Is there any difference or bias of stripsearch in terms of race when being arrested?

Faculty of Information, University of Toronto

INF2178: Experimental Design for Data Science

Qiyao Li 1009703670

Yiwen Ma 1009691576

Instructor: Shion Guha

April 16, 2023

Table of Content

1. Introduction

Strip search is a controversial police practice that involves an official performing intimate person search and inspecting their personal effects and body cavities. Thus, this type of search method is often seen as more intrusive than other types of searches, and their usage is normally restricted to circumstances in which there is a reasonable suspicion that the suspect is hiding contraband on their person. Though strip search is legalized in different nations and regions, it is still widely considered traumatic, unnecessary, and ineffective, particularly when performed on individuals who have not been convicted of a crime. Despite the constraints and limitations of using strip search practices, it is still evident that this method may be used disproportionately that affects certain race groups, specifically the black ethnicity.

According to Lemke (Lemke, 2022), black people make up around 10 percent of the City of Toronto's population and nearly one in three of the people who were strip-searched were black. Studies found that strip searches can also leave a long-term impact on the mental health of those who have undergone them, feelings like ashamed, humiliation, and anger often arise after searching. In addition, an article by Phan (2021) underscores the fact that Black individuals are more likely to experience prejudice in arrest due to their affiliation with street culture. They tend to be overrepresented in drug-related and economic cases, mainly because of poverty. Besides, Piquero and Brame also state that except for poverty, inequity and lack of opportunity, Black culture, which has close relationship with violence and drugs also prompts Black crime rate in the neighborhood (Piquero & Brame, 2008).

In the past 20 years, courts and agencies were trying to regulate how police perform strip searches in order to reduce the overall frequency of it since people often feel traumatic after being searched (Lemke, 2022). The supreme court has recognized this issue and stated that it is "likely to represent a disproportionate number of those who are arrested by police and subjected to personal searches, including strip searches" with the absence of statistics (Lemke, 2022).

1.1 Objective

It is quite clear that little research has been conducted on the relationship between the black race, occurrence categories, arrests and strip searches. This report aims to fill the gap in these areas by exploring factors that may be related to the disproportionate impact of strip searches on perceived race occurrence categories, especially on black individuals.

1.2 Methodology

This study will use secondary quantitative data obtained from Toronto Public Services which is considered official police records to explore the relationships between race, occurrence category, strip searches when at arrests. Tests will be used in this study including T-test, Power Analysis and ANCOVA test. A logistic regression model will be used to estimate the probability of being strip searched considering the predictors of perceived race, sex, and occurrence categories to verify the related research questions.

1.3 Research Question

The main research question of this study is that Is there any difference or bias of stripsearch in terms of race when being arrested? In addition to the main research question, the following sub-questions will also be discussed:

- Do Black individuals be treated differently when arrested with strip search practices compared to other races, particularly white and East/South Asian ethics?
- Is there a significant difference in means of the number of strip searches across different occurrence categories, controlling the total arrested number ?
- Are there associations between perceived race, strip search and occurrence category?

2. Literature Review

Discrimination against Black people during the arrest, strip search, and booking process is a significant problem that has been extensively studied by scholars and practitioners. This literature review analyzes three sources to provide key insights into the issue: a legal resource from FindLaw, a research paper from the National Institutes of Health, and a report from the Sentencing Project.

FindLaw's article states that strip searches are typically conducted on people being arrested and taken to jail to prevent the introduction of contraband. However, they can also be a form of humiliation or punishment, particularly for people arrested for minor offenses. The article notes that Black people are disproportionately affected by these types of searches and may be subject to racial profiling by law enforcement officers. Furthermore, according to the Kruse Law, a strip search should be deemed necessary to check for weapons or evidence to ensure a lawful arrest, but there should be reasonable and probable grounds to justify both the search and the arrest (Kruse Law Firm, n.d.).

The National Institutes of Health's research paper explores the impact of race-based data collection in the criminal justice system. The authors argue that Black people are overrepresented in the criminal justice system as victims and defendants due to systemic racism and discriminatory practices, such as racial profiling. The paper also highlights that Black people are more likely to receive bias on assault due to their street culture, have higher proportions in drug-related and robbery/theft cases because of poverty and low income, and more likely to have mental health issues and behaviors when arrested or searched (Phan, 2021). The consequences of discrimination include higher rates of arrest and incarceration, harsher sentences, and trauma from degrading treatment.

The Sentencing Project's report to the United Nations on racial disparities in the criminal justice system provides further evidence of the discrimination faced by Black people. The report highlights that Black people are more likely to be arrested and incarcerated than white people, even when they have committed the same crime, due to the prevalence of racial bias in the criminal justice system. The report also discusses the negative impact of strip searches on Black people as traumatizing, humiliating, and a form of punishment (Interactive, 2022).

Piquero and Brame's *Assessing the Race–Crime and Ethnicity–Crime Relationship in a Sample of Serious Adolescent Delinquents* argues that black people's street culture, which encompasses

attitudes and behaviors related to violence, respect, and masculinity, may influence the likelihood of engaging in criminal behavior. Except for factors, such as poverty, lack of opportunities, and racism, the street culture can promote criminal behavior by emphasizing aggression and a willingness to use violence to maintain respect and status. In addition, the authors also suggest that street culture may play a role in shaping the behavior of some Black youth in inner-city neighborhoods, but that this relationship is complex and influenced by a range of contextual factors (Piquero & Brame, 2008).

In summary, the literature on this topic suggests that Black people are subjected to discrimination and mistreatment during the arrest and booking process, including strip searches. This discrimination stems from poverty, lack of opportunities, racism, and the street culture of Black communities. Therefore, we will first verify if the arrest system is fair and just for all people, regardless of race and then focus on occurrence categories related to Black people's violence street culture and conduct further research to verify if the street culture is associated with criminal behavior and rates.

3. EDA

We first conducted Exploratory data analysis to figure out relationships among exploratory variables and detect any mistakes to help us select appropriate models for the next step.

3.1 Descriptive Statistics

Table 1. *The counts for categories of Sex, Perceived Race, Strip Search, and Occurrence Category.*

Variables	Categoryies	Counts
Sex	M	52502
	F	12576
	U	9
Perceived_Race	White	27635
	Black	17487
	Unknown or Legacy	5044
	East/Southeast Asian	4402
	South Asian	3603
	Middle-Eastern	3227
	Indigenous	1926
	Latino	1759
StripSearch	0	57287
	1	7800
Occurrence_Category	Assault	7724
	Assault & Other crimes against persons	7234
	Robbery & Theft	4580
	Robbery/Theft	3753
	Warrant	4378
	FTA/FTC/Compliance Check/Parollee	4246
	FTA/FTC, Compliance Check & Parollee	3877
	Police Category - Administrative	3885
	Drug Related	2751
	Other Statute	944
	Other Statute & Other Incident Type	2339
	Vehicle Related	768
	Vehicle Related (inc. Impaired)	1977
	Other Offence	1972
	Mischief & Fraud	1732
	Mischief	1321
	Impaired	1362
	Harassment/Threatening	1346
	Harassment & Threatening	1268
	Weapons	1106
	Weapons & Homicide	1105
	Break & Enter	1784
	Sexual Related Crime	839
	Sexual Related Crimes & Crimes Against Children	793
	LLA	586
	Police Category - Incident	562
	Fraud	474
	Mental Health	239
	Homicide	73
	Crimes against Children	69

Figure 1. *Line Graph Of Total Amount of Strip Searches in Arrest Location*

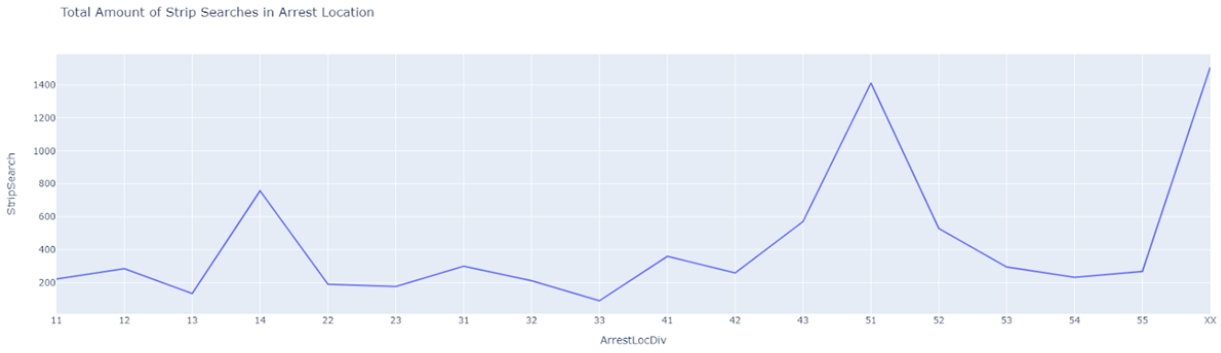


Figure 1 helps us to tell the trend of Strip Search happening in different Arrest Locations. Generally, different divisions have similar amounts of strip searches. The Division XX (the location is not geo-coded or the arrest took place outside of City of Toronto boundaries in other jurisdictions) has the highest strip search of 1500 cases. The Division 51 also has high strip searches of 1400 cases, and the Division 14 has obvious high strip searches of nearly 800 cases. The Division 33 has the lowest high strip search of 100 cases.

Figure 2. *Line Graph Of Total Amount of Strip Searches in Occurrence Category*

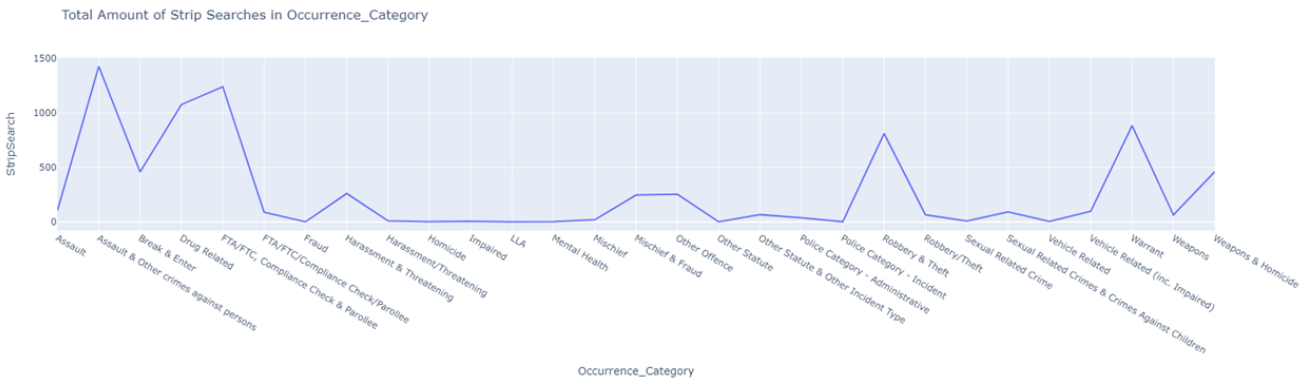


Figure 2 shows the trend of Strip Search happening on various Occurrence Categories. Most occurrence categories have low strip searches. The Assault & Other crimes against persons have highest strip search happenings of nearly 1500. The FTA/ FTC, Compliance Check & Parollee has the second highest strip searches of 1300. The Drug Related also has high strip search happenings of 1000. Then the Robbery & Theft and Warrant has an obvious peak of strip searches of nearly 800. This Line Graph helps us to figure out the high level occurrence category among all categories to conduct further analysis for the following steps.

