

Ouachita Baptist University

Scholarly Commons @ Ouachita

Honors Theses

Carl Goodson Honors Program

4-17-2024

Tree Canopy Cover Influences Habitat Use of Breeding Birds at Jack Mountain Wildlife Management Area

Kaleigh Thomas

Ouachita Baptist University

Follow this and additional works at: https://scholarlycommons.obu.edu/honors_theses



Part of the [Biodiversity Commons](#), [Forest Management Commons](#), and the [Ornithology Commons](#)

Recommended Citation

Thomas, Kaleigh, "Tree Canopy Cover Influences Habitat Use of Breeding Birds at Jack Mountain Wildlife Management Area" (2024). *Honors Theses*. 927.

https://scholarlycommons.obu.edu/honors_theses/927

This Thesis is brought to you for free and open access by the Carl Goodson Honors Program at Scholarly Commons @ Ouachita. It has been accepted for inclusion in Honors Theses by an authorized administrator of Scholarly Commons @ Ouachita. For more information, please contact mortensona@obu.edu.

THESIS:

**Tree Canopy Cover Influences Habitat Use of Breeding Birds at Jack Mountain Wildlife
Management Area**

Kaleigh N. Thomas

Ouachita Baptist University

THO68373@obu.edu

April 2024

Table of Contents

Abstract.....	4
Introduction	5
Methods.....	7
Field methods	7
Statistical analyses	8
Results.....	12
Discussion	19
Interdisciplinary Connection.....	20
References	23
Appendix	26

List of Tables and Figures

Figure 1: Map of 94 data collection points in Jack Mountain Wildlife Management Area	8
Table 1: Linear models used in the multiple regression for bird diversity	9
Table 2: Linear models used in the multiple regression for bird abundance.	10
Figure 3: Regression tree for variables affecting species diversity.....	11
Figure 4: Regression tree for variables affecting species abundance.....	12
Table 3: Linear models for species diversity organized by Δ AIC value	13
Table 4: Linear models for species abundance organized by Δ AIC value	14
Figure 5: Scatterplots showing the relationships between the explanatory variables and species diversity and abundance	15
Figure 6: Bar chart showing patient's who received discharge education	20

Abstract

Across the United States, bird populations have declined due to habitat loss. To better understand habitat use by birds, researchers observed populations at Jack Mountain Wildlife Management Area in southwest Arkansas. Students at Ouachita Baptist University surveyed 94 point count locations to estimate species diversity (total number of species observed) and species abundance (total number of individuals observed). At each point, students recorded the percentage of tree canopy cover, ground cover, midstory cover, and shrub cover. These variables were used as explanatory variables in multiple regression analyses to determine which variables were influential in explaining variation in species diversity and species abundance. Locations with higher species diversity had higher midstory cover ($P=0.003$) and lower canopy cover ($P=0.020$). Higher species abundance was found at locations with lower canopy cover ($P<0.001$). These findings suggest that the amount of canopy cover plays a substantial role in habitat use by breeding birds at Jack Mountain Wildlife Management Area.

Understanding how to use statistics software is useful to many disciplines including the field of nursing. As a nursing student, I have learned that the statistical analysis performed in this study is applicable to research projects completed in the field of healthcare. The statistical software environment R, and its interface RStudio 2, can be used by anyone performing statistical analysis, creating tables or graphs, and analyzing data. There are many packages associated within R (R Core Team, 2020) that would be beneficial to nurses and nurse management conducting performance improvement projects where they collect and analyze data pertaining to patient care, infection rates, and patient satisfaction.

Introduction

In the United States, the overall bird population is decreasing as more and more species populations decline (Rosenburg et al, 2019). Bird species are declining due to habitat loss caused by deforestation, climate change, and other human effects (Rosenburg et al, 2019). The loss of birds in habitats can lead to serious consequences within these habitats. Birds are responsible for seed dispersal, pollination of plants, and controlling insect populations (Rosenburg et al, 2019). Birds are a key component of connecting ecosystems in a way other animals do not. Many bird species migrate seasonally taking them to multiple ecosystems in the same year. The decline in species population and diversity is a reversible problem (Rosenburg et al., 2019). To preserve bird populations, it is essential that large areas with overlapping habitats be preserved for the use of wildlife (Dietz et al, 2020). The Jack Mountain Wildlife Management Area in southwestern Arkansas is a contiguous habitat, about 20,000 acres total (Ross Foundation, 2010). It is owned by the Ross Foundation and managed by the Arkansas Department of Fish and Game, with the goal of conservation of wildlife habitats. Previous research conducted by students from Ouachita Baptist University shows that the Jack Mountain Wildlife Management Area is made up of pine, mixed, and deciduous forests, with pine-dominated forests having more species diversity than deciduous or mixed woodlands (Pruett, 2022).

Using data collected from previous research students, I conducted a multivariate analysis on bird-species diversity and abundance to attempt to understand why pine habitats have greater diversity than other habitats. I examined the association between bird diversity and abundance with canopy coverage, shrub coverage, ground coverage, and midstory

coverage. I chose to measure these four variables because more complex habitats have been associated with higher diversity (Pruett, 2022). I performed statistical analyses using the software R (R Core Team, 2020), which is a free-statistical environment that can do a variety of statistical analyses and create many different types of charts and graphs (which are incorporated into this study).

Methods

Field methods

To calculate bird species diversity and abundance, research students at Ouachita Baptist University went to 94 randomly selected point locations within Jack Mountain Wildlife Management Area in southwestern Arkansas (Fig. 1). At each point, all species heard or seen within a 50 m radius during a ten-minute survey window were recorded to enable estimation of species diversity and abundance (Bibby et al. 2000). Surveys were conducted within four hours of sunrise (Huff et al. 2000).

Habitat complexity was also assessed at each of the point locations. Indices of vegetation structure were based on the following four measurements (Sam et al. 2019). Ground layer density (plants <0.5 m in height) was measured 15 times for each point using randomly chosen 1 m² plots. Shrub density (plants 1-3 m in height) and mid-story density were measured in 5 areas per point using scatter plots. Canopy density was measured by photographing up through the canopy at 5 locations randomly chosen near the point.

