# Factors that Affect the Number of Times Strip Searches in Minors - Based on the Dataset from Toronto Police Service

**Group 33**

Rimsha Rizvi

Yijie Shi

Faculty of Information

University of Toronto

INF2178: Experimental Design for Data Science

Prof. Shion Guha

February 28, 2023

## Introduction

Strip searches are a controversial and often invasive law enforcement practice used to detect hidden contraband or weapons. While strip searches are intended to be carried out only under certain circumstances and with proper oversight, concerns have been raised about their potential misuse or abuse, particularly among vulnerable populations such as minors or people of color. In recent years, reports have suggested that black and Latino individuals are at higher risk of being subjected to strip searches than other racial groups. This study aims to examine the relationship between the likelihood of strip searches and specific demographic characteristics such as age, race, and gender. Specifically, the report will investigate which suspicions are most likely to result in a strip search for minors, with the aim of shedding light on any potential issues or disparities related to the use of this practice. By analyzing data and patterns, this report aims to provide insight into any potential concerns or areas for improvement related to the use of strip searches in law enforcement, particularly among vulnerable populations.

## Literature Review

Strip checking, the practice of requiring individuals to remove some or all of their clothing for the purpose of security screening, has been a controversial and often debated topic for many years.

> *"I can't say no. No is not an option here. There is no choice, no consent, no opt out. My body is not mine.*
>
> *But of course, I can't express panic. I can't show the nausea and fear that has overtaken my body. Instead I just nod and obey her commands."*
> (McMurphy, 2018)

The excerpt written by McMurphy in the article describes the distressing emotional experience that a young college student endured during her first strip search. Strip checking can cause physical and psychological harm, particularly in minors, who may not fully understand the reasons for the screening and may be embarrassed or traumatized by the process. Furthermore, some studies have suggested that strip checking practices may be influenced by the race and gender of the individuals being screened.

> *"...concerned by ethnic disproportionality after the data showed that of children aged 10 to 17 who were strip-searched between 2018 and 2020, almost three out of five (58%) were black, as described by the officer."* (Weale & Dodd, 2022)

The data reveals that nearly three out of five children who were subjected to strip searches during this time period were black, according to an officer's description. This is a particularly sensitive issue, as it may have lasting emotional and psychological impacts on young people. Another source of concern is that certain ethnic or racial groups may be more likely than others to be subjected to strip searches, even when no reasonable suspicion of wrongdoing exists.The evidence suggests that people of color, particularly Black and Latin American individuals, may be disproportionately targeted for strip searches. For example, in the United States, a 2013 report by the American Civil Liberties Union found that black and Latinx drivers were more likely to be subjected to strip searches during traffic stops than white drivers. Similarly, in the UK, a recent report found that black children were disproportionately subjected to strip searches by police, even in cases where there was no evidence of criminal activity.

> *"The youngest person to be strip-searched was 12. Nearly half – 48% – of children strip-searched were from black and minority ethnic communities…Tobacco was the most common item found, with no recorded discoveries of drugs or knives…"*
> (Allison, 2013)

Based on the findings of the 2013 study, some young kids suspected of carrying knives or other dangerous weapons were discovered to only have tobacco in their possession after being subjected to a strip search. This raises concerns about the use of strip searches as a first resort, as well as the possibility that law enforcement will rely on stereotypes or assumptions about specific groups of people rather than conducting a thorough and fair investigation. It also emphasizes the importance of having clear guidelines and policies in place regarding the use of strip searches, as well as training and education on non-invasive methods of investigation and de-escalation.

In this report, we will be looking at various factors affecting the likelihood of strip checking in minors.

## Research Objective and Questions

Building on the findings from our literature review, our objective is to investigate whether there is a disparity in the likelihood of strip searches between Black or Latino individuals and other racial groups. We will also examine whether there is a correlation between age and the probability of being subjected to a strip search. Additionally, this study aims to identify which gender has the highest likelihood of being strip searched in law enforcement encounters.

As part of this study, we will also investigate the factors that may influence the likelihood of strip checking of minors in law enforcement encounters. Specifically, we will explore the correlation between the reason for a search and the probability of strip checking among minors. For instance, we will examine whether minors suspected of evidence possession or weapons possession are more likely to be strip searched compared to minors suspected of other offenses. Additionally, we will consider other factors such as age, race, and gender that may influence the probability of strip checking among minors.

- **RQ1:** Can demographic attributes such as age, race, and sex influence the probability of strip searches in law enforcement encounters? Furthermore, is it possible to determine which attribute has the strongest impact on the likelihood of such searches?
- **RQ2:** Is there a correlation between the reason for a search or offense (e.g., suspicion of evidence possession, suspicion of weapon possession, suspicion of assault or suspicion of assisting escape) and the likelihood of strip-searching in law enforcement encounters? Additionally, which type of suspicion is most likely to result in a strip search?

## Dataset Description

We will conduct our analysis using Arrest and Strip Searches data downloaded from Toronto Police Service - Public Safety Data Portal (Toronto Police Service, 2022). This dataset contains information on 65,276 instances of arrest and strip searches and includes 25 different features.

Three variables are numeric values that uniquely identify an occurrence: *EventID, ArrestID*, and *PersonID. PersonID* is a unique identifier for each individual, while *ArrestID* may occur multiple times for an individual. Additionally, multiple suspects may or may not be involved in a single event, which is uniquely identified by *EventID*.

*Arrest_Year* and *Arrest_Month* are categorical variables that specify the year and month in which the event occurred. *Occurrence_Category* is another categorical variable that specifies the type of offense the individual is suspected for.

There are 5 demographic variables that are all categorical in nature: *Perceived_Race, Sex, Age_group__at_arrest_, and Youth_at_arrest__under_18_years*. Additionally, there are 10 dichotomous variables, including *StripSearch* (whether the individual was strip checked or not), *Booked* (whether the person was booked within 24hrs of arrest or not), and *ItemsFound* (whether an item was found during strip checking or not).

