Market Guide for Augmented Analytics Tools

Published: 30 October 2019   **ID:** G00433836

Analyst(s): Austin Kronz

Augmented analytic capabilities are disrupting analytics and BI and data science and machine learning markets. Tools leverage ML/AI to transform how analytics content is developed, consumed and shared. Data and analytics leaders should plan to adopt augmented analytics as capabilities mature.

Key Findings

- Augmented analytics use machine learning (ML) to automate data preparation, insight discovery, model development and insight sharing for a broad range of business users, operational workers and citizen data scientists.

- Augmented analytics tools can identify the most important insights, based on statistical significance, and, in more advanced tools, users' preferences and business context/relevancy (location, role, time, etc.).

- Analytics and business intelligence (ABI) and data science and machine learning (DSML) solutions and platforms often complement augmented capabilities with natural language processing (NLP) and conversational interfaces, allowing all users to interact with data and insights without requiring advanced skills.

- Augmented analytics will make DSML model building accessible to new citizen data science roles (business analysts, developers and others) while making expert data scientists more productive, collaborative and less biased, freeing them for high-value tasks.

Recommendations

As a data and analytics leader planning to use augmented analytics to modernize solutions, you should:

- Explore opportunities to complement existing data and analytics initiatives by piloting augmented analytics for business extension projects, tasks that currently involve time-consuming manual analysis and high-value business problems.

- Observe the augmented analytics capabilities and roadmaps of established data and analytics providers, enterprise application vendors and startups. Assess the upfront setup, data

preparation, openness and explainability of models, as well as the number of variables supported, the range of algorithms and model accuracy.

- Develop a strategy to improve cross-skill-set collaboration which includes evolving roles, responsibilities and skills, and increase investments in data literacy, being cognizant of your organization's perceptions and appetite for things like AI.

Strategic Planning Assumptions

By 2021, augmented analytics will be a dominant driver of new purchases of analytics and business intelligence (BI) as well as data science and machine learning (DSML) platforms, and of embedded analytics.

By 2021, NLP and conversational analytics will boost analytics and BI adoption from 35% of employees to over 50%, including new classes of users, particularly front-office workers.

By 2021, 50% of analytical queries will be generated via search, NLP or voice, or will be automatically generated.

By 2025, a scarcity of data scientists will no longer hinder the adoption of data science and machine learning in organizations.

By 2021, automation of data science tasks will enable citizen data scientists to produce a higher volume of advanced analysis than specialized data scientists.

Market Definition

Augmented analytics is the use of machine learning (ML) automation and artificial intelligence (AI) techniques to augment human intelligence and contextual awareness. This is achieved by automating pieces of data preparation tasks, data analysis, insight generation and explanation within analytics and BI platforms (ABI), as well as many aspects of data science and machine learning (DSML) model development and consumption. As businesses become inundated with data, augmented analytics becomes crucial for presenting to users across the business only the insights important for them to make a decision or take an action at that moment. It drives less-biased decisions and more-impartial contextual awareness — transforming how users interact with data, make decisions and act on insights. It expands who has access to insights from analytics by delivering analytics anywhere and to everyone in the organization, and does so with less time, skill and interpretation bias than current manual approaches.

A similar trend involves augmented data management where AI and ML techniques are automating many aspects of database management, data quality, data integration, data cataloging, etc. This market guide focuses on vendors providing augmented analytics capabilities — specifically augmented ABI and augmented DSML tools.

Market Description

Augmented analytics capabilities will rapidly achieve mainstream adoption as a key feature of data preparation, modern analytics and BI and data science and machine learning platforms. More importantly, automated insights will also be embedded in enterprise applications and conversational analytics — and thereby reach beyond citizen and expert data scientists to augment the capabilities of operational workers to make better decisions by leveraging insights from advanced analytics.

Currently in analytics, content authors such as analysts, citizen data scientists and expert data scientists perform the following three data-to-insight-to-action activities iteratively to find meaningful insights:

