

Analyzing the Factors of Being Strip-Searched and the Number of Actions at Being Arrested

By

Jia Cao & Emily Clarke

INF2178H: Midterm

Group 31

Faculty of Information

1. Introduction

1.1 Research Questions/Objective

Our society has become increasingly diverse; law enforcement and order maintenance have, almost inevitably, generated sharp conflicts of interest. Tensions between the police and members of various groups, such as young people in general, lesbians and gay men, and gypsy and traveler groups have been rising in recent years. (Steffensmeier, D., et al., 2017) Although the use of stop and search has gained widespread attention, less is being paid to the interactions between the police and members of various groups.

Based on existing research, we are intrigued by examining the correlation between a few representative factors concerning the influence the chance of being strip-searched and the number of actions at being arrested. Notably, individuals' demographic characteristics significantly impact on strip searches decisions made by police officers. For instance, studies show that race, gender, age, DNA and cultural background will make an influence on strip-searching (Lombroso, 2006).

From the perspective of police service, the results of this study provide valuable information on the quality of decisions that are made when it comes to strip searches. It can help police develop policies that are more effective while reducing the likelihood of unnecessary searches.

Hypothetically, we argue that race and sex will make an influence on being strip-searched. Meanwhile, age group also will affect the number of actions at being arrested. Therefore, the following sections of the paper will include Research Hypothesis, Literature Review, Methodology, Results and Discussion, Limitations, and Conclusion to study the problem comprehensively.

1.1.1 Independent Variables

For the first research question, asking if race and sex influence the chance of being strip-searched, the independent variables are race and sex.

For the second research question, asking if age group and sex influence the number of actions at being arrested, the independent variables are age group at arrest and sex.

1.1.2 Dependent Variables

For the first research question, asking if race and sex influence the chance of being

strip-searched, the dependent variable is strip search.

For the second research question, asking if age group and sex influence the number of actions at being arrested, the dependent variable is actions at being arrested.

1.1.3 Data Analysis

Research Question One: *Do race and sex influence the chance of being strip-searched?*

Null Hypothesis (H0): There is no relationship between race, sex, and the chance of being strip-searched.

Alternative Hypothesis (H1): There is a relationship between race, sex, and the chance of being strip-searched.

Research Question Two: *Do age group and sex influence the number of actions at being arrested?*

Null Hypothesis (H0): There is no relationship between age group, sex, and the number of actions at arrest.

Alternative Hypothesis (H1) : There is a relationship between age group, sex, and the number of actions at arrest.

1.2 Literature Review

Criminal Man is widely regarded as the field of study that was founded by Cesare Lombroso. His theory about the origins of criminal behavior dominated the thinking of both Americans and Europeans during the late 19th century. Using evolutionary theories based on Darwinian principles, Lombroso proved that criminals are inferior to people who are honest, white, and women. He also reinforced the racial and sexual hierarchy prevailing in society. He proposed that the causes of crime are related to climate, ethnicity, culture, diet, heredity and age (Lombroso, 2006). His book, through several examples of tribal theft, illustrates the occurrence of crime must have ethnic influence at play. We cannot quantify these effects by specific numbers because one of the reasons is when we look at crime statistics, we find that there are always many complex causes of crime. For example, Spain, Lombardy, Dalmatia, Vvodina and Gorizia have very few cases of female criminality. But there are too many cases of women committing crimes in Slesia, Austria, and Russia's Baltic provinces (Lombroso, Cesare, et al., 2016). Though, in these places, not only the influence of race but also the influence of custom; In areas where men

and women are equally educated and where women intervene in men's struggles, female crime rates are similar to male crime rates. The Germanic peoples of the Austro-Hungarian Empire, especially the Salzbergs and Austrians, have been found to have a higher rate of juvenile delinquents than the Slavs, Italians, Gorizians, Tyrols, and Karnza, which can be explained in the same way (Lombroso, Cesare, et al., 2016).