Furthermore, there is a list of variables that record actions at arrest and reasons for strip checking, which are recorded as boolean values (0 for no and 1 for yes). These variables include *Actions_at_arrest___Concealed_i, Actions_at_arrest___Combative__, Actions_at_arrest___Resisted__d, Actions_at_arrest___Mental_inst, Actions_at_arrest___Assaulted_o,* and *Actions_at_arrest___Cooperative* for actions, and *SearchReason_CauseInjury, SearchReason_AssistEscape, SearchReason_PossessWeapons,* and *SearchReason_PossessEvidence* for search reasons.

Lastly, *ArrestLocDiv* is a categorical variable that stores a numeric value for the division where the arrest took place.

## Descriptive Statistics

The dataset records the data of arrests and strip searches in the Toronto area from 2020 to 2021. As a result of many empty data in 2021, we don't have enough datasets to compare the trend across the year. We have a total of 65,276 datasets and dropped the empty data before analysis. We created a new numerical variable that describes the number of times an arrestee with the same person ID has been strip checked between 2020 and 2021. The number of times ranges from 1 to 17. To further investigate the effect on the number of strip searches, we started our analysis with demographic attributes and 4 reasons for search as independent variables. In addition to the analysis of all persons, we focused on the age group of minors.

**Table 1. Descriptive Statistics for Variables in Analysis (N = 65,276)**

| Variables | Frequency | % |
|---|---|---|
| *Independent Variables* | | |

| | | |
|---|---|---|
| **Age (in years)** | | |
| under 17 | 3042 | 4.7 |
| 18 - 24 | 10041 | 15.3 |
| 25 - 34 | 20949 | 32.0 |
| 35 - 44 | 16242 | 24.9 |
| 45 - 54 | 9066 | 13.9 |
| 55 - 64 | 4590 | 7.0 |
| above 65 | 1322 | 2.0 |
| **Sex** | | |
| Female | 12617 | 19.4 |
| Male | 52650 | 80.6 |
| Unisex | 9 | 0.01 |
| **Perceived Race** | | |
| Black | 17526 | 26.8 |
| East/Southeast Asian | 4415 | 6.76 |
| Indigenous | 1934 | 2.96 |
| Latino | 1768 | 2.70 |
| Middle-Eastern | 3237 | 4.96 |
| South Asian | 3613 | 5.53 |
| Unknown or Legacy | 5056 | 7.75 |
| White | 27723 | 42.5 |
| **Search Reasons** | | |
| Cause Injury | | |
| No | 1828 | 2.8 |
| Yes | 5973 | 9.15 |
| Assist Escape | | |
| No | 5124 | 7.85 |
| Yes | 2677 | 4.10 |
| Possess Weapons | | |
| No | 4264 | 6.53 |
| Yes | 3537 | 5.42 |
| PossessEvidence | 4533 | 6.94 |
| No | 3268 | 5.01 |
| Yes | | |

| | **Mean** | **SD** | **Range** |
|---|---|---|---|

*Dependent Variable*

| | | | |
|---|---|---|---|
| Strip Searched | 2.50 | 2.24 | 1 - 17 |

In the beginning, we looked at the overall distribution of the number of strip searches through histograms (Figure 1), with most people concentrating on having one strip search. Only one outlier people with 17 times of strip searches. From table 1, we can see that the average number of times strip searches is at least twice per person. For variance, we can calculate from the standard deviation that $2.24^2 = 5.02$.

**Figure 1**

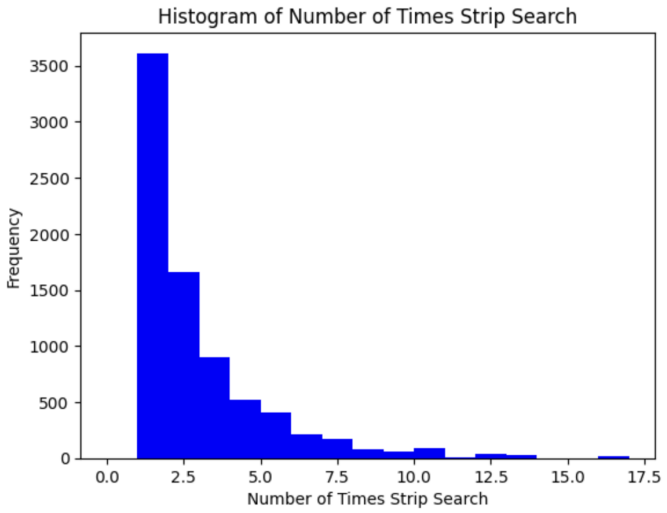
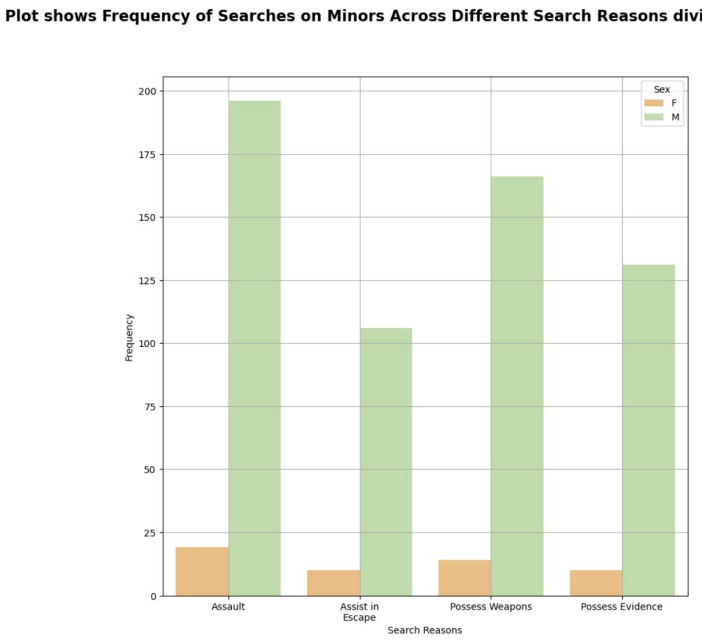
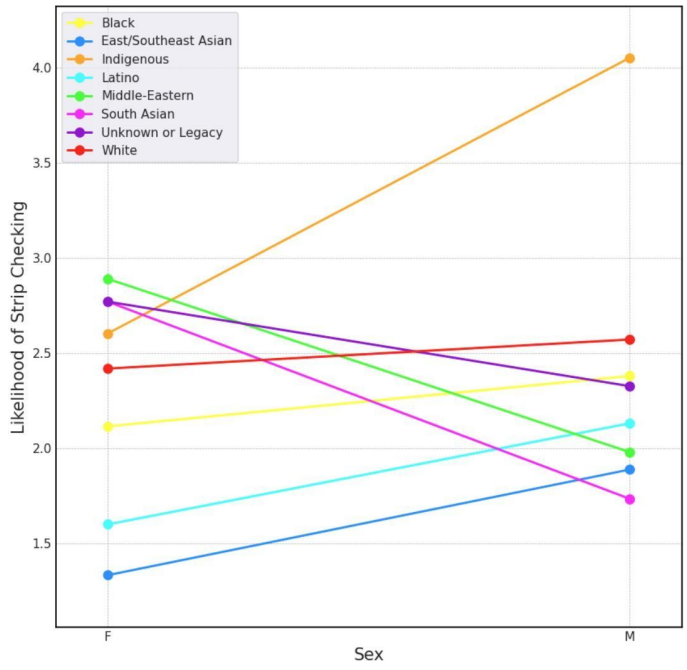


