

The Effect of Race, Gender and Age on the Strip Search Rate: Result from the GTA Arrest and Strip Searches Case Study

Kaiyang Bian: 1003138464
Ruiyue Ma: 1009705240

**A midterm paper submitted in conformity with the requirements
for the course of INF2178H
Faculty of Information
University of Toronto**

Introduction

As an intrusive and compulsive practice for criminals at arrest, strip searches have faced controversial issues related to criminal justice, fairness and human rights. According to the Supreme Court of Canada, the strip search was identified as “one of the most extreme exercises of police power” and “inherently humiliating and degrading”(R.v. Golden, 2001). Due to the invasive nature of this practice, it leads to the high chance of trauma for the suspects and prisoners. A study found that strip searches impose harm to people with a history of trauma, and nearly half of prisoners incarcerated in Canada have experienced childhood abuse(Canadian Civil Liberties Association, 2022). The strip searches provokes the traumatized experience of those, which might cause more serious psychological damage. In addition, those who were arrested or prisoned might be innocent, waiting for trial or release on bail. Under this circumstance, the strip searches result in a secondary victimization for the innocent.

On the other hand, strip searches are believed to be useful practice for the police to ensure the person was not concealing any dangerous weapons that may cause possible injury to themselves, officers and other individuals whom they may be held in custody with(Toronto Police Service, 2022). With this reason, administrative officials in Ontario have unrestricted authority to approve strip searches at any time and in any situation according to the Ministry of Correctional Services Act and regulations(Canadian Civil Liberties Association, 2022). The law and regulation gives police officers in Ontario permission to determine the search target, place and reason for conducting strip searches. The report from the Ontario Independent Police Review Director indicated that there were an excessive amount of unjustified and illegal strip searches conducted by police in Ontario (McNeilly, 2019).

With the rising concern on the impact of strip searches, it's necessary to understand the rationale and patterns of strip searches, in order to maintain criminal justice and protect individuals from social prejudices, racial or sexual discrimination. How does race influence the frequency of being subjected to a strip search during an arrest? Are marginalized groups more likely to be subjected to this intrusive practice due to social bias? Does the frequency of strip searches differ by gender and age groups? With deeper exploration into the research questions, this report tries

to investigate the patterns of strip searches activities on the arrested individual in Ontario, providing support for the legislation authorities in Ontario on maintaining criminal and social justice. More details regarding the selected indicators are discussed further in the literature review.

The dataset of Arrests and Strip Searches was retrieved from the Public Safety Data Portal by the Toronto Police Service on a total of nearly 65000 arrested individuals in the GTA areas. The project applies t-test, one-way ANOVA and two-way ANOVA models to study the relationship between predictors and outcomes.

Literature Review

The report from the Race and Identity-based Data Collection(2020) showed that the strip searches might enhance social bias as it targets racialized Canadians. The result found that one-third of people who were strip-searched at arrest are Black(Lemke, 2022), and Indigenous people were 1.3 times over-represented in strip searches(Race-Based Data Collection, 2020). The impact of racial profiling is not limited to countries of immigrants. Research conducted by the British Journal of Criminology (Newburn et al., 2004) has shown that the British police force frequently used the “stop and search’ power on the African-Caribbean arrestees regardless of their age and offense in North London for the period of 1999 to 2000. In addition, there were raising concerns on whether there was sexual bias on the subject of strip searches. A study from the Boston Law University Review indicated that most strip searches at arrests targeted men of colour (Cooper, 2013). In comparison to the 43 female cases, 124 male cases alleging illegal strip searches were found in the analysis of criminal cases recorded in the Quicklaw and CanLil databases referencing the Golden decision(Psutka& Sheehy, 2016). This report highlighted that males have a higher possibility of experiencing an illegal strip search, compared to females. Moreover, a statistical analysis showed that the arrestees who were identified as young adults, typically aged between 17 to 23, were more likely to undergo a strip search than other age groups(Ha, 2011).

Data Description

In our research project, we use the dataset that demonstrates information related to all arrests and strip searches conducted by the Toronto Police Service during the period of 2020 to 2010. The dataset contains a total of 65276 records of arrests and strip searches and 24 attributes related to information such as the arrest year and month, race, gender, age groups at arrest, arrest location division, strip search results, occurrence category, action at arrest and strip search reason. This dataset is accessible from the Public Safety Data Portal from Toronto Police Service collected in November 2022. The important attributes related to our study will be the strip search results, strip search reason, perceived race, sex and occurrence category. The original dataset consists of either ordinal data (such as Arrest ID, etc.), binary data in 0 or 1 format(strip search result, various search reasons etc.), or nominal data (perceived race, sex, age groups, occurrence category etc.). There are other missing values of some of the attributes, such as search reasons and items found. Therefore, data cleaning and preprocessing need to be conducted before our analysis.

Research Objective and Questions

Our study seeks to investigate if the strip searches rate differ by personal identities of the arrestees, which includes their race, gender and age at arrest. By conducting the analysis on relationship between strip searched frequency and those attributes, we aimed to study if the perceived racial and sexual bias, as well as the age discrimination could impose obstruction of criminal justice. Our research questions are constructed based on knowledge gained from our literature review and preliminary analysis of the data.

