



PULSE

Reducing Inefficiency and Increasing the Value of Analytics and Business Intelligence

By David Stodder

Sponsored by:





Reducing Inefficiency and Increasing the Value of Analytics and Business Intelligence

By David Stodder

Table of Contents

- The Pulse: Urgency for Improvement 3**
 - Spreadsheets: Ubiquitous and Challenging. 4
 - The State of Self-Service: What Users Can and Cannot Do 5
 - Time Loss and Inefficiency Due to Data Preparation. 7
- Selecting the Right Data Access Strategy 8**
 - Data Intelligence: Finding and Learning about Data 10
- Identifying and Managing Master Data 12**
- Governance: Both a Necessity and Value Driver 14**
 - Using Data Knowledge to Improve Governance 15
- Improving Sharing and Collaboration 15**
- Recommendations 18**
- Research Methodology and Demographics 19**

© 2018 by TDWI, a division of 1105 Media, Inc. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. Email requests or feedback to info@tdwi.org.

Product and company names mentioned herein may be trademarks and/or registered trademarks of their respective companies. Inclusion of a vendor, product, or service in TDWI research does not constitute an endorsement by TDWI or its management. Sponsorship of a publication should not be construed as an endorsement of the sponsor organization or validation of its claims.

This report is based on independent research and represents TDWI's findings; reader experience may differ. The information contained in this report was obtained from sources believed to be reliable at the time of publication. Features and specifications can and do change frequently; readers are encouraged to visit vendor websites for updated information. TDWI shall not be liable for any omissions or errors in the information in this report.

About the Author



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-driven decision making, increasing the value of BI and analytics, customer intelligence, visual analytics, BI/DW agility, mobile BI, big data, and information management. He has chaired TDWI conferences on BI agility and big data analytics. Stodder has provided thought leadership on BI, information management, and IT management for over two decades. He 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](https://twitter.com/dstodder) on Twitter, and on LinkedIn at [linkedin.com/in/davidstodder](https://www.linkedin.com/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 TDWI Pulse Reports

This series offers focused research and analysis of trending analytics, business intelligence, and data management issues facing organizations. The reports are designed to educate technical and business professionals and aid them in developing strategies for improvement. Research for the reports is conducted through surveys of professionals. To suggest a topic, 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 professionals who responded to our survey. Second, our report sponsor, 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

Datawatch sponsored the research and writing of this report.

The Pulse: Urgency for Improvement

In most industries today, it is essential for all personnel—from executives to managers to frontline sales and service staff—to use data effectively for both strategic decisions and operations.

Data is increasingly the lifeblood for a host of critical activities, including managing business performance, evaluating risk, engaging with customers, fine-tuning operational processes, meeting regulatory requirements, and innovating with products and services. Driven by business demand, self-service trends in analytics, data visualization, data preparation, and related information technologies are democratizing the use of data so users can do more on their own with less IT hand-holding and dependency.

Democratization of analytics and BI can quickly expose an organization's data management and governance shortcomings.



However, as democratization of analytics, reporting, and data visualization spreads, it can quickly expose an organization's shortcomings in data management and governance. At first, it might seem beneficial for users to blend diverse data sources on their own and develop creative visualizations—but only if they can trust the data enough to use their visualizations for decision making or to share analytics insights with colleagues and partners. If business users, data scientists, and analysts are spending too much time trying to find, prepare, and integrate data and not enough analyzing it, they are less likely to reap the benefits from investing in easier-to-use self-service technologies. Organizations' investments in BI and analytics will fall short of reaching their objectives.

Business and IT leaders need to come together to reduce inefficiencies and focus on how to improve data processes, governance, and technology implementations so their organizations can realize value from business intelligence (BI), analytics, and data management. This TDWI Pulse Report analyzes research findings about organizations' experiences and offers recommendations for addressing their challenges. We will examine areas such as data access strategy, master data management, and governance. This report will also discuss collaboration; given that few executives or managers make decisions alone, it is critical for technologies and methods to support joint development and sharing of analytics, visualizations, and other interactions with trusted data sources.

The report addresses concerns relevant to chief data officers (CDOs) and other business and IT leaders responsible for ensuring that their organizations can draw full value from data assets. Not all firms have a CDO, but the rising importance of data is spurring more organizations to either create this position or entrust "chief of data" responsibility to their chief information officers (CIOs), chief analytics officers (CAOs), or managers in charge of enterprise BI and data warehousing. Organizations are under pressure to improve the quality of data access and their knowledge base about data not just to satisfy human users but also to supply trusted data for automated artificial intelligence and machine learning algorithms and methods. Organizations need concerted leadership for effective governance, data stewardship, data quality improvement, and management to reduce inefficiency, redundancy, and errors.

Satisfaction with data and analytics projects is lacking. Overall, TDWI finds that in most organizations, there is significant room for improvement in data and analytics projects. When asked about their organizations' level of satisfaction with investment in data and analytics projects for meeting strategic goals such as enabling data-driven decision making or creating actionable customer intelligence, only 11% of research participants said they are "very satisfied," while 13% are not satisfied (figure not shown). The bulk of respondents were in the middle, but with the majority leaning toward satisfaction: nearly half (46%) are somewhat satisfied and one-quarter (25%) are somewhat unsatisfied.

Spreadsheets: Ubiquitous and Challenging

Just about every user has access to a spreadsheet application. Tools such as Microsoft Excel and Google Sheets are ubiquitous among users for viewing data, performing calculations, and creating simple visualizations. Users who do not have access to full-featured BI or visual analytics applications usually have little alternative but to work with readily available spreadsheets to see, analyze, prepare, and share data. However, with data volumes exploding beyond the normal capabilities of spreadsheets and interest rising in completing more sophisticated analysis and visualization, many users are frustrated with the limitations of spreadsheets and lack the specialized skills to customize spreadsheets and use advanced functionality.

Problems with data quality and consistency are the top concerns of spreadsheet use.



Spreadsheets' ubiquity will remain strong, so organizations need to make sure spreadsheet users are sourcing consistent, high-quality data. In Figure 1, we can see how research participants rank spreadsheet challenges that are most concerning to their organizations. At the top of the list is data quality and consistency, selected by 81% of participants. When spreadsheet users extract data from sources and manually prepare it—taking steps such as removing duplicate records, fixing errors in the data, and investigating anomalies and out-of-range values—they often do so in a personalized, inconsistent fashion. Errors can spread throughout visualizations and analysis.

