Strip Searches, Arrests, and Strip Search Rate Analysis

by

Junwei Shen
Peiwen Zhang

A midterm report submitted in conformity with the
requirements for the course of INF2178H

Faculty of
Information
University of
Toronto

Table of Contents

Introduction/background

According to the guidelines of the Toronto Police Service (TPS), an arrest can be performed under the following circumstances. Firstly, if a police officer has a reasonable belief that the person has committed or is about to commit a criminal offense. Secondly, if the individual is found to be committing, or there is a reasonable belief that they are about to commit, a breach of peace. Thirdly, if the officer has a belief on reasonable grounds that the individual is wanted on a warrant of arrest or committal that is valid within the territorial jurisdiction where the person is located. A strip search is defined as a search that includes the removal of part of the clothing and a visual inspection of the body performed by the police officer based on the description of the dataset in the TPSe Public Safety data portal (2021). It is done when an accused person is being "booked," or brought into the station and is often justified by the police as necessary to recover hidden weapons or drugs and to protect the accused, other inmates, or officers.

The TPS has specific guidelines and protocols for conducting strip searches, including requirements for the presence of at least two officers of the same gender as the person being searched and the use of appropriate search equipment. Regardless of those restrictions, the TPS has been a topic of concern and facing intense scrutiny over years for their use of strip searches. According to Gills (2019), "allegations of racial profiling, excessive force, and violations of human rights have led to calls for reform and greater transparency in police practices." The TPS has modified the code of conduct on strip searches on November 4th, 2021. Data has shown that around 27% of arrests resulted in a strip search before the policy change, while only around 4.9% of arrests resulted in a strip search after changing the policy (Gills, 2019).

In this report, our group will use the data from the TPS Public Safety data portal to analyze the patterns and trends of arrests, strip searches, and the percentage between arrests and strip searches. Various other relevant sources, including media articles, and government guidelines were referenced to provide a background on this topic. This report aims to provide an overview of the Toronto police arrests, strip searches, and strip search rate analysis, highlighting some key findings, contributing to ongoing discussions, and practice for data analysis.

1.1 Research Questions

What is the association between sex, perceived race, and the number of strip searches? What is the association between sex, perceived race, and the number of arrests? What is the

association between sex, perceived race, and strip search rate, the percentage between strip searches and arrests? Are there any interactions between sex and perceived racial groups with respect to the number of strip searches or arrests? Which racial groups might have a significant difference in the number of strip searches and arrests?

Literature Review

The study will make use of two prior studies pertaining to the matter of Toronto arrests and strip searches, in order to enhance the comprehension of the issue at hand and to clarify the objective of the research.

The first study is "A collection Impact: Interim report on the inquiry into racial profiling and racial discrimination of Black persons by the Toronto Police Service" (2018) by the institution of Ontario Human Rights Commission (OHRC), which involved a methodical analysis and examination of research conducted on the apprehension of black individuals during the years 2013 to 2017. The study conducted a qualitative analysis and found that a Black person in Toronto was nearly 20 times more likely than a White person to be involved in a fatal shooting by the TPS. In addition, in the data collected by the OHRC from the Special Investigations Unit (SIU), it was found that Black individuals were disproportionately represented in cases involving the use of force (28.8%), shootings (36%), deadly encounters (61.5%), and fatal shootings (70%). Despite only constituting 4.1% of Toronto's population, Black men were complainants in 25% of SIU cases involving sexual assault allegations against TPS officers. Overall, the report arrived at the conclusion that the TPS displays racial prejudice, and prompt action is imperative to confront systemic racism within the TPS and ensure that everyone is treated justly and equitably.

The second study "'Clearly, we were doing it wrong': Toronto police are doing far fewer strip searches under strict new rules, Interim chief says" (2020) by Wendy G. investigated the outcomes of strip searches after the TPS implemented new rules and guidelines. The article highlighted that the implementation of the new policy has resulted in a substantial reduction in the frequency of strip searches performed by the TPS and has fostered better relations with the community. Additionally, the article emphasized the importance of ongoing monitoring and supervision to ensure that the TPS follows the new policy and does not engage in any discriminatory conduct.

Dataset Description

2.1 Dataset

This report will reference the arrests and strip searches dataset, titled "Arrests_and_Strip_Searches_(RBDC-ARR-TBL-001)", from the TPS Public Safety data portal with around 65276 records containing the date range from the beginning of 2020 to the November of 2021. The dataset contains age at arrest, sex, perceived race, demographic information on the individual arrested, types of charges they faced, and the circumstances surrounding their arrest. Besides that, the dataset contained whether the arrested individuals were being stripped searched, the reasons for strip searching, and whether items were found from the strip search.

This dataset provided transparency and accountability regarding TPS's use of these practices and allowed for the analysis of trends and patterns over time. It is an essential resource for analyzing the arrests trend, strip search trend, and the rate trend between strip search and arrests. Click here to view the original dataset.

2.2 Data preparation

During the data preparation process, our group followed the procedure of removing duplicates, dealing with missing values, and standardizing and reformatting the dataset. We found 4 duplicated entries based on the distinct values of *eventID*, *arrestID*, and *personID*. For the Search Reason columns, we filled in missing values as 0. Additionally, we dropped entries that were null in the *Age_group__at_arrest_* column and *Perceived_Race* column since the percentage of missing values was relatively small. After cleaning all duplicate rows and missing values, we reformatted the dataset by counting all entries as the number of arrests, summing the number of strip searches as the total number of strip searches performed, and calculating the strip search rate using division of the total number of strip searches by the total number of arrests. We grouped by *Arrest_Month*, *Sex*, and *Perceived_Race* columns to perform these calculations.

2.3 Measurement
2.3.1 Independent Variables

After identifying the research problems, we have defined the independent variables of

this research is the month when the individual was arrested *(Arrest_Month)*, the sex identified by the individual *(Sex)*, and the race recorded by the individual *(Perceived_Race)*. All three of these independent variables are categorical variables with *Sex* having two categories, *Perceived_Race* having eight categories, and *Arrest_Month* having four categories.

2.3.2 Dependent Variables

The dependent variables for this project are the number of strip searches *(StripSearch)*, the number of total arrests *(Arrests)*, and the strip search rate *(portion)*. All three of these variables are continous.

Descriptive Analysis

3.1 Tables

Given that the dependent variables are both categorical data, a cross-tabulation is performed for a better understanding of the mean and standard deviation of the independent variables for each category.

Based on the table presented, it can be observed that individuals with race 'White' and 'Black' exhibit a higher mean and standard deviation.

