Analyzing the Racial Bias in Strip Searches

INF 2178 - Final Assignment

By: Group 42

Sophia Qiu, Kaicheng Huang

Introduction

Police arrest decisions have a profound impact on the individual being arrested. Arrests could result in negative consequences including but not limited to potential harm to reputation, loss of employment, and ultimately deprivation of liberty. It's imperative to ensure the fairness of arrest decisions to uphold the principles of justice and maintain public trust in law enforcement and the criminal justice system. In recent years, concerns have been raised on the use of strip search at time of arrest by Toronto police. Police statistics showed that the use of strip-search was becoming a routine procedure with Black and Indigenous individuals at more risk compared to others (The Canadian Press, 2019).

Many different influential factors can affect a police officer's decision to make an arrest and whether the arrested individual will be strip searched. Previous research has identified a link between individual demographics, socioeconomic status, and arrest decisions (Visher, 1983). Police officers may hold preconceived knowledge about the behaviour and characteristics of different offenders based on their group membership. Individuals exhibiting behaviours inconsistent with the existing schema are more likely to be arrested than those who confirmed the expectations. Newburn and colleagues (2004) found that arrestees of African-Caribbean origin were far more likely to be strip-searched compared to others. This effect remains significant even after controlling for age, gender and the types of crime. This pattern of conflict persists even after a decade of effort on regulating strip searches. Detainees may be subject to strip-search including cavity search as a routine procedure even in the absence of reasonable suspicion of contrabands (Ha, 2011). Experience of strip search may cause physical and mental harm to the recipient, this is especially concerning when the use of strip search is unnecessary and unjustified. These findings highlight the importance of understanding the effect demographic and social factors have on law enforcement practices. Using an arrest dataset published by Toronto Police Service, the current study examines the effect of perceived race and age in the likelihood of being strip searched at time of arrest.

The research questions for the present study are the following:

RQ1: Are there significant differences in the proportion of strip searches between different Perceived Race, while controlling for the age of individuals arrested?

Where the null hypothesis for this research question will be as the following:

Null hypothesis (H_0): There is no significant difference in the proportion of strip searches between different races while controlling for the age.

Alternative hypothesis (H_1): There is a significant difference in the proportion of strip searches between different races while controlling for the age.

RQ2: Does the perceived race of individuals arrested predict the likelihood of being subjected to a strip search, after controlling for age group, gender, and quarter of the arrest event?

Null hypothesis (H_0): There is no relationship between perceived race of individuals arrested and the likelihood of being subjected to a strip search, after controlling for age group, gender, and quarter of the arrest event.

Alternative hypothesis (H_1): There is a relationship between perceived race of individuals arrested and the likelihood of being subjected to a strip search, after controlling for age group, gender, and quarter of the arrest event.

Dataset

Data Overview

The present study is based on the Arrests and Strip Searches dataset obtained from Toronto Police Service Public Safety Data Portal (Toronto Police Service, 2022). The data contains comprehensive information on all arrests and strip searches conducted by the Toronto Police Service from 2020 to 2021. The dataset allows us to gain insight into patterns and trends surrounding arrests and strip search events, including the demographics of those who are arrested and searched. The demographic information contains the individual's personal ID number, perceived race, sex, age group, and the indicator of whether they are a minority. Indicators regarding the individuals were also recorded in the dataset, such as whether the police conducted a strip search on the individual at arrest, and whether the person was booked at a police station within 24 hours of the arrest event. Other information related to the arrest event was also indicated in the dataset, including the location of Division boundaries where the arrest took

place, the category of occurrence leading to the arrest, the individual's action at arrest, the research reason if the individual has been strip-searched, as well as the number of items found through the strip search. This dataset analyzes the frequency and distribution of arrests and strip searches in Toronto, while allowing for the examination of potential factors that may influence police decision-making in these instances.

