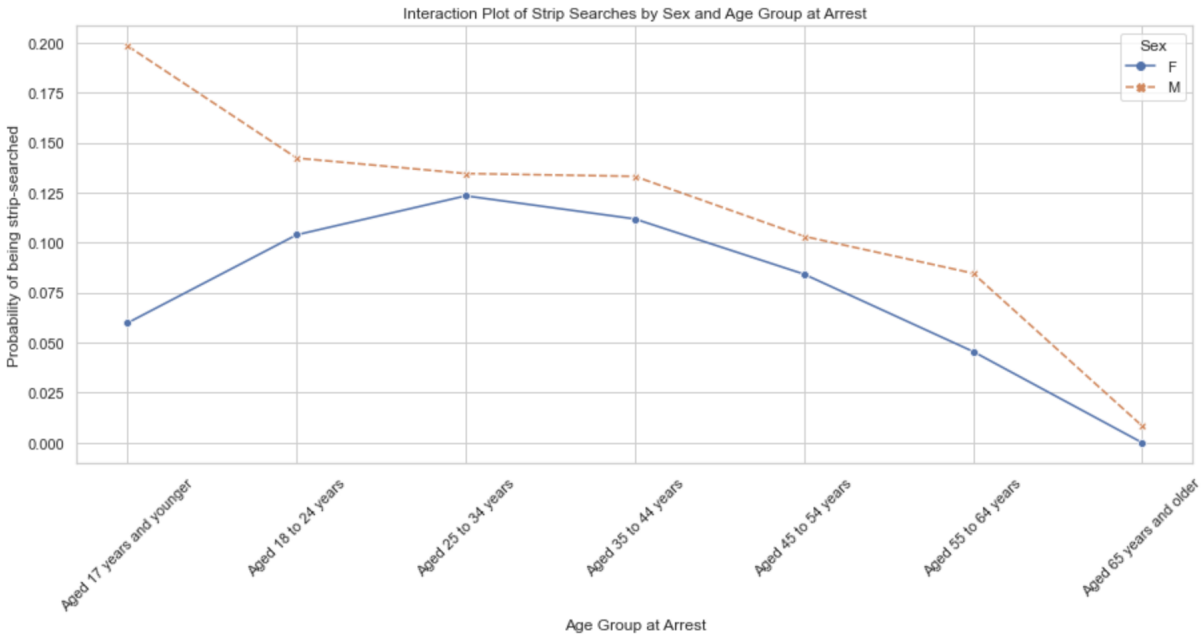

Figure 9, interaction plot of strip searches by sex and age group at arrest.

The interaction plot in figure 9 shows the relationship between strip searches, sex, and age group at the time of the arrest. Interestingly, the only intersections for male and female throughout the entire graph were in the "17 years and under" age group. The plot indicates that there is a significant interaction between sex and age group, where males in the younger age groups (18-24 and 25-34) have a higher mean number of strip searches compared to females in the same age groups. However, as the age group increases, the difference in the mean number of strip searches between males and females decreases, and the frequency of strip searches decrease overall. Overall, the interaction plot suggests that sex and age group have a significant impact on the frequency of strip searches conducted at the time of arrest, with younger males being more likely to be subjected to strip searches.

T-test Interpretation

Based on our findings from descriptive statistics and EDA, we came up with 5 t-tests to compare the means of different variables:

1. **Using t-test to compare the mean of age if they were strip searched**

   Comparing the mean age of individuals who were strip-searched versus those who were not, we conducted a t-test and found a p-value of 4.99×10^{-20}. This value is very small and much lower than the significance level of 0.05. Hence, we can conclude that there is a significant difference in the mean age between the two groups.

2. **Using t-test to compare the mean of age if they had actions at arrest**

   Similarly, we combined all types of actions at arrest into one variable and conducted a t-test to compare the mean age for individuals who were subject to actions at arrest versus those who were not. The t-test result revealed a p-value of 3.98×10^{-5}, indicating a significant difference in age between the two groups.

3. **Using t-test to compare the mean of age committing a crime if they are male or female**

   Furthermore, we conducted a t-test to compare the mean age of males and females who have been arrested. The p-value result was 1.98×10^{-33}, which provides sufficient evidence to reject the null hypothesis and conclude that there is a statistically significant difference in the mean age of males and females who have been arrested.

