**The Importance of Street Trees to Urban Avifauna:**

*A lesson exploring the relationships between urban forest and foraging birds*

This QUBES lesson is based on a study conducted by Wood and Esaian called “The importance of street trees to urban avifauna.” Using this as a framework, students will explore the relationships between street trees, foraging birds, and socioeconomics in the urban setting of Los Angeles, California. A generalized linear model with negative binomial distribution will be used to analyze edited data from the focal paper in RStudio. This lesson will demonstrate why these models are used, how their interpretation can guide the management of urban forest, and how this process leads to better conservation of biodiversity in urban spaces.

**Learning Objectives**

* Explore effects of urbanization, function of urban forests, the role of socioeconomic gradients, and urban avifauna as bioindicators.
* Analyze the relationship between street tree density and feeding bird density across socioeconomic groupings using a generalized linear model with negative binomial distribution in R.
* Discuss how the results from the analysis could guide planning strategies for the assessment and management of street trees and the conservation of biodiversity.

**Prerequisites**

* Understanding of basic statistical principles for linear regressions and generalized linear models with Poisson distribution.
* General understanding of RStudio software.

**Urbanization**

Urbanization is the process of converting a natural ecosystem to one dominated by human development. Urbanization has been intensifying globally over the last half century, and it is projected that urban areas will continue to expand. Although there are associated benefits for humans, there have been negative impacts on the environment. One of the problematic aspects of human development has been the use of impervious surfaces such as concrete structures and pavement that dominate the landscape. Use of these materials has caused declines in natural areas which has led to the loss of biodiversity.

Biodiversity declines due to this type of development can be attributed to habitat loss or degradation, declines in connectivity, alters in disturbance regimes, and more (Elmqvist, Zipperer, and Guneralp, 2016). Given the historical and current uses of urban spaces, navigating the challenges associated with urbanization will be important to the conservation of biodiversity.

A city with trees and mountains in the background

Description automatically generated with medium confidence

***Figure 1.*** *Example of an urbanized landscape advancing into natural areas. Image source: Cities4Forests.*

[Click here](https://www.prb.org/resources/urbanization-an-environmental-force-to-be-reckoned-with/) for a more in-depth look at the history and environmental interactions of urbanization.

Want to learn more about biodiversity? [Click here.](https://www.youtube.com/watch?v=GK_vRtHJZu4)

**Urban Forestry**

Urban forests are the collection of trees within the boundaries of a metropolitan area. These forests come in different shapes and sizes and include parks, gardens, river corridors, and more. A key component of urban forests are street trees. Street trees are considered public resources that are planted by municipalities rights-of-way (along public roads and sidewalks).

Diagram

Description automatically generated

Urban forests are a way to mitigate the negative impacts on biodiversity in urban areas. There are a variety of functional services associated with street trees which include improving urban aesthetic, increasing the quality of life of residents, and providing numerous environmental benefits. One of the goals of street tree plans is to provide habitat for wildlife which could contribute to biodiversity. Due to their importance, many cities have developed plans that work to promote, maintain, and provide an inventory of trees with-in a city’s boundary.

***Figure 2.*** *Benefits of a well-planned, well-managed urban forest. Image: Cities4Forests, 2019.*

**Luxury Effect Hypothesis**

While there are many services provided by street trees, it is important to note that these services are not distributed equally. The luxury effect hypothesis is defined as the positive relationship between affluence and organism diversity or activity in urban ecosystems. This trend has been found in cities around the world and affects a variety plants and wildlife. In these cases, urban forests and street trees are more prevalent in higher income communities than lower income communities. The primary reason this occurs is the costs associated with maintaining street trees as it is typically the responsibility of the property owner to cover these costs. This can cause an unequal distribution of environmental benefits for people and wildlife (Leong, Dunn, and Trautwein, 2018).

A street with cars and trees

Description automatically generated with low confidence

*Figure 4. Example images of trees in areas of differing socioeconomic status. Image derived from figure 2 of the focal paper.*

**Avifauna as Bioindicators**

A bioindicator is an organism whose status in an ecosystem is analyzed as an indication of the ecosystem's heath. Birds are often considered excellent biodiversity indicators because they are common, easy to detect, diverse, and can reach all areas of an ecosystem. Birds are also considered good representatives of their ecosystems due to their diverse use of resources such as habitat and food. When determining the quality of an ecosystem, researchers typically look at the richness (number of species) and abundance (number of individuals) of avian populations. The higher richness and abundance of birds in an area, the better quality of habitat and resources (Fraixedas, et al., 2020).

