Stop outcomes, day-to-night, in Oakland, CA

This section of data analysis incorporates our day variable with stop outcomes in the Oakland data set.

Set up

I use the following packages for this analysis. Suncalc, lutz (a cute acronym for "look up time zone"), and lubridate are necessary for calculating sunset and sunrise times. I also connect to our SQL server, but that code is excluded here because of privacy.

```
library(tidyverse)
library(RMySQL)
library(suncalc)
library(lutz)
library(lubridate)
```

Next, I query the entire Oakland data set to be available locally. I also clean the data by removing the entries that do not record data, parsing my date and time variables with lubridate, and creating a POSIX type time variable that will serve as an input for suncalc functions.

```
CAoak <- DBI::dbGetQuery(con, "SELECT * FROM CAoakland")

# Lubridate
CAoak <- CAoak %>%
  filter(!is.na(date)) %>%
  mutate(nice_date = ymd(date),
         nice_year = year(nice_date),
         nice_month = month(nice_date),
         nice_day = day(nice_date),
         nice_time = hms(time),
         nice_day_of_year = yday(date),

         #for sunset sunrise:
         posix_date_time = as.POSIXct(paste(nice_date, time), tz = "America/Los_Angeles", format = "%Y-
```

```
## Warning in .parse_hms(..., order = "HMS", quiet = quiet): Some strings
## failed to parse, or all strings are NAs
```

Set up sunset/sunrise times in CA Oakland

Next, I use the function we wrote, oursunriseset, and dplyr functions to create a new set of variables in the Oakland data. For each stop, we analyze the time of day and day of the year to categorize the stop as either occurring during the day (for which the new variable light would be "day") or during the night (for which light would be "night."

When I use the suncalc function getSunlightTimes, I let the input timezone be the Oakland timezone.

```
OaklandTZ <- lutz::tz_lookup_coords(37.75241810000001, -122.18087990000001, warn = F)

oursunriseset <- function(latitude, longitude, date, direction = c("sunrise", "sunset")) {

  date.lat.long <- data.frame(date = date, lat = latitude, lon = longitude)
```

```
  if(direction == "sunrise"){

    getSunlightTimes(data = date.lat.long, keep=direction, tz = OaklandTZ)$sunrise }else{
      getSunlightTimes(data = date.lat.long, keep=direction, tz = OaklandTZ)$sunset } }
```

Next, I run oursunriseset on the Oakland data set, using mutate to create the intermediate variables, sunrise and sunset, and the light variable. I remove the traffic stop observations for which time was not recorded.

```
# add light variable
CAoak <- CAoak %>%

  # use oursunriseset function to return posixct format sunrise and sunset times
  mutate(sunrise = oursunriseset(lat, lng, nice_date, direction = "sunrise"),
         sunset = oursunriseset(lat, lng, nice_date, direction = "sunset")) %>%

  mutate(light = ifelse(posix_date_time > sunrise & posix_date_time < sunset, "day", "night")) %>%

  # about 100 NA's to filter out

  filter(!is.na(light))

knitr::kable((CAoak %>% select(posix_date_time, subject_race, light) %>% head()))
```

posix_date_time	subject_race	light
2013-04-01 00:00:00	white	night
2013-04-01 00:01:00	black	night
2013-04-01 10:41:00	hispanic	day
2013-04-01 14:18:00	white	day
2013-04-01 00:15:00	black	night
2013-04-01 15:32:00	black	day

This short excerpt of the transformed Oakland data set shows how the time of stop and light variable correspond to one another.