4. **Using t-test to compare the mean of age in if they were booked**

   To compare the mean age of individuals who were booked (i.e., arrested) versus those who were not booked (i.e., not arrested), we conducted a t-test. The output revealed a p-value of 0.000398, which is much lower than the significance level of 0.05. Hence, we can conclude that there is a statistically significant difference in the mean age between those who were booked and those who were not.

5. **Using t-test to compare the mean of age in different year**

   Lastly, another t-test was conducted to compare the mean age of individuals arrested in the years 2020 and the year 2021. The p-value was 1.61×10^{-17}, which is less than 0.05. Therefore, we can conclude that there is a statistically significant difference in mean age between the two groups.

Methods

In this study, we aim to investigate the potential differences in the number of strip searches conducted by the Toronto Police based on various demographic factors and the circumstances of arrest. We primarily focus on the top 5 crime occurrence categories, including "Assault," "Assault & Other crimes against persons," "Robbery & Theft," "Warrant," and "FTA/FTC/Compliance Check/Parollee." By concentrating on these categories, we can provide a higher representation and minimize bias that could impact our results. We analyze perceived

race, gender, age, and action taken at arrest separately to identify their relationships with the number of strip searches.

To address our research questions, we employed three statistical tests: power analysis, logistic regression, and analysis of covariance (ANCOVA). However, In order to conduct the statistical tests, we would first have to state the null hypothesis (H0) and the alternative hypothesis (HA). The null hypothesis represents the status quo, that there is no significant difference between groups or variables. The alternative hypothesis represents what we are trying to show, which is that there is a statistically significant difference between groups or variables. The corresponding null hypothesis and alternative hypothesis for each research question are stated below:

Hypotheses for Research Question 1 (Power Analysis):

- Null Hypothesis (H0): There is no statistically significant difference in strip search rates between different racial, gender, and age groups. Any observed differences in strip search rates are due to random variation.
- Alternative Hypothesis (HA): There is a statistically significant difference in strip search rates between different racial, gender, and age groups. The observed differences in strip search rates are not solely due to random variation.

Hypotheses for Research Question 2 (ANCOVA):

- Null Hypothesis (H0): There are no significant differences in strip search rates between different racial, gender, and age groups after controlling for covariates such as the type of action taken at arrest.
- Alternative Hypothesis (HA): There are significant differences in strip search rates between different racial, gender, and age groups after controlling for covariates such as the type of action taken at arrest.

Hypotheses for Research Question 3 (Logistics Regression)

- Null Hypothesis (H0): There is no association between race, gender, age group, and type of action taken at arrest on the likelihood of a strip search being conducted.
- Alternative Hypothesis (HA): There is an association between race, gender, age group, and type of action taken at arrest on the likelihood of a strip search being conducted.

In the following sections, we describe the methodology for each of these tests and the code used to perform them.

Power Analysis

A power analysis is a crucial step in determining the appropriate sample size required to detect an effect, if it exists, with a desired level of statistical power. In this study, a power analysis was conducted using the FTestAnovaPower module from the statsmodels library. The power analysis was performed to detect a statistically significant difference in strip search rates

between different racial, gender, and age groups. The parameters for the power analysis were set as follows:

- Effect size: 0.05 (a small effect)
- Level of significance (alpha): 0.05 (indicating a 5% chance of committing a Type I error)
- Desired power: 0.8 (reflecting an 80% chance of detecting a true effect)

Based on these parameters, the power analysis results indicated that a minimum sample size of 20,445 per group was required to achieve the desired power of 80%.

Analysis of Covariance (ANCOVA)

ANCOVA is a statistical method used to compare the means of a dependent variable between different groups, while controlling for one or more covariates. In this study, an ANCOVA test was performed using the statsmodels library to assess whether there were significant differences in strip search rates between different racial, gender, and age groups after controlling for the type of action taken at arrest. The dependent variable was whether a strip search was conducted, and the independent variables included perceived race, gender, and age group, with the type of action taken at arrest as a covariate. To prepare the dataset for ANCOVA, the following steps were taken:

- Creating dummy variables for categorical factors, such as race, gender, and age group
- Cleaning column names to ensure compatibility with the statsmodels library

The multiple regression analysis was then conducted, and the results were reported, including estimates of the adjusted group means, F-values, p-values, and effect sizes. This information allows for an assessment of the extent to which the different demographic factors are related to strip search rates, after accounting for the influence of the type of action taken at arrest. Through the combination of power analysis, logistic regression, and ANCOVA, this study aimed to identify potential biases or discrimination in the Toronto Police's strip search practices and understand the underlying reasons behind any discrepancies found.

