Midterm Paper

Analysis on Arrest and Strip Searches

Group 37

LinLin Zhuang

YiYing Cao

Faculty of Information University of Toronto

INF2178: Experimental Design for Data Science

Professor Shion Guha

February 28, 2023

1 Introduction

1.1 Introduction

Strip searching is a technique that affords officers total control over a person, akin to the unequal power dynamics of a sexual assault (Lemke, 2022). The fact that this practice exposes a person's nakedness and violates their right to privacy makes it commonly seen as intrusive, intimidating and humiliating. Due to the potential violation of an individual's rights if not carried out properly, strip searches have received criticism from various voices and reports for their perceived targeting of specific groups. For example, strip search policies have been found to disproportionately affect women in jails. Though women account for only 8 percent of the prison population, women are on average far more likely to be strip searched than men (Dean, 2011). This disparity raises questions about gender discrimination and the potential for mistreatment of vulnerable populations.

Moreover, one major criticism of strip search policies is that they disproportionately impact specific racial groups. In recent years, police use of 'stop and search' has emerged as a key area of concern (Newburn, 2004). The disproportionate application of this power against young black men has been described as 'the most glaring example of an abuse of police powers (Newburn, 2004). Officers that engage in the practice of stop and search may be accused of using racial profiling and abusing their authority by having people stopped and searched without a warrant or other valid reason.

Another area of concern is the impact of strip searching on individuals under the age of 18. The children's commissioner for England has denounced the Metropolitan police's record on child protection after new data revealed that 650 children were strip-searched over a two-year period and the majority were found to be innocent of the suspicions against them (The Guardian, 2022). Children are traumatised by this procedure, which is not only intrusive but also may have long-term psychological implications.

This study examines the frequency of strip searches in relation to age group, perceived race, and sex among individuals in police custody. The dataset used in this study was obtained from the Toronto Police Service Public Safety Data Portal, and includes information on all arrests and strip searches. We conducted a one-way ANOVA to determine if there are significant differences in the

frequency of strip searches among different age groups. Additionally, we conducted a two-way ANOVA to investigate the combined effects of perceived race and sex on the frequency of strip searches. The results of this study can inform discussions on the potential bias and discrimination present in strip search practices in police custody.

1.2 Literature Review

Research on the use of strip-search powers in police stations has demonstrated that certain demographic groups are disproportionately targeted. Specifically, Keeton (2015) found that Afro-Caribbean arrestees were more likely to be subjected to strip searches, even after controlling for factors such as sex, age, reason for arrest, and charge. Building on this research, we aim to investigate the potential for age discrimination during strip searches.

The issue of racial discrimination in the justice system has received significant attention over the years. Data released by the Toronto Police Service on race-based strip searches reveals that Black people make up only around 10% of the city's population, yet one in every three people who were strip-searched were Black (Lemke, 2022). Similarly, concerns around sexual harassment and assault by police officers have also gained increased attention in recent years. The Toronto Police Service reported that around 60% of Toronto police employees have witnessed or been victims of sexual harassment (Woodward, 2022). Based on these findings, we hypothesize that there will be a significant difference in the frequency of strip searches conducted between different perceived races and sexes.

1.3 Research Objective and Questions

To investigate the existence of inequality in strip searches, we will utilize the Arrest and Strip Searches data from the Toronto Police Service to scrutinize the fairness of our justice system in this study. Our research aims to investigate the relationship between race, sex, age, and strip searches. To achieve this, we have formulated the following research questions and hypothesis based on our preliminary analysis of the dataset:

- Research question 1: Is there a significant difference in the frequency of strip searches conducted among different age groups at the time of arrest?

- One-way ANOVA Hypothesis:
  H0 (Null Hypothesis): All age groups have equal mean frequencies of strip searches conducted at the time of arrest.
  HA (Alternative Hypothesis): At least one age group has a different mean frequency of strip searches conducted at the time of arrest than the other age groups.

- Research question 2: Is there a significant difference in the frequency of strip searches conducted between the perceived race and sex?
- Two-way ANOVA Hypothesis:
  H0 (Null Hypothesis): All groups (perceived race and sex) have equal mean frequencies of strip searches conducted.
  HA (Alternative Hypothesis): At least one group (perceived race or sex) has a different mean frequency of strip searches conducted than the other groups.

We believe that addressing these questions will provide insights into potential biases within our justice system, particularly with regard to strip searches.

1.4 Dataset Description

Our project utilizes the Arrests_and_Strip_Searches_(RBDC-ARR-TBL-001) dataset, which provides comprehensive information on all arrests and strip searches carried out by the Toronto Police Service. The dataset contains personal information on 65,276 individuals, including their race, sex, age group at the time of arrest, and whether they were still considered youth. To process the dataset for our project, we group the independent variables by PersonID and aggregate the number of strip searches as the variable is binary. This results in a dataset with 41,154 unique individuals who have been arrested, excluding cases of multiple arrests. Additionally, the dataset includes the month and year of the arrest, the crime they were convicted of, the location of arrest, and action at arrest. The dataset could be found through the following link: https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/about.

1.5 Missing Values

The dataset contained some missing value. For some arrests, the location could not be geo-coded or the arrest took place outside of City of Toronto boundaries in other jurisdictions; these are indicated by XX (Toronto Police Service, 2022). Furthermore, column that contain missing values includes “Arrest ID”, “Perceived Race”, “Age group at arrest”, “Occurrence Category”, “Search Reason Cause Injury”, “Search Reason Assist Escape”, “Search Reason Possess Weapons”, “Search Reason Possess Evidence”, and “Items Found”.

1.6 Dependent variable

The dependent variable chosen for this study is "StripSearch", which indicates whether an individual has undergone a strip search or not. This variable is binary, with "Yes" represented by "1" and "No" represented by "0". However, since ANOVA requires a continuous dependent variable, we aggregated the number of strip searches by grouping the independent variables by PersonID. This approach allowed us to transform the binary StripSearch variable into a continuous variable, which we used as our dependent variable in the analysis.

