

Business Dimensional Modeling: A New Visual Approach

Qazi Muhammed Arfeen and Muhammad Nadeem
SZABIST
Karachi, Pakistan

Abstract: A diagramming technique which reduces the ever increasing complexity and number of elements included in the dimensional model called the Business Dimensional Modeling is proposed with some major changes which will visually aid its designers, modelers, decision makers to help understand the overall system being developed well. It also helps Business users to well understand what the needs are there and how their needs will be transferred to develop the system.

Business Dimensional Models help in capturing the business professionals' perspective of their business. This model is separately used in the design of a data warehouse in order to ensure support for the different types of reporting and analysis that need to be performed and at the same time is understandable to the end user.

Keywords: Business Dimensional Modelling

1. INTRODUCTION

There are generally two kinds of operations that can be performed by relational databases. One type of databases are called Transactions Intensive Systems (Online Transaction Processing) known as OLTP which are used for transaction (insert, update, delete intensive) purpose only and the other type of databases called Query Intensive Systems (Online Analytical Processing) OLAP which are used to query data for searching means (query intensive).

Data warehousing is a relational database that is designed to query and analysis rather than for transaction processing. These kinds of databases usually contain historical data which is derived from many sources mainly transactional data. [1]

1.1 Types of Modelling Techniques

There are mainly two types of modeling techniques: ER Modeling and Dimensional Modeling. ER modeling is normally suitable for Online Transaction Processing (OLTP) databases which have three basic graphical symbols used to elaborate the diagram namely Entity, Relationships and Attributes. The other kind of

modeling technique Dimensional Modeling is preferred when we are working on data warehousing projects, which is the prime focus of this research study as well. They use rather different notations to help its user prepare a better model best suited for any given data warehouse project. These are Facts, Dimensions, and Measures (Variables)

1.2 DETAILS OF DIMENSIONAL MODELING TECHNIQUES

Before discussing different schemas, each notation is discussed briefly.

Facts: A fact is a collection of related items, consisting of measures and context data. Each fact typically represents a business item, a business transaction, or an event that can be used in analyzing the business or business processes.

Dimensions: A dimension is a collection of members or units of the same types of view. Dimensions determine the contextual background for the facts.

Measures: A measure is a numeric attribute of a fact, representing the performance or behavior of the business relative to the dimension. A measure can be determined by combination of members of the dimensions and is located on facts.

Dimensional Modeling is the only viable technique for databases that are designed to support end-user queries in a data warehouse. When it comes to designing any model, data warehouse modelers normally choose one of the two schemas: Star Schema or Snowflake Schema.

Star Schema: The star schema is the simplest data warehouse schema. It is called a star schema because the diagram resembles a star, with points radiating from a center. This schema is considered to be the simplest of the data warehousing schemas [2]. The reason for calling it a Star is due to the structure itself. In the design of a Star Schema, the center of the star consists of one or more fact tables and the points of the star are the dimension tables as shown in Figure 1.

Snowflake Schema: Snowflake gets its name because the diagram resembles a snowflake in shape. A snowflake schema is a way of arranging tables in a relational database such that the entity relationship diagram ends up looking like a snowflake. At the center of the schema is the fact tables which are connected to multiple dimension tables but with a change, in a snowflake schema, the dimension tables are normalized. A snowflake schema is a way of arranging tables in a relational database such that the entity relationship diagram resembles a snowflake in shape. At the center of the schema are the fact tables which are connected to multiple dimension tables. Thus snowflake simplifies to a star schema when relatively few dimensions are used [3].