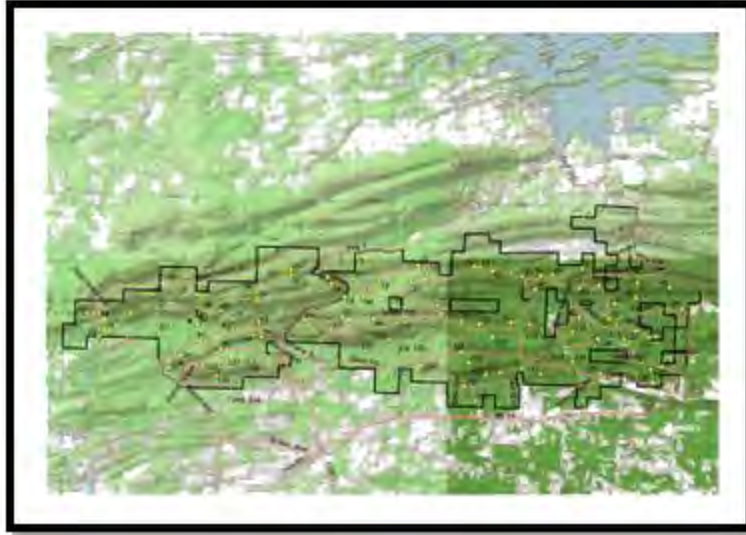


Figure 1: Map of the 94 data collection points (yellow dots) at Jack Mountain Wildlife Management Area in southwestern Arkansas. Black line denotes the boundary of the wildlife management area.

Statistical analyses

A series of statistical tests was performed on the data collected at Jack Mountain Wildlife Management Area to examine the data for correlation between each of the explanatory variables (ground coverage, shrub coverage, canopy coverage, and midstory coverage) and the response variables (bird diversity and abundance). Prior to completing a multivariate analysis, we had to determine if any of the explanatory variables correlated with each other.

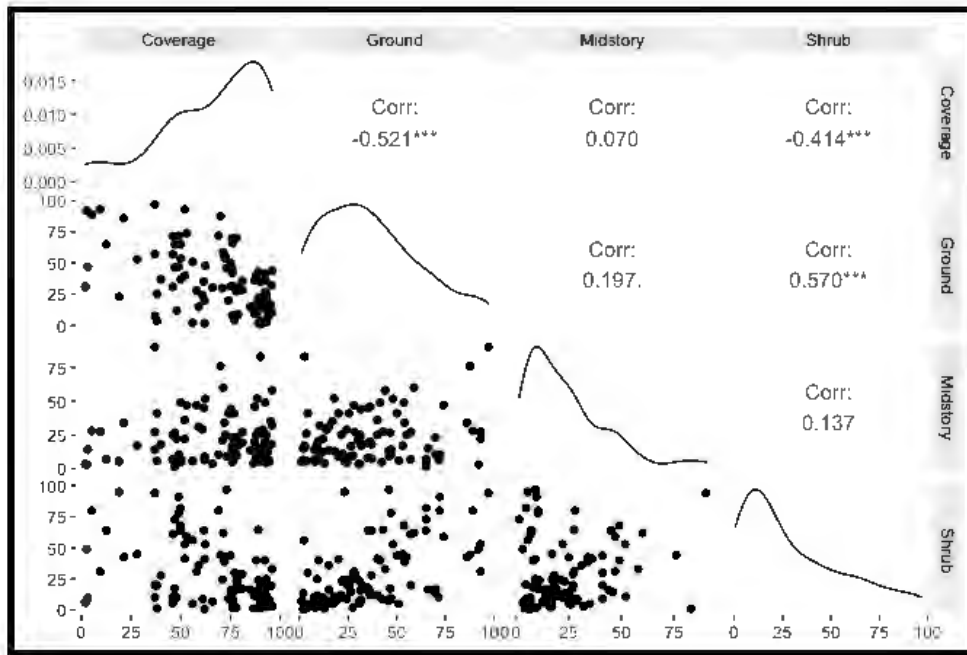


Figure 2: Scatter-plot matrices that show the correlations between explanatory variables. The asterisks indicate the significance of the correlation between two variables. Significance was found when $P < 0.001$ (R Core Team, 2020). The symbol “***” indicates $P < 0.001$ (R Core Team, 2020). The numbers signify the correlation coefficient between the variables (a higher number indicates a strong correlation).

Using the function “ggpairs” (Emerson et al, 2012) in R (R Core Team, 2020), a scatter-plot matrix was developed (Figure 2). The scatter-plot matrix shows correlations between the explanatory variables and the significance of each correlation using the correlation coefficient. The three significant correlations (ground and shrub, coverage and shrub, and coverage and ground) all display a moderate association (Asuero et al., 2007), that is, when the r value is between 0.40 and 0.59 (Asuero et al., 2007). All three correlations all have r values falling within that range. Ground and shrub has an r-value of 0.570, coverage and shrub has an r-value of -0.414 , and coverage and ground has an r-value of - 0.521. The functions “caret” (Kuhn,

2021) and “car” (Fox and Weisburg, 2019) were used to determine the collinearity of ground coverage, shrub coverage, midstory coverage, and canopy coverage. The values from the test were relatively small (ground coverage= 1.8214; canopy coverage= 1.4930; midstory coverage= 1.0978; shrub coverage= 1.5917), which shows that the correlation found in the scatter-plot matrix is not substantial, so all habitat variables were used as explanatory variables in a linear model (Stine, 1995).

I used the package “tree” (Ripley, 2021) in R (R Core Team, 2020) to create two regression tree models (Figure 3 and 4, page 12). Regression trees use raw data to determine the most significant variables in relation to how they impact species abundance and diversity. I used the results from the regression trees (Figure 3 and 4, page 12) to inform the linear models I chose to test (Table 1) (Crawley, 2015). Table 1 lists the linear models used for bird diversity, while Table 2 lists the linear models used for bird abundance.

Short-hand Names of Linear Models	Full Names of Linear Models
G	Ground Coverage
C	Canopy Coverage
G,C	Ground Coverage and Canopy Coverage
G,S	Ground Coverage and Shrub Coverage
S,C	Shrub Coverage and Canopy Coverage
G,C,S	Ground Coverage, Canopy Coverage, and Shrub Coverage
G,S,M	Ground Coverage, Shrub Coverage, and Midstory Coverage
C*S	Interaction between Canopy Coverage and Shrub Coverage
C*M	Interaction between Canopy Coverage and Midstory Coverage
G*C	Interaction between Ground Coverage and Canopy Coverage

Table 1: Table 1 lists the full names of the linear models used in the multiple regression for bird diversity.

Short-hand Names of Linear Models	Full Names of Linear Models
G	Ground Coverage
C	Canopy Coverage
G,C	Ground Coverage and Canopy Coverage
G,M	Ground Coverage and Midstory Coverage
G,S	Ground Coverage and Shrub Coverage
C,M	Canopy Coverage and Midstory Coverage
G,M,C	Ground Coverage, Midstory Coverage, and Canopy Coverage
G,M,S	Ground Coverage, Midstory Coverage, and Shrub Coverage
G*C	Interaction between Ground Coverage and Canopy Coverage
C*S	Interaction between Canopy Coverage and Shrub Coverage

Table 2: Table 2 lists the full names of the linear models used in the multiple regression for bird abundance.

