

Comparative Study Of Storage Of Responses Of Questionnaires In Databases

Mrs. Ranjana Ingolikar
Dept. of Comp. Sci.
SFS College
RTM Nagpur University

Mrs. Rasika Khandal
MCA Department
SRPCE, Nagpur

Rahul Mohare
DMIMS
RTM Nagpur University

Abstract

The purpose of this paper is to identify and explore the way of storing questionnaire responses in the relational data model. To explore the way of storing responses, help has been taken of a questionnaire designed by a business management researcher. The simple excel response sheet is redefined by using relational data model terminology. A comparative study of storage of responses of questionnaires in MS-Excel, MS-Access and RDBMS databases is made. It has been found that the database RDBMS is better option of storing responses of questionnaires.

information they want from it. They let the database management system software take care of the described data structures for storing the data and retrieval procedures for answering queries by using query engine [1]. Some major advantages of relational data model are:

- Using RDBMS tables can be created for storing questionnaire data and each table has one primary key and foreign keys.
- The primary key establishes the relationship between the tables using foreign key. A foreign key represents the relationship with other table.
- Referential integrity ensures that data is consistently defined across the tables. It means accurate data of each respondent is present in the table.

1. Introduction

1.1. Questionnaire

A questionnaire is one type of data collection method, which contains formalized set of structured type of questions and to obtain responses from respondents. Respondents were asked to mark their responses. Usually novices store their responses in the structured excel sheet. Here an attempt is made to associate relational data model with the storage of questionnaire responses. Business Management researcher whose case study was considered for research had used excel response sheet for storing responses. It is redefined by using relational data model approach.

1.2. Relational Data Model

In relational model of a database, data is represented in terms of tuples or rows, made up of atomic values and grouped them into relations. The purpose of the relational model is to provide a declarative and predefined syntax for specifying data and queries. Users directly state what information the database contains and what

Here it is proposed that the relational model design can be used to partition the questionnaire into different relations. The partitioning is necessarily a lossless join decomposition of the questionnaire. To illustrate the concept of partitioning, the following case study has been considered.

2. Case Study

The management researcher Mrs. Soma Sharma, Research scholar of RTM Nagpur University, in her Ph.D. thesis "A Study of Micro Finance as a tool for alleviating urban poverty with reference to Nagpur City" has done her work in MS-Excel response sheet for storing responses.

In this paper the best way of storing given responses in the relational data model is proposed. The questionnaire contains 35 structured questions which are multiple choice questions filled from respondents and whose responses were stored in

excel which are to be partitioned into relations using MS-Access database, which is nothing but relational database model and is used to store data in form of relations.

Multiple-Choice Questions: In which the management researcher provides a choice of answers and respondents are asked to select only one alternative from the given choices.

2.1 Problems/Limitations of MS-Excel sheet

The researcher used MS-Excel sheet for storing responses. With reference to the case study considered for research, storing responses in the MS-Excel datasheet is not an efficient way. It has the following major limitations:

1. Data types and default values of data types are not supported by excel.
2. Field width or size of field is not supported by excel.
3. Validation rules and validation text can not applied.
4. Text formatting is not supported.
5. MS-Excel does not give the facility of data query processing.
6. Data present in an MS-Excel does not support referential integrity rules for establishing relationship between data.
7. There are limitations in data storage size. Eg. In the new version of MS-Excel2007, the limitations are 1,048,576 rows by 16,384 columns. [a][b].

Looking at the above limitations or problems of MS-Excel datasheet. We propose the partitioning questionnaire and use of relational data model for storing responses.

2.2 Partition of Questionnaire

To store the responses of respondent to the questionnaire the following tables are used:

1. RESPONDENT_DETAILS
2. RESPONDENT_FAMILY_DETAILS
3. HSG_LOAN_DETAILS
4. MICRO_FINANCE_DETAILS
5. HSG_DETAILS

Relations are generated by partitioning the fields using lossless join decomposition. Each relation is in 3rd Normal Form because it contains non-

redundant data of each respondent. The relation has number of fields with specific data type and their corresponding values which are responded by the respondent.

Suppose the relation RESPONDENT_FAMILY_DETAILS has primary key RSPD_FM_ID and foreign key RSPD_ID which is a primary key of relation RESPONDENT_DETAILS. The foreign key field is used to establish the relationship between relations.

Since a respondent is responding to the questionnaire, it is proposed to generate a unique respondent number for each respondent which will act as a primary key for master table whose references can be taken in some other tables. 5 such relations are formed RESPONDENT_DETAILS which is a master table and four child tables are RESPONDENT_FAMILY_DETAILS, HSG_LOAN_DETAILS, MICRO_FINANCE_DETAILS and HSG_DETAILS with corresponding attributes and definition of primary key and foreign key.

3. Lossless Join Decomposition

The concept of a Lossless-Join Decomposition eliminates redundancy safely from databases, preserving the original data [2].

