

ASSP_CPUE_finalGraphs.rmd

EKelsey & ADuVall

7/8/2020

This is v.2020-09-11

STORM-PETREL MISTNETTING CPUE GRAPHS FOR FINAL NFWF REPORT

this script creates final graphs representing CPUE across years, between sites, and related to other variables

```
### SET WORKING DIRECTORY
# setwd("~/WERC-SC/ASSP_share")

## LOAD LIBRARIES
library(tidyr)
library(dplyr)
library(lubridate)
library(hms)
library(tidyverse)
library(ggplot2)
library(EnvStats)
library(here)
library(calecopal)

### READ IN DATA
banding <- readRDS(here("Working", "captures.RDS"))
metadata <- readRDS(here("Working", "cpue.RDS"))

## Filter data to ASSP species and remove SNRs
ASSP <- group_by(.data = banding) %>%
  filter(species == "ASSP" & recapture != "SNR") %>%
  #filter(band_no != "notbanded") %>%
  ungroup() # 3815 observations

unbanded <- banding %>%
  filter(species == "ASSP" & band_no == "notbanded") # 20 unbanded ASSP
```

summary statistics

CPUE

```
# CPUEraw
summary(metadata$CPUEraw)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
## 0.00282 0.03313 0.05255 0.06372 0.08795 0.26316      35

# CPUEstd
summary(metadata$CPUEstd)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
## 0.00357 0.03817 0.06310      Inf 0.10136      Inf      37

## inf value

# broodpatch frequency
summary(metadata$BPfreq_Y)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
## 0.0000 0.3000 0.4667 0.4824 0.6364 1.0000      16

summary(metadata$BPfreq_N)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
## 0.0000 0.3636 0.5333 0.5176 0.7000 1.0000      16

sites <- metadata %>%
  group_by(island_code, site_name) %>%
  summarise(n = n()) %>%
  arrange(desc(n))

## 'summarise()' regrouping output by 'island_code' (override with '.groups' argument)

months <- metadata %>%
  group_by(session_month) %>%
  summarise(n = n()) %>%
  arrange(desc(n))

## 'summarise()' ungrouping output (override with '.groups' argument)
```

Captures

```
species <- banding %>%
  group_by(species) %>%
  summarize(n=n()) %>%
  arrange(desc(n))

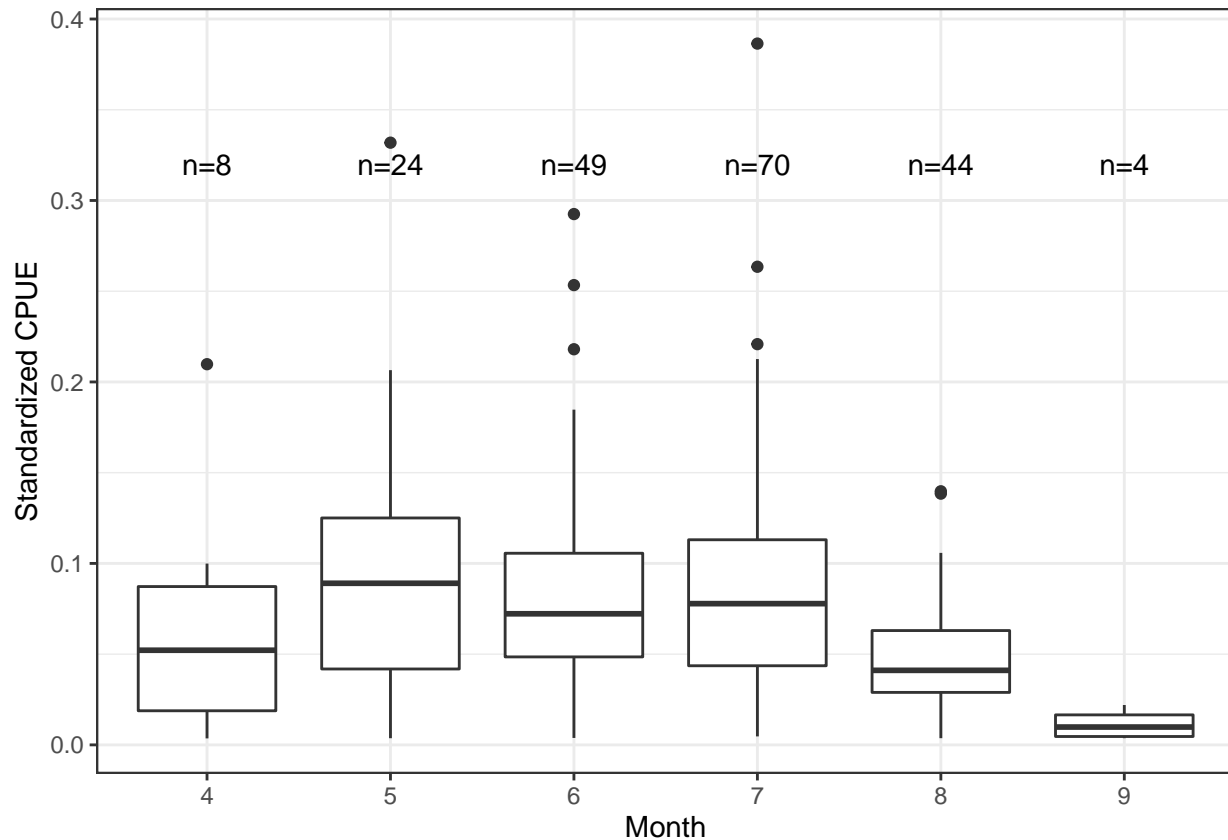
## 'summarise()' ungrouping output (override with '.groups' argument)
```

CPUE graphs

month and year

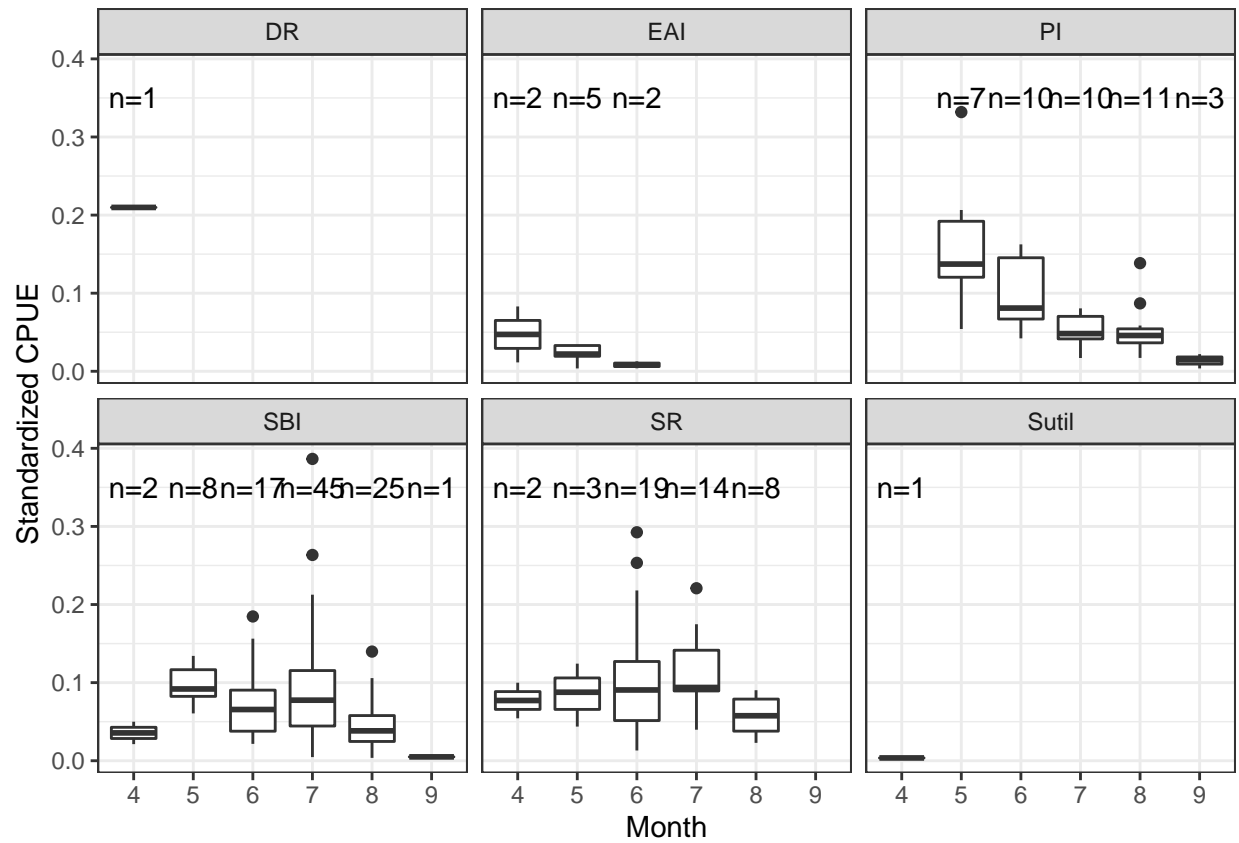
MONTH

```
month <- ggplot(metadata, aes(as.character(session_month), CPUEstd)) +  
  geom_boxplot() +  
  stat_n_text(y.pos = 0.32) +  
  xlab("Month") + ylab("Standardized CPUE") +  
  theme_bw()  
month
```

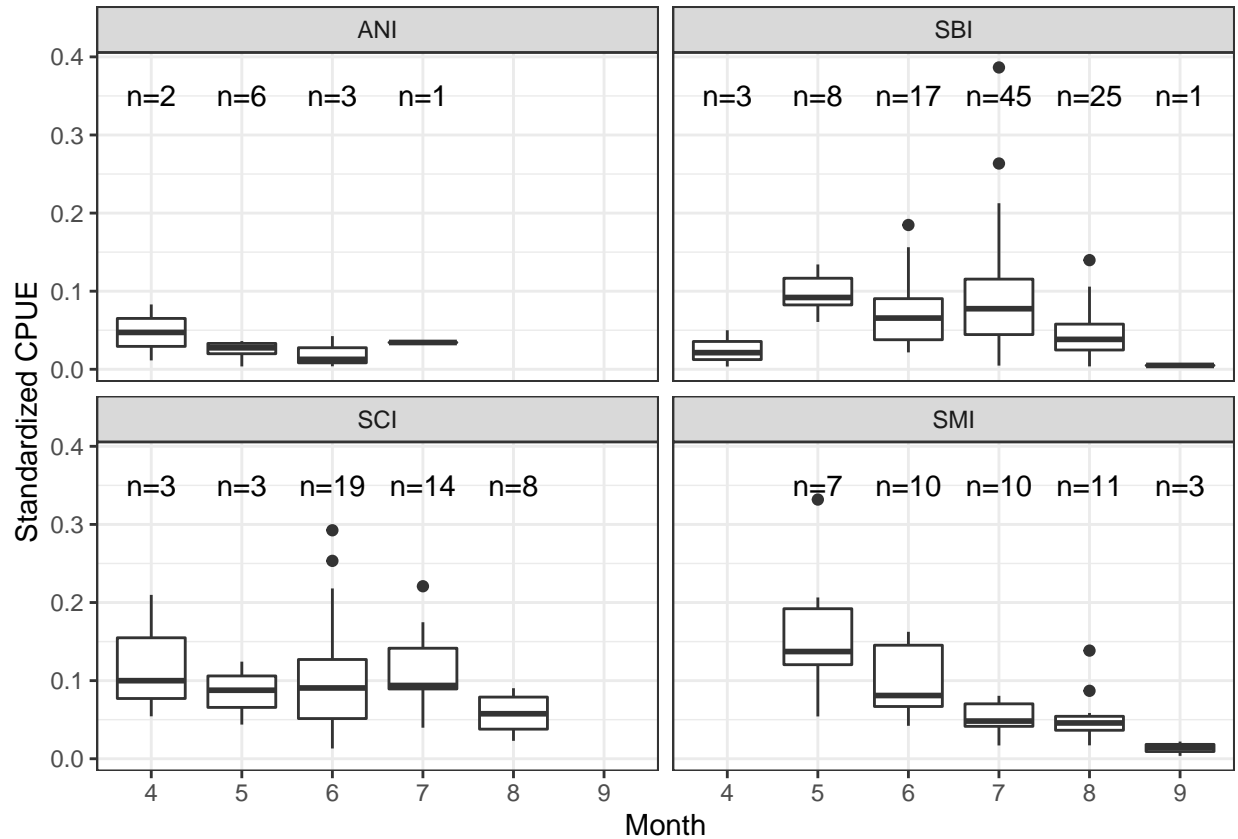


This graph show how catch rates varied across months

```
# sub-island  
month_subisl <- ggplot(metadata, aes(as.character(session_month), CPUEstd)) +  
  geom_boxplot() +  
  facet_wrap(~subisland_code) +  
  xlab("Month") + ylab("Standardized CPUE") +  
  stat_n_text(y.pos = 0.35) +  
  theme_bw()  
month_subisl %>% subset(metadata, subisland_code %in% c("DR", "EAI", "PI", "SBI", "SR", "Sutil"))
```



```
# island
month_isl <- ggplot(metadata, aes(as.character(session_month), CPUEstd)) +
  geom_boxplot() +
  facet_wrap(~island_code) +
  xlab("Month") + ylab("Standardized CPUE") +
  stat_n_text(y.pos = 0.35) +
  theme_bw()
month_isl
```



This graph shows catch rates by month for all locations combined and then each sub-island and island separately

