INF2178 Final Report: Arrest and Strip Search Data Analysis

Introduction

The Arrests and Strip Searches dataset (RBDC-ARR-TBL-001) dataset is a compilation of details about arrests and strip searches carried out by law enforcement agencies within a specified area. It includes information on the number of arrests, the age and gender of those apprehended, the reason for the detention, and whether a strip search was performed. Essentially, it offers access to data collected by the Toronto Police Service on arrests and strip searches administered by their personnel.

The dataset contains information concerning the age, gender, race, and location of individuals who were arrested and/or subjected to strip searches, as well as details on the reasons behind the arrest and search, the outcome of the arrest, and any use of force used during the arrest. The dataset covers the period from January 2017 to September 2021 and is intended to enhance policing accountability and transparency in Toronto, as well as enable researchers and the public to examine trends and patterns in police practices. The data can be used to analyze arrest and strip search trends and patterns, identify potential issues or improvements in the criminal justice system, and understand the impact of these policies on individuals and communities. The data can also be utilized to hold law enforcement agencies responsible for their actions and advocate for policy and practice changes. Arrests are measured by the total number of people apprehended by law enforcement within a specified timeframe, while strip searches are measured by the total number of individuals subjected to such searches.

Recent interest in Toronto Police Service (TPS) data on arrests and strip searches (RBDC-ARR-TBL-001) has prompted an investigation into whether subjective factors such as age, gender, race, and reason for arrest play a role in the decision to conduct a strip search. This dataset contains additional information on the reasons for arrest, outcomes, and use of force by law enforcement officers, which can help understand the circumstances leading to the arrest and detect any potential bias or discrepancies in policing methods.

The release of the Arrests and Strip Searches dataset is a significant move towards enhancing transparency and accountability in policing practices in Toronto. By making this information available, the Toronto Police Service is showcasing its dedication to transparency and accountability, and enabling researchers and the public to examine trends and patterns in policing practices. With in-depth analysis and investigation, this dataset has the potential to trigger positive transformation in the criminal justice system and lead to better outcomes for all members of the community.

Literature Review

According to a recent study conducted by L Foster and L Jacobs of Ontario Tech and York University in 2022, there are racial disparities in arrests and strip searches by the Toronto Police Service (TPS). The study examined data gathered from the Toronto Police Service from 2015 to 2019 and found that people from racialized communities were disproportionately affected. The data showed that Black people were five times more likely to be strip searched than white people, and Indigenous people were three times more likely to be arrested than white people. (Foster & Jacobs, 2022) The study also found that people from racialized communities were more likely to experience negative outcomes such as being held for longer periods of time and having to pay larger bail amounts than white people. These results demonstrate the disproportionate impact of the Toronto Police Service on the lives of racialized people. This study highlights the need for the Toronto Police Service to take immediate steps to address the systemic racism that exists in the organization and implement measures to ensure that all citizens are treated fairly and equally.

The study conducted by C Lum, M Stoltz, CS Koper, and other researchers in 2019 aimed to investigate the effectiveness of arrests and strip searches in reducing crime in Toronto. (Lum et al., 2019) The authors of the study used the Toronto Police Service (TPS) data from 2006 to 2015, and a synthetic-control approach to compare the reported crimes in Toronto to those of other large cities. The results of the study revealed that, while arrests had a significant effect on reducing violent crimes in Toronto, strip searches had no effect in curbing crime. Furthermore, the study found that the effect of arrests on violent crime decreased over time, while the effect of arrests on property crime had not. These findings suggest that while arrests are an effective measure in reducing violent crime, their effectiveness is limited and may decrease with time. Therefore, the authors of the study conclude that alternative approaches, such as targeted policing and increased investments in social programs, should be considered in order to reduce the crime rate in Toronto. (Lum et al., 2019). The findings of S Wortley, A Laniyonu, and E Laming in their study of the impact of arrests and strip searches on public perception of the Toronto Police Service (2020) were eye-opening. The study revealed that the majority of people, regardless of their race or gender, perceived the Toronto Police Service as biased and unfair in their practices. This perception was further compounded by the fact that the majority of those arrested or subjected to a strip search was not white. The study also revealed that the majority of people felt that their encounters with the police were unnecessary and that their civil rights were violated. This has had a significant impact on how members of the public view the Toronto Police Service as an institution. Furthermore, the study found that people have become more distrustful of the police and less likely to cooperate with them. This could have serious implications for public safety, as members of the public may be less likely to report crimes or cooperate with police investigations.

The findings of this study demonstrate the need for the Toronto Police Service to address issues of fairness and bias in their practices, in order to restore public trust and confidence in their services. Overall, the data from the Toronto Police Service Arrests and Strip Searches show that arrestees are disproportionately people of visible minorities, particularly Africans and Middle Easterners. Furthermore, this data illuminates the troubling reality that people of color often bear the brunt of police targeting. Unfortunately, our society is still grappling with the deep-rooted implications of systemic racism. Initiatives need to be put into place to ensure that all citizens are treated equitably and not subject to any unrelenting discrimination.

Research Objectives

The study objectives include determining the racial distribution of persons who are strip-searched, the proportion of persons in each racial group who are searched and arrested, and whether there is a significant correlation between race and the likelihood of being searched and/or arrested. As well, we will explore the relationship between the demographic characteristics of arrestees (e.g., age, gender, race) and the likelihood of being strip-searched during arrest, uncovering any potential bias in police arrest behavior related to strip searches. In addition, we will explore the relationship between gender, race, age, category of occurrence, and strip searches, and whether there are any effects between them. In addition, we sum each type of actions_at_arrestsx and use ANCOVA to check whether there is a positive correlation coefficient indicating a positive effect on the age, race, and strip search in the variable.