Data Cleaning

We conducted a comprehensive data cleaning process to ensure the accuracy, completeness, and consistency of the existing dataset. We created a new column, "Age_group," to classify individuals as either youth or adults, based on the "Youth_at_arrest__under_18_years" column in the original column. Rows with a value of "Not a youth" in the original dataset were assigned to "Adult" in the new column, and rows with values "Youth (aged 17 and younger)" or "Youth (aged 17 years and under)" are combined and assigned to "Youth". We then created a new column to indicate the gender of each individual and grouped the data by gender, enabling us to count the number of arrests in each group. We also created a new column to record the quarter in which the individual was arrested and grouped the data by the quarter of each arrest event, providing a quarterly view of the number of arrests. Subsequently, we created a new column, "Race", to categorize the individuals based on their perceived race and combined "South Asian" and "East/Southeast Asian" into a new category "Asian". Then, we convert "Age_group__at_arrest_" into numerical variables by calculating the mean of the age group interval.

We then created a new data frame for further analysis and removed all null values from the dataset to ensure the validity of the data. The new DataFrame contains eleven columns and 60033 rows with no null values. Overall, these cleaning processes enhanced the quality of the dataset and facilitated further analyses.

Exploratory Data Analysis

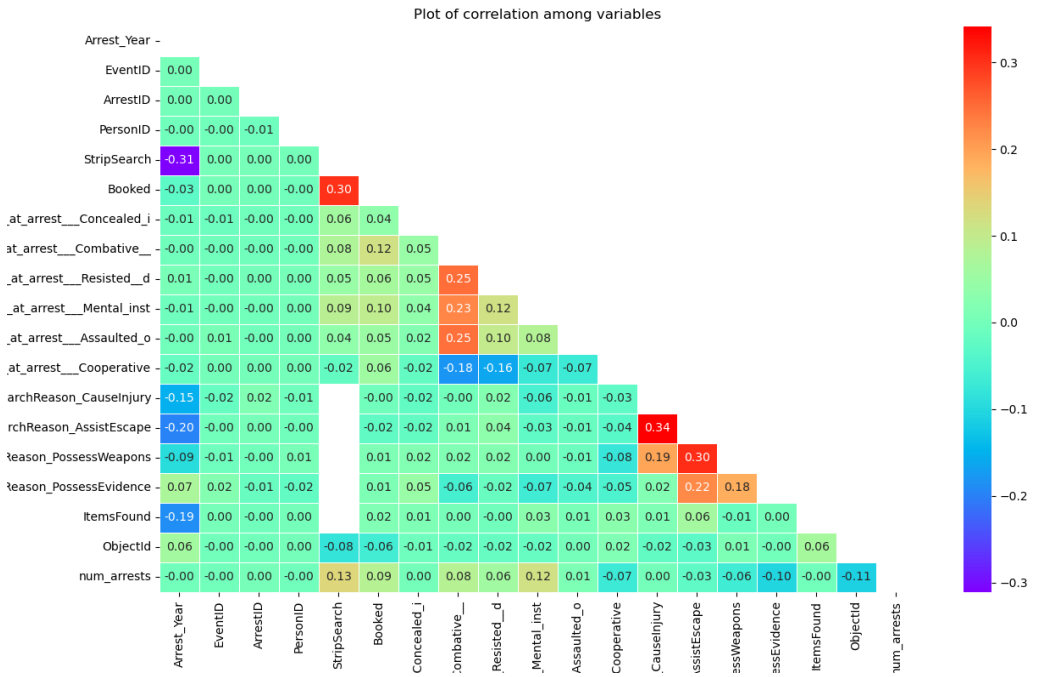
Correlation

A correlation heatmap was plotted to discover the correlation among the variables in the dataset (see Figure 1). However, due to the lack of continuous variables in the data frame, we faced challenges finding the significant association between different variables.

One thing worth noting on the heatmap is that "Booked" is positively associated with "StripSearch" with a correlation of 0.3. Yet, the dataset description stated that all individuals who have been strip-searched are booked within 24 hours. Thus, this inconsistency in the dataset reduces the validity of the correlation between "Booked" and "StripSearch."

