



# Predictive Analytics: The Hurwitz Victory Index Report

Fern Halper, Ph.D



**HURWITZ  
& ASSOCIATES**  
Insight to Action



## Table of Contents

Introduction	3
Market Trends for Predictive Analytics	4
What is the Victory Index?	6
Victory Index Methodology	6
Key Metrics	6
Data Sources	8
Vendor Selection	9
Results	9
Market Perspective	10
Victors . . . . .	10
Leaders. . . . .	11
Customer Perspective	12
Victors . . . . .	13
Leaders. . . . .	13
Challengers . . . . .	14
Vendor Assessments	14
Angoss . . . . .	15
KXEN . . . . .	18
IBM (SPSS) . . . . .	19
Megaputer . . . . .	22
Pegasystems . . . . .	24
Pitney Bowes . . . . .	26
Rapid-I . . . . .	28
Revolution Analytics	31
SAP . . . . .	32
SAS . . . . .	34
StatSoft . . . . .	38
TIBCO. . . . .	40

Related Research: Advanced Analytics: End User Experiences and Future Plans (2011)  
Find it online at [www.hurwitz.com](http://www.hurwitz.com)

## Introduction

Predictive analytics is being embraced at an increasing rate by organizations that need to gain actionable and forward-looking insight from their data. While much of the statistics and data mining technology for predictive analytics has been around for decades, over the past several years the market has become red hot.

What has changed? Twenty years ago, statisticians in companies were able to predict who might drop a service using survival analysis or machine learning techniques. However, it was difficult to persuade other people in the organization that such analytics could be used to provide competitive advantage. For one thing, it was difficult to obtain the computational power needed to interpret data that kept changing through time. Additionally, predictive analytics was the realm of statisticians and mathematicians. Generally, this analysis was performed using some sort of scripting language. Model output could be hard to understand. Finally, it could be nearly impossible to convince a call center, for example, that this analysis needed to be “operationalized” using rules or scores derived from the models as part of a process to reach out to at risk users. It would have also been difficult to implement.

Today, adoption of predictive analytics has increased for a number of reasons including a better understanding of the value of the technology and the availability of compute power. Economic factors are also a driving force in utilizing predictive analytics for business as companies realize that simply looking in the rear view mirror to gain insight and make decisions is not enough to remain competitive. Companies want to better understand what actions their customers might take. They want to better predict failures in their infrastructure. The uses for predictive analytics are extensive and growing. Many vendors have made it a point to try to make predictive analytics more “user friendly” by automating some model building capabilities and providing information that is more easily understood by business users. Implementing predictive analytics as part of a business process has also become more popular as the software tools and techniques and the hardware to support this kind of deployment become more available.

Predictive analytics has become a key component of a highly competitive company’s analytics arsenal. Hurwitz & Associates defines predictive analytics as:

*A statistical or data mining solution consisting of algorithms and techniques which can be used on both structured and unstructured data (together or individually) to determine future outcomes. It can be deployed for prediction, optimization, forecasting, simulation, and many other uses.*

Table 1, on the next page, highlights some of the popular uses for predictive analytics.

*Today, adoption of predictive analytics has increased for a number of reasons including a better understanding of the value of the technology and the availability of compute power.*



**Table 1: Use cases for predictive analytics**

Area	Use Case
Marketing analytics	Predicting consumer behavior, churn analysis, consumption analysis, propensity to spend
Business analytics	Economic forecasting, business improvements, risk analysis, financial modeling
Industry specific analytics	Reliability assessment (i.e. predicting failure in machines), analytics situational awareness, behavior (defense), investment analysis, fraud identification (insurance, finance), predicting disabilities from claims (insurance), finding patterns in health related data (medical)

*End users and vendors alike see great advantages to incorporating predictive models in their business processes. For example, statisticians at an insurance company might build a model that predicts the likelihood of a claim being fraudulent.*

### Market Trends for Predictive Analytics

As end users grow to understand the value of predictive analytics, the market itself is evolving to include solutions targeted at different kinds of users and to deal with more data, and different deployment models. Hurwitz & Associates sees the following trends in the predictive analytics market:

- Providing solutions across the user spectrum.** Whereas the traditional user of predictive analytics was a statistician or other quantitative analyst, there is a change occurring in user type. A recent Hurwitz & Associates survey on Advanced Analytics (2011) indicates a shift towards business analysts as users of these products and services. This shift has spawned different philosophies and strategies among analytics vendors. Some are providing more user-friendly Graphical User Interfaces (GUIs) and automating their process of building predictive analytics models. For example, a number of vendors assessed in this report including KXEN, IBM SPSS, Pitney Bowes, SAS, and StatSoft, offer this capability. Vendors such as TIBCO are providing ways to share the results of a complex analysis via interactive mashups or other Web interfaces. These strategies can help non-technical users build and make use of a certain level of predictive analytics. Other vendors, such as Angoss, try to reduce the complexity of predictive analytics by offering both an on-premises version of the software as well as the results in a Software as a Service (SaaS) framework for the end user.
- Operationalizing models.** End users and vendors alike see great advantages to incorporating predictive models in their business processes. For example, statisticians at an insurance company might build a model that predicts the likelihood of a claim being fraudulent. The model, along with some decision rules could be included in the company's claims processing system in order to flag claims with a high probability of fraud. These claims would be sent to an investigation unit for further review. In other cases, the model itself might not be as apparent to the end user. For example, a model could be built to predict customers who are good targets for upselling when they call into a call center. The call center agent, while on the phone with the customer, would receive a messages on specific additional products to sell to this customer. The agent might not even know that a predictive model was working behind the scenes to make this recommendation. Companies such



as IBM SPSS, Pegasystems, StatSoft and SAS are offering solutions to support this kind of deployment which might include model scoring as well as rules for decision-making.

- **Supporting unstructured data analysis.** End users realize that they can gain significant insight from mining unstructured (i.e. text) data and that this unstructured data, when used in conjunction with structured data can provide a big lift to predictive models. Some companies, such as Megaputer Intelligence have been providing this ability for years. In another example, SAS bought Teragram (a text analytics vendor) several years ago to analyze unstructured data and it has integrated this capability into its product line. A number of other vendors profiled in this report also have this capability, including IBM SPSS and StatSoft.
- **Open source.** Open source solutions are becoming increasingly important to the predictive analytics market because they enable a wide community to engage in innovation. Open source solutions are often used in academic institutions (at no cost), so students coming into the workforce are already familiar with them. But the openness of the software can also mean it is less accessible to non-technical users. Vendors are trying to change that. For example, predictive analytics software vendors such as TIBCO, SAS and StatSoft have started incorporating these open source languages in their solutions and new entrants are emerging to provide software and services on top of these languages. For instance, open source R is a free software environment for statistical computation and graphics. It consists of a language plus a run-time environment with graphics, a debugger, access to certain system functions, and the ability to run programs stored in script files.<sup>1</sup> Revolution Analytics, a vendor profiled in this report, wraps its software and services around R to make it easier to use and provide better performance. Rapid-I, another vendor profiled in this report, builds its product on an open source model. Other vendors in this space include Pentaho.
- **Big Data.** As companies gather greater volumes of disparate kinds of data (i.e. structured and unstructured) they are looking for solutions that can scale to analyze this increased volume. Real time analysis of large amounts of data is also becoming more prevalent. For example, companies might use this capability for real time ad placement. Or, a hospital might be interested in analyzing patient data from medical devices in real time. The buzz from vendors about “Big Data” solutions is growing louder. For instance, SAS is working with EMC Greenplum and Teradata for dedicated high-performance analytic appliances. IBM recently bought big data vendor Netezza. SAP just announced general availability of SAP HANA for big data. Other providers, such as Alpine Miner are purpose built for analyzing big data.
- **Model management.** More vendors are beginning to offer solutions that help manage the development and deployment of models. There is a movement away from simply utilizing a directory structure to providing more robust solutions that can, for instance, help manage model versioning or alert an analyst when a model gets stale. Vendors like IBM SPSS, SAS and StatSoft currently offer these capabilities. Others, like Pitney Bowes, say they will be providing them in the near future.
- **Performance.** As the amount of data continues to increase, companies are asking for higher performance systems. Many vendors are upgrading from supporting only 32 bit to supporting 64 bit machines.

*There is a movement away from simply utilizing a directory structure to providing more robust solutions that can, for instance, help manage model versioning or alert an analyst when a model gets stale.*

<sup>1</sup> Hornik (2011), “The R FAQ”, www.r-project.org, ISBN 3-900051-08-9



- **In-database analytics.** A trend that began a number of years ago but has picked up steam is in-database analytics. The idea behind in-database analytics is to make analytics more efficient by bringing the analytics to the data rather than the other way around. This means that users don't have to extract data from a data appliance or warehouse to run models; the models run in the data store. SAS is a big proponent of this. IBM SPSS and KXEN also support this model.
- **Analytic access.** Another trend that began several years ago and is still growing is the ability for users to access insight from multiple devices including desktop and mobile devices (i.e. laptops, tablets, smart phones, and the like).

These are only some of the trends that are making predictive analytics a vital part of the analytics market landscape. Companies looking to effectively compete should consider making predictive analytics part of their analytics plan.

## What is the Victory Index?

The Hurwitz & Associates Victory Index is a market research assessment tool, developed by Hurwitz & Associates that analyzes vendors across four dimensions: Vision, Viability, Validity and Value. What is unique about the Victory Index? Hurwitz & Associates takes a holistic view of the value and benefit of important technologies. We assess not just the technical capability of the technology but its ability to provide tangible value to the business. For our inaugural Victory Index report we have selected the field of predictive analytics. Intensive research by our analysts on the business value of this technology supports our conclusion of great potential for continuing innovation and market growth in this arena.

## Victory Index Methodology

### Key Metrics

In order to provide a comprehensive understanding of the market and the vendors, Hurwitz & Associates applies a rigorous methodology that employs approximately 50 attributes across four dimensions, which are analyzed with a weighted algorithm. These dimensions are explained below and the specific metrics are presented in Table 2.

- **Vision: The strength of the company's strategy.** Attributes evaluated here include: Is the vision not only realistic but also compelling? It is not enough to simply have a good vision; a company must also have a well-designed road map that can support this vision. Vision attributes also include more tactical aspects of the company's strategy such as a technology platform that can scale, well articulated messaging, and positioning. A key component of this dimension is clarity: it must be clear what business problem the company is solving for which customers.
- **Viability: The strength and vitality of the company in the market.** There is often not a direct correlation between the market strength of a company and its revenue or the number of years that it has been in business. This

*What is unique about the Victory Index? Hurwitz & Associates takes a holistic view of the value and benefit of important technologies. We assess not just the technical capability of the technology but its ability to provide tangible value to the business.*



is especially true when a market is emerging. Sometimes an emerging company can become a giant within a few years if they solve a complicated customer problem. Hurwitz & Associates evaluates a complex set of financial and company specific attributes that, when combined, help to determine the viability of a company. Attributes include financial ratios, customer adoption rates, intellectual property, strength of management team, and strength of partnerships. We are also concerned with the vitality of the company in a particular market. For example, a large company might be strong and vital in a certain product area, but may be lagging in the predictive analytics space because of limited investment or market push.

- **Validity: The strength of the product that the company delivers to customers.** There is a distinction between the positioning of a product and its ability to satisfy customer requirements. Therefore, the Victory Index analyzes how well a company executes on its promises. This part of the Index examines the features and functions delivered. It looks at how deep the functionality is and how effective the product is at evolving based on changing customer requirements. Other important attributes include ease of use, innovation, how well the product integrates with other technologies that customers need, and how well it incorporates important de facto and de jure standards.
- **Value: The advantage the technology provides to customers.** Even if a product is well designed, it must be able to help businesses achieve their business goals. Goals range from gaining insight about customers in order to be more competitive, to using the technology to increase revenue. A key attribute that is measured in this dimension is how well the product supports companies in meeting their objectives.

