Data Analysis on Arrests and Strip Searches

Chenyang Huan

Yueran Yang

INF2178H LEC 0101

Experimental Design for Data Science

Shion Guha

February 27, 2023

Introduction

In Canada, crime and violence are consistently considered as one of the top-of-mind concerns among Canadians, with only the unemployment rate being a greater concern (Government of Canada, 2022). Therefore, the government is devoted to addressing and preventing crime and violence problems as these are widely recognized as significant priorities. Nevertheless, when it comes to policing practices, some critics argue that police officers exhibit a higher propensity to arrest Black civilians than White civilians, especially for minor offences, when faced with evidence of illegal activity (Wortley & Jung, 2020). According to the British Journal of Criminology, striking disparities are evident in relation to ethnicity regarding the rate of strip research (Newburn & Shiner & Hayman, 2004). In light of the importance of equality in the criminal justice system, we will examine the relationships between different race groups and strip searches in our project.

Since arrests and searches are related to a wide range of criminal offences, analyzing the "Arrests and Strip Searches" dataset could provide valuable insights related to crime and violence in the city of Toronto. By analyzing the "Arrests and Strip Searches" dataset, we can identify potential biases or disparities in the criminal justice system, for example, whether a particular racial group are more likely to be arrested or whether certain racial groups are more likely to be targeted for strip searches. We hypothesize that the analysis of the "Arrests and Strip Searches" dataset will help policymakers and law enforcement agencies determine how to enhance fairness in law enforcement practices and equality in criminal justice system. For our analysis, we will use the 2020-2021 arrests and strip searches dataset provided by Toronto Police Service.

Literature Review

For decades, there have been calls to acknowledge and address systemic racism in policing. According to a study, Canadian official crime statistics reflect the fact that Black people have a higher level of engagement in criminal activity compared to other racial groups. They are particularly over-represented in drug-related offences, gang activity and street-level violence. The study suggests that the disproportionate deployment of police in Black communities, resulting in Black people being subject to higher levels of police surveillance compared to other racial groups.

Due to the strict surveillance, Black offenders are more likely to be identified and arrested in comparison to other offenders (Wortley & Jung, 2020).

Through an analysis of over 10,000 drug possession arrests conducted by the Toronto Police Service from 1996 to 2001, another study revealed that Black individuals (38%) were significantly more likely than White people (23%) to be taken to the police station for drug possession (Rankin et al, 2002).

In Particular, concerns have been raised regarding the role of race in strip search decisions. Such concerns center on whether a disproportionate number of racialized individuals are being subjected to strip searches, and whether race impacts the decision-making process for conducting a strip search (McNeilly, 2019). The "Breaking the Golden Rule: A Review of Police Strip Searches in Ontario" report indicates that, according to the Ontario Independent Police Review Director, African Canadians and Aboriginal people are more susceptible to being arrested and subjected to strip searches at a disproportionately higher rate than other Canadians, due to their overrepresentation in the criminal justice system (McNeilly, 2019).

Statistics Canada's 2016 Census Profile shows that in Toronto, just 9% of the population self-identifies as black, while across Ontario, only 2.8% of the population identifies as Indigenous (Statistics Canada, 2017). However, based on the race-related statistics provided by the Ontario Provincial Police and Toronto Police Service, Aboriginal people accounted for 14.58% of individuals subjected to strip searches in Ontario, and in Toronto, 27.67% of strip search cases involved black individuals (McNeilly, 2019).

Moreover, a report from the Toronto Police Service on strip searches found that Indigenous people who were arrested were overrepresented in strip searches by a factor of 1.3 times when comparing their proportion of arrests to their proportion in strip searches. Additionally, when people were arrested for crimes against individuals such as assault, Indigenous, Latino, Middle Eastern, and South Asian individuals were found to be overrepresented in strip searches. Similarly, for offenses related to weapons, Black, Latino, and Middle Eastern individuals were also found to be overrepresented in strip searches (Toronto Police Service, 2020).

Furthermore, strip searches are usually conducted by arresting officers when they have reasonable grounds to believe that the arrested person is uncooperative or concealing something that poses a threat to their safety, for example, a weapon. If the arrested person is resistant to cooperate, strip searches may be conducted with force under necessary circumstances. Based on the research, force was noted in less than 2% of strip searches, and in those cases, the arrested person was belligerent and refused to comply with the search (McNeilly, 2019).

Research Objective and Questions

Our study will seek to examine how the perceived race of the arrested people, the action on arrest is cooperative or not, and their interaction effect, impact the number of arrests and strip search conducted.

Our project aims to investigate the research questions that we have formulated based on our literature review and initial examination of the dataset. The section on descriptive Statistics and T-tests below provides an overview of our preliminary exploration of the data.

RQ1: Is there any difference in the mean number of arrests across different races?

RQ2: Does perceived race and action on arrest (cooperative or not) affect the number of arrests?

RQ3: Does perceived race and action on arrest (cooperative or not) affect strip searches?

RQ1 would be investigated via one-way ANOVA while RQ2 and RQ3 would be examined through two-way ANOVA.

Exploratory Data Analysis

Table1: Sample of Arrests and Strip Searches Dataset

	Arrest_Year	Arrest_Month	EventID	ArrestID	PersonID	Perceived_Race	Sex	Age_group__at_arrest_	Youth_at_arrest__under_18_years
0	2020	July-Sept	1005907	6017884.0	326622	White	M	Aged 35 to 44 years	Not a youth
1	2020	July-Sept	1014562	6056669.0	326622	White	M	Aged 35 to 44 years	Not a youth
2	2020	Oct-Dec	1029922	6057065.0	326622	Unknown or Legacy	M	Aged 35 to 44 years	Not a youth
3	2021	Jan-Mar	1052190	6029059.0	327535	Black	M	Aged 25 to 34 years	Not a youth
4	2021	Jan-Mar	1015512	6040372.0	327535	South Asian	M	Aged 25 to 34 years	Not a youth

5 rows × 25 columns

This table displays a sample of the first five rows of the Arrests and Strip Searches dataset (Table 1). In order to ensure the accuracy and reliability of the dataset, we performed several data cleaning procedures to prepare the data for analysis.

To begin with, we conducted an initial check of the dataset to identify any missing values. We found that there are some missing values in "ArrestID", "Perceived_Race" "Age_group__at_arrest_", and "Occurrence_Category". Of particular concern were the 4 variables related to SearchReason, which contained over 57,000 missing values. In light of this significant amount of missing values, we would exclude those four variables from our analysis. Subsequently, we removed all the missing values in "ArrestID", "Perceived_Race" "Age_group__at_arrest_", and "Occurrence_Category" and filled all null values in the 4 RearchReason variables with 0 to eliminate the potential impact of these variables on our analysis.

