Analyzing Strip Search Rates

The Impact of Arrest Factors
&
Individual Characteristics

By:

Group 44

Yiran Zheng

Jiahao Li

for the course of INF2178H

Faculty of Information

University of Toronto

Table of Contents

1. Introduction

1.1 Background

In recent decades, Canada has experienced a significant increase in its immigrant population, contributing to greater racial and ethnic diversity across the country (Statistics Canada, 2021). While this diversity is often celebrated as a strength, it has also been accompanied by concerns about racial discrimination and unequal treatment of people of color, particularly in the context of law enforcement (Wortley & Owusu-Bempah, 2011). Similarly, gender discrimination may exist within the criminal justice system, resulting in disparate treatment of men and women during arrests and searches (Daly & Chesney-Lind, 1988; Franklin & Fearn, 2008).

Research suggests that men and women may engage in different types of criminal behavior and be subject to different arrest patterns (Steffensmeier & Allan, 1996). For example, men are more likely to be arrested for violent offenses, while women are more often involved in property or drug-related crimes (Daly, 1994; Kruttschnitt & Carbone-Lopez, 2006). Additionally, age is another factor that may influence arrest rates and the likelihood of being searched during an arrest (Farrington, Loeber, & Howell, 2012). In Canada, some common types of crimes that lead to arrests and searches include drug offenses, property crimes, and violent offenses (Boyce, Cotter, & Perreault, 2014). The attitude of the police during an arrest and search can be affected by the level of cooperation exhibited by the offender. Research has shown that individuals who cooperate with law enforcement during an arrest are generally treated more leniently than those who do not (Mastrofski et al., 1996; Reisig et al., 2004).

Police officers face potential risks during arrests and searches, including the possibility of injury or assault. As a result, concerns about officer safety may influence their decisions regarding the use of force, arrests, and searches (Kaminski, DiGiovanni, & Downs, 2004; Terrill & Paoline, 2016). Understanding these complex relationships between demographic factors, police attitudes, and law enforcement practices is essential for promoting more equitable and effective policing strategies.

1.2 Research Objective and Questions

The aim of this study is to explore the relationships between gender, race and police safety with respect to arrests and strip searches. Specifically, we investigate three research questions:

- **ANCOVA Research Question 1:** Is there a significant interaction effect of gender and the number of arrests (dependent variable), while controlling for the likelihood of getting strip searched given race and age in the corresponding arrest location?

- **Logistics Research Question 2:** What is the relationship between the strip search rates of an individual's gender and race in the corresponding region and the likelihood of being strip-searched during an arrest?

- **Logistics Research Question 3:** What is the relationship between the strip search rates of an individual's likelihood of being strip-searched during an arrest in the corresponding region given their arrest occurrence category and whether they cooperated at arrest?

By examining these research questions, we aim to gain a deeper understanding of the potential disparities in policing practices and identify areas for improvement (Fridell, 2017; Novak, 2004). Our findings can help inform policy discussions and contribute to the development of more equitable and effective law enforcement strategies that ensure both public safety and the fair treatment of all individuals, regardless of their gender, race, or immigration status (Epp, 2014; Nix & Wolfe, 2017).

2 Literature Review

The role of gender in policing has been a subject of considerable interest among scholars. While some studies suggest that women are less likely to be arrested than men due to differences in offending behavior (Daly & Chesney-Lind, 1988), others argue that gender stereotypes and societal expectations may influence policing practices (Schulhofer, Tyler, & Huq, 2011). For example, female suspects may be perceived as less dangerous or less likely to commit serious offenses, leading to fewer arrests (Steffensmeier & Demuth, 2006). Additionally, researchers have found that race and age can further intersect with gender, resulting in differential treatment by law enforcement (Gelman, Fagan, & Kiss, 2007; Spohn & Beichner, 2000).

Several studies have explored the relationship between strip search rates and demographic factors such as gender and race. For instance, a report by the New York Civil Liberties Union (2012) found that minority individuals were disproportionately subjected to strip searches during police encounters. Similarly, research has shown that women, especially women of color, may be more likely to experience invasive searches during police interactions (Harcourt, 2003; Kahan, 2008). Regional variations in strip search rates have also been documented, suggesting that local policing practices and community contexts can influence the prevalence of such searches (Lopez, 2017; Smith & Petrocelli, 2001).

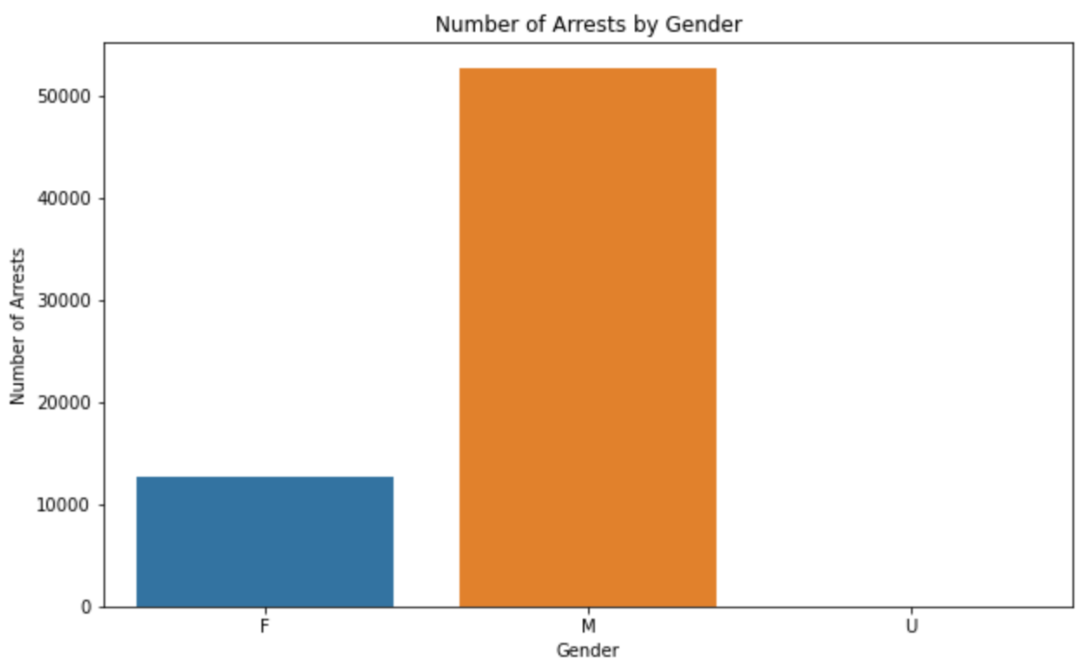
Several studies have examined the relationship between the nature of the offense or occurrence category and the likelihood of being strip-searched during an arrest. Research indicates that individuals arrested for more severe crimes, such as drug offenses or violent crimes, may be more likely to undergo strip searches (Del Carmen, 2004; Mustard, 2001). Furthermore, the suspect's behavior during the arrest, such as their level of cooperation, has been found to be an important determinant of whether a strip search is conducted (Mastrofski, Snipes, & Supina, 1996; Reisig, McCluskey, Mastrofski, & Terrill, 2004). These findings highlight the importance of considering the interaction between offense type and suspect behavior when examining the likelihood of being strip-searched during an arrest.

3. Exploratory Data Analysis

3.1 Descriptive Statistics

The dataset RBDC-ARR-TBL-001 on the Toronto Police Service Public Safety Data Portal provides arrest-related data for 2020 and 2021, including basic information about the arrestee such as race, gender, age, and status at the time of arrest, as well as police records of actions and reasons for searches.

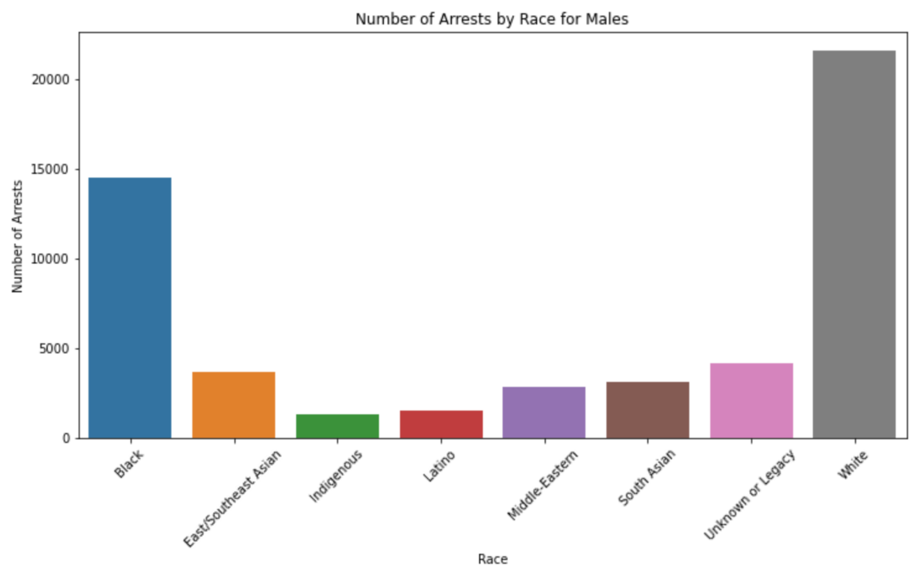
First, to derive a relationship between gender and the number of arrests, we produced Figure 1: Number of Arrests by Gender. This bar chart shows the total number of arrests from each gender over a two-year period. This graph clearly shows that the total number of male arrests far exceeds that of females, while the number of cases with gender 'U' is too small and will be taken out in the subsequent analysis.