*...a large company might be strong and vital in a certain product area, but may be lagging in the predictive analytics space because of limited investment or market push.*

**Table 2: Victory Index attributes**

Dimension	Attributes
Vision	<ul style="list-style-type: none"> <li>• Vision: clarity of vision, practicality of vision, a compelling vision, target market alignment, product direction and road map, leadership team vision</li> <li>• Messaging: strength of message, key differentiator value</li> <li>• Direct feedback from customers regarding vision</li> </ul>
Viability	<ul style="list-style-type: none"> <li>• Viability measures: revenue/employee, net profit, current assets/current liability, revenue growth, customer growth</li> <li>• Vitality: strength of intellectual capital, strength of leadership team, customer adoption, customer stability, competitive strength in a particular market, market presence, channel strategy</li> <li>• Direct feedback from customers regarding viability/vitality</li> </ul>
Validity	<ul style="list-style-type: none"> <li>• Direct customer measures: breadth and depth of functionality, technical benefit, tools and technology, technical value, performance, scalability, flexibility</li> <li>• Product functionality: data types supported, data preparation, model management and support, integration capabilities, model export capabilities, databases supported, in-database functionality, models supported, and other unique features</li> </ul>



---

Value	<ul style="list-style-type: none"> <li>• Direct customer measures: customer satisfaction, customer support from customer's perspective, ease of use, product innovation, business benefit as determined by customer, time to value, value/price</li> <li>• Social media metrics around customer loyalty and brand value</li> </ul>
-------	--

---

The Vision, and Viability indicators are further grouped under the heading of Market Strength since these metrics, analyzed together, give a good indication of the strength of the offering to the market. The Value and Validity measures are also analyzed together since they give a good indication of how customers view the offerings.

**Data Sources**

Data for scoring comes from multiple sources:

- **Vendor surveys:** An extensive survey was designed to obtain information on each vendor's vision, strategy, products, financial, and company stability, as well as what the company believes to be its value. Hurwitz & Associates also conducted briefings with each vendor to further augment and understand this information. In several cases where the vendor chose not to complete a survey, information obtained from other data sources were compiled in order to understand these attributes and a written summary was provided to the vendor for fact check and review.
- **Vendor references:** Each vendor supplied 3-5 names of clients who provided input regarding the value and validity of the vendor products to Hurwitz & Associates. Hurwitz & Associates either sent a survey to each of these clients or conducted a phone conversation with them, or both. Given that many companies use more than one predictive analytics solution, these conversations were also used to understand where one vendor's product provided value in comparison to other vendor solutions.
- **Online Victory Index Survey:** In addition to speaking with vendor references, Hurwitz & Associates conducted several online surveys to gather direct customer feedback regarding vendor products. We compiled results from hundreds of companies regarding how they were using vendor products, what they were using them for, and their experience with the product. We also asked respondents to rate the vendor's products across the four Victory Index dimensions listed above. Survey results provided an objective view of each vendor's offering.
- **Social media:** A unique aspect of the Victory Index is that it uses social media data to help inform results. Social media was used, when appropriate, to determine brand media dominance, brand loyalty and strength, and brand value. The commercial product, NetBase ([www.netbase.com](http://www.netbase.com)), was used to perform this analysis. Netbase Social Media Insight & Analysis is used by marketing teams to understand online brand equity, analyze and compare customer passion, and generate insights that answer "why" questions. The platform is used by hundreds of corporate customers, and was developed in partnership with five of the top 10 consumer packaged goods companies, including Coca-Cola and Kraft.
- **Hurwitz inquiries:** Input from discussions with Hurwitz clients was used in the study.

*A unique aspect of the Victory Index is that it uses social media data to help inform results.*



- Other secondary sources included investment reports and public filings among others.

## Vendor Selection

Unlike other assessment methodologies that limit the vendors to be included based on size, we felt it was important to present potential buyers with a range of different solution options. Whereas, in the past, predictive analytics solutions were predominately used by larger enterprises, some of the solutions on the market today have a price point that makes them feasible for use by smaller companies, too. Therefore, we have included vendors that may be new to the market, but have offerings worthy of consideration from buyers. These include vendors that are offering cloud solutions in the form of SaaS as well as vendors that are utilizing open source solutions as a basis for their offerings. We have also included companies that may have a strong presence among data miners, but have not yet appeared on the radar screen of the business buyer.

We believe all of the companies that we selected are strong contenders in the Victory Index.

Additionally, we have decided not to include Business Intelligence (BI) vendors in the report, such as IBM Cognos or SAP Business Objects, unless the vendor has a named, specific predictive analytics offering (i.e. SAP Predictive Workbench, SAS Enterprise Miner). We did this for several reasons. First, predictive analytics is often included under the BI umbrella, although the technologies tend to be different. We wanted to focus on players with a focus on predictive analytics so as not to cloud the issue. Second, often the audience for BI and predictive analytics is different. BI might be an IT buying decision, but predictive analytics deals mostly with the line of business. So, while many BI vendors offer forecasting capabilities, we felt companies exploring predictive analytics products were looking for solutions from a particular class of vendor. Of course, predictive analytics are often used in conjunction with BI offerings.

The vendors profiled in this study include Angoss, KXEN, IBM SPSS, Megaputer Intelligence, Pegasystems, Pitney Bowes, Rapid-I, Revolution Analytics, SAP, SAS, StatSoft, and TIBCO Spotfire. Some of these vendors offer solutions that are broad based, statistical and data mining solutions that can be used for a variety of use cases. These vendors often offer targeted solutions in particular areas. Others, such as KXEN, Pegasystems, and Pitney Bowes are more customer-relationship-problem focused.

## Results

Results are based on a thorough analysis of the data described above. Vendor scores result from use of a weighted algorithm across the four dimensions of the Victory Index. The results appear in two charts. Both are scatter plots which illustrate where each vendor scored relative to the median scores of the group. Figure 1 (Market Strength) is a plot of Vision vs. Viability. Figure 2 (Customer

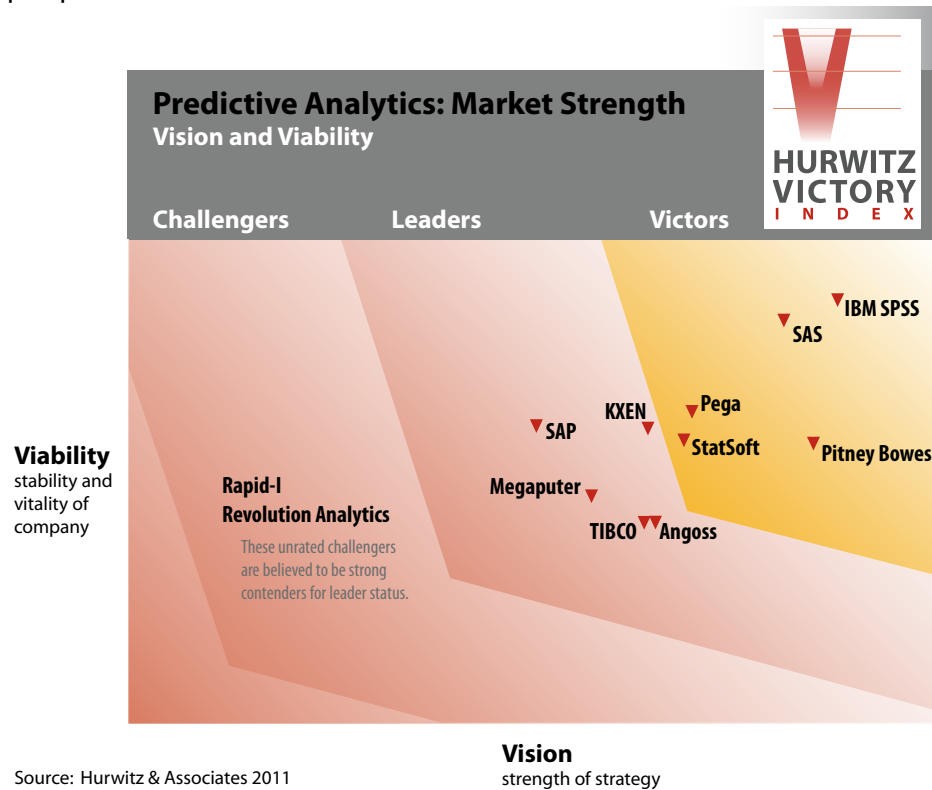
*Whereas, in the past, predictive analytics solutions were predominately used by larger enterprises, some of the solutions on the market today have a price point that makes them feasible for use by smaller companies, too.*



View) is a plot of Validity vs. Value. Some companies were winners across both categories. These are designated as Double Victor. These companies include IBM (SPSS), SAS, StatSoft, and Pegasystems. Complete vendor profiles appear in the next section.

**Market Perspective**

Figure 1, below, illustrates the results of the Victory Index from the market perspective.



*The Victors demonstrate significant strength in brand and have a strong market presence in predictive analytics.*

**Victors**

The Victors demonstrate significant strength in brand and have a strong market presence in predictive analytics. Their financials indicate stability and they have a compelling vision and strong strategy for their technology. These companies have a strong vision for their predictive analytics solution and their customers resonate with this vision. Their customers believe these companies are innovative in the predictive analytics space and their solutions provide strong value. Each Victor has a combined Vision and Viability score that ranks at or above the median scores across all the attributes measured in these two categories. IBM SPSS, SAS, StatSoft, Pegasystems, and Pitney Bowes received scores designating them as Victors from a Market Perspective.

- **IBM SPSS.** IBM SPSS is the overall Victor in this category. IBM has made no secret of its major strategic thrust into the analytics market. SPSS is a critical component in what IBM calls its Smarter Planet strategy. Smarter Planet holds a vision of the world as a more interconnected, instrumented, and intelligent place. For companies to be successful in this type of environment

requires a new emphasis on analytics. SPSS has deep predictive technical capabilities and is also focused on operationalizing predictive analytics. With a strong leadership team from both IBM and SPSS in place, growth that is exceeding expectations, and its ability to innovate, the company scored at the top of this category.

- **SAS.** SAS is also exceedingly strong from a market perspective. With its focus on industry solutions to solve business problems it provides a clear, pragmatic vision to the market. SAS has exhibited strong new customer growth and is making strides to differentiate its large set of products and solutions. It is fully vested in the predictive analytics space and it has a market leading approach to high performance analytics.
- **Pitney Bowes.** Portrait Software was recently acquired by Pitney Bowes and sits within its enterprise software division. Portrait's Customer Analytics and Interaction technologies play an important role in the company's Customer Communication Management vision. It still maintains its key strategic focus on B2C CRM and is clearly focused on predictive models for customer interaction management. While there is still an integration risk with its parent company, Portrait has had very healthy revenue growth.
- **Pegasystems.** Pegasystems, a market leader in Business Process Management (BPM), is focused on using its rules-based decision management solutions together with predictive analytics to help customers improve business value. Despite market concern about Pega's spending in 2010, the company still had healthy revenue growth.
- **StatSoft.** The StatSoft vision is one of providing deep statistical capabilities in a practical, easy to use, and low cost solution. The company makes it easy for customers to do business with them. Its customers totally resonate with this vision. The brand is best known among data miners, who feel a strong connection with the company. StatSoft is stepping up its marketing efforts to appeal to business users with its ease of use and low cost.

## Leaders

Leaders demonstrate solid brand and financial stability and have a good vision for their predictive analytics solution. However, companies in this category did not achieve Victor status for two main reasons. First, some Leaders are highly viable organizations, but their vision for predictive analytics is not as clearly articulated and well-executed as the companies rated as Victors. Second, some of the Leader companies recently dealt with management or other business changes that have hurt their ability to execute their predictive analytics strategy. These companies have strong technology and are in a great position to improve their Victory Index rating if they strengthen the clarity of their vision for predictive analytics. The Leaders are Angoss, KXEN, Megaputer Intelligence, SAP, and TIBCO.

- **Angoss.** After a tumultuous year shedding some of its marketing services as well as hiring a new senior leadership team, Angoss is now clearly focused on predictive analytics. Its strategy to provide both an on-premises and a services-based solution is a good one because it opens up options for companies that may not have predictive analytics skills in-house.