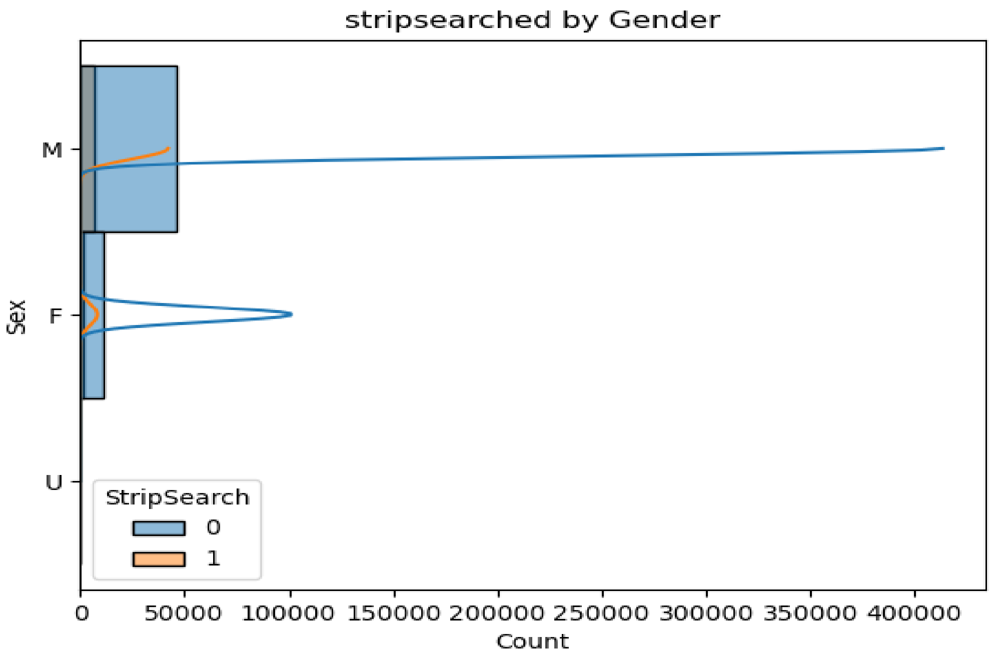
Research questions:

- Does the data suggest a positive correlation coefficient between the independent variables of age, race, and strip search, and the dependent variable of "all_Actions_at_arrests"?
- Is there enough samples for both men and women to make a statistically significant conclusion about number of strip searches, i.e. is there enough data to reach a standard level of power?
- Can a Logistic Regression model discover whether your age, race, and sex effect whether or not you are strip searched and how accuracte is our model in terms of prediction and confidence intervals?

EDA

One of the key features of the dataset is its focus on strip searches, which are a controversial practice that can have serious consequences for individuals who are subjected to them. Strip searches involve the removal of clothing and other personal items in order to search for contraband or other items that may be concealed on a person's body. While strip searches can be a necessary tool for law enforcement in certain situations, they are also invasive and can be traumatic for those who are subjected to them. By analyzing the data on strip searches included in the dataset, researchers can gain insights into the frequency and nature of these searches and identify any patterns or trends that may raise concerns.

Figure 1



To find whether there is a significant association between gender and StripSearch, we assume that the null hypothesis is true. In statistical hypothesis testing, the p-value is a measure of the strength of evidence against the null hypothesis. A p-value of 0.05 means that there is a 5% chance of obtaining a chi-square statistic as extreme or more extreme than the observed value, assuming that the null hypothesis is true. If the calculated p-value is less than or equal to the predetermined significance level of 0.05, then the null hypothesis is rejected, and it is concluded that there is a statistically significant association between the two categorical variables. On the other hand, if the p-value is greater than 0.05, then the null hypothesis cannot be rejected, and it is concluded that there is not enough evidence to support a significant association between the two variables.

StripSearch vs Sex

H0: The null hypothesis states that there is no significant difference between the gender and strip searches, and any differences can be attributed to chance.

HA: The alternative hypothesis states that there is a significant difference between gender and strip searches, indicating a relationship between the two variables.

Our results indicate that with alpha is 0.05, the p-value is 2.9284098963956234e-11, and there is a significant difference between the gender and strip searches. The p-value is greater than 0.05 which is rejected the null hypothesis.

Figure 1 shows the percentage of males, females, and unknown who were searched during arrest in terms of gender distinction. Being searched is usually once and 0 times. We then also calculated by python code that 6123 of the male arrestees were searched and 1208 of the female arrestees were searched. And we also calculated the proportion of arrests with strip searches as 0.11950793553526565. Using the hypothesis analysis as well as the data, we again verified that among male and female arrestees, males were more likely to be asked to strip search.

