**Examining the Role of Race and Gender in Strip Search Patterns**

Faculty of Information

INF2178H: Experimental Design for Data Science

Supervisor: Professor Shion Guha

Team 23: Yanke Mo,

Date: April 14, 2023

**Table of Content**

## Introduction

Strip Search is an aggressive procedure that entitles law enforcement agencies to search suspected individuals. It is highly controversial and concern raising due to its significant influence on whoever has been searched. Among all the causing factors of the strip search, race, gender, history of booked are particularly contentious. Based on the findings we concluded in the previous research, we discovered that the Toronto Police Service has conducted strip searches disproportionately regarding the arrestees' race and gender. The association between arrestees' race and gender with respect to search patterns is subject to unjust treatment conducted in the criminal justice system of Toronto. To improve the overall quality of the previous research and enhance the precision of the previous discoveries, this research will build upon the primary findings of unequal treatment in Toronto Law enforcement action, while considering the external impact of booking history. In this case, this research seeks to adopt statistical measures of logistic regression, one-way ANCOVA, and Tukey test to explore the data released by Toronto Police Service (Toronto Police Service, 2022). By conducting data analysis and statistical analysis, the researcher expects to make valuable exploration and improve the understanding of on racial and gender representation in strip searches.

## Literature Review

In the Toronto Criminal Justice System, over-representation among communities of color is continuously emphasized. Based on the news of "Toronto police's race-based data on use of force, strip searches highlighted" reported by CTV producer Chris Fox, races including black, indigenous, and Middle East were significantly overrepresented among the majority of offense types and call types (Fox, 2022). Fox uncovers that the Black community is 2.3 times more likely to be subjected to having firearms pointed at them, even when no weapons are perceived to be in their possession(Fox, 2022). This community was also overrepresented in general enforcement actions practiced by Toronto police during 2020; they were 2.2 times more likely to be involved in legal enforcement actions such as strip searches. Similar experiences were shared by indigenous communities and Middle East communities; by the factors of 1.6 times and 1.3 times they were more likely to be exposed to various types of legal enforcement actions (Fox, 2022).

Based upon this irrefutable evidence and public racial accusation, the Ontario Human Right Commission(OHRC) announced that the disproportionate use of force and enforcement action against black community is acknowledged (Ontario Human Rights Commission (OHRC)). The OHRC is committed to working with the police and the Police Services Board to root out racial discrimination in policing against black community.

**Data Description**

The research explores the pattern of strip search in Toronto Police Service using the data of Arrest and Strip Searches (RBDC-ARR-TBL-001) released by Toronto Police Service in 2022(Toronto Police Service, 2022) . The source of data can be approached via the link: [Arrests and Strip Searches (RBDC-ARR-TBL-001) | Toronto Police Service Public Safety Data Portal]) Among a total of 65,276 arrestee records, the attributes of arrest location, age of arrestee, arrestee gender, arrestee race, strip search, and booking record within 24 hours are highly relevant to the interest of the research. Majority of the attributes are stored as binary variables, categorical variables which scale from multiple levels, and numeric integer variables. The attribute of the strip search, which is defined as a search on individuals enforced by law, is explained through binary variables of 1 (Yes Strip Search) and 0 (No Strip Search). The gender attribute is represented by binary variables of Sex for female (F) and male (M), and the race attribute is described by a categorical variable which contains eight racial categories. The attribute of Booked represents a formal record that is created for arrestees who are involved in police custody. It is noted by a binary variable of 1(Yes booked) and 0 (No booked) in the dataset.

In exploring the association of strip search patterns specifically related with race and gender, the dataset allows the research to develop valuable insight into the criminal justice system in Toronto, while controlling the influential attributes constant . The attributes of race and gender, in particular, enable the research to conduct strip search studies based on racial and gender disparities, and the research finding is expected to provide beneficial discovery on the ongoing disparities occurring in the criminal justice system.

**Research Objective and Question**

The interest of this study falls on potential racial and gender disparities in the Toronto Police service. Supported by the literature pieces "Toronto Police's race-based data on the use of force, strip searches highlighted" and "OHRC statement on Toronto Police Service announcement on race-based data collection findings | Ontario Human Rights Commission," the research intends to uncover the disproportionate use of force in conjunction with racial and gender issues, particularly in the context of strip search.

To enhance the precision of the previous findings, this study aims to narrow the scope of the research by specifically focussing on the impacts of race and gender on search search pattern. Considering that the criminal background represented by variables of Booked is found to be highly related to the strip search pattern, this study will remove the influence of Booked by holding it constant, and keeping the research scope of the strip search pattern regarding race and gender. To examine the research interest above, the research conducts statistical testing and hypothesis on subsequent research questions:

1. Is there a difference in the frequency of strip searches conducted by the police officer on individuals of different races, after controlling for the individual's booked record?
2. Does the individual's race and gender significantly predict the likelihood of being subjected to a strip search by the police office

## Descriptive Statistic

The descriptive statistic incorporates data cleaning, univariate analysis, bivariate analysis, and small multiple figures creation based on the source of Arrest and Strip Searches (RBDC-ARR-TBL-001). The adoption of descriptive tools enables the researcher to obtain a professional understanding of the underlying patterns and distribution of the dataset, all of which will serve as a foundation for subsequent statistical testing and hypothesis analysis.

### *1. Data Cleanning*

The dataset of Arrest and Strip Searches, which has been made available through the Toronto police portal, contains 24 attributes in total. To filter out the required attributes for testing and hypothesis, the research collects arrest year, person ID, sex, perceived race, strip search, booked, and age as independent variables. The filtering process involves using '*groupby'*, which allows python to select and package the required attributes. During the attributes filtering, the research notices that the value of strip search is explained by the binary variables of 1 and 0. To convert the strip search as a continuous variable that counts the frequency of searching per individual, the research sums up the number of strip searches based on PersonID.

