

Implicit Racial Bias on the Discretionary Nature of Toronto Police Arrests and the Likelihood of Arrests Based on Different Types of Demeanors

by

Eva Li
Purui Zhang

A final paper submitted in conformity with the requirements
for the course of INF 2178 H

Faculty of Information
University of Toronto

Table of Contents

1. Introduction	3
2. Literature Review	3
2.1. Racial Disparity in Police Arrest	3
2.2. Role of Sex and Age on Arrest Outcomes	5
2.3. Encounter-specific Variables on Arrest Outcomes	7
3. Exploratory Data Analysis	7
3.1. Descriptive Analysis	7
3.2. T-tests	14
3.2.1. T-test: Average Number of Arrests in 2020 vs 2021	14
3.2.2. T-test: Average Number of Arrests in 2020 vs 2021 Among Black Minorities	15
3.2.3. T-test: Average Number of Arrests in 2020 vs 2021 Among Indigenous Minorities	16
3.2.4. T-test: Average Number of Arrests of Suspects Concealing Items vs Not Concealing Items	17
3.2.5. T-test: Average Number of Arrests of Suspects Performing Combative Behaviour vs Suspects Not Performing Combative Behaviour.	18
3.2.6. T-test: Average Number of Arrests of Suspects Resisting vs Not Resisting	19
3.2.7. T-test: Average Number of Arrests of Suspects with Mental Instabilities or Suicidal Tendencies vs Suspects without Mental Instabilities or Suicidal Tendencies	20
3.2.8. T-test: Average Number of Arrests of Suspects Assaulting Officers vs Suspects Not Assaulting Officers	21
3.2.9. T-test: Average Number of Arrests of Suspects Cooperating vs Suspects Not Cooperating.	22
3.2.10. T-test: Average Number of Arrests of Males vs Females.	24
3.2.11. T-test: Average Number of Arrests of Juveniles vs Adults.	25
4. Research Design & Methodology	26
4.1. Dataset Description	26
4.2. Research Questions	26
4.3. Data Analysis Method	27
4.3.1. Analysis of Covariance	27
4.3.2. Analysis of Logistic Regression	28
5. Results & Findings	29
5.1. One-way ANCOVA: Adjusted Mean Arrests Across Perceived Races	29
Hypothesis Model	29
Assumptions Check	29
Interpreting ANCOVA Results	31
5.2. Logistic Regression: Chosen factors on Likelihood of Arrest	32
Results	32
Prediction Interval Plot	35
Confusion Matrix Plot	36
6. Discussion and Limitations	36
7. Conclusions	38
References	40

1. Introduction

Ethnic discrimination in police arrests has been a contentious issue, particularly since 2020. However, ethnicity is not the only demographic feature that shows significant differences among different groups in police decision-making on arrests. To address concerns of potential bias and discrimination, it is crucial to first understand the phenomenon and current context. With this understanding, policies can be developed to increase transparency and gradually reduce discrimination. Given the importance of police arrests, as they may affect one's record and have lasting impacts on an individual's development, it is essential to investigate whether factors beyond criminal behavior influence officers' decisions to make arrests. This report aims to explore the variation in police arrest decisions, including factors such as an individual's race, age, sex, and demeanor during the arrest. To achieve this, we analyzed the "Arrests and Strip Searches" dataset from the Toronto Police Service website, which contains arrest records from April 2021 to December 2022. Through exploratory data analysis and several t-tests, we confirmed our research questions. The first question examined how an individual's ethnicity affects the number of arrests, controlling for previous arrests, using ANCOVA. The second question used logistic regression to study the relationship between an individual's likelihood of being arrested, their age, gender, and level of cooperation or aggression during the arrest. Our findings indicate a significant difference in the adjusted mean number of arrests by perceived race, with Black and Indigenous individuals experiencing disproportionate policing. Individuals perceived as Black were arrested significantly more than all other minority groups, except for Indigenous individuals who were arrested on average 0.261 times more than their Black counterparts. Additionally, we found that if an individual performed a hostile or disrespectful demeanor during an arrest, their likelihood of being arrested was higher than those who did not. In conclusion, this report sheds light on the ethnic and behavioral factors that influence police arrest decisions in the Great Toronto Area. While our findings have limitations, they provide a foundation for future research and the development of evidence-based policies to address discrimination and bias in police arrests.

2. Literature Review

2.1. Racial Disparity in Police Arrest

On June 15, 2022, then Toronto Police Chief, James Ramer, issued an apology for the disproportionate policing over the Black, Indigenous, and members of other radicalized populations, for the over-representation of the use-of-force incidents and strip searches among these groups (Fanfair, 2022). This was in response to the release of the race-based data collected in 2020 that looked at the law enforcement's perception of an individual's race in reportable use-of-force incidents and strip searches (Race and Identity-based Data Collection Strategy, n.d.), which showed evidence of biased policing towards minorities. Such findings were further corroborated by subsequent interim report by the Ontario Human Rights Commission's (OHRC) inquiry into racial profiling and racial discrimination of Black persons by the Toronto Police Service (TPS) which confirmed evidence of "disproportionate burden of law enforcement" against members of the Black community despite representing only 8.8% of the Toronto population at the time of the report (Ontario Human Rights Commission, 2020). More specifically, the analysis saw Black people representing almost a third of the total charges with respect to "out-of-sight" driving, possession of cannabis, and 25% of all Special Investigations Unit (SIU) cases resulting in death, serious injury or allegations of sexual assault.

Per the 2017 provincial Anti-Racism Act – the enacted bill has since mandated the collection of race-based data by organizations under the public sector including law enforcement (Race and Identity-Based Data Collection, n.d.). The Toronto Police Service (TPS) has since introduced the Race and Identity-Based Data Collection (RBDC) Strategy has a key component to TPS's commitment to Equity and Transparency in an effort to promote fair and unbiased policing to order to address the systemic bias that is embedded in many of our institutional practices. Such an approach can serve as an informative proxy to better understand the scope of the racial differential that exists in law enforcement. However, it is also important to acknowledge why such an intervention is necessary extends beyond addressing the systemic discrimination and inequalities that exist in the criminal justice system. That is, a criminal charge can have profound consequences on the accused that include but not limited to the following: temporary detention, curtailment of freedom prior to trial, fines, probation, potential incarceration, as well as long-term implication on future employment, education, and volunteer opportunities (Wortley & Jung, 2020). Nonetheless, necessary actions are required to circumvent potential wrongful and unjust burden of law. Among this is understanding the discretionary nature and its role in the racial disparity in police arrests. What makes this of critical importance is that such point of contact often serves as the gateway for an individual into the criminal justice system (Spencer et al., 2016).

The discretionary nature of police arrests has long been a topic of discussion within the judicial system and its implication on exercising fair and unbiased law enforcement (Smith et al., 1984). And that such discretion is strongly influenced by an individual's characteristics with respect to race, sex, and age. As a result, such disparity may very well be the product of bias by actors of law enforcement with profound implications on the accused, especially among the underserved and marginalized populations (New York University School of Law, n.d.). In other words, racial minorities are more susceptible to biased law enforcement practices resulting in the observed disproportionate representation among the arrested population (D'alessio & Stolzenberg, 2003). However, and understandably, we also cannot discount the effect of the accused's demeanor and history of prior contact with police at the time of arrest. Rather, some have argued that such aforementioned factors with respect to an individual's characteristics have "exhibited much less evidentiary consistency in explaining officer use-of-force and arrests" (Headley et al., 2020). Instead, they are confounded by situational-specific variables such as presence of a weapon and demeanor, such as resistance to arrest, during the encounter in influencing an officer's discretion in use-of-force and subsequent arrest outcomes. As such, these variables can have implications on an officer's perceptions of potential danger and criminal behaviour of the individual in question (Headley et al., 2020), which can be argued to have been shaped by the already-existing bias within law-enforcement. In fact, in a study published in 2016 by Spencer, Charbonneau, and Glaser, the authors have highlighted the role of implicit bias that "operate outside of conscious awareness and control" in policing that has manifested across all domains of law enforcement. That is, there is pervasive use of racial stereotypes that influence arrest outcomes, especially under the context of ambiguous circumstances that tend to result in bias and subjective judgement (Spencer et al., 2016; Schirrmester et al., 2020). This further highlights the role of racial identity in eliciting biased application of law and therefore, arrest outcome. For example, it has been well-documented that there exists a racial disparity in police arrests against Black males which can be attributed to the preconceived notion, often perpetuated by media, that a Black person is more criminally inclined and aggressive than their White counterparts; the result of such bias is reflected in the disproportionate number of black person

among juvenile and adult offenders, especially in the American criminal justice system (D'alessio et al., 2003; Piquero et al., 2008; Spencer et al., 2016). In fact, in a study conducted by Schleiden, Soloski, Milstead, and Rhynehart (2020) on the racial disparities in arrests found that by legal adulthood, Black males were at 8% increased risk for an arrest compared to their White counterparts, and 11% increase by the age of 23.

Two competing theories have since been proposed in describing the existing racial disparity: (1) Differential Involvement Hypothesis, and (2) Differential Selection Hypothesis (Piquero et al., 2008). The former suggest that such overrepresentation is attributed to the higher rates of committing an offence resulting in the observed increased proportion of arrests and prosecution of the population in question. Those who subscribe to such theory have attributed this due to “differences in social factors resulting in differences in crime patterns between races” (e.g., socioeconomic status, poverty, and segregation as a by-product of institutional practices and years of racial oppression) (D’alessio & Stolzenberg, 2003). However, the latter describe the racial disparity to be a result of the prejudice and discrimination that is embedded in law enforcement practices. Nonetheless, Wortley and Jung (2020), have suggested that there has been growing evidence that both theories hold true and work in parallel that result in the observed exaggerated arrest rates among minorities.

In the current study, we will further investigate the existing racial disparity in Toronto Police Service (TPS) practices using the Race and Identity-Based Data provided to the public. As contact with law enforcement serves as the point of entry into the criminal justice system, we will look for evidence of pervasive racial prejudice and discrimination in institutional practices on the number of arrests made by law enforcement in 2021, while controlling for previous accounts of arrests recorded in 2020. Such an approach will allow us to better discern and isolate the role of perceived race, as a reflection of the implicit bias, on arrest outcomes that are independent of previous arrest history and criminal activity, thus disentangling the effect of perceived race and identify potential implicit bias in law enforcement practices by controlling for such potential confounding effects. However, while we recognize the role of perceived race and the influence of racial stereotypes in the discretionary nature of arrest outcome, we further acknowledge that this cannot be studied in isolation. We have previously mentioned how police discretion is strongly influenced by race, sex, and age. Furthermore, this can often be confounded by situational-specific variables such as presence of a weapon and demeanor at the time of arrests (e.g., resistance to arrest, combative, or cooperative) in influencing an officer’s discretion in use-of-force and subsequent arrest outcomes. As a result, we will further investigate the likelihood of arrest with respect to such factors in better discerning the implicit bias in law enforcement practices. The following section will further introduce the role of age, sex and encounter-specific factors on arrest outcome.

2.2. Role of Sex and Age on Arrest Outcomes

The role of gender in police arrest decisions has been a topic of interest in criminology due to the significant gender differences in arrest rates. FBI Crime in the United States data from 2017 shows that men account for 73% of total arrests, while women account for only 27%. However, this gender gap has decreased from a much wider gap of 84% versus 16% in 1980 (*Policing Women: Race and gender disparities*, 2019). Meanwhile, teenage crime also plays a big role in the overall criminal system. This literature review aims to provide an overview of existing research on gender differences and juvenile differences in police arrests.

One reason for the higher proportion of male arrests is that men are statistically more likely to commit criminal acts than women, which may be due to biological differences. Research suggests that women tend to display a self-focused ruminative style, while men tend to use more distracting strategies, leading to outward-directed aggression and more violent criminal behavior (Staniloiu & Markowitsch, 2012). Additionally, institutional issues may also contribute to gender disparities in arrests. The Canadian criminal justice system, for example, has been primarily designed for male offenders, leading to potentially biased treatment of female offenders who may be more likely to be tolerated in certain circumstances (Statistics Canada, 2009).

In conclusion, gender differences in police arrests are a multifaceted issue influenced by various factors, including inherent biological differences and biased treatment in the criminal justice system. Further research is needed to better understand the extent of gender differences in arrests and to develop strategies to address these differences. It is important to recognize that societal attitudes towards gender and potential biases in the criminal justice system may contribute to disparities in arrests, with women potentially being treated more leniently than men.

Juvenile delinquency is another pervasive issue worldwide, and numerous studies have investigated the underlying causes of this phenomenon to prevent criminal behavior and promote positive outcomes for young people. While law enforcement policies vary among countries, typically, there is greater leniency for youth offenders compared to adult offenders. This review proceeds to provide an overview of key findings related to police arrests among youth and non-youth populations.

