

## EDITORIAL

## Ten simple rules for getting started with statistics in graduate school

Rachel A. Zitomer<sup>1</sup>, Jessica Karr<sup>2</sup>, Mark Kerstens<sup>3</sup>, Lindsey Perry<sup>4</sup>, Kayla Ruth<sup>4</sup>, Lindsay Adrean<sup>3</sup>, Suzanne Austin<sup>2</sup>, Jamie Cornelius<sup>2</sup>, Jonathan Dachenhaus<sup>3</sup>, Jonathan Dinkins<sup>4</sup>, Alan Harrington<sup>4</sup>, Hankyu Kim<sup>1</sup>, Terrah Owens<sup>4</sup>, Claire Revekant<sup>4</sup>, Vanessa Schroeder<sup>4</sup>, Chelsea Sink<sup>5</sup>, Jonathon J. Valente<sup>3,6</sup>, Ethan Woodis<sup>3</sup>, James W. Rivers<sup>3\*</sup>

**1** Department of Forest Ecosystems and Society, Oregon State University, Corvallis, Oregon, United States of America, **2** Department of Integrative Biology, Oregon State University, Corvallis, Oregon, United States of America, **3** Department of Forest Engineering, Resources, and Management, Oregon State University, Corvallis, Oregon, United States of America, **4** Department of Animal and Rangeland Sciences, Oregon State University, Corvallis, Oregon, United States of America, **5** Department of Fisheries, Wildlife, and Conservation Sciences, Oregon State University, Corvallis, Oregon, United States of America, **6** Smithsonian Conservation Biology Institute, Migratory Bird Center, National Zoological Park, Washington, DC, United States of America

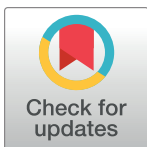
✉ Current address: American Bird Conservancy, Veneta, Oregon, United States of America

\* [jim.rivers@oregonstate.edu](mailto:jim.rivers@oregonstate.edu)

## Introduction

Graduate school is often a time of enormous professional growth, and for many students, it is the first time they receive advanced training in statistics. Despite the foundational importance of statistics in many fields, learning proper statistical approaches can be challenging for beginning graduate students, especially for those that lack prior experience in this subject area. This can occur for several reasons but is often related to competing interests for students' limited time at the start of a graduate program. New graduate students are typically working with a new advisor, taking on new research, and/or helping to instruct a class for the first time, in addition to taking advanced courses, applying for grants and fellowships, and settling into their new professional setting. Graduate school also marks a transition to self-directed learning that may diverge strongly from undergraduate degree programs, which can also bring new challenges. In short, the start of graduate school can be a demanding time for students with much to learn in a relatively short time period.

New graduate students often begin their program by taking an overview course in statistics such as basic statistical methods, experimental design, or linear modeling. However, in the collective experience of our group—which includes graduate students, research assistants, post-doctoral researchers, and university professors in the field of natural resources—we have found that learning to apply statistical analyses correctly requires course-based learning about specific analysis methods, as well as a broader understanding of the philosophy of applying statistical approaches in research. We have found that the latter topic, on which we focus here, often gets less attention in graduate courses because it does not fit neatly into any single course. Nevertheless, topics related to this theme regularly arise throughout graduate school and beyond and thus highlight its importance for learning how to design, undertake, and complete high-quality research. Although there have been several excellent reviews in the 10 simple rules series on topics that are linked to the broader points we raise here, such as guidelines for using statistics [1], managing data [2], and learning to program [3], none have been focused on a discussion of broad conceptual topics that relate to applying statistical approaches in



## OPEN ACCESS

**Citation:** Zitomer RA, Karr J, Kerstens M, Perry L, Ruth K, Adrean L, et al. (2022) Ten simple rules for getting started with statistics in graduate school. *PLoS Comput Biol* 18(4): e1010033. <https://doi.org/10.1371/journal.pcbi.1010033>

**Published:** April 21, 2022

**Copyright:** © 2022 Zitomer et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The authors received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

research with new graduate students as a primary audience. Therefore, we offer what we view as 10 important rules for new students to reflect on as they begin to build their statistical skill-sets in graduate school. It is important to note that some of the rules we offer cannot be met in all circumstances; thus, we view them as rules to which graduate students should aspire, recognizing that, at times, they may be bent in the name of pragmatism. Finally, although the primary audience for the rules we provide are new graduate students, we note that the principles we raise are broadly applicable to researchers in a wide range of disciplines. Indeed, we posit that some advisors of new graduate students might themselves benefit from a periodic review of these rules!

