Lab 03: Data Exploration and Deterministic Functions

David Murillo

September 27, 2022

Contents

Load packages

```
library(psych)
library(here)
```

```
## Warning: package 'here' was built under R version 4.1.3
```

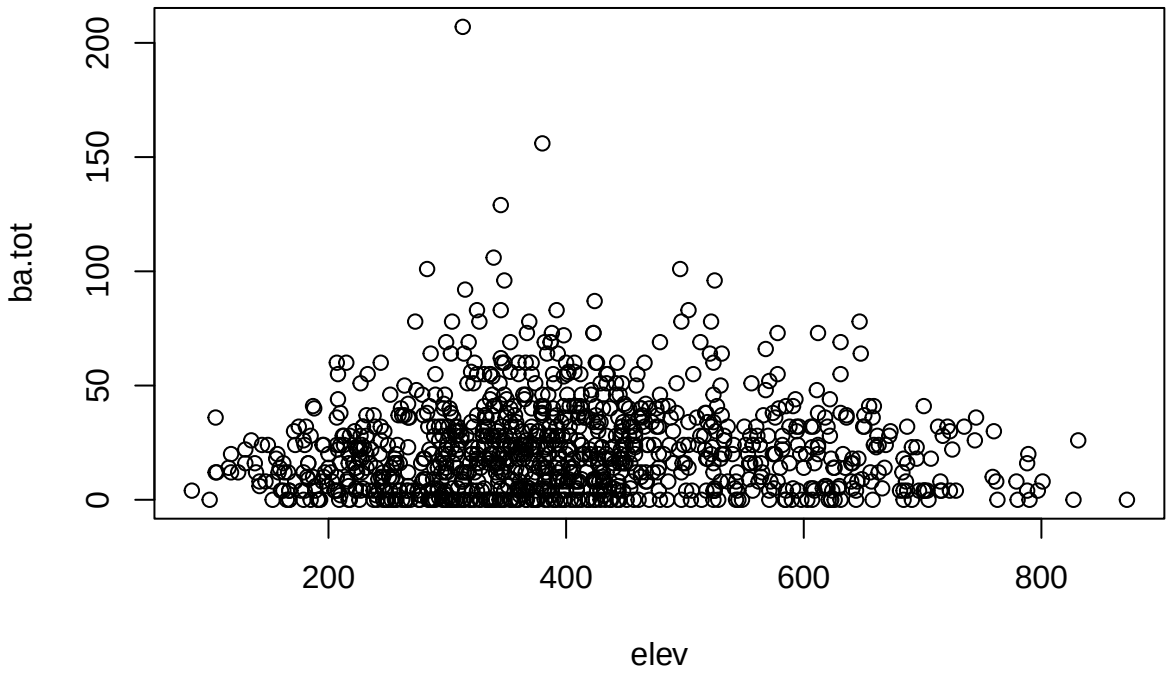
```
## here() starts at C:/Users/Pc/OneDrive - University of Massachusetts/Umass/Classes/Fall2022/DataAna/e
```

Load datasets

```
dat_bird <- read.csv(here("data", "bird.sta.csv"))
dat_habitat <- read.csv(here("data", "hab.sta.csv"))
```

