INF2178H Experimental Design for Data Science

Final Project

Study on Factors Affecting the Likelihood of Being Strip-Searched

By

Ziqing Ni, Yingpeng Zhang

Instructor: Dr. Shion Guha

Faculty of Information

University of Toronto

Table of Contents

1. Introduction

A strip search is a thorough search of an individual's body, including their clothing and personal belongings, to look for any items or contraband that may be concealed. This type of search is typically conducted by law enforcement officers or security personnel, such as policemen, when they believe some people may have prohibited items, such as drugs or weapons. However, Strip searches can be highly intrusive and cause significant emotional distress to the person being searched.

Strip search policies vary widely in different countries. Some countries, such as the United States, have specific guidelines on when and how strip searches should be conducted. But in other countries, strip searches are more common and invasive. Human rights will be used against strip searches. The European Convention on Human Rights prohibits inhumane and degrading strip searches (Mavronicola & Webster, 2023). Section 51 of the UK Immigration Act states that a person in immigration detention or prison may be strip-searched for Citizenship documents (Spalding, 2021).

Strip searches generally select law enforcement officers of the same gender. However, how to define transgender people creates additional complexity in the execution of strip searches. Due to the failure of the binary gender system, especially in Canada, the policing of the body has become an acute issue, not only for the police but also for society as a whole (Kirkup, 2009). The Supreme Court of Canada called strip searches *"one of the most extreme exercises of police power"* and *"inherently humiliating and degrading"*. In Ontario, there are no legislated safeguards restricting these highly invasive searches. However, there may be legislated safeguards in federal prisons and many other provinces (*Strip Searches in Ontario Prisons – CCLA, 2022*).

Through the analysis of the Toronto Police Service's release of race-based data on strip searches as part of its Race and Identity-based Data Collection Strategy, Monika points out that although black people account for 10% of the population in the Toronto area, about 1 /3 of the people who were strip-searched are black people. And about 1/3 of the Aboriginal people arrested accepted strip searches (Lemke, n.d.). The effectiveness of strip-searched has been met with skepticism,

and possible racialized issues deserve attention. Thus, in our study, we will do some empirical analysis on the strip search in Toronto and find out if the strip search is abused in the policy system.

2. Literature Review

In strip searches, human rights, race, gender, and other aspects are examined. The approach requires a person to disrobe or undergo a cavity search, which can violate their privacy and dignity. Increasingly, it is acknowledged that strip searches, especially when they are discriminatory or abusive, are damaging. Hence, legal and policy frameworks have been established to guarantee that strip searches respect human rights and promote dignity.

Smith (2022) conducted a literature review on cognitive-behavioral therapy for anxiety disorders. This study explored how this treatment alleviates symptoms and enhances the patient's quality of life. Smith has shown that cognitive-behavioral therapy effectively treats anxiety disorders like generalized anxiety disorder, panic disorder, and social anxiety disorder. Mavronicola and Webster (2023) investigate the relationship between the European Court of Human Rights' strip search case law and Article 3's bar on cruel and humiliating treatment. The authors investigate the Supreme Court's context-sensitive approach to determining allowed and prohibited conduct, as well as the variables that render strip searches either harsh and demeaning or compatible with human dignity. This demonstrates how the Court balances the necessity and propriety of prison action. Gilmartin and Scull (2022) analyze the responses of Republican women and the Catholic Church to Northern Ireland's strip searches in the 1980s and 1990s. The authors point out that the women viewed the practice as a gendered weapon used to demoralize and humiliate them and the nationalist group. In addition, they examine how, despite condemning republican violence, the Catholic Church came to view strip-searching as a violation of fundamental human rights. The authors demonstrate how some Catholic clergymen blended pastoral care for their parishioners with political statements against the British government, raising questions about the role of the Church in the Northern Ireland conflict.

Several provinces in Canada have varied legislation regarding strip searches. Gorman (2022) discusses whether or not strip searches are permissible in Canada based on the court review. It

concentrates on the case of Golden, in which the illegality of the search was ruled, and other examples such as Ali and Tim. The Golden case established guidelines for strip searches, including the need for a valid arrest and a reasonable suspicion that evidence could be concealed on the body of the arrested individual. Since Golden, this article explores how Canadian courts have interpreted and implemented these standards. It also raises fears that Golden police may not adhere to the regulations when conducting strip searches. The author summarizes the Key principles, including the necessity for a valid reason and the requirement that the search does not infringe on section 8 of the Canadian Charter of Rights and Freedoms.

Ottawa police sergeant Steven Desjourdy was the first officer in Canada to be convicted of sexual assault on the basis of an illegal strip search of a woman, which might be deemed "sexual assault by the state." Sexual assault cases provide numerous barriers for all complainants, but when the accused is a police officer conducting his duties, these obstacles are almost insurmountable. Given that Desjourdy was white and SB was a violent and dangerous Black Canadian woman, the likelihood of racism loomed large (Gorman, 2022).

It can be observed that there is a correlation between gender, race, and strip searches, and that strip searches conducted on women and people of non-white races frequently have negative social consequences. Due to Canada's observable seasonal variations, the crime rate similarly varies with the seasons. Thus, strip searches may also exhibit seasonal variations. According to Linning's research, the number of crimes fluctuates with the seasons, with the majority occurring in the summer. Yet, the analysis demonstrates that seasonal shifts in crime patterns are not uniform across all forms of crime and locations (Linning, 2015). His findings could impact crime prevention and police operations. If politicians are aware of how climate influences crime trends, they can spend resources more effectively and devise more targeted strategies to combat crime.

There might also be a connection between the age of the person being searched and the strip search. According to Grewcock and Sentas (2019), nearly half (45%) of all documented strip searches conducted in the 2017/18 fiscal year were conducted on young people aged 25 years or younger, and strip searches took place in a variety of settings. In New South Wales, there has been an uptick in recent years. Many studies have also been done on strip searches of children

(Donahoe, 2010; Liptak & Souter, 2009). According to the research discussed before, there may also be a connection between one's age and the likelihood of being subjected to a strip search.

3. Research Objects and Questions

A strip search checks an individual's body, clothing, and belongings for hidden objects or contraband. When authorities suspect someone has drugs or weapons, they undertake this type of search. Strip searches can be unpleasant and upsetting. Some countries have specific rules on when and how to strip search. Strip searches are against the EU Human Rights Convention.