Table 3.1: Mean and Standard Deviation for each dependent variable

index	Sex	Perceived_Race	portion_mean	portion_std	StripSearch_mean	StripSearch_std	Arrests_mean	Arrests_std
0	F	Black	0.0825	0.055	63.75	45.67548576643712	752.0	87.47190024992787
1	F	East/Southeast Asian	0.034999999999999996	0.020816659994661327	6.25	4.112987559751022	186.25	17.55704986607944
2	F	Indigenous	0.1175	0.049916597106239795	18.5	9.0	157.25	19.771612647092464
3	F	Latino	0.035	0.02380476142847617	2.5	1.7320508075688772	66.0	15.534906930308056
4	F	Middle-Eastern	0.0475	0.032015621187164243	4.5	3.1091263510296048	91.0	13.391539617733775
5	F	South Asian	0.05	0.03915780041490244	6.75	6.499999999999999	127.25	25.03830398942122
6	F	Unknown or Legacy	0.095	0.06454972243679027	24.0	19.270011243726177	231.5	60.55575942881073
7	F	White	0.1225	0.0618465843842649	194.5	114.80563284670893	1540.5	188.64163555977422
8	M	Black	0.1475	0.07365459931328117	544.5	289.56691799996764	3627.0	286.5670369506351
9	M	East/Southeast Asian	0.09	0.04690415759823429	79.0	39.370039370059054	916.75	80.11814193219077
10	M	Indigenous	0.18	0.09380831519646858	58.0	30.994623189622207	325.75	25.617376914898998
11	M	Latino	0.0825	0.04349329450233296	30.5	17.253019059476713	375.75	44.76885077819175
12	M	Middle-Eastern	0.075	0.045092497528228955	52.5	32.39855963876584	718.25	32.59217697546452
13	M	South Asian	0.075	0.045092497528228955	57.5	34.79942528261063	776.0	39.22584182228175
14	M	Unknown or Legacy	0.10250000000000001	0.05188127472091127	110.0	59.486693189877904	1031.5	121.47564913732023
15	M	White	0.1275	0.06849574196011506	696.75	386.6776219712402	5386.0	321.87885920016555

3.2 Bar Plots

Bar plots were utilized to illustrate the temporal changes in the total number of strip searches, number of arrests, and average strip search rate, thereby providing a visual depiction of the aforementioned variables. According to Figure 3.2, the majority of strip searches and arrests were carried out between January and March, whereas the number of strip searches and arrests

was at its lowest between April and June. The average strip search rate also displayed a similar trend to the overall number of strip searches during this period.

Figure 3.2: Bar plot of the total number of Arrests per month, the total number of searches per month, and average strip search percentage per month
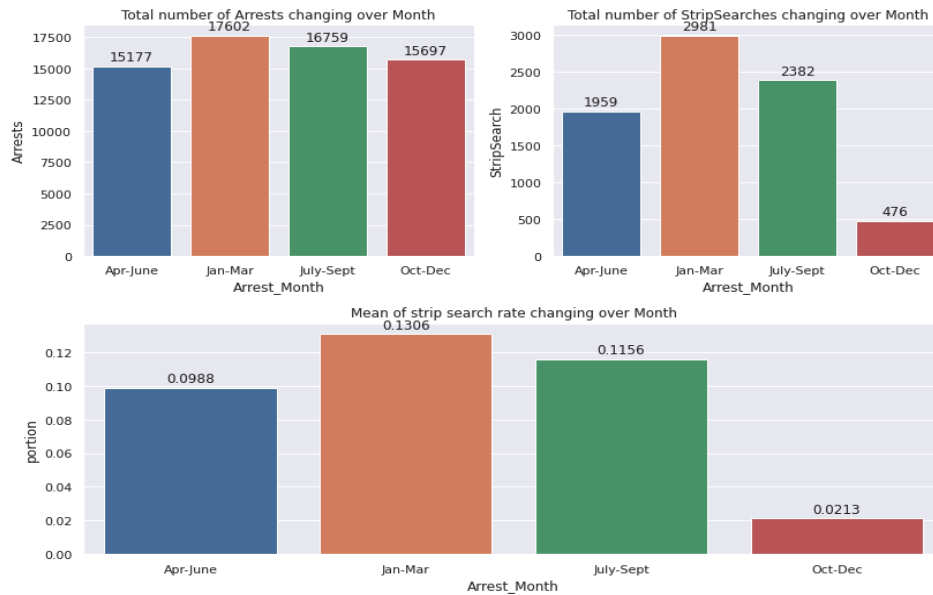
3.3 BoxPlots

Box plots were generated to explore the distribution of each dependent variable across various categories, which allowed for a more comprehensive examination of central tendency, symmetry, skewness, and outliers.

As we can see from Figure 3.3, the boxplot of the *StripSearch* variable within the female group has a positive skew due to the fact that the median is near the start of the plot and the first whisker is shorter than the last whisker. The strip search under the male category has a higher median than the female category. In the plot of the strip searches with race breakdown, evidence showed that the 'White' and 'Black' racial groups have obvious differences from other racial groups. More investigation into this is required further. Moreover, the minimum value of strip searches under the 'East/Southeast Asian', 'Indigenous', 'Latino', 'Middle-Eastern', and 'South Asian' categories is equal to the lower quartile. The results from the Arrests variable boxplots are similar to those above, as shown in Figure 3.4. From Figure 3.5, the median difference in strip search rate under the race category has decreased.