**Figure 1**

*Grouping the required attributes*

```
df3 = df_new.groupby(['Arrest_Year', 'PersonID', 'Sex', 'Perceived_Race']).agg({'StripSearch':'sum', 'Booked':'sum', 'Age':'max'})
df3.head(25)
```

| Arrest_Year | PersonID | Sex | Perceived_Race | StripSearch | Booked | Age |
|---|---|---|---|---|---|---|
| 2020 | 300000 | M | East/Southeast Asian | 0 | 0 | 40.0 |
| 2020 | 300001 | F | White | 0 | 1 | 17.0 |
| 2020 | 300002 | M | White | 0 | 0 | 40.0 |
| 2020 | 300003 | M | White | 1 | 3 | 30.0 |
| 2020 | 300004 | M | Black | 0 | 1 | 30.0 |
| 2020 | 300005 | M | South Asian | 0 | 0 | 50.0 |
| 2020 | 300006 | M | Middle-Eastern | 0 | 1 | 60.0 |
| 2020 | 300007 | M | South Asian | 1 | 3 | 30.0 |
| 2020 | 300008 | M | Black | 0 | 1 | 30.0 |
| 2020 | 300009 | M | White | 0 | 0 | 40.0 |
| 2020 | 300012 | M | White | 0 | 1 | 50.0 |
| 2020 | 300014 | M | East/Southeast Asian | 0 | 0 | 60.0 |
| 2020 | 300015 | F | Unknown or Legacy | 1 | 1 | 30.0 |

The study further develops research concern on the age of arrestees. To enhance the calculability of the age attribute, the research converts the age group (e.g. Aged 25 to 34 years) into the average age of the group. After completing all the processes, the dataset is qualified for data visualization.

**Figure 2**

*Replacing age strings into age values*

```
#Convert Age into categorical numbers
df2['Age_group__at_arrest_'] = df2['Age_group__at_arrest_'].replace({'Aged 25 to 34 years': 30,
                                                                    'Aged 35 to 44 years': 40,
                                                                    'Aged 18 to 24 years': 21,
                                                                    'Aged 45 to 54 years': 50,
                                                                    'Aged 55 to 64 years': 60,
                                                                    'Aged 65 and older': 65,
                                                                    'Aged 65 years and older': 65,
                                                                    'Aged 17 years and younger': 17,
                                                                    'Aged 17 years and under': 17,})
df2
```

None
Like what you see? Visit the data table notebook to learn more about interactive tables.

1 to 25 of 20000 entr

| index | Arrest_Year | Arrest_Month | EventID | ArrestID | PersonID | Perceived_Race | Sex | Age_group__at_arrest_ | Youth_at_arrest__under_18_years |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 2020 | July-Sept | 1005907 | 6017884.0 | 326622.0 | White | M | 40.0 | Not a youth |
| 1 | 2020 | July-Sept | 1014562 | 6056669.0 | 326622.0 | White | M | 40.0 | Not a youth |
| 2 | 2020 | Oct-Dec | 1029922 | 6057065.0 | 326622.0 | Unknown or Legacy | M | 40.0 | Not a youth |

To allow the researcher to conduct logistic regression, the independent variables of Strip Search are converted into dummy variables, and the table that contains the dummy variable of strip search is stored as df_3.

**Figure 3**

*Converting strip search into dummy variables*

```
#Convert Strip Search back to Dummy, people who get search account as 1, haven't have any search search get 0.
df3['StripSearch'] = np.where(df3['StripSearch'] == 0, 0, 1)
df3
```

| | Arrest_Year | PersonID | Sex | Perceived_Race | StripSearch | Booked | Age |
|---|---|---|---|---|---|---|---|
| **0** | 2020 | 300000 | M | East/Southeast Asian | 0 | 0 | 40.0 |
| **1** | 2020 | 300001 | F | White | 0 | 1 | 17.0 |
| **2** | 2020 | 300002 | M | White | 0 | 0 | 40.0 |
| **3** | 2020 | 300003 | M | White | 1 | 3 | 30.0 |
| **4** | 2020 | 300004 | M | Black | 0 | 1 | 30.0 |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **45276** | 2021 | 337332 | M | Unknown or Legacy | 0 | 1 | 40.0 |
| **45277** | 2021 | 337334 | F | Latino | 0 | 0 | 17.0 |
| **45278** | 2021 | 337339 | M | White | 0 | 1 | 40.0 |
| **45279** | 2021 | 337342 | M | White | 0 | 0 | 17.0 |
| **45280** | 2021 | 337345 | M | Black | 0 | 0 | 65.0 |

The independent variables of gender and perceived race have also been converted into binary variables and categorical variables respectively. The process of converting the variable involves creating a new dummy variable dataframe and combining the new dataframe into the original dataframe.

**Figure 4**

*Converting Race and Gender into dummy variable and categorical variable*

```
[54] # converting Sex to binary data
     df5 = pd.get_dummies(df3["Perceived_Race"])
     print(df5)
```

```
       0  1  2  3  4  5  6  7
0      1  0  0  0  0  0  0  0
1      0  1  0  0  0  0  0  0
2      0  1  0  0  0  0  0  0
3      0  1  0  0  0  0  0  0
4      0  0  1  0  0  0  0  0
...   .. .. .. .. .. .. .. ..
45276  0  0  0  0  0  1  0  0
45277  0  0  0  0  0  0  1  0
45278  0  1  0  0  0  0  0  0
45279  0  1  0  0  0  0  0  0
45280  0  0  1  0  0  0  0  0

[45272 rows x 8 columns]
```

```
[55] #Combine all the dummies into original data
     df3 = pd.concat([df3, df4, df5], axis=1)
     df3
```

| | Arrest_Year | PersonID | Sex | Perceived_Race | StripSearch | Booked | Age | F | M | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 2020 | 300000 | M | 0 | 0 | 0 | 40.0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **1** | 2020 | 300001 | F | 1 | 0 | 1 | 17.0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| **2** | 2020 | 300002 | M | 1 | 0 | 0 | 40.0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |

At the same time, to satisfy the requirement of ANCOVA test which requires the dependent variable to perform the continuous variables, the researcher stores the continuous value of strip search in another df_7.

