

# Statistical Analysis Investigating Possible Intersections of Race and Gender with Strip Searching Practices by the Toronto Police Service.

By Jiawei Liu and Vlada Gorchkova

## 1 Introduction

When one is arrested they may be subject to a strip search performed by the arresting officers, however, this is a controversial and potentially humiliating practice and it can be seen as a violation of autonomy, individual privacy, and human rights. This practice is usually performed on arrestees or inmates, and involves all clothing to be removed for a visual inspection by an officer, intended to search for contraband or weapons that may be hidden on or within the body. The Toronto Police Service (TPS), which is responsible for patrolling and enforcing law in the Toronto Area, includes this practice as one potential procedure that occurs when someone is arrested. However, in recent decades, it has been suggested by activists and research alike that the TPS may be unfairly biased against individuals of certain racial, gender, or class backgrounds [1]. The purpose of this paper is to investigate if this bias against certain communities extends to TPS' likelihood to strip search an individual, and the implications that this may have for Toronto, the TPS, and the justice system in Canada at large.

First, we will provide a literature review on strip searching, the Toronto Police Service's history with discrimination, and the questions this may pose for the dataset we will be analyzing. Following, we will introduce the dataset, and discuss how it was prepared for analysis. Finally, we will analyze the dataset according to research questions that we prepared as an outcome of the literature review, and discuss the findings, followed by their implications.

## 2 Literature Review

Generally speaking, many have investigated the impact of racism (specifically anti-Black racism) in policing, leading to the over-policing of Black communities, and ultimately larger proportions of inmates from these communities [9]. In Canada, specifically, the policing of Indigenous communities and furthermore the sexualization of Indigenous women when policed, has also been disproportionate compared to White populations in Canada [11], meaning that there are different (and sometimes intersecting) identities that can affect the way that a person will be treated by Toronto Police officers.

Regarding strip searches specifically, some activists argue that strip searches can be considered a form of state-sanctioned sexual assault, as inmates are unable to consent to the procedure [7]. In the case of arrestees, they have not yet been processed or booked, and yet may have the procedure conducted on them, regardless of any proved criminal wrongdoing. In Toronto, this may be a particular issue as about 56% of the population of Toronto identified as visible minorities in 2021, with 9% of the population as Indigenous [4], meaning that there are a high number of people that are vulnerable in the face of over-policing and the use of force disproportionate to white communities. Recent reports suggest this conclusion, indicating that Black individuals are more likely to be strip searched compared to others [8], and as such, this is something worth investigating in our assessment as well.

Another angle to the disparate impact on communities that has been discussed in recent years is the use of police services in response to mental health crises or incidents. With the death of Regis Korchinski-Paquet [15], Ejaz Choudry [14], and Bobby Ramroop [10], many are now calling into question whether TPS behaves with disproportionate or inappropriate force when responding to calls regarding those suffering a mental health incident. Disability justice activists argue that racialized individuals are especially vulnerable when experiencing these incidents [5], and in 2022, special response teams were formed in Toronto to provide the opportunity for non-police related responses [13]. Ultimately, the behavior of arrestees should not impact their treatment. Whether one is behaving erratically or otherwise, they should be treated with reasonable fairness and dignity, and instead only logical reasons should be used to evaluate further procedures and processing, including strip searches. As the dataset records mental health incidents as one of the actions taken at arrest, this could be worth investigating.

### **3 Research Questions**

#### ***3.1 Research Questions***

1. Whether the number of actions that an arrestee takes when being arrested differs between arrestees that have been strip searched and what ways this may interact with race?
2. Whether the number of actions that an arrestee takes when being arrested differs between arrestees that have been strip searched and what ways this may interact with gender?
3. Whether the number of search reasons provided for a strip search vary depending on the gender of an arrestee?

4. Whether the number of search reasons provided for a strip search vary depending on the race of an arrestee?
5. Whether the intersection of race and gender has any unique interaction effects.

### **3.2 Hypotheses**

1. The race and number of actions that an arrestee takes will have an impact on their likelihood to be strip-searched.
2. The gender and number of actions that an arrestee takes will have an impact on their likelihood to be strip-searched.
3. The number of search reasons will vary depending on the gender of the arrestee that has been strip searched.
4. The number of search reasons will vary depending on the race of the arrestee that has been strip searched.
5. There will be interaction effects between race and gender.

## **4 Dataset**

To further investigate this issue, we will be analyzing a dataset including all Toronto Police Service data related to strip searches following arrest from 2020 to 2021 [12], provided by a Private Member on the TPS Open Data Portal. Variables that are recorded when someone is arrested include whether they were strip searched, their race, sex, offense, age group, among other demographics variables.

Furthermore, there are two other kinds of variables that are of interest in this analysis. The first is Search Reasons that are provided for conducting the strip search, and these are listed as a boolean for one or more of four different reasons (Caused Injury, Assisted Escape, Possessed Weapons, and Possessed Evidence). The second is Actions at Arrest, also listed as a boolean for one or more of five different kinds of actions (Combative, violent or spitter/biter; Resisted, defensive or escape risk; Mental instability or possibly suicidal; Assaulted officer; and Cooperative). To enhance the comprehensiveness of our data analysis, we have introduced two additional columns in our dataset. These columns are named "total actions at arrest" and "total strip search reasons," and they serve to summarize the values of "Actions at Arrest" and "Search Reasons," respectively.

Regarding race and sex characteristics, race is recorded as any of the following categories: Black, East/Southeast Asian, Indigenous, Latino, Middle-Eastern, South Asian, White, and Unknown or Legacy. Sex is recorded as either M(ale), F(emale), or U(known).

## 5 Exploratory Data Analysis

### 5.1 Descriptive Statistics

#### 5.1.1 Barplot to show the frequency of main variables

To explore the data, we began by creating barplots to visualize the frequency distribution of the related variables, including Race, Sex, and StripSearch, which allowed us to quickly identify any patterns or imbalances in the distribution of these variables in our dataset.

