

## Lesson 2: Multidimensional Analysis

### CONTENTS

#### Objectives

#### Introduction

#### 2.1 Dimension Attributes

##### 2.1.1 Key Attribute

#### 2.2 Dimension Hierarchy

##### 2.2.1 Type of Hierarchy

#### 2.3 Summary

#### 2.4 Keywords

#### 2.5 Review Questions

#### 2.6 Further Readings

### Objectives

After studying this unit, you will be able to:

Recognize the Dimension Attributes

Summarize the Dimension Hierarchy

Identify the Type of Hierarchy

### Introduction

In statistics and related fields, multidimensional analysis is a data analysis process that groups data into two or more categories: data dimensions and measurements. To show this, let us take the case of a football game. A data set which comprises of the number of wins for one cricket team every year for many years could be categorized into a single dimensional or longitudinal data set.

Another data set which comprises of the number of wins many different cricket groups inside a year can be under a single dimensional or traverse sectional data set. A single data set that comprises of the number of wins for diverse cricket teams over numerous years could be comprised in a two-dimensional data set.

Multi-Dimensional analysis is an Informational analysis on data which takes into account numerous distinct connections, each of which comprises a dimension. For example, a retail analyst may want to understand the connections amidst sales by district, by quarter, by demographic circulation or by product. Multi-dimensional analysis will yield outcomes for these complex relationships.

### 2.1 Dimension Attributes

A dimension consists of members.



Example: The members of a product dimension are the individual products.

Members have attributes to identify them.

Notes



Example: Some possible attributes for a product dimension could be the product code, colour, and size.

If the dimension is defined as a hierarchy, the lower levels of the hierarchy must also have an attribute that identifies the parent of each member. Information about each dimension is stored in one or more dimension tables.

### 2.1.1 Key Attribute

Each dimension contains a key attribute. Each attribute is bound to have 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.



**Caution** Typically, the key attribute represents the primary key column or columns in the dimension table.

An attribute can also be bound to one or more additional columns for a specific task.



Example: An attribute's Name property determines the name that appears to the user for each attribute member and this property can be bound to a calculated column in the data source view.

Table 2.1 shows dimension attribute properties.

Table 2.1: Dimension Attribute Properties

Property	Description
Attribute Hierarchy Display Folder	Identifies the folder in which to display the associated attribute hierarchy to end users.
Attribute Hierarchy Enabled	Determines whether an attribute hierarchy is generated by Analysis Services for the attribute. If the attribute hierarchy is not enabled, the attribute cannot be used in a user-defined hierarchy and the attribute hierarchy cannot be referenced in Multidimensional Expressions (MDX) statements.
Attribute Hierarchy Optimized State	Determines the level of optimization applied to the attribute hierarchy. By default, an attribute hierarchy is Fully Optimized, which means that Analysis Services builds indexes for the attribute hierarchy to improve query performance.  The other option, Not Optimized, means that no indexes are built for the attribute hierarchy. Using Not Optimized is useful if the attribute hierarchy is used for purposes other than querying, because no additional indexes are built for the attribute. Other uses for an attribute hierarchy can be helping to order another attribute.
Attribute Hierarchy Ordered	Determines whether the associated attribute hierarchy is ordered. The default value is True. However, if an attribute hierarchy will not be used for querying, you can save processing time by changing the value of this property to False.
Attribute Hierarchy Visible	Determines whether the attribute hierarchy is visible to client applications. The default value is True. However, if an attribute hierarchy will not be used for querying, you can save processing time by changing the value of this property to False.

Contd....

