Demographic Factors Impacting Formal Booking and Strip Search Based on PSDP

By

Group 62

Mengqi Gao (1004233938)

Caiyi/Chelsea Yang (1004497207)

For the course of INF2178: Experimental Design for Data Science

Faculty of Information

University of Toronto

Introduction

The dataset "Arrests and Strip Searches" published from Public Safety Data Portal on November 10, 2022, contains 65,276 records with 24 attributes. This dataset contains data regarding all arrests and strip searches, as well as indicators indicating whether a person was booked at a police station within 24 hours of the arrest event.

The dignity and rights of the individual being arrested need to be respected and the strip search needs to be conducted in the least intrusive manner possible. From a police perspective, whether belonging to different groups affects the likelihood that an arrestee will be booked or strip searched.

By examining the data related to demographic factors of individuals being arrested, this research seeks to provide valuable insights for police officers to precisely conduct strip searches and formal bookings, by determine the influence of demographic factors on booked or strip search, including Perceived Race, Sex and Age group(at arrest), in consideration of promoting fair and unbiased law enforcement practices and fostering greater public accountability.

Literature Review

The article "Strip-Searching of Women in Canada: Wrongs and Rights"(Michelle P. & Elizabeth S., 2016) examines the legal and ethical issues surrounding the strip-searching of women in Canada. They believe that strip searches are a serious violation of women's rights, especially when they are performed without proper reason or in an insulting, humiliating manner. In the case of R.v. Golden in 2001(R. V. Golden), the Supreme Court of Canada ruled that justice strip searches should not be carried out as a routine policy. The court also stipulates the legal standards that must be met for a legal strip search, along with 11 safeguards that police must adhere to when carrying out such searches. Although the law provides protection for women's rights, there are cases of non-compliance in implementation. Psychological damage can occur during the strip search. The authors suggest that, for women who have experienced sexual violence or abuse, strip searches may cause trauma and even exacerbate existing psychological problems. This article advocates that, under clear and reasonable circumstances, strip searches should be carried out in a respectful manner, and minimized the harm to the search as much as possible. At the same time, it is necessary to strengthen supervision and accountability during the body search process to ensure respect for women's deprivation rights.

The jurisprudence of strip searches(McNeilly, G., 2019) shows that when the police have no evidence that the person suspected of crime is a threat to themselves, others or the facility, or there is no reason to believe that the person suspected of crime has concealed items on the body, the court usually justice that the strip search is unreasonable. For some specific situations, such as people who have hidden drugs in their bodies or have a history of violent behavior, the court justifies that the strip search is reasonable.

Exploratory Data Analysis

1. Dataset

This research is based on a dataset of "Arrests_and_Strip_Searches_(RBDC-ARR-TBL-001).csv" which contains data collected from the Public Safety Data Portal of Toronto Police Service. This dataset has sufficient information related to all arrests and strip searches, including 65,276 records with 24 variables. The dataset is available through Toronto Police Service official website with the following link: https://data.torontopolice.on.ca/datasets/TorontoPS::arrests-and-strip-searches-rbdc-arr-tbl-001/explore. In this research we choose 5 variables for analysis, a detailed description of each variable is shown in the Table 1 below.

Table 1.

Data Description

Variables	**Unique**	**Type**
Perceived_Race	8	Text
Sex	3	Text
Age_group__at_arrest_	9	Text
StripSearch	2	Number
Booked	2	Number

2. Measurement

2.1 Independent Variables

In this research, "Perceived_Race, Sex, Age_group__at_arrest_" were chosen to analyze the effect on the person who was booked and was strip searched respectively. Which are all categorical variables, with eight, three and nine different categories(Appendix Table A).

2.2 Dependent Variables

According to the research question, the dependent variables were chosen as "Booked" and "StripSearch"(Appendix Table B), which indicate whether a person was booked at police station

within 24 hours and a person was strip searched. Determining the influence of variables on these indicators is the main focus of this research.

3. Data analysis

For this research, the data analysis was divided into three parts. In the first part, in order to reduce the impact of unknown factors on the data, we cleared all the null values and the information of people whose Sex is shown as unknown, 65,239 records was left and the description is shown in the Table 2 below.

Table 2. Descriptive Statistics

	Perceived_Race	Sex	Age_group__at_arrest_	StripSearch	Booked
Valid	65239	65239	65239	65239	65239
Missing	0	0	0	0	0
Mean				0.120	0.528
Std. Deviation				0.324	0.499
Minimum				0.000	0.000
Maximum				1.000	1.000

Note. Not all values are available for *Nominal Text* variables

The second part is plots analysis, the histogram of Booked and StripSearch(Figure 1), shows that people who have been booked are slightly more than those who are not and the count, the number of people being strip searched is much smaller than those not been strip searched . In addition, the count plots between each independent variable and dependent variable(Figure 2-7) are shown below, which compares the quantitative relationship of the situation that has been booked or strip searched between different perceived races, genders, and age groups.

