An exploration of the role of Race on Strip Searches Conducted by the Toronto Police Services

by

Shuang (Sharon) Liang, Tashfia Nidhi

A final paper submitted in conformity with the requirements
for the course of INF2178

Faculty of Information
University of Toronto

Table of Contents

1.Introduction

1.1 Overview

Since the dawn of mankind, power dynamics have played a vital role in shaping societal practise of law enforcement. Law enforcement professionals such as police officers are vested with both immense responsibility and power to maintain societal safety and order. One of the many powers police officers do hold is the power to conduct a strip search on an arrested or detained individual. In Canadian policing, a strip-search is referred to as "the removal or rearrangement of some or all of the clothing of a person as to permit a visual inspection of a person's private areas, namely genitals, buttocks, breasts… or undergarments" (Lemke, 2022). These highly intrusive police-led strip searches are conducted usually in the absence of consent of the searched person and not only does this practise strip away clothes from arrested individuals but it also strips away their dignity and autonomy. A 2019 report by the Ontario Independent Police Review Director indicated that Toronto Police Services have some of the highest strip search conducting rates in the nation as between 37% to 43% of all arrestees between 2014 and 2016 were strip searched (McNeilly, 2018). Several studies have highlighted the disproportionately high representation of Black and Indigenous ethnic demographics in datasets about police practises such as carding and strip searches. In light of these crucial race-related information about police practises, this research paper is aimed at utilizing a Toronto Public Service dataset to explore the role race plays both single handedly and concurrently with other factors in regards to the conduction of strip searches. Key findings from this research paper includes the evidence of differences in strip-searches based on race as well as the evidence of an

interaction effect existing between race and occurrence category with respect to strip-search counts.

1.2 Literature Review

Research on racial and gender biases in policing practises have been conducted since decades, yet, the disparities and disproportionate representation and exploitation has persisted. Canada is no stranger to these sorts of statistics. Various studies have led to the findings that indigenous group (particularly females) and black racial demographic face the highest rates of intrusive police practises such as strip searches. “Reforming the Strip Search in Canadian Law and Practice” by Benjamin L. Berger (2009) found that strip searches are frequently used by polices in Canada even in cases where it’s not legally justified. In 2016, the Ontario Human Rights Commission (OHRC) released a report on the over-representation of Indigenous peoples in Ontario’s criminal justice system. The report also found that Indigenous peoples in Ontario are over-represented in strip search statistics, and that they are more likely to be subjected to a strip-search than non-Indigenous people (Robinson, 2016). The gruesome statistics associated with strip searches are not limited to Ontario. A 2018 report by BC Civil Liberties Association found that strip-searches were still being conducted in degrading and inhumane manner for women (Mazur, 2021). A 2018 Toronto Star report highlighted that despite Black people making up 8.3% of Toronto’s demographic between 2010-2017, they accounted for 25.8% of all carding incidents, thus highlighting another police practise where there is evidence of discrimination (Rankin & Gillis, 2018).

1.3 Dataset Description

This project utilizes a dataset openly released by the Toronto Police Service in November 2022 containing various information on the arrests and strip searches for 65 276 arrested individuals. This data was collected over the span of 2020 to 2021. There are a total of 24 attributes in the dataset, including demographic identifiers like perceived race, age category, sex as well as other information about the arrest such as the occurrence category, search reasons, actions at arrest and whether or not the booked individual was strip searched. The format of data varies across the various variables; binary variables are used to indicate booking, strip search events, the different search reasons and actions at arrest. Categorical text variables are utilized for all demographic identifiers and occurrence categories. ID variables, arrest location, and the year of arrest were conveyed via numerical variables.

1.4 Research Questions

As a result of all the race related revelations encountered while conducting background research as well as due to personal interest, the overarching research topic for this paper is regarding the role race plays both single-handedly and concurrently with other factors in regard to the conduction of strip searches. The two specific research questions that will be explored and answered in this paper are:

1) How does racial group influence the likelihood of an arrested individual being strip-searched?

2) How does racial group and occurrence category interact to influence the likelihood of an arrested individual being strip-searched?

2.Exploratory Data Analysis

2.1 Descriptive Statistics - General

To gain a better understanding of the processed and cleaned dataset, a summary chart (figure 1) showing the number of values for each non-numerical variable was generated. As evident in figure 1, there are a total of 2573 rows, 8 different racial categories, 3 different gender groups, 9 age-group categories and 31 occurrence causes.

	Perceived_Race	Sex	Age_group__at_arrest_	Occurrence_Category
count	2573	2573	2573	2573
unique	8	3	9	31
top	White	M	Aged 25 to 34 years	Warrant
freq	423	1511	452	114

Figure 1: Summarized Table of Non Numerical variables

2.2 Race related EDA

Since a significant amount of background research heralded findings revolving around the role of race in influencing police practice, the perceived race was chosen to be the pivotal variable of interest in this paper. As a result, a considerable part of the EDA was aimed at better understanding the racial makeup and relationships within this particular dataset by visualizing various aspects of the racial data especially as it related to strip searches.