*[Leaders] have strong technology and are in a great position to improve their Victory Index rating if they strengthen the clarity of their vision for predictive analytics.*

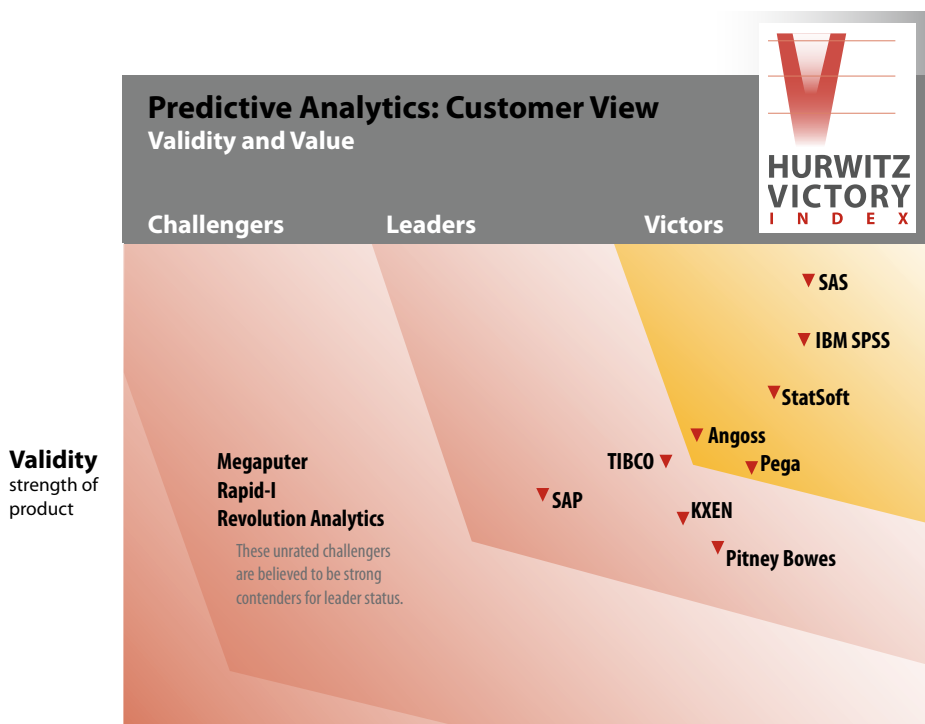


- **KXEN.** KXEN is clearly focused on automating predictive modeling for both large and small companies utilizing a factory approach to its model building. With a strong base of more than 400 customers and a strong value proposition in predicting customer behavior, it recently secured \$8M in series D funding to further accelerate vertical market development.
- **Megaputer Intelligence.** Megaputer, with its service focus on analyzing both structured and unstructured data together in predictive models has a good strategy moving forward. The company, which now has some solid partnerships in place with larger consulting companies such as Ernst & Young, is positioned to take advantage of this growing market space.
- **SAP.** SAP, a market leader in business intelligence is not developing its own innovations in predictive analytics as it has just renewed an OEM arrangement with IBM SPSS, which gives it a subset of SPSS capability. The company's recently announced products around Big Data may change its predictive analytics push.
- **TIBCO Spotfire.** The Spotfire brand is a clear leader in the visualization market and the combination of predictive analytics and visualization is quite powerful. Its strategy, to provide what it terms "visual free-dimensional analysis" (i.e. non OLAP cubes) to explore the results of models developed by statisticians, makes sense for many companies. It has an impressive and powerful tool set and needs to let the market know more about it.

*The Predictive Analytics Customer View chart measures both the customers' perspective on how well client objectives are met, as well as the strength of the products.*

### Customer Perspective

Figure 2, below illustrates the results of the Victory Index from the Customer Perspective



Source: Hurwitz & Associates 2011



**HURWITZ & ASSOCIATES**  
Insight to Action

## Victors

The Victors demonstrate superior technical and business value, technology and tools, customer support, and overall value as evidenced in customer survey scores and interviews with customers. Victors also have significant depth (and often breadth) of functionality and overall strong customer satisfaction scores. The combined scores across Value and Validity rank at or above the median scores in this category. SAS, IBM SPSS, StatSoft, Angoss, and Pegasystems are Victors in the Customer Perspective category.

- **SAS.** SAS is the overall Victor in this category. SAS received the highest scores across all companies for customer satisfaction, scalability of the solution technology and tools, and technical value. The sheer breadth and depth of its functionality across the predictive analytics lifecycle continues to impress customers with complex problems. Some customers expressed concern over the price of the SAS solution, but appreciate the strong business and technical value.
- **IBM SPSS.** SPSS scored close to SAS and we expect that as the IBM transition completes, the gap will continue to narrow. IBM SPSS received the highest scores among the larger full-featured vendors for ease of use and new functionality rollout. It ranked high in breadth and depth of functionality, robustness of software, and ease of deployment.
- **StatSoft.** StatSoft also received high marks from customers. Customers felt that the tools and technology were strong and that the value/price was excellent. In fact, StatSoft ranked first in value/price among all of the vendors in the study. Customers felt that StatSoft had the right balance of features for the cost. Customers also gave the company good grades for support and for ease of use of the interface.
- **Angoss.** Angoss received high marks from customers in flexibility of its software and ease of use. While not a classic statistics or data mining company, it received praise for its decision tree interface which many end users found easy to use.
- **Pegasystems.** Pegasystems received good scores across the board. Customers like Pega's tools and technologies and felt that it provided a good ROI. The Pegasystems solution is targeted to CRM and recognizing this, customers provided a lower score for solution flexibility and completeness.

## Leaders

Leaders have a solid product as indicated by customer survey scores. Customers are generally satisfied with their solutions and believe that they provide good value. However, the scores were not as high as the companies that achieved Victor status. The summary for each Leader below highlights the areas where the company scored well and points out one or two areas where customers would like to see improvement.

- **KXEN.** In keeping with its focus on helping business users create models, KXEN received high overall scores for ROI and business value. Customers liked its model automation features. Survey results indicate that customers would like to see KXEN improve its level of responsiveness to their support requests.

*The Victors demonstrate superior technical and business value, technology and tools, customer support, and overall value as evidenced in customer survey scores and interviews with customers.*



- **Pitney Bowes.** Pitney Bowes also received good scores across the board. Customers feel Portrait Software is easy to use and easy to deploy. In addition, the company provides good customer support and good business value. Survey results indicate that customers would like to see improvements in flexibility and depth of the software.
- **SAP.** Customers who were using SAP for business intelligence and used SAP for predictive analytics liked the ease of use of the interface and the flexibility of the software. They believed it provided good business value. The main areas of customer concern include the cost of the software, slow rollout of new functionality, and ease of deployment.
- **TIBCO Spotfire.** Customers liked the fact that the Spotfire product integrated data management, data preparation, statistical analysis, and visualization. Those who were using it to create interactive Web-based solutions felt that these solutions were easy to create. Survey results indicate that customers would like to see some improvement in support.

*As buyers look to purchase predictive analytics software they need to begin by asking some fundamental questions about the business problem(s) they are trying to solve.*

## Challengers

Challengers are viewed as potential Leaders or Victors, but are either too new to the market to obtain a good read on their strength or there is not enough information to place them in one of the other categories. Many of these vendors are open source providers which speaks to the importance that open source can play in the future of predictive analytics.

- **Rapid-I.** Rapid-I, an open source vendor, provides a market-leading environment called Rapid Miner that is freeware to individual users. Its strategy is to provide innovation and ease of use in open source analytics to enterprises through services. This makes sense because companies using open source solutions such as R, often complain that while it contains powerful algorithms, the interface is not user friendly, it is hard to learn, and deploy. In other words, it is not really enterprise ready.
- **Revolution Analytics.** Revolution Analytics, a new entrant to the market, has leveraged powerful marketing techniques to differentiate itself from some of the other vendors, offering solutions around open source R. Now, the company needs to continue to work to execute this vision.
- **Megaputer.** Megaputer provides a strong solution for customers who need to integrate structured and unstructured data for predictive analysis purposes. However, customers did not provide enough information to formally place the company into the Victor or Leader category from a customer perspective.

## Vendor Assessments

As buyers look to purchase predictive analytics software they need to begin by asking some fundamental questions about the business problem(s) they are trying to solve. It is also important for customers to evaluate employee analytics skill sets, budget, deployment models, and feature requirements prior to making a purchase decision. In the vendor assessments below, Hurwitz & Associates provides details regarding each vendor's products including target customers, horizontal and vertical solutions, licensing models, platforms supported, and key differentiators. We also provide information about how the vendor handles



key aspects of the predictive analytics process including data preparation, data formats and databases supported, type of data and analysis supported, model management and deployment, as well as whether the company supports other features such as in-database mining.

It is important to note that multi-vendor use is common (especially among larger firms). Companies often purchase more than one product for use. For example, a company might use one product because of its strong data preparation capabilities and another because of its ease of use for building decision trees. Or, a company may purchase a core statistics product to explore data and another product to do predictive modeling. Often buyers in companies will purchase products because they are familiar with them from another job or from academic training. However, they may also purchase an additional product because it is strong in their particular vertical market.

### Angoss

Company: Angoss Software

Public Website: [www.angoss.com](http://www.angoss.com)



Supports companies looking for a decision tree solution that is easy and quick for analysts who don't want to code. With its focus on SaaS deployments, companies without highly skilled analysts, especially in the financial vertical, who need more complex predictive functionality, might also consider Angoss.

There has been a lot of positive change recently at Angoss Software, a Canadian based company, founded in 1984, and best known for its risk management and marketing and sales analytics products. The past year has been a transformative one for the company. Although its core predictive analytics business continues to thrive, Angoss has replaced much of its senior management team and is exiting some business areas not core to its focus. Under its new leadership, the company has clarified its vision and strategy. It is returning to its roots in business intelligence and predictive analytics with a goal of increasing the use of predictive analytics by providing both an on-premises and a services-based solution. In addition to its core products, KnowledgeSEEKER and KnowledgeSTUDIO, Angoss has also been providing a SaaS solution for companies that don't have skilled analysts on board to perform predictive analysis. It has started to operationalize the determination of risk, fraud, and other kinds of predictive analysis it offers and provide this as a service to customers in a SaaS model. In fact, the SaaS model is a fundamental plank in the company's go-to-market strategy.

According to customers who use its on-premises software, KnowledgeSEEKER and KnowledgeSTUDIO represent excellent decision tree solutions that are easy and quick for analysts who don't want to code. Business analysts find these products fairly easy to use, as well. While Angoss products are not classical

*There has been a lot of positive change recently at Angoss Software, a Canadian based company, founded in 1984, and best known for its risk management and marketing and sales analytics products.*



statistical packages in the traditional sense, users like the fact that decision trees are easy to understand and communicate. The company also got good overall scores for the flexibility of its software and its time to value.

### Angoss Overview

Products	<ul style="list-style-type: none"> <li>• KnowledgeSEEKER version 7.5: Data mining and predictive analytics software that offers data profiling, advanced data visualization, decision tree, and model performance evaluation capabilities that enable both business and expert users to perform segmentation and classification analysis and determine relationships between variables in business data sets.</li> <li>• KnowledgeSTUDIO version 7.5: An advanced data mining and predictive analytics software for all phases of the model development and deployment cycle – profiling, exploration, modeling, validation, deployment (scoring), monitoring and building score cards.</li> <li>• StrategyBUILDER version 7.5: A standard module for KnowledgeSEEKER and KnowledgeSTUDIO that offers the ability to design, verify, and deploy the necessary predictive and business rules to execute sales, marketing or risk strategies.</li> <li>• In-Database Analytics Add-on: for KnowledgeSEEKER and KnowledgeSTUDIO that allows users to perform data mining tasks directly in the source database without importing tables into Angoss data files.</li> <li>• KnowledgeSERVER Scoring Engine version 7.5: KnowledgeSERVER is a deployment environment for Angoss models. It can be used for batch or real-time scoring of Angoss models.</li> </ul>
Target User/ company size	Business users, quantitative analysts, and statisticians/all company sizes with introduction of SaaS solutions
Verticals supported	Financial services, telecommunications, health sciences
Horizontal solutions	Risk Management, marketing and sales analytics
How offered	On-premises and SaaS, consulting services provided
Differentiator(s)	<ul style="list-style-type: none"> <li>• Offers a SaaS service where customer provides the data and Angoss performs analysis and provides information back to the client. This is an operational service.</li> <li>• StrategyBUILDER is a unique tool designed to build business rules from the output of either KnowledgeSTUDIO or KnowledgeSEEKER.</li> <li>• Very intuitive decision tree functionality.</li> </ul>

*[Angoss] ... allows users to perform data mining tasks directly in the source database without importing tables into Angoss data files.*