Newburn, T. et al. (2004) published a paper examining police power over strip search, which is a relatively unexamined issue. They use data collected from a single police force in north London from May 1999 to September 2000. Then analyze the racial difference in using strip searches during this period. Which presents that race and sex influence the chance of being strip-searched. The article written by Steffensmeier, D., & Streifel, C. (1991) proposed that criminology has long established that the involvement of people in crime decreases with age. It also states that males are more likely to offend at every age. His other research paper, "*Age and the Distribution of Crime*," analyzes the distribution of age-related crime across different crime categories by using the FBI's crime reports database for the 1940s, 1960s, and 1980s to compare the various characteristics of the age-related crime curve. In this research, he suggests that the age-related crime curve will likely continue to change. One of the most significant changes was the increasing number of young offenders. The variations in the age distribution of crime types support a traditional sociological notion that the crime rate generally decreases throughout a person's life. However, it can also peak later or decline more slowly.

In summary, existing literature reveals that the core subjects of sociology and law and criminal justice are race, gender, ethnicity, and age. There is a wide range of studies on criminal sentencing that examine the joint and primary effects of gender and race (2017).

The goal of this study is to analyze the collective effects of various factors on the probability. The hypothesis proposed is, **race and sex influence the chance of being strip-searched & age group and sex influence the number of actions at being arrested.**

2. Exploratory Data Analysis

2.1 Descriptive Statistics

To investigate the influence that race and sex have on the chance of being strip-searched, we produced Figure 1. The bar chart shows that the searches for both black and white were much

more frequent than those for other races (including east/southeast asian, indigenous, latino, middle-eastern, south asian, and unknown or legacy). White males have the highest number of strip searches, with over 20,000. Black males follow closely behind with a number of strip searches just below 15,000 for their demographic. Except for the unknown race groups ranked in third place, the totals of strip searches for East/Southeast Asian, Middle-Eastern, South Asian and Indigenous groups were similar.

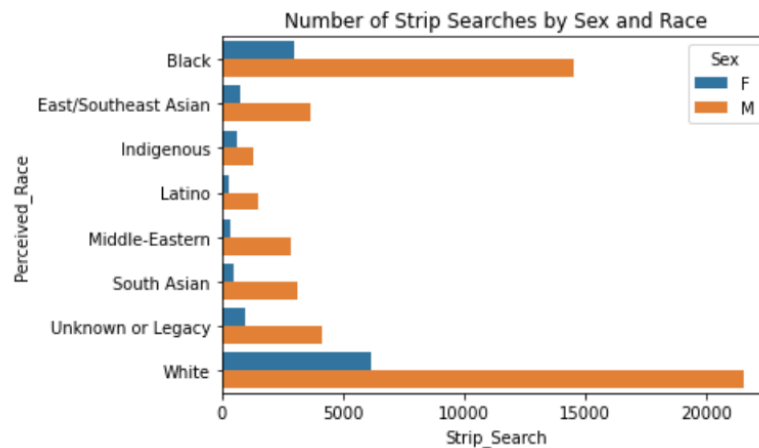


Figure. 1

On the whole, as Figure 1 suggests, men are more likely to have been strip searched than women. When solely looking at the women category, white females take the lead with a little over 5,000 strip searches, and black women take second with a little less than 5,000 strip searches. The remaining female races show low instances of strip searches, with the latina group being the lowest.

