INF 2178 Final Project

Factors that Influence the Situation of Strip Search Conducted by Police in Toronto

---

Master of Information
Faculty of Information
INF2178H: Experimental Design for Data Science
Professor Shion Guha

Colab:https://colab.research.google.com/drive/185ncsyNdKeTzRf7FerxrsaQQ5f2yXIuT?usp=sharing

By Group 55:
Jiaqi Han(peterjiaqihan@mail.utoronto.ca)
Jia Zeng(jiazeng.zeng@mail.utoronto.ca)

Table of Contents

1. Introduction

1.1 Crime in Toronto and Strip Search

As the largest city in Canada, Toronto is not immune to different kinds of crime, and the situation has tended to become worse in the past two years. According to the Toronto Police Service data portal, in 2022 the city's overall major crime increased by 17.2 per cent, with different percentages of increases in different kinds of crimes [2]. For instance, there was a 28.5 percent of increase in robbery, and 44.2 percent of increase in auto thefts [2]. Many of the major crime indicators (MCI) were increasing according to the data released by Toronto Police Service [6]. Meanwhile, several violent crimes shocked the residents in Toronto at the end of year 2022, and some people started to question and complain about Toronto's safety issues since they kept seeing news reports about the new crimes [2]. Moreover, another argument points out that the issue of increasing crime doesn't only appear in Toronto, instead it is more like a national problem which has various causes and catalysts [3]. Thus, it is important to study the data about crime records in the past two years to determine patterns and relationships among the factors which could bring us more knowledge to understand this social issue.

Strip searching is a kind of contemporary practice of police and security institutions. While doing a strip search, the officer will have total control over a person. This practice is critiqued by people, which makes it a controversial topic.

In this project, the dataset used was posted by Toronto Police Service. It is from the public safety data portal, which records information regarding all arrests and strip searches. Some factors recorded, such as the gender, age groups, race, actions at arrest and the reasons of being searched can be useful to study. Detailed information about the dataset will be provided in the Methods section of this paper. Two research questions were selected as the questions to explore with the data, and more information will be introduced in Section 1.3. The dataset was cleaned to make it suitable for further research, and exploratory data analysis was conducted to provide a first look of all the data, and give information about descriptive statistics. Moreover, hypotheses for the t-tests were set and examined by the tests to

determine whether significant difference exists between the means of various groups. Finally, the data was further analyzed by a logistic regression test and an ANCOVA test to answer the research questions.

1.2 Literature Review

The government of Toronto also released some reports about violent crime classified by the factors mentioned above. In one attachment released by the government website, trends in police-reported violent crime by different factors were recorded. For instance, one of the factors is age group. The report shows that, from 2009 to 2017, the rate of firearm-related violent crime was always higher for youth between 12 and 17 years old than the rate for adults (older than 18) [1]. It indicates that the rate of certain kinds of violent crime may differ by whether the suspects are youth under 18 years old. Thus, this factor is worthy to be further explored. Another point that is related to this dataset is about strip search. Strip search is a kind of search conducted by the police, and the process may involve removing some or all clothing, and doing visual inspections of the body [5]. According to Toronto Police Service's report, in 2020 about 53.5% of arrests led to a booking, and about 42% of these bookings finally caused a strip search [4]. This ratio was not low, which means the number of strip search cases could be high. The report also mentioned that the ratios of strip searches were different based on some demographic factors (e.g. 46% were perceived as White and 31% were perceived as Black) which might contain the issue of over-representation [4].

In Canada. Strip searching means when an arrested or detained person is compelled by police officers to expose their most intimate body areas. The police defined it as "the removal or rearrangement of some or all of the clothing of a person" so that they can do a visual inspection of that person's body areas, even if they are private[7]. It may make people worry about the abuse of strip searching. Although it's under the supervision of the police, it may be conducted without the consent of the person who is being searched[7]. Therefore, strip search would be an interesting topic to focus on.

1.3 Research Question of Interest

This study focuses on two research questions. Based on the information provided by the dataset and knowledge we retrieved from our literature review, we focused on how some factors of the people arrested would affect the situation of their strip searches conducted by police officers after they are arrested. Thus, we included the following two research questions.

Research Question of Interest 1 - used for Logistic Regression analysis - What is the association between sex and the likelihood of finding an item during a strip search, and how does the search reason of assisting with escape impact this association among individuals arrested in the city of Toronto in 2020 and 2021?

Research Question of Interest 2 - used for ANCOVA - Does being cooperative or not when being arrested(independent variable) have a significant effect on the number of reasons of being strip searched after people are arrested(dependent variable), after controlling for the effect of age (covariate)?

2. Exploratory Data Analysis

2.0 About the Raw Dataset

The dataset was released by Toronto Police Service's Public Safety Data Portal in a CSV format. The dataset contains 65,275 entries and 25 columns. Each entry is a record of a person arrested, and the columns contain information about the time (year and season) of each arrest, and the arrested people's identity information and arrest record details (strip search records, booking, actions at arrest, reasons for being searched, items found, and object ID).

