BARC

BARC Spotlight

An introduction to visual data science

Author: Shawn Rogers, CEO BARC US

Publication: July, 2025

Research provided by Spotfire

Index

BARC

An Introduction to Visual Data Science

Visual data science platforms

Visual data science platforms bridge the gap left by traditional analytics and business intelligence (ABI) platforms. ABI solutions have incorporated advanced analytics powered by machine learning (ML) for many years. But these solutions often struggle to easily combine data science algorithms with specialized visual insights through interactive workflows. These systems enable enterprise analytics professionals to address a broad range of general-purpose analytics for descriptive, diagnostic, predictive and prescriptive analytics workloads.

The challenge for ABI solutions is the focus on an 80/20 approach to the market. By focusing on the broad needs of the 80 percent, ABI solutions can lower the cost of entry for users who need highly commoditized analytics, presented as simple dashboards and reports. At the same time, these solutions strategically ignore the 20 percent of workloads that are driven by domain experts and require specialized visualizations and workflows. Serving user communities with broad analysis and reporting capabilities certainly has value, but it also creates a gap in the market for companies that require deeper insights driven by data science within specialized industries. This is where visual data science augments and extends analytic capabilities in a strategic way.

Visual data science platforms focus on sophisticated workflows, targeting the 20 percent of projects that require advanced visualization and analytic capabilities. They bring together data scientists, business users, and subject matter experts in a collaborative, visual-first environment -leveraging data science insights within a single platform. This approach avoids long project timelines and collaboration gaps that often exist between data science and machine learning (DSML) platforms and ABI solutions.

Bridging the gap

Historically, ABI solutions, along with traditional DSML platforms, have not creatively engaged new user personas and often focus only on a variety of business users. Domain experts are a hybrid version of business users with deep domain expertise – such as doctors in healthcare companies, geologists in energy companies, or quality engineers in manufacturing. They require specialized data and visual tools from traditional business users along with specific analytic needs.

The gap grew wider as data science professionals started to deliver insights to the business in the early 2000s these were generally project-oriented and executed independently with limited integration to existing platforms, systems, and teams. Traditional business users, along with the domain experts, queued outside their door patiently waiting for data science bandwidth to help with projects, which greatly delayed project timelines and time to insights.

IT plays a role in the gap that visual data science can fill. Traditional and manual data provisioning for each of the groups mentioned reduces IT's time to innovate, as they try to keep pace with data and systems demands. Advanced business users added a new wrinkle to the gap with the advent of the citizen data science. These users understand the remarkable outputs from advanced analytics and mostly lack the time needed to create analyses. Additionally, citizen science users require user interfaces (UIs) that offer reusable workflows, models, and algorithms in drag-and-drop environments, with minimal need for coding in R or Python.

A practical bridge is required to close the gap. Visual data science enables data scientists, business users, domain experts, and IT to reduce the time spent on steps that curtail the success of these groups and speed them towards success.

Why now

Why should your organization investigate augmenting your analytics strategy with a visual data science platform today? Here are a couple leading reasons:

- **Demand for faster actions** – Between insights and action, there is a window of value. Scaling analytics, along with data science, requires speed and the ability to leverage the moments when the value is at its highest. If your systems are too slow, too siloed, and disparate, your company will miss this value/opportunity window repeatedly, costing the organization significant monetary loss. Deploying projects to production rapidly and intersecting with this opportunity is a significant reason for adding visual data science to your lineup.

- **Siloed solutions** – On average, companies are working with 3 to 4 ABI platforms. Combine this number with various data management solutions and data science platforms, and it's easy to see why so much time is wasted when coming together on projects. These siloes create a significant time block as well as a data and governance nightmare. In the world of compliance and regulatory pressures, the more siloed the workflows, the more likely mistakes and failures will occur. The comprehensive nature of visual data science platforms can be a valuable replacement for some data management tools and less feature rich ABI solutions.

- **Explosive data** - As data has grown by type, speed, and volume, it's powering ML-driven insights and providing deeper data science driven actions. Real-time data has graduated from niche utilization to a standard data type for enhanced analytics. IoT sensor data and streaming data are now providing vast value when combined with other important data types in the enterprise.

