The University of Toronto

Racial Disparities in Age at Arrest: An Analysis of Police Records by Perceived Race, Sex, and Arrest Data

By

Omer Imran and Zhen Wu (Group 60)

Faculty of Information

INF2178 Experimental Design for Data Science

Instructor: Shion Guha

April 15th, 2023

Colab Link

Introduction

Arrest data is a critical source of information for understanding the prevalence of criminal behavior and its impact on society. In recent years, there has been growing concern about the disproportionate representation of particular demographic groups, such as minorities and young people, in the criminal justice system. This has led to a renewed focus on factors contributing to criminal involvement, including age, race, and sex.

In this study, we conducted an exploratory data analysis (EDA) and statistical tests to examine the relationship between age at arrests and strip searches in a dataset of police records. Specifically, we were interested in investigating whether there were any differences in the counts of strip searches among different racial groups, sexes, ages, and time of the arrest. We first performed an EDA to explore the general data distribution and identify potential issues needing attention. Additionally, we performed a Power Analysis to determine the sample size needed to achieve a statistical significance level of 0.05, which is the probability of correctly rejecting the null hypothesis when it is false. We conducted a series of analyses on a dataset of individuals who were subject to strip searches after arrest, including ANCOVA as well as exploratory data analysis and t-tests. Logistic regression was also used to examine the relationship between strip searches and categorical predictor variables which allowed us to estimate the odds ratio of a particular group being subjected to a strip search. By combining these analytical methods, we aimed to provide a comprehensive understanding of the arrest and strip search data and shed light on potential areas for further research and policy improvements. Specifically, we will explore the following research questions:

1. Is there significant differences in age at which individuals of different races are subjected to strip searches during an arrest with time as a control variable?
2. Are there any statistically significant differences in the mean age of arrest between different groups defined by sex with time as a control variable?
3. Are sex, race, and age significant factors that affect chances of being strip searched?

Literature Review

In this literature review, we explore three research questions related to the age at arrest, race, sex, and strip searches during police arrests. The first question asks whether there is a significant difference in the age at arrest for individuals of different races that are subjected to strip search with time as a control variable. The second question examines whether there are significant differences in the age at which individuals of different races and gender are arrested with time as a control variable. Finally, the third question explores whether sex, race and age are significant factors in determining

chances of getting strip searched. These questions are essential because they address issues of fairness and potential biases in the criminal justice system and could help identify areas where reform is needed.

Using strip searches during arrests has been controversial, with concerns about its potential for discriminatory treatment. Research has shown that strip-search powers are used disproportionately against Afro-Caribbean arrestees, even after controlling for factors like sex, age, the reason for arrest, and charge (Keeton, 2015). In addition, Latinos are more likely to be subjected to strip searches during police stops if officers do not take extra steps to confirm the validity of systems reporting the names of illegal immigrants, as false positives are more likely to be triggered by Hispanic last names. It is not only adults who are subjected to discriminatory strip searches. Black and Latino schoolchildren are also often subjected to arbitrary discipline. Studies have found that these minority groups are more likely to be referred to the police for infractions, arrested, and suspended than their White counterparts. Furthermore, research shows that the age at which individuals of different races are subjected to strip searches also varies. For instance, the author found that Black and Latino arrestees were more likely to be strip-searched younger than their White counterparts, reflecting a disparate impact of strip-search practices in the age of colorblind racism.

Criminal activities and time of year pose an interesting relationship. Researchers have studied whether there is seasonality in the level of crime in a given year. Andresen and Malleson (2013) studied this in the context of Vancouver in 2001 and found that criminal activities tend to increase and peak near the summer months i.e. July and August and then begin to fall. McDowall et al. (2012) found similar results: most criminal offenses tend to occur in the summer months with a fall in the winter months with the exception of robbery. This is a phenomenon that we would like to control for in our study.

Additionally, according to a study by Newburn, Shiner, and Hayman (2004), male arrestees were almost twice as likely to be strip-searched as female arrestees. Age was also a factor, with the rate of strip searches declining as age increased, particularly among adult arrestees. Juveniles had a relatively low rate of strip-search, which could be due to protective legislation. The study also found that ethnicity played a more significant role in the outcomes of an arrest than either sex or age. African-Caribbean arrestees were the most likely to be charged and less likely to receive a caution, reprimand, or warning, while white arrestees were the most likely to receive these outcomes. The study noted that receiving official action other than a charge, caution, reprimand, or warning was associated with a particularly low rate of strip searches. Among ethnic groups, Arabs had a relatively high rate of strip searches, with Asian and Mediterranean groups having lower rates.

The relationship between mean age at arrest and race and sex has interested researchers studying criminal justice. One study examined the over-representation of Black people in single charge cases in Toronto (Chan & Chunn, 2017). Despite making up only 8.8% of the population, Black people represented 28.8% of the cases in the dataset, making them 3.3 times more likely to

appear in single charge cases than their representation in the general population would predict. Further data analysis revealed that Black males have much higher charge rates than males from all other racial groups within each offense category. Black males represent only 4% of Toronto's population but are involved in almost a third of the charges captured by the data request. This means that Black males are 7.3 times more likely to appear in the charge dataset than their representation in the general population would predict.

In terms of overall charge rates conducted by the same study, the rate for Black males (29,694 per 100,000) is 4.5 times higher than the rate for White males (6,673 per 100,000) and 7.5 times higher than the rate for males from other racial minority backgrounds (3,935 per 100,000). The rate for Black women (2,805 per 100,000) is 2.4 times higher than the rate for White women (1,159 per 100,000) and 6.2 times higher than the rate for women from other racial minority groups (454 per 100,000) (Chan & Chunn, 2017). It is important to note that Black racial identity still increases the likelihood of arrest after controlling for frequency and seriousness of the offending, gender, age, socioeconomic status, family characteristics, and other risk factors.