**Figure 5**

*Store the continuous value of  Strip Search in df_7*

```
df7 = df_new.groupby(['Arrest_Year', 'PersonID', 'Sex', 'Perceived_Race'],as_index=False).agg({'StripSearch':'sum', 'Booked':'sum', 'Age':'max
df7.head(25)
```

| | Arrest_Year | PersonID | Sex | Perceived_Race | StripSearch | Booked | Age |
|---|---|---|---|---|---|---|---|
| 0 | 2020 | 300000 | M | East/Southeast Asian | 0 | 0 | 40.0 |
| 1 | 2020 | 300001 | F | White | 0 | 1 | 17.0 |
| 2 | 2020 | 300002 | M | White | 0 | 0 | 40.0 |
| 3 | 2020 | 300003 | M | White | 1 | 3 | 30.0 |
| 4 | 2020 | 300004 | M | Black | 0 | 1 | 30.0 |
| 5 | 2020 | 300005 | M | South Asian | 0 | 0 | 50.0 |

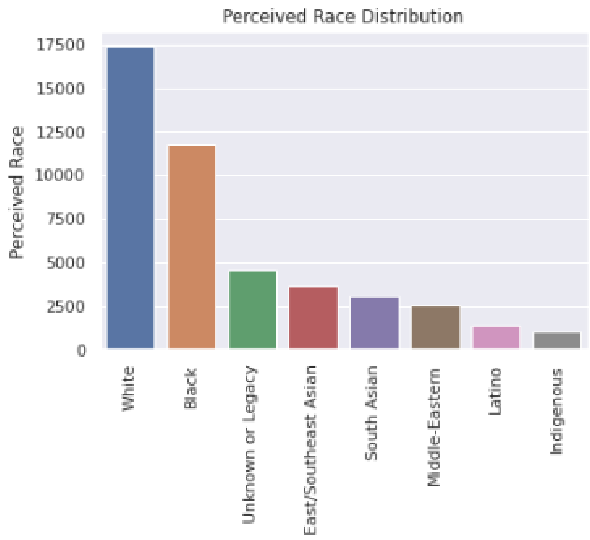
To perform the statistical testing in great precision, the researcher drops the missing value by using *dropna()* for df3 and df7. By now, the dataset is cleaned and is ready to perform statistical tests and analysis.

## 2. *Univariate Analysis*

The initial stage of data visualization involves adopting a histogram based on the perceived race of the arrestee (figure 3). According to the histogram of perceived race distribution, white arrestees and black arrestees perform outstandingly in comparison with other races. This performance reveals that legal involvement is commonly observed among white community and the black community, whereas all the other race communities have relatively low involvement in legal action.

**Figure 6**

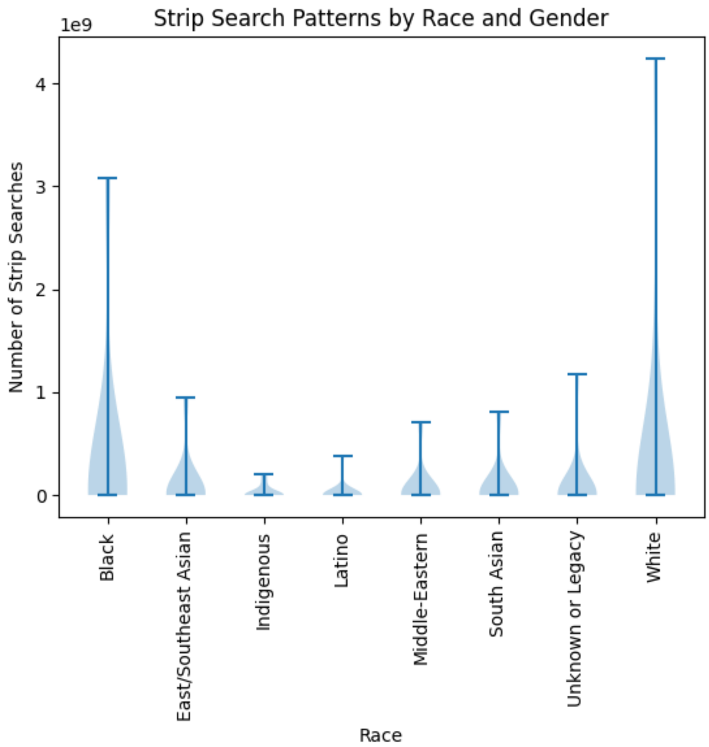
*Histogram of race distribution in all strip search events*




The analysis of the histogram proves against some assumptions raised in the previous research. Indigenous communities and Middle-Eastern communities are expected to have higher rates of legal involvement in the literature study. Yet, the histogram disapproves of this statement and argues that white communities and black communities have the highest involvement in legal enforcement. In particular, white communities are assumed to enjoy more tolerance in legal enforcement, where the histogram suggests that white communities have the highest legal issue among all races.

**Figure 7**

*Histogram of races distribution in all strip search events*




The violin plot (figure 4) conducted on strip searches and races expresses a similar argument to the histogram above. White communities and black communities were actively performing in strip searches, and a significant number of white and black arrestees had at least 1 or 2 times strip search experiences. Other than that, the majority of arrestees have no strip search experience regardless of race; the mean number of strip searches for all races lies at 0.

