3 Methodology of the development of a data warehouse as an intermediate phase in the extraction of knowledge from databases

This chapter includes the practical aspects of modeling and data warehouse implementation. These are described and illustrated with cases of the most important elements of this process, paying special attention to the analysis of requirements of a future user, the scalability, applied tools, reasons for failures and directions of development of effectively-implemented data warehouses.

3.1 Evaluation of requirements, needs and capabilities of a data warehouse user as a key factor for successful implementation

A data warehouse, the process of its development and implementation or adaptation to the needs of a specific customer are particularly complex issues, determined by methodological standards and substantive conditions. Unlike many other IT tools, such as office software packages, accounting or operating systems, data warehouses are products which require effort focused on a specific object of implementation. This means that the largest portion of the costs connected with the development of a data warehouse solution to introduce business intelligence tools in an organization is related to the design, planning and implementation of the work [1]. At present, some off-the-shelf products have been introduced on the market, which aspire to be called industry data warehouses; however, these have a very narrow scope of application and are much more suitable for small organizations with limited information needs. As a matter of fact, each successful implementation of a data warehouse is different and leads to an individual, customised solution intended only for the organization which had it developed. Therefore, a key element — which has a decisive effect on the success or failure of a data warehouse implementation, is a careful analysis of the user’s requirements, needs and capabilities.

It must be stressed that the process of analysis and implementation of a solution like this should equally involve the supplier of the data warehouse and its user. A standard of partnership should be established, as it often determines the success of the entire project. The analysis, as the key element, should be performed before a given product or supplier is selected, and its main result should be a decision based on the choice of the right solution. This often leads to a seemingly increased cost, as the user of the future data warehouse has to work with a few prospective suppliers at such an early stage of the project. However, experience shows that this kind of approach actually results in substantial savings, with regards to time and funds at subsequent stages, which allows for the smooth execution of the project and compensation of the higher expenses incurred at the analytical stage.

There is also an option to perform the analysis with an independent consultancy, which will help the future user to define specific needs and measure them against his/her capabilities, and then make the right choice concerning the solution and the supplier. The SAS Institute – one of the leaders in the data warehouse market – has suggested a five-phase model for the development of a data warehouse [2]:

- Phase 1 – analyzing the profitability and usefulness of the data warehouse project;
- Phase 2 – gathering requirements;
- Phase 3 – modeling and designing;
- Phase 4 – implementation;
- Phase 5 – reviewing.

Most of data warehouse project management models used today are principally in line with the idea of the SAS Institute. However, it is noteworthy that in the case of the SME sector, Phase 2 should be extended to include user preferences and should be followed by the selection of the scope and supplier of the right solution. This kind of approach is determined by the peculiarity of SME businesses where key specialists, who are a means of achieving the business objectives by the company, are often very busy and find it hard to completely alter their habits and preferences as it pertains to IT solutions. People like that may substantially hinder the implementation process, using their position in the company hierarchy to obstruct or even block any changes. Therefore, as early as the analytical stage, it must be checked whether such situations are likely to occur during the implementation stage and whether a compromise can be found to satisfy both the company’s information needs and requirements of the personnel without overturning their daily routines.

In the case of SMEs, other decisive elements are the organizational, infrastructural and financial capabilities of the data warehouse user. These elements must be determined and measured against the user’s needs at the analytical stage, so that a solution with a relevant level of sophistication and appropriate price can be offered. An avoidable – yet frequently occurring – situation is that the solution is selected because of some “good intentions”, on the basis on recommendations, current trends or style, or even because the seller proved to be very persuasive. Making choices on such grounds usually leads to the failure of the project, or – in the best-case scenario – a stressful and lengthy implementation process. This process happens when the chosen solution fails to meet the needs and capabilities of the company.

It is not uncommon that the project budget is only enough to buy the tools but not to cover the costs of implementation or to purchase the necessary equipment, or else a cheap and simplified solution is bought which cannot provide the company with a chance to progress. Except the SME-specific elements of the initial analysis, the first stage should also contain other universal aspects which are included in the first two out of ten stages of a data warehouse life cycle described in Building a Data Warehouse by Vidette Poe, Patricia Klauer and Stephen Brobst [3]:

**Planning:**
- Definition and/or clarification of the project scope;
- Definition of the required technical measures;
- Definition of the participants and responsibilities;
- Definition of the tasks and their expected results;
- Definition of the resources and deadlines;
- Definition of the final effects of the project.

