September 5, 2024

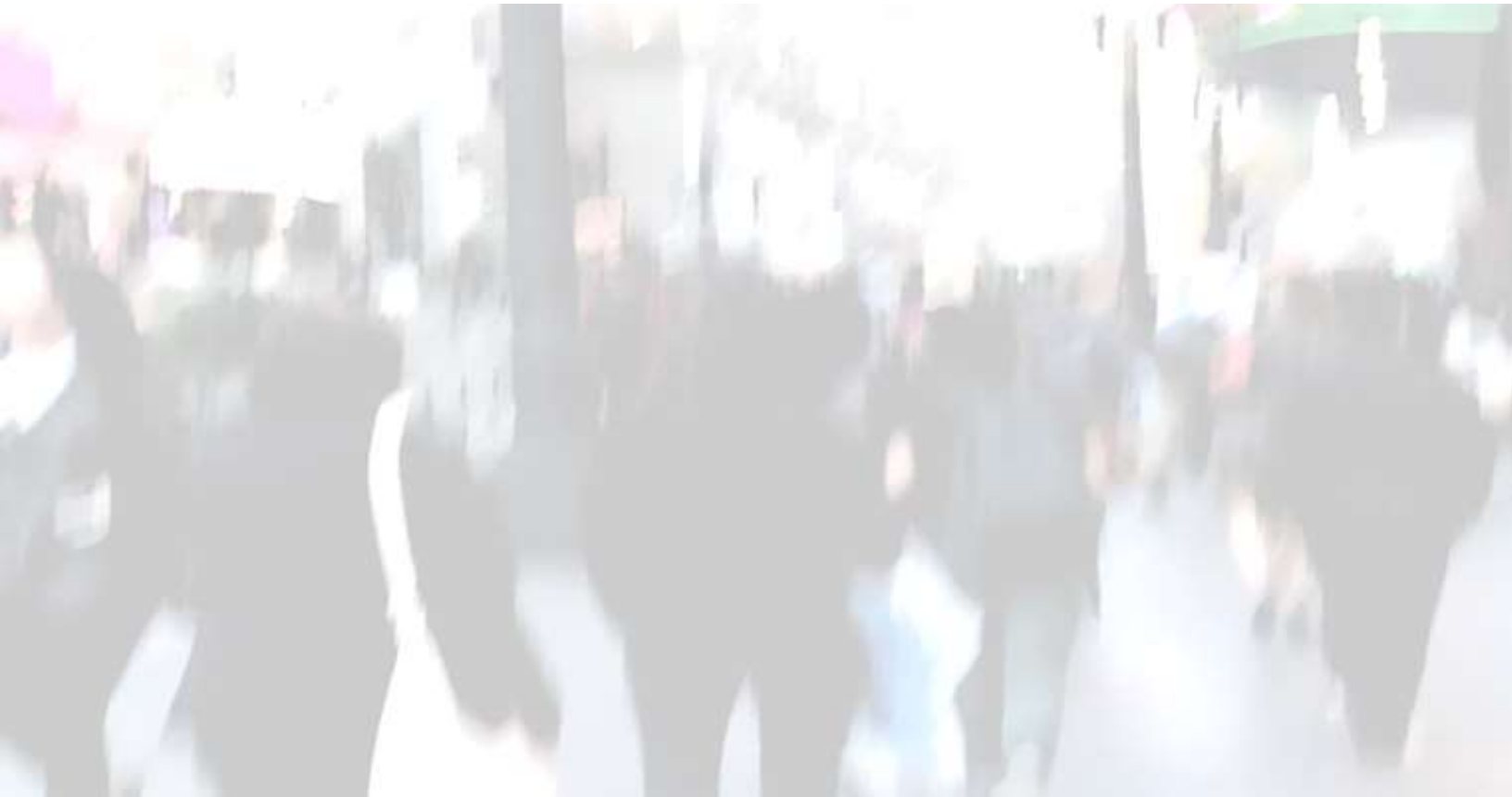
Dresner Advisory Services, LLC

2024 Edition

AI, Data Science, and Machine Learning Market Study

Wisdom of Crowds® Series

Licensed to Palantir



Disclaimer:

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

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

Definitions

Business Intelligence Defined

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

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

Howard Dresner, *The Performance Management Revolution: Business Results Through Insight and Action* (John Wiley & Sons, 2007).

AI, Data Science, and Machine Learning Defined

“AI, data science, and machine learning” includes statistics, modeling, machine learning, neural networks, and data mining to analyze facts to make predictions about future or otherwise unknown events.

Introduction

As we enter our 18th year as an independent, objective, and data-driven research organization, our commitment to providing insights into data, analytics, and performance management remains steadfast. This year's report marks the 11th edition of our annual research into AI, Data Science, and Machine Learning (AI / DS / ML), reflecting the growing significance of these technologies as organizations seek to enhance operations, improve forecasting, and drive innovation.

Our research shows that while the current deployment of AI, DS, and ML use cases remains limited but diverse, and there is a strong expectation of future growth. Key use cases such as customer segmentation, demand forecasting, and customer lifetime value are highlighted as areas with significant potential.

However, despite historic gains, the momentum in AI, DS, and ML deployment plateaued over the last three years. Production deployments continue to be most prevalent in research and development, particularly within larger organizations, with healthcare and consumer services leading the way as early adopters.

A notable trend in this year's findings is the rapid expansion of generative AI. Organizations are increasingly incorporating generative AI into their operations, with a marked rise in production use, experimental applications, and planned adoption over the next 12 months.

As always, our research aims to provide valuable, actionable insights that empower organizations to make informed decisions in their AI, DS, and ML initiatives.

We hope you enjoy this report!

Best,

Howard Dresner
Chief Research Officer
Dresner Advisory Services

Contents

Benefits of the Study

The Dresner Advisory Services AI, Data Science, and Machine Learning Market Study provides a wealth of information and analysis, offering value to both consumers and producers of business intelligence technology and services.

Consumer Guide

As an objective source of industry research, consumers use the Dresner Advisory Services AI, Data Science, and Machine Learning Market Study to understand how their peers leverage and invest in business intelligence and related technologies.

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

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

Supplier Tool

Vendor Licensees use the Dresner Advisory Services AI, Data Science, and Machine Learning Market Study in several important ways. For example:

External Awareness

- Build awareness for the business intelligence market and supplier brand, citing Dresner Advisory Services AI, Data Science, and Machine Learning Market Study trends and vendor performance
- Create lead and demand generation for supplier offerings through association with the Dresner Advisory Services AI, Data Science, and Machine Learning Market Study findings, webinars, etc.

Internal Planning

- Refine internal product plans and align with market priorities and realities as identified in the Dresner Advisory Services AI, Data Science, and Machine Learning Market Study
- Better understand customer priorities, concerns, and issues
- Identify competitive pressures and opportunities

About Howard Dresner and Dresner Advisory Services

The Dresner Advisory Services AI, Data Science, and Machine Learning Market Study was conceived, designed, and executed by Dresner Advisory Services, LLC—an independent advisory firm—and Howard Dresner, its Founder and Chief Research Officer.

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

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

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

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

- Wisdom of Crowds® Flagship BI Market Study
- Active Data Architecture™
- Analytical Data Infrastructure
- Analytical Platforms
- Data and Analytics Governance
- Data Engineering
- Embedded Business Intelligence
- Generative AI
- ModelOps

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

About Jim Ericson

Jim Ericson is Vice President and Distinguished Analyst with Dresner Advisory Services.

Jim has served as a consultant and journalist who studies end-user management practices and industry trending in the data and information management fields.

From 2004 to 2013 he was the editorial director at *Information Management* magazine (formerly *DM Review*), where he created architectures for user and industry coverage for hundreds of contributors across the breadth of the data and information management industry.

As lead writer, he interviewed and profiled more than 100 CIOs, CTOs, and program directors in a program called “25 Top Information Managers.” His related feature articles earned ASBPE national bronze and multiple Mid-Atlantic region gold and silver awards for Technical Article and for Case History feature writing.

A panelist, interviewer, blogger, community liaison, conference co-chair, and speaker in the data-management community, he also sponsored and co-hosted a weekly podcast in continuous production for more than five years.

Jim’s earlier background as senior morning news producer at NBC/Mutual Radio Networks and as managing editor of MSNBC’s first Washington, D.C. online news bureau cemented his understanding of fact-finding, topical reporting, and serving broad audiences.

The Dresner Team

About Elizabeth Espinoza

Elizabeth is Research Director at Dresner Advisory and is responsible for the data preparation, analysis, and creation of charts for Dresner Advisory reports.

About Kathleen Goolsby

Kathleen is Senior Editor at Dresner Advisory ensuring the quality and consistency of all research publications.

About Danielle Guinebertiere

Danielle is the Director of Client Services at Dresner Advisory. She supports the ongoing research process through her work with executives at companies included in Dresner market reports.

About Michelle Whitson-Lorenzi

Michelle is Client Services Manager and is responsible for managing software company survey activity and our internal market research data.

Survey Method and Data Collection

As with all our Wisdom of Crowds® market studies, we constructed a survey instrument to collect data and used social media and crowdsourcing techniques to recruit participants.

Data Quality

We carefully scrutinized and verified all respondent entries to ensure that only qualified participants were included in the study.

Executive Summary

Executive Summary

- The role of *BI expert* is the most likely *constant* or *often* user of AI, DS, and ML, followed by *statistician / data scientist* and *business analyst.* User activity is at historically average or slightly lower levels (fig. 5-11).
- AI, DS, and ML ranks 22nd, and *cognitive BI / artificial intelligence-based BI* ranks 32nd among 63 topics under our study. Historic importance remains in a relatively flat range well above the level of *important* (fig. 12-18). Industry importance is at a surging all-time high, very-near *critical* consensus of importance (fig. 69).
- Current deployment of use cases for AI, DS, and ML remains very low but broad, with predictions of strong future uptake. The most important include c*ustomer segmentation*, *demand forecasting,* and *customer lifetime value* (fig. 19-26).
- Current deployment and future plans for AI, DS, and ML show historic gains that flattened during the last three years. In-production deployments remain highest in *R&D* and in larger organizations, with *healthcare* and *consumer services* early movers (fig. 27-33).
- Longevity of AI, DS, and ML usage, formerly found more often in longstanding programs, is giving way to more "newer" users in "younger" programs, most often found in North America (fig. 34-38).
- The most important features for AI, DS, and ML include *range of regression models*, *outlier detection*, and *model explain-ability.* Feature interest is above historic levels. (fig. 39-44). Industry feature support is very strong (fig. 70-71).
- The most important data preparation capabilities include *detecting duplicates or outliers*, *cleansing and enriching source data* and *set operations* (fig. 45-47). Industry support for data preparation is extremely strong (fig. 73).
- The top usability features include *support for easy iteration*, *low code / no code with the ability to modify* and *access to advanced analytics* (fig. 48-52). Industry support for usability is very strong (fig. 74).
- The most important scalability features are *in-database analytics* and *in-memory analytics* (fig. 53-56). Industry support is strong (fig. 75).
- The most important neural networks include *artificial neural network* and *transformer networks* (fig. 57-58). Industry support is incomplete (fig. 72).
- Current use and planned adoption of generative AI is expanding rapidly, with far more in-production, experimental use, and 12-month plans (fig. 59-64).
- The use of open-source AI, DS, and ML is increasing, but not advancing rapidly. Top features are *Databricks* and *Azure Data Factory*. Top technologies include *Pandas and Tensorflow* (fig. 65-67). (Industry support fig. 77-79.)
- Top data sources include *Snowflake*, *Amazon S3*, and *Postgres* (fig. 68, 80).
- AI, DS, and ML vendor ratings are on page 98.

Study Demographics

Study Demographics

Our study includes a cross-section of respondents across geographies, functions, organization sizes, and vertical industries. We believe that, unlike other industry research, this supports a more representative sample that is a better indicator of true market dynamics. We constructed cross-tab analyses using these demographics to identify and illustrate important industry trends.

Geographies

North America (which includes the United States, Canada, and Puerto Rico) represents the largest group in our 2024 user study, with about 52 percent of all respondents. EMEA is the second-largest group (29 percent), followed by Asia Pacific (13 percent) and Latin America (6 percent) (fig.1).

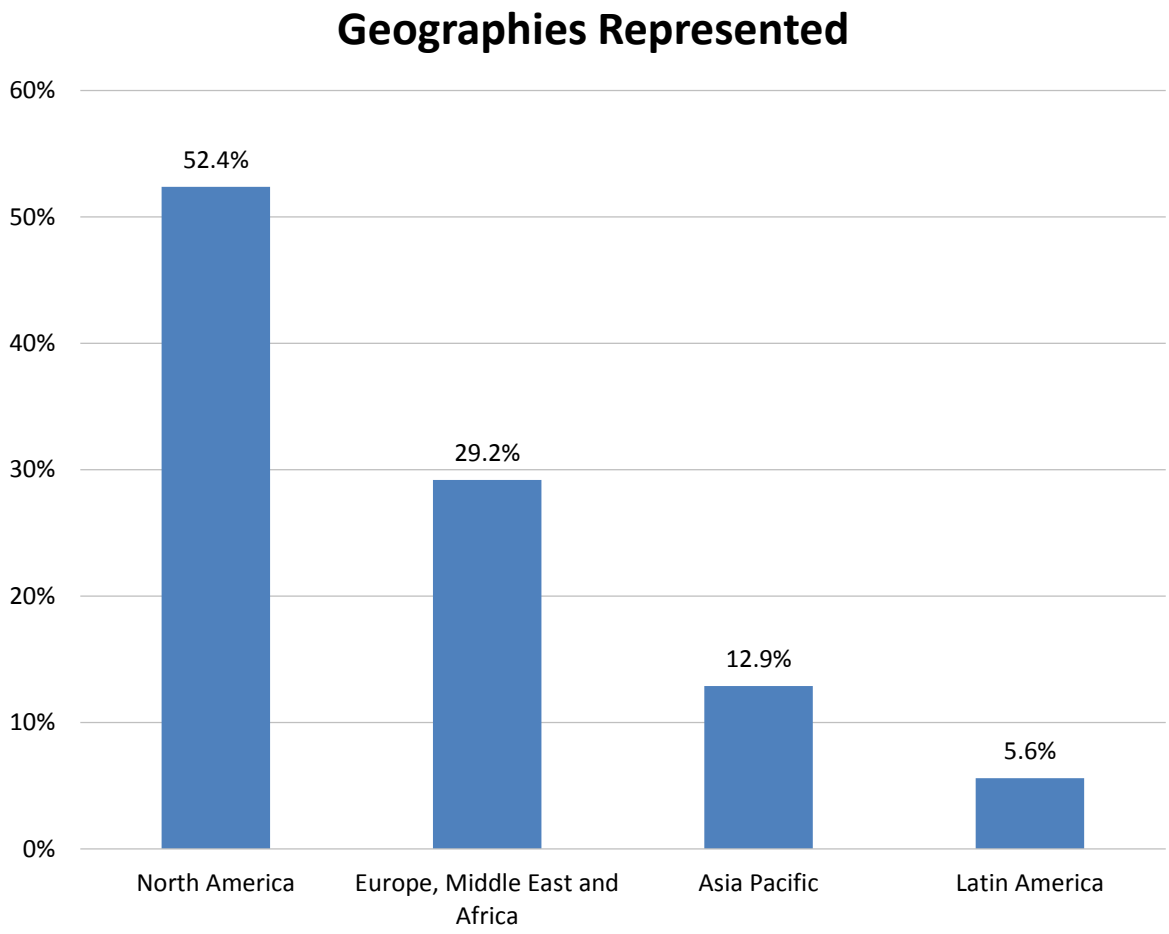


Figure 1 – Geographies represented

Functions

Our 2024 user sample includes a cross section of functional roles in respondent organizations. *Information technology* (25 percent) is the most represented, followed by *executive management* (20 percent), the *BICC* (17 percent), and *finance* (16 percent) (fig. 2).

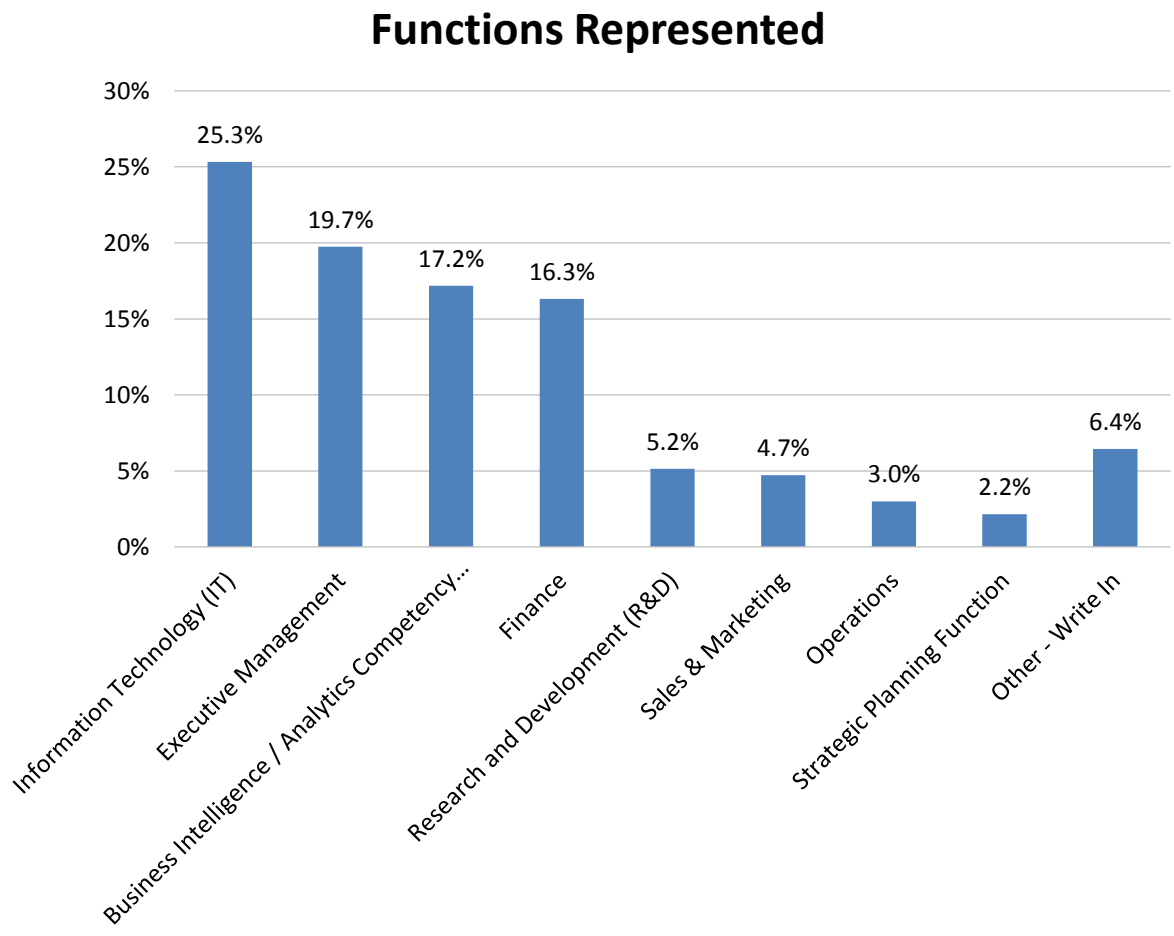


Figure 2 – Functions represented

Vertical Industries

Business services (19 percent), *technology* (17 percent), and *financial services* (15 percent) lead industry participation in our 2024 user sample (fig. 3). *Manufacturing* (14 percent), *healthcare* (11 percent), and *consumer services* (10 percent) are the next most represented.

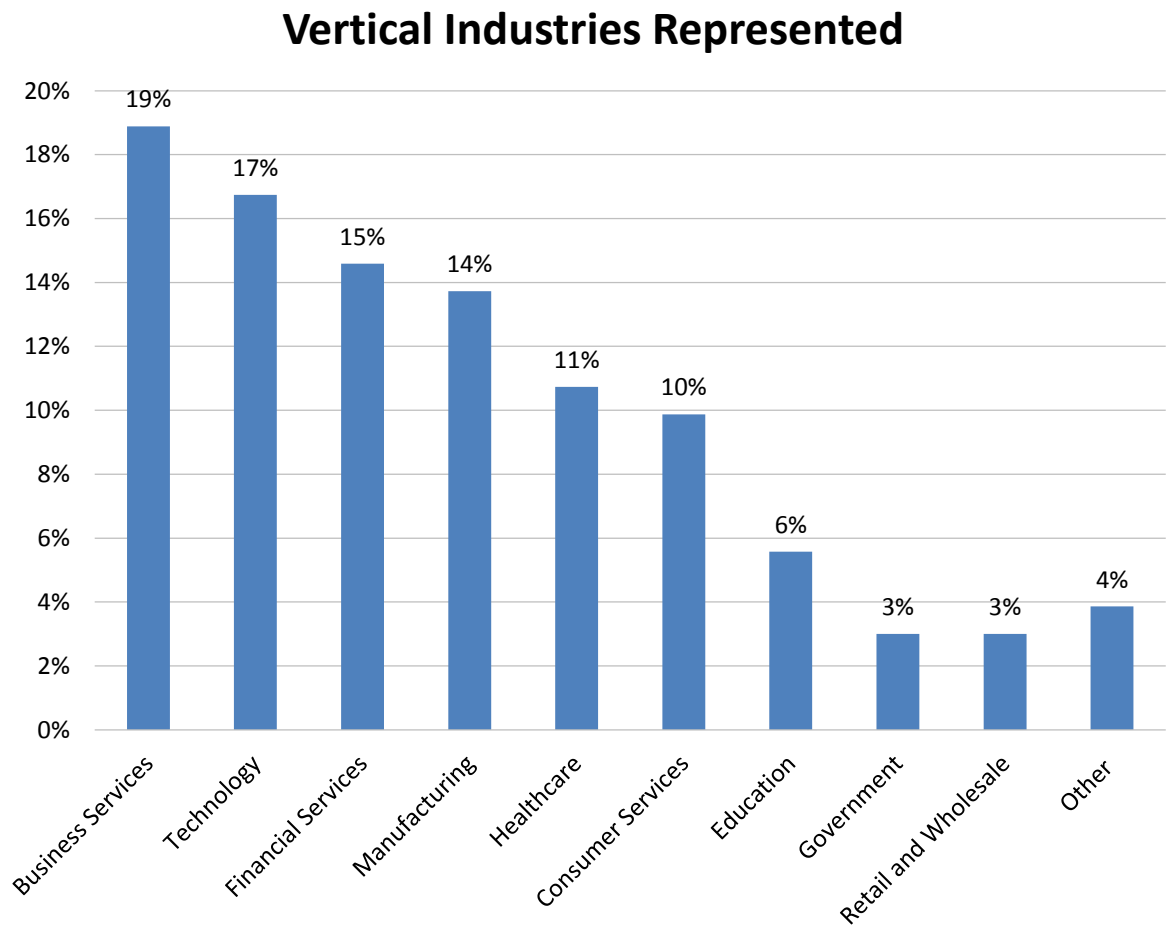


Figure 3 – Vertical industries represented

Organization Sizes

Our 2024 user sample includes a mix of small, medium, and large organizations (fig. 4). Small organizations (1-100 employees) and midsize organizations (101-1,000 employees) account for 26 percent and 31 percent of the sample, respectively. Large organizations (>1,000 employees) account for the remaining 44 percent. The chart below further breaks these organizations out by global headcount.

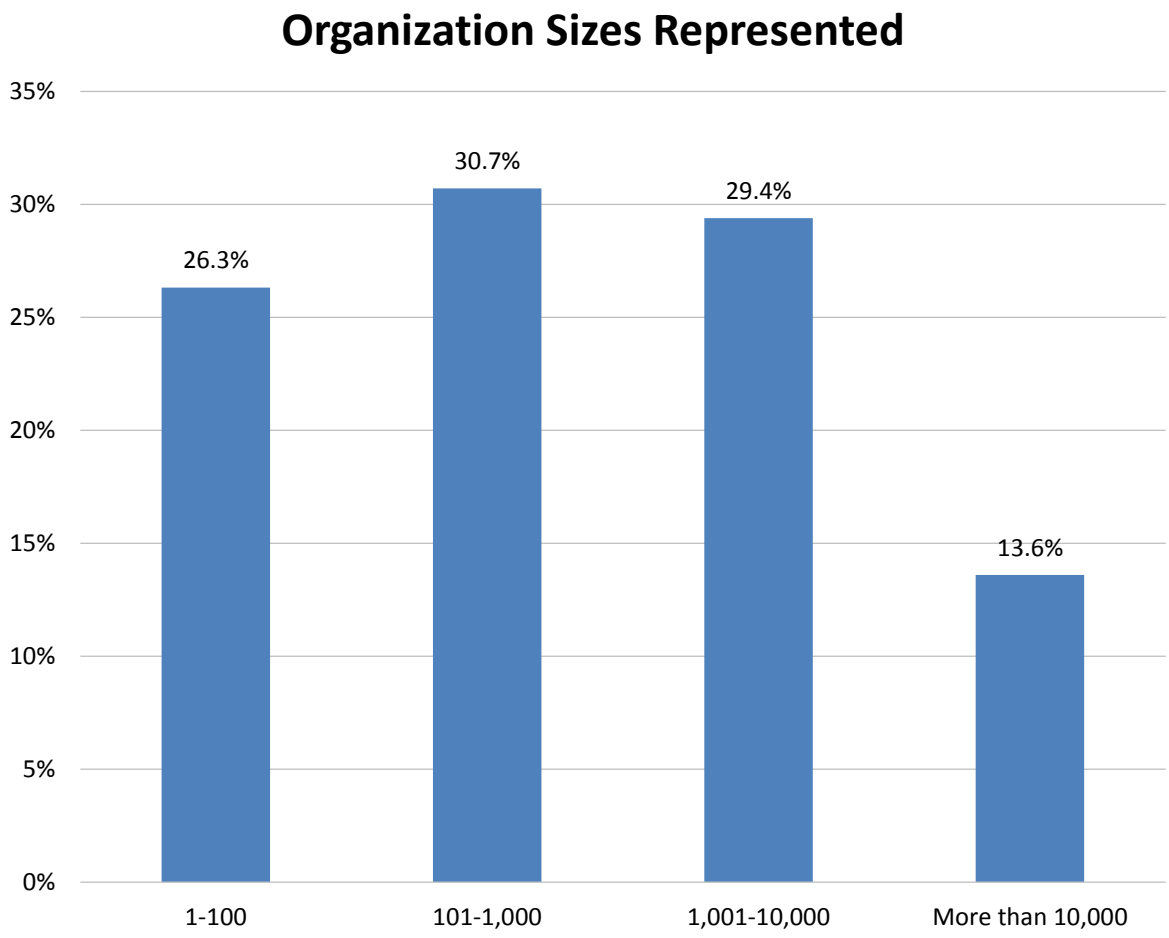


Figure 4 – Organization sizes represented

Analysis and Trends

Analysis: AI, Data Science, and Machine Learning

Users of AI, Data Science, and Machine Learning

Fig. 5 describes the functional users of AI, data science, and machine learning. We include *citizen data scientist*, a role that might overlap with other users but generally describes users that can generate models. In 2024, the role of *BI expert* is the most likely *constant* or *often* user of AI, data science, and machine learning (68 percent), followed by *statistician / data scientist* (65 percent), *business analyst* (59 percent), and *citizen data scientist* (34 percent). *Constant* use thereafter drops to below 20 percent among *IT staff, marketing analysts, financial analysts*, and *C-level executives*. AI, data science, and machine learning in 2024 is rarely outsourced, given that third-party consultant use is lowest of all functions sampled. Nonetheless, even the least likely user groups are more than 40 percent and top users up to 88 percent likely to use AI, data science, and machine learning at least *occasionally*.

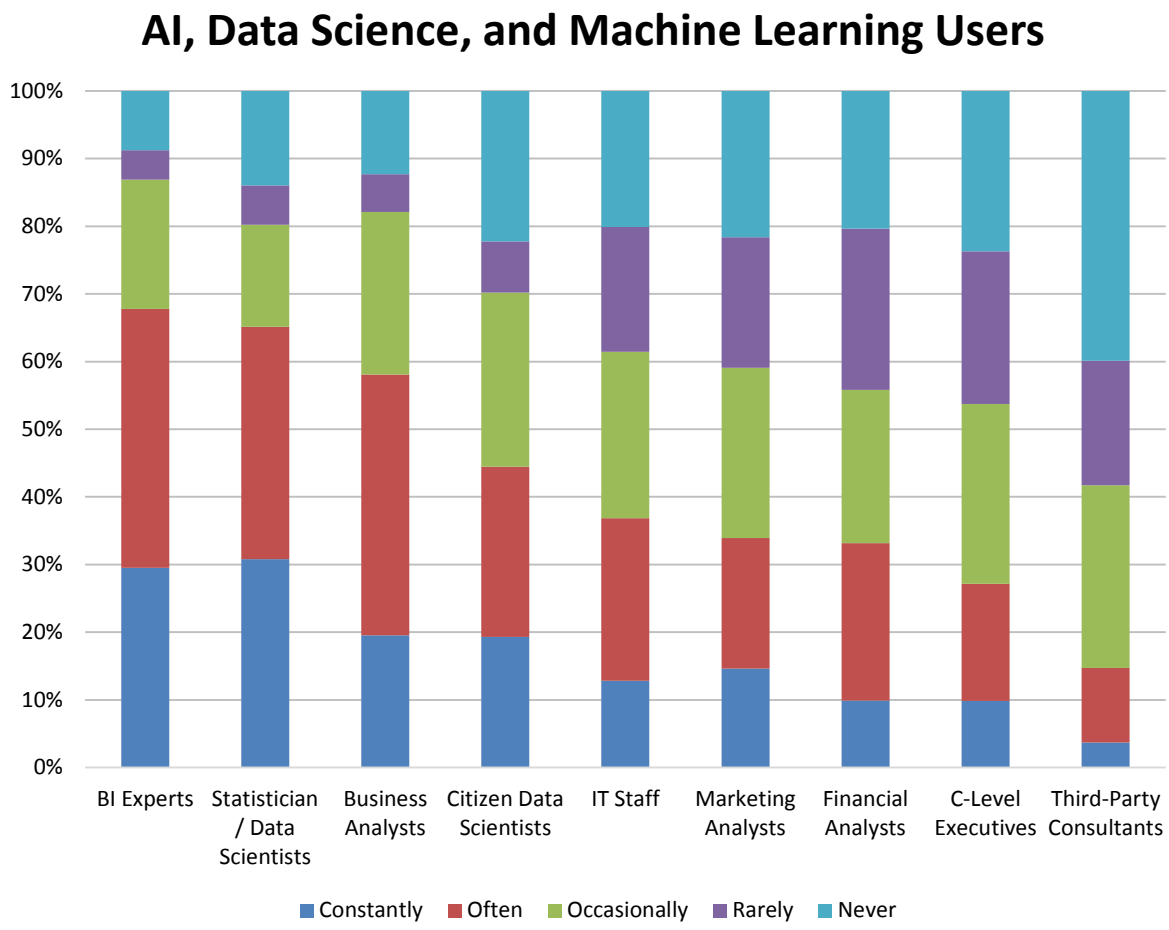


Figure 5 – AI, data science and machine learning users

Users of AI, Data Science, and Machine Learning Users 2018–2024

Viewed across seven years of data collection, most user roles for AI, data science, and machine learning report average to below-average historic levels of use and all user roles are below all-time-high levels of use (fig. 6). It is difficult to attribute declining historic use among some roles, notably *statistician / data scientist, citizen data scientist, financial analyst*, and *third-party consultants* to a particular cause. We observe that, in 2024, *BI expert* narrowly displaces *statistician / data scientist* as the top user role, and that *IT staff and marketing analysts* are more frequent users than in the past.

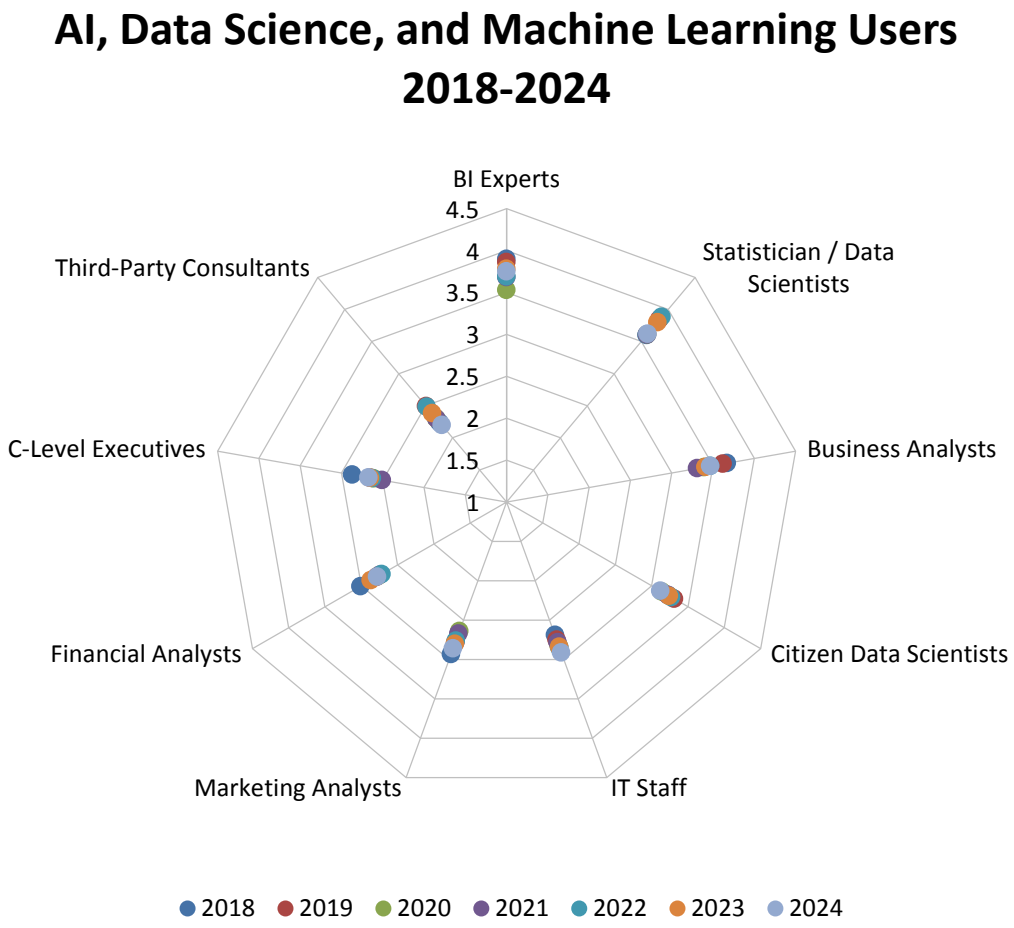


Figure 6 – AI, data science, and machine learning users 2018-2024

Users of AI, Data Science, and Machine Learning by Geography

Viewed by geography, all roles receive the highest 2024 importance scores from respondents in Latin America and Asia Pacific (fig. 7). Scores for the top seven choices (*BI expert*, *statistician / data scientist*, *business analyst*, *citizen data scientist, IT staff, marketing analysts*, *and financial analysts*) and all have scores corresponding to at least *important* across all regions. Notably, respondents in EMEA and North America are least likely to name *third-party consultants* and *executive management* as users and assign lower than *important* scores to these roles.

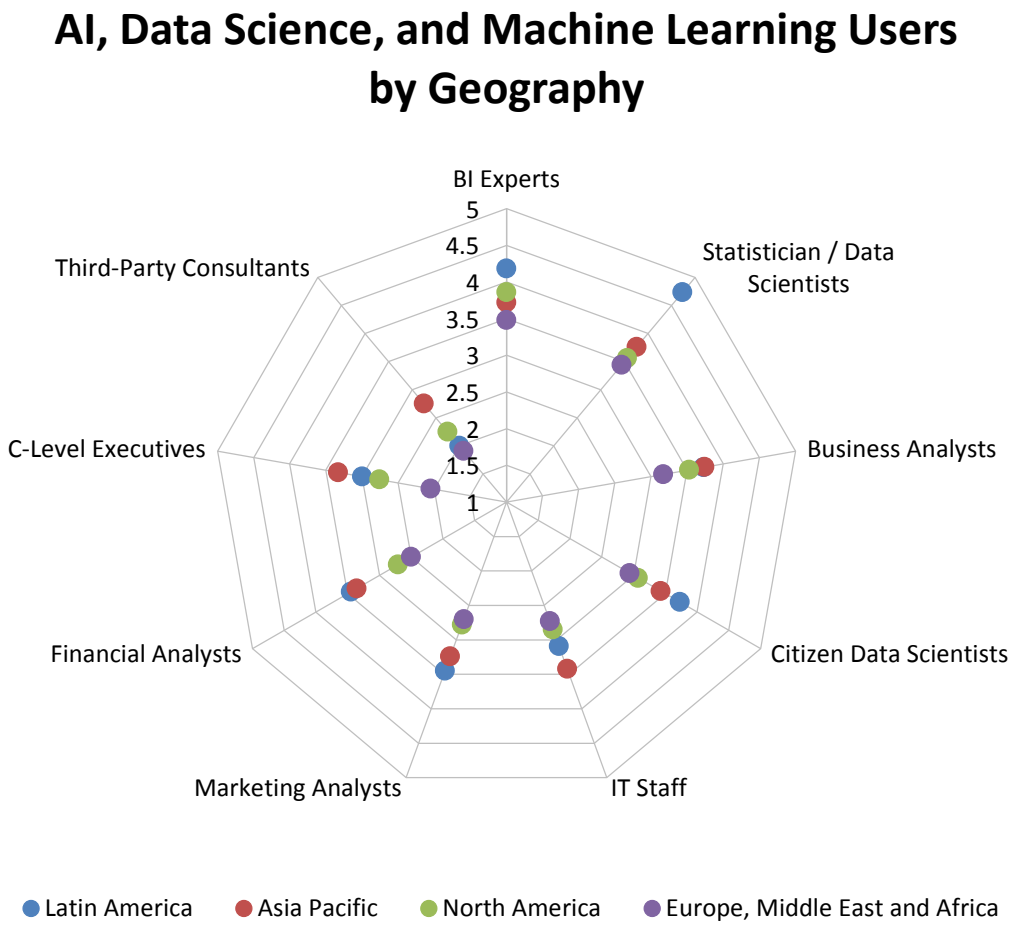


Figure 7 – AI, data science, and machine learning users by geography

Users of AI, Data Science, and Machine Learning by Organization Size

The likelihood of various and multiple user roles for AI, data science, and machine learning almost always correlates directly with increasing organization size in 2024 (fig. 8). Roles that are noticeably more important in very large organizations (> 10,000 employees), include *citizen data scientist* and *IT staff.* In very large organizations, *BI expert* is the only role with the distinction of greater than 4.0 criticality, signifying *very important*. Small (1-100 employees) and midsize (101-1,000 employees) organizations criticality begin to fall below the level of *important* with *citizen data scientist, IT staff,* and various departmental *analyst* and *executive* roles.

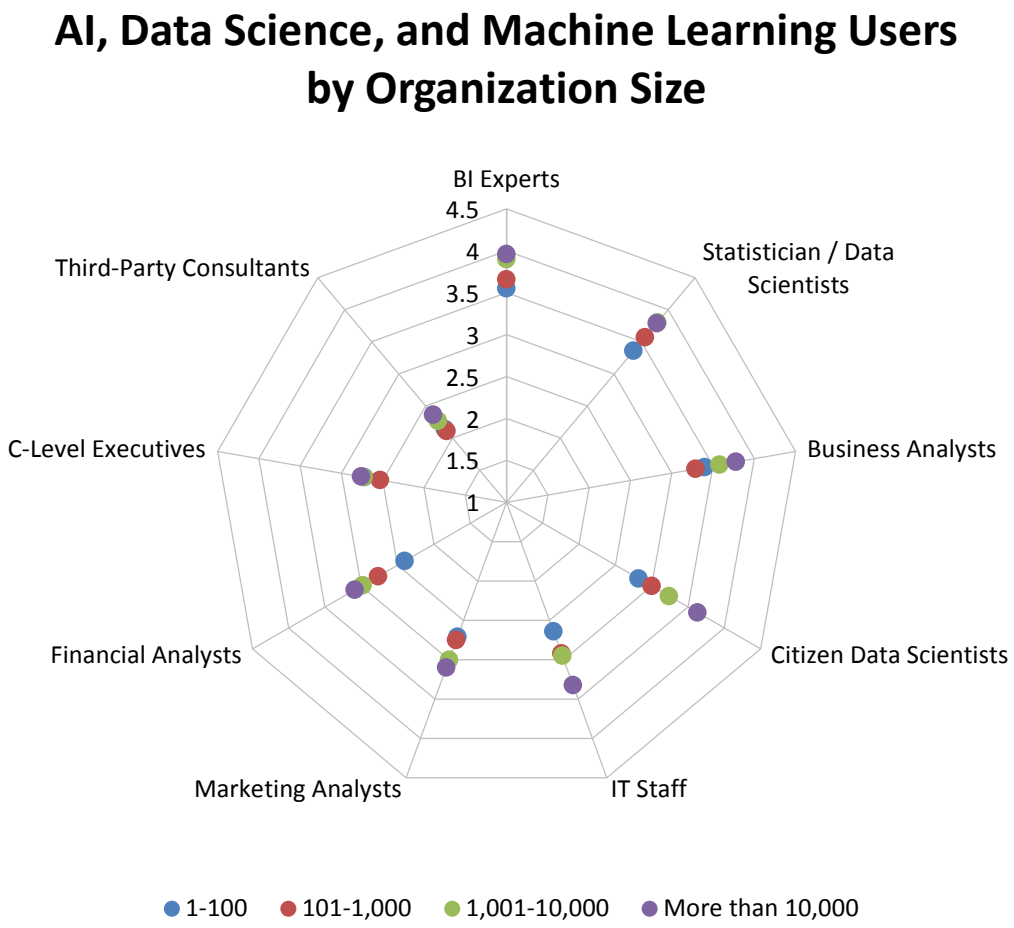


Figure 8 – AI, data science, and machine learning users by organization size

Users of AI, Data Science, and Machine Learning by Industry

Taken in sum, multiple industries assign an uneven mix of high to medium and occasionally lower usage by roles for AI, data science, and machine learning in 2024 (fig. 9). This year, the highest overall mean scores are reported by respondents in *consumer services, technology*, and *business services.* That said, respondents in *healthcare* assign the highest scores (4.0, *very important*) to the two top roles, *BI experts* and *statistician / data scientist,* but average or lower scores to most other roles. *Government* and *education* respondents are generally least interested in nearly all user roles.

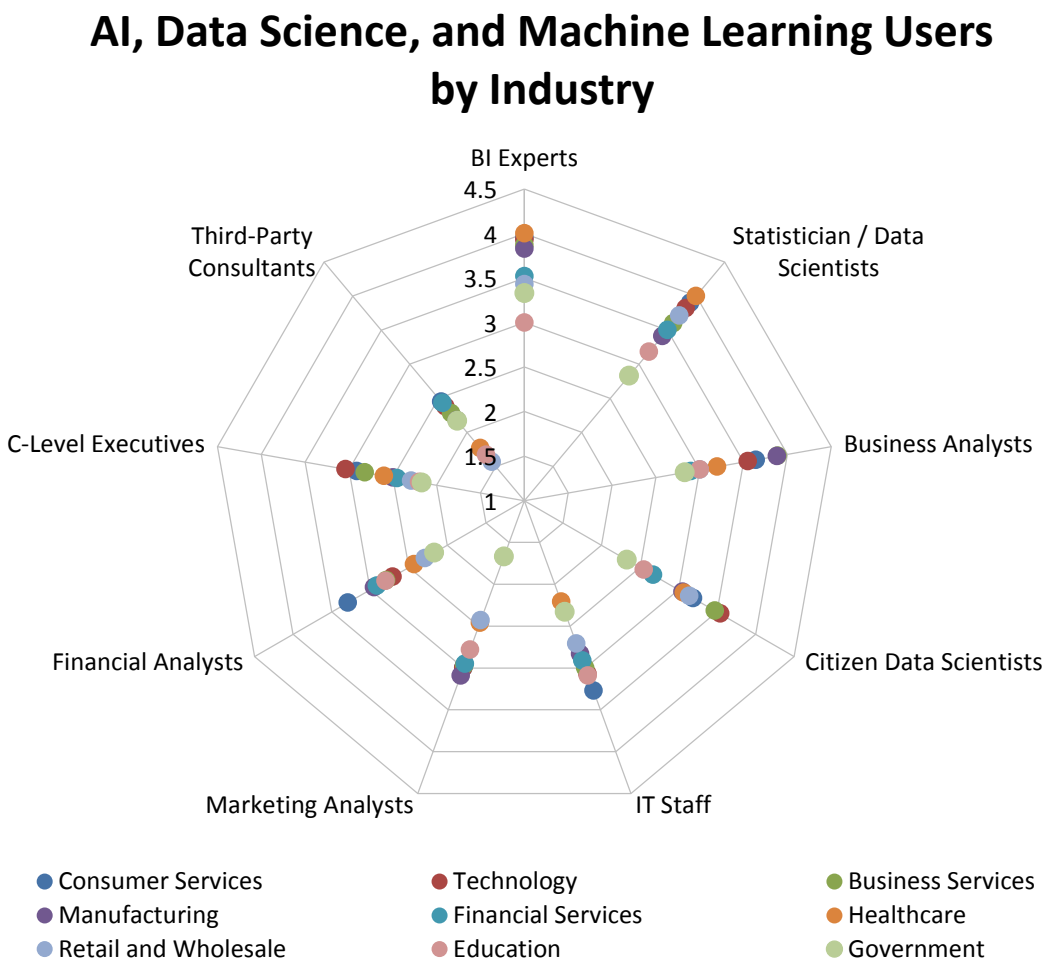


Figure 9 – AI, data science, and machine learning users by industry

Users of AI, Data Science, and Machine Learning by Company Age

The frequency of data science and machine learning use by role varies unevenly by company age in 2024 (fig. 10). This year, *BI* experts are (at least equally) the most likely active users in all organizations regardless of age. Organizations of 11-16 years appear to be the most likely active users across multiple roles, with the highest or near-high scores for *BI expert, statistician / data scientist, business analyst, marketing analyst, financial analyst, C-level executive*, and *third-party consultants. Statistician / data scientist* use (which held the overall top spot last year), is somewhat more common in organizations of 5-10 years. Younger organizations of less than five years are more likely to name *BI experts* and *business analysts* as AI, data science, and machine learning users (compared to other roles).