Users' preparation processes are often not repeatable (62% cite this as a concern), with transformations aimed at specific, localized needs. Renaming data items typically increases data chaos; nearly two-thirds (63%) say that inconsistent documentation of data definitions and metadata is a concern. Many spreadsheet users do not follow common governance practices and rules for regulatory compliance, something that 50% of research participants say is an issue.

Many of the concerns cited in Figure 1 speak to spreadsheet users' frustrations in accomplishing objectives with reporting and analytics. Nearly half (47%) say that slow and error-prone consolidation for report generation is an issue; 45% cite limitations on data exploration and discovery. Dependence on spreadsheets and the data chaos they can spawn make it hard for organizations to be agile. Users have to reinvent laborious data preparation and transformation processes to fit new requirements. Almost half (43%) say that lack of flexibility for new business requirements is a concern.

Which of the following issues, if any, are most concerning to your organization when users are working with spreadsheets to visualize, analyze, prepare, and share data and analytics?

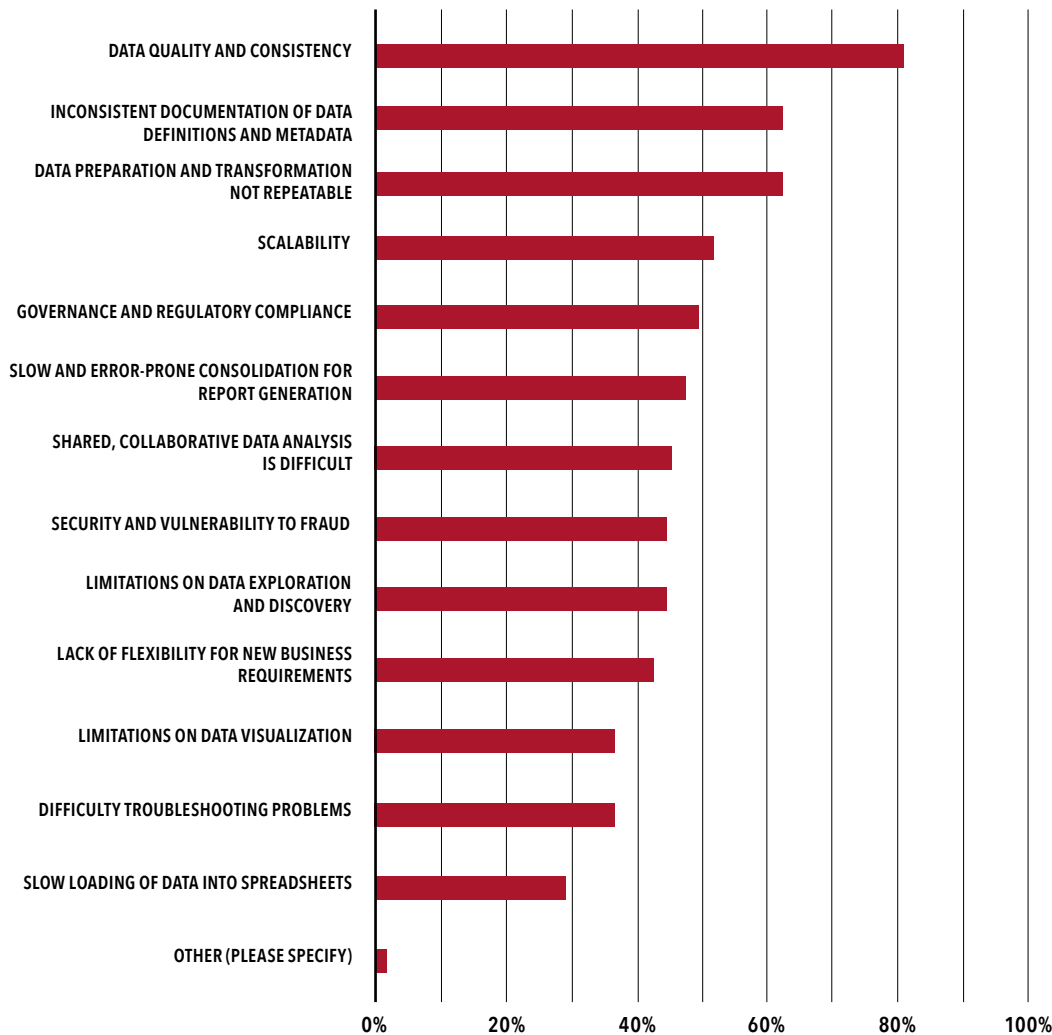


Figure 1. Based on 234 respondents. Multiple responses allowed.

The State of Self-Service: What Users Can and Cannot Do

As troublesome as spreadsheet problems are, users of all types of tools and applications can be equally frustrated if they are too dependent on IT developers and administrators for their data access, analysis, and visualization. Slow IT development and deployment of BI and analytics application functionality is an impediment to organizations determined to use data to drive more of their decisions and actions. Many users find traditional IT-centric enterprise BI systems to be too restrictive and hard to adjust to new requirements. An increasing number of users want functionality that goes beyond consumption of data via canned reports and dashboards. Thus, despite concerns about the information management issues that we will address in this report, it is not surprising that TDWI research continues to find strong interest in furthering the adoption of self-service BI, visual analytics, and data preparation.

Of course, what constitutes self-service varies. To improve information management effectively, organizations need to know what users can, cannot, and would like to do on their own. To discover which self-service attributes are most common, we asked research participants what users and analysts in their organizations can do without close IT support. Four out of five participants say that users and analysts can create reports in a self-service fashion (82%; figure not shown).

Users turn to self-service analytics and BI because they want functionality that goes beyond canned reports—and they cannot wait for IT to deliver it.



Although this result shows the popularity of user control over reporting, our case-study research often finds that not all users necessarily want or need the ability to develop reports from scratch. Primarily, users want self-service capabilities for interactive reporting, including the ability to select filters, personalize charts, and determine how they want to drill down into or slice-and-dice data to examine it more deeply. About two-thirds (65%) of research participants in this study say that users and analysts in their organizations can create dashboards and visualizations without close IT support, which illustrates the growth of easier-to-use functionality for tailoring data interaction. This suggests that organizations need information management to function behind the scenes to make it easier for users and analysts to personalize how they interact with the data.