The principles of operation of the data warehouse should also be predetermined (e.g. the strategies of data integration, content refreshing, data archiving, or administration of metadata); that is important, because any adopted solutions affect the technical infrastructure and thus the project costs. Gathering requirements with regards to the necessary data and (initial) modeling, which should provide answers to the following questions:

- What kind of data is currently used and what would the users like to have?
- What is the level of detail and scope of the summary required by the users?
- How would the users like to have their queries answered (in what format) and what kind of access tools would they like to use?
The feedback should provide a basis for a data model. Two solutions are possible here: the development of a logic data model, corresponding to the project scope and including correlations, size, attributes and candidates for foreign and primary keys; the development of a multi-dimensional corporate model, using diagrams to present facts, dimensions, hierarchies, correlations and candidates for foreign and primary keys.

The following elements should also be included in the planning process: a system to access and protect the data, bearing in mind that a consolidated data warehouse will store information that is both essential and sensitive to the corporation; the scope and sources of external information which will be uploaded into the warehouse. It often turns out that the end user has no influence on the form and structure of this kind of information and so it must undergo preliminary – often manual – processing, which means that competent personnel have to be selected and trained for the job. This element also applies to the requirements for standardization, indispensable for the system to operate and exchange information in virtual organizations. The scope of the changes in recording and accounting systems, from which the data will be extracted, through the attachment of necessary attributes to collected information, or through the exchange of the systems, if the upgrades cannot be made.

According to the authors’ long-term experiences in the field of implementation of IT solutions, an initial analysis is particularly important for the success of any project; however, in the case of a data warehouse, its importance is paramount and has a decisive effect on the final outcome.

3.2 Selection of the right scale and tools for the solution

The analysis discussion, presented in the previous chapter, stressed the importance of the scale of the solution which is particularly valid for the SME sector where companies have different financial capabilities. There are quite large organizations with substantial funds, and they may assign to implement advanced solutions, but there are also small, often family-run businesses, whose finances are much more modest. Nevertheless, all these organizations have information needs that can be satisfied, depending on their requirements and capabilities, using solutions and tools of the appropriate scale.

Practical implementations of BI systems, including data warehouses based on MS SQL Server and MS SQL Analysis Services, have contributed to the development of the following three models, varied in terms of scale and considering different needs and capabilities of the users and relevant software:

- a standard solution, where the user interface is based on standard tools included in office software suites (such as Excel or Access);
- a high standard solution, where the user interface is based on standard tools included in office software suites (Excel, Access) and subscription using file sharing portals;
- a specialised solution, which meets the highest requirements and uses dedicated BI tools.

The above-listed system solutions use the MS SQL Server and MS SQL Analysis Services as the storage engines, which ensure full scalability. Using the same database engine in all kinds of solutions allows the smooth transition from a lower-level solution to a higher one, keeping most of the previously completed elements, which optimizes the costs connected with the development of a BI system. In the first type of solution, the entire scope of the works connected with the design and configuration of the database and its data feed tools is executed, according to the user’s needs.
A complete database structure is created with a system of authorizations to access the data, queries and multi-dimensional OLAP cubes to present the data. The user interface utilizes readily-available tools included in office software suites (Excel, Access) and contains tables, floating reports and mechanisms allowing the import and attachment of data from databases.

Sometimes, specialised tools, offered by ERP system suppliers, are also used to perform economic and business analyses in the office applications through direct access to the databases of the systems. Popular programs such as InfoMaker or Crystal Reports, supporting report writing, can be used there. An advantage of this solution is a low cost of implementation and the use of a working environment with which data warehouse users are familiar. Thus, analytical tools can be developed in the established form as well; the only difference being that the data is fed automatically, not manually. This is particularly important in light of the issue referred to in the previous paragraph, i.e. the users’ habits. When implementing this kind of solution, the collaboration with the users usually runs smoothly, as they do not fear the introduction of the new tool which is very similar to the applications they already know and use. At the same time, users appreciate the speed of operation and the availability of information on demand.

Yet, the solution has certain drawbacks, the most significant of them being the difficulty of administration and the possibility of an uncontrolled leak of data from the organization. Both disadvantages result from the fact that the client applications (office suites) are installed on local computers, practically outside the control of the system administrator. This situation often leads to different results obtained by individual users due to modifications (sometimes erroneous) they make in their spreadsheets. Statements and summaries presented in the form of spreadsheets or reports, stored as regular files in local or even network-based shared user resources can be easily copied onto removable data carriers and used contrary to the interest of the company.