## Notes

Custom Rollup Column	Specifies the column that defines a custom rollup formula.
Custom Rollup Properties Column	Specifies the column that contains the properties of a custom rollup formula.
Default Member	Specifies a Multidimensional Expressions (MDX) expression that defines the default measure for the attribute.
Description	Contains the description of the attribute.
Discretization Bucket Count	Contains the number of buckets into which to discretize.
Discretization Method	Defines the method to use for discretization.
Estimated Count	Specifies the estimated number of members in the attribute. Until you run the Aggregation Design Wizard, the default value is zero. Either you can allow the wizard to count the number of records or you can enter an estimated value. Enter a value manually if you know the number of members and want to save the time that is required to query the database for the count. If you are working with a test subset of your production data, you can use the counts of your production data so that the aggregation design will be optimized for the production data instead of the test data.
Grouping Behaviour	A user defined value that provides a hint to client applications on how to group attributes.
ID	Contains the unique identifier (ID) of the dimension.
Instance Selection	Provides a hint to client applications about how a list of items should be displayed, based on the expected number of items in the list. The available options are as follows: <ul style="list-style-type: none"> <li>• None No hint is provided to the client application. This is the default value.</li> <li>• Drop Down The number of items is small enough to display in a drop-down list.</li> <li>• List The number of items is too large for a drop-down list, but does not require filtering.</li> <li>• Filtered List The number of items is large enough to require users to filter the items to be displayed.</li> <li>• Mandatory Filter The number of items is so large that the display must always be filtered.</li> </ul>
Is Aggregatable	Specifies whether the values of the attribute members can be aggregated. The default value is True, which means that the attribute hierarchy contains an (All) level. If the value for this property is False, the attribute hierarchy does not contain an (All) level.
Key Columns	Contains the column or columns that represent the key for the attribute, which is the column in the underlying relational table in the data source view to which the attribute is bound. The value of this column for each member is displayed to users unless a value is specified for the Name Column property.
Member Names Unique	Determines whether member names in the attribute hierarchy must be unique.
Members With Data	Used by parent attributes to determine whether to display data members for non-leaf members in the parent attribute. This property value is only used when the value of the Usage property is set to Parent. This means that a

Contd....

Notes

Members with Data Caption	parent-child hierarchy has been defined. The available options are as follows: <ul style="list-style-type: none"> <li>• Non-leaf Data Hidden Non-leaf data is hidden.</li> <li>• Non-leaf Data Visible Non-leaf data is visible.</li> </ul>
	Provides a template string that is used by parent attributes to create captions for system-generated data members in the parent attribute. This property value is only used when the value of the Usage property is set to Parent. This means that a parent-child hierarchy has been defined.
Name	Contains the user-friendly name of the attribute.
Name Column	Identifies the column that provides the name of the attribute that is displayed to users, instead of the value in the key column for the attribute. This column is used when the key column value for an attribute member is cryptic or not otherwise useful to the user, or when the key column is based on a composite key. The Name Column property is not used in parent-child hierarchies; instead, the Name Column property for child members is used as the member names in a parent-child hierarchy.
Naming Template	Defines how levels are named in a parent-child hierarchy constructed from the parent attribute. This property value is only used when the value of the Usage property is set to Parent. This means that a parent-child hierarchy has been defined.
Order By	Describes how to order the members that are contained in the attribute hierarchy. The default value is Name, which specifies that ordering of the attribute members is based on the value of the Name Column property, if any. Otherwise, members are ordered by the value of the key column. The available options are as follows: <ul style="list-style-type: none"> <li>• Name Column Order by the value of the Name Column property.</li> <li>• Key Order by the value of the key column of the attribute member.</li> <li>• Attribute Key Order by the value of the member key of a specified attribute, which must have an attribute relationship to the attribute.</li> <li>• Attribute Name Order by the value of the member name of a specified attribute, which must have an attribute relationship to the attribute.</li> </ul>
Order By Attribute	Identifies the attribute by which to order the members of the attribute hierarchy.
Root Member If	Determines how the root or topmost members of a parent-child hierarchy are identified. This property value is only used when the value of the Usage property is set to Parent. This means that a parent-child hierarchy has been defined. The default value is Parent Is Blank Self Or Missing, which means that only members that meet one or more of the conditions described for Parent Is Blank, Parent Is Self, or Parent Is Missing are treated as root members. The following values are also available: <ul style="list-style-type: none"> <li>• Parent Is Blank Only members with a null, a zero, or an empty string in the key column or columns are treated as root members.</li> <li>• Parent Is Self Only members with themselves as parents are treated as root members.</li> <li>• Parent Is Missing Only members with parents that cannot be found are treated as root members.</li> </ul>
Type	Contains the type of the attribute.
Unary Operator Column	Specifies the column that provides unary operators. It is a binding of Data Item type that defines the details of a column providing a unary operator.

Contd....