Our research finds that in most organizations users and analysts can access and query the data warehouse on their own, but less than half (44%) say that they can find and access *relevant* data in a self-service fashion. This result suggests that the data warehouse does not suffice as the single source of relevant data and personnel may encounter difficulty when they try to work on their own to access a wider range of relevant sources. We can see this difficulty in the research finding that only 28% of participants say that their users and analysts can access and analyze new data, including external data, without close IT support. Even fewer (16%) are able to query “big data” sources such as Hadoop clusters and data lakes. (Keep in mind that although data lakes are viewed as an important trend, only about a quarter of organizations in our research typically have a data lake.)

Just one in five (20%) said personnel in their organizations can identify trusted data sources on their own, and only 18% can determine data lineage—that is, who created the data set and where it came from—without close IT support. If users and analysts are to work effectively with self-service BI and analytics, they need to be confident that they can locate trusted data and know its lineage. For self-service to prosper, IT and/or the CDO function must help users by stewarding their experiences and pointing them to trusted, well-governed sources for their analysis.

Time Loss and Inefficiency Due to Data Preparation

No matter whether they are using spreadsheets, self-service BI and visual analytics, or applications with embedded BI or analytics functionality, TDWI finds that most users and analysts are spending the balance of their time on finding and preparing data rather than extracting value through visualization, reporting, and analysis. Nearly half (48%) of research participants say that personnel in their organizations are spending at least 61% of their time on finding and preparing data (see Figure 2). Many are likely applying manual efforts to combine, cleanse, and transform their data using spreadsheet-based processes that are hard to repeat or share. Thinking of the difficulties mentioned above, respondents are also likely spending considerable time looking for relevant data, only to find that the data is flawed.

Organizations that take steps to clarify, standardize, and automate processes for data preparation can significantly increase the time personnel can devote to valued-added activities such as analytics. Improving data preparation can also help further goals for greater self-service BI and analytics by making it easier to find and work with trusted data.

Thinking of your organization's most recent data and analytics projects, what percentage of the total time was spent finding and preparing the data compared to the time spent extracting value through visualization, reporting, or analysis?

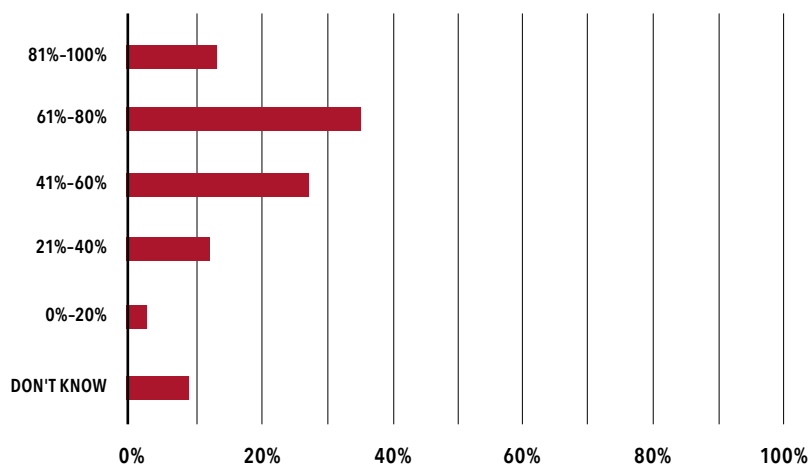


Figure 2. Based on 234 respondents.

Selecting the Right Data Access Strategy

Having good data access strategies is critical for users to get the most value from their data and for organizations to have efficient and effective information management. There is no one perfect way to access data for every circumstance. For some requirements, simple methods work; for requirements that are more analytically complex, demand integration of more data sources, or involve analysis of different types of data, users need more advanced solutions. With each type of data access strategy, organizations also need to consider the best way to balance users' voracious appetites for more data with the organization's governance rules, strategic priorities, and data platform performance considerations.

There is no one perfect way to access data; methods and technologies need to fit the requirements. However, with demand for more data and analytics, users need different ways of accessing data.



Which data access strategies are the most prevalent today, and which are in organizations' plans for the future? We asked participants about methods and technologies used in their firms to enable users to view, analyze, and interact with data (Figure 3). Not surprisingly, the most common model is for users to access data in desktop reports, files, spreadsheets, and/or databases (90%); about as common are users pulling data from sources into spreadsheets for analysis (89%). Most organizations also employ shared drives such as Microsoft SharePoint to access published data (78%) despite the weaknesses of such systems that were not built for data management, metadata management, and governance. Seventy percent say that users implement email to access and share data, a method that can also lead to data security and integrity problems if not well governed.

As noted earlier, spreadsheets, warts and all, are everywhere and will continue to be part of most environments along with desktop databases and file systems. Organizations should examine how they can contain the use of spreadsheet applications, desktop databases, and file systems to the purposes for which they are best suited and provide users with better solutions for their expanding analytics, visualization, and data preparation needs. These personal applications and systems are not going away, so an organization's overall data access strategy must ensure that users can access and extract well-governed, quality data to reduce errors and improve efficiency.

Many organizations plan to create central data catalogs. Regarding future plans—that is, what data access strategies organizations plan to adopt within the next three years—our research finds that the strategy with the largest planned use is a central data catalog or marketplace (44%). A data catalog typically contains metadata about database object definitions and where to find data sets. Thus, some regard a data catalog as synonymous with a metadata management system. Some of these systems include information about who produced the data set and other important characteristics that can explain data lineage.

A centralized version of a data catalog or metadata management system brings together trusted metadata from multiple systems into a single metadata management system. Such a shared resource can help users, administrators, auditors, and developers more easily locate quality data sources and know more about the data. Newer technologies enable organizations to automate development of a central data catalog or metadata repository and link metadata definitions to master data and business definitions, but the process still requires human guidance, especially to define metadata. Joint business and IT leadership expressed through a governance committee or a CDO can sustain support for this cross-functional resource.

Which of the following data access strategies does your organization currently implement to provide users with the ability to view, analyze, and interact with data, and which strategies do you plan to implement?

