Effect of Different Social Factors on Strip Search and Number of Total Arrest

INF2178
Experimental Design for Data Science

Professor Shion Guha
February 28, 2023

Statement of Individual Contribution:

Muqing Wang: Code, Introduction, Literature Review, EDA, Research Design and Method
Zerui Zhang: Code, Result, Discussion, Conclusion

Introduction **4**
Literature Review **4**
Strip Searches and Gender 4
Number of Arrests of An Individual on Age and Season Factor 4
Number of Arrests of An Individual on Race and Gender 5
Exploratory Data Analysis **5**
Data Cleaning 5
Descriptive Statistics 5
Data Visualization 6
Test of Normality: 9
Anderson-Darling Test 9
T-Test 9
Equal Variance Assumption Checks 9
Gender and Strip Search 9
Age Group and Total Arrest 10
Gender and Total Arrest 10
Research Design and Method **10**
Data Description 10
Research Objective 11
Anova Tests 11
One-Way ANOVA of Sex on Strip Search 11
Two-Way ANOVA of Age Group and Arrest Month on Total Arrest 11
Two-Way ANOVA of Sex and Perceived Race on Total Arrest 12
Post Hoc 12
Homogeneity of Variance Assumption Check 12
Interaction Plot 13
Results **13**
Test of Normality 13
Strip Search 13
Total Arrest 13
Research Question 1 13
Equal Variance Assumption Checks 13
T-Test of Gender and Strip Search 14
One Way Anova of Gender and Strip Search 14
Post Hoc Test 14
Table 1 14
Homogeneity of Variance Assumption check 14
Research Question 2 14
Equal Variance Assumption Checks 14
T-Test of Age Group and Total Arrest 15
Two Way Anova of Age Group and Arrest Month on Total Arrest 15

Table 2 15
Post Hoc Test 15
Table 3 15
Table 4 15
Homogeneity of Variance Assumption check 16
Interaction Plot 16
Research Question 3 16
Equal Variance Assumption Checks 16
T-Test of Gender and Total Arrest 16
Two-Way Anova of Sex and Perceived Race on Total Arrest 16
Table 5 17
Post Hoc Test 17
Table 6 17
Homogeneity of Variance Assumption check 17
Interaction Plot 17
Discussion **18**
Conclusion **18**
Reference **20**

Introduction

Effectiveness, equity and fairness in law enforcement have been long addressed issues in professional research and our daily lives. Law enforcement activities usually draw a lot of social media attention. For example, a police officer killed a suspect George Floyd during an arrest operation. This incident caused questions about effectiveness, equity and fairness. Many people suggested the victim was treated in such a way only because of his race. In order to find out if this concern and issue is happening, we designed three research questions to help us figure it out. Our analysis of strip searches can help us investigate whether there is a potential demographic bias in using strip searches, which can violate individuals' rights and have implications for equity and fairness in law enforcement.

Furthermore, our analysis of the number of total arrests for an individual can provide insights into the patterns of criminal activity and the effectiveness of law enforcement efforts over different demographic groups.

After the analysis using T-tests and ANOVA tests, we find out that police arrest operations indeed exist with some bias in gender, age group, and race.

Literature Review

Strip Searches and Gender

Some research studies have shown a significant difference in the proportion of strip searches conducted on male and female suspects. In general, female individuals have a lower probability of being strip searched than male individuals. Moreover, strip searches have become more extensive. Research 'Rethinking Strip Searches by NSW Police' helped prove the point. As indicated in the article, during the 2016-2017 and 2017-2018 financial years, strip searches rose by 18.5 percent. During the 2014-2015 and 2017-2018 financial years, they rose by 46.8 percent. Furthermore, the chance of females and males receiving a strip search is disproportionate. In the financial year 2016-2017, males have 51.2 percent more chances than females to receive a strip search. Moreover, during the financial year 2017-2018, males have 49.1 percent more chances than females to receive a strip search. (Grewcock & Sentas, 2019)

Number of Arrests of An Individual on Age and Season Factor

Several research studies show a significant difference in crime rate among youth and adults. Moreover, seasonal factors also have a significant impact on the crime rate.

According to "Crime and JusticeVolume 7", the 'age crime curve' indicates that the chance of people conducting criminal activities proliferates to its peak around approximately sixteen years old, followed by a gentle decrease to the age of sixty-four. (The data only provides information from eight to sixty-four years old.) It is almost the same for male and female criminals. (Farrington, 1986)

Furthermore, the 'Crime seasonality and its variations across space' analyze how seasonal patterns may affect crime in Vancouver CMA. Specifically, the article discusses the seasonal impact on crimes, including assaults, robbery, and theft. The plot shows that assault crime happens mainly in the third quarter of the year, which is approximately the end of spring and

the beginning of summer. Most theft crimes happen in the third and fourth quarter of the year, from the middle of summer to the middle of winter. Generally speaking, crime is at its lowest in the first quarter of the year, then proliferates from April to July. After that it remains at its peak from July to October, followed by a rapid decrease. (Andresen et al., 2013)

