Race and Crime

Analysis of the U.S. National Registry of Exonerations

Alejandro D. Hernandez

August 2024

Background

The pursuit of justice is a fundamental pillar of the United States legal system, aiming to hold accountability for the guilty and protection of the innocent. However, the justice system design induces wrongful convictions through eyewitness misidentification, false or misleading forensic evidence, perjury, official misconduct, and inadequate legal defense. The consequences of wrongful conviction are serious and often irreversible, resulting in loss of liberty and damage to the reputation of individuals and their families. Increased scrutiny of criminal cases and efforts of dedicated legal advocates have facilitated the *exoneration* (reversal of conviction and clearing of charges) of numerous individuals who were unjustly convicted. Exoneration brings attention to shortcomings of the criminal justice system and consideration of innocents who are currently imprisoned.

The National Registry of Exonerations is a project from the University of California, Irvine; University of Michigan; and Michigan State University which gathers detailed information on every exoneration in the United States since 1989. Their extensive registry includes personal demographics and describes criminal sentences, time served, and pathway to exoneration. This wealth of information allows researchers, policy-makers, legal practitioners, and the public to analyze trends, identify systemic issues within the criminal justice system, and advocate for reforms aimed at preventing future wrongful convictions.

Objective

In this study, we examine the National Registry of Exonerations from the University of Michigan to analyze individuals exonerated in the United States between 1989 and 2024 through descriptive statistics and statistical modeling. This document and our work are in progress.

Which offenses do we examine?

Table 1: List of high-frequency offenses, by race

Offense	**Overall**, N = 2,640	**Black**, N = 1,415	**White**, N = 785	**Hispanic**, N = 373	**Asian**, N =
Murder	875 (33%)	496 (35%)	226 (29%)	137 (37%)	4 (15%)
Drug Possession or Sale	601 (23%)	433 (31%)	85 (11%)	78 (21%)	3 (11%)
Child Sex Abuse	228 (8.6%)	60 (4.2%)	138 (18%)	24 (6.4%)	3 (11%)
Sexual Assault	195 (7.4%)	101 (7.1%)	73 (9.3%)	17 (4.6%)	0 (0%)
Robbery	142 (5.4%)	94 (6.6%)	23 (2.9%)	22 (5.9%)	0 (0%)
Assault	98 (3.7%)	40 (2.8%)	33 (4.2%)	23 (6.2%)	0 (0%)

Table 1 shows the highest-frequency criminal offenses across groups of race. For example, we read the top-left-most cell as reporting that, among the 2,640 individuals in the National Registry of Exonerations whose alleged crime occurred during or after 1989, 875 or 33% were exonerated from a Murder conviction. Percents are column-wise.

Table 2: Total years spent incarcerated and convicted, by offense

Offense	Incarcerated (17,086 total)	Convicted (24,318 total)
Murder	10,979 (64%)	12,071 (50%)
Child Sex Abuse	1,504 (9%)	2,077 (9%)
Sexual Assault	1,480 (9%)	1,865 (8%)
Robbery	789 (5%)	1,080 (4%)
Attempted Murder	606 (4%)	744 (3%)
Drug Possession or Sale	546 (3%)	4,018 (17%)
Assault	227 (1%)	403 (2%)
Manslaughter	207 (1%)	359 (1%)
Kidnapping	153 (1%)	184 (1%)
Weapon Possession or Sale	126 (1%)	264 (1%)

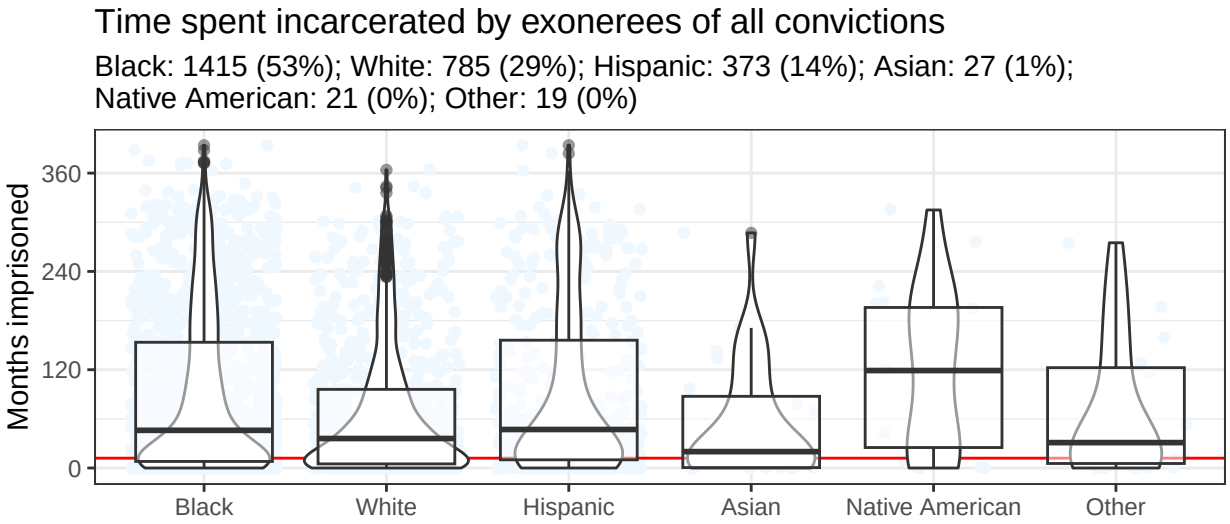
Table 2 shows the criminal offenses that account for the most time spent wrongfully incarcerated/convicted. For example, we read the top-left-most cell as reporting that, among the 17,086 years spent wrongfully incarcerated across all individuals in the National Registry of Exonerations whose alleged crime occurred during or after 1989, 10,979 or 64% were due to a Murder conviction. Again, percents are column-wise.

Table 2 was reproduced for each group of race (see Supplementary), which found that this order of offenses accounting for the most years lost to incarceration or conviction are similar across all subsets of races, especially the largest.

These tables motivate us to compare the experience of people convicted of these popular and severe offenses between groups of race, so as to measure the extent to which sentence/incarceration is associated with race, within the National Registry of Exonerations.

I. Distribution of time spent incarcerated and convicted, by offense and race

Overall



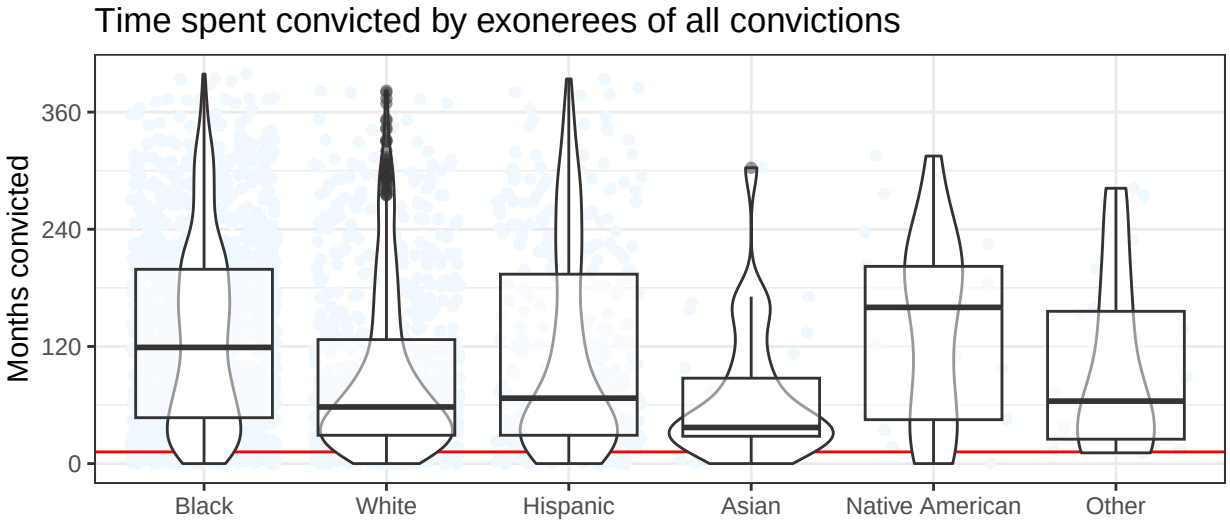


Table 3: Months spent convicted by exonerees of all convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	2,647	305,988	36	85	183	116	95
Black	1,415	186,239	47	119	199	132	95
White	785	70,687	29	58	127	90	84
Hispanic	373	41,805	29	67	194	112	105
Asian	27	1,831	28	37	88	68	68
Native American	21	2,902	45	160	202	138	95
Other	19	1,876	25	64	156	99	92
NA	7	648	30	88	113	93	87

Table 4: Months spent incarcerated by exonerees of all convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	2,647	216,965	8	42	133	82	93
Black	1,415	125,213	8	46	154	88	97
White	785	51,992	5	36	96	66	79
Hispanic	373	34,033	10	47	156	91	100
Asian	27	1,472	0	20	88	55	72
Native American	21	2,482	25	119	196	118	100
Other	19	1,335	6	31	122	70	87
NA	7	438	16	38	64	63	80