Exploratory Data Analysis

After performing various data cleaning processes on the arresting dataset and identifying key research questions, we narrowed our focus to specific patterns and variables, including race perceived at arrest, sex, occurrence, age, and actions at arrest, strip search, and items found. To further support our analysis, we conducted exploratory data analysis to identify correlations and connections between the variables. We hope to understand better the underlying factors contributing to arrest and criminal involvement patterns by delving deeper into the data and uncovering meaningful insights.

We observed specific patterns in the different categorical groups during our analysis of the arrest counts. Figure 1 depicts the counts of arrests among seven racial groups, where White had the highest number of arrests followed by Black, while Latino and Indigenous had the lowest.

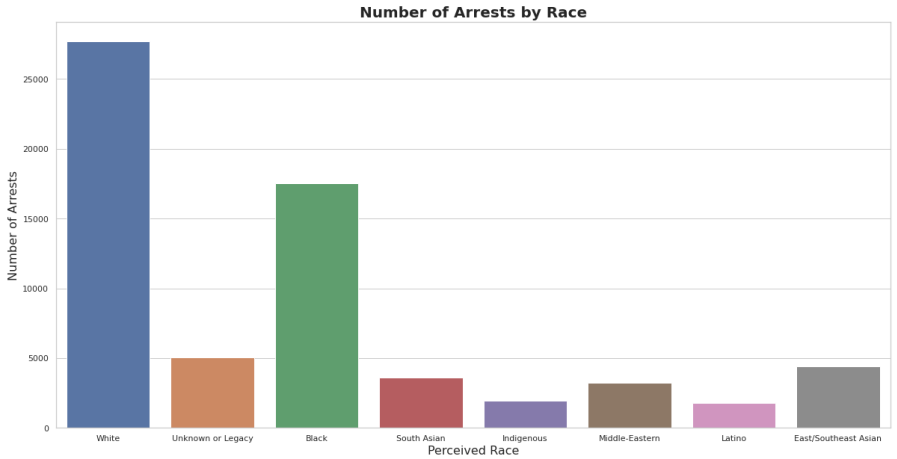

Figure 1: Number of Arrested by Race

To further understand these patterns, we performed an additional analysis which included sex, month, and age at arrest, as illustrated in Figure 2. Despite the similarity in the month of arrests, a clear correlation was observed between sex and age at arrest. The data revealed that males constituted around 80% of the total arrest counts, with over 50,000 records. Interestingly, most arrests occurred among individuals in their late 20s and early 30s, which aligns with our literature research on the relationship between physical abilities and criminal involvement.

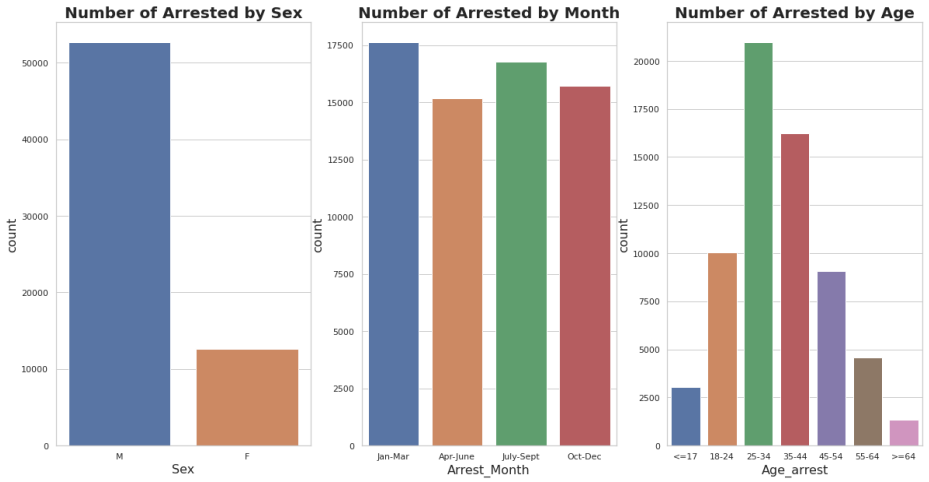

Figure 2: Number of Arrested by Sex, Month, and Age

In our research question, we were interested in understanding more about the arrest and strip search data. We first examined the differences in the counts of arrests among racial groups and wanted to see if a similar trend was present regarding the counts of strip searches. As shown in Figure 3, police officers conducted most of the strip searches on White and Black arrestees. However, when we compared the number of strip searches performed on these two groups with the number of arrests, we found that the proportion of strip-searched White arrestees was much lower than that of Black arrestees. In other words, Black arrestees were more likely to be strip-searched by police officers than White arrestees.

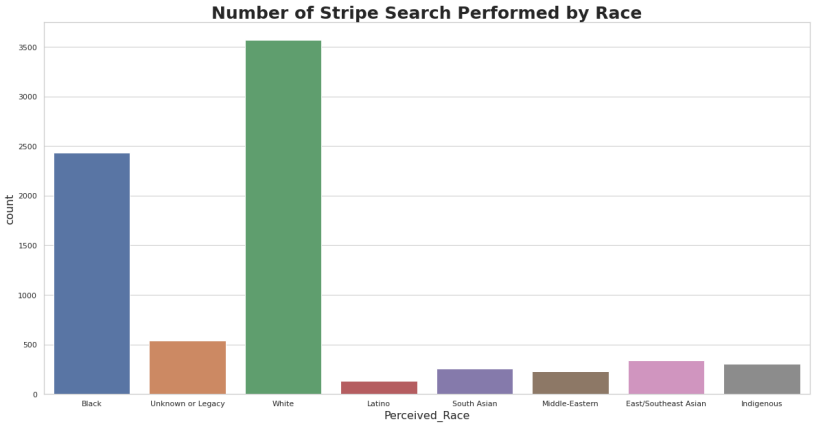

Figure 3: Number of Strips Search by Race

Furthermore, since we observed a difference in the number of arrests by sex and age, we performed additional visualizations using strip search data, as shown in Figure 4. Our analysis revealed that not only sex but also age at the time of arrest showed a similar trend regarding strip searches. This suggests that age and sex did not have a significant impact on the decision-making of police officers regarding strip searches. These findings suggest that the likelihood of being strip-searched by police officers may vary based on racial identity and that Black arrestees may be subject to strip searches at a higher rate than their White counterparts.

