A HARVARD BUSINESS REVIEW ANALYTIC SERVICES REPORT

Harvard Business Review

DATA BLENDING: A POWERFUL METHOD FOR FASTER, EASIER DECISIONS

sponsored by alteryx

SPONSOR PERSPECTIVE

We're in the middle of a significant shift in the way organizations blend and analyze data, affecting not only the way decisions are made, but also how businesses operate. Data analysts in the line of business, who formerly had no choice but to rely on highly skilled, highly paid data scientists for data access, can now take advantage of self-service technologies to achieve analytic independence. These technologies enable practically anyone—regardless of their technical proficiency—to be a data analyst and make data-driven business decisions.

Data blending is the driving force behind this movement, freeing up people throughout the organization to experiment with data. Rather than putting business data, such as valuable insights about customer behavior, into the hands of a few, companies can place this data into the hands of the many—those, for instance, who will use this insight to bring about a better customer experience.

Alteryx sponsored the study "Data Blending: A Powerful Method for Faster, Easier Decisions" by Harvard Business Review Analytic Services, to support the empowerment of these analysts. Harvard Business Review Analytic Services conducted a comprehensive survey of business professionals across multiple industries and departments, and the results are clear. Data fuels the critical business decisions that an organization needs to make and having the ability to blend all of the various sources of data together to discover insight can, and should, happen anywhere in the organization.

At Alteryx, we are leading the enablement of self-service for data analysts through an intuitive workflow for data blending and advanced analytics that allows them to gain deeper business insights in hours, not the weeks of traditional approaches. Analysts around the world at companies of all sizes are using Alteryx to easily blend internal, third-party, and cloud-based data, and then analyze it using spatial and predictive drag-and-drop tools. This is all done in a single workflow, with no programming required.

Some corporate technology managers might see this analytics-for-everyone trend as a threat to governance and security, much like the way the bring-your-own-device (BYOD) movement was viewed. But that would be a mistake, as it actually represents a huge opportunity for all. IT departments will continue their critical role of ensuring that certain highly private data is indeed secure, but by learning from mistakes rather than shutting down the data pipeline, they empower their organizations to make better, more informed business decisions.

We are excited by the results of this study, as it validates our goal to make data blending and analytics accessible to everyone. By overcoming the struggles of legacy technologies like spreadsheets that were not built to handle today's data requirements, organizations can build a more stable, data-driven culture and create an environment where business analysts revolutionize the way they use data to make decisions.

Rick Schultz
Senior Vice President, Marketing
Alteryx, Inc.

DATA BLENDING: A POWERFUL METHOD FOR FASTER, EASIER DECISIONS

DATA ANALYTICS, the science of examining an organization's raw data with the purpose of gleaning valuable business information from it, is becoming a critical tool for today's business managers.

New research from Harvard Business Review Analytic Services shows tremendous bottom-line potential from using the latest generation of powerful data analytics tools. Our global survey of 502 business leaders and executives, academics, and IT managers in more than 15 industry sectors shows that a vast majority of executives find data analytics important for their decision making. figure 1

This importance is due to the growing pressure that managers face to find new ways of squeezing more revenue and profits from existing operations. For example, market analyst firm IDC tells clients they should be able to double the productivity of their business analysts by providing them with the right tools and decision-making culture. "There's a genuine movement toward decision making based on data rather than heuristics today," says Bob Parker, a group vice president at IDC.

That's certainly the case at home-improvement retailer The Home Depot. "We're not building new stores at nearly the rate we used to," says Charles Coleman, a senior analyst at the company. "So to grow the company, we've come to rely on data analytics instead."

The trouble is, to make smart and fast business decisions from data, business managers must overcome multiple challenges while using multiple analytics tools. Among the big challenges is the growing number of data sources analysts must gain access to. For another, many of the analytical tools on which analysts rely are incapable of processing, or struggle to process. many of the data sources. These tools, including Microsoft's Excel application, have analysts manually poring over spreadsheets and creating complex formulas from scratch each time out. All that means business analysts rely heavily on their organizations' data scientists and IT professionals. And that, in turn, can mean pulling a support ticket and waiting in line.