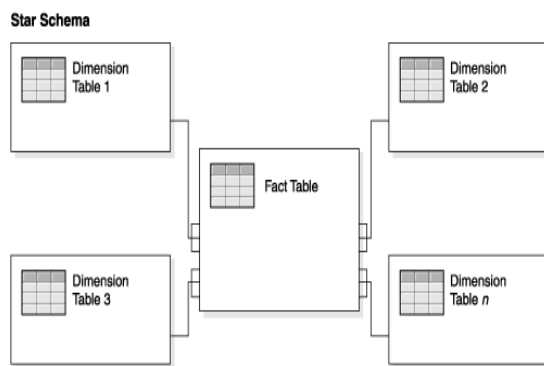


Figure 1. Star Schema

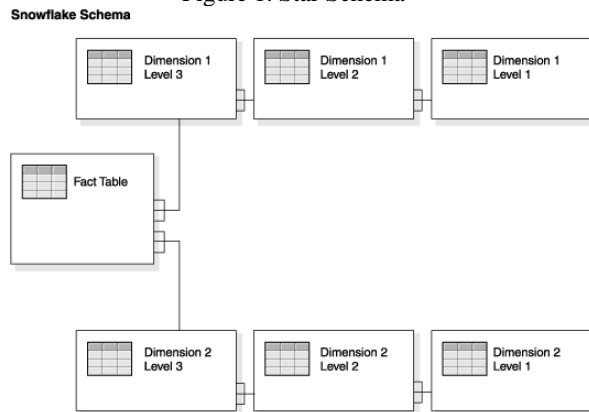


Figure 2. Snowflake Schema

1.3 A NEW DIMENSION: BUSINESS DIMENSION

Laura Reeves, coauthor of “The Data Warehouse Lifecycle Toolkit”, and a Cofounder of StarSoft Solutions has developed an approach to dimensional modeling and a diagramming technique. This dimensional modeling and diagramming technique will help to develop the model itself, facilitate discussion, foster communication, and manage expectations. This diagramming technique and dimensional modeling notations is known as Business Dimensional Modeling.

2. BUSINESS DIMENSIONAL MODELLING

The Business Dimensional Model helps capturing the business professionals’ perspective of their business. If asked directly, end users would not be able to define their business dimensional model. This model is specifically used in the design of a data warehouse in order to ensure support for the different types of reporting and analysis that need to be performed and at the same time understandable to the end user.

The idea to propose this model along with business community is that appropriate details are collected and diagrammed correctly, end the user can not only verify the model but may help them understand it and propose any future requirements. While using this model it is important to use business labels and not internal table or column names.

The main purpose of the business dimensional model is to separate the business discussion from technical implementation choices. The idea is to develop the dimensional model in a format that will be useful to the business, yet detailed enough to support table design decisions [4].

2.1 BUSINESS DIMENSIONAL MODELING NOTATION

Business Dimension Diagram (BDM): Figure 3 shows an example of a BDM. The highlighted rectangle “DAY” represents the lowest level of detail possible. The arrows show one-to-many relationships in a manner that is more comfortable for the users. Rectangles marked as “Attributes” and placed in a particular fashion that helps user understand drill up and drill down options and at the same time define hierarchies.

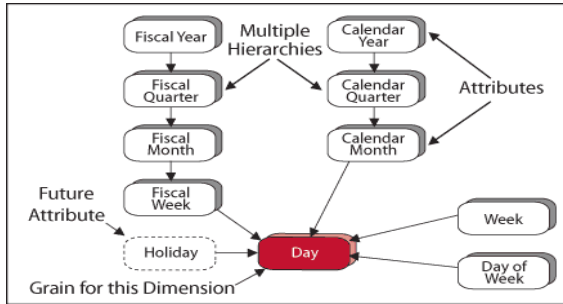


Figure 3. Example Business Dimension Diagram

If there is an element that the business community has requested that will not be a part of the initial implementation (because that data source is not being used or perhaps it is not captured anywhere), it can be noted in the diagram with a dotted box. This accomplishes two things:

- i. It notes where one thinks it should fit in the model as a starting place for the future.
- ii. It reminds all that it is a future element [4].

Business Fact Groups: The second part of the BDM is the fact groups. The fact groups are where the specifics about each fact are defined. The word "table" is purposely not used here because this scheme may implement the model using a multidimensional database technology. The business fact group represents facts that have identical dimensionality and grain [4].

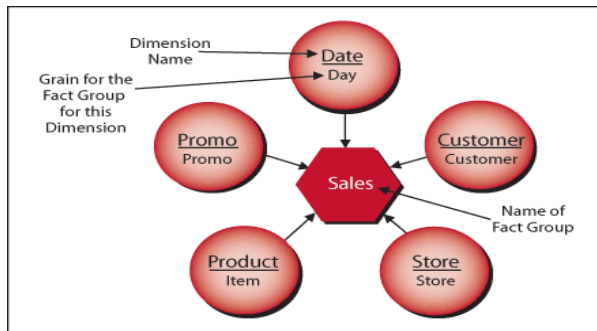


Figure 4. Example Business Fact Group

Figure 4 shows the name of the fact group at the center of the diagram. Each dimension that applies is included in the diagram and the specific grain is noted for each dimension [3].

2.2. PROPOSAL FOR BUSINESS FACT GROUPS

This gives an overall view of what needs to be developed and already well understood by the business community but as started by Laura Reese herself, "We are visual beings, and we remember things not just by the words, but often by the placement on a page." There is a limitation to this diagramming method, it only reminds about what is being developed, but not about that which needs to be done more to see the completion of the system. Recall the dotted rectangle in the BDM which explains that it is a future development task due to any reason there may be, if the dotted rectangle is introduced in the Data circle above, as in Figure 5.