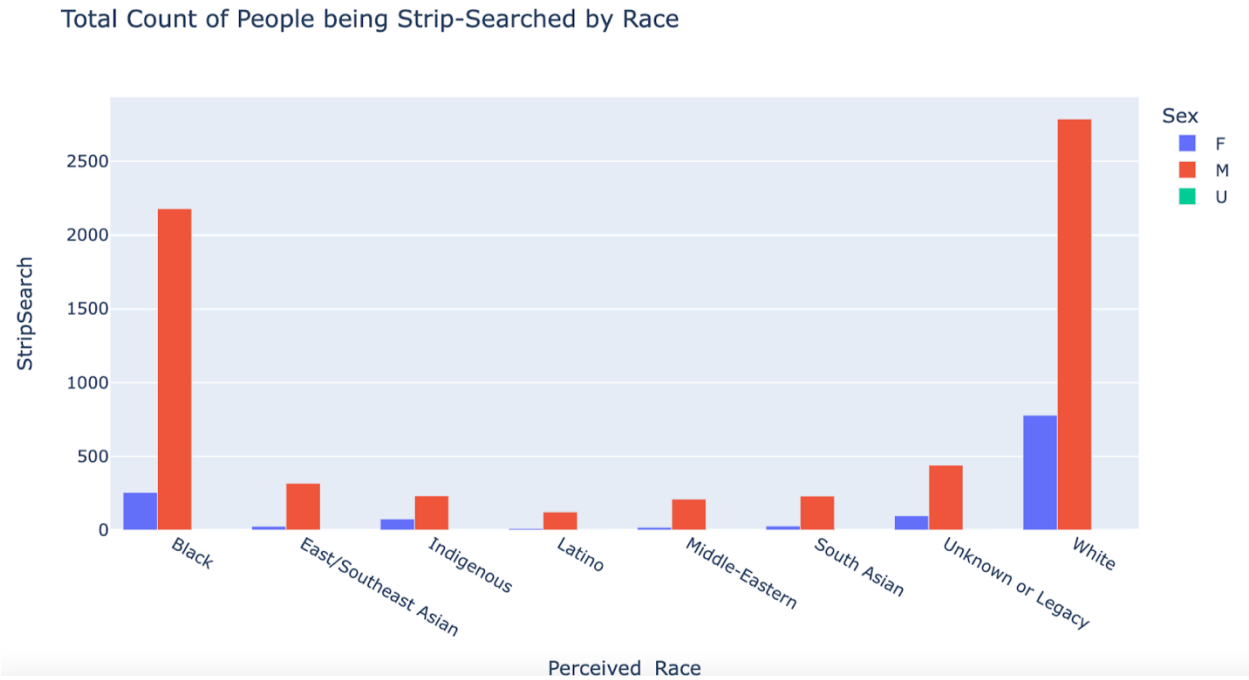

Figure 2: Bar Plot Illustrating the total count of strip search conducted per perceived race by gender

Figure 2 illustrates the total counts of strip searches conducted under each different perceived racial groups and gender. It can be observed that the number of males being strip searched is higher than the number of females across all racial groups. This is in line with Toronto crime statistics which shows males committing a greater percentage of crimes than females. (CITATION). The perceived race that had the highest strip search counts was White, followed by Black while Latino had the lowest total counts for both genders. Keeping Toronto's population demographic in mind where white ethnicity is the largest ethnic group, it is no surprise that the perceived 'white' race has the highest strip search count. The 2nd largest perceived race represented in this dataset is 'Black' which doesn't align with the city of Toronto's ethnic makeup where ethnic demographics such as East Asians and South Asians have a greater population than Black.

To gain a more comprehensive idea about the relationship between racial background and race, the mean, minimum, maximum, and standard deviation of the strip searches conducted for each perceived race (determined by looking at the strip counts for different gender, age group, occurrence category combination under each perceived race) was checked. This is illustrated below.

	StripSearch			
	mean	min	max	std
Perceived_Race				
Black	6.305699	0	177	18.707310
East/Southeast Asian	1.014881	0	22	2.802411
Indigenous	1.275000	0	21	3.053121
Latino	0.503817	0	14	1.513029
Middle-Eastern	0.805654	0	14	2.075107
South Asian	0.850993	0	23	2.667819
Unknown or Legacy	1.568915	0	40	4.193002
White	8.430260	0	168	22.396031

Figure 3: Descriptive Statistics for Perceived Race and Strip Search

Similar to the total strip search count bar plot above, Figure 3 indicates the perceived race 'White' as having the highest mean (M=8.43) strip search count across the different group combinations followed by 'Black' (M=6.31). Ignoring the unknown/legacy category under perceived race, it was interesting to note that the 'Indigenous' perceived race has the 3rd highest mean strip search count. However, this wasn't surprising since background research did highlight indigenous groups as one of the main demographics that are often over-represented in data associated with police practices.

These perceived race-related findings led to further exploration of the racial relationship associated with strip searches.

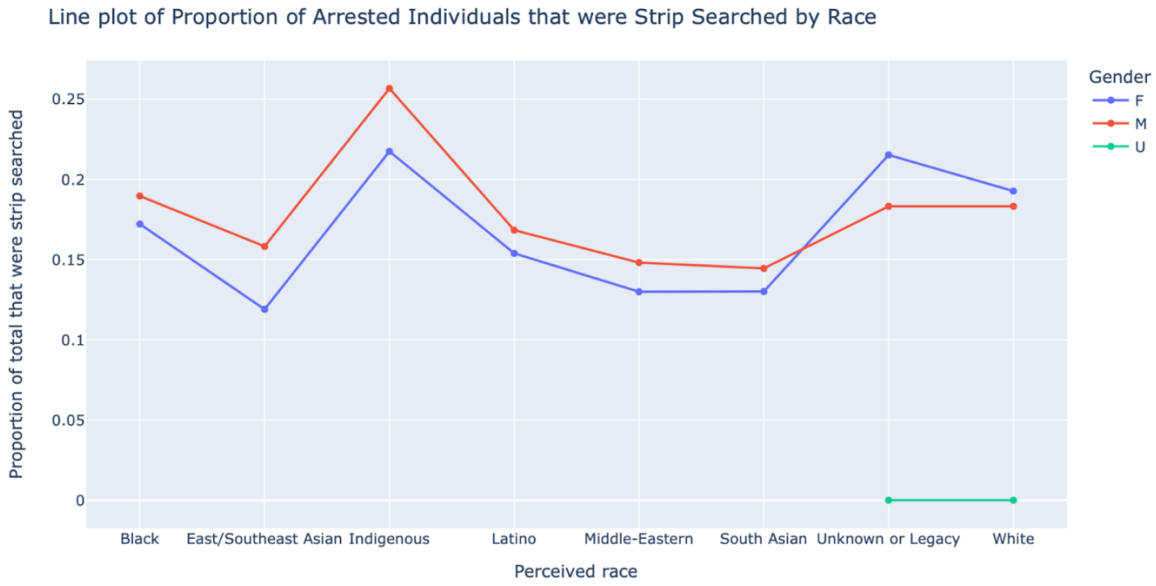

Figure 4: Line Plot displaying the proportions of arrested individuals that were strip searched.