Number of Arrests of An Individual on Race and Gender

Racism and Sexism has been a rising topic in the past several decades. Especially in sensitive areas like police activities, racism and sexism have drawn much social attention. The book "Gender, race, and crime: An analysis of urban arrest trends" analyzes the effect of gender and race on criminal activities. The authors found that females had lower arrest rates compared to males. The author also found that the male arrest rates were relatively high compared to their percentage in the population. It suggests that males are more likely to participate in criminal activities than females. Furthermore, the authors found that African Americans have a higher arrest rate than their percentage in the whole population. However, arrest rates for white, Hispanic, and Asian people were consistent with their percentage in the population. This may indicate that African Americans are more likely to get arrested than other racial groups. (CHILTON & DATESMAN, 1987)

Exploratory Data Analysis

Data Cleaning

Since the original dataset contains many variables irrelevant to our research questions. Therefore, we first selected several variables, including 'Sex', 'StripSearch', 'Youth_at_arrest__under_18_years', 'Arrest_Month', 'PersonID', 'Perceived_Race' to the working data frame. Then we start cleaning the misleading or inaccurate values. For example, the variable 'Youth_at_arrest__under_18_years' contains values 'Youth (aged 17 and younger)' and 'Youth (aged 17 years and under)', which means the same thing. As a result, we classified those two values as 'Youth' and the rest as 'Adults'. After this, we rename all the values of the 'Arrest_Month' to the first, second, third, and fourth quarters. Furthermore, we created a new column called 'half_year' and divided the four quarters into the first and second half to facilitate our further t-test. Last but not least, we created a new column called 'total_arrest', which calculates the total number of arrests of a suspect. This column is a continuous variable.

Descriptive Statistics

Based on the descriptive statistics we calculated on the new variable 'total_arrest', the mean of the 'total_arrest' is 4.13; on average, a suspect gets arrested 4.13 times from 2020 to 2021. The 25% quantile is 2.00, which means 25% of the values lie below 2.00, and 75% of the answers lie above 2.00. The 75% quantile is 5.00, which means 75% of the values lie below 5.00, and 25% of the answers lie above 5.00. The standard deviation measures how dispersed the data is in relation to the mean. In this case, the standard deviation is 5.40, which means the value of 'total_arrest' is dispersed compared to the mean. Looking at the mean of 4.13 and the median 2.00, we can tell without looking at the histogram that this variable does not follow a normal distribution, and the histogram would be right skewed.

Data Visualization

Figure 1 is the new variable 'total_arrest' histogram. The figure shows that the histogram is very right skewed, which aligns with our supposition in the descriptive statistic that the distribution of 'total_arrest' is right skewed. Furthermore, it is shown in the figure that time of arrests is at its peak at one arrest, followed by a rapid decrease in the number of total arrests.

Figure1: Histogram of total_arrest

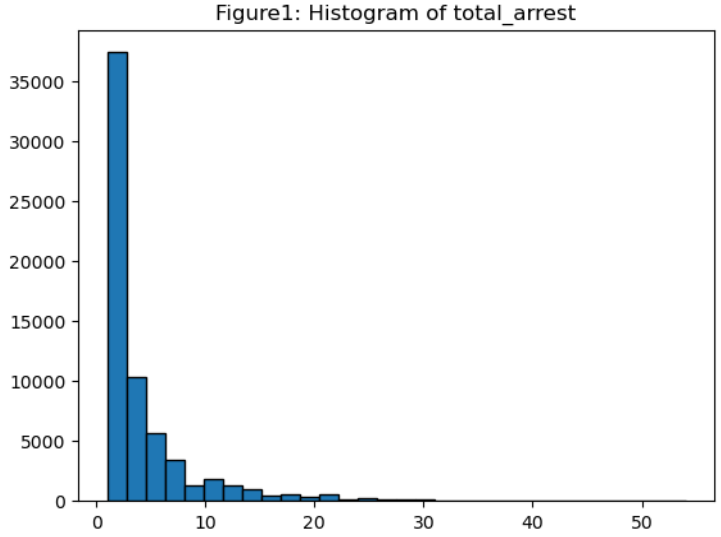


Figure 2 is the box plot of the 'total_arrest'. The box plot shows that the 25% quantile is approximately 1, the median is approximately 2, and the 75% quantile is around 5. This aligns with the result in the descriptive statistics section. Furthermore, the upper bar of the box plot represents the maximum (75% quantile + 1.5IQR). The maximum is approximately 11. However, we observe many outliers that lie beyond the maximum bar, and the highest reaches 54 times of total arrest.

Figure2: Boxplot showing distribution of total_arrest for criminals

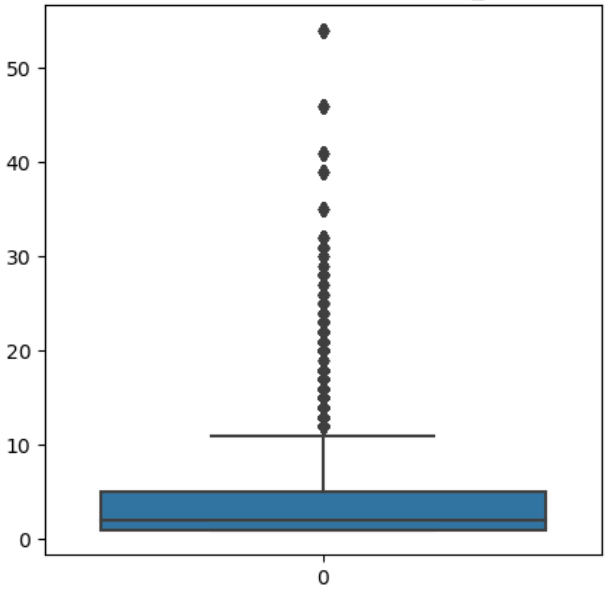


Figure 3 is the violin plot of age group and total_arrest. We split the data into adults and youth; then, we used different colours on the different levels of the 'Arrest_Month' variable.

