Final Project

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

April 16th 2023

1 Introduction

1.1 Introduction

Strip searching is a technique that gives officers significant control over a person, resembling the unequal power dynamics found in sexual assault cases (Lemke, 2022). This procedure, which exposes a person's body and invades their privacy, is frequently 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 traumatized 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 logistic regression analysis to determine the factors associated with the likelihood of an individual being subjected to a strip search during arrest, considering variables such as age group, perceived race, and sex. Additionally, we conducted a one-way

ANCOVA to investigate if there is a difference in the number of strip searches between perceived race groups after controlling for the total number of arrests. 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 strip-search powers in police stations has revealed that some demographic groups are disproportionately targeted. Keeton (2015) found that Afro-Caribbean arrestees were more likely to undergo strip searches, even after considering factors such as sex, age, reason for arrest, and charge. This result highlights potential racial disparities in strip search practices. Building on this research, we aim to examine if there is a difference in the number of strip searches among perceived race groups after accounting for the total number of arrests.

Racial discrimination in the justice system has been extensively investigated. For example, data from the Toronto Police Service on race-based strip searches showed that Black people represent only about 10% of the city's population, yet they constituted one in every three individuals strip-searched (Lemke, 2022). This finding supports the idea of potential racial bias in strip search practices. Likewise, concerns regarding sexual harassment and assault by police officers have increased in recent years. The Toronto Police Service reported that nearly 60% of its employees have either witnessed or experienced sexual harassment (Woodward, 2022). This information raises questions about the possibility of gender discrimination in strip search practices. Based on these findings, we hypothesize that age group, perceived race, and sex may be associated with the likelihood of an individual being subjected to a strip search during arrest.

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 based on our preliminary analysis of the dataset:

- Research Question 1 (Logistic Regression): What factors are associated with the likelihood of an individual being subjected to a strip search during arrest, considering variables such as age group, perceived race, and sex?
- Research Question 2 (One-way ANCOVA): Is there a difference in the number of strip searches between perceived race groups, after controlling for the total number of arrests?

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, and age group at the time of arrest. To perform a one-way ANCOVA analysis, the dataset is initially grouped by both 'PersonID' and 'Perceived_Race' variables. Then, the 'agg' function is employed to aggregate the data, calculating the total number of strip searches ('StripSearch', 'sum') and the count of arrests ('StripSearch', 'count') for each distinct combination of 'PersonID' and 'Perceived_Race'. This process yields a dataset consisting of 40,378 unique individuals who have experienced arrest. 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

In logistic regression analysis, the dependent variable chosen 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".

For the one-way ANCOVA analysis, the dependent variable selected is "TotalStripSearches", which represents the total number of strip searches conducted on each individual. As opposed to a binary variable, this continuous variable is suitable for ANOVA analysis. By grouping the dataset by 'PersonID' and 'Perceived_Race' and aggregating the strip searches, we effectively transformed the original binary "StripSearch" variable into a continuous one.

1.7 Independent variable

In the logistic regression analysis, the independent variables include "Age_group__at_arrest_", "Perceived_Race", and "Sex". The dataset features eight "Perceived_Race" categories: Black, East/Southeast Asian, Indigenous, Latino, Middle-Eastern, South Asian, Unknown or Legacy, and White. Additionally, there are three "Sex" categories: Male, Female, and Unknown. For clarity, we have excluded the Unknown category from our analysis. The "Age_group__at_arrest_" originally consisted of eight categories, but due to redundancy, we reorganized them into the following age groups: 17 years and under, 18 to 24 years, 25 to 34 years, 35 to 44 years, 45 to 54 years, 55 to 64 years, and 65 years and older.

In ANCOVA, there are typically two types of independent variables: categorical factors and continuous covariates. In our analysis, the categorical factor is "Perceived_Race", and the continuous covariate is "TotalArrests".

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.

Table 2: Strip Search and Perceived Race Contingency Table

Perceived_Race	StripSearch_0	StripSearch_1	Total
Black	15092	2434	17526
East/Southeast Asian	4074	341	4415
Indigenous	1628	306	1934
Latino	1635	132	1767
Middle-Eastern	3009	228	3237
South Asian	3356	257	3613
Unknown or Legacy	4517	536	5053
White	24152	3566	27718
Total	57463	7800	65263

Table 3: Percentage of strip search in each Perceived Race

Perceived Race	No	Yes
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 displays the distribution of strip searches across different perceived race categories. The rows represent various perceived race categories, while the columns indicate the binary outcome of strip searches (0 for "No" and 1 for "Yes"). The "Total" row and column display the sums of each respective row and column.

In Table 2 and 3, it is evident that White and Black individuals are the largest and second largest race groups subjected to strip searches, respectively. For White individuals, there were a total of 27,718 cases, of which 24,152 (87%) did not undergo strip searches (0), and 3,566 (13%) were subjected to strip searches (1). In the case of Black individuals, out of a total of 17,526 cases, 15,092 (86%) were not subjected to strip searches (0), while 2,434 (14%) underwent strip searches (1).

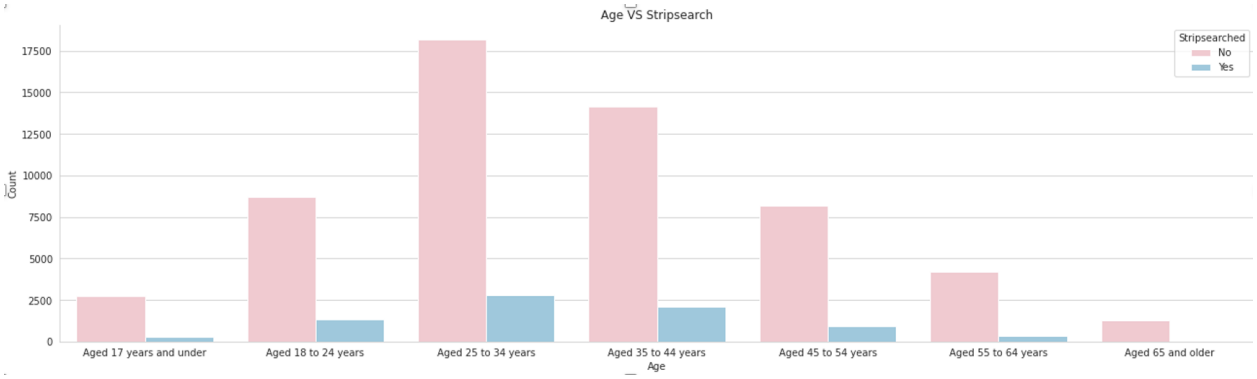

Figure 1: Distribution of Age Group by Sex

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.