Figure 3. *Barplot Of Race vs. Strip Search*

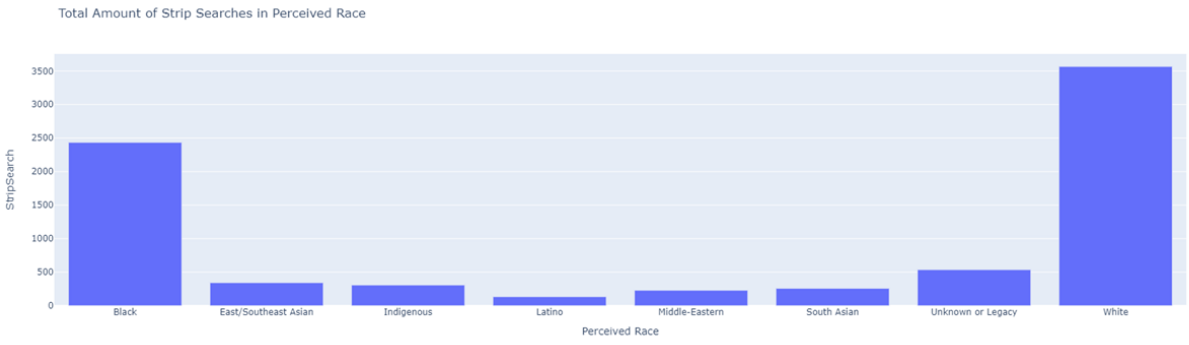
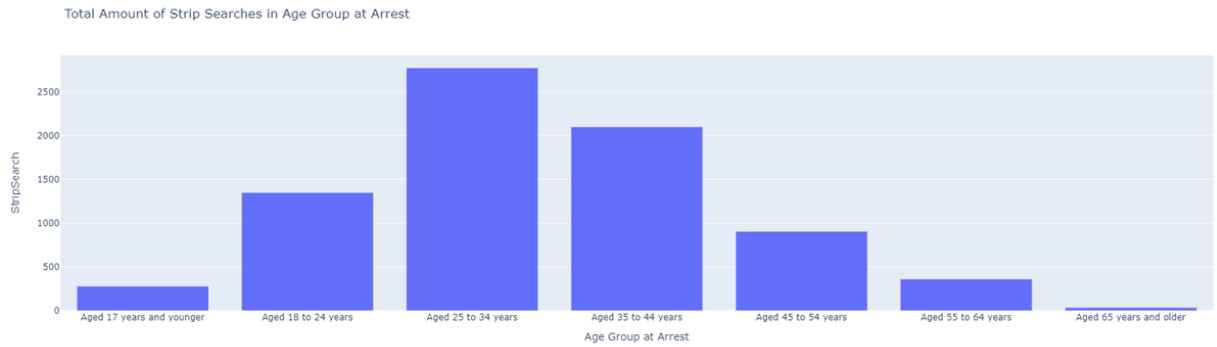


Figure 3 compares the strip search amount between each perceived race group. The White people have the highest strip search amount of 3500. The Black people have the second highest strip search of nearly 2500. In general, race groups except Black and White have average strip searches. We might conduct further analysis focusing on only Black and White groups.

Figure 4. *Barplot Of Age Group at Arrest vs. Strip Search*



We can compare the strip search amounts between each Age Group at Arrest by Figure 4. Aged 25 to 34 years people have the highest strip searches of nearly 2800. Aged 35 to 44 years people have the second highest strip searches of nearly 2100. Aged 65 years and older has the lowest strip searches. Strip searches mainly concentrate in the young and middle-aged people, aged 25 to 44 years.

Figure 5. *Histogram Of Total Amount Of Strip Searches In Perceived Race With Different Age Groups At Arrest*

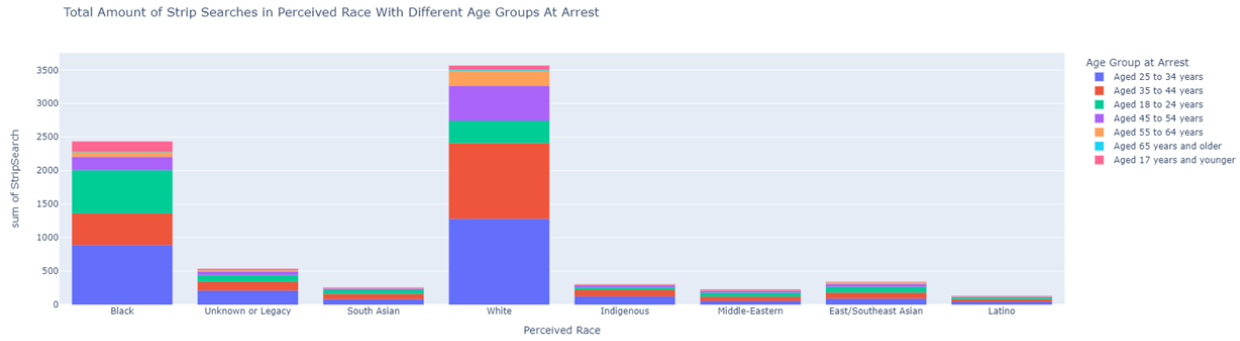
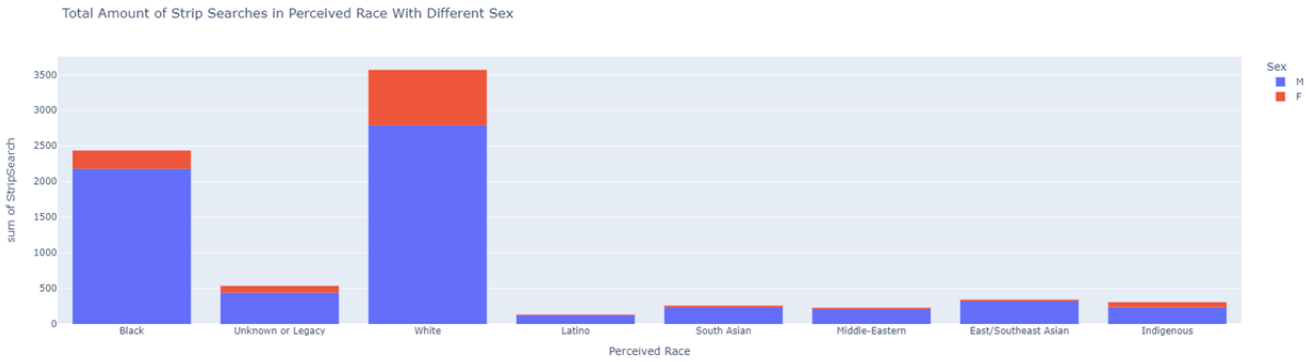


Figure 5 compares the distribution of strip searches of perceived race groups with different Age groups at arrest within each race. Aged 18 to 24 years, Aged 25 to 34 years, and Aged 35 to 44 years occupy the majority of strip searches within each race group. Black people have the highest juvenile strip searches, Aged 17 years and younger. Indigenous have the highest elderly crime, Aged 65 years and older. Although Black people have lower strip searches than White people, Black Aged 18 to 24 years people have higher strip searches than White Aged 18 to 24 years. Black people are committing crimes at a younger age.

Figure 6. *Histogram Of Total Amount Of Strip Searches In Perceived Race With Different Sex*



We can compare the distribution of strip searches among different perceived race groups with sex from Figure 6. The majority of strip searches for each race group are male, but White people have higher female strip searches. Black females have the second highest strip searches.

Figure 7. *Histogram Of Total Amount Of Strip Searches In Perceived Race With Different Occurrence Category*

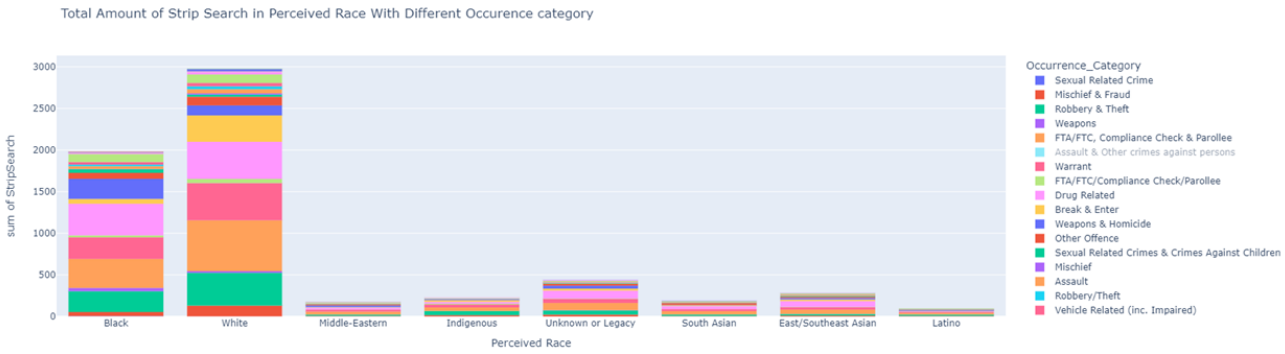


Figure 7. shows the distribution of strip searches among different perceived race groups within different occurrence categories. Drug Related, Warrant, Robbery & Theft, FTA/FTC, Compliance Check & Parollee, and Robbery & Theft occupy the majority of strip searches for almost all race groups. Weapons & Homicide and Drug Related occupy a higher proportion in Black people than other groups.

Figure 8.*Sub-Barplots Of Difference Actions At Arrest*

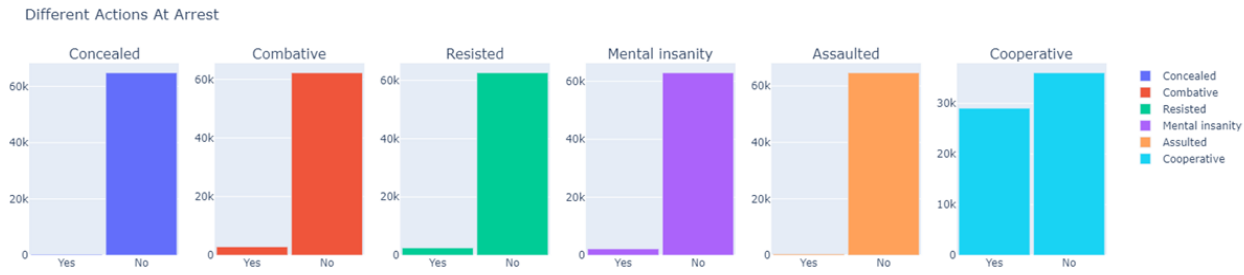
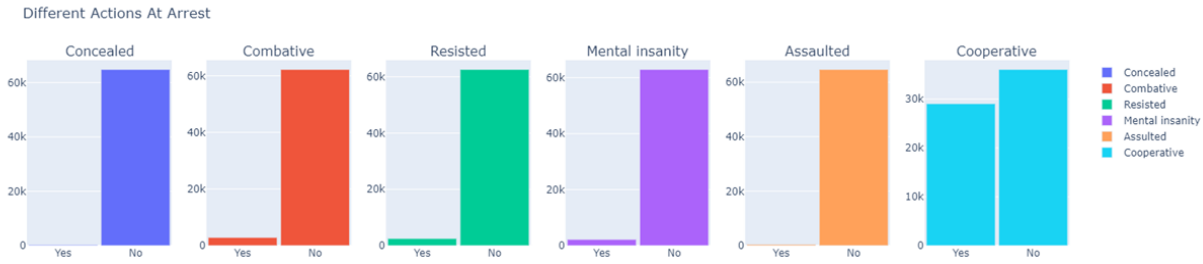


Figure 8. compares all different actions at arrest at the same time. Merely suspects choose to conceal or assault when being arrested. A fraction of suspects have combative, resisted behavior and mental insanity when being arrested. Although nearly half of suspects are cooperative while more than a half of suspects are not.

Figure 9. *Sub-Barplots Of Difference Search Reason*



We can compare all different search reasons at one time from Figure 9 that the highest suspects are searched due to the cause of injury. Lowest suspects are searched because of assisted escape. But in general, it has the average trend of being searched for the four reasons.

Figure 10 *Violin Plots - Number Of Strip Search and Perceived Race With Sex*

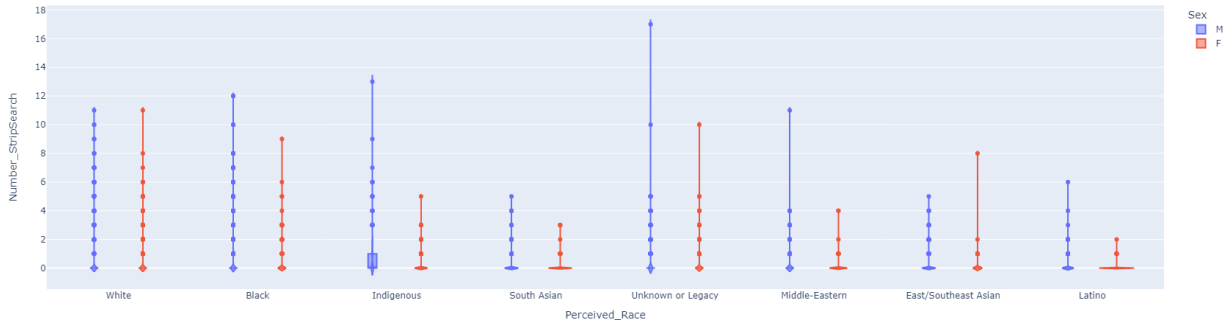


Figure 10. compares the number of strip search distributions of different racial groups in male and female using a violin plot. Most suspects do not have strip searches because almost all the widest parts are located at 0, indicating there is more data in that range. Latino female, Middle Eastern female, and South Asian female have a wider shape, indicating the data are more spread out while the number of strip searches in Black, White, and Indigenous are more concentrated around the median. In terms of outliers, Black, White, and Indigenous male have a higher maximum number of strip searches of 11, 12, and 13 respectively. Indigenous male has the most variability. The middle 50% of the number of Indigenous male's strip searches fall within a range of 2 times, from 0 to 2.

