

Rémi Thériault

From: yesica.yang@mdpi.com on behalf of Mathematics Editorial Office
<mathematics@mdpi.com>
Sent: Wednesday, March 22, 2023 9:59 PM
To: Rémi Thériault
Cc: Rémi Thériault; Mattan S. Ben-Shachar; Indrajeet Patil; Daniel Lüdecke; Brenton M. Wiernik; Dominique Makowski; Mathematics Editorial Office; Jeannie Liu
Subject: [Mathematics] Manuscript ID: mathematics-2265149 - Declined for Publication - Encourage Resubmission after Revisions

Dear Mr. Thériault,

I am writing to you concerning the manuscript you recently submitted to Mathematics. Based on the review reports, the manuscript is not suitable for publication in Mathematics in its present format. We encourage the resubmission of your manuscript after extensive revisions.

Significant revisions or new data are required before the manuscript can be considered for publication in Mathematics.

Manuscript ID: mathematics-2265149

Type of manuscript: Article

Title: Check your outliers! An accessible introduction to identifying statistical outliers in R with easystats

Authors: Rémi Thériault *, Mattan S. Ben-Shachar, Indrajeet Patil, Daniel Lüdecke, Brenton M. Wiernik, Dominique Makowski

Received: 21 February 2023

E-mails: theriault.remi@courrier.uqam.ca, mattansb@msbstats.info, patilindrajeet.science@gmail.com, d.luedecke@uke.de, brenton@wiernik.org, dom.makowski@gmail.com Submitted to section: Probability and Statistics, https://www.mdpi.com/journal/mathematics/sections/probability_and_statistics_theory

Advances in Statistical Computing

https://www.mdpi.com/journal/mathematics/special_issues/Advances_in_Statistical_Computing

https://susy.mdpi.com/user/manuscripts/review_info/33cbf34f213213b17592675ce62aec55

You can find the review reports at:

<https://susy.mdpi.com/user/manuscripts/resubmit/33cbf34f213213b17592675ce62aec55>

During resubmission, you must clearly indicate the ID (mathematics-2265149) of this manuscript. All changes must be highlighted and a cover letter with responses to reviewers' comments included. Note that the Editorial Office may send the paper to the same reviewers or invite new reviewers.

Please resubmit your revised manuscript through the following link:

https://susy.mdpi.com/user/manuscripts/upload?pre_hash_key=33cbf34f213213b17592675ce62aec55

Thanks again for submitting your work to Mathematics. If you have any questions, please contact the Editorial Office at mathematics@mdpi.com.

Kind regards,
Mathematics Editorial Office
mathematics@mdpi.com

[User Menu](#)
[Manuscript Information Overview](#)

Home		
(/user/myprofile)		
Manage	Manuscript ID	mathematics-2265149
Accounts	Status	Rejected
(/user/manage_accounts)	Article type	Article
Change	Title	Check your outliers! An accessible introduction to identifying statistical outliers in R with easystats
Password		
(/user/chgpwd)	Journal	<i>Mathematics</i> (https://www.mdpi.com/journal/mathematics)
Edit Profile	Section	Probability and Statistics
(/user/edit)		(https://www.mdpi.com/journal/mathematics/sections/probability_and_statistics_theory)
Logout		
(/user/logout)	Special Issue	Advances in Statistical Computing
		(https://www.mdpi.com/journal/mathematics/special_issues/Advances_in_Statistical_Computing)
Submissions		

Submissions
Menu

Submit		practices regarding the diagnosis and treatment of outliers, an additional difficulty arises concerning the mathematical implementation of the recommended methods. In this paper, we provide an overview of current recommendations and best practices and demonstrate how they can easily and conveniently be implemented in the R statistical computing software, using the <code>*{performance}* package of the <code>*easystats* ecosystem. We cover univariate, multivariate, and model-based statistical outlier detection methods, their recommended threshold, standard output, and plotting methods. We conclude with recommendations on the handling of outliers: the different theoretical types of outliers, whether to exclude or winsorize them, and the importance of transparency.</code></code>
Manuscript (/user/manuscripts/upload)		
Display		
Submitted		
Manuscripts (/user/manuscripts/status)		
Display Co-		
Authored		
Manuscripts (/user/manuscripts/co-authored)	Keywords	univariate outliers; multivariate outliers; robust detection methods; R; easystats
English Editing (/user/pre_english_article/status)	Manuscript	manuscript.zip
Discount	File	(/user/manuscripts/displayFile/33cbf34f213213b17592675ce62aec55)
Vouchers (/user/discount_voucher)	PDF File	manuscript.pdf
Invoices (/user/invoices)		(/user/manuscripts/displayFile/33cbf34f213213b17592675ce62aec55/latest_pdf)



