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

INF2178
Experimental Design for Data Science

Professor Shion Guha
April 16, 2023

Statement of Individual Contribution:

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

Introduction

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

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

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

Literature Review

Racial Bias of Strip Searches

Strip Search has always been a sensitive topic due to the potentially humiliating nature of the procedure. Strip searches involve the removal of an individual's clothing to search for weapons or prohibited items. Strip Searches are highly invasive and can make individuals feel exposed and vulnerable.

Strip searches are a legal way to perform detection against prohibited items possessed by suspects only if a law enforcement agent reasonably suspects that the individual possesses a weapon or restricted items. Furthermore, a report document is needed after officers perform strip searches on individuals. The paper should include the reason and justification for the strip search. Only having the above requirements can make the strip search legal.

However, miss usage of racial profiling by law enforcement officers can lead to individuals from certain ethnic groups being targeted for strip searches more frequently. Racial profiling or discrimination, or racism, occurs when police officers use an individual's race or ethnicity as a factor in their decision to stop, search, or arrest them rather than relying on evidence of criminal behaviour.

A study, “Race, Crime and Injustice?” by Tim Newburn, Michael Shiner, and Stephanie Hayman, reveals the secret behind crime injustice and different ethnic groups. The authors specifically discussed using strip searches by officers on other racial groups to see whether criminal injustice exists within the justice system. (Newburn, 2004)

The study's authors sought to investigate the factors influencing the likelihood of an individual being subjected to a strip search upon arrest. Using a multivariate analysis, they considered various factors, including the arrestees' age, ethnicity, and the reason for their arrest. (Newburn, 2004)

The study revealed significant disparities in the probability of being strip-searched across different ethnic groups. For instance, African-Caribbean individuals were found to be at a considerably higher risk of being strip-searched, with their probability being twice as high as that of white individuals. (Newburn, 2004)

On the other hand, the research found that individuals of Arabic or Asian descent were less likely to be subjected to strip searches than white arrestees. While this may seem like a positive finding, it still raises questions about the equitable application of law enforcement procedures and whether biases may also be at play in these instances. (Newburn, 2004)

In conclusion, the study highlights the importance of understanding the factors contributing to the unequal application of strip searches among different ethnic groups. It emphasizes the need for further research and policy reform to ensure that law enforcement practices are equitable and unbiased, protecting the rights and dignity of all individuals, regardless of their ethnicity.

Number of Arrests of An Individual on Race and Gender

Over the past several decades, the issues of racism and sexism have gained significant attention, particularly in sensitive areas such as police activities. These societal issues have raised concerns about fairness and equality within law enforcement and the criminal justice system. The book “Gender, Race, and Crime: An Analysis of Urban Arrest Trends” by Chilton and Datesman looks into the influence of gender and race on criminal activities and arrest rates. (CHILTON & DATESMAN, 1987)

Chilton and Datesman's study discovered that females had lower arrest rates than males, which may suggest that males are more likely to engage in criminal activities than females. (CHILTON & DATESMAN, 1987) Social expectations and stereotypes for different races may also influence this difference. These biases can result in the other treatment of males and females by law enforcement and the criminal justice system.

Furthermore, the authors found that African Americans had a higher arrest rate than their percentage in the overall population. In contrast, arrest rates for white, Hispanic, and Asian individuals were proportionate to their representation in the population. This disparity in arrest rates may indicate that African Americans are more likely to be arrested than other racial groups. (CHILTON & DATESMAN, 1987)

Several factors could contribute to the differences in arrest frequencies among different racial and ethnic groups, including socio-economic factors, systemic discrimination, and implicit

biases within law enforcement agencies. It is crucial to address these underlying issues to create a more equitable and just criminal justice system.

In conclusion, the book “Gender, Race, and Crime: An Analysis of Urban Arrest Trends” by Chilton and Datesman sheds light on the potential influence of gender and race on criminal activities and arrest rates. While the findings may indicate that males are more likely to participate in illegal activities than females and that African Americans have higher arrest rates than other racial groups, it is essential to consider the broader societal context and potential biases that may contribute to these differences. By addressing the causes of racism and sexism, we as human beings can work towards creating a more equitable and just society environment.