In addition, we split the two-year timeframe (2020-2021) into eight quarters, each representing a distinct three-month interval. Although the time periods are discrete, the number of arrests in each quarter can be treated as a continuous variable for our subsequent analysis. After that, we created a dictionary to map the values in the "Occurrence_Category" to standardized categories and replace the original values in the column with their corresponding standardized categories. For instance, we replaced "Police Category - Administrative" and "Police Category - Incident" with "police". By combining those values into a single standardized category, we can simplify our exploratory data analysis and improve the consistency of the data.

Descriptive Statistics

To gain a better understanding of the dataset, we initially grouped the number of arrests by sex. First, we created a new DataFrame by selecting the columns 'Arrest_Year', 'Arrest_Month', 'Sex', 'StripSearch', and 'ObjectId'. We then applied an aggregate function to count the number of values of the 'ObjectId' column for each group. As a result, we produced the following count chart showing the number of arrests grouped by female, male and unknown sex (Figure 1).

Figure 1: Number of Arrests by Sex

Number of Arrests by Sex
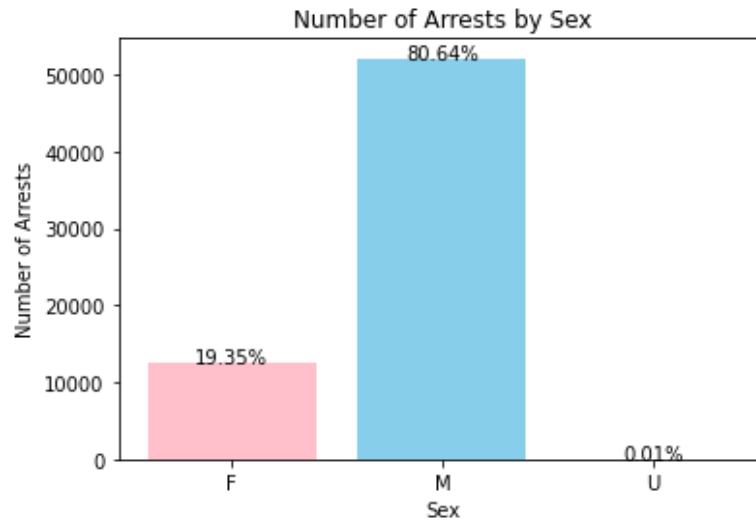
Since the y-axis of this figure was scaled in increments of 10,000, it is difficult to discern the number of arrests for unknown sex (9) from the chart. The reason for the y-axis scale is that the number of male (52,106) and female (12,500) arrests is significantly higher than the number of arrests for individuals with unknown sex. Therefore, we added the percentage for each sex group. As shown, 80.64% of the arrested individuals were male, 19.35% were female, and the unknown sex group accounted for only 0.01%.

By grouping the data by strip search status, we generated a bar chart to visualize the number of arrested individuals who were subjected to strip search (Figure 2).

Figure 2: Number of Arrests by Strip Search Status (Yes or No)

Number of Arrests by Strip Search Status (Yes or No)
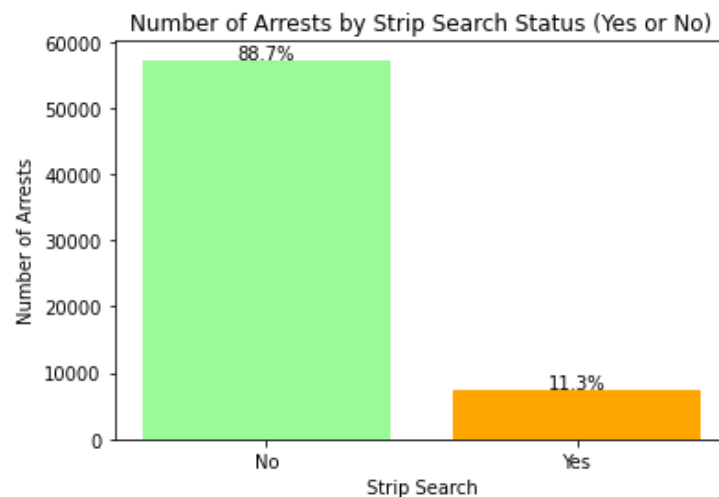
As we can see from the bar chart above, 11.3% of the arrested individuals were subjected to the strip search. Based on the dataset, we found that 1,208 females and 6,123 males were subjected to a strip search after being arrested, while no people from unknown sex group were subjected to a strip search. The chart below provides an overview of the number of arrests based on the sex and strip search status (Figure 3).

Figure 3: Number of Arrests by Sex and Strip Search Status
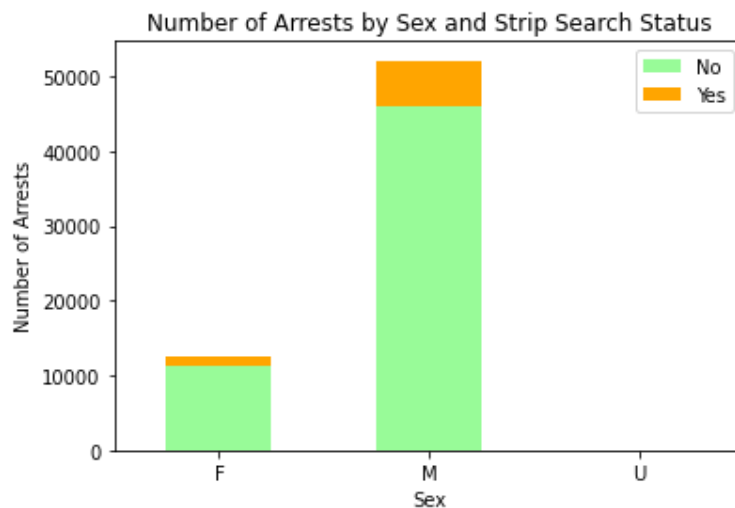
Similarly, we assigned the 8 different race groups with numbers 1 to 8 and created the following bar chart to illustrate the number of arrests by perceived race (Figure 4).

Figure 4: Number of Arrests by Perceived Race
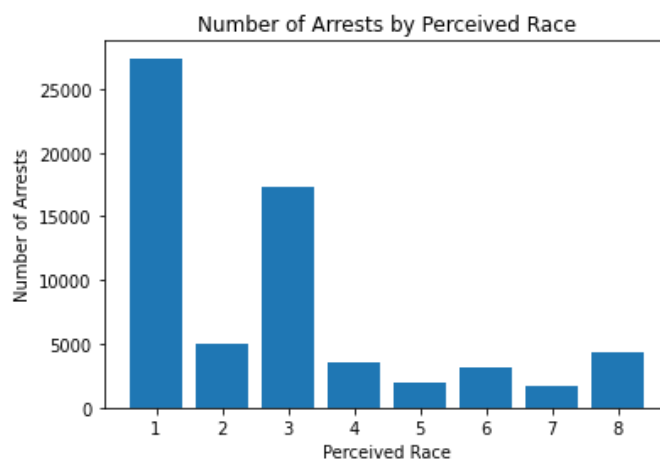
The above graph displays the perceived race categories and their corresponding numerical labels: 1 represents White, 2 represents Unknown or Legacy, 3 represents Black, 4 represents South Asian, 5 represents Indigenous, 6 represents Middle Eastern, 7 represents Latino, and 8 represents East/Southeast Asian. Based on the graph, it is evident that White and Black are the racial groups with the highest number of arrests, while Latino are arrested the least.