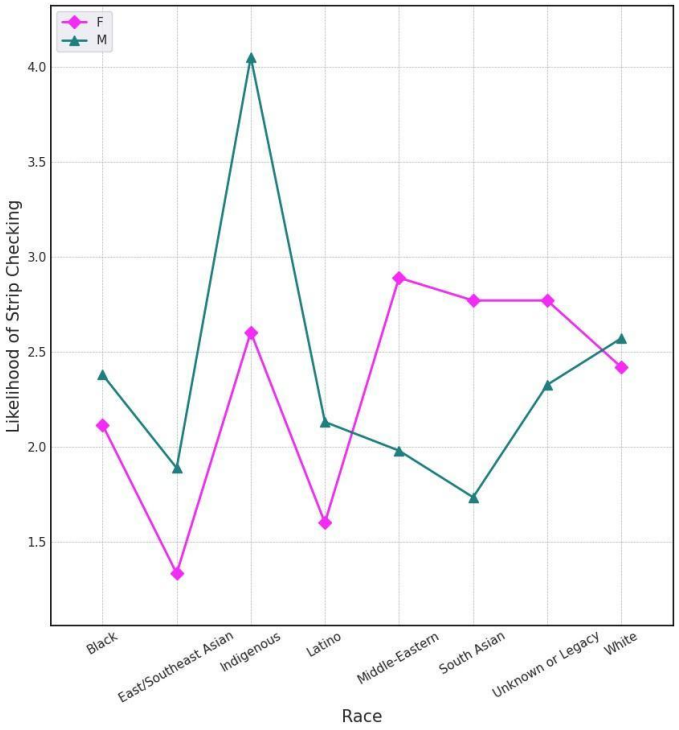

**Figure 2**



From the bar chart in Figure 2, we would see there had large differences in the frequency of searches between the males and females aged 17 and younger. Male minors have a higher frequency of searches for all four search reasons than female minors.

**Figure 3 Interaction Plots to show the mean number of times strip searches by sex and race**



These interaction plots(Figure 3) visualized basic information on relationships between sex and race of arrestees. We can see that the average strip search for males is higher than for females. Black males have the highest likelihood of strip checking than other races. There was a big different mean of the likelihood of strip checking between the Indigenous females and males.

## Results

We designed many t-tests with categorical variables to find how they are related. Comparing the differences between the average of two sample sets. The following contents are all the tests we did.

**For research question 1:**

**T-Test**

***Sex with the number of times strip search***

We created the numerical outcome (dependent) variable The number of Times Strip Search by the categorical (independent) variable Sex. We calculated the mean number of times strip searches for males and females. The result we got is that the mean number of times is higher for the person who is male. The null and alternative hypotheses we tested:

> Null hypothesis: $H_0: \mu_1 = \mu_2$ ( The population means of the two independent groups, male and female are equal)
>
> Alternative Hypothesis:$H_1: \mu_1 \neq \mu_2$ (The population means of the two independent groups, male and female are not equal)

From the T-test output, the mean number of times strip searches for males (mean=2.51,sd=2.27) is higher than the mean number of times for females (mean=2.45,sd=2.08). We can see that the t-statistic is 0.09 and the p-value is 0.38 which is greater than 0.05 (alpha level). We can not reject the null hypothesis. As a result, there is no significant difference in the mean number of times strip searches between the males and the females.

***Race with the number of times strip search***

In this data, race is divided into 9 categories, and we selected only the 3 groups with the highest number of people, which are Black or Non-Black, South Asian or Non-South Asian, and White or Non-White.

*Black or Non-Black*

We created the numerical outcome (dependent) variable The number of Times Strip Search by the categorical (independent) variable Race. We calculated the mean number of times strip searches for Black and Non-Black. The result we got is that the mean number of times is higher for the person who is Non-Black. The null and alternative hypotheses we tested:

Null hypothesis: $H_0: \mu_1 = \mu_2$ ( The population means of the two independent groups, Black people and Non-Black people are equal)

Alternative Hypothesis:$H_1: \mu_1 \neq \mu_2$ (The population means of the two independent groups, Black and Non-Black are not equal)

From the T-test output, the mean number of times strip searches for Non-Black (mean=2.55,sd=2.25) is higher than the mean number of times for Black (mean=2.39,sd=2.22). We can see that the t-statistic is -2.89, the p-value is 0.004 which is less than 0.05 (alpha level) and the 95% confidence interval [-0.26,-0.05]. We can reject the null hypothesis. As a result, there is a significant difference in the mean number of times strip searches between black and non-black.

*South Asian or Non-South Asian*

Similarly, we calculated the mean number of times strip searches for South Asian and Non-South Asian. The result we got is that the mean number of times is higher for the person who is Non-South Asian. The null and alternative hypotheses we tested:

Null hypothesis: $H_0: \mu_1 = \mu_2$ ( The population means of the two independent groups, South Asian people and Non-South Asian people are equal)

Alternative Hypothesis:$H_1: \mu_1 \neq \mu_2$ (The population means of the two independent groups, South Asian people and Non-South Asian people are not equal)

From the T-test output, the mean number of times strip searches for Non-South Asian (mean=2.52,sd=2.26) is higher than the mean number of times for South Asian (mean=1.87,sd=1.50). We can see that the t-statistic is -6.72 and the p-value is 9.48e-11 which is less than 0.05 (alpha level) and the 95% confidence interval [-0.84,-0.46]. We can reject the null hypothesis. As a result, there is a significant difference in the mean number of times strip searches between South Asian and non-South Asian.