RQ1: *Are there any differences in strip search rate for different perceived races (such as white, black, indigenous, southeast asian, etc.)?*

RQ2: *Are there any significant differences in strip search rate for different age groups (such as aged 17 and below, aged 18 to 24, etc.) and sex?*

Those research questions will help us understand if there is any social bias that affects Ontario police officers on the decision to conduct search searches on arrestees, which might contribute to the further study of criminal justice.

Research Design and Methods

This dataset is retrieved from the Public Safety Data Portal of Toronto Police Service, which consists of a total of 65276 samples of the arrest and strip searches records during the year of 2020 to 2021. As our research questions focus on analyzing the effect of arrestees' personal identity (race, gender and age groups) on the strip search result, we aim to study the relationship between strip searches rate and the attributes of `perceived_race`, `sex` and `age_group_at_arrest` for each arrestee. To satisfy the requirement of t-test and ANOVA test, we decided to create a continuous variable called strip search rate, which acted as the dependent variable in our study. Since the strip searches rate is not given in the original dataset, we calculate the strip searches rate for individual using the following formula:

$$\text{Strip searches rate} = \text{Total count of strip search} / \text{Total count of arrest}$$

As the dataset records all of the arrest events, which includes the chance of one arrestee that might be arrested multiple times during the given time period. During the data preprocessing phase, we first calculate the total count of the arrest event of each arrestee via the unique personal ID. We found that the total number of arrestees is 37347. The sample is large enough for us to conduct the analysis. Followed by calculating the total frequency of being strip searched, we come up with the results of strip searches rate as the dependent variable for t-test and ANOVA analysis.

RQ1: To conduct the analysis of this research question, we use a one-way ANOVA test. According to the purpose of the study, the independent variable is the perceived race, which is a categorical variable that consists of different levels. The relationships can be expressed as:

$$y(\text{strip search rate}) = \beta_0 + \beta_1(\text{race}) + e$$

RQ2: To analyze if the strip search rate is affected by age and gender, we conduct a two-way ANOVA test. In this case, the independent variables are sex (male and female) and different age groups with a total of 7 age levels, including aged 17 and below, aged 18 to 24, aged 25 to 34, aged 35 to 44, aged 45 to 54, aged 55 to 64 and aged 65 and above. Both sex and age groups are categorical variables that consist of different levels, which are not dependent on one another.

$$y(\text{strip search rate}) = \beta_0 + \beta_1(\text{sex}) + \beta_2(\text{age group}) + e$$

We will then set our hypothesis for each of the tests. Based on the result of one-way ANOVA and two-way ANOVA, interpretation can be conducted on how strip search rates differ depending on the age groups and sex of the arrestee.

Descriptive Statistics

	White	27723
	Black	17526
	Unknown or Legacy	5056
	East/Southeast Asian	4415
	South Asian	3613
M	Middle-Eastern	3237
F	Indigenous	1934
U	Latino	1768
Name: Sex, dtype: int64	Name: Perceived_Race, dtype: int64	
Aged 25 to 34 years	20949	
Aged 35 to 44 years	16242	
Aged 18 to 24 years	10041	
Aged 45 to 54 years	9066	
Aged 55 to 64 years	4590	
Aged 17 years and younger	3042	
Aged 65 years and older	1322	
Name: Age_group_at_arrest_, dtype: int64		

(Figure 1.0)

According to the original dataset, we did some data cleaning to impute missing values and deleted the age group with the same meaning. Then these are the categorical values obtained from the cleaned dataset. There are men, women and unisex in the Sex category. There are also 8 race groups and 7 age groups for us to do the exploratory data analysis.

Perceived_Race	count	mean	Actions_at_arrest__Cooperative					
			std	min	25%	50%	75%	max
Black	17526.0	0.425996	0.494507	0.0	0.0	0.0	1.0	1.0
East/Southeast Asian	4415.0	0.513703	0.499869	0.0	0.0	1.0	1.0	1.0
Indigenous	1934.0	0.394002	0.488762	0.0	0.0	0.0	1.0	1.0
Latino	1768.0	0.507353	0.500087	0.0	0.0	1.0	1.0	1.0
Middle-Eastern	3237.0	0.467717	0.499034	0.0	0.0	0.0	1.0	1.0
South Asian	3613.0	0.461389	0.498576	0.0	0.0	0.0	1.0	1.0
Unknown or Legacy	5056.0	0.438884	0.496300	0.0	0.0	0.0	1.0	1.0
White	27723.0	0.444216	0.496887	0.0	0.0	0.0	1.0	1.0

(Figure 1.1)