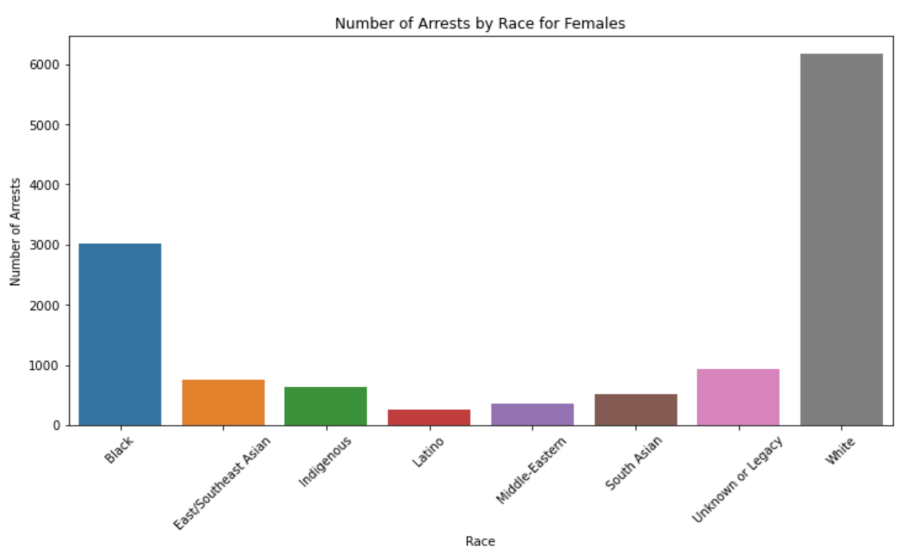
(Figure. 1: Number of Arrests by Gender)

In order to more accurately observe the relationship between total arrests and gender, we explored some variables that may affect the relationship between total arrests and gender. For this purpose, we made Figure 2: Number of Arrests by Race for Males and Figure 3: Number of

Arrests by Race for Females. From these two graphs, we can observe that the percentage of total crime is significantly higher for males than females in East/Southeast Asia, Latino, Middle-Eastern and South Asia, while the percentage of total crime is significantly higher for Indigenous females than males. We conclude that race will be a variable that affects the relationship between total arrests and gender, and we will designate race as a covariate variable in ANCOVA analysis.

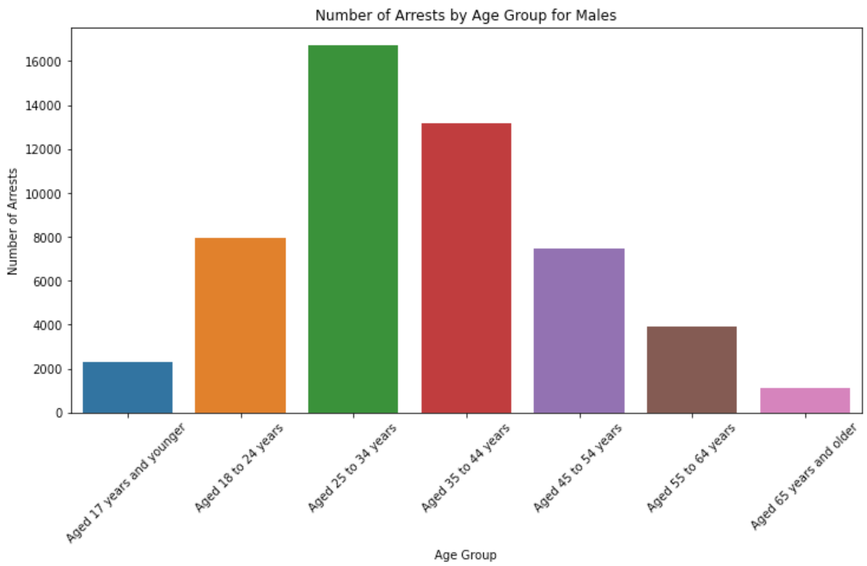


(Figure. 2: Number of Arrests by Race for Males)

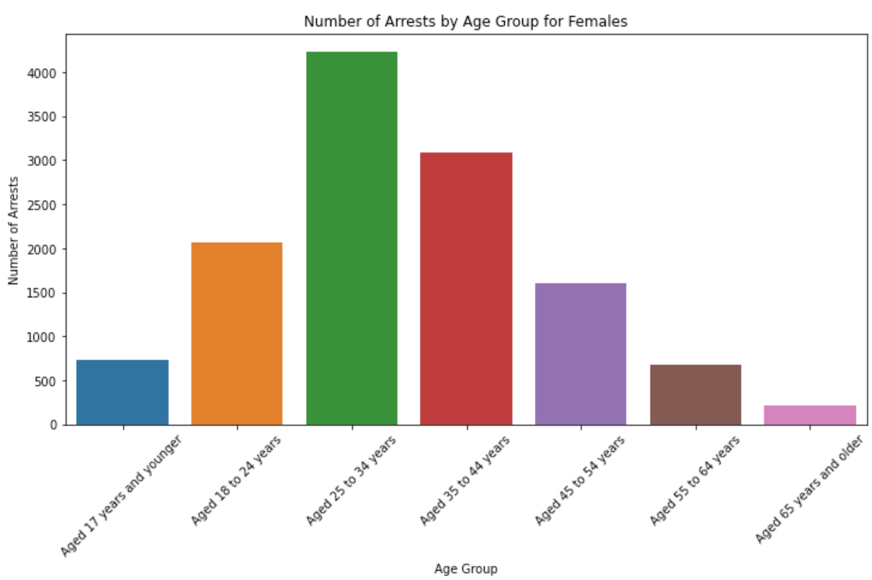


(Figure. 3: Number of Arrests by Race for Females)

To further investigate the variables that may affect the relationship between gender and total number of arrests, we produced Figure 4: Number of Arrests by Age Group for Males and Figure 5: Number of Arrests by Age Group for Females. From these two graphs, we can observe that the percentage of total crime is significantly higher for males than females in age groups: Aged 35 to 44 years, Aged 55 to 64 years, Aged 65 years and older, while the percentage of total crime is significantly higher for females than males in the age group: Aged 17 years and younger. We conclude that age will be a variable that affects the relationship between total arrests and gender, and we will also designate age as a covariate variable in ANCOVA analysis.



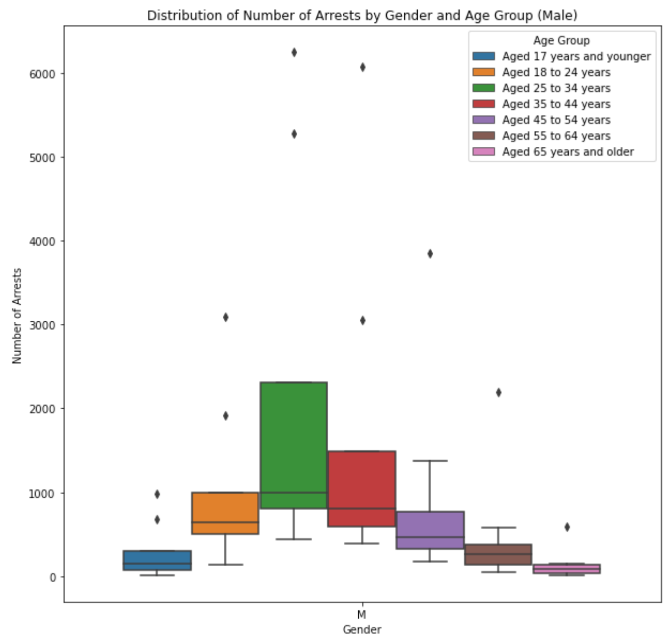
(Figure. 4: Number of Arrests by Age Group for Males)



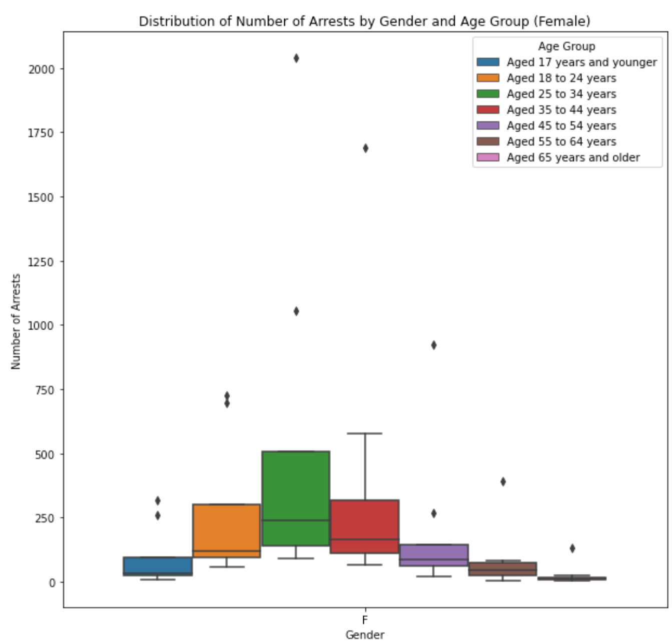
(Figure. 5: Number of Arrests by Age Group for Females)

To further investigate the relationship between the total number of arrests and gender, we produced Figure 6: Distribution of Number of Arrests by Gender and Age Group (Male) and

Figure 7: Distribution of Number of Arrests by Gender and Age Group (Female). From these two box plots we can see that the distribution number of crimes is higher for males than for females in all age groups, and the outlier appears more frequently for males than for females.