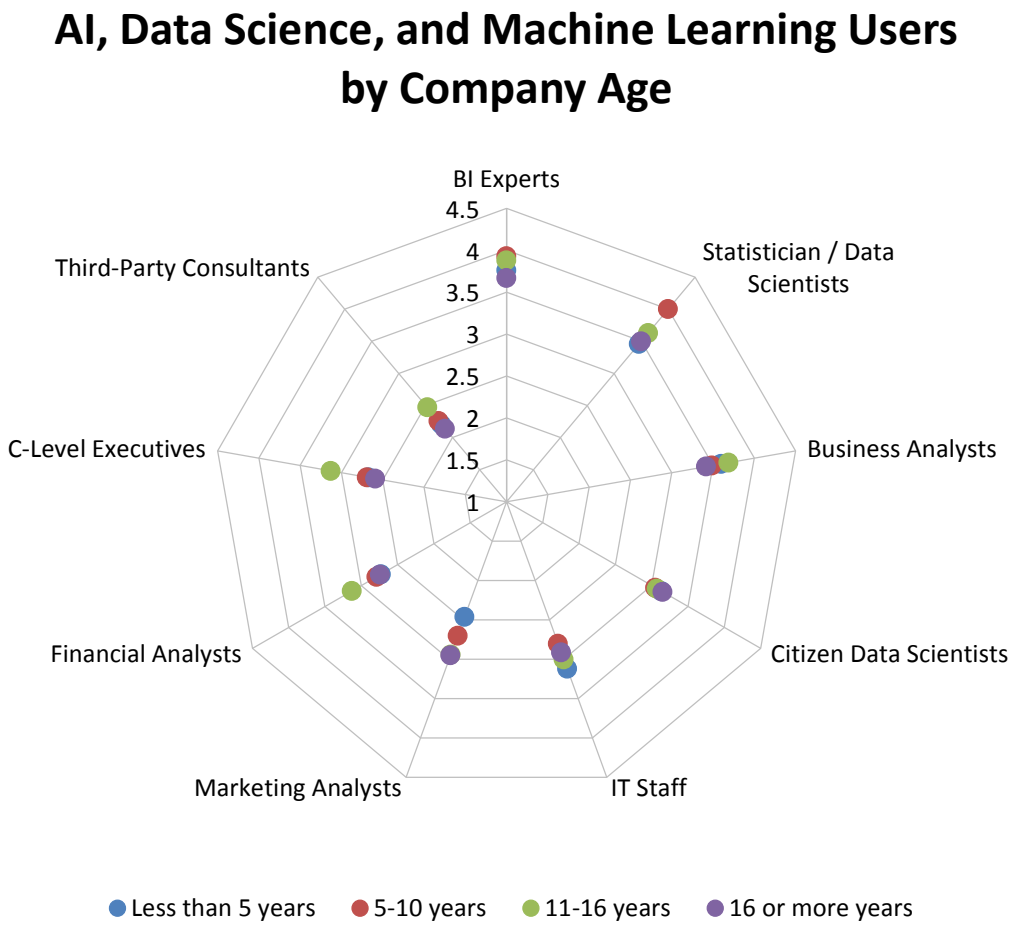


Figure 10 – AI, data science and machine learning users by company age

Users of AI, Data Science, and Machine Learning by Success with BI

The use of AI, data science, and machine learning variably correlates to success with BI according to role and frequency of use in 2024 (fig. 11). Organizations that are *completely successful* or *somewhat successful* with BI, lead or rank equally highest in the top four user roles of *BI expert*, *statistician / data scientist*, *business analyst*, and *citizen data scientist,* along with *financial analysts. Completely successful* BI organizations place lower emphasis on *IT staff* and *marketing analysts* compared to less-successful peers. Findings are further colored by other demographics such as organization size (see previous chart). Other results, such as lower success correlated with the use of *marketing analysts,* might indicate early attempts or missteps in use by that function.

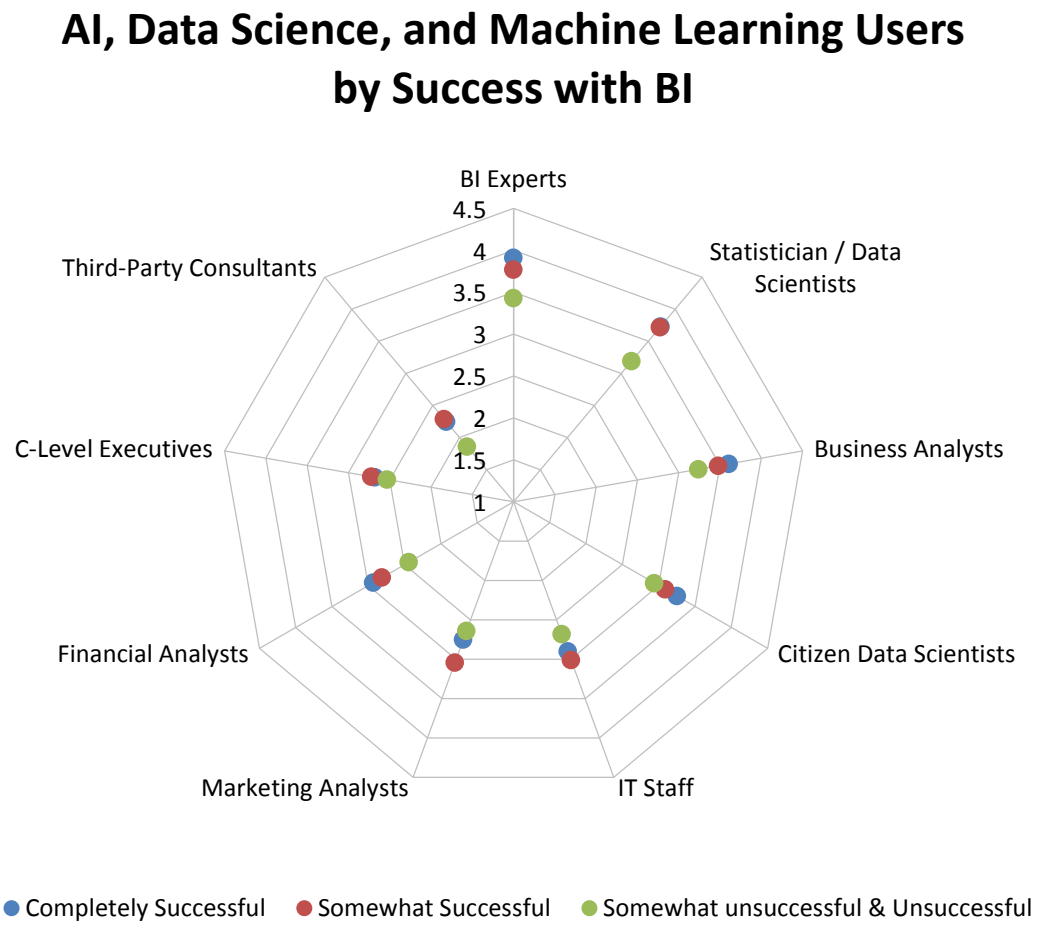


Figure 11 – AI, data science and machine learning users by success with BI

Importance of AI, Data Science, and Machine Learning

AI, data science, and *machine learning* ranks 22nd, and *cognitive BI / artificial intelligence-based BI* ranks 32nd among 63 topics under our study, as defined in our 2024 user survey. Generative AI ranks 42nd (fig. 12).

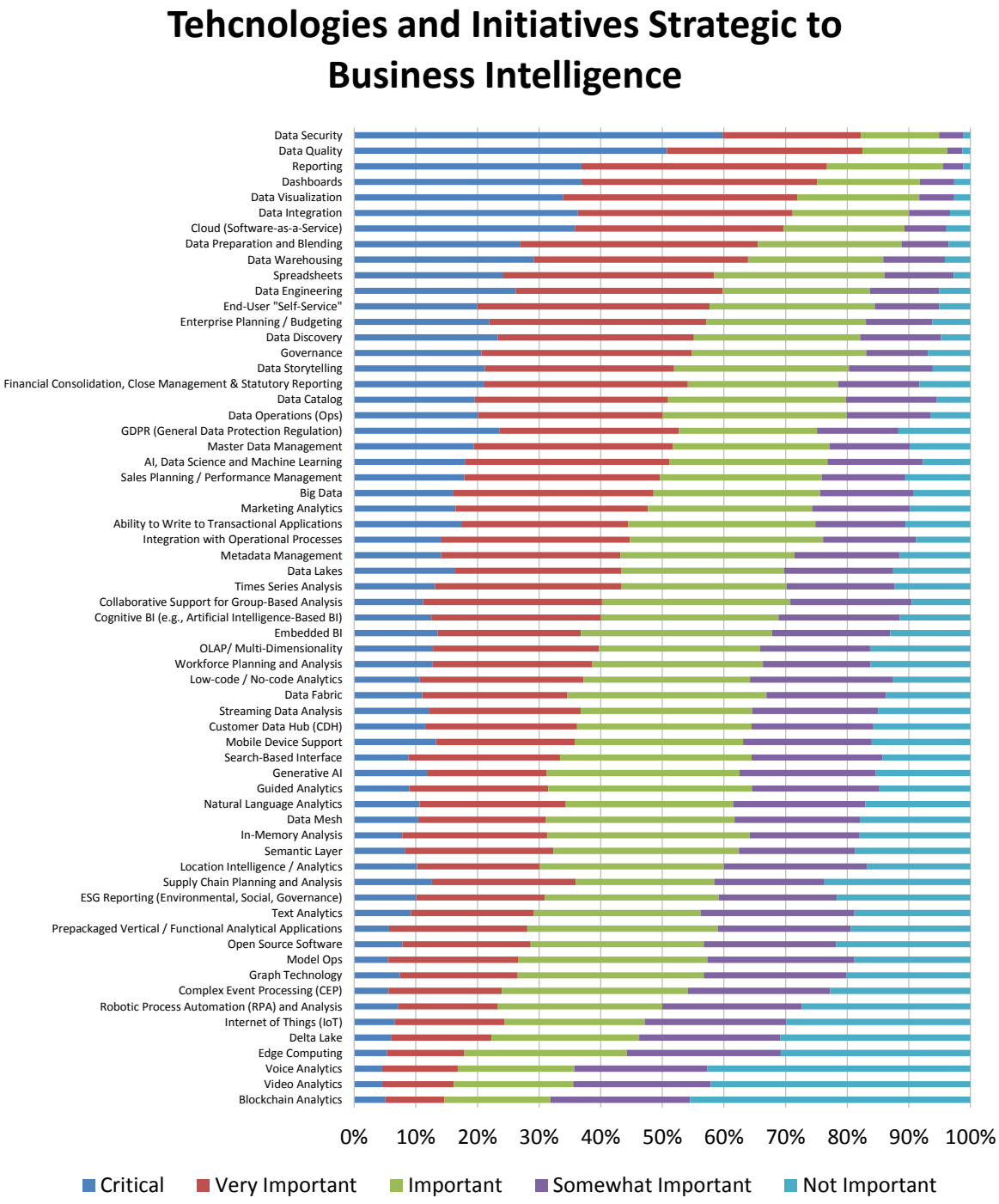


Figure 12 – Technologies and initiatives strategic to business intelligence

Importance of AI, Data Science, and Machine Learning 2014-2023

In our 2024 study, the weighted-mean perceived importance of AI, data science, and machine learning stands at 3.6, slightly higher than last year and slightly lower than the historic high of 3.7 in 2022 (fig. 13). A narrow range of 3.4-3.7 held from 2018 to the current year. This sustained measure of interest, not yet to the level of *very important,* likely holds room for future growth amid market interest and capex spending. This year, 82 percent consider the topic to be, at minimum, *important*, compared to 79 percent in 2023 and 83 percent in 2022. 2024's 3.6 weighted mean is above midway between *important* and *very important* on our scoring scale. Compared to the 3.4-3.7 scores of the last seven years, earlier 2014-2017 weighted-mean scores stood lower at 3.2-3.4, a bit above a level of *important*. (Also see industry importance, fig. 69.)

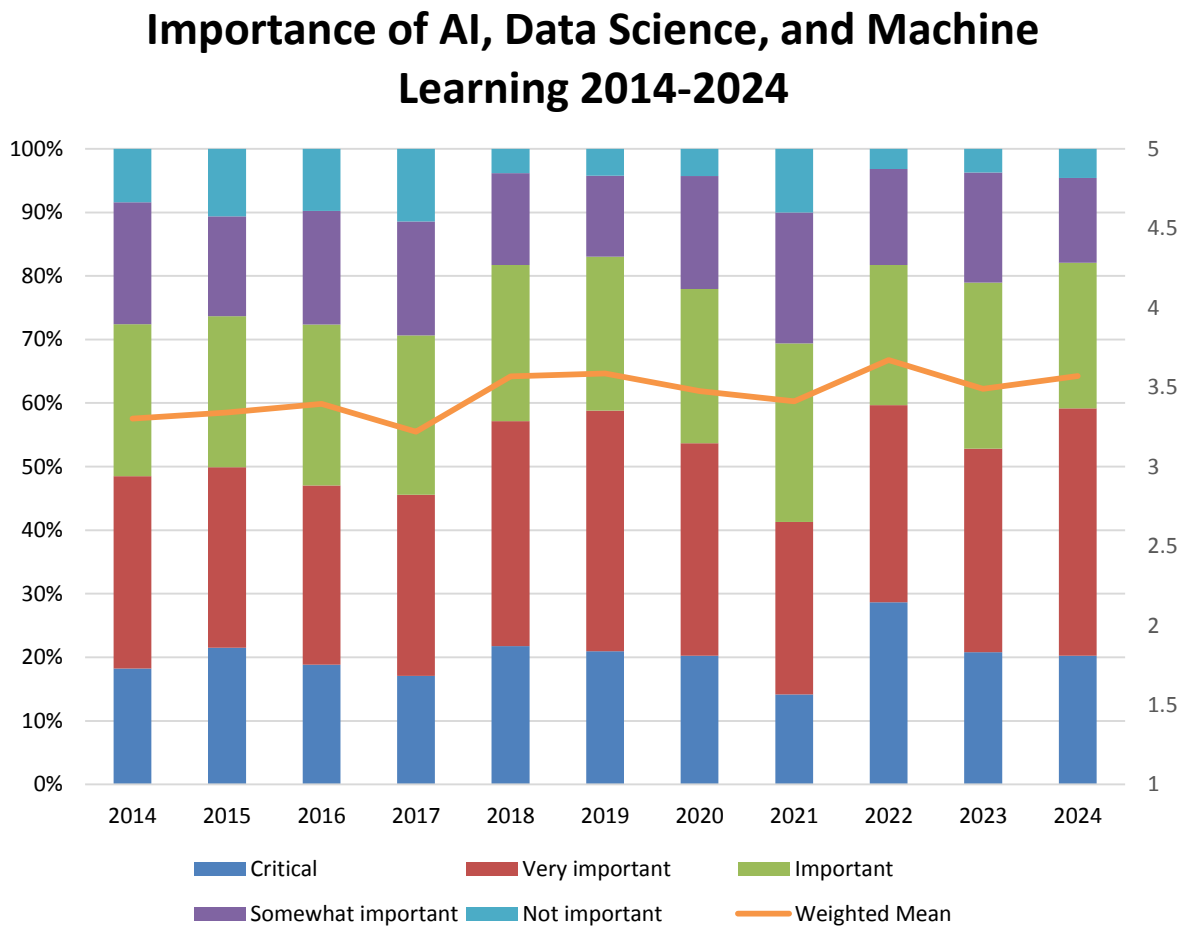


Figure 13 – Importance of AI, data science, and machine learning 2014-2024

Importance of AI, Data Science, and Machine Learning by Region

Interest in AI, data science, and machine learning varies by geography in 2024, with weighted-mean interest highest in Latin America (3.9), followed by North America (3.6), EMEA (3.6), and Asia Pacific (3.6) (fig. 14). Thus, all regions consider AI, data science, and machine learning to be higher than midway between a level signifying *important* and *very important*. Levels of *critical* importance are also lowest in Asia Pacific (17 percent), compared to North America (20 percent) and EMEA (23 percent).

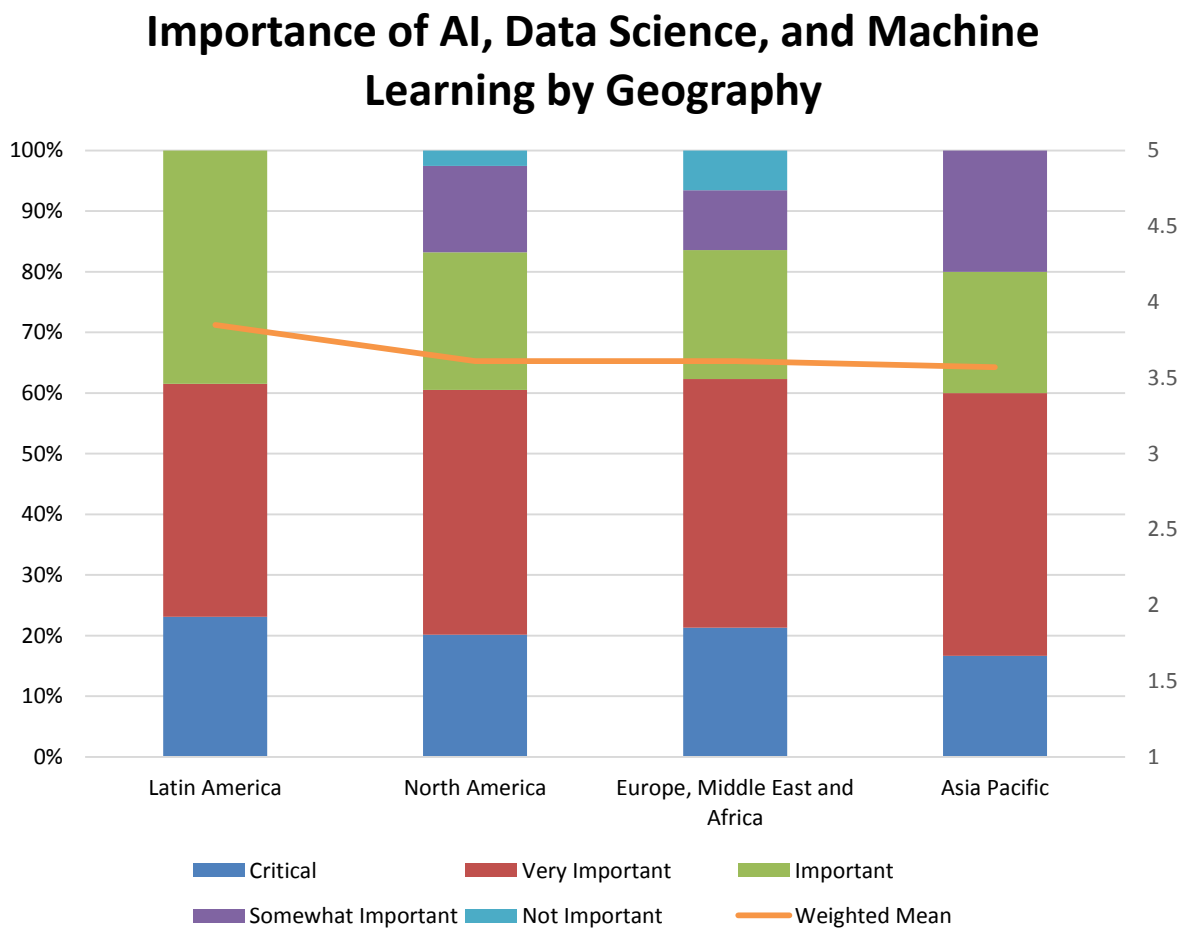


Figure 14 – Importance of AI, data science, and machine learning by geography

Importance of AI, Data Science, and Machine Learning by Function

Interest in AI, data science, and machine learning is above or well above the level of *important* across all functions in 2024 (fig. 15). This year, sentiment is highest in development and evaluative roles including *R&D* (weighted mean 4.2, greater than *very important)* and *executive management* (weighted mean 3.8, near *very important*). *BICC IT* and *sales and marketing* interest is next highest at 3.6, higher than midway between *important* and *very important.* All four leading interest by function are indicative of urgency, possibly incipient development and deployment of AI, DS, and ML. Least interested by role are respondents in *finance* (3.2) and *operations (*3.4), which, as mentioned, are still above the level of *important.*

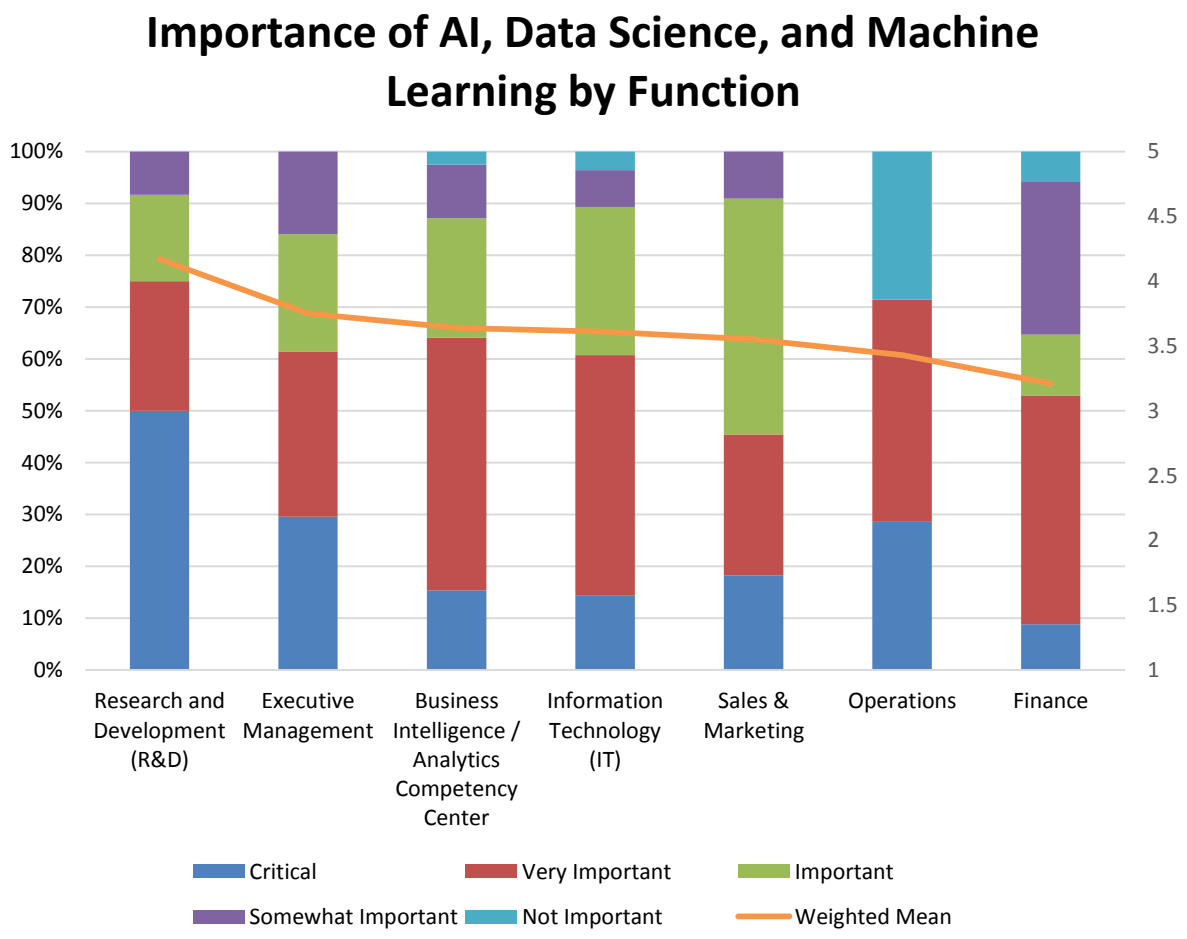


Figure 15 – Importance of AI, data science, and machine learning by function

Importance of AI, Data Science, and Machine Learning by Industry

Among vertical industries sampled in 2024, all except *retail and wholesale* give AI, data science, and machine learning somewhat or far higher than *important* scores ranging from 3.3-3.9 (fig. 16). This year, respondents in *technology* give the highest score, *above* 3.9, near *very important*, followed by *business services* (3.8)*, consumer services* (3.7)*, healthcare* (3.7), and *education* (3.6), *manufacturing (*3.5), *government* (3.4), *financial services* (3.3), and aforementioned *retail and wholesale* (2.9).

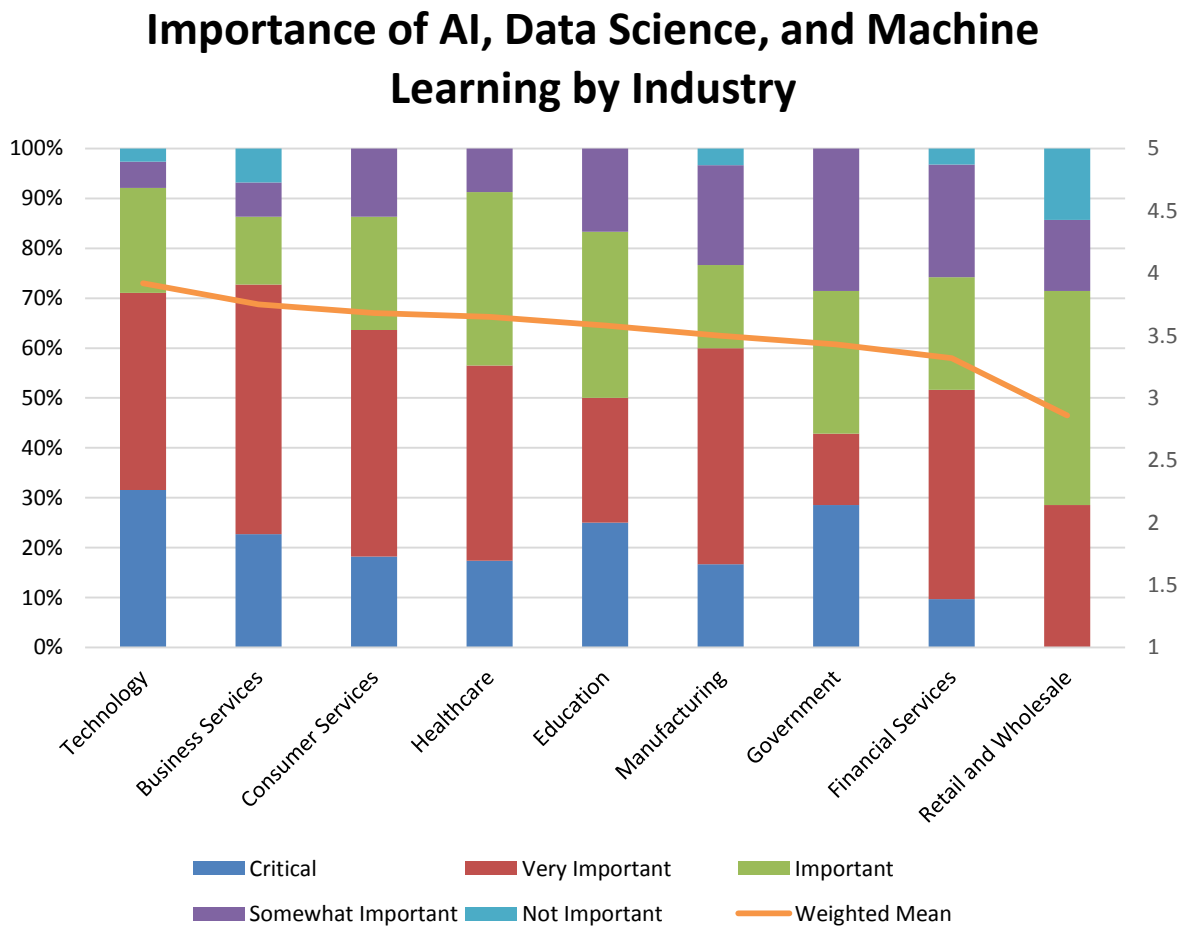


Figure 16 – Importance of AI, data science, and machine learning by industry

Importance of AI, Data Science, and Machine Learning by Organization Size

In 2024, the importance of AI, data science, and machine learning is broad across organizations of different size, and narrowly highest (weighted-mean 3.7) in small (1-100 employees) and very large (> 10,000 employees) organizations (fig. 17). This finding represents a widening of interest that was more confined to very large organizations in 2023. This year, the highest levels of *critical* interest are found in small organizations (29 percent), followed by very large (26 percent, midsize (20 percent), and large (11 percent) organizations. Nonetheless, weighted-mean interest is in a confined range of 3.5-3.7 across all organizations of any size.

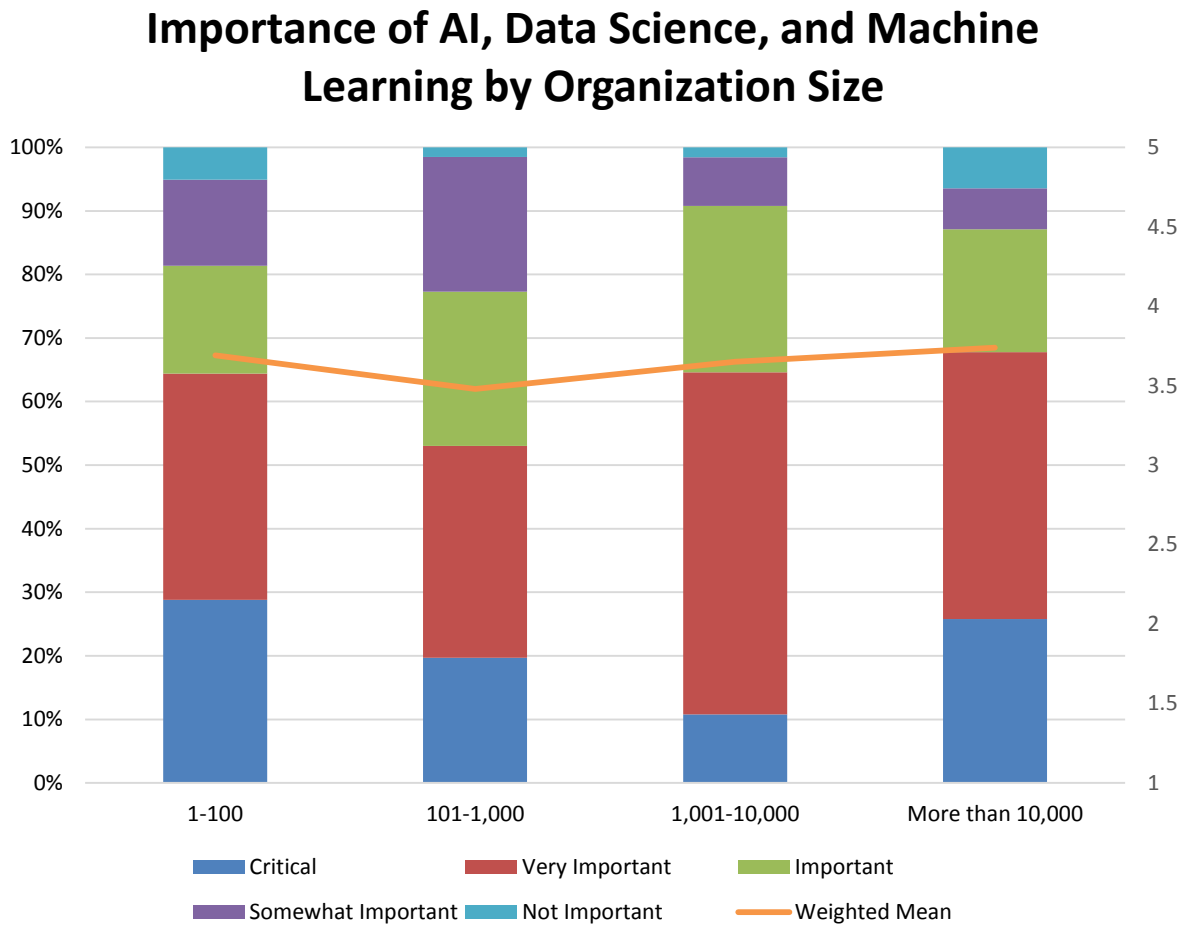


Figure 17 – Importance of AI, data science, and machine learning by organization size

Importance of AI, Data Science, and Machine Learning by Business Intelligence Success

The perceived importance of AI, data science, and machine learning correlates with organizational success with BI in 2024 (fig. 18). This is most dramatically visible when examining *critical* AI, DS, and ML importance in organizations with different levels of BI success. This year, *completely successful* BI organizations are 23 percent likely to say AI, data science, and machine learning are *critical*, compared to 21 percent in *somewhat successful* and just 13 percent of *somewhat unsuccessful* or *unsuccessful* organizations. We observe a similar if less pronounced correlation in combined *critical* and *very important* scores, and an inverse relationship among the few organizations that say AI, DS, and ML are *not important.*

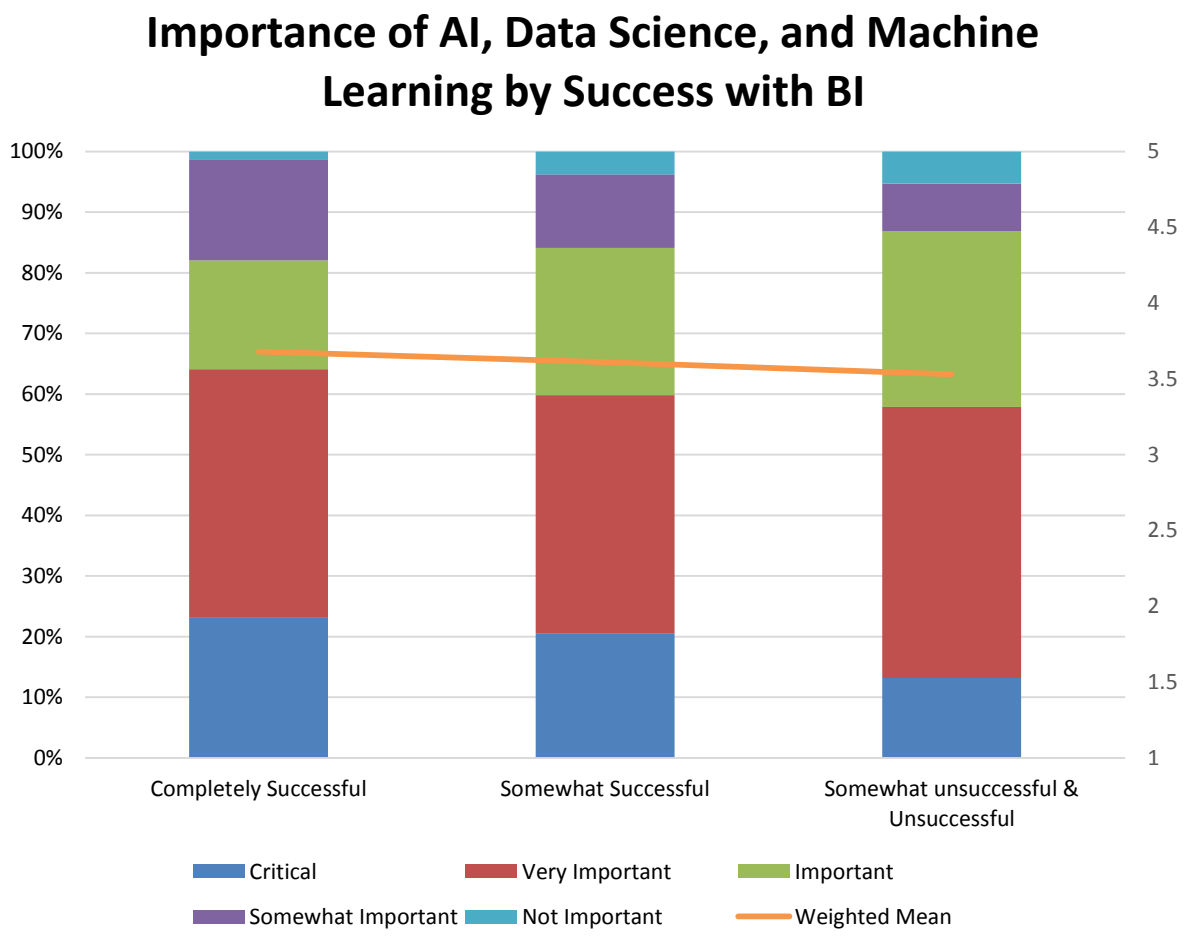


Figure 18 – Importance of AI, data science, and machine learning by success with BI

Use Cases for AI, Data Science, and Machine Learning

On the much-anticipated execution end of AI, data science, and machine learning, we asked respondents to describe the relevance of 13 different use cases for the technology (fig. 19). We can describe *current* usage as quite low across all use cases (6-19 percent) and future plans perspective, with less than 50 percent uptake for any use case expected in the next 12 months. The top six use cases by weighted mean expect to see close to 50 percent or greater uptake within 24 months. These include *customer segmentation* (58 percent), *demand forecasting* (60 percent), *customer lifetime value* (53 percent), *price optimization* (49 percent), *risk management* (48 percent), and *quality assurance* (50 percent). Even so, we describe current interest as "across the board," with more than half the top early-stage uses applied to roles and processes involving sales, marketing, planning, and forecasting. We frequently see large or the largest plurality of responses scores invariably applied to *don't know*, a reminder of the immaturity of AI, data science, and machine learning use.

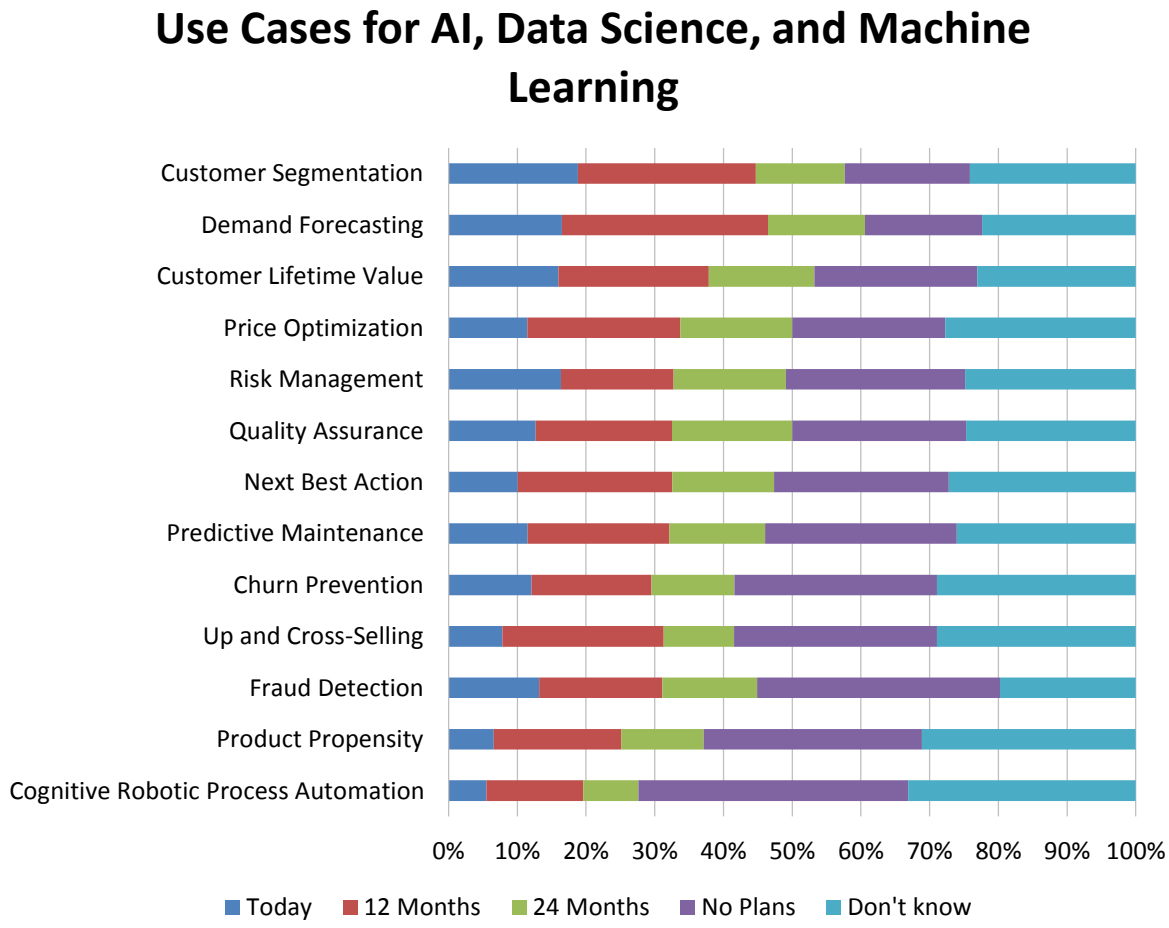


Figure 19 – Use cases for AI, data science, and machine learning

Use Cases for AI, Data Science, and Machine Learning by Geography

(Note: Figs. 19-25 all sample user response choices measured on a five-point scale by *currently use, 12 month, 24 month, no plans*, *and don't know.)*

We observe regional preferences regarding the current and future use of AI, data science, and machine learning in 2024 (fig. 20). In the great majority of cases, Asia Pacific leads in use case importance, though EMEA respondents notably post high or equally high use case scores for *customer segmentation and customer lifetime value*. Asia Pacific respondents lead all remaining use cases by importance, with outsized relative interest in areas including *demand forecasting, quality assurance, next-best action, up and cross selling*, and *fraud detection.* Most often, Latin America and North America report the lowest weighted-mean interest. While interesting, we consider this representation a very early snapshot of regional use.

