# Study on Demographic Attributes Affecting Arrest Action and Strip Search Reasons

By

Group 27

Jiazhou Bi, Yubing Chi, Bohao Cui

A mid-term paper submitted in conformity with the requirements

for the course of INF2178H

Faculty of Information

University of Toronto

**1.     Introduction**

1.1     Overview and Literature Review

Police arrests and strip searches are contentious social topics because sometimes the police need to enforce such actions on suspects in order to maintain social order and protect other lawful citizens. However, on the other side, these actions can be seen as intrusive, degrading, and violating the individual's human rights, especially when they happen without reasonable justification. A previous study has recognized that discretion plays a significant role in police arrest decision-making (Walker, 1993). On top of that, much research was systematically conducted to identify many potential factors affecting these decision-making processes. These factors include but are not limited to the suspects' demeanor, the seriousness of the offense, victims' requests, race and sex of suspect and victim, and the occurrence neighborhood (Black, 1971; Brooks, 1986; Klinger, 1997; Lundman, 1994; Mastrofski, Snipes, & Parks, 2000; Mastrofski, Snipes, & Supina, 1996; Smith, 1986; Smith & Visher, 1981; Visher, 1983; Worden, 1989). Moreover, strip searches are considered to be more privacy-violating and dignity-hurting. Even though strip searches can sometimes yield critical evidence leading to criminal charges, Grewcock and Sentas (2019) argued that there are multiple disadvantages related to strip searches, and we should rethink our decisions before proceeding. Furthermore, Liptak and Souter (2009) recorded a court case in which the judge ruled that children's rights were violated by strip searches in that case, which led to future debates on strip searches. Therefore, it would be meaningful to see if different factors are associated with the arrest reactions, and with reasons justifying the strip searches.

1.2     Research Objective and Questions

The main objective of this study is to explore the relationships between demographic attributes, such as perceived race, and sex of the arrestees, and their displayed actions when they are arrested, as well as the reasons for them being strip-searched if they were strip-searched. As a result, we have formulated two research questions based on the knowledge gained in the initial analysis of the dataset (see the Descriptive statistics and T-test section below for our initial exploration of the data):

- RQ1: How do the perceived sex, and race of the arrestees, affect the number of negative actions conducted by the arrestees? That is, are arrestees from a particular sex, or race group, more likely to display negative actions when they are arrested?

- RQ2: How do the perceived sex, and race of the arrestees, affect their number of reasons for being strip searched?

We believe these research questions can show the relationships between different variables in the arrest dataset and identify any potential biases or discriminatory practices that could impact the fairness and impartiality of law enforcement.

**2.     Exploratory Data Analysis**

2.1     Univariate analyses:

First, we conducted some univariate analyses to get the descriptive statistics of our variables. Our first variable is the Perceived Race of the arrestee. In total, we have 52,650 male arrestees and 12,617 female arrestees from 2020 to 2021 as shown in the table below. We have removed sex labeled as "*unknown*", and the rationale for this operation is further explained in the Method section later in this paper. We have noticed that there are much more males arrested than females from 2020 to 2021, so we think there might be a gender difference in how individual suspects/arrestees are treated. Based on this difference, as well as our literature review regarding sex differences within the criminal justice system, we decided to choose sex as one of the independent variables for our study.

| Sex | count |
|-----|-------|
| M | 52650 |
| F | 12617 |

We have plotted this piece of information into a bar chart to show the sex difference in the number of arrestees more directly:

Arrestee Number based on Sex
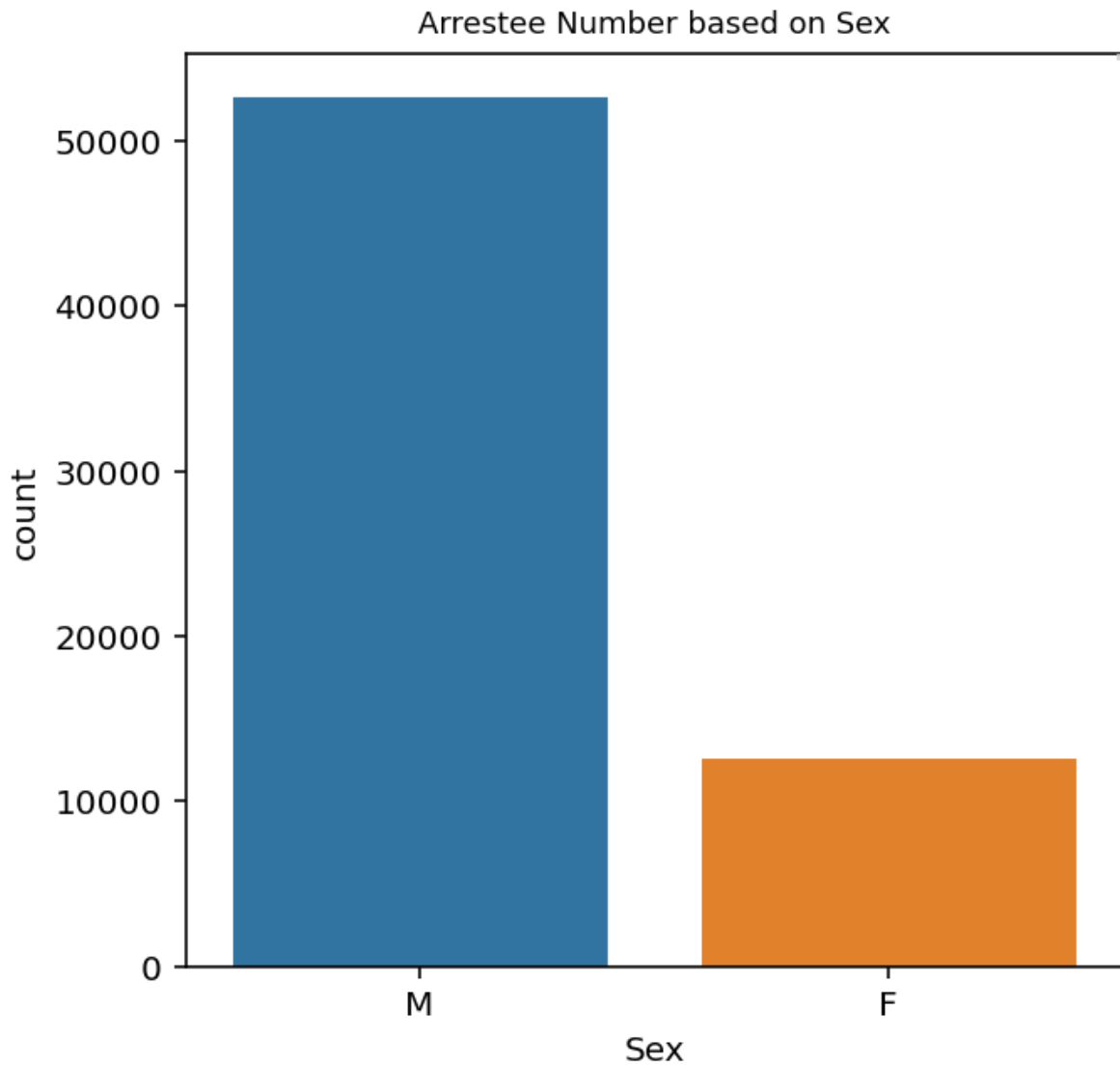
Our second independent variable is the perceived race of the arrestee. In this dataset, there are 8 different values for this attribute. They are *Black, East/Southeast Asian, Indigenous, Latino, Middle-Eastern, South Asian, Unknown or Legacy, or White*. From 2020 to 2021, there were, from the most to the least in number, 27,718 white arrestees, 17,526 black arrestees, 5,053 unknown or legacy arrestees, 4415 East/Southeast Asian arrestees, 3,616 South Asian arrestees, 3,237 Middle-Eastern arrestees, 1,934 Indigenous arrestees, and 1,767 Latino arrestees. The table of the number of arrestees grouped by perceived race is shown below.