Figure 1

Correlation Plot Among Each Variable

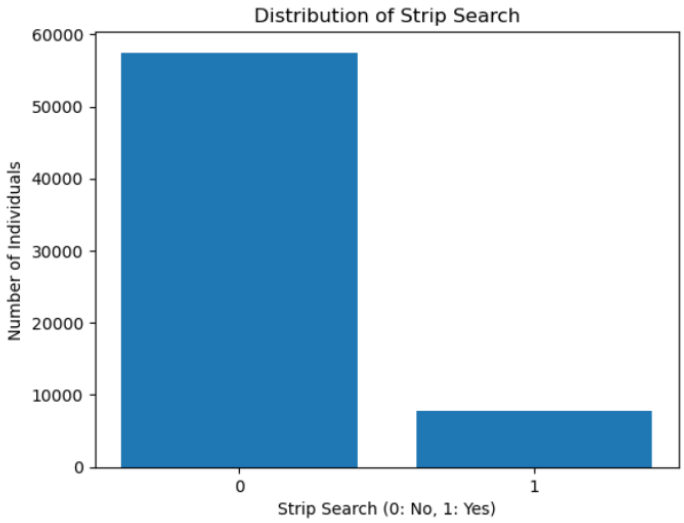


T-Tests and Distributions

A histogram was mapped to visualize the total number of individuals who were strip searched (n=7801) versus those who were not (n=57475) (see Figure 2).

Figure 2

Histogram of Strip Search Distribution



Subsequently, we conducted t-tests to reveal the significant difference between different age groups in the proportion of strip searches t(60032)=-4.83, P<.001. The result suggested that adults (M = 0.12) have a significantly higher proportion of strip search compared to youth (M = 0.09). A bar graph was plotted to visualize the distribution of the proportion of strip search by age groups (see Figure 3). Another t-test was conducted to reveal the significant difference between different genders in the proportion of strip searches t(60032)=7.06, P<.001. The result suggested that adults (M = 0.12) have a significantly higher proportion of strip search compared to youth (M = 0.10). A bar graph was plotted to visualize the distribution of the proportion of strip search by genders (see Figure 3). Lastly, a histogram was plotted to visualize the age distribution (see Figure 4).

Figure 3

Distribution of the Proportion of Strip Search by Different Age, Genders, Quarters, and Races.

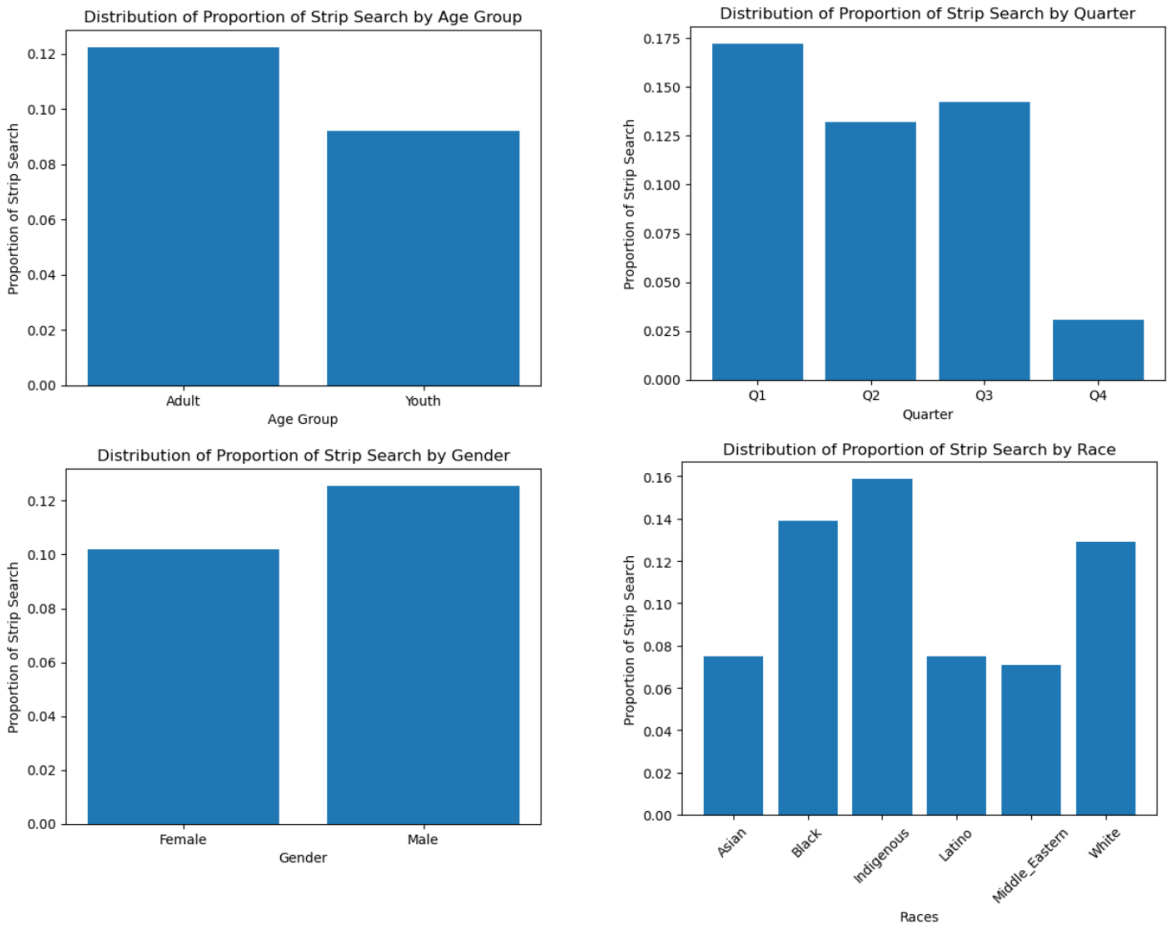
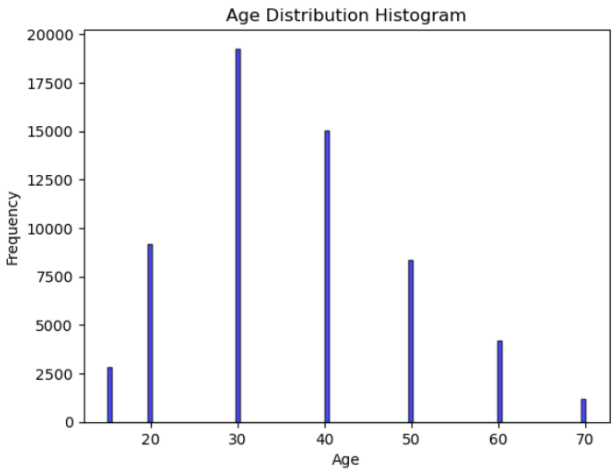