From the figure, we can see that for adults, the season factor does not have visible effects on total arrest since the median and the quartile range are approximately the same. Furthermore, adults, throughout all months, all have the same highest probability of getting arrested one time. However for youths, in the first quarter of the year, youths have a visible higher probability of getting arrested one time.

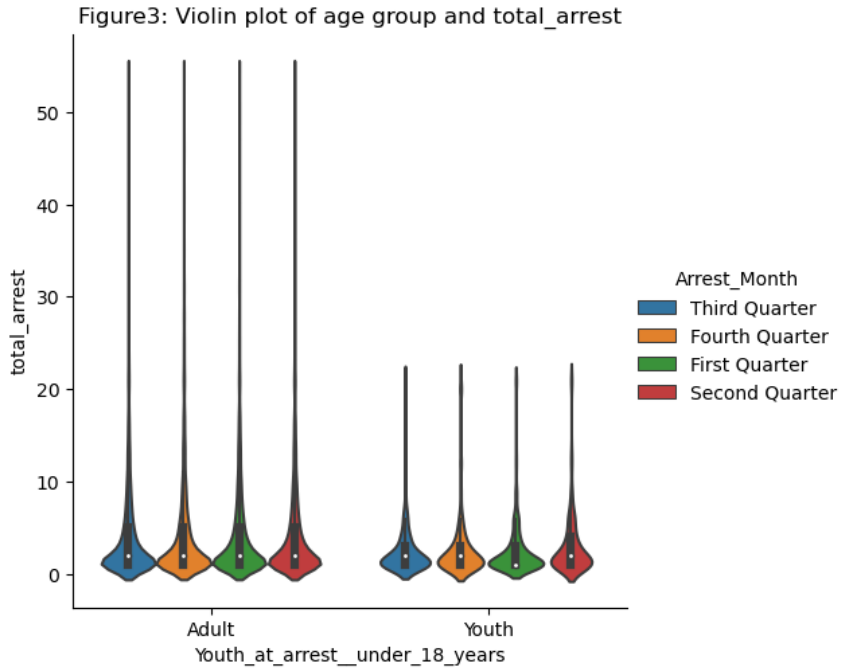


Figure 4 is the box plot of 'Sex' and 'total_arrest'. After seeing the plot, we find that males and females have approximately the same distribution. The only difference is the distribution of outliers. For the unidentified sex, since their percentage in the sample is relatively small compared to the two other sex, it is understandable that the maximum is lower than the other two.

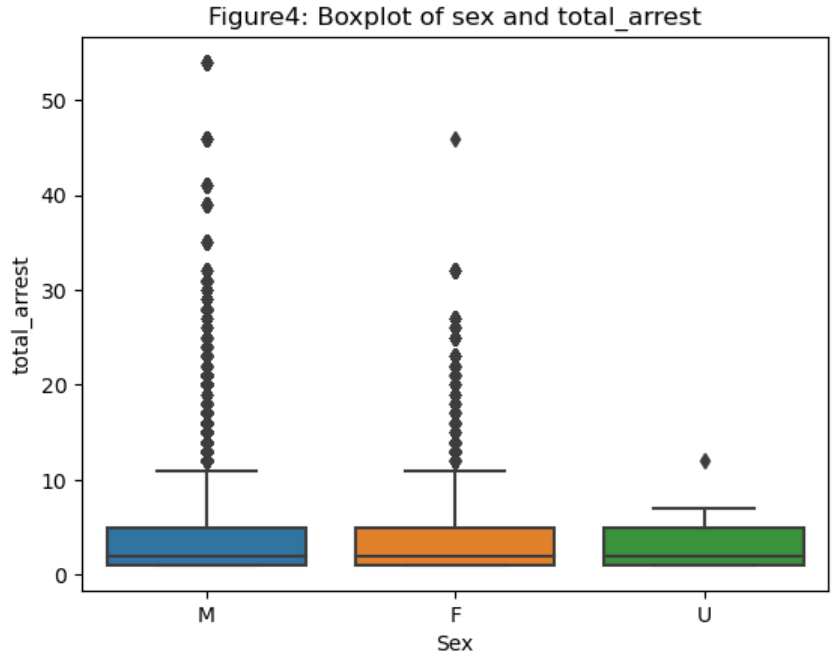


Figure 5 is the box plot of arrest month and total arrest. We see no significant difference among the four quarters; the distributions are almost identical to Figure 2.

