INF1343 Group 16 Writeup

Rongjia Bao, Xiaozhou (Andy) Ye

2023-02-22

Contents

1. Introduction

The criminal justice system is intended to maintain public safety and social order. However, there are concerns regarding possible biases and discrimination in the way it functions, particularly in the treatment of individuals during the arrest and booking process. One specific area of concern is the use of strip searches, which can be incredibly invasive and humiliating for those subjected to them. There have been more than 40 federal lawsuits filed against Oakland and its police officers, alleging an illegal policy of strip-searching people in public (Nick, 2011).

This data analysis project seeks to investigate two research questions relating to strip searches during arrests. Firstly, we aim to examine whether there is a significant difference in the average number of strip searches conducted by race. Secondly, we will investigate whether there are significant differences in the average number of strip searches when arresting individuals of varying genders and age groups. Research questions are stated at the end of the EDA section.

The analysis will use a dataset obtained from the Toronto Police Service Public Safety Data Portal, which contains information on the number of strip searches conducted during arrests, as well as the race, gender, and age group of the individuals searched. The dataset covers multiple years and features a large sample of individuals from diverse backgrounds.

The findings of this project will provide important insights into any potential disparities in the use of strip searches during arrests. They may have significant implications for efforts to reform the criminal justice system and promote fairness and equity in law enforcement practices.

2. Exploratory Data Analysis

This dataset was obtained from the Toronto Police Service Public Safety Data Portal and contains information on all arrests and strip searches. It consists of 65,276 samples and 25 variables. However, after cleaning the dataset, the total number of observations reduced to 224 rows and 6 variables. A glimpse of the dataset is demonstrated below:

Table 1: Post-cleaning dataset

Year	Race	Sex	AgeGroup	Youth	StripSearch	Arrest
2020	Black	F	Aged 17 years and under	Youth	11	195
2020	Black	F	Aged 18 to 24 years	Not a youth	78	366
2020	Black	F	Aged 25 to 34 years	Not a youth	79	494

The dataset provides information on the number of individuals who were strip-searched and arrested for each **Year**, **Race**, **Sex**, and **Age Group** combination during the years 2020 and 2021. The races included in the dataset are Black, East/Southeast Asian, Indigenous, Latino, Middle-Eastern, South Asian, White, and Unknown or Legacy. For the purpose of this study, genders other than male or female were excluded, and the age groups included are Aged 17 years and under, Aged 18 to 24 years, Aged 25 to 34 years, Aged 35 to 44 years, Aged 45 to 54 years, Aged 55 to 64 years, and Aged 65 years and older. The dependent variables in this study are **StripSearch** and **Arrest**, which are the only two continuous variables in this table. The variable **Youth** was added as an independent variable to support the analysis, based on the age of the individual.

2.1 Descriptive Statistics

Descriptive analysis is a statistical method used to summarize and describe the main characteristics of a dataset. As the dependent variables of interest in this study are **StripSearch** and **Arrest**, descriptive statistics were performed to gain a better understanding of these variables. Table 2 shows that both variables have a relatively small mean compared to the range and high standard deviation. This may suggest the presence of outliers within the dataset.

Table 2: Descriptive Statistics

	Strip Search	Arrest
N	224	224
Median	3	87.5
Mean	34.82	291.25
Std. Deviation	108.56	549.73
Range	858	3225

2.2 Data Visualization

To investigate the underlying relationships between the potential explanatory variables and the dependent variables, several diagrams were plotted. One common diagram used when it comes to dealing with catagorical data is boxplot. A boxplot is a visual tool used to represent the distribution of a numerical variable. It provides a clear picture of the central tendency, spread, and any possible outliers of the data. Figure 1 depicts a boxplot representing the number of StripSearch for each perceived race group.

