Racial Disparity and the Intersectionality of Race, Sex and Being youth or Not on Arrest Outcome by Toronto Police

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

Ethnicity discrimination in police arrests has been a controversial topic, especially since 2020. However, regarding police decisions on arrests, ethnicity is not the only demographic feature that shows a significant difference among different groups. In order to solve the concerns of potential bias and discrimination, we need to understand the phenomenon and current context first. With a solid understanding of the overall situation, more policies can be made to provide transparency and gradually reduce discrimination. This report aims to study the amount of variation in average police arrests concerning combinations of several common demographic features, including ethnicity, gender, and age division. The report uses the "Arrests and Strip Searches" dataset downloaded from the Toronto Police Service website in the public safety data portal to understand the police arrest situation in the Great Toronto Area. The dataset contains arrest records from April 2021 to December 2022. The report provides background information on previous police arrest differences concerning ethnicities, genders, and juveniles. Then the report performs Exploratory Data Analysis to study the primary trends of the dataset, as well as conducts t-tests to build the most intuitive perception of the variables of interest. The following are two research questions. The first research question studies solely the mean arrest differences among eight races using One-way ANOVA, and the second research question study the interaction effect between race, sex, and whether youth or not in affecting the mean number of arrests using Two-way ANOVA. There are three sub-questions in the Two-way ANOVA, corresponding to three combinations of three categories interacting in a group of two. Moreover, for the variable where significant differences exist, the Tucky test is applied to investigate further where the differences are between groups. The results on the two research questions indicate that from the dataset we used, there do exist mean differences of arrests between 8 races, and the combination of being youth or not and races. Furthermore, the test results indicate that the Indigenous group tends to have the highest arrest means than the other groups, followed by the White and Black population. Whereas the adult group has a higher average number of arrests than the youth group, the race differences keep the same order in decreasing order, from the Indigenous group to the White group and the last Black group. Finally, the report discusses its limitation due to the lack of context. Simply studying the number of arrests despite the criminal categories and severities can introduce biases from other perspectives.

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; Schirrmeister 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 disproportion 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 police serves as the point of entry into the criminal justice system, we will look at the implication of the embedded race differential in institutional practices on the number of arrests made by law enforcement in 2022. 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. Many studies have focused on the racial differential in the context of Black males, and we have previously mentioned how police discretion is strongly influenced by race, sex, and age. As a result, we will further introduce a additional independent variables with respect to sex and age in understanding how the combination two social constructs: race and sex, may intersect to contribute to outcome of Toronto police arrests. The following section will further introduce the role of sex and age, as isolated variables on arrest outcomes.

2.2. Sex and Police Arrests

Sex differences exist in many aspects of life, so do the criminology. In 2017, mem make up for 73% of total arrests, almost three times than women with a proportion of 27%, according to data from the FBI Crime in the United States. And such gap has already shrunk compared to a larger difference, with a ratio between men and women: 84% vs 16%, back in 1980. (*Policing Women: Race and gender disparities,* 2019) This literature review will provide an overview of existing findings on gender differences in police arrests.

What keeps consistent with the higher proportion of total arrests for males is that statistics have revealed that men commit more criminal acts than women. This difference could possibly result from biological difference between men and women. According to Nolen-Hoeksema (2000), women are more likely than men to display a self-focused ruminative style while men tend to employ more distracting strategies. Therefore, when it comes to performing

aggression, women might prefer to express sin inward directed forms, whereas men are more likely to choose outward directed ways and leading more violence criminals as a result. (Staniloiu & Markowitsch, 2012)

However, in addition to the different likelihood of crime due to neurophysiological differences between men and women, some studies claim that institutional issues are also a potential influencing factor. Statistics said that due to the long-term low criminal rate of women and girls, they will face a Canadian criminal justice system mainly designed for male criminal population when conducting a crime, which might cause a higher tendency of being tolerated, where males would be arrested under the same context. (Statistics Canada, 2009)

In conclusion, sex differences in police arrests are a complex issue that led by many factors within a long period of the history, including their inherent differences in the biological perspective, and biased treatment in the criminal justice system. This report will further validate this phenomenon with the data extracted from Toronto Police Service (TPS) and explore its interaction with other demographic features including races and age groups (being youth or not) in affecting an individual's average number of arrests.

2.3 Juveniles in Police Arrests

Juvenile delinquency has always been a common problem in all countries. So far, much research has been done to study the deeper reasons for juvenile delinquency to prevent criminals and contribute more positively to the youth group. Though law enforcement varies across countries, policies on police arrests typically differentiate between youth and non-youth groups, allowing for more tolerance of youth criminal acts. This literature review will provide an overview of core findings on comparing police arrests between the youth and non-youth populations.

Before going deeply into the problem scope, we must clarify that youth indicates people aged 12 to 17. Children below 12 do not hold lawful responsibilities for their behaviour in Canada. From Statistics Canada, the youth-related crime rate was lower than for young adults between 18 to 24, whereas it was higher than the adults at or above 25. Referring to Statistics Canada 2022 police-reported data, the youth at 17 have the most significant rate of being accused of a crime, not only within the youth group but also among the overall age groups. Though the youth population criminal rate is inconsistent when compared to adults of different age ranges, research has pointed out that, on average, youth face a lower likelihood of being arrested than non-youth groups. (Piquero,2008)

Referring to Statistics Canada (2022), among youth crimes, one critical feature is that this group is more likely to be involved in crimes involving more than one accused, also known as co-offending, compared to their adult counterparts. Regarding the major categories of youth crimes, theft of $5,000 and under, assault, and drug use (mainly Cannabis possession) are the top three types, and drug use is mainly conducted by aged youth than the younger group. Research by Wildeman and Turney (2014) pointed out that children growing up in families suffering poverty, low education experience, and parents using drugs tend to show higher volatility in performing criminal acts.

