Techno Science
International Scientific
Journal of Technical
Sciences

OLAP Technology and Its Application through Business Data Analysis and Database Multidimensionality

MAID OMEROVIĆ, ALJO DELIĆ, AHMED PALIĆ

FACULTY OF TECHNICAL STUDIES, UNIVERSITY OF TRAVNIK, BOSNIA AND HERZEGOVINA

ABSTRACT

Quality and timely information is the basis for making business decisions by the company's management. In their day-to-day work, managers are surrounded by an abundance of different data and information that can be more or less useful. It is only by using modern software tools for business data analysis that managers can have quality and fast insight into business activities. Analytical data processing systems *(OLAP)* enable managers to, independently, without programming knowledge, and in real time, preform a multidimensional data analysis, while obtaining information in clear tabular and graphic forms. Likewise, the significance of this paper lies in presenting the practical application of analytical data processing tools on an example of a business activity related to calculating incomes.

Keywords: data analysis, data warehouse, query, OLAP cube, multidimensional approach, pivot table, pivot chart

OLAP Systems

Indicating the significance of OLAP technology, E. F. Codd, PhD, states that:
"Structure without corresponding operators or inferencing techniques is rather like anatomy without physiology."

Online Analytical Processing (OLAP) is intended for online analyses and reporting, and it should enable the end user to ask any business question, use any data from the company for analysis, as well as provide the option of unlimited reporting. They enable a simple synthesis, analysis and consolidation of data. They are used for an intuitive, fast and flexible manipulation of transaction data. They support complex analyses conducted by analysts and enable data analysis from different perspectives (business dimensions). *OLAP* systems as data warehouses use multidimensionality and denormalisation. Data are placed in the so-called data cubes where each side of the cube is called a dimension (data category). The *OLAP* system interface enables the user to independently perform analytical operations and obtain an overview and business graphics without being familiar with programming and the database structure. According to their architecture, *OLAP* systems are divided into: *MOLAP* – multidimensional *OLAP*, *ROLAP* – relational *OLAP* and *HOLAP* – hybrid *OLAP. MOLAP and ROLAP* differ in the way they physically store data. In MOLAP systems, data are stored in a multidimensional structure (data cube), while in the case of *ROLAP* systems, data are stored in relational databases. *ROLAP* systems are

open systems since they provide direct access to data from tables, and there is no need for data duplication. The advantage of *MOLAP* systems lies in the fact that they provide excellent system performance when it comes to previously calculated data (aggregations).

Preparing Data for Business Analysis

In the previous part of the paper, we have seen that analytical data processing can be conducted directly on the relational database *(ROLAP)* or over the prepared data warehouse *(MOLAP)*. At this point in the paper, we will present the manner in which the data warehouse is formed by using the programme Microsoft Query, with the goal to conduct an analysis of the data prepared in such a way.

Prior to forming a query, it is necessary to primarily know what kind of data (information) we wish to receive on the basis of the observed business process *(Income calculation)*. Having in mind the available structure of the *income calculation* database, and for the company's management purposes, we can request, among other things, the following synthetic data: an overview of summed income per years and months, an overview of summed income per qualifications, departments and municipalities, an overview of average income per qualifications and departments, income structure (income on the basis of the labour cost and the time spent at work, compensations, remunerations) per months, departments, cost centres, remuneration structure

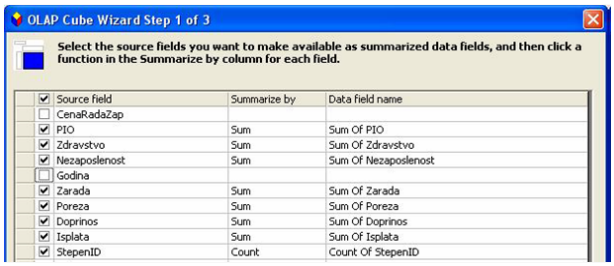


Image 1. Selecting Summarised Data Fields

Cijena rada zaposlenika – employee labour cost,
PIO – pension and disability insurance,
Zdravstvo – health care,

Nezaposlenost – unemployment, Godina – year,
Zarada – income, Poreza – taxes, Doprinos – contribution,
Isplata – payment, Stepen ID –Degree ID

(night work, work during holidays, overtime, ...) per departments, divisions, etc.
The above stated examples of the requested data can be combined, obtaining a much higher number of possible data which can be of interest for the observed business process. Likewise, other reports can be generated by introducing new dimensions and aggregations.

