

Proportion of cut trees in harvested plots

Tim White

June 2024

Packages and data

```
library(readxl)
library(tidyverse)

orig_all_data <- read_excel("data/data_raw.xlsx", sheet = "All_data_Categories")

plot_theme <- theme(plot.background = element_rect(fill = "white"),
                     plot.title = element_blank(),
                     plot.subtitle = element_text(family="sans", face="plain"),
                     axis.title.x = element_text(family="sans", face="bold"),
                     axis.title.y = element_text(family="sans", face="bold"),
                     axis.text.x = element_text(family="sans", face="plain"),
                     axis.text.y = element_text(family="sans", face="plain"),
                     panel.background = element_rect(fill="white"),
                     panel.grid.major.x = element_blank(),
                     panel.grid.major.y = element_line(color="gainsboro"),
                     panel.grid.minor = element_blank(),
                     axis.ticks = element_blank(),
                     legend.background = element_rect(color="black", fill = "white"),
                     legend.position = c(0.9, 0.85),
                     legend.title = element_text(color = "black", face = "bold", hjust = 0.5),
                     legend.text = element_text(color = "black"))
```

Clean data

Note that “cortado” and “cut” are used interchangeably throughout this document.

```
# Remove unnecessary columns from orig_all_data
prop_cortado_data <- orig_all_data %>%
  select(plot_id = `Plot #`,
         tree_id = `ID #`,
         stem_id = `Stem #`,
         harvested = Harvested,
         observation = Observacion) %>%
  mutate(observation = replace_na(observation, ""))
  mutate(harvested = fct_recode(harvested,
                                "yes" = "Yes", "no" = "No"))
```

```
# Create new data frame to compute proportion of cortado trees in each plot
prop_cortado <- prop_cortado_data %>%
  group_by(plot_id) %>%
  summarize(harvested = unique(harvested),
            num_cortado = sum(observation == "cortado"),
            num_total = diff(range(tree_id)) + 1,
            proportion_cortado = num_cortado/num_total) %>%
  mutate(plot_id = as.factor(plot_id),
         harvested = as.factor(harvested))
```

Summary and figures

Numerical summaries

```
# We find that the proportion of trees in harvested plots that are cortado ranges from 0% to 25%
# The mean proportion is 10.6%, and the median proportion is 9.1%
prop_cortado %>%
  filter(harvested == "yes") %>%
  pull(proportion_cortado) %>% summary()

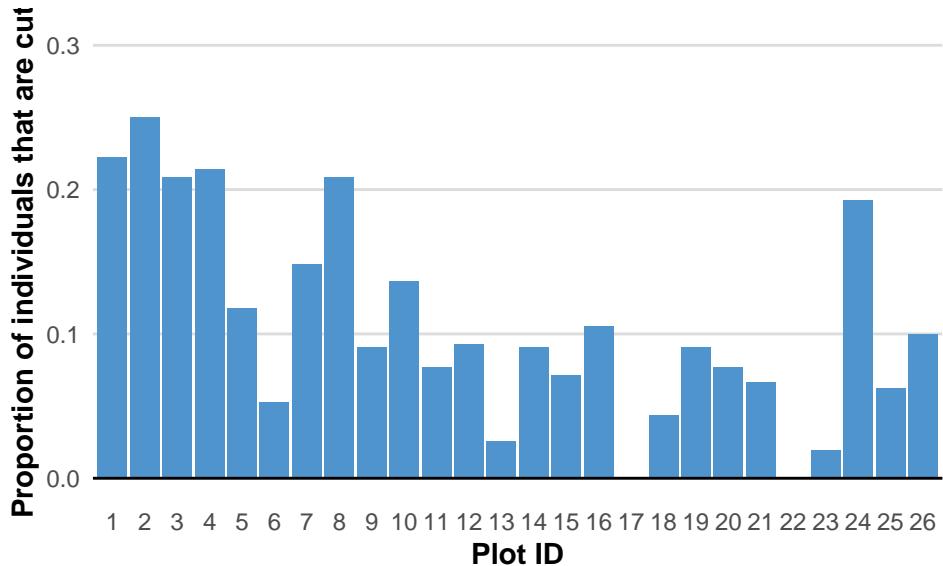
##      Min. 1st Qu. Median      Mean 3rd Qu.      Max.
## 0.00000 0.06354 0.09091 0.10629 0.14520 0.25000

# In all harvested plots, there are 66 cortado trees and 698 total trees for a proportion of 9.5%
prop_cortado %>%
  filter(harvested == "yes") %>%
  summarize(sum(num_cortado) / sum(num_total))

## # A tibble: 1 x 1
##   `sum(num_cortado)/sum(num_total)`
##   <dbl>
## 1 0.0946
```

Bar plots

```
propcut_fig1 <- prop_cortado %>%
  filter(harvested == "yes") %>%
  ggplot() +
  geom_bar(aes(x = as.factor(1:26), y = proportion_cortado),
            stat = "identity", fill = "steelblue3") +
  ylim(c(0, 0.3)) +
  plot_theme +
  geom_hline(yintercept = 0) +
  labs(x = "Plot ID", y = "Proportion of individuals that are cut")
ggsave("figures/propcut_fig1.png", height = 4, width = 6)
propcut_fig1
```



```
propcut_fig2 <- prop_cortado %>%
  filter(harvested == "yes") %>%
  mutate(num_not_cortado = num_total - num_cortado) %>%
  pivot_longer(cols = c(num_cortado, num_not_cortado), names_prefix = "num_",
               names_to = "status", values_to = "num_trees") %>%
  select(plot_id, status, num_trees) %>%
  mutate(status = as.factor(ifelse(status == "cortado", "Cut", "Not cut"))) %>%
  mutate(plot_id = as.factor(rep(1:26, each = 2))) %>%
  ggplot() +
  geom_col(aes(x = plot_id, y = num_trees,
               fill = fct_relevel(status, "Not cut", "Cut"))) +
  ylim(0, 60) +
  scale_fill_brewer(palette = "Paired") +
  plot_theme + theme(legend.position = c(0.1, 0.85)) +
  labs(title = "Cut and uncut individuals in harvested plots",
       x = "Plot ID", y = "Number of individuals") +
  geom_hline(yintercept = 0) +
  guides(fill = guide_legend(title = "Status"))
ggsave("figures/propcut_fig2.png", height = 4, width = 6)
propcut_fig2
```