- **Value** – The impact of analytics has been more than proven over the last several decades. Visual data science platforms represent that final part of the journey towards analytic value, delivering on the gap left by some traditional ABI and DSML platforms.

The traditional model of IT providing ABI solutions to business users, while excluding domain experts and siloing data scientists on disconnected platforms, no longer works. Companies are augmenting and extending their analytics strategy with visual data science platforms to meet the needs of fast insights/actions, eliminating siloed solutions, and meeting the challenges of explosive data to deliver better business value.

Critical components of visual data science platforms

Visual data science platforms can differ, but it's critical that the following four components act as the foundation to the solution.

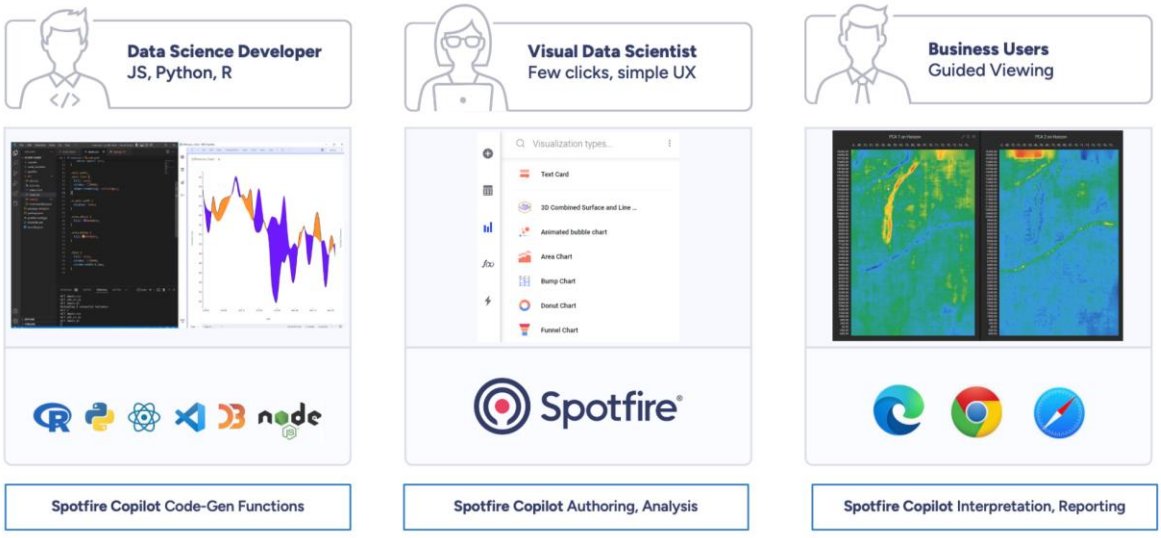
- **Visual-first analytics** – Highly interactive visualizations are the cornerstone for collaboration and speed the process for insights. They act as the primary user interface in visual data science platforms, helping users to spot trends, patterns, clusters, and outliers. This visual representation of the data, process, code, and analysis is the collaboration palette for the project stakeholders. The visualization enables users to align around the data narrative much faster than text or rows and columns. These domain-specific visualizations obscure difficult to understand R and Python code while quickly delivering the insights driven by it. A visual-first approach arms the various constituents with initial insights and the ability to explore, add more data, and optimize the analytics directly from the visualization. With all parties working from the same location, it speeds project time, reduces rework, and enables all experts to contribute within the same moments, avoiding traditional time wasting back and forth between parties. Lastly, these visualizations are often highly specific analytic pain points for the customer's business, think 3D surface and line plots and continuous process verification (CPV) charts, not simple bar charts.

- **Industry and domain focused** – Visual data science platforms deliver purpose built and vertical specific visualizations and analytic workflows that deliver significant value for specific industries. Highly regulated industries, such as pharmaceutical manufacturing, computer chip manufacturing, and the energy sector, demand highly specific business insights for mission critical projects. Visual data science platforms generally include preconfigured industry-specific visualizations, workflows, accelerators, and fast-start modules to deliver value quickly. Generally, these platforms align easily with the types of data used by these specific industries, such as streaming sensor data, and Supervisory Control and Data Acquisition (SCADA) systems in manufacturing, and geological data in the energy sector. By quickly and easily including this data, users can analyze what matters and communicate in the language of the business and domain experts.