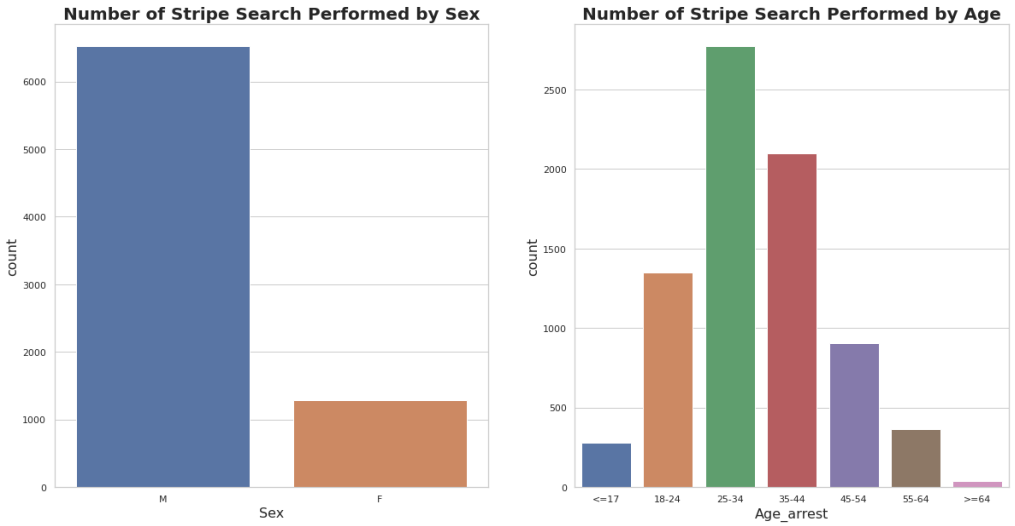

Figure 4: Number of Strip Search by Sex and Age

Figures 5 and 6 show that cooperative action at arrest took up the most significant portion of the data, with counts of almost 30,000. However, despite the lack of resistance or aggression, around 3,000 of these cooperative arrests still resulted in strip searches, which accounts for approximately 10% of the total arrests in the same action. This finding suggests that the actions at arrest may not be the primary factor in determining whether an individual will be subject to a strip search. These results align with previous studies on the subject, which have indicated that strip searches are often carried out based on various factors, including the discretion of individual officers, the policies and procedures of the law enforcement agency, and the circumstances of the arrest. Therefore, further research is needed to understand better the factors influencing the decision to perform a strip search and to develop more consistent and equitable policies and procedures surrounding this controversial practice.