2.1 EDA for Logistic Regression

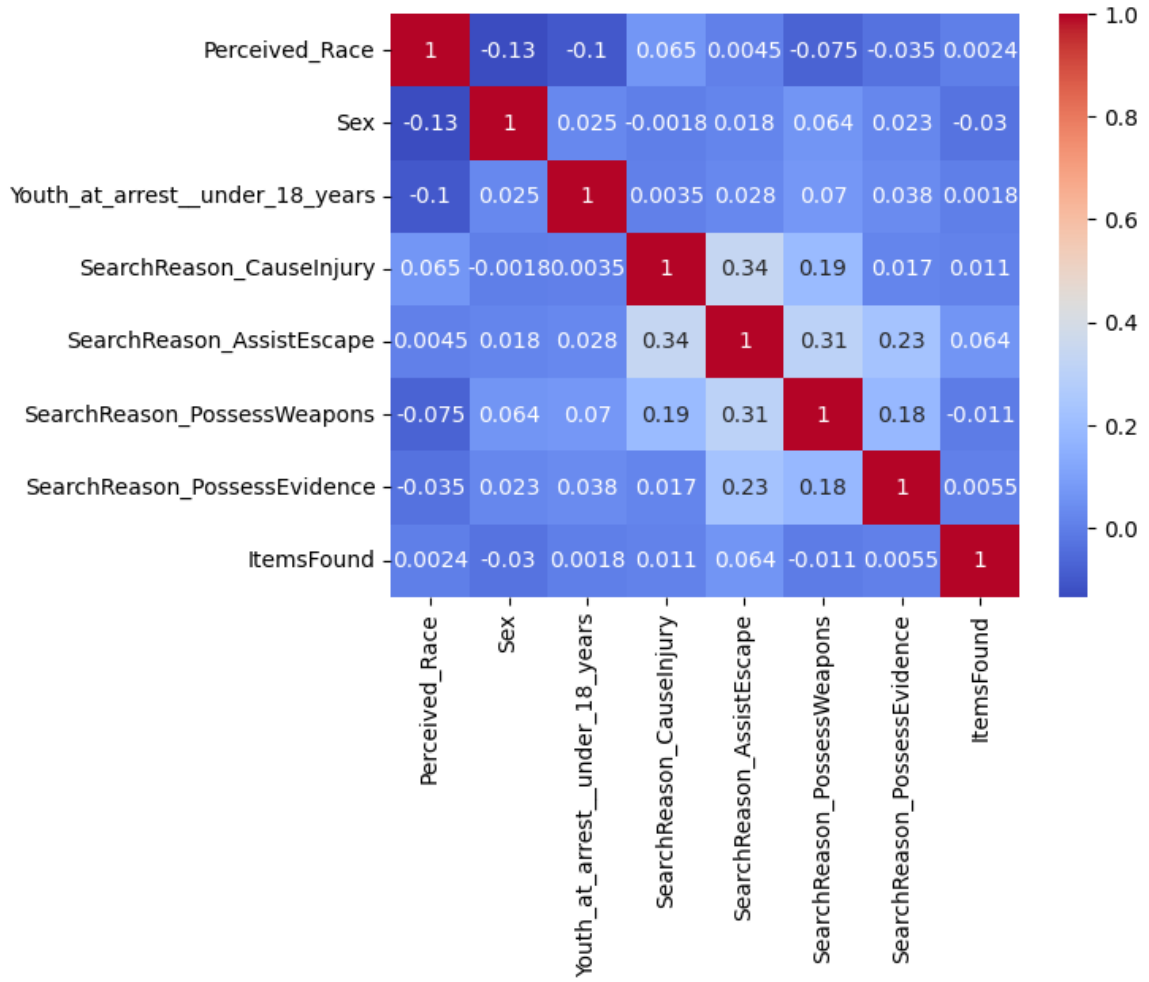
2.1.1 Dataset Clean and Description

Similar to our ANCOVA test, our logistic regression still uses the dataset released by Toronto Police Service's Public Safety Data Portal. However, the dataset cleaning process is different, since our logistic regression requires variables different from the ANCOVA.

We aim to investigate the underlying reasons and relationships that contribute to the occurrence of "ItemsFound." This term refers to the discovery of any dangerous items during a strip search of an arrested individual. Thus, we establish a condition where 'StripSearch' equals 1, indicating that the strip search was successful. To identify potential causes, the first step involves constructing a correlation diagram. Based on conventional wisdom, we have chosen 'Perceived_Race,' 'Sex,' 'Youth_at_arrest__under_18_years,' 'SearchReason_CauseInjury,' 'SearchReason_AssistEscape,' 'SearchReason_PossessWeapons,' and 'SearchReason_PossessEvidence' as independent variables.

Prior to constructing the correlation diagram, it is essential to cleanse the relevant column in the dataset. Given the sufficiently large sample size, all blank and unidentified variables are eliminated, all categorical independent variables are signed to a numerical number, there is no change for the dependent variable since it is already a dummy variable with the satisfaction of the logistic regression condition, and then we built the correlation matrix.

The correlation matrix shows below:



In determining the strength of correlation, a value closer to 1 indicates a stronger relationship. Consequently, the variables 'Sex' and 'SearchReason_AssistEscape' have emerged as the most potent causal factors in the row labeled "ItemsFound." As a result, we have decided to utilize 'Sex' and 'SearchReason_AssistEscape' as independent variables.

Now the sample size is 7264. There is a necessary assumption that should be checked here (also included in 2.1.2) to ensure the proper sample size can be used for a logistic regression - a balanced dataset.

ItemsFound	probability
0 = failure = no item found	0.626101
1 = success = there is an item found	0.373899

From the results above, this tells us that the 'ItemsFound' column is imbalanced, as 62.6% of the values are 0 and only 37.4% are 1. Class imbalance can pose challenges for machine

learning models, as they may be biased towards the majority class and perform poorly on the minority class. To handle class imbalance, we use SMOTE to achieve normalization. We generate new samples of the minority class using SMOTE and return the resampled training data with an equal number of positive and negative examples. Now our dataset is ready for logistic regression analysis.

2.1.2 Logistic Regression Assumption Check

- Balanced Sample Set

  We already checked the condition in 2.1.1, the revised sample set satisfied the balanced condition.

- Dependent Variable (DV) is Binary

  We choose “ItemsFound” as our DV, which only contains 0 or 1, therefore, it is a dummy variable that satisfies the condition.

- Independence of observations

  We assume that the observations are sampled independently.

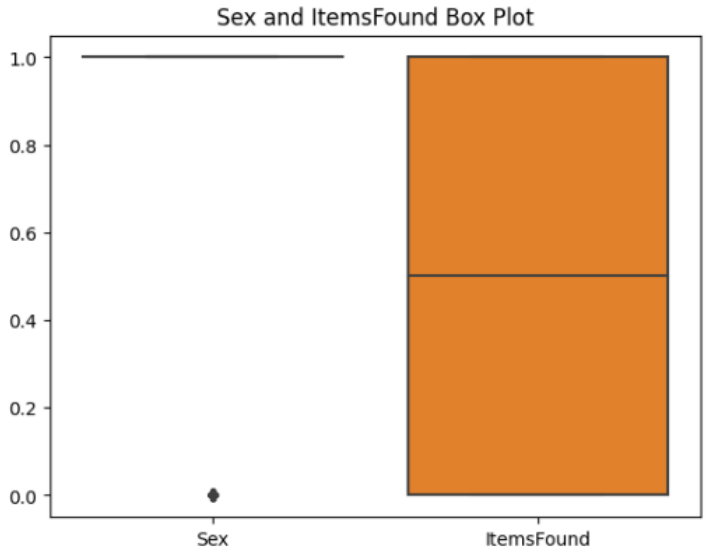
- Assumption of Continuous IVs being Linearly Related to the Log Odds

  Our IVs are all categorical.

- Assumption of Absence of Multicollinearity

  We checked the correlation between possible IVs in 2.1.1.

- Assumption of Lack of outliers



There is no evidence that could conclude the variables satisfied the assumption of lack of outliers, since most of the sex are male and there are extremely far outliers on the sex.

2.2 EDA for ANCOVA

2.2.1 Dataset Clean and Descriptive statistic data

Our second research question needs to use ‘Actions_at_arrest___Cooperative’, ‘Age_group__at_arrest_’, and the sum of search reasons for every record. Firstly, since we only needed to focus on the people who were strip searched, we cleaned the data by removing people who were not searched, which means removing the rows that contain “0” in the “StripSearch” column. In the raw dataset, there are four columns which record whether the people arrested were strip searched for four different reasons respectively. The columns contain 1 or 0, where 1 means the arrested person was strip searched for this reason, and 0 means the person was not searched for this reason. To calculate the number of search reasons as a continuous variable, we added the four search reasons together and created a new column called "Num_of_Search_Reason" to represent the sum of search reasons. For instance, if a person has ‘1’ for two of the search reason columns, his record in the new "Num_of_Search_Reason" column would be ‘2’. Next, we converted the values in the age group columns to continuous variables by taking the lower bound age of every age group, and created a new column called ‘Age’ to store the data. For example, for the age group ‘Aged 35 to 44 years’, the value in the new column would be 35. Finally, we abstract the column ‘Actions_at_arrest___Cooperative’ from the raw dataset, which contains either ‘0’ or ‘1’. Having ‘1’ in this column represents that the people arrested were cooperative when they were arrested, and having ‘0’ means the arrested people were not cooperative, and might conduct negative behaviours when they were arrested. We made this column to be categorical and let ‘1’ to represent being cooperative and let ‘0’ to represent being uncooperative.

