

# USING OLAP DATA CUBES IN BUSINESS INTELLIGENCE

Marian Pompiliu CRISTESCU  
marian.cristescu@ulbsibiu.ro

“LUCIAN BLAGA” UNIVERSITY OF SIBIU, ROMANIA

## ABSTRACT:

*The purpose of this paper is to demonstrate that it is possible to develop business intelligence projects in big and medium-size organizations, only with Microsoft products, used in accordance with standard OLAP cube technology, and presented possible alternatives, in relation with the requested functions.*

## KEYWORDS:

*Data cubes, OLAP technology, Business Intelligence, Microsoft SQL Server 2008, Data Mining*

## 1. Introduction

In the business environment the facts and situations have to be divided in superordinated and subordinated in such a way that entrepreneurs can discover how and why a certain event took place. Because OLAP cubes are also configured like this, data is easier to find. Similar to the hierarchy, OLAP cubes are dependent on the cubes interconnection.

Those cubes that are associated one with another not only allow the user to find information regarding a product, event or situation, but they allow the user to see data which interacts with the first cube as well. Cubes are created with the help of Analysis Services, the tables being selected through a wizard.

The dimensions and facts can either

be defined, either be selected at the proposal of the Wizard. Analysis Services also offer prefabricated dimensions, like day, month, quarter, semester, etc.

## 2. The OLAP Standard and Data Cubes

OLAP cubes are used to organize information for later use. OLAP cubes use the OLAP technology – Online Analytical Processing – to perform various tasks, usually regarding the processing and representation of information. OLAP cubes are good for distribution, marketing, management reporting, business process management, budgetary, forecast, billing and database analysis (Microsoft Corporation, 2010a).

The software programs that use OLAP cubes can automatically search information regarding a certain event,

product, service or project. For example: the biggest cube in the world currently has 1.4 terabytes and its average response time at each consultation is 1.2 seconds.

OLAP cubes are quite complex, but at the same time easy to utilize. The OLAP cube allows the user to access important data from many perspectives. The facts which are introduced in an OLAP cube are called “standards” and they are distributed in layers called “dimensions” (Microsoft Corporation, 2010b).

There are two essential aspects of OLAP cubes from which the entire software is formed. These aspects are the hierarchy and the binding in networks of the cubes. The OLAP cubes are extremely dependent to a structured hierarchy.

This hierarchy comprises a series of documents which are superordered and subordered. Which form a long list of lists. To create a Business Score Card (Schultz, Knuth & Pruß, 2008), only the temporal dimension was necessary. Dimensions and facts are linked one to the other through the “Date” field.

In the example below the “Date” field is valid as primary key in the “Date” table, “Date (Ship Date)” and “Date (Delivery Date)”

Date)”. In Figure no. 1, the relationship between the tables is easily noticeable.

The relationships are defined with the aid of relational data banks. This cube was created with the purpose of creating KPIs and reports as a demonstration and receives relevant data in this sense.

### 3. Key Performance Indicators

In the economic terminology, a KPI is a standard quantifiable to measure the success by the enterprise. In Analysis Services a KPI is a list of calculations which associate with a group of measurement in a cube, that are used to evaluate business success (Kemper, Mehanna, & Unger, 2004).

Typically, these calculations are a combination of “Multidimensional Expression” (MDX) expressions or calculated elements. KPIs also have additional data that show information regarding the mode in which client applications have additional metadata, which contain information about how the results of the KPIs should look (Kemper, Mehanna, & Unger, 2004).