The second solution combines the advantages of the first one while eliminating its drawbacks, and consists of the use of both the standard tools described above and an information portal with a file subscription mechanism. This solution brings additional benefits for more complex organizations with a distributed structure, where the required IT tools are more sophisticated than an MS Office software package. At the same time, it allows for differentiation of the access rights to data and modification of calculation formulas for individual users. When using a portal-based environment, access to all reports and spreadsheets is supervised by the portal whose functions allow the determination of the levels of authorization for users or their groups. As a Business Intelligence class tool, the portal also makes it possible to define groups of reports, summaries and data, which only authorized users can access, down to the level of individual cells. In a complex organization, it is a must that the same strict observance of the principle of working on an official and approved version of the data and reports for all organization units occur. The portal allows the development and the publication of reports in stages, which ensure that only approved and updated documents are available to work on.

An additional facility is the support of communication among individual users of the portal. For example, if an authorized user decides that his or her file contains certain significant information, they can make it available to other authorized users with relevant comments. The portal is equipped with an advanced reporting environment which includes such functions as data aggregation on different levels, using floating tables, or data mining (from general to specific). Most portals also incorporate mechanisms which facilitate data interpretation: charts, indices, semaphores, or data trees. The functionalities of portals include file subscription mechanisms, which provide portal users with a library of the most current, approved and valid models of operating procedures and documents, which goes beyond the traditional scope of use of a data
This kind of portal is widely represented on the market; however, they can also be developed by internal IT teams, using open source tools. Thanks to the use of Internet technology and standard browsers, the portals do not generate significant expenses for an IT infrastructure, while ensuring that the principal requirement of a distributed organization – remote working – is satisfied. A portal-based solution can be implemented starting from the initial stage of data warehouse implementation, or at a subsequent stage when the standard solution (of the first type) has been introduced. In the second situation, it is very important that all previously prepared elements of the database and user interfaces are employed. This kind of solution organizes and secures the elements developed in the first-type solution and provides new possibilities without risking a loss of the previously achieved advantages.

The third solution is intended for organizations with a lot of information needs, where information is the main tool to model the future, and the analytical department is extensive and represented in each significant organizational unit, when process control replaces unit control. These kinds of solutions are IT tools which enhance the efficiency of the information process itself, in particular, including budget planning. This solution utilizes specialized Business Intelligence software with all the functionalities offered by the second solution, ensuring the compliance of the presented data with actual transactional data, the consistency of processed versions of reports and summaries, classified access to various user rights and supporting work in a distributed system, but also providing the users with advanced data visualization tools and efficiently supporting information processes, e.g. budget reconciliation. The software provides excellent support where information is not generated in a single unit or by one user, but where reliable information for management requires the collaboration of a number of organizational units and to generate a separate process. Regardless of the direction of the information flow (upward or downward), the software accelerates the exchange of information and allows one to make arrangements and carry out negotiations among individual units electronically, and the definition of frameworks of new processes and information rules is simple and does not take up much time.

This class of software guarantees a sophisticated and universal reporting system, which can be used to generate formal management reports, often using a similar interface and flexibility to those offered by the popular MS Excel program. The drill-down data mining technique, the data slicing and dicing, floating table views, management cockpits and indices enable the user to develop new, valuable reports, including a balanced scorecard, which is not available in the two previously discussed solutions. Databases in this type of software are multi-dimensional OLAP cubes, compatible with the solutions described above. The systems may also use summaries and reports developed using a standard MS Office package. Therefore, the implementation of this kind of solution can mark another stage in the development of Business Intelligence systems and data warehouses in the organization, based on previously developed elements. The cost of implementation of this solution is much higher than in the first two discussed cases, but if properly implemented in a complex organization, a solution like this will undoubtedly prove beneficial.