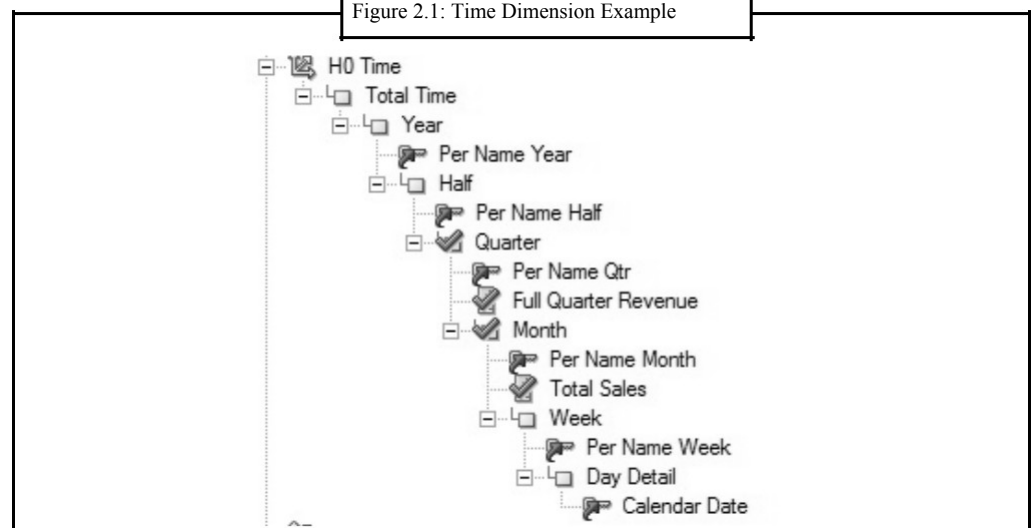
## Notes

Usage	<p>Describes how an attribute is used.</p> <p>The available options are as follows:</p> <p>Regular The attribute is a regular attribute. This is the default value.</p> <p>Key The attribute is a key attribute.</p> <p>Parent The attribute is a parent attribute.</p>
Value Column	<p>Identifies the column that provides the value of the attribute. If the Name Column element of the attribute is specified, the same Data Item values are used as default values for the Value Column element. If the Name Column element of the attribute is not specified and the Key Columns collection of the attribute contains a single Key Column element representing a key column with a string data type, the same Data Item values are used as default values for the Value Column element.</p>

## 2.2 Dimension Hierarchy

A hierarchy is a set of parent-child relationships between attributes within a dimension. These hierarchy attributes are also known as levels. Example, the Time Dimension can have Total, Year, Quarter, Month and Date as its levels as shown in Figure 2.1:

Figure 2.1: Time Dimension Example



### 2.2.1 Type of Hierarchy

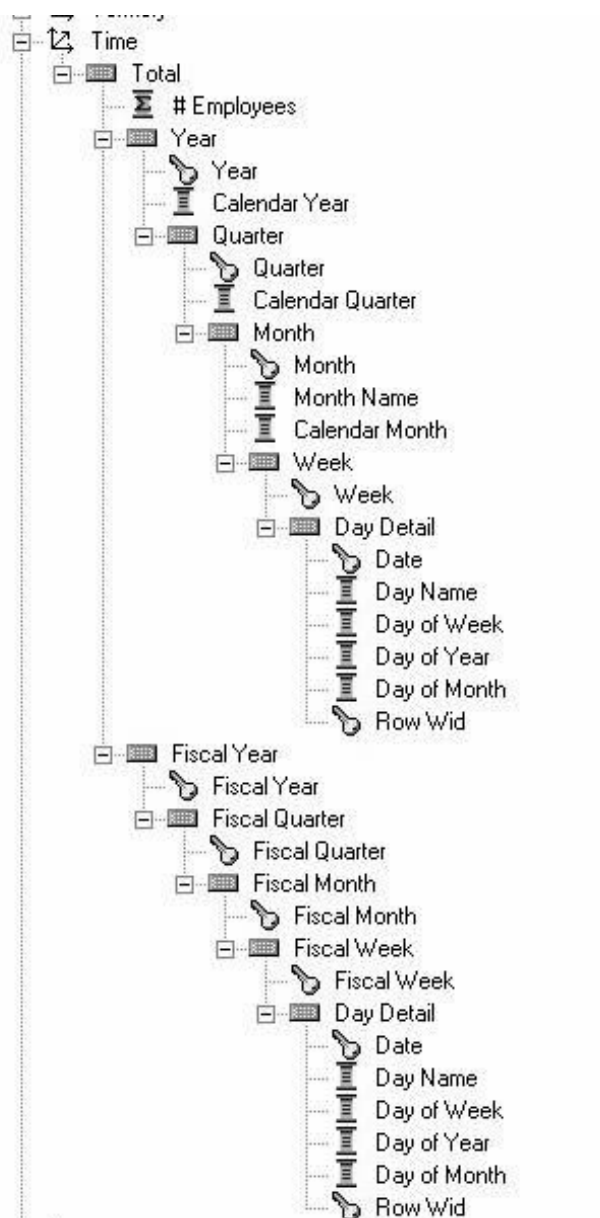
Level-based: This type of hierarchy consists of an ordered set of two or more levels.



