



University of Toronto

INF2178 Midterm Project

Group Number: Group 50

By

Yijun Lin # 1003404620

Keying Zhu #1003790953

INF 2178H S: Experimental Design for Data Science

Professor: Shion Guha

February 27th, 2023

Table of Content

Section 1: Introduction	2
1.1 Background Introduction	2
1.2 Literature Review	2
1.3 Dataset Description	3
1.4 Research Questions/Objective	3
Section 2: Tidy Data Procedure	3
Section 3: Exploratory Data Analysis (EDA)	6
3.1 Descriptive Statistics	9
3.2 T-tests	12
3.2.1 Finalized sub-dataset 1	13
3.2.2 Finalized sub-dataset 2	13
Section 4: Method	14
4.1 ANOVA Tests	14
4.1.1 One-way ANOVA	14
4.1.2 Two-way ANOVA	15
4.2 Tukey's HSD	16
Section 5: Results/ Findings (includes interaction plots)	17
5.1 Finalized sub-dataset 1	17
5.2 Finalized sub-dataset 2	18
Section 6: Discussion and Conclusion	20
Reference	21
Appendix	21

Section 1: Introduction

1.1 Background Introduction

The Toronto Police Service (TPS) is the primary law enforcement agency responsible for policing the city of Toronto, Ontario, Canada. TPS has a clear mission to provide safety and security for Toronto residents and visitors while adhering to the principles of community policing, accountability, and professionalism. There is still room for improvement at TPS, and one of the most pressing issues facing TPS is racial profiling. Critics argue that even when there is no evidence of criminal activity, Toronto police disproportionately target BIPOC (black, Indigenous, and people of color) individuals, particularly Black and Indigenous individuals, for stops, searches, and arrests. This has resulted in a loss of trust between TPS and these communities, undermining the effectiveness of community policing and undermining public confidence in law enforcement. In recent years, TPS has come under increasing scrutiny and criticism over issues such as racial profiling, excessive use of force, and lack of accountability.

Therefore, the analysis in this report will be based on the dataset "Arrests and Strip Searches (RBDC-ARR-TBL-001)" provided by the Toronto Police Service Public Safety Data Portal. Through the research question: Are Police Actions That Lead to Strip Searches Influenced by Perceived Race? Is there a correlation between the frequency of strip searches and the policy situation of COVID-19, as well as year factors such as gender? Does the sex of the arrested person affect how cooperative they are in the arrest process? Also, does the time of quarter/yearly affect the level of cooperation of individuals during capture? Analyze the relationship between the police and the people to make further suggestions to improve the existing TPS system.

In this report, there will be six sections: Introduction, Tidy data procedure, EDA (Exploratory Data Analysis), Method, Results, and Conclusion including the discussion. The main topic of this report is proposing suggestions based on the existing data set to make improvements in the Toronto Police Service, which includes two aspects: the attitude of the police towards the masses and the attitude of the masses towards the police.

1.2 Literature Review

There is a history of racial bias in Canadian policing, including documented systemic racism in the criminal justice system. Black and Indigenous people in Canada are disproportionately targeted by police and are more likely to experience excessive use of force. (Owusu-Bempah, 2014) The authors' research found evidence of racial profiling and discrimination by Toronto police. For example, a 2008 report by the Ontario Human Rights Commission found that blacks were overrepresented in Toronto police use of force cases (Owusu-Bempah, 2014). Despite the evidence, the Toronto Police Service has consistently denied allegations of racial bias. The article examines the reasons why police officers may refuse to acknowledge the existence of racial bias, including the impact on police morale and the likelihood of litigation.

The history of systemic racism and racial discrimination in Canada highlights the disproportionate incarceration of Black people in the criminal justice system. The article *Data, denials, and confusion: The racial profiling debate in Toronto* reveals evidence of police racial profiling, discrimination, and harassment of black men (Wortley & Tanner, 2003). By examining the psychological impact of police interactions on black men, the authors show that these experiences lead to fear, anxiety, and distrust of police officers, as well as reduced trust in the criminal justice system as a whole.

1.3 Dataset Description

The data set "Arrests and Strip Searches (RBDC-ARR-TBL-001)" used in this report contains 65,276 data items collected and provided by the Toronto Police Service Public Safety Data Portal. The dataset contains information on 24 characteristics, including race, gender, age, and other demographic data, of individuals, arrested and strip-searched by TPS officers between January 1, 2020, and December 31, 2021 (Toronto Police Service, 2022). The dataset includes information about each arrest and strip search, such as the reason for the arrest, where the individual was arrested, and the type of crime charged. The dataset also includes information on individuals' level of cooperation at the time of arrest, as well as the date and time of arrest and strip search. The scope and limitations of this data set are that there are many missing values and too many categorical variables. This dataset can be accessed via <https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/> explore and anyone has permission to view it (Toronto Police Service, 2022). The appendix section provided a full list of the attributes. Data is mostly provided in either yes(1) or no(0) format, text-based format, or numerical integer format.

1.4 Research Questions/Objective

Here below are our group's main topics with two research questions.

Main topic: Improvement of Toronto Police Service System

Research Question 1: If there exist biased treatments on police actions which result in strip-searching based on perceived race? Do the factors of the year (considering the situation for policies in covid19) and sex have an impact on the strip-search frequency?

Research Question 2: Does gender play a role in cooperation levels while being arrested? Does quarterly/yearly time also influence the people's cooperation attitude during the arrest?

Section 2: Tidy Data Procedure

Before the real application of the dataset, in order to better research the related research topics, the data cleaning procedure is applied first. By reading the dataset into Python and checking the null value in each column, Table 1 shows the number of null values contained in each column which is listed below.

Column Name	Null Value Counts
Arrest_Year	0
Arrest_Month	0
Event ID	0
Arrest ID	469
Person ID	0
Perceived Race	4
Sex	0
Age_group_at_arrest	24
Younth_at_arrest_under_18_years	0
ArrestLocDiv	0
StripSearch	0
Booked	0
Occurrence Category	165
Actions_at_arrest_concealed_i	0
Actions_at_arrest_Combative_	0
Actions_at_arrest_Resisted_d	0
Actions_at_arrest_Mental_inst	0
Actions_at_arrest_Assaulted_o	0
Actions_at_arrest_Cooperative	0
SearchReason_CauseInjury	57475

SearchReason_AssistEscape	57475
SearchReason_PossessWeapons	57475
SearchReason_PossessEvidence	57475
ItemsFound	57475

Table 1 Counts of missing value for all columns

Based on the table 1 result, since column Perceived_Race includes 4 null values which are relatively small values which might not highly affect the statistical result. The 4 missing values will be dropped from the original datasets. Also, columns relate to search reasons and ItemsFound since it contains too many null values, it is hard for us to make appropriate research questions and will make inaccurate statistical analysis results hence they will be removed from the dataset.

