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, policymakers, 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 = 27	**Native American**, N = 21	**Other**, N = 19
Murder	875 (33%)	496 (35%)	226 (29%)	137 (37%)	4 (15%)	8 (38%)	4 (21%)
Drug Possession or Sale	601 (23%)	433 (31%)	85 (11%)	78 (21%)	3 (11%)	1 (4.8%)	1 (5.3%)

Offense	**Overall**, N = 2,640	**Black**, N = 1,415	**White**, N = 785	**Hispanic**, N = 373	**Asian**, N = 27	**Native American**, N = 21	**Other**, N = 19
Child Sex Abuse	228 (8.6%)	60 (4.2%)	138 (18%)	24 (6.4%)	3 (11%)	2 (9.5%)	1 (5.3%)
Sexual Assault	195 (7.4%)	101 (7.1%)	73 (9.3%)	17 (4.6%)	0 (0%)	2 (9.5%)	2 (11%)
Robbery	142 (5.4%)	94 (6.6%)	23 (2.9%)	22 (5.9%)	0 (0%)	1 (4.8%)	2 (11%)
Assault	98 (3.7%)	40 (2.8%)	33 (4.2%)	23 (6.2%)	0 (0%)	1 (4.8%)	1 (5.3%)

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%)

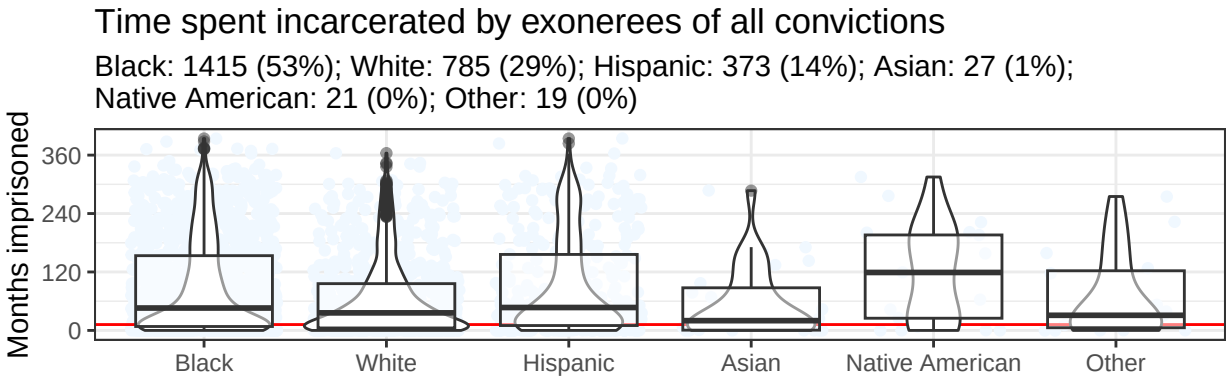
Table 2 shows the criminal offenses that account for the most time spent wrongfully incarcerated/convicted. For example, we read the top-left-most cell as reporting that, among the 17,086 years spent wrongfully incarcerated across all individuals in the National Registry of Exonerations whose alleged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

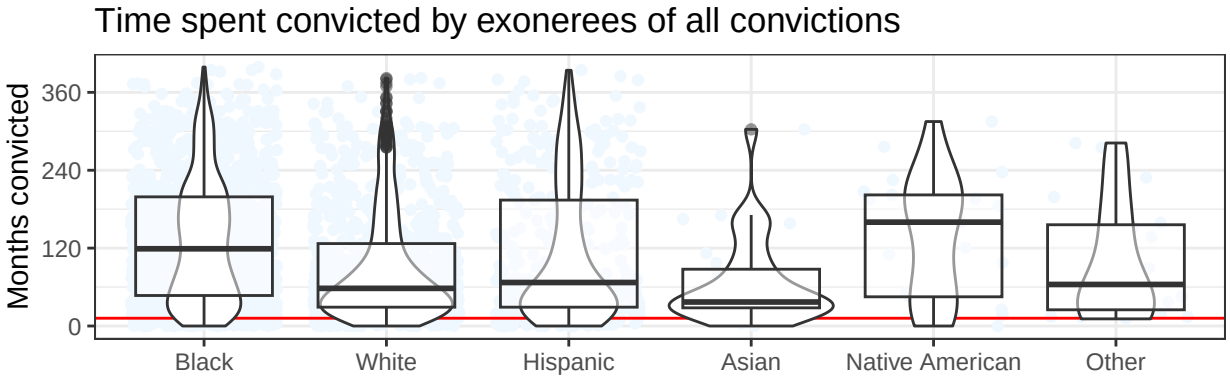
Table 2 was reproduced for each group of race (see Supplementary), which found that this order of offenses accounting for the most years lost to incarceration or conviction are similar across all subsets of races, especially the largest.

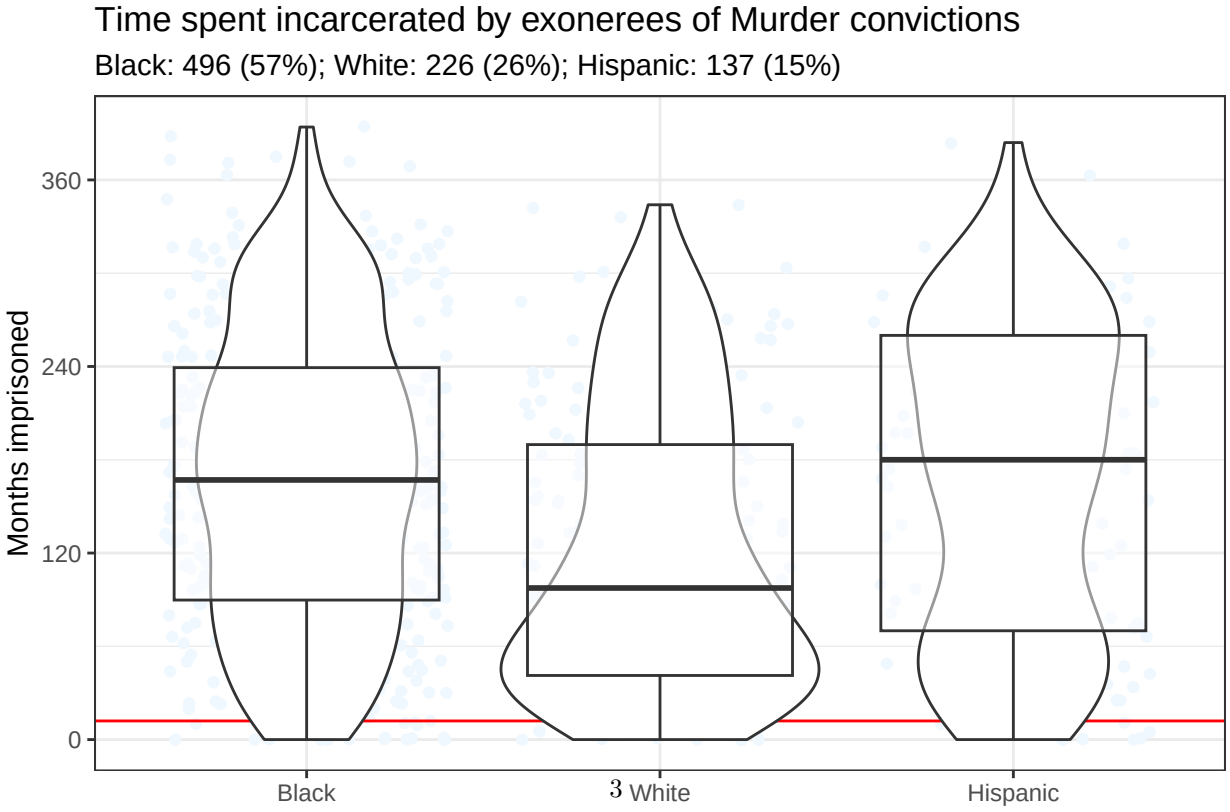
These tables motivate us to compare the experience of people convicted of these popular and severe offenses between groups of race, so as to measure the extent to which sentence/incarceration is associated with race, within the National Registry of Exonerations.