EDA

Data Cleaning

Since the original dataset contains many irrelevant variables to our research questions. Therefore, we first selected several variables, including ‘Sex’, ‘StripSearch’, ‘Youth_at_arrest__under_18_years’, 'PersonID', and 'Perceived_Race', for the working data frame. The above four variables (except `PersonID`) are included in the research questions.

Then we start cleaning the misleading or inaccurate values. For example, the variable ‘Youth_at_arrest__under_18_years’ contains values ‘Youth (aged 17 and younger)’ and ’Youth (aged 17 years and under)’, which means the same thing. As a result, we classified those two values as ‘Youth’ and the rest as ‘Adults’.

After this, we divided the perceived race into minority and majority. The logic behind how we draw the line between minority and majority people who are classified as ‘White’, or ‘Aboriginal’ is considered as the majority. Other races, including ‘black’, ‘east Asian’ etc, are considered a minority.

Last but not least, we created a new column called ‘total_arrest’. The total arrests column uses the number of occurrences of each PersonID to calculate the number of suspect arrests. This column is a continuous variable. This new variable, `total_arrest` will contribute to analyzing both of the research questions we have for this project. As a result, the working data frame contains six columns (variables).

Descriptive Statistic

Based on the descriptive statistics we calculated on the new variable ‘total_arrest’, the mean of the ‘total_arrest’ is 4.13; on average, a suspect gets arrested 4.13 times from 2020 to 2021. The 25% quantile is 2.00, which means 25% of the values lie below 2.00, and 75% of the answers lie above 2.00. The 75% quantile is 5.00, representing 75% of the values lying below 5.00 and 25% of the answers lying above 5.00. The standard deviation measures how

dispersed the data is to the mean. In this case, the standard deviation is 5.40, which means the value of 'total_arrest' is distributed compared to the mean. Looking at the mean of 4.13 and the median of 2.00, we can tell without looking at the histogram that this variable does not follow a normal distribution, and the histogram would be right skewed.

The descriptive statistic of the variable `Perceived_Race`after splitting eight different ethnicity groups into two. The two groups had approximately closed counts (the majority being 29657, and the minority being 35615); rather than before, the number of white and Latino people differed by ten times. By doing this operation, we could avoid bias from the huge population gap.

Data Visualization

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

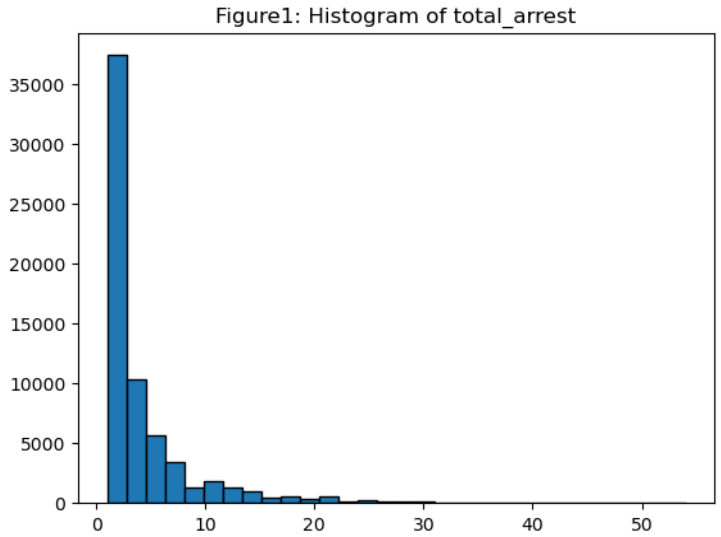


Figure 2 is the box plot of the 'total_arrest'. The box plot shows that the 25% quantile is approximately 1, the median is about two, and the 75% quantile is around five. This aligns with the result in the descriptive statistics section. Furthermore, the upper bar of the box plot represents the maximum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

