INFORMATION DECISION SUPPORT DATA · ACCESS
UNIVERSITY OF ILLINOIS

Finance EDW: NULL Values and the Data Warehouse

Last Updated: 21 May 2003

NULL data can be defined as missing or non-existent data, and is generally a hindrance to any analysis. A particular problem exists with relational databases, whereby SQL can automatically exclude rows having NULL values from query results. This can occur regardless of the intent of the query writer or analyst, and can cause considerable confusion. Such an occurrence may also cause reports to have incomplete data.

Nevertheless, NULL values are widely dispersed throughout the Banner source data for the Enterprise Data Warehouse (EDW). Where possible, the goal for the EDW is to transform each Banner NULL to another value so that the EDW can be more easily queried. There are a few situations, however, where it is advantageous or appropriate to leave data as NULL in the EDW. This document describes the NULL value-related transformations applied to Finance Increment data.

Character Data: Transform NULL to a single space

A simple transformation used in the EDW is to translate a NULL character value (e.g., a NULL in a code, name or description) to a single space. Since a single space is never used as a code, name, or description, this transformation preserves the original meaning of the NULL value as "Missing" or "Not Applicable".

However, there are a handful of columns in the Chart of Accounts Hierarchy tables for which character data is allowed to be NULL. These NULL values are not the result of NULL values in Banner. For example, there are columns for the hierarchy associated with each organization. If an organization is at a higher level in the hierarchy, at the college level for example, columns for the lower levels of the hierarchy, e.g., department, are set to NULL rather than space. This was done in order to facilitate UI-Integrate reports in which the report logic was simpler if NULLs were used instead of spaces.

Transformations of Numeric or Date/Time Data

Unlike character data, there are no universal transformations for numeric or date/time data, because there are no alternative values that can always be understood as "Missing" or "Not Applicable". In the EDW, numeric and date/time data is transformed from NULL only when a specific meaning can be assigned. Two noteworthy examples of such transformations are the following:

Implied document amounts – Banner will leave some amounts on financial documents NULL if they do not apply, but this can be safely translated to zero in the EDW. Tax on a purchase order or invoice voucher is one specific example.

Implied expiration date – In some cases where Banner allows history tracking, a NULL expiration date is assumed to mean the end of time. This can safely be translated to the EDW end-of-time value, i.e., 12/31/9999. A specific example of this situation is the

expiration date used with the history of financial managers assigned to funds, organizations or programs.

NULLs in EDW Ledger Summary and Detail Amounts

In the Banner financial ledgers, a single dollar amount is used with each accounting entry. Other data must be used to determine what that amount represents. For example, there is an indicator for debit or credit in the general ledger.

In the EDW such amounts are divided into separate columns, such as debit amount and credit amount. For a given entry, only one of these columns will contain data. The other column will be set to NULL.

This applies to the general, operating and encumbrance ledgers. In the operating ledger, a dollar amount will exclusively be a budget, revenue, expense, or encumbrance amount. In the encumbrance ledger, a dollar amount will exclusively be an original, adjustment, or liquidation amount. Again, only one of these columns will contain data; the others are set to NULL.