**Figure1:** Barplot to check frequency of Race

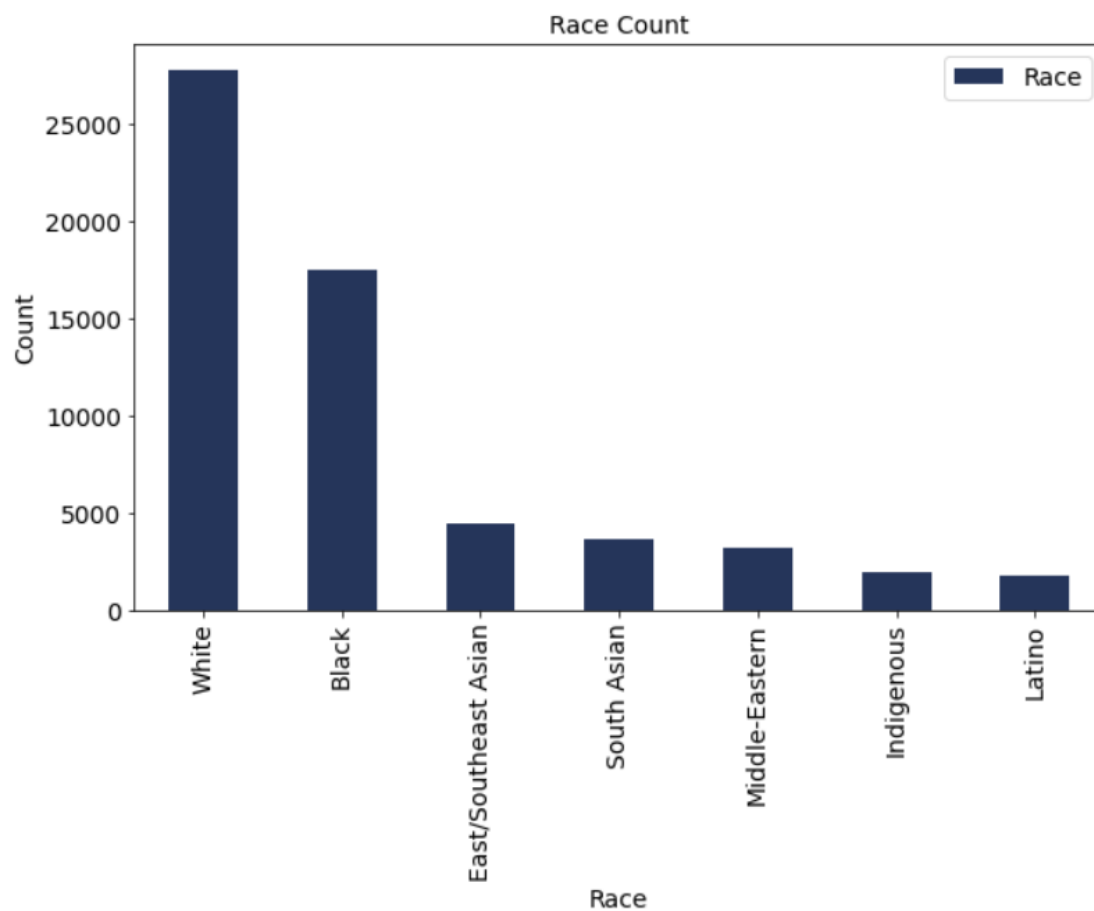


Figure1 contains information on the race of the recorded arrestees. We created this table to gain an understanding of the racial composition of our dataset. The plot shows 7 distinct categories, and the majority of recorded arrestees are White or Black people, with an amount of 27,718 and 17,526, respectively.

**Figure2:** Barplot to check frequency of Gender

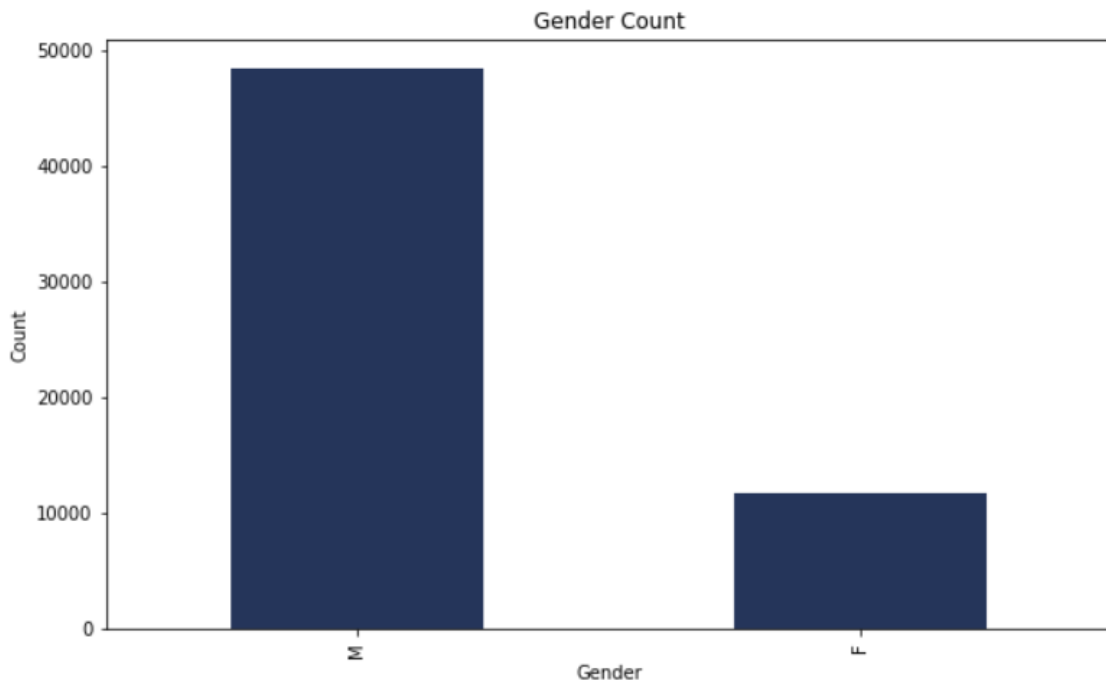


Figure2 displays the gender distribution of arrestees. There are 4 times more male arrestees (48523) than that of females (11691). The large difference in the number of male and female arrests highlights the potential impact of gender on the variable we are interested in, for example, the number of improper actions taken during an arrest. We will explore this relationship further later.