Looking at the proportions of arrested individuals that were strip searched for each of the racial groups provides comparative insight into the likelihood of an individual being strip searched based on their race. Surprisingly, neither White or Black racial groups had the highest

proportion, but rather it is the Indigenous group that has the highest proportion of being strip searched for both genders. This is in line with the background research. Another interesting revelation from this plot is that while a higher proportion of arrested white males were strip searched in comparison to arrested black males, a higher proportion of arrested black females were strip searched in comparison to arrested white females. This is also in line with background research that revealed females from minority groups being over-represented in strip search data.

2.3 Occurrence Category EDA

Aside from racial groups, the occurrence category is another variable that could provide interesting insights into factors influencing strip search likelihood. There are a total of 31 occurrence categories.

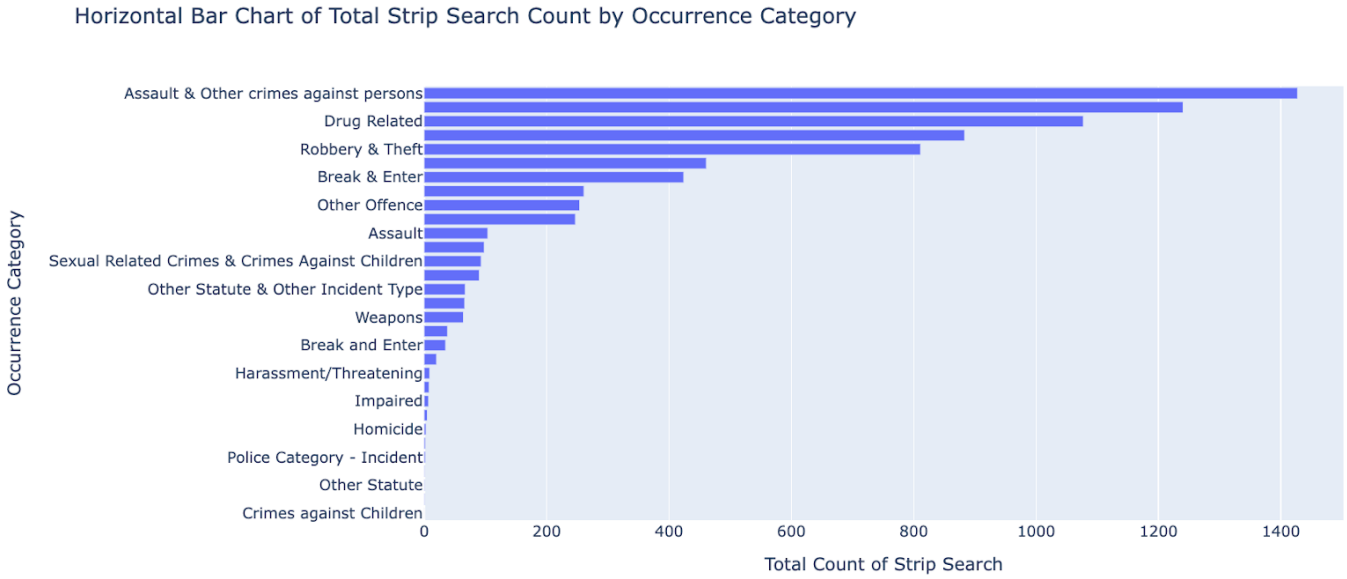

Figure 5: Bar Chart displaying the total strip search counts for the top 16 Occurrence Categories

As illustrated in figure 5, the highest number of strip search counts were conducted for individuals arrested for assault and other crimes against persons, followed by drug-related arrests and then robbery and theft.

2.4 T-tests

To further explore the relationship between the different perceived races on strip search practise, a series of T-tests were conducted to determine whether a statistically significant difference exists between the strip search count means of different perceived races. Before the tests were run, it was ensured that the following assumptions were fulfilled: 1) normality assumption; 2) independence of error; 3) A two-level explanatory variable that's nominal and 4) a continuous outcome variable.

2.4.1 Strip search counts between Black racial background and all other racial groups

As a result of background findings regarding the biases against Black individuals in Canadian police practise as well as the previous bar graphs highlighting the strip counts of Black ethnic group to be relatively high compared to other minority groups, the initial Welch's T-test conducted is aimed at checking whether there is a statistically significant difference between the mean strip-search count (outcome variable) of Black in comparison to the remainder of the 7 racial groups. The hypothesis that are being tested are:

H0 (Null Hypothesis): The mean strip-search counts between black individuals is the same as the combined mean strip-search counts for all other racial groups.

HA (Alternative Hypothesis): There is a statistically significant difference in the mean strip-search counts between Black individuals and other races.

The Welch's T-test results indicate a p-value of 9.66e-05 which is lower than a critical value of 0.05 (95% confidence interval). This leads to the rejection of the null hypothesis. This indicates that there indeed is a statistically significant difference in the mean strip search count of Black individuals (M=6.3, SD=18.7) and all other racial groups (M=2.5, SD = 10.6).

