ANALYSING BANK DATA USING OLAP

Eva Mièovská
Investment and Development Bank

Introduction

Database applications in bank systems solve two basic spheres of problems:

1. to supply daily bank service (e.g., to maintain client accounts)

2. to perform a huge analysis for obtaining details for evaluation of a bank activity and for decision making

The first mentioned domain of problems represents a task for transaction processing in OLTP systems (On Line Transaction Processing) [1]. Recording new balance on client accounts in a bank is one example of this type of transactions.

A bank service can represent a work of several applications, where every one of them can operate on an independent OLTP database. For example, in our bank all data about client accounts and data from The General Ledger are processed by two applications and stored in separated databases.

The second mentioned problem is a task for DSS systems (Decision Support System). This kind of systems solves a role of data synthesis from various input databases and summarisation of these data according to required criteria. This represents a base level for preparing details for large data analysis, which should be executed in real time (e.g., a balance on risky loan accounts, a comparison of expensive and cost accounts in the whole bank, individual branches or cost centers).

The original idea of RDBMS (Relational Database Management System) was to provide a solution of both mentioned tasks on the same database. At this time, Commercial DBMSs do not offer adequate support for simultaneous resolution for both problems.

This situation led to a necessity to use a different access to data presentation and their performance. This was a beginning of multidimensional databases, which allow to prepare multidimensional views on the same data. OLAP systems (On Line Analytical Processing) are implementation of a multidimensional data access.

In the first part of this paper we will shortly say something about applying OLAP criteria in the SAS® environment. Hereby we will mention about various data analysing components in this environment. Next we will concentrate on application development phases and in the last we will describe some of our bank applications in more details.
1. OLAP in the SAS environment

E. F. Codd defined 12 rules for evaluating OLAP products [2]. SAS/EIS Motore® arose as an extension of the SAS module EIS®. Its main rule is to provide rapid development tool for OLAP applications. Some of the basic Motore attributes are:

- it makes possible prepared and presumarised data to cut through with defined dimensions, i.e. with multidimensional data views
- it contains tools for mapping a logical and physical database structure on individual sources; it also contains information about the origin of data stored in database and about all performed conversions to unified form
- it contains so-called control objects; these objects note down position information about data space; user can move through data following defined dimensions or across these dimensions without loosing previous data information
- a user with Motore applications gets grateful tool for output generating; he can convert an output design during the running of an application (a vertical or horizontal graph with 3D effect or without it, a pie graph, a bubble graph, a graphic table, a report); he can prepare printed outputs or save contents of actual screens for returning to them later; Motore allows to enter to some other environment (e.g., Excel) and return back without closing original application later; Motore also enables to change default colours and fonts

In the following part we would like to explain several input components of a creating data analysis process. Hereby we will mention about applying some OLAP criteria.

1.1. Input components of a data analysis process

Before creating various analytical surveys and comparisons it is needed to manipulate the data, that is to prepare the transaction data for analysis. The Investment and Development Bank (IRB) uses the bank transaction system Profile, which is based on the MUMPS database and for DSS development was chosen SAS system. Data transfer from MUMPS to SAS environment is made through text extracts. In this environment they are processed to a required form. Let us concentrate on Information base, SAS/EIS Motore, Metadata and Metabase as individual elements of application development process in the SAS system.

1.1.1. Information base

The word database is a general name for a place where data are stored under regular rules, so we could manipulate with the data anytime we need (e.g., to read and to change them, to add new data or to delete useless ones). During DSS development we store in advance prepared data into the database. This way created database we called Information base (further we will use only the abbreviation IB).

Applications developed in the SAS system enable to present single data in user terms. This IB attribute is provided by formats. Formats are connected to appropriate data and stored out of the data table. Data storing by this way saves a disk memory and accelerates work with data. It also allows to display them in the way required by user what the system supplies automatically.
The optimization OLAP rule is applied in a data consolidation process. This process consists of a preparation input data in advance and creation two types of tables: base and summary tables. A *base table* contains the most detail data. *Summary tables* that are derived from the base table, include some of summarised input data combinations so that they do not have to be calculated in runtime. This is the way how the data consolidation process forces a movement through dimensions.