To sum up, juvenile delinquency has shown a continuous decreasing trend in Canada from the 1990s till now. Understanding the phenomenon can build us a solid overview of juvenile delinquency and enable strengthening policies to prevent youth from committing crimes and build a safer and more friendly environment. This report will use the Toronto Police Services data to check if it is consistent with the literature review findings. It will also assess the

interaction between this factor and general demographic features like gender and race in affecting one's arrest amounts.

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 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 majority of the arrest completed over the span of 2020 to 2022 by Toronto Police Services (TPS) have been adult offenders, with only 4.7% of the arrest 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%). Figure 1A further depicts the average mean count of arrests with respect to sex and perceived race. We can see that, on average and among all minority groups, there appears to be higher number of arrests among the Indigenous population. This appears to be true across both males and females. Figure 1B further depicts the distribution of the counts of arrests with respect to sex and perceived race, and based on preliminary analysis of EDA, it appears that the distribution of counts of arrests is not normal and is right-skewed. Further analysis also found that majority of arrests have been a result of an assault and other crimes against persons (Figure 2). This is followed by incidences of robbery and theft and arrest warrants for detainment.

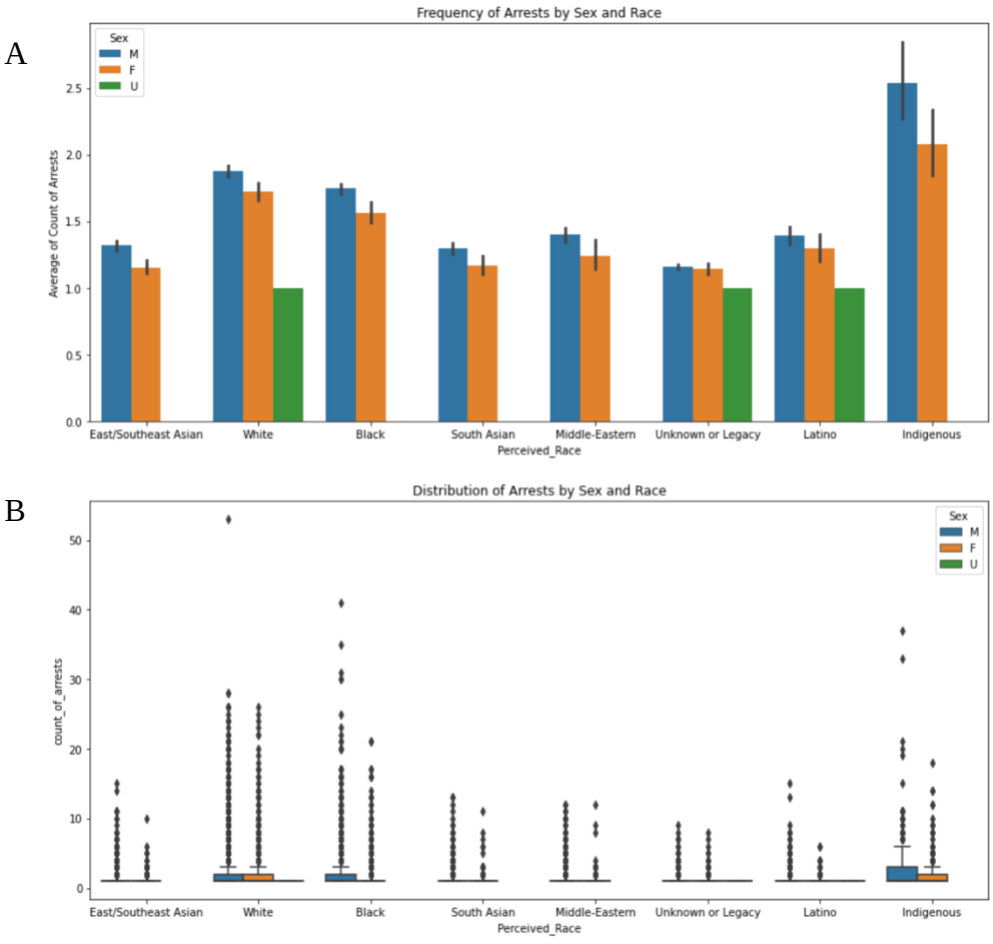

Figure 1. Frequency of Arrests by Sex and Perceived Race (A) and Distribution of Arrests by Sex and Perceived Race (B)