Los Angeles has a high diversity and abundance of birds that consist of hundreds of migratory and year-round species that utilize urban ecosystems throughout the annual cycle. Year-round residents are present in the ecosystem throughout the year, while migratory birds are only present for part of the year. In Los Angeles, migrants are present during the winter where they spend their time fueling up for spring migration. Migratory birds are of particular interest due to the complex factors that influence their populations and their significant declines globally (Gilroy, et al., 2016). Given that the goal of many urban forest plans is to provide habitat for wildlife, understanding the relationships between birds and street trees will be essential to their conservation.

  

***Figure 3.*** *Examples of birds surveyed from the focal paper. From left to right: Lesser Goldfinch, Orange-crowned Kinglet, and a Yellow-rumped Warbler. Source of images:* [*allaboutbirds.org.*](file:///C:\Users\sarah\Documents\ENVS%20603\allaboutbirds.org)

**The Data**

The data for this lesson comes from a study published by Wood and Esaian in 2020 called “The importance of street trees to urban avifauna”. In this study, the authors wanted to understand the importance of street trees to wildlife. They did so by measuring and identifying public street trees and documenting foraging behavior of birds across two winters in residential communities situated across socioeconomic gradients throughout Greater Los Angeles (LA).

Map

Description automatically generated

***Figure 5.*** *A total of 36 sites were surveyed in LA across a socioeconomic gradient of low, medium, and high-income residential areas. These images provide a visual of the walking routes surveyed and a map of their locations. Image derived from figure 2 of the focal paper.*

There are several objectives in this study, however we will be focusing on part of objective #2:

Assess the relationship between street tree density and foraging bird density across socioeconomic gradients.

To assess this relationship, we will run a generalized linear model in R with a negative binomial distribution. This is similar to what the authors did in their study, however this process has been simplified for ease of interpretation.

**Before continuing the lesson, read the “Data Information” file to better understand the data and how it has been edited.**

Before going through the analysis, it is important to understand that the edited data will not produce the exact same figure as those in the paper. The trends and outcomes will be very similar to what was found in the paper, but not identical. Also, the statistical analysis is meant to teach basic negative binomial regression. We will not be following the exact methodology that the authors used in their study, as it is more complex than what this lesson is designed to teach.

**Negative Binomial Regression**

Negative binomial regression is a type of generalized linear model that has a negative binomial distribution (or family). This distribution is like Poisson in that it describes the probabilities of occurrence of positive whole numbers (aka count data). Unlike Poisson distribution, the variance does not equal the mean indicating that overdispersion is present. Overdispersion is when the variance of the **response variable** is greater than is assumed by the model.

The variance of a negative binomial distribution is a function of its mean and has an additional parameter called the dispersion parameter (k). The dispersion parameter is a measure of how much a sample fluctuates around the mean value. In other words, it tells us how many times larger the variance is than the mean. This parameter does not affect the expected counts of the data but does affect the expected variance.

Overdispersion in count data is a problem because it underestimates standard error which leads to smaller p-values. Essentially, there is a higher risk of making a type I error. Overdispersion is common in real data sets. This could be due to; predictor variables that were not included in the model, underlying heterogeneity (variation) in the sample, or there may be an excess of zero counts (requires zero inflation models).

In the R portion of this lesson, we will be assessing the influence of street tree density on total bird density. Total bird density will be the response variable and total tree density will be the predictor variable. Since there cannot be negative or partial counts of birds, this variable is considered count data.

For a GLM Refresher: [Click here.](https://www.dataquest.io/blog/tutorial-poisson-regression-in-r/#:~:text=A%20Poisson%20Regression%20model%20is,form%20by%20some%20unknown%20parameters.)

**Check your Understanding**

Before continuing to the R portion of the lesson, answer the following questions:

1. What are challenges that urbanization poses for biodiversity?

Challenges posed by biodiversity include habitat loss and degradation, loss of disturbance regimes, and a decrease in connectivity.

1. What are the benefits of urban forests? Are there any drawbacks?

* There are many benefits to urban forests including recreation, health, resources for wildlife (food, habitat, cover, etc.), aesthetics, flood management, temperature management, and more.
* Potential drawbacks are associated with the costs of maintaining street trees that can lead to environmental inequity as described by the Luxury Effect Hypothesis. Geographic features of the area as well as climate are considerations that may limit where/how urban forests are managed. The types of trees that are planted also should be considered as native plants have more benefits than introduced species.

1. Given birds are used as bioindicators, what kinds of trends would we expect to see in urban areas that display the luxury effect hypothesis? Would there be differences in birds for areas with low, medium, or high incomes?

* Researchers look at the richness and abundance of birds in an area to determine the quality of habitat or health of the ecosystem. Therefore, if the luxury effect hypothesis is occurring in a urban area, then we would expect for there to be more birds present in the higher income areas than the lower income areas.

1. When is it appropriate to use a negative binomial regression?

* Negative binomial regression is used to model overdispersed count data. This is when the variance of the response variable is greater than the mean.