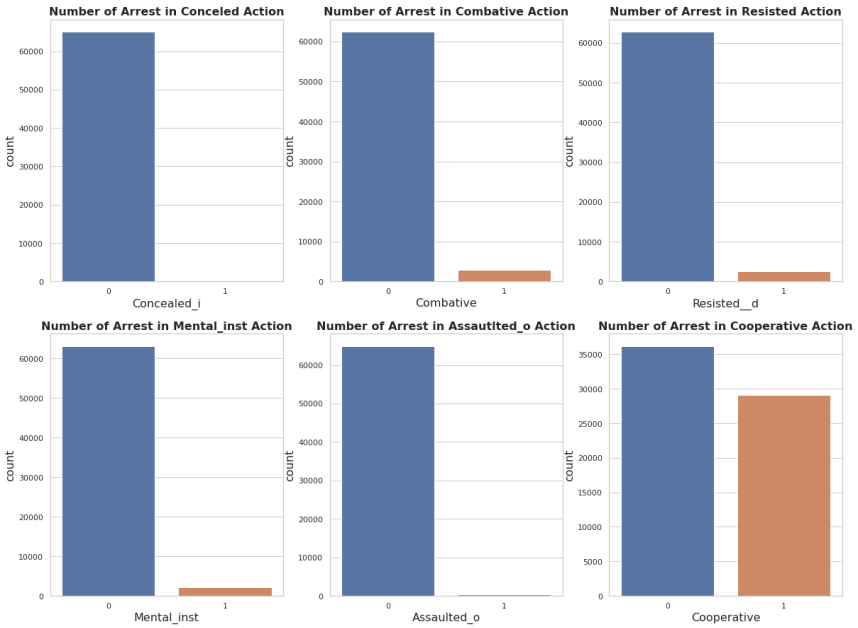

Figure 5: Counts of Arrest by Action

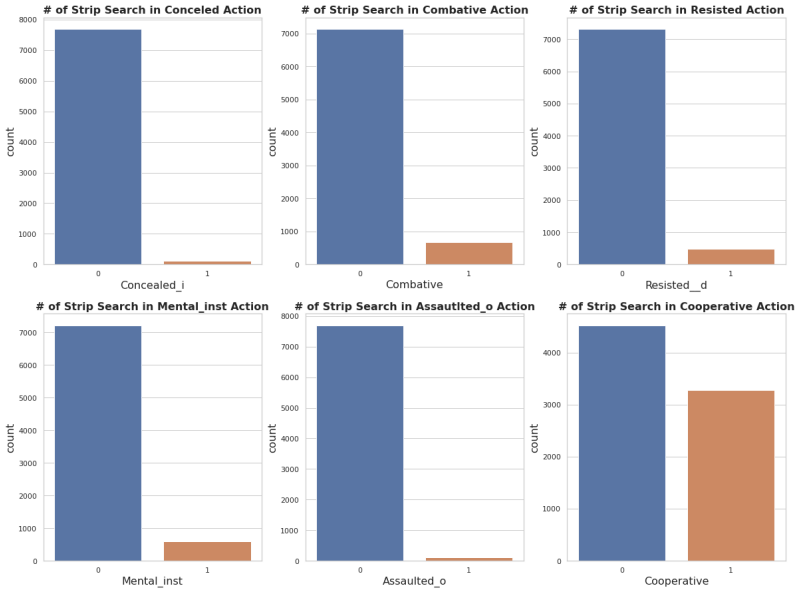

Figure 6: Counts of Strip Search by Action

Before proceeding with the statistical test, we examined the general distribution of the data to identify any significant variables that require attention. Figures 7 and 8 show a similar distribution of arrest data in sex, age, and race, which may suggest that police officers are making arrests based on the crime committed rather than the individual's demographics. In other words, the officers are not biased toward a particular gender, age group, or racial group when making arrests. However, this does not necessarily mean that there is no bias present in the criminal justice system. Other factors may contribute to the disparities in the criminal justice system, such as the over-policing of specific

neighborhoods or the differential treatment of certain crimes. Therefore, it is essential to consider multiple factors and conduct further analysis to understand the underlying causes of the observed patterns.

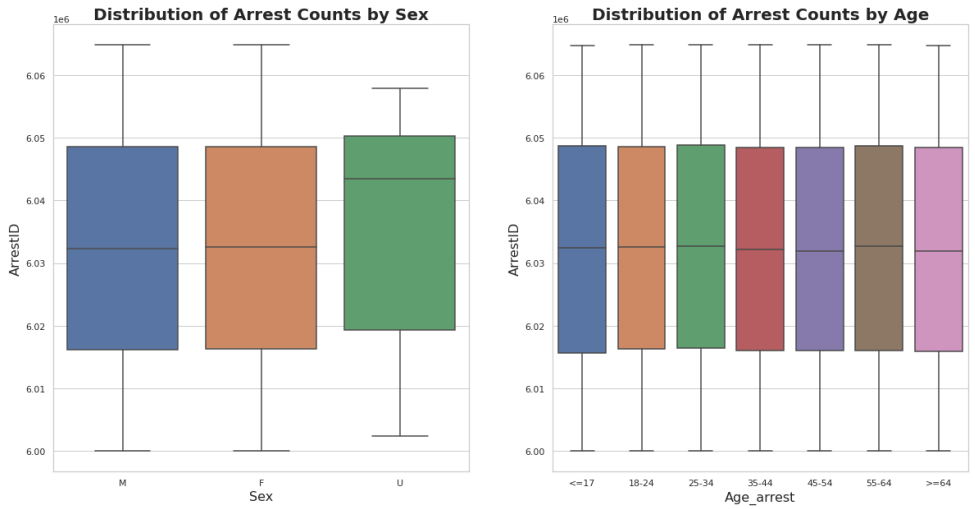

Figure 7: Distribution of Arrest by Sex and Age

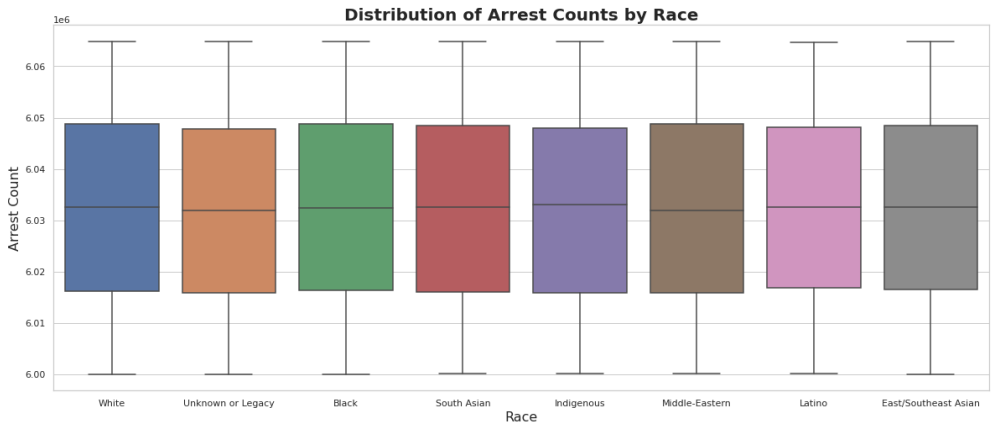

Figure 8: Distribution of Arrest by Race

T-tests

T-tests are deployed to test the hypotheses of means of a dependent variable depending upon the different levels of a categorical variable. They are handy in checking whether a change in the level of a categorical variable influences the continuous dependent variable. The only limitation is that a single t-test can only check the difference between two levels at a time – something that ANOVA overcomes.

Welch's t-test is the most robust t-test. It assumes unequal variances between the groups it is being tested on. And when the variances are equal, the results are the same as those of the Student

t-test with equal variances. That is why we exclusively ran Welch's t-tests. We used two-tailed t-tests only because we had no hypothesis about the direction of the difference.

Additionally, there were two assumptions to test for before running a t-test: homogeneity of variances and normality. The first assumption is considered false since we're using Welch's t-test. We ran a Shapiro test to check for normality, but all of our variables had more than 5000 rows, making the p-value inaccurate. Additionally, all of our variables had significant results on the Shapiro test, i.e., we rejected the null hypothesis that the distribution was normal for arrest age and variables we discuss below. Regardless of this limitation, we moved forward with our t-tests.

We considered the t-tests somewhat part of our EDA and, as such, executed a lot of them. A select few relevant to our research questions are presented below, along with their interpretations.

