Лекция 10

Тема « Cube Design »

Cube Design

In Chapter 5 you learned to create dimensions using the Dimension Wizard and to refine and enhance dimensions using the Dimension Designer. Dimensions eventually need to be part of your cube for you to analyze data across various dimension members. In previous chapters, you read about the Unified Dimensional Model (UDM). Now, prepare yourself for significantly more detail because all the fact and dimension tables you see when you ' re looking at a DSV in the Cube Designer comprise the UDM. Yes, the UDM is more than a multiple data - source cube on steroids, but to make it as clear as possible, think of the UDM as a cube for now. In this chapter you learn how to create cubes using the Cube Wizard and enhance the cube using the Cube Designer. You learn to add calculations to your cube that facilitate effective data analysis followed by analyzing the cube data itself in the Cube Designer.

The Unified Dimensional Model

To generate profits for a business, key strategic decisions need to be made based on likely factors such as having the right business model, targeting the right consumer group, pricing the product correctly, and marketing through optimal channels. To make the right decisions and achieve targeted growth you need to analyze data. The data can be past sales, expected sales, or even information from competitors. The phrase " Knowledge is power " is very fitting here because in the world of business, analyzing and comparing current sales against the expected sales helps executives make decisions directly aligned with the goals of the company. Such sales information is typically stored in a distributed fashion and must be collected from various sources. Executives making the business decisions typically do not have the capability to access the raw sales data spread across various locations and subsequently optimize it for their use. The decision - makers typically rely on the data that has already been aggregated into a form that is easy to understand and that facilitates the decision - making process. Presenting aggregated data to the decision - makers quickly is a key challenge for business intelligence providers. Analysis Services 2008 enables you to design a model that bridges the gap between the raw data and the information content that can be used for making business decisions. This model is called the Unified Dimensional Model (UDM).

The UDM is central to your Analysis Services database architecture. UDM is your friend because it helps you narrow the gap between end users and the data they need. Analysis Services provides you with features that help you design a model that will serve the needs of end users. UDM, as the name suggests, provides you with a way to bring data from multiple heterogeneous sources into a single model. The UDM buffers you from the difficulties of managing the integration of various data sources so you can build your model easily. It provides you with the best of the OLAP and relational worlds, exposing rich data and metadata for exploration and analysis.

Figure 6 - 1 shows the architecture of the Unified Dimensional Model that is implemented in Analysis Services 2008. As shown in the figure, the UDM helps you to integrate data from various data sources such as Oracle, SQL Server, DB2, Teradata, and flat files into a single model that merges the underlying schemas into a single schema. The end users do not necessarily have to view the entire schema of the UDM. Instead, they can view sections of the UDM relevant to their needs through the functionality provided by Analysis Services 2008 called *perspectives* .

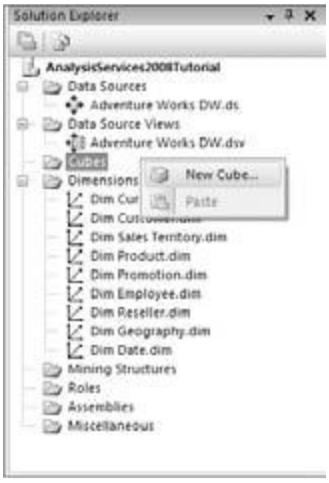
In the OLAP world, data analyzed by end users is often historical data that might be a few days, months, or even years old. However, the responses to the OLAP queries are typically returned within a few seconds. In the relational world the end users have instant access to the raw data but the responses to queries can take much longer, on the order of minutes. As mentioned earlier, the UDM merges the best of both the OLAP and relational worlds and provides the end users with real - time data with the query performance of the OLAP world. The UDM is able to provide the query performance of the OLAP world with the help of a feature in Analysis Services 2008 that creates a cache of the relational data source that also aggregates the data into an Analysis Services instance. During the time the cache is being built, the UDM retrieves the data directly from the data sources. As soon as the cache is available, the results are retrieved from the cache in response to relevant queries. Whenever there is a change in the underlying data source, the UDM receives a notification and appropriate updates are made to the cache based on the settings defined for cache updates.

