Effect of Multiple Factors on the Level of Cooperation and Mental Instability at the Time of Arrest

Laien Zhou
Yuelin Liu

A final paper submitted in conformity with the requirements
for the course of INF2178

Faculty of Information

University of Toronto

Table of Contents

1 Introduction

1.1 Background

Arrest and strip search procedures are important components of the nowadays criminal justice system, designed to maintain public safety and uphold the law. However, studies have shown that these procedures can disproportionately affect certain groups, particularly those who belong to racial minorities (Engel 2003). In Canada, racial profiling has been identified as a pervasive issue in policing, with racialized individuals more likely to be arrested by police officers than non-racialized individuals (O'Neil 2020). The Toronto Police Service has recognised the need to address systematic prejudice within its ranks and has instituted a number of programmes targeted at decreasing racial disparities in enforcement (O'Neil 2020). However, we are not sure how effective the efforts have been in reducing the impact of race on the arrest and strip search process. In addition to race, there are other variables such as age, gender, and location(in Toronto or not) at the time of arrest that may also affect the level of cooperation of individuals.

Additionally, the issue of mental instability and its impact on the criminal justice system has been a topic of concern for many years. Mental instability has been identified as a potential factor affecting an individual's level of cooperation during the arrest and strip search process. Studies have shown that individuals with mental health issues are disproportionately represented in the criminal justice system, with many of them facing arrest (Educaloi, 2020). Furthermore, individuals with mental health issues may exhibit behaviors that could be perceived as non-compliant or aggressive during the arrest and strip search process, leading to potential harm or use of force by law enforcement officers. Therefore, understanding how the factors like race, sex, location at arrest and age group affects the presence of mental instability during the arrest process is crucial in developing better strategies to address the issue of over-representation of individuals with mental health issues in the criminal justice system and to ensure the safety of both law enforcement officers and individuals with mental health issues.

To address these issues, this study will both focus on examining the relationship between various factors and the level of cooperation at the time of the arrest, and investigating the relationship between the presence of mental instability and various factors during the arrest process. By analyzing data from Toronto Police Service, we aim to provide new insights into the factors that influence the arrest and strip search process, and to help develop better strategies to reduce incidents of force during arrests, which would ultimately lead to a safer and more effective community policing approach.

1.2 Literature Review

Kaminski and Sorensen (1995) discovered that police officers who encounter non-White suspects are more likely to be attacked. Engel (2003) also found that non-White suspects were more likely to resist arrest when confronted by White officers. Walker (1999) explained this phenomenon by stating, “To the extent that officers stereotype young African-American males as potential suspects, they may provoke higher rates of antagonistic behavior that, in turn, results in higher rates of arrest” (pp. 226-227). In addition to race, gender, and age were also examined as possible factors of the level of cooperation during an arrest. According to the study, men have a higher rate of criminal behavior compared to women and that offending decreases with age (U.S. Department of Justice 2009).

Also, Bierie (2015) discovered that female offenders were less likely to assault police personnel than male offenders. Besides, two studies have discovered evidence to support the expected age group impacts. According to McCluskey et al. (1999), suspects who looked to be under the age of 21 were more likely to fight capture. Bierie (2015) also discovered that confrontations between police and criminals involving at least one juvenile were more likely to result in an officer being assaulted. Furthermore, Perreault (2019) reported that police services in rural areas served 16% of the population in Canadian provinces in 2017 but reported a higher percentage of violent crimes, property crimes, Criminal Code traffic offenses, and other Criminal Code violations compared to urban areas.

While at the same time, the relationship between race, sex, location at arrest, age group, and mental instability has been a topic of interest in criminology and psychology research. Studies have shown that individuals from certain racial groups are more likely to have mental health issues, and this may lead to higher levels of police contact (McGuire & Miranda, 2008). Research has also shown that sex can be a significant predictor of mental health disorders, with women being more likely to be diagnosed with depression and anxiety disorders than men (American Psychological Association, 2011). Location at the time of arrest has also been identified as a potential factor affecting mental health outcomes. One study found that individuals who were arrested in urban areas had higher rates of mental health disorders and were more likely to have had prior mental health treatment compared to those arrested in rural areas (Litman, 2021). Finally, age group has been shown to be a significant predictor of mental health disorders, with younger individuals being more likely to experience mental health issues such as depression and anxiety (Jurewicz, 2015).

1.3 Research Questions/Objective

The main objective of this study is to investigate the factors that influence the level of cooperation between individuals and police officers as well as the mental instability of the individuals at the time of the arrest. Specifically, we aim to answer the following 2 research questions:

1. Does any of ***perceived race, sex, location in Toronto or not*** *at arrest or* ***age group*** effects on the level of cooperation at the time of arrest controlled by ***threat index***?
2. How does ***perceived race***, ***sex***, ***location in Toronto or not*** at arrest and ***age group*** affect the probability of the individuals' mental instability at the time of the arrest?

2 Exploratory data analysis

2.1 Overview

The dataset used in this study is the "Arrests and Strip Searches" dataset (RBDC-ARR-TBL-001) from the Toronto Police Service Public Safety Data Portal. It provides relevant information about arrests and strip searches and is conducted by police officers within Toronto. A strip search is defined as a search that involves the removal of clothing and a visual examination of the body. The dataset includes 65,276 records, with 24 attributes including the year and month of arrest, the ID of arrest, the race, gender, age group of the suspects, the arrest location, and the actions taken during an arrest, etc. The location of the arrest is recorded at the Division level and refers to the area within the Division boundaries where the arrest took place. The age is the individual's age at the time of the arrest, as reported by the arresting officer.

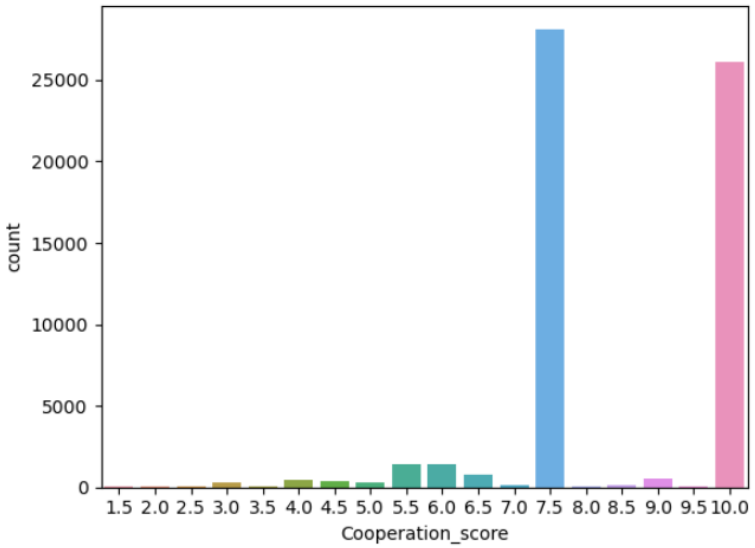
2.2 Descriptive statistics

For the problem to be studied in this paper, the dataset underwent the following data cleaning process. To begin with, the samples with a "U"(Uncertain) attribute under the *"Sex"* category were removed as their count was relatively small in comparison to other categories and would not be statistically significant. Additionally, the samples with an "Unknown or Legacy" attribute under the *"Perceived_Race"* category were removed. The category *'ArrestLocDiv'* has been renamed to *'intoronto'*, while we set its value to "1" for rows where it is not equal to "XX" and to "0" for rows where it is equal to "XX". This is done to represent whether an arrest location is in Toronto or not. For the *'Age_group__at_arrest_'* category, we reassigned the original age group values to a smaller set of standardized categories to make data analysis and interpretation easier. After that, we created a new continuous variable *"Cooperation_score"* (defined in the *Method* section). Another continuous variable is 'threat_index' (defined in the Method section as well). Lastly, all the samples containing "N/A" values were purged from the dataset to ensure accurate statistical calculations.

2.2.1 Cooperation Scores vs Independent variables

The figure below shows the distribution of **Cooperation Scores** (defined in the *Method* section) in the sample. Approximately half of the ratings are concentrated at a score of 7.5, with close to half but fewer concentrated at a score of 10. The remaining ratings have a similar normal distribution structure with centers around 5.5-6 and 9 for scores below 7.5 and between 7.5-10, respectively. These findings indicate that the majority of individuals in the dataset were willing to cooperate, as a score of 7.5 implies that the individuals did not show explicit cooperative behavior but also no obvious resistance at the time of arrest, while a score of 10 implies that the individuals cooperate fully with the police at the time of the arrest.