Figure 11 *Box Plots - Number Of Strip Search and Perceived Race With Occurrence Categories*

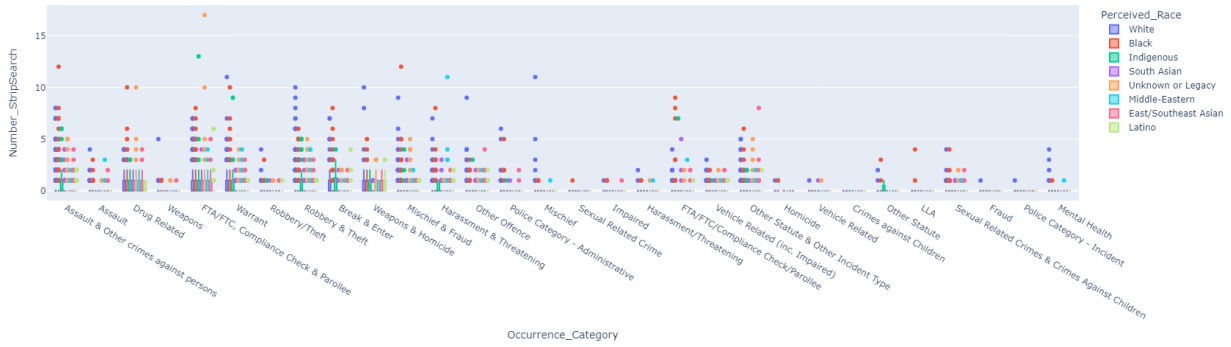


Figure 11. compares the number of strip search distributions of different racial groups in occurrence categories using a box plot. It is noticeable that the occurrence of drugs, FTA/FTC, warrant, weapons and homicide, and break and enter in Black, White, and Indigenous groups have a higher box height, indicating a wider range of values in the middle 50% of the dataset, which can indicate greater variability or dispersion of the data. In addition, white and black groups have extreme outliers on these occurrence categories. It is not conflicted with our knowledge in the literature review.

3.2 Power Analysis

Before conducting t-tests to examine if there is a significant difference between the means of two selected race groups, specifically black individuals compared to other race groups, we calculated the effect size of the explanatory variable using Cohen's D metric.

3.2.1 Perceived Race

Table 2. *Perceived Race Power Analysis Result – Black VS White*

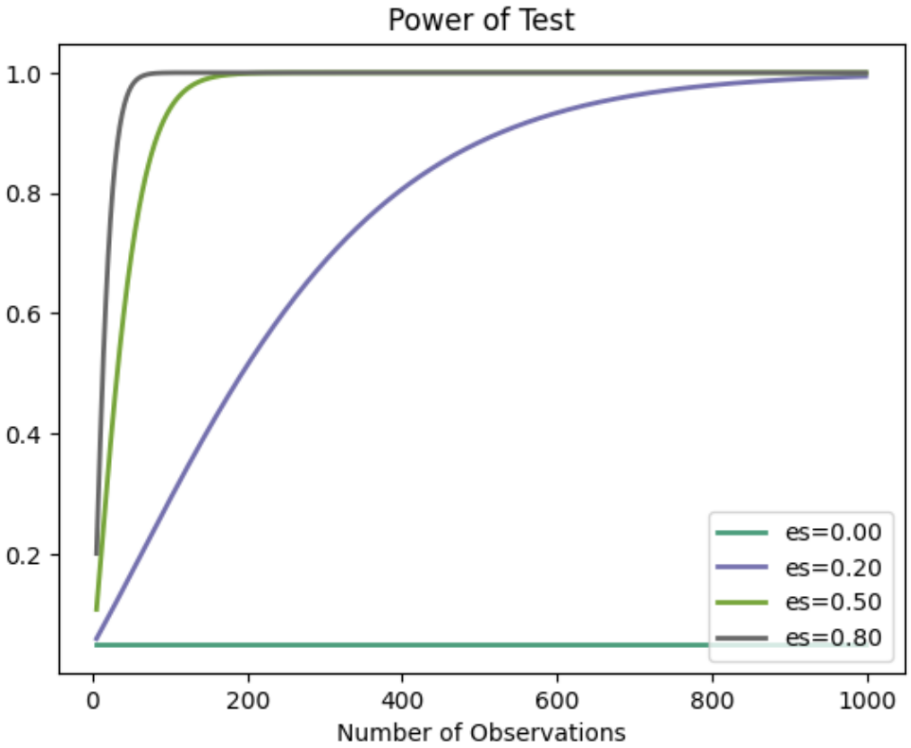
Black VS White	
Effect size (Cohen's D) for Strip Search	0.0012072019455060734
Sample Size of nobs1	9075178.243 needed for BL_ss

Actual size of BL_ss	9909
Sample Size of nobs2	13247800.312 needed for W_ss
Actual size of W_ss	14465

As summarized in the above table, the effect size of Cohen's D metric is 0.0012. According to Cohen's criteria, effect sizes can be categorized into small (0.2), medium (0.5), or large (0.8 and above). This number represents a minimal effect, as it is below Cohen's small effect standard. This suggests that there is a minimal difference between the means of black and white communities when facing strip search practices.

After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. This analysis indicated that a sample size of 13,247,801 was needed for the White group in the Perceived Race variable, while a sample size of 9,075,179 was needed for the Black group in the Perceived Race variable. This finding is significant as the sample sizes available in the dataset are only 14,465 and 9,909 for the respective groups, which can affect the reliability of the test results.

Figure 12. *Power Analysis Graph – Black VS White*



The above figure represents a comparison of effect size according to Cohen's criteria, the dark green line in the above graph shows the actual effect size of this test which is 0.0012. This line indicates minimal even hard-to-notice differences between the two chosen groups.

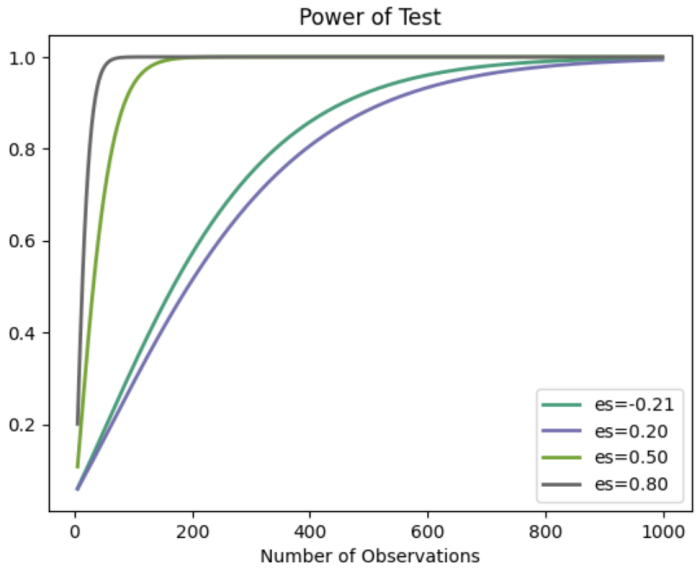
Table 3. *Perceived Race Power Analysis Result – Black VS East/South Asian*

Black VS East/Southeast Asian	
Effect size (Cohen's D) for Strip Search	-0.2147784226845171
Sample Size of nobs1	716.360 needed for BL_ss
Actual size of BL_ss	9909
Sample Size of nobs2	223.750 needed for ES_ss
Actual size of ES_ss	3095

As summarized in the above table, the effect size of Cohen's D metric is -0.2148. According to Cohen's criteria, this test result of effect size falls into the small effect region which is just a little above 0.2 standard. This number represents a small effect and suggests that there is a small difference between the means of Black and East/Southeast Asian communities when facing strip search practices.

After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. This analysis indicated that a sample size of 224 was needed for the East/South Asian group in the Perceived Race variable, while a sample size of 717 was needed for the Black group in the Perceived Race variable. This finding is significant as the sample sizes available in the dataset are 3,095 and 9,909 for the respective groups, which could lead to a more precise estimate of the test result.

Figure 13. *Power Analysis Graph – Black VS East/South Asian*



The above figure represents a comparison of effect size according to Cohen's criteria, the dark green line in the above graph shows the actual effect size of this test which is –0.2148 is just a little bit above small effect size line.

3.2.2 Sex

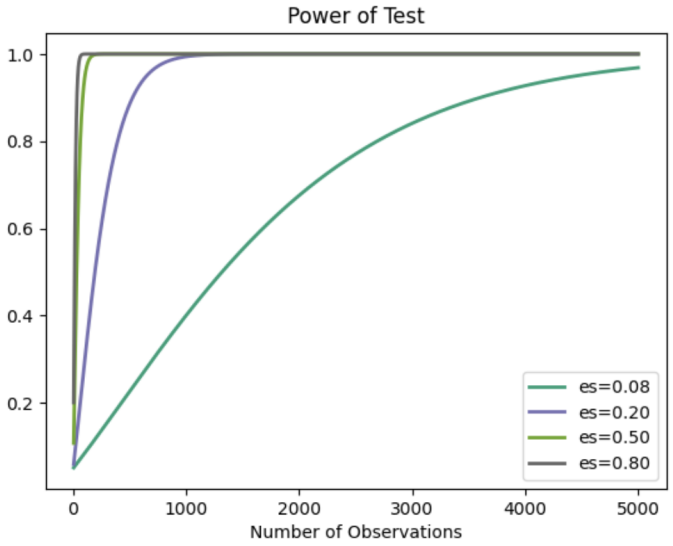
Table 4. *Sex Power Analysis Result – Female VS Male*

Female VS Male	
Effect size (Cohen's D) for Strip Search	0.07640487121055602
Sample Size of nobs1	1692.384 needed for F_ss
Actual size of F_ss	7646
Sample Size of nobs2	6548.415 needed for M_ss
Actual size of M_ss	29585

As summarized in the above table, the effect size of Cohen's D metric is 0.0764. According to Cohen's criteria, this test result of effect size is below Cohen's small effect standard. This suggests that there is a minimal difference between the means of female and male groups when facing strip search practices.

After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. This analysis indicated that a sample size of 6,549 was needed for the male group in the Perceived Race variable, while a sample size of 1,963 was needed for the female group in the Perceived Race variable. This finding is significant as the actual sample sizes available in the dataset are 29,585 and 7,646 for the respective groups, which could lead to a more precise estimate of the test result.

Figure 14. *Power Analysis Graph – Female VS Male*



The above figure represents a comparison of effect size according to Cohen's criteria, the dark green line in the above graph shows the actual effect size of this test which is 0.0764. The actual effect size line falls under the small effect range according to Cohen's standard, which indicates minimal even hard-to-notice differences between the two chosen groups.

Overall, there are minimal or hard-to-notice mean differences found in the Black and White group comparison and Female and male group comparisons, and only a small effect size was founded between the Black and East/South Asian group when facing strip search practices. Still, it is worth doing these tests before conducting t-tests and other tests. It is particularly useful when dealing with these large sample sizes, where even small or hard-to-notice differences may be statistically significant.

3.3 T-tests

The t-test enables us to further compare the means of strip searches between different groups.

3.3.1 T-tests - Number of Strip Search & Perceived Race (Black and White)

Since our research question is that Is there any difference or bias of stripsearch in terms of race when being arrested, specifically do Black individuals be treated differently when arrested with strip search practices compared to other races, particularly white ethics and is there a significant difference in means of the number of strip searches across different occurrence categories, we will first apply the t-test to compare the means of number of strip searches of black and white groups firstly.

- Null Hypothesis: the number of strip searches mean of black people has no difference from the number of strip searches mean of white people.

- Alternative Hypothesis: the number of strip searches mean of black people is different from the number of strip searches mean of white people.