1.7 Independent variable

The independent variables selected for this study are "Perceived race" and "Sex" for two-way ANOVA, and "Age group at arrest" for one-way ANOVA. The dataset includes eight "Perceived races," including Black, East/Southeast Asian, Indigenous, Latino, Middle-Eastern, South Asian, Unknown or Legacy, and White. In addition, there are three "Sex" categories: Male, Female, and Unknown. As a result, we have filtered out the Unknown for clarity. Lastly, the “Age group at arrest” were separated into eight categories, but there are some redundant so we reorganized the age categories. The updated “Age group at arrest” includes 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 and older.

2 EDA

2.1 Descriptive Statistics

Table 1: Demographic Characteristic

Gender	**Female**	**Male**	**Total**
Counts	N = 8309	N = 32845	N = 41154
Age group			
Aged 17 years and under	552 (6.64%)	1657 (5.04%)	2209 (5.37%)
Aged 18 to 24 years	1521 (18.31%)	5525 (16.82%)	7046 (17.12%)
Aged 25 to 34 years	2617 (31.50%)	9933 (30.24%)	12550 (30.50%)
Aged 35 to 44 years	1845 (22.20%)	7520 (22.90%)	9365 (22.76%)
Aged 45 to 54 years	1056 (12.71%)	4652 (14.16%)	5708 (13.87%)
Aged 55 to 64 years	534 (6.43%)	2652 (8.07%)	3186 (7.74%)
65 years or older	184 (2.21%)	906 (2.76%)	1090 (2.65%)
Race			
Black	1955 (23.53%)	8546 (26.02%)	10501 (25.52%)
East/Southeast Asian	645 (7.76%)	2811 (8.56%)	3456 (8.40%)
Indigenous	309 (3.72%)	527 (1.60%)	836 (2.03%)
Latino	206 (2.48%)	1099 (3.35%)	1305 (3.17%)
Middle-Eastern	295 (3.55%)	2083 (6.34%)	2378 (5.78%)
South Asian	437 (5.26%)	2417 (7.36%)	2854 (6.93%)
Unknown or Legacy	810 (9.75%)	3579 (10.90%)	4389 (10.66%)
White	3652 (43.95%)	11783 (35.87%)	15435 (37.51%)

Youth			
Youth (aged 17 and younger)	552 (6.64%)	1657 (5.05%)	2209 (5.37%)
Not a youth	7757 (93.36%)	31188 (94.95%)	38945 (94.63%)

Table 1 provides a comprehensive breakdown of the study participants based on their demographic characteristics and the dependent variable. A total of 41,154 participants were included in the study, with 8,309 being female and 32,845 being male. The age groups were divided into seven categories, with the majority falling between 25 to 34 years (30.50%). The perceived race was grouped into eight categories, with White being the largest group (37.51%), followed by Black (25.52%), and Indigenous being the smallest group (2.03%). The table also shows that 5.37% of the participants were aged 17 and under, while the remaining 94.63% were not considered youth. Overall, this table provides a clear overview of the distribution of participants in the study based on their gender, age, race, and youth status.

Figure 1: Distribution of Age Group by Sex

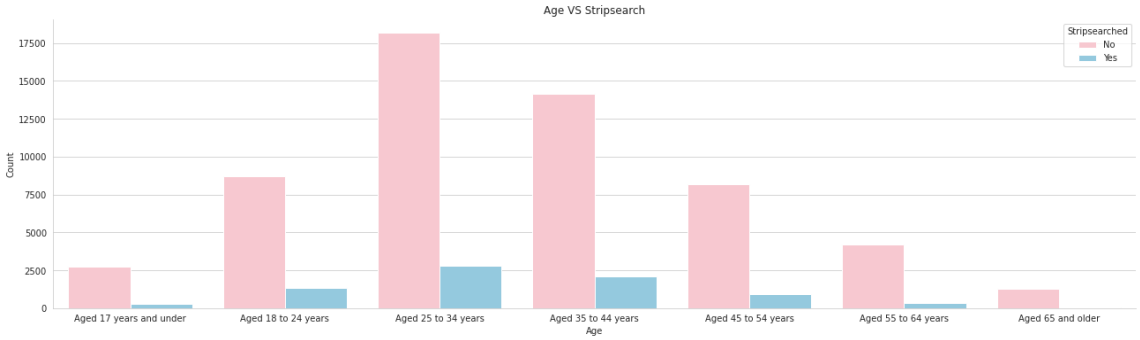


Figure 1 provides a visual representation of the frequency of strip searches by age group. The graph shows that the age group with the highest frequency of strip searches is aged 25 to 34 years, followed by the age group of 35 to 44 years. In contrast, the age group of 65 and older had the lowest frequency of strip searches. This pattern is consistent with the age distribution shown in Table 1. Therefore, the graph reinforces the findings from the table and provides a clear illustration of the relationship between age and strip searches.

Figure 2: Distribution of Perceived Race by Sex

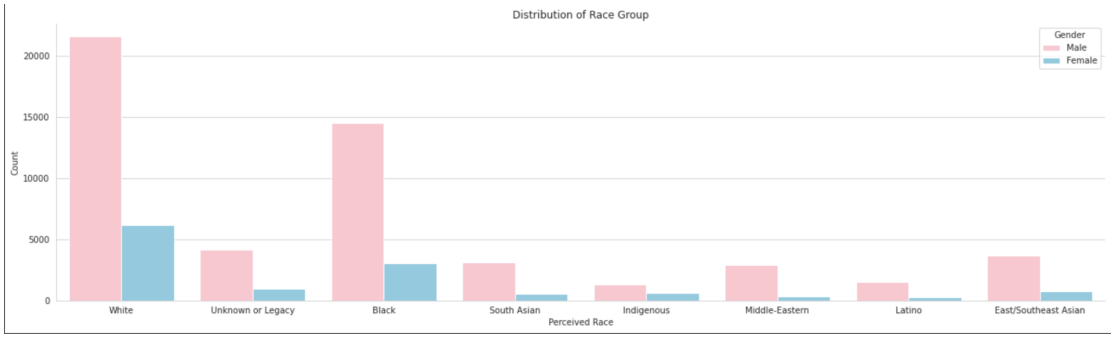


Figure 2 displays the frequency of strip searches by race, and it is consistent with the distribution of perceived race shown in Table 1. According to the graph, White people undergo strip searches the most frequently, with Black people coming in second. The group of Indigenous people experiences the fewest strip searches.

