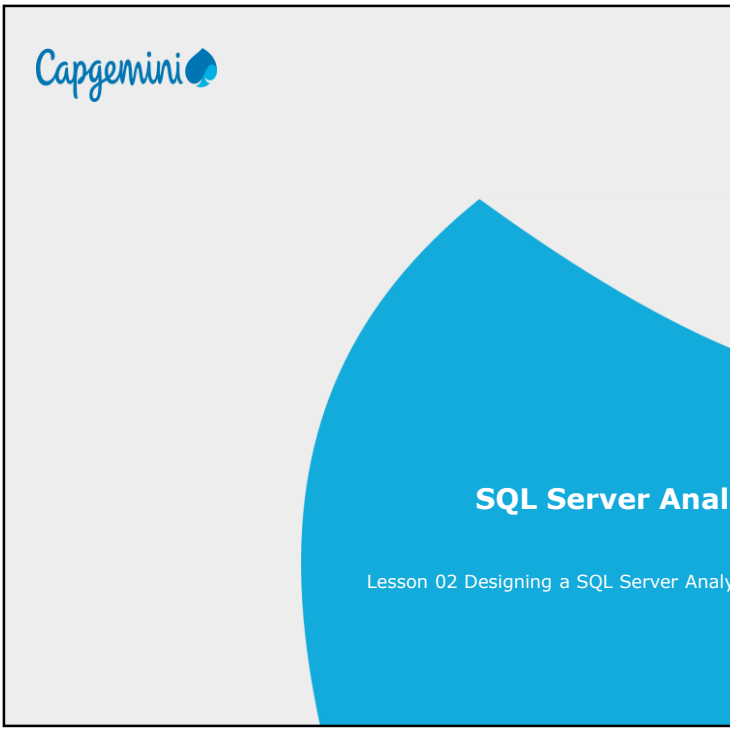
Instructor Notes:



Instructor Notes:

Lesson Objectives

On completion of this lesson on SSAS, you will learn:

- Components of Designing a Analysis Services Database
- Concept of Dimension
- Attributes and Hierarchy
- Concept of Cube
- Measures and Measure Groups
- Dimension Tables
- Fact Tables
- Star Schema and Snowflake Schema

Instructor Notes:

Designing a Analysis Services Database

Following are the major components of SSAS Database:

- Data Sources
- Data Source Views
- Dimensions
- Measures
- Cubes
- Dimension Tables
- Fact Tables

Instructor Notes:

Designing a Analysis Services Database

Data Sources:- A data source represents a connection that can be shared among OLAP cubes and data mining models in a project.

Data Source Views:- A data source view represents a subset of the data in a data source, and can also contain named queries and named calculations. Data source views can also be shared among multiple OLAP cubes and data mining models in a project.

Instructor Notes:

Designing a Analysis Services Database

Dimensions:-

- Dimensions are used in Analysis Services to organize data within cubes. Dimensions use a combination of hierarchies and attributes to represent levels of categorization within a cube.

Cubes:-

- A cube represents a set of measures that are grouped into measure groups and hierarchically organized by dimensions. Cubes are typically constructed from data retrieved from relational data sources, such as an OLTP database, a data warehouse, or a data mart.

Measures:-

- The measures are the usually numeric values that provide summaries at various different levels of aggregation.

What is an OLAP cube ?
An OLAP cube is a specially designed database that is optimized for reporting. While most databases designed for online transaction processing such as those used in claims processing are designed for efficiency in data storage, OLAP cubes are designed for efficiency in data retrieval. This means that the data is stored in such a way as to make it easy and efficient for reporting.
Regular “relational” databases treat all data into the database similarly, however OLAP cubes categorize data into "dimensions" and "measures". Measures represent items that are counted, summarized or aggregated, such as costs or units of service. **Dimensions** are variables by which measures are summarized, such as hospitals, physicians, or dates of service. Dimensions represent the variables by which measurement is performed, such as date, location, product code. This organization of data greatly facilitates the ability to formulate data requests based on real-life situations. In addition, many of queries that could be posed to the data are "pre-aggregated" in the database such that the answers have already been precalculated and can be reported without delay.
OLAP cubes can be thought of as extensions to the two-dimensional array of a spreadsheet. For example a company might wish to analyze some financial data by product, by time-period, by city, by type of revenue and cost, and by comparing actual data with a budget. These additional methods of analyzing the data are known as dimensions. Because there can be more than three dimensions in an OLAP system the term hypercube is sometimes used.
A **Measure** represents a column that contains quantifiable data, usually numeric, that can be aggregated. A measure is generally mapped to a column in a fact table.
You can also define a measure as a calculated member by using a Multidimensional Expressions (MDX) to provide a calculated value for a measure based on other measures in the cube. Calculated members add flexibility and analysis capability to a cube in Analysis Services.

