research

BEST PRACTICES REPORT

Q4 2017

What It Takes to Be Data-Driven

Technologies and Best Practices for Becoming a Smarter Organization

By Fern Halper and David Stodder

tdwi
Transforming Data With Intelligence™

Research Sponsors

Hortonworks

Looker

SAP

SAS

What It Takes to Be Data-Driven

Technologies and Best Practices for Becoming a Smarter Organization

By Fern Halper and David Stodder

Table of Contents

Research Methodology and Demographics . . . 3

Executive Summary . . . 4

An Introduction to the Data-Driven Business . . . 5

- Definitions and key characteristics for the data-driven business . . . 5
- Analytics technologies supporting data-driven decisions . . . 6
- Data management technologies supporting data-driven decisions . . . 8
- Making more decisions using data and analytics . . . 9
- Barriers to becoming data-driven . . . 9

The Current State of the Data-Driven Enterprise . . . 10

- Data-driven departments . . . 11
- Investing to become data-driven . . . 12
- Satisfaction with analytics strategy . . . 13
- Satisfaction with performance management . . . 14

Technology Strategies for Becoming Data-Driven . . . 16

- Data management strategies . . . 18
- Big data and open source technology strategies . . . 18
- Satisfaction with data management platforms . . . 19

Analytics Strategies for Becoming Data-Driven . . . 20

- Self-service and BI/analytics democratization . . . 21
- Ease of use and automated model building . . . 24
- Embedding analytics/real-time analytics for action . . . 25

Organizational Strategies for Becoming Data-Driven . . . 26

- Leadership strategies . . . 26
- Cultural strategies . . . 28
- Governance strategies for data-driven organizations . . . 29

Putting It All Together . . . 31

Vendor Solutions . . . 32

- Hortonworks . . . 32
- Looker . . . 33
- SAP . . . 33
- SAS . . . 33

Recommendations . . . 34

About the Authors

FERN HALPER, Ph.D., is vice president and senior director of TDWI Research for advanced analytics. She is well known in the analytics community, having been published hundreds of times on data mining and information technology over the past 20 years. Halper is also coauthor of several Dummies books on cloud computing and big data. She focuses on advanced analytics, including predictive analytics, text and social media analysis, machine learning, AI, cognitive computing, and big data analytics approaches. She has been a partner at industry analyst firm Hurwitz & Associates and a lead data analyst for Bell Labs. Her Ph.D. is from Texas A&M University. You can reach her by email (fhalper@tdwi.org), on Twitter (twitter.com/fhalper), and on LinkedIn (linkedin.com/in/fbhalper).

DAVID STODDER is senior director of TDWI Research for business intelligence. He focuses on providing research-based insight and best practices for organizations implementing BI, analytics, performance management, data discovery, data visualization, and related technologies and methods. He is the author of TDWI Best Practices Reports and Checklist Reports on data discovery, data visualization, customer analytics in the age of social media, BI/DW agility, mobile BI, and information management. Stodder has provided thought leadership on BI, information management, and IT management for over two decades. He has served as vice president and research director with Ventana Research, and he was the founding chief editor of *Intelligent Enterprise*, where he served as editorial director for nine years. You can reach him at dstodder@tdwi.org, @dstodder on Twitter, and on LinkedIn at linkedincom/in/davidstodder.

About TDWI Research

TDWI Research provides research and advice for data professionals worldwide. TDWI Research focuses exclusively on data management and analytics issues and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of the business and technical challenges surrounding the deployment and use of data management and analytics solutions. TDWI Research offers in-depth research reports, commentary, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.

About the TDWI Best Practices Reports Series

This series is designed to educate technical and business professionals about new business intelligence (BI) technologies, concepts, or approaches that address a significant problem or issue. Research for the reports is conducted via interviews with industry experts and leading-edge user companies and is supplemented by surveys of BI professionals. To support the program, TDWI seeks vendors that collectively wish to evangelize a new approach to solving BI problems or an emerging technology discipline. By banding together, sponsors can validate a new market niche and educate organizations about alternative solutions to critical BI issues. To suggest a topic that meets these requirements, please contact TDWI senior research directors Fern Halper (fhalper@tdwi.org), Philip Russom (prussom@tdwi.org), and David Stodder (dstodder@tdwi.org).

Acknowledgments

TDWI would like to thank many people who contributed to this report. First, we appreciate the many users who responded to our survey, especially those who agreed to our requests for phone interviews. Second, our report sponsors, who diligently reviewed outlines, survey questions, and report drafts. Finally, we would like to recognize TDWI's production team: James Powell, Peter Considine, Lindsay Stares, and Rod Gosser.

Sponsors

Hortonworks, Looker, SAP, and SAS sponsored the research and writing of this report.

Research Methodology and Demographics

Report purpose. This in-depth Best Practices Report examines how organizations become data-driven. It looks at patterns for building out infrastructure for managing data and driving analytics. It examines the best practices of those organizations that are data-driven.

Terminology. Business and IT leaders want to enable knowledge workers to make informed, evidence-based decisions—that is, to be able to formulate and answer business questions easily using relevant data, reports, visualizations, and analytics—and take action on these decisions. This is a data-driven organization.

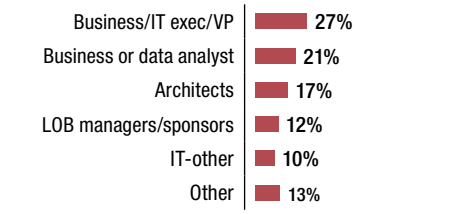
Survey methodology. In August 2017, TDWI sent an invitation via email to the business intelligence (BI) and data professionals in our database, asking them to complete an online survey. TDWI also posted the invitation online and in publications from TDWI and other firms. The survey collected responses from 289 respondents. A total of 173 respondents completed all questions. All responses are valuable and so are included in this report's data sample. This explains why the number of respondents varies per question.

Research methods. In addition to the survey, TDWI Research conducted telephone interviews with technical users, business sponsors, and BI/analytics experts. TDWI also received briefings from vendors that offer products and services related to these technologies.

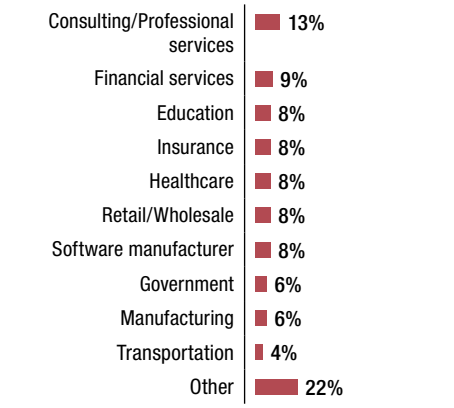
Survey demographics. The majority of survey respondents are business and IT executives (27%), followed by analysts (21%), architects (17%), and line-of-business (LOB) managers and sponsors (11%).

A wide range of industries is represented in this study. These include consulting (13%), financial services (9%), education, insurance, healthcare, retail, and healthcare (8%) Most survey respondents reside in the U.S. (62%) or Europe (12%). Respondents come from enterprises of all sizes.

Position

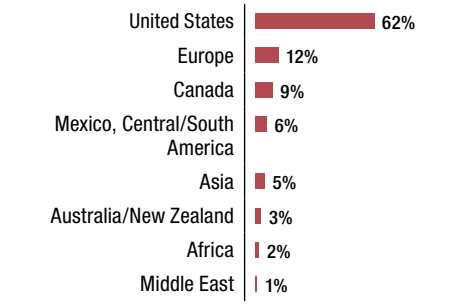


Industry

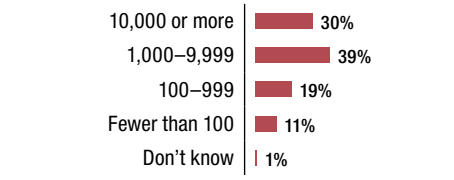


("Other" consists of multiple industries, each represented by less than 4% of respondents.)

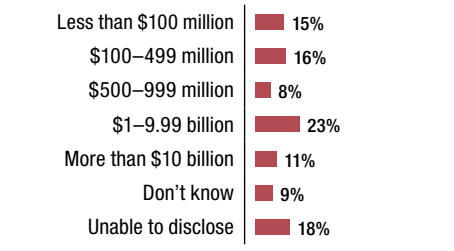
Geography



Number of Employees



Company Size by Revenue



Based on 173 respondents who completed every question in the survey.

Executive Summary

The term *data-driven* has been used in the market around BI and analytics for years. The majority of organizations realize that gut instinct alone is not enough to enable decisions that will drive success. Becoming a data-driven organization, however, has many dimensions. Although instituting performance management metrics and methods is often how organizations begin to develop an enterprise strategy for using data to drive decisions, becoming data-driven also requires rigorous technology, analytics, and organizational strategies.

One of the most popular strategies for becoming data-driven is to democratize analytics.

In this study, roughly a third of respondents do not believe they are close to being data-driven, and another third felt they are becoming more data-driven. Another third are on the fence. Respondents cite numerous barriers to becoming data-driven, including leadership, technology, and skills. They cite dissatisfaction with some of their current data management and analytics practices. There is certainly room for improvement in many areas. On the good news front, the majority (83%) are on the path to becoming more data-driven and they are investing to do so.

Organizations are using a range of analytics strategies to get there. One of the most popular strategies is to democratize analytics. A majority of survey respondents are using self-service visualization tools to help both business analysts and business users get the insights they need to drive decision making. Those that are data-driven are also using tools such as self-service data preparation. Early adopters are using other technologies as well, including more advanced analytics such as predictive analytics. Some are making use of easy-to-use advanced analytics tools that help automate model building. Although the vast number of respondents use analytics simply to inform their decisions, others are embedding analytics into systems and processes that drive action. However, more than half of our survey respondents are dissatisfied with user reliance on IT to accomplish analytics projects.

Although many aspects of data, BI, and analytics still seem to be IT-dominated, clearly some areas should be owned by IT rather than the business. For instance, IT is emerging as owning data-related issues and business owns the business questions. Collaboration between the two groups—critical for becoming a data-driven organization—is also improving.

This TDWI Best Practices Report examines how organizations become data-driven. It looks at patterns for building out infrastructure for managing data and driving analytics. It examines the best practices of those organizations that are data-driven. The report also documents strategies across three areas we believe are important in order to become data-driven. These include technology, analytics, and organizational strategies.

An Introduction to the Data-Driven Business

Definitions and key characteristics for the data-driven business

Organizations are at an important moment in analytics. A majority of businesses believe in the power of BI and analytics to help drive insight and value. Organizations are analyzing data to understand customer behavior, to improve operational efficiencies, for forecasting, and for a host of other use cases. TDWI research indicates that the vast majority of organizations are using technology such as visual analytics and dashboards to help them gain insight.

However, *gaining* insight and *using* that insight to make decisions are often two different things. Many organizations engage in various degrees of analyzing data, but they do not consistently use it to take action. As Thomas Edison said, "The value of an idea lies in the using of it."[1] Analyzing data for its own sake might be interesting and informative, but it doesn't necessarily provide value. Value emerges when an organization uses the results of its analysis (even if it is a probability) to help make decisions. In other words, an organization is data-driven when it uses data and analysis to help drive action—even if that action is a deliberate inaction.

Gaining insight and using that insight to make decisions are often two different things.

Data-driven decisions, of course, come in different shapes and sizes. There are strategic decisions that help to set the course for the organization. Data for strategic decision making might come from all over the company, and analysis might be performed to support a proposed strategy. This type of analysis is often future-focused and commissioned by senior executives. Tactical decisions are shorter in nature and are typically decisions made to address current issues. There are also day-to-day operational decisions made by employees. Finally, organizations today are often programming decisions that are repetitive and routine. For example, should we offer a credit card to this customer? What product recommendation should we make to a potential customer?

From a technology perspective, data-driven decisions may utilize multiple analytics tools and techniques. For example, one organization might use analytics to help make tactical decisions and take action on sales. This might happen manually using tools such as data visualization and forecasting. Another organization might be further along in their analytics sophistication. It might use IoT sensor data along with predictive models to determine when it should service its fleet—automating those decisions prior to a disruptive breakdown.