Figure 1. Histogram of Booked and StripSearch

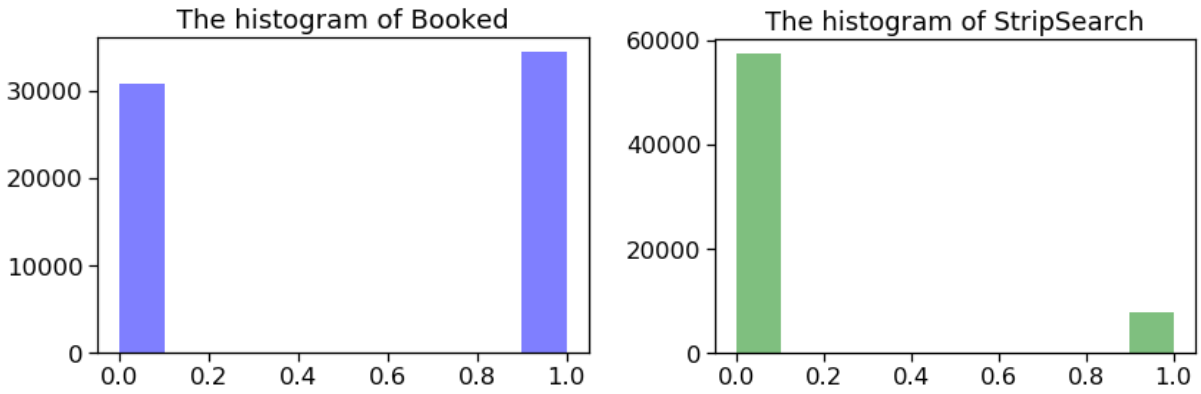


Figure 2. Countplot of Booked vs. Perceived_Race

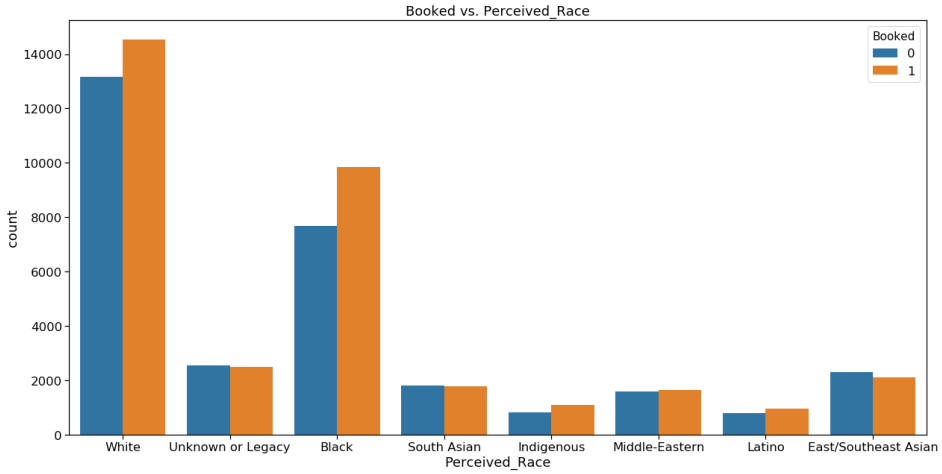


Figure 3. Countplot of Booked vs. Sex

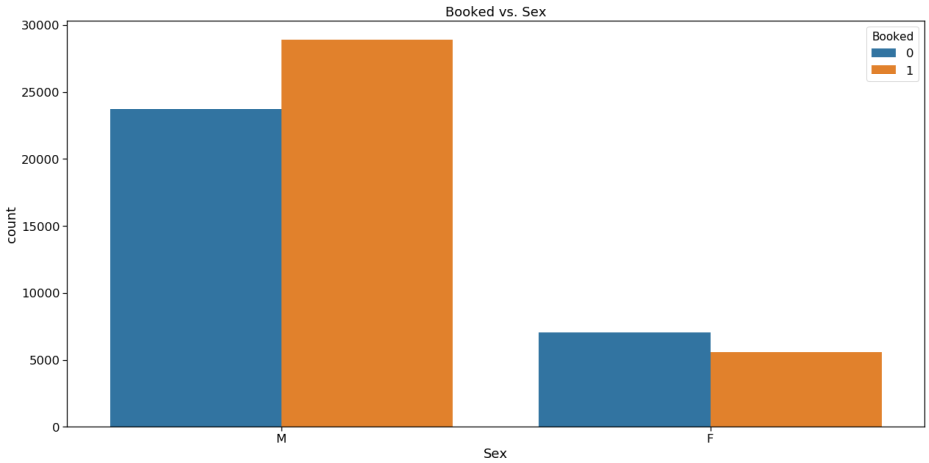


Figure 4. Countplot of Booked vs. Age_group__at_arrest_

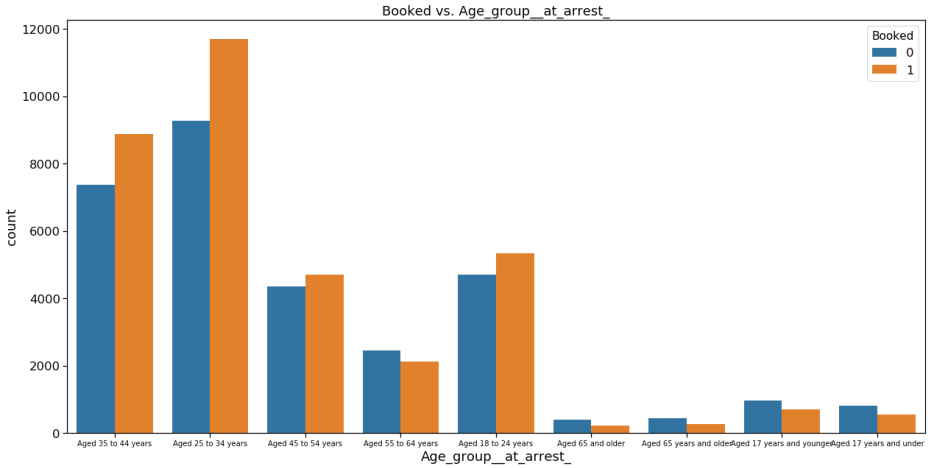


Figure 5. Countplot of StripSearch vs. Perceived_Race

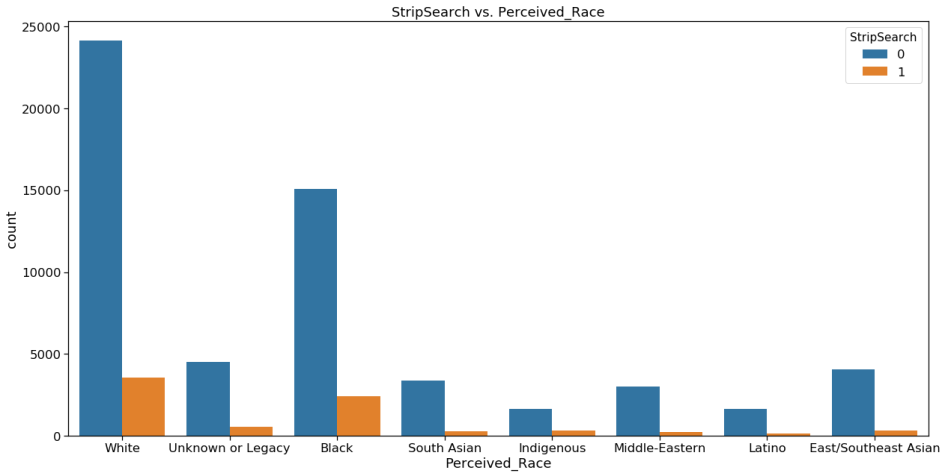


Figure 6. Countplot of StripSearch vs. Sex

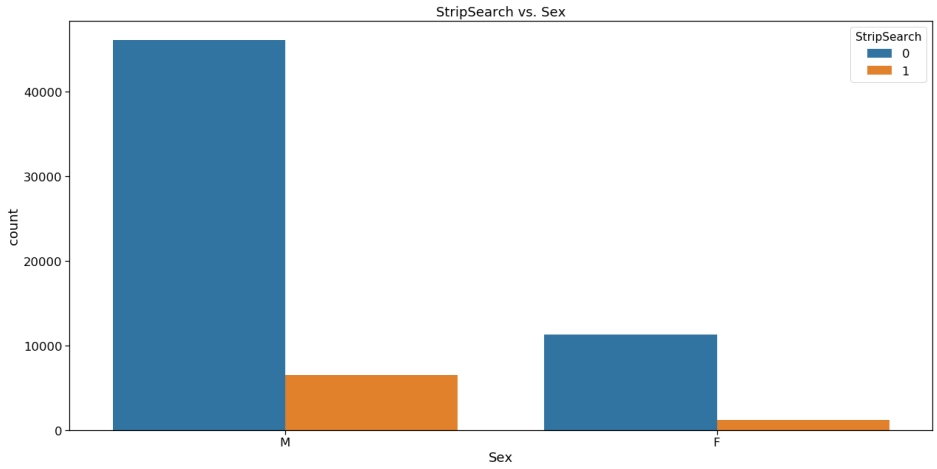
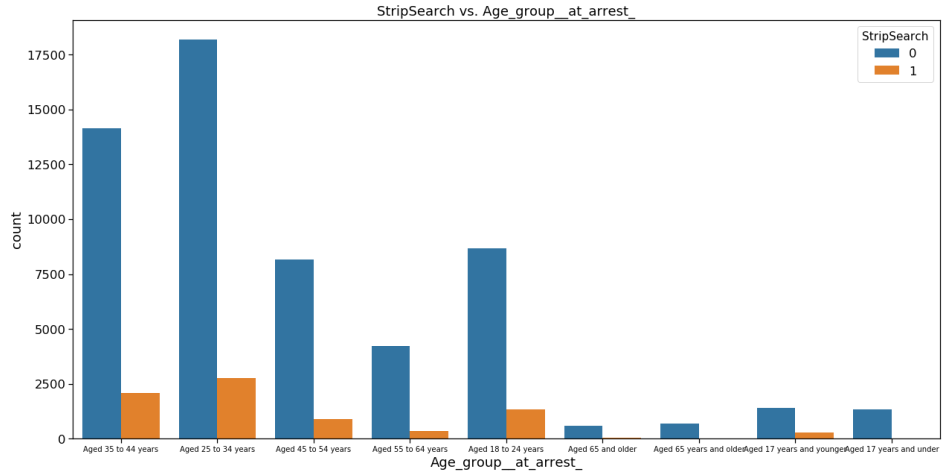
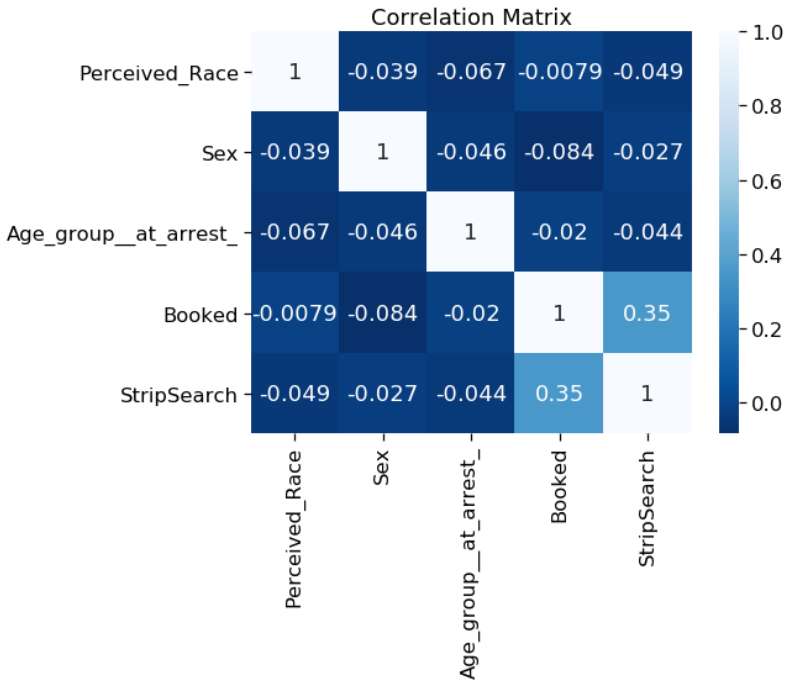


Figure 7. Countplot of StripSearch vs. Age_group__at_arrest_



For further analysis, we converted all variables of the object type into int, and merge the group 'Age 17 years and under'&'Age 17 years and younger' and 'Aged 65 and older'&'Aged 65 years and older', then made a correlation matrix(Figure 8) between these variables.

Figure 8. Correlation Matrix



The last part is Welch's T-test. Before running the Welch's T-test, we need to check the assumption of normality first. The test statistic and p-value of each group is shown in Table 3.

Table 3. Test of Normality (Shapiro-Wilk)

	W(Booked)	p(Booked)	W(StripSearch)	p(StripSearch)
Male	0.927	0.002	0.927	0.002
Female	0.936	0.006	0.036	0.005
White	0.929	0.296	0.946	0.496
Black	0.951	0.577	0.963	0.773
Unknown or Legacy	0.867	0.038	0.886	0.070
East/Southeast Asian	0.928	0.289	0.970	0.870
South Asian	0.926	0.270	0.959	0.700
Middle-Eastern	0.901	0.118	0.898	0.106
Indigenous	0.799	0.005	0.938	0.392