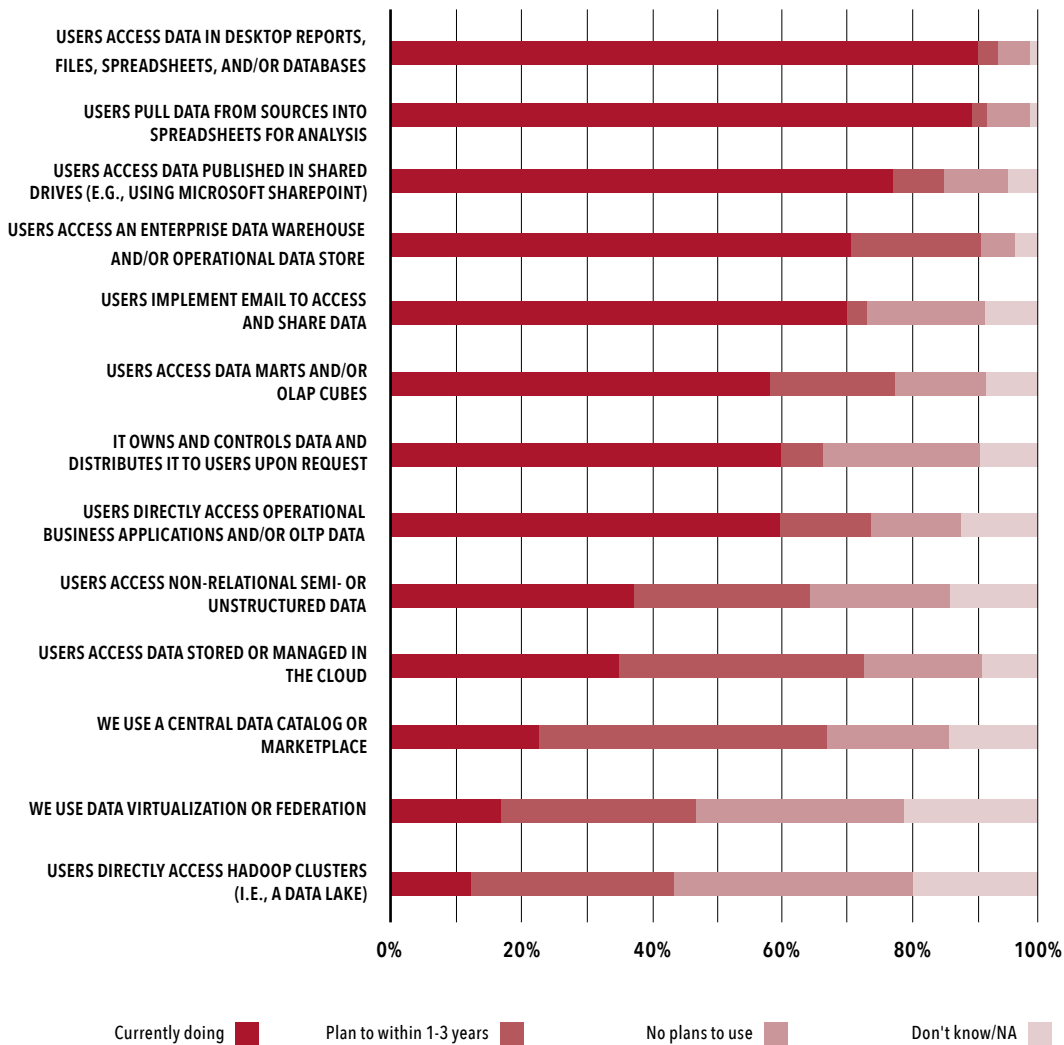


Figure 3. Based on 238 respondents. Ordered by current strategy use.

The research results show strong interest in accessing nonrelational semi- or unstructured data; 38% say personnel in their organizations currently do this and 27% plan to do so. However, a much smaller percentage say that users directly access Hadoop clusters and data lakes, where much big data is stored and managed; 13% currently enable this access and 31% plan to do so in one to three years. This suggests that among organizations participating in this research, most of those that use Hadoop clusters and data lakes provide users access to them through an intermediary platform, such as a data warehouse, BI system, file system, or analytics database. Organizations should ensure that their overall data access strategy addresses growing interest in analyzing the variety of data stored in big data systems (as well as the associated challenges).

With data intelligence, finding data includes discovery steps for evaluating sources; a data marketplace of curated sources can be helpful.



Data Intelligence: Finding and Learning about Data

Providing data access is important, but it must be part of a more complete data intelligence strategy that enables users, analysts, and data scientists to find data sources, discover if the sources are appropriate for their project, and tap into shared knowledge about the data. Finding data is about more than just establishing access to the data; it includes search, discovery, and analysis steps for evaluating data assets and whether the relationships between pieces of data merit further investigation.

One focus might be whether new data is appropriate for a particular business project and merits processes for ongoing analysis and reporting. Another might be to use semantic knowledge to discover patterns and trends in the data relationships across sources and suggest or predict other data that could be relevant. Data intelligence, therefore, focuses on understanding the value of the data and data relationships, often by understanding the semantic meaning of data and the significance of data relationships for a higher-level inquiry or objective.

We asked research participants what systems or service users and analysts in their organizations use to find and learn about data, analyze data relationships, and manage information assets. In Figure 4, we can see that after the ubiquitous spreadsheets, data warehouses (65%) and business intelligence applications (58%) were the most common (participants could select multiple answers). Additionally, nearly half (45%) are using a data mart or online analytical processing (OLAP) cube. These systems generally offer data that has been carefully selected, cleansed, profiled, and transformed by IT, and they can provide something close to a single view of the truth for certain types of requirements and from selected sources.

Quite a number of participants say that their organizations are using less formal ways of finding data and sharing knowledge about it, such as email (48%), word of mouth based on colleagues' suggestions (45%), and internal and external social networks (25% and 5% respectively). These methods generally depend on "tribal knowledge" that is passed from one employee to the next.

Which of the following systems or services do users and analysts in your organization employ to find and learn about the data, analyze data relationships, and manage information assets?

