Bias in Canada's Prison System[*]

Jaffa Romain

22 December 2020

Abstract

The Globe and Mail's 2020 investigation of the Federal prison system found that Black and Indigenous inmates are more likely to get worse scores in their risk assessments than white inmates, simply because of their race. To investigate this further, this report will show findings of an independent analysis conducted using observational data obtained from the Correctional Service of Canada. The same methodology used by the Globe and Mail will be used; a multi-level logistic regression model is used to observe the factors that contribute to a prisoner's score in order to investigate whether there truly is uncovered systemic bias in the risk assessments.

keywords: Risk Assessment, Logistic Regression, Correctional Service of Canada, bias, Globe and Mail

1   Introduction

For prisoners in Canada's federal justice system, it can often be difficult navigating through life behind bars. This is particularly true in the case of many Black and Indigenous Canadians, who are over-represented in the prison population; In the Office of the Correctional Investigators 2018-2019 Annual report, it was found that approximately 8% of federal offenders identified as Black despite only accounting for 3.5% of Canada's total population, while Indigenous offenders made up over 20% of the offender population, despite only accounting for about 5% of Canada's population [Zinger, 2020].

Risk Assessments are a crucial component that influence several conditions in the time served by a prisoner, and are utilized to aid correctional practitioners in determining the scope and severity of correctional interventions [Public Safety Canada, 2018]. According to the Globe and Mail in the article, "Bias Behind Bars", the initial report - known as the preliminary report - feeds into all future assessments, as well as any future decisions, such as where a prisoner will serve his sentence, their access to rehabilitation programs, how often they will meet with officers, visitation privileges, and even the chances of getting parole, the initial report feeds into all the evaluations to come, as well as any future decisions [Cardoso, 2020]. In 2018, the Globe and Mail set out to investigate how race affected a prisoner's time in prison using data from the Correctional Service of Canada (CSC). Through extensive research they found two key scores in the preliminary reports: an inmate's security classification, and their reintegration score, that factor heavily in an inmate's time in prison and their process integrating back into society. The Globe used the data obtained, accounting for variables including race, age, gender, and their most serious offense. Using logistic regression, the Globe studied the impact of these different risk assessment variables, to look at how race affects an inmate's security level, reintegration score, the reintegration score on reoffending. Through their analysis, they found that the over-represented prison population of Black and Indigenous inmates experienced bias in these risk assessments. For example, it was found that Black men are more likely to receive a "maximum" initial security rating than white men. This raises concern, as this rating is crucial for access to treatment programs. Indigenous men are more likely to receive the lowest reintegration potential score, a score that plays a significant role in parole decisions [Cardoso, 2020]. It was also found that the risk assessments overestimate the likelihood

---

[*]Code supporting this analysis is available at: https://github.com/jaffaromain/Bias-Behind-Bars

of recidivism in these two ethnic groups. These results lead to the conclusion that Black and Indigenous inmates are more likely to get worse scores than white inmates, solely based on their race.

In this report, using [R Core Team, 2019], the same data from the Correctional Service of Canada (CSC) is used in an attempt to replicate the results from the Globe and Mail, but under slightly different assumptions. Logistic regression models are used for assessing the two key two key scores in the preliminary reports: an inmate's security classification, and their reintegration score based on different factors including age and gender. Section 2 discusses the data used, including the methodology,the methodology and the strengths and weaknesses of the data. Additionally, the choice of variables is explained and some basic exploratory data analysis is presented. Section 3 discusses the model(s) implemented, as well as the implications of choosing logistic regression to model the data. This is then followed by the results in section 4, and then a discussion of the determined outcomes and notable findings, and what it means moving forward with regards to bias in Canada's prison system in section 5.

2 Data

To create the models, the same data used by the Globe and Mail was used. The Globe and Mail obtained this data from the Correctional Service of Canada (CSC) after requesting a freedom of interest request in 2018. The data set can be found with the Globe and Mail's accompanying methodology article. The data was cleaned in order to be utilized for this analysis using R [**?**]. One variable, `JUDGE`, was removed as it had null entries. Since we are only interested in federal offenders, those who were classified as being under provincial jurisdiction were removed. Given this cleaned data set, there were several notable variables of interest:

Race – This corresponds to the race/ethno-cultural background of the offender. This information is self reported by the offender at the time the offender is admitted.In the original data set, there were over 30 possible categories for race. Using Ontario's Data Standard for the Identification and Monitoring of Systemic Racism, different categories were merged [Government of Ontario, 2020]. For example, people who identified as Caribbean were grouped with those who identified as Black. In Figure 1, we can see the proportion of offenders by race. White, Black, and Indigenous inmates have the largest representation in the prison population.

Race Grouping – This summarizes the race of inmates into two categories. "Indigenous" corresponds to offenders with a Race of "First Nations", "Métis", or "Inuit". All other inmates are categorized as "Non Indigenous".

This report focuses on the disparities between white, Black and Indigenous inmates. Using the `Race Grouping` variable, inmates can be further consolidated into 4 categories, which can be seen Figure 2.

Gender – This corresponds to the gender of the offender at the time the data were extracted. Figure 3 shows the disparity between the number of male and female inmates.

Although female inmates represent a much smaller population of the data, it is important to note that female inmates have scores like their Dynamic score factor into the reintegration rating, unlike male inmates. This would suggest that the female inmate population could follow different trends compared to the male prison population, which will need to be addressed when modeling the data. As we can see in Figure 4, female inmates have a higher proportion of Indigenous inmates in comparison to the total population.

Age – This is the age of the offender, in years, at the time the data were extracted. An inmate can have multiple entries of this variable depending on how many fiscal years they spend imprisoned.

