## Example R script for reproducible analysis

### Analysis of change in Pileated woodpecker abundance versus habitat

#### and variation over time (2000,2005, 2010)

### Script written 11/21/2017

### by Nathan Brouwer brouwern@gmail.com

# This is a bare-bones script to carry out the creation of the dataframe and begin the regression/ANOVA analysis.

# To be complete this script MUST have additional comment added. Please see the associated vignette

# in the widlifeR package for more infomation.

#load devtools

library(devtools)

#download wildlifeR from github

install\_github("brouwern/wildlifeR")

#load wildlifeR into your current R sessions

library(wildlifeR)

### The USGS BBS data

data(BBS\_PA)

### Selecting a focal bird species

#### Determining the numeric species code

data(AOU\_species\_codes)

library(dplyr)

AOU\_species\_codes %>% filter(alpha.code == "PIWO")

BBS\_PA\_PIWO <- BBS\_PA %>% filter(Aou == "4050")

### Selecting a focal analysis question

### Formatting data for habitat analyses

#### Subsetting bird data

BBS\_PA\_PIWO\_2 <- BBS\_PA\_PIWO %>% filter(Year == 2006)

#### Preparing habitat data

data(BBS\_PA\_landcover\_1km)

BBS\_PA\_landcover\_1km\_2 <- BBS\_PA\_landcover\_1km %>%

select(Route, NLCD.41, NLCD.42, NLCD.43, SUM)

#### Calculating percentages of habitat

BBS\_PA\_landcover\_1km\_2$decid.percent <- BBS\_PA\_landcover\_1km\_2$NLCD.41/BBS\_PA\_landcover\_1km\_2$SUM

BBS\_PA\_landcover\_1km\_2$conifer.percent <- BBS\_PA\_landcover\_1km\_2$NLCD.42/BBS\_PA\_landcover\_1km\_2$SUM

BBS\_PA\_landcover\_1km\_2$mixed.forest.percent <- BBS\_PA\_landcover\_1km\_2$NLCD.42/BBS\_PA\_landcover\_1km\_2$SUM

#### Merging bird counts and habitat cover

BBS\_PA\_PIWO\_3 <- full\_join(BBS\_PA\_PIWO\_2 ,

BBS\_PA\_landcover\_1km\_2, by = "Route")

#### Filling in columns

BBS\_PA\_PIWO\_3$Year <- 2006

BBS\_PA\_PIWO\_3$Aou <- 6080

BBS\_PA\_PIWO\_3$name <- "PIWO"

#### Filling in NAs

BBS\_PA\_PIWO\_4 <- NA\_to\_zero(dat = BBS\_PA\_PIWO\_3,

column = "SpeciesTotal")

#compare output

#with NAs

summary(BBS\_PA\_PIWO\_3$SpeciesTotal)

#with NAs removed by set\_NA\_SpeciesTotal\_to\_zero()

summary(BBS\_PA\_PIWO\_4$SpeciesTotal)

## Save data

write.csv(BBS\_PA\_PIWO\_4, file = "PIWO\_vs\_forest\_cover.csv")

### Formatting data for analysis of change over time

library(Hmisc)

PIWO\_00\_05\_10 <- sample\_BBS\_routes(dat = BBS\_PA\_PIWO,

years = c(2000,2005,2010), aou.code = 4050)

library(ggplot2)

library(ggpubr)

#### Plot

ggerrorplot(data = PIWO\_00\_05\_10,

y = "SpeciesTotal",

x = "Year",

desc\_stat = "mean\_ci")

### save

write.csv(PIWO\_00\_05\_10, file = "PIWO\_00\_05\_10.csv")

## Data Exploration

### Exploring abundance-habitat data

#### Boxplots

## number of birds seen

plot1.spp.total <- ggboxplot(data = BBS\_PA\_PIWO\_4,

y = "SpeciesTotal", xlab = "")

## Deciduous landcover

plot2.decid.percent <- ggboxplot(data = BBS\_PA\_PIWO\_4,

y = "decid.percent", xlab = "")

## Coniferous landcover

plot3.confir.percent <- ggboxplot(data = BBS\_PA\_PIWO\_4,

y = "conifer.percent", xlab = "")

## Mixed forest landcover

plot4.mixed.percent <- ggboxplot(data = BBS\_PA\_PIWO\_4,

y = "mixed.forest.percent", xlab = "")

###

library(cowplot)

plot\_grid(plot1.spp.total, plot2.decid.percent,

plot3.confir.percent, plot4.mixed.percent,

labels = c("A", "B","C","D"))

#### Histograms

## number of birds seen

plot1.spp.total <- gghistogram(data = BBS\_PA\_PIWO\_4,

x = "SpeciesTotal", xlab = "")

## Deciduous landcover

plot2.decid.percent <- gghistogram(data = BBS\_PA\_PIWO\_4,

x = "decid.percent", xlab = "")

## Coniferous landcover

plot3.confir.percent <- gghistogram(data = BBS\_PA\_PIWO\_4,

x = "conifer.percent", xlab = "")

## Mixed forest landcover

plot4.mixed.percent <- gghistogram(data = BBS\_PA\_PIWO\_4,

x = "mixed.forest.percent", xlab = "")

## Layout the four plots in a grid.

library(cowplot)

plot\_grid(plot1.spp.total, plot2.decid.percent,

plot3.confir.percent, plot4.mixed.percent,

labels = c("A", "B", "C","D"))

#### Scatter plot

library(ggplot2)

library(ggpubr)

ggscatter(data = BBS\_PA\_PIWO\_4,

y = "SpeciesTotal", x = "conifer.percent") +

geom\_smooth(se = FALSE)

### Exploring abudance-time data

#### Boxplot

ggboxplot(data = PIWO\_00\_05\_10,

y = "SpeciesTotal", x = "Year")

#### Histograms

gghistogram(data = PIWO\_00\_05\_10,

x = "SpeciesTotal", facet.by = "Year")

#### Plotting means and error bars

ggerrorplot(data = PIWO\_00\_05\_10,

y = "SpeciesTotal", x = "Year",

desc\_stat = "mean\_ci")

## Data modeling

### Modeling habitat

m.null <- lm(SpeciesTotal ~ 1, BBS\_PA\_PIWO\_4)

m.decid <- lm(SpeciesTotal ~ decid.percent, BBS\_PA\_PIWO\_4)

m.mixed <- lm(SpeciesTotal ~ mixed.forest.percent, BBS\_PA\_PIWO\_4)

### Present results

#### Regression summary

library(broom)

tidy(m.null)

tidy(m.decid)

tidy(m.mixed)

#### Significance test

anova(m.null, m.decid)

anova(m.null, m.mixed)

#### Model comparison

library(bbmle)

AICtab(m.null, m.decid, m.mixed)

### Modeling variation over time

PIWO\_00\_05\_10$Year <- factor(PIWO\_00\_05\_10$Year)

m.time.null <- lm(SpeciesTotal ~ 1 , data = PIWO\_00\_05\_10)

m.time.year <- lm(SpeciesTotal ~ Year , data = PIWO\_00\_05\_10)

### Regression summary

library(broom)

tidy(m.time.null)

tidy(m.time.year)

### Model comparison

library(bbmle)

AICtab(m.time.null, m.time.year)

### Multiple comparisons

#### Refit model w/aov()

m.time.year.aov <- aov(SpeciesTotal ~ Year , data = PIWO\_00\_05\_10)

TukeyHSD(m.time.year.aov)

tukey.out <- TukeyHSD(m.time.year.aov)

plotTukeysHSD(tukey.out)