CHECK YOUR OUTLIERS! AN INTRODUCTION TO IDENTIFYING STATISTICAL OUTLIERS IN R WITH *EASYSTATS*

Rémi Thériault
Department of Psychology, Université du Québec à Montréal, Montréal, Québec, Canada

Mattan S. Ben-Shachar
Independent Researcher, Ramat Gan, Israel

Indrajeet Patil
Center for Humans and Machines, Max Planck Institute for Human Development, Berlin, Germany

Daniel Lüdecke
Institute of Medical Sociology, University Medical Center Hamburg-Eppendorf, Germany

Brenton M. Wiernik
Independent Researcher, Tampa, FL, USA

Dominique Makowski
School of Psychology, University of Sussex, Brighton, UK

October 23, 2023

Abstract

Beyond the challenge of keeping up-to-date with current best practices regarding the diagnosis and treatment of outliers, an additional difficulty arises concerning the mathematical implementation of the recommended methods. Here, 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 *{performance}* package of the *easystats* ecosystem. We cover univariate, multivariate, and model-based statistical outlier detection methods, their recommended threshold, standard output, and plotting methods. We conclude by reviewing the different theoretical types of outliers, whether to exclude or winsorize them, and the importance of transparency.

Keywords univariate outliers; multivariate outliers; robust detection methods; R; easystats