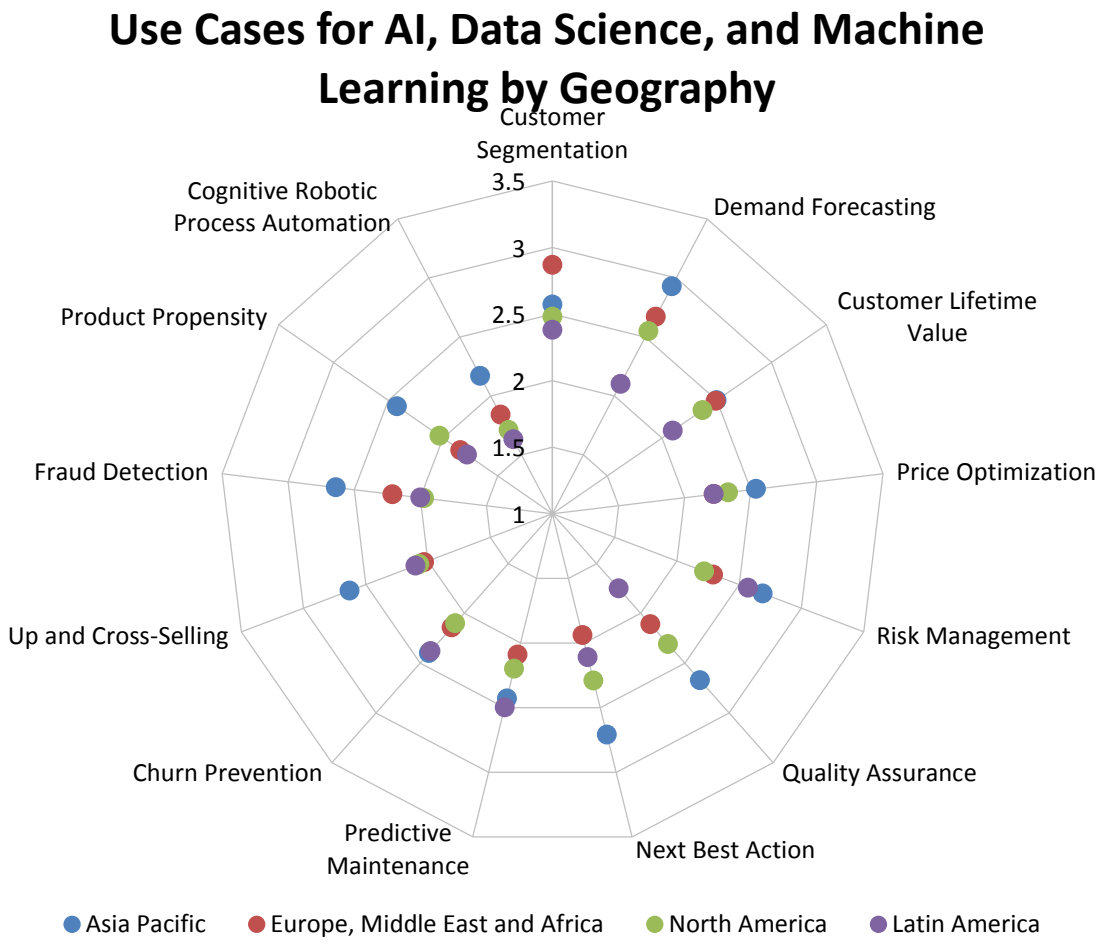


Figure 20 – Use cases for AI, data science, and machine learning by geography

Use Cases for AI, Data Science, and Machine Learning by Function

Current use and future uptake of use cases for AI, data science, and machine learning varies interestingly by function in 2024 (fig. 21). This year, *operations* leads interest in 10 of 13 use cases, with levels of at least *important* assigned to seven: *customer segmentation, customer lifetime value, price optimization, quality assurance, predictive maintenance, fraud detection*, and *product propensity. Finance* respondents lead interest in *demand forecasting,* report above-average interest in *customer segmentation,* but post lower scores for all other use cases. C*hurn prevention* resonates most with *R&D,* an indication of internal development. Relatively lower *executive management* uptake in most use cases may reflect a still-immature enterprise environment for AI, data science, and machine learning.

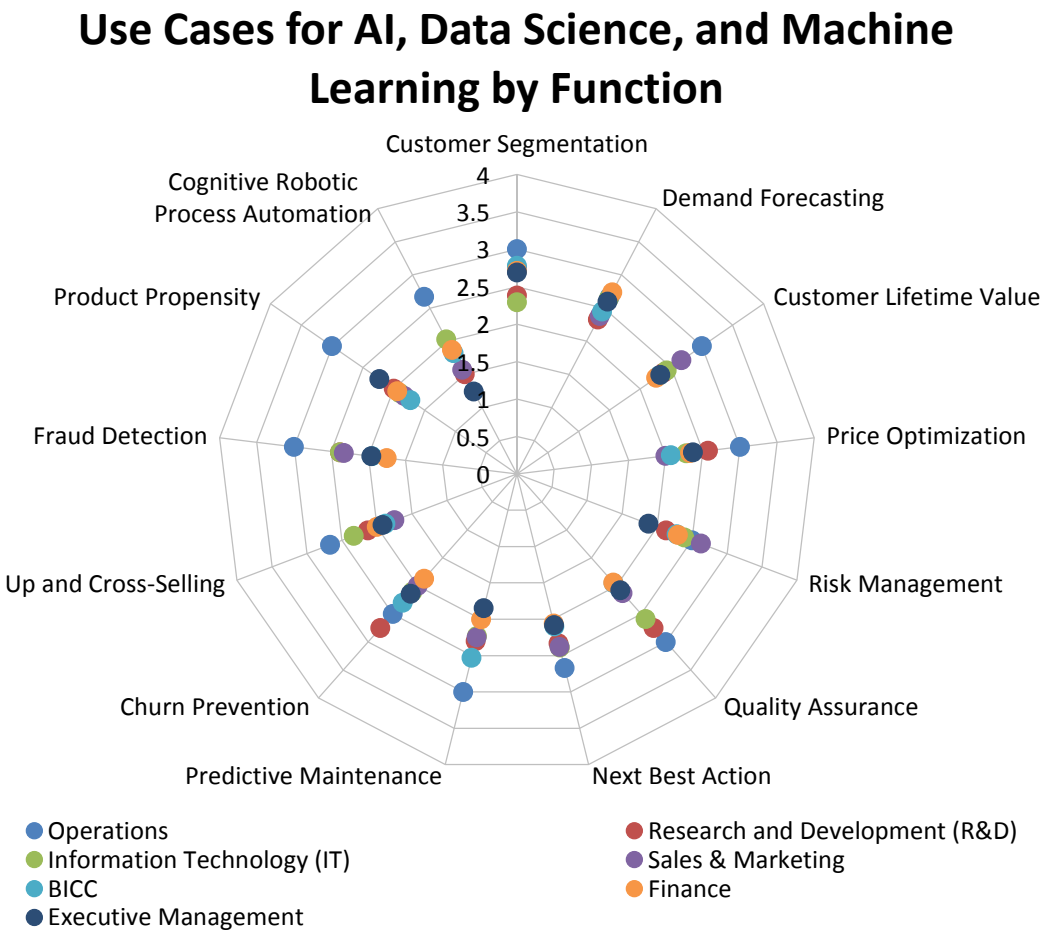


Figure 21 – Use cases for AI, data science, and machine learning by function

Use Cases for AI, Data Science, and Machine Learning by Industry

2024 current use and future uptake of use cases for AI, data science, and machine learning varies by industry, with early sentiment selective and notably strong among respondents in *financial services* (fig. 22). This year, *financial services* organizations report the highest or equally highest current/future uptake in 10 of 13 use case examples, often by wide margins. Use cases in financial services where criticality is near or above the midpoint of importance include *churn prevention, customer segmentation, fraud detection, price optimization*, *and risk management. Technology* organizations report the second-highest interest by weighted mean, with particular emphasis on *churn prevention, customer lifetime value*, *and up* and *cross-selling.* In 2024, *healthcare* and *retail and wholesale* report the lowest overall scores by weighted mean (though *retail* industry respondents show selective stronger interest in *customer lifetime value, customer segmentation*, and *demand forecasting*). The fact that only a handful of results exceed the midpoint of importance likely reflects uncertainty and limited experience with AI, DS. and ML use cases.

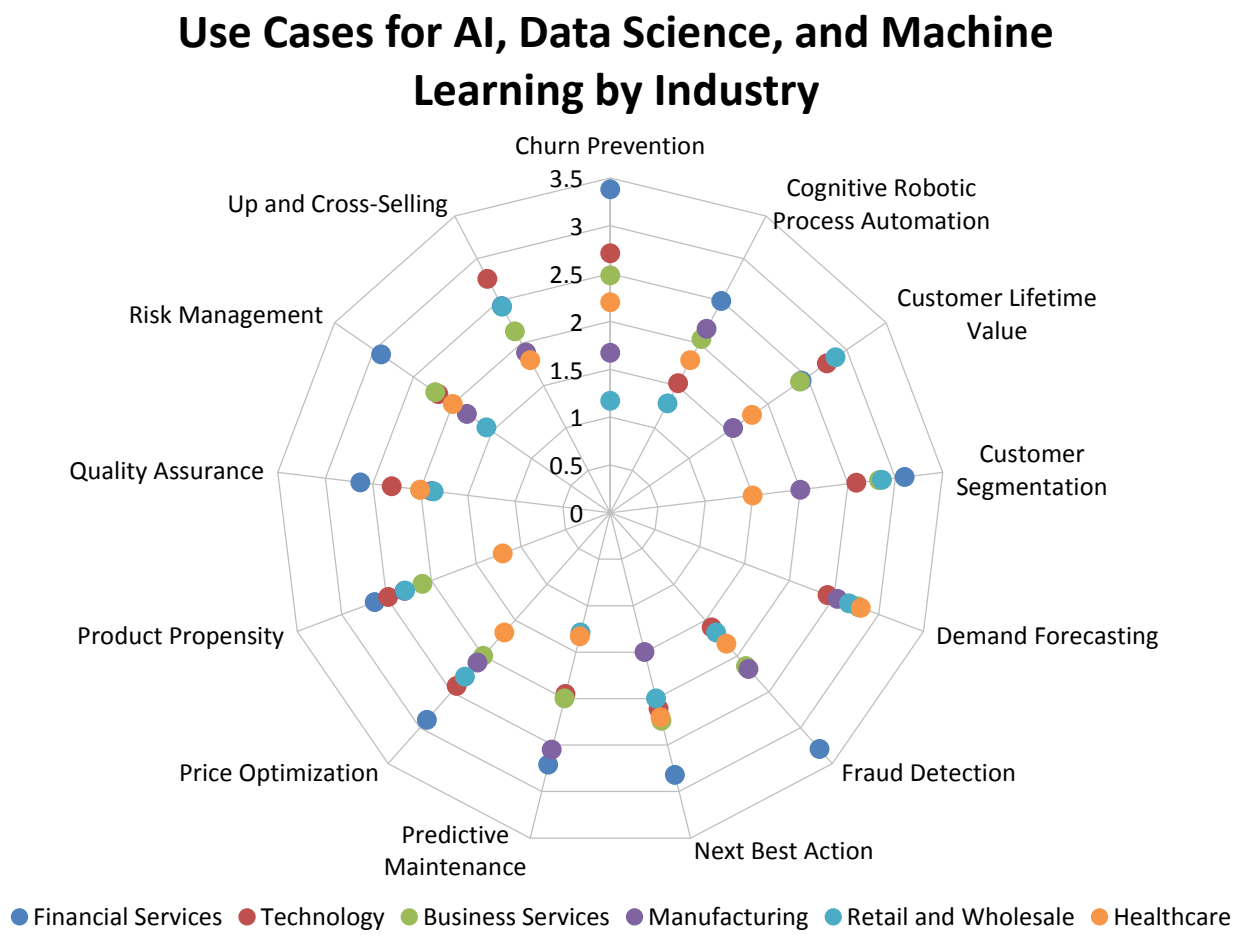


Figure 22 – Use cases for AI, data science, and machine learning by industry

Use Cases for AI, Data Science, and Machine Learning by Organization Size

Interest in use cases for AI, data science, and machine learning in 2024 is dominantly a very large organization (> 10,000 employees) phenomenon (fig. 23). This year, very large organizations lead uptake in all use cases, with standout areas of attention led by *demand forecasting* and *fraud detection*, followed by *customer segmentation, price optimization*, and *quality assurance*. Interest in all smaller organizations is visibly much lower, to the point of being surprising, with some pockets of interest but no use cases that rise near the midpoint of importance. We expect that, with more transparency and experience, all use cases will have situational relevance to organizations of any size.

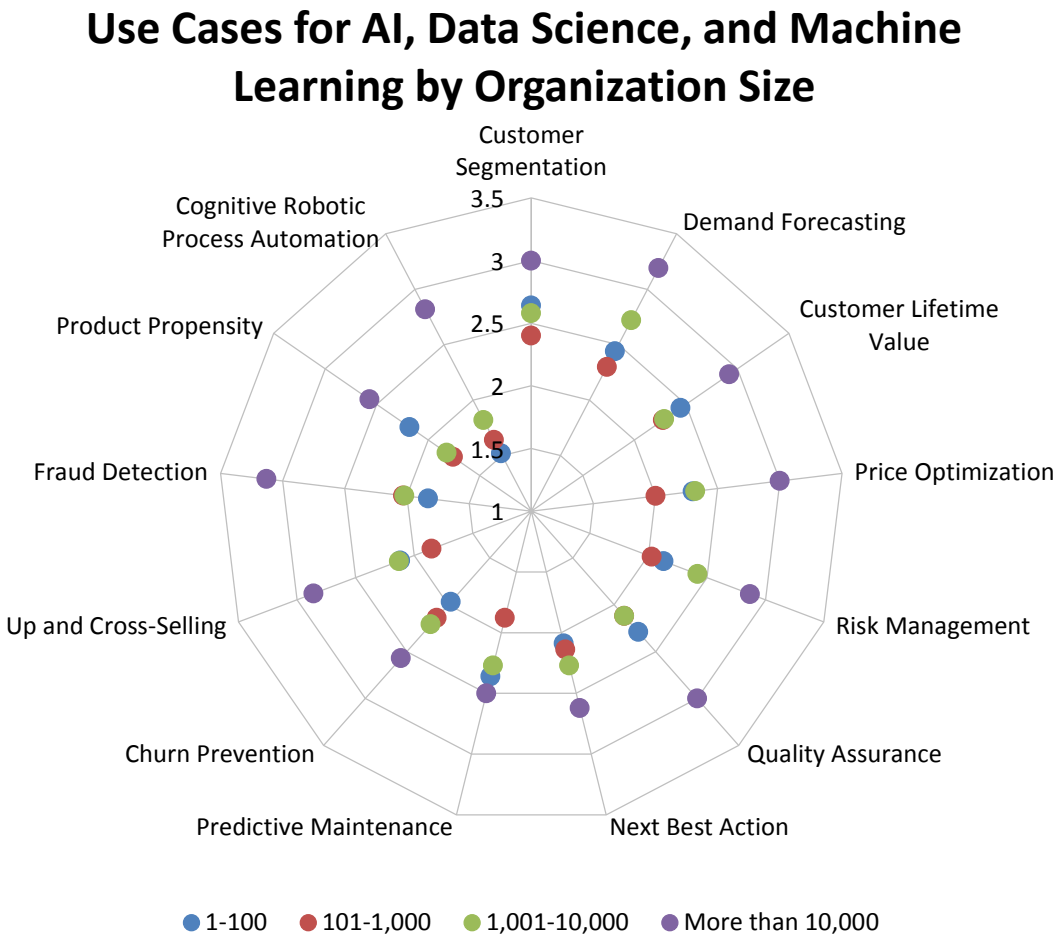


Figure 23 – Use cases for AI, data science, and machine learning by organization size

Use Cases for AI, Data Science, and Machine Learning by Company Age

Uptake of use cases for AI, data science, and machine learning varies unevenly but mostly skews toward middle-aged rather than the oldest or youngest organizations in 2024 (fig. 24). This year, for example, the oldest (16 years or more) organizations report the lowest or second-lowest current and future uptake, often well below the midpoint of importance. Likewise, the youngest organizations of less than five years, report relatively low urgency to most use cases. In contrast, organizations of 11-16 years lead or share highest or more imminent uptake in 10 of 13 use cases; and while organizations of 5-10 years are next most interested, only *customer* segmentation rises above the midpoint of importance in the latter group. Again, the fact that almost no use cases rise to the perceived midpoint of urgency is indicative of early days use and experience with AI, DS, and ML.

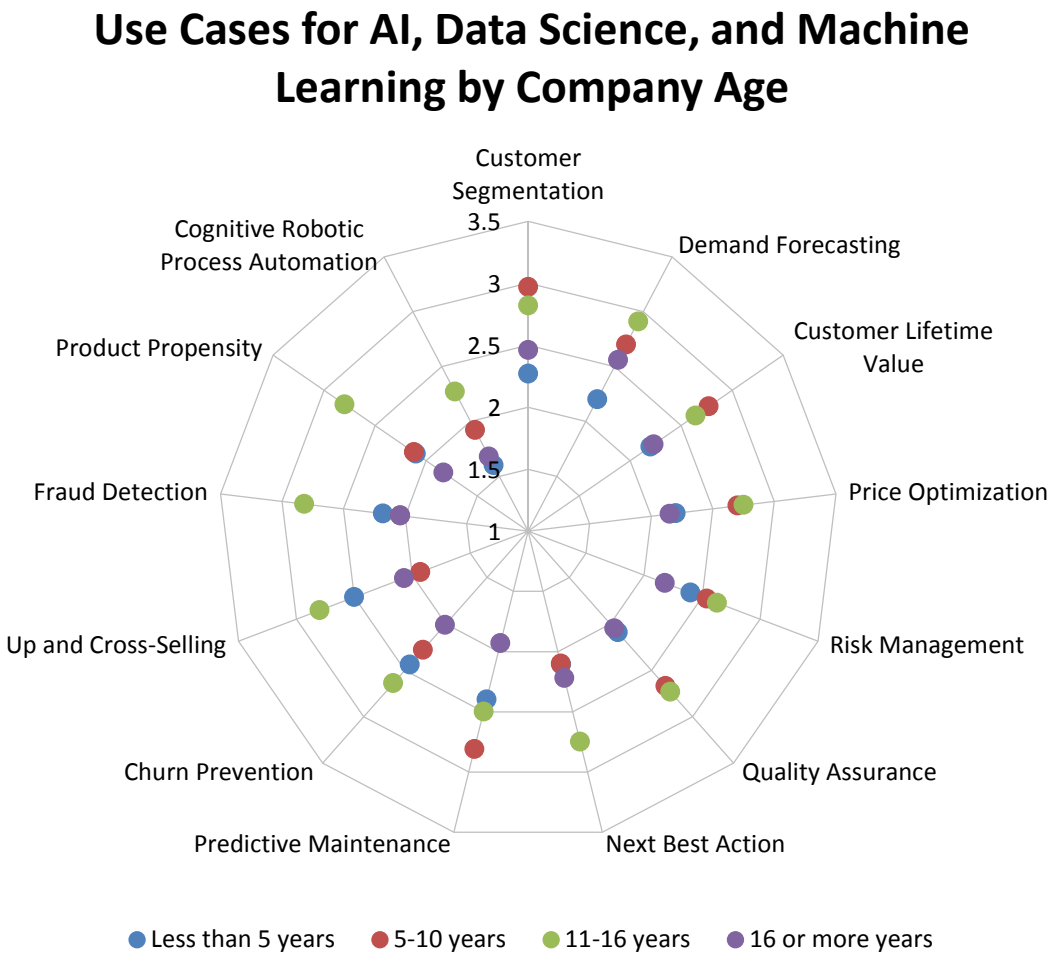


Figure 24 – Use cases for AI, data science, and machine learning by company age

Use Cases for AI, Data Science, and Machine Learning by Success with BI

Unlike some other measures, we observe a strong correlation between interest in use cases for AI, data science, and machine learning, and perceived success with business intelligence (fig. 25). Given limited experience and a likely shortage of measured outcomes associated with the technology, we might conclude this finding largely reflects higher attention to and anticipation of AI, DS, and ML in companies that consider themselves successful BI organizations. Even so, we see certain use cases most associated with perceived BI success that include *demand forecasting, customer segmentation*, and *risk management.* Again, we observe early-stage experience with AI, DS, and ML in the absence of use cases rising to the midpoint of urgency.

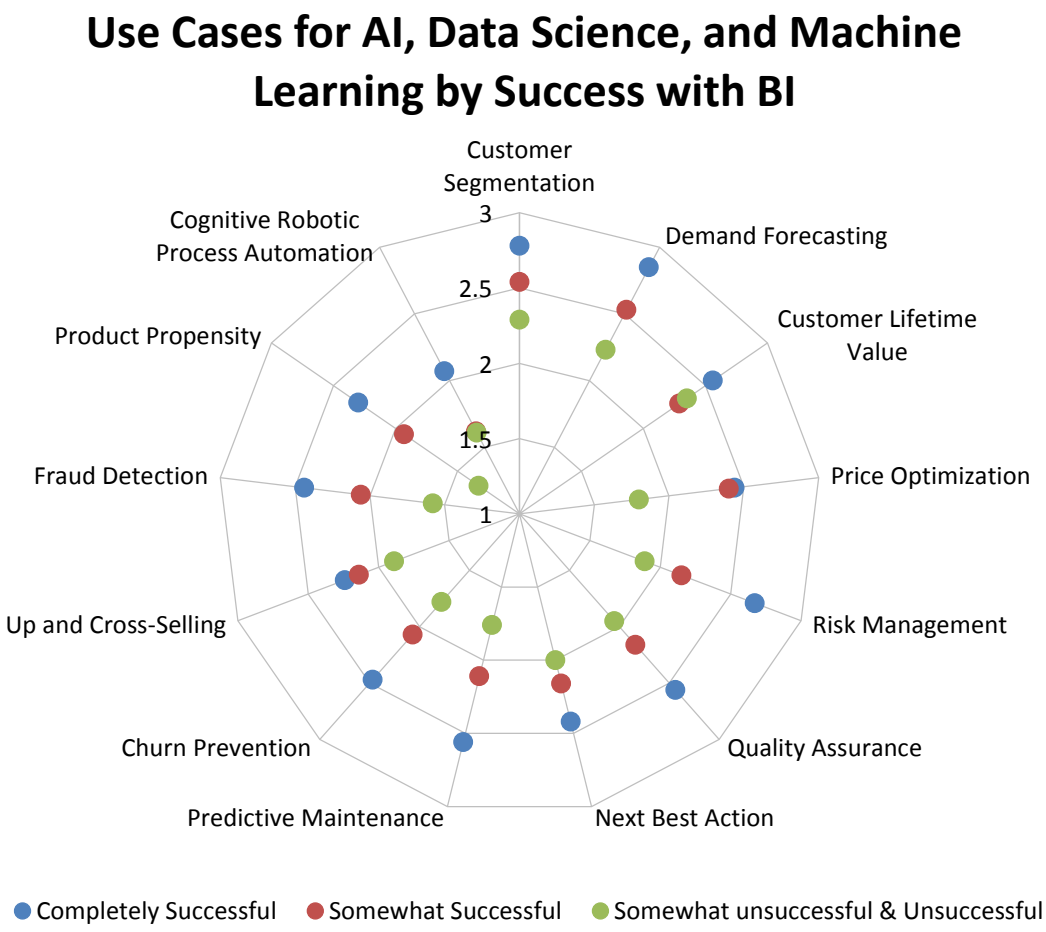


Figure 25 – Use cases for AI, data science, and machine learning by success with BI

Use Cases for AI, Data Science, and Machine Learning by Generative AI Adoption

Organizations that are more actively adopting or considering the use of generative AI prioritize some use cases over others (fig. 26). For example, those organizations with use cases *in production* are most likely to have prioritized *customer segmentation, customer lifetime value, demand forecasting*, *and quality assurance.* Among organizations that are *experimenting today, price optimization, churn prevention*, *and predictive maintenance* are the most common use cases. We nonetheless consider all use case scenarios, including *no plans*, to be subject to industry and other demographics, product/service maturity, and future needs.

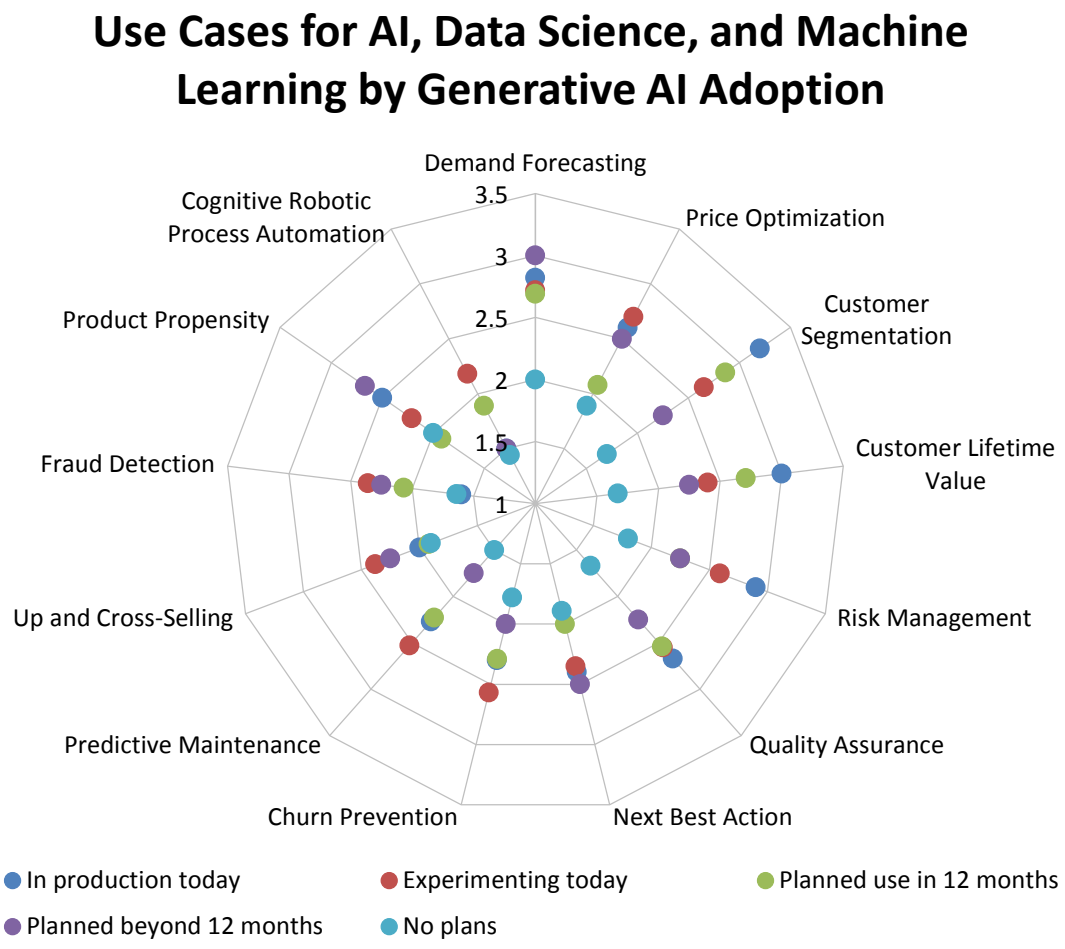


Figure 26 – Use cases for AI, data science, and machine learning by generative AI adoption

Deployment and Adoption Plans for AI, Data Science, and Machine Learning

Deployment of AI, Data Science, and Machine Learning 2016-2024

Across the last nine years of our study, we observed consideration of and actual deployment of AI, data science, and machine learning gain and then flatten or decrease slightly during the last three years (fig. 27). (Beginning in 2022, we divided "*yes, we use today"* into "*in production"* and "*in very limited ways,"* and combined weighted mean.) From a 2017 weighted-mean low of 3.0, to a 2021 high of 3.6, sentiment fell slightly to the 3.3 level seen in 2023 and again in 2024. Thus, despite recent flattening, sentiment remains above the level of *important* for the last eight years. Those with *no plans* grew to 17 percent in 2024 from an all-time low of 9 percent in 2022, possibly indicating some hesitancy or mixed success amid market events. Even so, we expect the ongoing momentum in the AI, data science, and machine learning space to bode well for wider adoption as more use cases and solutions become apparent (also see perceived importance, fig. 13).

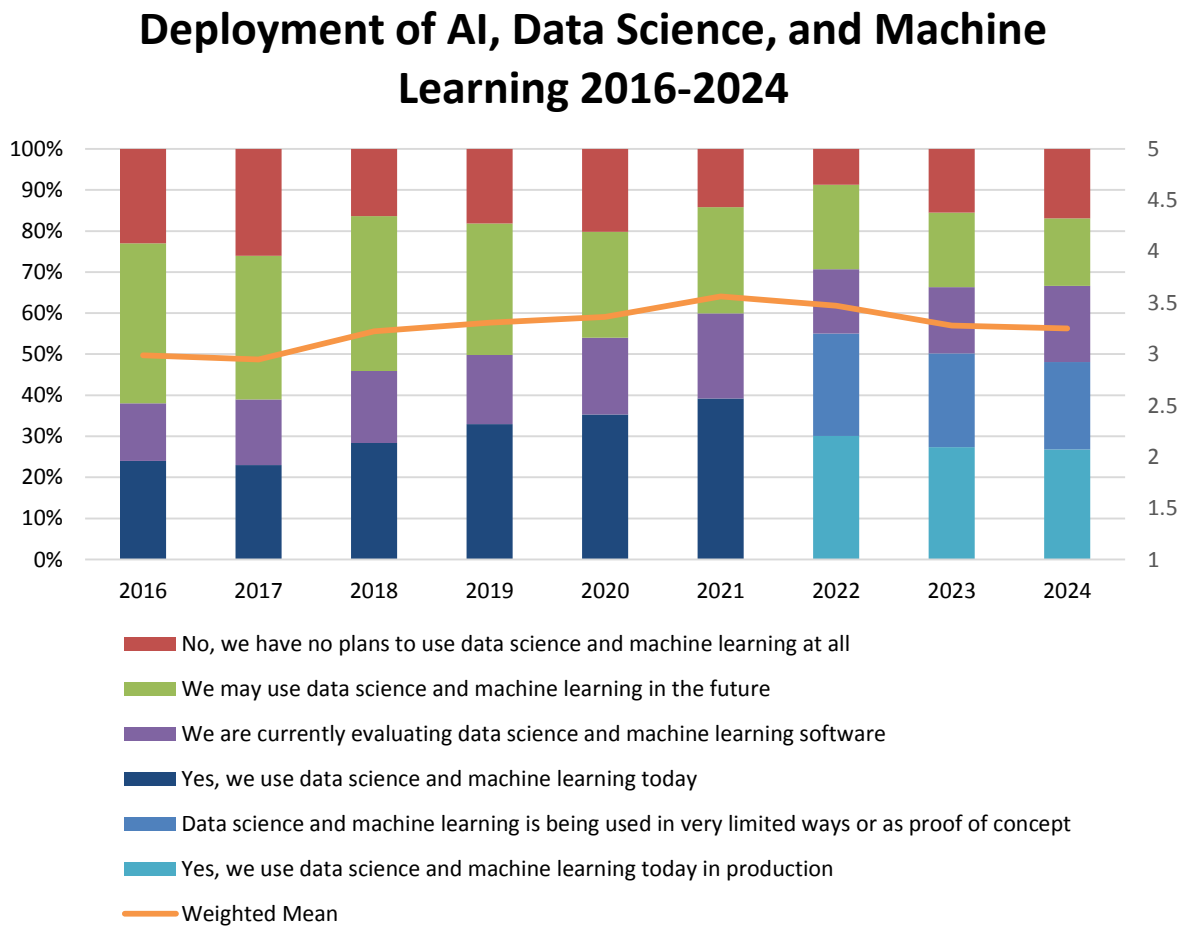


Figure 27 – Deployment of AI, data science, and machine learning 2016-2024

Deployment of AI, Data Science, and Machine Learning by Geography

Deployment rates of AI, data science, and machine learning vary by geography in 2024 with both current and weighted-mean use visibly highest in EMEA, followed by North America, Asia Pacific, and Latin America (fig. 28). The rate of *current in-production* use is similarly highest in Asia Pacific and EMEA (33-34 percent), followed by North America (23 percent) and Latin America (8 percent). All regions are committed to future use, with combined current users and evaluators most numerous in EMEA (70 percent) and North America (67 percent).

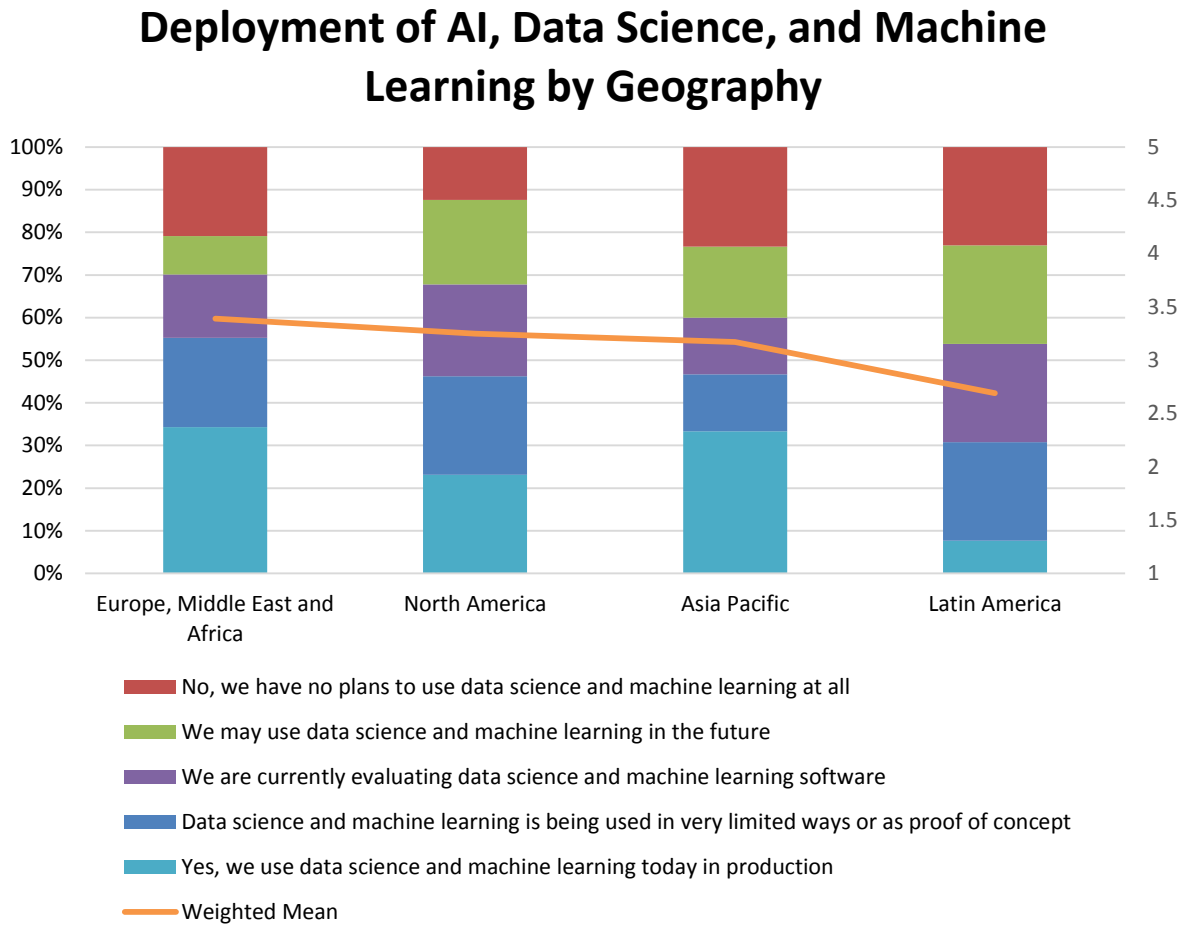


Figure 28 – Deployment of AI, data science, and machine learning by geography

Deployment of AI, Data Science, and Machine Learning by Function

Current deployment of AI, data science, and machine learning varies broadly according to function in 2024 (fig. 29). Current use in production is by far highest among respondents in *R&D* (about 50 percent), an indication of further development and testing. *Current* plus *very limited* use is also highest in *R&D* (75 percent), followed by *operations* (72 percent), and quickly trails off among other functions. After *R&D*, full production use is next highest (28-29 percent) among respondents in *operations, IT*, and the *business intelligence / analytics competency center*. Among the most reluctant, *executive management* (20 percent), *sales and marketing* (18 percent) and *finance* (16 percent), are least likely to report AI, DS. and ML in production use today, though these groups also predict likely future use.

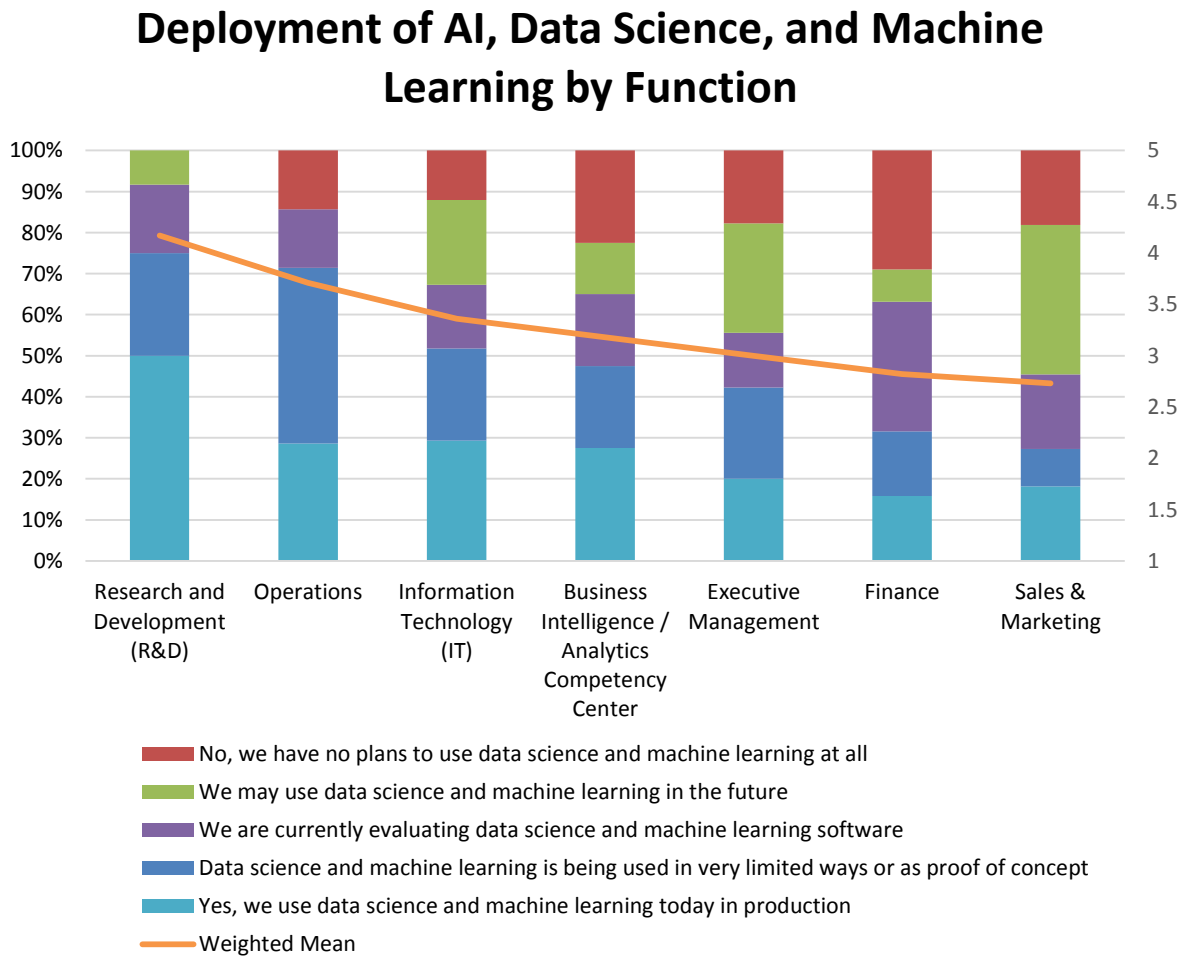


Figure 29 – Deployment of AI, data science, and machine learning by function

Deployment of AI, Data Science, and Machine Learning by Industry

Viewed by industry, respondents in *consumer services* and *healthcare* appear to have a first-mover edge in full *in-production* or *limited* use deployment of AI, data science, and machine learning (fig. 30). Interestingly, current in-production use is highest in *healthcare* (45 percent), though it is difficult to discern how this use might be distributed across research, clinical/operational, back office, and other areas. In 2024, *consumer services* reports 32 percent in-production use, while most other industries, including *business services, financial services, technology*, and *retail and wholesale* all report in-production use in the range of 26-29 percent. *Government, manufacturing*, and *education* have been slowest to embrace AI, DS, and ML, though at least 70 percent of respondents in any industry report they may use data science and machine learning in the future.

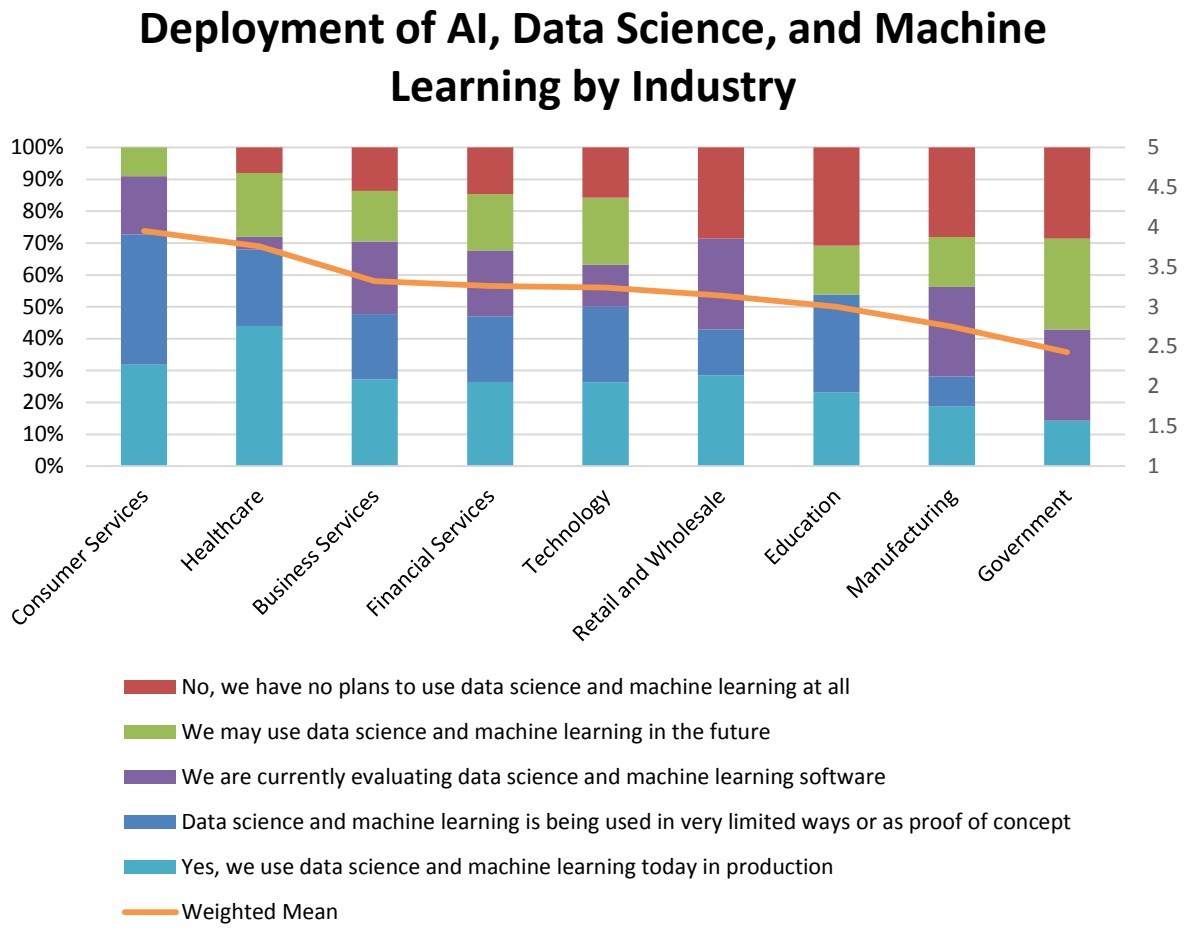


Figure 30 – Deployment of AI, data science, and machine learning by industry

Deployment of AI, Data Science, and Machine Learning by Organization Size

The number of deployments of AI, data science, and machine learning clearly increases with global headcount in 2024 and is by far highest in very large organizations (> 10,000 employees) (fig. 31). This year, 17 percent of small organizations (1-100 employees) report *in-production* deployments, compared to 20 percent at midsize (101-1,000 employees), 27 percent at large (1,001-10,000 employees), and 58 percent at very large organizations (> 10,000 employees). Combined *in production, limited use* and *evaluation* responses boost very large organization participation to 77 percent. Eighty percent of small and midsize organizations and 87-88 percent of large and very large organizations say at minimum that they may use AI, data science, and machine learning in the future.