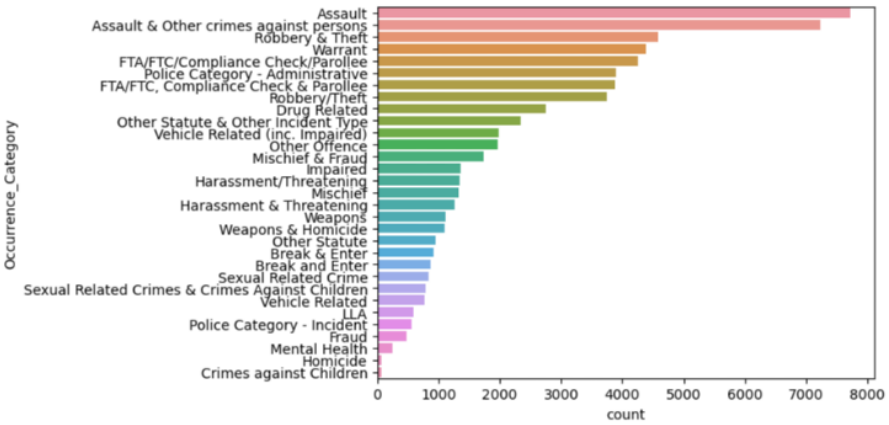

Figure 2. Frequency of Arrests by Occurrence Category

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, it appears that majority of arrest outcome have been cooperative across all perceived races. Results of the analysis also found that among all minority groups, it appears that in at least one-fourth of the arrests among the Black community, individuals tend to engage more non-cooperative behavior including resistance to arrest (31.8%), combative (29%), and assault of an officer (34.5%) (Table 2). Furthermore, Black person also represented 36.8% of the total arrests that involved concealment of a weapon.

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

	Concealed Weapon		**Combative**		**Resistance**	
	Count	**%**	**Count**	**%**	**Count**	**%**
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**	

	Count	%	Count	%	Count	%
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

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 incidences involving mental instability (25.3%) among females. Nonetheless, 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 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 arrests. 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

There are two major dependent variables in the following five t-tests, one is the total arrests of a person by "PersonID," and the other is the total number of strip searches of a person by "StripSearch." Both are ratio variables, and there are no missing values. Therefore, both variables satisfy the assumption requirement for dependent variable. All the samples are randomly selected among the overall population. The following report won't repeat checking these assumptions.

3.2.1. T-test on gender differences (male vs. female) w.r.t arrests

This test is to check if the average arrest amount differs by genders. To be more specific, the two involving groups are males and females.

a. Assumption check

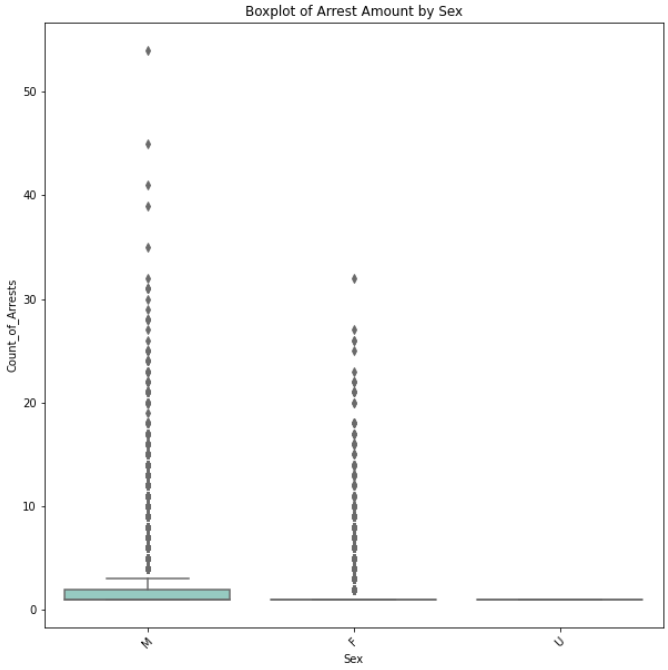

Figure 3. Boxplot of arrest number by sex

From the boxplot, the distributions of two groups are non-normal, and they are both right-skewed. However, as each group has a large sample size, we can approximate the sample means to be normal according to the central limit theorem. In addition, the t statistic will converge in probability to the normal distribution by the law of large numbers. Therefore, the normality assumption is satisfied. For homogeneity of variances, we apply Levene's test, and the test result (statistic= 28.20, pvalue=1.10*10^ (-7)) shows that the p value is less than 0.05, indicates that it violates the equal variance assumption. Therefore, we will apply Welch's test.

b. Hypothesis

Null hypothesis: H0: On average, males and females do not differ in the total number of arrests.
Alternative hypothesis: Ha: On average, males and females differ in the total number of arrests.

c. Test result & Conclusion

t-statistic	5.3848082092215455
p-value	7.381040021829354e-08

From the results, we could see that the t statistic is 5.38, and the p-value is 7.38^10(-8), which is smaller than 0.05, the default significance level. Therefore, we reject the null hypothesis and conclude that, on average, there is a significant difference in the total number of arrests between males and females.