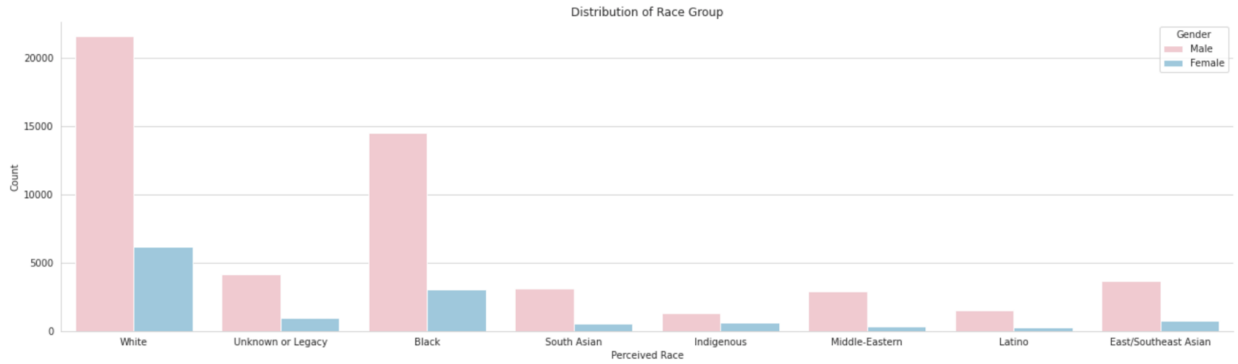

Figure 2: Distribution of Perceived Race by Sex

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.

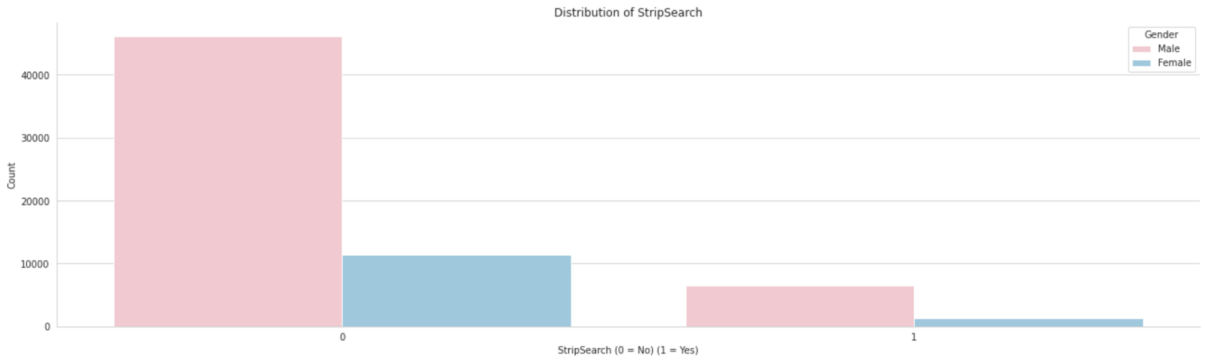

Figure 3: Distribution of Stripsearch by Sex

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.

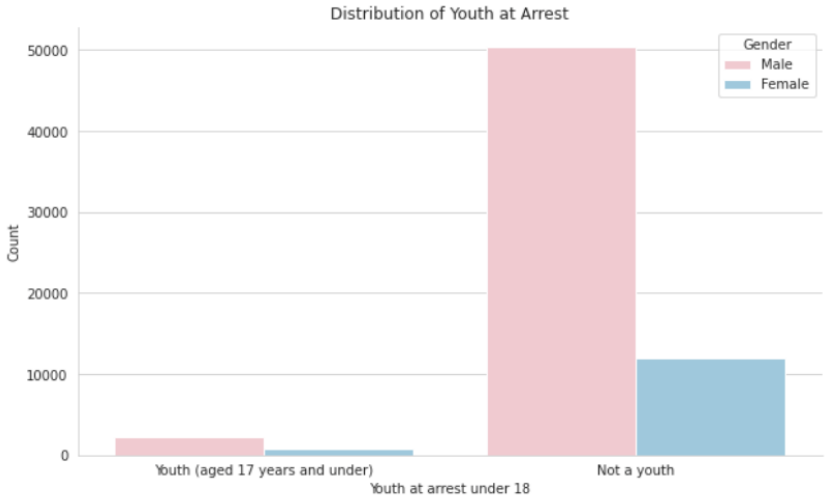

Figure 4: Distribution of Arrest under 18 years of age by Sex

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.

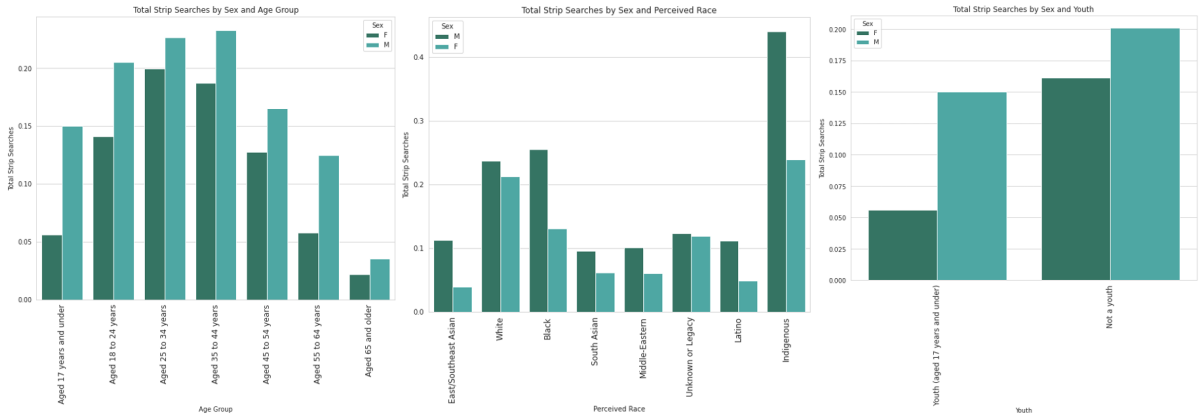

Figure 5: Distribution of Total Strip Search by Sex and Age Group, Perceived Race, and Youth

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.