1. Preparing the data

2. Finding patterns in the data and building models

3. Sharing and operationalizing findings from the data

The augmented analytics paradigm accelerates the time it takes to get accurate insights for business users. It augments their analysis by using ML algorithms to automate those three main parts of the data and analytics workflow (see Figures 1 and 2). How augmented analytics affects each component of the workflow is explored below.
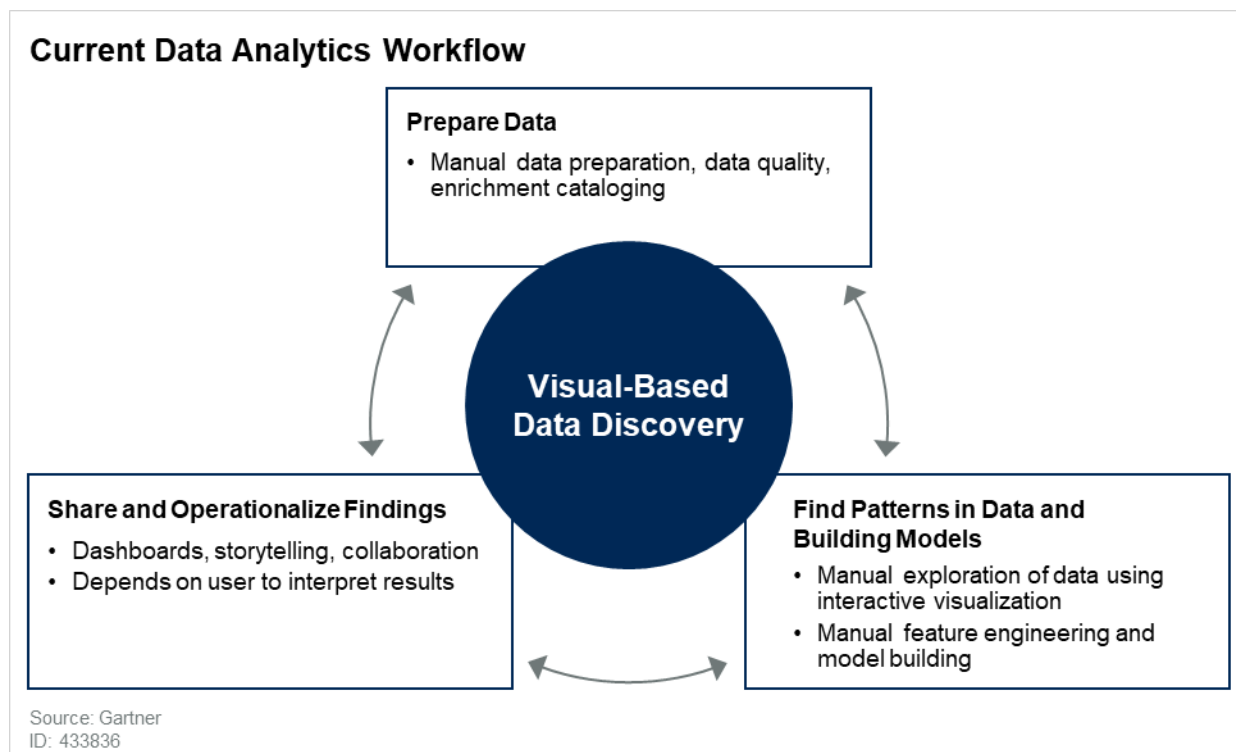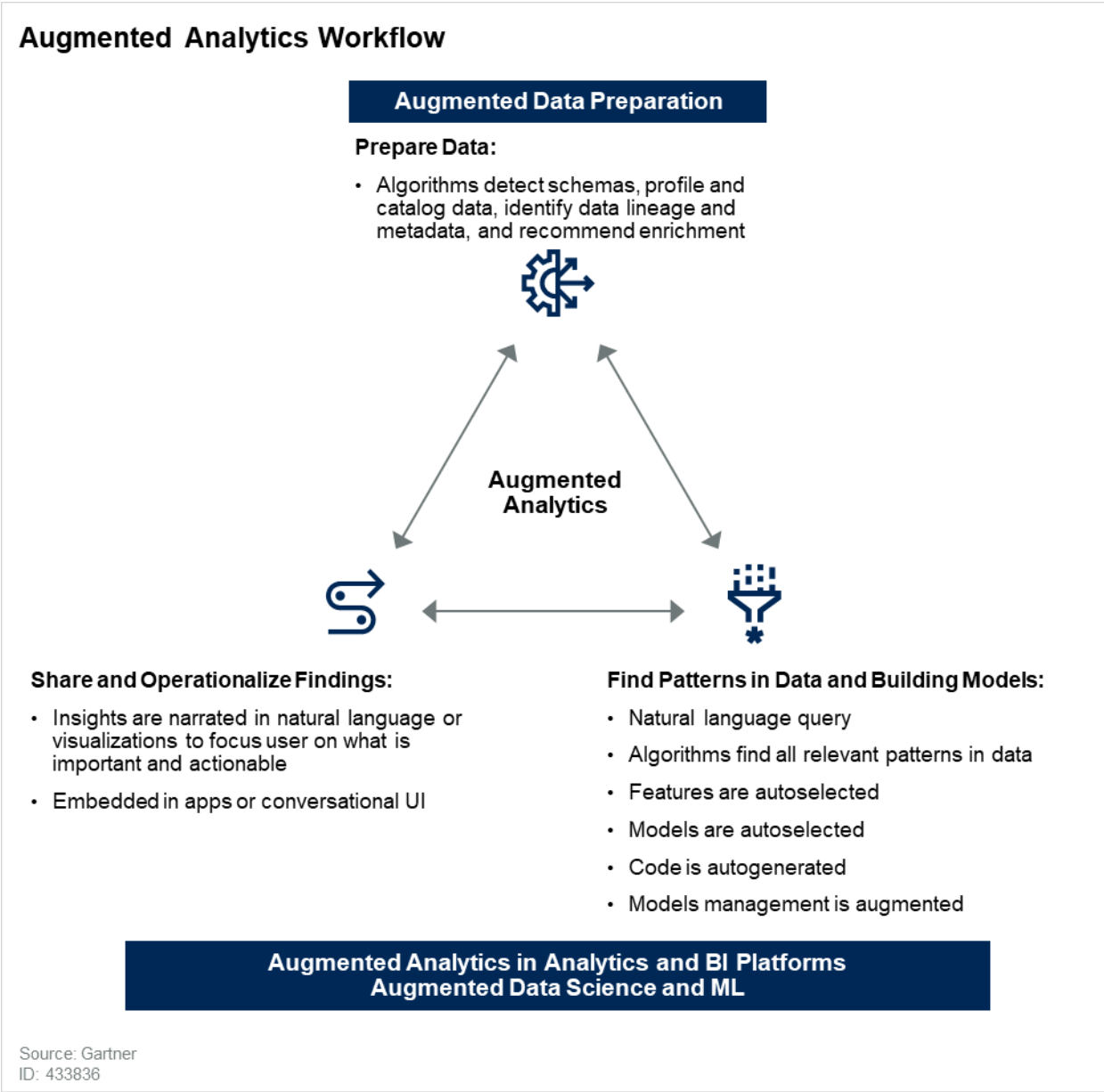
Figure 1. Current Data Analytics Workflow

Current Data Analytics Workflow

Prepare Data
- Manual data preparation, data quality, enrichment cataloging

Visual-Based Data Discovery

Share and Operationalize Findings
- Dashboards, storytelling, collaboration
- Depends on user to interpret results

Find Patterns in Data and Building Models
- Manual exploration of data using interactive visualization
- Manual feature engineering and model building

Source: Gartner
ID: 433836

Figure 2. Augmented Analytics Workflow

Augmented Analytics Workflow

Augmented Data Preparation

Prepare Data:

- Algorithms detect schemas, profile and catalog data, identify data lineage and metadata, and recommend enrichment

Augmented Analytics

Share and Operationalize Findings:

- Insights are narrated in natural language or visualizations to focus user on what is important and actionable
- Embedded in apps or conversational UI

Find Patterns in Data and Building Models:

- Natural language query
- Algorithms find all relevant patterns in data
- Features are autoselected
- Models are autoselected
- Code is autogenerated
- Models management is augmented