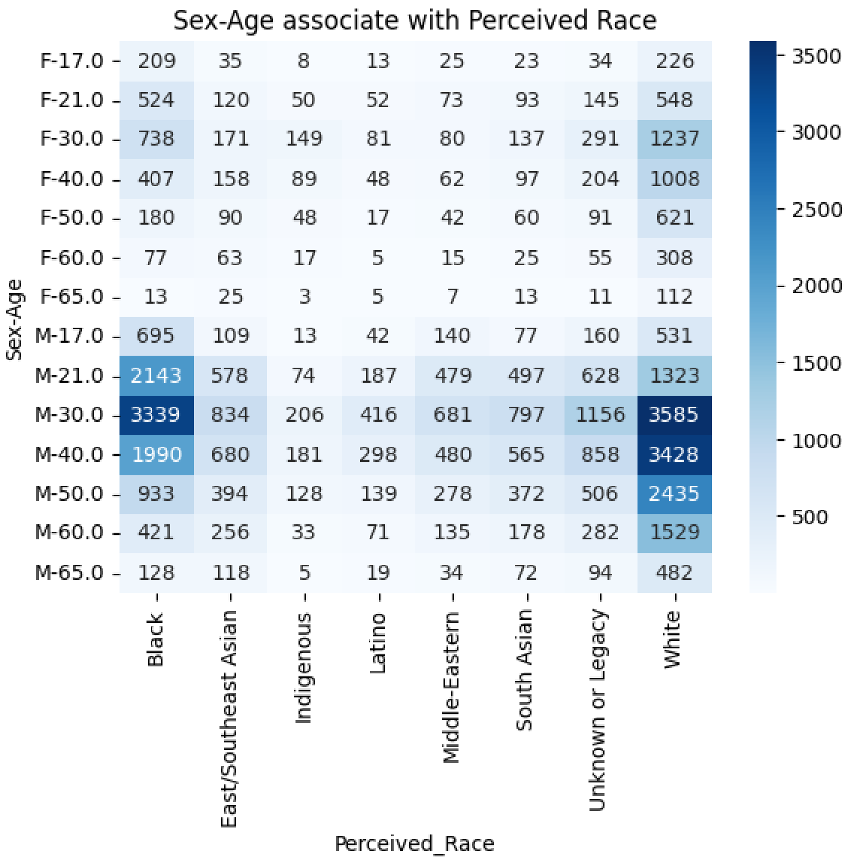
Considering that white communities account for the largest demographic population (43.5%) in Toronto (Demographics of Toronto, 2022), the research remains neutral on the strip search activities white communities participated in. Black communities have only a 9.6% demographic distribution in Toronto, but the outstanding performance of Black communities uncovers potential racial disparities occurring in Toronto's legal enforcement activities.

### *3. Small Multiple Figure*

The second stage of data visualization involves conducting a heatmap (figure 5) based on the perceived race of arrestees, gender of arrestees, and age of arrestees. According to the heatmap below, white male arrestees and black male arrestees perform actively in strip searches at the age of 30. White male arrestees express continuous active performance throughout the ages of 40 and 50, whereas black males have a significant decline in legal involvement until they reach the age of 40. Female performance in strip search remains relatively low in comparison with the male. Frequent involvements among females take place at the age of 30, where white female arrestees and black female arrestees exhibit high involvement in strip search. Based on the pattern, white communities and black communities are actively involved in strip search at the age of 30, while males are more likely to experience strip search in comparison with females. The heatmap exhibits valuable insight into the pattern of strip search with regard to race, gender, and age.

**Figure 8**

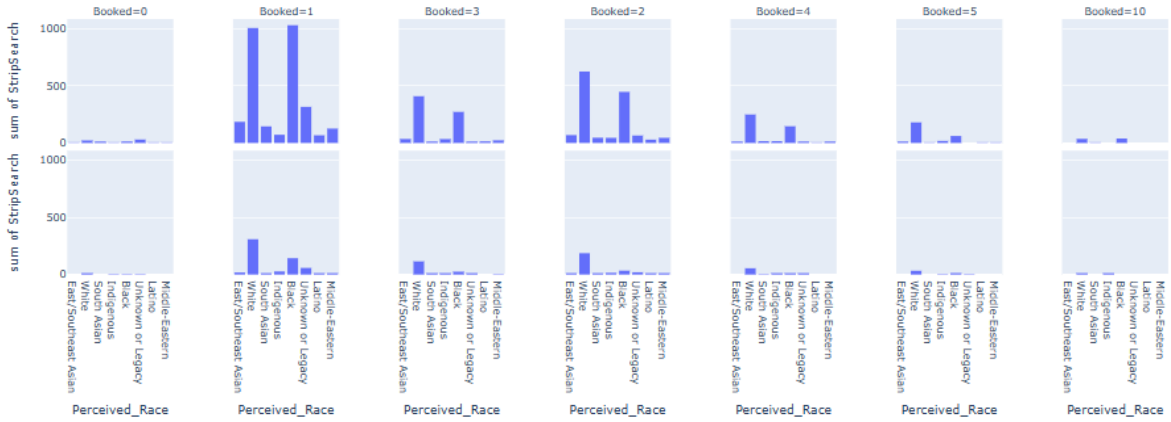
*Heatmap of pattern of strip search with regard to race, gender, and Age*




The small multiple histogram (figure 6) conducted on gender, race, and booked shares some other noteworthy patterns related to strip search. It captures that a significant number of strip searches often appear with an existing booked history. Arrestees regardless of gender performed a boost in strip search when they have history of booked, whereas individuals who do not have a history of booked have almost no strip search experience at all. This phenomenon may indicate the variable of Booked as an influential factor to the overall research.

**Figure 9**

*Small Multiple Histogram of strip search pattern with regard to race, gender, and Booked.*

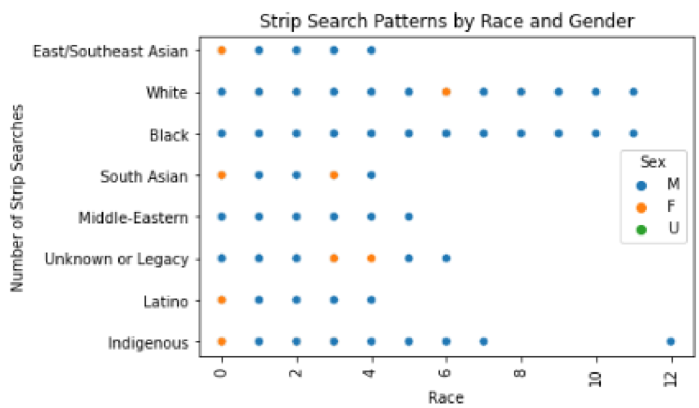



4. ***Bivariate Analysis - Scatter plot***

To exhibit the empirical association of strip search with regard to race and gender, the research conducts a scatter plot to obtain the pattern of strip search. The examination of the scatter plot indicates that strip searches are frequently conducted in both white and black communities, with a maximum number of searches reaching 12 times for both races. In contrast, other races generally experience a maximum number of strip searches at five times.