Figure 3: Distribution of Stripsearch by Sex

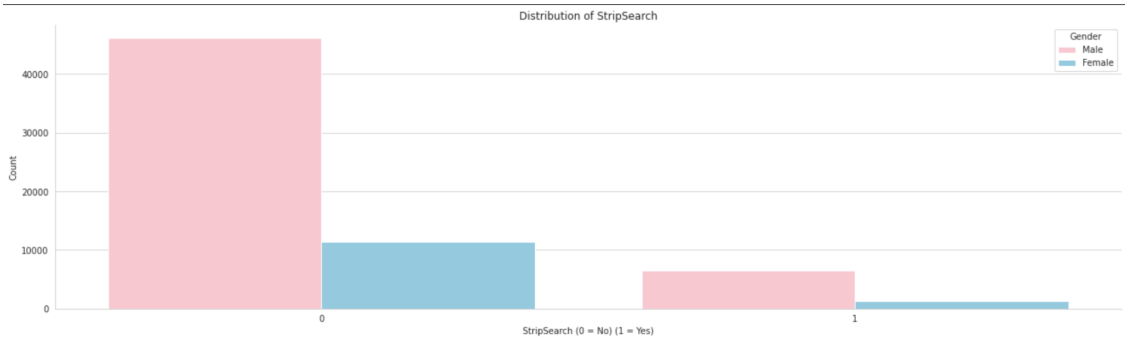


Figure 3 illustrates the gender distribution of the study participants who were arrested and whether or not they were subjected to a strip search. The graph highlights a significantly larger number of male participants who were arrested than female participants, with male arrests being twice as large as female arrests. Additionally, the number of males who were strip-searched was higher than females who were strip-searched, and the number of males who were not strip-searched was also higher than females who were not strip-searched.

Figure 4: Distribution of Arrest under 18 years of age by Sex

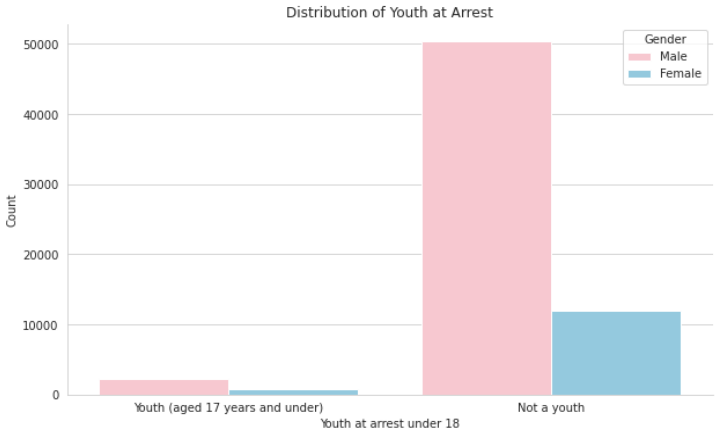


Figure 4 shows the gender distribution of whether the individual is youth or adult at the time of arrest. The graph stated that individuals who are classified as adults have a significantly higher number of arrests than those who are classified as youth for both males and females at the time of arrest.

Furthermore, we have generated additional bar plots to further demonstrate the correlation between those attributes and stripsearch. We did not include the attribute of Gender and Stripsearch as it has already shown in Figure 3.

Figure 5: Distribution of Total Strip Search by Sex and Age Group, Perceived Race, and Youth

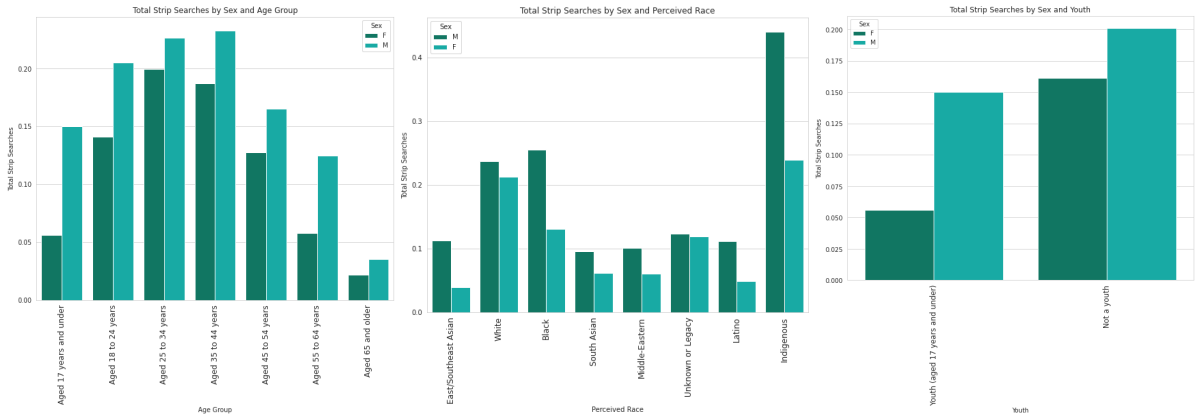


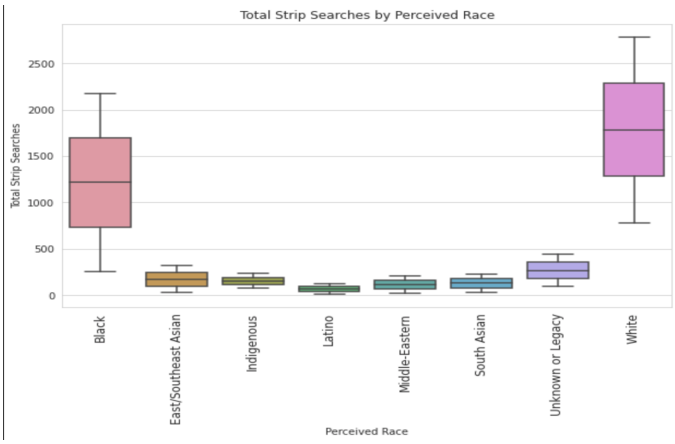
Figure 5 was created using the dataset that was grouped by PersonID. It's important to note that this grouping means that if the same individual was strip searched or arrested multiple times, it could inflate certain categories and make them appear higher than in the non-grouped dataset.