Preprints.org

You can put your paper online **immediately and before peer review** at Preprints.org (<https://www.preprints.org>), with the following benefits:

- Anyone can read and download your work immediately, before peer review is complete.
- Receive comments and feedback.
- Make your work citable via assignment of a digital object identifier.
- Immediate indexing by Google Scholar and other online databases.
- Papers are put online within 24 hours.
- A doi will be applied to your announced preprints automatically.

Upload to Preprints.org



data

Data is of paramount importance to scientific progress, yet most research data drowns in supplementary files or remains private. Enhancing the transparency of the data processes will help to render scientific research results reproducible and thus more accountable. Co-submit your methodical data processing articles or data descriptors for a linked data set in *Data* (<https://www.mdpi.com/journal/data>) journal to make your data more citable and reliable.

- Deposit your data set in an online repository, obtain the DOI number or link to the deposited data set.
- Download and use the Microsoft Word template (<https://www.mdpi.com/files/word-templates/data-template.dot>) or LaTeX template (<https://www.mdpi.com/authors/latex>) to prepare your data article.
- Upload and send your data article to the *Data* (<https://www.mdpi.com/journal/data>) journal here (/user/manuscripts/upload?form%5Bjournal_id%5D=176&form%5Barticle_type_id%5D=47).

Submit To Data (/user/manuscripts/upload?form%5Bjournal_id%5D=176&form%5Barticle_type_id%5D=47)

Author Information

Submitting Author Rémi Thériault

Corresponding Author Rémi Thériault

Author #1 Rémi Thériault

E-Mail theriault.remi@courrier.uqam.ca



Author #2	Mattan S. Ben-Shachar
E-Mail	mattansb@msbstats.info
Author #3	Indrajeet Patil
E-Mail	patilindrajeet.science@gmail.com
Author #4	Daniel Lüdecke
E-Mail	d.luedecke@uke.de
Author #5	Brenton M. Wiernik
E-Mail	brenton@wiernik.org
Author #6	Dominique Makowski
E-Mail	dom.makowski@gmail.com

Manuscript Information

Received Date	21 February 2023
Submission to First Decision (Days)	29
Submission to Publication (Days)	
Word Count	2098
Page Count	11

Editor Decision

Decision	Reject and encourage resubmission
Comments	Less convinced as a scientific research is the major concern. I cannot find a solid contribution for the statistical methodologies, new packages, or sound applications in this study. Almost all the obtained conclusions are known. The readers would be happy to see more convincing results from good real examples from the real world if only the well known statistical methods and packages are used in this study. The tutorial-like results using the examples in R are too weak for an academic paper. It is contributive if the authors can include real examples in different areas to show the power of the used package and present the insight findings, not just show how to use the packages. Then, resubmit the revision as a new paper for review.
Decision Date	22 March 2023



B

I

x^2

x_2

↺

Ω

🔗

A

▼

✎

▼

⌵

⌶

$\frac{1}{2}$

▼

:

▼

≡

≡

≡

≡

I_x

<>

P

0 WORDS

↗

Word / PDF
 Choose file
 No file chosen

Submit or Save as draft (submit later)

Review Report Form

Open Review

☒ I would not like to sign my review report

☐ I would like to sign my review report

Quality of English Language

☐ English very difficult to understand/incomprehensible

☐ Extensive editing of English language and style required

☐ Moderate English changes required

☐ English language and style are fine/minor spell check required

☒ I am not qualified to assess the quality of English in this paper

	Yes	Can be improved	Must be improved	Not applicable
Does the introduction provide sufficient background and include all relevant references?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Are all the cited references relevant to the research?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Is the research design appropriate?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Are the methods adequately described?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Are the results clearly presented?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Are the conclusions supported by the results?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comments and Suggestions for Authors

In this paper, the authors have shown how to search for outliers using the check_outliers() function of the {performance} package while following current good practices. This contribution will help researchers engage in good search practices while providing an outlier detection experience. They cover univariate, multivariate,

▼

and model-based statistical outlier detection methods, their recommended threshold, standard output, and plotting methods. The present paper represents the subject of check the outliers, An accessible introduction to identifying statistical outliers in R with easystats. The study design and methods appropriate for the research question. The results presented clearly and accurately. The authors logically explain the findings. I highly recommend the publication of this paper in Mathematics Journal.