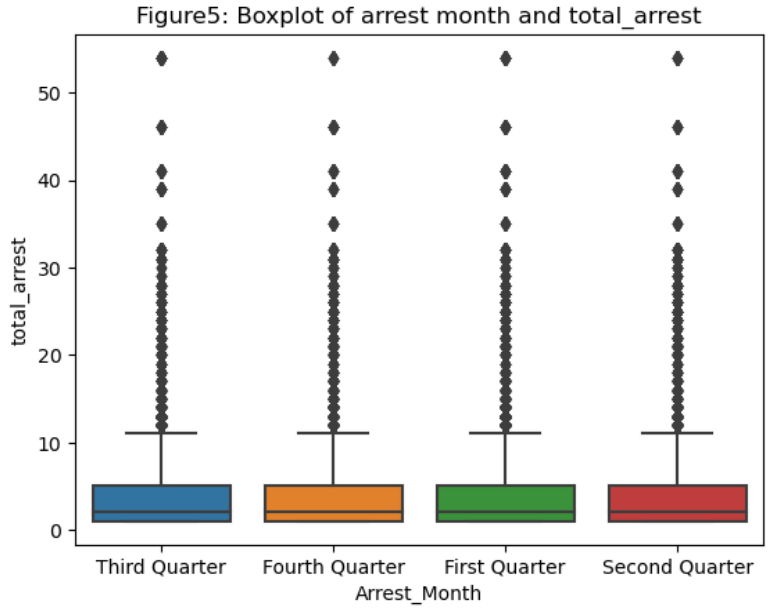
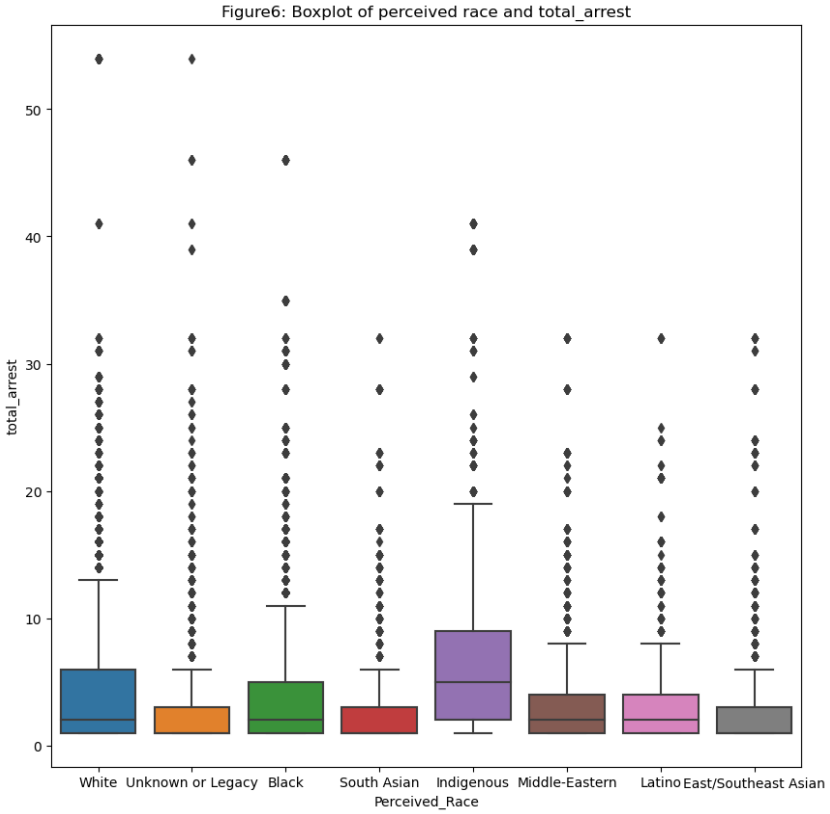


Figure 6 is the box plot of perceived race and total arrest. There are two significant differences compared to Figure 2. The first difference is that indigenous people have a higher total arrest time, with a median of around 5; 25% quantile is 2, and 75% quantile is around 9. From the graph, we can tell that indigenous people are generally arrested. The second difference is that south Asia and east/southeast Asia generally have lower arrest times.



Test of Normality:

Anderson-Darling Test

We will be using the Anderson-Darling test to test the normality of the two different dependent variables, 'StripSearch', and 'total_arrest'. The hypothesis of the Anderson-Darling test is as follows:

H0: The data distribution follows a normal distribution.

H1: The distribution of the data does not follow a normal distribution.

If the p-value is less than our chosen significance level (0.05), the null hypothesis would be rejected, which means there is enough evidence to suggest that the sample does not follow a normal distribution.

It is enough to look at the result of the two dependent variables rather than perform this test on every dependent variable group. This is because the population is a good representation of the sample groups.

T-Test

Equal Variance Assumption Checks

We will use the Levene test to test the equal variance assumption of the different groups before each of our t-tests. The Levene test is a method to determine whether the variances of two groups are equal.

The hypothesis of the Levene test is as follows. (A and B represent two groups in the t-test)

H0: The variance of the two groups is the same. (align with the equal variance assumption)

$(Var_A = Var_B)$

H1: The variance of the two groups is not the same. (the equal variance assumption is violated)

$(Var_A \neq Var_B)$

If the p-value is less than our chosen level of significance (0.05), the null hypothesis would be rejected, and this means that there is enough evidence to suggest that the equal variance assumption is violated.

Gender and Strip Search

We want to analyze if there is any difference in strip searches conducted and gender. Therefore, we performed two samples two-sided Welch t-tests on two different groups: male and female. The hypothesis is as follows:

H0: There is no difference in strip searches conducted between female and male suspects.

$(\mu_M = \mu_F)$

H1: There is a difference in strip searches conducted between female and male suspects.