The UDM also provides rich, high - end analytic support through which complex business calculations can be exploited. Such complex calculations can be extremely difficult to formulate in the relational world at the data - source level. Even if such calculations are defined on the relational data source, responses from OLAP - style queries against the relational data source might be really slow compared to responses from Analysis Services.
UDM natively interfaces to end - user clients through the XML for Analysis (XMLA) standard, which allows client tools to retrieve data from Analysis Services. Client tools such as Office Web Components (OWC) and Excel pivot tables allow the end users to create ad - hoc queries for data analysis. In addition, the UDM supports rich analytic features such as Key Performance Indicators (KPIs), Actions, and Translations that help surface the status of your business at any given time so that appropriate actions can be taken.
The UDM provides an efficient interface for detail - level reporting through dimension attributes that are common in the relational world. In addition to that, the UDM is easily understandable by a relational user. The ability to transform the UDM results into views that are helpful to end users and the ability to perform ad - hoc queries on data from high - level aggregations data to detail - level items make the UDM a powerful construct indeed. The UDM also allows you to design the model in the end user ' s language, which is needed in a global market.

Creating a Cube Using the Cube Wizard

Cubes are the principal objects of an OLAP database that help in data analysis. Cubes are multidimensional structures that are primarily composed of dimensions and facts. The data from a fact table that is stored within the cube for analysis are called *measures* . In Analysis Services 2008 you can store data from multiple fact tables within the same cube. In Chapter 2 you became familiar with the Cube Wizard and in this chapter you see more details of the Cube Wizard followed by refinements to your cube in the Cube Designer.
Similar to the Dimension Wizard you used in Chapter 5 , the Cube Wizard facilitates creation of cube objects from the DSV. For this exercise, you continue with the AnalysisServices2008Tutorial project you created in Chapter 5 , which contained the dimensions [Dim Geography], [Dim Employee], and [Dim Date]. To start with a clean slate, please delete the existing cube Adventure Works DW if it is still
there from Chapter 2 . To completely understand the functionality of the Cube Wizard, follow these steps to build a new cube:

1. Open the AnalysisServices2008Tutorial project from Chapter 5 . If the Adventure Works DW cube exists, delete the cube by right - clicking it in the Solution Explorer and selecting Delete.
2. Right - click the Cubes folder and select New Cube, as shown in Figure 6 - 2 . Click Next on the introduction page to proceed.

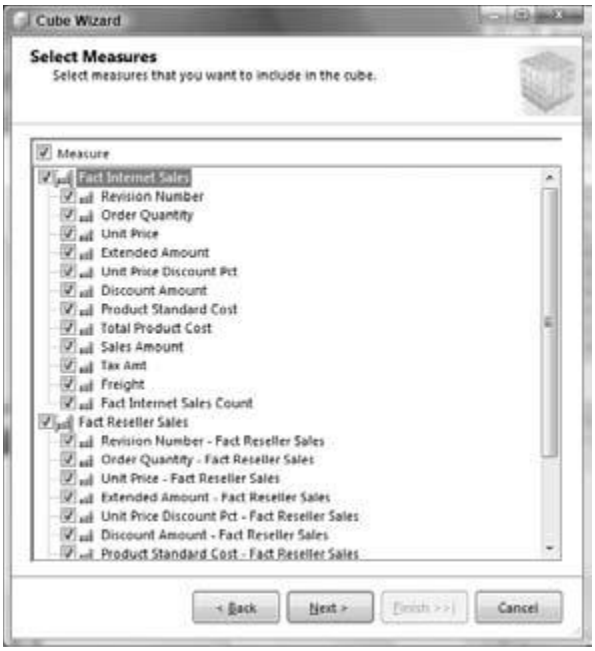


3. In the Select Creation Method page you have the option to build a cube from existing tables, create an empty cube, or create a cube based on a template and generate new tables in the data source. In this tutorial you build the cube from the existing tables in the Adventure Works DW data source. Click Next to proceed to the next step in the Cube Wizard.
4. The next page of the Cube Wizard is the Measure Group Tables selection page. If you have multiple DSVs, you need to select the DSV upon which you are creating the cube. In the current project you only have the Adventure Works DW DSV. You now must select one or more tables