Table 5. *t-test results – Number of Strip Search at arrest & Perceived Race (Black and White)*

		t	p-value	df	Mean	Std	95% Confidence Interval of the Difference	
Number of StripSearch	Perceived_Race Black	2.730783061500993 4	0.006321576386601609	34902	0.7214153988012909	1.6430558162449205	Lower	Upper
	Perceived_Race White				0.7637763666885629	1.5266440674694435	0.011956190630020368	0.07276574514452352

Checking the two tailed t-distribution table, when df is 34902, p is 0.05, the t is 1.96. The absolute value of the t value given by the t test above is 2.73, which is greater than the critical value found in the table, so we can reject the null hypothesis of the t-test and conclude that the results of the test are statistically significant. Besides, our p value of 0.006 is much smaller than 0.05, so we can reject the null hypothesis of no difference and say with a high degree of confidence that there is a true difference in group means. Also, we can see that the difference in means for our sample data is 0.04 (0.76 – 0.72), and the confidence interval shows that the true

difference in means is between 0.012 and 0.073. So, 95% of the time, the true difference in means will be different from 0.

In conclusion, the mean score of black people's number of strip searches is different from the mean score of white people's number of strip searches. We will conduct further research among perceived race groups.

3.3.2 T-tests - Number of Strip Search & Sex of Black

Next, we want to compare the means of number of strip searches of black male and black females continuously with t-test.

- Null Hypothesis: the number of strip searches mean of black male has no difference from the number of strip searches mean of females.

- Alternative Hypothesis: the number of strip searches mean of black male is different from the number of strip searches mean of black females.

Table 6. *t-test results – Number of Strip Search at arrest & Sex of Black*

		t	p-value	df	Mean	Std	95% Confidence Interval of the Difference	
Number of StripSearch	Black Male	10.819868175802178	5.088763342704695e-27	5657	0.4817786693413574	1.2377502702881396	Lower	Upper
	Black Female				0.7713251166353318	1.7112814205996583	0.23708536116432827	0.3420075334236205

Checking the two tailed t-distribution table, when df is 5657, p is 0.05, the t is 1.96. The t value given by the t test above is 10.82, which is greater than the critical value found in the table, so we can reject the null hypothesis of the t-test and conclude that the results of the test are statistically significant. Besides, our p value of 5.088763342704695e-27 is much smaller than 0.05, so we can reject the null hypothesis of no difference and say with a high degree of confidence that there is a true difference in group means. Also, we can see that the difference in means for our sample data is 0.29 (0.77 – 0.48), and the confidence interval shows that the true difference in means is between 0.237 and 0.342. So, 95% of the time, the true difference in means will be different from 0.

In conclusion, the mean score of black male's number of strip searches is different from the mean score of black females's number of strip searches.

3.3.3 T-tests - Number of Strip Search & Weapons & Homicide Occurrence in Black and White

From Figure 7. *Histogram Of Total Amount Of Strip Searches In Perceived Race With Different Occurrence Category,* we observed that Weapons & Homicide and Drug Related occupy a higher proportion in Black people than other groups. Then we first conduct a t-test to compare the means of number of strip searches of Weapons & Homicide Occurrence in Black and White.

- Null Hypothesis: the mean of number of strip search for Weapons & Homicide occurrence of black has no difference from the mean of number if strip search for Weapons & Homicide occurrence of white.

- Alternative Hypothesis: the mean of number of strip searches for Weapons & Homicide occurrence of black is different from the mean of number of strip searches for Weapons & Homicide occurrence of white.

Table 7. *t-test results – Number of Strip Search at arrest & Weapons & Homicide Occurrence in Black and White*

		t	p-value	df	Mean	Std	95% Confidence Interval of the Difference	
Number of StripSearch	Weapons & Homicide Occurrence in Black	0.933784146829081	0.350841704750572 1	528	0.800813008130081 300813	1.1755136458423632	Lower	Upper
	Weapons & Homicide Occurrence in Black				0.898412698412698 4	1.5989140486186502	-0.3029265820312859	0.10772720146605158

Checking the two tailed t-distribution table, when df is 528, p is 0.05, the t is 1.984. The t value given by the t test above is 0.934, which is smaller than the critical value found in the table, so we cannot reject the null hypothesis of the t-test and conclude that the results of the test are not statistically significant. Besides, our p value of 0.351 is greater than 0.05, so we cannot reject the null hypothesis of no difference and say that there is not a true difference in group means.

In conclusion, the mean score of Weapons & Homicide Occurrence of black's number of strip searches is not different from the mean score of Weapons & Homicide Occurrence of white's number of strip searches, which is conflicted with our observation from EDA graph and understanding from our literature review. Therefore, we will further explore if there is any difference among other race groups.

3.3.4 T-tests - Number of Strip Search & Drug Related Occurrence in Black and White

Then we conduct another t-test to compare the means of number of strip searches of Drug Related Occurrence in Black and White.

- Null Hypothesis: the mean of number of strip searches for Drug Related occurrences of black has no difference from the mean of number of strip searches for Drug Related occurrence of white.

- Alternative Hypothesis: the mean of number of strip searches for Drug Related occurrence of black is different from the mean of number of strip searches for Drug Related occurrence of white.

Table 8. - *t-test results – Number of Strip Search at arrest & Drug Related Occurrence in Black and White*

		t	p-value	df	Mean	Std	95% Confidence Interval of the Difference	
Number of StripSearch	Drug Related Occurrence in Black	1.484055191095985 6	0.13795562535731254	1953	0.8991507430997877	1.4407974385275928	Lower	Upper
	Drug Related Occurrence in White				0.9925650557620818	1.3755136107789807	-0.21686134720434724,	0.0300327218797591 53

Checking the two tailed t-distribution table, when df is 1953, p is 0.05, the t is 1.96. The t value given by the t test above is 1.48, which is smaller than the critical value found in the table, so we cannot reject the null hypothesis of the t-test and conclude that the results of the test are not statistically significant. Besides, our p value of 0.13 is greater than 0.05, so we cannot reject the null hypothesis of no difference and say with a high degree of confidence that there is no difference in group means.

In conclusion, the mean score of Drug Related Occurrence of black's strip searches has no difference from the mean score of Drug Related Occurrence of white's strip searches. This is conflicted with our observation from the previous EDA graph and our literature review. We need further research and analysis on Drug Related occurrences with perceived race groups.

3.3.5 T-tests - Number of Strip Search & Break and Enter Occurrence in Black and White

From the literature review, we also notice that Break and Enter also occupies a higher proportion in the occurrence category of strip search because of the street culture of the Black group. Then we first conduct a t-test to compare the means of strip searches of Break and Enter Occurrence in Black and White.

- Null hypothesis: the mean of number of strip search for Break and Enter occurrence of black and the mean of number of strip search for Break and Enter occurrence of white are not different

- ALternative hypothesis: the mean of number of strip search for Break and Enter occurrence of black and the mean of number of strip search for Break and Enter occurrence of white are different

Table 9. *t-test results – Number of Strip Search at arrest & Break and Enter Occurrence in Black and White*

		t	p-value	df	Mean	Std	95% Confidence Interval of the Difference	
Number of StripSearch	Break and Enter Occurrence in Black	2.869848874564793	0.004303850577108006	441	1.2535714285714286	2.046937067153468	Lower	Upper
	Break and Enter Occurrence in White				1.6492048643592143	2.0783915525759364	0.12469293379516838,	0.6665739377804032

Checking the two tailed t-distribution table, when df is 441, p is 0.004 the t is 1.98. The t value given by the t test above is 2.86, which is greater than the critical value found in the table, so we can reject the null hypothesis of the t-test and conclude that the results of the test are statistically significant. Besides, our p value of 0.0043 is smaller than 0.05, so we can reject the null hypothesis of no difference and say with a high degree of confidence that there is a true difference in group means. Also, we can see that the difference in means for our sample data is 0.39 (1.64 – 1.25), and the confidence interval shows that the true difference in means is between 0.13 and 0.67. So, 95% of the time, the true difference in means will be different from 0.

In conclusion, the mean score of Break and Enter Occurrence of black's number of strip searches is different from the mean score of Break and Enter Occurrence of white's number of strip searches, which is not conflicted with our understanding from literature review and the observation from EDA graph.

4. Method

4.1 Dataset Description

The dataset contains information about 65276 entries and 25 variables related to all arrests and strip searches. Based on the research fact that black people suffer from discrimination during the arrest, strip searches, and booking process, this paper tends to explain the booked and strip searches by the following factors. These are Age, Sex, Perceived Race (we only extract Black and White), Occurrence Category, Actions at arrest - Concealed items, Actions at arrest -

Combative, violent or spitter/biter, Actions at arrest - Resisted, defensive or escape risk, Actions at arrest - Mental instability or possibly suicidal, Actions at arrest - Assaulted officer, Actions at arrest – Cooperative, SearchReason-CauseInjury,SearchReason-AssistEscape, SearchReason-PossessWeapons, and SearchReason-PossessEvidence (2022). For this study, we analyzed Arrests and Strip Searches (RBDC-ARR-TBL-001) updated in 2022 and considered the demographic characteristics and arrested behaviors to illustrate the relationship between the strip search and book taking place and the above 14 factors.

- Outcome Variables - DV

The outcome variable of the dataset is "StripSearch". 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 a visual examination of their body. It has 0 (not receive strip search) and 1 (receive strip search). In probability theory, 0 and 1 are used to represent the probability of an event occurring, so we assume this dummy variable to be equivalent to a continuous variable. Therefore, we chose to use strip search as our outcome variable.

The "Number_StripSearch" variable has been created according to the "StipSearch" columns which represent the total amount of people who have been strip-searched while being arrested. Repeated values have been removed based on the unique Person ID column.

- Exploratory Variables - IV

Perceived_Race: perceived race is the race group suspects are treated by police when being arrested. It has eight categorical groups - White, Black, Unknown or Legacy, East/Southeast Asian, South Asian, Middle-Eastern, Indigenous, and Latino.

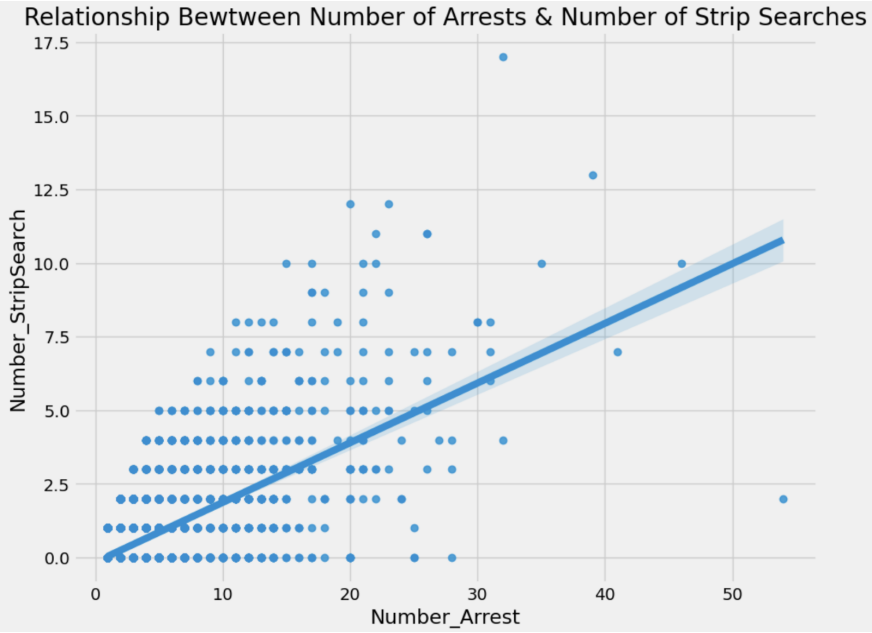
Occurence_Category: Occurrence categories are taxonomies or classifications of accidents and incidents happening at a high level to permit analysis of the data in support of safety (2011). It has 31 categorical levels, including Assault, Assault & Other crimes against persons, Robbery & Theft, Robbery/Theft, Warrant, FTA/FTC/Compliance Check/Parollee, FTA/FTC, Compliance Check & Parollee, Police Category - Administrative, Drug Related, Other Statute, Other Statute & Other Incident Type, Vehicle Related, Vehicle Related (inc. Impaired), Other Offence, Mischief & Fraud, Mischief, Impaired, Harassment/Threatening, Harassment & Threatening, Weapons, Weapons & Homicide, Break & Enter, Break and Enter, Sexual Related Crime, Sexual Related Crimes & Crimes Against Children, LLA, Police Category - Incident, Fraud, Mental Health, Homicide, and Crimes against Children, and NaN. We will clean the dataset to combine Break & Enter and Break and Enter together and remove NaN.

Sex: sex is the physical sexuality suspects are treated by police when being arrested. It has three categorical levels - Male, Female, and Unknown. We drop the "Unknown" during the data cleaning process.

5. ANCOVA Test

ANCOVA is used here to determine the differences between group means while controlling for the effects of covariates, the Number_Arrest. Prior to doing ANCOVA tests, a plot was used to determine whether there is a straightforward mathematical relationship (typically linear) between the control variable and the outcome. The below figure shows a linear relationship between variables Number_StripSearch and Number_Arrest.