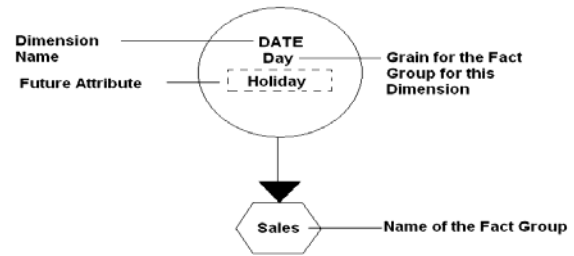


Figure 5. Future Attribute in BDM

It can be seen where this is actually leading to. As it is known that Sales in this diagram represents a Fact Group and there may be multiple Fact Groups in a system (Data Marts) by visually looking at the above diagram one can easily estimate what extra needs to be done to complete the system.

This addition to the diagram has another hidden feature. By properly versioning the diagrams and documentations to track the progress of work being done, once the HOLIDAY future attribute is developed, it can be removed from the diagram, and by comparing the two versions of the same diagram, the business community can also keep track of the progress that is being made at any given time.

2.3. PROPOSAL FOR BUSINESS DIMENSION DIAGRAM

The Red Book published by IBM with the title of the book "Dimensional Modeling in a Business Intelligence Environment" states that for applying any of the three kinds of changes to *slowly changing dimensions* "Such decisions cannot be made by the dimensional modeler. It is critical to involve the business users to identify

how they would like to see the changes that happen in various business entities.” The question is on where and how to involve the business community at that later stage with having said all about the terminology misunderstandings some time unavailability of required persons at the required hours.

It is proposed to introduce a factor in the Business Dimension Diagram that can identify the slowly changing values early in the cycle, and not only to identify the attributes which may be slowly changing but also take the advantage of the sessions being held to design BDM to identify which type of solution should be applied to any of those slowly changing attributes, as in Figure 6.

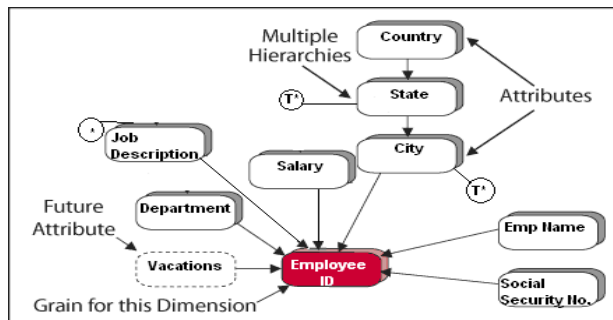


Figure 6. Proposed factor in BDM

The Employee Dimension is taken into consideration, the diagram remains almost the same with two visually minor but big changes. Additional notations are introduced:

- a. Letter T
- b. After the letter T an asterisk (*)
- c. Encapsulated within a Circle
- d. A Straight Line attaching the circle with the attribute
- e. A single asterisk (*)
- f. Encapsulated with a Circle

T* represents the three types of solutions that can be applied to the identified slowly changing values replacing the (*) with the appropriate number like for example T1, T2 or T3.

The standalone asterisk mark (*) identifies that this attribute is identified as a fast changing attribute and as we know that fast changing attributes are converted to Dimensions themselves, only gives an impression that the modeler will take care of all the asterisk (*) issues. It will also help draw a line between the dimension model and Business Dimension Model. Business

Dimension Modeling does not go in details of how things will be accomplished but what needs to be done is the main focus and this is the reason for not elaborating the diagram further instead keeping it simple with a confidence that the modeler will do his job best when it comes to implementation. The modeler will also be confident enough that all business as well as design decisions has already being made.

3. CONCLUSION

Dimensional Modeling importance in data warehousing can not be neglected. It is the de-facto industry standard which keeps on improving when it was first evolved. There are multiple technical issues where this dimensional model can be improved but with the Business Dimensional Modeling technique in place we can assure that the Business Community is well involved in the process of developing a system, which not only helps get a better picture of the requirements before going into the implementation phase but also increases the confidence in the modelers mind that there is less to be changed. This paper has only focused on the Visual aspect of this Dimensional Modeling Technique but there is still a lot of work that can be done on the documentation of such standards, version controlling, the processing of translating this Dimensional Model into Logical Layer.

REFERENCES

- [1] <http://www.lc.leidenuniv.nl/awcourse/oracle/server.920/a96520/concept.htm>, last accessed 7/12/2007.
- [2] http://download-uk.oracle.com/docs/cd/B19306_01/server.102/b14223/schemas.htm, last accessed 7/12/2007.
- [3] http://en.wikipedia.org/wiki/Snowflake_schema, last accessed 7/12/2007.
- [4] <http://www.dmreview.com/issues/20031101/7618-1.html>, last accessed 7/12/2007.