Platforms supported	<p>Standalone (desktop) configuration:</p> <ul style="list-style-type: none"> <li>• Windows XP Professional</li> <li>• Windows Vista</li> <li>• Windows 7</li> <li>• Windows Server 2003 and 2008</li> </ul> <p>Client/Server configurations:</p> <ul style="list-style-type: none"> <li>• Windows Server 2003 and 2008 on x86-32/64 architecture</li> <li>• AIX 5.2, 5.3, and 6.1 on POWER/PowerPC architectures</li> <li>• Solaris 8, 9 and 10 on SPARC architectures</li> <li>• Red Hat Enterprise Linux 4.x and 5.x on x86-32/64 architecture</li> </ul>
Predictive Analysis Families	<p>Classification (decision trees and neural networks), regression (linear and logistic regression), unsupervised learning (cluster analysis, principal component analysis), scorecard builder</p>

*Data import wizards with sampling options are provided for all supported data source types.*

### Key Features

- ETL: Data import wizards with sampling options are provided for all supported data source types (Text, Microsoft Excel, SAS, SPSS, and any database accessible via ODBC). The product set does not focus on data preparation although the company supports a range of tools for supporting new derived variables.
- Data formats/databases supported:
  - On Windows platforms: Text (delimited and fixed-width), Microsoft Excel, SAS, SPSS, and all databases accessible via ODBC.
  - On Unix and Linux, the same file types as on Windows, but supported database types are limited to the following: DB2, Oracle, Microsoft SQL Server, Teradata, Informix, Sybase, MySQL, dBase IV, and V, FoxPro, Greenplum, and PostgreSQL.
- Data limitations: unlimited.
- In-database mining: via additional module for its core KnowledgeSEEKER product and KnowledgeSTUDIO product that allows users to perform data mining tasks directly in the source database without importing tables into Angoss data files. The add-on is compatible with Teradata, SQL Server, and Netezza.
- Model export: PMML v3.2, SQL, XML, SAS, SPSS, Java, structured English.
- Model deployment: interactive deployment in KnowledgeSEEKER and KnowledgeSTUDIO, as well as batch and real-time scoring via KnowledgeSERVER, which supports the DMX scripting language (data mining extensions to SQL).
- Other: can utilize decision trees with continuous target variables, strategy trees, and assignment of treatments.



## KXEN

Company Name: KXEN  
Private

Website: [www.kxen.com](http://www.kxen.com)

KXEN provides value for companies that don't employ statisticians or companies that want to produce and deploy many models.



KXEN often uses the craftsman vs. factory floor analogy when it is trying to describe its philosophy about predictive analytics. Its goal is to provide non-statisticians with a quick (i.e. automated) way to build and rebuild hundreds or even thousands of models for analysis that changes rapidly. Examples of this might include models for micro-segmentation or campaign management. The philosophy is not about lovingly building each model with painstaking effort, it is about building lots of models that may change and deploying them into an organization quickly. The company provides its users with automated variable selection, data preparation, and model building, as well as a way to deploy these scoring models into data sources and manage them effectively.

Customers like the KXEN approach. While some commented that the solutions could be more complete in some areas, they like the fact that the product is easy to use and believe it is a good value for the price. Moreover, many believe that they are getting excellent business value from the product.

### KXEN Overview

Products	<ul style="list-style-type: none"> <li>• InfinitelInsight™ Explorer: Helps to prepare data for analysis, includes a semantic layer for advanced users to build reusable analytic records.</li> <li>• InfinitelInsight™ Modeler: provides automated model building for business users.</li> <li>• InfinitelInsight™ Scorer: used to score models and put them into production.</li> <li>• InfinitelInsight™ Factory: used for model management.</li> <li>• InfinitelInsight™ Social Network Analysis: used to understand social networks.</li> </ul>
Target User/ company size	Business user in any size company including small companies without statisticians and larger companies where there may be a predictive analytics bottleneck.
Verticals supported	Communications, financial services, hospitality, media, retail
Horizontal solutions	Churn, cross sell-upsell, customer experience, fraud analysis, risk analysis
How offered	On-premises license only

*Customers like the KXEN approach. While some commented that the solutions could be more complete in some areas, they like the fact that the product is easy to use and believe it is a good value for the price.*



Differentiator(s)	<ul style="list-style-type: none"> <li>• Supports organizations that need to operationalize the building of 100s or 1000s of models – a factory approach to predictive modeling.</li> <li>• Social network analysis.</li> </ul>
Platforms supported	<ul style="list-style-type: none"> <li>• Windows, Linux, Solaris, HP-UX, AIX</li> </ul>
Predictive Analysis Families	Association, classification, regression, time series

#### Key Features

- ETL: provides data preparation through KXEN Explorer.
- Data formats/databases supported: Teradata, Oracle, Microsoft SQL Server, IBM DB2, Netezza, Sybase IQ, Sybase ASE, MySQL, Microsoft Access, etc.). Access to major proprietary file formats (SAS, SPSS, Excel, etc.).
- In-database mining: SQL push to Teradata, Oracle, Netezza, IBM DB2.
- Model export: Java, SQL Code, SAS code, PMML.
- Model deployment/management: Through KXEN factory.
- Other: transforms text into root word coding, supports ensemble modeling.

#### IBM (SPSS)

Company Name: IBM  
Public

Website: [www.ibm.com/spss](http://www.ibm.com/spss)



Now with tighter linkage/integration with the IBM stack, SPSS continues to score high in ease of use, visualization, and ease of deployment.

When IBM acquired SPSS in 2009, its goal was to help grow its predictive analytics portfolio, and to integrate predictive analytics capabilities throughout its software, hardware, and services product portfolios, enterprise-wide— a strategic focus area for the company. SPSS, with its more than forty years of experience in predictive analytics has provided IBM with needed expertise in this area and IBM has been successfully advancing SPSS as an enterprise class solution. While it is part of the larger IBM, SPSS is still focused on making analytics easy to use as part of a business process. In other words, SPSS is focused on taking analytics to all kinds of users via easy-to-use interfaces and automation of model development. It is also providing the tools and techniques to help integrate these models into the decision management process. For example, its Decision Management solution combines predictive analytics with business rules, and “what if scenarios” to help to create and deploy predictive models in business processes such as predicting fraud or churn.



**HURWITZ  
& ASSOCIATES**  
Insight to Action

*Now with tighter linkage/ integration with the IBM stack, SPSS continues to score high in ease of use, visualization, and ease of deployment.*

SPSS supports a range of customers, from those who are using it on their desktop to those who are building much bigger server-based solutions with predictive analytics as part of a business process. Across the board, customers responding to the Victory Index survey were impressed with the breadth and depth of the software as well as its ease of use and the ease of deployment of the products. SPSS received the highest score among larger predictive analytics vendors in the ease of use category. Customers also liked the fact that SPSS provides “usable output” meaning that executives could understand the results. Additionally, SPSS got exceptionally high marks for rolling out new functionality with each release and for listening to their customers. Customers felt that SPSS could be used effectively by any size organization and provide sustained value.

*Customers also liked the fact that SPSS provides “usable output” meaning that executives could understand the results.*

### SPSS Overview

Products	<ul style="list-style-type: none"> <li>• IBM SPSS Modeler Premium v. 14.1: Data mining and text mining workbench.</li> <li>• IBM SPSS Statistics v. 19: Data and predictive analytics software.</li> <li>• IBM SPSS Data Collection v. 6: Used by survey and market researchers to achieve deeper understanding of people’s attitudes, preferences and behavior.</li> <li>• IBM SPSS Collaboration and Deployment Services v. 4.2: Used to manage, monitor, and deploy analytical assets.</li> <li>• IBM SPSS Decision Management v. 6: Decision-process framework that uses predictive analytics and business rules to operationalize decisions.</li> <li>• IBM Cognos Consumer Insights: Used by marketing professionals to understand customer demands and opinions expressed through social media.</li> </ul>
Target User/ company size	<p>SPSS targets different classes of users for its products. These include:</p> <ul style="list-style-type: none"> <li>• IBM SPSS Data Collection: business users, survey/market researchers, and information technology administrators.</li> <li>• IBM SPSS Statistics: business users, analysts and statistical programmers.</li> <li>• IBM SPSS Modeler: business users, analysts and data miners.</li> <li>• IBM SPSS Collaboration and Deployment Services: business users, analysts, data miners, database administrators and information technology administrators.</li> <li>• IBM SPSS Decision Management: business users and analysts.</li> </ul>
Verticals supported	<p>SPSS looks across three “pillars” of predictive analytics. These are customer analytics, operational analytics, and threat and risk analytics. Within each of these SPSS can support various verticals including: aerospace and defense, automotive, banking, chemicals and petroleum, communications, construction, consumer products, education, electronics, energy and utilities, financial markets, forest and paper, government, healthcare, insurance,</p>



	life sciences, media and entertainment, metals and mining, retail, telecommunications, travel and transportation
Horizontal solutions	<p>Via the three pillars, IBM supports the following horizontal solutions:</p> <ul style="list-style-type: none"> <li>• Customer Analytics: Up-sell/cross-sell, targeting, market basket analysis, churn prevention, customer segmentation, brand monitoring.</li> <li>• Operational Analytics: Security, predictive maintenance, assortment planning, condition monitoring, reverse logistics, allocation management.</li> <li>• Threat and Risk Assessment: claims fraud, credit-card fraud, insider threat, signal analysis, cyber security.</li> </ul>
How offered	License and SaaS model
Differentiator(s)	<ul style="list-style-type: none"> <li>• Ease of use of the interface.</li> <li>• Operationalizing predictive analytics in business processes.</li> <li>• In database scoring as well as SQL pushback.</li> </ul>
Platforms supported	<ul style="list-style-type: none"> <li>• Client: Windows, Mac (IBM SPSS Statistics only), Linux (IBM SPSS Statistics only)</li> <li>• Server: Windows, Linux, AIX, HP UX, Solaris, Linux for System Z</li> </ul>
Predictive Analysis Families	Association, classification, regression, time series, and more

*IBM SPSS Modeler offers numerous nodes to merge, join, filter, and transform data for modeling.*

### Key Features

- ETL: IBM SPSS Statistics and IBM SPSS Modeler provide a variety of pure ETL tools to access data, transform data, and load data back into a database or location. For example, IBM SPSS Statistics offers a database wizard that enables extraction of data. IBM SPSS Modeler offers numerous nodes to merge, join, filter, and transform data for modeling. Both provide the ability to write output and results into target files and databases. In addition, both of these products offer a set of automated and manual data preparation procedures that enable users to screen data for problematic fields and identify data that is not useful during analysis. IBM SPSS Statistics also enables analysts to generate and implement statistical transformations that improve the analysis of skewed data sets, and these transformations can be used within IBM SPSS Modeler to improve the development of predictive models.
- Data formats/databases supported: IBM supports a variety of data sources, including DB2 UDB for Windows/UNIX version, SQL Server, Oracle, Teradata, Netezza, DB2/400 / DB2 for zOS, Sybase IQ, HP Neoview, MySQL AB Enterprise Edition, Variable files, Flat files, SPSS Statistics (.SAV) files, SAS files, User input data, XML, HTML, RSS Feeds, PDF, MS-Excel, RTF, SPSS Data Collection, IBM Cognos BI.
- In-database mining: Yes (see differentiators, above).
- Model export: PMML 4.

- Model deployment/management: For organizations that want to leverage predictive analytics across the enterprise, IBM SPSS Statistics and IBM SPSS Modeler models can be used with IBM SPSS Collaboration and Deployment services for deployment into operational systems. Within a department or functional area, both IBM SPSS Statistics and IBM SPSS Modeler are used to generate models for local use. Both products have the ability to export models as XML (PMML), and this XML can be used in any PMML-enabled environment to score other data sets, transform other data sets and the like. IBM SPSS Decision Management applications provide business users with the ability to deploy the combination of predictive models and business rules into operational decision-making processes.
- Other: supports unstructured data, ensemble modeling, automated data preparation, automated model development.

## Megaputer

Company: Megaputer Intelligence  
Private

Website: [www.megaputer.com](http://www.megaputer.com)

Megaputer shines at the intersection of structured and unstructured data analysis.