Augmented Analytics in Analytics and BI Platforms
Augmented Data Science and ML

Source: Gartner
ID: 433836

Market Direction

Over the past 10 years, visual-based data discovery tools have disrupted the traditional BI market. Visual-based data discovery capabilities have become the defining feature of modern analytics and BI platforms. These easy-to-use tools enable users to assemble data rapidly, explore hypotheses visually and find new insights in data. They have transformed how business users explore data, in comparison with the IT-centric, semantic-layer-based approach of traditional BI platforms.
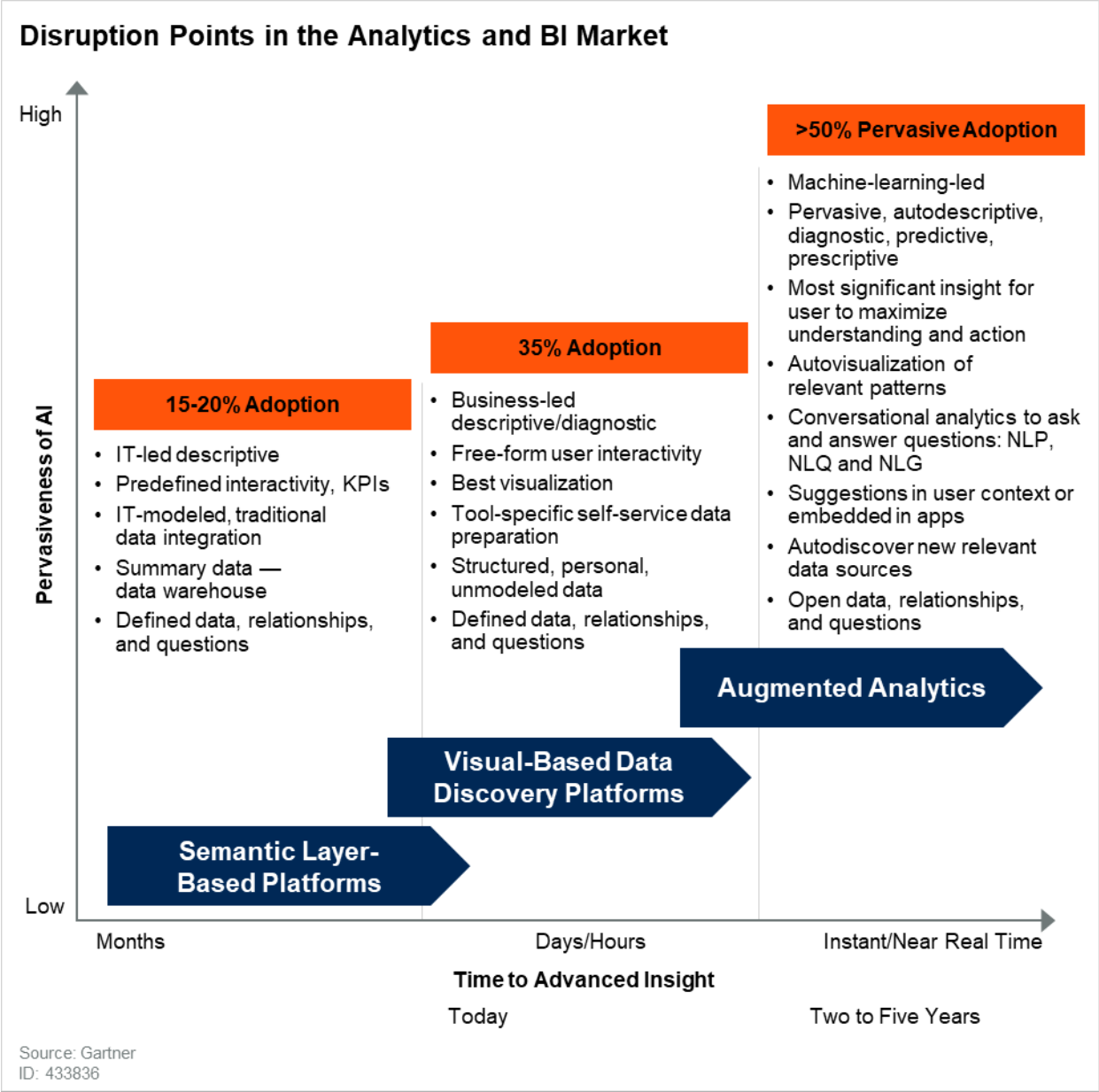
Visual-based data discovery features — now standard features of modern analytics and BI platforms — are easy to use, because users analyze data by creating visual queries to investigate hypotheses. But when data sizes and the number of variables is large, it is not possible for users to explore every possible pattern and combination, let alone determine whether their findings are the most relevant, significant and actionable. As a result, augmented analytics and business intelligence (BI) platforms are now the primary driver of the modern analytics and BI market (see "Forecast: Modern Business Intelligence Platforms by Selected Functionality, Worldwide, 2017-2022").

Moreover, the idea that "a picture is worth a thousand words" has long been assumed in the field of data and analytics. And rightly so, as visualizations are a powerful and consumable way to find and communicate patterns in data (more so than tables or lists). However, they only highlight visual relationships; they do not identify statistically significant findings. For example, a user may be able to see visual differences between bars on a bar chart, but unless the size differences are stark it would require further analysis to determine whether the differences are relevant, statistically significant and actionable. Moreover, finding insights from advanced analytics — a key aspirational goal for most companies, particularly those facing rapid digital and industry change — requires expert data science skills, which are extremely scarce.

Visual-based data discovery with manual interactive exploration using visualizations has been the defining feature of modern analytics and BI platforms. Now, however, augmented analytics characterized by ML/AI automation of the insight discovery, exploration, explanation, prediction and prescription process is a defining feature of new-generation analytics and BI platforms (see Figure 3).

Augmented analytics can reduce time-consuming exploration and the identification of false and irrelevant insights. Applying a range of algorithms and ensemble learning to data in parallel, and explaining actionable findings to users, reduces the risk of missing important insights in the data, when compared to manual exploration. It also can help optimize resulting decisions and actions by including in-line descriptions of underlying factors — or key drivers — causing a particular metric to deviate from a predetermined or statistical norm. This paradigm shift requires investment in data literacy throughout organizations, as insights are distributed to all employees (see "Toolkit: Enabling Data Literacy and Information as a Second Language").

Figure 3. Disruption Points in the Analytics and BI Market

Disruption Points in the Analytics and BI Market

>50% Pervasive Adoption

- Machine-learning-led
- Pervasive, autodescriptive, diagnostic, predictive, prescriptive
- Most significant insight for user to maximize understanding and action
- Autovisualization of relevant patterns
- Conversational analytics to ask and answer questions: NLP, NLQ and NLG
- Suggestions in user context or embedded in apps
- Autodiscover new relevant data sources
- Open data, relationships, and questions

35% Adoption

- Business-led descriptive/diagnostic
- Free-form user interactivity
- Best visualization
- Tool-specific self-service data preparation
- Structured, personal, unmodeled data
- Defined data, relationships, and questions

15-20% Adoption

- IT-led descriptive
- Predefined interactivity, KPIs
- IT-modeled, traditional data integration
- Summary data — data warehouse
- Defined data, relationships, and questions

Augmented Analytics

Visual-Based Data Discovery Platforms

Semantic Layer-Based Platforms

Pervasiveness of AI — High / Low

Months — Days/Hours — Instant/Near Real Time

Time to Advanced Insight

Today — Two to Five Years

Source: Gartner
ID: 433836

Market Analysis

Augmented analytics includes augmented data preparation, augmented analytics and business intelligence (augmented ABI), and augmented data science and machine learning (augmented DSML) segments. Table 1 contains a list of key capabilities seen in each segment — while this market guide is focused on the augmented ABI and augmented DSML segments, the tools may (and typically do) have some or all of the capabilities in one or multiple segments listed. For a more

detailed analysis, examples and use cases, see "Augmented Analytics Is the Future of Data and Analytics."

