Big Data Analytics
Assessing the Revolution in Big Data and Business Analytics

10 Best Practice Recommendations
Sponsored by Datameer
Ventana Research performed this research to determine attitudes toward and utilization of big data analytics. This document is based on our research and analysis of information provided by organizations that we deemed qualified to participate in this benchmark research.

This research was designed to investigate big data analytics practices and needs and potential benefits. It is not intended for use outside of this context and does not imply that organizations are guaranteed success by relying on these results to improve big data analytics. Moreover, gaining the most benefit from big data analytics requires an assessment of your organization’s unique needs to identify gaps and priorities for improvement.

The full report with detailed analysis is available for purchase. We can provide detailed insights on this benchmark research and advice on its relevance through the Ventana On-Demand research and advisory service. Assessment Services based on this benchmark research also are available.

We certify that Ventana Research wrote and edited this report independently, that the analysis contained herein is a faithful representation of our evaluation based on our experience with and knowledge of big data and analytics, and that the analysis and conclusions are entirely our own.
10 Best Practice Recommendations

This benchmark research reveals significant new insights into the evolving nature and use of big data analytics processes and systems. For organizations considering how to optimize their analytics and insights derived from them, we offer the following recommendations.

1. **Examine the experiences of early adopters of big data analytics.**

   The research shows that big data analytics is important or very important for four out of five organizations, but only about half (56%) are satisfied with their current efforts. We find satisfaction 50 percent more often among those that have had implementations for more than a year (66%) than those that began in the last 12 months (44%). Similarly, those using advanced tools are satisfied more often than those using legacy systems such as relational databases. Therefore, prepare to invest time in your own big data analytics deployment, and consider using tools that are designed to smooth the process and deliver specific results. Insist that vendors provide customer references that include best practices, ideally from companies in your type of industry or of your size and complexity, and ask them how they can accelerate and ease deployment. Learning from them can help your organization avoid pitfalls and shorten the time to value for its deployment.

2. **Define big data analytics clearly and establish communications among all stakeholders.**

   In nearly half of organizations participating in the research there are many different opinions about the meaning of big data analytics. Two of the three definitions most mentioned are that big data analytics involves analyzing data from all sources rather than just one (76%) and analyzing all of the data rather than just a sample of it (55%). We advise focusing your efforts, at least in beginning, on analyzing large amounts of data from many data sources. The research also shows that communication and knowledge sharing (ranked first by 24% and mentioned by 51%) is the most-often realized benefit of big data analytics, so from the beginning work to ensure that communication channels are in place, there is consensus on what big data analytics is and will do, and outcome expectations
are clearly defined. Doing so can help your organization achieve other benefits cited as important by research participants; topping that list are gaining competitive advantage (mentioned by 51%), making business processes more efficient (49%) and improving the customer experience and satisfaction (46%).

3. **Determine where you can profit most from big data analytics.**

Identify those business activities where application of big data analytics can produce important benefits. Three of the ones most often cited in the research are related to customers and revenue: enabling cross-selling and up-selling (36%), understanding the customer better (32%) and optimizing pricing (28%). Such a results-based focus may appeal to decision-makers considering an initiative. For example, recommendation engines have multiple uses including cross-selling and up-selling, which should be a priority for marketing. Improving forecasting (30%) and pricing should appeal to those in operations and finance. And optimizing operations through analysis of data such as log files is important for IT; one in four said they use big data analytics to optimize IT operations. Our analysis shows differing priorities based on organizations’ sizes and industries. Carefully consider your competitive environment and organizational needs and build a business case based on them.

4. **Address concerns that can hold back improvement and effectiveness.**

The research identifies a variety of organizational barriers to deploying big data analytics. The top reason most organizations do not plan to implement big data analytics or don’t know when they will do so is concern about security risks (39%); this need not be an issue, but it is wise to make it part of your evaluation to ensure you learn how to avoid it. Half of those that have already deployed big data analytics said difficulty in identifying return on investment (ROI) and defining the business case holds back making improvements in it; undertake advance research on both of these issues. The other most troublesome barriers to improvement, for two in five organizations, are lack of skills and data quality and information management. Thus you also should address people and information issues prior to implementing big data analytics. Determining the expected ROI and
defining a strong business case up front can prevent a loss of confidence once the initiative is under way.