MONTH AND ASSUMED BREEDERS

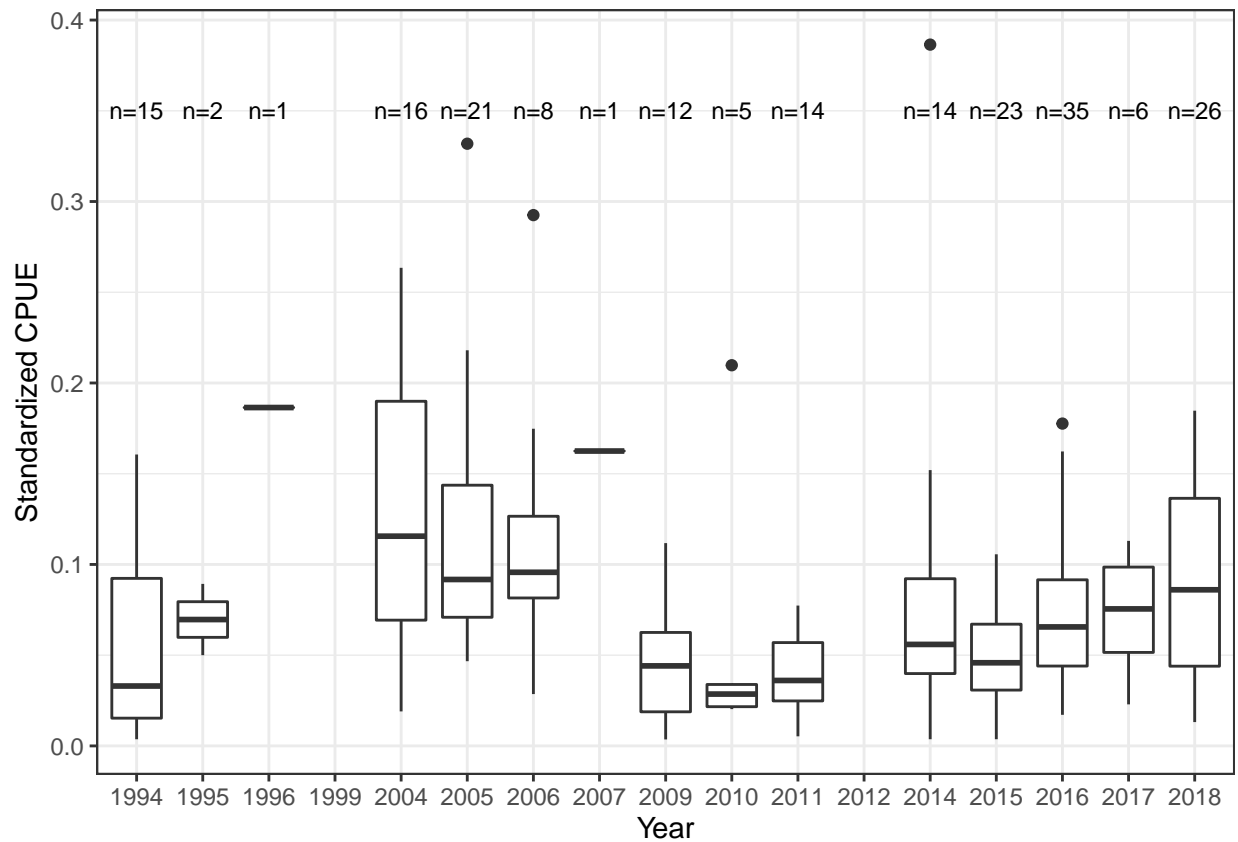
```
# metadata_brd <- metadata %>%
#   dplyr::select(CPUEstd, CPUEbrd_std, session_month)
# ## CPUEbrd_std are brood patches w/in the std effort v]
#
# metadata_brd <- gather(metadata_brd, "CPUEtype", "CPUE", CPUEstd:CPUEbrd_std)
#
# month_brd <- ggplot(metadata_brd, aes(session_month, CPUE, color = CPUEtype)) +
#   geom_boxplot() +
#   stat_n_text(y.pos = 0.3) +
#   xlab("Month") + ylab("Standardized CPUE") +
#   theme_bw()
# month_brd
```

This graphs shows the total CPUE each month, as well as the CPUE of assumed breeders (based on broodpatch score) NOTE: the samples sizes at the top are the sum of CPUEbrd_std and CPUEstd

YEAR

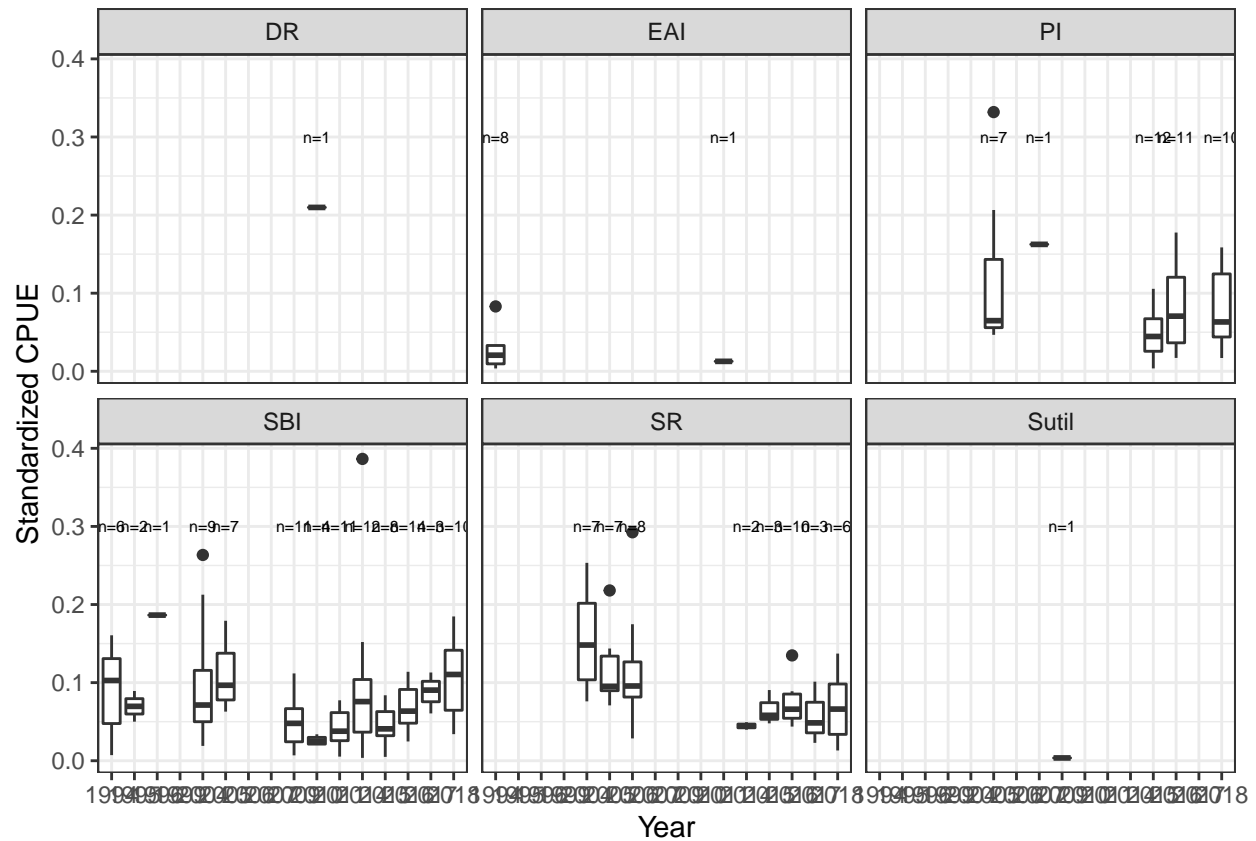
```
year <- ggplot(metadata, aes(as.character(session_year), CPUEstd)) +
  geom_boxplot() +
  xlab("Year") + ylab("Standardized CPUE") +
```

```
stat_n_text(size = 3, y.pos = 0.35) +
theme_bw()
year
```

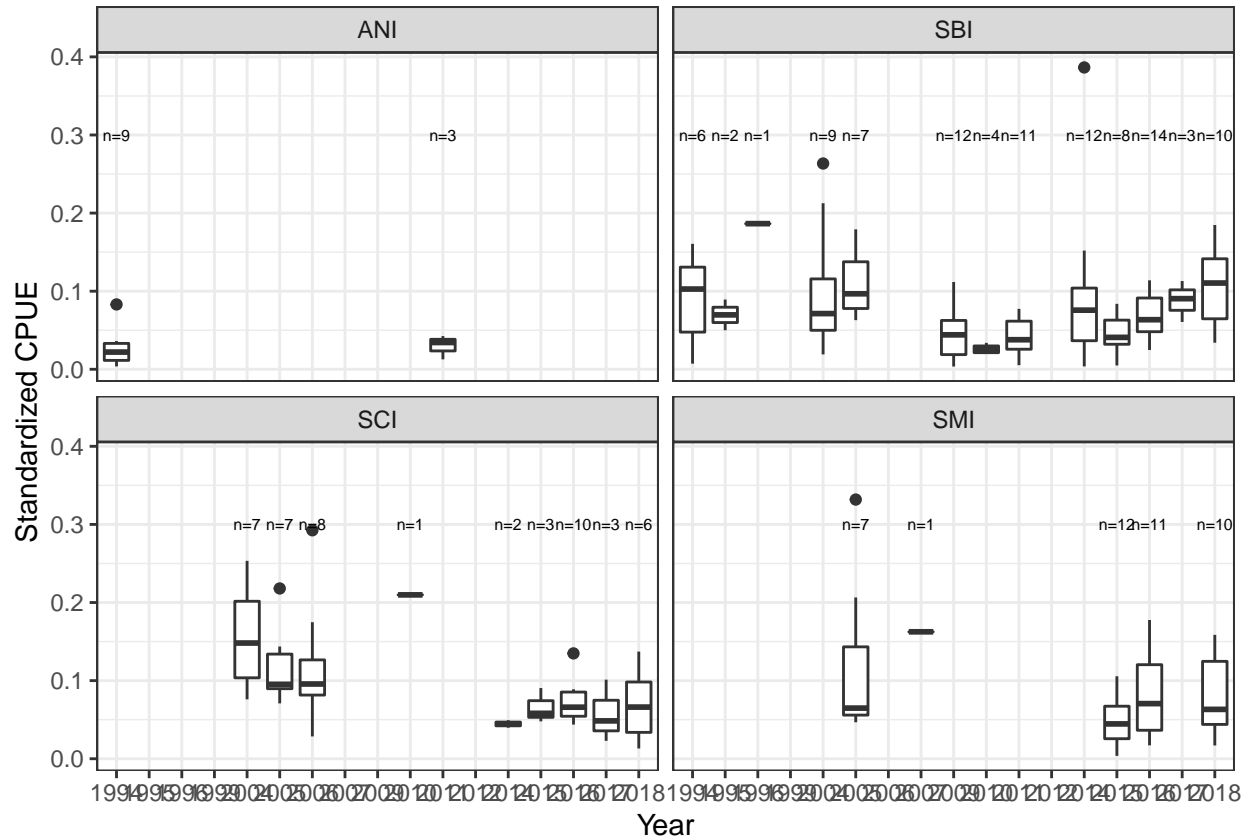


This graph show how catch rates varied across years for all sites combined

```
# by subisland
year_subisl <- ggplot(metadata, aes(as.character(session_year), CPUEstd)) +
  geom_boxplot() +
  facet_wrap(~subisland_code) +
  xlab("Year") + ylab("Standardized CPUE") +
  stat_n_text(size = 2, y.pos = 0.3) +
  theme(axis.text.x = element_text(angle = -45)) + # not working?
  theme_bw()
year_subisl %>% subset(metadata, subisland_code %in% c("DR", "EAI", "PI", "SBI", "SR", "Sutil"))
```



```
# by island
year_isl <- ggplot(metadata, aes(as.character(session_year), CPUEstd)) +
  geom_boxplot() +
  facet_wrap(~island_code) +
  xlab("Year") + ylab("Standardized CPUE") +
  stat_n_text(size = 2, y.pos = 0.3) +
  theme(axis.text.x = element_text(angle = -45)) + # not working?
  theme_bw()
year_isl
```



This graph shows how catch rates vary across years for sub-island separately NOTE: This graph is broken out by sub-island which is intermediate between “island” and “site”. Using “island” or “site” for this graph might ultimately make more sense

YEAR AND ASSUMED BREEDERS

```
# metadata_brd_yr <- metadata %>%
#   select(CPUEstd, CPUEbrd_std, year)
# metadata_brd_yr <- gather(metadata_brd_yr, "CPUEtype", "CPUE", CPUEstd:CPUEbrd_std)
#
# year_brd <- ggplot(metadata_brd_yr, aes(session_year, CPUE, color = CPUEtype)) +
#   geom_boxplot() +
#   xlab("Year") + ylab("Standardized CPUE") +
#   stat_n_text(size = 3, y.pos = 0.3) +
#   theme(axis.text.x = element_text(angle = -45))
# year_brd
```

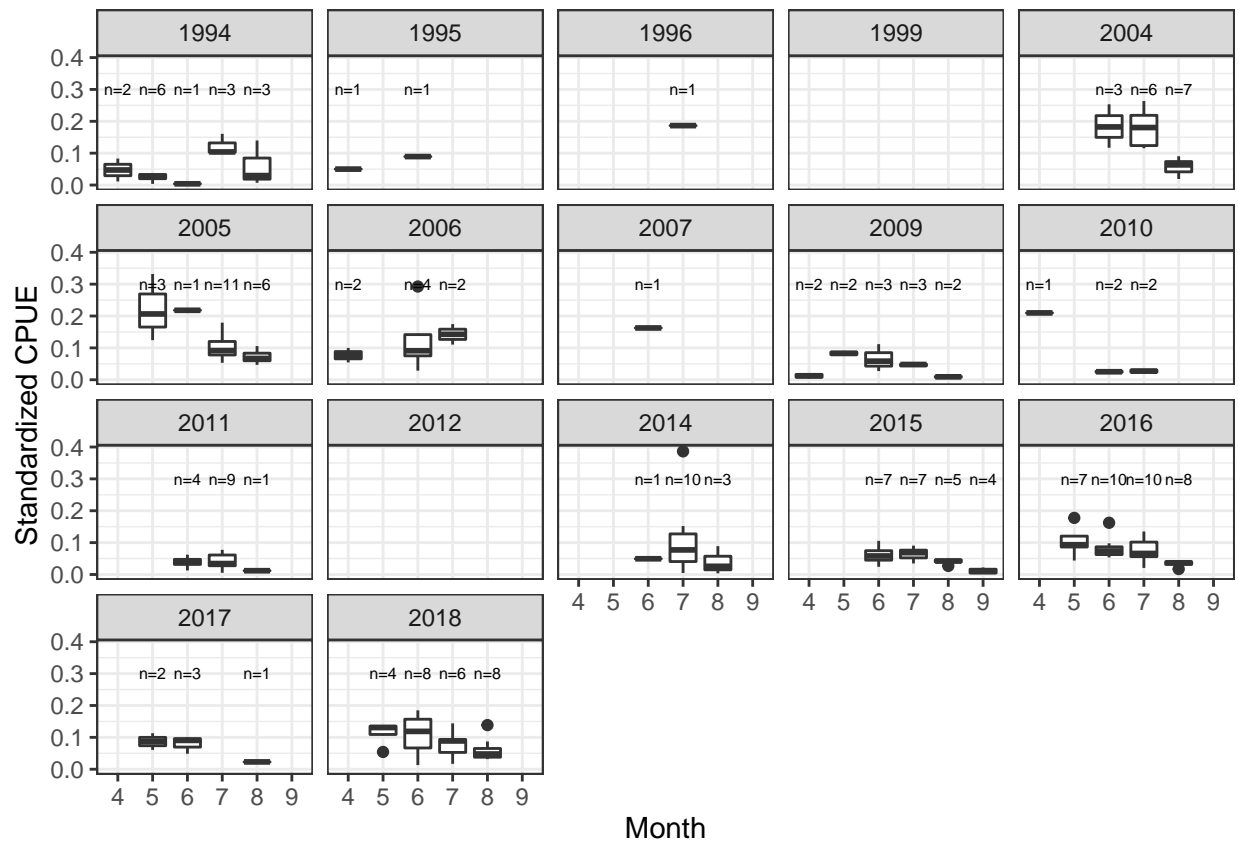
This graph shows annual CPUE compared to the CPUE of assumed breeders (based on broodpatch score) NOTE: the samples sizes at the top are the sum of CPUEbrd_std and CPUEstd

MONTH VS YEAR

```
month_year <- ggplot(metadata, aes(as.character(session_month), CPUEstd)) +
  geom_boxplot() +
  facet_wrap(~session_year) +
  xlab("Month") + ylab("Standardized CPUE") +
```