The multiple regression analysis required many models to determine which variables best explained the variation in species diversity and species abundance (Table 1). I ranked the models by AIC value (Akaike Information Criterion). The AIC value indicates how well the model fits the data (Awad, 1995). A lower AIC value indicates a better fitting model. I used the package “AICcmodavg” (Mazerolle, 2020), which took all the models and ranked them in order of AIC value.

After running the multiple regression, I created a series of scatter plots (Figure 5) to show the relationships between the explanatory variables and response variables. I used the package “ggplot2” (Wickham, 2016) in R (R Core Team, 2020) to create the plots.

Results

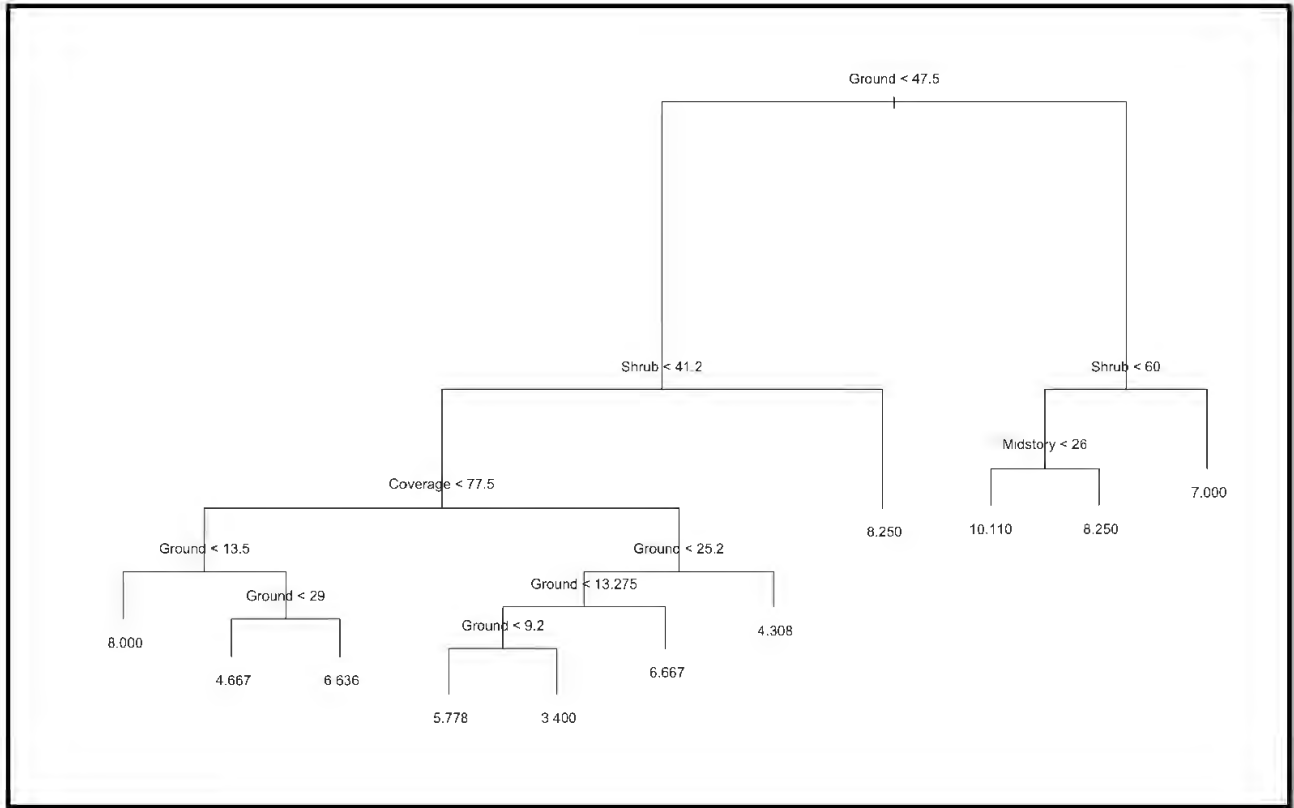


Figure 3: The regression-tree plots show the ranking of variables in order of significance in how they affect species diversity.

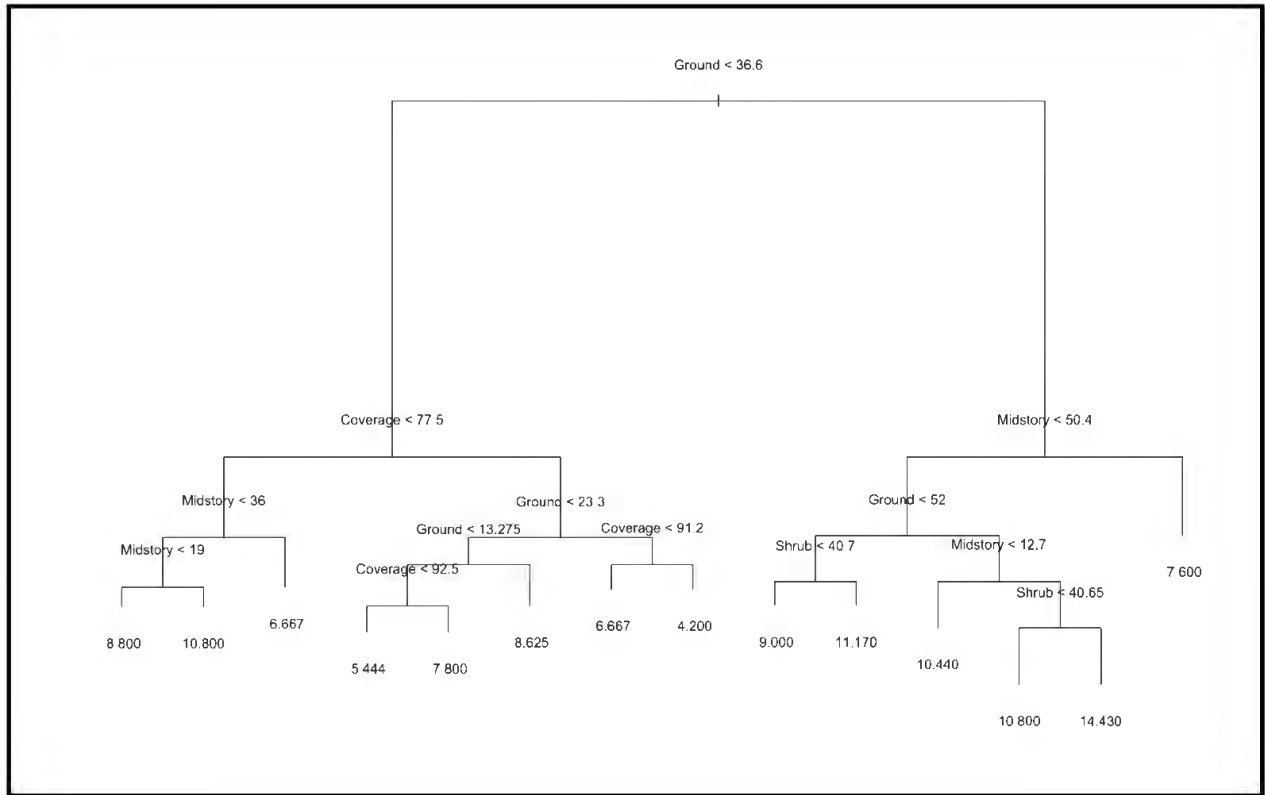


Figure 4: The regression-tree plots show the ranking of variables in order of significance in how they affect species abundance (bottom).