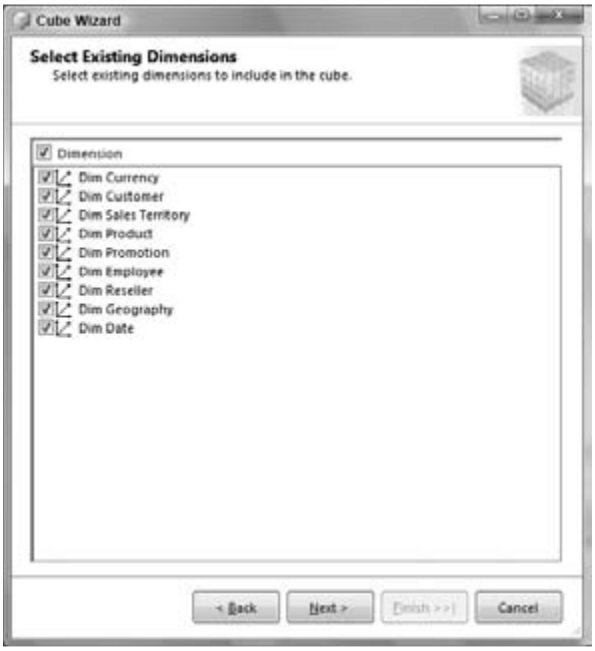
that will serve as fact tables for your Measure Group. The Suggest button on this screen can be used to have the Cube Wizard scan the DSV to detect the fact tables in the DSV and detect fact tables. Click the Suggest button to have the Cube Wizard automatically select potential Measure Group tables.
The Cube Wizard now scans the DSV to detect the fact and dimension tables in the DSV, automatically selects the candidate tables, and updates the list as shown in Figure 6 - 3 . Any table that has an outgoing relationship is identified as a candidate fact table, whereas a table that has an incoming relationship is detected as a dimension table.

5. You have the option to select or deselect a table as a fact or dimension table. The wizard has identified the FactInternetSales, FactResellerSales, and DimReseller tables as measure group tables. The DimReseller table was detected as a measure group table because there is an outgoing relationship from it. However, it will not be used as a measure group table in this example. Deselect the DimReseller table from being a fact table and click Next.

6. On the Select Measures page, the Cube Wizard shows all the columns from the fact tables that it detects as potential measures of the cube as shown in Figure 6 - 4 . The Cube Wizard does not select the primary and foreign keys in a table as measures. There is a one - to - one mapping between a column in the fact table and a measure in the cube. The Cube Wizard groups measures from a fact table under an object called a *measure group* . Therefore, by default, there will be one measure group for each fact table included in the cube. In the DSV you are using there are two fact tables, and therefore two measure groups named Fact Internet Sales and Fact Reseller Sales are created. You can select or deselect the measures you want to be part of the cube in this page. Use the default selection and click Next.