Figure2: Boxplot showing distribution of total_arrest for criminals

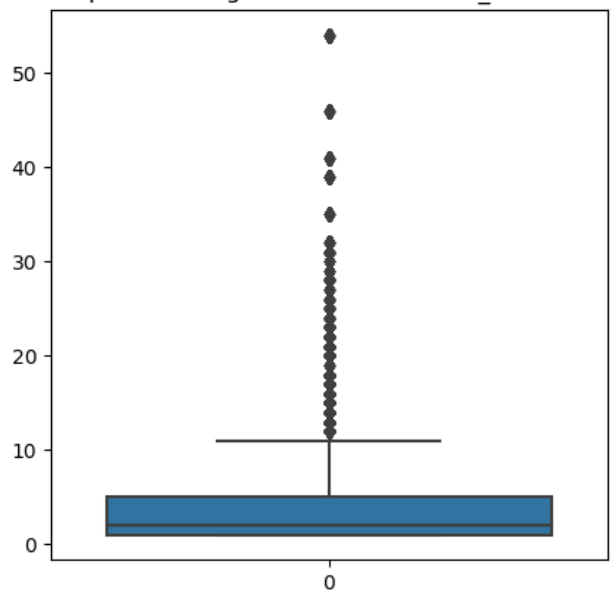


Figure 4 is the box plot of 'Sex' and 'total_arrest'. After seeing the plot, we find that males and females have approximately the same distribution. The only difference is the distribution of outliers. Male suspects have visibly more outliers (falls beyond the upper bar, 75% quantile + 1.5IQR).

Figure3: Boxplot of sex and total_arrest

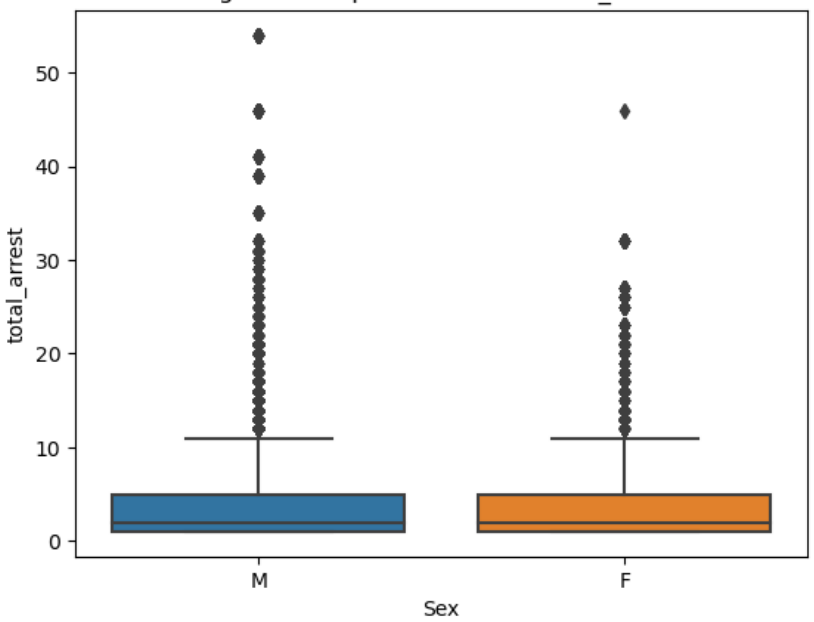


Figure 4 is the box plot of the `Perceived_Race` against the continuous variable `total_arrest`. Comparing two racial ethnicity groups, the twenty-five percentile is approximately the same, around one arrest. The median, seventy-five percentile, and upper bar differ significantly, even by visual inspection. The median of total arrests for the majority group is around four arrests, and the median of total arrests for the minority group is around two arrests. The seventy-five quantile is seven arrests for the majority and five for the minority. Furthermore, the outliers of minority groups are more than the majority group. This may indicate that

people in the minority group have fewer arrests than people in the majority group; however, some people commit far more crimes and get arrested.