After cleaning the original dataset, we produced a new individual dataset which only contained the variables that were required by the research question

```
RangeIndex: 7801 entries, 0 to 7800
Data columns (total 3 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   Actions_at_arrest___Cooperative  7801 non-null   object
 1   Age                             7801 non-null   int64
 2   Num_of_Search_Reason            7801 non-null   int64
dtypes: int64(2), object(1)
memory usage: 183.0+ KB
```

Figure 2.2.1 Basic Information of the New Dataset for Research Question 2

	Age	Num_of_Search_Reason
count	7801.000000	7801.000000
mean	30.088707	1.981156
std	10.360271	1.209695
min	17.000000	0.000000
25%	25.000000	1.000000
50%	25.000000	2.000000
75%	35.000000	3.000000
max	65.000000	4.000000

Table 2.2.1 Statistic Data for Continuous Variables in the New Dataset for Research Question 2

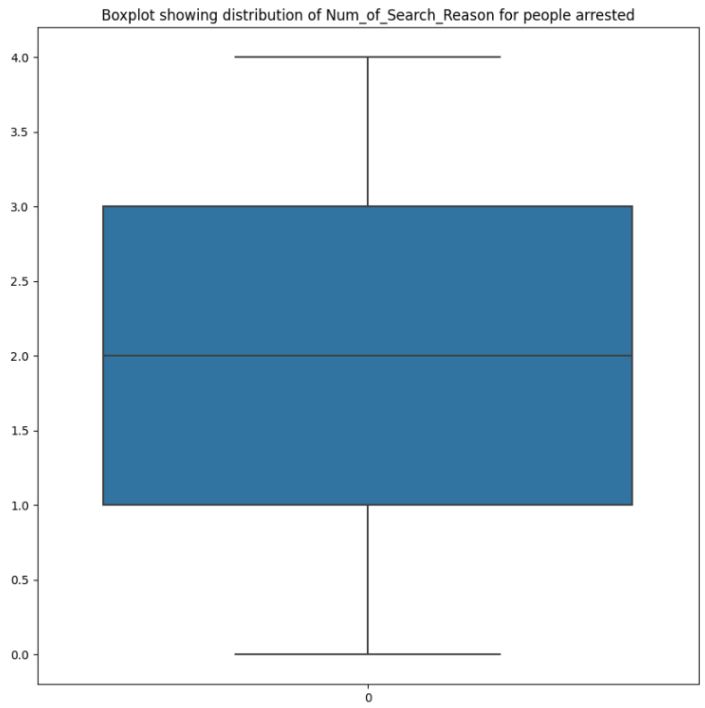


Figure 2.2.2 Boxplot for Number of Search Reasons for People Arrested

According to Figure 2.2.1, Figure 2.2.2, and Table 2.2.1, our new dataset for the second research question contains 7801 rows, and the columns include "Actions_at_arrest___Cooperative," "Age," and "Num_of_Search_Reason." The average number of search reasons is about 1.98, which is almost half of the total number of search reasons. Meanwhile, the 50 percent quantile is two. It may indicate that on average every person who was arrested and strip searched had two reasons for being searched. One insight from this point could be that usually the police would strip search the arrested people for two reasons, and the people being searched usually violated two laws or rules which caused them to be strip searched. However, we can see that the minimum value of search reasons is zero, which means the people arrested were searched without any reason, or their reasons were not recorded by the dataset. Moreover, for the covariate 'Age', the mean of it is about 30, the 50 percent quantile is 25, and the 75 percent quantile is 35. An insight provided by these results could be that for people who were arrested and strip searched, their age were mostly under 35. People who were older than 35 years old might have less chances of being strip searched by police officers, within the time and location limits of the dataset (Year 2020 and 2021, around the city of Toronto area).

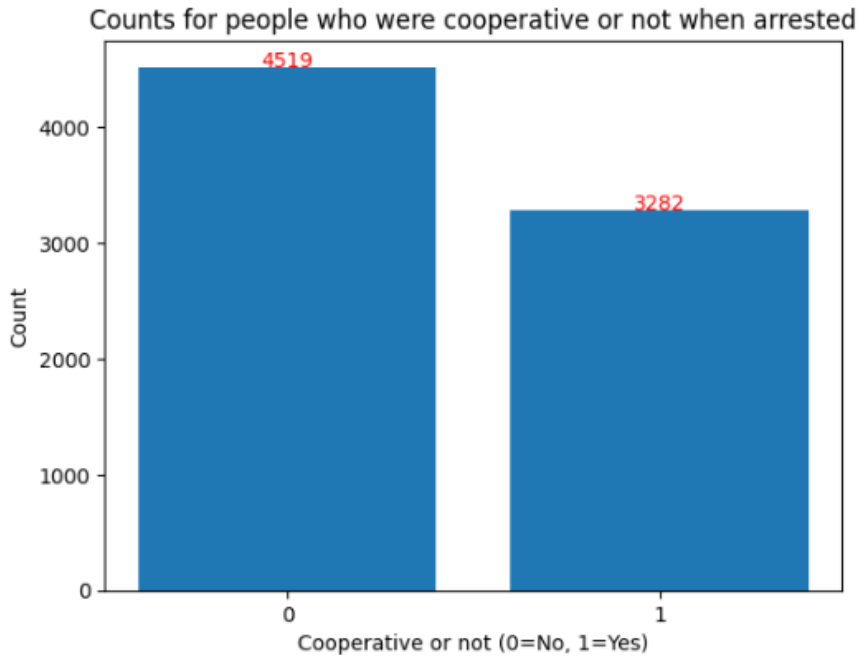


Figure 2.2.3 Counts for People who were Cooperative or Not when Arrested

We also compared the number of strip searched people who were cooperative or not when they were arrested. According to Figure 2.2.3, the number of people who were not cooperative while being arrested (4519) is higher than the number of cooperative people (3282). The difference between them is 1237, and it indicates that more people tended to be uncooperative when they were arrested by police officers.

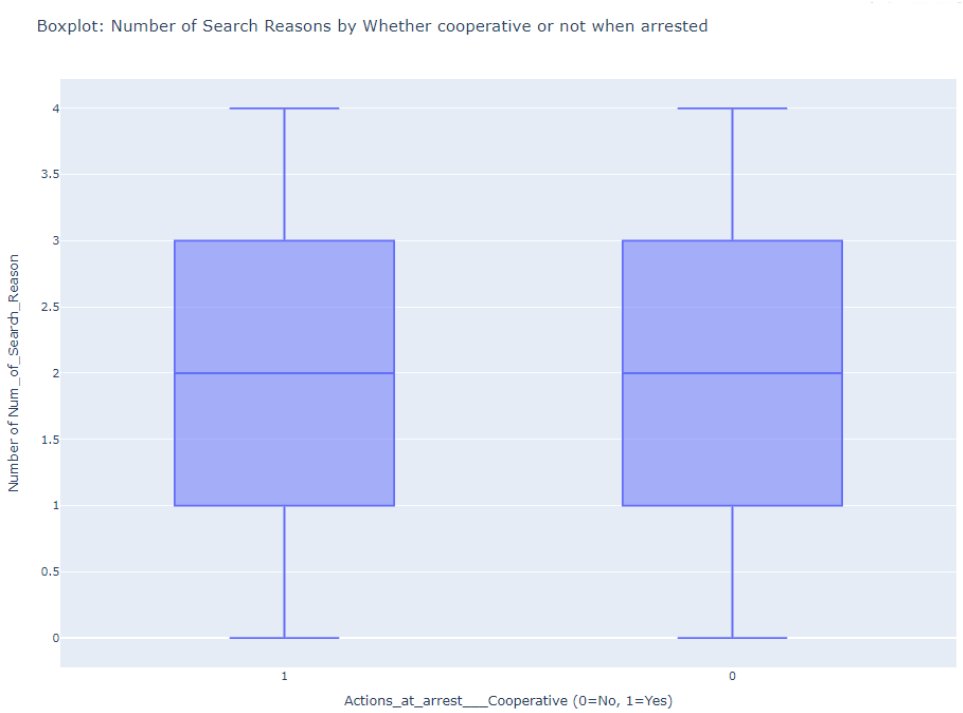


Figure 2.2.4 Boxplot for Number of Search Reasons by Whether being Cooperative or Not when Arrested

Next, we used boxplots to see the statistical results regarding the number of search reasons for cooperative and uncooperative people respectively. An interesting point we found is that, according to Figure 2.2.4, there is no difference between the two groups about search reason amount. Both the boxplots have the same max, min, percent quantiles, and there are no obvious outliers in the graph. Potentially it may represent that based on our new dataset, the number of search reasons for the two groups of people appears to be the same, if we don't consider other factors.