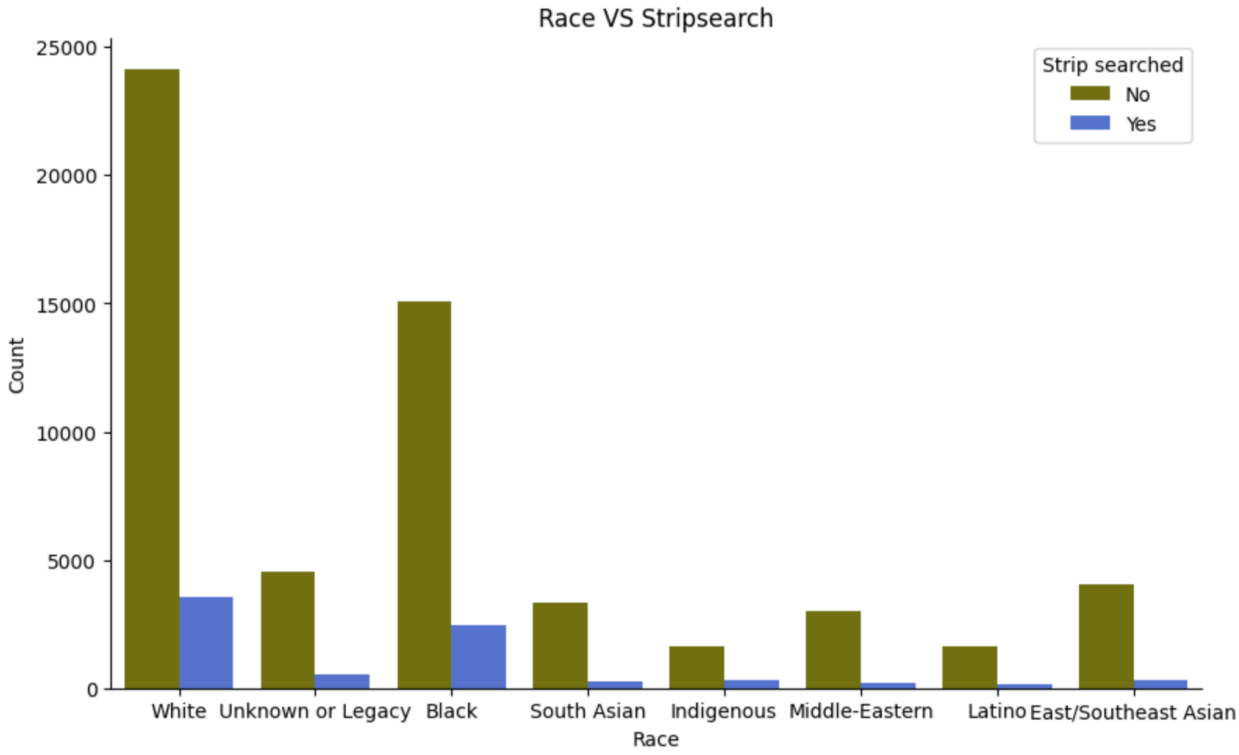

Figure 6: Distribution of Total Strip Searches by Perceived Race

Figure 6 shows the number of arrested individuals who have and have not been strip searched for every perceived race. As we can observe, white people were the race that experienced the most strip search in number. On the other hand, indigenous people were the race that experienced the least strip search in number. However, this distribution only shows the total count but it does not show the proportion and ratio.

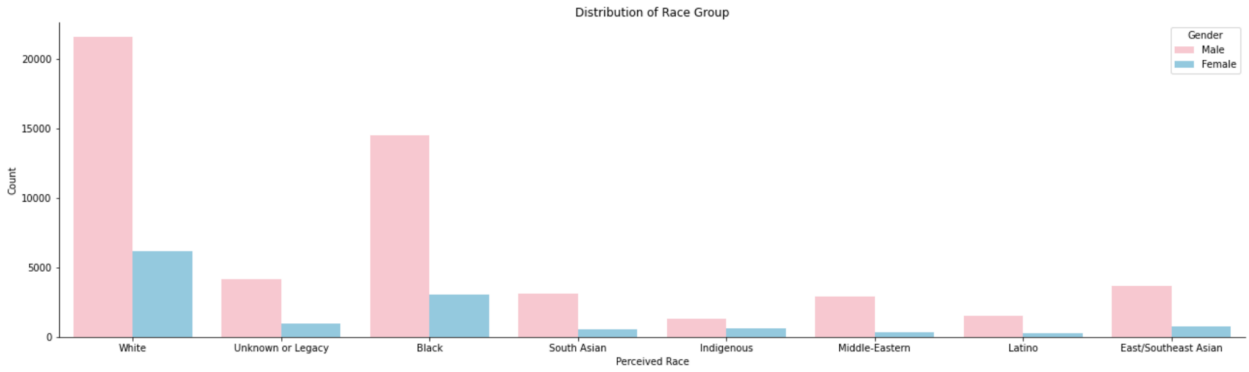

Figure 7: Distribution of Race group

As we mentioned above, Figure 6 only shows the total count instead of the proportion. In figure 7, the distribution shows the perceived race and its total number of arrested. As we can observe, the most number of arrested race groups is white and the least number of arrested race groups is Indigenous.