3.2.2.T-test on gender differences (male vs. female) w.r.t strip searches

This test is to check if the average strip search amount differs by genders. To be more specific, the two involving groups are males and females.

a. Assumption check

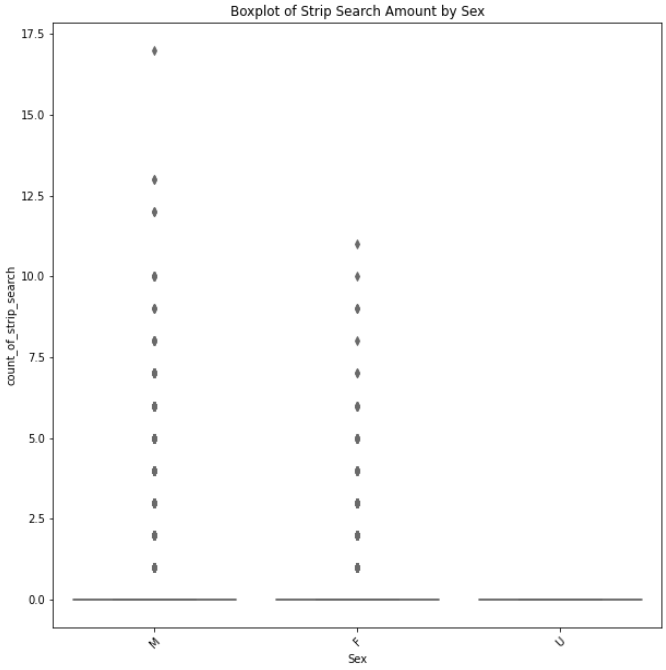

Figure 4. Boxplot of strip search by sex

From the boxplot, distributions of the two groups are right-skewed. According to the central limit theorem and law of large numbers, the non-normality has little effect on the test results. Levene's test checks variance homogeneity, and the test result (statistic=110.91, pvalue=6.74*10^ (-26)) shows that the p value is less than 0.05, indicating a violation to the equal variance assumption. Therefore, we will apply Welch's test.

b. Hypothesis

Null hypothesis: H0: On average, males and females do not differ in the total number of strip searches.

Alternative hypothesis: Ha: On average, males and females differ in the total number of strip searches.

c. Test result & Conclusion

t-statistic	6.495500992883376
p-value	8.566543261428306e-11

From the results, we could see that the t statistic is 6.50, and the p-value is 8.57^10(-11), which is smaller than 0.05, the default significance level. Therefore, we reject the null hypothesis and conclude that, on average, there is a significant difference in the total number of strip searches between males and females.

3.2.3. T-test on ethnicity differences (black vs. white) w.r.t arrests

This test is to check if the average arrest amount differs by ethnicity. To be more specific, the two involving groups are black and white.

a. Assumption check

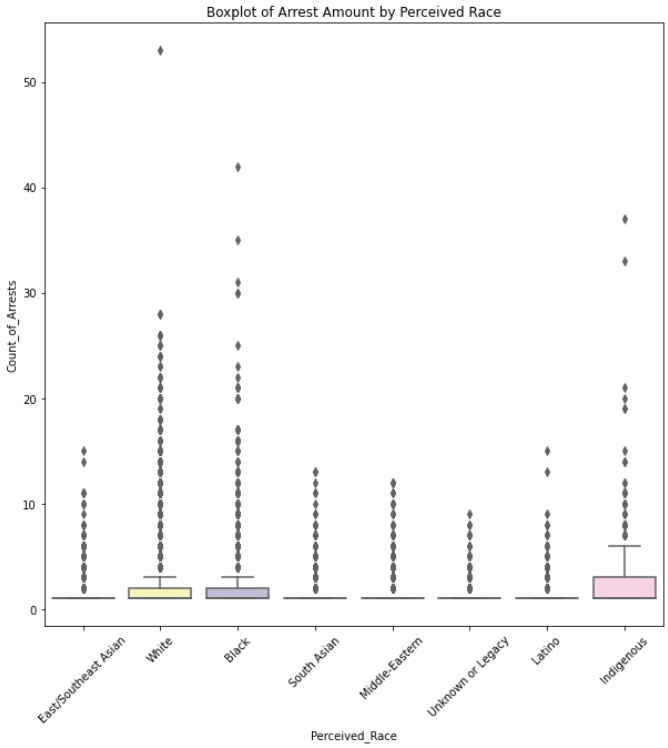

Figure 5. Boxplot of arrest number by race

From the boxplot, distributions of the two groups are right-skewed. According to the central limit theorem and law of large numbers, the non-normality has little effect on the test results. Levene's test checks variance homogeneity, and the test result (statistic=67.97, pvalue=1.74*10^ (--16)) shows that the p value is less than 0.05, indicating a violation to the equal variance assumption. Therefore, we will apply Welch's test.

b. Hypothesis

Null hypothesis: H0: On average, black and white do does not differ in the total number of arrests.

Alternative hypothesis: Ha: On average, black and white differ in the total number of arrests.

c. Test result & Conclusion

t-statistic	5.002415488701116
p-value	5.701721645535063e-07

From the results, we could see that the t statistic is 5.00, and the p-value is 5.70^10(-7), which is smaller than 0.05, the default significance level. Therefore, we reject the null hypothesis and conclude that, on average, there is a significant difference in the total number of arrests between black and white people.