**Figure 10**

*Scatter Plot of strip search distribution expressed by race and gender*

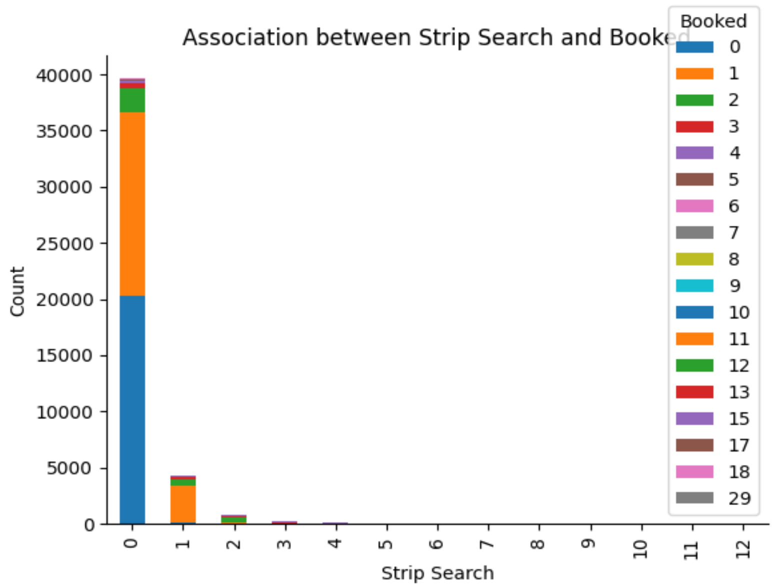


The scatter plot also captures unique findings that were not observed by previous visualizations. While the small multiple histogram expresses that white female arrestees have relatively high involvement in strip search, the scatter plot reveals that females from South Asian and unknown races also uncover active participation in strip search. Females from South Asia and unknown races had strip search experience three times higher than other races, and the majority of the females in other races have no strip search experience at all.

Lastly, the research has addressed the previous concern on the high association found between variables of strip search and booked. Based on the figure 8, the association between strip search and Booked are explored. Half of the arrestees who have not participated in the strip search have no booking history in the Toronto Police Service, whereas the majority of the arrestees who have strip search experience have more than 1 booking history.

**Figure 11**

*Association between Strip Search and Booked*




Based on all the exploratory data analysis conducted above, the research develops a comprehensive understanding of the impacts of race and gender on the number of strip searches. The visualization discovers that race and gender may shape the pattern of strip search, while the Booked can be highly associated with strip search.

### *5. T-test*

T-test is used to test the null hypothesis and determine whether the observed difference between the means is statistically significant. The null hypothesis for the independent t-test is that there is no significant difference between the means of two samples

(H0: $\mu = \mu 0$), while the alternative hypothesis means there is a statically significant difference (H1: $\mu \neq \mu 0$) (Mashra et al., 2019). In this case, our null hypotheses are 1) there is no difference in the mean frequency of strip searches between male and female. 2) there is no difference in the mean frequency of strip searches between perceived races: specifically test between White & Black, White & Indigenous, and Black & South Asian. Using the "stats.ttest_ind' function in the Scipy package and 'rp.ttest' in Researchpy package in Python, we got the same t-test results for each test group.

From the two-sample t-test for the comparison of strip searches by gender, the results showed that the mean frequency of strip searches for males ($M$=0.17, $SD$=0.58) was significantly higher than for females ($M$=0.14, $SD$=0.51), $t(45270)= 5.65$, $p < 0.05$, which suggest there is significant difference between the male and female on strip searches (Table 1).

**Table 1**

*t-test Comparison of strip searches by gender*

| | N | *Mean* | *SD* | *t-value* | *p-value* |
|---|---|---|---|---|---|
| Male | 36226 | 0.17 | 0.58 | 5.65 | 1.53e-08 |
| Female | 9046 | 0.14 | 0.51 | | |

From the two-sample t-test for the comparison of strip searches by perceived race (White & Black), the results showed no significant difference in the mean frequency of strip searches between White ($M$=0.21, $SD$=0.65) and Black ($M$=0.21, $SD$=0.62) arrestees, the p-value of t-test ($p$ = 0.88) is greater than the significance level alpha = 0.05.

**Table 2**

*t-test Comparison of strip searches by perceived race: White & Black*

| | N | *Mean* | *SD* | *t-value* | *p-value* |
|---|---|---|---|---|---|
| White | 17386 | 0.21 | 0.65 | -0.14 | 0.88 |
| Black | 11805 | 0.21 | 0.62 | | |

According to the results of the t-test analysis, there was a significant difference in the mean frequency of strip searches between Indigenous and White arrestees. Specifically, Indigenous arrestees had a higher mean frequency of strip searches ($M$=0.30, $SD$=0.85) compared to White arrestees ($M$=0.21, $SD$=0.65), $t(18390)$=-4.58, $p$<0.05 (Table **3**).

**Table 3**

*t-test Comparison of strip searches by perceived race: White & Indigenous*

| | N | *Mean* | *SD* | *t-value* | *p-value* |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| White | 17386 | 0.21 | 0.65 | -4.58 | 4.47e-06 |
| Indigenous | 1006 | 0.30 | 0.85 | | |

Table 4 presents the results of the t-test comparison of strip searches by perceived race for Black and South Asian groups. The mean frequency of strip searches for Black arrestees (*M*=0.21, *SD*=0.62) was significantly higher than for South Asian arrestees (*M*=0.08, *SD*=0.33), $t(13809)=10.25$, $p<0.05$, which suggest there is a significant difference in the mean frequency of strip searches between the Black and South Asian.

**Table 4**

*t-test Comparison of strip searches by perceived race: Black & South Asian*

| | N | *Mean* | *SD* | *t-value* | *p-value* |
|---|---|---|---|---|---|
| Black | 11805 | 0.21 | 0.62 | 10.25 | 1.31e-24 |
| South Asian | 2006 | 0.08 | 0.33 | | |

Before we ran the t-test, we conducted the Shapiro-Wilk test to check whether the data were normally distributed. However, due to the sample size, N > 5000, the p-value may not be accurate, which indicated $p = 0$, so we decided to keep conducting the t-tests (Appendix A).