BoxPlot of Number of Strip Searches Compared with Sex

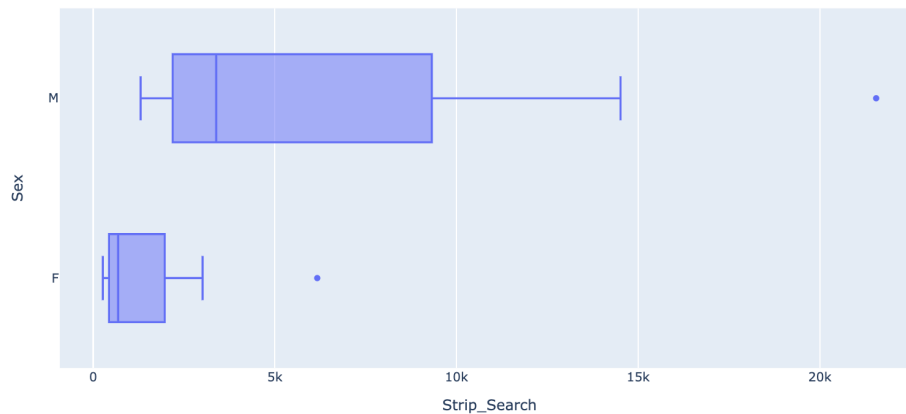


Figure. 2

Figure 2 shows a boxplot of the number of strip searches by sex. In the male category, it shows a max of 2788, min of 122, median of 274, first quartile of 220, and third quartile of 1,309.5 strip searches. In the female category, it shows a max of 778, min of 10, median of 50.5, first quartile of 21.5, and third quartile of 175.5 strip searches. Simply by looking at the median of both sexes, it is apparent that males (274) are more likely to be strip searched than females (50.5).

Histogram of Number of Actions at Arrest Compared with Sex and Age

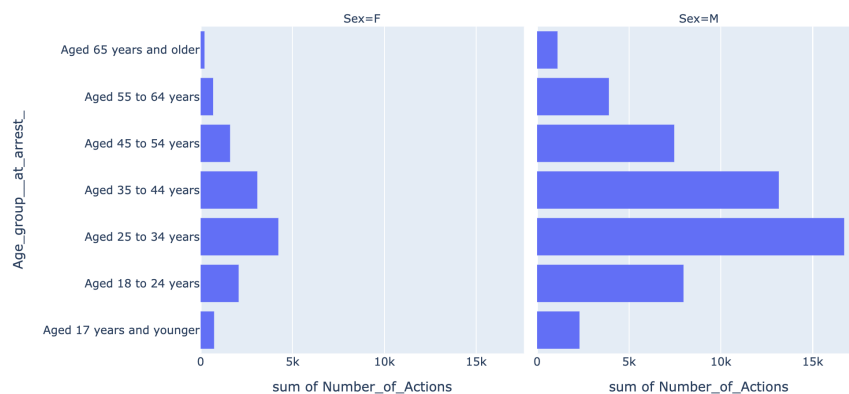


Figure. 3

To explore whether age group and sex influence the number of actions at being arrested, we created a histogram that separates sex and age according to the number of actions at arrest. From Figure. 3, we can easily find that the number of actions for men is surprisingly high, nearly four times as likely as women when facing arrest. The age group with the highest number of actions at arrest in both females and males are those aged 25 to 34 years. After this peak, a decline in

actions at arrest occurs as individuals get older. Though, while ages 35 to 44 is less than the peak of ages 25 to 34 years, it is still the second highest age category that tends to display resistive actions at arrest across both sexes. After these two peaks, resistive actions decline as individuals get older.

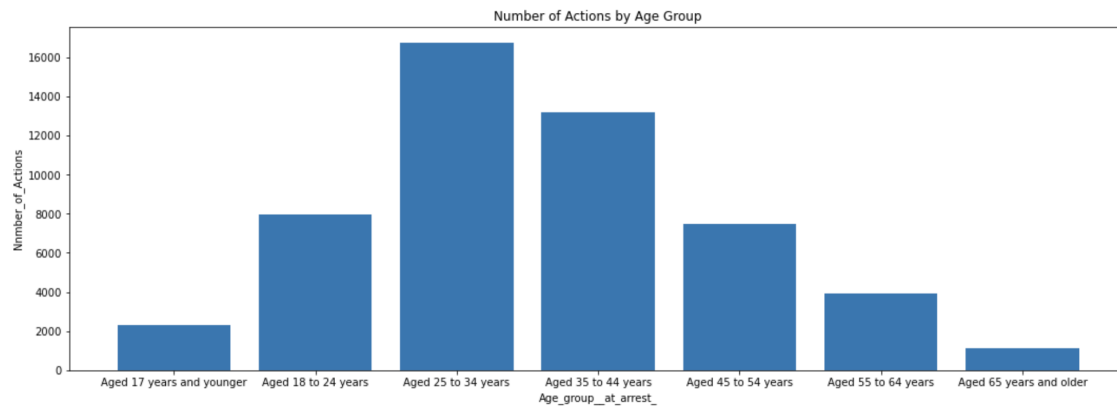


Figure. 4

To eliminate differences in the total number of actions by age, we produced Figure. 4. Regardless of sex, the group aged 25 to 34 years has the highest total number of people conducting actions when facing arrest. As the age increases, the total number of people conducting actions at the time of their arrest decreases. This difference may be influenced by the number of each age group. From Figure 4, the group of people aged 65 years and older has the lowest number of actions. That is to say, the older people get, the smaller the number of acts of resistance.