Figure 3.3: Strip Search with Sex Breakdown Boxplot, and Strip Search with Race breakdown Boxplot
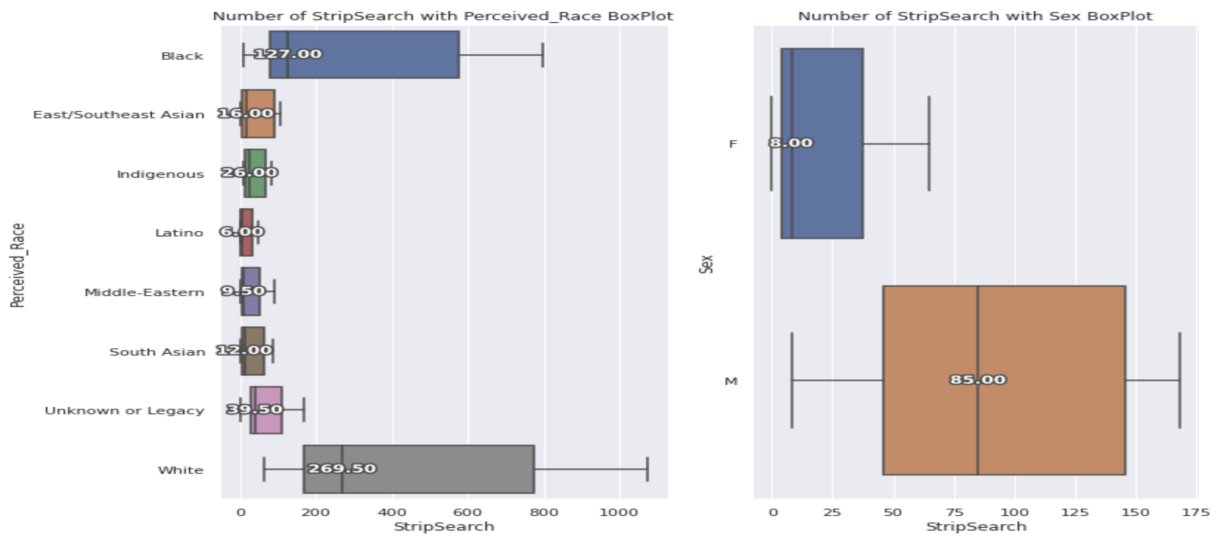
Figure 3.4: Arrests with Sex Breakdown Boxplot, and Arrests with Race breakdown Boxplot
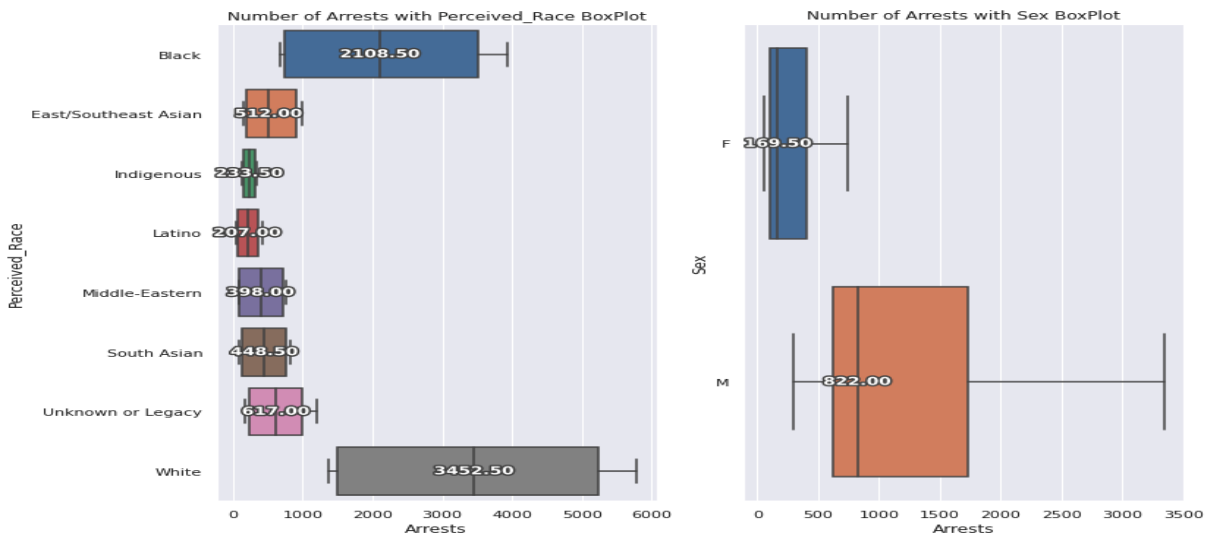
Figure 3.5: Strip Search Rate with Sex Breakdown Boxplot, and Strip Search Rate with Race breakdown Boxplot
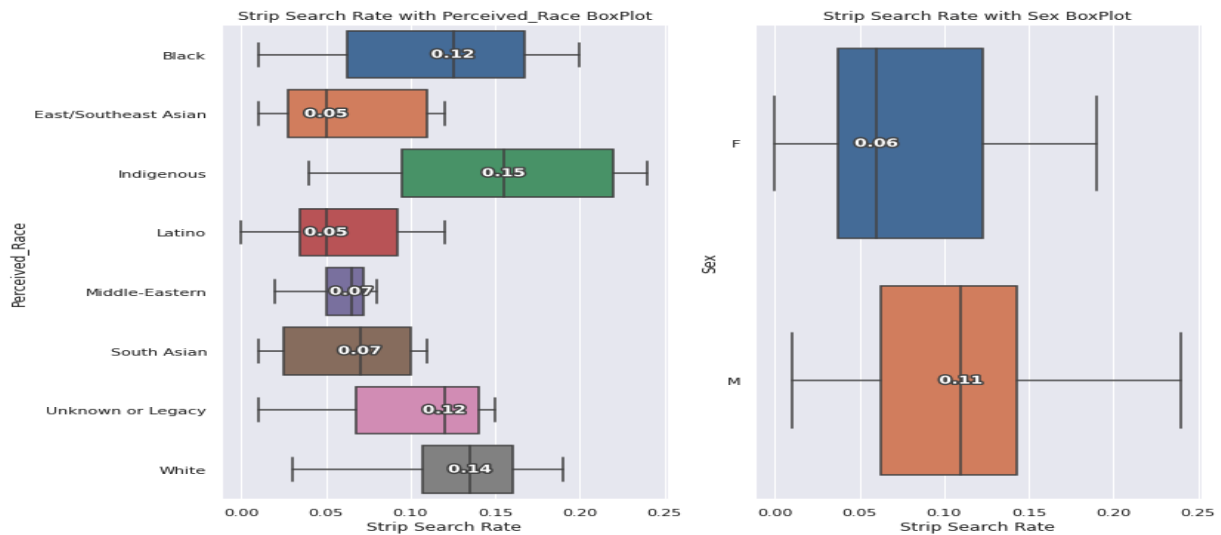
3.4 Scatter Plots

Considering that *StripSearch* and *Arrests* are ordinal variables, a scatter plot was generated to visualize the relationship between these two variables. Figure 3.6 demonstrates that a simple linear relationship might exist between these two dependent variables. However, we will not explore the relationship between these two variables in greater depth, as it falls outside the scope of the current research topics.

Figure 3.6: Scatter Plot between number of Strip Search and Arrests
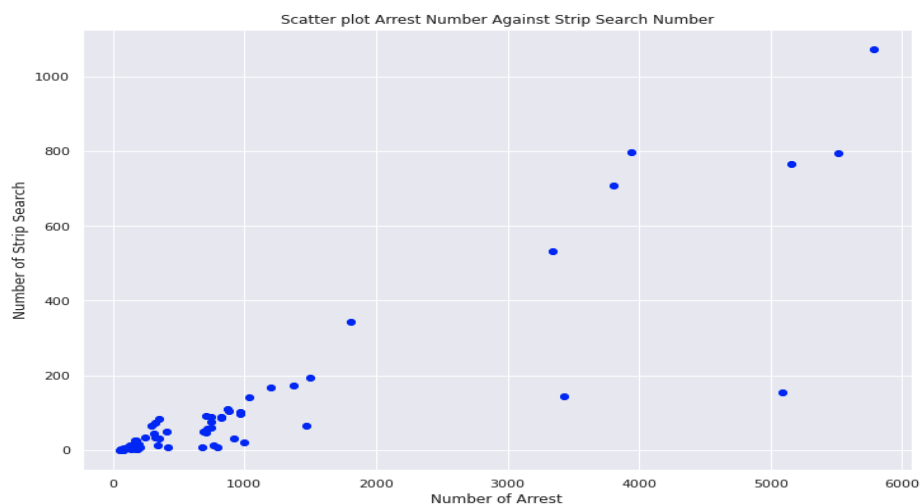
Methods

4.1 Research Design

Based on the research questions, this study put forth several hypotheses, such as "the average number of strip searches is equal for male and female," "the average number of arrests is equal for Black and White individuals," "the average number of strip searches does not differ between male and female in different racial groups," etc. Further establishment of the hypothesis and hypothesis testing would be specified in the following sections.