Table 1. Augmented Analytics Capabilities

Category	Example Capabilities	Additional Information
Augmented Data Prep	▪ Automated matching, joining, profiling, tagging and annotating data prior to data prep ▪ Sensitive attribute recognition ▪ Automate repetitive transformations and integrations ▪ Data quality and enrichment recommendations	Uses ML/AI automation to augment and accelerate data profiling and data quality, harmonization, modeling, manipulation, enrichment/inference, metadata development, and data cataloging (see "Market Guide for Data Preparation Tools"). This augmentation trend is also transforming all aspects of data management, including automating data integration and database/data lake administration.
Augmented ABI	▪ Automated descriptive insights ▪ Key driver analysis ▪ Segment and cluster identification ▪ Anomaly/outlier detection ▪ Forecasts/predictions ▪ Contextualized/relevant insights	Automatically finds, visualizes and narrates important findings (such as correlations, exceptions, clusters, drivers and predictions) in data that are relevant to users without requiring them to build models or write algorithms. Users may explore data via visualizations and conversational interfaces, including natural language query (NLQ) technologies, supported by natural language generation (NLG) narration and interpretation of results or the most statistically important findings in the user's context.
Augmented DSML	▪ Automated feature generation/selection ▪ Automated algorithm selection ▪ Automated model-tuning ▪ Automated model deployment and monitoring	Uses ML/AI to automate key aspects of data science and ML/AI modeling. Users may be business users or citizen data scientists looking to quickly prototype an idea. Highly skilled data scientists can also be more productive with less bias and have more time to focus on creative tasks and on building and operationalizing the most relevant models (see "How Augmented Machine Learning Is Democratizing Data Science").

Source: Gartner (October 2019)

Representative Vendors

The vendors listed in this Market Guide do not imply an exhaustive list. This section is intended to provide more understanding of the market and its offerings.

Market Introduction

The vendors and products listed and analyzed in this section are represented because they have achieved some level of visibility and traction in this market (see Note 1). Vendors are widely diverse in their capabilities. Gartner encourages data and analytics leaders to inspect the products under

consideration for the detailed functionalities included in each of the core capabilities as well as the potential broader capabilities outside the scope of this report (see "Critical Capabilities for Analytics and Business Intelligence Platforms" and "Critical Capabilities for Data Science and Machine Learning Platforms"). Table 2 contains an example list of vendors and the augmented capabilities they deliver by their respective product categories. It is possible, and increasingly common, for products to span both categories. See Note 3 for detailed definitions of each capability for this Market Guide.

Table 2. Augmented Analytics Capabilities by Vendor and Product Category

Vendor	Product, Service or Solution Name	Category	Specialist or Platform	Augmented ABI Category						Augmented DSML Category*			
				Automated Descriptive Insights	Key Driver Analysis	Auto-generated and Analyzed Segment/ Clustering	Automated Anomaly/ Outlier Alerting	Auto-generated and Analyzed Forecasts/ Predictions	Contextualized/ Relevant Insights	Automated Feature Generation/ Selection	Automated Algorithm Selection	Automated Model Packaging and Deployment	Automated Model Monitoring
Answer-Rocket	Answer-Rocket	Augmented ABI	Platform	X	X	X	X	X					
IBM	IBM Cognos Analytics	Augmented ABI	Platform	X	X								
Oracle	Oracle Analytics Cloud	Augmented ABI	Platform	X	X	X		X					
Outlier	Outlier	Augmented ABI	Specialist	X	X	X	X		X				
Qlik	Qlik Sense	Augmented ABI	Platform	X				X	X				
SAS	SAS Visual Analytics	Augmented ABI	Platform	X	X	X							
Sisense	Sisense	Augmented ABI	Platform	X				X					
Thought-Spot	Thought-Spot	Augmented ABI	Platform	X	X	X	X		X				

Vendor	Product, Service or Solution Name	Category	Specialist or Platform	Augmented ABI Category						Augmented DSML Category*			
				Automated Descriptive Insights	Key Driver Analysis	Autogenerated and Analyzed Segment/ Clustering	Automated Anomaly/ Outlier Alerting	Autogenerated and Analyzed Forecasts/ Predictions	Contextualized/ Relevant Insights	Automated Feature Generation/ Selection	Automated Algorithm Selection	Automated Model Packaging and Deployment	Automated Model Monitoring
TIBCO Software	TIBCO Spotfire	Augmented ABI	Platform	X	X								
Microsoft	Power BI	Augmented ABI and DSML	Platform	X	X	X			X	X	X		
Salesforce	Einstein Analytics Plus	Augmented ABI and DSML	Platform	X						X	X	X	
SAP	SAP Analytics Cloud	Augmented ABI and DSML	Platform	X	X					X	X		
Tellius	Tellius	Augmented ABI and DSML	Platform	X	X	X	X	X	X	X	X	X	
Aible	Aible	Augmented DSML	Specialist							X	X	X	X

Vendor	Product, Service or Solution Name	Category	Specialist or Platform	Augmented ABI Category						Augmented DSML Category*			
				Automated Descriptive Insights	Key Driver Analysis	Auto-generated and Analyzed Segment/Clustering	Automated Anomaly/Outlier Alerting	Auto-generated and Analyzed Forecasts/Predictions	Contextualized/Relevant Insights	Automated Feature Generation/Selection	Automated Algorithm Selection	Automated Model Packaging and Deployment	Automated Model Monitoring
Big Squid	Kraken	Augmented DSML	Specialist							X	X	X	
DataRobot	DataRobot	Augmented DSML	Platform							X	X	X	X
DataStories International	DataStories Platform	Augmented DSML	Specialist							X	X	X	
dotData	dotData Enterprise	Augmented DSML	Platform							X	X	X	
H2O.ai	H2O Driverless AI	Augmented DSML	Platform							X	X	X	X
	IBM Watson Studio	Augmented DSML	Platform							X	X	X	

Vendor	Product, Service or Solution Name	Category	Specialist or Platform	Augmented ABI Category						Augmented DSML Category*			
				Automated Descriptive Insights	Key Driver Analysis	Auto-generated and Analyzed Segment/Clustering	Automated Anomaly/Outlier Alerting	Auto-generated and Analyzed Forecasts/Predictions	Contextualized/Relevant Insights	Automated Feature Generation/Selection	Automated Algorithm Selection	Automated Model Packaging and Deployment	Automated Model Monitoring
Prevedere	ERIN	Augmented DSML	Specialist							X	X	X	
	SAS Visual Data Mining and Machine Learning	Augmented DSML	Platform							X	X		
	TIBCO Data Science	Augmented DSML	Platform							X	X	X	

*Note that many products delivering in the augmented DSML category may come with reports or explanations to assist citizen data scientists or even data scientists to better understand their data and automated predictive model outputs, but these are with respect to a predictive model building process (DSML) and therefore Xs will not be present under the augmented ABI capabilities category.