$(\mu_M \neq \mu_F)$

The t-test calculates the test statistic; this measures the difference between the means of the two samples (male and female) considering the variability within each sample. The t-statistic is compared to critical values from the t-distribution to determine the p-value; the P-value represents, assuming the null hypothesis is true, the probability (chance) of observing a result as extreme as the observed test statistic. Our selection of significance level is 0.05. If the

p-value is less than the significance level 0.05, we would reject the null hypothesis and conclude that the means of the two samples have significant differences.

Age Group and Total Arrest

We want to analyze if there is any difference in the number of total arrests and age group (youth and adults). Therefore, we performed two samples two-sided Welch t-tests on two different groups: youth and adults. The hypothesis is as follows:

H0: There is no difference in the number of total arrests between female and male suspects.

$(\mu_Y = \mu_A)$

H1: There is a difference in strip searches conducted between female and male suspects.

$(\mu_y \neq \mu_A)$

Our selection of significance level is 0.05. If the p-value is less than the significance level 0.05, we would reject the null hypothesis and conclude that there is a difference in searches for the number of total arrests between female and male suspects.

Gender and Total Arrest

We want to analyze if there is any difference in the number of total arrests and different gender groups (male and female). Therefore, we performed two samples two-sided Welch t-tests on two different groups: male and female. The hypothesis is as follows:

H0: There is no difference in the number of total arrests between female and male suspects.

$(\mu_m = \mu_F)$

H1: There is a difference in strip searches conducted between female and male suspects.

$(\mu_M \neq \mu_F)$

Our selection of significance level is 0.05. If the p-value is less than the significance level 0.05, we would reject the null hypothesis and conclude that there is a difference in the number of total arrests between female and male suspects.

Research Design and Method

Data Description

In this project, we will deal with the dataset provided by Toronto Police Service Public Safety Data Portal. The dataset can be found with the following link: https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/about. This dataset contains 65,276 records of arrests in the GTA area from 2020 to 2021. It also contains records and descriptive statistics of strip searches conducted by police officers. A strip search is a type of search performed by a law enforcement officer on an individual that involves the removal of some or all of their clothing and visually inspecting their body. Each of the 65,276 rows in the dataset represents an incident of arrest or strip search performed by the Toronto Police Service. The 24 columns provide information on the demographic and behavioural details of the person arrested or searched. The demographic information includes arrest month, person ID (the suspect's ID), perceived race (the suspect's perceived race), sex, age group, and occurrence category (the nature of the crime). The

behavioural information includes actions at arrest-Combative, actions at arrest-assaulted, actions at arrest-cooperative, and reasons to conduct strip searches. All variables in this data set are either categorical or binary (0 or 1). Due to the fact that the description of this dataset is limited, we assume that the person ID refers to the suspects' ID.

Research Objective

Our study will focus on finding out how gender affects strip searches and how gender, race, arrest month, and age group affect the number of arrests of an individual. Moreover, we want to see if the interaction aligns with the above literature. Therefore, we seek to investigate the following research questions. We gained insights into these questions from the above literature and our preliminary dataset analysis.

- Research Question 1: Is there a significant difference in the proportion of strip searches conducted on male and female individuals in Toronto?
- Research Question 2: Do the total arrests vary significantly by the arrest month and the person's age group, or are total arrests influenced by one or none of these variables?
- Research Question 3: Do the total arrests vary significantly by the perceived race and the sex of the person, or are total arrests influenced by one or none of these variables?

We believe that analyzing the above question can help us and readers understand how arrests are affected by different socioeconomic factors and how suspects' gender influences strip searches.

Anova Tests

One Way Anova of Sex on Strip Search

We want to analyze if there is any difference in strip search conducted and gender. Therefore, we performed One Way Anova in male and female groups and female. The hypothesis is as follows: (Since our independent variable is sex and only has two levels, the hypothesis is the same as the t-test.

H0: There is no difference in strip searches conducted between female and male suspects.

$(\mu_M = \mu_F)$

H1: There is a difference in strip searches conducted between female and male suspects.

$(\mu_M \neq \mu_F)$

If the p-value is less than the significance level 0.05, we would reject the null hypothesis and conclude that the means of the two samples have significant differences.

Two-Way Anova of Age Group and Arrest Month on Total Arrest

We want to analyze if there is any effect of age group and arrest month on the continuous variable total arrest. Therefore, we perform a two-way ANOVA of age group and arrest month on the total arrest. The hypothesis is as follows:

H0: There is no difference in the number of total arrests for any age group. No difference exists in the number of total arrests for any arrest month. There is no interaction effect on age group and arrest months.

H1: There is a difference in the number of total arrests for any age group. There is a difference in the number of total arrests for any arrest month. There is an interaction effect on gender and arrest months.
If the p-value for the age group is less than the significance level 0.05, we would reject the null hypothesis and conclude that the means of the two samples have significant differences.
If the p-value for arrest month is less than the significance level 0.05, we would reject the null hypothesis and conclude that the means of the two samples have significant differences.
If the p-value for the interaction term is less than the significance level 0.05, we would reject the null hypothesis and conclude that there is no interaction effect on the two independent variables.