Strip searches evaluate gender, race, and other aspects. Police and society are concerned about body policing. The Supreme Court of Canada deemed strip searches "excessive police power" and "humiliating and degrading." Empirical research demonstrates that coloured people are disproportionately strip searched. Strip searches are controversial and may have racial implications. Consequently, this study examines Toronto strip searches to evaluate if the policy system abuses them.

From the previous literature review and introduction, we decided to study which factors will affect the likelihood of being strip-searched from the dataset provided by the Toronto Police Service. The research questions are listed as the following:

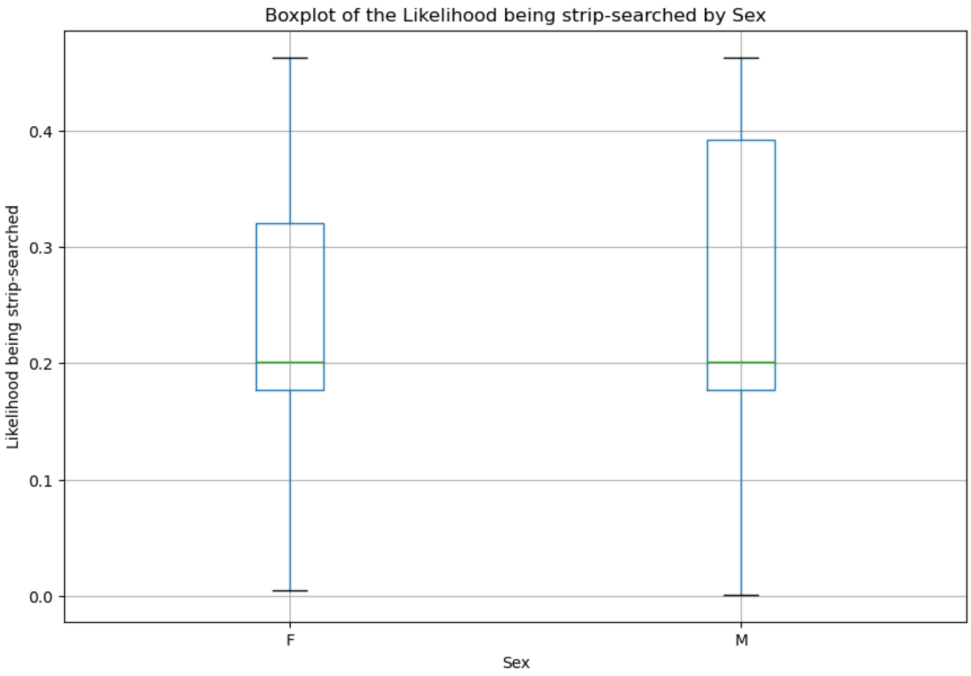
- RQ1: How does sex affect the likelihood of being strip-searched in occurrence categories?
- RQ2: How does perceived race (white vs. non-white) affect the likelihood of being strip-searched in occurrence categories?
- RQ3: How does age_group affect the likelihood of being strip-searched in occurrence categories?
- RQ4: Is there a significant difference in the likelihood of being strip-searched in each category between different gender groups after controlling for the estimation of age at arrested?
- RQ5: Can we predict the likelihood of being strip-searched based on sex_coded, race_coded, and Age_Estimated?

4. Exploratory Data Analysis

Table 1 shows the descriptive results for the likelihood of being strip-searched in occurrence categories by sex.

Table1. Likelihood of Being Strip-Searched in Occurrence Categories by Sex		
	Male	Female
Mean	0.252466	0.245305
Min	0.001059	0.005140
Max	0.462882	0.462882
Median	0.201596	0.201596
Std	0.119566	0.112956
IQR	0.21442	0.142761

Figure 1 shows the boxplot for the likelihood of being strip-searched in occurrence categories by sex.



From the summary table and the boxplot, we can see that group of females has a median of 0.201596, an IQR of 0.142761 (ranging from 0.177074 to 0.319835), and no outliers. The group

of males has a median of 0.201596, an IQR of 0.21442 (ranging from 0.177074 to 0.391494), and no outliers. We can infer that the likelihood of being strip-searched in occurrence categories for males is more spread out than for females. In other words, the likelihood of males being strip-searched in occurrence categories is more diverse and variable.

Before performing the t-test, we checked the following assumptions:

- Normality: The data is normally distributed.
- Homogeneity of variance: The variances of the two groups Males and Females are equal.
- Independence: The observations are independent of each other

To discover how sex affects the likelihood of being strip-searched in occurrence categories, we conduct a t-test; the null hypothesis is that there is no significant difference between the means of the likelihood of the two groups (males/females) being strip-searched. The alternative hypothesis is that the means of the likelihood of the two groups (males/females) being strip-searched are different. The t-test result is the following:

Table 2. The t-test Result for the Likelihood of Being Strip-Searched in Males/Females two groups	
t-statistic	2.0554354098057157
p-value	0.0399739624433901