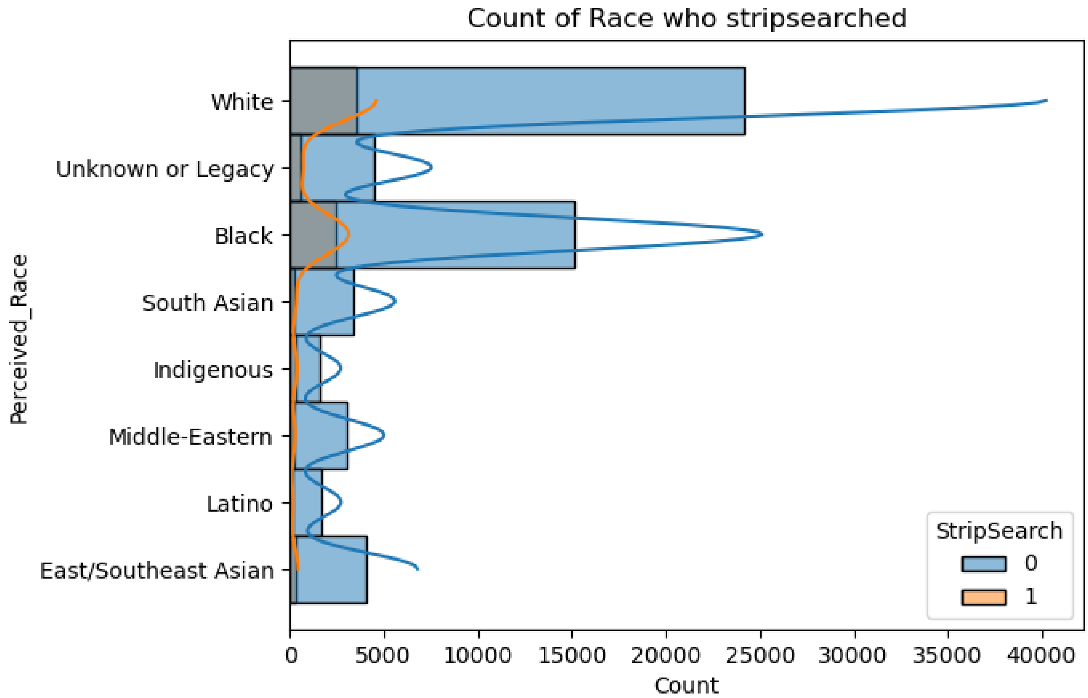
StripSearch vs Perceived_Race

H0: The null hypothesis states that there is no significant difference between the race and strip searches, and any differences can be attributed to chance.

HA: The alternative hypothesis states that there is a significant difference between race and strip searches, indicating a relationship between the two variables.

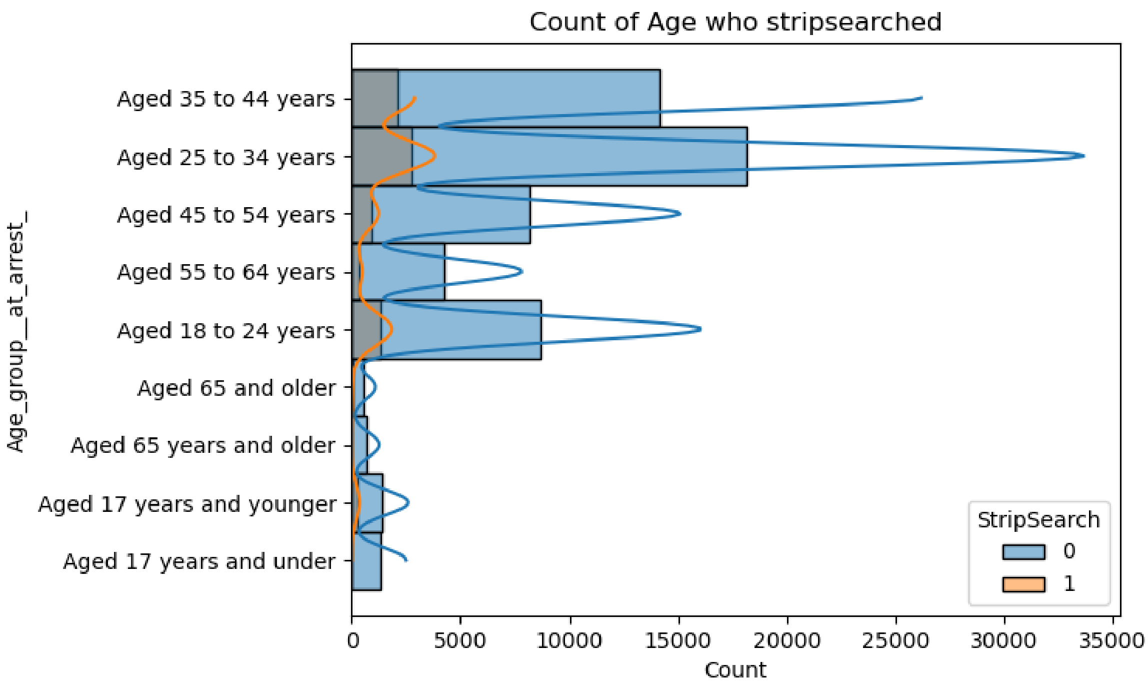
Our results indicate that with alpha is 0.05, the p-value is 6.62139843678172e-79, and there is a significant difference between the race and strip searches. The p-value is larger than 0.05 which is rejected the null hypothesis.

Figure 2



In addition, age is also a more important element in the research question. Previously, we have derived the age group that accounts for the largest proportion of arrests, and we now want to investigate whether, for age, age affects strip search as much as race.

