Grouping Attribute Values in a Dimensional Table Design: A Customized Approach for Teaching Business Analytical Applications Using Microsoft Business Intelligence Tools

Chang Liu[1], Charles Downing[2]

OM&IS Department, College of Business, Northern Illinois University

DeKalb, IL 60115, United States of America

[1]cliu@niu.edu; [2]cdowning@niu.edu

Abstract

Customization and flexibility are needed for drill-down and drill-through data analysis in a business intelligence cube design. This paper presents an easy way of customized grouping dimensional attribute members in Microsoft SQL Server Analysis Services. The approach described in this paper indicates that such customization yields quicker learning and more complete understanding of a cube design by students.

Keywords

Business Intelligence; Analytics; Cube; Dimension; Grouping; Customization; SQL Server; Analysis Services

Introduction

Business Intelligence (BI) is a broad category of technologies, applications, and processes for gathering, storing, accessing, and analyzing enterprise data to help its users make better decisions. BI benefits for organizations include improvement of operationalperformance, customer service, and identification of new business opportunities (Sabherwal and Becerra-Fernandez, 2011). From 2008 to 2011, BI and business analytics have been consistently ranked as top 5 technology priorities each year by CIOs in Gartner Inc.'s annual survey on IT executives. For 2012, BI was the highest rankedtechnology to help decision makers accessing organizational data and tuning raw data into actionable information (Gartner Inc., 2012). Therefore, there is a strong appeal to teach students to develop a BI solution to analyze enterprise data in a BI design course.

The Hands-on Environment

The lesson described in this paper was based on a semester long graduate level business intelligence applications and tools course taught in the spring of 2012 at a large, Midwestern university. All twenty-four students enrolled in this course are MIS majors. This was the second time that the school offered a BI design course in the MIS curriculum. The lead author has developed a set of ten labs for this course using Microsoft SQL Server Business Intelligence Development Studio including Analysis Services (SSAS) and Reporting Services (SSRS). Microsoft BI tools are one of the top popular academic alliance programs that are currently adopted by most universities offering BI in their curriculums (Wixom, Ariyachandra, Goul, Gray, Kulkarni, and Phillips-Wren, 2011). Other popular BI software packages include: IBM Cognos, SAP, Oracle, SAS on Demand for Academics, etc. Microsoft SSAS supports online analytical processing (OLAP) by letting users design, create, and manage multi-dimensional structures or cubes that contain data aggregated from data sources. Figure 1 shows the students' hands-on assignments

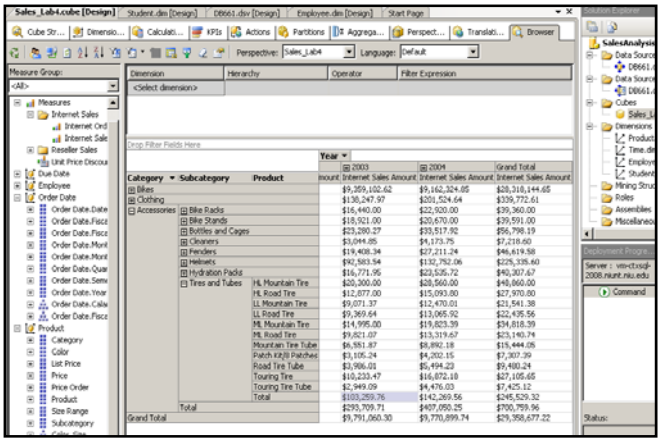

FIG. 1 MULTI-DIMENSIONAL STRUCTURE FOR A SALES CUBE DESIGN

from lab1 through lab 4 to complete a sales cube design to analyze order amount information based on calendar year, category, subcategory, and product dimensions. The data used in the hands-on exercises came from AdventureWorksDW, a sample data warehouse provided by Microsoft SSAS for BI application design scenarios. All Microsoft BI tools were installed in a virtual cloud computing environment called citrix, or Anywhere Lab as shown in Figure 2. This hands-on environment provides the students with access to the required software from anywhere and everywhere. They did not need to come to a computer lab to work on their lab assignments. In addition, the instructor did not need to reserve a computer lab and be limited by the number of computers in a lab. The students, however, were required to bring their own laptops to the class which had wireless access.