### **Rule 1: Start with a research question before data collection**

New graduate students often feel pressure to get started with their research as soon as possible and that includes collecting data toward a thesis or dissertation. Despite this, it is critical to start the process by formulating a research question well before data collection begins [4]. Extensive reading of the primary literature, as well as recent review papers and edited volumes, is key to learning about the scope of work that has—and has not—been conducted in a particular field, which helps to hone and focus research questions [4]. Although it may seem more efficient to collect data while developing research questions, doing so is suboptimal and should be avoided for several reasons. First, it reduces the amount of time available to undertake a literature review expansive enough to identify the key knowledge gap(s) that could be the focus of a graduate project. Second, collecting data prior to finalizing a research question makes it more likely that key variables that need to be measured remain unidentified, at least at the start. Finally, collecting and analyzing data while developing hypotheses can lead to hypothesizing after the results are known, also known as HARKing [5]. Although it may be acceptable to revise the original hypotheses prior to data analysis, it is problematic if hypotheses are changed after data analysis is conducted and then presented as if they were developed prior to data collection [5]. Importantly, HARKing goes against the hypothetico-deductive premise of conducting science [6] in which a researcher generates a falsifiable hypothesis and then collects data to test it because, by definition, a hypothesis that has been created based on the results of an analysis cannot be falsified by the same analysis. Additional pitfalls of HARKing include a reduction in the search for alternative explanations for phenomena and strong potential for hindsight confirmation bias [5]. These issues can limit replicability and should be avoided by ensuring that a priori hypotheses are tested and that any speculation that is undertaken after results are known is clearly stated as such. Preregistration of research, wherein a research plan is deposited in a public repository before a project is undertaken [7], is a useful way of ensuring that hypotheses and analytical approaches are determined prior to data analysis, rather than in response to it. This practice increases transparency in the research process and is becoming increasingly common in many scientific fields [8].

It is worth noting that some graduate students may work on a research topic that has already been selected by their research advisor, join a project that is already underway, and/or receive a dataset that has already been collected. In such cases, they may not be starting from “square one” in the research process. Nevertheless, it is still important to have a thorough understanding of the background literature and the research question(s) that motivated data collection. Thus, rather than assuming that such an “inherited” study lacks room for improvement, we recommend that students critically examine the choices made in data collection, understand their strengths and weaknesses, and independently assess whether the resulting data are sufficient to address the research question(s) being examined.

## Rule 2: Understand how manipulative experiments differ from descriptive studies

Studies come in many forms, with descriptive studies and manipulative experiments being among the most common. Of these, the randomized manipulative experiment is the gold standard in terms of inference [9], and in its simplest form, it is a representative group of experimental units (e.g., individuals, study sites) from a population that is randomly assigned to either a treatment or a control group. Units in the treatment group are subjected to an experimental manipulation, whereas those in the control group are not, and the average response across experimental units in the 2 groups is then compared. Because other factors are held constant, differences detected in the treatment group relative to the control group can be attributed to the experimental manipulation. In contrast, observational studies are those that describe relationships as they occur, without manipulation by a scientist, who in this case serves as a passive observer and not as a manipulator of the system under study. The differences between these 2 types of studies may seem subtle, but they have enormous importance for inference, and the strength of conclusions that can be drawn. Most importantly, randomized experiments can assign causation to the factor that is manipulated by a researcher, whereas descriptive studies are limited to providing correlative evidence of relationships [10]. It is critical to note that despite these differences, descriptive studies play an important role in generating knowledge; they are, in fact, the foundation for much of science as many systems do not allow for manipulation, especially across large scales [11]. However, their limitations must be understood relative to manipulative experiments so that proper inference is made (see Rule 3).