*White or Non-White*

Comparing the minority and majority, we calculated the mean number of times strip searches for White and Non-White. The result we got is that the mean number of times is higher for the person who is White. The null and alternative hypotheses we tested:

Null hypothesis: $H_0: \mu_1 = \mu_2$ ( The population means of the two independent groups, White people and Non-White people are equal)

Alternative Hypothesis:$H_1: \mu_1 \neq \mu_2$ (The population means of the two independent groups, White and Non-White are not equal)

From the T-test output, the mean number of times strip searches for White (mean=2.62,sd=2.19) is higher than the mean number of times for Non-White (mean=2.40,sd=2.28). We can see that the t-statistic is 4.42 and the p-value is 9.79e-06 which is less than 0.05 (alpha level) and the 95% confidence interval [0.12,0.32]. We can reject the null hypothesis. As a result, there is a significant difference in the mean number of times strip searches between White and Non-White.

***Age group with the number of times strip search***

Comparing the minors and adults, we calculated the mean number of times strip searches for the minors and adults. The result we got is that the mean number of times is higher for the person who is an adult. The null and alternative hypotheses we tested:

Null hypothesis: $H_0: \mu_1 = \mu_2$ ( The population means of the two independent groups, people aged 17 and under and people aged 18 and above are equal)

Alternative Hypothesis:$H_1: \mu_1 \neq \mu_2$ (The population means of the two

independent groups, people aged 17 and under and people aged 18 and above are not equal)

From the T-test output, the mean number of times strip searches for the person aged 18 and above (mean=2.50,sd=2.24) is higher than the mean number of times for the person aged 17 and under (mean=1.22,sd=0.67). we can see that the t-statistic is -5.71 and the p-value is 0.0004 which is less than 0.05 (alpha level) and the 95% confidence interval [-1.79, -0.76]. we can reject the null hypothesis. As a result, there is a significant difference in the mean number of times strip searches between people aged 17 and under and aged 18 and above.

**Two-way ANOVA**

From the t-test results of demographic attributes showed that there is a significant difference in the average number of times strip searches through age group and race; then We want to discover more about the impact through the two-way ANOVA and know how these two factors interact. The null and alternative hypotheses we tested:

***Age Group and Race***

Set 1:

Null hypothesis: $H_0: \mu_{\alpha 1} = \mu_{\alpha 2} = \mu_{\alpha 3} = \ldots = \mu_{\alpha 9}$ ( All means are equal under different age groups)

Alternative Hypothesis:$H_1: \mu_{\beta 1} \neq \mu_{\beta 2} \neq \mu_{\beta 3} \neq \ldots \neq \mu_{\beta 9}$ (Not all means are equal under different age groups)

Set 2:

Null hypothesis: $H_0: \mu_{\alpha 1} = \mu_{\alpha 2} = \mu_{\alpha 3} = \ldots = \mu_{\alpha 8}$ ( All means are equal under different race)

Alternative Hypothesis:$H_1: \mu_{\beta 1} \neq \mu_{\beta 2} \neq \mu_{\beta 3} \neq \ldots \neq \mu_{\beta 8}$ (Not all means are equal under different race)

Set 3:

Null hypothesis:$H_0$: There is no interaction between age groups and race

Alternative Hypothesis:$H_0$: There is an interaction between age groups and race

From the two-way ANOVA output, we can see that set 1 and set 2 of F-statistic are 34.62 and 13.80. For the p-value4.17e-09 and 3.26e-11, both are less than 0.05(alpha level). We can reject the null hypothesis. As a result, we have enough evidence that not all means are the same for different age groups and races. For set 3, the p-value is 1.44e-07, less than the alpha level of 0.05, we can reject the null hypothesis and get the result that there is a significant interaction effect between age groups and races on the number of times strip searches.

After the two-way ANOVA, we conducted pairwise comparison testing (details in Appendix) for race and age group. In race groups, the mean of the likelihood of strip searches has the most significant difference between South Asians and Blacks ($p< 0.05$). For age groups, people aged 18 to 24 years and people aged 25 to 34 years have the biggest mean difference ($p< 0.05$).

**For research question 2:**

**T-test**

Our goal was to identify which reason would influence a minor's likelihood of being strip searched and how to influence that likelihood.

***Causing injury as a searching reason and the likelihood of strip checking***

We calculated the mean likelihood of strip checking for causing injury as a searching reason and those who are not. The result we got is that the mean likelihood of strip checking is higher for the person who with causing injury. The null and alternative hypotheses we tested:

Null hypothesis: $H_0: \mu_1 = \mu_2$( There is no difference in the likelihood of strip checking between individuals suspected of assault and those who are not.)

Alternative Hypothesis:$H_1: \mu_1 \neq \mu_2$( There is a difference in the likelihood of strip checking between individuals suspected of assault and those who are not.)

From the T-test output, the mean likelihood of strip checking for causing injury as a reason (mean=1.39,sd=0.80) is higher than the likelihood of strip checking for those who are not (mean=1.24,sd=0.62). We can see that the t-statistic is 1.15 and the p-value is 0.25 which is greater than 0.05 (alpha level) and the 95% confidence interval [-0.07,0.38]. We can not reject the null hypothesis. As a result, there is no significant difference in the likelihood of strip checking between individuals who cause injury and those who do not.

***Assisting an escape as a searching reason and the likelihood of strip checking***

We calculated the mean likelihood of strip checking for assisting an escape as a searching reason and those who are not. The result we got is that the likelihood of strip checking is higher for the person who with assisting an escape. The null and alternative hypotheses we tested:

Null hypothesis: $H_0: \mu_1 = \mu_2$( There is no difference in the likelihood of strip checking between individuals with assisting an escape and those who are not.)