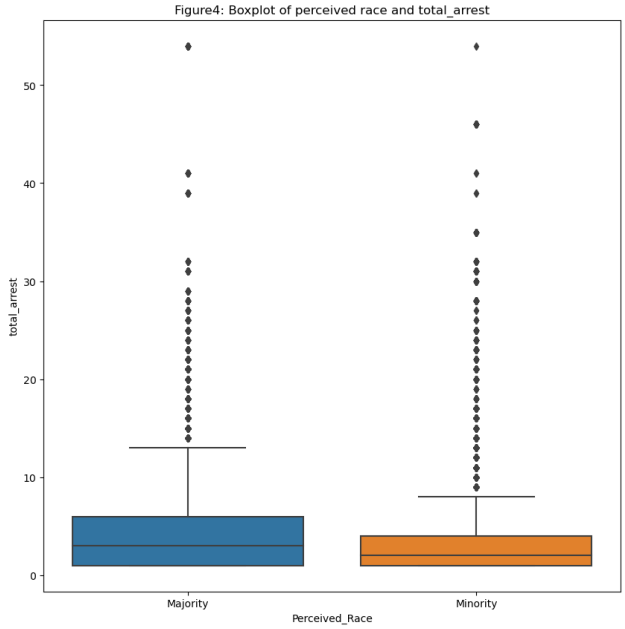
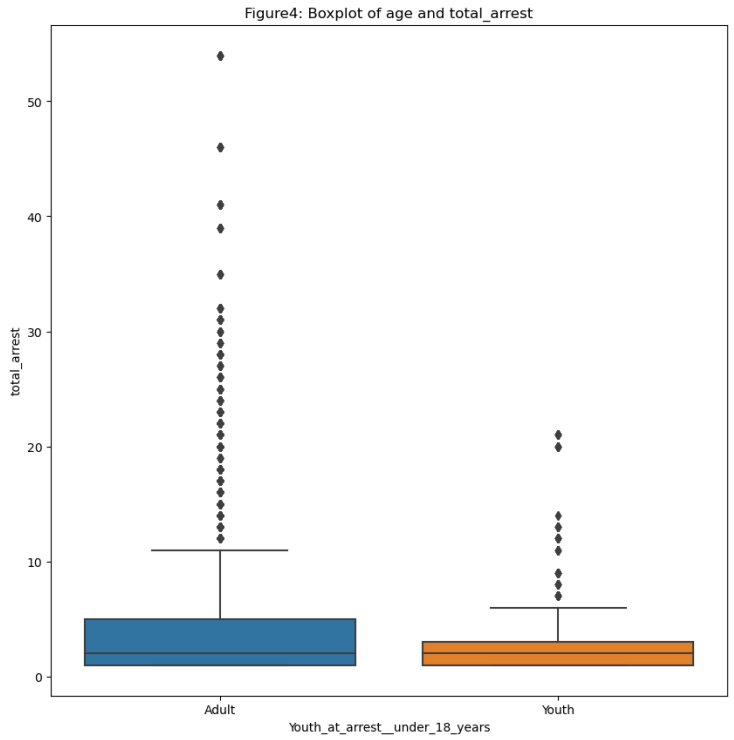


Figure 5 is the box plot of the two age groups against the continuous variable `total_arrest`. Comparing the two age groups, we see that the twenty-five percentile and median are approximately the same, around one and two arrests, respectively. At the same time, the seventy-five percentile and upper bar differ significantly, even just by visual inspection. The seventy-five quantile is around six arrests for adults and three for youth. Furthermore, the outliers of the youth group are less than the adult group. This may indicate that people in the youth group have fewer arrests than people in the adult group; however, some people commit far more crimes and get arrested.



Test of Normality:

Anderson-Darling Test

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

H0: The data distribution follows a normal distribution.

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

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

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

T-Test

Equal Variance Assumption Checks

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

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

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

$(Var_A = Var_B)$

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

$(Var_A \neq Var_B)$

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

Gender and Strip Search

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

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

$(\mu_M = \mu_F)$

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

$(\mu_M \neq \mu_F)$

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

Age Groups and Strip Search

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

H0: There is no difference in strip searches conducted between youth and adult suspects exists.

$(\mu_Y = \mu_A)$

H1: There is a difference in strip searches conducted between youth and adult suspects.

$(\mu_Y \neq \mu_A)$

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

Perceived Race and Strip Search

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

H0: There is no difference in strip searches conducted between majority and minority group suspects.

$(\mu_{maj} = \mu_{min})$

H1: There is a difference in strip searches conducted between majority group and minority group suspects.