It is important to clarify that youth refers to individuals aged 12 to 17, as children below the age of 12 do not bear legal responsibility for their behavior in Canada. According to Statistics Canada, the youth crime rate is lower than that of young adults between the ages of 18 and 24, but higher than that of adults aged 25 and older. In 2014, police-reported data revealed that 17-year-olds had the highest rate of being accused of a crime, not only within the youth group but also across all age groups. Although the youth crime rate varies when compared to adults of different age ranges, research suggests that, on average, youth are less likely to be arrested than non-youth groups (Piquero, 2008). Among youth crimes, co-offending, or involvement in crimes with more than one accused, is a critical feature that is more prevalent among this group than among adult offenders. The most frequent types of youth crimes include theft of \$5,000 and under, assault, and drug use, primarily cannabis possession. Additionally, drug use is more commonly observed among older youth than their younger counterparts. Wildeman and Turney (2014) found that children who grow up in impoverished families, with low levels of education, and parents who use drugs tend to be more prone to criminal behavior.

To sum up, the rate of juvenile delinquency in Canada has decreased continuously since the 1990s. Understanding this phenomenon provides a foundation for developing effective policies that can prevent youth from engaging in criminal behavior and create a safer and more supportive environment for young people.

As found earlier, the dataset from Toronto Police Services did not indicate significant arrest discrepancies between youth and non-youth groups. However, when age was combined with perceived race, several notable findings emerged. In this next stage, we will examine how additional factors may impact the likelihood of an individual being arrested, while controlling for age and gender.

2.3. Encounter-specific Variables on Arrest Outcomes

In the past four decades, research has shown that police actions, including arrests, can be influenced by various extralegal factors such as suspect demeanor (Lundman, 1996). In addition to demographic features, this study is interested in investigating how acquired behaviors, especially those exhibited prior to or during police arrests, impact police decisions on arrests. The interaction and demeanor between suspects and police can vary widely, including both positive and negative aspects, and this study aims to explore how police officers may respond differently to different attitudes. It is important to note that the findings of this review are generalized for all police officers and suspects, with no specific focus on the Toronto area. However, data from Toronto will be used later to validate these insights.

Lundman (1996) has previously noted that when disrespect and hostility are within legal boundaries and criminal behavior is partially under control, suspect demeanor can influence police actions. However, Lundman also pointed out that race and class can also impact police actions. Novak and Engel (2005) reported similar findings, but with more detailed insights, indicating that disrespect and hostility can increase the likelihood of a suspect being arrested by police officers. Suspects with mental disorders, on the other hand, are more likely to behave disrespectfully but have a smaller likelihood of being arrested due to the consideration of mental disadvantages as mitigating factors during encounters. However, there are different viewpoints on this topic. Smith and Visser (1981) argued that the influence of demeanor on police officers' arrest decisions may be overstated, as it fails to consider the complex decision-making processes that police officers engage in when deciding whether or not to make an arrest. Johnson (2016) further pointed out that police officers may react differently to suspects who behave disrespectfully, and he applied the displaced aggression theory in his experiments. Johnson found that police officers' mood prior to encountering suspects can play a role in responding to hostile behavior. Specifically, his experiments illustrated that police officers who are primed negatively are more likely to make arrest decisions than those who are primed positively. When police officers are primed negatively and simultaneously encounter suspects displaying hostile demeanor, they are most likely to make arrest decisions (Johnson, 2016).

3. Exploratory Data Analysis

3.1. Descriptive Analysis

In the current study, we present two analyses. First, exploratory data analysis (EDA) was conducted as an initial investigation into the data to identify potential anomalies and patterns in the dataset and explore potential relationships between variables. Descriptive statistics were performed to describe and characterize the observations in the data where we report the proportion of samples falling under each class of the categorical variables with respect to the individual in question during the time of the arrest. This includes repeated arrest of an individual with respect to measures of demographic factors (sex, perceived race, age group and whether they were a youth), and month of arrest (Table 1).

Table 1. Descriptive Statistics of Collected Data on Demographic Characteristics (N = 65,276)

Variables /Features	Categories	Count	Frequency (%)
Arrest Month	Jan-Mar	17,613	27.0
	Apr-June	15,185	23.3
	July-Sept	16,773	25.7
	Oct-Dec	15,705	24.1
Perceived Race	White	27,723	42.5
	Black	17,526	26.8
	Unknown or Legacy	5,060	7.8
	East/Southeast Asian	4,415	6.8
	South Asian	3,613	5.5
	Middle Eastern	3,237	5.0
	Indigenous	1,934	3.0
	Latino	1,768	2.7
Sex	Male	52,650	80.7
	Female	12,617	19.3
	Unknown	9	0.01
Age Group at Arrest	Aged 17 and under	3,042	4.7
	Aged 18 to 24 years	10,041	15.4
	Aged 25 to 34 years	20,949	32.1
	Aged 35 to 44 years	16,242	24.9
	Aged 45 to 54 years	9,066	13.9
	Aged 55 to 64 years	4,590	7.0
	Aged 65+	1,322	2.0
Youth at Arrest	Not a Youth	62,234	95.3
	Youth (17 and under)	3,042	4.7

Results of the analysis showed that the majority of the arrests completed over the span of 2020 to 2021 by Toronto Police Services (TPS) have been adult offenders, with only 4.7% of the arrests to be under the age of 18. More specifically, among the adult population, individuals between the ages of 25 and 44 are most represented in the dataset making up 57% of the total arrests by age group alone. Initial investigation also showed evidence of overrepresentation of males (n=52,650) in the arrested population by Toronto Police compared to females (n = 12,617) at 80.7%. With relevance to the purpose of the present study, while it appears that individuals of Caucasian descent were more represented in the current dataset at 42.5%, Black individuals continue to be the most over-represented person of color when it comes to arrests compared to other minority populations (26.8%). Moreover, if we break down the Race and Identity-Based Data into arrests made in 2020 and 2021, we do not see a significant difference in the demographic characteristics of the arrested population. Most notably, Black individuals remain the highest represented minority group among the arrested population, with a slight increase from 2020. Likewise, adults between the ages of 25 to 44 continue to make up the majority of arrests. The same observation can also be made among the male population. Due to the lack of significant differences, the remainder of the descriptive analysis will look at the Race and Identity-Based Data as a whole.

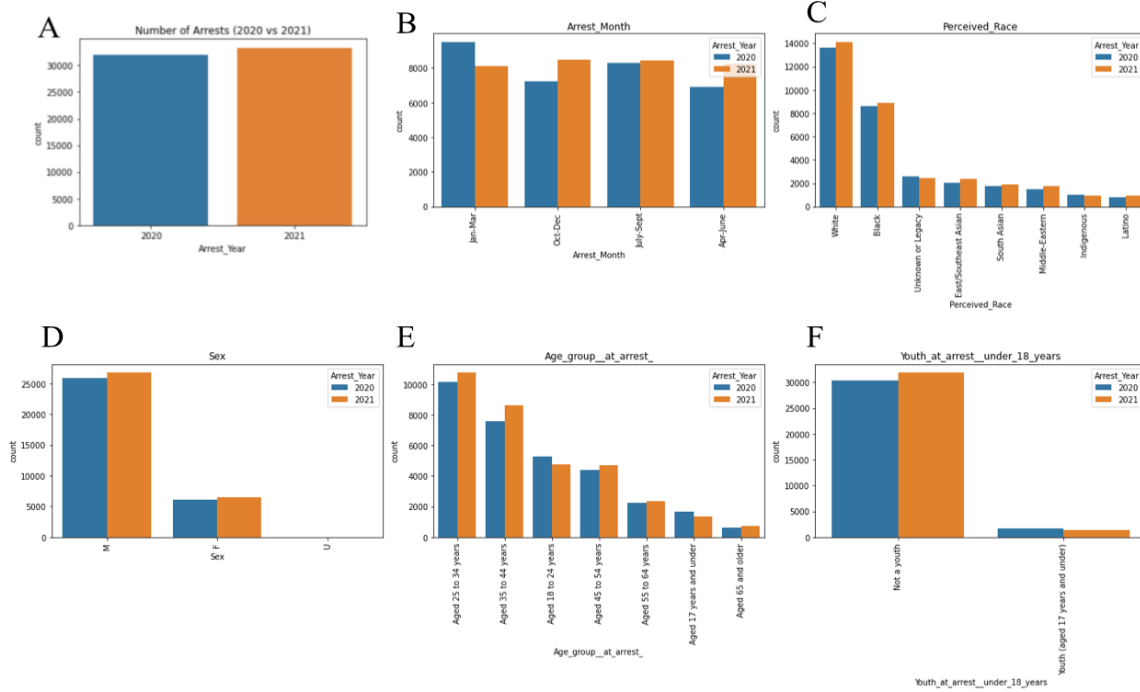


Figure 1. Overall frequency of arrests in 2020 and in 2021 (A). We further depict the frequency of arrest by arrest month (B), perceived race (C), sex (D), age group at the time of arrest (E), and youth status (F).

To further highlight two key demographic characteristics and its role in the discretionary nature of police arrests, we depict the average number of arrests with respect to perceived race (our variable of interest) and sex (one of our control variables). As illustrated in Figure 2A, we can see that the Indigenous population has the highest average number of arrests over 2020 and 2021 across all known racial subgroups and among minorities. This is true across both male and female populations. Such observations are followed by the Black racial subgroup. The distribution of the number of arrests per racial group can also be seen in Figure 2B.

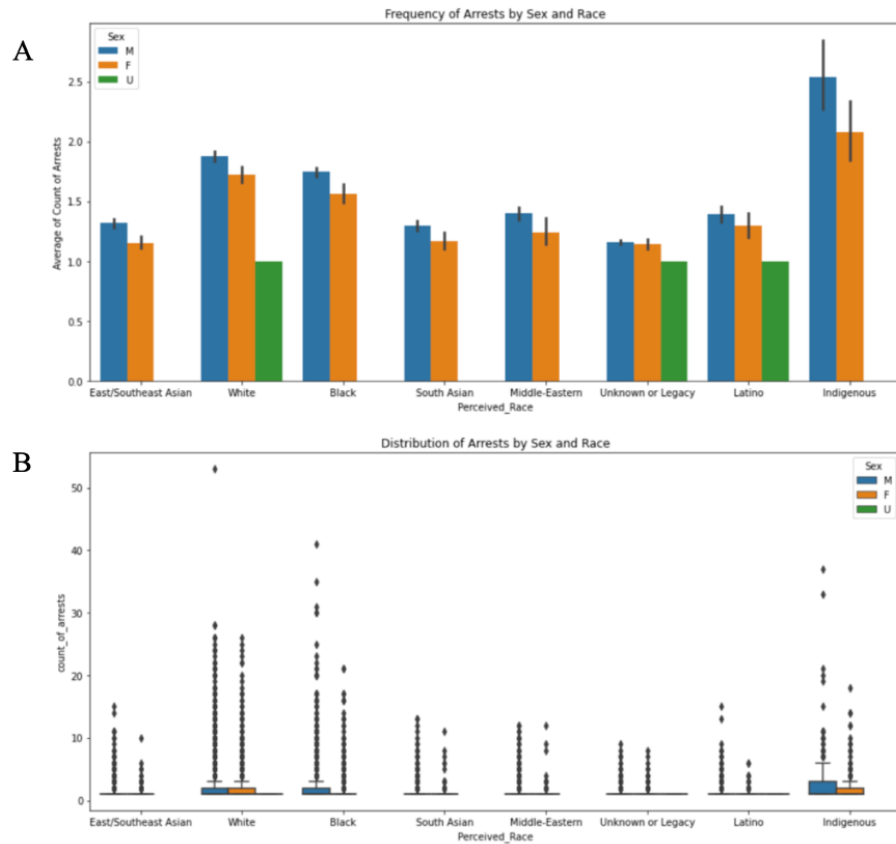


Figure 2. Frequency of Arrests by Sex and Perceived Race (A) and Distribution of Arrests by Sex and Perceived Race (B). All instances are heavily skewed to the right.

Further analysis also found that the majority of arrests have been a result of an assault and other crimes against persons (Figure 3). This is followed by incidences of robbery and theft and arrest warrants for detainment.

