



XLCubed

OLAP Dictionary

Introduction:

Why should you as an Excel user be interested in OLAP? Well think about it, the latest OLAP clients like XLCubed allow you to access OLAP databases from within Excel and take the burden of processing large amounts of data away from Excel and place it securely on a server where everyone shares one version of the truth. OLAP databases can be built from your existing accounting system's relational tables to provide advanced reporting and analysis. You could build a cube from your general ledger with Account, Department and Month dimensions. The key thing to appreciate is that your accounting system is designed for transactions (e.g. entering invoices) whereas OLAP databases are designed for reporting and analysis, hence the OLAP (On-line Analytical Processing) acronym. Given some of the horror stories floating about in the Excel Zone it is only a matter of time before OLAP becomes a consideration for many of you. For those of you who are new to the OLAP world there often seems to be some confusing nomenclature and a lack of hard and fast definitions. Unfortunately that is the nature of OLAP as it is very difficult to visualise and explain the concepts of multi-dimensionality in a world of limited dimensions... As we can't afford the Doctor Who special effects we have tried to provide some definitions that will hopefully at least clarify some of the terms you might hear being bandied about. You are going to hear a lot more about OLAP in the coming year as Microsoft push its latest OLAP database (Analysis Services 2005 comes for free as part of the Yukon release of SQLServer).

OLAP Definitions

Aggregate

OLAP Databases usually have hierarchical relationships within one or more dimensions. Aggregation is simply the calculation of the hierarchical values. For example, in a time dimension you might have aggregations at quarter and year based on the individual months. Normally the calculations will be summations, but this does not have to be the case.

Analysis, Multi-Dimensional

Multi-dimensional analysis is what is performed against an OLAP data source. It means that an organisation's data can be viewed from multiple perspectives that provide insight into an organisation that may not be provided by traditional reporting. A well-designed OLAP database will allow the analyst to explore his data by representing the actual hierarchies and relationships within the organisation. By optimising reporting against large datasets OLAP tools allow analysts to quickly gain insight into complex data.

Calculated Member

A calculated member is a member of a dimension whose value is calculated from other member's values using a specified formula. This gives huge flexibility to the data which can be provided in the underlying cubes.

Cell

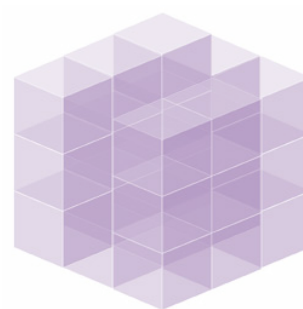
A Cell is the unique point at the intersection of a chosen member from each dimension of an OLAP database.

Children

Children in hierarchical terms are defined as members of a dimension that are required to produce an aggregate total for a parent member. Child members may themselves have children. A member may be a child of more than one parent, which need not be at the same hierarchical level – this allows complex hierarchical aggregations to exist within any dimension.

Cube

Cube is the generic, market friendly, term used to describe an OLAP database. Most organisations are multi-dimensional in the information they create. At any given time you might have simple hierarchies that can be expressed in two dimensions (like a spreadsheet), but once you want to look at those hierarchies over time you have immediately created a third dimension. A two dimensional array can be visualised as a spreadsheet with the dimensions on rows and columns. A three dimensional array can be visualised as a cube with each dimension being a side of the cube. Above three dimensions we have no metaphor, but that is how most organisations operate and the multi-dimensional aspect can be difficult to visualise. Once an organisation wants to view data across multiple dimensions (for example products by sales area by customer by time) an OLAP database becomes the quickest way to perform the analysis, particularly where large amounts of data are involved. Also referred to as Multi-Dimensional Arrays.



Dense

An OLAP database is considered to be dense if a high proportion of its possible combinations of dimension members contain actual data. See also Sparse.

Dimension

Dimensions are the part of the structure of a multi-dimensional array (OLAP Database) that contains members and defines their relationship. An OLAP database will contain multiple dimensions, each of which represents the users understanding of their organisation. A typical dimension would be **Time**, which may have a hierarchy of levels: **Year**, **Quarter**, and **Month** each with its own members e.g. 2004, Q1, January etc.

Drill Down/Up

Drilling Down or Up is a typical analytical technique used by OLAP client software to explore an OLAP database. It allows the user to navigate from the highest summary down the most detailed level of information. The drilling follows the path of the dimensional hierarchies.

Drill Through

Some OLAP client/database combinations allow the user to “drill through” to the source data that has been used to build the OLAP database. Typically an OLAP database is built on transactions from relational sources and the ability to get back to the relational data is invaluable.

Hierarchical Relationships

The member's of a dimension can be organised into parent-child relationships. This is usually done to provide a logical aggregation for data within the OLAP database, for example from Town to Region to Country as geographical hierarchy.

Leaf Node

The lowest level in a hierarchy.

Level

A term used to describe where a member resides within a hierarchy. Within a dimension's hierarchy members are said to be at the same level if they have the same number of descendants.

Member

A member is a discrete data item within a dimension. It is given a name or identifier to identify it and define its position within a dimension.

Multi-Dimensional Query Language

A Computer language that can be used to specify the data to be retrieved from a Multi-dimensional array (OLAP database). The execution of the query will result in the creation of a result set to represent anything from a cell to a slice or a sub-cube.

