Disclaimer:

This report should be used for informational purposes only. Vendor and product selections should be made based on multiple information sources, face-to-face meetings, customer reference checking, product demonstrations, and proof-of-concept applications.

The information contained in all Wisdom of Crowds® Market Study Reports reflects the opinions expressed in the online responses of individuals who chose to respond to our online questionnaire, and does not represent a scientific sampling of any kind. Dresner Advisory Services, LLC shall not be liable for the content of reports, study results, or for any damages incurred or alleged to be incurred by any of the companies included in the reports as a result of the content.

Reproduction and distribution of this publication in any form without prior written permission is forbidden.
Definitions

Business Intelligence Defined
Business intelligence (BI) is “knowledge gained through the access and analysis of business information.

Business intelligence tools and technologies include query and reporting, OLAP (online analytical processing), data mining and advanced analytics, end-user tools for ad hoc query and analysis, and dashboards for performance monitoring.”


Advanced and Predictive Analytics Defined
Advanced and Predictive Analytics Includes statistics, modeling, machine learning, and data mining to analyze facts to make predictions about future, or otherwise unknown, events.
Introduction
This year we celebrate the eighth anniversary of Dresner Advisory Services! Our thanks to all of you that have been with us along the way, encouraging and challenging us!

Since our founding in 2007, we have strived to offer a fresh, real-world and alternative perspective on the Business Intelligence (BI) market. We hope that you agree that we not only have succeeded in doing so but also continue to “raise the bar”—offering increasingly compelling research and greater value with each successive year!

Since we published our first Wisdom of Crowds® Business Intelligence Market Study in 2010, we have continued to expand our research offerings to include a variety of important topics including: Location Intelligence, Advanced and Predictive Analytics, Cloud Computing and BI, Collaborative Computing and BI, Embedded BI, BI Emerging Technologies, and Small & Mid-Sized Enterprise BI. During 2015 we have added to these topics with coverage for Enterprise Planning, End-User Data Preparation, and Big Data Analytics.

This publication marks our second annual Advanced and Predictive Analytics Market Study report. In it we continue to expand coverage of this important and developing topic—examining changing user perceptions and plans and current and future industry capabilities.

In closing, we’re very excited about both the market and our ability to continue to add substantial perspective and value to it!

Thanks for your support!

Best,

Howard Dresner
Chief Research Officer
Dresner Advisory Services
### Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definitions</td>
<td>3</td>
</tr>
<tr>
<td>Business Intelligence Defined</td>
<td>3</td>
</tr>
<tr>
<td>Advanced and Predictive Analytics Defined</td>
<td>3</td>
</tr>
<tr>
<td>Introduction</td>
<td>4</td>
</tr>
<tr>
<td>Benefits of the Study</td>
<td>8</td>
</tr>
<tr>
<td>Consumer Guide</td>
<td>8</td>
</tr>
<tr>
<td>Supplier Tool</td>
<td>8</td>
</tr>
<tr>
<td>External Awareness</td>
<td>8</td>
</tr>
<tr>
<td>Internal Planning</td>
<td>8</td>
</tr>
<tr>
<td>About Howard Dresner and Dresner Advisory Services</td>
<td>9</td>
</tr>
<tr>
<td>About Neil Raden</td>
<td>10</td>
</tr>
<tr>
<td>Survey Method and Data Collection</td>
<td>11</td>
</tr>
<tr>
<td>Data Collection</td>
<td>11</td>
</tr>
<tr>
<td>Data Quality</td>
<td>11</td>
</tr>
<tr>
<td>Executive Summary</td>
<td>13</td>
</tr>
<tr>
<td>Study Demographics</td>
<td>13</td>
</tr>
<tr>
<td>Geography</td>
<td>14</td>
</tr>
<tr>
<td>Functions</td>
<td>15</td>
</tr>
<tr>
<td>Vertical Industries</td>
<td>16</td>
</tr>
<tr>
<td>Organization Size</td>
<td>17</td>
</tr>
<tr>
<td>Analysis: Business Intelligence Users</td>
<td>19</td>
</tr>
<tr>
<td>Importance of Business Intelligence</td>
<td>19</td>
</tr>
<tr>
<td>Analysis: Advanced &amp; Predictive Analytics</td>
<td>22</td>
</tr>
<tr>
<td>Types of Users for Advanced and Predictive Analytics</td>
<td>22</td>
</tr>
<tr>
<td>2014 – 2015 User Comparison</td>
<td>27</td>
</tr>
<tr>
<td>Importance of Advanced and Predictive Analytics</td>
<td>28</td>
</tr>
<tr>
<td>Importance by Region</td>
<td>29</td>
</tr>
<tr>
<td>Importance by Function</td>
<td>30</td>
</tr>
<tr>
<td>Importance by Industry</td>
<td>31</td>
</tr>
<tr>
<td>Importance by Organization Size</td>
<td>32</td>
</tr>
</tbody>
</table>

http://www.dresneradvisory.com   Copyright 2015 – Dresner Advisory Services, LLC
## 2014 – 2015 APA Importance Comparison