In addition, we grouped the data based on the arrested people's perceived race and whether they were cooperative during the arrest. After that, we produced a bar chart with the eight race groups as the x-axis, and the number of arrests as the y-axis (Figure 5).

Figure 5: Number of Arrests by Perceived Race and Actions at Arrest_Cooperative
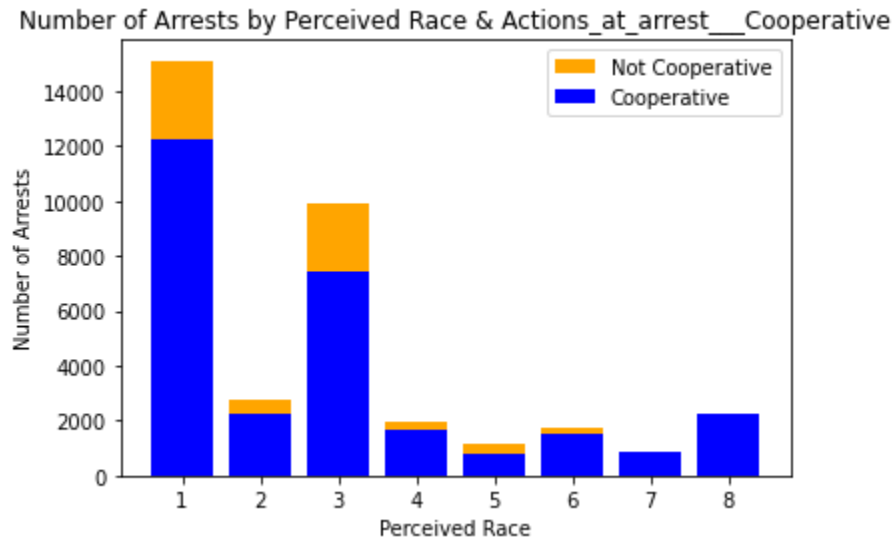
As shown, the plot displays the sum of arrests grouped by perceived race and denotes whether they cooperate or not respectively. We created a table that shows the total arrests of each race and their cooperation rate during the arrests (Table 2). We found that East/Southeast Asian has the highest cooperation rate (51.66%) while Indigenous people have the lowest cooperation rate (39.75%).

Table 2: Percentage of People Who Were Cooperative During the Arrests for Each Race Group

Perceived Race	Total Arrests	Actions_at_arrest___Cooperative	Actions_at_arrest___Cooperative %
Black	17352	7453	42.95
East/Southeast Asian	4388	2267	51.66
Indigenous	1907	758	39.75
Latino	1752	892	50.91
Middle-Eastern	3213	1509	46.97
South Asian	3594	1662	46.24
Unknown or Legacy	5002	2214	44.26
White	27407	12283	44.82

Moreover, we generated a bar chart to visualize the number of strip searches conducted on people of different races (Figure 6), and a table to display the percentage of people who were strip searched for each race group (Table 3).

Figure 6: Number of Strip Searches Conducted on People of Different Races
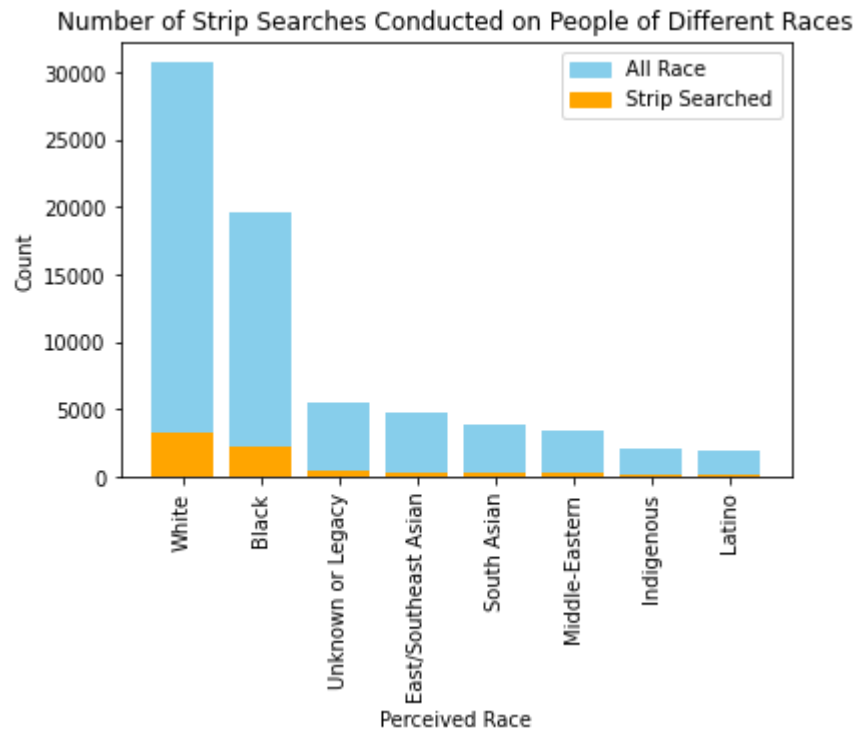
Table 3: Percentage of People Who Were Strip Searched for Each Race Group

Perceived Race	Total Arrests	Strip Searched	Strip Searched %
Black	17352	2299	13.25
East/Southeast Asian	4388	327	7.45
Indigenous	1907	287	15.05
Latino	1752	125	7.13
Middle-Eastern	3213	214	6.66
South Asian	3594	248	6.90
Unknown or Legacy	5002	493	9.86
White	27407	3338	12.18

It is obvious that Indigenous people have the highest strip search rate at 15.05%, followed by Black people at 13.25% and White people at 12.18%.