| Perceived_Race | count |
| --- | --- |
| White | 27718 |
| Black | 17526 |
| Unknown or Legacy | 5053 |
| East/Southeast Asian | 4415 |
| South Asian | 3613 |
| Middle-Eastern | 3237 |
| Indigenous | 1934 |
| Latino | 1767 |

We have also plotted this information into a bar chart. From the graph, we have noticed major differences between arrestees with different perceived races. Combining this piece of knowledge with the information we learned from our literature review, we chose perceived race as our second independent variable.

Arrestee Number based on Perceived Race
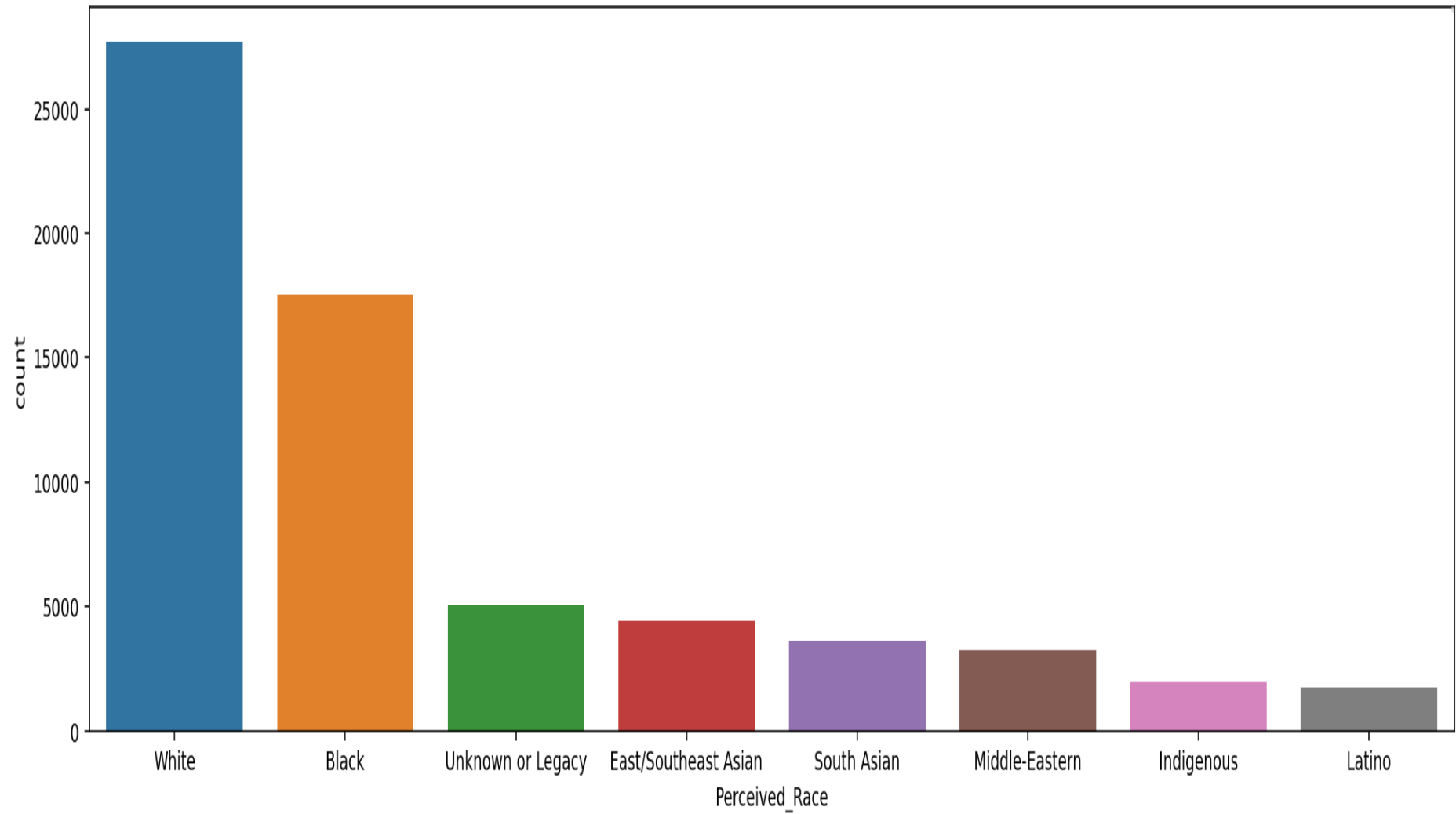
Our first dependent variable is aggregated negative actions during the arrest. We formulated this new attribute by summing all the arrest behaviors other than being cooperative. The details of the formation of this variable are explained in the Method-Variable section. The mean of this variable is 0.153247, and the median is 0. The central tendency of aggregated negative actions is shown below. As suggested, most arrestees displayed no negative behaviors during the arrest. The maximum amount of negative displayed during the arrest is 4.

```
count       65267.000000
mean            0.153247
std             0.534204
min             0.000000
25%             0.000000
50%             0.000000
75%             0.000000
max             4.000000
```