4.2 Brief Methods

By performing the data manipulation and EDA processes, the research identified the variables *Arrests* and *Strip Searches* as continuous outcomes, while the variables *Sex* and *Perceived Race* are identified as categorical predictors. The questions for investigation prove the necessity for proposing hypotheses and forming hypothesis testing. Specifically, the team proposed the following null hypotheses for hypothesis testing:

1. A. H_0: The mean numbers of strip searches between different sex groups are the same. (H_0: $\mu_{SS,M}=\mu_{SS,F}$)

   B. H_0: The mean numbers of strip searches between different perceived racial groups are the same. (H_0: $\mu_{SS,i}=\mu_{SS,j}$ for **all** i, j, where i, j belongs to the racial groups in the dataset and i$\neq$j)

2. A. H_0: The mean numbers of arrests between different sex groups are the same. (H_0: $\mu_{A,M}=\mu_{A,F}$)

   B. H_0: The mean numbers of arrests between different perceived racial groups are the same. (H_0: $\mu_{A,i}=\mu_{A,j}$ for **all** i, j, where i, j belongs to the racial groups in the dataset and i$\neq$ j)

3. A. H_0: The mean strip search rates between different sex groups are the same. (H_0: $\mu_{R,M}=\mu_{R,F}$)

   B. H_0: The mean strip search rates between different perceived racial groups are the same. (H_0: $\mu_{R,i}=\mu_{R,j}$ for **all** i, j, where i, j belongs to the racial groups in the dataset and i$\neq$j)

4. There is no significant difference between the means of number of strip searches or arrests of any two groups being compared. In detail, given any sex (Male or Female), the

change in racial groups would not affect the mean number of strip searches or arrests.
(H_0: $\mu_i - \mu_j = 0$, for **all** i and j, where μ_i and μ_j are the means of any two groups being compared)

5. A. H_0: There is no significant difference between the means of the number of strip searches between two specific racial groups.
(H_0: $\mu_{SS,i}=\mu_{SS,j}$ for **some** i, j, where i, j belongs to the racial groups in the dataset and i≠j)
B. H_0: There is no significant difference between the means of the number of arrests between two specific racial groups.
(H_0: $\mu_{A,i}=\mu_{A,j}$ for **some** i, j, where i, j belongs to the racial groups in the dataset and i≠j)

The project proposed several hypothesis testing methods such as the T-tests and the ANOVA (including one-way and two-way ANOVA) to compare means between groups or samples. There are some key differences between the two tests: 1) T-tests are used to compare the means of two groups or samples, while ANOVA is used to compare the means of three or more groups or samples. Moreover, for hypothesis 5, the report introduced Tukey's HSD for post-hoc tests to investigate the means of two different groups among various categories. The detailed design of T-tests, ANOVA, and Tukey's HSD will be explained in the following sections.

4.3 Welch's T-tests

The project will apply Welch's T-tests method to determine whether there is a significant difference between the means of two groups for each dependent variables, in our case, the means of males and females. The t-statistic of this project is calculated by dividing the means of strip searches and arrests between males and females by the standard error of the difference. The resulting t-value is then compared to a critical value (threshold α=0.05) based on the degrees of freedom and the significance level to determine whether there is a statistically significant difference between the two groups. Before running the Welch's T test, The following two assumptions were fullfilled: 1) Normality: The data in male and female groups follows a normal distribution, 2) Independence: The two groups are assumed to be independent of each other.

4.4 ANOVA

The project also applied ANOVA methods to study the difference in means of strip searches and arrests between different sex groups and perceived racial groups. ANOVA compares the means of the different groups, and it determines whether the means of the groups are significantly different from each other, based on the variance within and between the groups. That is, if the variation between the groups is significantly larger than the variation within the groups, it suggests that there are significant differences between the means of the groups. The project has included one-way ANOVA tests for the explanatory variables *Sex* and *Perceived_Race*, and two-way ANOVA tests for both the explanatory variables and their interaction term. The assumptions of normality and independence are fulfilled, while the homoscedasticity is relaxed and it will be discussed in limitations.

4.5 Tukey's HSD

It is important to notice that two-way ANOVA will only be able to verify if there are any significant differences among various racial groups, but it is not capable of comparing any two random selected racial groups. Therefore, the project taked Tukey's HSD approach to identify pairs of groups that are significantly different from each other, for example, strip search rates between Black and Latino, after an overall significant effect has been found by ANOVA.

Result and Findings

5.1 T-tests Analysis and Findings

Sex and Strip Searches T-tests Result

Table 5.1: T-test results for Sex and Strip Searches

	T	dof	alternative	p-val	CI95%	cohen-d	BF10	power
T-test	-3.07216	62	two-sided	0.003155	[-269.89, -57.11]	0.76804	11.854	0.856399

In Table 5.1, the T-test result for comparing means in the number of strip searches between male and female. Assume the threshold α=0.05, the p-value is 0.003155 and the 95% CI does not

include 0, the evidence suggests that we can reject the null hypothesis, arriving at the result of there is significant difference in means of the number of strip searches between male and female ($\mu_{SS,M} \neq \mu_{SS,F}$), meaning that if we randomly select samples of male and female from the dataset, they would likely to show a difference in means of the number of strip searches.

Sex and Arrests T-tests Result

Table 5.2: T-test results for Sex and Arrests

	T	dof	alternative	p-val	CI95%	cohen-d	BF10	power
T-test	3.875268	62	two-sided	0.00026	[605.53, 1895.78]	0.968817	97.26	0.968174

In Table 5.2, we have the T-test result for comparing means in the number of arrests between males and females. Assume the threshold α=0.05, the p-value is 0.00026 and the 95% CI does not include 0, the evidence suggests that we can reject the null hypothesis, arriving at the result of there is significant difference in means of the number of arrests between males and females ($\mu_{A,M} \neq \mu_{A,F}$), meaning that if we randomly select samples of males and females from the dataset, they would likely show a difference in means of the number of arrests.

Sex and Strip Search Rate T-tests Result

Table 5.3: T-test results for Sex and Strip Search Rate

	T	dof	alternative	p-val	CI95%	cohen-d	BF10	power
T-test	2.488122	62	two-sided	0.015543	[0.01, 0.07]	0.622031	3.293	0.687749

In Table 5.3, we have the T-test result for comparing means in the strip search rate between males and females. Assume the threshold α=0.05, the p-value is 0.015543 and the 95% CI does not include 0, the evidence suggests that we can reject the null hypothesis, arriving at the result of there is significant difference in means of the number of arrests between males and females ($\mu_{R,M} \neq \mu_{R,F}$), meaning that if we randomly select samples of males and females from the dataset, they would likely show a difference in means of the strip search rate.

5.2 Anova Analysis and Findings

Strip Searches ANOVA Analysis and Findings

Table 5.4: One-way ANOVA result between sex and strip searches

```
    Source          SS  DF           MS        F     p-unc        np2
0      Sex  4.277160e+05   1  427716.000000  9.43817  0.003155  0.132117
1   Within  2.809696e+06  62   45317.684476     NaN       NaN       NaN
```

In Table 5.4, we have the ANOVA result between sex and the number of strip searches. Assume the threshold $\alpha=0.05$, the p-value of 0.003155 suggests that we can reject the null hypothesis arriving at the result of there is significant difference in means of the number of strip searches between males and females ($\mu_{SS,M} \neq \mu_{SS,F}$), meaning that if we randomly select samples of males and females from the dataset, they would likely show a difference in means of the number of strip searches.