*Customers agree that where Megaputer stands out from other predictive analytics vendors is when it brings text and structured data together.*

This US based private company, founded in 1997 is probably best known for its strength in providing an integrated approach to analyzing structured and unstructured data. Megaputer was launched as a data mining company in 1997. In 1998, it acquired exclusive rights to distribute a third party text analytics application called TextAnalyst, which was one of the first text analysis systems on the market, and then developed text analytics capabilities for its flagship product PolyAnalyst. Given its heritage in both data mining and text analytics, the company frequently deploys solutions that bring together structured and unstructured data. Megaputer often works with its customers to build solutions, using its core product named PolyAnalyst, that address customer challenges and incorporates this solution in a client's standard business process. The company also provides its solutions in a SaaS model (using both structured and unstructured data), because it believes that sometimes companies don't need an entire solution, just regular output. This is especially true for companies that don't have quantitative analysts on staff. Megaputer also offers its own professional services to help build out solutions and has recently entered into a partnership with Ernst & Young. The company prides itself on its customer support, solving its customer's challenges in the most time and price efficient manner.

Customers agree that where Megaputer stands out from other predictive analytics vendors is when it brings text and structured data together. Examples include call center analysis or survey data. Megaputer also got high marks for



HURWITZ  
& ASSOCIATES  
Insight to Action

some of its specific functionality including its machine learning and link analysis. Customers also cited the firm’s ease of use in constructing models and the value for the price point.

### Megaputer Overview

Products	<ul style="list-style-type: none"> <li>• PolyAnalyst v6: provides both data and text mining delivered through one interface. PolyAnalyst offers built-in OLAP features and a report generator for creating graphical browser-based reports that summarize the results of the analysis for non-technical users.</li> <li>• PolyAnalyst Web Reports v6: delivers summaries of data analysis results to business users. Business users can view graphical outputs and interact with web reports through a standard web browser – no client software installation is necessary.</li> </ul>
Target User/ company size	Business and data analysts at primarily large and mid sized companies
Verticals supported	<ul style="list-style-type: none"> <li>• e-Discovery</li> <li>• Insurance (e.g. claims management)</li> <li>• Pharmaceutical (e.g. research data analysis)</li> <li>• Manufacturing; Manufacturers warranty claims analysis</li> </ul>
Horizontal solutions	Fraud detection, voice of the customer analysis, market research data analysis; social media data analysis; pharmaceutical research; subrogation prediction; and warranty and repair data analysis
How offered	On premises license and SaaS
Differentiator(s)	<ul style="list-style-type: none"> <li>• Has a long history of performing structured/unstructured analysis using tools it has internally developed.</li> <li>• Built-in OLAP functionality.</li> </ul>
Platforms supported	Microsoft Windows XP SP3/2003/Vista/2008/7
Predictive Analysis Families	Association, classification, regression, simulation, time series, social network analysis, distribution analysis, others

### Key Features

- ETL: Supports loading data from disparate data sources including; email archives and Exchange Server, websites, RSS feeds, documents from a file system (PDF, HTML, MS Word, PPT formats) – in addition to databases, spreadsheets and flat files. Supports creating new attributes, changing values, changing types, filtering, aggregation, consolidation, sampling, expanding and reconstructing transactions. Supports cleansing through anomaly detection, language identification, automated spelling correction, term replacement.
- Data formats/databases supported: Data sources: ODBC, OLEDB, CSV, Excel, Web, File system. Document Formats: PDF, ASCII, HTML, MS Word, MS RTF,

*Customers also cited [Megaputer’s] ease of use in constructing models and the value for the price point.*



RSS Feeds. Character Formats: ASCII, Latin-1, Double-byte, UTF. Also web data sources, direct email import, email server data source, social media sources, dataset links, and federated search.

- In-Database Mining: No. Had this but discontinued it for the present.
- Model export: C, SQL, and plan to start supporting PMML v.3.
- Model deployment/management: PolyAnalyst can store models to be deployed for scoring new data on a different system. The management of the models is done via structured file folders. Scheduling the execution of models and generating condition based alerts are available.
- Other: supports unstructured data; out of the box medical insurance fraud functionality called Medicop.

## Pegasystems

Company: Pegasystems  
Public

Website: [www.pegasystems.com](http://www.pegasystems.com)



Supports large companies focused on customer service and improving lifetime value of a customer; targeted to business users.

Pegasystems, headquartered in Cambridge Massachusetts is focused on automating customer-centric business processes in order to improve the lifetime value of a customer. Its vision is to make predictive and adaptive analytics as easy as possible for business people to use in order to drive measurable improvements in lifetime customer value. This requires that models optimize what the company refers to as "Next Best Actions." This is an action, for example, that a call center agent might take to upsell or cross-sell a customer. Pegasystems provides a unified offering that includes business rules and decision trees to help organizations plan, manage, and automate their policies and rules, their cross silo work processes, as well as their customer centric multi-channel interactions. Its goal is to allow business users to create intent-led predictive models that they can use to guide behavior. It offers pre-built models for business users.

The company is highly customer focused. Customers reported that they like the Pegasystems' technology and tools and said they get a good ROI from the product.

### Pegasystems Overview

Products	Pegasystems provides a unified platform, with out of the box solutions for process management, case management and multi-channel CRM (including unified inbound and out bound marketing automation, social media monitoring and response, web-based self-service and call center) that
----------	--



*[Pegasystems'] vision is to make predictive and adaptive analytics as easy as possible for business people to use in order to drive measurable improvements in lifetime customer value.*

includes predictive and adaptive analytics together with process and rules. The specific predictive and adaptive capabilities include:

- Decision Strategy Manager (DSM): Provides core decision management components including rule types to define and execute decision strategies, a UI to configure and sequence Next-Best-Action (NBA) logic and rule forms; the deployment of NBAs, and to import predictive models.
- Adaptive Conversation Advisor (ACA): UI for presenting and negotiating offers as well as to configure NBA user interface components. The UI is embedded into Pegasystems and can be embedded into 3rd-party call center applications.
- Adaptive Decision Manager (ADM): Used to configure, deploy and execute adaptive (self-learning) models. ADM operates as one component of DSM, and can be configured with other elements to define NBAs.
- Predictive Analytics Director (PAD): Used to define and build predictive strategies using historical data from Pegasystems and other data sources. Fit-for-purpose model factory versus model laboratory to meet predictive analytics demand for data-driven strategies.
- Visual Business Director (VBD): Used to measure, control, and optimize strategies using both historical and current data; visualize current performance and simulate strategy changes.
- CDM Bridge: Provides connection from CDM (legacy version of DSM) to Pegasystems. Also provides NBA decisions in Pegasystems from existing CDM system.

*A unique feature in [Pegasystems] offering is its ability to allow customers to move to the cloud and back to on-premises and vice versa without a penalty.*

Target User/ company size	Business users at Fortune 500 companies and public sector agencies
Verticals supported	Automotive, communications, energy (Oil & Gas), financial services, insurance, healthcare, life sciences, public sector
Horizontal solutions	CRM, BPM and case management, decision management & predictive analytics Note: Pega predictive models are used in critical business operations, like credit risk assessment, and have also been featured in scientific publications in areas varying from CRM, to clinical studies, and even hydrodynamics.
How offered	License model and pricing are offered on a per user basis. The company also offers provisioning in a private or public cloud. A unique feature in its offering is its ability to allow customers to move to the cloud and back to on-premises and vice versa without a penalty.
Differentiator(s)	<ul style="list-style-type: none"> <li>• Unified platform for customer.</li> <li>• Scale of use: Supports hundreds of predictive models (traditional and self-learning) to be executed in parallel.</li> </ul>



Platforms supported	Oracle, Websphere, JBoss, AIX, Z/OS, HP
Predictive Analysis Families	Decision trees, regression models, and score cards, proprietary non-linear models

### Key Features

- ETL: Data Preparation Director for business users who need to manipulate data without IT involvement (not a key feature).
- Data formats/databases supported: flat files, JDBC, Oracle, DB2, Teradata.
- In-database mining: not supported (No need and no gain for actionable models).
- Model export: Models are built as part of the company’s unified platform so this is not required. However, 3rd party models in PMML format can be imported.
- Model deployment: Deployment is unified since the platform is unified. Predictive models, stand-alone, or connected to decisions and processes, are instantly deployed. Many hundreds of predictive models can be executed in real time in support of a single decision (i.e. what’s the next best action for this customer).
- Other: The scale of use. Supports hundreds of predictive models (traditional and self-learning) to be executed in parallel; Fit-for-purpose: the predictive models come with meta-data that automates downstream (self) monitoring and auditing. The models can compare data they are exposed to, to the data used for their development, and flag discrepancies that over time might lead to inaccurate predictions.

*[Portrait Software] supports B2C marketers who are interested in using advanced and predictive analytics to understand and predict customer behavior and put that information to use in their marketing campaigns.*

### Pitney Bowes

Company: Pitney Bowes  
Public



Website: [www.portraitsoftware.com](http://www.portraitsoftware.com)

Supports B2C marketers who are interested in using advanced and predictive analytics to understand and predict customer behavior and put that information to use in their marketing campaigns.

Pitney Bowes acquired Portrait Software, a twenty-year old customer interaction and analytics company in 2010, in order to round out its own product offering in its Business Insight Software division. The vision of the combined group is to enable “lifetime customer relationships.” Portrait Software sits alongside other statistical products in an organization and gives marketers and others concerned with customer interaction the tools to analyze and predict behavior. It does this by providing a series of product modules that address different customer

relationship pain points. A unique aspect of the solution is what the company refers to as “Portrait Uplift” which is a module designed to help change future behavior; in other words, to predict which customers can be persuaded in certain campaigns. For example, in a churn analysis this would mean analyzing data to determine what would make groups of customers likely to stay. Additionally, according to Pitney Bowes, the shelf-life of customer interaction models is short. This means that there are always new models to be built. The software was designed for an organization that would need to build lots of models, in a kind of “model factory” mentality. Since the models need to be easy enough for non-statisticians to use, Pitney Bowes has built a lot of automation into its model-building process.

Customers in the Victory Index survey liked the fact that the tools in the suite are easy to use and understand and help them manage the customer interaction life cycle through a set of integrated modules. Pitney Bowes also got high marks for time to value as customers felt that the automation in data preparation and modeling made it easier to build models faster. Pitney Bowes did well among marketers in the business value category, as well.

*... according to Pitney Bowes, the shelf-life of customer interaction models is short. This means that there are always new models to be built. The software was designed for an organization that would need to build lots of models, in a kind of “model factory” mentality.*

### **Pitney Bowes/Portrait Software Overview**

Products	<p>Insight Tools:</p> <ul style="list-style-type: none"> <li>• Portrait Miner: A visual predictive analytics environment.</li> <li>• Portrait Uplift: Proprietary algorithms to understand how customer actions changed as a result of marketing activities, i.e. to predict “persuadability” only those that will respond to your message.</li> <li>• Portrait Optimizer: A contact optimization solution used to allocate marketing spend.</li> <li>• Portrait Self-Service Analytics: A browser based tool for the business user for customer data exploration.</li> </ul> <p>Action Tools:</p> <ul style="list-style-type: none"> <li>• Portrait Foundation: Customer process management work include: flow engine for fulfillment and provisioning.</li> <li>• Portrait Dialogue: Outbound campaign management.</li> <li>• Portrait Interaction Optimizer: inbound interaction management based on real-time decisioning and scoring.</li> </ul>
Target User/ company size	Companies focused on the B2C market, both mid market and larger companies
Verticals supported	Financial services, insurance, telecommunications, retail, publishing, travel and leisure, government
Horizontal solutions	A marketing focused solution dealing with retention, churn, cross selling, and marketing metrics
How offered	License and 1-3 year subscription
Differentiator(s)	<ul style="list-style-type: none"> <li>• Ease of use via automation of model building.</li> <li>• “Uplift modeling” (see above.)</li> <li>• Integrated solution from analysis to campaign management.</li> </ul>



Platforms supported	Client-server with Solaris or Windows servers (32 or native 64 bit) and Windows clients (XP, Vista, Windows 7) or desktop on Vista or Windows 7 (32 or 64 bit)
Predictive Analysis Families	Decision trees, linear & logistic regression (presented as "scorecards", a form of generalized additive model), clustering, association, & sequencing rule discovery, uplift modeling.