Source: Gartner (October 2019)

Vendor Profiles

Aible

Product: Aible

Category: Augmented DSML

Focus: Contextualized augmented DSML via business-user input optimized for business value, augmented data preperation

Website: www.aible.com

Aible enables business people to build machine learning models and to take into account the fact that all costs and benefits are not equal when assessing which model is best. Aible starts with how the model will impact your business and then assesses, ranks and allows for sensitive analysis of variables affecting the models based on business impact. The system uses business-oriented inputs given by a user to create the most impactful models based on those constraints, which users can adjust for what-if analysis after the models are built.

Aible has also introduced blueprints for common types of analyses — such as: maximize sales win rate, optimize marketing campaigns, predict customer churn, and predict inventory stock-out — to make it easier and quicker for business people to build and select models. Aible can also perform automated feature engineering to find new relationships in the data.

Once the Aible model is trained, it is then dockerized and can be deployed anywhere — in any application, on any cloud platform or on-premises, at the user's discretion. Aible can write predictions back to enterprise applications, such as Salesforce, for users to consume in the business applications which they work in.

AnswerRocket

Product: AnswerRocket

Category: Augmented ABI

Focus: Natural language search and automated insight generation combined with natural language generation (NLG)

Website: www.answerrocket.com

AnswerRocket is an augmented data discovery platform with native capabilities in NLQ, NLG and automated generation of insights. It enables users to ask questions and analyze data in natural language and then generates insights and a narration to explain the findings. Users can ask questions or enter keywords in a search box, or via voice. It delivers drill-down and full, advanced data exploration capabilities that can be shared or saved as a PDF, spreadsheet, or PowerPoint. As the user asks a question, AnswerRocket presents suggestions and surfaces relevant stories related

to the search topic. The queries can be sophisticated; for example, "How did my brand perform?" will automatically run a gradient boost algorithm to analyze and determine brand performance drivers. Dashboards can also be scheduled and refreshed on demand.

AnswerRocket can be run on-premises or in the cloud via major public cloud providers, and a number of supported data management solutions. The platform also supports the ability to plug in third-party AI and ML frameworks such as scikit-learn and TensorFlow. AnswerRocket has a focus (tailored content) around the consumer-packaged goods, and also supports the retail and wholesale, financial services, media and entertainment, and e-commerce verticals.

Big Squid

Product: Kraken

Category: Augmented DSML

Focus: Augmented DSML for business users, integrated what-if scenarios

Website: www.bigsquid.com

Big Squid's Kraken product augments many pieces of the machine learning model creation process, traditionally only able to be executed by expert-level data scientists. With Kraken, citizen data scientists and analysts can quickly connect to data and even other BI tools to begin building models.

Once connected, users simply have to choose a variable of interest and Kraken will handle the algorithm selection, model training and testing. Kraken supports univariate time series, regression and classification models. A visual representation of key drivers affecting the target variable is automatically displayed in rank order, allowing business users and other non-data scientists to quickly see what has the greatest impact on the target variable.

Data scientists have the ability to dive deeper into the analysis to see which algorithms and models were attempted, their model score and other model performance statistics.

DataRobot

Product: DataRobot

Category: Augmented DSML

Focus: Augmented DSML for citizen data scientists and data scientists

Website: www.datarobot.com/platform

DataRobot is an augmented data science and machine learning platform that automates key tasks across the end-to-end data science pipeline. The platform not only enables data scientists to work efficiently, but also enables citizen data scientists to build models quickly and easily.

DataRobot can connect to a multitude of data formats both on-premises and in the cloud and automatically visualize things like descriptive statistics and missing values. The platform will automatically perform operations like one-hot encoding, missing value imputation, text mining, standardization, and data partitioning. The platform transforms data into features that are optimized for each algorithm, and also allows users to include their own custom features if they have built them. DataRobot Automated Time Series integrates best practices in time series modeling, and can automatically detect stationarity and seasonality, as well as transform the target and implement back-testing to achieve the highest possible accuracy.

Operationalization of ML models is delivered through DataRobot MLOps. The solution allows for easy deployment on modern architectures like Kubernetes and Spark. DataRobot MLOps monitors for data drift, various ML model metrics, and infrastructure monitoring/alerting. Additional capabilities allow users to frequently update models, test new or competitive models and change applications on the fly without disrupting business applications or sacrificing governance best practices.

DataStories International

Product: DataStories

Category: Augmented DSML

Focus: Automated insight generation in the form of an interactive narrative

Website: datastories.com

DataStories leverages ML and AI techniques to automatically generate insights communicated through interactive stories. DataStories is an augmented DSML vendor that allows business domain experts, scientists and engineers to automatically analyze sensor and enterprise data to optimize complex R&D and manufacturing processes and performance. Whereas most augmented DSML platforms return models, DataStories takes a different approach — automatically generating interactive data stories so that domain experts can explore automated insights, key drivers of a KPI, and correlations and summary statistics of a data source. Data scientists can also leverage the underlying autogenerated predictive model.

Once data is loaded into the system, DataStories automatically runs a range of algorithms on the data. Model accuracy metrics and visualizations are displayed, along with the ability to upload new test data to see how the model performs. Results are presented in an interactive narrative that can be used to further explore and run sensitivity analyses on variables to optimize KPIs, or the story can be exported to Microsoft PowerPoint. The models can be exported to C/C++, Excel, Matlab, Python or R. Interactive reports can be shared through the organization, and interactive graphics can be exported as browser (HTML) files. DataStories SDK integrates with the Jupyter Notebook environment.

dotData

Product: dotData Enterprise

Category: Augmented DSML

Focus: AI-powered feature engineering including the ability to connect to multiple types of data beyond tabular datasets

Website: dotdata.com/dotdata-enterprise

The dotData Platform automates the entire data science process, by supporting multiple data formats, including relational data sources, and allowing for automation of 100% of the data science cycle. This includes automated generation of features as well as automated selection & optimization of machine learning models. ML models can be deployed into production using API-based integration.

dotData's AI-powered feature engineering derives features based not only on mathematical transformations, but also by blending domain expertise to maximize the feature importance and increase both model accuracy and relevancy. It automatically transforms source tables with complex relationships into a single "feature" table and makes the table ready for machine learning — typically exploring millions of features for predictive modeling.

The platform tests a multitude of different algorithms and automatically ranks the models based on either predictive power or various accuracy metrics. Easy-to-understand visualizations are generated for users to help them understand the importance and impact of the features in the model. API templates are automatically created to help easily move selected models to production.

H2O.ai

Product: H2O Driverless AI

Category: Augmented DSML

Focus: Ability to deploy explainable feature engineering and model selection pipeline in compact production-ready code, time series modeling, and NLP

Website: www.h2o.ai

H2O.ai offers the open-source H2O platform, along with additional open-source components including Sparkling Water for Spark integration, H2O4GPU and H2O AutoML. H2O.ai also offers a commercial product called H2O Driverless AI that uses automated machine learning.