- **Seamless data interaction** – Interacting with critical data without needing to switch to other platforms or requiring coding is an incredible time-saver and collaboration tool. Visual data science platforms also need to deliver the ability to transform, enrich, access, model, explore, prepare, and clean data from within the platform. Openness to all enterprise data sources (cloud, hybrid and on-premises), API's, and system and data connectors is required along with maintaining connectivity and history with data and AI governance platforms. This increased agility helps with experimentation, discovery, and can impact data quality while enabling faster troubleshooting in projects. Overall, seamless data interaction eliminates the traditional friction between IT, business users, domain experts and data scientists. Reducing wasted cycles and fast-tracking insights.

- **One Platform for all** – Visual data science platforms offer a single shared environment that allows users to deliver the entire spectrum of analytics. This unified UI promotes low and no-code workflows for domain experts and business users while delivering the unique and powerful tools required by data science practitioners. Powerful R, Python interfaces need to be embedded into the platform for development, AI code generation, and interactive operations on the visualizations. The platform needs to have easy to use collaboration features that engage all personas working on the platform, including comments, annotations, shared views etc. The single platform reduces solution hopping from ABI-style tools to data management tools to data science tools savings time and expense. Additionally, centralizing the platform allows for easier definition, process, and workflow sharing, as well as amplifying reuse of vetted data science models and standards. Lastly, the one platform simplifies deployment and management of sophisticated analytic projects.

- **GenAI as a copilot across roles** - From a simple chat interface, Generative Artificial Intelligence (GenAI) acts as a powerful assistant — copilot, companion, or sidekick—capable of remarkably sophisticated tasks. For developers, it generates code for data science workflows and visualizations. For visual data scientists, it configures visualizations, guides analysis sequences, and interrogates data across virtual views and source systems. For consumers, it explains insights, builds narratives, and turns complex analysis into accessible storytelling.

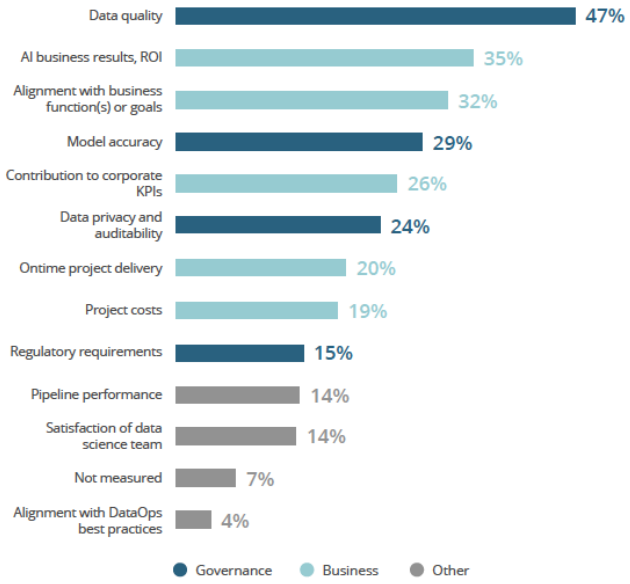


Considerations and challenges

Introducing a visual data science platform into an already busy environment that includes a mix of ABI tools, DSML platforms and a complex data infrastructure is not without challenges. However, the immediate return on investment (ROI) by activating visual data science can far outweigh the work and expense. Several strategic items to keep in mind:

- **Governance, security, and compliance** – Visual data science platforms require integration with existing data governance platforms. Aligning with these systems will eliminate audit and compliance gaps and regulatory mistakes. Visibility, lineage, and control are critical for enabling the new platform to perform.