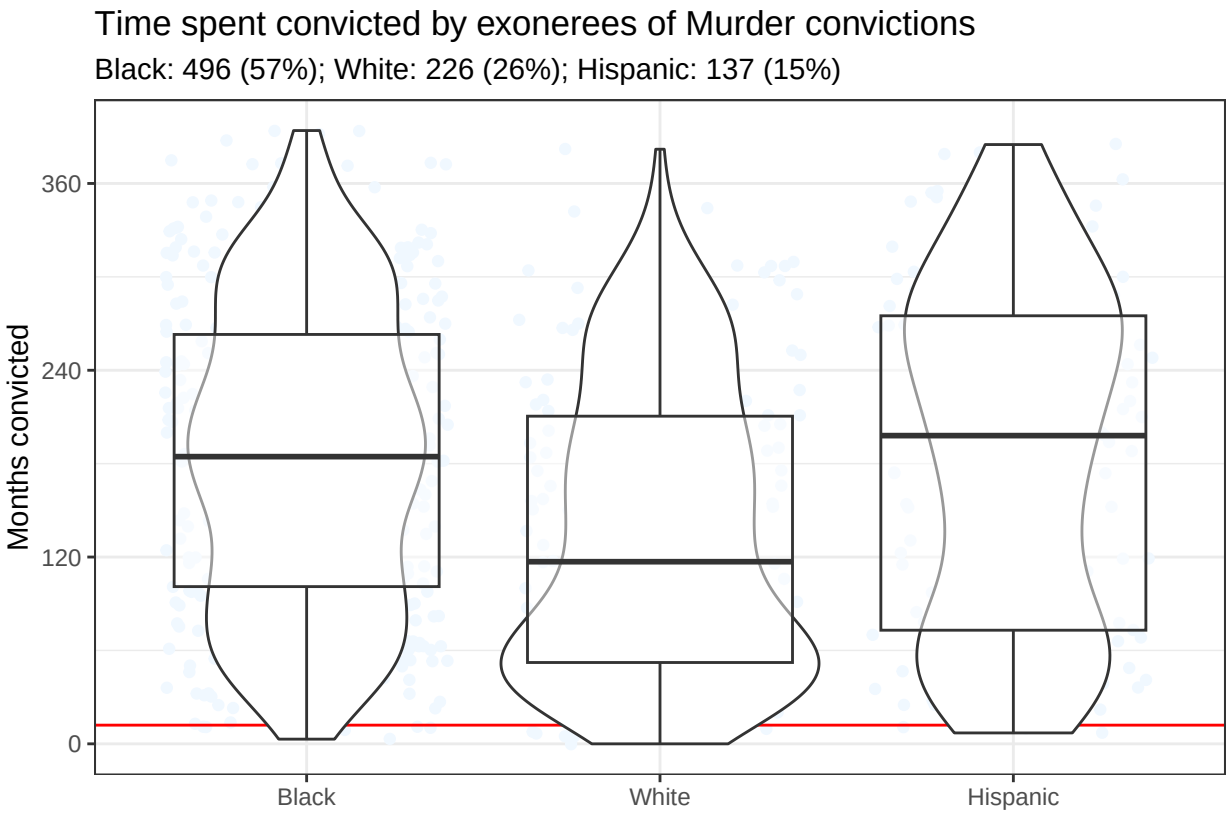
Analysis

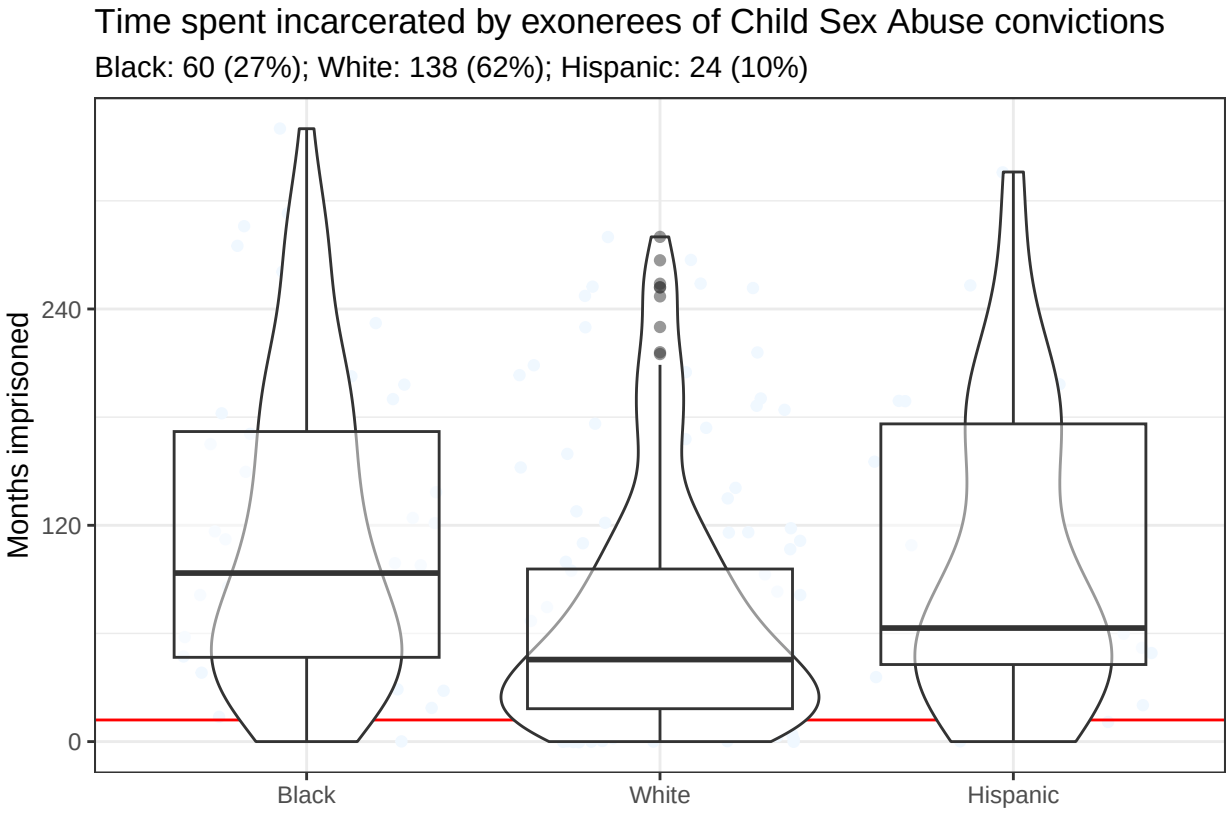
Distribution of years spent incarcerated and convicted, by offense and race

