White Paper Excerpt

I. Overview

Organizations can collect and store questionnaire data (answers to questions) within one or more tables in a data warehouse. Questionnaire data may comprise any set of questions that are pertinent to an organization's operation. Questionnaire data may be obtained through marketing surveys, employment assessments, recruiting tests and customer service evaluations.

The means by which data is organized can impact the performance of queries directed to the data warehouse to retrieve information. Two methods for organizing and accessing questionnaire data will be presented and analyzed from a performance and scalability perspective. According to the first method, questionnaire data is organized into columns, with each column corresponding to a particular question and each element within a column corresponding to an answer to the question. In the second method, data is organized into rows, with each row corresponding to a particular question and answer combination.

In this white paper, hypothetical questionnaire data is gathered by a high-tech store located within a particular metropolitan area. A user can query the data warehouse to gauge, for example, which consumer answered "yes" to one question and "no" to another.

II. Method 1

According to the first method, questionnaire data is organized into columns, with each column corresponding to a particular question. Each element within a column corresponds to an answer to a question. Table 1 (CONSUMER_FACT_1) includes hypothetical answers to questions presented in Table 2 (LU_QUESTION).

Table 1: CONSUMER_FACT_1

CONSUMER_ID	CONSUMER_DESC	Q1	Q2	Q3	Q4	Q5
1	Sally Sherman	yes	yes	yes	yes	Sony High Definition Television
2	Bob Wadha	no	no	yes	yes	Sony Playstation 3
3	Shawn Rickson	no	no	yes	yes	Microsoft Xbox 360
4	Paul Barry	no	no	yes	yes	Apple iPod
5	Joe Sandburg	yes	no	yes	yes	Microsoft Xbox 360
6	Al Raza	yes	no	yes	no	Sony High Definition Television
7	Frank Bonds	yes	no	yes	yes	Apple iPod
8	Allison Weiss	no	no	no	yes	Sony Playstation 3
9	Horatio Iglesia	no	no	yes	yes	Microsoft Xbox 360
10	Xing Zu	no	no	yes	yes	Sony High Definition Television

Table 2: LU_QUESTION

QUESTION_ID	QUESTION_DESC
Q1	Are you married?
Q2	Do you own a house?
Q3	Do you own a car?
Q4	Do you have a job?
Q5	If you could purchase only one of the following, which would you choose? Apple Ipod, Microsoft Xbox 360, Sony High Definition Television, or Sony Playstation 3?

To illustrate how a user (*e.g.*, store manager) may access CONSUMER_FACT_1, two examples will be presented and the SQL for each test case will be provided.

Example I: The store manager wants to gauge which consumers are married (*i.e.*, answer to Q1 is "yes") and would purchase a Microsoft Xbox 360 (*i.e.*, answer to Q5 is "Microsoft Xbox 360"). The SQL for example I is shown below. The data warehouse returned one entry, which is shown in Table 3...