Offender Security Level - The offender security level is used as an indicator to match inmates to institutions and treatment programs. This assessed using a tool known as the Custody Rating Scale [Cardoso, 2020]. The security level can be set to three levels: minimum, medium, or maximum. An inmate who receives a maximum security level will complete their sentence in a maximum-security facility.
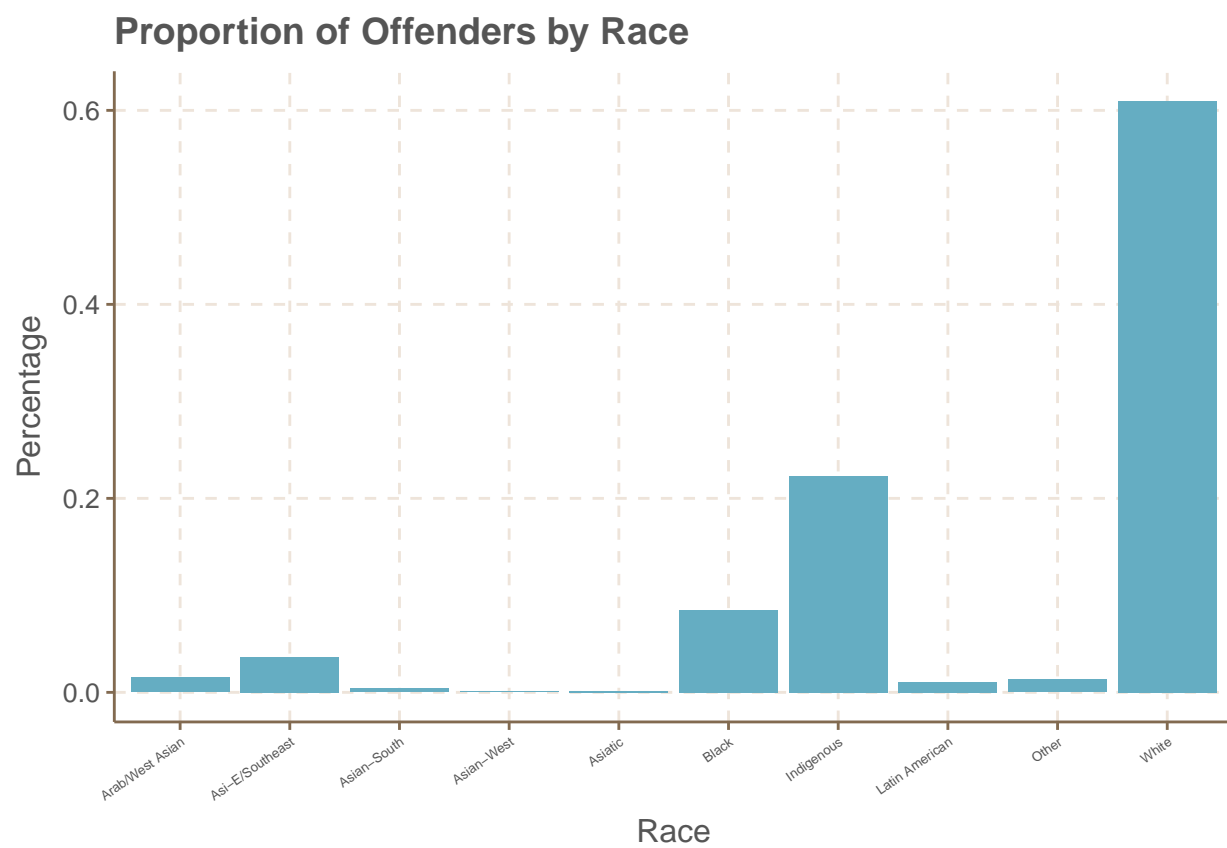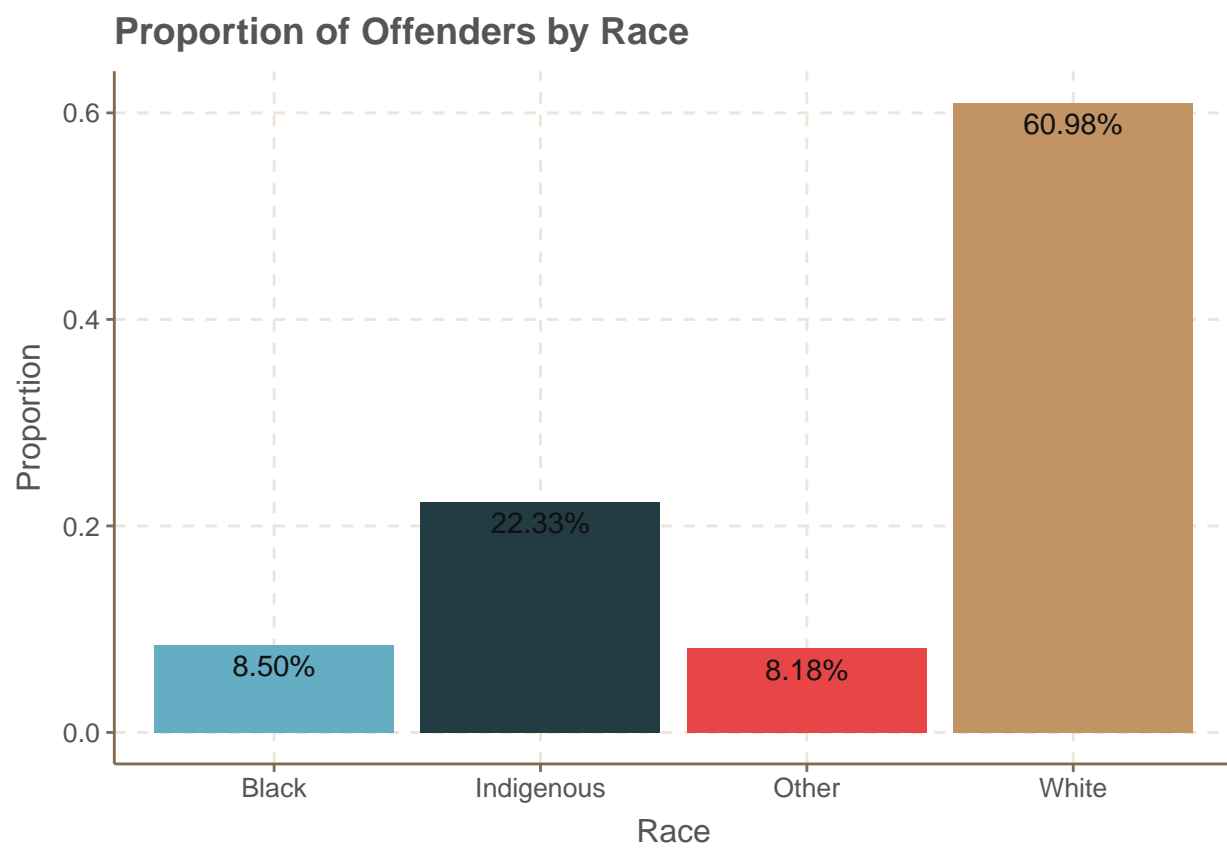
Figure 1: Offenders By Race