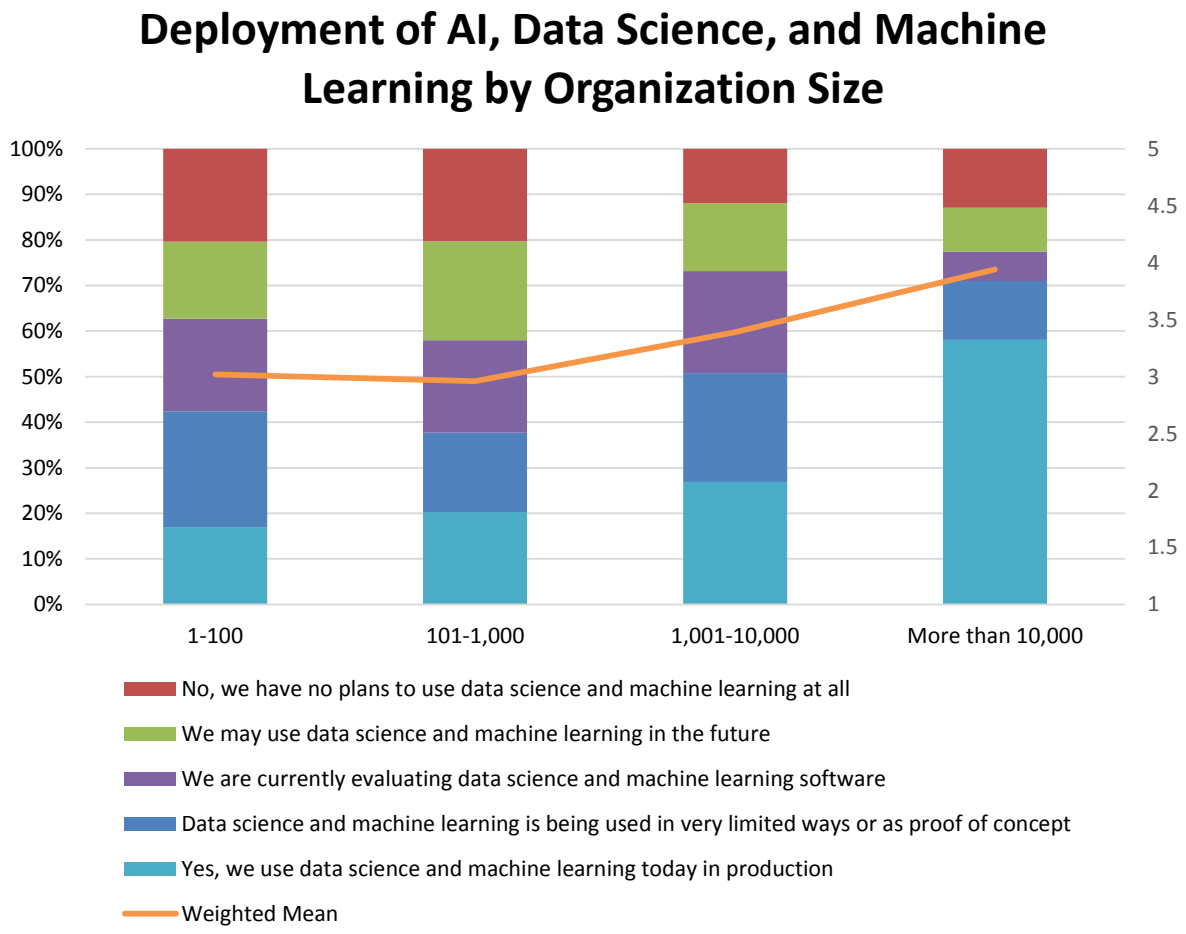


Figure 31 – Deployment of AI, data science, and machine learning by organization size

Deployment of AI, Data Science, and Machine Learning by Success with BI

Deployment of AI, data science, and machine learning correlates to success with BI in 2024 (fig. 32). Organizations that say they are *completely successful* with business intelligence are 31 percent likely to report *in-production* use of AI, DS, and ML, compared to 26 percent of *somewhat successful* and 18 percent of *somewhat unsuccessful and unsuccessful* organizations. Conversely, organizations that are *somewhat unsuccessful and unsuccessful* with BI are far more likely to say they have *no plans* (21 percent) or *may use in the future* (26 percent).

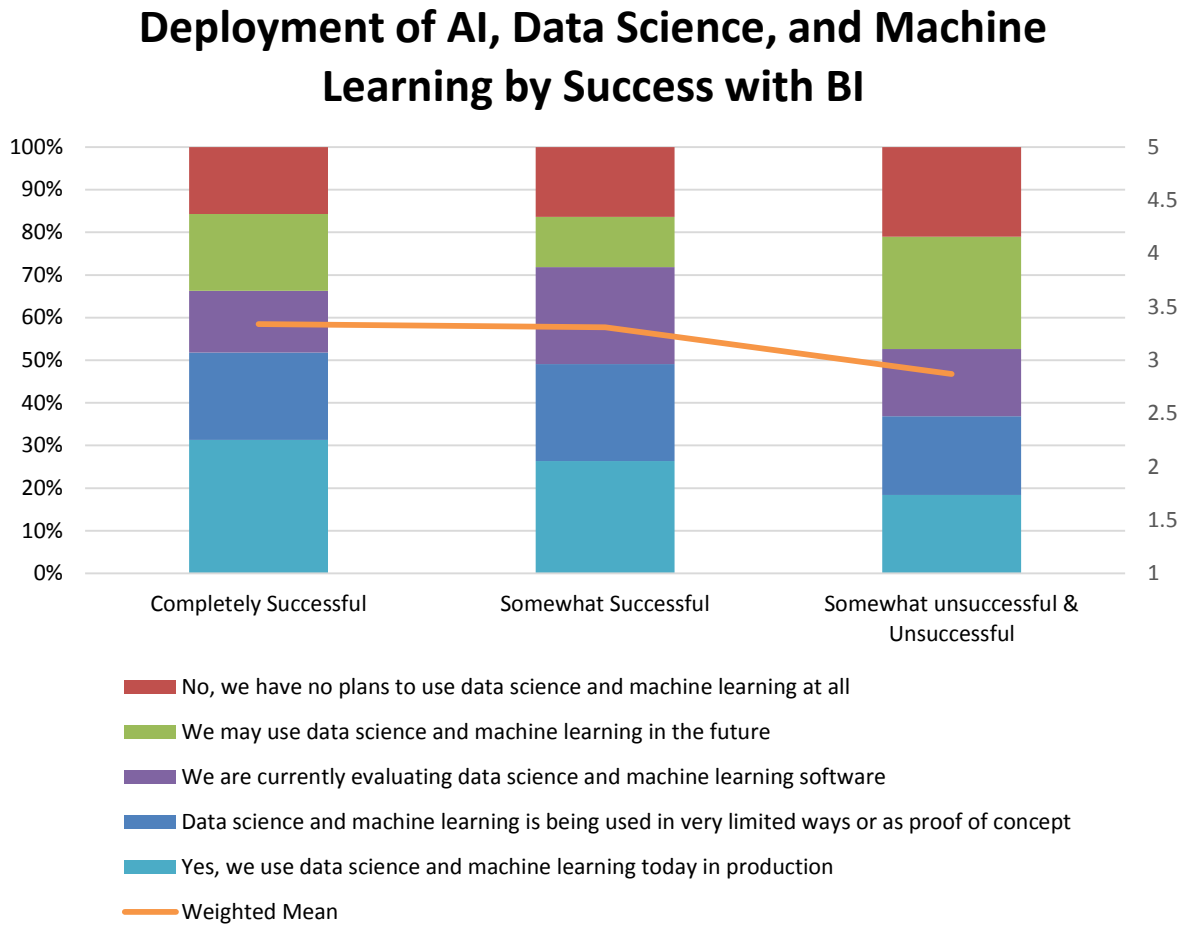


Figure 32 – Deployment of AI, data science, and machine learning by success with BI

Deployment of AI, Data Science, and Machine Learning by Data Literacy

The perceived level of data literacy clearly correlates with the likelihood of current or future deployment of AI, data science, and machine learning in 2024 (fig. 33). Organizations that report *high* data literacy are 37 percent likely to be current in-production users, compared to 27 percent of *moderate*, and 19 percent of *low and very low* literacy. The correlation is true in reverse among respondents with *no plans* or those that *may use* data science and machine learning in the future. Though high data literacy is often a perceived rather than quantified measure, we find that industries and roles that claim greater data management immersion, experience, and focus historically report higher data literacy.

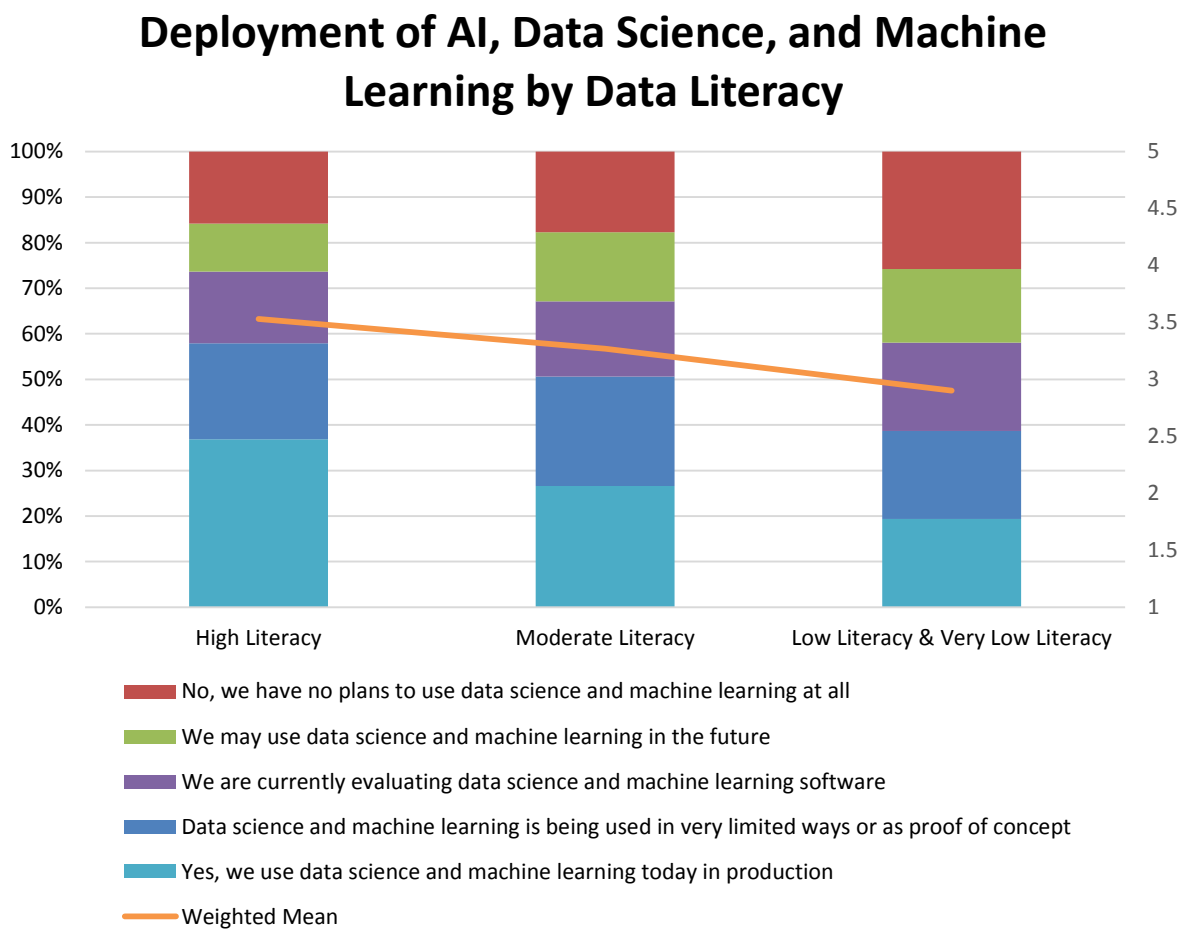


Figure 33 – Deployment of AI, data science, and machine learning by data literacy

Longevity of AI, Data Science, and Machine Learning

Longevity of AI, Data Science, and Machine Learning 2018-2024

We see indicators of AI, data science, and machine learning adoption reflected in the longevity of current programs (fig. 34). At a glance, the last seven years in sum appears to show an increasing number of longstanding (> five years) organizations using AI, data science, and machine learning, and fewer one to two years and less than one-year users. Since longevity reached a high in 2022, we have nonetheless seen a two-year rebound in "younger" longevity of less than one, one to two, two to three, and three to five year-year users. Thus, we see proportionately more "newer" users of AI, data science, and machine learning since 2022 that may represent a second wave of adoption and future maturity.

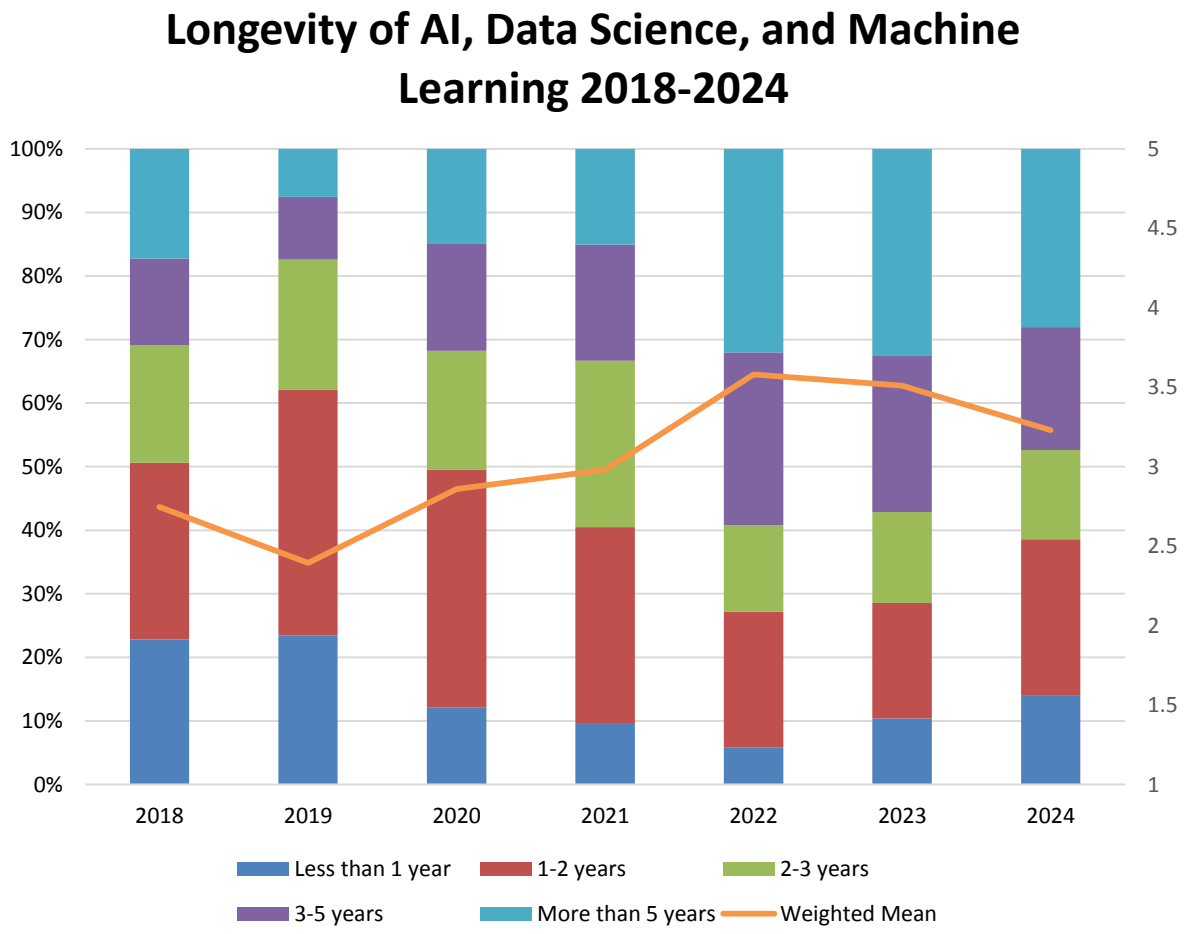


Figure 34 – Longevity of AI, data science, and machine learning 2018-2024

Longevity of AI, Data Science, and Machine Learning by Geography

Our 2024 study gives a mixed message of the longevity of AI, data science, and machine learning by region (fig. 35). The longest maturity of *more than five years* is somewhat comparable, ranging from 25 percent in EMEA, to 27 percent in North America, and 30 percent in Asia Pacific. While the percentages of older users is fairly consistent, we see the greatest *less than one year* adoption in North America and EMEA, and more mid-age and higher longevity in Asia Pacific.

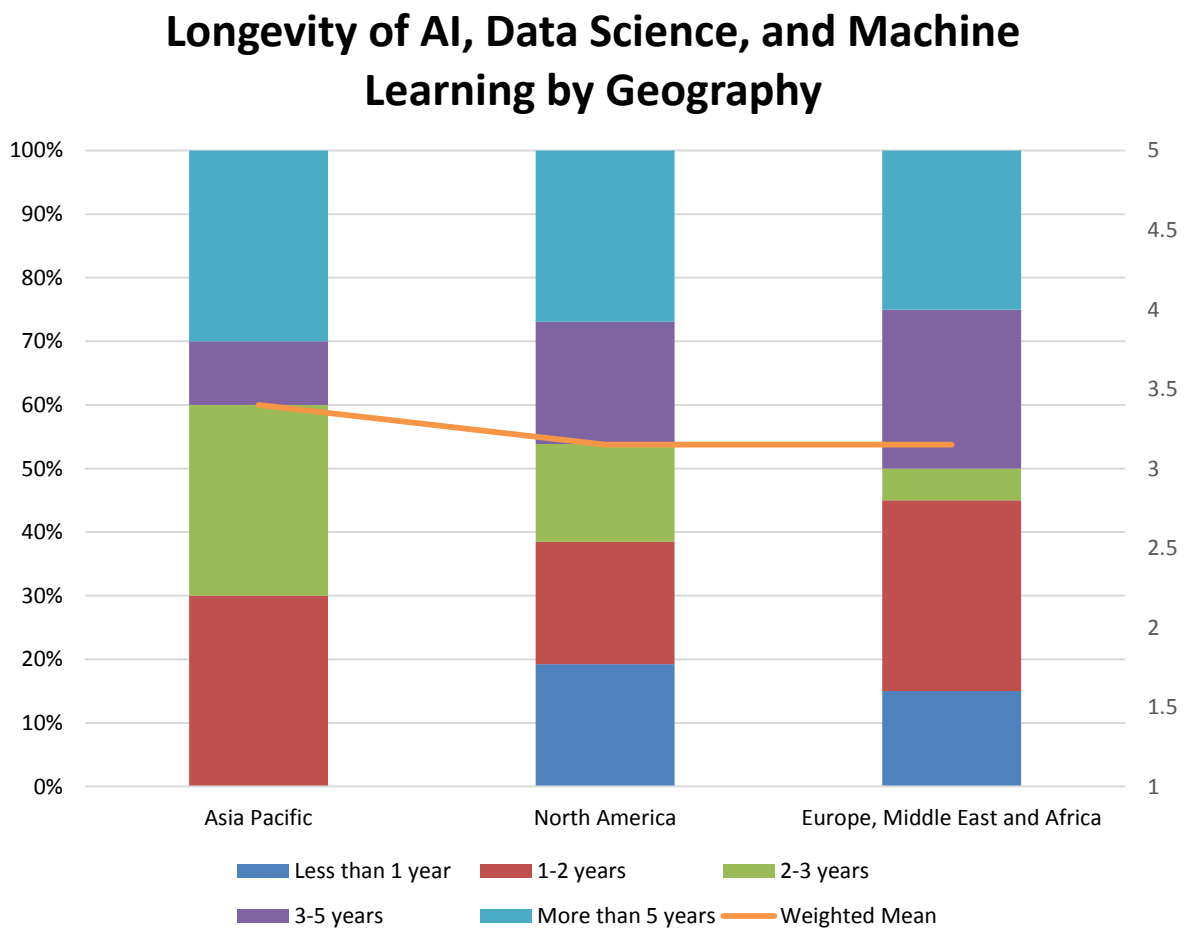


Figure 35 – Longevity of data science and machine learning by geography

Longevity of AI, Data Science, and Machine Learning by Function

The longevity of AI, data science, and machine learning efforts varies dramatically by function in 2024, often by a considerable margin (fig. 36). Much as we saw in the case of deployment by function (fig. 29), longevity is greatest in *R&D* and is also high among *IT* respondents, an indication of ongoing technical/analytical development, testing, and deployment. Use history is newer and stronger among respondents in *executive management* and even the BICC, an indication of incipient interest, recent focus, and resource investment, and a possibly steep learning curve. Fully 60 percent of *BICC* and 66 percent of *executive management* user experience is two years or less.

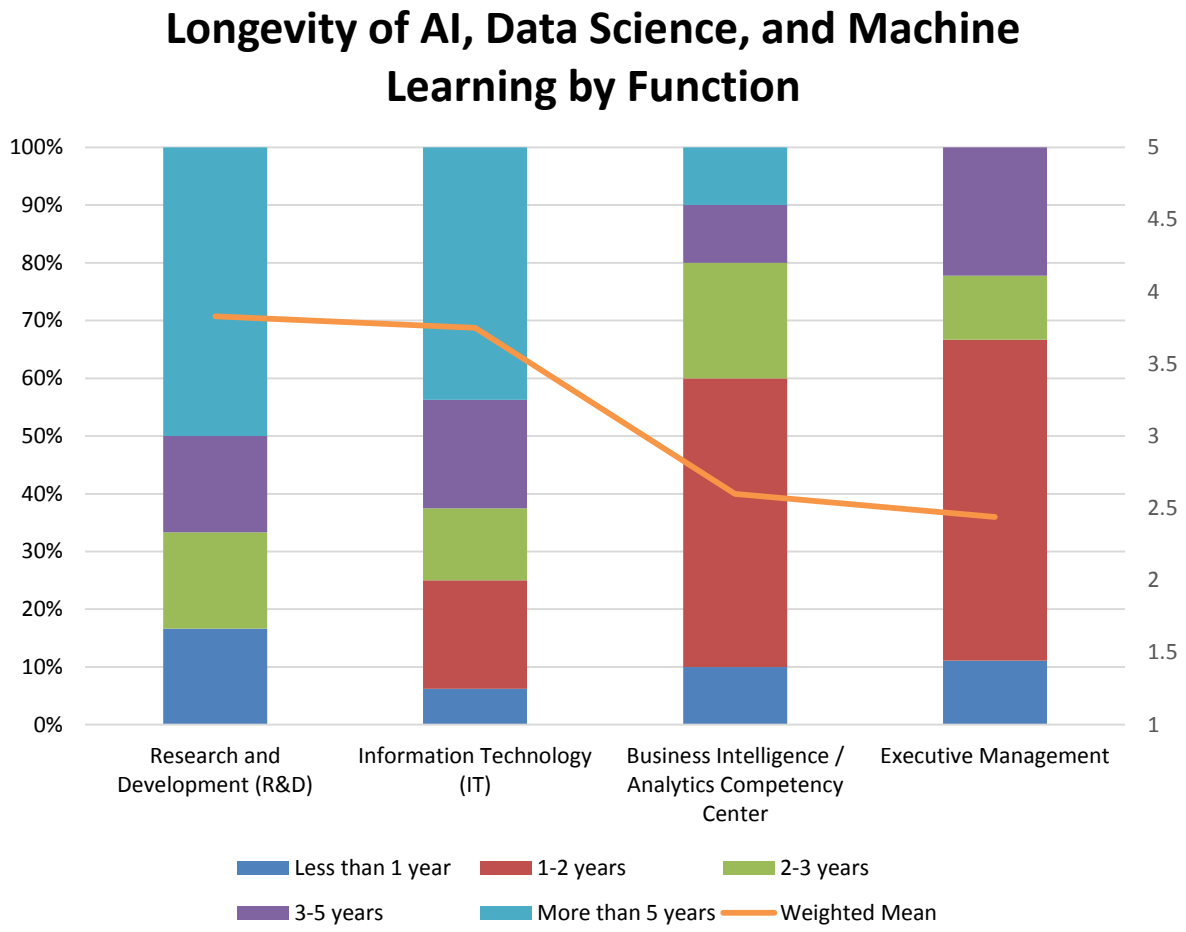


Figure 36 – Longevity of AI, data science, and machine learning by function

Longevity of AI, Data Science, and Machine Learning by Industry

Viewed by industry, *manufacturing, financial services*, and *technology* respondents give the highest weighted-mean score (3.8, 3.6 and 3.4 respectively) for longevity of AI, data science, and machine learning in 2024 (fig. 37). *Financial services* respondents also report the most mature use of more than five years (38 percent). Newer investment of less than one year, or one to two years is highest in *healthcare* (60 percent), *business services* (54 percent), and *consumer services* (50 percent).

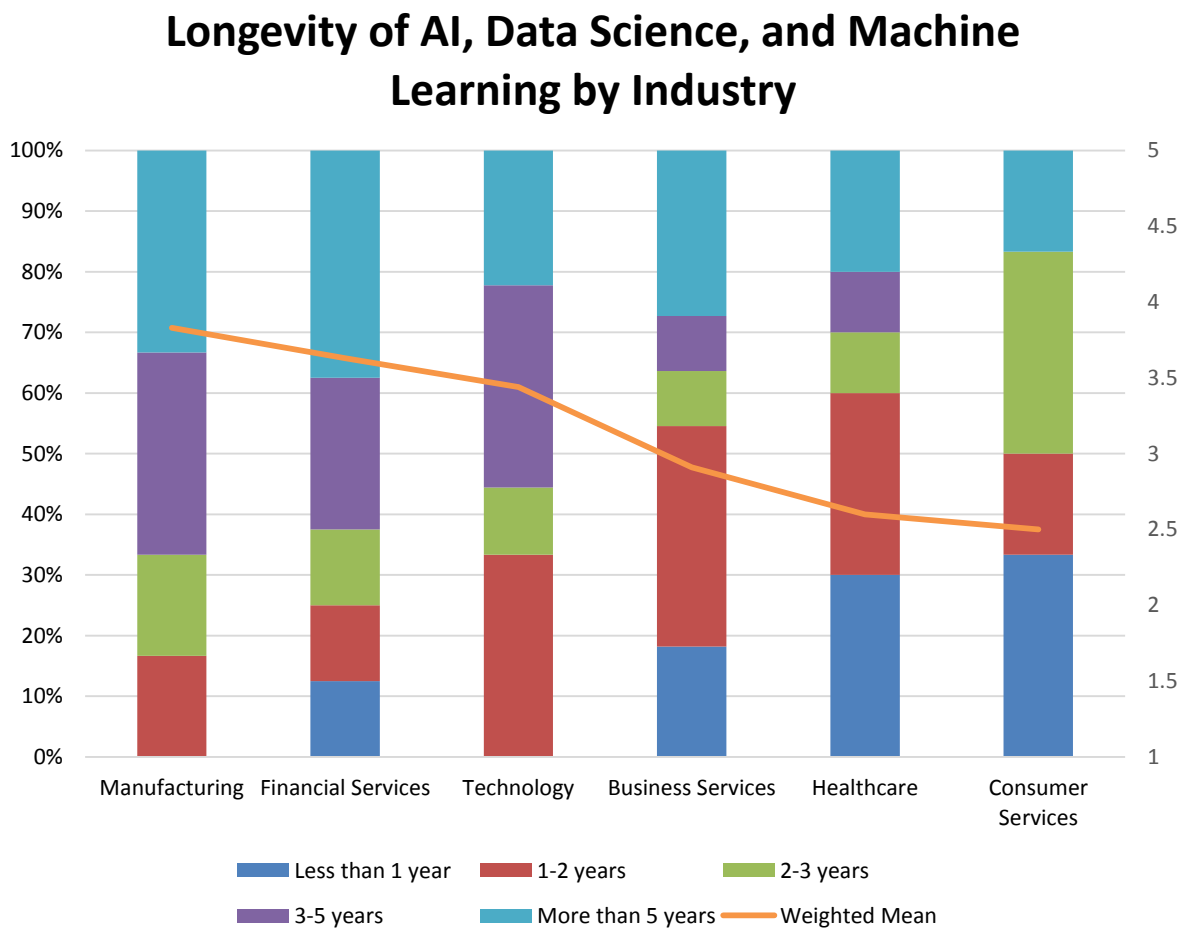


Figure 37 – Longevity of AI, data science, and machine learning by industry

Longevity of AI, Data Science, and Machine Learning by Organization Size

The longevity of the use of AI, data science, and machine learning clearly increases with organization size in 2024 (fig. 38). This finding is unsurprising in the wake of very-large organization dominance of deployment and other measures. Fully half of very-large organization involvement is > *5-year* experience, which is more than twice the incidence at large (1,001-10,000 employees) organizations and up to 4.5 the rate at in smaller midsize (101,-1,000 employees), and small (1-100 employees) peer organizations. We also observe the steepest *less than one-year* uptake in large organizations and by far the steepest uptake of shorter and up to three-year longevity in small organizations.

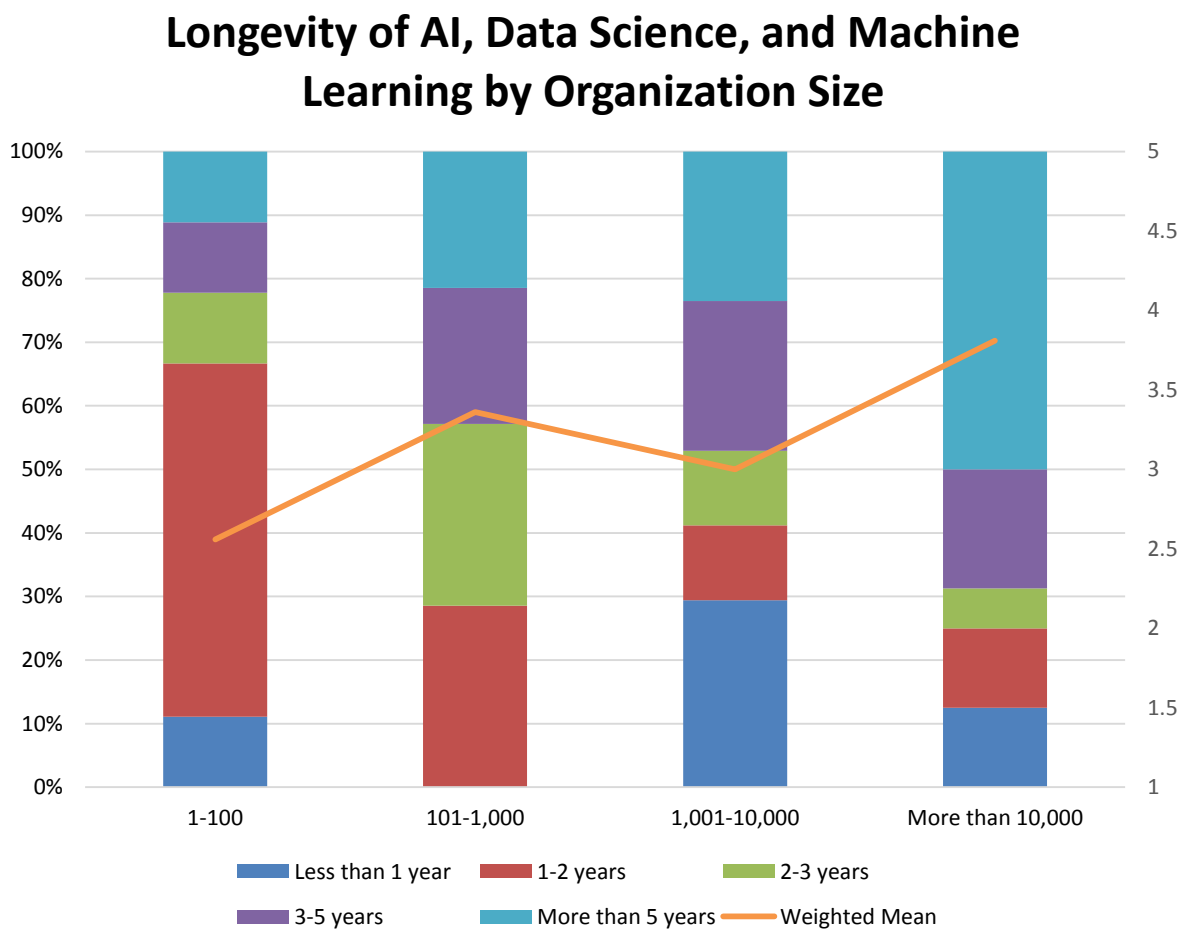


Figure 38 – Longevity of AI, data science, and machine learning by organization size

Features for AI, Data Science and Machine Learning

Respondents express significant interest in the full range of feature requirements for AI, data science, and machine learning in 2024, where all but two of 23 sampled features are at least *important* to two-thirds (67 percent) or far more (up to 83 percent of) respondents (fig. 39). The most important among these support traditional statistical methods: *range of regression models*, *outlier detection*, *model explain-ability, model management and governance,* and *optimization (e.g., linear programming)*, all of which are at least important to more than 80 percent of all respondents. (Also see industry support for features, fig. 70-71.)

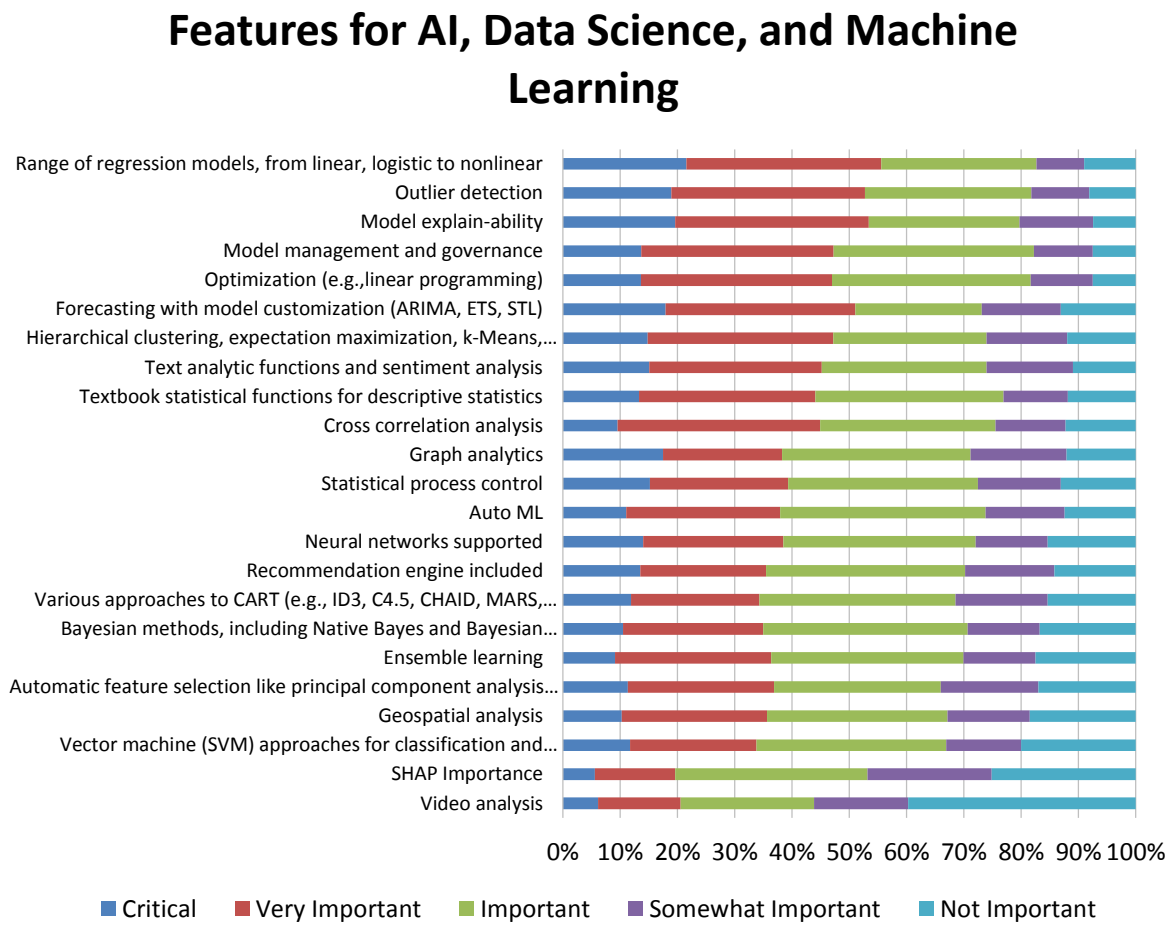


Figure 39 – Features for AI, data science, and machine learning

Features for AI, Data Science, and Machine learning 2014-2024

Fig. 40 shows respondent interest in feature requirements for AI, data science, and machine learning across 11 years of our data collection. Amid ebbs and flows, we observe 2024 interest often above historic average levels, with two all-time highs recorded for *model management and governance* (weighted mean 3.4), and *text analytics* (3.1). Top pick *range of regression models* receives the top 2024 score (3.5). This view of feature importance over time shows strong growth since our inaugural 2014 study and sustained relevance of a majority of all features near or above the value signifying *important* in 2024.

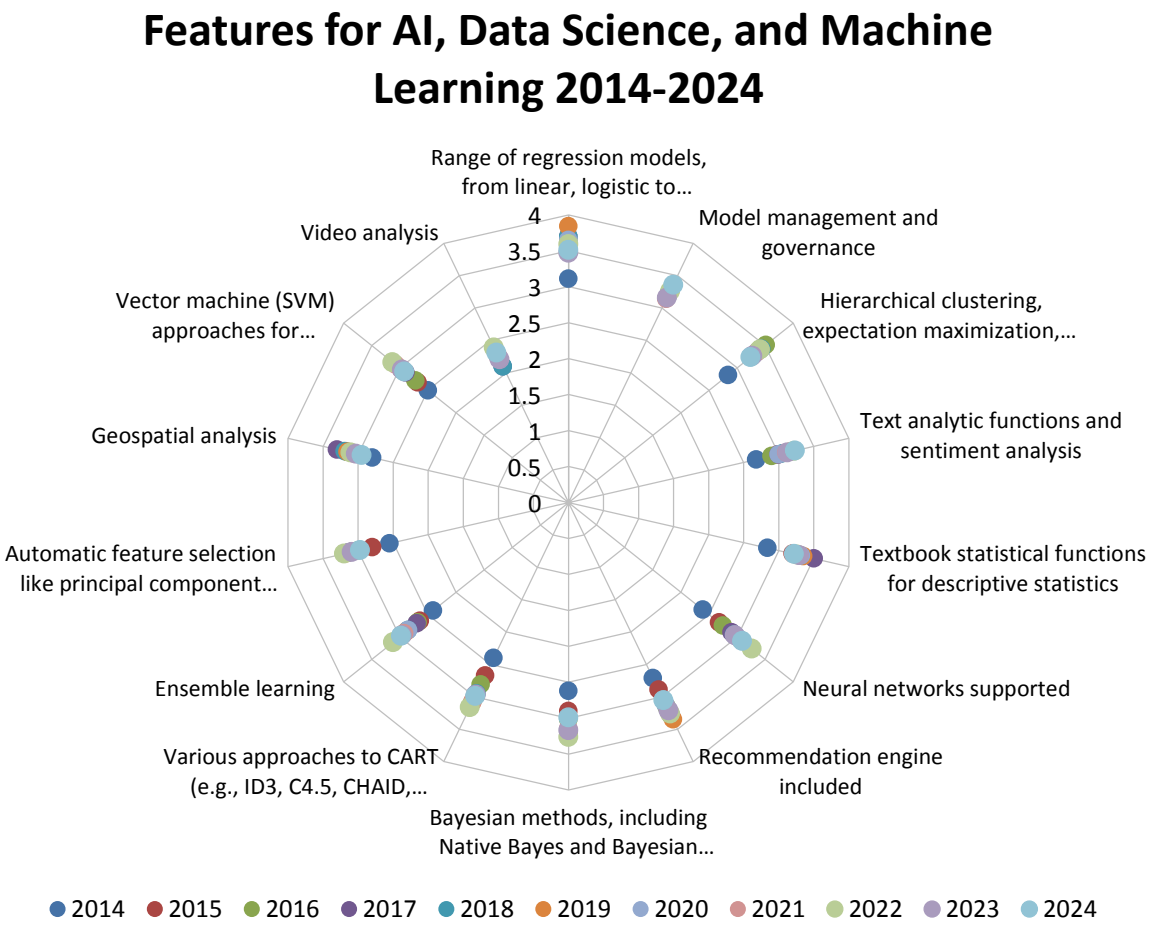


Figure 40 – Features for AI, data science, and machine learning 2014-2024

Features for AI, Data Science, and Machine learning by Geography

In 2024, many top features for addressing AI, data science, and machine learning see somewhat clustered importance across regions, and the top nine features are all at least at the level signifying *important* across all geographies (fig. 41). The top feature, *range of regression models,* is especially tightly clustered across geographies, with uniform importance midway roughly between *important* and *very important.* Among top features, *model explain-ability* and *outlier detection* narrowly receive the highest or equally highest scores from North America respondents. All remaining features receive the highest marks from respondents in Asia Pacific or Latin America. EMEA respondents most often report the lowest importance scores, especially for lower-ranked features.

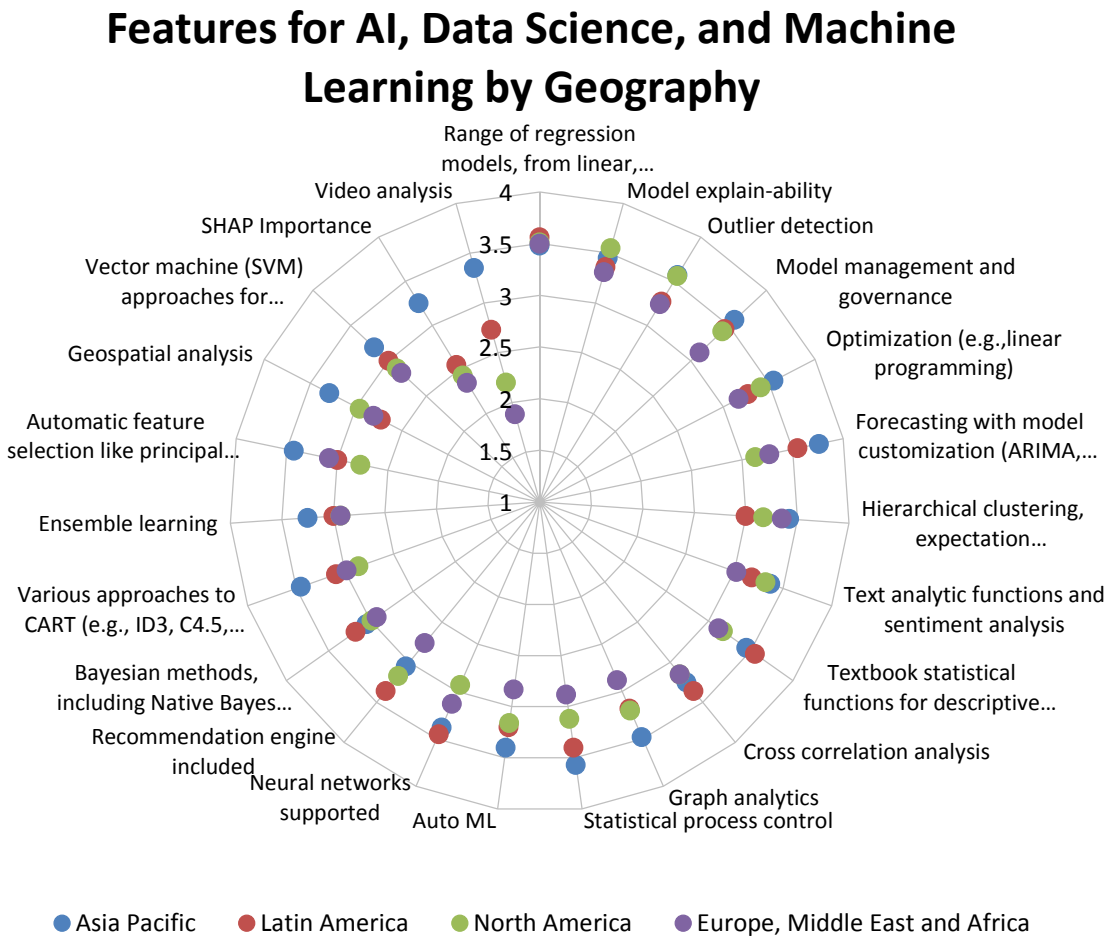


Figure 41 – Features for AI, data science, and machine learning by geography

Features for AI, Data Science, and Machine learning by Function

Viewed by function, *R&D* respondents most often assign the highest importance to most features for AI, data science, and machine learning in 2024 (fig. 42). As we observe in other measures dominated by *R&D* interest, this is a likely reflection of ongoing technical/analytical development, testing and deployment. *R&D* often shows outsized interest in top-ranked as well as training feature priorities, an indication of the discrete needs of organizations in specialized environments. *BICC* respondents are the next most represented, another indicator of internal deployment and demand. *Executive management* interest is also high, a finding which usually confirms enterprise attention and investment. Feature interest is lowest overall in *operations, IT*, and *finance*.