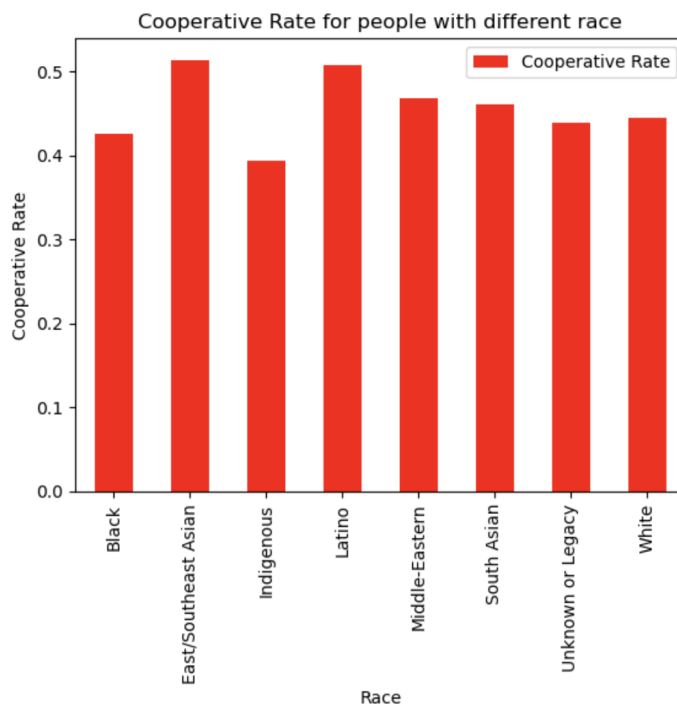
Sex	count	mean	Actions_at_arrest__Cooperative					
			std	min	25%	50%	75%	max
F	12617.0	0.430372	0.495148	0.0	0.0	0.0	1.0	1.0
M	52650.0	0.449687	0.497467	0.0	0.0	0.0	1.0	1.0
U	9.0	0.333333	0.500000	0.0	0.0	0.0	1.0	1.0

(Figure 1.2)

Age_group_at_arrest_	count	mean	std	Actions_at_arrest__Cooperative				
				min	25%	50%	75%	max
Aged 17 years and under	1361.0	0.456282	0.498268	0.0	0.0	0.0	1.0	1.0
Aged 17 years and younger	1681.0	0.472933	0.499415	0.0	0.0	0.0	1.0	1.0
Aged 18 to 24 years	10041.0	0.439996	0.496411	0.0	0.0	0.0	1.0	1.0
Aged 25 to 34 years	20949.0	0.435964	0.495894	0.0	0.0	0.0	1.0	1.0
Aged 35 to 44 years	16242.0	0.444219	0.496894	0.0	0.0	0.0	1.0	1.0
Aged 45 to 54 years	9066.0	0.461725	0.498560	0.0	0.0	0.0	1.0	1.0
Aged 55 to 64 years	4590.0	0.455120	0.498036	0.0	0.0	0.0	1.0	1.0
Aged 65 and older	624.0	0.496795	0.500391	0.0	0.0	0.0	1.0	1.0
Aged 65 years and older	698.0	0.479943	0.499956	0.0	0.0	0.0	1.0	1.0

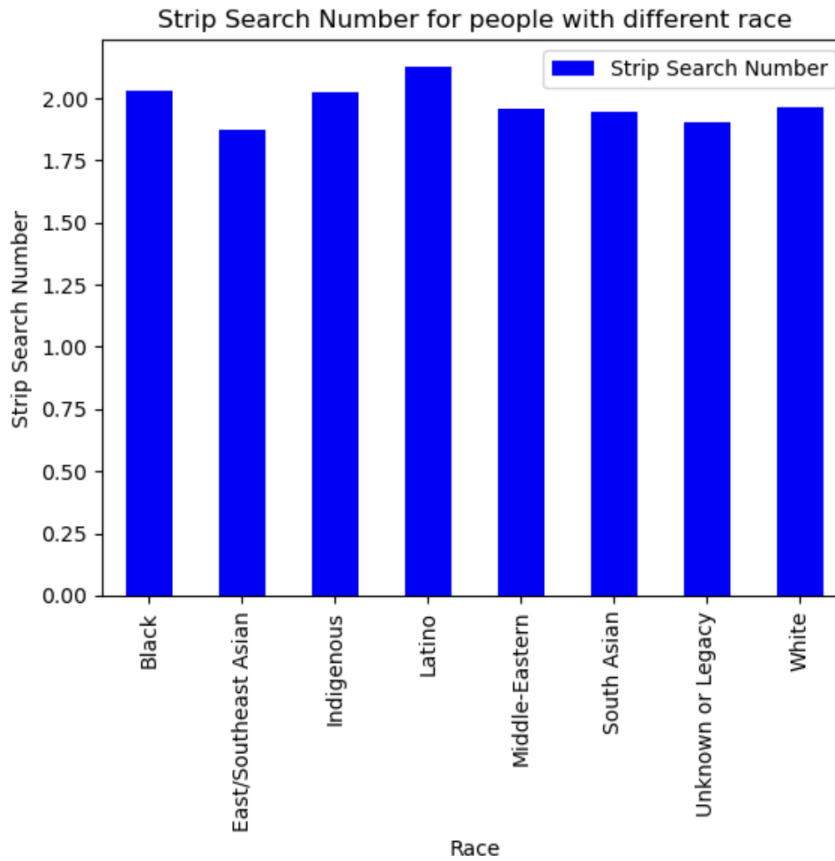
(Figure 1.3)

Then we did three analyses on the cooperative rate amongst these 3 categorical values, and found the East/southeast Asian group has the highest cooperative rate in the race group, men has a slightly higher cooperative rate at arrest than women, and people who are 65 years older are more likely to be cooperative at arrest. This



(Figure 1.4)

The above Figure 1.4 showed the cooperative rate amongst race groups.