Two-Way Anova of Sex and Perceived Race on Total Arrest

We want to analyze if there is any effect of perceived race and gender on the continuous variable total arrest. Therefore, we perform a two-way ANOVA of age group and arrest month on the total arrest. The hypothesis is as follows:
H0: There is no difference in the number of total arrests for any gender. There is no difference in the number of total arrests for any race. There is no interaction effect on gender and race.
H1: There is a difference in the number of total arrests for any gender. There is a difference in the number of total arrests for any race. There is an interaction effect on gender and race.
If the p-value for gender is less than the significance level 0.05, we would reject the null hypothesis and conclude that the means of the two samples have significant differences.
If the p-value for race is less than the significance level 0.05, we would reject the null hypothesis and conclude that the means of the two samples have significant differences.
If the p-value for the interaction term is less than the significance level 0.05, we would reject the null hypothesis and conclude that there is no interaction effect on the two independent variables.

Post Hoc

Post hoc tests are conducted after the ANOVA tests to determine which group means are significantly different if the result in ANOVA is significant. We choose to use Tucky's HSD to perform Post Hoc Test. The hypothesis of Tucky's HSD is as follows. (i,j represent different groups)
H0: There is no significant difference in the mean of group i and group j.
$(\mu_i = \mu_j)$
H1: There is a significant difference in the mean of group i and group j.
$(\mu_i \neq \mu_j)$
If the p-value for a specific pair of group i, j is less than the significance level 0.05, we would reject the null hypothesis, and conclude that the means of the two groups have significant differences.

Homogeneity of Variance Assumption Check

We will use the Bartlett test to check the homogeneity of variance assumption.

The hypothesis of the Levene test is as follows. (A and B represent two groups in the t-test)
H0: The variance between the two groups is the same. (align with the equal variance assumption)
($Var_A = Var_B$)
H1: The variance of the two groups is not the same. (the equal variance assumption is violated)
($Var_A \neq Var_B$)
If the p-value is less than our chosen level of significance (0.05), the null hypothesis would be rejected, and this means that there is enough evidence to suggest that the equal variance assumption is violated.

Interaction Plot
Interaction plots will be used to see if there are any interactions between the independent variables. The interaction plot uses the number of total arrests for the y-axis. It consists of two or more lines representing the level of one independent variable. The x-axis is another independent variable. Furthermore, if the lines in the interaction plot are crossed, this represents an interaction between the two independent variables. If the lines in the interaction plot are parallel, there is no interaction between the two independent variables.

Result

Test of Normality

Strip Search
The Anderson-Darling test was used to test whether the data follows a normal distribution. The test statistic was 20664.22, with a corresponding critical value 0.79 (significance level 0.05). Since the test statistic is significantly larger than the critical value of 0.79, we should reject the null hypothesis that the data follows a normal distribution.

Total Arrest
The Anderson-Darling test was used to test whether the data follows a normal distribution. The test statistic was 7688.85, with a corresponding critical value 0.78 (significance level 0.05). Since the test statistic is significantly larger than the critical value of 0.05, we should reject the null hypothesis that the data follows a normal distribution.

Research Question 1
Is there a significant difference in the proportion of strip searches conducted on male and female individuals in Toronto?

Equal Variance Assumption Checks
Prior to conducting a t-test, the assumption of the equality of variances was tested using a Levene test. The results of the Levene test indicated that the variances between the group of strip search to male criminals and the group of strip search to female criminals were

significantly different since the p-value is 6.10*10-12, which is smaller than the significance level 0.05.

T-Test of Gender and Strip Search

A Welch's t-test was conducted to compare the mean scores between the strip searches of male criminals and the strip searches of female criminals. The t-statistic was 7.25 with a p-value less than 0.05, indicating a significant difference between the two groups. Therefore, the strip search of male and female criminals was rejected because there was no difference between the two groups.

One-Way ANOVA of Gender and Strip Search

A one-way ANOVA was conducted to compare the means of the independent group: sex, on a continuous variable: strip search. The one-way ANOVA test revealed at least one gender with a significant difference in strip search between three different sex groups(male, female, unidentified. Since the p-value is smaller than the significance level 0.05. ($F = 24.27$, $p < 0.05$). The results of the ANOVA test indicated a significant effect of sex on strip search ($F = 24.27$, $p < 0.05$). Therefore, the null hypothesis of no significant differences in strip search between the three sex groups was rejected.

Post-Hoc Test

Post-hoc tests using the Tukey HSD test were conducted to determine which sex groups differed significantly—based on the result indicated that the proportion of strip searches on females was significantly different from males, with a p-value < 0.05. No other significant differences were found between the groups. See Table 1.

Table 1

Group1	Group2	Meandiff	P	Lower	Upper	Reject
F	M	0.02	0.0	0.01	0.03	True
F	U	-0.10	0.61	-0.36	0.15	False
M	U	-0.12	0.49	-0.38	0.13	False

Homogeneity of Variance Assumption check

The assumption of equal variances was checked using Levene's test. It was found to be violated ($p < 0.05$).

Research Question 2

Does the total arrests vary significantly by the arrest month and the person's age group, or are total arrests influenced by one or none of these variables?

Equal Variance Assumption Checks

Before conducting a t-test, the assumption of the equality of variances was tested using a Levene test. The results of the Levene test indicated that the variances between the youth and adult groups are significantly different (p = 1.57*10-52).