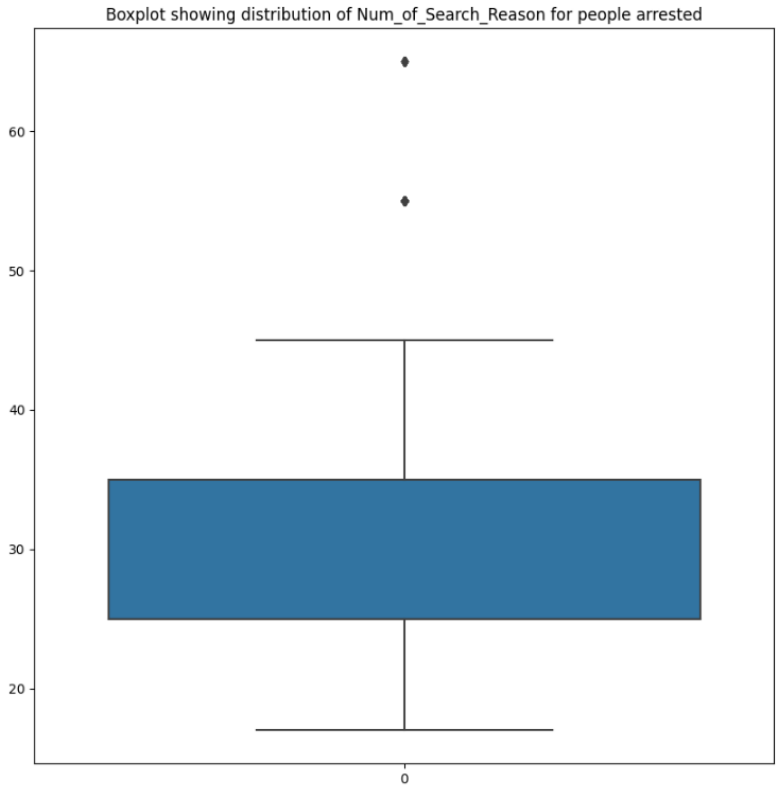


Figure 2.2.5 Boxplot for Age

Figure 2.2.5 is a visualized version of the results provided by Table 2.2.1 regarding the distribution of ages for the people involved in this new dataset. The figure again proved that people who were arrested and strip searched tended to be young or middle-aged for most of the time. Relatively few people were above 45 years old, and there are only two outliers that are above 50 years old.

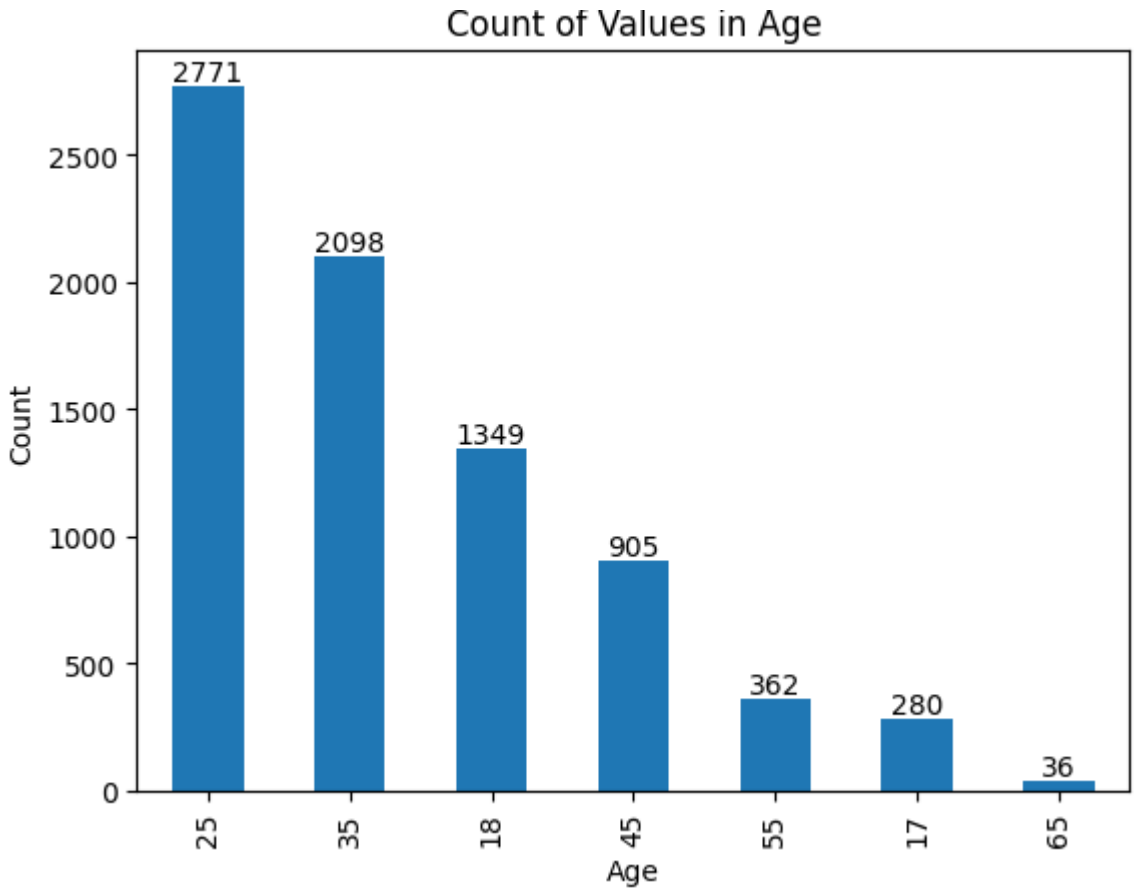


Figure 2.2.6 Bar Chart for Age

To check the number of people arrested and strip searched in different ages, we produced a bar chart to check the counts data. According to FIgure 2.2.6, people in the age of 25 years old is the group with the most counts (2771), followed by people who are 35 years old (2098). Young 18-year-old adults are the group with the third highest counts, followed by middle-aged people who are 45 years old (905). The rest of the groups (age 55, youth who are 17 years old, and elderly people who are 65 years old) have much lower counts (362, 280 and 36 respectively). It proves the results in the previous paragraph which states that people who were arrested and strip searched tended to be young or middle-aged for most of the time.