## Rule 3: Understand the limits of inference

Inference in research involves drawing conclusions about a population based on observations of a sample of that population using inductive reasoning. The validity of inductive reasoning is dependent on the premise that the sample is representative of the population of interest [12]. The set of situations to which the conclusions of a study can be generalized is known as its scope of inference [13]. The ideal for most researchers is to apply the results of research as broadly as possible. However, the reasonable scope of inference for a research project is typically constrained by the realities of data collection. Thus, defining the scope of inference for a study involves thinking carefully about the chosen sampling approach, potential sources of bias, and acknowledging how they limit the generalizability of a study. Likewise, adopting an approach that is too narrow when defining the scope of inference for a study can also be limiting. Some graduate students may focus so intently on the specifics of their study that they fail to consider the larger range of conditions that it might represent. In some cases, the scope of inference for a study may be broadened considerably if there is evidence that the conditions in which sampling occur are typical of a larger geographic area or timeframe. In other cases, studies with a relatively narrow scope of inference might fill a gap in the existing body of research on their subject that allows more generalizable conclusions to be drawn. Thus, defining the scope of inference for a study requires a striking a balance where the scope is narrow enough that findings are appropriately applicable, yet broad enough to provide knowledge that extends beyond a single study system. Getting feedback from colleagues and peers about what is a reasonable scope of inference for a particular study is often a useful exercise (see Rule 10).

## Rule 4: Start learning a statistical programming language early

Although some statistical programs are “canned” and do not require extensive knowledge of programming to perform statistical analysis, our experience has found that graduate students

who learn a programming language early in their program become more proficient with data analysis throughout graduate school and beyond. Although it does take time, learning a programming language helps to solidify statistical concepts because it requires creating all of the steps needed to undertake an analysis, in marked contrast to statistical programs operated through a graphical user interface. In addition, learning how to program in a commonly used language increases efficiency and accessibility of analyses [14], especially when well-annotated scripts are created that can save time when code is reused in subsequent analyses. Using a statistical programming language can also increase reproducibility as scripts not only enable data manipulation and statistical testing but also record the nature and sequence of these steps [15]. Archiving all versions of scripts and datasets—including raw data—is a straightforward and efficient way to track changes made throughout the analysis process [16]. It is important to note that programming scripts should not be the only mechanism for ensuring reproducibility [16], but they are particularly useful for keeping a record of each step in the analysis so that important details are not lost.

We suggest that new students consult literature and talk with others in their field to find out which programming languages are commonly used and to understand their strengths and weaknesses with different types of data or analytical approaches before choosing the language(s) on which to focus [3]. For example, one programming language that is free and used widely in the natural resource sciences [17] is the R statistical environment (<https://www.r-project.org>). Because it is an open-source language, there are many packages that are created for specific analyses and freely available to researchers. There are also extensive tutorials and a rich online public support community that can provide relevant code and troubleshoot errors that may arise. Regardless of whether one uses R or another programming language (e.g., SAS, Python), the work that new graduate students put into learning programming early in graduate school will pay dividends later as they move beyond graduate school and into the next phase of their career.

### **Rule 5: Set up databases in a tidy format before data collection**

Graduate students vary in how they obtain the data they use for their graduate research. Although some students are fortunate to receive a well-curated database in its final form, most end up spending a large amount of time wrangling their data. Indeed, it is estimated that up to 80% of data analysis time is spent cleaning and preparing data prior to conducting any statistical analyses [18]. Although different types of data analyses often require different data formats, there are some general rules to keep in mind when developing a database, and they should be implemented prior to data collection. First, standardize data entry through the establishment of metadata that, at a minimum, describes the actual names of the variables found within a database (e.g., “body\_mass”), provides a definition for each variable (e.g., “body mass measurement taken with an electronic scale to  $\pm 0.01$  g”), and reports possible entries for each variable (e.g., “positive continuous values”). Including this information is critical so that anyone in possession of the data can clearly understand individual variables and how they were defined. Additionally, structure databases so that each variable forms a column, each observation forms a row, and each type of observational unit forms a table [19]. This structure—referred to as “tidy data”—is an ideal format for storing data that facilitates statistical analyses. Wickham [19] extolls the many virtues of this approach, and his paper on tidy data is a must-read for any new graduate student. Storing data in a tidy format also make it easier to reproduce results, when returning to previously used datasets and when sharing data with other researchers, both of which are critical for building scientific knowledge [1].