After dropping null values for selected columns, next the value counts for each column are applied to check whether the columns have the text format of null or unknown values and consider doing extra editions with that. By checking the value counted table based on Table 2, there are only two columns which need to take into consideration extra cleaning procedures. Column Perceived_Race contains the category of Unknown and Legacy with 5056 rows as the third largest categories in the current column of this dataset. Since it is reasonable to have situations when the polices cannot distinguish exact peoples' race and it has a relatively large share in race group, this category is kept. The Sex column consists of three categories, female, male and unisex, by comparing unisex with the other two categories, it only has 9 rows which makes the sample biased and might be hard to get statistical analysis results with other groups. Unisex is excluded from current research topics.

---- Perceived_Race ---	
Name	Value counts
White	27723
Black	17526
Unknown or Legacy	5056
East/Southeast Asian	4415
South Asian	3613
Middle-Eastern	3237

Indigenous	1934
Latino	1768
---- Sex ---	
Name	Value counts
M	52647
F	12626
U	9

Table 2 Patically value counted table for all columns

Based on the dataset description, some of the columns are in text format which is hard for data visualization purposes such as race, sex, age group and more. The mutation for those groups from 1 to n (regards to the number of categories in this column) into numerical categorical variables will help us consider the selection for variables in research questions based on correlation matrix result and form subset data frames. For null value columns that have not dropped from columns, it will convert to the number 999 and consider them as outliers.

Section 3: Exploratory Data Analysis (EDA)

Since the application of the data cleaning procedure in section 2, the overview of data visualization for the current dataset will help us better understand which columns are suitable for the research questions and the approaches in creating subsets. A correlation matrix (Figure 1) shows the relationship between each variable as the majority of variables have relatively weak relationships with each other for all correlation coefficients are smaller than ± 0.5 . Reviewing back to two research questions, StripSearch and Actions_at_arrest_Cooperative are two columns about the main topics and also as starting points of researching, by checking their relationship with others. Perceived_Race and sex have a weak relationship with strip-searching hence the arrest_year is added as an external factor for research question one as it has the highest negative correlation coefficient with a value of -0.31. For Actions_at_arrest_Cooperative, as all variables are not highly correlated with it and there does not exist a comparatively stronger relationship included, the selections for research question two will rely on the general thinking with background research information.

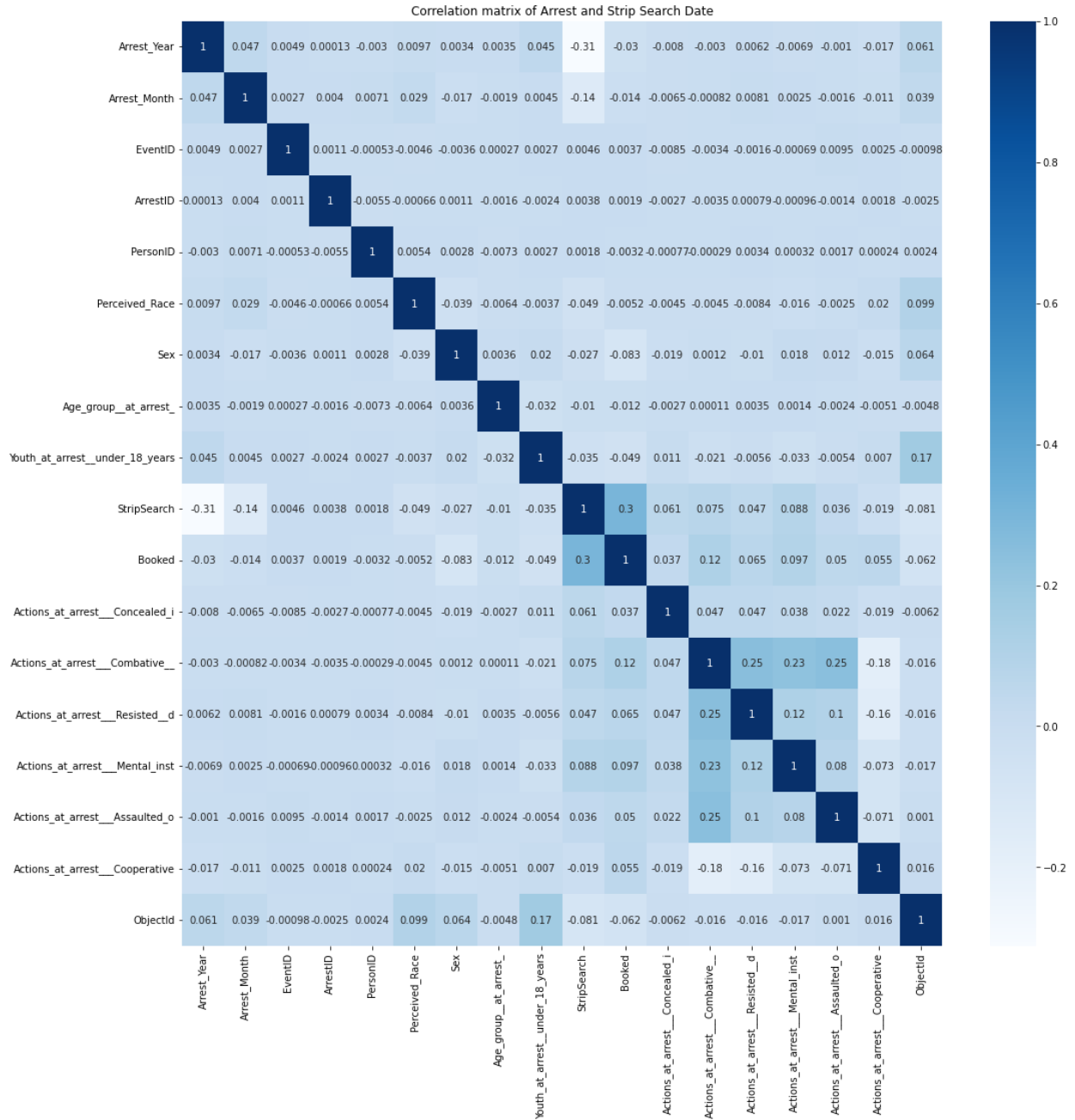


Figure 1 Correlation Matrix

As the confirmation in columns selections and general ideas about this data, columns StripSearch and Actions_at_arrest_Cooperative are composed of binary variables which generate difficulties for statistical results based on t-test and ANOVA, data merging to create new columns that are made by continuous numerical values is the next step target in this section. The columns that are associated with research question 1 are Arrest_Year, Perceived Race, Sex and StripSearch. For research question 1, the main goal of this question focuses on the situation when people of different races are being strip searched hence the counting of frequency for already strip search

actions with different races will be used as dependent variables. Add the purpose of diversity research in this dataset, external control factors arrest year and sex are used with races to count the frequency. Table 3 practically shows the finalized sub-dataset with 32 rows and 4 columns that will be used for research question 1.