All of these organizations are data-driven to some degree, although the companies might be at different stages of analytics maturity. In fact, there are many dimensions to being data-driven. These include technology as well as non-technology factors such as organizational, cultural, and governance considerations. As we will see, often it isn't technology but other issues that get in the way of becoming data-driven. Becoming data-driven is both an organizational ambition and an imperative that involves culture as well as technology. It is about using data to take decisive action in addition to building relationships and trust around both the data itself and how people look at the data and perform analysis. Figure 1 lists some important attributes of the data-driven organization.

Often, it isn't the technology but other issues that get in the way of becoming data-driven.

Technology Characteristics	Organizational Characteristics
• Integrated data management to support analytics development and deployment • Good data quality • Integrated analytics strategy using a range of tooling • Operationalized analytics/analytics in production driving actions • Monitored health of analytics output and feedback to continually improve	• Goal-driven • Strong leadership • Transparent • Empowering • Self-service culture • Skills-based • Collaborative • Governance-oriented

Figure 1. *Some attributes of the data-driven organization.*

[1] http://quotationsbook.com/quote/20071/

Analytics technologies supporting data-driven decisions

Numerous technologies can support data-driven decisions across a range of use cases. Typically, when organizations think about analytics technologies to help make data-driven decisions, they are thinking about tools such as:

Performance management has been a major focus of BI and data warehousing.

Performance management. Performance management methods and tools help users focus data access, analysis, and, ultimately, performance on important objectives and project goals. Objectives usually range from broader goals such as improving customer satisfaction to narrowly defined goals oriented specifically to an employee's responsibilities, such as sales quotas or completing a process by a deadline. Performance management has therefore long been a major focus of BI and data warehousing to enable business users to work with data-driven metrics effectively to guide decisions and actions. Instead of giving users voluminous reports that require them to search for relevant data, performance metrics can make it easier for users to consume data within a context and access the most important data faster.

To define key performance indicators (KPIs) and other metrics accurately and consistently, executives and managers need to examine BI reports and analytics. Data warehouses and data marts often play an important role in providing access to carefully cleansed and structured data to support performance management metrics. Data quality and consistency are critical to the success of metrics because if users do not trust the data, they will not trust the metrics.

With well-defined metrics that are up-to-date, performance management can help organizations align decisions and actions with objectives. The objectives provide valuable context for analyzing data. Decision makers thus need technology that gives them a means of interacting with data relevant to the performance metrics. If users are unable to examine the data—that is, if they only receive canned reports or reports that are too old—they cannot use data effectively to improve performance. This can render the metrics less valuable for guiding decisions and actions.

Dashboards have evolved to serve as a user's portal to performance metrics and a range of BI reports.

Dashboards and scorecards. Dashboards and scorecards are commonly how organizations communicate performance management objectives. Most BI solutions support dashboards on mobile devices, natively or via the Web, as well as on desktops, laptops, and workstations. Early on, organizations deployed dashboards in tandem with performance management scorecards, which are a type of visualization designed for tracking metrics, often within the context of a management methodology. Dashboards have evolved to serve a broader purpose as the user's portal (or "dashboard," as in a car) for not just performance metrics but the broader range of visualizations, including graphs, heat maps, and gauges, and text feeds.

Dashboards can be standalone applications or deployed as an embedded feature within an application, such as an ERP or CRM system. Organizations also develop them to serve specific vertical industry, line-of-business, or departmental requirements. The increased specialization can make dashboards more relevant; however, highly specialized dashboards often multiply in organizations, forcing users to go from one to the next as they switch applications. Organizations will often try to consolidate dashboards as part of a reduction in the number of data and front-end silos.

Dashboards need to be easy to use and provide enough context so that users are not confused or misled. The key issue is whether users understand what they see and can take action based on the information. The learning curve must fit the capabilities of the users, who may not be data-savvy and need to understand the information within the context of their subject matter expertise or role in a process. To support daily, operational decisions, the information conveyed by dashboards must be timely enough to fit users' requirements. TDWI research finds organizations vary in the frequency with which they update dashboards, from once a day to multiple times a day or more frequently depending on the purpose and their BI platform's capabilities.

Self-service visual analytics and data discovery. The focus of solutions in this area is to address business user needs that go beyond data consumption. Users want to analyze data, but they don't want to leave the easier-to-use interface and graphical experience typified by a dashboard and manually write queries or programs. Newer solution providers found a market opportunity because of the frustration typical enterprise BI users experienced with inflexible applications that could not support the kind of ad hoc data interaction needed to do more advanced analytics. Slow IT deployment of BI application functionality and access to new data also frustrated users. Visual analytics and data discovery solutions offered greater self-service functionality than older enterprise BI solutions.

With these solutions, users can pursue data insights via modern graphical interfaces and visualizations, including dashboards. Self-service capabilities enable them to choose data sets, query the data, and create visualizations on their own. Most solutions come with libraries of visual representation types such as charts, heat maps, and scatterplots; many solutions allow users to expand their options by importing visualizations from outside sources, including open source. The tools are also useful for easier, small-scale testing of prototypes and analytics models before deployment at a bigger scale and against larger data sources.

Data preparation and integration. Advances in data preparation are critical to enabling users to advance with visual analytics solutions as they seek to access a greater variety of data sources. Data preparation involves a range of processes that begin with data ingestion and collection and run through quality improvement and transformation. These processes are often slow and complicated and require significant manual effort, which can prevent BI and analytics from playing a more integral role in daily decision making. Users typically must wait for IT to prepare the data or take on the task themselves with substandard tools and less-consistent methods, which can introduce errors and inconsistencies.

Advances in data preparation are critical to enabling users to work with visual analytics solutions and drive decisions.

Self-service functionality is evolving for data preparation and integration—using terms such as data blending, wrangling, and munging—to enable users to explore data and choose data sets that fit their BI and visual analytics processes. Self-service functionality is also maturing for use in the development of data catalogs, glossaries, and metadata repositories. These are critical to enabling users to gain complete views of data and to share both data and insights based on the data with others.

Advanced analytics. Advanced analytics includes tools such as predictive modeling and machine learning. These tools can help users find patterns in data that can drive decisions. For instance, predictive analytics might identify which customers might be at risk of dropping a service. Organizations can use that insight along with other analysis to make decisions regarding how to retain that customer. TDWI research indicates that there is considerable excitement around advanced analytics tools such as predictive analytics and machine learning to help users make better decisions. In fact, TDWI survey respondents often cite better decision making as a top driver for these kinds of tools.[2] These tools have a place in the emerging area of automation, which can also help an organization become more data-driven.

There is much excitement around predictive analytics to help drive better decision making.

Automation. There are two flavors of automation worth mentioning here. The first is automating analytics for decision making. Vendors are offering a range of tools that automate the analytics life cycle, from data preparation to model building. These use advanced analytics such as machine learning or natural language processing embedded into software products. For example, some software provides an NLP interface to allow business users to ask questions using natural language interaction. Other software automatically corrects for data quality while other solutions direct attention to salient points or provide automated predictive model building. These products can help organizations come to insight faster for decision making.

[2] See, for instance, the 2017 TDWI Best Practices Report on AI, machine learning, and NLP at tdwi.org/bpreports

Other software provides automation for analytics in production. This includes embedding analytics into dashboards to alert users if there is a problem—for instance, on the factory floor. Organizations use predictive analytics in production to help alert personnel or systems to issues such as fraud. Systems can route potentially fraudulent transactions to a special investigation unit for further processing. Firms can use automated and embedded analytics in small, repeatable decisions such as those found in recommendation engines or credit scoring for authorizing a new credit card. This is often referred to as decision management. We discuss automation further in the Analytics Strategies section of this report.

Data management technologies supporting data-driven decisions

The on-premises data warehouse has been supporting enterprise BI and reporting needs for some time. These data warehouses are typically used when the kinds of questions and the output are known. However, as organizations move to analyze disparate data types or realize that they need to make more real-time decisions, they will most likely adopt other data management platforms. This is all part of what TDWI refers to as an MDA (multiplatform data architecture)—i.e., an eclectic mix of old and new data managed on traditional and modern platforms. Some of the newer platforms that are part of this architecture include:

Data lakes. A data lake is a collection of often diverse data that can scale to tens or hundreds of petabytes in size. Although the warehouse is used for reporting and other kinds of BI, such as visualization or dashboards, the data lake is often used for advanced analytics. Although the majority of data lakes are Hadoop-based, they need not be. Previous TDWI research indicates that this is a top area for growth in terms of platforms supporting advanced analytics for decision making.[3] In fact, TDWI research indicates that the data lake and the relational warehouse are complementary, which is why an increasing number of data warehouse teams deploy both and integrate them tightly.

NoSQL. NoSQL (loosely defined as "Not only SQL") is an umbrella term for technologies that specialize in nonrelational storage and retrieval of data. These include key-value store, graph databases, content or document database systems, and nonrelational columnar databases. With these technologies, users can gain new perspectives on data relationships, examine context around BI reports and KPIs, and search and analyze huge volumes of multistructured data including, for example, unstructured big data generated by customer behavior and social media.

Hybrid database technology. Along with the ability to access live data from business applications and transaction databases, in-memory computing is an important supporting technology for visual analytics and discovery. In-memory computing exploits larger random access memory (RAM) space for holding data and performing analytics. Data platform solutions often pair in-memory computing with compression technology and columnar databases to support faster, compute-intensive analysis against large volumes of detailed data rather than just samples or summaries.

An emerging class of databases combine in-memory technology and a scale-out architecture and bring together transaction/operational processing and reporting and analytical processing. This means that analysis can occur as soon as the data hits the database, making it useful for real-time data insights that support true real-time decisions. Examples include applications for financial trading systems, business activity monitoring, utility grid monitoring, e-commerce product recommendations, and facility surveillance. Many of these databases are also ACID compliant.

The cloud is an important piece of this evolving environment. More often, these platforms are offered in the cloud, and so it makes sense to analyze and act on data in the cloud.

[3] See, for instance, the 2017 TDWI Best Practices Report on AI, machine learning, and NLP at tdwi.org/bpreports

Why it is important to make more decisions using data and analytics

Instinct alone does not make for good decisions. Data and analytics—fact-based analyses of trusted data—can help organizations become more successful.

Studies support this hypothesis. For example, more than five years ago, Andrew McAfee and Erik Brynjolfsson from MIT found that the more companies characterized themselves as data-driven, the better they performed on objective measures of financial and operational results. In particular, companies in the top third of their industry in the use of data-driven decision making were, on average, 5% more productive and 6% more profitable than their competitors were.[4] At TDWI, we repeatedly see that those organizations that use analytics and become sophisticated with their analyses are more likely to measure top- and bottom-line impact than those that do not. It is a success cycle—as an organization uses analytics for insight and action, it gets better results and those better results drive success. The company then starts to build on its success, perhaps bringing in new analytics techniques and people to manage them. The company continues to build on its progress, becoming more sophisticated and more successful.

At TDWI, we see that those organizations that use analytics and become more sophisticated with their analyses are more likely to measure top- and bottom-line impact than those that do not.

Ultimately, companies that use data to drive decisions do not just become more profitable—they also become smarter. An organization is smart when it is willing to try new ways to improve and to turn problems on their heads and innovate with analytics. As the culture becomes more data-driven and companies begin to experience analytics success, they start to do things differently and smarter. This can lead to more top-line impact. For instance, they may migrate from using analytics and data to help drive decisions to using analytics to make certain kinds of decisions automatically. This is all part of what TDWI refers to as the analytics success cycle.

Barriers to becoming data-driven

Although becoming data-driven does provide value, it still poses challenges to many organizations. In fact, regardless of how far along organizations are in their journey to become data-driven, they still cite the following five barriers to improving the use of data and analytics to drive decisions, manage daily operations, and develop strategy (Figure 2). Interestingly, the lack of software solutions ranks at the bottom of the list of barriers.

Lack of business executive support is often a top barrier to becoming data-driven.