Figure 7: Correlation Heatmap
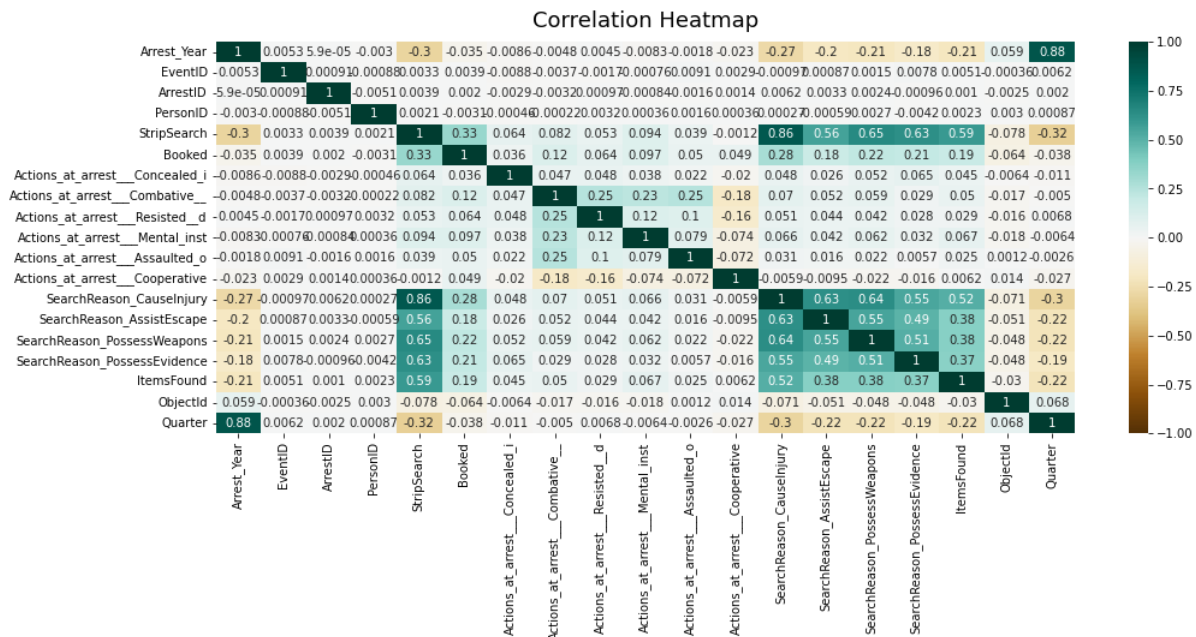
Furthermore, we generated a correlation heatmap (Figure 7) to visualize the relationships between the different attributes of the dataset. The heatmap indicated a strong correlation between strip searches and the reason for the search, particularly when the reason was to cause injury. This

finding is consistent with what we discovered during our literature review, which is that strip searches are often carried out when an arrested individual could pose a threat to cause injury.

T-Test

Before performing a t-test, we need to check the normality of the data as well as the equality of variances between the two groups being compared. Once the data meets these assumptions, we can use a t-test to compare the means of the two groups based on the number of arrests in each quarter.

T-Test #1: Mean Arrests Comparison between Sex

H_0 (Null Hypothesis): The mean number of arrests among sex groups is equal.

H_a (Alternative Hypothesis): The mean number of arrests among sex groups is not equal.

Our result indicates that the mean number of arrests for males (M=6513.25, SD=490.29) is higher than the mean number of arrests for females (M=1562.5, SD=244.09). Shapiro tests were conducted on males and females' total arrests grouped by year and month respectively. The p-value is 0.45 and 0.35 which are much greater than the benchmark 0.05, thus the samples satisfy normality. The p-value of Levene's test is reported as 0.23 and establishes the assumption of equal variance. The P-value of the T-test evidenced the means of arrests between sex is significantly different and it's consistent with the Figure1 obtained from EDA, we have enough evidence to reject the null hypothesis that means of arrest between sex are not equal.

T-Test #2: Mean of Cooperative Arrests between Perceived Races

H_0 (Null Hypothesis): The mean of cooperative arrests is equal.

H_a (Alternative Hypothesis): The mean of cooperative arrests is not equal.

Among all races, only white and unknown races fail normality. A loop was applied to carry out T-Tests between the perceived races with equal sample sizes. Note that there is not enough evidence to prove the mean of cooperative arrests is significantly different between unknown races and East/Southeast Asian, South Asian and Middle Eastern, Indigenous and Latino.

T-Test #3: Comparison of Drug-related Arrests' Mean between Two Age Groups: 18-24 vs. 45-54

We extracted the number of drug-related arrests for two different age groups: 18 to 24 years old and 45 to 54 years old and grouped the number of arrests by the eight quarters in which they occurred. Finally, we generated a line chart showing the number of drug-related arrests for each age group by quarter (Figure 8).

Figure 8: Drug Arrests by Quarter
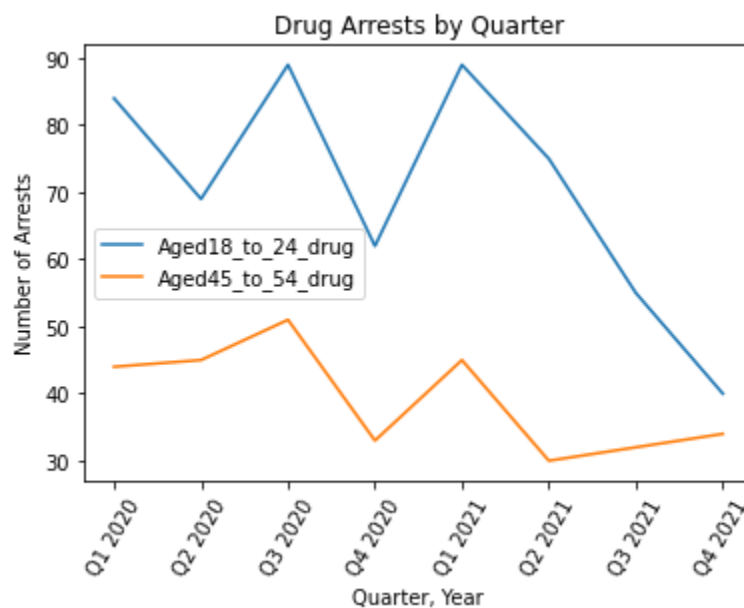
Our result indicates that the mean number of arrests for 18-24 age group (M=70.38, SD=16.32) is higher than the mean number of arrests for 45-54 age group (M=39.25, SD=7.34). We used levene function and found the variances of drug-related arrests of the two age groups are equal (P-value = 0.063699). Hence, we conducted an independent sample t-test to analyze whether the drug-related arrests differ between two age groups. The hypothesis being tested are the following:

H_0 (Null Hypothesis): The mean number of drug-related arrests per quarter for the two age groups (18-24 and 45-54 years old) is equal.

H_a (Alternative Hypothesis): The mean number of drug-related arrests per quarter for the two age groups (18-24 and 45-54 years old) is not equal.