	Arrest_Year	Perceived_Race	Sex	Strip_Search_counts
0	2020	Black	F	239
1	2020	Black	M	1984
2	2020	East/SouthEast Asian	F	23
3	2020	East/SouthEast Asian	M	272
4	2020	Indigenous	F	68
5	2020	Indigenous	M	218
6	2020	Latino	F	10
7	2020	Latino	M	116
8	2020	Middle-Eastern	F	15
9	2020	Middle-Eastern	M	191
10	2020	South-Asian	F	26
11	2020	South-Asian	M	215
12	2020	Unknown or Legacy	F	92
13	2020	Unknown or Legacy	M	405
14	2020	White	F	698
15	2020	White	M	2542

Table 3 Finalized data for research question 1(Partially)

The same logic can be applied by creating a sub-dataset with research question 2 as columns Arrest_Year, Arrest_Month, Sex and Actions_at_arrest_Cooperative are picked by filtering the frequency counts for people who have cooperative attitudes during the arrest period. The finalized sub-dataset(16 rows with 4 columns) is listed in Table 4 below.

	Arrest_Year	Arrest_Month	Sex	Cooperative_counts
0	2020	Apr-June	F	493
1	2020	Apr-June	M	2601
2	2020	Jan-Mar	F	983
3	2020	Jan-Mar	M	3497
4	2020	July-Sept	F	665
5	2020	July-Sept	M	3074
6	2020	Oct-Dec	F	573
7	2020	Oct-Dec	M	2654
8	2021	Apr-June	F	649
9	2021	Apr-June	M	2929
10	2021	Jan-Mar	F	673
11	2021	Jan-Mar	M	2944
12	2021	July-Sept	F	690
13	2021	July-Sept	M	2936
14	2021	Oct-Dec	F	704
15	2021	Oct-Dec	M	3040

Table 4 Finalized data for research question 2

3.1 Descriptive Statistics

To be more familiar with the two sub-datasets, we produce the following charts and tables.

Figure 2 indicates the number of strip searches that happen in each race for a total of 8 types. It is unsurprising to see black and white races take the majority of strip search counts because compared to other races, white races took up the highest portion of Toronto's population and the black race was the group with racism issues/events in recent years which add the meaningfulness to this area of research. For other race groups, since they have relatively small counts to compare with, it is hard to see a clear trend or assumptions for this moment. Table 5 shows the mean search counts by different races as an add-on to Figure 2 which also indicates the white race has the highest mean with the highest standard deviation representing higher spread out. By adding

sex as another independent variable, a boxplot (figure 3) is created with no outliers. All female groups have lower strip search frequency than males which shows the differences between sex. For those 6 races except white and black, the female groups have less frequency of checking which shows as short boxplots.

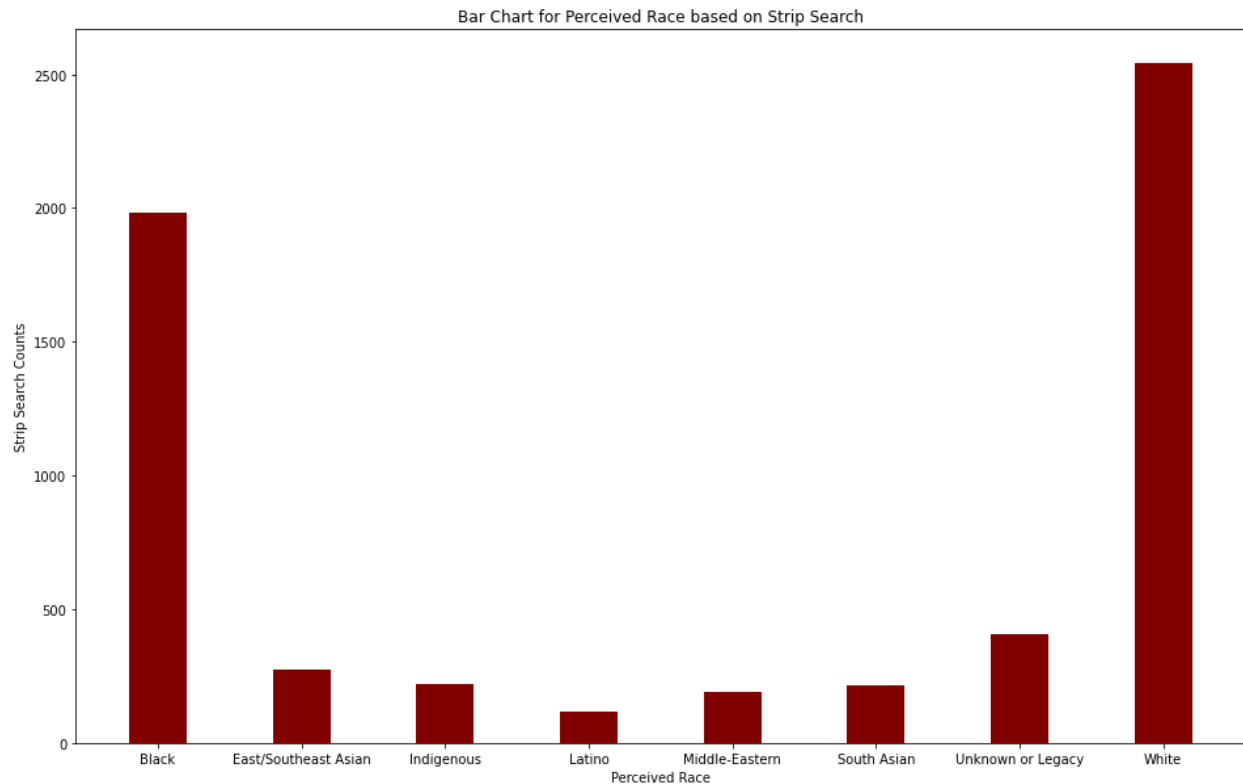


Figure 2 Bar chart for Perceived Race based on Strip Search Counts

	Strip Search Counts						
	Count	Mean	Standard Deviation	Min	25%	75%	Max
Perceived Race							
Black	4	608.5	922.06	16	150.25	675.25	1984.0
East/SouthEast Asian	4	85.25	125.68	2	17.75	101.0	272.0
Indigenous	4	76.5	98.27	6	12.0	105.5	218.0
Latino	4	33.0	55.49	0	4.5	36.5	116.0
Middle-Eastern	4	57.0	89.59	3	12.0	62.0	191.0