Latino	0.764	0.002	0.917	0.196
Age(<17)	0.959	0.652	0.798	0.003
Age(18-24)	0.938	0.326	0.935	0.288
Age(25-34)	0.908	0.108	0.950	0.487
Age(35-44)	0.911	0.122	0.978	0.947
Age(45-54)	0.923	0.186	0.952	0.536
Age(55-64)	0.927	0.221	0.941	0.365
Age(>65)	0.919	0.163	0.723	0.0003

Note. Significant results suggest a deviation from normality.

The groups that neither of the variables violates the assumption of normality can continue with the analysis, by comparing the p-value with 0.05. Groups grouped by perceived race with similar count value in the countplot and age-grouped groups at similar ages were selected for further examination.

We computed the mean of Booked and StripSearch for groups. The hypothesis being tested are following:

- **H0(Null Hypothesis):** The means between different groups are equal.
- **H1(Alternative Hypothesis):** The means between different groups are different.

Following are groups that with result pass the Welch's T-test(Table 4).

Table 4. Independent Samples T-Test

		t	p
Booked	Black vs. East/Southeast Asian	2.609	0.015
	Age(<17) vs. Age(18-24)	-3.323	0.003
StripSearch	Middle-Eastern vs. Indigenous	-3.277	0.005
	Indigenous vs. Latino	2.235	0.036

Note. Welch's t-test.

The results show that the mean of Booked of Black vs. East/Southeast Asian and Age(<17) vs. Age(18-24) is different, and the mean of StripSearch of Middle-Eastern vs. Indigenous and Indigenous vs. Latino is different. The p-value is less than the alpha determined as 0.05, which is statistically significant. Therefore, we can reject the null hypothesis that mean between groups are equal.