Creating a Data Warehouse - *OLAP CUBE*

Creating *OLAP* cubes as data warehouses accomplishes the following advantages in data analysis: easier access to large databases by working with their subsets, a reduction of time necessary for data processing by performing the necessary data aggregations in advance, facilitated creation of reports due to the previously defined dimensions and levels of data overview.

Based upon the created query, we can create *OLAP* cubes by activating the option *CreateOLAP Cube* from the MS Query menu. Creating *OLAP cubes* is conducted in three steps:

1. ***Selecting Summarised Data Fields.***
   The fields forming summarised data and mathematical operations which will be conducted on the fields are determined. The fields which are not used for summarised data represent the candidates for cube dimensions.
2. ***Defining the Dimensions of the OLAP Cube.***
   Selecting the fields which will form the cube dimensions out of the remaining ones (candidates for the dimensions). In this step, the dimension hierarchy is also formed.
   The dimension hierarchy should enable multiple levels of specificity, depending on the user's needs in the process of data analysis.
3. ***Selecting the type of the OLAP Cube.***
   The option of keeping a cube file with all the data was selected, forming the data warehouse. When saving, a special file of the *.OQY (OLAP query)* type keeps the information on the defined aggregations and cube dimensions, and, in an individual file of the *.CUB* (cube) type, the data warehouse is kept for the needs of the multidimensional analysis.

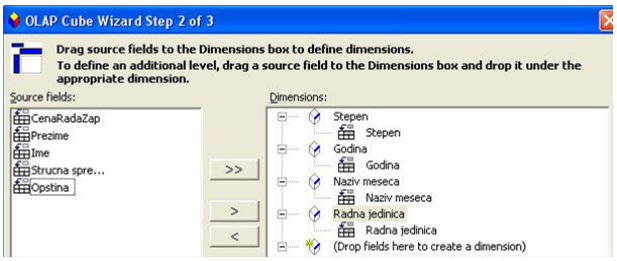


Image 2. Defining the Dimensions of the OLAP Cube.

Cijena rada zaposlenika - employee labour cost,
Prezime – last name. Ime – first name,
Stručna sprema – qualification,
Opština – municipality,

Stepen – degree,
Godina – year,
Naziv mjeseca – month name,
Radna jedinica – department

Data Analysis Management in Excel

Upon preparation, the data analysis management can be performed in the Excel programme by applying *Pivot tables and Pivot charts*, where Excel is attributed the role of *OLAP* interface.

A pivot table serves as a tabular representation of multidimensional data within which summarised data can be shown at any level of specificity.
The user has the option to form reports for a certain point of view reading the data (the selected level of specificity and dimension arrangement). Creating pivot tables is performed through the menu *Data/ PivotTable and PivotChart Report*. By activating this option, we start the Wizard and select the data source, type and data file, as well as the position of the pivot table in the Excel worksheet.
Namely, there is a choice between three different types of databases as source data for the analysis by way of a pivot table:

• Databases – selecting one of the provided relational databases with a collection of primitive data sorted into tables (normalised data);

• Queries – selecting a query formed in the MS Query programme (denormalised table);

• OLAP Cubes – selecting a cube file with the defined summarised data and candidates for the dimensions

The first two options, *Databases and Queries*, represent the application of a *ROLAP* system where there is no formation of a data warehouse and the "live data" are accessed from the database. By selecting the option *OLAP Cubes*, we perform a multidimensional analysis over the data warehouse, representing the *MOLAP* system.

Image 4. provides the appearance of a formed pivot table where the employees' income distribution is presented per months and qualifications.