South Asian	4	64.25	101.01	1	11.5	73.25	215.0
Unknown or Legacy	4	134.0	184.31	4	27.25	170.25	405.0
White	4	891.5	1130.9	80	204.5	1159.0	2542.0

Table 5 Summary Statistics for Strip Search Counts by race

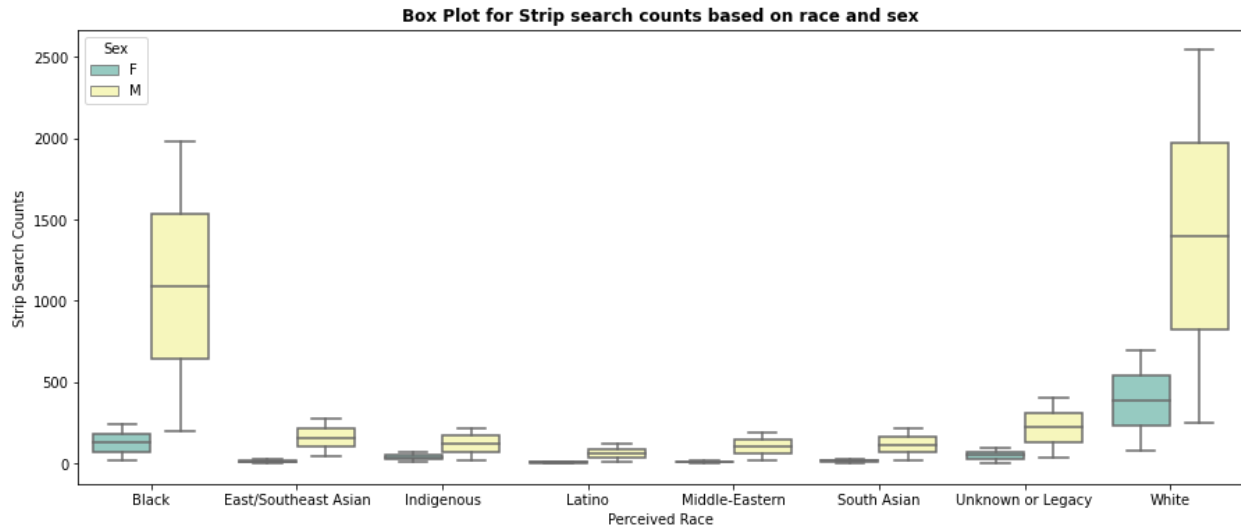


Figure 3 Boxplot for Strip Search count based on race and sex

The following figures or tables will link to research question two, figure 4 shows the counts of cooperative attitude at arrest based on sex, it is surprising to see male have higher cooperative counts for females as it reserves the general thinking for race difference. Table 6 shows the summary statistical information for cooperative counts based on arrest month as Jan-Mar has the highest mean with the highest standard deviation, and no strong issues are discovered in IQR.

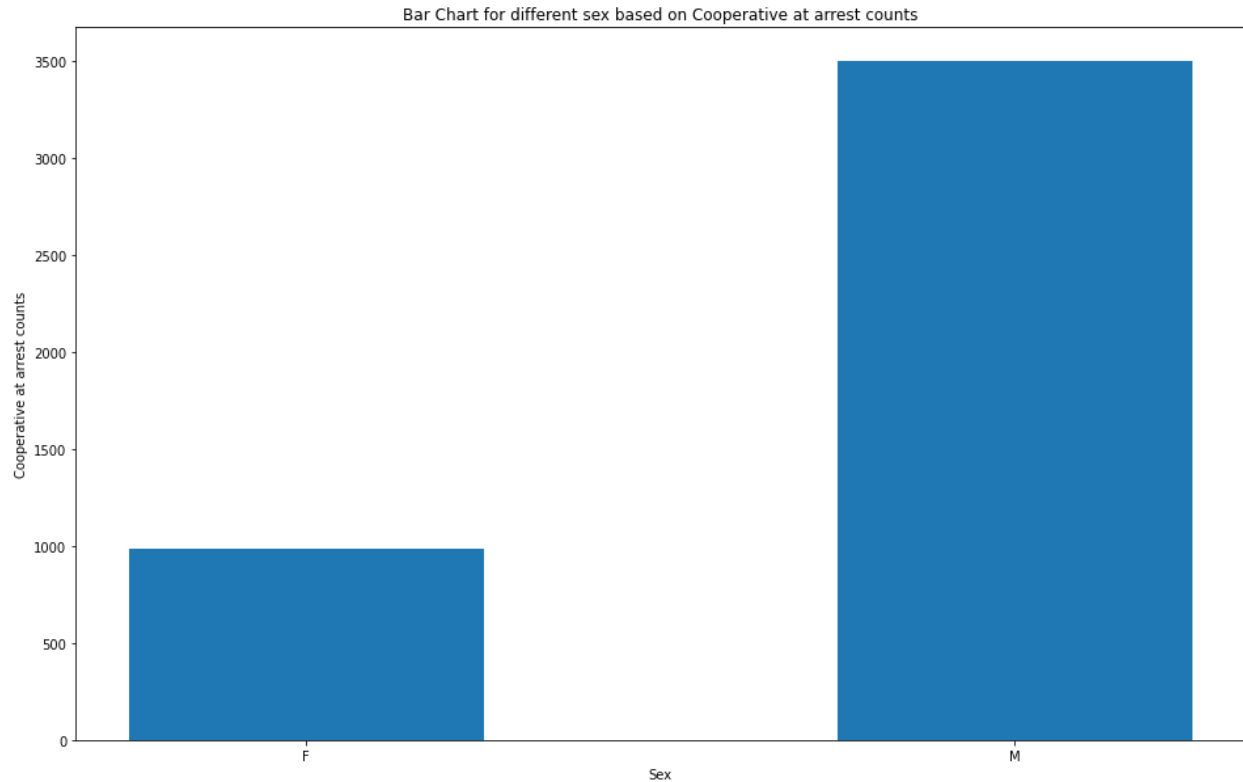


Figure 4 Bar chart for Perceived Race based on Strip Search Counts

	Cooperative Counts						
	Count	Mean	Standard Deviation	Min	25%	75%	Max
Arrest Month							
Apr-June	4	1668.0	1275.36	493	610.0	1625.0	2929.0
Jan-Mar	4	2024.25	1405.35	673	905.5	1963.5	3497.0
July-Sept	4	1841.25	1345.0	665	683.75	1813.0	3074.0
Oct-Dec	4	1742.75	1285.89	573	671.25	1679.0	3040.0

Table 6 Summary Statistics for Cooperative at arrest Counts by arrest month

Figure 5 demonstrates how month and sex behave based on the counts of cooperative actions at arrest, it is obvious to see female groups have smaller cooperative counts than male. For July-Sept, it has relatively short box plots which means they have a similar attitude toward

cooperation in their groups. For Jan-Mar, it has a higher position of boxplot for two genders which shows the differences between other groups and can be further discussed in the following sections. There is no outlier for this boxplot.