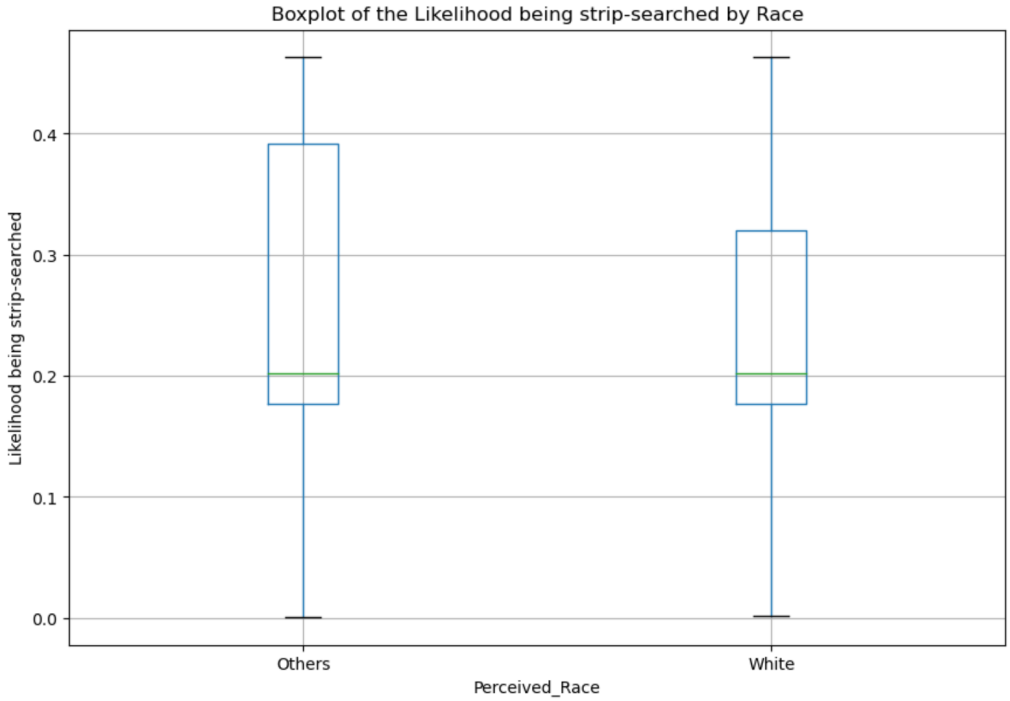
The p-value is about 0.04, which is less than the significance level of 0.05, we reject the null hypothesis and conclude that there is a significant difference between the means of the likelihood of the two groups(males/females). In other words, sex affects the likelihood of being strip-searched in occurrence categories.

Next, we reclassify the perceived race by white and others. We will discover how the likelihood of being strip-searched in occurrence categories changes for these two groups.

Table 3 shows the descriptive results for the likelihood of being strip-searched in occurrence categories by perceived race.

Table 3. Likelihood of Being Strip-Searched in Occurrence Categories by Perceived Race		
	White	Others
Mean	0.251179	0.251391
Min	0.002110	0.001059
Max	0.462882	0.462882
Median	0.201596	0.201596
Std	0.121124	0.116320
IQR	0.142761	0.21442

Figure 2 shows the boxplot for the likelihood of being strip-searched in occurrence categories by perceived race.



We can see that group of others has a median of 0.201596, an IQR of 0.21442 (ranging from 0.177074 to 0.391494), and no outliers. The group of white has a median of 0.201596, an IQR of 0.142761 (ranging from 0.177074 to 0.319835), and no outliers. We can infer that the likelihood of being strip-searched in occurrence categories for other races is more spread out than for white.

In other words, the likelihood of other races being strip-searched in occurrence categories is more diverse and variable.

Before performing the t-test, we checked the following assumptions:

- Normality: The data is normally distributed.
- Homogeneity of variance: The variances of the two groups White/Other races are equal.
- Independence: The observations are independent of each other

To discover how perceived race affects the likelihood of being strip-searched in occurrence categories, we conduct a t-test; the null hypothesis is that there is no significant difference between the means of the likelihood of the two groups (white/other races) being strip-searched. The alternative hypothesis is that the means of the likelihood of the two groups (white/other races) being strip-searched are different. The t-test result is the following:

Table 4. The t-test Result for the Likelihood of Being Strip-Searched in White/Other races two groups	
t-statistic	-0.07841619141073725
p-value	0.9374990968539462

The p-value is about 0.94, which is far greater than the significance level of 0.05, we fail to reject the null hypothesis and conclude that there is not enough evidence to support a significant difference between the means of the likelihood of being strip-searched of white and other races two groups.

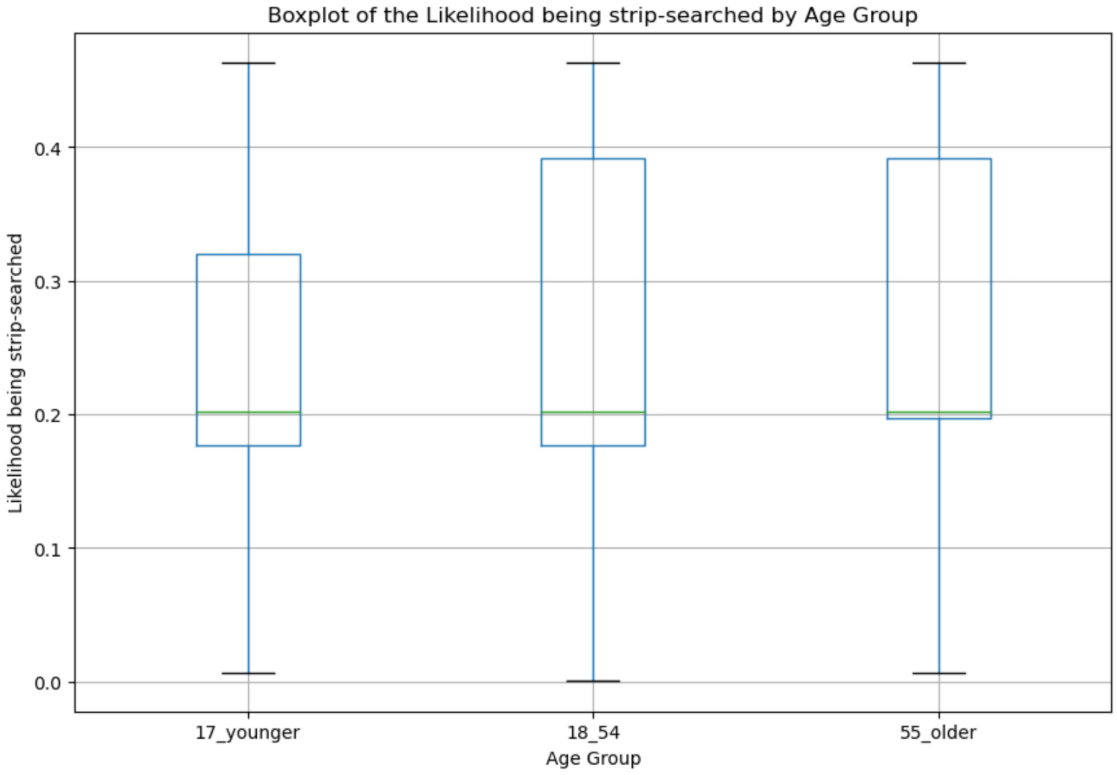
Next, we regrouped the age group into Aged 17 years old or younger, Aged 18 to 54 years old, and Aged 55 or older to conduct the research question effectively.