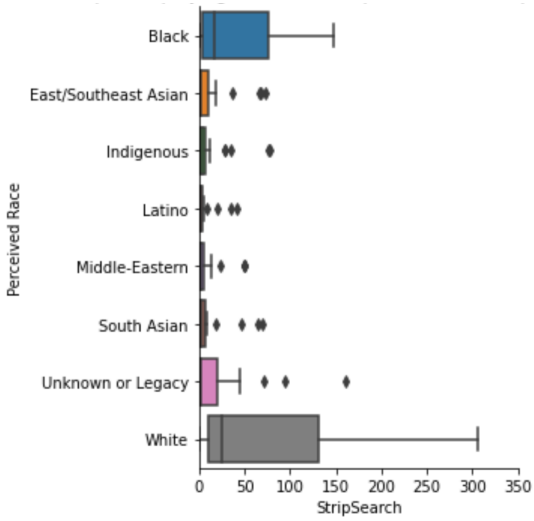

Figure 1: Boxplot displaying number of stripsearch for each race

From the plot, it is apparent that White individuals have the highest median value, which is indicated by the horizontal line within the box. The boxplot's width for the White group is also the widest, indicating the largest spread of data within that group. The whiskers, which extend from the box, are longest for the White group as well, further demonstrating their wider distribution of values.

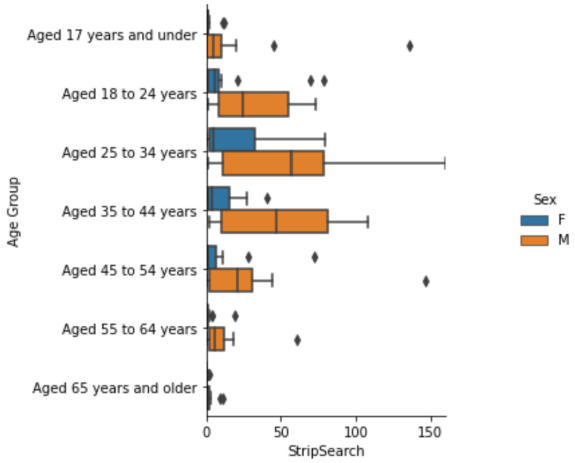

Figure 2: Boxplot displaying number of stripsearch for each age group

Figure 2 illustrates the distribution of the number of strip searches by age groups. The age group between 25 to 34 years old has the highest median value, which is represented by the horizontal line inside the box. The boxplot also shows that this age group has the widest box and longest whiskers, indicating the largest spread of data among all age groups. On the other hand, the

age group of 65 years and older has the narrowest box and shortest whiskers than other age groups, indicating the smallest spread of data. This age group has the tightest distribution of values, which is reflected in the boxplot.

The boxplot also shows the presence of outliers, which are data points that are significantly different from the rest of the group. Outliers are shown as individual points that are outside of the whiskers. For both of these plots, all age groups have outliers, except some were removed to make the visualization more appealing. Overall, these boxplot provides a clear and concise representation of the distribution of strip search data among different age groups. They highlight the differences in the central tendency and spread of the data, as well as the presence of any outliers.

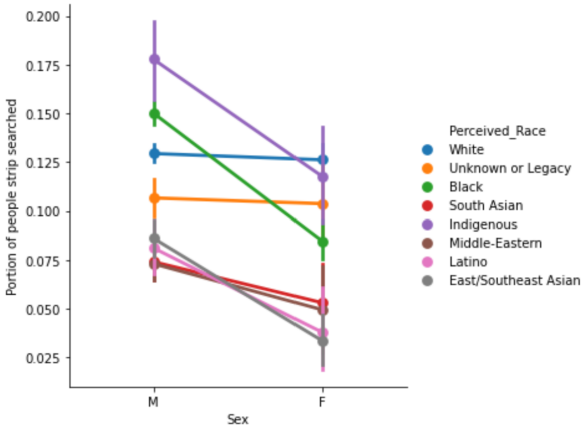

Figure 3: Pinplot displaying portion of people being strip searched

A point plot is a type of data visualization used to display the relationship between variables. In a portion point plot, there are two categorical variables: perceived race and sex. Each point on the plot represents the proportion or percentage of one variable within a level of the other variable. For instance, the proportion of males who were strip-searched within each perceived race group or the proportion of females who were strip-searched within each perceived race group. The position of the point on the y-axis indicates the proportion or percentage, while the position on the x-axis indicates the category of the variable. In this point plot, the portion of people who were strip-searched for each perceived race is displayed by sex. Upon observation, it is apparent that there are significant differences in the proportion of individuals who were strip-searched across sex and perceived race. Specifically, Latinos and Blacks have the largest differences in the portion of being strip-searched across sex, with higher proportions of males in both groups being strip-searched compared to females. In contrast, for the white group, the proportions of individuals who were strip-searched are about the same across sex.