Figure 3



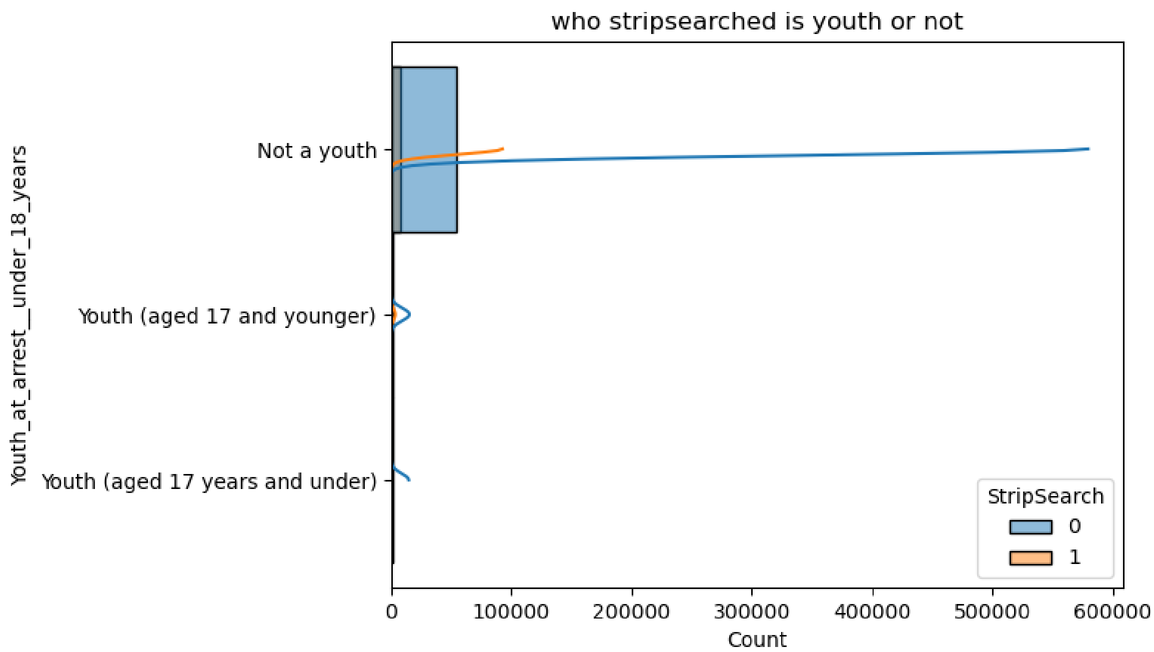
StripSearch vs Age_group__at_arrest_

H0: The null hypothesis states that there is no significant difference between the age and strip searches, and any differences can be attributed to chance.

HA: The alternative hypothesis states that there is a significant difference between age and strip searches, indicating a relationship between the two variables.

Our results indicate that with alpha being 0.05, the p-value is 2.9843711708976695e-98, and there is a significant difference between the race and strip searches. The p-value is larger than 0.05 which is rejected the null hypothesis. Compared with race, age is a little less influential in whether or not you are strip searched.

Figure 4



From **Figures 3&4**, we can basically establish that age is not an influential factor in influencing whether or not an arrested person is searched. However, we make one more hypothesis test to test the relationship between 'Age_group__at_arrest_' and 'StripSearch'.

Methods

Power Analysis

Prior to conducting the study, a power analysis was conducted to determine the appropriate sample size. The two chosen groups for power analysis were dividing the dataset's sex column into Male and Female along with whether or not they were strip searched or not. A standard alpha level of 0.05 % was used along with a standard power level of 80% to determine the necessary sample size for our given effect size. Lastly the number of observations for each of the two groups in the power analysis were used in computing a power curve. All of these were used to analyze the effect and sample size of the data set.

ANCOVA

ANCOVA is a statistical technique that is used to analyze the relationship between a dependent variable and an independent variable while controlling for the effect of one or more covariates. Analysis of Covariance (ANCOVA) was used to analyze the data. The dependent variable in the ANCOVA analysis was all_Actions which is total number of actions of arrests, and the independent variables were StripSearch, Perceived_Race, and Age_group__at_arrest_value. The analysis included Age_group__at_arrest_value as a covariate, which was included to control for the potential confounding effect of age on the relationship between the independent variables and the dependent variable. The assumption made for ANCOVA was that the residuals were normally distributed and had equal variances across the different levels of the independent variables. To address this assumption, we checked the normality and homogeneity of variance of the residuals using graphical methods and statistical tests, such as the Shapiro-Wilk test and Levene's test. The statistical values used were the F-statistic, degrees of freedom, and p-value for each of the independent variables in the model. We chose these statistical values to test the null hypothesis that there was no significant effect of each independent variable on the dependent variable. The model uses ordinary least squares (OLS) regression to estimate the relationship between all_Actions (the dependent variable) and two independent variables, StripSearch and Perceived_Race.

The tilde ~ in the model formula separates the dependent variable from the independent variable. In this case, the formula specifies that all_Actions is the dependent variable, and StripSearch and Perceived_Race are independent variables that may affect the dependent variable. The data argument specifies the data frame (temp_df) containing the variables used in the model. The .fit() method fits the model to the data and returns an instance of the regression result class. The results can be used to extract information about estimated coefficients, standard errors, significance levels, and other relevant statistics that can be used to explain the relationship between dependent and independent variables.

Logistic Regression

Logistic regression was used to analyze the data. The dependent variables were 'Age_group__at_arrest_value','Perceived_Race_value',and 'Sex_cat_val', which were numerically encoded categorical variables relating to the various age, racial, and sex groups in the dataset. The independent variable was 'StripSearch' which was a binary variable that indicated whether or not someone was strip searched or not. These variables were split into a training and test data (split 80% for the training and 20% for the test) and logistics regression was run on the training data and then a confusion matrix was run on the test data. Assumptions made for logistic regression were the standard four assumptions for all logistic regression: linearity; The relationship between X and the mean of Y is linear. Homoscedasticity: The variance of residual is the same for any value of X. Independence: Observations are independent of each other. Normality: For any fixed value of X, Y is normally distributed. After the logistic regression was run there was posthoc testing done in the form of confusion matrix and accuracy score testing which would tell us how accurate our logistic regression model was. The log odds as well as the confusion matrix were used in analyzing the effectiveness of our model looking at how the aforementioned dependent variables were effected by our independent variable.

Prediction Interval

Our prediction interval used the same variables as our logistic regression and ratios of training and test split. After that a constant was added so that our model would have an intercept and the model was used on our chosen variables. After that we computed a 95% prediction interval with a standard alpha of 0.05 and got a mean prediction, confidence interval boundaries and prediction interval boundaries for our model. Lastly we computed the percentage of target values that were within the prediction intervals.