Table 5 shows the descriptive results for the likelihood of being strip-searched in occurrence categories by age group.

Table 5. Likelihood of Being Strip-Searched in Occurrence Categories by Age_Group			
	17_younger	Aged 18 to 54	55_older

Mean	0.246568	0.251112	0.257880
Min	0.006686	0.001059	0.006510
Max	0.462882	0.462882	0.462882
Median	0.201596	0.201596	0.201596
Std	0.111083	0.119130	0.112780
IQR	0.142761	0.21442	0.194258

Figure 3 shows the boxplot for the likelihood of being strip-searched in occurrence categories by age group.



We can see that group Aged 17 years old or younger has a median of 0.201596, an IQR of 0.142761 (ranging from 0.177074 to 0.319835), and no outliers. The Aged 18 to 54 years old group has a median of 0.201596, an IQR of 0.21442 (ranging from 0.177074 to 0.391494), and no outliers. The Aged 55 or older group has a median of 0.201596, an IQR of 0.194258 (ranging from 0.197236 to 0.391494), and no outliers. We can infer that the likelihood of being strip-searched in occurrence categories for Aged 18 to 54 years old is more spread out than for

Aged 55 or older, than for Aged 17 years old or younger. In other words, the likelihood of Aged 18 to 54 years old being strip-searched in occurrence categories is the most diverse and variable among the three groups.

Before performing the t-test, we checked the following assumptions:

- Normality: The data is normally distributed.
- Homogeneity of variance: The variances of the two age groups Aged 17 years old or younger/Aged 18 to 54 years old, Aged 17 years old or younger/Aged 55 or older, Aged 18 to 54 years old/Aged 55 or older, are equal.
- Independence: The observations are independent of each other

To discover how different age group affects the likelihood of being strip-searched in occurrence categories, we conduct a series of t-test; firstly, the null hypothesis is that there is no significant difference between the means of the likelihood of the two groups (Aged 17 years old or younger/Aged 18 to 54 years old) being strip-searched. The alternative hypothesis is that the means of the likelihood of the two groups (Aged 17 years old or younger/Aged 18 to 54 years old) being strip-searched are different. The t-test result is the following:

Table 6. The t-test Result for the Likelihood of Being Strip-Searched in Aged 17 years old or younger/Aged 18 to 54 years old two groups	
t-statistic	-0.6695032561876009
p-value	0.5036813652308522

The p-value is about 0.5, which is far greater than the significance level 0.05, we fail to reject the null hypothesis and conclude that there is not enough evidence to support a significant difference between the means of the likelihood of being strip-searched of Aged 17 years old or younger/Aged 18 to 54 years old two groups.

Secondly, the null hypothesis is that there is no significant difference between the means of the likelihood of the two groups (Aged 17 years old or younger/Aged 55 or older) being strip-searched. The alternative hypothesis is that the means of the likelihood of the two groups

(Aged 17 years old or younger/Aged 55 or older) being strip-searched are different. The t-test result is the following:

Table 7. The t-test Result for the Likelihood of Being Strip-Searched in Aged 17 years old or younger/Aged 55 or older two groups	
t-statistic	-1.2972882283935105
p-value	0.19502544002537764

The p-value is about 0.2, which is greater than the significance level 0.05, we fail to reject the null hypothesis and conclude that there is not enough evidence to support a significant difference between the means of the likelihood of being strip-searched of Aged 17 years old or younger/Aged 55 or older two groups.

Thirdly, the null hypothesis is that there is no significant difference between the means of the likelihood of the two groups (Aged 18 to 54 years old /Aged 55 or older) being strip-searched. The alternative hypothesis is that the means of the likelihood of the two groups (Aged 18 to 54 years old /Aged 55 or older) being strip-searched are different. The t-test result is the following:

Table 8. The t-test Result for the Likelihood of Being Strip-Searched in Aged 18 to 54 years old/Aged 55 or older two groups	
t-statistic	-1.1614808094494713
p-value	0.2460649246888497

The p-value is about 0.25, which is greater than the significance level 0.05, we fail to reject the null hypothesis and conclude that there is not enough evidence to support a significant difference between the means of the likelihood of being strip-searched of Aged 18 to 54 years old/Aged 55 or older two groups.

5. Methods

1. Power analysis

Compute the power graph of the particular equation and discuss the relevant components, such as effect size, sample size - gender group

$$power = 1 - \beta = P(Z \geq z_{\alpha/2} - \theta/\sqrt{1/n_1 + 1/n_2}) + P(Z < -z_{\alpha/2} - \theta/\sqrt{1/n_1 + 1/n_2})$$

Where $\theta = (\mu_1 - \mu_2)/\sigma$ is the effect size. n_1 is the sample size in group 1, and n_2 is the sample size in group 2. When $n_1 = n_2$, the power can get its maximum when n = $n_1 + n_2$ is fixed.

2. Analysis of covariance ANCOVA

ANCOVA is used to test the effects of categorical variables on a continuous variable, controlling for the other continuous variables, and the control variables are called the “covariates”.

ANCOVA can use multiple regression methods to estimate the effects. In multiple regression, the categorical variables are defined as dummies. k-1 dummies are needed for a categorical variable with k levels.

The ANOVA model is in the form:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

And the regression model is:

$$y_i = \mu + \beta(x_i - \bar{x}) + \epsilon_i$$

ANCOVA is the combination of ANOVA and regression, so its linear model is

$$y_{ij} = \mu + \beta(x_{ij} - \bar{x}..) + \epsilon_{ij}$$

The assumptions for the ANCOVA are:

1. The covariate X are fixed without error, and it is independent to the treatment.

2. The regression of Y on X, is linear and independent of the treatment.
3. The residuals are normally and independent with zero mean and common variance.

3. Logistic regression

In real life, we frequently need to differentiate between distinct categories, such as determining whether an email is spam, whether the police will conduct a body search if they encounter a prisoner, and so on. The logistic regression method is frequently used for yes/no estimation.