2.4.2 Strip search counts between Indigenous racial background and all other racial groups

With a long history of prejudice and injustice against Canada's Indigenous demographic by the police system, it was of practical interest to look at whether the mean strip search counts varied between the indigenous racial group in comparison to all other racial groups. As a result, a 2nd Welch's T-test was conducted. The hypotheses being tested are:

H0 (Null Hypothesis): The mean strip-search counts between indigenous individuals is the same as the combined mean strip-search counts for all other racial groups.

HA (Alternative Hypothesis): There is a statistically significant difference in the mean strip-search counts between Indigenous individuals and other races.

The T-test results indicate a p-value of 5.59e-09 which is below the critical value of 0.05 thus leading to the rejection of the null hypothesis. Therefore, it can be said that there is a statistically significant difference in the mean strip search count of indigenous individuals (M=1.28, SD = 3.05) and all other racial groups (M=3.21, SD = 12.8).

2.4.3 Strip search counts between Black + Indigenous racial backgrounds and all other racial groups

Since background research revealed Black and Indigenous ethnic groups as the two most likely groups to experience higher rates of incarceration, and intrusive police practices, this Welch's T-test was conducted to determine whether there is a statistically significant difference in the mean strip search count between indigenous + Black racial groups compared to all other racial groups.

H0 (Null Hypothesis): The mean strip-search counts between arrested Black or Indigenous individuals is the same as the combined mean strip-search counts for all other racial groups.

HA (Alternative Hypothesis): There is a statistically significant difference in the mean strip-search counts between arrested black or indigenous individuals and other races.

Tne T-test results indicate a p-value of 0.0064 which is below the critical value of 0.05 thus leading to the rejection of the null hypothesis. This indicates that there indeed is a statistically significant different in the mean strip search count for black and indigenous demographic combined (M = 4.38, SD = 15.0) and all other racial groups (M = 2.60, SD = 11.1).

2.5 Remarks about EDA

This series of Welch's T-tests conducted does illustrate a statistically significant difference in the strip search counts between several racial groups. This was no surprise based on the background research where it was learnt that an arrested individual's perceived race does play an essential role in police treatment towards them. To further confirm and consolidate the findings of these preliminary tests, the first research question of this project will explore how racial groups influence the likelihood of being strip searched. Furthermore, this paper will also explore the relationship between the Occurrence category and race with respect to the strip searches.

3.Method

Based on the exploratory data analysis, background research and personal interest, the primary focus of this paper will be on the relationship between race and strip search conduction. The table below summarizes the research questions and the specific methods utilized to answer each research question.

Research Question	**Method**
How does racial group influence the likelihood of an arrested individual being strip-searched?	- One Way Anova - Post Hoc Test (Tukey's HSD)
How does racial group and occurrence category interact to influence the likelihood of an arrested individual being strip-searched?	- Two Way Anova - Post Hoc Test (Tukey's HSD) - Interaction Effect Plot

Figure 6: Table of Research Questions and selected method of exploration

2.6 Anova Assumptions

In order to answer the two research questions, both one-way and two-way Anova were conducted. The assumptions for conducting an anova are: 1) normality of sample, 2) sample independence, 3) equal variance across groups and 4) a continuous outcome variable.

2.7 Research Question 1

2.7.1 One-Way Anova Overview

In order to answer the first research question, an One-Way anova was conducted where the perceived race behaved as the categorical predictor variable and the strip search counts behaved as the continuous response variable. The findings of the Anova will reveal insights into whether the racial differences in strip search counts are statistically significant.

2.7.2 About the Categorical Predictor Variable

There are 8 different racial groups under the perceived race variable (the predictor variable). The 8 races are White, Black, East/Southeast Asian, Indigenous, Middle Eastern, Latino, South Asian and Unknown.

2.7.3 Hypothesis Testing

The following hypotheses are being tested:

H0 (Null hypothesis 1): There is no difference in the mean strip search counts between the different perceived racial groups.

H1 (Alternate Hypothesis 1): There is a difference in the mean strip search counts between at least 2 of the different perceived racial groups.

2.7.4 Checking Assumption

Before conducting the Anova, the homogeneity of variance assumption was checked which revealed a ratio of 1.43. Since 1.43 < 2.00, this variance homogeneity assumption is met.

2.7.5 Anova Results

	df	sum_sq	mean_sq	F	PR(>F)
C(Perceived_Race)	7.0	23815.384566	3402.197795	24.160518	4.749578e-32
Residual	2565.0	361194.128066	140.816424	NaN	NaN

Figure 7: Summarized Table of One-Way Anova Results

As illustrated in Figure ___, the degree of freedom in this particular Anova test was 7 and a F value of 24.16 was attained. This relatively large F value means that the ratio of explained variance to unexplained variance is greater which elucidates towards the rejection of the null hypothesis. A p-value of 4.75e-32 was also attained which is the p-value associated with this particular F-distribution. As this value is much lower than a critical P value of 0.05 (indicating 95% confidence interval), it can be said that there is a statistically significant difference in the mean strip search counts between at least 2 of the different perceived racial groups.

2.7.6 Post-hoc Test

Since the One-way Anova heralded a statistically significant difference in the average strip search count means across the racial groups, a post-hoc test, specifically the Tukey's HSD test was conducted to determine which particular pair of racial groups have statistically significant difference in their mean strip search counts. The Tukey's HSD test was chosen over other post-hoc tests since it is both robust and conservative as it has a corrective feature to prevent type 1 error.

