



February 26, 2026

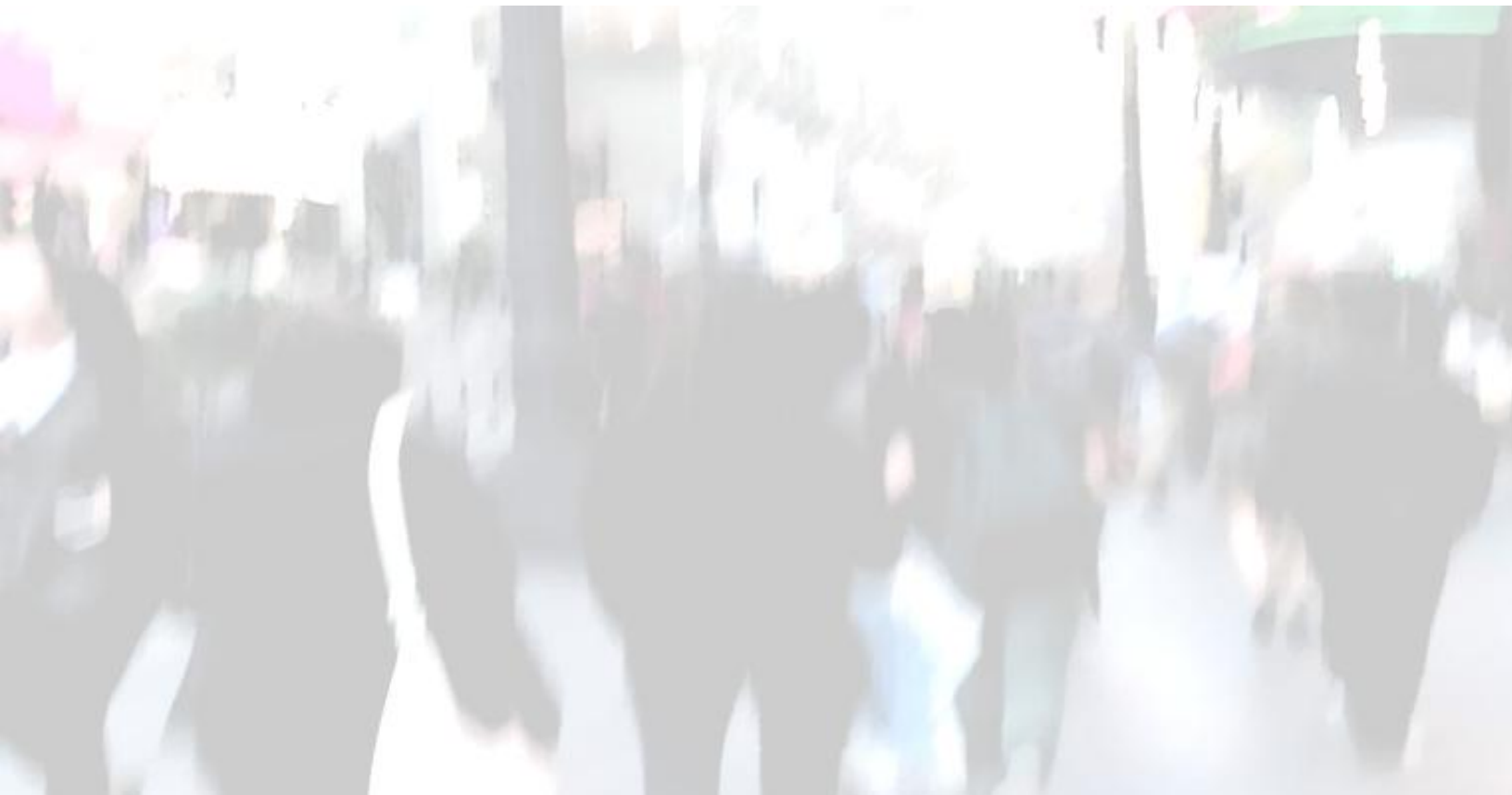
Dresner Advisory Services, LLC

2026 Edition

Data Engineering Market Study

Wisdom of Crowds® Series

Licensed to Cloud Software Group



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 reflect 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 its content.

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

Definitions

Business Intelligence Defined

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

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

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

Data Engineering Defined

Data engineering is a component of analytical data infrastructure (ADI) which includes the best practices and technology capabilities to develop engineered data workflows and pipelines to and between operational and analytic data management infrastructures. Data engineering includes the requirements and priorities for data orchestration, integration, and transformations including advanced analytics in the data engineering pipeline workflow.

For this report, we asked about data engineering development and deployment capabilities used to develop, debug, schedule, secure, govern, and run data workflows for a range of use cases, both analytical and operational in nature. We also sought insight on the role of data engineering within the broader context of analytical data infrastructure.

Analytical Data Infrastructure Defined

An analytical data infrastructure (ADI) is the data workflow and technology ecosystem responsible for the ingestion, persistence, and transformation of data from source systems for downstream consumption by analysts and analytic workflows/use cases and AI agents. Architecturally, it is defined by a range of components (including data engineering), which are available from a variety of vendors, with best-in-class components differing according to analytic data infrastructure architectures and use cases.

Introduction

As Dresner Advisory Services approaches its 19th year in 2026, we present the 12th annual edition of the Data Engineering Market Study. Since our founding in 2007, our research has focused on independent, data-driven insight drawn directly from organizations that evaluate, purchase, and use these solutions.

Data engineering is a foundational component of analytical data infrastructure (ADI), providing the capabilities required to design, build, and operate data workflows and pipelines across operational and analytic environments. It encompasses requirements for data orchestration, integration, and transformation, including support for advanced analytics within the data engineering workflow.

This report examines the development and deployment capabilities used to design, debug, schedule, secure, govern, and run data workflows across analytical and operational use cases. It also evaluates the role of data engineering within the broader ADI ecosystem, which supports the ingestion, persistence, and transformation of data for downstream analytic, application, and AI-driven consumption.

This perspective is intended to support leaders seeking a fact-based understanding of data engineering's role within modern analytical data infrastructure and to inform decisions grounded in peer experience.

We welcome the opportunity to continue the dialogue as you explore the findings in this report.

Best,



Chief Research Officer
Dresner Advisory Services

Contents

Definitions	3
Business Intelligence Defined	3
Data Engineering Defined	3
Analytical Data Infrastructure Defined	3
Introduction	4
Benefits of the Study	7
Consumer Guide	7
Supplier Tool	7
External Awareness	7
Internal Planning	7
About Howard Dresner and Dresner Advisory Services	8
The Dresner Team	9
About Elizabeth Espinoza	9
About Sherry Fairchok	9
About Danielle Guinebertiere	9
About Michelle Whitson-Lorenzi	9
Survey Method and Data Collection	9
Data Quality	9
Executive Summary	11
Study Demographics	14
Geography	14
Functions	15
Vertical Industries	16
Organization Size	17
Analysis and Trends	19
Data Engineering: Key Enabler of Modern Analytic Data Infrastructure	19
Importance of Data Engineering	21
Current Adoption and Future Plans for Data Engineering	29
Use Cases for Data Engineering	34

ADI Capabilities Relevant for Data Engineering.....	39
Data Engineering Features for Processing and Transforming Data.....	43
ADI Management Features Relevant for Data Engineering	46
Data Sources and Targets for Data Engineering	50
ADI Integration Techniques Relevant for Data Engineering.....	55
Deployment of Data Engineering Capabilities.....	59
Licensing of Data Engineering Capabilities.....	63
Sourcing of Data Engineering Capabilities.....	67
Industry and Vendor Analysis.....	72
Industry Support for Data Engineering Capabilities Within ADI.....	73
Industry Support of Sources and Targets.....	74
Industry Support of Interface Types	75
Industry Support for Development and Usability Features	76
Industry Support for Data Processing and Transformation Features	77
Industry Support for Single Product Data Engineering Functionality.....	78
Vendor Ratings	80
Other Dresner Advisory Services Research Reports	81
Appendix: 2026 Data Engineering Survey Instrument.....	82

Benefits of the Study

The 2026 Dresner Advisory Services Data Engineering Market Study provides a wealth of information and analysis, offering value to both consumers and producers of related technology and services.

Consumer Guide

The Dresner Advisory Data Engineering Market Study is an objective source of industry research that helps consumers to understand how their peers leverage and invest in data engineering and related technologies.

From our unique vendor performance measurement system, users glean key insights into software supplier performance, which enables:

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

Supplier Tool

Vendor licensees use the Dresner Advisory Services Data Engineering Market Study in several important ways:

External Awareness

- Build awareness for markets and supplier brands, citing Dresner Advisory Services Data Engineering Market Study market trends and vendor performance.
- Gain lead and demand generation for supplier offerings through association with Dresner Advisory Services Data Engineering Market Study brand, findings, webinars, etc.

Internal Planning

- Refine internal product plans and align with market priorities and realities identified within the Dresner Advisory Services Data Engineering 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 Data Engineering Market Study was conceived, designed, and executed by Dresner Advisory Services, LLC, an independent advisory firm, and Howard Dresner, its president, founder, and chief research officer.

Howard Dresner is one of the foremost thought leaders in business intelligence and performance management, having coined the term “business intelligence” in 1989. He



has published two books on the subject, *The Performance Management Revolution – Business Results through Insight and Action* (John Wiley & Sons, Nov. 2007) and *Profiles in Performance – Business Intelligence Journeys and the Roadmap for Change* (John Wiley & Sons, Nov. 2009). He lectures at forums around the world and is often cited by the business and trade press.

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

Howard has conducted and directed numerous in-depth primary research studies over the past three decades and he and the Dresner team are experts 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
- AI, Data and Analytics Governance
- AI Development Platforms
- Agentic AI Assisted Analytics
- Agentic AI Automation Platforms
- Analytical Data Infrastructure
- Analytical Data products
- ModelOps
- Semantic Layer and Data Virtualization

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

The Dresner Team

About Elizabeth Espinoza

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

About Sherry Fairchok

Sherry is senior editor at Dresner Advisory, ensuring the quality and consistency of all research publications.

About Danielle Guinebertiere

Danielle is vice president 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 director of research operations 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

Strategic Importance and Market Trajectory

Foundational Role in ADI: Organizations are modernizing analytical data infrastructure to support complex BI and AI initiatives. (see “Special Report: The State of Analytic Data Infrastructure in 2026”). Data engineering serves as the backbone of this infrastructure, providing the essential capabilities for building and managing the pipelines that drive these strategic investments.

Escalating Business Priority: The perceived importance of data engineering has risen steadily since 2020. It is now classified as critical by a majority of organizations, particularly those of larger scale. Data suggests a direct correlation between a high priority on data engineering and the magnitude of ROI realized from BI investments.

Market Adoption and Sourcing Strategies

Vendor Selection and Sourcing Preferences: The market reflects a clear preference for best-of-breed sourcing, though the ability to integrate with a strategic ADI platform remains a core requirement. Preferences vary by geography; for instance, Asia Pacific leans toward best of breed, while Latin American organizations prioritize tight integration with primary ADI vendors.

Broadening Global Adoption: A majority of organizations have already deployed data engineering technology, with significant expansion planned over the next two years. Current adoption is most mature in North America and EMEA, specifically within the healthcare, manufacturing, and financial services sectors.

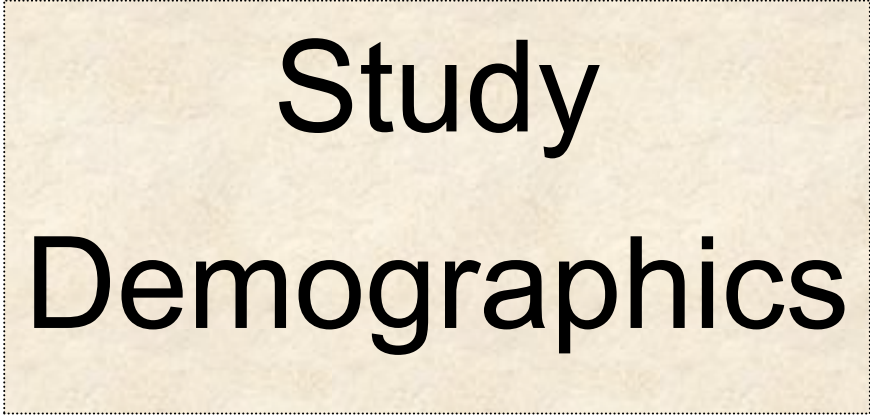
Functional Requirements and Integration

Priority on Ease of Development: Users are prioritizing features that streamline data flow creation, such as graphical development environments, no-code or code-friendly transformations, and AI-assisted capabilities. While modern features are sought after, the market continues to require robust traditional ETL functionalities like complex grouping, scheduling, and monitoring.

Diverse Integration Requirements: Integration flexibility is a key driver of BI success. Preferences are centered on API-based connectivity, ETL/ELT data flows, and standard SQL interfaces. While cloud-native and metadata exchange techniques are currently in the minority, a diverse integration portfolio correlates with higher levels of organizational BI achievement.

Operational Scope and Data Complexity

Expansion Beyond Traditional BI: While data warehousing and ad hoc queries remain primary focuses, data engineering is increasingly utilized for operational use cases. This includes master data delivery, system migrations, and external data sharing. Vendors must align their messaging to address these broader operational requirements.



Study
Demographics

Study Demographics

Study participants provide a cross-section of responses across geographies, functions, organization sizes, and vertical industries. We constructed cross-tab analyses to identify and illustrate important industry preference, priorities, and trends in the data.

Geography

North America, which includes the United States, Canada, and Puerto Rico, represents the largest group of respondents (at 53%), followed by EMEA (23%). Asia Pacific and Latin America account for the balance (25%) of respondents (**Error! Reference source not found.**).

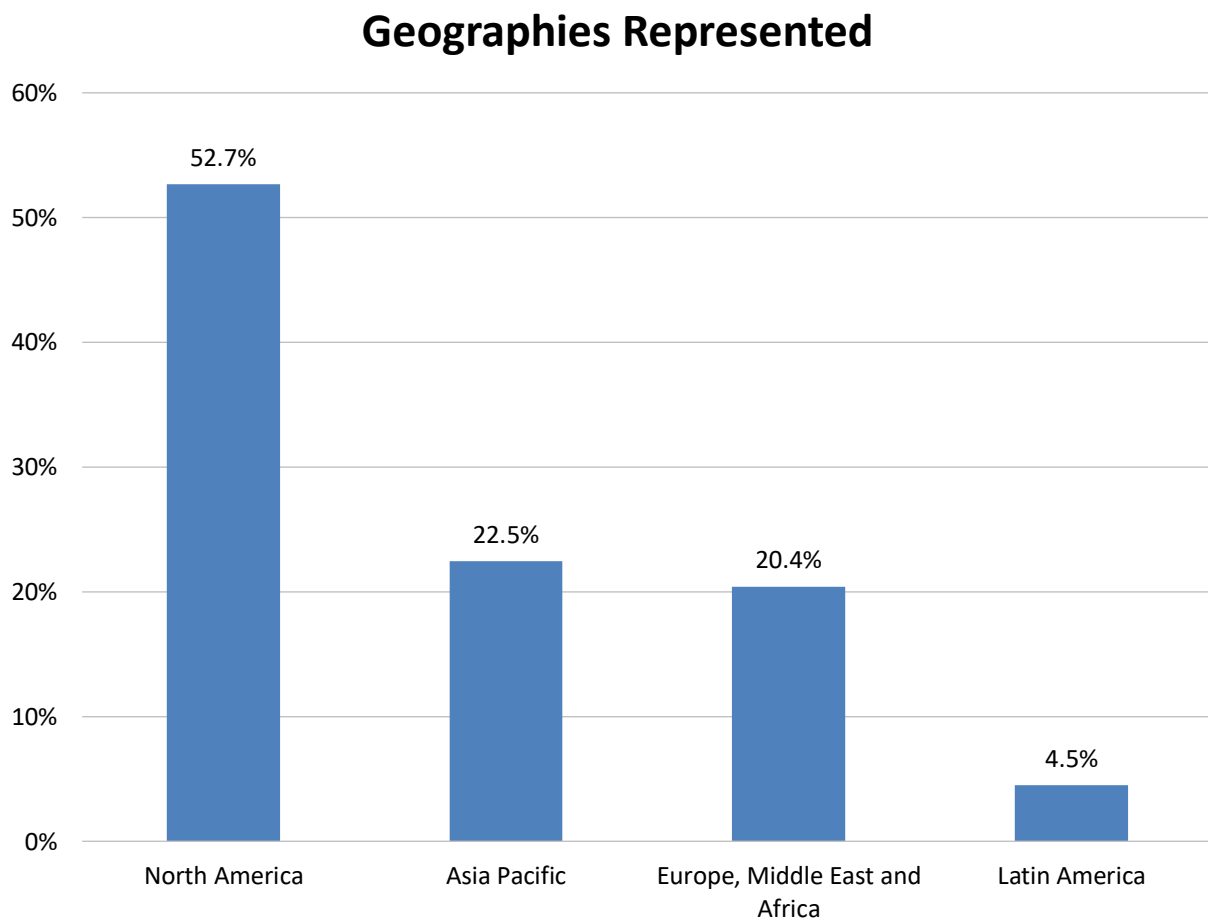


Figure 1 -- Geographies Represented

Functions

In 2026, respondents from IT functions account for 38% of our sample, followed by the Business Intelligence/Analytics Competency Center (BICC) at 19% of the respondents (fig. 2). Together, these two functions make up a majority of the sample, at an aggregate 57%. The remainder of respondents represent a wide variety of functions including data science, R&D, finance, operations, executive management and more.

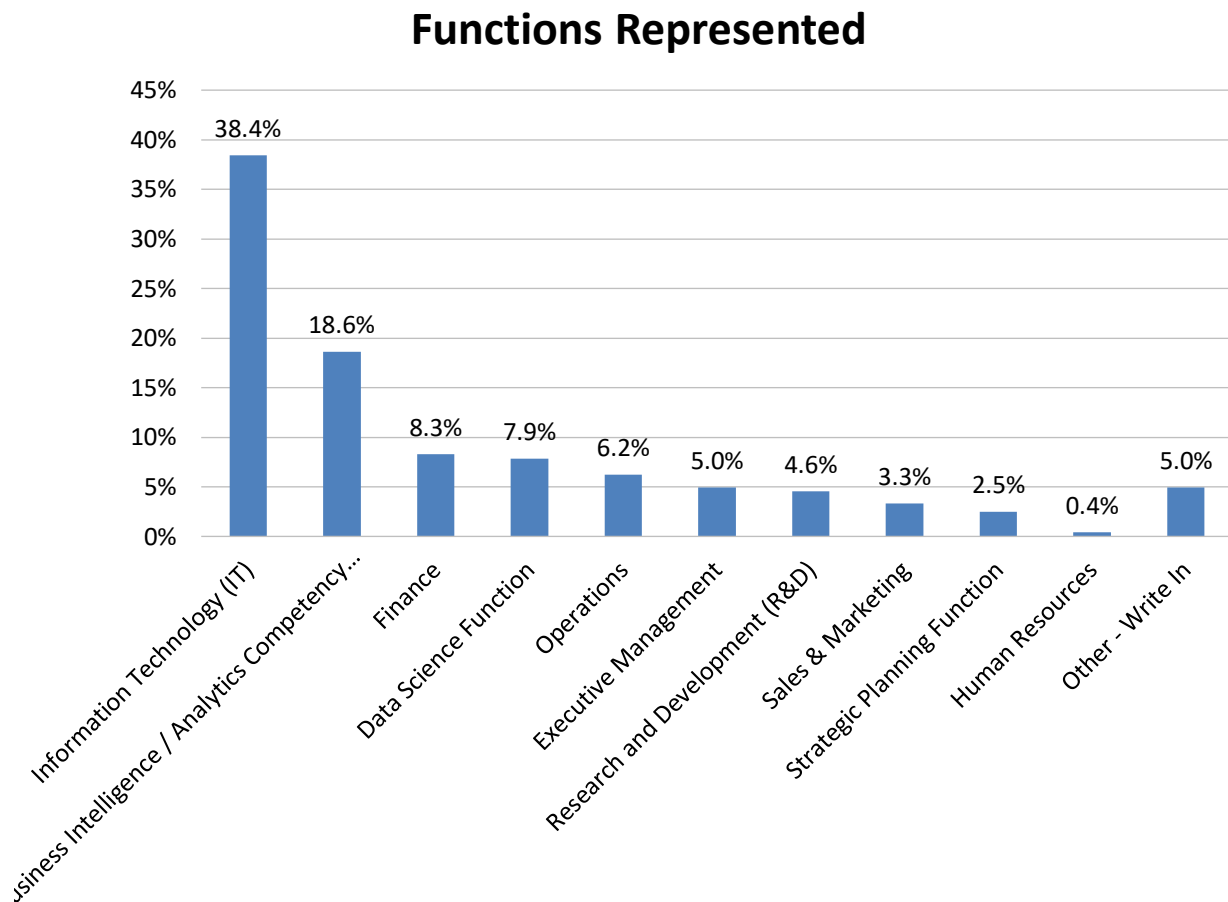


Figure 2 - Functions Represented

Vertical Industries

In 2026, business services (which includes consulting, telecommunications, and transportation) leads the vertical industry distribution of respondents at 23% (fig. 3), followed by manufacturing (19%), financial services and technology (15% each), and consumer services (8%). A range of other industries rounds out the sample, each representing 5% or less.

Vertical Industries Represented

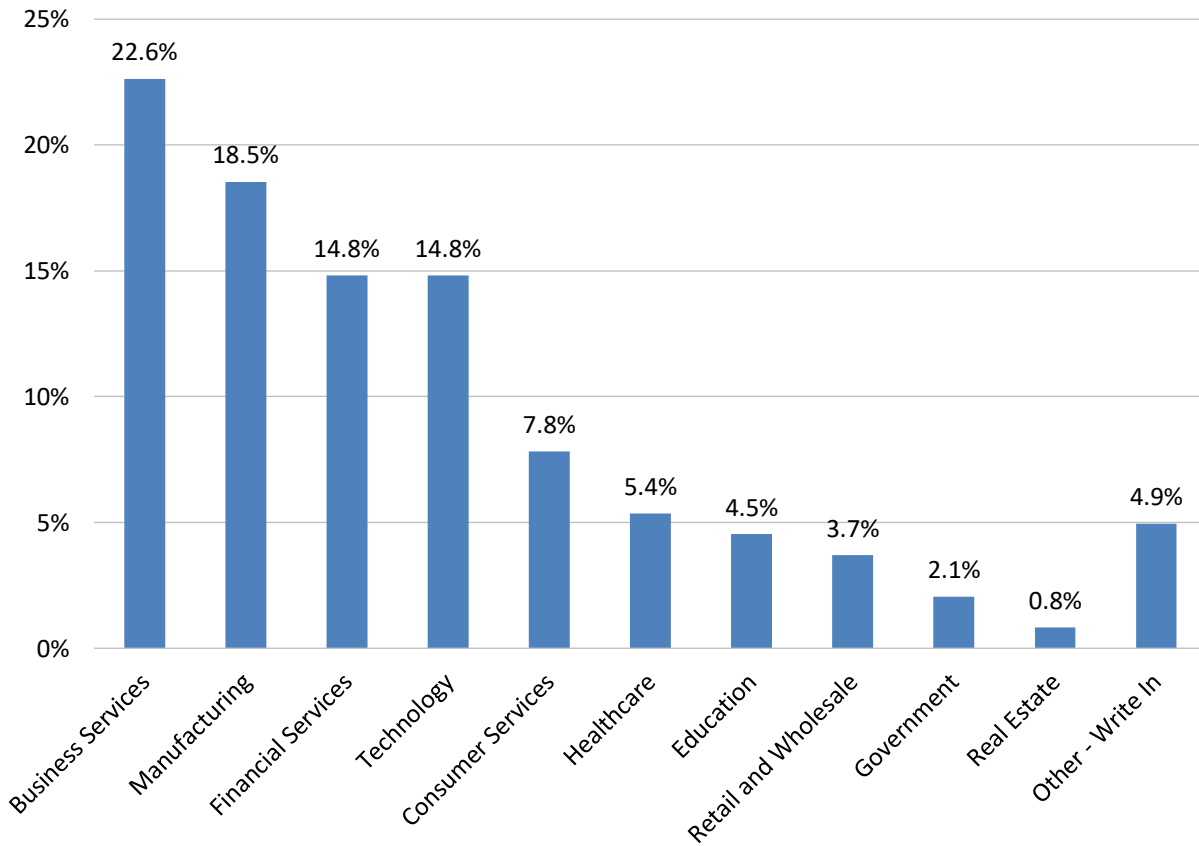


Figure 3 -- Vertical Industries Represented

Organization Size

In 2026, our survey includes small organizations (1-100 employees), midsize organizations (101-1,000 employees), large organizations (more than 1,000 employees) and very large organizations (over 10,000 employees). This year, small organizations account for 20% of our sample, midsize organizations account for 26%, and organizations with 1,001-10,000 employees account for 27% (fig. 4). Twenty-seven percent of the respondents are from organizations with more than 10,000 employees.