Logistic Regression Analysis

Logistic regression is a statistical method used to analyze the relationship between a binary outcome variable (e.g., whether a strip search was conducted) and one or more predictor variables (in this case, demographic factors). In this study, the logistic regression analysis was performed using the statsmodels library. The dependent variable was whether a strip search was conducted (coded as 1 for yes and 0 for no), and the independent variables included perceived race, gender, age group, and type of action taken at arrest. To prepare the dataset for analysis, the following steps were taken:

- Creating dummy variables for categorical factors, such as race, gender, and age group
- Removing unnecessary columns from the dataset
- Splitting the data into training (80%) and test (20%) subsets

The logistic regression model was fitted to the training data, and its performance was assessed using the classification report from the sklearn library. This report includes metrics such as precision, recall, F1-score, and overall accuracy, providing a comprehensive evaluation of the model's predictive ability.

Results & Findings

A power analysis was conducted to determine the minimum sample size required for this study. The analysis was based on an effect size of 0.05, a significance level (alpha) of 0.05, and a desired statistical power of 0.8. The study considered 8 racial groups, 3 gender groups, and 7 age groups, resulting in a total of 168 unique groups. The power analysis revealed that a minimum sample size of 20,445 per group would be required to achieve the desired statistical power. The dataset used in this study, however, consists of 27,800 observations.

Given that the dataset provides a substantial number of observations for analysis, it is reasonable to assume that there should be a sufficient amount of data to draw meaningful conclusions. Nonetheless, it is important to note that the results of the study may still be affected by the smaller sample size compared to the ideal minimum. While the analysis can still be conducted, the findings may not be as robust as they would be with a larger sample size. Researchers should interpret the results with caution and consider collecting more data to achieve a higher statistical power.

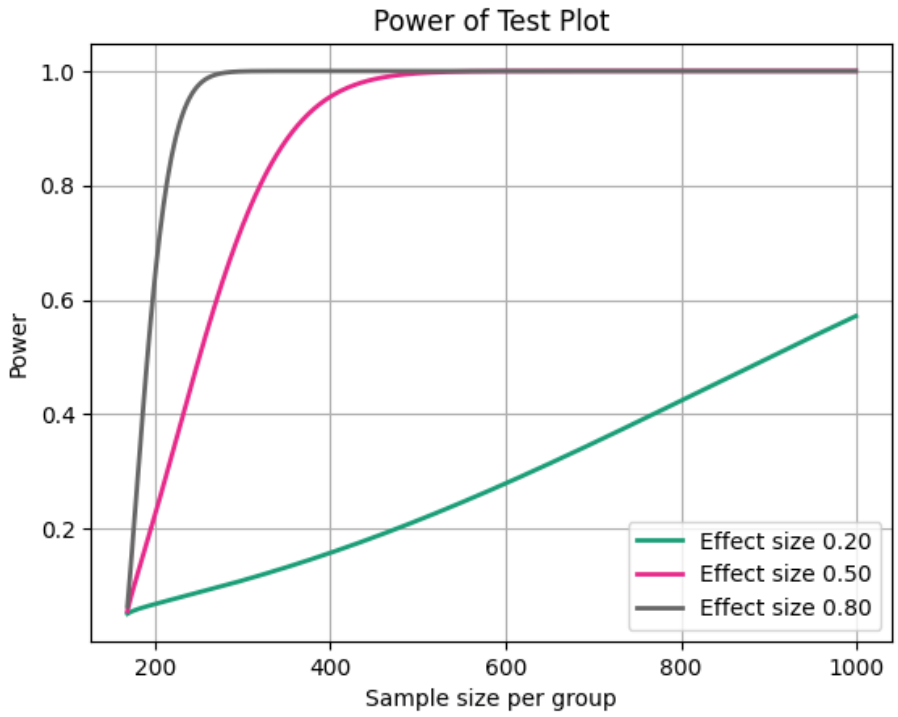

Figure 10, power of test plot.