**Figure3:** Barplot to check whether strip searched

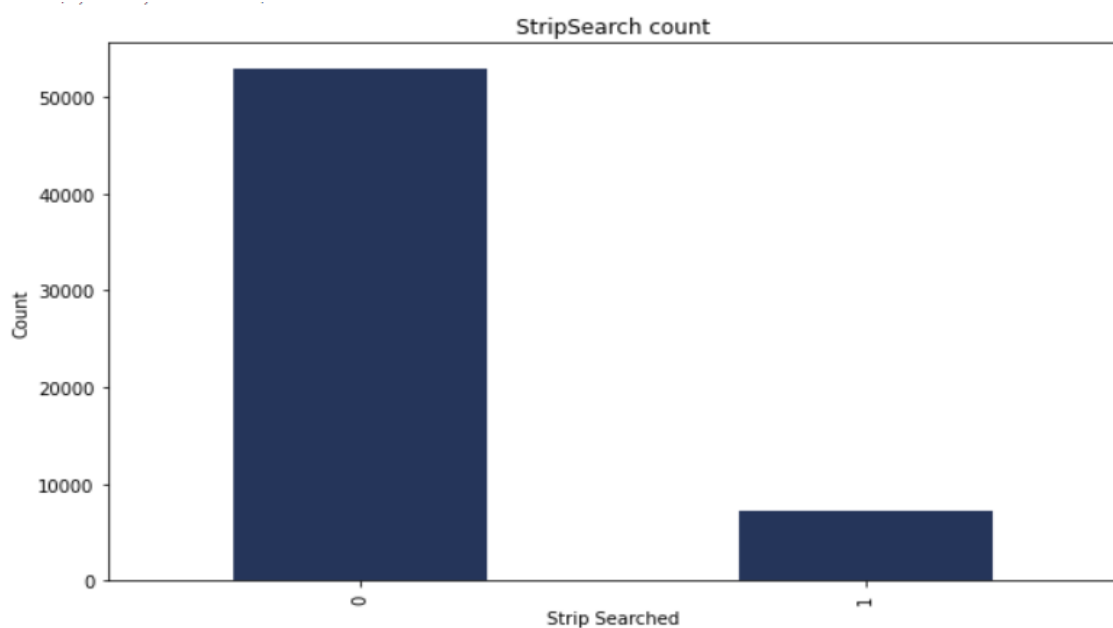


Figure3 shows the number of arrestees who will be strip searched is fewer than that without strip searches. Then, for the people who are strip searched, we wondered whether race and gender will impact the probability to be strip searched. Therefore, several pivot tables were created.

### 5.1.2 Pivot table for Strip search probability based on Race/Sex

**Table1:** probability of StripSearch by Race

**Table2:** probability of StripSearch by Sex

StripSearch	
Race	
Black	0.138879
East/Southeast Asian	0.077237
Indigenous	0.158221
Latino	0.074703
Middle-Eastern	0.070436
South Asian	0.071132
White	0.128653

StripSearch	
Sex	
F	0.101531
M	0.125260

Table1 and Table2 are two pivot tables that display the probability of being strip-searched based on race and sex. We found for the Indigenous category, they are most likely (15.8%) to be strip-searched after being arrested. Followed by the White and the Black, the probabilities are 13.9% and 12.9%, respectively. However, from figure one, the amount of Indigenous arrestees is quite small, so the highest probability here is not convincing. In Table2, the probability of strip search for female arrestees (0.101531) is a little bit lower than that of males (0.125260), in other words, gender might be a factor to impact strip search.

### 5.1.3 Boxplot

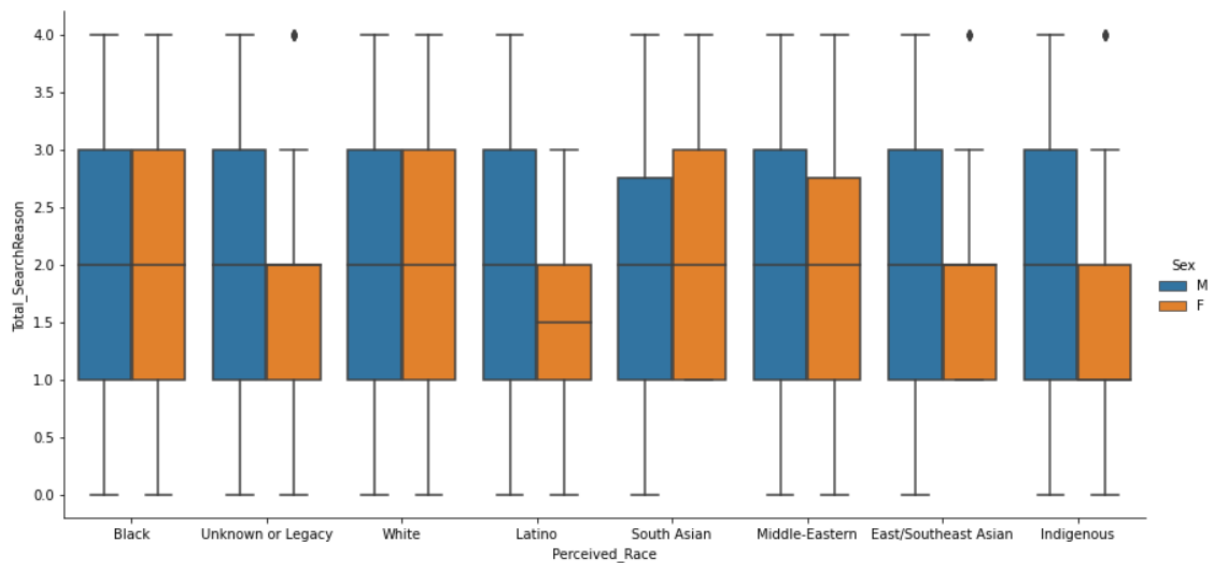
Since we are interested in whether the number of search reasons provided for a strip search vary depending on the race/gender of an arrestee, Figure 4 below visualizes the relationship between gender, race, and the total number of strip search reasons for arrestees. The range of search reasons is from 0 to 4.