We assume that the success probability of an event is P and define odds as P/(1-P). In logistic regression, we estimate the log odd value first, then calculate the probability P of success under a given X. The logistic regression model is:

$$log(P/(1 - P) = X\beta$$

Where X are the factors that are related to the probability of success. The estimated β_i means that one more unit of the X_i, the average change of the log odd ratio on the success. After we get the estimated β, the probability to be successful will be

$$\hat{P} = \frac{e^{X\hat{\beta}}}{1+e^{X\hat{\beta}}}$$

The basic assumption for the logistic regression is: independent of errors, linearity in the logit for continuous variable X, no multicollinearity and lack of strong influential points (Stoltzfus, 2011).

6. Results

6.1 Power Analysis

The effect size for the likelihood of being strip-searched in each occurrence category is as follows:

Table 9. Effect Size Calculation	
Effect size (Cohen's D)	0.02176077818674804

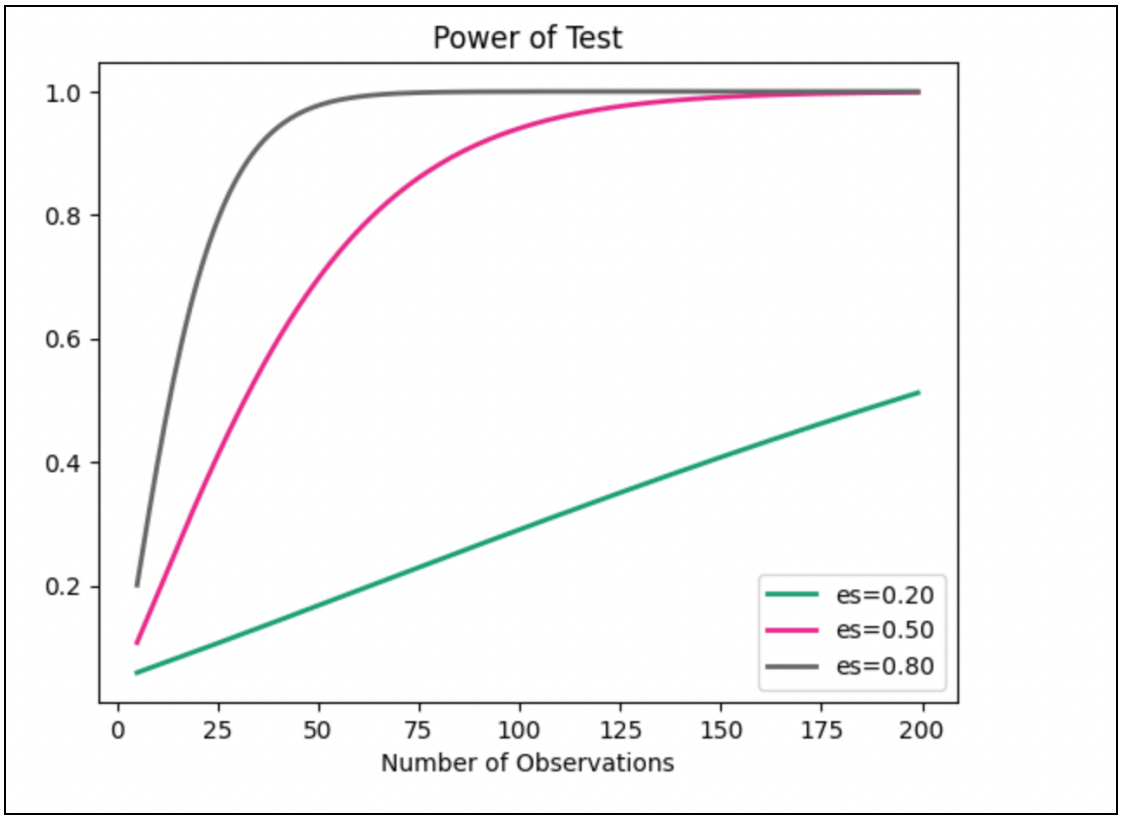
We chose to split our dataset into two groups in terms of gender, male and female. After creating two different datasets containing the likelihood of being strip-searched in each occurrence category, we conducted the power analysis. Prior to computing a t-test to analyze whether the likelihood of being strip-searched in occurrence categories(outcome variable) differed between males and females (two-level explanatory variable), we calculated the effect size of the explanatory variable using Cohen's D metric, which was 0.02, which is considered a small effect according to Cohen's guidelines. It means that the difference between males' likelihood of being strip-searched in occurrence categories and females' likelihood of being strip-searched in occurrence categories is small and may not have much significance.

The required sample size for each group is calculated and shown as follows:

Table 10. Sample Size		
	Required Sample Size	Actual Sample Size
Male	85751.403	52518
Female	20547.158	12584

After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. The results indicated that a sample size of 85751 was required for the male group, while a sample size of 20547 was required for the female group. This is significant because the sample sizes provided in the dataset are 52518 and 12584 respectively, which impacts the reliability of the results. In addition, we tested the power of our proposed sample size, which is 1.

Figure 4 depicts the power curve.



We plotted the power curve, which shows the statistical power of a hypothesis test as a function of sample size or effect size, which can be interpreted differently depending on the effect size used. When using effect sizes of 0.2, 0.5, and 0.8, the black curve reaching high power quickly indicates high power and a small sample size requirement for the larger effect sizes. The green curve reaching high power only at larger sample sizes or effect sizes indicates low power and a larger sample size or effect size requirement for the smaller effect sizes. The position of the curve relative to the alpha level can indicate the likelihood of producing false positives or false negatives.

6.2 One-way ANCOVA

We used one-way ANCOVA to conduct the research, is there a significant difference in the likelihood of being strip-searched in each category between different gender group after controlling for the estimation of age at arrested?

Before conducting one-way ANCOVA, we checked some key assumptions of ANCOVA. We changed the variable "Age_group__at_arrest_" to the estimated age in terms of each age group, in this way, we altered this categorical variable to a continuous one.

- Normality: The dependent variable "Occurrence_Category_Strip_Search_Likelihood" is normally distributed within each group for each level of the covariate "Age_Estimated"
- Homogeneity of variance: The variance of the dependent variable "Occurrence_Category_Strip_Search_Likelihood" equals across all groups, for each level of the covariate "Age_Estimated".
- Linearity: There is a linear relationship between the dependent variable "Occurrence_Category_Strip_Search_Likelihood" and the covariate "Age_Estimated".
- Independence: The observations are independent of each other.
- Homogeneity of regression slopes: The relationship between the covariate "Age_Estimated" and the dependent variable "Occurrence_Category_Strip_Search_Likelihood" is the same for all sex groups.