T-Test of Age Group and Total Arrest

A Welch's t-test was conducted to compare the mean scores between the group of Youth_at_arrest_under_18 to arrest_month. The t-statistic was -27.22, with a p-value less than 0.05, indicating a significant difference between the two groups. Therefore, Youth_at_arrest_under_18 to arrest_month, that there is no difference in the number of total arrests between the two independent variables, was rejected.

Two-Way Anova of Age Group and Arrest Month on Total Arrest

A two-way ANOVA was conducted, and the results are in Table 2. The ANOVA revealed a significant main effect of Youth_at_arrest_under_18 (F = 253.15, p = 6.81*10^-57), a significant main effect of arrest_month (F = 3.10, p = 2.57*10^-02). and a insignificant effect between Youth_at_arrest_under_18 and arrest_month (F = 0.99, p = 3.98*10^-01). Based on these results, we reject the null hypothesis for both Youth_at_arrest_under_18 and arrest_month, indicating a significant difference in mean scores between the groups for both factors. We cannot reject the null hypothesis that no interaction exists between Youth_at_arrest_under_18 and arrest-month. The two independent variables are independent of each other.

Table 2

	sum_sq	df	F	P
Youth_at_arrest_under_18	7.36*10^3	1.0	253.16	6.81*10^-57
Arrest_month	2.70*10^2	3.0	3.10	2.57*10^-2
Youth_at_arrest_under_18:Arrest_month	8.61*10^1	3.0	0.99	3.98*10^-1
Residual	1.90*10^6	65268.0	NaN	NaN

Post-Hoc Test

To further investigate the differences among groups, we conducted post-hoc tests using the Tukey's HSD method separately for Youth and arrest month. See Table 3 for youth at arrest and table 4 for arrest month. For the arrest month, we found significant differences between the fourth and second quarters (mean difference = 0.19, $p < 0.05$), and no other significant differences were found between the groups. For youth at arrest, we found a significant difference between adults and youth (mean difference = -1.60, $p < 0.05$).

Table 3

Group1	Group2	Meandiff	P	Lower	Upper	Reject

Adult	Youth	-1.60	0.0	-1.79	-1.40	True

Table 4

Group1	Group2	Meandiff	P	Lower	Upper	Reject
First Quarter	Fourth Quarter	-0.10	0.33	-0.25	0.05	False
First Quarter	Second Quarter	0.088	0.46	-0.07	0.24	False
First Quarter	Third Quarter	-0.04	0.92	-0.19	0.11	False
Fourth Quarter	Second Quarter	0.19	0.01	0.03	0.35	True
Fourth Quarter	Third Quarter	0.06	0.72	-0.09	0.22	False
Second Quarter	Third Quarter	-0.13	0.16	-0.28	0.03	False

Homogeneity of Variance Assumption check

To check the assumption of homogeneity of variance, Bartlett's test was conducted for each factor and their interaction. For both Youth_at_arrest_under_18 and arrest_month, Bartlett's test was significant ($p < 0.05$), suggesting that the assumption of homogeneity of variance was violated for both factors.

Interaction Plot

An interaction plot was generated to visualize the relationship between total arrest, arrest time, and youth at arrest. The plot revealed a parallel interaction between youth and adults, indicating no interaction between arrest_month on total arrests for youth and adult criminals.

Research Question 3

Does the total arrests vary significantly by the perceived race and the sex of the person, or are total arrests influenced by one or none of these variables?

Equal Variance Assumption Checks

Prior to conducting a t-test, the assumption of the equality of variances was tested using a Levene test. The results of the Levene test indicated that the variances between the group of Sex and total_arrest were significantly different ($p = 1.57*10^{-52}$).

T-Test of Gender and Total Arrest

A Welch's t-test was conducted to compare the mean scores between the group of Sex and total_arrest. The t-statistic was 5.5, with a p-value less than 0.01, indicating a significant difference between the two groups. Therefore, Sex to total_arrest, that there is no difference between the two groups was rejected.

Two-Way Anova of Sex and Perceived Race on Total Arrest

A two-way ANOVA was conducted, and the results are presented in Table 3. The ANOVA revealed a significant main effect of Sex (F = 43.05, p = 2.06*10^-19), a significant main effect of Race (F = 165.19, p = 1.11*10^-106), and a significant interaction between Sex and Race (F = 12.33, p =1.14*10^-21).

Based on these results, we reject the null hypothesis for both Sex and Race, indicating a significant difference in mean scores between the groups for both factors. The significant interaction between Sex and Race suggests that the effect of one factor on the outcome variable depends on the level of the other factor.

Table 5

	sum_sq	df	F	P
Sex	2.43*10^3	2.0	43.05	2.06*10^-19
Perceived_Race	3.26*10^4	7.0	165.19	1.11*10^-106
Perceived_Race:Sex	4.87*10^3	14.0	12.33	1.14*10^-21
Residual	1.84*10^6	65253.0	NaN	NaN

Post Hoc Test

To further investigate the differences among groups, we conducted post-hoc tests using the Tukey's HSD method separately for Perceived race and sex. See Table 5 for sex and table 6 for perceived race. For sex, we found significant differences between the females and males (mean difference = 0.28, $p < 0.05$), and no other significant differences were found between the groups.