Figure 4

Age Distribution Histogram



Shapiro-Wilk Test

A Shapiro-Wilk test was conducted to assess the normality of the distribution of the dataset. The test statistic is 0.616 and the p-value is less than 0.05, which suggests the null hypothesis should be rejected, and the dataset is not normally distributed. However, the accuracy of the Shapiro-Wilk test may be affected when the sample size is larger than 5000, the p-value may inflate and be inaccurate with the current dataset (n = 65277).

Methods

The results obtained from the descriptive statistics further guided our direction to investigate our research questions.

RQ1: Are there significant differences in the proportion of strip searches between different Perceived Race, while controlling for the age of individuals arrested?

From the EDA above, we could not conclude whether there is any statistical significance of the proportion of strip searches between different perceived races while taking the age of the individuals arrested into account. Therefore, we planned to investigate RQ1 using one-way Analysis of Covariance (ANCOVA) to test if there is any significant difference in the proportion of strip searches between different perceived races while controlling for the age of the individuals arrested. The one-way ANCOVA examines the relationship between a continuous dependent variable (proportion of strip searches) and a categorical independent variable (perceived race) while controlling for the effect of a continuous covariate (age of individuals

arrested). Conducting an ANCOVA helps us to isolate the effect of perceived race on the proportion of strip searches by controlling for the age of individuals arrested. Since age may be a confounding variable that influences both perceived race and the proportion of strip searches, the effect of perceived race on the proportion of strip searches can be estimated more effectively by including age as a covariate in the analysis.

RQ2: Does the perceived race of individuals arrested predict the likelihood of being subjected to a strip search, after controlling for age group, gender, and quarter of the arrest event?

We planned to conduct a logistic regression to model the relationship between strip search and perceived race while controlling for other factors (age group, gender, and quarter). By conducting a logistic regression, we will examine whether perceived race is a significant predictor of the likelihood of being a subject to a strip search during an arrest event.

Power Analysis

Subsequently, we will conduct a power analysis to identify the minimum sample size required. The primary objective is to identify how much data is required to detect a significant difference in the proportion of strip searches, so that we can achieve sufficient statistical power without false negative results.

Results and Findings

RQ1: Are there significant differences in the proportion of strip searches between different Perceived Race, while controlling for the age group of individuals arrested?