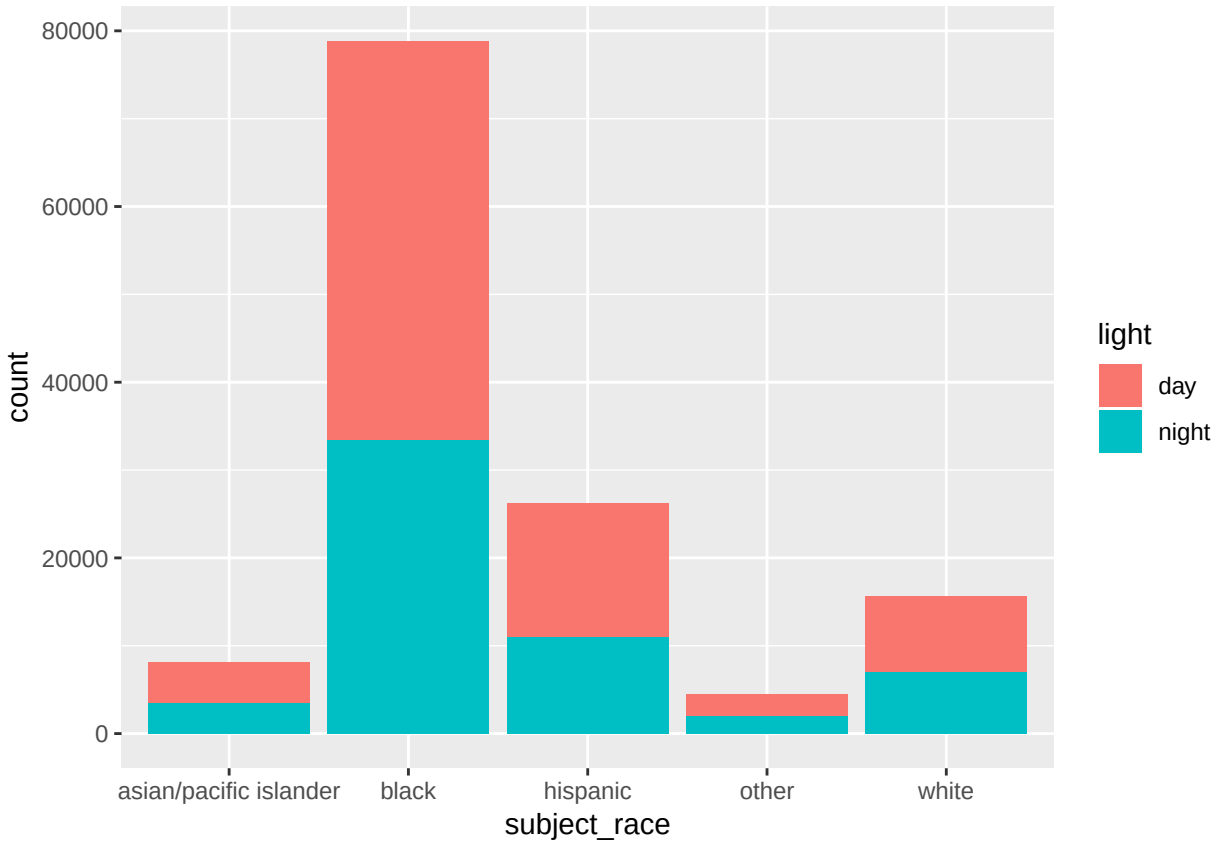
Stop outcomes, by race and day/night

With the set up done, we can begin plotting stop outcomes related to day and night!

Counts, by race

The bar chart shows the number of stops per race occurring during the day and night.

```
# see how drivers are stopped by race and light in absolute counts
ggplot(data = CAoak) +
  geom_bar(mapping = aes(x = subject_race, fill = light))
```




```
ggsave("night and day stop counts by race.png")
```

```
## Saving 6.5 x 4.5 in image
```

The bar chart relays effects of a slew of variables, such as driving behavior of a time of day, time of year, and racial group; traffic and patrol behavior for a given time and location is also relayed. Thus, again, we cannot make definitive statements about racial profiling.