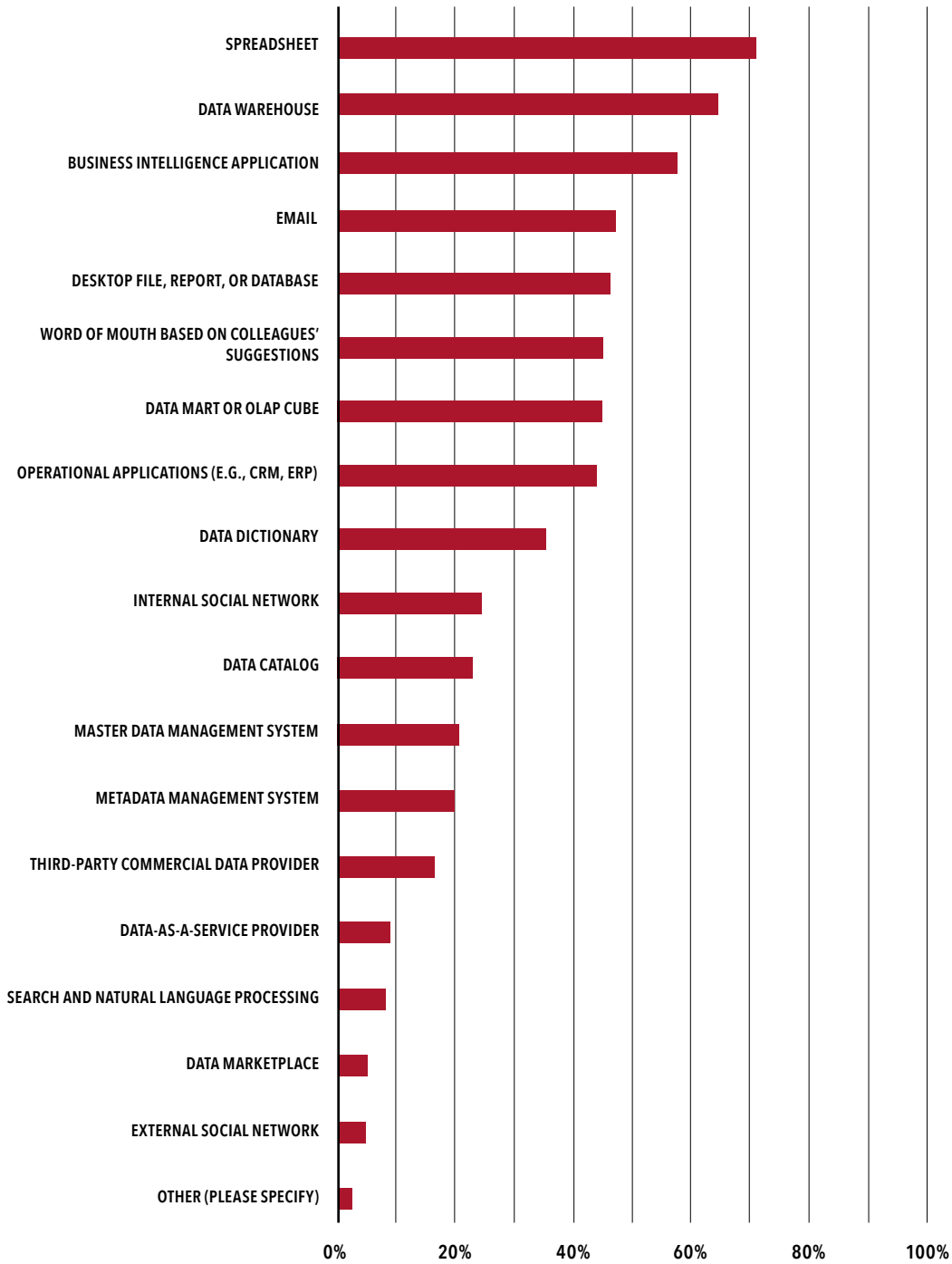


Figure 4. Based on 231 respondents. Multiple responses allowed.

Although tribal knowledge is valuable, organizations usually lack formal processes for capturing it. Typically, the knowledge is haphazardly documented, if at all. Depending on such methods of sharing knowledge about data can leave organizations vulnerable to inconsistencies as well as governance and regulatory problems.

Systems for formalizing and sharing knowledge about data include master data management (MDM) and metadata management, which about a fifth of research participants report that their organizations are using (21% and 20%, respectively). Beyond purpose-built MDM, other types of tools and services offer functionality for data cataloging, metadata and master data management, data preparation, and gathering information about data. We will look at this subject more closely in the next section.

Master data management can help organizations identify and manage “gold” reference data and understand data relationships across sources.



With the advent of cloud data services and the Web, a newer approach gaining traction is a data marketplace. This can make it easier for users to find data and make knowledge, annotation and comments, and reviews of data sources available to colleagues and partners. A data marketplace can offer trusted, carefully curated sources and/or other less well-understood sources. In our research, we find that only 5% of research participants say that their organizations are currently using a data marketplace. Some organizations are using data services such as from third-party commercial data providers (16%) and data-as-a-service providers (9%). Some of these provide tools for data discovery, metadata and master data integration, and data preparation.

Identifying and Managing Master Data

TDWI research finds that only a minority of organizations have deployed a dedicated MDM or metadata management system. MDM systems (or systems that have MDM functionality) enable organizations to identify and manage “gold” reference copies of data sourced from multiple systems—or, using a registry, enable organizations to do so in a federated fashion. A common use case is an organization that has significant customer data in multiple systems and uses MDM to establish a single, comprehensive view.

A focus on master data signals an organization’s commitment to an ongoing process of improving and sharing knowledge about data, including its quality and consistency for specific data-driven business projects. Organizations can use MDM to improve quality and reduce duplication at the source systems themselves. The processes and systems help organizations streamline discovery and documentation of how data in one system relates to data in other systems, particularly as they support business definitions of customers, patients, products, or other objects of interest.

Do you feel that an MDM platform would help resolve your organization's data access and governance concerns?

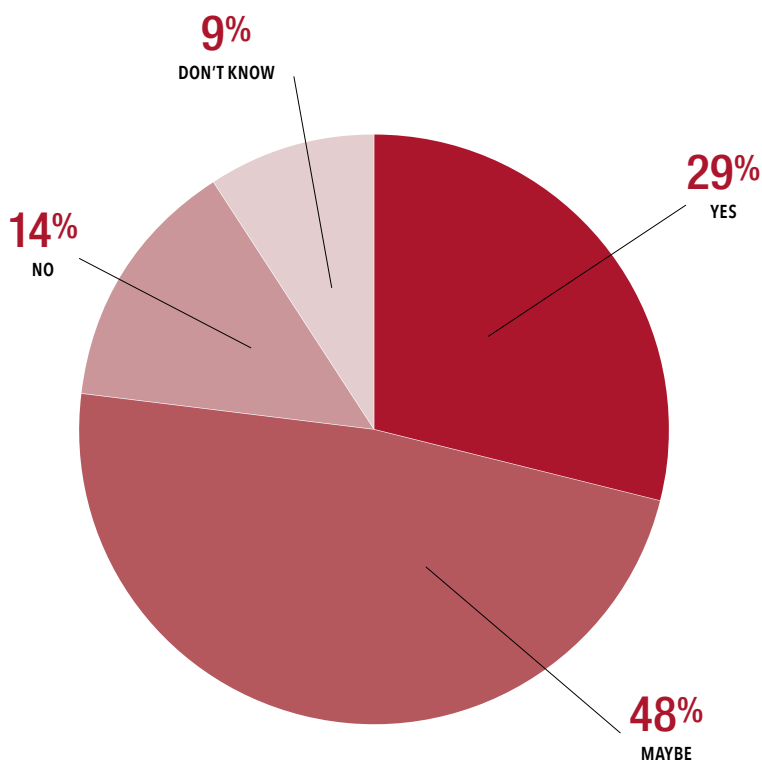


Figure 5. Based on 229 respondents.