Our results indicate the P-value is 0.000412, which is less than 0.05. Therefore, we can reject the null hypothesis. There is a significant difference in the mean number of drug-related arrests per quarter between the two age groups: 18-24 and 45-54 years old. As we can see from the above chart, the number of drug-related arrests is higher in the 18-24 age group than in the 45-54 age group.

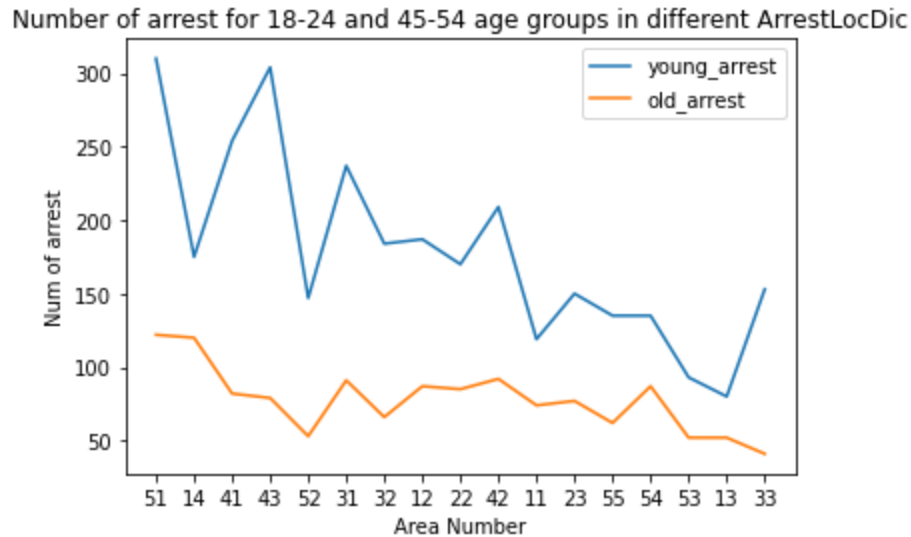## T-Test #4: Comparison of Mean of Arrests in Different Arrest Location between The Two Age Groups: 18-24 vs. 45-54

Our result indicates that the mean number of arrests for 18-24 age group in different locations (M=178.94, SD=64.2) is higher than the mean number of arrests for 45-54 age group in different locations (M=77.76, SD=21.77). After we check the assumptions such as normality and equal variance, we conducted an independent sample t-test to analyze whether the number of arrests in different locations differ between two age groups. The hypothesis being tested are the following:

H_0 (Null Hypothesis): The mean number of arrests in different arrest location for the two age groups (18-24 and 45-54 years old) is equal.

H_a (Alternative Hypothesis): The mean number of arrests in different arrest location for the two age groups (18-24 and 45-54 years old) is not equal.

As the p-value is 1.18e-6, which is less than 0.05. Therefore, we can reject the null hypothesis. This indicates that there is a significant difference in the number of arrests in different locations differ between two age groups. The line chart below also clearly shows that the number of arrests in 18-24 age group is higher than in the 45-54 age group across all locations (Figure 9).

Figure 9: Number of Arrest for 18-24 and 45-54 Age Groups in Different ArrestLocDic

Number of arrest for 18-24 and 45-54 age groups in different ArrestLocDic

T-Test #5: Comparison of Mean of Arrests in Location 41 and 43

We are interested in whether there is a significant difference in the mean number of arrests among different locations, so we randomly selected location 41 and 43 to do our analysis.

Our result indicates that the mean number of arrests in location 41 (M=330, SD=33.29) is higher than the mean number of arrests in location 43 (M=325.62, SD=49.86). The line chart below displays the number of arrests in ArrestLocDic 41 and 43 over eight quarters (Figure 10). As we can see from the chart, the number of arrests in location 41 started off lower than that in location 43, but in the last few quarters, the number of arrests in location 42 surpassed that in location 43. We need to conduct a t-test to determine whether there is a significant difference in the mean number of arrests between the location 41 and 43.

Figure 10: Number of Arrest in ArrestLocDic 41 and 43 over Eight Quarters
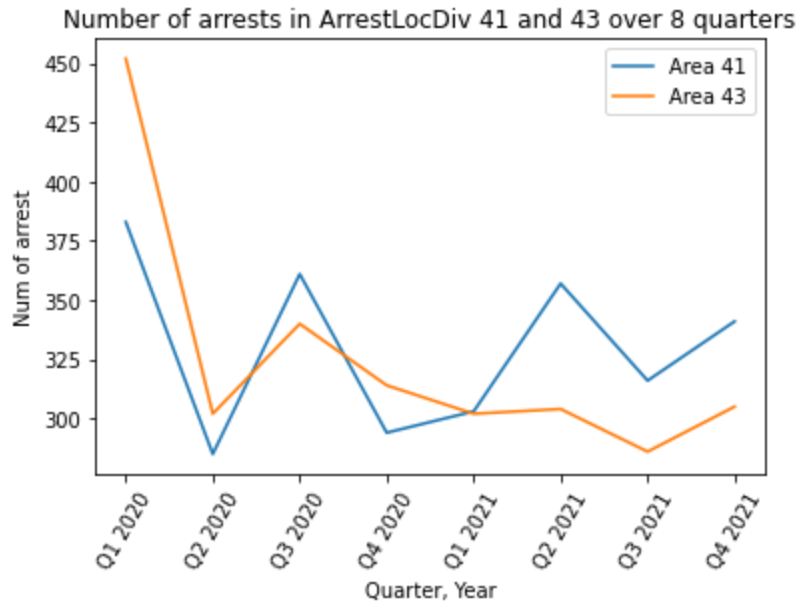
Number of arrests in ArrestLocDiv 41 and 43 over 8 quarters

After we check the assumptions such as normality and equal variance, we conducted an independent sample t-test to compare the mean number of arrests between two different locations, 41 and 43. The hypothesis being tested are the following:

H_0 (Null Hypothesis): The mean number of arrests in both locations is equal.

H_a (Alternative Hypothesis): The mean number of arrests in both locations is not equal.

The results of the t-test indicated that the P-value is 0.85, which is much greater than the significance level of 0.05. Therefore, we cannot reject the null hypothesis and the findings indicate that there is no significant difference in the mean number of arrests between location 41 and 43.

Research Design and Methods

Dataset Description

In our proposed project, we will use a dataset that includes information related to all arrests and strip searches that took place in the city of Toronto between 2021 and 2022. The dataset was collected by Toronto Police Service, which contains 65,276 records and the various attributes include arrest dates, perceived race, sex, age group, reasons for arrest, strip search, actions at arrest, search reasons, item found, etc. The dataset is available on Toronto Police Service Public Safety

Data Portal and could be found through the following link:
. We have provided a full list of the attributes in the appendix. Data is provided in either text format, in binaries 0 or 1 format or in numeric integer format.