Method

1. Research Objective

Based upon the literature reviews and exploratory data analysis, we aim to study how sex, age group at arrest, perceived race and if any, interactions between them influence the chance of being subjected to a strip search, as well as a formal booking at the police station within 24 hours. A comprehensive set of detailed research questions are listed below to guide the direction of our entire project.

2. Research Questions

- Research Question 1: To what extent do demographic factors differ among individuals who are arrested? In this context, demographic factors are specifically sex, age group at arrest and perceived race. If there exists any influential demographic factors to the chance of being subjected to a strip search and a formal booking, which specific groups are significantly different from others?
- Research Question 2: Will different combinations and interactions of demographic factors affect the chance of being subjected to a strip search? Will different combinations of demographic factors affect the chance of being subjected to a formal booking?

3. Research Design and Methods

After investigating descriptive statistics and Welch's t-tests, we are determined to utilize one-way ANOVA tests and two-way ANOVA tests, as well as Tukey's HSD tests to further address our research questions. In the subsequent text, we managed to elaborate on how we used the applied methods to examine each research question.

- **Digesting Research Question 1:** Through Welch's t-test results and heatmap, we have identified specific groups in which there is a significant difference respectively in mean strip search score and mean booked score. We conducted four one-way ANOVA tests on perceived race and age group at arrest to determine if there were any significant differences in the means of three or more groups. One possible approach to identifying

which perceived race and age group at arrest have significantly different mean strip search score and mean booked score is to utilize one of the post-hoc tests, which is Tukey's HSD test. Conclusively, we could fully understand research question 1 after these methods.

- **Digesting Research Question 2:** To uncover whether different combinations and interactions of demographic factors affect the chance of being subjected to a strip search, we used two two-way ANOVA tests to determine (1) if there is a statistically significant interaction between the two independent variables (perceived race and age group at arrest) and the dependent variable (mean strip search score). (2) if there is a statistically significant interaction between the two independent variables (perceived race and age group at arrest) and the dependent variable (mean booked score). Again, we ran through Tukey's HSD test after each ANOVA test.

4. Preprocessing

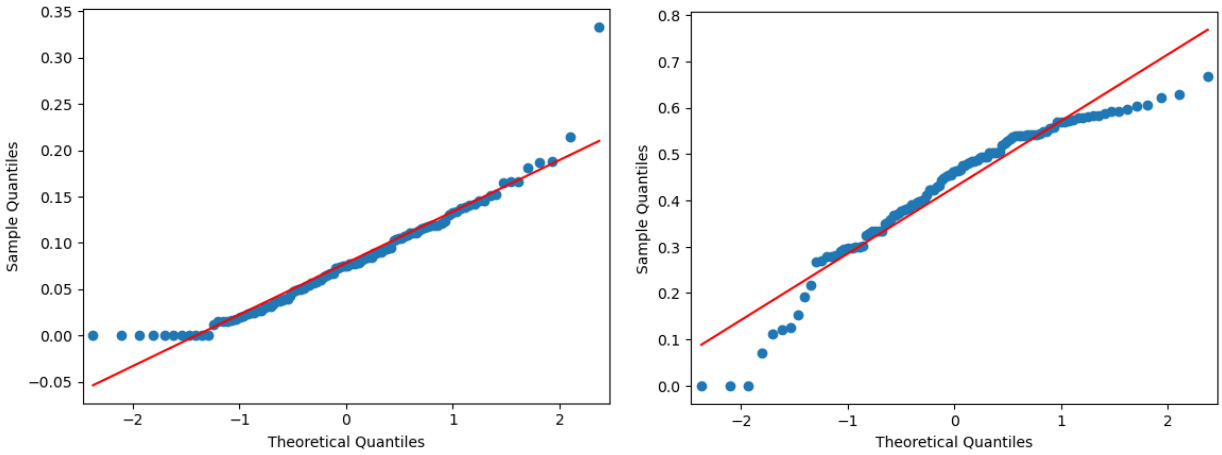
Prior to essentially performing the tests, we need to clean the data to eliminate any missing values, duplicate groups and variables that would not contribute to our research this time. To that end, we removed all the empty values from sex, perceived race, age group at arrest, strip search and formal booking, and only the five variables mentioned beforehand are kept. Specifically, we replaced 'Aged 65 years and older' with 'Aged 65 and older', 'Aged 17 years and under' with 'Aged 17 and younger' and 'Aged 17 years and younger' with 'Aged 17 and younger' in the age group at arrest. In addition, Sex 'U' was removed because the data is unknown and of no notable importance.

In order to conduct one-way ANOVA tests to compare the means of groups in the dataset, a continuous variable as the response variable must be present for the test to be valid. Nonetheless, both strip search and formal booking are binary variables that render the variance of data not meaningful. With respect to this particular context, we created two new variables called mean strip search score and mean booked score which are obtained from grouping by sex, perceived race and age group at arrest and calculating the mean strip search and formal booking for each group. This significantly helps us to generate continuous variables and form a new dataset with

only distinct combinations of groups, containing 112 rows and 5 columns of data (See Appendix Table E).

Figure 9

Q-Q Plots for mean strip search score and mean booked score



Note: Left is for mean strip search score, right is for mean booked score

Table 5. *Levene's test results for mean strip search score homogeneity of variances (Alpha = 0.05)*

Independent Variable	Statistic	P-value
Sex	0.17	0.67
Perceived Race	1.74	0.11
Age Group at Arrest	0.90	0.50

Note: all p-values > 0.05.

Table 6. *Levene's test results for mean booked score homogeneity of variances (Alpha = 0.05)*

Independent Variable	Statistic	P-value

Sex	11.32	0.01
Perceived Race	0.60	0.75
Age Group at Arrest	1.98	0.076

Note: only p-value of Sex < 0.05.

Firstly by checking the three assumptions of ANOVA test, independence, normality and equal variance, we are able to minimize possible bias results from the ANOVA test. All groups are mutually exclusive due to the fact that the new dataset is grouped by a few variables and cleaned at the very start. As Figure 9 displays, residuals mostly follow the straight line and only a few at the tail deviate, thus we deduced that mean strip search score and mean booked score are normally distributed. Lastly, according to Table 5 and 6, the non-significant results of the Levene's test for homogeneity of variances suggest that there are no statistically significant differences in variability between the groups. Only sex gives a significant p-value, therefore we would not consider sex. As such, none of the three assumptions are violated, so we would proceed to one-way ANOVA tests at this stage.

