

Measuring the Business Value of Data Quality

Published: 10 October 2011

Analyst(s): Ted Friedman, Michael Smith

Research shows that 40% of the anticipated value of all business initiatives is never achieved. Poor data quality in both the planning and execution phases of these initiatives is a primary cause. Poor data quality also effects operational efficiency, risk mitigation and agility by compromising the decisions made in each of these areas.

Key Findings

- Poor data quality is a primary reason for 40% of all business initiatives failing to achieve their targeted benefits.
- Data quality effects overall labor productivity by as much as a 20%.
- As more business processes become automated, data quality becomes the rate limiting factor for overall process quality.

Recommendations

Business leaders and IT leaders focused on data quality improvement and information governance should:

- Measure the business value of improved data quality by focusing on business processes, investment decisions and overall productivity.
- Qualify the business value of improved data quality using business metrics that are correlated with financial outcomes.
- Share the findings of this research with your finance department to receive feedback and guidance on where to begin identifying opportunities for increased business value from improved data quality.

Table of Contents

Analysis.....	2
---------------	---

The Impact of Data Quality on Business Processes.....3
 The Impact of Data Quality on Productivity.....3
 The Impact of Data Quality on Decision Making.....5
 Process for Measuring the Specific Business Value of Data Quality.....5
 Example of Measuring Business Value of Data Quality.....8
 Bottom Line..... 10
 Recommended Reading..... 10

List of Figures

Figure 1. The Productivity Benefits From IT.....4
 Figure 2. The Gartner Business Value Model.....7

Analysis

Organizations are discovering that data quality deficiencies have a significant impact on their most strategic business initiatives, often holding them back from achieving the growth, agility and competitiveness that they desire. In addition to challenges with growth and agility, compliance and transparency pressures increasingly bring data quality issues to the fore — it is no longer acceptable to ignore flaws in data, and organizations must prove the accuracy of information that they report to auditors, regulators and the public. Given the current economic challenges and budgetary pressures facing most organizations, there is a substantial desire to eradicate quality issues in data as a way to reduce costs and improve efficiency (see "Findings From Primary Research Study: Organizations Perceive Significant Cost Impact From Data Quality Issues").

Historically, a minority of organizations have taken a proactive approach to managing data quality — the majority have endured the pains of poor-quality data and dealt with it in a reactive manner. However, with recognition of the impact of poor-quality data increasing as a result of all of these contemporary business drivers, many are undertaking improvement efforts within the context of individual projects or occasionally on a broader, enterprisewide basis.

They can address these challenges through the discipline of data quality, which includes the approaches, organizational models, techniques and technologies used to assess and ensure the "fitness for purpose" of all types of data for use by the various applications and business processes across the enterprise (see "Key Issues for Data Management and Integration Initiatives, 2010"). While technology plays a key role in data quality improvement, changes in work processes and behavior of people are critical.

The demand for data quality best practices is maturing — more organizations are focused on fine-tuning their efforts by increasing the precision with which they measure and monitor data quality, the breadth of their scope (looking across multiple domains, processes and groups) and creating data quality-specific roles and organizational structures. Tools for measuring, monitoring, and

improving data quality can help enterprises improve effectiveness and deploy controls in approaching data quality challenges.

How can CIOs and CFOs quantify the impact on their organizations? Each organization is different, but many different examples can be seen across a variety of industries (see Note 1). CIOs can use similar examples that are relevant to their organization to quantify the financial impact of improving data quality and gain buy-in from business leaders in sales, marketing, customer service and other parts of the organization. One of the new realities of the global economic environment is the desire of business executives to manage risk more effectively. This has created the need to shift from "gut-feel" to fact-based decision making (see "Survey Analysis: How Executives Use Business Metrics"). Gartner conducted a survey of business executives before and after the financial crisis of 2008. The results showed a 30% increase in the use of business metrics as the basis for setting strategic direction. So what is the measurable impact of data quality?

The Impact of Data Quality on Business Processes

The effects of data quality on business processes can be estimated based on Six Sigma. Harry and Schroeder estimate that the average successful business process maintains 3.5 sigma value (see Note 2). This equates to 22,800 defects per million. At this rate, the average cost of quality is 20% of the overall business process costs. For example, if the business process were sales and a company spent \$200 million to maintain its sales process throughout the year, then at 3.5 sigma the cost of quality would be \$40 million annually. In the past 10 years data quality has begun to set the basis for process quality.