Figure 1: Distribution of the Cooperation Scores



The four box plots below show the distribution of the dependent variables (*Cooperation_Scores*) when grouped by 4 different independent variables, respectively:

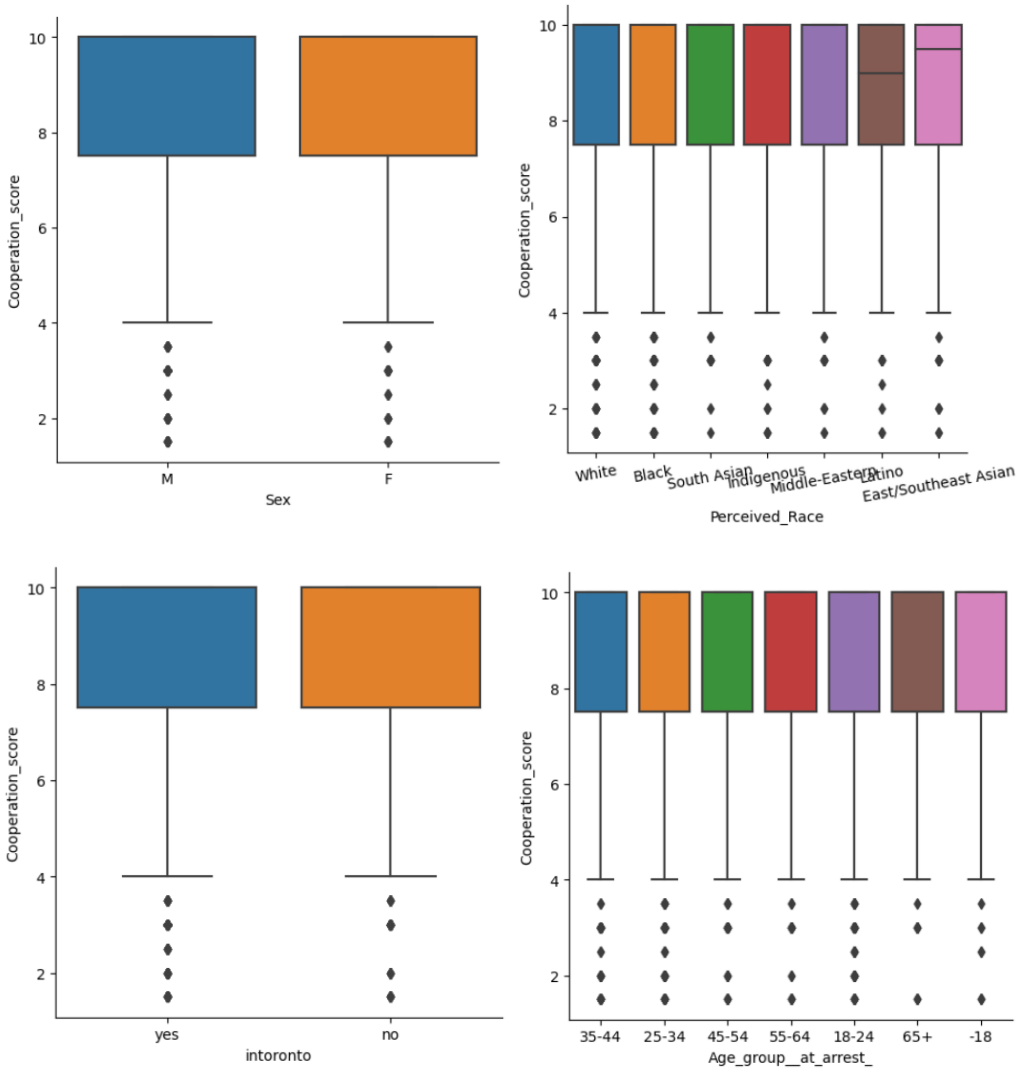
- In the top left plot, the distribution of cooperation scores for individuals of different genders is presented. It can be observed that the boxes for both genders are concentrated in the interval of 7.5 to 10 points, which corresponds to the main concentration of Cooperation scores in Figure 1. The cases with scores below 4 points for both genders are classified as outliers.

- The box plot in the upper right corner displays the distribution of cooperation scores across various races. The boxes for all races are concentrated in the range of 7.5 to 10 points, with cases below 4 points classified as outliers. However, there are noticeable differences in the mean cooperation scores among different race groups. Latino individuals have a lower mean cooperation score compared to other races, while individuals from East/Southeast Asian backgrounds have a higher mean cooperation score than other races. Moreover, the distribution of outliers for cooperation scores varies among different race groups.

- The box plot in the lower left corner displays the distribution of cooperation scores for individuals who were in the Toronto area at the time of arrest and those who were not. Both boxes are concentrated within the 7.5-10 point range, but the distribution of outliers (under a score of 4) for both groups is different. Outliers for individuals not in the Toronto area are relatively more widely spread than those for individuals in the Toronto area.

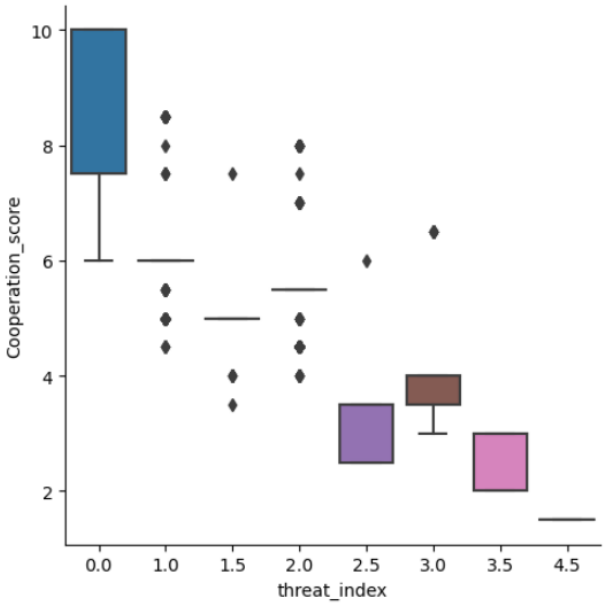
- The bottom right box plot displays the distribution of cooperation scores for individuals in different age groups. The boxes for all age groups are mainly distributed in the range of 7.5-10 points, with scores below 4 points being categorized as outliers. However, the distribution of outliers varies among the age groups, with the greatest distance between outliers observed in individuals over 65 years old, and a relatively distant distribution of outliers in those aged 45-64 years old.

Figure 2: Boxplots of Sex, Perceived_Race, Intoronto (location), and Age_group_at_arrest with Cooperation_Score



The box plot below shows the distribution of the dependent variable (*Cooperation_Scores*) when grouped by the independent variable *threat_index*. It can be observed that approximately half of observations are centered at 0, and the majority of the other half observations are located among [2.5, 3.5]. Most of the cases with threat index among [1, 2] and above 3.5 are classified as outliers.

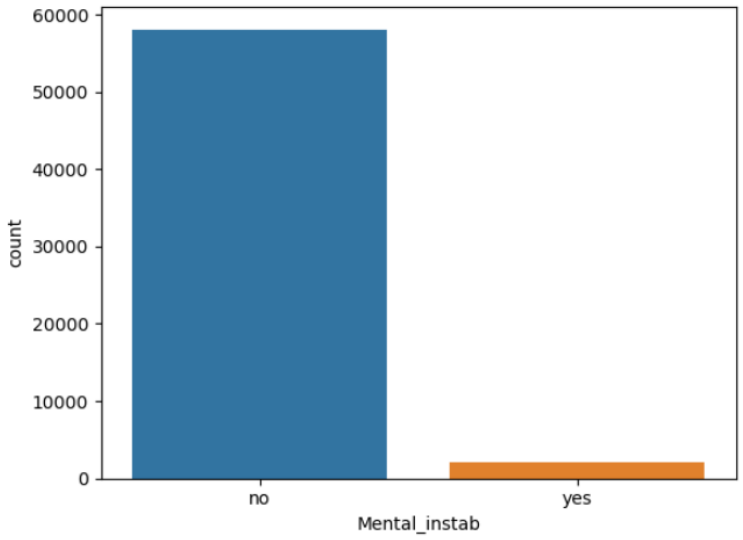
Figure 3: Boxplots of threat_index (control variable) with Cooperation_Score



2.2.2 Mental Instability at arrest vs Independent variables

The figure below shows the distribution of **Mental Instability** (defined in the *Method* section) in the sample. The majority of observations do not exhibit mental instability at the time of arrest, while only a small proportion of observations display mental instability at the time of arrest.

Figure 4: Distribution of the Mental Instability



The four barplots below displays the distribution of the dependent variable (*Mental Instability*) when grouped by 4 different independent variables, respectively:

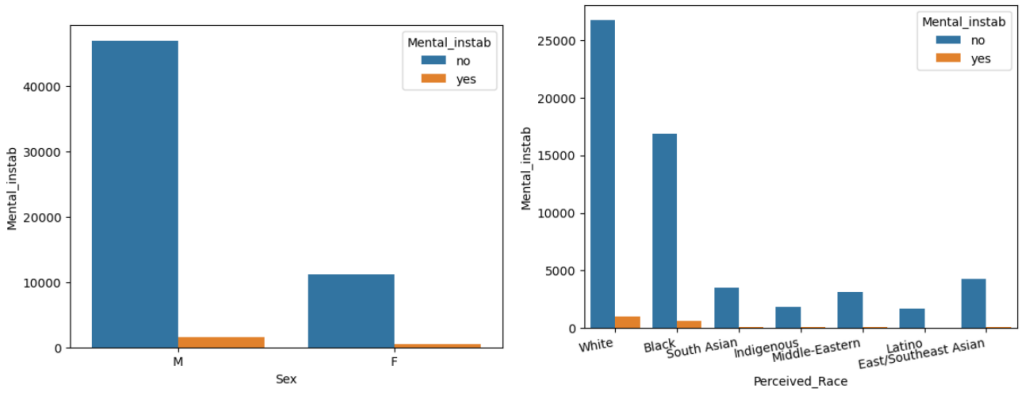
- In the top left plot, the distribution of mental instability for individuals of different genders is presented. It can be observed that for both genders the majority of the observations does not show the mental instability at arrest, which corresponds to the distribution of mental instability in Figure 4.

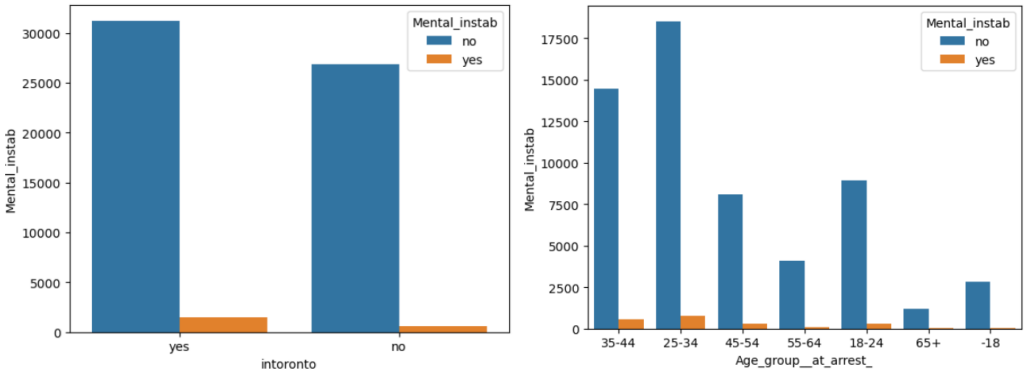
- The bar plot in the upper right corner displays the distribution of mental instability across various races. Each of the 7 categories of race shows a pattern that most observations in the specific group do not show the mental instability. However, the percentage of individuals showing mental instability among the white group and black group is much higher than that in other groups.