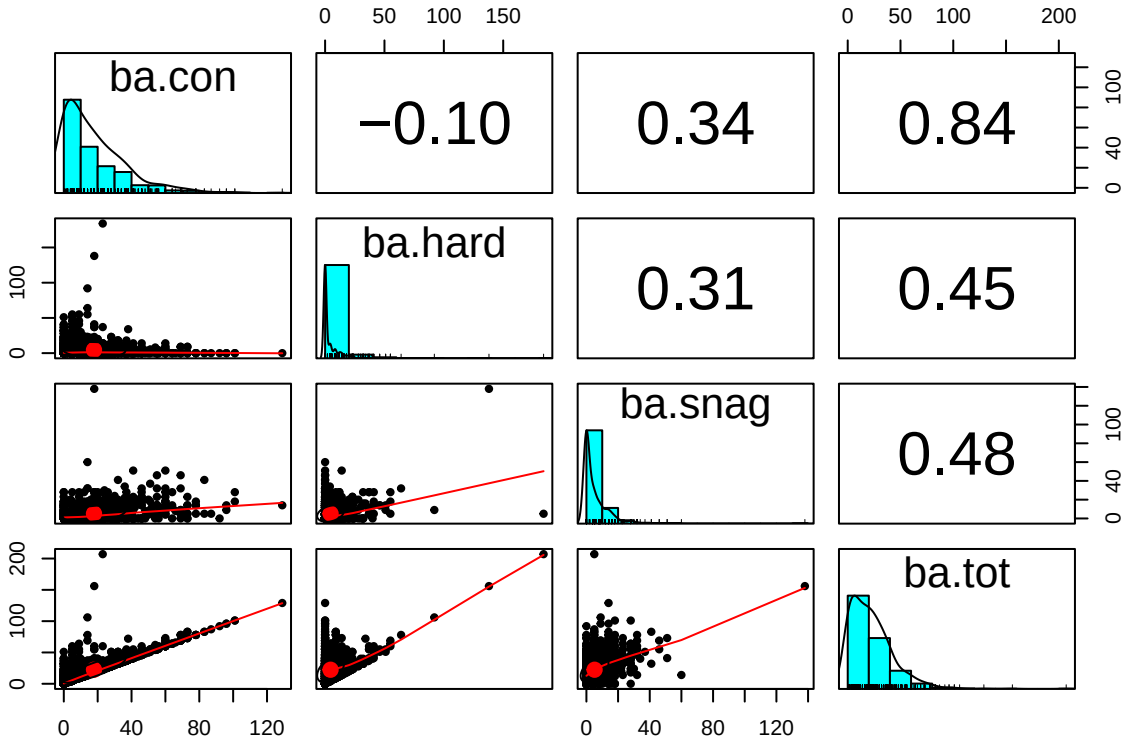
Merge data

```
dat_all <- merge(dat_habitat, dat_bird)

plot(ba.tot ~ elev, data = dat_all)
```



1. Use the pair plot function from psych to create a pair plot of the three terrain variables and basal area from the lecture questions.

```
pairs.panels(dat_all[,c("ba.con", "ba.hard","ba.snag", "ba.tot")])
```



2. Choose two bird species and create plots of presence/absence (on the y-axis) and basal area (on the x axes).

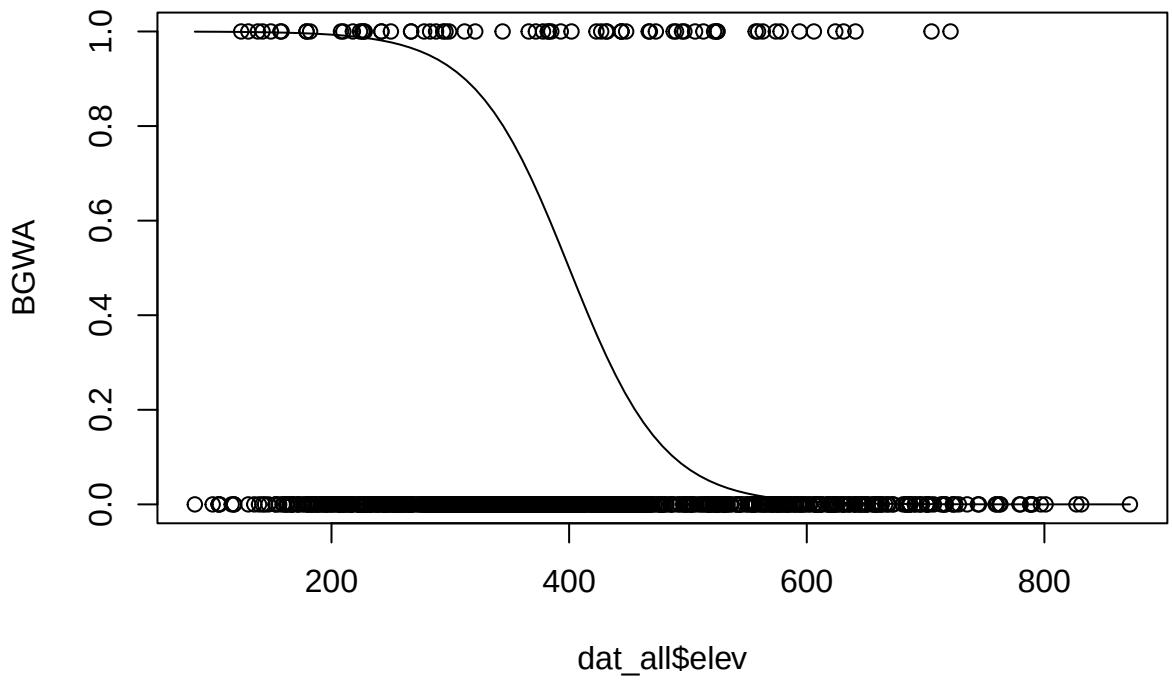
```
BGWA <- as.numeric(dat_all$BGWA) > 0

SWTH <- as.numeric(dat_all$SWTH) > 0
```