Table 11. Check Data Type for Running One-way ANCOVA		
Sex	Predictor	Categorical Variable
Age_estimated	Covariate	Continuous Variable
likelihood of being strip-searched in each occurrence category	Dependent Variable	Continuous Variable

Table 12. One-way ANCOVA Result						
	Source	SS	DF	F	p-unc	np2
0	Sex	0.130534	2	4.153933	1.570670e-02	0.000128
1	Age_Estimated	2.755201	1	175.354911	5.642759e-40	0.002687
2	Residual	1022.593413	65083			

From the ANCOVA result, the p-value for sex is less than 0.05, we can reject the null hypothesis that each gender results in the same likelihood of being strip-searched in each occurrence category, after controlling for the estimated age at arrest.

6.3 Logistic Regression

To conduct the logistic regression, we encoded the variables “Sex” and “Perceived_Race” and treated them as dummy variables. Since we are interested in the difference in the likelihood of being strip-searched between white people and other races, we encoded the other races to the same number and split the variables “Perceived_Race” into two groups. We checked the following assumptions of the logistic regression model:

- Linearity: The relationship between the independent variables “sex_coded”, “race_coded”, “Age_Estimated” and the log odds of the dependent variable “StripSearch” is linear.
- Independence of errors: The errors or residuals in the logistic regression model is independent of each other.
- No multicollinearity: There should be no high correlation among the independent variables.
- Binary dependent variable: The dependent variable “StripSearch” is binary.

The formula of our logistic regression is “StripSearch ~ sex_coded + race_coded + Age_Estimated”. And the logistic regression result shows as follows:

Table 13. Logistic Regression Result

	coef	Std err	z	P>|z|	[0.025	0.975]
Intercept	-1.2742	0.059	-21.632	0.000	-1.390	-1.159
sex_coded	-0.2734	0.036	-7.494	0.000	-0.345	-0.202
race_coded	0.2246	0.028	8.091	0.000	0.170	0.279
Age_Estimated	-0.0161	0.001	-13.086	0.000	-0.019	-0.014
Pseudo R-squ: 0.006702						

Table 14. Odds Ratio	
	Odds Ratio
Intercept	0.279664
sex_coded	0.760760
race_coded	1.251841
Age_Estimated	0.983988

We perform a logistic regression to examine the effects of sex, race, and age, on the likelihood of being strip-searched. Since the p-values of these three variables are less than 0.05, which indicates that the corresponding coefficient is statistically significant and we reject the null hypothesis, meaning that there is evidence to suggest that the variables are associated with the outcome variable. Since the odds ratio of sex_coded is 0.76, which is less than 1, if sex_coded increases, the probability of being strip-searched occurring decreases. Since the odds ratio of race_coded is 1.25, which is greater than 1, if race_coded increases, the probability of being strip-searched occurring increases. Since the odds ratio of Age_Estimated is 0.983988, which is less than 1, if Age_Estimated increases, the probability of being strip-searched occurring decreases.

Table 15. Confidence Interval		
	Lower CI	Upper CI
Intercept	0.249172	0.313887
sex_coded	0.708251	0.817161
race_coded	1.185546	1.321842
Age_Estimated	0.981612	0.986370

With 95% confidence, we can say that the true population coefficient for sex_coded lies between 0.708251 and 0.817161. With 95% confidence, we can say that the true population coefficient for race_coded lies between 1.185546 and 1.321842. With 95% confidence, we can say that the true population coefficient for Age_Estimated lies between 0.981612 and 0.986370.

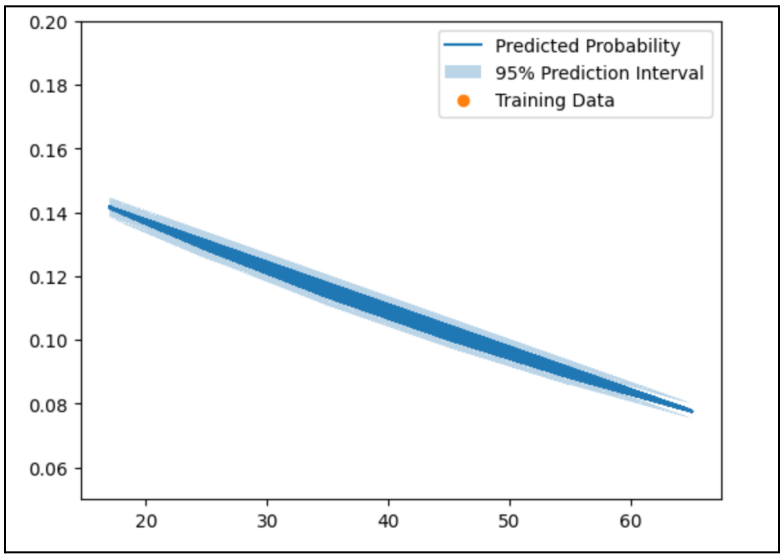
Table 16. Test Accuracy	
Test Accuracy	0.8783129753399401

Since the test accuracy is 0.878, which is higher than 80%, it indicates that the model is able to accurately predict the outcome variable “the likelihood of being strip-searched” for a majority of cases in the test set.

Table 17. Confusion Matrix		
	Predicted Positive	Predicted Negative
True Positive	11433	0
True Negative	1584	0

There were 11433 cases where the model correctly predicted that a criminal suspect had been strip-searched (predicted positive) and a criminal suspect had actually been strip-searched (true positive). There were 1584 cases where the model predicted that a criminal suspect had not been strip-searched (predicted negative), but a criminal suspect had actually been strip-searched (false negative).

Figure 5. Prediction Interval



The 95% prediction interval represents the range of values within which we can expect a new observation to fall with 95% confidence. The predicted probability in the prediction interval is the point estimate of the probability of the new observation falling within this range. Since the prediction interval is narrow, it indicates that our prediction is much more precise.