$(\mu_{maj} \neq \mu_{min})$

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

Research Design and Method

Data Description

In this project, we will deal with the dataset provided by Toronto Police Service Public Safety Data Portal. The dataset can be found with the following link: https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/about. This dataset contains 65,276 records of arrests in the GTA area from 2020 to 2021. It also contains records and descriptive statistics of strip searches conducted by police officers. A strip search is a type of search performed by a law enforcement officer on an individual that involves the removal of some or all of their clothing and visually inspecting their body. Each of the 65,276 rows in the dataset represents an incident of arrest or strip search performed by the Toronto Police Service. The 24 columns provide information on the demographic and behavioural details of the person arrested or searched. The demographic information includes arrest month, person ID (the suspect's ID), perceived race (the suspect's perceived race), sex, age group, and occurrence category (the nature of the crime). The behavioural information includes actions at arrest-Combative, actions at arrest-assaulted, actions at arrest-cooperative, and reasons to conduct strip searches. All variables in this data set are either categorical or binary (0 or 1). Due to the fact that the description of this dataset is limited, we assume that the person ID refers to the suspects' ID.

Research Objectives

Our study will focus on finding out how demographical bias affects the decision-making process of police officers upon making an arrest. Moreover, we want to see if the interaction aligns with the above literature. Therefore, we seek to investigate the following research

questions. We gained insights into these questions from the above literature and our preliminary dataset analysis.

- Research Question 1: What is the relationship between demographic status (gender, perceived race, and age group) and the likelihood of being conducted strip searches?
- Research Question 2: Does the perceived race significantly affect the total number of arrests while holding strip search constant?

The first research question focuses on analyzing how demographic status, including gender, perceived race and age group affect the chance of being conducted strip searches. The second research question examines whether there is a potential bias in law enforcement based on perceived race. Studies have shown disparities in arrest rates across different racial and ethnic groups. Some of these disparities may be due to implicit or explicit bias among law enforcement officers. Both research questions can contribute to a better understanding of potential biases within the law enforcement systems and, therefore, can help identify areas of improvement.

ANCOVA

ANCOVA (Analysis of Covariance) is a statistical method incorporating variance analysis (ANOVA) with regression analysis. ANCOVA is used to evaluate the effect of a categorical independent variable on a continuous dependent variable while controlling for the influence of one or more continuous covariates.

For our research questions, we will choose the number of arrests as the dependent variable for our research questions. The perceived race variable is the dependent variable. The strip search variable is the controlled variable.

By holding the strip search constant, we control for a potential confounding variable that may affect the relationship between perceived race and the number of arrests. Strip searches are invasive and potentially humiliating procedures that may be more likely to be conducted on certain racial or ethnic groups, leading to higher arrest rates for those groups. By controlling for a strip search, we can examine the effect of perceived race on the total number of arrests while holding constant a potentially confounding variable.

When analyzing the result of ANCOVA, we should look at P-values, the coefficient of independent variables and covariates. When examining p-values and F statistics, if the p-value is smaller than 0.05 or the F statistic is significant. We should reject the null hypothesis and accept that at least one of the independent variables (perceived race) or covariate (strip search) has a significant effect on the dependent variable (total arrests).
When evaluating coefficients, a positive coefficient means that an increase in the value of the independent variable (perceived race) or covariate (strip search) is associated with an increase in the value of the dependent variable (total arrests). A negative coefficient indicates that an increase in the value of the independent variable or covariate is associated with a decrease in the value of the dependent variable.

Power Analysis

Power analysis is a statistical method used to determine the minimum sample size needed for a study to detect a statistically significant effect with a certain confidence level. It involves calculating the statistical power of a study. Statistical power is the probability of correctly rejecting the null hypothesis when it is false.

Furthermore, we should look at the power plot. A power plot is a graph that shows the relationship between the sample size, the effect size, and the statistical power. It can be used to visualize the trade-off between the sample size and the statistical power of a study and to determine the minimum sample size required to achieve a certain level of power. The x-axis represents the sample size, the y-axis represents the statistical power, and curves usually represent the effect size.

Logistic Regression

Logistic regression is a statistical method used for analyzing the relationship between a binary dependent variable (StripSearch) and one or more independent variables (age group, gender, and perceived race). In other words, it helps predict the probability of a strip search based on the predictor variables mentioned above. Logistic regression is particularly useful when the relationship between the independent and dependent variables is not linear and the dependent variable is binary. In our case, the dependent variable strip search, independent variable age group, perceived race, and gender are all binary variables.

The null hypothesis (H0) assumes no relationship between the independent and dependent variables. This means the coefficients of the independent variables in the model are equal to zero.

The alternative hypothesis (H1) is that there is a significant relationship between the independent variables and the dependent variable. This means that at least one of the coefficients is not equal to zero.