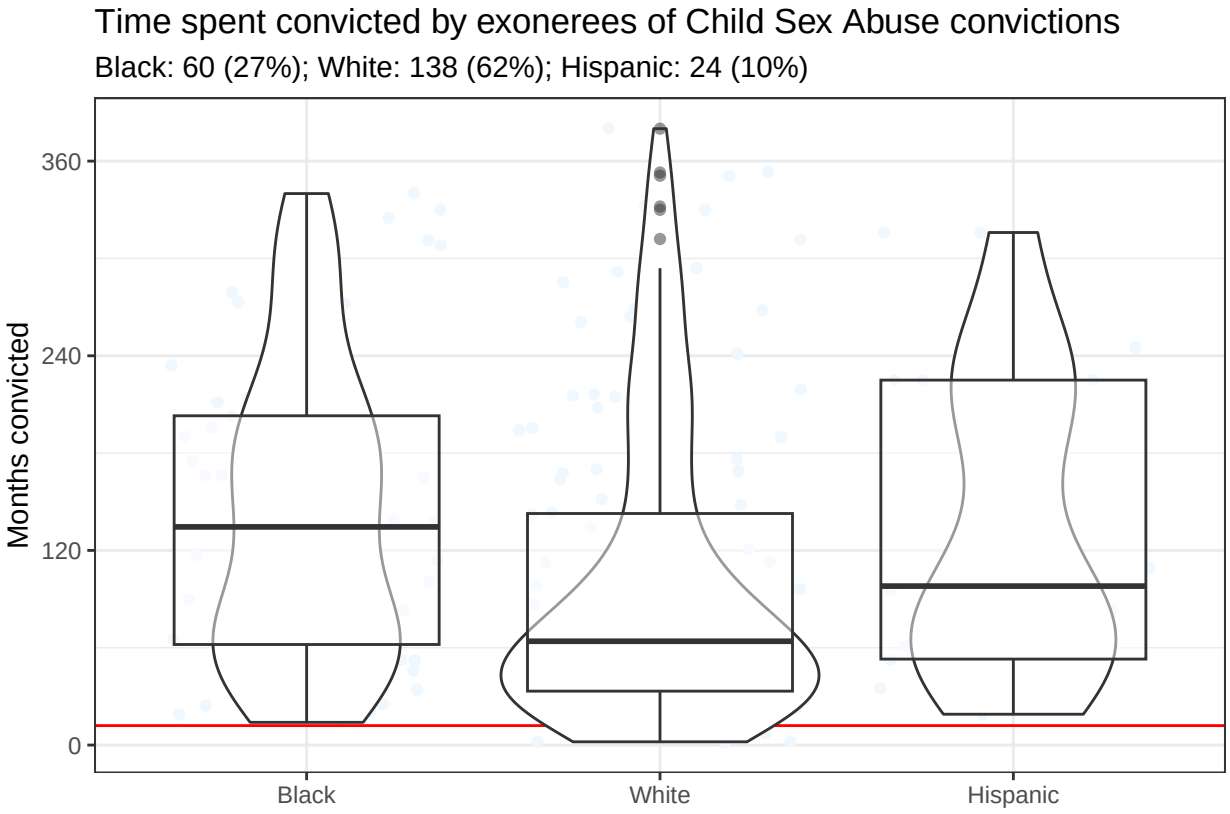


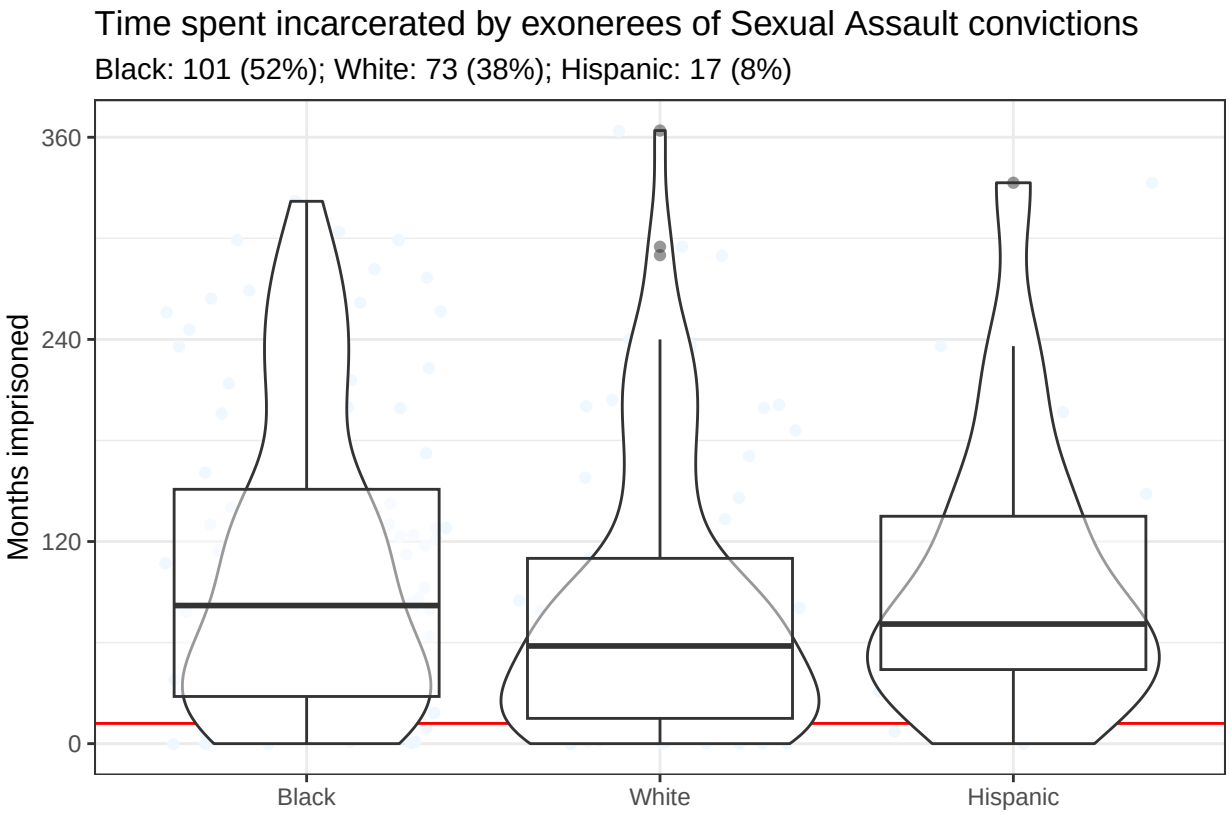


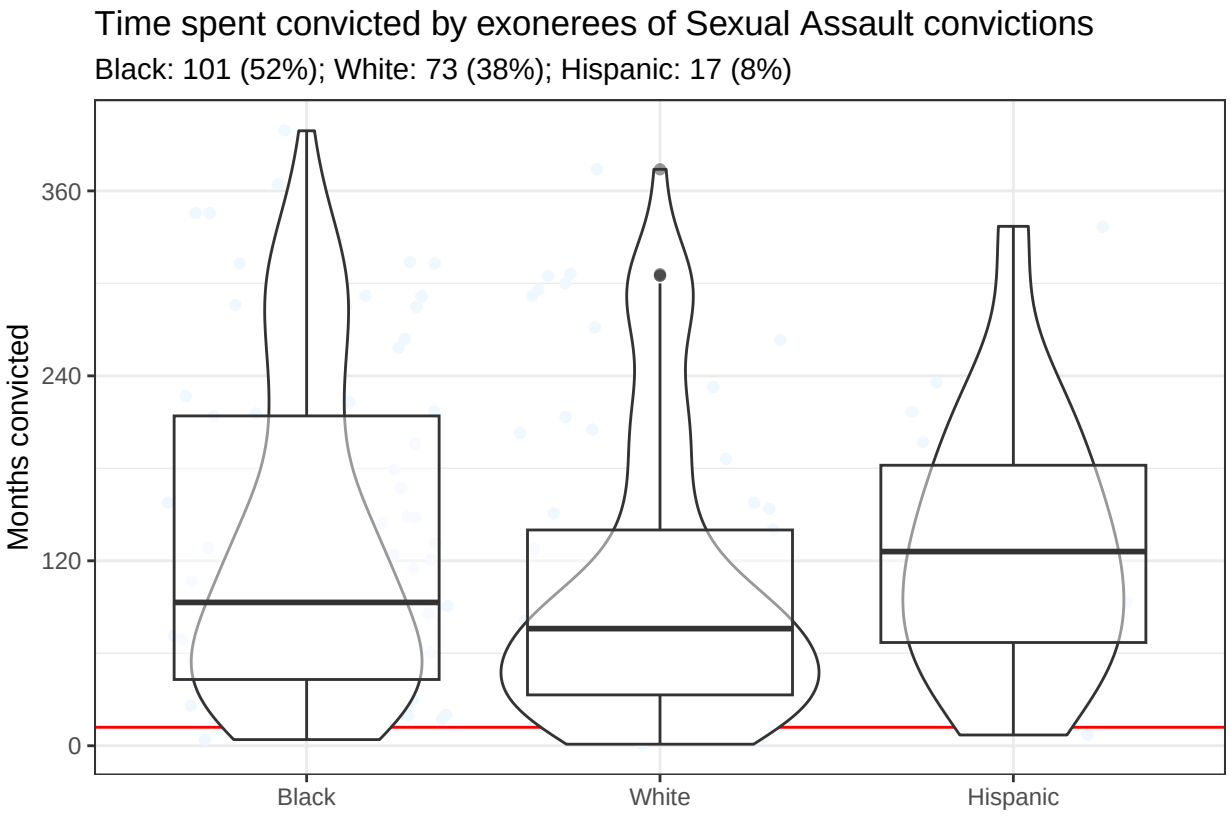


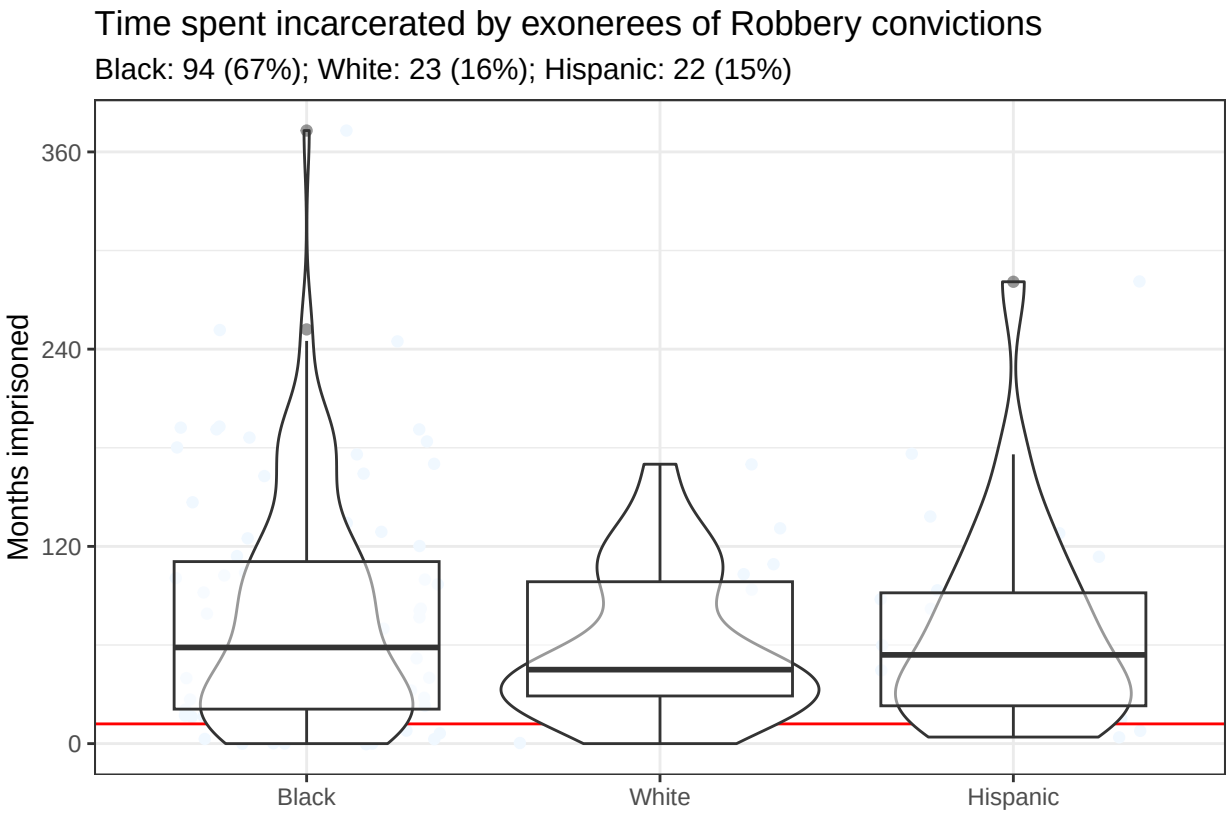

Time spent convicted by exonerees of Murder convictions
Black: 496 (57%); White: 226 (26%); Hispanic: 137 (15%)