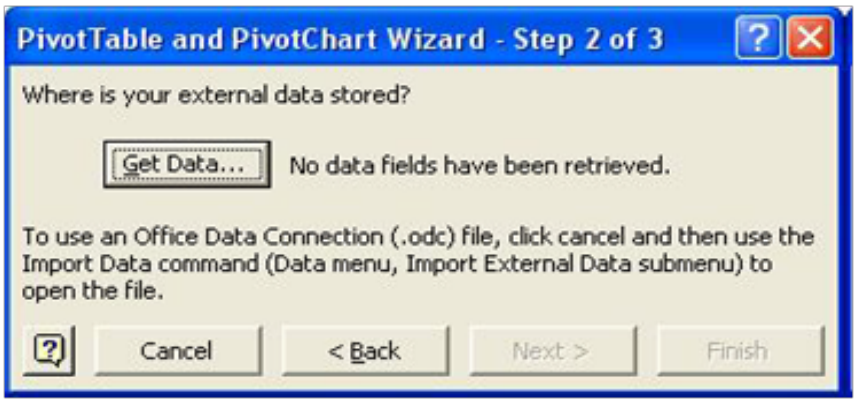


Image 3. Selecting the Data Source for the Analysis

Table 1. Pivot Table of the Income Distribution per Months and Qualifications

Year	2005							
Sum of Income	Degree							
Month name	**NK**	**KV**	**VKV**	**SSS**	**VŠS**	**VSS**	**MR**	**Grand Total**
January	30,775.50	36,485.00	20,340.00	52,972.50	42,162.00	149,265.00	58,185.00	390,160.50
February	33,340.13	39,503.75	22,035.00	57,386.66	45,675.50	161,698.88	63,033.75	422,673.88
March	24,368.75	42,542.50	23,730.00	61,801.25	49,189.00	174,137.25	67,882.50	443,651.25
April	25,413.13	44,365.75	24,747.00	64,449.88	51,297.10	208,537.28	70,791.75	489,601.88
May	27,858.00	48,620.00	27,120.00	86,630.00	56,216.00	228,534.00	77,580.00	552,550.00
June	26,979.69	47,100.63	26,272.50	83,922.81	54,459.25	221,392.31	75,155.63	535,282.81
July	27,153.75	47,404.50	26,442.00	84,464.25	54,810.60	222,820.65	75,640.50	538,736.25
August	28,242.34	49,304.93	27,502.05	87,850.40	57,007.94	231,753.47	78,672.91	560,334.05
Grand Total	**224,123.27**	**355,307.06**	**198,188.55**	**579,477.96**	**410,817.39**	**1,598,134.34**	**566,942.03**	**3,932,990.61**

A graphic overview of the data from the table can often be a better way of providing an insight into the data, and it can help in analysing data and drawing conclusions. When working with pivot tables through the *PivotTable toolbar*, clicking the option *Chart Wizard* automatically provides a ready-made chart, predefined as a *Column type.*

Image 5. provides an example of a pivot table with an overview of the employees' average and total income per departments and qualifications.
The data are sorted according to the average income per department in a descending order - from the highest to the lowest.

Image 5. Pivot Table of the Total and Average Income per Departments and Qualifications

Year	**2005**								
Month	**(All)**								
		Degree							
Department	Data	**NK**	**KV**	**VKV**	**SSS**	**VŠS**	**VSS**	**MR**	**Grand Total**
Company management	Average of income2						32,148.09	37,982.49	**34,092.89**
	Sum of income						514,369.45	303,859.88	**818,229.33**
Department of development	Average of income2						27,367.85	32,885.27	**30,126.56**
	Sum of income						218,942.81	263,082.15	**482,024.96**
Finances and commerce	Average of income2						29,070.58		**29.070.58**
	Sum of income						232,564.62		**232,564.62**
Manufacture	Average of income2	13,058.30	22,206.69	24,773.57		25,372.81	30,455,54		**23,326.96**
	Sum of income	130,582.95	355,307.06	198,188.55		202,982.49	395,921.99		**1,282,983.05**
Accounting	Average of income2				21,617.50	25,979.36			**23,071.45**
	Sum of income				345,879.95	207,834.90			**553,714.85**
General affair	Average of income2	11,692.54					29,541.93		**20,617.24**
	Sum of income	93,540.32					236,335.46		**329,875.78**
Transportation	Average of income2				19,466.50				**19,466.50**
	Sum of income				233,598.01				**233,598.01**
Totak Average of income2		**12,451.29**	**22,206.69**	**24,773.57**	**20,695.64**	**25,676.09**	**30,153.48**	**35,433.88**	**25,374.13**
Total Sum of income		**224,123.27**	**355,307.06**	**198,188.55**	**579,477.96**	**410,817.39**	**1,598,134.34**	**566,942.03**	**3,932,990.61**

A Summary of the OLAP Technology Application-Related Analysis

This paper provides an overview of the business data analysis by applying an OLAP system on an example of a business activity related to calculating the employees' income, which is mandatory for all legal entities.