Overall, this point plot provides a clear and concise representation of the relationship between

perceived race, sex, and the proportion of individuals who were strip-searched. It highlights the differences in the proportion of individuals who were strip-searched between different demographic groups.

2.3 Hypothesis Test

A t-test is a statistical test used to determine if there is a significant difference between the means of two groups of data. It is a parametric test that assumes the data is normally distributed and the variances of the two groups are equal. In this data analysis, a t-test was employed to investigate whether the number of strip searches and arrests differed significantly by race, gender, year, and whether they were youth (aged 17 or below). The statistical significance level, also known as alpha, was set to 0.05. The results of the t-test provide a p-value, which indicates the probability of obtaining the observed difference between the two groups by chance alone. If the p-value is less than the significance level, typically set at 0.05, then we can conclude that there is a significant difference between the means of the two groups.

2.3.1 T-test of StripSearch between sex

In this section, a t-test was performed to compare the mean number of strip searches between male and female sex groups. To achieve this, we formulated the following null and alternative hypotheses:

H_0 : The number of StripSearch upon arrest is the same for Men and Women

H_1 : The number of StripSearch upon arrest is NOT the same for Men and Women

Table 3: T-test of StripSearch between sex

	t	df	p-value
StripSearch	3.292	223	0.001

The t-test results indicate that there is a statistically significant difference in the mean number of strip searches between male and female sex groups. The t statistic of 3.292 and the p-value of 0.001 suggest that the observed difference in means is unlikely to have occurred by chance. Since the p-value (0.001) is smaller than the significance level (0.05), we can reject the null hypothesis and conclude that the strip search differs significantly between men and women.

2.3.2 T-test of Arrest between sex

Similarly, for comparing the mean number of arrests between male and female sex groups, the following null and alternative hypotheses were formulated:

H_0 : The number of Arrest is the same for Men and Women

H_1 : The number of Arrest is NOT the same for Men and Women

Table 4: T-test of Arrest between sex

	t	df	p-value
Arrest	5.134	223	< 0.001

The t-test results indicate that there is a statistically significant difference in the mean number of strip searches between male and female sex groups. The t statistic of 5.134 and the p-value of less than 0.05 suggest that we can reject the null hypothesis and conclude that the number of arrests differs significantly between men and women.

2.3.3 T-test of StripSearch between year

For comparing the mean number of strip searches between year 2020 and 2021, the following null and alternative hypotheses were formulated:

H_0 : The number of StripSearch upon arrest is the same for 2020 and 2021

H_1 : The number of StripSearch upon arrest is NOT the same for 2020 and 2021

Table 5: T-test of StripSearch between year

	t	df	p-value
Arrest	4.094	223	< 0.001

The t-test results indicate that there is a statistically significant difference in the mean number of strip searches between years. The p-value of less than 0.001 suggest that we can reject the null hypothesis and conclude that the strip search differs significantly between 2020 and 2021.

2.3.4 T-test of Arrest between year

For comparing the mean number of arrests between year 2020 and 2021, the following null and alternative hypotheses were formulated:

H_0 : The number of Arrest is the same for 2020 and 2021

H_1 : The number of Arrest is NOT the same for 2020 and 2021

Table 6: T-test of Arrest between year

	t	df	p-value
StripSearch	-0.16	223	0.873

The t-test results indicate that there is no statistically significant difference in the mean number of arrests between years. The p-value of 0.873 suggest that we do not have enough evidence to reject the null hypothesis.

2.3.5 T-test of StripSearch between youth and adult

For comparing the mean number of strip searches between youth and adult, the following null and alternative hypotheses were formulated:

H_0 : The number of StripSearch upon arrest is the same for youth and adult

H_1 : The number of StripSearch upon arrest is NOT the same for youth and adult

Table 7: T-test of StripSearch between youth and adult

	t	df	p-value
StripSearch	-1.471	223	0.143

The t-test results indicate that there is no statistically significant difference in the mean number of strip searches between youth and adult. The p-value of 0.143 suggest that we do not have enough evidence to reject the null hypothesis.