The Power of Test Plot visually represents the relationship between the sample size per group and the statistical power of the test for different effect sizes. In this case where the effect sizes are 0.2, 0.5, and 0.8, with the sample sizes range from 5 to 1000. The plot demonstrates that as the sample size per group increases, the power of the test also increases, ultimately resulting in a higher likelihood of detecting a true effect when one exists. For smaller effect sizes, such as 0.2, a larger sample size per group is needed to achieve the desired level of power (0.8). Conversely, for larger effect sizes, such as 0.8, a smaller sample size per group is sufficient to reach the same power level. The plot also shows that the power curves are steeper for larger effect sizes, which means that even small increases in the sample size per group can lead to substantial gains in power for larger effect sizes.

From the power analysis performed earlier, a minimum sample size of 20,445 per group is required to achieve a power of 0.8 with an effect size of 0.05 and a significance level of 0.05. This indicates that a relatively large sample size per group is needed to detect small differences in strip search rates between the different racial, gender, and age groups with a high level of confidence.

On the other hand, the analysis of covariance (ANCOVA) conducted for this study aimed to determine the association between strip search rates and factors such as race, gender, and age group, while controlling for the type of action taken at arrest. The ANCOVA model showed an adjusted R-squared value of 0.033, indicating that approximately 3.3% of the variance in strip search rates can be explained by the included variables.

The results of the ANCOVA test indicate several significant associations:

1. Actions at arrest (coef = 0.1555, $p < 0.001$): A positive and significant association was found between the number of actions at arrest and the rate of strip searches.
2. Age group (various coefficients and p-values): Statistically significant differences were observed in strip search rates across age groups, with the largest differences observed between the 65 years and older group compared to the reference group (coef = -0.1120, $p < 0.001$) and the 55-64 years age group (coef = -0.0330, $p = 0.004$).
3. Race (various coefficients and p-values): Several racial groups were found to have statistically significant differences in strip search rates compared to the reference group, with East Southeast Asian (coef = -0.0629, $p < 0.001$), Indigenous (coef =

0.0615, p < 0.001), Latino (coef = -0.0571, p < 0.001), Middle Eastern (coef = -0.0630, p < 0.001), South Asian (coef = -0.0539, p < 0.001), and Unknown or Legacy (coef = -0.0273, p = 0.001) groups showing significant differences.

4. Gender (various coefficients and p-values): A significant association was found between gender and strip search rates, with males showing a higher rate of strip searches compared to the reference group (coef = 0.0315, p < 0.001). However, the 'U' (unknown) gender category did not show a statistically significant association with strip search rates (coef = -0.1728, p = 0.279).

In addition, a logistic regression analysis was performed to investigate the relationship between strip search rates and factors such as race, gender, and age group, while controlling for the type of action taken at arrest. The logistic regression model yielded a pseudo R-squared value of 0.04461, indicating that the model explains about 4.5% of the variance in strip search rates.

Several significant associations were found in the logistic regression analysis:

1. Actions at arrest (coef = 1.1178, p < 0.001): A positive and significant association was observed between the number of actions at arrest and the odds of being strip searched.
2. Age group (various coefficients and p-values): Statistically significant differences were detected in the odds of being strip searched across age groups. Notably, the 45-54 years age group (coef = -0.2555, p = 0.033), 55-64 years age group (coef =

-0.4830, p = 0.001), and 65 years and older group (coef = -3.3819, p = 0.001) showed lower odds of being strip searched compared to the reference group.

3. Race (various coefficients and p-values): Several racial groups demonstrated significant differences in the odds of being strip searched compared to the reference group, including East Southeast Asian (coef = -0.8175, $p < 0.001$), Indigenous (coef = 0.5254, $p < 0.001$), Latino (coef = -0.6123, $p < 0.001$), Middle Eastern (coef = -0.7007, $p < 0.001$), South Asian (coef = -0.5781, $p < 0.001$), and Unknown or Legacy (coef = -0.3147, p = 0.001) groups.
4. Gender (various coefficients and p-values): A significant association was found between gender and the odds of being strip searched, with males having higher odds compared to the reference group (coef = 0.3760, $p < 0.001$). The 'U' (unknown) gender category did not show a statistically significant association (coef = -14.5198, p = 0.993).

Model assessment using the test dataset revealed an overall accuracy of 0.88. However, the model showed poor performance in predicting strip searches among the minority class (class 1), with a recall and f1-score of 0.00. This indicates that the model's ability to accurately predict individuals who were strip searched is limited.