The activity of calculating incomes is an especially sensitive issue in companies which have limited funds at their disposal for payments on a monthly level: companies under the jurisdiction of municipal authorities, public companies and budget users.

Due to the legal form of calculating incomes in the part related to compensations, remunerations for night work, work during holidays and overtime, abuse on part of some employees can occur, directly reflecting the income of other workers since the funds are limited. In such cases, data analysis is of special significance for the management on both the synthetic and analytic levels.

The above mentioned issue in the part related to income calculation can be interesting from the aspect of a deeper, highly specialised analysis which can serve as the subject of a future paper.

Considering the spatial limitation, this paper depicts one of the possible approaches in the data analysis management. It should be noted that, in the business process of income calculation, a legal form dominates and has to be fulfilled both in the calculation and in the system of standardised, prescribed reporting.

Analytical data processing with the goal of providing the company's management with a more complete illustration of the structure and cost allocation from different aspects represents one upgrade.

It is possible to further expand and develop it by introducing new attributes on a primitive level and by applying other analytical processing systems *(Data mining)*.

Database multidimensionality

Data warehouse is based on a multidimensional model which represents data in the form of the so-called n-dimensional data cubes.
Data cubes are defined with:

• a fact table: mostly contains numerical data and foreign keys to dimension tables;

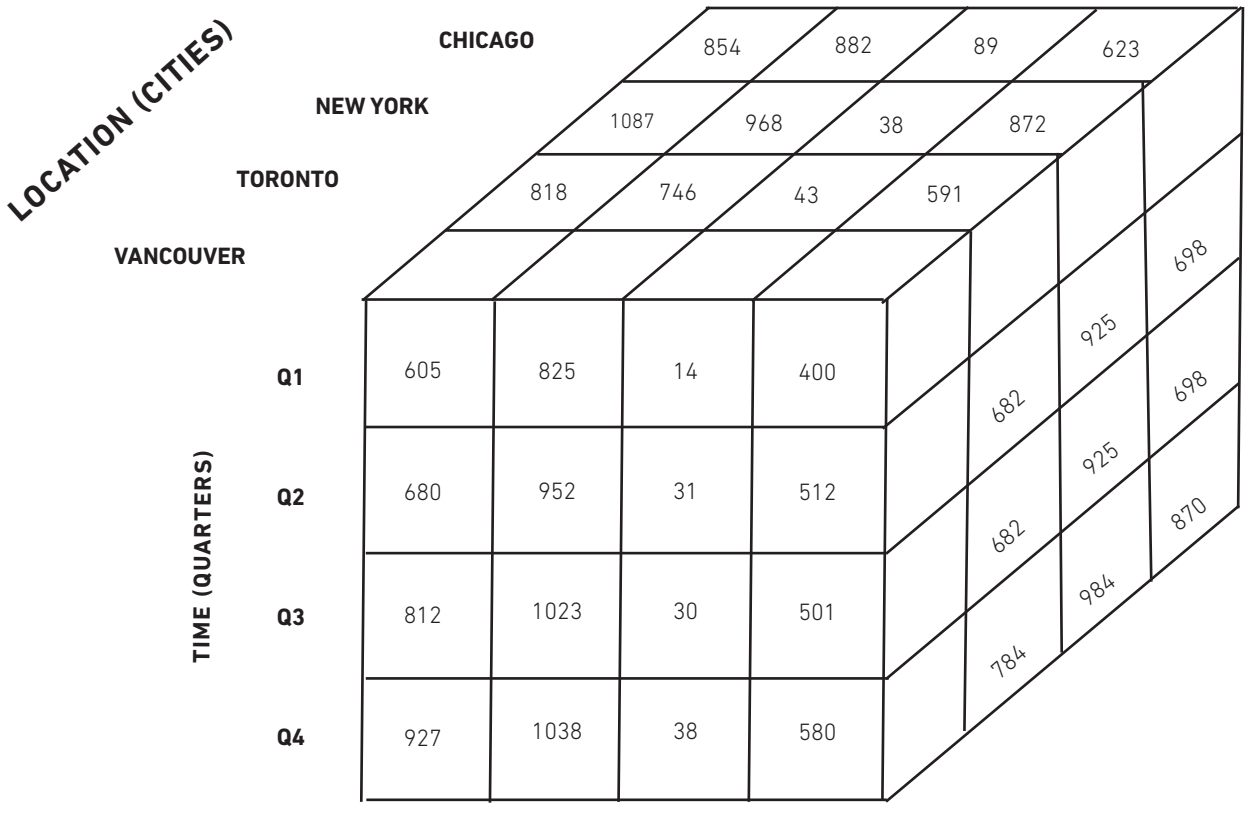
• dimension tables: enable a representation of data in several ways.