7. Discussion

A t-test is a hypothesis test that allows us to compare the means of groups, such as male vs. female, white people vs. other races, and age groups to determine whether they are statistically different from each other in terms of the likelihood of being strip-searched in each occurrence category. It provides us with a t-statistic, which is used to calculate a p-value. The p-value tells us the probability of observing a difference as large as, or larger than, the one observed in our data, assuming that the null hypothesis is true. Since the sex factor is significant in terms of the p-value from t-test. In other words, sex affects the likelihood of being strip-searched in occurrence categories. The purpose of using power analysis after getting a significant t-test result is to determine the probability that the test will correctly reject the null hypothesis, given specific effect size, sample size, and significance level. The power is low in our research, which indicates that the sample size may have been too small to detect the effect size of interest, and the results may be less reliable. We should collect more data to validate our findings. From our one-way ANCOVA, the main advantage of ANCOVA over ANOVA is that it can increase the precision of the estimates and reduce error variance by controlling for the covariate “Age_estimated”. It can lead to increased power and sensitivity, allowing us to detect smaller effects than would be possible without controlling for the covariate “Age_estimated”. After controlling for the estimated age at arrest, the sex factor will impact the likelihood of being strip-searched in each occurrence category. We will conduct more ANCOVAs for analyzing differences between groups while controlling for other continuous variables in further case studies. The purpose of running logistic regression is to model the relationship between a binary dependent variable and one or more independent variables. It is a powerful tool for analyzing the relationship between the dependent variable and the independent variables, and can be used to make predictions about the likelihood of being strip-searched, based on the values of the independent variables, “sex_coded”, “race_coded”, “Age_Estimated”. All three variables are significant in terms of the p-values of the logistic regression. From the odds ratio of each variable, we obtain the

information on the magnitude and direction of the association between the independent variable and the dependent variable. The confusion matrix provides a detailed breakdown of the model's predictions, including true positives, true negatives, false positives, and false negatives. These values can be used to calculate various performance metrics that reflect the accuracy, precision, and recall of the model. In our case studies, the sections of the false negative and true negative are 0, which is abnormal. We should alter some predictors in our logistic regression model to check this issue. In general, the likelihood of being strip-searched should be affected by several predictors. We found that sex, age, and race factors may differ in the likelihood of being strip-searched by increasing our sample size, the result helps us to analyze racism and sexism issues in GTA.

8. Conclusion

Based on the strip search data released by the Toronto Police Service during the arrest process, this paper studies the impact of age, gender and race on the probability of strip search under different types of arrests. First of all, in different types of arrest data, the proportion of people of different genders and races who were strip-searched was counted. Then the two-sample t-test method was used to compare whether there were gender and race differences in the strip search. The results showed that there were significant differences in gender, while race The difference is not apparent. 25.25% of men were strip-searched during arrest, while only 24.53% of women were caught. The purpose of the police conducting a strip search is to eliminate possible risks. Women are less dangerous than men, and women may bring allegations of sexual harassment during a strip search. Therefore, the police are more cautious when conducting strip searches for women. In the groups of teenagers, adults and seniors, the average strip search ratio of teenagers is lower than that of adults and seniors. However, the gap between them is small, and there is no noticeable difference.

Doing power analysis on the above two-sample t-test, we found that the power of the hypothesis test results is greater as the sample increases and the effect size increases. Effect size reflects the difference between groups. If the difference is slight, we need to collect more samples in order to distinguish them. The above analysis studies the impact of a single variable on strip search. Using age and gender together to estimate the probability of strip rate, we found that the older

you are, the easier it is to be strip-searched, and men are more likely to be strip-searched than women. These results are consistent with the hypothesis test of a single variable. Finally, we analyze the relationship between gender, age, race and other factors and strip search for individuals. The results show that gender, age, and race will all have an impact on strip searches. The above strong evidence shows significant race, gender and age differences in strip searches. Furthermore, this difference may be considered a manifestation of discrimination and needs to be taken seriously.

Finally, there are some things that could be improved in the analysis of this paper. For example, the R-square in the ANCOVA analysis results is only 0.67%, which has the limited ability to explain strip searches. At the same time, in the hypothesis test and subsequent regression analysis, the normality assumption is not well satisfied, resulting in the risk of invalid hypothesis tests and regression results. For the above limitation, it will be one of the problems that future research will focus on solving.

References

Donahoe, D. R. (2010). Strip Searches of Students: Addressing the Undressing of Children in Schools and Redressing the Fourth Amendment Violations. *Mo. L. REv.*, *75*, 1123.

Gilmartin, N., & Scull, M. M. (2022). Republican women and Catholic Church responses to the strip searching of female prisoners in Northern Ireland, 1982–92. Women's History Review, 31(5), 721-740.

Gorman, W. K. (2022). Strip Searches in Canada and the Constitutional Right to Be Free from Unreasonable Searches and Seizures. Ct. Rev., 58, 84.

Grewcock, M., & Sentas, V. (2019). Rethinking strip searches by NSW Police.

Kirkup, K. (2009). Indocile bodies: gender identity and strip searches in canadian criminal law. *Canadian Journal of Law and Society/La Revue Canadienne Droit et Société*, *24*(1), 107-125.

Lemke, M. (n.d.). Strip searches are ineffective, unnecessary and target racialized Canadians. The Conversation. https://theconversation.com/strip-searches-are-ineffective-unnecessary-and%20-target%20-racialized-canadians-185187

Linning, S. J. (2015). Crime seasonality and the micro-spatial patterns of property crime in Vancouver, BC and Ottawa, ON. Journal of Criminal Justice, 43(6), 544-555.

Liptak, A., & Souter, J. D. H. (2009). Supreme Court says child's rights violated by strip search. *New York Times*.

Marmo, M. (2019). Strip Searching: Seeking the Truth 'in'and 'on'the Regular Migrant's Body. Crimmigration in Australia: Law, Politics, and Society, 197-220

Mavronicola, N., & Webster, E. (2023). Strip searches through the lens of the prohibition of inhuman and degrading treatment in European human rights law. In *Body Searches and Imprisonment* (pp. 67-99). Cham: Springer International Publishing.

Spalding, A. (2021). Strip‐Searching for Nationality Documents. *The Modern Law Review*, *84*(3), 456-476.