### Rule 6: Understand the form of data

Once data are in hand (note that data is a plural term and datum is singular), whether they are collected directly or obtained indirectly, it is essential to have a good understanding of those data and how they should be analyzed. Introductory statistics courses typically focus on general linear models (e.g.,  $t$  tests, linear regression, and analysis of variance) that assume that measured responses are normally and independently distributed around the modeled mean. The normality assumption in particular can be a source of great anxiety for many new graduate students when they discover their residuals do not follow such a form, often leading to arcane data transformations or even abandonment of certain data that have been collected. Thus, it is critical to recognize that techniques are now widely available for analyzing data with a variety of forms, including those that are binary (e.g., logistic regression and Fisher exact test) or are based on counts (e.g., Poisson and negative binomial regression) or proportions (e.g., Dirichlet regression). Of course, each of these approaches comes with its own assumptions, and, therefore, it is critical to ensure that they are met regardless of the method chosen. In some cases, nonparametric approaches, which tend to have fewer assumptions but also have lower statistical power, may be reasonable alternatives. Regardless, being able to understand the form of the data one has and identify an appropriate underlying distribution when evaluating statistical hypotheses with those data are critically important skills for new graduate students given the many analytical options available today.

In many sampling situations, the assumption of independence that underlies general linear models, as well as many other analytical techniques, may also be violated. Sometimes, nonindependence is purposeful; for example, a researcher may be interested in how herbicide exposure affects salamander body weight over time. In this case, multiple body weight measurements on the same salamander are not independent of one another, yet they are required to answer the research question. In other cases, nonindependence can be a side effect of sampling constraints, such as when samples collected closer together in space or time are more likely to have similar values due to environmental homogeneity at local scales. Several techniques have been developed to account for these sources of nonindependence, including mixed model analysis [20] and diverse methods for handling spatial and temporal autocorrelation [21,22]. New students should be aware that nonindependence in data must be recognized and accounted for, and failing to do so can lead to pseudoreplication and other issues related to data interpretation [23].

### Rule 7: Understand what a $p$ -value is and what it is not

$P$ -values have been frequently overemphasized and misinterpreted in the scientific literature [24]. Technically defined, a  $p$ -value is the probability of obtaining an observed dataset if the null statistical hypothesis were true. Typically, a null hypothesis postulates the absence of an effect; that is, that there is no difference between groups being compared or no association between variables being evaluated [25]. A very small  $p$ -value therefore suggests that it is unlikely one would have obtained the observed results if there truly was no effect.  $p$ -Values are traditionally compared to a threshold value, alpha ( $\alpha$ ), which is defined a priori and serves to delineate the acceptable probability of mistakenly rejecting the null hypothesis when it is true [26]. This approach frequently results in interpretation of an observed effect as “significant” or “nonsignificant” if the  $p$ -value is smaller or larger than  $\alpha$ , respectively [27].

It is critically important for new graduate students to recognize that a  $p$ -value is a function of the magnitude of an effect, the variability in a response, and the sample size of a dataset. Because of this, even trivial effects may be deemed statistically “significant” when sample sizes are sufficiently large [25]. Conversely, strong effects may not be reflected in the outcome of

statistical tests when sample sizes are small or when variability in a response is large. For example, 2 studies could estimate similar effect sizes yet differ in the precision of their estimates—for example, if one study has a small error estimate and the error estimate of the other study is large—leading one to conclude the first study detected a statistically “significant” effect, whereas the second study was found to have no effect at all [26]. In reality, there is no conflict between these 2 hypothetical studies regarding the observed effect, yet adopting the use of statistical “significance” when describing their findings suggests that such a conflict exists, which can lead to erroneous conclusions [26]. Therefore, new graduate students should take care to appreciate that  $p$ -values should not be taken as an indicator of whether the effect of a particular variable is strong or has importance to the system under study. Instead,  $p$ -values are merely a means of quantifying how certain one can be that there is no effect, given the type and quantity of data that have been collected. The American Statistical Association’s statement on  $p$ -values provides excellent guidance on the use and interpretation of  $p$ -values and clarification of common mistakes and is recommended reading for all new graduate students [27].