We have also drawn a box plot of this variable. But because 90.58% (59,121 out of 65,267 arrestees) displayed no negative behaviors during the arrest, the box shrank into a line concentrated on 0 as shown below, and arrestees with any negative behaviors are considered outliers.
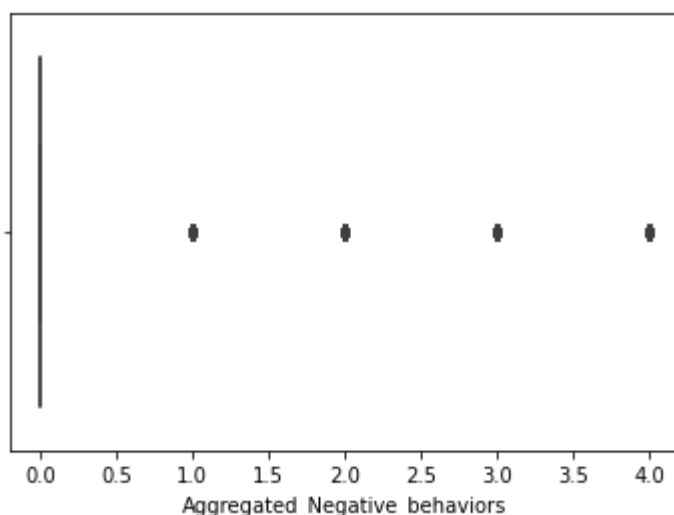
Our second dependent variable is aggregated reason for strip searches. We formulated this attribute by summing all the strip search reasons. The details of the formation of this variable are explained in the Method-Variable section. There were a total of 7,801 arrestees strip-searched. The mean aggregated search reason is 1.98, and the standard deviation is 1.21. The range of this variable is [0,4], and the details of this variable are shown below.

```
count        7801.000000
mean            1.981156
std             1.209695
min             0.000000
25%             1.000000
50%             2.000000
75%             3.000000
max             4.000000
```

We also drew a box plot for this variable, and the result is centered on 2 search reasons, and most arrestees being searched fall between 1 to 3 search reasons.
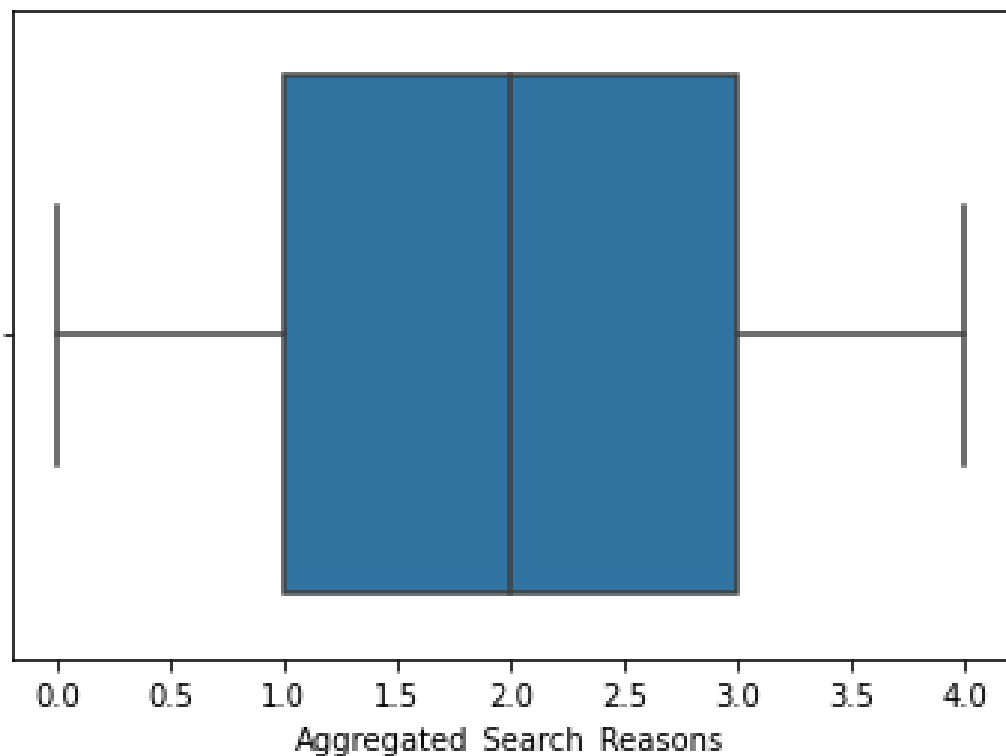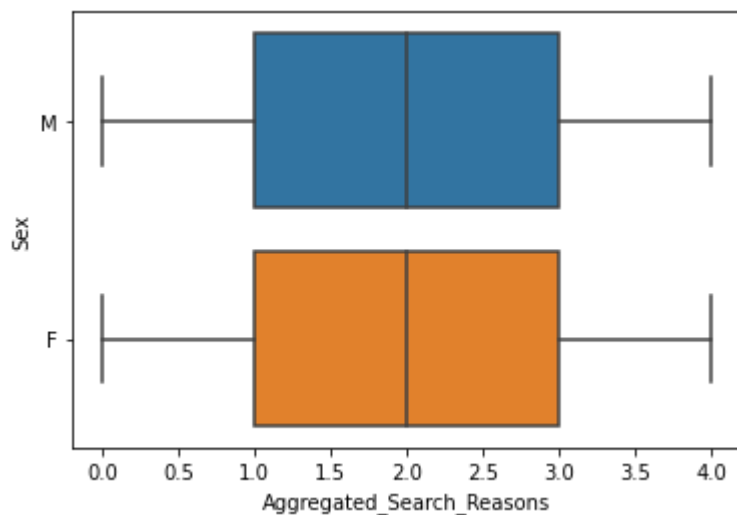
## 2.2      Bivariate Analyses:

### 2.2.1 Paired Box-Plot

We would like to see if males and females were considered to have different amounts of reasons for being strip-searched. The mean for male aggregated strip-search reasons is 2.0044, and for females is 1.8628. We have drawn a box plot for their aggregated strip-search reasons, respectively.



They look identical on the graph, but their means are different. This may be due to the values for this attribute being discrete, and a 0.15 mean difference cannot be reflected on a box plot. As a result, we have conducted a t-test to test if there are any mean differences between male and female aggregated strip-search reasons.