Months convicted
360
240
120
0
Black
White
Hispanic

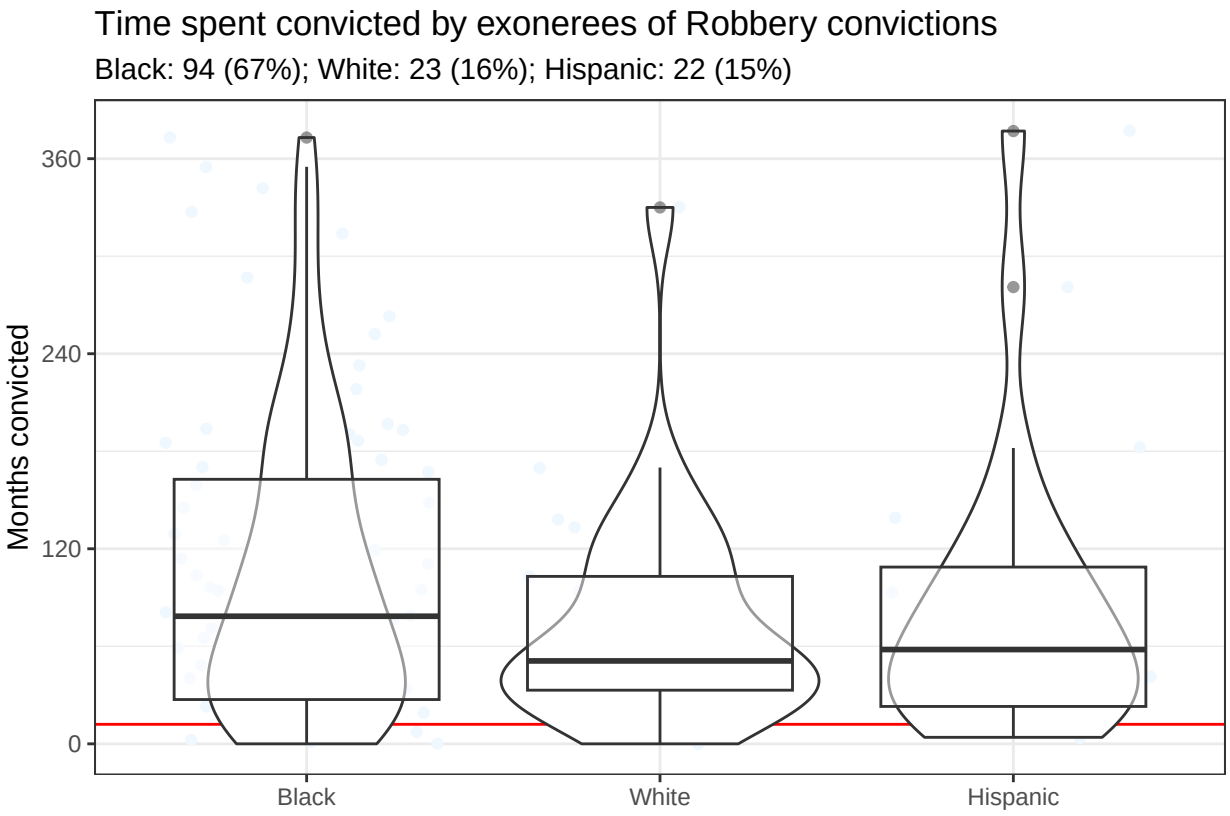

Time spent incarcerated by exonerees of Child Sex Abuse convictions
Black: 60 (27%); White: 138 (62%); Hispanic: 24 (10%)
Months imprisoned
240
120
0
Black
White
Hispanic