Visually inspect the plots and fit logistic curves using the parameterization functions provided above.

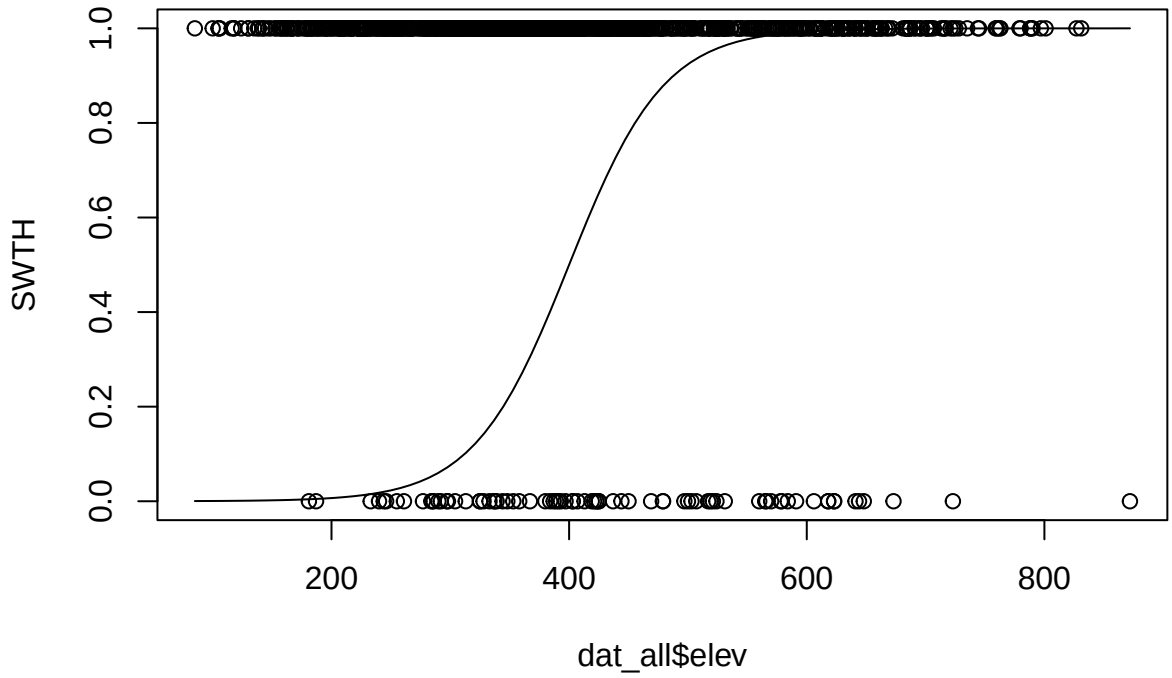
```
# Function to calculate the logistic parameter a given the slope and midpoint
get_logistic_param_a = function(slope, midpoint)
{
  b = slope / 4
  return (-midpoint * (slope / 4))
}

# Function to calculate the logistic parameter b given the slope
get_logistic_param_b = function(slope)
{
  return (slope / 4)
}


# Calculate the value of the logistic function at x, given the parameters a and b.
logistic = function(x, a, b)
```

```
{
  val = exp(a + b * x)
  return(val / (1 + val))
}

# Calculate the value of the logistic function at x, given a slope and midpoint.
logistic_midpoint_slope = function(x, midpoint, slope)
{
  b = get_logistic_param_b(slope)
  a = get_logistic_param_a(slope, midpoint)
  return(logistic(x, a, b))
}
```

```
plot(dat_all$elev, BGWA)
curve(logistic_midpoint_slope(x, midpoint = 400, slope = -0.1), add = TRUE)
```




```
plot(dat_all$elev, SWTH)
curve(logistic_midpoint_slope(x, midpoint = 400, slope = 0.1), add = TRUE)
```