2.7.7 Tukey's HSD Results

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Black	East/Southeast Asian	-5.2908	0.001	-7.9765	-2.6051	True
Black	Indigenous	-5.0307	0.001	-7.9897	-2.0717	True
Black	Latino	-5.8019	0.001	-8.6832	-2.9205	True
Black	Middle-Eastern	-5.5	0.001	-8.317	-2.6831	True
Black	South Asian	-5.4547	0.001	-8.2201	-2.6894	True
Black	Unknown or Legacy	-4.7368	0.001	-7.4119	-2.0616	True
Black	White	2.1246	0.1775	-0.4092	4.6583	False
East/Southeast Asian	Indigenous	0.2601	0.9	-2.7821	3.3023	False
East/Southeast Asian	Latino	-0.5111	0.9	-3.4778	2.4557	False
East/Southeast Asian	Middle-Eastern	-0.2092	0.9	-3.1135	2.695	False
East/Southeast Asian	South Asian	-0.1639	0.9	-3.0181	2.6903	False
East/Southeast Asian	Unknown or Legacy	0.554	0.9	-2.2129	3.321	False
East/Southeast Asian	White	7.4154	0.001	4.7849	10.0459	True
Indigenous	Latino	-0.7712	0.9	-3.9874	2.4451	False
Indigenous	Middle-Eastern	-0.4693	0.9	-3.628	2.6893	False
Indigenous	South Asian	-0.424	0.9	-3.5367	2.6887	False
Indigenous	Unknown or Legacy	0.2939	0.9	-2.739	3.3268	False
Indigenous	White	7.1553	0.001	4.2463	10.0642	True
Latino	Middle-Eastern	0.3018	0.9	-2.7842	3.3879	False
Latino	South Asian	0.3472	0.9	-2.6919	3.3862	False
Latino	Unknown or Legacy	1.0651	0.9	-1.8921	4.0223	False
Latino	White	7.9264	0.001	5.0965	10.7564	True
Middle-Eastern	South Asian	0.0453	0.9	-2.9327	3.0234	False
Middle-Eastern	Unknown or Legacy	0.7633	0.9	-2.1312	3.6578	False
Middle-Eastern	White	7.6246	0.001	4.8603	10.3889	True
South Asian	Unknown or Legacy	0.7179	0.9	-2.1264	3.5622	False
South Asian	White	7.5793	0.001	4.8675	10.291	True
Unknown or Legacy	White	6.8613	0.001	4.2416	9.481	True

Figure 8: Tukey's HSD test results for One Way Anova (Mean strip search count by Racial categories)

To determine which pairs of perceived racial groups have a statistically significant difference in their mean strip search counts, the 'reject' column of the Tukey's HSD test table was utilized since a True in this column indicates the rejection of the null hypothesis while a false indicates the failure to reject the null hypothesis. Some highlighted racial group pairings that did have a statistically significant difference in their mean strip search counts include Indigenous and white, Black and indigenous, East/Southeast Asian and Black. Surprisingly, white racial group appeared to have a statistically significant difference in the mean strip counts with all other races except for Black.

2.7.8 Limitations

Although the anova tests heralded statistically significant results, there are some limitations that compromises the validity of these results. As a result of the dataset's nature, the different groups being tested were not normally distributed. The Normality Assumption Check heralded the following probability plots (figure 9). As evident in these plots, the regression line doesn't go through most of the data points thus indicating that the data for the different racial demographics weren't normally distributed, thus failing one of the essential assumptions for an Anova.

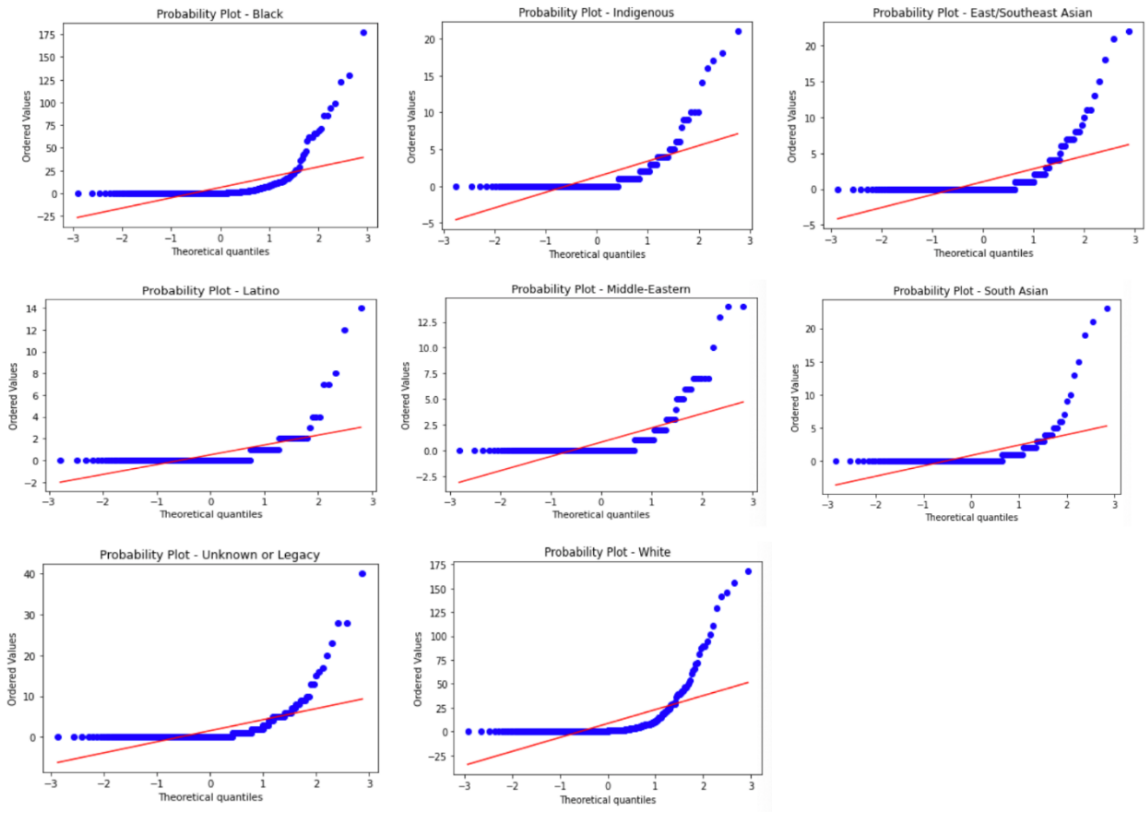

Figure 9: Probability plots for Normality Assumption Check for Perceived Race variable

2.8 Research Question 2

2.8.1 Two-Way Anova Overview

Since the aim of the 2nd research question is to explore the combined effect of perceived race and occurrence category on strip search counts, a 2-way Anova was conducted. For this 2-way Anova, the racial group and occurrence category behaved as the two categorical predictor variables while the strip search counts behaved as the continuous response variable.