Table 5.5: One-way ANOVA result between perceived race and strip searches

```
          Source          SS  DF           MS        F     p-unc  \
0  Perceived_Race  1.457535e+06   7  208219.348214  6.551174  0.000012
1          Within  1.779877e+06  56   31783.517857       NaN       NaN

        np2
0  0.450216
1       NaN
```

In Table 5.5, we have the ANOVA result between perceived race and the number of strip searches. Assume the threshold $\alpha=0.05$, the p-value of 0.000012 suggests that we can reject the null hypothesis, arriving at the result of there is significant difference in means of the number of strip searches between all racial groups ($\mu_{SS,i} \neq \mu_{SS,j}$ for **all** i, j, where i, j belongs to the racial groups in the dataset and $i \neq j$), meaning that if we randomly select samples from all racial groups from the dataset, at least some of them would show a difference in means of the number of strip searches.

Table 5.6: Tukey's HSD of sex with respect to strip searches

```
Tukey HSD results for Sex with respect to StripSearch:
Multiple Comparison of Means - Tukey HSD, FWER=0.05
===================================================
group1 group2 meandiff p-adj  lower   upper  reject
---------------------------------------------------
    F      M    163.5 0.0032 57.115 269.885   True
---------------------------------------------------
```

In Table 5.6, we have Tukey's HSD result between sex and the number of strip searches. Assuming the threshold $\alpha=0.05$, the p-value of 0.0032 suggests that we can reject the null

hypothesis, arriving at the result of there is significant difference in means of the number of strip searches between males and females ($\mu_{SS,M} \neq \mu_{SS,F}$). It supports the result of the One-way ANOVA test by showing more details that male has 163.5 more mean strip search number than female.

Table 5.7: Tukey's HSD of perceived race with respect to strip searches

```
Tukey HSD results for Perceived_Race with respect to StripSearch:
            Multiple Comparison of Means - Tukey HSD, FWER=0.05
==============================================================================
      group1            group2      meandiff p-adj    lower    upper   reject
------------------------------------------------------------------------------
           Black East/Southeast Asian   -261.5 0.0849 -542.1369  19.1369  False
           Black           Indigenous -265.875 0.0755 -546.5119  14.7619  False
           Black               Latino -287.625 0.0408 -568.2619  -6.9881   True
           Black       Middle-Eastern -275.625 0.0577 -556.2619   5.0119  False
           Black          South Asian   -272.0 0.0638 -552.6369   8.6369  False
           Black    Unknown or Legacy -237.125 0.1569 -517.7619  43.5119  False
           Black                White    141.5 0.7556 -139.1369 422.1369  False
East/Southeast Asian         Indigenous   -4.375    1.0 -285.0119 276.2619  False
East/Southeast Asian             Latino  -26.125    1.0 -306.7619 254.5119  False
East/Southeast Asian     Middle-Eastern  -14.125    1.0 -294.7619 266.5119  False
East/Southeast Asian        South Asian    -10.5    1.0 -291.1369 270.1369  False
East/Southeast Asian  Unknown or Legacy   24.375    1.0 -256.2619 305.0119  False
East/Southeast Asian              White    403.0 0.0008  122.3631 683.6369   True
      Indigenous             Latino   -21.75    1.0 -302.3869 258.8869  False
      Indigenous     Middle-Eastern    -9.75    1.0 -290.3869 270.8869  False
      Indigenous        South Asian   -6.125    1.0 -286.7619 274.5119  False
      Indigenous  Unknown or Legacy    28.75    1.0 -251.8869 309.3869  False
      Indigenous              White  407.375 0.0007  126.7381 688.0119   True
          Latino     Middle-Eastern     12.0    1.0 -268.6369 292.6369  False
          Latino        South Asian   15.625    1.0 -265.0119 296.2619  False
          Latino  Unknown or Legacy     50.5 0.9991 -230.1369 331.1369  False
          Latino              White  429.125 0.0003  148.4881 709.7619   True
  Middle-Eastern        South Asian    3.625    1.0 -277.0119 284.2619  False
  Middle-Eastern  Unknown or Legacy     38.5 0.9999 -242.1369 319.1369  False
  Middle-Eastern              White  417.125 0.0005  136.4881 697.7619   True
     South Asian  Unknown or Legacy   34.875 0.9999 -245.7619 315.5119  False
     South Asian              White    413.5 0.0005  132.8631 694.1369   True
Unknown or Legacy              White  378.625  0.002   97.9881 659.2619   True
------------------------------------------------------------------------------
```

In Table 5.7, we have Tukey's HSD result between perceived race pairs and the number of strip searches. It is worth noting that the means of strip searches between 1) Black and Latino, 2) East/Southeast Asian and White, 3) Indigenous and White, 4) Latino and White, 5) Middle Eastern and White, 6) South Asian and White all have significant differences. Specifically, 1) the racial group Black is likely to have 287.625 more mean strip searches than Latino, 2) the racial group East/Southeast Asian and White is likely to have 403 fewer mean strip searches than White, 3) The racial group Indigenous is likely to have 407.375 fewer mean strip searches than White, 4) The racial group Latino is likely to have 429.125 fewer mean strip searches than White, 5) The racial group Middle-Eastern is likely to have 417.125 fewer mean strip searches than White, 6) The racial group South Asian is likely to have 413.5 fewer mean strip searches than White.

Table 5.8: Two-way ANOVA result for perceived race, sex with respect to strip searches

	sum_sq	df	F	PR(>F)
C(Sex)	4.277160e+05	1.0	26.549000	4.785296e-06
C(Perceived_Race)	1.457535e+06	7.0	12.924500	3.335448e-09
C(Sex):C(Perceived_Race)	5.788600e+05	7.0	5.132963	2.079854e-04
Residual	7.733010e+05	48.0	NaN	NaN

In table 5.8, we have the Two-way ANOVA result for perceived race, sex with respect to strip searches. The model was built using Ordinary Least Square. First of all, the p-values of the effects from sex and perceived race are less than the threshold $\alpha=0.05$, which suggests that their effects are significant to the number of strip searches. Secondly, with the p-value of the interaction term less than the threshold $\alpha=0.05$, it suggests that there are some interactions between sex and perceived race with respect to the number of strip searches. That is, the effect of sex on the number of strip searches depend on the level of perceived race, by selecting a different race, the number of strip searches will likely be different.

Arrests Analysis ANOVA and Findings

Table 5.9: One-way ANOVA result between sex and arrests

	Source	SS	DF	MS	F	p-unc	np2
0	Sex	2.502626e+07	1	2.502626e+07	15.017704	0.00026	0.19499
1	Within	1.033199e+08	62	1.666450e+06	NaN	NaN	NaN

In Table 5.9, we have the ANOVA result between sex and the number of arrests. Assume the threshold $\alpha=0.05$, the p-value of 0.00026 suggests that we can reject the null hypothesis, arriving at the result of there is significant difference in means of the number of arrests between males and females ($\mu_{A,M} \neq \mu_{A,F}$), meaning that if we randomly select samples of males and females from the dataset, they would likely show a difference in means of the number of arrests.