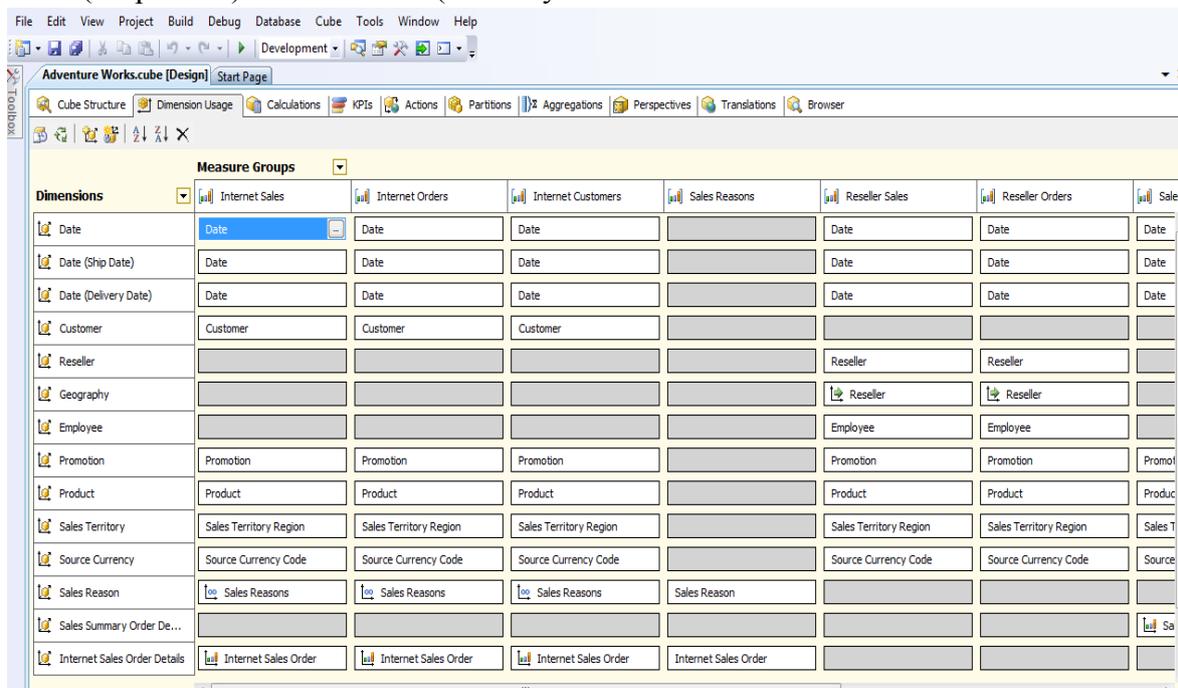


Figure no. 1. Visual Studio 2008 – The use of the dimension with the registration card

General KPI expressions:

- *The purpose* – a MDX of a numerical expression or a calculation that indicates the target value of the KPIs

- *The value* – a MDX of a numerical expression that indicates the current value of the KPIs

- *The status* – a MDX expression that represents the state of the KPIs in a certain moment. The status of a MDX expression should return a normalized value between -1 and 1. The values lower or equal to -1 are interpreted as “bad” or “low”, a value of zero (0) is interpreted as “acceptable” or “average” and values higher or equal to 1 are interpreted as “good” or “big”

- *The trend* – a MDX expression that evaluates the value of KPIs throughout time. The trend can be any criterion based on time which is useful in a business context. The MDX trend expression makes it easier for the business user to find out if the KPI values are improving or not throughout time.

- *Status display* – a visual element that generates the displays of status for a KPI

- *The trend barometer* – a visual element that offers a quick display of the trend for a KPI

- *The display folder* – the folder in which the KPI cube is found if a user searches in the cube

- *Superordered KPI* – a reference to an existing KPI which uses the value of the subordinate KPIs in the within the pale of calculation of the subordinate KPIs

- *The share* – a MDX numerical expression that indicates a relative significance of a KPI. If the KPI is assigned to a superordered KPI, the share is used for the proportional correction of the subordinate KPI value in the calculation of the value of the superordered KPI.

First, we will add a KPI with the name “ProfitPercentage” at the cube, selecting the *New KPI* button or we can click right on the mouse from the KPI organization window and *New KPI*.

In KPI Designer, which opens, we enter the name of the KPI as *ProfitPercentage*. We add the standard for the measurement group *SalesFact*, selecting the associated measurement group *Drop down*. Now we enter an expression to evaluate the characteristic value of the KPI. We enter the following expression in the expression field of the value:

$$\frac{[Measures].[Profit]}{[Measures].[Product Cost]}$$

**OR**

$$\frac{([Measures].[Sales Amount])}{[Measures].[Product Cost]-1}$$

We have obtained the expression above with the help of the logic that ProfitPercentage is calculated as the difference between distribution and costs, divided by costs. Here [Measures].[Profit] is “Calculated Member” according to the calculation below:

```
CREATE MEMBER
CURRENTCUBE.[MEASURES].[Profit]
AS SUM
([Date].[Calendar Year - Qtr -
Month].currentmember,
[Measures].[Sales Amount]-
[Measures].[Product Cost]),
FORMAT_STRING = "#,#",
VISIBLE = 1;
```

Now we set the objective of the KPI at a constant. We introduce 0.25 in the target expression field, which means that an increase by 25 % of the turnover is expected. The next step is to write an expression to compare the value with the target. Now we introduce the next MDX expression in the expression status field:

```
CASE
WHEN KPIValue("ProfitPercentage")/
KPIGoal("ProfitPercentage") >= .9
THEN 1
WHEN KPIValue("ProfitPercentage")/KP
IGoal("ProfitPercentage") < .8
THEN -1
ELSE 0
END
```

In this case a ratio of 0.9 would be seen as good. This means that, if the value reaches at least 90 % of the target, it can be assumed that it is successful, that is why the value will be displayed at 1. If the value/target is lower or equal to 0.8, then the status becomes -1 and it is assumed that it is a failure.

