

Benchmarks for BI and Data Warehousing Success

By Hugh J. Watson and Thilini Ariyachandra

People and companies always want to know how they are doing. That's why they keep score, whether it is for golf or for business performance. The scores show how people and companies are doing over time, against goals or in comparison to others. Scores provide feedback and incentive to improve performance.

In business, benchmarks are especially useful. They are helpful in answering the question: How are we doing, especially in comparison to other companies? This is true in the business intelligence (BI) and data warehousing world. People want to know how successful their BI and data warehousing initiatives are in comparison to other companies.

We conducted a study (promoted by *DM Review*) that created metrics used for assessing the success of a data warehouse architecture and a company's use of BI. We collected survey data from 454 companies that can be used for benchmarking purposes.

In this article, we discuss how the success measures were selected, what the success metrics are, the survey data that was selected, the benchmark data for the metrics and the use of the metrics.

BI and Data Warehousing Success Metrics

Both practitioners and academics have a long-standing interest in measuring success. BI practitioners need to know how successful they are and communicate this assessment to management. Academics need to be able to

measure success in their research.

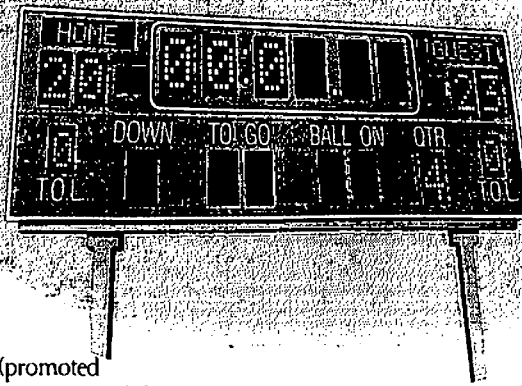
To develop the success metrics, we reviewed practitioner and academic literature, and interviewed 20 leading authorities on BI and data warehousing. Two major success constructs emerged as relevant: product measures and development measures. Each of these constructs has component parts.

Product Measures

- Information quality: The data warehouse should provide accurate, complete and consistent information.
 - System quality: The data warehouse should be flexible, scalable and able to integrate data.
 - Individual impacts: Users should be able to quickly and easily access data; think about, ask questions, and explore issues in new ways; and improve their decision-making because of the data warehouse and BI.
 - Organizational impacts: The data warehouse and BI should meet the business requirements; facilitate the use of BI; support the accomplishment of strategic business objectives; enable improvements in business processes; lead to high, quantifiable ROI; and improve communication and cooperation across organizational units.

Development Measures

- Development cost: The cost of developing and maintaining the data warehouse should be appropriate.
 - Development time: The time to develop the initial version of the data warehouse should be appropriate.
- These are not the only possible measures (e.g., system usage), but these were the ones most germane to our research. Some of these measures relate specifically to the data



warehouse, such as information quality and the warehouse development cost. Others, such as individual and organizational impacts, measure the success of both the data warehouse and the data access tools (e.g., Excel, Business Objects) and BI applications (e.g., EIS) that use the warehouse. Collectively, these measures provide a useful set of success metrics for the data warehouse and BI.

Data Collection and Analysis

Once again using the literature and experts, questions were developed for the success measures. In the survey instrument, respondents were asked to indicate (on a seven-point scale, with one being strongly disagree and seven strongly agree) the success of their BI and data warehousing initiatives. The job positions of the respondents were relatively evenly distributed over data warehouse managers, data warehouse staff members, IS managers and independent consultants/system integrators. The latter were asked to complete the survey with one of their clients in mind. The companies included in the survey ranged from small (i.e., less than \$10 million in revenues) to large (i.e., in excess of \$10 billion). Most of the companies are located in the United States and represent a variety of industries, with financial services providing the most responses. The respondents are believed to be representative of the companies that use BI and data warehousing.

In addition to questions about the success metrics, respondents were asked to indicate their company's data warehouse architecture. Five choices were given:

- 1) Independent data marts.
- 2) Bus architecture with conformed dimensions (bus architecture).
- 3) Hub and spoke (i.e., Corporate Information Factory).
- 4) Centralized (i.e., no dependent data marts), and
- 5) Federated.

The independent data marts architecture scored lowest on the success metrics, followed by the federated architecture. Interestingly, the bus, hub-and-spoke and centralized architectures scored similarly on the success metrics for information

Success Metrics	Benchmark Scores
Information Quality (accuracy, completeness, consistency)	
1. Your data warehouse supports queries and reports with few data errors.	5.16
2. Your data warehouse supports the level of data correctness needed for its intended purpose.	5.46
3. The data values in the data warehouse correctly represent the real-world objects and events being described.	5.63
4. Your data warehouse includes data about all the business processes and subject areas that are required by users and applications.	4.45
5. All necessary decision support data is available within the data warehouse.	4.21
6. Your data warehouse provides all the data needed by users and applications.	4.19
7. Your data warehouse provides a single version of the truth.	5.35
8. Your data warehouse reduces data inconsistencies.	5.44
9. Your data warehouse architecture provides a single system of record for decision-support data.	5.44
System Quality (flexibility, scalability, integration)	
10. Your data warehouse architecture makes it is easy to add new business processes and subject areas.	5.38
11. Your data warehouse provides the capability to satisfy new requirements quickly.	5.19
12. Your data warehouse provides the capability to easily support future application needs.	5.28
13. Your data warehouse architecture is scalable to handle increases in the number of users without negatively impacting system performance.	5.40
14. Your data warehouse is scalable to handle increases in the complexity and number of simultaneous queries without degrading system performance.	5.04
15. Your data warehouse architecture is easily scalable to handle increases in the volume of data.	5.43
16. Your data warehouse supports and facilitates the integration of data from multiple systems.	5.90
17. Your data warehouse supports and facilitates the integration of internal and external data sources.	5.43
18. Your data warehouse supports and facilitates the integration of all needed data around primary keys.	5.43
Individual Impacts	
19. End users are using the data warehouse that was implemented.	5.64
20. Users can access the data more easily and quickly because of the data warehouse.	5.91
21. The data in the data warehouse is easy and intuitive for users to understand and use.	5.03
22. Your data warehouse enables users to think about, ask questions and explore issues in ways that were previously not possible.	5.58
23. Your data warehouse has improved decision-making capabilities of end users.	5.76
Organizational Impacts	
24. Your data warehouse has met the business requirements for which it was implemented.	5.50
25. Your data warehouse has greatly facilitated the use of business intelligence.	5.44
26. Your data warehouse has enabled improvements in business processes.	5.31
27. Your data warehouse has supported the achievement of strategic business objectives.	5.39
28. Your data warehouse has led to high and measurable ROI.	4.53
29. Your data warehouse has improved communication and cooperation across organizational units.	4.97

Figure 1: Benchmark Scores for Information Quality, System Quality, Individual Impacts and Organizational Impacts