H2O Driverless AI can automatically engineer new features for a given dataset. H2O Driverless AI can also automatically generate both Python scoring pipelines and new ultra-low-latency automatic scoring pipelines. The new automatic scoring pipeline is a unique technology that deploys all feature engineering and the winning machine learning model in a highly optimized, low-latency production-ready Java code or C++ (with Python & R runtimes) that can be deployed anywhere.

H2O Driverless AI delivers time series capabilities to optimize for almost any prediction time window. It incorporates data from numerous predictors, handles structured character data as well as high-cardinality categorical variables, and handles gaps in time series data and other missing

values. It also offers full explainability for every model produced using Shapley, K-LIME, PDP, disparate impact analysis and more.

IBM

Product: IBM Cognos Analytics; IBM Watson Studio

Category: Augmented ABI: Augmented DSML

Focus: Key driver analysis, natural language query with native NLG (IBM Cognos Analytics). Rapid ML experimentation for citizen data scientists (IBM Watson Studio)

Website: www.ibm.com/products/cognos-analytics

IBM Cognos Analytics combines production reporting capabilities with self-service dashboards and ad hoc analysis within one modern analytics and BI platform. IBM Cognos Analytics version 11.1 brought the former augmented analytics capabilities of IBM Watson Analytics into IBM Cognos Analytics. IBM Cognos Analytics is available both on-premises or as a hosted solution on the IBM Cloud.

The latest release brings augmented analytics capabilities, with the ability to explore what drives a particular metric. A unique network chart will display the relative importance of each driver. An AI assistant provides a conversational interface for users to pose questions, and to suggest other data sources and insights. An NLG capability has also been implemented.

The new AutoAI feature in IBM Watson Studio allows citizen data scientists to connect .csv files and automatically build a series of machine learning models. The AutoAI feature currently only supports classification and regression models. Models can be packaged and accessed via other IBM sources such as IBM Watson Machine Learning and IBM Watson OpenScale.

Microsoft

Product: Power BI

Category: Augmented ABI and DSML

Focus: Key driver analysis, augmented DSML via point and click ML model building available within Power BI

Website: powerbi.microsoft.com

Microsoft offers data preparation, visual-based data discovery, interactive dashboards and augmented analytics via a single product in Power BI.

Power BI Quick Insights can be run to automatically identify basic insights, correlations and potential outliers via a set of visualizations. AI visualizations in Power BI automatically surface key drivers affecting a given metric as well as segment data into interesting groups using decision trees.

Analysts are able to enhance their visualizations with forecasting and clustering. The Q&A feature also allows users to explore data in natural language.

Power BI integrates with Python, R and Azure ML. Power BI now also includes augmented DSML features including point and click options to build and use regression and classification models inside Power BI via dataflows. Features are automatically generated and up to 50 iterations of different models are tested and scored to find the best one for that use case. A report is created to display key model metrics and make the models more transparent. The model can be used to score new data in Power BI on a schedule and provide predictions as well as row-level explanations.

Oracle

Product: Oracle Analytics Cloud

Category: Augmented ABI

Focus: Integrated augmented analytics and NLG via an OEM of Yseop as well as an augmented analytics mobile application

Website: www.oracle.com/solutions/business-analytics/analytics-cloud.html

Oracle Analytics Cloud (OAC) includes integrated data preparation, data discovery (with advanced exploration and augmented analytics), interactive dashboards and scenario/what-if planning options. Oracle added augmented visualization using ML to uncover what drives results for attributes and measures, and to identify key statistical correlations and anomalies.

Augmented analytics features include automated insight generation and robust NLG. Within a view or dashboard, when selecting a variable in the query panel, insights are automatically generated via the "explain" feature to identify significant segments, clusters, drivers, outliers and anomalies. This also includes NLQ and narration of findings. These autogenerated insights can be added to a dashboard or explored further.

Oracle has recently announced availability of OAC with Autonomous Data Warehouse, which delivers in-database analytics and machine learning through Oracle Machine Learning (OML). OML provides capabilities for building, training, deploying and managing machine learning models and, through the upcoming OML for Python API, will also include advanced AutoML capabilities for feature and algorithm selection, and hyperparameter tuning. OML capabilities can be consumed directly from Oracle Analytics via SQL or REST APIs.

Outlier

Product: Outlier

Category: Augmented ABI

Focus: Autonomous anomaly detection

Website: outlier.ai

Outlier is an augmented analytics platform that monitors business data for unexpected changes and important patterns such as anomalies, trends, relationships and emerging behaviors. Instead of preparing dashboards or running queries when questions arise, Outlier combines automated pattern detection across any metrics or data with NLG. This alerts users to insights specific to their needs and provides an NLQ interface for further exploration of relationships and patterns.

Unlike traditional anomaly detection software, which identifies relevant spikes over time in a specific metric, Outlier also identifies emerging trends and patterns. For example, Outlier is able to look for new relationships across thousands of dimensions such as revenue, customer interest, website traffic, ad word spend, and warehouse demands across product lines to find opportunities and problems. Outlier learns user preferences based on their behavior and shares only the most important findings or stories (out of thousands of possible patterns) with each user.

Prevedere

Product: ERIN Predictive Analytics Cloud

Category: Augmented DSML

Focus: Time series modeling and forecasting, automatically identify correlations among internal and external data for predictive models

Website: www.prevedere.com

Prevedere is a time series modeling and forecasting platform that leverages augmented analytics to automate correlation, modeling and back-testing. The External Real-Time Insights (ERIN) Engine constantly monitors curated global external datasets — including government, consumer and purchasing indicators — as well as internal data brought from organizations to identify each company's unique set of leading indicators.

The service allows users to narrow the suggested list of indicators to the ones with the best fit, based on correlation thresholds as well as category tags associated with each indicator. Users can then automatically generate a forecasting model. Visualizations and metrics for model accuracy, and component contribution and relative importance of the explanatory variables are displayed as part of the output.

The platform has its own visualization capabilities and has a web API for use with Microsoft Power BI, Salesforce and Tableau, as well as an API for planning tools such as those from Oracle and SAP. There are plans to include proactive model monitoring to alert users when models may have degraded. Common customer use cases include strategic planning, marketing planning, sales planning, financial planning, demand planning and complex analyses involving internal and external data.

Qlik

Product: Qlik Sense

Category: Augmented ABI

Focus: Automated insight generation, conversational chatbot

Website: www.qlik.com/us/products/qlik-sense

Qlik Sense is a modern analytics and BI platform continuing to expand its augmented capabilities. Qlik Insight Advisor uses the Qlik Cognitive Engine to automatically create a set of visualizations. This was further improved with the added ability for users to adjust the autogenerated visualizations and the option for the Cognitive engine to learn from users' interactions with the product. Newer features include visual indicators for the expected degree of relevancy of an autogenerated visualization.