The tree models (Figures 3 and 4) were created to help determine how to group the explanatory variables for the linear models. For both diversity and abundance, ground coverage was determined as the most important variable because it is at the top of the tree model. For diversity, shrub coverage was the second most important variable because it was right below ground coverage on both branches of the tree. For species abundance, coverage and midstory were the next most important variables. Using the data from the tree plots, I made linear models and ranked them in order of increasing AIC value (Tables 3 and 4).

Models	R ²	AIC	Δ AIC	AICcWt
G,C	0.1661	428.340	0.00	0.330
G*C	0.1755	429.500	1.16	0.180
G	0.1280	430.360	2.02	0.120
G,C,S	0.1661	430.670	2.23	0.110
C	0.1246	430.720	2.38	0.100
S,C	0.1321	432.090	3.75	0.050
G,S	0.1287	432.470	4.13	0.040
C*S	0.1401	433.450	5.11	0.030
G,S,M	0.1392	433.560	5.22	0.020
C*M	0.1247	435.130	6.79	0.010

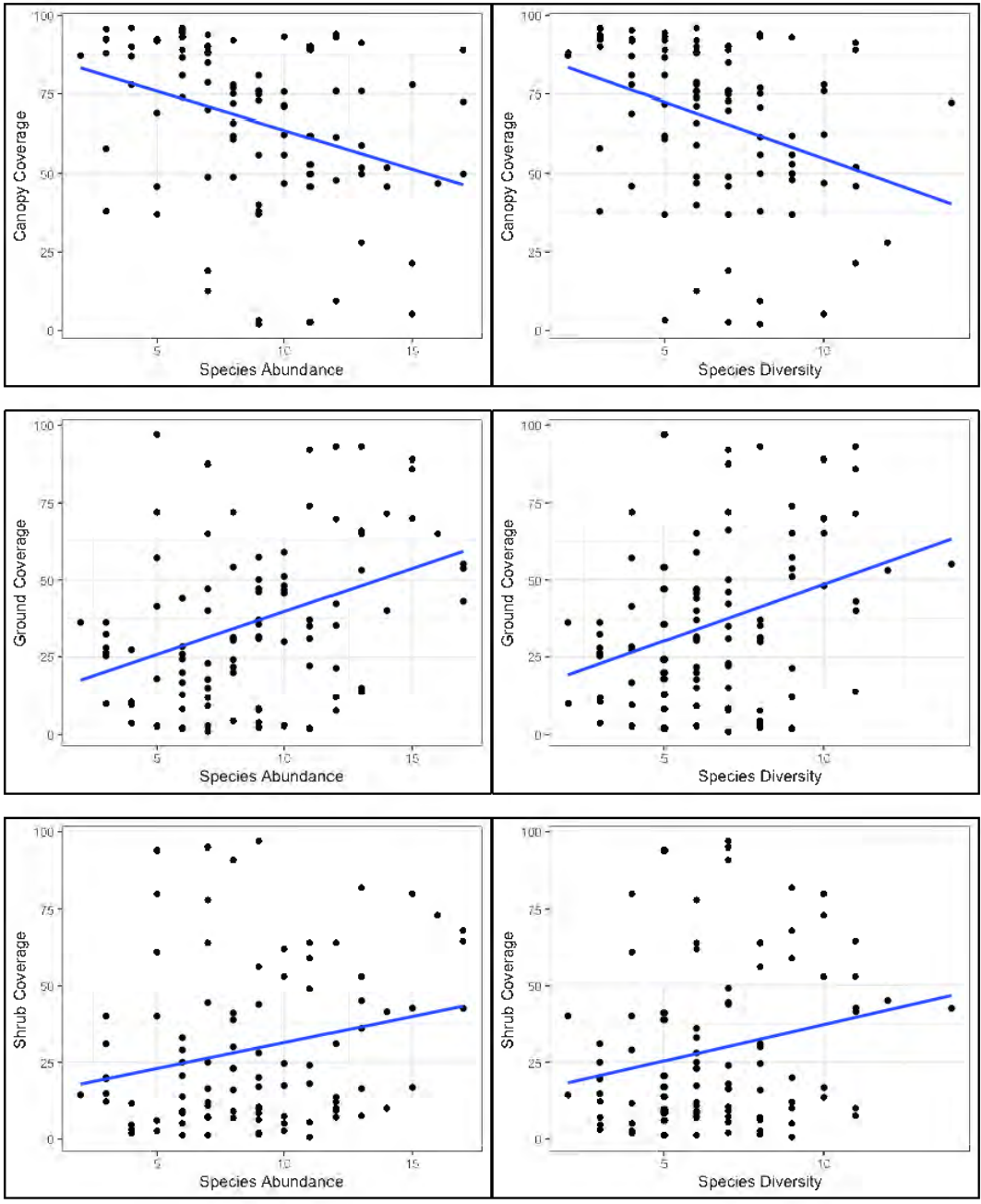
Table 3: The models above were used to determine the best linear model for species diversity. These models were chosen based on Δ AIC value. We included the models with a Δ AIC less than 7 because these models together accounted for 100% of the variation (AICcWT). These were the best models based on AIC value. G= ground coverage; M= midstory coverage; C= canopy coverage; and S= shrub coverage.

Model	R ²	AIC	Δ AIC	AICcWt
G,M	0.181	494.600	0.000	0.220
G,M,C	0.181	494.730	0.130	0.200
G,C	0.181	494.750	0.150	0.200
G	0.181	495.700	1.100	0.130
G*C	0.184	496.420	1.820	0.090
G,M,S	0.181	496.800	2.200	0.070
G,S	0.152	497.870	3.270	0.040
C	0.119	499.210	4.620	0.020
C,M	0.124	500.920	6.330	0.010
C*S	0.143	501.110	6.510	0.010

Table 4: The models above were used to determine the best linear model for species abundance. These models were chosen based on Δ AIC value. These were the best models based on AIC value. We included the models with a Δ AIC less than 7 because these models together accounted for 100% of the variation (AICcWT). G= ground coverage; M= midstory coverage; C= canopy coverage; and S= shrub coverage.