Alternative Hypothesis:$H_1: \mu_1 \neq \mu_2$( There is a difference in the likelihood of strip checking between individuals with assisting an escape and those who are not.)

From the T-test output, the mean likelihood of strip checking for assisting an escape as a reason (mean=1.51,sd=0.93) is higher than the likelihood of strip checking for those who are not (mean=1.22,sd=0.56). We can see that the t-statistic is 2.84 and the p-value is 0.005 which is less than 0.05 (alpha level) and the 95% confidence interval [0.08,0.50]. We can reject the null hypothesis. As a result, there is a significant difference in the likelihood of strip checking between individuals who assist an escape and those who do not.

***Possessing evidence as a searching reason and the likelihood of strip checking***

We calculated the mean likelihood of strip checking for possessing evidence as a searching reason and those who are not. The result we got is that the likelihood of strip checking is higher for the person who with possessing evidence. The null and alternative hypotheses we tested:

Null hypothesis: $H_0: \mu_1 = \mu_2$( There is no difference in the likelihood of strip

checking between individuals with possessing evidence and those who are not.)

Alternative Hypothesis:$H_1$: $\mu_1 \neq \mu_2$( There is a difference in the likelihood of strip checking between individuals with possessing evidence and those who are not.)

From the T-test output, the mean likelihood of strip checking for possessing evidence as a reason (mean=1.48,sd=0.86) is higher than the likelihood of strip checking for those who are not (mean=1.21,sd=0.61). We can see that the t-statistic is 2.58 and the p-value is 0.01 which is less than 0.05 (alpha level) and the 95% confidence interval [0.07, 0.47]. We can reject the null hypothesis. As a result, there is a significant difference in the likelihood of strip checking between individuals with possessing evidence and those who do not.

***Possessing weapons as a searching reason and the likelihood of strip checking***

We calculated the mean likelihood of strip checking for possessing weapons as a searching reason and those who are not. The result we got is that the likelihood of strip checking is lower for the person who with possessing weapons. The null and alternative hypotheses we tested:

Null hypothesis: $H_0$: $\mu_1 = \mu_2$( There is no difference in the likelihood of strip checking between individuals with possessing weapons and those who are not.)

Alternative Hypothesis:$H_1$: $\mu_1 \neq \mu_2$( There is a difference in the likelihood of strip checking between individuals with possessing weapons and those who are not.)

From the T-test output, the mean likelihood of strip checking for possessing weapons as a reason (mean=1.36,sd=0.75) is lower than the likelihood of strip checking for those who are not (mean=1.37,sd=0.83). We can see that the t-statistic is -0.08 and the p-value is 0.93 which is greater than 0.05 (alpha level) and the 95% confidence interval [-0.24, 0.22]. We can not reject the null hypothesis. As a result, there is no significant difference in the likelihood of strip checking between individuals with possessing evidence and those who do not.

**One-way ANOVA**

We combined the 4 types of search reasons into one categorical variable (a factor with four types) in order to deeply learn how these research reasons affect the likelihood of strip checking in minors (quantitative outcome). The null and alternative hypotheses we tested:

Null hypothesis: $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ ( All means are equal)

Alternative Hypothesis: $H_1$: at least one mean differs from the overall mean

From the one-way ANOVA result, we can see that the F-statistic is 0.67 and the p-value is 0.57 which is greater than 0.05(alpha level). We can not reject the null hypothesis. As a result, we have enough evidence that all means are the same for different types of search reasons. The search reasons are not the important factor to effect the likelihood of strip searches in minors.

## Discussion and Conclusion

Through a series of analyses and comparisons in this project. Initially, we started by looking at the effect of demographic characteristics on the probability of strip searches. It was determined that different age groups and different races have significant statistical significance in strip searches. As part of our RQ2, we chose to investigate the relationship between the reason for the search and the likelihood of strip searches among minors. As a result of our analysis, we found two search reasons (assisting an escape and possessing evidence) that influence the likelihood of strip searches. We dropped many missing data in 2021, and the remaining dataset amount does not allow us to analyze the two-year trend for this project. Through some literature readings, we learned that there is unfair treatment in strip searches and the number of strip searches is increasing. Hopefully, our findings will help people pay more attention to minorities and policymakers can take action to develop new rules to decrease unreasonable strip searches.

## References


Allison, E. (2013, March 3). *43000 strip-searches carried out on children as young as 12*. The Guardian. Retrieved February 27, 2023, from https://www.theguardian.com/society/2013/mar/03/43000-strip-searches-children

McMurphy, E. (2018, February 23). *Is Strip-Searching a Form of Sexual Abuse?* Rooted in Rights. Retrieved February 27, 2023, from https://rootedinrights.org/is-strip-searching-a-form-of-sexual-abuse/

Toronto Police Service. (2022, November 10). *Arrests and Strip Searches (RBDC-ARR-TBL-001)*. Public Safety Data Portal. https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/about

Weale, S., & Dodd, V. (2022, August 7). *Revealed: Met police strip-searched 650 children in two-year period*. The Guardian. Retrieved February 27, 2023, from https://www.theguardian.com/uk-news/2022/aug/08/police-data-raises-alarm-over-welfare-of-strip-searched-children