### 2.2.2   Line Graph

We have drawn two line graphs to see if our independent variables have effects on our dependent variable for each research question. The graphs are attached below. It looks like both sex and perceived race have a significant impact on aggregated negative behaviors and aggregated search reasons, with significant interaction from the appearance of the graphs. We will conduct further analysis to confirm these effects.

Aggregated Negative Behaviors by Sex and Perceived Race

Aggregated Search Reasons by Sex and Perceived Race
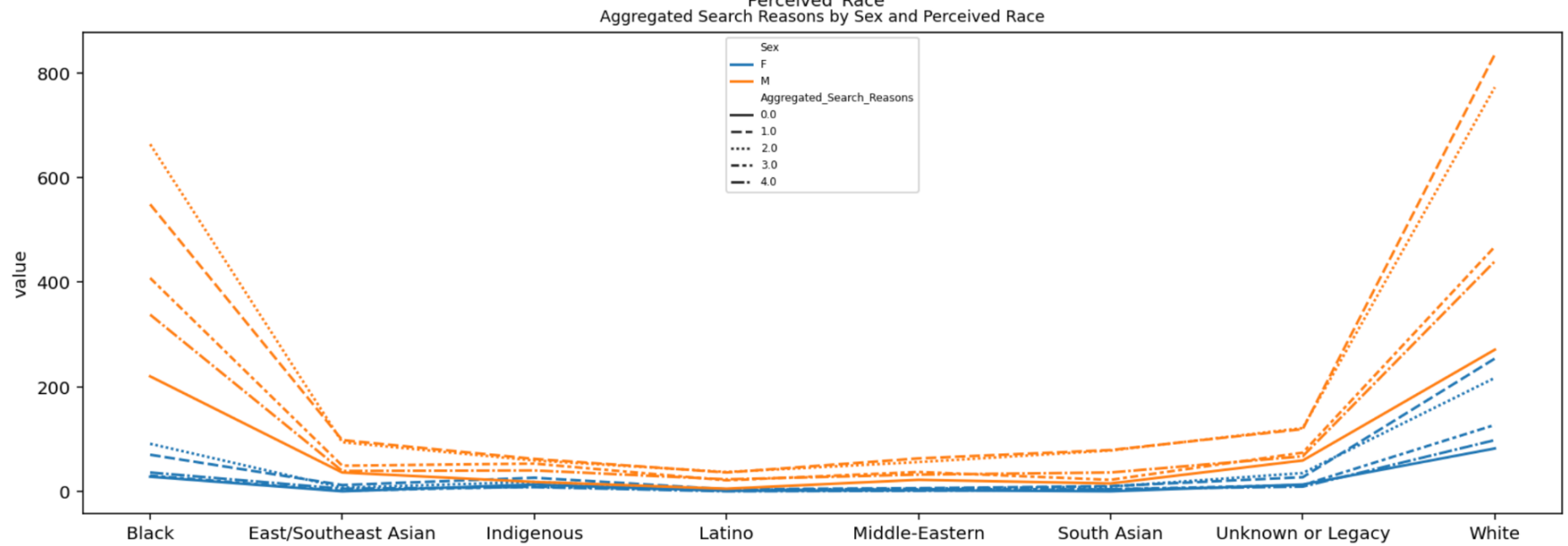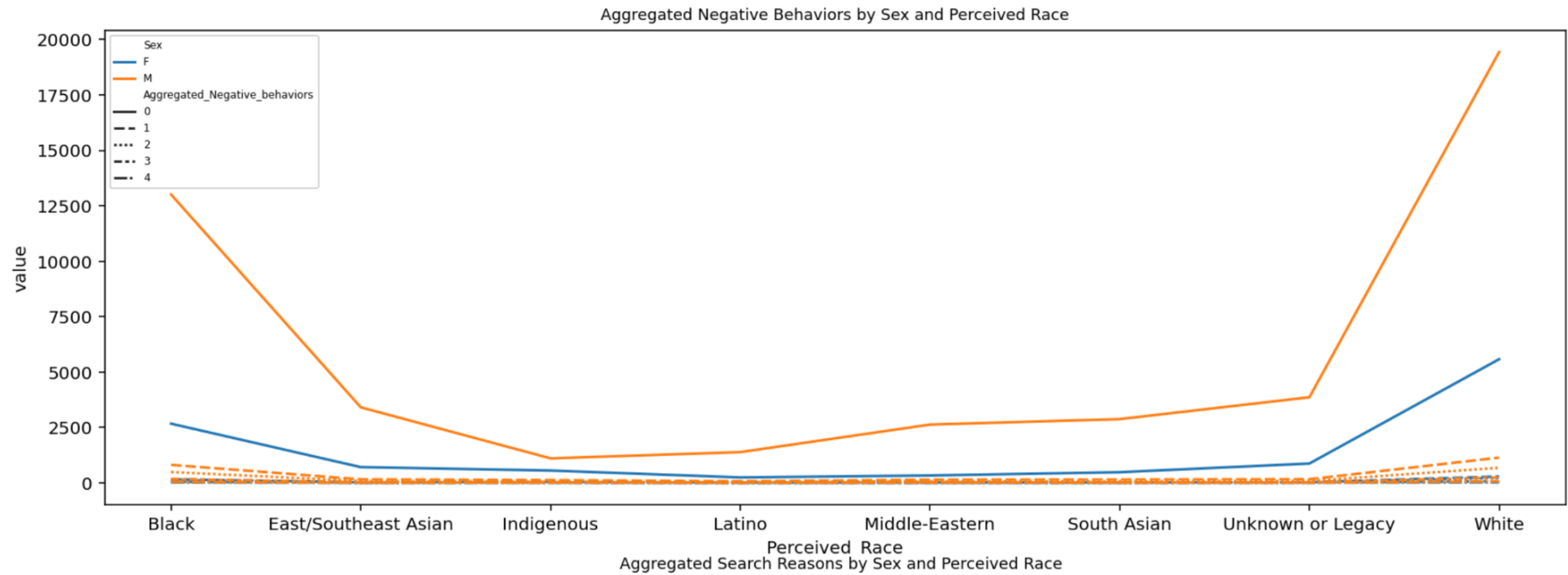
2.3     T-test:

Due to sex differences in the criminal justice system mentioned above, we would like to conduct a t-test to see if there exist any differences in reasons for strip searches between male and female arrestees if they were strip-searched. Therefore, our research hypotheses are

- Ho: Males and females are perceived to have the same amount of reasons to be strip searched.
- Ha: Males are more likely to be perceived to have more reasons to be strip searched than females.