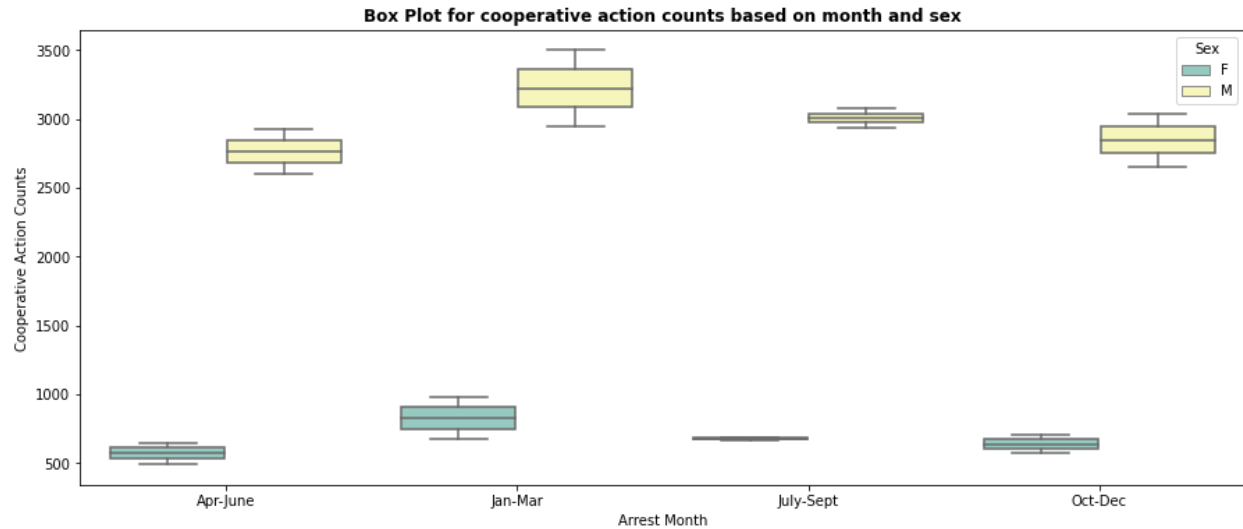


Figure 5 Bar chart for cooperative action counts based on month and sex

3.2 T-tests

Even the procedure of mutation helps us understand the correlation between numerical variables with categorical variables, Welch's t-test is a more straightforward way to show relationships in different groups for categorical variables and for double confirmation with the choice of our current sub-datasets. We check the assumptions line by line for our sub-datasets and they are all satisfied: 1) nominal two-level explanatory variable; 2) quantitative variable as outcome variable; 3) normality and 4) independence of errors. For Welch's t-test we launch, there are no assumptions with equal variances among residuals. The paragraphs below demonstrate some useful statistical results extracted from the t-test. As there are two research questions with different sub-datasets, the t-test result explanation will be split into two main parts.

3.2.1 Finalized sub-dataset 1

Sex and Strip Search Counts

For sex and strip search counts, we use Welch's T-test to analyze whether the population mean for strip search(outcome variable) actions will differ in sex groups(male and female as explanatory variable). The hypothesis is stated as followed:

H_0 (Null Hypothesis): The population means for strip search counts for two independent groups: male and female are equal.

H_1 (Alternative Hypothesis): The population means for strip search counts for two independent groups: male and female are not equal.

By comparing significance level at 0.05, the t-test statistics result is -1.719 with p-value of 0.0959. As $0.0959 > 0.05$, there do not exist statistically significant differences. Therefore, we failed to reject the null hypothesis.

Year and Strip Search Counts

For year and strip search counts, we use Welch's T-test to analyze whether the population mean for strip search(outcome variable) actions will differ in different two years (2020 and 2021 as explanatory variables). The hypothesis is stated as followed:

H_0 (Null Hypothesis): The population means for strip search counts for two independent groups: 2020 and 2021 are equal.

H_1 (Alternative Hypothesis): The population means for strip search counts for two independent groups: 2020 and 2021 are not equal.

By comparing significance level at 0.05, the t-test statistics result is 2.166 with p-value of 0.0384. As $0.0384 < 0.05$, there exist statistically significant differences. Therefore, we can reject the null hypothesis and conclude there are statistical differences between the year 2020 and year 2021 groups with the population mean for the strip search action.

3.2.2 Finalized sub-dataset 2

Sex and Cooperation Counts

For sex and cooperation counts, we use Welch's T-test to analyze whether the population mean for counts of cooperation at arrest (outcome variable) will differ in sex groups(male and female as explanatory variable). The hypothesis is stated as followed:

H_0 (Null Hypothesis): The population means for cooperation at arrest counts for two independent groups: male and female are equal.

H_1 (Alternative Hypothesis): The population means for cooperation at arrest counts for two independent groups: male and female are not equal.

By comparing significance level at 0.05, the t-test statistics result is 20.816 with p-value of $6.249e-12$. As $6.248e-12 < 0.05$, there exist statistically significant differences. Therefore, we reject the null hypothesis and conclude there are statistical differences between male and female groups with the population mean for cooperation counts.

Year and Cooperation Counts

For year and cooperation counts, we use Welch's T-test to analyze whether the population mean for counts of cooperation at arrest(outcome variable) will differ in two years (2020 and 2021 as explanatory variable). The hypothesis is stated as followed:

H_0 (Null Hypothesis): The population means for cooperation at arrest counts for two independent groups: 2020 and 2021 are equal.

H_1 (Alternative Hypothesis): The population means for cooperation arrest counts for two independent groups: 2020 and 2021 are not equal.

By comparing significance level at 0.05, the t-test statistics result is 0.005 with p-value of 0.996. As $0.996 > 0.05$, there do not exist statistically significant differences. Therefore, we failed to reject the null hypothesis.

Section 4: Method

4.1 ANOVA Tests

4.1.1 One-way ANOVA

One-way Analysis of variance is used when there is one explanatory variable that has two or more levels with only one level of a qualitative outcome. Before the one-way ANOVA is applied to research questions, all assumptions that are listed are checked in detail: 1) normality; 2) equal variance; 3) independence of errors. The paragraphs below will show all the hypotheses for one-way ANOVA tests and detailed interpretation will present in the results section.

Finalized sub-dataset 1

Perceived Race and Strip Search Counts

For perceived race and strip search counts, we use one-way ANOVA to analyze whether the population mean for strip search(outcome variable) actions will differ in race groups(eight(8) different races as explanatory variable). The hypothesis is stated as followed:

H_0 (Null Hypothesis): The population means for strip search counts for eight(8) race groups are equal. ($\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7 = \mu_8$)

H_1 (Alternative Hypothesis): The population means for strip search counts for eight(8) race groups are not all equal.