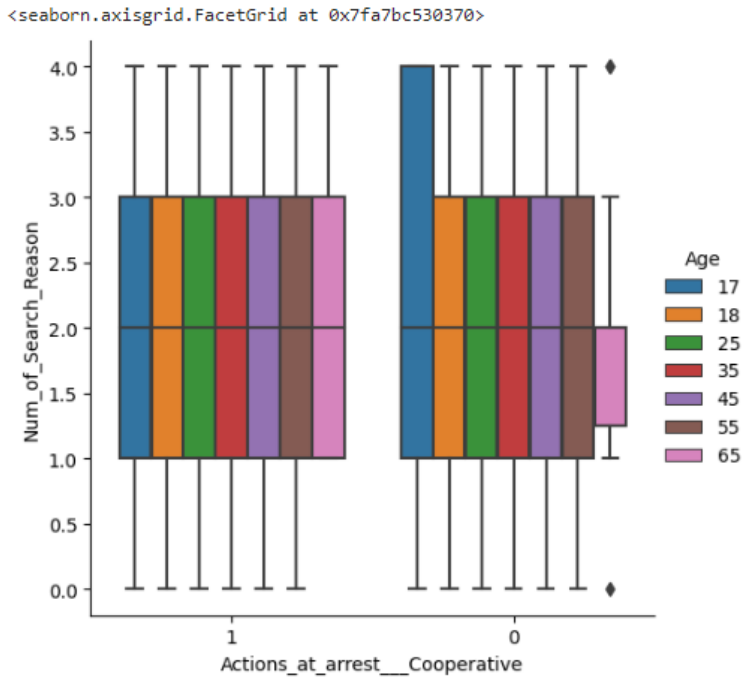


Figure 2.2.7 Boxplot for Number of Search Reason for Coop and Uncoop People Based on Age

To better understand the data on hand, including the covariate, we made side-by-side boxplots to check the situation of the search reason amount for people being cooperative and uncooperative based on their age. Figure 2.2.7 shows that for the two groups of people, the boxplots about search reasons are mostly the same for most of the ages. However, for uncooperative people who are 17 years old, the 75 percent quantile is higher than all other boxplots. For uncooperative people who are 65 years old, the 25 percent quantile is higher, and the 75 percent quantile is lower than all other boxplots. It may indicate that uncooperative young teenagers may be strip searched by the police officers for more reasons, and elderly people may be searched for less reasons.

2.2.2 T-Test

For further exploration of the data, we conducted t-tests for our variables to see if the differences are significant between the levels.

Research Question 2 contains one independent variable, which is whether the arrested people were cooperative or not when they were arrested. A t-test was conducted to see if the means of the outcome variable for each of them are statistically significant. Our null hypothesis is that the average number of reasons for being strip searched by police officers are equal

between cooperative and uncooperative people. Our alternative hypothesis is the average number of reasons for being strip searched for the arrested cooperative and uncooperative people are not equal. The t-test result shows that the p-value is 1.1246060950304406e-1, which is lower than the alpha level (0.05), 95% CI [0.1330, 0.2407]. Thus, we can reject our null hypothesis and conclude that the difference between the two means are statistically significant. The average number of reasons for being strip searched for cooperative people (M = 1.8729, SD = 1.1748) is lower than the number for uncooperative people (M = 2.0597, SD = 1.2286).

3. Methodology

3.1 Logistic Regression

3.1.1 Interest

Research Question of Interest 1: What is the association between sex and the likelihood of finding an item during a strip search, and how does the search reason of assisting with escape impact this association among individuals arrested in the city of Toronto in 2020 and 2021?

3.1.2 Fundamental Elements of Statistics Analysis

- Observational Unit:
  The individual arrest record from the city of Toronto in 2020 and 2021.
- Condition:
  The arrested individual did a strip research.

- DV - Dependent Variable (Categorical & Binary)
  Whether the individual arrest records an item found after a strip search or not from the city of Toronto in 2020 and 2021.
  Suppose DV is a success or a failure represented as 0 or 1.
  1 - success : There is an item found for the individual after a strip search.
  0 - failure : There is not an item found for the individual after a strip search.

- IV - Independent Variable
  Sex with 2 levels - F or M
  (*F=Female, M=Male)
  SearchReason_AssistEscape with 2 levels - 0 or 1

(*0=failure=the assist escape is not the search reason, 1=success=the assist escape is the search reason)

3.1.3 Logistic Regression Model and Interpreting Coefficients

Since we state variables in 3.1.2,

Y|X ~ Bernoulli(π), where π = P(success)

then, E(Y|X) = π.

Logistic Regression Model:

$$\log(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k$$

Odds of a success for A:

$$odds_A = \frac{\pi}{1-\pi}\Big|_{X=A}$$

Odds ratio of a success of A to B:

$$\frac{odds_A}{odds_B} = \frac{\frac{\pi}{1-\pi}|_{X=A}}{\frac{\pi}{1-\pi}|_{X=B}}$$

If X_k increases by 1 unit and when all other variables are held constant, the odds of a success changes by a multiplicative factor of e^{β_k}.

3.1.4 Train Test Split

The formula of the logistic regression is:

ItemsFound ~ Sex + SearchReason_AssistEscape

The summary of the model:

```
Optimization terminated successfully.
         Current function value: 0.690581
         Iterations 4
                           Logit Regression Results
==============================================================================
Dep. Variable:             ItemsFound   No. Observations:                 7362
Model:                          Logit   Df Residuals:                     7359
Method:                           MLE   Df Model:                            2
Date:                Tue, 11 Apr 2023   Pseudo R-squ.:                0.003703
Time:                        00:52:37   Log-Likelihood:                -5084.1
converged:                       True   LL-Null:                       -5102.9
Covariance Type:            nonrobust   LLR p-value:                 6.224e-09
==============================================================================================
                                 coef    std err          z      P>|z|      [0.025      0.975]
----------------------------------------------------------------------------------------------
Intercept                     -0.0360      0.060     -0.597      0.550      -0.154       0.082
Sex                           -0.0803      0.064     -1.264      0.206      -0.205       0.044
SearchReason_AssistEscape      0.2959      0.049      6.030      0.000       0.200       0.392
==============================================================================================
```

Interpretation of the above coefficients:

- The Intercept:

  The intercept is -0.0360, which represents the log odds of finding an item during a strip search for an individual who is male (the reference category for sex) and whose search reason is not assisting with escape (the reference category for search reason).

- Sex:

  The coefficient for sex is -0.0803, indicating that increasing the sex of the individual by one unit (from male to female) decreases the predicted log of the odds of finding an item during a strip search by 0.0803, holding the search reason constant. However, this effect is not statistically significant (p-value = 0.206), indicating that there is no evidence of a relationship between sex and the likelihood of finding an item during a strip search.

- SearchReason_AssistEscape:

  The coefficient for the search reason of assisting with escape is 0.2959, means when holding the sex constant, increasing the search reason of assisting with escape by one unit (from 0 to 1) increases the predicted log of the odds of finding an item during a strip search by 0.2959, which indicates that individuals whose search reason is assisting with escape are more likely to have an item found during a strip search compared to those whose search reason is not assisting with escape.This effect is statistically significant (p-value < 0.001), indicating that there is a relationship between the search reason and the likelihood of finding an item during a strip search.

The pseudo R-squared:

Pseudo R-squared measures the proportion of variation in the dependent variable that is explained by the independent variables in the model. In this case, the pseudo R-squared of 0.003703 indicates that the independent variables of sex and search reason of assisting with escape explain only a small proportion of the variation in the dependent variable of whether or not an item was found during a strip search, suggesting that the model has limited ability to explain the variability in the outcome variable.

Odds Ratio (interpreted in 3.1.3):

Column Name	Odds Ratio
Intercept	0.964646
Sex	0.922851

SearchReason_AssistEscape	1.344319

The logistic regression results provide the estimates of the regression coefficients, standard errors, and p-values for each independent variable. The odds ratio is derived from the logistic regression coefficients and represents the multiplicative effect of each independent variable on the odds of the dependent variable.

In our cases, the logistic regression results show that Sex is not statistically significant, while SearchReason_AssistEscape is statistically significant at a p-value of 0.000. The odds ratio indicates that the odds of finding an item for a female arrestee are 0.92 times the odds of finding an item for a male arrestee, holding the search reason constant. The odds ratio also shows that the odds of finding an item for an individual who was strip-searched for the reason of assisting escape are 1.34 times the odds of finding an item for an individual who was strip-searched for other reasons, holding the sex constant.

Therefore, there is enough evidence for us to conclude that the logistic regression results and the odds ratio are consistent with each other, with the odds ratio providing a more interpretable measure of the effect size of each independent variable on the odds of the dependent variable.