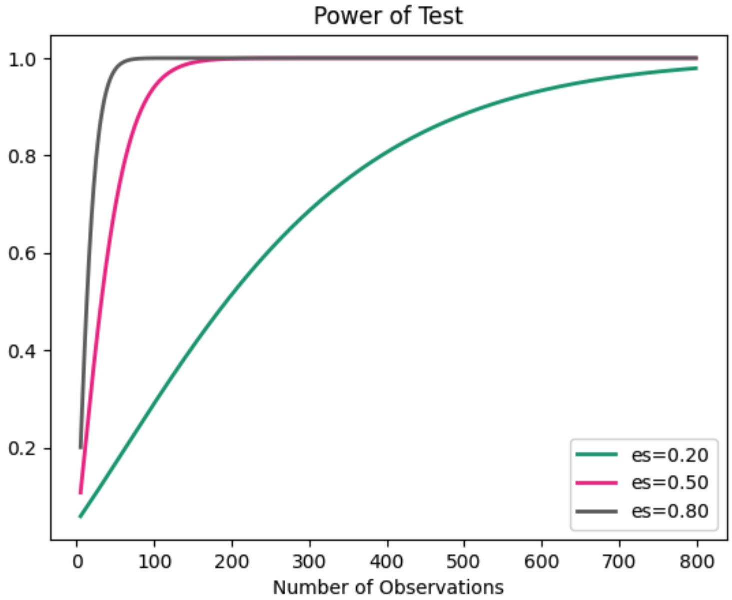
### *6. Power Analysis*

Power analysis is often used to determine the sample size needed to detect a statistically significant effect, given a certain significance level, effect size, and power (Brownlee, 2018). By conducting power analysis, it helps to ensure that the study has a sufficient sample size to detect a significant effect if one exists. With a sufficient sample size, we could detect Type I errors (the null hypothesis is rejected when it is actually true, false positive) and Type II errors (fails to reject a null hypothesis that is actually false, false-negative) to avoid incorrect conclusions.

A power analysis is a crucial tool for researchers to determine the minimum sample size required for a study and to assess the statistical power of the study. The power analysis has four main components: 1) Effect size; 2) Sample size, which refers to the number of observations in each group; 3) Significance level; and 4) Statistical power. In this study, we used a significance level of 0.05 and a statistical power of 0.80 to calculate the required sample size for each group and the effect size for hypothesis testing. Based on the calculated effect size and sample size, we found that a minimum of 37 samples were required to observe an effect of the desired size, with an effect size of -0.66. Compare with the actual sample size the model currently has, it is sufficient to support the following analysis and we have chance of detecting significant effects. Additionally, we plotted a power curve to demonstrate the impact of sample size on statistical power for three different effect sizes. The plot showed that there is a point of diminishing returns in terms of statistical power, occurring between 50 to 350 observations, in a large effect (Figure 12).

**Figure 12**

*Statistical power at variable effect sized a=0.05*




## Methods

In the previous sections, the research captures the key patterns relevant to the research interest. The t-test and exploratory data analysis observed that strip search patterns are shaped by the race and gender of the arrestees, while the variables of Booked History are highly associated with strip search. To narrow the scope of the research and improve overall precision, the research will conduct ANCOVA test to confirm the association between race and strip search, while removing the impacts of Booked by setting it as the covariate in ANCOVA. If confirmed, the research will use logistic regression analysis to examine how gender and race affect this relationship. Overall, the use of ANCOVA and logistic regression allow the research to explore the association between variables, while controlling the influential factor constant.

**Results**

1. **ANCOVA**

ANCOVA is used to analyze the relationship between the means of two or more independent groups while controlling for the effects of covariates on the dependent variable (Khammar et al., 2020). Similar to ANOVA, the null hypothesis for ANCOVA is that there is no significant difference between the means of the groups, but it considers the effect of a covariate on the dependent variable. By incorporating the continuous covariate into the analysis, ANCOVA can adjust for its influence on the dependent variable and reduce the potential impact on the result, which results in more accurate estimates of the effect and increase statistical power by controlling the covariate.

The purpose of this study is to investigate whether there is a difference in the frequency of strip searches conducted by police officers on individuals of different races, after controlling for the individual's booked record. Specifically, the study aims to examine the effects of perceived race as an independent variable on the frequency of strip searches, with the individual's booked status serving as the covariate. To test the null hypothesis that there is no significant difference in the frequency of strip searches based on race, a one-way ANCOVA analysis will be conducted. Based on the ANCOVA results, it suggest that there was a difference (at the $p < 0.05$ level) in the frequency of strip searches conducted by police officers based on the individual's race, after controlling for the effect of Booked, ($F$ (7, 45241) = 14.614, $p < 0.05$).

**Table 5**

*Results of One-way ANCOVA*

| | Sum of Squares | *df* | *F* | *p* |
|---|---|---|---|---|
| Perceived_Race | 24.36 | 7 | 14.614 | 3.81e-19 |
| Booked | 3931.61 | 1 | 16508.75 | 0 |
| Residual | 10774.29 | 45241 | | |

**Post Hoc Test**

However, one-way ANCOVA wasn’t able to identify whether there are significant differences in the dependent variable among the different groups while controlling for the covariate. Therefore, an post-hot test was conducted for pairwise comparisons using a Tukey HSD test. This test examined the differences between groups, revealing which pairs of groups had statistically significant differences in means, and minimizing the chance for Type I error in hypothesis testing and yielding a more accurate estimation of significance (S. Lee & D. Lee, 2018). The result of the Tukey HSD test (Table 6) shows that there were eighteen group pairs whose strip searches were statistically different from each other (all $p < 0.5$). Furthermore, a total of ten group pairs showed no significant differences from each other (all $p > 0.5$).