Table 5.10: One-way ANOVA result between perceived race and arrests

	Source	SS	DF	MS	F	p-unc	\
0	Perceived_Race	7.723217e+07	7	1.103317e+07	12.087828	2.752245e-09	
1	Within	5.111401e+07	56	9.127501e+05	NaN	NaN	

	np2
0	0.601749
1	NaN

In Table 5.10, we have the ANOVA result between perceived race and the number of arrests. Assume the threshold $\alpha=0.05$, the p-value of 2.75×10^{-9} suggests that we can reject the

null hypothesis, arriving at the result of there is significant difference in means of the number of arrests between all racial groups ($\mu_{A,i} \neq \mu_{A,j}$ for **all** i, j, where i, j belongs to the racial groups in the dataset and i≠j), meaning that if we randomly select samples from all racial groups from the dataset, at least some of them would show a difference in means of the number of arrests.

Table 5.11: Tukey's HSD of sex with respect to arrests

```
Tukey HSD results for Sex with respect to Arrests:
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
=========================================================
group1 group2  meandiff p-adj   lower     upper    reject
---------------------------------------------------------
    F       M 1250.6562 0.0003 605.5329 1895.7796   True
---------------------------------------------------------
```

In Table 5.11, we have Tukey's HSD result between sex and the number of arrests. Assuming the threshold α=0.05, the p-value of 0.0003 suggests that we can reject the null hypothesis, arriving at the result of there is significant difference in means of the number of arrests between males and females ($\mu_{A,M} \neq \mu_{A,F}$). It supports the result of the One-way ANOVA test by showing more details that male has 11250.6562 more mean arrest number than female.

Table 5.12: Tukey's HSD of perceived race with respect to arrests

```
Tukey HSD results for Perceived_Race with respect to Arrests:
            Multiple Comparison of Means - Tukey HSD, FWER=0.05
==============================================================================
      group1              group2       meandiff p-adj   lower      upper   reject
------------------------------------------------------------------------------
              Black East/Southeast Asian  -1638.0 0.0236 -3141.9025 -134.0975  True
              Black          Indigenous   -1948.0 0.0034 -3451.9025 -444.0975  True
              Black              Latino  -1968.625  0.003 -3472.5275 -464.7225  True
              Black      Middle-Eastern -1784.875 0.0098 -3288.7775 -280.9725  True
              Black         South Asian -1737.875  0.013 -3241.7775 -233.9725  True
              Black   Unknown or Legacy   -1558.0 0.0372 -3061.9025  -54.0975  True
              Black               White   1273.75 0.1548  -230.1525 2777.6525 False
East/Southeast Asian      Indigenous    -310.0 0.9979 -1813.9025 1193.9025 False
East/Southeast Asian          Latino  -330.625 0.9969 -1834.5275 1173.2775 False
East/Southeast Asian  Middle-Eastern  -146.875    1.0 -1650.7775 1357.0275 False
East/Southeast Asian     South Asian   -99.875    1.0 -1603.7775 1404.0275 False
East/Southeast Asian Unknown or Legacy    80.0    1.0 -1423.9025 1583.9025 False
East/Southeast Asian           White   2911.75    0.0  1407.8475 4415.6525  True
         Indigenous          Latino   -20.625    1.0 -1524.5275 1483.2775 False
         Indigenous  Middle-Eastern   163.125    1.0 -1340.7775 1667.0275 False
         Indigenous     South Asian   210.125 0.9998 -1293.7775 1714.0275 False
         Indigenous Unknown or Legacy   390.0 0.9915 -1113.9025 1893.9025 False
         Indigenous           White   3221.75    0.0  1717.8475 4725.6525  True
             Latino  Middle-Eastern   183.75 0.9999 -1320.1525 1687.6525 False
             Latino     South Asian   230.75 0.9997 -1273.1525 1734.6525 False
             Latino Unknown or Legacy  410.625 0.9884 -1093.2775 1914.5275 False
             Latino           White  3242.375    0.0  1738.4725 4746.2775  True
     Middle-Eastern     South Asian    47.0    1.0 -1456.9025 1550.9025 False
     Middle-Eastern Unknown or Legacy  226.875 0.9997 -1277.0275 1730.7775 False
     Middle-Eastern           White  3058.625    0.0  1554.7225 4562.5275  True
        South Asian Unknown or Legacy  179.875 0.9999 -1324.0275 1683.7775 False
        South Asian           White  3011.625    0.0  1507.7225 4515.5275  True
  Unknown or Legacy           White   2831.75    0.0  1327.8475 4335.6525  True
------------------------------------------------------------------------------
```

In Table 5.12, we have Tukey's HSD result between perceived race pairs and the number of

arrests. It is worth noting that the means of arrests between 1) Black and all racial groups except White, 2) East/Southeast Asian and White, 3) Indigenous and White, 4) Latino and White, 5) Middle Eastern and White, 6) South Asian and White all have significant differences. Specifically, 1) the racial group Black is likely to have more mean strip searches than all racial groups except White, 2) the racial group East/Southeast Asian and White is likely to have 2911.75 fewer mean arrests than White, 3) The racial group Indigenous is likely to have 3221.75 fewer mean arrests than White, 4) The racial group Latino is likely to have 3242.375 fewer mean arrests than White, 5) The racial group Middle-Eastern is likely to have 3058.625 fewer mean arrests than White, 6) The racial group South Asian is likely to have 3011.625 fewer mean arrests than White.

Table 5.13: Two-way ANOVA result for perceived race, sex with respect to arrests

	sum_sq	df	F	PR(>F)
C(Sex)	2.502626e+07	1.0	1535.270278	4.226101e-38
C(Perceived_Race)	7.723217e+07	7.0	676.844846	1.059587e-45
C(Sex):C(Perceived_Race)	2.530531e+07	7.0	221.769856	2.640375e-34
Residual	7.824422e+05	48.0	NaN	NaN

In table 5.13, we have the Two-way ANOVA result for perceived race, sex with respect to arrests. The model was built using Ordinary Least Square. First of all, the p-value of the effect from sex and perceived race is less than the threshold $\alpha=0.05$, which suggests that their effect is significant to the number of arrests. Secondly, with the p-value of the interaction term less than the threshold $\alpha=0.05$, it suggests that there are some interactions between sex and perceived race with respect to the number of arrests. That is, the effect of sex on the number of arrests depend on the level of perceived race, by selecting a different race, the number of arrests will likely be different.

Strip Search Rate ANOVA Analysis and Findings

Table 5.14: One-way ANOVA result between sex and strip search rate

	Source	SS	DF	MS	F	p-unc	np2
0	Sex	0.021756	1	0.021756	6.190752	0.015543	0.090786
1	Within	0.217888	62	0.003514	NaN	NaN	NaN

In Table 5.14, we have the ANOVA result between sex and the strip search rate. Assume the threshold $\alpha=0.05$, the p-value of 0.015543 suggests that we can reject the null hypothesis,

arriving at the result of there is significant difference in means of the strip search rate between male and female ($\mu_{R,M} \neq \mu_{R,F}$), meaning that if we randomly select samples of male and female from the dataset, they would likely to show a difference in means of strip search rate.