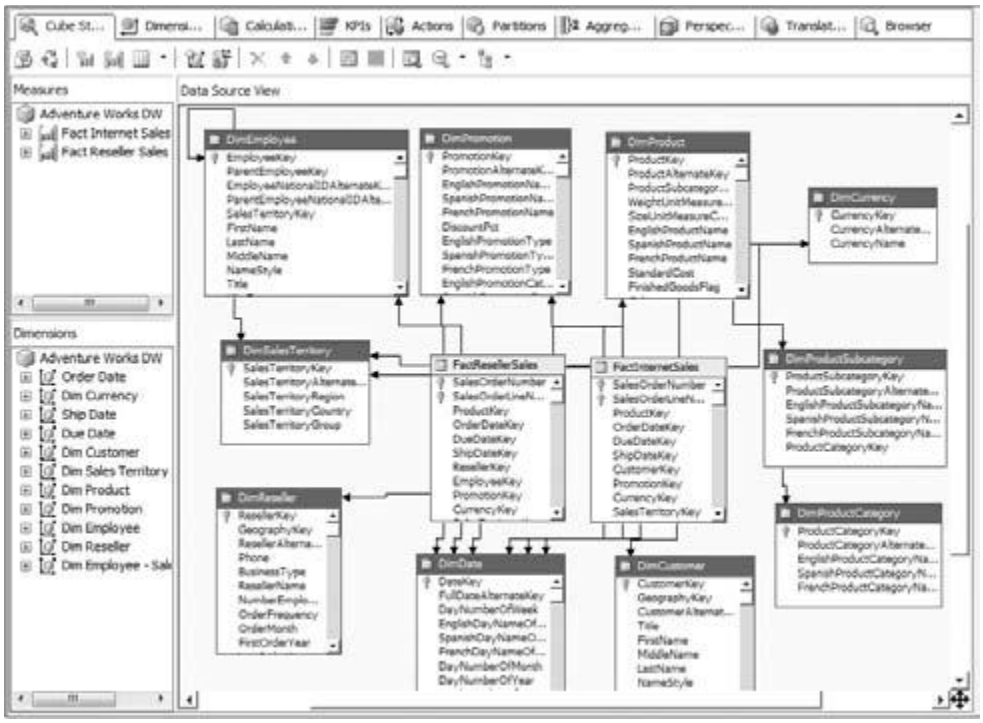


7. In the Select Existing Dimensions page (Figure 6 - 5), the Cube Wizard displays a list of all existing dimensions defined in the project. Accept the selection of all the dimensions and click Next.



9. In the final page of the Cube Wizard (shown in Figure 6 - 7) you can specify the name of the cube to be created and review the measure groups, measures, dimensions, attributes, and hierarchies. Use the default name Adventure Works DW suggested by the Cube Wizard and click Finish.

The Cube Wizard creates the cube after you click the Finish button. The created Adventure Works DW cube is opened within the Cube Designer as shown in Figure 6 - 8 . The Cube Designer contains several pages that help perform specific operations that will refine the initial cube created by the Cube Wizard. The default page is the Cube Structure page as shown in Figure 6 - 8 . In the Cube Structure page you can see three panes that show the Measures, Dimensions, and the Data Source View. The Data Source View contains all the tables that are part of the cube. Operations such as adding or deleting tables in the DSV and zooming in or out with the DSV Designer are possible within the cube Data Source View pane.
The Dimensions pane shows the dimensions that are part of the current cube and the Measures pane shows the cube ' s measure groups and measures. You can add or delete measures and dimensions in the Cube Structure view. The dimensions within the cube shown in the Dimensions pane are called cube dimensions. You can have multiple instances of the shared dimensions of the database within a cube. For example, the fact tables FactInternetSales, and FactResellerSales have a relationship with the Dim Date dimension through Order Date, Ship Date, and Due Date. Hence you can see three cube dimensions Ship Date, Due Date, and Order Date in the Dimensions pane, which refer to the Dim Date dimension.
A dimension such as Dim Date, which plays the role of three cube dimensions, is called a role playing dimension. You learn more about role playing dimensions in Chapters 8 and 9 . Within the Dimensions pane you can see the Hierarchies and Attributes of each dimension under separate folders when you expand each dimension.



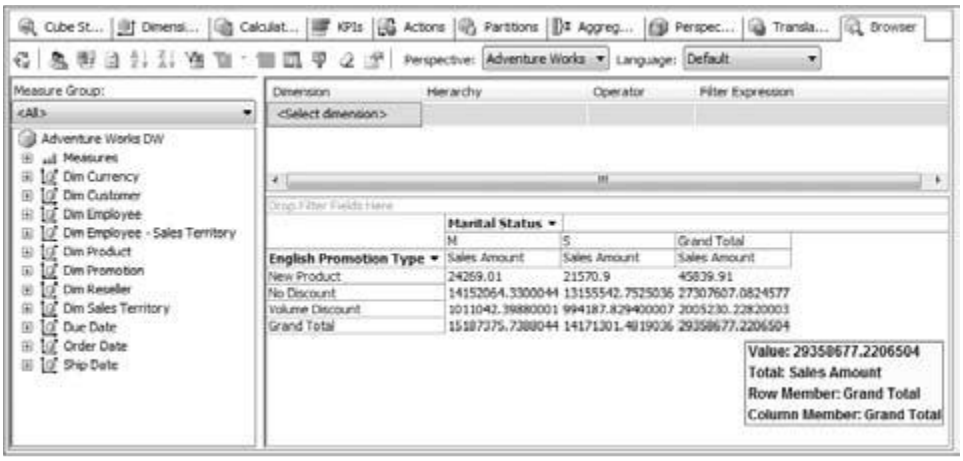
So far you have created an Analysis Services database containing the Adventure Works DW cube. You have to deploy the project to the Analysis Services instance so that you can analyze the data within the cube. You can deploy the project to the server in one of the following ways:

1. Select Debug Start Debugging from the menu.
2. Right - click the database AnalysisServices2008Tutorial in the Solution Explorer and select Deploy.
3. Right - click the Adventure Works DW cube and choose Process — from which you will first be prompted to deploy the project, followed by Process dialog to process the cube.
4. Use the shortcut key F5 to deploy and process.