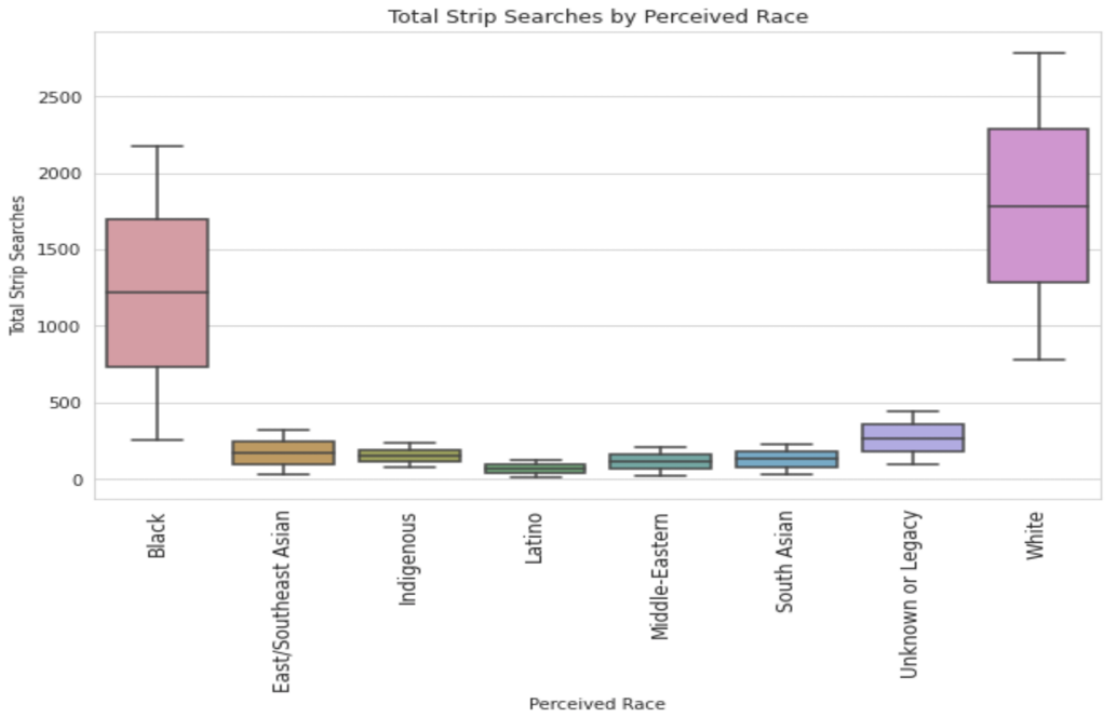

Figure 8: Box plot of Total Strip Searches by Perceived Race

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 8 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 3 reveals that Indigenous people have the highest percentage (15.82%) of experiencing strip searches. Table 3 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.

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.

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

Based on our findings mentioned in the Descriptive Statistics and T-tests sections, we will use inferential statistical tests to explore our research questions. To further explore our research analysis, we will perform power analysis, ANCOVA, and Logistic Regression to deepen our findings.

The power analysis will determine the sample size required to detect a certain effect size with a certain level of confidence. It aims to help us determine the likelihood of finding a true difference between groups in their study. It also ensures that research studies are designed with adequate statistical power and the results are meaningful and reliable.

The ANCOVA is a statistical technique for analyzing the relationship between a dependent variable and one or more independent variables while controlling for the effects of one or more covariates. It allows us to better understand the relationships between variables by accounting the effects of other variables that may influence the outcome.

Logistic regression is a statistical technique used to analyze the relationship between a categorical dependent variable and one or more independent variables. It models the probabilities of an event occurring based on a set of predictors or independent variables. It helps us to predict an event occurring based on one or more independent variables.

4 Results & Finding

4.1 Power Analysis

Power analysis is a statistical technique utilized to estimate the likelihood of identifying a genuine effect in a research study based on a given sample size and statistical test. In our power analysis, we aim to further discover the relation for White people and Black people in Perceived race (independent variable) and Strip search (dependent variable). The significant level was set at 0.05. We first calculated the effect size using the Cohen's d to find the standardized difference between the two means in units of standard deviation. The results state that the variables have an effect size of 0.03. The generated result is smaller than 0.2, which indicates that the relationship being investigated is small and may not have much practical significance.

Then, we calculated the sample size to find the number of participants included in the study. The result indicates that we have a sample size of 17310, which is fairly large. A large sample size generally leads to greater statistical power because it increases the precision of the estimate and reduces the influence of random error.

The power analysis generated a result of 0.94, this indicates that there is a 94% chance of correctly rejecting the null hypothesis if the alternative hypothesis is true. In other words, if there is a true difference between the two groups being compared, this power analysis suggests that the study has a high probability of detecting it. Overall, a power of 0.94 is considered to be a high level of statistical power, indicating that the study is well-designed to detect a meaningful effect size.

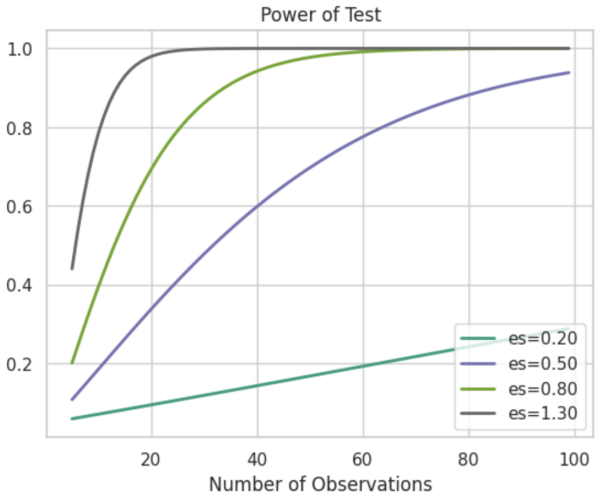

Figure 9 Power of Test

4.1 ANCOVA

4.1.1 Assumption Check

Before conducting the one-way ANCOVA, we need to verify that the data meet the assumptions required for it.

Normality: Figure 10 is the Normal Q-Q plot, which is used to examine whether the residuals are normally distributed. Although some dots at two ends are dispersed, we see that most of the residual points follow the straight dashed line. Therefore, we may assume that residuals are normally distributed.

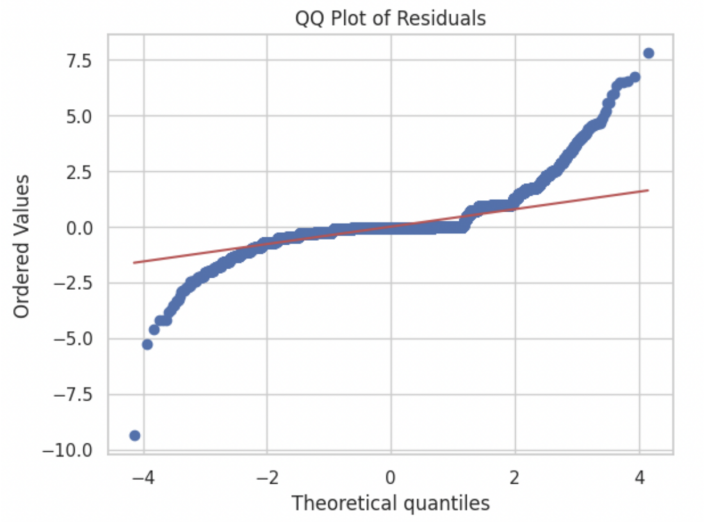

Figure 10 QQ Plot

Homogeneity of variances: Given the Levene's test statistic of 69.284 and a p-value of 5.565e-100, the p-value is much smaller than a typical significance level (e.g., 0.05). This suggests that the assumption of homogeneity of variances may not hold for this dataset, as there is significant evidence that the variances of TotalStripSearches are not equal across the different Perceived_Race groups.

