



2016

# Data & Analytics Survey

Exclusive Research from



### **Big Data and Its Use Cases Keep Growing**

Companies still consider data-driven initiatives potentially transformative, and younger employees say the transformation is coming soon.

As companies continue to collect an ever-expanding amount of data, they're actively seeking more ways to derive additional value from it. While enterprises with more than 1,000 employees are farther down the road than smaller organizations (SMBs) in implementing these data-driven projects, companies of all sizes are investigating new analytic tools to tackle an array of business goals. What's more, in what may indicate a generational shift in IT expectations, younger participants in an IDG Enterprise survey of 724 IT decision-makers who are involved in big data initiatives are dramatically more likely than their elders to feel big data will fundamentally alter the way business is done in the very near future.

Based on total respondents, more than half (53%) of their companies are currently implementing or planning to implement data-driven projects within the next 12 months — that is, projects specifically undertaken with the goal of generating greater value from existing data. Out of the projects that are underway or in the planning stages, 26% are already implemented, 14% are in the process of implementation or pilot testing and 13% are planning implementation in the next 12 months. Another 8% are considering a data-driven project, and 8% say they're likely to pursue one, although they're currently struggling to find the right strategy or solutions.

Among respondents who have no plans for a data-driven project, more than a third (38%) say they have no need to derive additional information or analytics from existing data. However, 22% cite budgetary constraints, and 21% say they don't have enough data.

Given their greater resources, enterprises are more likely than SMBs to have already launched or to be in the process of launching a data-driven project. SMBs are more likely to be struggling to identify the right strategy or solution for a data-driven project, or to have no plans to pursue one. The reasons why are contradictory, though. On the one hand, many SMBs report that they don't need any further information from the data they have. On the other, they're five times more likely than enterprises to say they lack sufficient data.

### **The Emerging Generation Gap**

Overall, those who are implementing data-driven projects, or plan to do so, feel big data offers enormous opportunities. Across all company sizes, 78% agree or strongly agree that the collection and analysis of big data has the potential to

fundamentally change the way their company does business in the next one to three years, and 71% agree or strongly agree that big data will create new revenue opportunities and/or lines of business for their company in the same timeframe. While only 58% agree or strongly agree that their company is positioned well to capitalize on big data's potential, two-thirds (65%) are optimistic that their organization's big data ecosystem will begin to take, or change, form in the next year.

Interestingly, though, these responses differ notably depending on the age of the respondent. Those aged 18-34 are vastly more likely than other age groups to strongly agree on both the transformative potential of big data and their companies' readiness to take advantage of it. Meanwhile, those over 55 years of age are significantly more likely than other age groups to disagree that big data will open up new revenue opportunities and/or lines of business in the near future. They're also far more unsure about whether and how their big data ecosystem will change in the next 12 months.

These age-linked differences may be attributable to younger employees being more comfortable with the latest technologies and more inured to the inevitability of technology-driven disruption. On the other hand, older respondents have seen many supposedly transformational technologies come and go throughout their careers. It's possible that they're simply less willing to predict that any particular trend — even one as far-reaching and multifaceted as big data analysis — will be a source of fundamental change.

### **Data Goals and Pain Points**

Respondents' top three business objectives in investing in data-driven initiatives are improving customer relationships (55%), making the business more data-focused (53%), and changing the way operations are organized (41%). Unsurprisingly, given budget realities, SMBs are significantly more likely than enterprises to cite reorganizing operations and optimizing the supply chain.

The top three challenges organizations hope to solve through data-driven initiatives are finding correlations across multiple disparate data sources (60%), predicting customer behavior (47%), and predicting product or service sales (42%). While identifying computer security risks, analyzing high-scale machine data and predicting fraud and financial risk were lower on the list of challenges overall, they were higher priorities for enterprises. And interestingly, respondents aged 18-34 were significantly more likely than others to mention the need to analyze social network comments for consumer sentiment — possibly reflecting how tightly participation in social media is woven into the lives of this age group.

Yet implementing these initiatives can itself be a challenge. In fact, 90% of respondents say they've experienced pain in one of these areas: data access and

analysis ranks as the number one headache instigator (38%), data transformation (17%), data creation and collection (13%), data migration (13%), and data storage (10%).

### **Where Data Comes From**

Respondents say that on average, 54% of the data their organization views and analyzes is generated internally, while 25% is generated externally, and 21% comes from a combination of the two.

The types of data collected vary by company size. Enterprises are more likely to collect transactional data, machine-generated/sensor data, government and public domain data, and data from security monitoring. Smaller organizations, by comparison, tend to collect email, data from third-party databases, social media, and statistics from news media. Regardless of company size, though, the top three sources of data are sales and financial transactions (56%), leads and sales contacts from customer databases (51%), and email and productivity applications (tied at 39%).

Due to its disorganized model and lack of a pre-defined database, managing unstructured data (emails, word documents, presentations, etc.) creates a challenge for organizations. Only 17% of organizations say that unstructured data is a primary focus of their data-driven initiatives, while nearly half (45%) call it one of their biggest challenges, and 31% say it is a problem that they have under control. On the other hand, another 17% say it's not on their short list of priorities, and 12% say it's not a problem.