### Key Features

- ETL: provides capabilities for importing and exporting data from source systems (RDBMS, flat files, third-party formats such as SAS and SPSS) and for creating a customer view used for analysis (sort, join, merge, aggregate, transform). Includes a library of functions for creating behavioral attributes from transaction and event data. These results can be deployed as data (lists, scores) directly back to source systems, or as automatically generated code representing the segment rules or scoring formulae (e.g. SQL, SAS), or for real-time scoring in Portrait Interaction Optimizer. Pitney Bowes also now bundles the Sagent DataFlow tool to provide a broader ETL and reporting capability and is actively developing several tighter integration points.
- Data formats/databases supported: Oracle, MS-SQL, Teradata, any ODBC source, plus a broad range of fixed and delimited flat-files and numerous third-party formats.
- In-database mining: No. The company takes the position that this is a good, but impractical idea.
- Model export: SQL, SAS, PMML.
- Model deployment/ model management: The focus of Portrait Software is to enable business users to create many models with a short shelf life. The company is currently creating a collaborative and mass model management system that will provide version control and workforce sign off.
- Other: is directly integrated with other Pitney Bowes customer interaction software. For example, contact optimization is directly integrated to campaign planning software.

### Rapid-I

Company: Rapid-I  
Private

Website: [www.rapid-i.com](http://www.rapid-i.com)

Supports companies with analytics expertise in-house, looking for a cost effective solution



*Rapid-I ... was founded on the principles of "flexibility and freedom" by the founders of the open-source machine learning, data mining, and predictive analytics project RapidMiner.*

Predictive analytics vendor Rapid-I, headquartered in Dortmund, Germany, was founded on the principles of "flexibility and freedom" by the founders of the open-source machine learning, data mining, and predictive analytics project



**HURWITZ  
& ASSOCIATES**  
Insight to Action

RapidMiner (formerly YALE (Yet Another Learning Environment)). The company has built its software “on top of” its open source solution RapidMiner (formerly YALE) and provides extensions that allow users to also integrate the two open-source analytics solutions Weka and R. In this way, the company feels that RapidMiner and RapidAnalytics provide the power of the three most widely used open source predictive analytics solutions on the market.

**Rapid-I Overview**

Products	<ul style="list-style-type: none"> <li>• RapidMiner (version 5): Open source data mining software previously known as Yale with extension for integrating R and Weka. Available as a stand-alone analysis application or as an engine to plug into other applications. Includes 100 plus modeling techniques and modules for all steps of the analysis process from data extraction and transformation (ETL), analysis, modeling, evaluation, visualization, reporting, prediction, and deployment.</li> <li>• RapidAnalytics (version 1): Server based version of RapidMiner that provides user management and user rights management, model management, automated process scheduling, collaboration support with shared resources, web services, and reporting.</li> <li>• RapidNet (version 1): Provides a visual interface to visualize and interactively explore networked and relational data.</li> <li>• RapidSentryzer (version 2): Social media monitoring and analysis solution for unstructured text analytics utilizing social media and other online sources as input.</li> <li>• RapidDoc (version 1): Web services based solution for document classification.</li> </ul>
Target User/ company size	<p>All companies, all sizes, but mostly trained statisticians or other quantitative analysts. Specifically, RapidMiner:</p> <ul style="list-style-type: none"> <li>• Predictive analytics process designers</li> <li>• Statisticians/mathematicians/data mining experts</li> <li>• IT departments</li> <li>• Savy analysts in other departments including marketing, sales, market research, controlling, credit risk, fraud detection and prevention, manufacturing, finance, etc.</li> <li>• Recipients and readers of RapidMiner reports: in all departments</li> </ul> <p>RapidAnalytics: analysts, report designers, business users  RapidNet: analysts and business users  RapidSentryzer: analysts and business users  RapidDoc: analysts and IT</p>
Verticals supported	<p>Automotive, aviation and transportation, biotech, electronics, financial, IT, insurance, manufacturing, market research, media, pharmaceutical, retail, telecommunications, utilities, government</p>

*...[Rapid-I]... feels that RapidMiner and RapidAnalytics provide the power of the three most widely used open source predictive analytics solutions on the market.*



Horizontal solutions	Frequent application areas of the Rapid-I solutions: direct marketing optimization, market basket analysis, product recommender systems, up- and cross-selling, churn prevention, increasing customer retention and loyalty, credit risk scoring and prevention, fraud detection and prevention, online media monitoring, online market research, social media monitoring, customer insight generation, market intelligence, competitive intelligence, product innovation and public relations optimization, manufacturing machine failure forecasting and prevention, predictive maintenance, optimization of production and manufacturing processes, forecasting of sales volumes, energy consumption, utility prices, exchanges rates, stock market prices, web analytics, click path analysis, web site optimization, best ad selection, link recommendations web mining and information extraction and aggregation from web pages and online media
How offered	On-premises and SaaS
Differentiator(s)	<ul style="list-style-type: none"> <li>• Open source platform.</li> <li>• Provides consulting, training, support, customization, and custom made extensions.</li> </ul>
Platforms supported	Microsoft Windows (XP, Vista, Windows 7), Linux, Apple OS X, SUN Solaris, IBM AIX, HP UX, Unix, and all other operating systems with Java 1.6 or higher available
Predictive Analysis Families	Association, classification, clustering, regression, information extraction, time series, text mining, audio mining, image mining, and more

*Rapid-I features include: audio and image analysis, text analytics, and integration to R using RapidMiner/RapidAnalytics extensions.*

### Key Features

- ETL: RapidMiner and RapidAnalytics have built-in ETL and data preparation functionality. RapidSentry uses built-in web crawlers and automated information extraction from web pages. RapidDoc uses a web service based data import.
- Data formats/databases supported: Excel, Access, Oracle, IBM DB2, Microsoft SQL, Sybase, Ingres, MySQL, Postgres, Teradata, JDBC/ODBC, SPSS, dBase, text files, audio files and more.
- In-Database Mining: RapidAnalytics Enterprise Edition offers in-database mining for all standard SQL databases including for example Ingres VectorWise, MySQL, Oracle, etc.
- Model export: PMML, RML (RapidMiner XML Format for Models and Processes).
- Model deployment: within RapidMiner (interactively or via command line or via Java library) or on RapidAnalytics servers or as RapidAnalytics web services or via PMML in third party tools.
- Other: audio and image analysis, text analytics, also integrates to R using RapidMiner/RapidAnalytics extensions.



## Revolution Analytics

Company Name: Revolution Analytics  
Private



Website: [www.revolutionanalytics.com](http://www.revolutionanalytics.com)

Supports data analysts and data scientists looking for leading edge open source software methods for analyzing big data or performing near real time predictive analytics

*[Revolution Analytics]... is providing functionality to handle the performance problems of base R and to handle very large amounts of data (also known as big data) and to quickly process it using out-of-memory parallel algorithms.*

Revolution Analytics, a US based company formed in 2007, is capitalizing on the popularity of the open source R statistical computing language which, according to the company, is now used by several million analysts worldwide. With co-founder Norman Nie (the former chairman and CEO of SPSS) at the helm, Revolution Analytics hopes to become the "R Company" and, according to the company, make Revolution a "world-leading predictive analytics company while continuously improving and contributing to the R experience for the open source community." As a commercial provider of the R statistical computing language, the company plans to provide compelling features at about half the cost of leading statistical packages. Its strategy is to build value added applications around R in order to make it more user friendly and robust. To that end, Revolution Analytics is concentrating on three areas. First, it is providing an easy-to-use interface (a visual GUI) that will make predictive analytics more accessible to the novice user. Second, it is providing functionality to handle the performance problems of base R and to handle very large amounts of data (also known as big data) and to quickly process it using out-of-memory parallel algorithms. Finally, it is providing the features, services, and technical support, and training to make the package robust enough for enterprise use.

In the Victory Index survey, customers said they liked the performance of the product, which the company benchmarked to be 5-8 times faster than R.

### Revolution Analytics Overview

Products	<p>Revolution R Enterprise v 4.3: Production-grade analytics software built upon open source R statistics language, which includes the modules:</p> <ul style="list-style-type: none"> <li>• RevoScaleR: Big Data statistical solution.</li> <li>• RevoDeployR: Enterprise scale integration, administration, and deployment platform for R.</li> </ul>
Target User/ company size	<p>Revolution is targeting companies of all sizes. It sees two kinds of users: 1) analysts, statisticians, data scientists looking for cost effective alternatives to traditional statistical packages and 2) companies that are analytically aware, that may not have quantitative talent, and are priced out of market with traditional software. These might include those companies who are converting software or startups.</p>



Verticals supported	Finance, life sciences, manufacturing, retail telecom, social media, government
Horizontal solutions	Statistical data analysis, predictive analytics, data mining, machine learning
How offered	On-premises subscription based
Differentiator(s)	<ul style="list-style-type: none"> <li>• Cost</li> <li>• Big Data focused</li> <li>• Open-source community</li> </ul>
Platforms supported	Microsoft Windows; Red Hat Enterprise Linux
Predictive Analysis Families	Association, classification, information extraction, regression, optimization, simulation, time series and all capabilities included in R and packages from CRAN

### Key Features

- ETL: Data Step with R functionality, interfaces with third-party ETL tools (e.g. Talend).
- Data formats/databases supported: ASCII, SAS, SPSS, ODBC/JDBC, Hadoop (HDFS, Hbase, Streaming), Netezza, XDF, plus custom interfaces via R programming/packages.
- In-database mining: Planned for Hadoop and Netezza.
- Model export: R, PMML.
- Model deployment/management: Deployment via Server edition of Revolution R Enterprise. RevoDeployR Web services platform for integration and management.
- Other: supports unstructured data analysis, ensemble modeling, social network analysis, and many other methods from open-source R packages.
- User interfaces: R language IDE, Web-based GUI available 2012.

### SAP

Company Name: SAP  
Public

Website: [www.sap.com](http://www.sap.com)

If SAP or SAP BusinessObjects is already deployed in an organization with highly skilled employees, then it is worth considering SAP Predictive Workbench.



*Revolution Analytics  
Differentiator(s)*

- Cost
- Big Data focused
- Open-source community

Historically, predictive analytics has not been at the forefront of SAP's BI strategy. SAP's philosophy has been to offer the benefits of predictive analytics to all



**HURWITZ  
& ASSOCIATES**  
Insight to Action

of its customers, as an extension of BI. It does this, in part, through an OEM relationship with IBM SPSS modeler. It calls this product the SAP BusinessObjects Predictive Workbench. Predictive Workbench works together with SAP BusinessObjects BI. It is basically targeted at customers with statistical know-how who have already deployed SAP. Additionally, while SAP has not been aggressive in seeking the predictive analytics market, per se, it recently announced that predictive analytics would play a role in its new road map for in-memory computing (via the HANA in-Memory appliance) which will deal with processing massive amounts of data in real time. It will also be part of some of its business process products. So, we may be hearing more from SAP in the future regarding predictive analytics.

Customers responding to the Victory Index survey had a mixed reaction to SAP predictive analytics. Often, they reported choosing SAP predictive analytics because they are a SAP/Business Objects shop, even if the product was not the best technical choice. While some respondents in the survey had complaints about customer support, overall, companies have seen good business and technical value from the Predictive Workbench. In general, overall user satisfaction has been above average.