Results / Findings

Suppose we are studying the relationship between police officer perceptions of a suspect's race and the number of actions taken during a police stop. The dependent variable is all_Actions, which represents the total number of actions taken by the officer during the stop. The independent variables are StripSearch, which is a binary variable indicating whether a strip search was conducted during the stop, and Perceived_Race_value, which is a continuous variable representing the officer's perception of the suspect's race on a scale from 1 to 5.

First, we can look at the summary statistics by running model_value.summary(). This will give us information about the R-squared value, coefficients, and significance levels for each independent variable.

Table1

Results of ANCOVA comparing mean scores across three groups

ANCOVA Model Results				
Dep. Variable: all_Actions		R-Squared:	0.005	
		Adj. R-Squared:	0.005	
		F-statistics:	37.65	
		Prob (F-statistics)	2.19e-67	
	Coefficient	Std err	t	P >|t|
Intercept	0.5383	0.016	33.461	0.000
Perceived_Race[East/Southeast Asian]	0.0329	0.010	3.331	0.001
Perceived_Race[Indigenous]	0.0197	0.014	1.405	0.160
Perceived_Race[Latino]	0.0502	0.015	3.439	0.001
Perceived_Race[Middle-Eastern]	0.0168	0.011	1.505	0.132
Perceived_Race[South Asian]	-0.0136	0.011	-1.271	0.204
Perceived_Race[Unknown or legacy]	-0.0435	0.009	-4.654	0.000
Perceived_Race[White]	0.0025	0.006	0.431	0.666
Age_group_at_arrest[age 17 and younger]	-0.0278	0.021	-1.304	0.192
Age_group_at_arrest[age 17 to 24 years]	0.0086	0.017	0.509	0.611
Age_group_at_arrest[age 25 to 34 years old]	0.0217	0.016	1.323	0.186
Age_group_at_arrest[age 35 to 44 years old]	0.0228	0.017	1.381	0.167
Age_group_at_arrest[age 45 to 54 years old]	0.0261	0.017	1.530	0.126
Age_group_at_arrest[age 55 to 64 years old]	0.0086	0.018	0.477	0.633
Age_group_at_arrest[age 65 and older]	0.0236	0.028	0.831	0.406
Age_group_at_arrest[age 65 years and older]	0.0263	0.027	0.967	0.334
StripSearch	0.1174	0.007	16.559	0.000

Note. Standard Errors assume that the covariance matrix of the errors is correc tly specified.

From this output, we can see that the R-squared value is 0.005, which means that about 0.05% of the variance in all_Actions can be explained by the independent variables. We can also see that both StripSearch and each Perceived_Race value have significant coefficients, with StripSearch having a coefficient of 0.1174. This indicates that StripSearch is positively associated with the number of actions taken during the stop. A regression coefficient represents the change in the dependent variable that is associated with a one-unit change in the independent variable while holding all other variables in the model constant. In this case, the dependent variable is all_Actions, which represents the total number of actions taken during the stop, and the two independent variables are StripSearch and Perceived_Race_value. A positive coefficient indicates a positive association between the independent variable and the dependent variable, while a negative coefficient indicates a negative association. In this case, we can see that most races value, age ranges other than the youth range of 17 and younger, and StripSearch has positive coefficients, which indicates that an increase in the value of either variable is associated with an increase in the number of actions taken during the stop.

In other words, the results suggest that when officers perform a strip search or perceive an individual as belonging to a particular race, they are more likely to take a higher number of actions during the stop. However, it's important to note that correlation does not imply causation, and there may be other factors at play that are contributing to the relationship between these variables and the number of actions taken during the stop.

F-statistics is a statistical measure that is often used to assess the significance of the differences between group means or the overall fit of a regression model. When conducting an analysis of variance (ANOVA) or regression analysis, the F-statistic is calculated as the ratio of the between-group or regression sum of squares to the within-group or residual sum of squares. To analyze the F-statistic, you will typically compare the computed value of F to the critical value of F at a given significance level (e.g., alpha = .05). If the computed value of F is greater than the critical value of F, then you can reject the null hypothesis and conclude that there is a significant difference between groups or that the regression model provides a good fit to the data. In addition to interpreting the F-value itself, it is also important to examine the associated degrees of freedom and p-value. The degrees of freedom reflect the number of observations and parameters in the analysis and are used to calculate the critical value of F. The p-value reflects the probability of observing the computed F-value or a more extreme value if the null hypothesis were true. A p-value less than .05 indicates that the computed F-value is statistically significant at the 5% level. We can make assumption that

The variance of the dependent variable is constant across different values of the independent variables.

Overall, the F-statistic is a useful tool for evaluating the significance of group differences or regression models and can provide valuable information about the relationships between variables in your data. Let's say we have conducted an ANOVA analysis and obtained an F-statistic of 37.65 with a p-value of 2.17e-67. This F-statistic tells us the ratio of the variance between the groups to the variance within the groups. A high F-value indicates that the variance between the groups is greater than the variance within the groups, which suggests that in the table1 all groups are significantly different from each other.

The F-statistic tests the null hypothesis that all of the regression coefficients in the model are equal to zero. A large F-statistic (i.e., a value much larger than 1) suggests that the variation in the dependent variable explained by the model is greater than would be expected by chance alone. The p-value associated with the F-statistic represents the probability of obtaining an F-statistic as large or larger than the observed value if the null hypothesis were true. In general, a p-value less than 0.05 is considered statistically significant and provides evidence against the null hypothesis. From the table above, we can conclude that the overall model is statistically significant and that the probability of obtaining an F-statistic as large or larger than 37.65 if the null hypothesis were true is only 2.17e-67. This suggests that we can reject the null hypothesis and conclude that at least one of the regression coefficients in the model is not equal to zero.