First, we examined the relationship between gender and total search reasons based on race. For the majority of races, the distributions of the total search reasons for males and females

were different, while for the two largest racial groups, Black and White, the data distributions were almost the same.

Next, we looked at the relationship between race and total search reasons based on gender. Among almost all races, 50% of the data (Interquartile range) was distributed from 1-3 search reasons, while only IQR of females for Latino, East/Southeast Asian and Indigenous were from 1-2 search reasons, and 4 search reasons tend to be considered as outliers for them.

**Figure4:** boxplot for Race, gender and total search Reasons.



## 5.2 Chi-square test

### 5.2.1 Chi-square test: Race and StripSearch

In this case, Race and Strip Search are both categorical variables, since a t-test requires continuous data, it is not appropriate for testing the relationship between categorical and dummy variables. Therefore, to determine if there is a significant association between Race and Strip Search, we conducted a chi-square test, which is the appropriate statistical test for this scenario. The hypothesis were:

$H_0$ : There is no significant association between Race and Strip Search.

$H_a$ : There is a significant association between Race and Strip Search.

For the result, we got a chi-square statistic of 371.404, and a p-value of  $3.908 \times 10^{-77}$  which was smaller than 0.05. Therefore, the null hypothesis can be rejected, and we can conclude there was a significant association between race and being strip searched.

### **5.2.2 Chi-square test: Sex and StripSearch**

Similar to the case before, a chi-square test was conducted for Sex and StripSearch. The hypothesis were:

$H_0$ : There is no significant association between Sex and Strip Search.

$H_a$ : There is a significant association between Sex and Strip Search.

We received a chi-square statistic of 49.776, and the p-value of  $1.723 \times 10^{-12}$  which was smaller than 0.05. Therefore, the null hypothesis can be rejected, and we can conclude there was a significant association between Sex and being strip searched.

## **5.3 T-Tests**

### **5.3.1 T-Test: Race and Total actions at arrest**

Since it was a two sample t-test, we chose White and Black these two representative races (the majority of arrestees). The purpose was to explore whether there is a significant difference in the mean total actions at arrest for people who are strip searched between White and Black arrestees.

Prior to conducting the t-test, we performed a Levene test to evaluate the homogeneity of variances between the two groups. The hypothesis were:

$H_0$ : The variance between groups of White and Black are equal.

$H_a$ : The variance between groups of White and Black are unequal.

The resulting p-value was  $1.006 \times 10^{-5}$ , smaller than the significance level of 0.05, indicating that there was enough evidence to reject the null hypothesis, and there was a significant difference to the variances of the two groups. Therefore, we should conduct a welch t-test. The hypothesis for welch t-test were:

$H_0$ : The sample means of the total number of actions at arrest for White and Black arrestees are equal.

$H_a$ : The sample means of the total number of actions at arrest for White and Black arrestees are not equal



**Table3:** Welch's t-test result for Race and Total actions at arrest

Race	sample means	sd.	t-statistic	p-value	CI	dof
White	0.230	0.576	-4.312	$1.649 \times 10^{-5}$	[-0.103,-0.039]	4763
Black	0.301	0.657				

The results from Table3 indicate that the mean of total actions at arrest for the White (M=0.2302, SD=0.576) is slightly lower than that of the Black (M=0.301, SD=0.657). With alpha established at 0.05, this is a statistically significant difference as the p-value ( $1.649 \times 10^{-5}$ ) is less than 0.05, and 95% CI [-0.103,-0.039]. Therefore, we can reject the null hypothesis that there is a significant difference in the number of total actions at arrest between the White and the Black who are strip-searched.

### 5.3.2 T-Test: Sex and Total actions at arrest

To explore the relationship between gender and the total number of actions at arrest, we conducted a similar Levene test to check homogeneity of variances, the resulting p-value was 0.419, which was greater than the significance level of 0.05, indicating that we did not have enough evidence to reject null hypothesis, and the variances of the two groups can be assumed to be equal. Then proceeded to conduct a normal two sample t-test. Our hypotheses for the t-test were:

$H_0$ : The sample means of the total number of actions at arrest for Female and Male arrestees are equal.

$H_a$ : The sample means of the total number of actions at arrest for Female and Male arrestees are not equal

**Table4:** T test result for Gender and Total actions at arrest

Sex	sample means	sd.	t-statistic	p-value	CI	dof
Female	0.242	0.594	-0.808	0.419	[-0.051,0.021]	7798
Male	0.257	0.608				

The results from Table4 indicate that the mean of total actions at arrest for females (M=0.242, SD=0.594) is slightly lower than that of the Black (M=0.257, SD=0.608). With alpha established at 0.05, there's not enough evidence to say the sample means are

statistically significant difference as the p-value (0.419) is larger than 0.05, and 95% CI [-0.051,0.021]. Therefore, we can not reject the null hypothesis and there does not exist a significant difference in the number of total actions between females and males who are strip-searched.

### 5.3.3 T-Test: Race and Total Search reasons

To investigate the relationship between race and the total number of strip search reasons, we still chose to focus on the White and Black racial groups in our analysis. The resulting p-value from levene's test was 0.165, which was greater than the significance level of 0.05, indicating that the variances of the two groups can be assumed to be equal. We then proceeded to conduct a normal two-sample t-test. Our hypotheses for the t-test were:

$H_0$ : The sample means of the total number of reasons to perform strip search for White and Black arrestees are equal.