Results and Findings

Table 7. *One-way ANOVA test results (Alpha = 0.05)*

Dependent Variable	Independent Variable	Statistic	P-value
Mean Strip Search Score	Perceived Race	6.53	0.0036

Note: The p-value < 0.05. Perceived Race includes Middle-Eastern, Indigenous and Latino.

- **H0(Null Hypothesis):** Middle-Eastern, Indigenous and Latino all have the same population mean.
- **H1(Alternative Hypothesis):** At least one group has a different population mean than others.

The small p-value (0.0036) suggests that there is evidence to reject the null hypothesis. We can conclude that there is a significant difference between the group means of mean strip search score of perceived race. This is further supported by the high F-statistic (6.53), which highlights the large variation between the group means.

Table 8. *Multiple Comparison of Means - Tukey's HSD Test (Alpha = 0.05)*

Group 1	Group 2	Mean Difference	P-value	Lower	Upper	Reject
Indigenous	Latino	-0.061	0.034	-0.117	-0.004	True
Indigenous	Middle-Eastern	-0.081	0.00036	-0.137	-0.024	True
Latino	Middle-Eastern	-0.02	0.67	-0.0765	0.0365	False

A statistically significant difference between the Indigenous and the Latino is found due to the small p-value (0.034), the mean of the Indigenous is 0.061 units smaller than the mean of the East/Southeast Asian with a 95% confidence interval of (-0.117, -0.004). A statistically significant difference between the Indigenous and the Middle-Eastern is found due to the small p-value (0.00036), the mean of the Indigenous is 0.081 units smaller than the mean of the Middle-Eastern with a 95% confidence interval of (-0.137, -0.024).

Table 9. *One-way ANOVA test results (Alpha = 0.05)*

Dependent Variable	Independent Variable	Statistic	P-value
Mean Strip Search Score	Age Group at Arrest	3.02	0.093

Note: The p-value > 0.05. Age Group at Arrest includes Age 45 to 54 years and Age 55 to 64 years.

- **H0(Null Hypothesis):** Age 45 to 54 years and Age 55 to 64 years have the same population mean.
- **H1(Alternative Hypothesis):** At least one group has a different population mean than others.

The large p-value (0.093) suggests that there is not enough evidence to reject the null hypothesis. We can conclude that there is no significant difference between the group means of mean strip search score of age group at arrest. This is further supported by the low F-statistic (3.02), which highlights the small variation between the group means.

Table 10. *Multiple Comparison of Means - Tukey's HSD Test (Alpha = 0.05)*

Group 1	Group 2	Mean Difference	P-value	Lower	Upper	Reject
Aged 45 to 54 years	Aged 55 to 64 years	-0.023	0.093	-0.051	0.0041	False

No statistically significant difference between the Aged 45 to 54 years and the Aged 55 to 64 years is found due to the large p-value (0.093), the mean of the Aged 45 to 54 years is 0.023 units smaller than the mean of the Aged 55 to 64 years with a 95% confidence interval of (-0.051, 0.0041).

Table 11. *One-way ANOVA test results (Alpha = 0.05)*

Dependent Variable	Independent Variable	Statistic	P-value
Mean Booked Score	Perceived Race	6.81	0.015

Note: The p-value < 0.05. Perceived Race includes Black and East/Southeast Asian.

- **H0(Null Hypothesis):** Black and East/Southeast Asian have the same population mean.
- **H1(Alternative Hypothesis):** At least one group has a different population mean than others.

The small p-value (0.015) suggests that there is evidence to reject the null hypothesis. We can conclude that there is a significant difference between the group means of mean booked score of perceived race. This is further supported by the high F-statistic (6.81), which highlights the large variation between the group means.

Table 12. *Multiple Comparison of Means - Tukey's HSD Test (Alpha = 0.05)*

Group 1	Group 2	Mean Difference	P-value	Lower	Upper	Reject
Black	East/Southeast Asian	-0.10	0.015	-0.19	-0.022	True

A statistically significant difference between the Black and the East/Southeast Asian is found due to the small p-value (0.015), the mean of the Indigenous is 0.10 units smaller than the mean of the East/Southeast Asian with a 95% confidence interval of (-0.19, -0.022).

Table 13. *One-way ANOVA test results (Alpha = 0.05)*

Dependent Variable	Independent Variable	Statistic	P-value
Mean Booked Score	Age Group at Arrest	11.04	0.0024

Note: The p-value < 0.05. Age Group at Arrest includes Age 17 and younger, Age 18 to 24 years.

- **H0(Null Hypothesis):** Age 17 and younger and Age 18 to 24 years have the same population mean.
- **H1(Alternative Hypothesis):** At least one group has a different population mean than others.

The very small p-value (0.0024) suggests that there is enough evidence to reject the null hypothesis. We can conclude that there is a significant difference between the group means of mean booked score of age group at arrest. This is further supported by the high F-statistic (11.04), which highlights the large variation between the group means.

Table 14. *Multiple Comparison of Means - Tukey's HSD Test (Alpha = 0.05)*

Group 1	Group 2	Mean Difference	P-value	Lower	Upper	Reject
Aged 17 and younger	Aged 18 to 24 years	0.15	0.0024	0.059	0.25	True

A strong statistically significant difference between the Aged 17 and younger and the Aged 18 to 24 years is found due to the smallest p-value (0.0024), the mean of the Aged 17 and younger is

0.15 units smaller than the mean of the Aged 18 to 24 years with a 95% confidence interval of (0.059, 0.25).

Table 15. *Two-way ANOVA test results based on perceived race, age group at arrest and their interaction term*

Cases	**Sum of Squares**	**df**	**F**	**P-value**
Perceived Race	0.012	2	4.57	0.062
Age Group at Arrest	0.0071	1	0.56	0.48
Perceived Race ✻ Age Group at Arrest	0.000081	2	0.032	0.97
Residuals	0.0077	6		

Note. Type III Sum of Squares. Dependent Variable is mean strip search score.

Table 15 shows that no significant difference is found in the mean strip search score among perceived race ($F = 4.57$, $p = 0.062 > 0.05$), age groups at arrest ($F = 0.56$, $p = 0.48 > 0.05$) or in the interaction effect between perceived race and age group at arrest ($F = 0.032$, $p = 0.97 > 0.05$).

Tukey's HSD test was also conducted but demonstrated insignificant differences in the combinations of perceived race and age group at arrest, thus no interaction effect occurred. The output is shown in the Appendix Table C.