The first distribution (Figure 5) shows the total strip searches by sex and age group at arrest. From this distribution, we can state that the age group between 35 and 54 have the highest number of strip searches for males and the age group between 65 and older have the lowest number of strip searches for male. On the other hand, the age group between 25 to 34 have the highest number of strip searches for females and the age group between 65 and older have the lowest number of strip searches for females.

The second distribution (Figure 5) shows the total strip searches by sex and perceived race. As we can observe, Indigenous males have the highest number of strip searches, and South Asian males have the lowest number of strip searches. This is in total contrast to Figure 2, where Indigenous people are shown to be the smallest group. Indigenous females also have the highest number of strip searches whereas East/Southeast Asian females have the lowest number of strip searches.

The third distribution in Figure 5 shows the total strip searches by sex and whether the individual is youth at time of arrest. The distribution mentions that both adult males and females have a higher number of strip searches than youth who are 17 years and older at the time of arrest.

Figure 6: Box plot of Total Strip Searches by Perceived Race



Furthermore, we created a box plot to investigate the correlation between perceived race and strip search, as perceived race is one of our key attributes for further analysis. Figure 6 displays the total number of strip searches conducted in each race. The box plot shapes indicate that perceived race is normally distributed. As shown in the plot, White people have the highest

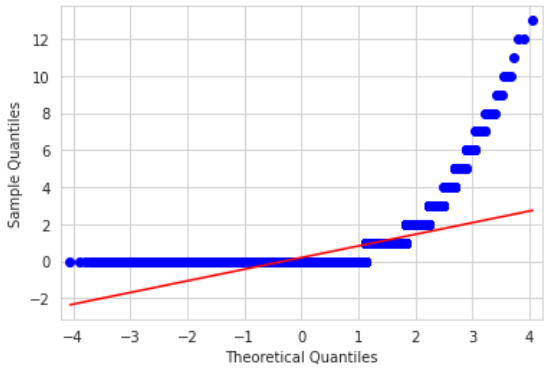
number of strip searches with a median of around 1750 total strip searches, whereas Latinos have the lowest number of strip searches.

However, to determine which race is more likely to be strip searched, we need to investigate the percentage of strip searches within its entire racial population. Although Figure 6 shows that White people have the highest number of strip searches, Table 2 reveals that Indigenous people have the highest percentage (15.82%) of experiencing strip searches. Table 2 displays the percentage of both strip searched and non-strip searched individuals in each perceived race. We can observe that the percentage of black people who were arrested and experienced strip searches is 13.89%, which is the second highest among all other races. In contrast, Middle-Eastern people have the lowest percentage (7.04%) of experiencing strip search among all other races.

	No	Yes
Perceived_Race		
Black	86.11%	13.89%
East/Southeast Asian	92.28%	7.72%
Indigenous	84.18%	15.82%
Latino	92.53%	7.47%
Middle-Eastern	92.96%	7.04%
South Asian	92.89%	7.11%
Unknown or Legacy	89.39%	10.61%
White	87.13%	12.87%

Table 2: Percentage of strip search in each Perceived Race

Figure 7 Normal Q-Q plot



We generated the Normal Q-Q plot to test whether the residuals are normally distributed. As we can observe, only some of the residual points followed the straight dashed line. The residual

points at the two ends are very off from the straight dashed line, especially the right end. As a result, we can conclude that the residuals are not normally distributed.

2.2 T-test

Gender (Male and Female) and Strip Search (Two-sided T-test)

• H0 (Null Hypothesis): The average number of strip searches for males is equal to the average number of strip searches for females.

• Ha (Alternative Hypothesis): The average number of strip searches for males is different from the average number of strip searches for females.

Our result indicates that the t-statistic is 5.88 and the p-value is 4.16e-09. Since the p-value is less than the chosen significance level of alpha = 0.05, with 95% CI [0.03, 0.06], we reject the null hypothesis and conclude that the average number of strip searches for males is different from the average number of strip searches for females.

Gender (Male and Female) and Strip Search (One-sided, greater than test)

• H0 (Null Hypothesis): The average number of strip searches for males and females is equal to or less than the average number of strip searches for females.

• Ha (Alternative Hypothesis): The average number of strip searches for males is greater than the average number of strip searches for females.

Our result indicates that the t-statistic is 5.88 and the p-value is 2.08e-09. Since the p-value is less than the chosen significance level of alpha = 0.05, we reject the null hypothesis and conclude that the average number of strip searches for males is greater than the average number of strip searches for females.

Gender (Male and Female) and Strip Search (One-sided, less than test)

• H0 (Null Hypothesis): The average number of strip searches for males and females is equal to or greater than the average number of strip searches for females.

• Ha (Alternative Hypothesis): The average number of strip searches for males is less than the average number of strip searches for females.

Our result indicates that the t-statistic is 5.88 and the p-value is 0.99. Since the p-value is greater than the chosen significance level of alpha = 0.05, we do not have enough evidence to reject the

null hypothesis and conclude that the average number of strip searches for males and females is equal to or greater than the average number of strip searches for females.

Race(Black and White) and Strip Search (Two-sided T-test)

• H0 (Null Hypothesis): The average number of strip searches performed on Black people is equal to the average number of strip searches performed on White people.

• Ha (Alternative Hypothesis): The average number of strip searches performed on Black people is different from the average number of strip searches performed on White people.

Our result indicates that the t-statistic is 0.14 and the p-value is 0.89. Since the p-value is greater than the chosen significance level of alpha = 0.05, with 95% CI [-0.02, 0.02], we do not have enough evidence to reject the null hypothesis and conclude that the average number of strip searches performed on Black people is equal to the average number of strip searches performed on White people.

Race(Black and White) and Strip Search (One-sided, greater than test)

• H0 (Null Hypothesis): The average number of strip searches performed on Black people is equal to or less than the average number of strip searches performed on White people.

• Ha (Alternative Hypothesis): The average number of strip searches performed on Black people is greater than the average number of strip searches performed on White people.

Our result indicates that the t-statistic is 0.14 and the p-value is 0.44. Since the p-value is greater than the chosen significance level of alpha = 0.05, we do not have enough evidence to reject the null hypothesis and conclude that the average number of strip searches performed on Black people is equal to or less than the average number of strip searches performed on White people.

Race(Black and White) and Strip Search (One-sided, less than test)

• H0 (Null Hypothesis): The average number of strip searches performed on Black people is equal to or greater than the average number of strip searches performed on White people.