2.2 T-Test

2.2.1 T-Test Hypothesis

Strip Search and Sex

Null Hypothesis (H0): There is no relationship between sex and the chance of being strip-searched.

Alternative Hypothesis (H1): There is a relationship between sex and the chance of being strip-searched.

Number of Actions and Sex

Null Hypothesis (H0): There is no relationship between sex and the number of actions at arrest.

Alternative Hypothesis (H1): There is a relationship between sex and the number of actions at arrest.

2.2.2 T-Test Results for Strip Search and Sex

The results of the t-test comparing the means between strip search and sex indicate that the mean of strip searches of males (814.625) is higher than that of females (160.375). Considering that alpha is fixed at 0.05, this is not a statistically significant difference as the p-value (0.125) because it is greater than 0.05, with a 95% CI [-1536.54, 228.04]. Therefore, we can not reasonably reject the null hypothesis as there is not sufficient evidence that suggests there is a relationship between Strip Search and Sex.

2.2.3 T-Test Results for Actions at Arrest and Sex

The results of the t-test comparing the means of actions at arrest and sex suggest that the mean of actions at arrest is higher in males (7519.142) and lower in females (1801.285). Considering the alpha is 0.05, the p-value (0.038) offers that these variables are statistically significant as it is lower than 0.05, 95% CI [-11048.09, -387.615]. Therefore, we can reject the null hypothesis as there is a relationship between Actions at Arrest and Sex.

3. Methods

3.1 Dataset

This dataset was pulled from the Public Safety Data Portal of the Toronto Police Service. This website report provides a collection of arrest-related information for the years 2020 and 2021, called “*Arrests and Strip Searches (RBDC-ARR-TBL-001)*”. The dataset was published on November 10, 2022. It provides arrest-related data from years 2020 and 2021, which consists of 24 variables and 65,276 observations.

Within the scope of this analysis, we will be using variables *sex*, *perceived race*, *age group*, *strip search*, and *number of actions at arrest*.

Sex refers to the arrestee's biological sex. Sex and gender should not be used interchangeably. In the data processing component, we decided to remove the data with sex, 'U', because this data volume was too small and they had no actions at the time of being arrested or strip searched. Sex is separated into two (2) categories, male and female.

Age Group at Arrest refers to the age of the individual at the time of being arrested. Age categories ‘17 and under’ and ‘65 and older’ were renamed to state years in their titles to be consistent with the other age group categories. Age group consists of seven (7) categories, aged

17 years and younger, aged 18 to 24 years, aged 25 to 34 years, aged 35 to 44 years, aged 45 to 54 years, and aged 65 years and older.

Perceived Race refers to a person's race as categorized by another individual, in this case a police officer at the Toronto Police. The race categories include eight (8) components, including black, east/southeast asian, indigenous, latino, middle-eastern, south asian, unknown or legacy, and white. Unknown or legacy was kept in this category considering it has a significant number of strip searches associated with it.

Strip Search refers to a type of police procedure involving the officer taking off an individual's clothes and visually inspecting their body.

Actions at Arrest refers to an individual displaying negative or hostile actions at the time of their arrest. Police officers label an action at arrest if they have a concealed item, are combative, are defensive or attempt to escape, are mentally unstable, or have assaulted the officer. Each of these variables were grouped in with one another to create a robust continuous variable for our analysis.

3.2 Measurement

When we conduct statistical tests for research question one and research question two, we will be using a combination of one-way ANOVA and two-way ANOVA to determine if there is a relationship between the independent and dependent variables in each respective research question. If the p-value is significant, then a Tukey HSD test will be conducted to dissect the mean differences across the variables from the tests. Lastly, an interaction plot for each research question will be created to delve deeper into the behaviors of the independent variables compared with the dependent variable.