3.2.4. T-test on ethnicity differences (indigenous vs. white) w.r.t arrests

This test is to check if the average arrest amount differs by ethnicity. To be more specific, the two involving groups are indigenous and white.

a. Assumption check

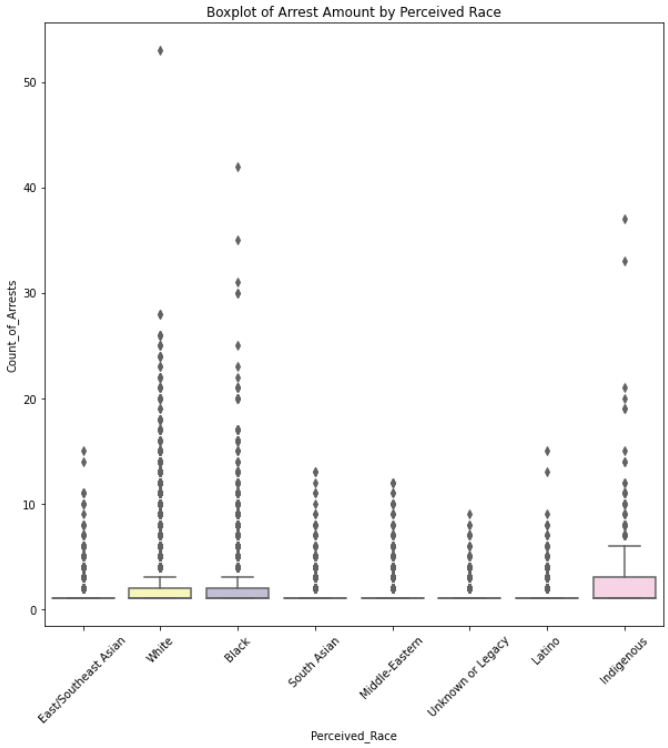

Figure 6. Boxplot of arrest number by race

From the boxplot, distributions of the indigenous and white groups are right-skewed. According to the central limit theorem and law of large numbers, the non-normality has little effect on the test results. Levene's test checks variance homogeneity, and the test result (statistic=59.83, pvalue=1.10*10^ (-14)) shows that the p value is less than 0.05, indicating a violation to the equal variance assumption. Therefore, we will apply Welch's test.

b. Hypothesis

Null hypothesis: H0: On average, indigenous and white do does not differ in the total number of arrests.

Alternative hypothesis: Ha: On average, indigenous and white differ in the total number of arrests.

c. Test result & Conclusion

t-statistic	5.094407347210731
p-value	4.2978709354381165e-07

From the results, the t statistic is 5.09, and the p-value is 8.57^10(-11), which is smaller than 0.05, the default significance level. Therefore, we reject the null hypothesis and conclude that, on average, there is a significant difference in the total number of arrests between indigenous and white people.

3.2.5. T-test on age differences (youth vs. non-youth) w.r.t arrests

This test is to check if the average arrest amount differs by youth-or-not.

a. Assumption check

In conducting this test, the report has recoded the dataset for the youth group division. From the file, the variable “Youth_at_arrest__under_18_years” has three subgroups, whereas two groups are presenting the same group, “Youth (aged 17 and younger)” and “Youth (aged 17 years and under)”, so we consolidated them into one group of being youth despite the annotations.

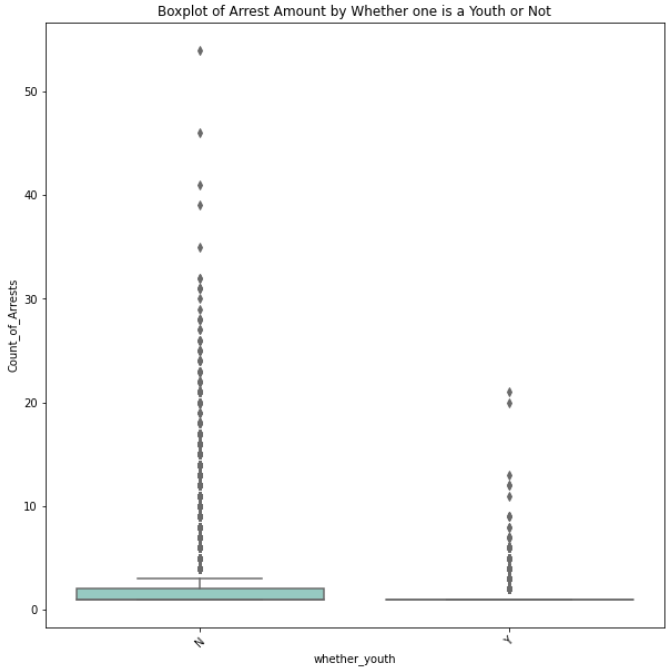

Figure 7. Boxplot of arrest number by youth status

From the boxplot, distributions of the two groups are right-skewed. According to the central limit theorem and law of large numbers, the non-normality has little effect on the test results. Levene's test checks variance homogeneity, and the test result (statistic=100.01, pvalue=1.62*10^(-23)) shows that the p value is less than 0.05, indicating a violation to the equal variance assumption. Therefore, we will apply Welch's test.

b. Hypothesis

Null hypothesis: H0: On average, youth and non-youth do not differ in the total number of arrests.

Alternative hypothesis: Ha: On average, youth and non-youth differ in the total number of arrests.

c. Test result & Conclusion

t-statistic	**10.233359483537276**
p-value	3.688918451540259e-24

From the results, the t statistic is 10.23, and the p-value is 3.69^10(-24), which is smaller than 0.05, the default significance level. Therefore, we reject the null hypothesis and conclude that, on average, there is a significant difference in the total number of arrests between youth and non-youth.

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 with 37,347 unique individuals who have been arrested at least once. 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.