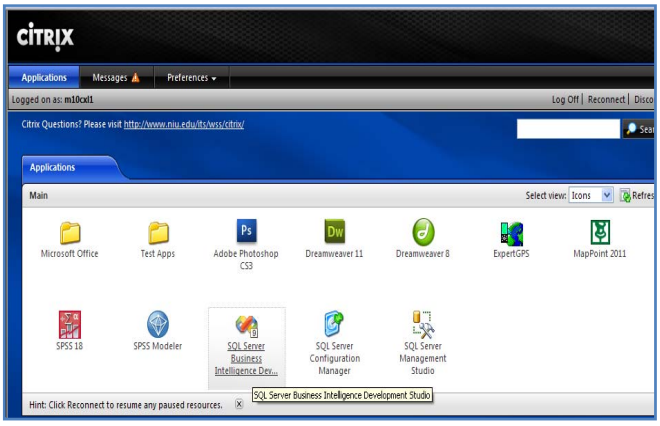

FIG. 2 THE HANDS-ON ENVIRONMENT FOR ACCESSING BI SOFTWARE

The Teaching Approach

This teaching tip describes lab assignment 5, which provides students with experience in customizing grouping dimensional attribute values in Microsoft SSAS. In particular, the lab requires students to split "list price" values of products into customized ranges for browsing. The purpose of the lab is to allow students to customize grouping values for adimensional attribute, i.e., price range, in this example so that they could analyze data with customizedgrouping criteria with flxeibility.

In general, students might expect that Microsoft SSAS would provide a built-in feature for grouping attribute values. It is indeed that values of a dimensional attribute can be grouped through a process called discretization in Microsoft SSAS. Figure 3 shows the discretization property settings for grouping options.

"DiscretizationMethod" Setting	Description
None	Displays the members.
Automatic	Selects the method that best represents the data: either the EqualAreas method or the Clusters method.
EqualAreas	Tries to divide the members in the attribute into groups that contain an equal number of members.
Clusters	Tries to divide the members in the attribute into groups by sampling the training data, initializing to a number of random points, and running several iterations of the Expectation-Maximization (EM) clustering algorithm.

FIG. 3 DISCRETIZATION PROPERTY SETTING

(Source: Microsoft TechNet)

The students and some professional Microsoft SSAS designers have reported that the discretization method of SQL Analysis Services is hard to understand and fails to provide an expected direct grouping result. For example, during a lab exercise, students were instructed to use the discretization method to group the "list price" information into 10 different ranges. Figure 4 shows the property settings for the dimension attribute of "list price" using the "EqualAreas" discretization method for 10 the Discretization Bucket Count value. Figure 5 presents the result of the "list price" displayed in the SSAS browser.