4.0 Results and Discussion

4.1 Test Results of Research Question 1

4.1.1 One-Way ANOVA

The one-way ANOVA associated with research question one (1) analyzed the independent variable sex and dependent variable strip-search. From the one-way ANOVA test in Appendix A,

it was apparent that the sum of squares is $1.001551e+08$, the degrees of freedom is 1.0, the F statistic is 3.416025, and the p-value associated with the F statistic is 0.085798. From this, we can tell that the p-value is above 0.05. While the p-value is only off by 0.02, it is still considered non-significant. Therefore, we cannot conclude that a significant difference exists between strip-search and sex.

4.1.2 Two-Way ANOVA

The two-way ANOVA for research question 1 analyzes strip searches across perceived race and sex. Interestingly, when analyzed across these variables, we have very different outcomes as compared to the one-way ANOVA. Variable, C(Perceived_Race), indicates that the sum of squares is 42.425430, the degrees of freedom is 7, the F statistic is 58.046535, and the p-value for the F statistic is $1.939933e-83$. Variable, C(Sex), indicates that the sum of squares is 7.000364, the degrees of freedom is 1, the F statistic is 67.045357, and the p-value is $2.700757e-16$. Variables, C(Perceived_Race):C(Sex), indicate that the sum of squares is 7.786977, degrees of freedom is 7, the F statistic is 10.654154, and the p-value is $1.778725e-13$. All this to say, each p-value indicated for perceived race, sex, and combined as compared to strip search indicate significance. Meaning, since the p-value in each case is below 0.05, it is indicative of having statistical significance.

4.1.3 Tukey HSD Test

Considering the one-way ANOVA shows a non-significant p-value, a Tukey test is not necessary. See the limitation section as to why a Tukey test was not conducted for the two-way ANOVA. The interaction effects of the ANOVA variables will be further dissected in the interaction plot section.

4.1.4 Interaction Plot

The interaction plot shown in Figure 5, shows the mean of strip search compared to race and sex. The findings from Figure 5 are quite interesting considering males across all races are more likely to be strip searched than their female counterparts. There is an exception to this claim, considering that white females are more likely to be strip searched than males of east/southeast asian, indigenous, latino, middle-eastern, south asian, and unknown or legacy race groups.

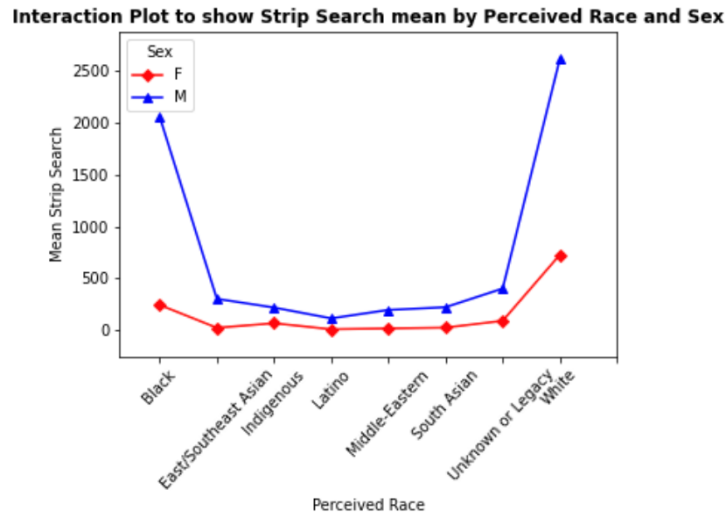


Figure. 5

Additionally, there are two races that have a high likelihood of being strip searched, which are white and black races. Those of the white race are more likely to be strip searched than those of the black race to be strip searched, with black individuals coming in a close second. To illustrate, white males are the demographic more likely to be strip searched with a mean strip search of approximately 2,700 and black males are the second most likely demographic to be strip searched with a strip search mean of approximately 2,200.