ANCOVA

A one-way ANCOVA was conducted to examine the effect of race on the likelihood of police conducting strip searches at time of arrest while controlling for age. The main effect of race on the likelihood of strip-search was significant, $F(5,60026) = 74.74$, $p < .001$, indicating that the likelihood of strip search being conducted at time of arrest significantly differed based on perceived race. The result also revealed a significant main effect of age on the likelihood of strip-search, $F(1, 60026) = 121.49$, $p < .001$, indicating that the likelihood of strip search being conducted at time of arrest significantly differed based on the age of the suspect.The covariate, age, accounted for a significant amount of variance in strip search proportions, with higher age associated with higher proportion of strip searches. No significant effects were found for the interaction between race and age, indicating that the proportion of strip searches significantly differs between different perceived races after controlling for the age.

Table 1

One-way ANCOVA results for RQ1

Predictor	Sum of Squares	*df*	F	*p*
C (Race)	39.427	5.00	74.742	<.001
Age	12.817	1.00	121.489	<.001
Residual	6332.849	60026.00		

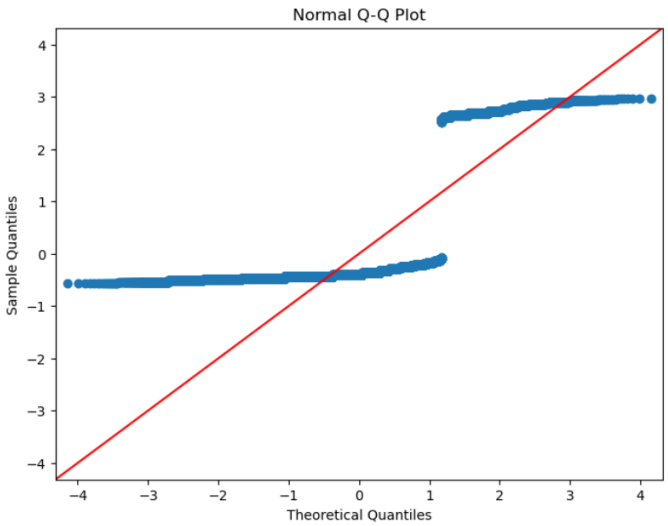
Normality Test

A quantile-quantile plot (Q-Q plot) was generated to assess the normality of the data, where the points are located more closely to the diagonal line indicating the data is more

normally distributed. The Q-Q plot showed non-linear points, representing that the data deviates from the normal distribution and is skewed to the left (see Figure 5).

Figure 5

Normal QQ Plot for One-Way ANCOVA

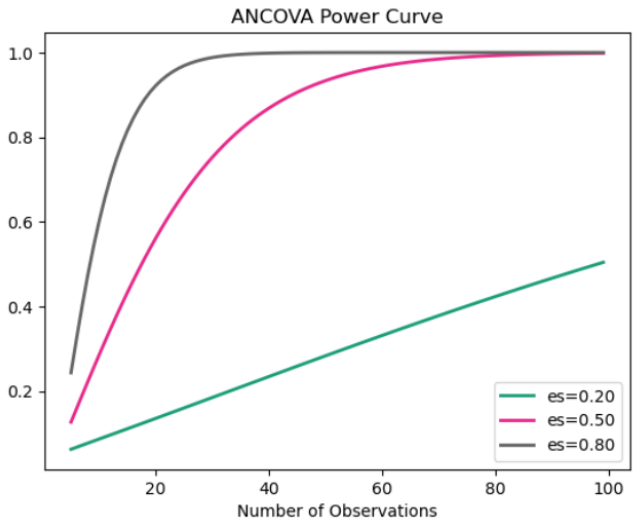


Power Analysis

A minimum effect size of 0.006 was calculated in the one-way ANCOVA, with a minimum sample size of 205025. Then a power curve was plotted to visualize the relationship between the effect size and number of observations (see Figure 6). The observation revealed that sample size is positively associated with power, meaning that as the likelihood of accepting the alternative hypothesis increases, the sample size also increases accordingly.

Figure 6

ANCOVA Power Curve



RQ2: Does the perceived race of individuals arrested predict the likelihood of being subjected to a strip search, after controlling for age group, gender, and quarter of the arrest event?