- **Data quality** – Data quality is a challenge for most companies. Visual-first platforms need clean, high-quality data to fuel performance. It's not just a race to insights; it's a race to trust. Quality data plays a key role in that process and data quality should be prioritized when introducing a visual data science platform into your ecosystem. In a recent BARC research study[1] data quality was selected by 47 percent of 333 global respondents as their top metric for success with data governance.



- **Performance and scale** – Confirmation of fast compute and scalability is important when deploying visual data science. Organizations can utilize cloud-native compute environments to respond to system demands, leverage query optimizing technology, and test with real-world workloads. Interactive visual-first environments coupled with predictive and other ML outputs will tax a compute ecosystem. Ensure you have the right technology to empower the platform.

Deploying a visual data science platform isn't plug-and-play, but with smart planning and clear alignment to business goals, it can unify teams, speed decisions, and unlock more value from your data investments.

Use Case Examples

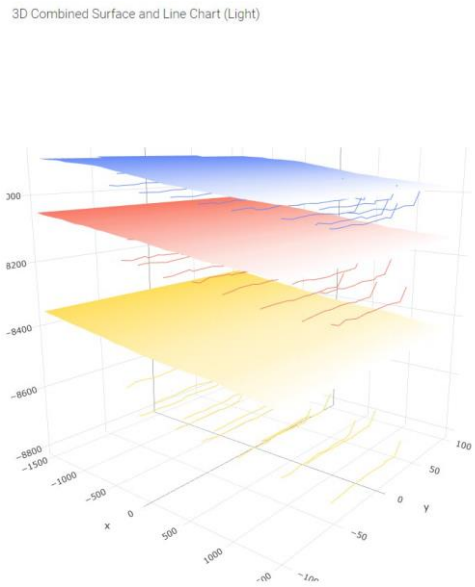
Upstream energy

The energy sector, especially upstream exploration and production, offers a glimpse into strategic use cases that are uncommon and driven by highly skilled domain experts (scientists) with unique data sources. Think about the cost of oil well placement—leveraging geological, seismic, economic, supervisory control, and data acquisition (SCADA) systems along with other real-time equipment and production data.

The operational cost of selecting optimal sites easily runs into the millions of dollars. Selecting the wrong site can be devastating for an energy company. These data-intensive decisions require the collaboration of visual data scientists, energy domain experts, geologists, and line-of-business executives at enterprise scale.

The visual-first approach of visual data science platforms plays a highly valuable role in these mission-critical projects. The example below, a 3D combined surface and line chart, serves as the framework for visualizing spatial relationships with interactive filters and measuring tools, facilitating insight, collaboration, and data management. This provides a single location for these expert personas to work together and make faster, more informed decisions.

[1] BARC Research – Preparing and Delivering Data for AI, Adoption Trends, Requirements and Best Practices



Spotfire 3D Mod diagram Spotfire®

3D combined surface Mod

Manufacturing

Reducing costs, improving quality, and optimizing operations are top strategic initiatives for companies in the manufacturing sector. Predicting supply chain failures, monitoring machine and tool maintenance requirements, and identifying compliance and regulatory issues are difficult challenges if the teams responsible are working on disparate platforms and data. Smart factories incorporate all these workflows, and visual data science platforms bring teams and data together to enable them.

Manufacturing use cases are highly dependent on timely data from product measurements, machines, processes, and third-party data sources to deliver value. Insights not delivered in a timely manner quickly lose their value. Historical analysis can be valuable, but real-time and predictive insights drive the operational value of a smart factory. An excellent example is continuous process verification (CPV), which provides immediate insights into quality variations and often relies on a visual-first approach to identify multivariate anomalies in process parameters and other hidden issues. Consistently tracking quality provides the data necessary to remain compliant with regulatory guidelines. This is non-negotiable in industries like pharmaceuticals, where manufacturers are held to strict US Food and Drug Administration (FDA) guidelines, such as 21 CFR Part 11. Data trustworthiness is a cornerstone of this FDA regulation. CPV is often part of a digital twin environment that allows domain experts, data scientists, and other stakeholders to clearly see and predict issues in manufacturing processes.

These systems work quickly and generally deliver more value when they are collaborative, provide easy integration with data (sensors and streaming data), and leverage advanced analytics from data science teams. These insights fuel quicker root cause analysis and save factories downtime and money.