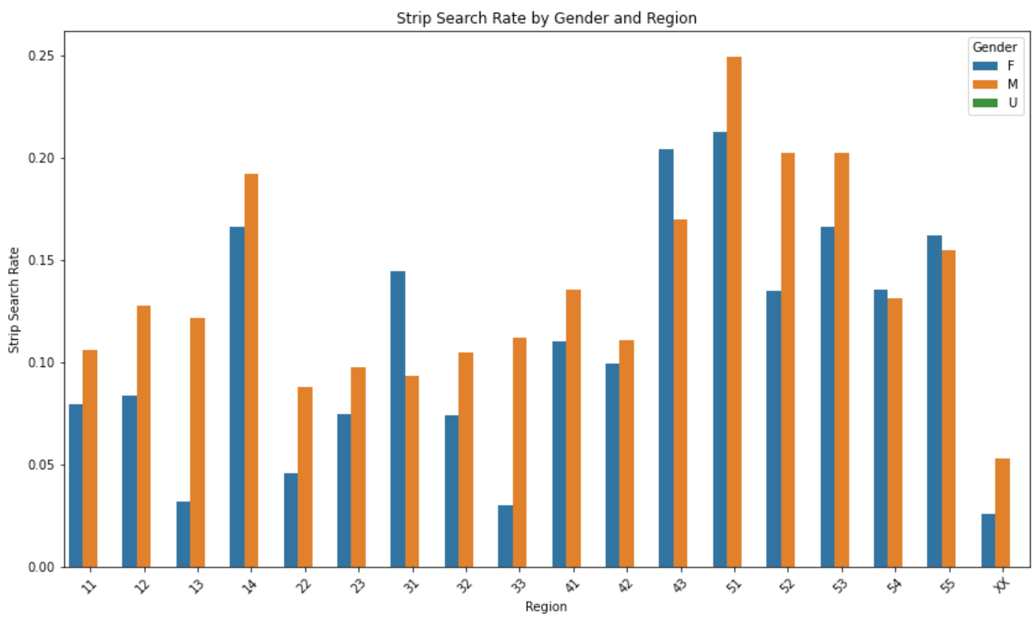
(Figure. 6: Distribution of Number of Arrests by Gender and Age Group (Male))



(Figure. 7: Distribution of Number of Arrests by Gender and Age Group (Female))

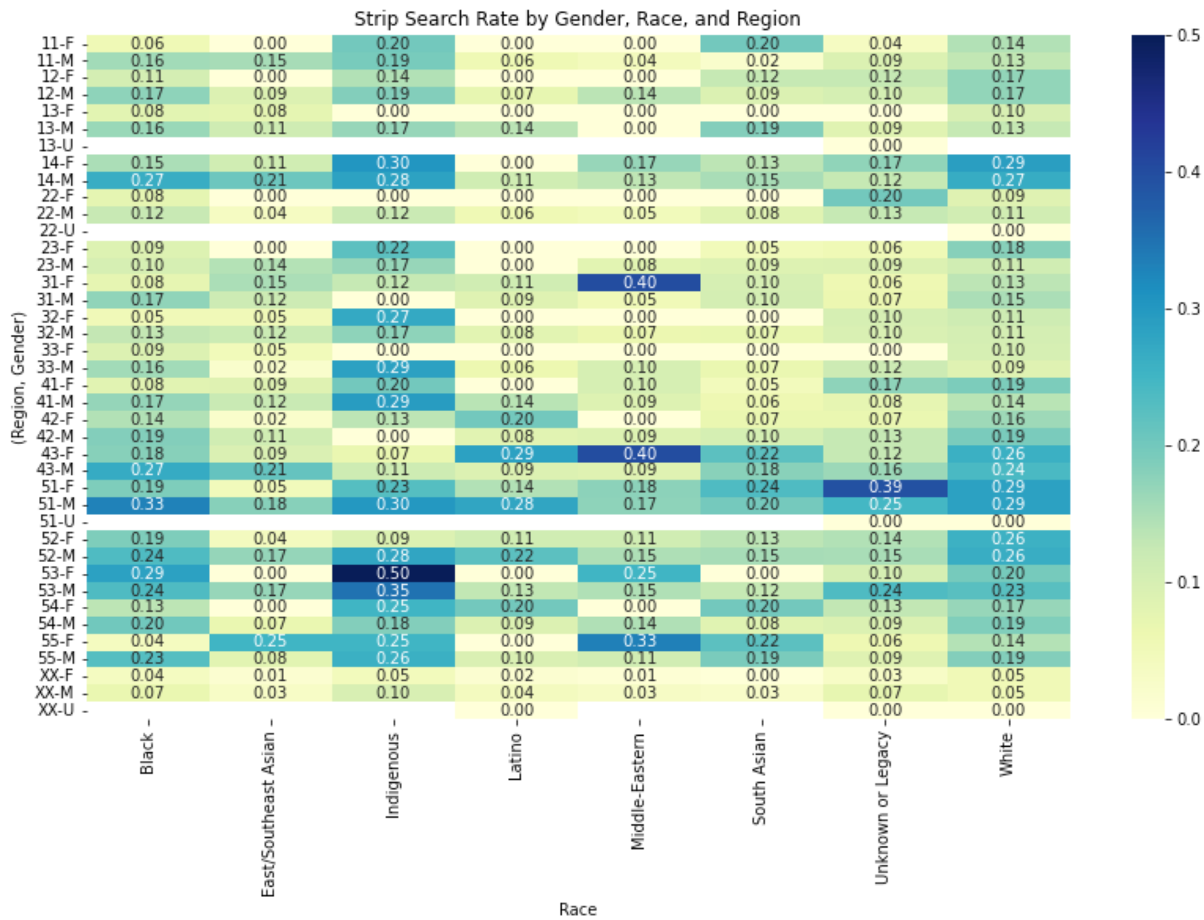
STUDY ON FACTORS AFFECTING STRIP SEARCHES & ARRESTS

To investigate the relationship between strip search rates of an individual's gender and race in the corresponding region and the likelihood of being strip-searched during an arrest, we visualized the strip search rates as Figure 8: Strip Search Rate by Gender and Region. The graph shows that men are more likely to be searched than women in most regions of arrest, with men more than twice as likely to be searched as women in Region 13, Region 22 and Region 33. However, in Region 31, Region 43 and Region 55, women are more likely to be strip-searched than men.



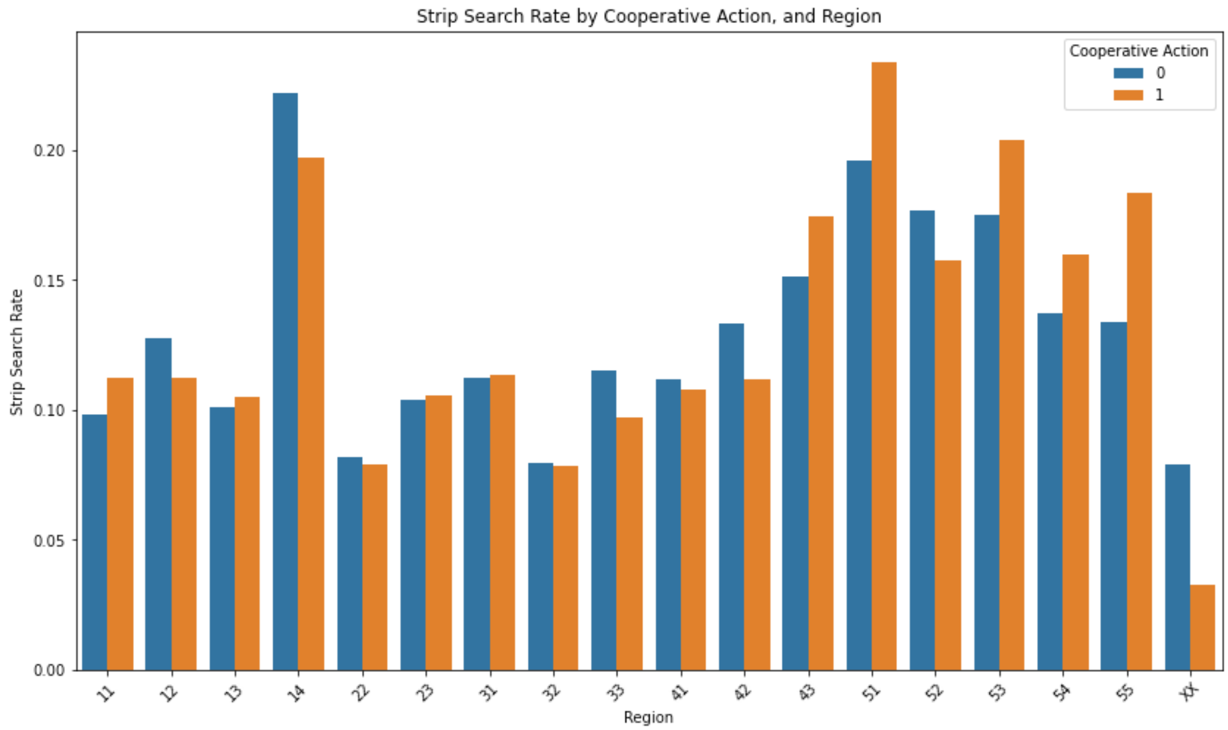
(Figure. 8: Strip Search Rate by Gender and Region)

To further explore the relationship between the probability of strip-searching and gender, region, and race, we produced Figure 9: Strip Search Rate by Gender, Race, and Region. From the heatmap we can see that race can be a factor in the police's judgment of whether or not to conduct a strip search when people of the same gender in the same area are arrested. The most obvious is that Indigenous people, regardless of gender, are susceptible to strip searches in most regions, with female Indigenous people in Region 33 having a 50% strip search rate. Middle Eastern females had a significantly higher probability of being strip-searched in Region 31 and 43 than in other regions, with a 40% probability.



(Figure. 9: Strip Search Rate by Gender, Race, and Region)

To explore the probability that an individual cooperates when arrested, area versus being strip-searched, we produced Figure 10: Strip Search Rate by Cooperative Action, and Region. From the graph we can see that in most areas, the presence or absence of cooperative behavior is not significantly related to the probability of being strip-searched, but outside of GTA and the Region 14, the absence of cooperative behavior results in a higher probability of being strip-searched, while in Region 51, 53, 54, and 55, those with cooperative behavior have a higher probability of being strip-searched at the time of arrest.



(Figure. 10: Strip Search Rate by Cooperative Action, and Region)

To further explore the relationship between the probability of being searched and region, cooperative action and occurrence category, we produced Figure 11: Strip Search Rate by Occurrence Category, Cooperative Action, and Region. As we can see from the heatmap, the probability of being searched is high in most regions when the crime types are Break & Enter, Drug Related, Robbery & Theft, and Weapons & Homicide, with or without cooperative actions. In addition, there are some specific situations that result in 100% being strip-searched, such as homicide crimes without cooperative action in Region 33 and mental health crimes with cooperative action in Region 51.