Business processes are becoming digital. Over the past 10 years, information technology has been used to automate transaction that go well beyond accounting. Automation of sales, product development, manufacturing, and recruitment processes, to name a few, is expanding each year to lower costs. When processes are automated to the maximum extent possible, data quality issues become the limiting factor in maximizing process quality.

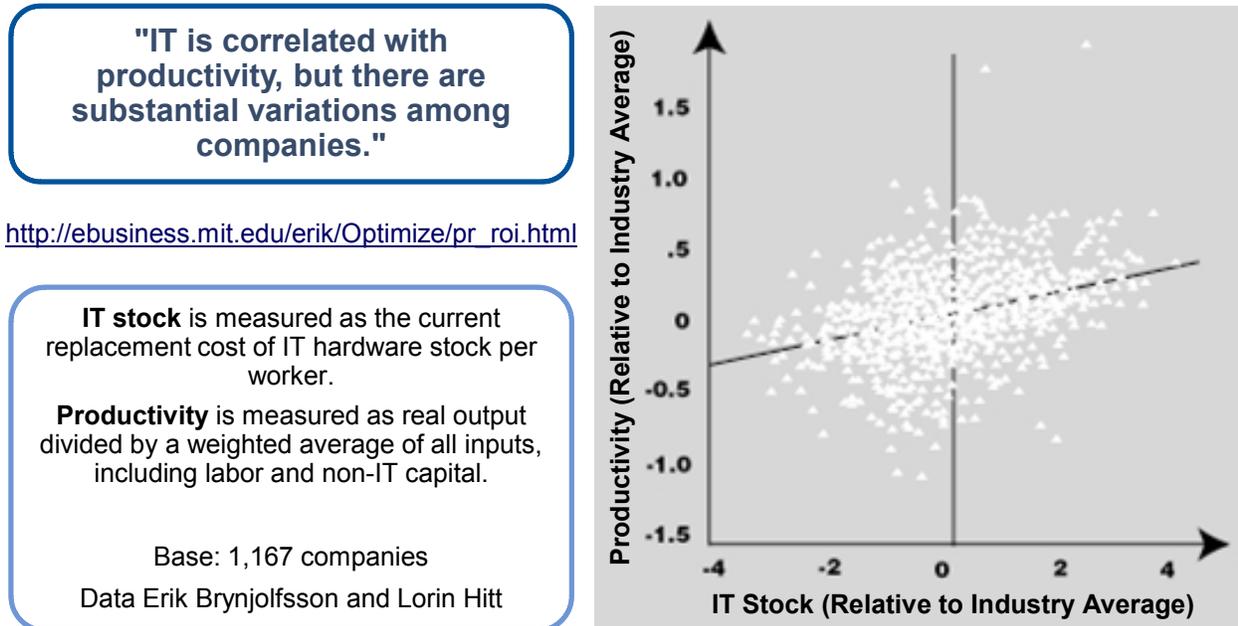
Consider the following example. A \$10 billion consumer goods distribution company recently implemented a sales force automation system. All leads, qualified leads, sales resources, and campaign strategies are recorded in the system. The quality of prospect and product data are critical to the overall quality of the sales process. Data quality defects become process quality defects as mistakes are made when data is wrong. It does not take much. To maintain average process quality of 3.5 sigma, data quality must remain below 22,800 defects per million, or 97.72%, in highly automated processes such as automated material handling systems. The 20% cost of quality results from flaws in data. In effect, data quality becomes the limiting factor in process quality.

The Impact of Data Quality on Productivity

The top half of Figure 1 shows the impact of information technology on overall labor productivity in the U.S. for the past 10 years. The Bureau of Labor Statistics estimates that 70% of the annual productivity growth of 2.7%, or 1.9%, comes from IT. This is the good news; the bad news is

shown on the bottom half of Figure 1. The productivity benefits are not shared equally; in fact, there exists substantial variation. Why is it that one enterprise can benefit while others do not?

Figure 1. The Productivity Benefits From IT



Source: Gartner (October 2011)

One reason is poor data quality. Through client interactions and studies, Gartner continues to observe a high correlation between significant data quality issues and degraded productivity of key resources. A survey of business executives in your company can be used to estimate the impact of poor data quality on labor productivity.

The average 1.9% increase in labor productivity from IT saves approximately \$21.7 million for the average enterprise. If poor data quality affects 10% of this savings, the result is an annual \$2.2 million loss on that saving.