- The bar plot in the lower left corner displays the distribution of mental instability for individuals who were in the Toronto area at the time of arrest and those who were not. The two categories show the same pattern that the majority of observations inside the group does not show the mental instability at arrest, while the percentage of individuals showing mental instability in the "inside toronto" group is higher than that in the other group.

- The bottom right bar plot displays the distribution of mental instability for individuals in different age groups. Each of the age groups show the same pattern that most individuals inside the group does not show the mental instability ar arrest, while the percentage of individuals showing mental instability in Young and middle-aged (25-54) are obviously higher than that in other age groups.

Figure 5: Barplots of Sex, Perceived_Race, Intoronto (location), and Age_group_at_arrest_ with Mental_Instability





2.3 Power analysis

2.3.1 Power tests on independent variables vs Cooperation Score

1) Sex vs Cooperation Score

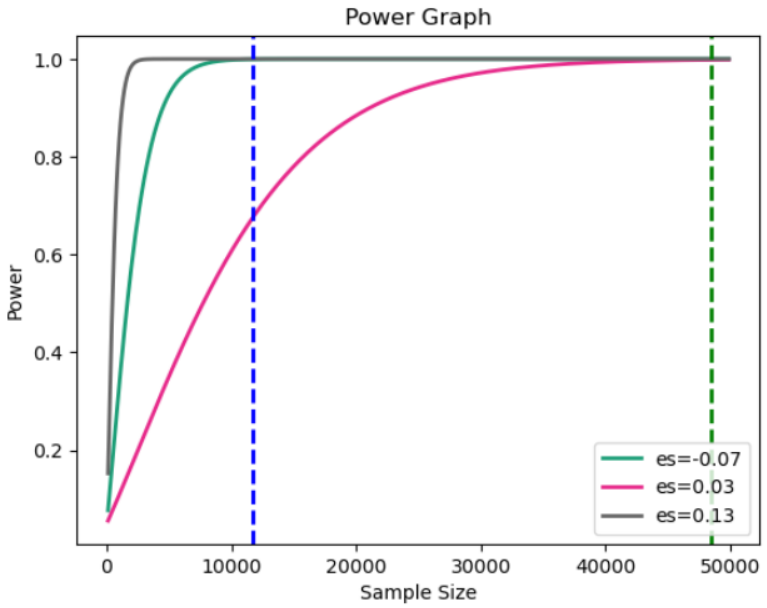
Table 1: Power Tests on Sex vs Cooperation score

	Cooperation score	
	Sample size needed	Actual size
male	9774.166	48505
female	40583.453	11682
Effect size (Cohen's D)	0.0315	

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Prior to computing a t-test to analyze whether an individual's cooperation score (outcome variable) differed between sex, male and female (two-level explanatory variable), we calculated the effect size of the explanatory variable using Cohen's D metric, which was 0.0315. The goal of the analysis is to determine the sample size needed to detect a significant effect at a power of 0.8 and a significance level of 0.05.
- After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. According to the results, the sample size needed for male scores is 9774, and the sample size needed for female scores is 40583. It is PARTIALLY SIGNIFICANT since the sample size provided in the dataset are 48505 and 11682, respectively, which impacts the reliability of the results. This means that the actual sample sizes are already larger than the required sample sizes, which is good for increasing the power of the analysis.
- This power analysis suggests that the sample sizes for male and female scores are adequate to detect a small effect size at a power of 0.8 and a significance level of 0.05.

Figure 6: Power graph of Sex with Cooperation score



2) Intoronto vs Cooperation Score

Table 2: Power tests on Intoronto vs Cooperation score

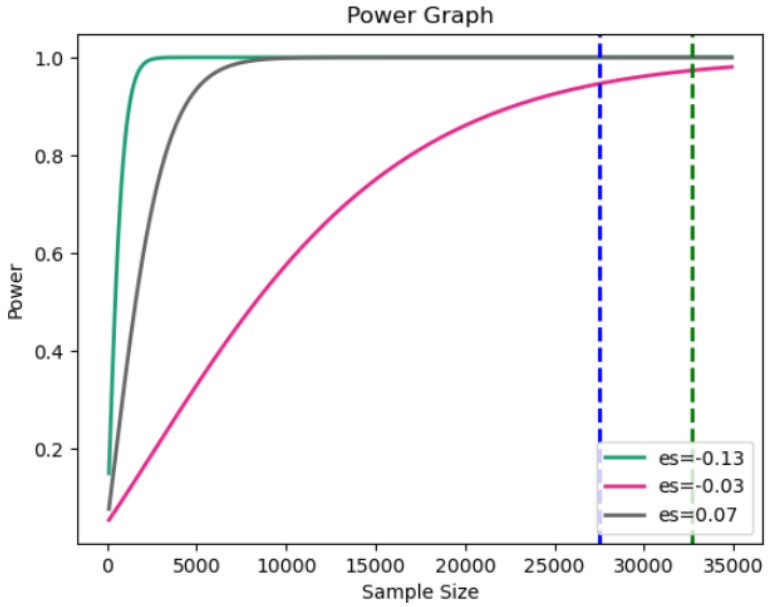
	Cooperation score	
	Sample size needed	Actual size
inside toronto	15593.145	32694
outside toronto	18542.985	27493
Effect size (Cohen's D)	-0.0304	

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- The goal of the analysis is to determine the sample size needed to detect a significant effect at a power of 0.8 and a significance level of 0.05.
- Prior to computing a t-test to analyze whether an individual's cooperation score (outcome variable) differed between arrested inside toronto and arrested outside toronto (two-level explanatory variable), we calculated the effect size of the explanatory variable using Cohen's D metric, which was -0.0304.
- After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. The results indicated that a sample size of 15593 was required for individuals inside toronto, while a sample size of 18542 was required for individuals outside toronto. It is SIGNIFICANT since the actual sample size provided in the dataset are 32694 and 27493, respectively, which impacts the reliability of the

results. This means that the actual sample sizes are larger than the required sample sizes, which is good for increasing the power of the analysis.

Figure 7: Power graph of Intoronto with Cooperation score



3) Perceived_Race vs Cooperation Score

Table 3: Power tests on Perceived_Race vs Cooperation score

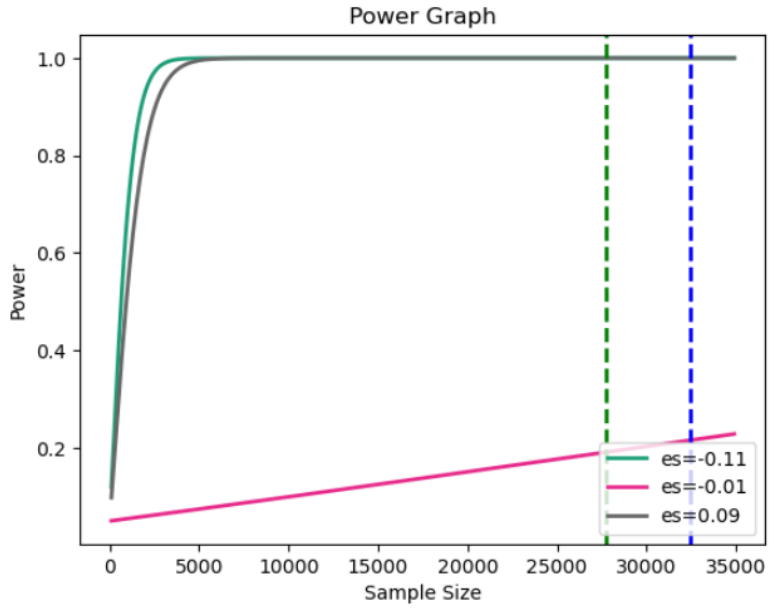
	Cooperation score	
	Sample size needed	Actual size
white	201847.230	27708
not white	172196.898	32479
Effect size (Cohen's D)	-0.0091	

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- The aim of the analysis is to determine the sample size needed to detect a significant effect at a power of 0.8 and a significance level of 0.05.
- Prior to computing a t-test to analyze whether an individual's cooperation score (outcome variable) differed between white and non-white individuals (two-level explanatory variable), we calculated the effect size of the explanatory variable using Cohen's D metric, which was -0.0091.
- After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. The results indicated that a sample size of 201847 was required for white individuals, while a sample size of 172196 was

required for non-white individuals. It is NOT SIGNIFICANT since the actual sample size provided in the dataset are 27708 and 32479, which impacts the reliability of the results.