- **Lack of business executive support/corporate strategy.** Trust and culture are two big roadblocks for any organization looking to utilize analytics. Typically, it is easier to move forward with analytics if a business executive is on board. This person can help set the tone and the vision for the effort as well as provide necessary funding and organizational support. TDWI has seen numerous cases where an organization only moved forward in analytics once there was a new executive in place to help drive the effort. That is not to say it can't be done without the support of a senior executive, but it may take longer. In this survey, 42% of respondents cited this lack of support as a barrier.
- **Difficulty accessing and integrating all relevant data.** It is hard for an organization to achieve insights upon which to take action if it cannot access the right data for analyses. Sometimes this is a technology challenge: the organization doesn't have the tools to integrate data. Other times it is a skills issue in terms of putting the right people in place to bring together data silos. This is often an organizational issue: either the funding doesn't exist to make this happen or the politics are such that users do not share data. In our survey, 37% of respondents cited this as a barrier.
- **Lack of skills to build out BI/analytics to support decision making or action.** A skill shortage always ranks high on the list of challenges in any analytics-related effort. The required skills are focused on both data management and analytics. They can include the

[4] https://hbr.org/2012/10/big-data-the-management-revolution

skills to perform visual analytics using self-service tools as well as building out predictive models. This can also include back-end processes such as embedding analytics into applications or putting models into production. In this survey, 34% of respondents cited this as a barrier.

- **Data quality is insufficient.** Data quality is very important when you're making data-driven decisions. As the old saying goes, garbage in, garbage out. Accuracy, integrity, timeliness, reliability, and consistency are all important here. Twenty-seven percent of respondents said this was a problem.
- **Data governance and security concerns.** A number of these issues relate to data governance. TDWI research indicates that organizations are most successful with analytics when business and IT work together to provide policies and guidance for governing data. This includes determining a common vocabulary and definitions—these also play into data quality as well as policies about who can access what data—which are important for obtaining relevant data for analysis. In this survey, 27% of respondents believed this was a barrier to becoming data-driven.

We will address best practices for dealing with these and other challenges later in this report.

In your organization, which of the following factors present the biggest barrier to being data-driven?

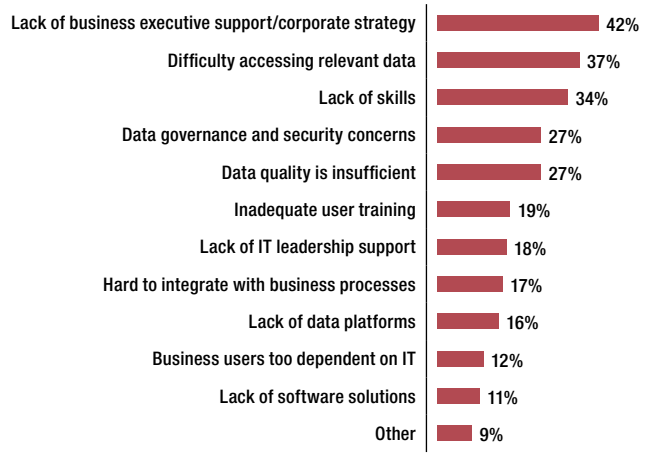


Figure 2: *Based on 264 respondents. Note that "software solutions" is not among the top 5 factors.*

The Current State of the Data-Driven Enterprise

Thirty-one percent of respondents felt they were fairly close or very close to being data-driven.

To understand the current state of the data-driven enterprise, we asked respondents a series of questions, the first of which was, "How close is your organization to reaching its goals for being a data-driven organization?" Respondents answered the question on a scale from one to seven, where one was not close at all and seven was very, very close. Thirty-nine percent of respondents did not believe they were close to being data-driven; 31% felt they were fairly, very, or extremely close to being data-driven; however, only 11% believed that they were very close or extremely close. The rest (29%) were on the fence (all not shown). In this report, when we look only at the group of respondents who felt that their firms were not close to being data-driven, we will refer to that group as the "not-data-driven group." Conversely, when we call out the group of respondents who believe that their firm *is* close to being data-driven, we will refer to it as the "data-driven group."

Interestingly, although many respondents do not consider their organizations close to being data-driven, they do believe that they are moving in the right direction to enable more effective use of data and analytics in decision making. In fact, 83% of respondents were somewhat or very confident that they are moving in the right direction (not shown). This movement should include employing organizational best practices as well as the use of technology for analytics.

Data-driven departments

We asked respondents which three departments in their companies are the most advanced in their ability to use data and analytics to drive decisions, manage operations, and develop strategy (Figure 3). There was a wide range of answers.

Finance, IT, operations, and marketing are the most data-driven departments.

Finance, IT, operations, and marketing rank highest. Not surprisingly, finance, IT, operations, and marketing rank at the top of the list among all respondents, although only about 20-30% of respondents rank them in the top three. It makes sense that respondents would regard their finance and IT departments as data-driven; the nature of the work requires it. Previous TDWI research indicates that marketing is often one of the first areas to adopt analytics in a company so that managers can better understand customers or drive smarter campaigns. This group hopefully uses its analytics to become more data-driven. Likewise, BI and analytics are important in operations and TDWI research indicates that operations often ranks at the top of the list of organizations using advanced analytics to improve efficiencies.

Executive management ranks mid-range. Interestingly, only 17% of respondents cited their executive management as being in the top three departments that use data to drive decisions. This suggests that there are often problems in getting executives on board. For instance, a 2016 Ernst and Young study found that analytics can upset the balance of power in the C-suite. There are many personalities who may feel vulnerable when someone else owns the analytics that demonstrates how well their teams are performing.[5] It is up to the CEO to help drive this effort among executives. As we will see in the next section, investing in technology for executives to become more data-driven is a priority among many respondents.

Which departments in your company are the most advanced in their ability to use data and analytics to drive decisions?

Figure 3. *Based on 284 respondents choosing their top three departments.*

[5] https://hbr.org/2016/08/the-reason-so-many-analytics-efforts-fall-short#comment-section

Other departments ranked very low. Some departments were less common choices in terms of their ability to use data and analytics to drive decisions. These include logistics (3%), supply chain (8%), and HR/workforce management (6%). Also ranking low on the list was GRC (governance, risk, and compliance) at 10%.

Investing to become data-driven

Over three-fourths (76%) of respondents anticipate that their budgets will increase somewhat or significantly over the next few years.

We asked respondents whether they anticipated that their organization's budget would increase, decrease, or remain about the same for technologies, services, and training to become data-driven. The good news is that organizations realize that they need to invest more to be successful. The vast majority (76%, not shown) said they anticipated that their budgets to become more data-driven would increase somewhat or significantly over the next three years. Less than 3% felt their budgets would decrease (not shown).

We asked respondents about recent significant investments in BI, analytics, and data management technology and services that help to support data-driven decisions (Figure 4). Top investments focus around supporting decisions for executives, customers, and operations. For example, 28% of the respondents are putting a significant investment in improving executives' views of performance. This may involve interactive dashboards that include performance metrics, mentioned earlier. Twenty-nine percent of respondents stated that their organization was investing in developing and applying predictive insights. That is good news and points to the excitement around predictive analytics also mentioned earlier.

Many organizations realize the power of this technology to help them in understanding behaviors (customer, operations, patient, etc.) that can drive better decisions. As we will see, that doesn't mean these organizations are necessarily putting these models into production. However, they may be using them to inform decisions, such as how to price an insurance policy based on risk profiles or what customers are at risk of churn. Other important areas for investment include optimizing operations and process management (27%) and customer-focused data and analytics efforts such as gaining a complete view of the customer (24%) and improving customer experience and engagement (20%).

Which of the following objectives are the focus of your organization's most recent significant investment in BI, analytics, or data management?

***Figure 4**. Based on 264 respondents. Respondents could select up to three responses.*

Satisfaction with analytics strategy

Organizations can be more successful in becoming data-driven if they develop a broader analytics strategy that addresses the range of concerns that arise rather than confronting them piecemeal. What are the most important goals and considerations that organizations should address in their analytics strategy? How satisfied are organizations currently with how they are addressing them?

We asked respondents to rate the level of satisfaction among users and managers with important aspects of their organization's analytics strategy (Figure 5). We found the highest satisfaction is with accuracy of insights; 64% expressed some level of satisfaction, although just 10% said they were very satisfied. Not all types of analytics are or need to be accurate; depending on the quality of the data and analytics requirements, analysts may expect a significant margin of error and plan to do a fair amount of testing to pick a winner in a study. However, for finance and some types of operational reporting-based analyses, accuracy is vital. The results suggest that organizations have more experience with these types of BI and analytics and that the majority is reasonably satisfied with the accuracy of insights users and managers are receiving. This is in line with the fact that finance is a top area for data-driven decisions, mentioned above.

Nearly two in three respondents (64%) are satisfied with the accuracy of their insights, although only 10% are very satisfied.

Model management, reproducibility, reuse, and taking action need attention. Just over half (54%) of respondents said users and managers are satisfied to at least some degree with their ability to reproduce analysis; 13% said they were very satisfied and 40% said they were somewhat satisfied. To feel confident about using analytics for decision making, users often want to see if the insights remain relatively constant even as the data is refreshed or as the study is repeated perhaps a month or two later. Decision makers will want to know if the results change when they apply the same methods and models to a different data set. Thus, reproducibility and the ability to reuse code and models should be encouraged through an organization's analytics strategy. Only 31% of respondents indicated that users and managers are satisfied with reuse of code for analytics. Model management, which is critical to successful analytics strategies, can support reproduction and reuse. However, in our survey, less than a third (30%) expressed overall satisfaction with how models were managed, deployed, and refreshed.

Just over half of respondents said users and managers are satisfied with their ability to reproduce analysis.

An analytics strategy must cover not only the development of models and code for analytics but also whether the organization effectively takes action on the analytics. Just over half (52%) showed some level of satisfaction with users' and managers' ability to act on insights. This indicates that there is significant room for improvement in making analytics actionable. Confidence is a large part of whether personnel will make analytics part of their decision making, but timeliness is also a key ingredient. Too often, data insights do not reach decision makers at the time when they would be most valuable. We found that only a little over half (52%) of respondents report satisfaction with the timeliness of analysis and just 7% said their users and manager are "very" satisfied. Analytics strategies therefore should cover how to reduce the latency between the development of insights and when they are actionable for executives, managers, and other users. We discuss some of the ways analytics can drive action in more depth later in this report.

Expansion, scale, and user self-reliance are trouble spots. These three areas, critical to the democratization of analytics, appear to be where respondents have the least satisfaction with their organizations' analytics strategies. More than half (57%) said their users and managers are either somewhat or very dissatisfied with self-reliance, i.e., the ability to accomplish analytics projects with less IT involvement. Where users are highly reliant on IT, expansion is limited by IT's capacity to work on users' projects. Consequently, IT becomes a bottleneck, allowing only a select group of users to engage in advanced analytics beyond spreadsheets and BI reporting.

Trouble spots include expansion, scale, and user self-reliance.

In our survey, 47% are either somewhat or very dissatisfied with the use of analytics by a wide group of people in their organizations. In addition, because analytics projects often demand examination and interaction with as much data as possible, supporting the scalability of analytics to large amounts of data must be part of an analytics strategy. We found that nearly half of respondents (46%) regard their users as dissatisfied with such scalability in their organizations.