Conclusion

Visual data science platforms are a unique and valuable addition to existing analytic strategies. They add productivity and enable cross-functional analytics by "closing the gap" and driving collaboration among highly talented stakeholders across the organization. Addressing the opportunity to deliver faster and interactive outcomes with visual data science is a fast path to greater revenue and process optimization for all data-driven companies.

About Spotfire

Contact info

Global Headquarters
Spotfire, a business unit of
Cloud Software Group
851 Cypress Creek Rd.
Fort Lauderdale, FL
33309, United States

Research and Development
Sweden (Göteborg
Första Långgatan 26
413 28 Göteborg
Sweden
Spotfire.com

Spotfire, a business unit of Cloud Software Group, goes beyond basic rearview dashboards to offer a single visual analytics platform for data exploration and real-time decisions.

The Spotfire® visual data science platform combines advanced analytics and interactive visualizations. Visual data science is a powerful visual-first approach to data science that combines data wrangling, visualization, modeling, and predictive analytics in a single, collaborative platform. It breaks down silos between tools and teams—freeing data scientists from time-consuming data prep and giving them more time to focus on meaningful analysis. At the same time, business users can easily access insights with just a few clicks. Built for data-driven innovation, Spotfire champions this visual-first approach to data science, making it fast and easy to explore, visualize, and operationalize data across the enterprise. It combines market-leading visual analytics, data science, and data wrangling to allow experts to analyze data at-rest, in-motion, and at-scale—solving problems that require human creativity and modern computing.

Whether you're building predictive models, analyzing real-time streaming data, or uncovering trends through interactive dashboards, Spotfire brings data science and business insight together—visually and intuitively. Enjoy a point-and-click user experience where engineers, scientists, and other experts can rapidly build sophisticated, interactive visual data science applications and deploy those to thousands of end-users, enhancing business-critical operations and daily decision-making. Trusted by leading organizations worldwide, Spotfire helps teams accelerate discovery, scale impact, and unlock the full value of their data.

Empowering data-driven decisions through visual data science

Spotfire is a visual data science platform designed for scientists, engineers, and domain experts to rapidly analyze complex data and derive actionable insights for business operations.

Spotfire is widely used to address challenging industrial science and engineering problems. Some high-value examples include optimizing yield in high-tech manufacturing, accelerating pharmaceutical drug discovery and development, and addressing geoscience, exploration, production, and supply challenges in the energy sector.

Spotfire provides an immersive user experience with analytics tools at the fingertips of the visual data scientist—to drill down, up, and sideways across data domains of interest. The Spotfire hybrid-in-memory data engine is fast, and with "scheduled updates" and "data-on-demand," big data can be loaded in and out of memory rapidly. Data virtualization capabilities offer a dedicated semantic data layer, featuring query optimization, intelligent caching, and a catalog of virtual views managed within Spotfire.

Data science tools are built into Spotfire and are extensible with Python and R. Geospatial analysis methods include interactive multi-layer maps, enabling DSML calculations within and between layers. Methods for spatial joins, choropleth maps, contours, heatmaps, networks, and hotspots can be invoked from interactive user markings. Time series methods are available for at-rest and event-stream data, with live data and alerts on all Spotfire visualizations.

The visualization palette is rich and can be extended with JavaScript in an interactive Spotfire development environment. A large palette of specialized visualizations and data science functions for industrial science applications is included.

Spotfire includes built-in AI engines for variable association, visualization, and data prep recommendations, along with RAG LLM infrastructure for interactive chat and Agentic AI for data discovery. Spotfire visual data scientists and developers can populate Spotfire data science functions with Python code, and business analysts can interrogate data to obtain insights, reports, and explanations.

Spotfire enables all users to rapidly discover insights and take actions to update business strategy and operations. Spotfire visual data scientists and developers can create full-featured, interactive applications in days versus months compared to other tools. Spotfire is pioneering a new data and AI approach to business strategy and operations with its latest visual data science platform.

Michael O'Connell,

Chief Analytics Officer

Spotfire