Table 5.15: One-way ANOVA result between perceived race and strip search rate

```
        Source      SS  DF       MS        F    p-unc      np2
0  Perceived_Race  0.069394   7  0.009913  3.260793  0.005646  0.28957
1          Within  0.170250  56  0.003040       NaN       NaN      NaN
```

In Table 5.15, we have the ANOVA result between perceived race and the strip search rate. Assume the threshold $\alpha=0.05$, the p-value of 0.005646 suggests that we can reject the null hypothesis, arriving at the result of there is significant difference in means of the strip search rate between all racial groups ($\mu_{R,i} \neq \mu_{R,j}$ for **all** i, j, where i, j belongs to the racial groups in the dataset and $i \neq j$), meaning that if we randomly select samples from all racial groups from the dataset, at least some of them would show a difference in means of the strip search rate.

Table 5.16: Tukey's HSD of sex with respect to strip search rate

```
Tukey HSD results for Sex with respect to portion:
Multiple Comparison of Means - Tukey HSD, FWER=0.05
===================================================
group1 group2 meandiff p-adj  lower  upper  reject
---------------------------------------------------
    F      M   0.0369 0.0155 0.0072 0.0665   True
---------------------------------------------------
```

In Table 5.16, we have Tukey's HSD result between sex and the strip search rate. Assuming the threshold $\alpha=0.05$, the p-value of 0.0155 suggests that we can reject the null hypothesis, arriving at the result of there is significant difference in means of the strip search rate between males and females ($\mu_{R,M} \neq \mu_{R,F}$). It supports the result of the One-way ANOVA test by showing more details that male has 0.0369 more mean strip search rate than female.

Table 5.17: Tukey's HSD of perceived race with respect to strip search rate

```
Tukey HSD results for perceived race with respect to portion:
          Multiple Comparison of Means - Tukey HSD, FWER=0.05
==================================================================================
       group1              group2      meandiff p-adj   lower    upper  reject
----------------------------------------------------------------------------------
              Black East/Southeast Asian  -0.0525 0.5539 -0.1393   0.0343  False
              Black            Indigenous   0.0338 0.9212  -0.053   0.1205  False
              Black                Latino  -0.0563 0.4651   -0.143   0.0305  False
              Black        Middle-Eastern  -0.0538  0.524 -0.1405    0.033  False
              Black            South Asian  -0.0525 0.5539 -0.1393   0.0343  False
              Black     Unknown or Legacy  -0.0163 0.9989   -0.103   0.0705  False
              Black                 White     0.01    1.0  -0.0768   0.0968  False
East/Southeast Asian          Indigenous   0.0863 0.0526  -0.0005    0.173  False
East/Southeast Asian              Latino  -0.0037    1.0  -0.0905    0.083  False
East/Southeast Asian      Middle-Eastern  -0.0012    1.0   -0.088   0.0855  False
East/Southeast Asian          South Asian     0.0    1.0  -0.0868   0.0868  False
East/Southeast Asian   Unknown or Legacy   0.0363 0.8894  -0.0505    0.123  False
East/Southeast Asian               White   0.0625 0.3297  -0.0243   0.1493  False
         Indigenous               Latino    -0.09 0.0369  -0.1768  -0.0032   True
         Indigenous       Middle-Eastern  -0.0875 0.0468  -0.1743  -0.0007   True
         Indigenous          South Asian  -0.0863 0.0526   -0.173   0.0005  False
         Indigenous   Unknown or Legacy    -0.05 0.6137  -0.1368   0.0368  False
         Indigenous                White  -0.0238 0.9883  -0.1105    0.063  False
             Latino       Middle-Eastern   0.0025    1.0  -0.0843   0.0893  False
             Latino          South Asian   0.0037    1.0   -0.083   0.0905  False
             Latino   Unknown or Legacy     0.04 0.8293  -0.0468   0.1268  False
             Latino                White   0.0662 0.2601  -0.0205    0.153  False
     Middle-Eastern          South Asian   0.0012    1.0  -0.0855    0.088  False
     Middle-Eastern   Unknown or Legacy   0.0375  0.871  -0.0493   0.1243  False
     Middle-Eastern                White   0.0638 0.3054   -0.023   0.1505  False
        South Asian   Unknown or Legacy   0.0363 0.8894  -0.0505    0.123  False
        South Asian                White   0.0625 0.3297  -0.0243   0.1493  False
  Unknown or Legacy                White   0.0262 0.9791  -0.0605    0.113  False
----------------------------------------------------------------------------------
```

In Table 5.17, we have Tukey's HSD result between perceived race pairs and the strip search rate. It is worth noting that the means of arrests between 1) Indigenous and Latino, 2) Indigenous and Middle Eastern have a significant difference in the mean strip search rate. Specifically, 1) the racial group Indigenous is likely to have 0.09 more mean strip search rate than Latino, 2) the racial group Indigenous is likely to have 0.0863 more mean strip search rate than Middle Eastern.

Table 5.18: Two-way ANOVA result for perceived race, sex with respect to strip search rate

	sum_sq	df	F	PR(>F)
C(Sex)	0.021756	1.0	7.432740	0.008915
C(Perceived_Race)	0.069394	7.0	3.386782	0.005110
C(Sex):C(Perceived_Race)	0.007994	7.0	0.390137	0.903576
Residual	0.140500	48.0	NaN	NaN

In table 5.18, we have the Two-way ANOVA result for perceived race, sex with respect to the strip search rate. The model was built using Ordinary Least Square. First of all, the p-value of the effect from sex and perceived race is less than the threshold $\alpha=0.05$, which suggests that their effect is significant to the strip search rate. Secondly, with the p-value of the interaction term

greater than the threshold $\alpha=0.05$, it does not reject the null hypothesis, meaning it suggests that there are no interactions between sex and perceived race with respect to the strip search rate. That is, the effect of sex on the strip search rate does not depend on the level of perceived race, by selecting a different race, the strip search rate would not be different.

5.3 Interaction Plots

Although the interaction plots may not reveal any statistically significant differences, they can provide valuable insights for better understanding the data.

In Figure 5.19, the plot showed: 1) males have higher mean strip search numbers than females; 2) there is a noticeable big difference in mean strip search numbers between black males and black females, as well as white males and white females; 3) there is a noticeable big difference in mean strip search numbers between Black and the rest of the racial groups, as well as White and the rest of the racial groups in both male samples. However, this difference is not too significant in female samples in comparison. The interaction plot for Arrests showed similar pattern with the strip search plot as shown in Figure 5.20.