While some study indicates that other factors such as age and gender also make impacts whether or not the person who was arrested was subjected to a strip search, the paper did not extensively examine these factors.

ANOVA Test Method

One-Way ANOVA Test: Perceived Race & Number of Arrests

H_0 (Null Hypothesis): The mean number of arrests across difference races is equal.

H_a (Alternative Hypothesis): The mean number of arrests across difference races is not equal.

$Y_{ij} = \mu_i + \epsilon_{ij}$ where Y_{ij} is the jth observation of total arrests in the ith perceived race group, μ_i is the mean response of the ith perceived race group and ϵ_{ij} is the random error.

One-way ANOVA Test will be performed to examine the difference between means number of arrests among different perceived races.

Two-Way ANOVA Test #1: Perceived Race, Actions_at_arrest-cooperative & Number of Arrests

H_0 (Null Hypothesis): There is no significant interaction effect or main effect of perceived race, cooperation on number of arrests.

H_a (Alternative Hypothesis): There is a significant interaction effect or main effect of perceived race, cooperation on number of arrests.

$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$ where Y_{ijk} is the response for the kth observation in the jth level of cooperative or not and i^{th} level of perceived race, μ is the overall mean of

total arrests, τ_i represents the effect of the ith level of perceived race, β_j is the jth level of cooperative or not, $(\tau\beta)_{ij}$ is the interaction effect between the ith level of perceived race and j^{th} level of cooperative or not and ϵ_{ij} is the random error.

Two-Way ANOVA Test #2: Perceived Race, Actions_at_arrest-cooperative & Strip Search

H_0 (Null Hypothesis): There is no significant interaction effect or main effect of perceived race, cooperation on number of strip searches.

H_a (Alternative Hypothesis): There is a significant interaction effect or main effect of perceived race, cooperation on number of strip searches.

$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$ where Y_{ijk} is the response for the kth observation in the jth level of cooperative or not and i^{th} level of perceived race, μ is the overall mean of total strip searches, τ_i represents the effect of the ith level of perceived race, β_j is the jth level of cooperative or not, $(\tau\beta)_{ij}$ is the interaction effect between the ith level of perceived race and j^{th} level of cooperative or not and ϵ_{ij} is the random error.

If the interaction effects are found to be statistically significant after conducting the Two-way ANOVA tests mentioned above, we will use the Tukey HSD test as a post-hoc test to further analyze our research questions.

Results

RQ1: Is there any difference in the mean number of arrests across different races?

The reported p-value is less than the threshold of 0.05 which shows we have sufficient evidence to prove that at least one race's mean number of arrests is different from the other groups. This result is consistent with the loop results of the T-test conducted in the previous section.

The one-way ANOVA test we performed indicated a significant difference in the mean number of arrests among different perceived races. Therefore, we can conclude that the number of arrests may be associated with the perceived race of the individuals. The findings are consistent with our

literature review, which has shown that people from certain racial groups are more likely to be overrepresented in the criminal justice system.

RQ2: Does perceived race and action on arrest (cooperative or not) affect the number of arrests?

Two-Way ANOVA Test #1: Perceived Race, Actions_at_arrest-cooperative & Number of Arrests

Table 3: Two-way ANOVA results for the effects of Perceived Race and Cooperation Level on Number of Arrests

	sum_sq	df	F	PR(>F)
C(Perceived_Race)	3.778557e+07	7.0	763.741623	2.195836e-91
C(Actions_at_arrest___Cooperative)	3.340509e+05	1.0	47.264077	3.708155e-10
C(Perceived_Race):C(Actions_at_arrest___Cooperative)	5.828127e+05	7.0	11.780114	3.617279e-11
Residual	7.915886e+05	112.0	NaN	NaN

We used Two-Way ANOVA Test to inspect the mean effect of perceived race, whether the arrested people are cooperative or not and their interaction effect on the total arrests (Table 3). Based on the test results, perceived race, cooperative or not at arrests and their interaction effects are all statistically significant, as evidenced by their P-values being smaller than 0.05. Among them, the effect of perceived race seems to be the most significant, with a high F statistic, while the interaction effect was relatively small. To further investigate the effects, a Tukey HSD is carried out.

Post-hoc Test # 1

Since the interaction effect being statistically significant, a Tukey HSD was conducted to compare all possible pairs of means (Table 4). The result suggests that the comparison between strip searched or not was not significant, while all possible comparisons of means of perceived race were significant except for two groups. These findings suggest that there are significant differences in the mean values of the dependent variable among different groups based on their perceived race.

Table 4: Tukey test for Perceived Race and Cooperation Level

```
Tukey test for Perceived Race:
    Multiple Comparison of Means - Tukey HSD, FWER=0.05
=============================================================
group1 group2  meandiff  p-adj    lower       upper     reject
-------------------------------------------------------------
     1      2 -1400.3125  0.001 -1530.4567 -1270.1683   True
     1      3  -628.4375  0.001  -758.5817  -498.2933   True
     1      4 -1488.3125  0.001 -1618.4567 -1358.1683   True
     1      5  -1593.75   0.001 -1723.8942 -1463.6058   True
     1      6  -1512.125  0.001 -1642.2692 -1381.9808   True
     1      7 -1603.4375  0.001 -1733.5817 -1473.2933   True
     1      8 -1438.6875  0.001 -1568.8317 -1308.5433   True
     2      3   771.875   0.001   641.7308   902.0192   True
     2      4     -88.0  0.4325  -218.1442    42.1442  False
     2      5  -193.4375  0.001  -323.5817   -63.2933   True
     2      6 -111.8125  0.1485  -241.9567    18.3317  False
     2      7  -203.125   0.001  -333.2692   -72.9808   True
     2      8   -38.375     0.9  -168.5192    91.7692  False
     3      4  -859.875   0.001  -990.0192  -729.7308   True
     3      5 -965.3125   0.001 -1095.4567  -835.1683   True
     3      6 -883.6875   0.001 -1013.8317  -753.5433   True
     3      7   -975.0    0.001 -1105.1442  -844.8558   True
     3      8  -810.25    0.001  -940.3942  -680.1058   True
     4      5 -105.4375 0.2058  -235.5817    24.7067  False
     4      6  -23.8125     0.9  -153.9567   106.3317  False
     4      7 -115.125  0.1245  -245.2692    15.0192  False
     4      8    49.625     0.9   -80.5192   179.7692  False
     5      6    81.625 0.5257   -48.5192   211.7692  False
     5      7   -9.6875     0.9  -139.8317   120.4567  False
     5      8  155.0625 0.0083    24.9183   285.2067   True
     6      7  -91.3125 0.3823  -221.4567    38.8317  False
     6      8   73.4375 0.6396   -56.7067   203.5817  False
     7      8    164.75 0.0038    34.6058   294.8942   True
-------------------------------------------------------------


Tukey test for Cooperation:
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
======================================================
group1 group2  meandiff p-adj    lower     upper  reject
------------------------------------------------------
     0      1 -102.1719 0.3018 -297.2014 92.8576  False
------------------------------------------------------
```

The two-way ANOVA test we performed revealed that perceived race, cooperative or not during arrests and their interaction effects are all statistically significant in relation to the total number of arrests, which is consistent to the interaction plot where the lines are not parallel (Figure 11). The lines trend in the same direction, indicating a possible positive interaction effect. To further investigate the effects, we conducted a Tukey HSD test, which suggests that there are significant differences in the number of arrests based on perceived race among most groups, while the level of cooperation during arrests did not differ significantly.