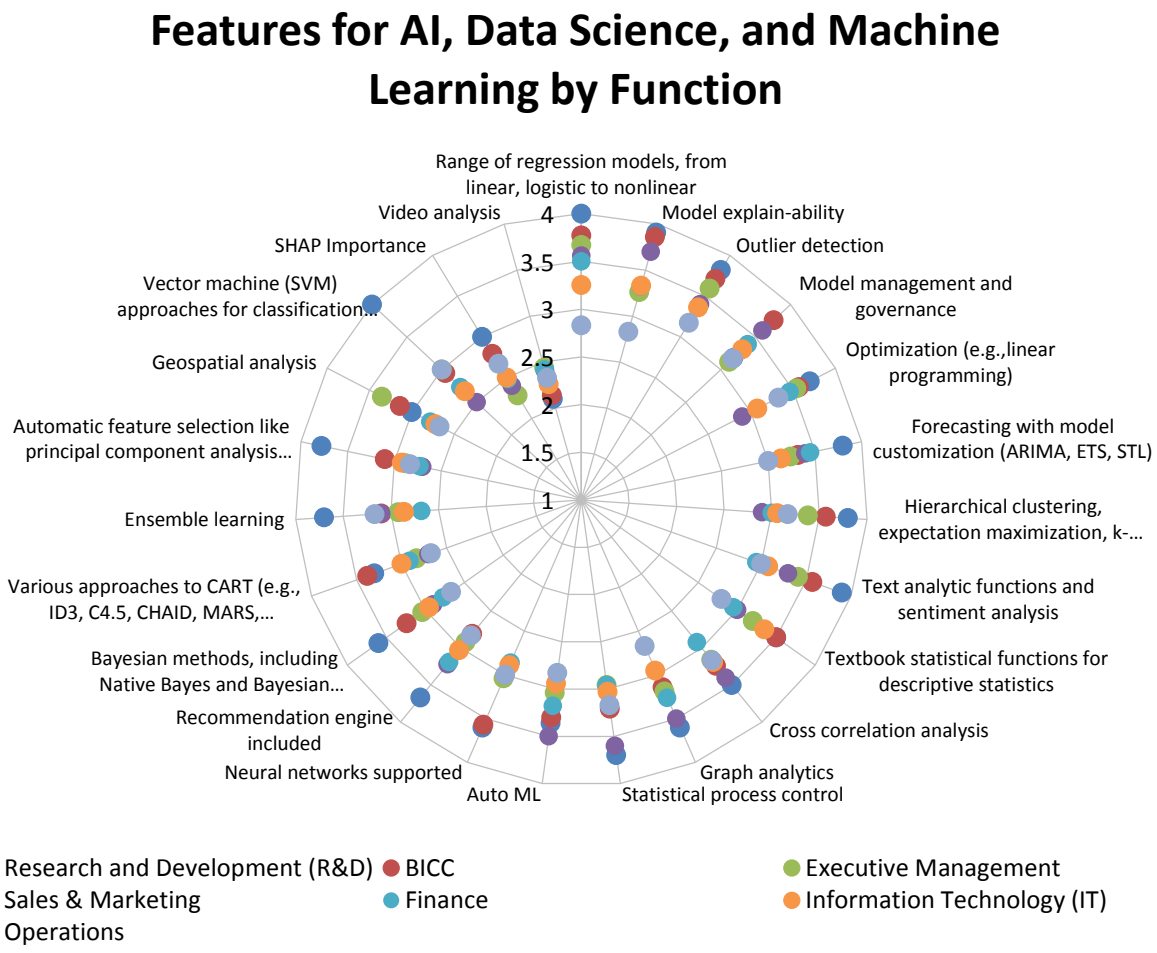


Figure 42 – Features for AI, data science and machine learning by function

Features for AI, Data Science, and Machine learning by Industry

Interest in features for AI, data science, and machine learning varies selectively by industry in 2024, with a mix of high and low interests (fig. 43). This year, the most-represented industries by weighted mean include *consumer services, financial services*, and *healthcare.* For example, *consumer services* is strongly represented in top features *range of regression models, outlier detection, optimization,* and *forecasting with model customization*. *Financial services* respondents lead interest in *model explain-ability, outlier detection*, and *model management and governance*. *Healthcare* respondents narrowly lead top pick *range of regression models*. *Retail and wholesale, education,* and *manufacturing* respondents report the lowest overall scores by weighted mean but also show compartmentalized interest in select areas of higher feature interest.

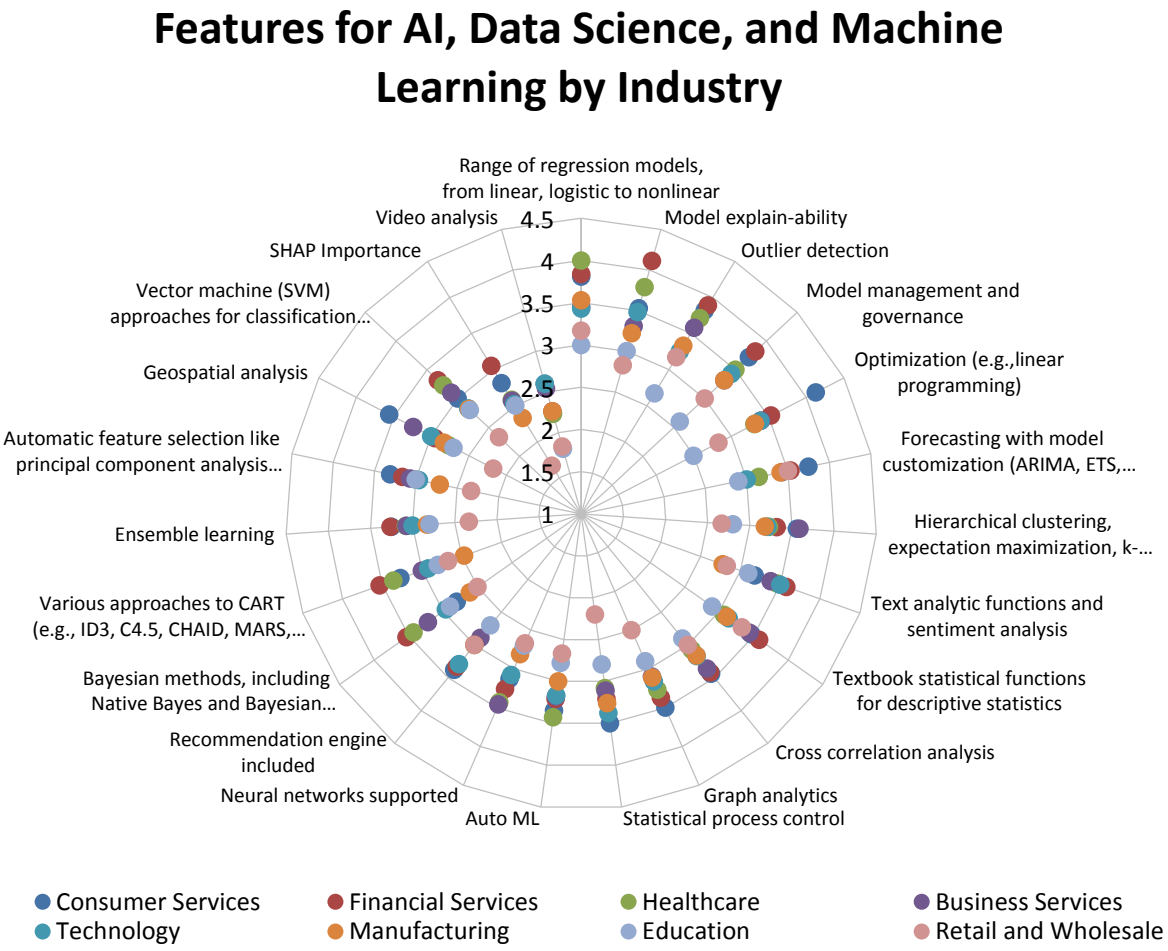


Figure 43 – Features for AI, data science, and machine learning by industry

Features for AI, Data Science, and Machine learning by Organization Size

Interest in analytical features for AI, data science, and machine learning consistently correlates with increasing organization size in 2024 (fig. 44). Indeed, very large organizations (> 10,000 employees) lead interest in every feature sampled this year, with some features, including *forecasting with model customization, hierarchical clustering, textbook statistical analysis, cross-correlation analysis*, and nearly all lower-ranked features receiving outsized importance scores compared to all smaller peer organizations. Features that show the most clustered importance across organizations of different size include *model explain-ability, graph analysis*, and *statistical process control*.

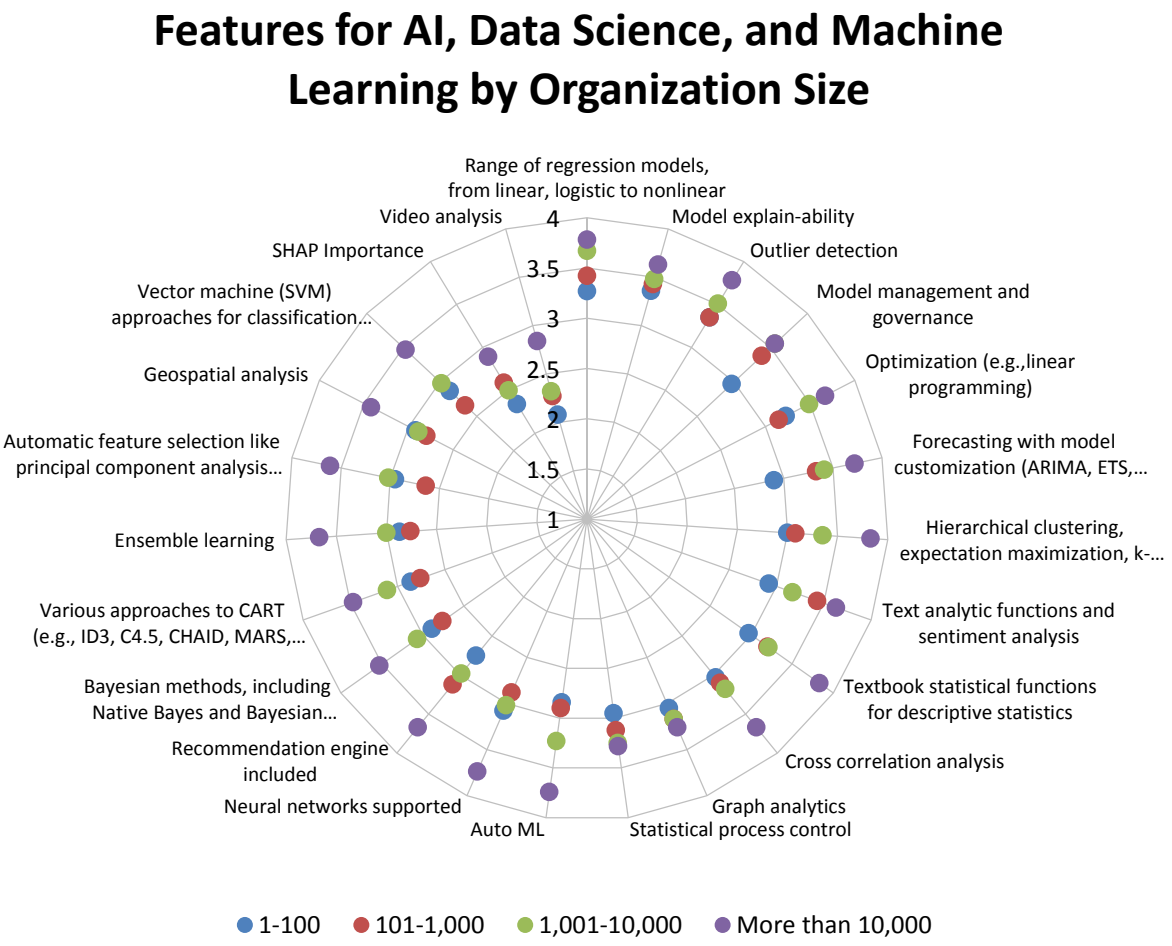


Figure 44 – Features for AI, data science, and machine learning by organization size

Data-Preparation Plans for AI, Data Science, and Machine learning

Our study addresses a detailed set of 16 data-preparation capabilities that support AI, data science, and machine learning (figs. 45-47).

All 16 data-preparation features earn very high attention among 2024 respondents (fig. 45). Features with the highest *critica*l or *very important* importance (from 52 percent up to 68 percent) this year include *detecting duplicates or outliers*, *cleansing and enriching source data*, *set operations*, *support for data type conversions, data lineage, profiling and quality, complex filtering*, and *data flows for multi-step transformations.* A second tier of importance includes features for *support for cutting, merging, and replacing values; text analytics and enrichment; ML-driven data prep recommendations*; and *semi-structured extraction and manipulation.* In total, all features are at least *important* to about 70 percent or far more respondents. (Also see industry support, fig. 73.)

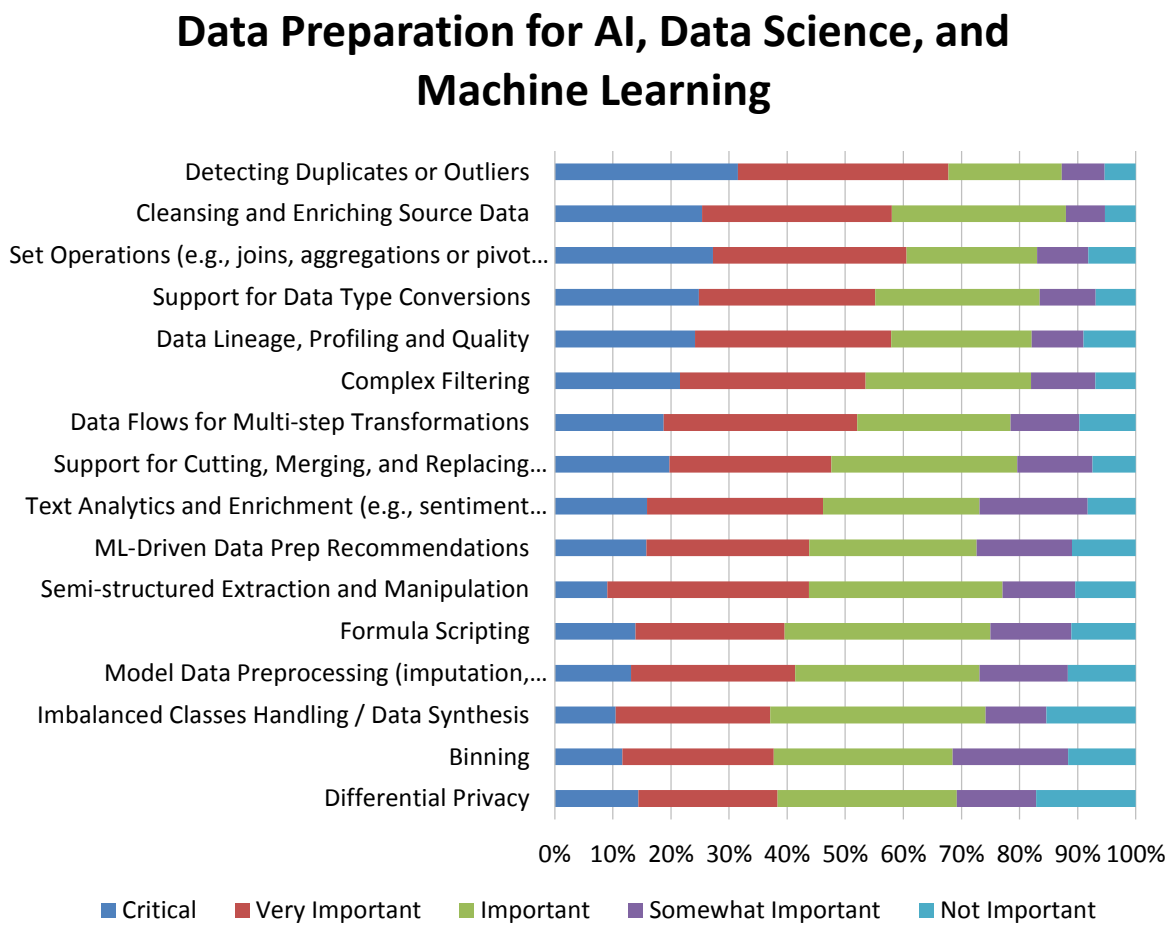


Figure 45 – Data preparation for AI, data science, and machine learning

Data Preparation for AI, Data Science, and Machine Learning by Industry

Interest in data preparation for AI, data science, and machine learning distributes inconsistently across vertical industries, with the highest scores by weighted mean seen in *consumer services, financial services, technology*, and *healthcare* (fig. 46). This year, *healthcare* respondents give their lone *very important* score to *range of regression models; consumer services* respondents assign a greater than *very important* score to *optimization (e.g., linear programming);* and *financial services* respondents assign a greater than *very important* score to *model explain-ability.* Nearly all other scores outside respondents in *retail and wholesale* and *education* are consistently above the level of *important* across all features.

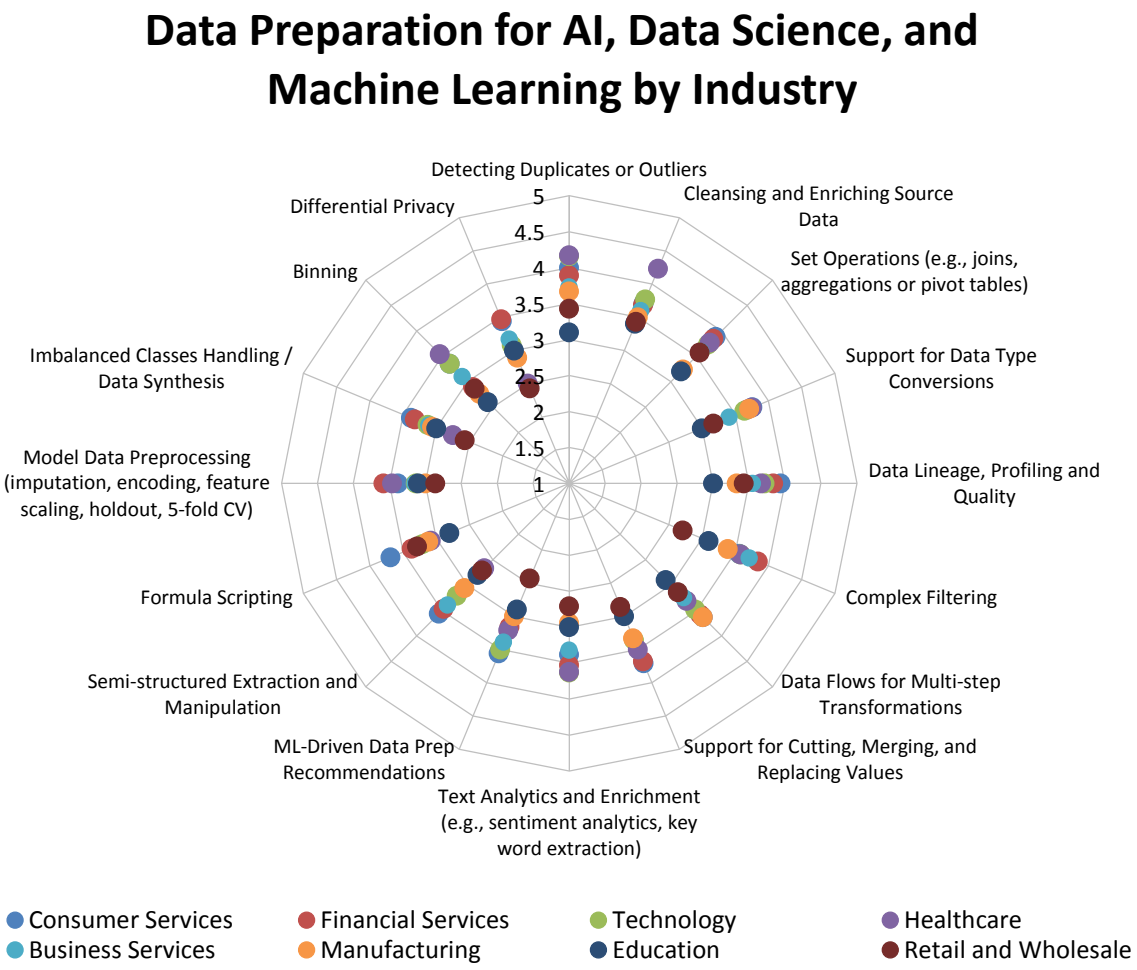


Figure 46 – Data preparation for AI, data science, and machine learning by industry

Data Preparation for AI, Data Science, and Machine Learning by Organization Size

In 2024, very large organizations (>10,000 employees) and large organizations (1,001-10,000 employees) most often give the highest scores for data-preparation features, though results are not uniform and are more tightly clustered compared to other measures by organization size (fig. 47). Said another way, feature interest and importance is more broadly applicable to organizations of different size compared to levels of deployment or importance. The top eight data preparation features are at least *important* to all organizations regardless of size.

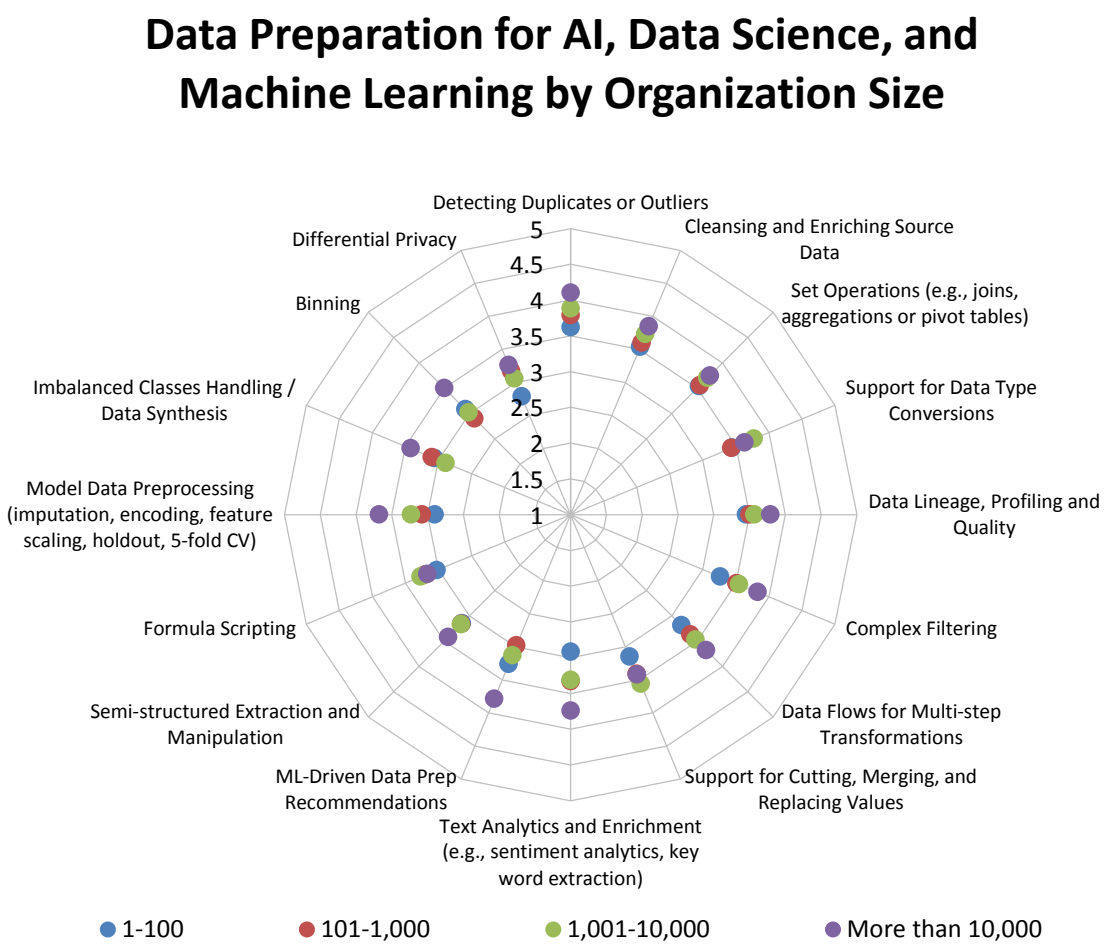


Figure 47 – Data preparation for AI, data science, and machine learning by organization size

Usability for AI, Data Science, and Machine Learning

Our study addresses a detailed set of 16 usability benefits that support AI, data science, and machine learning activities and processes (figs. 48-52).

Usability features generally address process or activity simplification and automation and, without exception, respondents give them high importance scores. All 18 2024 criteria we sample are, at minimum, *important* to at least 63 percent and as many as 86 percent of respondents (fig. 48). The top four features (*support for easy iteration*, *low code / no code with the ability to modify, access to advanced analytics*, and *python support*) are either *critical* or *very important* to half or more respondents, and other select usability features share similar support. Even the most esoteric usability features near the bottom of our list are at least *somewhat important* to close to 80 percent or more respondents (also see industry support for usability tools, fig. 74).

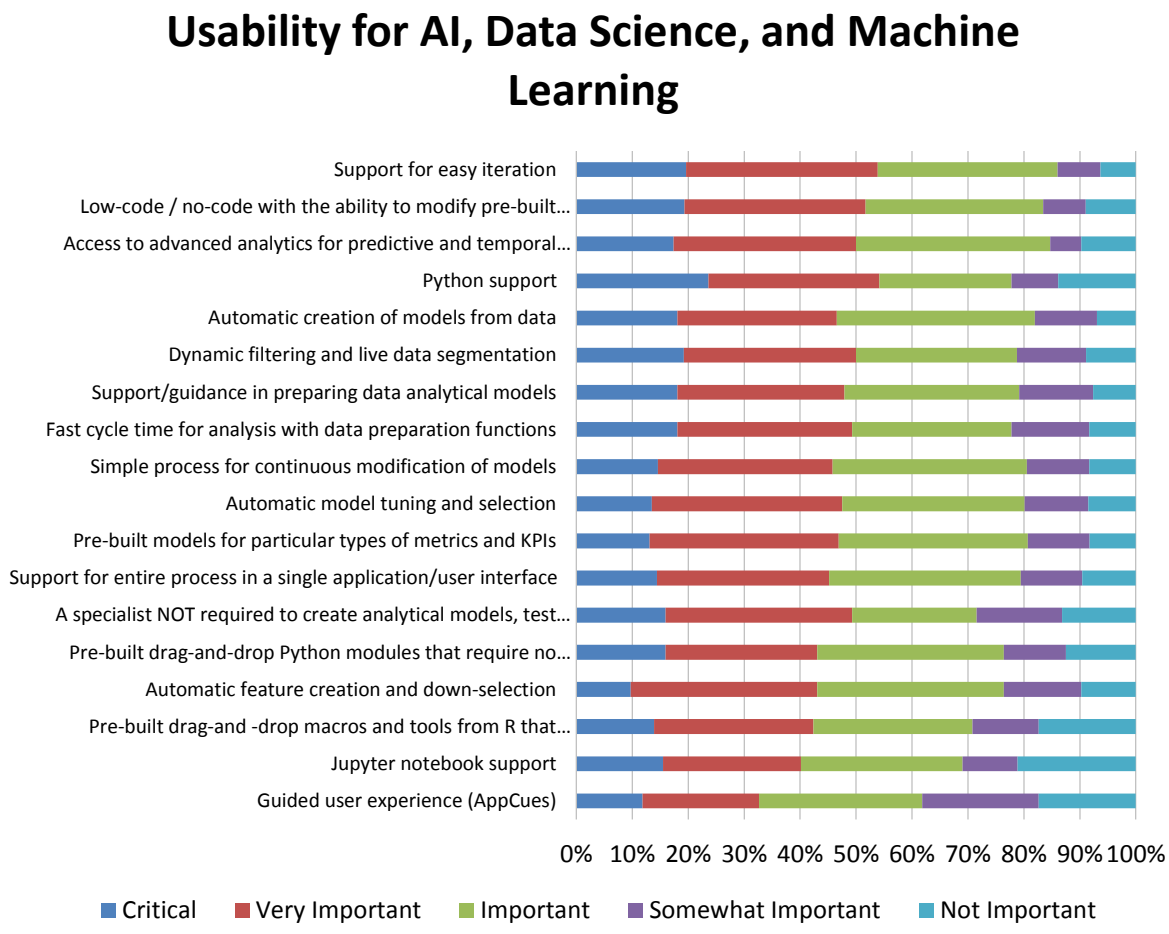


Figure 48 – Usability for AI, data science, and machine learning

Usability of AI, Data Science, and Machine Learning 2018-2024

Interest in most usability features for AI, data science, and machine learning mostly continues a five-year trend of decline in 2024, but only in degrees relative to historic measures (fig. 49). This year, only *automatic creation of models from data* shows relatively high momentum and interest. Perhaps more important, all usability features remain well above the level of *important* to respondents. But compared to peak interest seen in 2018-2019, sentiment toward nearly all usability features for AI, data science, and machine learning declines narrowly and gradually from 2020 to 2024 on an admittedly compressed scale.

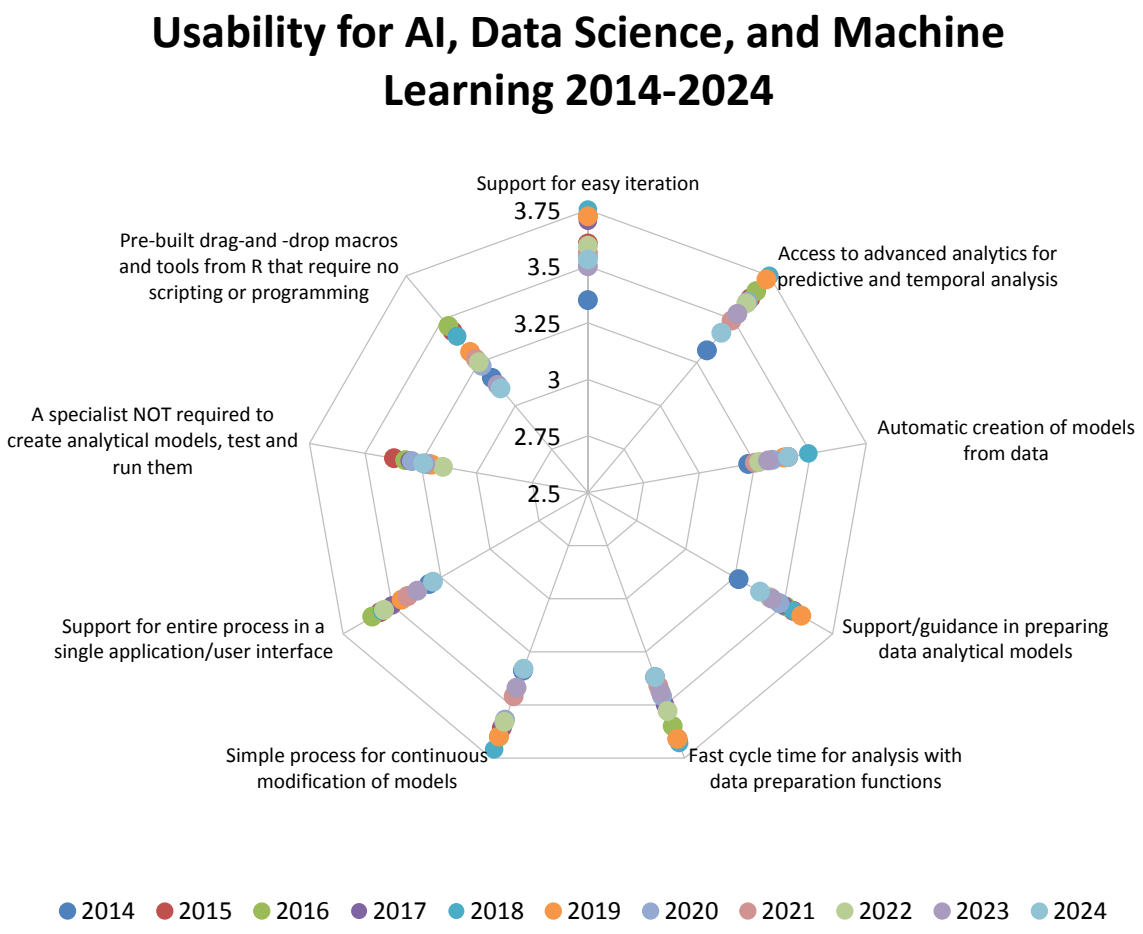


Figure 49 – Usability for AI, data science, and machine learning 2014-2024

Usability of AI, Data Science, and Machine Learning by Geography

The perceived importance of usability features for AI, data science, and machine learning in 2024 varies by geography, with regional interest that is nonetheless almost always lowest in EMEA (fig. 50). Unlike other geographic measures that tend to favor Asia Pacific and Latin America, North America respondents lead positive sentiment toward multiple usability features this year. Even so, the top 12 of 16 feature scores are above the level of *important* in all geographies. The most common interest (tightest clustering) across all geographies is for *low code / no code with the ability to modify,* and *access to advanced analytics for predictive and temporal analysis*.

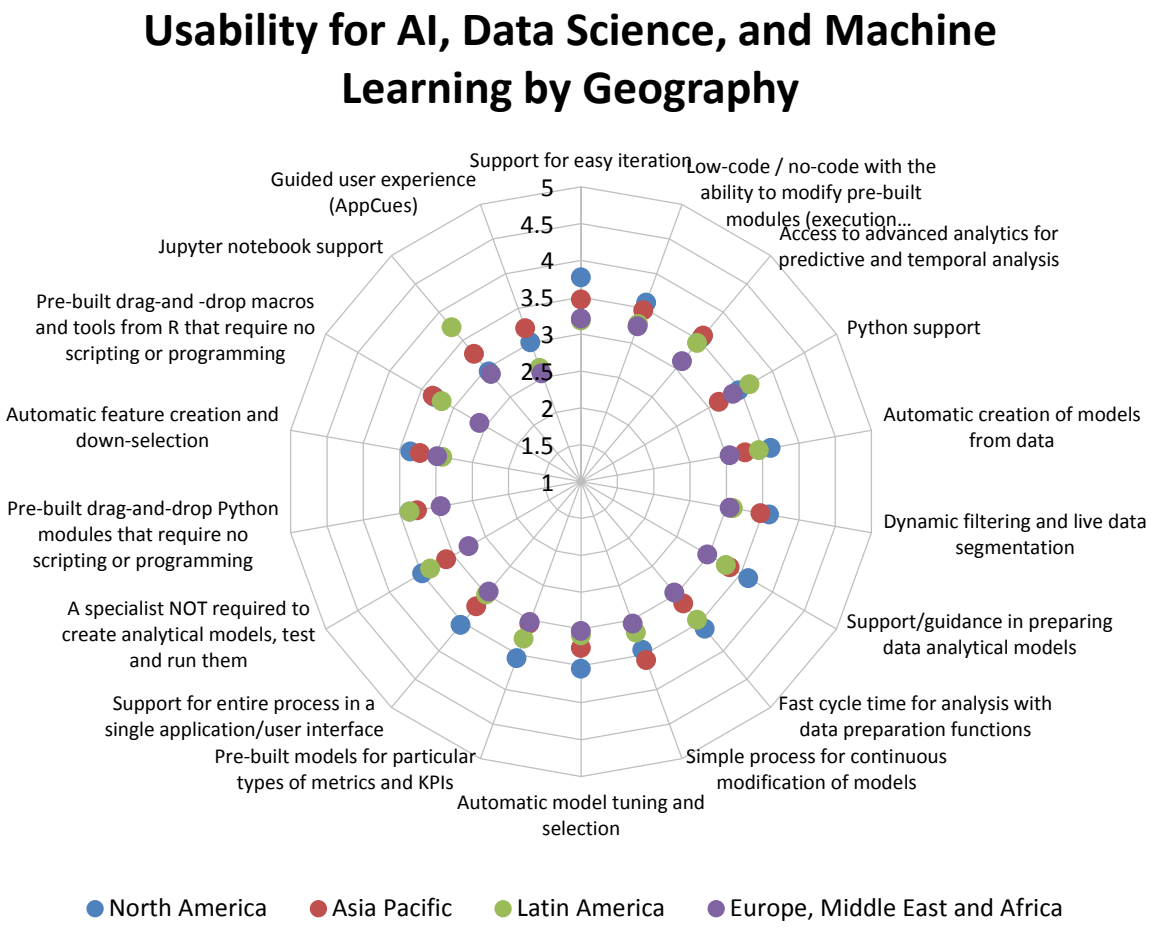


Figure 50 – Usability for AI, data science, and machine learning by geography

Usability of AI, Data Science, and Machine Learning by Function

Viewed by function, combined *executive management* and *R&D* interest suggest ongoing momentum and investment in AI, data science, and machine learning (fig. 51). *Executive management* respondents visibly lead interest in the top two usability features, *support for easy iteration* and *low code / no code*, and, along with *R&D*, post high or equally high scores for *automatic creation of models, dynamic filtering, support / guidance in preparing data analytical models* and *fast cycle time analysis. Sales and marketing* post the next highest overall scores, with high interest in multiple discrete usability features. *Finance* and *IT* are most often least interested in usability features.

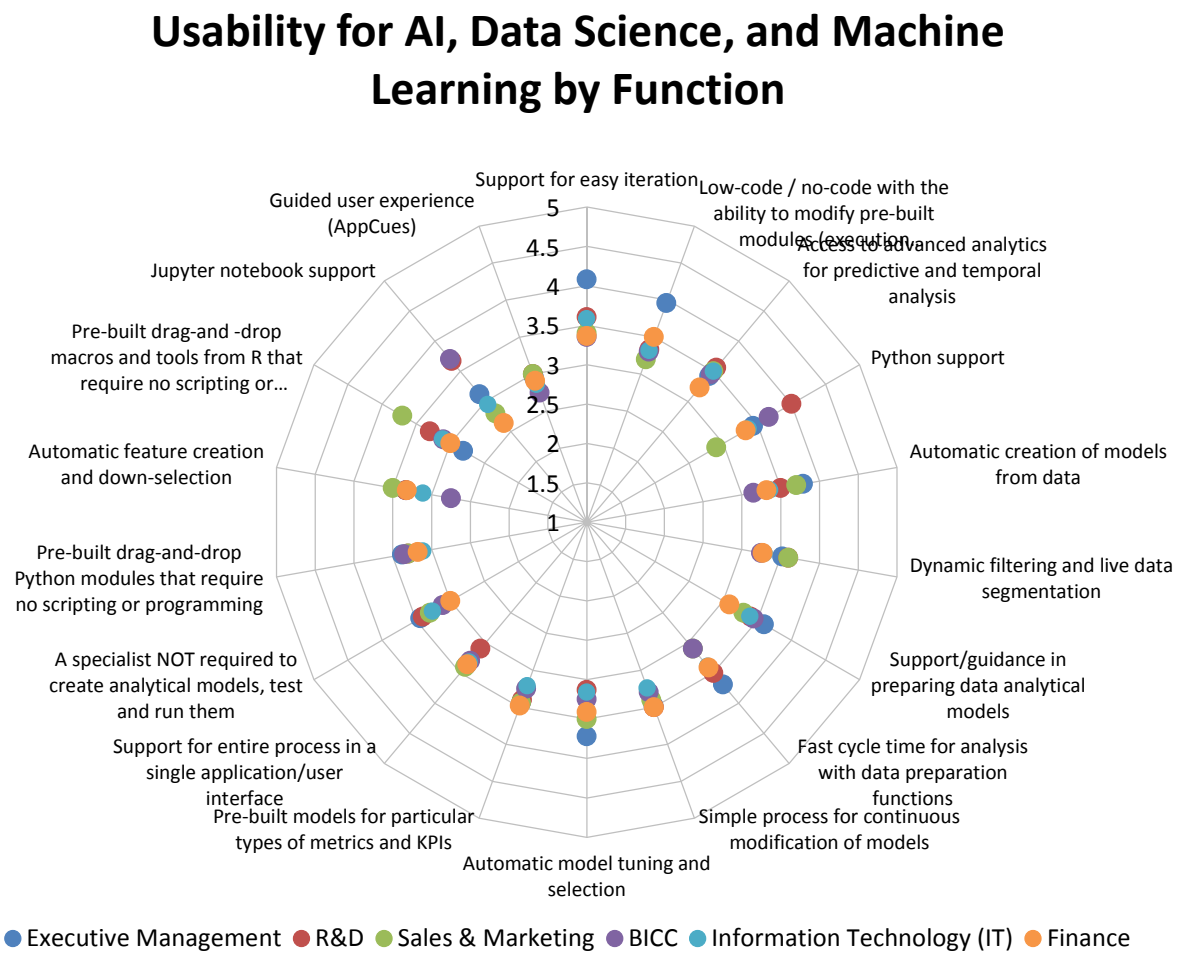


Figure 51 – Usability for AI, data science, and machine learning by function

Usability of AI, Data Science, and Machine Learning by Organization Size

Interest in usability features for AI, data science, and machine learning increases with global headcount to an observable if not always dramatic degree (fig. 52). In every case, interest in features is highest at very large organizations (>10,000 employees), almost always followed by large organizations (1,001-10,000 employees). Small (1-100 employees) or midsize (101-1,000 employees) organization interest is most often lowest overall. At least 14 of 18 usability features are at least *important* to organizations of any size.

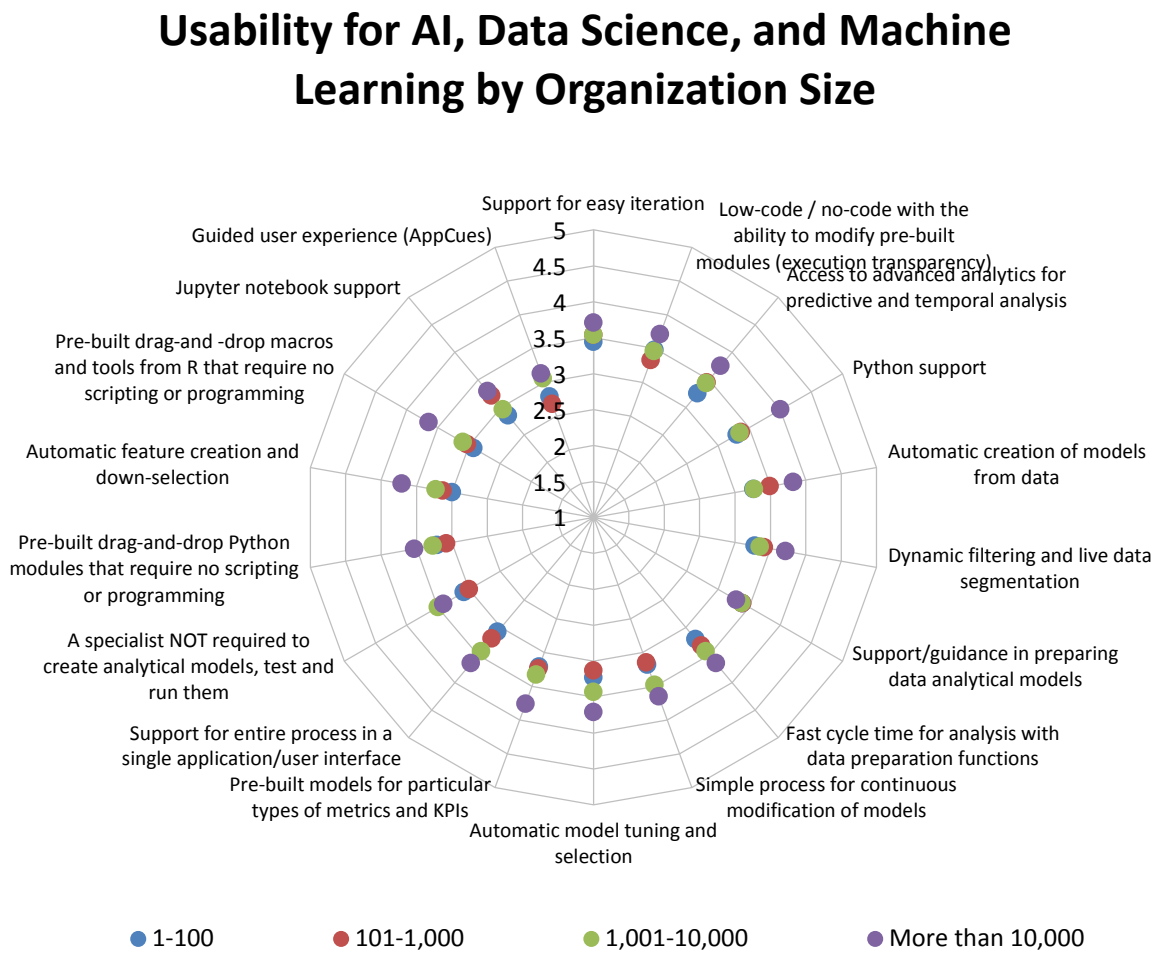


Figure 52 – Usability for AI, data science, and machine learning by organization size

Scalability of Data Science and Machine Learning

Our study addresses respondents' interest in a set of scalability technologies and architectures that support AI, data science, and machine learning (figs. 53-56). All 12 features sampled in 2024 are at least *important* to close to half or far more (up to 80 percent) of respondents (fig. 53). This year (like the last two years), two features, *in-database analytics* and *in-memory analytics,* are most important, after which we observe rising year-over-year interest in second tier features including *horizontal scaling, vertical scaling, multi-tenant cloud services, GPU acceleration, code generation supported*, and *h*ybrid / cloud bursting. (Also see industry support for scalability, fig. 75.)