Adoption Plans for Advanced and Predictive Analytics

Deployment Plans for Advanced and Predictive Analytics

Adoption Plans by Industry

Adoption Plans by Organization Size

Adoption Plans by Function

Adoption Plans by Geography

Feature Requirements of Advanced and Predictive Analytics

Feature Requirements by Geography

Feature Requirements by Industry

Feature Requirements by Function

Feature Requirements by Organization Size

2014 to 2015 Analytic Features Comparison

Data Preparation Plans for Advanced and Predictive Analytics

Data Preparation by Industry

Data Preparation by Organization Size

Data Preparation by Function

Data Preparation by Geography

2014 to 2015 Data Preparation Comparison

Usability for Advanced and Predictive Analytics

Usability by Organization Size

Usability by Function

2014 to 2015 Usability Comparison

Scalability of Advanced and Predictive Analytics

Scalability by Region

Scalability by Industry

Scalability by Function

Scalability by Organization Size

2014 to 2015 Scalability Comparison

Industry and Vendor Analysis

Industry Importance of Advanced and Predictive Analytics

http://www.dresneradvisory.com

Copyright 2015 – Dresner Advisory Services, LLC
Benefits of the Study
The DAS Advanced and Predictive Analytics Market Study provides a wealth of information and analysis—offering value to both consumers and producers of Business Intelligence technology and services.

Consumer Guide
As an objective source of industry research, consumers use the DAS Advanced and Predictive Analytics Market Study to understand how their peers leverage and invest in business intelligence and related technologies.

Using relevant criteria to evaluate vendors and products, users glean key insights into software supplier performance, enabling:

- Comparisons of current vendor performance to industry norms
- Identification and selection of new vendors

Supplier Tool
Vendor Licensees use the DAS Advanced and Predictive Analytics Market Study in several important ways. For example:

External Awareness
- Build awareness for the business intelligence market and supplier brand, citing DAS Advanced and Predictive Analytics Market Study trends and vendor performance
- Create lead and demand-generation for supplier offerings through association with the DAS Advanced and Predictive Analytics Market Study findings, webinars, etc.

Internal Planning
- Refine internal product plans and align with market priorities and realities as identified in the DAS Advanced and Predictive Analytics Market Study
- Better understand customer priorities, concerns, and issues
- Identify competitive pressures and opportunities
About Howard Dresner and Dresner Advisory Services
The DAS Business Intelligence Competency Center Market Study was conceived, designed, and executed by Dresner Advisory Services, LLC, an independent advisory firm, and Howard Dresner, its president, founder and chief research officer.

Howard Dresner is one of the foremost thought leaders in business intelligence and performance management, having coined the term “Business Intelligence” in 1989. He has published two books on the subject, The Performance Management Revolution – Business Results through Insight and Action (John Wiley & Sons, Nov. 2007) and Profiles in Performance – Business Intelligence Journeys and the Roadmap for Change (John Wiley & Sons, Nov. 2009). He lectures at forums around the world and is often cited by the business and trade press.

Prior to Dresner Advisory Services, Howard served as chief strategy officer at Hyperion Solutions and was a research fellow at Gartner, where he led its business intelligence research practice for 13 years.

Howard has conducted and directed numerous in-depth primary research studies over the past two decades and is an expert in analyzing these markets.

Through the Wisdom of Crowds® Business Intelligence market research reports, we engage with a global community to redefine how research is created and shared. Other research reports include:

- Wisdom of Crowds “Flagship” Business Intelligence Market Study
- Advanced and Predictive Analytics
- Business Intelligence Competency Center
- Cloud Computing and Business Intelligence
- Collaborative Computing and Business Intelligence
- End User Data Preparation

Howard (www.twitter.com/howarddresner) conducts a weekly Twitter “tweetchat” on Fridays at 1:00 p.m. ET. The hashtag is #BIWisdom. During these live events the #BIWisdom “tribe” discusses a wide range of business intelligence topics.

You can find more information about Dresner Advisory Services at www.dresneradvisory.com.
About Neil Raden

Neil Raden is the Founder and Principal Analyst of Hired Brains Research, a provider of research, consulting and advisory services in business intelligence, data management, advanced analytics and decision management.

Neil serves as a trusted advisor to executives in fields as diverse as healthcare, nuclear waste management, cosmetics marketing, and many in between. He maintains his reputation for professional leadership through contributions to technical and business publications, research reports and participation in industry events.


In 1998, he formed Hired Brains Magazine, a high-end monthly for professional consultants. Neil’s blog appears at http://hiredbrains.wordpress.com. He is a regular contributor to expert sites such as LinkedIn Groups, Focus, Quora, and eBizQ, as well as recent contributed articles to Forbes (print) magazine. Neil was also an early Wikipedia editor and administrator in technology, healthcare and mathematics. He can be reached at nraden@hiredbrains.com.
Survey Method and Data Collection
As with all of our Wisdom of Crowds market studies, we constructed a survey instrument to collect data and used social media and crowd-sourcing techniques to recruit participants.

We expanded data collection to include our own research community of 3,000 organizations as well as vendors’ customer communities.

Data Collection
We received a total of 885 responses for this study. This is the second year we have conducted our Advanced and Predictive Analytics survey.