Figure 8: Power graph of Perceived race with Cooperation score



4) Age_group__at_arrest vs Cooperation Score

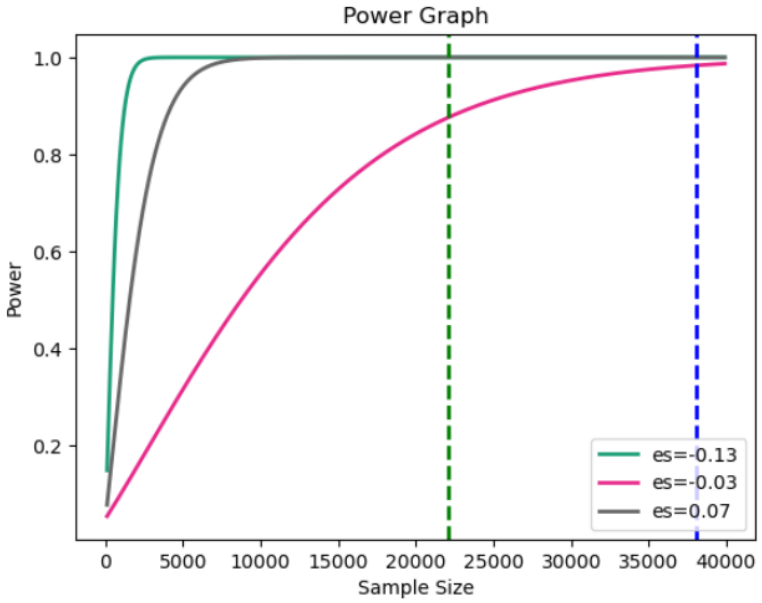
Table 4: Power tests on Age group vs Cooperation score

Cooperation score		
	Sample size needed	Actual size
< 35	24281.021	22124
≥ 35	14113.268	38063
Effect size (Cohen's D)	-0.0296	

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Prior to computing a t-test to analyze whether an individual's cooperation score (outcome variable) differed between individuals less than 35 and equal or older than 35 (two-level explanatory variable), we calculated the effect size of the explanatory variable using Cohen's D metric, which was -0.0296.
- After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. The results indicated that a sample size of 24281 was required for individuals less than 35, while a sample size of 14113 was required for individuals equal or older than 35. It is SIGNIFICANT since the sample size provided in the dataset are 22124 (close to the required nob) and 38063, which impacts the reliability of the results.

Figure 9: Power graph of Age group with Cooperation score



2.3.2 Power tests on independent variables vs Mental Instability

1) Sex vs Mental instability

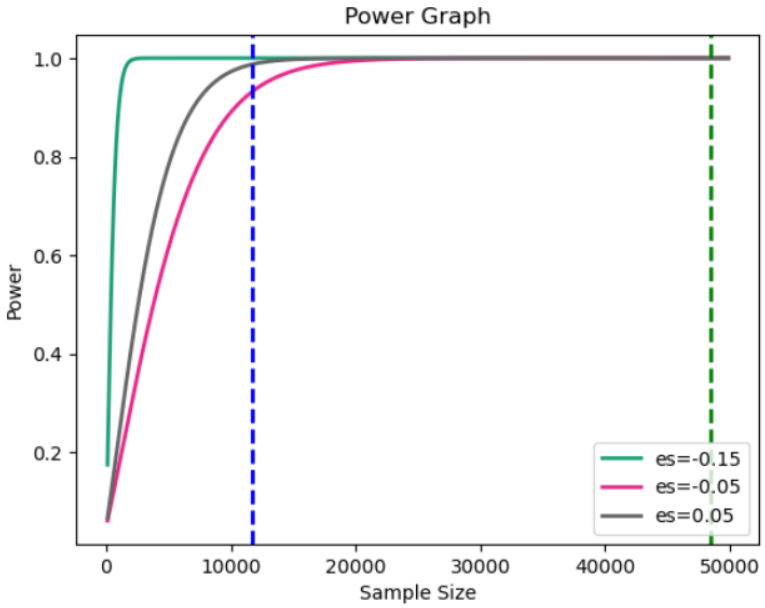
Table 5: Power tests on Sex vs Mental Instability

	Mental Instability	
	Sample size needed	Actual size
male	4787.451	48505
female	19878.043	11682
Effect size (Cohen's D)	-0.0451	

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Prior to computing a t-test to analyze whether an individual's Mental Instability (outcome variable) differed between male and female (two-level explanatory variable), we calculated the effect size of the explanatory variable using Cohen's D metric, which was -0.0451.
- After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. The results indicated that a sample size of 4787 was required for male, while a sample size of 19878 was required for female. It is PARTIALLY SIGNIFICANT since the sample size provided in the dataset are 48505 and 11682, which impacts the reliability of the results.

Figure 10: Power graph of Sex with Mental Instability



2) Intoronto vs Mental instability

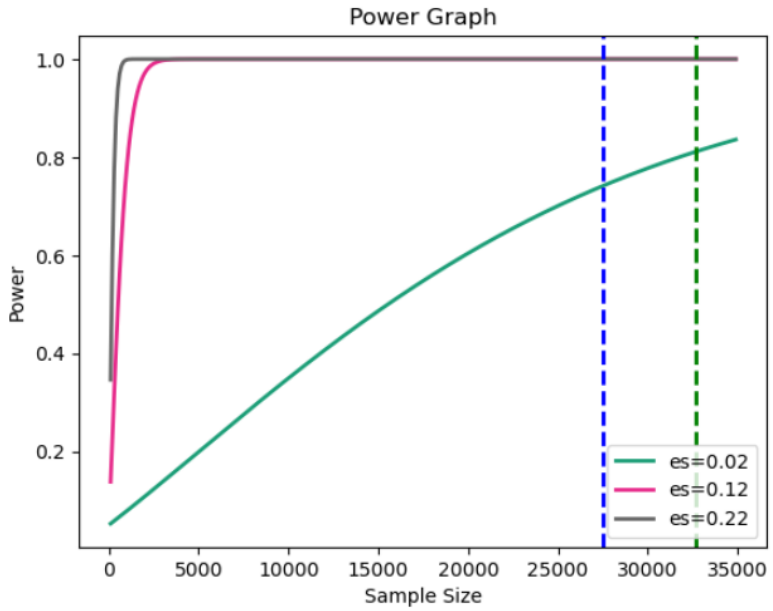
Table 6: Power tests on Intoronto vs Mental Instability

	Mental Instability	
	Sample size needed	Actual size
inside toronto	967.836	32694
outside toronto	1150.927	27493
Effect size (Cohen's D)	0.1222	

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Prior to computing a t-test to analyze whether an individual's Mental Instability (outcome variable) differed between individuals inside toronto and outside toronto (two-level explanatory variable), we calculated the effect size of the explanatory variable using Cohen's D metric, which was 0.1222.
- After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. The results indicated that a sample size of 967 was required for individuals inside toronto, while a sample size of 1150 was required for individuals outside toronto. It is SIGNIFICANT since the sample size provided in the dataset are 32694 and 27493, which impacts the reliability of the results.

Figure 11: Power graph of Intoronto with Mental Instability



3) Perceived_Race vs Mental instability

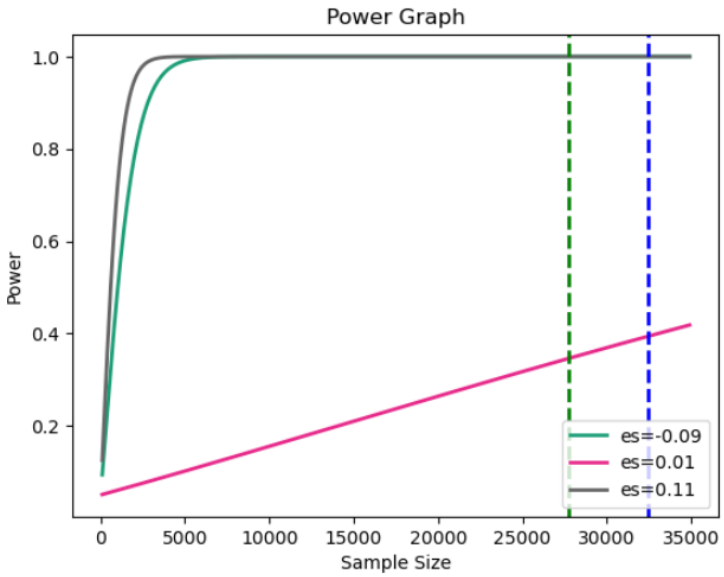
Table 7: Power tests on Perceived race vs Mental Instability

	Mental Instability	
	Sample size needed	Actual size
white	96763.585	27708
not white	82549.506	32479
Effect size (Cohen's D)	0.0132	

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Prior to computing a t-test to analyze whether an individual's Mental Instability (outcome variable) differed between white and non-white individuals (two-level explanatory variable), we calculated the effect size of the explanatory variable using Cohen's D metric, which was 0.0132.
- After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. The results indicated that a sample size of 96763 was required for white individuals, while a sample size of 82549 was required for non_white individuals. It is NOT SIGNIFICANT since the sample size provided in the dataset are 27708 and 32479, which impacts the reliability of the results.

Figure 11: Power graph of Perceived race with Mental Instability



4) Age_group__at_arrest vs Mental instability

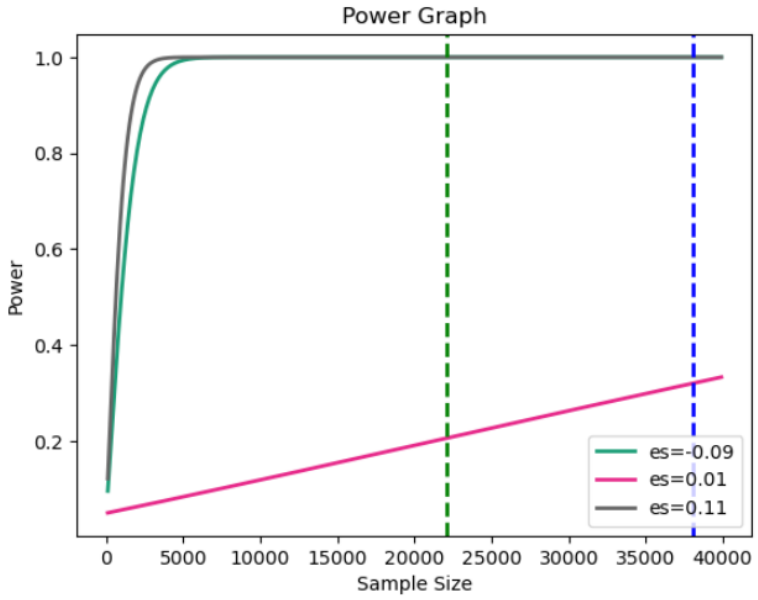
Table 8: Power tests on Age group vs Mental Instability

	Mental Instability	
	Sample size needed	Actual size
< 35	182267.661	22124
≥ 35	105942.509	38063
Effect size (Cohen's D)	0.0108	

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Prior to computing a t-test to analyze whether an individual's Mental Instability (outcome variable) differed between individuals less than 35 and equal or older than 35 (two-level explanatory variable), we calculated the effect size of the explanatory variable using Cohen's D metric, which was 0.0108.
- After obtaining the effect size, the required sample size was computed using the obtained effect size and establishing the statistical power at 80%. The results indicated that a sample size of 182267 was required for individuals less than 35, while a sample size of 105942 was required for individuals equal or older than 35. It is NOT SIGNIFICANT since the sample size provided in the dataset are 22124 and 38063, which impacts the reliability of the results.