### **Rule 8: Learn how statistical power can influence results**

In any study, there are constraints on the amount of data that can be collected. Therefore, it is important to understand the role of statistical power and sample size in influencing results, as well as the scope of inference [28] (see Rule 3). Statistical power is the probability of correctly rejecting a statistical null hypothesis that is false, and this probability increases as both sample size and effect size increase. Thus, it may be possible to detect differences between 2 populations even with small sample sizes if effect sizes are large, whereas large sample sizes might lead to detecting very small statistical differences [29]. Given these considerations, it is essential to turn a critical eye when examining results to the extent to which methodological effects (e.g., sample size) may be driving the results relative to the strength of the response in the system (i.e., effect size). Conversely, when sample sizes are small and effect sizes are either small or data are highly variable, statistical power to detect differences is low [25]. In such instances, it is important to recognize the old adage that “the absence of evidence is not evidence of absence” and that an effect may be present even though low statistical power does not allow researchers to detect it within a study. Furthermore, reporting such negative results should also reflect this possibility; for example, it is preferable to state that no differences were detected in a study rather than stating that no differences were present. Although retrospective power analysis is generally viewed unfavorably, power analysis may, in some instances, be useful when applied a priori to determine what sample size is necessary to detect the minimum effect size of interest given a specified expected variance [28].

### **Rule 9: Appreciate the importance of effect size**

As noted above in Rule 7, statistical null hypotheses posit that there is no difference between groups being compared or no association between variables being studied [25]. In practice, however, there is almost always some expected difference in measures taken between any 2 populations, however small, simply due to sampling effects and random variability [30]. Thus, the purpose of performing statistical analysis is not usually to determine whether 2 groups are absolutely the same, but rather whether the difference between them (i.e., the effect size) is meaningful within the context of the study system.

The effect size is the magnitude of a relationship between variables or a difference among groups [31], and it is typically presented with confidence intervals that represent the degree of uncertainty around the estimate [24]. In most cases, estimates of effect size are more informative than  $p$ -values because they provide context for the magnitude and direction of an effect



and therefore should be emphasized when evaluating and reporting results [30]. For example, stating that a species of interest was “ $3.2 \times$  (95% CI: [1.3, 4.6]) more likely to use a treatment site than a control site” is more informative than stating that “there was a significant effect ( $p < 0.01$ ) of treatment on site use.” The former approach highlights the magnitude of the effect, which, in turn, may help resource managers evaluate whether the biological response is worth the cost of implementing similar treatments in the future. In other words, interpreting the results in the context of the system under study is what really matters, without undue focus on statistical outcomes.

### Rule 10: Don't fly solo

If you are a new graduate student who feels a bit overwhelmed by the previous rules outlined here, this last rule is especially for you. Although it may seem like you are embarking on a singular odyssey by yourself, graduate school is really a time to develop connections with new peers, colleagues, and mentors. As such, it is important to develop and make use of a new professional support network throughout graduate school, including in the exploration of statistical analysis techniques and theory. We have found that statistics can be especially challenging for new graduate students and that learning new techniques and programming languages is best facilitated by collaboration and conversation with others. In addition to the benefits that derive from working with other students, there are often other outlets that can provide help when learning statistics. For example, some universities have statistical consultants that provide tailored one-on-one help that can provide input and assistance on myriad topics. With regard to working with consultants, one oft-repeated piece of advice is to seek help from statisticians before data collection begins, as that is the phase of the research process in which they can best help to remedy problems that may be present in a study; when the time has come for data analysis, it is often too late. More broadly, there is a wealth of online tutorials on statistical methods and programming languages, as well as free online courses, which combine statistics and programming and extensive online help communities for many programming languages. In short, make sure you tap into the resources that are available so that when—not if—you get stuck you will have a support system that can help get you back on track.