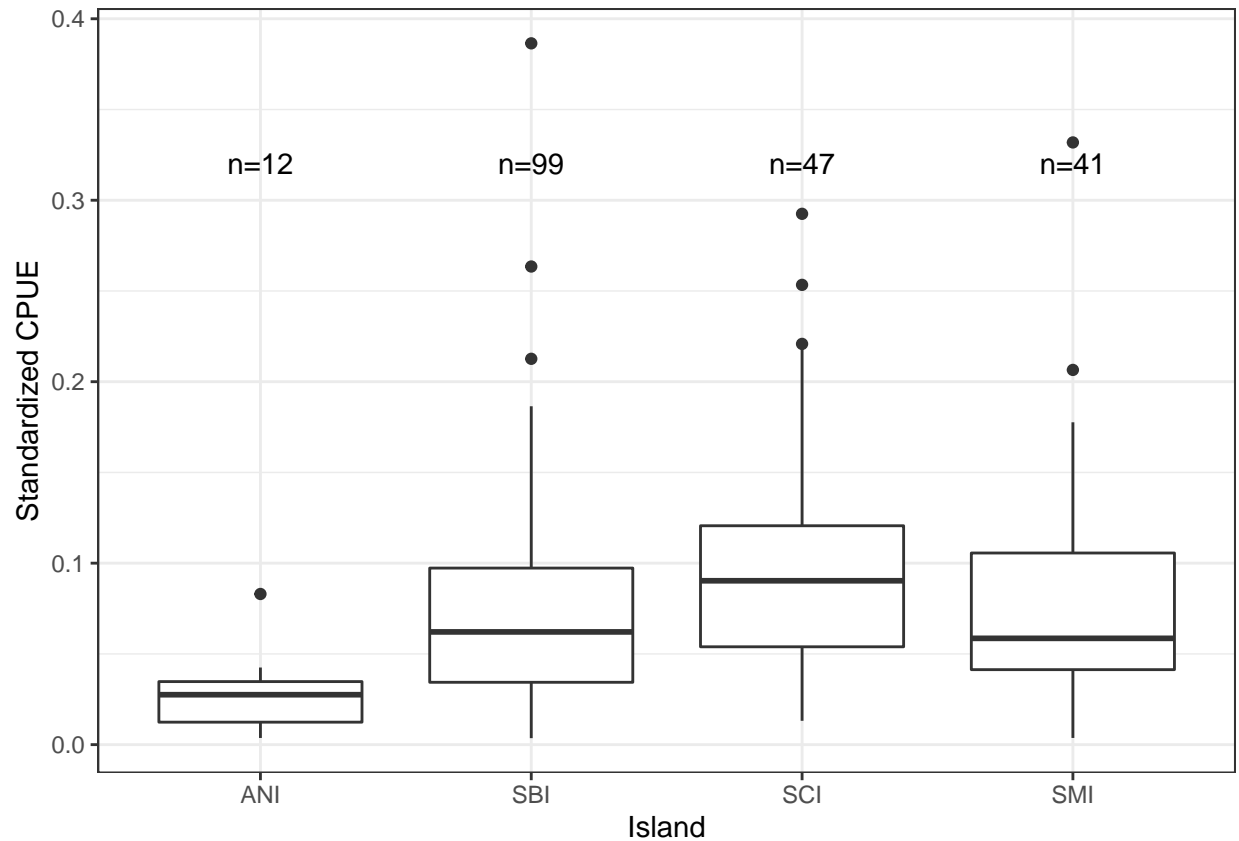
```
stat_n_text(size = 2, y.pos = 0.3) +
theme_bw()
month_year
```



This graph shows how catches varied across months, for each year separately. Most years do not have enough data to show much of a pattern

CATCHES AT ISLANDS AND SITES THROUGH TIME

```
island <- ggplot(metadata, aes(island_code, CPUEstd)) +
  geom_boxplot() +
  xlab("Island") + ylab("Standardized CPUE") +
  stat_n_text(size = 4, y.pos = 0.32) +
  theme_bw()
island
```

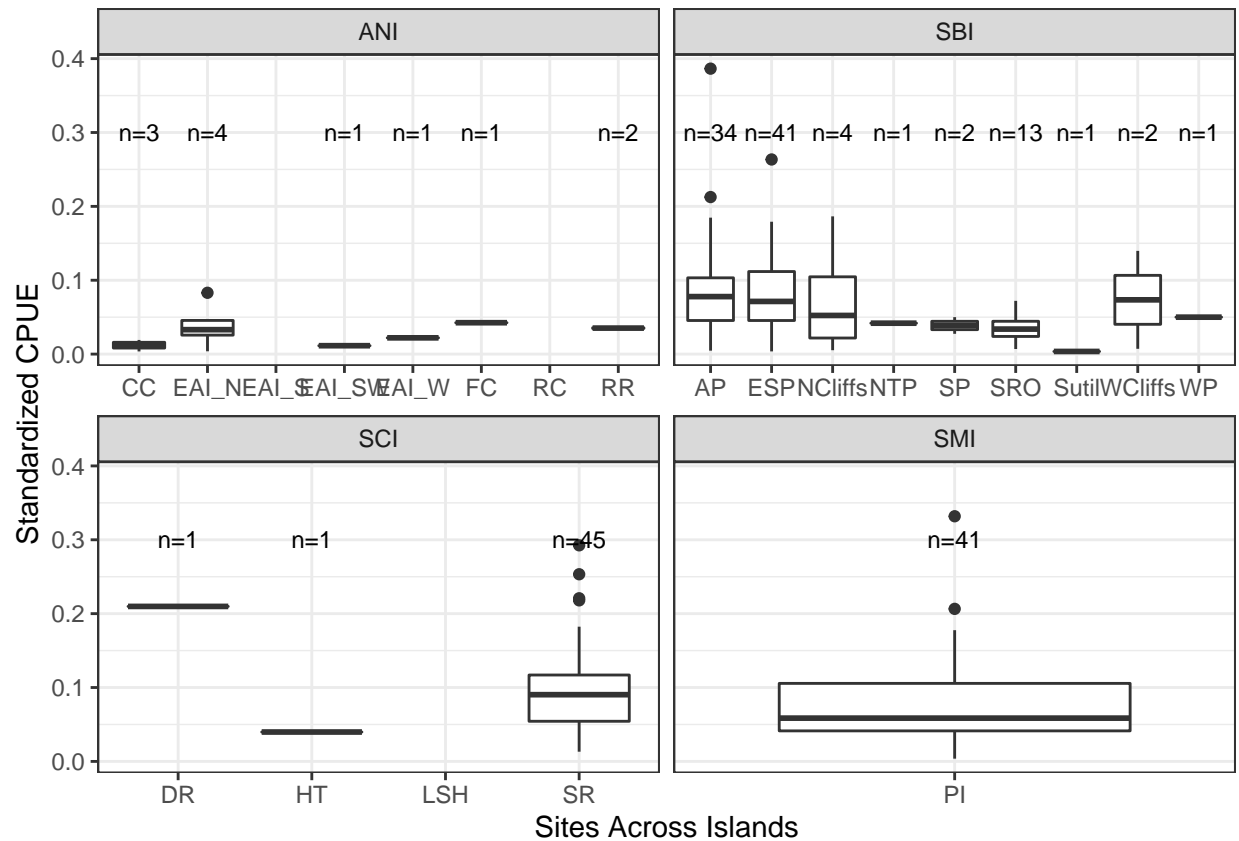


This graphs show how catch rates varied across islands.

```
site <- ggplot(metadata, aes(site_code, CPUEstd)) +
  geom_boxplot() +
  facet_wrap(~island_code, scales = "free_x") +
  xlab("Sites Across Islands") + ylab("Standardized CPUE") +
  stat_n_text(size = 3, y.pos = 0.3) +
  theme_bw()
site # %>% subset(metadata, site_code %in% c("DR", "EAI", "PI", "SBI", "SR", "Sutil"))
```

Warning: Removed 38 rows containing non-finite values (stat_boxplot).

Warning: Removed 38 rows containing non-finite values (stat_n_text).

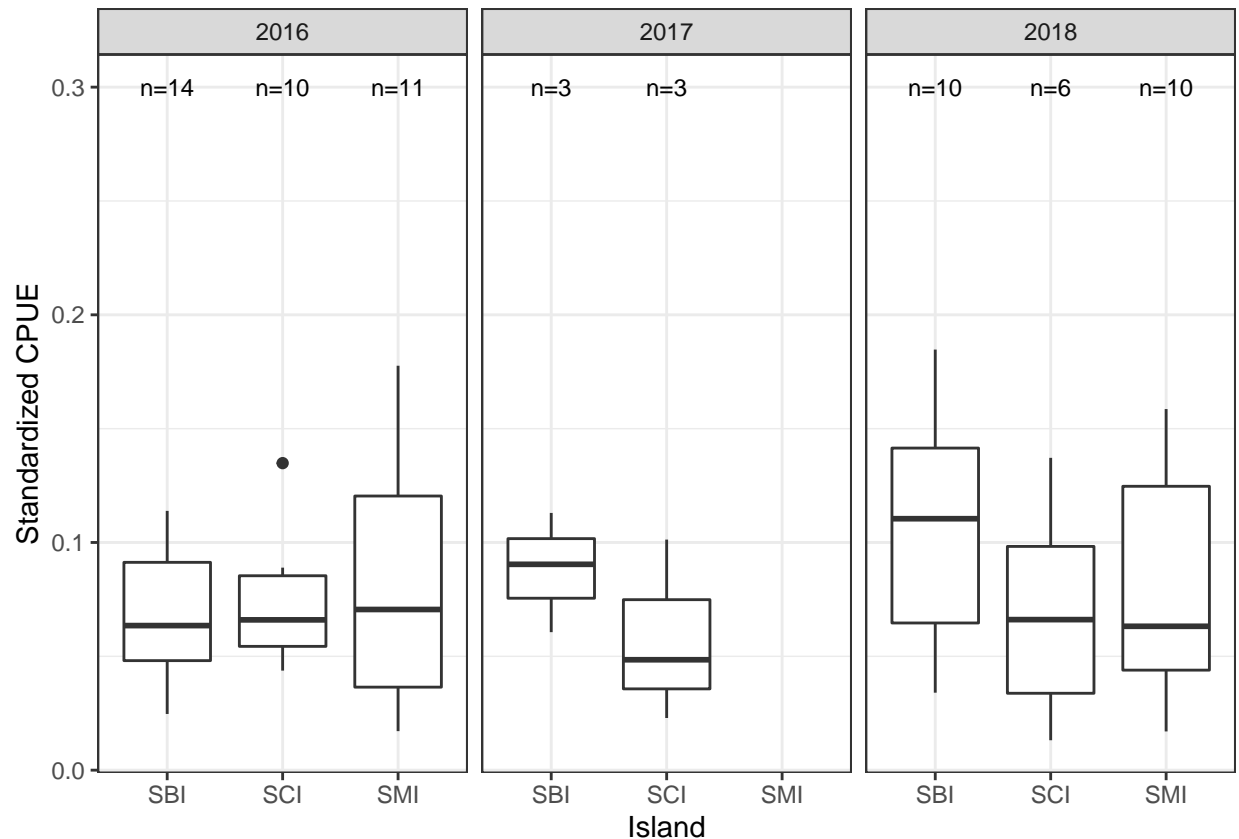


This graph shows how catch rates varied across all sites within islands.

```
isl_yr <- ggplot(metadata, aes(island_code, CPUEstd)) +
  geom_boxplot() +
  facet_wrap(~session_year) +
  xlab("Island") + ylab("Standardized CPUE") +
  stat_n_text(size = 3, y.pos = 0.3) +
  theme_bw()
isl_yr %>% subset(metadata, session_year %in% c("2016", "2017", "2018"))
```

Warning: Removed 3 rows containing non-finite values (stat_boxplot).

Warning: Removed 3 rows containing non-finite values (stat_n_text).

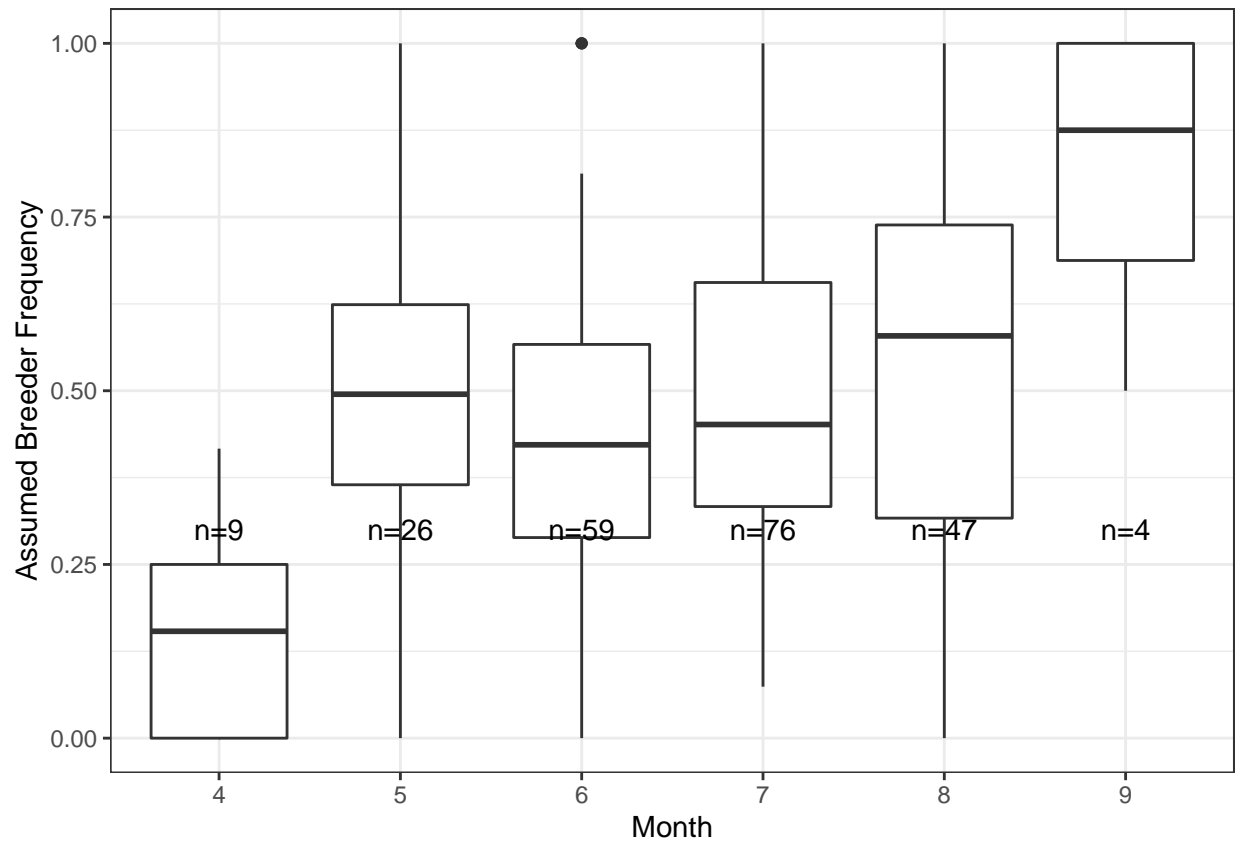


This graph shows how catches varied across islands in 2016 - 2018, the years in which acoustic data was collected. This graph is comparable to the call rate outputs graphs being generated by Kerry.

