Homework 5

Luc-Tanton Tran

2025-01-25

Instructions

Answer the following questions and/or complete the exercises in RMarkdown. Please embed all of your code and push the final work to your repository. Your report should be organized, clean, and run free from errors. Remember, you must remove the `#` for any included code chunks to run.

Load the tidyverse

```
library("tidyverse")
library("janitor")
```

Data

For this assignment, we will use data from a study on vertebrate community composition and impacts from defaunation in Gabon, Africa. One thing to notice is that the data include 24 separate transects. Each transect represents a path through different forest management areas.

Reference: Koerner SE, Poulsen JR, Blanchard EJ, Okouyi J, Clark CJ. Vertebrate community composition and diversity declines along a defaunation gradient radiating from rural villages in Gabon. *Journal of Applied Ecology*. 2016. This paper, along with a description of the variables is included inside the data folder.

1. Load `IvindoData_DryadVersion.csv` and store it as a new object called `gabon`.

```
gabon <- read.csv("data/IvindoData_DryadVersion.csv")
```

2. Use one or more of the summary functions you have learned to get an idea of the structure of the data.

```
glimpse(gabon)
```

```
## Rows: 24
## Columns: 26
## $ TransectID             <int> 1, 2, 2, 3, 4, 5, 6, 7, 8, 9, 13, 14, 15, 16, ~
## $ Distance               <dbl> 7.14, 17.31, 18.32, 20.85, 15.95, 17.47, 24.06~
## $ HuntCat                <chr> "Moderate", "None", "None", "None", "None", "N~
## $ NumHouseholds          <int> 54, 54, 29, 29, 29, 29, 29, 54, 25, 73, 46, 56~
## $ LandUse                <chr> "Park", "Park", "Park", "Logging", "Park", "Pa~
## $ Veg_Rich               <dbl> 16.67, 15.75, 16.88, 12.44, 17.13, 16.50, 14.7~
## $ Veg_Stems              <dbl> 31.20, 37.44, 32.33, 29.39, 36.00, 29.22, 31.2~
## $ Veg_liana              <dbl> 5.78, 13.25, 4.75, 9.78, 13.25, 12.88, 8.38, 8~
## $ Veg_DBH                <dbl> 49.57, 34.59, 42.82, 36.62, 41.52, 44.07, 51.2~
## $ Veg_Canopy             <dbl> 3.78, 3.75, 3.43, 3.75, 3.88, 2.50, 4.00, 4.00~
## $ Veg_Understory         <dbl> 2.89, 3.88, 3.00, 2.75, 3.25, 3.00, 2.38, 2.71~
## $ RA_Apes                <dbl> 1.87, 0.00, 4.49, 12.93, 0.00, 2.48, 3.78, 6.1~
## $ RA_Birds               <dbl> 52.66, 52.17, 37.44, 59.29, 52.62, 38.64, 42.6~
```

```
## $ RA_Elephant             <dbl> 0.00, 0.86, 1.33, 0.56, 1.00, 0.00, 1.11, 0.43~
## $ RA_Monkeys              <dbl> 38.59, 28.53, 41.82, 19.85, 41.34, 43.29, 46.2~
## $ RA_Rodent               <dbl> 4.22, 6.04, 1.06, 3.66, 2.52, 1.83, 3.10, 1.26~
## $ RA_Ungulate             <dbl> 2.66, 12.41, 13.86, 3.71, 2.53, 13.75, 3.10, 8~
## $ Rich_AllSpecies         <int> 22, 20, 22, 19, 20, 22, 23, 19, 19, 19, 21, 22~
## $ Evenness_AllSpecies     <dbl> 0.793, 0.773, 0.740, 0.681, 0.811, 0.786, 0.81~
## $ Diversity_AllSpecies    <dbl> 2.452, 2.314, 2.288, 2.006, 2.431, 2.429, 2.56~
## $ Rich_BirdSpecies        <int> 11, 10, 11, 8, 8, 10, 11, 11, 11, 9, 11, 11, 1~
## $ Evenness_BirdSpecies    <dbl> 0.732, 0.704, 0.688, 0.559, 0.799, 0.771, 0.80~
## $ Diversity_BirdSpecies   <dbl> 1.756, 1.620, 1.649, 1.162, 1.660, 1.775, 1.92~
## $ Rich_MammalSpecies      <int> 11, 10, 11, 11, 12, 12, 12, 8, 8, 10, 10, 11, ~
## $ Evenness_MammalSpecies  <dbl> 0.736, 0.705, 0.650, 0.619, 0.736, 0.694, 0.77~
## $ Diversity_MammalSpecies <dbl> 1.764, 1.624, 1.558, 1.484, 1.829, 1.725, 1.92~
```

3. Use `mutate()` Change the variables `HuntCat`, `LandUse`, and `TransectID` to factors.

```
gabon1 <- gabon |>
  #As a note, I was reading the textbook and it advises the use of "|>" as a pipe instead of "%>%".
  #I changed my default pipe (Ctrl+Shift+M) to "|>" rather than "%>%" for this reason,
  #and it will show up as such throughout the remainder of this assignment.
  mutate(HuntCat = as.factor(HuntCat),
         LandUse = as.factor(LandUse),
         TransectID = as.factor(TransectID))
```

4. Use `filter` to make three new dataframes focused only on 1. national parks, 2. logging concessions, and 3. neither. Have a look at the README in the data folder so you understand the variables.

```
gabon_park <- gabon1 |>
  filter(LandUse == "Park")
gabon_logging <- gabon |>
  filter(LandUse == "Logging")
gabon_neither <- gabon1 |>
  filter(LandUse == "Neither")
```

5. How many transects are recorded for each land use type?

```
table(gabon1$LandUse)
```

```
##
## Logging Neither    Park
##      13       4       7
```

6. For which land use type (national parks, logging, or neither) is average all species diversity the greatest?

```
mean(gabon_park$Diversity_AllSpecies)
```

```
## [1] 2.425143
```

```
mean(gabon_logging$Diversity_AllSpecies)
```

```
## [1] 2.232538
```

```
mean(gabon_neither$Diversity_AllSpecies)
```

```
## [1] 2.3575
```

National Parks have the highest diversity across all species.