Strip Searches in Ontario Prisons - CCLA. (2022, June 20). CCLA. https://ccla.org/major-cases-and-reports/strip-searches-in-ontario-prisons/

Huang, W. (n.d.). ANCOVA. https://www.lehigh.edu/~wh02/ancova.html

Stoltzfus J. C. (2011). Logistic regression: a brief primer. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine, 18(10), 1099–1104. https://doi.org/10.1111/j.1553-2712.2011.01185.x

List of Appendices

1. List of Attributes:

Name	Type
Arrest Year	Number
Arrest Month	Text
EventID	Number
ArrestID	Number
PersonID	Number
Perceived Race	Text
Sex	Text
Age group (at arrest)	Text
Youth at arrest (under 18 years)	Text
ArrestLocDiv	Text
StripSearch	Number
Booked	Number
Occurrence Category	Text
Actions at arrest - Concealed items	Number
Actions at arrest - Combative, violent or spitter/biter	Number
Actions at arrest - Resisted, defensive or escape risk	Number
Actions at arrest - Mental instability or possibly suicidal	Number
Actions at arrest - Assaulted officer	Number
Actions at arrest - Cooperative	Number
SearchReason-CauseInjury	Text
SearchReason-AssistEscape	Text
SearchReason-PossessWeapons	Text

SearchReason-PossessEvidence	Text
ItemsFound	Text

2. One-way ANOVA and Tukey's HSD Result:

```
One-way ANOVA: s = 10.653688182900858, p = 5.499643668690072e-07
The p-value is <0.001 (***)
         Multiple Comparison of Means - Tukey HSD, FWER=0.05
=============================================================================
     group1          group2     meandiff p-adj   lower   upper  reject
-----------------------------------------------------------------------------
 month_Apr_June   month_Jan_Mar  -0.0132 0.0007 -0.0221 -0.0044   True
 month_Apr_June month_July_Sept   0.0034 0.7809 -0.0059  0.0127  False
 month_Apr_June   month_Oct_Dec  -0.0128 0.1486 -0.0283  0.0028  False
  month_Jan_Mar month_July_Sept   0.0166    0.0  0.0083   0.025   True
  month_Jan_Mar   month_Oct_Dec   0.0004 0.9999 -0.0146  0.0154  False
month_July_Sept   month_Oct_Dec  -0.0162 0.0327 -0.0315 -0.0009   True
-----------------------------------------------------------------------------
```

3. Two-way ANOVA and Tukey's HSD Result:

```
                                            sum_sq       df          F  \
C(Age_group__at_arrest_)                  1.305639      2.0  41.488567
C(Perceived_Race)                         0.114509      1.0   7.277386
C(Age_group__at_arrest_):C(Perceived_Race)  0.068630      2.0   2.180817
Residual                                1023.982003  65077.0        NaN

                                                PR(>F)
C(Age_group__at_arrest_)                    9.845139e-19
C(Perceived_Race)                           6.984606e-03
C(Age_group__at_arrest_):C(Perceived_Race)  1.129575e-01
Residual                                          NaN
          Multiple Comparison of Means - Tukey HSD, FWER=0.05
=============================================================================
     group1           group2     meandiff p-adj   lower   upper  reject
-----------------------------------------------------------------------------
17_youngerOthers 17_youngerWhite  -0.0068 0.7364 -0.0209  0.0072  False
17_youngerOthers     18_54Others  -0.0156    0.0 -0.0236 -0.0075   True
17_youngerOthers      18_54White  -0.0127 0.0001 -0.0209 -0.0045   True
17_youngerOthers  55_olderOthers   -0.029    0.0 -0.0395 -0.0185   True
17_youngerOthers   55_olderWhite  -0.0237    0.0 -0.0337 -0.0137   True
 17_youngerWhite     18_54Others  -0.0088 0.2811 -0.0206  0.0031  False
 17_youngerWhite      18_54White  -0.0059 0.7209 -0.0178   0.006  False
 17_youngerWhite  55_olderOthers  -0.0222    0.0 -0.0358 -0.0086   True
 17_youngerWhite   55_olderWhite  -0.0169 0.0036 -0.0301 -0.0037   True
     18_54Others      18_54White   0.0029 0.0793 -0.0002  0.0059  False
     18_54Others  55_olderOthers  -0.0134    0.0 -0.0207 -0.0061   True
     18_54Others   55_olderWhite  -0.0082  0.005 -0.0147 -0.0016   True
      18_54White  55_olderOthers  -0.0163    0.0 -0.0237 -0.0089   True
      18_54White   55_olderWhite   -0.011    0.0 -0.0177 -0.0044   True
  55_olderOthers   55_olderWhite   0.0052 0.6022 -0.0041  0.0146  False
-----------------------------------------------------------------------------
```

4. One-way ANCOVA:

1 to 3 of 3 entries Filter

index	Source	SS	DF	F	p-unc	np2
0	Sex	0.13078852009173464	2	4.161865756931466	0.015582611649653352	0.0001278856010150063
1	Age_Estimated	2.755872010325408	1	175.39107166614437	5.541398957838312e-40	0.002687804893827294
2	Residual	1022.5685541243372	65079	NaN	NaN	NaN

Show 25 per page

5. Logit Regression:

Logit Regression Results

Dep. Variable:	StripSearch	**No. Observations:**	52066
Model:	Logit	**Df Residuals:**	52062
Method:	MLE	**Df Model:**	3
Date:	Sun, 16 Apr 2023	**Pseudo R-squ.:**	0.006702
Time:	03:57:37	**Log-Likelihood:**	-18911.
converged:	True	**LL-Null:**	-19039.
Covariance Type:	nonrobust	**LLR p-value:**	4.883e-55

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	-1.2742	0.059	-21.632	0.000	-1.390	-1.159
sex_coded	-0.2734	0.036	-7.494	0.000	-0.345	-0.202
race_coded	0.2246	0.028	8.091	0.000	0.170	0.279
Age_Estimated	-0.0161	0.001	-13.086	0.000	-0.019	-0.014

6. Lower CI & Upper CI:

1 to 4 of 4 entries Filter

index	Lower CI	Upper CI	OR
Intercept	0.24917218866785631	0.3138867528815644	0.2796638146227466
sex_coded	0.7082510369044484	0.8171614074425554	0.7607597611200847
race_coded	1.1855464032848282	1.32184174997494	1.2518405379258626
Age_Estimated	0.9816119051355711	0.9863696671872869	0.9839879105840944

7. Confusion Matrix:

```
Confusion Matrix :
 [[11433     0]
 [ 1584     0]]
```