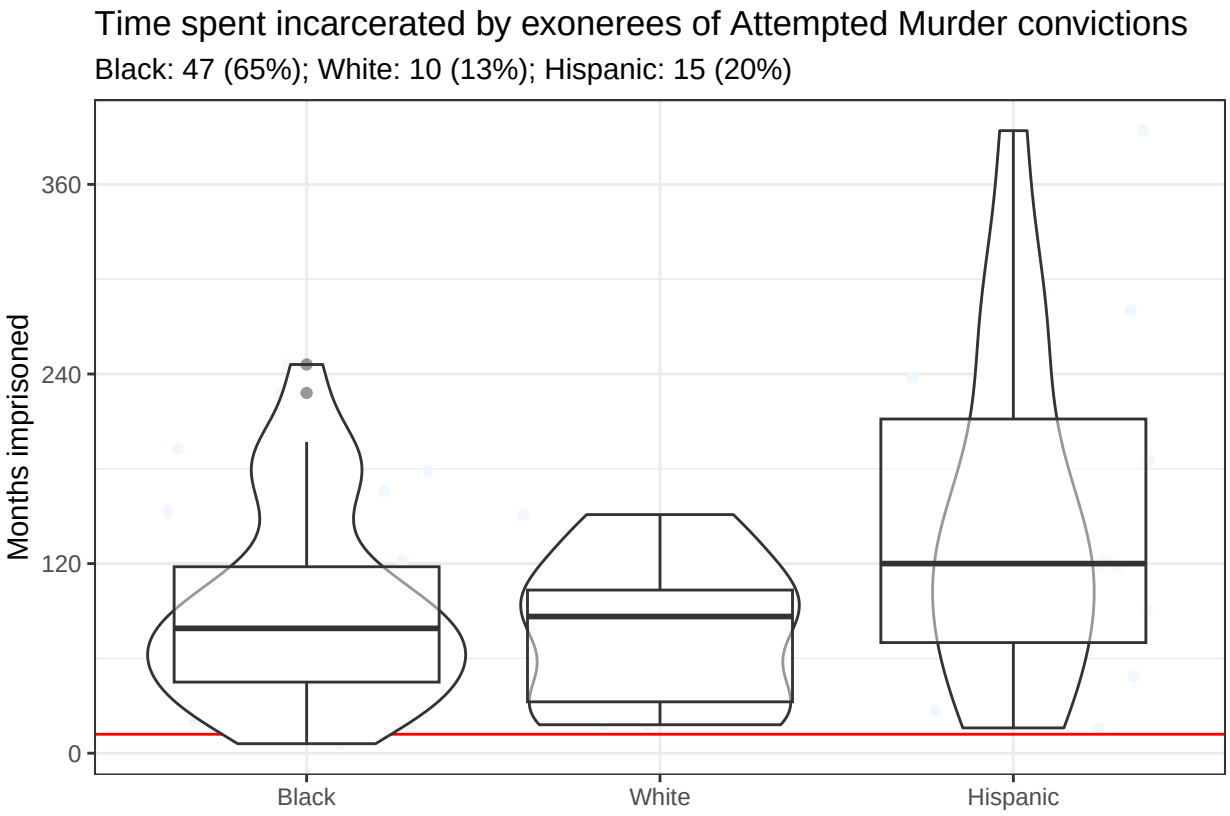

Time spent convicted by exonerees of Child Sex Abuse convictions
Black: 60 (27%); White: 138 (62%); Hispanic: 24 (10%)
Months convicted
360
240
120
0
Black
White
Hispanic

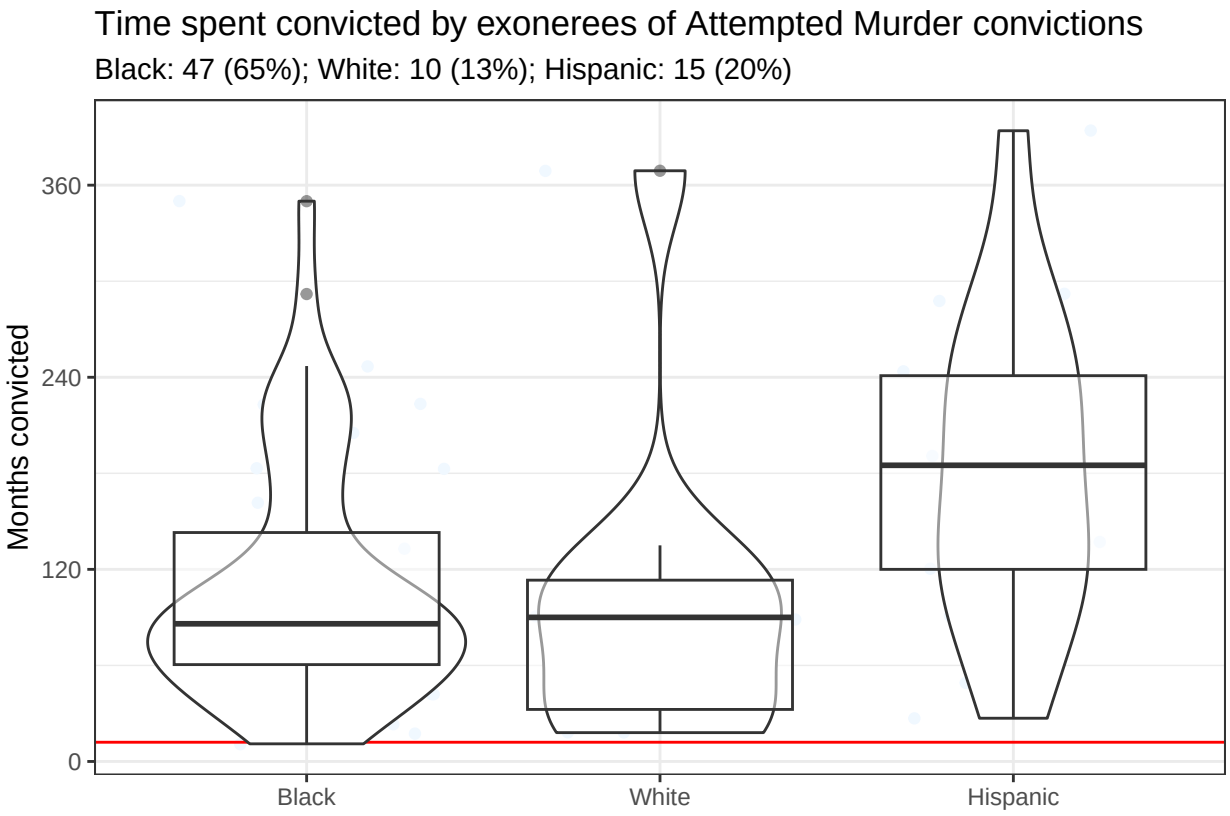

Time spent incarcerated by exonerees of Sexual Assault convictions
Black: 101 (52%); White: 73 (38%); Hispanic: 17 (8%)
Months imprisoned
360
240
120
0
Black
White
Hispanic