First, we used the following code to run the t-test for the mean difference of aggregated strip search reasons between male and female arrestees

```
t_test_sex_search=Strip_S_table[['Sex','Aggregated_Search_Reasons']]
t_test_sex_search.head()
```

```
group1 = t_test_sex_search[t_test_sex_search['Sex']=='F']
group2 = t_test_sex_search[t_test_sex_search['Sex']=='M']
ttest_ind(group1['Aggregated_Search_Reasons'], group2['Aggregated_Search_Reasons'])
```

The result we got is the following

```
Ttest indResult(statistic=-3.836623571021187, pvalue=0.00012572162046602353)
```

It shows that at a 5% level of significance, our p-value is 0.0125%. We have enough evidence to reject the null hypothesis. Male arrestees are perceived to have more reasons to be strip-searched than their female counterparts. Therefore, we should dig deeper to see if adding another independent variable will affect its impact on the dependent variable.

## 3.     Method

3.1     Dataset Description

In our project, we will use a dataset that shows arrest information such as demographic attributes of arrestees (race, sex, age at arrest), reasons for arrest, time of arrest, strip searches, actions during the arrest, and so on. The dataset is publicly available on the Toronto Police Service Public Safety Data Portal. This dataset is being updated consistently and it was last updated on Nov. 10th, 2022 as of Feb 17th, 2023. Currently contains information on 65,276 arrestees that were arrested from 2020 to 2021. The data was collected

to record different information related to all arrests and strip searches conducted by the Toronto Police Service. Arrests that took place outside of the city boundary were also recorded if they were completed by the Toronto Police Service. Some key demographic attributes of the arrestee include their age at arrest, sex, and race. Other notable information is the date of arrest, and arrest location, if the arrestees were booked at a police station within 24 hours of the arrest, and the reason for the arrest. Furthermore, the behaviours displayed by the arrestee during the occurrence, and the reasons for the strip search. Different actions are displayed during the arrest, and reasons for strip searches were coded in binary format. Other variables were coded in string and/or datetime format that fits their respective type of information. We have attached all the attributes in the appendix.

### 3.2    Data Cleaning

3.2.1    *Data error:* "Due to issues with the booking template, there may be some records where a person was strip-searched, but the data does not indicate a booking (i.e., value = 0); in those cases, the user should presume a booking took place. (Toronto Police Service)" The above paragraph comes from the explanation of Booked on the official website of the Toronto Police. If the suspect is stripped for inspection, then Booked should be 1. However, there are some cases in the data where the stripped inspection is true, but the value of Booked is 0. We used the following code to correct this error:

```
ass table.loc[ass table["StripSearch"] == 1, "Booked"] = "1"
```

3.2.2    *Inconsistent format:* From the picture below, we can see that the police did not use a uniform format when recording the age information of the arrested suspects. We can see that there are two terms that indicate that the suspect is 17 years old or younger, and we will uniformly name them "Youth (aged 17 and younger)".

| Youth (aged 17 and younger) | Youth (aged 17 years and under) |
| --- | --- |
| 184.0 | NaN |
| 11.0 | NaN |
| 24.0 | NaN |
| 1.0 | NaN |

The same problem applies to records of suspects aged sixty-five and older. We used the same method for modification.

| Aged 65 and older | Aged 65 years and older |
| --- | --- |
| 108.0 | NaN |
| 516.0 | NaN |
| NaN | NaN |
| NaN | 102.0 |
| NaN | 596.0 |
| NaN | NaN |

We used the following codes to modify these inconsistencies:

```
ass_table = ass_table.replace('Aged 65 years and older','Aged 65 and older')
ass_table = ass_table.replace('Aged 17 years and under','Aged 17 years and younger')
```

3.2.3    *Empty value:* The police arrest data uses 1 and 0 to indicate whether the suspect has committed a certain behavior, 1 means yes, and 0 means no. But in the data record, only 1 is recorded, and some of the behaviors that have not been performed are filled with 0, and some have nothing. The data with empty values looked like the following:

| Break and Enter | Crimes against Children | Drug Related | ... | Police Category – Incident | Robbery & Theft | Robbery/Theft | Sexual Related Crime | Sexual Related Crimes & Crimes Against Children | Vehicle Related | Vehicle Related (inc. Impaired) | War |
|---|---|---|---|---|---|---|---|---|---|---|---|
| NaN | NaN | 33.0 | ... | NaN | 255.0 | NaN | NaN | 7.0 | NaN | 47.0 | |
| NaN | NaN | 28.0 | ... | NaN | 28.0 | NaN | NaN | 5.0 | NaN | 3.0 | |
| NaN | NaN | 7.0 | ... | NaN | 83.0 | NaN | NaN | 1.0 | NaN | 17.0 | |
| NaN | NaN | 5.0 | ... | NaN | 1.0 | NaN | NaN | NaN | NaN | 1.0 | |
| NaN | NaN | 10.0 | ... | NaN | 30.0 | NaN | NaN | NaN | NaN | 3.0 | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| 2.0 | NaN | 13.0 | ... | NaN | NaN | 2.0 | NaN | NaN | NaN | NaN | |
| 448.0 | 26.0 | 286.0 | ... | 203.0 | NaN | 1206.0 | 231.0 | NaN | 200.0 | NaN | |
| 19.0 | NaN | 54.0 | ... | 2.0 | NaN | 24.0 | 2.0 | NaN | NaN | NaN | |
| NaN | NaN | NaN | ... | NaN | NaN | 1.0 | NaN | NaN | NaN | NaN | |
| NaN | NaN | NaN | ... | NaN | NaN | NaN | NaN | NaN | NaN | NaN | |

For the convenience of analysis and statistics, we uniformly replace the null values with 0 with the following codes:

```
Strip_Search_Sex=Strip_Search_Sex.fillna(0)
```

Now all the NaN are replaced with 0, making our analysis possible.

3.2.4  *Data removal:* The dataset contained three values for sex. They are male, female, and unknown. We have removed unknowns since there are only 9 occurrences (compared to 52,650 males and 12,617 females), which accounts for roughly only 0.014% of the subjects. Removing this additional level will not only make the statistical analysis more valid (too small sample size) but also make the interpretation more straightforward. Therefore, we have removed the unknowns using the following code:

```
arrest_f_table = ass_table.loc[ass_table['Sex'] == "F"]
arrest_m_table = ass_table.loc[ass_table['Sex'] == "M"]
```

The code above only includes subjects with the attribute "sex" equal to either "F" (female) or "M" (male), leaving out the subjects with unknown sex.