Figure 10. *Interaction Plot to show mean strip search score by perceived race and age group at arrest*

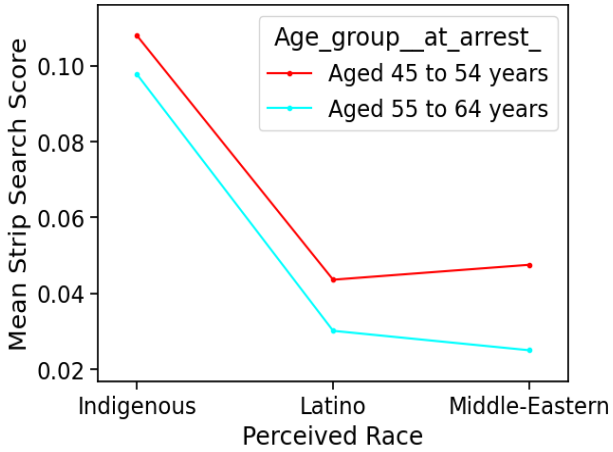


Figure 10 does not provide any information on statistically significant difference because of its non-parallel feature, this interaction plot depicts: (1) Aged 45 to 54 years Indigenous, Latino and Middle-Eastern people have a higher mean strip search score compared to Aged 55 to 64 years Indigenous, Latino and Middle-Eastern people; (2) Aged 45 to 54 years and Aged 55 to 64 years Latino and Middle-Eastern people seem to have a much larger mean strip search score gap compared to Aged 45 to 54 years and Aged 55 to 64 years Indigenous people.

Table 16. *Two-way ANOVA test results based on perceived race, age group at arrest and their interaction term*

Cases	**Sum of Squares**	**df**	**F**	**P-value**
Perceived Race	0.076	2	0.73	0.52
Age Group at Arrest	0.022	1	0.41	0.54
Perceived Race ✻ Age Group at Arrest	0.020	2	0.19	0.83
Residuals	0.31	6		

Note. Type III Sum of Squares. Dependent Variable is mean booked score.

Table 16 shows that no significant difference is found in the mean booked score among perceived race ($F = 0.73$, $p = 0.52 > 0.05$), age groups at arrest ($F = 0.41$, $p = 0.54 > 0.05$) or in the interaction effect between perceived race and age group at arrest ($F = 0.19$, $p = 0.83 > 0.05$).

Tukey's HSD test was also conducted but demonstrated insignificant differences in the combinations of perceived race and age group at arrest, thus no interaction effect occurred. The output is shown in the Appendix Table D.

Figure 11. *interaction Plot to show mean booked score by perceived race and age group at arrest*

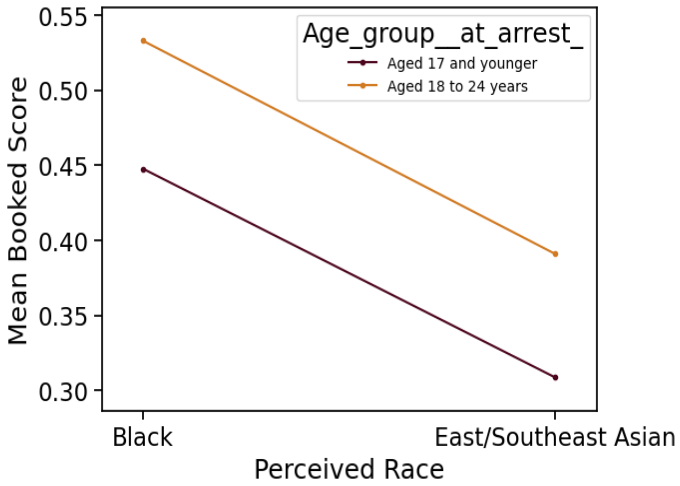


Figure 11 does demonstrate information on statistically significant difference because of its parallel feature, this interaction plot depicts: (1) Aged 17 and younger Black and East/Southeast Asian people and have higher mean booked score compared to Aged 18 to 24 years Black and East/Southeast Asian people ; (2) This is contradictory to the two-way ANOVA result we obtained previously, one possible reason can be the information shown as statistically significant difference in this interaction plot may be due to chance, supporting by both two-way ANOVA table and Tukey's HSD test failing to reject the null hypothesis.

Discussion and Conclusion

This research is based on the arrest and strip search dataset from the toronto public service, which contains 65,276 records. This time we only took into account demographic factors, we only kept three independent variables, which are Perceived Race, Sex and Age Group at Arrest. Youth at Arrest was not considered given the reason that it overlapped with the concept of Age Group at Arrest. After assessing normality assumptions and performing Welch's T-test across different groups for these three variables, we continued to analyze formal booking specifically for Black vs. East/Southeast Asian and Age(<17) vs. Age(18-24) groups. On the other hand, we focused on analyzing Middle-Eastern vs. Indigenous and Indigenous vs. Latino groups on strip search.

In addition, we used one-way ANOVA test, Tukey's HSD test and two-way ANOVA test for further examining how strip search and formal booking are impacted by specific groups of Perceived Race and Age Group at Arrest. The results from one-way ANOVA tests revealed that Latino and Middle-East people differ from other groups of individuals at arrest in a strip search. With conducting two-way ANOVA tests and drawing interaction plots, no different combinations or interactions between demographic factors affect the chance of being subjected to a strip search. In terms of being subjected to a formal booking, one-way ANOVA test results demonstrated that Black and East/Southeast people aged 24 and younger (Aged 17 and younger and Aged 18 to 24 years) differ from other groups of individuals at arrest. Also, no different combinations or interactions between demographic factors affect the chance of being subjected to a formal booking.

Ultimately, it is worth mentioning some limitations and future plans for the study. Sex did not pass the normality checking before Welch's T-test. Therefore, we will consider applying other statistical approaches to investigate whether there exists a gender bias in the future. In future research, we would also like to study whether actions at arrest have an impact on execution of strip search, as well as whether different kinds of strip search reasons are influenced by demographic factors. Next, we will construct a predicted model for this dataset and implement another real life dataset in this model to evaluate if our conclusions are trustworthy.