Qlik's Qlik Insight Bot is the result of their acquisition of CrunchBot, extending their augmented functionality to include conversational interfaces using NLP/NLG. Users are able to ask questions in natural language and the system will respond with autogenerated charts, dimension or measure lists, forecasts or predictions. Selections made via Qlik Insight Bot are persisted to the Qlik environment, and each new question maintains the context of the ongoing conversation.

Salesforce

Product: Einstein Analytics Plus

Category: Augmented ABI and DSML

Focus: Automated insights on datasets and Salesforce reports, transparent augmented DSML open to multiple user personas

Website: www.salesforce.com/products/einstein-analytics/overview

Salesforce offers Einstein Analytics Plus which includes the Einstein Analytics Platform, Einstein Prediction Builder and Discovery, and Data Insights (AI-automated insights for Salesforce reports). Einstein Analytics Plus uses AI and machine learning to generate suggestions for how to clean and prepare data.

Data Insights can automatically visualize and narrate important insights about a Salesforce report such as descriptive analytics, variances and correlations. This is also available to any datasets in Einstein Analytics Plus via the Einstein Discovery capabilities. Users can generate predictive models without requiring them to manually build models or write algorithms. This augmented DSML capability handles feature and model selection, and it also adds an additional tab to the automated insights explaining "why" that metric has changed. This includes key driver visualizations for the most important variables affecting the target with additional natural language explanation.

The underlying model is also open to data scientists via R code to validate. Salesforce is extending these capabilities by adding AI model bias detection and model explainability and accountability features. With the recent acquisition of Tableau, Salesforce will have access to additional natural language search functionality via Tableau's Ask Data offering.

SAP

Product: SAP Analytics Cloud

Category: Augmented ABI and DSML

Focus: Automated insight generation, key driver analysis, and augmented DSML

Website: www.sap.com/products/cloud-analytics.html

SAP Analytics Cloud is a purely cloud-based deployment running on SAP HANA within the wider SAP Cloud Platform, which is the underlying architecture. SAP Analytics Cloud combines data discovery, predictive analytics and planning and budgeting in an integrated package.

SAP Analytics Cloud's Smart Insights feature enables business users to gain insights automatically, including basic variances, correlations and the identification of key contributors. The insights are visualized with automatically selected graphs and explain the most important drivers of a particular metric. There is an "unexpected values" tab for outliers that have been identified, once the Smart Discovery feature has been activated. SAP has also added a natural language Q&A feature called Search to Insight.

SAP Smart Predict allows business users or citizen data scientists to build predictive models without being a data science expert. Users can select whether they are solving a classification, regression, or time series question, and Smart Predict will automatically generate the appropriate features and algorithm needed to solve that use case.

SAS

Product: SAS Visual Analytics; Visual Data Mining and Machine Learning

Category: Augmented ABI; Augmented DSML

Focus: Automated insight generation including key driver analysis and natural language generation (SAS Visual Analytics), template-based ML model creation for citizen data scientists and data scientists (SAS Visual Data Mining and Machine Learning)

Website: www.sas.com/en_us/software/visual-analytics.html

SAS Visual Analytics on SAS Viya combines reporting, data preparation, visual exploration and dashboards in a single product. SAS Visual Analytics is the core of SAS solutions and the company's data science product: SAS Visual Data Mining and Machine Learning (VDMML).

SAS has added automated insight generation via their Automated Explanation feature. Here, the product displays the most important variables influencing a metric with native NLG. Automated Explanation surfaces basic insights, variances, and trends, as well as using a series of machine learning models including decision trees to look for statistically relevant segments found in the data. Menu-driven advanced analytics, such as forecasting, clustering, correlation matrices and decision trees, have been core to the product since its first release. The underlying models used to

automatically generate visual representations of clusters, forecasts and predictions can be exported as code to be transferred to a data science environment or via an API for operationalization.

VDMML allows citizen data scientists and data scientists to take advantage of templates that automatically execute feature engineering and model selection for predictive model building. Other SAS offerings such as SAS Model Manager assist organizations with the deployment and ongoing monitoring of predictive models.

Sisense

Product: Sisense

Category: Augmented ABI

Focus: Automated anomaly detection

Website: www.sisense.com

Sisense offers data preparation, analytics and visual exploration of complex data mashups.

Augmented capabilities, such as Sisense Pulse, continue to make the Sisense user experience simple while analytically relevant, allowing a broader range of users to interact with the tool. Sisense Pulse is an alerting engine that uses ML techniques to detect anomalies and notify users when something needs their attention.

Sisense suggests possible relevant dimensions and measures that users may want to analyze. In addition, their Insight Miner feature allows users to automatically uncover relationships or correlations among their data with a single click.

Tellius

Product: Tellius

Category: Augmented ABI and DSML

Focus: Automated insight generation (including key driver analysis), natural language search combined with NLG, augmented DSML including predictive model selection

Website: www.tellius.com

Tellius is an augmented analytics platform that combines natural language search and conversational interface, automated discovery of insights, and augmented DSML with NLG explanation of results. The system supports complex analytical search terms such as "on the fly" comparisons; calculations; trends; growth; ranking, such as top and bottom performers; and insights such as "drivers" and "why" analyses.

Tellius leverages ML to automatically generate visual insights that highlight anomalies, correlations, underlying trends and change drivers. Recommendations are based on the user's question, what's statistically important in the data, and the user's learned behavior.

Tellius also has features that enable the citizen data scientist to create, train, modify and operationalize predictive models with APIs using augmented DSML features. Explainable AI is one of the core focus areas for Tellius. The platform will show which attributes (or features) most impacted a particular insight and model. Insights can be placed on a dashboard for further exploration, using typical interactivity approaches such as filtering, sorting, drilling and others, or via additional questioning and automated insight generation iteration.

ThoughtSpot

Product: ThoughtSpot

Category: Augmented ABI

Focus: Natural language search, automated insight generation

Website: www.thoughtspot.com

ThoughtSpot provides visual exploration, dashboards and augmented analytics in a single platform. Data is loaded in-memory into a massively parallel processing (MPP) engine and indexed for fast query performance. The search-based interface was an initial differentiator, but this has expanded to include automatic insight generation via SpotIQ, proactive alerting, and voice-driven queries via SearchIQ.

SpotIQ is pervasive throughout ThoughtSpot and can be invoked from entire datasets to specific charts or even specific subsets of data within a chart. The system will automatically generate insights and visualizations such as correlations, trends, and segments within the data, based on statistical relevancy and contextualization from prior user feedback and usage behavior. These visualizations can be pinned to dashboards. A new snapshotting feature will also compare how results may have changed over time to identify changes in behavior (such as a spike in sales or defects), with an alert sent via email or mobile message.