### **Data and Dollars, by Size**

Given the volume and variety of data organizations deal with, and the expense of managing it, it's no wonder respondents have such high hopes — and a hard time achieving them. Almost a quarter of respondents (19%) are managing less than a terabyte of data, while only 7% are managing more than a petabyte. Although the average company manages 162.9TB of data, the average enterprise has 347.56TB of data, seven times as much data as the average SMB with 47.81TB.

Similarly, organizations of all sizes expect the amount of data to increase considerably in a relatively short timeframe. The average company expects its data volumes to grow to 247.1TB, an increase of nearly 52%, in the next 12 to 18 months. The average enterprise's data will grow slower than that, increasing by only 33% to 461.25TB. The average SMB organization expects to see its data volumes more than double, growing 138% to 113.82TB. However, don't put too much focus on this number; in 2015 small organizations said their data volume would grow to 110.6TB by 2016.

Under the circumstances, it's no wonder the average respondent plans to spend \$7.93 million on data-driven initiatives in the next 12 months. Enterprises plan to invest a significant amount more than SMBs, with an average expected spending of \$13.9 million versus \$4.3 million for SMBs. Despite this jump in spending, almost as many respondents expect their IT budget allocation for data-driven initiatives to remain the same in the next 12 to 18 months (35%) as expect it to rise (44%). This holds true across all company sizes.

Ninety percent of respondents expect to have to invest in the data center to keep up with data-driven initiatives in the next 12 to 18 months, primarily in storage capacity, data security, and cloud management. Similarly, 89% expect to invest in solutions to derive further business value from data, with data analytics, data/application integration, and data visualization at the top of the list. No one big data technology takes the lead, though: 46% plan to invest in some part of the Hadoop ecosystem in the next year, 41% plan to invest in a NoSQL solution, and 39% cite some other technology. Hadoop is the clear winner among enterprise organizations – 52% have plans to adopt some form of Hadoop ecosystem versus 42% SMB, while 44% of SMBs have plans to invest in NoSQL compared to 38% of enterprises.

### **Big Data, Big Security Concerns**

While raw data is valuable, the results of analysis are even more so — and enterprise organizations are more likely to appreciate the results of their analysis compared to SMBs. Nonetheless, only 39% of respondents overall say their company is securing the output of big data with alternate or additional security measures, while 44% say the measures they're taking to secure existing data are adequate to protect the results of analysis, too. The most popular approaches to securing big data output are identity and access controls (59%) and data encryption (52%), with data segregation (42%) a distant third.

Interestingly, conformance with industry standards for security is low on the overall list of criteria respondents cite as important when evaluating an offering from a data and analytics vendor — far behind ease of integration, ease of use, support and services, and scalability. However, enterprise organizations have this vendor criteria tied as the third most important factor. It's possible that some companies, especially SMBs, are trusting their vendors to handle security concerns. The average company's big data storage and analytics tools are just 18% cloud-based, but SMBs are nearly twice as likely as enterprises to have 50% or more of their big data storage and analytics in the cloud.

It's also possible that companies don't feel the data they're collecting requires further protection, since 68% either do not collect sensitive data (e.g. financial information, personally identifying information, or medical records) or restrict how much they collect. Among that group, half cite the need to meet compliance

mandates, while 38% have concerns about their ability to store this data securely, and 37% say they don't need to collect sensitive data. Enterprises are far more likely to limit data collection for compliance purposes, while SMBs are more than twice as likely as larger companies to say they don't need sensitive data in the first place.

Whether it be security concerns, overabundance, or inability to effectively analyze data, there are multiple reasons why organizations may seem stagnant when it comes to data and analytics progress. Nonetheless, big data and analytics terminology has become mainstream and is no longer in the “hype” stage. In order to stay ahead in this time of digital disruption and implement effective tech projects, organizations are and will continue to invest in their raw numbers and create actionable data.

### **Methodology**

IDG Enterprise conducted its 2015 Data & Analytics survey to gain a better understanding of organizations' big data initiatives, investments and strategies. A total of 1,060 respondents participated in the survey and results in this report are based off of 724 respondents who reported their organizations are currently implementing, planning or considering big data projects. Audience members of six IDG Enterprise brands (CIO, Computerworld, CSO, InfoWorld, ITworld and Network World) were asked to take the survey via pop - up and email invitations.

A broad range of industries are represented including high tech (18 percent), education (11 percent), financial services (11 percent), manufacturing (9 percent), government (9 percent), professional services (7 percent), healthcare (7 percent), and telecommunications and utilities (6 percent). Two out of five respondents (41 percent) work in organizations with 1,000 or more employees while 58 percent work in companies with fewer than 1,000 employees (one percent indicated they were not sure).

Over a third (47 percent) of respondents indicated they are a top IT executive at their company or business unit/location. Fifty-two percent of respondents hold executive IT management titles, 17 percent hold mid - level IT management titles, 14 percent hold IT professional titles, 12 percent hold business management titles and 5 percent hold other titles.

The margin of error for a sample size of 724 is +/- 3.6 percentage points. The margin of error for a sample size of 1,060 is +/- 3.0 percentage points. For the purposes of this report, "enterprise" refers to organizations with \$1 billion or more in annual revenues, "mid - market" refers to organizations with \$100 - 999 million in annual revenues and "small" refers to organizations with less than \$100 million in annual revenues. For questions requiring respondents to select a single answer, percentages may not sum to 100 percent due to rounding.