$H_a$ : The sample means of the total number of reasons to perform a strip search for White and Black arrestees are not equal

**Table5:** t test result for Race and Total Search reasons

Race	sample means	sd.	t-statistic	p-value	CI	dof
White	1.965	1.213	-2.018	0.044	[-0.127,-0.002]	5998
Black	2.029	1.206				

The results from Table5 indicate that the mean total strip search reasons for the White (M=1.965, SD=1.213) is slightly lower than that of the Black(M=2.029, SD=1.206). With alpha established at 0.05, this is a statistically significant difference as the p-value (0.044) is less than 0.05, and 95% CI [-0.127,-0.002]. Therefore, we can reject the null hypothesis that there is a significant difference in the number of total reasons for strip search between the White and the Black.

### 5.3.4 T-Test: Gender and Total Search reasons

To explore the relationship between gender and the total number of strip search reasons, similarly, we conducted a Levene test at first, the resulting p-value was 0.150, which was greater than the significance level of 0.05, indicating that the variances of the two groups can be assumed to be equal. Then proceeded to conduct a normal two sample t-test. Our hypotheses for the t-test were:

H<sub>0</sub>: The sample means of the total number of reasons to perform strip search for Female and Male arrestees are equal.

H<sub>a</sub>: The sample means of the total number of reasons to perform a strip search Female and Male arrestees are not equal

**Table6:** t test result for Gender and Total Search reasons

Sex	sample means	sd.	t-statistic	p-value	CI	dof
Female	1.863	1.176	-3.832	0.0001	[-0.214,-0.069]	7798
Male	2.004	1.215				

The results from Table6 indicate that the mean total strip search reasons for females (M=1.863, SD=1.176) is slightly lower than that of the Black(M=2.004, SD=1.215). With alpha established at 0.05, this is a statistically significant difference as the p-value (0.0001) is less than 0.05, and 95% CI [-0.214,-0.069]. Therefore, we can reject the null hypothesis that there is a significant difference in the number of total reasons for strip search between females and males.

## 6 Research Design and Methods

### 6.1 Research Design

As the dataset was externally generated, we were unable to enact data collection processes of our own, and had only the resulting dataset to work with. To begin our analysis, we had to clean the dataset to make it intelligible, for instance by removing data points with the “U” marker for sex (of which there were only 6) or by changing the ‘*Youth\_at\_arrest\_\_under\_18\_years*’ parameter into a boolean value called ‘*Minor*’.

Following this, in the course of exploratory data analysis, we determined that feature engineering needed to be performed and created two new parameters, as demonstrated in the Exploratory Data Analysis. The first was ‘*Total Search Reasons*’ that recorded the number of search reasons provided to conduct the search, and the second was ‘*Total Actions at Arrest*’ that recorded the number of negative actions that the arrestee performed in the course of arrest. This would create variables that could undergo t-testing and further statistical analysis.

## 6.2 Methods

In our analysis, statistical significance was determined at an alpha level of 0.05. There was no weighting of factors in determining statistical significance - the limitations of such an approach are discussed in section 8.2. T-tests were conducted to identify patterns in the dataset worthy of investigation and One-Way and Two-Way ANOVA were conducted to identify subsequent differences in more granular detail.

## 7 Results and Discussion

### 7.1 One-Way ANOVA and Post-Hoc Tests

#### 7.1.1 Race and Actions at Arrest

Following the results of the t-test in the exploratory data analysis, we were interested in determining if there exist differences in mean between all race categories that are recorded in this dataset. We conducted a one-way ANOVA comparison, with hypotheses as follows:

$H_0$ : There are no differences in mean actions at arrest between any races

$H_A$ : There exists differences in mean actions at arrest between any of the races

#### ONE WAY ANOVA COMPARISON

$\approx$	SUM OF SQUARES	DEGREES OF FREEDOM	F	P VALUE (P(>F))
PERCEIVED RACE	12.474214	7.0	4.87638	0.000017
RESIDUAL	2847.518265	7792.0	NaN	NaN

#### MULTIPLE COMPARISON OF MEANS USING TUKEY'S HSD, FWER=0.05

For the sake of brevity, only the means for which the null hypothesis can be rejected are demonstrated here. For all others the confidence interval includes 0, with a p-value that was not significant and so were not detectably different.

GROUP 1	GROUP 2	MEAN DIFF	P-ADJ	LOWER	UPPER	REJECT
Black	East/Southeast Asian	-0.1252	0.0082	-0.2313	-0.0192	True
Black	Unknown or Legacy	-0.0885	0.0451	-0.1759	-0.001	True
Black	White	-0.0709	0.001	-0.1191	-0.0227	True

<b>East/Southeast Asian</b>	Indigenous	0.1639	0.0134	0.0196	0.30282	True
<b>Indigenous</b>	White	-0.1096	0.0481	-0.2188	-0.005	True

This means and 95% confidence intervals for all the groups identified are as follows:

RACE	MEAN	LOWER LIMIT	UPPER LIMIT
<b>BLACK</b>	0.3012	0.2750	0.3273
<b>WHITE</b>	0.2302	0.2113	0.2492
<b>EAST/SOUTHEAST ASIAN</b>	0.1760	0.1238	0.2280
<b>INDIGENOUS</b>	0.3398	0.2593	0.4205
<b>UNKNOWN OR LEGACY</b>	0.2127	0.1643	0.2610

The lowest of these means is East/Southeast Asian, meaning that compared to Indigenous or Black arrestees, there is a statistically significant difference in the number of negative actions that an East/Southeast Asian arrestee demonstrated that resulted in a strip search. This indicates a potential bias against East/Southeast Asian detainees, if there were fewer negative actions, but such an invasive procedure was conducted.

Considering the variances are not equal according to Levene's test, and so ideally Welch's ANOVA or a similar metric should be used. To account for this to the best of our abilities, we set the ANOVA function we used to the most robust metric we could.

### 7.1.2 Sex and Actions at Arrest

As the t-tests did not indicate that there were significant differences, it was likely that ANOVA would not suggest any as well.