(Figure 1.5)

The Figure 1.5 demonstrated the strip search number differences among each race.

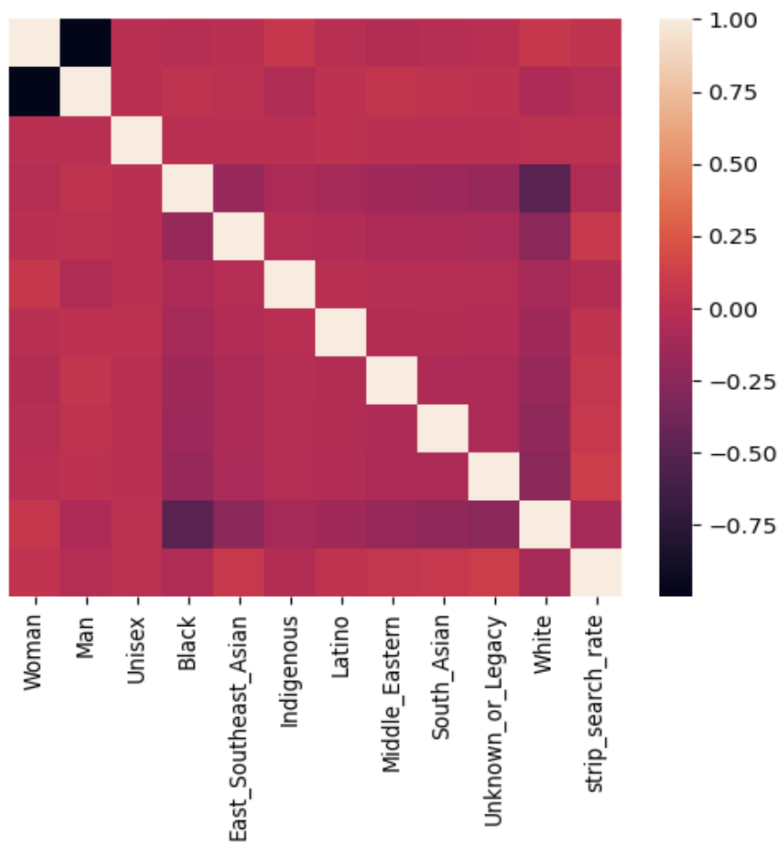
	Sex	Race	arrest_counts	Age_group_at_arrest_	strip_counts	strip_search_rate
0	M	East_Southeast_Asian	1	Aged 35 to 44 years	1	1.0
1	F	White	1	Aged 17 years and younger	1	1.0
2	M	White	1	Aged 35 to 44 years	1	1.0
3	M	Black	1	Aged 25 to 34 years	1	1.0
4	M	South_Asian	1	Aged 45 to 54 years	1	1.0
...
34037	M	White	5	Aged 17 years and younger	1	0.2
34038	M	White	1	Aged 35 to 44 years	1	1.0
34039	M	Black	1	Aged 18 to 24 years	1	1.0
34040	M	Black	1	Aged 65 years and older	1	1.0
34041	M	White	1	Aged 18 to 24 years	1	1.0

34042 rows × 6 columns

(Figure 1.6)