**Table 6**

*Results of Tukey HSD for One-way ANCOVA*

| Group 1 | Group 2 | meandiff | p-adj | Lower | Upper |
|---|---|---|---|---|---|
| 1 | 2 | -0.11 | 0.0 | -0.15 | -0.07 |
| | 3 | 0.10 | 0.0 | 0.04 | 0.15 |
| | 4 | -0.11 | 0.0 | -0.16 | -0.06 |
| | 5 | -0.12 | 0.0 | -0.15 | -0.07 |
| | 6 | -0.12 | 0.0 | -0.15 | -0.08 |
| | 7 | -0.09 | 0.0 | -0.11 | -0.05 |
| 2 | 3 | 0.21 | 0.0 | 0.15 | 0.27 |
| | 8 | 0.11 | 0.0 | 0.08 | 0.14 |
| 3 | 4 | -0.2 | 0.0 | -0.28 | -0.13 |
| | 5 | -0.21 | 0.0 | -0.27 | -0.14 |
| | 6 | -0.21 | 0.0 | -0.28 | -0.15 |
| | 7 | -0.18 | 0.0 | -0.24 | -0.12 |
| | 8 | -0.09 | 0.0 | -0.15 | -0.04 |
| 4 | 8 | 0.11 | 0.0 | 0.06 | 0.15 |
| 5 | 8 | 0.12 | 0.0 | 0.09 | 0.15 |
| 6 | 8 | 0.12 | 0.0 | 0.09 | 0.15 |
| 7 | 8 | 0.09 | 0.0 | 0.06 | 0.12 |

**1-Black, 2-East/South East Asian, 3-Indigenous, 4-Latino, 5-Middle-Eastern, 6-South Adian, 7-Unknown or Legacy, 8-White*

## 2. Logistic Regression

Logistic regression used to model the relationship between a binary dependent variable and one or more independent variables. In this study, we aim to explore whether an individual has been strip searched based on their underlying social identity. This research question stems from our earlier finding that there was a difference in the frequency of strip searches conducted by police officers based on the individual's race, even after controlling for the effect of being booked. Specifically, strip seaches was converted to binary dependent variable indicating whether a strip search occurred (0 = no, 1 = yes), while race and gender served as independent variables. By using logistics regression, we hope to identify any significant predictors of strip searches and examine the relationship between social identity and the likelihood of being subjected to a strip search by a police officer.

The model result suggests that there is a statistically significant relationship between an individual's race, gender, and the likelihood of being subjected to a strip search by a police officer. The result indicated individuals who identify as East/Southeast Asian, Latino, Middle-Eastern, South Asian, and Unknown or Legacy are significantly less likely to be strip searched compared to White, and Indigenous. Furthermore, the model suggests that gender had a significant positive effect on the strip searches by police officers. Males are more likely to be strip searched compared to females, with a statistically significant coefficient of 0.339 ($p<0.001$). This finding is consistent with previous research conducted in the midterm project that has shown a gender bias in police officers' decisions to conduct strip searches.

**Table 7**

*Logistic Regression Results*

| | coef | *std err* | *z* | P>\|z\| | 95% CI |
|---|---|---|---|---|---|
| Intercept | -3.612 | 0.093 | -38.837 | 0.00 | [-3.795 - -3.430] |
| **Race** | | | | | |
| East/Southeast Asian | -0.997 | 0.173 | -5.775 | 0.00 | [-1.336 - -0.659] |
| Indigenous | 0.416 | 0.169 | 2.734 | 0.00 | [0.131 - 0.792] |
| Latino | -1.042 | 0.275 | -3.79 | 0.00 | [-1.581- 0.504] |
| Middle-Eastern | -1.117 | 0.213 | -5.26 | 0.00 | [-1.534 - -0.7] |
| South Asian | -1.272 | 0.213 | -5.98 | 0.00 | [-1.689 - -0.855] |
| Unknown or Legacy | -1.084 | 0.213 | -6.63 | 0.00 | [-1.404 - -0.763] |
| White | 0.079 | 0.072 | 1.107 | 0.268 | [-0.062 - 0.221] |
| **Gender** | | | | | |
| Sex | 0.339 | 0.086 | 3.935 | 0.00 | [0.17 - 0.508] |

### Odds Ratio

To determine the effect of a predictor variable on the odds of an outcome occuring, we then calculated odds ratio for the logistic regression model. The odds ratio for each category of the "Perceived race" and "Sex" show either a postive or negative association with the strip searches. For instance, the odds of the strip searches occuring for individuals perceived as Indiigenous are 1.581 times the odds of rhe strip searches occuring for the reference group, indicating a positive association. Conversely, the odds of the strip searches occuring for individuals perceived as Latino are 0.26 times the odds of the strip searches for the reference group, indicating a negative association (Table 8).

