# ESTIMATING IN NON-PROBABILITY SURVEYS WITH R

**M. Rueda[1] and R. Ferri-García[2] and L. Castro-Martín[3]**

[1] University of Granada, Spain
e-mail: mrueda@ugr.es

[2] University of Granada, Spain
e-mail: rferri@ugr.es

[3] Andalusian School of Public Health, Spain
e-mail: luiscastro193@gmail.com

## Abstract

The convenience of non-probability surveys, often in the form of online surveys, is widely known. They allow for an easy, cheap and efficient way of recollecting data. However, researchers are also aware of the important bias problems which are associated with these kinds of methodologies. A wide variety of methods have been proposed in the last years in order to reduce the bias which such surveys imply. In practice, however, the application of those methods may become a difficult task. This course will show how to correctly apply them with R using the NonProbEst package, which includes state-of-the-art techniques and machine learning models. It includes basic concepts as well as important considerations in order to obtain optimal results, such as the application of hyperparameter optimization processes. A diverse set of alternatives will be covered so it can be adapted to any context, including Propensity Score Adjusting (also known as Propensity Weighting), Statistical Matching and Model Based approaches (also known as Mass Imputation), Model Assisted, Model Calibrated and variance estimation via Jackknife. Even though previous knowledge of the programming language R is required, the contents are presented in a clear and easy to apply way.

**Keywords:** Non-probability sampling, Propensity Score Adjustment, Statistical Matching, model-based estimators, Propensity Weighting, Mass Imputation