• Ha (Alternative Hypothesis): The average number of strip searches performed on Black people is less than the average number of strip searches performed on White people.

Our result indicates that the t-statistic is 0.14 and the p-value is 0.56. Since the p-value is greater than the chosen significance level of alpha = 0.05, we do not have enough evidence to reject the

null hypothesis and conclude that the average number of strip searches performed on Black people is equal to or greater than the average number of strip searches performed on White people.

Youth at Arerst (Youth and Not Youth) and Strip Search (Two-sided T-test)

• H0 (Null Hypothesis): The average number of strip searches for individuals who are youth is equal to the average number of individuals who are not youth at the time of arrest.

• Ha (Alternative Hypothesis): The average number of strip searches for individuals who are youth is different from the average number of individuals who are not youth at the time of arrest.

Our result indicates that the t-statistic is -5.03and the p-value is 4.80e-07. Since the p-value is less than the chosen significance level of alpha = 0.05, with 95% CI [-0.1, - 0.06], we reject the null hypothesis and conclude that the average number of strip searches for individuals who are youth is different from the average number of individuals who are not youth at the time of arrest.

Youth at Arerst (Youth and Not Youth) and Strip Search (One-sided, greater than test)

• H0 (Null Hypothesis): The average number of strip searches performed on individuals who are youth is equal to or less than the average number of strip searches performed on individuals who are not youth at the time of arrest.

• Ha (Alternative Hypothesis): The average number of strip searches performed on individuals who are youth is greater than the average number of strip searches performed on individuals who are not youth at the time of arrest.

Our result indicates that the t-statistic is -5.03 and the p-value is 0.99. Since the p-value is greater than the chosen significance level of alpha = 0.05, we do not have enough evidence to reject the null hypothesis and conclude that the average number of strip searches performed on individuals who are youth is equal to or less than the average number of strip searches performed on individuals who are not youth at the time of arrest.

Youth at Arerst (Youth and Not Youth) and Strip Search (One-sided, less than test)

• H0 (Null Hypothesis): The average number of strip searches performed on individuals who are youth is equal to or greater than the average number of strip searches performed on individuals who are not youth at the time of arrest.

• Ha (Alternative Hypothesis): The average number of strip searches performed on individuals who are youth is less than the average number of strip searches performed on individuals who are not youth at the time of arrest.

Our result indicates that the t-statistic is -5.03 and the p-value is 2.40. Since the p-value is greater than the chosen significance level of alpha = 0.05, we do not have enough evidence to reject the null hypothesis and conclude that the average number of strip searches performed on individuals who are youth is equal to or greater than the average number of strip searches performed on individuals who are not youth at the time of arrest.

Age at Arrest (Aged 18 to 24 & Aged 35 to 44) and Strip Search (Two-sided T-test)

• H0 (Null Hypothesis): The average number of strip searches for individuals aged 18 to 24 years is equal to the average number of individuals aged 35 to 44 years at the time of arrest.

• Ha (Alternative Hypothesis): The average number of strip searches for individuals aged 18 to 24 years is different from the average number of individuals aged 35 to 44 years at the time of arrest.

Our result indicates that the t-statistic is -3.25 and the p-value is 0.001. Since the p-value is less than the chosen significance level of alpha = 0.05, with 95% CI [-0.06, -0.02], we reject the null hypothesis and conclude that the average number of strip searches for individuals aged 18 to 24 years is different to the average number of individuals aged 35 to 44 years at the time of arrest.

Age at Arrest (Aged 18 to 24 & Aged 35 to 44) and Strip Search (One-sided, greater than test)

• H0 (Null Hypothesis): The average number of strip searches performed on individuals aged 18 to 24 years is equal to or less than the average number of strip searches performed on individuals aged 35 to 44 years at the time of arrest.

• Ha (Alternative Hypothesis): The average number of strip searches performed on individuals 18 to 24 years is greater than the average number of strip searches performed on individuals aged 35 to 44 years at the time of arrest.

Our result indicates that the t-statistic is -3.25 and the p-value is 0.99. Since the p-value is greater than the chosen significance level of alpha = 0.05, we do not have enough evidence to reject the null hypothesis and conclude that the average number of strip searches performed on individuals

aged 18 to 24 years is equal to or less than the average number of strip searches performed on individuals aged 35 to 44 years at the time of arrest.

Age at Arrest (Aged 18 to 24 & Aged 35 to 44) and Strip Search (One-sided, less than test)

• H0 (Null Hypothesis): The average number of strip searches performed on individuals aged 18 to 24 years is equal to or greater than the average number of strip searches performed on individuals aged 35 to 44 years at the time of arrest.

• Ha (Alternative Hypothesis): The average number of strip searches performed on individuals aged 18 to 24 years is less than the average number of strip searches performed on individuals aged 35 to 44 years at the time of arrest.

Our result indicates that the t-statistic is -3.25 and the p-value is 0.0006. Since the p-value is less than the chosen significance level of alpha = 0.05, we will reject the null hypothesis and conclude that the average number of strip searches performed on individuals aged 18 to 24 years is less than the average number of strip searches performed on individuals aged 35 to 44 years at the time of arrest.

Race at Arrest (Black and Non-black) and Strip Search (Two-sided T-test)

• H0 (Null Hypothesis): The average number of strip searches for individuals who are black is equal to the average number of individuals who are non-black at the time of arrest.

• Ha (Alternative Hypothesis): The average number of strip searches for individuals who are black is different from the average number of individuals who are non-black at the time of arrest.

Our result indicates that the t-statistic is 8.33 and the p-value is 8.51e-17. Since the p-value is less than the chosen significance level of alpha = 0.05, with 95% CI [0.04, 0.07], we will reject the null hypothesis and conclude that the average number of strip searches for individuals who are black is different from the average number of individuals who are non-black at the time of arrest.

Race at Arrest (Black and Other Race) and Strip Search (One-sided, greater than test)

• H0 (Null Hypothesis): The average number of strip searches performed on individuals who are black is equal to or less than the average number of strip searches performed on individuals who are non-black at the time of arrest.

• Ha (Alternative Hypothesis): The average number of strip searches performed on individuals who are black is greater than the average number of strip searches performed on individuals who are non-black at the time of arrest.