Murder

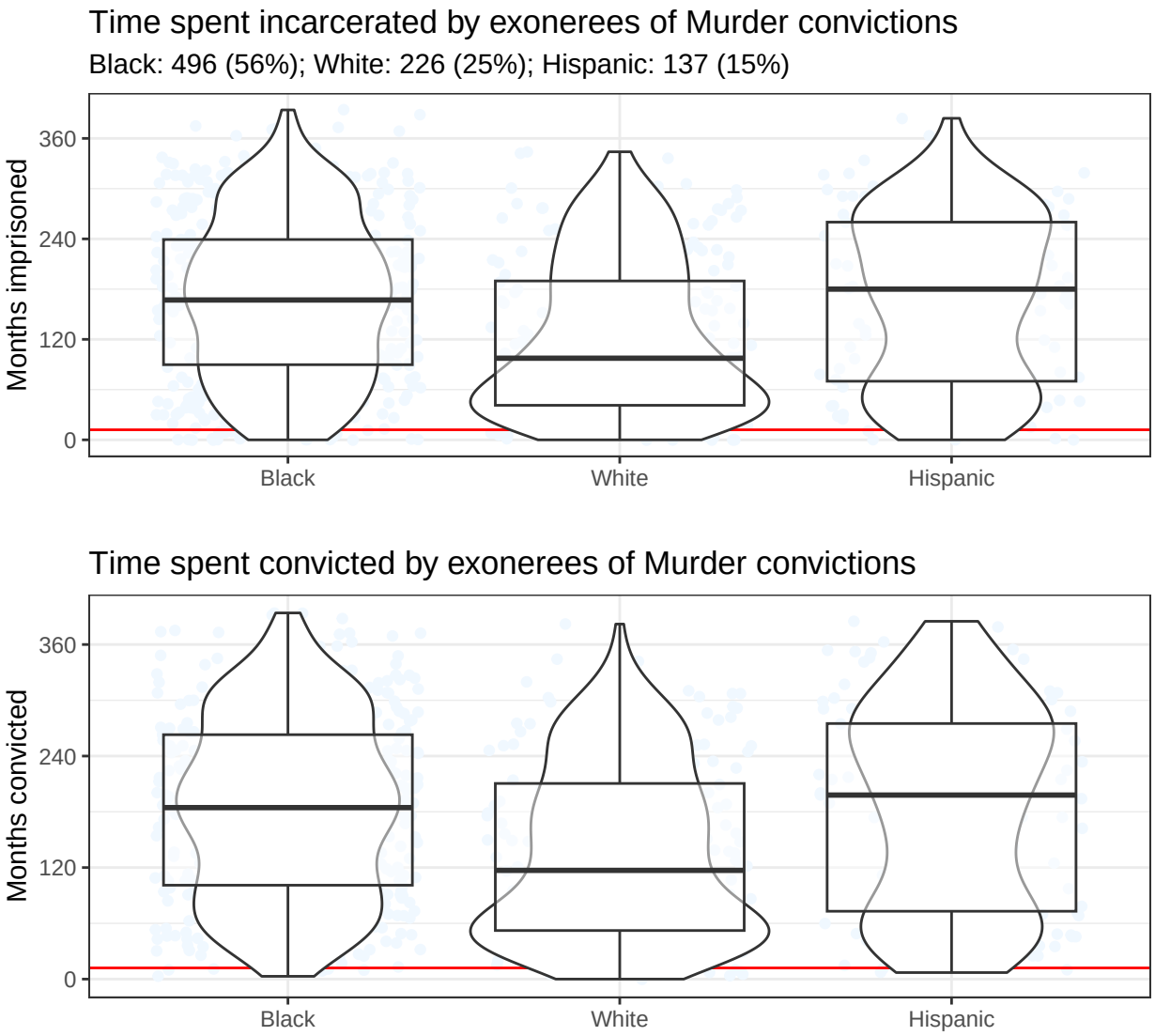


Table 5: Months spent convicted by exonerees of Murder convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	875	149,755	80	172	252	171	99
Black	496	90,875	101	184	263	183	96
White	226	30,545	52	117	210	135	92
Hispanic	137	25,302	73	198	275	185	109
Native American	8	1,811	196	204	238	226	45
Asian	4	580	90	104	160	145	107
Other	4	642	110	168	218	160	103

Table 6: Months spent incarcerated by exonerees of Murder convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	875	136,532	68	155	232	156	97
Black	496	83,243	90	167	239	168	95
White	226	27,472	41	98	190	122	92
Hispanic	137	22,832	70	180	260	167	102
Native American	8	1,796	196	202	236	224	46
Asian	4	558	90	102	151	140	100
Other	4	631	104	162	216	158	104

Child sex abuse

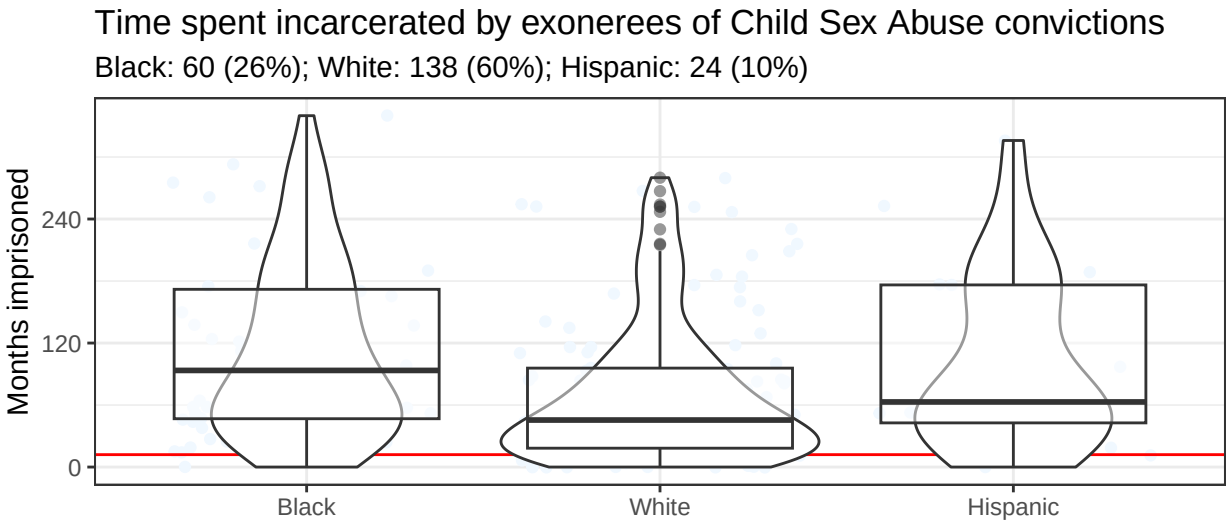


Table 7: Months spent convicted by exonerees of Child Sex Abuse convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	228	26,149	44	80	178	115	93
White	138	13,835	33	64	143	100	92
Black	60	8,538	62	134	203	142	92
Hispanic	24	3,253	53	98	225	136	93
Asian	3	144	34	43	59	48	25
Native American	2	334	132	167	202	167	99
Other	1	45	45	45	45	45	NA

Table 8: Months spent incarcerated by exonerees of Child Sex Abuse convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	228	19,261	26	56	120	84	77
White	138	9,707	18	46	96	70	69
Black	60	6,854	47	94	172	114	85
Hispanic	24	2,439	43	63	176	102	85
Asian	3	73	15	24	34	24	19
Native American	2	149	52	74	97	74	63
Other	1	39	39	39	39	39	NA

Sexual assault

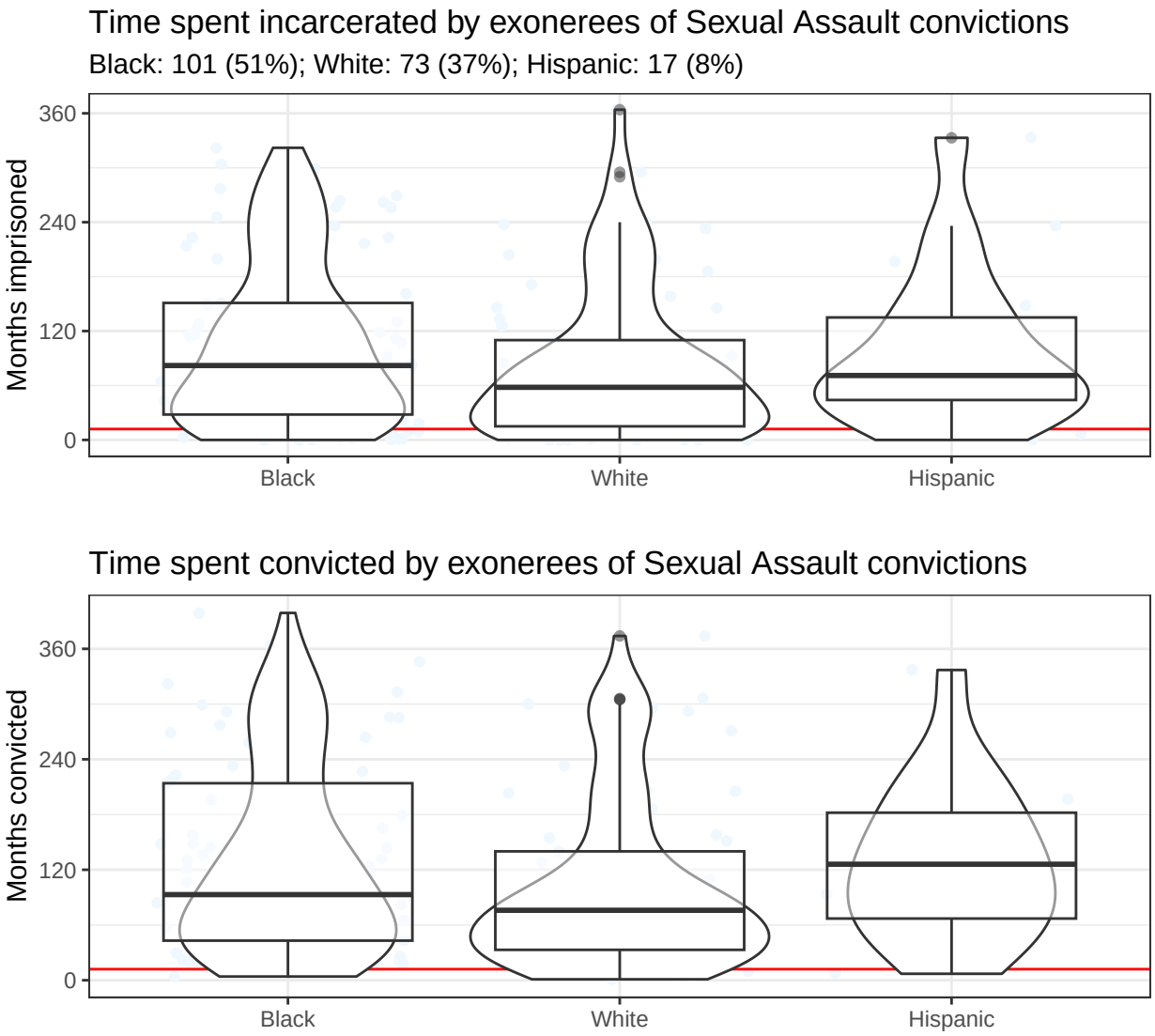


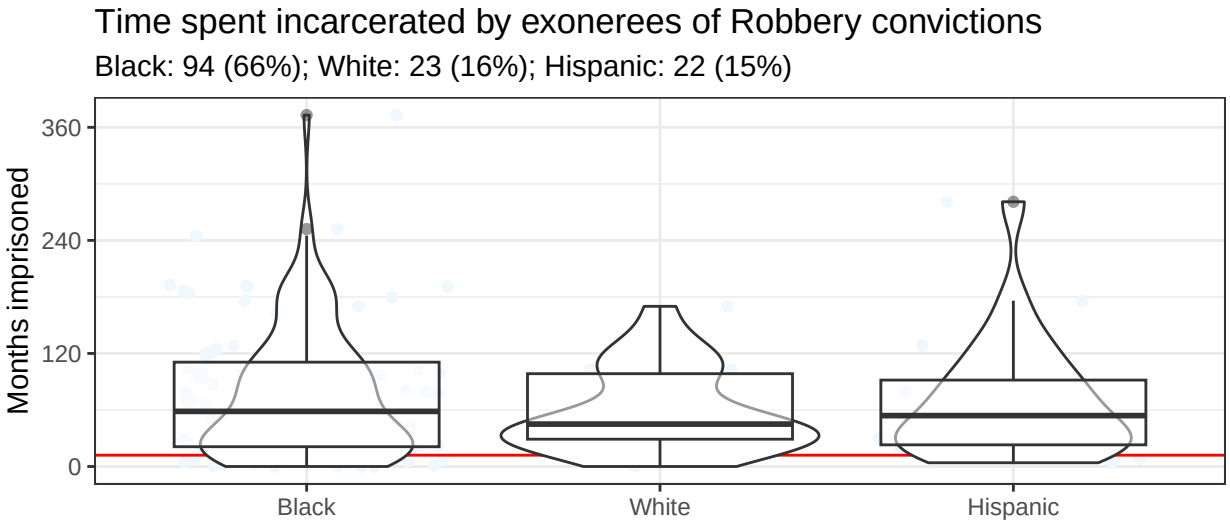
Table 9: Months spent convicted by exonerees of Sexual Assault convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	195	23,466	40	86	180	120	99
Black	101	13,263	43	93	214	131	104
White	73	7,426	33	76	140	102	93
Hispanic	17	2,205	67	126	182	130	85
Native American	2	270	122	135	148	135	35
Other	2	302	112	151	190	151	110