Figure 15. *Power Analysis Graph – Female VS Male*



5.1 ANCOVA Test - Perceived Race

Research Question: Do strip search practices vary across different racial groups?

- Null Hypothesis: There is no significant difference in the mean number of strip searches between the different perceived race groups when controlling for the number of arrests.

- Alternative Hypothesis: There is a significant difference in the mean number of strip searches between the different perceived race groups when controlling for the number of arrests.

Table 10. *ANCOVA results – Number_StripSearch & Perceived Race*

Source	SS	DF	F	p-unc	np2
Perceived_Race	21.514958	7	12.278243	8.762727e-16	0.002304
Number_Arrest	6044.292151	1	24145.666378	0.000000e+00	0.393459
Residual	9317.640644	37222	NaN	NaN	NaN

The ANCOVA test with a p-unc (uncorrected p-value) value of 8.762727e-16 (＜0.05) for Perceived Race suggests that there is a significant difference in mean values of strip search practices among different perceived race groups while controlling for the total number of arrests. Therefore, we can reject the null hypothesis. Our initial hypothesis was that perceived race could predict whether someone would be strip searched during an arrest. From our results, there is indeed a statistically significant relationship between perceived race and strip search practices. These findings provide valuable insights that will be further elaborated on in the results and findings section.

5.2 ANCOVA Test - Occurrence Category

Research Question: Do strip search practices vary across different racial groups?

- Null Hypothesis: There is no significant difference in the mean number of strip searches between the different occurrence categories when controlling for the number of arrests.

- Alternative Hypothesis: There is a significant difference in the mean number of strip searches between the different occurrence categories when controlling for the number of arrests.

Table 11. *ANCOVA results – Number_StripSearch & Occurrence Category*

Source	SS	DF	F	p-unc	np2
Occurrence_Category	588.171356	29	86.216802	＜0.001	0.062979
Number_Arrest	5453.742413	1	23183.588502	＜0.001	0.383939
Residual	8750.984246	37200	NaN	NaN	NaN

After conducting an ANCOVA test with a p-unc (uncorrected p-value) value of ＜0.001 (＜0.05) for Occurrence Category and controlling for the total number of arrests, the null hypothesis that the mean value between strip search practices for different occurrence categories is equal, can be rejected. Our initial hypothesis was that different occurrence categories would predict whether an individual would be strip searched during an arrest. Our results indicate that there is a statistically significant relationship between occurrence category and strip search. It is worth noting that the Python test result showed an actual p-unc value of 0.0, which typically implies that the p-value is extremely small but not necessarily zero. Therefore, it is important to interpret a p-value of 0.0 as a strong indication of statistical significance. The recorded p-unc value in this study is ＜0.001, and we will provide a detailed explanation of our findings in the results and findings section.

6. Chi-Square

The chi-square test is used to determine whether there is an association between two categorical variables and if the occurrence of one variable is related to the occurrence of the other variable.

We adopted the chi-square test to exam our sub research question: Are there associations between perceived race, strip search and occurrence category?

6.1 Perceived Race VS Strip Search

Research Question: Is there an association between perceived race and strip search?

- Null hypothesis: Perceived Race and Strip Search are independent.
- Alternative hypothesis: Perceived Race and Strip Search are dependent.

Table 12. *Chi-Square results – Perceived Race & StripSearch*

Perceived Race VS Strip Search	
Chi-squared statistic:	356.12485254652904
P-value	6.017709683374717e-73

The p-value of the test is 6.017709683374717e-73, which indicates significant evidence against the null hypothesis. In other words, it suggests that Perceived Race and Strip Search are not independent variables, but rather have a significant association on statistical level. A p-value of less than 0.05 is often considered statistically significant, therefore this p-value is significantly below the significance threshold. As a result, we would reject the null hypothesis and accept the alternative hypothesis that Perceived Race and Strip Search are dependent.

6.2 Perceived Race VS Occurrence Category

Research Question: Is there an association between perceived race and occurrence category?

- Null hypothesis: Perceived Race and Occurrence Category are independent.
- Alternative hypothesis: Perceived Race and Occurrence Category are dependent.

Table 13. *Chi-Square results – Perceived Race & Occurrence Category*

Perceived Race VS Occurrence Category	
Chi-squared statistic:	2751.6587949730747
P-value	<0.001

The p-value of the test in Python is 0.0, indicating that the p-value is less than the lowest value that the computer can represent, which is generally approximately 2.2e-16. As a result, the p-value is reported as <0.001. The p-value of <0.001 implies that there is substantial evidence against the null hypothesis and that the alternative hypothesis should be adopted. In other words, the p-value indicates a very high relationship between Perceived Race and Occurrence Category, implying that these two variables are dependent.

6.3 Strip Search VS Occurrence Category

Research Question: Is there an association between strip search and occurrence category?

- Null hypothesis: Strip Search and Occurrence Category are independent.
- Alternative hypothesis: Strip Search and Occurrence Category are dependent.

Table 14. *Chi-Square results – Strip Search & Occurrence Category*

Strip Search & Occurrence Category	
Chi-squared statistic:	8547.072981535563
P-value	<0.001

The p-value of the test shows in python is 0.0, which indicates that the p-value is smaller than the smallest value that can be represented by the computer, which is usually around 2.2e-16. Therefore, the p-value is reported as <0.001. The p-value is <0.001 indicates that there is significant evidence against the null hypothesis and alternative hypothesis should be accepted. In other words, the p-value suggests there is a very strong association between Strip Search and Occurrence Category, and we can conclude that these two variables are dependent.

In conclusion, the chi-square test findings suggest that the incidence of the selected exploratory variables, namely perceived race, strip search, and occurrence category, has a statistically significant association. These findings support our study objectives by providing strong evidence that there are substantial connections between these factors. Overall, research implies that perceptions of race and occurrence category have a substantial impact in deciding whether people are subjected to strip searches during police enforcement practices.

7. Logistic Regression

Based on the Chi-square results, we found that the occurrence category is associated with perceived race. Therefore, we perform a logistic regression to examine the effects of Sex, Perceived_Race_Black, Perceived_Race_East_Southeast_Asian, Perceived_Race_Indigenous, Perceived_Race_Latino, Perceived_Race_Middle_Eastern,Perceived_Race_South_Asian, Perceived_Race_White,Occurrence_Category_Drug_Related, Occurrence_Category_Weapons_Homicide, Occurrence_Category_Break_Enter. The logistic regression results suggest that several features are significantly associated with the likelihood of suspects being strip searched. Most of the features are statistically significant except for Perceived_Race_Latino.

Figure 16. *Logit Regression Results*

Logistic Regression Results						
	coef	std err	z	P>|z|	[0.025, 0.975]	
Intercept	-2.4821	0.061	-40.796	0	-2.601	-2.363
Sex	-0.2031	0.041	-4.967	0	-0.283	-0.123
Perceived_Race_Black	0.3486	0.066	5.276	0	0.219	0.478
Perceived_Race_East_Southeast_Asian	-0.3106	0.093	-3.324	0.001	-0.494	-0.127
Perceived_Race_Indigenous	0.6777	0.098	6.906	0	0.485	0.87
Perceived_Race_Latino	-0.1328	0.124	-1.072	0.284	-0.376	0.11
Perceived_Race_Middle_Eastern	-0.3421	0.106	-3.238	0.001	-0.549	-0.135
Perceived_Race_South_Asian	-0.2703	0.099	-2.72	0.007	-0.465	-0.076
Perceived_Race_White	0.3271	0.064	5.094	0	0.201	0.453
Occurrence_Category_Drug_Related	1.8488	0.051	36.463	0	1.749	1.948
Occurrence_Category_Weapons_Homicide	1.9447	0.076	25.494	0	1.795	2.094
Occurrence_Category_Break_Enter	1.0791	0.071	15.211	0	0.94	1.218

Interpretation of the above coefficients:

Compared with female suspects, the predicted log of the odds of male suspects being strip searched are 0.2031 higher, controlling for other features ($\beta = 0.2031$, $p < 0.05$). Sex is statistically significant to strip search. This finding is consistent with our t-test results, suggesting that sex is associated with the likelihood of suspects being strip searched.

Compared with non-black suspects, the predicted log of the odds of suspects perceived as black being strip searched are 0.3486 higher, controlling for other features ($\beta = 0.3486$, $p < 0.001$). The perceived race black is statistically significant to strip search. This finding is consistent with our ANOVA results, suggesting that whether the perceived race is black is associated with the likelihood of suspects being strip searched.

Compared with non-east southeast asian suspects, the predicted log of the odds of southeast asian suspects being strip searched are 0.3106 lower, controlling for other features. The perceived race southeast asian is statistically significant to strip search, which is not conflicted with our ANOVA results. Whether the perceived race is east southeast asian is associated with likelihood suspects being strip searched.

Compared with non-indigenous suspects, the predicted log of the odds of indigenous suspects being strip searched are 0.6777 higher, controlling for other features. The perceived race indigenous is statistically significant to strip search, which is not conflicted with our ANOVA results. Whether the perceived race is east indigenous is associated with likelihood suspects being strip searched.

Compared with non-latino suspects, the predicted log of the odds of latino suspects being strip searched are 0.1328 lower, controlling for other features. The perceived race latino is statistically significant to strip search, which is not conflicted with our ANOVA results. Whether the perceived race is latino is not associated with likelihood suspects being strip searched.

Compared with non-middle eastern suspects, the predicted log of the odds of middle eastern suspects being strip searched are 0.3421 lower, controlling for other features. The perceived race middle eastern is statistically significant to strip search, which is not conflicted with our ANOVA

results. Whether the perceived race is middle eastern is associated with likelihood suspects being strip searched.

Compared with non-south asian suspects, the predicted log of the odds of south asian suspects being strip searched are 0.2703 lower, controlling for other features. The perceived race south asian is statistically significant to strip search, which is not conflicted with our ANOVA results. Whether the perceived race is south asian is associated with likelihood suspects being strip searched.

Compared with non-white suspects, the predicted log of the odds of white suspects being strip searched are 0.3271 lower, controlling for other features. The perceived race white is statistically significant to strip search, which is not conflicted with our t-test results. Whether the perceived race is white is associated with likelihood suspects being strip searched.

Compared with suspects who do not have drug related occurrence, the predicted log of the odds of suspects who are strip searched with drug related occurrence are 1.8488 higher, controlling for other features. The occurrence category drug related is statistically significant to strip search, which is not conflicted with our t-test results. Whether the occurrence category is drug related is associated with likelihood suspects being strip searched.

Compared with suspects who do not have weapons and homicide occurrence, the predicted log of the odds of suspects who are strip searched with weapons and homicide occurrence are 1.9447 higher, controlling for other features. The occurrence category weapons and homicide is statistically significant to strip search, which is not conflicted with our t-test results. Whether the occurrence category is weapons and homicide is associated with likelihood suspects being strip searched.

Compared with suspects who do not have break and enter occurrence, the predicted log of the odds of suspects who are strip searched with break and enter occurrence are 1.0791 higher, controlling for other features. The occurrence category break and enter is statistically significant to strip search, which is not conflicted with our t-test results. Whether the occurrence category is break and enter is associated with likelihood suspects being strip searched.

Figure 17. *Odds Ratio of Independent Variables*

Odds Ratio of Logit Regression Model	
	Odds ratio
Intercept	0.083568
Sex	0.816174
Perceived_Race_Black	1.417045
Perceived_Race_East_Southeast_Asian	0.732984
Perceived_Race_Indigenous	1.96935
Perceived_Race_Latino	0.875636
Perceived_Race_Middle_Eastern	0.710265
Perceived_Race_South_Asian	0.763185
Perceived_Race_White	1.386888
Occurrence_Category_Drug_Related	6.352061
Occurrence_Category_Weapons_Homicide	6.991492
Occurrence_Category_Break_Enter	2.942077

Interpretation of Odds Ratio:

The odds ratio (OR) of sex is smaller than 1, the probability of female suspects being strip searched is smaller than the probability of male suspects being strip searched.

In this logistic regression model, with 95% confidence interval, the confidence interval for Sex is [0.753, 0.884], it means that there is a 95% probability that the true coefficient for the sex lies between 0.753 and 0.884. Since sex is a dummy variable, it indicates that whether the suspect's sex is male or female is associated with a change in the log-odds of the being strip searched outcome between 0.2 and 0.3, holding all other predictors constant.

The odds ratio (OR) of Perceived Race Black is larger than 1, the probability of black suspects being strip searched is larger than the probability of non-black suspects being strip searched.

