# Graphical Exploratory Data Analysis Using Halfspace Depth

Ivan Mizera

The contribution centers around the implementation of two graphical exploratory tools based on different variants of halfspace depth. The first one is the `bagplot`, a bivariate generalization of the univariate boxplot proposed by Rousseeuw, Ruts, and Tukey (1999). The second one is the `lsdplot`, proposed by Mizera and Müller (2004). Some other, related procedures are explained, implemented and discussed as well.

While data depth in general, and halfspace depth in particular, offers a lot of data-analytic inspiration, it is the two-dimensional setting that is best shielded from various algorithmic curses of dimensionality haunting the general concept, and offers a visual perspective that might attract even conservative mainstream practicioners. Hence the choice of the subject, whose second part was driven by an evident desire for self-promotion, not that unknown from various R contributed packages; however, the first part is an attempted penance in the form of public service, not that unknown from various R contributions even more. In order to give the same care to both biological and adoptive child, a new contouring algorithm for location halfspace depth (rather than the existing S code cut and pasted) had to be developed in collaboration with David Eppstein. As a result, we aim at giving the useR not a quick brew, but a considerably matured product, in which not only `lsdplot`, but also `bagplot|` is capable of swallowing and digesting hundreds of thousands datapoints in mere seconds.

## References

Peter J. Rousseeuw, Ida Ruts, and John W. Tukey (1999). The bagplot: A bivariate boxplot. *The American Statistician*, **53**, 382–387.

Ivan Mizera and Christine H. Müller (2004). Location-scale depth (with discussion and reejoinder), *Journal of the American Statistical Association*, **99**, 949–989.