Finalized sub-dataset 2

Arrest Month and Cooperation Counts

For arrest month and cooperation counts, we use one-way ANOVA to analyze whether the population mean for counts of cooperation attitude at arrest(outcome variable) will differ in

arrest month groups(4 different quarters as explanatory variable). The hypothesis is stated as followed:

H₀(Null Hypothesis): The population means for cooperation at arrest for four(4) arrest month groups are equal. ($\mu_1 = \mu_2 = \mu_3 = \mu_4$)

H₁(Alternative Hypothesis): The population means for cooperation at arrest for four(4) arrest month groups are not all equal.

4.1.2 Two-way ANOVA

Two-way ANOVA is a suitable analysis method with one quantitative outcome and two explanatory variables. Two-way ANOVA can help us know if there exist interactions between two explanatory variables. All two-way ANOVA assumptions are the same as one-way ANOVA and are checked before the use of this method. The paragraphs below will show all the hypotheses for different two-way ANOVA tests and detailed interpretation will present in the results section.

Finalized sub-dataset 1

Perceived Race, Sex and Strip Search Counts

For perceived race, sex and strip search counts, we use two-way ANOVA to analyze whether the population mean for counts of strip search(outcome variable) action will differ in perceived race and sex(8 different races and two sex as explanatory variables). The hypotheses are stated as followed:

Hypothesis 1

H₀(Null Hypothesis): The population means for strip search counts for male and female groups are equal.

H₁(Alternative Hypothesis): The population means for strip search counts at arrest for male and female groups are not equal.

Hypothesis 2

H₀(Null Hypothesis): The population means for strip search counts at arrest for four(4) arrest month groups are equal.

H₁(Alternative Hypothesis): The population means for strip search counts at arrest for male and female groups are not equal.

Hypothesis 3

H_0 (Null Hypothesis): The effect of race does not depend on the effect on sex which means there is no interaction effect of these two variables.

H_1 (Alternative Hypothesis): The effect of race depends on the effect on sex which means there is an interaction effect of these two variables.

Sex, Arrest Month and Cooperation Counts

For sex, arrest month and cooperation counts, we use two-way ANOVA to analyze whether the population mean for counts of strip search(outcome variable) will differ in sex and arrest months(2 different sex and 4 different arrest months as explanatory variables). The hypotheses are stated as followed:

Hypothesis 1

H_0 (Null Hypothesis): The population means for cooperation at arrest for male and female groups are equal.

H_1 (Alternative Hypothesis): The population means for cooperation at arrest for male and female groups are not all equal.

Hypothesis 2

H_0 (Null Hypothesis): The population means for cooperation at arrest for male and female groups are equal.

H_1 (Alternative Hypothesis): The population means for cooperation at arrest for male and female groups are not equal.

Hypothesis 3

H_0 (Null Hypothesis): The effect of sex does not depend on the effect on arrest month which means there is no interaction effect of these two variables.

H_1 (Alternative Hypothesis): The effect of sex depends on the effect on arrest month which means there is interaction effect of these two variables.

4.2 Tukey's HSD

This paper uses Tukey's HSD as post-hoc contrast since Tukey's HSD can find the exact groups that have different means. Tukey's HSD is used for all one-way ANOVA analyses and there are four assumptions related to Tukey's HSD: 1) normality; 2) Equal variance; 3) independence of

errors; 4) approximately equal size for sample size in each group(homogeneity). As the research procedure of two-way ANOVA is based on the one-way ANOVA, the Tukey might not be applied for current 2-way ANOVA but results from Tukey for one-way ANOVA help us infer the results for 2-way ANOVA.

Section 5: Results/ Findings (includes interaction plots)

5.1 Finalized sub-dataset 1

One-way ANOVA

Perceived Race and Strip Search Counts

For perceived race and strip search counts, the test statistics is 1.509 with p-value of 0.212. By comparing significance level at 0.05(alpha = 0.05), with p value of 0.212 is greater than 0.05, there does not exist statistically differences and we fail to reject the null hypothesis that all eight race groups of the population means are equal. Therefore, conclude that differences One-way ANOVA in means for strip search is not caused by the different race groups.

Tukey's Test

Perceived Race and Strip Search Counts

Even if we fail to reject the null hypothesis, we still conduct a Tukey's test to see whether one or more of the groups are rejected. The results of Tukey's test indicate that all groups fail to reject with the significant level of 0.05, this means there does not exist a group with 2 races comparison that has different means in strip counts. Detailed Tukey's HSD test results are attached in the appendix for section 7.1.

Two-way ANOVA

Perceived Race and Strip Search Counts

The Two-way ANOVA test is performed to analyze the effect of perceived race and sex on strip search counts. As table 7 shows the summary statistics of perceived race, sex and strip search counts, by comparing significance level at 0.05, our result indicates that the p-value(0.767) is greater than 0.05, it revealed that there was not a significant statistical interaction between the perceived race and sex. Thus, we fail to reject the null hypothesis 3 that the effect of race does not depend on the effect on race which means there is no interaction effect of these two variables. Additionally, simple main effects analysis showed that perceived race's p-value (0.255) is greater than 0.05 and arrest year's p-value (0.104) is also greater than 0.05, which implies that there were not significant statistical differences on the total strip search counts for different perceived race group, we fail to reject the null hypothesis 1 and hypothesis 2.

	Sum of Squares	df	F	Pr(>F)

Perceived Race	2917095.500	7.0	1.444954	0.255282
Sex	856086.125	1.0	2.968375	0.104175
Perceived Race: Sex	1159032.375	7.0	0.574115	0.766762
Residual	4614436.000	16.0	NaN	NaN

Table 7 2-Way ANOVA statistical result for Perceived Race, sex and Strip Search Counts

Interaction Plot

Perceived Race and Strip Search Counts

As Figure 7 of the interaction plots for Perceived Race, sex and Strip Search Counts shows below, since it was parallel then there were no interactions at any level of the graph, it implies that the effect of perceived race on the strip search counts are same at different levels of the sex. The results are expected, as the P-value from the ANOVA test told us there was not a significant interaction between them.