3.1. How Lossless Join Decomposition is achieved

When the relational data model is designed for questionnaire, relations are partitioned or decomposed into BCNF or 3rd NF which is always lossless in order to eliminate duplication from tuples. Normalization performs the major role when a relation is decomposed. Functional Dependency is the information in a relation which depends on other information in the same relation.

Example: In the above case study, each respondent gives his personal information by filling the respondent sheet. Following information gathered from respondent is:

RESPONDENT = {RSPD_ID, RSPD_NO, RSPD_Nm, RSPD_Age, RSPD_Edu, RSPD_Status, RSPD_Occupation, No_Fm_Mem, No_Earning_Mem, Monthly_Income, Housing_Status}

Some of the functional dependencies formed between fields are:

$F = \{RSPD_NO \rightarrow RSPD_Nm;$
 $No_Fm_Mem \rightarrow Monthly_Income;$
 $RSPD_Status \rightarrow Monthly_Income\}$

Now, the relation RESPONDENT is further decomposed into

RESPONDENT_DETAILS = {RSPD_ID,
 RSPD_NO, RSPD_Nm, RSPD_Age, RSPD_Edu,
 RSPD_Status, RSPD_Occupation}

RESPONDENT_FAMILY_DETAILS =
 {RSPD_Fm_ID, RSPD_ID,
 RSPD_NO, No_Fm_Mem, No_Earning_Mem,
 Monthly_Income, Housing_Status}

Hence, the relations RESPONDENT_DETAILS and RESPONDENT_FAMILY_DETAILS have achieved lossless join decomposition because join of decomposed relations should be able to create the original relation (i.e. no spurious tuples are present in the relation) [3].

3.2. How this approach served the purpose of applying this method

By applying this method, partitioning the questionnaire and their responses are stored into MS-Access. Using this approach we decompose the given MS-Excel data sheet into the number of relations which serves the lossless join decomposition. After decomposition each relation of RDBMS serves normalization i.e. is in the 3rd NF which reduces the redundancies from relations and gives the accurate and consistent data to us.

For storing responses of questionnaire, we used MS-Access which gives advantage in query processing.

4. Benefits of RDBMS data model

When questionnaire data is stored in MS-Access, it gives better performance, while storing respondent data in the database, establishing relationship between tables, executing complex queries using RDBMS query processor engine rather than MS-Excel sheet. Some of the benefits are given below:

1. Data can be stored in the form of relation.
2. Because of normalization, only valid or accurate information can be fetched.

3. It support all data types with their default values and field width
 Eg. for “numeric” data type, field width is long number; for “text” data type, field width is 255 characters long.
4. Validation rules and validation text can be applied on given field.
5. Query processing can be done easily because of powerful query processor engine.
6. RDBMS offers data consistency across the tables using referential integrity.
7. If required the data from MS-Access can be easily exported to MS-Excel effortlessly.

5. Problems/Limitation of RDBMS data model

1. There is physical limit to how “big” the database can be. Therefore because of the physical limit, as MS-Access databases grow, the retrieval of results will become slow.
2. In MS-Access complex queries give inconsistent results.
3. MS-Access does not support object data type like audio/video.

6. Conclusion

From the above case study we conclude that:

1. MS-Access data model is better fit in storing response of questionnaire data as compare to MS-Excel datasheet.
2. Use of loss-less join decomposition guarantees that the originality of the response sheet is maintained.
3. Storing data in relational data model and defining relations among the table will produce the transactional data which proves to be easy to handle.
4. Referential integrity represents good relationship between relations, where it is not supported by MS-Excel.
5. Using MS-Access query processor, fast query results are generated which is not supported by MS-Excel.
6. To define data types for each column or fields with their size this is not possible in case of MS-Excel.
7. Validation rules are also applied on each field in MS-Access, but not in case of MS-Excel datasheet.

7. References

- [1] E. F. Codd, "Relational Database: A Practical Foundation for Productivity," 1981 ACM Turing Award Lecture, Communications of the ACM 25(2), 1982, pp. 109-117
- [2] Vardi, M. Y. (1984). A Note on Lossless Database Decompositions, Inform. Process. Lett. 18,257-260
- [3] Aho, A. V., Beeri, C. and Ullman, J. D. (1979). "The Theory of Joins in Relational Databases", ACM Trans. Database Systems, 4, 297 ± 314.

- [4] R. Elmasri, S. Navathe, *Fundamentals of Database Systems, Fourth Edition*, Addison-Wesley, 2003
- [5] R. van der Lans, *Introduction to SQL: Mastering the Relational Database Language, Fourth Edition*, Addison-Wesley, 2006

8. Reference Sites:

[a]<http://spreadsheetpage.com>

[b][http://wiki.answers.com/Q/Total number of row and column of m.s.excel](http://wiki.answers.com/Q/Total_number_of_row_and_column_of_m.s.excel)

IJERT