Table 10: Months spent incarcerated by exonerees of Sexual Assault convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	195	18,771	26	72	142	96	88
Black	101	10,613	28	82	151	105	90
White	73	6,002	15	58	110	82	84
Hispanic	17	1,691	44	71	135	99	89
Native American	2	182	77	91	105	91	40
Other	2	283	101	142	182	142	115

Robbery



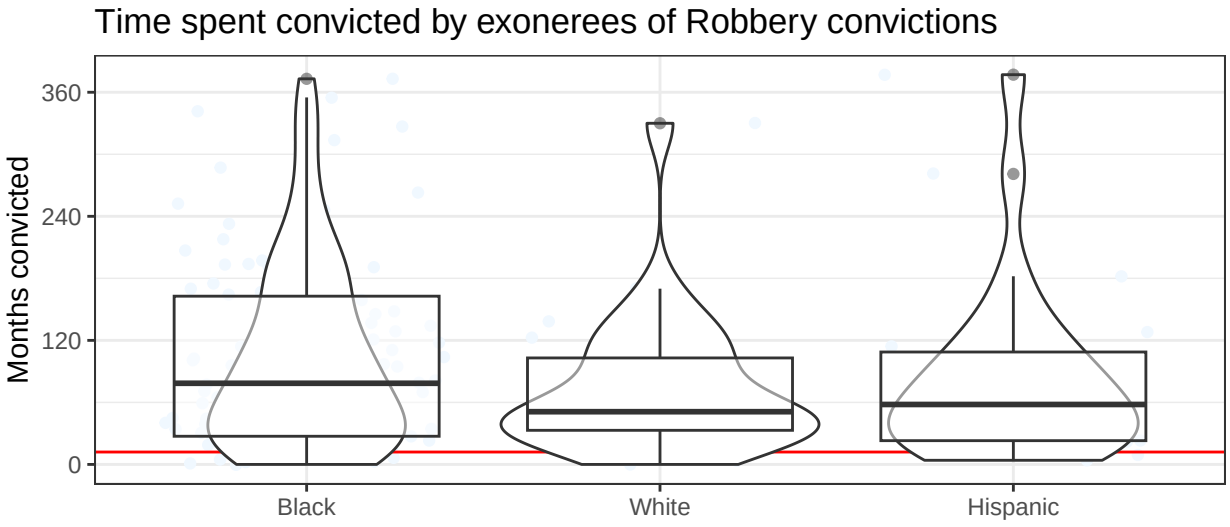


Table 11: Months spent convicted by exonerees of Robbery convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	142	13,748	28	72	136	97	88
Black	94	9,777	27	78	163	104	92
White	23	1,742	33	51	103	76	72
Hispanic	22	1,938	23	58	109	88	92
Other	2	170	48	85	122	85	105
Native American	1	121	121	121	121	121	NA

Table 12: Months spent incarcerated by exonerees of Robbery convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	142	10,248	23	50	104	72	66
Black	94	7,060	21	58	111	75	70
White	23	1,329	29	45	98	58	45
Hispanic	22	1,568	23	54	92	71	66
Other	2	170	48	85	122	85	105
Native American	1	121	121	121	121	121	NA

Attempted murder

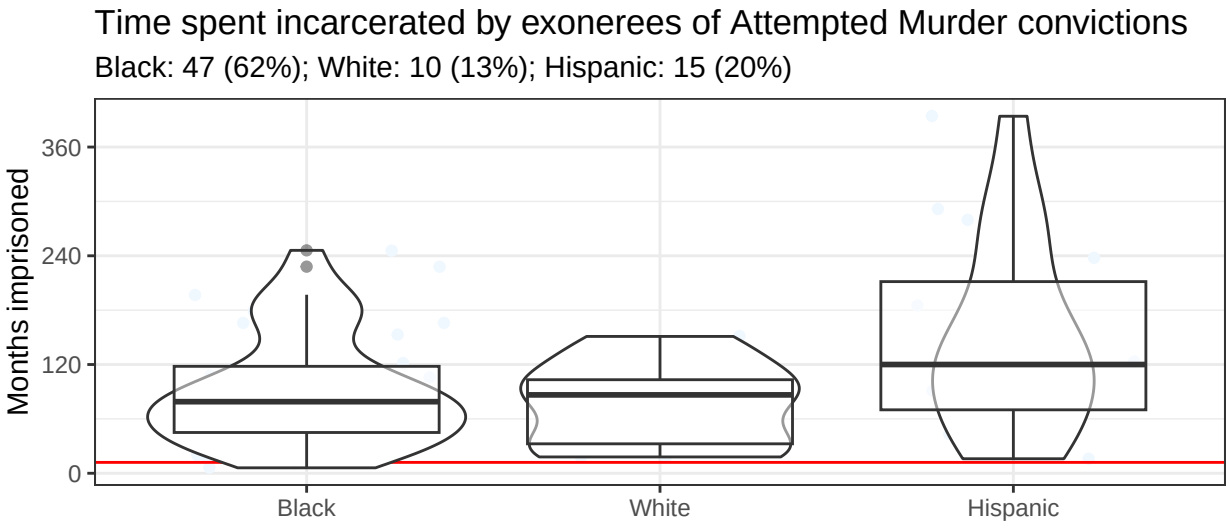


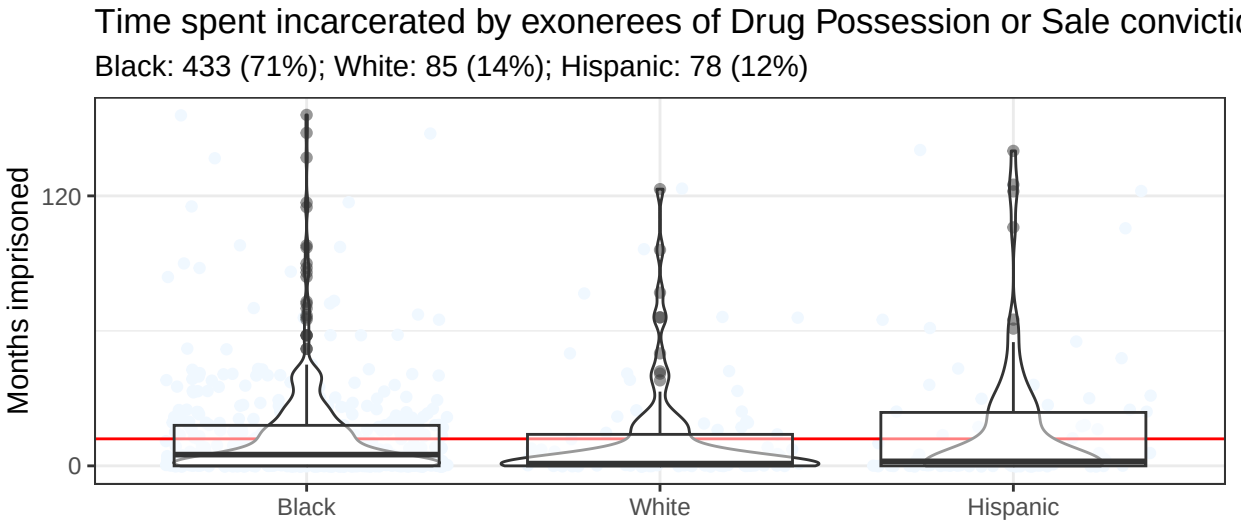
Table 13: Months spent convicted by exonerees of Attempted Murder convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	75	9,319	62	98	183	124	88
Black	47	5,118	60	86	143	109	77
Hispanic	15	2,721	120	185	241	181	100
White	10	1,003	32	90	113	100	104
Asian	2	292	140	146	152	146	17
Native American	1	185	185	185	185	185	NA

Table 14: Months spent incarcerated by exonerees of Attempted Murder convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	75	7,680	45	87	139	102	75
Black	47	4,249	45	79	118	90	61
Hispanic	15	2,238	70	120	212	149	109
White	10	758	32	86	103	76	48
Asian	2	277	136	138	141	138	6
Native American	1	158	158	158	158	158	NA

Drug possession or sale



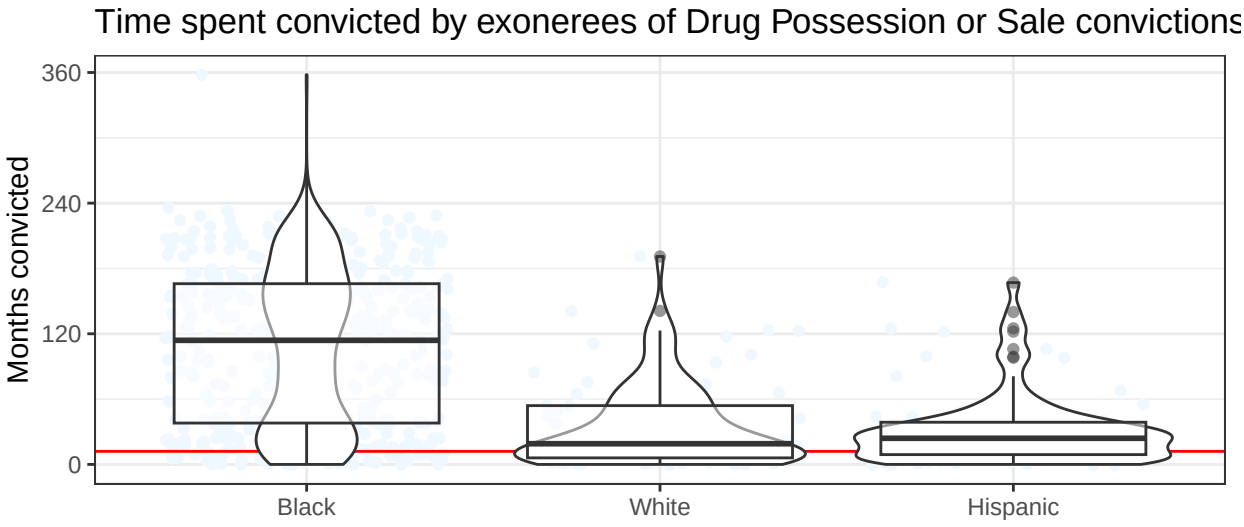


Table 15: Months spent convicted by exonerees of Drug Possession or Sale convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	604	51,247	19	65	149	85	72
Black	433	45,608	38	114	166	105	73
White	85	2,899	6	19	54	34	38
Hispanic	78	2,471	9	24	39	32	34
Asian	3	47	6	11	24	16	18
NA	3	203	54	88	91	68	41
Native American	1	0	0	0	0	0	NA
Other	1	19	19	19	19	19	NA