Homogeneity of regression slopes: The interaction p-value (0.0007978599535934704) is smaller than the typical significance level of 0.05. This indicates that there is a statistically significant interaction between the perceived race and total arrests in predicting the total strip searches. This result suggests that the relationship between total arrests and total strip searches is not the same for all perceived race groups, and the homogeneity of regression slopes assumption may not hold for this dataset.

Linearity: Figure 11 displays scatter plots for each perceived race group, with total arrests on the x-axis and total strip searches on the y-axis. Upon examination, the relationship between total arrests and total strip searches appears to be nonlinear within all of the race groups. This observation suggests that the linearity assumption may not hold in this dataset.

Independence of observations: Given the nature of the dataset, it is reasonable to assume that the observations are independent. Each row in the dataset represents a unique individual with their respective perceived race, total strip searches, and total arrests. Assuming that each arrest and strip search incident is unrelated to the others and that the data collection process

doesn't introduce any dependencies, the assumption of independence of observations is likely satisfied. In other words, the occurrence of one arrest and strip search incident does not influence or depend on the occurrence of another arrest and strip search incident.

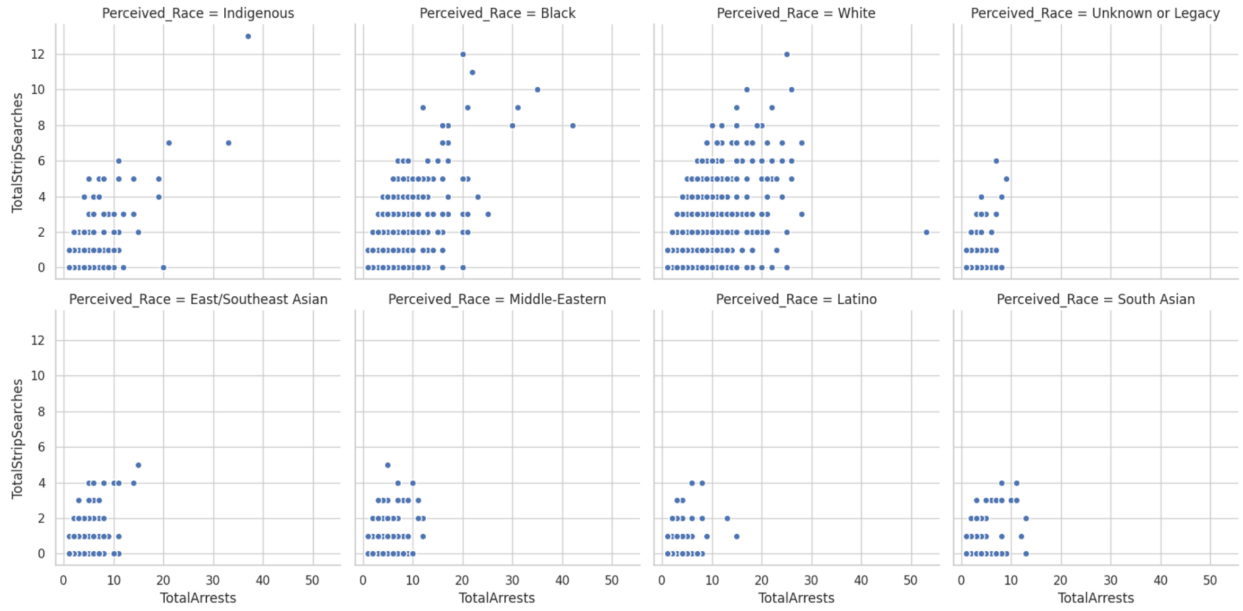

Figure 11 Scatter Plots for Perceived Race

4.1.2 ANCOVA Result

Table 4 OLS Regression Results

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	-0.1350	0.005	-24.699	0.000	-0.146	-0.124
C(Perceived_Race)[T.East/Southeast Asian]	-0.0459	0.010	-4.688	0.000	-0.065	-0.027
C(Perceived_Race)[T.Indigenous]	-0.0060	0.018	-0.333	0.739	-0.041	0.029
C(Perceived_Race)[T.Latino]	-0.0620	0.015	-4.219	0.000	-0.091	-0.033

C(Perceived_Race)[T.Middle-Eastern]	-0.0682	0.011	-6.001	0.000	-0.090	-0.046
C(Perceived_Race)[T.South Asian]	-0.0521	0.011	-4.940	0.000	-0.073	-0.031
C(Perceived_Race)[T.Unknown or Legacy]	0.0058	0.009	0.648	0.517	-0.012	0.023
C(Perceived_Race)[T.White]	-0.0288	0.006	-4.538	0.000	-0.041	-0.016
TotalArrests	0.2177	0.001	154.171	0.000	0.215	0.220

The intercept coefficient (-0.1350) represents the expected value of the dependent variable when all other variables in the model are equal to zero. The coefficients for each level of the Perceived_Race variable indicate the expected change in the dependent variable for each level of Perceived_Race, relative to the reference category. For example, compared to the reference category, individuals perceived as East/Southeast Asian have an expected decrease in the dependent variable of 0.0459 units, individuals perceived as Indigenous have an expected decrease of 0.0060 units, and individuals perceived as Latino have an expected decrease of 0.0620 units. The coefficient for TotalArrests indicates that, for each additional unit of TotalArrests, the expected value of the dependent variable increases by 0.2177 units.

4.2 Logistic Regression

4.2.1 Assumption Check

Before conducting the logistic regression, we need to verify that the data meet the assumptions required for it.

Independence of observations: Given the dataset's nature, it is reasonable to assume that the observations are independent, as each record represents a unique arrest and strip search incident.