3.3     Variables

As mentioned before, this database was initially collected and created by the Toronto Police Service. The spreadsheet contains the records of 65,276 occurrences of arrests at that time. We have manipulated the data to fit our research question. In our study, the main variables are the following:

*perceived sex*: the perceived biological sex of the arrestee

*race*: the perceived race of the arrestee

*strip search:* if the arrestee is strip searched

*actions at arrest - concealed items:* if the arrestee had concealed items during the arrest

*actions at arrest - combative, violent or spitter/biter:* if the arrestee had combative, violent, spitting, or biting behaviors during the arrest

*actions at arrest - resisted, defensive or escape risk:* if the arrestee showed resistant, defensive, or escaping signs during the arrest

*actions at arrest - mental instability or possibly suicidal:* if the arrestee displayed mental instability or suicidal tendency during the arrest

*actions at arrest - assaulted officer:* if the arrestee assaulted or attempted to assault the officer during the arrest

*aggregated negative behaviors:* the total counts of any negative behaviors during the arrest, including *concealed items, combative, violent or spitter/biter, resisted, defensive or escape risk, mental instability or possibly suicidal, assaulted officer*

*search reason - cause injury:* the arrestee was strip searched because they caused injury

*search reason - assist escape:* the arrestee was strip searched because they were trying to escape

*search reason - possess weapons:* the arrestee was strip searched because they were believed to possess weapons

*search reason - possess evidence:* the arrestee was strip-searched because they were believed to possess evidence

*aggregated search reasons:* the total counts of any strip search reasons, including *cause injury, assist escape, possess weapons, possess evidence*

3.4     Measurement

Variables and Value Categories

| Variable | Categories |
|---|---|
| **Independent Variables** | |
| Perceived Race (perceived_race, nominal) | The perceived race of the arrestee (Black, East/Southeast Asian, Indigenous, Latino, Middle-Eastern, South Asian, Unknown or Legacy, White). |
| Sex (sex, nominal) | The biological sex of the arrestee (Male, Female, Unknown [removed from analysis]). |
| **Dependent Variables** | |
| Aggregated Negative Behaviors (aggregated_negative_behaviors, ratio) | The sum of all the behaviors during arrest other than being cooperative (ranging from 0 to 5). |
| Aggregated Strip Search Reasons (aggregated_search_reasons, ratio) | The sum of all the strip search reasons (ranging from 0 to 4). |

*Source*: Arrests and Strip Searches (RBDC-ARR-TBL-001) by Toronto Police Service

3.4.1   Perceived Race

The first independent variable is the perceived race of the arrestee (*perceived_race*, nominal). This one describes the perceived race of the arrestee, which is a demographical

attribute of the arrestee commonly used to identify an individual. It can take any value from the following: *Black, East/Southeast Asian, Indigenous, Latino, Middle-Eastern, South Asian, Unknown or Legacy, or White*.

### 3.4.2 Sex

The second independent variable is sex (*sex*, nominal). This one describes the biological sex of the arrestee. In many social settings, it can take many different values. In our dataset, the set of this variable is *Male, Female, or Unknown*. We have excluded *Unknown* from our dataset because of the reason mentioned in the Data Cleaning section. We have to disclaim that this method of coding an individual's sex is used by the Toronto Police Service. We do not support or oppose this method of coding an individual's sex.

### 3.4.3 Aggregated Negative Behaviors

The first dependent variable is the total count of each arrestee's negative behaviors (if any) displayed during the arrest (*aggregated_negative_behaviors*, ratio). We have summed all the behaviors during the arrest other than being cooperative to formulate this variable (i.e., *aggregated_negative_behaviors = actions at arrest - concealed items + actions at arrest - combative, violent or spitter/biter + actions at arrest - resisted, defensive or escape risk + actions at arrest - mental instability or possibly suicidal + actions at arrest - assaulted officer*). In theory, the range of this variable is from 0 to 5. In our dataset, the range is 0 to 4.

### 3.4.4 Aggregated Search Reasons

The second dependent variable is the total count of each arrestee's reasons for being strip-searched if they were strip-searched (*aggregated_search_reasons*, ratio). We have summed all the strip search reasons to formulate this variable (i.e., *aggregated_search_reasons = search reason - cause injury + search reason - assist escape + search reason - possess weapons + search reason - possess evidence)*. The range of this variable is from 0 to 4.

### 3.5 ANOVA Tests

We completed two two-way ANOVA analyses for our research questions. In the first ANOVA analysis, the two independent variables are the perceived races of the arrestees and the sexes of the arrestees. The dependent variable is their aggregated negative behaviors. In the second ANOVA analysis, the two independent variables are the same as the first ANOVA

test, namely the perceived races of the arrestees and the sexes of the arrestees. The dependent variable in this analysis is the aggregated strip-search reasons. We have also conducted post-hoc tests to find out which combination(s) of the independent variable is significantly different from the others, and checked the assumptions of ANOVA tests.

## 4.    Results

4.1    The Effect of Perceived Race and Sex on Aggregated Negative Behaviors

We have conducted a two-way ANOVA on the effect of perceived race and sex on aggregated negative behaviors.

4.1.1    Assumption Checks

4.1.1.1 Independent Observations

Because some of the arrestees may act in groups, we cannot conclude that the observations are independent. In our study, we have assumed that the observations are independent so that we can proceed with our two-way ANOVA tests.

4.1.1.2 Homogeneity of Variance

We have conducted Levene's test to check if the assumption of equal variance is met. The results of Levene's test for perceived race and sex (in order) are pasted below

```
Levene's test: statistic=27.056, p-value=0.000
```

```
Levene's test: statistic=3.008, p-value=0.083
```

Because the p-value for the first test is smaller than 0.001, we reject the null hypothesis of Levene's test. The assumption of equal variance is violated. A common alternative for this violation is to apply Welch's ANOVA because it is not sensitive to unequal variances. However, it is not suitable for two-way ANOVAs. For the sake of this project, we will proceed with a regular two-way ANOVA analysis.

4.1.1.3 Normality

We also did a Shapiro-Wilk normality test to check the normality assumption for ANOVA. The result is pasted below

`Assumption of normality is violated.`

However, ANOVA is considered to be robust against the normality assumption if the sample size is large enough. In our case, it is fairly safe to conclude that this violation of normality does not have a significant impact on our analysis because our sample size is enough (more than 60,000 observations).