Organization Sizes Represented

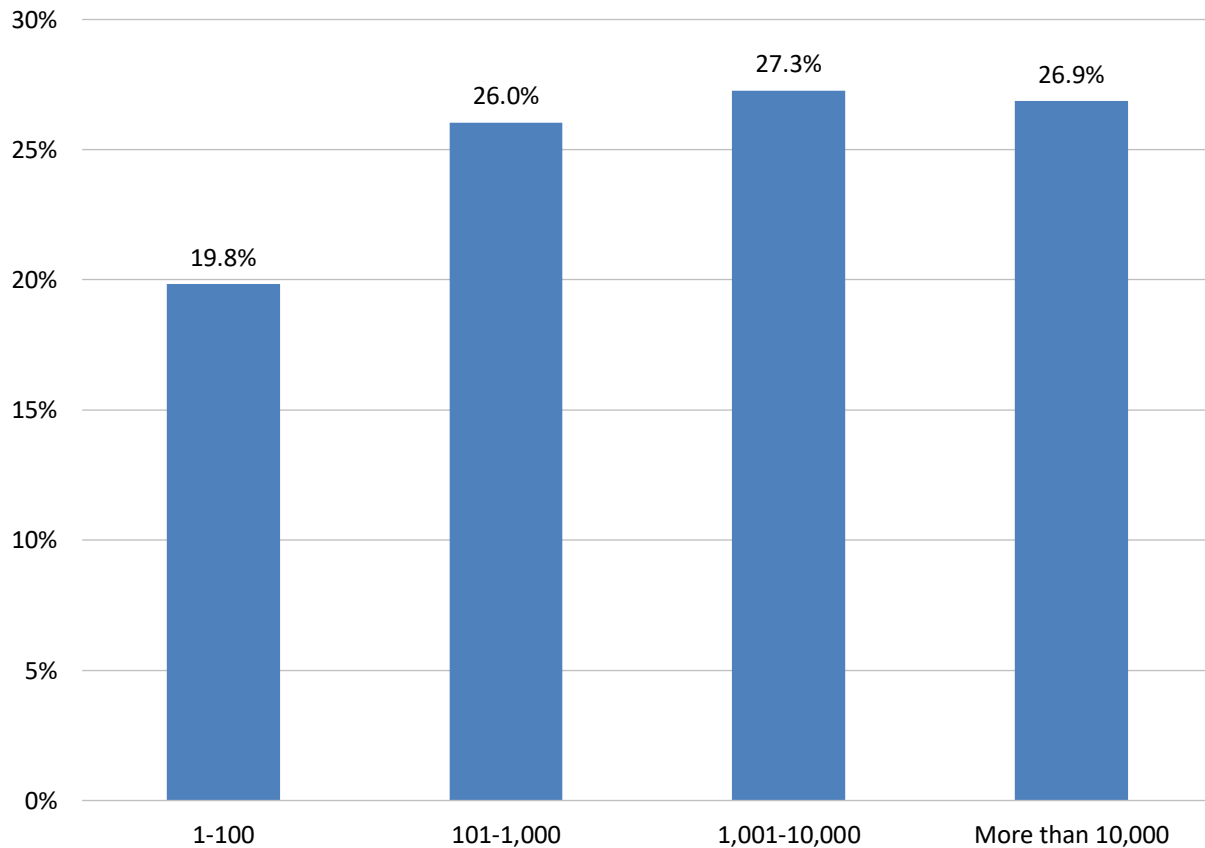


Figure 4 – Organization Sizes Represented



Analysis & Trends

Analysis and Trends

Data Engineering: Key Enabler of Modern Analytic Data Infrastructure

An analytical data infrastructure (ADI) is the data workflow and technology ecosystem responsible for the ingestion, persistence, and transformation of data from source systems for downstream consumption by analysts and analytic workflows and AI agents. Architecturally, it is defined by a range of components working together to support a diverse range of requirements, data types, platforms and use cases. Data engineering is an increasingly critical enabler of a successful ADI, with 82% of respondents indicating they perceive it as important (fig. 5).

Importance of ADI Components

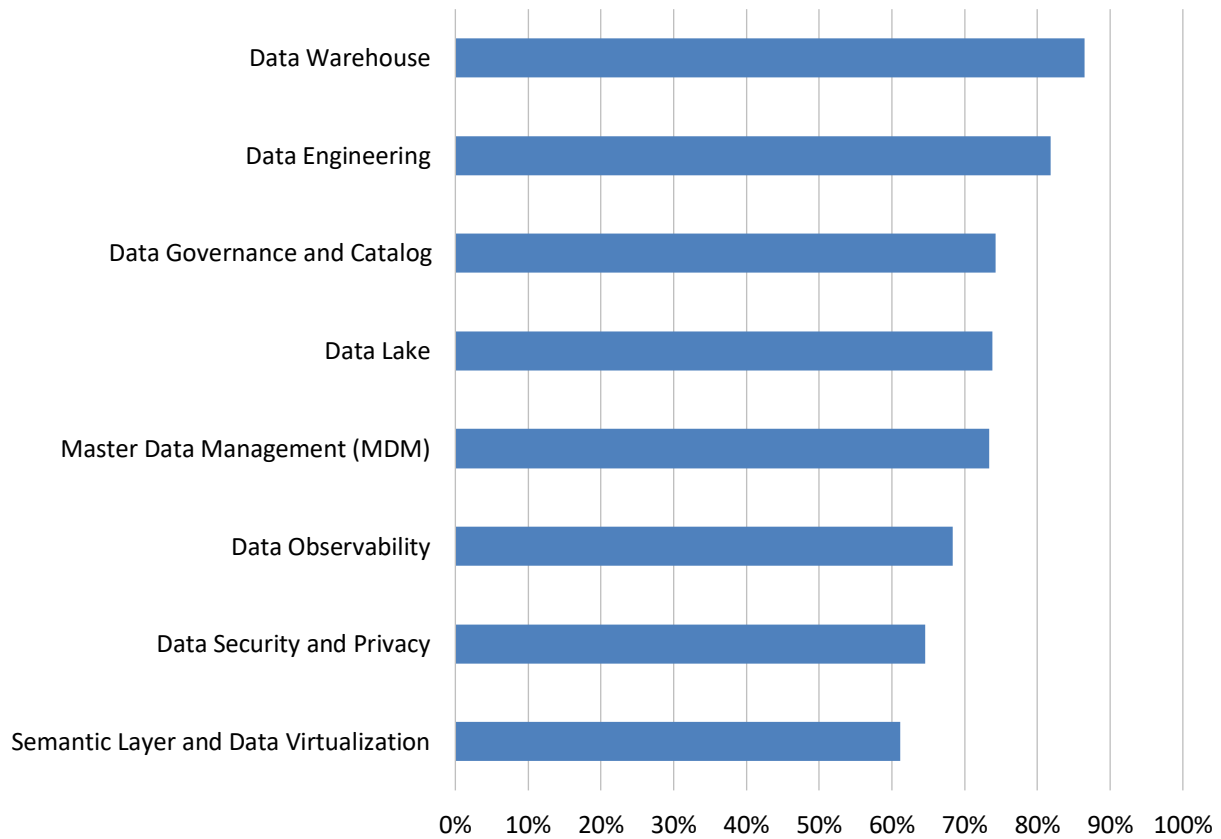


Figure 5 – Importance of ADI Components

2026 Data Engineering Market Study

A critical insight about the role of data engineering within ADI is its importance for nearly any use case an organization is trying to support. Next to data warehouse capabilities, survey respondents rate data engineering as the most relevant component across all use cases (table 1). At least 33% of respondents identified its importance for every use case presented, with this rate rising to 40%-50% for critical use cases such as executive dashboards, customer retention, financial planning and customer 360/personalization. This testifies to the already established and continually rising priority of data engineering capabilities within the industry.

	Data Warehouse	Data Engineering	Data Governance and Catalog	Data Lake	Master Data Management (MDM)	Data Observability	Data Security and Privacy	Semantic Layer and Data Virtualization
Executive Dashboards & KPI Reporting	50.6%	46.4%	31.7%	27.4%	30.8%	30.8%	32.1%	32.9%
Churn Prediction & Customer Retention	44.3%	50.2%	25.7%	30.4%	23.6%	21.9%	27.4%	21.1%
Financial Planning & Analysis (FP&A)	48.5%	43.9%	33.3%	24.1%	29.1%	24.5%	26.6%	25.7%
Customer 360 & Personalization	43.9%	40.5%	22.4%	29.1%	30.8%	22.4%	27.9%	22.4%
Sales & Revenue Forecasting	40.5%	38.0%	25.7%	24.5%	30.0%	27.4%	25.7%	28.3%
Data Monetization & External Data Products	36.7%	39.2%	27.9%	32.9%	24.9%	21.9%	27.9%	20.7%
Operational Performance Monitoring	35.4%	37.6%	26.6%	29.1%	21.1%	30.8%	20.3%	25.3%
Fraud Detection & Risk Management	29.5%	35.0%	27.4%	29.5%	22.4%	26.6%	37.6%	18.1%
Human Capital Analytics	35.0%	35.4%	24.9%	26.2%	27.9%	21.9%	26.2%	21.5%
Regulatory Compliance & Audit Readiness	32.9%	33.3%	37.1%	21.5%	29.1%	28.7%	36.3%	18.1%
Product Analytics & Feature Usage	35.4%	36.3%	22.8%	26.6%	21.9%	26.2%	18.6%	19.4%
Marketing Attribution & Campaign ROI	35.0%	34.2%	18.6%	27.4%	24.5%	21.1%	19.8%	25.3%
R&D and Experimentation Tracking	32.9%	35.9%	23.2%	24.9%	22.8%	25.3%	22.8%	20.7%
IoT Analytics & Predictive Maintenance	26.2%	35.9%	16.0%	29.5%	18.6%	26.2%	21.1%	16.5%
Supply Chain Optimization	36.7%	35.4%	23.2%	24.9%	24.1%	23.6%	21.1%	22.8%
	16.0%	24.7%		33.3%		42.0%		50.6%
Percentage of All Respondents								

Table 1 – Importance of ADI Components Across Use Cases

Importance of Data Engineering

We asked respondents about the importance of data engineering, regardless of use case within the ADI universe. In our 2026 survey, 48% of respondents indicate data engineering is of critical importance (fig. 6) while 31% indicate it is very important and another 16% see it as important. Less than 5% indicate that data engineering is somewhat important or not important.

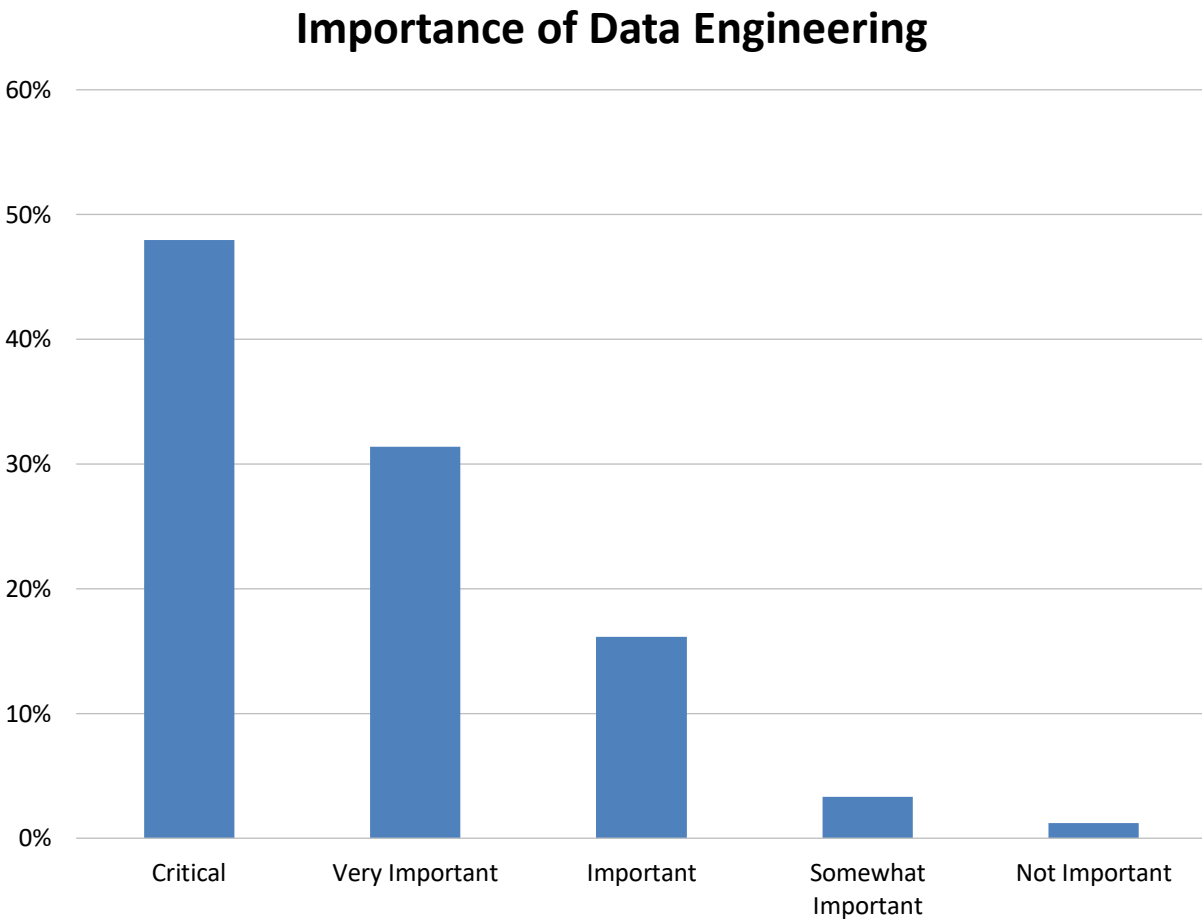


Figure 6 – Importance of Data Engineering

Data engineering continues its year-over-year trend of escalating in importance, driven by an increase in critical importance responses (fig. 7). For 2026, 48% of respondents rated it as critical, representing the next step in a steady increase in criticality since 2019 (when only 28% rated it as such). A full 80% view it as either critical or very important at present, and we expect data engineering will remain highly important for some time to come as organizations move to hybrid and multi-cloud deployments that require more complex data engineering workflows.

Importance of Data Engineering 2015-2026

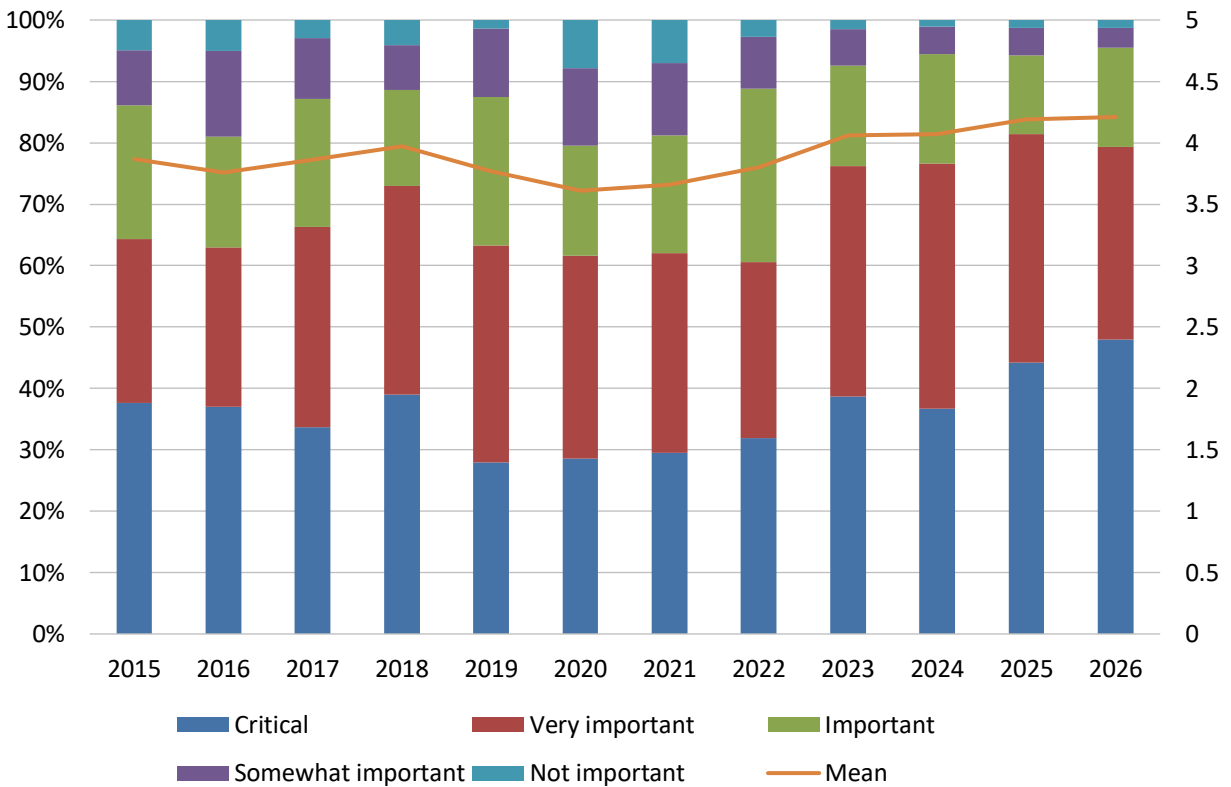


Figure 7 – Importance of Data Engineering 2015-2026

This year, respondents from strategic planning functions emphatically place the highest overall importance on data engineering, with 70% rating it critical and 100% viewing it as at least very important (fig. 8). The data science function views data engineering as nearly the same priority, with 65% rating it critical and 90% seeing it as at least very important. Other functions showing the greatest perception of data engineering’s importance include finance, the BICC and IT. Notably, a majority of respondents across all functions rate it as critical or very important.

Importance of Data Engineering by Function

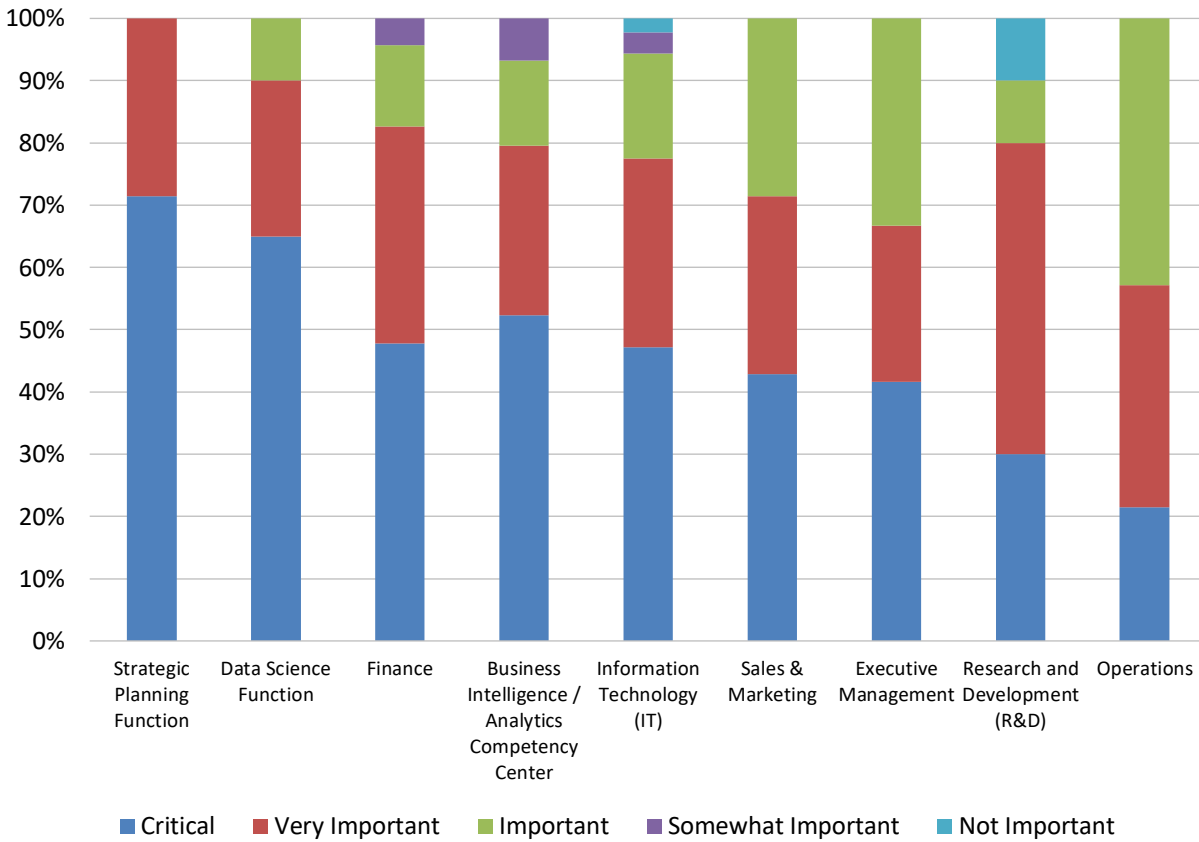


Figure 8 – Importance of Data Engineering by Function

Data engineering importance appears somewhat correlated to organization size, with larger organizations more likely to rate it as critical or very important. In 2026, 72% of small organizations with 1-100 employees viewed data engineering as critical or very important (fig. 9). This metric increases to 78% for organizations with 101-1,000 employees, 81% for organizations with 1,001-10,000 employees, and 85% for the largest organizations with more than 10,000 employees. Also notable is that a strong majority, at 85% or higher, see data engineering as important, very important or critical across all organization size categories.

Importance of Data Engineering by Organization Size

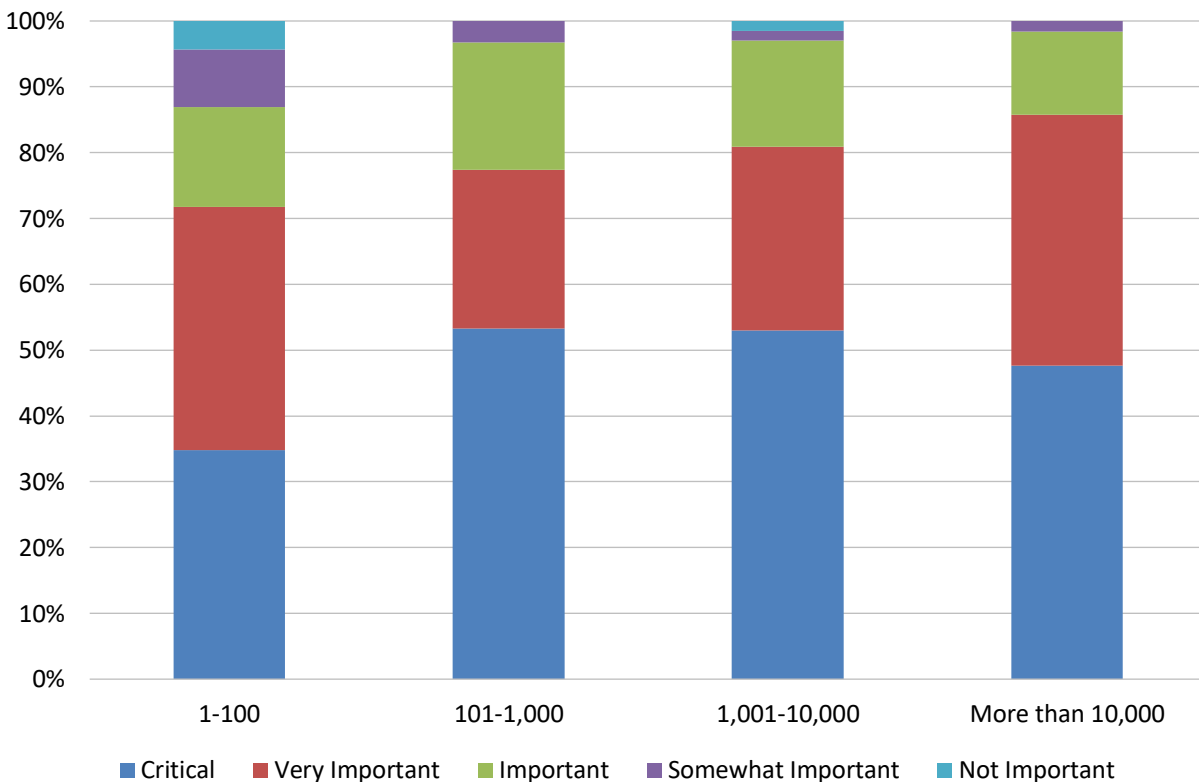


Figure 9 – Importance of Data Engineering by Organization Size

Although all industries place a high priority on data engineering, respondents from data-intensive industries undergoing significant change and modernization, such as healthcare, retail and wholesale and financial services, give it the highest relative priority. These three industries all rate it as critical or very important with 84% frequency or higher (fig. 10). Manufacturing, consumer services and technology organizations follow closely in their perceived importance of data engineering. Business services and education rank the importance of data engineering lowest among all industries, although a majority of organizations in those sectors still see it as critical or very important.