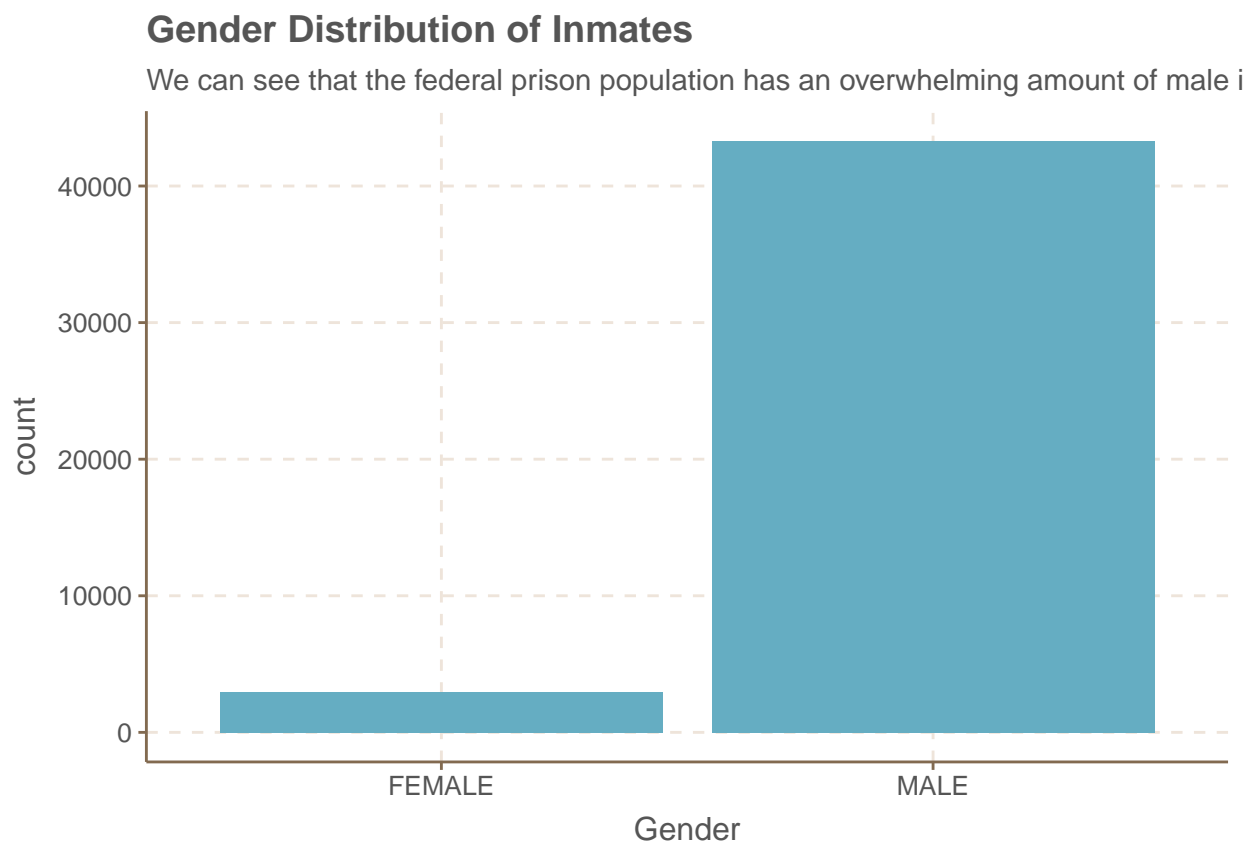Figure 2: Offenders by 4 Categories of Race

Gender Distribution of Inmates

We can see that the federal prison population has an overwhelming amount of male i

