

Read, write, format Excel 2007 (xlsx) files

Adrian A. Drăgulescu*

*Contact author: adrian.dragulescu@gmail.com

Keywords: Excel 2007, rJava, xlsx

The package `xlsx`[1] makes possible to interact with Excel 2007 files from R. While a power R user usually does not need to use Excel or even avoids it altogether, there are cases when being able to generate Excel output or to read Excel files into R is useful. For example, in an office environment where you need to collaborate with co-workers who use Excel as their primary tool. Or, to use Excel's formatting capabilities for reporting small to medium sized data sets.

The approach taken in the `xlsx` package is to use a proven, existing API between Java and Excel 2007 and use the `rJava` [2] package to link Java and R. The advantage of this approach is that the code on the R side is compact, easy to maintain and extend, even for people with little Java experience. All the heavy lifting of reading/writing xlsx files is being done in Java. The Java code used by `xlsx` is a project of the Apache Software Foundation[3]. So our approach benefits from a mature software project with many developers, test suites, and users that report issues on the Java side. While it is possible to interact directly with xlsx files from R as shown by the package `RExcelXML`[4], doing this in R is a big task that our approach avoids.

With the package `xlsx`, besides reading and writing xlsx files, you can programatically control cell properties to set text color and background color, set a cell font and data format. You can add borders, hide/unhide sheets, add/remove rows, add/remove sheets, etc. You can split panes, freeze panes, auto size columns, merge cell regions, work with cell comments and more. You can control the printer setup by setting the landscape mode, setting the number of copies, adjust the page margins, or if the copies should be in color.

References

- [1] Adrian A. Drăgulescu (2010). *Read, write, format Excel 2007 (xlsx) files*. R package version 0.1.1. <http://CRAN.R-project.org/package=xlsx>
- [2] Simon Urbanek (2009). *rJava: Low-level R to Java interface*. R package version 0.8-1. <http://CRAN.R-project.org/package=rJava>
- [3] Apache POI project <http://poi.apache.org/>.
- [4] Duncan Temple Lang (2009). *RExcelXML: Read and manipulate new-style (Office '07) Excel files*. <http://www.omegahat.org/RExcelXML/>