# Justifying a Data Warehouse Project

**D. J. Power**

What is the return on investment for a proposed data warehouse project? What is the payback period? What is the opportunity cost? What are the anticipated benefits? What can we do with a data warehouse that we can't do with our current information systems? Do our competitors have a data warehouse? You may be asking these questions about a data warehouse DSS project. Are you receiving satisfactory answers? If you hesitated and said "maybe" that is not surprising because justifying a data warehouse project can be very difficult. Why?

For many years, business professors have been discussing the issues surrounding financial evaluation of capital expenditure projects. We don't all agree. Typical tools recommended are ROI, NPV, and discounted cash flow. These tools are closely tied to the capital budgeting process. They are intended to provide a rational allocation of capital.

Because managers are asked to spend funds on a data warehouse project, anticipated results and benefits should be quantified so that the requested expenditure can be evaluated in comparable units. But in a data warehouse project it is difficult to quantify the results and benefits. We are basically making gross estimates and guesses. A financial analysis is especially difficult because the costs are uncertain and many of the benefits are qualitative and intangible.

A number of alternative tools are available for evaluating data warehouse projects. Value analysis is an evaluation of "soft" benefits such as improving staff productivity, improving the speed of strategic actions, enhancing a company's competitive advantage, or improving access to data. A related approach, the business value added approach "measures IT contribution not in dollars, but by its support of key goals and metrics of functional groups (check Maglitta, 1997; Keen, 1982)". A third alternative, the scenario approach attempts to project what decision making will be like when the data warehouse is in place and hence speculate on how the company will benefit. All of these qualitative approaches have pluses and minuses, but each can be improved by understanding the upside and downside of a data warehouse project. Let's briefly review what managers are reporting in companies that have implemented data warehouses.

**Upside potential**

The primary benefit of data warehouses should be improved decisions. This intangible benefit presumes that managers will change their decision processes and actually use a data warehouse. In a recent Sentry Market survey, 30% of respondents identified "access to data" as the biggest benefit of a data warehouse. Other important benefits included: improved data accuracy; gaining competitive advantage; better control of data; better data consistency; decentralization of data; cost savings; and less reliance on legacy systems. Few respondents thought that there would be cost savings.

According to Theresa Rigney, Sentry Market Research's 1996 Software Market Survey "shows that 55% of the respondents polled declared that data warehousing is vital to their organizations' business objectives." The following case examples identify reducing the number of

programmers, reducing staff hours spent preparing report, and improved sales staff productivity as tangible benefits of a specific data warehouse implementation. "Your results may vary!"

All of the data warehouse vendors and the Data Warehousing Institute (DWI) have glowing reports of data warehouse successes at their web sites. For example, to justify their data warehouse investment, UPS managers "looked at both tactical and strategic benefits. The biggest tactical benefit was a sharp decrease in the cost of programmers. Roughly half of the 130 programmers on staff were not required after the data warehouse was up and running. Users developed their own reports instead of ordering new features from the IT department. That savings alone justified the investment in the data warehouse."

In 1992-93, Ertl Co., Inc., Dyersville, Iowa, a manufacturer of replica toys, built a data warehouse that eliminated 18,000 hours a year in report creation work. This amounted to several hundreds of thousands of dollars a year in savings (see the DWI case study).

A Hewlett-Packard case study quotes Greg Friedrichs, grocery H-E-B grocery company's manager of decision support systems, "There's no question that our new decision support system represents a significant investment, and a leap of faith. We can't point to specific sales results that we've achieved by having this system. But on the other hand, none of our users can imagine how they ever did their work without it."

In a mini-case study of MCI titled *SOLD Delivers Quantifiable Results* , Flanagan and Safdie quote an MCI manager. "The best measure of the quality of our calling lists, and of the service we provide when we deliver them, is salesman productivity--sales per individual," according to Hacker. "Since we've modernized our system, that figure has grown 200 to 300 percent. What's more, this has all happened when our sales force grew from 300 to 2,000 strong. Just keeping up with growth would have been impossible before this system, let alone improving the quality of our deliverable". So we need to ask if improved staff productivity can be expected.

These qualitative and survey results are positive, but we also need to examine the major quantitative analysis that is being discussed by vendors and consultants conducted by IDC. In late 1995 and early 1996, Stephen Graham, vice president of software research at International Data Corporation (IDC) Canada, evaluated 62 organizations that had had a data warehouse in operation for at least 6 months. The report from the study is titled "The Foundations of Wisdom: A Study of the Financial Impact of Data Warehousing". The title phrase "foundations of wisdom" causes me some concerns, but let's review some of the major finding.