The right type of solution should be chosen at the analytical stage and should follow the defined user needs and capabilities (also financial). If, as a result of a thorough analysis, the first or the second solution is chosen, the database components of the final project should be designed and prepared in such a way as to allow the possibility of an easy and trouble-free further development and upgrade to a higher-tier solution in the future, without doing the same groundwork that it is necessary to do at the initial stage. This choice, however, requires the implementing entity to know all classes of possible solutions, regardless of the type to be implemented according to the initial project. Therefore, a future user of a data warehouse or Business Intelligence solutions should
verify the relevant expertise and experience of the supplier already at the analytical stage. The process of selecting the right solution must be preceded by an extensive evaluation; however, there are certain symptoms that indicate the most suitable scale for the circumstances: In the standard solution, where the user interface is based on common tools offered by office software suites (Excel, Access), and these symptoms are the following:

1. In the standard solution, where the user interface is based on common tools offered by office software suites (Excel, Access), there are the following:
   - a small budget;
   - expected limited involvement of senior and middle management in the implementation process and data warehouse operation;
   - lack of IT personnel to administer the solution, or their inadequate level of competence;
   - considerable attachment to currently used information tools among prospective users;
   - a strong reluctance of prospective users to change;
   - the absence of teamwork within the organization;
   - the absence of a distributed structure;
   - the application of reports within one unit/position only;
   - limited requirements for visualization;
   - no pressure to ensure data protection;
   - information plays an auxiliary role in business processes.

2. In the high standard solution, where the user interface is based on common tools offered by office software suites (Excel, Access) and on portal-based file subscription mechanisms, the factors are the following:
   - a small to medium budget;
   - expected small to medium involvement of senior and middle management in the implementation process and data warehouse operation;
   - presence of IT personnel able to administer the solution or planned outsourcing of the service;
   - considerable attachment to currently used information tools among prospective users;
   - no reluctance of prospective users to change;
   - the presence of teamwork within the organization;
   - the presence of a distributed structure;
   - the application of reports in a number of units/positions;
   - average requirements for visualization;
   - focus on data protection;
   - information plays an auxiliary role in business processes.

3. The top standard, specialized solutions, based on dedicated BI tools could be implemented in the following circumstances:
   - a large budget;
   - the expected substantial involvement of senior and middle management in the implementation process and data warehouse operation;
   - the presence of IT personnel capable of administering the solution or planned outsourcing of the service;
   - little attachment to currently used information tools among prospective users, or management supports the effort to overcome any problem that may arise from such an attachment;
- the enthusiasm of prospective users for new solutions;
- the presence of advanced teamwork in complex information and management processes;
- the presence of a distributed structure and user involvement from various locations in a common information process;
- the application and development of reports in a number of units/positions and the presence of multi-stage processes of information development in different units/positions;
- substantial requirements for visualization;
- a focus on data protection;
- the information and its processes are some of the principal elements in the building of a competitive advantage.

Besides choosing the right scale of the intended solution, it is very important to make sure the selected solution is flexible and open. Although more flexible and open solutions are more labour-intensive at the implementation stage, they can be successfully adapted to the requirements of a specific customer and developed even further. This fact has been observed in POZKAL, whose case is discussed in the chapter “Examples of application of Business Intelligence (BI) tools in knowledge acquisition and management”.

3.3 Reasons for the failure of data warehouse implementations

Literary sources identify two groups of the most common reasons for the failure of implementation of data warehouse projects as a management support tool. The first group points to problems in the implementation process, whereas the second one finds fault with the quality of the data. The most common reasons for failure in the first group are the following [4]:
- not to have the support of the management;
- lack of good justification of the project and collaboration with future users;
- poorly specified objective and final product;
- lack of project execution criteria and standards;
- poor planning, monitoring and supervision of the project’s execution process;
- not having made provisions in the project for threats and contingencies;
- lack of a project modification procedure.

The most common reasons for failure in the second group are [3]:
- incomplete data (missing records, fields in source data sets or in the warehouse);
- incorrect records (duplicated records, incorrect calculations, wrong coding, e.g. different codes for different years, incorrect data in source data sets, incorrect data attributes);
- incomprehensible data (incorrect formatting, errors while loading source data, e.g. numerical data loaded as text, incoherent content of records);
- data incompatibility (different coding in different source systems, changeable meaning of the code in time, duplicate data, e.g. one customer’s transactions assigned to a number of objects in a warehouse, inconsistency of warehouse codes with database codes, inconsistency of names and addresses, incompatible dates – a frequent fault due to different formats used in the source systems, incompatibility of business rules, i.e. different methods of value determination, incompatible use of special characters, space, empty fields, incompatible primary and foreign keys in floating base tables, lack of data
synchronization because of supplying the warehouse with correlated data from different source systems at different times).