Figure 3: Offenders By Gender

Figure 4: Offenders by Race and Gender

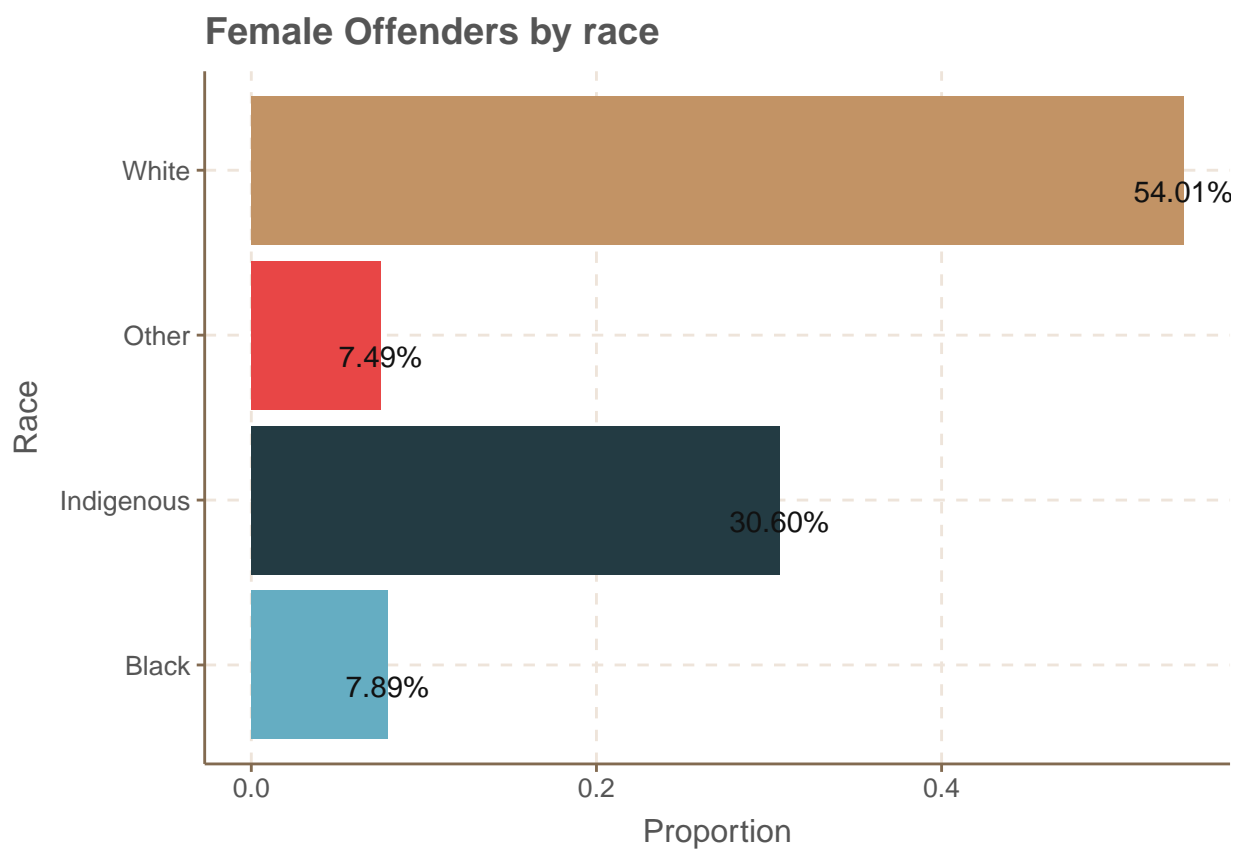# Inmate Security Level Score

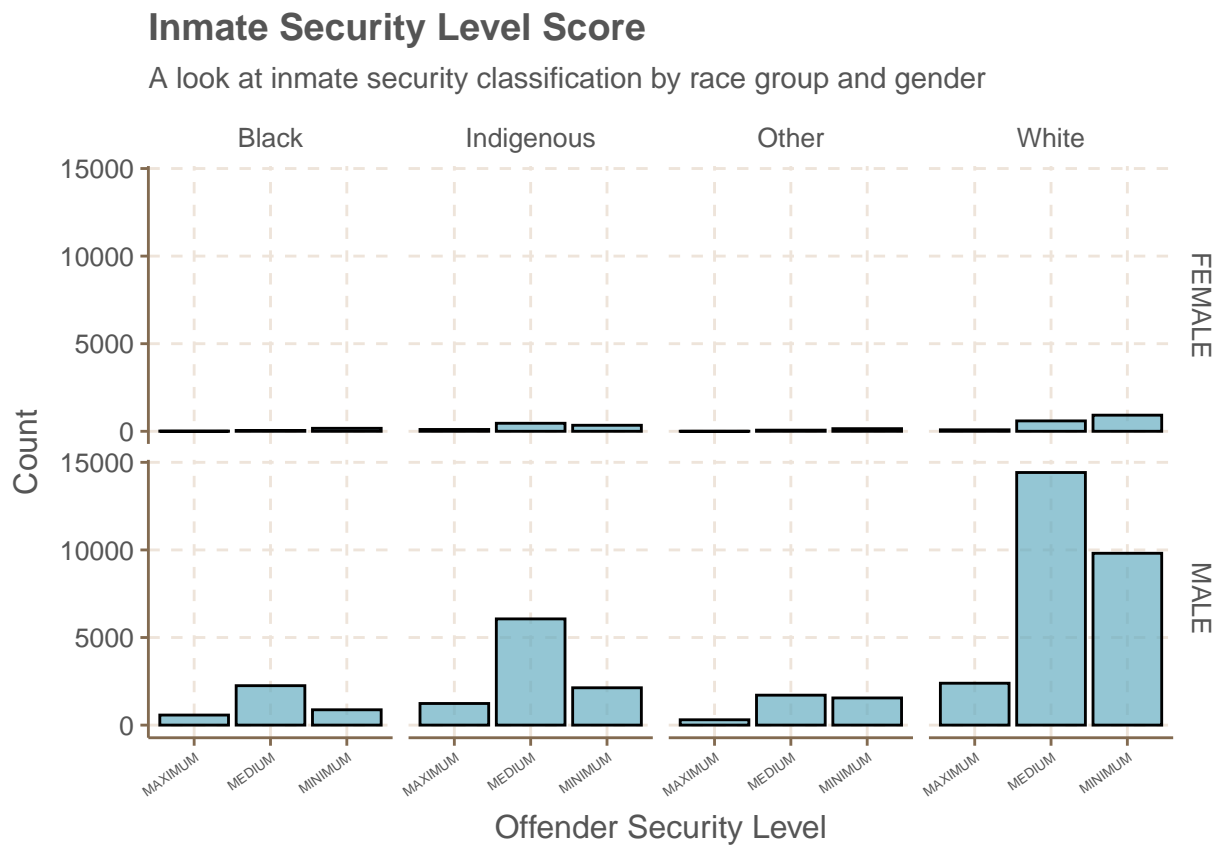A look at inmate security classification by race group and gender

Figure 5: Distribution of Inmate Security Level Scores

Reintegration Potential - An inmate's reintegration score factors heavily near the end of an inmate's sentence, and plays a large role in parole hearings. This score estimates the likelihood of an inmate successfully reintegrating into society with committing a new offense.

