# RKTeaching: A new R package for teaching Statistics

**Alfredo Sánchez-Alberca**[1,*]

1. San Pablo CEU University
*Contact author: asalber@ceu.es

Step by step, *R* is becoming one of the main software used by the scientific community for data analysis, displacing other giants like *SPSS*, *SAS* or *STATA*. *R* has a lot of strengths, most of them as a consequence of being open source, but its main weakness is the lack of a mature Graphical User Interface (GUI), making it a little intimidating for students and beginners. That is the main reason that R has not spread yet to other public, especially in the field of education. Fortunately, in the last years some GUI are emerging to overcome this drawback, as for example **R commander**, **JGR**, **RStudio** or **RKWard**, but they are still not as user friendly as commercial GUIs.

One of the most promising GUI is **RKWard** [Rödiger 2012]. RKWard has a lot of advantages over its competitors. First, it is open source and multi platform. Second, it is based in KDE and QT graphics libraries, much more modern than the tcl/tk libraries that uses **R commander**. And third, the most important, unlike **RStudio**, it is highly expandable and customizable by means of plugins. In addition, **RKWard** is not just a GUI, but a complete development environment that has been designed both for beginners and experts.

For all these reasons, I decided three years ago to develop a new *R* package for teaching statistics based on **RKWard**. My main goal was to simplify and facilitate the use of *R* to reduce the learning curve. The new package is called **RKTeaching** and, at this moment, it includes menus, dialogs and procedures for data preprocessing (filtering, recoding, weighting), frequency tabulation (grouped and non grouped data), plotting (bar charts, pie charts, histograms, box charts, scatter plots), descriptive statistics (mean, median, mode, percentiles, variance, unbiased variance, std deviation, unbiased std deviation, coefficient of variation, intercuartile range, skewness coefficient, kurtosis coefficient), probability distributions (Binomial, Poison, Uniform, Normal, Chi square, Student t, Fisher's F, exponential), parametric tests (T test for one mean and mean comparison (independent and paired samples), F test for variances comparison, Normal and Binomial tests for one proportion and proportions comparison, ANOVA for multiple factors and repeated measures, sample size calculation for estimating means and proportions), non parametric tests (Kolmogorov-Smirnov, Shapiro Wilks, Mann Whitney U, Wilcoxon, Kruskal Wallis, Friedman, Chi square), concordance tests (intraclass correlation coefficient, Cohen's Kappa), regression and correlation (linear, non linear, logistic, models comparison, predictions), simulations (coin tosses, dice roll, small numbers law, central limit theorem). All these menus include optional assistance that guide the user step by step trough the analysis and help interpret the results. In addition, some procedures have an extended version where all the calculations and formulas used in the analysis are displayed in order to help students to understand the procedure.

In the last two years we have used this package to teach statistics to Medical, Pharmacy, Psychology, Biotechnology, Optics, Nursing and Nutritional sciences degrees, with very satisfying results. To assess the achievements we conducted an experiment to compare **RKTeaching** with previous software. The assessment by the students clearly reflects its ease of use and learning compared to *SPSS* [Sánchez 2012].

## References

Rödiger, S et al. (2012). RKWard: A Comprehensive Graphical User Interface and Integrated Development Environment for Statistical Analysis with R. *Journal of Statistical Software 49(9)*, 1–34.

Sánchez-Alberca, A. (2012). RKTeaching: un paquete de R para la enseanza de Estad´. In *III Jornadas de Intercambio de Experiencias de Innovacin Educativa en Estad´(Valencia, Spain)*, pp. 136–147.