When you deploy the project to the Analysis Services instance, BIDS sends an XMLA request containing object definitions to the instance of the Analysis Services server selected in the project. By default the Analysis Services project is deployed to the default instance of Analysis Services on your machine. The object definitions are of the cubes and dimensions you created. If you have installed Analysis Services 2008 as a named instance, you need to change the deployment server name. Then BIDS sends another request to process the objects within the database.

Browsing Cubes

Now that you deployed the cube to an Analysis Services instance, switch the BIDS Cube Designer view to the Browser page. In the Browser page you will see three panes: a Measure Group pane, a Filter pane, and a Data pane along with a toolbar as shown in Figure 6 - 9 .



The Measure Group pane, at the left - side of the Cube Browser, shows the measure groups (includes measures) and the dimensions that are related to each measure group (includes attributes and hierarchies) of the cube. The Data pane, at the bottom right, uses the Office Web Components (OWC) control used for analyzing multidimensional data. You can drag and drop hierarchies on the rows and/ or columns and measures in the data area to analyze the data. Indeed, you can have multiple hierarchies in the row and column areas. The OWC also has a filter area (above the column area) that can be used to filter the data being analyzed. You can slice the data you want to analyze based on specific members of a hierarchy.

The top - right pane also allows you to filter your multidimensional data for analysis, but it has additional filtering options above those provided by the OWC control. Whereas the filter area in OWC allows you to select or deselect members of a hierarchy, the Filter pane allows you to perform comparison operations like equal, not equal, contains, in, not in, begins with, range operations, and any MDX expression. With the help of the filter functionality in the Filter pane and OWC, you will be able to analyze your multidimensional data.

Suppose you want to analyze the Internet sales of products based on the promotions offered to customers and the marital status of those customers. First you would need to drag and drop [Dim Promotion].[English Promotion Type] from the Measure Group pane to the OWC rows area. You will learn the MDX statements that are generated by the OWC in this section. SQL Server Profiler in SQL Server 2008 has the ability to trace the MDX statements that are sent to Analysis Services instances. For more information on how to obtain traces please refer to the section on using SQL Server Profiler in Chapter 15 .

The first statement sent from OWC to Analysis Service instance is:

```
DROP VISUAL TOTALS for [Adventure Works DW]
```

When you drag and drop members of a hierarchy on a row (or column), or a measure in the fields area, OWC creates a row (or column) called Grand Total, which will automatically provide totals of the measure value for that hierarchy. By default the OWC in the Cube Browser shows you the totals of visible members of the hierarchy in the OWC. This is called Visual Totals because the total is calculated only for the members that are visible in the Browser. You have the option of disabling Visual Totals in the OWC. To do so, right - click OWC, select Commands and Options, click the Report tab, and select the " All Items (including hidden items) " option for the " Calculate totals based on " section. When the OWC calculates totals based on visual items only, it uses Visual Totals. The preceding MDX statement, DROP VISUAL TOTALS for [Adventure Works DW], removes references to visual totals from cells and clears

the memory cache for visual totals thereby ensuring that values for current members selected will be accurate. The full syntax for the Drop Visual Totals statement is shown in the following code snippet. You can optionally specify the MDX set upon which the visual totals will be dropped. If the MDX set expression is not specified, visual totals are dropped for the entire cube.

```
DROP VISUAL TOTALS FOR < cube name > [ON ' < MDX set expression > ']
```

The second statement sent to the Analysis Services server by OWC is:

```
CREATE SESSION
SET [Adventure Works DW].[ {7868741D-072F-458A-8A8D-EA3FED4A3FA7}
Pivot0Axis1Set0] AS
'
{
{ [Dim Promotion].[English Promotion Type].[All] },
AddCalculatedMembers([Dim Promotion].[English Promotion Type].[English
```

```
Promotion Type].MEMBERS)
}
‘
```

This statement creates a set called {7868741D - 072F - 458A - 8A8D - EA3FED4A3FA7}Pivot0Axis1Set0, which contains the members of the hierarchy [Dim Promotion].[English Promotion Type]. Because OWC creates the queries in an automated manner, it dynamically creates a session name that includes a descriptive name (Pivot0Axis1Set0) and the session ID. The Analysis Services instance allows you to create sets and other MDX objects within a specific scope. You can create objects within the scope of the database or within the scope of your connection. In the preceding statement OWC creates the set within the scope of the current session and the set will be available only for this specific session.