ISLAND AND ASSUMED BREEDERS

```
# metadata_brd_isl <- metadata %>%
#   select(CPUEstd, CPUEbrd_std, island_code)
# metadata_brd_isl <- gather(metadata_brd_isl, "CPUEtype", "CPUE", CPUEstd:CPUEbrd_std)
#
# island_brd <- ggplot(metadata_brd_isl, aes(island_code, CPUE, color = CPUEtype)) +
#   geom_boxplot() +
#   xlab("Island") + ylab("Standardized CPUE of Assumed Breeders") +
#   stat_n_text(size = 4, y.pos = 0.3) +
#   theme_bw()
# island_brd # %>% subset(metadata, subisland_code %in% c("DR", "EAI", "PI", "SBI", "SR", "Sutil"))

month_BP <- ggplot(metadata, aes(as.character(session_month), BPfreq_Y)) +
  geom_boxplot() +
  stat_n_text(y.pos = 0.3) +
  xlab("Month") + ylab("Assumed Breeder Frequency") +
  theme_bw()
month_BP
```



```
## need to fix text position
```

This graph shows the average CPUE at on each island and also the average CPUE of assumed breeders (based on broodpatch score) NOTE: the samples sizes at the top are the sum of CPUEbrd_std and CPUEstd

```
metadata_effort <- metadata %>%
  select(site_code, session_ID, app_sunset, std_ending, lat, long)

catches <- banding %>%
  left_join(metadata_effort, by = c("site_code", "session_ID", "lat", "long")) %>%
  mutate(catchPastSS = app_sunset - capture_date)

# std = if_else(std_ending > capture_date, "1", "0"),
# #          captureT = hms(capture_date),
```

CATCH TIMES

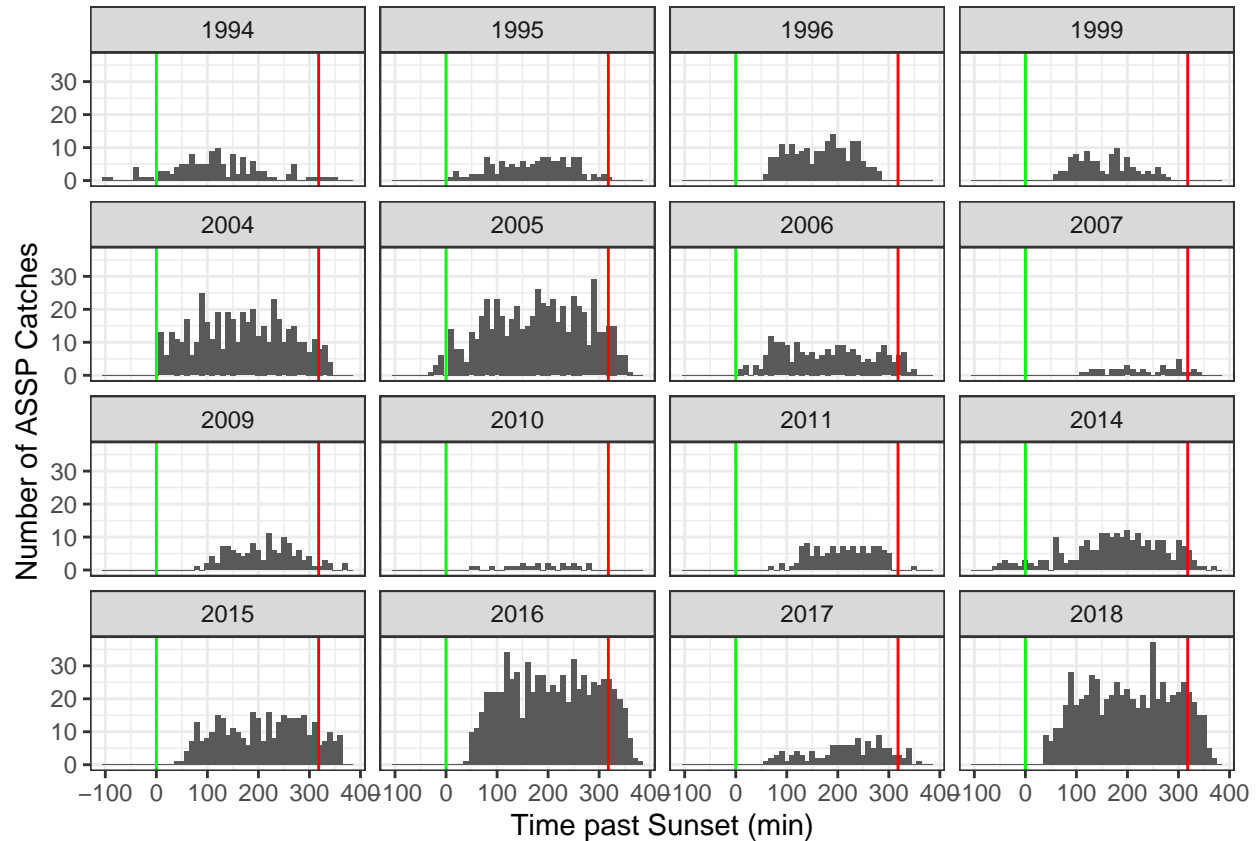
catch times vs. 5.3 hours post SS, subset by year

```
endT_yr <- ggplot(catches, aes(catchPastSS)) +
  geom_histogram(binwidth = 10) +
  geom_vline(xintercept = 318, color = "red") +
  geom_vline(xintercept = 0, color = "green") +
  xlab("Time past Sunset (min)") + ylab("Number of ASSP Catches") +
  facet_wrap(~ session_year) +
```

```
theme_bw()
endT_yr
```

```
## Don't know how to automatically pick scale for object of type difftime. Defaulting to continuous.
```

```
## Warning: Removed 29 rows containing non-finite values (stat_bin).
```



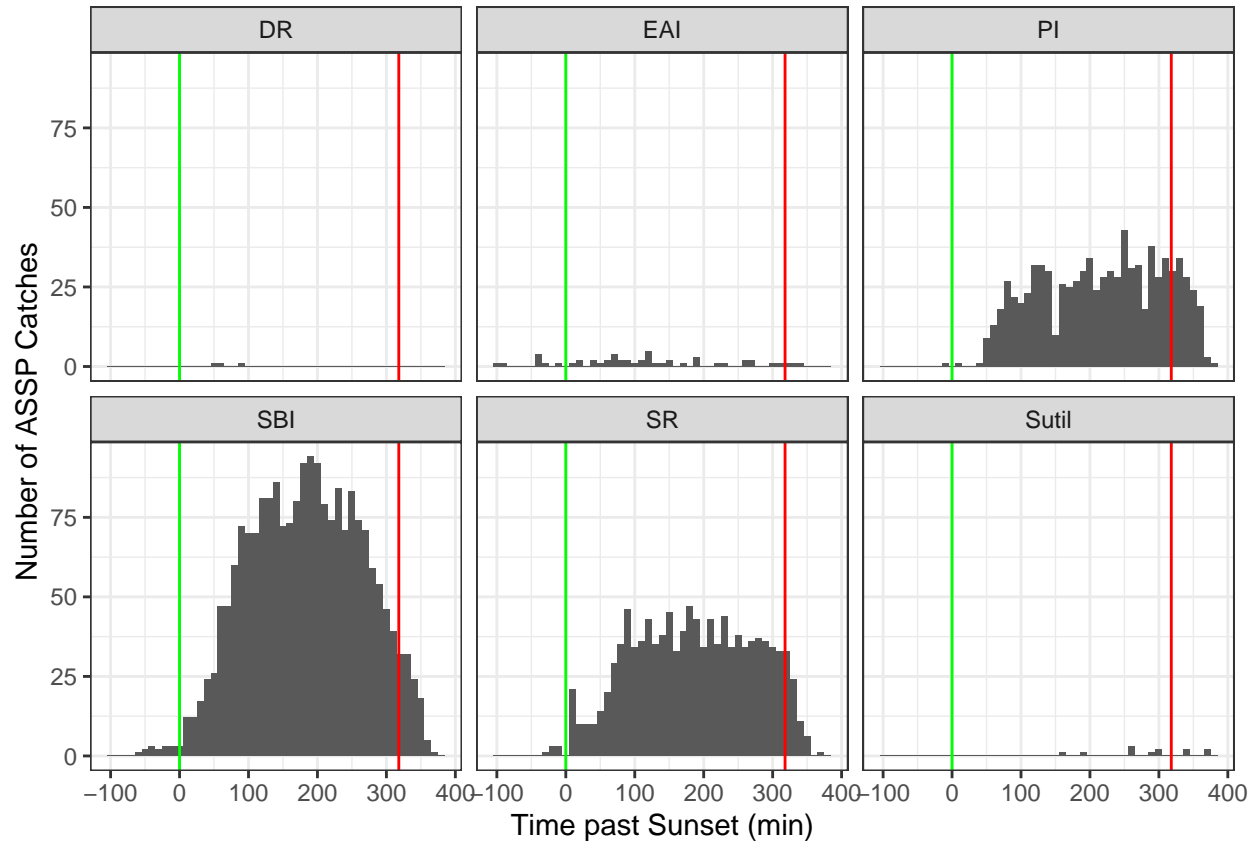
This graph shows the time after sunset of ASSP catches, broken out by year. The green vertical line indicates sunset, the red vertical line indicates 5.3 hours after sunset, the cutoff time indicated by Adams (2016). NOTE: in some years there were catches recorded before sunset. This probably isn't accurate and is a bug in how time was calculated in R. Needs review.

catch times vs. 5.3 hours post SS, subset by island

```
endT_isl <- ggplot(catches, aes(catchPastSS)) +
  geom_histogram(binwidth = 10) +
  geom_vline(xintercept = 318, color = "red") +
  geom_vline(xintercept = 0, color = "green") +
  xlab("Time past Sunset (min)") + ylab("Number of ASSP Catches") +
  facet_wrap(~ subisland_code) +
  theme_bw()
endT_isl %>% subset(catches, subisland_code %in% c("DR", "EAI", "PI", "SBI", "SR", "Sutil"))
```

```
## Don't know how to automatically pick scale for object of type difftime. Defaulting to continuous.
```

```
## Warning: Removed 1 rows containing non-finite values (stat_bin).
```



This graph shows the time after sunset of ASSP catches, broken out by sub-island. The green vertical line indicates sunset, the red vertical line indicates 5.3 hours after sunset, the cutoff time indicated by Adams (2016). NOTE: in some years there were catches recorded before sunset. This probably isn't accurate and is a bug in how time was calculated in R. Needs review.

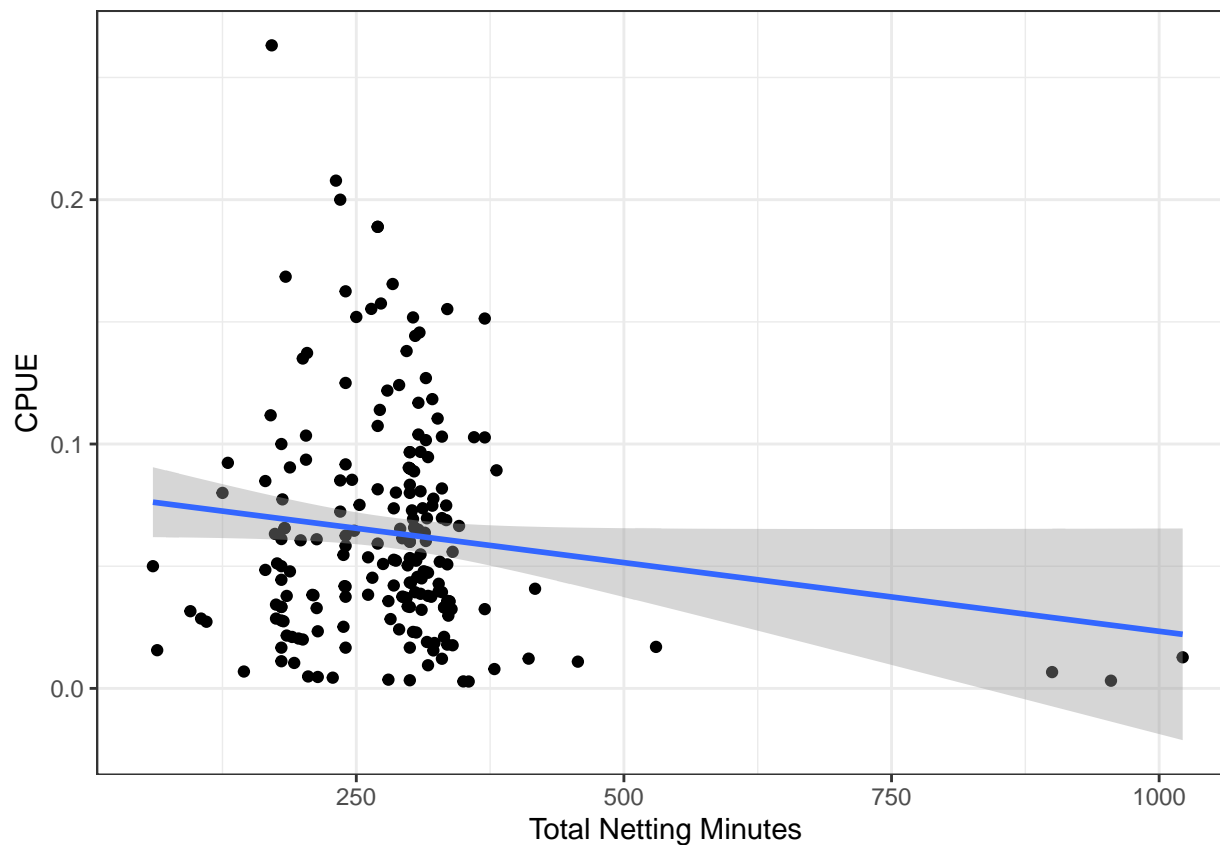
CPUE vs. cumulative mintues

```
CPUE_min <- ggplot(metadata, aes(min, CPUEraw)) +
  geom_point() +
  geom_smooth(method = 'lm') +
  xlab("Total Netting Minutes") + ylab("CPUE") +
  theme_bw()
CPUE_min
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