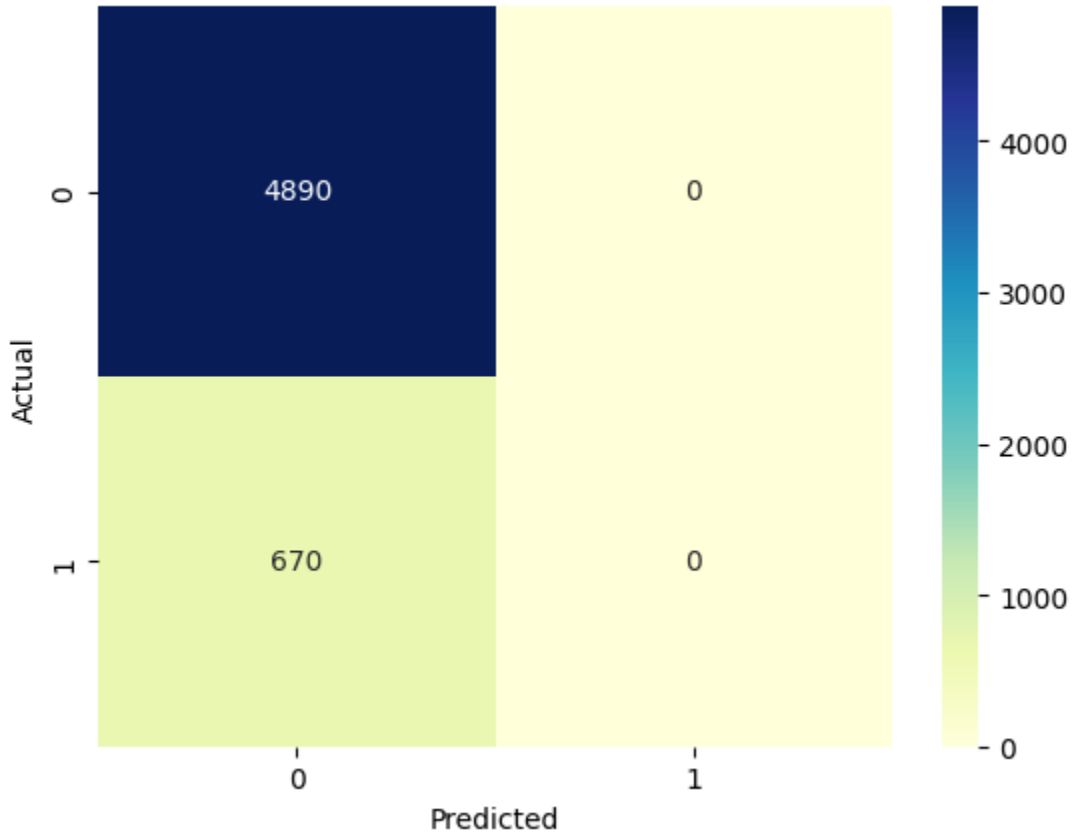

Figure 11, Confusion matrix based on the logistic regression model.

The confusion matrix provides a visual representation of the performance of the logistic regression model in predicting strip searches based on actions at arrest, perceived race, sex, and age group. In this matrix, the rows represent the actual outcomes, while the columns represent the predicted outcomes. The matrix has two rows and two columns, corresponding to the two possible outcomes: "Strip Search" (1) and "No Strip Search" (0). The top-left cell represents the number of true negatives (TN), where the model correctly predicted that no strip search would occur (4,890 cases). The top-right cell represents the false positives (FP), where the model incorrectly predicted a strip search would occur (0 cases). The bottom-left cell represents the

false negatives (FN), where the model incorrectly predicted that no strip search would occur (670 cases). The bottom-right cell represents the true positives (TP), where the model correctly predicted that a strip search would occur (0 cases). The model's accuracy, calculated as (TP + TN) / (TP + TN + FP + FN), is 88%, indicating that it correctly predicted the outcome in 88% of the test cases. However, the precision and recall for strip searches are both 0, suggesting that the model fails to predict any strip searches correctly. The confusion matrix highlights that the model may not be well-suited for predicting strip searches, despite its overall accuracy, as it fails to capture the true positives.