Test 1: Sex and Arrest Age

Null Hypothesis: there is no difference in the mean arrest age of males and females
Alternative Hypothesis: there is a difference in the mean arrest age of males and females

Our results gave a p-value < 0.05, indicating significant results hence we reject the null that there are no differences in mean arrest age of males and females. The mean arrest age for males was 36.36, with a standard deviation of 12.39 years. The mean arrest age for females was 34.95, with a standard deviation of 11.94 years. We are 95% confident that the difference in the means is between 1.18 and 1.64 years.

More analysis could be possible if we compare each gender with each other – more on this in the ANOVA section.

Test 2: Race and Arrest Age

Null Hypothesis: there is no difference in the mean arrest age of white and black individuals
Alternative Hypothesis: there is a difference in the mean arrest age of white and black individuals

Our results gave a p-value < 0.05, indicating significant results hence we reject the null that there are no differences in mean arrest age of white and black people. The mean arrest age for a white person was 38.65, with a standard deviation of 12.29 years. The mean arrest age for females was 32.48, with a standard deviation of 11.29 years. We are 95% confident that the difference in the means is between 5.94 and 6.38 years.

We ran more similar tests comparing means of white to south Asians, white and indigenous. The results aren't presented here because we do a deep dive into them in the ANOVA section.

Test 3: Strip search and Arrest Age

Null Hypothesis: there is no difference in the mean arrest age of people who were strip-searched and those who were not
Null Hypothesis: there is a difference in the mean arrest age of people who were strip-searched, and those who were not

Our results gave a p-value < 0.05 indicating significant results hence we reject the null that there are no differences in mean arrest age of people who were strip searched and those who were not. The mean arrest age for strip-searched individuals was 34.54, with a standard deviation of 10.96 years. The mean arrest age for those not strip searched was 36.30, with a standard deviation of 12.47 years. We are 95% confident that the difference is between 1.50 and 2.03 years.

This result seems to answer our research question but we will further investigate this in ANOVAs. The next two t-tests are also indicative of why we need two-way ANOVAs: t-tests can look into differences at multiple levels of two categorical variables.

Test 4: Black, Gender, and Arrest Age

Null Hypothesis: there is no difference in the mean arrest age of black males and females
Alternative Hypothesis: there is a difference in the mean arrest age of black males and females

Our results gave a p-value < 0.05 indicating significant results hence we reject the null that there are no differences in mean arrest age of black males and females. The mean arrest age for black males is 32.75, with a standard deviation of 11.34 years, while the mean arrest age for black females is 31.20, with a standard deviation of 10.94 years. We are 95% confident that the difference is between 1.12 and 1.97 years.

In summary, all these t-tests did allude to what the answers to our research questions will be. Still, as noted earlier, t-tests can only examine the difference of means between two categorical variable groups. It cannot look at the difference of means of various levels of a categorical variable or assess the interaction of two categorical variables. As a result, we move on to ANOVAs for a comprehensive answer to our research questions.

Methods

The dataset used in this study includes information related to all arrests and strip searches in the City of Toronto. The data is provided by the Toronto Police Service and can be accessed through the following link:

Toronto Police Service. 2021. Arrests and Strip Searches (RBDC-ARR-TBL-001). [Dataset]. Toronto Police Service.
https://doi.org/10.21232/TorontoPS-arrests-and-strip-searches-rbdc-arr-tbl-001

The dataset includes 25 variables such as the age of the person arrested, their perceived race, sex, location of the arrest, and whether or not the person was strip-searched. 37,347 arrestees were recorded in the dataset, with a total of 64,805 arrest records. Most variables have binary values, such as Stripsearch, Actions at arrest, search reason, and items found. Strip searches, as one of the binary values, involve the removal of some or all clothing and a visual inspection of the body, which is represented as either 1 or 0. Categorical variables include arrest year, month, perceived race, sex, age group, location, and occurrence. There are 8 racial and 7 age groups perceived in the data separated into male, female, or undefined. It is also worth noting that there are 469 null values under Arrest ID.

Prior to conducting ANCOVAs and logistic regression, we conducted three power analyses based on our research questions. For gender differences, we calculated the effect size of the explanatory variable using Cohen's d metric, which was 0.14. We then computed the required sample size using the obtained effect size and established the statistical power at 80%. As shown in Table 1, the results indicated that a sample size of 505 was required for females, while a sample size of 2437 was required for males. The actual sample size for both males and females in our dataset was 14509 and 3009, respectively, which exceeds the required sample size. At the same time, as shown in Figure 9, we need to have over 750 observations to achieve 80% power at 0.14 effect size. Therefore, we can be confident in the reliability of our results.

(Alpha=0.05, Power=0.8)	**Male**	**Female**
Effect Size	0.14	
Sample Size	2437	505
Actual Size	14509	3009