2.8.2 About the Categorical Predictor Variables

-Perceived race → There are 8 different racial groups under the perceived race variable (the predictor variable).
-Occurrence category → There are 31 different types of occurrences under the occurrence category variable.

2.8.3 Hypotheses Testing

There are 3 sets of hypotheses being tested with this two-way Anova. They are:

Racial groups

H0 (Null hypothesis 1): There is no difference in the mean strip search counts between the different perceived racial groups.

H1 (Alternate Hypothesis 1): There is a difference in the mean strip search counts between at least 1 of the different perceived racial groups.

Occurrence category:

H0 (Null hypothesis 2): There is no difference in the mean strip search counts between the different occurrence categories.

H1 (Alternate Hypothesis 2): There is a difference in the mean strip search counts between at least 1 of the occurrence categories.

Interaction:

H0 (Null hypothesis 3): There are no statistically significant differences in the mean strip count for the different perceived races that can be explained by the occurrence category and vice versa. There is no interaction.

H1 (Alternate hypothesis 3): There is atleast one statistically significant difference in the mean strip count for the different perceived races that can be explained by the occurrence category and vice versa. There is interaction.

2.8.4 Anova Results

	sum_sq	df	F	PR(>F)
C(Perceived_Race)	25824.927551	7.0	33.041153	4.619568e-44
C(Occurrence_Category)	45449.806286	30.0	13.568283	1.042411e-61
C(Perceived_Race):C(Occurrence_Category)	56141.839924	210.0	2.394316	1.412906e-22
Residual	259602.481856	2325.0	NaN	NaN

Figure 10: Summarized Two-Way Anova Results

The perceived race predictor variable had a degree of freedom of 7, an F statistic of 33.04 and a p value of 4.62e-44. As this p value is much lower than the critical p value of 0.05 (indicating 95% confidence interval), the null hypothesis is rejected. Therefore, it can be said that there is a statistically significant difference in the mean strip search counts between at least 2 of the different perceived racial groups.

The occurrence category predictor variable had a degree of freedom of 7, an F value of 13. 57 and a p value of 1.04e-61. This relatively large F value means that the ratio of explained variance to unexplained variance is great which elucidates towards the rejection of the null hypothesis. Since the p value is also lower than the critical p value of 0.05 (indicating 95% confidence interval), the null hypothesis is rejected. Therefore, it can be said that there is a statistically significant difference in the mean strip search counts between at least 2 of the different occurrence categories

For the interaction term, the degree of freedom is 210, the F value is 2.39 and the p value is 1.41e-22. As this p value is lower than 0.05, the null hypothesis is rejected. Therefore, it can be said that there is an interaction effect occurring where the mean strip search count for one of the two predictor variables is influenced by the other predictor variable. This anova finding is pivotal in understanding the interrelated role racial categories and occurrence category plays together in influencing an arrested individuals' likelihood of being strip-searched.

2.8.5 Post-hoc test

Since the Two-way Anova indicated a statistically significant difference in mean strip-search counts for both predictor variables, the Tukey's HSD test was conducted as a post-hoc test to determine which particular pairs within each categorical variables differ from each other with respect to their mean strip search counts. The Tukey's HSD test was chosen over other post-hoc tests since it is both robust and conservative as it has a corrective feature to prevent type 1 error.

2.8.6 Tukey's HSD result for Perceived Race

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Black	East/Southeast Asian	-5.2908	0.001	-7.9765	-2.6051	True
Black	Indigenous	-5.0307	0.001	-7.9897	-2.0717	True
Black	Latino	-5.8019	0.001	-8.6832	-2.9205	True
Black	Middle-Eastern	-5.5	0.001	-8.317	-2.6831	True
Black	South Asian	-5.4547	0.001	-8.2201	-2.6894	True
Black	Unknown or Legacy	-4.7368	0.001	-7.4119	-2.0616	True
Black	White	2.1246	0.1775	-0.4092	4.6583	False
East/Southeast Asian	Indigenous	0.2601	0.9	-2.7821	3.3023	False
East/Southeast Asian	Latino	-0.5111	0.9	-3.4778	2.4557	False
East/Southeast Asian	Middle-Eastern	-0.2092	0.9	-3.1135	2.695	False
East/Southeast Asian	South Asian	-0.1639	0.9	-3.0181	2.6903	False
East/Southeast Asian	Unknown or Legacy	0.554	0.9	-2.2129	3.321	False
East/Southeast Asian	White	7.4154	0.001	4.7849	10.0459	True
Indigenous	Latino	-0.7712	0.9	-3.9874	2.4451	False
Indigenous	Middle-Eastern	-0.4693	0.9	-3.628	2.6893	False
Indigenous	South Asian	-0.424	0.9	-3.5367	2.6887	False
Indigenous	Unknown or Legacy	0.2939	0.9	-2.739	3.3268	False
Indigenous	White	7.1553	0.001	4.2463	10.0642	True
Latino	Middle-Eastern	0.3018	0.9	-2.7842	3.3879	False
Latino	South Asian	0.3472	0.9	-2.6919	3.3862	False
Latino	Unknown or Legacy	1.0651	0.9	-1.8921	4.0223	False
Latino	White	7.9264	0.001	5.0965	10.7564	True
Middle-Eastern	South Asian	0.0453	0.9	-2.9327	3.0234	False
Middle-Eastern	Unknown or Legacy	0.7633	0.9	-2.1312	3.6578	False
Middle-Eastern	White	7.6246	0.001	4.8603	10.3889	True
South Asian	Unknown or Legacy	0.7179	0.9	-2.1264	3.5622	False
South Asian	White	7.5793	0.001	4.8675	10.291	True
Unknown or Legacy	White	6.8613	0.001	4.2416	9.481	True