(Figure. 11: Strip Search Rate by Occurrence Category, Cooperative Action, and Region)

3.2 T-test

To illustrate the difference between each group mathematically, we can use Welch's t-tests to compare the likelihood of getting strip searched given race or age, separately for males and females in the corresponding arrest location. We checked these assumptions before running the t-test:

1. **Nominal two level explanatory variable:** a categorical variable with two levels or categories that do not have any inherent numerical meaning.
2. **Quantitative Outcome Variable:** a quantitative outcome variable refers to a variable that is measured on a continuous scale and used as the dependent variable in the statistical analysis.

3. **Independence:** the data in each group must be independent of each other. The observations in one group should not be related or dependent on the observations in the other group.
4. **Normality:** the data in each group must be normally distributed. The data follows a bell-shaped curve, where the majority of the data points are near the mean, and the distribution is symmetric.

Assumptions 1, 2, and 3 are fulfilled. The data in each group are not normally distributed. However, the sample size is large enough, the t-test can still be robust to violations of the normality assumption and variance assumption. Therefore, we still ran Welch's t-test.

Race/Age and the Number of Arrests

We performed Welch's t-test separately for race or age. Our goal was to determine whether there was a significant difference in the likelihood of getting strip searched given race or age_group for males and females in the corresponding arrest location. To achieve this, we used the following two hypotheses in our two Welch's t-test analysis.

H_0 (Null Hypothesis): The likelihood of getting strip searched given race/age_group for the two independent groups (male and female) are equal.

H_1 (Alternate Hypothesis): The likelihood of getting strip searched given race/age_group for the two independent groups (male and female) are not equal.

Based on the obtained results, the p-value for the test of 'Age_group' is 0.075 and the p-value for the test of 'Perceived_Race' is 0.362. Therefore, we fail to reject the null hypothesis for both variables at the chosen significance level (e.g., 0.05). This means that there is no statistically significant difference in the means of the two groups being compared (males and females) for either 'Age_group' or 'Perceived_Race'.

4. Method

4.1 Dataset Description

The dataset we will use in this project is called RBDC-ARR-TBL-001, which was published by Toronto Police Service Public Safety Data Portal on November 10th, 2022. It refers to the information about arrests and strip searches in a particular jurisdiction from 2021 to 2022, and demographic information of the arrestee. The attributes of arrests and strip searches in the dataset include the arrestee's age, gender, race, event ID, person ID, arrest ID, occurrence category, time and location of the arrest, actions at the arrest, reason for the arrest, reason for the strip search, and whether any contraband or illegal items were found during the search. The dataset contains 65277 times of arrests and strip searches happened in the City of Toronto, some of them outside of City of Toronto boundaries in other jurisdictions. The data is provided in binary format (0 or 1), string format, and categorical format. The dataset is available on Toronto Police Service website(https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/about).

4.2 Independent Variable

Our project contains five independent variables. According to the content of attributes shown on Toronto Police Service website, the detailed information is listed as follows. We made some changes for these variables while cleaning the data.

ANCOVA
Sex (Sex). The dataset records each arrestee's sex as: "F", "M", "U". Its level of measurement is categorical. Its level of measurement is categorical. The "U" has been removed as a gender label since the number of individuals with that label is rather small compared to those labeled as male or female.

Logistic Regression
Sex (Sex). The dataset records each arrestee's sex as: "F", "M", "U". Its level of measurement is categorical. Its level of measurement is categorical. The "U" has been removed as a gender label since the number of individuals with that label is rather small compared to those labeled as male or female.

Perceived Race (Perceived_Race). The dataset records each arrestee's race as: "'White", "Unknown or Legacy", "Black", "South Asian", "Indigenous", "Middle-Eastern", "Latino", "East/Southeast Asian", "nan". Its level of measurement is nominal. We removed the data in this column that contains nan.

Occurrence Category(Occurrence_Category). It refers to the type of arrest occurrence, which are categorized into 31 types.
Actions at arrest - Cooperative(Actions_at_arrest___Cooperative). The dataset records whether the arrestee cooperated with the police when they were arrested.

To perform logistic regression on the four independent variables mentioned above, we made the following modification: we calculated the rate of arrestees who were subjected to strip searches within each group of interest (such as age group or race group) within each arrest location(mentioned below in controlling variable) and used this rate as an independent variable in the analysis.

4.3 Covariate Variable

In the ANCOVA, we used the following two covariate variables.

Perceived Race (Perceived_Race). The dataset records each arrestee's race as: "'White", "Unknown or Legacy", "Black", "South Asian", "Indigenous", "Middle-Eastern", "Latino", "East/Southeast Asian", "nan". We removed the data in this column that contains nan. Its level of measurement is categorical.

Age group (Age_group_at_arrest_). It refers to which age group the arrestee belongs to at arrest. The dataset includes nine age groups: 'Aged 17 years and under', 'Aged 18 to 24 years', 'Aged 25 to 34 years', 'Aged 35 to 44 years', 'Aged 45 to 54 years', 'Aged 55 to 64 years', 'Aged 65 years and older', 'Aged 65 and older'. We combined the 'Aged 65 and older' and 'Aged 65 years and older' groups together and renamed them 'Aged 65 years and older'. The 'Aged 17 years and under' and 'Aged 17 years and younger' groups are combined as 'Aged 17 years and younger''. Its level of measurement is categorical.

For ANCOVA, we also mutate the two covariate similar with what we have done for logistic regression: we calculated the rate of arrestees who were subjected to strip searches within each group of interest within each arrest location(mentioned below in controlling variable) and used this rate as an independent variable in the analysis.

4.4 Controlling Variable

For the logistic regression, we used one controlling variable for both models.

Arrest Location (ArrestLocDiv). It refers to where the arrest took place within Division boundaries. There are 15 location codes, one of them is letters (XX), the others are numbers. XX means the location could not be geo-coded or the arrest took place outside of City of Toronto boundaries in other jurisdictions. Its level of measurement is categorical.

4.5 Dependent Variable

We used two dependent variables in this project. We list the detailed information as follows.

The Number of Arrests (counts). The 'Number of Arrests' variable represents the count of individuals who were arrested for different groups categorized by gender, race, and age group. It is a numerical variable that measures the frequency of arrests within each group. We used the 'groupby()' function to group the data by these variables and calculated the count of arrests for each group. This function automatically removes all missing values (NA) from the dataset.

StripSearch (StripSearch). The variable indicates whether an individual was subjected to a strip search or not. A value of 1 indicates that the individual was strip searched, while a value of 0 indicates that they were not.

For research question 1, we construct a new dataframe copied from the original data set but only with columns (‘Perceived_Race’, Sex’, ‘Age_group__at_arrest_’, ‘ArrestLocDiv’, ‘StripSearch’). And for Logistic regression (research question 2 & 3), we construct two new dataframes, one with ('Perceived_Race', ‘Sex’, ‘ArrestLocDiv’, ‘StripSearch’) and another one with (‘Occurrence_Category’, ‘Actions_at_arrest___Cooperative’, ‘ArrestLocDiv’, ‘StripSearch’).

4.6 ANCOVA

After conducting Exploratory Data Analysis and Welch's t-test, we designed an ANCOVA test to explore our **Research Question 1: Is there a significant interaction effect of gender and the number of arrests (dependent variable), while controlling for the likelihood of getting strip searched given race and age in the corresponding arrest location?**

4.6.1 Assumption Check

Performing a t-test before ANCOVA provides useful information about the differences between two independent groups on a dependent variable and indicates the validity of covariate variables. However, it is still necessary to check for the assumptions of ANCOVA. Checking for assumptions is important for accurate interpretation of ANCOVA results and for ensuring the validity of the analysis. Hence, we checked these assumptions before running the ANCOVA test:

1. **Continuous outcome variable:** a quantitative outcome variable refers to a variable that is measured on a continuous scale and used as the dependent variable in the statistical analysis.

2. **Categorical variables with more than two levels:** a categorical variable with two levels or categories that do not have any inherent numerical meaning.

3. **Constant variance:** the variance of the residuals should be constant across all levels of the covariate. In other words, the spread of the residuals should be roughly the same for all values of the covariate.

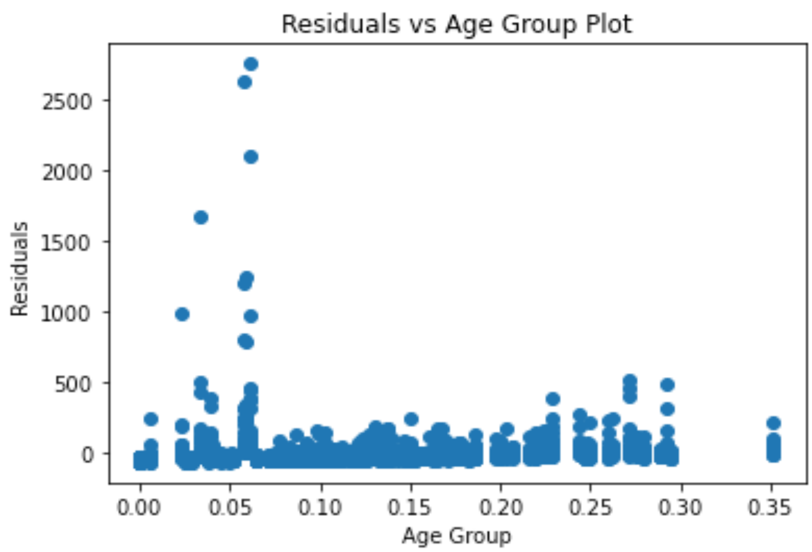
4. **Normality:** the dependent variable should be normally distributed within each group defined by the categorical variable.

5. **Independence of error:** there should be no systematic patterns or correlations among the residuals, which could indicate that the model is not capturing all relevant factors that affect the outcome variable.