### 4.1.2 ANOVA Results

As mentioned above, we have conducted an interaction model studying the effects of the perceived race and sex of the arrestee and their aggregated negative behaviors during the arrest. The result looks like the following

```
                            df        sum_sq    mean_sq          F  \
C(Sex)                     1.0      0.859223   0.859223   3.019575
C(Perceived_Race)          7.0     53.332993   7.618999  26.775522
C(Sex):C(Perceived_Race)   7.0      4.840069   0.691438   2.429929
Residual               65247.0  18566.092977   0.284551        NaN

                              PR(>F)
C(Sex)                  8.226976e-02
C(Perceived_Race)       5.959743e-37
C(Sex):C(Perceived_Race) 1.734509e-02
Residual                         NaN
```

According to the results, at a 5% significance level, we can conclude that sex (p-value = 0.0821) does not have a significant effect on the arrestee's aggregated negative behaviors. The perceived race of the arrestee (p-value < 0.001), however, does have a significant effect on the arrestee's aggregated negative behaviors. Its effect depends on the sex of the arrestee, as suggested by the significant interaction (p-value = 0.0173) between the two variables.

### 4.1.3 Post-Hoc Test

Because we found that perceived race and the interaction between perceived race and sex have a significant effect on aggregated negative behaviors, we have done a Tukey's HSD test to see which group of perceived race is different from the other groups. We found that Black arrestees do differ significantly from all other arrestees other than White arrestees; East/Southeast Asian arrestees differ significantly from Latino, and White arrestees; Indigenous arrestees differ significantly from all other arrestees other than the Latinos;

```
          Black East/Southeast Asian  -0.0688  0.001 -0.0967   -0.041   True
          Black            Indigenous   0.0565  0.001  0.0169   0.0962   True
          Black                Latino  -0.0495 0.0062 -0.0908  -0.0082   True
          Black        Middle-Eastern  -0.0331 0.0323 -0.0648  -0.0015   True
          Black           South Asian  -0.0661  0.001 -0.0963  -0.0359   True
          Black     Unknown or Legacy  -0.0735  0.001 -0.0999   -0.047   True
          Black                 White  -0.0142 0.1277 -0.0302   0.0018  False
          Black                   nan  -0.1746    0.9  -1.002   0.6529  False
East/Southeast Asian           Indigenous   0.1254  0.001  0.0802   0.1705   True
East/Southeast Asian               Latino   0.0193    0.9 -0.0273   0.0659  False
East/Southeast Asian       Middle-Eastern   0.0357   0.09 -0.0026    0.074  False
East/Southeast Asian          South Asian   0.0027    0.9 -0.0344   0.0398  False
East/Southeast Asian    Unknown or Legacy  -0.0046    0.9 -0.0387   0.0294  False
East/Southeast Asian                White   0.0546  0.001  0.0278   0.0814   True
East/Southeast Asian                  nan  -0.1058    0.9 -0.9335    0.722  False
     Indigenous                Latino  -0.1061  0.001 -0.1605  -0.0516   True
     Indigenous        Middle-Eastern  -0.0896  0.001 -0.1372  -0.0421   True
     Indigenous           South Asian  -0.1226  0.001 -0.1693   -0.076   True
     Indigenous     Unknown or Legacy    -0.13  0.001 -0.1742  -0.0858   True
     Indigenous                 White  -0.0707  0.001 -0.1096  -0.0318   True
     Indigenous                   nan  -0.2311    0.9 -1.0593   0.5971  False
         Latino        Middle-Eastern   0.0164    0.9 -0.0325   0.0654  False
         Latino           South Asian  -0.0166    0.9 -0.0646   0.0315  False
         Latino     Unknown or Legacy  -0.0239 0.7652 -0.0697   0.0218  False
         Latino                 White   0.0353 0.1478 -0.0053   0.0759  False
         Latino                   nan  -0.1251    0.9 -0.9534   0.7032  False
 Middle-Eastern           South Asian   -0.033 0.2057  -0.073   0.0071  False
 Middle-Eastern     Unknown or Legacy  -0.0404 0.0222 -0.0776  -0.0031   True
 Middle-Eastern                 White   0.0189 0.5925 -0.0118   0.0496  False
 Middle-Eastern                   nan  -0.1415    0.9 -0.9694   0.6864  False
    South Asian     Unknown or Legacy  -0.0074    0.9 -0.0434   0.0287  False
    South Asian                 White   0.0519  0.001  0.0226   0.0812   True
    South Asian                   nan  -0.1085    0.9 -0.9363   0.7193  False
Unknown or Legacy                 White   0.0593  0.001   0.034   0.0846   True
Unknown or Legacy                   nan  -0.1011    0.9 -0.9288   0.7266  False
          White                   nan  -0.1604    0.9 -0.9878    0.667  False
```

Middle-Eastern arrestees differ significantly from Unknown or Legacy arrestees; South Asian arrestees differ significantly from White arrestees; and lastly, Unknow or Legacy arrestees differ significantly from White arrestees. The detailed result is shown above.

4.2     The Effect of Perceived Race and Sex on Aggregated Search Reasons

We have conducted a two-way ANOVA on the effect of perceived race and sex on aggregated strip search reasons.

4.2.1   Assumption Checks

4.2.1.1 Independent Observations

Same as the previous research question, because some of the arrestees may act in groups, we cannot conclude that the observations are independent of each other. In our study,

we have assumed that the observations are independent so that we can proceed with our two-way ANOVA tests.

4.2.1.2 Homogeneity of Variance

We have conducted Levene's test to check if the assumption of equal variance is met. The results of Levene's test for perceived race and sex (in order) are pasted below

```
Levene's test: statistic=1.165, p-value=0.319
```

```
Levene's test: statistic=2.075, p-value=0.150
```

Because the p-values for the tests are greater than 0.05, we failed to reject the null hypothesis of Levene's test. The assumption of equal variance is not violated. It is safe to proceed with the ANOVA test.

4.2.1.3 Normality

We have also done a Shapiro-Wilk normality test to check the normality assumption for ANOVA. The result is pasted below

```
Assumption of normality is violated.
```

Same as previously stated, ANOVA is considered to be robust against the normality assumption if the sample size is large enough. In our case, it is fairly safe to conclude that this violation of normality does not have a significant impact on our analysis because our sample size is enough (more than 6,000 observations for this research question).