```
## Warning: Removed 35 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 35 rows containing missing values (geom_point).
```



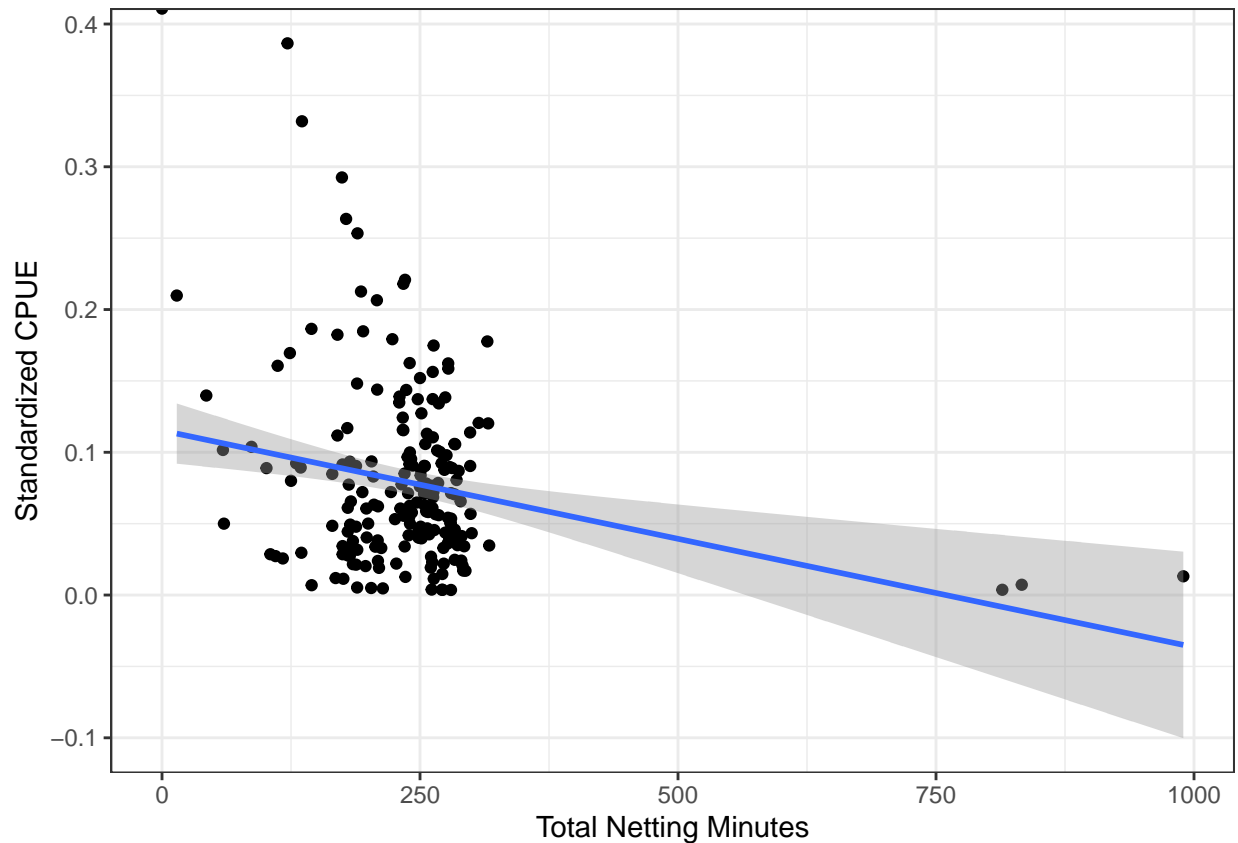
This graphs show the CPUE in relation to the number of minutes the net was open. The blue line shows the linear regression of these points.

```
CPUE_minstd <- ggplot(metadata, aes(min_std, CPUEstd)) +
  geom_point() +
  geom_smooth(method = 'lm') +
  xlab("Total Netting Minutes") + ylab("Standardized CPUE") +
  theme_bw()
CPUE_minstd
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

```
## Warning: Removed 38 rows containing non-finite values (stat_smooth).
```

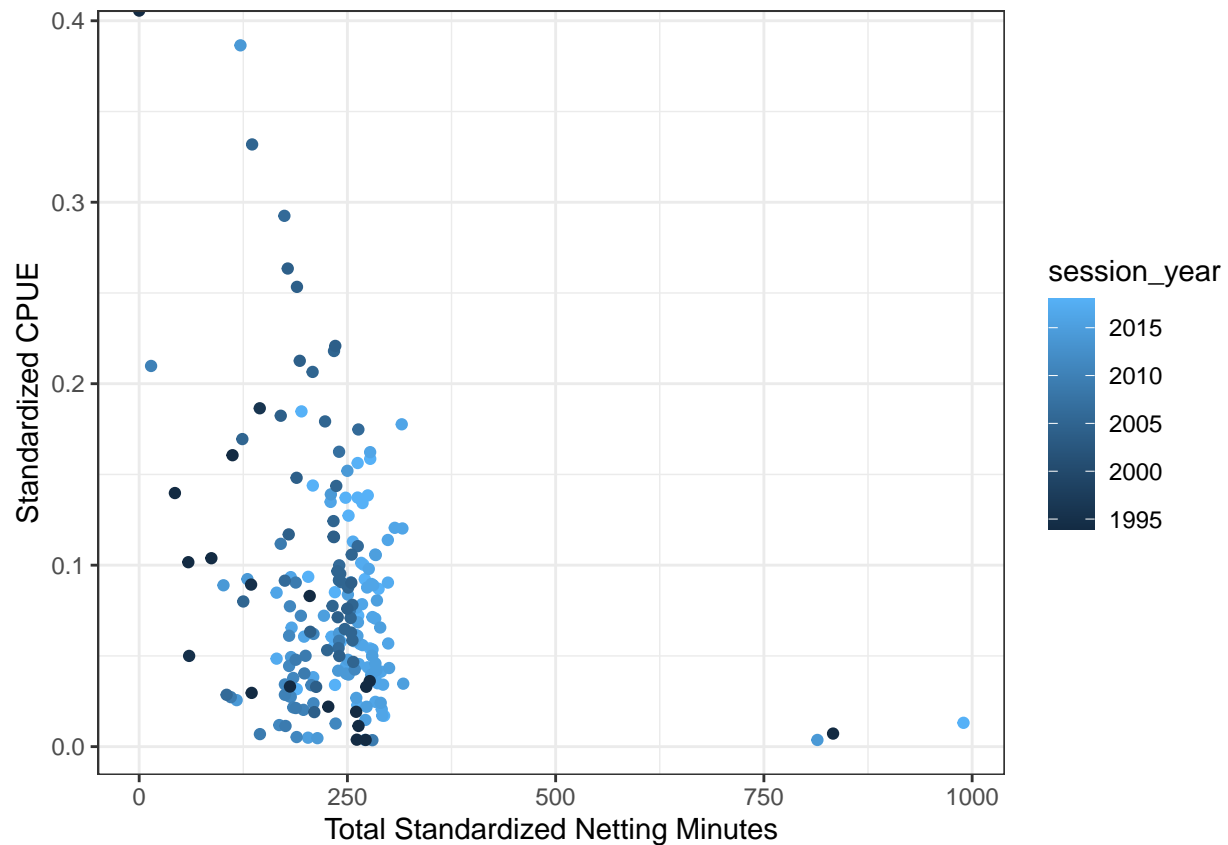
```
## Warning: Removed 37 rows containing missing values (geom_point).
```

this graph shows the standardized CPUE (until 5.3 hours after sunset) in relation to the total number of standardized minutes the net was open. The blue line shows the linear regression of these points.

```
CPUE_minstd_c <- ggplot(metadata, aes(min_std, CPUEstd, color = session_year)) + #
  geom_point() +
  # geom_smooth(method = 'lm') +
  xlab("Total Standardized Netting Minutes") + ylab("Standardized CPUE") +
  # scale_fill_brewer(palette="Dark2")
  theme_bw()
CPUE_minstd_c
```

```
## Warning: Removed 37 rows containing missing values (geom_point).
```



This graph shows the same scatter plot of standardized CPUE, broken out by color

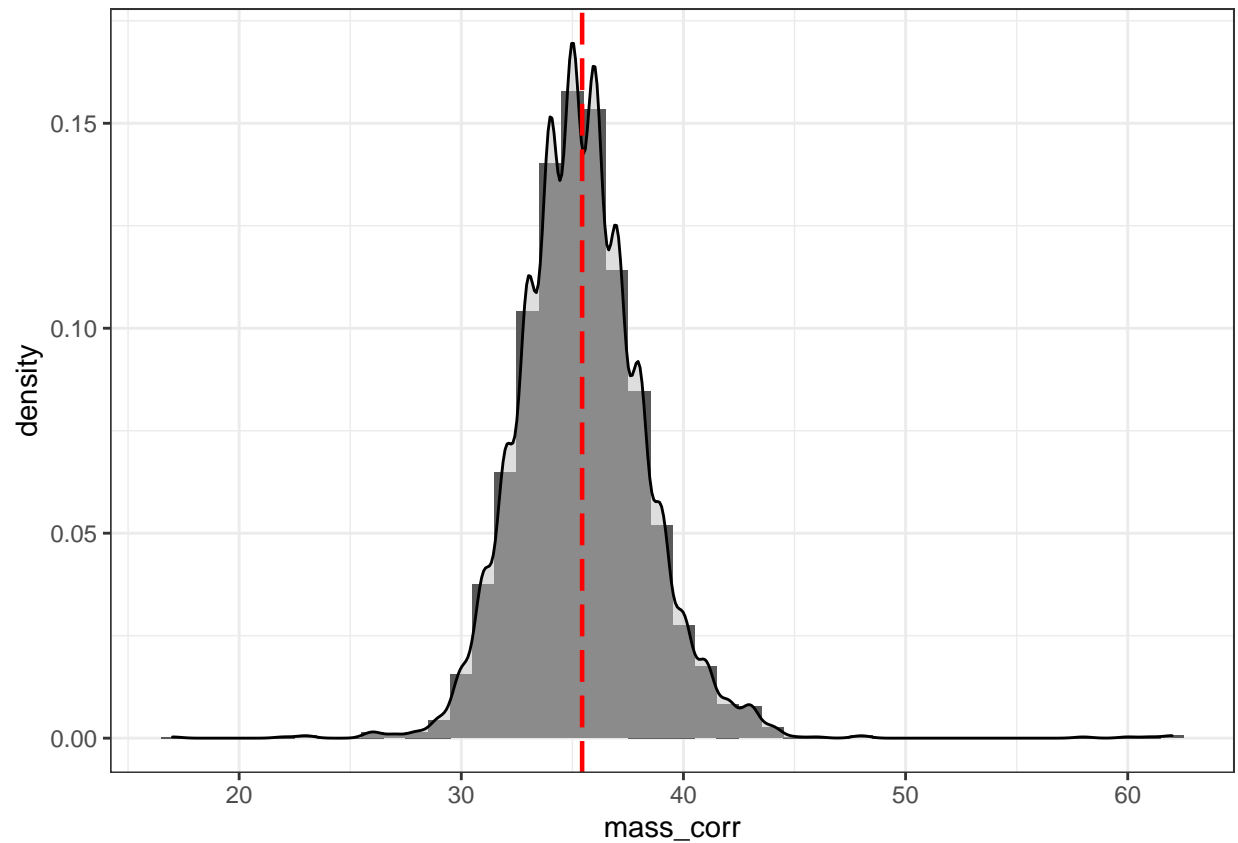
MORPHOMETRIC DATA

Morphometrics across islands

```
ggplot(data = ASSP, aes(x = mass_corr), na.rm = TRUE) +
  geom_histogram(aes(y = ..density..), binwidth = 1) +
  geom_density(alpha = .5, fill = "gray") +
  geom_vline(aes(xintercept = mean(mass_corr, na.rm = TRUE)),
    colour = "red", linetype = "longdash", size = .8) +
  theme_bw()
```

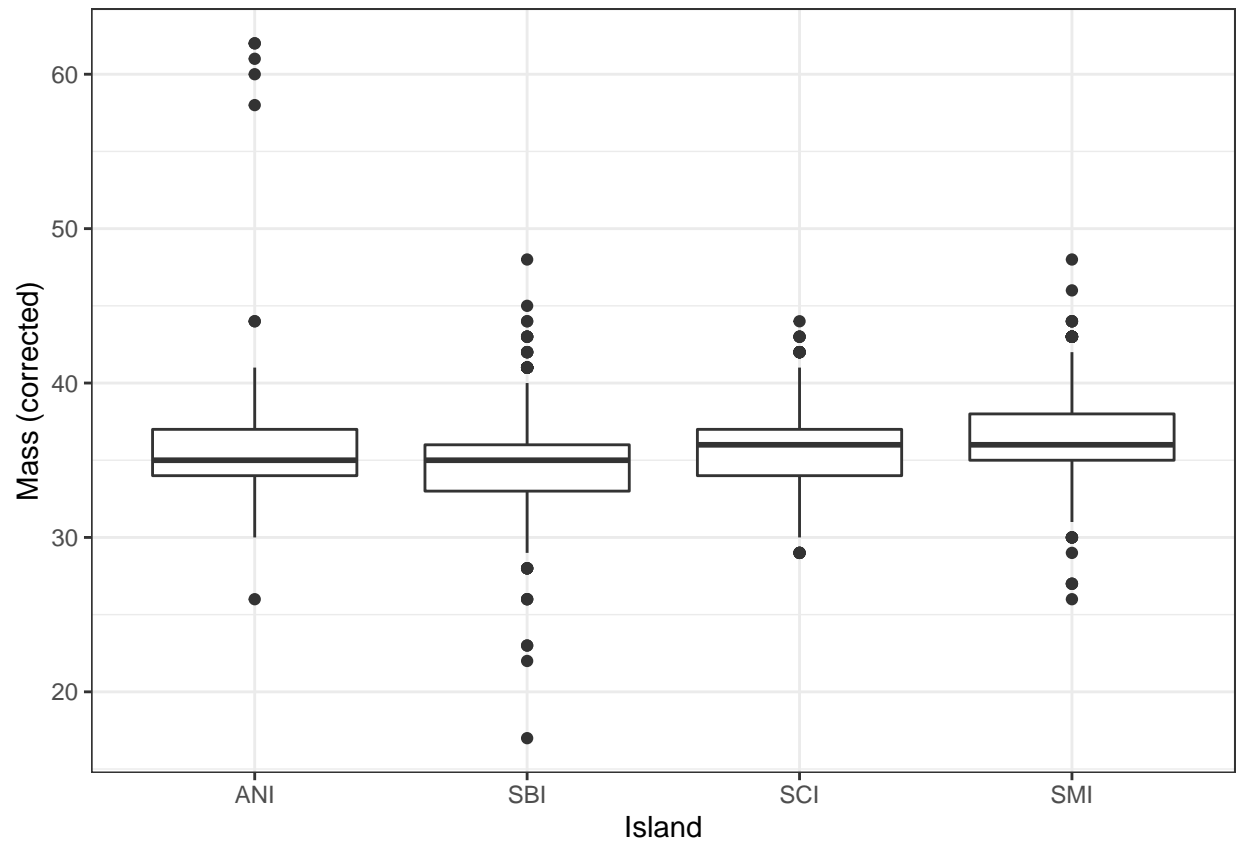
```
## Warning: Removed 450 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 450 rows containing non-finite values (stat_density).
```