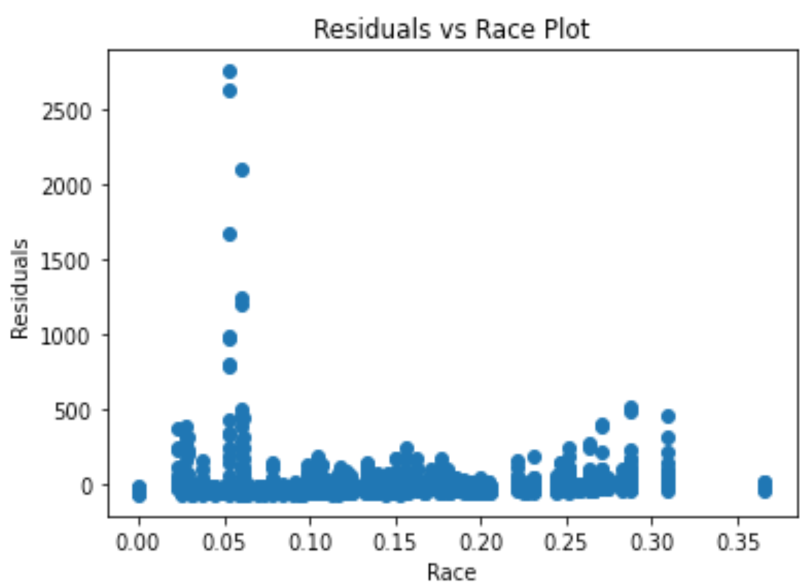
6. **Linearity:** at each level of categorical independent variable, the covariate should be linearly related to the dependent variable.

7. **Homogeneity of regression slopes (covariate coefficients):** the relationship between the covariate and the dependent variable should be the same across all groups defined by the categorical variable.

Assumptions 1 and 2 are fulfilled as the number of arrests is a continuous variable, and sex is a categorical variable with two categories. To test for Assumption 3 (homoscedasticity), we built a linear model of counts ~ Perceived_Race + Sex + Age_group__at_arrest_ and plotted two graphs: "Residuals vs Age Group Plot (Figure. 12)" and "Residuals vs Race Plot (Figure. 13)".



(Figure. 12: Residuals vs Age Group Plot)

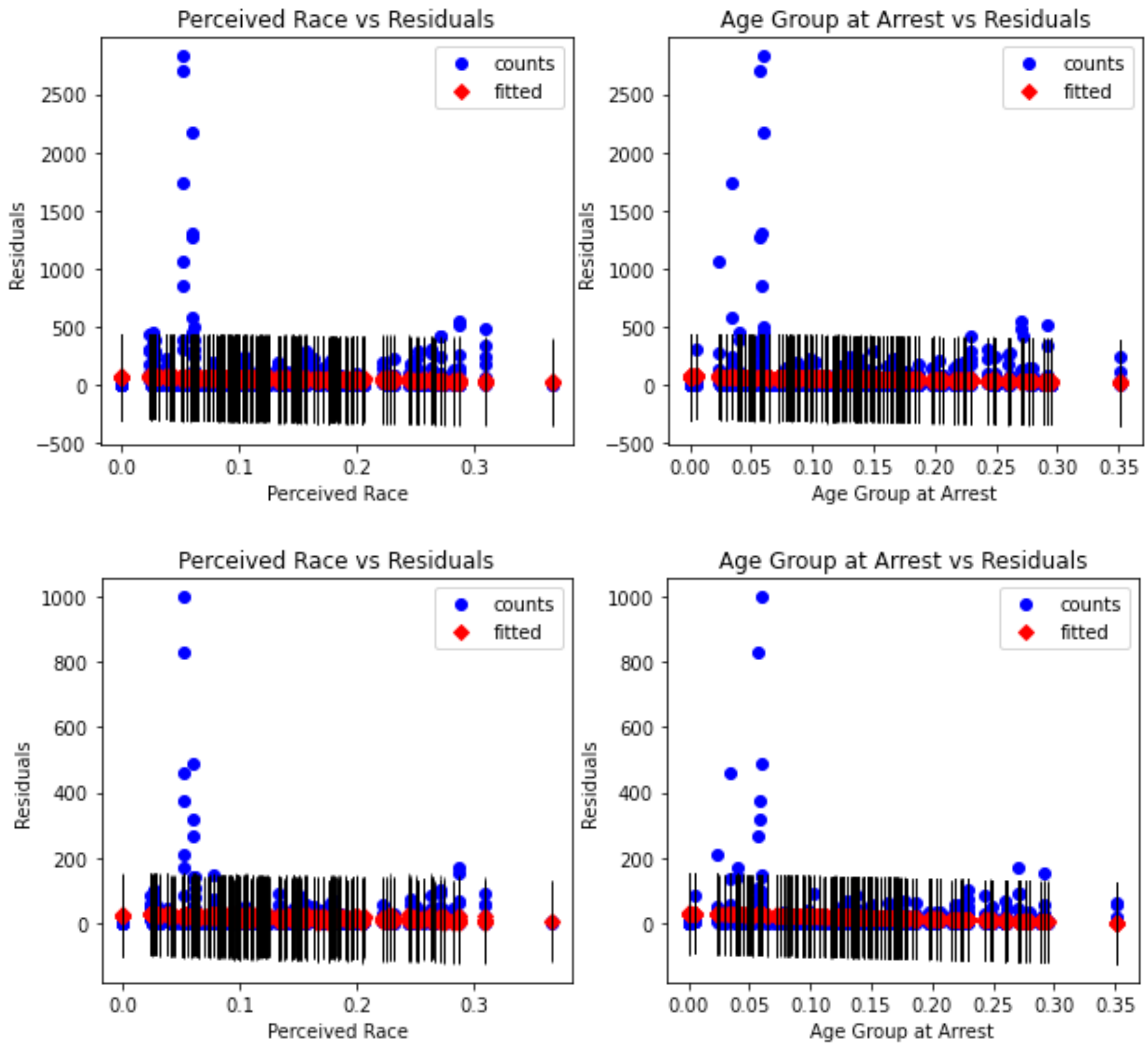


(Figure. 13: Residuals vs Race Plot)

By observing the pattern of the plot, the variance seems to be constant for most of the data, but there are some exceptions around 0.05 for both covariates. Therefore, the assumption of variance homogeneity is not fully met. To test for Assumption 4 (normality of residuals), we conducted a t-test and found that the data in each group are not normally distributed. However, since the sample size is large, we can still conduct further research using this model. To test for Assumption 5 (independence of residuals), we plotted "Residuals vs Predicted Value". Although there is a slightly descending trend in the plot (Figure. 14), the assumption of independence of

(Figure. 14: Independence of error)

errors are not significantly violated, and we can still conduct our ANCOVA. To test for Assumption 6 (linearity), we built two new models: counts ~ Perceived_Race + Age_group__at_arrest_ using either male or female data. We then plotted the residuals and the likelihood of interest (i.e., Race and Age_group), and found a significant outlier around 0.05 for all plots ((Figure. 15)). Therefore, the assumption of linearity is not fully met. To test for Assumption 7 (no multicollinearity), we examined the correlation matrix of the independent variables and found no significant correlation between them, see ((Figure. 15)). Thus, the assumption is met.



(Figure. 15: Residuals Against Each Independent Variable)

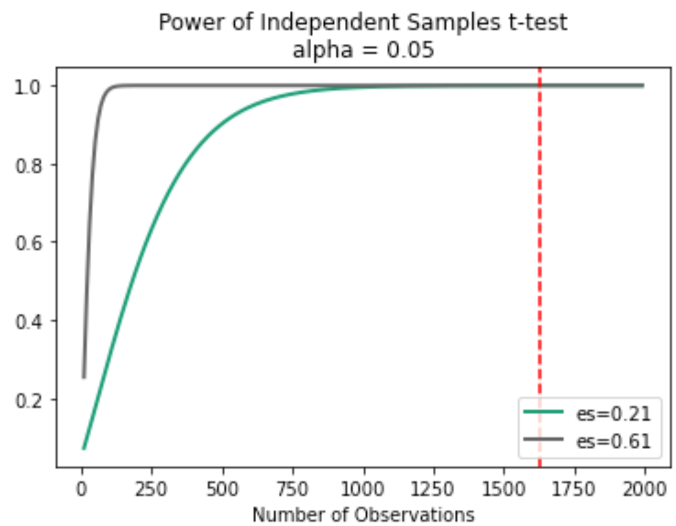
Overall, the model has some violations of the assumptions, but it is still suitable for further research. However, it is important to interpret the results with caution, and consider potential sources of bias or confounding factors.

4.6.2 Power Analysis

After verifying the assumptions, it is crucial to perform a power analysis to determine the appropriate sample size needed to detect a statistically significant effect size with a given level of confidence. To determine the statistical power of the ANCOVA for examining the interaction effect of gender and the number of arrests, a power analysis was conducted. The power analysis was performed using the Python package statsmodels, specifically the TTestIndPower function.

The data used for the power analysis was a subset of the original dataset that included only the variables of interest: Perceived_Race, Sex, Age_group__at_arrest_, and counts. The data was cleaned by removing missing values and excluding records with a gender of 'U'. The Age_group__at_arrest_ variable was also recorded to simplify the age categories. The mean counts of arrests for male and female groups were calculated, along with their respective standard deviations. The effect sizes were then calculated as the absolute difference between the mean counts of arrests for males and females divided by the standard deviation of each gender group. A range of sample sizes from 10 to 2000 in increments of 10 were simulated for each effect size using the TTestIndPower function. The power of the test was calculated for an alpha level of 0.05.

In the Figure. 16, the red dotted line indicates the number of our observations, which is larger than the number of observations required for both effect sizes. The results of the power analysis indicated that for an effect size of 0.61, a sample size of approximately 100 would be required to achieve a power of 0.9. For an effect size of 0.21, a sample size of approximately 1000 would be required to achieve a power of 0.9. The power analysis was visualized using a plot that showed the relationship between sample size, effect size, and statistical power. Thus, the sample can be considered large enough for both effect sizes.



(Figure. 16: Power of Independent Samples t-test\nalpha = 0.05)

4.6.3 Test

The specific ANCOVA model being used here has the following components:

- Dependent Variable: The dependent variable is "counts", which is a continuous variable representing the number of arrests.
- Covariate Variable: The continuous covariates are "Perceived_Race" and "Age_group__at_arrest_".

These variables are mutated as mentioned in 3.3 and will be used to control for their effects on the dependent variable, and to test for any interaction effects with the independent variable.

- Between: The between-subjects factor is "Sex", which is a categorical variable with two levels (male and female). This variable will be used to compare the means of the dependent variable between the two groups.