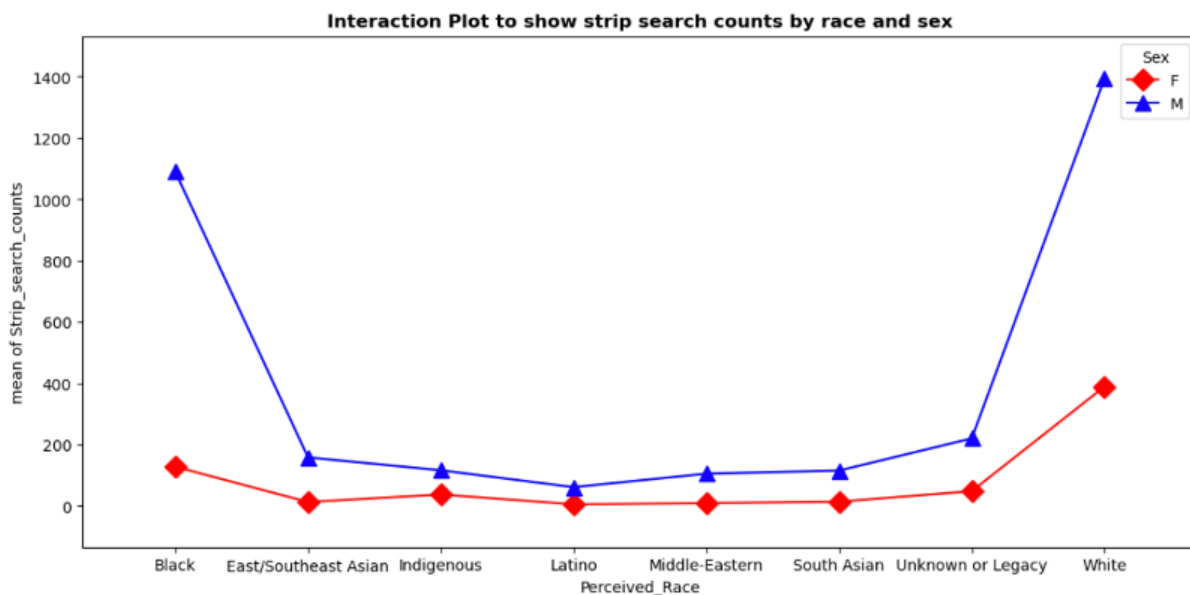


Figure 7 The interaction plots for Perceived Race, sex and Strip Search Counts

5.2 Finalized sub-dataset 2

One-way ANOVA

Arrest Month and Cooperation Counts

For arrest month and cooperation counts, by comparing significance level at 0.05, with p-value(0.983) greater than 0.05, there does not exist statistical differences and we fail to reject the null hypothesis that all the population means for cooperation four arrest month groups are equal. Therefore, conclude that differences in One-way ANOVA in mean for cooperation counts is not caused by the different arrest month groups.

Tukey's Test

Sex, Arrest Month and Cooperation Counts

Even if we fail to reject the null hypothesis, we still conduct a Tukey's test to see whether one or more of the groups are rejected. The results of Tukey's test indicate that all groups fail to reject with the significant level of 0.05, this means there does not exist a group with 2 arrest months comparison that has different means in cooperation at arrest. Detailed Tukey's HSD test results are attached in the appendix for section 7.2.

Two-way ANOVA

Sex, Arrest Month and Cooperation Counts

Two-way ANOVA test was performed to analyze the effect of sex and arrest month on total cooperation counts. As table 8 shows the summary statistics Sex, Arrest Month and Cooperation Counts. By comparing significance level at 0.05, since the p-value(8.91e-01) less than 0.05, it revealed that there was a significant statistical interaction between the Arrest month and sex. Thus, we could reject the null hypothesis that the effect of sex does not depend on the effect on the arrest month which means there is no interaction effect of these two variables and conclude that the effect of sex depends on the effect on the arrest month. Additionally, simple main effects analysis showed that arrest month's p-value (1.77e-01) is less than 0.05 which implies that there was a significant statistical effect on the total cooperation counts, sex's p-value(2.30e-08) is less than 0.05 which also implies that there was a significant statistical effect on the total cooperation counts.

	Sum of Squares	df	F	Pr(>F)
Arrest Month	2.849507e+05	3.0	2.111758	1.770237e-01
Sex	2.080500e+07	1.0	462.555009	2.300125e-08
Arrest Month: Sex	2.741919e+04	3.0	0.203202	8.913329e-01
Residual	3.598275e+05	8.0	NaN	NaN

Table 8 2-Way ANOVA statistical result for Sex, Arrest Month and Cooperation Counts

Interaction Plot

Sex, Arrest Month and Cooperation Counts

As Figure 8 below showed the interaction plots for Arrest month, sex and cooperation counts since it was parallel then there were no interactions which indicate that the plot does not provide the information of the statistically significant differences, it implies that females have a higher mean of total cooperation counts more than male. However, it reverses the analysis results we got from 2-way ANOVA with these two factors. It might be due to the sample differences between two gender groups which will be detailed explained in the conclusion and discussion part. The figure also showed that female had a significantly higher mean of total cooperation counts between January and March.

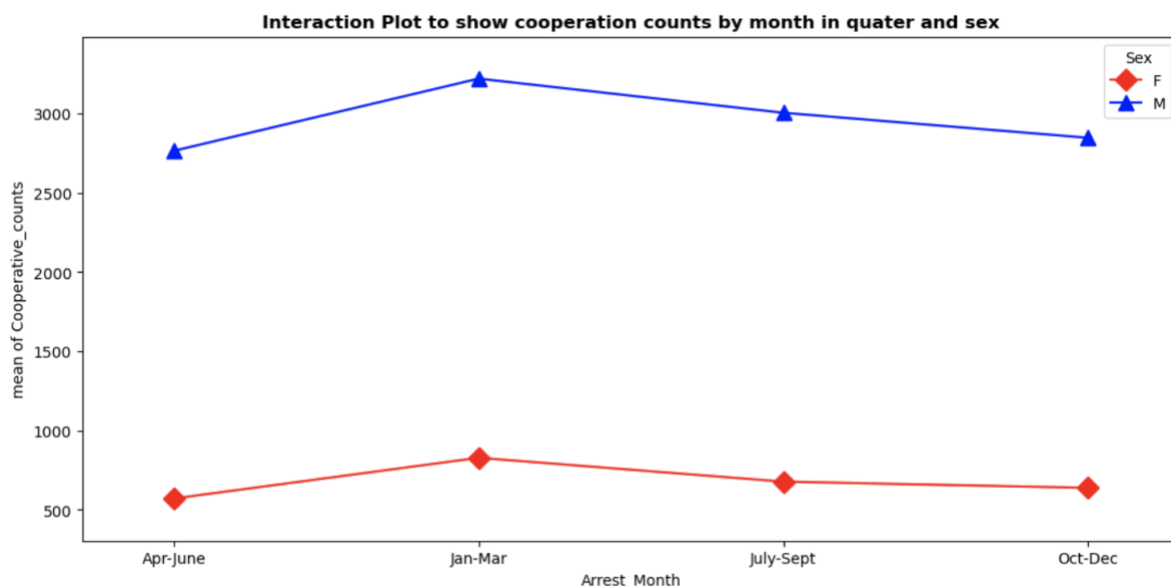


Figure 8 Interaction Plot to show cooperation counts by month in quarter and sex

Section 6: Discussion and Conclusion

The findings suggest that there is no evidence of biased treatments in police actions resulting in strip-searching based on perceived race. Moreover, the results of the analysis show that factors such as year and sex do not have a significant impact on the strip-search frequency. To sum up, in those two years (year 2020 and 2021), there is no strong biased behavior towards police actions for mass and we currently can say there is no extra improvement for police policy for police behavior or actions in Toronto Police Service based on current dataset results.