Table 5 Some of the Result of Strip Search Counts for Each Race

Race	**Sex**	**Age Group at Arrest**	**Counts**
South Asian	Female	55 to 64 Years old	0
Latino	Female	55 to 64 Years old	0

Unknown or Legacy	Female	65 and older	0
East/Southeast Asian	Female	65 and older	0
Black	Female	65 and older	0
White	Male	45 to 54 Years old	437
Black	Male	18 to 24 Years old	560
White	Male	25 to 34 Years old	942

Large sample size: From the output on Table 4, it seems that some groups have many strip searches (e.g., White Males aged 25 to 34 years have 942 events), while some other groups have very few or no strip searches (e.g., South Asian Females aged 55 to 64 years have 0 events). However, considering that our dataset has a large overall sample size (65,276 individuals) and many groups have a substantial number of events, it is likely that the assumption of a large sample size is fulfilled. The logistic regression model should be able to provide reliable results for the groups with a sufficient number of events.

Linearity of logit: This assumption is relevant for continuous independent variables but doesn't apply to categorical ones. Our logistic regression analysis only involves categorical independent variables, so this assumption isn't applicable in our case.

Table 6 Correlation Matrix for Perceived Race (East/Southeast Asian)

Correlation Matrix	Perceived Race East/Southeast Asian
East/Southeast Asian	1.000000
Indigenous	- 0.046937
Latino	-0.044516

Middle Eastern	-0.061237
South Asian	-0.065212
Unknown	-0.077441
White	-0.230076
Sex Male	0.016896
Age_group__at_arrest__Aged_18_to_24_	0.027614
Age_group__at_arrest__Aged_25_to_34_years	-0.026562
Age_group__at_arrest__Aged_35_to_44_years	-0.000588
Age_group__at_arrest__Aged_45_to_54_years	-0.005745
Age_group__at_arrest__Aged_55_to_64_years	0.014982
Age_group__at_arrest__Aged_64_and_older	0.021928

Table 7 VIF Table

Variance	Features	VIF
0	East/Southeast Asian	0.051230

1	Indigenous	0.020615
2	Latino	0.102379
3	Middle Eastern	0.021969
4	South Asian	0.050285
5	Unknown	0.065985
6	White	0.026194
7	Sex Male	0.014911
8	Age_group__at_arrest__Aged_18_to_24_	0.021478
9	Age_group__at_arrest__Aged_25_to_34_years	0.023437
10	Age_group__at_arrest__Aged_35_to_44_years	0.023014
11	Age_group__at_arrest__Aged_45_to_54_years	0.032142
12	Age_group__at_arrest__Aged_55_to_64_years	0.024347
13	Age_group__at_arrest__Aged_64_and_older	0.042157

Table 8 Tolerance Value

0	19.519628

1	48.509528
2	9.767652
3	45.518890
4	19.886614
5	15.154928
6	38.176729
7	67.066751
8	46.559746
9	42.668397
10	43.451778
11	31.112369
12	41.072902
13	23.720936

Absence of multicollinearity: Correlation coefficient results generated in Table 6, none of the correlation coefficients seem to be above the threshold (usually 0.8 or 0.9) that would indicate severe multicollinearity. Based on the VIF values in Table 7, there does not appear to be multicollinearity among the variables in our dataset. A common rule of thumb is to consider a

VIF value of 10 or higher as indicative of multicollinearity, and all of our VIF values are well below that threshold. Similarly, for the tolerance values in Table 8, there does not appear to be multicollinearity among the variables in our dataset. A common rule of thumb is to consider a tolerance value of 0.1 or lower as indicative of multicollinearity, and all of our tolerance values are well above that threshold. Overall, based on the above results, the independent variables are not highly correlated with each other and the regression coefficients can be considered reliable.

4.2.2 Logistic Regression Result

Table 9 Logistic Regression Table

Intercept	Coef	Std Err	z	p>|z|	[0.025	0.975]
East/Southeast Asian	-2.3934	0.077	-30.965	0	-2.545	-2.242
Indigenous	-0.5954	0.068	-8.735	0	-0.729	-0.462
Latino	0.2265	0.074	3.068	0.002	.082	0.371
Middle Eastern	-0.5965	0.101	-5.930	0	-0.794	-0.399
South Asian	-0.7884	0.082	-9.660	0	-0.948	-0.628
Unknown	-0.7126	0.076	-9.402	0	-0.861	-0.564
White	-0.2880	0.057	-5.027	0	-0.400	-0.176
Sex Male	-0.0122	0.033	-0.374	0.709	-0.076	0.052
Age_group__at_arrest__Aged_18_to_24_	0.3173	0.037	8.593	0	0.245	0.582
Age_group__at_arrest__Aged_25_to_34_years	0.4316	0.077	5.618	0	0.281	0.517
Age_group__at_arrest__Aged_35_to_44_years	0.3735	0.073	5.105	0	0.230	0.486
Age_group__at_arrest__Aged_45_to_54_years	0.3396	0.075	4.555	0.473	0.193	0.214
Age_group__at_arrest__Aged_55_to_64_years	0.0575	0.080	0.718	0.003	-0.099	-0.099

Age_group__at_arrest__Aged_65_and_older	-0.2849	0.095	-3.011	0	-0.470	-1.074

According to Table 9, we can observe the following results:

- Intercept: The intercept value of -2.3934 represents the log odds of experiencing a strip search for a female individual who is Blak, and belongs to the age group "Aged 17 years and under."
- Perceived_Race_East_Southeast_Asian: Holding all other variables constant, the log odds of experiencing a strip search for individuals of East Southeast Asian descent are lower by 0.5954 compared to Black individuals.
- Perceived_Race_Indigenous: Holding all other variables constant, the log odds of experiencing a strip search for Indigenous individuals are higher by 0.2265 compared to Black individuals.
- Perceived_Race_Latino: Holding all other variables constant, the log odds of experiencing a strip search for Latino individuals are lower by 0.5965 compared to Black individuals.
- Perceived_Race_Middle_Eastern: Holding all other variables constant, the log odds of experiencing a strip search for Middle Eastern individuals are lower by 0.7884 compared to Black individuals.
- Perceived_Race_South_Asian: Holding all other variables constant, the log odds of experiencing a strip search for South Asian individuals are lower by 0.7126 compared to Black individuals.
- Perceived_Race_Unknown_or_Legacy: Holding all other variables constant, the log odds of experiencing a strip search for individuals with unknown or legacy perceived race are lower by 0.2880 compared to Black individuals.
- Perceived_Race_White: Holding all other variables constant, the log odds of experiencing a strip search for White individuals do not significantly differ from those of Black individuals (coef: -0.0122, p-value: 0.709).
- Sex_M: Holding all other variables constant, the log odds of experiencing a strip search for males are higher by 0.3173 compared to females.