The ANCOVA will estimate the adjusted means for each level of "Sex", controlling for the effects of "Perceived_Race" and "Age_group__at_arrest_". The analysis will also test for any interaction effects between "Sex" and the covariates. This will provide insights into whether the effects of "Perceived_Race" and "Age_group__at_arrest_" on the dependent variable differ between males and females.

4.7 Logistic Regression

After conducting ANCOVA, we want to find out if the likelihood of an individual being strip searched given their age group and race would have some mathematical relationship between other variables, so we generated a logistic regression model to explore research question 2.

Research Question 2: What is the relationship between the strip search rates of an individual's gender and race in the corresponding region and the likelihood of being strip-searched during an arrest?

However, the results showed that the relationship was not clear, and the model's classification was not meaningful. Therefore, we sought to identify other independent variables that may be more strongly correlated with the outcome of strip searches. To explore this, we developed another logistic regression model for research question 3.

Research Question 3: What is the relationship between the strip search rates of an individual's likelihood of being strip-searched during an arrest in the corresponding region given their arrest occurrence category and whether they cooperated at arrest?

4.7.1 Assumption Check

Before generating logistics models, we checked these assumptions:

1. **Linearity of the logit:** the linear relationship between the continuous independent variables and log odds of the dependent variable

2. **No multicollinearity:** The independent variables should not be highly correlated with each other.

3. **No influential outliers**

4. **The dependent variable is binary or ordinal:** the dependent variable only has two or a limited number of discrete categories.

5. **The observations are independent of each other:** each observation in the dataset is independent of each other.

6. **Independence of error:** the residuals should not be correlated with each other.

7. **Large sample size:** the sample size should be large enough to provide sufficient power to detect significant effects.

After conducting an analysis of the logistic regression models, it appears that some of the model assumptions were not fully met. For assumption 1, which requires linearity of the logit, it was found that there was no clear linearity of the logit for either model, despite not being able to check this assumption before building the model. However, assumption 2 was met as the variance inflation factor for all variables in both models were found to be less than 10. For assumption 3, which relates to outliers, it was determined that there were no outliers as the independent variables were all distributed uniformly between 0 and 1. Assumption 4 was met as the dependent variable was binary, taking on values of either 0 or 1. Assumption 5 was also met, as each observation in the dataset was recorded independently and no duplicates or other issues were present. However, assumption 6 was not met, as there was a pattern in the residuals of both models, indicating a lack of independence of the errors. Finally, assumption 7 was met, as the sample size was sufficiently large with 65275 observations.
Despite not meeting all of the model assumptions, it is still important to build and analyze the logistic models to explore our research questions. However, it is necessary to interpret the results with caution and consider the potential impact of violating the model assumptions. Future research could explore alternative modelling techniques or further investigate the underlying factors contributing to the violation of certain assumptions in this analysis.

4.7.2 Model Creation

For both logistic regressions, we use 80% of the observations as training data and use 20% of the observations as testing data for testing accuracy.

For research question 2, the model is built using the logit () function of the statsmodels. formula. API package based on the formula StripSearch ~ Perceived_Race + Sex. Perceived_Race and Sex are mutated as mentioned in 3.2. This model is used to predict the probability of an individual being subjected to a strip search based on the likelihood of being strip-searched during an arrest at the corresponding location given their race and gender.

For research question 3, the model is also built using the logit () function of statsmodels. formula.API package but based on the formula StripSearch ~ Occurrence_Category + Actions_at_arrest___Cooperative. Occurrence_Category and Actions_at_arrest___Cooperative are mutated as mentioned in 3.2. This model is used to predict the probability of an individual being subjected to a strip search based on the likelihood of being strip-searched during an arrest at the corresponding location given their occurrence category and whether they are cooperative during arrest.

5. Results/Findings

5.1 ANCOVA

The ANCOVA table shows the relationship between dependent variable: number of arrests, independent variable: Sex and other two covariate variables: Perceived_Race and Age_group_at_arrest_ (these two covariate variables are mutated as mentioned in 3.3). The table includes four rows, each representing a source of variance: Sex (gender), Perceived_Race (perceived race), Age_group_at_arrest_ (age group at arrest), and Residual. The first three rows correspond to the effects of the corresponding variables on the dependent variable, while the fourth row represents the error term.

Table 1: ANCOVA Test of Perceived Race and Age Group

Source	SS	DF	F	p-unc	Np2
Sex	5.580529e+05	1	26.063386	3.693409e-07	0.015815
Perceived_Race	1.568199e+04	1	0.732414	3.922283e-01	0.000451
Age_group__at_arrest_	6.198837e+04	1	2.895114	8.904015e-02	0.001782
Residual	3.472925e+07	1662	NaN	NaN	NaN

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

Based on the table, the only significant effect is for the variable "Sex" (gender), which has a p-value of 3.693409e-07, indicating that gender has a statistically significant effect on the number of arrests while controlling for the effects of perceived race and age group at arrest. However, the effects of perceived race and age group at arrest are not significant at the conventional alpha level of 0.05, suggesting that they do not have a significant impact on the number of arrests while controlling for gender.

5.2 Logistic Regression

Research Question 2:

Table 2: Logistic Test Results of Sex and Perceived Race

Dep. Variable:	StripSearch	**No. Observations:**	52210
Model:	Logit	**Df Residuals:**	52207
Method:	MLE	**Df Model:**	2
Date:	Wed, 12 Apr 2023	**Pseudo R-squ.:**	0.08221
Time:	02:09:07	**Log-Likelihood:**	-17593.
converged:	True	**LL-Null:**	-19169.
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
Intercept	-3.2176	0.030	-107.523	0.000	-3.276	-3.159
Sex	1.1837	0.466	2.538	0.011	0.269	2.098
Perceived_Race	7.4451	0.432	17.245	0.000	6.599	8.291

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

For research question 2, the table shows the logistic regression results for the association between StripSearch and two independent variables, Sex and Perceived_Race. The intercept is statistically significant, indicating a baseline level of the dependent variable. Sex has a statistically significant association with StripSearch(p-value = 0.011), while Perceived_Race has a highly significant association with StripSearch(p-value = 1.2157151736017803e-66).

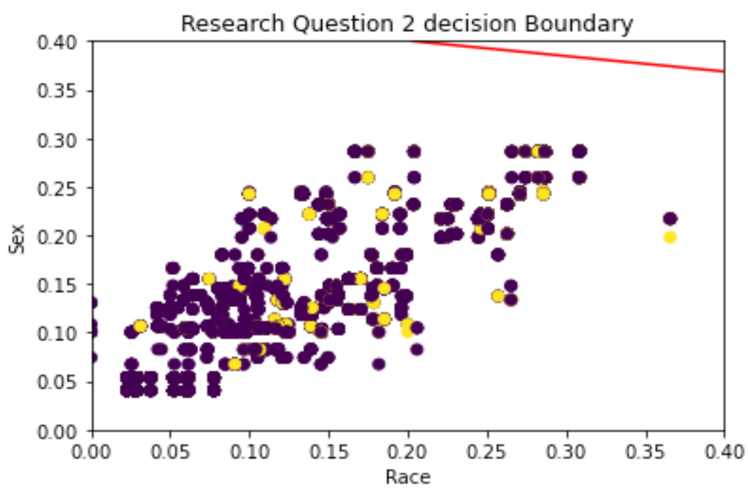
After constructing the model, we test the accuracy of the logistic regression, and the result is high at 0.882. However, further analysis using the confusion matrix (See Table. 3) shows that 11,524 of the predictions were true Positive (TP), meaning that the model correctly predicted that these instances belong to the positive class. There were no false negatives (FN), meaning that the model correctly predicted all of the positive instances. However, there were also no true negatives (TN) and 1529 false positives (FP), which means that the model did not correctly predict any of the negative instances. It is impossible to calculate the odds ratio for this model as FN and TP are 0.

Table 3: Confusion Matrix

		Prediction Results	
		Positive (PP)	Negative (PN)
Actual	Positive (P)	11524(TP)	0(FN)
Observations	Negative (N)	1529(FP)	0(TN)

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

The decision boundary plot revealed that the model predicted all observations as not being strip searched (0), despite the significant associations found in the logistic regression, see Figure.17. Therefore, the logistic model is meaningless given independent variables (Perceived_Race and Sex), the confidence interval is also meaningless, we did not plot it, but the coefficients of the confidence interval is showed below.



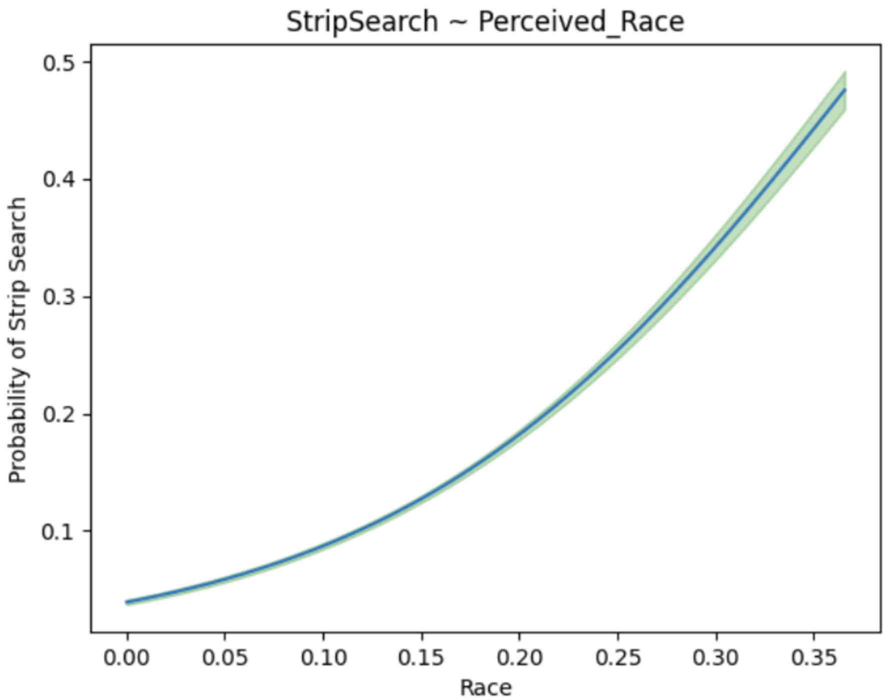
(Figure. 17: Research Question 2 decision Boundary)

Table 4: CI for Sex and Perceived_Race

	Lower CI	Upper CI	OR
Intercept	-3.276288	-3.158984	-3.217636
Sex	0.269493	2.098002	1.183747
Perceived_Race	6.598949	8.291261	7.445105

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

The logistic model for research question 2 involves two independent numerical variables, making it infeasible to display the prediction interval in a two-dimensional graph. However, a graph depicting the logistic regression for the dependent variable of research question 2 and one of its independent variables, 'Perceived_Race', is presented below. The green region on the graph represents the 95% confidence interval of the probability of undergoing a strip search.