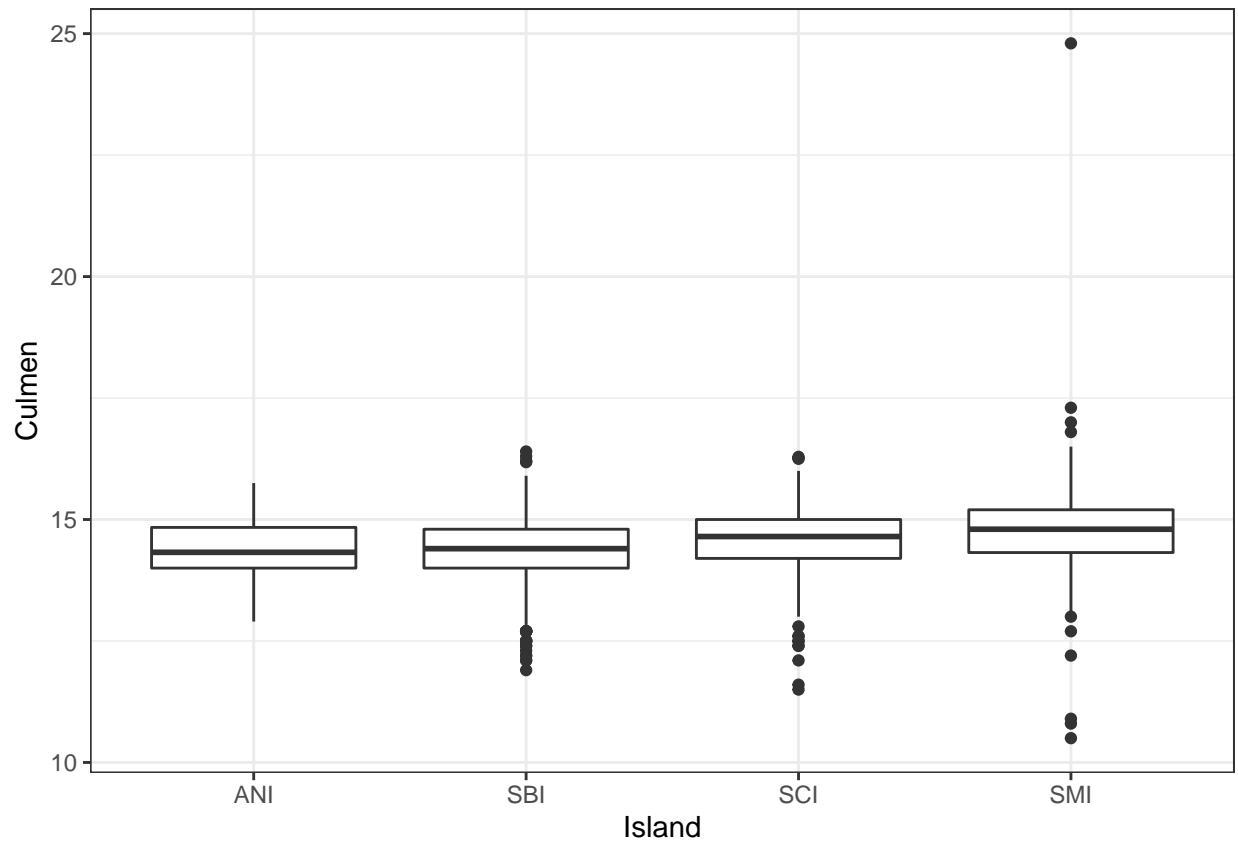
```
mass <- ggplot(data = ASSP, aes(x = island_code, y = mass_corr)) +
  geom_boxplot() +
  xlab("Island") + ylab("Mass (corrected)") +
  theme_bw()
mass
```

```
## Warning: Removed 450 rows containing non-finite values (stat_boxplot).
```



```
culmen <- ggplot(data = ASSP, aes(x = island_code, y = culmen)) +
  geom_boxplot() +
  xlab("Island") + ylab("Culmen") +
  theme_bw()
culmen
```

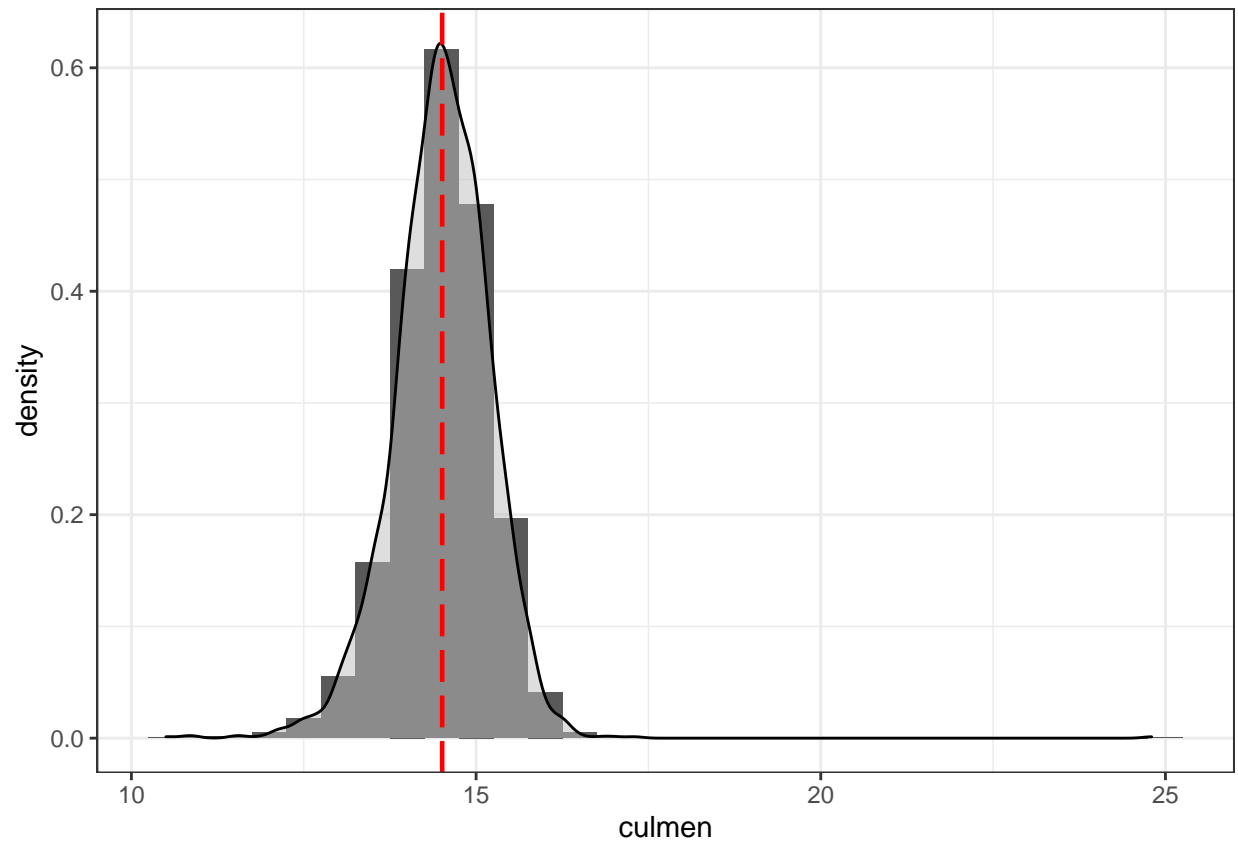
```
## Warning: Removed 1209 rows containing non-finite values (stat_boxplot).
```



```
ggplot(data = ASSP, aes(x = culmen), na.rm = TRUE) +
  geom_histogram(aes(y = ..density..), binwidth = 0.5) +
  geom_density(alpha = .5, fill = "gray") +
  geom_vline(aes(xintercept = mean(culmen, na.rm = TRUE)),
    colour = "red", linetype = "longdash", size = .8) +
  theme_bw()
```

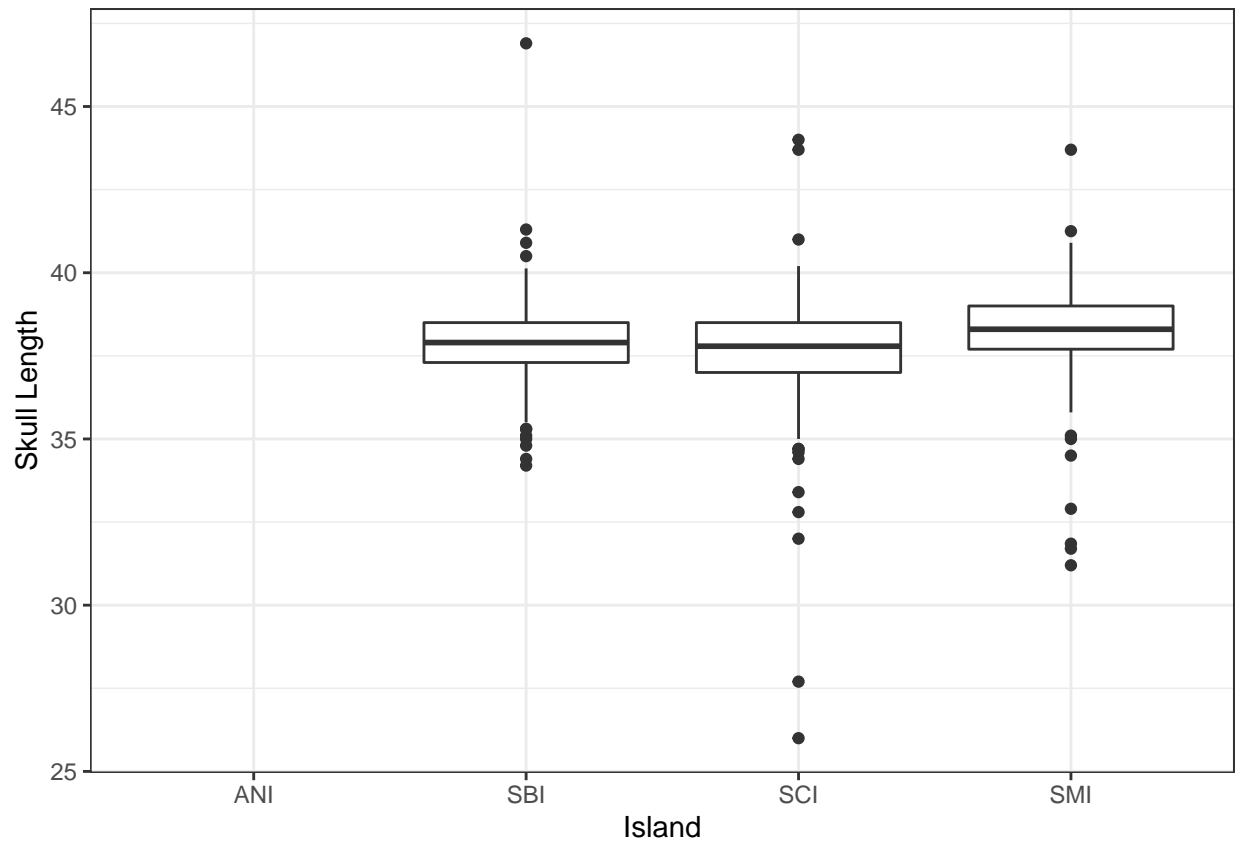
```
## Warning: Removed 1209 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 1209 rows containing non-finite values (stat_density).
```



```
skulllength <- ggplot(data = ASSP, aes(x = island_code, y = skull_length)) +  
  geom_boxplot() +  
  xlab("Island") + ylab("Skull Length") +  
  theme_bw()  
skulllength
```

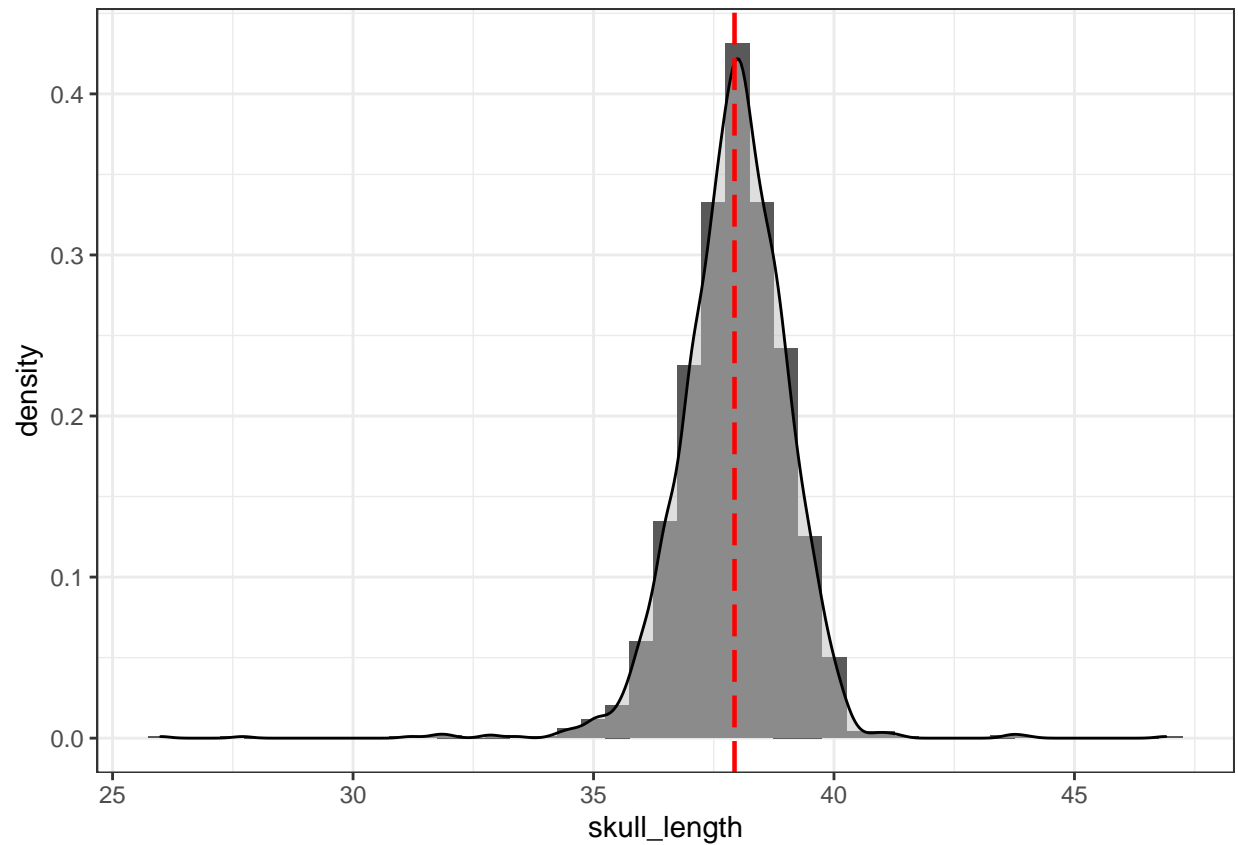
```
## Warning: Removed 1775 rows containing non-finite values (stat_boxplot).
```



```
ggplot(data = ASSP, aes(x = skull_length), na.rm = TRUE) +
  geom_histogram(aes(y = ..density..), binwidth = 0.5) +
  geom_density(alpha = .5, fill = "gray") +
  geom_vline(aes(xintercept = mean(skull_length, na.rm = TRUE)),
    colour = "red", linetype = "longdash", size = .8) +
  theme_bw()
```