1.1.2. SAS/EIS Motore and OLAP

Thereinbefore we mentioned that the main Motore role was to provide a rapid development tool for OLAP applications. *Motore* (OLAP Engine) enables to connect consolidated data with defined dimensions quite easily. Dimensions in a completed application are displayed by special objects to navigate a movement through multidimensional data space. Motore is set up in order to access and provide data and application displays them. We can say Motore represents an intermediary between requirements coming from a screen and matching data.

Not less important Motore role is to supply a performing required values calculations (they could be defined by statistic formulas) and display them in runtime from any dimension level. This is the main idea of the Generic Dimensionality OLAP rule.

1.1.3. Metadata and metabase

During DSS system it is necessary to pay attention not only on input and stored data. Storage of actual and historical information about data is also very important. This information, so called *metadata*, represent a link between various components of IB.

Information stored in IB can be divided into several following groups:

- a structure and a content of IB
- connections between input and output data
- extracts and transformations history
- information about summary tables
- data relationship information
- defined dimensions and their hierarchy
- data and file access and ownership
- data origin history and table updating

*Metabase* is a place where metadata are stored. The SAS system offers tools for automatic metabase manipulation. It means that every modification in IB causes corresponding changes in all relative metabase items. Metadata maintenance is the key element of all data life cycle in IB.

Metabase and metadata satisfy OLAP rules about accessibility and consistent reporting performance in OLAP products.
2. Development and presentation of bank applications

The bank transaction database represents an important source for analyses of the state of a bank and its branches, analyses of individual products or a structure of clients. Analyses of data depend on the possibility to change the data views. We can use the same data according to various aspects or views (e.g., a view according to a time, an account balance viewed by individual foreign currencies, a geographical view, a view according to account numbers). In the following part we will discuss generally about individual development phases of building applications. Then we will focus on several examples of prototype applications built in the IRB.

2.1. Application development phases

Application development process with the use of the SAS/EIS Motore can be divided into several phases:

• **a problem specification**

Although every phase of the development process has a key importance, we would like to emphasise the role of the first one. The misleading in the initial specification of the problem causes the fact, that the developers are later forced to rebuild the whole development process from the beginning.

• **data transformation**

This phase consists of transferring data from a transaction database into the SAS environment. After transferring the data enter a process of a transformation to obtain the structure needed for the final application. An example of the transformation data process is discussed in details in my colleague’s paper called ‘Preparation of the Balance Sheet for OLAP applications’.

• **data consolidation**

This phase includes a creation of base and summary tables and their optimalization according to the need of disk space and runtime.

• **Motore definition**

Motore definition ensures the direct connection between defined dimensions and consolidated data.

• **screen design**

In this phase a developer creates the graphic interface of the applications according to the defined Motore.

• **application testing and tuning**

Testing and tuning of applications ensures the full adaptation of the application to the users needs.
Each of the specified phases requires an adequate attention from a developer. Moreover, it is necessary to respect goals of the individual phases. On the other hand it is essential to remember that the application should be built as a part of an integrated system. The end user plays an important role mainly in the first and the last two mentioned development phases.

2.2. **Presentation of some bank applications**

The specification phase ran partially with consultants of PHARE project which was orientated on help in development Decision Support System (DSS) in IRB Slovakia. Presented applications represent prototype bank application examples. They were built on data from the transaction database that reflect the General Ledger. By content these data are appropriate for analyses of the following bank products and economic ratios:

- bank loans
- client deposits
- the state of covering bank assets by bank liabilities
- Balance Sheet
- Profit and Loss Account

Our bank applications have been developed in the SAS system V6.11 and SAS/EIS Motore V4007. Their presentation shows several applied ways of customisation Motore to obtain a required effect. This type of customisation is e.g., a calculation of a variable value in runtime or preventing a value presentation even though it really exists among chosen dimension data.

All text variables are displayed through formats. That means the texts are not components of base and summary data tables.

In the next part we will describe individual presented applications (used data are unreal).