The linear models were ranked from lowest AIC value to highest AIC value (Tables 3 and 4). The best linear model was chosen based on the lowest AIC value. The best models for species diversity (AIC=428.34) and for species abundance (AIC= 494.6) both have the lowest AIC values. The Δ AIC value indicates how similar the model is to the best fitting model. The models with a Δ AIC less than two are considered the best fitting models. There were multiple models that had a Δ AIC less than two for species diversity. One model includes ground cover and canopy cover and the other model includes the interaction between the two variables. The model that explains the most variation in species diversity includes ground coverage and canopy coverage, explaining 33% of the variation in the data (AICcWT=0.33).

For species abundance, there were several models that had a Δ AIC value less than two. Those models include ground and midstory; ground coverage, midstory coverage, and canopy coverage; ground coverage and canopy coverage; ground coverage; and the interaction between ground coverage and canopy coverage. These could all be considered a “best fitting model” based on their Δ AIC values. The best fitted model for species abundance included ground coverage and mid-story coverage, explaining 22% of the variation in the data (AICcWT=0.22). The r^2 values indicate how much variation each model explains. The best fitting model for species diversity has an r^2 value of 0.1661. The best fitting model for species abundance has an r^2 value of 0.1806.



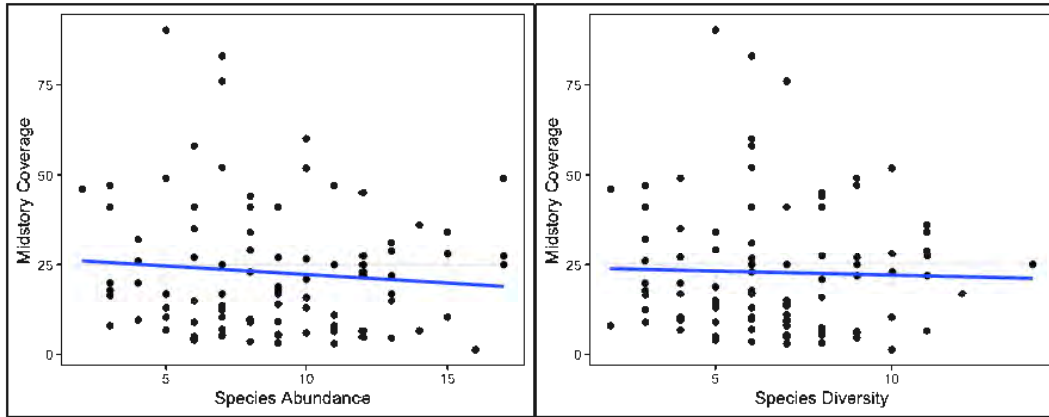


Figure 5: Scatter plots showing the relationship between the explanatory variables and species abundance and diversity. On the top row, from left to right: canopy coverage and species abundance ($P < 0.001$; $r^2 = 0.128$), canopy coverage and species diversity ($P < 0.001$; $r^2 = 0.1515$), ground coverage and species abundance ($P < 0.001$; $r^2 = 0.0048$), ground coverage and species diversity ($P < 0.001$; $r^2 = 0.1192$), shrub coverage and species abundance ($P = 0.029$; $r^2 = 0.050$), shrub coverage and species diversity ($P = 0.0246$; $r^2 = 0.0536$), midstory coverage and species abundance ($P = 0.7805$; $r^2 < 0.001$), midstory coverage and species diversity ($P = 0.3851$; $r^2 = 0.008$).

The scatterplots show species abundance and diversity and their relationship to canopy coverage, ground coverage, midstory coverage, and shrub coverage (Figure 4), as these variables were found in the best fitted models. The scatterplots were created using the best linear models (ground coverage and canopy coverage for species diversity; ground coverage and midstory coverage for species abundance). There is a strong negative correlation between canopy coverage and species abundance ($p < 0.001$), and there is a strong negative correlation between canopy coverage and species diversity ($p < 0.001$). This means the number of species and the number of birds increased as the canopy coverage decreased. There is a strong positive correlation between ground coverage and species abundance ($p < 0.001$), and there is a strong positive correlation between ground coverage and species diversity ($p < 0.001$). There are more birds and a greater variety of species when there is more ground coverage. There is a positive correlation between shrub coverage and species abundance ($p = 0.029$), and there is a positive correlation between shrub coverage and species diversity ($p = 0.0246$). There is not a significant

correlation between mid-story coverage and species abundance ($p= 0.7805$), or mid-story coverage and species diversity ($p= 0.3851$).

Discussion

Understanding that canopy coverage, ground coverage, and shrub coverage are three significant factors in bird diversity and abundance allows conservationists to make plans that create optimal environments for the birds living in the area. It is also important for conservationists to understand how these variables interact with one another. Based on the multivariate analysis, the three most significant factors influencing bird abundance and diversity were canopy coverage, shrub coverage, and ground coverage. As canopy coverage decreases, the variety of species and the abundance of birds increase. As ground and shrub coverage increases, the variety of species and the number of birds increase. This makes sense because more canopy coverage allows for less sunlight to support the midstory, ground, and shrub coverage. Less canopy coverage allows for light to enter the forest, which promotes growth underneath the canopy level (Lewandowski et al., 2020). Previous studies have found that even one tree being removed from the canopy allows sunlight to reach the lower levels of the forest, creating more bird diversity and abundance (Lewandowski et al., 2020). Using this information, researchers can advocate for better environmental protection in forests to allow for optimal habitats for the birds nesting there. For example, to create optimal conditions for promoting bird diversity and abundance, people could remove already dead or diseased trees from the forest to create spaces in the canopy coverage (Lewandowski et al., 2020). By creating a small break in the canopy coverage, conservationists could promote the creation of bird habitats in that area (Lewandowski et al., 2020).

Interdisciplinary Connection

The software used in the research in this study (R (R Core Team, 2020) can be used in research for other disciplines. The healthcare industry is constantly changing and improving as the result of research conducted in various hospitals, clinics, and laboratories. Utilizing the R (R Core Team, 2020) statistical environment would be beneficial for people in a variety of disciplines, including the field of nursing. In particular, nurse managers and supervisors are often responsible for conducting performance improvement projects that help with patient outcomes. For example, if a particular floor in the hospital has a high number of patients that develop secondary infections while in the hospital, the nurse managers might be tasked with determining a way to reduce that number through various methods. The research projects nurses complete can be amplified and improved by conducting statistical analysis on the results.