Figure 11: Tukey's HSD test result for Racial categories

The Tukey's HSD results for racial groups resemble the Tukey's HSD test results obtained for the one way anova used to answer research question #1.

2.8.7 Tukey's HSD result for Occurrence Category

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Assault	Assault & Other crimes against persons	12.5081	0.001	6.5481	18.4682	True
Assault	Break & Enter	5.3742	0.4161	-1.4084	12.1569	False
Assault	Break and Enter	-0.4238	0.9	-7.2382	6.3906	False
Assault	Crimes against Children	-0.9541	0.9	-9.8475	7.9392	False
Assault	Drug Related	9.9247	0.001	3.8588	15.9906	True
Assault	FTA/FTC, Compliance Check & Parollee	11.3231	0.001	5.2887	17.3574	True
Assault	FTA/FTC/Compliance Check/Parollee	-0.0358	0.9	-6.1179	6.0463	False
Assault	Fraud	-0.94	0.9	-7.6033	5.7232	False
Assault	Harassment & Threatening	2.0808	0.9	-4.2208	8.3823	False
Assault	Harassment/Threatening	-0.8552	0.9	-7.0593	5.3488	False
Assault	Homicide	-0.8541	0.9	-9.8621	8.1539	False
Assault	Impaired	-0.8666	0.9	-7.2989	5.5657	False
Assault	LLA	-0.939	0.9	-7.7534	5.8754	False
Assault	Mental Health	-0.919	0.9	-8.0607	6.2226	False
Assault	Mischief	-0.7319	0.9	-6.9547	5.4909	False
Assault	Mischief & Fraud	1.7307	0.9	-4.455	7.9163	False
Assault	Other Offence	1.5607	0.9	-4.4736	7.5951	False
Assault	Other Statute	-0.9421	0.9	-7.307	5.4228	False
Assault	Other Statute & Other Incident Type	-0.2337	0.9	-6.4013	5.9339	False
Assault	Police Category - Administrative	-0.5499	0.9	-6.6997	5.6	False
Assault	Police Category - Incident	-0.9275	0.9	-7.4823	5.6273	False
Assault	Robbery & Theft	6.4186	0.0151	0.5138	12.3234	True
Assault	Robbery/Theft	-0.3486	0.9	-6.2669	5.5697	False

Figure 12: Tukey's HSD test result for Occurrence Category
*The full list is not shown for space reasons

The 2nd Tukey's HSD test conducted for this research question looks at the specific occurrence categories that have statistically significant differences in their mean strip search counts. The 'reject' column of this Tukey's HSD test result table is comprised of both true and false values which indicates that while there are statistically significant differences in the mean strip-search counts between some of the occurrence categories, there are also pairs of occurrence categories where there are no statistically significant differences in their mean strip search counts.

2.8.8 Limitations

Since the chances of an error in the hypothesis tests increases with increasing number of elements in a predictor variable, the presence of 31 elements in the occurrence category increases the likelihood of this error. Furthermore, as revealed while conducting the one way anova for research question 1, the data is not normally distributed which is a key assumption for Anova. These limitations hinders the strength of the claims made based on the Anova results.