Example: A time hierarchy might have three levels for Year, Quarter, and Month.

Level-based hierarchies can also contain parent-child relationships. This type of dimension hierarchy levels allow to perform aggregate navigation and configure level-based measure calculations.

Figure 2.2: Level based Hierarchy Example



**Notes**

Also it supports special type of level-based dimension for unbalanced and Skip-level hierarchy.

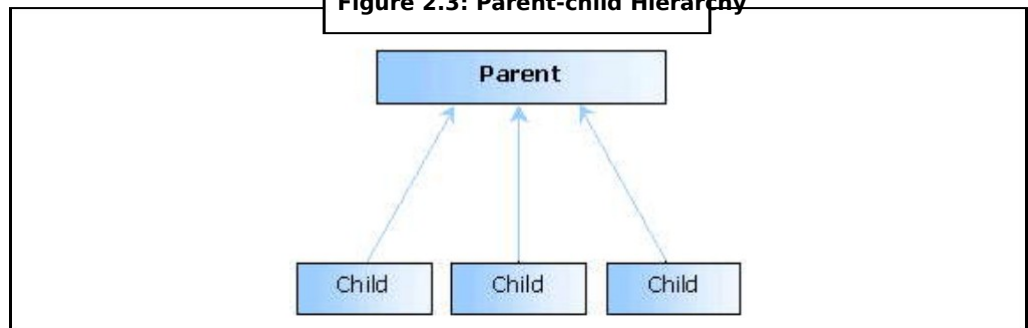
It also supports time dimension to provide special functionality for modelling time series data.

**Parent-child:** A parent-child hierarchy is a hierarchy in a standard dimension that contains a parent attribute. A parent attribute describes a self-referencing relationship, or self-join, within a dimension main table. It is actually value-based hierarchy. It consists of values that define the hierarchy in a parent-child relationship (Figure 2.3).



*Example:* An employee hierarchy might have no levels but might just contain names of employees who are managed by other employees. Employees may have titles like Vice President and then Vice Presidents might report to other Vice Presidents and they can be at different depths in the hierarchy.

**Figure 2.3: Parent-child Hierarchy**



**Source:** <http://3d.recoil.org/nojavascript/Images/Parent-Child.gif>

In addition to above discussed two level of hierarchy it can be of following type as well:

**Ragged Hierarchy:** A hierarchy in which all the lowest-level members do not have the same level of depth is ragged hierarchy.



*Example:* A time hierarchy might be having current month data at the day level, the previous month's data at the month level, and the previous 10 year's data at the quarter level.

It is also known as an unbalanced hierarchy.

**Skip-level:** A hierarchy in which certain members do not have values for certain higher levels are known as skip-level hierarchy.



*Example:* In India, Delhi city does not belong to another state (it belongs to Delhi as a state itself).

What matters is that users can still navigate from the country level (India) to Delhi (city level) and below without the need for a state level.

**User-defined:** These are user-defined hierarchies of attributes that are used in service of Microsoft SQL Server to arrange the members of a dimension into hierarchical structures and provide navigation paths in a form of cube.

For example, the Table 2.2 defines a dimension table for a time dimension.