Table2

Results of ANCOVA comparing mean scores across three groups

ANCOVA Model Results				
Dep. Variable: all_Actions				
	Coefficient	Std err	t	P >|t|
Perceived_Race_value	0.0013	0.001	1.265	0.206
Age_group_at_arrest	-0.0009	0.001	-0.697	0.486
StripSearch	0.1164	0.007	19.502	0.000

When conducting a linear regression analysis, the standard error (std err), t-value, and p-value are used to assess the significance of the relationship between the independent variable(s) and the dependent variable. The standard error is a measure of the precision of the estimate of the regression coefficient. It represents the average distance that the sample estimates fall from the true population parameter. A smaller standard error indicates a more precise estimate. The t-value measures the number of standard errors the estimated coefficient is away from zero. It indicates whether the estimated coefficient is statistically significant or not. A larger absolute value of the t-value indicates a more significant relationship between the independent variable and the dependent variable. From Table 1, it is easy to find that the positive and negative coefficients are consistent with the positive and negative t-values. Among the variables in the test, only Perceived_Race[South Asian], Perceived_Race[unknown or legacy], and Age_group_at_arrest[age 17 and younger] have negative values, which indicates that they and the independent variable all_actions in which the relationship is not significant. The other variables, however, are more significantly related to the independent variables. The p-value (p(t)) is the probability of observing a t-value as extreme as the one calculated from the sample, assuming the null hypothesis is true. A p-value less than 0.05 is typically used as the threshold for determining statistical significance. If the p-value is less than 0.05, we can reject the null hypothesis and conclude that there is a significant relationship between the independent variable and the dependent variable. What's more we can use the std err, t-value, and p-value to assess the significance and precision of the regression coefficients and to make inferences about the relationship between the independent variable(s) and the dependent variable.

In summary, we can conclude that both Perceived_Race and StripSearch have a positive correlation coefficient, indicating a positive effect on the variables. However, for Age_group_at_arrest, the relationship with all_actionsd does not clearly indicate a positive effect on the variable.