Figure 12: Power graph of Age group with Mental Instability



2.4 T-Tests

T-test is used to compare the means of two groups for each factor, including 1) Sex, Perceived_Race, Intoronto (location), Age_group_at_arrest_ and threat_index vs Cooperation Score and 2) Sex, Perceived_Race, Intoronto (location), and Age_group_at_arrest_ vs Mental Instability.

Prior to conducting T-tests, the assumptions of the T-tests have been checked. The assumptions of T-test includes three aspects:

1) **Independent samples**: The two samples being compared should be independent of each other.
2) **Normality**: The data in each group should be approximately normally distributed.
3) **Homogeneity of variances**: The variances of the two groups should be approximately equal.

Assumption 1 is guaranteed by our experimental design, assumption 2 and 3 are checked prior to the T-test but not presented here based on the length of the paper.

2.4.1 T-tests on independent variables vs Cooperation Score

1) Sex vs Cooperation Score

- H0 (null hypothesis): There is no difference in cooperation score for male and female.
- Ha (alternative hypothesis): There is a significant difference in cooperation scores for male and female.

Table 9: T-test on Sex vs Cooperation score

Cooperation score			
	Mean		SD
male	8.4239		1.5548
female	8.3748		1.5643
satatistic	3.0514	p-value	0.0023
CI	(0.0176, 0.0807)	DoF	17659.60

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Interpretation: The results indicate that the mean cooperation score for male (M=8.4239, SD=1.5548) is higher than the mean cooperation score for female (M=8.3748, SD=1.5643). With alpha established at 0.05, this is a statistically significant difference as the p-value (0.0023) is less than 0.05, 95% CI [0.0176, 0.0807].
- Therefore, we can REJECT the null hypothesis that there is no difference in cooperation scores for male and female.

2) Intoronto vs Cooperation Score

- H0 (null hypothesis): There is no difference in cooperation score for individuals arrested inside Toronto and outside Toronto.
- Ha (alternative hypothesis): There is a significant difference in cooperation scores for individuals arrested inside Toronto and outside Toronto.

Table 10: T-test on In_toronto vs Cooperation score

Cooperation score			
	Mean		SD
inside toronto	8.3928		1.6443
outside toronto	8.4401		1.4453
satatistic	-3.7616	p-value	0.0002
CI	(-0.0721, -0.0227)	DoF	60067.41

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Interpretation: The results indicate that the mean cooperation score for individuals arrested inside Toronto (M=8.3928, SD=1.6443) is lower than the mean cooperation score for individuals arrested inside Toronto (M=8.4401, SD=1.4453). With alpha established at 0.05, this is a statistically significant difference as the p-value (0.0002) is less than 0.05, 95% CI [-0.0721, -0.0227].

- Therefore, we can REJECT the null hypothesis that there is no difference in cooperation score for individuals arrested inside Toronto and outside Toronto.

3) Perceived_Race vs Cooperation Score

- H0: There is no difference in cooperation score for white and non-white individuals.
- Ha: There is a significant difference in cooperation scores for white and non-white individuals.

Table 11: T-test on Perceived race vs Cooperation score

	Cooperation score		
	Mean	SD	
white	8.4067	1.5566	
not white	8.4210	1.5569	
satatistic	-1.1238	p-value	0.2611
CI	(-0.0393, 0.0106)	DoF	58703.24

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Interpretation: The results indicate that the mean cooperation score for white individuals (M=8.4067, SD=1.5566) is lower than the mean cooperation score for non-white individuals (M=8.4210, SD=1.5569). With alpha established at 0.05, this is NOT a statistically significant difference as the p-value (0.2611) is greater than 0.05, 95% CI [-0.0393, 0.0106].
- Therefore, we CANNOT reject the null hypothesis that there is no difference in cooperation score for white and non-white individuals.

4) Age_group__at_arrest vs Cooperation Score

- H0 (null hypothesis): There is no difference in cooperation score for individuals younger than 35 and older or equal to 35.
- Ha (alternative hypothesis): There is a significant difference in cooperation score for individuals younger than 35 and older or equal to 35.

Table 12: T-test on Age group vs Cooperation score

	Cooperation score	
	Mean	SD
< 35	8.3852	1.5650

≥ 35	8.4314		1.5517
satatistic	-3.4998	p-value	0.0005
CI	(-0.0720, -0.0203)	DoF	45914.73

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Interpretation: The results indicate that the mean cooperation score for individuals younger than 35 (M=8.3852, SD=1.5650) is lower than the mean cooperation score for individuals older or equal to 35 (M=8.4314, SD=1.5517). With alpha established at 0.05, this is a statistically significant difference as the p-value (0.0005) is less than 0.05, 95% CI [-0.0720, -0.0203].
- Therefore, we can REJECT the null hypothesis that there is no difference in cooperation score for individuals younger than 35 and older or equal to 35.

5) threat_index (control variable) vs Cooperation Score

- H0 (null hypothesis): There is no difference in cooperation score for individuals with a threat index greater than 2 and equal or less than 2.
- Ha (alternative hypothesis): There is a significant difference in cooperation score for individuals with a threat index greater than 2 and equal or less than 2.

Table 13: T-test on Threat index vs Cooperation score

Cooperation score			
	Mean		SD
≤ 2	0.0920		0.3832
＞2	3.2205		0.4631
satatistic	-203.1536	p-value	0.0000
CI	(-3.1587, -3.0983)	DoF	932.3768

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Interpretation: The results indicate that the mean cooperation score for individuals with a threat index greater than 2 (M=3.2205, SD=0.4631) is greater than the mean cooperation score for individuals equal or less than 2 (M=0.0920, SD=0.3832). With alpha established at 0.05, this is a statistically significant difference as the p-value (0.0000) is less than 0.05, 95% CI [-3.1587, -3.0983].

- Therefore, we can REJECT the null hypothesis that there is no difference in cooperation score for individuals with a threat index greater than 2 and equal or less than 2.

2.4.2 T-tests on independent variables vs Mental Instability

1) Sex vs Mental instability

- H0 (null hypothesis): There is no difference in mental instability probability for male and female.
- Ha (alternative hypothesis): There is a significant difference in mental instability probability for male and female.

Table 14: T-test on Sex vs Mental instability

	Mental instability		
	Mean	SD	
male	0.0328	0.1782	
female	0.0410	0.1985	
satatistic	-4.0986	p-value	0.0004
CI	(-0.0121, -0.0042)	DoF	16510.14

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Interpretation: The results indicate that the mean mental instability probability for male (M=0.0328, SD=0.1782) is higher than the mean mental instability probability for female (M=0.0410, SD=0.1985). With alpha established at 0.05, this is a statistically significant difference as the p-value (0.0004) is less than 0.05, 95% CI [-0.0121, -0.0042].
- Therefore, we can REJECT the null hypothesis that there is no difference in mental instability probability for male and female.

2) Intoronto vs Mental instability

- H0 (null hypothesis): There is no difference in mental instability probability for individuals arrested inside Toronto and outside Toronto.
- Ha (alternative hypothesis): There is a significant difference in mental instability probability for individuals arrested inside Toronto and outside Toronto.

Table 15: T-test on In_toronto vs Mental instability

	Mental instability

	Mean		SD
inside toronto	0.0446		0.2064
outside toronto	0.0223		0.1478
satatistic	15.3601	p-value	0.00001
CI	(0.0194, 0.0250)	DoF	58747.61

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Interpretation: The results indicate that the mean mental instability probability for individuals arrested inside Toronto (M=0.0446, SD=0.2064) is lower than the mean mental instability probability for individuals arrested inside Toronto (M=0.0223, SD=0.1478). With alpha established at 0.05, this is a statistically significant difference as the p-value (0.00001) is less than 0.05, 95% CI [0.0194, 0.0250].
- Therefore, we can REJECT the null hypothesis that there is no difference in mental instability probability for individuals arrested inside Toronto and outside Toronto.

3) Perceived_Race vs Mental instability

- H0: There is no difference in mental instability probability for white and non-white individuals.
- Ha: There is a significant difference in mental instability probability for white and non-white individuals.

Table 16: T-test on Perceived race vs Mental instability

Mental instability			
	Mean		SD
white	0.0357		0.0333
not white	0.1857		0.1795
satatistic	1.6187	p-value	0.1054
CI	(-0.0005, 0.0053)	DoF	58031.65

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Interpretation: The results indicate that the mean mental instability probability for white individuals (M=0.0357, SD=0.0333) is lower than the mean mental instability probability for non-white individuals (M=0.1857, SD=0.1795). With alpha established at 0.05, this is NOT a

statistically significant difference as the p-value (0.1054) is greater than 0.05, 95% CI [-0.0005, 0.0053].

- Therefore, we CANNOT reject the null hypothesis that there is no difference in mental instability for white and non-white individuals.

4) Age_group__at_arrest vs Mental instability

- H0 (null hypothesis): There is no difference in mental instability probability for individuals younger than 35 and older or equal to 35.
- Ha (alternative hypothesis): There is a significant difference in mental instability probability for individuals younger than 35 and older or equal to 35.

