**August,** 10, 2021

The Editor

*The R Journal*

Dear Editor,

We would like to thank you for the letter dated 14/06/2021, and the opportunity to resubmit a revised copy of the manuscript “Measuring the Extent and Patterns of Urban Shrinkage for Small Towns Using R”. We would also like to take this opportunity to express our thanks to the reviewers for their time and effort dedicated to providing feedback on our manuscript, as well as for the insightful comments on and valuable improvements to our paper.

We believe the reviewers suggestions considerably improved our revised manuscript, which you will find uploaded alongside this document. Following this letter are the editor and reviewer comments with our responses in italics, including how and where the text was modified. Changes made in the manuscript are marked using track changes. The revision has been developed in consultation with all co-authors, and each author has given approval to the final form of this revision.

We very much hope the revised manuscript is accepted for publication in the R Journal.

Sincerely yours,

Cristiana Vîlcea on behalf of the authors

**R Journal Review**

Measuring the Extent and Patterns of Urban Shrinkage for Small Towns Using R

**Review 1**

**Overview**:

An interesting topic, overall, is discussed in this article, while showcasing the construction of composite indicators in an attempt to explain urban shrinkage in small towns of Romania, that would be a good addition to the published literature regarding respective domain experts that use R. Principal component analysis was used for this purpose, in a sensible manner, maintaining the general methodology easy to comprehend for readers with prior knowledge of unsupervised learning and dimensionality reduction procedures.

*As R is not yet a common language used by most of the geographers in their studies, we intend to show that it could be integrated in geographic research in order to bring new scientific approaches, different from those used before. The authors try to show that it is a language with applications in different geographical domains, due to its multiple existing packages.*

From a more technical perspective, the pipeline followed was unclear making the reader lose track of the order that results are presented; maybe the authors should revise the language used to improve readability (see Article). Base R’s prcomp() is indeed a complete function that applies Principal component analysis via singular value decomposition, so external libraries were not necessary.

*Thank you for pointing that out. The reviewer is correct. We tried to address these issues taking into account your critique and improved readability.*

**Article**

Structure and content

Overall, text contains much information especially within the Introduction section. Maybe the reader would find it more helpful if paragraphs, regarding the background of this research, were more well-structured (one topic each and roughly equal in length), were not one-sentenced and there was a logical evolution of evidence.

*Thank you for this comment. This section was changed considerably; it no longer includes loose references and information and it now focuses only on the causes and extent of the phenomenon, as well as the indicators of urban shrinkage, which are the basis for the present study. We have followed the reviewer’s suggestion to have well-structured paragraphs, roughly equal in length, focusing on one topic each. Please see the section Introduction in the article.*

Data / Methodology

1. I found the term “*statistical data*” in “Data and methods” section, confusing. Is there a chance that you meant to use the term “Official Statistics” taking into account that data were provided by an official body?

*Thank you for pointing that out. The reviewer is correct, we meant to use the term ‘official statistics’. We replaced it accordingly.*

1. Authors also mention that “*statistical data was important to calculate indicators and construct composite indices*.”. Please describe how these indicators were calculated.

*We thank the reviewer for raising that point. We have added a brief explanation for calculating these indicators, in the first paragraph of the section Data and methods, which reads as:*

*„The multidimensional concept of shrinking cities has a comprehensive meaning, as the drivers which determine this process are complex (demographic, economic, social), therefore official statistics from the last three censuses (1992, 2002, 2011) and from 2018 were used to calculate indicators (applying simple standard formulae in population geography regarding the demographic and economic phenomena such as population natural increase, feminization, migration, ageing, unemployment, i.e. k1- k6) and construct composite indices. All indices used in the analysis of this phenomenon were based on their importance and relationship to each other and to the concept itself.*

1. Missing values, along with sample size, are usually the most significant factors that limit the results of a research, affecting both statistical analyses and the interpretation of results. Inspecting the given dataset, authors imputed missing values with zero; however, this is not mentioned in the text. Although the frequency of missingness was not high and zero imputing would hardly affect the final results, it would be useful to see a few lines addressing the authors’ thoughts behind this approach.

*Reviewers are right. We have added a brief explanation regarding why we chose this approach, which now reads as.*

*„Thus, in order to perform the calculation without errors caused by missing values, we imputed these missing values with zero, as the frequency of missingness was not high and it was observed that zero imputing does not affect the results of the study“.*

1. Report the version of R used rather than the version of RStudio and please cite it according to the Journal’s guidelines.

*We replaced the reference as suggested by the reviewer.*

1. Is variable standardization what authors mean in page 5? If yes, the “equally weighted” methodology from (Becker, et al., 2017) does not reflect on this kind of variable scaling. (They mention specifically: “*All seven components are weighted equally with a weight of 1/7*”).

*We refer to variable standardization that is performed* *by prcomp() with the option scale=true. We also wanted to point out that, all indicators have the same importance or weight. We rephased the paragraph to eliminate all misunderstandings.*

1. Authors mention that “*As the variables have different measurement units, the correlation matrix was used to weight all the variables equally*”. I could not replicate the principal component loadings when I run *data.frame(t(cor(pca$x, data[, 1:ncol(data)])))* in the author’s custom function prepare\_biplot\_graphic(). I was able to get close to the values shown in table 4 by running *round(pca$rotation, 3)* -but again, there seems to be a slight deference in rounding. Please confirm that replication is possible with the respective code.

*Thank you for pointing that out. The reviewer is correct, there were some rounding issues. We checked and corrected data. It should work now. We also placed the summary values extracted and used in the tables included in the article (summary and pca\_loadings) in the research data sheet.*

1. The methodology used to forecast Composite Shrinking Index for the year 2030 is not discussed anywhere in the article. In order to avoid further confusion, a small respective part is needed.

*We thank the reviewer for raising that point. We have added a brief explanation for the methodology used to forecast Composite Shrinking Index for the year 2030.*

*Based on the values calculated for Composite Shrinking Index for 1992, 2002, 2011 and 2018 (see sheets aggregation\_1992, column agregat\_92; aggregation\_2002, column agregat\_02; aggregation\_2011, column agregat\_11 and aggregation\_2018, column agregat\_18) we used the forecast function in Excel to predict the values for 2030 (sheet aggregation summary). We also corrected the values, as we observed some minor mistakes in the computed values.*

1. Instead of pasting code, please cite the .R file from the supplementary materials.

*At the reviewer’s recommendation we deleted the code within the article and made the reference to the supplementary materials.*