How many total number of Gray Jays were observed in all of the sampling sites.

```
sum(dat_all$GRJA)
```

```
## [1] 181
```

Show the R code you used to perform the calculation.

Calculate the total number of sampling sites in which Gray Jays were observed.

```
GRJA_sta <- sum(as.numeric(dat_all$GRJA > 0))
GRJA_sta
```

```
## [1] 110
```

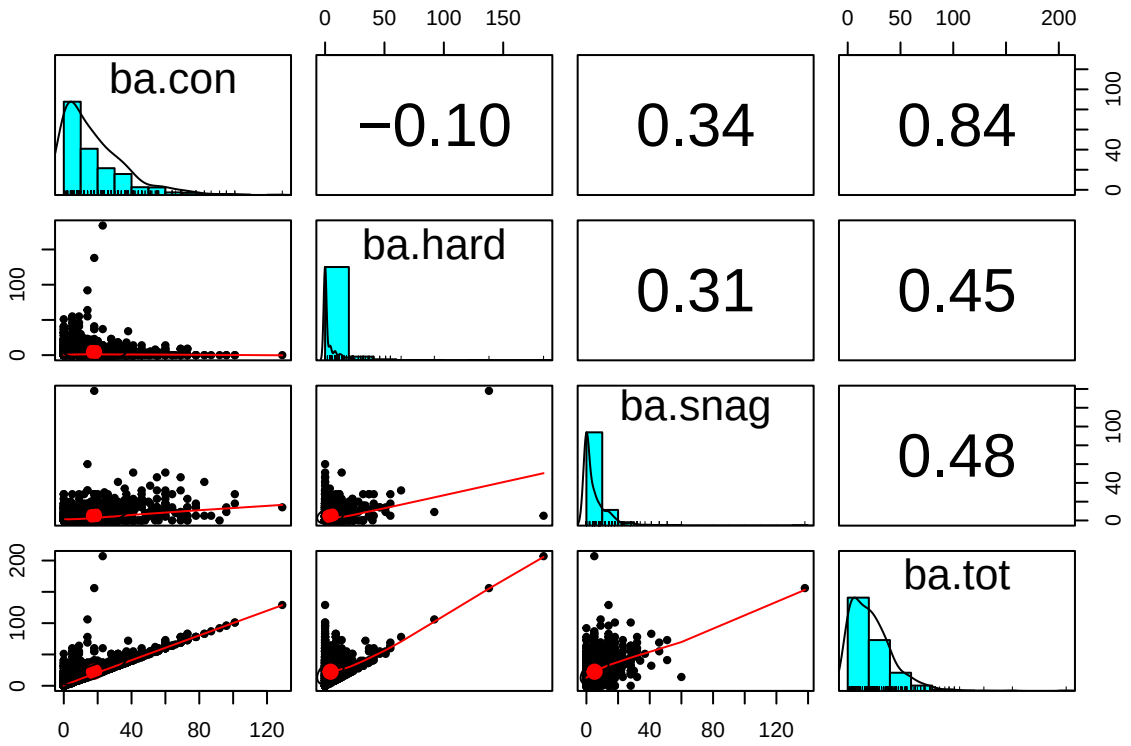
Questions

Q1 (1 pt.): What is basal area, and how is it measured?

Answer:

Q2 (2 pts.): Include a figure of your terrain/basal area pairplot.

