

BIO252 Coursework assignment 2020 – Data Analysis Report

I briefly introduce the topic, detail the assigned task, and provide a detailed explanation of the marking scheme. Please refer also to Lecture 10 and the Handouts and detailed instructions provided during the Help & Feedback session.

The topic

Understanding how and why biodiversity varies over space is one of the main aims of ecological research. There are well documented geographical patterns of biodiversity change, most notably latitudinal and elevational gradients in biodiversity (Sanders & Rahbek 2012). Especially elevational gradients of biodiversity have attracted the interest of ecologists since the times of Darwin, Wallace and von Humboldt, as collecting data is more tractable, e.g. by locating transects along elevational gradients in mountains (Lomolino 2001). A combination of ecological and evolutionary processes are supposed to drive elevational clines in biodiversity, such as changes in temperature with elevation (Laiolo, Pato & Obeso 2018). Thanks to the increasing availability of data in online databases, there has been an increase in the last decades of studies quantifying and comparing elevational biodiversity gradients between geographic areas and taxonomic groups. The increase in use of online research data is a general trend in ecological research, hence the use and analysis of such data is an important transferrable skill.

The task

I would like you to select one of the two datasets provided for the assignment and use it to quantify and test if biodiversity changes across an elevational gradient and if this relationship changes between two biogeographic regions -- the Atlantic Forest and the Cerrado regions. For both datasets, provide a graphical and statistical analysis, choosing an appropriate statistical model for the latter. For both datasets then do a second analysis, to quantify and test if the elevational biodiversity gradient differs between taxonomic groups (Butterflies: subfamilies) or with the method used to collect the species data (birds dataset).

Once you have selected the data and question/hypothesis you prefer (refer also to the handout for the Help session), I would like you to briefly introduce the background (read the background papers provided), clearly formulate the questions and aims, then provide a full description of the Methods (Data and Statistical methods) and present the results, including figure(s) and Table(s). Then briefly discuss your conclusions, list the cited literature, as well as provide a brief overall summary/abstract. Finally, append the R code, following the scheme suggested in the R codes provided for preparing the data for analysis.

Marking scheme

The aim of the BIO252 module is to introduce the students to the basics of scientific data analysis using the R Software Environment, focusing especially on ecological data (broadly defined), to learn how to translate a biological question into the appropriate quantitative formulation and choose the most appropriate statistical

method to analyse the data. Understanding issues related to reproducibility of the methods and results and rigorous presentation of the findings (graphical and numerical information) are also key aims.

The key guiding principle is the ‘question-driven’ approach (*sensu* Fieberg & Börger 2012 J. Mammal. 93 (4), 890-902), based on first formulating relevant biological question(s), then translate into the appropriate quantitative formulation (statistical models and parameters), which then guides the choice of the most appropriate inferential approach to analyse the data. Only then are models fitted to data, the model fit evaluated and inferences derived. Equally important is to document the full workflow in a reproducible manner.

Consequently, the key marking criterion will be if the student has been able to apply the above approach, using one of the two datasets available on Canvas. To further enhance the transferrable skills learned with the assignment, the structure of the report to write for the assignment will match that of the dissertation project, or similar reports students will encounter in the current and future years, as well as during work life. The main focus will be on clearly defining the question(s)/aim(s) addressed in the Introduction, properly describing the methods used in the Methods section, and concisely and fully provide the results, comprising between 2-4 figures (maximum) and 1-2 Tables, complete with proper figure/table legends and figure annotation. Similarly important will be to include an appendix with the complete, and properly formatted and documented, R code. The focus is to produce a complete yet concise report of the analyses and results; hence, the Introduction and Discussion sections should be very brief and concise.

For the references, use the formatting required by the Journal ‘Methods in Ecology and Evolution’ – it is suggested to use a reference manager such as EndNote/Mendeley/Zotero to do so. If instead this will be done manually, the specifications are:

References in the text should be separated by a semi-colon. Papers with more than three authors should be abbreviated: (Manel et al. 1999). Work with the same first author and date should be coded by letters: (Thompson et al. 1991a,b). References should be listed in chronological order within the text. The references in the list should be in alphabetical order with the journal name in full. The format should be as follows:

-Begon, M., Harper, J.L. & Townsend, C.R. (1996) Ecology: Individuals, Populations and Communities, 3rd edn. Blackwell Science, Oxford.

