

Visual Interfaces to Data

Chris Stolte
Chief Development Officer
Tableau Software
837 N 34th St., Suite 400
Seattle, WA 98103 USA
+1.206.633.3400

cstolte@tableausoftware.com

ABSTRACT

Easy-to-use visual interfaces to data can broadly expand the audience for databases. Domain experts rather than database experts can engage in rapid-fire Q&A sessions with the data. Visual interfaces can provide a medium for story-telling, debate, and conversations about the data. They can also put new and challenging demands on the capabilities of traditional relational databases.

In this talk, I will describe our formal language-based approach to visual analysis and how the use of a formal language enables us to build user experiences that more effectively support the process of analysis. Tableau's VizQL algebra is a declarative language for succinctly describing visual representations of data and analytics operations on the data. A VizQL statement compiles into the SQL or MDX queries necessary to generate the view and into the graphical commands to render the interactive view of the data. Our easy-to-use drag-and-drop user experiences for analysis and visual interface authoring are built on top of VizQL.

In addition to supporting the process of analysis, a formal language-based approach provides a basis for reasoning about the structure of views and the space of possible views. This in turn enables the development of powerful new analytic capabilities, such as automatic presentation of structured data, visual authoring of statistical models, and view-based calculation, which we demonstrate.

I will also discuss the challenges we have faced in getting relational databases "in the wild" to effectively support visual analysis for the average business or scientific user. The challenges range from the technical to the political. Traditional relational databases, both for OLTP and OLAP, often require sophisticated data modeling and data management expertise, optimize for

performance based on known workloads, and are designed for scaling to large databases sizes (e.g. PB or TB) on clusters of machines rather than reducing analytic latency using limited hardware. I will describe our approaches to building a database focused on providing interactive query performance on tens or hundreds of millions of rows of data with little or no data modeling (physical or logical) and running on a typical knowledge worker desktop machine.

Finally, I will discuss the changing landscape of interfaces to databases. The original interface to the database was transactional in focus: Many users read and make atomic changes to a small number of rows in a large database. In recent decades, powerful analytic use cases have emerged focused on the study and analysis of massive amounts of data by relatively small numbers of power users. The emergence of easily authored visual interfaces to public and private data changes will enable a new style of database usage. Millions of users performing analytics on thousands of data sets all hosted in the cloud with usage demonstrating the familiar long-tail distribution. Everyone will become an author and all interfaces will enable analytics.

Categories and Subject Descriptors

H.5.0 [Information Systems]: Information Interfaces and Presentation.

General Terms

Design, Human Factors.

Keywords

Visual analytics.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SIGMOD'10, June 6–10, 2010, Indianapolis, Indiana, USA.

Copyright 2010 ACM 978-1-4503-0032-2/10/06...\$10.00.