Reference

Michelle Psutka and Elizabeth A Sheehy, Strip-Searching of Women in Canada: Wrongs and Rights, 2016 94-2 *Canadian Bar Review* 241, 2016 CanLIIDocs 152, <https://canlii.ca/t/74q>, retrieved on 2023-02-22

McNeilly, G. (2019, March). *Summary of Ontario Jurisprudence Involving Strip Searches*. Retrieved February 22, 2023, from https://www.oiprd.on.ca/wp-content/uploads/OIPRD-Summary-of-Ontario-Jurisprudence-Involving-Strip-Searches-Accessible.pdf

R. V. Golden - SCC cases. (n.d.). Retrieved February 22, 2023, from https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/1924/index.do

Appendix

Table A. Describe of independent variables

```
White                   27708
Black                   17518
Unknown or Legacy        5052
East/Southeast Asian     4412
South Asian              3613
Middle-Eastern           3237
Indigenous               1932
Latino                   1767
Name: Perceived_Race, dtype: int64
M    52631
F    12608
Name: Sex, dtype: int64
Aged 25 to 34 years          20944
Aged 35 to 44 years          16240
Aged 18 to 24 years          10038
Aged 45 to 54 years           9065
Aged 55 to 64 years           4588
Aged 17 years and younger     1681
Aged 17 years and under       1361
Aged 65 years and older        698
Aged 65 and older              624
Name: Age_group__at_arrest_, dtype: int64
```

Table B. Describe of dependent variables

	Booked	StripSearch
count	65239.000000	65239.000000
mean	0.528426	0.119560
std	0.499195	0.324449
min	0.000000	0.000000
25%	0.000000	0.000000
50%	1.000000	0.000000
75%	1.000000	0.000000
max	1.000000	1.000000

Table C. Multiple Comparison of Means - Tukey's HSD Test (Alpha = 0.05)

Indigenous/Latino/Middle-Eastern & Age(45-54)/Age(55-64)

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
(Indigenous , Aged 45 to 54 years)	(Indigenous , Aged 55 to 64 years)	-0.0102	0.9996	-0.1526	0.1322	False
(Indigenous , Aged 45 to 54 years)	(Latino , Aged 45 to 54 years)	-0.0643	0.5279	-0.2066	0.0781	False
(Indigenous , Aged 45 to 54 years)	(Latino , Aged 55 to 64 years)	-0.0777	0.3641	-0.2201	0.0646	False
(Indigenous , Aged 45 to 54 years)	(Middle-Eastern , Aged 45 to 54 years)	-0.0603	0.5821	-0.2027	0.082	False
(Indigenous , Aged 45 to 54 years)	(Middle-Eastern , Aged 55 to 64 years)	-0.0828	0.3129	-0.2252	0.0595	False
(Indigenous , Aged 55 to 64 years)	(Latino , Aged 45 to 54 years)	-0.0541	0.6715	-0.1964	0.0883	False
(Indigenous , Aged 55 to 64 years)	(Latino , Aged 55 to 64 years)	-0.0675	0.4846	-0.2099	0.0748	False
(Indigenous , Aged 55 to 64 years)	(Middle-Eastern , Aged 45 to 54 years)	-0.0501	0.7274	-0.1925	0.0922	False
(Indigenous , Aged 55 to 64 years)	(Middle-Eastern , Aged 55 to 64 years)	-0.0726	0.4211	-0.215	0.0697	False
(Latino , Aged 45 to 54 years)	(Latino , Aged 55 to 64 years)	-0.0135	0.9985	-0.1558	0.1289	False
(Latino , Aged 45 to 54 years)	(Middle-Eastern , Aged 45 to 54 years)	0.0039	1.0	-0.1384	0.1463	False
(Latino , Aged 45 to 54 years)	(Middle-Eastern , Aged 55 to 64 years)	-0.0186	0.9932	-0.1609	0.1238	False
(Latino , Aged 55 to 64 years)	(Middle-Eastern , Aged 45 to 54 years)	0.0174	0.9949	-0.125	0.1598	False
(Latino , Aged 55 to 64 years)	(Middle-Eastern , Aged 55 to 64 years)	-0.0051	1.0	-0.1475	0.1372	False
(Middle-Eastern , Aged 45 to 54 years)	(Middle-Eastern , Aged 55 to 64 years)	-0.0225	0.9841	-0.1649	0.1198	False

Table D. Multiple Comparison of Means - Tukey's HSD Test (Alpha = 0.05)

Black/East-Southeast Asian & Age(<17)/Age(18-24)

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
(Black , Aged 17 and younger)	(Black , Aged 18 to 24 years)	0.0853	0.872	-0.3767	0.5473	False
(Black , Aged 17 and younger)	(East/Southeast Asian , Aged 17 and younger)	-0.1391	0.645	-0.6012	0.3229	False
(Black , Aged 17 and younger)	(East/Southeast Asian , Aged 18 to 24 years)	-0.0568	0.9548	-0.5188	0.4052	False
(Black , Aged 18 to 24 years)	(East/Southeast Asian , Aged 17 and younger)	-0.2245	0.3289	-0.6865	0.2376	False
(Black , Aged 18 to 24 years)	(East/Southeast Asian , Aged 18 to 24 years)	-0.1421	0.6319	-0.6042	0.3199	False
(East/Southeast Asian , Aged 17 and younger)	(East/Southeast Asian , Aged 18 to 24 years)	0.0823	0.8825	-0.3797	0.5444	False

Table E. Cleaned Dataset with 112 rows and 5 columns

	Sex	Perceived_Race	Age_group__at_arrest_	StripSearch	Booked
0	F	Black	Aged 17 and younger	0.037975	0.367089
1	F	Black	Aged 18 to 24 years	0.118911	0.462751
2	F	Black	Aged 25 to 34 years	0.080416	0.475875
3	F	Black	Aged 35 to 44 years	0.072917	0.463542
4	F	Black	Aged 45 to 54 years	0.116541	0.447368
5	F	Black	Aged 55 to 64 years	0.024096	0.301205
6	F	Black	Aged 65 and older	0.000000	0.384615
7	F	East/Southeast Asian	Aged 17 and younger	0.024390	0.268293
8	F	East/Southeast Asian	Aged 18 to 24 years	0.037594	0.278195
9	F	East/Southeast Asian	Aged 25 to 34 years	0.026316	0.378947
10	F	East/Southeast Asian	Aged 35 to 44 years	0.037037	0.328042
11	F	East/Southeast Asian	Aged 45 to 54 years	0.030928	0.298969
12	F	East/Southeast Asian	Aged 55 to 64 years	0.057971	0.217391
13	F	East/Southeast Asian	Aged 65 and older	0.000000	0.192308
14	F	Indigenous	Aged 17 and younger	0.111111	0.111111
15	F	Indigenous	Aged 18 to 24 years	0.114943	0.597701
16	F	Indigenous	Aged 25 to 34 years	0.137809	0.547703
17	F	Indigenous	Aged 35 to 44 years	0.102941	0.463235
18	F	Indigenous	Aged 45 to 54 years	0.076923	0.474359
19	F	Indigenous	Aged 55 to 64 years	0.121212	0.484848
20	F	Indigenous	Aged 65 and older	0.000000	0.333333
21	F	Latino	Aged 17 and younger	0.000000	0.071429
22	F	Latino	Aged 18 to 24 years	0.084746	0.491525
23	F	Latino	Aged 25 to 34 years	0.021053	0.505263
24	F	Latino	Aged 35 to 44 years	0.015625	0.421875
25	F	Latino	Aged 45 to 54 years	0.047619	0.380952
26	F	Latino	Aged 55 to 64 years	0.000000	0.000000
27	F	Latino	Aged 65 and older	0.166667	0.333333
28	F	Middle-Eastern	Aged 17 and younger	0.033333	0.300000