Table 16: Months spent incarcerated by exonerees of Drug Possession or Sale convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	604	8,410	0	5	18	14	23
Black	433	5,957	0	5	18	14	22
White	85	971	0	1	14	11	22
Hispanic	78	1,329	0	2	24	17	30
Asian	3	4	0	0	2	1	2
NA	3	149	38	56	64	50	26
Native American	1	0	0	0	0	0	NA
Other	1	0	0	0	0	0	NA

Assault

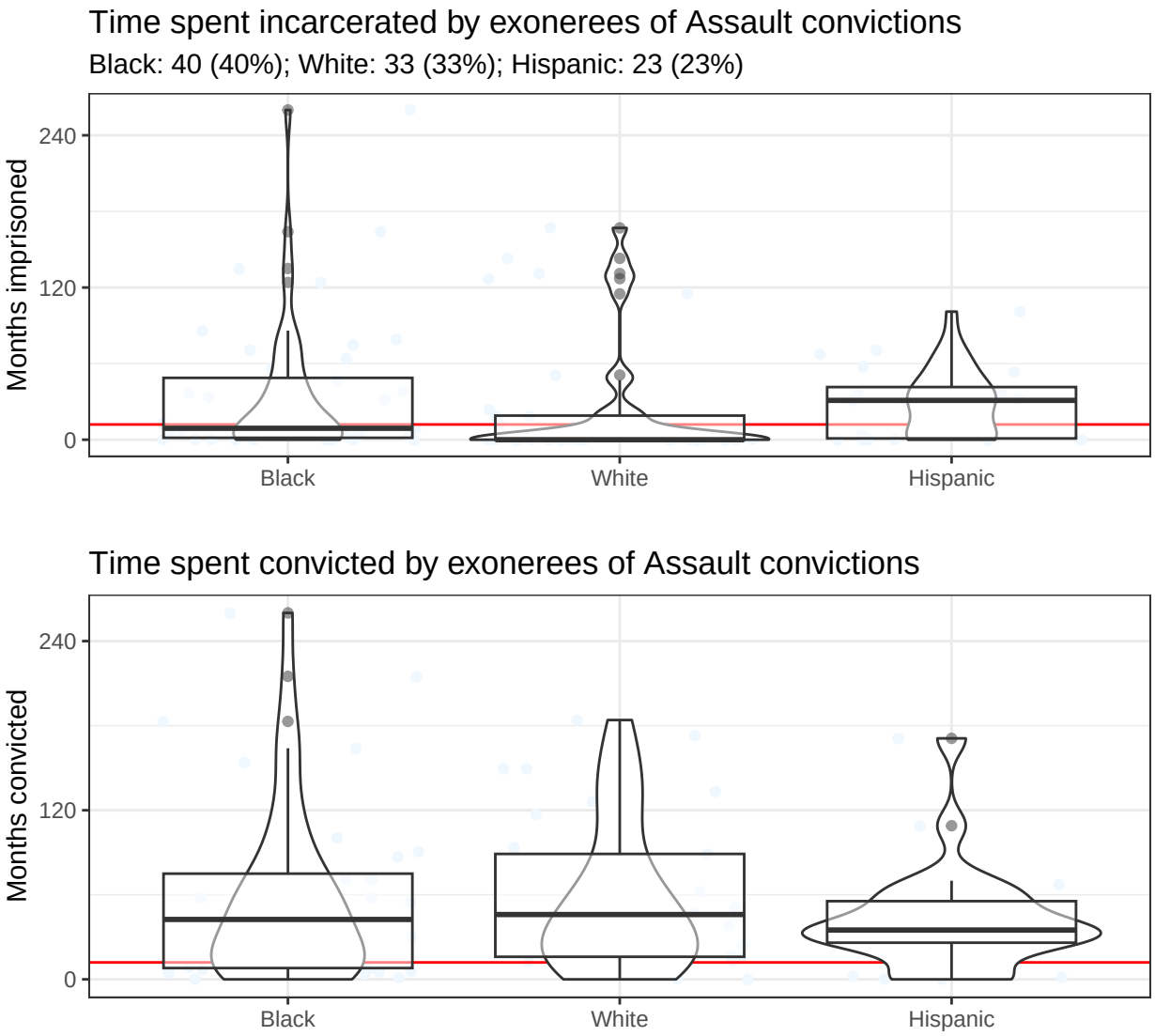


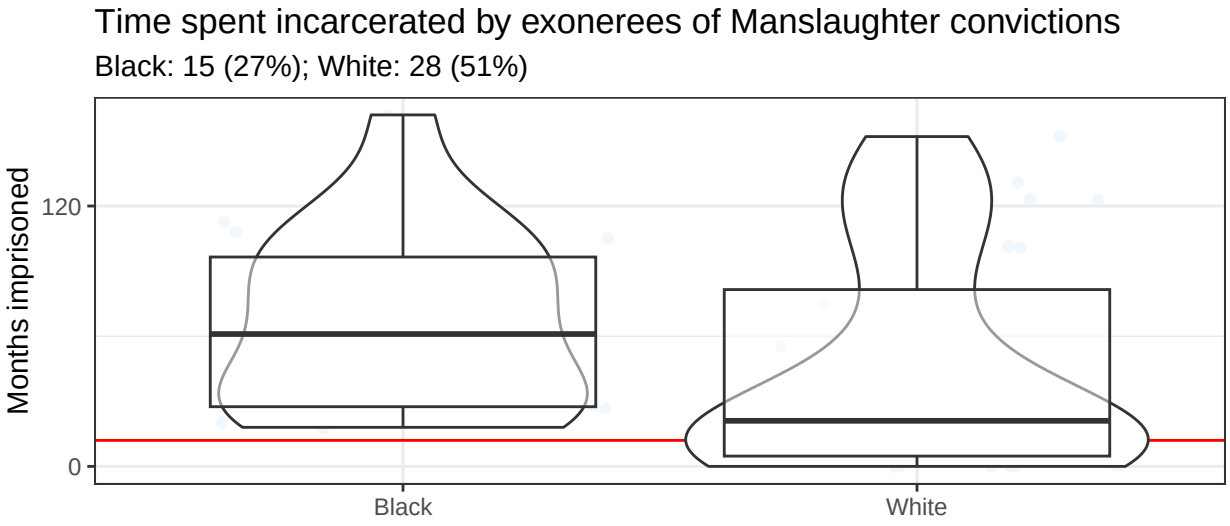
Table 17: Months spent convicted by exonerees of Assault convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	98	5,347	15	38	70	55	54
Black	40	2,351	8	42	75	59	63
White	33	1,918	16	46	89	58	53
Hispanic	23	989	26	35	56	43	38
Native American	1	25	25	25	25	25	NA
Other	1	64	64	64	64	64	NA

Table 18: Months spent incarcerated by exonerees of Assault convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	98	3,074	0	8	44	31	47
Black	40	1,462	2	9	49	37	55
White	33	866	0	0	19	26	50
Hispanic	23	668	1	31	42	29	28
Native American	1	25	25	25	25	25	NA
Other	1	53	53	53	53	53	NA

Manslaughter



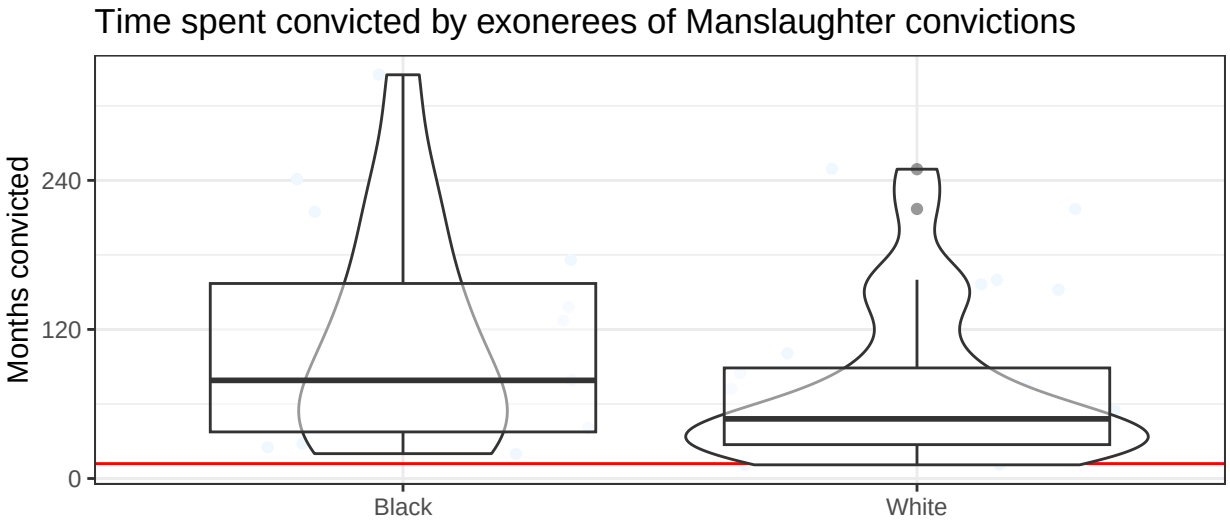


Table 19: Months spent convicted by exonerees of Manslaughter convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	54	4,592	32	61	120	85	75
White	28	1,976	27	48	89	71	64
Black	15	1,685	38	79	157	112	92
Hispanic	8	571	67	67	74	71	30
Asian	1	33	33	33	33	33	NA
Native American	1	45	45	45	45	45	NA
Other	1	282	282	282	282	282	NA

Table 20: Months spent incarcerated by exonerees of Manslaughter convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	54	2,750	17	30	83	51	46
White	28	1,250	5	21	82	45	51
Black	15	1,003	28	61	96	67	43
Hispanic	8	464	54	67	67	58	27
Asian	1	33	33	33	33	33	NA
Native American	1	0	0	0	0	0	NA
Other	1	0	0	0	0	0	NA

II. Distribution of age at events, by offense and race

Supplementary

i. Which offenses account for the most time spent wrongfully incarcerated and convicted for each race?