Although we currently only see about one in five organizations having an MDM platform overall, if we focus only on large organizations with 10,000 or more employees, the percentage rises to one-third (33%; figure not shown). Among all participants, however, 31% say their organization plans to deploy one within one to three years. Still, nearly the same percentage say that they are interested but have no current plans to deploy an MDM system (29%) and 8% are not interested. Focusing again only on large organizations, the “no plans” and “not interested” percentages fall to 22% and 2%, respectively. Because they often have multiple sources containing related data, large organizations are generally the most interested in MDM and able to deploy systems and processes. However, the advent of cloud-based, as-a-service options makes functionality for defining and managing master data an option for organizations that ordinarily do not have the budget for on-premises systems.

Following up on the questions of MDM interest and deployment, we asked research participants if they felt that an MDM platform would help resolve their organization's concerns about data access and governance (Figure 5). Nearly a third (29%) said yes; the majority answered “maybe” (48%) and only 14% said no. Filtering the results to see only large organizations with 10,000 or more employees, the percentage saying “yes” rises to 34%, with “maybe” staying nearly the same at 47%; 19% said “no” (no figures shown). This shows the greater urgency among larger organizations to identify and manage master data.

Governance: Both a Necessity and Value Driver

Governance is important primarily for overseeing sensitive data such as customers' personally identifiable information (PII) and monitoring how it flows through and outside of the organization. Governance rules and policies should spell out how personnel and systems protect sensitive data and determine who has access to it. Governance is therefore critical to regulatory adherence, such as to the European Parliament and Council of the European Union's General Data Protection Regulation (GDPR).

However, governance can and should extend beyond data protection to include data stewardship. This responsibility focuses on sharing knowledge about data, mentoring users, and developing quality standards for the selection of data and its use for BI, analytics, and visualization. As analytics adoption grows, some organizations are expanding governance and stewardship to cover analytics workflows, including the development and scoring of models. These organizations are able to use governance and stewardship methods to improve quality, reduce redundancy, and improve the efficiency of data access and analytics processes as they scale up, all of which can increase the value generated by BI and analytics projects.

To learn more about the state of governance, we asked research participants to indicate how accurately each of a series of statements reflects how governance is implemented in their organizations. Most agreed with the statement that "governance rules cover data security and privacy"; 27% said it was very accurate and 41% said it was somewhat accurate (figure not shown). About the same percentages indicated that it was accurate to say that their governance focuses primarily on regulatory compliance.

Nearly half (48%) agreed that governance in their organization covers only corporate assets such as the enterprise data warehouse; just 23% said that governance covers Hadoop clusters and big data lakes (note, however, that 42% said they didn't know or that the question was not applicable, likely because they do not have Hadoop clusters or a data lake). As organizations expand their data architecture to such big data platforms, they need to reset governance rules and policies so they cover data systems that lie outside the enterprise data warehouse and other corporate assets if those systems are holding sensitive data. Stewardship practices, including mentoring by data analysts and scientists, can help business users determine the quality of data sources that are not directly managed by IT.

Half (50%) agreed with the statement that their governance "is 'tribal'; we have no formal governance." Organizations where this is the case need to take steps toward formal governance, first to secure oversight of sensitive data in BI and analytics projects and second to look at how governance could help improve stewardship and the creation of value from data assets. Fewer respondents (43%) said that their organization has a centralized governance committee, an institution that brings business and IT leadership together to set governance rules and policies.

Using Data Knowledge to Improve Governance

For better governance and stewardship, organizations should make use of the rich information developed about data sources through implementation of MDM, metadata management systems, and other solutions such as data preparation tools, policy engines, and semantic integration systems. Some modern data preparation solutions use AI, particularly machine learning, to gain knowledge about data sources and users' requirements so they can automatically offer recommendations as well as perform duties such as data cleansing, enrichment, normalization, and filtering as appropriate to the project. Governance policies and oversight will be more effective, accurate, and consistent if they tap these sources of knowledge and can apply master data definitions across sources to monitor sensitive information.

Only a minority of research participants say that data stewards guide users and analysts in selecting data sets. Taking on this responsibility would be helpful.



These knowledge-generating systems and solutions can, in particular, improve data lineage tracking: that is, identifying where the data came from, where it moves, and what happens to the data over time. Understanding data lineage is key to users' confidence in analytics conclusions and visualizations; among the first things executives typically want to know when examining analytics and visualizations are the origin of the data and what has been done to it on the way to reaching an analytics conclusion.

We asked research participants how confident users and analysts in their organizations were in their visibility into the lineage of data used in reports, visualizations, and analytics. Nearly two in five (38%) say that they are somewhat confident, but 33% say they are not too confident (figure not shown). A significant portion (18%) say they are not at all confident, and just 8% say they are very confident (4% do not know). Obviously, for most organizations, there is room for improvement.

Improving Sharing and Collaboration

Governance and stewardship can support efforts to improve collaboration and sharing on data, analytics, BI, and visualizations. Through governance and stewardship committees and communication, users and analysts can exchange insights about the quality and relevance of data sources for particular projects. Data stewards can help users avoid redundant data preparation and analysis efforts by monitoring whether data preparation routines, analytics models, or other artifacts already exist. Governance and stewardship managers can lead efforts to create common libraries of visualizations and analytics models.

For this report, TDWI examined the intersection of governance and the sharing of knowledge about data and models at research participants' organizations. Figure 6 shows how participants evaluated the accuracy of a series of statements describing governance and sharing. We can see that in the majority of participants' organizations, creators of analytics models know the data

Please indicate the accuracy of the following statements in terms of how well they reflect governance and the sharing of knowledge about data and models among personnel in your organization.