# Appendix

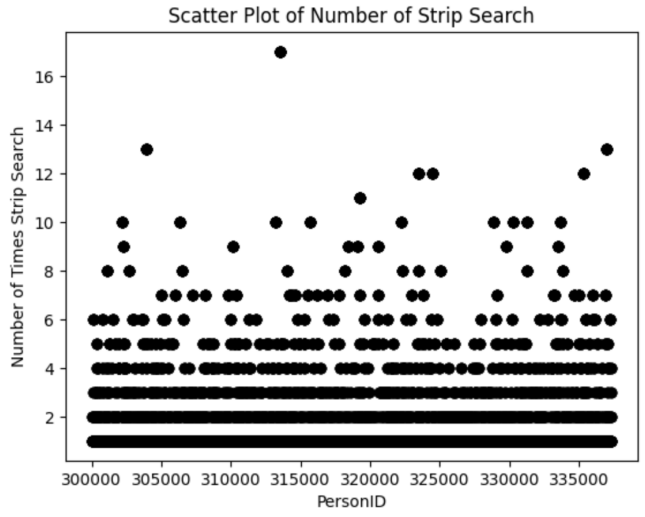
## Data Description

```
Data columns (total 25 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   Arrest_Year                     65276 non-null  int64
 1   Arrest_Month                    65276 non-null  object
 2   EventID                         65276 non-null  int64
 3   ArrestID                        64807 non-null  float64
 4   PersonID                        65276 non-null  int64
 5   Perceived_Race                  65272 non-null  object
 6   Sex                             65276 non-null  object
 7   Age_group__at_arrest_           65252 non-null  object
 8   Youth_at_arrest__under_18_years 65276 non-null  object
 9   ArrestLocDiv                    65276 non-null  object
 10  StripSearch                     65276 non-null  int64
 11  Booked                          65276 non-null  int64
 12  Occurrence_Category             65111 non-null  object
 13  Actions_at_arrest___Concealed_i 65276 non-null  int64
 14  Actions_at_arrest___Combative__ 65276 non-null  int64
 15  Actions_at_arrest___Resisted__d 65276 non-null  int64
 16  Actions_at_arrest___Mental_inst 65276 non-null  int64
 17  Actions_at_arrest___Assaulted_o 65276 non-null  int64
 18  Actions_at_arrest___Cooperative 65276 non-null  int64
 19  SearchReason_CauseInjury        7801 non-null   float64
 20  SearchReason_AssistEscape       7801 non-null   float64
 21  SearchReason_PossessWeapons     7801 non-null   float64
 22  SearchReason_PossessEvidence    7801 non-null   float64
 23  ItemsFound                      7801 non-null   float64
 24  ObjectId                        65276 non-null  int64
dtypes: float64(6), int64(12), object(7)
```

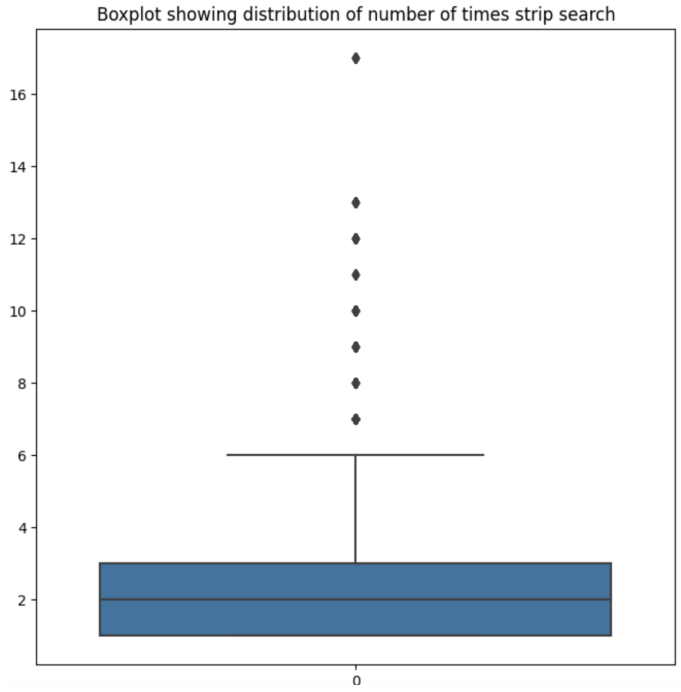
### Frequency table for Perceived_Race:

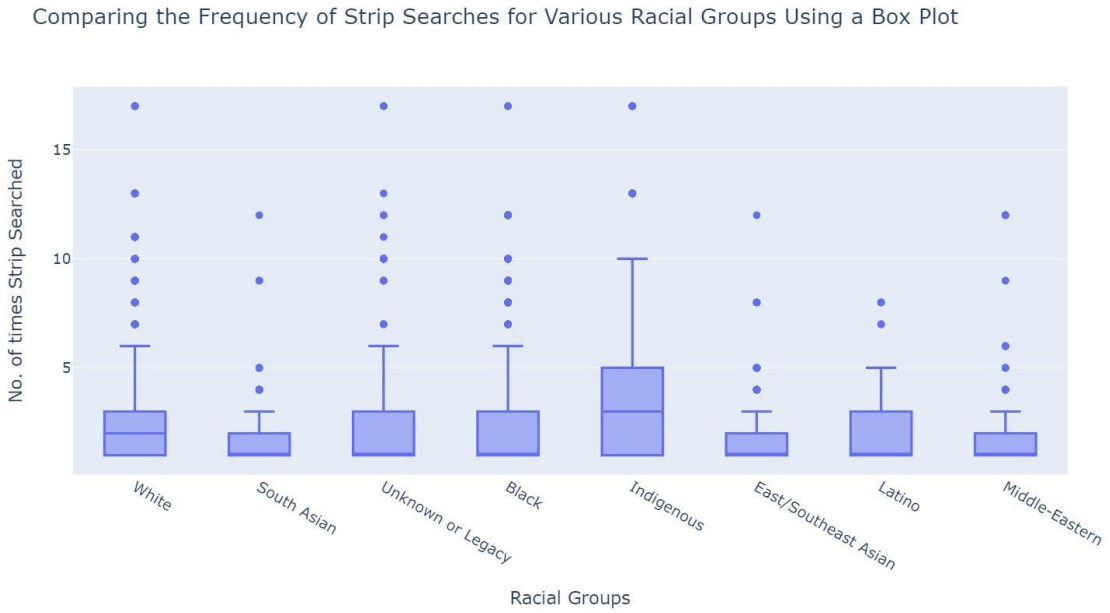
```
White                    27723
Black                    17526
Unknown or Legacy         5056
East/Southeast Asian      4415
South Asian               3613
Middle-Eastern            3237
Indigenous                1934
Latino                    1768
Name: Perceived_Race, dtype: int64
```

### Scatter plot



## Boxplots


Boxplot showing distribution of number of times strip search
16
14
12
10
8
6
4
2
0


Comparing the Frequency of Strip Searches for Various Racial Groups Using a Box Plot
No. of times Strip Searched
15
10
5
White
South Asian
Unknown or Legacy
Black
Indigenous
East/Southeast Asian
Latino
Middle-Eastern
Racial Groups