Importance of Data Engineering by Industry

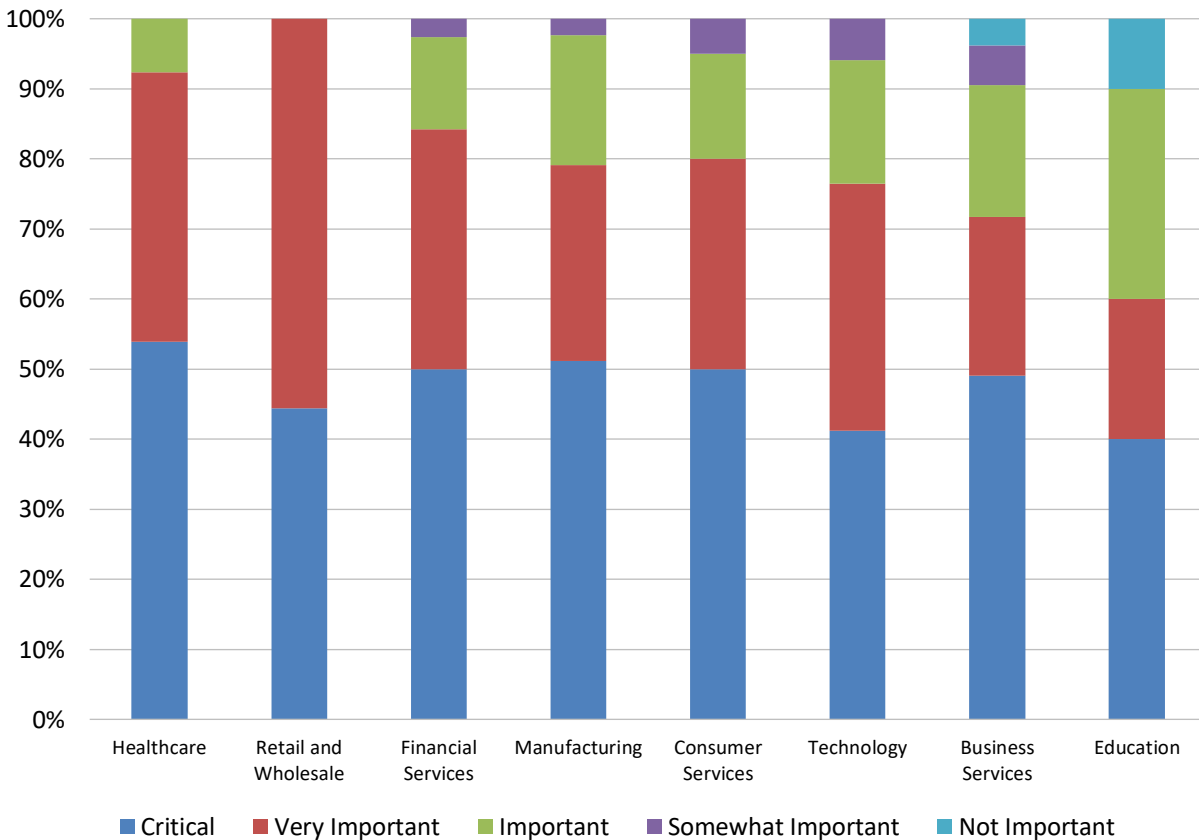


Figure 10 – Importance of Data Engineering by Industry

Modern BI use cases require more sophistication and flexibility to support complex data pipelines and address more challenging data transformation and provisioning needs. As such, it makes sense that organizations treating data engineering as an important capability supporting their BI programs enjoy greater success. A full 88% of organizations with extremely successful BI initiatives also rate data engineering as critical or very important (fig. 11), as do 82% of those with very successful BI initiatives. In contrast, those organizations with moderately successful or somewhat unsuccessful BI efforts only rate data engineering as critical or very important 72% and 65% of the time, respectively.

Importance of Data Engineering by BI Success

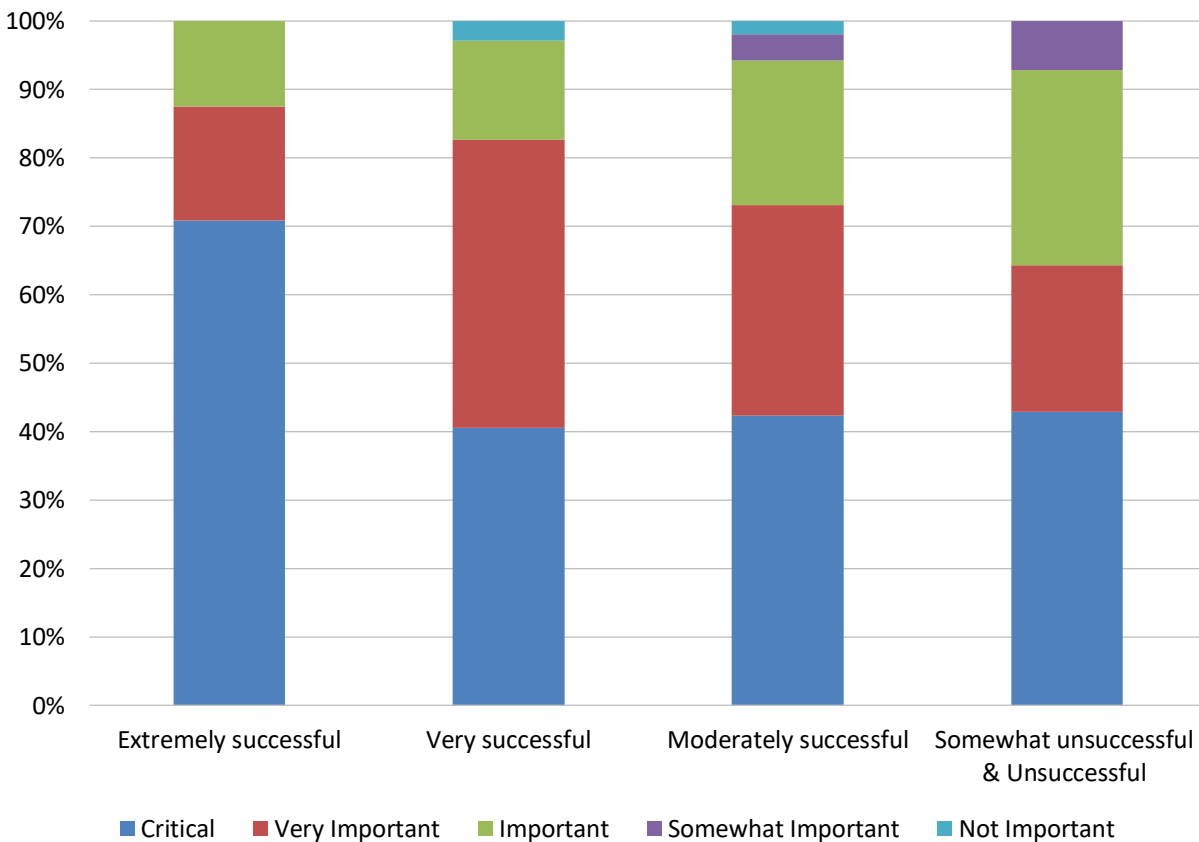


Figure 11 – Importance of Data Engineering by BI Success

In a similar manner, we observe a clear correlation between organizations' views of data engineering importance and their ability to achieve ROI from BI investments. Those organizations capturing extremely high ROI of over 50% see data engineering as critical with greater than 90% frequency (fig. 12). Those returning a very high and transformative ROI level, between 24% and 50%, rate data engineering as critical 53% of the time. Organizations getting high and significant ROI levels of 12% and those returning a good ROI of 6% view data engineering as critical with around 35% frequency.

Importance of Data Engineering by BI ROI

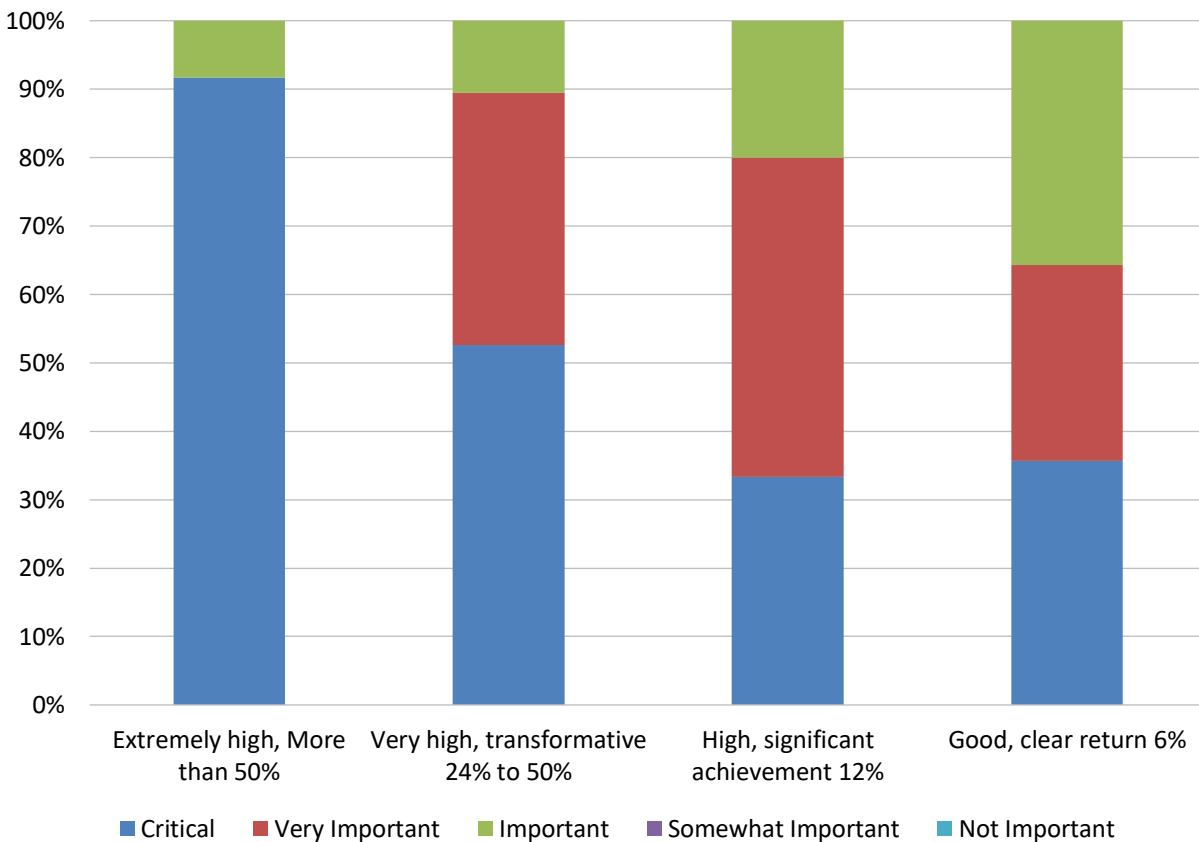


Figure 12 – Importance of Data Engineering by BI ROI

Impending requirements for more robust data pipelines to feed AI models are also driving data engineering demand. Organizations perceiving AI as a current and critical area of investment are more likely to rate data engineering importance highly. Those that see AI as a cornerstone of their current business strategy overwhelmingly rate data engineering as critical or very important, at 88% (fig. 13). In contrast, organizations for which AI plays an important or exploratory role see data engineering as critical or very important far less often, at levels of 76% and 59% respectively.

Importance of Data Engineering by AI Strategic Role

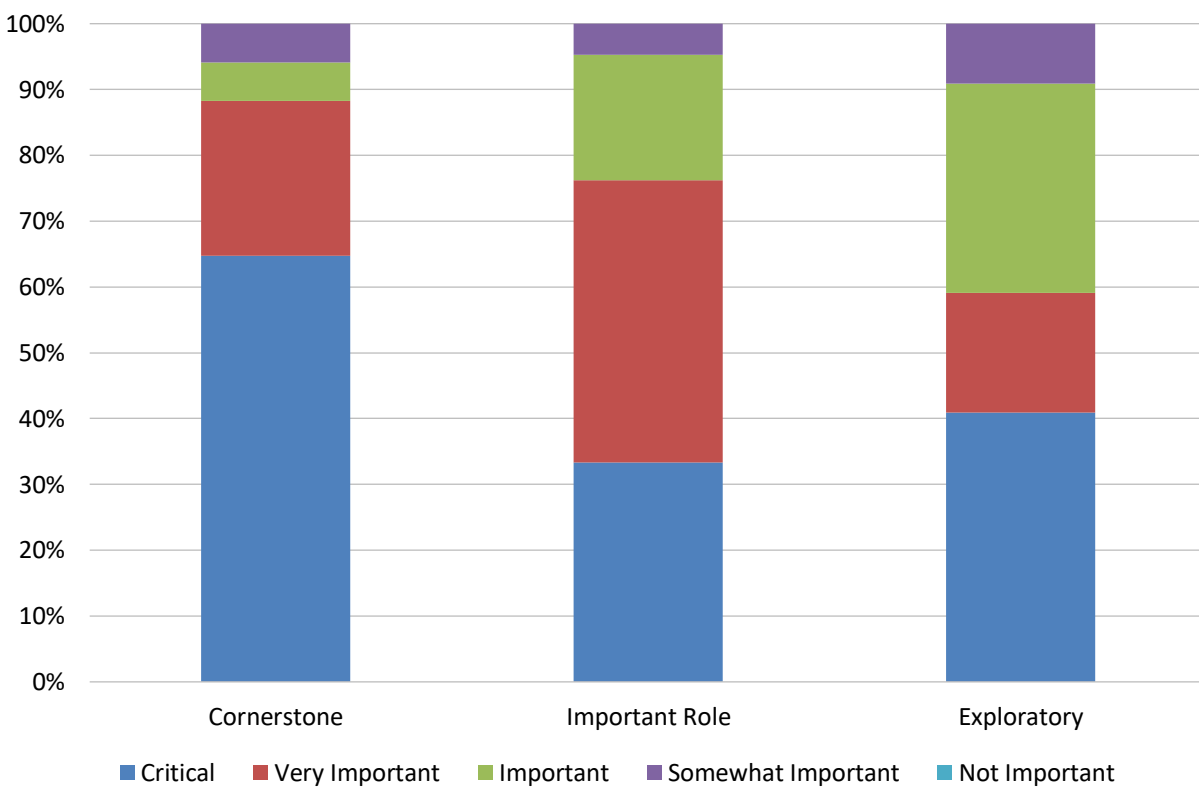


Figure 13 – Importance of Data Engineering by AI Strategic Role

Current Adoption and Future Plans for Data Engineering

A majority of organizations already have data engineering capabilities in place, and the rate of adoption is set to grow over the next two years. Fifty-nine percent of organizations indicated that they have deployed data engineering today (fig. 14). Another 20% plan to do so in the next 12 months, and a further 12% anticipate doing so in 24 months. Only 9% of respondents indicated no plans for data engineering.

Data Engineering Adoption

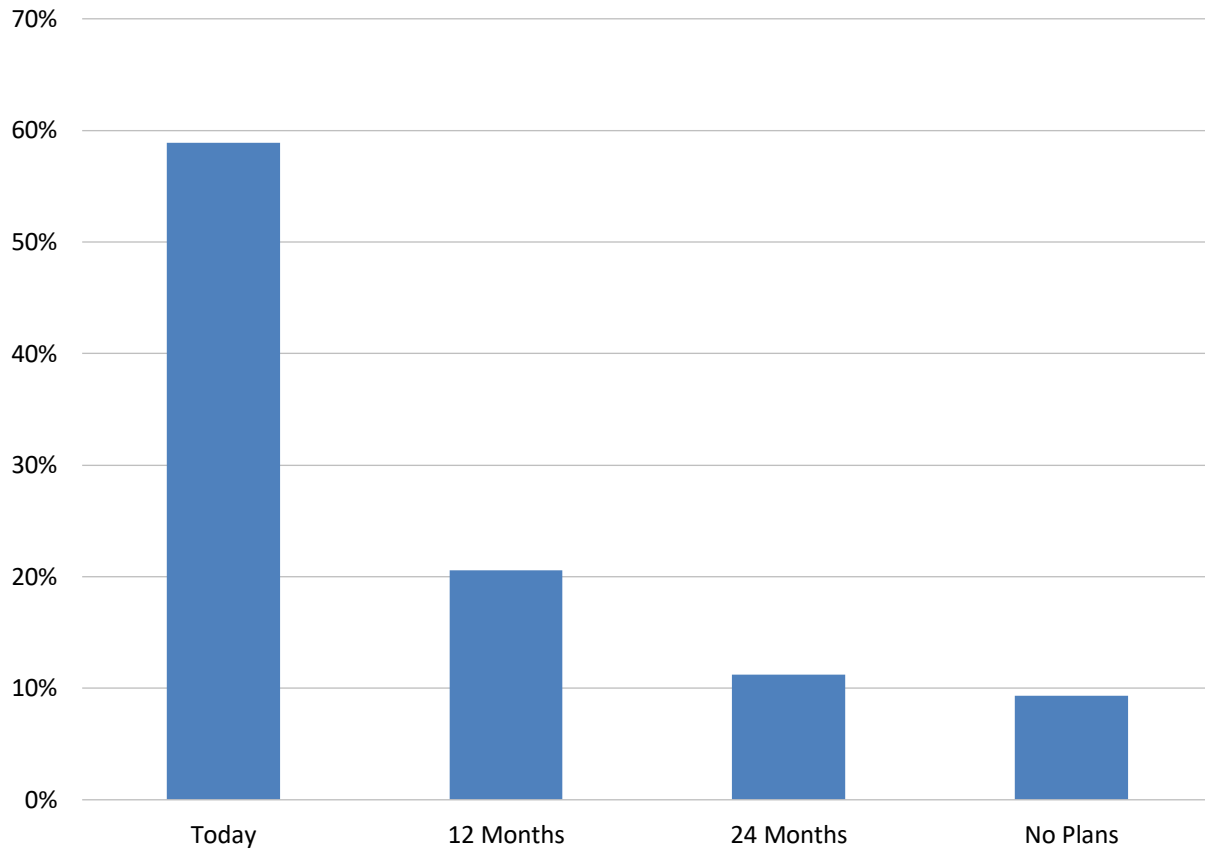


Figure 14 – Data Engineering Adoption

Analyzed by geography, a strong majority of organizations in all regions have either already adopted data engineering or plan to do so within 12 months. However, EMEA and North America organizations show the greatest current adoption levels with 64% and 65% of those regions, respectively, having it in place today (fig. 15). Latin American organizations are the next-highest level of current adoption at 55%. Asia Pacific organizations reflect the lowest levels of adoption today, at only 42%.

Data Engineering Adoption by Geography

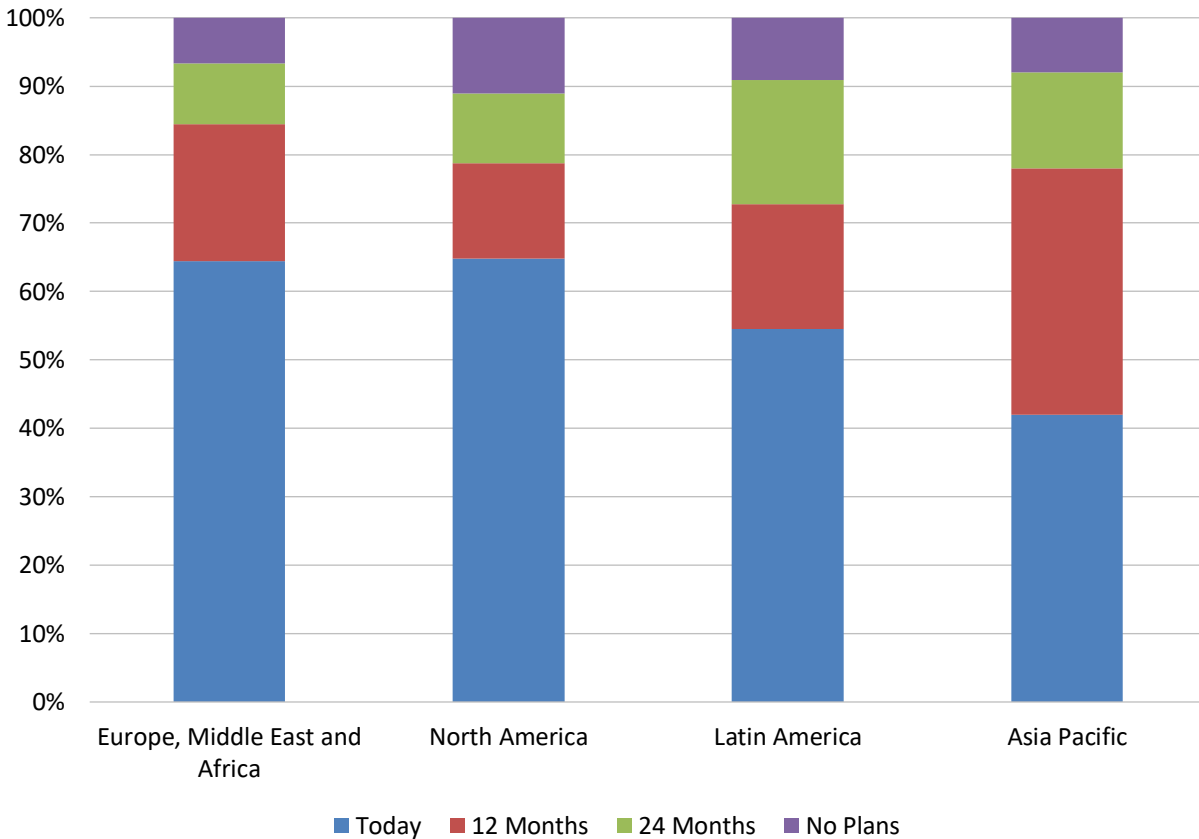


Figure 15 – Data Engineering Adoption by Geography

Levels of data engineering adoption vary significantly across industry sectors, although a majority of organizations in most industries report adoption today. Healthcare leads other industries in data engineering adoption: 83% of organizations attest to adoption today, with another 17% planning to adopt in 12 months (fig. 16). All other industries show more variability, with manufacturing, business services, financial services and retail/wholesale all reporting around 60% current adoption. The technology and education sectors appear to be lagging in data engineering adoption, at only 45% and 40% today, respectively.