2-D data cube

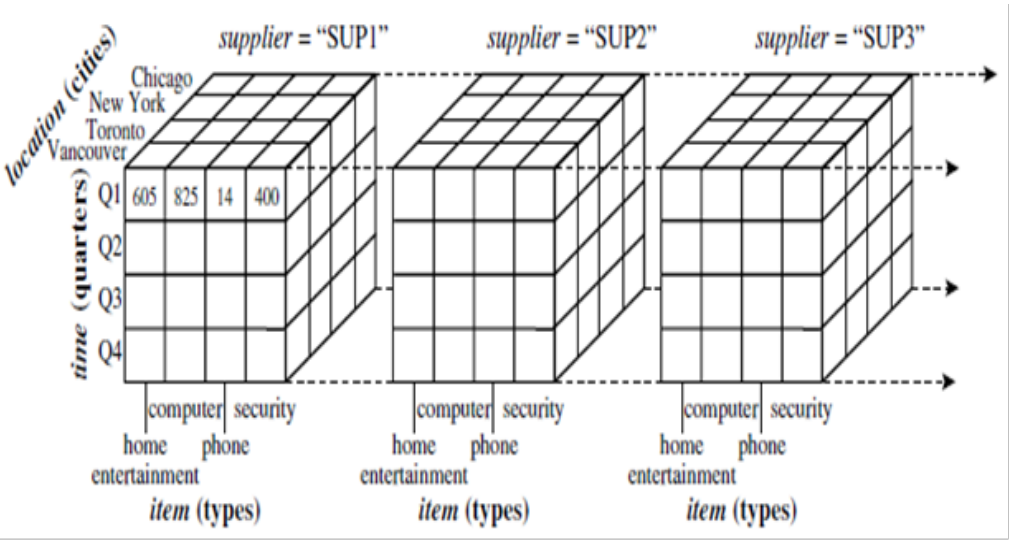
LOCATION = "VANCOUVER"				
TIME (TYPE)	ITEM (TYPE)			
	HOME ENTERTAINMENT	COMPUTER	PHONE	SECURITY
Q1	605	825	14	400
Q2	680	952	31	512
Q3	812	1023	30	501
Q4	927	1038	38	580