All the other values that reach the areas mentioned above are seen as neutral (0). We choose a type of image from the Drop Down status display to integrate the status that shows the KPI. The next characteristic that has to be configured is the trend. We approach a pseudo-notion:

```

CASE
  WHEN IsEmpty(x) THEN 0
  WHEN  $v - vX \geq 0.25$  THEN 1
  WHEN  $v - vX < -0.1$  THEN -1
  ELSE 0
END

```

Here “x” is the current period of time from the previous year. For example, if the current time period is 20-May-2016, then “x” will be 20-May-2015. “v” is the value of the KPI which would indicate in standard mode value for the current period of time. “vX” is the value of KPI for the same period of time from the previous year.

The expression first checks if “x” is empty, indicating 0 (no change in trend). Then it checks further if the current period increased by at least 25 %, which indicates 1 (a positive trend). If there is a regression of more than 10 % compared to last year, a negative trend is displayed or -1. All the other results don’t indicate a change in trend (or 0).

We introduce the following MDX expression in the field of the trend text:

```

CASE
  WHEN IsEmpty(ParallelPeriod([Date].
    [Calendar Year - Qtr - Month].
    [Calendar Year], 1, [Date].
    [Calendar Year - Qtr - Month].
    CurrentMember)) THEN 0
  WHEN KPIValue("ProfitPercentage")-
    (KPIValue("ProfitPercentage"),Par
    allelPeriod([Date].[Calendar Year -
    Qtr - Month].[Calendar Year],1,[Date].
    [Calendar Year - Qtr - Month].
    CurrentMember)) >= 0.25 THEN 1
  WHEN KPIValue("ProfitPercentage") -
    KPIValue("ProfitPercentage"),
    ParallelPeriod([Date].[Calendar Year
    - Qtr - Month].[Calendar Year],1,
    [Date].[Calendar Year - Qtr - Month].
    CurrentMember)) < -0.1 THEN -1
  ELSE 0
END

```

In the MDX expression above „X” is replaced with *ParallelPeriod([Date].[Calendar Year - Qtr - Month].[Calendar Year], 1,[Date].[Calendar Year - Qtr - Month].CurrentMember)*. *ParallelPeriod* is a function which indicates a member from a previous period in the same relative as an indicated element. The „V” variable is replaced with *KPIValue("ProfitPercentage")* and „vX” is replaced with *(KPIValue("ProfitPercentage"), ParallelPeriod([Date].[Calendar Year - Qtr - Month].[Calendar Year], 1,[Date].[Calendar Year - Qtr - Month].CurrentMember))*. We choose a suitable type of image from the Drop Down of the trend indicator to integrate Trends for the display of the KPI. In Figure no. 2 an example of how KPI Designer should look is presented:

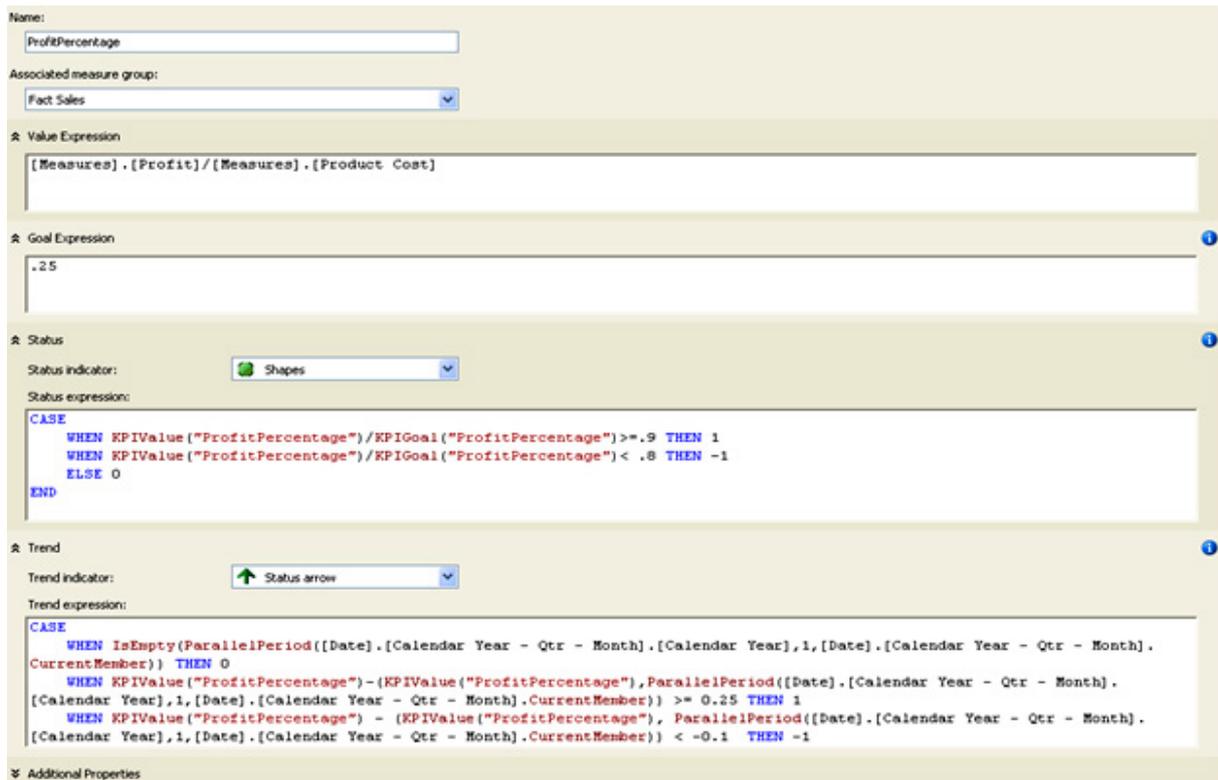


Figure no. 2. Visual Studio 2008 – KPI Registration Card – insight into project

The KPI was finalized in this moment and next the cube has to be executed. In Figure no. 3 the KPI built with the help of the KPI browser from BIDS is presented. The KPI can, nevertheless, be seen in Microsoft Office 2007 or Excel 2010.

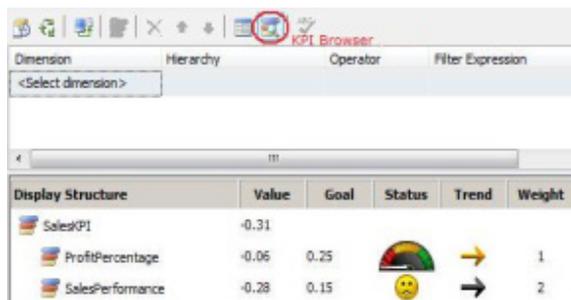


Figure no. 3. Visual Studio 2008 – KPI Registration Card – Layout representation

#### 4. Data Mining

Usually, Data Mining (sometimes called Data or Knowledge Discovery) is the process of data analysis from various perspectives and their synthesis into useful information that are used for the increase of incomes or the decrease of costs (Microsoft Corporation, 2010b). The software for Data

Mining is an option from the series of analytical tools for analysis of the data. It allows the users to analyze, to classify data on many categories from many measurements or different angles and to identify links. From a technical point of view the Data Mining process represents the discovery of connections or models between dozens of fields in the big relational data banks (Microsoft Corporation, 2010b).

Even though Data Mining is a relatively new notion, the technology is not new. The companies have used for years high performance computers to see and analyze market studies with large volumes of data. Of course, the continuous innovations in what concerns computing capacity, hard disk and statistic software have drastically increased the accuracy of the analysis while the costs have decreased.

Data Mining is first of all used these days by companies that put the accent on the user, like retail, financial services, communications and marketing organizations. Data Mining makes it easier for these companies to determine the links between the internal factors like price or

product positioning, or the qualification of the personnel and external factors like the economic indicators, competition as the effect on the turnover, the level of customer satisfaction and the profits of the companies.

- *SQL Server Data Mining for cloud* – Microsoft SQL Server Data Mining for BI Labs Clouds project is a collection of services of Data Mining based on cloud having as foundation the Data Mining platform as part of SQL Server Analysis Services (Microsoft Corporation, 2010b). These services serve a series of machines of Analysis Services which are configured for Data Mining on the client data. The services are exposed as Windows Communication Foundation services in cloud. The users can access these services through a client based on the browser which sends the client data to servers and accesses forecasts.