Nesting

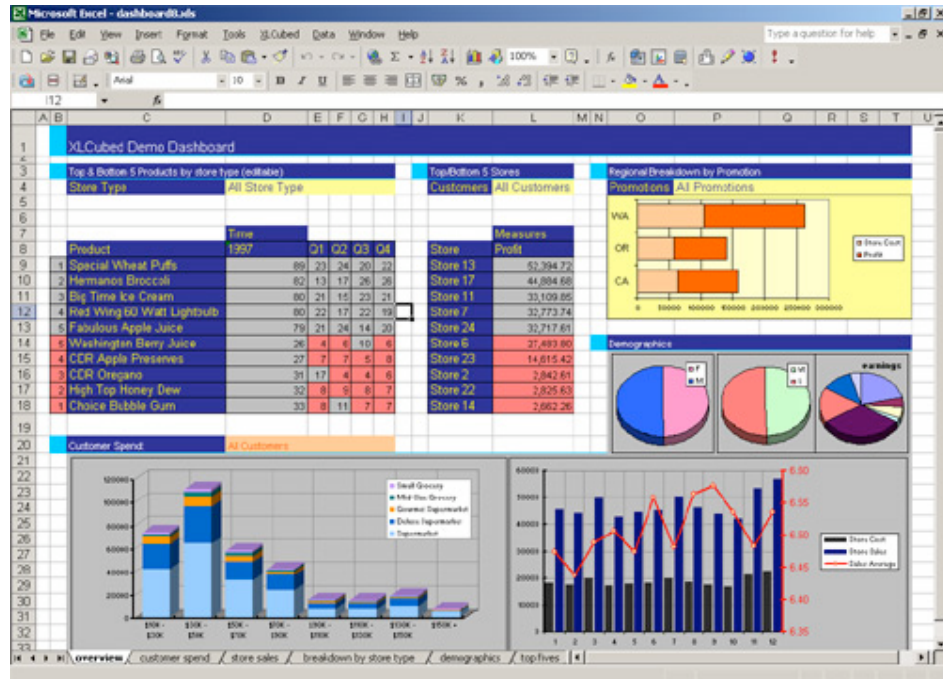
Nesting is a technique used in OLAP client tools to show dimensionality in a two dimensional report. Nesting describes the stacking of dimensions in multiple header rows or columns. It is easy to show in an example:

		Q1				Q2			
		Male		Female		Male		Female	
		Apples	Pears	Apples	Pears	Apples	Pears	Apples	Pears
Canada	Profit	X	X	X	X	X	X	X	X
	Unit Sales	X	X	X	X	X	X	X	X
USA	Profit	X	X	X	X	X	X	X	X
	Unit Sales	X	X	X	X	X	X	X	X
Europe	Profit	X	X	X	X	X	X	X	X
	Unit Sales	X	X	X	X	X	X	X	X

Both the rows and columns are nested.

OLAP Client or front-end

Software that allows users to connect to an OLAP database and allows for the data in the cube to be retrieved and analysed. Usually contain functionality that allows the power of the OLAP database to be exploited, for example ranking, sorting and filtering capability. OLAP clients may operate independently as self-contained software or enable other software to talk to the OLAP database, for example XLcubed within Excel.



an example report created within Excel using XLcubed

Many OLAP databases come with their own OLAP client but more specialised client software is often available from 3rd party vendors.

Parent

A member that exists one level up a hierarchy from another member is often referred to as a parent member. The parent member's value is usually an aggregation of its child member's values.

Pivot

To Pivot a report refers to the action of changing the dimensional orientation, or more simply swapping the rows with columns! It can also be used to describe actions in an OLAP client that allow dimensions to be swapped from the header to the body of a report.

Pre-Calculated Data

Pre-calculated data is data that has been calculated prior to its being requested in a query for example. This is done to improve adhoc query response, but obviously requires more storage space and processing time to build the OLAP database.

Roll-up

See Aggregate

Slice

A slice is a subset of a multi-dimensional array (OLAP database) that corresponds to specified values of one or more dimensions. You can visualise a slice as a page with a header that specifies the dimension values and a body that displays the values requested. Dimension values are either specified or assumed depending on the OLAP client.

Slice and Dice

A generic term used to describe typical analysis activity on an OLAP database. As it is impossible to view a multi-dimensional array it is necessary to view slices, the term "slice and dice" refers to concept of cutting through the imaginary cube.

Sparse

An OLAP database is considered to be sparse if only a small proportion of its possible combinations of dimension members contain actual data. See also Dense.

About XLCubed Limited

XLCubed Limited produce two products, an Excel Edition and a Web Edition. They are designed to work together seamlessly to provide a comprehensive environment to analyse and publish OLAP data in the familiar environments of Excel and the Web.

The Excel add-in was produced so that finance users felt immediately at home. XLCubed Limited has customers ranging from the world's largest Oil Companies and Banks all the way to a UK based 6th form college. In our experience end users don't care what the technology is, whether its OLAP, Excel and XLCubed it's a case of can they get their job done quicker, faster, within budget and with minimal outlay.

Find out more and download an evaluation copy of the software on <http://www.xlcubed.com>