Data Engineering Adoption by Industry

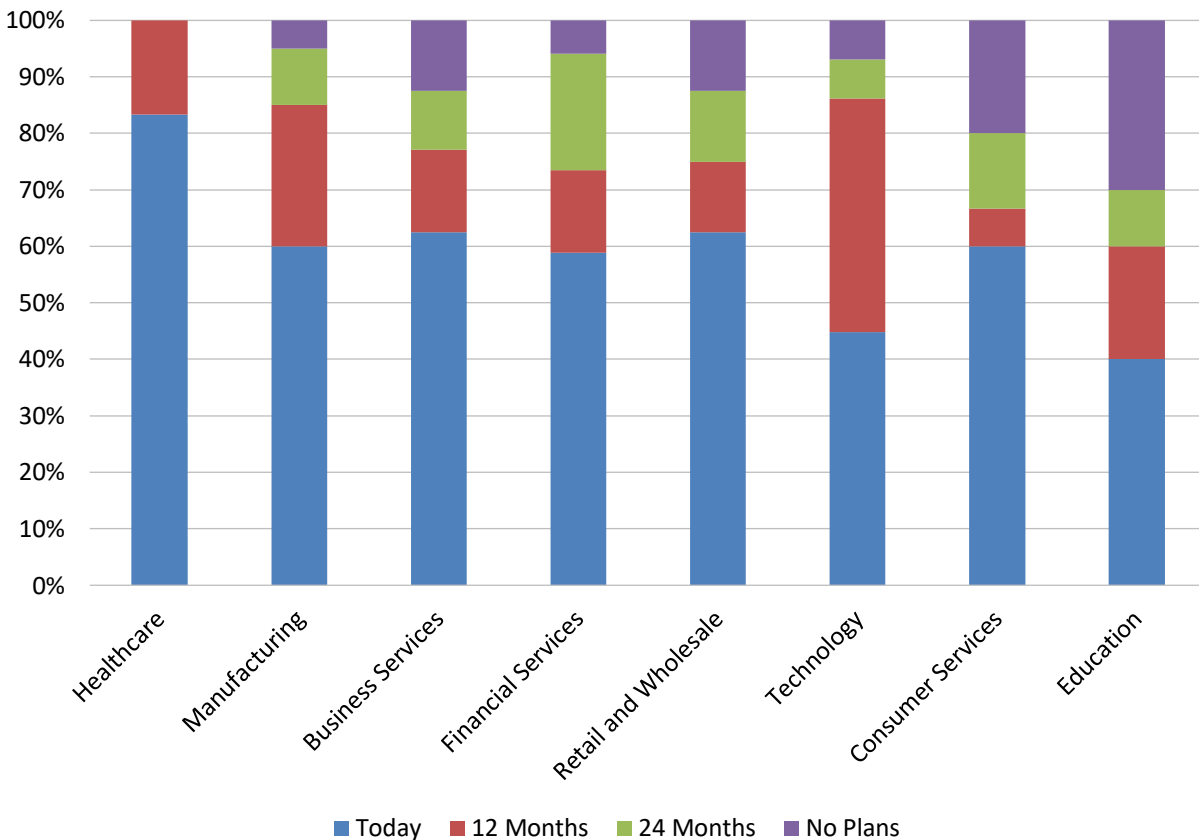


Figure 16 – Data Engineering Adoption by Industry

From a business function perspective, it is not surprising to see the highest adoption rates of data engineering within functions whose business processes and analytic needs are data-intensive and fast-moving. Strategic planning, the BICC and data science functions show the strongest current adoption, with 84%, 80% and 76% of respondents in those functions, respectively, reporting adoption today (fig. 17). Executive management also stands out with 64% current adoption. All other functions lag, although roughly half of respondents in those functions report adoption today. Looking at planned adoption within 12 months, all functions except R&D will show a strong majority of adoption in a year's time.

Data Engineering Adoption by Function

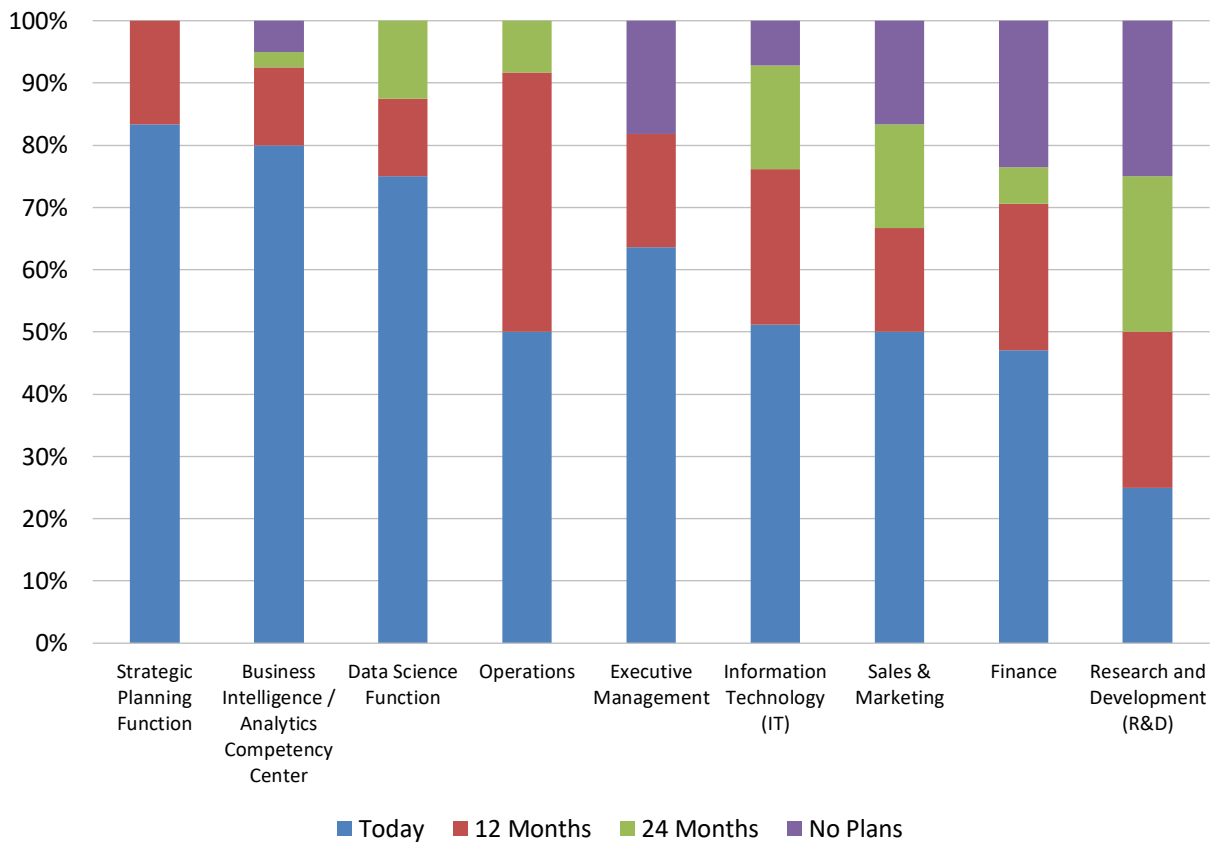


Figure 17 – Data Engineering Adoption by Function

The bigger an organization gets, the greater the number of data sources, critical business processes, and complexity of data management. This is why organization size and data engineering adoption show a clear correlation (fig. 18). Among organizations in our sample, 47% of small organizations with 1-100 employees have deployed data engineering capabilities today. This rises to 55% for medium-sized organizations with 101-1,000 employees. At the 1,001-10,000-employee level, 57% of organizations have adopted data engineering today. And 72% of the largest organizations, with more than 10,000 employees, have done so.

Data Engineering Adoption by Organization Size

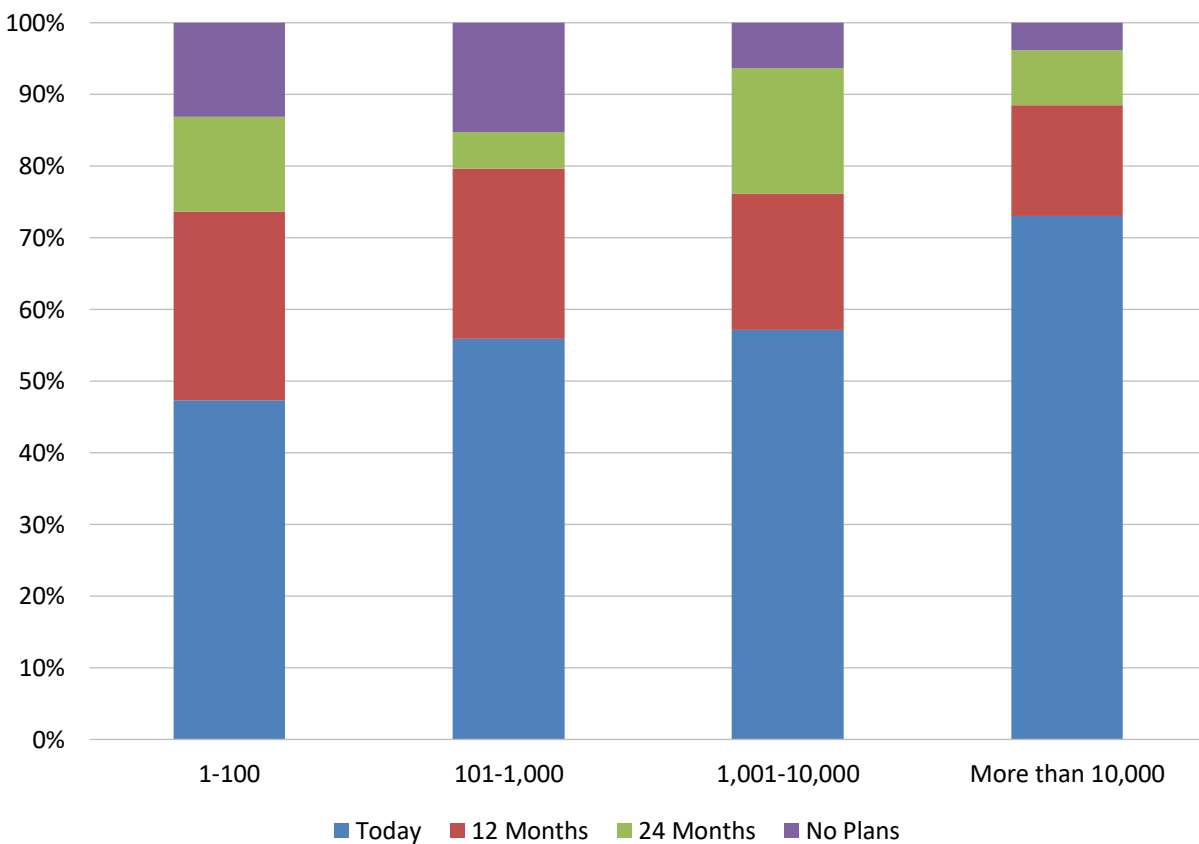


Figure 18 – Data Engineering Adoption by Organization Size

Use Cases for Data Engineering

Data engineering capabilities are used for an ever-expanding range of use cases in support of BI, analytics and AI. Consistent with traditional data management priorities, data delivery that supports data warehousing comprises the greatest activity among all use cases, with 33% of organizations indicating that more than 60% of their data engineering work is applied here, and 58% indicating at least 41% or their work is focused here (fig. 19). A similar use case, supporting data integration and transformation for data discovery and query, is next most popular, with approximately 47% of organizations stating that it comprises 41% or more of their data engineering work. A range of other use cases, including many that are not purely BI- and analytics-focused, are also highly relevant.

Prioritization of Data Engineering Use Cases

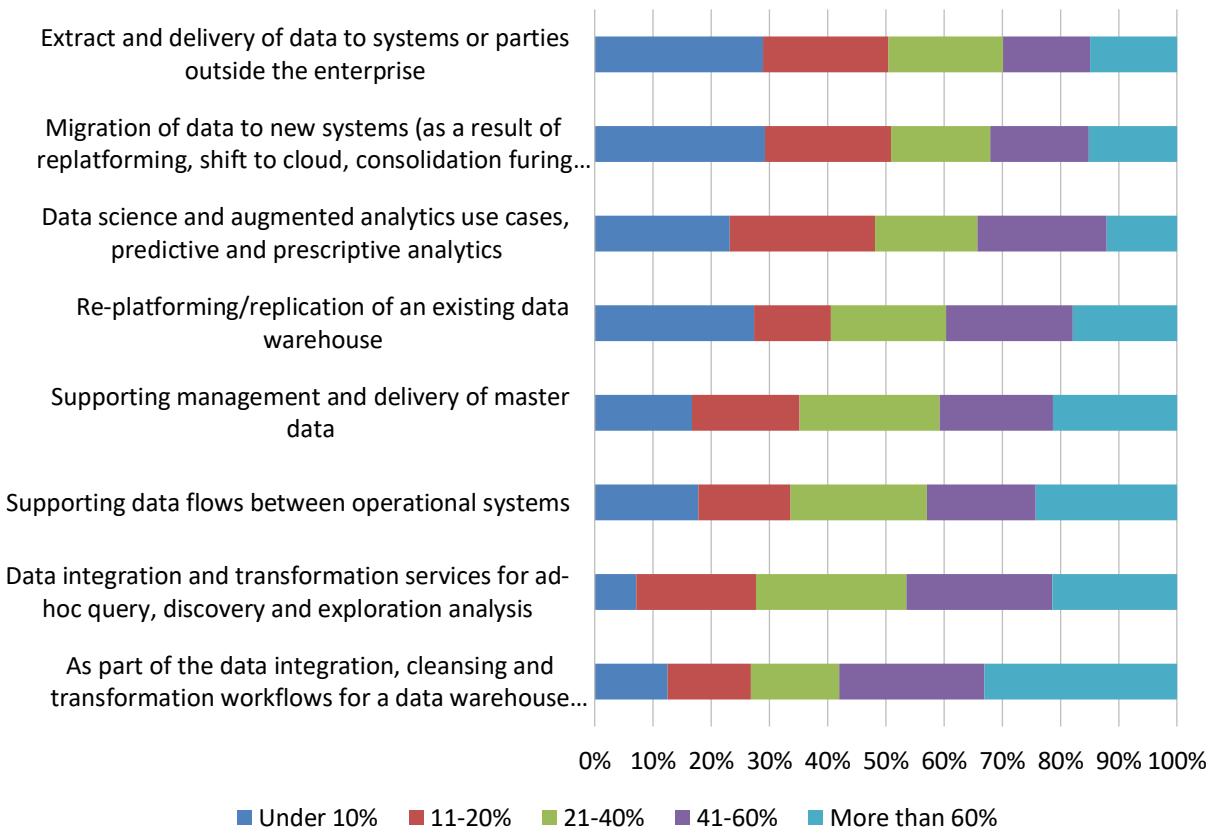


Figure 19 – Prioritization of Data Engineering Use Cases

While the use cases of data warehousing and data discovery/query have been mainstays of data engineering activity for several years, it is interesting to note the rise of other use cases, many of which are more operational in nature (fig. 20). In 2025 and 2026, activity around operational data flows, master data management, system migrations and external data delivery, for example, appears to be garnering more attention for data engineering capabilities. The rate at which organizations identify the importance of data engineering for the master data management use case grew by 30% from 2025 to 2026. The perceived importance of data engineering for external data delivery and operational data flow use cases increased by 20% and 15%, respectively, during the same period.

Data Engineering Use Cases 2022-2026

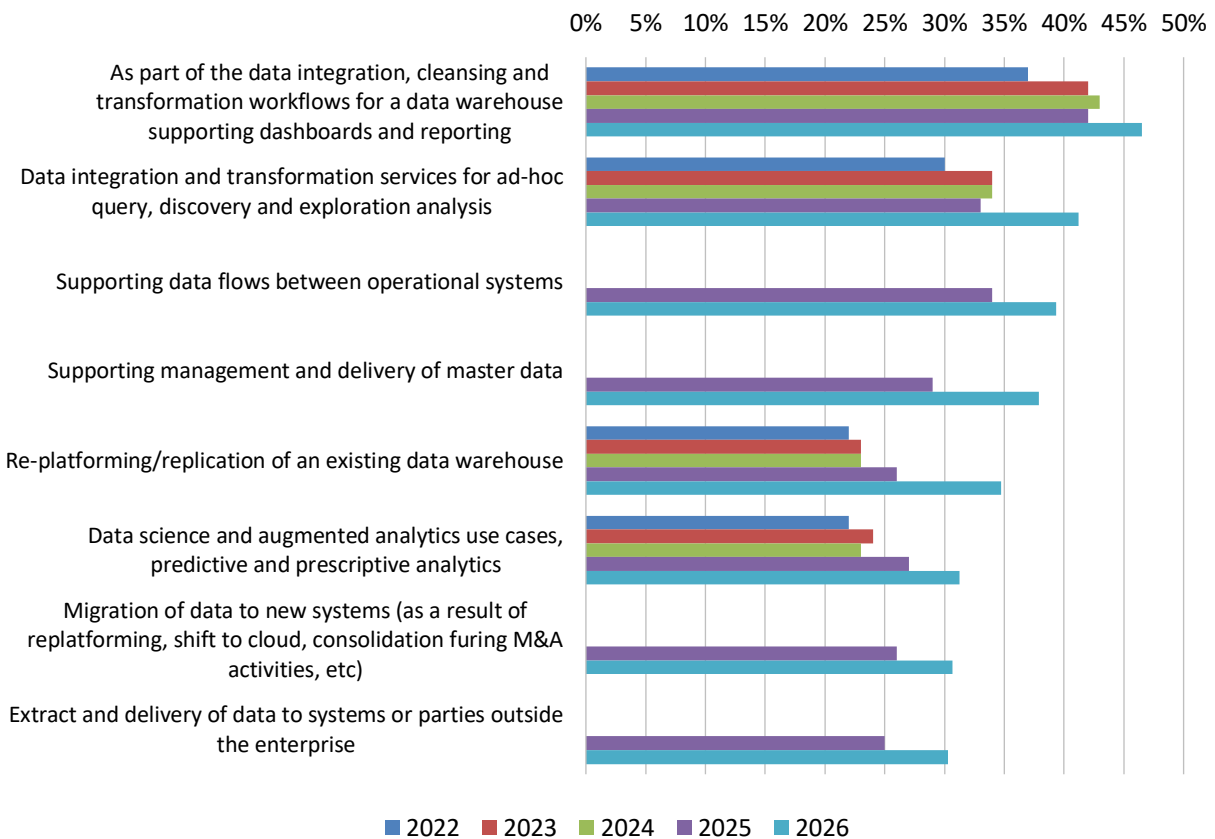


Figure 20 – Data Engineering Use Cases 2022-2026

We filtered data-engineering usage according to the highest percentage of use case responses and geographic regions (fig. 21). The percentage of data engineering activity against the key use cases varies by geography. North America and EMEA show similar percentages in data engineering investments for supporting main BI-related use cases such as data warehousing, ad hoc query and data science. Latin American organizations emphasize data integration and transformation in support of ad hoc query, master data management and data science. Asia Pacific organizations emphasize data integration and transformation in support of ad hoc query, master data management and data science.

Data Engineering Use Cases by Geography

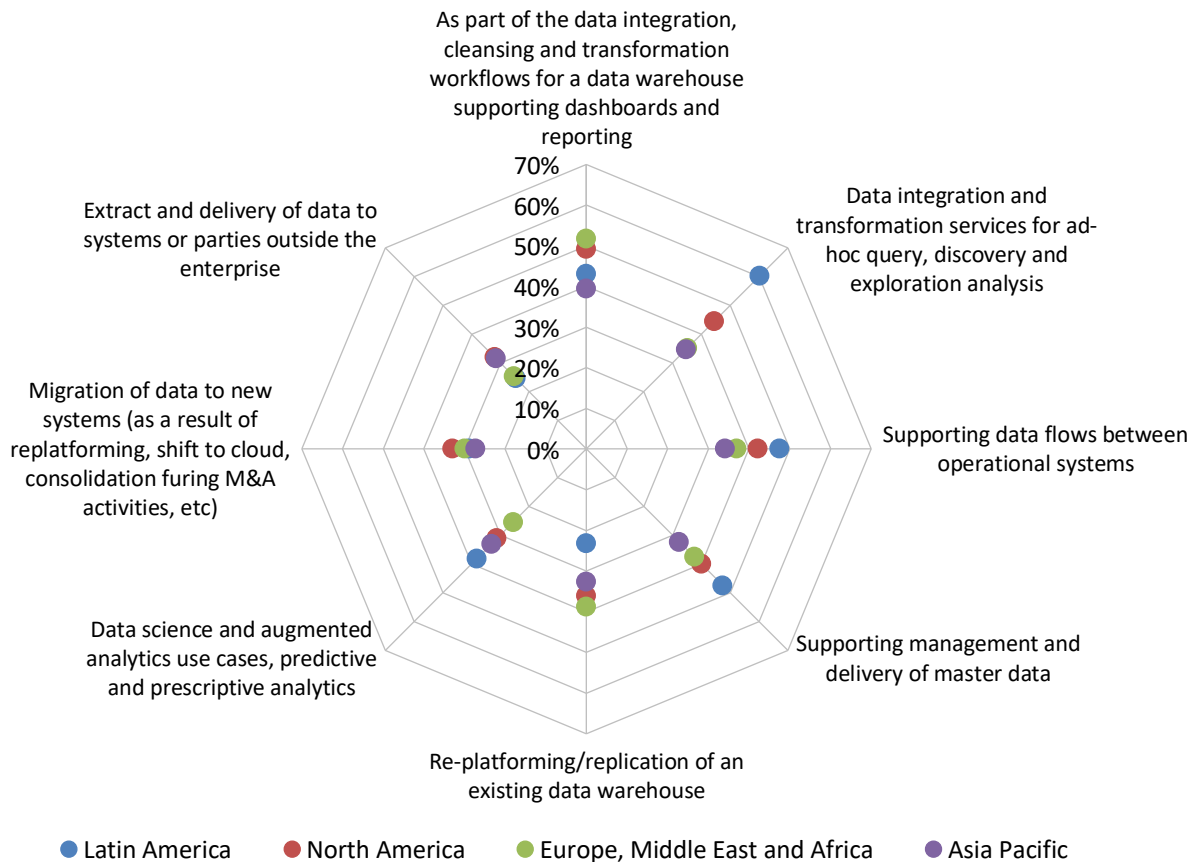


Figure 21 - Data Engineering Use Cases by Geography

Data engineering activity varies significantly across industry sectors. Nearly all industries place strong emphasis on the “supporting the data warehouse” use case, followed closely by “data integration and transformation in support of ad hoc query and data discovery” (fig. 22). Data science-related use cases are also a consistently high priority across most industries. Emphasis on other key use cases, such as external data delivery, master data management, and data flow between operational systems, varies more widely across industries. Industries such as manufacturing and healthcare focus substantial data engineering efforts on these use cases, while financial services and technology organizations do so to a much lesser degree. In keeping with the overall perception of data engineering importance in the education sector, in that industry all use cases garner minimal attention.

Data Engineering Use Cases by Industry

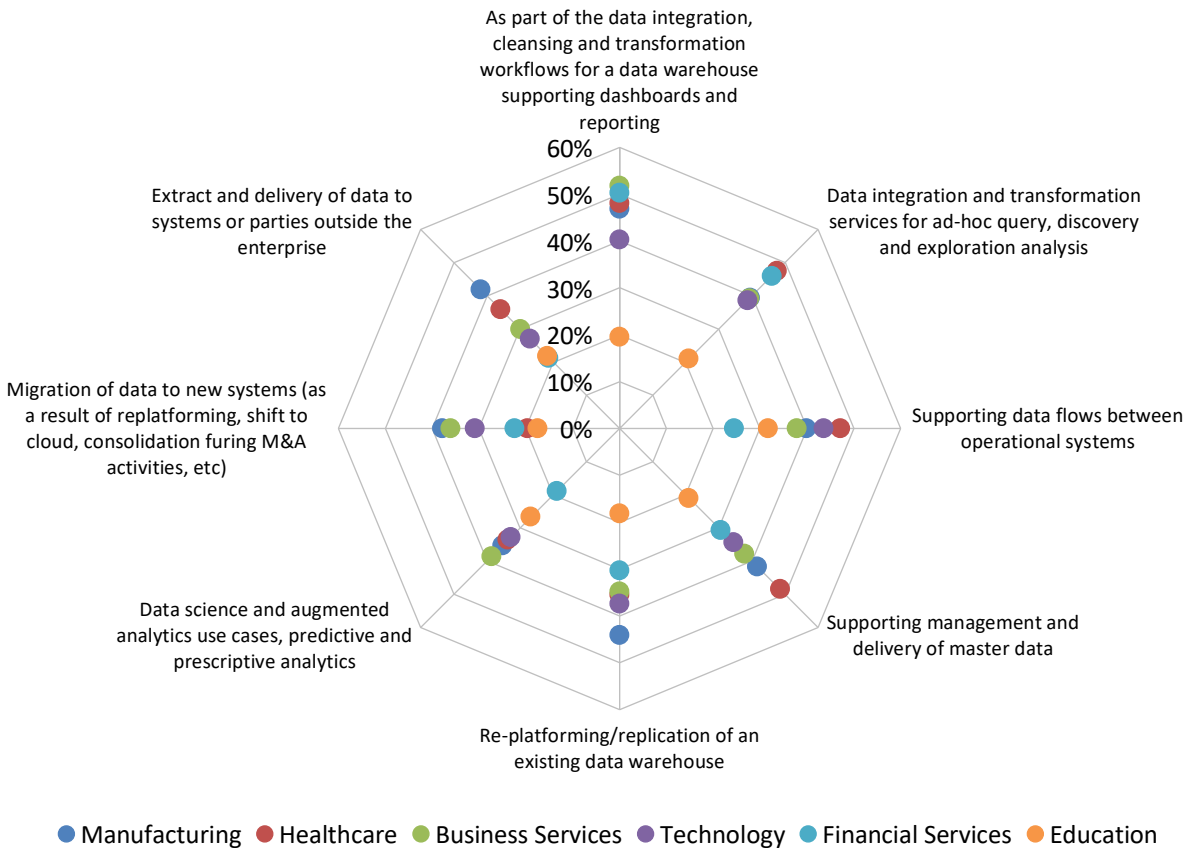


Figure 22 – Data Engineering Use Cases by Industry

Application of data engineering to specific use cases varies even more significantly by business function. Data engineering for data warehousing, data integration for ad hoc query, operational data flows and master data management are prominent use cases for respondents in most functions (fig. 23). The finance function particularly stands out for these use cases. The data science function prioritizes operational data flows, while the operations function most heavily emphasizes replatforming of existing data warehouses. The R&D function places the highest priority on the data science and augmented analytic use case.