$H_0$ : There is no difference in total actions at arrest means between the sexes

$H_A$ : There exists a difference in total actions at arrest means between the sexes

#### ONE WAY ANOVA COMPARISON

	SUM OF SQUARES	DF	F	P VALUE
SEX	0.24625	1.0	0.671406	0.412587

RESIDUAL	2860.129498	7798.0	NaN	NaN
----------	-------------	--------	-----	-----

The ANOVA comparison confirmed our suspicions, there was no significant difference in means between the total search reason means between the sexes. As such, no post-hoc tests were conducted.

### 7.1.3 Race and Total Search Reasons

As the t-tests did not indicate that there were significant differences, it was likely that ANOVA would not suggest any as well.

$H_0$ : There is no difference in total search reason means between any races

$H_A$ : There exists a difference in means between any races

#### ONE WAY ANOVA COMPARISON

	SUM OF SQUARES	DF	F	P VALUE
RACE	21.18.128132	7.0	1.770831	0.0884433
RESIDUAL	11295.330340	7792.0	NaN	NaN

The ANOVA comparison confirmed our suspicions, there was no significant difference in means between the total search reason means between any racial group when taking  $\alpha=0.05$  as the level of significance. As such, no post-hoc tests were conducted.

### 7.1.4 Sex and Total Search Reasons

Following the results of the t-test in the exploratory data analysis, a one-way anova comparison was conducted. The hypotheses were as follows:

$H_0$ : There is no difference in total search reason means between sexes

$H_A$ : There exists a difference in total search reason means between sexes

#### ONE WAY ANOVA COMPARISON

	SUM OF SQUARES	DF	F	P VALUE
SEX	21.455509	1.0	14.686968	0.000128

RESIDUAL	11391.736285	7798.0	NaN	NaN
----------	--------------	--------	-----	-----

As the ANOVA comparison yielded a F statistic and corresponding p-value of statistical significance, a comparison of means using Tukey's HSD was conducted.

**MULTIPLE COMPARISON OF MEANS USING TUKEY'S HSD, FWER=0.05**

GROUP 1	GROUP 2	MEAN DIFF	P-ADJ	LOWER	UPPER	REJECT
F	M	0.1415	0.001	0.0691	0.2138	True

While men are more likely to be strip searched going by the relative frequency, when strip searching women, there are fewer reasons (at a statistically significant level) provided for conducting the strip search.

The average mean difference between the two means is ~0.1415 (95% confidence interval 0.0691, 0.2138). The mean search reasons for women is ~1.8628 (95% confidence interval [CI], 1.7983-1.9272; P < 0.05), while the mean for men is ~ 2.0043 (95% confidence interval [CI], 1.9748-2.03380; P <0.05).

## 7.2 Two-Way ANOVA and Post-Hoc Tests

Two-way ANOVA was conducted to determine the possible interaction between sex and race in regards to Total Actions at Arrest and Total Search Reasons

**TWO WAY ANOVA COMPARISON of TOTAL ACTIONS AT ARREST**

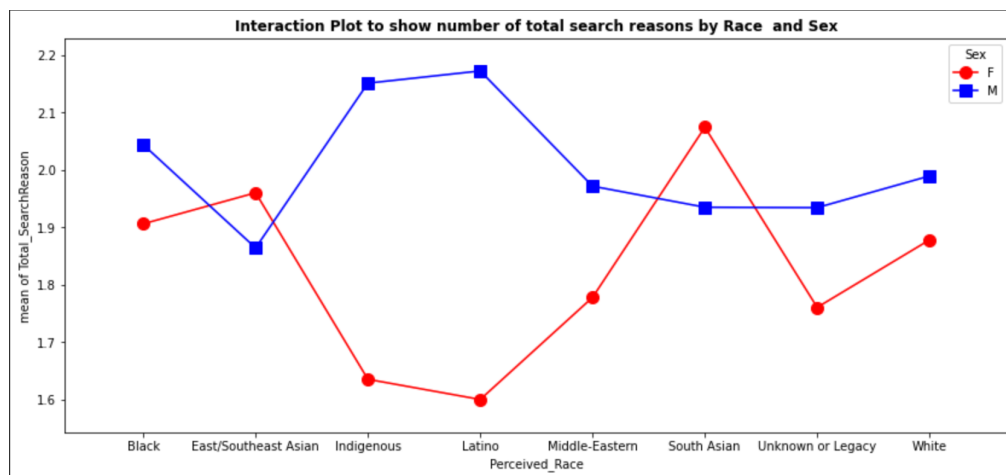
≈	SUM OF SQUARES	DEGREES OF FREEDOM	F	P VALUE (P(>F))
C(PERCEIVED RACE)	12.456495	7.0	4.869146	0.000017
C(SEX)	0.307261	1.0	0.840741	0.359213
C(PERCEIVED_RACE):C(SEX)	2.667133	7.0	1.042561	0.398640
RESIDUAL	2844.774504	7784.0	NaN	NaN

As there is no interaction effect that is statistically significant, no post-hoc tests were conducted.

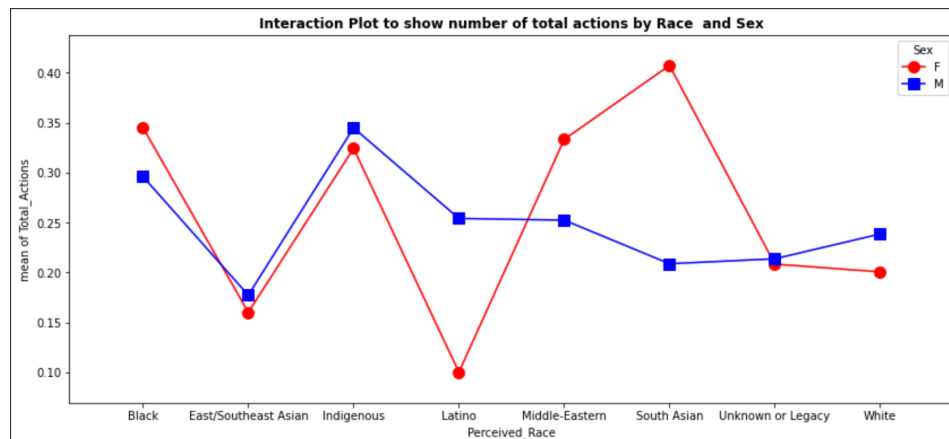
## TWO WAY ANOVA COMPARISON of TOTAL SEARCH REASONS