Submission Date	21 February 2023
Date of this review	03 Mar 2023 23:24:46

© 1996-2023 MDPI (Basel, Switzerland) unless otherwise stated

Disclaimer **Terms and Conditions**
(<https://www.mdpi.com/about/terms-and-conditions>)
Privacy Policy (<https://www.mdpi.com/about/privacy>)



B

I

x^2

x_2

↺

Ω

🔗

A

▼

✎

▼

≡

≡

$\frac{1}{2}$

▼

≡

▼

≡

≡

≡

≡

I_x

<>

P

0 WORDS

Word / PDF
 Choose file
 No file chosen

Submit or Save as draft (submit later)

Review Report Form

Open Review

☒ I would not like to sign my review report

☐ I would like to sign my review report

Quality of English Language

☐ English very difficult to understand/incomprehensible

☐ Extensive editing of English language and style required

☐ Moderate English changes required

☒ English language and style are fine/minor spell check required

☐ I am not qualified to assess the quality of English in this paper

	Yes	Can be improved	Must be improved	Not applicable
Does the introduction provide sufficient background and include all relevant references?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Are all the cited references relevant to the research?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Is the research design appropriate?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Are the methods adequately described?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Are the results clearly presented?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Are the conclusions supported by the results?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comments and Suggestions for Authors
 The authors propose the illustration of an R library for detecting outliers. The paper is quite clear, however, the following remarks are provided to improve the article

- The word accessible in the title should be removed.



- The authors should clarify whether the functions implemented in the library were implemented only for continuous variables or other types of variables are also considered.

- The author's reference to the Normal distribution on page 2, line 31 is unclear. They should clarify or delete this sentence.

- The reference to the mathematical score is unclear on page 2, line 36. I think this reference should be removed.

A critical review on the analysis of outliers is proposed in this article which must be cited

Riani, M., & Atkinson, A. C. (2020). Robust regression methods in machine learning. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(2), e1359.

2

- Page 3, lines 101-102: It is unclear what the authors mean when referring to BCA, a technique used in non-parametric bootstrapping, but the paper does not mention this.

- It is necessary to change the x-axis labels of all the figures as the numbers overlap; they should be written in a smaller font or otherwise eliminate some of them since they are not readable. I am talking about the figures referring to the histogram of the outliers.

- Page seven, line 146, it is better to write association instead of relationship.

- Page 7, line 154, the authors should specify what they mean by compatible regression models.

- Page 7, code of the first chunk, the authors use the iqr method, but this is not discussed above. They must add details of all the methods they show in the output.

- Page 9, line 247, it is difficult to follow the discussion for those unfamiliar with the R datawizard package. What does it do? How does the proposed function tie in with those in the package?

- Page 10, line 272, what does "Registration" mean?

- Check the reference style because it is not uniform



B

I

x^2

x_2

$\frac{1}{2}$

Ω

A

I_x

<>

P

0 WORDS

Word / PDF

Choose file
 No file chosen

Submit or Save as draft (submit later)

Review Report Form

- Open Review

☒ I would not like to sign my review report

☐ I would like to sign my review report
- Quality of English Language

☐ English very difficult to understand/incomprehensible

☐ Extensive editing of English language and style required

☐ Moderate English changes required

☒ English language and style are fine/minor spell check required

☐ I am not qualified to assess the quality of English in this paper

	Yes	Can be improved	Must be improved	Not applicable
Does the introduction provide sufficient background and include all relevant references?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Are all the cited references relevant to the research?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Is the research design appropriate?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Are the methods adequately described?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Are the results clearly presented?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Are the conclusions supported by the results?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Comments and Suggestions for Authors
 The paper outlines techniques for processing data with inhomogeneities, such as outliers. However, the style of the article does not correspond to scientific research. The work is educational and auxiliary in nature and is more in line with educational publications, including popular scientific ones.