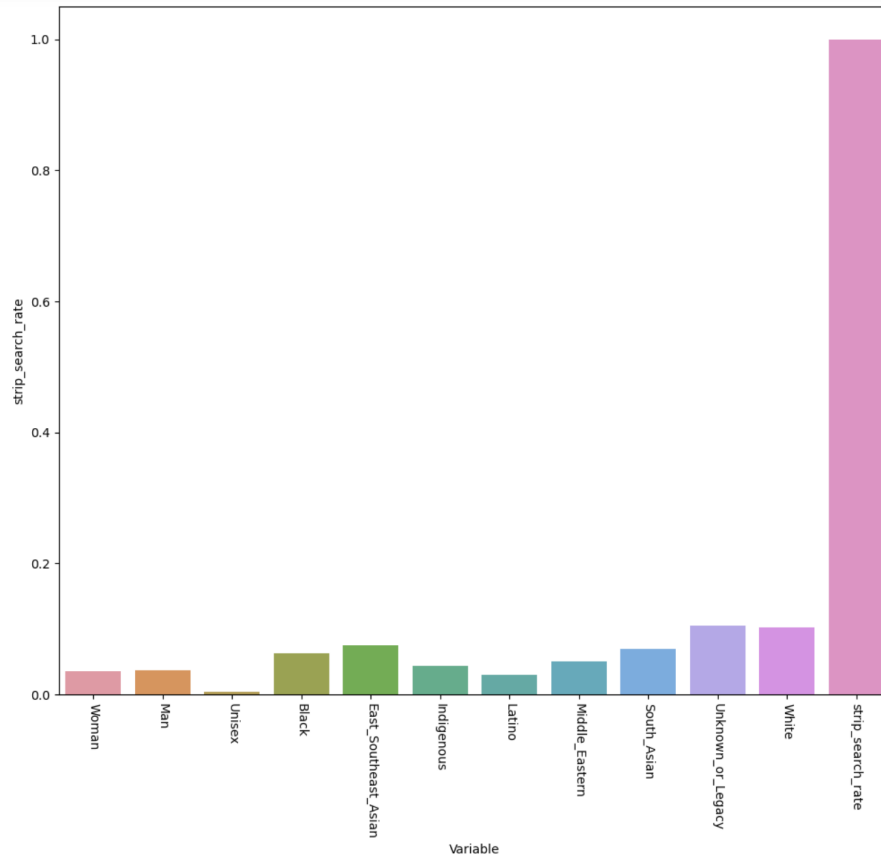
Then we found that some people were arrested several times, and we then decided to further explore the percentage of their strip search rate. We deducted the dataset and created a new

dataframe containing each person only one time using a person ID. The strip search rate is conducted using times being arrested divided by times being strip searched. After data cleaning to remove the people with multiple sex or races, the final data frame has 34041 rows and 6 columns. We used the one-hot encoded method to set dummy variables for the sex and race category. Then used the concat method to combine these data frames together.



(Figure 1.7)

This is the heatmap generated from the seaborn to see the correlation of different categorical attributes (individual variables from gender and race) and strip search rate.



(Figure 1.8)

The barchart of Figure 1.8 is generated to compare the percentage of different variables to strip search rate.

The reason for us to do the heat map and bar chart is to do data virtualization for people to better understand the relationship of different categorical data to the strip search rate. This helps the audience to know the result of the data exploration.

T-test Result-Table 1.0

First IV	Second IV	DV	T statistics	P value
White (Race)	Black (Race)	Strip Search Rate	-1.27	0.2039
White (Race)	Indigenous (Race)	Strip Search Rate	4.666	3.0981 e-6
White (Race)	East/southeast Asian (Race)	Strip Search Rate	-17.44	1.468 e-67

Black (Race)	East/southeast Asian (Race)	Strip Search Rate	-16.880	3.291 e-63
Black (Race)	Indigenous (Race)	Strip Search Rate	5.177	2.301 e-07

To further explore the result of strip search rate on different race groups, we conducted the t-test by choosing different pairs of race groups to compare with. We tried to figure out if the strip search rates differ between the selected ethnic groups. The independent variables are different race groups shown on the table 1.0, which were randomly sampled and they were independent from each other. Since the size of each race group is large enough (>1000 for all chosen ethnic groups), we can be fairly tolerant with the data normality requirement as the sample mean will be approximately normal according to the Central Limit Theorem. The following results showed the relationship of strip search rate between two different races.

White and Black

H0 (Null Hypothesis): The strip search rate is the same for black people and white people.

H1 (Alternative Hypothesis): The strip search rate is different for black people and white people.

From the above result, we found that the T statistics is -1.27 . The p-value is 0.2039, which is larger than 0.05. Thus, we fail to reject the null hypothesis. We conclude there is no significant difference in strip search rate between black people and white people .

White and Indigenous

H0 (Null Hypothesis): The strip search rate is the same for Indigenous people and white people.

H1 (Alternative Hypothesis): The strip search rate is different for black people and white people.

From the above result, we found that the T statistics is 4.666 . The p-value is 3.0981 e-6, which is much smaller than 0.05. Thus, we can reject the null hypothesis. We conclude there is a significant difference in strip search rate between indigenous people and white people .

White and East/southeast Asian

H0 (Null Hypothesis): The strip search rate is the same for east/southeast asian people and white people .

H1 (Alternative Hypothesis): The strip search rate is different for east/southeast asian people and white people .

From the above result, we found that the T statistics is -17.44 . The p-value is 1.468 e-67, which is much smaller than 0.05. Thus, we can reject the null hypothesis. We conclude there is a significant difference in strip search rate between east/southeast asian people and white people.

Black and East/southeast Asian

H0 (Null Hypothesis): The strip search rate is the same for east/southeast asian people and black people .

H1 (Alternative Hypothesis): The strip search rate is different for east/southeast asian people and black people .

From the above result, we found that the T statistics is -16.880 . The p-value is 3.291 e-63, which is much smaller than 0.05. Thus, we can reject the null hypothesis. We conclude there is a significant difference in strip search rate between east/southeast asian people and black people.

Black and Indigenous

H0 (Null Hypothesis): The strip search rate is the same for Indigenous people and black people .

H1 (Alternative Hypothesis): The strip search rate is different for Indigenous asian people and black people .