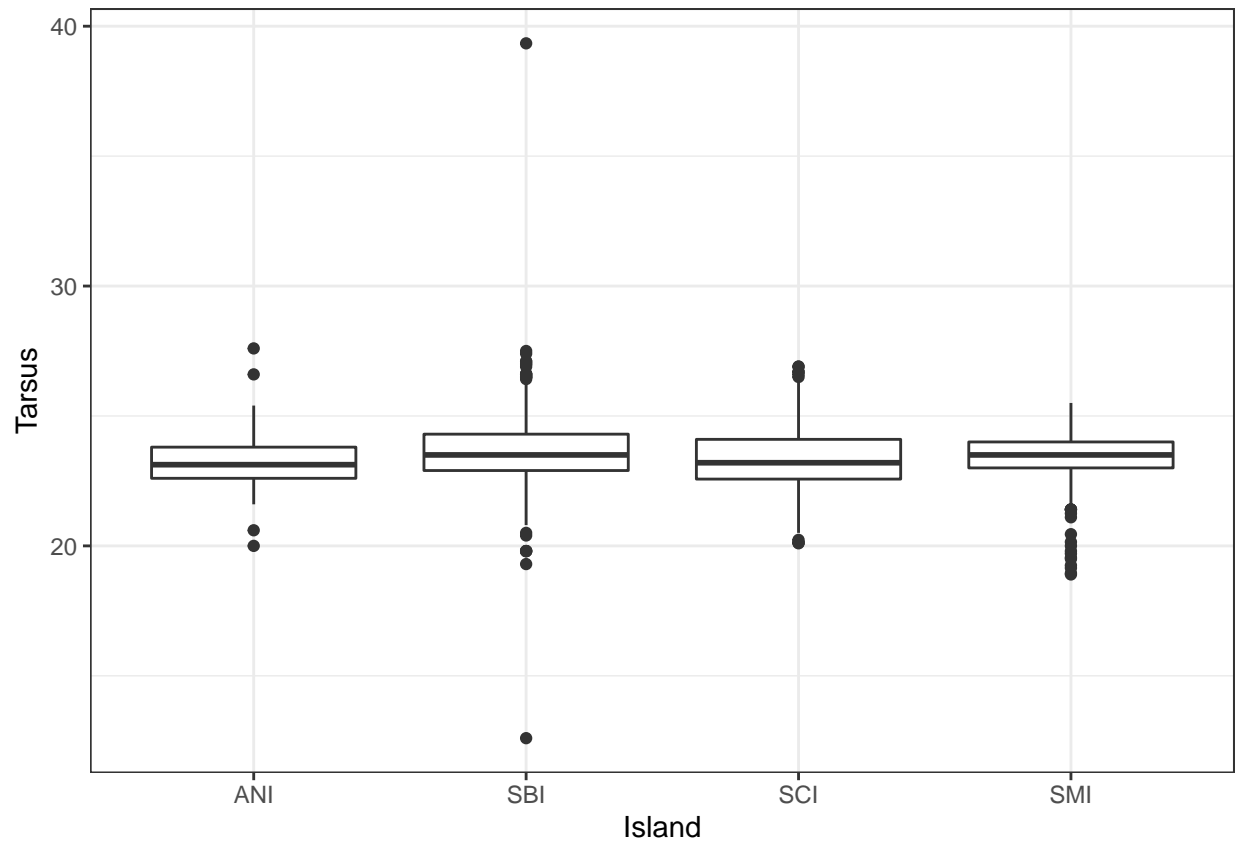
```
## Warning: Removed 1775 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 1775 rows containing non-finite values (stat_density).
```



```
tarsus <- ggplot(data = ASSP, aes(x = island_code, y = tarsus)) +  
  geom_boxplot() +  
  xlab("Island") + ylab("Tarsus") +  
  theme_bw()  
tarsus
```

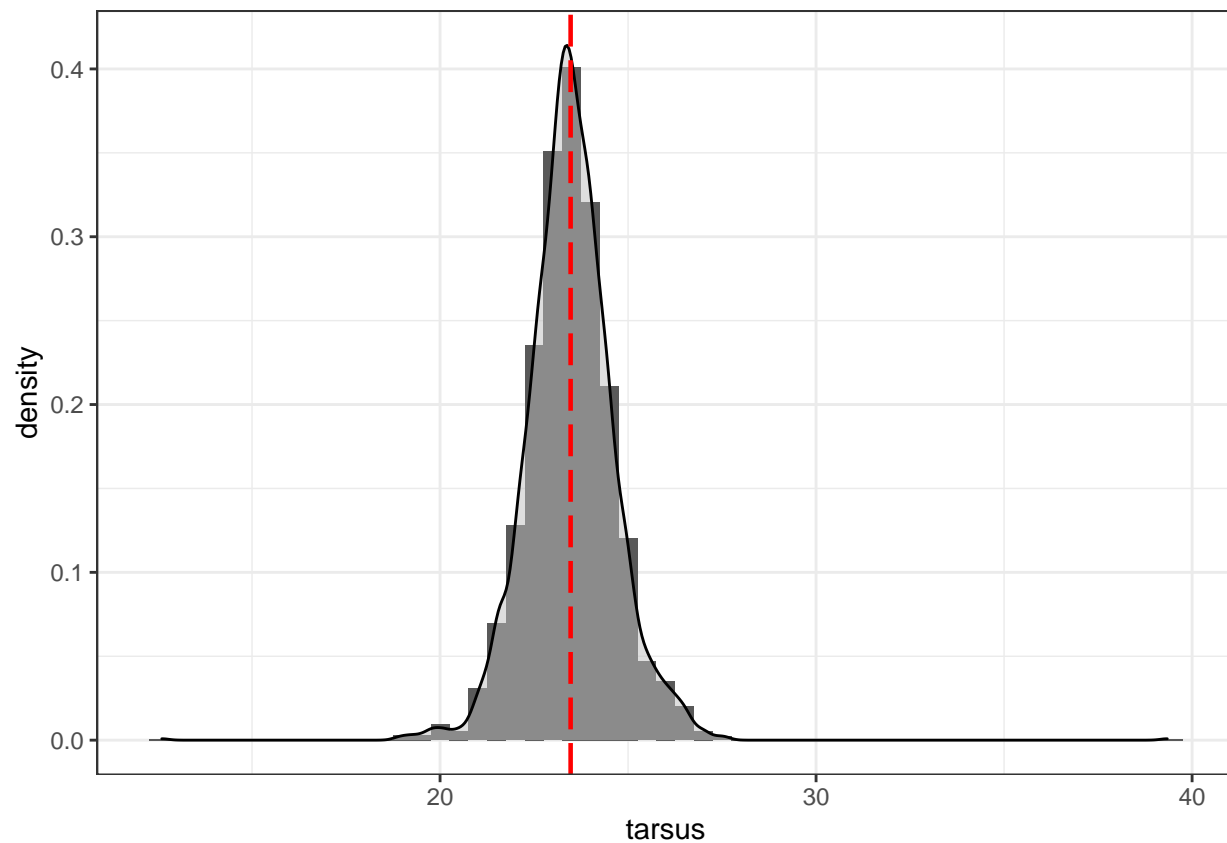
```
## Warning: Removed 1208 rows containing non-finite values (stat_boxplot).
```

```
ggplot(data = ASSP, aes(x = tarsus), na.rm = TRUE) +
  geom_histogram(aes(y = ..density..), binwidth = 0.5) +
  geom_density(alpha = .5, fill = "gray") +
  geom_vline(aes(xintercept = mean(tarsus, na.rm = TRUE)),
    colour = "red", linetype = "longdash", size = .8) +
  theme_bw()
```

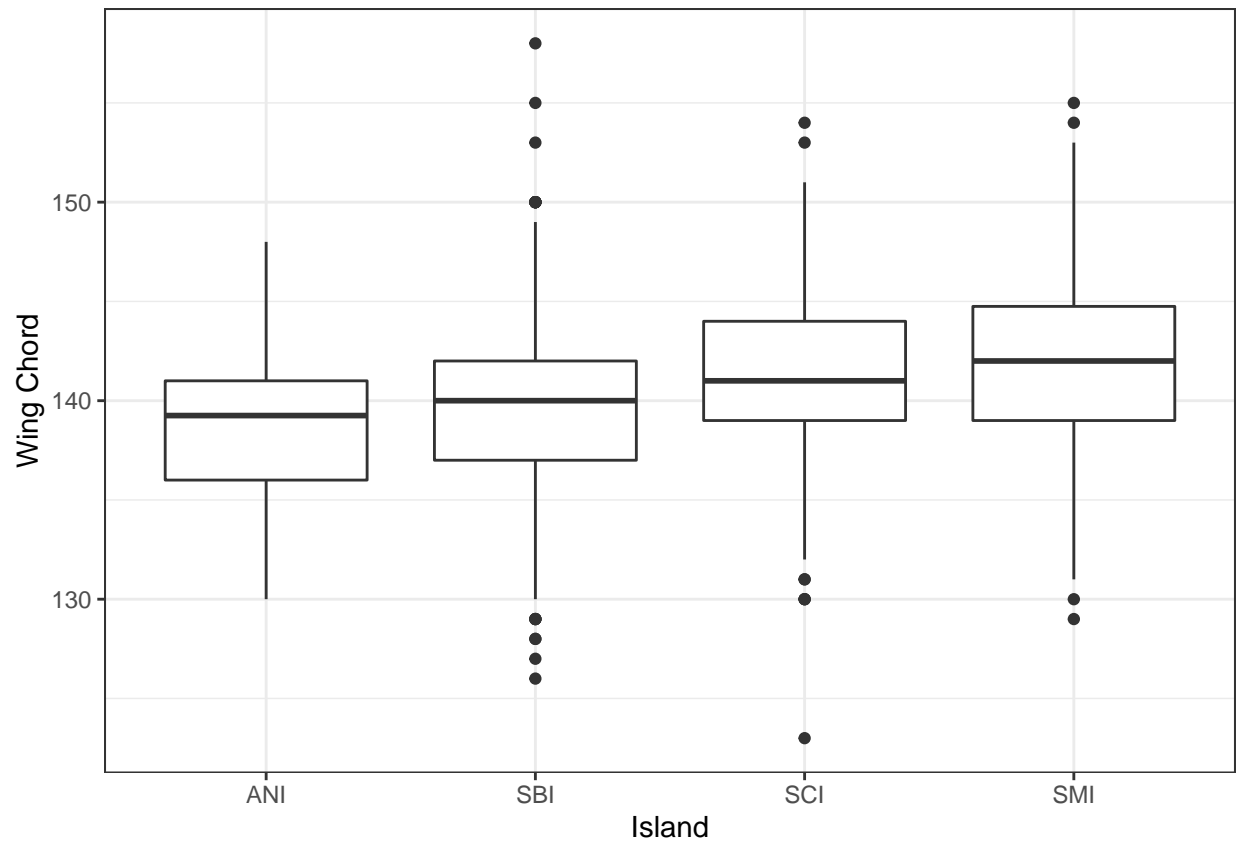
```
## Warning: Removed 1208 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 1208 rows containing non-finite values (stat_density).
```



```
wing <- ggplot(data = ASSP, aes(x = island_code, y = wing)) +  
  geom_boxplot() +  
  xlab("Island") + ylab("Wing Chord") +  
  theme_bw()  
wing
```

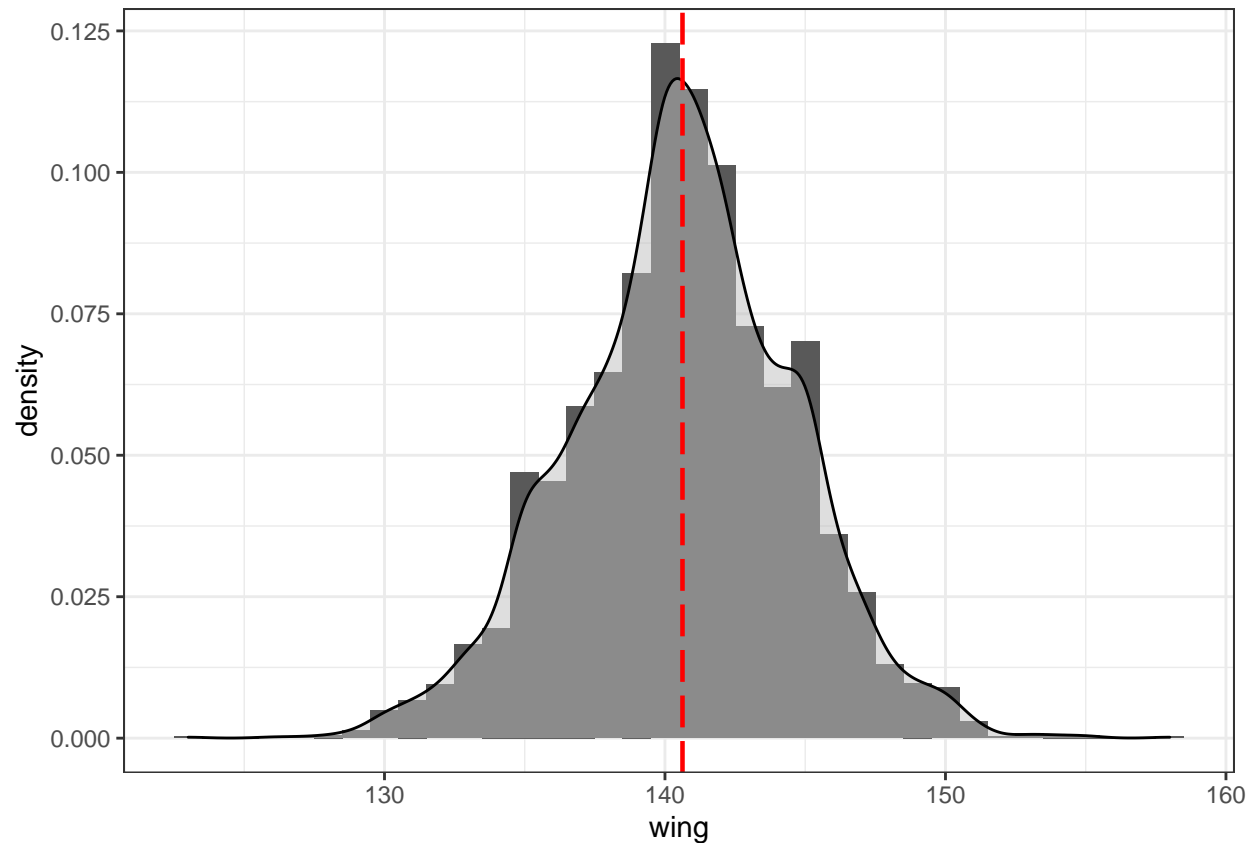
```
## Warning: Removed 118 rows containing non-finite values (stat_boxplot).
```



```
ggplot(data = ASSP, aes(x = wing), na.rm = TRUE) +
  geom_histogram(aes(y = ..density..), binwidth = 1) +
  geom_density(alpha = .5, fill = "gray") +
  geom_vline(aes(xintercept = mean(wing, na.rm = TRUE)),
    colour = "red", linetype = "longdash", size = .8) +
  theme_bw()
```

```
## Warning: Removed 118 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 118 rows containing non-finite values (stat_density).
```



Wing Chord

```
summary(ASSP$wing)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
##    123.0  138.0   141.0   140.6  143.0   158.0     118
```