Our result indicates that the t-statistic is 8.33 and the p-value is 4.26e-17. Since the p-value is less than the chosen significance level of alpha = 0.05, we will reject the null hypothesis and conclude that the average number of strip searches performed on individuals who are black is greater than the average number of strip searches performed on individuals who are non-black at the time of arrest.

Race at Arrest (Black and Other Race) and Strip Search (One-sided, less than test)

• H0 (Null Hypothesis): The average number of strip searches performed on individuals who are black is equal to or greater than the average number of strip searches performed on individuals who are non-black at the time of arrest.

• Ha (Alternative Hypothesis): The average number of strip searches performed on individuals who are black is less than the average number of strip searches performed on individuals who are non-black at the time of arrest.

Our result indicates that the t-statistic is 8.33 and the p-value is 0.99. Since the p-value is greater than the chosen significance level of alpha = 0.05, we do not have enough evidence to reject the null hypothesis and conclude that the average number of strip searches performed on individuals who are black is equal to or greater than the average number of strip searches performed on individuals who are non-black at the time of arrest.

3 Method

3.1 One way Anova (OLS)

We have conducted a one-way ANOVA to find out whether there is a statistically significant difference in the mean strip search frequencies between the different age groups at the time of arrest. In this case, the hypothesis will be the following:

• H0 (Null Hypothesis): All age groups have equal mean frequencies of strip searches conducted at the time of arrest.

• HA (Alternative Hypothesis): At least one age group has a different mean frequency of strip searches conducted at the time of arrest than the other age groups.

	sum_sq	df	F	PR(>F)
C(Age_group__at_arrest_)	97.043627	6.0	35.496419	4.343393e-43
Residual	17391.243210	38168.0	NaN	NaN

Table 3: One-Way ANOVA

In Table 3, our result indicates that the F-statistic is 35.49 and the p-value is 4.34e-43. Since the p-value is less than the chosen significance level of alpha = 0.05, we reject the null hypothesis and conclude that there is a significant difference in the mean strip search frequencies between the different age groups at the time of arrest.

3.2 Tukey's HSD

We performed a one-way ANOVA to determine if there is a significant difference in strip search frequency across age groups at the time of arrest. However, ANOVA does not provide information about which specific age groups are significantly different from each other. To address this, we conducted Tukey's HSD test, which is a post-hoc test that allows us to compare all possible pairs of age groups and determine if there are significant differences between them.

Table 4 summarizes the results of Tukey's HSD test. Each row represents a pair of age groups, and the columns provide information on the mean difference between the groups (meandiff), as well as the lower and upper bounds of the confidence interval for the difference. If the value under the "reject" column is "True," it means that we can reject the null hypothesis of no difference between the means of the two age groups at the given significance level (in this case, 0.05).

As an example, the row that compares the age group of 17 years and under to the 18-24 age group. The mean difference between these groups is 0.0718, and the confidence interval for the difference ranges from 0.0218 to 0.1219. The "True" value in the reject column indicates that there is a significant difference in the mean strip search frequency between these two age groups. On the other hand, the row that compares the age group of 25-34 to the 35-44 age group shows a mean difference of 0.0054, with a confidence interval ranging from -0.023 to 0.0337. However, the "False" value in the reject column indicates that the null hypothesis of no difference between the means cannot be rejected at the 0.05 significance level.

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Aged 17 years and under	Aged 18 to 24 years	0.0718	0.001	0.0218	0.1219	True
Aged 17 years and under	Aged 25 to 34 years	0.1049	0.001	0.0575	0.1523	True
Aged 17 years and under	Aged 35 to 44 years	0.1103	0.001	0.0616	0.1589	True
Aged 17 years and under	Aged 45 to 54 years	0.0359	0.3816	-0.0156	0.0873	False
Aged 17 years and under	Aged 55 to 64 years	-0.0148	0.9	-0.0715	0.0418	False
Aged 17 years and under	Aged 65 and older	-0.1005	0.0015	-0.1756	-0.0254	True
Aged 18 to 24 years	Aged 25 to 34 years	0.033	0.026	0.0023	0.0638	True
Aged 18 to 24 years	Aged 35 to 44 years	0.0384	0.0095	0.0057	0.0711	True
Aged 18 to 24 years	Aged 45 to 54 years	-0.036	0.0595	-0.0728	0.0008	False
Aged 18 to 24 years	Aged 55 to 64 years	-0.0867	0.001	-0.1304	-0.0429	True
Aged 18 to 24 years	Aged 65 and older	-0.1723	0.001	-0.2382	-0.1065	True
Aged 25 to 34 years	Aged 35 to 44 years	0.0054	0.9	-0.023	0.0337	False
Aged 25 to 34 years	Aged 45 to 54 years	-0.069	0.001	-0.102	-0.0361	True
Aged 25 to 34 years	Aged 55 to 64 years	-0.1197	0.001	-0.1604	-0.0791	True
Aged 25 to 34 years	Aged 65 and older	-0.2054	0.001	-0.2692	-0.1415	True
Aged 35 to 44 years	Aged 45 to 54 years	-0.0744	0.001	-0.1092	-0.0397	True
Aged 35 to 44 years	Aged 55 to 64 years	-0.1251	0.001	-0.1672	-0.083	True
Aged 35 to 44 years	Aged 65 and older	-0.2108	0.001	-0.2755	-0.146	True
Aged 45 to 54 years	Aged 55 to 64 years	-0.0507	0.017	-0.096	-0.0054	True
Aged 45 to 54 years	Aged 65 and older	-0.1363	0.001	-0.2033	-0.0694	True
Aged 55 to 64 years	Aged 65 and older	-0.0857	0.0069	-0.1567	-0.0147	True

Table 4: Tukey HSD (age)

3.3 Two-Way ANOVA

We have also conducted a two-way ANOVA to find out whether there is a statistically significant difference in the mean strip search frequencies between the perceived race and sex at the time of arrest. In this case, the hypothesis will be the following:

• H0 (Null Hypothesis): All groups (perceived race and sex) have equal mean frequencies of strip searches conducted.