This report focuses on two socio-demographic features, the races, and genders of the person being arrested. There are eight categories under the categorical variable "perceived_race," with seven major races and a group indicating "unknown or legacy." For genders, there are three categories, including male, female, and a combined gender-diverse group, which is indicated by "U" in the file. The dependent variable is a continuous variable of the total number of a person being arrested, which is calculated by summing up the times of a PersonID appearing in the file. From the exploratory data analysis, we could see that the number of a person being arrested is naturally right skewed, indicating that the majority of the group only has a few arrest records, leaving the small portion having more than average arrest records.

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. 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 in the criminal justice system (Spencer et al.,

2016), and the potential subsequent implications that can follow. We also discussed how the discretionary nature of police arrests can be heavily influenced by race, age, and sex (Smith et al., 1984). Therefore, 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.

As a result, the present study aims to investigate (1) the race differential in Toronto police arrests, and (2) how race, as a social construct, intersects with an individual's sex in influencing arrest outcomes, considering one's age maturity, measured by non-youth and youth. Therefore, we introduce the following research questions to be addressed:

1. How does an individual's perceived race influence the number of police arrests?

2. How does an individual's social identity with respect to perceived race, sex, and youth status, work together to influence the number of police outcomes?

4.3. Data Analysis Method

To investigate the racial disparity in Toronto Police arrests, a one-way Analysis of Variance (ANOVA) will be conducted in determining whether significant differences exist in the average number of police arrests between each racial subgroup of the perceived race among the arrested population recorded by Toronto Police. This allows for the comparison of a continuous dependent measure with respect to the number of police arrests from 2020 to 2022, across multiple classes of an independent categorical variable with respect to the perceived race of the detained individual.

Second, as previously addressed – since we recognize that the discretionary nature of police arrests is heavily influenced by race, sex, and age, we will be conducting three separate two-way ANOVAs to determine: (1) how an individual's perceived race with respect to their sex influence the mean number of police arrests, (2) how an individual's sex and whether being a youth influence the mean number of police arrests, and (3) how an individual's perceived race and whether being a youth influence the mean number of police arrests.

Due to the nature of ANOVA, a post-hoc test is to be conducted for each ANOVA to identify where exactly the differentials lie with respect to the mean number of police arrests, should significance be identified. The Tukey Honest Significant Difference (HSD) test will be performed to compare the means of all possible pairwise groups of a categorical independent variable to identify which groups are significantly different from each other after controlling for multiple comparisons. This enables control over the inflation of making a Type 1 Error in falsely rejecting a null hypothesis when it is true. Results of Tukey statistical test will also provide a quantified measure of the magnitude with respect to the mean difference in the number of police arrests for all pairwise comparison between each group.

3. Results & Findings

5.1. One-way ANOVA: Racial Disparity in Police Arrests

Hypothesis Model: To investigate whether significant differences exist in the mean number of arrests between perceived racial groups by Toronto Police Service (TPS) from 2020 to

2022, a Welch one-way ANOVA of unequal variance was conducted. The following hypothesis model is as followed:

- **Null Hypothesis (H_0)**: There is no significant difference in the mean number of police arrests by perceived race.

- **Alternative Hypothesis (H_A)**: There is a significant difference in the mean number of police arrests by perceived race.

Assumptions Check: Prior to running the ANOVA, the following assumptions and conditions were checked: (1) normality of continuous dependent variable (number of police arrests) with respect to each racial subgroup of perceived race, and (2) homogeneity of variances using the Kolmogorov-Smirnov test, and Levene's test respectively.

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. However, for the purpose of this report, we will continue to proceed with the one-way ANOVA analysis. More specifically, a Welch one-way ANOVA will be used. The Levene's test of homogeneity was found to be significant at an alpha of 0.05 (statistic: 145.0, p-value<0.001), which assumes equal variance as its null. Therefore, a Welch one-way ANOVA for unequal variance will be performed.

ANOVA Results: Results of the bivariate analysis found that there was significant evidence to suggest that there exists a difference in the mean number of arrests among the eight racial groups defined by TPS (Black, White, East/Southeast Asian, South Asian, Indigenous, Latino, Unknown/Legacy) at an alpha significance level of 0.05 (F=254.4, p-value=<0.001) (Table 5). Therefore, we can reject the null hypothesis in which no significant differences exist in the mean number of police arrests by perceived race. To determine which of the racial subgroups differ with respect to the mean number of arrests, a post-hoc Tukey HSD was performed (Table 6).

Subsequent Tukey HSD analysis found that across all races with respect to the Black population, there is a higher mean number of arrests among Black persons, with the exception of the Indigenous population which found that the latter to have on average, 0.649 more arrests than the Black racial subgroup (p-value=0.001) at a 95% confidence interval (CI = [0.467, 0.831]) and an significance level of 0.05. This is also true with respect to the pairwise comparison of the Blacks to Whites, which saw a mean difference of 0.126 (p-value=0.001, CI=[0.062, 0.190]). A similar observation can be observed with respect to the Indigenous population. In all pairwise instances of all other racial subgroups in the Tukey analysis, individuals who were perceived to be Indigenous were found to have a higher significant number of arrests than all other racial subsets, including Black person as previously mentioned, as well as the White subgroup (mean difference = -0.523, p-value=0.001, CI=[-0.703, -0.343]).