**Outline of R Lesson**

* Visualize the data
* Create a histogram to observe the distribution of the data.
* To test for normality, run a Shapiro-Wilkes test.
* Both of these tests demonstrate that we are working with count data that does not follow a normal distribution.
* Run a generalized linear model with Poisson distribution
* This is why the prerequisite of previous experience with GLM’s is required.
* This lesson is building off of this previous knowledge.
* Using a Poisson Distribution is a natural response of researchers to this type of data as it is count data that does not follow a normal distribution (the authors in the paper started with this). However, there are ways to check if this model is a good fit or not and it is valuable to point this out to students.
* It also makes for a useful comparison when looking at negative binomial regression (NBR). Looking at the two model outputs demonstrates that the overdispersion is accounted in NBR by the dispersion parameter (k) which leads to a standard error that is larger. Given this larger standard error, we know that p-values are more accurate than in the Poisson (for this dataset).
* Check for overdispersion
* There are several ways to check for overdispersion in data. It is a bit repetitive given that the Poisson model output indicates overdispersion, however it is still valuable as it gives more specific information on how overdispersed the data is.
* Using the MASS package in R, run a generalized linear model with a negative binomial distribution to assess the relationship between street tree density and density of birds.
* The important takeaways from this output are the dispersion parameter, estimate, standard error, p-values, and residual deviance on degrees of freedom. The estimate should be the same (or very similar) as the one from the Poisson output. Standard error and p-values should be larger in the negative binomial model. The values for residual deviance on degrees of freedom should be closer together than in the Poisson model because this model is accounting for overdispersion (via the dispersion parameter) and is a better fit. The dispersion parameter (theta) is a shape parameter. The larger value of theta indicates more spread (dispersion) in the data than a smaller value.
* Build a scatterplot depicting the negative binomial regression.
* The scatter plot is a useful visualization tool for this relationship. There should be a positive relationship depicted by the data with more high-income points at the higher densities of trees. This is somewhat depicting the luxury effect hypothesis explained above (more on this later on). There are more birds present with higher densities of trees. What is even more interesting is just how many more trees (and therefore birds) there appears to be in high income areas than in lower income areas (generally). There is a moderate correlation occurring between the two variables (not necessarily strong or weak).
* Remember: these values are standardized to 1km walking routes.
* Note: The authors added a quadratic term to their plots because it better fit the trend of their data. The plot that I created did not appear to need one, so I did not add it. The data does not match up perfectly to the authors in the paper, however these skills are still useful (and the interpretation is similar).

Chart, scatter chart

Description automatically generated

* Create boxplots to visualize the relationship between socioeconomic group, tree density, and bird density.
* This is how we can look at the influence of socioeconomic grouping on tree density and bird density specifically. Important to note that further statistical analysis is needed to properly address these relationships, however they are still valuable to look at because you can get a general idea of these relationships.
* One of the code chunks asks the students to create a boxplot based off an example provided in the R code. Below is an example of what this should look like.

Chart, box and whisker chart

Description automatically generated

* Note: Coolors Palette Generator is a user-friendly way of generating colors (and their code for R). The link provided in the R code should take you to a page that appears like the image below (the colors might be different). To toggle through different palettes, hit the space bar. To use a color in R, just copy and paste the 6-digit code into the appropriate space. For an example, look at the code in the image above. The last three colors of this palette were used to create the boxplots.

Graphical user interface, treemap chart

Description automatically generated with medium confidence

* Answer questions throughout the lesson to check your understanding of the material.
* See the below sections for more guidance on the questions.

Need help with R? [Click here.](https://www.statmethods.net/r-tutorial/index.html)

**Questions from R lesson:**

1. What are the values for residual deviance and the associated degrees of freedom? Is overdispersion present in this data? Why might this be an issue?

* Residual deviance has a value of 201.29 on 34 degrees of freedom. There is overdispersion present in this data. This is an issue because it could be underestimating standard error which leads to smaller p-values giving a significant result where there may be none (Type 1 error).

1. Compare the negative binomial model to the Poisson. How is the negative binomial model accounting for the overdispersion present in the data? (Hint: consider coefficients) Is this model a better fit? Why?

* This model is accounting for overdispersion via the dispersion parameter. This is one of the main differences between Poisson and negative binomial modeling. Another difference between this model and the Poisson is the standard error and p-values. The standard error in the negative binomial is higher than in the Poisson model which makes sense because there is more uncertainty in the data due to the variability. This also provides more accurate p-values (which happen to still be significant).
* This model is a better fit. We can tell because of the residual deviance and the degrees of freedom being closer together in value (37 on 34 df) than in the Poisson model.