Data Engineering Use Cases by Function

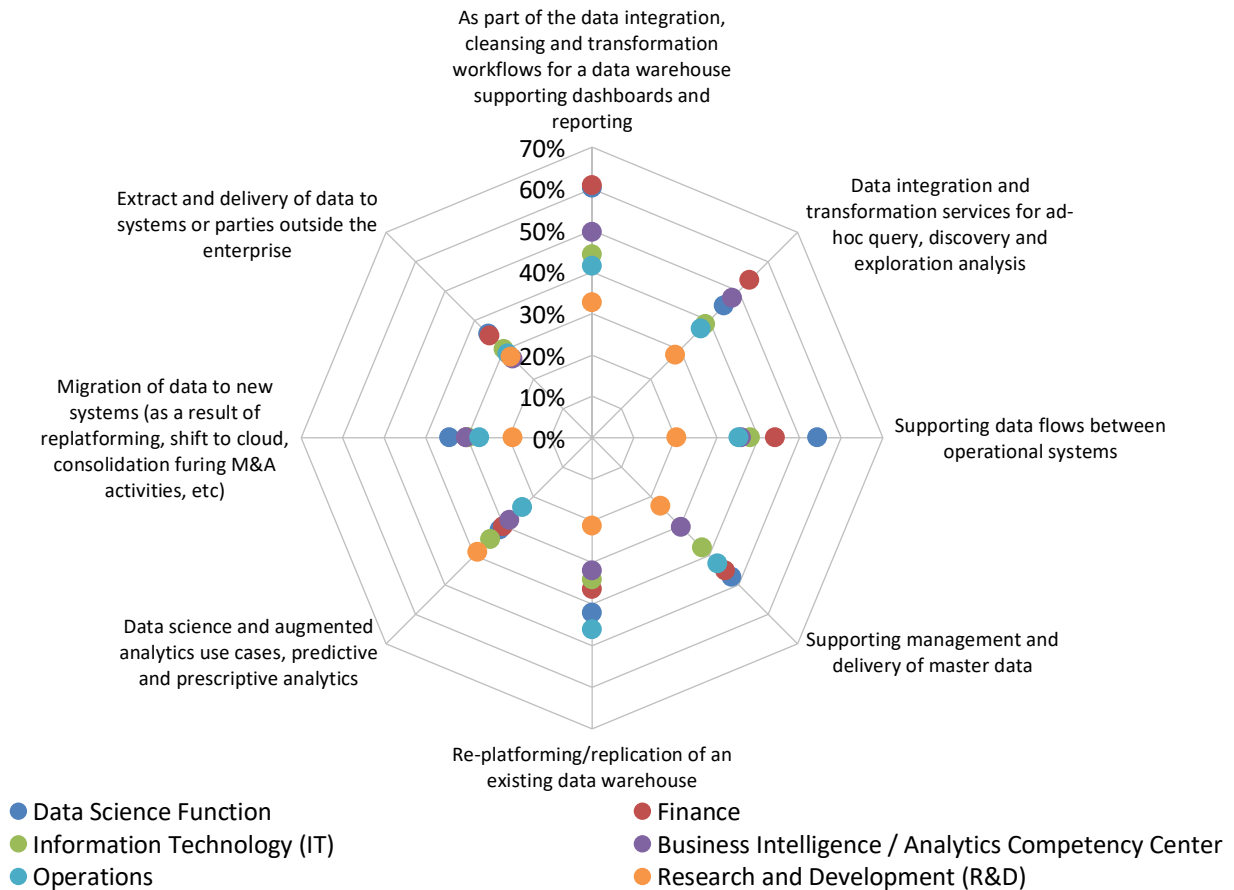


Figure 23 – Data Engineering Use Cases by Function

ADI Capabilities Relevant for Data Engineering

Analytic data infrastructure spans a wide range of capabilities involving nearly all aspects of data management. We asked survey respondents to identify which ADI capabilities are most important to them in their data engineering work (fig. 24). Organizations rate development-oriented capabilities the highest, with 63% indicating graphical workflow building is important, 56% indicating debugging and monitoring is important, and 54% stating no-code and code-friendly transformations are important. Various integration, management and security capabilities were rated as important by around 50% of respondents. Relatively new to the ADI arena, AI-assisted capabilities for mapping and modeling, data quality and anomaly detection were rated important by around 40% or less of respondents.

Data Engineering Capabilities

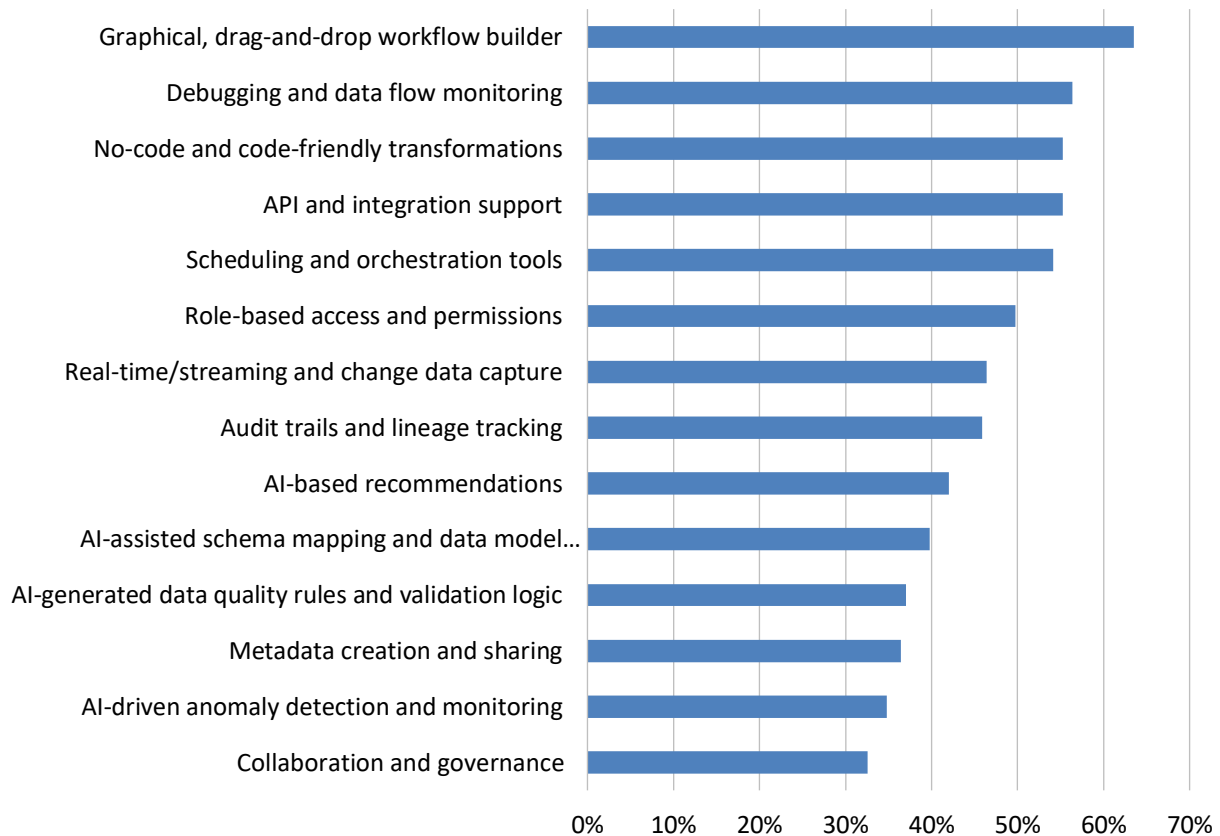


Figure 24 - Data Engineering Capabilities

Customer priorities for data engineering capabilities vary significantly across regions. Graphical workflow development and debugging/monitoring capabilities are important globally, according to 45% or more of customers in all regions (fig. 25). Regional differences emerge regarding all other capabilities. For example, 60% of Asia Pacific organizations view API and integration support as important, while only 47% of EMEA organizations and 27% of Latin American organizations do so. North American organizations rate scheduling and orchestration, role-based access, audit trails and AI-assisted schemas and data models higher than do all other regions, at about 60% frequency for each capability. Latin American organizations stand out for their low prioritization of most governance and AI-assisted capabilities (all at 20% or lower frequency).

Data Engineering Capabilities by Geography

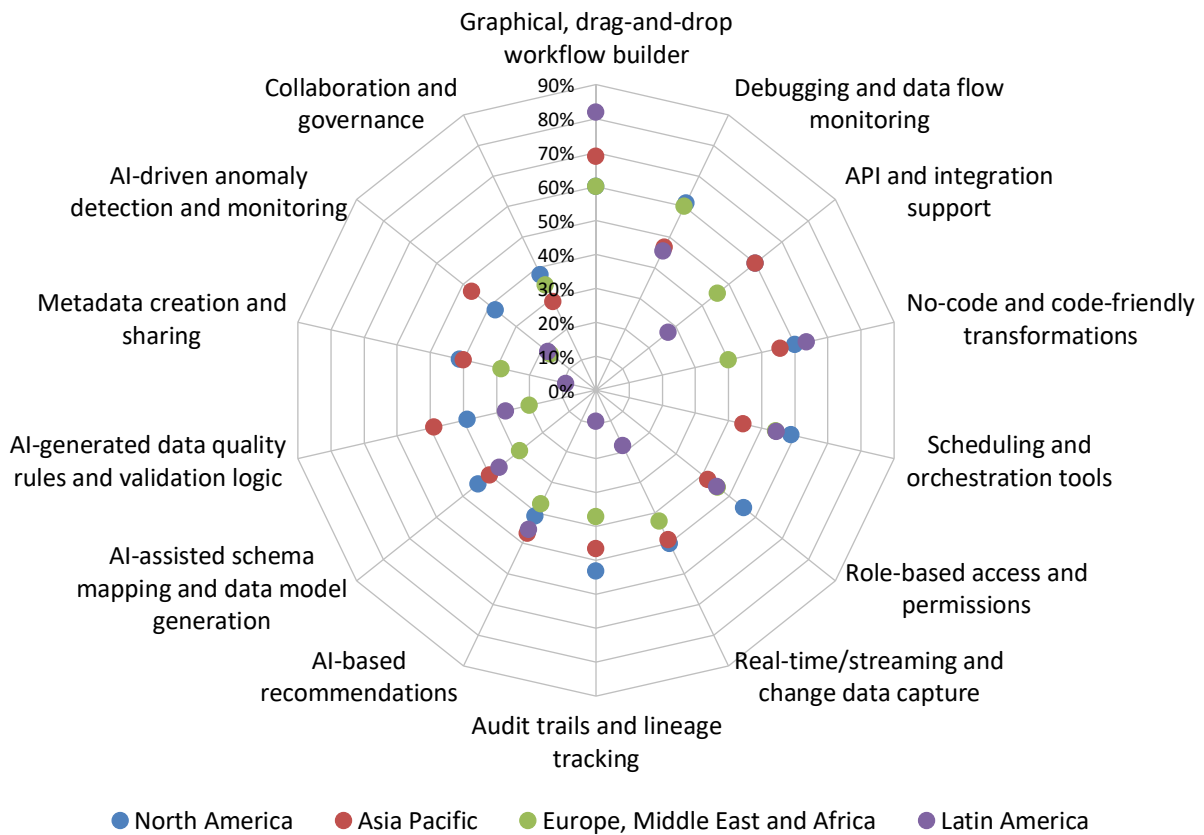
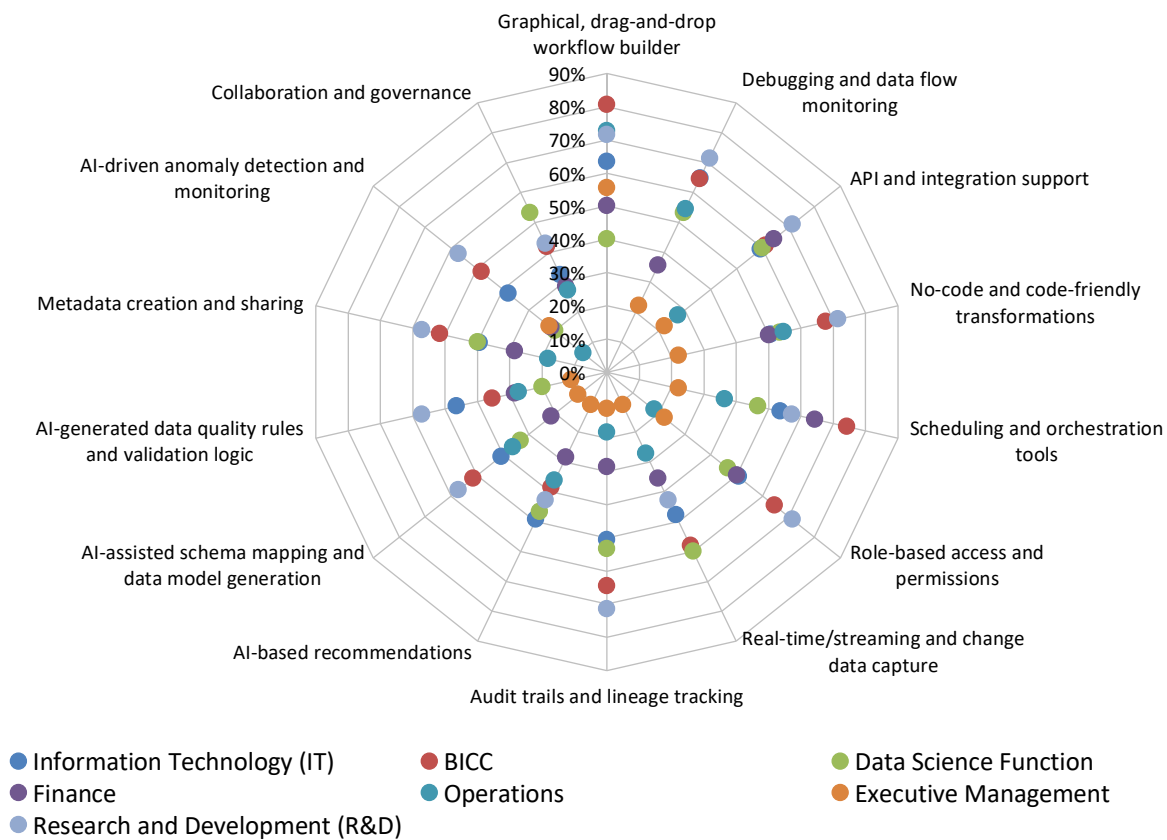


Figure 25 - Data Engineering Capabilities by Geography

A wider variation in perception of data engineering capability importance is apparent across business functions. Over 60% of BICCs consider the top five capabilities important (with over 80% of BICCs viewing graphical capabilities as important; fig. 26). Over 70% of R&D functions likewise rate these top capabilities as important. By contrast, the IT, data science, finance and operations functions see these capabilities as important somewhat less frequently, ranging from 40%-70% of the time. Executive management is least likely to rate all data engineering capabilities as important (with the exception of graphical development), doing so at 25% or lower frequency.

Data Engineering Capabilities by Function



The level of importance organizations place on data engineering capabilities is somewhat linked to their degree of success with BI. For example, organizations that are extremely successful, very successful or moderately successful with BI indicate graphical development capabilities are important at 57%, 63% and 62% frequency, respectively. Only 33% of those that are somewhat unsuccessful or unsuccessful with BI do so (fig. 26). This same pattern—albeit at different frequency levels—holds true for many of the other capabilities, including API and integration support, scheduling and orchestration, AI-generated data quality, and AI-driven anomaly detection. The data also shows a correlation: for example, extremely successful organizations more frequently place greater importance on metadata, AI-based recommendations, and audit trails and lineage, than do those in other success categories.

Data Engineering Capabilities by BI Success

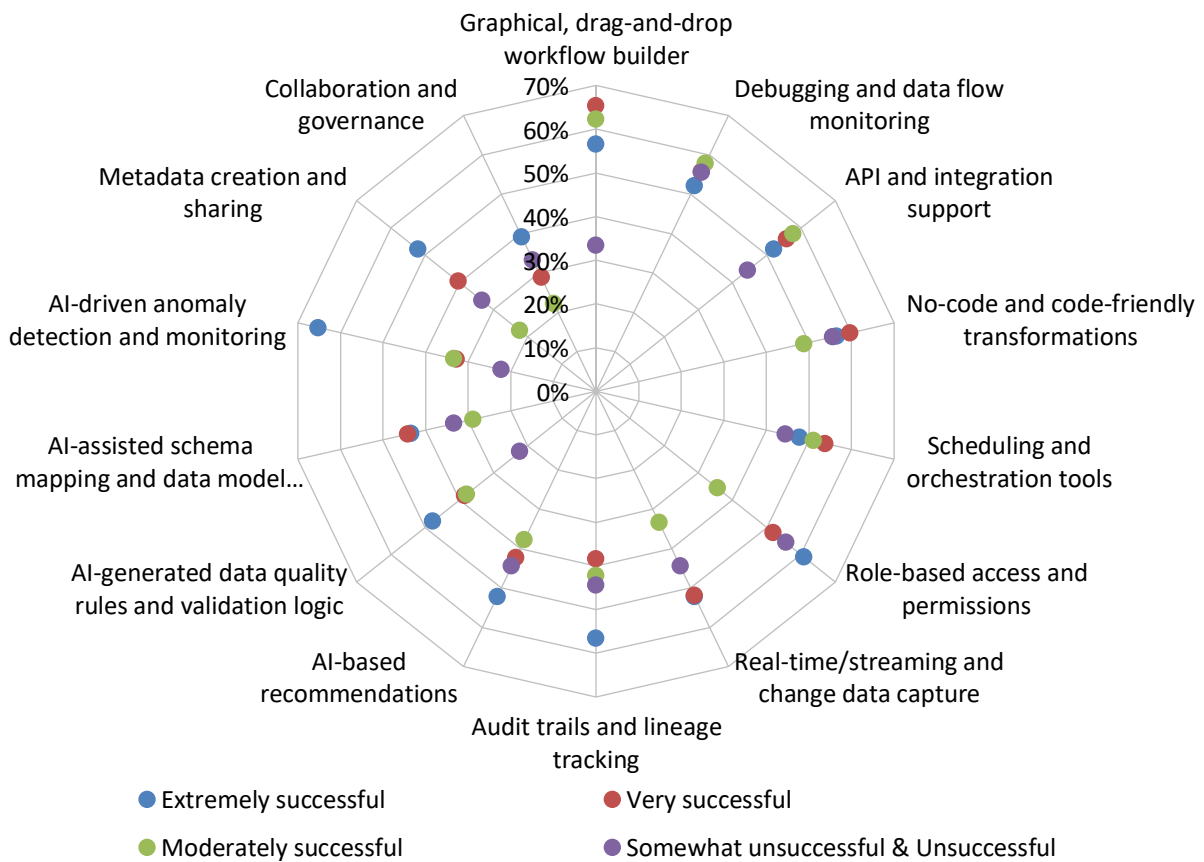


Figure 26 - Data Engineering Capabilities by BI Success

Data Engineering Features for Processing and Transforming Data

When assessing the relative importance of data engineering features for processing and transforming data, most organizations prioritize foundational features. Features for aggregation, workflow, job monitoring and scheduling, reshaping data and deriving new data are rated as critical by more than 40% of respondents, and at least very important by 75% or more (fig. 27). All features in this very long list (except for support for Apache and Kafka) are considered at least important by over 60% of respondents. This speaks to the declining demand for big data-specific technologies.

Data Engineering Features for Processing and Transforming Data

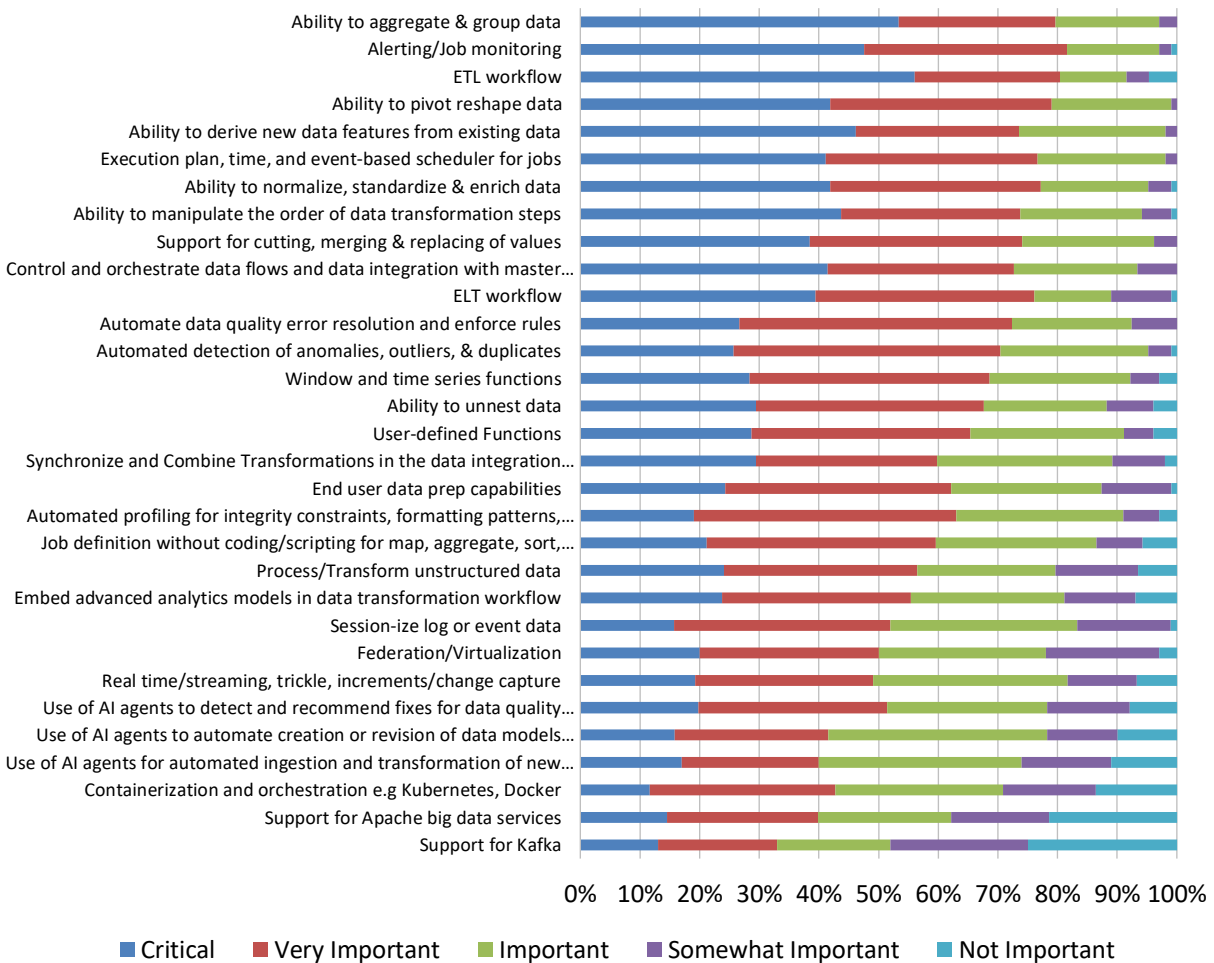


Figure 27 – Data Engineering Features for Processing and Transforming Data

Through 2023-2026, customer prioritization of features for processing and transforming data has not shifted significantly, aside from the continued rise in priority of all related data engineering features (fig. 28). Alerting/job monitoring saw a large increase in priority, driving this capability much higher in 2026 than in 2025, reversing the downward trend from 2023. Other capabilities prioritized higher in 2026 include pivoting and reshaping, ability to manipulate the order of data transformation steps, and automating data quality (also reversing a slight downward trend in prior years).