Data Quality
We carefully scrutinized and verified all respondent entries to ensure that the study includes only qualified participants.
Executive Summary
Executive Summary

- There is a substantial gap in adoption of advanced and predictive analytics (APA) between the industry-messaging machine of vendors, media, and analysts, and the actual people who generate analytical output in organizations. While awareness of its importance is high—89 percent of respondents (p. 28-33) attach at least some importance to advanced and predictive analytics that varies by function, headcount, etc. Adoption and practice are far from universal (p. 34-39).
- APA tools are early in the stages of deployment (though not development, since many of these tools have been available for decades). Just 27 percent report current use that varies by role, industry, headcount, etc. (p. 39-47). At least as many have no plans at all.
- In our sample, higher education, telecommunications, financial services, and retail have slightly higher-than-average adoption plans (p. 36). Fraud and churn analysis in telecommunications and financial services and customer analytics and logistics in retail are some of the most common use cases cited for advanced and predictive analytics.
- The most popular analytic features support traditional statistical methods: regression models, hierarchical clustering, and textbook functions (p. 40-45).
- We noticed that data preparation issues are uniformly important: respondents have high and growing interest in a broad set of data preparation features, led by set operations and support/automation for data quality support (p. 46-51).
- The importance of usability categories is uniform, with specific features attracting equal interest. Most respondents, led by front-office sales/marketing/executives say analytic usability automation and self-service are very important (p. 52-55).
- Among scalable technologies and architectures, in-database analytics and in-memory analytics attract the highest interest among respondents, with close to 60 percent describing them as “critical” or “very important.” In-Hadoop analytics are the third-most cited scalable analytic platform but are notably less “critical” or “very important” to respondents (p. 56-61).
- APA practitioners very likely are separated in function from many of the respondents in this survey. Data scientists are most likely to be “constant” users (p. 22-27). Deep quantitative analytics are often separated from Line-of-Business groups; the extent and sophistication of analytics already in place in organizations may be understated in these responses, as well as rankings of important/unimportant features.
- What the survey represents is the state and trajectory of analytics among those who are already producing and using BI in their organizations and will likely gravitate to more complex modeling through training, packages from third parties, and more useful tools.
Study Demographics
In its second year, we have 885 respondents, which is a 30 percent increase from last year’s report (679 respondents). The study participants include a rich cross-section of data populations across geographies, functions, organization size, and vertical industries. We believe that, unlike other industry research, this supports a more representative sample that is a better indicator of true market dynamics. We constructed cross-tab analyses using these demographics to identify and illustrate important industry trends.

Geography
The survey respondents represent a broad span of geographies. North America, which includes the United States, Canada, and Puerto Rico, represents the largest group with 48 percent of all respondents. Globally, the second largest group is EMEA, with participation of 33 percent (fig. 1).

![Geographies Represented](http://www.dresneradvisory.com)
Functions
Similar to our last study, information technology is the lead function and represents about one-third of respondents. Information technology, executive management, and the business intelligence competency center (BICC) together represent about two-thirds of respondents (fig. 2).

![Functions Represented](image-url)
Vertical Industries
Technology, consulting, and healthcare lead industry participation in 2015 and combine for 32 percent of our sample (fig. 3). Consulting respondents also provide insight into the partner ecosystem for BI vendors.

Figure 3 = Vertical industries represented
Organization Size
We received strong participation from a balanced mix of small, medium, and large organizations (fig. 4). We broke the smallest organizations into two groups (1-25 and 26-100 employees) to better separate shops and proprietorships from other small businesses. Small (1-100) organizations represent 31 percent of our sample, midsized (101-1,000) account for 24 percent, and large organizations (1,000 and greater) reflect 41 percent of the sample.

![Organization Sizes Represented](http://www.dresneradvisory.com)
Analysis and Trends
Analysis: Business Intelligence Users

Importance of Business Intelligence
We begin with Business Intelligence importance because BI shares common traits with advanced and predictive analytics that include data availability, usability, and scalability. Close to 90 percent of respondents completely or somewhat agree that their BI efforts have been successful. Based on this record of acceptance and confidence, we believe that, informally, organizations view advanced and predictive analytics as building on existing BI efforts and as the potential next step in a progression towards better overall insights and capabilities.

Figure 5 - Success with business intelligence
Business Intelligence is an area of continued focus and investment across our respondent population. Though questions persist about the governance and application of BI throughout organizations, respondents are nonetheless confident in their BI work. One-quarter of respondents see their organization operating at the highest standard of “data as truth.” Another 42 percent report that the next highest state, a common understanding of data, is available across their organizations (fig. 6).

**Figure 6 - Business intelligence and the state of data**

- **We have multiple, inconsistent data sources with conflicting semantics and data. Information is generally unreliable and distrusted**: 12%
- **Consistent data is available at a departmental level. Conflicting, functional views of data causes confusion and disagreement**: 22%
- **A common view of enterprise data is available. However, parochial views and semantics are used to support specific positions**: 42%
- **Data as “truth” - A common view of enterprise data is available with common application of data, filters, rules, and semantics**: 24%
Introduced in 2014, Action on Insight is Dresner Advisory’s high-level self-assessment of BI best (and worst) practices. In different contexts, Action on Insight measures how well organizations put data to use above and beyond passive reporting/observation.