Instructor Notes:

Designing a Analysis Services Database

Dimension Table:- A Relational table in Data ware house which is used to define a dimension for a cube. This table stores all the dimension members and their details

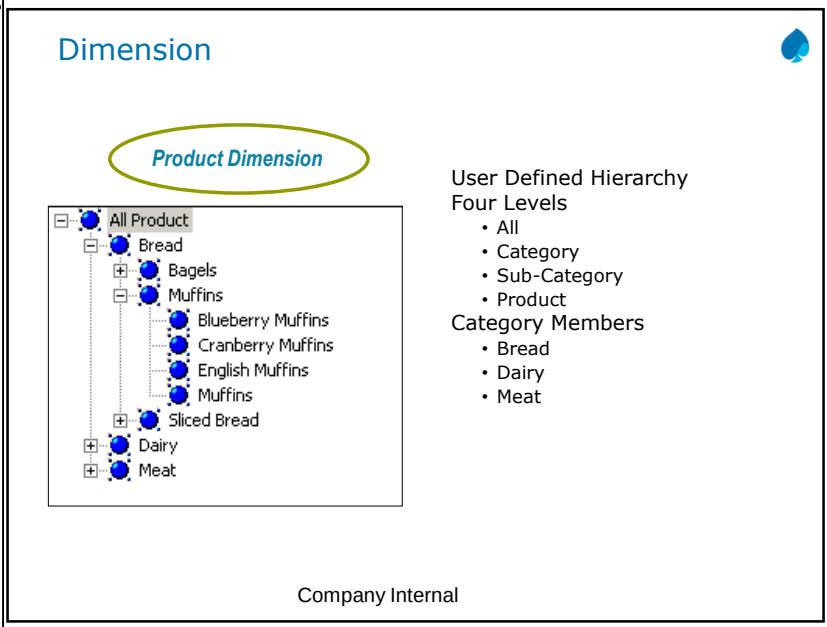
- For example a Product table in data ware house would define a dimension called product.

Fact Table:- A Relational table in Data ware house which contains facts about business (usually numerical). The fact table has data which is the source of creating measures.

Company Internal

Business Intelligence Development Studio is the environment that you will use to develop Online Analytical Processing (OLAP) cubes and data mining models in SQL Server Analysis Services. Business Intelligence Development Studio is the Microsoft Visual Studio 2008 environment with enhancements that are specific to business intelligence solutions. For more information about the general features of Business Intelligence Development Studio

Instructor Notes:

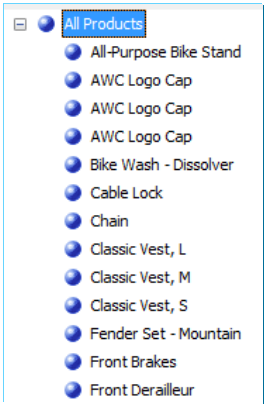


In Microsoft SQL Server Analysis Services, dimensions are a fundamental component of cubes. Dimensions organize data with relation to an area of interest, such as customers, stores, or employees, to users. Dimensions in Analysis Services contain attributes that correspond to columns in dimension tables. These attributes appear as attribute hierarchies and can be organized into user-defined hierarchies, or can be defined as parent-child hierarchies based on columns in the underlying dimension table. Hierarchies are used to organize measures that are contained in a cube.

All Microsoft SQL Server Analysis Services dimensions are groups of attributes based on columns from tables or views in a data source view. Dimensions exist independent of a cube, can be used in multiple cubes, can be used multiple times in a single cube. A dimension that exists independent of a cube is called a database dimension and an instance of a database dimension within a cube is called a cube dimension.