Table 17: T-test on Age group vs Mental instability

	Mental instability		
	Mean	SD	
< 35	0.0357	0.1855	
≥ 35	0.0337	0.1805	
satatistic	1.2710	p-value	0.2037
CI	(-0.0010, -0.0203)	DoF	45210.32

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

- Interpretation: The results indicate that the mean mental instability probability for individuals younger than 35 (M=0.0357, SD=0.1855) is lower than the mean mental instability probability for individuals older or equal to 35 (M=0.0337, SD=0.1805). With alpha established at 0.05, this is a statistically significant difference as the p-value (0.2037) is greater than 0.05, 95% CI [-0.0010, -0.0203].
- Therefore, we CANNOT reject the null hypothesis that there is no difference in mental instability probability for individuals younger than 35 and older or equal to 35.

3 Method

3.1 Dataset

Table18. Variables and Value Categories	
Variable	Categories
Independent Variables	
1. Sex	Male 48505 Female 11682
2. Age_group__at_arrest_	-18 2825 18-24 9183 25-34 19299 35-44 15048 45-54 8395 55-64 4225 65+ 1212
3. in Toronto or not (intoronto)	Yes 32694 No 27493
4. Perceived_Race	White 27708 Black 17518 East/Southeast Asian 4412 South Asian 3613 Middle-Eastern 3237 Indigenous 1932 Latino 1767
Control Variable	
1. threat_index	Domain: [0, 4.5] Total count: 60187
Dependent Variable	
1. Cooperation_score	Domain: [0,10] Total count: 60187
2. Mental_instab	Yes 2074 No 58113

Source: Arrests and Strip Searches (RBDC-ARR-TBL-001)

The given dataset contains information about the arrests and strip searches in Toronto. In total, the dataset contains 60,187 observations. It has two types of variables: independent variables and dependent variables. The independent variables include:

- Sex: categorical variable with two categories (Male and Female) representing the gender of the individual

- Age_group_at_arrest: categorical variable with seven categories representing the age group of the arrestee at the time of arrest
- in Toronto or not (in_toronto): binary variable indicating whether the arrest occurred in Toronto or not
- Perceived_Race: categorical variable with seven categories representing the perceived race of the arrestee
- threat_index: continuous variable representing the perceived level of threat posed by the arrestee on a scale of 0 to 4.5

The dependent variables are:

- Cooperation_score: continuous variable representing the level of cooperation shown by the arrestee during the arrest and strip search process on a scale of 0 to 10.
- Mental_instab: binary variable indicates whether the individual had a history of mental instability and includes two categories - "Yes" (2074) and "No" (58113).

3.2 Measurement

3.2.1 Dependent Variables

1) Cooperation_score

The ***Cooperation_score*** (cooperation score at arrest) is a composite index that is intended to reflect the level of cooperation that a suspect displayed during the process of being arrested. The index is calculated based on the values assigned to each of the six behaviors at arrest.

Cooperation Score = 7.5 + 2.5 * df["Actions_at_arrest___Cooperative"] - 0.5 * df["Actions_at_arrest___Concealed_i"] - 2 * df["Actions_at_arrest___Combative__"] - 1.5 * df["Actions_at_arrest___Resisted__d"] - 1 * df["Actions_at_arrest___Mental_inst"] - 2.5 * df["Actions_at_arrest___Assaulted_o"]

There are six specific types of actions at arrest:

1. Actions at arrest - Concealed items
2. Actions at arrest - Combative, violent or spitter/biter
3. Actions at arrest - Resisted, defensive or escape risk.
4. Actions at arrest - Mental instability or possibly suicidal
5. Actions at arrest - Assaulted officer
6. Actions at arrest – Cooperative

The cooperation coefficient has a linear positive relationship with the level of cooperation, meaning that the higher the index, the more cooperative the suspect was with the police. Conversely, lower indexes indicate less cooperation.

Each action at arrest is given a value from [-2.5, 2.5], depending on their level of cooperation (+) and level of threat (-). The cooperation score is set to be 0 to 10, with a starting value of 7.5. It is then adjusted based on the values assigned to each of the six behaviors at arrest, with each behavior being paired with a different coefficient based on the classification of the egregiousness and impact of the behavior. For example, if a suspect concealed items, this would result in a decrease in the cooperation score of 0.5. If the inmate was combative, violent or a spitter/biter, this would result in a decrease of 2.0 in the cooperation score. The final cooperation score is the sum of the starting value of 7.5 and the adjustments made based on each of the six behaviors at arrest.

Limitation: It's important to note that this scoring system is subjective and based on the judgment of the person assigning the values to each behavior. The values assigned to each behavior and the coefficients used to adjust the cooperation score may be adjusted based on the specific needs of the organization using the scoring system.

2) Mental_instab

The ***Mental_instab*** (mental instability at arrest) refers to whether the individuals show mental instability at arrest, with two categories: “Yes” and “No”.

The number of individuals showing mental instability at arrest is 2074, while the number of individuals not showing mental instability at arrest is 58113.

3.2.2 Independent Variables

The independent variables in this study include "Sex", "Age_group_at_arrest_", "in Toronto or not (intoronto)" and "Perceived_Race":

1) ***Sex*** refers to the gender of the individuals in the study, with two categories: Male and Female. The number of males in the study is 48505, while the number of females is 11682.
2) ***Age_group_at_arrest_*** refers to the age range of individuals at the time of their arrest. The age ranges included in the study are -18, 18-24, 25-34, 35-44, 45-54, 55-64, and 65+. The number of individuals in each age range are 2825, 9183, 19299, 15048, 8395, 4225, and 1212, respectively.

3) ***intoronto*** refers to whether the individuals were arrested in Toronto or not, with two categories: "Yes" and "No." The number of individuals arrested in Toronto is 32694, while the number of individuals not arrested in Toronto is 27493.
4) ***Perceived_Race*** refers to the race of the individuals in the study. The races included are White, Black, East/Southeast Asian, South Asian, Middle-Eastern, Indigenous, and Latino. The number of individuals in each race are 27708, 17518, 4412, 3613, 3237, 1932, and 1767, respectively.

3.2.3 Control Variable

1) threat_index

threat_index is a composite index that is intended to reflect the level of threat that a suspect displayed during the process of being arrested. The index is calculated based on the value assigned to three of the six behaviors at arrest.
Threat Index = 2 * df['Actions_at_arrest___Combative__'] + 1 * df['Actions_at_arrest___Resisted__d'] + 1.5 * df['Actions_at_arrest___Assaulted_o']
All of the three behaviors refer to the same as those in the section of ***3.2.1 Dependent Variables 1)***:

1. Actions at arrest - Combative, violent or spitter/biter
2. Actions at arrest - Resisted, defensive or escape risk.
3. Actions at arrest - Assaulted officer.

The threat index has a linear positive relationship with the level of threat, meaning that the higher the index, the more of a threat the suspect has. Conversely, lower indexes indicate less threat.
Each of three actions is given a value from {1, 1.5, 2}, depending on their level of threat. The threat index is set to be 0 to 4.5. For example, if a suspect has all of the three threatening actions, the threat index of him would be 1+1.5+2=4.5 If a suspect has none of the three threatening actions, the threat index of him would be 0.

Limitation: It's important to note that this index system is subjective and based on the judgment of the person assigning the values to each behavior. The values assigned to each behavior and the coefficients used to adjust the cooperation score may be adjusted based on the specific needs of the organization using the index system.

These independent variables will be analyzed to determine their impact on the dependent variable.

3.3 Analysis approach overview on two hypotheses

3.3.1 Overview of ANCOVA test

In our study, we used ANCOVA (analysis of covariance) to analyze the relationship between sex, perceived race, age group, arrested location, and cooperation score at arrest, while controlling for the effect of the threat index level. ANCOVA is a statistical technique that combines the principles of ANOVA and regression analysis, allowing us to examine the relationship between a continuous dependent variable (cooperation score) and one or more categorical independent variables, while controlling for the effects of one or more continuous variables (threat index).
In this study, we performed four one-way ANCOVA tests, one for each of the four independent variables. For each test, we set the null hypothesis as individuals from all categories of the independent variable have the same cooperation score at arrest, after controlling for the threat index. We calculated the F-value and p-value for each independent variable, and interpreted the results to determine if we could reject the null hypothesis.

3.3.2 Overview of Logistic Regression

Logistic regression is a statistical technique used to analyze the relationship between one or more independent variables and a binary outcome variable. In our study, we used logistic regression to analyze the effects of sex, perceived race, age group, and arrested location on the likelihood that individuals showed a mental instability at arrest.
We used a binary logistic regression to estimate the probability of individuals showing a mental instability at arrest as a function of each demographic variable, controlling for other demographic variables. We set a significance level of $p < 0.05$ for all analyses, and we reported the odds ratios with their corresponding 95% confidence intervals as measures of effect size. For each predictor variable, we calculated the odds ratio, which represents the change in odds of the outcome occurring for a unit change in the predictor variable. We also calculated the p-value for each predictor variable, and interpreted the results to determine if we could reject the null hypothesis. Overall, logistic regression allowed us to identify which independent variables had a significant effect on the likelihood of individuals showing a mental instability at arrest.

4 Results/ Findings

4.1 One-way ANCOVA result on Hypothesis 1

4.1.1 Sex vs Cooperation score

In this section, we wanted to investigate the relationship between sex and cooperation score, while controlling for the potential confounding effect of threat index.
Null hypothesis: Individuals from all sex categories have the same cooperation score at arrest on average after accounting for their threat index level.