### SAP Overview

Products	• SAP BusinessObjects Predictive Workbench
Target User/ company size	Statistician or data miner in large organization where the organization is an SAP customer
Verticals supported	None directly targeted. The most common are telecommunications, banking, insurance, retail, utilities, and government.
Horizontal solutions	These are supported via native capabilities in SAP's CRM solution.
How offered	On-premises licensing
Differentiator(s)	• Integrated with SAP and SAP BusinessObjects data sources
Platforms supported	• Windows 32 bit
Predictive Analysis Families	<ul style="list-style-type: none"> <li>• Association: Apriori, GRI, Carma, Sequence.</li> <li>• Classification: CART, QUEST, CHAID, C5.0, Regression, Neural Networks, Logistic Regression, SVM, KNN.</li> <li>• Information extraction (i.e. for unstructured data): Available in SAP BusinessObjects Text Analysis.</li> <li>• Regression: MLR, General Linear Model, Logistic, Cox.</li> <li>• Optimization: Appropriate algorithms have optimization features.</li> <li>• Simulation: Available in SAP BusinessObjects Dashboards.</li> <li>• Time Series: Exponential Smoothing, ARIMA, Regression.</li> <li>• Other: PCA, discriminate analysis, segmentation, K means, hierarchical agglomeration, Kohonen SOMs.</li> </ul>

*...while SAP has not been aggressive in seeking the predictive analytics market, per se, it recently announced that predictive analytics would play a role in its new road map for in-memory computing via the HANA in-Memory appliance...*



## Key Features

- ETL: SAP has a range of ETL tools that can be used with the Predictive Workbench. Additionally, the Predictive Workbench contains data preparation functions specific for analysis, such as binning, normalization etc.
- Data formats/databases supported: Any ODBC source, including Microsoft SQL Server, DB2, Oracle, BOBJ Universes, and SAP BW. Also supports CSV, XLS, and others.
- In-database mining: Support for in-database algorithms in Oracle 11g, IBM DB2, and Microsoft Analysis Services. In addition, algorithms in HANA (the new SAP In-Memory Appliance).
- Model export: C, PMML (Version 3.2), SQL, Text, HTML.
- Model deployment/management: Predictive Workbench is a desktop tool so individual users manage it. Predictive Workbench Server installs a Windows Service which is managed as a Windows Service. The interactive modeling process can be scripted and run 'in batch'. Scored or predicted data can be exported. Models can be deployed as stated above.
- Other: The Predictive Workbench supports ensemble modeling with auto nodes for the general classes of problem.

*SAS, with its broad and deep predictive analytics product set, has consistently been a leader in the advanced analytics market. Its dual goal is to meet customer needs, industry by industry and to provide continuous innovation around analytics.*

## SAS

Company Name: SAS  
Private

Website: [www.sas.com](http://www.sas.com)



With strong depth and breath of functionality across the predictive analytics lifecycle, SAS supports customers with increasing volumes of complex data across a wide variety of vertical industries.

SAS, with its broad and deep predictive analytics product set, has consistently been a leader in the advanced analytics market. According to Dr. Jim Goodnight, CEO, "No problem should be too difficult for SAS to solve." The company's objective is to help organizations anticipate opportunity, empower action, and drive impact. Its dual goal is to meet customer needs, industry by industry and to provide continuous innovation around analytics. As such, the company's strategy has evolved from providing a robust set of analytics capabilities to include a strong focus on industry solutions to solve business problems. SAS now divides its application space into two parts: customer applications and industry focused models. If you ask anyone from SAS a question about SAS product(s) he or she will likely ask about the nature of the business problem that you are trying to solve. SAS views predictive analytics as a corporate asset with its own lifecycle and has built the software to support this view. SAS is also innovating around high performance analytics to address big data and complex business problems.



**HURWITZ  
& ASSOCIATES**  
Insight to Action

While some respondents to the Victory Index survey complained about the price of the product, SAS still received the highest overall customer satisfaction ratings among predictive analytics vendors in the Victory Index. Customers liked the fact that the company listens to them regarding product updates. In fact, SAS puts a large percent of its profits back into R&D to help satisfy its customers. SAS also got extremely high marks among customers for its technical excellence, the breadth and depth of its functionality, and the comprehensiveness and scalability of its solutions.

## SAS Overview

### Products

- SAS Enterprise Miner 7.1: Data mining process to create predictive and descriptive models based on analysis of vast amounts of data from across the enterprise.
- SAS Model Manager 3.1: A patented, secure model repository complemented by a metadata structure and project templates that is used to create, manage, monitor and deploy analytical models.
- JMP Pro 9.0: Provides data visualization, exploratory data analysis and data mining in an in-memory, desktop-based environment. It includes advanced statistical techniques as well as core data mining functionality. JMP Pro is integrated with R.
- SAS Scoring Accelerator for Databases: Used to publish scoring models created in SAS Enterprise Miner into database-specific functions to be deployed and then executed directly within databases listed below.
- SAS Analytics Accelerator 2.1 for Teradata: Supports generating descriptive statistics, statistical transformations, dimension reduction, variable selection, time series, logistic regression, multiple linear regression and scoring of these SAS models directly in Teradata without requiring data extraction.
- Credit Scoring for SAS Enterprise Miner 7.1: Is designed for the in-house creation of credit scorecards that assist in evaluating the credit worthiness of customers.
- SAS/ETS 9.3 (Econometrics and Time Series): Provides time series, forecasting and econometric techniques that enable modeling, forecasting and simulation of business processes.
- SAS/STAT 9.3: Provides a range of statistics from traditional statistical analysis of variance and predictive modeling to exact methods and statistical visualization techniques.
- SAS/IML 9.3: A matrix programming language for programmers, statisticians, researchers and high-end analysts. The SAS/IML Studio interface provides interactive programming and exploratory data analysis. Also supports R code.
- SAS Forecast Server 4.1: Generates large quantities of high-quality forecasts quickly and automatically. SAS Forecast

*Customers liked the fact that [SAS] listens to them regarding product updates. In fact, SAS puts a large percent of its profits back into R&D to help satisfy its customers.*



- Server bundles the SAS High-Performance Forecasting engine with the SAS Forecast Studio GUI.
- SAS/OR 9.3: Provides optimization, simulation, and project scheduling techniques to identify the actions that will produce the best results. Key components include OPTMODEL (optimization modeling) and SAS Simulation Studio (discrete event simulation).
  - SAS Enterprise Miner for Desktop 7.1: A complete data-mining workbench that runs entirely within the confines of a Windows PC. It is designed for quantitative analysts in small to medium-sized organizations, or those who work independently in departments.
  - SAS Rapid Predictive Modeler 1.3: Designed for business analysts and subject-matter experts with limited statistical expertise to generate predictive models.

Target User/ company size	SAS Predictive Analytics products are targeted to organizations of all sizes, including small, mid-market and large. Potential users might include: <ul style="list-style-type: none"> <li>• Analytics professionals and team leads: statisticians, modelers, data miners, researchers, scientists, engineers, etc.</li> <li>• Business Analysts, customer/marketing insight analysts, risk analysts, operations research analysts, etc.</li> <li>• IT/Database Staff: Chief Scoring Officer, SAS programmers, ETL developers, database professionals, etc.</li> <li>• Management Team: Chief Analytics Officer, VP or Director of Analytics, Data Mining Team Lead, etc.</li> </ul>
Verticals supported	Casinos, communications, education, financial services, government, health insurance, health care providers, hotels, insurance, life sciences, manufacturing, media, oil & gas, retail, travel and transportation, and utilities
Horizontal solutions	Customer intelligence, data management, fraud prevention and detection, risk management, IT management, performance management, supply chain intelligence, human capital management and sustainability management and compliance
How offered	SAS offers a variety of licensing options to suit the needs of enterprises, government entities, non-profits, educational institutions and individuals. Licensing options include: enterprise licensing, named-user, server capacity-based/CPU-based, value-based, home-office, per-use/month and yearly subscription.
Differentiator(s)	<ul style="list-style-type: none"> <li>• Completeness of solution: Users can develop, deploy, and manage models against structured, semi-structured, and unstructured data.</li> <li>• Technical breadth and depth.</li> <li>• Data preparation capabilities.</li> </ul>

*SAS offers a variety of licensing options to suit the needs of enterprises, government entities, non-profits, educational institutions and individuals.*



Platforms supported	<ul style="list-style-type: none"> <li>• Client environment: Linux; Microsoft Windows environments including Windows Vista and Windows 7</li> <li>• Server environment: AIX; HP-UX; Linux; Solaris; OpenVMS; and Microsoft: Windows XP Professional, Windows Server 2003 and 2008, Windows Vista, and Windows 7</li> <li>• Server &amp; Desktop Virtualization environment: VMware GSX Server, MSFT Windows Terminal Serv., Citrix Pres. Server, Citrix XenApp</li> <li>• JMP Pro is also available on Apple OS X</li> <li>• Mobile Devices: iPad, iPhone, Blackberry and Windows Mobile Devices are supported</li> <li>• Select SAS products (e.g. SAS Enterprise Miner) are automatically tailored for job submission in a grid-computing environment</li> </ul>
Predictive Analysis Families	Association, classification, information extraction, regression, optimization, simulation, time series, more

### Key Features

- ETL: SAS offers a variety of ways to handle data integration and data preparation tasks. Core analytic data preparation tasks including collecting appropriate data, sampling, data partitioning, transposing, appending, removing duplicate records, selecting, editing and filtering variables, creating dummy variables, optimal variable binning, string matching function, statistical outlier detection, summarization, aggregating data attributes, etc. are supported. SAS also provides data profiling capabilities to inspect and address data for errors, inconsistencies, redundancies, and incomplete information. SAS provides data cleansing capabilities to correct, parse, standardize, identify, match, cluster data, and impute otherwise missing data in multiple ways. SAS provides ETL and ELT capabilities that can access a large array of data sources; extract, cleanse, transform, conform, aggregate, load, and manage data.
- Data formats/databases supported: SAS supports the collection and aggregation of information from many typical data sources (through SAS/ACCESS engines). These include general purpose relational databases: IBM DB2, Informix, Oracle, Microsoft SQL Server, Sybase, ODBC, OLEDB, and MySQL; Data Warehouses/Appliances: Aster Data, Greenplum (EMC), HP Neoview, Netezza, Teradata; Non-relational/other sources: ADABAS, Datacom, CA-IDMS, IMS, PC files, System 2000, VSAM, SOAP, RSS, WSDL, Microsoft Exchange, and JMP data; SAS supports bulk loading and SQL pass-through capabilities. SAS provides the integration of survey, market baskets, call detail records, clinical trial data, genetic sources, Web transactions, GIS spatial data, time-stamped transactional data, and many more sources as inputs to the modeling process.
- In-database mining: In database scoring is supported for Teradata, IBM Netezza, IBM DB2, Greenplum (EMC), and Aster Data.
- Model export: SAS Enterprise Miner supports PMML Version 4.0 for model score code generation. The score code is optimized to include the minimal

*SAS provides data cleansing capabilities to correct, parse, standardize, identify, match, cluster data, and impute otherwise missing data in multiple ways.*



required set of variable inputs to support fast and efficient deployment. Scoring code is also translated as native scoring functions for Teradata, IBM DB2, IBM Netezza, Greenplum (EMC) and Aster Data (Teradata). SAS provides Java and .NET in-database scoring APIs. C scoring functions can be exported as UDFs in DBMS and embed Java scoring code into applications like credit origination and rules engines.

- Model deployment/management: Using SAS Model Manger as the analytic deployment hub, models can be deployed in a SAS environment taking full advantage of the underlying SAS platform and its metadata management capabilities. Models can also be deployed directly via a number of in-database partnerships, using the SAS Scoring Accelerator. Using the Scoring Accelerator, users can directly deploy scoring logic developed in SAS into their native operational database, such as Teradata, IBM Netezza, IBM DB2, Asterdata, and Greenplum.
- Other: supports unstructured data, ensemble modeling, automation, hierarchical market basket analysis, and social network analysis.

*[StatSoft] customers report that the product is quite comprehensive with an intuitive user interface, exceptional graphics capabilities, and a great user support team. They also like the quality of the documentation.*

## StatSoft

Company Name: StatSoft  
Private

Website: [www.statsoft.com](http://www.statsoft.com)



The best predictive analytics solutions from a company you may have never heard of.

A US based company founded in 1984, StatSoft is probably best known among data miners for its award-winning STATISTICA product. The company is also recognized for its use of advanced analytics in quality control systems. Business users might not have heard of the company because historically, it tended to be more technology than market focused. Many of StatSoft's executive team have been with the business since its inception. The company's vision around its products consists of five core principles that it strives to meet. According to StatSoft executives, these are, "ease of use, comprehensiveness, fully integrated, lowest TCO, easy to do business with." The people at StatSoft pride themselves on providing their customers with an experience that has "no surprises." There are fourteen thousand functions in one product; it has a single user interface that was built from the ground up. StatSoft is a strong Microsoft and Intel partner and it adheres to these standards.

Customer reaction to the product is certainly in line with StatSoft's core values. Customers report that the product is quite comprehensive with an intuitive user interface, exceptional graphics capabilities, and a great user support team. They also like the quality of the documentation. On the quality control side, customers praise the company's quality control software and graphics.