- Age_group__at_arrest__Aged_18_to_24_years: Holding all other variables constant, the log odds of experiencing a strip search for individuals aged 18 to 24 years are higher by 0.4316 compared to individuals aged 17 years and under.
- Age_group__at_arrest__Aged_25_to_34_years: Holding all other variables constant, the log odds of experiencing a strip search for individuals aged 25 to 34 years are higher by 0.3735 compared to individuals aged 17 years and under.
- Age_group__at_arrest__Aged_35_to_44_years: Holding all other variables constant, the log odds of experiencing a strip search for individuals aged 35 to 44 years are higher by 0.3396 compared to individuals aged 17 years and under.
- Age_group__at_arrest__Aged_45_to_54_years: Holding all other variables constant, the log odds of experiencing a strip search for individuals aged 45 to 54 years do not significantly differ from those of individuals aged 17 years and under (coef: 0.0575, p-value: 0.473).
- Age_group__at_arrest__Aged_55_to_64_years: Holding all other variables constant, the log odds of experiencing a strip search for individuals aged 55 to 64 years are lower by 0.2849 compared to individuals aged 17 years and under.
- Age_group__at_arrest__Aged_65_and_older: Holding all other variables constant, the log odds of experiencing a strip search for individuals aged 65 and older are lower by 1.5014 compared to individuals aged 17 years and under.

Table 10 Logistic Regression Result Table

Covariate	**Lower CI**	**Upper CI**	**Odds Ratio**
Intercept	0.078477	0.106251	0.091314
Perceived_Race_East_Southeast_Asian	0.482418	0.630165	0.551365
Perceived_Race_Indigenous	1.085223	1.449510	1.254209
Perceived_Race_Latino	0. 452202	0. 670751	0. 550740

Perceived_Race_Middle_Eastern	0. 387377	0. 533425	0. 454573
Perceived_Race_South_Asian	0. 422655	0. 568884	0. 490349
Perceived_Race_Unknown_or_Legacy	0. 670107	0. 838852	0. 749747
Perceived_Race_White	0. 926632	1. 053166	0. 987875
Sex_M	1. 277494	1. 476439	1. 373369
Age_group__at_arrest__Aged_18_to_24_	1. 324455	1. 789878	1. 539679
Age_group__at_arrest__Aged_25_to_34_years	1. 258749	1. 676918	1. 452866
Age_group__at_arrest__Aged_35_to_44_years	1. 213445	1. 625362	1. 404382
Age_group__at_arrest__Aged_45_to_54_years	0. 905355	1. 239232	1. 059219
Age_group__at_arrest__Aged_55_to_64_years	0. 624795	0. 905317	0. 752089
Age_group__at_arrest__Aged_64_and_older	0. 145283	0.341757	0. 222826

We perform a logistic regression odds ratios to examine the effects of perceived race, sex, and age group at arrest on the likelihood of experiencing a strip search. From Table 10, we can observe that most features are statistically significant.

- Compared to Black individuals, East Southeast Asian individuals have a lower likelihood of experiencing a strip search. The odds ratio is approximately 0.55. For East Southeast

Asian individuals, the odds of experiencing a strip search are about 0.55 times those of Black individuals.

- Compared to Black individuals, Indigenous individuals have a higher likelihood of experiencing a strip search. The odds ratio is approximately 1.25. The odds of experiencing a strip search for Indigenous individuals are about 1.25 times those of Black individuals.
- Compared to Black individuals, Latino individuals have a decreased likelihood of experiencing a strip search. The odds ratio is approximately 0.55. For Latino individuals, the odds of experiencing a strip search are about 0.55 times those of Black individuals.
- Compared to Black individuals, Middle Eastern individuals have a decreased likelihood of experiencing a strip search. The odds ratio is approximately 0.45. For Middle Eastern individuals, the odds of experiencing a strip search are about 0.45 times those of Black individuals.
- Compared to Black individuals, South Asian individuals have a decreased likelihood of experiencing a strip search. The odds ratio is approximately 0.49. For South Asian individuals, the odds of experiencing a strip search are about 0.49 times those of Black individuals.
- Compared to Black individuals, individuals with unknown or legacy perceived race have a decreased likelihood of experiencing a strip search. The odds ratio is approximately 0.75. For individuals with unknown or legacy perceived race, the odds of experiencing a strip search are about 0.75 times those of Black individuals.
- Males have a higher likelihood of experiencing a strip search compared to females. The odds ratio is approximately 1.37. For males, the odds of experiencing a strip search are about 1.37 times those of females.
- Compared to individuals aged 17 years and under, individuals aged 18 to 24 years have a higher likelihood of experiencing a strip search. The odds ratio is approximately 1.54. For individuals aged 18 to 24 years, the odds of experiencing a strip search are about 1.54 times those of individuals aged 17 years and under.
- Compared to individuals aged 17 years and under, individuals aged 25 to 34 years have a higher likelihood of experiencing a strip search. The odds ratio is approximately 1.45. For

individuals aged 25 to 34 years, the odds of experiencing a strip search are about 1.45 times those of individuals aged 17 years and under.

- Compared to individuals aged 17 years and under, individuals aged 35 to 44 years have a higher likelihood of experiencing a strip search. The odds ratio is approximately 1.40. For individuals aged 35 to 44 years, the odds of experiencing a strip search are about 1.40 times those of individuals aged 17 years and under.
- Compared to individuals aged 17 years and under, individuals aged 55 to 64 years have a decreased likelihood of experiencing a strip search. The odds ratio is approximately 0.75. For individuals aged 55 to 64 years, the odds of experiencing a strip search are about 0.75 times those of individuals aged 17 years and under.
- Compared to individuals aged 17 years and under, individuals aged 65 and older have a significantly decreased likelihood of experiencing a strip search. The odds ratio is approximately 0.22. For individuals aged 65 and older, the odds of experiencing a strip search are about 0.22 times those of individuals aged 17 years and under.