Please rate satisfaction with the following aspects of your organization's current analytics strategy

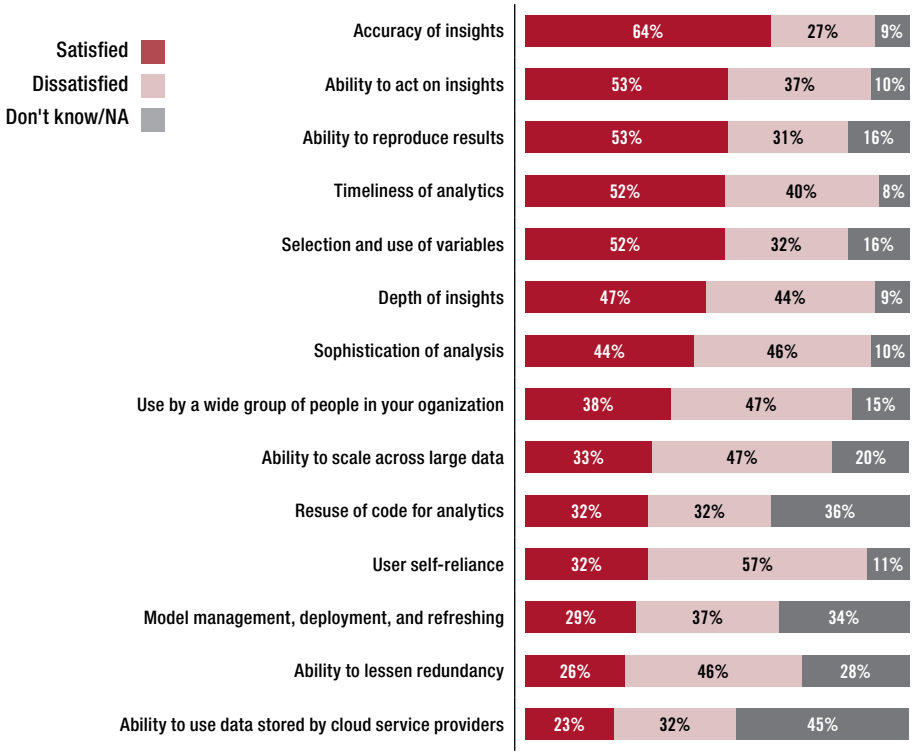


Figure 5. *Based on 183 respondents.*

Satisfaction with performance management

As noted earlier, instituting performance management metrics and methods is often how organizations begin to develop an enterprise strategy for using data to drive decisions. We examined the level of organizations' satisfaction with objectives common to performance management implementations and found that the highest level is with the alignment of KPIs with the organization's goals (43%, Figure 6). Although not an overwhelming percentage, this result shows that many organizations are confident that KPIs accurately communicate what that organization wants to accomplish.

Less than a third of respondents are satisfied that their organizations take action based on performance metrics.

However, only 11% said users and managers are satisfied that they measure the right number of metrics, and not too many; so although alignment may be relatively good, there is dissatisfaction with implementation of metrics and making sure users and managers are not overwhelmed by having to track too many of them. In addition, less than a third (29%) said users and managers were satisfied that their organizations took action based on performance metrics.

For this question, the second-highest level of satisfaction was with respondents who said their users and managers are satisfied with the use of BI and analytics tools to support performance management (39%). The ability to use such tools effectively is important to data-driven performance management. Organizations frequently deploy dashboards to enable users to see and interact with data relevant to performance metrics using visualizations such as charts and gauges. The third-highest percentage of respondents (36%) indicated satisfaction with dashboards for performance management—not as high a percentage as expected, given the level of experience many organizations have with dashboards. One issue may be frustration with the inability to personalize dashboard visualizations and data access; only 18% indicated satisfaction with this aspect of dashboard deployments. In addition, only about one-quarter (26%) said users and managers were satisfied that they have the right data to support metrics and management.

Satisfaction with the use of analytics for performance management is low. Our research finds that analytics do not currently play a satisfactory role in most organizations' performance management initiatives. Only 23% of respondents said their users and managers are satisfied with the use of analytics to define KPIs and metrics, and just 20% indicated that they are satisfied with the use of analytics to interpret performance results.

Anecdotally, however, we find that some leading-edge organizations are seeking to increase the role of analytics in the development of metrics and interpretation of results. For example, an organization might use predictive analytics on machine data to develop threshold metrics for when computer systems or other equipment might fail. They could use predictive analytics to examine streams of machine data for patterns and anomalies. Public health agencies can use the analysis of large data volumes to create data-driven metrics that measure population health and proactively mitigate the spread of diseases; they can also use metrics to monitor how well healthcare service providers in affected areas are responding to a disease and whether they need to reallocate their resources to different locations.

Regarding performance amanagement metrics and methods, which of the following objectives are being accomplished by your organization to the satisfaction of users and managers?

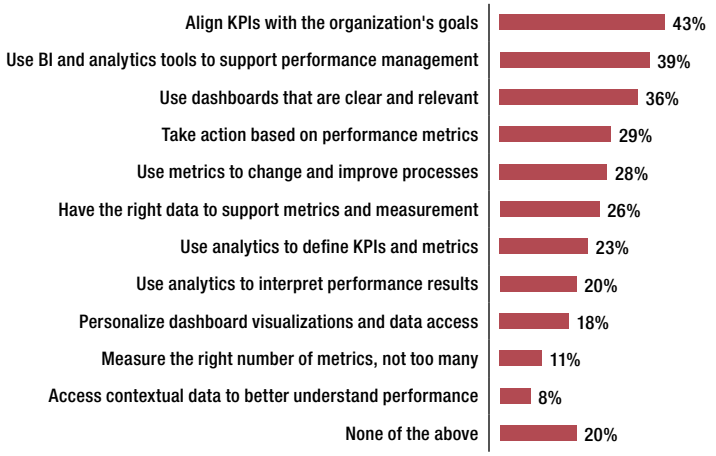


Figure 6. *Based on 192 respondents. Multiple answers allowed.*

Technology Strategies for Becoming Data-Driven

Organizations employ a variety of technologies and practices to realize decision-making value from data, from the simplest to the most advanced technologies. As expected, spreadsheets are ubiquitous; 88% of respondents said they currently use them to drive decisions (Figure 7). Although offering many capabilities and usually supplied as part of users' existing personal desktop applications package, spreadsheets can be challenging for users who want to do more than basic data analysis and visualization.

Nearly three-quarters (71%) of respondents are using data visualization, supporting the trend toward easier-to-use, self-service analytics.

Additionally, as spreadsheets proliferate throughout organizations, they can create headaches for IT data management regarding data quality, consistency, governance, and analytics standardization. Many organizations prefer to steer users toward BI reporting applications, which the great majority of respondents are also currently using (83%).

Offering evidence of the strong trend toward easier-to-use, self-service, visual BI, analytics, and data discovery tools, our research finds that nearly three-quarters of respondents (71%) are using visualization. Another 20% are planning to use visualization within three years. As discussed, performance management is one of the key applications of BI and data warehousing systems; we find that 47% of respondents are currently using BI and analytics technologies to support scorecards for performance management; 30% plan to use these technologies for this purpose within three years.

Search engines, which are important on their own for finding disparate and varied data and content, have increasingly become a component integrated into BI and visual analytics solutions. However, only 26% of respondents' organizations currently use them. Slightly fewer (21%) plan to use search engines within three years; 25% have no plans to use them and 28% said they did not know or found them not applicable.

Technologies need to enable organizations to forecast, predict, and perform simulations. One of the most important—and common—activities organizations do with BI and analytics technologies, including spreadsheets, is forecasting. Nearly half (46%) of respondents said they are currently using these technologies for forecasting, and 39% plan to do it within three years. Applications and platforms must be flexible for forecasting because methods can vary depending on the data available and the purpose of the analysis. Forecasting could be for budgeting, developing strategy, or for informing operational decision makers about factors that could influence immediate outcomes. Visualization is highly useful for both consumption and analysis of forecasts, particularly for time series analysis because it enables users to discover insights and act upon them more easily.

Data management and computing platforms must have capabilities to crunch through data quickly and support multiple iterations, especially as organizations move toward predictive analytics on big data that is voluminous and varied. One-third of respondents are currently using BI and analytics technologies for predictive analytics and 48% are planning to use them for this purpose within three years, which was the highest planned use percentage among all the technologies we listed in this survey question.

Although forecasting and predictive analytics are similar—some regard forecasting as simply a kind of predictive analytics—in practice, predictive analytics involves broader modes of inquiry into the behavior of customers, machines, or other topics of interest, whereas forecasting is a

more traditional activity focused on numbers. Predictive analytics pushes data requirements beyond those that are typical of forecasting.

One in five (20%) respondents is using BI and analytics technologies to perform what-if simulations, and 35% plan to do so within three years (21% have no plans to do so). Forecasting and predictive analysis can be useful for what-if analysis and the development of simulations to evaluate and predict outcomes if certain decisions are made differently. Decision makers want to see what would happen if they were to increase or decrease resources for particular operations and processes, change pricing, alter variables that affect cost/benefit ratios, and more. Some data management and analytics platforms are strong enough today to enable organizations to create fuller simulations involving many replications and iterations through the data. By applying data visualization technology along with such platforms, users can explore simulations from different perspectives and at macro or micro levels.

Which of the following types of BI and analytics technologies do you use to drive decisions? Currently? In the next 3 years?

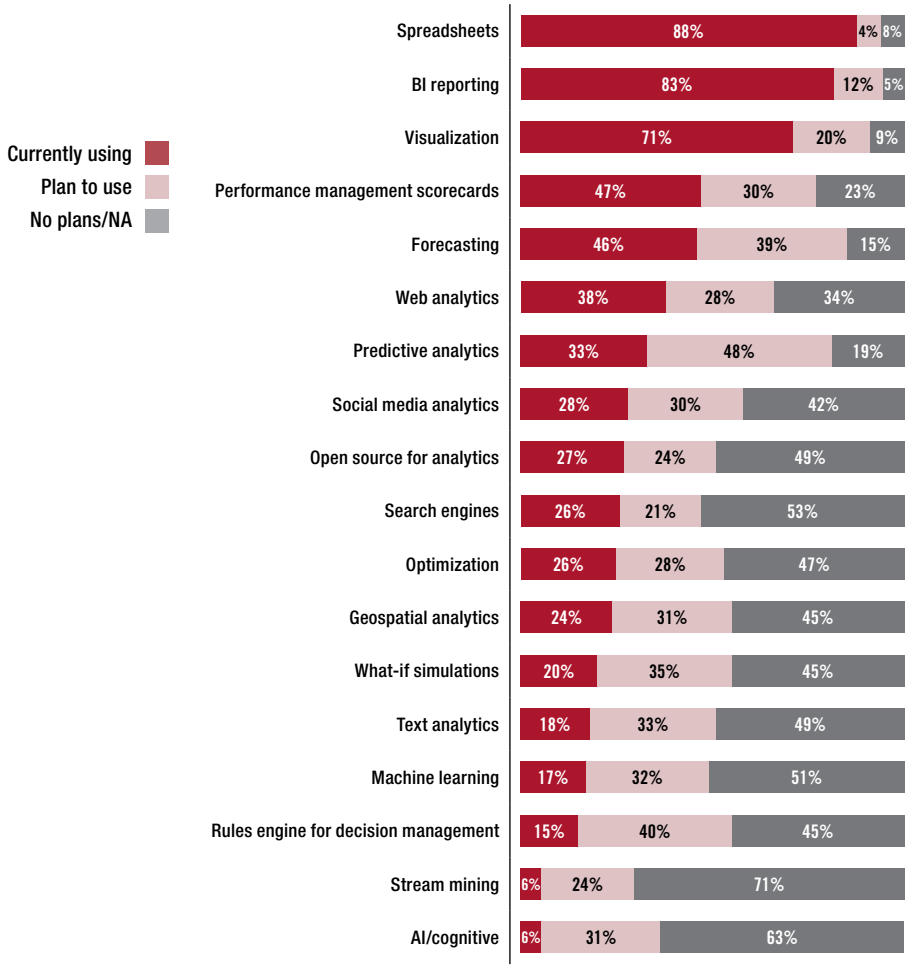


Figure 7. *Based on 186 respondents.*

Data management strategies

As stated previously, many organizations are moving toward a multiplatform data architecture strategy in order to support data-driven decision making. Data is becoming diverse and its frequency is increasing; this data includes traditional data from applications and newer Web applications as well as real-time IoT data. To deal with this data complexity, data management solutions are also becoming more heterogeneous and complex. The multiplatform architecture is an outgrowth of this complexity. Frequently, companies will keep their data warehouse for reporting or multidimensional OLAP but use other platforms for the advanced mining of this diverse data.

This whole environment then is an MDA (multiplatform data architecture). The architecture is a large-scale design pattern and includes numerous, diverse platform types woven together in a fabric. This includes the DBMS and newer systems: cloud-based databases, columnar databases, other specialized databases, and systems that employ Hadoop or Apache Spark. In previous TDWI research, we've seen that only 19% of respondents believe that the data warehouse can support their advanced analytics efforts.[6] Developing a multiplatform architecture is an important data management strategy for organizations that wish to advance their analytics efforts as part of their data-driven strategy. As we will see, it is still the early days for these MDAs.