From the above result, we found that the T statistics is 5.177 . The p-value is 2.301 e-07, which is much smaller than 0.05. Thus, we can reject the null hypothesis. We conclude there is a significant difference in strip search rate between Indigenous people and black people.

The reason to do these T tests is to compare the mean difference between the strip search rate between each race, and have a better clue which race is more likely to be strip searched.

Results and Interpretation

With the preliminary exploratory data analysis, we conducted the ANOVA test to explore the result of our research questions. We found that there are strong relationships between strip search rate and different personal identity of the arrestee. Specifically, race and age of the arrestee showed a significant impact on the strip search rate. The detailed interpretation was shown in the following section.

Impact of Race on Strip Search Rate

We tried to determine whether there was a statistically-significant difference between the mean strip search rate (dependent variable) for different races including White, Black, East/Southeast Asian, Indigenous, Latino, Middle-Eastern, South Asian and Unknown or Legacy (8-level categorical variable). The following hypotheses are tested:

H0 (Null Hypothesis): The population means among all eight different races are equal.

H1 (Alternative Hypothesis): The population means among all eight different races are not equal.

At least one mean is different from the others.

Figure 2.0 : Result of One-way ANOVA on the Effect of Race on Strip Search Rate

	sum_sq	df	F	PR(>F)
Race	66.523241	7.0	179.055325	1.432096e-261
Residual	1806.346713	34034.0	NaN	NaN

From the above result, we found that the F-value is 179.06 . The p-value is 1.43 e-261, which is much smaller than 0.05. Thus, we can reject the null hypothesis. We conclude there is a significant difference in strip search rate between different races.

Multiple Comparison of Means - Tukey HSD, FWER=0.05						
group1	group2	meandiff	p-adj	lower	upper	reject
Black	East_Southeast_Asian	0.0845	0.001	0.0695	0.0995	True
Black	Indigenous	-0.0583	0.001	-0.0895	-0.0272	True
Black	Latino	0.0656	0.001	0.0424	0.0888	True
Black	Middle_Eastern	0.0768	0.001	0.0589	0.0946	True
Black	South_Asian	0.0888	0.001	0.0722	0.1054	True
Black	Unknown_or_Legacy	0.109	0.001	0.0939	0.124	True
Black	White	-0.0044	0.835	-0.0138	0.005	False
East_Southeast_Asian	Indigenous	-0.1428	0.001	-0.1759	-0.1098	True
East_Southeast_Asian	Latino	-0.0189	0.3324	-0.0446	0.0068	False
East_Southeast_Asian	Middle_Eastern	-0.0077	0.9	-0.0287	0.0132	False
East_Southeast_Asian	South_Asian	0.0043	0.9	-0.0156	0.0242	False
East_Southeast_Asian	Unknown_or_Legacy	0.0244	0.0017	0.0059	0.043	True
East_Southeast_Asian	White	-0.0889	0.001	-0.1033	-0.0744	True
Indigenous	Latino	0.1239	0.001	0.0864	0.1614	True
Indigenous	Middle_Eastern	0.1351	0.001	0.1007	0.1695	True
Indigenous	South_Asian	0.1472	0.001	0.1134	0.1809	True
Indigenous	Unknown_or_Legacy	0.1673	0.001	0.1343	0.2003	True
Indigenous	White	0.054	0.001	0.0231	0.0849	True
Latino	Middle_Eastern	0.0112	0.9	-0.0163	0.0386	False
Latino	South_Asian	0.0232	0.1413	-0.0034	0.0499	False
Latino	Unknown_or_Legacy	0.0434	0.001	0.0177	0.0691	True
Latino	White	-0.07	0.001	-0.0929	-0.0471	True
Middle_Eastern	South_Asian	0.012	0.6931	-0.0101	0.0342	False
Middle_Eastern	Unknown_or_Legacy	0.0322	0.001	0.0112	0.0532	True
Middle_Eastern	White	-0.0811	0.001	-0.0985	-0.0637	True
South_Asian	Unknown_or_Legacy	0.0201	0.0449	0.0002	0.04	True
South_Asian	White	-0.0932	0.001	-0.1093	-0.0771	True
Unknown_or_Legacy	White	-0.1133	0.001	-0.1278	-0.0989	True

Figure 2.1 Tukey multiple comparison of means for different race groups

From the post-hoc test results from Figure 2.1, we selected the groups with the most significant differences. Excluding the Unknown_or_Legacy group, we found that the Indigenous people have overall high significant differences around 0.1 ($p < 0.01$) from other race groups, based on the significance level of 0.05. In addition, there are significant differences between Black and South Asian for the mean difference of 0.089, Black and East/Southeast Asian whereas the mean difference is 0.085. Moreover, there are significant differences between South Asian and White for the difference of -0.09, East/Southeast Asian and White for the difference of -0.089, and Middle Eastern and White for the difference of -0.08.

On the other hand, Some groups showed no significant difference, including Black and White, East/Southeast Asian and Latino, East/Southeast Asian and Middle Eastern, East/Southeast Asian and South Asian , and Latino and Middle Eastern. These groups do not differ significantly from each other at the 0.05 level of significance.