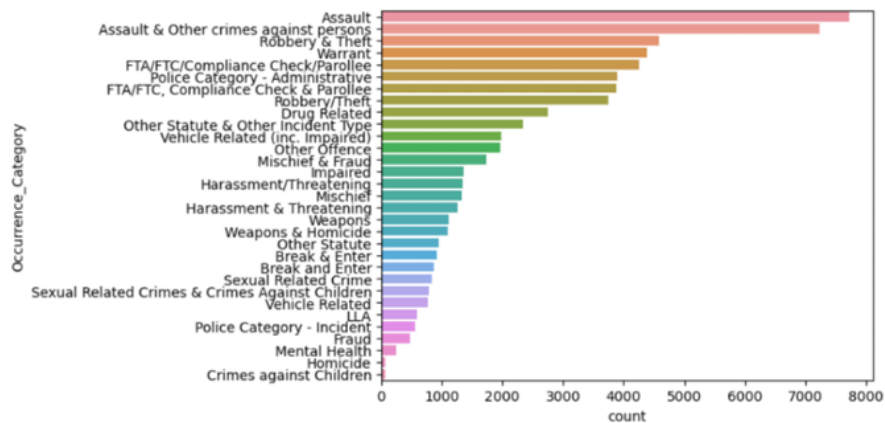


Figure 3. Frequency of Arrests by Occurrence Category

Encounter-Specific Variables

As previously discussed in the literature review, we found that the discretionary nature of an arrest outcome is heavily included by race, age, and sex. However, studies have also acknowledged the role of situational- or encounter-specific variables that contribute to the arrest outcome. As a result, we conducted further analysis into the frequency of occurrence with respect to encounter-specific variables by perceived race and sex, respectively. Such variables include presence of a concealed weapon and the specific behavior exhibited by the individual in question at the time of arrest, specifically whether they resisted arrests or assaulted a law enforcement officer during detainment. As illustrated in Table 2, results of the descriptive analysis found that among all minority groups, at least one-fourth of the total arrests having involved hostile behaviour including resistance to arrest, combative, and assault of an officer, involved an arrest of a Black individual at 31.8%, 29%, and 34.5%, respectively. Furthermore, Black persons also represented 36.8% of the total arrests that involved concealment of a weapon over the course of 2020 and 2021.

Table 2: Frequency of Encounter-Specific Variables by Perceived Race

	Concealed Weapon		Combative		Resistance	
	<u>Count</u>	<u>%</u>	<u>Count</u>	<u>%</u>	<u>Count</u>	<u>%</u>
Black	98	36.8	836	29	796	31.8
East/Southeast Asian	15	5.6	123	4.3	117	4.7
Indigenous	7	2.6	162	5.6	122	4.9
Latino	4	1.5	72	2.5	61	2.4
Middle eastern	13	4.9	122	4.2	103	4.1
South Asian	12	4.5	126	4.4	90	3.6
Unknown Or Legacy	14	5.3	154	5.3	140	5.6
White	103	38.7	1286	44.6	1073	42.9
	Mental Instability		Assault		Cooperative	
	<u>Count</u>	<u>%</u>	<u>Count</u>	<u>%</u>	<u>Count</u>	<u>%</u>
Black	665	30.5	143	34.5	7466	25.6
East/Southeast Asian	106	4.9	10	2.4	2268	7.8
Indigenous	78	3.6	23	5.5	762	2.6
Latino	42	1.9	12	2.9	897	3.1
Middle eastern	110	5	16	3.9	1514	5.2
South Asian	82	3.8	14	3.4	1667	5.7
Unknown Or Legacy	103	4.7	26	6.3	2220	7.6
White	993	45.6	171	41.2	12315	42.3

Such findings are further corroborated in Figure 4, which illustrates the breakdown of actions taken at arrests within each racial subgroup. While, majority of arrest outcomes have been cooperative across all perceived races, Indigenous and Black populations had the least relative proportion of cooperative arrests made within those populations at 66% and 74.6%, respectively. More specifically, both racial minorities had the highest relative frequency of having engaged in more hostile behaviour with 14% of the total arrests made on the Indigenous community having displayed combative behaviour and 10.6% resisting arrest. With respect to total arrests made on Black persons over the course of 2020 and 2021, 8.4% were combative and 8.0% resisted arrest.

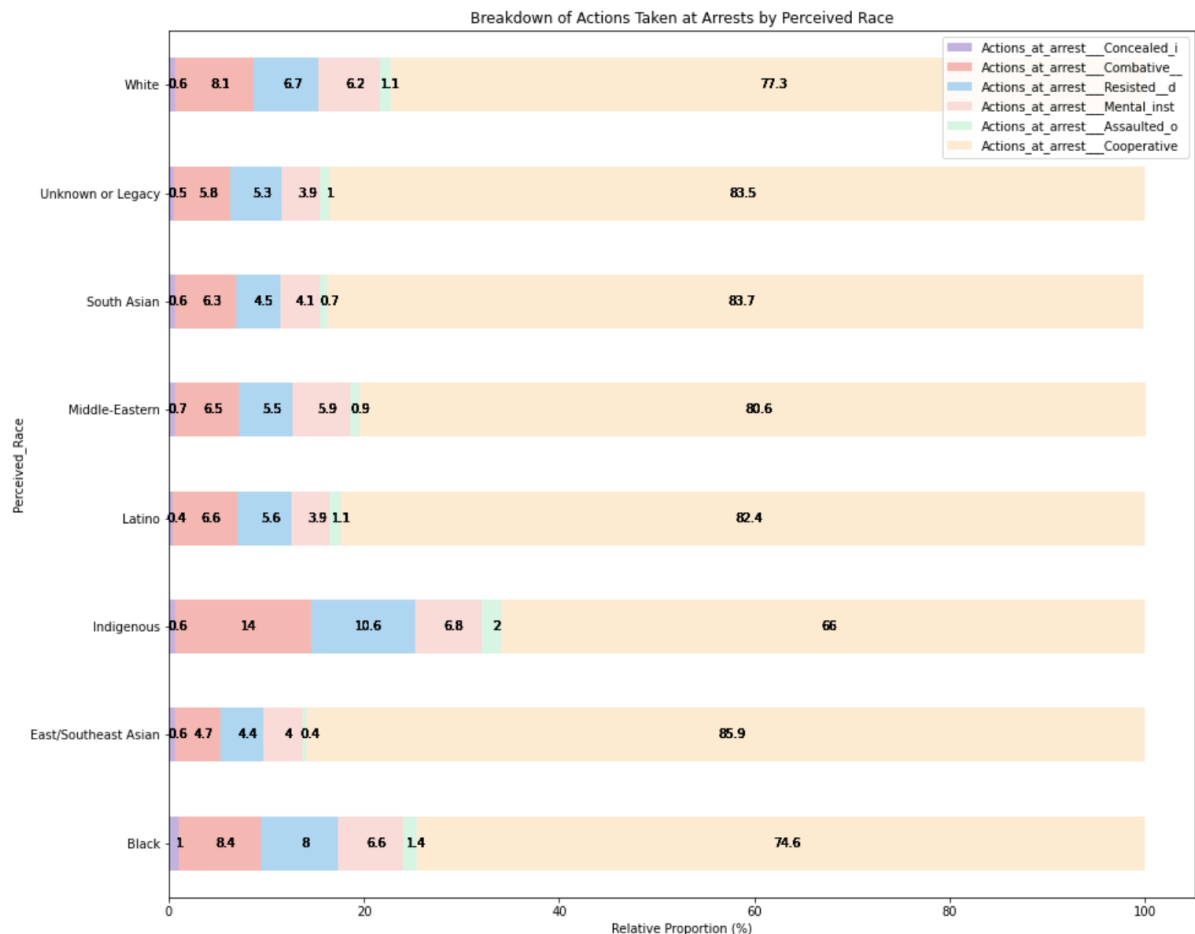


Figure 4. Breakdown of Actions Taken During Arrest Within Each Racial Group

If we further break the frequency of occurrence with respect to encounter-specific variables by sex (Table 3), it can be evidently observed that males represented majority of the total Toronto Police arrests that involved a concealed weapon (92.1%), and of the total arrests involving aggressive behavior with a law enforcement officer during detainment including assault against an officer (74.7%), resistance against an arrest (82.7%), and combative (80.4%). Important to note, while females appear to be less likely to possess a concealed weapon (7.9%), we see an increase in the frequency of arrests involving assault to an officer (25.3%) and

incidents involving mental instability (25.3%) among females. Nonetheless, the majority of arrests have been cooperative.

Table 3. Frequency of Encounter-Specific Variables by Sex

	Male		Female		Unspecified	
	Count	%	Count	%	Count	%
Concealed Weapon	245	92.1	21	7.9	0	0
Combative	2317	80.4	563	19.5	1	0.04
Resistance	2069	82.7	433	17.3	0	0
Mental Instability	1672	76.7	505	23.2	2	0.9
Assault	310	74.7	105	25.3	0	0
Cooperative	23676	81.3	5430	18.7	3	0.01

Arrests Leading to Strip Searches and Bookings

The 2020 race-based data reported by the Toronto Police Services (TPS), saw an over-representation of the use-of-force incidents and strip searches among these groups. To further find incidences of racial discrimination beyond initial detainment by a law-enforcement officer, we further analyzed the proportion of arrests leading to strip searches and bookings by TPS, with respect to an individual's perceived race and Sex. As shown in Table 4 and in congruence with previous TPS reporting described in the literature review, we can see that members of Indigenous (15.8%) and Black (13.9%) community are more likely to be strip searched as a result of an arrest. These same groups were also found to result in most bookings with 56.1% of arrests of Black persons leading to a booking, and 56.9% of Indigenous arrests.

Table 4. Proportion of arrests leading to bookings and strip searches

Feature	Booked	Strip Searches	Proportion of Arrests (%)	Proportion of Booked (%)	Proportion of Strip Searches (%)
Perceived Race					
Black (n=17,518)	9833	2434	26.8	56.1	13.9
East/Southeast Asian (n=4412)	2110	341	6.8	47.8	7.7
Indigenous (n=1932)	1099	306	3	56.9	15.8
Latino (n=1768)	974	132	2.7	55.1	7.5
Middle Eastern (n=3237)	1649	228	5	50.9	7
South Asian (n=3613)	1789	257	5.5	49.5	7.1
Unknown or Legacy (n=5059)	2491	537	7.8	49.2	10.6
White (27,713)	14532	3566	42.5	52.4	12

Sex					
Female (n=12,609)	5581	1283	19.3	44.3	10.2
Male (n=52,634)	28894	6518	80.7	54.9	12.4
Unspecified (n=9)	2	0	0	22.2	0

3.2. T-tests

3.2.1. T-test: Average Number of Arrests in 2020 vs 2021

As previously mentioned, the first objective of this study is to investigate the racial disparity of Toronto police arrests. However, to better discern and isolate the effect of perceived race on arrest outcome, we want to study this relationship independent of previous arrest history. As the present study uses data pertaining to arrests occurring in both 2020 and 2021, we will focus our investigation on arrest outcomes in 2021, while controlling for arrests made in 2020. To do so, we first want to determine whether there exists significant differences in the number of arrests made in 2020 and 2021. Therefore, we conduct a Welch's two-sample t-test to identify such significance. The following hypotheses are as followed:

- **Null Hypothesis:** There is no significant difference in the average number of arrests in 2020 and 2021.
- **Alternative Hypothesis:** There is a significant difference in the average number of arrests in 2020 and 2021.

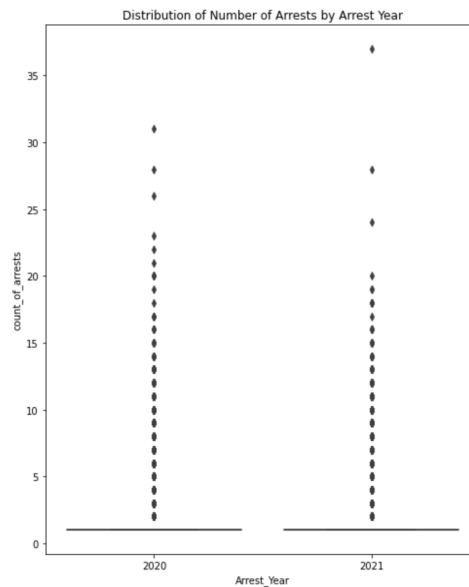


Figure 5. Distribution of Number of Arrest by Arrest Year

Assumptions check: Prior to conducting a two-sample t-test, the following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests) with respect to each arrest year, and (2) homogeneity of variances using Levene's test. Results of the box plot depict a non-normal distribution of the count of arrests by arrest year (Figure 5). However, as the sample sizes (n) are sufficiently large, such that each group has an $n > 50$, we can reasonably assume normality. Results of the Levene's test also found that the homogeneity of variances was violated (statistic = 7.176, p.value = 0.007) at a 95% confidence level. Therefore, a Welch two-sample t-test for unequal variances was to be conducted.

Results: At an alpha significance of 0.05 and a confidence level of 95%, results of the Welch's two-sample t-test found that there is sufficient evidence to suggest that there is a difference in the average number of police arrests between 2020 and 2021 (statistic = 84.074, p-value < 0.01). Therefore, we reject the null and conclude that there is a difference in the average arrest numbers which should be further explored.