Table 19. One way ANCOVA on sex vs cooperation score controlled threat_index

Source	SS	DF	F	p-unc	np2
Sex	23.863475	1	15.597172	0.000078	0.000259
threat_index	53758.339558	1	35136.462318	0.000000	0.368614
Residual	92080.752997	60184	NaN	NaN	NaN

- **Statistical interpretation:** Interpretation p-unc = “uncorrected p-value” for Sex is less than 0.05, indicating a statistically significant relationship between sex and cooperation score when controlling for threat index. We can REJECT the null hypothesis that individuals from all sex categories have the same cooperation score at arrest on average after accounting for their threat index level, even after controlling for the threat index at arrest of individuals.
- **Practical interpretation:** We hypothesized that sex would be able to predict an individual's cooperation score at arrest. From our result, we see that there is a statistically significant relationship between sex and the cooperation score when controlling for individuals’ threat index. This finding suggests that there may be inherent differences between males and females in how they respond to law enforcement encounters, which should be taken into consideration when developing policies and training programs for law enforcement officers.

4.1.2 Perceived race vs Cooperation score

Null hypothesis: Individuals from all perceived race groups have the same cooperation score at arrest on average after accounting for their threat index level.

Table 20. One way ANCOVA on perceived race vs cooperation score controlled threat_index

Source	SS	DF	F	p-unc	np2
Perceived_Race	205.818396	6	22.463016	1.358838e-26	0.002235
threat_index	53356.952429	1	34940.261543	0.000000e+00	0.367331
Residual	91898.798076	60179	NaN	NaN	NaN

- **Statistical interpretation:** Interpretation p-unc = “uncorrected p-value” for Perceived race is less than 0.05. There is a statistically significant relationship between an individual's perceived race and their cooperation score at the time of arrest after accounting for their threat index level. We can REJECT the null hypothesis that individuals from all perceived race groups have the same cooperation score at arrest on average after accounting for their threat index level, even after controlling for the threat index at arrest of individuals.
- **Practical interpretation:** We hypothesized that perceived race would be able to predict an individual's cooperation score at arrest. From our result, we see that there is a statistically significant relationship between perceived race and the cooperation score when controlling for individuals’ threat index. In practical terms, this means that there may be a bias in the way that law enforcement officers interact with individuals of different perceived races that may affect their level of cooperation during an arrest. Further research is needed to better understand this relationship and to develop strategies to mitigate the impact of bias on law enforcement outcomes.

4.1.3 Age group vs Cooperation score

Null hypothesis: Individuals from all age groups have the same cooperation score at arrest on average after accounting for their threat index level.

Table 21. One way ANCOVA on age group vs cooperation score controlled threat_index

Source	SS	DF	F	p-unc	np2
Age_group__at_arrest_	42.283587	6	4.606632	0.00011	0.000459
threat_index	53612.061995	1	35044.954626	0.00000	0.368027
Residual	92062.332885	60179	NaN	NaN	NaN

- **Statistical interpretation:** Interpretation p-unc = “uncorrected p-value” for Age group is less than 0.05. We can REJECT the null hypothesis that individuals from all age groups have the same cooperation score at arrest on average after accounting for their threat index level, even

after controlling for the threat index at arrest of individuals. The effect size, measured by np2, is very small at 0.000459. This indicates that age group accounts for a very small amount of variance in cooperation score, with most of the variance being accounted for by the threat index.

- **Practical interpretation:** We hypothesized that age group would be able to predict an individual's cooperation score at arrest. From our result, we see that there is a statistically significant relationship between age group and the cooperation score when controlling for individuals' threat index. Practically speaking, this could have implications for law enforcement and criminal justice policies, as it suggests that individuals of different age groups may require different approaches to ensure their cooperation during an arrest. Overall, the results suggest that age group is a weak predictor of cooperation score at arrest, but still has a statistically significant relationship with the outcome.

4.1.4 Intoronto vs Cooperation score

Null hypothesis: Individuals from all arrested location groups have the same cooperation score at arrest on average after accounting for their threat index level.

Table 22. One way ANCOVA on intoronto vs cooperation score controlled threat_index

Source	SS	DF	F	p-unc	np2
intoronto	149.955252	1	98.145181	4.049891e-23	0.001628
threat_index	53873.630746	1	35260.100464	0.000000e+00	0.369432
Residual	91954.661221	60184	NaN	NaN	NaN

- **Statistical interpretation:** Interpretation p-unc = "uncorrected p-value" for intoronto is less than 0.05. This indicates that there is a statistically significant relationship between arrested in Toronto or not and the cooperation score when controlling for individuals' threat index. We can REJECT the null hypothesis that individuals from all arrested location groups have the same cooperation score at arrest on average after accounting for their threat index level, even after controlling for the threat index at arrest of individuals.
- **Practical interpretation:** We hypothesized that arrested in toronto or not would be able to predict an individual's cooperation score at arrest. From our result, we see that there is a statistically significant relationship between arrested in toronto or not and the cooperation score when controlling for individuals' threat index. Practically speaking, The researchers may further investigate the reasons behind this relationship, such as differences in police tactics or community relations in Toronto versus other locations.

4.2 Logistic Regression result on Hypothesis 2

Table 23. The original Logit Regression results on hypothesis 2

Logit Regression Results						
	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
Intercept	-5.285	0.257	-20.58	0.000	-5.789	-4.782
Sex[T.M]	-0.269	0.054	-4.99	0.000	-0.375	-0.164
Perceived_Race[T.East/Southeast Asian]	-0.454	0.107	-4.26	0.000	-0.663	-0.246
Perceived_Race[T.Indigenous]	-0.036	0.123	-0.29	0.766	-0.278	0.205
Perceived_Race[T.Latino]	-0.512	0.161	-3.17	0.002	-0.829	-0.196
Perceived_Race[T.Middle-Eastern]	-0.078	0.105	-0.74	0.454	-0.285	0.127
Perceived_Race[T.South Asian]	-0.522	0.119	-4.39	0.000	-0.756	-0.289
Perceived_Race[T.White]	-0.088	0.053	-1.68	0.093	-0.191	0.015
Age_group__at_arrest_[T.18-24]	1.696	0.258	6.56	0.000	1.190	2.203
Age_group__at_arrest_[T.25-34]	1.983	0.254	7.81	0.000	1.486	2.480
Age_group__at_arrest_[T.35-44]	1.958	0.255	7.68	0.000	1.459	2.457
Age_group__at_arrest_[T.45-54]	1.849	0.258	7.15	0.000	1.343	2.356
Age_group__at_arrest_[T.55-64]	1.590	0.270	5.89	0.000	1.062	2.119
Age_group__at_arrest_[T.65+]	1.532	0.316	4.84	0.000	0.912	2.152
intoronto	0.703	0.049	14.37	0.000	0.608	0.799

Since the above regression table is not too interpretable, the odds ratio is illustrated as follows to make the features more interpretable, where

$$Odds\ ratio = \frac{Odds\ after\ a\ unit\ change\ in\ the\ predictor}{Original\ odds}$$

If the odds are the same across groups, the odds ratio (OR) will be 1.0. If not, the OR will be larger or smaller than one. OR>1, if the predictor increases then the probability of outcome occurring increases. When OR<1, if predictor increases, then probability of outcome occurring decreases

Table 24. The Odds ratio results of Logit Regression on hypothesis 2

Odds ratio			
	Lower CI	**Upper CI**	**OR**
Intercept	0.003062	0.008377	0.005064

Sex[T.M]	0.687417	0.849014	0.763954
Perceived_Race[T.East/Southeast Asian]	0.515281	0.782284	0.634899
Perceived_Race[T.Indigenous]	0.757264	1.227139	0.963986
Perceived_Race[T.Latino]	0.436540	0.822105	0.599068
Perceived_Race[T.Middle-Eastern]	0.752010	1.135983	0.924268
Perceived_Race[T.South Asian]	0.469775	0.748696	0.593059
Perceived_Race[T.White]	0.825875	1.014792	0.915474
Age_group__at_arrest_[T.18-24]	3.288263	9.049775	5.455093
Age_group__at_arrest_[T.25-34]	4.418219	11.944130	7.264419
Age_group__at_arrest_[T.35-44]	4.300086	11.674594	7.085320
Age_group__at_arrest_[T.45-54]	3.830628	10.550677	6.357336
Age_group__at_arrest_[T.55-64]	2.890856	8.321915	4.904840
Age_group__at_arrest_[T.65+]	2.489618	8.601017	4.627445
intoronto	1.835896	2.224266	2.020772

We perform a logistic regression to examine the effects of Sex, Perceived race, Age group, Arrested in Toronto or not, on the likelihood that individuals show a mental instability at arrest. All of the **statistically significant** features are interpreted as follows:

- Sex[T.M]: Compared with individuals who are females, the odds ratio of getting the individuals showing a mental instability at arrest who are males decreases by about 0.764 times.
- Perceived_Race[T.East/Southeast Asian]: Compared with individuals who are black, the odds ratio of getting the individuals showing a mental instability at arrest who are East/Southeast Asian decreases by about 0.635 times.
- Perceived_Race[T.Latino]: Compared with individuals who are black, the odds ratio of getting the individuals showing a mental instability at arrest who are East/Southeast Asian decreases by about 0.599 times.
- Perceived_Race[T.South Asian]: Compared with individuals who are black, the odds ratio of getting the individuals showing a mental instability at arrest who are East/Southeast Asian decreases by about 0.593 times.
- Age_group__at_arrest_[T.18-24]: Compared with individuals who are less than 18, the odds ratio of getting the individuals showing a mental instability at arrest who are between 18 and 24 increases by about 5.455 times.

- Age_group__at_arrest_[T.25-34]: Compared with individuals who are less than 18, the odds ratio of getting the individuals showing a mental instability at arrest who are between 25 and 34 increases by about 7.264 times.
- Age_group__at_arrest_[T.35-44]: Compared with individuals who are less than 18, the odds ratio of getting the individuals showing a mental instability at arrest who are between 35 and 44 increases by about 7.085 times.
- Age_group__at_arrest_[T.45-54]: Compared with individuals who are less than 18, the odds ratio of getting the individuals showing a mental instability at arrest who are between 45 and 54 increases by about 6.357 times.
- Age_group__at_arrest_[T.55-64]: Compared with individuals who are less than 18, the odds ratio of getting the individuals showing a mental instability at arrest who are between 55 and 64 increases by about 4.905 times.
- Age_group__at_arrest_[T.65+]: Compared with individuals who are less than 18, the odds ratio of getting the individuals showing a mental instability at arrest who are older than 65 increases by about 4.627 times.
- intoronto: Compared with individuals who are arrested outside Toronto, the odds ratio of getting the individuals showing a mental instability at arrest who are arrested inside Toronto increases by about 4.627 times.

