COMING SEPTEMBER 15, 2017

# A Software Toolkit for Data Management

Authors: Kim Chantala, Jean Robinson, Helen Smith, RTI International

## Abstract

We have developed a suite of data tools with the SAS language to help keep data preparation on schedule and within budget. This toolkit provides a remarkably easy, low-cost way to create codebooks, master lists of SAS data sets for a project, reports of variables needing special investigation, and data crosswalks showing the relationship of variables across datasets. Traditionally, these documents are produced at the end of a project with a great deal of programming or are manually produced, but our tools allow the programmer to seamlessly create these documents at any time during the data preparation task. We have found that producing these documents early in data collection improves data quality and communication between the data collection team and client. This toolkit provides a comprehensive way to document and review not only datasets you create, but also datasets that you receive, especially if they do not have good documentation.