

Adding Business Intelligence to Paradata: The Blaise Audit Trail

Joel Devonshire; University of Michigan, Survey Research Center

Previously, the Survey Research Center (SRC) enhanced the usability of Blaise Audit Trail (ADT) data by creating a process by which all ADT files were parsed and loaded into relational database tables. While this provided significant efficiency and security-related benefits, analysis of the ADT data still presented challenges due to the sheer size of the data (every entry into a Blaise field represents a new row of data, creating millions of records). Moreover, because audit trail data remains largely unexplored as a paradata resource, analysts and survey managers do not necessarily know what questions they want to ask, or what questions are possible to answer, with ADT data. This led us to explore the potential of Business Intelligence (BI), particularly the use of Online Analytical Processing, or OLAP, to create more user-friendly analysis and reporting tools for the raw audit trail data.

OLAP designs allow for multi-dimensional, pre-aggregated data, and the key advantage it provides is broad, interactive, ad hoc data exploration and visualization with minimal processing at runtime. Design decisions made up front, such as key group-by variables, dimension hierarchies, and levels/types of data aggregation allow data to be quickly “rolled up” into parent categories or “drilled down” to finer levels of detail, and complex query-building can be greatly simplified.

This paper and presentation will demonstrate how we designed an OLAP cube for ADT data, and will provide examples of flexible analysis and reporting, including interactive pivot tables and web-based visualization. Discussion will focus on how such tools can aid in responsive survey design, make analysis of massive data sets less daunting, and provide efficiencies in IT support for our clients.