Figure 5: *Power Test for Men and Strip Search*

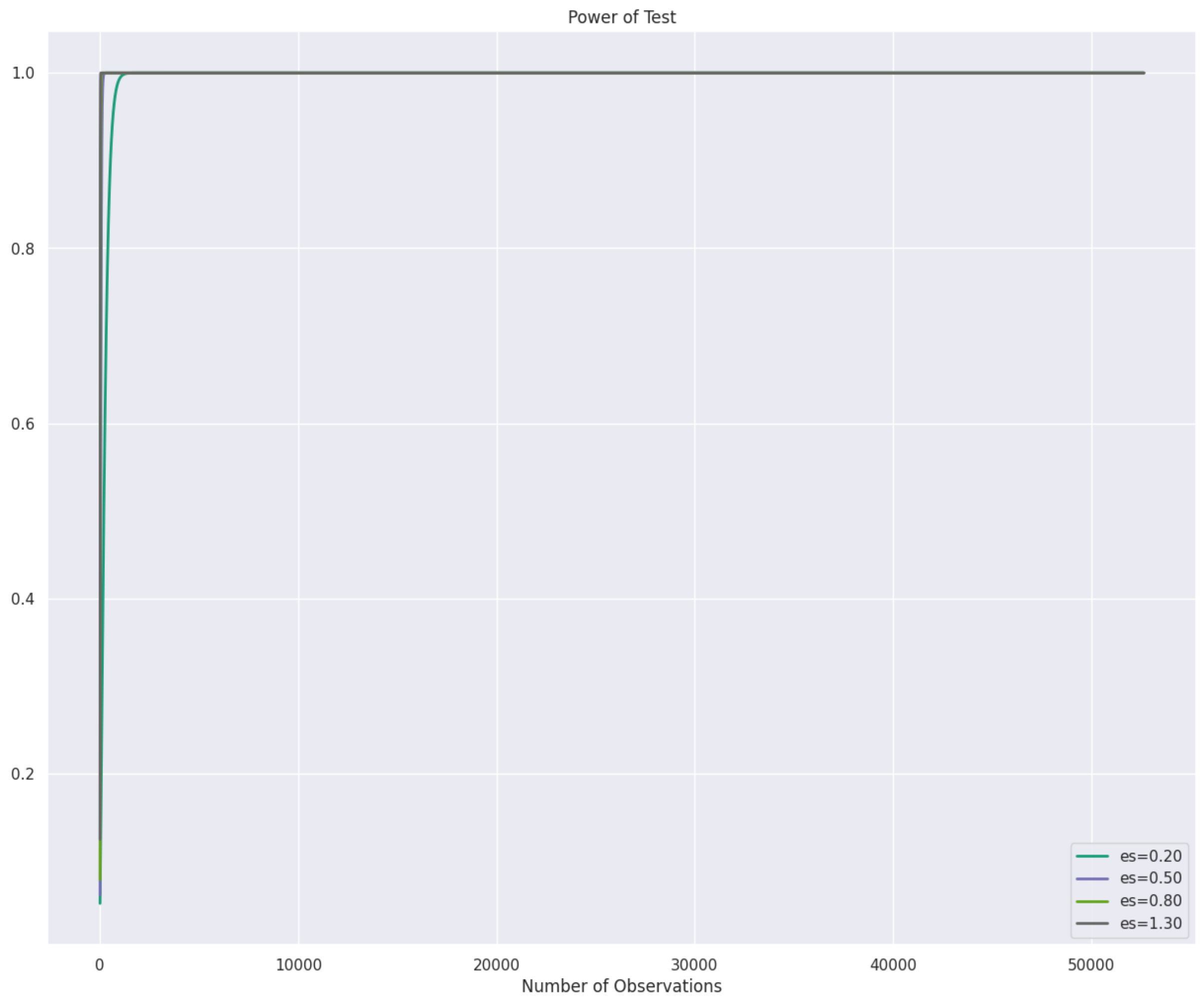
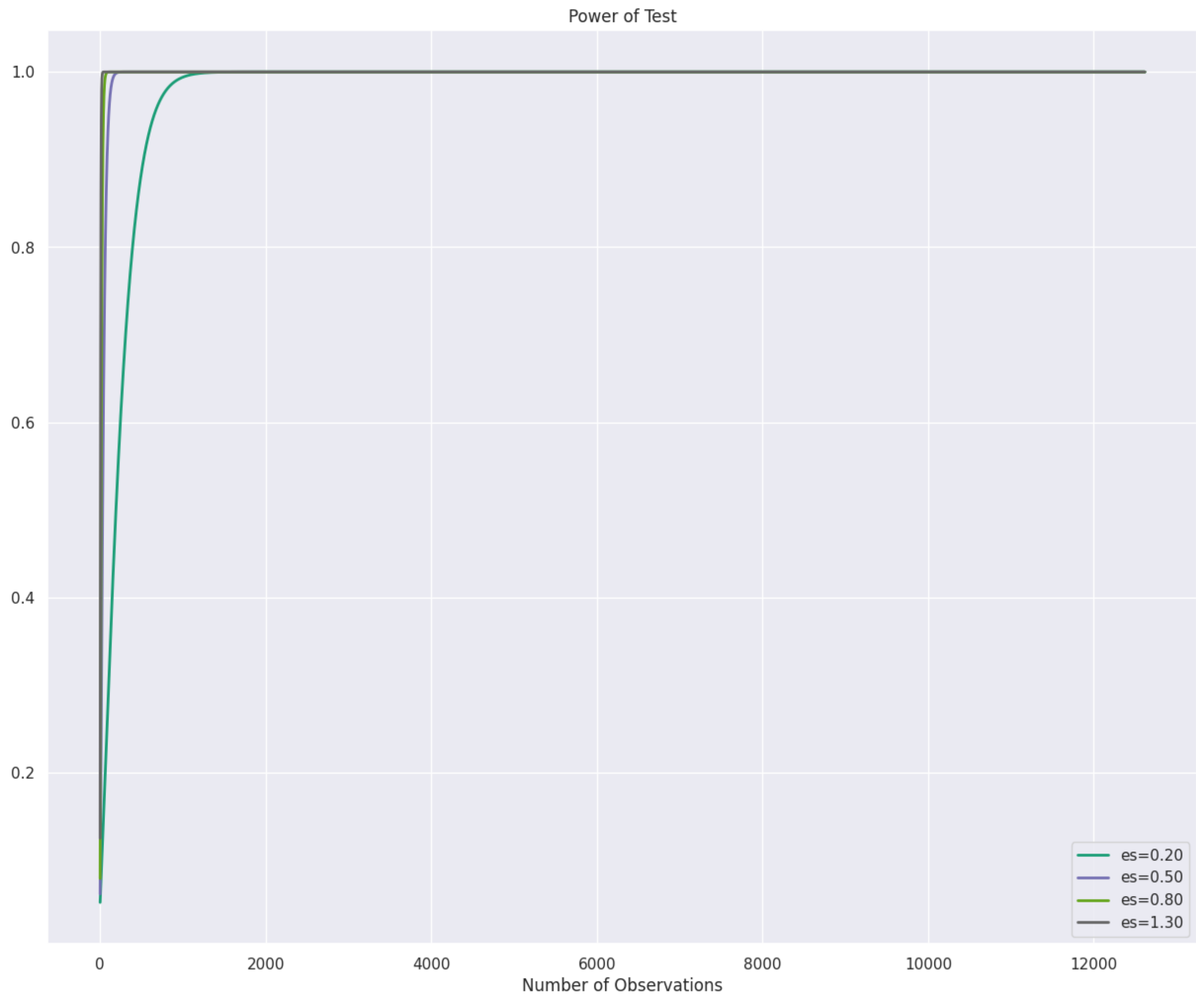


Figure 6: *Power test for women and strip search*



The Power analysis that was conducted on the variables mentioned in Methods found an effect size of 0.07, which indicates that the actual difference in means is very small between the two groups. Or that there was little difference in whether they were a man or a woman in how often they were strip searched by police. Next the number of sample sizes needed for that effect size were computed and compared to the actual sample size of the dataset. For men the sample needed to be 2093 people and for women the sample needed to be 8736 people for an effect size of 0.07, an alpha of 0.05 and a power of 80%. The actual size of those respective samples were 52650 men and 12617 which gave a power curve value of 1.0 for both the male and female sex groups. Lastly looking at Figures 5 and 6 for our selected variables and sample sizes we can conclude that our dataset has enough data to make reliable conclusions about how often men and women are strip searched.

Table 3: *Logistic Regression*

Variables	coeff,	P>|z|	Lower CI	Upper CI	Odds Ratio
Intercept	-2.3455	0	0.0864	0.1062	0.0958
Age_group__at_arrest_value	0.0713	0	1.0576	1.0904	1.0739
Percieved_Race_value	0.0115	0.089	0.9982	1.0251	1.0116
Sex_cat_val	-0.2298	0	0.7403	0.8531	0.7947

Table 4: *Confusion Matrix:*

11451	0
1605	0

Next looking at logistic regression, focusing on the odds ratio for each of the dependent variables there was an odds ratio of approximately 0.01 for the intercept, approximately 1.07 for the Age_group__at_arrest_value, approximately 1.01 for the Percieved_Race_value, and approximately 0.79 for the Sex_cat_val. This indicates for the Age_group__at_arrest_value variable that there is a slightly increased chance of being strip searched as you transition from age group to age group (or as you get older) but its closeness to 1 means it is likely not to a significant degree or that the odds are even with or without the dependent variable. For the Perceived_Race_value it is much the same as the previous variable, if you are not part of the Unknown race reference category it is still about even logs odds one way or another to be strip searched by the police with the log odds of 1.01 being close to 1. The sex_cat_val's log odds indicate that women have a less likely chance to be strip searched by the police compared to men because their log odds are less than 1. Lastly looking at the confusion matrix it gave no false positive or negative for our model and it had a test accuracy of approximately 88%. However it seems likely that this is inaccurate because of the generally smaller sample size of those people who were strip searched compared to the overall sample size of the dataset.