3.2.2. T-test: Average Number of Arrests in 2020 vs 2021 Among Black Minorities

Previous published analyses on the Race and Identity-based data found alarming disproportionate policing of Black minorities, including the 2020 interim report by the Ontario Human Rights Commission's (OHRC) inquiry into racial profiling and racial discrimination of Black persons by the Toronto Police Service (TPS) by Wortley and Jung. To further highlight potential increases in the average number of arrests on Black persons from 2020 to 2021, we conduct an upper, one-tailed two-sample t-test. The hypothesis for such question is as followed:

- **Null Hypothesis:** Among the Black minority racial subgroup, the average number of arrests in 2021 is not higher than in 2020.
- **Alternative Hypothesis:** Among the Black minority racial subgroup, there is a higher average number of arrests in 2021 than in 2020.

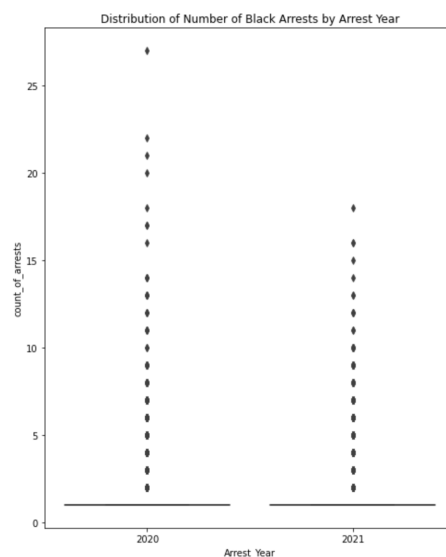


Figure 6: Distribution of Number of Black Arrest by Arrest Year

Assumptions check: The following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests among Black persons) with respect to each arrest year, and (2) homogeneity of variances using Levene's test. Results of the box plot depict a non-normal distribution of the count of arrests by arrest year (Figure 6). However, as the sample sizes (n) are sufficiently large, such that each group has an $n > 50$, we can reasonably assume normality. Results of the Levene's test also found that the homogeneity of variances to be satisfied (statistic = 1.936, p.value = 0.164) at a 95% confidence level. Therefore, we were able to proceed with a two-sample t-test for equal variances.

Results: At an alpha significance of 0.05 and a confidence level of 95%, results of the upper, one-tailed two-sample t-test found that there is sufficient evidence to suggest that there is a higher number of Black arrests in 2021 than in 2020 (statistic = 39.038, p-value < 0.01). Therefore, we reject the null and conclude that there is a higher average number of Black arrests in 2021 than in 2020.

3.2.3. T-test: Average Number of Arrests in 2020 vs 2021 Among Indigenous Minorities

Similar to the Black minority subgroup, the Indigenous has also been subject to disproportionate policing by Toronto Police Services (TPS). To highlight potential increases in the average number of arrests on Indigenous persons from 2020 to 2021, we conduct an upper, one-tailed two-sample t-test. The hypothesis for such question is as followed:

- **Null Hypothesis:** Among the Indigenous minority racial subgroup, the average number of arrests in 2021 is not higher than in 2021.
- **Alternative Hypothesis:** Among the Indigenous minority racial subgroup, there is a higher average number of arrests in 2021 than in 2020.

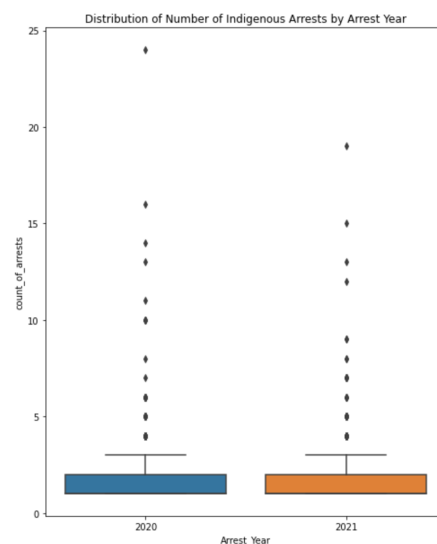


Figure 7: Distribution of Number of Indigenous Arrest by Arrest Year

Assumptions check: The following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests among Indigenous persons) with respect to each arrest year, and (2) homogeneity of variances using Levene's test. Results of the box plot depict a non-normal distribution of the count of arrests by arrest year (Figure 7). However, as the sample sizes (n) are sufficiently large, such that each group has an $n > 50$, we can reasonably assume normality. Results of the Levene's test also found that the homogeneity of variances to be unsatisfied (statistic = 0.008, p.value = 0.931) at a 95% confidence level. Therefore, we were able to proceed with a two-sample t-test for equal variances.

Results: At an alpha significance of 0.05 and a confidence level of 95%, results of the upper, one-tailed two-sample t-test found that there is sufficient evidence to suggest that there is a higher number of Black arrests in 2021 than in 2020 (statistic = 6.848, p-value < 0.01). Therefore, we reject the null and conclude that there is a higher average number of Indigenous arrests in 2021 than in 2020.

3.2.4. T-test: Average Number of Arrests of Suspects Concealing Items vs Not Concealing Items

The literature review has indicated that the disrespectful and hostile behavior of suspects could potentially increase their chances of being arrested. However, it remains unclear how specific disrespectful or hostile actions contribute to the variation in the number of arrests. To address this gap, we will analyze our dataset by categorizing suspect behavior into distinct aspects, such as concealing items, combative actions, resistance, mental instability or suicidal tendencies, assaulting officers, and cooperation. We will examine how each aspect affects the average number of arrests and investigate potential correlations between these behaviors and arrest outcomes. To highlight potential differences in the average number of arrests between suspects concealed items and didn't conceal items, we conduct an upper, one-tailed two-sample t-test. The hypothesis for such question is as followed:

- **Null Hypothesis:** There is no significant difference in the average number of police arrests between suspects concealing items or not.
- **Alternative Hypothesis:** There is a significant difference in the average number of police arrests between suspects concealing items or not.

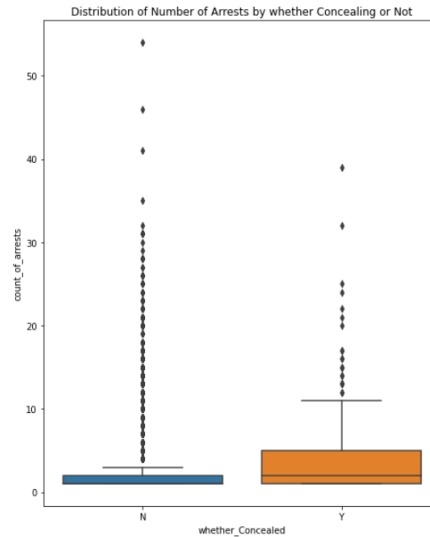


Figure 8: Distribution of Number of Arrest by Whether Suspect Concealed Items or Not

Assumptions check: The following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests) with respect to all suspects, including those who concealed items and not concealed items, and (2) homogeneity of variances using Levene's test. Results of the box plot depict a non-normal distribution of the count of arrests by concealed items or not (Figure 8). However, as the sample sizes (n) are sufficiently large, such that each group has an $n > 50$, we can reasonably assume normality. Results of the Levene's test also found that the homogeneity of variances was violated (statistic = 541.831, p -value = 0.000) at a 95% confidence level. Therefore, we conduct a Welch's two-sample t-test to identify such significance.

Results: At an alpha significance of 0.05 and a confidence level of 95%, results of the Welch's two-sample t-test found that there is sufficient evidence to suggest that there is a difference in the average number of police arrests between suspects who concealed items versus who didn't (statistic = 7.767, p -value < 0.01). Therefore, we reject the null and conclude that there is a difference in the average arrest numbers which should be further explored.

3.2.5. *T-test: Average Number of Arrests of Suspects Performing Combative Behaviour vs Suspects Not Performing Combative Behaviour:*

Performing combative, violent or spitter/biter behaviour is a kind of hostile actions to police officers, and we will check how the average number of arrests on suspects performing such behaviour differ from those who didn't do so, and we conduct an upper, one-tailed two-sample t-test. The hypothesis for this question is as followed:

- **Null Hypothesis:** There is no significant difference in the average number of police arrests between suspects whether performing combative, violent or spitter/biter behaviour or not.

- **Alternative Hypothesis:** There is a significant difference in the average number of police arrests between suspects whether performing combative, violent or spitter/biter behaviour or not.

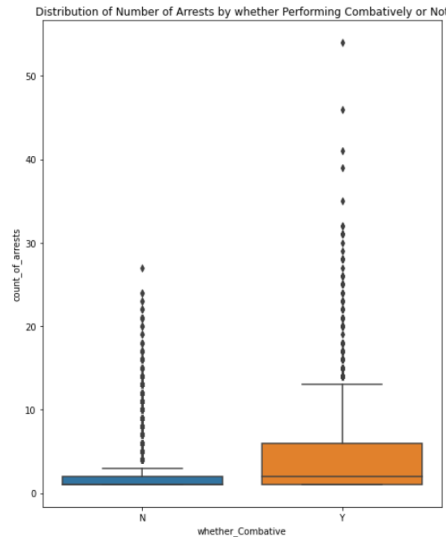


Figure 9: Distribution of Number of Arrest by Whether Suspect Performed Combative Behaviour or Not

Assumptions check: The following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests) with respect to all suspects, including those who performed combative actions and didn't perform combative actions, and (2) homogeneity of variances using Levene's test. Results of the box plot depict a non-normal distribution of the count of arrests by whether the suspect performed combative actions or not (Figure 9). However, as the sample sizes (n) are sufficiently large, such that each group has an $n > 50$, we can reasonably assume normality. Results of the Levene's test also found that the homogeneity of variances was violated (statistic = 6226.413, p -value = 0.000) at a 95% confidence level. Therefore, we conduct a Welch's two-sample t-test to identify such significance.

Results: At an alpha significance of 0.05 and a confidence level of 95%, results of the Welch's two-sample t-test found that there is sufficient evidence to suggest that there is a difference in the average number of police arrests between suspects who performed combative actions versus who didn't (statistic = 26.704, p -value < 0.01). Therefore, we reject the null and conclude that there is a difference in the average arrest numbers which should be further explored.

3.2.6. T-test: Average Number of Arrests of Suspects Resisting vs Not Resisting

Similarly, resisting, defending or even performing the tendency of escaping is regarded as a hostile action towards police officers, and we'll proceed to check if the average number of arrests differ based on these specific actions, we conduct an upper, one-tailed two-sample t-test. The hypothesis for this question is as followed:

- **Null Hypothesis:** There is no significant difference in the average number of police arrests between suspects who resisted, indicated defensive and escaping risk or not.
- **Alternative Hypothesis:** There is a significant difference in the average number of police arrests between suspects who resisted, indicated defensive and escaping risk or not.

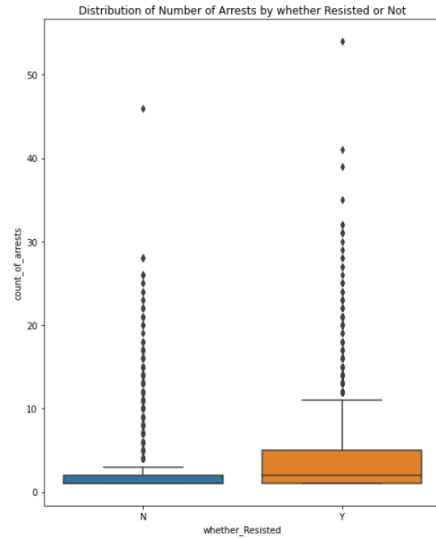


Figure 10: Distribution of Number of Arrest by Whether Suspect Resisted or Not

Assumptions check: The following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests) with respect to all suspects, including those who resisted and didn't resist, and (2) homogeneity of variances using Levene's test. Results of the box plot depict a non-normal distribution of the count of arrests by whether the suspect resisted or not (Figure 10). However, as the sample sizes (n) are sufficiently large, such that each group has an $n > 50$, we can reasonably assume normality. Results of the Levene's test also found that the homogeneity of variances was violated (statistic = 5343.500, p -value = 0.000) at a 95% confidence level. Therefore, we conduct a Welch's two-sample t-test to identify such significance.

Results: At an alpha significance of 0.05 and a confidence level of 95%, results of the Welch's two-sample t-test found that there is sufficient evidence to suggest that there is a difference in the average number of police arrests between suspects who resisted versus who didn't resist (statistic = 24.724, p -value < 0.01). Therefore, we reject the null and conclude that there is a difference in the average arrest numbers which should be further explored.

3.2.7. T-test: Average Number of Arrests of Suspects with Mental Instabilities or Suicidal Tendencies vs Suspects without Mental Instabilities or Suicidal Tendencies

As we've mentioned in the literature review, suspects with mental instabilities are more likely to perform disrespectful or even hostile actions. So we'll check if the average number of arrests differ between the group with mental illness and the groups without such illness, and we conduct an upper, one-tailed two-sample t-test. The hypothesis for this question is as followed:

- **Null Hypothesis:** There is no significant difference in the average number of police arrests between suspects who have mental instability or suicidal tendencies or not.
- **Alternative Hypothesis:** There is a significant difference in the average number of police arrests between suspects who have mental instability or suicidal tendencies or not.

