http://webcache.googleusercontent.com/search?q=cache:c90N5lDtG0sJ:www.kimballgroup.com/2008/07/drill-down-to-ask-why-part-2/+&cd=3&hl=en&ct=clnk&gl=us

# Drill Down to Ask Why, Part 2

By Ralph Kimball

Last month, I described analytic application stages that are useful to think about the architecture of a data warehouse (DW)/business intelligence (BI) system. We considered the example of an airline fare planner who is tasked with determining why the yield has dropped. Remember that in this case yield refers to a key performance indicator (KPI) signifying the revenue or profits from a fixed, perishable resource. In this month's column, I'll continue with this example and explore how to implement analytic capabilities to explore "why" in a real data warehouse/business intelligence system. The fare planner may ask "why" in five ways:

**1. Give me more detail.**

In a dimensional data warehouse environment where all business process subject areas are built as fact tables and dimension tables on the lowest level atomic data, drilling down is easily accomplished by adding a row header (grouping column) to the report query from any of the attached dimensions. In your yield report, if you assume the report is calculated from a database of individual boarding passes, then drilling down by date, time of day, aircraft type or fare class is simply a user interface gesture that drags the new grouping attribute from the relevant dimension into the query. Note that this general form of drilling down does not require a predetermined or predeclared hierarchy. As I have been pointing out for 15 years, drilling down has nothing to do with hierarchies!

**2. Give me a comparison.**

In the example, if the route is the row header of the yield report, then a common comparison mode would be to add one or more numeric columns to the report, each perhaps with a different date. Thus, the yields at different dates could be compared. This can be accomplished in several different ways by your BI tool, but as the comparisons become more complex, a drill-across approach to sort merging separate queries is the most practical and scalable. See my April and May DM Review articles for more information on drill-across techniques. Keep in mind that comparisons do not need to be presented solely as text and number rectangular reports. Sometimes comparisons are more effective graphically, especially with time series data.

**3. Let me search for other factors.**

In order to jump to a separate database, the context of the current report query must be captured and used as input to the new database. When the user selects a specific row or cell in the original report, that detailed context should be used. For instance, if the June 2008 yield on the San Jose to Chicago route is selected by the user, then June 2008, San Jose and Chicago should be carried as constraints on the next query. Perhaps the next query focuses on weather. This scenario presents some challenges to the data warehouse/business intelligence designer. The user can't jump to another database if it isn't there. Thus, the designer needs to anticipate and provide possible targets for this kind of jump. Also, the BI tool should support the context capture steps described in this section so that the process of jumping to a new database is as effortless as possible.

**4. Tell me what explains the variance.**

In order to feed data mining, the data warehouse/business intelligence designer must again anticipate the database resources that will be needed when these kinds of analyses are performed. Practical information about how to build these kinds of data mining interfaces, including decision trees, is available in Michael Berry and Gordon Linoff's book *Data Mining Techniques For Marketing, Sales and Customer Relationship Management.*

**5. Search the Web.**

Finally, if you have successfully searched for other factors (number three), you need to be able to transfer the context of the exception you have identified into a Google or Yahoo! query. If you are dubious that this would be of much value, search for "airline yield 2007 versus 2006." This can be a paradigm-shift kind of experience. What we need is a one-button jump from a report cell to a browser.

In this column, we have reminded ourselves of the true goal of the data warehouse/business intelligence system: assisting the end user in making decisions. The key capability is providing the most flexible and comprehensive ways to drill down and ask why.