Instructor Notes:

Dimension



Default Attribute Hierarchy

Company Internal

In Microsoft SQL Server Analysis Services, dimensions are a fundamental component of cubes. Dimensions organize data with relation to an area of interest, such as customers, stores, or employees, to users. Dimensions in Analysis Services contain attributes that correspond to columns in dimension tables. These attributes appear as attribute hierarchies and can be organized into user-defined hierarchies, or can be defined as parent-child hierarchies based on columns in the underlying dimension table. Hierarchies are used to organize measures that are contained in a cube.

All Microsoft SQL Server Analysis Services dimensions are groups of attributes based on columns from tables or views in a data source view. Dimensions exist independent of a cube, can be used in multiple cubes, can be used multiple times in a single cube. A dimension that exists independent of a cube is called a database dimension and an instance of a database dimension within a cube is called a cube dimension.

Instructor Notes:

OLAP Dimension

A database dimension is a collection of related objects, called attributes, which can be used to provide information about fact data in one or more cubes.

For example, typical attributes in a product dimension might be product name, product category, product line, product size, and product price.

By default, these attributes are visible as attribute hierarchies and can be used to understand the fact data in a cube.

Attributes can be organized into user-defined hierarchies that provide navigational paths to assist users when browsing the data in a cube

Dimension is a term in data management and data warehousing that refers to logical groupings of data such as geographical location, customer information, or product information. Slowly Changing Dimensions (SCDs) are dimensions that have data that changes slowly, rather than changing on a time-based, regular schedule.[1]
For example, you may have a dimension in your database that tracks the sales records of your company's salespeople

Instructor Notes:

OLAP Dimension

KeyAttribute:

- Each dimension contains a key attribute. Each attribute is bound to one or more columns in a dimension table.
- The key attribute is the attribute in a dimension that identifies the columns in the dimension main table that are used in foreign key relationships to the fact table.
- Typically, the key attribute represents the primary key column or columns in the dimension table

Attribute Hierarchies:

- By default, attribute members are organized into two level hierarchies, consisting of a leaf level and an All level.
- Attributes can be, and typically are, arranged into user-defined hierarchies that provide the drill-down paths by which users can browse the data in the measure groups to which the attribute is related.

Instructor Notes:

Cube, Dimension and Measures

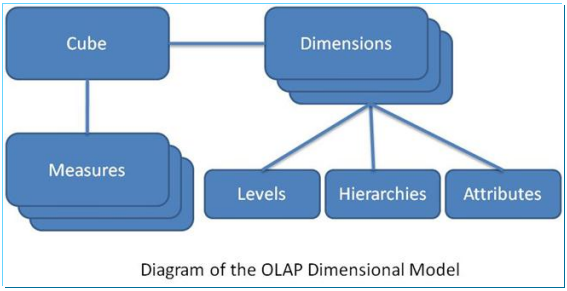


Diagram of the OLAP Dimensional Model

A Cube object is composed of dimensions, and measure groups.

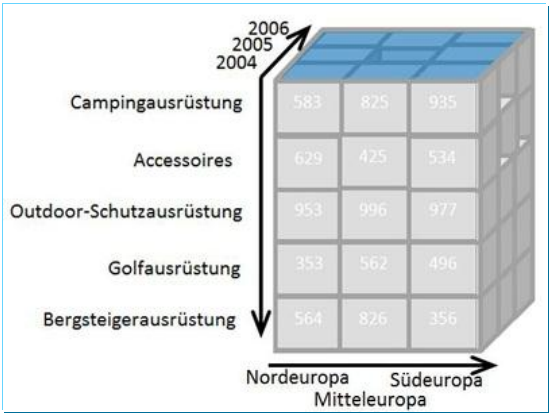
Dimensions are the actual set of dimensions used in the cube. All dimensions have to be defined in the dimensions collection of the database before being referenced in the cube.