Big data and open source technology strategies

Open source technologies are becoming important elements of the MDA strategy. In recent years, many organizations have turned to open source technologies such as Apache Hadoop, MapReduce, and Spark to develop distributed big data and analytics platforms that can handle the requirements of predictive and other forms of advanced analytics. In other research studies, TDWI has found that many organizations supplement existing data warehouses with Hadoop "ecosystem" technologies, which include those that implement Spark core and component libraries for data processing, data streaming, machine learning, and more. Although data warehouses tend to support traditional BI applications, big data and analytics platforms serve the needs of data scientists and other analysts who demand specialized software and data processing capabilities that allow them to engage in deeper, computationally intensive analytics of large data sets. Big data platforms also make it easier to manage and access multistructured or unstructured information drawn from sources other than standard business applications, such as Web and user search behavior or sensor and machine data.

Forty percent of respondents said their organizations plan to increase the use of open source technologies.

We asked research participants if their organizations are planning to increase, keep about the same, or decrease use of big data platforms and data lakes that employ open source technologies. Significantly, the largest percentage (40%, not shown) said their organizations plan to increase such use. The percentage rises to 48% when we filter survey results to see only respondents who said their organizations plan to increase investment in technologies, services, and training for enabling data-driven and data-informed decisions, management of operations, and development of strategy.

A much smaller percentage (14%) said their organizations would keep investment in open source technologies about the same, which suggests that organizations do not want just to maintain a small investment in these technologies without enhancing them over time; only 2% said their organizations were planning to decrease use of open source technologies. A much larger percentage of respondents (34%) said they do not use these technologies (all not shown).

[6] See the 2017 TDWI Best Practices Report on AI, machine learning, and NLP at tdwi.org/bpreports.

Satisfaction with data management platforms

Of course, BI and analytics depend on strong data management platforms. We examined the level of user and manager satisfaction in data management platforms for BI and analytics, whether they reside on premises or in the cloud (Figure 8). Interestingly, the highest level of satisfaction was with support for large numbers of users; 51% are either very or somewhat satisfied. This suggests that these organizations are happy with their platform's support for expanded enterprise BI.

However, fewer are satisfied with their platform's ability to scale to handle unexpected requirements; just 28% are either very or somewhat satisfied, and about half (51%) are dissatisfied. This indicates that for these respondents, their data platforms are better suited to standard requirements such as reporting but less suited to ad hoc querying and analytics that go beyond well-defined parameters. Less than half (41%) of respondents said their users and managers are satisfied with data availability for analytics workloads.

Half (50%) of respondents are satisfied with their data management platform's ability to provide data quality and consistency. This suggests room for improvement. For organizations accessing semi- and unstructured content, satisfaction with this capability in their data platforms drops to 21%, while nearly double that percentage (41%) are unsatisfied (38% said they didn't know or regarded this aspect as not applicable). Initial data quality and consistency are often lower for semi- and unstructured content than for structured data coming from business applications.

Half of respondents are satisfied with the ability of their data management platform to provide good data quality.

If we filter results to see only respondents who are satisfied with how well their data management platforms provide access to semi- and unstructured content, we can see that a good percentage are already investing in open source platforms such as Hadoop. We find that 31% are keeping their investment in Hadoop and other open source systems unchanged and 39% plan to increase use of these systems. Looking at all respondents, however, we can see that satisfaction with open source technologies is still somewhat low. Only 23% are satisfied, while nearly the same number (25%) are unsatisfied. Most respondents (52%) answered that either they did not know or they found open source technologies not applicable to them, which suggests that this group of respondents is not currently using Hadoop or other open source technologies for their data platforms.

Satisfaction—and experience with—multiplatform data architecture is somewhat low. Our research finds that only 23% of respondents said their users and managers are satisfied with their ability to view or access all relevant data from both on-premises and cloud-based systems. As organizations exploit cloud platforms for storing and managing data, most need to bridge these with their existing on-premises systems. Multiplatform data architectures are immature at this point, as these research results indicate, but will be important for organizations moving forward.

How satisifed are users and managers with the following aspects of data management platforms either on premises or in the cloud to support analytics?

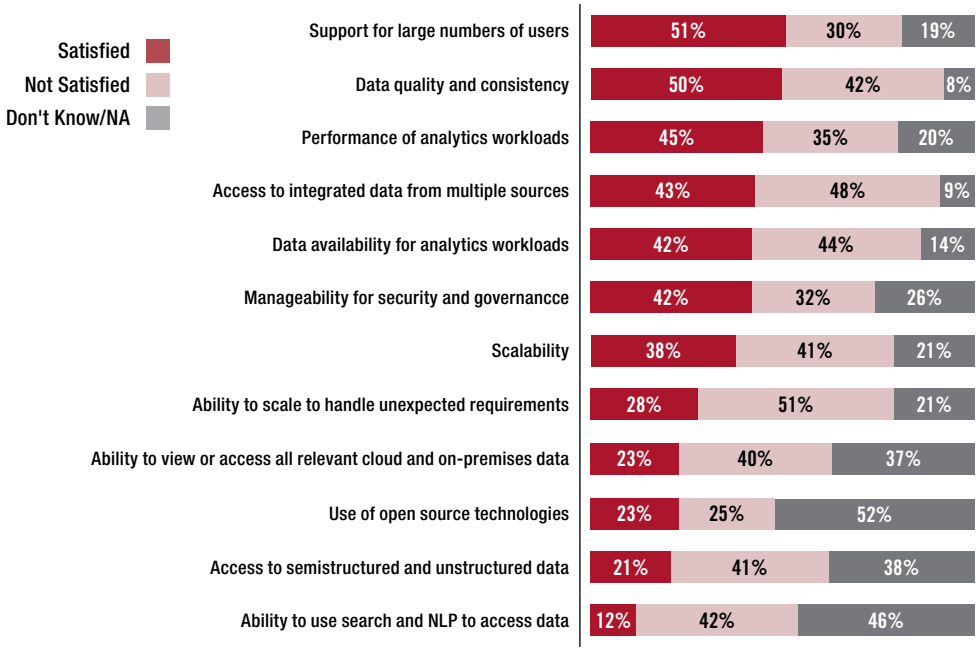


Figure 8. *Based on 175 respondents.*

Analytics Strategies for Becoming Data-Driven

Analytics is at the core of becoming a data-driven organization. What good is data if it is not analyzed for action? Given this, we asked respondents about a series of analytics strategies that might be deployed in their organization. As illustrated in Figure 9, research participants view expanding the use of visual analytics and automating data preparation as their two top strategies for becoming data-driven. In fact, these strategies win over the others by a factor of two to one in the data-driven group.

Visual analytics is viewed as an important strategy for becoming data-driven.

As illustrated earlier in this report, most organizations regard the move to self-service visual analytics as an important analytics strategy for becoming data-driven. This is the case across all respondents, with no significant difference between groups. Other analytics strategies include making use of easy-to-use advanced analytics, embedding analytics for action, and automation.

Which of the following analytics strategies is your organization using to become more data-driven?

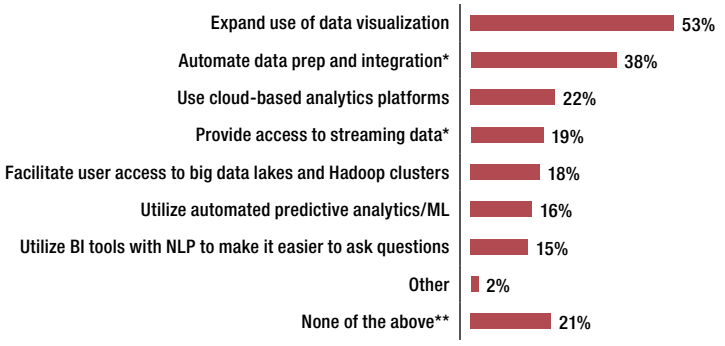


Figure 9. *Based on 193 respondents. Multiple responses allowed. * denotes that the data-driven group is more likely to use this technique. ** denotes the not-data-driven group is more likely to use this technique.*

Self-service and BI/analytics democratization

Fifty-six percent of respondents said users are able to access and query data in a data warehouse.

Technology trends enabling greater ease of use and self-reliance with BI and analytics continue to reshape how users interact with data and apply insights to decisions. Visual BI and analytics solutions make it possible for business users to be more productive with data because the solutions do not require users to engage in manual work themselves or depend on IT developers to enable data access, write queries, create visualizations, and build dashboards. Earlier eras of enterprise BI deployments may have "democratized" some reporting, data access, and analysis, but newer waves of technology take capabilities further and put users in more control of their workspaces.

However, although there is a lot of excitement and hype surrounding self-service BI and analytics, most users are still working with traditional BI. To gauge progress with self-service capabilities, we asked participants what activities users are able to perform in a largely self-service or self-reliant fashion, without IT doing most of the work for them (Figure 10). The largest percentage (56%) said users are able to access and query data in a data warehouse; in other words, they interact with data provisioned and curated for them by IT. With this traditional arrangement, users can rest assured that they are using professionally collected, cleansed, and transformed data, done according to their stated requirements.

Many organizations prefer this approach so they can both provide users with good data and protect both the data warehouse *and* the sources that feed the warehouse. The drawback is that if users want to access new data, make changes to how they want the data transformed, or perform unanticipated queries, they must wait for IT to revise the data warehouse and its related processes. Users can be tempted to set up their own data marts as a faster and more flexible alternative, which typically leads to data management problems associated with disconnected data silos.

Just under half (45%) of respondents said users can create their own visualizations, including dashboards, and almost a third (31%) can perform data exploration and discovery. The percentages rise somewhat when we filter to see only the respondents who indicated that they are confident that their organizations are moving in the right direction to enable the most effective use of data and analytics for making decisions. Given that these are core technology capabilities of most self-service visual analytics and data discovery solutions, the results tell us that as use of self-service capabilities grows, organizations appear to gain greater confidence in their direction toward enabling effective use of data and analytics for decisions.

USER STORY HIGHMARK HEALTH BUILDS A STRONG DATA FOUNDATION FOR DIVERSE ANALYTICS

Highmark Health, headquartered in Pennsylvania, has grown in recent years to become one of the largest health and wellness organizations in the U.S. It functions as an insurer, with Highmark Health plans; as a provider, with the Allegheny Health Network; and as a services organization with HMHS. The company's mode of growth and its variety of businesses present tremendous challenges in integrating data for analytics, says Bob Gladden, VP of enterprise analytics at Highmark Health. "When we talk about becoming data-driven, we're referencing pulling data from these fairly disparate organizations together to drive decisions differently than we have in the past."

To address these challenges, Highmark is developing an enterprise analytics platform using solutions from SAS, Informatica, and open source Hadoop ecosystem technologies. At Highmark's organizations, "there's a significant amount data that's being created as opposed to just being captured," says Gladden. For example, reducing the rate of unplanned patient readmission after discharge is a key goal for all healthcare providers; the readmission rate is an important measure of the provider's quality of patient care. "To measure this, we have to determine whether a hospital stay is actually a readmission," Gladden explains. "This information does not show up in a claim. The only way you'd know is by going through a calculation process using algorithms to determine whether a readmission has happened. Our project, using SAS as the transformation language, is to take all of that derived— or, if you, will "smart" data— and incorporate it back into the overall enterprise analytics platform."

Gladden is a strong advocate of creating a platform designed specifically for analytics, not one focused on just operational, reporting-centric BI. "We need to organize the data to meet analytics needs, from users asking basic questions up to data scientists. This will seriously reduce the data preparation overhead that encumbers analytics. Trying to do both on the same platform tends to meet nobody's needs."

Although data science is important, Gladden recommends that organizations not overlook opportunities for less-sophisticated types of analytics. "Everyone is always looking at the fruit at the top of the tree— the very advanced analytics—but the lower-hanging branches are often ignored. What all types of analytics have in common, however, is the need for a strong data foundation. This is necessary to give decision makers confidence in the analytics."

Gladden says it is also imperative for organizations to evaluate whether they can properly maintain and update analytics models. "Folks can spin up a lot of interesting models, but there's no way you'd move them into production mode" without confidence in model management.

A quarter of respondents can perform data preparation steps on their own.