Organizations sampled have a high estimation of their ability to execute action on insight. More than one-quarter of respondents report the highest state of “closed-loop processes” (fig. 7). Another 59 percent have achieved the next-highest state of ad hoc action on insights; just 4 percent say insights are rarely leveraged.
Analysis: Advanced & Predictive Analytics

Types of Users for Advanced and Predictive Analytics
Business intelligence experts, business analysts, and statisticians/data scientists are early adopters of advanced and predictive analytics, with more than 50 percent of each group using analytics, at minimum, “occasionally” (Fig. 8). Data scientists are most likely to be “constant” users. Other groups are far less-frequent users: about 60 percent of marketing analysts and executives, and 70 percent of third-party consultants rarely or never use advanced/predictive tools.

Figure 8 - Users of advanced and predictive analytics
Statisticians and data scientists are consistently the most likely “constant” users of advanced and predictive analytics, led by Asia Pacific, EMEA, and North America (fig. 9). Other prominent internal users (BI expert/analyst/executive) break somewhat inconsistently across geographic regions. Executive and marketing analyst use is weakest in EMEA.

**Figure 9 - Users of advanced and predictive analytics by geography**

![Users of Advanced and Predictive Analytics by Geography](image-url)
Advanced and predictive analytics user types either support or report up to a variety of business roles. As shown in fig. 10, statisticians/data scientists are “constant” or “often” analytic users typically supporting the business intelligence competency center (BICC), followed by sales and executive management. BI expert users are likely found supporting BICC and executive management functions. Business analyst users are most in demand among sales and executive management. Marketing and executive users of APA mostly have bespoke roles in their own domain but also are found in support of sales and work to support the competency center also. Third-party consultant users take lesser roles but most often are engaged in support of executives, marketing, and sales.

![Figure 10 - Users of advanced and predictive analytics by function](http://www.dresneradvisory.com)
Small (1-100 employees) and very large (10,000+) organizations consistently have the most users of advanced and predictive analytics across analyst, executive, and third-party functions (fig. 11). Again, the most constant users in organizations, specifically statisticians/data scientists, are found in organizations of all sizes.
In our sample, insurance followed by higher education, financial services, and healthcare account for the greatest number of statistician/data scientist “constant” and “often” users (fig. 12). Insurance, healthcare, and higher education also present the most constant and frequent users among BI experts. Retail & wholesale and manufacturing are laggards by comparison.

Figure 12 - Users of advanced and predictive analytics by industry

http://www.dresneradvisory.com   Copyright 2015 – Dresner Advisory Services, LLC
2014 – 2015 User Comparison
Between 2014 and 2015, the number of users for advanced and predictive analytics tools increased among statisticians / data scientists but decreased among most other user groups (fig. 13). Fewer business analysts (-5 percent) and BI experts, marketing analysts and executives were “constant” users in 2015 than in the previous year. There are more “never” and “rarely” users in 2015 for all roles except statisticians/data scientists.

Figure 13 - Advanced and predictive analytics users (2014 to 2015)
Importance of Advanced and Predictive Analytics

Globally, 89 percent of respondents attach, at minimum, some importance to advanced and predictive analytics (fig. 14). Almost 50 percent say that advanced analytic tools are at least “very important” to their mission. Just 11 percent say advanced and predictive analytics are “not important.”

![Importance of Advanced and Predictive Analytics](figure14.png)

Figure 14 - Importance of advanced and predictive analytics
Importance by Region
Asia Pacific, North America, and EMEA have a similarly favorable response to advanced and predictive analytics. Between 85 and 91 percent of respondents in these regions consider the tools and technology at least "somewhat important" (fig. 15). More than 25 percent of Asia-Pacific respondents (the most in our sample) find advanced analytics “critical.”

Figure 15 - Importance of advanced and predictive analytics by geography
Importance by Function
Marketing, sales, and executive management have the strongest positive sentiment overall toward APA (fig. 16). Marketing is likewise most likely to find analytics “critical,” well ahead of similar sentiment in other functions. IT and finance are comparatively more likely to find advanced analytics only “somewhat important” or “not important.”

Figure 16 - Importance of advanced and predictive analytics by function
Importance by Industry
Between 75 and 85 percent of all industries sampled consider APA, at minimum, “important” (fig. 17). By weighted mean, telecommunications, closely followed by education and healthcare attach the highest importance. These three industries are also most likely to consider the technology “critical.” Financial services and retail and wholesale have a slightly lesser take on the importance of the technologies; but half still consider analytics, at minimum, “very important.” Manufacturing trails all industries in measures of perceived importance.

Figure 17 - Importance of advanced and predictive analytics by industry
Importance by Organization Size
Very large organizations (10,000+ employees) attach the greatest importance to APA, with 57 percent saying they are at least “very important” (fig. 18). Mid-sized (100-1,000) and some large (1,000-10,000) organizations are about 50 percent likely to say analytics are, at minimum, very important. Small organizations trail mid-sized and larger peers by only small percentages.