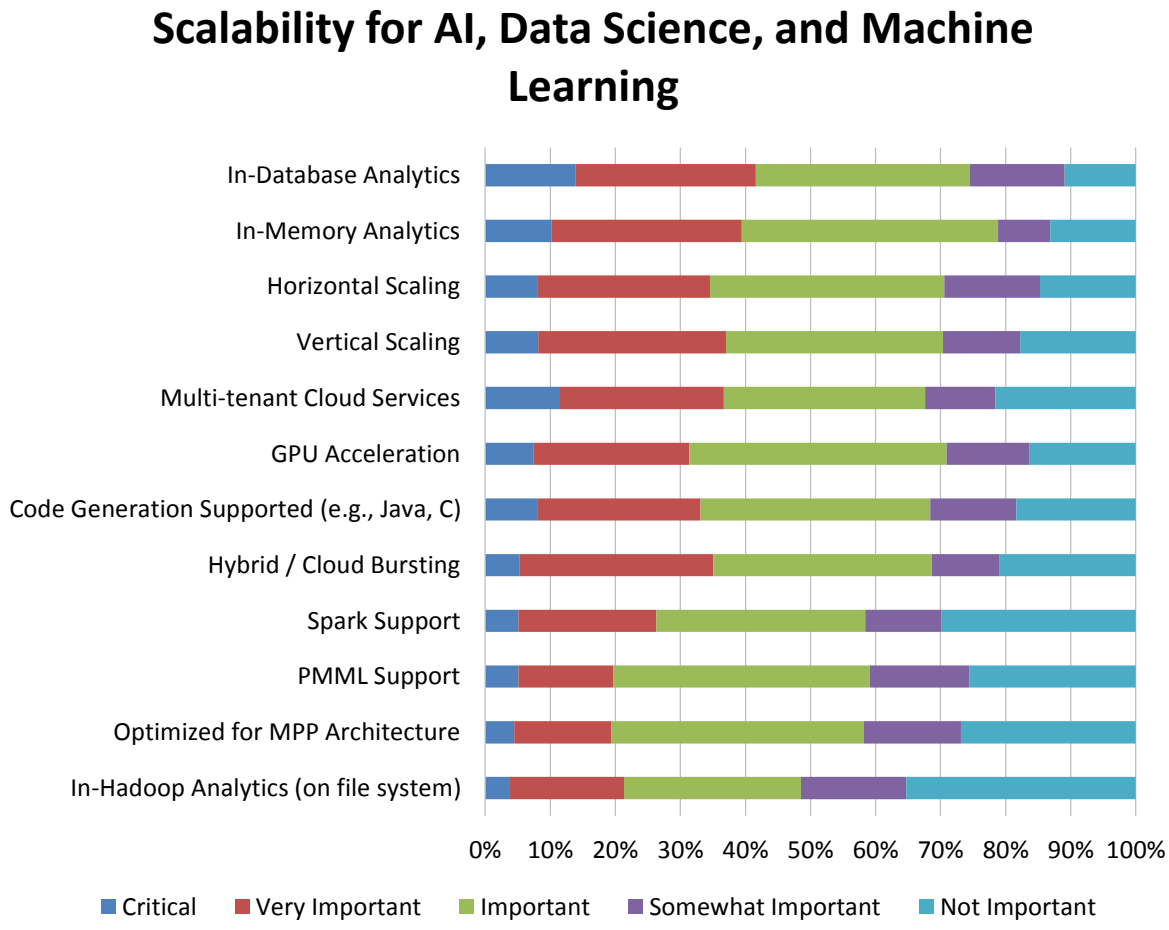


Figure 53 – Scalability for AI, data science, and machine learning

Scalability for AI, Data Science, and Machine Learning 2021-2023

Over time and in 2024, we observe mostly minor advancers and decliners in the 11 scalability features for AI, data science, and machine learning (fig. 54). Year-over-year advancers in this year include *GPU acceleration, horizontal scaling, hybrid / cloud bursting, code generation supported* and *PMML support*. Decliners in 2024 include current and long-time category leaders *in-memory analytics* and *in-database analytics,* along with *multitenant cloud services, Spark support, optimized for MPP architecture*, and *in-Hadoop analytics*. Rankings over time are largely sustained, with narrow changes that appear to show ongoing expansion and adoption of additional scalability features for AI, data science, and machine learning.

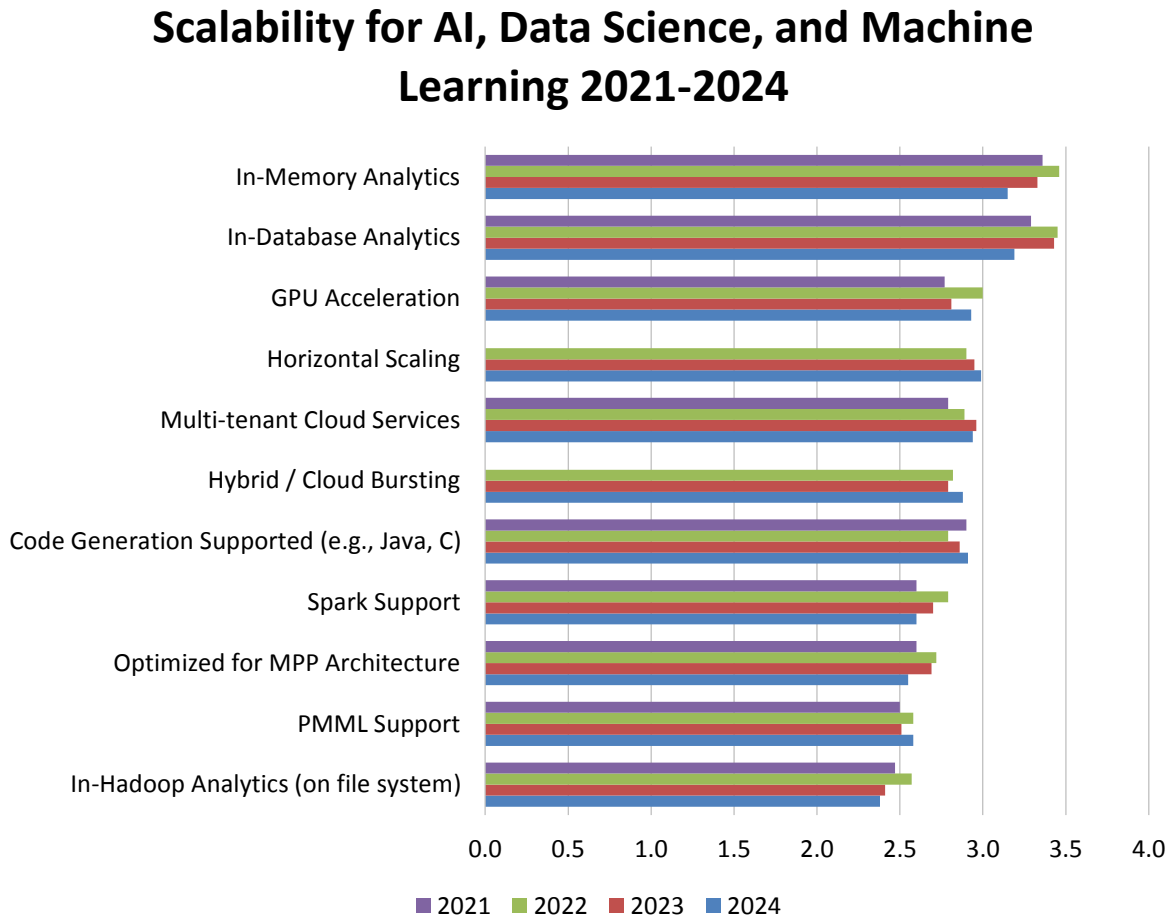


Figure 54 – Scalability for AI, data science and machine learning 2021-2024

Scalability for AI, Data Science, and Machine Learning by Function

Functional interest in scalability features for AI, data science, and machine learning is highest by weighted mean in *sales and marketing, R&D, IT*, and the *BICC*, a mix that implies front-office attention and high development/deployment interest and/or incipient demand (fig. 55). In 2024, *sales and marketing* reports the highest or equally highest interest in a range of conventional and more esoteric areas including *in-database analytics, vertical scaling, code generation supported, hybrid / cloud bursting, Spark support*, and *optimized for MPP architecture. R&D* respondents are also highly interested in *in-database analytics*, as well as *in-memory analytics, multi-tenant cloud services*, and *PMML support*. *BICC* respondents are most attentive to frontline *in-database analytics* and *in-memory analytics*, and pay less attention to other areas. *Finance* and *executive management* are least interested in scalability for AI, data science, and machine learning.

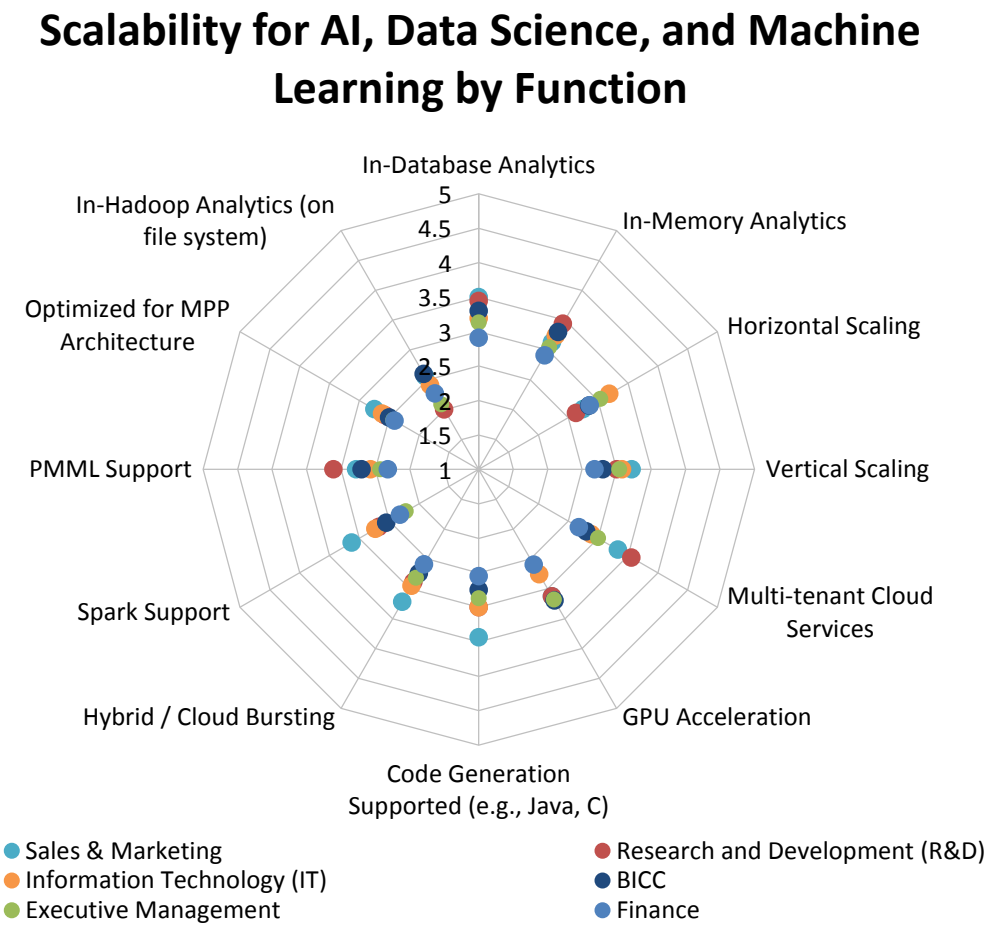


Figure 55 – Scalability of data science and machine learning by function

Scalability for AI, Data Science, and Machine Learning by Organization Size

In 2024 (and in all earlier studies), interest in scalability capabilities for data science and machine learning correlates strongly to global headcount (fig. 56). Very large organizations (> 10,000 employees) lead all12 categories, although interest clusters well across most scalability features, particularly the top 10. *In-database analytics* and *in-memory analytics* remain the areas of greatest universal interest across organizations of any size. In most cases, interest is second-highest in large organizations (1,001-10,000 employees), followed by either midsize (101-1,000 employees) or small (1-100 employees) organizations.

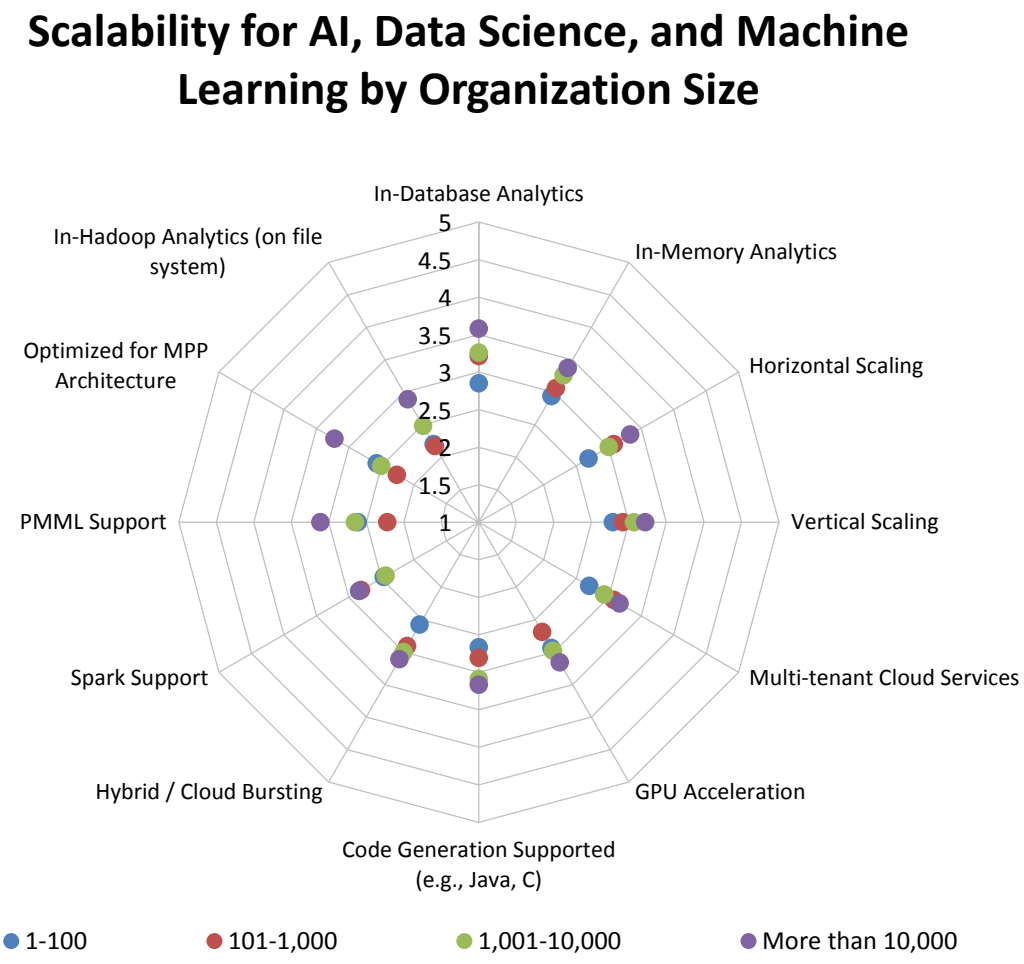


Figure 56 – Scalability of data science and machine learning by organization size

Neural Networks for AI, Data Science, and Machine Learning

We asked organizations to gauge their interest in 11 types or aspects of neural networks in the context of AI, data science, and machine learning (fig. 57). The top four picks in 2024 are *artificial neural network, transformer network, convolutional neural networks*, and *feed-forward deep learning.* Each of these four picks is seen as *critical* or *very important* to 20-23 percent of respondents, and at least *important* to 39-41 percent of respondents. Among remaining networks, *long short-term memory* and *recursive neural networks* are most popular. Also worth noting, between 40-49 percent of respondents *don't know* the importance of any of the neural network choices, though just 6-13 percent think any are *not important*. (Also see industry support, fig. 72)

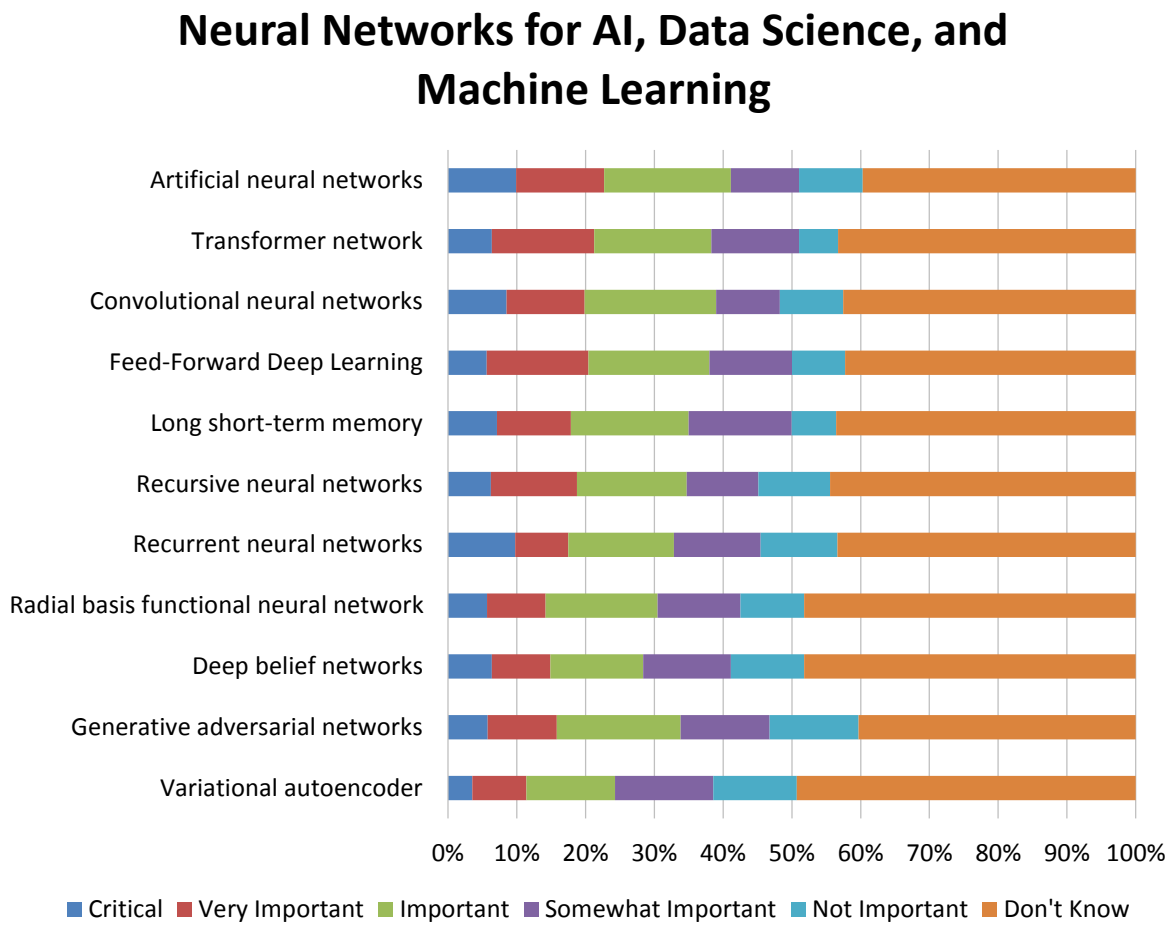


Figure 57 – Neural networks for AI, data science, and machine learning

Neural Networks for AI, Data Science, and Machine Learning by Industry

Attitudes toward types of neural networks in the context of AI, data science, and machine learning vary by industry, with 2024 weighted-mean interest led by respondents in *manufacturing, consumer services, technology*, and *business services* (fig. 58). This year, for example, respondents in *manufacturing* lead or nearly lead the greatest interest in *artificial neural networks*, *recursive neural networks, recurrent neural networks, radial basis functional neural networks* and *variable autoencoder. Consumer services* respondents show high especially high interest in *feed-forward deep learning* and *long short-term memory*, as well as *transformer networks* and *artificial neural networks*. Technology respondents narrowly lead all industries in interest in *artificial neural networks* and *deep belief networks*. *Retail and wholesale* respondents are least interested in neural networks by industry

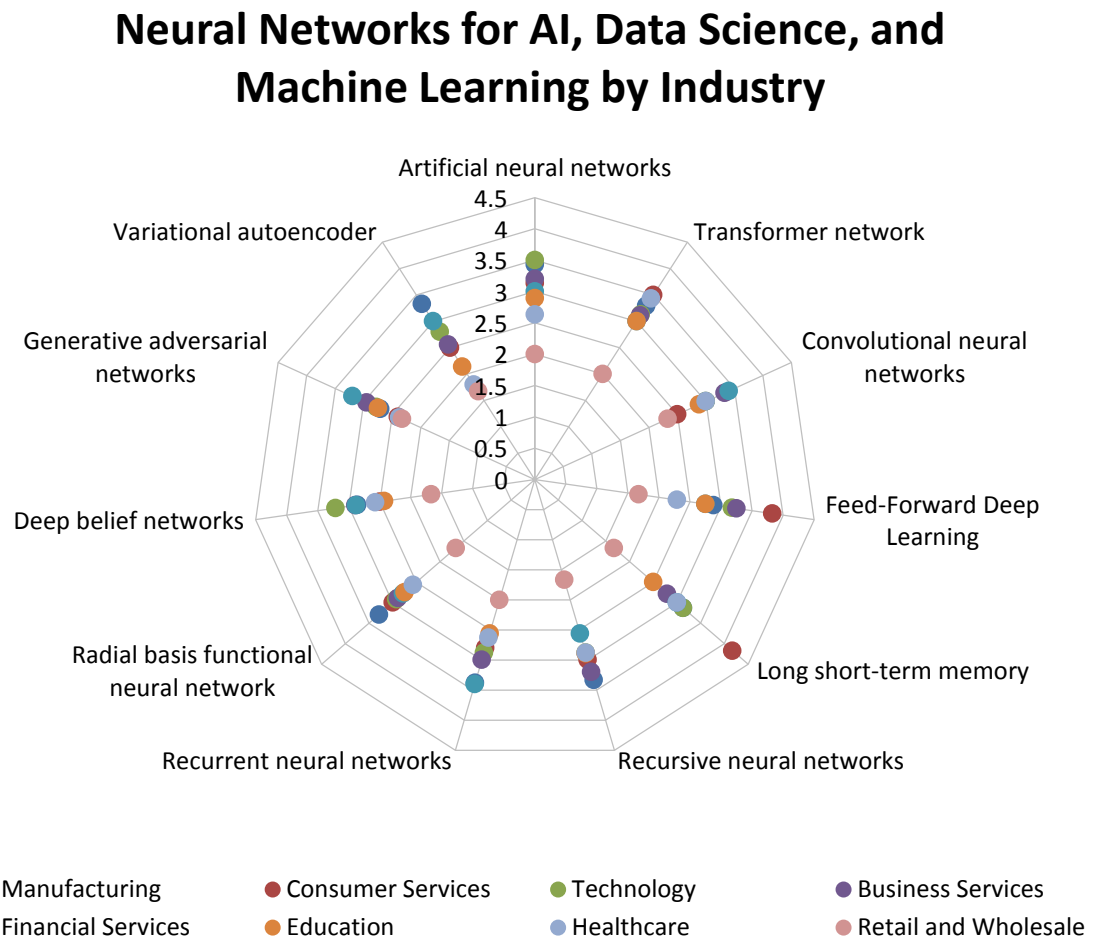


Figure 58 – Neural networks for AI, data science, and machine learning by industry

Generative AI

Generative AI uses machine-learning algorithms (e.g., neural networks) to create large language models (LLMs), enabling the generation of new and original data, images, text, or programming code that best approximates to the training data used. In a rapidly expanding class of products and services, prominent examples such as ChatGPT are available as a stand-alone offering or extension with and plug-ins for their enterprise applications—including many data and analytics products.

Year-over-year, 2024 current and future use of generative AI is expanding rapidly, with far more in-production, experimental use, and 12-month plans than in 2023 (fig. 59). In-production and experimental use grew to 49 percent this year, compared to 29 percent in 2023. In addition, *don't know*, and *no plans* respondents declined dramatically, from a net 58 percent in 2023 to 28 percent in 2024.

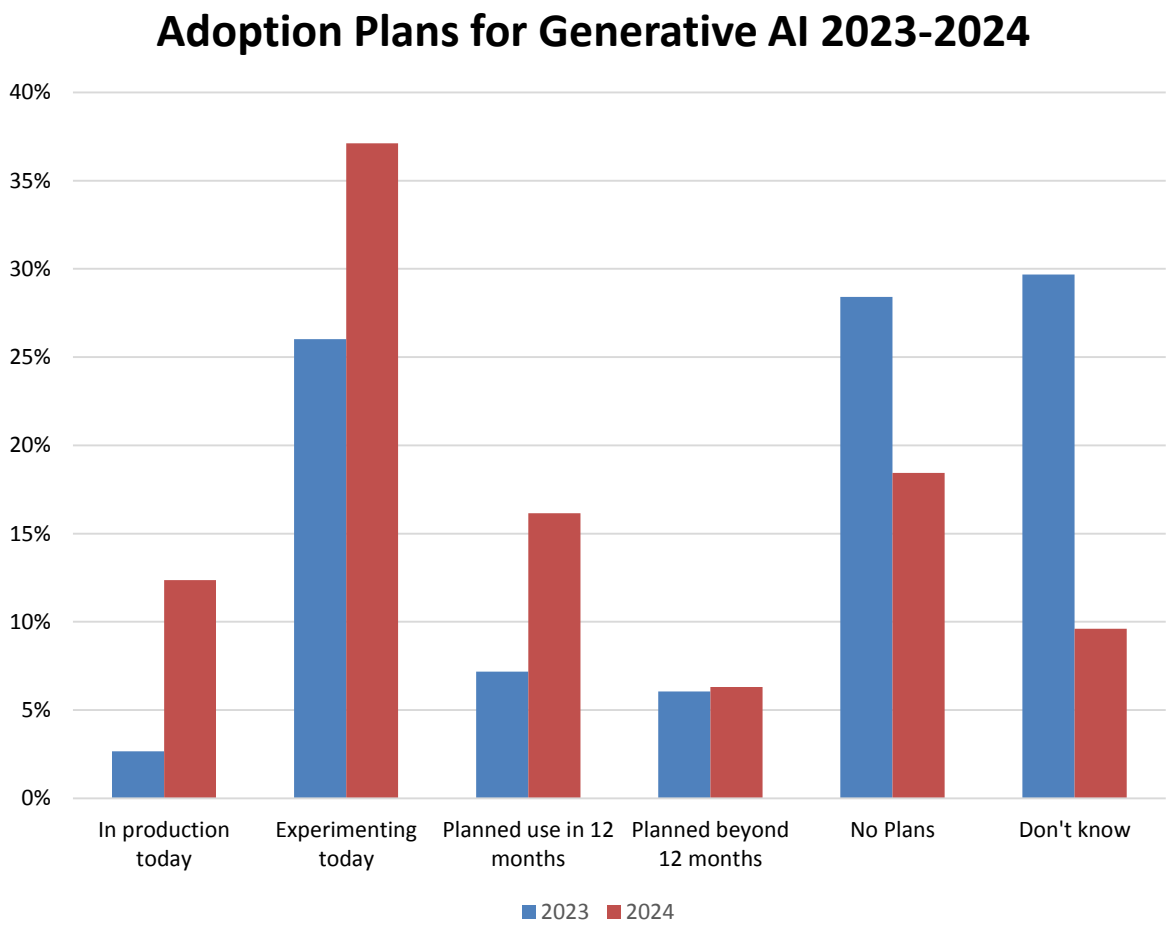


Figure 59 – Adoption plans for generative AI 2023-2024

Generative AI by Geography

Adoption plans for generative AI vary by geography in 2024, with current use and projections highest in Latin America (weighted mean 3.5), followed by North America (3.3), Asia Pacific (3.1), and finally EMEA (3.0) (fig. 60). This year, admittedly nascent in-production adoption is highest in North America (15 percent), followed by Latin America (11 percent), Asia Pacific (9 percent), and EMEA (8 percent). *Don't know* responses tellingly declined to just 7-11 percent of responses.

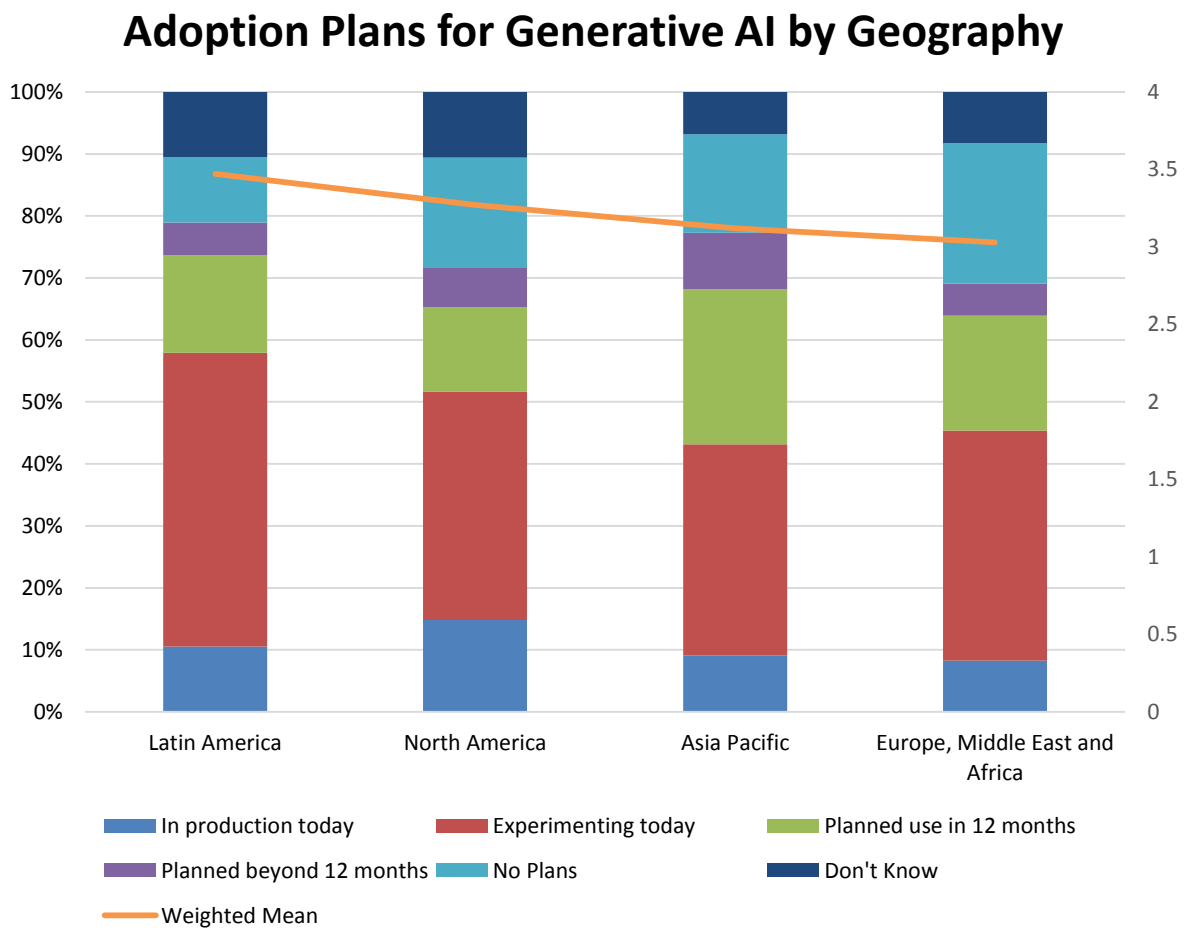


Figure 60 – Adoption plans for generative AI by geography

Generative AI by Function

A view of adoption plans by functions reveals new front-office interest in 2024 that is highest in *sales and marketing*, followed by more expected strength in the BICC, R&D, and, as we also found last year, in *executive management* (fig. 61). *In-production sales and marketing* use reaches 33 percent this year, well higher than any other function, and 75 percent when combined with *experimental use.* As mentioned, strong *executive management* use coupled with *R&D* and *BICC* activity indicates both active use and incipient demand. As was the case in 2023, *finance* reports the lowest adoption plans for generative AI by function. (Not shown, in-production use in 2023 was 5 percent or less across all functions).

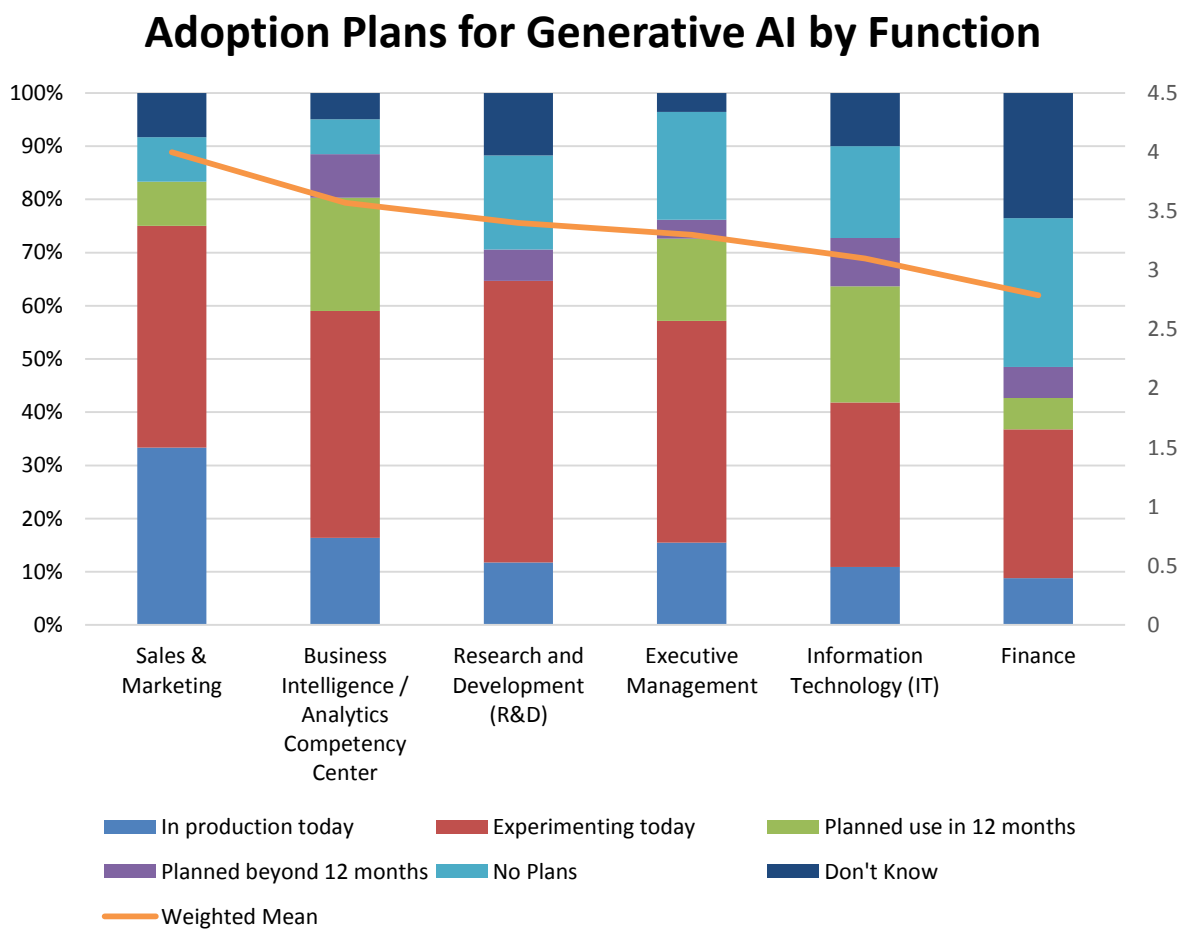


Figure 61 – Adoption plans for generative AI by function

Generative AI by Industry

Though in-production use remains very low, 2024 plans for adoption of generative AI are very strong across most industries with the exception of two: *retail and wholesale*, and, more surprisingly, *financial services* (fig. 62). This year, *consumer services* reports the strongest plans by weighted mean (3.6, above midway between very *important* and *very important*), followed by *healthcare* (3.5), *technology* (3.4), *education* (3.3), *business services* (3.2), and *manufacturing* (3.1). *Technology* respondents report the highest *in-production* use today (19 percent), well above all other industries.

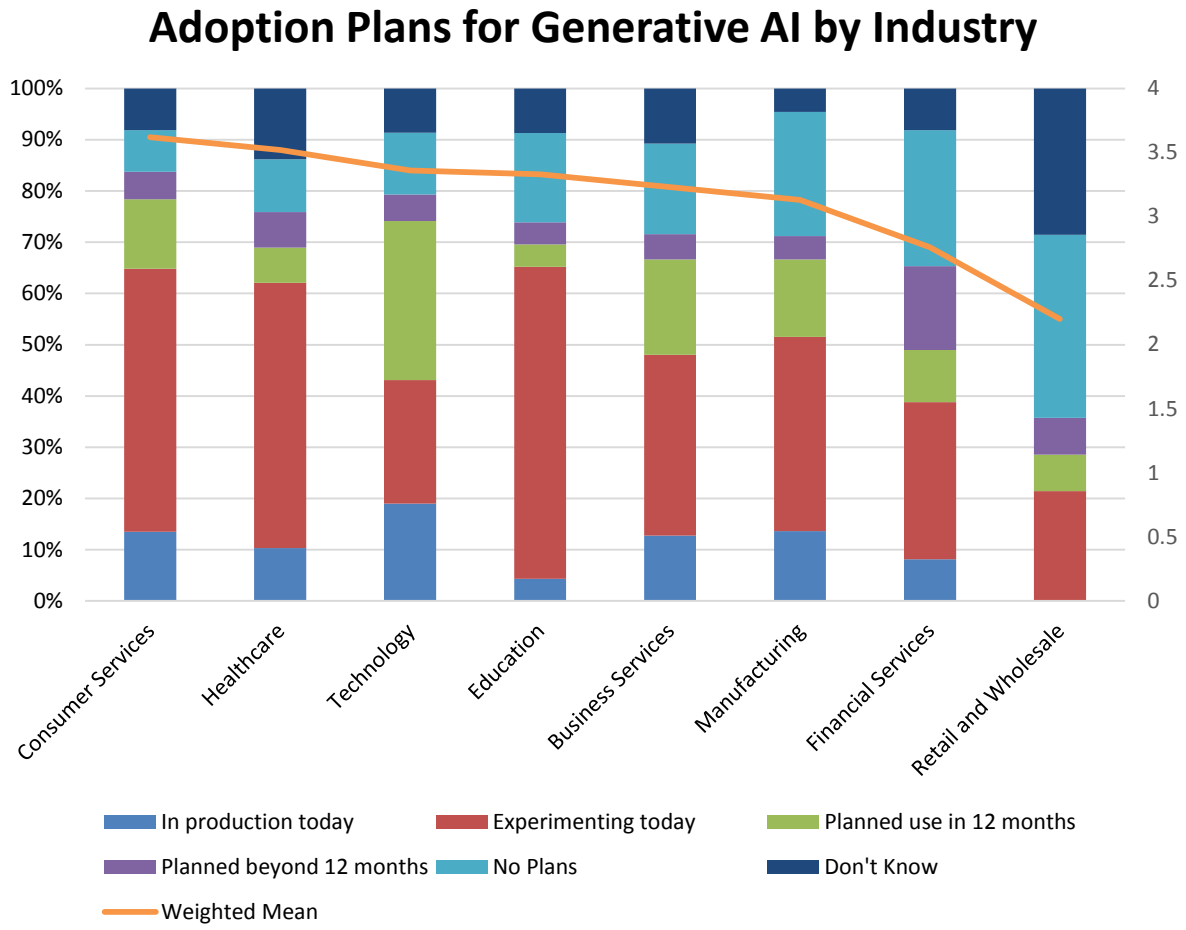


Figure 62 – Adoption plans for generative AI by industry

Generative AI by Organization Size

Viewed by organization size, adoption plans for generative AI increase with global headcount in 2024 (fig. 63). The largest organizations (> 10,000 employees) report the highest adoption by weighted mean (3.6) followed by large (1,001-10,000 employees) at 3.2, midsize (101-1,000 employees), also at 3.2 and small (1-100 employees) organizations at 3.0. Very large organizations also report the highest in-production use (18 percent), compared to 13 percent at small, 12 percent at large, and 9 percent at midsize organizations. Between 31-43 percent of organizations of any size are experimenting with generative AI.

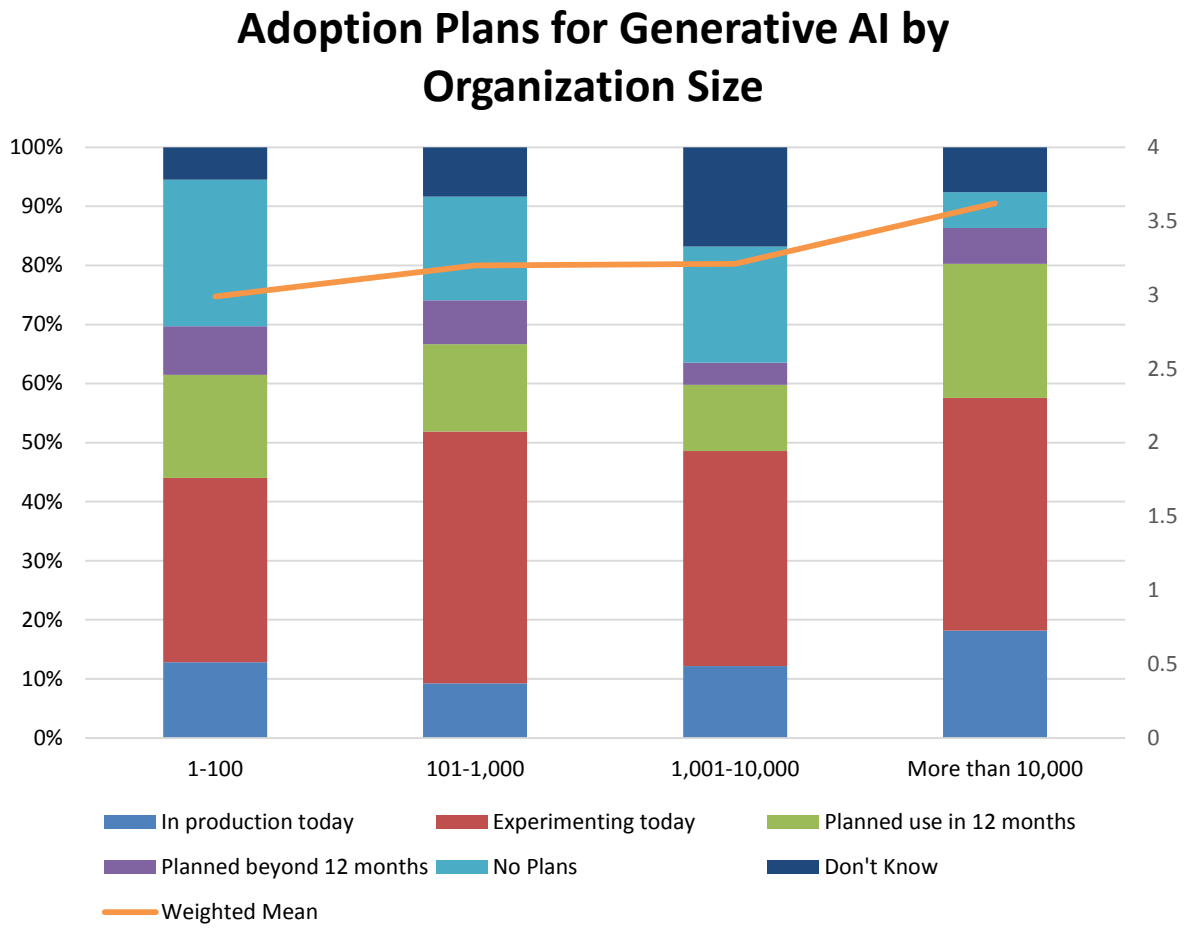


Figure 63 – Adoption plans for generative AI by organization size

Generative AI by Data Literacy

Organizations that perceive the highest accomplishment in the pursuit of data literacy are most likely to be adopters of generative AI by measures of both weighted mean and current use (fig. 64). In 2024, organizations with *extremely high* data literacy are well more likely (23 percent) to be in-production users of generative AI than are *high literacy* (15 percent), *moderate literacy* (11 percent), or *low and very low* literacy (10 percent). The implications of this finding are intuitive: those that are able to explore, understand, and communicate with data in a meaningful way are likely most able to adopt, learn, and advance the use of generative AI. Even so, more than half of respondents at *low or very low* data literacy organizations are either using in-production, or experimenting with generative AI.