Finally, OWC sends the following query to retrieve and show the members of the hierarchy [Dim Promotion].[English Promotion Type]:

```
SELECT
NON EMPTY [{7868741D-072F-458A-8A8D-EA3FED4A3FA7}
Pivot0Axis1Set0]
DIMENSION PROPERTIES MEMBER_NAME, PARENT_UNIQUE_NAME ON COLUMNS
FROM [Adventure Works DW]
CELL PROPERTIES VALUE, FORMATTED_VALUE, FORE_COLOR, BACK_COLOR
```

Next, drag and drop [Dim Customer].[Marital Status] from the Measure Group pane to the OWC columns area. OWC now sends a series of MDX statements followed by an MDX query to retrieve the members on rows and columns. The following code shows the sequence of MDX statements sent by the OWC to the Analysis Services instance. First the OWC drops visual totals followed by creating two sets for the members of the hierarchies selected on rows and columns of the OWC. OWC then queries the members from the created sets and finally drops the earlier set Pivot0Axis1Set0 because OWC has created new sets for members on rows and columns of the OWC.

```
Drop visual totals for [Adventure Works DW]
CREATE SESSION
SET [Adventure Works DW].[{7868741D-072F-458A-8A8D-EA3FED4A3FA7}
Pivot1Axis0Set0] AS
‘
{
{ [Dim Customer].[Marital Status].[All] },
AddCalculatedMembers([Dim Customer].[ Marital Status].[ Marital
Status].MEMBERS)
}
‘
SET [Adventure Works DW].[{7868741D-072F-458A-8A8D-EA3FED4A3FA7}
Pivot1Axis1Set0] AS
‘
{
{ [Dim Promotion].[English Promotion Type].[All] },
AddCalculatedMembers([Dim Promotion].[English Promotion Type].[English
Promotion Type].MEMBERS)
}
‘
SELECT
NON EMPTY [{7868741D-072F-458A-8A8D-EA3FED4A3FA7}Pivot1Axis0Set0]
DIMENSION PROPERTIES MEMBER_NAME, PARENT_UNIQUE_NAME ON COLUMNS,
NON EMPTY [{7868741D-072F-458A-8A8D-EA3FED4A3FA7}Pivot1Axis1Set0]
DIMENSION PROPERTIES MEMBER_NAME, PARENT_UNIQUE_NAME ON ROWS
FROM [Adventure Works DW]
CELL PROPERTIES VALUE, FORMATTED_VALUE, FORE_COLOR, BACK_COLOR
DROP SET [Adventure Works DW].[{7868741D-072F-458A-8A8DEA3FED4A3FA7}
Pivot0Axis1Set0]
```

Finally, drag and drop the measure [Sales Amount] from the Fact Internet Sales measure group to the Drop Totals or Detail Fields Here area of the OWC pane. The OWC once again generates statements to drop existing sets and create new sets for members on rows and columns. These sets are used in the query to retrieve the measure data along with the properties of the cells. The cell properties returned by the instance of Analysis Services are used by the OWC to display values. From the query you can see that the properties of formatted values, foreground colors, and background colors are being retrieved by the OWC. The OWC uses the formatted value to display the cell values. The statements and query sent to Analysis Services by the OWC are shown here:

```
Drop visual totals for [Adventure Works DW]
CREATE SESSION
```

```
SET [Adventure Works DW].[{7868741D-072F-458A-8A8D-EA3FED4A3FA7}
Pivot2Axis0Set0] AS
‘
{
{ [Dim Customer].[Marital Status].[All] },
AddCalculatedMembers([Dim Customer].[ Marital Status].[ Marital
Status].MEMBERS)
}
‘
SET [Adventure Works DW].[{7868741D-072F-458A-8A8D-EA3FED4A3FA7}
Pivot2Axis1Set0] AS
‘
{
{ [Dim Promotion].[English Promotion Type].[All] },
AddCalculatedMembers([Dim Promotion].[English Promotion Type].[English
Promotion Type].MEMBERS)
}
‘
SELECT
NON EMPTY [{7868741D-072F-458A-8A8D-EA3FED4A3FA7}Pivot2Axis0Set0]
DIMENSION PROPERTIES MEMBER_NAME, PARENT_UNIQUE_NAME ON COLUMNS,
NON EMPTY [{7868741D-072F-458A-8A8D-EA3FED4A3FA7}Pivot2Axis1Set0]
DIMENSION PROPERTIES MEMBER_NAME, PARENT_UNIQUE_NAME ON ROWS,
{
[Measures].[Sales Amount]
}
ON PAGES
FROM [Adventure Works DW]
CELL PROPERTIES VALUE, FORMATTED_VALUE, FORE_COLOR, BACK_COLOR
DROP SET [Adventure Works DW].[{7868741D-072F-458A-8A8DEA3FED4A3FA7}
Pivot1Axis0Set0]
DROP SET [Adventure Works DW].[{7868741D-072F-458A-8A8DEA3FED4A3FA7}
Pivot1Axis1Set0]
```