Figure 18 - Importance of advanced and predictive analytics by organization size
2014 – 2015 APA Importance Comparison

When we compare our 2014 and 2015 sample populations, the perceived importance of advanced and predictive analytics remains fairly constant between the two groups (fig. 19). Grades of “very important” and “important” changed only slightly year over year. Slightly more respondents consider the technology “critical” in 2015, but slightly more also say it is only “somewhat important.”

![Importance of Advanced and Predictive Analytics 2014 to 2015](image)

*Figure 19 - Importance of advanced and predictive analytics 2014 to 2015*
Adoption Plans for Advanced and Predictive Analytics

Deployment Plans for Advanced and Predictive Analytics
Penetration and use of advanced and predictive analytic tools remains very low, with current users accounting for just 27 percent of our sample. Forty-four percent of organizations are either using or evaluating the technologies; but 37 percent are only considering advanced analytics and 19 percent have no plans for their use (fig. 20).

![Current Deployment of Advanced and Predictive Analytics](image-url)

*Figure 20 - Current deployment of advanced and predictive analytics*
Though 73 percent of organizations have current or future plans for adoption, the uptake of advanced and predictive analytics has been slow (fig. 21). 2015 adoption (most likely to be currently budgeted) is reported by just 17 percent or organizations. One-third have no plans for this year or 2016 and more than one-quarter have no future plans at all.

![Deployment Plans for Advanced and Predictive Analytics](http://www.dresneradvisory.com)

---

**Figure 21 - Deployment plans for advanced and predictive analytics**
Adoption Plans by Industry
By industry, 2015 and 2016 adoption of advanced and predictive analytics is strongest in higher education, telecommunications, and financial services (fig. 22). Both retail & wholesale and healthcare plan accelerated adoption in 2016. More than half of all organizations will adopt after 2016 or have no current plans to adopt.

Figure 22 - Deployment plans for advanced and predictive analytics by industry
Adoption Plans by Organization Size

Large organizations (> 1,000 employees) are generally most likely to adopt advanced and predictive analytics in 2015 and beyond and least likely to have no adoption plans (fig. 23). Small (1-100) and mid-sized (100-1,000) organizations will see increasing percentages but slowing growth in adoption in future timeframes. Mid-sized organizations lead in 2015 adoption rates, though current year adoption by organizations of any size is just 15 to 20 percent.

Figure 23 - Adoption plans for advanced and predictive analytics by organization size
Adoption Plans by Function

By function, marketing and BICC report the greatest 2015, 2016, and ensuing years’ deployment plans for APA (fig. 24). By comparison, these departments are four to five times more likely than sales or finance to adopt this year. We believe marketing deployments are due in part to campaigns/launches and the use of demographic and social media. BICC adoption is likely linked to advanced services that serve various other departments and lines of business. Adoption beyond 2016 becomes somewhat more comparable across departments, though 30 to 40 percent of sales and finance respondents report no plans.

Figure 24 - Adoption plans for advanced and predictive analytics by function
Adoption Plans by Geography
Near-term deployment of advanced and predictive analytics will be strongest in Asia Pacific, where 25 percent will adopt this year, followed by EMEA (fig. 25). North America and Latin America are least likely to adopt this year.

Figure 25 - Adoption plans for advanced and predictive analytics by geography
Feature Requirements of Advanced and Predictive Analytics

Respondents expressed interest in a broad range of feature requirements for advanced and predictive analytics (fig. 26). The most popular among these support traditional statistical methods: regression models, hierarchical clustering, and textbook statistical functions and are, at a minimum, "important" to at least three-quarters of respondents. Geospatial analysis, highly associated with mapping, populations, demographics and other Web-generated data, is the next most required feature. A variety of search, automation, standards-based, and esoteric features also attract interest, more specifically by geography (fig. 27), industry (fig. 28), and other variables. The rather large percentages of “somewhat important” and “unimportant” responses in key types of quantitative analysis are probably an indication that the survey base does not yet possess deep experience with predictive analysis.

---

**Figure 26 - Features for advanced and predictive analytics**

- Range of regression models, from linear, logistic to nonlinear
- Hierarchical clustering, expectation maximization, k-Means, and variants of...
- Textbook statistical functions for descriptive statistics
- Geospatial analysis
- Text analytic functions and sentiment analysis
- Bayesian methods, including Naive Bayes and Bayesian Networks
- Recommendation engine included
- Automatic feature selection like principal component analysis (PCA)
- Vector machine (SVM) approaches for classification and estimation
- Neural networks supported
- Various approaches to CART (e.g. ID3, C4.5, CHAID, MARS, random forests, ...)
- Ensemble learning

---

http://www.dresneradvisory.com  Copyright 2015 – Dresner Advisory Services, LLC
Feature Requirements by Geography

By region, regression models and hierarchical clustering attract the most interest across all industries and functions (fig. 27). Asia Pacific shows the greatest regional interest in a large number of features sampled, including textbook statistical functions, text analytics, automatic feature selection, vector approaches and approaches to CART. Perhaps noticeably, North America is least interested in neural networks and EMEA is least drawn to a mix of features including Bayesian methods, geospatial analysis, and textbook functions.