Table 1: Power Analyst for Male and Female

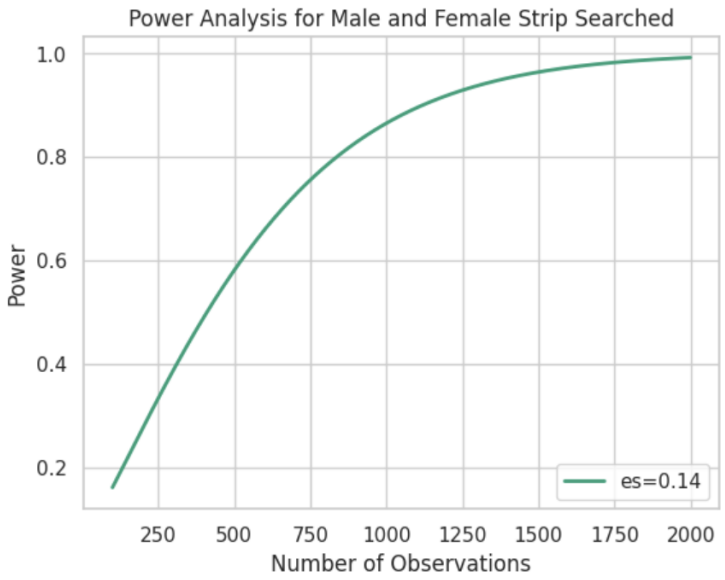

Figure 9: Power Curve for Male and Female

We also conducted power analyses to determine the required sample size for analyzing the effect of race on strip searches in White and Black race groups, using Cohen's d metric. The effect size we obtained was 0.53, which was much higher than the effect size for gender. The required sample size for White and Black race groups, with a statistical power of 80%, were calculated as 68 and 4, respectively as shown in Table 2. The actual sample size for White and Black race groups in our dataset were 21546 and 14509, respectively. As shown in Figure 10, the observation needed to achieve 80% of power at 0.54 effect size is around 55. Thus, we have a sufficient number of observations to make meaningful analyses. It's important to note that the larger effect size for race compared to gender suggests that race may be a more influential factor in determining whether an individual will be subjected to arrest.

(Alpha=0.05, Power=0.8)	**White**	**Black**
Effect Size	0.54	
Sample Size	68	46
Actual Size	21546	14509

Table 2: Power Analyst for White and Black

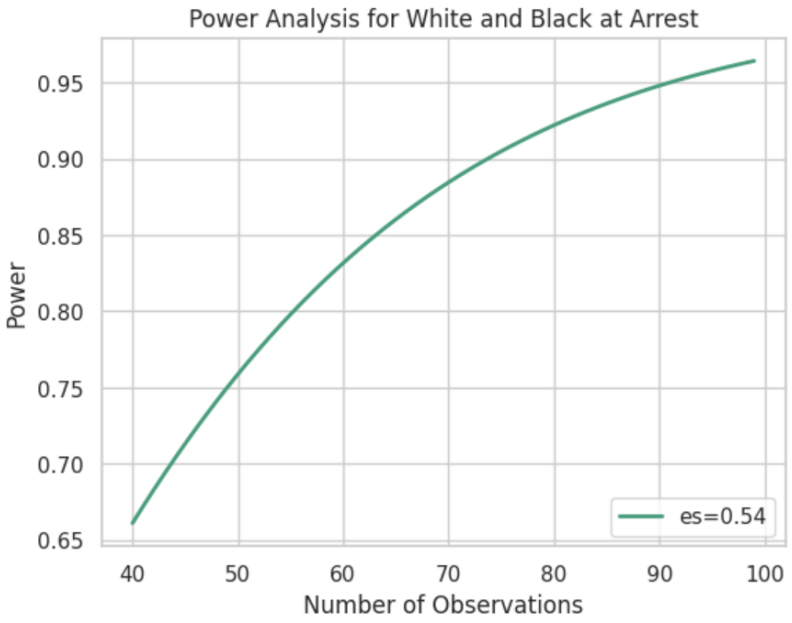

Figure 10: Power Curve for White and Black

Furthermore, we also conducted a power analysis to determine the required sample size for analyzing the relationship between strip searches and the total number of arrests in our dataset. Among the dataset, as shown in Table 3, 7801 individuals were subjected to a strip search, while the remaining 57451 were not. Using Cohen's d metric, we obtained an effect size of 0.14 and computed the required sample size for a statistical power of 80%. The results indicated that a sample size of 432 was required for the strip-searched group, while a sample size of 3185 was required for the non-strip-searched group. As shown in Figure 11, the observation needs to exceed around 800 to be able to have a sufficient result with power over 80%. Since the actual sample sizes for both groups in our dataset exceed the required sample sizes, we have a sufficient number of observations to conduct meaningful analyses.

(Alpha=0.05, Power=0.8)	**Stripped**	**Not Stripped**
Effect Size	0.14	
Sample Size	432	3185
Actual Size	7801	57451

Table 3: Power Analyst for Stripped and Not-stripped

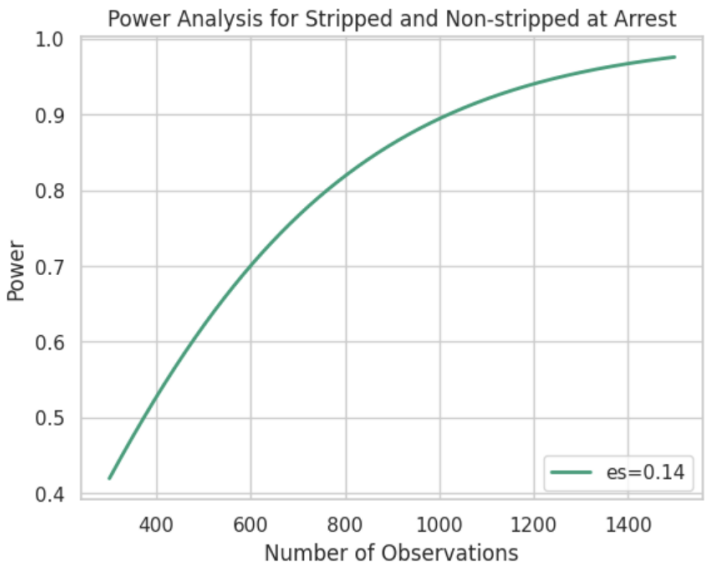

Figure 11: Power Analyst for Stripped and Not-stripped

After thoroughly exploring the EDA section and power analysis above, we used ANCOVAs for answering our first two research questions. ANCOVAs allow us to test for significant differences in the continuous dependent variable with independent variables that can be both categorical and continuous. Logistic Regression was employed to answer the last research question. In logistic regression we predict a categorical variable using both continuous and categorical variables as independent variables. Additionally, note that the train-test split was 80-20 and stratified to ensure the same balance of y in both sets.