29	F	Middle-Eastern	Aged 18 to 24 years	0.063158	0.452632
30	F	Middle-Eastern	Aged 25 to 34 years	0.056180	0.359551
31	F	Middle-Eastern	Aged 35 to 44 years	0.065789	0.289474
32	F	Middle-Eastern	Aged 45 to 54 years	0.019608	0.294118
33	F	Middle-Eastern	Aged 55 to 64 years	0.000000	0.125000
34	F	Middle-Eastern	Aged 65 and older	0.000000	0.000000
35	F	South Asian	Aged 17 and younger	0.040000	0.120000
36	F	South Asian	Aged 18 to 24 years	0.074766	0.411215
37	F	South Asian	Aged 25 to 34 years	0.064935	0.422078
38	F	South Asian	Aged 35 to 44 years	0.049587	0.396694
39	F	South Asian	Aged 45 to 54 years	0.015873	0.333333
40	F	South Asian	Aged 55 to 64 years	0.000000	0.269231
41	F	South Asian	Aged 65 and older	0.076923	0.153846
42	F	Unknown or Legacy	Aged 17 and younger	0.027027	0.297297
43	F	Unknown or Legacy	Aged 18 to 24 years	0.132530	0.481928
44	F	Unknown or Legacy	Aged 25 to 34 years	0.118380	0.429907
45	F	Unknown or Legacy	Aged 35 to 44 years	0.107296	0.394850
46	F	Unknown or Legacy	Aged 45 to 54 years	0.089109	0.376238
47	F	Unknown or Legacy	Aged 55 to 64 years	0.017544	0.280702
48	F	Unknown or Legacy	Aged 65 and older	0.000000	0.000000
49	F	White	Aged 17 and younger	0.053640	0.279693
50	F	White	Aged 18 to 24 years	0.104972	0.432320
51	F	White	Aged 25 to 34 years	0.165605	0.502695
52	F	White	Aged 35 to 44 years	0.145056	0.502664
53	F	White	Aged 45 to 54 years	0.089730	0.454054
54	F	White	Aged 55 to 64 years	0.051151	0.368286
55	F	White	Aged 65 and older	0.015267	0.297710
56	M	Black	Aged 17 and younger	0.141700	0.528340
57	M	Black	Aged 18 to 24 years	0.181054	0.603298
58	M	Black	Aged 25 to 34 years	0.152417	0.604739
59	M	Black	Aged 35 to 44 years	0.141311	0.582951

60	M	Black	Aged 45 to 54 years	0.123268	0.557257
61	M	Black	Aged 55 to 64 years	0.111301	0.486301
62	M	Black	Aged 65 and older	0.067568	0.479730
63	M	East/Southeast Asian	Aged 17 and younger	0.054264	0.348837
64	M	East/Southeast Asian	Aged 18 to 24 years	0.118326	0.503608
65	M	East/Southeast Asian	Aged 25 to 34 years	0.083252	0.540646
66	M	East/Southeast Asian	Aged 35 to 44 years	0.092531	0.545151
67	M	East/Southeast Asian	Aged 45 to 54 years	0.078351	0.492784
68	M	East/Southeast Asian	Aged 55 to 64 years	0.060510	0.493631
69	M	East/Southeast Asian	Aged 65 and older	0.015625	0.390625
70	M	Indigenous	Aged 17 and younger	0.333333	0.666667
71	M	Indigenous	Aged 18 to 24 years	0.214815	0.622222
72	M	Indigenous	Aged 25 to 34 years	0.186788	0.592255
73	M	Indigenous	Aged 35 to 44 years	0.187783	0.583710
74	M	Indigenous	Aged 45 to 54 years	0.138756	0.588517
75	M	Indigenous	Aged 55 to 64 years	0.074074	0.629630
76	M	Indigenous	Aged 65 and older	0.000000	0.555556
77	M	Latino	Aged 17 and younger	0.166667	0.574074
78	M	Latino	Aged 18 to 24 years	0.084000	0.540000
79	M	Latino	Aged 25 to 34 years	0.078947	0.580827
80	M	Latino	Aged 35 to 44 years	0.093264	0.593264
81	M	Latino	Aged 45 to 54 years	0.039548	0.536723
82	M	Latino	Aged 55 to 64 years	0.060241	0.578313
83	M	Latino	Aged 65 and older	0.095238	0.571429
84	M	Middle-Eastern	Aged 17 and younger	0.074713	0.402299
85	M	Middle-Eastern	Aged 18 to 24 years	0.090000	0.531667
86	M	Middle-Eastern	Aged 25 to 34 years	0.056478	0.539313
87	M	Middle-Eastern	Aged 35 to 44 years	0.088647	0.542768
88	M	Middle-Eastern	Aged 45 to 54 years	0.075419	0.569832
89	M	Middle-Eastern	Aged 55 to 64 years	0.050000	0.518750
90	M	Middle-Eastern	Aged 65 and older	0.000000	0.400000

91	M	South Asian	Aged 17 and younger	0.104651	0.372093
92	M	South Asian	Aged 18 to 24 years	0.084034	0.465546
93	M	South Asian	Aged 25 to 34 years	0.073045	0.540123
94	M	South Asian	Aged 35 to 44 years	0.103546	0.548936
95	M	South Asian	Aged 45 to 54 years	0.042793	0.542793
96	M	South Asian	Aged 55 to 64 years	0.031963	0.502283
97	M	South Asian	Aged 65 and older	0.012048	0.325301
98	M	Unknown or Legacy	Aged 17 and younger	0.111111	0.422222
99	M	Unknown or Legacy	Aged 18 to 24 years	0.119013	0.494920
100	M	Unknown or Legacy	Aged 25 to 34 years	0.129909	0.579305
101	M	Unknown or Legacy	Aged 35 to 44 years	0.108446	0.523462
102	M	Unknown or Legacy	Aged 45 to 54 years	0.077329	0.455185
103	M	Unknown or Legacy	Aged 55 to 64 years	0.049020	0.444444
104	M	Unknown or Legacy	Aged 65 and older	0.030303	0.333333
105	M	White	Aged 17 and younger	0.067350	0.352855
106	M	White	Aged 18 to 24 years	0.133751	0.540752
107	M	White	Aged 25 to 34 years	0.150768	0.569142
108	M	White	Aged 35 to 44 years	0.145703	0.568983
109	M	White	Aged 45 to 54 years	0.113506	0.542597
110	M	White	Aged 55 to 64 years	0.095064	0.484461
111	M	White	Aged 65 and older	0.023769	0.390492