HURWITZ  
& ASSOCIATES  
Insight to Action

## StatSoft Overview

Products	<ul style="list-style-type: none"> <li>• STATISTICA Enterprise (v. 10): the Server platform for delivering analytic and BI applications to departments and divisions within the enterprise via centrally-managed queries, analysis templates, report templates, and dashboards.</li> <li>• STATISTICA Data Miner (v. 10): Solution for model development, deployment (solution publishing), and management and database/data warehouse integration.</li> <li>• STATISTICA Text Miner (v. 10): Solution for retrieval, pre-processing, and analytic/interpretive mining procedures for unstructured text data.</li> <li>• STATISTICA Live Score (v. 10): the Server platform for integrating predictive modeling into business processes and line-of-business applications, real-time scoring.</li> </ul>
Target User/ company size	<p>Both business users and quantitative staff in any size company</p> <ul style="list-style-type: none"> <li>• Business users: have Data Miner Recipes as a step-by-step guided user interface for the application of predictive analytics as well as integrated recommendations and scores displayed in line of business applications via the STATISTICA Live Score product, Dashboards, and scheduled reports</li> <li>• Quantitative staff: can configure own models</li> </ul>
Verticals supported	<ul style="list-style-type: none"> <li>• Power generation (e.g., emissions reduction)</li> <li>• Manufacturing (e.g., root cause analysis and process optimization),</li> <li>• Pharmaceuticals/biopharmaceuticals (e.g., root cause analysis, multivariate statistical process control, product stability and shelf life estimation, compliance reporting),</li> <li>• Insurance (e.g., fraud detection, predictive claims),</li> <li>• Financial and banking applications, credit scoring</li> <li>• Others: Semiconductors, Chemical, Petrochemical, Food Processing, Automotive, Heavy Equipment, Telecom, R&amp;D</li> </ul>
Horizontal solutions	Churn, root cause analysis, emissions(CO2) reduction, fraud detection, product shelf life estimation (consumer products and medicines), batch genealogy (Pharmaceuticals), social media analysis and others
How offered	On-premises (annual licensing and maintenance) only
Differentiator(s)	<ul style="list-style-type: none"> <li>• Comprehensive set of capabilities.</li> <li>• Reasonable Price.</li> <li>• Quality of support team.</li> </ul>
Platforms supported	Windows, including all of the current Windows server (Windows Server 2008) and client (Windows 7) operating systems
Predictive Analysis Families	Association, classification (decision trees and neural networks), regression, optimization, simulation, and others

### *StatSoft's differentiators:*

- *Comprehensive set of capabilities*
- *Reasonable Price*
- *Quality of support team*



## Key Features

- ETL: Includes a set of configurations for data access, including a metadata layer that contains the role of particular data elements in the analytics. Supports pre-processing capabilities including missing data cleaning, duplicate detection, handling outliers, handling sparse data, calculated variables, recoding, date/time lagging. StatSoft can also integrate with 3rd party best-of-breed ETL tools. Optional ETL module for more sophisticated data pre-processing and for the scripting and scheduling of automated data pre-processing batch tasks (e.g., variable creation) is available.
- Data formats/databases supported: any standard relational database (e.g., SQL Server, Oracle, or any database that supports the OLE DB or ODBC standards) and databases that support multidimensional cubes such as SAP Business Warehouse. These database connections can be one-time queries or auto-refreshing queries, as new data are collected. STATISTICA offers IDP (In-place Database Processing) data access technology that allows it to process data directly from external databases as if data resided in STATISTICA native data files, which significant performance gains. STATISTICA also imports data from a variety of flat file formats including Excel, text (CSV or other), SAS, Minitab, JMP, and SPSS.
- In-database mining: Support for in-database scoring, derived variable creation and segmentation.
- Model export: C#, C, SQL, Java, SAS, PMML (version is model dependent), Visual Basic, Database Stored Procedure, Teradata.
- Model deployment/management: provides deployment via code for either batch-scoring in databases or compiling into a third party application for batch or real-time scoring, and the STATISTICA Live Score platform for the management of versions of models, linking to the respective real-time data sources, and a service-oriented architecture for delivering real-time predictions to line of business applications.
- Other: supports unstructured data, ensemble modeling, native 64-bit multi-processor version for enhanced performance, STATISTICA Data Miner Recipes®, an interactive work flow from data importing & pre-processing through to the automated evaluation of multiple models. It is designed as an aid to users who may have less experience with predictive analytics.

*[StatSoft] supports pre-processing capabilities including missing data cleaning, duplicate detection, handling outliers, handling sparse data, calculated variables, recoding, date/time lagging.*

## TIBCO

Company: TIBCO  
Public

Website: <http://spotfire.TIBCO.com>

Companies looking to provide business users with interactive visualization combined with the results of predictive models to enrich the information available for analysis should consider TIBCO Spotfire.



HURWITZ  
& ASSOCIATES  
Insight to Action

TIBCO, perhaps best known in the analytics world for its market leading Spotfire visualization product, acquired the statistics and data mining company Insightful in 2008. Insightful developed the S+ family of statistics products that grew out of work done at Bell Laboratories. It also has interfaces to open source R. The company plans to continue to enhance S+ while embracing R. TIBCO's view is that the combination and integration of visual and statistical analysis is what differentiates the product from its competitors. The visualization component of Spotfire allows those analysts building a model to first explore their data visually, then use the data mining tools available in Spotfire (Spotfire Miner, Spotfire S+, or R) to develop and deploy predictive models to business end users who can consume and interact with the results via the Spotfire visual interface. Spotfire enables these analysts to explore and then communicate their own findings to these business users, and to provide stats-powered applications and tools to those users, via its one click Web publishing capability, so that they can utilize the analyses themselves. Given its heritage in integration, TIBCO also enables users to put to use the results of predictive models in business processes.

*End-users liked the visualization component of Spotfire and the fact that Spotfire is tightly integrated with TIBCO's data mining and deployment software.*

End-users liked the visualization component of Spotfire and the fact that Spotfire is tightly integrated with TIBCO's data mining and deployment software. Customers reported that they are effectively using the Spotfire platform to build predictive models and then easily deploy mashups for use in an analysis Web portal. Additionally, users said that the rules can be fairly easily incorporated to flexibly operationalize a business process. TIBCO also got high marks for new functionality roll out, value/price, and technical value.

## TIBCO Overview

---

### Products

- TIBCO Spotfire Professional: Used to build and deploy analytic applications over the Web or to perform ad hoc analysis using visualization. TIBCO also provides an add on called TIBCO Spotfire Metrics that provides more advanced statistical techniques around clustering, pattern matching, and correlations (as well as features to support application creation).
- TIBCO Spotfire Webplayer: Allows users to visualize and interact with analysis created with Spotfire Professional via a compatible browser.
- TIBCO Spotfire Statistics Services: Allows results from Spotfire S+, Spotfire Miner and R to be integrated into Spotfire.
- TIBCO Spotfire S+: Used by statisticians and other experts to develop models that can be integrated with Spotfire through Spotfire Statistics Services.
- TIBCO Spotfire Miner: Enables predictive analytics via a visual drag and drop interface. Work flows can be extended by integrating scripts from TIBCO Spotfire S+ and/or deployed using Spotfire Statistics Services.
- TIBCO Spotfire S+FinMetrics: Specifically for financial institutions for econometrics, time series and asset pricing methods, portfolio analysis, risk modeling, capital modeling.



- TIBCO Spotfire S+NuOPT: Focused on solving very large optimization problems such as portfolio optimization, nonlinear and robust statistical modeling and circuit optimization.

Target User/ company size	Both statisticians and business users
Vertical Solutions	Life sciences research and clinical development, manufacturing, financial services, government. Also, consumer goods, energy, telecommunications, and more
Horizontal solutions	Spotfire Network Analytics, operations analytics, Spotfire for Spend Analysis (Procurement reporting and analysis), Spotfire for Sales & Marketing
How offered	Spotfire pricing is based on the named user model and a three-year term license.
Differentiator(s)	<ul style="list-style-type: none"> <li>• Interactive visualization without the use of cubes.</li> <li>• Model deployment - one click publishing to Web clients.</li> <li>• Strong integration capabilities.</li> </ul>
Platforms supported	TIBCO Spotfire Miner, S+ (desktop) and Spotfire Statistics Services run on Solaris, Red Hat Enterprise, SLES (SuSE Linux), and all variants of Windows. In addition, TIBCO Spotfire Professional, Web Player Server, and Clinical Graphics (client) run on Windows.
Predictive Analysis Families	Association, classification, information extraction, regression, optimization, simulation, time series and more. Contains over 4000 analytic functions built-in, and an active community of statisticians and developers periodically contributing new methods.

#### Key Features

- ETL: TIBCO Spotfire Miner and S+ support the following operations for data preparation: creating, simulating, capturing/writing, profiling, cleansing (missing values, duplicate detection, outlier detection), matching, comparing, merging/joining, transforming and encoding (translating coded values, deriving/calculating values, partitioning, sampling, filtering, shuffling, sorting, aggregating, transposing/pivoting, column splitting, normalizing, validating), augmenting, and loading. The addition of TIBCO Collaborative Information Manager (CIM) provides a single vendor solution for ETL and master data management. External vendor data preparation tools supported include Informatica and SAS.
- Data formats/databases supported: TIBCO Spotfire Miner and S+ support data access from relational databases through native drivers for Oracle, DB2, SQL Server, Sybase, as well as non-native access through ODBC and JDBC. Miner and S+ also support data access to file-based data sources such as Spotfire STDF, SAS, Access, Excel, SPSS, Matlab, Minitab, Stata, Systat, Lotus, dBase, FoxPro, Paradox, Quattro Pro, and Gauss, in addition to more generic

#### TIBCO's Differentiators:

- Interactive visualization without the use of cubes
- Model deployment - one click publishing to Web clients
- Strong integration capabilities



delimited and fixed format text data files. Through integration with Spotfire Application Data Services, Miner and S+ gain access to data retrieval from enterprise data sources, including SAP BI, SAP ERP, Salesforce.com, Siebel eBusiness, Oracle E-Business, and SOAP Web Services.

- Model export: C code and as PMML v. 2.1 (which can be converted to later versions using the PMML converter provided by Zementis).
- Publishing/model deployment: The primary method for deploying models is by publishing them to Spotfire Statistics Services. Through Spotfire Statistics Services, the models can be automated or scheduled, integrated into Spotfire applications deployed through Spotfire Professional or Spotfire Web Player, or integrated into other applications through Web Services. TIBCO also offers what it refers to as its “one-click publishing” capability; i.e. users of TIBCO Spotfire Professional just save their work on the Spotfire Server and can publish it in for business users using one click. These business users then have access to this enriched information via the Spotfire interface.
- Interactive visualization as part of Spotfire, users receive all of the functionality associated with this product.
- Other: supports unstructured data.

*Through [TIBCO's] Spotfire Statistics Services ... models can be automated or scheduled, integrated into Spotfire applications deployed through Spotfire Professional or Spotfire Web Player, or integrated into other applications through Web Services.*

For information about reprints please contact [info.request@hurwitz.com](mailto:info.request@hurwitz.com)



## About Hurwitz & Associates

Hurwitz & Associates is a consulting, market research and analyst firm that focuses on how technology solutions solve real world business problems. The firm's research concentrates on disruptive technologies, such as Cloud Computing, Service Oriented Architecture and Web 2.0, Service Management, Information Management, and Social and Collaborative Computing. We help our customers understand how these technologies are reshaping the market and how they can apply them to meet business objectives. The team provides direct customer research, competitive analysis, actionable strategic advice, and thought leadership. Additional information on Hurwitz & Associates can be found at [www.hurwitz.com](http://www.hurwitz.com).



© Copyright 2011, Hurwitz & Associates

All rights reserved. No part of this publication may be reproduced or stored in a retrieval system or transmitted in any form or by any means, without the prior written permission of the copyright holder. Hurwitz & Associates is the sole copyright owner of this publication. All trademarks herein are the property of their respective owners.

This report is based on independent research and input from a random sampling of users. This report represents our findings, some user experiences may vary from our conclusions. Hurwitz & Associates did not receive financial or other benefits in connection with the Victory Index while preparing the report. The information contained in this report has been obtained from sources believed to be reliable. Hurwitz & Associates shall not be liable for any omissions or errors in the information contained in this report. All opinions in this report are subject to change without notice.

175 Highland Avenue, 3rd Floor • Needham, MA 02494 • Tel: 617-597-1724  
www.hurwitz.com