Most notably, there appears to be no significant differences where the East/Southeast Asian subgroup has a higher mean number of arrests. In fact, for all pairwise comparisons involving East/Southeast Asians that are significant, results of the Tukey found only significant differences with Black (mean difference = -0.391, p-value = 0.001, CI=[-0.491, -0.292]) which showed evidence Black racial subgroup having a higher mean number of arrests. Similarly, when

compared to the White subgroup, the latter which saw a 0.518 higher mean number of arrests than East/Southeast Asians (p-value = 0.001, [CI = 0.422, 0.613]).

Nonetheless, results of this analysis suggests that among all minority groups, members the Black and Indigenous community are found to be more frequently arrested on average than other racial minorities including East/Southeast Asians, South Asian, and Latino. More importantly, we further see evidence of only the Indigenous minority population being arrested more, with respect to the mean difference, than the White population.

Table 5. Welch One-way ANOVA for Unequal Variances

	Sum Sq	**Df**	**F value**	**Pr(>F)**
Perceived Race	4971.5	7.0	254.4	<0.001***
Residuals	114967.5	41174		

Table 6. Tukey HSD for all Significant Pairwise Comparison of Perceived Race (alpha=0.05)

group1	**group2**	**meandiff**	**p-adj**	**lower**	**upper**	**reject**
Black	East/Southeast Asian	-0.3913	0.001	-0.4906	-0.292	True
Black	Indigenous	0.649	0.001	0.4668	0.8312	True
Black	Latino	-0.3149	0.001	-0.4634	-0.1663	True
Black	Middle-Eastern	-0.3058	0.001	-0.4208	-0.1907	True
Black	South Asian	-0.403	0.001	-0.5098	-0.2961	True
Black	Unknown or Legacy	-0.517	0.001	-0.608	-0.4261	True
Black	White	0.1264	0.001	0.0623	0.1904	True
East/Southeast Asian	Indigenous	1.0403	0.001	0.8449	1.2357	True
East/Southeast Asian	White	0.5177	0.001	0.4224	0.613	True
Indigenous	Latino	-0.9638	0.001	-1.1883	-0.7394	True
Indigenous	Middle-Eastern	-0.9547	0.001	-1.1586	-0.7509	True

Indigenous	South Asian	-1.0519	0.001	-1.2513	-0.8526	True
Indigenous	Unknown or Legacy	-1.166	0.001	-1.3573	-0.9747	True
Indigenous	White	-0.5226	0.001	-0.7026	-0.3425	True
Latino	Unknown or Legacy	-0.2022	0.0031	-0.3617	-0.0426	True
Latino	White	0.4412	0.001	0.2953	0.5871	True
Middle-Eastern	Unknown or Legacy	-0.2112	0.001	-0.3402	-0.0823	True
Middle-Eastern	White	0.4322	0.001	0.3206	0.5437	True
South Asian	White	0.5293	0.001	0.4262	0.6325	True
Unknown or Legacy	White	0.6434	0.001	0.5568	0.73	True

5.2. Two-way ANOVA: Sex, Race and Youth-status Interaction Effect in Police Arrests

Table 7. Combined Two-way ANOVA for Perceived Race and Sex, and Youth Status

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
C(Sex)	2.0	56.237923	28.118961	9.604222	6.759716e-05 ***
C(Perceived_Race)	7.0	3087.13755	441.019651	150.633247	1.662327e-220 ***
C(whether_youth)	1.0	134.842054	134.842054	46.056216	1.164870e-11 ***
C(Sex):C(Perceived_Race)	14.0	33.344656	2.381761	0.813507	6.552287e-01
C(whether_youth):C(Perceived_Race)	7.0	110.198709	15.742673	5.377016	3.571977e-06 ***
C(Sex):C(whether_youth)	2.0	4.008840	2.004420	0.68462	5.042860e-01

a. How do an individual's perceived race and sex together influence the number of police arrests?

- **Null hypothesis:** All perceived race and sex combinations are equal in the number of arrests on average.
- **Alternative hypothesis:** There is at least one combination of perceived race and sex that is different from others in terms of the number of arrests on average.

Conclusion:

From the Two-way ANOVA test results, we can see that for the interaction effect between sex and perceived race, the p-value is 0.66, which is greater than 0.05. Therefore, we fail to reject the null hypothesis that all the combinations of sex and races are equal in the number of arrests on average. And we won't further apply the Tukey test.

b. How do an individual's sex and whether being a youth together influence the number of police arrests?

- **Null hypothesis:** All sex and status of youth combinations are equal in the number of arrests on average.
- **Alternative hypothesis:** There is at least one combination of sex and status of youth that is different from others in terms of the number of arrests on average.

Conclusion:

As for the interaction effect between sex and whether one is youth or not, its p-value is 0.50, which is greater than 0.05. We fail to reject the null hypothesis of the second research question, that all the combinations of sex and the status of youth are equal in the number of arrests on average. Still, there is no Tukey test for it.

c. How do an individual's perceived race and whether being a youth together influence the number of police arrests?

- **Null hypothesis:** All perceived race and status of youth combinations are equal in the number of arrests on average.
- **Alternative hypothesis:** There is at least one combination of perceived race and status of youth that is different from others in terms of the number of arrests on average.