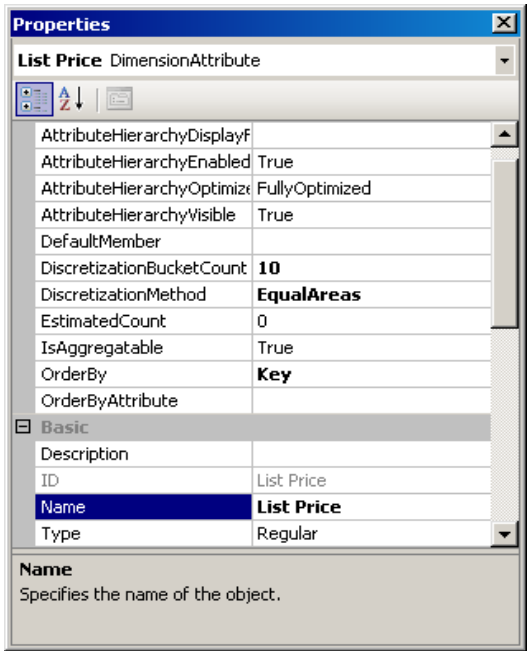

FIG. 4 GROUPING DIMENSION ATTRIBUTE USING DISCRETIZATION METHOD

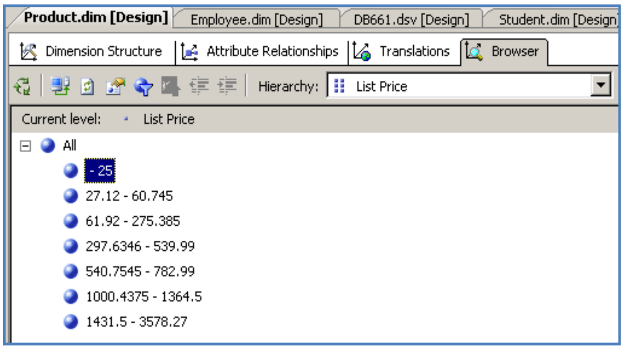

FIG. 5 THE "LIST PRICE" DISPLAY BASED ON FIGURE 4 SETTINGS

Students quickly realized that although the Discretization Bucket Count property specifies the number of buckets or ranges required by a user or a designer under a selected discretization method, i.e., Automatic, Cluster, or EqualAreas, the number of final buckets or ranges may NOT be the same as the Discretization Bucket Count value. The number of groups does not generally meet the specified value in that property. This is because those algorithms can only try to approximate number of buckets to the specified Discretization Bucket Count property value (Janus and Fouche, 2010).

The students also experienced similar results of not getting the exact numbers of grouping they specified when they exercised "Authomatic" and "Cluster" discretization methods. As a matter of fact, results of using "Authomatic" and "Cluster" discretization methods are quite similar to the one displayed in Figure 5.

StandardCost	FinishedGoodsFlag	Color	SafetyStockLevel	ReorderPoint	ListPrice
884.7083	☑	Red	100	75	1457.99
884.7083	☑	Red	100	75	1457.99
884.7083	☑	Red	100	75	1457.99
884.7083	☑	Red	100	75	1457.99
884.7083	☑	Red	100	75	1457.99
413.1463	☑	Red	100	75	699.0982
486.7066	☑	Red	100	75	782.99
413.1463	☑	Red	100	75	699.0982
486.7066	☑	Red	100	75	782.99
413.1463	☑	Red	100	75	699.0982
486.7066	☑	Red	100	75	782.99
413.1463	☑	Red	100	75	699.0982
486.7066	☑	Red	100	75	782.99
413.1463	☑	Red	100	75	699.0982
486.7066	☑	Red	100	75	782.99
413.1463	☑	Red	100	75	699.0982
486.7066	☑	Red	100	75	782.99
413.1463	☑	Black	100	75	699.0982
486.7066	☑	Black	100	75	782.99
413.1463	☑	Black	100	75	699.0982

FIG. 6 LIST PRICE VALUES IN A DIMENSIONAL TABLE

All students were disappointed and felt frustrated at this point since they could not build a "list price" range that they could have easily manipulated. The rangecreated by SSAS obviously did not meet their expectations. The instructor then asked the students toexamine the data source view of the "list price" attribute again in the product dimensional table as shown in Figure 6.

Students were challenged to come up with a solution to create a new attribute in the same data source viewof SSAS to contain customizedgroup valuesfor the list price. They were excited to learn that there is a better approach they could use to group dimensional attribute values. This better approach is to create a new named column which can be derived from existing values. Using this approach,a user or adesigner of a BI application can create concatenated attributes (e.g., full customer name consisting of first name, middle name, and last name), adjusted attributevalues (e.g., service length from date of hire), and/or various calculated values.

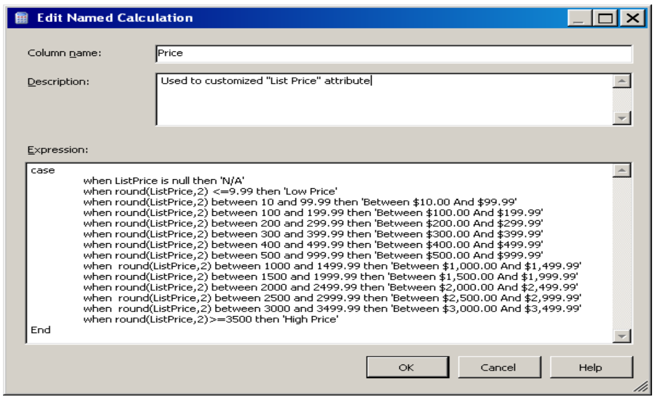

FIG. 7 CREATING A "PRICE" ATTRIBUTE FOR GROUPING "LIST PRICE" VALUES

Figure 7 shows the creation of a new attribute named "Price" using the CASE Expressions. The CASE Expressions are a flexible wayto modify the representation of data in SQL Server database application environment. The statements in Figure 7 allow the students to specify the grouping options based on their needs. The "round" function used here in conjunction with the CASE Expressions was to round the "list price" field to the number of decimals specified to ensure accurate calculations. Figure 8 displays the group values of "Price" generated from this approach.