3-D data cube

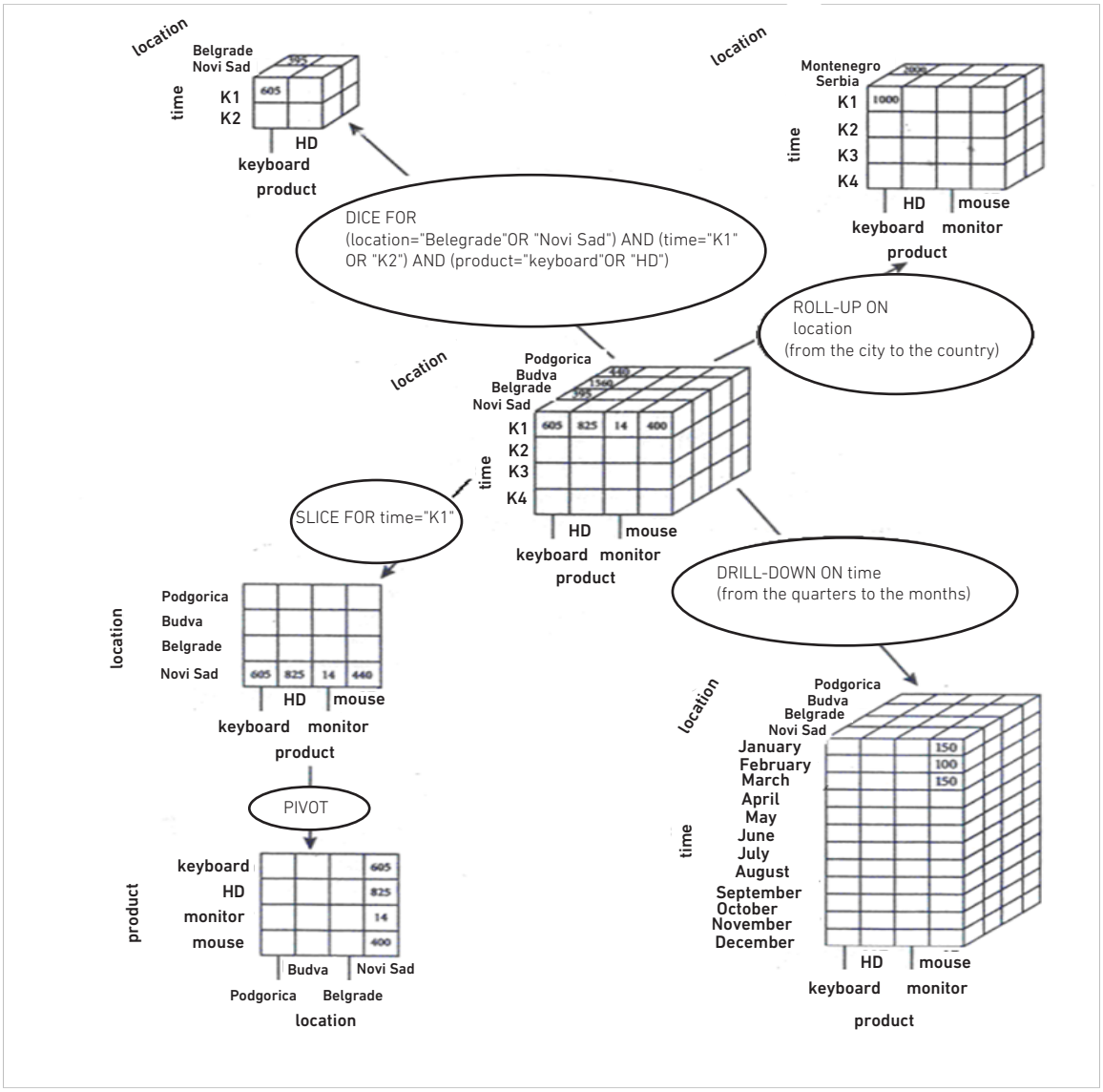
	location = "Chicago"				location = "New York"				location = "Toronto"				location = "Vancouver"			
	item				item				item				item			
	home				home				home				home			
time	ent.	comp.	phone	sec.	ent.	comp.	phone	sec.	ent.	comp.	phone	sec.	ent.	comp.	phone	sec.
Q1	854	882	89	623	1087	968	38	872	818	746	43	591	605	825	14	400
Q2	943	890	64	698	1130	1024	41	925	894	769	52	682	680	952	31	512
Q3	1032	924	59	789	1034	1048	45	1002	940	795	58	728	812	1023	30	501
Q4	1129	992	63	870	1142	1091	54	984	978	864	59	784	927	1038	38	580