Inmate Reintegration Scores

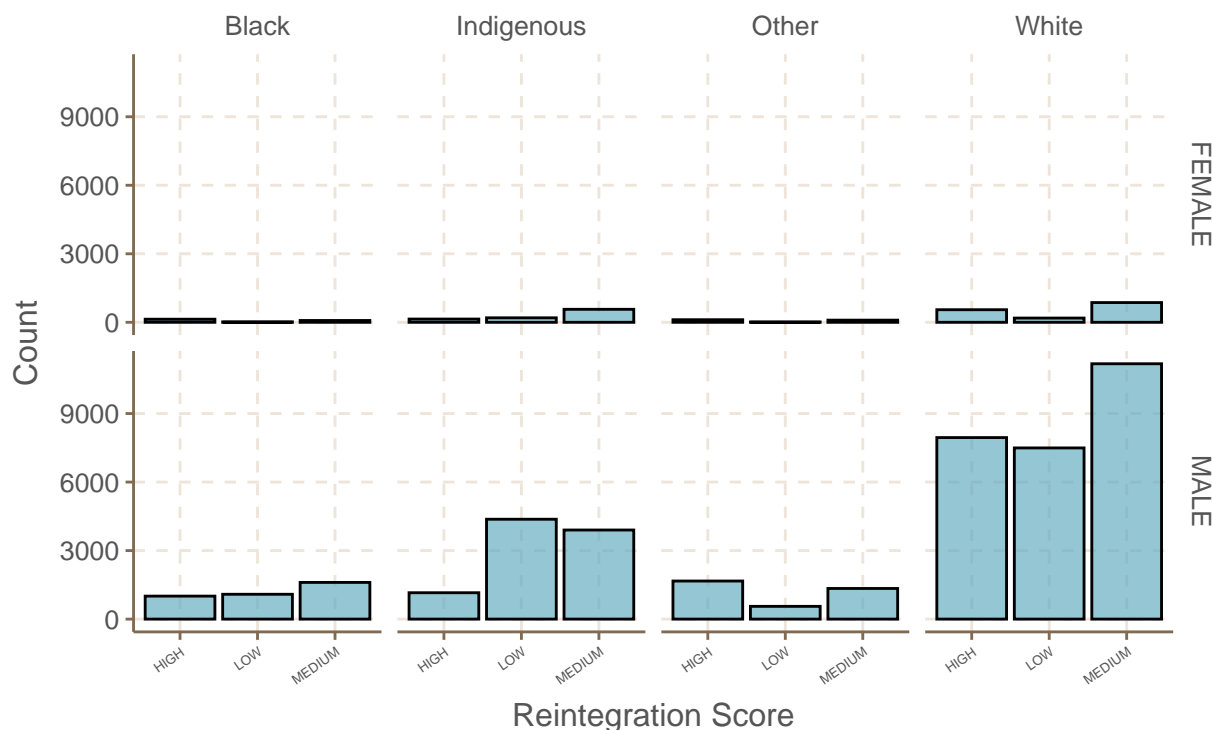A look at Reintegration Scores by race group and gender

Figure 6: Distribution of Inmate Reintegration Scores

2.1 Methodology

The Correctional Service of Canada data set is a cross-sectional time series data set. That is, it is data derived from federal inmate data over time. It contains information from the agency's database of inmates, providing a snapshot of 7 years' worth of entries (from 2012 to 2018). The data set initially contained 744 958 rows and 26 variables, documenting information of 50 166 people who are either in custody or supervised leave. All observations are logged records of an inmates, risk assessment score, age, gender, province, and more. The survey had a high response rate, with 52.4% completion. Even with this response rate, to reduce non-response bias, survey weights were adjusted to account for non-response as much as possible using 2016 Census data. The data was enhanced by combining responses with information from personal and household tax data, giving more information to work with and draw more conclusions from. A minimum sample size was also used to ensure an acceptable amount of variability at the stratum level.

2.2 Strengths and Weaknesses

One of the key strengths of this data set is its data collection method and confidentiality. The data set maintains the anonymity inmates by generating unique IDs for each inmate. Using the unique IDs, we are able to identify multiple sentences and offenses of each inmate. As this data is collected directly from the

government agency responsible for federal prisons, we can expect that the data set is an accurate reflection of the federal prison population.

One limitation when working with the data set is the ability to interpret results from the data. Although offenders are identified with a unique ID, inmates can have multiple sentence IDs and offenses. This also leads to there being multiple rows of entries for identifiers such as race for each inmate. An issue that develops as a result is the interpretation of different observations. The Globe and Mail found that the data set contains 2 267 cases of inmate data where the inmates' race changes over time. Although this only represents roughly 5% of the data set, it makes it unclear which fiscal year accurately represents the inmates. Race is self-reported in the data set but it is unclear as what the specifications are for this category when filling out the report.

3   Model

4   The logistic Model

To attempt to replicate the Globe and Mail's analysis, we will be focusing on the two key scores in the risk assessments: offender security score and reintegration score. To model these scores we will use the same methodology; using logistic regression. A logistic regression model allows us to model the relationship between a binary response variable and categorical/continuous predictor variables. In this case, our response variables are the key scores we aim to predict, setting 1 to be a maximum score, and 0 to be a minimum or medium score. Using this model allows us to predict the probability of an occurrence by fitting the data to a logit functions and does not require the assumption of the normality of independent variables. One trade off when choosing a logistic regression model is that the output of the model is the odd ratio of an event. A large odds ratio tells us that the large chance of someone getting a maximum score only means that the chance of a particular event is much greater than that of the reference group. An alternative method that could have been used is a linear additive model. However, this would limit the ability to model any key interaction terms, and assumptions such as the normality of errors would have to be met.