**Table 8**

*Odds Ratio*

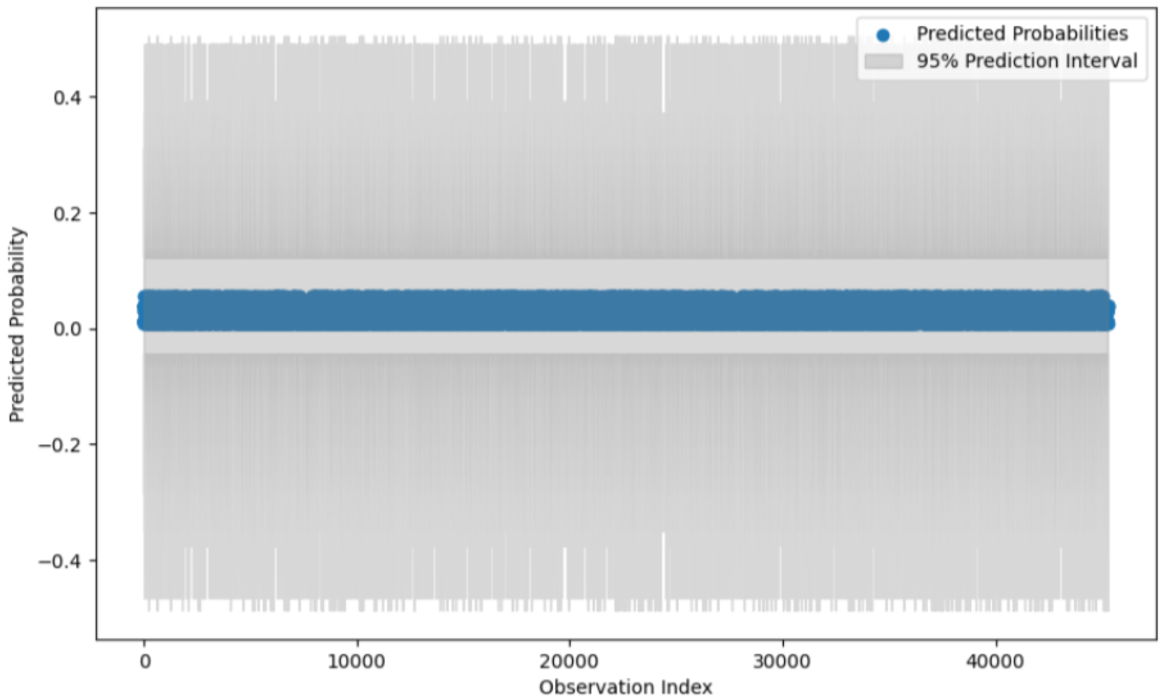
| | Odds Ratio |
|---|---|
| Intercept | 0.026 |
| **Race** | |
| East/Southeast Asian | 0.39 |
| Indigenous | 1.58 |
| Latino | 0.26 |
| Middle-Eastern | 0.36 |
| South Asian | 0.27 |
| Unknown or Legacy | 0.34 |
| White | 1.14 |
| **Gender** | |
| Sex | 1.39 |

### Prediction Interval

The prediction interval defines the range of levels which the future observation of our dataset is expected to fall. According to the prediction interval we plot, the majority of the lower bound and the upper bound lie between the prediction probability of (0.00, 0.1), suggesting that the 95% confidence interval of the logistic model falls in between. Overall, the prediction interval helps to provide a measure of uncertainty about the future observations of a dataset and is an essential tool for statistical analysis.

**Figure 13**

*95% Prediction Interval for Prediction Probability*




**Discussion**

The results of the ANCOVA test demonstrate that an individual's race is a significant factor in determining the frequency of strip searches conducted by police officers, even after controlling for booking records. And the Logistic regression model suggests that an individual's social identity can influence the likelihood of being subjected to a strip search by law enforcement. In this section, we will dive deeper into the discussion of the limitations of the study, the implications of the findings, and discuss the opportunity for future research.

During the study analyze phase, our primary focus was to approach the research question with a high level of reliability and accuracy. But one limitation of this study is that we did not use QQ plot or Shapiro test to check for normality of the variables. Since the basic criterion for running these tests is that the data should be continuous, we could not apply these tests to most of the variables in our dataset, such as perceived race, sex, and age, as they are categorical. This limitation could potentially affect the accuracy of our analysis as normality is an important assumption for many statistical tests. However, we addressed this limitation by utilizing other methods such as histograms and violin plots to examine the distribution of these categorical variables in the EDA stage. While these methods provide a visual representation of the distribution of the categorical variables and allow us to identify any potential skewness or outliers, they may not be as precise as normality tests in detecting deviations from normality. Therefore, it is important to interpret the results of our study with caution and acknowledge this limitation. Other than that, we discovered the prediction interval we plotted matches poorly with the confidence interval we determined previously. The non-match of the confidence interval and the prediction interval might address an inaccurate data modeling we fall to detect in our research.

In addition, we realized the uneven distribution of races in the dataset after conducting the histogram and violin test. The number of white and black individuals in the dataset is much higher than other groups, which may have influenced the results. It is possible that the overrepresentation of white and black individuals in the dataset could have biased the results. The findings may not be representative of the entire population, and caution should be taken when generalizing the results to other contexts or populations with different racial distributions. Future studies could address this limitation by ensuring a more representative sample of the population.

Despite those limitations, this study has significant implications for addressing bias and discrimination in law enforcement. The fact that there are significant association between race, gender, and strip searches, suggests that police officials may be influenced by unconscious biases or stereotypes when deciding whether to conduct a strip search on the arrestees. In addition, it highlighted the need for conducting more research in this area to investigate further potential factors contributing to bias and discrimination in law enforcement to promote fairness and equality in law enforcement practices.

## Conclusion

Overall, the findings of this research focus on a concerning issue in the Toronto Police service, highlighting the disproportionate use of force against individuals with different race and gender in the context of strip search. By conducting advanced statistical analysis, this research has produced more reliable and comprehensive results based on the previous research findings. The ANCOVA test and logistic regression validated the association between strip search patterns and the arrestee's race and gender, while controlling for the variable of Booked. The establishment of these valuable findings confirms the unjust treatment and racial disparities raised by previous public discourse. Based on the findings, a pressing need for impartial treatment of all races and gender should be promoted in law enforcement activities,  thereby enhancing accountability and fairness of the criminal justice system of Toronto.

**Reference**


Brownlee, Jason. "A Gentle Introduction to Statistical Power and Power Analysis in Python - MachineLearningMastery.com." *Machine Learning Mastery*, 13 July 2018, https://machinelearningmastery.com/statistical-power-and-power-analysis-in-python/. Accessed 15 April 2023.

Fox, Chris. "Toronto police's race-based data on use of force, strip searches highlighted." *CTV News Toronto*, 15 June 2022, https://toronto.ctvnews.ca/toronto-police-s-race-based-data-on-use-of-force-strip-searches-highlighted-1.5947778. Accessed 15 April 2023.

Ontario Human Rights Commission (OHRC). "OHRC statement on Toronto Police Service announcement on race-based data collection findings | Ontario Human Rights Commission." *Ontario Human Rights Commission* |, 16 June 2022, https://www.ohrc.on.ca/en/news_centre/ohrc-statement-toronto-police-service-announcement-race-based-data-collection-findings. Accessed 3 April 2023.

Wikipedia. "Demographics of Toronto." *Wikipedia*, https://en.wikipedia.org/wiki/Demographics_of_Toronto. Accessed 15 April 2023.