Table 2 has been recreated to show how the criminal offenses that account for the most time spent wrongfully incarcerated/convicted compare between groups of race.

Table 21: Total years spent incarcerated and convicted by Black exonerees

Offense	Incarcerated (9,887 total)	Convicted (14,887 total)
Murder	6,708 (68%)	7,348 (49%)
Sexual Assault	837 (8%)	1,055 (7%)
Robbery	549 (6%)	773 (5%)
Child Sex Abuse	543 (5%)	680 (5%)
Drug Possession or Sale	377 (4%)	3,620 (24%)

Table 22: Total years spent incarcerated and convicted by White exonerees

Offense	Incarcerated (4,054 total)	Convicted (5,543 total)
Murder	2,190 (54%)	2,434 (44%)
Child Sex Abuse	750 (19%)	1,097 (20%)
Sexual Assault	473 (12%)	586 (11%)
Robbery	99 (2%)	135 (2%)
Manslaughter	93 (2%)	153 (3%)

Table 23: Total years spent incarcerated and convicted by Hispanic exonerees

Offense	Incarcerated (2,694 total)	Convicted (3,314 total)
Murder	1,841 (68%)	2,043 (62%)
Child Sex Abuse	192 (7%)	259 (8%)
Attempted Murder	180 (7%)	221 (7%)
Sexual Assault	133 (5%)	177 (5%)
Robbery	118 (4%)	149 (4%)

Table 24: Total years spent incarcerated and convicted by Asian exonerees

Offense	Incarcerated (113 total)	Convicted (141 total)
Murder	44 (39%)	47 (33%)
Supporting Terrorism	27 (24%)	27 (19%)
Attempted Murder	22 (19%)	24 (17%)
Attempt, Violent	6 (5%)	6 (4%)
Child Sex Abuse	5 (4%)	11 (8%)

Table 25: Total years spent incarcerated and convicted by Native American exonerees

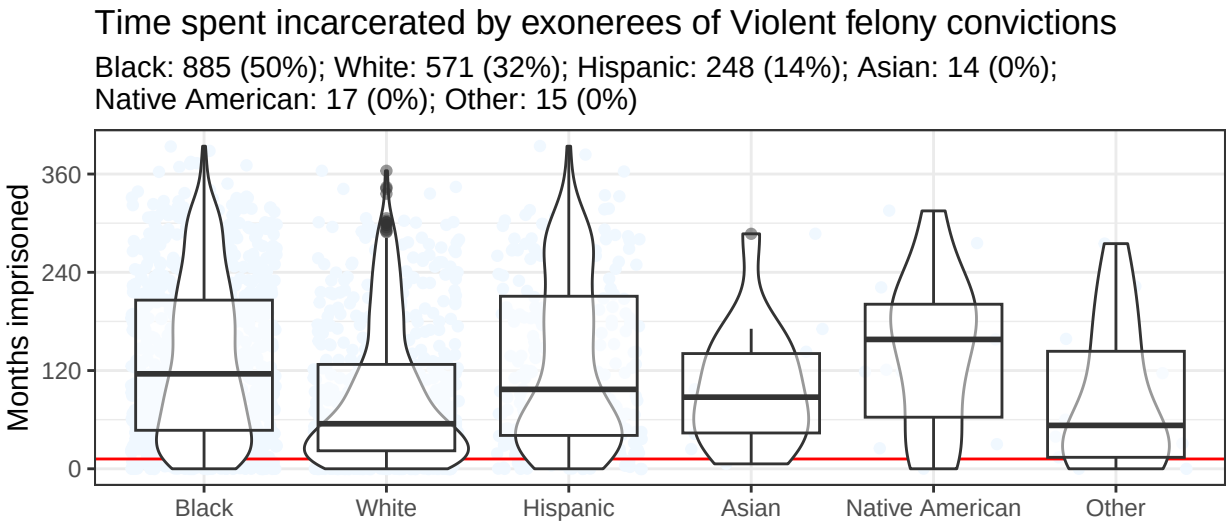
Offense	Incarcerated (199 total)	Convicted (234 total)
Murder	146 (73%)	148 (63%)
Sexual Assault	14 (7%)	22 (9%)
Attempted Murder	13 (7%)	15 (6%)
Child Sex Abuse	11 (6%)	27 (12%)
Robbery	10 (5%)	10 (4%)

Table 26: Total years spent incarcerated and convicted by Other exonerees

Offense	Incarcerated (105 total)	Convicted (148 total)
Murder	50 (48%)	51 (34%)
Sexual Assault	23 (22%)	25 (17%)
Robbery	13 (12%)	13 (9%)
Kidnapping	9 (9%)	12 (8%)
Assault	4 (4%)	5 (3%)

ii. How does the distribution of time spent incarcerated/convicted for a particular class of offense compare between races?

Violent felony



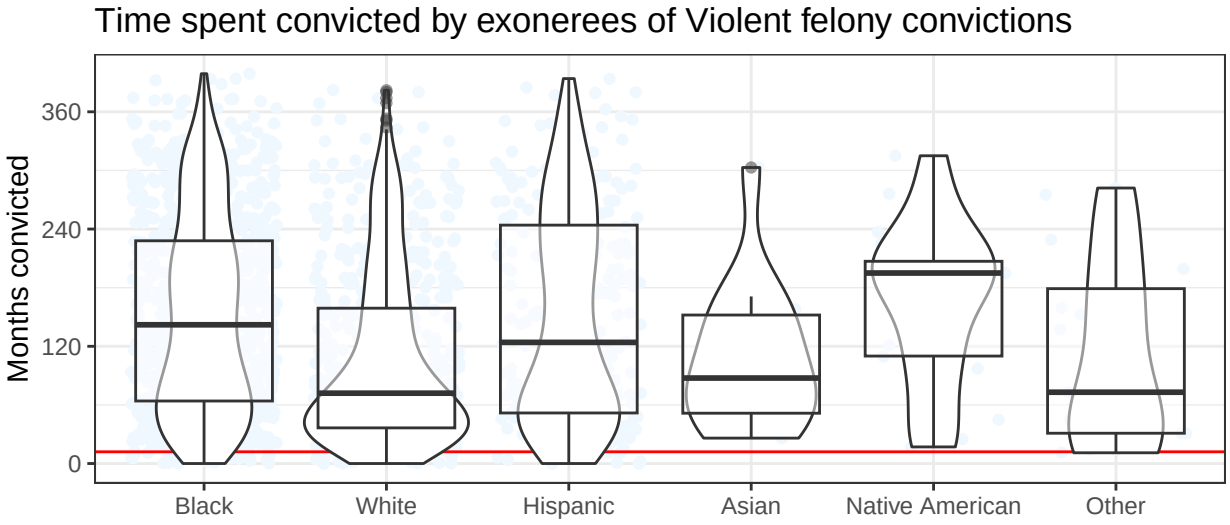


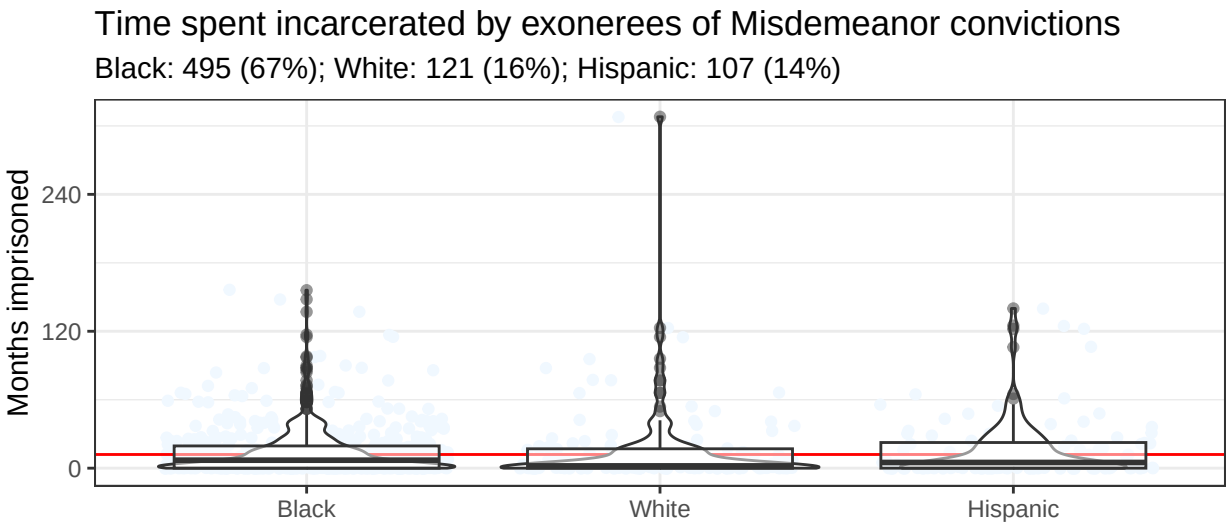
Table 27: Months spent convicted by exonerees of Violent felony convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	1,751	239,790	51	112	214	137	101
Black	885	135,593	64	142	228	153	101
White	571	60,824	36	72	159	107	90
Hispanic	248	37,207	52	124	244	150	108
Native American	17	2,808	110	195	207	165	85
Other	15	1,717	31	73	179	114	97
Asian	14	1,509	52	88	152	108	75
NA	1	132	132	132	132	132	NA

Table 28: Months spent incarcerated by exonerees of Violent felony convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	1,751	203,268	34	90	187	116	96
Black	885	117,425	47	116	206	133	97
White	571	48,763	22	55	128	85	83
Hispanic	248	31,921	41	97	211	129	102
Native American	17	2,439	63	158	201	143	94
Other	15	1,321	14	53	144	88	90
Asian	14	1,395	44	88	141	100	75
NA	1	4	4	4	4	4	NA