Overall, the logistic regression analysis suggests that strip search rates are significantly associated with the type of action taken at arrest, age group, race, and gender. However, the model's performance in predicting strip searches is limited, and further refinements may be necessary.

Conclusion

In this study, we aimed to investigate potential differences in strip search rates conducted by the Toronto Police based on various demographic factors and arrest circumstances. Our analysis was guided by three research questions.

For our first research question, we determined the minimum sample size required to detect a statistically significant difference in strip search rates between different racial, gender,

and age groups. We conducted a power analysis, which considered 8 racial groups, 3 gender groups, and 7 age groups, totalling 168 unique groups. Based on the power analysis, a minimum sample size of 20,445 per group is needed to achieve the desired level of confidence in our results. Our dataset includes 27,800 observations, which is more than the required minimum sample size of 20,445 per group. This suggests that our findings are likely to be robust, as the sample size is sufficient to detect differences in strip search rates between the different racial, gender, and age groups with a high level of confidence. The Power of Test Plot visually demonstrates the relationship between the sample size per group and the statistical power of the test for different effect sizes. It shows that as the sample size per group increases, the power of the test also increases, resulting in a higher likelihood of detecting a true effect when one exists. Smaller effect sizes require larger sample sizes to achieve the desired level of confidence, while larger effect sizes can reach the same confidence level with smaller sample sizes. In our case, the sample size per group is adequate to detect small differences in strip search rates between the different racial, gender, and age groups with a high level of confidence.

Regarding our second research question, we conducted an analysis of covariance (ANCOVA) to examine whether there were significant differences in strip search rates between different racial, gender, and age groups after controlling for covariates such as the type of action taken at arrest, which showed that the variables included in our model account for about 3.3% of the variance in strip search rates. The ANCOVA results indicate that there are indeed significant associations between strip search rates and the factors under investigation. For actions at arrest, there was a positive and significant relationship, meaning that as the number of actions at arrest

increases, so does the rate of strip searches. Age group differences were also found to be significant, with older age groups showing the largest differences in strip search rates compared to the reference group. Furthermore, our analysis revealed that race plays a role in strip search rates, as several racial groups showed significant differences when compared to the reference group. For gender, males had a higher rate of strip searches than the reference group, while the 'unknown' gender category did not exhibit a significant association with strip search rates. Overall, our findings suggest that differences in strip search rates exist between various racial, gender, and age groups, even when controlling for factors such as the type of action taken at arrest. This highlights the importance of considering demographic factors and other relevant variables when examining strip search practices and their potential impact on different segments of the population.

For the third research question, we investigated the association between race, gender, age group, and type of action taken at arrest on the likelihood of a strip search being conducted. we conducted a logistic regression analysis to examine the association between race, gender, age group, and type of action taken at arrest on the likelihood of a strip search being conducted. The model explains approximately 4.5% of the variance in strip search rates. Our analysis revealed several significant associations. First, we found a positive relationship between the number of actions at arrest and the odds of being strip searched. Second, differences in the odds of being strip searched were observed across age groups, with older age groups generally showing lower odds compared to the reference group. Third, race also appeared to play a role in the likelihood of a strip search, as several racial groups demonstrated significant differences in the odds of

being strip searched compared to the reference group. Lastly, gender was significantly associated with the odds of being strip searched, with males having higher odds compared to the reference group, while the 'unknown' gender category did not exhibit a significant association. When assessing the model's performance using the test dataset, we found an overall accuracy of 88%. However, the model showed poor performance in predicting strip searches among the minority class, with a recall and f1-score of 0. This indicates that the model's ability to accurately predict individuals who were strip-searched is limited. Overall, our logistic regression analysis suggests that the likelihood of a strip search being conducted is significantly associated with factors such as the type of action taken at arrest, age group, race, and gender. However, the model's performance in predicting strip searches is limited, and further refinements may be necessary to improve its accuracy in identifying individuals who were subject to strip searches.

In conclusion, this study investigated potential differences in strip search rates conducted by the Toronto Police based on demographic factors and arrest circumstances. Our findings, based on sufficiently large sample size, suggest that there are indeed differences in strip search rates across racial, gender, and age groups, even after controlling for factors such as the type of action taken at arrest. Furthermore, our logistic regression analysis demonstrates that the likelihood of a strip search being conducted is significantly associated with these demographic factors and the type of action taken at arrest. This study underscores the importance of considering demographic factors and other relevant variables when examining strip search practices and their potential impact on different segments of the population. By shedding light on the disparities in strip search rates among various groups, this research can help inform policy

discussions and the development of fair and effective law enforcement practices, ultimately contributing to fostering trust and transparency between law enforcement agencies and the communities they serve.