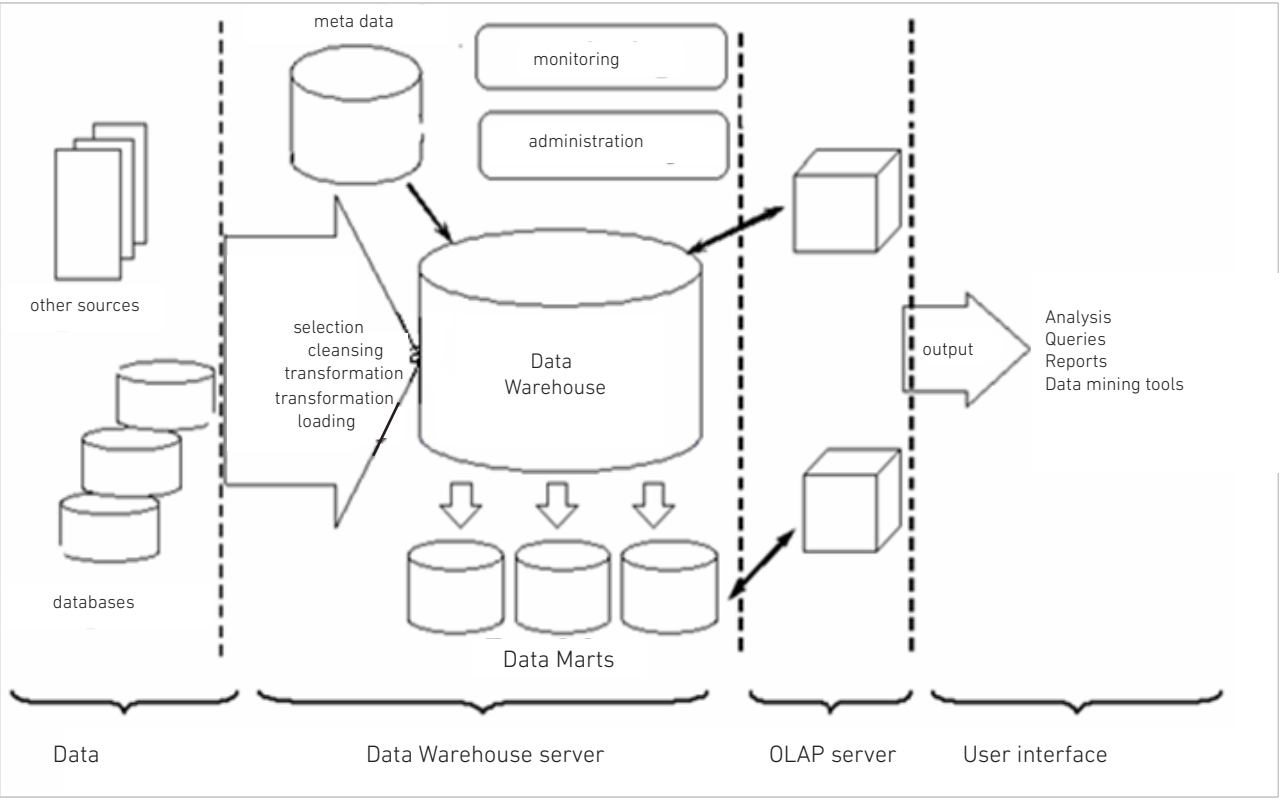
4-D data cube



5-D data cube



Data Warehouse System Architecture



Pivot Tables and Microsoft Analysis Services

The notion of pivot tables implies a tool for performing different types of aggregations according to different levels and, most frequently, we encounter them as an additional strength in the programmes for cross calculations, typically represented by MS Excel. Pivot tables can easily be transformed into visual reports. By applying pivot tables, one can materialize their effort in forming data structures and processing that data by delivering an effective report.
In fact, a pivot table is a technical term for one of the implementations related to the *OLAP* concept which is, in everyday life, drawn nearest to the user. There are hardly any people who spend their working day without an installed programme for cross calculations.

A pivot table illustrates aggregations related to a data set. If you make a change in the data, the pivot table should reflect it. The changes will not happen automatically. You need to select the whole area of the pivot table, right-click somewhere on the surface of the pivot table and, from a context menu, instruct the pivot table to refresh the data.
Rest assured that the pivot table mechanism will perform with complete accuracy. The data will be modified in accordance with the changes made.

```
Sub RefreshData()
    Dim iP As Integer
    Application.DisplayAlerts = False
    For iP = 1 To ActiveSheet.PivotTables.Count
        ActiveSheet.PivotTables(iP).RefreshTable
    Next
    Application.DisplayAlerts = True
End Sub
```

If you are prepared to deal with macros, then the following part of the programme code could be of benefit:

Microsoft offers a significantly stronger weapon in implementing the *OLAP* concept. The solution is called Microsoft Analysis Services. In line with the complete demands of the *OLAP* concept, Microsoft provides an *OLAP* solution in the form of a classic *OLAP* sever.