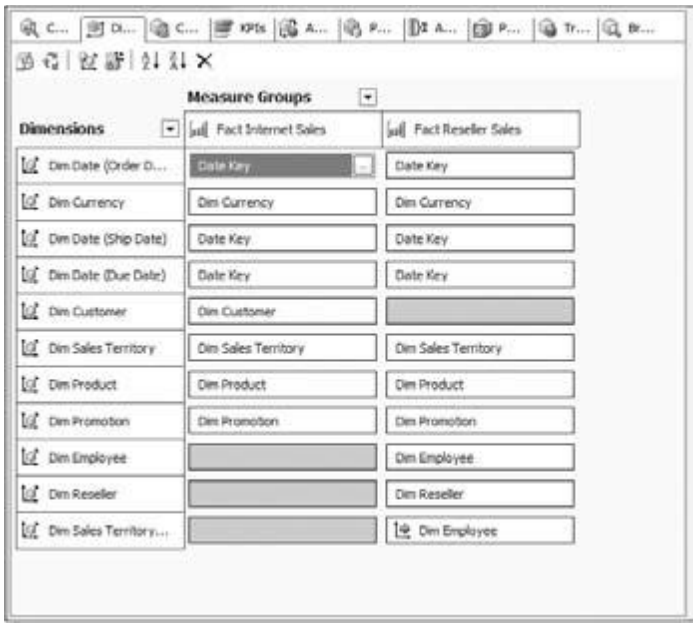
If you hover over a particular cell you can see the cell values without formatting, along with the row and column member values that correspond to that cell as shown in Figure 6 - 9 .

Cube Dimensions

The Cube Wizard helps you create your cube object from the DSV by creating appropriate dimension objects. The wizard detects the relationships between dimension tables and fact tables in the DSV, creates appropriate dimensions if needed, and establishes appropriate relationships between the dimensions and measure groups within the cube. As mentioned in the previous section, a cube contains an instance of the database dimension referred to as cube dimension. There can be multiple instances of a database dimension within a cube. There exists a relationship between the cube dimension and the measure groups within the cube. In this section you learn about various types of relationships between the cube dimensions and the measure groups within, as well as refine the Adventure Works DW cube created by Cube Wizard by adding a new dimension.

The Cube Wizard establishes relationships between the measure groups and cube dimensions based on its analysis of relationships in the DSV. You might have to refine these relationships based on your business needs. You can change these relationships in the Dimension Usage tab of the cube editor. If you switch to the Dimension Usage tab you will see the dimensions, measure groups of the cube, and the relationships between them, as shown in Figure 6 - 10 .

The cube dimensions and measure groups are represented in a matrix format as rows and columns, respectively, where the relationship between them corresponds to the intersection cell. The intersection cell shows the dimension type along with the attribute that is used in the relationship to join.


Measure Groups
Dimensions
Fact Internet Sales
Fact Reseller Sales
Dim Date (Order D...
Date Key
Date Key
Dim Currency
Dim Currency
Dim Currency
Dim Date (Ship Date)
Date Key
Date Key
Dim Date (Due Date)
Date Key
Date Key
Dim Customer
Dim Customer
Dim Sales Territory
Dim Sales Territory
Dim Sales Territory
Dim Product
Dim Product
Dim Product
Dim Promotion
Dim Promotion
Dim Promotion
Dim Employee
Dim Employee
Dim Reseller
Dim Reseller
Dim Sales Territory...
Dim Employee