Misdemeanor



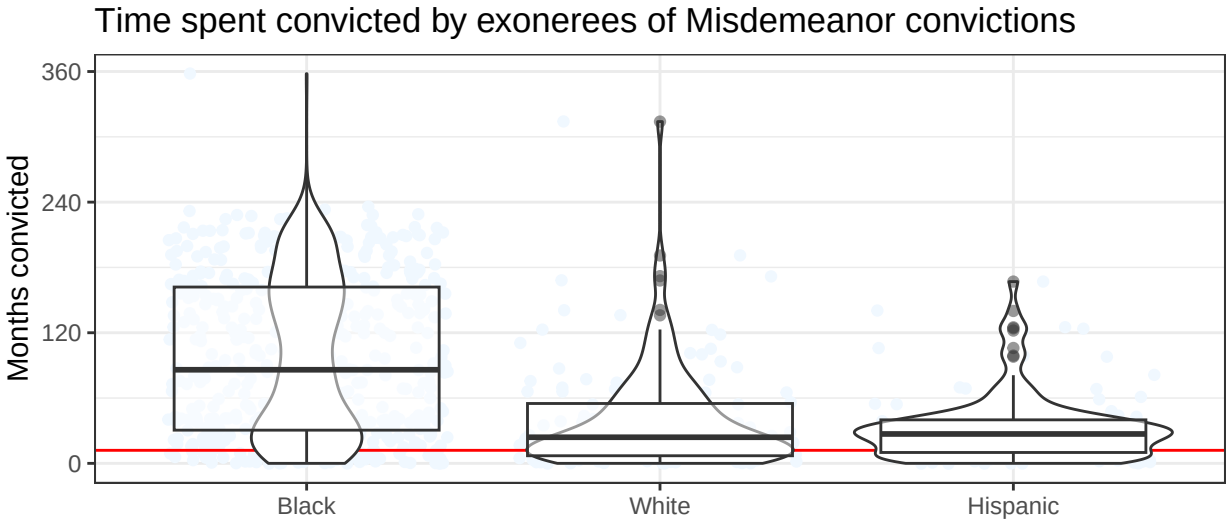


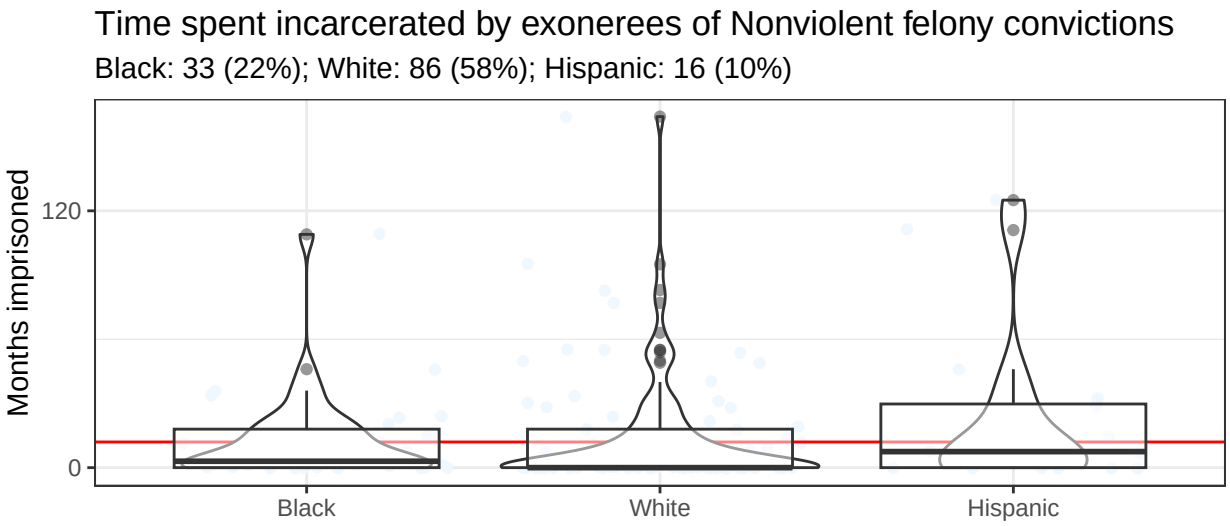
Table 29: Months spent convicted by exonerees of Misdemeanor convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	735	57,582	18	48	138	78	70
Black	495	48,606	30	86	162	98	73
White	121	4,893	7	24	55	40	48
Hispanic	107	3,465	10	27	40	32	32
Asian	5	95	11	11	36	19	17
NA	5	504	38	88	94	101	96
Native American	1	0	0	0	0	0	NA
Other	1	19	19	19	19	19	NA

Table 30: Months spent incarcerated by exonerees of Misdemeanor convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	735	11,389	0	5	19	15	27
Black	495	7,325	0	7	20	15	22
White	121	1,909	0	2	17	16	36
Hispanic	107	1,728	0	5	22	16	27
Asian	5	5	0	0	1	1	2
NA	5	422	38	56	72	84	86
Native American	1	0	0	0	0	0	NA
Other	1	0	0	0	0	0	NA

Nonviolent felony



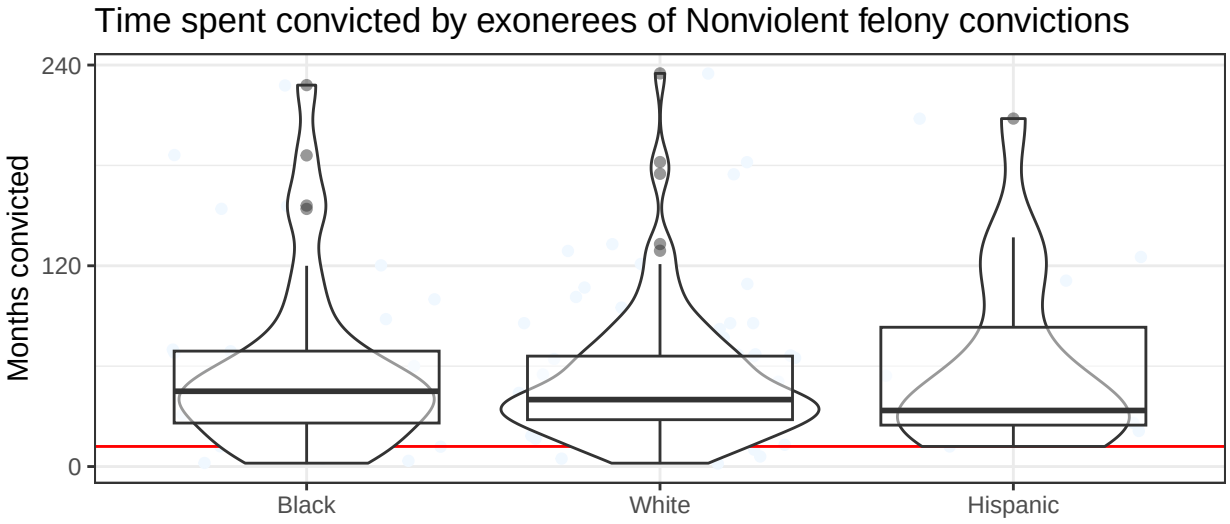


Table 31: Months spent convicted by exonerees of Nonviolent felony convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	148	7,883	26	40	66	53	45
White	86	4,496	28	40	66	52	40
Black	33	1,968	26	45	69	60	54
Hispanic	16	989	25	34	83	62	55
Asian	8	227	25	30	34	28	12
Other	3	140	26	37	63	47	38
Native American	1	51	51	51	51	51	NA
NA	1	12	12	12	12	12	NA

Table 32: Months spent incarcerated by exonerees of Nonviolent felony convictions

Race	N	Sum	Quartile25	Quartile50	Quartile75	Mean	SD
All	148	2,169	0	2	18	15	26
White	86	1,218	0	0	18	14	26
Black	33	426	0	3	18	13	21
Hispanic	16	384	0	8	30	24	39
Asian	8	72	0	8	18	9	10
Other	3	14	0	0	7	5	8
Native American	1	43	43	43	43	43	NA
NA	1	12	12	12	12	12	NA

Misconduct

This class of offense had zero months of times spent incarcerated or convicted. This doesn't necessarily mean zero time, only zero complete months, because our current variable for time is measured in months, as an integer.

End of report.