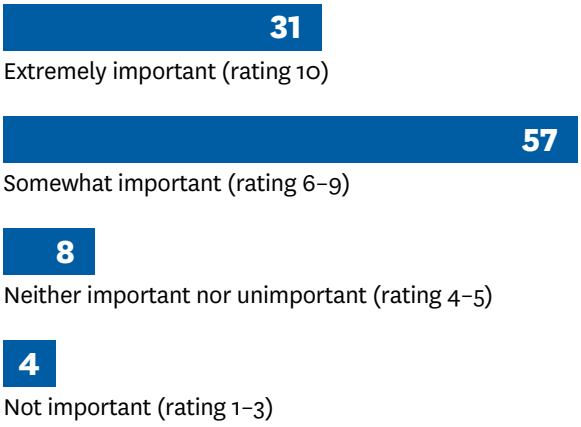
A new solution offers help. Known as "data blending," it's essentially a data analytics methodology and tool set combined. Data blending gives business analysts self-service solutions to cleanse, enrich, and join together the data they require. This can all be delivered in an automated process. Even better, the data can be taken from any number of disparate areas, both internally and externally. Data blending solutions also offer newfound speed and ease-of-use through graphical drag-and-drop interfaces that allow the user to see how data is transformed through every step in the process and give the user the ability to drill down on details. All this empowers managers to make informed and intelligent business decisions quickly and without the need to lean on a data scientist or IT professional.

FIGURE 1

DATA ANALYTICS MATTER

Percentage rating how important data analytics is to their organization's decisions and/or processes.

On a scale of 1 to 10 where 1 = not at all important and 10 = extremely important



RISING DATA SOURCES

The extent of the data source challenge is remarkable. For typical organizations, the number of data sources can range from a minimum of five to, in many cases, more than 15. Worse, nearly eight out of 10 survey respondents predict this number will increase within the next two years, with almost a third predicting a substantial increase. figure 2

The data source challenges are myriad. Data today can be found in both internal and external sources. Internal sources include everything from spreadsheets to data warehouses and data marts containing information from enterprise ERP and CRM applications. Many organizations also acquire data from third parties, including marketing service providers, trade associations, and governments. As for external data sources, these can include cloud-based enterprise applications such as Salesforce.com and Marketo, social media, and even data generated from mobile devices.

Some organizations go even further. At department-store chain Belk Inc., business analysts use geographical information—along with customer feedback data from online surveys, social media, call centers, and emails—to get a full customer view. "Analyzing that view helps us determine the customer's 'path to purchase,' " says Anu Brookins, Belk's VP of customer insights and analytics.

OLDER TOOLS CAN'T COPE

With all this data coming in from all these sources, old-school solutions are no longer sufficient. That's because these tools were designed for a far simpler data environment, one that featured fewer data types, with most data either raw or highly structured. They certainly weren't designed to handle unstructured data, such as social media and videos, or data from the cloud and mobile devices.

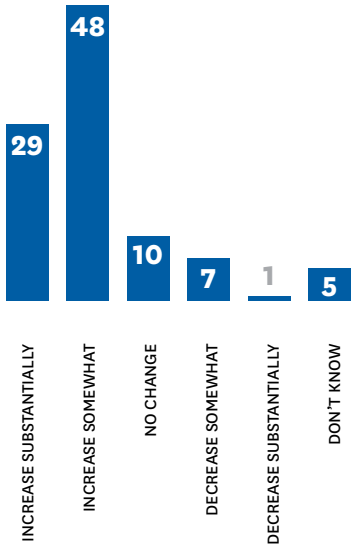
FIGURE 2

MANY DATA SOURCES NOW—AND MORE COMING

Percentage saying how many different data sources their organization currently accesses for most of its operational decision making.

Percentage rating how they expect the number of data sources to change over the next two years.

5 TO 9	33%
10 TO 14	13%
15 OR MORE	18%
DON'T KNOW	9%
FEWER THAN 5	27%



After Microsoft's Excel spreadsheet application, the most prevalent tools used for analysis include analytics tools developed in-house, third-party tools, and even some cloud-based ones. figure 3

While Excel is the preferred spreadsheet application for basic calculations and manipulating numeric data, its design dates back well before the current age of big data. Today, accessing large numbers of data sources, and cleansing, processing, and blending both internal and external data from multiple sources, are de facto requirements. Excel wasn't designed to do that. As a result, whenever data is joined from multiple sources, Excel requires tedious and laborious formulas and manual tweaks that leave the resulting spreadsheet highly susceptible to errors. "Excel and in-house development tools tend to hit the point of diminishing returns fairly quickly," says Parker of IDC.

Little surprise, then, that only 7 percent of respondents rate their current data-analysis tools as very or extremely effective. "The dirty little secret of big data," says Tom Davenport, a professor of IT and management at Babson College and the author or co-author of 17 books on those topics, "is that most data analysts spend the vast majority of their time cleaning and integrating data—not actually analyzing it."

In his other role as a senior advisor to Deloitte Analytics, Davenport sees many data analysts today manually poring over Excel spreadsheets all day, then generating the most basic cut-and-paste reports to the departments that need business insight. "Some companies with a global reach can have hundreds of data analysts," he says, "which equates to a huge waste of time."

Further, as these business analysts struggle with multiple data types and sources, and the old-school tools, they often have no choice but to lean on in-house data professionals. These IT professionals can help analysts both gain access to the data they need and squeeze more out of their Excel tools, but they often have to leverage much more sophisticated tools for combining data into a useful form for analysis.