4.1.5 Brief Discussion

The data analysis conducted in this section involves a statistical investigation of the relationship between strip searches and the independent variables of sex and perceived race. The analysis includes both one-way and two-way ANOVA tests, as well as an interaction plot. The one-way ANOVA test showed that there was no significant difference between strip searches and sex. Although strip search and sex did not achieve statistical significance, it displayed an intriguing trend that merits further investigation in future studies. However, the two-way ANOVA test revealed that perceived race, sex, and their interaction were all significant factors in predicting strip searches. The interaction plot showed that males across all races were more likely to be strip searched than their female counterparts, but white females were more likely to be strip searched than males of several other race groups. Additionally, the plot showed that white and black races had a higher likelihood of being strip searched than other races, with white males being the most likely demographic to be strip searched and black males being the second most likely demographic to be strip searched.

4.2 Test Results of Research Question 2

4.2.1 One-Way ANOVA

The one-way ANOVA analyzed the number of actions at arrest compared with sex. The sum of squares is 1.144286e+08, degrees of freedom is 1, F statistic is 6.527416, and the p-value is 0.025238. The p-value is less than 0.05, which means that it is statistically significant. A post-hoc test will be conducted and will be further discussed in section 4.2.3.

4.2.2 Two-Way ANOVA

The two-way ANOVA for research question 2 analyzes the number of actions at arrest across age group and sex. The C(Age_group__at_arrest_) states that the sum of squares is 11.636744, the degrees of freedom is 7, the F statistic is 4.857, and the p-value is 0.000648. The p-value for age group compared with actions is statistically significant. The C(Sex) variable suggests that the sum of squares is -0.000233, degrees of freedom is 2, F statistic is -0.00034, and the p-value is 1. The p-value for sex across actions is not statistically significant. When comparing both C(Age_group__at_arrest_):C(Sex), the sum of squares is 10.3348, degrees of freedom is 14, F statistic is 2.156, and p-value is 0.013954. When comparing the two independent variables across the dependent variable, the p-value is significant.

4.2.3 Tukey HSD Test

A Tukey HSD Test was conducted for the number of actions at arrest across sex. The mean difference between females and males is 5717.8571, the adjusted p-value is 0.0252, the lower is 841.64, and the upper is 10594.069. All this to say, actions at arrest and sex are statistically significant and we can reject the null hypothesis.

A Tukey HSD Test could not be performed on the two-way ANOVA, see limitation section for an elaboration.

4.2.4 Interaction Plot

Figure 6 displays an interaction plot that illustrates the mean number of actions at arrest compared by age group and sex. It should be noted that the x-axis ticks are shifted one tick to the right, which is improper formatting. Figure 6 will be analyzed with this in mind, and will attribute the age group to the plotted points one tick to the left on the x-axis.

In both male and female groups, the peak of number of actions at arrest occurs within the age group of 25 to 34 years old with males with over 15,000 actions at arrest and females with a little

less than 5,000 actions at arrest. Once this peak is hit at ages 25 to 34 years, there is a gradual decline as males and females get older.

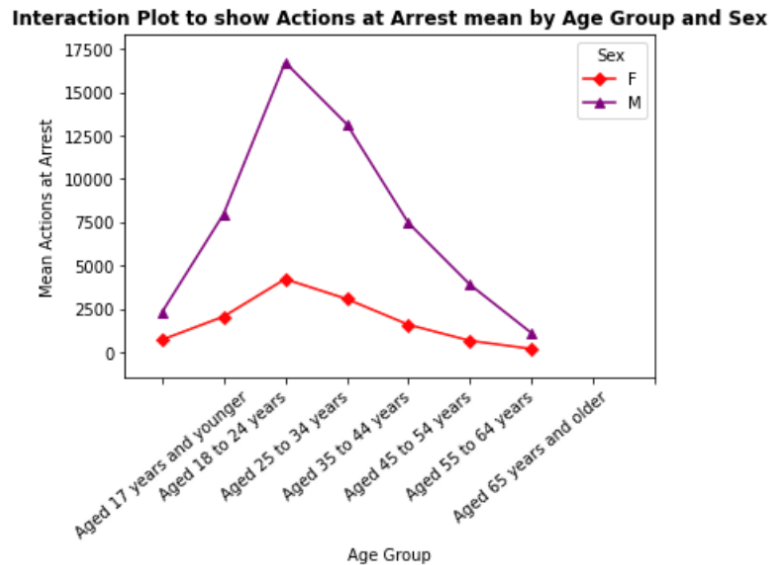


Figure. 6

Moreover, males are more likely than females to display unfavorable actions at the time of their arrests. The only exception to this statement is that females aged 25 to 34 years of age are more likely to display adverse actions at the time of their arrest than males aged 55 years and older. Therefore, it could be said that younger individuals are more likely to display a hostile action at the time of their arrest than older individuals. Additionally, males are more likely to show these actions than females.