(Figure. 18: Logsitic Regression：StripSearch ~ Perceived_race)

A new logistic regression model was developed in research question 3 to identify independent variables that may be more strongly correlated with the outcome of strip searches.

Research Question 3:

The results show that both Occurrence_Category and Actions_at_arrest___Cooperative are statistically significant predictors of StripSearch. Specifically, for a one-unit increase in Occurrence_Category, the log odds of StripSearch increase by 7.5449 units ($z = 70.465$, $p < 0.001$). On the other hand, for a one-unit increase in Actions_at_arrest___Cooperative, the log odds of StripSearch decrease by -0.8907 units ($z = -3.645$, $p < 0.001$).
The Pseudo R-squared value of 0.2580 indicates that the model explains about 26% of the variability in StripSearch. The LLR p-value of 0.000 indicates that the model as a whole is statistically significant.

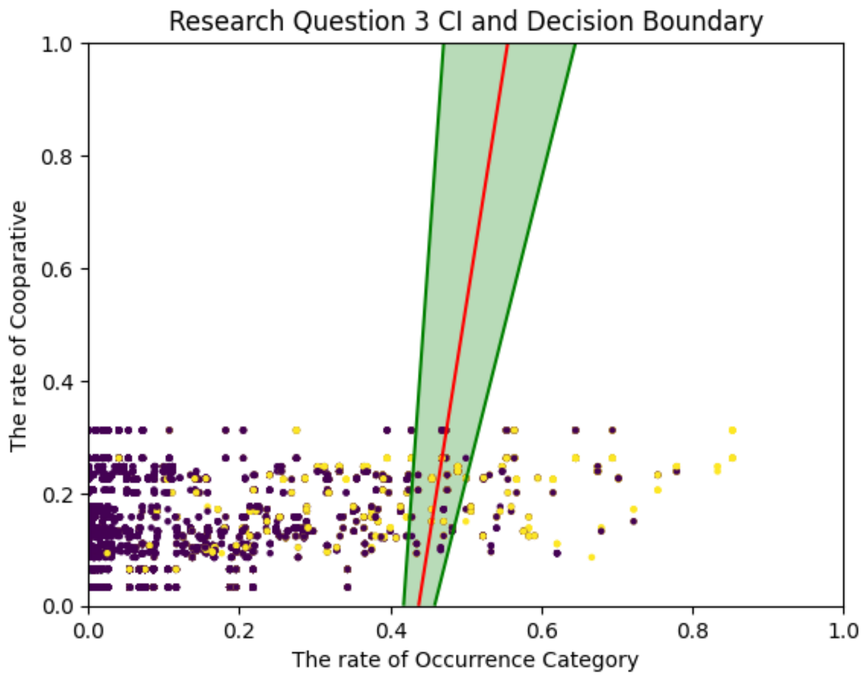
Table 5: Logistic Test Results of Occurrence Category and Actions at Arrest

Dep. Variable:	StripSearch	**No. Observations:**	52088
Model:	Logit	**Df Residuals:**	52085
Method:	MLE	**Df Model:**	2
Date:	Wed, 12 Apr 2023	**Pseudo R-squ.:**	0.2580
Time:	02:33:29	**Log-Likelihood:**	-14237.
converged:	True	**LL-Null:**	-19187.
Covariance Type:	nonrobust	**LLR p-value:**	0.000

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
Intercept	-3.2999	0.032	-102.852	0.000	-3.363	-3.237
Occurrence_Category	7.5449	0.107	70.465	0.000	7.335	7.755
Actions_at_arrest___Cooperative	-0.8907	0.244	-3.645	0.000	-1.370	-0.412

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

The new logistic regression model showed a slight improvement in accuracy compared to the previous model, with an increase from 0.882 to 0.889. However, upon further analysis by plotting the decision boundary (See Figure. 19). The green region in a logistic regression plot represents the confidence interval for the estimated coefficients. It is bounded by two lines, which correspond to the upper and lower limits of the confidence interval at a specified level of significance, typically 0.05, corresponding to a 95% confidence level. However, it is important to note that the level of significance used to calculate the confidence interval can vary depending on the study design and research question. We found that this model is significantly better than the previous one, as it was able to effectively separate the observations into two distinct sections.



(Figure. 19: Research Question 3 CI and Decision Boundary)

A similar result is also shown in the confusion matrix. The model made a total of 12,023 predictions. Of these, 11,156 were true positive (TP), meaning the model correctly predicted that 11,156 instances belong to the positive class. 355 were false negatives (FN), meaning the model incorrectly predicted that 355 instances belonged to the negative class. 1,088 were false positives (FP), meaning the model incorrectly predicted that 1,088 instances did not belong to the negative class when they actually did. Finally, 424 were true negatives (TN), meaning the model correctly predicted that 424 instances belonged to the negative class. Based on the confusion matrix table, the odds ratio for this model is OR = 12.2466. The independent variable is associated with higher odds of the dependent variable.

Table 6: Confusion Matrix

		Prediction Results	
		Positive (PP)	Negative (PN)
Actual	Positive (P)	11156(TP)	355(FN)
Observations	Negative (N)	1088(FP)	424(TN)

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

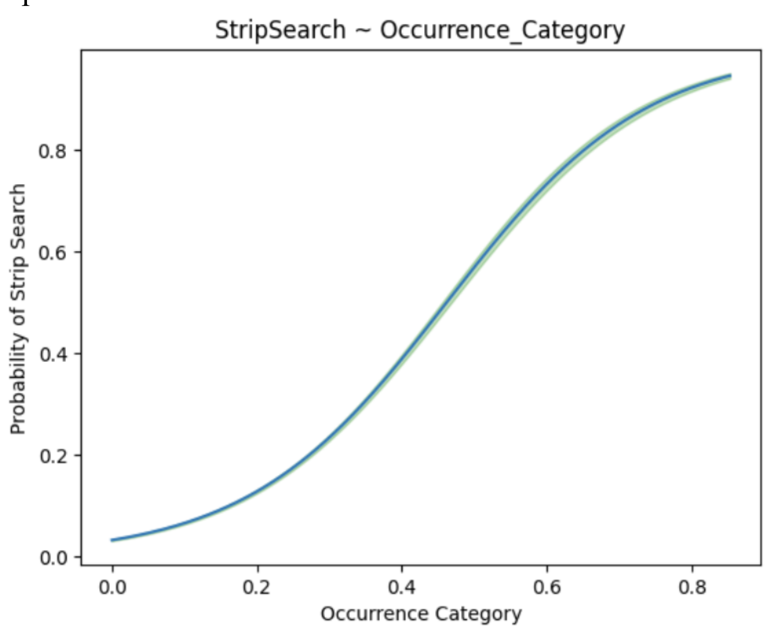
The coefficient of a confidence interval for the logistic regression of research question 3 is shown below:

Table 7: CI for Occurrence_Category and Action_at_arrest__Cooperative

	Lower CI	Upper CI	OR
Intercept	-3.362757	-3.236991	-3.299874
Occurrence_Category	7.335011	7.754728	7.544869
Action_at_arrest__Cooperative	-1.369587	-0.411804	-0.890695

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

Similar to research question 2, the logistic model for research question 3 involves two independent numerical variables, making it unfeasible to display the prediction interval in a two-dimensional graph. However, a graph depicting the logistic regression for the dependent variable of research question 3 and one of its independent variables, 'Occurrence_Category ', is presented below. The green region on the graph represents the 95% confidence interval of the probability of undergoing a strip search.



(Figure. 20: Logsitic Regression： StripSearch ~ Occurrence_Category)

6. Discussion

Based on our results, we found some of these factors do affect the rate of searches or the number of arrests happening from 2020 to 2021. The results suggest that gender, perceived race, occurrence category, and actions at arrest-cooperative are significant predictors of the likelihood of being strip-searched during an arrest.

We designed research question 1 to investigate the interaction effect of gender and the number of arrests, while controlling for perceived race and age group at the time of arrest. The ANCOVA analysis revealed that gender significantly influences the number of arrests, while the perceived race and age group at the time of arrest did not have a significant effect on the number of arrests. The finding that gender has a significant effect on the number of arrests is consistent with previous research in the field, which suggests that males are more likely to be arrested than females. The result suggests that gender is an essential factor to consider when examining the number of arrests, and it may be useful for policymakers and law enforcement officials to consider gender-based differences in the context of arrest. On the other hand, the non-significant effect of perceived race and age group at the time of arrest on the number of arrests is somewhat surprising. This finding seems to contradict previous background studies that have shown that race and age are significant predictors of arrest rates. However, it is important to note that the current study controlled for the likelihood of getting strip searched given race and age in the corresponding arrest location, which may have contributed to the non-significant effect.