Logistic Regression

We conducted a logistic regression to examine the effect of race and age on predicting the likelihood of strip search at time of arrest while controlling for age group, gender, and quarter (see Table 2). The odds ratios were also calculated to indicate the odds of being strip searched among the groups of interest (see Table 3). The model was statistically significant, $\chi 2(10) = 1666.56$, $p < .001$, indicating that the perceived race variables significantly predicted the likelihood of strip search. The model explained 5.46% (pseudo $R^2 = .055$) of the variance in strip search, and correctly classified 87.62% of cases, indicating good predictive accuracy.

Results indicated that being Black (B = 0.68, $p < .001$), Indigenous (B = 0.87, $p < .001$), and White (B = 0.57, $p < .001$) were positively associated with strip search, whereas being Latino (B = -0.01, p = .897) and Middle Eastern (B = -0.09, p = .328) were not significantly associated with StripSearch. The odds ratios indicate that the odds of being strip-searched were higher for Black individuals (OR=7.21, 95% CI [1.777, 2.194]), Indigenous individuals (OR=10.96, 95% CI [2.022, 2.836]), and White individuals (OR=5.84, 95% CI [1.594, 1.954]),

Latino individuals (OR=2.68, 95% CI [0.788, 1.233]), and Middle Eastern individuals (OR=2.504, 95% CI [0.764, 1.094]) compared to the reference category Asian.

Additionally, being in Quarter 2 (B = -0.33, p < .001), Quarter 3 (B = -0.23, p < .001), and Quarter 4 (B = -1.87, p < .001) were negatively associated with strip search, whereas being in Quarter 1 was the reference category. The odds of being strip-searched were higher for individuals who were arrested in Quarter 4 (OR=1.167, 95% CI [0.137, 0.173]) compared to the reference categories. Moreover, odds of being strip-searched were higher for individuals who were arrested in Quarter 2 (OR=2.048, 95% CI [0.667, 0.771]) and in Quarter 3 (OR=2.221, 95% CI [0.745, 0.854]) compared to Quarter 1.

Being in the youth age group (B = -0.41, p < .001) was negatively associated with strip search compared to the adult age group. The odds of being strip-searched were higher for individuals in the adult age group (OR=1.944, 95% CI [0.571, 0.774]).

Lastly, being male (B = 0.27, p < .001) was positively associated with strip search compared to being female. The odds of being strip-searched were higher for male individuals (OR=3.713, 95% CI [1.217, 1.414]) compared to female individuals.

Table 2

Logistic Regression results for RQ2

Group	B	SE	*z*	*p*	[0.025	0.975]
Intercept	-2.2884	0.061	-37.520	<.001	-2.408	-2.169
Race[T.Black]	0.6805	0.054	12.653	<.001	0.575	0.786
Race[T.Indigenous]	0.8732	0.086	10.110	<.001	0.704	1.042
Race[T.Latino]	-0.0148	0.114	-0.129	0.897	-0.239	0.209
Race[T.Middle_Eastern]	-0.0895	0.091	-0.979	0.328	-0.269	0.090
Race[T.White]	0.5680	0.052	10.935	<.001	0.466	0.670
Age_group[T.Youth]	-0.4082	0.077	-5.270	<.001	-0.560	-0.256

Quarter[T.Q2]	-0.3326	0.037	-9.053	<.001	-0.405	-0.261
Quarter[T.Q3]	-0.2258	0.035	-6.455	<.001	-0.294	-0.157
Quarter[T.Q4]	-1.8701	0.059	-31.880	<.001	-1.985	-1.755
Gender[T.Male]	0.2714	0.038	7.115	<.001	0.197	0.346