Figure 11: Interaction between Perceived Race and Cooperation on Number of Arrests

Interaction between Perceived Race and Cooperation on Number of Arrests
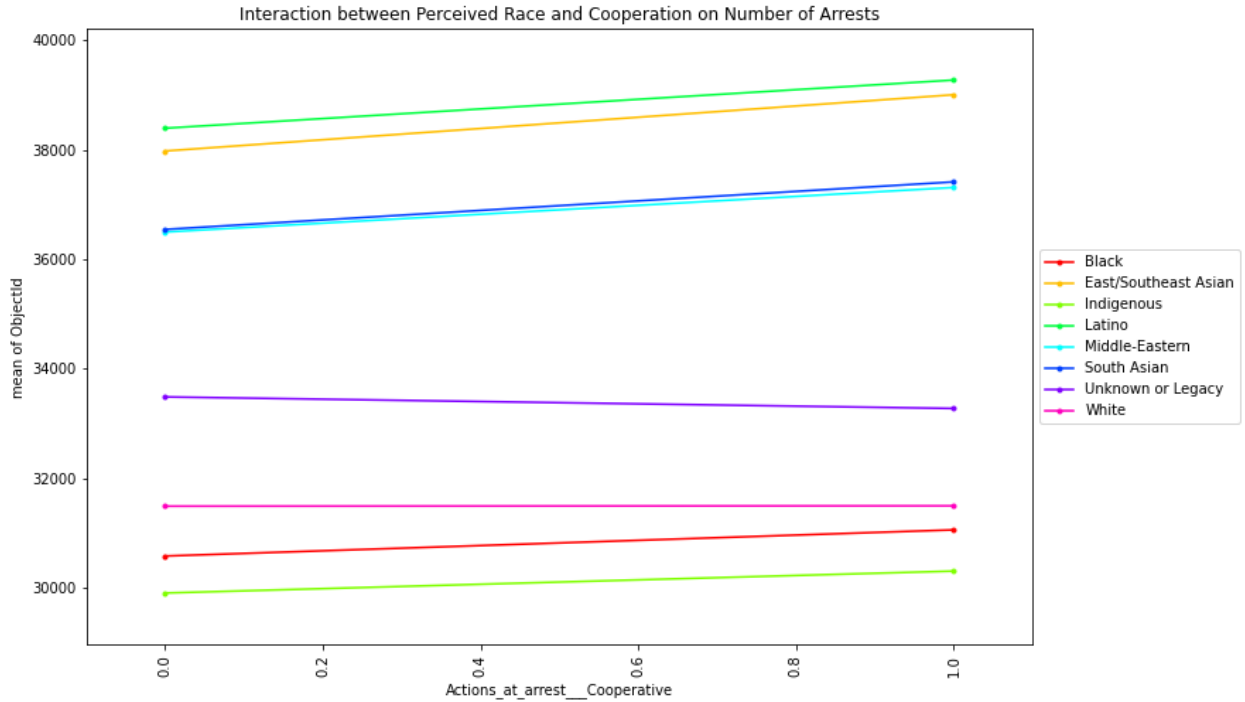
RQ3: Does perceived race and action on arrest (cooperative or not) affect strip searches?

Two-Way ANOVA Test #2: Perceived Race, Actions_at_arrest-cooperative & Strip Search

Table 5: Two-way ANOVA results for the effects of Perceived Race and Cooperation Level on Strip Search

	sum_sq	df	F	PR(>F)
C(Perceived_Race)	6.195917e+05	7.0	7.997475	9.171208e-08
C(Actions_at_arrest__Cooperative)	4.830548e+03	1.0	0.436457	5.102842e-01
C(Perceived_Race):C(Actions_at_arrest__Cooperative)	9.776079e+03	7.0	0.126186	9.962794e-01
Residual	1.162101e+06	105.0	NaN	NaN

The results of the two-way ANOVA with interaction shows that there is a significant effect of perceived race on the number of strip search, while cooperative or not and the interaction effect are not significant (Table 5). This suggests that the level of the dependent variable differs significantly among different levels of perceived races.

Post-hoc Test # 2

**Table 6: Tukey test for Perceived Race and Strip Search**

```
Tukey test for Perceived Race:
 Multiple Comparison of Means - Tukey HSD, FWER=0.05
===========================================================
group1 group2  meandiff p-adj    lower     upper    reject
-----------------------------------------------------------
    1      2 -177.8125  0.001 -289.2379 -66.3871   True
    1      3  -64.9375 0.6051 -176.3629  46.4879   False
    1      4 -190.9107  0.001  -306.247 -75.5745   True
    1      5  -188.125  0.001 -303.4612 -72.7888   True
    1      6 -194.3583  0.001 -307.6256 -81.0911   True
    1      7 -199.6964  0.001 -315.0327 -84.3602   True
    1      8 -188.1875  0.001 -299.6129 -76.7621   True
    2      3   112.875 0.0448    1.4496 224.3004   True
    2      4  -13.0982    0.9 -128.4345  102.238   False
    2      5  -10.3125    0.9 -125.6487 105.0237   False
    2      6  -16.5458    0.9 -129.8131  96.7214   False
    2      7  -21.8839    0.9 -137.2202  93.4523   False
    2      8   -10.375    0.9 -121.8004 101.0504   False
    3      4 -125.9732  0.022 -241.3095  -10.637   True
    3      5 -123.1875 0.0275 -238.5237  -7.8513   True
    3      6 -129.4208 0.0136 -242.6881 -16.1536   True
    3      7 -134.7589 0.0105 -250.0952 -19.4227   True
    3      8  -123.25 0.0193 -234.6754 -11.8246   True
    4      5    2.7857    0.9  -116.333 121.9045   False
    4      6   -3.4476    0.9 -120.5642  113.669   False
    4      7   -8.7857    0.9 -127.9045  110.333   False
    4      8    2.7232    0.9  -112.613 118.0595   False
    5      6   -6.2333    0.9   -123.35 110.8833   False
    5      7  -11.5714    0.9 -130.6902 107.5473   False
    5      8   -0.0625    0.9 -115.3987 115.2737   False
    6      7   -5.3381    0.9 -122.4547 111.7785   False
    6      8    6.1708    0.9 -107.0964 119.4381   False
    7      8   11.5089    0.9 -103.8273 126.8452   False
-----------------------------------------------------------

Tukey test for Strip Search:
 Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====================================================
group1 group2 meandiff p-adj   lower     upper  reject
-----------------------------------------------------
    0      1  -11.677 0.6112 -55.8514 32.4973  False
```