Business Intelligence specialists, Wiecka and Kuźmierz from the SAS Institute, have determined the reasons behind the above-described situations [3]:

- a lack of knowledge of the company’s own information resources (where the data sources are and what kind of data they provide);
- low quality of the data (bad data management, “dead souls” in data sets, numerous changes in the coding of previously obtained data referring to the same quantities, unclear division of code sets and symbols);
- lack of technical resources to implement a data warehouse (inadequate computing power, lack of required disk space);
- wrong definition of the project priority (missing sponsor, bottlenecks in the design process);
- lack of focus on the essence of the project (too much distraction and attention to minor implementation details);
- variety of project implementation tools (it is much more useful to have one tool to carry out the tasks from the beginning to the end of the project than to integrate tools which are incompatible or hardly compatible by definition).

On the other hand, Gorawski and Konopacki, specialists in Decision Support Systems (DSS), which are based on data warehouses, provide the following reasons [3]:

- unclear or inaccurate determination of the expected outcome of the project;
- failure to meet the customer and future user’s expectations by the implementation team, most frequently due to inadequate understanding on one or both sides;
- lack of a clear-cut objective and undefined measures of evaluation;
- lack of awareness that a DDS project is never completed because business evaluation methods, organizational structures, opinion-forming factors etc. change, which calls for the occasional correction of the logic model, and thus, e.g. of the physical model, requires new data extraction (as a result, the customer is often surprised that a lot of money was spent on a data warehouse which then has to be improved);
- extraction errors – often caused by inherent data defects in source databases;
- incorrect organization of project execution which leads to numerous clashes, misunderstandings, a loss of time and unnecessary discussions;
- mistakes in the selection of programming tools, OLAP environment or the warehouse structure.

Experiences gained while implementing data warehouses in SME sector organizations show that the above-described reasons and causes of failure can be substantially eliminated by performing a thorough preliminary analysis and using its findings to select the right scale solution. The ultimate cause of numerous problems that occur in the implementation of information systems based on data warehouses is a lack of proper analysis or a superficial approach to this important stage, and the ensuing choice of a solution which does not fit the organizational and financial needs and capabilities of the organization. The reasons for this could be found on both sides of the process of selection of the right data warehouse management tool. The supplier of data warehouse systems often offers only one type of solution, which it is endorsed to the future users, taking advantage of their ignorance or limited knowledge of the subject, failing to consider its little relevance in the given circumstances.
A situation like this has been illustrated in the case of the POZKAL printing house (see the section concerning the choice of solution in “Examples of the application of Business Intelligence (BI) tools in knowledge acquisition and management”). However, the user was knowledgeable enough to make the right choice. On the other hand, future users often make up their mind on the basis of incidental premises, such as media news or current fashions, without analyzing or considering their own needs and capabilities. When the choice is based on such premises, positive results are unlikely, even if both sides of the project become fully involved in the implementation process, because, on one hand, the incorrectly selected solution will not provide the tools that are necessary to execute the project successfully or it is too expensive and, on the other hand, most of the budget is spent on the software, making its efficient implementation impossible.

3.4 Directions for the evolution of an organization’s information needs after the successful implementation of a data warehouse

In the contemporary market of interorganisational exchange, the information needs of an organization are constantly on the increase; therefore the development of IT solutions applied in Business Intelligence is a continuous process. After the successful implementation of an information system based on a data warehouse and OLAP, new information needs often emerge within the company. The ways to satisfy the new needs usually follow one of the directions described below:
- The development of the operated OLAP data warehouse through the incorporation of new summaries and reports into the existing solution, or
- The application of new information technologies, most often those such as data mining, workflow, management cockpits and event monitoring, process knowledge, predictive analysis and alerting, which often require the transition to a higher level system (e.g. from information portals to specialized BI software), if the solution implemented before was simpler.

One of the studies which presents the evolution to date and the future direction of the development of Business Analytics solutions (including Business Intelligence) is a report published by IDC in 2007 (Fig. 3.1), which identifies the 5 stages of development of BA systems [5]:
- Stage 1 – reporting from production systems (operating systems) and statistics;
- Stage 2 – queries, OLAP processing and data mining;
- Stage 3 – Business Intelligence suites and analytical applications;
- Stage 4 – decision process automation, DPA;
- Stage 5 – intelligent process automation, IPA.