2.8.9 Interaction Effect Plot

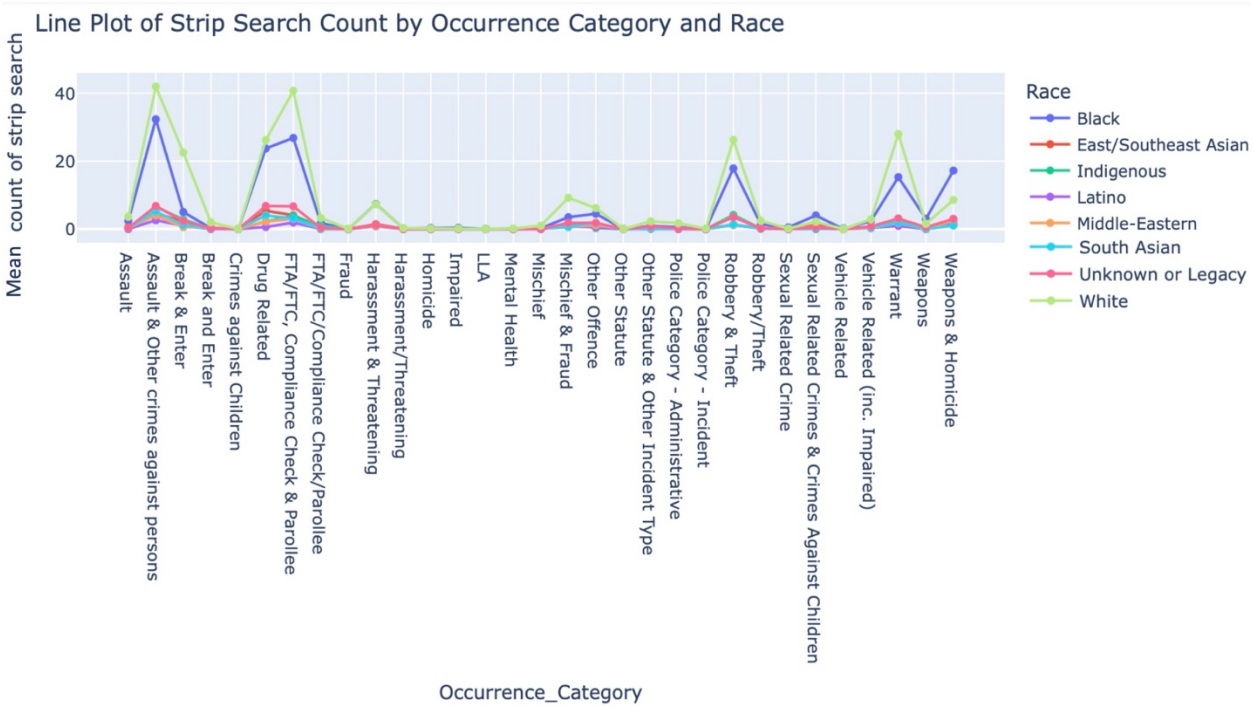

Figure 13: Interaction Effect Plot for Perceived Race and Occurrence Category with respect to Mean strip search count

A parallel line on the interaction plot indicates no interaction present while an intersecting point elucidates the presence of an interaction. In this plot, the highest interaction between perceived race and occurrence categories can be observed when the demographic is black and the occurrence category is Assault & other crimes against persons. This means that the highest number of strip searches conducted where when a Black individual was arrested for Assault or other crimes against persons or for FTA/FTC Compliance Check & Parollee. There is also notable interaction when the demographic is white, and the occurrence category is assault and other crimes against persons.

4.Discussion

The primary focus of this paper was to explore the influence of perceived race on the likelihood of being strip searched. There were several key findings that came out of the various statistical tests conducted. Not only are these findings essential in understanding this dataset, but they provide valuable insights into Toronto's policing practises. To answer the first research question (how does racial group influence an arrested individual's likelihood of being strip searched?), the preliminary t-tests, one-way Anova and Tukey's HSD post-hoc test needs to be taken into consideration. Based on background research as well as the preliminary T-tests and the strip search proportion graph (figure 4) produced, it was no surprise that there were statistically significant differences in mean strip search counts between the different races. While the highest

mean strip search count was observed in white arrestees, the difference in mean strip search count between white and black arrestees was not statistically significant, thus illustrating how despite Black individuals making up only 8.5% of Toronto's population, their strip search count is on par with white ethnic group who makes up 50.2% of Toronto's population (City of Toronto, 2021).

The second research question was aimed at exploring the combined interaction effect of occurrence category and perceived race on strip search. The statistically significant anova result did indicate that there is a interaction effect between occurrence category and perceived race on the strip search counts. Law enforcement practises such as police-led strip searches are multifactorial events that cannot be pinpointed to a singular predictor or cause, so it is important to look at various predictors to develop a comprehensive understanding of these practises.

Due to a lack of high quality data for other predictor variables as well as the complexities associated with conducting Anovas with more than 3 independent variables, the combined interaction effect of only 2 predictor variables were explored in this paper. However, for future investigation, the interaction effect of multiple predictor variables can be explored. Unfortunately, this is not the only limitation of this study. For both Anovas conducted, the assumptions weren't met as the data was not normally distributed. This hinders the strengths of the claims that were made based on the Anova results as there is an increased likelihood of error in hypothesis testing.

Thematic findings across the various statistical tests conducted includes the over-representation of Black and Indigenous individuals in strip search data relative to their demographic composition in the city of Toronto. There are various socio-economic and cultural factors that guide police interactions with different demographics. This research proposes similar findings as various other research exploring racial biases in police practises across North America. There are various societal and political implications of these findings. Studies such as this one indicates the need for acknowledgement of these biases as well as the necessity for proactive policies to combat the racial over-representation and biases that exist in policing practises.

5.Conclusion

With recent resurgence of movements like Black Lives Matter, there has been an increased light shone on various controversial policing practises, specifically policing practises on minority groups. The aim of this research was to investigate the role of race in strip searches conducted in Toronto over 2020 and 2021. The first research question was aimed to investigate whether there is a difference in mean strip search counts based on the perceived race of the arrested individuals. Statistical exploration leads to finding that there indeed is a difference in strip search count based on race as certain minority groups (Black, Indigenous) were seemed to be over-represented relative to their respective ethnic makeup in Toronto. The 2nd research question was aimed at exploring whether there is an interaction effect between Occurrence category and perceived race on the strip search count. The statistical tests conducted revealed that there indeed was an interaction effect present between these 2 factors. Due to the scope of

this project, other factors that may impact the conduction of strip searches weren't explored. To build on to the findings of this paper, those other factors such as age category can be explored as well. Studies of this nature are essential for providing insight into law enforcement practises within society and the findings can be used to guide policy changes in these areas.

6.Citations

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

Mazur, D. (2021, December 10). *Policing indigenous communities*. BC Civil Liberties Association. Retrieved February 28, 2023, from https://bccla.org/2017/12/policing-indigenous-communities/

Rankin, J., & Gillis, W. (2018, December 10). *Black people 'grossly overrepresented,' more likely to be hurt or killed by Toronto police, racial profiling report finds*. thestar.com. Retrieved February 28, 2023, from https://www.thestar.com/news/gta/2018/12/10/blacks-grossly-overrepresented-more-likely-to-be-hurt-or-killed-in-interactions-with-toronto-police-racial-profiling-interim-report-finds.html

Robinson, R. (2016, June 22). *2016 Report*. Office of the Independent Police Review Director. Retrieved February 28, 2023, from https://www.oiprd.on.ca/news/annual-reports-2016