A significant portion of users engage in self-service data preparation. Data preparation is vital to BI and analytics, and as noted earlier, technology advances are enabling users to do more data preparation on their own with tools that are a significant improvement over the usual user self-service option, which is typically manual work involving personal spreadsheets. In our research, one-quarter (25%) of respondents said users could perform data preparation steps, including for data quality, consistency, and transformation, on their own. Just 19% can integrate or blend data on their own. These results indicate that we are still in the early stages of maturity with self-service data preparation. As noted earlier, TDWI research finds that most users continue to be reliant on IT for data preparation.

Many organizations actually prefer to have aspects of data preparation, integration, and metadata cataloging performed centrally rather than by individual users. Their concern is that if everyone is doing data preparation processes their own way, it will increase data chaos. Thus, as organizations evaluate self-service data preparation solutions, business and IT leaders will need to sort out which data preparation activities users can safely and productively do themselves and which ones a centralized IT function should perform.

A center of excellence or governance committee composed of business and IT stakeholders can be a good forum for discussing the direction of data preparation. Technology evaluation processes for data preparation need to take into account the development of solutions that incorporate machine learning and other AI techniques to automate data preparation, integration, and metadata management. These newer solutions can help both users and IT improve the speed of data preparation, including the preparation of voluminous and varied big data for BI and analytics.

Only small percentages of users are performing advanced analytics steps on their own. Our research finds that most users do not perform advanced analytics steps in a self-service fashion. This is not a surprising result because many of these steps require specialized expertise to perform. One in five respondents (20%) said users are able to perform some level of predictive analysis on their own without IT handholding. Somewhat fewer (15%) said their users could develop, test, and score analytics models without IT help. We discuss advanced analytics more in the next section.

Just 10% of respondents said users can access and query data on Hadoop, Spark, or other big data platforms in a self-service fashion. This last result speaks to the general lack of experience in enabling users to have direct access to big data from BI and visual analytics solutions; mostly data scientists and analysts with programming experience access and interact with big data. However, as technologies (including from the open source community) for SQL on Hadoop and fast, interactive access with big data mature, it will become increasingly possible for users to work directly with data in data lakes and hubs rather than having to move the data into a data mart or data warehouse.

In your organization, which of the following activities are users able to perform in a largely self-service fashion?

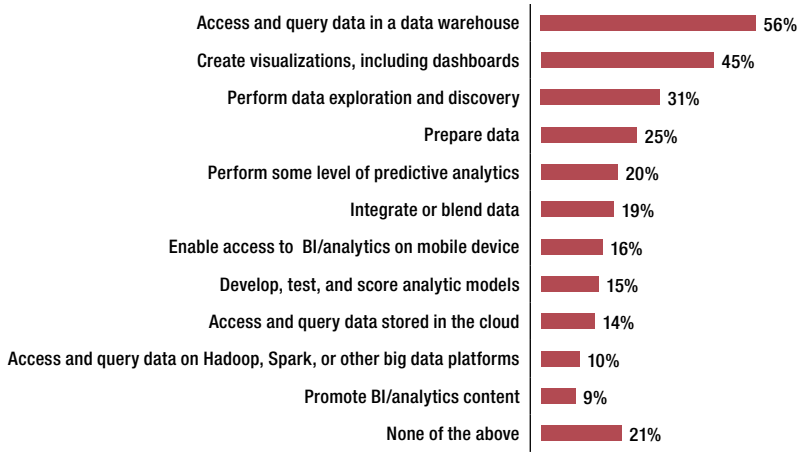


Figure 10. *Based on 198 respondents. Multiple responses allowed.*

Ease of use and automated model building

Vendors are trying to make their tools easier to use by automating the model building process.

Self-service visual analytics, mentioned above, helps to put analytics into the hands of everyone who needs it. Clearly, one aspect of democratization is making tools easier to use in terms of helping to support data-driven decisions. Ease of use does not stop at data preparation and visual analytics. With analytics talent in demand—especially in advanced analytics such as model building—vendors are trying to make their tools easier to use by automating advanced analytics. This includes automating techniques such as predictive modeling or providing natural language interfaces for asking questions.

The idea is to help business analysts and business users perform more sophisticated analysis—such as building models—without the help of statisticians, data scientists, or other highly quantitative staff. All the user needs to do is to specify an outcome variable of interest and the software does the rest. Some software even uses unsupervised approaches to surface interesting patterns. As mentioned, only a small percent are using these tools, but we expect that to grow as vendors continue to build out their offerings.

This move by vendors has given rise to the notion of the citizen data scientist. These are members of an organization who may not be formally trained in statistics or math but perform advanced analytics using automated advanced analytics software. Although this strategy is one that organizations are planning to use to deliver predictive models, the strategy can be both good and bad. On the one hand, if business analysts can expand their skill set that is a good thing. If a tool is easy to use, so much the better. Nonetheless, it is not good for someone to use highly advanced analytics tools without any training. There are many outlets for training, including online courses, user conferences, and industry conferences. For instance, TDWI offers data science boot camps that introduce business analysts and business users to the fundamentals of predictive modeling. In this way, the citizen data scientist can understand the techniques in order to be able to interpret and defend results.

The same holds true for a developer looking to embed a model into an application. With the popularity of open source and analytics marketplaces, it can be easy for a developer who does not know the theory behind predictive modeling or machine learning simply to select a model to embed in his or her application. The developer might test the model and it may seem to work. However, it is incumbent upon the developer to make sure that he or she understands the techniques that are going into these applications.

It is incumbent on the organization to put controls in place before any model goes into production.

Likewise, it is also incumbent on the organization to put controls in place before any model goes into production. For instance, some organizations use data scientists and statisticians working with business analysts to act as the control point. Other organizations have a comprehensive checklist of items they must complete before a model can go into production.

In our survey, we found no one group to be ultimately responsible for evaluating and deploying analytics models. The responsibility fell to IT (35%), the CDO/CAO/or chief data science officer (29%), or the business (27%, all not shown).

Embedding analytics/real-time analytics for action

Embedding analytics involves integrating actionable insights into systems and business processes used to make decisions.

Embedding analytics involves integrating actionable insights into systems and business processes used to make decisions. These systems might be automated or provide manual, actionable insights. It can be useful to think about a continuum for operationalizing and embedding analytics which goes from static (such as in a presentation or a dashboard) to interactive, to real time, to pervasive and automated. Some common examples of embedded analytics include embedding analytics in operational dashboards for frontline personnel, embedded analytics in websites, or embedding analytics into systems that either provide information for a manual action or take automated action.

We asked respondents what kinds of embedded analytics they use today. As illustrated in Figure 11, 43% of respondents embed analytics into dashboards, and 23% of respondents embed analytics into operational systems. These percentages rise to 60% and 30% respectively in the data-driven group (not shown). Not many of the respondents are embedding analytics into websites, mobile applications, or point applications, and less than one-quarter (23%, not shown) even have applications that provide alerts that they can act on. Most respondents look at the results of their analysis and then use this to take action (54%, not shown).

What kinds of embedded analytics do you make use of today? (Please select all that apply.)

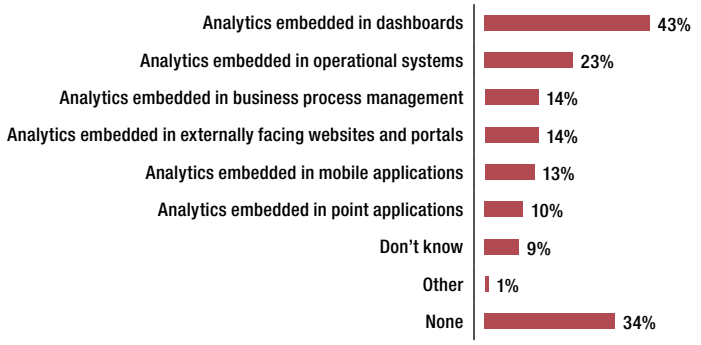


Figure 11. *Based on 173 respondents. Multiple responses allowed.*

Only about half of the respondents update their data daily or more often than daily.

Additionally, we asked respondents how often they collect data in their most frequently updated operationalized/embedded analytics application—be that a dashboard or a model put into production. About a quarter (29%) update their data once a day. Another quarter (23%) update their data either multiple times a day or more frequently. The rest do not operationalize any of their analytics, don't know, or update their data weekly or monthly (all not shown).

Although these respondents are early in this part of their data-driven analytics journey, as data becomes more voluminous and too real-time to take action manually, organizations will more often need to embed analytics into processes to take action (only 10% do this today, not shown). This can be an effective strategy to help organizations become more data- and analytics-driven.

For example, this might include preventive maintenance. A utility company might analyze data coming from many sensors that are measuring temperature, barometric pressure, wind speed, and precipitation. If a certain threshold is reached (say, the temperature gets too hot), an asset can be taken offline automatically. Other examples include recommendation engines and taking action on cart abandonment. Automation might include making numerous small decisions repeatedly. For instance, a telecommunications company might analyze call detail records in real time and execute personalized campaigns based on phone usage to proactively offer incremental increases to data plans. These are small tactical decisions that can be made over and over, and they can add up to high revenue.[7] As analytics becomes operationalized and embedded, it becomes more pervasive. As these examples illustrate, it may ultimately become more programmatic and more real-time. Therefore, although it is not necessarily the case today, organizations should start thinking about how they will staff and fund these kinds of projects in the future.

Organizational Strategies for Becoming Data-Driven

Technology and analytics strategies are critical for organizations to implement to become data-driven. For instance, as we saw above, it is important to provide a solid data infrastructure and access to data of high quality. Self-service is also important to help drive decisions because it empowers business users to ask and answer questions. Embedding and operationalizing analytics helps organizations move toward systematically or even programmatically driving decisions.

As stated earlier in this report, however, often organizational and people issues are the top reasons companies struggle to become data-driven. These include some of the barriers listed at the beginning of this report, such as leadership and governance strategies, as well as issues around changing people's perceptions, gaining trust, and motivating team members to use analytics for decisions.

Leadership strategies

Leadership is critical to any analytics effort. Leaders are responsible for ensuring things get done and they set the vision and tone for the organization. They can help evangelize data-driven concepts and put the funding and organizational structures in place to facilitate data-driven decisions. They are also role models. If leaders are data-driven, it can help others to become data-driven, too. These leaders include those in IT (e.g., the CIO) as well as leaders on the analytics front such as chief analytics officers and the VP of analytics. Of course, this also includes business leadership. Each position has a role to play.

[7] This is the field of decision management. For examples, see books by James Taylor, Neil Raden, and Tom Debvoise.

For instance, IT leadership should own the data strategy. That includes creating the enterprise data strategy, increasing data sharing, dealing with master data management, and owning data quality. Previous TDWI research indicates that in an analytically mature organization, IT owns data management.[8] The good news is that in many of the responses to this survey, IT *is* responsible for these functions. In fact, when we asked which executives or managers are ultimately responsible for these initiatives, IT was the leader for data-related initiatives.

IT/DevOps should also own deploying analytics models into production and should monitor these models once in production. Although IT might have ultimate responsibility, that doesn't mean that IT works alone. IT will need business input, as well as input from any other analytics organization, as part of this process. This should be a collaborative effort. In fact, about a third of respondents in this survey stated that IT and the business work together in the governance process (see Figure 12).

Business leadership is ultimately responsible for aligning projects with organizational objectives. Although collaboration between business and IT is essential, it is the business that needs to own this because, for the most part, they own the business decisions. In this survey, 50% of respondents do work in collaboration between business and IT to align projects with corporate strategies. Slightly less than half (42%) establish joint business and IT accountability for projects. These are steps in the right direction.

Half of respondents collaborate between business and IT.

Which of the following goals and initiatives are pursued jointly by business and IT leadership?

Figure 12. *Based on 202 responses. Multiple responses allowed.*

Of course, there are leaders in all parts of the company. They are not just C-level executives. As Tom Davenport, Jeanne Harris, and RobertMorrison point out in their book, *Analytics at Work*, "Any employee can move an organization in a more analytical direction." These people can be managers, directors, and VPs. They set goals, build out networks, and find the right leverage points.[9] They are knowledgeable, persistent, and trustworthy. Yes, they also look for funding. As Figure 13 illustrates, over 50% of managers are spending time making the case to corporate leadership to invest in BI and analytics. This can be time well spent because success in one department, such as marketing, can jumpstart the analytics success cycle mentioned at the beginning of this report. Respondents don't necessarily *like* doing this, but it can be quite important.