Data Engineering Features for Processing and Transforming Data - 2023-2026

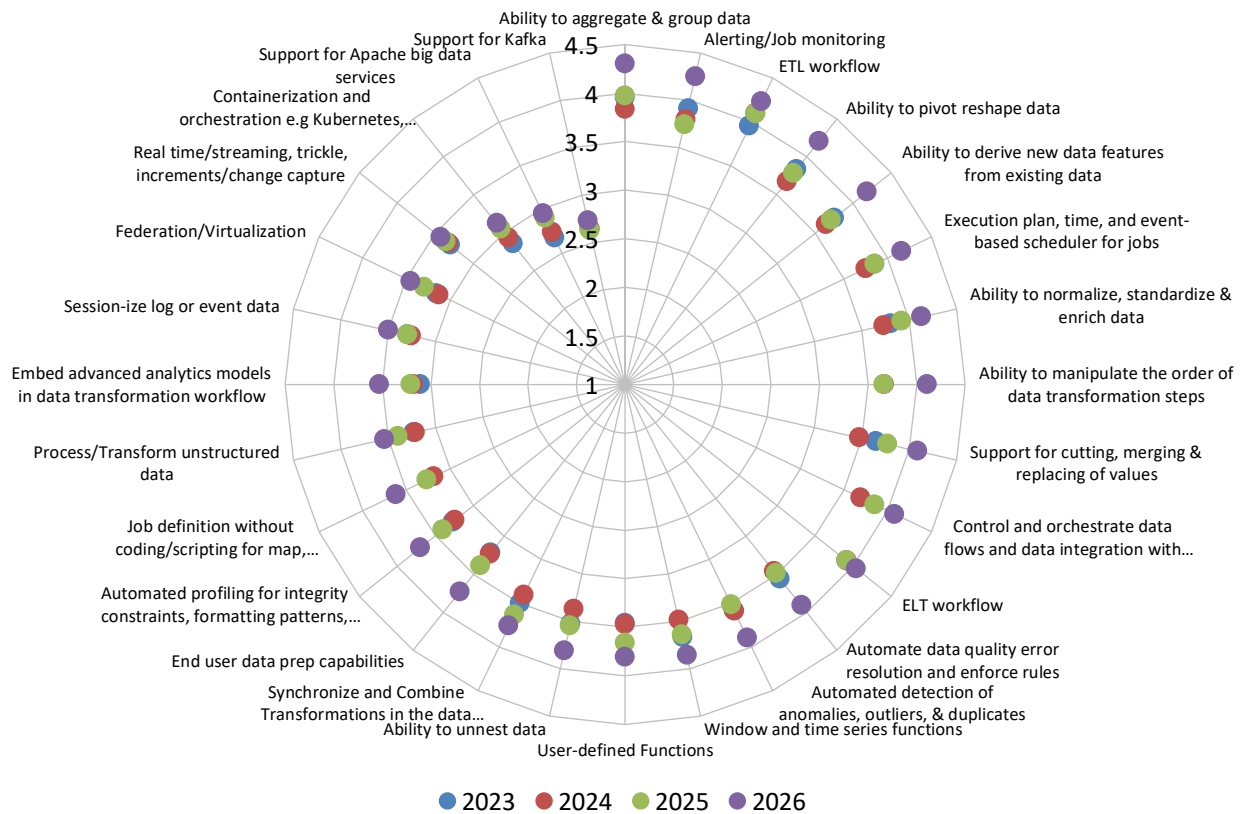


Figure 28 – Data Engineering Features for Processing and Transforming Data - 2023-2026

An organization’s level of BI success correlates strongly with the priority it places on data engineering features for processing and transforming data. Those organizations rating themselves as extremely successful with BI also assigned a higher priority than their peers with lower BI success levels to virtually all features in this class (fig. 29). Exceptions to this trend include aggregation and grouping, ETL workflow, automating data quality, and automated detection of anomalies. However, extremely successful organizations had priorities nearly as high in these cases—higher than the other success segments.

Data Engineering Features for Processing and Transforming Data by BI Success

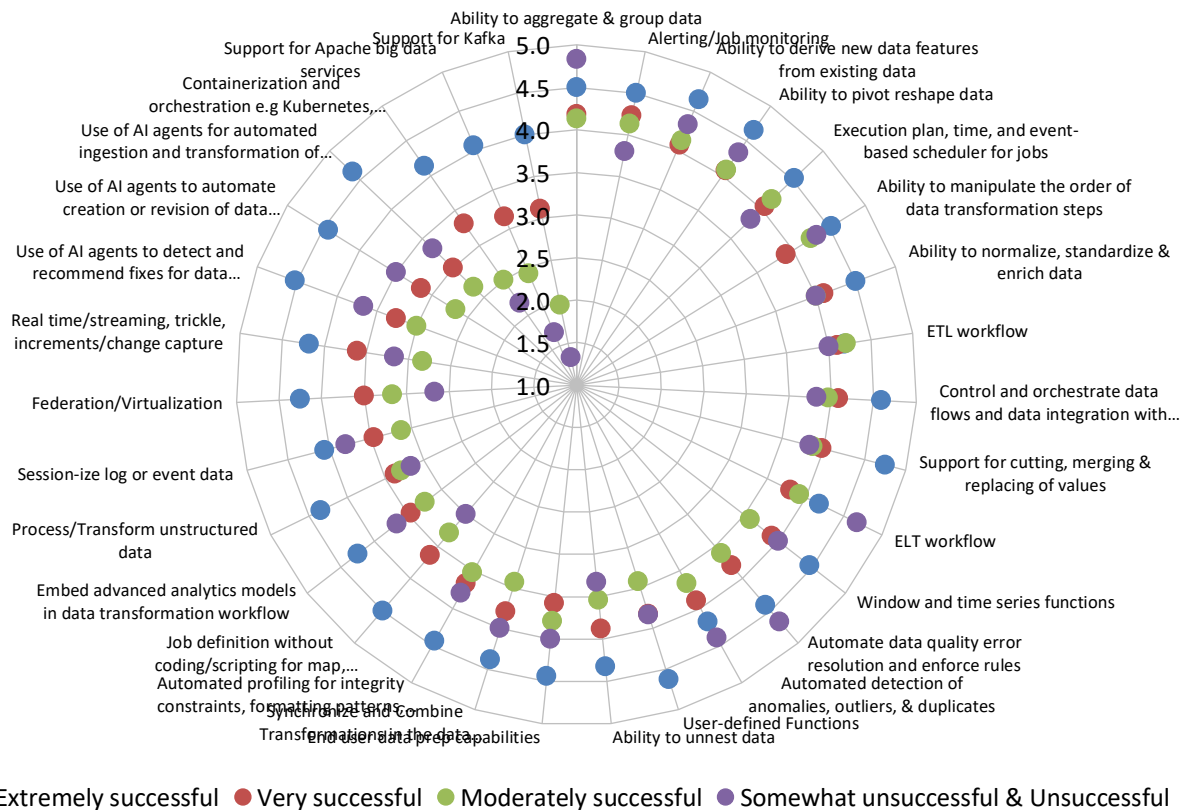


Figure 29 – Data Engineering Features for Processing and Transforming Data by BI Success

ADI Management Features Relevant for Data Engineering

ADI capabilities also include management features that are important for customers doing data engineering work. Customers indicate varying perceptions of importance for such features (fig. 30). They viewed security as the highest priority, with 72% indicating it is important for data engineering. Fifty-four percent rate quality assurance as important, with a nearly equal amount (at 53%) identifying cost monitoring and optimization as important. All other features are viewed as important by a minority of respondents, with lifecycle management highest in this group at 47%.

Data Engineering Management Features

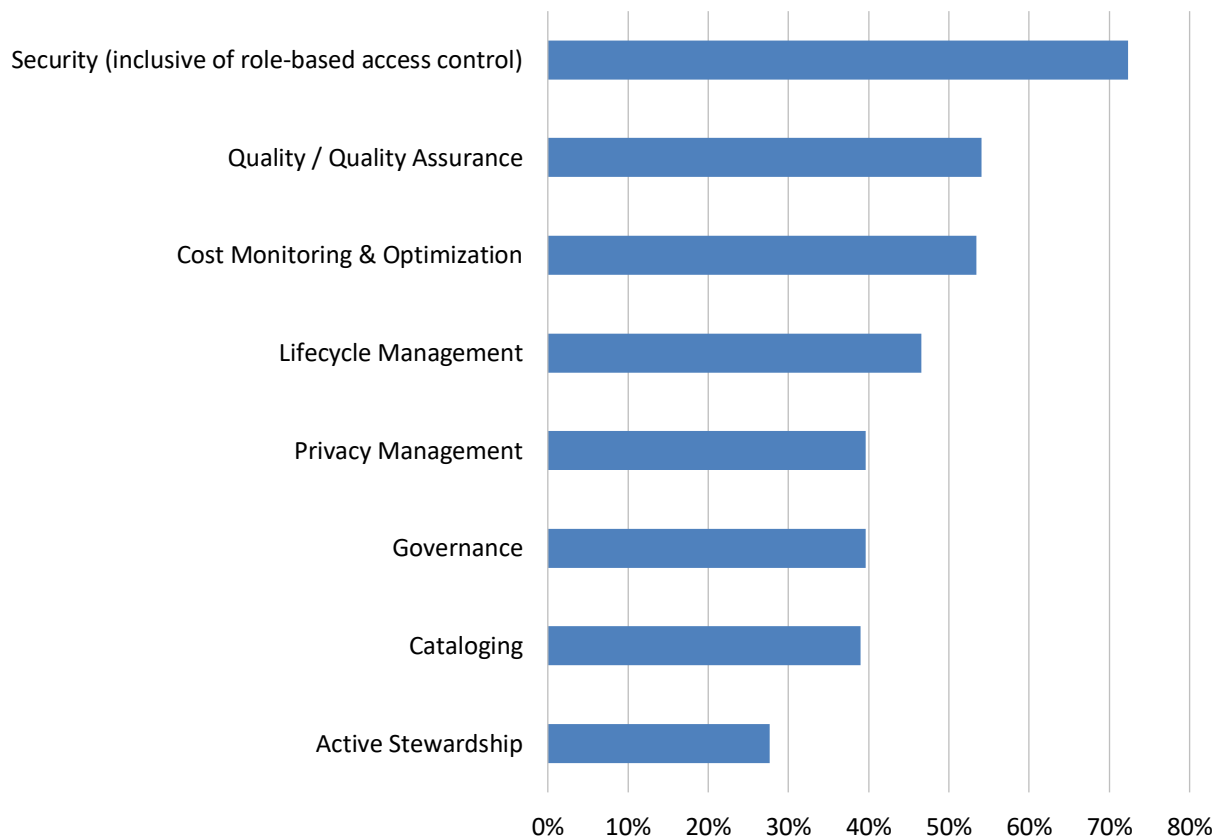


Figure 30 – Data Engineering Management Features

Customers across different regions view the importance of data engineering management features similarly (fig. 31). All regions rate security as important with very high frequency (70%-77%). Other features like quality assurance are perceived differently across the regions, with 60% of EMEA organizations considering it important but only 30% of Latin American organizations saying the same. Cost monitoring and optimization reflects a similar range; however, Latin American organizations most often view this feature as important (70%) while Asia Pacific organizations are least likely to do so (44%). Regions are fairly well aligned on the importance of all other management features, with data showing a much narrower spread.

Data Engineering Management Features by Geography

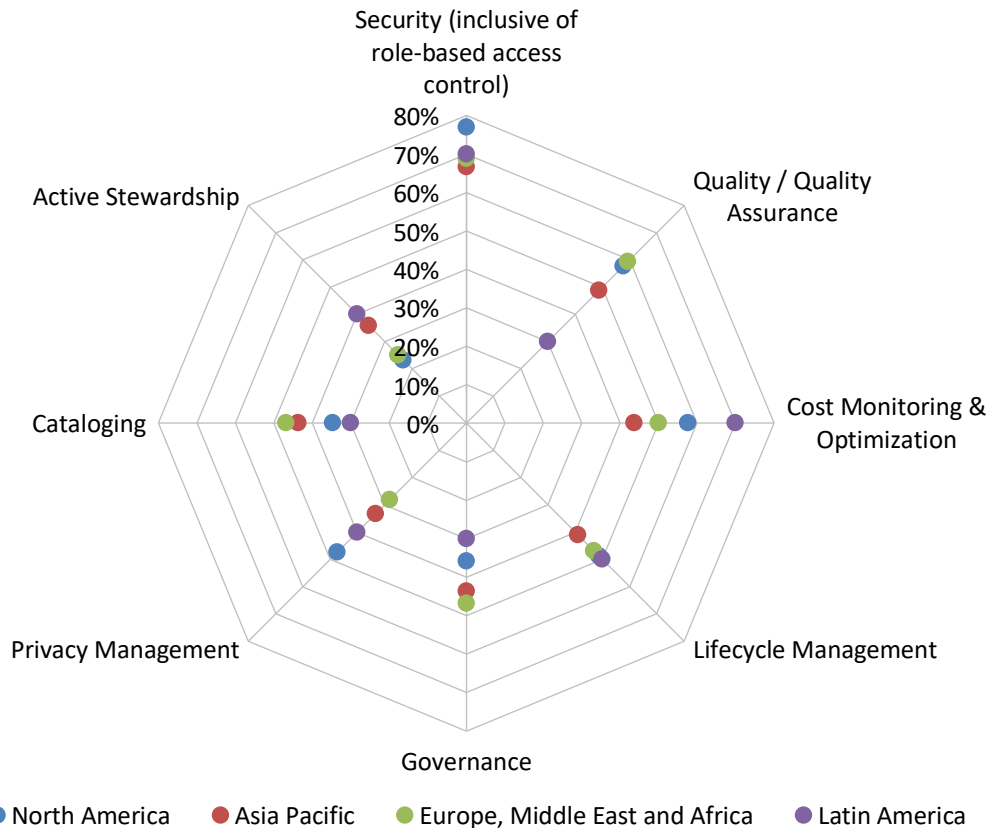


Figure 31 – Data Engineering Management Features by Geography

Looking at data engineering management feature importance by industry reveals more significant variations (fig. 32). Manufacturing organizations are more likely than other sectors to perceive management features as important. This applies to most features in this category, with more than 57% of manufacturing organizations indicating importance for all except active stewardship. Healthcare and consumer services stand out for their frequency in rating many features as important, while the opposite is true for consumer services, which was the least likely sector to indicate importance for several features.

Data Engineering Management Features by Industry

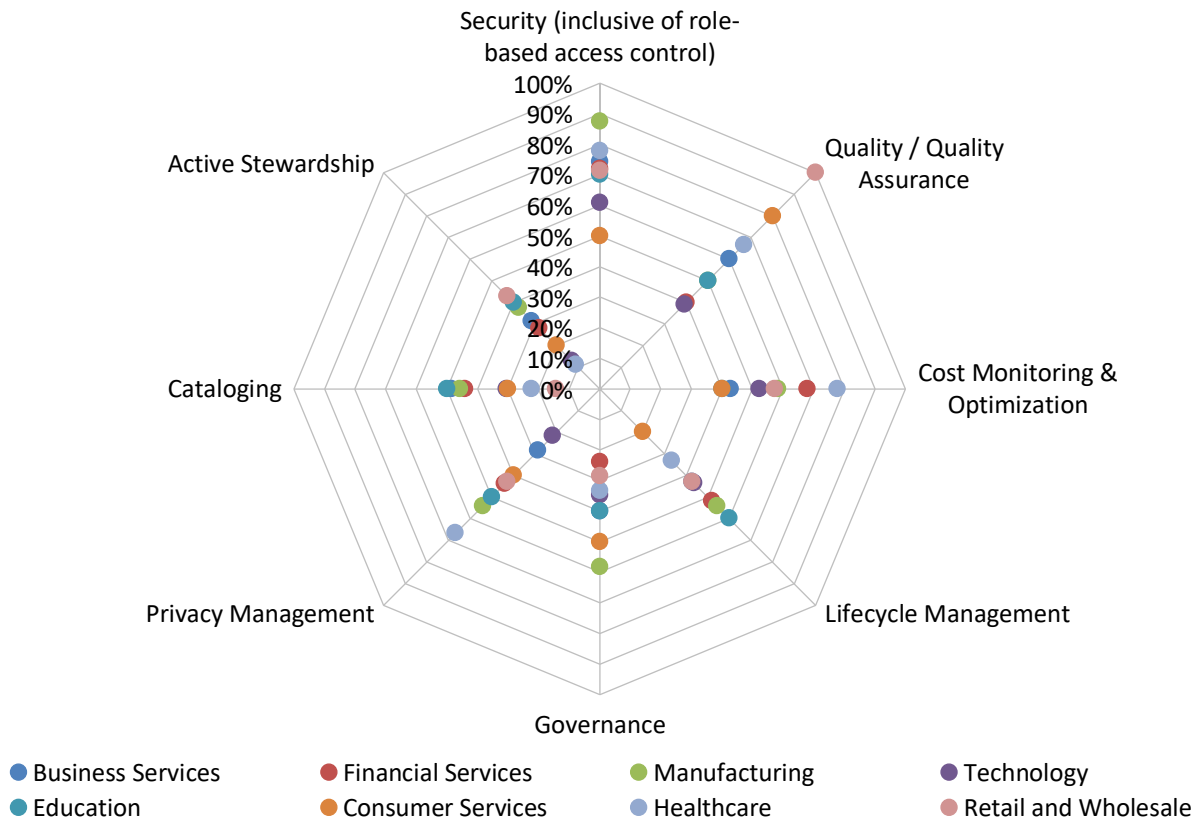


Figure 32 - Data Engineering Management Features by Industry

Organizations with more BI success are more likely to consider management features important to their data engineering work, though this is not the case for all such features (fig. 33). For example, 44% of extremely successful organizations consider privacy management important, while only 29% of moderately successful organizations do. Forty-four percent of extremely successful organizations also consider governance important, while only 33% of somewhat unsuccessful or unsuccessful organizations do. Similarly, 39% of extremely successful organizations see active stewardship as important compared to 11% of somewhat unsuccessful or unsuccessful organizations.

Data Engineering Management Features by BI Success

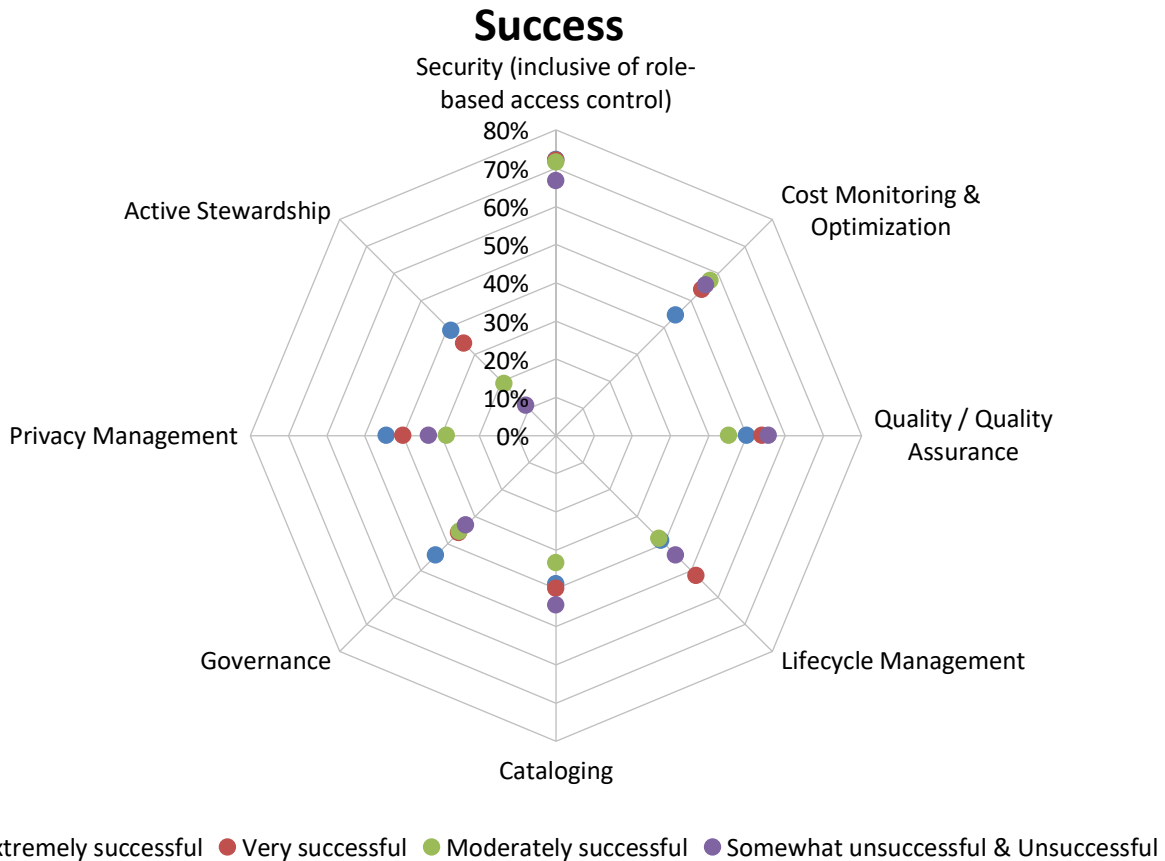


Figure 33 - Data Engineering Management Features by BI Success

Data Sources and Targets for Data Engineering

Most organizations support a large, complex landscape of systems and data platforms and wish to leverage an increasingly diverse set of data types as part of data pipelines built with data engineering technology and processes. These data sources and targets may be internal or external. Unsurprisingly, organizations generally indicate they have larger numbers of internal than external data sources (fig. 34). Fifty-one percent of those surveyed stated they work with 41 or more internal sources, and 14% said they have 81-100 internal sources, while only 44% reported having 41 or more external sources and 11% reported 81-100 external sources.

Number of Data Engineering Source Systems

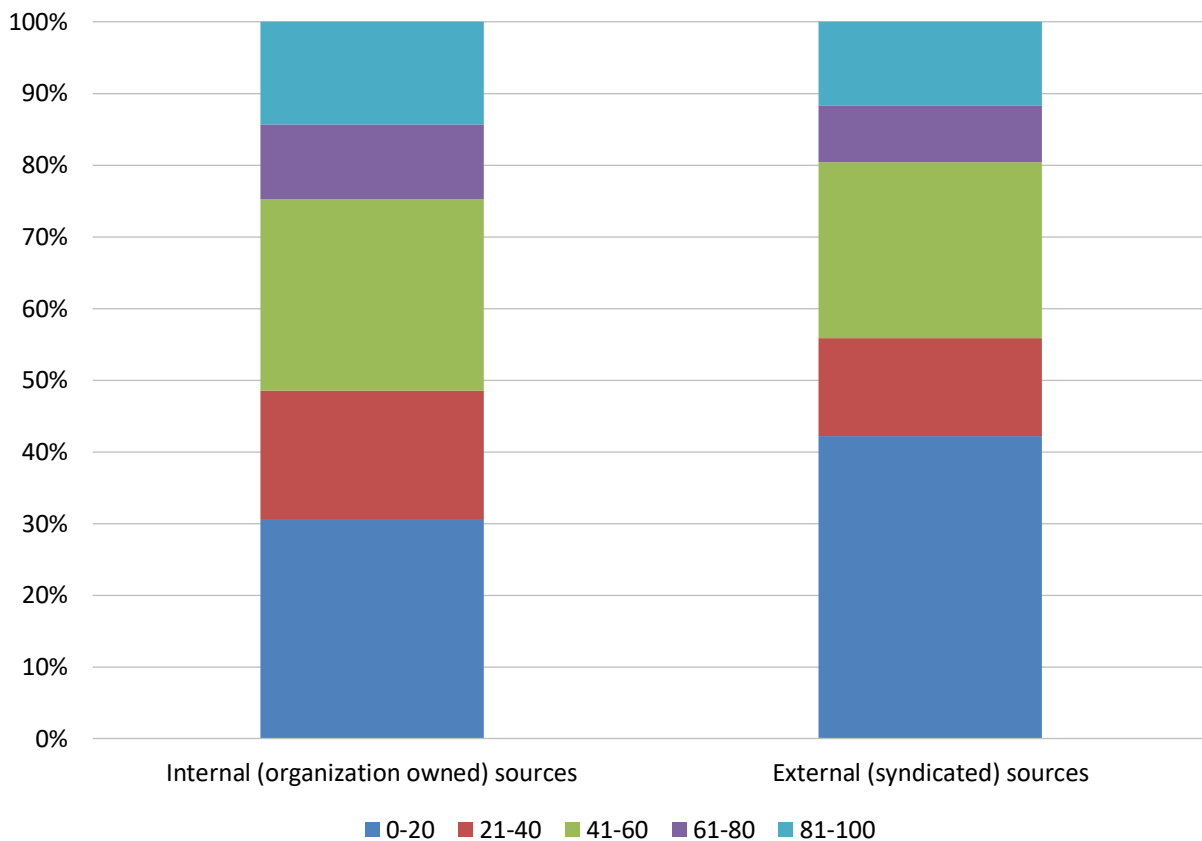


Figure 34 – Number of Data Engineering Source Systems

The pattern of larger numbers of internal source systems than external ones holds true across regions—except for in Latin America (fig. 35). Asia Pacific organizations on average deal with the greatest numbers of both internal (52) and external (50) sources. Latin American organizations average in the upper 30s for total sources, split between internal (36) and external (39). EMEA shows the largest disparity in source locations, with internal sources (44) outpacing external (28). North America represents a rather typical balance between internal (39) and external (31).

