SimpleR: Taking on the "Evil Empire" by Developing Applications for Non-statistical Users

The "Standard Model" for R software development – the R package – assumes:

1. Many independent users;
2. Some degree of statistical understanding by users;
3. Access and functionality within the overall R environment;
4. Persistence over time;
5. (Usually) A command line interface.

This model serves serious data analysts well, and R provides many tools to facilitate such development. We argue here that R can also serve another and potentially much larger community of engineers and scientists for whom Excel® (or similar software) is now the primary tool for data analysis and statistical graphics. These folks need:

1. Narrow, often single purpose "template" analyses;
2. Rapid development/rapid discard – needs disappear as soon as a project is completed or change radically when technology changes;
3. Software customized for a few – even one – users;
4. A simple GUI interface requiring minimal documentation and learning;
5. Graphs as the primary output.

Excel is their default because (a) it's there and they know it; (b) they don't know about or can't implement better methods.

R can change this state of affairs. R is open source, has superb graphics, and is easily embedded into web served applications using R2HTML, Rpad, Rserve, Rzope, etc.. Alternatively, it can be modified for single use applications through a GUI such as Rcmdr, gWidgets, or by simply modifying the R menu structure. The key is to present the user with a simple interface and readily interpretable output, even if the underlying analysis is complex.

We discuss our strategy for developing such applications, which relies on the global workspace as the software environment, thus avoiding the unnecessary (for us) overhead of packages. We give an example in actual use at Genentech and discuss the pros and cons of this approach.