Code Appendix

```
### ------------------------------------------------------------
### Setup
knitr::opts_chunk$set(echo = FALSE,
                      message = FALSE,
                      warning = FALSE)
knitr::opts_knit$set(root.dir = paste(head(unlist(strsplit(getwd(), "/")), -1),
                                      collapse = "/"))
# options(knitr.kable.NA = '-')
labs = knitr::all_labels()
labs = labs[!labs %in% c("setup", "allcode")]

# import required libraries
library(tidyverse)
library(gridExtra)
library(gtsummary)
library(gt)
library(knitr)
library(kableExtra)
### ------------------------------------------------------------
### Data Loading

# clear workspace
rm(list = ls())

# load data
exonerees <- read.csv("data/exonerees1989-may2024.csv")

# correct Crime Category to be a factor
crime_levels <- c("Violent felony", "Misdemeanor", "Nonviolent felony",
                  "Misconduct")
exonerees$Crime.Category <- factor(exonerees$Crime.Category,
                                   levels = crime_levels)

# correct Age Group variables to be factors
age_levels = c("0 - 17", "18 - 29", "30 - 39", "40 - 49", "50 - 64", "65 - 100")
exonerees$Age.Group.at.Crime <- factor(exonerees$Age.Group.at.Crime,
                                       levels = age_levels)
exonerees$Age.Group.at.Conviction <- factor(exonerees$Age.Group.at.Conviction,
                                       levels = age_levels)
exonerees$Age.Group.at.Release <- factor(exonerees$Age.Group.at.Release,
                                       levels = age_levels)
exonerees$Age.Group.at.Exoneration <- factor(exonerees$Age.Group.at.Exoneration,
                                       levels = age_levels)

# correct Race to be a factor
race_levels <- c("Black", "White", "Hispanic", "Asian", "Native American",
                 "Other")
exonerees$Race <- factor(exonerees$Race, levels = race_levels)

# correct Sex to be a factor
exonerees$Sex <- as.factor(exonerees$Sex)
```

```r
# correct binary tags to also be factors
tag_vars <- c("F.MFE","FC","ILD","P.FA","DNA","MWID","OM","CV","IO","SA","CIU",
              "NC","P","H","JV","CDC","F","JI","M","CSH","SBS","A","FED","PH",
              "BM")
exonerees <- exonerees %>% mutate_at(tag_vars, as.factor)
# make sure everything looks good
summary(exonerees)
dim(exonerees)
### ------------------------------------------------------------
### Ordered Sample Size Tables
exonerees %>%
  select(Race, Worst.Crime.Display) %>%
  filter(!is.na(Race)) %>%
  # ordered frequency of offenses, overall and by race
  tbl_summary(
    by = Race,
    statistic = list(
      all_categorical() ~ "{n} ({p}%)")) %>%   # count (percent)
  add_overall() %>%
  bold_labels() %>%
  modify_header(label ~ "Offense") %>%
  modify_spanning_header(c(stat_1, stat_2) ~ "Race") %>%
  # convert to a data frame and arrange by a new overall count column
  as.data.frame() %>%
  mutate(overall.count = parse_number(`**Overall**, N = 2,640`)) %>%
  arrange(desc(overall.count)) %>%
  select(-overall.count) %>%  # remove the overall count column
  # only present the 6 most frequent offenses
  head() %>%
  # table styling
  knitr::kable(
    format = "simple",
    caption = "List of high-frequency offenses, by race")
### ------------------------------------------------------------
### Time Served Table

# write column headers now, for later use
header_Incarcerated <- paste0(
  "Incarcerated (",
  formatC(sum(exonerees$Years.Incarcerated), big.mark = ","),
  " total)")
header_Convicted <- paste0(
  "Convicted (",
  formatC(sum(exonerees$Years.Convicted), big.mark = ","),
  " total)")

# time served incarcerated and convicted (ordered), by offense
exonerees %>%
  group_by(Offense = Worst.Crime.Display) %>%
  # summarize total years
  summarise(Incarcerated = sum(Years.Incarcerated),
            Convicted = sum(Years.Convicted)) %>%
  # calculate percents
```

```
  mutate(Incarcerated.perc = round(100*Incarcerated / sum(Incarcerated)),
         Convicted.perc = round(100*Convicted / sum(Convicted))) %>%
  # arrange and filter out offenses that collectively have not served 90+ years
  arrange(desc(Incarcerated)) %>%
  filter(Incarcerated > 90) %>%
  # # reformat count to character
  mutate(Incarcerated = formatC(Incarcerated, big.mark = ","),
         Convicted = formatC(Convicted, big.mark = ",")) %>%
  # combine count and percent
  mutate(Incarcerated = paste0(Incarcerated, " (", Incarcerated.perc, "%)"),
         Convicted = paste0(Convicted, " (", Convicted.perc, "%)")) %>%
  # remove unused columns
  select(-contains(".perc")) %>%
  # table styling
  knitr::kable(
    format = "simple",
    # rename header to include total time across all offenses
    col.names = c("Offense", header_Incarcerated, header_Convicted),
    caption = "Total years spent incarcerated and convicted, by offense")
### ------------------------------------------------------------
### Distribution of time spent incarcerated and convicted, by offense and race
## crime, by total time spent incarcerated
plot_times <- function (data, crime_filter = NA) {

  ## filter the data by crime (optional)
  if (is.na(crime_filter)) {
    crime_filter <- "all" # this is used later for labeling
  } else {
    data <- data %>%
      filter(Worst.Crime.Display == crime_filter) %>%
      filter(!is.na(Months.Incarcerated) & !is.na(Months.Convicted))
  }

  ## only plot groups of race with 10 or more members present
  # get counts
  counts <- data %>%
    group_by(Race) %>%
    summarise(n = n()) %>%
    mutate(prop = n / sum(n)) %>%
    filter(n >= 10)
  # filter race groups with too few members
  plot_data <- data %>% filter(Race %in% counts$Race)

  ## write the subtitle of the plot, which shows counts of race groups
  gg_subtitle <- ""
  included_races <- unique(plot_data$Race)
  for (i in 1:length(included_races)) {
    race <- race_levels[i]
    count <- nrow(plot_data %>% filter(Race == race))
    gg_subtitle <- paste0(
      gg_subtitle, # previous groups
      # title, count, and percent (from data with all race groups)
      # ensure percent is over 'data' not the race filtered 'plot_data'
```

```
    race, ": ", count, " (", floor(100*count/nrow(data)), "%)")
  # add a semi-colon if the list isn't done
  if(i < length(included_races))
    gg_subtitle <- paste0(gg_subtitle, "; ")
  # add a new line after the fourth race group
  if (i == 4)
    gg_subtitle <- paste0(gg_subtitle, "\n")
  }

## plot distribution of time spent incarcerated
gg_incarated <- plot_data %>%
  ggplot(aes(y = Months.Incarcerated, x = Race)) +
  # jitter points (these must be overlayed by the horizontal line)
  geom_jitter(width = 0.4, color = "aliceblue") +
  # marker at 1 year
  geom_hline(yintercept = 12, color = "red") +
  # time spent wrongfully imprisoned
  geom_violin() +
  geom_boxplot(alpha = 0.5) +
  # plot specifics
  scale_y_continuous(breaks = seq(0, 400, 120)) +
  theme_bw() +
  labs(
    x = "",
    y = "Months imprisoned",
    title = paste("Time spent incarcerated by exonerees of",
                  crime_filter, "convictions"),
    subtitle = gg_subtitle
    )

## plot distribution of time spent convicted
gg_convicted <- plot_data %>%
  ggplot(aes(y = Months.Convicted, x = Race)) +
  # jitter points (these must be overlayed by the horizontal line)
  geom_jitter(width = 0.4, color = "aliceblue") +
  # marker at 1 year
  geom_hline(yintercept = 12, color = "red") +
  # time spent wrongfully convicted
  geom_violin() +
  geom_boxplot(alpha = 0.5) +
  # plot specifics
  scale_y_continuous(breaks = seq(0, 400, 120)) +
  theme_bw() +
  labs(
    x = "",
    y = "Months convicted",
    title = paste("Time spent convicted by exonerees of",
                  crime_filter, "convictions"),
    subtitle = gg_subtitle
    )

## tabular distribution of time spent convicted
tbl_convicted <- data %>%
```

```r
  # calculate summary statistics by race
  group_by(across(Race)) %>%
  summarize(
    N = n(),
    Sum = sum(Months.Convicted),
    Quartile25 = quantile(Months.Convicted, 0.25, na.rm = TRUE),
    Quartile50 = median(Months.Convicted, na.rm = TRUE),
    Quartile75 = quantile(Months.Convicted, 0.75, na.rm = TRUE),
    Mean = mean(Months.Convicted, na.rm = TRUE),
    SD = sd(Months.Convicted, na.rm = TRUE)
  ) %>%
  # append a last row for summary statistics, regardless of race
  add_row(
    data %>%
      summarize(
        Race = "All",
        N = n(),
        Sum = sum(Months.Convicted, na.rm = TRUE),
        Quartile25 = quantile(Months.Convicted, 0.25, na.rm = TRUE),
        Quartile50 = median(Months.Convicted, na.rm = TRUE),
        Quartile75 = quantile(Months.Convicted, 0.75, na.rm = TRUE),
        Mean = mean(Months.Convicted, na.rm = TRUE),
        SD = sd(Months.Convicted, na.rm = TRUE)
        )
    ) %>%
  # arrange by sample size (beginning with overall)
  arrange(desc(N)) %>%
  # table styling
  knitr::kable(
    format = "latex",
    booktabs = TRUE,
    format.args = list(big.mark = ","),
    digits = 0,
    caption = paste("Months spent convicted by exonerees of", crime_filter,
                    "convictions")
    ) %>%
  # bold overall row
  kableExtra::row_spec(1, background = "#D3D3D3") %>%
  kableExtra::kable_styling(latex_options = "hold_position")


## tabular distribution of time spent incarcerated
tbl_incarcerated <- data %>%
  # calculate summary statistics by race
  group_by(across(Race)) %>%
  summarize(
    N = n(),
    Sum = sum(Months.Incarcerated),
    Quartile25 = quantile(Months.Incarcerated, 0.25, na.rm = TRUE),
    Quartile50 = median(Months.Incarcerated, na.rm = TRUE),
    Quartile75 = quantile(Months.Incarcerated, 0.75, na.rm = TRUE),
    Mean = mean(Months.Incarcerated, na.rm = TRUE),
    SD = sd(Months.Incarcerated, na.rm = TRUE)
```

```
    ) %>%
    # append a last row for summary statistics, regardless of race
    add_row(
      data %>%
        summarize(
          Race = "All",
          N = n(),
          Sum = sum(Months.Incarcerated, na.rm = TRUE),
          Quartile25 = quantile(Months.Incarcerated, 0.25, na.rm = TRUE),
          Quartile50 = median(Months.Incarcerated, na.rm = TRUE),
          Quartile75 = quantile(Months.Incarcerated, 0.75, na.rm = TRUE),
          Mean = mean(Months.Incarcerated, na.rm = TRUE),
          SD = sd(Months.Incarcerated, na.rm = TRUE)
          )
      ) %>%
    # arrange by sample size (beginning with overall)
    arrange(desc(N)) %>%
    # table styling
    knitr::kable(
      format = "latex",
      booktabs = TRUE,
      format.args = list(big.mark = ","),
      digits = 0,
      caption = paste("Months spent incarcerated by exonerees of", crime_filter,
                      "convictions")
      ) %>%
    # bold overall row
    kableExtra::row_spec(1, background = "#D3D3D3") %>%
    kableExtra::kable_styling(latex_options = "hold_position")


  ## return plots and tables
  return(list(
    plot_conviction = gg_convicted,
    plot_incarceration = gg_incarated,
    table_convicted = tbl_convicted,
    table_incarcerated = tbl_incarcerated)
    )
  }
### ------------------------------------------------------------
## Overall stats
overall_stats <- plot_times(exonerees)
# plots
overall_stats <- plot_times(exonerees)
gridExtra::grid.arrange(overall_stats$plot_incarceration,
                        overall_stats$plot_conviction + labs(subtitle = NULL))
# tables
overall_stats$table_convicted
overall_stats$table_incarcerated
### ------------------------------------------------------------
## Murder stats
mdr_stats <- plot_times(exonerees, crime_filter = "Murder")
# plots
```

```r
gridExtra::grid.arrange(mdr_stats$plot_incarceration,
                        mdr_stats$plot_conviction + labs(subtitle = NULL))