Assuming the null hypothesis is true, the P-value represents the probability (chance) of observing a result as extreme as the observed test statistic. Our selection of significance level is 0.05. If the p-value is less than the significance level of 0.05, we would reject the null hypothesis and conclude that at least one of the coefficients is not equal to zero.

Accuracy Score

The accuracy score is a simple and easily understandable way to evaluate the performance of a logistic regression model. It represents the proportion of correct predictions made by the model out of the total number of predictions. In other words, the accuracy score tells you how often the model is correct in its predictions. The output is expressed as a value between 0 and 1. A score of 1 indicates that the model has 100% of predictions correctly, and a score of 0 means that the model has not made any correct predictions.

Confusion Matrix

A confusion matrix is a table that helps us visualize the performance of a classification model by showing the number of correct and incorrect predictions for each class. There are four classes, including

True Positives (TP): The number of times the model correctly predicted the positive class (e.g., a strip search occurred).

True Negatives (TN): The number of times the model correctly predicted the negative class (e.g., no strip search occurred).

False Positives (FP): The number of times the model incorrectly predicted the positive class (e.g., predicted a strip search, but none occurred).

False Negatives (FN): The number of times the model incorrectly predicted the negative class (e.g., predicted no strip search, but one occurred).

Prediction Interval

A prediction interval is a range in which an observation is likely to fall with a certain confidence level (usually 95%). It provides an estimate of the uncertainty around a single predicted value. It is commonly used in regression models, including logistic regression, to give the range within which the predicted values may vary.

Assumption Checks

Independence of observation

The independence of observation assumption can be examined by investigating the study design or the data collection method.

Test of Normality

The test of normality was conducted before running the T-Tests. Although the same test is required for ANCOVA and logistic regression, running the test once is enough.

Homogeneity of Variance Assumption Check

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

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

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

$(Var_A = Var_B)$

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

$(Var_A \neq Var_B)$

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

Homogeneity of Regression Assumption Check

ANCOVA is the assumption that the regression slopes of the covariate are equal across the different levels of the categorical independent variable.

Results

Test of Normality

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

Research Question 2

What is the relationship between demographic status (gender, perceived race, and age group) and the likelihood of being conducted a strip search?

T-test of gender and strip search

A Welch's t-test was conducted to compare the mean scores between the strip searches of male criminals and the strip searches of female criminals. The t-statistic was 7.25 with a p-value less than 0.05, indicating a significant difference between the two groups. Therefore, the strip search of male and female criminals was rejected because there were differences between the two groups.

T-test of age and strip search

A Welch's t-test was conducted to compare the mean scores between the different age groups and strip searches. The t-statistic was 5.33, with a p-value less than 0.05, indicating a significant difference between the two groups. Therefore, the null hypothesis was rejected because there were differences between the two variables.

T-test of perceived race and strip search

A Welch's t-test was conducted to compare the mean scores between the different race groups and strip searches. The t-statistic was -7.90 with a p-value less than 0.05, indicating a significant difference between the two groups. Therefore, the null hypothesis was rejected because there were differences between the two variables.

Power analysis of gender and strip search

In order to identify a significant effect of gender on strip search outcomes, we conducted a power analysis to determine the necessary sample size. Cohen's d was calculated as the effect size for Gender vs StripSearch, resulting in a value of 0.07. In order to achieve 80% power at a 0.05 alpha level, we determined that a sample size of nobs1 = 2097 participants would be required to detect the main effect of gender. Additionally, a sample size of nobs2 = 8753 participants would be required to detect the interaction effect of gender and strip search.

Upon data collection, the sample size for the first group was determined to be 52659, while for the second group, it was 12617. The power curve, which is based on these actual sample sizes, reveals that the test's power is exceptionally high (close to 1.0) for both groups. This indicates that the study has a high likelihood of detecting the small effect size.

Power analysis of age and strip search

In order to identify a significant effect of age on strip search outcomes, we conducted a power analysis to determine the necessary sample size. Cohen's d was calculated as the effect size for Age vs StripSearch, resulting in a value of 0.09. In order to achieve 80% power at a 0.05 alpha level, we determined that a sample size of nobs1 = 21353 participants would be required to detect the main effect of gender. Additionally, a sample size of nobs2 = 1043 participants would be required to detect the interaction effect of age and strip search.