More than 60 percent of respondents say their business managers are at least somewhat dependent on their IT managers, data scientists, data managers, and other dedicated analytics professionals. But fewer than 2 percent say they're extremely satisfied with those IT professionals' efforts. And a much larger group—nearly 40 percent—say they're only moderately satisfied.

What's more, depending on data pros is risky; as a group, they're becoming scarce. The McKinsey Global Institute predicts that the United States could face a shortage of nearly 200,000 skilled data pros by 2018. Adds Parker of IDC: "These kinds of data professionals can be hard to find, quantify, and keep."

FIGURE 3

SPREADSHEETS STILL DOMINATE

Percentage that say their organization uses the following tools for data analysis.

Select all that apply.

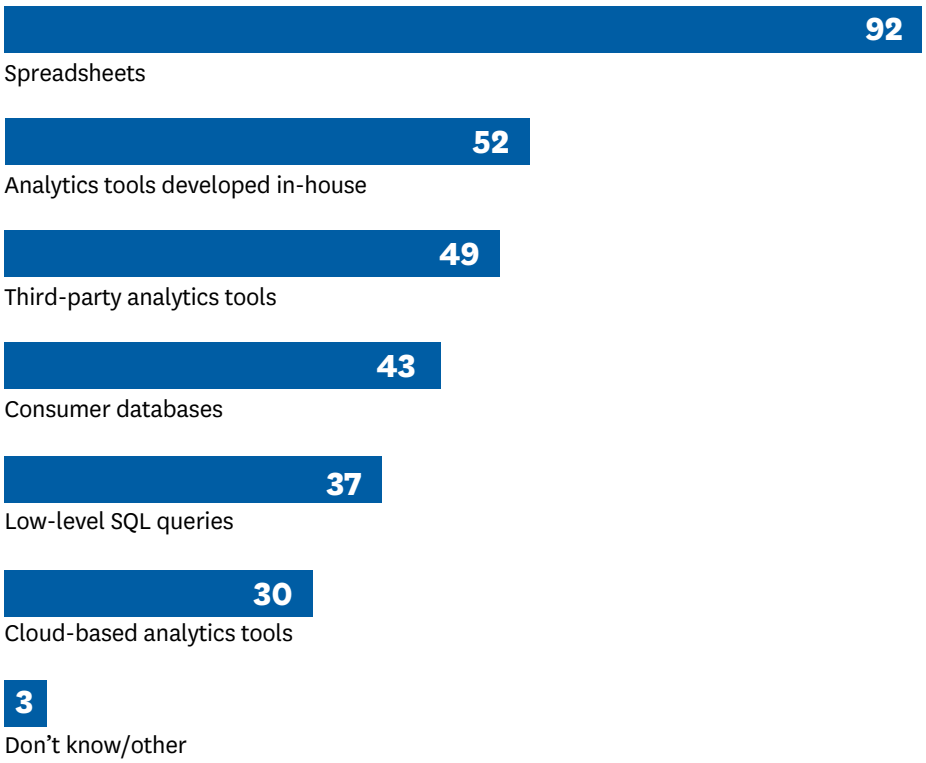
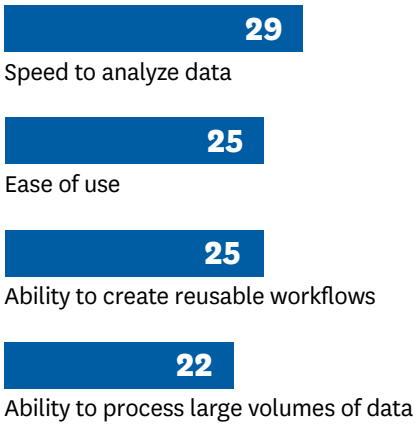


FIGURE 4

SPEED, EASE, VOLUME, REUSE

Percentage of managers who describe the features of tools used to analyze data for business decisions as extremely important.



A NEW SOLUTION: DATA BLENDING

Add it all together, and business analysts are asking for better tools to help them provide the kinds of information managers and executives need to make better decisions. Such a solution needs to be "self service"—that is, one where the analysts no longer need to rely on centralized data staff for help. Roughly a quarter of respondents rank key features of these tools as extremely important, including the speed at which they can analyze data; the volume of data that can be processed; the ease-of-use of the tools, such as having an intuitive graphical user interface; and the ability to drag and drop specific data preparation and data-joining tools to create workflows that can be used repeatedly without tedious manual work. figure 4

The good news is that the new methodology known as data blending is making such solutions possible. Already, business analysts across several industries are working with data blending tools that include powerful data preparation capabilities. These include data extraction, cleansing, blending, and enrichment—all right on the desktop. And with no need for costly help from IT professionals.