Average Enterprise:

Annual Revenue: \$2,857 billion

Labor Costs: \$1,143 billion

Average Annual IT Productivity: 1.9%

Benefit From IT Productivity: \$21.7 million

10% of IT Productivity From Data Quality: \$2.12 million

The Impact of Data Quality on Decision Making

Our research, and that of the Standish Group, shows that 40% of projects succeed (achieve targeted benefits), 40% are challenged (achieve half the benefit) and 20% fail completely (no benefit). This results in an average 60% benefit realization.

Benefit realization rate = $(40\% \times 100\%) + (40\% \times 50\%) + (20\% \times 0\%) = 60\%$.

This means that 40% of the benefits targeted in the annual capital budget are never realized. There are many reasons for projects not achieving targeted benefits, but consider the impact of data quality on the assumptions and projections made. One way to assess this in your company is to take the list of approved capital projects for the current year. Ask the members of the approving authority (capital expenditure committee) how accurate they believe the data was that was used to make the assumptions and projections in the business case for each project. Also ask them for their estimate of the impact that the level of data quality will have on benefit realization. A 10% impact of data quality on benefit realization would mean \$2.1 million per year for the average enterprise.

Average Enterprise:

Annual Revenue: \$5 billion

IT Spend: \$175 million (3.5% of revenue)

IT Capex: \$52.5 million (30% of total IT spend)

Benefit Realization: 60% (weighted average for benefit realization, see above)

Lost Benefits: $40\% \times \$46.2 \text{ million} = \21 million

10% Impact From Data Quality: \$2.1 million

Process for Measuring the Specific Business Value of Data Quality

Measuring the impact of data quality on business processes and capital investments requires that we use both accounting and non-accounting metrics. Accounting metrics are required by various regulators around the world for reporting the financial position of an enterprise. This is true of both public and private sector organizations. They are quite useful for comparison purposes because the definitions are clear and unambiguous.

Much of the value of information technology, like the value of data quality, is consumed in building and growing branding, intellectual property, unique business processes and knowledge. Yet this value cannot be measured with accounting metrics. Recognizing this measurement gap, Gartner spent several years developing a set of extensions to accounting metrics called the Gartner Business Value Model (see "The Gartner Business Value Model: A Framework for Measuring Business Performance"). The metrics in the model measure activities and events that precede accounting results and are correlated with actual accounting results. Each metric in the model is mathematically linked with an accounting metric in either the balance sheet or the income statement. This is how "buy-in" with senior business executives, particularly financial executives, is

achieved. By showing these executives how the non-accounting metrics are correlated to and have a causal relationship with accounting results, the extension becomes part of the lexicon for business value.

Using extensions like those defined in the Gartner Business Value Model (GBVM), we can more precisely measure the value of data quality. Figure 2 is a graphical representation of the model.

Figure 2. The Gartner Business Value Model

Business Aspect	Desired Business Outcome	Key Risk Indicators			
Demand Management	Market Responsiveness	Target Market Index	Market Coverage Index	Market Share Index	Opportunity/Threat Index
		Product Portfolio Index	Channel Profitability Index	Configureability Index	
	Sales Effectiveness	Sales Opportunity Index	Sales Cycle Index	Sales Close Index	Sales Price Index
		Cost of Sales Index	Forecast Accuracy	Customer Retention Index	
	Product Development Effectiveness	New Products Index	Feature Function Index	Time-to-Market Index	R&D Success Index
Supply Management	Customer Responsiveness	On-Time Delivery	Order Fill Rate	Material Quality	Service Accuracy
		Service Performance	Customer Care Performance	Agreement Effectiveness	Transformation Ratio
	Supplier Effectiveness	Supplier On-Time Delivery	Supplier Order Fill Rate	Supplier Material Quality	Supplier Service Accuracy
		Supplier Service Performance	Supplier Care Performance	Supplier Agreement Effectiveness	Supplier Transformation Ratio
	Operational Efficiency	Cash-to-Cash Cycle Time	Conversion Cost	Asset Utilization	Sigma Value
Support Services	Human Resources Responsiveness	Recruitment Effectiveness Index	Benefits Administration Index	Skills Inventory Index	Employee Training Index
		HR Advisory Index	HR Total Cost Index		
	Information Technology Responsiveness	Systems Performance	IT Support Performance	Partnership Ratio	Service Level Effectiveness
		New Projects Index	IT Total Cost Index		
	Finance & Regulatory Responsiveness	Compliance Index	Accuracy Index	Advisory Index	Cost of Service Index

Source: Gartner (October 2011)

The following 10 steps summarize the process for measuring the business value of data quality.