The odds ratio (OR) of Perceived Race East Southeast Asian is smaller than 1, the probability of East Southeast Asian suspects being strip searched is smaller than the probability of non-East Southeast Asian suspects being strip searched.

The odds ratio (OR) of Perceived Race Indigenous is larger than 1, the probability of Indigenous suspects being strip searched is larger than the probability of non-Indigenous suspects being strip searched.

The odds ratio (OR) of Perceived Race Latino is smaller than 1, the probability of Latino suspects being strip searched is smaller than the probability of non-Latino suspects being strip searched.

The odds ratio (OR) of Perceived Race Middle Eastern is smaller than 1, the probability of Middle Eastern suspects being strip searched is smaller than the probability of non-Middle Eastern suspects being strip searched.

The odds ratio (OR) of Perceived Race South Asian is smaller than 1, the probability of South Asian suspects being strip searched is smaller than the probability of non-South Asian suspects being strip searched.

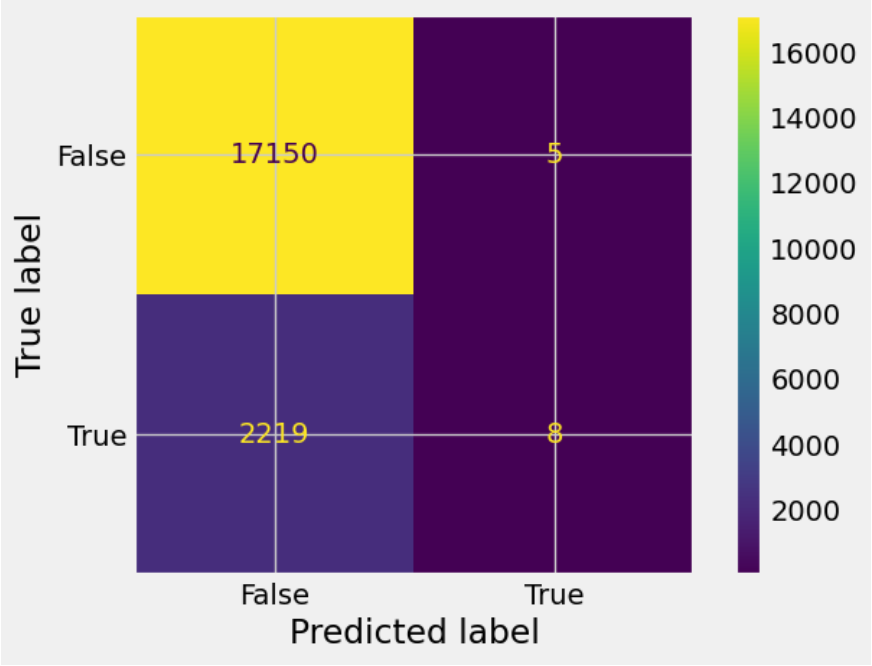
The odds ratio (OR) of Perceived Race White is larger than 1, the probability of white suspects being strip searched is larger than the probability of non-white suspects being strip searched.

The odds ratio (OR) of Occurrence Category Drug Related is larger than 1, the probability of suspects being strip searched with drug related occurrence is larger than the probability of suspects being strip searched without drug related occurrence.

The odds ratio (OR) of Occurrence Category Weapons and homicide is larger than 1, the probability of suspects being strip searched with weapons and homicide occurrence is larger than the probability of suspects being strip searched without weapons and homicide occurrence.

The odds ratio (OR) of Occurrence Category Break and Enter is larger than 1, the probability of suspects being strip searched with Break and Enter occurrence is larger than the probability of suspects being strip searched without Break and Enter occurrence.

Figure 18. *Confusion Matrix*



Interpret Confusion Matrix

Top-Left Quadrant shows the outcome of True Negative (TN), displaying 17150 instances were incorrectly classified as negative.It means 17150 instances who were strip searched were incorrectly classified as not being strip searched.

Top-Right Quadrant shows the outcome of False Positive (FP), displaying 5 instances were correctly classified as negative.It means 5 instances who were not strip searched were incorrectly classified as being stripsearched.

Bottom-Left Quadrant shows the outcome of False Negative (FN), displaying 2219 instances were incorrectly classified as positive. It means 2219 instances who were not strip searched were correctly classified as not being strip searched.

Bottom-Right Quadrant shows the outcome of True Positive (TP), displaying 8 instances were correctly classified as positive. It means 8 instances who were strip searched were correctly classified as being strip searched.

Using these values, we can calculate the performance metrics:

Accuracy: (17150 + 8) / (5 + 8 + 2219 + 17150) = 0.88

Precision: 8 / (8 + 5) = 0.62

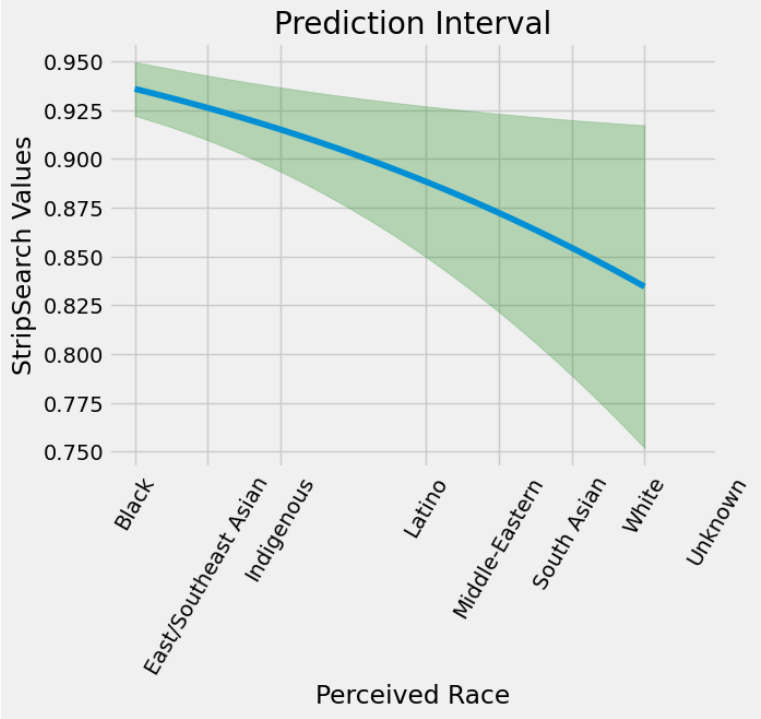
Recall (Sensitivity or True Positive Rate): 8 / (8 + 2219) = 0.0035

F1 Score: 2 * (0.62 * 0.0035) / (0.62 + 0.0035) = 0.0069

From this example, we can see that the model has an accuracy of 88%, which means that 88% of the instances in the test set were correctly classified by the model. The precision of 0.62 indicates that the model is not making many false positive predictions, while the recall of 0.0035 indicates that the model is not capturing most of the positive instances. The F1 Score of 0.0035 is the harmonic mean of precision and recall and provides a balance between the two metrics.

Initially, we decompose the Perceived_Race into groups of dummy variables to determine the relationship between the respective race group as independent variables and the binary dependent variable StripSearch. However, we also would like to compare among race groups. Therefore, to get the prediction Interval, we conduct another logit regression model making the race independent variable as a categorical variable with 8 levels.

Figure 19. *Prediction Interval*



The prediction interval plot helps us to view both the expected value of our predictions - StripSearch(the blue line) and the uncertainty associated with those predictions (the green shaded area).the model predicted values - Strip Search values with their associated intervals of uncertainty. In this plot, the blue line represents the estimated StripSearch values given the perceived race groups. The green shaded area around the line represents the prediction interval, which indicates the range of possible values for a new observation. We can predict the black people have the highest stripsearch probability of 0.94. And the green shaded area around the line is quite thin, indicating a low degree of uncertainty in our prediction. However, the white people have a low stripsearch probability of 0.85. And the green shaded area around the line is quite wide, indicating a high degree of uncertainty in our prediction.

8. Results and Findings

8.1 ANCOVA Test

Two ANCOVA tests have been conducted in the previous sections, where we are able to reject all null hypotheses and can conclude that, when controlling the total number of arrests, there are significant mean differences of being strip searched within different race groups and within different occurrence categories.

A strip search is not a mandatory procedure when an individual is arrested. However, our findings also reveal that while controlling the total arrests amounts, there are significant mean differences between various race groups, which raises important questions about potential biases and discrimination during law enforcement practices. The power analysis section of our study also sheds light on the mean differences between black individuals and white individuals, as well as those of East/South Asian descent. These results underscore the urgent need for further investigation and intervention to ensure that all individuals, regardless of their perceived race, are treated fairly and equitably throughout the criminal justice system.

Moreover, there are also significant mean differences found in different occurrence categories when facing strip search practices. This means that certain occurrence categories may be more likely to result in strip searches, even if the total number of arrests is the same. Strip searches may be conducted depending on the severity of the offense, the behavior of the arrestees, and the policies and procedures of the arresting agencies. Still, the test results indicate that there are possibilities that officers may be more likely to assume that individuals in certain occurrence categories are more likely to be carrying contraband, and therefore more likely to conduct strip searches in these cases. This could be due to biases or stereotypes about certain groups of people. Ultimately, our tests highlight the need for ongoing efforts to promote justice and equality in all aspects of law enforcement and criminal justice practices.

8.2 Chi-Square

The findings suggest that there are significant associations between perceived race, strip search, and occurrence category. This means that the occurrence of strip search practices and occurrence category vary depending on the perceived race of the individual being arrested. This raises important questions about potential biases and discrimination during law enforcement practices. This highlights the importance of recognizing and addressing potential biases and discrimination in law enforcement practices. The findings suggest that individuals of certain perceived races may be disproportionately targeted for strip searches, and that occurrence category may also play a role in determining who is subjected to these invasive practices. It is important for law and regulation agencies to be aware of these potential biases and take steps to ensure that all individuals are treated fairly and justly, regardless of their perceived race or the circumstances surrounding their arrest. It signals the need for further investigation and intervention to ensure fair and just treatment for all individuals, regardless of their perceived race for strip search processes when at arrest. It is critical that law enforcement and regulatory organizations recognize these possible prejudices and take efforts to ensure that all persons are treated fairly and justly, regardless of their perceived race or the circumstances surrounding their arrest. It

indicates the need for further investigation and intervention to ensure fair and just treatment for all individuals, regardless of their perceived race for strip search processes when at arrest.

8.3 Logistic Regression

Based on the logistic regression analysis, we found that several factors are significantly associated with the likelihood of suspects being strip searched, including Sex, Perceived Race Black, Perceived Race East Southeast Asian, Perceived Race Indigenous, Perceived Race Middle Eastern, Perceived Race South Asian, Perceived Race White, Occurrence Category Drug Related, Occurrence Category Weapons & Homicide, and Occurrence Category Break & Enter. The only feature that was not statistically significant in this analysis was the Perceived Race Latino. This finding can verify both our research question one: Black individuals are treated differently when arrested with strip search practices compared to other races" and research question two: there is a significant difference in means of the number of strip searches across different occurrence categories, controlling the total arrested number.

The logit regression results indicate that male suspects have a higher likelihood of being strip searched compared to female suspects. Additionally, suspects perceived as Black, Indigenous, and White have a higher likelihood of being strip searched compared to other race groups. Regarding the occurrence categories, suspects involved in drug-related occurrences, weapons, and homicide occurrences, and break and enter occurrences have a higher likelihood of being strip searched compared to suspects not involved in these occurrences. These findings suggest that there are significant disparities in the likelihood of suspects being strip searched based on their sex, perceived race, and the type of occurrence they are involved in.

To further investigate the relationship between perceived race and the likelihood of being strip searched, we conducted another logit regression model with perceived race as a categorical variable with 8 levels. The prediction interval plot helps us to view both the expected value of our predictions (StripSearch) and the uncertainty associated with those predictions. It shows that black people have the highest strip search probability and a low degree of uncertainty in our prediction. On the other hand, white people have a low strip search probability and a high degree of uncertainty in our estimation. Therefore, it is not conflicted with our literature review findings and answers our research questions again.

Overall, the results suggest that perceived race, sex, and occurrence category variables are significant predictors of the likelihood of suspects being strip searched. The odds ratios provide insight into the direction and magnitude of the effects of these variables, while the confusion matrix and prediction interval provide useful information for evaluating the performance of the model and understanding the relationship between perceived race and the likelihood of being strip searched. It is important to note that correlation does not imply causation and further research is needed to establish causal relationships between these variables and the outcome of strip search.