4.2.2    ANOVA Results

We have done a two-way ANOVA with interaction to see the effects of sex and perceived race on the aggregated search reasons, the result is shown below

|  | df | sum_sq | mean_sq | F | PR(>F) |
|---|---|---|---|---|---|
| C(Sex) | 1.0 | 21.455509 | 21.455509 | 14.699189 | 0.000127 |
| C(Perceived_Race) | 7.0 | 16.448488 | 2.349784 | 1.609839 | 0.127497 |
| C(Sex):C(Perceived_Race) | 7.0 | 13.457549 | 1.922507 | 1.317111 | 0.237430 |
| Residual | 7784.0 | 11361.830249 | 1.459639 | NaN | NaN |

According to the results, at a 5% significance level, we can conclude that sex (p-value = 0.000127) does have a significant effect on the arrestee's aggregated strip-search reasons. The perceived race of the arrestee (p-value = 0.1275), however, does not have a significant effect on the arrestee's aggregated strip-search reasons. Furthermore, there is no interaction between these two variables, as the interaction between them (p-value = 0.2374) is not statistically significant.

4.2.3 Post-Hoc Test

Because only sex is a significant variable and it has two groups, there is no need for a post-hoc test to see which group is different from the other. We have done a Tukey's HSD test anyways and the result is shown below.

```
Multiple Comparison of Means - Tukey HSD, FWER=0.05
===================================================
group1 group2 meandiff p-adj lower  upper reject
---------------------------------------------------
    F      M   0.1416 0.001 0.0693 0.214   True
---------------------------------------------------
```

## 5.    Conclusion

Through this study, we have addressed two main research questions. Firstly, we investigated the potential relationship between demographic attributes of arrestees (such as perceived race and sex) and any negative behaviors they exhibited at the time of the arrest. By conducting a Two-way ANOVA and graphical analysis, we were able to draw conclusions. We found that most arrestees did not display negative behaviors during the arrest, which is positive. However, among those who did exhibit negative behaviors, it appears that gender is not associated with negative behaviors, while some races(such as white and black) are coorelated with more negative behaviors. For the interpretation of this question, the explanatory variables are the gender and race of arrestees, and the outcome is the extent of negative behavior during an arrest. And we found that there was no significant interaction between gender and negative behavior during the arrest. However, race was identified as a critical factor in determining the extent of negative behavior during an arrest. It means that there were significant differences in the extent of negative behavior between people of different races.

Secondly, we examined the relationship between the demographic attributes (perceived race and sex) of the arrestees and the reasons behind the decision to perform strip searches. The results indicate that most arrestees are subjected to strip searches for two reasons. Furthermore, it appears that race is not associated with negative behaviors, while difference of gender(female and male) are coorelated with the reason for strip searches. For the interpretation of this question, the explanatory variables are gender and race, and the outcome is the reasons for strip searches among arrestees. We found that there were significant differences in the reasons for strip searches between men and women. And there were no significant differences in the reasons for strip searches between people of different races.

The validity of these conclusions is confirmed by checking several assumptions. Firstly, we assumed that the observations are independent to conduct our two-way ANOVA tests. Then, we performed Levene's test to verify if the assumption of equal variance is met, and a Shapiro-Wilk normality test to confirm the normality assumption for ANOVA.

## 6.     Discussion

From a data processing and collection perspective, we believe that it is not a wise approach to convert the various negative behaviors displayed by the arrestees and the reasons for their strip searches into qualitative data for analysis. Additionally, the data obtained in the real world during the study was not ideal, and we had to undertake data processing and compromises. We discovered that there were inconsistent formats and empty values present in the dataset, indicating that the data collection process may have been either inaccurate or incomplete. As such, we recommend that future data collection processes be made more comprehensive, with a view to enhancing the accuracy of the results. This may include implementing more robust and standardized data collection procedures, utilizing advanced technology to automate data collection, and ensuring that data is validated and checked for completeness and accuracy before analysis. Ultimately, taking these steps can help to ensure that the data collected is of the highest quality, which is essential for drawing accurate conclusions and making informed decisions.

From a broader societal perspective, these results highlight the importance of safeguarding equal treatment and protection for all individuals during law enforcement practices. Law enforcement agencies must avoid unfairly treating arrested persons based on

their race or gender, and instead, strive to uphold diverse and inclusive policies that better serve and represent the community as a whole. Furthermore, the results of this study suggest that there is a need for further exploration into the specific factors that contribute to gender's significant impact on the reasons for being strip-searched. By better understanding these factors, we can develop more targeted interventions and policies to address any disparities and ensure that all individuals are treated fairly and with dignity.

References

Black, D. (1971). The social organization of arrest. Stanford Law Review, 23, 1087.

Brooks, L. W. (1986). Determinants of police orientations and their impact on discretionary police behavior. Doctoral dissertation, College Park: University of Maryland.

Grewcock, M., & Sentas, V. (2019). Rethinking strip searches by NSW Police.

Klinger, D. A. (1997). Negotiating order in patrol work: An ecological theory of police response to deviance. Criminology, 35, 277-305.

Liptak, A., & Souter, J. D. H. (2009). Supreme Court says child's rights violated by strip search. *New York Times*.

Lundman, R. J. (1994). Demeanor or crime? The Midwest city police-citizen encounters study. Criminology, 32, 631-656.

Mastrofski, S. D., Snipes, J. B., & Parks, R. B. (2000). The helping hand of the law: Police control of citizens on request. Criminology, 38, 307-342.

Mastrofski, S. D., Snipes, J. B., & Supina, A. E. (1996). Compliance on demand: The public's response to specific police requests. Journal of Research in Crime and Delinquency, 33, 269- 305.

Smith, D. A. (1986). The neighborhood context of police behavior. In A. Reiss & M. Tonry (Eds.), Communities and crime (pp. 313-341). Chicago: University of Chicago Press.

Smith, D. A., & Visher, C. A. (1981). Street level justice: Situational determinants of police arrest decisions. Social Problems, 29, 167-177.

Visher, C. A. (1983). Gender, police arrest decisions, and notions of chivalry. *Criminology*, *21*(1), 5-28.

Walker, S. (1993). Taming the system: The control of discretion in criminal justice, 1950-1990. New York: Oxford University Press.

Worden, R. E. (1989). Situational and attitudinal explanations of police behavior: A theoretical reappraisal and empirical assessment. Law and Society Review, 23, 667-710.