Two models will be built using a subset of the cleaned data. The first model will look at how race affects an inmate's security level. The second model will look at how race affects reintegration score. The two models will be run separately for men and women.

To look at how well the models fit the data set, two standard metrics will be used. The first is the metric AUC. AUC is used to establish how well our model is able to distinguish between two possible outcomes [Cardoso, 2020]. In this case, how well the model is able to distinguish the key scores. An AUC of 0.5 means the model performs about as well as a coin toss. As AUC increases, we expect performance to increase, with high validity in a model being at 0.71 [Cardoso, 2020]. The fitted models had AUCs ranging from

Each model willfollow the following equation:

$$Pr(y_i = 1) = logit^-1(X_i * \beta)$$

where $y_i = 1$ represents our response variables. This will be either the reintegration score, security score, or an inmate reoffending. $Pr(y_i = 1)$ is the probability of the response variable occuring. The $X_i\beta$ are the linear predictors, where β represents the fitted coefficients for each independent variable X_i in the model. The coefficient values signify the mean variable changes in a one unit shift in the given variable β, which represents each race category.

The models are fit using the `glm` function. Tables made to show the results were made using `kable` from `knitr`. Model results were tidied with `broom`.

Table 1: Model 1 Coefficents: Male Inmates

	Coefficient	P-value
(Intercept)	-2.097	0.000
race_groupBlack	0.543	0.000
race_groupIndigenous	0.327	0.000
race_groupOther	-0.020	0.766

Table 2: Model 1 Coefficents: Female Inmates

	Coefficient	P-value
(Intercept)	-2.435	0.000
race_groupBlack	0.084	0.782
race_groupIndigenous	0.610	0.000
race_groupOther	-1.061	0.041

4.1 Model 1: Effect of Race on an Inmate's Security Level

This model will look at the likelihood of an inmate receiving a maximum secure score. The data for this model will be a subset of the cleaned data set. Since the inmates are only assessed with this score when they have just begun their sentence, this subset of data will use an inmate who are in custody and their minimum age. This subset of data gives us 36 187 inmate cases.

4.2 Model 1 Males Subset:

The final fitted model has the following equation:

$$Pr(MaximumScore) = -2.097 + 0.543x_{black} + 0.327x_{Indigenous} - 0.02x_{Other}$$

Model 1 Females Subset:

The final fitted model has the following equation:

$$Pr(MaximumScore) = -2.435 + 0.084x_{black} + 0.61x_{Indigenous} - 1.061x_{Other}$$

4.3 Model 2: Effects of Race on Reintegration Score

This model will look at the likelihood of an inmate receiving a maximum reintegration score. The data for this model will be a subset of the cleaned data set. The subset of data will only look at inmates in custody. This subset of data gives us 91 732 inmate cases.

4.4 Model 2 Males Subset:

The final fitted model has the following equation:

$$Pr(MaximumScore) = -1.265 + 0.078x_{black} - 1.1115x_{Indigenous} + 0.747x_{Other}$$

4.5 Model 2 Females Subset:

The final fitted model has the following equation:

$$Pr(MaximumScore) = -1.202 + 1.190x_{black} - 1.009x_{Indigenous} + 0.628x_{Other}$$

Table 3: Model 2 Coefficents: Male Inmates

	Coefficient	P-value
(Intercept)	-1.265	0.000
race_groupBlack	0.078	0.088
race_groupIndigenous	-1.115	0.000
race_groupOther	0.747	0.000

Table 4: Model 2 Coefficents: Female Inmates

	Coefficient	P-value
(Intercept)	-1.202	0.000
race_groupBlack	1.190	0.000
race_groupIndigenous	-1.009	0.000
race_groupOther	0.628	0.001

We can see from table 5 the AUC for each model. The value for AUC ranges from 0.55-0.67. This varies from the results found in the analysis done by the Globe and Mail, where the AUC values were higher. The range of AUCs found here indicate that the models have weak to moderate validity. Although the models perform better than an average coin toss, the scores indicate that there is room for improvement.

5 Results

In the Globe and Mail's Analysis, a significance level of 0.05 was set as a threshold for determining statistical significance in results. Any value or estimate that is statistically significance is likely to not have occurred randomly, but is instead likely to have occurred due to a specific cause. For our models, we can look at the p-values of each of our coefficients to determine the statistical significance of each coefficient. When analyzing p-values, we assume the null hypothesis that there is no relationship between a given factor and the probability of getting a maximum score. P-values less than the significance level suggests strong evidence against our null hypothesis, as the p-value is the probability of observing a value as or more extreme than our coefficient. Both models uses the race group `White` as a reference level.

For the first model that estimates the Security score, both the Black men and Indigenous men had statistically significant p-values. The race category Other, had a high p-value of 0.766, indicating that inmates being in this category had no discernible effect from being a white man. Using the `jtools` package, we are able to plot Figure **??**, which shows that both black males have an increased likelihood of receiving a maximum security score, as well as Indigenous women. In this report, black men were 54% more likely to receive a maximum security score than white men. Indigenous men were 32% more likely. For the female inmates, Black inmates did not have statistically significant coefficients. The coefficient for Indigenous was statistically significant, and signified a 60% likelihood of receiving a maximum score compared to a white female inmate.