Average Data Engineering Source Systems by Geography

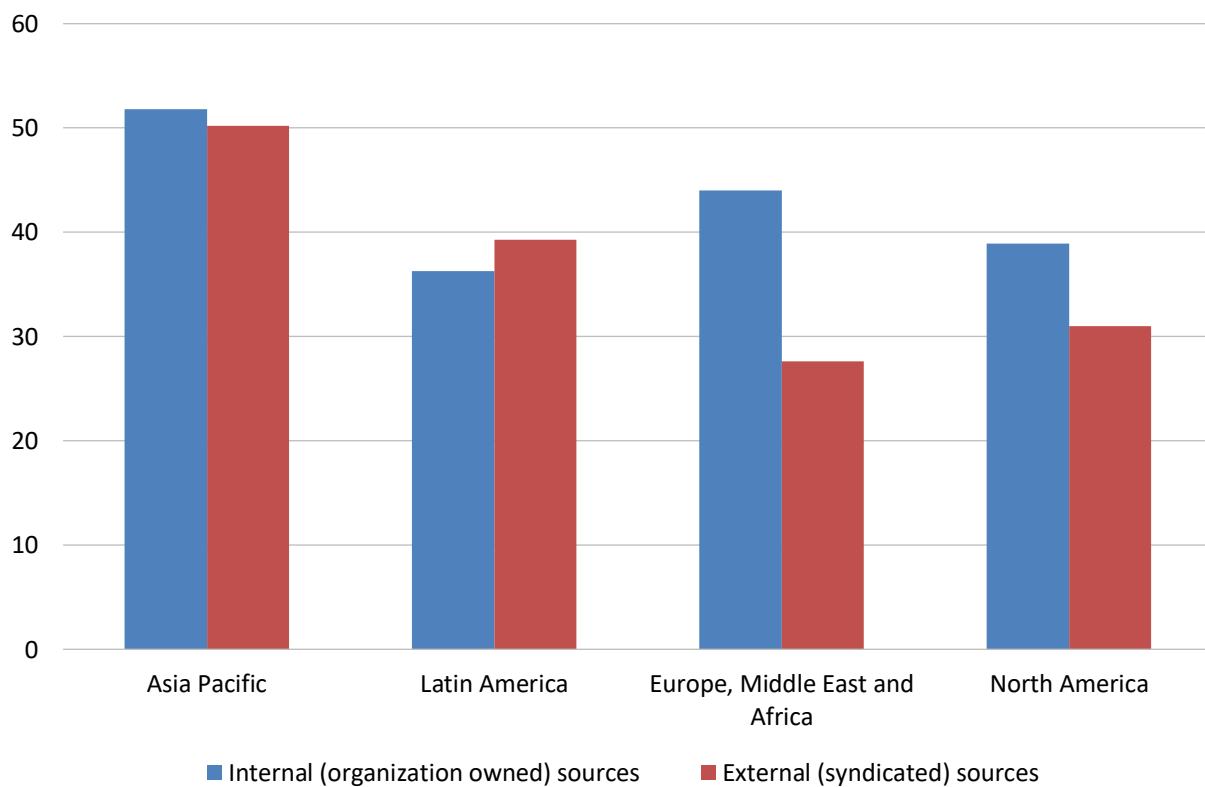


Figure 35 – Average Data Engineering Source Systems by Geography

We observe a nearly equal split between internal and external sources in many industry sectors, but significant variability in the overall numbers of sources involved in data engineering (fig. 36). Education handles the largest number of sources overall on average, split between 59 internal and 62 external—notably, the only sector reporting a larger average for external sources. Manufacturing also stands out for large numbers of sources, with an average of 53 internal and 50 external. The magnitude of total sources is similar across other sectors, although technology and business services have a much more equal balance of internal and external that do financial services and healthcare (where the number of internal sources significantly eclipses external).

Average Data Engineering Source Systems by Industry

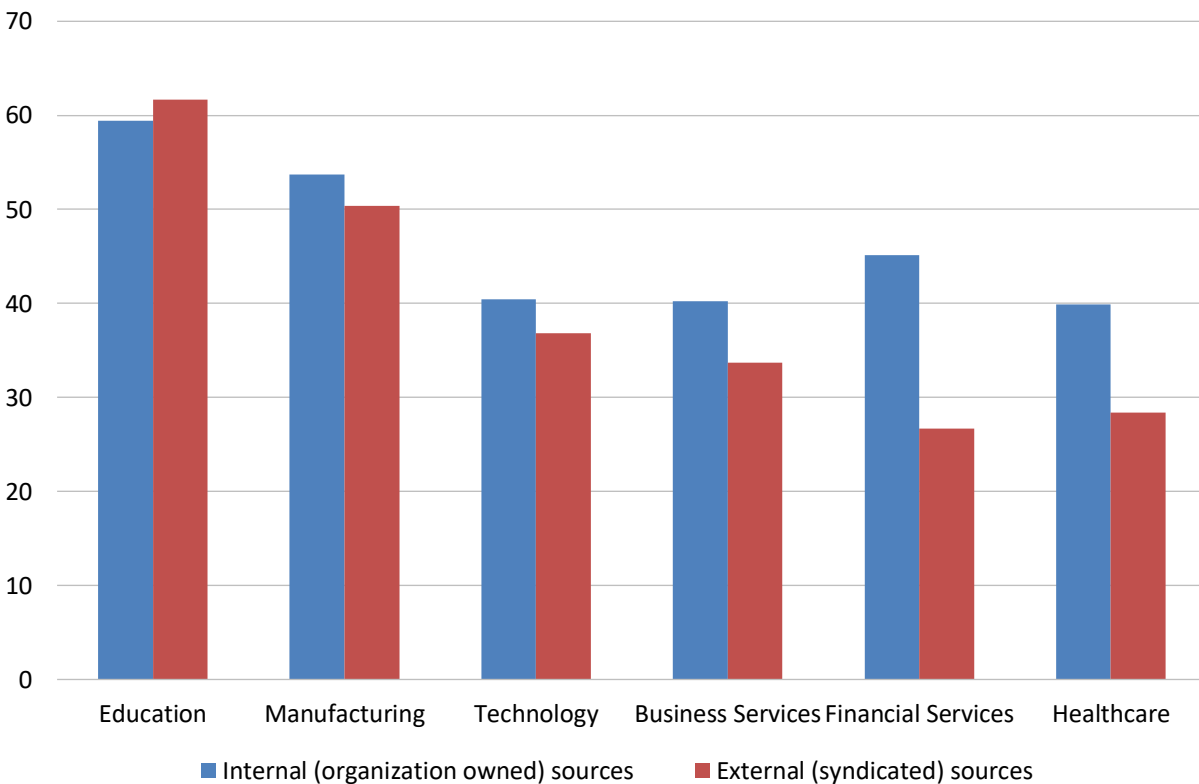


Figure 36 – Average Data Engineering Source Systems by Industry

As expected, data engineering processes involve more sources as organization size increases (fig. 37). Small organizations with 1-100 employees work with an average of 23 internal sources and 23 external sources and are the only organization size range reporting more external than internal. Medium-sized organizations with 101-1,000 employees average around 47 internal sources and 37 external sources in their data engineering work. Larger organizations with 1,001-10,000 employees report a similar total magnitude, but a nearly equal balance of internal and external sources, averaging about 40 for both. And the largest organizations with over 10,000 employees use the largest average number of internal sources, at 53, while also typically engaging with 40 external sources.

Average Data Engineering Source Systems by Organization Size

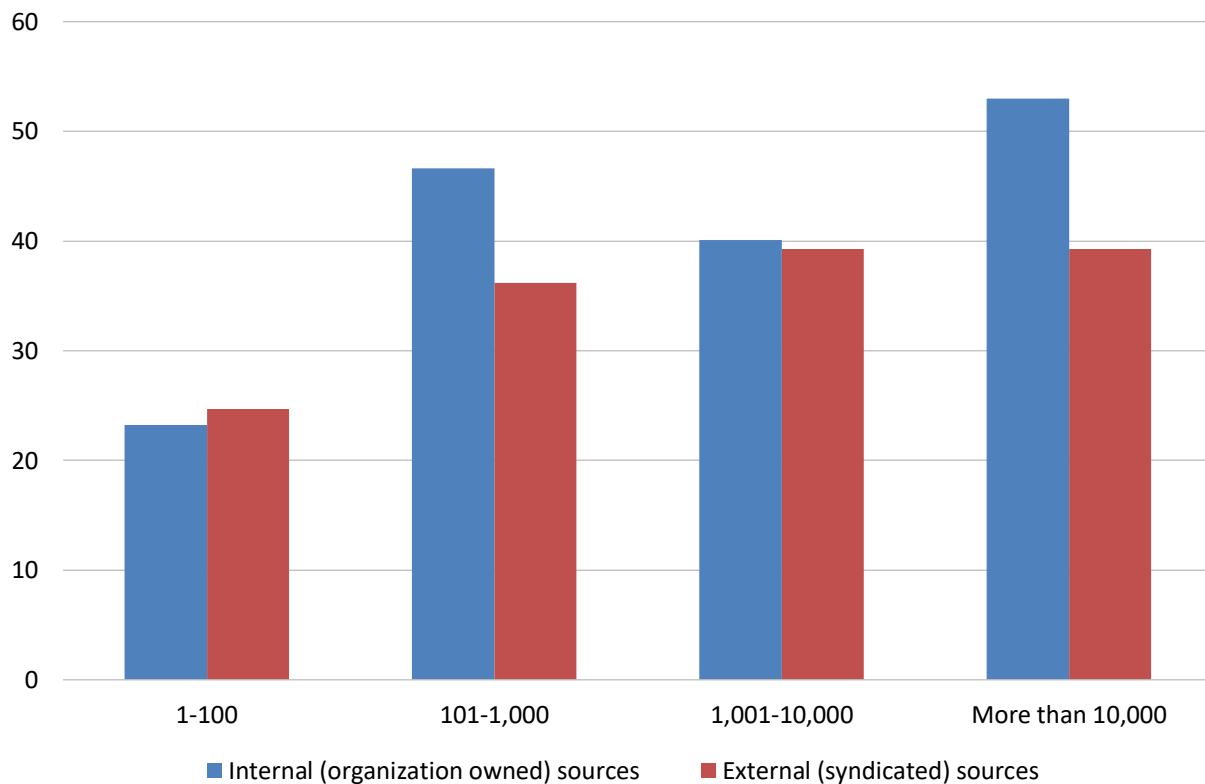


Figure 37 - Average Data Engineering Source Systems by Organization Size

Modern BI and analytics use cases supporting strategic business outcomes increasingly require innovative, complex and diverse combinations of data. In line with this trend, we see a clear correlation between the number of data sources involved in data engineering efforts and the degree of an organization’s success with BI—the more data sources involved, the higher the success level (fig. 38). Organizations that state they are extremely successful with BI report an average of 58 internal and 59 external sources in their data engineering work, far more than their less-successful peers. Very successful organizations report 42 internal and 38 external sources. Moderately successful organizations tap 45 internal sources and 30 external sources. Those organizations stating they are somewhat unsuccessful or unsuccessful consume the smallest numbers of both internal and external sources, at 34 and 27 respectively.

Average Data Engineering Source Systems by BI Success

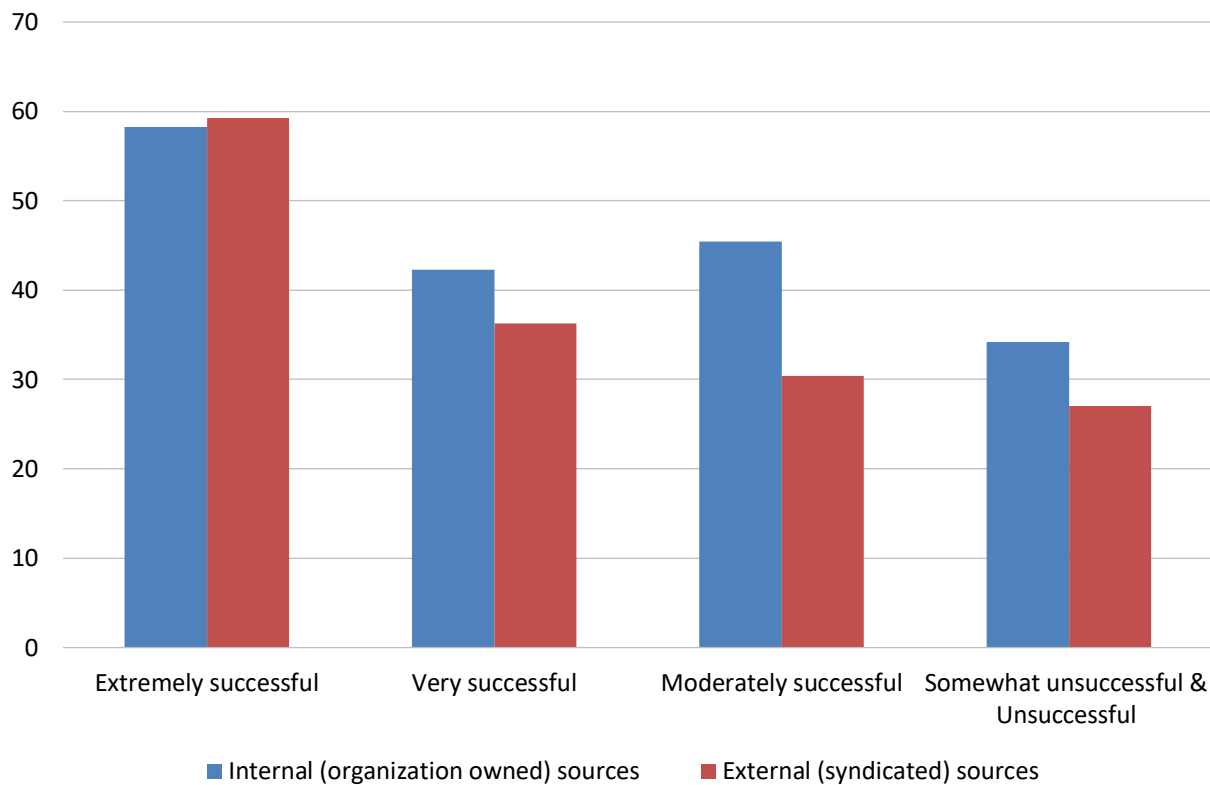


Figure 38 - Average Data Engineering Source Systems by BI Success

ADI Integration Techniques Relevant for Data Engineering

ADI components can be integrated with each other and with non-ADI systems and technologies through a wide variety of techniques. Our respondents deem some techniques more important for their data engineering work than others (fig. 39). API-based integration tops their priorities, with 71% of respondents rating it as important, followed by ETL/ELT at 66% and SQL query interfaces at 63%. Workflow orchestration is the only other technique viewed as important by most respondents, at 55%. All other techniques were viewed as important by 50% or less of our sample.

Data Engineering Integration Techniques

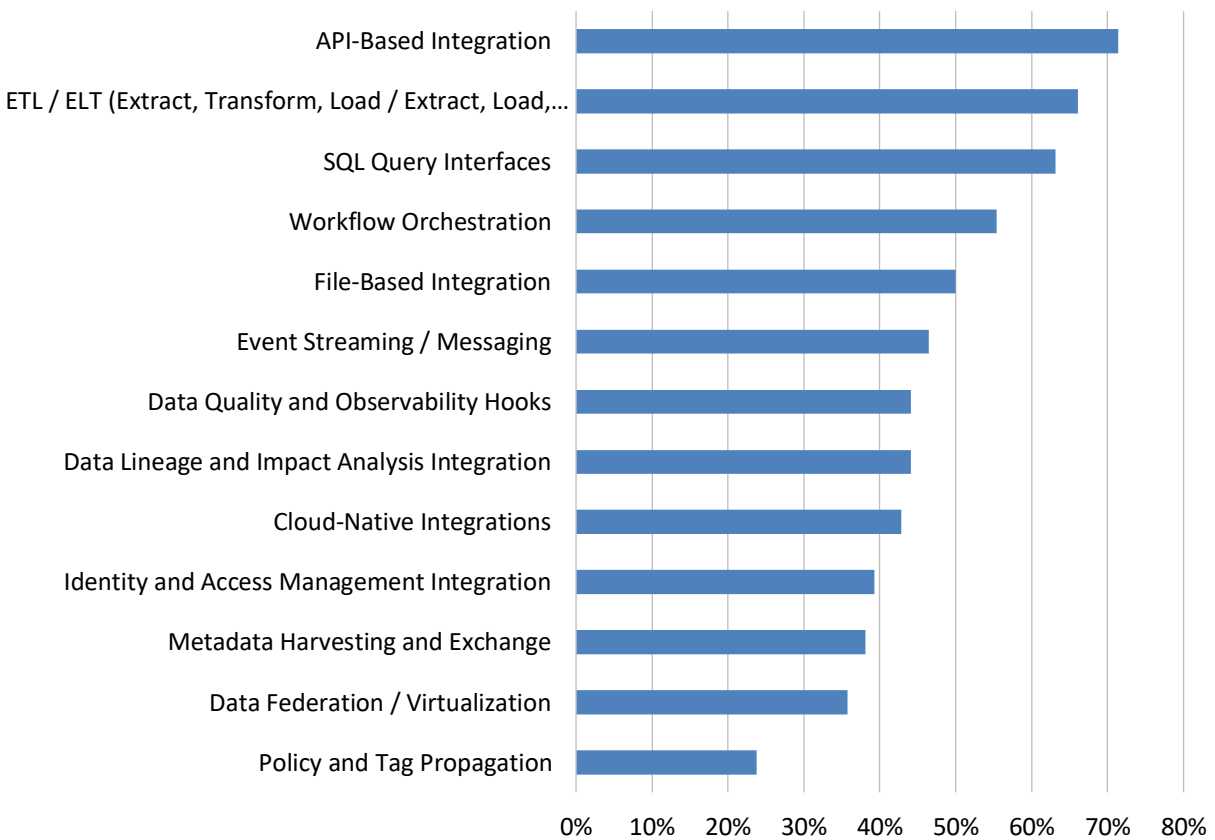


Figure 39 - Data Engineering Integration Techniques

Preferences for data engineering integration techniques vary across regions. While API-based integration is deemed important by a majority of organizations in all regions, North America does so with a much higher frequency (83%) than other regions (fig. 40). The same is generally true for ETL/ELT and cloud-native integration approaches. Asia Pacific organizations more often rate modern integration techniques as important, doing so more frequently than any other region for data virtualization (48%), metadata harvesting (50%) and event streaming (55%). EMEA organizations show a balanced preference across most integration techniques, rating the majority as important 30%-50% of the time.

Data Engineering Integration Techniques by Geography

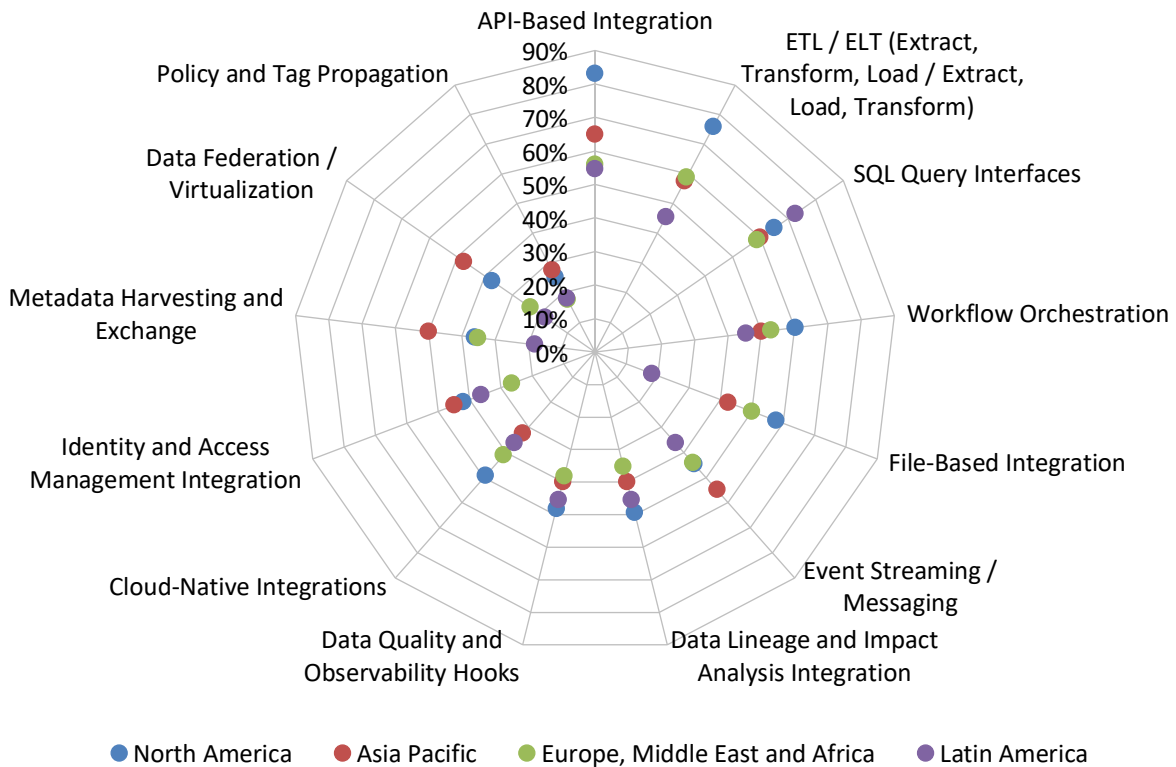


Figure 40 - Data Engineering Integration Techniques by Geography

Data engineering integration preferences vary significantly across industries. The priorities of business services, healthcare, financial and consumer services organizations generally align with the overall market, indicating API-based, ETL/ELT, SQL interfaces and workflow orchestration techniques are “important” with over 50% frequency (fig. 41). In contrast, sectors such as retail/wholesale and education place a higher importance on data quality and observability hooks, identity and access management, metadata harvesting and data virtualization than do their peers in other industries. This year’s sample also shows technology organizations prioritize data integration techniques such as ETL/ELT and event streaming well above nearly all other integration techniques.

