Imputation of Missing Values with the R Package VIM

Matthias Templ[1,2,3,*]**, Alexander Kowarik**[1,3]

1. Statistics Austria & Vienna University of Technology
2. Vienna University of Technology
3. data-analysis OG
[*]Contact author: matthias.templ@gmail.com

Keywords: Missing Values, Imputation, Survey, R

The package **VIM** [4, 3] is developed to explore and analyze the structure of missing values in data using visualization methods, to impute these missing values with the built-in imputation methods and to verify the imputation process using visualization tools, as well as to produce high-quality graphics for publications.

The most common imputation methods such as hotdeck, kNN and regression imputation are implemented, but also more advanced methods like iterative robust regression estimation are available.

A point- and click graphical user interface has been developed to give access to these methods and tools to users with limited R skills. It is available in the package **VIMGUI** [2]. All important methods are supported by the flexible point- and click-interface.

The presentation describes the application of the methods available in the package **VIM** and demonstrates the usage of the graphical user interface of the **VIMGUI** package. Special attention is also given to the VIM integration of survey [1] objects.

References

[1] Lumley, T. (2012). *survey: analysis of complex survey samples*. R package version 3.28-2.

[2] Schopfhauser, S., M. Templ, A. Alfons, A. Kowarik, and B. Prantner (2013). *VIMGUI: Visualization and Imputation of Missing Values*. R package version 0.9.0.

[3] Templ, M., A. Alfons, and P. Filzmoser (2012). Exploring incomplete data using visualization techniques. *Advances in Data Analysis and Classification 6*, 29–47.

[4] Templ, M., A. Alfons, A. Kowarik, and B. Prantner (2013). *VIM: Visualization and Imputation of Missing Values*. R package version 4.0.0.