# tables
mdr_stats$table_convicted
mdr_stats$table_incarcerated
### ------------------------------------------------------------
# Child sex abuse
csa_stats <- plot_times(exonerees, crime_filter = "Child Sex Abuse")
# plots
gridExtra::grid.arrange(csa_stats$plot_incarceration,
                        csa_stats$plot_conviction + labs(subtitle = NULL))
# tables
csa_stats$table_convicted
csa_stats$table_incarcerated
### ------------------------------------------------------------
# Sexual assault
sxa_stats <- plot_times(exonerees, crime_filter = "Sexual Assault")
# plots
gridExtra::grid.arrange(sxa_stats$plot_incarceration,
                        sxa_stats$plot_conviction + labs(subtitle = NULL))
# tables
sxa_stats$table_convicted
sxa_stats$table_incarcerated
### ------------------------------------------------------------
# Robbery
rby_stats <- plot_times(exonerees, crime_filter = "Robbery")
# plots
gridExtra::grid.arrange(rby_stats$plot_incarceration,
                        rby_stats$plot_conviction + labs(subtitle = NULL))
# tables
rby_stats$table_convicted
rby_stats$table_incarcerated
### ------------------------------------------------------------
# Attempted murder
atm_stats <- plot_times(exonerees, crime_filter = "Attempted Murder")
# plots
gridExtra::grid.arrange(atm_stats$plot_incarceration,
                        atm_stats$plot_conviction + labs(subtitle = NULL))
# tables
atm_stats$table_convicted
atm_stats$table_incarcerated
### ------------------------------------------------------------
# Drug possession or sale
dps_stats <- plot_times(exonerees, crime_filter = "Drug Possession or Sale")
# plots
gridExtra::grid.arrange(dps_stats$plot_incarceration,
                        dps_stats$plot_conviction + labs(subtitle = NULL))
# tables
dps_stats$table_convicted
dps_stats$table_incarcerated
### ------------------------------------------------------------
# Assault
asa_stats <- plot_times(exonerees, crime_filter = "Assault")
```

```r
# plots
gridExtra::grid.arrange(asa_stats$plot_incarceration,
                        asa_stats$plot_conviction + labs(subtitle = NULL))
# tables
asa_stats$table_convicted
asa_stats$table_incarcerated
### ------------------------------------------------------------
# Manslaughter
msl_stats <- plot_times(exonerees, crime_filter = "Manslaughter")
# plots
gridExtra::grid.arrange(msl_stats$plot_incarceration,
                        msl_stats$plot_conviction + labs(subtitle = NULL))
# tables
msl_stats$table_convicted
msl_stats$table_incarcerated
### ------------------------------------------------------------
### Supplementary Time Served Table

time_served_table <- function (race_filter) {
  # filter exonerees by race
  data <- exonerees %>% filter(Race == race_filter)

  # write column headers now, for later use
  header_Incarcerated <- paste0("Incarcerated (",
                         formatC(sum(data$Years.Incarcerated), big.mark = ","),
                         " total)")
  header_Convicted <- paste0("Convicted (",
                         formatC(sum(data$Years.Convicted), big.mark = ","),
                        " total)")

  # time served incarcerated and convicted (ordered), by offense
  table <- data %>%
    group_by(Offense = Worst.Crime.Display) %>%
    # summarize total years
    summarise(Incarcerated = sum(Years.Incarcerated),
              Convicted = sum(Years.Convicted)) %>%
    # calculate percents
    mutate(Incarcerated.perc = round(100*Incarcerated / sum(Incarcerated)),
           Convicted.perc = round(100*Convicted / sum(Convicted))) %>%
    # arrange by total years spent incarcerated
    arrange(desc(Incarcerated)) %>%
    # reformat count to character
    mutate(Incarcerated = formatC(Incarcerated, big.mark = ","),
           Convicted = formatC(Convicted, big.mark = ",")) %>%
    # combine count and percent
    mutate(Incarcerated = paste0(Incarcerated, " (", Incarcerated.perc, "%)"),
           Convicted = paste0(Convicted, " (", Convicted.perc, "%)")) %>%
    # remove unused columns
    select(-contains(".perc")) %>%
    # select top 5
    head(5) %>%
    # table styling
    knitr::kable(
```

```
      format = "simple",
      # rename header to include total time across all offenses
      col.names = c("Offense", header_Incarcerated, header_Convicted),
      caption = paste("Total years spent incarcerated and convicted by",
                      race_filter, "exonerees"))
  return(table)
}

# iterate over all groups of race
time_served_table("Black")
time_served_table("White")
time_served_table("Hispanic")
time_served_table("Asian")
time_served_table("Native American")
time_served_table("Other")
### ------------------------------------------------------------
### Distribution of time spent incarcerated and convicted, by class of offense
### and race, by total time spent incarcerated
plot_times2 <- function (data, crime_filter = NA) {

  ## filter the data by crime (optional)
  if (is.na(crime_filter)) {
    crime_filter <- "all" # this is used later for labeling
  } else {
    data <- data %>%
      filter(Crime.Category == crime_filter) %>%
      filter(!is.na(Months.Incarcerated) & !is.na(Months.Convicted))
  }

  ## only plot groups of race with 10 or more members present
  # get counts
  counts <- data %>%
    group_by(Race) %>%
    summarise(n = n()) %>%
    mutate(prop = n / sum(n)) %>%
    filter(n >= 10)
  # filter race groups with too few members
  plot_data <- data %>% filter(Race %in% counts$Race)

  ## write the subtitle of the plot, which shows counts of race groups
  gg_subtitle <- ""
  included_races <- unique(plot_data$Race)
  for (i in 1:length(included_races)) {
    race <- race_levels[i]
    count <- nrow(plot_data %>% filter(Race == race))
    gg_subtitle <- paste0(
      gg_subtitle, # previous groups
      # title, count, and percent (from data with all race groups)
      # ensure percent is over 'data' not the race filtered 'plot_data'
      race, ": ", count, " (", floor(100*count/nrow(data)), "%)")
    # add a semi-colon if the list isn't done
    if(i < length(included_races))
      gg_subtitle <- paste0(gg_subtitle, "; ")
```

```
  # add a new line after the fourth race group
  if (i == 4)
    gg_subtitle <- paste0(gg_subtitle, "\n")
  }

## plot distribution of time spent incarcerated
gg_incarated <- plot_data %>%
  ggplot(aes(y = Months.Incarcerated, x = Race)) +
  # jitter points (these must be overlayed by the horizontal line)
  geom_jitter(width = 0.4, color = "aliceblue") +
  # marker at 1 year
  geom_hline(yintercept = 12, color = "red") +
  # time spent wrongfully imprisoned
  geom_violin() +
  geom_boxplot(alpha = 0.5) +
  # plot specifics
  scale_y_continuous(breaks = seq(0, 400, 120)) +
  theme_bw() +
  labs(
    x = "",
    y = "Months imprisoned",
    title = paste("Time spent incarcerated by exonerees of",
                  crime_filter, "convictions"),
    subtitle = gg_subtitle
    )

## plot distribution of time spent convicted
gg_convicted <- plot_data %>%
  ggplot(aes(y = Months.Convicted, x = Race)) +
  # jitter points (these must be overlayed by the horizontal line)
  geom_jitter(width = 0.4, color = "aliceblue") +
  # marker at 1 year
  geom_hline(yintercept = 12, color = "red") +
  # time spent wrongfully convicted
  geom_violin() +
  geom_boxplot(alpha = 0.5) +
  # plot specifics
  scale_y_continuous(breaks = seq(0, 400, 120)) +
  theme_bw() +
  labs(
    x = "",
    y = "Months convicted",
    title = paste("Time spent convicted by exonerees of",
                  crime_filter, "convictions"),
    subtitle = gg_subtitle
    )

## tabular distribution of time spent convicted
tbl_convicted <- data %>%
  # calculate summary statistics by race
  group_by(across(Race)) %>%
  summarize(
    N = n(),
```

```
    Sum = sum(Months.Convicted),
    Quartile25 = quantile(Months.Convicted, 0.25, na.rm = TRUE),
    Quartile50 = median(Months.Convicted, na.rm = TRUE),
    Quartile75 = quantile(Months.Convicted, 0.75, na.rm = TRUE),
    Mean = mean(Months.Convicted, na.rm = TRUE),
    SD = sd(Months.Convicted, na.rm = TRUE)
    ) %>%
  # append a last row for summary statistics, regardless of race
  add_row(
    data %>%
      summarize(
        Race = "All",
        N = n(),
        Sum = sum(Months.Convicted, na.rm = TRUE),
        Quartile25 = quantile(Months.Convicted, 0.25, na.rm = TRUE),
        Quartile50 = median(Months.Convicted, na.rm = TRUE),
        Quartile75 = quantile(Months.Convicted, 0.75, na.rm = TRUE),
        Mean = mean(Months.Convicted, na.rm = TRUE),
        SD = sd(Months.Convicted, na.rm = TRUE)
        )
    ) %>%
  # arrange by sample size (beginning with overall)
  arrange(desc(N)) %>%
  # table styling
  knitr::kable(
    format = "latex",
    booktabs = TRUE,
    format.args = list(big.mark = ","),
    digits = 0,
    caption = paste("Months spent convicted by exonerees of", crime_filter,
                    "convictions")
    ) %>%
  # bold overall row
  kableExtra::row_spec(1, background = "#D3D3D3") %>%
  kableExtra::kable_styling(latex_options = "hold_position")



## tabular distribution of time spent incarcerated
tbl_incarcerated <- data %>%
  # calculate summary statistics by race
  group_by(across(Race)) %>%
  summarize(
    N = n(),
    Sum = sum(Months.Incarcerated),
    Quartile25 = quantile(Months.Incarcerated, 0.25, na.rm = TRUE),
    Quartile50 = median(Months.Incarcerated, na.rm = TRUE),
    Quartile75 = quantile(Months.Incarcerated, 0.75, na.rm = TRUE),
    Mean = mean(Months.Incarcerated, na.rm = TRUE),
    SD = sd(Months.Incarcerated, na.rm = TRUE)
  ) %>%
  # append a last row for summary statistics, regardless of race
  add_row(
```

```
      data %>%
        summarize(
          Race = "All",
          N = n(),
          Sum = sum(Months.Incarcerated, na.rm = TRUE),
          Quartile25 = quantile(Months.Incarcerated, 0.25, na.rm = TRUE),
          Quartile50 = median(Months.Incarcerated, na.rm = TRUE),
          Quartile75 = quantile(Months.Incarcerated, 0.75, na.rm = TRUE),
          Mean = mean(Months.Incarcerated, na.rm = TRUE),
          SD = sd(Months.Incarcerated, na.rm = TRUE)
          )
      ) %>%
    # arrange by sample size (beginning with overall)
    arrange(desc(N)) %>%
    # table styling
    knitr::kable(
      format = "latex",
      booktabs = TRUE,
      format.args = list(big.mark = ","),
      digits = 0,
      caption = paste("Months spent incarcerated by exonerees of", crime_filter,
                      "convictions")
      ) %>%
    # bold overall row
    kableExtra::row_spec(1, background = "#D3D3D3") %>%
    kableExtra::kable_styling(latex_options = "hold_position")


  ## return plots and tables
  return(list(
    plot_conviction = gg_convicted,
    plot_incarceration = gg_incarated,
    table_convicted = tbl_convicted,
    table_incarcerated = tbl_incarcerated)
    )
  }
## Violent felony
vlf_stats <- plot_times2(exonerees, crime_filter = "Violent felony")
# plots
gridExtra::grid.arrange(vlf_stats$plot_incarceration,
                        vlf_stats$plot_conviction + labs(subtitle = NULL))
# tables
vlf_stats$table_convicted
vlf_stats$table_incarcerated
## Misdemeanor
msd_stats <- plot_times2(exonerees, crime_filter = "Misdemeanor")
# plots
gridExtra::grid.arrange(msd_stats$plot_incarceration,
                        msd_stats$plot_conviction + labs(subtitle = NULL))
# tables
msd_stats$table_convicted
msd_stats$table_incarcerated
## Nonviolent felony
```

```r
nvf_stats <- plot_times2(exonerees, crime_filter = "Nonviolent felony")
# plots
gridExtra::grid.arrange(nvf_stats$plot_incarceration,
                        nvf_stats$plot_conviction + labs(subtitle = NULL))
# tables
nvf_stats$table_convicted
nvf_stats$table_incarcerated
```

End of document