[8] See, for instance, the TDWI Analytics Maturity Model at tdwi.org/assessments

[9] For more information on the traits of analytics leaders, see Tom Davenport ,Jeanne Harris, and Robert Morrison, *Analytics at Work* (Harvard Business School Publishing Corporation, 2010).

USER STORY BUILDING A DATA CULTURE OF TRUST AT NATIONAL LIFE

At 168 years old, National Life is one of the oldest insurance companies in the U.S., offering a wide range of financial solutions, including life insurance, annuities, and investments. When the time came to adjust to a data-driven, fact-based, decision-making culture, consolidating multiple data sources and systems required thinking outside the box.

According to Noah Sassaman, a data architect at the company, until a few years ago the company was heavily spreadsheet centric. "We have an immense amount of data and everyone had their own 'data guy' who would get them the data they needed. We would trust that the information was accurate, but you never knew."

About three years ago, during an annual review, the company was tasked with providing information above and beyond what was previously asked for. It bogged down many of the company's systems he worked with. This experience solidified in Sassaman's mind that efficiencies were needed, especially on the data preparation side. He wrote a proposal and secured funding for a commercial data preparation product.

At the same time, the company was using a data visualization tool and having difficulty pulling data from multiple sources. Sassaman started working with the business analysts, and since then, "We've had one 'aha' moment after another in terms of solving significant issues the analysts were facing. There is something addictive about that. When you can work with analysts and help them come up with a solution for connecting data sources, it changes their outlook. It makes them open to whatever is next."

National Life has now built a data team to work with both business analysts and Data IT. The ultimate goal is for the analyst to enter questions into an interface and work themselves through qualifying questions to provide a data set.

Sassaman stresses that this is a methodical rollout within a complete change of culture. "We are making a change to let curiosity drive the analyst for fact-based decision making. We are sharing these 'aha' moments with people we interact with daily, making sure everyone can do this. A light goes off and they get it. Instead of hours to prepare data, it now takes minutes. Instead of fighting with the data, now you know it is correct. The process snowballs and grows. It is about empowering people, expanding the circle, and growing it out further—and then further."

According to Sassaman, in order to break down barriers it is important to take time to step back from both the analysts and Data IT and get to know people on a personal level. It is this foundational step that creates depth to the relationships so each party is more willing to give and take.

Cultural strategies

Only 30% of respondents stated that their BI and IT leadership jointly articulate and share successes.

We've seen that respondents state that they feel they are moving in the right direction to establish a culture that is data-driven, but many are still not there. When asked specific questions about various aspects of an organization's data-driven culture, most responses were neutral. For example, the majority of respondents answered "somewhat" to questions about aligning KPIs with business objectives or measuring the value of analytics based on impact to business outcomes (not shown). Only about 29% of respondents stated that their BI and IT leadership jointly articulate and share successes (see Figure 12). Less than half of the respondents have embedded analytics as part of a process that is ongoing (Figure 11).

As outlined in Figure 1, there are certain technology and organizational characteristics of data-driven companies. Organizationally, changing culture often comes down to dealing with

individuals and personalities in an organization. Some best practices associated with changing individual behavior include:

- **Incentives.** Only 7% of respondents use incentives such as bonuses or recognition for those who apply BI and analytics to decisions that deliver measurable output from efforts (Figure 13). Organizations need to focus more on incentivizing individuals to make use of analytics. This cuts across the organization, from the business manager who might be using self-service analytics to the call-center agent who may use the results of an analysis to suggest a next-best offer to a customer who calls. Incentive plans should be well-thought-out and attractive to specific roles.
- **Empowerment.** The organization must empower users to apply data and analytics to solving business problems. It is not enough to supply users with self-service tools so they can be independent of IT. Increasing users' freedom is vital, especially if their starting point is little or no independence from IT. However, users also need what IT can supply in terms of structure, governance, and facilitation of training and skills development.
- **Analytics literacy.** Developing skills and analytics literacy is critical for success. This is the case for those performing analysis as well as those using the results of analysis as part of an operational business process. With literacy comes trust, which is also a key for success. How can organizations expect a business user to perform an analysis if she doesn't know how to use a tool or think critically about data? How can a call center agent use the results of an analysis if he hasn't bought into the idea or doesn't understand what the results mean? Knowledge helps people to feel empowered. Yet, in this survey, respondents did not rate training as a top consideration for building an analytics culture. For example, only 26% of respondents encourage their staff to attend training. Only 20% plan to increase their training budget.

What steps are managers taking in your organization to develop a data-driven culture?

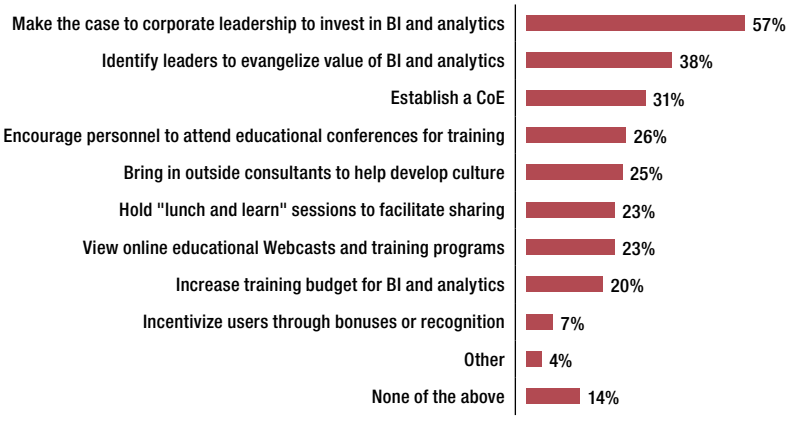


Figure 13. *Based on 230 respondents. Multiple responses allowed.*

Governance strategies for data-driven organizations

Governance rules and policies must set out how an organization protects sensitive data as they collect, report on, analyze, and share it. As organizations become more data-driven and democratize BI and analytics through self-service technologies, they need to ensure that they are protecting sensitive data. Sensitive data includes personally identifiable information (PII) and other data covered by regulations, but also data and information about finances, strategies,

Over a third (35%) of respondents stated that they do not have governance in place.

and intellectual property that the organization wants to protect. Organizations therefore need to align and integrate governance with data security and privacy procedures.

In some organizations, governance extends beyond data protection to include data stewardship, even if that term is not explicitly used. Stewardship aspects of data governance set out responsibility and accountability for the quality of data assets. It can include data analysts and data scientists mentoring business users in how to choose the right data sets and adopt procedures for maintaining data quality as they use and share data insights, including through dashboards and other visualizations. As analytics grows, some organizations are expanding governance and stewardship to cover analytics workflows, including the development and scoring of models. The intent is to use governance and stewardship to help establish standards for quality.

We asked research participants to indicate which data governance attributes best describe their organization's approach to data governance (Figure 14). The largest percentage (39%) indicated that in their organizations, governance includes data security and privacy rules, which, as we noted, is the primary responsibility of governance. About one-quarter (24%) said their governance sets users' role-specific access to data. Surprisingly, the second largest percentage (35%) said they do not have any data governance. TDWI recommends that organizations that do not have governance should make it a priority to set rules and policies that at least cover collection and use of sensitive data.

> **USER STORY GROWING GOVERNANCE IN THE FINANCIAL SECTOR**
>
> Financial regulatory data is critical for banks and other financial institutions, whether the data is for regulatory reporting or stress testing to ensure there's enough capital to deal with adverse developments. For business executives to make business decisions, they need to have a solid understanding of where they are. That means they have to have high-quality data.
>
> In forming a data governance group, a financial entity might first focus on regulatory data. A small data governance team may work best to get the project started. Meeting regularly with business units (perhaps weekly or monthly), the team can review new and outstanding data issues (such as exceptions or new or impending government regulations). Once the team establishes itself, it may be beneficial to have an executive steering committee focused on data issues that can break ties and provide overall project guidance. The steering committee could also develop a data certification process.
>
> Governance teams are responsible for establishing and enforcing policies and processes, including how the enterprise will manage data according to data quality metrics. The team must also define who "owns" the data. Those policies, once approved by the steering committee and management, can be circulated throughout the institution, with management emphasizing that the policies are required reading for every employee—front office and back office alike.
>
> Sharing the written policy is key to ensuring that all employees understand and follow the policy. Everyone has to buy into the process for it to succeed.

Nearly one-third (30%) of respondents have a centralized governance committee.

Nearly one-third (30%) said they have a centralized governance committee. Our research finds that IT usually heads up governance committees, although business stakeholders are involved. Setting up a committee that includes IT, data security, regulatory, and business stakeholders is a good way to start identifying key governance requirements and putting plans in place to enforce rules and policies. Regulatory compliance is a key focus for 28% of respondents. Regulatory concerns are spreading across many industries, particularly as more organizations interact directly over the Internet with customers and therefore manage potentially sensitive

information about them. Thus, this percentage is likely to rise in the future. Data ownership can often become a contentious issue in organizations; our research finds that 17% of respondents say their governance policies document data ownership.

One-quarter (25%) of respondents integrate data stewardship with data governance, and 17% said their governance covers analytics and data visualizations. These organizations are at the forefront of expanding notions of governance to data stewardship to improve not only oversight of sensitive data but also the quality of artifacts such as visualizations and analytics models.

Few organizations govern cloud-based data. Finally, our research finds that only 12% of respondents said their governance covers cloud-based data. Organizations using cloud platforms for data storage and management need to make sure they extend governance policies and rules to cover sensitive data in the cloud. Otherwise, cloud platforms can become a point of governance vulnerability.

Which of the following statements accurately describes attributes of your organization's data governance?

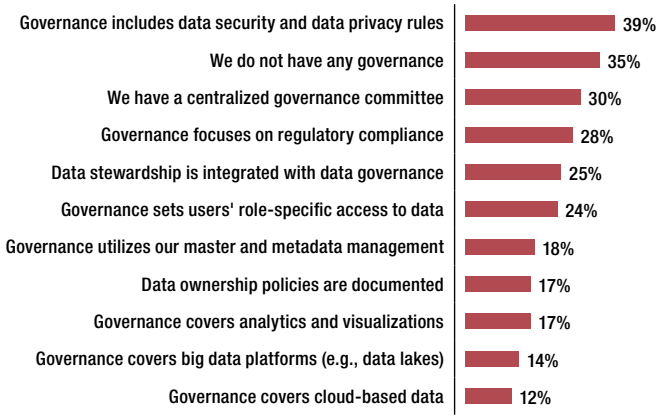


Figure 14. *Based on 181 respondents. Multiple responses allowed.*

Putting It All Together

Throughout this report we have focused on technology and organizational best practices for becoming more data-driven. We have explored the potential impact of self-service, application intelligence, and automation. We have discussed organizational imperatives such as leadership, culture, and governance. We conclude with the benefits that data-driven respondents have seen from their efforts.

Data-driven benefits should include tangible benefits such as increased productivity, improved operational efficiencies that drive cost savings, and top-line growth that delivers higher profits. We asked respondents about the benefits their organization has experienced from steps they have taken to become more data-driven. Figure 15 illustrates the responses for the two groups: the data-driven group and the group that believes it is not data-driven. The sample sizes are relatively small, but there are some significant differences in important outcomes.

First, those who are not data-driven were much more likely not to realize any significant value from any efforts to become data-driven. This group is just not that far along, even though they believe they are moving in the right direction. On the other hand, the group that believes it is data-driven was more likely to cite benefits such as more efficient and effective operations, higher customer satisfaction and loyalty, and higher top-line growth. There is a significant difference between the two groups in these variables.

What are the most important business benefits that your organization has obtained from taking steps to become data-driven?