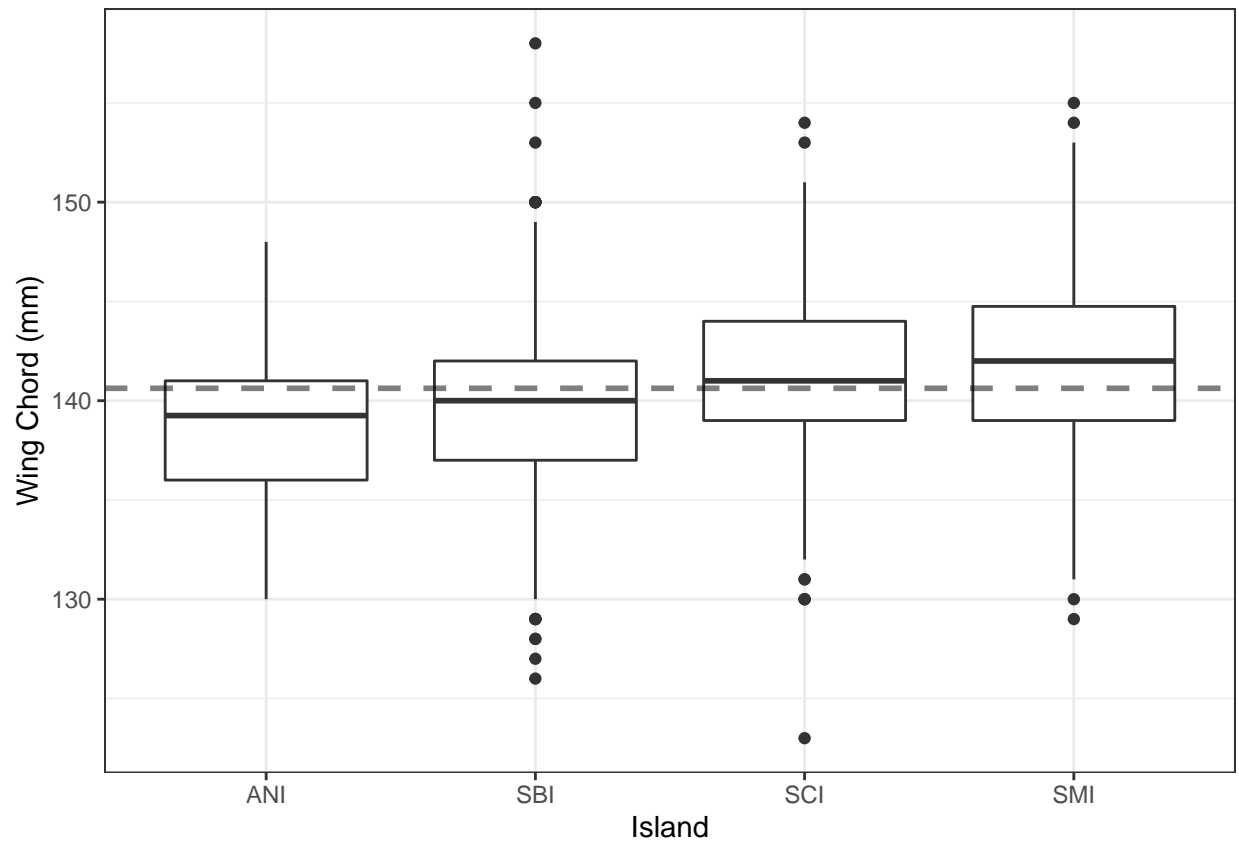
```
# Look for indication of different methods used to measure wing chord
# (e.g, flattened wing chord versus relaxed wing chord).
```

```
avwing <- mean(ASSP$wing, na.rm = TRUE)
```

```
# by island
```

```
wing.isl <- ggplot(ASSP, aes(x = island_code, y = wing)) +
  geom_boxplot() +
  xlab("Island") + ylab("Wing Chord (mm)") +
  geom_hline(yintercept = avwing, linetype = "dashed", color = "black", size = 1, alpha = 0.5) +
  theme_bw()
wing.isl
```

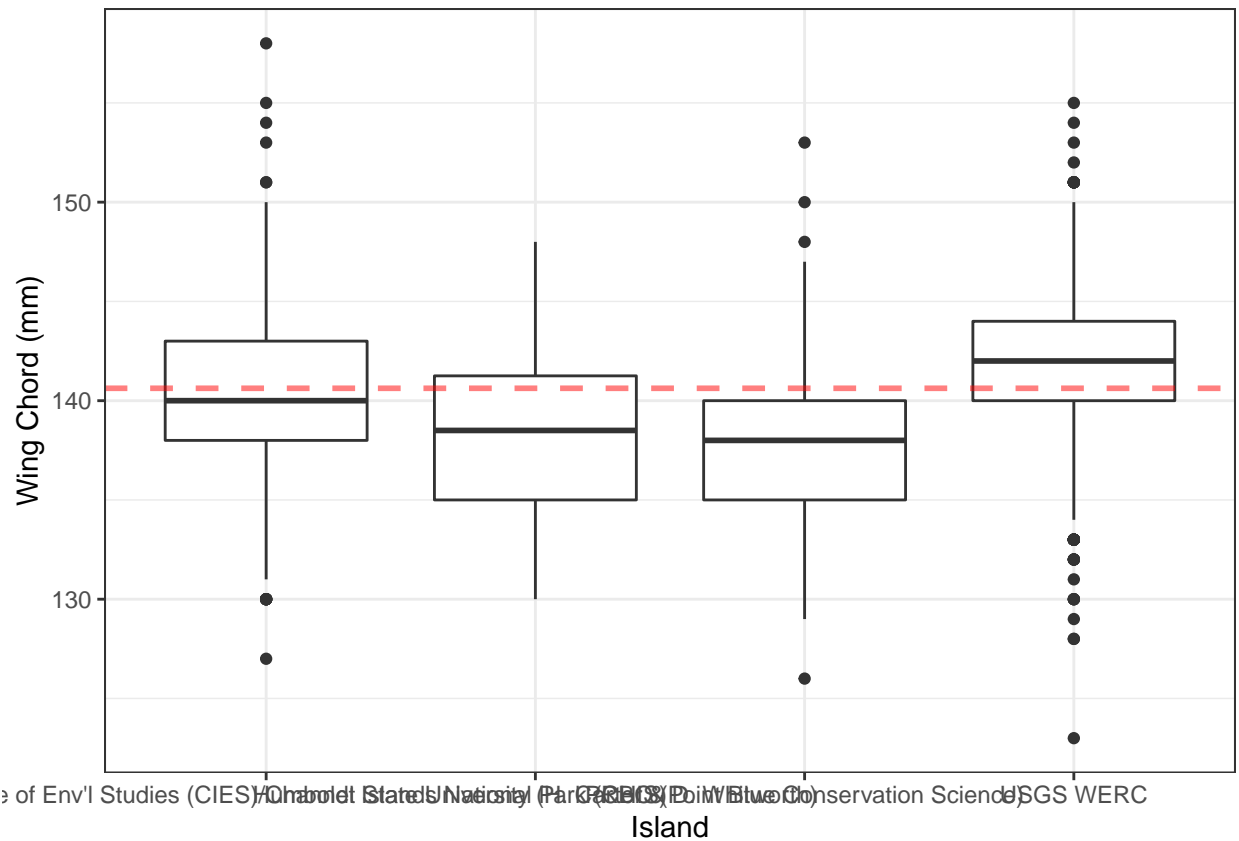
```
## Warning: Removed 118 rows containing non-finite values (stat_boxplot).
```



```
# by organization and island
ASSPorg <- inner_join(ASSP, metadata, by = "session_ID")

wing.org <- ggplot(ASSPorg, aes(x = org, y = wing)) +
  geom_boxplot() +
  xlab("Island") + ylab("Wing Chord (mm)") +
  geom_hline(yintercept = avwing, linetype = "dashed", color = "red", size = 1, alpha = 0.5) +
  theme_bw()
wing.org
```

```
## Warning: Removed 118 rows containing non-finite values (stat_boxplot).
```



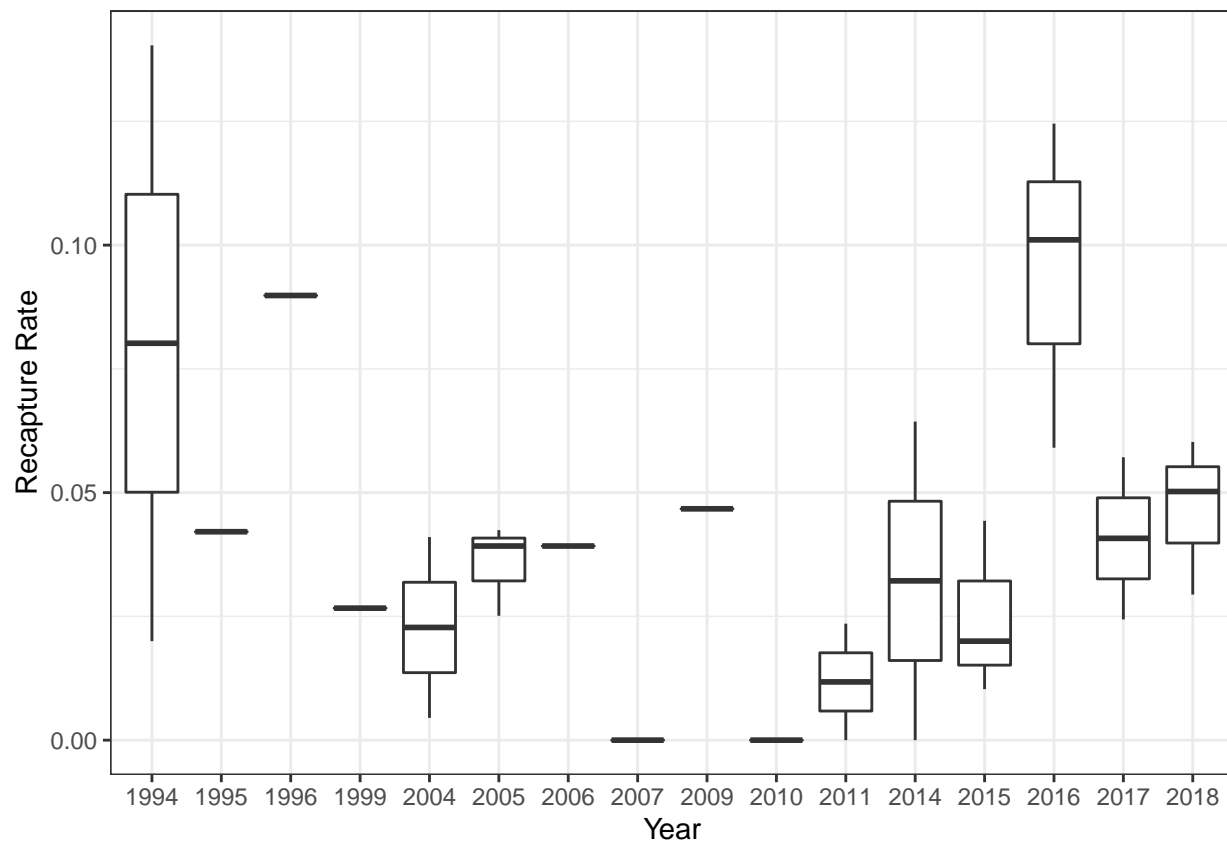
```
## how to fix org names?
```

RECAPTURES

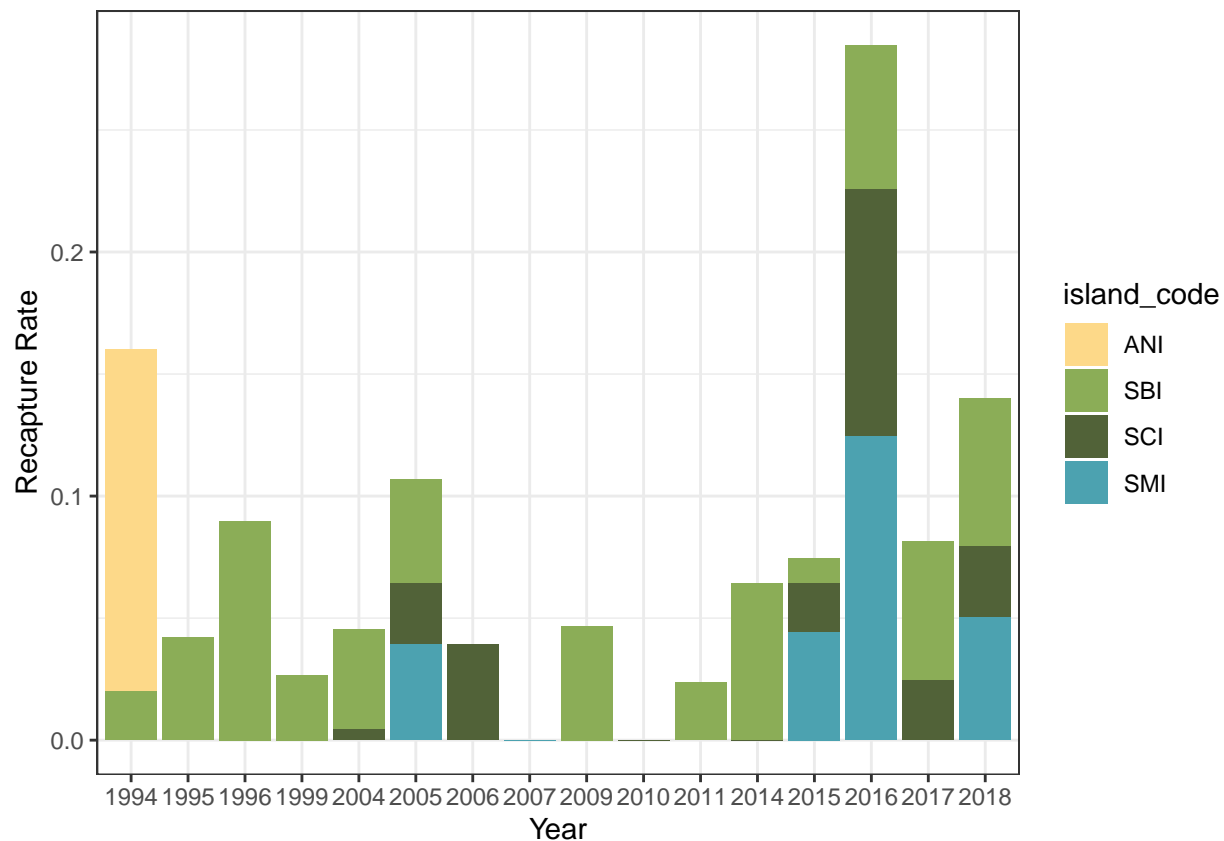
```
recap <- group_by(.data = ASSP, session_year, island_code) %>%
  summarise(no_captures = n(),
            recaps = sum(recapture == "Y"),
            ratio = recaps/no_captures)
```

```
## 'summarise()' regrouping output by 'session_year' (override with '.groups' argument)
```

```
# recapture rate over time
ggplot(recap, aes(x = as.character(session_year), y = ratio)) +
  geom_boxplot() +
  xlab("Year") + ylab("Recapture Rate") +
  theme_bw()
```



```
ggplot(recap, aes(x = as.character(session_year), y = ratio, fill = island_code)) +
  geom_col() +
  xlab("Year") + ylab("Recapture Rate") +
  scale_fill_manual(values = cal_palette("sierra2")) +
  theme_bw()
```



BROOD PATCH

```
summary(ASSP$BP)
```

```
##      Length      Class      Mode 
##      3815 character character
```

```
unique(ASSP$BP)
```

```
## [1] "4.5" "1" "1.5" "3" "5" "2" "4" NA "0" "B" "PD" "D"
```

```
## Brood Patch and Assumed Breeders
```

```
monthCatches <- metadata %>%
```

```
  group_by(session_month, island_code) %>%
  tally()
```

```
ggplot(metadata, aes(x = as.character(session_month), y = BPfreq_Y)) +
  geom_boxplot(na.rm = TRUE) +
  ylab("Frequency of Assumed Breeders") +
  facet_wrap(~island_code, ncol = 1) +
  xlab("Month") +
  theme_bw()
```