R (R Core Team, 2020) is the idea software for nurse managers to use if they only have a basic understanding of statistics and need to perform statistical analysis, or if they simply need software to make charts and graphs. R (R Core Team, 2020) and its companion software RStudio is open access, simple to use, and could be incorporated into the nursing school curriculum or offered as a continuing-education class for Registered Nurses.

R (R Core Team, 2020) is already being used in nursing. For example, the R (R Core Team, 2020) package “shiny” is useful in helping nurses create graphs for their research (Heinsberg, 2022). Most nursing managers do not need to use R (R Core Team, 2020) for any major statistical analysis for what is required of their performance improvement projects.

However, statistical analysis can strengthen the impact of their data. Nurses may also need to perform statistical analysis when creating budgets, analyzing staffing, and looking at patient outcomes.

I conducted a performance improvement project through my Leadership and Management class at Encore, a location of Arkansas Heart Hospital. For the project, I worked with nurse leaders on ways to improve patient education on newly prescribed medications at the time of discharge from the emergency department. To begin the project, I examined 20 patient charts with one of the nurse managers to see if there was a lack of documented patient education. We analyzed if the patients signed a document confirming that they received discharge education on their newly prescribed medications. For this project, I used R (R Core Team, 2020) to create a bar graph (Tutorialspoint) to show the difference between the number of patients who received discharge education and the patients who did not receive discharge education.

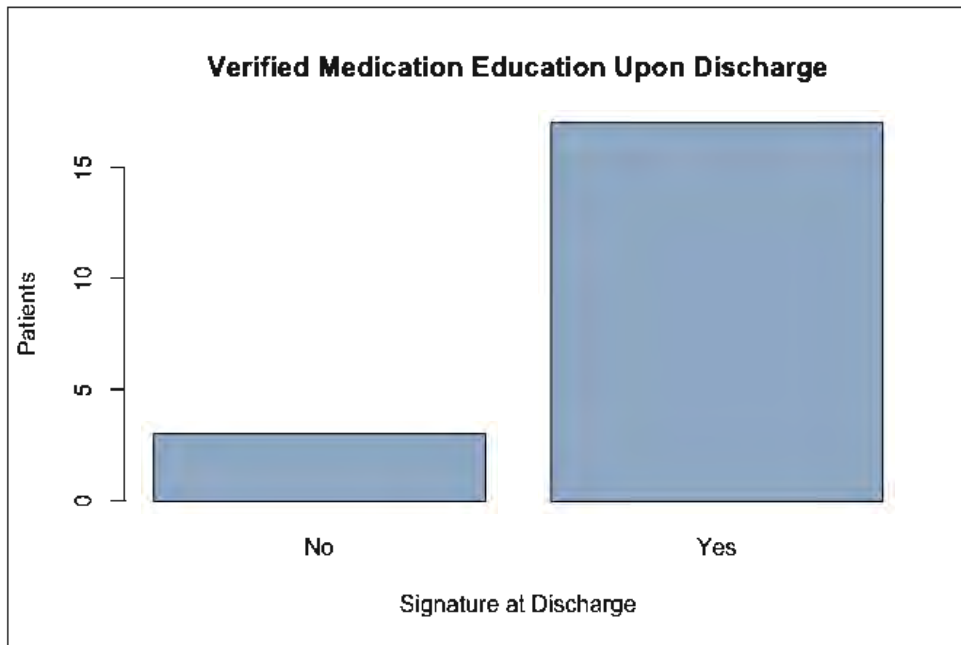


Figure 6: Bar graph showing the number of patient charts that contained a patient signature confirming discharge education was completed and the patient charts that did not include a signature confirming discharge education was completed.

To create the bar graph, I used code from Tutorialspoint.com, a website that has code for many different functions in R (R Core Team, 2020), including bar graphs. R (R Core Team, 2020) enables the user to select the color for the graphics. I selected the color “slategray3” (Frazier, n.d) to create the color above. Being able to change the color of the graphics makes R (R Core Team, 2020) more appealing to students of many disciplines. The bar graph (Figure 5) shows that most charts that were audited had patient signatures confirming they received discharge education. I also completed the Chi-squared Goodness of Fit test using R (R Core Team, 2020). The resulting p-value was 0.001745. This indicates that the amount of people who signed that they received discharge education is significantly greater than those who did not. Using R (R Core Team, 2020) for the Performance Improvement Project is just one example of how R (R Core Team, 2020) could benefit the medical field.

References

- Asuero, A. G., et al. "The Correlation Coefficient: An Overview." *Critical Reviews in Analytical Chemistry*, vol. 36, no. 1, 12 Jan. 2007, pp. 41–59.
- Awad, Adnan M. "PROPERTIES OF THE AKAIKE INFORMATION CRITERION." *Microelectron. Reliab.*, vol. 36, no. 4, 18 July 1995, pp. 457–464.
- Bibby, C.J., Burgess, N.D., Hill, D.A. and Mustoe, S. 2000. Bird census techniques. Elsevier.
- Crawley, Michael J. *Statistics: An Introduction Using R*. Wiley, 2015.
- Dietz, M.S., Belote, R.T., Gage, J. and Hahn, B.A. 2020. An assessment of vulnerable wildlife, their habitats, and protected areas in the contiguous United States. *Biological Conservation* 248: 108646
- Emerson, John W., Green, Walton A., Schloerke, Barret., Crowley, Jason., Cook, Dianne., Hofmann, Heike., Wickham, Hadley. The Generalized Pairs Plot. *Journal of Computational and Graphical Statistics*, vol. 22, no. 1, pp. 79-91, 2012.
- Fox, John. Weisburg, Sanford (2019). *An {R} Companion to Applied Regression*, Third Edition. Thousand Oaks CA: Sage.
- Frazier, M. (n.d.). *R color cheatsheet*. National Center for Ecological Analysis and Synthesis- University of California Santa Barbara.
<https://www.nceas.ucsb.edu/sites/default/files/2020-04/colorPaletteCheatsheet.pdf>

Heinsberg, L. W., Koleck, T. A., Ray, M., Weeks, D. E., & Conley, Y. P. (2022, August). *Advancing nursing research through interactive data visualization with R shiny*. Biological research for nursing. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9900251/>

Huff, M.H., Bettinger, K.A., Ferguson, H.L., Brown, M.J. and Altman, B., 2000. A habitat-based point-count protocol for terrestrial birds, emphasizing Washington and Oregon. UNITED STATES DEPARTMENT OF AGRICULTURE FOREST SERVICE GENERAL TECHNICAL REPORT PNW.

Kuhn, Max (2021). caret: Classification and Regression Training. R package version 6.0-90.