Table 3

Odds Ratio results for RQ2

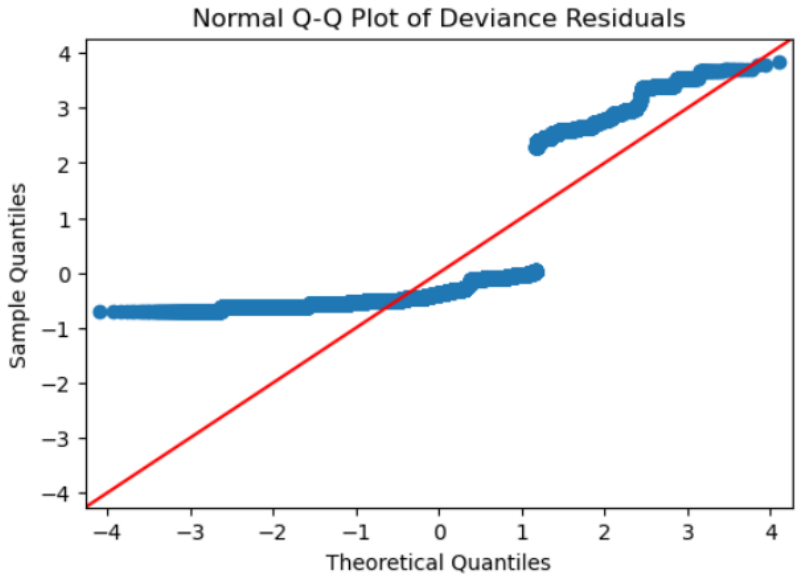
Group	Lower CI	Upper CI	*OR*
Intercept	0.090000	0.114307	1.106750
Race[T.Black]	1.777350	2.194507	7.206219
Race[T.Indigenous]	2.021660	2.836276	10.963532
Race[T.Latino]	0.787592	1.232729	2.678712
Race[T.Middle_Eastern]	0.764404	1.093810	2.495257
Race[T.White]	1.593954	1.953948	5.840377
Age_group[T.Youth]	0.571237	0.773870	1.944254
Quarter[T.Q2]	0.667233	0.770596	2.048391
Quarter[T.Q3]	0.744973	0.854473	2.220754
Quarter[T.Q4]	0.137369	0.172884	1.166616
Gender[T.Male]	1.217294	1.413597	3.712772

Normality Test

A quantile-quantile plot (Q-Q plot) was generated to assess the normality of the data, where the points are located more closely to the diagonal line indicating the data is more normally distributed. The Q-Q plot showed non-linear points, representing that the data deviates from the normal distribution and is skewed to the left (see Figure 7).

Figure 7

Normal QQ Plot for Logistic Regression

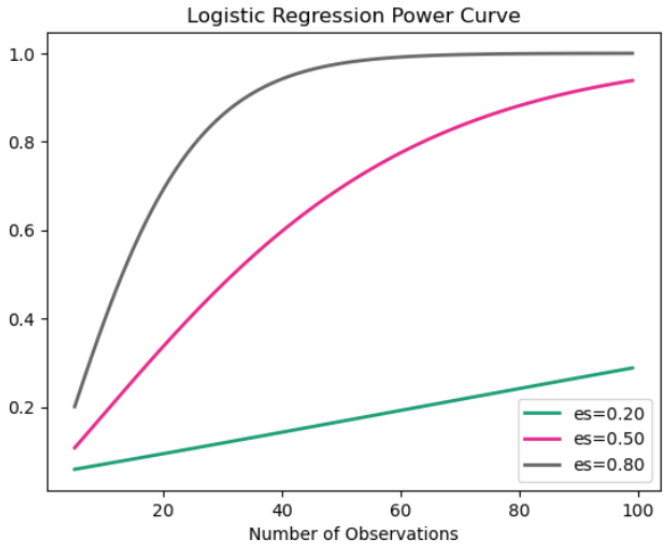


Power Analysis

A minimum effect size of 0.008 was calculated in the logistic regression analysis, with a minimum sample size of 236647. Then a power curve was plotted to visualize the relationship between the effect size and number of observations (see Figure 8). The observation revealed that sample size is positively associated with power, meaning that as the likelihood of accepting the alternative hypothesis increases, the sample size also increases accordingly.

Figure 8

Logistic Regression Power Curve



Discussion

The one-way ANCOVA results indicated that both race and age have a significant main effect on the proportion of strip searches during arrest events, suggesting that there is a significant difference in the likelihood of strip searches being conducted at the time of arrest based on the perceived race and age of the suspect. The covariate, age, was found to account for a significant amount of variance in strip search proportions, with higher age associated with a higher proportion of strip searches. Since no significant effects were found for the interaction between race and age, implying that the effect of race on the likelihood of strip search did not vary across different ages. Therefore, the proportion of being strip-searched has a significant difference for different individuals of different races even after controlling for age.