Similarly, a Tukey HSD was used to compare the means of strip search for different groups of perceived races (Table 6). The test revealed that there were significant differences in the mean number of arrests for all pairs of races, except for groups 2 and 3. Specifically, the mean number of strip search for White was significantly lower than that of all other racial groups. The mean number of arrests for individuals in South Asian, Indigenous, Middle Eastern, Latino, East/Southeast Asian were also significantly lower than that of Unknown and Black. There was no significant difference in the mean number of strip search between Black and Unknown people.

The results indicate that there is no association between cooperation and strip searches. In addition, the finding suggest that the race of an individual may play a role in the number of strip searches made by the police, with certain racial groups being more likely to be subjected to strip searches than others. Note that the lines in the interaction plot are not parallel, however we found no significant main effect of cooperation, this could be due to random variation in the data or other unmeasured factor and small sample size (Figure 12).

Figure 12: Interaction between Perceived Race and Cooperation on Strip Search
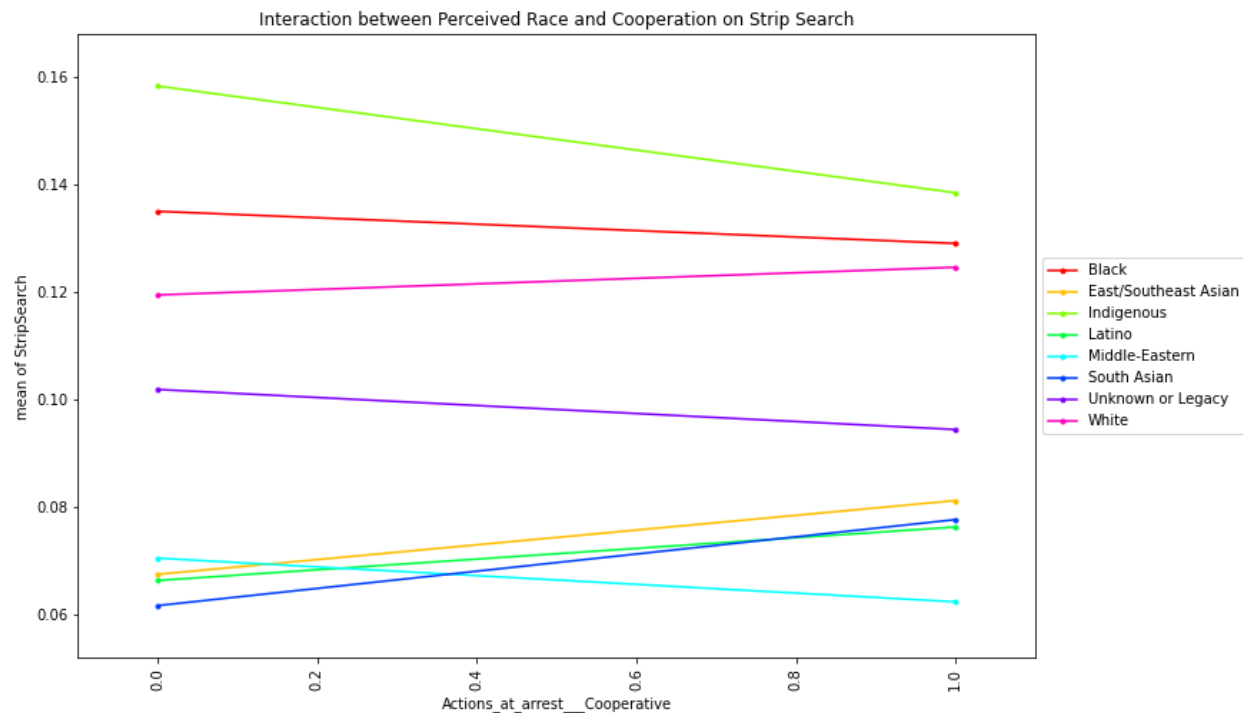
Preliminary Suppositions and Implications

Our project aims to investigate the how the combination of features affects on strip searches and arrests. This project will provide insight into how perceived race and cooperation behaviour interact with the number of arrests and strip research respectively. Knowledge gained from this project will contribute to existing literature and help police and policymakers improve equality in the criminal justice system.

Discussion and Conclusion

The test results of our analyses revealed significant differences in the mean number of arrests and the likelihood of being subjected to a strip search across different racial groups. Specially, we found that certain racial groups were more likely to be arrested and subjected to strip searches than other groups. The finding highlights the need for further investigation into potential inequalities in the criminal justice system.

Reference

Government of Canada, D. of J. (2022, August 17). *Public perception of crime and justice in Canada: A review of opinion polls*. Canadians' Concerns in Relation to Crime. Retrieved February 22, 2023, from https://www.justice.gc.ca/eng/rp-pr/csj-sjc/crime/rr01_1/p1.html

McNeilly, G. (2019, March). Breaking the golden rule - office of the independent police review director. Office of the Independent Police Review Director . Retrieved February 24, 2023, from https://www.oiprd.on.ca/wp-content/uploads/OIPRD_Breaking-the-Golden-Rule_Report_Accessible.pdf

Newburn, T., Shiner, M., & Hayman, S. (2004). RACE, CRIME AND INJUSTICE? Strip Search and the Treatment of Suspects in Custody. The British Journal of Criminology, 44(5), 677–694. http://www.jstor.org/stable/23639161

Rankin, Jim, Quinn, Jennifer, Shephard, Michelle Simmie, Scott and Duncanson, John. 2002. "Singled Out: An Investigation into Race and Crime." Toronto Star. October 19: A1.

Statistics Canada. (2017). Toronto, C [Census subdivision], Ontario and Ontario [Province] (table). Census Profile. 2016 Census. Statistics Canada Catalogue no. 98-316-X2016001. Ottawa. Released November 29, 2017. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E

Toronto Police Service (2020). Online reporting-Strip Searches: Measurement & Outcomes. Online Reporting- Toronto Police Service. Retrieved February 24, 2023, from https://www.tps.ca/services/online-reporting/

Wortley , S., & Jung, M. (2020). (rep.). Racial Disparity in Arrests and Charges: An analysis of arrest and charge data from the Toronto Police Service (pp. 4–5). Toronto, Ontario: Government of Ontario.