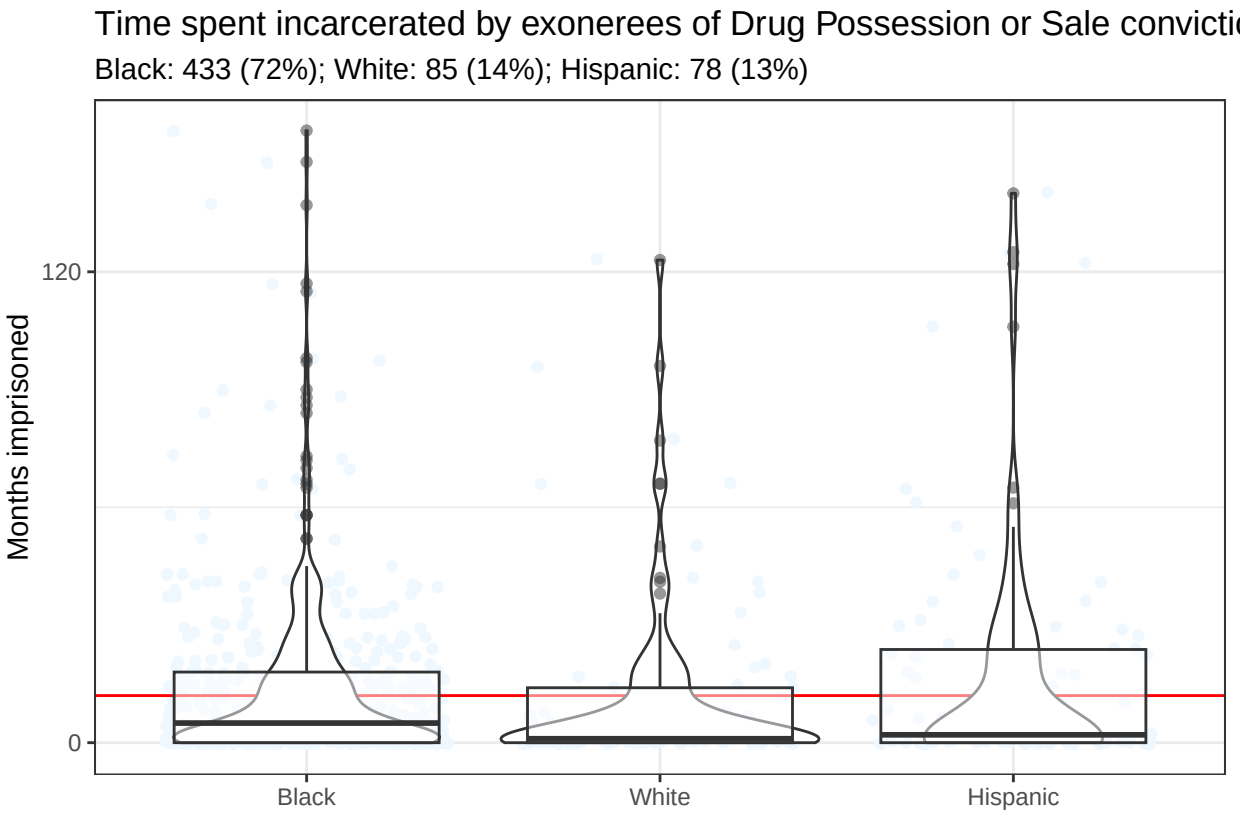

Time spent convicted by exonerees of Sexual Assault convictions
Black: 101 (52%); White: 73 (38%); Hispanic: 17 (8%)
Months convicted
360
240
120
0
Black
White
Hispanic

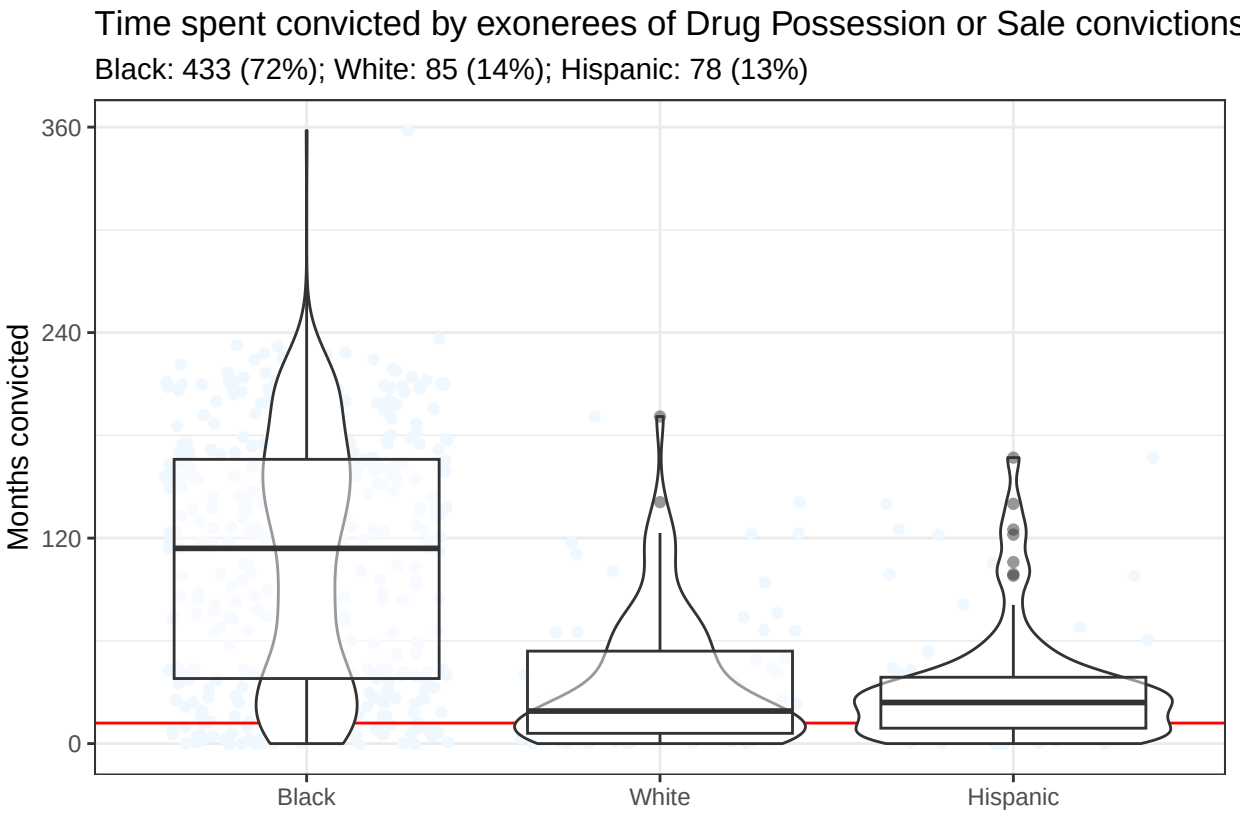

Time spent incarcerated by exonerees of Robbery convictions
Black: 94 (67%); White: 23 (16%); Hispanic: 22 (15%)
Months imprisoned
360
240
120
0
Black
White
Hispanic

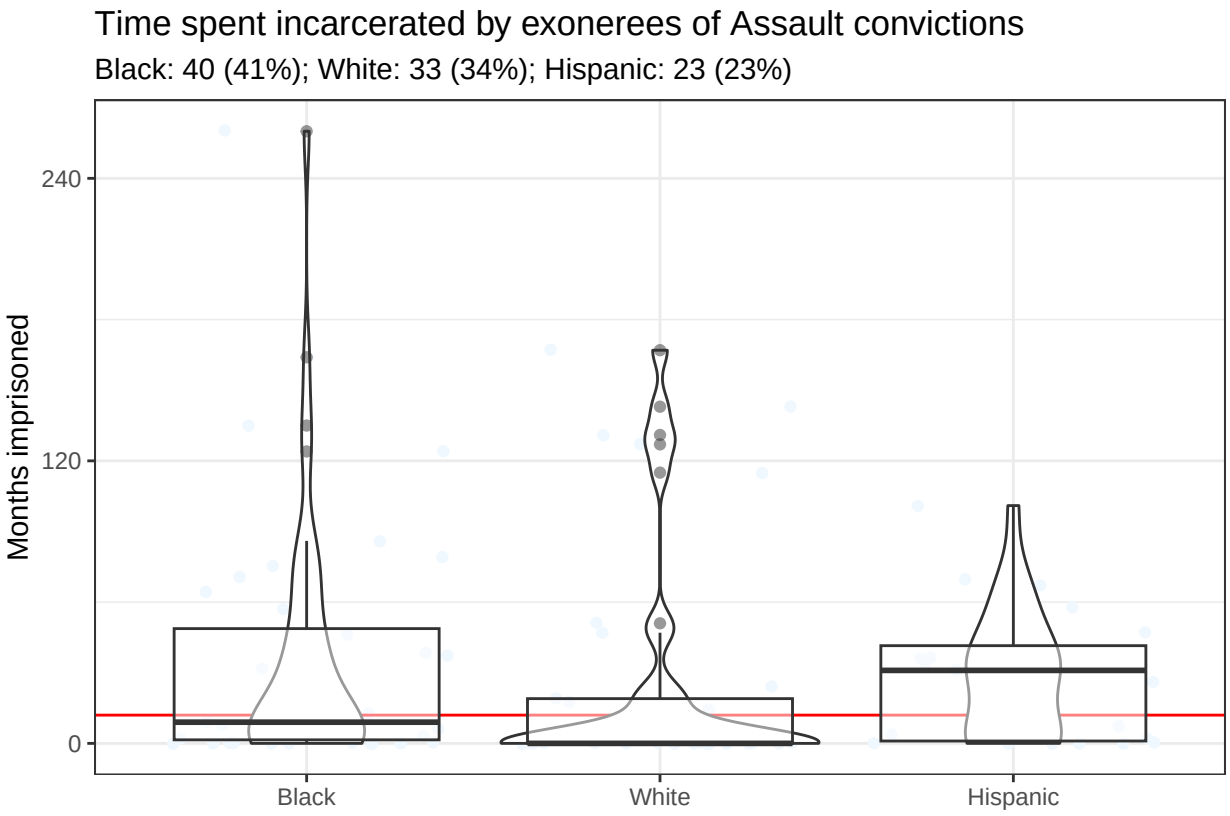

Time spent convicted by exonerees of Robbery convictions
Black: 94 (67%); White: 23 (16%); Hispanic: 22 (15%)
Months convicted
360
240
120
0
Black
White
Hispanic