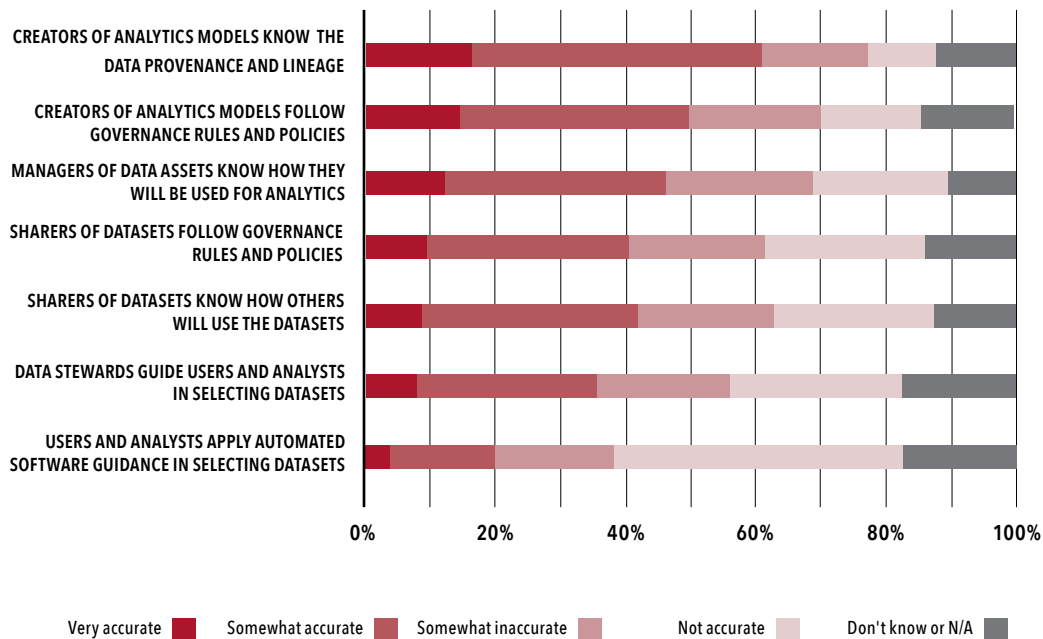


Figure 6. Based on 214 respondents.

provenance and lineage; 62% said this statement was either very or somewhat accurate. These results at first seem puzzling given what we learned earlier regarding confidence in the visibility into data lineage. However, note here that the statement focuses on the “creators of analytics models”—that is, data scientists and specialists in analytics model development. These personnel appear to be more confident than the broader group of business users and analysts. We also found that half (50%) of research participants say that the statement “creators of analytics models follow governance rules and policies” is accurate, which is positive news, although it also tells us that in half of organizations surveyed, this is not the case.

Some of the results in Figure 6 show where organizations could improve. Only 36% of research participants said the statement “data stewards guide users and analysts in selecting data sets” was either very or somewhat accurate. This is a key area where stewardship can be helpful. In addition, just 9% said it was very accurate to say that “sharers of data sets follow governance rules and policies”; 32% said the statement was somewhat accurate. Obviously, it would be wise for such organizations to inform sharers of data sets why it is important to follow governance rules and policies, particularly to safeguard sensitive data.

Which of the following actions are users and analysts in your organization able to accomplish through working collaboratively across individuals, teams, and cross-functional departments?

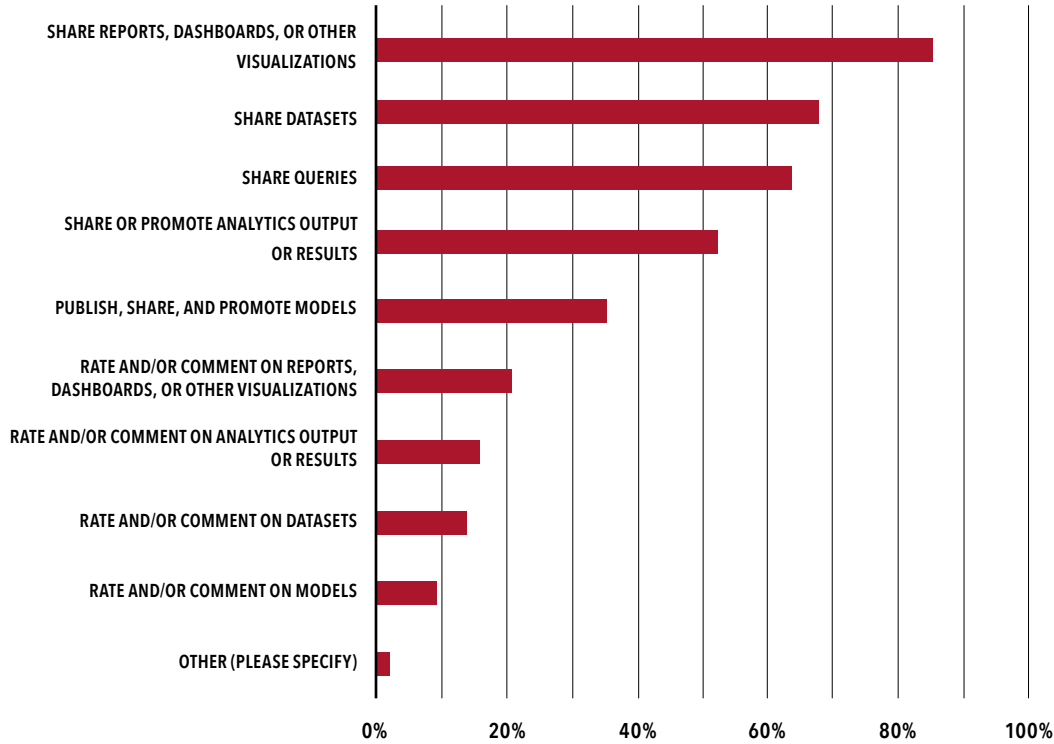


Figure 7. Based on 218 respondents. Multiple responses allowed.

Sharing visualizations is common; rating or commenting on them much less so. Figure 7 offers a snapshot of what actions the majority of users and analysts can accomplish through working collaboratively. We can see that most research participants (85%) report that sharing reports, dashboards, or other visualizations is common. More than two-thirds (68%) share data sets and nearly as many (64%) share queries. However, fewer say that users and analysts in their organizations rate and/or comment on what they share, which is a helpful capability for collaboration. Only 22% rate and/or comment on reports, dashboards, or other visualizations. Just 16% say that users and analysts in their organizations rate and/or comment on analytics output or results.