Lewandowski, P., Przepióra, F., & Ciach, M. (2020, October 25). *Single dead trees matter: Small-scale canopy gaps increase the species richness, diversity and abundance of birds breeding in a temperate deciduous forest*. Forest Ecology and Management. <https://www.sciencedirect.com/science/article/abs/pii/S0378112720314626>

Mazerolle, Marc J. (2020) AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c). R package version 2.3-1.

Pruett, C. L. 2022. Evaluation of breeding bird distributions and habitat use on Ross Foundation Property at Jack Mountain WMA. Final Report Ross Foundation, Arkadelphia, AR.

R - bar charts. Tutorialspoint. (n.d.). https://www.tutorialspoint.com/r/r_bar_charts.htm

Recreational Opportunities. The Ross Foundation. (2010). https://rossfoundation.us/recreational_opportunities.html#:~:text=Additionally%2C%20the%20Ross%20Foundation%20also,Jack%20Mountain%20Wildlife%20Management%20Area.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. R version 4.0.3 (2020-10-10)

Ripley, Brian. (2021). tree: Classification and Regression Trees. R package version 1.0-41.

Rosenberg, K.V., Dokter, A.M., Blancher, P.J., Sauer, J.R., Smith, A.C., Smith, P.A., Stanton, J.C., Panjabi, A., Helft, L., Parr, M. and Marra, P.P., 2019. Decline of the North American avifauna. *Science*, 366(6461), pp.120-124.

<https://www.science.org/doi/10.1126/science.aaw1313>

Sam, K., Koane, B., Bardos, D.C., Jeppy, S. and Novotny, V. 2019. Species richness of birds along a complete rain forest elevational gradient in the tropics: Habitat complexity and food resources matter. *Journal of Biogeography* 46: 279-290.

Stine, Robert A. "Graphical Interpretation of Variance Inflation Factors." *The American Statistician* , vol. 49, no. 1, Feb. 1995, pp. 53–56.

Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Appendix

The following code was put into R Studio 2 (R Core Team, 2020) to provide statistical analysis and graphics. The code is provided so that others may rerun the data set or insert their own data sets. Any statements after a hashtag symbol (#) include notes on how to run the code and what the purpose of the code is.

```
model1<-  
  lm(Complex_birds$Species_Diversity~Complex_birds$Date+Complex_birds$Coverage+C  
    omplex_birds$Ground+Complex_birds$Midstory+Complex_birds$Shrub)  
summary(model1) #Input data into Model 1 and then provides standard deviations and t-values  
                which showed significance between Diversity and Coverage. The data shows GREATER  
                significance between Diversity and Ground.  
par(mfrow=c(2,2))  
plot(model1) #Plots data and shows the data is normal.  
X<-Complex_birds[,7:10] #includes columns: point, coverage, ground, midstory, shrub  
library(GGally)  
library(ggplot2)  
ggpairs(X) #This graphs Figure 2. Figure 2 shows correlation between: ground and coverage;  
           shrub and coverage; shrub and ground. All three are highly correlated with "***".  
library(corpcor)  
cor2pcor(cov(X)) # just shows graph in numerical format  
#removed coverage  
model2<-  
  lm(Complex_birds$Species_Diversity~Complex_birds$Date+Complex_birds$Ground+Co  
    mplex_birds$Midstory+Complex_birds$Shrub)  
summary(model2)
```

```

#removed ground
model3<-
  lm(Complex_birds$Species_Diversity~Complex_birds$Date+Complex_birds$Midstory+Complex_birds$Shrub)
summary(model3)
#removed shrub
model4<-lm(Complex_birds$Species_Diversity~Complex_birds$Date+Complex_birds$Midstory)
summary(model4) # Created different data models to compare variables.
anova(model1,model2,model3,model4)
plot(model1,model2,model3,model4)
#AIC
AIC(model1)
BIC(model1)
AIC(model2)
AIC(model3)
AIC(model4) #This package determined the AIC values of various models comparing variables
and species diversity. This says Model 1 is the best because it has the lowest AIC value
(432.0512)
model5<-
  lm(Complex_birds$Species_Diversity~Complex_birds$Shrub+Complex_birds$Ground+Complex_birds$Coverage) #Model with variables previously removed
summary(model5)
#The following models (A-D) were created comparing species abundance and the variables.
modelA<-
  lm(Complex_birds$Species_Abundance~Complex_birds$Date+Complex_birds$Coverage
+Complex_birds$Ground+Complex_birds$Midstory+Complex_birds$Shrub)
summary(modelA)
plot(modelA)
modelB<-
  lm(Complex_birds$Species_Abundance~Complex_birds$Date+Complex_birds$Coverage

```

```
+Complex_birds$Midstory+Complex_birds$Shrub) # Ground was removed because it
was the only variable that showed correlation with species abundance.
```

```
summary(modelB)
```

```
modelC<-
```

```
lm(Complex_birds$Species_Abundance~Complex_birds$Date+Complex_birds$Midstory
+Complex_birds$Shrub) # The variable canopy coverage was removed because it was
the only variable that showed correlation with species abundance.
```

```
summary(modelC)
```

```
modelD<-
```

```
lm(Complex_birds$Species_Abundance~Complex_birds$Date+Complex_birds$Midstory
) # The variable shrub coverage was removed because it was the only variable that
showed correlation with species abundance
```

```
summary(modelD)
```

```
# The AIC values of models A through D were calculated.
```

```
AIC(modelA)
```

```
BIC(modelA)
```

```
AIC(modelB)
```

```
AIC(modelC)
```

```
AIC(modelD)
```

```
#Best model is model 1
```

```
#Variance inflation factors in the package "car". You need to open "car" to do this function.
```

```
library(caret)
```

```
library(car)
```

```
car::vif(model1)
```

```
model1<-
```

```
lm(Complex_birds$Species_Diversity~Complex_birds$Date+Complex_birds$Coverage+C
omplex_birds$Ground+Complex_birds$Midstory+Complex_birds$Shrub)
```

```
par(mfrow=c(1,1))
```

```
library(tree)
```

```
modeltree<-tree(Species_Diversity~.,data=Complex_birds)
```

```

plot(modeltree)
text(modeltree)
#Remove abundance. only use explanatory variable and response variable that we are using.
diversity<-subset(Complex_birds,select=-c(Species_Abundance))
diversity
#removing date and species abundance
diversity2<-subset(Complex_birds2,select=-c(Date))
diversity2
modeltree2<-tree(Species_Diversity~.,data=diversity2)
plot(modeltree2) #This created the tree model comparing variables contributing bird species
                diversity.
text(modeltree2)
modeltree1<-tree(Species_Diversity~.,data=Complex_birds2)
plot(modeltree1)
text(modeltree1)
diversity3<-subset(Complex_birds,select=-
                  c(Species_Diversity,Date,HC_index,Point,Shannon_all,Shannon_no_Cov,Habitats))
diversity3
modeltree3<-tree(Species_Abundance~.,data=diversity3)
plot(modeltree3) #This created the tree model comparing variables contributing bird species
                abundance.
text(modeltree3)
# The most important explanatory variable is Ground and the threshold value separating low
  and high values of ground is 36.6 (%). The right hand branch of the tree

library(mgcv)
modelGAM<-
  gam(diversity2$Species_Diversity~s(diversity2$Coverage)+s(diversity2$Ground)+s(diver
    sity2$Midstory)+s(diversity2$Shrub))
plot(modelGAM)