The confidence interval table:

	Lower CI	Upper CI	OR
Intercept	0.857154	1.085618	0.964646
Sex	0.814834	1.045187	0.922851
SearchReason_AssistEscape	1.221056	1.480024	1.344319

The presented confidence interval table provides additional information that corroborates the findings reported in the summary of the logistic regression model and odds ratio table. The table's purpose is to enable the verification of the model's results by providing the uncertainty intervals associated with the estimated coefficients and their corresponding odds ratios.

The prediction interval table:

	mean	mean_se	mean_ci_lower	mean_ci_upper	obs_ci_lower	obs_ci_upper
39825	0.56473	0.016622	0.532146	0.597314	-0.413640	1.543100
42494	0.54476	0.010119	0.524924	0.564596	-0.433268	1.522788
51810	0.56473	0.016622	0.532146	0.597314	-0.413640	1.543100
48818	0.56473	0.016622	0.532146	0.597314	-0.413640	1.543100
625	0.56473	0.016622	0.532146	0.597314	-0.413640	1.543100

Prediction intervals tell us a range of values the target can take for a given record. We can see the lower and upper boundary of the prediction interval from the "mean_ci_lower" and "mean_ci_upper" columns, respectively. Since we choose alpha = 0.05, it is a 95% prediction interval.

3.1.5 Accuracy and Confusion Matrix

The test accuracy is 0.5560908465244322, which means that the model that we established to check whether an item found after a strip search correctly classified 55.61% of the individuals in the test set.
with a confusion matrix:

Confusion Matrix		Actual Value	
		Positive(1)	Negative(0)
Predicted Value	Positive(1)	TP(True Positive): Predicted Positive and it is true. 578	FP(False Positive): Predicted Positive and it is false. 289
	Negative(0)	FN(False Negative): Predicted Negative and it is false. 356	TN(True Negative): Predicted Negative and it is true. 230

Our model train and test the sample set and then identities the following results out of 1023 total cases:

- True Positives (TP): there are 578, the actual number of individuals with an item found after a strip search that were correctly identified as having one.

- False Positives (FP): there are 289, the number of individuals without an item found after a strip search that were incorrectly identified as having one.
- False Negatives (FN): there are 356, the number of individuals with an item found after a strip search that were incorrectly identified as not having one.
- True Negatives (TN): there are 230, the actual number of individuals without an item found after a strip search that were correctly identified as not having one.

Recall:

$$\frac{TP}{TP+FN} = \frac{578}{578+356} = 0.619$$

The proportion of all the actual successful cases that were predicted correctly by the model is 61.9%.

Precision:

$$\frac{TP}{TP+FP} = \frac{578}{578+289} = 0.667$$

The proportion of all predicted positive cases that were predicted correctly by the model is 66.7%.

Accuracy:

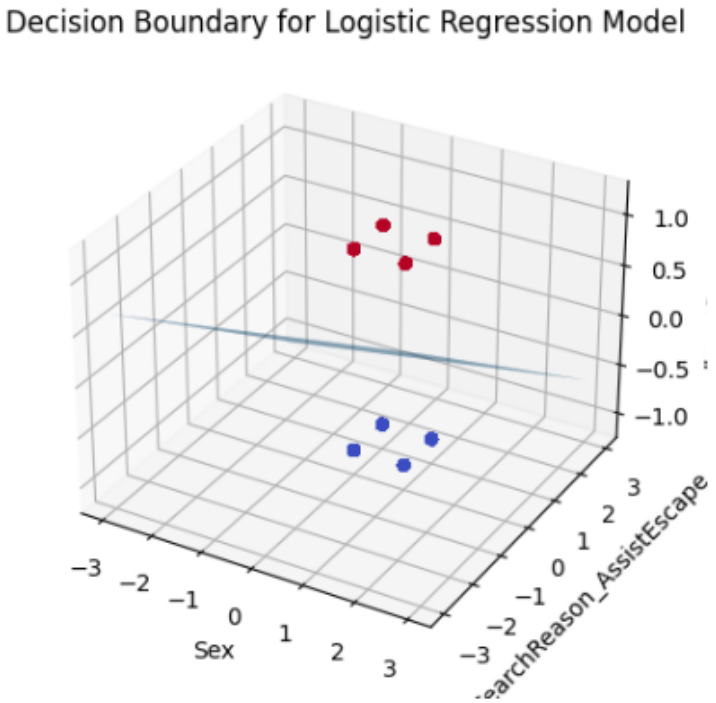
From all positive and negative cases, the proportion of all the model has predicted correctly is 55.61%.

3.1.6 Prediction Interval

Prediction intervals are used to represent the uncertainty in the predicted outcome of a continuous variable. It does not provide a continuous prediction interval as in the case of linear regression. Since all variables in our model are categorical, which cannot be represented as a continuous distribution, prediction intervals cannot be computed for them.

Instead of it, we use a decision boundary graph to plot the data to visualize the results of the classification that our model did. The decision boundary is the boundary that separates the two classes in a binary classification problem. In logistic regression, the decision boundary is a linear equation that separates the data points into two regions based on their predicted class probabilities. Points on one side of the boundary are predicted to belong to one class, while points on the other side are predicted to belong to the other class.

In the 3D plot you created, the decision boundary is represented by the surface that separates the blue and red points. Points above the surface are predicted to belong to the "ItemsFound=1" class, while points below the surface are predicted to belong to the "ItemsFound=0" class.



3.2 ANCOVA

Interest: Does being cooperative or not when being arrested(independent variable) have a significant effect on the number of reasons of being strip searched after people are arrested(dependent variable), after controlling for the effect of age (covariate)?

3.2.1 Power Analysis

Before doing the ANCOVA test, we conducted power analysis which is a statistical technique that can determine the sample size required to detect the significant effect in our study, with a given level of statistical power. Thus, our goals include the effect size, which is the size of the difference in the dependent variable (number of search reasons)'s means between the levels of the independent variable, after controlling for the covariate. We also calculated the sample sizes required for the two levels in our independent variable (cooperative and uncooperative). The alpha level is set to 0.05, and the power is set to 0.8, which is usually used by similar

kinds of studies. After getting the ideal sample sizes, we calculated the statistical power based on our settings for alpha level, effect size, and sample sizes.

Effect Size: 0.1549

Cooperative or Not	Ideal Sample Size	Actual Sample Size	Power
Yes	780	3282	1.0
No	566	4519	1.0

Table 3.2.1 Power Analysis Results

Table 3.2.1 summaries the power analysis results provided by our codes. The effect size given by Cohen's D codes is about 0.1549 (with an alpha level of 0.05 and power of 0.8). According to the results, to get an ideal level of power for the ANCOVA, we need to have 780 samples for people who were cooperative when arrested, and the actual size is 3282, which means we have enough samples for gaining enough power. We need to have 566 samples of people who were not cooperative when arrested, and the actual size is 4519, which means we also have enough samples for this level. The statistical power for both levels achieved 1.0, which means we have enough statistical power to continue the ANCOVA test.