The logistic regression analysis for research question 2 explores the relationship between an individual's gender and race in the corresponding region and the likelihood of being strip-searched during an arrest. The results showed that both gender and perceived race had significant associations with the likelihood of being strip-searched. One possible explanation for this finding is that different racial groups and gender may be subject to varying levels of scrutiny or suspicion from law enforcement officers, which could also impact the rate of strip searches. Nonetheless, the limitations of the experiment and data collection must be taken into consideration when interpreting these results. Specifically, gender was found to have a statistically significant association with strip search, while perceived race had a highly significant association with strip search. However, the logistic regression model's accuracy was high at 0.882, but further analysis using the confusion matrix revealed that the model did not correctly predict any of the negative instances. The decision boundary plot also revealed that the model predicted all observations as not being strip searched (0), despite the significant associations found in the logistic regression. Therefore, the logistic model is meaningless, and a new logistic regression model was developed in research question 3 to identify independent variables that may be more strongly correlated with the outcome of strip searches.

The results of logistic regression analysis for research question 3 suggest that both Occurrence_Category and Actions_at_arrest___Cooperative are significant predictors of StripSearch. The findings demonstrate that as the Occurrence_Category increases, the likelihood of being strip-searched also increases significantly. Conversely, as Actions_at_arrest___Cooperative increase, the probability of being strip-searched decreases significantly. These results imply that the circumstances surrounding an individual's arrest and their behavior during the arrest play a crucial role in determining whether they are subjected to strip searches. The Pseudo R-squared value of 0.2580 indicates that the model explains approximately 26% of the variability in StripSearch. Although this value may not seem very high, it is within the acceptable range for logistic regression models. This new logistic regression model outperforms the previous model in terms of accuracy, with an increase from 0.882 to 0.889. The decision boundary plot and confidence interval demonstrate that the model effectively separates the observations into two distinct sections. Furthermore, the confusion matrix shows that the model accurately predicted the majority of instances, with 11,156 true negatives and 424 true positives. However, there were also 1,088 false negatives and 355 false positives, suggesting that the model still has some room for improvement. These results have important implications for law enforcement agencies and policymakers, as they underscore the need for clear guidelines and regulations surrounding strip searches and their use.

Our three research questions examined the relationships between different factors and the likelihood of being strip-searched during an arrest. The results of research question 1 showed that gender had a statistically significant effect on the number of arrests, while perceived race and age group at arrest did not have significant impacts. Our findings suggest that gender is an important predictor of arrest rates, while the perceived race and age group at the time of arrest may not have a significant impact when controlling for other factors. Further research is needed to confirm these results and explore other factors that may influence arrest rates. The results of research question 2 indicated that both gender and perceived race were statistically significant predictors of strip search, with perceived race having a highly significant association. However, the model developed in research question 2 had a poor performance as it predicted all observations as not being strip-searched. To improve the model's performance, research question 3 added the variables of occurrence category and actions at arrest-cooperative to the logistic regression model. The results indicated that both variables were statistically significant predictors of strip search. The new model demonstrated a slight improvement in accuracy and effectively separated the observations into two distinct sections. It is possible that other independent variables may have a stronger correlation with the likelihood of being strip-searched during an arrest than an individual's gender and race in the corresponding region. The models developed in the study may have limitations, and other variables not examined in the study may also be significant predictors. Hence, future research should continue to investigate the factors that contribute to the likelihood of being strip-searched during an arrest to improve the accuracy of predictive models and inform policy decisions.

The discussion highlights the significant effect of gender on the number of arrests and the importance of addressing gender-based differences in arrest contexts. Law enforcement agencies should provide gender-sensitive training to officers to reduce potential biases, while policymakers should promote gender equality in the criminal justice system. Although perceived race and age group at arrest did not significantly affect the number of arrests, continuous monitoring of these factors is crucial to prevent potential discrimination. Policymakers and law enforcement agencies must maintain transparency in their practices and communicate openly with the public to address concerns related to racial and age-based profiling.

Moreover, the study found that occurrence category and cooperative actions at arrest are significant predictors of strip searches, emphasizing the need for clear guidelines and regulations. Law enforcement agencies should establish strict protocols for conducting strip searches, respecting individuals' dignity and rights, and training officers to assess the necessity of a strip search based on circumstances and behavior. It is crucial to exercise caution when interpreting the results of this study, as there were limitations to the experiment and data collection. These limitations may include issues related to sample size, measurement accuracy, and potential confounding variables that were not accounted for in the study design.

In the dataset, there are some biases in the recording of the data itself, for example, some arrests happened in places that could not be geo-coded and the arrests that took place outside of the City of Toronto boundaries in other jurisdictions are all indicated by XX. However, the place that could not be geo-coded may belong to the City of Toronto. The data recorded in this way will have errors, further leading to bias in the results of the data analysis done with that data. Besides, all independent variables and dependent variables we used in this project are not normally distributed. The reason we insist on doing the test is because the sample size is large enough to remain robust to assumption violations. Therefore, our results of ANCOVA and logistic regression are potentially affected, which means the bias may exist. It is important to acknowledge the models' limitations and that unexamined factors may influence the likelihood of strip searches. Future research should explore factors contributing to strip searches during arrests and potential biases in law enforcement practices. This will enable policymakers and law enforcement agencies to develop more effective strategies for addressing potential inequalities and ensuring fair treatment within the criminal justice system.

7. Conclusion

In conclusion, this study aimed to investigate the relationship between gender, perceived race, and other factors in the context of arrest and strip searches in Canada from 2020 to 2021. The results demonstrated that gender plays a significant role in arrest rates, emphasizing the need to address gender-based differences in the criminal justice system. Additionally, the study found that occurrence category and cooperative actions at arrest significantly influence the likelihood of being strip-searched during an arrest. These findings suggest that law enforcement agencies should provide officers with gender-sensitive training, establish clear guidelines for conducting strip searches, and maintain transparency in their practices. Policymakers should promote gender equality in the criminal justice system and continuously monitor factors such as race and age to prevent potential discrimination.

Although the models developed in the study provide valuable insights, it is crucial to acknowledge their limitations. Future research should continue to explore factors that contribute to the likelihood of being strip-searched during an arrest and potential biases in law enforcement practices. By doing so, researchers can help to inform policy decisions and guide the development of more effective strategies to ensure fair treatment and equality within the criminal justice system. Ultimately, addressing these issues is not only essential for fostering trust between law enforcement agencies and the communities they serve but also for promoting a just and equitable society.

References

Boyce, J., Cotter, A., & Perreault, S. (2014). Police-reported crime statistics in Canada, 2013. Juristat, Catalogue no. 85-002-X.

Daly, K. (1994). Gender, crime, and punishment: Is justice gender-blind, or are men and women offenders treated differently by the courts? Yale University Press.

Daly, K., & Chesney-Lind, M. (1988). Feminism and criminology. Justice Quarterly, 5(4), 497-538.

Del Carmen, R. V. (2004). Racial profiling revisited: "Driving while black" and investigatory stops. Criminology & Public Policy, 3(1), 25-31.

Epp, C. R. (2014). Pulled Over: How Police Stops Define Race and Citizenship. University of Chicago Press.

Farrington, D. P., Loeber, R., & Howell, J. C. (2012). Young adult offenders: The need for more effective legislative options and justice processing. Criminology & Public Policy, 11(4), 729-750.

Fridell, L. (2017). Producing bias-free policing: A science-based approach. Springer.

Gelman, A., Fagan, J., & Kiss, A. (2007). An analysis of the New York City Police Department's “stop-and-frisk” policy in the context of claims of racial bias. Journal of the American Statistical Association, 102(479), 813-823.

Harcourt, B. E. (2003). Rethinking racial profiling: A critique of the economics, civil liberties, and constitutional literature, and of criminal profiling more generally. The University of Chicago Law Review, 1275-1381.

Kahan, D. M. (2008). Two conceptions of emotion in risk regulation. The University of Chicago Law Review, 647-681.

Kaminski, R. J., DiGiovanni, C., & Downs, R. (2004). The use of force between the police and persons with impaired judgment. Police Quarterly, 7(3), 311-338.

Lopez, M. P. (2017). Racial and ethnic bias in New York City stop-and-frisk policing: An exploratory analysis. Race and Justice, 7(2), 98-123.

Mastrofski, S. D., Snipes, J. B., & Supina, A. E. (1996). Compliance on demand: The public's response to specific police requests. Journal of Research in Crime and Delinquency, 33(3), 269-305.

New York Civil Liberties Union. (2012). Stop-and-frisk during the Bloomberg administration. Retrieved from https://www.nyclu.org/sites/default/files/publications/stopandfrisk_briefer_final_0.pdf

Nix, J., & Wolfe, S. E. (2017). The impact of negative publicity on police self-legitimacy. Justice Quarterly, 34(1), 84-108.

Novak, K. J. (2004). Disparity and racial profiling in traffic enforcement. Police Quarterly, 7(1), 65-87.

Reisig, M. D., McCluskey, J. D., Mastrofski, S. D., & Terrill, W. (2004). Suspect disrespect toward the police.

Schulhofer, S. J., Tyler, T. R., & Huq, A. Z. (2011). American policing at a crossroads: Unsustainable policies and the procedural justice alternative. Journal of Criminal Law and Criminology, 101(2), 335-374.

Smith, M. R., & Petrocelli, M. (2001). Racial profiling? A multivariate analysis of police traffic stop data. Police Quarterly, 4(1), 4-27.

Spohn, C., & Beichner, D. (2000). Is preferential treatment of female offenders a thing of the past? A multisite study of gender, race, and imprisonment. Criminal Justice Policy Review, 11(2), 149-184.

Steffensmeier, D., & Allan, E. (1996). Gender and crime: Toward a gendered theory of female offending. Annual Review of Sociology, 22, 459-487.

Steffensmeier, D., & Demuth, S. (2006). Does gender modify the effects of race-ethnicity on criminal sanctioning? Sentences for male and female white, black, and Hispanic defendants. Journal of Quantitative Criminology, 22(3), 241-261.

Statistics Canada. (2021). Immigration and ethnocultural diversity in Canada. Retrieved from https://www12.statcan.gc.ca/nhs-enm/2011/as-sa/99-010-x/99-010-x2011001-eng.cfm

Terrill, W., & Paoline, E. A. (2016). Police use of less lethal force: Does administrative policy matter? Justice Quarterly, 33(2), 336-366.

Wortley, S., & Owusu-Bempah, A. (2011). The usual suspects: Police stop and search practices in Canada. Policing and Society, 21(4), 395-407.