4.2.4 Brief Discussion

This data analysis involves conducting one-way and two-way ANOVA tests and a Tukey HSD test to examine the relationship between the number of actions at arrest, sex, and age group. The results show that sex and age group are statistically significant predictors of the number of actions at arrest, with younger males being more likely to display adverse actions at the time of their arrest. The two-way ANOVA conducted for research question 2 revealed that age group has a statistically significant effect on the number of actions at arrest, while sex does not. However, when considering the interaction effect between age group and sex, there was a significant difference in the number of actions at arrest.

5.0 Limitations

One limitation that arose during the coding process included difficulties running a two-way ANOVA on the cleaned data labeled “df_1” for research question one and “action_count_1” for research question two. It stated that there were Infs or Nans in the dataset. After running `df.isna()` to check which entries within the data frame were Na, it still could not tell me. Then, I attempted to convert the independent variables into dummy variables using `get dummies`, which was difficult to perform on a large independent variable, such as race (in research question one) or age (in research question two). Therefore, I used the group that was not grouped using the `groupby()` function to run the two-way ANOVA, which ultimately means the two-way ANOVA tests may not be as accurate.

Another limitation that appeared during the coding process was the inability to perform a Tukey test on my two-way ANOVA tests for research question one and research question two. It gave me an error code that my variables could not have Infs or NaNs. It is reasonable to assume that there were infinite variables in the perceived race column, considering we defined NaNs as 0 when preprocessing the data when using `isnull()`, `fillna()`, and `dropna()`.

To suggest further research, future research could focus on replicating the study using a larger and more diverse sample, exploring different variables or factors, or investigating the potential impact of interventions or policies aimed at addressing trends within police institutions.

6.0 Conclusion

Ultimately, this report provides valuable insights into the factors of being strip-searched and the number of actions being arrested. Through the analysis of the Toronto police dataset, our research aimed to answer two core research questions, which are:

1. Do race and sex influence the chance of being strip-searched?
2. Do age groups and sex influence the number of actions at being arrested?

The analysis highlights the importance of sound statistical methods, such as ANOVA tests and post-hoc tests, in generating reliable and accurate conclusions. For research question 1, the two-way ANOVA revealed that there was a statistically significant interaction effect between perceived race and sex on the likelihood of being strip searched, while the one-way ANOVA did not find a significant difference between sex and being strip searched. The interaction plot

showed that white and black races had the highest likelihood of being strip searched, with white males being the most likely demographic to be strip searched. For research question 2, the two-way ANOVA found a statistically significant interaction effect between age group and sex on the number of actions at arrest, while the main effects of age group and sex were not significant. This suggests that the relationship between age group and number of actions at arrest depends on the sex of the individual.

Through careful examination of the data, we have identified significant trends, patterns, and relationships that help us understand the underlying factors driving the Toronto Police Department. Our findings not only provide a better understanding of the data, but also offer useful information that can be used to inform decision-making and policy formulation with respect to trends within police institutions.

References

Lombroso, C., Gibsonand, M., Rafter, N. H., & Seymour, M. (2006). *Criminal Man*. Duke University Press. <https://doi.org/10.2307/j.ctv11vc7kd>

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

Steffensmeier, D., & Streifel, C. (1991). Age, Gender, and Crime across Three Historical Periods: 1935, 1960, and 1985. *Social Forces*, 69(3), 869–894. <https://doi.org/10.2307/2579479>

Steffensmeier, D. J., Allan, E. A., Harer, M. D., & Streifel, C. (1989). Age and the Distribution of Crime. *American Journal of Sociology*, 94(4), 803–831. <http://www.jstor.org/stable/2780859>

Steffensmeier, D., Painter-Davis, N., & Ulmer, J. (2017). Intersectionality of Race, Ethnicity, Gender, and Age on Criminal Punishment. *Sociological Perspectives*, 60(4), 810–833. <https://www.jstor.org/stable/26579835>

Bowling, B. and Phillips, C. (2002), *Racism, Crime and Justice*. Harlow: Longman.