Table 6

Group1	Group2	Meandiff	P	Lower	Upper	Reject
F	M	0.28	0.0	0.15	0.40	True
F	U	-0.24	0.99	-4.46	3.98	False
M	U	-0.52	0.9551	-4.74	3.70	False

Homogeneity of Variance Assumption check

To check the assumption of homogeneity of variance, Bartlett's test was conducted for each factor and their interaction. For Sex, Bartlett's test was significant ($p < 0.05$), suggesting that the assumption of homogeneity of variance was violated. For Race, Bartlett's test was insignificant, suggesting that the assumption of homogeneity of variance was not violated.

Interaction Plot

An interaction plot was generated to visualize the relationship between total arrest, perceived race, and sex. The plot revealed a significant crossover interaction between total arrested and

perceived race, indicating that the effect of perceived race on total arrested depends on different race groups.

Discussion

In this study, there were several limitations. The two dependent variables, strip searches and the number of total arrests do not follow a normal distribution. However, the normally distributed dependent variable is a key assumption of t-tests and ANOVA tests. This may affect the accuracy of our analysis. In our research question one: Is there a significant difference in the proportion of strip searches conducted on male and female individuals in Toronto? We had limitations about the study that may only consider a limited time frame, which may not capture the entire picture. Also, in our study, we did not account for other variables that may influence the occurrence of strip searches, such as the offence's severity or the individual's behaviour during the arrest. In research question two: Do the total arrests vary significantly by the arrest month and the person's age group, or are total arrests influenced by one or none of these variables? We faced limitations. The study did not account for other variables that may influence the total number of arrests, such as the location of the arrest or the type of offence. In research question three: Do the total arrests vary significantly by both the perceived race and the sex of the person, or are total arrests influenced by one or none of these variables? The study may not account for other variables that may influence the total number of arrests, such as the location of the arrest or the type of offence. Additionally, the study may rely on self-reported data, which may not accurately reflect the perceived race or sex of the individuals involved.

Here are some improvements we came up with that will be helpful for future studies. Future studies could include a larger sample size to increase the generalizability of the findings. Also, researchers could collect additional information on the circumstances surrounding each strip search to understand better the factors that may influence the occurrence of strip searches. Future studies could include a more comprehensive set of variables to understand better the factors influencing the number of arrests. Additionally, researchers could use a larger sample size to increase the generalizability of the findings. Future studies could use a more comprehensive set of variables to understand better the factors that influence the total number of arrests. Researchers could collect data from multiple sources to ensure the accuracy of the perceived race and sex of the individuals involved. Moreover, researchers could use statistical methods to control for potential confounding variables, such as the type of offence, to understand better the relationship between perceived race/sex and total arrests.

Conclusion

The present study aimed to investigate three research questions related to strip searches, arrests by age group and month, and arrests by perceived race and sex. Our findings shed light on important aspects of policing practices and offer insights into potential disparities that must be addressed.

First, our results showed a significant difference in the proportion of strip searches conducted on male and female individuals in Toronto. Specifically, our t-test result revealed that the proportion of strip searches conducted on males and females significantly differs. This finding is consistent with previous research on strip searches and raises concerns about potential gender-based biases in policing practices.

Second, our study also examined the relationship between total arrests, the arrest month, and the person's age group. Our t-test and two-way ANOVA results indicated a significant variation in the number of arrests by month and age group. Specifically, our analysis revealed that the number of arrests was higher in certain quarters, such as the first and second quarters, specifically in adults. These findings suggest that policing practices may be influenced by factors beyond the specific behaviours of individuals.

Finally, our study examined the relationship between total arrests and both the perceived race and sex of the person. Our results from the t-test and two-way ANOVA results indicated a significant variation in the total number of arrests by perceived race and sex. Specifically, our analysis revealed that individuals perceived as Black, Asian and White were arrested at a higher rate than those other perceived race groups. Black had the highest rate of the arrest. These findings highlight racial biases in policing practices and the need for further investigation and action to address them.

Our study contributes to the growing body of research on policing practices and their impact on equity and fairness for various groups of individuals. Our findings suggest that disparities in policing practices exist and may be influenced by various factors, including gender, age, and race. These findings call for greater attention to and action on issues related to policing practices and their impact on marginalized communities. Further research is needed to explore these disparities' underlying causes and develop evidence-based strategies to address them.

Reference

Grewcock, M., & Sentas, V. (2019, August 22). *Rethinking strip searches by NSW Police*. Michael Grewcock, Vicki Sentas. Retrieved February 25, 2023, from https://apo.org.au/node/254776?utm_source=APO-feed&utm_medium=RSS&utm_campaign=rss-all

Farrington, D. P. (1986). *Age and crime: Crime and justice: Vol 7*. University of Chicago Press Journals. Retrieved February 26, 2023, from https://www.journals.uchicago.edu/doi/abs/10.1086/449114

Andresen, M. A., & Malleson, N. (2013, July 4). *Crime seasonality and its variations across space*. Applied Geography. Retrieved February 26, 2023, from https://www.sciencedirect.com/science/article/pii/S0143622813001410?casa_token=d6fGY0Y1Iw0AAAAA%3AjLXkxmBgo6pTtqE3jZfxGcX0TCX0c1u__9MfCkkZ8L9l_vw2RfwzQ9KD3PbKnEuZlHKSdlxY1IEq

CHILTON, R. O. L. A. N. D., & DATESMAN, S. U. S. A. N. K. (1987). Gender, race, and crime: *Gender & Society*, *1*(2), 152–171. https://doi.org/10.1177/0891243287001002003