ThoughtSpot's platform allows users to indicate whether an autogenerated insight was beneficial or interesting using thumbs up/down icons. The system will learn from this feedback and better tailor the automated insights for that particular user in the future.

TIBCO Software

Product: TIBCO Spotfire; TIBCO Data Science

Category: Augmented ABI; Augmented DSML

Focus: Automated insight generation and some augmented data preparation capabilities (TIBCO Spotfire)

Website: www.tibco.com/products/tibco-spotfire

TIBCO Spotfire offers extensive capabilities for analytics dashboards, interactive visualization and data preparation in a single design tool and workflow. It also offers flexible processing options, either in-memory or in-database. Through acquisitions and OEM relationships, TIBCO Software has expanded its capabilities to include data science, ML and streaming analytics, location intelligence, data cataloging, and data virtualization. Its latest release has added automated insight generation, NLQ and augmented visual data preparation. It also includes enhancements to real-time analytics, and a single, simplified user interface.

TIBCO Spotfire will automatically analyze a user's data and suggest visualizations that are most relevant, based on the distributions and correlations of the data in the system. Users can drill into a suggested visualization of interest to have TIBCO Spotfire generate similar analyses. The same capability exists when searching the data via natural language.

TIBCO Data Science offers an AutoML node available as part of a ML pipeline interface that allows citizen data scientists to connect to modeled data (or as the next step in a data prep pipeline) that will automatically generate features, tune and select the best model of the multiple options the tool has run.

Market Recommendations

- Familiarize yourself with and monitor the augmented analytics capabilities and roadmaps of your analytics and BI, data science and ML, and self-service data preparation platform vendors, as well as emerging startups as they mature. Do so particularly in terms of the upfront setup required, the data preparation required, the types of data that can be analyzed, the types and range of algorithms supported, and the accuracy and explainability of the findings.

- Launch an augmented analytics pilot to assess viability and prove value. Start with a small list of specific business problems that are currently tackled manually and are time-intensive or prone to bias.

- Look for opportunities to use sandboxing and free trials to test and explore how augmented analytics complements existing data integration, analytics and BI, and data science initiatives.

- Investigate your current offerings being utilized today in the ABI and DSML space, as augmented analytics is here and will continue to be rolled out in the platforms you already have deployed over the next one to three years. Plan and invest in upskilling business people with data literacy.

- Engage both business analysts and data scientists in learning about and incorporating augmented analytics tools into the analytic process, in an effort to identify the best division of responsibility between the roles.

- Recognize that these tools will mature and evolve over the next couple of years. Consult "Cool Vendors in Analytics" and "Cool Vendors in Data Science and Machine Learning" and "Other Vendors to Consider for Modern Analytics and BI." Also, monitor moves by the vendors in "Magic Quadrant for Analytics and Business Intelligence Solutions" and "Magic Quadrant for Data Science and Machine Learning Platforms"

Gartner Recommended Reading

Some documents may not be available as part of your current Gartner subscription.

"Augmented Analytics Is the Future of Data and Analytics"

"Build a Comprehensive Ecosystem for Citizen Data Science to Drive Impactful Analytics"

"Augmented Analytics Feature Definition Framework"

"Cool Vendors in Analytics"

"Cool Vendors in Data Science and Machine Learning"

"Hype Cycle for Analytics and Business Intelligence, 2019"

"Hype Cycle for Data Science and Machine Learning, 2019"

"Magic Quadrant for Analytics and Business Intelligence Solutions"

"Magic Quadrant for Data Science and Machine Learning Platforms"

Note 1 Representative Vendor Selection

The 20 vendors named in this guide were selected to represent two of the three key pillars of augmented analytics solutions, as discussed in the Market Analysis section: augmented analytics and BI (augmented ABI) and augmented data science and machine learning (augmented DSML). The vendors selected best demonstrate the breadth of possible capabilities offered in the current market, both by vendors adding augmented capabilities to existing ABI or DSML platforms, as well as those building entire platforms around augmentation.

Note 2 Gartner's Initial Market Coverage

This Market Guide provides Gartner's initial coverage of the augmented analytics tools market and focuses on market definition, rationale and dynamics.

Note 3 Capability Definition

Augmented ABI and DSML tools can automatically or seamlessly deliver the following:

Table 3. Augmented Analytics Market Guide Capabilities Definitions

Category	Capability	Definition
Augmented ABI	Automated Descriptive Insights	The platform automatically surfaces basic insights such as variances, associations, correlations, or trends from a column or dataset before a content author begins actively exploring the data. Typically displayed as a brief natural language description or sample visualizations.
Augmented ABI	Key Driver Analysis	As part of a business user or business analyst workflow, the platform can automatically identify the most important key drivers (or most important attributes) of a given metric in a dataset.
Augmented ABI	Automated Anomaly/ Outlier Alerting	Automated alerting, notification or proactive display of outliers/anomalies based on changes in the business and data, in context of their role and business process. This is not manual threshold-based alerting or visual depiction of statistical outliers within a dashboard object.
Augmented ABI	Autogenerated and Analyzed Segments/ Cluster	The platform automatically finds new segments or clusters in a dataset. Menu-driven clustering as part of the visual data discovery process would not be included under this feature.
Augmented ABI	Autogenerated Forecasts/Predictions	The platform automatically creates a forecast or prediction. Menu-driven forecasting as part of the visual data discovery process would not be included under this feature.
Augmented ABI	Contextualized/ Relevant Insights	Automated insight generation that incorporates explicit and/or implicit usage and feedback data from users to display content that is most relevant for their role and use cases.
Augmented DSML	Automated Feature Generation/Selection	Automatically identifying the best types of data or variables to be used as part of the predictive model building process. Performing feature transformation, combinations, and creation automatically.
Augmented DSML	Automated Algorithm Selection	The process of selecting the appropriate algorithm to fit a given use case. Multiple algorithms may be tested, scored and ranked, ultimately choosing the most accurate model. Adjusting the hyperparameters of an algorithm to improve accuracy and optimize the predictive model.
Augmented DSML	Automated Model Packaging/ Deployment	The ease and speed with which the user can move models from a developer environment to a production environment or embed them in a business process. Platforms should support the ability to create APIs or containers (such as code, PMML and packaged apps) that can be used for faster deployment in business scenarios.
Augmented DSML	Automated Model Monitoring	The ability to track the performance of models in production to ensure the relationships found during development are still valid, and that the model is predicting well.

Source: Gartner (October 2019)

GARTNER HEADQUARTERS

Corporate Headquarters
56 Top Gallant Road
Stamford, CT 06902-7700
USA
+1 203 964 0096

Regional Headquarters
AUSTRALIA
BRAZIL
JAPAN
UNITED KINGDOM

For a complete list of worldwide locations,
visit http://www.gartner.com/technology/about.jsp