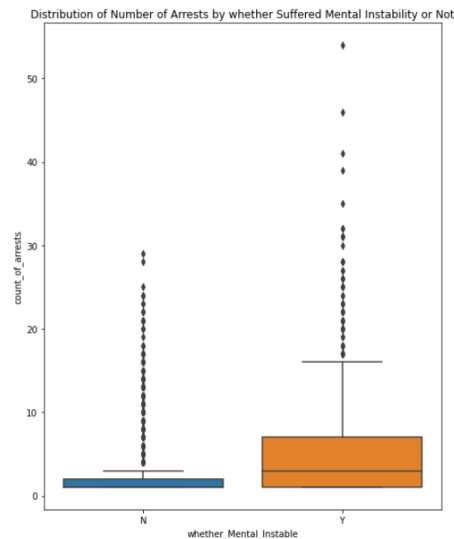


Figure 11: Distribution of Number of Arrest by Whether Suspect Had Mental Instabilities or Not

Assumptions check: The following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests) with respect to all suspects, including those with and without mental instabilities or suicidal tendencies, and (2) homogeneity of variances using Levene's test. Results of the box plot depict a non-normal distribution of the count of arrests by whether the suspect had mental instabilities or not (Figure 11). However, as the sample sizes (n) are sufficiently large, such that each group has an $n > 50$, we can reasonably assume normality. Results of the Levene's test also found that the homogeneity of variances was violated (statistic = 6360.450, p -value = 0.000) at a 95% confidence level. Therefore, we conduct a Welch's two-sample t-test to identify such significance.

Results: At an alpha significance of 0.05 and a confidence level of 95%, results of the Welch's two-sample t-test found that there is sufficient evidence to suggest that there is a difference in the average number of police arrests between suspects who concealed items versus who didn't (statistic = 25.984, p -value < 0.01). Therefore, we reject the null and conclude that there is a difference in the average arrest numbers which should be further explored.

3.2.8. T-test: Average Number of Arrests of Suspects Assaulting Officers vs Suspects Not Assaulting Officers

The other hostile action is directly assaulting officers, and we'll check if the average number of arrests differ due to this action, and we conduct an upper, one-tailed two-sample t-test. The hypothesis for such question is as followed:

- **Null Hypothesis:** There is no significant difference in the average number of police arrests between suspects who assaulted officers or not.
- **Alternative Hypothesis:** There is a significant difference in the average number of police arrests between suspects who assaulted officers or not.

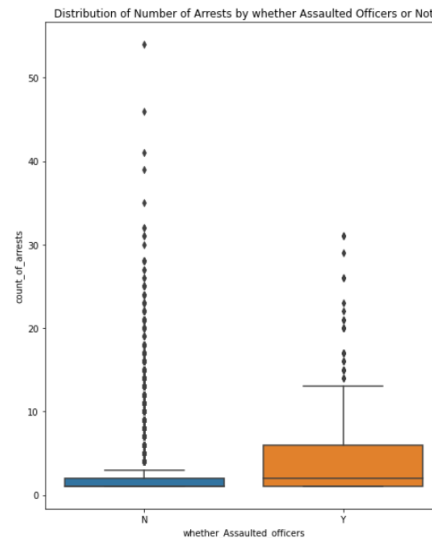


Figure 12: Distribution of Number of Arrest by Whether Suspect Assaulted Officers or Not

Assumptions check: The following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests) with respect to all suspects, including those who assaulted officers and who didn't, and (2) homogeneity of variances using Levene's test. Results of the box plot depict a non-normal distribution of the count of arrests by whether the suspect assaulted officers or not (Figure 12). However, as the sample sizes (n) are sufficiently large, such that each group has an $n > 50$, we can reasonably assume normality. Results of the Levene's test also found that the homogeneity of variances was violated (statistic = 768.347, $p.value = 0.000$) at a 95% confidence level. Therefore, we conduct a Welch's two-sample t-test to identify such significance.

Results: At an alpha significance of 0.05 and a confidence level of 95%, results of the Welch's two-sample t-test found that there is sufficient evidence to suggest that there is a difference in the average number of police arrests between suspects who assaulted officers and who didn't (statistic = 10.419, $p-value < 0.01$). Therefore, we reject the null and conclude that there is a difference in the average arrest numbers which should be further explored.

3.2.9. T-test: Average Number of Arrests of Suspects Cooperating vs Suspects Not Cooperating.

Finally, we aim to investigate whether the level of cooperation exhibited by suspects during the arrest process has an impact on the average number of arrests in a given group. Still, we conduct an upper, one-tailed two-sample t-test. The hypothesis for such question is as followed:

- **Null Hypothesis:** There is no significant difference in the average number of police arrests between suspects who cooperated or not.
- **Alternative Hypothesis:** There is a significant difference in the average number of police arrests between suspects who cooperated or not.

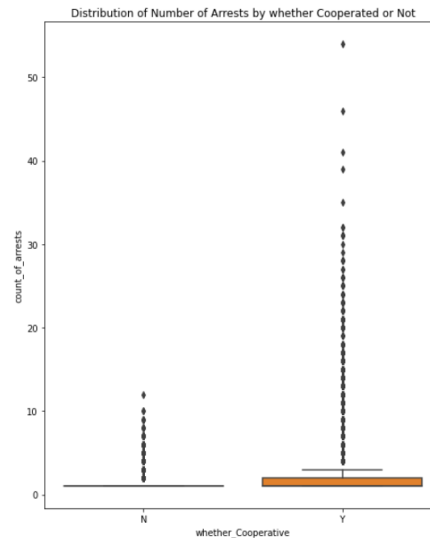


Figure 13: Distribution of Number of Arrest by Whether Suspect Cooperated or Not

Assumptions check: The following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests) with respect to all suspects, including those who cooperated and who didn't cooperate, and (2) homogeneity of variances using Levene's test. Results of the box plot depict a non-normal distribution of the count of arrests by whether the suspect cooperated or not (Figure 13). However, as the sample sizes (n) are sufficiently large, such that each group has an $n > 50$, we can reasonably assume normality. Results of the Levene's test also found that the homogeneity of variances was violated (statistic = 4138.082, p -value = 0.000) at a 95% confidence level. Therefore, we conduct a Welch's two-sample t-test to identify such significance.

Results: At an alpha significance of 0.05 and a confidence level of 95%, results of the Welch's two-sample t-test found that there is sufficient evidence to suggest that there is a difference in the average number of police arrests between suspects who cooperated and who didn't (statistic = 51.153, p -value < 0.01). Therefore, we reject the null and conclude that there is a difference in the average arrest numbers which should be further explored.

3.2.10. T-test: Average Number of Arrests of Males vs Females.

Additionally, we would like to check if the average number of arrests differ between males and females. Still, we conduct an upper, one-tailed two-sample t-test. The hypothesis for such question is as followed:

- **Null Hypothesis:** There is no significant difference in the average number of police arrests between males and females.
- **Alternative Hypothesis:** There is a significant difference in the average number of police arrests between males and females.

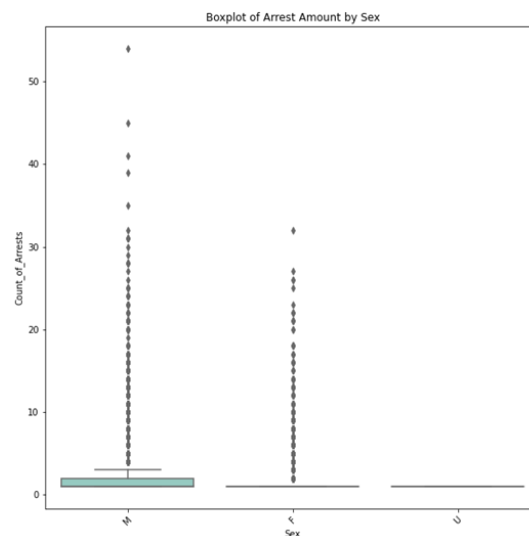


Figure 14: Distribution of Number of Arrest by Sex

Assumptions check: The following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests) with respect to males, females and undefined groups, and (2) homogeneity of variances using Levene's test. Results of the box plot depict a non-normal distribution of the count of arrests by sex (Figure 14). However, as the sample sizes (n) are sufficiently large, such that each group has an $n > 50$, we can reasonably assume normality. Additionally, since the sample size of the undefined group is too small, we remove it from our t-test. Results of the Levene's test also found that the homogeneity of variances was violated (statistic = 28.20, p.value = 0.000) at a 95% confidence level. Therefore, we conduct a Welch's two-sample t-test to identify such significance.

Results: At an alpha significance of 0.05 and a confidence level of 95%, results of the Welch's two-sample t-test found that there is sufficient evidence to suggest that there is a difference in the average number of police arrests between males and females (statistic = 5.385, p-value < 0.01). Therefore, we reject the null and conclude that there is a difference in the average arrest numbers which should be further explored.

3.2.11. T-test: Average Number of Arrests of Juveniles vs Adults.

Lastly, we check if there is an average number of arrests based on the age feature, and in our case, we split that into two groups with juveniles and adults separately. Still, we conduct an upper, one-tailed two-sample t-test. The hypothesis for such question is as followed:

- **Null Hypothesis:** There is no significant difference in the average number of police arrests between juveniles and adults.
- **Alternative Hypothesis:** There is a significant difference in the average number of police arrests between juveniles and adults.

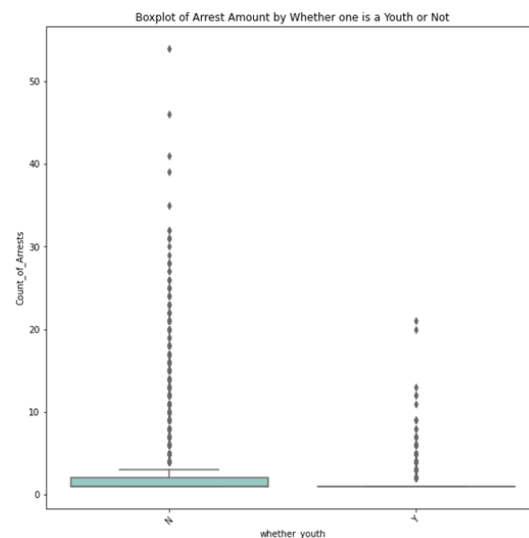


Figure 15: Distribution of Number of Arrest between Juveniles and Adults

Assumptions check: The following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests) with respect to all age groups, including juveniles and adults, and (2) homogeneity of variances using Levene's test. Results of the box plot depict a non-normal distribution of the count of arrests by whether an individual is juvenile or not (Figure 15). However, as the sample sizes (n) are sufficiently large, such that each group has an $n > 50$, we can reasonably assume normality. Results of the Levene's test also found that the homogeneity of variances was violated (statistic = 212.606, p -value = 0.000) at a 95% confidence level. Therefore, we conduct a Welch's two-sample t-test to identify such significance.

Results: At an alpha significance of 0.05 and a confidence level of 95%, results of the Welch's two-sample t-test found that there is sufficient evidence to suggest that there is a difference in the average number of police arrests between juveniles and adults (statistic = -18.33, p -value < 0.01). Therefore, we reject the null and conclude that there is a difference in the average arrest numbers which should be further explored.

4. Research Design & Methodology

4.1. Dataset Description

The dataset of the research project is extracted from Toronto Police Service Website, which documents information related to all arrests and strip searches in the Great Toronto Area, covering the period from April 2020 to December 2021. The dataset is publicly available and can be used to gain an overview of the arrests, especially strip search occurrences between 2020 and 2021. Learning these insights may also support the Toronto Police in arranging police resources more efficiently. There are 65,276 records in the dataset. Each record indicates an arrest, and it also contains many details, including the unique arrest ID, some socio-demographic features of the person being arrested, arrest-related information such as occurrence type, location, whether strip search and booking occurred, actions at arrests, and reasons of search. Referring to the official website explanation, a strip search refers to a search conducted by a police officer on a person, which includes the removal of some or all clothing and a visual inspection of the body.

The overall dataset is clear and intuitive except for the following two sections: Firstly, the "Booked" variable is a binary categorical variable, with 1 indicating a booking and 0 otherwise. However, from the official abstract, when the corresponding strip search occurred, a booking indeed took place, whereas the recording could be a wrong 0. Secondly, in the two variables relating to age, "Age_group__at_arrest_" and "Youth_at_arrest__under_18_years", there seem to be two subgroups indicating the same groups. Under the former, "Aged 17 years and under" and "Aged 17 years and younger" are two groups, whereas "Youth (aged 17 and younger)" and "Youth (aged 17 years and under)" are two groups under "Youth_at_arrest__under_18_years". As we are not using booking in our research, it won't be a problem in our study. As for whether one is youth, even the two youth groups are having different annotations, they both belong to the group of youth, so we recode them as the same group, and doing so doesn't affect our study and findings.