The IDC study showed an average 401% three-year Return on Investment for 52 companies that were not deemed outliers. E. B. Baatz notes "The astonishing numbers are well-documented and compelling enough to make any CEO embark into the data warehousing venue without executing an exhaustive, possibly time-consuming ROI analysis first." In "Case-in-Point" at acxiom.com, the IDC results are also reviewed and summarized. Supposedly, the data warehouse projects had a "mean payback period of 2.31 years. Over 90 percent reported a three-year return on investment exceeding 40 percent. Half reported returns of greater than 160 percent, and one quarter over 600 percent. European companies averaged 340 percent ROI, compared with 440 percent among North American firms.

At the SAS web site the white paper titled "Data Warehousing ROI Tops 400%" (check http://www.sas.com/new/dwsurvey/intro.html) summarizes the IDC findings. The SAS paper notes "Its goal was to assess the financial impact of data warehousing; its results are nothing short of astounding."

In "Digesting the ROI Paradox," E. B. Baatz argues "the decision to invest in data warehouses seems a lot like the ATM example, with CIOs scrambling to define the ROI and CEOs leaping without the data... The ultimate ROI, if anyone were so enslaved by numbers enough to take the time and trouble to measure it, is astronomically high."

**Downside potential**

I have read the parts of the IDC study that are available free from a number of the study sponsors. Also, I read the white paper summary. The report can be ordered from IDC Canada by calling Jacquie McLean at 416 369-0033. The methodology used in the study raises concerns for academic researchers. First, the companies Stephen Graham and IDC studied were self-selected by vendors and the IT professional services firms who sponsored the study. Second, the participating firms had to be willing to be identified. Both of these research design elements bias the study so that it is much more likely to find results in favor of building a data warehouse. Despite this bias, the study seems thoughtful and the authors report some downside potential of data warehouse projects. Also, it should be remembered that the authors examined ROI for completed projects. A retrospective calculation of ROI or payback is much easier and more reliable than an a priori estimate that might be included in a data warehouse feasibility analysis.

Theresa Rigney cautions "be wary of studies like the IDC report. Sure, the 62 completed data warehousing projects examined by IDC reported a great ROI, but most projects out there aren't yet completed. In fact, the costs and complexity of an enterprisewide data warehousing project have capsized a lot of pilot efforts. IT must be aware of these costs -- in time and effort as well as hardware and software -- before it undertakes a full-scale data warehousing project." She argues there "are no ways of predicting or quantifying ROI. The real payback comes from getting information assets out to the users in the best way possible."

In "Case-in-Point" at acxiom.com, the IDC results show eight companies reported negative ROI results, which they attributed to either extraordinarily high costs, low warehouse usage, or a large undertaking requiring longer than three years for payback." Also, IDC "found that discrete implementations -- those developed to support one business need -- realized higher average ROI (532 percent in 1.57 years vs. 321 percent in 2.73 years) than broader, enterprise initiatives. However, the authors were quick to point out that discrete, demand-driven warehouses are in some respects easier to develop."

On a related downside note, Flanagan and Safdie state "An organization's return is measured in productivity and market gains which accrue from the use of the warehouse. Therefore, organizations whose warehouses are unused have the same denominator (costs) as those whose products are well used, but have a much smaller, or nonexistent numerator (benefits)."

On the human upside, Winkler notes that a "data warehouse enables analysts to do for themselves what was previously done by intermediaries." But she goes on to state "The downside of decision-empowerment falls on the shoulders of middle-managers. Their role of data filter is obsolete (check http://www.sentrytech.com/smrwps21.htm)". This possible downside issue is too complex for a brief analysis. Ask yourself if eliminating middle-managers as a data filter and making that role obsolete is a plus or minus.

Finally, some data warehouse projects are failures. According to Flanagan and Safdie, Earl Hadden and Associates found that "85% of current data warehouse projects fail to meet their intended objectives, and that 40% don't even get off the ground". John Ladley, Meta Group Program Director for Application Delivery Strategies, says "We are past the stage where every data warehouse is a success story ... believe 40% of the warehousing projects aren't meeting expectations, and 10% are abject failures."

**Conclusions**

My first conclusion is that a detailed qualitative analysis of a proposed data warehouse is the most that managers can reasonably expect. Although in some situations, financial analysis tools can be useful decision aids, in a data warehouse project their use provides only the appearance of accuracy and precision. When making a data warehousing project decision, managers should generally ask "What are the expected results and benefits?" rather than "What is the anticipated return on investment (ROI)?". Managers should not demand a positive ROI from a data warehouse project, but they must demand positive results.

