

Incarceration Analysis

Robert Petit

February, 2022

Using the NLSY97 data, we can investigate some information about incarcerations among the participants. The first notable trends show up when we simply look at the mean number of months spent incarcerated by race and gender, as seen in Figure 1.

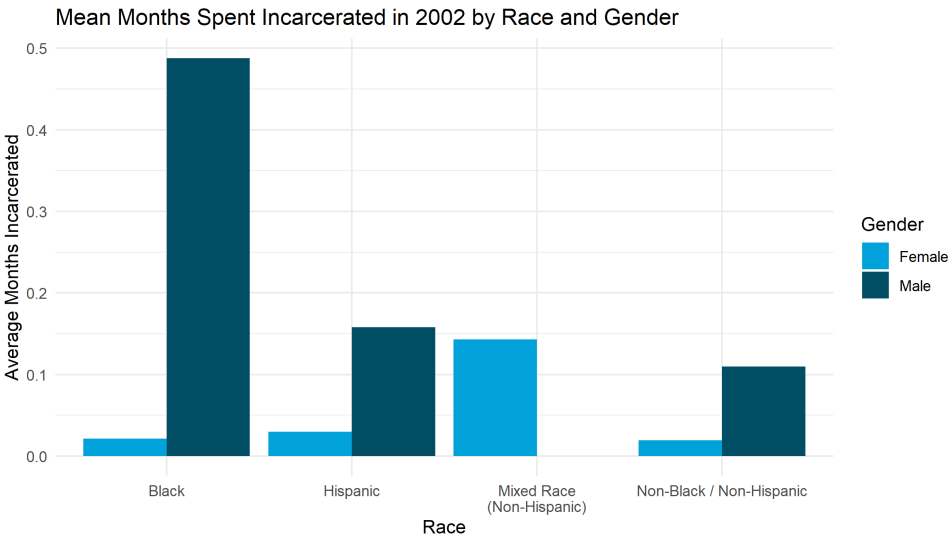


Figure 1: Months spent incarcerated, by race and gender

These trends are not incredibly surprising, the number of months spent incarcerated during a year is dramatically higher for black males than any other combination of race and gender; it is more than triple that of the next highest category, seen more precisely in table 1.

Table 1: Mean Months Incarcerated 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

A very simple regression also gives some estimation of the magnitude of these effects.

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

	<i>Dependent variable:</i>
	Months Incarcerated 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 ²	0.015
Adjusted R ²	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	

Since black females are the omitted category, in order to understand the expectation of black males related to other combinations we have to observe the intercept and male coefficient simultaneously. As it turns out, these are the *only* two positive coefficients which captures the information of the above graphs more accurately. It should also be noted that these values are all statistically significant to a high degree, with p-values less than 0.01.

Ultimately, this data is very consistent with the common understanding of incarceration in the US. It varies significantly and consistently based on race. It should be noted that these are all correlational observations, and making a causal claim requires more complex observations.