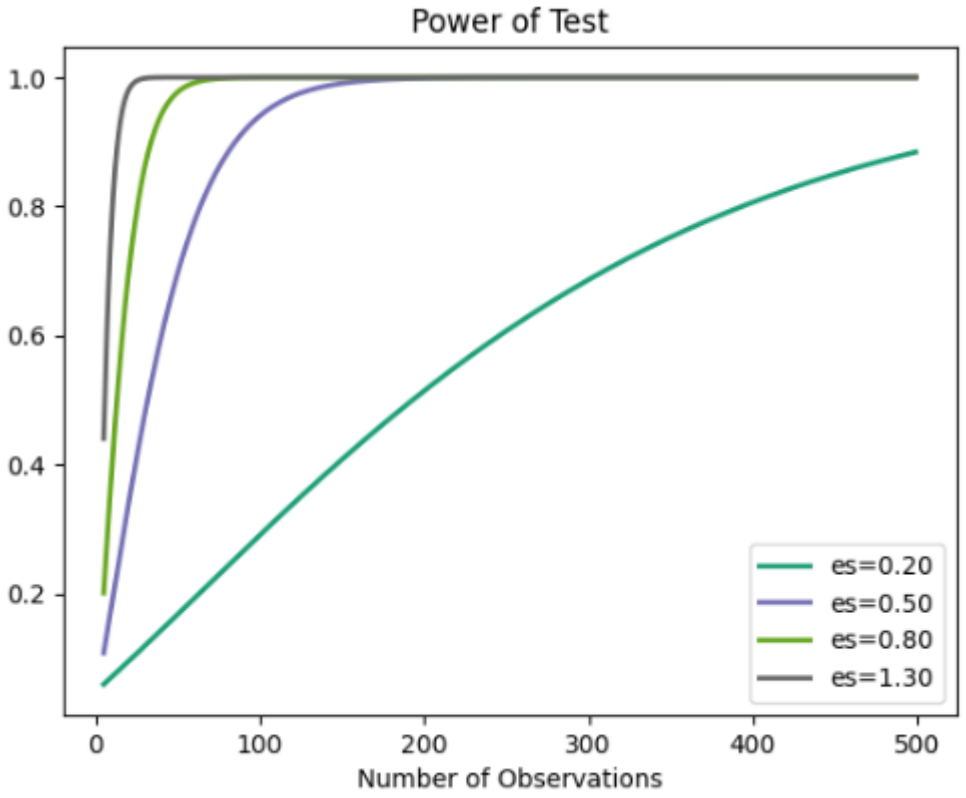


Figure 3.2.1 Power Curve

Based on our power analysis settings, we also created a power curve, which is demonstrated by Figure 3.2.1. According to the graph, lower effect size requires larger sample sizes to gain enough power. For instance, if the effect size is 1.3, it only needs about less than 100 samples to achieve a power that is higher than 0.8. However, if the effect size is 0.2, it may require more than 500 samples to reach the 0.8 standard. In our dataset, there are more than 1000 samples for both Cooperative (3282) and Uncooperative (4519), which indicates that our dataset provides reliable amounts of samples for our study.

3.2.2 ANCOVA Test

In our study, we used Analysis of Covariance (ANCOVA), which is a statistical method used to analyze the relationship between an outcome variable and one or multiple explanatory variables, while controlling for the effects of one or more covariates. It can be used to compare the means of two groups while controlling the effects of another or more continuous variables.

For our second research question, the variable setting is displayed below:

Type	Name	Description
Independent Variable	Actions_at_arrest___Cooperative	Contains 0 or 1. (1 means being cooperative when arrested, 0 means not).
Outcome Variable	Num_of_Search_Reason	Calculated from the sum of the search reasons held by each arrested person
Covariate	Age	Continuous variable converted from the age groups by taking the lower bound value.

Our hypothesis for ANCOVA are listed below:

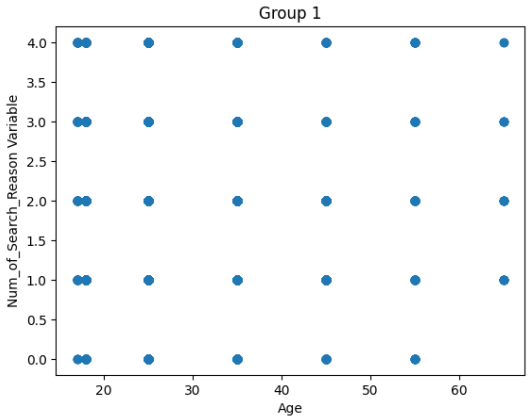
Null Hypothesis: The average number of reasons for being strip searched by police officers are the same for people who are cooperative or uncooperative when they are arrested.

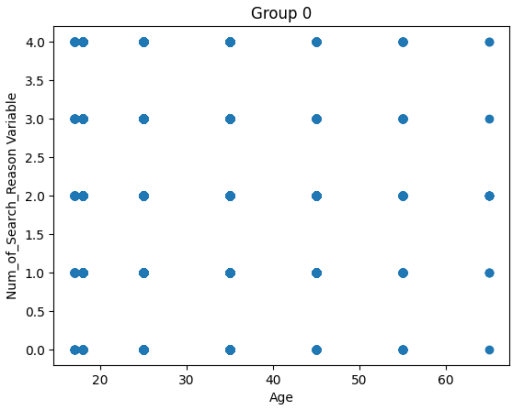
Alternative hypothesis: The average number of reasons for being strip searched by police officers are different for people who are cooperative or uncooperative when they are arrested.

3.2.3 Assumption Check

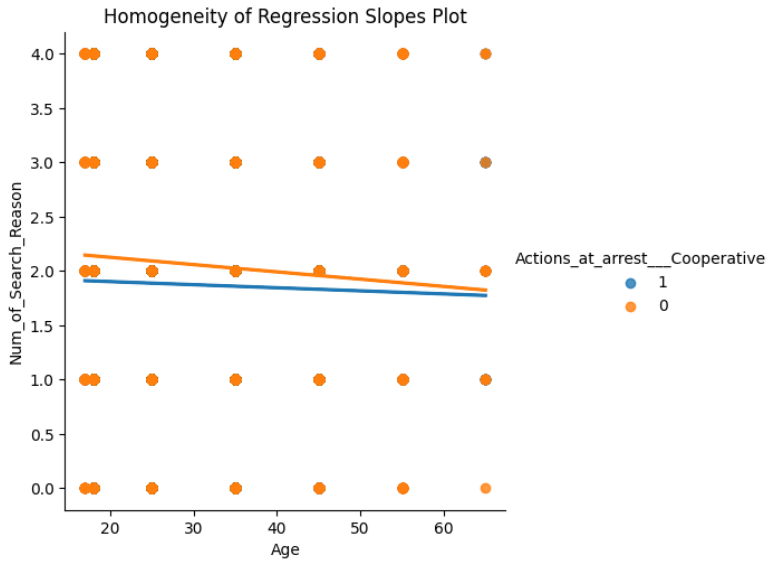
Another important step we did is checking the assumptions for ANCOVA before starting doing the test. It is important to check the assumptions for ANCOVA because violations of these assumptions may lead to biased or inaccurate results. If the assumptions of ANCOVA are not met, the statistical inferences drawn from the analysis may not be valid, and the estimated effects of the independent variable and covariate(s) may be negatively impacted. We checked the following five key assumptions:

1. Normality: For an ANCOVA test, The dependent variable (the outcome variable) should be normally distributed in each group for each level of the covariate. We conducted the Shapiro-Wilk test and got zero for the p-values for both levels in our independent variable. It might indicate that the null hypothesis of normality was rejected. The data may not be normally distributed.
2. Independence: The observations should be independent of each other. We can assume this assumption is satisfied because of the nature of the raw dataset and our data cleaning method. There would be no systematic relationships between them.
3. Homogeneity of variance: The variance of the dependent variable should be equal across all groups and levels of the covariate. We conducted a Levene's test, which produced a p-value of 0.0014. It suggests that this assumption may be violated because the p-value is lower than 0.05, and we can't reject the null hypothesis. The variances may be not equal.
4. Linearity: For each level of the independent variable, there is a linear relationship between the dependent variable and the covariate. We produced two scatter plots for the two groups in the independent variable, and the plots in the graphs are mostly evenly spread. It may indicate that this assumption is met.