![Analytical Features by Geography](current_page)

**Figure 27 - Analytic features by geography**
Feature Requirements by Industry

In our sample, respondents in the insurance industry are especially drawn to regression models, geospatial analysis, and Bayesian methods (fig. 28). Education is also most drawn to regression models along with textbook functions, Bayesian methods, and approaches to CART. While typically an industry laggard, healthcare tracks rather closely with financial services in mid-tier interest in the full range of analytic features. Manufacturing trails all industries in nearly all categories of feature interest.

![Analytical Features by Vertical Industry](http://www.dresneradvisory.com)
Feature Requirements by Function
By function, finance is most noticeably the laggard in all areas of feature interest (fig. 29). Executives, not a typical audience for advanced and predictive analytics, show a somewhat surprising high interest in several advanced features and even lead in a few categories. Marketing reports the highest interest in automatic feature selection and vector machine approaches.

Figure 29 - Analytic features by function
Feature Requirements by Organization Size

Different sizes of organizations have a fairly consistent view of the importance of analytical features that mostly ranges between a mean of 3 and 3.5 (fig. 30). Very large (10,000+) organizations have slightly greater overall interest in a range of features, but especially outstrip small enterprise (1-100) interest in areas including hierarchical clustering, geospatial, Bayesian methods, vector machine approaches, and neural networks.

Figure 30 - Analytic features by organization size
2014 to 2015 Analytic Features Comparison
Between 2014 and 2015, interest increased in all advanced and predictive analytic features in our study. Fig. 31 ranks interest in analytical features by growth year over year. By this measure, the most desired features, regression models, hierarchical clustering, and textbook statistical functions also grew the most incremental interest year over year.

![Analytical Features 2014 to 2015](http://www.dresneradvisory.com)

**Figure 31 - Analytical features 2014 to 2015**
Data Preparation Plans for Advanced and Predictive Analytics

For the second year, our study addresses a detailed set of data preparation features that support advanced and predictive analytic activities and processes (figs. 32-41).

Almost 40 percent of respondents view set operations (joins, aggregations, pivots) as the most critical component of data preparation plans (fig. 32). Features related to data quality, detection of duplicates and cleansing/enrichment are the next most popular features. Overall, however, respondents attach high importance to all data preparation features sampled; 60 to 70 percent or more say all features are, at minimum, “very important.” Fewer than 10 percent say any categories are “unimportant.”

![Data Preparation for Advanced and Predictive Analytics](http://www.dresneradvisory.com)
Data Preparation by Industry

Across vertical industries, respondents have similar interest in several data preparation processes (fig. 33). Mean values, mostly between 3.5 and 4, reflect above-average interest in nearly all categories and industries. Respondents in telecommunications report especially high interest in several preparation practices, notably complex filtering and detection of duplicates. As found elsewhere, manufacturing is uniformly least interested in data preparation methods and features.

Figure 33 - Data preparation by industry
Data Preparation by Organization Size

Figure 34 illustrates attitudes toward data preparation activities and processes in organizations of different size. As noted elsewhere in our report, small and large organizations report greater levels of interest in data preparation than mid-sized peers. However, interest in unique manipulations and processes is very steady across the four most popular categories: set operations, detection of duplicates, cleansing and enrichment, and complex filtering.

Figure 34 - Data preparation by organization size
Data Preparation by Function
By function, front-office marketing, executive management and, separately, the BICC report the highest interest in data preparation activities and processes. Finance shows the least interest (fig. 39). IT and sales interest fall somewhere between the other groups. (A mean value of 3.0 is considered "important" and 4.0 is "very important.")
Data Preparation by Geography
Interest in data preparation activities and processes is fairly consistent across geographic regions with minor variations (fig. 36). EMEA shows slightly greater attention to set operations, cleansing and enrichment, and support for data conversions. North America’s greatest emphasis is set operations and complex filtering. Asia Pacific leads other regions in support for cutting and merging. Latin America trails slightly in all categories of interest.

Figure 36 - Data preparation by geography
**2014 to 2015 Data Preparation Comparison**

Across all activities and processes measured, interest in data preparation increased between 2014 and 2015. Chart 37 ranks importance by year-over-year change in sentiment. By this measure, interest in set operations, support for cutting and merging, and cleansing and enrichment of source data are growing in interest most quickly.

![Data Preparation 2014 to 2015](image)

*Figure 37 - Data preparation 2014 to 2015*
Usability for Advanced and Predictive Analytics

For the second year, our study addresses a detailed set of usability benefits that support APA activities and processes (figs. 38-47).

Usability features generally address process or activity automation and streamlining. At least half of all respondents consider each usability feature tested “very important.” Among these, the top choices are fast cycle times for data preparation, access to advanced analytics, a simple process for continuous modification, and support for easy iteration (fig. 38). The top features considered “critical” include fast cycle times for data preparation (22 percent) and pre-built, drag-and-drop menus (21 percent).