• HA (Alternative Hypothesis): At least one group (perceived race or sex) has a different mean frequency of strip searches conducted than the other groups.

	sum_sq	df	F	PR(>F)
C(Perceived_Race)	198.584238	7.0	72.893792	2.396116e-105
C(Sex)	23.059903	1.0	59.251763	1.418850e-14
C(Perceived_Race):C(Sex)	17.456880	7.0	6.407851	1.473674e-07
Residual	15714.126369	40377.0	NaN	NaN

Table 5: Two-Way ANOVA

In Table 5, our result indicates that the F-statistic is 6.41 and the p-value is 1.47e-07. Since the p-value is less than the chosen significance level of alpha = 0.05, we reject the null hypothesis and conclude that at least one group (perceived race or sex) has a different mean frequency of strip searches conducted than the other groups

Table 6 of the Tukey's HSD result shows that there is a significant difference in the mean strip search frequency between the black group and other groups. However, there is no significant difference between black and white groups. Regarding the interaction effect shown in Table 7, we concatenated the Perceived Race and Sex variables to compare the mean strip search frequencies between all possible combinations of Perceived Race and Sex. The table shows a significant difference in the mean strip search frequency between black females and both white females and males. Additionally, there is a significant difference in the mean strip search frequency between black males and white females, but no significant difference between black males and white males.

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Black	East/Southeast Asian	-0.1382	0.001	-0.1756	-0.1009	True
Black	Indigenous	0.1373	0.001	0.0684	0.2061	True
Black	Latino	-0.1349	0.001	-0.1909	-0.0788	True
Black	Middle-Eastern	-0.1406	0.001	-0.184	-0.0973	True
Black	South Asian	-0.147	0.001	-0.1872	-0.1068	True
Black	Unknown or Legacy	-0.1151	0.001	-0.1493	-0.0808	True
Black	White	-0.0013	0.9	-0.0255	0.023	False
East/Southeast Asian	Indigenous	0.2755	0.001	0.2018	0.3492	True
East/Southeast Asian	Latino	0.0034	0.9	-0.0586	0.0653	False
East/Southeast Asian	Middle-Eastern	-0.0024	0.9	-0.0531	0.0483	False
East/Southeast Asian	South Asian	-0.0087	0.9	-0.0568	0.0394	False
East/Southeast Asian	Unknown or Legacy	0.0232	0.7071	-0.02	0.0664	False
East/Southeast Asian	White	0.137	0.001	0.1011	0.1728	True
Indigenous	Latino	-0.2721	0.001	-0.3569	-0.1873	True
Indigenous	Middle-Eastern	-0.2779	0.001	-0.3548	-0.201	True
Indigenous	South Asian	-0.2842	0.001	-0.3594	-0.209	True
Indigenous	Unknown or Legacy	-0.2523	0.001	-0.3245	-0.1801	True
Indigenous	White	-0.1385	0.001	-0.2066	-0.0705	True
Latino	Middle-Eastern	-0.0058	0.9	-0.0715	0.0599	False
Latino	South Asian	-0.0121	0.9	-0.0758	0.0516	False
Latino	Unknown or Legacy	0.0198	0.9	-0.0403	0.0799	False
Latino	White	0.1336	0.001	0.0785	0.1886	True
Middle-Eastern	South Asian	-0.0063	0.9	-0.0592	0.0465	False
Middle-Eastern	Unknown or Legacy	0.0256	0.7226	-0.0229	0.074	False
Middle-Eastern	White	0.1394	0.001	0.0974	0.1814	True
South Asian	Unknown or Legacy	0.0319	0.4062	-0.0138	0.0776	False
South Asian	White	0.1457	0.001	0.1069	0.1845	True
Unknown or Legacy	White	0.1138	0.001	0.0812	0.1463	True

Table 6: Tukey's HSD (race)

```
Multiple Comparison of Means - Tukey HSD, FWER=0.05
=================================================
group1 group2 meandiff p-adj lower  upper  reject
-------------------------------------------------
     F      M   0.0457 0.001 0.0305 0.0609   True
-------------------------------------------------
```

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
BlackF	BlackM	0.1293	0.001	0.0751	0.1834	True
BlackF	East/Southeast AsianF	-0.094	0.0721	-0.1913	0.0033	False
BlackF	East/Southeast AsianM	-0.0192	0.9	-0.0826	0.0443	False
BlackF	IndigenousF	0.1115	0.2197	-0.0206	0.2436	False
BlackF	IndigenousM	0.3195	0.001	0.2133	0.4257	True
BlackF	LatinoF	-0.0835	0.9	-0.2412	0.0743	False
BlackF	LatinoM	-0.0198	0.9	-0.1011	0.0615	False
BlackF	Middle-EasternF	-0.0713	0.9	-0.2054	0.0628	False
BlackF	Middle-EasternM	-0.0306	0.9	-0.0984	0.0373	False
BlackF	South AsianF	-0.0708	0.7091	-0.1842	0.0426	False
BlackF	South AsianM	-0.0367	0.8506	-0.1022	0.0288	False
BlackF	Unknown or LegacyF	-0.0141	0.9	-0.1037	0.0755	False
BlackF	Unknown or LegacyM	-0.0091	0.9	-0.0697	0.0514	False
BlackF	WhiteF	0.0846	0.001	0.0242	0.1451	True
BlackF	WhiteM	0.1096	0.001	0.057	0.1623	True
BlackM	East/Southeast AsianF	-0.2232	0.001	-0.3106	-0.1359	True
BlackM	East/Southeast AsianM	-0.1484	0.001	-0.1952	-0.1016	True
BlackM	IndigenousF	-0.0178	0.9	-0.1428	0.1072	False
BlackM	IndigenousM	0.1902	0.001	0.093	0.2875	True
BlackM	LatinoF	-0.2127	0.001	-0.3646	-0.0609	True
BlackM	LatinoM	-0.149	0.001	-0.2182	-0.0799	True
BlackM	Middle-EasternF	-0.2006	0.001	-0.3276	-0.0735	True
BlackM	Middle-EasternM	-0.1598	0.001	-0.2125	-0.1071	True
BlackM	South AsianF	-0.2001	0.001	-0.3051	-0.095	True
BlackM	South AsianM	-0.166	0.001	-0.2155	-0.1164	True
BlackM	Unknown or LegacyF	-0.1433	0.001	-0.2221	-0.0646	True
BlackM	Unknown or LegacyM	-0.1384	0.001	-0.1812	-0.0956	True
BlackM	WhiteF	-0.0446	0.0304	-0.0873	-0.0019	True
BlackM	WhiteM	-0.0196	0.6812	-0.0504	0.0112	False