5. **Exploit skills available now in the organization and develop others needed for big data analytics.**

A significant percentage (39%) said lack of skills impedes improvement in big data analytics, but the research also shows more satisfaction when advanced tools are used. Deploying standardized commercial software can reduce the need for in-house expertise in analytics. We recommend a two-pronged approach: Exploit the skills you already have (as in using SQL and spreadsheets) while adding advanced skills designed for big data and analytics. Organizations that use specialized roles such as data scientist or data miner reported improvement in activities from using big data analytics more often (88%) than those that are led by IT (73%) or lines of business (78%). On the other hand new tools can help SQL and spreadsheet users apply their skills to analytics, and for advanced analytics other tools remove the complexity of custom coding. In any case, be sure to include training in your implementation plans. In the research the most important training needs are for applying analytics to business problems (in 54% of organizations), the use of big data analytics tools (53%), analytics concepts and techniques (46%) and visualizing big data (41%). Therefore concentrate on developing skills in visualization, mathematics and statistics.

6. **Realize that big data analytics requires a variety of information sources.**

Half or more of organizations use an array of different types of data for their analytics: transactional data from applications, external data sources, content such as documents and Web pages, and event-centric data. Consider which mix of these each of your different user groups need to access and how to integrate them for analysis. Significantly fewer (40%) organizations are satisfied with the integration of information than are satisfied with the process overall (56%). Keep in mind that external data sources coming from cloud-based applications, social media data and third-party sources are increasingly used for big data analytics and often require different integration approaches than internal data types. Having high-quality data in compatible formats is a necessary precursor to analysis. Don’t
underestimate the importance of it, and find tools that can assist in the process.

7. Evaluate advanced tools such as Hadoop and in-memory systems for big data analytics.

Our analysis finds more significant improvement in organizational activities for those using in-memory systems (50%), Hadoop (42%) and real-time dashboards (39%) for big data analytics than other tools. Yet the tools most commonly used are not these but rather relational database management systems (RDBMSs) and flat files, which cannot exploit the power of big data. We strongly recommend evaluating tools designed for big data analytics. In-memory systems provide speed and flexibility and should be considered to enable functions such as real-time dashboards, visual analytics and data discovery, and analytic sandbox environments. Two-thirds of users of this technology said that real-time analytics is an important benefit of it. These functions can shorten analytic time to value, which many organizations consider an important metric for big data analytics. While in-memory technology is limited in the size of data it can handle, Hadoop provides a parallelized scalable architecture for big data. In addition, users can apply standard SQL approaches to implement analysis in Hadoop, which more than half do. Consider Hadoop for tasks such as ETL offloading, recommendation engines and fraud analytics, in which large-scale models must be frequently updated and scored.

8. Deploy analytic models that meet specific needs and update them regularly.

The research reveals that effective organizations manage their modeling processes by building and deploying them in a timely manner to meet specific needs. Most regularly update the analytic models: Seven in 10 do it weekly or more often, and half (54%) do it daily or more frequently. Those that update their models based on business need have more significant improvement in outcomes than those that do not (73% vs. 50%). Consider modern approaches to deploying big data analytics, including in-database approaches, that provide tight integration with the underlying architecture and enable timely deployments. The timing for model updates depends on changes in your business environment and the goals of the model, but we conclude from the research that organizations should update their
models at least once a month. Similarly, we recommend scoring models at least once a week (four out of five score more often than that) and considering modern methods for scoring such as PMML (which 28% currently use) and in-database approaches (62%). Also consider the advantages in time saved and accuracy of machine learning techniques such as closed-loop validation processes and continuous model updates over more manual modeling approaches.