Results

ANCOVA

We conducted two ANCOVA tests. For the first ANCOVA, the dependent variable is "Age_arrest," the covariate is "Arrest_Month," and the between-subjects factor is "Perceived_Race." As shown in Table 4, this analysis showed that the uncorrected p-value for "Perceived_Race" was less than 0.05, indicating a statistically significant difference in the mean age at strip search among the different perceived race groups after controlling for the covariate "Arrest_Month."

Source	SS	DF	F	p-unc	np2
Perceived_Race	50574.757765	7	63.52494	2.4335e-89	0.053994
Arrest_Month	1249.942451	1	10.99002	9.2022e-04	0.001409

Residual	886104.03387	7791	NaN	NaN	NaN

Table 4: ANCOVA for Age and Perceived_race at Strip search

For the second ANCOVA, the dependent variable is "Age_arrest," the covariate is "Arrest_Month," and the between-subjects factor is "Sex." As shown in Table 5:the uncorrected p-value for "Sex" was 2.4586e-29, which is less than 0.05 and therefore indicates a statistically significant difference in the mean age at arrest between males and females, after controlling for the effect of "Arrest_Month." We hypothesized that there is a relationship between sex and age at arrest, after accounting for the effect of arrest month.

Source	**SS**	**DF**	**F**	**p-unc**	**np2**
Sex	1.996494e+04	2	65.94191	2.4586e-29	0.002017
Arrest_Month	5.761744e+03	1	65.94191	6.8977e-10	0.000583
Residual	9.876821e+06	65244	NaN	NaN	NaN

Table 5: ANCOVA for Age and Sex

Logistic Regression

We broke down the variables by one-hot encoding. The reference category is Black Female. The results are below, with only the significant variables.

Variable	**Coeff**	**Std Err**	**Lower CI**	**Upper CI**	**Odd-Ratio**
Intercept	-1.674	0.053	0.168	0.207	0.18
Age	-0.0129	0.001	0.984	0.989	0.987
M	0.301	0.037	1.25	1.45	1.35
SouthEastAsian	-0.596	0.067	0.482	0.628	0.55
Indigenous	0.292	0.074	1.159	1.547	1.33
Latino	-0.678	0.105	0.412	0.623	0.50
MiddleEastern	-0.728	0.080	0.412	0.601	0.48
SouthAsian	-0.654	0.075	0.448	0.601	0.51

Unknown	-0.260	0.057	0.689	0.862	0.77

Table 6: One-hot encoding result

The one-hot encoded column for White and U i.e. Unknown gender came out to be insignificant. In terms of the results we see the following,

- Increasing your age by one unit, decreases your odds of getting strip searched by 1.3%
- If you are a male, then your odds of being strip searched increase by 35%
- If you are South Asian, your odds of being strip searched decrease by 45%
- If you are Indigenous, your odds of being strip searched increase by 33%
- If you are Latino, your odds of being strip searched decrease by 50%
- If you are Middle Eastern, your odds of being strip searched decrease by 52%
- If you are from a Legacy race, your odds of being strip searched decreases by 23%

All of these results showcase that being Black will increase your odds of being strip searched except in the case of Indigenous. Moreover note, all of these changes assume that when one variable changes the rest are kept constant.

Now, in terms of the effectiveness of the model, we calculated its accuracy and f1-score on both the test dataset and train dataset. The results below in Table 7 indicate high accuracy but a terrible f1-score. Our model is essentially predicting everything to 0 i.e. no strip search. This is still resulting in a high accuracy because of the fact the 0 represents ~88% of the dataset.

Metric	**Training Data**	**Test Data**
Accuracy	88.045%	88.045%
F1 Score	0	0

Table 7: Metrics for Logistic Regression Performance

The confusion matrix for training data in Table 8 and for test data in Table 9 also indicates the same weakness as presented above.

Train Data	**Predicted 0**	**Predicted 1**
Actually 0	45,958	0
Actually 1	6,240	0

Table 8: Confusion Matrix for Training Data

Test Data	**Predicted 0**	**Predicted 1**
Actually 0	11,490	0
Actually 1	1,560	0

Table 9: Confusion Matrix for Test Data

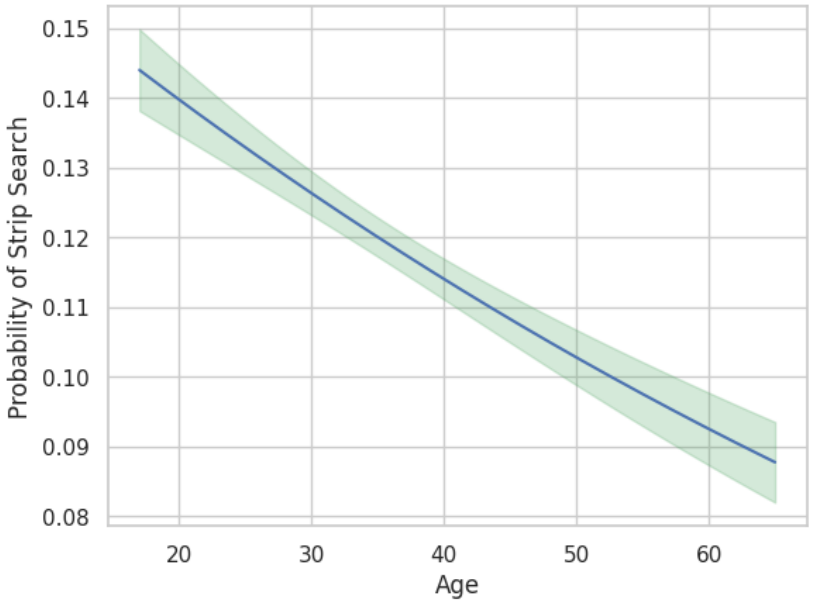

Figure 12: Prediction Interval for Age and Probability of Strip Search

Additionally, we looked at the prediction interval for Age and the odds of getting strip searched. As Figure 12 indicates, the probability of strip search falls with age, similar to what our logistic regression found. The prediction interval is close to the regression line as well indicating that the prediction interval is not huge, i.e. our results will be accurate.