Table 7: Tukey's HSD (race & sex)

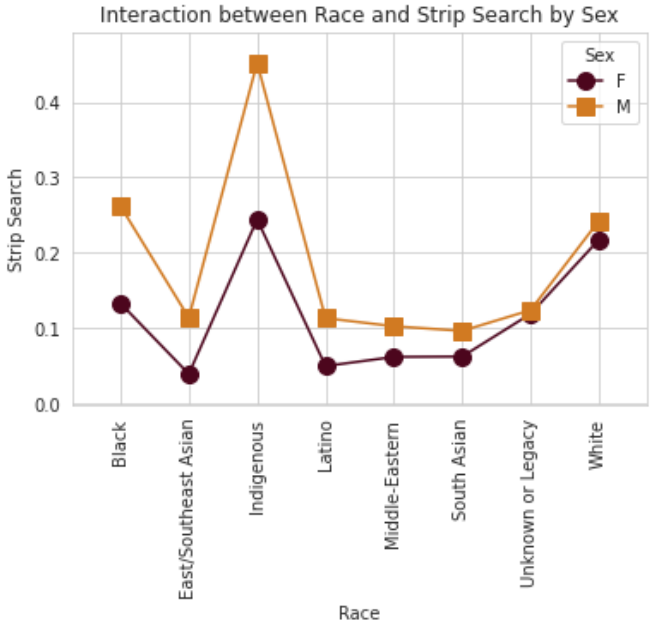
4 Results & Finding

We have conducted a thorough analysis to address our two research questions. Firstly, we used a one-way ANOVA to investigate if there was a significant difference in mean strip search frequencies among different age groups. The obtained p-value was less than 0.05, leading us to reject the null hypothesis and conclude that at least one age group has a different mean frequency of strip searches at the time of arrest than the others. We used Tukey's HSD test to determine there is a significant difference between the age group of 17 and under and some other age groups (e.g. age 35 - 44). Our analysis confirmed our hypothesis for the first research question. Additionally, we found that the difference between male youth and female youth who were strip searched is significantly higher than the difference between adult males and adult females who were strip searched. Secondly, we conducted a two-way ANOVA to investigate if there was a significant difference in mean strip search frequencies between the perceived race and sex at the time of arrest. The obtained p-value was less than 0.05, leading to the result that there is a significant difference in the mean strip search frequencies between the perceived race and sex at the time of arrest. The Tukey's HSD shows us that there is no significant difference between black and white groups, but there is a significant difference in the mean strip search frequency between black females and both white females and males. Furthermore, we observed that Indigenous people have the highest

likelihood of experiencing strip searches at 15.8%, which is the highest among all perceived races. Moreover, our plot indicates that Indigenous males are most likely to be strip searched, and East/SouthEast Asian females are least likely to be strip searched. In conclusion, our analysis of the Arrests and Strip Searches dataset revealed significant differences in mean strip search frequencies among different groups. Therefore, individuals belonging to certain groups are more likely to experience strip searches by the Toronto Police Department.

4.1 Interaction Plot

Figure 8 Interaction between Race and Strip Search by Sex



We generated the interaction plot (Figure 8) to see the interaction between the perceived race and strip search for both males and females. From this plot, we can interpret that males, on average, have a higher chance of experiencing strip searches than females. Also, we can state that Indigenous males have the highest chance of experiencing strip searches whereas East/Southeast Asian females have the lowest chance of experiencing strip searches.

5 Discussion & Conclusion

Our objective was to investigate whether there is any bias in our justice system towards particular groups by examining the correlation between those groups and strip searches

conducted at the time of arrest. From our study, we have validated our first hypothesis that there will be a significant difference in the frequency of strip searches conducted among different age groups at the time of arrest. Furthermore, we also verified our second hypothesis that there will be a significant difference in the frequency of strip searches conducted between the perceived race and sex. As a result, the findings show that there is bias towards certain groups in strip searching.

However, the analysis and findings could not be applied to every scenario because there are many limitations in our dataset. First of all, this data was only collected in Toronto, which the conclusions drawn from the study may not be representative of Ontario or Canada as a whole. Moreover, we encountered difficulty in understanding the precise meanings of certain attributes in the dataset due to the absence of a codebook. For example, the division of the arrest location was presented by numbers instead of geographic names, which makes it hard for us to conduct analysis and make predictions based on the geographic locations. Additionally, the recording of an individual's race is also not always clear. For instance, individuals who have been arrested multiple times have been recorded as a different perceived race each time they were arrested. Last but not the least, there are many missing data in this dataset. We had to discard certain individuals' data due to the presence of missing values in some of the attributes.

Reference

Arrests and strip searches (RBDC-arr-TBL-001). (n.d.). Retrieved February 26, 2023, from https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/about

Dean, L. (2011, May 25). Strip searches pervasive despite huge settlements. Retrieved February 26, 2023, from https://www.huffpost.com/entry/strip-searches-pervasive_n_241544

Keeton, A. (2015). Strip searching in the age of colorblind racism: The disparate impact of florence v. Board of Chosen Freeholders of the County of Burlington. *Michigan Journal of Race & Law,* (21.1), 55. doi:10.36643/mjrl.21.1.strip

Lemke, M. (2022, July 18). Policing toronto: Strip searching in a divided city - the bullet. Retrieved February 26, 2023, from https://socialistproject.ca/2022/07/policing-toronto-strip-searching-in-a-divided-city/

Newburn, T. (2004). Race, crime and injustice?: Strip Search and the treatment of suspects in custody. *British Journal of Criminology, 44*(5), 677-694. doi:10.1093/bjc/azh043

Revealed: Met police strip-searched 650 children in two-year period. (2022, August 07). Retrieved February 26, 2023, from https://www.theguardian.com/uk-news/2022/aug/08/police-data-raises-alarm-over-welfare-of-strip-searched-children

Woodward, J. (2022, December 14). Another Toronto officer alleges discrimination in complaint against Police Association. Retrieved February 28, 2023, from https://toronto.ctvnews.ca/another-toronto-officer-alleges-discrimination-in-complaint-against-police-association-1.6193518