2.3.6 T-test of Arrest between youth and adult

For comparing the mean number of arrests between youth and adult, the following null and alternative hypotheses were formulated:

H_0 : The number of Arrest is the same for youth and adult

H_1 : The number of Arrest is NOT the same for youth and adult

Table 8: T-test of Arrest between youth and adult

	t	df	p-value
Arrest	-2.199	223	0.029

The t-test results indicate that there is a statistically significant difference in the mean number of arrests between youth and adult. The p-value of less than 0.029 suggest that we can reject the null hypothesis and conclude that the arrest differs significantly between youth and adult.

2.3.7 T-test of SrtripSearch among black and non-black

To investigate whether there is a significant difference in the mean number of strip searches between black and non-black populations, the following null and alternative hypotheses were formulated:

H_0 : The number of StripSearch is the same for black and non-black

H_1 : The number of StripSearch is NOT the same for black and non-black

Table 9: T-test of StripSearch between black and non-black

	t	df	p-value
StripSearch	1.453	223	0.148

The t-test results indicate that there is no statistically significant difference in the mean number of strip searches between youth and adult. The p-value of 0.148 suggest that we do not have enough evidence to reject the null hypothesis.

2.4 Research Questions

Based on the investigations, we have identified several explanatory variables that may have an impact on the dependent variables, specifically the strip search. Specifically, we have found that sex and age group are potentially important factors, as well as perceived race. In light of these findings, we have formulated the following research questions:

- Does the perceived race of an individual significantly affect his/her likelihood of being strip searched when arrested?
- Do sex and the age group of an individual significantly affect his/her likelihood of being strip searched when arrested?

By exploring these research questions, we hope to gain a deeper understanding of how strip search is influenced by various demographic factors, and to identify any potential disparities or inequalities that may exist within the criminal justice system.

3. Method

3.1 ANOVA

ANOVA, which refers to Analysis of Variance, is a statistical technique used to determine whether there are significant differences among the means of two or more groups of data. The ANOVA method compares the variation between groups with the variation within groups to determine whether the differences in means are statistically significant or are likely to have occurred by chance.

The null hypothesis in ANOVA is that there is no significant difference among the means of the groups being compared. The alternative hypothesis is that there is at least one significant difference among the means.

The ANOVA technique computes an F statistic, which is the ratio of the variation between groups to the variation within groups. This is demonstrated in the equation below:

$$F = \frac{\mathrm{MS}_{\mathrm{between}}}{\mathrm{MS}_{\mathrm{within}}}$$

where MS refers to mean squared and is equal to the sum of squares divided by the degrees of freedom. If the F-statistic is large enough, it suggests that there are significant differences between the groups. It is a powerful tool for analyzing data and drawing meaningful conclusions about the differences between groups.

ANOVA can be categorized into two types, namely one-way ANOVA and two-way ANOVA. One-way ANOVA is utilized to determine whether there are significant differences in the means of a continuous outcome variable among two or more groups. It compares the variability between groups with the variability within groups and assumes that the groups are independent, the populations have equal variances, and the data are normally distributed. The null hypothesis of one-way ANOVA is that there are no significant differences between the means of the groups being compared. If the null hypothesis is rejected, it implies that there is a significant difference between at least two of the groups.

On the other hand, two-way ANOVA is an extension of one-way ANOVA that incorporates two independent variables, each with two or more levels, and their interaction effect on the outcome variable. It can be used to examine the significance of the differences in means between groups that are defined by two factors. The null hypothesis of two-way ANOVA is that there is no significant interaction effect between the two independent variables on the outcome variable. If the null hypothesis is rejected, it indicates that there is a significant interaction effect between the two independent variables on the outcome variable.

3.2 Tukey's HSD

Tukey's HSD, which stands for Honestly Significant Difference, is a type of post-hoc statistical test that is commonly utilized to determine if there are significant differences between the means of multiple groups. This test is frequently applied in an analysis of variance (ANOVA) when the null hypothesis has been rejected, indicating that there is a substantial variation between at least two of the group means.