Table 5: AUCs for Reintegration and Security Score Models

models	AUCs
Model1 Male	0.55
Model1 Female	0.59
Model2 Male	0.61
Model2 Female	0.67
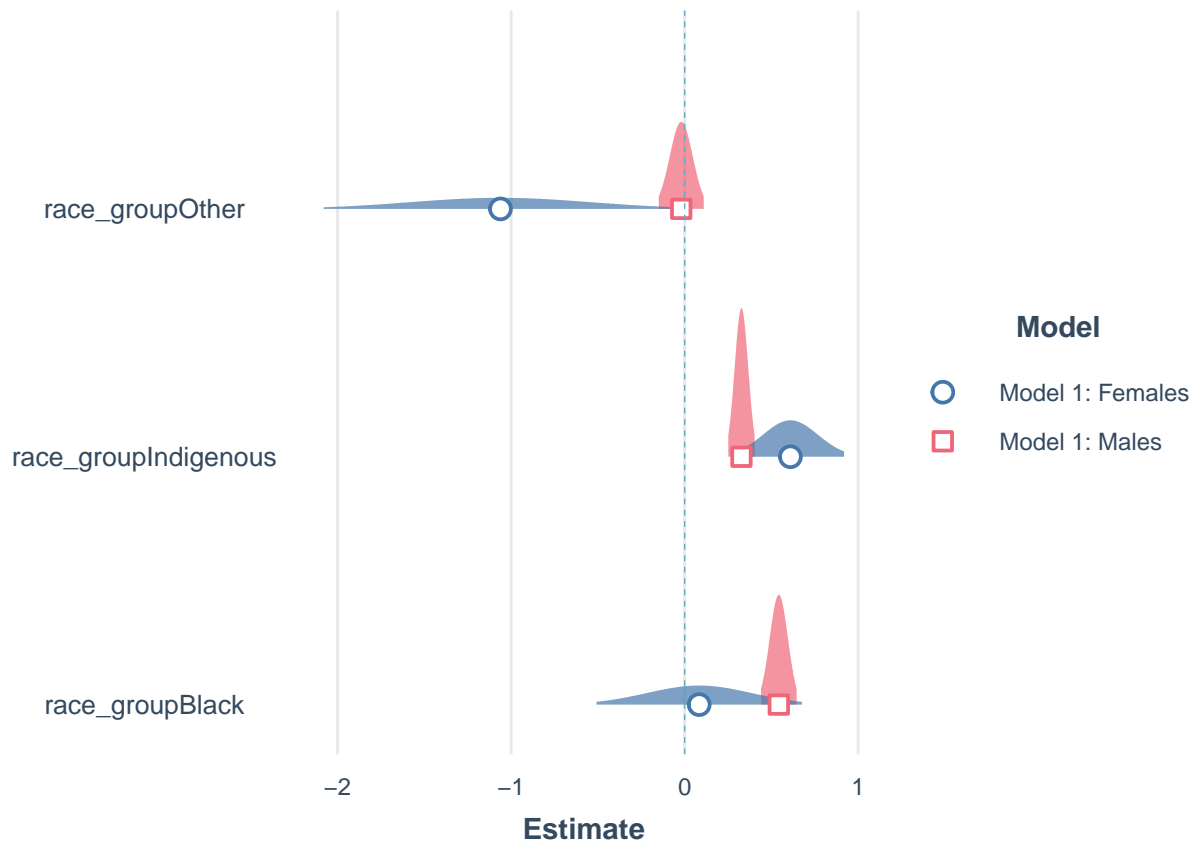
For the second model that estimates the Reintegration score of inmates, the Indigenous men, and inmates categorized as other had statistically significant p-values.the coefficient for Black mean, had a p-value of 0.088, which is above the significance threshold, indicating that inmates being in this category had no discernible effect from being a white man. Figure 7 shows the coefficients of the models for both male and female inmates. For male inmates, the likelihood of Indigenous inmates receiving a maximum security score decreases. For those categorized as other, the likelihood increases. For the female inmates, all coefficients were statistically significant. For Black female inmates, the likelihood increases, while for Indigenous inmates, likelihood decreases. For those categorized as Other, the likelihood increases.

6 Discussion

2020 has seen a wave of protests and riots against the systemic racism, particularly in the United States, where there have been monumental protests against police brutality, such as the George Floyd protests. In Canada, where pride is taken in being a diverse country, racial equity directly impacts a large number of its citizens identifying in minority groups. The Globe and Mail article, "Bias Behind Bars" bring to light racial discrimination in Canada's federal system. The article examines the risk assessments used by correctional institutions in Canada that are in charge of the inmates. The article concluded that there were netted statistically significant results for Indigenous women and Black men that suggest racial bias in these risk assessments.