Discussion

The results of our study provide important insights into the potential differences in strip search rates between different racial, gender, and age groups after controlling for covariates such as the type of action taken at arrest. Our findings also reveal associations between race, gender, age group, and type of action taken at arrest on the likelihood of a strip search being conducted. Our analysis indicates that there are significant differences in strip search rates among different racial groups even after controlling for covariates. This suggests that certain racial groups may be disproportionately subjected to strip searches. Although the reason for this disparity cannot be definitively determined from our study, it could be attributed to factors such as implicit bias or discriminatory practices within the Toronto Police Services. Similarly, our findings reveal that gender and age also play a role in the likelihood of a strip search being conducted. Males were found to be more likely to be subjected to a strip search, which might be due to gender stereotypes or perceived threat levels associated with different genders. Age also showed a significant association, with older individuals being less likely to be subjected to strip searches compared to younger individuals. This could be because older individuals are perceived as less likely to engage in criminal activities or pose a threat to the officers.

It is important to acknowledge the limitations of our study while interpreting the results. First, our analysis is based on observational data, which prevents us from establishing causality between the variables. Second, although we controlled for a number of covariates in our analysis, there may be unmeasured confounding factors that influence the likelihood of a strip search being conducted, such as the nature of the specific arrest incident or the officers involved in the arrest. Third, the generalizability of our findings may be limited due to the focus on Toronto Police Services data; our results might not be applicable to other jurisdictions or policing agencies. Last, the confusion matrix reveals several limitations of the logistic regression model for predicting strip searches based on actions at arrest, perceived race, sex, and age group. The model appears to be biased towards the majority class ("No Strip Search"), possibly due to an imbalanced dataset, leading to a misleading accuracy of 88%. However, the precision and recall for strip searches are both 0, indicating the model's failure to predict any strip searches correctly. This highlights the importance of using multiple performance metrics, particularly when dealing with imbalanced datasets. The model may have overfit the training data, resulting in poor generalization on the test data. To address these limitations, resampling techniques can be used to balance the dataset, more suitable performance metrics can be adopted, regularization techniques can be applied to reduce overfitting, and alternative, more complex models can be experimented with to improve strip search prediction.

Despite these limitations, our study provides valuable insights into the potential disparities in strip search practices among different demographic groups and arrest circumstances. The findings can serve as a starting point for further research, as well as inform

policy discussions and potential interventions aimed at promoting equitable policing practices. Future research could explore the reasons behind the observed disparities, such as the role of implicit bias or specific policies in place that may contribute to the differences in strip search rates. Additionally, longitudinal studies could help to understand trends in strip search practices over time and evaluate the effectiveness of interventions aimed at reducing disparities.

Reference

[1] Rethinking strip searches by NSW Police (Grewcock & Sentas, 2019)
https://apo.org.au/node/254776?utm_source=APO-feed&utm_medium=RSS&utm_campaign=rss-all

[2] Gendered and Racialized Violence, Strip Searches, Sexual Assault and Abuse of Prosecutorial Power (Tanovich, 2011)
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1760142

[3] Independent Expert Assessment Report on the Toronto Police Service's Race-Based Data Collection Strategy, Phase I (Foster & Jacobs, 2022)
https://www.tps.ca/media/filer_public/1f/e7/1fe76fd7-1bde-4c4d-bd90-091a5ccba5da/2bea48b6-40a7-4f43-ba92-ce59df850c73.pdf

[4] Strip-Searching of Women in Canada: Wrongs and Rights (Psutka & Sheehy)
https://cbr.cba.org/index.php/cbr/article/view/4376

[5] Indocile Bodies: Gender Identity and Strip Searches in Canadian Criminal Law (Kirkup, 2014)
https://www.cambridge.org/core/journals/canadian-journal-of-law-and-society-la-revue-canadienne-droit-et-societe/article/indocile-bodies-gender-identity-and-strip-searches-in-canadian-criminal-law/289DBCCEC6C549D4A021CD7FACBEAB9E