7. Use `filter` to find the transect that has the highest relative abundance of elephants. What land use type is this? Use `arrange()` to sort your results.

```
gabon1 |>
  filter(RA_Elephant == max(RA_Elephant)) |>
  arrange(RA_Elephant)
```

```
##   TransectID Distance HuntCat NumHouseholds LandUse Veg_Rich Veg_Stems
## 1         18    18.85    None            17 Logging    11.75        37
##   Veg_liana Veg_DBH Veg_Canopy Veg_Understory RA_Apes RA_Birds RA_Elephant
## 1     12.25   42.92       3.29           3.14    0.51    57.41         2.3
##   RA_Monkeys RA_Rodent RA_Ungulate Rich_AllSpecies Evenness_AllSpecies
## 1      35.13      2.09        2.56              20               0.789
##   Diversity_AllSpecies Rich_BirdSpecies Evenness_BirdSpecies
## 1                2.363               11                0.765
##   Diversity_BirdSpecies Rich_MammalSpecies Evenness_MammalSpecies
## 1                 1.833                  9                  0.692
##   Diversity_MammalSpecies
## 1                    1.52
```

The transect with the highest relative abundance is TransectID 18, with Logging land use type.

8. Use `filter` to find all transects that have greater than 15 tree species or a breast height diameter between 50 and 60cm.

```
gabon1 |>
  select(TransectID, LandUse, Veg_Rich, Veg_DBH) |>
  filter(Veg_Rich > 15 | between(Veg_DBH, 50, 60))
```

```
##    TransectID LandUse Veg_Rich Veg_DBH
## 1           1    Park    16.67   49.57
## 2           2    Park    15.75   34.59
## 3           2    Park    16.88   42.82
## 4           4    Park    17.13   41.52
## 5           5    Park    16.50   44.07
## 6           6    Park    14.75   51.22
## 7           9 Logging    16.00   69.30
## 8          14 Logging    15.75   52.12
## 9          15 Neither    10.88   54.77
## 10         17 Neither    14.25   57.71
## 11         21 Neither    16.25   50.36
## 12         22 Logging    17.13   39.31
## 13         24    Park    16.75   44.21
## 14         26 Logging    18.75   35.58
```

9.Which transects and land use types have more than 10 tree species and 10 mammal species? Use `arrange()` to sort by the number of tree species.

```
gabon1 |>
  select(TransectID, LandUse, Veg_Rich, Rich_MammalSpecies) |>
  filter(Veg_Rich > 10 & Rich_MammalSpecies > 10) |>
  arrange(Veg_Rich)
```

```
##   TransectID LandUse Veg_Rich Rich_MammalSpecies
## 1          3 Logging    12.44                11
## 2          6    Park    14.75                12
## 3         14 Logging    15.75                11
## 4          5    Park    16.50                12
```

```
## 5          1    Park    16.67                11
## 6         24    Park    16.75                12
## 7          2    Park    16.88                11
## 8          4    Park    17.13                12
## 9         22 Logging    17.13                12
```

10. Explore the data! Develop one question on your own that includes at least two lines of code.

Proposed question: Write code that can help you decide which transect (with RA_Birds >= 50%) has the most open (least covered) canopy. Assume that the only variables we care about are Transect_ID, LandUse, RA_Birds, and Veg_Canopy.

```
gabon1 |>
 select(TransectID, LandUse, RA_Birds, Veg_Canopy) |>
 filter(RA_Birds >= 50) |>
 arrange(Veg_Canopy)
```

```
##    TransectID LandUse RA_Birds Veg_Canopy
## 1           8 Neither    73.06       3.00
## 2          15 Neither    85.03       3.13
## 3          21 Neither    68.15       3.13
## 4           9 Logging    85.01       3.25
## 5          17 Neither    57.82       3.25
## 6          22 Logging    55.55       3.25
## 7          18 Logging    57.41       3.29
## 8          27 Logging    68.25       3.33
## 9          26 Logging    55.03       3.38
## 10         25 Logging    62.04       3.40
## 11         19 Logging    67.25       3.43
## 12         13 Logging    74.40       3.50
## 13         14 Logging    66.56       3.57
## 14         20 Logging    57.97       3.63
## 15          2    Park    52.17       3.75
## 16          3 Logging    59.29       3.75
## 17         16 Logging    72.99       3.75
## 18          1    Park    52.66       3.78
## 19          4    Park    52.62       3.88
```

It seems to show that Transect #8, with RA_Birds of 73.06%, has the most open canopy in the set of Transects with RA_Birds >= 50%.

Knit and Upload

Please knit your work as a .pdf or .html file and upload to Canvas. Homework is due before the start of the next lab. No late work is accepted. Make sure to use the formatting conventions of RMarkdown to make your report neat and clean!