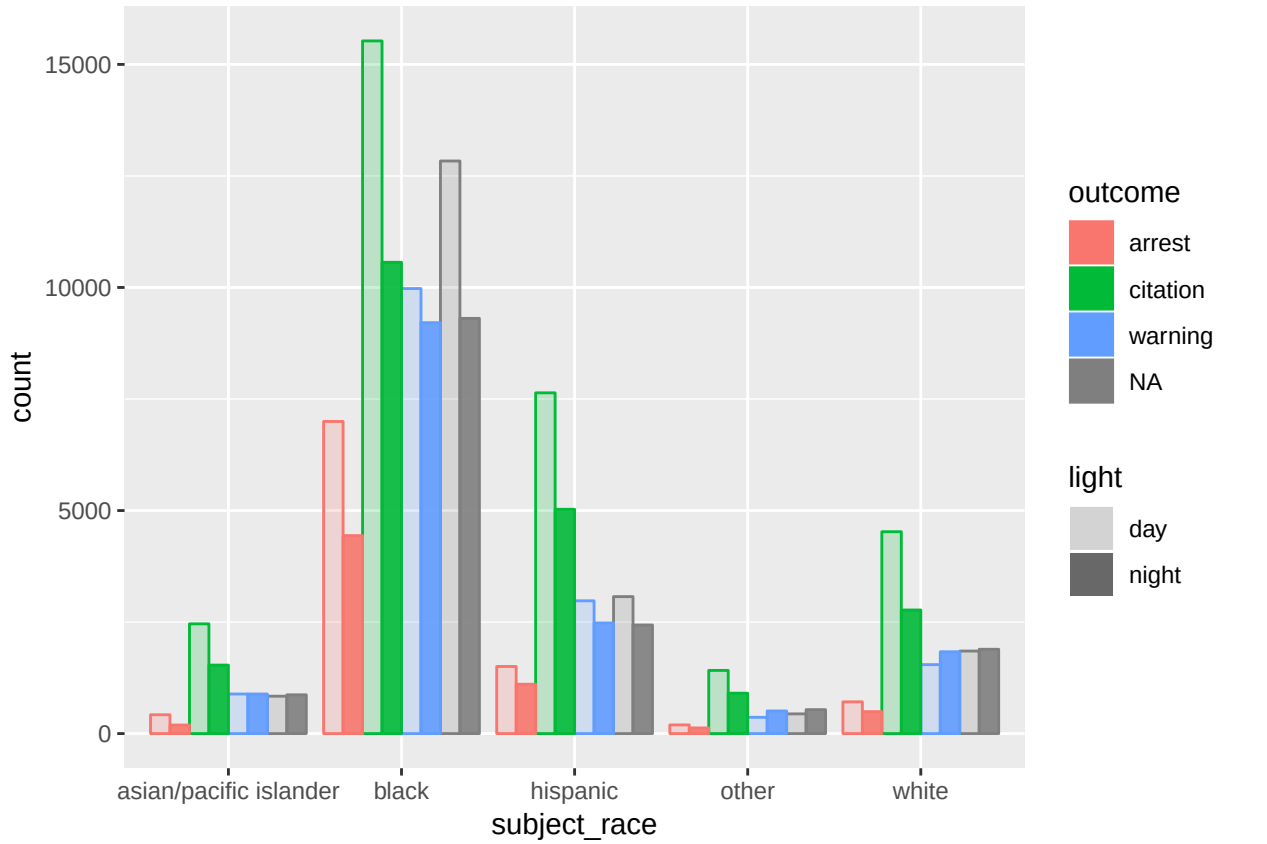
As in the veil of darkness assumption, the race of a motorist is harder to discern during the night than during the day. However, just because the day-night bars in each racial group looks to be about equal neither corroborate nor denies racial profiling. We are without information on the true traffic infringement rate per racial group in different times of the day.

Day/night outcomes of stop by race

Next, I examine how outcome of stop, race, and day/night interact. I pipe the Oakland data set into ggplot, allowing the color of each bar to represent a traffic stop outcome and the transparency of the bars to denote day and night stops.

```
# how are outcomes of stops affected by race and time of day?

CAoak %>%
  ggplot(aes(x = subject_race, fill = outcome, color = outcome, alpha = light)) +
  geom_bar(position = "dodge") +
  scale_alpha_manual(values = c(.2, .9))
```




```
ggsave("outcome and race and light.png")
```

```
## Saving 6.5 x 4.5 in image
```

Observations:

- For each type of outcome and race, stops during the day outnumber stops during the night. This may result from the Oakland police department's patrol patterns, driver behavior during the day, and/or another factor.
- Missing data regarding traffic outcome does not look to be evenly distributed among stopped motorists – missing data during the day and night look to be about the same for all racial groups except for Black.
- Citations look to be the most frequent outcome of traffic stops for all racial groups, followed by warning, then arrest (not including missing data.)
- Future directions could look more deeply into the probabilities of each of these outcomes for different racial groups, as the nature of a stricter outcome (being issued a citation rather than a warning) may not be uniformly distributed among racial groups. Furthermore, incorporating the light variable in this direction could be illuminating: are traffic patrols more lenient during the day or during the night once a motorist is stopped? Does the extent of this leniency differ for different racial groups?