9. **Utilize cross-functional teams that include a data science resource.**

The research shows that IT organizations (37%) are most often involved in designing and deploying big data analytics systems, followed by cross-functional teams (22%), data scientists or data miners (20%) and line-of-business analysts (15%). We believe that all these roles should participate in buying decisions for big data analytics, as should representatives from the departments that will deploy the analytics as well. More than three-fourths (78%) of participants said that collaboration is important to big data analytics, so it makes sense to institute it from the start. Remember also that communication and knowledge sharing is a highly ranked benefit of these initiatives. The research shows that when IT takes the lead, satisfaction is much less than when a specialized resource is involved. We suggest creating cross-functional teams, ideally including a data scientist, to manage the process; this is likely to produce the most significant impact on organizational activities. Furthermore, consider vendors that employ such specialized talent and can help lead the project from an analytical perspective rather than a technological one.

10. **Take a broad view of big data analytics deployment.**

We recommend considering multiple perspectives before deploying big data analytics. Include in your assessment the type of analytics to be deployed, the specific technologies to be included, the processes involved in customizing models for different purposes and the people and skills that must be involved to enable success. Consider a variety of technological approaches such as in-database analytics, in-memory, Hadoop, open source data-mining libraries and third-party analytics suppliers. These options can improve outcomes associated with big data analytics initiatives as well as supplement the expertise in your organization. An increasing number
(44%) of organizations plan to purchase a dedicated or packaged application when they deploy big data analytics. Fewer of them (36%) will choose custom builds than did those that have already implemented systems (54%). Consider evaluating packages that fit the deployment rather than creating a completely custom approach. Also understand how you want to deploy analytic models into the relevant lines of business. This may require integration with a variety of tools, including visualization, business intelligence, and business process and decision management.
About Ventana Research

Ventana Research is the most authoritative and respected benchmark business technology research and advisory services firm. We provide insight and expert guidance on mainstream and disruptive technologies through a unique set of research-based offerings including benchmark research and technology evaluation assessments, education workshops and our research and advisory services, Ventana On-Demand. Our unparalleled understanding of the role of technology in optimizing business processes and performance and our best practices guidance are rooted in our rigorous research-based benchmarking of people, processes, information and technology across business and IT functions in every industry. This benchmark research plus our market coverage and in-depth knowledge of hundreds of technology providers means we can deliver education and expertise to our clients to increase the value they derive from technology investments while reducing time, cost and risk.

Ventana Research provides the most comprehensive analyst and research coverage in the industry; business and IT professionals worldwide are members of our community and benefit from Ventana Research’s insights, as do highly regarded media and association partners around the globe. Our views and analyses are distributed daily through blogs and social media channels including Twitter, Facebook, LinkedIn and Google+.

To learn how Ventana Research advances the maturity of organizations’ use of information and technology through benchmark research, education and advisory services, visit www.ventanaresearch.com.
Appendix: About This Benchmark Research

Ventana Research designed this benchmark research for business and IT managers connected with managing or using big data systems and business analytics. The research was conducted from July through October 2013. Applying our standard methodology and quality assurance criteria, we identified 240 qualified participants. They represent a range of organization sizes: 32 percent from very large companies (having 10,000 or more employees), 28 percent from large companies (with 1,000 to 9,999 employees), 25 percent from midsize companies (with 100 to 999 employees), and 15 percent from small companies (with fewer than 100 employees). A large majority (81%) of these companies are located or headquartered in North America, although many of these are global organizations operating worldwide. Among industry categories, companies that provide services accounted for 43 percent, those in manufacturing for 27 percent and those in finance, insurance and real estate for 18 percent. Government, education and nonprofits accounted for the remaining 13 percent. Categorized by their job title, 18 percent are executives, 10 percent are in management, and the majority (63%) are what we term users in the lines of business. By functional area, 58 percent work in business units and 37 percent in IT. (More demographic detail about the participants is available in the full research report.)

These Best Practice Recommendations are drawn from the full Ventana Research Benchmark Research report. The full report is available for purchase, payable by check or credit card. Advice and focused guidance based on this benchmark research can be purchased through our Ventana On-Demand service. For more information about the full Benchmark Research report or assessment of your organization using our Maturity Index methodology, please contact us at sales@ventanaresearch.com.