Table 2.2: Dimension Table for Time Dimension

Notes

Year	Quarter	Month
1999	Quarter 1	Jan
1999	Quarter 1	Feb
1999	Quarter 1	Mar
1999	Quarter 2	Apr
1999	Quarter 2	May
1999	Quarter 2	Jun
1999	Quarter 3	Jul
1999	Quarter 3	Aug
1999	Quarter 3	Sep
1999	Quarter 4	Oct
1999	Quarter 4	Nov
1999	Quarter 4	Dec



*Did u know?*

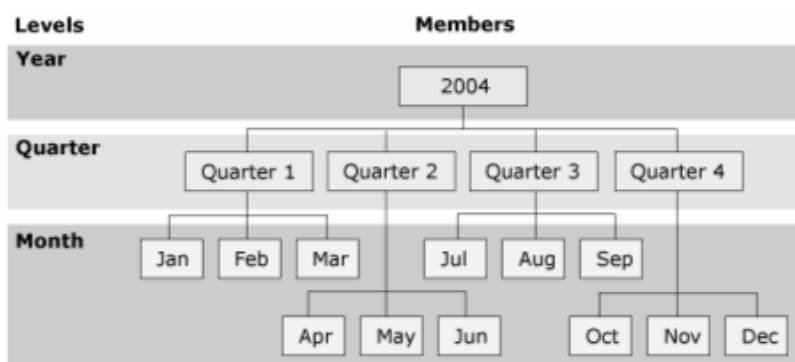
The Year, Quarter, and Month attributes are used to construct a user-defined hierarchy, named Calendar, in the time dimension.



*Task* Prepare a table hierarchy of employees according to their salary.

The relationship between the levels and members of the Calendar dimension is shown in Figure 2.4.

Figure 2.4: Relationship between the Levels and Members of the Calendar Dimension



*Notes* Any hierarchy other than the default two-level attribute hierarchy is called a user-defined hierarchy.

## Notes

### Self Assessment

State whether the following statements are True or False:

A hierarchy is a set of parent-child relationships between attributes within a dimension.

Level-based hierarchies can also contain parent-child relationships.

A Ragged Hierarchy describes a self-referencing relationship, or self-join, within a dimension main table.

Parent attribute is also known as an unbalanced hierarchy.

Skip-level is hierarchies of attributes that are used in service of Microsoft SQL Server to arrange the members of a dimension into hierarchical structures.

User-defined is a hierarchy in which certain members do not have values for certain higher levels are known as skip-level hierarchy.

Ragged Hierarchy consists of values that define the hierarchy in a parent-child relationship.

## Notes

### 2.3 Summary

If the dimension is defined as a hierarchy, the lower levels of the hierarchy must also have an attribute that identifies the parent of each member.

Typically, the key attribute represents the primary key column or columns in the dimension table.

A hierarchy is a set of parent-child relationships between attributes within a dimension. These hierarchy attributes are also known as levels.

Level-based hierarchies can contain parent-child relationships.

A parent-child hierarchy is a hierarchy in a standard dimension that contains a parent attribute.

A hierarchy in which all the lowest-level members do not have the same level of depth is ragged hierarchy.

The Year, Quarter, and Month attributes are used to construct a user-defined hierarchy, named Calendar, in the time dimension.

### 2.4 Keywords

**Hierarchy:** A hierarchy is a set of parent-child relationships between attributes within a dimension.

**Key attribute:** 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.

**Level-based:** This type of hierarchy consists of an ordered set of two or more levels.

Parent-child: A parent-child hierarchy is a hierarchy in a standard dimension that contains a parent attribute.

Ragged Hierarchy: A hierarchy in which all the lowest-level members do not have the same level of depth is ragged hierarchy.

Skip-level: A hierarchy in which certain members do not have values for certain higher levels are known as skip-level hierarchy.

User-defined: These are user-defined hierarchies of attributes that are used in service of Microsoft SQL.

## 2.5 Review Questions

What is the multi-dimensional analysis?

Discuss about the key attribute of dimension attributes.

Briefly explain the dimension attribute properties.

What is the dimension hierarchy? Explain with example.

Discuss are the various type of hierarchy.

“Parent-child is actually value-based hierarchy”. Comment.

What is ragged hierarchy? Give the suitable example.

Explain the relationship between the levels and members of the calendar dimension.

## Answers: Self-Assessment

Notes

- |                                  |                     |
|----------------------------------|---------------------|
| 1. Multi-Dimensional analysis    | 2. Dimension tables |
| 3. Primary key column or columns | 4. Attribute        |
| 5. True                          | 6. True             |
| 7. False                         | 8. False            |
| 9. False                         | 10. False           |
| False                            |                     |