Improving how an organization’s personnel collaborate on data and analytics is important to realizing value from data and analytics because projects—and the decisions based on BI and analytics—typically involve teams. Such teams may include personnel across functions or departments, which magnifies the need for shared knowledge about data, such as master data definitions. Governance and stewardship should facilitate good communication to reduce confusion. TDWI often finds that delays in communication lead to errors that hold organizations back and require rework and duplication of effort. This can prevent organizations from realizing the full potential of data assets and analytics.

Recommendations

In closing, we offer the following recommendations for reducing inefficiency and realizing value from data and analytics.

Takes steps to improve data quality, consistency, and completeness. TDWI research finds that these issues are the leading obstacles to achieving success with data and analytics projects. Taking steps toward improvement will increase confidence that projects will deliver better data-driven decision making.

Evaluate alternatives to spreadsheets. TDWI research finds that many users and analysts rely on spreadsheets for their reporting, analysis, visualization, and data preparation. This is understandable because spreadsheets are so common, but their use can lead to inefficiency and governance problems. Determine where it makes sense to use alternative solutions.

Improve data access strategies so that they fit project needs. TDWI research finds that users and analysts often depend on spreadsheets, desktop databases, and primitive file systems for their data access. These solutions may be adequate for some projects, but not all (particularly those that need to scale up to more data and go deeper into analytics). Make sure that appropriate technologies and methods are available.

Create and maintain a central data catalog. Providing users, analysts, and data scientists with a shared, accurate resource of knowledge about the data will help them immeasurably by making it easier to find and access relevant data. Organizations can use the data catalog as a foundation to create a more comprehensive resource that catalogs usable reports, analytics (including models and formulas), visualizations, and other data artifacts to enable personnel to share these as well. Such a catalog will help organizations reduce errors and duplication of effort.

Apply knowledge about data to governance and stewardship. MDM, metadata management, data catalogs, data preparation, and other solutions and methods that gather and apply knowledge about data should play a critical role in increasing the efficiency and accuracy of governance and enabling smarter data stewardship. Make sure that committees responsible for governance and stewardship make full use of this knowledge.

Support teams by encouraging and supporting sharing and collaboration. TDWI research finds that organizations need to devote more attention to improving how teams collaborate on data analytics and share knowledge. We advise making team support an important part of governance and stewardship.

Research Methodology and Demographics

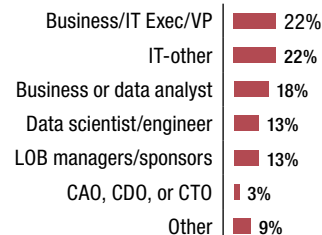
Report purpose. This TDWI Pulse Report focuses on information management inefficiencies that can hurt analytics and BI projects and result in user dissatisfaction. The report discusses data access strategies, master data management, governance, and stewardship as critical areas where improvements have potential for reducing inefficiencies and increasing value. The report offers recommendations for action.

Survey methodology. In February 2018, TDWI sent an invitation via email to the analytics, 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 263 respondents. All responses are valuable and so are included in this report’s data sample. Not every respondent answered every question, which is why the number of respondents varies per question.

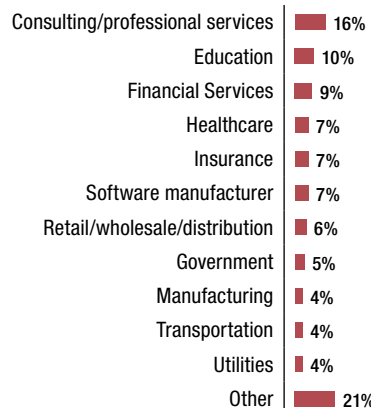
Survey demographics. Survey respondents came from both business and IT. The largest groups were business and IT executives and vice presidents (22%) and IT titles, including application managers (22%), followed by business and data analysts (18%), data scientists and engineers (13%), line of business managers (also 13%), and chief analytics, data, and technical officers (3%).

A wide range of industries is represented in this study, including consulting and professional services (16%), education (10%), financial services (9%), and healthcare, insurance, and software manufacturers (all 7%). Most survey respondents reside in the U.S. (68%) with other regions accounting for 32%. Respondents come from enterprises of all sizes.

Position

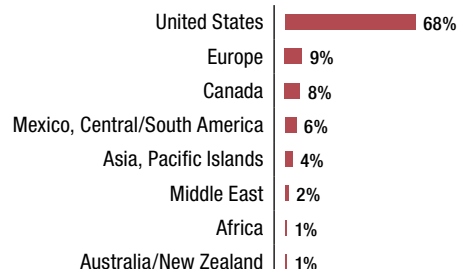


Industry

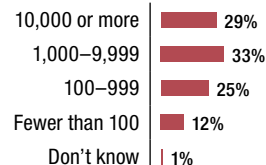


(“Other” consists of multiple industries, each represented by less than 4% of respondents.)

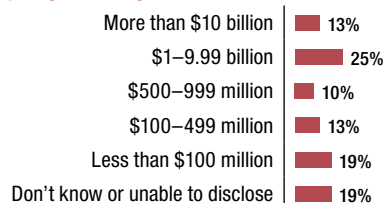
Geography



Number of Employees



Company Size by Revenue



Based on 263 respondents.

Sponsor



Datawatch Corporation is the data intelligence solutions provider that will fuel your business. Only Datawatch can confidently position individuals and organizations to master all data—no matter the origin, format, or narrative—resulting in faster time to insight. Datawatch solutions are architected to drive the use of more data, foster more trust, and incorporate more minds into analytics and reporting projects. With over 25 years in business, organizations of all sizes in more than 100 countries worldwide use Datawatch products, including 93 of the *Fortune* 100.

Datawatch is the maker of Monarch Swarm, a browser-based enterprise data preparation solution that empowers individuals and organizations to intelligently tap into data to drive faster insight and better value. It is designed to allow people of any skill level to unlock more data, regardless of the source or the structure, and make that data usable. Swarm provides more trust in data by providing clear lineage and evidence of integrity, as well as organizational governance controls. It also allows more minds to participate in the analytics process by powering cross-team sharing and collaboration in a centralized marketplace where users can publish their output to any analytics or reporting platform.



research

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.



**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