The dataset contains six situational variables indicating different actions taken by suspects during their arrest. These include concealing items (recorded as "Actions_at_arrest__Concealed_i"), combative or violent behavior (recorded as "Actions_at_arrest__Combative_"), resisting, defensive or attempting to escape (recorded as "Actions_at_arrest__Resisted_d"), displaying signs of mental instability or suicidal tendencies (recorded as "Actions_at_arrest__Mental_inst"), and assaulting officers (recorded as "Actions_at_arrest__Assaulted_o"). The last variable, "Actions_at_arrest__Cooperative", indicates a cooperative attitude and is the only variable with a positive connotation. All variables are binary with a value of 1 indicating "yes" and 0 indicating "no". There are no missing values, and the variables can be used directly.

Overall, the dataset provides detailed information related to arrests and strip searches in the Greater Toronto Area, which can be used to study the criminal distributions by socio-demographic features and those situational variables. Learning that can explicitly present the phenomenon and better support police resource allocation.

4.2. Research Questions

We have previously discussed in the literature review of the role in which law enforcement officers play as the point of entry into the criminal justice system (Spencer et al., 2016), and the potential subsequent implications that can follow. We also discussed the

discretionary nature of police arrests and the implications in which pervasive racial stereotypes and prejudice can have on exercising fair and unbiased law enforcement as a by-product of other demographic characteristics (age and sex) and situational-specific factors. As such, this serves as a critical point in addressing any existing discrimination in the system that can otherwise lead to wrongful convictions with profound life-time consequences. Therefore, in response to the alarming disproportionate policing over Black, Indigenous, and other marginalized populations, as well as the significant consequences of the aftermath, the present study aims to uncover the race differential and potential impact from demeanors in the Toronto police arrests.

The present study has two objectives: (1) to isolate the role of perceived race, as a reflection of implicit bias, on arrest outcomes that are independent of previous arrest history, as we recognize that people who have been previously arrested are more likely to be arrested; and (2) how the likelihood of being arrested is influenced by demographic characteristics (age, sex) and actions at arrests (i.e., encounter-specific variables such as resistance to arrests and assault on a law enforcement officer). Therefore, we formally introduce the following questions to be addressed in this study:

1. How does an individual's perceived race influence the number of police arrests, while controlling for previous arrests?
2. What is the relationship between demographic features including age and sex, and level of cooperation or aggression during an arrest, and the likelihood of an individual being arrested?

4.3. Data Analysis Method

4.3.1. Analysis of Covariance

To better discern and isolate the role of perceived race, as a reflection of the implicit bias, on arrest outcomes that are independent of previous arrest history and criminal activity, a one-way Analysis of Covariance (ANCOVA) will be conducted in determining whether significance differences exists in the average number of police arrests between each racial subgroup of perceived race among the arrested population recorded by Toronto Police. More specifically, we will be investigating the race differential on the number of arrests made by law enforcement in 2021, while controlling for previous accounts of arrests recorded in 2020, as we recognize that those who have been previously arrested are more likely to have repeated arrests. Therefore, this will allow for the comparison of the mean number of arrests in 2021 adjusted for arrests made in 2020, across multiple classes of an independent categorical variable with respect to the perceived race of the detained individual in question. Additionally, we also took the first instance of an arrested individual where there were multiple occurrences in the data that varied by perceived race.

Prior to conducting the ANCOVA, a power analysis was performed in determining the optimal sample size to observe an effect size of at least 0.2, 0.5, and 0.8, at an alpha significance of 0.05 and confidence level of 95% at which the aforementioned test is to be conducted at in investigating the relationship between perceived race and the average number of arrests while controlling for previous arrest history. We also aim for a high statistical power of 0.8, corresponding to the likelihood of finding an effect if there is an effect to be found. This also

means a lower probability of making a Type II Error of failing to detect our null hypothesis when it is indeed false. Figure 14 further depicts the statistical power of conducting our test as a function of effect size and sample size. Evidently, we can see that the minimum required sample size in order to observe even the smallest magnitude of effect with respect to an effect size of 0.2 and a statistical power of 0.8 is estimated to be 200 at a confidence level of 95% involving eight racial subgroups in our independent variable, perceived race. Our current sample size using the Race and Identified-Based data published by Toronto Police Service (TPS) consists of 21,614 unique arrests in 2021 after accounting for duplicate and missing records, thereby exceeding the required minimum sample size required to detect the smallest effect. Therefore, we are confident that our test has sufficient power that can otherwise lead to false negatives and will proceed with our ANCOVA model.

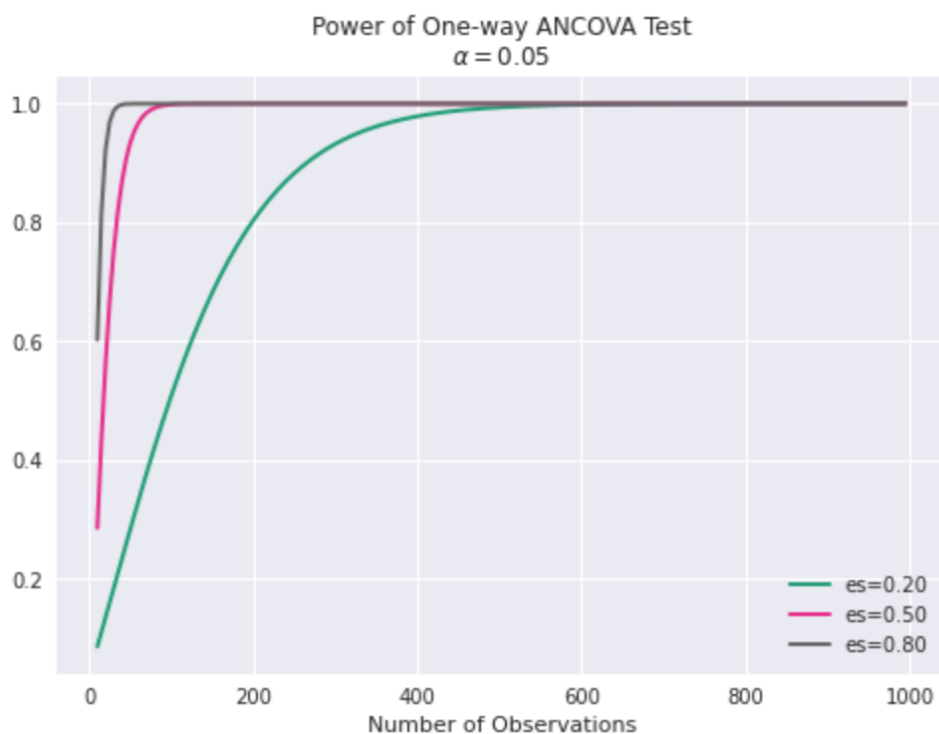


Figure 16. Power Curve

4.3.2. Analysis of Logistic Regression

In our second research question, we sought to explore the relationship between an individual's likelihood of being arrested and various factors including demographic features and demeanor during arrests. To accomplish this, we employed logistic regression as our statistical method. Initially, we had a dataset consisting of 65,276 records of arrests. To make our analysis more informative, we integrated the records based on unique person IDs, resulting in a dataset with 40,423 individuals and a count of arrests as a continuous variable.

To apply logistic regression, we transformed the count of arrests into a binary variable indicating whether an individual was arrested or not. To do this, we checked the distribution of the count of arrests and found it to be right-skewed. We then used the median as a boundary, such that any individual with a count of arrests greater than the median was considered to be arrested, while those with a count less than or equal to the median were considered not to be arrested.

In addition to the binary outcome variable, we included several predictor variables in our analysis, such as an individual's sex, age, and perceived race, as well as their record of demeanor during arrests. When summing the count of arrests and the records of different types of demeanor, we took the existence of a certain type of behavior as a predictor variable and examined how it affected an individual's probability of being arrested. To prepare demographic features, we implemented one-hot encoding for the perceived race and sex variables to transform them into binary variables, which could be directly applied to the logistic regression analysis. For the age variable, we split the data into two categories: adult and juvenile.

Overall, our analysis sought to shed light on the relationship between demographic and demeanor-related factors and an individual's likelihood of being arrested, with the goal of informing efforts to promote greater equity and fairness in the criminal justice system.

5. Results & Findings

5.1. One-way ANCOVA: Adjusted Mean Arrests Across Perceived Races

Hypothesis Model

To investigate whether significant differences exist in the mean number of arrests between perceived racial groups by Toronto Police Service (TPS) in 2021, while controlling for the number of previous arrests, a one-way ANCOVA was conducted. The following hypothesis model is as followed:

- **Null Hypothesis (H_0):** There is no significant difference in the average number of arrests by perceived race in 2021, while controlling for the number of arrests in the previous year in 2020.
- **Alternative Hypothesis (H_A):** There is a significant difference in the average number of arrests by perceived race in 2021, while controlling for the number of arrests in the previous year in 2020.

Assumptions Check

Prior to running the ANCOVA, the following assumptions and conditions were checked: (1) normality of the continuous dependent variable (number of police arrests in 2021) with respect to each racial subgroup of perceived race, and (2) homogeneity of variances using the Kolmogorov-Smirnov test, and Levene's test respectively. With the addition of the covariate, the following assumptions were also verified: (3) linearity between the covariate and the dependent variable at each level of the perceived race, and (4) homogeneity of regression slopes which assumes the effect between the covariate and the dependent variable to be the same across all levels of the independent variable. Results of the Kolmogorov-Smirnov and Levene's test, as

well as the test of homogeneity of regression slopes through significant interaction between the independent variable of perceived race and the covariate using ANOVA is summarized in Table 5. All tests were conducted at an alpha significance of 0.05.

Table 5. Summary of Kolmogorov-Smirnov, Levene, and ANOVA Interaction Results

	Statistic	p-value*
Kolmogorov-Smirnov**	0.841	<0.001
Levene	81.078	<0.001
ANOVA (Interaction Term)	9.736	<0.001

*Tested at an alpha significance of 0.05 (95% confidence level)

**For all racial subgroups in perceived race

Results of the Kolmogorov-Smirnov test found the normality to be violated for all racial subset (for all subsets: statistic = 0.841, $p < 0.001$) at a 95% confidence level. The test assumes as its null that the sample has the same distribution as the reference distribution, which we define as normal. Since the resulting p-value is less than 0.05, we therefore reject the null, such that the sample does not follow a normal distribution. Furthermore, results of the Levene's test indicate violation of the homogeneity of variance (statistic: 81.078, $p\text{-value} < 0.001$), which assumes equal variance as its null. This results in the violation of our second assumption.

With respect to the covariate, the test assumes a relationship between the number of arrests made in 2021 and the number of previous arrests recorded in 2020 to be linear. Figure 15 depicts a scatterplot of the covariate and the dependent variable for each racial subgroup of perceived race which evidently demonstrates violation of our third assumption. Lastly, results of the homogeneity of regression slopes with respect to the covariate also showed violation of such assumption (statistic=9.736, $p\text{-value} < 0.001$) (Table 5). This was determined by verifying the significance of the interaction term between the covariate and perceived race using ANOVA.

All assumptions were violated; however, for the purpose of this report, we will continue to proceed with the one-way ANCOVA analysis.

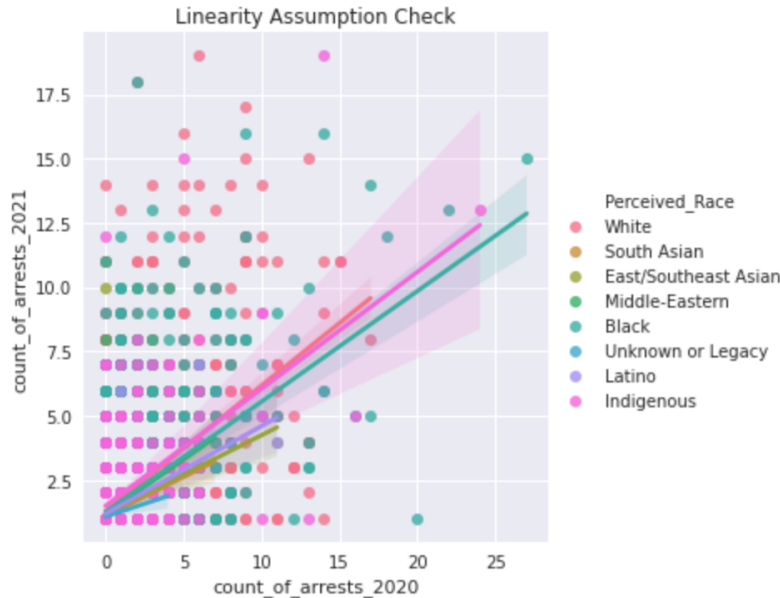


Figure 17. Linear assumption check of the number of arrests made in 2020 (covariate) and number of arrests made in 2021 (dependent variable) across all levels of perceived race (independent variable)

Interpreting ANCOVA Results

Results of the Analysis of Covariance (ANCOVA) found that there was sufficient evidence to suggest that there exists a difference in the adjusted mean number of arrests among the eight perceived racial subgroups defined by Toronto Police Service (TPS) with respect to Black, White, East/Southeast Asian, South Asian, Indigenous, and the Latino community. TPS also defines the category Unknown/Legacy for individuals whose ethnic background appears ambiguous. This was found to be significant at an alpha significance level of 0.05 (F-statistic = 909.6, $p\text{-value} < 0.001$). Therefore, we can reject the null hypothesis in which there are no significant differences in the adjusted mean number of police arrests in 2021 by perceived race, even when controlling for the number of arrests recorded in 2020.