# Post-hoc: Pairwise Comparison Test Results

**Post-hoc: Pairwise Comparison Test**

[30]:
```
pair_t = model_1.t_test_pairwise('C(Perceived_Race)')
pair_t.result_frame
```

Out[30]:

| | coef | std err | t | P>\|t\| | Conf. Int. Low | Conf. Int. Upp. | pvalue-hs | reject-hs |
|---|---|---|---|---|---|---|---|---|
| White-Black | 0.155452 | 0.095949 | 1.620149 | 1.052411e-01 | -0.032634 | 0.343539 | 6.711026e-01 | False |
| Latino-Black | -0.048241 | 0.339415 | -0.142130 | 8.869809e-01 | -0.713587 | 0.617104 | 9.985564e-01 | False |
| South Asian-Black | -0.776917 | 0.255070 | -3.045891 | 2.327727e-03 | -1.276924 | -0.276910 | 4.331238e-02 | True |
| Middle-Eastern-Black | -0.616423 | 0.302788 | -2.035824 | 4.180206e-02 | -1.209969 | -0.022877 | 4.259919e-01 | False |
| Unknown or Legacy-Black | -0.011661 | 0.168619 | -0.069156 | 9.448670e-01 | -0.342199 | 0.318877 | 9.985564e-01 | False |
| East/Southeast Asian-Black | -1.094201 | 0.243101 | -4.501012 | 6.861980e-06 | -1.570745 | -0.617657 | 1.715354e-04 | True |
| Indigenous-Black | 0.780271 | 0.212953 | 3.664045 | 2.499221e-04 | 0.362825 | 1.197718 | 5.235267e-03 | True |
| Latino-White | -0.203693 | 0.336962 | -0.604500 | 5.455291e-01 | -0.864229 | 0.456843 | 9.806121e-01 | False |
| South Asian-White | -0.932369 | 0.251796 | -3.702871 | 2.146561e-04 | -1.425958 | -0.438780 | 4.711806e-03 | True |
| Middle-Eastern-White | -0.771875 | 0.300035 | -2.572617 | 1.011176e-02 | -1.360025 | -0.183725 | 1.442928e-01 | False |
| Unknown or Legacy-White | -0.167113 | 0.163624 | -1.021326 | 3.071320e-01 | -0.487860 | 0.153634 | 9.233427e-01 | False |
| East/Southeast Asian-White | -1.249653 | 0.239663 | -5.214200 | 1.893896e-07 | -1.719458 | -0.779848 | 5.113508e-06 | True |
| Indigenous-White | 0.624819 | 0.209021 | 2.989271 | 2.805225e-03 | 0.215082 | 1.034556 | 4.930788e-02 | True |
| South Asian-Latino | -0.728676 | 0.411580 | -1.770437 | 7.669375e-02 | -1.535483 | 0.078132 | 5.842775e-01 | False |
| Middle-Eastern-Latino | -0.568182 | 0.442739 | -1.283335 | 1.994132e-01 | -1.436069 | 0.299705 | 8.648936e-01 | False |
| Unknown or Legacy-Latino | 0.036580 | 0.364375 | 0.100391 | 9.200363e-01 | -0.677693 | 0.750854 | 9.985564e-01 | False |
| East/Southeast Asian-Latino | -1.045960 | 0.404271 | -2.587274 | 9.691918e-03 | -1.838440 | -0.253479 | 1.442928e-01 | False |
| Indigenous-Latino | 0.828512 | 0.386893 | 2.141452 | 3.226871e-02 | 0.070098 | 1.586927 | 3.682184e-01 | False |
| Middle-Eastern-South Asian | 0.160494 | 0.381937 | 0.420211 | 6.743433e-01 | -0.588205 | 0.909193 | 9.887529e-01 | False |
| Unknown or Legacy-South Asian | 0.765256 | 0.287450 | 2.662224 | 7.778665e-03 | 0.201776 | 1.328735 | 1.243198e-01 | False |
| East/Southeast Asian-South Asian | -0.317284 | 0.336591 | -0.942641 | 3.458942e-01 | -0.977093 | 0.342525 | 9.233427e-01 | False |
| Indigenous-South Asian | 1.557188 | 0.315506 | 4.935520 | 8.160718e-07 | 0.938710 | 2.175666 | 2.121765e-05 | True |
| Unknown or Legacy-Middle-Eastern | 0.604762 | 0.330525 | 1.829699 | 6.733342e-02 | -0.043157 | 1.252681 | 5.667716e-01 | False |
| East/Southeast Asian-Middle-Eastern | -0.477778 | 0.374049 | -1.277312 | 2.015303e-01 | -1.211016 | 0.255460 | 8.648936e-01 | False |
| Indigenous-Middle-Eastern | 1.396694 | 0.355196 | 3.932183 | 8.491242e-05 | 0.700415 | 2.092974 | 2.035909e-03 | True |
| East/Southeast Asian-Unknown or Legacy | -1.082540 | 0.276884 | -3.909727 | 9.318944e-05 | -1.625307 | -0.539773 | 2.141161e-03 | True |
| Indigenous-Unknown or Legacy | 0.791932 | 0.250830 | 3.157248 | 1.598772e-03 | 0.300238 | 1.283627 | 3.149442e-02 | True |
| Indigenous-East/Southeast Asian | 1.874472 | 0.305911 | 6.127510 | 9.363369e-10 | 1.274804 | 2.474140 | 2.621743e-08 | True |