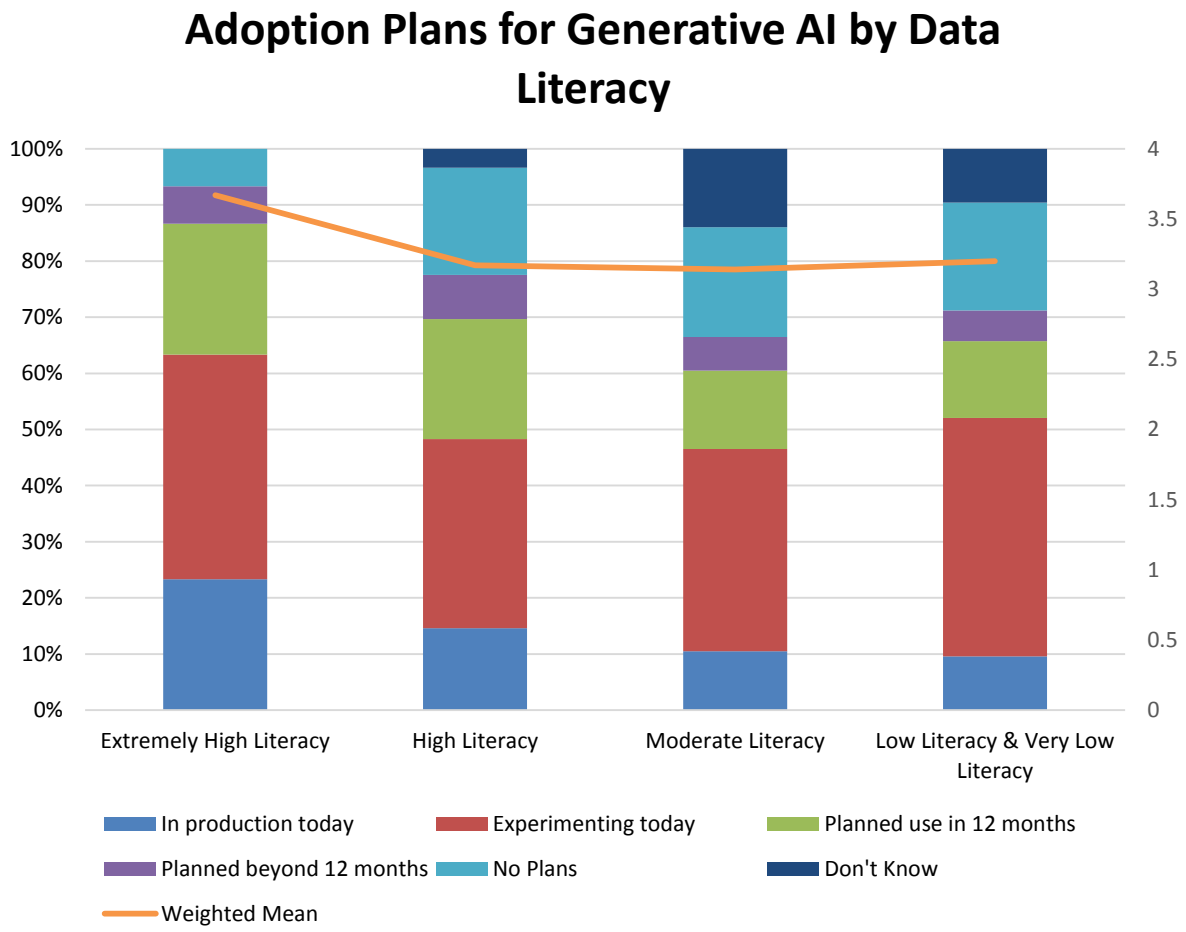


Figure 64 – Adoption plans for generative AI by data literacy

Open Source and AI, Data Science, and Machine Learning

Our last four annual studies of open-source AI, data science, and machine learning track a level of interest that is increasing, consistently at or above the level of *important,* but not advancing rapidly (fig. 65). Between 2021 and 2023, 60-70 percent reported the topics to be, at minimum, *important*, though this summary finding (along with standalone scores of *critical* importance), declined somewhat between 2022 and 2023. In 2024, sentiment again reverses to another high, with 72 percent considering the topic at least *important. A*s a relatively new set of mainstream topics, sentiment is consistently favorable, if not conclusive, toward the future impact of these technologies. Also, the hype and excitement surrounding AI settled somewhat as immediate implications and use cases are still being digested. In sum, the gap between conception and execution of artificial intelligence remains an unexhausted and incomplete discipline that we do not expect to recede in importance. (Also see industry support, fig. 79).

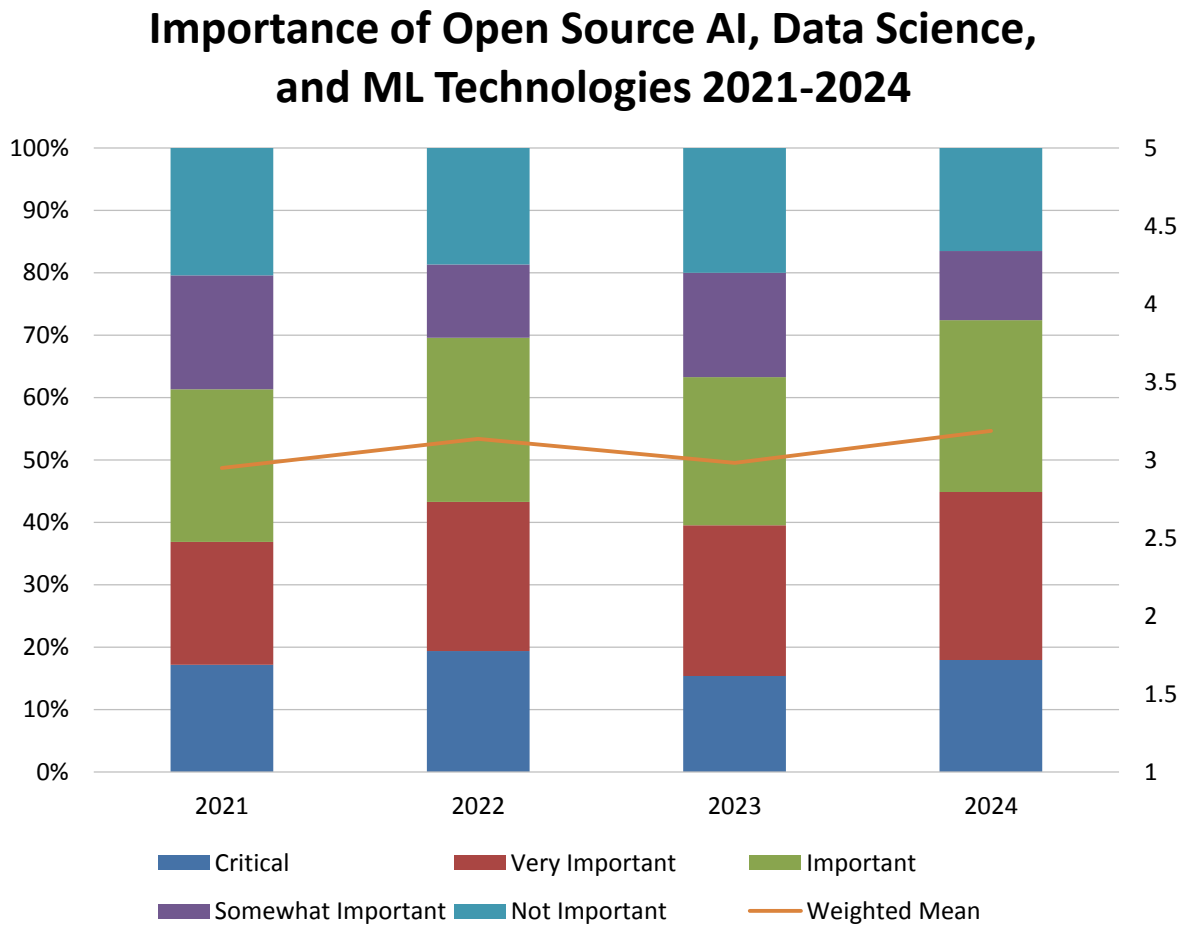


Figure 65 – Importance of open source AI, data science, and ML technologies 2021-2024

Open-Source Features for AI, Data Science, and Machine Learning

We asked organizations to gauge their interest in 17 (two new added in 2024) open-source features for AI, data science, and machine learning (fig. 67). The top four features (*Databricks*, [up from 3rd in 2023], *Azure data factory*, *support for Apache big data services*, and *Google Dataflow [up from 5th in 2023]*) are at least *important* to 55-59 percent of respondents. The top 12 open-source features are at least *important* to more than 40 percent of respondents. All but three of 17 services are, at minimum, *somewhat important* to half or more respondents.

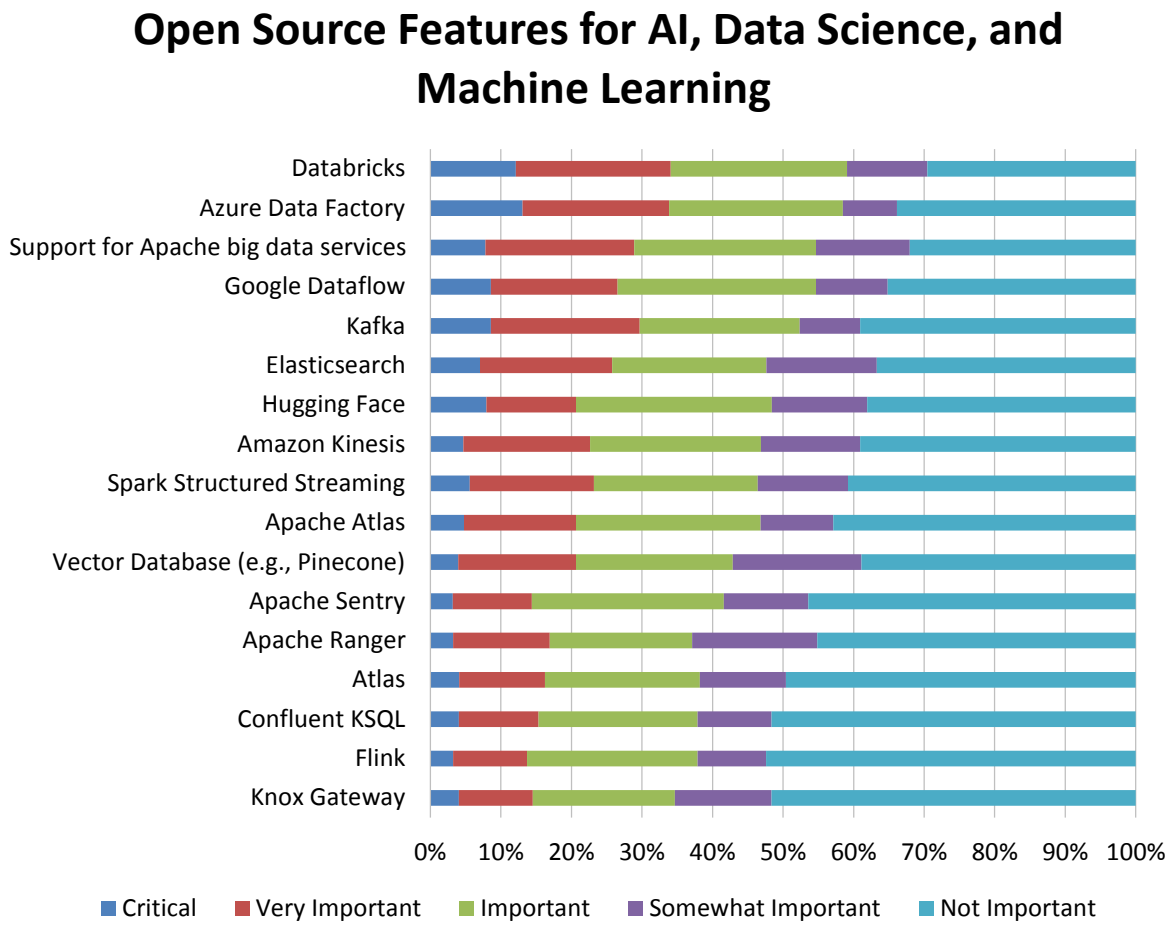


Figure 66 – Open source features for AI, data science, and machine learning

Open-Source AI, Statistical, and Machine Learning Technologies

We asked organizations to gauge their interest in 17 (six new added since 2022) AI, statistical, and machine learning technologies (fig. 67). Within this category, all technologies have relevant importance, led by *Pandas*, *Tensorflow*, *R language*, and *Pytorch,* all of which are, at minimum, important to 61-64 percent of respondents. *Anaconda*, *scikit-learn*, *MLflow*, *Keras*, and *Spark Mlib* are the next most important. The top 11 technologies are at least *somewhat important* to half or far more respondents. (Also see industry importance, fig. 67.)

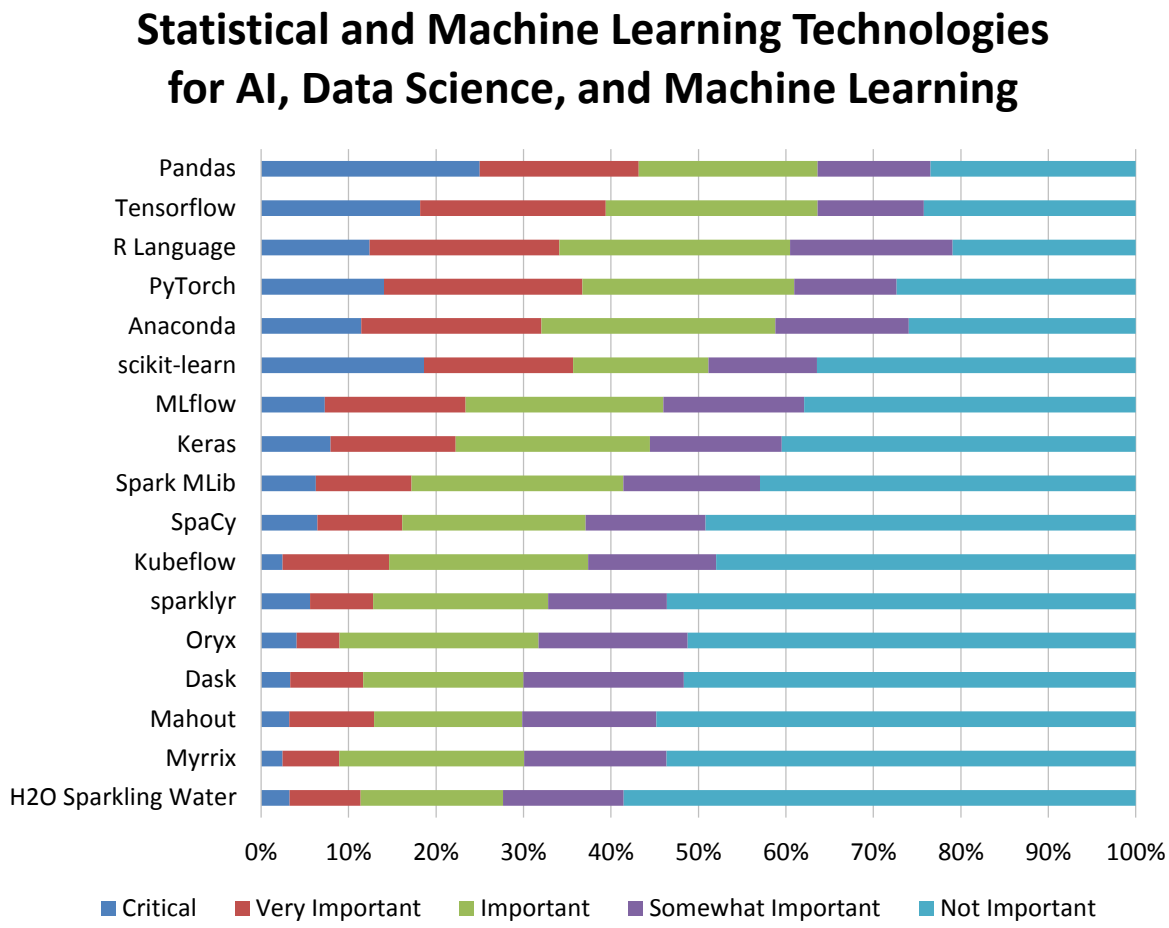


Figure 67 –Statistical and machine learning technologies for AI, data science, and machine learning

Data Sources for AI, Data Science, and Machine Learning

We asked organizations to gauge their interest in 33 (six new added since 2022) data sources for AI, data science, and machine learning (fig. 68). Within this category, all choices have relevance to respondents, and all are at least *somewhat important* to 40 percent or far more respondents. Top picks *Snowflake, Amazon S3,* and *Postgres* are seen as *critical* to about 20 percent or more respondents. Many other technologies of interest represent diverse methods and data management practices. (Also see industry importance, fig. 80.)

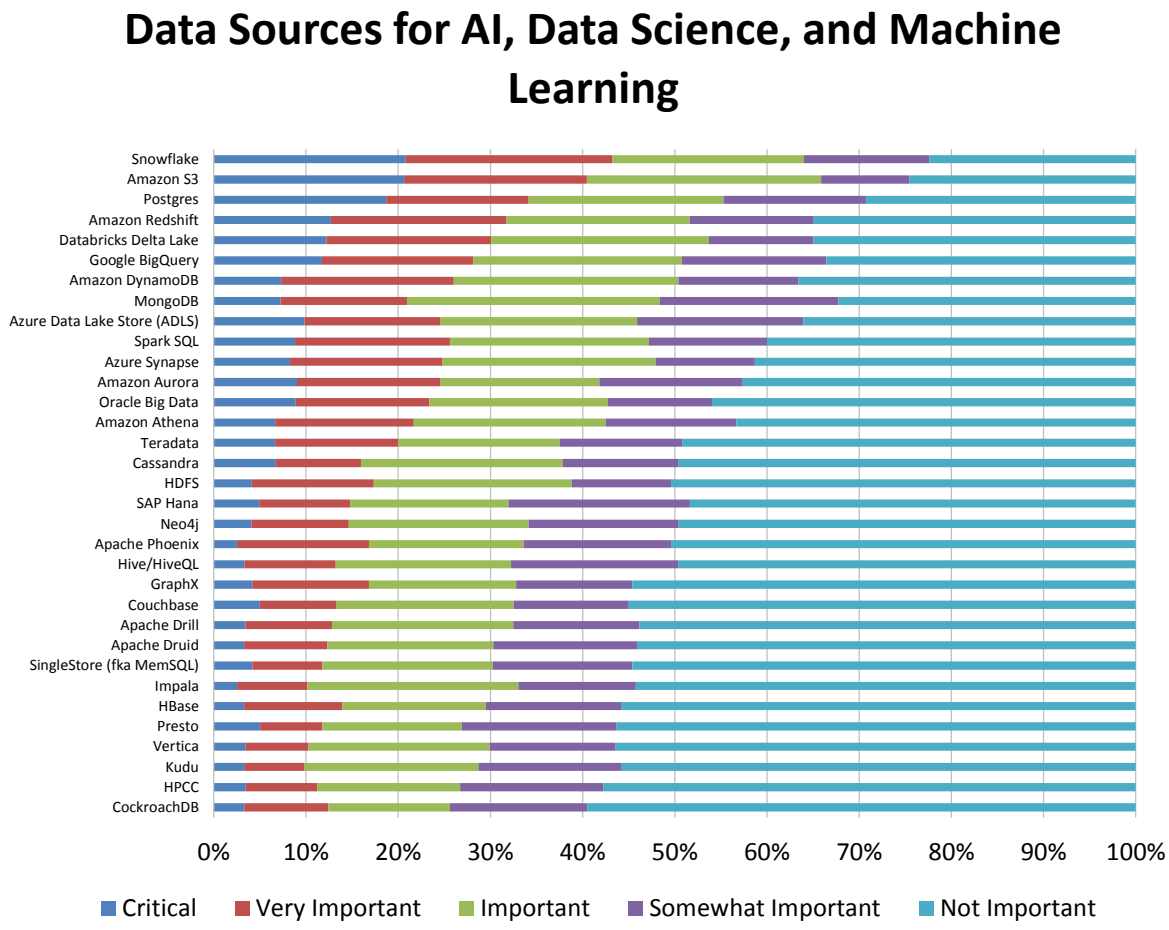


Figure 68 – Data sources for AI, data science, and machine learning

Industry and Vendor Analysis

Industry and Vendor Analysis

Industry Importance of AI, Data Science and Machine Learning

Across 11 years of study, we observed a mostly steady rise in industry sentiment toward AI, data science, and machine learning that is strikingly resurgent in 2024 (fig. 69). Industry sentiment in 2024 stands at weighted-mean 4.9, a rarely seen consensus of effectively *critical* sentiment. This year's pinnacle achievement comes despite a recent flattening and decline during 2022 and 2023, a time at which we speculated whether we had seen "peak" AI, data science, and machine learning enthusiasm attention. This was clearly not the case. We can point to very high enterprise and industry investment in artificial intelligence and generative AI, LLMs, analytical solutions, and the stretched supply of processing chips generating demand in 2024. How long this sentiment rises or is sustained is uncertain and, not unusually, industry excitement is well ahead of user sentiment (fig. 13). As we have observed, mainstreaming data science and machine learning should pave a path to more widely applicable solutions from analytical vendors, hyperscalers, and chip providers.

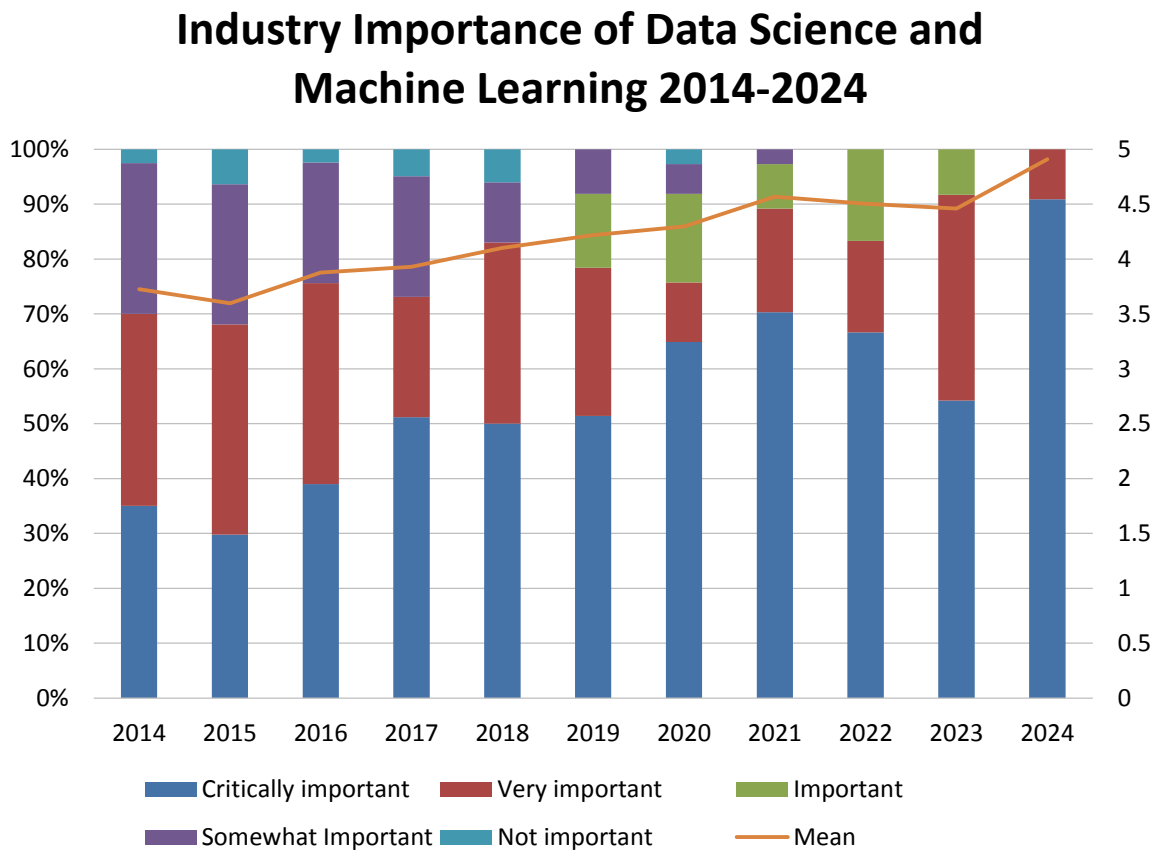


Figure 69 – Industry importance of AI, data science, and machine learning 2014-2024

Industry Support for Analytical Features and Functions

Industry respondents very strongly support multiple analytical features and functions in 2024, led by core *range of regression models*, *textbook statistical functions, outlier detection*, *cross correlation analysis*, and *auto ML* (fig. 70). All four features are at 85-95 percent currently supported and nearing full maturity. Support tapers only somewhat for f*orecasting with model customization, hierarchical clustering*, and *model explain-ability,* all of which will see investment to greater than 90 percent support in 12 months. Somewhat remarkably, all but one of 20 analytical features sampled in 2024 are at least 60 percent currently supported, and all of these are expected to be at least 75 percent and up to 100 percent supported in future time frames. Vendor support and investment aligns well with, and far ahead of, user feature requirements this year (fig. 39).

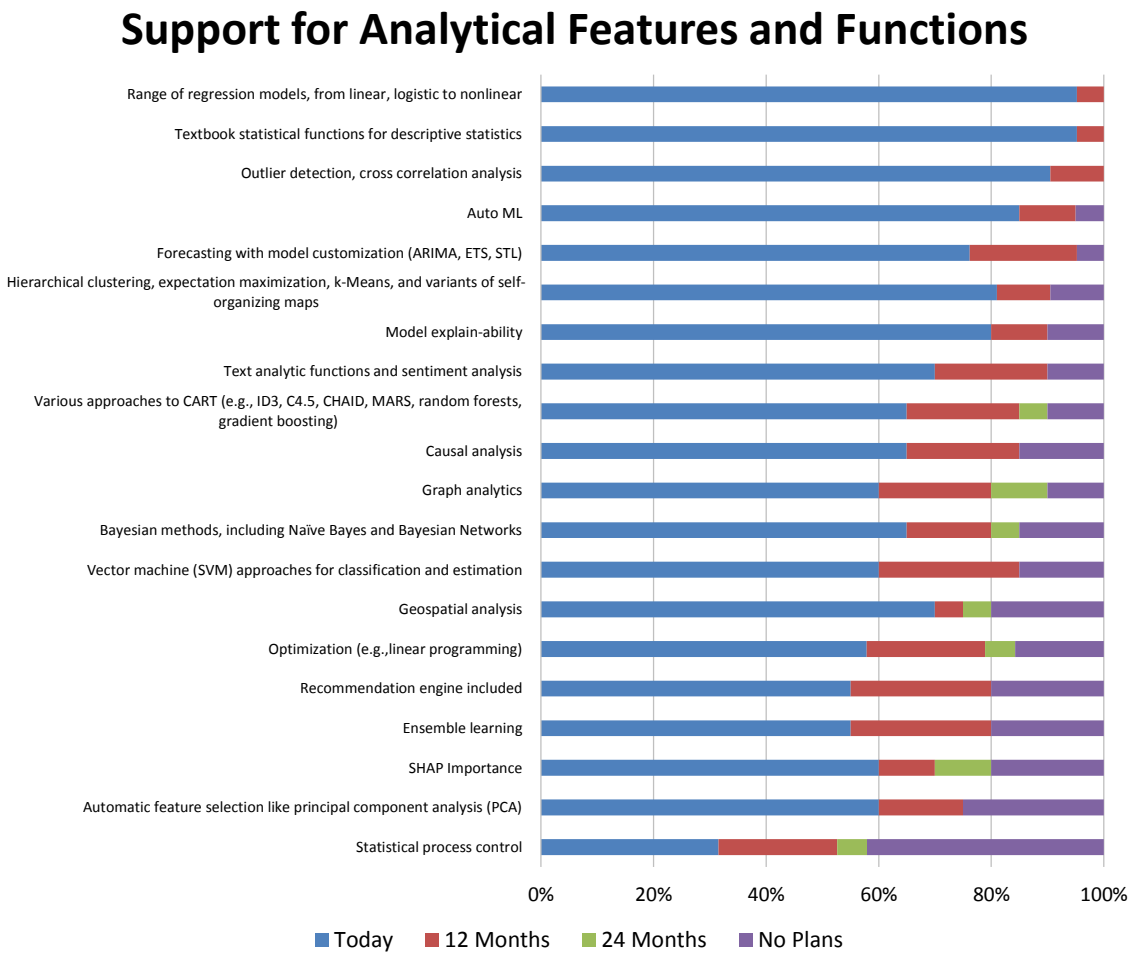


Figure 70 – Industry support for analytical features and functions

Viewed across 10 years of data gathering, industry analytical feature and function support gradual increasing momentum that jumped sharply in 2022 and has extended into 2023 and 2024 (fig. 71). The top five features (*textbook statistical functions, range*

of regression models, outlier detection, auto ML, and *hierarchical clustering),* are all supported by 85-95 percent of our industry sample, and all but one feature is supported by half or more vendors in our sample. Between 2015 and 2021, the industry reported an uncertain pattern of increases and decreases in investment support for analytical features and functions. This year, some feature support falls below peak 2022 or 2023 levels, but we also observe new highs in areas including *textbook statistical functions, range of regression models, auto-ML, model explain-ability,* and some lower-ranked features and functions. Maturing momentum in analytical features and functions more than supports the history of user demand (fig. 40).

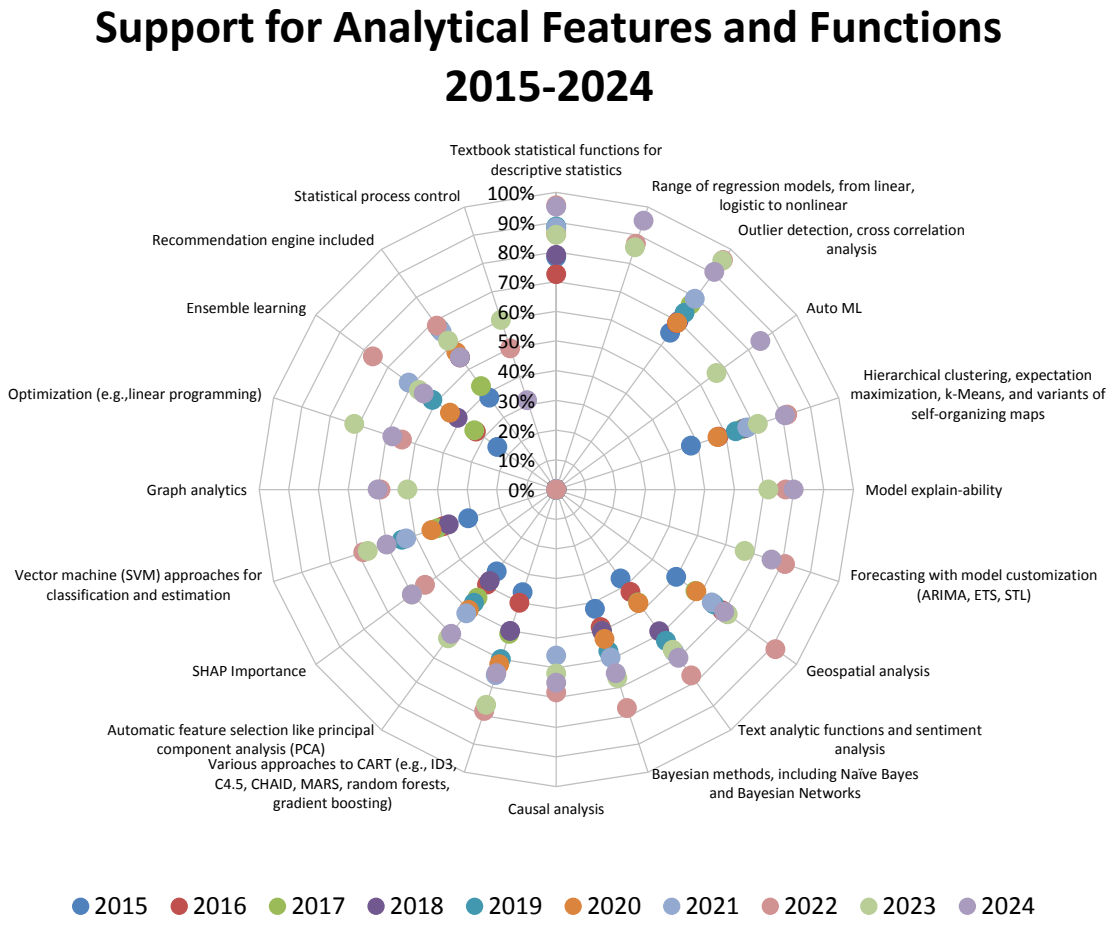


Figure 71 – Support for analytical features and functions 2015-2024

Industry Support for Neural Networks

We asked the vendor community to describe their support for 11 types of neural networks (four added since 2022), and find varying levels of support in 2024 and some

future investment planned (fig. 72). This year, current support is highest for *artificial neural networks* (63 percent) and *transformer network* (58 percent). A second tier or *long short-term memory, convolutional neural networks*, and *generative adversarial networks is currently supported by* 42-47 percent of our industry sample. Twelve-month investment is predicted to be strongest for *convolutional neural networks* (16 percent), *variational autoencoder* (13 percent), *radial basis* (13 percent), *recurrent* (12 percent), and *recursive* (12 percent) neural networks.(Also see user importance, fig. 57.)

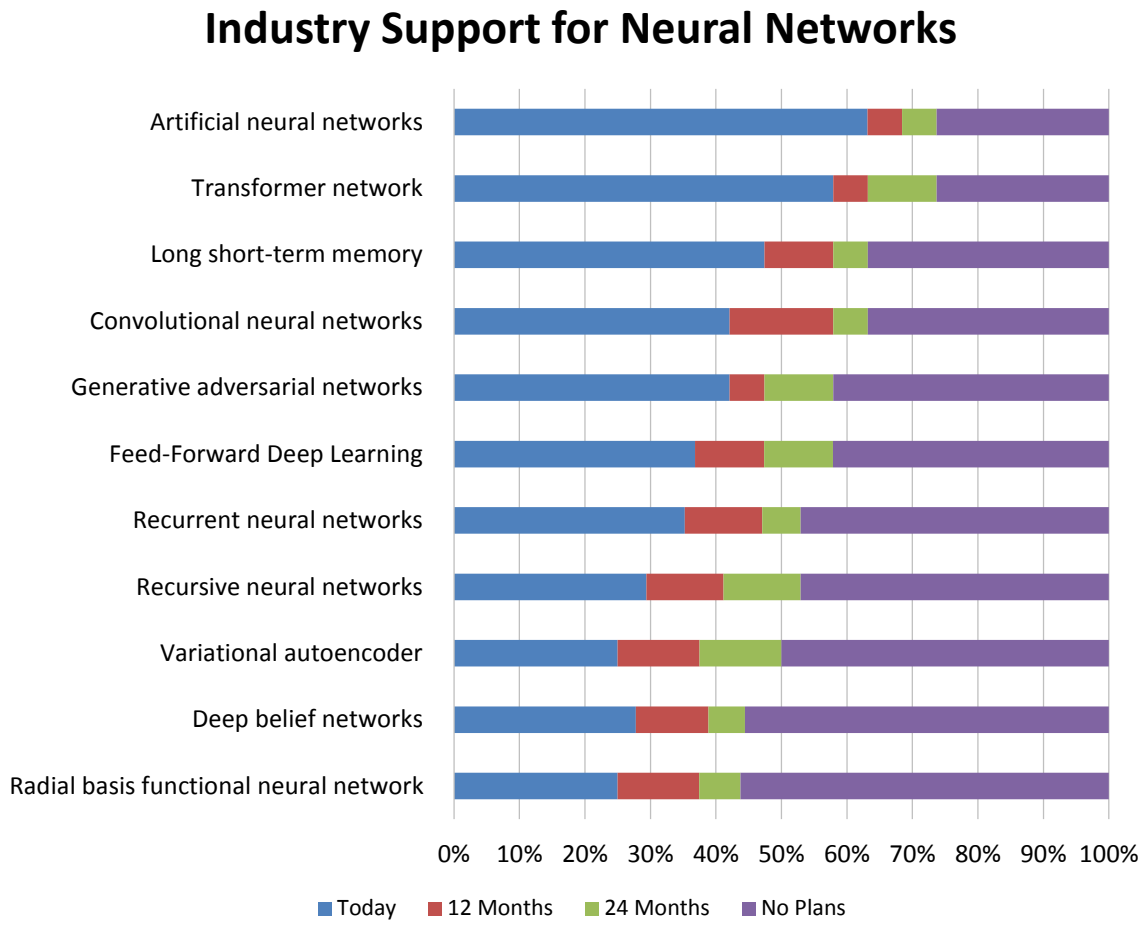


Figure 72 – Industry support for neural networks

Industry Support for Data Preparation

Data preparation, a core step in the data science and machine learning feature generation process, has near complete and effectively mature support today among software vendors in our 2024 industry survey (fig. 73). Current support for the top nine features (*complex filtering*, *cleansing and enrichment*, *formula scripting, data flows for multi-step transformations*, binning, *set operations, support for cutting merging and replacing, support for data type conversions*, and *detection of duplicates or outliers*) is at 95-100 percent. The top 14 (of 16) features all have at nearly 80 percent or far more current support, and all have 60 percent or more with some future investment expected. Industry support at current levels is ahead of user feature requirements (fig. 45).

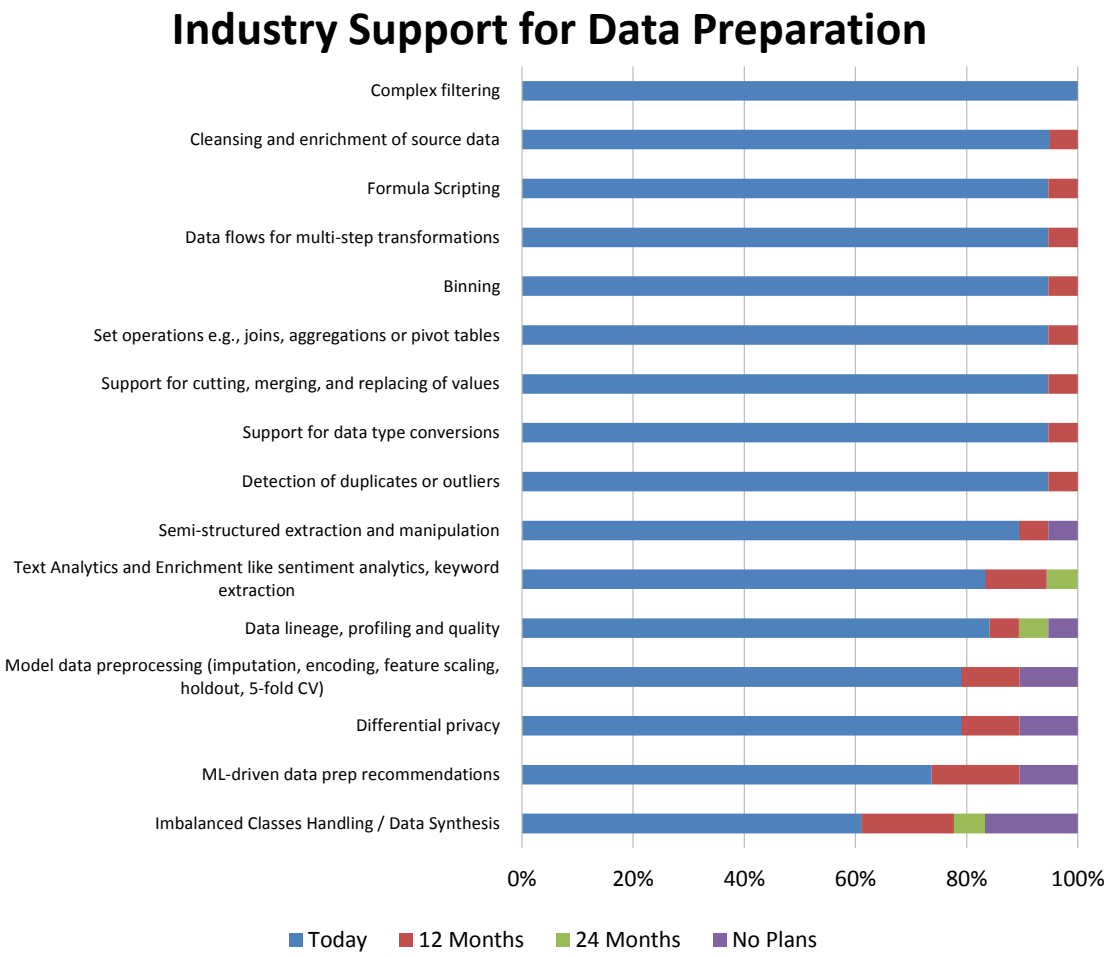


Figure 73 – Industry support for data preparation

Industry Support for Tool Usability

AI, data science, and machine learning tool usability features have very strong and broad industry support in 2024 (fig. 74). Ten of 18 sampled features, led by *automatic creation of models from data* and *access to advanced analytics* are currently supported by 84-95 percent of our industry sample. All features are currently supported by at least two-thirds (67 percent) of our industry sample, and vendors further expect future investments to raise support near or above 90 percent for 14 leading features. Again, industry support well exceeds user criticality scores for every usability feature (fig. 48).

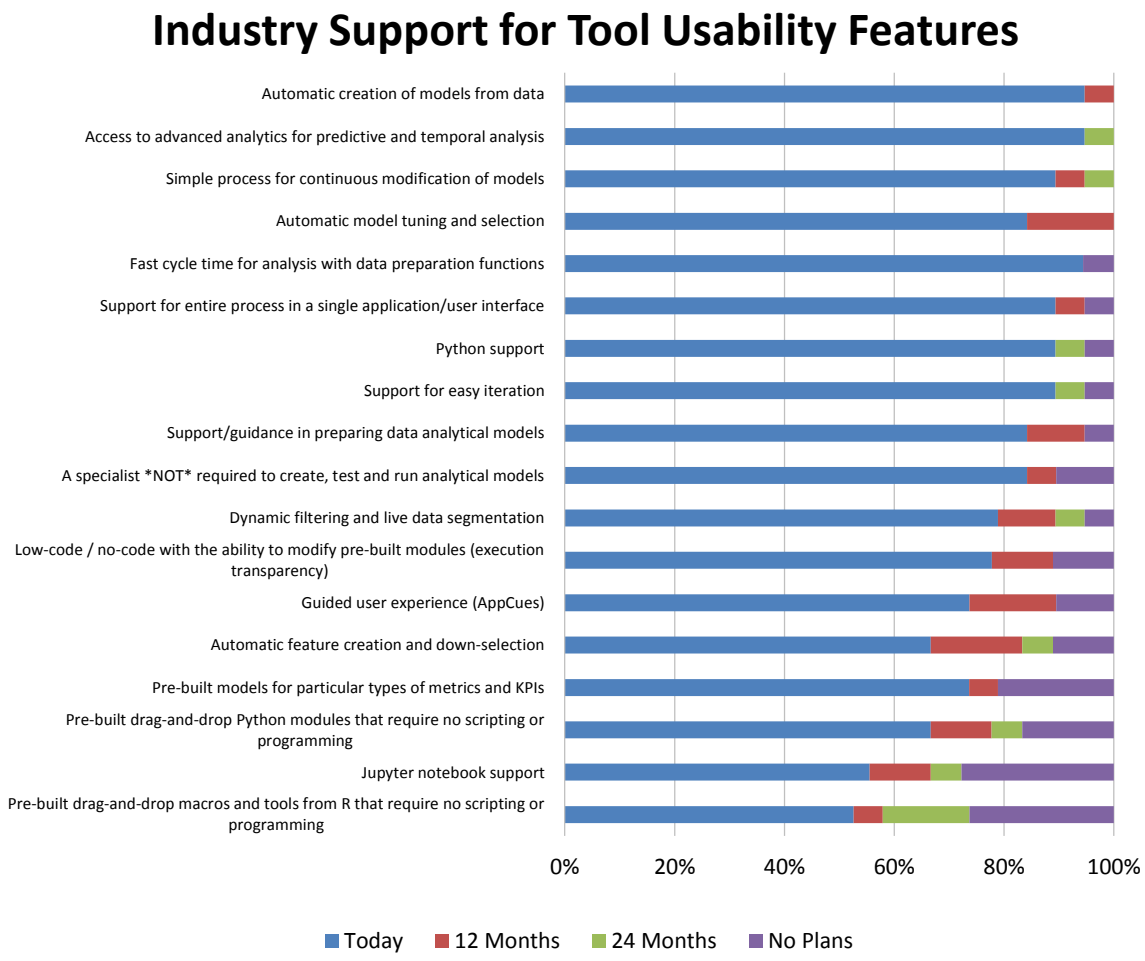


Figure 74 – Industry support for tool usability features

Industry Support for Scalability

Scalability of AI, data science, and machine learning can involve multiple different technologies and services to address high data volumes, large numbers of users, data variety, or analytic throughput. In our 2024 study, current industry support is 100 percent for in-memory analytics, 94 percent for horizontal scaling, 90 percent for *multi-tenant cloud services*, and 90 percent for *in-database analytics* (fig. 75). All four features (along with vertical scaling), are top-five scalability requirements for users (fig. 53). Other features supported by half or more majority of industry respondents include *horizontal scaling*, *Spark support*, *optimized for MPP*, *in-Hadoop analytics*, and *hybrid / cloud bursting*.