Table 6 summarizes the results of the ANCOVA test where we further breakdown the implication of perceived race on the adjusted mean number of arrests in 2021. Most evidently, results of ANCOVA test further corroborate findings reported by the interim report by the Ontario Human Rights Commission's (OHRC) which identified disproportionate policing on the Black and Indigenous community. Results of our ANCOVA model found that individuals who were perceived to be Black were arrested on average 1.23 times in 2021, while controlling for previous arrest history and criminal activity reported in 2020 (coefficient = 1.231, $p < 0.001$, CI = [1.205, 1.258]), and were found to be statistically significantly higher compared to all other minority groups with the exception of the Indigenous population which were found to be arrested on average 0.261 ($p < 0.001$, [0.164, 0.358]) more than their Black counterparts in 2021 even after adjusting for previous arrest history. Such results are also consistent with initial exploratory analysis which found Indigenous members of the Toronto community to experience a higher average number of arrests over 2020 and 2021 for both males and females across all racial subgroups (Figure 2A). Additionally, we also see that compared to their Black counterparts individuals who were perceived as Latino were found to be arrested 0.076 less times

(p-value=0.06, CI=[-0.154, 0.003]) ; however, this was not found to be significant at an alpha significance 0.05.

Table 6. ANCOVA Results on the Adjusted Mean Number of Arrests in 2021 by Perceived Race

	Coefficient	Standard Error	p-value	Confidence Interval
Intercept	1.231	0.014	<0.001*	[1.205, 1.258]
East/Southeast Asian	-0.08	0.027	0.003*	[-0.133, -0.027]
Indigenous	0.261	0.049	<0.001*	[0.164, 0.358]
Latino	-0.076	0.040	0.06	[-0.154, 0.003]
Middle Eastern	-0.065	0.032	0.042*	[-0.127, -0.002]
South Asian	-0.156	0.031	<0.001*	[-0.217, -0.096]
Unknown or Legacy	-0.172	0.027	<0.001*	[-0.226, -0.118]
White	0.082	0.017	<0.001*	[0.047, 0.116]
Covariate [±]	0.453	0.006	<0.001*	[0.442, 0.464]

Note: Results are interpreted in reference to the adjusted mean number of arrests in 2021 among Black persons.

() Significant at alpha 0.05.*

(+) Adjusted Mean Number of Arrests in 2021 at baseline with Black as the reference category.

(±) Number of arrests in 2020.

Nonetheless, we can see evidence of higher disproportionate policing in 2021 resulting in an arrest by Toronto law enforcement among individuals perceived as Black or Indigenous even after controlling for previous arrest history and criminal activity reported in 2020. Results of all other racial minorities including East/Southeast Asian (coefficient=-0.08, p-value=0.003, CI=[-0.133, -0.027]), Middle Eastern (coefficient=-0.065, p-value=0.042, CI=[-0.127, -0.002]) and South Asian (coefficient=-0.156, p-value=<0.001, CI=[-0.217, -0.096]) were all found to have been arrested statistically significantly less on average compared to Blacks, with individuals perceived as South Asian having the greatest difference (Table 6).

5.2. Logistic Regression: Chosen factors on Likelihood of Arrest

Results

Table 7. Result of the logistic regression model predicting the likelihood of being arrested

Model	Log Odds					Odds Ratio	
	Coef	S.E.	z	p	2.5%	97.5%	exp(Coef)
Constant	-2.9887	0.046	-65.131	0.000	-3.079	-2.899	0.050
Concealed Weapon	0.8049	0.197	4.087	0.000	0.419	1.191	2.237
Combative	1.6261	0.070	23.199	0.000	1.489	1.764	5.084
Resisted	1.4794	0.069	21.585	0.000	1.345	1.614	4.390
Mental Instability	1.2958	0.082	15.723	0.000	1.134	1.457	3.654
Assault	0.3456	0.171	2.022	0.043	0.011	0.681	1.413
Cooperative	1.8851	0.029	65.289	0.000	1.829	1.942	6.587
Sex (M)	0.2338	0.040	5.794	0.000	0.155	0.313	1.263
Youth (<18 years old)	-0.3066	0.070	-4.366	0.000	-0.444	-0.169	0.736

We conducted a logistic regression analysis to investigate the impact of demographic features and behaviors on the probability of being arrested. Instead of log odds, we utilized odds ratios to compare the correlations between different independent variables and arrest outcomes, as odds ratios are easier to interpret. There are two demographic features, sex and youth; and 6 types of demeanors during arrests, and from the summary table, all the results were statistically significant with $P < 0.05$, as shown in the table 7.

The results indicated that individuals who concealed items had a higher likelihood of being arrested compared to those who did not conceal items. The odds ratio for people who concealed items was 123% higher than its counterpart of people who didn't conceal items, controlling for other actions and demographic features such as age and gender. The coefficient for concealing items was statistically significant with a p-value close to 0.00, supporting our literature review that disrespectful behavior correlates with a higher likelihood of being arrested.

Similarly, individuals who engaged in combative, violent behavior, or spitting/biting had a 408% higher odds ratio of being arrested, holding other variables constant. The coefficient for this action was statistically significant with a p-value close to 0.00, aligning with our previous finding that hostile behavior is associated with a higher possibility of being arrested.

Regarding resisting, defending, or attempting to escape, the predicted odds ratio for being arrested was 339% higher for individuals who displayed such behavior compared to those who did not, with everything else the same. The coefficient for this action was considered statistically

significant, as its p-value was less than 0.05, consistent with our previous research that people who behave hostilely are more likely to be arrested.

We also examined the impact of mental instabilities on the likelihood of being arrested. Our previous findings indicated that individuals with mental instabilities were less likely to be arrested due to the natural disadvantage of this group. However, our data indicated that individuals with mental instabilities had a 265% higher odds ratio of being arrested compared to those without mental instabilities. This conclusion was statistically significant, as its p-value was less than 0.05.

Then we come to check the case for assaulting officers. The coefficient suggests that the group that engages in this behavior has a slightly higher likelihood of being arrested compared to the group that does not, with its odds ratio increasing by 41%. Compared to the previous four actions, behaving this action alone increases the least of the odds ratio. With the p-value of 0.043, this conclusion is still considered as statistically significant.

Unlike the previous five negative actions, the last behavior indicated a positive attitude of individuals towards cooperation during arrest. The odds ratio for individuals who cooperated was 6.587 times that of those who did not, meaning that individuals who cooperated had a higher likelihood of being arrested, when the other conditions are the same. This finding was statistically significant, with a p-value smaller than 0.05.

We have also examined the influence of gender and age on the likelihood of being arrested. Our data shows that, when controlling for other variables, males have a higher odds ratio of being arrested compared to females. Specifically, the odds ratio for males being arrested is 1.26 times of that for females. This finding is consistent with our previous research, which suggested that females were less likely to be arrested. The p-value for this result is less than 0.05, indicating that it is statistically significant.

Regarding age, our analysis reveals that being a youth is associated with a lower odds ratio of being arrested compared to being an adult. The coefficient for youth indicates that their odds ratio of being arrested is 73.6% of that for adults, controlling for other variables. In other words, youths are less likely to be arrested than adults.

Prediction Interval Plot

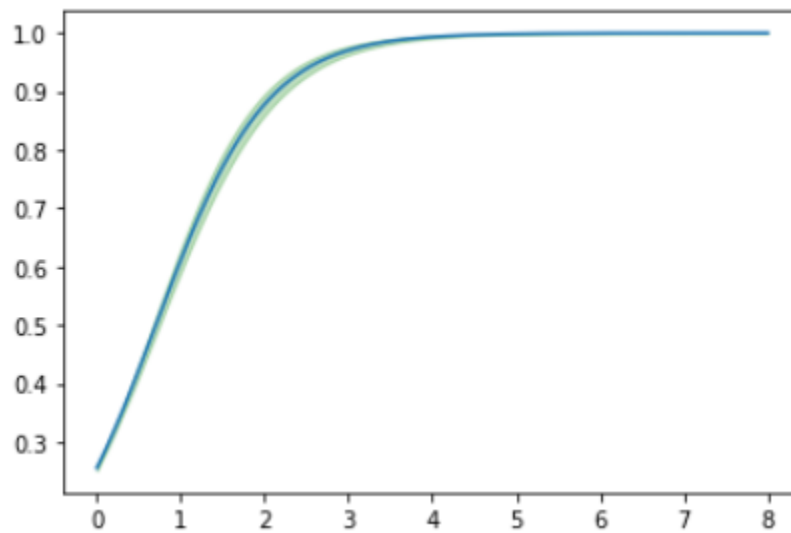


Figure 18. Prediction Interval

Figure18 displays the prediction interval of our model. The blue line represents the predicted probabilities, while the green shaded region represents the upper and lower limits of the prediction interval.

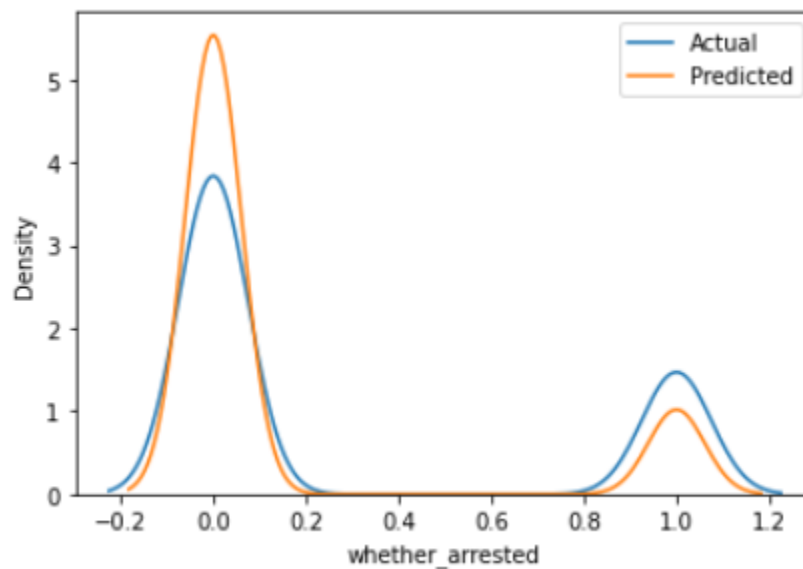


Figure 19. Prediction Overview

Referring to Figure19, we can see that our model tends to predict not being arrested more often, with the prediction numbers higher than the actual not being arrested. As for the prediction of being arrested, our model may under-predict the actual records of being arrested.

Confusion Matrix Plot

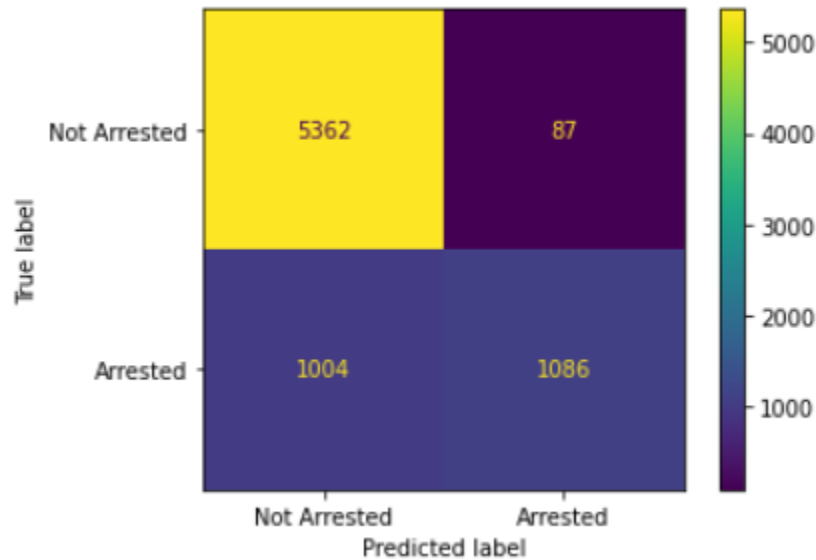


Figure 20. Confusion Matrix Plot of the Logistic Regression Model

From the confusion matrix plot(Figure 20), we can see that the overall accuracy rate of our model is at a decent level, but we do have a high proportion of false negatives, in which we wrongly take the actual arrests records as not being arrested.