Table 5: *Prediction intervals*

Record #	Mean	Mean SE	Lower CI boudnary	Upper CI boundary	Prediction Interval lower boundary	Prediction Interval upper boundary
35295	0.1025	0.0033	0.0960	0.1090	-0.5301	0.7352
59786	0.1203	0.0016	0.1172	0.1235	-0.5123	0.7529
34813	0.1346	0.0021	0.1305	0.1388	-0.4980	0.7673
64564	0.1149	0.0037	0.1076	0.1222	-0.5177	0.7476
53112	0.1298	0.0027	0.1246	0.1350	-0.5028	0.7625

Lastly looking at Prediction intervals for 5 common Record # (row and column values) values from Table 5 the mean target values can range from approximately -0.5 to approximately 0.75 depending on the record. As well using this set of upper and lower prediction interval boundaries to compute the percentage of values in the test data that were within the prediction interval values gave the percentage of approximately 87% with an alpha of 0.05 (making it a 95% prediction interval).

Discussion

For our logistic regression results we found that our only significant result in terms of log odds was that being female meant that someone had a lesser chance of being stripped searched by the police. Otherwise the log odds were very close to 1 or even odds which meant that our other dependent variables had little effect on being strip searched. While our confusion matrix had no false positive or negative values it is likely because while there were sufficient sample sizes for men and women (according to our power analysis) the amount of those who were strip searched compared to those who were in the dataset overall was likely a much more uneven ratio. This sort of difficulty can be inferred from the notes attached to the dataset, specifically the note that someone may be included in the dataset without being booked or strip searched which could increase the ‘noise’ of the dataset. Lastly, our prediction interval was fairly good in terms of total accuracy being 87% but the aforementioned difficulties of our logistic regression still apply. After all despite its accuracy of 88% the lower percentage of those in the overall dataset who were strip searched likely

threw off the model causing it to predict that being strip searched was always significant regardless. This makes the accuracy of the prediction interval less useful overall, as well as the fact that there were categorical values used and there were few usable continuous variables. Ultimately much the same as we concluded with our midterm the general lack of continuous variables and especially useful ones mean that using statistical methods, without further data, on this data set is not very useful beyond deducing broad trends.
As mentioned earlier in our report, the main constraint of our study was the absence of continuous data in the dataset, which we have mentioned before. The process of cleaning the data was challenging, and we had reservations about the accuracy of specific numbers in our tests, which limited our analysis. To ensure the accuracy of our conclusions, it is advisable to examine the data and code more thoroughly for further study. Once the dataset has been cleaned, we will continue with deeper research.

Conclusion

By conducting ANCOVA, Power Tests, Logistic Regression, and Prediction intervals on the same dataset and variables that were analyzed in the midterm, we discovered that the Power analysis provided us with a sufficient amount of data to draw some sort of conclusion regarding the dataset. However the Logistic Regression and Prediction interval found that of our variables only sex (that is being female) had a significant effect on the log odds of being strip searched. As well as that despite the high accuracy of our model, the disproportionate amount of people being strip searched in the dataset overall and the lacking amount of continuous numerical data mean it is difficult to make confident conclusions from our model. In conclusion without more data, in terms of breadth and depth as well as more continuous variables, or a more complex set of models to analyze the data with it is hard to make any significant correlational or causational conclusions about this data set with confidence. However, the general trends that were captured with EDA and that were suggested by our models do imply exciting grounds for future data collection, entry, and analysis.

What's more, our research also focus on the relationship between police officer perceptions of a suspect's race and the number of actions taken during a police stop shows that there is a significant association between the two independent variables, StripSearch and Perceived_Race_value, and the dependent variable, all_Actions. The results suggest that when officers perform a strip search or perceive an individual as belonging to a particular race, they are more likely to take a higher number of actions during the stop. However, it is important to note that correlation does not imply causation, and there may be other factors at play that are contributing to the relationship between these variables and the number of actions taken during the stop. Overall, our study sheds light on the complex and multifaceted nature of police arrests and strip-searched and emphasizes the need for further research to better understand the factors that influence the actions taken by police officers during these encounters.

References:

Lum, Stoltz, Koper, & Scherer. (2019). *Research on body‐worn cameras - lum - wiley online library*. Research on body-worn cameras. Retrieved February 27, 2023, from https://onlinelibrary.wiley.com/doi/full/10.1111/1745-9133.12412

Wortley, Laniyonu, & Laming. (n.d.). *Use of force by the Toronto Police Service Final Report*. Use of force by the Toronto Police Service . Retrieved February 27, 2023, from https://falconers.ca/wp-content/uploads/2022/03/Use-of-force-by-the-Toronto-Police-Service-Final-report.pdf

Foster, & Jacobs. (2022). *Independent expert assessment report: Toronto Police Service Race ... - TPS*. https://www.tps.ca/media/filer_public/1f/e7/1fe76fd7-1bde-4c4d-bd90-091a5ccba5da/2bea48b6-40a7-4f43-ba92-ce59df850c73.pdf. Retrieved February 27, 2023, from https://www.tps.ca/media/filer_public/1f/e7/1fe76fd7-1bde-4c4d-bd90-091a5ccba5da/2bea48b6-40a7-4f43-ba92-ce59df850c73.pdf