![Usability for Advanced and Predictive Analytics](image-url)

Figure 38 - Usability for advanced and predictive analytics
Usability by Organization Size
As exhibited elsewhere (figs. 23, 32, 37), interest in usability features at small and very large organizations outstrips interest in mid-sized enterprises (fig. 39). Among small enterprises, the greatest interest is seen in support for an entire process in a single application, support, and guidance for preparing analytical models and pre-built, drag-and-drop menus. Very large (>10,000) organization interest is highest in fast cycle time analysis and access to advanced analytics. Mid-sized organization interest is below mean value for every usability feature tested.
Usability by Function

By function, sales generally posts the most above-mean scores to most usability features tested and easily ranks highest in time-sensitive areas of automation including support for fast cycle times, support for analytical models, prebuilt macros, and automatic creation of models (fig. 40). Executive management and marketing also have above-mean regard for quick-turning and user friendly front-office usability features that is consistently much higher than seen in IT. Interest in the finance department is especially low, mostly below a figure considered “important” and far below the mean.

Figure 40 - Usability by function
2014 to 2015 Usability Comparison
Interest in a full range of usability features grew significantly between 2014 and 2015. We consider this chart a representation of the growth in demand for user independence and automation, one that will be echoed importantly in different vendor capabilities. Fig. 41 ranks usability features by growth year over year. By this measure, fast cycle time is the most important support feature for respondents in 2014 and 2015. Access to advanced analytics supported its “hot topic” status by moving into second place, ahead of support for easy iteration and simple process for continuous iteration of models.

![Usability 2014 to 2015](image)

Figure 41 - Usability 2014 to 2015
Scalability of Advanced and Predictive Analytics
For the second year, our study addresses respondent interest in a set of scalable technologies and architectures that support advanced and predictive analytics (figs. 42-49).

In-database analytics and in-memory analytics are the scalable technologies that attract the highest interest among respondents, with close to 60 percent describing them as “critical” or “very important” (fig. 42). In-Hadoop analytics are the third-most cited scalable analytic platform but are notably less “critical” or “very important” to respondents. Three-quarters or more of all respondents say all scalable analytic technologies including massively parallel processing and predictive model markup language are, at minimum, “somewhat important.”
Scalability by Region

High interest in in-database and in-memory analytics holds consistently across geographies (fig. 43). Asia Pacific, followed by North America and EMEA, has the highest regard for in-database analytics; Latin America places the greatest emphasis on in-memory. Asia Pacific reports the most above-mean interest in all scalable technologies and architectures sampled.
Scalability by Industry
Respondent interest in scalable analytic technologies is somewhat consistent across industries sampled (fig. 44). Insurance is slightly more interested than other industries in in-database analytics. Financial services, insurance, and education report the most interest in in-memory analytics. Financial services reports the most categories of above-average interest. Healthcare reports slightly above-mean interest in MPP and PMML support. Manufacturing trails interest by industry.
Scalability by Function

By function, sales, and executive management appear slightly more enthusiastic toward in-database, in memory, and in-Hadoop analytics. In this view, IT’s interest is also high in all categories but leads other lines of business in architectural support for MPP and PMML. As we noted in other measures, finance trails other departments in interest.

Figure 45 - Scalability by function
Scalability by Organization Size

As in other measures, small and large organizations have the greatest interest in scalable analytic features and architectures (fig. 46). Small organizations score highest in in-memory analytics, which may be considered de facto “enterprise” tools at small scale. Very large organizations somewhat expectedly show the greatest interest in “big data” and scalable architectures for data mining including Hadoop, MPP, and PMML.

Figure 46 - Scalability by organization size
2014 to 2015 Scalability Comparison
Between 2014 and 2015, interest in all scalable analytic features and architectures increased only marginally, even for “hot button” topics such as in-memory and Hadoop. Fig 47 ranks feature interest ranked by year-over-year change. By this measure, incremental growth in feature interest was about the same for all the categories sampled and falls between a measure of “somewhat important” (2.5 on our scale) and “important” (3.0).

![Scalability 2014 to 2015](image_url)

Figure 47 - Scalability 2014 to 2015
Industry and Vendor Analysis
Industry and Vendor Analysis

Industry Importance of Advanced and Predictive Analytics
The greatest majority of vendors consider advanced and predictive analytics to be either “critically” or “very” important. That said, more than a quarter of vendors place a relatively low priority upon it, suggesting a developing market (fig. 48).

Figure 48 - Industry importance of advanced and predictive analytics
Industry Support for Analytical Features and Functions

Characteristic of a developing market, respondents indicate strong support for basic functionality (e.g., descriptive statistics) with far more limited support for advanced capabilities such as neural networks (fig. 49).

Figure 49 - Industry support for analytical features and functions
Industry Support for Data Preparation

Data preparation, a key step in the analytical process, has strong support among surveyed software vendors today (fig. 50).

![Industry Support for Data Preparation](image-url)

Figure 50 - Industry support for data preparation
Industry Support for Tool Usability

With business analysts as a key target for this capability, usability has heightened importance. While a majority of vendors offer a core of these features, several (e.g., prebuilt macros, auto-creation of models) are less common (fig. 51).