ListPrice	Size	StartDate	EndDate	Status	Price
1457.99	58	2001-07-01 00:00:00Z	2002-06-3...		Between $1,000.00 And $1,499.99
1457.99	60	2001-07-01 00:00:00Z	2002-06-3...		Between $1,000.00 And $1,499.99
1457.99	44	2001-07-01 00:00:00Z	2002-06-3...		Between $1,000.00 And $1,499.99
1457.99	48	2001-07-01 00:00:00Z	2002-06-3...		Between $1,000.00 And $1,499.99
1457.99	52	2001-07-01 00:00:00Z	2002-06-3...		Between $1,000.00 And $1,499.99
699.0982	58	2001-07-01 00:00:00Z	2002-06-3...		Between $500.00 And $999.99
782.99	58	2002-07-01 00:00:00Z	2003-06-3...		Between $500.00 And $999.99
699.0982	60	2001-07-01 00:00:00Z	2002-06-3...		Between $500.00 And $999.99
782.99	60	2002-07-01 00:00:00Z	2003-06-3...		Between $500.00 And $999.99
699.0982	62	2001-07-01 00:00:00Z	2002-06-3...		Between $500.00 And $999.99
782.99	62	2002-07-01 00:00:00Z	2003-06-3...		Between $500.00 And $999.99
699.0982	44	2001-07-01 00:00:00Z	2002-06-3...		Between $500.00 And $999.99
782.99	44	2002-07-01 00:00:00Z	2003-06-3...		Between $500.00 And $999.99
699.0982	48	2001-07-01 00:00:00Z	2002-06-3...		Between $500.00 And $999.99
782.99	48	2002-07-01 00:00:00Z	2003-06-3...		Between $500.00 And $999.99
699.0982	52	2001-07-01 00:00:00Z	2002-06-3...		Between $500.00 And $999.99
782.99	52	2002-07-01 00:00:00Z	2003-06-3...		Between $500.00 And $999.99
699.0982	58	2001-07-01 00:00:00Z	2002-06-3...		Between $500.00 And $999.99
782.99	58	2002-07-01 00:00:00Z	2003-06-3...		Between $500.00 And $999.99
699.0982	60	2001-07-01 00:00:00Z	2002-06-3...		Between $500.00 And $999.99

FIG. 8 PRICE VALUES IN A DIMENSIONAL TABLE

Next, the students learned to add the newly created"Price" attribute to the product dimension. They completed this task by simply dragging the "Price" attribute from the data source view of Microsoft SSAS into the attribute panel in the product dimension structure design. In addition, they ensured the right display order for the newly created "Price" dimensional attribute. Figure 9 presents the result of "Price" display in the SSAS browser.

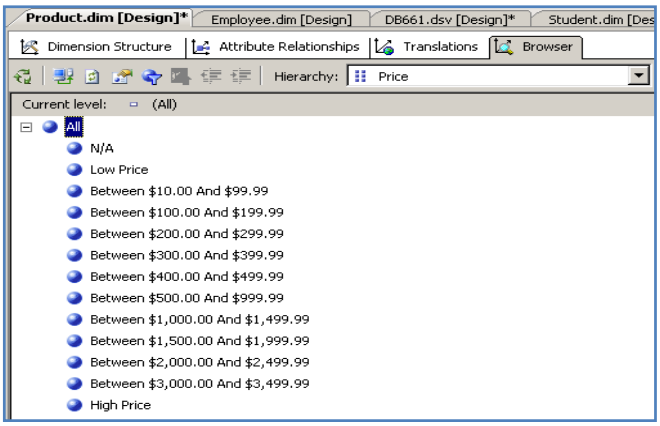

FIG. 9 "THE PRICE" DISPLAY BASED ON FIGURE 7 SETTINGS

Discussion

This paper explained a customized way to group dimensional attribute values in Microsoft SSAS. Based on feedback from the students, the authors believe that the approach described in this paper indicates such customization yields quicker learning and more complete understanding of a cube design by students.