### Concluding thoughts

Graduate school is an exciting time, but it can also be stressful due to its finite nature. In many fields, developing an understanding of statistics is a key part of the professional growth of graduate students, and it takes many forms: classes, workshops, self-study, and more. Given its central importance, and the expanding body of statistical science, our collective experience has found that researchers must continue to learn about statistical methods throughout their career, and, thus, developing a solid base early in one's graduate program is essential for continued growth and development. It is therefore our expectation that these rules, when taken together, will help make that process a little less daunting for new students as they develop a deeper understanding of the statistical methods they will use during graduate school and throughout the rest of their careers.

### References

1. Kass RE, Caffo BS, Davidian M, Meng XL, Yu B, Reid N. Ten Simple Rules for Effective Statistical Practice. *PLoS Comput Biol*. 2016; 12(6):1–8. <https://doi.org/10.1371/journal.pcbi.1004961> PMID: 27281180

2. Hart EM, Barmby P, LeBauer D, Michonneau F, Mount S, Mulrooney P, et al. Ten Simple Rules for Digital Data Storage. *PLoS Comput Biol*. 2016 Oct 1; 12(10):e1005097. <https://doi.org/10.1371/journal.pcbi.1005097> PMID: 27764088
3. Carey MA, Papin JA. Ten simple rules for biologists learning to program. *PLoS Comput Biol*. 2018 Jan 1; 14(1):e1005871. <https://doi.org/10.1371/journal.pcbi.1005871> PMID: 29300745
4. John FA V, Keane AM, Jones JPG, Milner-Gulland EJ FORUM: Robust study design is as important on the social as it is on the ecological side of applied ecological research. *J Appl Ecol*. 2014 Dec 1; 51(6):1479–85. <https://doi.org/10.1111/1365-2664.12352>
5. Kerr NL. HARKing: hypothesizing after the results are known. *Pers Soc Psychol Rev*. 1998 Dec 21; 2(3):196–217. [https://doi.org/10.1207/s15327957pspr0203\\_4](https://doi.org/10.1207/s15327957pspr0203_4) PMID: 15647155
6. Popper KR. *The Logic of Scientific Discovery*. New York, NY: Hutchinson & Co.; 1959. <https://doi.org/10.17485/ijst/2016/v9i10/88981>
7. Ioannidis JPA. Pre-registration of mathematical models. *Math Biosci*. 2022 Mar 1; 345:108782. <https://doi.org/10.1016/j.mbs.2022.108782> PMID: 35090877
8. Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. The preregistration revolution. *Proc Natl Acad Sci U S A*. 2018 Mar 13; 115(11):2600–6. <https://doi.org/10.1073/pnas.1708274114> PMID: 29531091
9. Diamond JM. Overview: Laboratory Experiments, Field Experiments, and Natural Experiments. In: Diamond JM, Case TJ, editors. *Community Ecology*. New York: Harper & Row; 1986. p. 3–22.
10. Cunningham RB, Lindenmayer DB. Approaches to Landscape Scale Inference and Study Design. 2016. <https://doi.org/10.1007/s40823-016-0019-4>
11. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ*. 1996 May 11; 312(7040):1215–8. <https://doi.org/10.1136/bmj.312.7040.1215> PMID: 8634569
12. Randomization Greenland S., Statistics, and Causal Inference. *Epidemiology*. 1990; 1(6):421–9. <https://doi.org/10.1097/00001648-199011000-00003> PMID: 2090279
13. Ganio LM, Puettmann KJ. Design Challenges in Large-Scale Management Experiments. In: Peterson CE, Maguire DA, editors. *Balancing Ecosystem Values: Innovative Experiments for Sustainable Forestry*. Portland, OR: U.S. Department of Agriculture; 2004. p. 35–9. <https://doi.org/10.1007/s10151-004-0093-1> PMID: 15654535