Upon data collection, it was discovered that the actual sample size for the first group (AS) was 62234, and for the second group (YS), it was 3042. The power curve, which is based on these actual sample sizes, indicates that the test's power is significantly high (close to 1.0) for the first group. This suggests that the study has a high probability of detecting the small effect size in that group. However, the power of the test for the second group is relatively low, indicating that the study may encounter difficulty in detecting the effect size in that group.

Power analysis of perceived race and strip search

In order to identify a significant effect of perceived race on strip search outcomes, we conducted a power analysis to determine the necessary sample size. Cohen's d was calculated as the effect size for Perceived Race vs StripSearch, resulting in a value of 0.16. In order to achieve 80% power at a 0.05 alpha level, we determined that a sample size of nobs1 = 47183 participants would be required to detect the main effect of gender. Additionally, a sample size of nobs2 = 18089 participants would be required to detect the interaction effect of perceived race and strip search.

Upon data collection, it was discovered that the first group (AS) had an actual sample size of 62234, while the second group (YS) had 3042. The power curve, based on these sample sizes, reveals that the test's power is significantly high (close to 1.0) for the first group, suggesting that the study has a high chance of detecting the small effect size in that group. However, the power of the test for the second group is relatively low, indicating that the study may face challenges in detecting the effect size in that group.

Logistic Regression

We used logistic regression analysis to investigate the relationship between "StripSearch" as the dependent variable and three independent variables: "Sex", "Perceived_Race", and "Youth_at_arrest__under_18_years". The analysis involved 52,217 observations and employed the maximum likelihood estimation (MLE) method. The results revealed a significant connection between the dependent variable and the independent variables, as demonstrated by the pseudo-R-squared value of 0.01.

Further analysis showed that the intercept had a negative coefficient of -2.06, indicating that the log odds of StripSearch decreased by a factor of 0.13, keeping all other variables constant. Sex had a positive coefficient of 0.25, implying that males had 1.279 times higher odds of being strip-searched compared to females. Perceived_Race had a negative coefficient of -0.55, indicating that belonging to a minority group was associated with 0.57 times lower odds of being strip-searched compared to non-minority groups. Youth_at_arrest__under_18_years had a negative coefficient of -0.36, suggesting that being under 18 years old at the time of arrest was associated with 0.70 times lower odds of being strip-searched.

All coefficients were significant at $p < 0.05$, and we also reported the 95% confidence intervals and odds ratios for each variable. Overall, the findings suggest that gender, perceived race, and youth at the time of arrest are critical predictors of strip search outcomes.

Logistic Regression Result

Dep. Variable	Strip Search	No. Observation	52217
Model	Logit	Df Residual	52213
Method	MLE	Df Model	3
Date	Sun, 16 Apr 2023	Pseudo R-squ	0.003
Time	10:32:29	Log-Likelihood	-18959
Converged	True	LL-Null	-19020
Covariance Type	Nonrobust	LLR P-value	5.45*10^26

	coef	st error	z	P	[0.025	0.975]
Intercept	-2.07	0.04	-59.21	0.00	-2.14	-2.01
Sex	0.23	0.04	6.43	0.00	0.16	0.31
Perceived race	-0.21	0.03	-7.65	0.00	-0.26	-0.16
Youth at arrest under 18	-0.32	0.07	-4.38	0.00	-0.47	-0.18

Odd Ratio

	Lower CI	Upper CI	OR

Intercept	0.12	0.13	0.13
Sex	1.18	1.36	1.26
Perceived Race	0.80	0.86	0.81
Youth at arrest under 18	0.63	0.84	0.73

Predicted Response

Accuracy Score,
The logistic regression model, achieved an accuracy score of 0.8771, which means that the model predicted the outcome correctly for 87.71% of the test data set.

Confusion Matrix

According to the confusion matrix (table 2), the model accurately predicted 11,451 instances where a "strip search" was not conducted, resulting in true negatives. However, in 1,605 cases where a "strip search" was conducted, the model failed to predict it, resulting in false negatives. Surprisingly, no true positives or false positives were identified in the test data set.

Table 2

11451	0
1605	0

Research Question 1

Does the perceived race significantly affect the total number of arrests while holding strip search constant?