Regarding Research Question 2, the findings suggest that gender does play a role in cooperation levels during arrests, with males being more cooperative than males. Additionally, the quarterly time also appears to influence people's cooperative attitude during the arrest, with people being more likely to cooperate during the months of January to March. It is often assumed that females

are socialized to be more submissive and inclusive. At the same time, research suggests females are also more likely to be victims of intimate partner violence and may be more afraid of being physically harmed during arrest. However, the result shows male is more cooperative than females, which may be caused by the sample population distribution not representative of the overall population, since the number of male in the sample is significantly higher than that of females or the males in this particular sample are more cooperative. The reason for people being more cooperative during a specific season might link to the weather differences. Findings may be related to seasonal changes in crime rates or changes in social norms and expectations for specific times in the year(quarterly). On the other hand, by comparing yearly, there is no obvious difference for people's attitude toward police. Under the background of 2020 and 2021 with consideration of covid19 policies, people's attitude towards police service remains the same in those two years which might need further research and comparison with before covid19 yearly data. Based on the current result from the dataset, the improvement might be related to add more psychological policies such as creating a more comfortable atmosphere with reducing the tension feelings towards the public. This might need more professional consultation from psychologists.

It is important to note that these findings are based on the available data and the limitations such as the small sample sizes of certain race groups, as well as the lack of specificity in race categories, the abundance of binary data, and severely uneven sample sex groups, which limits the extent to which more complex relationships can be examined of the study, and further research is needed to provide a more comprehensive understanding of the factors that influence police actions and cooperation levels during arrests. By addressing issues of bias and improving police-community relationships will require a multi-faceted approach, including policy changes, training, and community engagement. For example, police departments can review and revise their policies and procedures to ensure that they align with best practices for promoting equity and reducing bias.

Reference

- Owusu-Bempah, A. (2014, January 1). [PDF] *black males' perceptions of and experiences with the police in Toronto: Semantic scholar*. [PDF] Black Males' Perceptions of and Experiences with the Police in Toronto | Semantic Scholar. Retrieved February 23, 2023, from <https://www.semanticscholar.org/paper/Black-Males'-Perceptions-of-and-Experiences-with-in-Owusu-Bempah/627d0f7a6a013bf668593f1c4c0131adbfc5581>
- Toronto Police Service. (2022, November). *Arrests and strip searches (RBDC-arr-TBL-001)*. Arrests and Strip Searches. Retrieved February 23, 2023, from <https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/explore>
- Wortley, S., & Tanner, J. (2003, July). *Data, denials, and confusion: The racial profiling debate in Toronto*. Retrieved February 23, 2023, from <https://utpjournals.press/doi/10.3138/cjccj.45.3.367>

Appendix

Table of attributes

7.1 Tukey's Test of One-way ANOVA for finalized sub-dataset1

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Black	East/Southeast Asian	-523.25	0.8449	-1754.0868	707.5868	False
Black	Indigenous	-532.0	0.8339	-1762.8368	698.8368	False
Black	Latino	-575.5	0.7744	-1806.3368	655.3368	False
Black	Middle-Eastern	-551.5	0.8082	-1782.3368	679.3368	False
Black	South Asian	-544.25	0.818	-1775.0868	686.5868	False
Black	Unknown or Legacy	-474.5	0.8987	-1705.3368	756.3368	False
Black	White	283.0	0.9937	-947.8368	1513.8368	False
East/Southeast Asian	Indigenous	-8.75	1.0	-1239.5868	1222.0868	False
East/Southeast Asian	Latino	-52.25	1.0	-1283.0868	1178.5868	False
East/Southeast Asian	Middle-Eastern	-28.25	1.0	-1259.0868	1202.5868	False
East/Southeast Asian	South Asian	-21.0	1.0	-1251.8368	1209.8368	False
East/Southeast Asian	Unknown or Legacy	48.75	1.0	-1182.0868	1279.5868	False
East/Southeast Asian	White	806.25	0.4029	-424.5868	2037.0868	False
Indigenous	Latino	-43.5	1.0	-1274.3368	1187.3368	False
Indigenous	Middle-Eastern	-19.5	1.0	-1250.3368	1211.3368	False
Indigenous	South Asian	-12.25	1.0	-1243.0868	1218.5868	False
Indigenous	Unknown or Legacy	57.5	1.0	-1173.3368	1288.3368	False
Indigenous	White	815.0	0.3899	-415.8368	2045.8368	False
Latino	Middle-Eastern	24.0	1.0	-1206.8368	1254.8368	False
Latino	South Asian	31.25	1.0	-1199.5868	1262.0868	False
Latino	Unknown or Legacy	101.0	1.0	-1129.8368	1331.8368	False
Latino	White	858.5	0.3285	-372.3368	2089.3368	False
Middle-Eastern	South Asian	7.25	1.0	-1223.5868	1238.0868	False
Middle-Eastern	Unknown or Legacy	77.0	1.0	-1153.8368	1307.8368	False
Middle-Eastern	White	834.5	0.3616	-396.3368	2065.3368	False
South Asian	Unknown or Legacy	69.75	1.0	-1161.0868	1300.5868	False
South Asian	White	827.25	0.372	-403.5868	2058.0868	False
Unknown or Legacy	White	757.5	0.479	-473.3368	1988.3368	False

7.2 Tukey's Test of One-way ANOVA for finalized sub-dataset2

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Apr-June	Jan-Mar	356.25	0.9806	-2433.5857	3146.0857	False
Apr-June	July-Sept	173.25	0.9976	-2616.5857	2963.0857	False
Apr-June	Oct-Dec	74.75	0.9998	-2715.0857	2864.5857	False
Jan-Mar	July-Sept	-183.0	0.9972	-2972.8357	2606.8357	False
Jan-Mar	Oct-Dec	-281.5	0.9902	-3071.3357	2508.3357	False
July-Sept	Oct-Dec	-98.5	0.9996	-2888.3357	2691.3357	False

Google Colab Link for related codes:

<https://colab.research.google.com/drive/1gjcACYcimEO2kfiKxtadO9llpAdU-T67?usp=sharing>