Within the relational database Microsoft SQL server 2005, a software placed at the central layer of the three-tier architecture is delivered.
The commercial name is *Microsoft SQL Server* Analysis Services *(SSAS).* This product enables a complete *OLAP* and Datamining functionality which is demanded of a product belonging to the business intelligence applications.

Microsoft SQL Server Analysis Services *(SSAS)* provides the application of *OLAP* by allowing the user to design, create and manage multidimensional structures containing aggregate data taken from other data sources, primarily transaction relational databases.

Conclusion

OLAP enables a simple synthesis, analysis and consolidation of data. They are used for an intuitive, fast and flexible manipulation of transaction data. They support complex analyses conducted by analysts and enable data analysis from different perspectives (business dimensions).

The example of applying OLAP technology in practical life is clearly and precisely stated which is most important nowadays, especially its application in business activities where each of us should at least know the basics. Afterwards, the database multidimensionality is illustrated in a very interesting way.

We can clearly conclude that we have accomplished the goal of this paper which is to demonstrate to the reader, without providing superfluous data and information, what the multidimensional database and OLAP technology really are.

OLAP is not an expense, OLAP is a necessity.

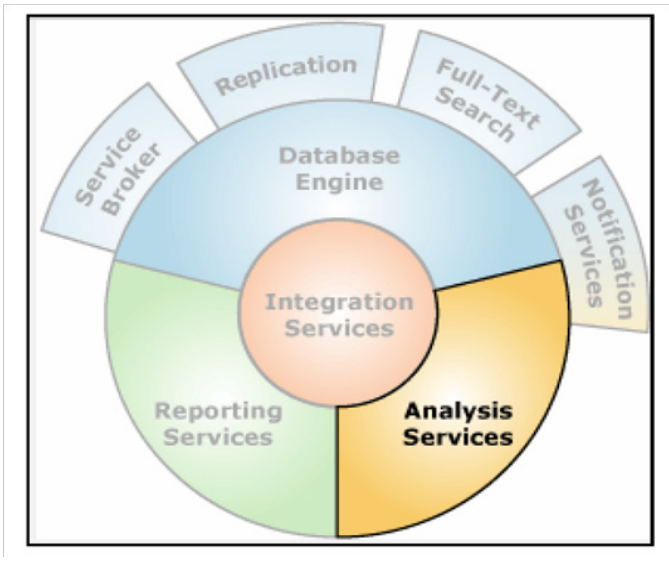


Image 6. Microsoft Analysis Services

References

1. Business Data Analysis Using OLAP Systems, International Scientific Conference MANAGEMENT2010, Kruševac (Serbia), 17th – 18th March 2010.
2. Lectures from FER (Faculty of Electrical Engineering and Computing, Zagreb), link: https://www.fer.unizg.hr/_download/repository/OSP1_Skocir.pdf, University of Zagreb, Faculty of Electrical Engineering and Computing (Accessed 24thJuly 2019).

Techno Science
International Scientific
Journal of Technical
Sciences

OLAP tehnologija i primjena istih kroz poslovnu analizu podataka i višedimnezionalnost baze podataka

SAŽETAK

Kvalitetna i pravovremena informacija je osnov za donošenje poslovnih odluka od strane menadžmenta preduzeća. Menadžeri su u svom svakodnevnom radu okruženi mnoštvom različitih podataka i informacija koje mogu biti manje ili više upotrebljive. Samo uz korišćenje savremenih softverskih alata za poslovnu analizu podataka menadžeri mogu imati kvalitetan i brz uvid u poslovanje. Sistemi za analitičko procesiranje podataka *(OLAP)* omogućavaju menadžerima da, samostalno, bez znanja programiranja, u realnom vremenu, obavljaju višedimenzionu analizu podataka dobijajući pritom informacije u preglednim tabelarnim i grafičkim formama. Značaj ovog rada je isto tako u prikazu praktične primjene alata za analitičku obradu podataka na primjeru poslovne aktivnosti obračuna zarada.

Ključne riječi: analiza podataka, skladište podataka, upit, *OLAP* kocka, višedimenzionalni pristup, pivot tabela, pivot grafikon

Received: November 29, 2018 / Accepted: December 07, 2018
Correspondence to: mr. Aljo Delic, University of Travnik, Faculty of Technical Studies, Travnik,
Bosnia and Herzegovina
E-mail:aljo.delic@sapna.ba