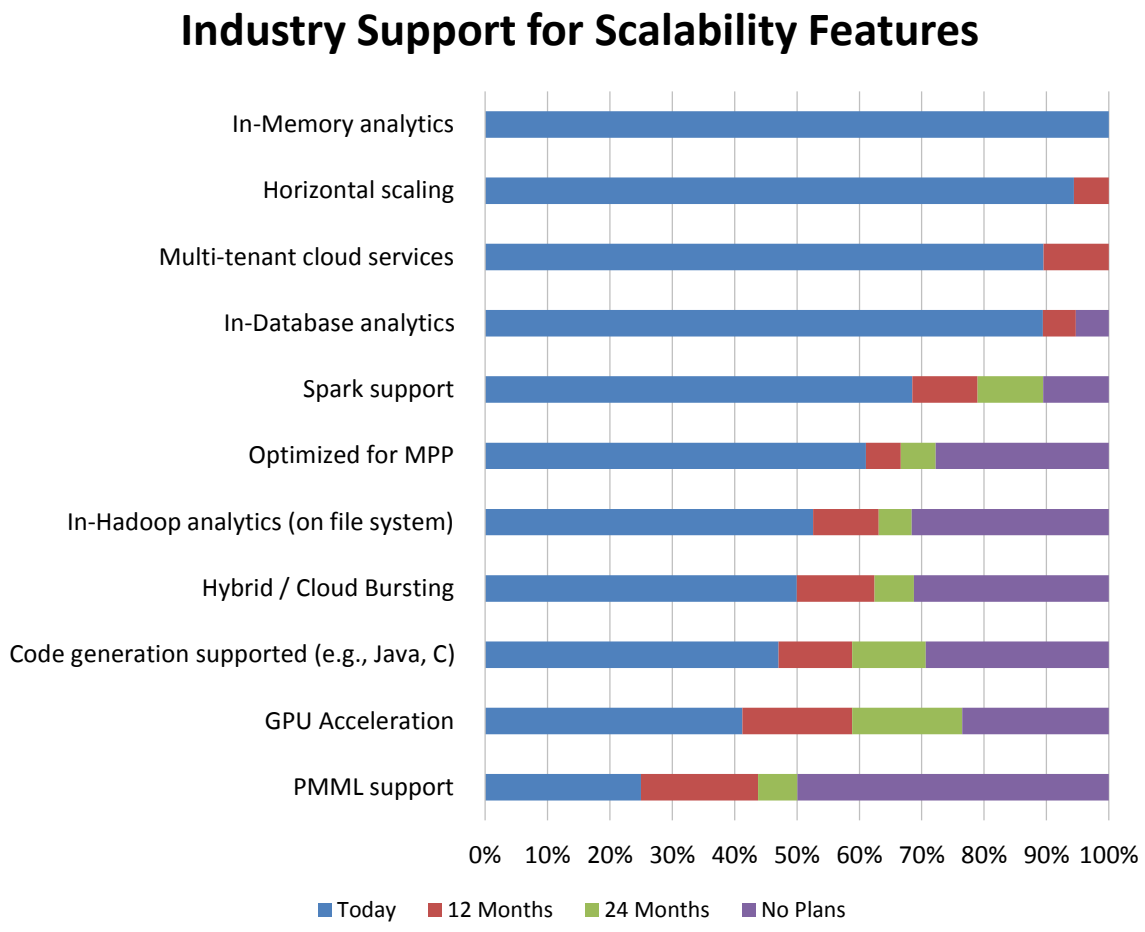


Figure 75 – Industry support for scalability features

Industry Sources of AI, Data Science, and Machine Learning Capabilities

In 2024, industry respondents most support single platforms but signal openness by supporting analytical capabilities via single or multiple products with a homogeneity of third party / open source (versus all proprietary) features. This year, the largest percentage of industry respondents (50 percent) best describe their analytical capabilities as *single product required*, *some features sourced from third party / open source* (50 percent) (fig. 77). Half as many offer multiple products with third party or open source features, and far fewer offer platforms or suites with all-proprietary features.

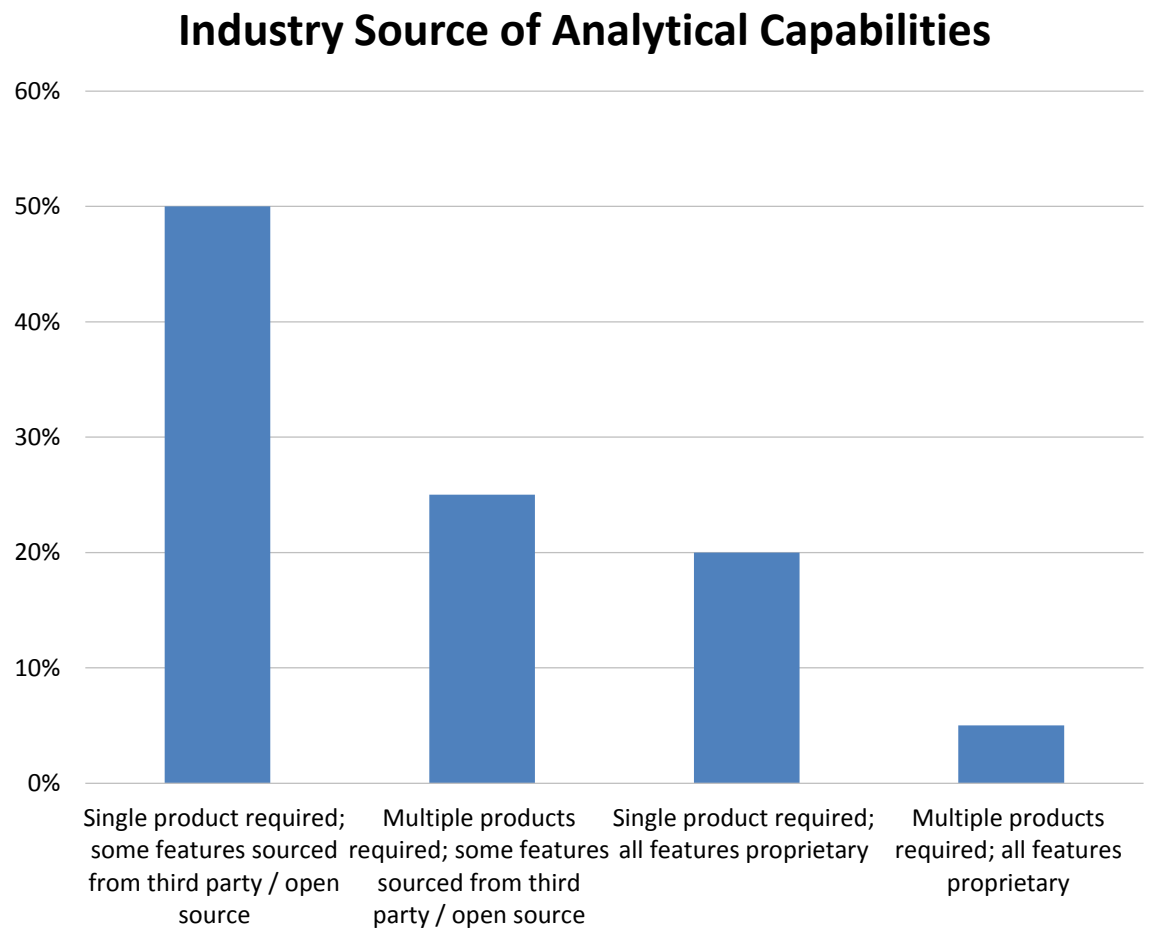


Figure 76 – Industry source of analytical capabilities

Open-Source Support

Industry respondents historically place a high level of importance on open-source technology, a sentiment that once again carries strongly into 2024 (fig. 77). This finding, which relates to the broad developer community, can also be used to compare the industry importance of open source to user sentiment toward customizable versus pre-built features (fig. 66) for AI, data science, and machine learning technologies (fig. 67). In our current-year sentiment snapshot, we observe ongoing peak sentiment in *critically important* industry sentiment (50 percent) and sustained *very important* and *important* sentiment that collectively accounts for 100 percent of industry sentiment in 2024.

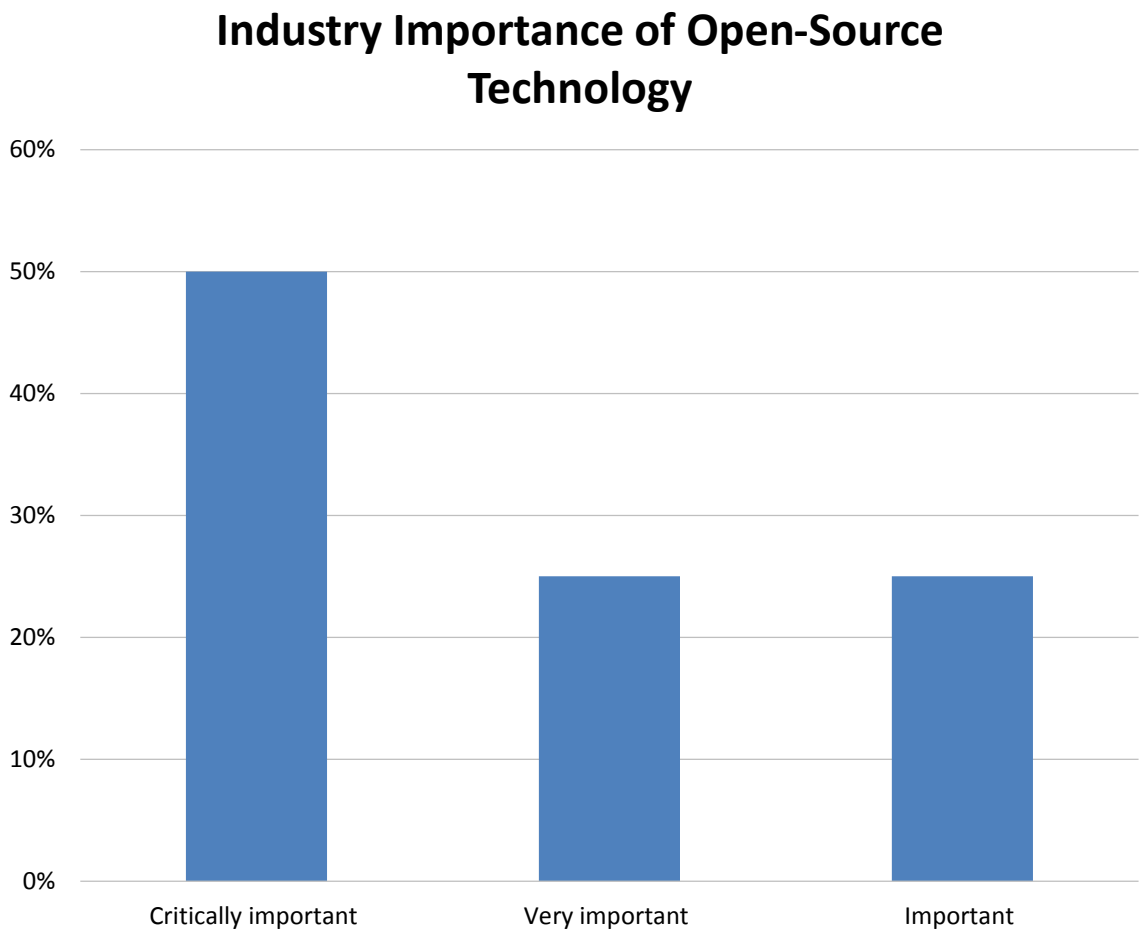


Figure 77 – Industry importance of open-source technology

Industry Support for Open-Source Infrastructure Technologies

Industry support for open-source infrastructure technologies in 2024 is broad and increasing year over year but remains selective, arguably immature, and in some cases limited (fig. 78). This year, we observe that *Databricks* and *Spark* are most supported today at 71 percent and 72 percent, respectively. The next most currently supported open-source infrastructure technologies are *Hugging Face, Elasticsearch, Kafka*, and *Atlas*, all of which have 60 percent or greater industry support. No other technologies reach the level of 50 percent support in 2024, though vendors predict future investment, mostly in a 24-month time frame, that will raise numbers, particularly for *Pinecone vector database* and *Amazon Kinesis*. Still, our vendor community has *no plans* as of now to raise eight of 16 infrastructure technologies to greater than 50 percent support over any time frame. (Also see user importance, fig. 66.)

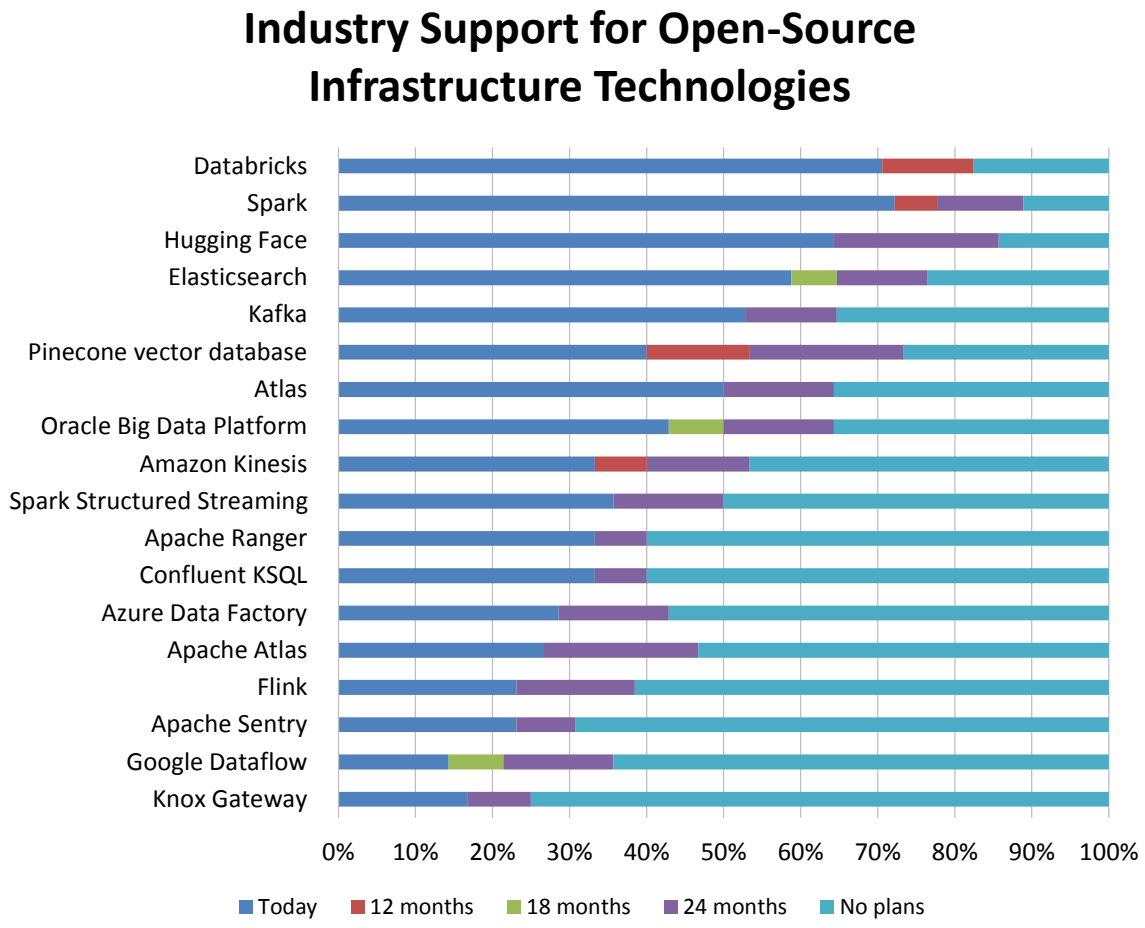


Figure 78 – Industry support for open-source infrastructure technologies

Industry Support for Open-Source Analytical and Machine Learning Technologies

The vendor/provider industry in 2024 is selectively invested in support for open-source analytical and machine learning technologies, with boundaries and mostly minor plans for more future investment (fig. 79). Four technologies have 75-82 percent current industry support, led by *R language*, *Pandas*, *Tensorflow*, and *PyTorch.* A second tier consisting of *scikit-learn* (71 percent), *SpaCy* (67 percent), and *Anaconda* (65 percent) are the next most supported in 2024. New 12- and 24-month investment plans are limited and strongest for *H2O Sparking Water* (19 percent), *Spark MLib* (17 percent), *sparklyr* (13 percent), and *MLflow* (12 percent). Current support and future plans appear to align with user requirements (fig. 67).

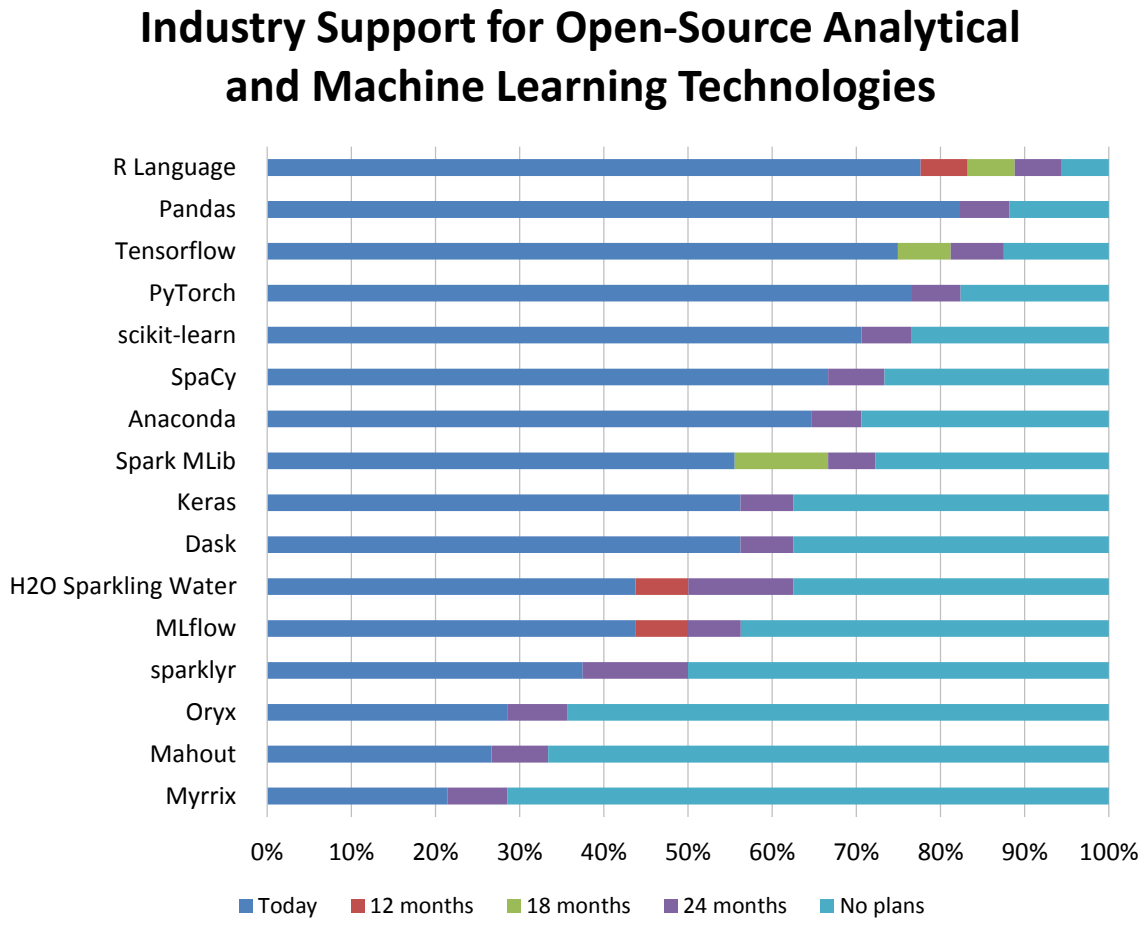


Figure 79 – Industry support for open-source analytics and machine learning technologies

Industry Support for Analytical Data Sources

Our industry sample supports a broad range of analytical data sources in 2024, with current support highest for *Mongo DB* (94 percent), *SAP Hana* (94 percent), *Snowflake* (90 percent), *Amazon S3* (89 percent), and *Amazon Redshift* (89 percent) (fig. 80). All but six data sources are supported by more than half of our industry sample, and some lesser current sources are expected to gather significant investment in future time frames. (Also see user importance, fig. 68.)

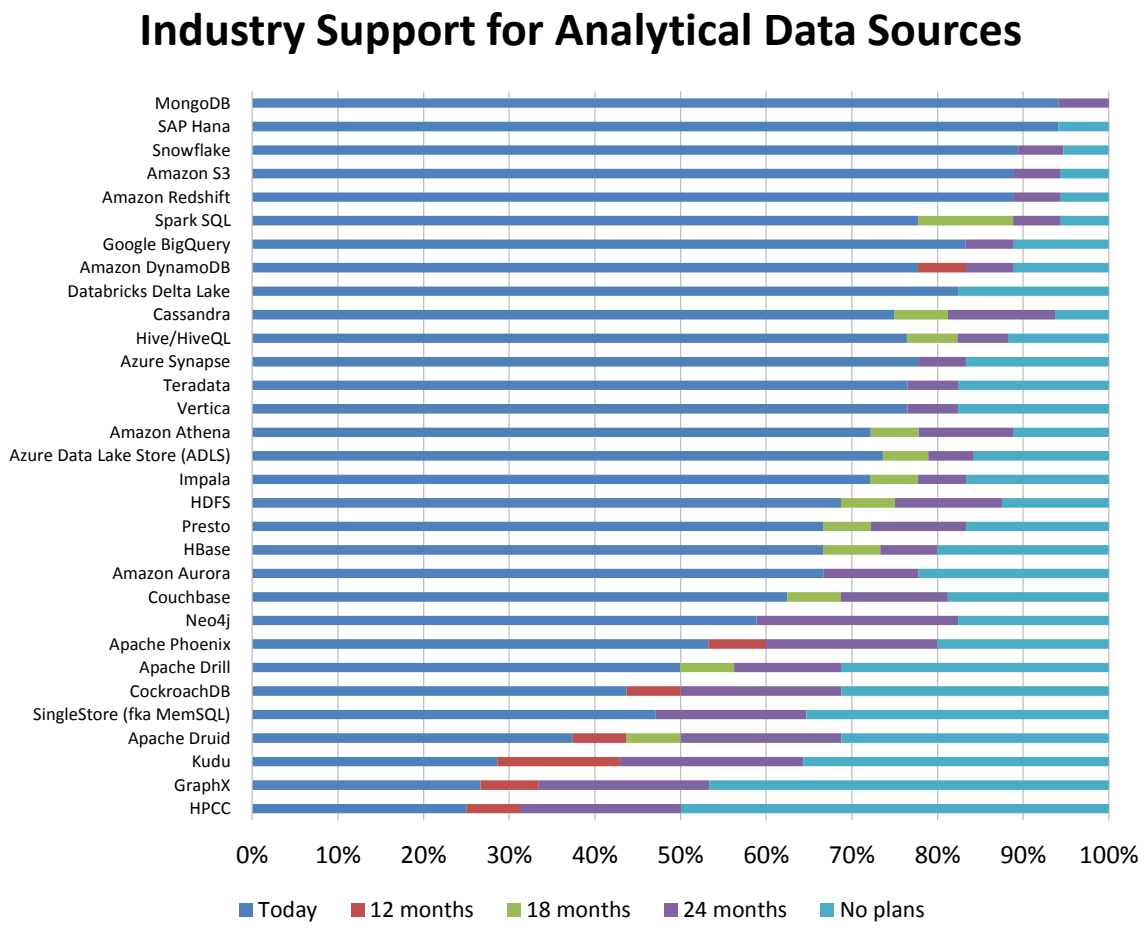


Figure 80 – Industry support for analytical data sources

AI, Data Science, and Machine Learning Vendor Ratings

In rating the vendors, we consider analytical features and functions, neural networks, data preparation, usability, ModelOps, scalability, open source support, and access to data sources (fig. 81). It is important to scrutinize all rating categories and match vendor strengths to use cases and requirements.

Top-ranked vendors include Domino Data Lab (1st), Palantir (1st), Altair (2nd), Dataiku (3rd), Google (4th), and SAS (5th).

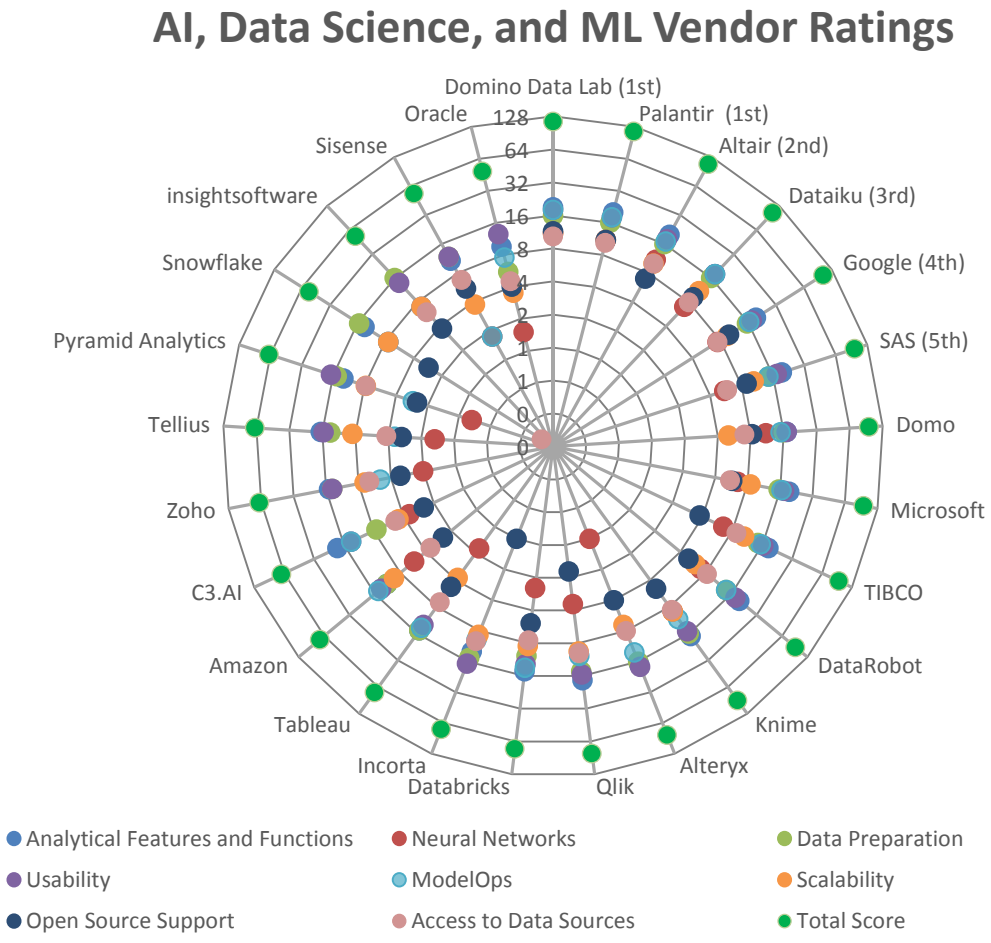


Figure 81 – AI, data science, and machine learning vendor ratings

*A logarithmic scale is used for the scoring chart to address skewness towards larger values

Other Dresner Advisory Services Research Reports

- Wisdom of Crowds® "Flagship" Business Intelligence Market Study
- Active Data Architecture™
- Analytical Data Infrastructure
- Analytical Platforms
- Cloud Computing and Business Intelligence
- Collective Insights®
- Data Catalog
- Data Engineering
- Data and Analytics Governance
- Embedded Business Intelligence
- Enterprise Performance Management
- ESG Reporting
- Financial Consolidations, Close Management, and Reporting
- Generative AI
- Guided Analytics®
- Master Data Management
- ModelOps
- Sales Performance Management
- Self-Service Business Intelligence
- Small and Mid-Sized Enterprise Business Intelligence
- Small and Mid-Sized Enterprise Performance Management
- Supply Chain Planning and Analysis
- Workforce Planning and Analysis

Appendix: AI, Data Science and Machine Learning Survey Instrument

Please enter your contact information below

First Name*: _____

Last Name*: _____

Title: _____

Company Name*: _____

Street Address: _____

City: _____

State: _____

Zip: _____

Country: _____

Email Address*: _____

Phone Number: _____

URL: _____

What major geography do you reside in?*

( ) North America

( ) Europe, Middle East and Africa

( ) Latin America

( ) Asia Pacific

Please identify your primary industry*

( ) Advertising

( ) Aerospace

( ) Agriculture

( ) Apparel & accessories

( ) Automotive

( ) Aviation

( ) Biotechnology

( ) Broadcasting

( ) Business services

( ) Chemical

( ) Construction

( ) Consulting

( ) Consumer products

( ) Defense

( ) Distribution & logistics

( ) Education (Higher Ed)

( ) Education (K-12)

( ) Energy

( ) Entertainment and leisure

( ) Executive search

( ) Federal government

( ) Financial services

( ) Food, beverage and tobacco

( ) Healthcare

( ) Hospitality

( ) Insurance

( ) Legal

( ) Manufacturing

( ) Mining

( ) Motion picture and video

( ) Not for profit

( ) Pharmaceuticals

( ) Publishing

( ) Real estate

( ) Retail & wholesale

( ) Sports

( ) State and local government

( ) Technology

( ) Telecommunications

( ) Transportation

( ) Utilities

( ) Other - Write In: _____

How many employees does your company employ worldwide?

( ) 1 - 100

( ) 101 - 1,000

( ) 1,001 - 2,000

( ) 2,001 - 5,000

( ) 5,001 - 10,000

( ) More than 10,000

What function do you report into?*

( ) Business Intelligence Competency Center

( ) Executive management

( ) Faculty (Education)

( ) Finance

( ) Human resources

( ) Information Technology (IT)

( ) Manufacturing

( ) Marketing

( ) Medical staff (Healthcare)

( ) Operations

( ) Research and development (R&D)

( ) Sales

( ) Strategic planning function

( ) Supply chain

( ) Other - Write In: _____

Does your organization use or intend to use data science and machine learning?

( ) Yes, we use data science and machine learning today in production

( ) Data science and machine learning is being used in very limited ways or as proof of concept

( ) We are currently evaluating data science and machine learning software

( ) No, we have no plans to use data science and machine learning at all

( ) We may use data science and machine learning in the future

What are your plans for AI, data science and machine learning in the future?

( ) Will Adopt this Year

( ) Will Adopt Next Year

( ) Will Adopt Beyond Next Year

How long have data science and machine learning been in use in your organization?

( ) Less than 1 year

( ) 1-2 years

( ) 2-3 years

( ) 3-5 years

( ) More than 5 years

Which kinds of users use (or will use) data science and machine learning within your organization?

	Constantly	**Often**	**Occasionally**	**Rarely**	**Never**
BI expert	( )	( )	( )	( )	( )

Business Analyst	( )	( )	( )	( )	( )
"Citizen" Data Scientist	( )	( )	( )	( )	( )
Executive	( )	( )	( )	( )	( )
Financial Analyst	( )	( )	( )	( )	( )
IT Staff	( )	( )	( )	( )	( )
Marketing Analyst	( )	( )	( )	( )	( )
Statistician / Data Scientist	( )	( )	( )	( )	( )
Third-Party Consultant	( )	( )	( )	( )	( )

How is data science and machine learning being used in your organization?

	Today	**12 Months**	**24 Months**	**No Plans**	**Don't know**
Churn Prevention	( )	( )	( )	( )	( )
Cognitive Robotic Process Automation	( )	( )	( )	( )	( )
Customer Lifetime	( )	( )	( )	( )	( )

Value					
Customer Segmentation	( )	( )	( )	( )	( )
Demand Forecasting	( )	( )	( )	( )	( )
Fraud Detection	( )	( )	( )	( )	( )
Next Best Action	( )	( )	( )	( )	( )
Predictive Maintenance	( )	( )	( )	( )	( )
Price Optimization	( )	( )	( )	( )	( )
Product Propensity	( )	( )	( )	( )	( )
Quality Assurance	( )	( )	( )	( )	( )
Risk Management	( )	( )	( )	( )	( )
Up and Cross-Selling	( )	( )	( )	( )	( )

Analytical Features: Which of the following features are important for AI, data science and machine learning?

	Critical	**Very Important**	**Important**	**Somewhat Important**	**Not Important**

Automatic feature selection like principal component analysis (PCA)	( )	( )	( )	( )	( )
Text analytic functions and sentiment analysis	( )	( )	( )	( )	( )
Various approaches to CART (e.g., ID3, C4.5, CHAID, MARS, random forests, gradient boosting)	( )	( )	( )	( )	( )
Vector machine (SVM) approaches for classification and estimation	( )	( )	( )	( )	( )
Neural networks supported	( )	( )	( )	( )	( )
Geospatial analysis	( )	( )	( )	( )	( )
Range of regression models, from linear, logistic to nonlinear	( )	( )	( )	( )	( )
Recommendation engine included	( )	( )	( )	( )	( )
Hierarchical clustering, expectation	( )	( )	( )	( )	( )

maximization, k-Means, and variants of self-organizing maps					
Textbook statistical functions for descriptive statistics	( )	( )	( )	( )	( )
Bayesian methods, including Naïve Bayes and Bayesian Networks	( )	( )	( )	( )	( )
Ensemble learning	( )	( )	( )	( )	( )
Video analysis	( )	( )	( )	( )	( )
Model management and governance	( )	( )	( )	( )	( )
Auto ML	( )	( )	( )	( )	( )
Model explain-ability	( )	( )	( )	( )	( )
Graph analytics	( )	( )	( )	( )	( )
Forecasting with model customization (ARIMA, ETS, STL)	( )	( )	( )	( )	( )
Optimization (e.g.,linear programming)	( )	( )	( )	( )	( )

Outlier detection	( )	( )	( )	( )	( )
Cross correlation analysis	( )	( )	( )	( )	( )
SHAP Importance	( )	( )	( )	( )	( )
Statistical process control	( )	( )	( )	( )	( )

Which types of neural networks are most important to your organization?

	Critical	**Very Important**	**Important**	**Somewhat Important**	**Not Important**	**Don't Know**
Artificial neural network	( )	( )	( )	( )	( )	( )
Convolutional neural networks	( )	( )	( )	( )	( )	( )
Long short-term memory	( )	( )	( )	( )	( )	( )
Recursive neural networks	( )	( )	( )	( )	( )	( )
Deep learning neural networks	( )	( )	( )	( )	( )	( )
Adversarial	( )	( )	( )	( )	( )	( )

neural networks						

Data Preparation: For AI, data science and machine learning, which data preparation capabilities are important?

	Critical	**Very Important**	**Important**	**Somewhat Important**	**Not Important**
Support for Cutting, Merging, and Replacing Values	( )	( )	( )	( )	( )
Complex Filtering	( )	( )	( )	( )	( )
Support for Data Type Conversions	( )	( )	( )	( )	( )
Cleansing and Enriching Source Data	( )	( )	( )	( )	( )
Detecting Duplicates or Outliers	( )	( )	( )	( )	( )
Set Operations (e.g., joins, aggregations or pivot tables)	( )	( )	( )	( )	( )
Data Flows for Multi-step Transformations	( )	( )	( )	( )	( )
Data Lineage, Profiling and Quality	( )	( )	( )	( )	( )

Differential Privacy	( )	( )	( )	( )	( )
Formula Scripting	( )	( )	( )	( )	( )
Imbalanced Classes Handling / Data Synthesis	( )	( )	( )	( )	( )
ML-Driven Data Prep Recommendations	( )	( )	( )	( )	( )
Model Data Preprocessing (imputation, encoding, feature scaling, holdout, 5-fold CV)	( )	( )	( )	( )	( )
Semi-structured Extraction and Manipulation	( )	( )	( )	( )	( )
Text Analytics and Enrichment (e.g., sentiment analytics, key word extraction)	( )	( )	( )	( )	( )
Binning	( )	( )	( )	( )	( )

Usability: Which usability features are important for AI, data science and machine learning?

	Critical	**Very Important**	**Important**	**Somewhat Important**	**Not Important**
A specialist NOT	( )	( )	( )	( )	( )

required to create analytical models, test and run them					
Support/guidance in preparing data analytical models	( )	( )	( )	( )	( )
Automatic creation of models from data	( )	( )	( )	( )	( )
Fast cycle time for analysis with data preparation functions	( )	( )	( )	( )	( )
Access to advanced analytics for predictive and temporal analysis	( )	( )	( )	( )	( )
Support for easy iteration	( )	( )	( )	( )	( )
Simple process for continuous modification of models	( )	( )	( )	( )	( )
Pre-built drag- and -drop macros and tools from R that require no scripting or programming	( )	( )	( )	( )	( )
Support for entire process in a single application/user	( )	( )	( )	( )	( )

interface					
Automatic feature creation and down-selection	( )	( )	( )	( )	( )
Automatic model tuning and selection	( )	( )	( )	( )	( )
Dynamic filtering and live data segmentation	( )	( )	( )	( )	( )
Guided user experience (AppCues)	( )	( )	( )	( )	( )
Low-code / no-code with the ability to modify pre-built modules (execution transparency)	( )	( )	( )	( )	( )
Pre-built drag-and-drop Python modules that require no scripting or programming	( )	( )	( )	( )	( )
Pre-built models for particular types of metrics and KPIs	( )	( )	( )	( )	( )

Scalability: Which scalability features are important for AI, data science and machine learning?

	Critical	Very Important	Important	Somewhat Important	Not Important
In-Memory Analytics	( )	( )	( )	( )	( )
In-Database Analytics	( )	( )	( )	( )	( )
In-Hadoop Analytics (on file system)	( )	( )	( )	( )	( )
Optimized for MPP Architecture	( )	( )	( )	( )	( )
Multi-tenant Cloud Services	( )	( )	( )	( )	( )
PMML Support	( )	( )	( )	( )	( )
Code Generation Supported (e.g., Java, C)	( )	( )	( )	( )	( )
GPU Acceleration	( )	( )	( )	( )	( )
Spark Support	( )	( )	( )	( )	( )
Hybrid / Cloud Bursting	( )	( )	( )	( )	( )
Horizontal	( )	( )	( )	( )	( )

Scaling					

What is the importance of big data / open source technologies (e.g., Apache) and architecture as a part of your data science and machine learning strategies?

( ) Critical

( ) Very Important

( ) Important

( ) Somewhat Important

( ) Not Important

Which open source / big data infrastructure features are important for AI, data science and machine learning?

	Critical	**Very Important**	**Important**	**Somewhat Important**	**Not Important**
Spark	( )	( )	( )	( )	( )
Atlas	( )	( )	( )	( )	( )
Knox Gateway	( )	( )	( )	( )	( )
Kafka	( )	( )	( )	( )	( )
Confluent KSQL	( )	( )	( )	( )	( )
Flink	( )	( )	( )	( )	( )
Amazon Kinesis	( )	( )	( )	( )	( )
Google	( )	( )	( )	( )	( )

Dataflow					
Apache Ranger	( )	( )	( )	( )	( )
Apache Sentry	( )	( )	( )	( )	( )
Spark Structured Streaming	( )	( )	( )	( )	( )
Apache Atlas	( )	( )	( )	( )	( )
Azure Data Factory	( )	( )	( )	( )	( )
Databricks	( )	( )	( )	( )	( )
Elasticsearch	( )	( )	( )	( )	( )

Which open source / big data deployment technologies are important for AI, data science and machine learning?

	Critical	**Very Important**	**Important**	**Somewhat Important**	**Not Important**
Microservices architecture	( )	( )	( )	( )	( )
Kubernetes	( )	( )	( )	( )	( )
Yarn	( )	( )	( )	( )	( )
Mesos	( )	( )	( )	( )	( )
Docker swarm	( )	( )	( )	( )	( )
Nomad	( )	( )	( )	( )	( )

Consul	( )	( )	( )	( )	( )
Etcd	( )	( )	( )	( )	( )
Zookeeper	( )	( )	( )	( )	( )
Zipkin	( )	( )	( )	( )	( )
Prometheus	( )	( )	( )	( )	( )
Grafana	( )	( )	( )	( )	( )
FluentD	( )	( )	( )	( )	( )
Akka	( )	( )	( )	( )	( )

Which open source / big data statistical and machine learning technologies are important for AI, data science and machine learning?

	Critical	**Very Important**	**Important**	**Somewhat Important**	**Not Important**
Mahout	( )	( )	( )	( )	( )
R Language	( )	( )	( )	( )	( )
Oryx	( )	( )	( )	( )	( )
Myrrix	( )	( )	( )	( )	( )
Spark MLib	( )	( )	( )	( )	( )
scikit-learn	( )	( )	( )	( )	( )
Tensorflow	( )	( )	( )	( )	( )
MLflow	( )	( )	( )	( )	( )

PyTorch	( )	( )	( )	( )	( )
Kubeflow	( )	( )	( )	( )	( )
Pandas	( )	( )	( )	( )	( )
Dask	( )	( )	( )	( )	( )
Anaconda	( )	( )	( )	( )	( )
H2O Sparkling Water	( )	( )	( )	( )	( )
Keras	( )	( )	( )	( )	( )
sparklyr	( )	( )	( )	( )	( )
SpaCy	( )	( )	( )	( )	( )

Which data sources are important for AI, data science and machine learning?

	Critical	**Very Important**	**Important**	**Somewhat Important**	**Not Important**
Google BigQuery	( )	( )	( )	( )	( )
HBase	( )	( )	( )	( )	( )
HDFS	( )	( )	( )	( )	( )
Hive/HiveQL	( )	( )	( )	( )	( )
Impala	( )	( )	( )	( )	( )
MongoDB	( )	( )	( )	( )	( )
Amazon	( )	( )	( )	( )	( )

Redshift					
Spark SQL	( )	( )	( )	( )	( )
Couchbase	( )	( )	( )	( )	( )
Cassandra	( )	( )	( )	( )	( )
Amazon S3	( )	( )	( )	( )	( )
Neo4j	( )	( )	( )	( )	( )
Presto	( )	( )	( )	( )	( )
Kudu	( )	( )	( )	( )	( )
Amazon DynamoDB	( )	( )	( )	( )	( )
Azure Data Lake Store (ADLS)	( )	( )	( )	( )	( )
Apache Drill	( )	( )	( )	( )	( )
SingleStore (fka MemSQL)	( )	( )	( )	( )	( )
Snowflake	( )	( )	( )	( )	( )
SAP Hana	( )	( )	( )	( )	( )
Apache Phoenix	( )	( )	( )	( )	( )
Vertica	( )	( )	( )	( )	( )
Teradata	( )	( )	( )	( )	( )
CockroachDB	( )	( )	( )	( )	( )

Apache Druid	( )	( )	( )	( )	( )
HPCC	( )	( )	( )	( )	( )
GraphX	( )	( )	( )	( )	( )
Oracle Big Data	( )	( )	( )	( )	( )
Amazon Athena	( )	( )	( )	( )	( )
Amazon Aurora	( )	( )	( )	( )	( )
Azure Synapse	( )	( )	( )	( )	( )
Databricks Delta Lake	( )	( )	( )	( )	( )