

Part 14: RM4Es Summary and RM4Es software introduction

Here in part 14, we summarize our modeling process and provide a RM4Es™ solution together with an introduction of the RM4Es™ software.

Throughout the discussion in previous parts, we have treated model building as a process, and specifically as a RM4Es process. Through specifying equations, conducting estimations, performing evaluations, then going back to re-specify equations or re-estimate new equations, we can finally reach a stage where we are satisfied with our evaluations and can make explanations to the users of our modeling process.

Data analysis has been defined as a repetitive, cyclical search for understanding by many scholars (Lubinsky & Pregibon 1988, Young & Smith 1991). In other words, data analysis is a process that involves a sequence of repetitive actions. The order of these actions, as defined as statistical strategy by some scholars (Harrell 2001), has huge impacts on the results of any data analysis, as demonstrated by a few experiments such as that of professor Faraway (J. J. Faraway 2004).

We name the above mentioned sequence of data analysis actions as research flow as it is a term familiar to business analysts. In previous parts, we have used an RM4Es framework to represent all kinds of research flows for regression modeling.

Based on the above mentioned RM4Es representation framework, we have developed an RM4Es software system that allows users to create research flows, to open saved research flows and modify them, then to run any research flows in the open source statistical software R. Research flows can fully represent all statistical strategies (Faraway 2004, Harrell 2001). Therefore, the RM4Es system allows analysts and researchers to manage their statistical strategies, so that they can re-use them and improve them in the future.

At the same time, our RM4Es software will provide features facilitating research flow mapping, and intelligent guidance for research flow formation. As claimed by professor Young and professor Lubinsky (Young & Lubinsky 1995), in an environment where research flows can be structured and related intelligent guidance can be provided, data analysis productivity, accuracy, accessibility and satisfaction will all be improved. The users of our RM4Es framework are confirming this claim with their experience. Students of predictive modeling also found that they learned fast and learned more than before, with an RM4Es software system in place.

Statistical software packages are widely available, but all of them only handle statistical tactical computation. The RM4Es system as discussed in this note performs computing for statistical strategies that represents a new direction for statistical computing and has been considered as an important breakthrough.

Specifically, the RM4Es software is a research flow management system. A research flow is a collection of analytical steps such as regression model specification and

ordinary least squares estimation, as ordered by research logic. For example, specifying a multiple regression equation, then selecting dependent variable and independent variables, then choosing ordinary least square estimation method, then evaluating R-square, and finally explaining the results is a simple but common used research flow.

With this RM4Es system, users can easily create a research flow by drag and drop RM4Es components, and run it to get desired results by clicking on RUN. The created research flows can be saved for future use or for sharing with colleagues. The advanced features of the RM4Es system will provide log file management functions and intelligent guidance to help designing research flows.

In general, this RM4Es system allows researchers, analysts and research managers to manage research and knowledge systematically. It will increase research productivity and research quality. It will also help research organizations to meet regulatory requirements. Specifically, systematically managing research flows can improve work efficiency for researchers and analysts. Then, information saved in logfile management folders may be used for professors to help advising students or to meet corporation management & regulatory requirements.