4. Results

4.1 One-way ANOVA for the effect of Race on StripSearch

To investigate the relationship between perceived race and strip search using analysis of variance, the following hypotheses were formulated:

H_0 : There is NO significant difference in the likelihood of strip search between perceived race.

H_1 : There is a significant difference in the likelihood of strip search between perceived race.

These hypotheses can be tested using an ANOVA to compare the means of the perceived race groups and determine if there is a statistically significant difference between them in terms of the likelihood of being subjected to a strip search. If the null hypothesis is rejected, it would

indicate that there is evidence to support the alternative hypothesis, suggesting that perceived race does have a significant effect on the likelihood of a strip search.

Table 10: One-way ANOVA for Race and StripSearch

	Sum of Squares	df	F	Pr(>F)
C(Race)	416727.9	7	5.815	0.000003
Residual	2211421	216	NaN	NaN

The F-statistic assesses whether the means of the perceived race groups differ more than would be expected by chance. A high F-statistic, along with a low p-value, suggests strong evidence for a significant difference in the means of the perceived race groups.

With a p-value less than 0.001, it is unlikely that the observed differences in the means of the perceived race groups were due to chance. This implies that the null hypothesis, which proposes no significant difference in the likelihood of strip search between perceived race, can be rejected. Thus, the alternative hypothesis, which asserts that there is a significant difference in the likelihood of strip search between perceived race, is supported.

Consequently, the one-way ANOVA result indicates that perceived race significantly affects the likelihood of a strip search. Further examination, like post-hoc tests, may be conducted to identify which perceived race group(s) significantly influence the likelihood of a strip search. The output of Tukey's Honestly Significant Difference (HSD) test, which is used to identify which pairs of groups have significantly different means in a one-way ANOVA, is demonstrated on the next page.

Figure 4 displays pairwise comparisons between groups, including mean differences, p-values, and confidence intervals. If the p-value is less than the significance level, usually 0.05, the mean differences are considered statistically significant, and the corresponding pairs are marked as "True" in the rightmost column. For instance, the final row indicates that the average probability of a strip search for the "Unknown or Legacy" perceived race group significantly differs from the mean likelihood for the "White" perceived race group, as reflected by the low p-value of 0.0022. Conversely, the high p-value of 0.1099 implies that there is no significant difference in the mean likelihood of a strip search for the "Black" and "East.Southeast Asian" groups.

Note that "False" entries in the rightmost column do not necessarily imply no difference between the corresponding group means. It merely indicates that the difference is not statistically significant by Tukey's HSD test.

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Black	East.Southeast Asian	-74.75	0.1099	-157.5223	8.0223	False
Black	Indigenous	-76.0	0.0977	-158.7723	6.7723	False
Black	Latino	-82.2143	0.053	-164.9866	0.558	False
Black	Middle-Eastern	-78.7857	0.0752	-161.558	3.9866	False
Black	South Asian	-77.75	0.0831	-160.5223	5.0223	False
Black	Unknown or Legacy	-67.7857	0.1981	-150.558	14.9866	False
Black	White	40.4286	0.7846	-42.3437	123.2009	False
East.Southeast Asian	Indigenous	-1.25	0.9	-84.0223	81.5223	False
East.Southeast Asian	Latino	-7.4643	0.9	-90.2366	75.308	False
East.Southeast Asian	Middle-Eastern	-4.0357	0.9	-86.808	78.7366	False
East.Southeast Asian	South Asian	-3.0	0.9	-85.7723	79.7723	False
East.Southeast Asian	Unknown or Legacy	6.9643	0.9	-75.808	89.7366	False
East.Southeast Asian	White	115.1786	0.001	32.4063	197.9509	True
Indigenous	Latino	-6.2143	0.9	-88.9866	76.558	False
Indigenous	Middle-Eastern	-2.7857	0.9	-85.558	79.9866	False
Indigenous	South Asian	-1.75	0.9	-84.5223	81.0223	False
Indigenous	Unknown or Legacy	8.2143	0.9	-74.558	90.9866	False
Indigenous	White	116.4286	0.001	33.6563	199.2009	True
Latino	Middle-Eastern	3.4286	0.9	-79.3437	86.2009	False
Latino	South Asian	4.4643	0.9	-78.308	87.2366	False
Latino	Unknown or Legacy	14.4286	0.9	-68.3437	97.2009	False
Latino	White	122.6429	0.001	39.8705	205.4152	True
Middle-Eastern	South Asian	1.0357	0.9	-81.7366	83.808	False
Middle-Eastern	Unknown or Legacy	11.0	0.9	-71.7723	93.7723	False
Middle-Eastern	White	119.2143	0.001	36.442	201.9866	True
South Asian	Unknown or Legacy	9.9643	0.9	-72.808	92.7366	False
South Asian	White	118.1786	0.001	35.4063	200.9509	True
Unknown or Legacy	White	108.2143	0.0022	25.442	190.9866	True