```

```

fullmodel<-
  lm(diversity2$Species_Diversity~diversity2$Ground*diversity2$Shrub*diversity2$Cover
age*diversity2$Midstory+l(diversity2$Coverage^2)+l(diversity2$Ground^2)+l(diversity2
$Midstory^2)+l(diversity2$Shrub^2))

summary(fullmodel)

mod2<-update(fullmodel,~.-
  diversity2$Shrub:diversity2$Coverage:diversity2$Midstory:diversity2$Ground)

summary(mod2)

mod3<-update(mod2,~.-diversity2$Shrub:diversity2$Coverage:diversity2$Midstory)

summary(mod3)

mod4<-update(mod3,~.-diversity2$Ground:diversity2$Shrub)

summary(mod4)

mod5<-update(mod4,~.-l(diversity2$Shrub^2))

summary(mod5)

mod6<- update(mod5,~.-diversity2$Ground:diversity2$Shrub:diversity2$Coverage)

summary(mod6)

mod7<-update(mod6,~.-diversity2$Shrub:diversity2$Coverage)

summary(mod7)

mod8<-update(mod7,~.-l(diversity2$Midstory^2))

summary(mod8)

mod9<-update(mod8,~.-diversity2$Ground:diversity2$Coverage)

summary(mod9)

mod10<-update(mod9,~.-l(diversity2$Ground^2))

summary(mod10)

mod11<-update(mod10,~.-diversity2$Ground)

summary(mod11)

mod12<-update(mod11,~.-diversity2$Shrub)

summary(mod12) #everything is significant for species diversity

```

```
#AIC for Species Diversity
```

```
AIC(fullmodel)
```

```
AIC(mod2)
```

```
AIC(mod3)
```

```
AIC(mod4)
```

```
AIC(mod5)
```

```
AIC(mod6)
```

```
AIC(mod7)
```

```
AIC(mod8)
```

```
AIC(mod9)
```

```
AIC(mod10)
```

```
AIC(mod11)
```

```
AIC(mod12)
```

```
#repeated for species abundance
```

```
abundancemodel<-
```

```
  lm(diversity3$Species_Abundance~diversity3$Ground*diversity3$Shrub*diversity3$Coverage*diversity3$Midstory+l(diversity3$Coverage^2)+l(diversity3$Ground^2)+l(diversity3$Midstory^2)+l(diversity3$Shrub^2))
```

```
summary(abundancemodel)
```

```
modA<-update(abundancemodel,~.diversity3$Shrub:diversity3$Coverage:diversity3$Midstory)
```

```
summary(modA)
```

```
modB<-update(modA,~.-diversity3$Ground:diversity3$Shrub:diversity3$Coverage)
```

```
summary(modB)
```

```
modC<-update(modB,~.-diversity3$Ground:diversity3$Shrub)
```

```
summary(modC)
```

```
modD<-update(modC,~.-l(diversity3$Shrub^2))
```

```
summary(modD)
```



```

modE<-update(modD,~.-
    diversity3$Ground:diversity3$Shrub:diversity3$Coverage:diversity3$Midstory)
summary(modE)
modF<-update(modE,~.-diversity3$Shrub:diversity3$Coverage)
summary(modF)
modG<-update(modF,~.-I(diversity3$Midstory^2))
summary(modG)
modH<-update(modG,~.-diversity3$Coverage)
summary(modH)
modI<-update(modH,~.-diversity3$Ground:diversity3$Midstory)
summary(modI)
modJ<-update(modI,~.-diversity3$Midstory:diversity3$Coverage)
summary(modJ)
modK<-update(modJ,~.-diversity3$Midstory)
summary(modK)
modL<-update(modK,~.-diversity3$Ground)
summary(modL)
modM<-update(modL,~.-diversity3$Shrub)
summary(modM)
modN<-update(modM,~.-I(diversity3$Ground^2))
summary(modN)
modO<-update(modN,~.-diversity3$Ground:diversity3$Shrub:diversity3$Midstory)
summary(modO)
modP<-update(modO,~.-diversity3$Shrub:diversity3$Midstory)
summary(modP) #CONTINUE COPYING OUTPUT FROM HERE
#AIC for abundance
AIC(modA)

```

AIC(modB)

AIC(modC)

AIC(modD)

AIC(modE)

AIC(modF)

AIC(modG)

AIC(modH)

AIC(modI)

AIC(modJ)

AIC(modK)

AIC(modL)

AIC(modM)

AIC(modN)

AIC(modO)

AIC(modP)

#comparing AIC values

library(AICcmodavg)

models.diversity<-

```
list(fullmodel,mod2,mod3,mod4,mod5,mod6,mod7,mod8,mod9,mod10,mod11,mod12)
```

models.diversity

models.diversity.names<-

```
c('fullmodel','mod2','mod3','mod4','mod5','mod6','mod7','mod8','mod9','mod10','mod11','mod12')
```

aictab(cand.set = models.diversity,modnames=models.diversity.names)

```

models.abundance<-
  list(modA,modB,modC,modD,modE,modF,modG,modH,modI,modJ,modK,modL,modM,
  modN,modO,modP)

models.abundance

models.abundance.names<-
  c('modA','modB','modC','modD','modE','modF','modG','modH','modI','modJ','modK','mo
  dL','modM','modN','modO','modP')

aictab(cand.set = models.abundance,modnames=models.abundance.names)

plot(modL)

plot(Complex_birds2$Species_Diversity,Complex_birds2$Coverage)

plot(1/(Complex_birds2$Coverage^2),Complex_birds2$Species_Diversity)

```

The following code was used to make the bar-graph and do the Chi-squared test in the paper's interdisciplinary connection portion.

```

library(ggplot2)

L <-c(3,17)

barplot(L)

M <- c("No","Yes")

barplot(L,names.arg = M,xlab ="Signature at Discharge",ylab =
  "Patients",col="slategray3",main="Verified Medication Education Upon Discharge")

chisq.test(PI)

```