9. Discussion

9.1 Limitation

This dataset contains a sufficient amount of information with several columns and rows, but no numerical data can be discovered in it. Initially, we wish to integrate all search reason categories and rank them in severity order, assuming that they are continuous numerical variables. Because this variable may offer us some insight into if there are underlying relationships between perceived race groups and the reasons why people are strip searched. Yet, there isn't enough research to back up our assumptions and provide a guide on the severity of each search reason, which might render this assumption useless and so contribute nothing to the research subject. Therefore, we decided to assume the strip search column as a numerical value where it represents the probability of getting strip searched. The probability of this event that is certain to occur is represented by 1, while the probability of an event that is impossible to occur is represented by 0. Moreover, we also created new variables for the ANCOVA test since it requires continuous variables, namely Number_StripSearch and Number_Arrest.

Furthermore, another limitation this study has was the sample size issue with power analysis between the Black and White races on strip search practices. The effect size for the difference in strip search practices between black and white individuals is very small, as indicated by Cohen's D metric of 0.0012. This means that the difference between the means of black and white communities is minimal. However, even though the effect size is small, a large sample size is still required to detect this minor difference with statistical power. The power analysis in these scenarios suggested that sample sizes of more than 13 million for the white group and more than 9 million for the black group would be required to attain a statistical power of 80%.

However, the required sample size of 13 million for the white group and 9 million for the black group are likely too large and impractical to obtain in reality and the statistical power of the analysis can only be limited, and the results will be reliable indeed. Therefore, in reality, we suggest that researchers can explore other options to improve the reliability of their results, such as increasing the sample size within the available dataset or utilizing other statistical methods that are less affected by small sample sizes. Additionally, researchers can use their study as a starting point for future research to continue exploring the relationship between the variables of interest with a larger and more diverse sample size.

9.2 Future Work

As mentioned in previous sections, studies have shown that strip searches are considered traumatic experiences for individuals and can lead to long-lasting emotional and psychological damage. As a result, there is an urgent need for more study to better understand the experiences of various racial groups during strip searches and to identify viable solutions that might mitigate the harm caused.

Furthermore, a study done by Reventlow (2020) argued that while statistics can give significant insights into the amount and character of racism, relying solely on data collection is insufficient to address and mitigate the systematic bias in racism. Comprehensive plans that comprehend the

complexities of problems and involve various parties and techniques are essential to accomplish long-term change. For example, by collaborating with people of color and other marginalized communities and emphasizing their needs and viewpoints, society may develop more effective ways for combating racism and establishing a more inclusive environment (Reventlow, 2020).

10. Conclusion

This study is an attempt to explore the question of “Is there any difference or bias of stripsearch in terms of race when being arrested?” We are particularly interested in sub research questions of comparing stripsearch of black to white and East/ South Asian, number of strip searches among different occurrence categories, and the associated relationship between perceived race and occurrence categories. Regardless of the limits, there is still considerable evidence to demonstrate that there are differences and biases of strip search when being arrested among different racial groups. Black, white, and indigenous communities display significant association with being strip searched than other racial groups. In addition, during our exploratory study, we verify the association between perceived race and occurrence categories. And we also verify the significant associations of strip searches and occurrence categories, specifically drugs, weapons and homicide,and break and enter, which is corresponding to our literature review. However, the study has limitations such as the assumption of categorical variables as numerical values and small sample sizes that affect statistical power. Future research should focus on understanding the experiences of various racial groups during strip searches and identifying viable solutions to mitigate the harm caused. Additionally, it's important to implement comprehensive plans that involve various parties and techniques to address and combat racism in law enforcement practices. Only then can we hope to build a system that truly serves the needs and interests of all members of the society.

Reference

AAT Bioquest. (2022, May 19). What are the assumptions of ANCOVA? AAT Bioquest. Retrieved April 14, 2023, from https://www.aatbio.com/resources/faq-frequently-asked-questions/what-are-the-assumptions-of-ancova

Bhandari, P. (2022, November 17). What is effect size and why does it matter? (examples). Scribbr. Retrieved April 14, 2023, from https://www.scribbr.com/statistics/effect-size/#:~:text=Effect%20size%20tells%20you%20how,size%20indicates%20limited%20practical%20applications

CBCnews. (2022, June 16). *Toronto police use more force against black people with little explanation, data shows | CBC News*. CBCnews. Retrieved February 26, 2023, from https://www.cbc.ca/news/canada/toronto/toronto-police-race-based-data-use-force-strip-searches-1.6489151

Commercial Aviation Safety. (2011, October 1). *Aviation occurrence categories definitions and usage notes - ICAO*. Retrieved February 26, 2023, from https://www.icao.int/SAM/Documents/2017-SSP-GUY/CICTT%20Occurrence%20Category.pdf

Findlaw. (2019, March 19). *Police booking procedure*. Findlaw. Retrieved February 21, 2023, from https://www.findlaw.com/criminal/criminal-procedure/booking.html

Interactive. (2022, November 2). *Report to the United Nations on racial disparities in the U.S. criminal justice system*. The Sentencing Project. Retrieved February 21, 2023, from https://www.sentencingproject.org/reports/report-to-the-united-nations-on-racial-disparities-in-the-u-s-criminal-justice-system/

Kruse Law Firm. (n.d.). *When can police do a 'Pat Down' search*. Kruse Law Firm. Retrieved February 21, 2023, from https://www.kruselaw.ca/library/when-can-police-do-a-pat-down-search-kruse-law.cfm#:~:text=When%20associated%20with%20a%20lawful,a%20same%20gender%20police%20officer.

Lemke, M. (2022, June 17). *Strip searches are ineffective, unnecessary and target racialized Canadians*. Phys.org. Retrieved February 26, 2023, from https://phys.org/news/2022-06-ineffective-unnecessary-racialized-canadians.html

Phan, M. (2021, August 18). *Race-based data in the Criminal Justice System*. Maytree. Retrieved February 27, 2023, from https://maytree.com/publications/race-based-data-in-the-criminal-justice-system/

Piquero, A. R., & Brame, R. W. (2008). Assessing the race–crime and ethnicity–crime relationship in a sample of serious adolescent delinquents. *Crime & Delinquency*, *54*(3), 390–422. https://doi.org/10.1177/0011128707307219

Reventlow, N. J. (2020, July 29). *Data collection is not the solution for Europe's racism problem*. Racism | Al Jazeera. Retrieved February 26, 2023, from https://www.aljazeera.com/opinions/2020/7/29/data-collection-is-not-the-solution-for-europes-racism-problem/

Toronto Police Service. (2022, October 10). Arrests and Strip Searches (RBDC-ARR-TBL-001). Toronto. Retrieved February 21, 2023, from https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/about.

Appendices

ANCOVA Test Result

Numer_StripSearch & Perceived Race

OLS Regression Results

Dep. Variable:	Number_StripSearch	R-squared:	0.399
Model:	OLS	Adj. R-squared:	0.399
Method:	Least Squares	F-statistic:	3095.
Date:	Sat, 15 Apr 2023	Prob (F-statistic):	0.00
Time:	00:28:32	Log-Likelihood:	-27042.
No. Observations:	37231	AIC:	5.410e+04
Df Residuals:	37222	BIC:	5.418e+04
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	-0.1233	0.006	-22.336	0.000	-0.134	-0.112
Perceived_Race[T.East/Southeast Asian]	-0.0569	0.010	-5.519	0.000	-0.077	-0.037
Perceived_Race[T.Indigenous]	-0.0348	0.020	-1.730	0.084	-0.074	0.005
Perceived_Race[T.Latino]	-0.0817	0.016	-5.117	0.000	-0.113	-0.050
Perceived_Race[T.Middle-Eastern]	-0.0679	0.012	-5.565	0.000	-0.092	-0.044
Perceived_Race[T.South Asian]	-0.0726	0.011	-6.434	0.000	-0.095	-0.050
Perceived_Race[T.Unknown or Legacy]	-0.0338	0.010	-3.440	0.001	-0.053	-0.015
Perceived_Race[T.White]	-0.0277	0.007	-4.240	0.000	-0.040	-0.015
Number_Arrest	0.2019	0.001	155.389	0.000	0.199	0.204

Omnibus:	25527.771	Durbin-Watson:	2.011
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1649823.828
Skew:	2.632	Prob(JB):	0.00
Kurtosis:	35.184	Cond. No.	22.2

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Number_StripSearch & Occurrence Category

OLS Regression Results

Dep. Variable:	Number_StripSearch	R-squared:	0.436
Model:	OLS	Adj. R-squared:	0.436
Method:	Least Squares	F-statistic:	958.5
Date:	Sat, 15 Apr 2023	Prob (F-statistic):	0.00
Time:	00:28:33	Log-Likelihood:	-25874.
No. Observations:	37231	AIC:	5.181e+04
Df Residuals:	37200	BIC:	5.207e+04
Df Model:	30		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	-0.2363	0.007	-33.363	0.000	-0.250	-0.222
Occurrence_Category[T.Assault & Other crimes against persons]	0.1189	0.010	12.480	0.000	0.100	0.138
Occurrence_Category[T.Break & Enter]	0.1844	0.020	8.996	0.000	0.144	0.225
Occurrence_Category[T.Crimes against Children]	7.878e-05	0.067	0.001	0.999	-0.131	0.131
Occurrence_Category[T.Drug Related]	0.4441	0.013	33.038	0.000	0.418	0.470
Occurrence_Category[T.FTA/FTC, Compliance Check & Parollee]	0.3433	0.014	23.951	0.000	0.315	0.371
Occurrence_Category[T.FTA/FTC/Compliance Check/Parollee]	-0.0326	0.021	-1.579	0.114	-0.073	0.008
Occurrence_Category[T.Fraud]	-0.0032	0.027	-0.120	0.905	-0.056	0.050
Occurrence_Category[T.Harassment & Threatening]	0.1445	0.018	8.054	0.000	0.109	0.180
Occurrence_Category[T.Harassment/Threatening]	-0.0195	0.019	-1.037	0.300	-0.056	0.017
Occurrence_Category[T.Homicide]	0.0665	0.070	0.945	0.344	-0.071	0.204
Occurrence_Category[T.Impaired]	0.0295	0.016	1.873	0.061	-0.001	0.060
Occurrence_Category[T.LLA]	0.0046	0.030	0.154	0.878	-0.053	0.063
Occurrence_Category[T.Mental Health]	-0.0157	0.043	-0.368	0.713	-0.099	0.068
Occurrence_Category[T.Mischief]	-0.0233	0.021	-1.113	0.266	-0.064	0.018
Occurrence_Category[T.Mischief & Fraud]	0.0560	0.016	3.487	0.000	0.025	0.088
Occurrence_Category[T.Other Offence]	0.1058	0.016	6.745	0.000	0.075	0.137
Occurrence_Category[T.Other Statute]	-0.0376	0.023	-1.652	0.098	-0.082	0.007
Occurrence_Category[T.Other Statute & Other Incident Type]	-0.0159	0.014	-1.121	0.262	-0.044	0.012
Occurrence_Category[T.Police Category - Administrative]	-0.0221	0.017	-1.288	0.198	-0.056	0.012
Occurrence_Category[T.Police Category - Incident]	-0.0106	0.027	-0.390	0.697	-0.064	0.043
Occurrence_Category[T.Robbery & Theft]	0.1091	0.011	9.734	0.000	0.087	0.131
Occurrence_Category[T.Robbery/Theft]	-0.0183	0.013	-1.359	0.174	-0.045	0.008
Occurrence_Category[T.Sexual Related Crime]	0.0044	0.021	0.209	0.834	-0.036	0.045
Occurrence_Category[T.Sexual Related Crimes & Crimes Against Children]	0.0632	0.020	3.114	0.002	0.023	0.103
Occurrence_Category[T.Vehicle Related]	0.0253	0.021	1.223	0.221	-0.015	0.066
Occurrence_Category[T.Vehicle Related (inc. Impaired)]	0.0435	0.013	3.253	0.001	0.017	0.070
Occurrence_Category[T.Warrant]	0.1451	0.012	11.804	0.000	0.121	0.169
Occurrence_Category[T.Weapons]	0.0567	0.020	2.880	0.004	0.018	0.095
Occurrence_Category[T.Weapons & Homicide]	0.4408	0.018	24.199	0.000	0.405	0.477
Number_Arrest	0.1956	0.001	152.262	0.000	0.193	0.198

Omnibus:	25894.269	Durbin-Watson:	2.017
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1951470.314
Skew:	2.643	Prob(JB):	0.00
Kurtosis:	38.072	Cond. No.	77.1