Adopting the same methodology as the Globe and Mail, the results shows statistically significant results that suggests bias against black men, Indigenous men, and Indigenous women, although they differed slightly from the Globe and Mail's findings. These findings support claims of racial bias in Canada's prison, and raises questions as to whether the risk assessments themselves are designed in such a way that encourages bias against these particular groups. The risk assessments were initially developed in the 1980s and 1990s using
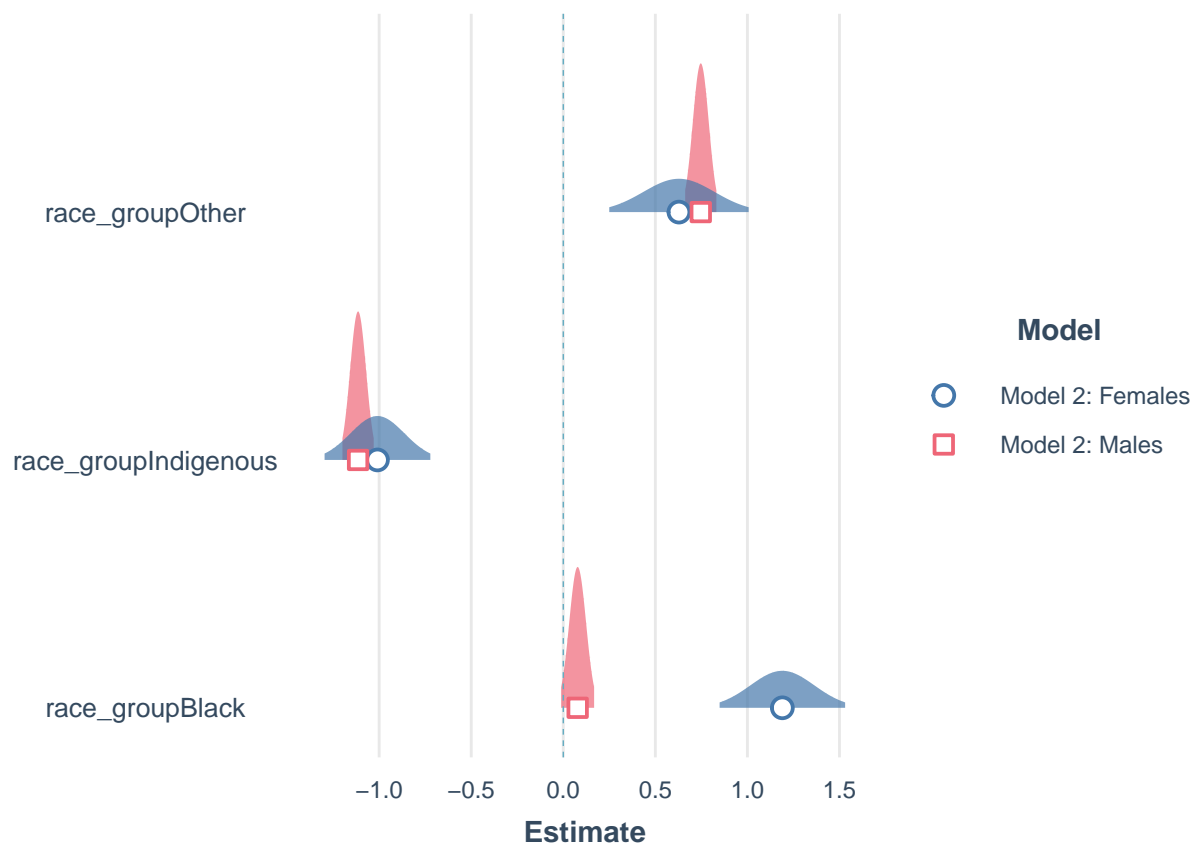
Figure 7: Model Coefficient for Reintegration Score Models

the Correctional Service of Canada's own administrative data [Cardoso, 2020]. Consequently, any bias in the collection of administrative data would be present in the current risk assessments. The risk assessments are subjective by nature, as the questions are mixed with numerical scores and scores based on a parole officer's judgment [Cardoso, 2020]. This makes it possible to have variance in how inmates are scored due to the lack of standardized assessments. Potential bias can have significant implications for Black and Indigenous inmates. The Offender Security Level score directly impacts a prisoner's time in a facility. An inmate given a worse score loses access to specific treatment programs, and certain types of jobs.

6.1 Weaknesses and Next Steps

The main goal of this analysis was to attempt to replicate the Globe and Mail article. However, a major limitation was the lack of reproducibility with respects to the original report. This was first apparent when cleaning the data, where the number of entries differed from that of the data used in the article, despite following the same methodology. Many steps were vague. For example, when grouping inmates by race, it was suggested to use Ontario's Data Standard for System racism. However, there were categories such as Multirac/Ethnic, that were grouped into other that represent a large percentage of the prison population. This made it possible for the data set to differ. When modelling, the data subsets used in the models were significantly different from the data used in the article, which may justify the disparity in model performance.

While the findings of this report are consistent with the Globe and Mail study, the model performance results indicates that the accuracy of the model could be improved. Models used one predictor variable with two separate subsets. The initial data collection segregated the data of male and female prisoners, despite a lack of analysis to demonstrate that the two groups behaved independently. It is possible that the 'gender' variable could be used as an interaction term in the model.

7 References

data set: Tom Cardoso.Bias behind bars: A Globe investigation finds a prison system stacked against Black and In-digenous inmates. The Globe and Mail, Canada, 2020. URL https://www.theglobeandmail.com/canada/article-investigation-racial-bias-in-canadian-prison-risk-assessments/

References

Tom Cardoso. *Bias behind bars: A Globe investigation finds a prison system stacked against Black and Indigenous inmates.* The Globe and Mail, Canada, 2020. URL https://www.theglobeandmail.com/canada/article-investigation-racial-bias-in-canadian-prison-risk-assessments/.

Government of Ontario. *Data Standards for the Identification and Monitoring of Systemic Racism*, 2020. URL https://www.ontario.ca/document/data-standards-identification-and-monitoring-systemic-racism.

Public Safety Canada. *Offender Risk Assessment Practices Vary Across Canada.* Public Safety Canada, Ottawa, Canada, 2018. URL https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/2017-s015/index-en.aspx.

R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2019. URL https://www.R-project.org.

Hadley Wickham and Jim Hester. *readr: Read Rectangular Text Data*, 2020. URL https://CRAN.R-project.org/package=readr. R package version 1.4.0.

Hadley Wickham, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo, and Hiroaki Yutani. Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686, 2019. doi: 10.21105/joss.01686.

Hadley Wickham, Romain François, Lionel Henry, and Kirill Müller. *dplyr: A Grammar of Data Manipulation*, 2020. URL https://CRAN.R-project.org/package=dplyr. R package version 1.0.2.

Ivan Zinger. *Office of the Correctional Investigator Annual Report 2019-2020*. The Globe and Mail, Ottawa, Canada, 2020. URL https://www.oci-bec.gc.ca/cnt/rpt/annrpt/annrpt20192020-eng.aspx.