Figure 4: Tukey's HSD Results

4.2 Two-way ANOVA for the effect of Sex and Age Group on StripSearch

In order to examine the potential impact of sex and age group on the likelihood of strip search as the dependent variable, a two-way analysis of variance was conducted. To guide this analysis, the following hypotheses were developed:

H_0 : There is NO significant interaction effect between sex and age group on the likelihood of being subjected to a strip search

H_1 : There is a significant interaction effect between sex and age group on the likelihood of being subjected to a strip search

Using the ordinary least square functions, we were able to evaluate the F-statistics of the afore-

mentioned two-way ANOVA and obtained the following results for the independent variables of sex and age group, as well as their interaction term.

Table 11: Two-way ANOVA for Sex and Age Group against StripSearch

	Sum of Squares	df	F	Pr(>F)
C(Sex)	122298	1	11.491	0.0008
C(AgeGroup)	194812	6	3.051	0.007
C(Sex):C(AgeGroup)	66085	6	1.192	0.312
Residual	2234953	210	NaN	NaN

Table 11 shows the results of the two-way ANOVA for the impact of sex and age group on the likelihood of strip search. According to the obtained results, it is evident that the probability of being subjected to a strip search is significantly influenced by both sex ($F(1, 210) = 11.49,\ p = 0.0008$) and age group ($F(6, 210) = 3.05,\ p = 0.007$), indicating significant main effects. However, the interaction between sex and age group does not have a significant impact on the dependent variable of strip search ($F(6, 210) = 1.19,\ p = 0.312$). This suggests that the combined effect of sex and age group on the likelihood of strip search is not considerably different from the effects of each variable alone.

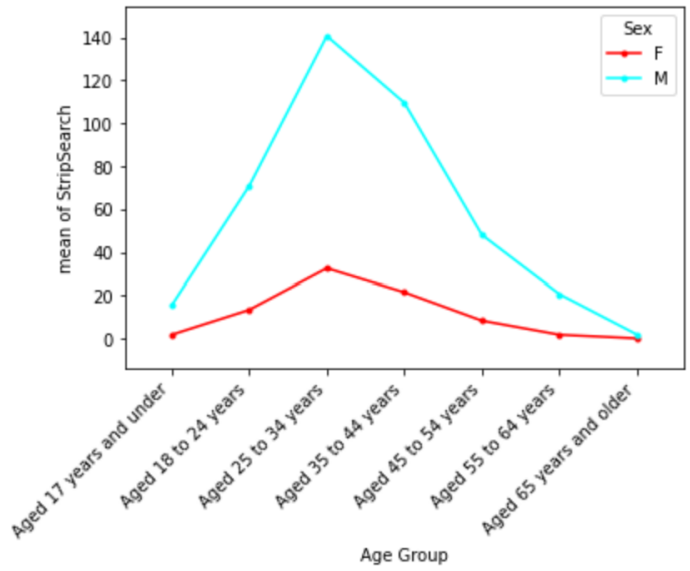

Figure 5: Interaction plot for two-way ANOVA results

Although the interaction plot does not provide information on whether the differences observed are statistically significant, it did reveal several interesting patterns. Firstly, the plot indicated that males are more likely to be subjected to a strip search compared to females. Additionally, the plot showed that there was a noticeable difference in the average likelihood of being strip searched between individuals aged 25 to 34 years old who differed in sex. Lastly, the plot revealed that there was very little difference in the average likelihood of being strip searched between individuals aged 65 years and older who differed in sex. Overall, these observations suggest that both sex and age group are important factors to consider when examining the likelihood of being strip searched.