**RQ2: Likelihood of Stop Check based on various search reasons**

6]:

| | coef | std err | t | P>\|t\| | Conf. Int. Low | Conf. Int. Upp. | pvalue-hs | reject-hs |
|---|---|---|---|---|---|---|---|---|
| Aged 35 to 44 years-Aged 25 to 34 years | -0.130165 | 0.125076 | -1.040686 | 0.298054 | -0.375348 | 0.115018 | 0.996526 | False |
| Aged 18 to 24 years-Aged 25 to 34 years | -0.547994 | 0.113773 | -4.816537 | 0.000001 | -0.771020 | -0.324967 | 0.000054 | True |
| Aged 45 to 54 years-Aged 25 to 34 years | 0.268577 | 0.171995 | 1.561540 | 0.118437 | -0.068580 | 0.605734 | 0.929154 | False |
| Aged 55 to 64 years-Aged 25 to 34 years | -0.944781 | 0.278425 | -3.393305 | 0.000694 | -1.490570 | -0.398993 | 0.021292 | True |
| Aged 65 and older-Aged 25 to 34 years | -1.505312 | 0.736264 | -2.044527 | 0.040935 | -2.948589 | -0.062035 | 0.648280 | False |
| Aged 65 years and older-Aged 25 to 34 years | -0.616423 | 2.198932 | -0.280328 | 0.779233 | -4.926924 | 3.694078 | 0.999963 | False |
| Aged 17 years and younger-Aged 25 to 34 years | -0.732069 | 0.195676 | -3.741228 | 0.000184 | -1.115647 | -0.348491 | 0.005885 | True |
| Aged 17 years and under-Aged 25 to 34 years | -1.616423 | 0.985600 | -1.640040 | 0.101037 | -3.548465 | 0.315619 | 0.903990 | False |
| Aged 18 to 24 years-Aged 35 to 44 years | -0.417829 | 0.133126 | -3.138590 | 0.001704 | -0.678792 | -0.156865 | 0.049877 | True |
| Aged 45 to 54 years-Aged 35 to 44 years | 0.398742 | 0.185366 | 2.151109 | 0.031498 | 0.035375 | 0.762109 | 0.564881 | False |
| Aged 55 to 64 years-Aged 35 to 44 years | -0.814616 | 0.286878 | -2.839595 | 0.004529 | -1.376974 | -0.252259 | 0.119351 | False |
| Aged 65 and older-Aged 35 to 44 years | -1.375147 | 0.739502 | -1.859558 | 0.062986 | -2.824771 | 0.074477 | 0.776045 | False |
| Aged 65 years and older-Aged 35 to 44 years | -0.486258 | 2.200018 | -0.221025 | 0.825079 | -4.798889 | 3.826373 | 0.999963 | False |
| Aged 17 years and younger-Aged 35 to 44 years | -0.601904 | 0.207527 | -2.900368 | 0.003738 | -1.008713 | -0.195095 | 0.102909 | False |
| Aged 17 years and under-Aged 35 to 44 years | -1.486258 | 0.988021 | -1.504278 | 0.132551 | -3.423046 | 0.450530 | 0.941805 | False |
| Aged 45 to 54 years-Aged 18 to 24 years | 0.816571 | 0.177935 | 4.589157 | 0.000005 | 0.467770 | 1.165371 | 0.000158 | True |
| Aged 55 to 64 years-Aged 18 to 24 years | -0.396787 | 0.282133 | -1.406384 | 0.159650 | -0.949845 | 0.156270 | 0.963294 | False |
| Aged 65 and older-Aged 18 to 24 years | -0.957318 | 0.737674 | -1.297752 | 0.194411 | -2.403359 | 0.488723 | 0.979581 | False |
| Aged 65 years and older-Aged 18 to 24 years | -0.068429 | 2.199404 | -0.031113 | 0.975181 | -4.379857 | 4.242998 | 0.999963 | False |
| Aged 17 years and younger-Aged 18 to 24 years | -0.184075 | 0.200917 | -0.916176 | 0.359603 | -0.577927 | 0.209776 | 0.998049 | False |
| Aged 17 years and under-Aged 18 to 24 years | -1.068429 | 0.986654 | -1.082882 | 0.278895 | -3.002537 | 0.865679 | 0.996145 | False |
| Aged 55 to 64 years-Aged 45 to 54 years | -1.213358 | 0.310221 | -3.911273 | 0.000093 | -1.821475 | -0.605242 | 0.003051 | True |
| Aged 65 and older-Aged 45 to 54 years | -1.773889 | 0.748867 | -2.368764 | 0.017872 | -3.241870 | -0.305908 | 0.385477 | False |
| Aged 65 years and older-Aged 45 to 54 years | -0.885000 | 2.203184 | -0.401691 | 0.687922 | -5.203836 | 3.433836 | 0.999963 | False |
| Aged 17 years and younger-Aged 45 to 54 years | -1.000646 | 0.238758 | -4.191040 | 0.000028 | -1.468677 | -0.532615 | 0.000954 | True |
| Aged 17 years and under-Aged 45 to 54 years | -1.885000 | 0.995049 | -1.894379 | 0.058212 | -3.835566 | 0.065566 | 0.762931 | False |
| Aged 65 and older-Aged 55 to 64 years | -0.560531 | 0.780218 | -0.718429 | 0.472515 | -2.089968 | 0.968907 | 0.999536 | False |
| Aged 65 years and older-Aged 55 to 64 years | 0.328358 | 2.214036 | 0.148308 | 0.882104 | -4.011752 | 4.668468 | 0.999963 | False |
| Aged 17 years and younger-Aged 55 to 64 years | 0.212712 | 0.323950 | 0.656619 | 0.511445 | -0.422318 | 0.847742 | 0.999607 | False |
| Aged 17 years and under-Aged 55 to 64 years | -0.671642 | 1.018853 | -0.659214 | 0.509778 | -2.668869 | 1.325586 | 0.999607 | False |
| Aged 65 years and older-Aged 65 and older | 0.888889 | 2.316575 | 0.383708 | 0.701205 | -3.652225 | 5.430003 | 0.999963 | False |
| Aged 17 years and younger-Aged 65 and older | 0.773243 | 0.754658 | 1.024627 | 0.305571 | -0.706091 | 2.252576 | 0.996526 | False |
| Aged 17 years and under-Aged 65 and older | -0.111111 | 1.225816 | -0.090643 | 0.927779 | -2.514043 | 2.291821 | 0.999963 | False |
| Aged 17 years and younger-Aged 65 years and older | -0.115646 | 2.205159 | -0.052444 | 0.958177 | -4.438354 | 4.207061 | 0.999963 | False |
| Aged 17 years and under-Aged 65 years and older | -1.000000 | 2.407456 | -0.415376 | 0.677878 | -5.719264 | 3.719264 | 0.999963 | False |
| Aged 17 years and under-Aged 17 years and younger | -0.884354 | 0.999415 | -0.884872 | 0.376253 | -2.843477 | 1.074770 | 0.998049 | False |