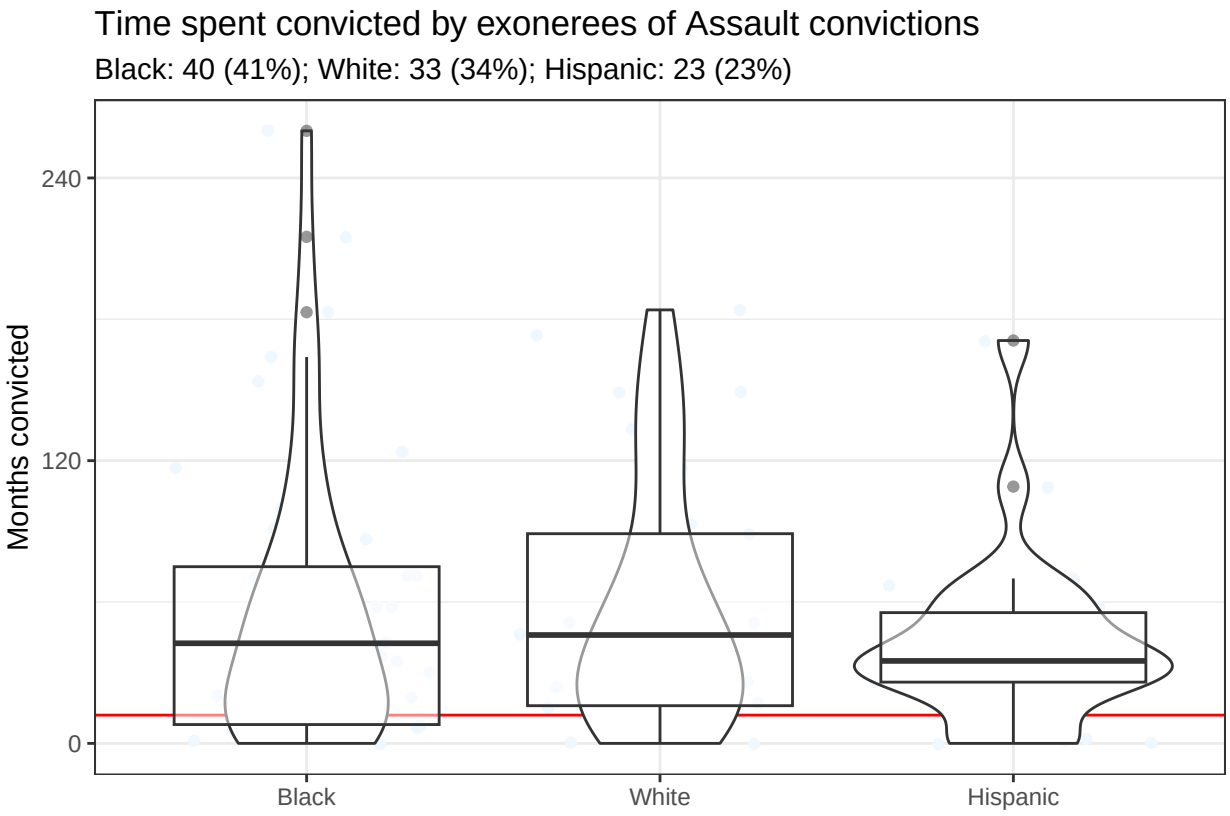

Time spent incarcerated by exonerees of Attempted Murder convictions
Black: 47 (65%); White: 10 (13%); Hispanic: 15 (20%)
Months imprisoned
360
240
120
0
Black
White
Hispanic

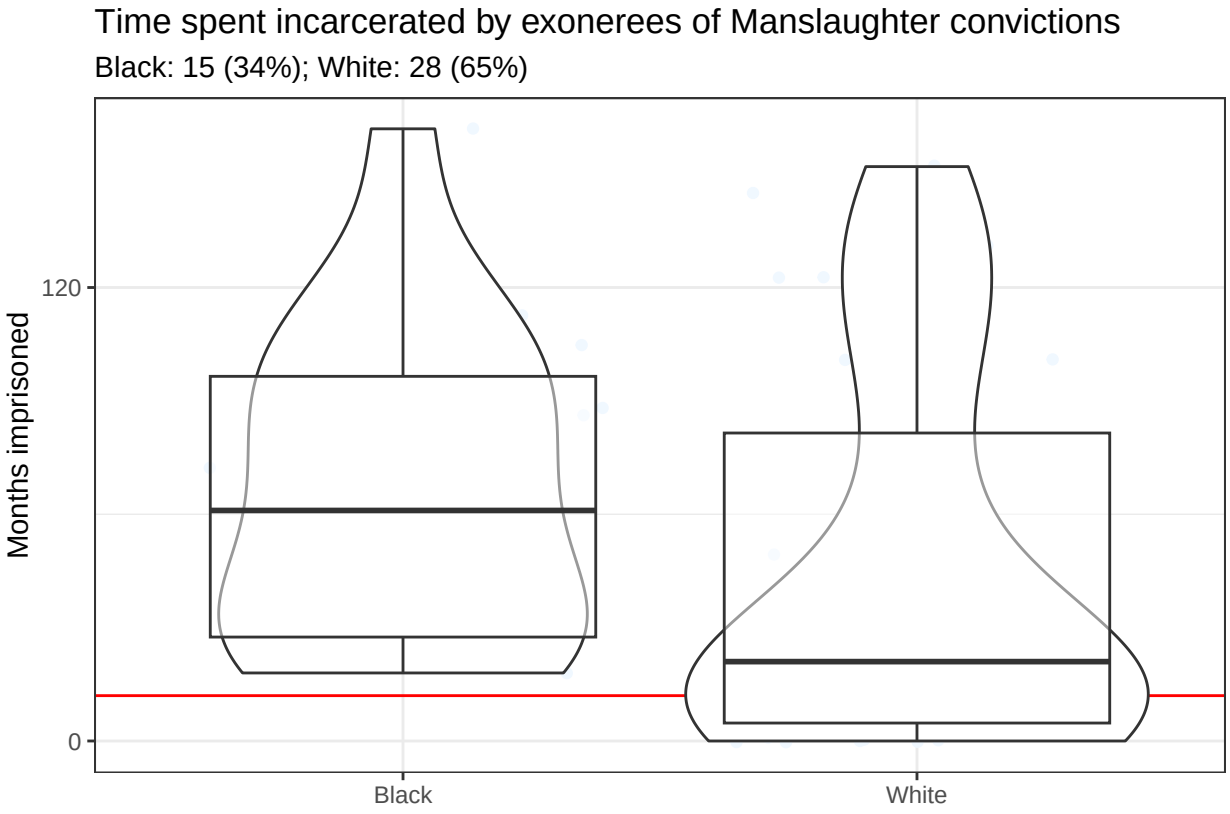

Time spent convicted by exonerees of Attempted Murder convictions
Black: 47 (65%); White: 10 (13%); Hispanic: 15 (20%)
Months convicted
360
240
120
0
Black
White
Hispanic