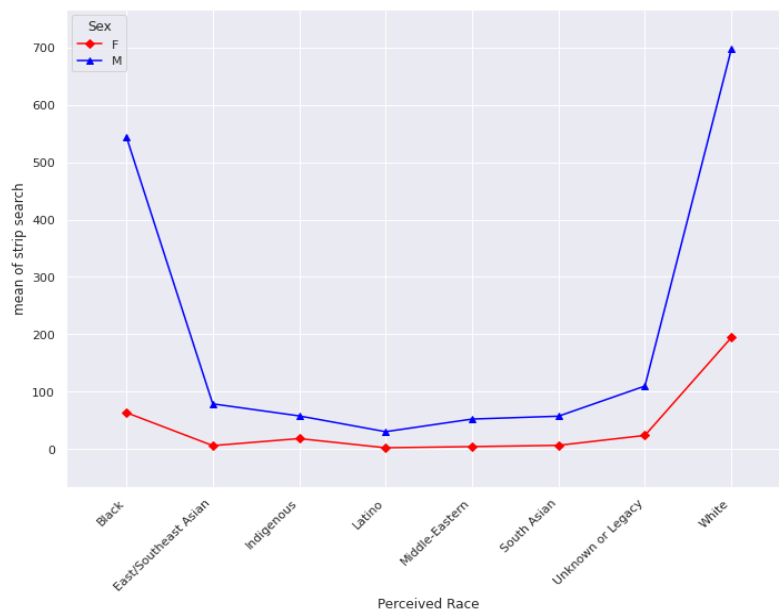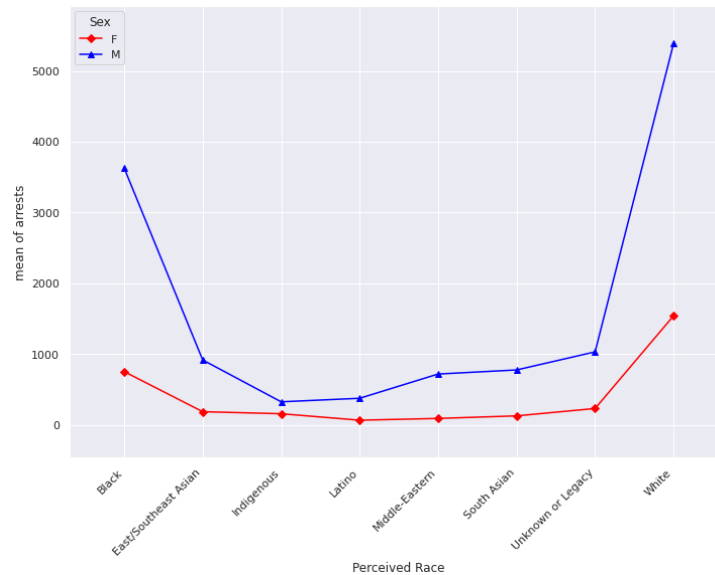*Figure 5.19: Interaction plot of Sex, Perceived Race, and Strip Searches*

Figure 5.20: Interaction plot of Sex, Perceived Race, and Arrests

As for the strip search rate, in Firgure 5.21, the following three observations were made. 1) males have a higher mean strip search rate than females; 2) there is a noticeable big difference in mean strip search rate between black males and females, East/Southeast Asian males and females, as well as white Indigenous females; 3) the two lines indicating males and females are almost "parallel" to each other, suggesting that there might not be interactions between perceived race and sex with respect to mean strip search rate, this aligns with the result we got from Two-way ANOVA test in table 5.18.

Figure 5.21: Interaction plot of Sex, Perceived Race, and Strip Search Rate
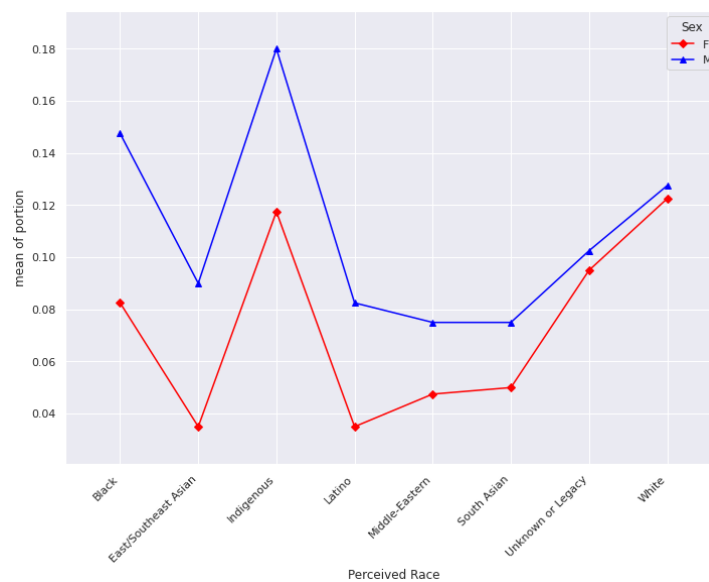
Discussion and Conclusions

6.1 Limitations

This report may have limitations, including statistical hypothesis and causal inferece. Three assumptions need to be fulfilled for the ANOVA tests, while the dataset only filled two of those. The homoscedasticity is not made in the dataset. One of the plausible reasons might be the error in the data restructure process. While statistical analysis can identity patterns and associations between variable, it cannot establish causality. That said, police will not perform the strip searches because the individual have 'White' race or 'Male' gender. Besides that, there are other factors not included in the analysis may be responsible for the observed relationships.

6.2 Conclusion

In conclusion, the project provided an investigation of the arrests and strip searches dataset provided by the TPS Public Safety data portal. The team performed analysis and visualizations of the patterns and trends of arrests, strip searches, and the percentage between arrests and strip searches, following the reaesrch procedure of drawing research questions, data cleaning and preparation, descriptive analysis, and statistical analysis.

According to the result from One-Way ANOVA and T-tests, different sex and perceived race would likely have different mean strip search numbers, mean arrest numbers, and mean strip search rate. The Tukey's HSD and interaction plots suggested that there are some significant differences between males and females in terms of their mean strip search numbers, mean arrest numbers, and mean strip search rate. Additionally, the Tukey's HSD has revealed that while the total amount of searches and arrests is high in racial groups Black and White, "Indigenous" having the highest value of strip search rate. The difference of mean strip search rates between "Indigenous" and "Latino", and "Indigenous and Middle-Eastern" is significant as well. Moreover, The two-way ANOVA suggests the interaction term sex: race is significant with respect to the mean number of strip searches and arrests, but it is not significant with respect to mean strip search rate, this can also be proved by the interaction plot Figure 5.2.

Reference

Toronto Police Service. (2021). Chapter 1 -Arrest & Release.
    https://www.tps.ca/media/filer_public/9b/d2/9bd29690-bfcf-42df-9480-7ba0d0ec6068/08
    da6409-baec-4b98-afe5-458c1e8d01b4.pdf

Arrests and strip searches (RBDC-arr-TBL-001). Toronto Police Service Public Safety Data
    Portal. (n.d.). Retrieved February 21, 2023, from
    https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tb
    l-001/about

Gills, W. (2019). Toronto police conduct "far too many" strip searches, says damning report from
    Ontario's police watchdog. Thestar.com.
    https://www.thestar.com/news/gta/2019/03/21/toronto-police-conduct-far-too-many-strip-
    searches-says-damning-report-from-ontarios-police-complaints-director.html

Ontario Human Rights Commission. (2018). Interim report on the inquiry into racial profiling
    and racial discrimination of Black persons by the Toronto Police Service.
    https://www.ohrc.on.ca/sites/default/files/TPS%20Inquiry_Interim%20Report%20EN%2
    0FINAL%20DESIGNED%20for%20remed_3_0.pdf#overlay-context=en/news_centre/o
    hrc-interim-report-toronto-police-service-inquiry-shows-disturbing-results

Gill, W. (2020, November 4). "Clearly, we were doing it wrong": Toronto police are doing far
    fewer strip searches under strict new rules, Interim chief says. Thestar.com.
    https://www.thestar.com/news/gta/2020/11/23/clearly-we-were-doing-it-wrong-toronto-p
    olice-doing-far-fewer-strip-searches-under-new-rules-interim-chief-says.html