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

Equal Variance Assumption Check
Prior to conducting an ANCOVA-test, the assumption of the equality of variances was tested using a Levene test. The results of the Levene test indicated that the variances between all

race groups were significantly different since the p-value is 4.05*10-203, which is smaller than the significance level of 0.05.

ANCOVA

We conducted an analysis of covariance (ANCOVA) with Perceived_Race and StripSearch as the independent variables and a continuous variable as the dependent variable, and the results are presented in Table1.

The analysis indicated that Perceived_Race was a significant predictor of the dependent variable, with an F-value of $F(1.65) = 914.85$, $p < 0.001$, and a partial eta squared value of 0.0138. Similarly, StripSearch was also a significant predictor of the dependent variable, with an F-value of $F(1.65) = 1061.98$, $p < 0.001$, and a partial eta squared value of 0.02.

The residual sum of squares (SS) was 1.85*10^6, and the degrees of freedom (DF) was 65269. However, no F-value or p-value was computed for the residual. In summary, these findings suggest that both Perceived_Race and StripSearch have a substantial impact on the dependent variable.

Table 1

	Source	SS	DF	F	P
0	Preceived Race	2.84*10^4	1	1008.91	9.93*10^-220
1	Strip Search	3.22*10^4	1	1144.13	1.21*10^248
2	Residual	1.84*10^6	65269	NaN	NaN

Discussion

The purpose of this study was to explore the correlation between demographic factors, such as gender, perceived race, and age group, and the likelihood of being strip-searched during a police arrest. Additionally, this study aimed to determine if perceived race had a significant impact on the total number of arrests while keeping strip search constant.

The results of the study indicate that demographic status has a significant impact on the probability of being strip-searched during a police arrest. Specifically, the study found that males were more likely to undergo strip searches than females. Additionally, individuals who perceived themselves as belonging to a racial or ethnic minority group were more likely to be strip-searched compared to those who identified as White. Younger individuals were also more likely to be strip-searched compared to older individuals.

These findings are consistent with prior research, which has shown that particular demographic groups are more likely to experience strip searches during police arrests.

However, this study provides further insight into the specific demographic factors that contribute to this phenomenon.

Regarding the second research question, the results indicate that perceived race has a significant impact on the total number of arrests, even when controlling for the occurrence of strip searches. The study found that individuals who identified as racial or ethnic minorities were arrested more frequently than their White counterparts.

These findings are concerning, as they suggest that racial prejudice could influence police practices, even when specific procedures such as strip searches are held constant. Therefore, it is crucial for law enforcement agencies to address this issue and take necessary measures to ensure that all individuals are treated equitably and without bias during the arrest process.

Overall, this study provides valuable insights into the correlation between demographic status and police practices like strip searches. These findings underscore the necessity for further research in this area to better comprehend the factors contributing to these disparities and develop effective approaches to address them.

Conclusion

To sum up, this study delved into the connection between demographic status (gender, perceived race, and age group) and the possibility of being strip-searched during an arrest. The logistic regression analysis demonstrated that gender, perceived race, and youth at the time of arrest were critical predictors of strip search outcomes. Specifically, males had a 1.279 times higher likelihood of being strip-searched than females, while belonging to a minority group was associated with 0.57 times lower odds of being strip-searched compared to non-minority groups. Being under 18 years old at the time of arrest was also associated with 0.70 times lower odds of being strip-searched. These findings highlight the importance of police officers being aware of these demographic factors and avoiding discriminatory practices when conducting strip searches.

The second research question aimed to investigate whether perceived race had a significant effect on the total number of arrests while holding strip search constant. The ANCOVA analysis revealed that both perceived race and strip search had a substantial impact on the dependent variable. These findings suggest that perceived race may be a contributing factor to the number of arrests made, even when controlling for the effects of the strip search.

Overall, this study emphasizes the significance of comprehending how demographic factors, such as gender, perceived race, and youth at the time of arrest, influence police practices, particularly concerning strip searches. Policymakers and law enforcement agencies should take note of these findings and work towards developing policies and procedures that prevent discrimination and ensure equitable treatment for all individuals in the criminal justice system.

Reference

Newburn, T. (2004). Race, crime and injustice?: Strip Search and the treatment of suspects in custody. *British Journal of Criminology*, *44*(5), 677–694. https://doi.org/10.1093/bjc/azh043

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