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Chi-Square Test Result

Perceived Race & StripSearch

```
observed frequencies:
 StripSearch               0     1
Perceived_Race
Black                  15053  2299
East/Southeast Asian    4061   327
Indigenous              1620   287
Latino                  1626   125
Middle-Eastern          2999   214
South Asian             3346   248
Unknown or Legacy       4506   493
White                  24064  3338

Chi-squared statistic: 356.12485254652904
p-value: 6.017709683374717e-73
Degrees of freedom: 7
Expected frequencies:
 [[15383.0263443   1968.9736557 ]
 [ 3890.08296443   497.91703557]
 [ 1690.60807046   216.39192954]
 [ 1552.30976999   198.69023001]
 [ 2848.41307309   364.58692691]
 [ 3186.18007615   407.81992385]
 [ 4431.75130793   567.24869207]
 [24292.62839365  3109.37160635]]
```

Perceived Race & Occurrence_Category

```
observed frequencies:
 Occurrence_Category   Assault  Assault & Other crimes against persons  \
Perceived_Race
Black                    2127                                    1962
East/Southeast Asian      648                                     595
Indigenous                186                                     220
Latino                    284                                     235
Middle-Eastern            475                                     378
South Asian               490                                     452
Unknown or Legacy         598                                     578
White                    2916                                    2758

Occurrence_Category   Break & Enter  Crimes against Children  Drug Related  \
Perceived_Race
Black                           280                        9           942
East/Southeast Asian             81                        9           209
Indigenous                       68                        1            57
Latino                           36                        2            29
Middle-Eastern                   41                        2            82
South Asian                      40                        6           114
Unknown or Legacy               100                       13           224
White                          1069                       27          1076

Occurrence_Category   FTA/FTC, Compliance Check & Parollee  \
Perceived_Race
Black                                                1009
East/Southeast Asian                                  197
Indigenous                                            130
Latino                                                 89
Middle-Eastern                                        176
South Asian                                           189
Unknown or Legacy                                     288
White                                                1670
```

```
Occurrence_Category   FTA/FTC/Compliance Check/Parollee  Fraud  \
Perceived_Race
Black                                              1122    118
East/Southeast Asian                                225     57
Indigenous                                          162      2
Latino                                              113     14
Middle-Eastern                                      193     21
South Asian                                         210     35
Unknown or Legacy                                   247     70
White                                              1968    156

Occurrence_Category   Harassment & Threatening  Harassment/Threatening  ...  \
Perceived_Race                                                          ...
Black                                      328                     338  ...
East/Southeast Asian                        63                     102  ...
Indigenous                                  22                      29  ...
Latino                                      41                      43  ...
Middle-Eastern                              84                     105  ...
South Asian                                 90                      71  ...
Unknown or Legacy                          106                     122  ...
White                                      522                     534  ...

Occurrence_Category   Police Category - Incident  Robbery & Theft  \
Perceived_Race
Black                                        144             1064
East/Southeast Asian                          54              280
Indigenous                                    14              149
Latino                                        15               87
Middle-Eastern                                33              204
South Asian                                   24              255
Unknown or Legacy                             42              341
White                                        235             2110
```

Occurrence_Category	Robbery/Theft	Sexual Related Crime	\
Perceived_Race			
Black	916	240	
East/Southeast Asian	238	68	
Indigenous	126	8	
Latino	78	42	
Middle-Eastern	167	75	
South Asian	177	62	
Unknown or Legacy	282	100	
White	1763	244	

Occurrence_Category	Sexual Related Crimes & Crimes Against Children	\
Perceived_Race		
Black	200	
East/Southeast Asian	59	
Indigenous	9	
Latino	26	
Middle-Eastern	51	
South Asian	66	
Unknown or Legacy	103	
White	269	

Occurrence_Category	Vehicle Related	Vehicle Related (inc. Impaired)	\
Perceived_Race			
Black	221	500	
East/Southeast Asian	80	174	
Indigenous	4	8	
Latino	31	65	
Middle-Eastern	62	125	
South Asian	72	212	
Unknown or Legacy	65	214	
White	233	679	

Occurrence_Category	Warrant	Weapons	Weapons & Homicide
Perceived_Race			
Black	1235	495	492
East/Southeast Asian	200	51	50
Indigenous	170	25	23
Latino	99	27	25
Middle-Eastern	159	55	63
South Asian	140	70	31
Unknown or Legacy	273	82	84
White	2083	300	315

```
Chi-squared statistic: 2751.6587949730747
p-value: 0.0
Degrees of freedom: 203
Expected frequencies:
 [[2.07452633e+03 1.92788063e+03 4.60617899e+02 1.85321487e+01
  7.34034238e+02 1.00664483e+03 1.13878711e+03 1.27039222e+02
  3.37338823e+02 3.60974027e+02 1.93378943e+01 3.65808501e+02
  1.57388973e+02 6.41910658e+01 3.54528062e+02 4.59812154e+02
  5.26420456e+02 2.53541281e+02 6.26870074e+02 1.04344055e+03
  1.50674427e+02 1.20593258e+03 1.00637625e+03 2.25340185e+02
  2.10299601e+02 2.06270873e+02 5.30986348e+02 1.17074835e+03
  2.96782961e+02 2.90874160e+02]
 [5.24609355e+02 4.87525369e+02 1.16481751e+02 4.68643779e+00
  1.85623688e+02 2.54561867e+02 2.87978206e+02 3.21258707e+01
  8.53067517e+01 9.12836579e+01 4.89019596e+00 9.25062069e+01
  3.98007615e+01 1.62327338e+01 8.96535925e+01 1.16277993e+02
  1.33122001e+02 6.41159025e+01 1.58523852e+02 2.63866824e+02
  3.81027768e+01 3.04958053e+02 2.54493948e+02 5.69843668e+01
  5.31808810e+01 5.21620902e+01 1.34276631e+02 2.96060614e+02
  7.50509241e+01 7.35566975e+01]
 [2.27992261e+02 2.11875770e+02 5.06223106e+01 2.03669938e+00
  8.06710058e+01 1.10631149e+02 1.25153701e+02 1.39617218e+01
  3.70738322e+01 3.96713618e+01 2.12525152e+00 4.02026747e+01
  1.72971860e+01 7.05465437e+00 3.89629446e+01 5.05337585e+01
  5.78540693e+01 2.78644089e+01 6.88935703e+01 1.14675030e+02
  1.65592515e+01 1.32533046e+02 1.10601631e+02 2.47650837e+01
  2.31121103e+01 2.26693496e+01 5.83558648e+01 1.28666269e+02
  3.26167074e+01 3.19673250e+01]
 [2.09341609e+02 1.94543510e+02 4.64812092e+01 1.87008947e+00
  7.40718045e+01 1.01581092e+02 1.14915643e+02 1.28195988e+01
  3.40410488e+01 3.64260905e+01 1.95139770e+00 3.69139399e+01
  1.58822091e+01 6.47755626e+00 3.57756246e+01 4.63999009e+01
  5.31213819e+01 2.55849921e+01 6.32578089e+01 1.05294168e+02
  1.52046404e+01 1.21691329e+02 1.01553989e+02 2.27392038e+01
  2.12214500e+01 2.08149088e+01 5.35821286e+01 1.18140869e+02
  2.99485342e+01 2.93522738e+01]
 [3.84131691e+02 3.56977897e+02 8.52907625e+01 3.43152339e+00
  1.35918165e+02 1.86396372e+02 2.10864626e+02 2.35233415e+01
  6.24636721e+01 6.68401077e+01 3.58072006e+00 6.77352877e+01
  2.91430827e+01 1.18860013e+01 6.56465344e+01 8.51415658e+01
  9.74751571e+01 4.69472185e+01 1.16075009e+02 1.93209686e+02
  2.78997771e+01 2.23297681e+02 1.86346640e+02 4.17253351e+01
  3.89403306e+01 3.81943473e+01 9.83206049e+01 2.16782760e+02
  5.49541064e+01 5.38599975e+01]


[4.29682321e+02 3.99308609e+02 9.54046064e+01 3.83843606e+00
 1.52035446e+02 2.08499396e+02 2.35869114e+02 2.63127573e+01
 6.98706622e+01 7.47660589e+01 4.00532458e+00 7.57673900e+01
 3.25988917e+01 1.32954524e+01 7.34309507e+01 9.52377179e+01
 1.09033836e+02 5.25142556e+01 1.29839272e+02 2.16120639e+02
 3.12081540e+01 2.49776491e+02 2.08443767e+02 4.66731573e+01
 4.35579048e+01 4.27234622e+01 1.09979538e+02 2.42489026e+02
 6.14706064e+01 6.02467573e+01]
[5.97657741e+02 5.55410055e+02 1.32701065e+02 5.33899328e+00
 2.11470560e+02 2.90007925e+02 3.28077268e+02 3.65991858e+01
 9.71851531e+01 1.03994304e+02 5.57112343e+00 1.05387085e+02
 4.53427545e+01 1.84930347e+01 1.02137263e+02 1.32468935e+02
 1.51658360e+02 7.30436182e+01 1.80597251e+02 3.00608535e+02
 4.34083367e+01 3.47421447e+02 2.89930548e+02 6.49190632e+01
 6.05859672e+01 5.94253165e+01 1.52973764e+02 3.37285097e+02
 8.55012692e+01 8.37989815e+01]
[3.27605869e+03 3.04447816e+03 7.27400396e+02 2.92656719e+01
 1.15917509e+03 1.58967737e+03 1.79835433e+03 2.00618302e+02
 5.32720057e+02 5.70044392e+02 3.05380924e+01 5.77678915e+02
 2.48546141e+02 1.01369501e+02 5.59865028e+02 7.26127976e+02
 8.31314739e+02 4.00388323e+02 9.89943163e+02 1.64778457e+03
 2.37942637e+02 1.90438938e+03 1.58925323e+03 3.55853605e+02
 3.32101755e+02 3.25739653e+02 8.38525122e+02 1.84882701e+03
 4.68674891e+02 4.59343807e+02]]
```

Occurrence_Category & StripSearch

```
observed frequencies:
 Occurrence_Category  Assault  Assault & Other crimes against persons  \
StripSearch
0                       7620                                    5807
1                        104                                    1371

Occurrence_Category  Break & Enter  Crimes against Children  Drug Related  \
StripSearch
0                            1323                       69          1674
1                             392                        0          1059

Occurrence_Category  FTA/FTC, Compliance Check & Parollee  \
StripSearch
0                                                   2636
1                                                   1112

Occurrence_Category  FTA/FTC/Compliance Check/Parollee  Fraud  \
StripSearch
0                                               4154    473
1                                                 86      0

Occurrence_Category  Harassment & Threatening  Harassment/Threatening  ...  \
StripSearch                                                           ...
0                                       1007                    1337  ...
1                                        249                       7  ...

Occurrence_Category  Police Category - Incident  Robbery & Theft  \
StripSearch
0                                          560             3768
1                                            1              722

Occurrence_Category  Robbery/Theft  Sexual Related Crime  \
StripSearch
0                            3685                   831
1                              62                     8

Occurrence_Category  Sexual Related Crimes & Crimes Against Children  \
StripSearch
0                                                              699
1                                                               84


 Occurrence_Category  Vehicle Related  Vehicle Related (inc. Impaired)  \
 StripSearch
 0                               763                              1879
 1                                 5                                98

 Occurrence_Category  Warrant  Weapons  Weapons & Homicide
 StripSearch
 0                       3493     1042                644
 1                        866       63                439
```

```
Chi-squared statistic: 8547.072981535563
p-value: 0.0
Degrees of freedom: 29
Expected frequencies:
 [[6847.53892827 6363.49487664 1520.39477757   61.17040213 2422.87984088
  3322.70532149 3758.8768845   419.32753924 1113.47862428 1191.49305018
    63.82998483 1207.45054639  519.50515432  211.88008854 1170.21638857
  1517.73519487 1737.59403151  836.88202334 2069.15534161 3444.15959818
   497.34196514 3980.50877627 3321.81879392  743.7966288   694.15108504
   680.85317153 1752.66500015 3864.37366498  979.61296164  960.10935517]
 [ 876.46107173  814.50512336  194.60522243    7.82959787  310.12015912
   425.29467851  481.1231155    53.67246076  142.52137572  152.50694982
     8.17001517  154.54945361   66.49484568   27.11991146  149.78361143
   194.26480513  222.40596849  107.11797666  264.84465839  440.84040182
    63.65803486  509.49122373  425.18120608   95.2033712    88.84891496
    87.14682847  224.33499985  494.62633502  125.38703836  122.89064483]]
```