2.2.1. **Daily balance of loan accounts**

The main aim of this application is to provide an effective tool for loan account analysis according to the General Ledger. We defined four base navigation classes represented by the following **dimensions** for moving on data:

- time (day)
- region (IRB, main branch, cost center)
- currency
- loan type (long term loan, middle term loan, short term loan,…)

The application has been built from two overlapping frames. From **visualisation classes** have been chosen:

- vertical 3D chart for a graphical representation (Figure 1)
- Graphic Table which represents the same data in table form as the previous vertical 3D chart (Figure 2)
2.2.2. Changes in the balance of deposits within 6 days period

This application is specified for analysis a balance of the General Ledger deposit accounts within a period of the previous 6 days. Unlike the previous application this one displays a value of 6 variables all at once. Moving on data is provided by the next dimensions:

- region (IRB, main branch, cost center)
- currency
- deposit account type (term account, current account,…)

The application has been created by two overlapping frames with the following visualisation classes:

- vertical chart without 3D effect for a graphical representation (Figure 3)
- Graphic Table representing the same data in the table form as the vertical chart with 6 data columns

![Figure 3](image-url)

2.2.3. The state of covering of bank assets by liabilities

The important measure of the state of the bank is a relation between bank assets and liabilities. Movement on data is managed by three defined dimensions:

- time (day)
- region (IRB, main branch, cost center)
The first screen of application is created by two frames with next visualisation classes:
• vertical chart without 3D effect shows the balance of loans and deposits on the right side of the screen (Figure 4)

• Critical Success Factor located on the left side of the screen, displays the percentage of relation between loans and deposits shown in the chart on the right

The number showing this relation is calculated in runtime from displayed data about loans and deposits. It was neccessary to do the particular interference to generated Motore code.

This introductory application is connected through displayed buttons with two extra applications. Each of them enables detailed analysis of account balances of the General Ledger that concern on separated analysis of loans or deposits. Each of these applications is connected with another one which enables a view of analysed data in a pie chart (Figure 5). Both pie charts are interesting for displaying data which are concerned only on foreign currencies although the data tables also contain data for home currency. It was neccecssary to make changes in generated Motore Code for gaining this effect.

### 2.2.4. Balance Sheet

The application represents one of the most important economic ratios of the state of the bank. The basic prescription for this type of economic sheet is prescribed by The National Bank. For the needs of the bank management one more summarised level of the official prescription has been prepared. It is possible to expand from the highest summarisation level to more detailed information. The part of the first screen of the application is a button for opening the application which enables the movement from certain lines of the Balance Sheet down to suitable the General Ledger accounts. This enables the detailed view of bank Balance Sheet.

The application enables to create the Balance Sheet in runtime not only for all IRB, but also for any main branch or cost center. Everything is made for each of the previous six days separately. The movement on data is defined by the following dimensions:

• time (day)
• region (IRB, main branch, cost center)
• summarisation level
• a structure of the General Ledger accounts

Chosen visualisation class:

• Graphic table

### 2.2.5. Profit and Loss Account

Another important economic ratio of a bank state is the Profit and Loss Account. In this application like it is in Balance Sheet it is possible to expand from the highest levels of summarisation into more detailed levels. Also this application offers a chance to show the view on the most detailed data in the General Ledger accounts. The Profit and Loss Account can be created for whole bank in runtime or for its any branch or cost center for previous six days. The movement on data is defined by the following dimensions:

• time (day)
• region (IRB, main branch, cost center)
• summarisation level
• a structure of the General Ledger accounts

Chosen visualisation class:
• Graphic table

Conclusion

Bank transaction databases are excellent source of information for management. Their usage is connected with the opportunity of effective and appropriate access to stored data. It is not only important to know how to collect and store data effectively but also to know how to chose and transform them into required information.

OLAP systems simplified a structure of multidimensional databases. Data coming from heterogeneous sources enter the common process of creating base table and its summary tables. More simple database structure allows to increase a capacity of system dealing with the multidimensional view on analysed data. It is the main assumption for providing complex bank data analyses as the base of the DSS system development.

In IRB we have been working on development of bank applications in the SAS System for almost one year. Although a staff capacity is insufficient (about 2.3 man) we succeeded in gaining the main part of development tasks using the SAS/EIS Motore.

After our practical experience with tools of the SAS System we would like to say that they belong to the tools that provide meeting the requirement of the fast high quality bank manager information. It is done in the way which a user would welcome.

References:

Address for correspondence:
Investièná a rozvojová banka, a.s.
Tallerova 1, 818 55 Bratislava
Slovak Republic
Phone: +421 7 389 3220 Fax: +421 7 323 048

1 The registered name of the IRB bank is: Investièná a rozvojová banka, a.s., Slovak republic
SAS, SAS/EIS, SAS/EIS Motore are registered trademarks of SAS Institute Inc., Cary, NC, USA