

# Hidden Class HW1 Report

Yuxin Feng

Spring 2022

## 1 Data Collection

According to the NLSY97 data, I collected the information and filtered unimportant data. By grouping race and gender, we get a clean version of NLSY97 data for more analysis.

## 2 Figure

Figure 1 tells us the mean number of incarcerations happened in 2002 according to Race and Gender. We know that the number of woman and man has great difference. Usually numbers of woman are smaller than man. (We have some lack data in man - Mixed Races.) In addition, when paying attention to races, black man has larger numbers than any other races.

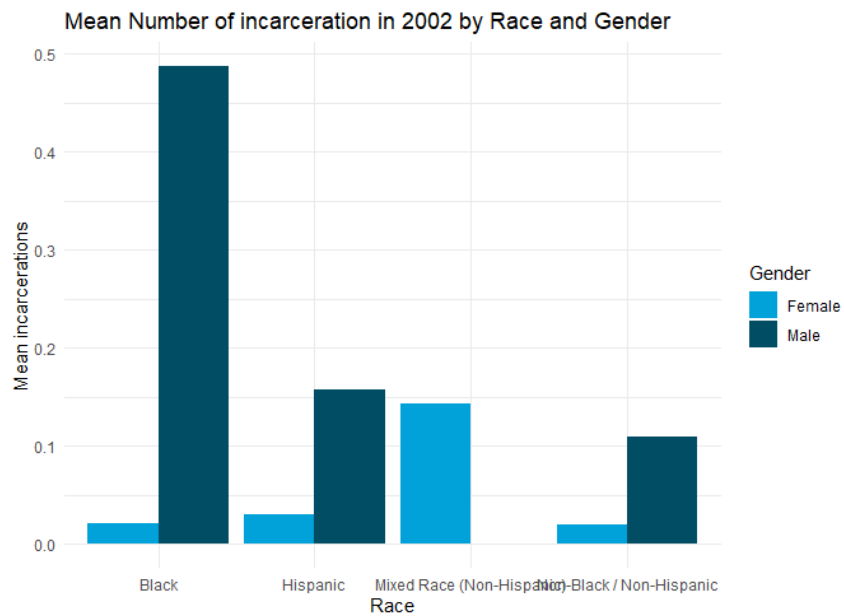


Figure 1: Mean Number of incarcerations in 2002 by Race and Gender

### 3 Table

We can have a better view of the exact probabilities of incarcerations from the table. It has the same qualitative result as the picture we saw above.

Table 1: Mean incarcerations in 2002 by Race and Gender

Gender	Black	Hispanic	Mixed Race Non Hispanic	Non Black Non Hispanic
Female	0.0211268	0.0298013	0.1428571	0.0193192
Male	0.4876712	0.1579509	0.0000000	0.1099476

The second table gives us a result of the regression between black women and other variables. The result tells us that the correlation between black women and black men is positive, while the others are positive. Also, the regression has a very small R square, which means that it may not be a good fit for the data.

Table 2: Regression Output. Omitted category is Black Females.

	<i>Dependent variable:</i> incarcerations in 2002
Hispanic	−0.159*** (0.038)
Mixed Race (Non-Hispanic)	−0.174** (0.083)
Non-Black / Non-Hispanic	−0.189*** (0.035)
Male	0.194*** (0.022)
Constant	0.155*** (0.026)
Observations	8,621
R <sup>2</sup>	0.015
Adjusted R <sup>2</sup>	0.014
Residual Std. Error	1.019 (df = 8616)
F Statistic	32.033*** (df = 4; 8616)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01