-Tuytens, F.A.M. (1999) The consequences of social perturbation caused by badger removal for the control of bovine tuberculosis in cattle: a study of behaviour, population dynamics and epidemiology. PhD thesis, University of Oxford, Oxford.

-McArthur, W.M. (1993) History of landscape development. Reintegrating Fragmented Landscapes (eds R.J. Hobbs & D.A. Saunders), pp. 10-22. Springer Verlag, Berlin.

-Hill, M.O., Roy, D.B., Mountford, J.O. & Bunce, R.G.H. (2000) Extending Ellenberg's indicator values to a new area: an algorithmic approach. Journal of Applied Ecology, 37, 3-15.

-R Development Core Team (2009) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. URL <http://www.R-project.org> [accessed 18 November 2009]

(link available here: <http://www.methodsinecologyandevolution.org/view/0/authorGuidelines.html#Formatting> → see under ‘REFERENCES’)

Produce one single file (format as accepted by Turnitin), with a title page including your name and student ID (using the usual cover sheet), followed by a page with a brief Abstract (max 200 words), followed by the Main Body, with Tables and Figures inserted in the text and the references after the Discussion section, followed by an Appendix with the R codes. The final document should be submitted via Blackboard following the Turnitin submission link. Deadline is 4th January 2021.

References

- Laiolo, P., Pato, J. & Obeso, J.R. (2018) Ecological and evolutionary drivers of the elevational gradient of diversity. *Ecology Letters*, **21**, 1022-1032.
- Lomolino, M.V. (2001) Elevation gradients of species-density: historical and prospective views. *Global Ecology and Biogeography*, **10**, 3-13.
- Sanders, N.J. & Rahbek, C. (2012) The patterns and causes of elevational diversity gradients. *Ecography*, **35**, 1-3.

Feedback and Marks

General feedback

Excellent link between biological questions and analytical methods. Excellent execution and presentation of methods and analysis. Evidence of extra reading (ideally through additional knowledge and references). Very good basic scientific style and logic and no errors or repetition. Excellent use of technical terminology (>70 %).

☐

Good link between biological questions and analytical methods. Good execution and presentation of methods and analysis. Some evidence of extra reading (ideally through additional knowledge and references). Good basic scientific style and logic and no errors or repetition. Good use of technical terminology (60 - 70%).

☐

Fair link between biological questions and analytical methods. Fair execution and presentation of methods and analysis. No evidence of extra reading. Fair basic scientific style and logic, with some errors or repetition. Fair use of technical terminology (50 – 60%).

☐

Limited coverage and poor structure. Poor choice of analytical methods with little or incorrect justification. Poor or partial execution and presentation of analysis. No evidence of extra reading. Significant amount of basic errors in scientific style and logic (< 50 %).

☐

	Marks	
Summary - Does the abstract provide a concise and useful summary of the work, in relation to the title? - Max 200 words. - Are the question/hypotheses, data, methods and results clearly summarised? - Is the take home message clear? - Standard of scientific English and general presentation.	5	
Introduction - Is the introduction adequate and relevant to the topic of the title? - Does it make appropriate use of the literature? Provide at least 2-3 references. - Are the questions/ hypotheses of the work clear and properly formulated? - Standard of scientific English and general presentation.	5	
Methods - Complete documentation of the data. - Complete description of the procedures/methods used, including criteria for selecting/subsetting/manipulating data, so that other researchers would be able to replicate the study. - Complete documentation of the statistical methods, including the software and versions used (incl. R and additional R libraries). - Correct choice of the statistical methods. (IMPORTANT!) - Standard of scientific English and general presentation.	35	
Results - Results for all questions/hypotheses are written completely, but concisely. All relevant quantitative information underpinning the results is provided, supported by appropriate statistics. - The Results provided fully and correctly answer the aims and questions, using correct statistical methods. - Use of Table and Figures to complement the written part of the Results. Do not repeat the same information in Tables in Figures. Each table and figure should be self-explanatory, with appropriate legends. - Standard of scientific English and general presentation.	30	
Discussion - Critically evaluates the strength of the results and adequacy of the data and methods employed. - Briefly discusses the broader implications/relevance of the study and may outline avenues for future studies.	5	
References Consistently and properly formatted. Do follow the requirements of the Journal 'Methods in Ecology and Evolution'.	5	
Appendix/R code - The script is well documented. - The code runs without errors and is complete. - The script is well organized and legible, with consistent use of indentation, spacing, etc. (e.g.: https://google.github.io/styleguide/Rguide.xml).	15	
Total	100	