$\approx$	SUM OF SQUARES	DEGREES OF FREEDOM	F	P VALUE (P(>F))
C(PERCEIVED RACE)	116.463711	7.0	1.611329	0.127075
C(SEX)	20.588048	1.0	14.104890	0.000174
C(PERCEIVED_RACE):C(SEX)	14.290093	7.0	1.39594	0.200960
RESIDUAL	11361.830249	7784.0	NaN	NaN

As there is no interaction effect that is statistically significant, no post-hoc tests were conducted. Nonetheless, we generated interaction plots to investigate if there were any unique interactions that could be investigated in next steps.



**Figure 5** Demonstrates an interaction plot of race and sex with total search reasons provided by the arresting officer. While overall, the mean for women appears to be below that of men, it intersects for Asian populations. In general, we had found that women have a lower search reason mean, but for Asian women this is not the case. Further research should be done.





**Figure 6** Demonstrates an interaction plot with total actions by Race and Sex. While this area did not reach any statistically significant conclusions in the populations that are significantly different here (South Asian and Latino), this might warrant further investigation with more data.

### **7.3 Discussion**

Ultimately, we found significant results to answer two research questions, and were able to reject the null hypothesis. We discovered that there are statistically significant differences in means between sexes in terms of there being fewer mean reasons to conduct a strip search provided for women than men. A way of interpreting this is that a strip search is more often conducted on women with less probable cause or grounds to do so. As gender is a protected class, this is reason for concern - and as discussed in the literature review, considered by some to be a form of state-sanctioned sexual assault [7]. Considering that this is a very invasive procedure, if a regular search is escalated to a strip search for fewer reasons that could be grounds for concern.

Another finding discovered was that regarding race and actions at arrest. The highest mean negative actions at arrest were in the Black and Indigenous categories. They were statistically significant in their difference from East/Southeast Asian, Unknown/Legacy, and White racial groups for Black arrestees, and East/Southeast Asian and White detainees for Indigenous arrestees. The lowest mean was East/Southeast Asian, and it differed significantly from Black and Indigenous arrestees. A way of interpreting this is that given fewer negative actions at arrest on average, an East/Southeast Asian arrestee might still be searched in such an invasive and dehumanizing way, that might be undeserved given less resistance, combatance, or violence from the arrestee. Interpreting the high means of Black and Indigenous race groups is more challenging however.

Given that Indigenous Canadians and Black Canadians are overrepresented in the Canadian correctional system [2] [3], and that there is a history of racial discrimination by the Toronto Police [9], it is hard to interpret this data, as the data collection process itself is subject to the arresting officer's judgment. This metric is based on the arresting officer's perception of the actions, which might very well be biased, as even a Black or Indigenous individual cooperating with an officer, may sometimes be perceived as violent or resisting arrest [6]. So the abnormally high means should be investigated further, and at this stage, it would be inappropriate for us to make any claims.

## **8 Conclusion, Limitations, and Next Steps**

### ***8.1 Conclusion***

To conclude, we investigated the intersections of race, gender, actions at arrest, and search reasons as they connect to individuals that have been strip searched throughout this paper. We began with exploratory data analysis to compare sample means, and then conducted one-way and two-way ANOVA to investigate connections that might be present in terms of race and gender of arrestees that have been strip searched. We discovered that sample means differed between men and women and their total search reasons, with women searched with fewer mean search reasons (MEAN: 1.8628 (95% [CI], 1.7983-1.9272;  $P < 0.05$ )) than men (MEAN: 2.0043 (95% [CI], 1.9748-2.03380;  $P < 0.05$ )). We also discovered that there exist differences in the sample means of certain races when counting their total actions at arrest. East/Southeast Asians that were strip searched had the fewest total actions at arrest (MEAN: 0.1760, 95% [CI], 0.1238-0.2280;  $P < 0.05$ ) while Black (MEAN: 0.3012, 95% [CI], 0.2750-0.3273) and Indigenous (MEAN: 0.3398, 95% [CI], 0.2593-0.4205) people had the highest mean actions at arrest. Overall, these findings are difficult to interpret, but we can conclude that women will be searched with fewer mean search reasons, and East/Southeast Asians that were strip searched demonstrated fewer negative actions at arrest on average.

### ***8.2 Limitations***

There are several limitations to the analyses that were conducted.

The first limitation speaks to how robust the collected data is. For some parameters, it could be subject to the arresting officers' discretion and so could result in biased data. For instance where one may be upset at being arrested, an officer might record that the arrestee is combative or violent when they merely voice their displeasure, as mentioned in the discussion. There is no way of knowing how an arrestee was actually behaving at the time of their arrest. Similarly, the reasons for initiating a search could be inaccurate or the number of search reasons could be inflated, as an officer might have no reason to believe that a particular arrestee has a weapon, but is likely not required to defend their reason before conducting the search. As such, they are free to provide as many or as few reasons as they like for the search, as there is no 'rulebook' for such reasons specifically, but instead relies on the arresting officer's better judgment, which as determined in the literature review, can be biased.

The second and third limitations are in the realm of analysis. In some instances, we had to conduct Welch's t-tests due to unequal variances, but were unsure how to correct for this in an ANOVA. We corrected for this the best we could by setting `robust="hc3"` in the function, but were unable to use Welch's ANOVA or any similar measure. The other limitation is that in creating the new variables '*Total Search Reasons*' and '*Actions at Arrest*', we did not weigh any of the individual actions or search reasons more strongly than the others. That is to say, they were summed up, and so search reasons or actions that might be considered 'more dangerous' from the arresting officers point of view were weighted equally in the final count. For instance, resisting arrest vs. being combative are likely of differing severity. We did not weigh them, however, as we did not have any background information/research that we could rely on to inform the weighting criteria, but intuition suggests that differing categories would be taken into consideration differentially by an officer.