Conclusion:

From the Two-way ANOVA test results, the p-value for the interaction effect between perceived race and whether being a youth is 3.57*10^ (-6), which is less than 0.05. Therefore, we reject the null hypothesis that all perceived race and status of youth combinations are equal in the number of arrests on average. Furthermore, we will apply the Tukey test on it to explore how do the mean number of arrests vary between these combinations.

Table 8. Tukey Test Result for the interaction effect between races and youth status

	reject1	reject2	Total sum
N / Indigenous	12.0	2.0	14.0
N / White	6.0	7.0	13.0
N / Black	10.0	0.0	10.0
N / Unknown or Legacy	2.0	5.0	7.0

N / East/Southeast Asian	4.0	1.0	5.0

From the Tukey test result, we can conclude that the mean values are significant different, if the "Total sum" does not fall in the reject region. And from the table, we can see that all the combinations are significantly different in the mean number of arrests, and all of them are non-youth. To be more specific, "non-youth Indigenous", "non-youth Black" and "non-youth White" are the three most significantly different groups than the rest.

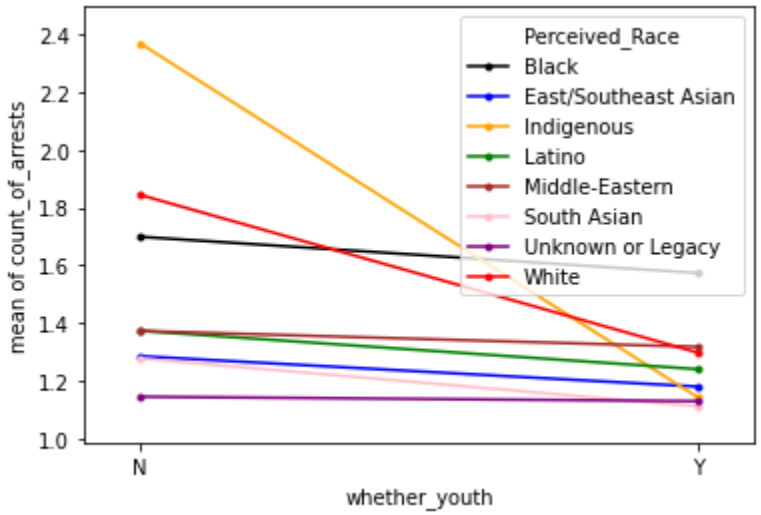

Figure 8. Interaction Plot

Finally, we visualize the interaction effect between youth status and race on the mean number of arrests for a more straightforward view. From the graph, it's clear that overall adults are having higher mean values in the number of arrests, according to the downward sloping lines from the non-youth group to the youth group. Though the adult Indigenous people has the largest mean value of arrest numbers, its youth counterpart shows an extreme opposite trend, reaching almost the lowest mean value of arrests among all the races in the youth group. Similar phenomenon appears on the White group, and the difference of mean arrest number between the youth and adult group is the second largest. As for the Black group, the difference does exist but smaller than the counterparts of Indigenous and White group. The remaining groups are having much smaller differences between adults and the youth group, due to the relatively lower mean values even in their adult groups.

4. Discussion and Conclusions

From the one-way ANOVA test result, we conclude that there exist differences between mean number of arrests among the eight ethnicities. Consistent with existing findings with respect to previous reports on the race-based data collected by Toronto Police, results of the Tukey HSD analysis further corroborated findings in which there appears greater burden 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, with the exception of individuals who were perceived to be Indigenous and disregarding the other factors such as action categories and attitudes during the arrests. Similar results were also observed with the Indigenous group when compared to all other minorities.

Therefore, there appears to be disparity in the number of arrests with respect to the Black and Indigenous populations when it comes to people of color.

From the two-way ANOVA test result, we fail to find the differences in the interaction between gender and races and the interaction between gender and being a youth combination in their mean number of being arrested. However, the test result does prove that there exist differences in the mean number of arrests between different combinations of races and whether the individual is youth or not. Further from the Tucky test, we conclude that overall non-youth tend to have a higher mean number of being arrested, whereas the mean number of arrests shows a decreasing trend from Indigenous to White, and finally, to the Black populations.

Despite the findings of the report, there are also limitations of the study. From the descriptive analytics and box plots, even there are three classes under the variable gender, the amount of gender-diverse group is very limited, therefore when we studied the gender differences, we gave up this group and focused only on the males and females. The same limitation applies in the ethnicity. As we only have relatively large amount data in the group of black, white and indigenous populations, our t-tests only tested on the three groups in terms of ethnicity study. Beyond that, our model is lacking the context of more details about each arrest record. For example, whereas the file does include the categories of criminal acts, we didn't incorporate this factor into our model, which may lead a biased result in studying the arrest amount differences with respect to ethnicity. Even if the average number of being arrested is the same among different groups, the unshown factor of categories contains information about the severity of criminals. Failing to take this into model consideration may mislead a wrong perception on average number of arrests among combinations of demographic features.

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

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

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology, 109(3),* 504-511.

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., & 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

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

Statistics Canada. (2009). *Female Offenders in Canada* (No. 85-002-XIE, Vol. 28, no. 1). Ottawa, ON.

Statistics Canada. (2022). Police-reported crime statistics in Canada, 2021 (No. 85-002-X). https://www150.statcan.gc.ca/n1/en/pub/85-002-x/2022001/article/00013-eng.pdf?st=AMC7fo29

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