Fitzgerald, M. (1999), *Final Report into Stop and Search*. London: Metropolitan Police.

Goldstein, H. (1977), *Categorizing and Structuring Discretion*, in *Policing a Free Society*, 93-130. Cambridge, MA: Ballinger Publishing Company.

Home Office (2003), *Statistics on Race and the Criminal Justice System: A Home Office Publication Under s. 95 of the Criminal Justice Act 1991*, 2002. London: Home Office.

Jefferson, T. and Walker, M. (1992), 'Ethnic Minorities in the Criminal Justice System', *Criminal Law Review*, 81: 140, 83-95.

Judge, T. (1986), 'The Provisions in Practice', in J. Benyon and C. Bourn, eds, *The Police: Powers, Procedures and Proprieties*. Oxford: Pergamon Press.

Appendices

Appendix A

One-Way ANOVA for RQ1 (Strip Search across Sex)

	sum_sq	df	F	PR(>F)
C(Sex)	1.001551e+08	1.0	3.416025	0.085798
Residual	4.104685e+08	14.0	NaN	NaN

Appendix B

Two-Way ANOVA for RQ1 (Strip Searches across Perceived Race and Sex)

	sum_sq	df	F	PR(>F)
C(Perceived_Race)	42.425430	7.0	58.046535	1.939933e-83
C(Sex)	7.000364	1.0	67.045357	2.700757e-16
C(Perceived_Race):C(Sex)	7.786977	7.0	10.654154	1.778725e-13
Residual	6812.593171	65247.0	NaN	NaN

Appendix C

One-Way ANOVA for RQ2 (Actions across Sex)

	sum_sq	df	F	PR(>F)
C(Sex)	1.144286e+08	1.0	6.527416	0.025238
Residual	2.103655e+08	12.0	NaN	NaN

Appendix D

Tukey HSD Test for RQ2 (Actions across Sex)

```
Tukey HSD for Number of Actions at Arrest across Sex
Multiple Comparison of Means - Tukey HSD, FWER=0.05
=====
group1 group2 meandiff p-adj  lower    upper  reject
-----
      F      M 5717.8571 0.0252 841.6453 10594.069   True
-----
```

Appendix E

Two-Way ANOVA for RQ2 (Arrest Actions across Age Group and Sex)

	sum_sq	df	F	PR(>F)
C(Sex)	-6.969632e-09	2.0	-1.018179e-08	1.000000
C(Age_group_at_arrest_)	7.970273e+00	7.0	3.326745e+00	0.005253
C(Sex):C(Age_group_at_arrest_)	1.033487e+01	14.0	2.156858e+00	0.013954
Residual	2.233449e+04	65256.0	NaN	NaN

Appendix F

T-Test Output for Research Question 1 (Strip Search and Sex)

Ttest results for strip search: Ttest_indResult(statistic=-1.7147422430880237, pvalue=0.12533757126252384)

Mean Strip Search by Females and Males: 160.375 814.625

Standard deviation Strip Search by Females and Males: 262.1389694145346 1046.8481862237713

Confidence interval for Strip Search: (-1536.5409608306613, 228.04096083066133)

DOF for Strip Search: 7.87441900834675

Appendix G

T-Test Output for Research Question 2 (Actions and Sex)

Ttest results for Actions at Arrest: Ttest_indResult(statistic=-2.554880915690823, pvalue=0.03895978707835791)

Mean Number of Actions at Arrest by Females and Males: 1801.2857142857142 7519.142857142857
Standard deviation Actions by Females and Males: 1448.6176990825559 5741.291446720892

DOF for Actions at Arrest: 6.760874438245906

Confidence interval for Actions at Arrest: (-11048.098578073623, -387.61570764066437)