In practice, when the user, following the successful implementation of a data warehouse, is satisfied with the effects, he or she starts to see further information needs and is willing to invest in the development of this kind of system. Failed implementations, those which involve numerous problems, or those which hardly satisfy the user’s current information needs, naturally make the user wary of such systems and further development is, in these cases, rather unlikely, or it takes a long time before making another attempt at improving the existing structure or implementing a new BI system. The conclusion is that a properly performed initial implementation is the driving force behind further developments. The main reason could be the fact that the user can quickly see the benefits and savings connected with the application of a data warehouse and thus, seeks new opportunities to improve his/her business in terms of information systems.
The experiences gained so far by the author’s team in the implementation of BI solutions indicate that the most common ways to expand the system consisted of the introduction of workgroup and collaboration methods, management cockpits, event monitoring, process knowledge and predictive analyses. A very interesting case was an implementation at a Bydgoszcz-based printing company where the implemented solution developed practically in copybook fashion. After the introduction of a data warehouse with ad-hoc queries and OLAP summary tables using basic tools requirements, it quickly appeared as a more sophisticated method of presenting the results (visualization and management cockpits).

Figure 3.1 Development of Business Analytics systems according to IDC


When this need was satisfied, new needs for event monitoring and knowledge of processes emerged. The company director was interested in short summary reports presenting the most significant (from the point of view of management) information, which were to be sent to his mobile phone and in the real-time monitoring of the efficiency of production (event monitoring). The sales department developed its information needs with regards to processing production orders (process knowledge). It must be noted here that the daily workflow within the company is very dynamic with more than a hundred production orders activated each day on a number of production lines.

In the meantime, an automatic information exchange system was implemented to monitor the progress: on one hand, for order fulfillment, inventory balance and production capacities, and on
the other hand, order planning for key accounts. A new electronic document management system was added (collaboration, workflow and predictive analysis). Finally, the way in which key information was transmitted to the Director was also changed. The old standard reporting method concerning the company’s internal information was replaced with a system in which only unusual events or situations that occur, or are likely to occur, in major processes (predicting and alerting) were notified. This solution was supplemented with relevant alerts addressed to sales personnel, the sales director, the production manager and the financial director of the company. The implementation of these solutions required impeccable collaboration between the user and the implementing entity as well as mutual trust and a good understanding of the user’s business.

Another interesting example can be seen in the implementation of a price calculation system at a different local printing company in Bydgoszcz. The customer’s problem consisted of the pricing of RFQs by sales personnel. Because of the complexity of the printing process, a salesman was unable to define the price of a product on his own, even attempting a rough estimate. Therefore, whenever he obtained an order specification (usually as variants) from a prospective customer, he had to ask the costing department for information, which prolonged the business process and incurred higher costs of sales service. After a thorough analysis of the problem, it turned out that even though practically every enquiry was different; its components were often repeated. The solution was based on a data mining method in which relevant records were searched according to the content (model). The system divides the variable RFQs into individual components and, on the basis of the data stored in the data warehouse, finds the most recent occurrences of the components that match and then assembles them using current variables (such as the cost of materials, labour, currency exchange rates). This method allows an initial calculation for a majority of RFQs.

An additional advantage of the system is the fact that its efficiency improves as the amount of actual calculation data increases in the data warehouse. When a data warehouse is launched, the most frequently implemented element is, undoubtedly, the estimation of sales forecasts using a predictive data mining method. The solution is very popular and often offered as part of an ERP system, or even lower class transaction-processing systems. Despite the fact that the supporting tools have a great popularity and availability, the solution deserves a few words.

In transactional systems, sales planning usually relies on historical sales data, possibly supplemented by set values of a forecasted market growth, macroeconomic tendencies and assumptions of the management, and considers a periodicity. Yet, the problem is much more complex and – in order to develop a reliable estimate – one has to get into the details of the specific circumstances in which the company operates.

For example, in a foodstuffs manufacturing company, where imported raw materials were used in the production; such additional elements had to be considered as disturbances in the historical data caused by promotional actions, stocking up of new stores (hypermarkets/superstores and retail chains), specific patterns of maritime supplies of commodities from far away countries, and the fulfillment of long-term contracts with suppliers. Then again, in a company manufacturing herbal drugs, besides disturbed historical sales data, one had to include contracting herbs which could be only picked at a specific time and places (lack of equivalents or sources of ad-choc replenishment in the event of shortages). In a company dealing with the distribution of colour cosmetics and fragrances, fashion trends had to be considered (a last year’s hit may not sell next year). The examples given above are only some of a bigger number of possible factors which are often so idiosyncratic that a very detailed analysis of historical data is required at the preliminary stage, along with an insight into the user’s specific circumstances. An efficient support of the
planning processes is only possible using a well-designed and properly implemented data warehouse.

**Bibliography**