The logistic regression conducted in this study revealed that perceived race significantly predicted the likelihood of strip search, after controlling for age group, gender, and quarter of the year. Specifically, being Black, Indigenous, and White were positively associated with strip searches. This finding is consistent with prior research that has demonstrated racial disparities in policing practices, particularly with regard to strip searches (cite, year). Additionally, the results indicated that being in Quarters 2, 3, and 4 was negatively associated with strip searches, whereas being in Quarter 1 was the reference category. This finding suggests that seasonality may play a role in the likelihood of strip searches during an arrest. Further, being in the youth

age group was negatively associated with strip search compared to the adult age group. The odds of being strip-searched were higher for individuals in the adult age group. Lastly, being male was positively associated with strip search compared to being female, with the odds of being strip-searched being higher for male individuals. These findings indicate that men and adults are more likely to be subjected to strip searches during an arrest.

The present study has certain limitations that should be taken into account when interpreting the results. One limitation is that the age was categorized into intervals (e.g. 'Aged 17 years and younger', 'Aged 18 to 24 years'). It is difficult to conduct ANCOVA using a covariant without precise numbers. We took the mean for each category (e.g. consider all individuals from 'Aged 18 to 24 years' is aged 21.5 years) before proceeding to the analysis. The lack of information on the specific ages of the adult group may limit the ability to control the age precisely. Therefore, a follow-up study that further differentiates the age group by years is suggested to draw a more robust conclusion for the control of age in predicting the proportion of strip search.

Another limitation is the inconsistencies identified in the dataset. Specifically, some rows showed the same personID, suggesting that the same individual was involved in multiple arrest events. However, those rows contain different perceived races with the same personID, indicating a lack of consistency across the dataset. Moreover, the dataset contains a miscode in the “Booked” column. Based on the description, all individuals who have been strip-searched are booked within 24 hours. However, the data contains rows that showed positive in strip search and negative in booked. These discrepancies in the data raise concerns about the reliability and validity of the dataset and may limit the extent to which accurate conclusions can be drawn from the data.

Another important limitation of the present study is that both one-way ANCOVA and logistic regression results do not meet the assumption of normal distribution. These violations of assumptions can potentially lead to inaccurate conclusions or misleading results, raising questions about the reliability and validity of the findings. In particular, violations of the normality assumption may affect the accuracy of the p-values, which may result in either type I or type II errors. Future research with a dataset that addresses these issues is suggested to improve the accuracy and generalizability of the findings.

The last limitation is that the minimum effect sizes obtained in the one-way ANCOVA and logistic regression were relatively small (0.006 and 0.008 respectively), and large minimum sample sizes (205025 and 236647 respectively) were required to detect those effect sizes with adequate power. Given that the dataset had only 60033 rows, it may be difficult to obtain such a large sample size in practice, which could limit the generalizability of the findings.

Conclusion

In conclusion, this study investigated the differences in the proportion of strip search between different races while controlling for the effects of age. The results showed that there was a significant difference in the proportion of strip searches between different races even after controlling for age. The present study also examined whether race predicts the likelihood of being strip-searched while controlling for age group, gender, and quarter of the year. The results revealed that being Black, Indigenous, or White individuals are positively associated with the likelihood of being strip-searched at the time of arrest. These findings highlight the importance of addressing racial disparities in law enforcement, particularly with regard to strip searches. It is suggested that law enforcement agencies and policymakers with increased training and awareness regarding the impact of implicit bias and discrimination on their decision-making processes. Additionally, further research is needed to understand the role of seasonality and age in the likelihood of strip search during arrest, and to identify potential interventions to reduce disparities in this area of law enforcement.

References

Ha, D. (2010). Blanket Policies for Strip Searching Pretrial Detainees: An Interdisciplinary Argument for Reasonableness. Fordham L. Rev., 79, 2721.

Newburn, T., Shiner, M., & Hayman, S. (2004). Race, Crime and Injustice? Strip Search and the Treatment of Suspects in Custody 1. British Journal of Criminology, 44(5), 677-694. Chicago

The Canadian Press. (2019). Ontario's police watchdog says strip searches are happening far too often. CBC News.

Toronto Police Service. (2022). Arrests and Strip Searches (RBDC-ARR-TBL-001). PUBLIC SAFETY DATA PORTAL.

Visher, C. A. (1983). Gender, police arrest decisions, and notions of chivalry. Criminology, 21(1), 5-28.