5. Homogeneity of regression slopes assumption: relationship between the dependent variable and the covariate is the same across all levels of the independent variable. We produced a slope plot which shows that the two lines are not completely parallel, so we can say this assumption is not perfectly met.



4. Results/Findings

4.1 Logistic Regression

Based on the logistic regression analysis, the coefficient for Sex is -0.0803, indicating that there is no significant relationship between Sex and the likelihood of finding an item during a strip search (p-value = 0.206). However, the coefficient for the search reason of assisting with escape is 0.2959, indicating that individuals whose search reason is assisting with escape are more likely to have an item found during a strip search compared to those whose search reason is not assisting with escape (p-value < 0.001). The pseudo R-squared value of 0.003703 indicates that the model has a limited ability to explain the variability in the outcome variable, and the test accuracy is 55.61%.

Additionally, the model identified 578 true positives, 289 false positives, 356 false negatives, and 230 true negatives. The recall value is 0.619, indicating that 61.9% of all actual

successful cases were correctly predicted by the model, while the accuracy value is 0.556, indicating that the model correctly classified 55.6% of all cases.
For more detailed interpretation, please see section 3.1.4, 3.1.5 and 3.1.6.

4.2 ANCOVA Results

Source	SS	DF	F	p-unc	np2
Actions_at_arrest___Cooperative	64.745462	1	44.574464	2.616441e-11	0.005684
Age	21.106323	1	14.530795	1.389489e-04	0.001860
Residual	11326.779157	7798	NaN	NaN	NaN

Table 4.2.1 ANCOVA Results

Table 4.2.1 shows the results of our ANCOVA test. According to the table, the p-unc (uncorrected p-value) for the first source (Actions_at_arrest___Cooperative) is 2.616441e-11, which is significantly smaller than the alpha level (0.05). It indicates that whether being cooperative or not while getting arrested has a significant effect on the amount of reasons for being strip searched after the arrest. The p-unc value for the second source (Age) is 1.389489e-04, which is also lower than the alpha level. It suggests that the covariate ‘Age’ has a significant effect on the amount of reasons for being strip searched. Thus, we can reject the null hypothesis. In general, differences in the number of strip search reasons between cooperative and uncooperative arrested people are significantly different, while controlling for the effect of their ages as the covariate.

5. Discussion

5.1 Logistic Regression

The use of a logistic regression model to analyze the relationship between sex, search reason, and the likelihood of finding an item during a strip search may be limited by several factors.

1. All the independent variables are categorical, which may reduce the model's ability to capture the complexity of the relationship between the variables.

2. The correlation between the dependent variable (items found) and the independent variables (sex and search reason) is weak, which means that the model may not have strong predictive power.
3. The sample is imbalanced, which can lead to biased results. To address this issue, SMOTE was used to generate synthetic samples, which may have introduced some limitations, such as overfitting or underestimation of the variance.
4. The presence of extreme outliers in the sample for the sex variable may impact the accuracy of the results and their generalizability to the population.

Overall, these limitations suggest that the results of the logistic regression model should be interpreted with caution and may not fully represent the association between sex, search reason, and the likelihood of finding an item during a strip search in the city of Toronto in 2020 and 2021. Further research using a larger and more representative sample and more sophisticated statistical methods may be needed to better understand this relationship.
To be able to address the limitation posed by the imbalanced sample set, it is recommended to explore additional data sources, such as resurvey and re-administrative data, to augment the existing database and improve its sufficiency for analysis.
Given the categorical nature of all variables and the presence of outliers, logistic regression may not be the optimal model for analysis. Alternative classification models could be considered viable options to overcome the limitations of the logistic regression model. To be able to determine a suitable model that fits the parameters of the database, further exploration and experimentation are required.

5.2 ANCOVA Test

Our study determines that having age as the covariate, people will be strip searched for different amounts of reasons based on whether they choose to cooperate with the police officers or not when they are arrested. It makes sense since police officers may pay more attention to the suspects that are not cooperative and even have negative behaviours such as escaping and resisting. In our ANCOVA study, although we successfully determined the influence and effects that the independent variable has on the outcome variable, the results are threatened by several limitations. For instance, our assumption check section identified the violation of some of the assumptions for ANCOVA, such as Normality and Homogeneity

of Variance. These violations may cause some issues such as decreased statistical power of our study, inaccurate results and estimates, and more chances of having errors.
To reduce the effects of these issues, we need to implement different solutions such as using different kinds of tests and regressions, increasing the size of our samples, and transforming some of the data if necessary. It is important to realize that no tests and studies can be perfect since the raw data from the real world can be very different, but we are supposed to make our methods to be as reasonable and perfect as possible.

6. Conclusion

6.1 Logistic Regression

The study aimed to investigate the association between sex, search reason, and the probability of finding an item during a strip search of individuals arrested in Toronto in 2020 and 2021. The results showed that individuals who were searched for the reason of assisting with escape were more likely to have an item found during a strip search compared to those who were not searched for this reason. However, there was no evidence of a relationship between sex and the probability of finding an item during a strip search.
We established a model to identify whether there is an item found for an individual after a strip search, and its accuracy is 55.61%.
The model used to analyze the data had some limitations, such as only using categorical variables, having weak predictive power, and potentially being impacted by imbalanced data and extreme outliers. Therefore, the findings should be interpreted with caution and may not fully represent the actual association in the population.

6.2 ANCOVA Test

The study aimed to investigate whether being cooperative or not when being arrested by police officers in Toronto have a significant effect on the number of reasons for being strip searched after people are arrested, controlling for the effect of age. The uncorrected p-value for the independent variable is 2.616441e-11, which is lower than the alpha level (0.05). We cleaned the data for the test, conducted EDA, t-test, and also power analysis, and finally got the results suggesting that people who are cooperative may have different average number of reasons for being strip searched by police officers with people who choose to be uncooperative. One insight from it is that it would be interesting to explore Toronto's police

system's policies about conducting strip searches when the officers are facing different kinds of situations, and it can also be associated with statistical study about suspects' behaviours when they are facing arrests. However, the violations of some assumptions for ANCOVA suggest that we need to treat the results carefully, and more studies need to be conducted to explore further about the research question.

7. References

[1] City of Toronto. N.d. Attachment 1: Community Violence in Toronto - Trends and Selected Demographic Patterns (pp. 9). Retrieved from https://www.toronto.ca/legdocs/mmis/2019/hl/bgrd/backgroundfile-139316.pdf

[2] Landau, J. December 2022. Toronto residents questioning the city's safety amid wave of shocking crimes. blogTO. Retrieved from https://www.blogto.com/city/2022/12/crime-wave-toronto-questioning-safety/

[3] Russ, G. February 16, 2023. It's not just Toronto. Violent crime is a national problem. Thehub. Retrieved from https://thehub.ca/2023-02-16/its-not-just-toronto-violent-crime-is-a-national-problem/

[4] Toronto Police Service. n.d. Strip Searches: Measurement & Outcomes. Toronto Police Service. Retrieved from https://www.tps.ca/media/filer_public/a6/67/a667e545-ac03-49ac-b6d4-2ae0eca30efd/92bbc7d7-c866-4d04-9a0c-aba55b1a8e01.pdf

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

[6] Toronto Police Service. 2023. TPS Crime Statistics – Major Crime Indicators. Toronto Police Service Public Safety Data Portal. Retrieved from https://data.torontopolice.on.ca/pages/major-crime-indicators

[7] Lemke M. July 6, 2022. Policing Toronto: Strip Searching in a Divided City. The Bullet. Retrieved from https://socialistproject.ca/2022/07/policing-toronto-strip-searching-in-a-divided-city/