Discussion

Through EDA, t-tests, and one-way and two-way ANOVAs, we examined the research questions we began with. With a combination of all these tools, we could answer our questions. Let's dive into them one by one again.

Is there significant differences in age at which individuals of different races are subjected to strip searches during an arrest with time as a control variable?

This question was answered with ANCOVA analyst. The significant p-value for perceived race in the ANCOVA suggests that there are differences in the mean age at arrest among individuals of different perceived races who were strip searched after arrest, even after controlling for the effect of arrest month. These findings are consistent with previous research and analyst from t-test which indicated that there is a 1.5 to 2.03 years difference in average age in strip search. This finding highlights the ongoing issue of racial disparities within the criminal justice system and is consistent with previous research that has shown that individuals from certain racial and ethnic groups are more likely to be involved in the criminal justice system and to experience harsher treatment than their white counterparts (Chan & Chunn, 2017). By identifying and addressing the factors that contribute to these disparities, it may be possible to develop more equitable and just policies and practices within the criminal justice system.

Are there any statistically significant differences in the mean age of arrest between different groups defined by sex with time as a control variable?

Research shows that individuals of color are more likely to be subject to various forms of harsh treatment within the system, such as strip searches, while males are more likely to be involved in the system and to experience harsher treatment than females. Building on these findings, you conducted a subsequent analysis of sex and age at arrest, with the goal of better understanding the factors that contribute to disparities in the criminal justice system. Therefore, we conducted an ANCOVA analysis which showed that there were significant differences in the mean age at arrest between males and females, even after controlling for the effect of arrest month. This finding highlights the need to better understand the ways in which sex interacts with other factors, such as race and ethnicity, to create disparities within the system.

Are sex, race, and age significant factors that affect chances of being strip searched?

All of our variables turned out to be significant except Unknown gender and White individuals. Our results indicated that males are 35% more likely to get arrested which isn't as big as the odds indicated by Newburn, Shiner, and Hayman (2004). On the other hand, as age increased, the odds decreased which is opposite of what Newburn, Shiner, and Hayman (2004) detected. This may be caused by spatial differences across their data and ours. While Black individuals have the relatively higher odds, Indigenous individuals had the highest odds of being strip searched. These results indicate that while all the variables were significant and had noticeable differences in odds, there is no one-size-fits-all conclusion for strip searches, race, gender, and age. It is recommended to further study this with more data.

Limitations

One of the limitations of this study that we would like to point out is that the perceived race of an individual is determined by the police officer. What this results in is that if a person is arrested multiple times by different police officers then the person's perceived race may be different. For example, Person ID 327535 was first identified as Black on their first two arrests, and then South Asian on their third arrest. This is a major limitation of our data that we had to bear with during our study. There were plenty of outliers in terms of the number of arrests, but one interesting data point that stood out was the maximum number of arrests: one individual was arrested a staggering 54 times.

Moreover, it is important to note that the analysis cannot establish causation. While the results suggest that there may be a relationship between certain factors and disparities in the criminal justice system, it's possible that other factors not included in the analysis may be driving the observed patterns. For example, factors such as geographic location may play a role but were not included in the analysis. We could make our analysis more visual by mapping the arrest location with the Police division that made the arrest. That would have shown if there were any arrest biases by neighborhood or police activity.

Conclusion

It appears that there are significant differences in age and perceived race among individuals who are subjected to strip searches after arrest. The logistic regression analysis also revealed significant differences in odds based on various demographic factors. These findings suggest that disparities may exist in the criminal justice system, particularly in the context of post-arrest procedures. Additionally, our power analysis indicated that the sample size was sufficient to achieve our target, providing more confidence in the results obtained from the subsequent analyses.

It's important to note, however, that there are limitations to these analyses. The accuracy of the perceived race variable may be questionable, and other factors that were not included in the analysis may also contribute to disparities. Additionally, the analyses cannot establish complete causation.

Overall, these findings highlight the need for further research and examination of the factors that contribute to disparities in the criminal justice system. It's important to ensure that individuals are not subjected to discriminatory practices, and that procedures such as strip searches are conducted fairly and equitably.

References

Chan, J., & Chunn, J. (2017). Racialization and criminalization: Examining the overrepresentation of Black Canadians in the criminal justice system. Canadian Journal of Criminology and Criminal Justice, 59(2), 165-186. https://doi.org/10.3138/cjccj.2016.E02

Ulmer, J. T., & Steffensmeier, D. (2014). The Age and Crime Relationship: Social Variation, Social Explanations. Retrieved from https://www.sagepub.com/sites/default/files/upm-binaries/60294_Chapter_23.pdf

Keeton, A. (2015). Strip Searching in the Age of Colorblind Racism: The Disparate Impact of Florence v. Board of Chosen Freeholders of the County of Burlington. Michigan Journal of Race and Law, 21. https://repository.law.umich.edu/cgi/viewcontent.cgi?article=1050&context=mjrl

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. https://doi.org/10.1093/bjc/azh036

Toronto Police Service. 2021. Arrests and Strip Searches (RBDC-ARR-TBL-001). [Dataset]. Toronto Police Service. https://doi.org/10.21232/TorontoPS-arrests-and-strip-searches-rbdc-arr-tbl-001

McDowall, D., Loftin, C., & Pate, M. (2012). Seasonal Cycles in Crime, and Their Variability. Journal of Quantitative Criminology, 28(3), 389–410. https://doi.org/10.1007/s10940-011-9145-7

Andresen, M. A., & Malleson, N. (2013). Crime seasonality and its variations across space. Applied Geography, 43, 25–35.

https://doi.org/10.1016/j.apgeog.2013.06.007