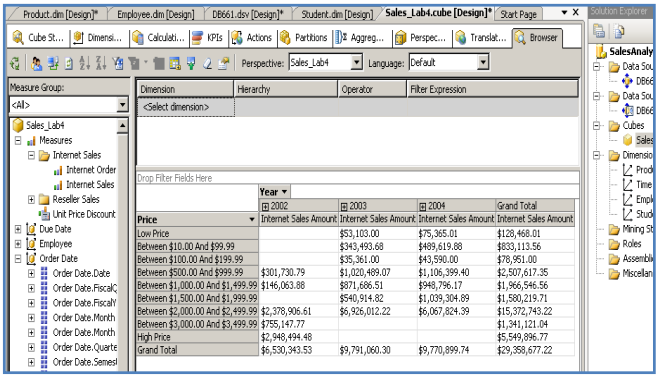

FIG. 10 A CUBE DESIGN WITH PRICE AND YEAR DIMENSIONS

For example, immediately after the above lab exercise, the students were asked in class to create a cube to answer a question of sales amounts of products priced between $3,000 and $3,499.99 in the year of 2002. Results indicated that ALL students could correctly set up a cube as shown in Figure 10 without any additional help from the instructor. Moreover, a homework assignment was given to ask the students to re-group the "list price" options with $100 increments. Again, students finished the assignment easily by using the approach explained in this paper. The students felt strongly from their course feedback that the approach described in this paper provides an easy and a flexible way to customize grouping attribute values in a dimensional table design. As a result of lessons learned, the authors would like to provide the following suggestions for other faculty to teach a business intelligence application class using Microsoft BI tools:

- Prerequisites for students should include the knowledge of database concepts and SQL statements. Students can learn quickly and in more depth if they have these skills.
- Prior to implmenting a cube design, it is important for students to fully understand the star-schema based dimensional modeling technique. Building a one-to-many relationship from each dimension table to the central fact table is the key for students to understand how to build a measurement based on a specified dimension.
- It would be very helpful to ask students to build cubes to analyze measurement values such as order amount, cost, profit, etc., for each dimension created. It helps students improve their skills for analyzing data.
- Students should be challenged to further explore using SSAS to design multi level groupingoptions. The example used in this paper employs only one level of grouping. However, it can be easily extended to include multi-level customized groupings with building a hierarchy in SSAS. This will lead to a better understanding of dimensional structure design for drill-down and drill-through in a cube.

REFERENCES

Gartner Inc. (2012). http:// www. gartner. com/ it/ page. jsp?id=1897514.

Janus, P. and Fouche, G. "Pro SQL Server 2008 Analysis Services."Apress, New York, NY.2011

Microsoft TechNet, "Grouping Attribute Members (Discretization)", Accessed March 12, 2012. http: // technet.microsoft.com/en-us/ library/ ms 174810 (d= printer).aspx

Sabherwal, R. and Becerra-Fernandez, "I. Business

Intelligence, Practices, Technologies, and Management.", John Wiley & Sons, Inc., New Jersey, NJ. 2011

Wixom, B.; Ariyachandra, T.; Goul, M.; Gray, P.; Kulkarni, U.; and Phillips-Wren, G. (2011). "The Current State of Business Intelligence in Academia," Communications of the Association for Information Systems: Vol. 2011, No. 29, 299-312.

Chang Liu, DBA, is a Professor of MIS and Department Chair at Northern Illinois University. He received his doctor of business administration from Mississippi State University in 1997. He teaches database and electronic commerce courses and received MIS Graduate Teaching Awards in 2002, 2003, and 2006. His research works published at Information & Management (I&M), International Journal of Electronic Commerce and Business Media, Journal of Global Information Management, Journal of Internet Research, Journal of Computer Information Systems, Mid-American Journal of Business, International Journal of Mobile Communications, Journal of International Technology and Information Management, and Journal of Informatics Education Research.

Charles E. Downing, Ph.D, is the Presidential Teaching Professor in the Operations Management and Information Systems Department of the College of Business at Northern Illinois University. He received his Ph.D. from Northwestern University. He researches, teaches and consults in topics such as Systems Development, Systems Analysis and Design, Information Technology as a business process integrator (Supply Chain and portals), and Business Intelligence. His articles have appeared in Decision Sciences, Communications of the ACM, Decision Support Systems, Information & Management, Journal of Global Information Management, The Journal Information Technology Management, Information Systems Management, Journal of End UserComputing, Journal of Educational Technology Systems, European Journal of Operational Research, and Journal of the Operation Research Societ.