1. What does the figure above indicate about the relationship between street trees and bird density? Does there appear to be a pattern associated with socioeconomic groups?

* There is a moderate positive relationship between street tree density and total bird density. There appear to be more points for higher income areas at higher densities of trees and birds while lower and medium income points appear to be at lower densities of trees and birds.

1. Interpret the boxplots depicting street tree density and total bird density. Do you think that there is support for the luxury effect hypothesis in this study?

* Street tree density appears to be different across all three of the socioeconomic groups. The lowest density of trees being in the low-income group and the highest appearing in the highest income group. The low- and high-income groups appear to be the most differentiated from one another. Further testing would need to be conducted to know for sure if there are significant differences between all three groups.
* Total bird density shows similar trends in that the lowest density of birds is in the low-income group while the highest density is in the high-income group. These two groups are also the most differentiated from one another. In this case, medium income density looks very similar to low income but different from the high income group. however further testing would need to be conducted to know where there are significant differences between the three groups.
* It appears that there is support for the luxury effect hypothesis in this study based on the above observations.

1. How could the results from this model be used to guide management of urban forests and conservation of biodiversity? Consider the output from the model and the observations from the plots in your answer.

* The model indicates a significant, positive relationship between density of street trees and total bird density. The more street trees present along the walking route, the more birds that are present. It appears that these densities are not distributed equally across all walking routes as there are more trees, and therefore birds, among routes in higher income areas than lower. If one of our goals as city planners is to provide habitat for wildlife, then our efforts should be focused in lower income areas as these areas have the greatest need for them.

1. What are other ways that we can support conservation of biodiversity in (or around) urban spaces?

* Street trees are one great way to support biodiversity, but there are many others. Vegetation, in general, is essential for supporting biodiversity. Having more of it around whether it is in greenspaces such as parks, community gardens, or even in our own backyards provides numerous benefits to wildlife. The type of vegetation also matters as native species of trees and other plants provide more support to our native species than ornamentals. We can also support environmental agendas politically in our local governments encouraging equitable action to reduce carbon emissions, support renewable energy sources, increase green spaces, and more.

**Lesson Summary**

Key Takeaways

This lesson was made up of several interrelated goals.

* To promote a better understanding of the role that urban forests play in promoting biodiversity in urban areas.
* Explore how indirect factors like income can influence the distribution of trees, leading to unequal environmental benefits for wildlife and people.
* Why birds are used to determine the quality of an environment.
* How to use and interpret negative binomial regression in R to provide information for guiding future management of trees and to enhance conservation in urban ecosystems.

**Citations**

1. Cities4Forests. Urban Forests for Healthier Cities: Policy, Planning, Regulations, and Institutional Arrangements. 1–48 (2019).

2. Döring, M. Interpreting Generalized Linear Models. *Data Science Blog* https://www.datascienceblog.net/post/machine-learning/interpreting\_generalized\_linear\_models/ (2018).

3. Elmqvist, T., Zipperer, W. C. & Güneralp, B. Urbanization, habitat loss and biodiversity decline: Solution pathways to break the cycle. *Routledge Handb. Urban. Glob. Environ. Chang.* 139–151 (2015) doi:10.4324/9781315849256-20.

4. Ford, C. Getting started with Negative Binomial Regression Modeling. *University of Virginia Library* https://data.library.virginia.edu/getting-started-with-negative-binomial-regression-modeling/ (2016).

5. Fraixedas, S. *et al.* A state-of-the-art review on birds as indicators of biodiversity: Advances, challenges, and future directions. *Ecol. Indic.* **118**, 106728 (2020).

6. Leong, M., Dunn, R. R. & Trautwein, M. D. Biodiversity and socioeconomics in the city: a review of the luxury effect. *Biol. Lett.* **14**, (2018).

7. Lillis, D. Generalized Linear Models in R, Part 7: Checking for Overdispersion in Count Regression. *The Analysis Factor* <https://www.theanalysisfactor.com/glm-r-overdispersion-count-regression/>.

8. NEGATIVE BINOMIAL REGRESSION | R DATA ANALYSIS EXAMPLES. *UCLA: Advanced Research Computing, Statistical Methods, and Data Analytics* https://stats.oarc.ucla.edu/r/dae/negative-binomial-regression/ (2021).

9. Variance Ratio Meaning & Definition. *MBA Skool* https://www.mbaskool.com/business-concepts/statistics/7363-variance-ratio.html (2020).

10. Wood, E.M. & Esaian, S. Data from: The importance of street trees to urban avifauna, *Dryad*, Dataset, <https://doi.org/10.5061/dryad.qfttdz0d6> (2020).

11. Wood, E. M. & Esaian, S. The importance of street trees to urban avifauna. *Ecol. Appl.* **30**, 1–20 (2020).