These odds ratios are considered significant because their respective p-values are less than the conventional significance level of 0.05 in Table 9. The odds ratios for "Perceived_Race_White" and "Age_group__at_arrest__Aged_45_to_54_years" are not statistically significant.

Furthermore, we assessed the odds ratio along with the lower and upper confidence intervals in Table 10, which can help to understand statistical significance. If the confidence interval includes 1, then the result is not statistically significant and vice versa. Based on Table 9, we observe that for the intercept, the odds ratio is 0.091314, with a 95% confidence interval ranging from 0.078477 to 0.106251. From the table, we can also see that the 95% confidence interval ranges from [Lower CI] to [Upper CI] for race, sex, and age group. To avoid redundancy from the previously mentioned odds ratios results, we will only pick a few examples. For perceived race in East Southeast Asians, the odds ratio is 0.551365, with a 95% confidence interval ranging from 0.482418 to 0.630165. This means that we are 95% confident that the true odds ratio lies within this range and suggests a statistically significant decrease in the likelihood of a strip search compared to the reference group. For the perceived race of White individuals, the 95% confidence interval ranges from 0.927 to 1.053, suggesting no statistically significant

difference in the likelihood of a strip search compared to the reference group. For the age group of 18 to 24 years, the 95% confidence interval ranges from 1.324 to 1.790, suggesting a statistically significant increase in the likelihood of a strip search compared to the reference age group.

Additionally, we have assessed the accuracy score of the logistic regression model, to measure how well our model is correctly classifying the true outcomes. In this case, the accuracy is approximately 87.75%, which means the model is correctly predicting the outcomes for about 87.75% of the test dataset.

Last, from the confusion matrix in Figure 10, we can observe that:

- 11,449 True Negatives (TN): The model correctly predicted negative outcomes (i.e., no strip search) 11,449 times.
- 1,599 False Positives (FP): The model incorrectly predicted positive outcomes (i.e., strip search) 1,599 times.
- 0 True Positives (TP): The model correctly predicted positive outcomes (i.e., strip search) 0 times.
- 0 False Negatives (FN): The model incorrectly predicted negative outcomes (i.e., no strip search) 0 times.

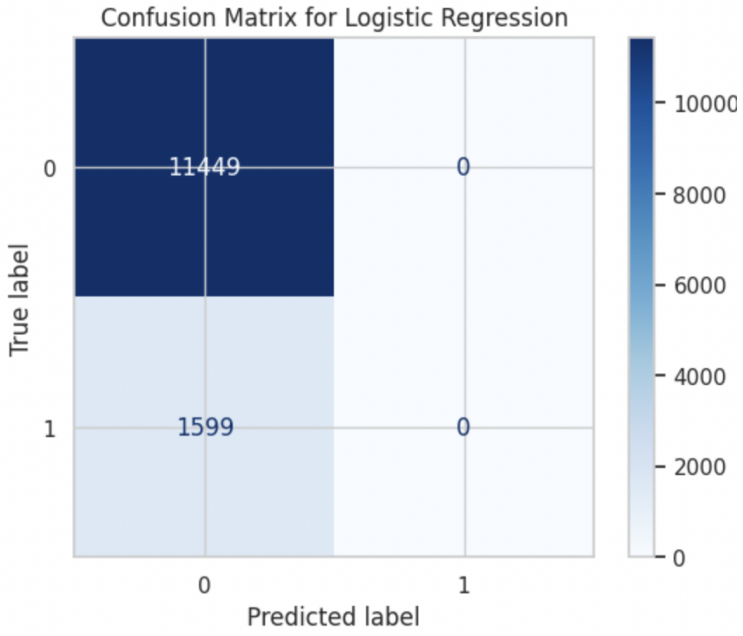

Figure 12 Confusion Matrix for Logistic Regression

From the confusion matrix, it appears that the model is only predicting negative outcomes (no strip search) and is not able to correctly identify any positive outcomes (strip search).

5 Discussion

The findings of a study examining how perceived race, year, and sex impact the likelihood of being subjected to a strip search during arrests are reported in the Results section. This section is organized into three subsections, which discuss the power analysis, ANCOVA, and logistic regression results.

In the Power Analysis subsection, the effect size for strip search was calculated using Cohen's d, and it was found to be 0.03, indicating a weak relationship between the variables. The study also computed the necessary sample size for each racial group and concluded that the dataset's large sample sizes provided more statistical power than initially anticipated. The power graph displayed a pattern where increasing effect sizes corresponded with a higher rate of increase in statistical power of the test, indicating that larger effect sizes have a stronger association with the likelihood of detecting a significant effect in the study.

In the ANCOVA section, the results stated that there was a statistically significant relationship between perceived race and the probability of being strip searched. This finding supports the hypothesis that perceived race is a significant predictor of the likelihood of being strip searched during arrests.

In the Logistic Regression section, the study investigated the effect of age and race on the likelihood of being strip searched. The results indicated that both age and race were significant predictors of the odds of being strip searched. The confusion matrix result also suggests that the confusion matrix model might not be well-suited for identifying strip searches and could benefit from further refinement or the use of other predictive techniques. For example, the dataset can use synthetic data generation methods like SMOTE

6 Limitation

Firstly, the data for this study was collected only in Toronto, so the conclusions drawn from the study may not be representative of Ontario or Canada as a whole.

Secondly, while White and Black individuals account for the largest and second-largest race groups subjected to strip searches, it is important to consider other factors such as arrest rates, population proportions, and the nature of the offenses when interpreting these numbers in a broader context.

Thirdly, although our dataset satisfies the assumption of a large sample size, it is important to note that the reliability of the results might be limited for groups with very few or no events.

Fourthly, the confusion matrix and contingency table results indicate that the dataset is unbalanced, as the number of individuals who have been strip-searched is much higher than those who have not been strip-searched. This imbalance can consequently impact the performance and reliability of machine learning models.

Fifthly, 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 difficult to conduct analysis and make predictions based on geographic locations.

Lastly, the recording of an individual's race is 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.

7 Conclusion

To sum up, the objective of this research was to examine how the age group, perceived race, and sex affect the likelihood of being subjected to strip searches during arrests. By utilizing power analysis, ANCOVA, and logistic regression, this study shed light on the connections between these variables and their effect on the probability of strip searches.

As a result, by acknowledging the impact of racial bias, temporal trends, and gender disparities on the utilization of strip searches, law enforcement agencies can strive towards promoting fairness and equity in the criminal justice system.

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