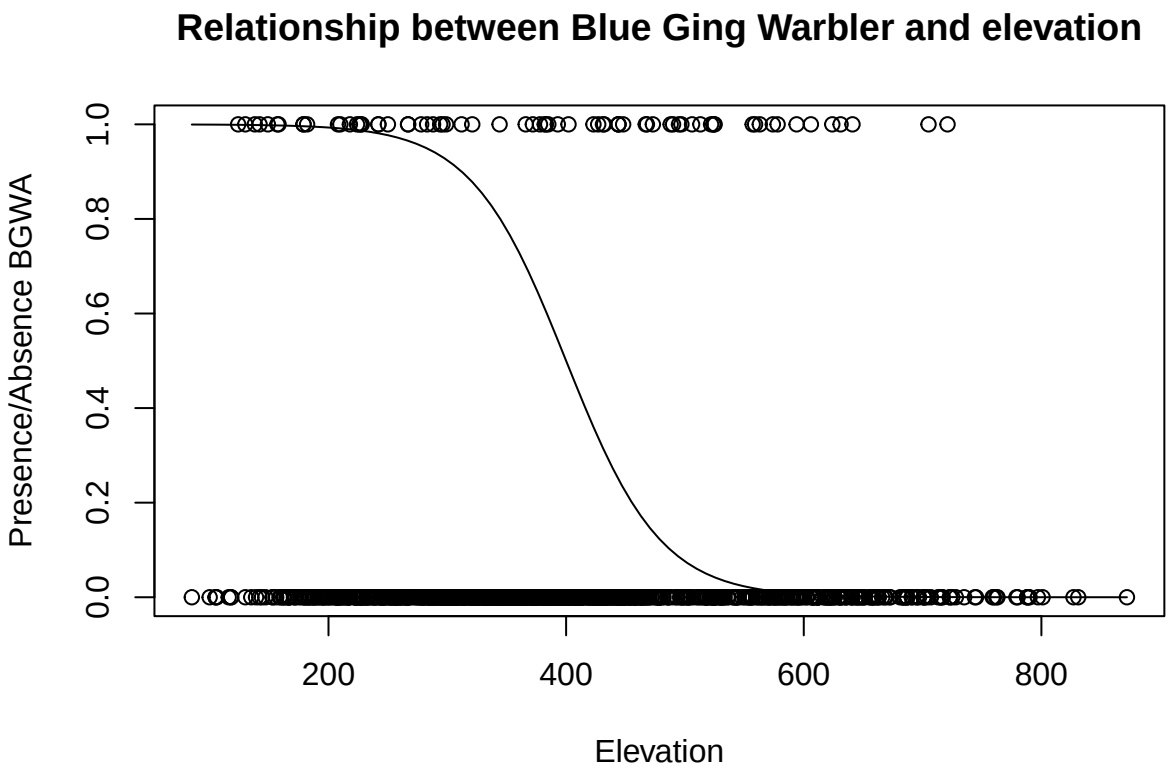
```
pairs.panels(dat_all[,c("ba.con", "ba.hard","ba.snag", "ba.tot")])
```



Consider the first bird species you chose to examine in the walkthrough.

Q3 (1 pt.): Include a figure of your logistic function plot. Your figure must include the name of the bird species, appropriate title, axes, etc.

```
plot(dat_all$elev, BGWA, main = "Relationship between Blue Ging Warbler and elevation", ylab = "Presenc

curve(logistic_midpoint_slope(x, midpoint = 400, slope = -0.1), add = TRUE)
```



Q4 (3 pts.): Qualitatively describe the bird's presence/absence patterns in terms of basal area (or your other chosen predictor). Your answer should make reference to your fitted logistic model plot. Some questions you might consider are:

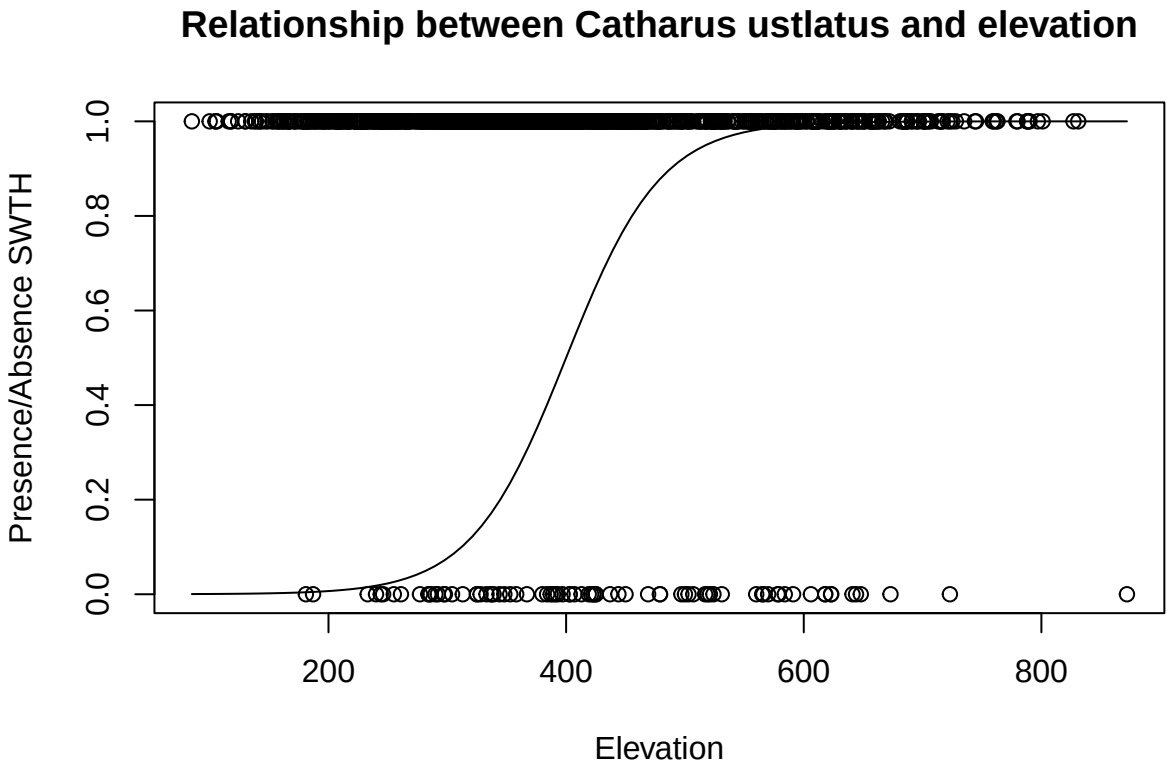
Answer:

Does the bird species seem to prefer areas with high or low tree cover? Does the bird species prefer low or high elevations? (for example, if you used elevation instead of basal area) Does a logistic model seem like a good fit?

Consider the second bird species you chose to examine in the walkthrough.

Q5 (1 pt.): Include a figure of your logistic function plot. Your figure must include the name of the bird species, appropriate title, axes, etc.

```
plot(dat_all$elev, SWTH, main = "Relationship between Catharus ustlatus and elevation", ylab = "Presenc
curve(logistic_midpoint_slope(x, midpoint = 400, slope = 0.1), add = TRUE)
```



Q6 (3 pts.): Qualitatively describe the bird's presence/absence patterns in terms of basal area (or your other chosen predictor). Your answer should make reference to your fitted logistic model plot. Some questions you might consider are:

Answer:

Does the bird species seem to prefer areas with high or low tree cover? Does the bird species prefer low or high elevations? (for example, if you used elevation instead of basal area) Does a logistic model seem like a good fit?

Q7 (1 pt.): How many total number of Gray Jays were observed in all of the sampling sites.

Answer: 181

Q8 (2 pts.): Show the R code you used to perform the calculation.

```
sum(dat_all$GRJA)
```

```
## [1] 181
```

Q9 (1 pt.): Calculate the total number of sampling sites in which Gray Jays were observed.

Answer: 110

Q10 (2 pts.): Include the R code you used to perform the presence/absence calculation.

```
GRJA_sta <- sum(as.numeric(dat_all$GRJA > 0))
GRJA_sta
```

```
## [1] 110
```