A final limitation is that we chose an alpha value of 0.05. In some instances, if we had chosen another value, and therefore another power, for instance with an alpha value of 0.1, then for some sections (for instance race and search reasons), we would have found statistically significant results. As this is sensitive data, we chose a more conservative statistical power so that we do not erroneously commit a Type 1 error.

### 8.3 Next Steps

Next steps would be to better connect additional datasets regarding police use of force, as well as map the overall population of Toronto to this data. This data could better be analyzed by seeing the proportions of arrested populations to general Toronto populations and seeing if there are any covariances in the data. Additionally, the insights presented in the interaction plots ought to be investigated further with additional data, as in this dataset, no statistically significant results were found.

## References

[https://www.tps.ca/media/filer\\_public/d4/7f/d47fd156-41f5-4ec0-a2b1-1002d9682baf/06b1320e-15a0-48a2-a676-af88640fc5f6.pdf](https://www.tps.ca/media/filer_public/d4/7f/d47fd156-41f5-4ec0-a2b1-1002d9682baf/06b1320e-15a0-48a2-a676-af88640fc5f6.pdf)

- [1] Deloitte LLP. 2022. *Workplace Well-Being, Harassment and Discrimination Review: Report prepared for The Toronto Police Services Board and the Toronto Police Service*. Retrieved from [https://www.tps.ca/media/filer\\_public/d4/7f/d47fd156-41f5-4ec0-a2b1-1002d9682baf/06b1320e-15a0-48a2-a676-af88640fc5f6.pdf](https://www.tps.ca/media/filer_public/d4/7f/d47fd156-41f5-4ec0-a2b1-1002d9682baf/06b1320e-15a0-48a2-a676-af88640fc5f6.pdf)
- [2] Department of Justice Government of Canada. 2016. 2. Statistical Overview on the Overrepresentation of Indigenous Persons in the Canadian Correctional System and Legislative Reforms to Address the Problem - Spotlight on Gladue: Challenges, Experiences, and Possibilities in Canada's Criminal Justice System. Retrieved February 27, 2023 from <https://www.justice.gc.ca/eng/rp-pr/jr/gladue/p2.html>

- [3] Department of Justice Government of Canada. 2022. Overrepresentation of Black People in the Canadian Criminal Justice System. Retrieved February 27, 2023 from <https://www.justice.gc.ca/eng/rp-pr/jr/obpccjs-spnsjpc/index.html>
- [4] Statistics Canada Government of Canada. 2022. Profile table, Census Profile, 2021 Census of Population - Toronto, City (C) [Census subdivision], Ontario. Retrieved February 23, 2023 from <https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E>
- [5] Kelly Hayes. 2023. Disability Justice Organizers Dream Big and Resist a Culture of Disposability. *Truthout*. Retrieved February 23, 2023 from <https://truthout.org/audio/disability-justice-organizers-dream-big-and-resist-a-culture-of-disposability/>
- [6] Schott Holmes. 2017. Resisting Arrest and Racism - The Crime of "Disrespect." *UMKC Law Rev.* (September 2017).
- [7] Jessica Hutchison. 2020. "It's Sexual Assault. It's Barbaric": Strip Searching in Women's Prisons as State-Inflicted Sexual Assault. *Affilia* 35, 2 (May 2020), 160–176. DOI:<https://doi.org/10.1177/0886109919878274>
- [8] Monika Lemke. 2022. Policing Toronto: Strip Searching in a Divided City - The Bullet. *Socialist Project*. Retrieved February 23, 2023 from <https://socialistproject.ca/2022/07/policing-toronto-strip-searching-in-a-divided-city/>
- [9] Joseph Mensah, David Firang, Christopher J. Williams, and Michelle Afrifah. 2021. RACIAL DISCRIMINATION IN THE CANADIAN CRIMINAL JUSTICE SYSTEM: HOW ANTI-BLACK RACISM BY THE TORONTO POLICE HARMS US ALL. *Can. Soc. Work Rev. Rev. Can. Serv. Soc.* 38, 2 (2021), 63–86. DOI:<https://doi.org/10.7202/1086120ar>
- [10] Shanifa Nasser. Family of Toronto man who fell to his death says police locked them in room, didn't let them speak to him | CBC News. *CBC News*. Retrieved February 23, 2023 from <https://www.cbc.ca/news/canada/toronto/taresh-bobby-ramroop-fall-death-1.6630688>
- [11] Pamela Palmater. 2016. Shining Light on the Dark Places: Addressing Police Racism and Sexualized Violence against Indigenous Women and Girls in the National Inquiry. *Can. J. Women Law* 28, 2 (July 2016), 253–284. DOI:<https://doi.org/10.3138/cjwl.28.2.253>
- [12] Private Member. Arrests and Strip Searches (RBDC-ARR-TBL-001). *Toronto Police Service Public Safety Portal*. Retrieved February 23, 2023 from <https://data.torontopolice.on.ca/pages/open-data>
- [13] Nadine Yousif. 2022. Toronto's first-ever mental health crisis response teams — without police — to launch in March. *thestar.com*. Retrieved February 23, 2023 from <https://www.thestar.com/news/gta/2022/01/19/torontos-first-ever-mental-health-crisis-response-teams-without-police-to-launch-in-march.html>
- [14] 2020. Police kill Canadian man during mental health check. *BBC News*. Retrieved February 23, 2023 from <https://www.bbc.com/news/world-us-canada-53141449>
- [15] Special Investigations Unit -- Director's Report Details, Case Number: 20-TCD-124. *Special Investigations Unit*. Retrieved February 23, 2023 from [https://www.siu.on.ca/en/directors\\_report\\_details.php?drid=908](https://www.siu.on.ca/en/directors_report_details.php?drid=908)