Measure groups are sets of measures in the cube. A measure group is a collection of measures that have a common data source view and a common set of dimensions. A measure group is the unit of process for measures; measure groups can be processed individually and then browsed

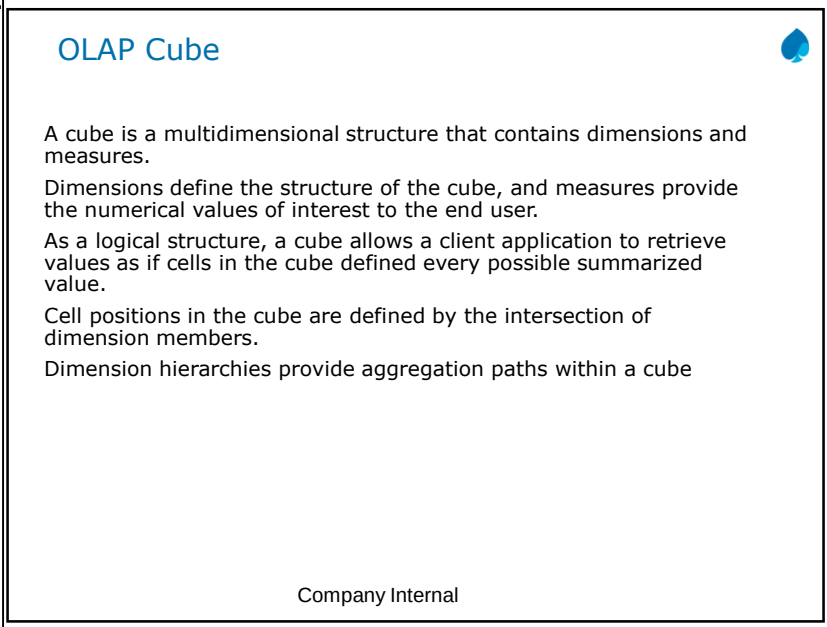
Instructor Notes:

Cube, Dimension and Measures

Product Category and Region Wise Sales across Year timeline



Instructor Notes:



- Each cube can contain multiple fact tables. A collection of measures exposed by a fact table is referred to as a measure group.

- Instead of building several small cubes and exposing combined functionality through virtual cubes MSAS 2008 allows a cube architect to build a large cube (with multiple measure groups)

- Cubes are developed using Business Intelligence Development Studio (BIDS) an add-on Visual Studio 2008. Cube development doesn't have to be associated with Analysis Server instance; you can develop cubes on small development machines and subsequently deploy them to an Analysis Server.

- Cubes are administered using SQL Server Management Studio (SSMS). You can backup MSAS databases, configure MSAS properties, and process cubes using SSMS. However, you cannot modify cube or dimension structures with SSMS.

Instructor Notes:

Star Schema and Snowflake Schema Structure

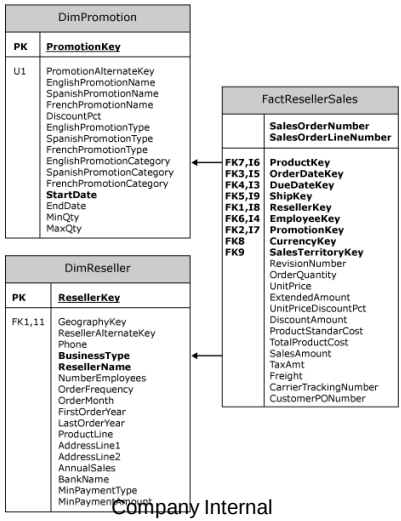
The structure of a dimension is driven by the structure of the underlying dimension table or tables.

Star Schema:

- The simplest structure is called a star schema, where each dimension is based on a single dimension table that is directly linked to the fact table by a primary key - foreign key relationship.
- Less no. of foreign keys and hence lesser query execution time
- Has redundant data

Instructor Notes:

Star Schema



Instructor Notes:

Star Schema and Snowflake Schema Structure

Snowflake Schema:

- A more complex structure called a snowflake schema could be created where dimension is based on attributes from columns in multiple tables linked to each other and ultimately to the fact table by primary key - foreign key relationships
- No Redundancy of data
- More foreign keys-and hence more query execution time

Instructor Notes:

Snowflake Schema

Company Internal

Instructor Notes:

Explain the lesson coverage

Summary

In this lesson, you learnt:

- Components of Designing a Analysis Services Data
- Concept of Dimension
- Attributes and Hierarchy
- Concept of Cube
- Measures and Measure Groups
- Dimension Tables
- Fact Tables
- Star Schema and Snowflake Schema