6. Discussion and Limitations

Consistent with existing findings with respect to previous reports on the race-based data collected by Toronto Police and the interim report by the Ontario Human Rights Commissions (OHRC), results of the one-way ANCOVA further corroborated findings in which there appears greater burden of law enforcement among the Black and Indigenous communities. More specifically with respect to all other minorities, Black showed evidence of higher-than-average number of police arrests compared to all other racial subgroups, excluding Latinos which was not found to be significant, and with the exception of individuals who were perceived to be Indigenous. Such observation is observed even while controlling for previous arrest history and criminal activity reported in 2020. However, the results of the model are irrespective of other demographic characteristics, reason for arrest, and encounter-specific variables which were further addressed in our logistic model in further analyzing the product of such confounding variables on the implicit bias of law enforcement officers in influencing the discretionary nature of police arrests and therefore, the likelihood of being arrested based on the aforementioned characteristics. Nonetheless, results of the ANCOVA model suggest evidence of racial disparity in the average number of arrests occurring in 2021 with respect to individuals who were perceived as Black and Indigenous when it comes to people of color despite accounting for previous arrest history and criminal activity. As such, the model provides an initial indicator of the potential pervasive racial stereotype and prejudice that may be embedded in Toronto law enforcement practices and should be addressed.

Our logistic model has provided further evidence that certain types of disrespectful or hostile actions increase the likelihood of being arrested. All of the five negative actions, including concealing items, combative or violent behavior, spitting or biting, resisting or attempting to escape, suffering mental instabilities, and assaulting officers, have all shown statistically significant results, indicating that they lead to a higher chance of arrest based on our dataset. It is important to note that our findings regarding people with mental instabilities contradict our research findings in the literature review. Our model indicates that those with mental instabilities or suicidal tendencies are more likely to be arrested, which may require further investigation to determine the reasons behind this discrepancy.

Moreover, our findings on the impact of performing cooperative actions are inconsistent with our expectations. While we assumed that officers would be less likely to arrest individuals who perform cooperative actions, our model showed no significant difference between the likelihood of arrest for cooperative versus non-cooperative behavior. This could be due to bias resulting from the significantly smaller sample size of non-cooperative individuals in our dataset, or it is possible that cooperative actions have little impact on police arrest decisions.

Overall, our study has revealed that certain types of negative behavior increase the likelihood of being arrested, while the impact of positive behavior on police arrest decisions remains unclear and requires further investigation.

Limitations

Despite significant findings in our investigation, we recognize the limitations of the current study. We approached this study under the assumption of the Differential Selection Hypothesis previously mentioned in our literature review which attributes the race differential in law enforcement practices to the pervasive prejudice and discrimination that influence the discretionary nature of police arrests. To better isolate and discern the implicit bias of law enforcement officers with respect to perceived race, we controlled for previous arrest history and criminal activity as those who have been previously arrested are more likely to be repeat offenders in our ANCOVA model. However, as the collection of Race and Identity-Based data by Toronto Police Services (TPS) was only implemented in 2020, the extent to which we can confidently say that implicit racial bias persists even after controlling for previous arrests history and criminal activity is limited by the data provided by TPS. As such, we were only able to account for each distinct offender's criminal conduct recorded in 2020. To fully discern the presence of racial bias in the discretionary nature of Toronto police arrests, a more extensive and detailed account of an offender's criminal activity extending beyond their previous encounter with law enforcement in 2020 would be ideal to increase the confidence of such conclusions when controlling for previous arrest history. Furthermore, with respect to the Differential Involvement Hypothesis, we recognize that there may be other confounding variables that may also contribute to the disproportionate policing of certain minorities making it difficult to fully attribute the race differential in Toronto police arrest to be a result of discrimination and prejudice against certain racial minorities. More specifically, our study fails to account for the potential differences in crime rate with respect to each racial subgroup. Similarly, we also fail to account for the differential crime rate with respect to each jurisdiction of various TPS divisions, such that there may have been higher crime rates in certain jurisdictions leading to increased policing activity resulting in higher arrests for all racial groups. Nonetheless, findings of this

current study and the ANOVA model present valuable information in providing a basis for further investigations in addressing the systemic bias in law enforcement practices. Further studies would require additional considerations in synthesizing more information about previous arrest history and criminal activity with respect to each racial subgroup, as well as control of the differential crime rate that may contribute to the observed race differential with respect to each racial subgroup and police division.

Additionally, we also like to address the limitation for our logistic model. Specifically, in preprocessing the data for our logistic regression analysis, we used the sum of arrests and the median as a boundary to determine if an individual was arrested or not. This approach assumes that individuals with fewer arrests are not arrested, while those with more arrests are arrested. However, this is not an ideal method for analyzing the dataset with respect to logistic regression. Instead, we need a dataset where individuals with suspicious behavior may be booked by the police but ultimately not arrested. Furthermore, in our logistic regression, we summed the indicator of whether individuals performed certain types of actions, regardless of the relationship between each arrest and behavior record. For example, an individual may have been arrested three times but only concealed items once. In this case, we treated it as one arrest record with one concealing item record, rather than three arrest records with one having the record of concealing items. Our approach allowed us to understand the relationship between demeanor and arrests in a general sense, but it missed many details and did not provide precise and definitive results. To address this issue, we need to obtain another dataset where individuals behaved similarly but were not ultimately arrested. Then, we can study each record with detailed and precise information, rather than summing up the data. This would significantly improve the accuracy of our results. Lastly, while our model has a high accuracy rate, we also have a high false negative rate. Our prediction can only identify half of the true arrest records. Therefore, there is still much room for improvement in our study.

7. Conclusions

We focus our attention on arrests as it serves as a point of entry into the criminal justice system and therefore serves as a critical point in addressing any existing discrimination in the system that can otherwise lead to wrongful convictions with profound life-time consequences. Results of our ANCOVA model suggest evidence of implicit racial bias among Toronto law enforcement officers that persists even when controlling for previous arrest history and criminal activity. More specifically, we saw a higher disproportionate average number of arrests among the Black and Indigenous community when compared to all other minority groups, consistent with findings previously reported by the Ontario Human Rights Commission and discussed in the literature review.

Furthermore, based on our logistic regression analysis, we found evidence that certain types of disrespectful or hostile behaviors, including concealing items, resisting arrest, exhibiting combative behavior, or demonstrating suicidal tendencies, are positively correlated with the likelihood of being arrested. This finding is consistent with our literature review. However, we were unable to determine the extent to which assaulting officers contributed to the likelihood of being arrested, as our dataset did not provide sufficient information in this regard. While our analysis suggested that positive behaviors such as cooperation might also increase the likelihood

of being arrested, we question this conclusion due to the disparity in sample sizes. Furthermore, our analysis indicated that, after controlling for other variables, males are less likely to be arrested than females, and juveniles have a lower likelihood of being arrested than adults.

In future research, we plan to investigate whether there is an interaction between these behaviors and ethnicities, which may further influence officers' decisions to make arrests. Additionally, future research will also need to take into consideration the differential crime rate that may contribute to the observed race differential with respect to each racial subgroup and police division on Toronto Police arrests.

References

- D'alessio, S.J., & Stolzenberg, L. (2003). Race and the Probability of Arrest. *Social Forces*, 81, 1381 - 1397.
- Etheline. (2022, January 19). Race-based data collection, analysis and public reporting. Toronto Police Services Board. Retrieved February 23, 2023, from <https://www.tpsb.ca/policies-by-laws/board-policies/177-race-based-data-collection-analysis-and-public-reporting>
- Fanfair, R. (2022, June 15). *Race-Based Data Shows Over-Policing*. Race-based data shows over-policing. Retrieved February 23, 2023, from <https://www.tps.ca/media-centre/stories/race-based-data-shows-over-policing/>
- Headley, A. M., & Wright, J. E. (2020). Is Representation Enough? Racial Disparities in Levels of Force and Arrests by Police. *Public Administration Review*, 80(6), 1051–1062. <https://doi.org/10.1111/puar.13225>
- Johnson, R. R. (2016). Suspect demeanor and arrest: A triggered displacement of aggression explanation. *American Journal of Criminal Justice*, 42(1), 170–187. <https://doi.org/10.1007/s12103-016-9352-8>
- Lundman, R. J. (1996). Demeanor and Arrest: Additional Evidence from Previously Unpublished Data. *Journal of Research in Crime and Delinquency*, 33(3), 306–323. <https://doi.org/10.1177/0022427896033003002>
- New York University School of Law. (n.d.). Race and Discretion in the Criminal Legal System | NYU School of Law. Retrieved February 24, 2023, from <https://www.law.nyu.edu/centers/race-inequality-law/our-work/race-and-discretion-in-criminal-legal-system>
- Novak, K. J., & Engel, R. S. (2005). Disentangling the influence of suspects' demeanor and mental disorder on arrest. *Policing: An International Journal of Police Strategies & Management*, 28(3), 493–512. <https://doi.org/10.1108/13639510510614573>
- Ontario Human Rights Commission. (2020, August 10). New OHRC report confirms black people disproportionately arrested, charged, subjected to use of force by Toronto police. New OHRC report confirms Black people disproportionately arrested, charged, subjected to use of force by Toronto police. Ontario Human Rights Commission. Retrieved February 23, 2023, from https://www.ohrc.on.ca/en/news_centre/new-ohrc-report-confirms-black-people-disproportionately-arrested-charged-subjected-use-force
- Piquero, A. R., & Brame, R. W. (2008). Assessing the Race-Crime and Ethnicity-Crime Relationship in a Sample of Serious Adolescent Delinquents. *Crime and delinquency*, 54(3), 390–422. <https://doi.org/10.1177/0011128707307219>
- Piquero, A. R. (2008). Disproportionate minority contact. *The Future of Children*, 18(2), 59–79. <https://doi.org/10.1353/foc.0.0013>
- Policing Women: Race and gender disparities in police stops, searches, and use of force (2019, May 14). Prison Policy Initiative. <https://www.prisonpolicy.org/blog/2019/05/14/policingwomen/>
- Race and Identity-based Data Collection Strategy. (n.d.). Toronto Police Service Public Safety Data Portal. Retrieved February 24, 2023, from <https://data.torontopolice.on.ca/pages/race-based-data>
- Race and Identity-Based Data Collection. (n.d.). Toronto Police Service. Retrieved February 24, 2023, from <https://www.tps.ca/race-identity-based-data-collection/>

- Schirrmeister, E., Göhring, A., & Warnke, P. (2020). Psychological biases and heuristics in the context of foresight and scenario processes. *Futures & Foresight Science*, 2(2).
<https://doi.org/10.1002/ffo2.31>
- Schleiden, C., Soloski, K. L., Milstead, K., & Rhynehart, A. (2020). Racial Disparities in Arrests: A Race Specific Model Explaining Arrest Rates Across Black and White Young Adults. *Child and Adolescent Social Work Journal*, 37(1), 1–14.
<https://doi.org/10.1007/s10560-019-00618-7>
- Smith, D. A., & Visher, C. A. (1981). Street-Level Justice: Situational Determinants of Police Arrest Decisions. *Social Problems*, 29(2), 167–177. <https://doi.org/10.2307/800422>
- Smith, D. A., Visher, C. A., & Davidson, L. A. (1984). Equity and Discretionary Justice: The Influence of Race on Police Arrest Decisions. *Journal of Criminal Law & Criminology*, 75(1), 234–249. <https://doi-org.myaccess.library.utoronto.ca/10.2307/1143211>
- Spencer, K. B., Charbonneau, A. K., & Glaser, J. (2016). Implicit Bias and Policing. *Social and Personality Psychology Compass*, 10(1), 50–63. <https://doi.org/10.1111/spc3.12210>
- Statistics Canada. (2009). Female Offenders in Canada (No. 85-002-XIE, Vol. 28, no. 1). Ottawa, ON.
- Staniloiu, A., & Markowitsch, H. (2012). Gender differences in violence and aggression – a neurobiological perspective. *Procedia - Social and Behavioral Sciences*, 33, 1032–1036.
<https://doi.org/10.1016/j.sbspro.2012.01.279>
- Wildeman, C., & Turney, K. (2014). Positive, negative, or null? The effects of maternal incarceration on children's behavioral problems. *Demography*, 51(3), 1041–1068.
<https://doi.org/10.1007/s13524-014-0291-z>
- Wortley, S., & Jung, M. (2020, July). Racial Disparity in Arrests and Charges: An analysis of arrests and charge data from the Toronto Police Service. Ontario Human Rights Commission. Retrieved February 23, 2023, from
<https://www.ohrc.on.ca/sites/default/files/Racial%20Disparity%20in%20Arrests%20and%20Charges%20TPS.pdf#overlay-context=en/disparate-impact-second-interim-report-inquiry-racial-profiling-and-racial-discrimination-black>