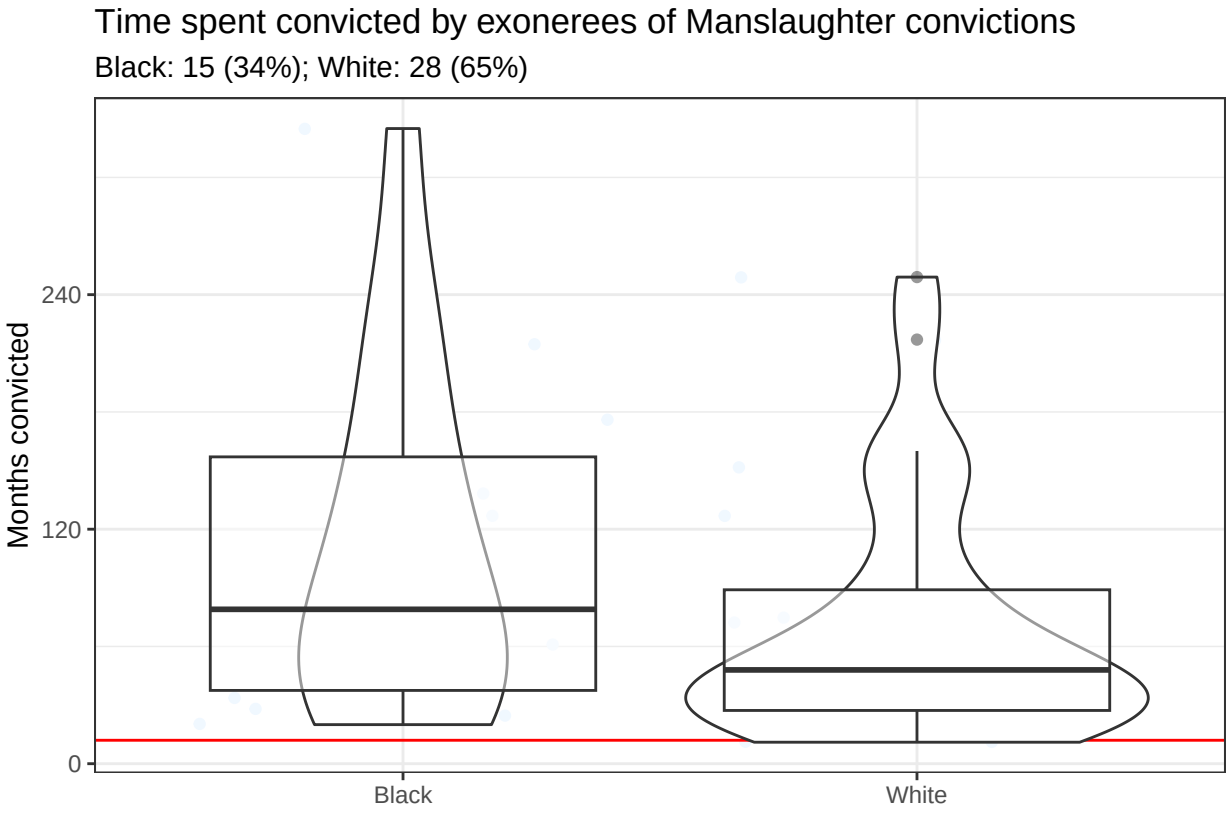

Time spent incarcerated by exonerees of Drug Possession or Sale convicti
Black: 433 (72%); White: 85 (14%); Hispanic: 78 (13%)
Months imprisoned
120
0
Black
White
Hispanic


Time spent convicted by exonerees of Drug Possession or Sale convictions
Black: 433 (72%); White: 85 (14%); Hispanic: 78 (13%)
Months convicted
360
240
120
0
Black
White
Hispanic


Time spent incarcerated by exonerees of Assault convictions
Black: 40 (41%); White: 33 (34%); Hispanic: 23 (23%)
Months imprisoned
240
120
0
Black
White
Hispanic


Time spent convicted by exonerees of Assault convictions
Black: 40 (41%); White: 33 (34%); Hispanic: 23 (23%)
Months convicted
240
120
0
Black
White
Hispanic


Time spent incarcerated by exonerees of Manslaughter convictions
Black: 15 (34%); White: 28 (65%)
Months imprisoned
120
0
Black
White



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 3: 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 4: 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 5: 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 6: 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 7: 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 8: 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%)

Offense	Incarcerated (105 total)	Convicted (148 total)
Kidnapping	9 (9%)	12 (8%)
Assault	4 (4%)	5 (3%)

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", "llm_appendix", "allcode")]

library(tidyverse)
library(gridExtra)
library(gtsummary)
library(gt)
library(knitr)
### -------------------------------------------------------------
### Data Loading
rm(list = ls()) # clear workspace
# 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)

# 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)
```

```r
# make sure everything looks good
summary(exonerees)
dim(exonerees)
### ------------------------------------------------------------
### Ordered Sample Size Tables
# offense frequency (ordered), overall and by race
exonerees %>%
  select(Race, Worst.Crime.Display) %>%
  filter(!is.na(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
  head() %>%
  kable(caption = "List of high-frequency offenses, by race")
### ------------------------------------------------------------
### Time Served Table
# set headers to columns, 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
  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")) %>%
  kable(
    # 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 years 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)
  }

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

  # write the subtitle of the plot, which shows counts of race groups
  gg_subtitle <- ""
  included_races <- unique(data$Race)
  for (i in 1:length(included_races)) {
    race <- race_levels[i]
    count <- nrow(data %>% filter(Race == race))
    gg_subtitle <- paste0(
      gg_subtitle, # previous groups
      # title, count, and percent
      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 <- data %>%
    filter(!is.na(Months.Incarcerated)) %>%
    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
    ylab("Months imprisoned") + xlab("") +
    scale_y_continuous(breaks = seq(0, 400, 120)) +
    theme_bw() +
    labs(
    title = paste("Time spent incarcerated by exonerees of", crime_filter,
                  "convictions"),
    subtitle = gg_subtitle)

  # plot distribution of time spent convicted
  gg_convicted <- data %>%
    filter(!is.na(Months.Convicted)) %>%
    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
    ylab("Months convicted") + xlab("") +
    scale_y_continuous(breaks = seq(0, 400, 120)) +
    theme_bw() +
    labs(
    title = paste("Time spent convicted by exonerees of", crime_filter,
                  "convictions"),
    subtitle = gg_subtitle)


  return(list(incarceration_times = gg_incarated,
              conviction_times = gg_convicted))
}

# Overall
gridExtra::grid.arrange(
  plot_times(exonerees)$incarceration_times,
  plot_times(exonerees)$conviction_times + labs(subtitle = NULL))

# Murder
plot_times(exonerees, crime_filter = "Murder")$incarceration_times
plot_times(exonerees, crime_filter = "Murder")$conviction_times

# Child sex abuse
plot_times(exonerees, crime_filter = "Child Sex Abuse")$incarceration_times
plot_times(exonerees, crime_filter = "Child Sex Abuse")$conviction_times
```

```r
# Sexual assault
plot_times(exonerees, crime_filter = "Sexual Assault")$incarceration_times
plot_times(exonerees, crime_filter = "Sexual Assault")$conviction_times

# Robbery
plot_times(exonerees, crime_filter = "Robbery")$incarceration_times
plot_times(exonerees, crime_filter = "Robbery")$conviction_times

# Attempted murder
plot_times(exonerees, crime_filter = "Attempted Murder")$incarceration_times
plot_times(exonerees, crime_filter = "Attempted Murder")$conviction_times

# Drug possession or sale
plot_times(exonerees, crime_filter = "Drug Possession or Sale")$incarceration_times
plot_times(exonerees, crime_filter = "Drug Possession or Sale")$conviction_times

# Assault
plot_times(exonerees, crime_filter = "Assault")$incarceration_times
plot_times(exonerees, crime_filter = "Assault")$conviction_times

# Manslaughter
plot_times(exonerees, crime_filter = "Manslaughter")$incarceration_times
plot_times(exonerees, crime_filter = "Manslaughter")$conviction_times
### ------------------------------------------------------------
### Supplementary Time Served Table

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

  # set headers to columns, 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
    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) %>%
    # pretty print
    kable(
      # 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")
```

Document in progress.