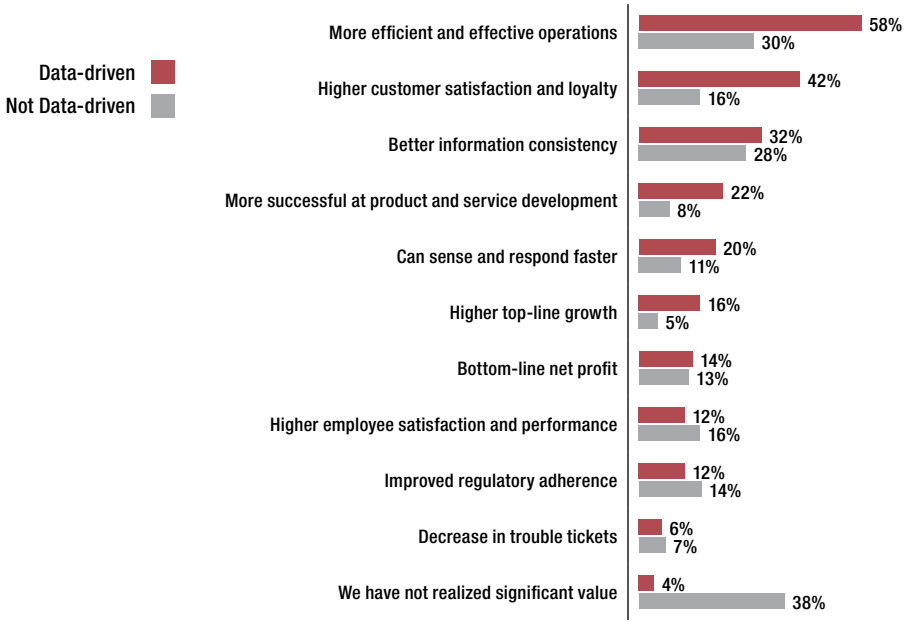


Figure 15. *Based on 50 data-driven and 76 not-data-driven companies. Three responses allowed.*

Vendor Solutions

The firms that sponsored this report are among the leaders in BI, analytics, and data management. To get a sense of the direction of the industry as a whole, this section looks at the portfolio of these vendors. (Note: The vendors and products mentioned here are representative; the list is not intended to be comprehensive.)

Hortonworks

Hortonworks provides open source-based data management platforms, services, and solutions. The company is a major contributor to Apache open source projects, including Apache Hadoop, NiFi, Spark, Ranger, and Atlas. The Hortonworks Data Platform (HDP) is powered by the company's Apache Hadoop distribution, as well as other open source technologies, and is designed to scale out for distributed clusters and manage data lakes. Hortonworks also provides an Apache NiFi-based stream-processing system called Hortonworks Data Flow (HDF). Although a separate product, HDF can run on top of HDP to collect, curate, and transport data from multiple services in real time for streaming analytics.

In September 2017, Hortonworks introduced its DataPlane Service (DPS), which extends the company's data management, security, and governance offerings across federated data fabrics to include systems such as enterprise data warehouses. DPS can be applied to both on-premises and cloud-based systems, and organizations can use it to spin up cloud or on-premises clusters. DPS uses a Web-based app store approach that allows organizations to use APIs to plug in data management and processing systems other than those from Hortonworks for management of data and workloads wherever they reside.

Looker

Looker offers a Web-based platform for enabling business users to go beyond standard reporting to find and explore data in a self-service fashion. It is frequently deployed in the cloud but can be implemented on premises. Working with a shared data model, the Looker platform addresses data integration, transformation, and governance as well as self-service data exploration and visualization.

A core component of Looker is LookML, an abstraction layer for user-friendly development of models and for producing enhanced SQL queries that have greater expression and better performance and that can be reused in part or as a whole. Looker can hide complexity from nontechnical users but does not aim to be a black box. Users can see the LookML and the SQL it produces to debug it, if necessary. Instead of requiring a separate, intermediate data store, Looker supports in-database processing, which enables transformations at query time and improves exploration agility. Looker speeds deployment by providing prebuilt Looker Blocks that sit on top of data systems. The Looker Blocks Directory provides a place where Looker and partners publish prewritten, reusable code that can be deployed immediately

SAP

SAP recently introduced its analytics framework for digital innovation called Leonardo. Leonardo is not a product; it is a framework for applying AI and machine learning to business projects, which includes analytics technologies as well as applications and services. The framework sits on a multicloud infrastructure and includes a foundational data management layer (SAP Hana, Data Hub, or Vora), the SAP cloud platform, Leonardo technologies (including machine learning and blockchain), and services based on design-thinking principles.

One such foundational layer for SAP Leonardo is the new SAP Data Hub. The Data Hub is a way to orchestrate algorithm execution on data in motion or at rest. The Data Hub provides a visual design environment to create data pipelines and an open data architecture that works across Hadoop data lakes, cloud object storage, relational databases, enterprise applications, and more.

SAS

SAS has focused on solving business problems using analytics since 1976. The company offers software for the entire analytics life cycle, from data access and preparation to production and management, within one application. Within the SAS platform are engines designed to address a wide diversity of analytics use cases and end-user needs, from visual analytics to machine learning and NLP.

SAS supports the process of embedding analytics for operational decision making. SAS can embed analytics in-memory, in-cloud, in-Hadoop, in-stream, in-device, and in-database. It also has a decision management product called SAS Decision Manager for automating repeatable operational decisions and event streaming for automating decisions in low-latency systems.

The company is investing heavily on productivity, accessibility and driving efficiency across the analytics life cycle. For instance, it is providing natural language interaction to its visual data mining, machine learning, and text analytics products in an upcoming release. The company also has plans to release a data discovery product, which will bring AI to data preparation and assessment, automating the ETL data profiling process.

Recommendations

In closing, we offer a list of 11 best practices for becoming data-driven along with comments about why each is important. Think of these best practices as recommendations that can guide your organization in its strategy to become data-driven.

Start small. Start small but think strategically. It may be that your organization starts to develop data-driven decisions using a small cloud analytics service. That is a good first step, but as you enjoy more success, you will need to think more strategically. That will include thinking through your platform, data management, and skills.

Apply analytics to performance management. Although many organizations have been using BI and data warehousing to support performance management, fewer apply analytics. Organizations should use analytics to improve the way they derive, define, and apply metrics. They could also use analytics to explore data associated with metrics to predict future results and be proactive in making adjustments to achieve better outcomes.

Bring big data platforms and data warehousing together into a single strategy. Our research finds that healthy percentages of organizations either are currently or are planning to develop BI and analytics on Apache Hadoop and Spark ecosystem technologies. As these systems grow in size and importance, organizations will need a strategy for how they fit with the existing data warehouse and BI systems. To avoid delays and disconnects, data-driven organizations need to create a comprehensive, multiplatform data architecture that provides a strategy for integrating all platforms.

Operationalize analytics to drive action. A hallmark of a data-driven organization is that it acts on the analytics it performs. These analytics are becoming part of a business process, whether that process is manual (such as checking a dashboard) or automated inside a system (such as a recommendation engine).

Understand when to automate. Some analytics for decision making don't necessarily require people be involved once the initial model building is done, such as a triage application for fraud. Organizations can use decision management applications in cases where they need to make small, repeatable decisions. Some of the vendors profiled in this report offer decision management applications.

Monitor the models. Using more advanced analytics to make and execute on data-driven decisions can be powerful. However, it is extremely important to make sure you monitor the models to ensure that old, stale models are not chugging along in production.

Think about the people. At the end of the day, your organization and the people in it will have to buy into the data-driven mindset. That means thinking through their motivations and incentivizing them to get on board. This may ultimately include many parts of the organization, from business users to those on the factory floor.

Training is important to build skills. Becoming data-driven is more than technology. Organizations need to help personnel develop skills such as how to use tools, engage in critical thinking, and use the output of embedded analytics. This requires training. That training can happen in-house or via a third party, but it is important. Remember, even if a tool is easy to use, that still means you need to use the tool well.

Self-service doesn't mean no governance. Self-service analytics can be empowering and a path to becoming data-driven, but if not handled properly, self-service can become the next iteration of the "spreadmart." That means putting in place solid data management and governance practices. It may also be important that organizations put controls in place if the analysis will be put into use in a business process.

Extend stewardship to analytics workflows and processes. As organizations devote more and more energy to analytics, stewardship needs to cover not just data management and BI but analytics as well. Organizations should identify stewards who can help users and analysts develop and deploy quality models and visualizations, including dashboards, so they meet an organizational standard of excellence.

Modernize data preparation for speed and agility. Data preparation processes are often slow, tedious, and consume too much time. For organizations to use data effectively to drive decisions, data preparation processes must take advantage of trends toward smarter and more automated technologies that use machine learning. Organizations should evaluate new data preparation technologies, including self-service data preparation to support self-service BI, visual analytics, and advanced analytics with big data.

Modernize the data platform. A multiplatform data architecture makes sense for more advanced analytics in support of data-driven decisions. Although the data warehouse is not going anywhere anytime soon, it makes sense to start thinking about what comes next in terms of extending the data management environment to support advanced analytics. MDAs might include the data warehouse, Hadoop, and other platforms both on premises and in a public cloud.

www.hortonworks.com

Hortonworks is a leading innovator in the industry, creating, distributing, and supporting enterprise-ready open data platforms and modern data applications. Our mission is to manage the world's data. We have a single-minded focus on driving innovation in open source communities such as Apache Hadoop, NiFi, and Spark. We, along with our 1600+ partners, provide the expertise, training, and services that allow our customers to unlock transformational value for their organizations across any line of business. Our connected data platforms powers modern data applications that deliver actionable intelligence from all data: data-in-motion and data-at-rest. We are Powering the Future of Data.

https://looker.com/

Looker is a complete data platform that offers data analytics and business insights to every department and easily integrates into applications to deliver data directly into the decision-making process. Data is no longer just a place to find answers. It's the place where ideas originate, and when everyone in your company is looking at the same numbers, sharing the same truth, they'll be able to collectively make smarter, more informed decisions. This is the core of our mission at Looker: to inspire everyone to embrace their curiosity, dig deeper, and keep asking questions.

The company is powering data-driven cultures at more than 1,200 industry-leading and innovative companies such as Sony, Amazon, The Economist, Spotify, Etsy, Lyft, and Kickstarter.

Looker is headquartered in Santa Cruz, California, with offices in San Francisco, New York, Chicago, Boulder, London, and Dublin, Ireland. Investors include CapitalG, Kleiner Perkins Caufield & Byers, Meritech Capital Partners, Redpoint Ventures, First Round Capital, Sapphire Ventures, and Goldman Sachs.

For more information, connect with us on LinkedIn, Twitter, Facebook, and YouTube or visit looker.com.

www.sap.com

As a market leader in enterprise application software, SAP is at the center of today's business and technology revolution. Our innovations enable approximately 300,000 customers worldwide to work together more efficiently and use business insight more effectively. SAP helps organizations of all sizes and industries overcome the complexities that plague our businesses, our jobs, and our lives. With Run Simple as our operating principle, SAP's nearly 77,000 employees focus on a singular purpose that inspires us every day: to help the world run better and improve people's lives.

sas.com/platform

SAS offers software for the entire analytics life cycle, including enterprise platform solutions for data management, discovery, and the deployment of analytics. As a global leader in advanced analytics, machine learning, and data mining, the company offers a comprehensive suite of software for structured, unstructured, and streaming data analysis for all users in an organization, ranging from the expert data scientist and business analyst to the end consumer of the analytical results. Combining the power of the SAS Platform with open source technologies enables you to unify disparate toolsets and analytics assets into a streamlined, collaborative environment that fosters productivity, business agility, and tangible outcomes. These results are easy to deploy and based on automated creation of all required assets for operationalizing and embedding analytics, including model governance and management, model documentation, and model monitoring. SAS supports the process of embedding advanced analytics for automating operational decision actions. Business and IT can collaborate in an environment to design, build, and execute decision flows that combine business rules and advanced analytics for the automation of operational business decisions. With scenario testing and direct integration with operational data, governed by integrated workflow, organizations can reduce deployment times in execution environments.

Learn more at sas.com/platform.

TDWI Research provides research and advice for data professionals worldwide. TDWI Research focuses exclusively on business intelligence, data warehousing, and analytics issues and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of the business and technical challenges surrounding the deployment and use of business intelligence, data warehousing, and analytics solutions. TDWI Research offers in-depth research reports, commentary, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.

tdwi

Transforming Data With Intelligence™

555 S. Renton Village Place, Ste. 700
Renton, WA 98057-3295

T 425.277.9126
F 425.687.2842
E info@tdwi.org

tdwi.org