Impact of Age and Sex on Strip Search Rate

For the second research question, we tried to analyze whether there was a statistically-significant difference in strip search rate(outcome variable) by age groups (total of 7-level categorical variables) and by sex (male, female and unisex). The following hypotheses are tested:

The main effect of sex on strip search rate

H0 (Null Hypothesis): The mean of strip search rate between different sex are equal.

H1 (Alternative Hypothesis): The mean of strip search rate between different sex are unequal.

The main effect of age group on strip search rate

H0 (Null Hypothesis): The mean of strip search rate among all seven age groups are equal.

H1 (Alternative Hypothesis): The mean of strip search rate among all seven age groups are not equal. At least one mean is different from the others.

The interaction effect between sex and age group on strip search rate

H0 (Null Hypothesis): The mean of strip search rate on the interaction effect between sex and age group are equal.

H1 (Alternative Hypothesis): The mean of strip search rate on the interaction effect between sex and age group are not equal.

Figure 2.2 : Result of Two-way ANOVA on the Effect of Age group and Sex on Strip Search Rate

	sum_sq	df	F	PR(>F)
C(Sex)	0.000001	2.0	0.000013	9.970827e-01
C(Age_group__at_arrest_)	15.228232	6.0	46.443907	5.486996e-39
C(Sex):C(Age_group__at_arrest_)	0.617196	12.0	0.941180	4.875186e-01
Residual	1859.377728	34025.0	NaN	NaN

According to the result above(Figure 2.2), we discovered that the sex does not have a statistically significant impact on the strip search rate, as evidenced by the F-value very close to 0 and it has a very high p-value of 9.97 e-0.1. As the p-value is greater than 0.05, we reject the null hypothesis and conclude that sex does not have significant effect on the strip search rates. On the other hand, the result showed that the F-value for age group is 46.44, with a p-value of 5.49 e-39 (p-value<0.05). Therefore, we reject the null hypothesis. This result showed that age group has a significant effect on the strip search rate. Finally, the result suggested that there is no significant effect between sex and age group on the strip search rate, as it has the F-value of 0.94 and p-value for 4.88 e-0.1 (>0.05). We fail to reject the null hypothesis in this case and conclude that the interaction between sex and age does not have an impact on the stirp search rate.

Multiple Comparison of Means - Tukey HSD, FWER=0.05						
group1	group2	meandiff	p-adj	lower	upper	reject
Aged 17 years and younger	Aged 18 to 24 years	0.0089	0.7622	-0.0095	0.0272	False
Aged 17 years and younger	Aged 25 to 34 years	-0.024	0.001	-0.0414	-0.0067	True
Aged 17 years and younger	Aged 35 to 44 years	-0.0338	0.001	-0.0517	-0.016	True
Aged 17 years and younger	Aged 45 to 54 years	-0.0228	0.0065	-0.0416	-0.004	True
Aged 17 years and younger	Aged 55 to 64 years	-0.0025	0.9	-0.0231	0.0182	False
Aged 17 years and younger	Aged 65 years and older	0.0349	0.0025	0.008	0.0618	True
Aged 18 to 24 years	Aged 25 to 34 years	-0.0329	0.001	-0.0442	-0.0216	True
Aged 18 to 24 years	Aged 35 to 44 years	-0.0427	0.001	-0.0548	-0.0307	True
Aged 18 to 24 years	Aged 45 to 54 years	-0.0317	0.001	-0.0452	-0.0182	True
Aged 18 to 24 years	Aged 55 to 64 years	-0.0113	0.3519	-0.0272	0.0046	False
Aged 18 to 24 years	Aged 65 years and older	0.026	0.0187	0.0025	0.0495	True
Aged 25 to 34 years	Aged 35 to 44 years	-0.0098	0.0826	-0.0203	0.0006	False
Aged 25 to 34 years	Aged 45 to 54 years	0.0012	0.9	-0.0109	0.0133	False
Aged 25 to 34 years	Aged 55 to 64 years	0.0216	0.001	0.0068	0.0363	True
Aged 25 to 34 years	Aged 65 years and older	0.0589	0.001	0.0362	0.0816	True
Aged 35 to 44 years	Aged 45 to 54 years	0.011	0.1425	-0.0017	0.0238	False
Aged 35 to 44 years	Aged 55 to 64 years	0.0314	0.001	0.0161	0.0467	True
Aged 35 to 44 years	Aged 65 years and older	0.0687	0.001	0.0457	0.0918	True
Aged 45 to 54 years	Aged 55 to 64 years	0.0204	0.0048	0.0039	0.0368	True
Aged 45 to 54 years	Aged 65 years and older	0.0577	0.001	0.0339	0.0815	True
Aged 55 to 64 years	Aged 65 years and older	0.0373	0.001	0.0121	0.0626	True