- *Data Mining with Power Pivot 2010* – Excel offers a convincing interface and common for Microsoft Data Mining. With new functions through Power Pivot,

business users can use a well-conceived infrastructure of the Microsoft technologies. This presentation is user friendly and offers interactive demos which activate Data Mining through these technologies.

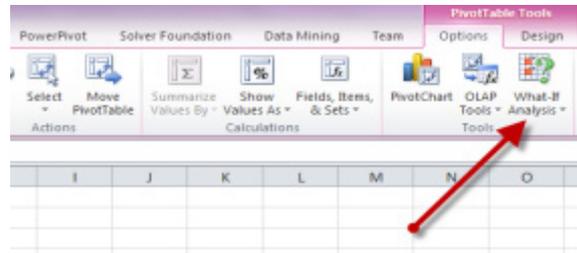


Figure no. 4. Data Mining with Power Pivot 2010

- *Microsoft SQL Server 2008 Data Mining-Add-Ins for Microsoft Office 2007* – Microsoft SQL Server 2008 Data mining-Add-Ins for Microsoft Office 2007 helps users to exploit the advantages of SQL Server 2008 Predictive Analytics in Office Excel 2007 and Office Visio 2007 (Microsoft Corporation, 2010b).

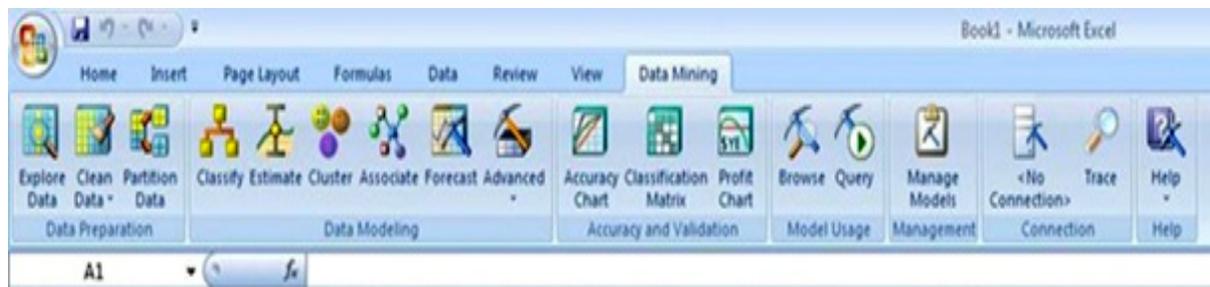


Figure no. 5. Microsoft SQL Server 2008 Data Mining-Add-Ins for Microsoft Office 2007

The package contains the following components (Microsoft Corporation, 2010b):

- The Analyze-Tools table for Excel: this Add-In offers friendly tasks with users which are operated by SQL Server 2008 Data Mining for significant analysis in spreadsheet in Excel data;

- Data Mining-Client for Excel: with this Add-in, Data Mining models in Excel 2007 can be built, tested, examined and managed or with spreadsheet or with external data through the SQL Server 2008 Analysis Services instance;

- Data Mining models for Visio: this

Add-in eases the construction and distribution of Mining models as Visio 2007 drawings.

## 5. Conclusions

The BI notion is sometimes used as a synonym for competitive intelligence because both support decision making.

BI technologies, processes and applications most usually analyze structured internal data and business processes, and competitive intelligence collects, analyzes and spreads information

with a theme target in what concerns the competitors of an enterprise.

The BI solutions have a growing demand because the data available are in fast expansion and the companies look for analysis and knowledge that can bring them

a competitive advantage in a hard economic environment.

BI has the purpose of supporting better business decisions or a better decisional process.

## REFERENCES

Kemper, H.G., Mehanna, W. & Unger, C. (2004). *Business Intelligence Grundlagen und Praktische Anwendungen*, Vieweg.

Microsoft Corporation. (2010a). *Online Documentation SQL Server 2008*, Microsoft Corporation.

Microsoft Corporation. (2010b). *Microsoft Trainings – und Zertifizierungsteam: Entwerfen und Implementieren von OLAP Lösungen mit Microsoft SQL Server 2008*, Microsoft Corporation.

Schultz, M.B., Knuth, J. & Pruß, V. (2008). *Microsoft SQL Server 2008 Reporting Services – Das Praxisbuch*, Microsoft Press.