When I began writing this commentary I considered arguing that today, in Fall 1997, a data warehouse is a basic business requirement. And therefore executives in most companies, large and small, have no choice but to fund data warehouse projects. I think it is true that data warehouses are a basic business requirement, but I'm not willing to let the IS/IT staff "off the hook" that easily. I'm concerned that some IS/IT managers who want to build data warehouses have insuffient knowledge about them. Please be wary of IS/IT managers who have difficulty explaining in detail the benefits of a data warehouse project. So my more conservative conclusion is that investigating a data warehouse project is a basic business requirement.

A word of warning... Learning enough to understand and explain the benefits of a data warehouse may be expensive. IS/IT managers will need to do more than read a book. IS/IT staff should actually work with a prototype data warehouse. The IT unit may want to hire a consultant; staff should attend seminars and talk to vendors. The process of learning about data warehousing will be time consuming and costly. Companies may need to spend a few hundred thousand dollars on a prototype or a departmental data mart and that is a significant investment. In firms with multimillion dollar IS/IT budgets a prototype data warehouse or data mart project is needed and it should be viewed as "a learning experience". General managers need to spend enough money on a data warehouse project so that IS/IT and business managers can learn about warehouses and can evaluate the benefits. Building a data mart or data warehouse is an employee and corporate development experience.

Your benefits from a data warehouse project will vary depending upon where your company is on the IS/IT learning curve. If your company still relies on a legacy system developed in Cobol, the benefits could be very large, but the implementation risk is high. So start small; focus on anticipated results **NOT** ROI; but do start a data warehouse project.

**References**


1. Applegate, L. M., F. W. McFarlan, and J. L. McKenney, **Corporate Information Systems Management: Text and Cases**, Irwin, 1996, Ch. 11 "A Portfolio Approach to Information Technology Development" .
2. Baatz, E.B. "Digesting the ROI Paradox," *CIO Online*, Oct. 1, 1996. URL http://www.cio.com.
3. Datamation White paper, "The Evolution of Data Warehousing", Datamation, 1997, URL http://www.datamation.com/PlugIn/whitepapers/data_warehouse2/intro.html
4. Data Warehouse Institute (DWI) Case Study: Ertl Company, URL http://www.dw-institute.com/cases/ertl.htm.
5. Flanagan, T. and E. Safdie, "Data Warehouse Technical Guide", Sybase, URL http://www.sybase.com/products/dataware/techguide.html.
6. Hewlett-Packard, **H-E-B** grocery company case study, URL http://hpcc920.external.hp.com/retail/sucheb5.html.
7. International Data Corporation, "The Foundations of Wisdom: A Study of the Financial Impact of Data Warehousing", 1996, 36 Toronto Street, Toronto, Ontario, Canada M5C 2C5, phone 416-369-0033.
8. "IDC Tracks Warehouse ROI. Does data warehousing pay?", *Case-in-Point* , Vol 3 No 3, May/June 1997, http://www.acxiom.com/cip-v3n3-e.htm.
9. Keen, J. "Turn 'soft' benefits into hard savings", *Datamation*, September 1997, URL http://www.datamation.com/PlugIn/issues/1997/september/09roi.html
10. Keen, P. G. W. "'Value Analysis: Justifying Decision Support Systems"," *MIS Quarterly*. March, 1981. pp. 1-15.
11. Maglitta, J.E. "Managing Beyond ROI," *Computerworld Online*, Oct. 27, 1997, URL http://www2.computerworld.com/home/online9697.nsf/All/971027man_roi1.
12. Rigney, T., "Advice to MIS: Think Big, Start Small", Sentry Market Research, Special Editorial Supplement to *Client/Server Computing* and *Software Magazine*, URL http://www.sentrytech.com/smrwps30.htm.
13. SAS, "Data Warehousing ROI Tops 400%", http://www.sas.com/new/dwsurvey/intro.html
14. Sentry Market Research, Special Editorial Supplement *to Client/Server Computing and Software Magazine* check URL http://www.sentrytech.com/smrwp4.htm.
15. Winkler, C., HUMAN FACTORS: Surviving the Earthquake , SMR *Data Warehousing Directions*, URL http://www.sentrytech.com/smrwps21.htm.



Please cite this article as:

Power, D. J. "Justifying a Data Warehouse Project", Part I and II. DSstar, The On-Line Executive Journal for Data-Intensive Decision Support, February 3 and 10, 1998: Vol. 2, No. 5 and 6.