1. Choose an area within your enterprise where data quality is an issue. Using the GBVM, sales might be an example.
2. Engage the management team in that area to explore the potential benefits that improvements in data quality may have. Include a representative from finance in the discussion, someone familiar with the business area being explored. Explain this process to these stakeholders and ask for their involvement as outlined below.
3. Select between two and six metrics from that sales area that are leading indicators of revenue. The GBVM is ideal for this purpose.
4. Baseline the current performance levels, using the definitions, for each metric selected.
5. Discuss the current state of data quality in the area being explored, and the expectations for data quality held by business stakeholders. Current state should be assessed across a range of data quality dimensions, such as validity, completeness, consistency and accuracy. Also discuss how this level of data quality affects decision making in the area being explored. Estimate the impact of various levels of improvements to data quality on the metrics selected. These are estimates and assumptions will need to be documented.
6. Determine the value of improvements in data quality by converting improvements in performance of the metrics selected using the financial sensitivity calculations. The financial sensitivity calculations were developed as part of the research for the GBVM and must be validated with the financial representative on your team. This has been done hundreds of times with financial representatives at Gartner clients. A common result is for the representative to tweak the calculation based on specific characteristics found in the enterprise, but acceptance of the calculation is achieved.
7. Determine the cost for achieving the targeted improvements in data quality with the stakeholder. Information technology professionals usually take the lead on this step.
8. Build the business case for improvements in data quality using the benefits determined in step 6 and the costs determined in step 7. Be sure to include all assumptions.
9. Present the business case with the value quantified to the approving authority. Be sure to have all stakeholders with you.
10. Implement the solution, including a change management focus to ensure data quality controls, metrics, and the roles participating in data quality improvement evolve over time as business requirements evolve.

Example of Measuring Business Value of Data Quality

The following is a simple example of applying the above process to data quality concerns in a sales process. While not relevant for all organizations in all industries, this example is meant to demonstrate the basic steps and principles of applying the GBVM concepts to identifying and addressing data quality issues in a typical enterprise. The figures cited in the example are fictitious, but are representative for efforts of this type.

1. For this example, we will focus on the sales area. Sales growth is an important objective, and many organizations will want to link their data quality focus to this objective.
2. The key stakeholders that must participate in this effort will be IT leaders focused on data quality, executives responsible for the sales organization and process, and a representative from finance.
3. After general discussion, the stakeholders agree that sales opportunity index (SOI), forecast accuracy and customer retention are the key leading indicators of sales growth. These will be the metrics that will be used to measure the business value of data quality.
4. The baseline for current performance, using the definitions of these sales-related metrics in the GBVM was:
 - SOI = 0.1
 - Forecast accuracy = 0.40
 - Customer retention = 0.70
5. The team then identified specific ways in which data quality issues negatively impacted these key metrics. Their main findings showed:
 - SOI was weak due to poor quality of prospect data (demographics, credit history, past purchase history, and so on) which caused fewer leads to be qualified than expected.
 - Forecast accuracy was low due to poor quality of inventory data (for example, duplicate stock-keeping unit numbers) and incomplete/missing sales history data.
 - Customer retention was low because a lack of quality in customer returns data inhibited the ability to identify dissatisfied customers.
6. In order to determine the business impact of these data quality issues, the team then used specific data quality dimensions to quantify the magnitude of data quality flaws affecting each metric.
 - Related to SOI, the quality of prospect data was measured to be 50% complete (based on the definition of completeness requiring valid values for 10 specific attributes about a prospect) and 30% accurate (in many cases, the revenue and employee size figures for a prospect were measured to be inaccurate).
 - Related to forecast accuracy, the product master was found to have 10% duplicate records. Combined with the impact of some missing sales history data, this contributed to an overall quality level of 70% for key data used in forecasting.
 - Related to customer retention, only 60% of returns had accurate customer details, prohibiting the other 40% from being linked to customer accounts.
 - Using the financial sensitivity calculations in the GBVM, the finance representative determined that these data quality flaws created a loss of \$1 million of sales growth opportunity.

7. The team estimated that the cost for resolving these data quality issues would be \$300,000. This was based on the expected time and effort to implement required changes to existing systems, as well as for training of pertinent individuals involved in the related business processes.
8. Using the costs and benefits identified, the finance representative determined that there was clearly a solid business case for implementing these changes.
9. The team then presented the business case to the approving authorities and received funding and support to execute the effort.
10. Led by IT, modifications to various sales and inventory systems were made to introduce improved data quality controls. This eliminated the possibilities of the identified data quality issues from occurring in the future. In addition, sales and inventory personnel were trained in data quality best practices and given specific work tasks to continually monitor and validate data quality levels. Ongoing, the team will measure both the specific data quality metrics and the key sales growth metrics to ensure the changes are having a positive impact.