![Industry Support for Tool Usability Features](image)

Figure 51 - Industry support for tool usability features
Industry Support for Scalability

Scalability takes a number of forms. Massive volumes of data, large communities of users, and support for enterprise-class and real-time applications are several of them. Currently most vendors offer support for large sets of data via in-memory with declining support for large user communities and enterprise-scale applications (fig. 52).

Figure 52 - Industry support for scalability features
Advanced and Predictive Vendor Ratings

In rating the vendors, we considered core advanced and predictive features, data preparation, usability, scalability, and integration. To be ranked, we required a minimum score of 24/37. As a result, we ranked 13.

The top vendors for advanced and predictive analytics in 2015 include SAP and SAS, which are tied for first place, RapidMiner in second place, Dell Software (Statsoft) in third place, IBM in fourth place, and Birst, Oracle, and TIBCO tied for fifth place (fig. 53).

![Advanced and Predictive Analytics Vendor Ratings](image_url)
Appendix: Advanced and Predictive Analytics Survey Instrument

In this section we would like to ask a few questions related to Cloud (Software-as-a-Service - SaaS), Collaborative Business Intelligence and Advanced & Predictive Analytics.

Advanced & Predictive Analytics includes statistics, modeling, machine learning, and data mining to analyze facts to make predictions about future, or otherwise unknown, events.

Cloud Business Intelligence utilizes a cloud-oriented architecture (e.g., elasticity, multi-tenancy) to host BI applications and tools as a service.

Collaborative Business Intelligence couples BI with capabilities for co-creation, commenting, and sharing (among others) to enable group-based decision-making.

1. How important are Cloud-based, Collaborative Business Intelligence and Predictive Analytics to your organization?

<table>
<thead>
<tr>
<th>Importance of Cloud-based BI?</th>
<th>Critical</th>
<th>Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Not Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance of Collaborative BI?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Importance of Advanced &amp; Predictive Analytics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Section: Advanced & Predictive Analytics: includes statistics, modeling, machine learning, and data mining to analyze facts to make predictions about future, or otherwise unknown, events.

2. Do you currently have Advanced and Predictive Analytics Deployed within your organization today?

- Yes
- No

3. Which kinds of users use Advanced & Predictive Analytics within your organization?

<table>
<thead>
<tr>
<th>Constantly</th>
<th>Often</th>
<th>Occasionally</th>
<th>Rarely</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI-expert</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marketing Analyst</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd Party Consultant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business Analyst</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistician/Data Scientist</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 4. Analytical Features: Which of the following features are important for Advanced & Predictive Analytics?

<table>
<thead>
<tr>
<th>Feature</th>
<th>Critical</th>
<th>Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Unimportant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatic feature selection like principal component analysis (PCA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendation engine included</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Various approaches to CART (e.g. ID3, C4.5, CHAID, MARS, random forests, gradient boosting)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range of regression models, from linear, logistic to nonlinear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text analytic functions and sentiment analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hierarchical clustering, expectation maximization, k-Means, and variants of self-organizing maps</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geospatial analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bayesian methods, including Naive Bayes and Bayesian Networks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neural networks supported</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ensemble learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Textbook statistical functions for descriptive statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vector machine (SVM) approaches for classification and estimation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 5. Data Preparation: For Advanced & Predictive Analytics, which data preparation capabilities are important?

<table>
<thead>
<tr>
<th>Capability</th>
<th>Critical</th>
<th>Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Unimportant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex filtering</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cleansing and enrichment of source data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detection of duplicates or outliers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support for cutting, merging, and replacing of values</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set operations e.g., joins, aggregations or pivot tables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support for data type conversions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 6. Usability: Which usability features are important for Advanced & Predictive Analytics?

<table>
<thead>
<tr>
<th>Feature</th>
<th>Critical</th>
<th>Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Unimportant</th>
</tr>
</thead>
<tbody>
<tr>
<td>A specialist &quot;NOT&quot; required to create analytical models, test and run them</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support/guidance in preparing data analytical models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Automatic creation of reports from data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast cycle time for analysis with data preparation functions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access to advanced analytics for predictive and temporal analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support for easy iteration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple process for continuous modification of models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pres-built drag and drop macros and tools from R that require no scripting or programming</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support for entire process in a single application/user interface</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
7. Scalability: Which scalability features are important for Advanced & Predictive Analytics?

<table>
<thead>
<tr>
<th>Feature</th>
<th>Critical</th>
<th>Very Important</th>
<th>Important</th>
<th>Somewhat Important</th>
<th>Unimportant</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Memory analytics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-Database analytics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-Hadoop analytics (on file system)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support for MPP architecture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMML Support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Other Dresner Advisory Services Research Reports

- Wisdom of Crowds “Flagship” Business Intelligence Market Study
- Advanced and Predictive Analytics
- Business Intelligence Competency Center
- Cloud Computing and Business Intelligence
- Collaborative Computing and Business Intelligence
- Embedded Business Intelligence
- End User Data Preparation
- Enterprise Planning
- Location Intelligence
- Mobile Computing and Business Intelligence
- Small and Mid-sized Enterprise Business Intelligence