Data blending is the process of combining data from multiple sources to create an actionable analytic data set either for business decision making or for driving a specific business process. Data blending is needed when an organization's data-management processes and infrastructure are insufficient for bringing together specific datasets required by business groups. For example, data blending can readily bring together disparate data like customer information from a cloud sales automation system such as Salesforce.com, clickstream Web data stored in an open-source big data file system, and lists of data from Excel. Data blending differs from data integration and data warehousing in that its primary use is not to create a single unified version of the truth that's stored in data, but rather to build an analytic dataset that helps business analysts answer specific business questions.

Data blending differs from data integration and data warehousing in that its primary use is not to create a single unified version of the truth that's stored in data, but rather to build an analytic dataset that helps business analysts answer specific business questions.

Bank of America Merrill Lynch, one of the largest U.S. investment banks, has deployed a data blending solution to provide self-service analytical tools for its hundreds of business analysts. These analysts can now generate more in-depth and predictive client-profitability analyses. They've also gained true workflow automation, moving far beyond manual processes. And they've improved their regulatory reporting. All these gains have also reduced the firm's costs markedly.

Data blending has also found a home at The Home Depot. The company originally sought to improve business analysis of its stock-keeping units, or SKUs. That's the number retailers assign to a product to identify its price, product option, and manufacturer. At any given moment, The Home Depot uses approximately 160,000 SKUs in its more than 2,000 stores. Until recently, analyzing that information involved manually retrieving data from Excel spreadsheets. But now, thanks to its data blending solution, The Home Depot analysts can combine that mountain of data in an automated fashion. They can also do the work in just a fraction of the time it formerly took: a mere 30 to 60 minutes, down dramatically from the previous two weeks.

For The Home Depot, this greater speed is critical. The company can now more effectively price its merchandise according to factors that include competition and market elasticity. Senior analyst Coleman says that has not only saved thousands of hours of work, but also improved the company's bottom line by as much as 4 percent. "As we speed our rate of analysis," he says, "we can give ourselves time to work through other analytics projects and further develop data strategies."

Retailer Belk now uses data blending in every corporate group serving its 300 stores. Its merchandising group can now generate deep, predictive reports that precisely identify customer demographics and what customers are likely to buy. Belk's marketing group, meanwhile, uses data blending to determine which ad campaigns have been most—and least—effective. And the real estate group uses data blending to generate reports in record time to anticipate the financial impact when a store is remodeled. "Across the organization, we're now asking more and more sophisticated questions," says Belk VP Brookins. "This helps us understand how we can provide more value to our customers and increase the loyalty and retention of our customers—which ultimately leads to company growth."

It's not just financial services and retailers that are using data blending. Cardinal Health Inc., which provides medical products to more than three quarters of all U.S. hospitals, is another. Looking to empower its sales and marketing analysts, Cardinal Health deployed a data blending tool. The solution's built-in automation has dramatically sped up its work. Before data blending, Cardinal generated 100 customer reports after a month of manually compiling information. Now, with data blending, that has jumped to more than 2,000 reports, each packed with charts, graphs, and predictive business insights. And Cardinal can generate them all in under an hour.

How far might data blending and its associated tools go in the future? Advisory firm Ovum predicts that self-service tools will soon "enable a whole new universe of users." One big area could be genomics healthcare, where clinicians deploy such tools to analyze the generated data from a person's DNA to detect, diagnose, and treat various forms of cancer.

Another area ripe for data blending is the Internet of Things. This catchall category comprises the vast array of machines, sensors, and wearable devices generating reams of information, and it's one of the main sources of big data. As many as 25 billion "things" will be connected by 2020, predicts Gartner. Making that connectivity even smarter will be what data futurist Davenport calls the "analytics of things." That's where these data blending solutions would help make sense of the mountains of data these connected sensors amass.

Fortunately, companies don't have to wait for the benefits of data blending. They're here and now. "When I walk through our stores," says Coleman of The Home Depot, "I see the impact data blending has made. I can see product changes right on the shelf. And that's really exciting."

METHODOLOGY AND PARTICIPANT PROFILE

Harvard Business Review Analytic Services surveyed a total of 502 individuals. About one-quarter (26 percent) were from organizations with over 10,000 employees, another 26 percent were from organizations with from 1,000 to 10,000 employees, and 40 percent were from organizations with fewer than 500 employees. Respondents' own operating units were primarily in North America (45 percent), followed by Asia/Pacific (20 percent), Europe (19 percent), Latin America (9 percent), and the Middle East/Africa (6 percent). Participants represent a wide range of industries and functional areas, with 17 percent saying they work in general/executive management, 13 percent in sales/business development, and 9 percent in marketing/communications.

hbr.org/hbr-analytic-services

Harvard Business Review