Figure 2.3 Tukey multiple comparison of means for different age groups

According to the result, we selected the top 5 age groups that are significantly different from others. As the result indicated, there are significant differences between the group aged 35 to 44 and group aged 65 and above for the significant difference of 0.069, followed by aged 24 to 34 and aged 65 and above for the significant difference of 0.060. The third one is between the aged 45 to 54 and age 65 and above for the significant difference of 0.058. In addition, There are significant differences between aged 18 to 24 and aged 35 to 44(meandiff=-0.043), and between aged 55 to 64 and 65 and above(meandiff=0.037).

```
Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====
group1 group2 meandiff p-adj lower upper reject
-----
F      M    -0.0217  0.001  -0.029 -0.0144  True
F      U     0.0956  0.7395  -0.2217  0.4129  False
M      U     0.1173  0.6477  -0.1999  0.4345  False
-----
```

Figure 2.4 Tukey multiple comparison of means for different sex

The Tukey HSD test shows that male and female do have significant differences in strip search rate as the p-value is less than 0.05, then we have enough evidence to reject the null hypothesis that male and female have equally the same strip search rate. However, although the absolute value of mean difference between male and unisex, and between female and unisex is larger than the difference between male and female, the sample size is too small so we fail to reject the null hypothesis that male and female has the same strip search rate as unisex since the p-value is 0.7495 and 0.6477 which are both larger than 0.05.

Discussion and Conclusion

The findings of this project implicates some limitations. As we discovered that there might be some discrepancy on the original dataset, as some arrestees who shared the same personal ID were recorded under different genders or different race groups for various arrest events. We suspect that there might be some errors for the record of this data. Also, there are large portions of unknown values, such as the unknown race, which might significantly affect the result of our study. There are some improvements that could be made for our future studies. Given the

possible influence of other factors on the strip searches result for individual arrestees, the accuracy of the strip searches outcome can be improved if we conduct other statistical analysis, such as logistic regression, to control other variables including occurrence category or arrestee's action at arrest.

In conclusion, our findings show that certain personal identity of the arrestee including race and gender have significant impact on the strip search rate, whereas sex has little impact on the strip search rate for this dataset if we include the effect of unisex. We use quantitative methods to analyze the result of our research questions. The results of RQ1 showed that it's necessary to study the effect of race on strip search rate, as we discovered that there is a significant difference between the perceived race of the arrestee and the chance of being strip searched by the police officer. In addition, RQ2 examined the impacts of other personal attributes on the effect of strip search rates. As the result shown, we discovered the age group also has a significant impact on the strip search rates. Lastly, there is a significant difference between the strip search rate for male and females. These results are aligned with some of the findings from previous research and studies that race, age and sex has significant impacts on strip search rates. These findings from our study will contribute to the social and criminal justices, which is important for the legislation authorities in Ontario to study the fairness and ethics on the appropriate use of strip search in the current criminal system.

References

- Canadian Civil Liberties Association. (2022, June 20). *Strip searches in Ontario prisons*. CCLA. Retrieved February 24, 2023, from <https://ccla.org/major-cases-and-reports/strip-searches-in-ontario-prisons/#:~:text=Ontario's%20law%20does%20the%20opposite,reasons%20to%20justify%20strip%20searches.>
- Cooper, R. F. (2013). We Are Always Already Imprisoned: Hyper-Incarceration and Black Male Identity Performance . In *Boston University Law Review* (Vol. 93, Ser. 1185, p. 1202). essay, Boston University School of Law.
- Ha, D. (2011). Blanket policies for strip searching pretrial detainees: an interdisciplinary argument for reasonableness. *Fordham Law Review*, 79(6), 2721-2760.
- Lemke , M. (2023, January 3). *Strip searches are ineffective, unnecessary and target racialized Canadians*. The Conversation. Retrieved February 24, 2023, from <https://theconversation.com/strip-searches-are-ineffective-unnecessary-and-target-racialized-canadians-185187>
- McNeilly, G. (2019, March). *Breaking the golden rule - a review of police strip searches in Ontario*. Office of the Independent Police Review Director. Retrieved February 24, 2023, from https://www.oiprd.on.ca/wp-content/uploads/OIPRD_Breaking-the-Golden-Rule_Report_Accessible.pdf
- Newburn, T.,Shiner, M. & Hayman, S. (2004). Race, Crime and Injustice? *Strip Search and the Treatment of Suspects in Custody*, *The British Journal of Criminology*, 44(5), 677–694, Retrieved February 24,2021, from <https://doi.org/10.1093/bjc/azh043>
- Psutka,M. & Sheehy. (2016). Strip-Searching of Women in Canada: Wrongs and Rights, 2016 94-2 *Canadian Bar Review* 241, 2016 CanLIIDocs 152, Retrieved February 24, 2023, from <https://canlii.ca/t/74q>,
- R. v. Golden, 2001 SCC 83 (CanLII), [2001] 3 SCR 679, Retrieved February 24, 2023, from <https://canlii.ca/t/51xm>
- Toronto Police Service. (2022, July 15). *2020 race and identity-based data collection findings*. 2020 RBDC Findings. Retrieved February 24, 2023, from <https://www.tps.ca/race-identity-based-data-collection/2020-rbdc-findings/>