Bottom Line

Data quality issues are a perfect example of the business cliché, "You can't manage what you don't measure." Because most organizations do not make the effort to measure the quality of their data in any objective or quantitative way, they fail at building a business case for formal data quality improvement efforts, expend energy on problems that are perceived rather than real, and never have a clear understanding of whether their efforts are making a difference. Even worse, organizations make (often erroneous) assumptions about the state of their data and continue to experience inefficiencies, excessive costs, compliance risks and customer satisfaction issues as a result. In effect, data quality in their business goes unmanaged.

A metrics-based approach to assessing data quality helps remove the assumptions, politics and emotion often associated with this issue, thereby giving organizations a factual basis on which to justify, focus and monitor their efforts. In addition, the identification, communication and analysis of metrics on an ongoing basis provides a tangible indication to the organization that the data quality issue is important to the business. With this solid foundation of facts and focus, rather than perceptions and apathy, organizations can truly begin to manage the quality of their data and will reap significant benefits as a result.

Recommended Reading

Some documents may not be available as part of your current Gartner subscription.

"The Gartner Business Value Model: A Framework for Measuring Business Performance"

"Findings From Primary Research Study: Data Quality Issues Create Significant Cost, Yet Often Go Unmeasured"

"Strategic Focus on Data Quality Yields Big Benefits for BT"

"Case Study: Aera Energy's Comprehensive Focus on Data Quality Generates Competitive Advantage"

"Case Study: Smith & Nephew Focuses on Data Governance as First Step in MDM Program"

Note 1 Example of Quantifying the Impact of Data Quality

By decreasing the amount of returned mail by 10%, a healthcare plan with 500,000 providers can realize \$400,000 in savings over three years (source: ["The importance of data quality in producing savings,"](#) Healthcare Finance News). SiriusDecisions, a sales and marketing research firm, quantifies data quality using the 1-10-100 rule, which says "It takes \$1 to verify a record as it's entered, \$10 to cleanse and de-dupe it and \$100 if nothing is done, as the ramifications of the mistakes are felt over and over again." On the revenue side of the equation, a data quality strategy and targeted data quality improvement efforts that solve conflicts at the source can lead to a 25% increase in converting inquiries to marketing-qualified leads (source: ["Data Quality Practices Boost Revenue by 66 Percent,"](#) destinationCRM.com).

Note 2 Six Sigma

M. Harry and R. Schroeder, "Six Sigma: The Breakthrough Management Strategy Revolutionizing the World's Top Corporations," Doubleday Business, January 2000.

Regional Headquarters

Corporate Headquarters

56 Top Gallant Road
Stamford, CT 06902-7700
USA
+1 203 964 0096

European Headquarters

Tamesis
The Glanty
Egham
Surrey, TW20 9AW
UNITED KINGDOM
+44 1784 431611

Asia/Pacific Headquarters

Gartner Australasia Pty. Ltd.
Level 9, 141 Walker Street
North Sydney
New South Wales 2060
AUSTRALIA
+61 2 9459 4600

Japan Headquarters

Gartner Japan Ltd.
Aobadai Hills, 6F
7-7, Aobadai, 4-chome
Meguro-ku, Tokyo 153-0042
JAPAN
+81 3 3481 3670

Latin America Headquarters

Gartner do Brazil
Av. das Nações Unidas, 12551
9° andar—World Trade Center
04578-903—São Paulo SP
BRAZIL
+55 11 3443 1509

© 2011 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. or its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. The information contained in this publication has been obtained from sources believed to be reliable. Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information and shall have no liability for errors, omissions or inadequacies in such information. This publication consists of the opinions of Gartner's research organization and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice. Although Gartner research may include a discussion of related legal issues, Gartner does not provide legal advice or services and its research should not be construed or used as such. Gartner is a public company, and its shareholders may include firms and funds that have financial interests in entities covered in Gartner research. Gartner's Board of Directors may include senior managers of these firms or funds. Gartner research is produced independently by its research organization without input or influence from these firms, funds or their managers. For further information on the independence and integrity of Gartner research, see "Guiding Principles on Independence and Objectivity" on its website, http://www.gartner.com/technology/about/ombudsman/omb_guide2.jsp.