Test accuracy:
In this step, the data is splitted into two groups: the train dataset and the test dataset, each dataset contains x (independent variables) and y (dependent variable). Then we run the prediction function to get the prediction of y based on the x-test dataset and compare it with our y-test data using the "acuracy_score" function, which shows that the test accuracy of this logistic regression model is approximately 0.9638.

5 Discussion

5.1 Limitations

The findings of this project have to be seen in light of some limitations.

- Firstly, the dataset used in this study is limited to a single city, Toronto, which means that the results may not be generalizable to other locations. And it only covers a relatively short period of 2 years (*'Arrest_Year'*), which may not capture long-term trends or changes over time.
- Besides, there are some issues with the accuracy of the data used in the study. The dataset only relies on self-reported data from officers, which may be biased or incomplete. For

example, it includes indicators of whether a person was booked at a police station within 24 hours following an arrest, but due to issues with the booking template, there are records where a person was strip searched but the data does not indicate a booking. This could impact the validity of the study's findings. Additionally, some of the data in the *'ArrestLocDiv'* column is marked as 'XX'. It is unclear whether this means that the data is an accurate representation that it is outside of the City of Toronto or it just cannot be geo-coded. The age of the person arrested is given as a range, not the exact age, and the arrest month *'Arrest_Month'* is also given as a period of months. These could impact the accuracy of the study's findings and limit the ability to draw strong conclusions from the data.

- Another limitation of the study is related to the calculation of the cooperation score and the threat index level. The scoring and calculating system used in the study is subjective and is fully based on the judgment of the person assigning the values to each behavior at arrest. The coefficients used to adjust the cooperation score may be adjusted based on the specific needs of the law enforcement agency using the system.
- Besides, the study was limited to variables that were available in the dataset, and did not include other potentially relevant variables such as prior criminal history or mental health status. It also did not explore the impact of interactions between variables, such as the intersection of race and gender, which may have important implications for the outcomes at the time of arrest.
- Finally, there are some missing values for the *Search Reason* and *'ItemsFound'* columns of the data, which could lead to incomplete or biased results. Overall, these limitations should be taken into consideration when interpreting the findings of the study and applying them to make policy decisions.

It is important to keep these limitations in mind when interpreting the results of this study, and to use them as a basis for future research that can further explore the complex relationships between demographic variables and outcomes at the time of arrest.

5.2 Further Improvement

To address the limitations, several improvements could be made for our future studies.

Firstly, we may use a larger and more diverse dataset that includes data from multiple cities or jurisdictions. This would help to ensure that the findings are more generalizable and can be applied to a wider range of contexts. Additionally, future studies could aim to collect more detailed and accurate data, such as through direct observation or interviews with arrestees, to provide a more comprehensive picture of the factors that influence outcomes at the time of arrest.

Secondly, we may use more objective measures for calculating scores such as the cooperation score and the mental instability index. This could involve developing standardized criteria and assessment tools to ensure that these scores are calculated in a consistent and reliable manner across different law enforcement agencies and jurisdictions. For future studies, we could also investigate additional variables that may impact the cooperation scores and the threat index. By doing so, it could provide more comprehensive and accurate insights into the dynamics of arrests and strip searches in the criminal justice system.

In addition, future research could aim to explore the interactions between demographic variables and other potentially relevant factors, such as prior criminal history, mental health status, and the circumstances surrounding the arrest. This could provide a more nuanced understanding of the factors that impact outcomes at the time of arrest and could help to inform more targeted interventions and policies.

Overall, there is a need for further research that can build upon the findings of this study and provide a more comprehensive understanding of the complex factors that influence outcomes at the time of arrest. By addressing the limitations of this study and improving upon its findings, we can work towards developing more equitable and effective policies and practices for law enforcement agencies and criminal justice systems.

6 Conclusion

This study aims to study the relationship between the level of cooperation and mental instability at the time of arrest, and various indicators, including the race, gender, age group of the individual, threat index and whether the location in Toronto or not at the time of the arrest, based on the 'Arrests and Strip Searches' dataset from the Toronto Police Service.

The two research questions we had were: 1. Does any of ***perceived race, sex, location in Toronto or not*** *at arrest or* ***age group*** effects on the level of cooperation at the time of arrest controlled by ***threat index***? 2. How does ***perceived race***, ***sex***, ***location in Toronto or not*** at arrest and ***age group*** affect the probability of the individuals' mental instability at the time of the arrest? Exploratory data analysis, Power analysis, T-tests, ANCOVA tests, and Logistic Regression tests were conducted in the study.

The ANCOVA analysis was conducted to examine the relationship between the cooperation score and demographic variables, while controlling for the effect of the threat index. The results indicated that age group and perceived race were significant predictors of the cooperation score, with older individuals and individuals who were perceived as Black or Indigenous having lower cooperation scores. However, the effect sizes were small. Sex and being arrested in Toronto were not

significant predictors of the cooperation score. The results suggest that demographic factors play a role in the level of cooperation shown by individuals at the time of arrest, but the effect sizes are relatively small.

The logistic regression analysis was conducted to examine the effects of sex, perceived race, age group, and being arrested in Toronto on the likelihood of showing mental instability at the time of arrest. The results indicated that all variables except for sex were significant predictors of mental instability. Compared to Black individuals, those who were perceived as East/Southeast Asian, Latino, or South Asian had lower odds of showing mental instability. Older age groups (25 and older) had higher odds of showing mental instability than those who were less than 18 years old. Being arrested in Toronto was also associated with higher odds of showing mental instability. The effect sizes were moderate to large, indicating that demographic factors play a substantial role in mental instability at the time of arrest.

Overall, this study provides valuable insights into the factors that impact the level of cooperation and the mental instability condition during arrests. There are several limitations to the dataset used, including the reliance on self-reported data from officers, and the subjective calculation of the cooperation score and the threat index. Future research could address these limitations by expanding the scope of the dataset, using larger sample sizes, and considering alternative data collection methods. The scoring and index calculating system could also be refined to improve its validity and reliability.

In conclusion, the findings from the ANCOVA and logistic regression analyzes both provide insight into the factors that contribute to cooperation and mental instability at the time of arrest. The small effect sizes observed in the ANCOVA analysis indicate that the threat index is a more important predictor of cooperation than demographic variables. However, the moderate to large effect sizes observed in the logistic regression analysis suggest that demographic variables play a substantial role in mental instability at the time of arrest.

These findings can inform policies and interventions aimed at improving the outcomes of arrests and reducing the incidence of mental instability among arrestees. By understanding the factors that influence cooperation and mental instability during arrests, we can work towards improving trust and cooperation between law enforcement and the community, which is essential for promoting safety and well-being for all. We hope that our findings will inform future research and policy efforts aimed at improving the outcomes and experiences of individuals who are arrested.

References

American Psychological Association. (2011, August 18). *Study finds sex differences in mental illness*. https://www.apa.org/news/press/releases/2011/08/mental-illness

Bierie, David M. 2015. "Assault of Police." *Crime & Delinquency*. doi: 10.1177/0011128715574977.

Educaloi. (2020, November 12). *What happens when someone accused of a crime has a mental illness?* Éducaloi. Retrieved April 10, 2023, from https://educaloi.qc.ca/en/capsules/what-happens-when-someone-accused-crime-mental-illness

Engel, R. S. (2003). *Explaining suspects' resistance and disrespect toward police.* Journal of Criminal Justice, 31, 475-492.

Jurewicz I. (2015). Mental health in young adults and adolescents - supporting general physicians to provide holistic care. *Clinical medicine (London, England)*, *15*(2), 151–154. https://doi.org/10.7861/clinmedicine.15-2-151

Kaminski, R., & Sorensen, D. (1995). *A multivariate analysis of individual, situational and environmental factors associated with police assault injuries.* American Journal of Police, 14, 3-48.

Litman, T. (2021, November 4). *Urban sanity and its impact on Mental Health*. Urbanet. Retrieved April 10, 2023, from https://www.urbanet.info/urban-sanity/

McCluskey, J. D., S. D. Mastrofski, and R. B. Parks. (1999). "To Acquiesce or Rebel: Predicting Citizen Compliance with Police Requests." *Police Quarterly* 2:389-416.

McGuire, T. G., & Miranda, J. (2008). New evidence regarding racial and ethnic disparities in mental health: policy implications. *Health affairs (Project Hope)*, *27*(2), 393–403. https://doi.org/10.1377/hlthaff.27.2.393

O'Neil, J., & Hanson, A. (2020). Toronto police launch review of use of force, racial profiling in wake of protests. *CBC News*. https://www.cbc.ca/news/canada/toronto/toronto-police-review-use-of-force-racial-profiling-1.5623215

Perreault, Samuel. (2019). *Police-reported crime in rural and urban areas in the Canadian provinces, 2017.* Statistics Canada. https://www150.statcan.gc.ca/n1/pub/85-002-x/2019001/article/00009-eng.htm

U.S. Department of Justice, Federal Bureau of Investigation. (2009). *Uniform Crime Reporting Program Data: Arrests by Age, Sex, and Race, 2009*. ICPSR30761-v1. Ann Arbor, MI:

Inter-university Consortium for Political and Social Research [distributor], 2011-09-30. doi: 10.3886/ICPSR30761.

Walker, S. (1999). *The police in America (3rd ed.)*. Boston: McGraw-Hill.