14. Bodner K, Brimacombe C, Chenery ES, Greiner A, McLeod AM, Penk SR, et al. Ten simple rules for tackling your first mathematical models: A guide for graduate students by graduate students. *PLoS Comput Biol*. 2021 Jan 14; 17(1):e1008539. <https://doi.org/10.1371/journal.pcbi.1008539> PMID: 33444343
15. Sandve GK, Nekrutenko A, Taylor J, Hovig E. Ten Simple Rules for Reproducible Computational Research. *PLoS Comput Biol*. 2013 Oct; 9(10):e1003285. <https://doi.org/10.1371/journal.pcbi.1003285> PMID: 24204232
16. Ten SS. Simple Rules for a Computational Biologist's Laboratory Notebook. *PLoS Comput Biol*. 2015; 11(9):e1004385. <https://doi.org/10.1371/journal.pcbi.1004385> PMID: 26356732
17. Lai J, Lortie CJ, Muenchen RA, Yang J, Ma K. Evaluating the popularity of R in ecology. *Ecosphere*. 2019 Jan 1; 10(1):e02567. <https://doi.org/10.1002/ECS2.2567>
18. Dasu T, Johnson T. *Exploratory data mining and data cleaning*. Vol. 479. John Wiley & Sons; 2003.
19. Wickham H. Tidy Data. *J Stat Softw*. 2014 Sep 12; 59(10):1–23. <https://doi.org/10.18637/JSS.V059.I10>
20. Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, et al. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol Evol*. 2009 Mar 1; 24(3):127–35. <https://doi.org/10.1016/j.tree.2008.10.008> PMID: 19185386
21. Dormann CF, McPherson JM, Araújo MB, Bivand R, Bolliger J, Carl G, et al. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography (Cop)*. 2007 Oct 1; 30(5):609–28. <https://doi.org/10.1111/J.2007.0906-7590.05171.X>
22. Betts MG, Ganio LM, Huso MMP, Som NA, Huettmann F, Bowman J, et al. Comment on “Methods to account for spatial autocorrelation in the analysis of species distributional data: a review.” *Ecography (Cop)*. 2009 Jun 1; 32(3):374–8. <https://doi.org/10.1111/J.1600-0587.2008.05562.X>
23. Hurlbert SH. Pseudoreplication and the Design of Ecological Field Experiments. *Ecol Monogr*. 1984 Jun 1; 54(2):187–211. <https://doi.org/10.2307/1942661>
24. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol*. 2016 May 21; 31(4):337–50. <https://doi.org/10.1007/s10654-016-0149-3> PMID: 27209009
25. Johnson DH. The Insignificance of Statistical Significance Testing. *J Wildl Manag*. 1999; 63(3):763–72.
26. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature*. 2019 Mar 20; 567(7748):305–7. <https://doi.org/10.1038/d41586-019-00857-9> PMID: 30894741



27. Wasserstein RL, Lazar NA. The ASA Statement on p-Values: Context, Process, and Purpose. *Am Stat*. 2016 Apr 2; 70(2):129–33. <https://doi.org/10.1080/00031305.2016.1154108>
28. Steidl RJ, Hayes JP, Schaubert E. Statistical Power Analysis in Wildlife Research. *J Wildl Manag*. 1997 Apr; 61(2):270–9.
29. Peterman RM. Statistical Power Analysis can Improve Fisheries Research and Management. *Can J Fish Aquat Sci*. 1990; 47(1):2–15. <https://doi.org/10.1139/F90-001>
30. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biol Rev*. 2007; 82(4):591–605. <https://doi.org/10.1111/j.1469-185X.2007.00027.x> PMID: [17944619](https://pubmed.ncbi.nlm.nih.gov/17944619/)
31. Sullivan GM, Feinn R. Using Effect Size—or Why the P Value Is Not Enough. *J Grad Med Educ*. 2012 Sep 1; 4(3):279. <https://doi.org/10.4300/JGME-D-12-00156.1> PMID: [23997866](https://pubmed.ncbi.nlm.nih.gov/23997866/)