Discussion

The result presented is from two different ANOVA tests that aim to investigate the effect of race and the interaction between sex and age group on the likelihood of strip search. In both cases, the null hypothesis states that there is no significant difference, and the alternative hypothesis suggests a significant difference.

The first test, a one-way ANOVA, examines the impact of race on the likelihood of strip search. The result indicates that the null hypothesis can be rejected, and there is strong evidence to support the alternative hypothesis that perceived race significantly affects the likelihood of a strip search. However, further examination is required to determine which race groups significantly influence the probability of a strip search. Tukey's Honestly Significant Difference (HSD) test is used for this purpose. The pairwise comparisons suggest that the "Unknown or Legacy" group has a significantly different likelihood of a strip search than the "White" group. Still, there is no significant difference between the "Black" and "East.Southeast Asian" groups. It is essential to note that the lack of significance in the HSD test does not necessarily mean no difference between groups, only that the difference is not statistically significant.

The second test, a two-way ANOVA, aims to investigate the potential impact of sex and age group on the likelihood of strip search. The result indicates that both sex and age group significantly affect the probability of a strip search, with the interaction term between them being non-significant. This implies that the impact of sex and age group on the likelihood of strip search is not dependent on each other.

In conclusion, both ANOVA tests reveal significant effects on the likelihood of strip search. The first indicates the impact of perceived race, while the second shows the impact of sex and age group. These findings are useful in developing policies and procedures to ensure that strip search practices are fair and unbiased. Additionally, it highlights the importance of ongoing monitoring and evaluation of these practices to identify and address any disparities that may arise.

Conclusion

The objective of this project is to analyze data to determine if there are any potential biases or discrimination in strip searches during arrests, focusing on differences by race, gender, and age groups. The dataset used was obtained from the Toronto Police Service Public Safety Data Portal. The findings of this study may have significant implications for promoting fairness and equity in law enforcement practices.

To address the research questions of whether perceived race, sex, and age group affect the likelihood of strip searches during arrests, the study used two types of ANOVA: one-way ANOVA and two-way ANOVA. One-way ANOVA determines significant differences in the means of a continuous outcome variable among two or more groups, while two-way ANOVA incorporates two independent variables with two or more levels and their interaction effect on the outcome variable. The post-hoc statistical test Tukey's HSD was used in one-way ANOVA to determine significant differences between the means of multiple groups, and an interaction plot was used in two-way ANOVA to interpret the effect of two explanatory variables on the dependent variable.

The study conducted two ANOVA analyses to investigate the effects of race and demographic factors on the likelihood of strip searches in Toronto. The first analysis, a one-way ANOVA, examined the relationship between perceived race and strip search. The null hypothesis that proposed no significant difference in the likelihood of strip search between perceived race was rejected ($p<0.001$), indicating that perceived race significantly affects the likelihood of strip search. The second analysis, a two-way ANOVA, examined the effects of sex and age group on the likelihood of strip search. The results showed significant main effects of both sex ($p=0.0008$) and age group ($p=0.007$) on the probability of being subjected to a strip search. However, the interaction between sex and age group was not significant.

Overall, the results showed that perceived race, sex, and age group significantly affect the probability of strip search, while the interaction between sex and age group was found to be non-significant. The "Unknown or Legacy" group was found to have a significantly different likelihood of strip search than the "White" group, and the impact of sex and age group on strip search was not dependent on each other. These findings suggest that demographic factors such as race, sex, and age group may play a role in the likelihood of strip searches in Toronto. Further investigation is needed to better understand and address issues of potential discrimination and bias in law enforcement practices.

References

- MCCANN, N. I. C. K. (n.d.). Oakland whacked for illegal strip searches. Courthouse News Service. Retrieved February 24, 2023, from https://www.courthousenews.com/oakland-whacked-for-illegal-strip-searches/