Data Engineering Integration Techniques by Industry

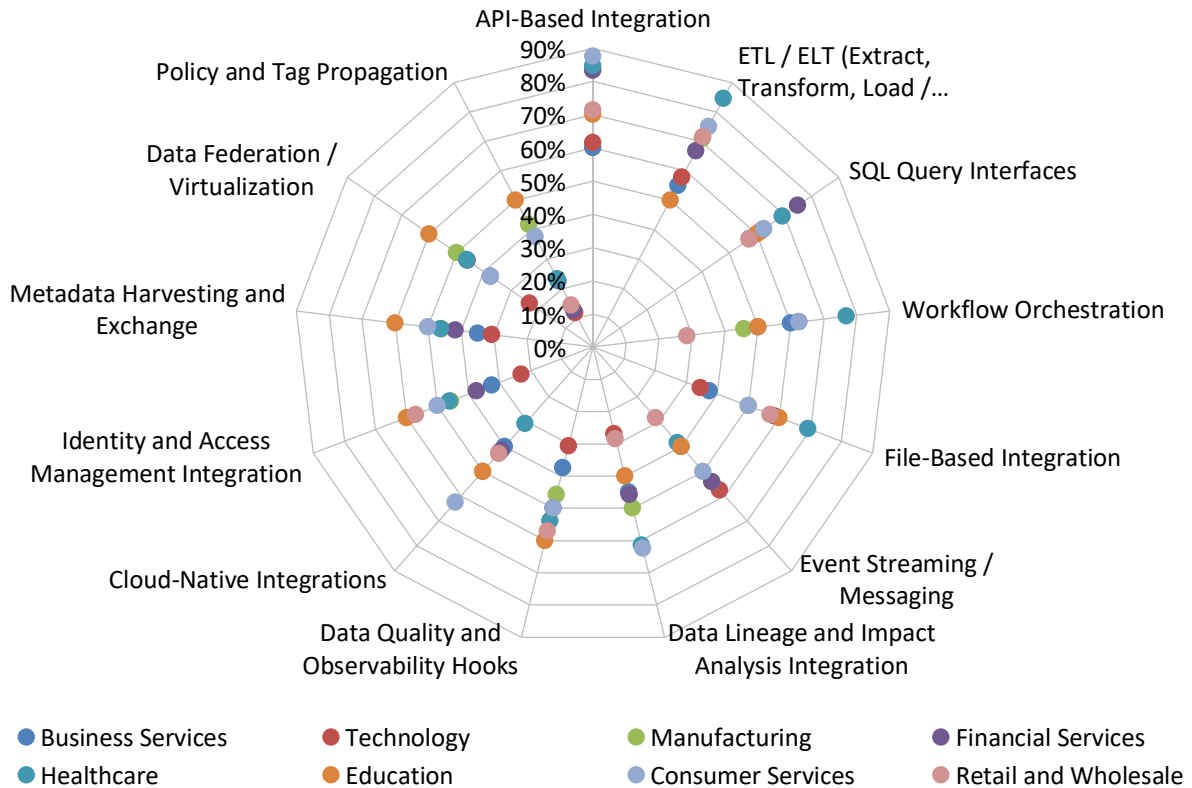


Figure 41 - Data Engineering Integration Techniques by Industry

Survey data shows that organizations with the greatest degree of BI success tend to employ a diverse set of integration techniques in their data engineering work (fig. 42). Those rating their BI efforts as extremely successful more frequently rate nearly all integration techniques as more important than do their peers, leading with API-based integration (80%), workflow orchestration (61%), event streaming (57%), data lineage (57%), metadata harvesting (57%) and data virtualization (48%). These extremely successful organizations perceive the more modern styles of integration as highly important and appear to be de-emphasizing older styles (ETL/ELT, SQL query, file-based) by less frequently rating these styles as important than do their peers.

Data Engineering Integration Techniques by BI Success

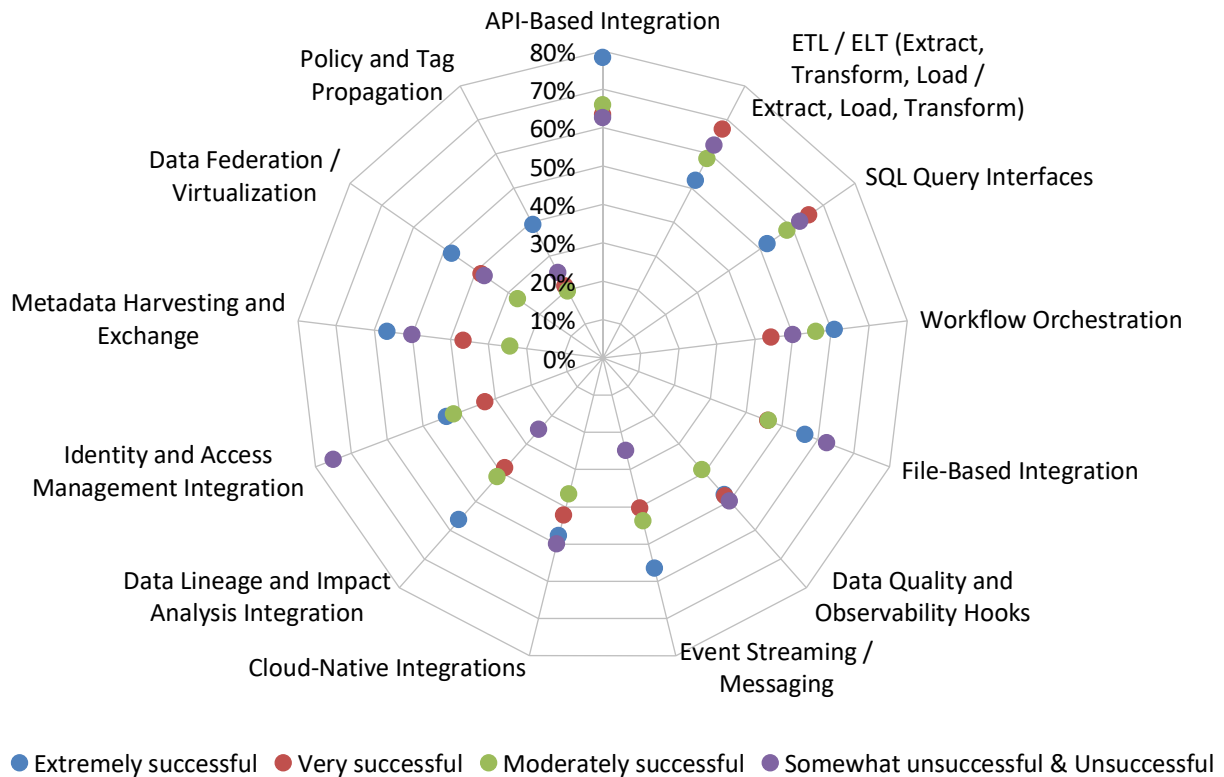


Figure 42 - Data Engineering Integration Techniques by BI Success

Deployment of Data Engineering Capabilities

Organizations implementing analytical data infrastructure components can choose from a range of deployment options. For data engineering capabilities within ADI, they continue to place a high priority on hybrid deployment, with 52% of respondents indicating it is important (fig. 43). They far less frequently, at 37%, cite private cloud, the second most popular deployment option, followed closely by public cloud at 36%. The remaining deployment options of multi-cloud, platform-independent and on-premises all hold minority positions, being rated as important by less than one third of respondents.

Data Engineering Deployment Options

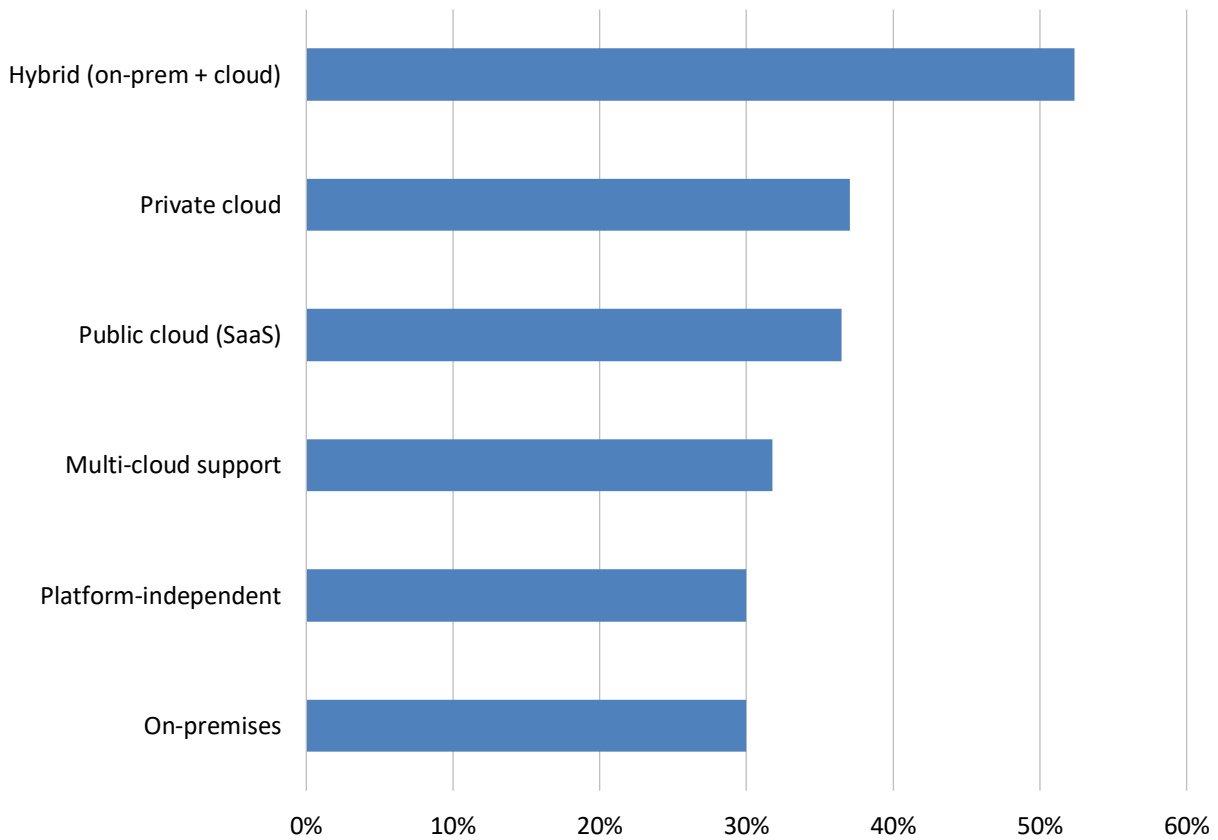


Figure 43 – Data Engineering Deployment Options

Preferences around deployment options for data engineering differ significantly across regions. Asia Pacific organizations more frequently consider a wider range of options as important compared to their peers, leading with hybrid at 80% and platform-independent at 50% (fig. 44). In contrast, EMEA organizations rate hybrid as important only 34% of the time and rate no other option as important with more than 45% frequency. North American organizations show a balanced view across most options. Hybrid is their top preference, with 47% rating it as important.

Data Engineering Deployment Options by Geography

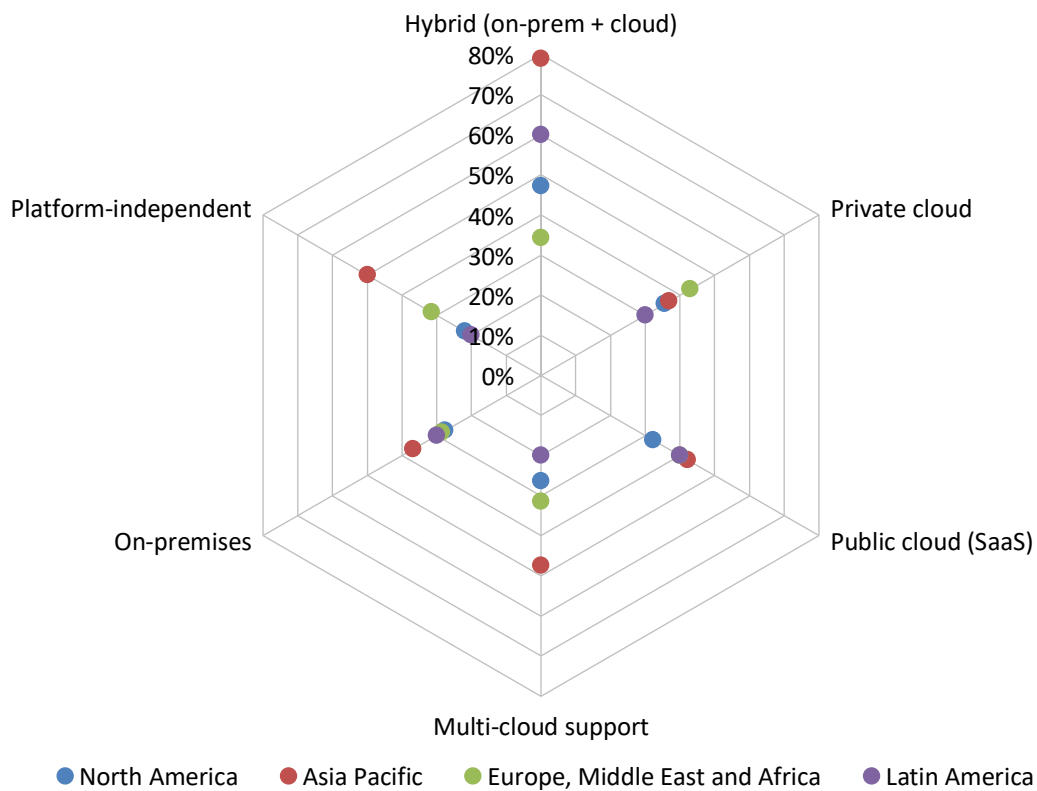


Figure 44 – Data Engineering Deployment Options by Geography

Preferences for data engineering deployment options are even more variable across industries, reflecting continued demand for diversity. The education sector strongly prioritizes hybrid and on-premises deployments, with 71% of organizations rating each as important (fig. 45). In contrast, healthcare organizations place most importance on public cloud deployment, rating it as important with 55% frequency. Financial services sees a wider range of options as important, led by hybrid at 68%, followed strongly by platform-independent at 52%, private cloud at 44%, and public cloud at 40%.

Data Engineering Deployment Options by Industry

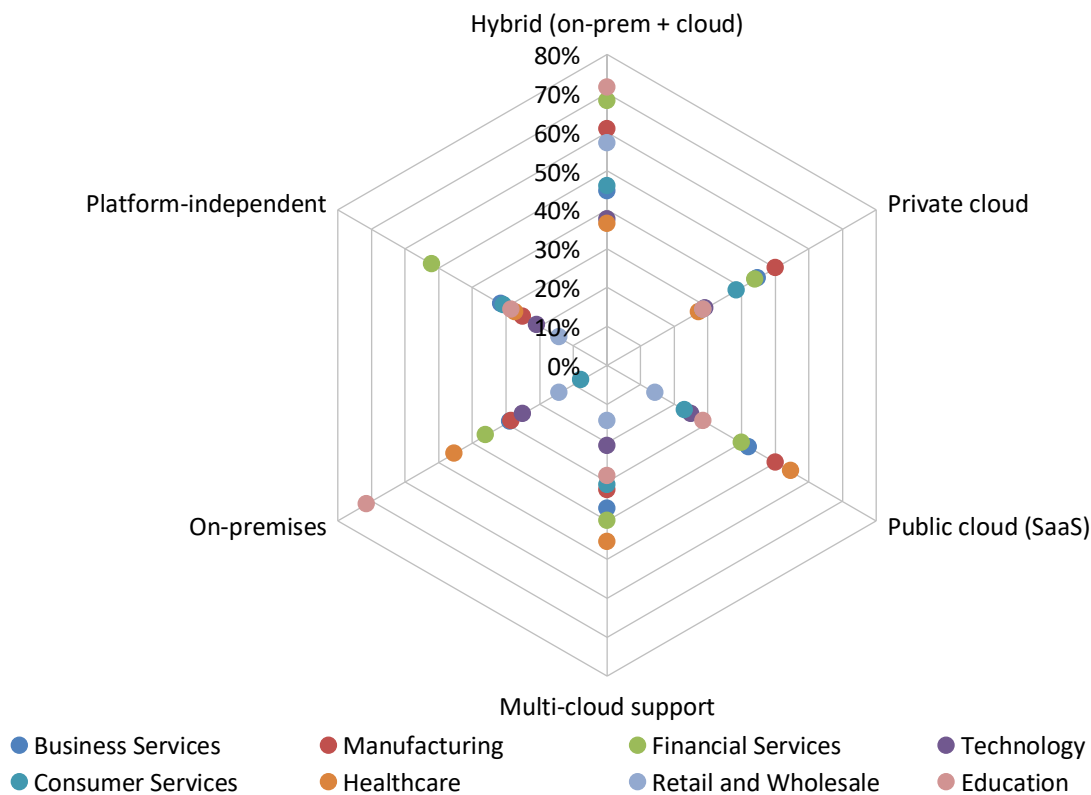


Figure 45 – Data Engineering Deployment Options by Industry

As with integration techniques, organizations that use a diverse range of deployment options for data engineering tend to also experience higher levels of success in their BI efforts (fig. 46). Those rating themselves as extremely successful with BI frequently assign importance to hybrid (60%), public cloud (45%), multi-cloud (41%), on-premises (45%) and platform independent (50%) deployment, highlighting the need for diverse styles. Their peers reporting lower levels of BI success tend to emphasize one or two deployment options far more than all others. For example, 56% of organizations that are somewhat unsuccessful or unsuccessful indicate hybrid is important but identify all other options as important with less than 40% frequency.

Data Engineering Deployment Options by BI Success

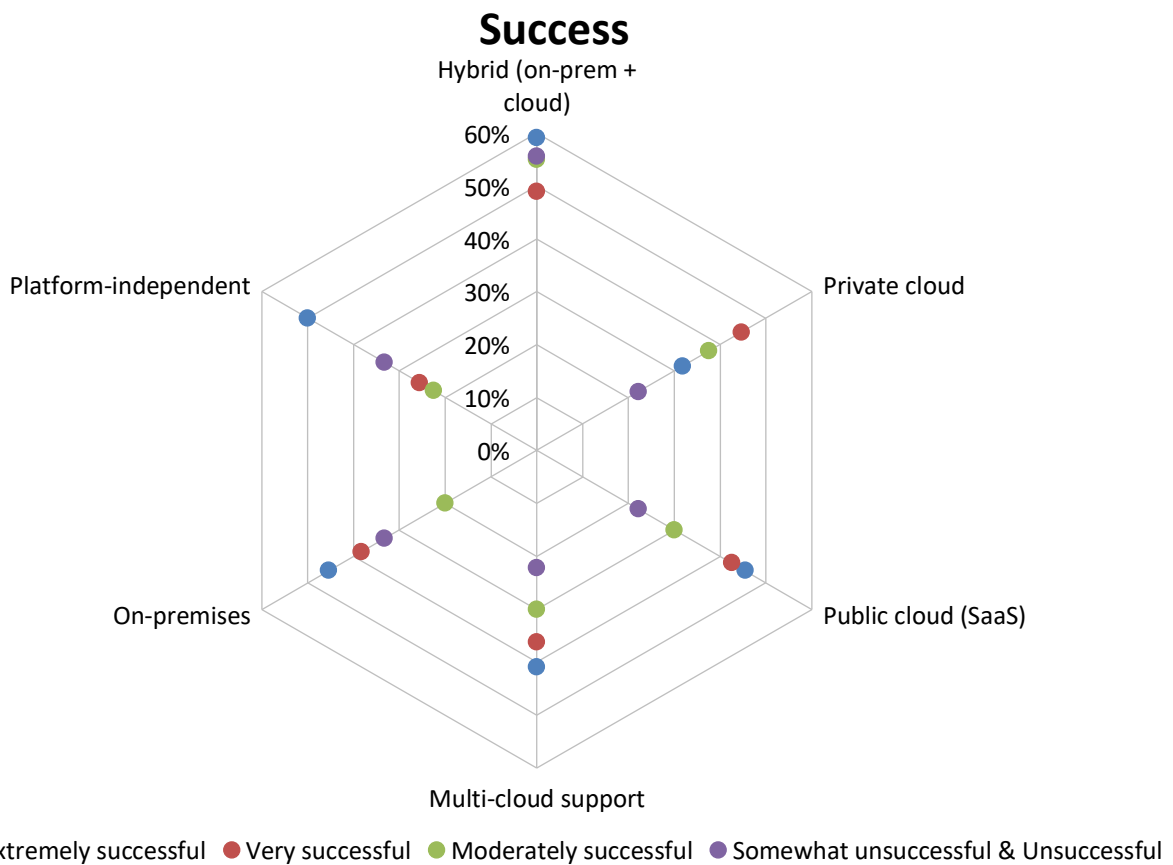


Figure 46 - Data Engineering Deployment Options by BI Success

Licensing of Data Engineering Capabilities

Our respondents value a range of licensing options for data engineering technology (fig. 47). User-based licensing models are most popular, with a slight majority of 51% of organizations indicating they prefer this approach. Models based on computing resources consumed are the second most-preferred, indicated as such by 46% of respondents. Concurrent use, open source and data volume-based approaches are less-favored in our sample, with less than 40% of organizations indicating these are preferred options.

Data Engineering Licensing Preferences

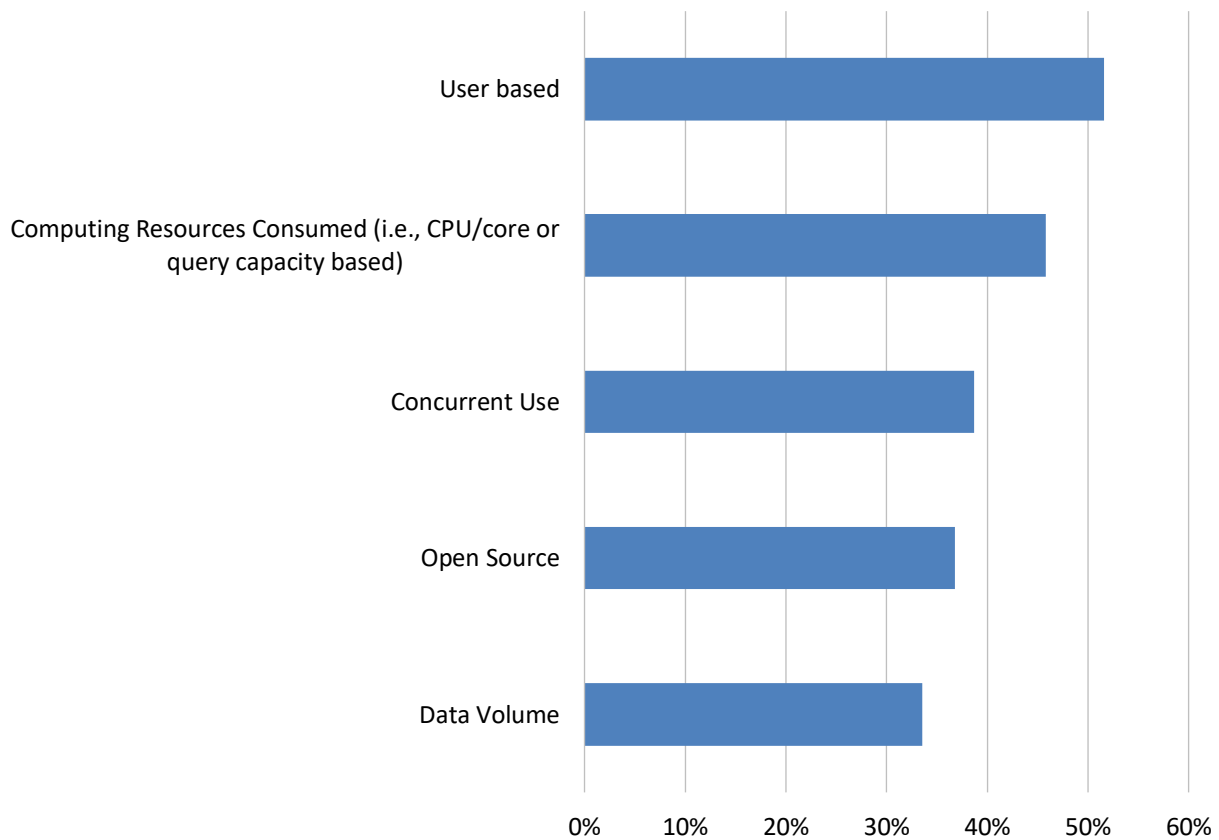


Figure 47 - Data Engineering Licensing Preferences

Similar to other aspects of data engineering technology, preferences for licensing models vary substantially across regions (fig. 48). Asia Pacific organizations seem to accept all options, indicating preference frequencies of 50% or higher across the full set. In contrast, Latin American organizations prefer user-based and resource-based approaches at a 63% rate and all other options far less. North American and EMEA organizations, like Asia Pacific, show an openness to diverse options, although at much lower levels—below 50% for all licensing model approaches.

Data Engineering Licensing Preferences by Geography

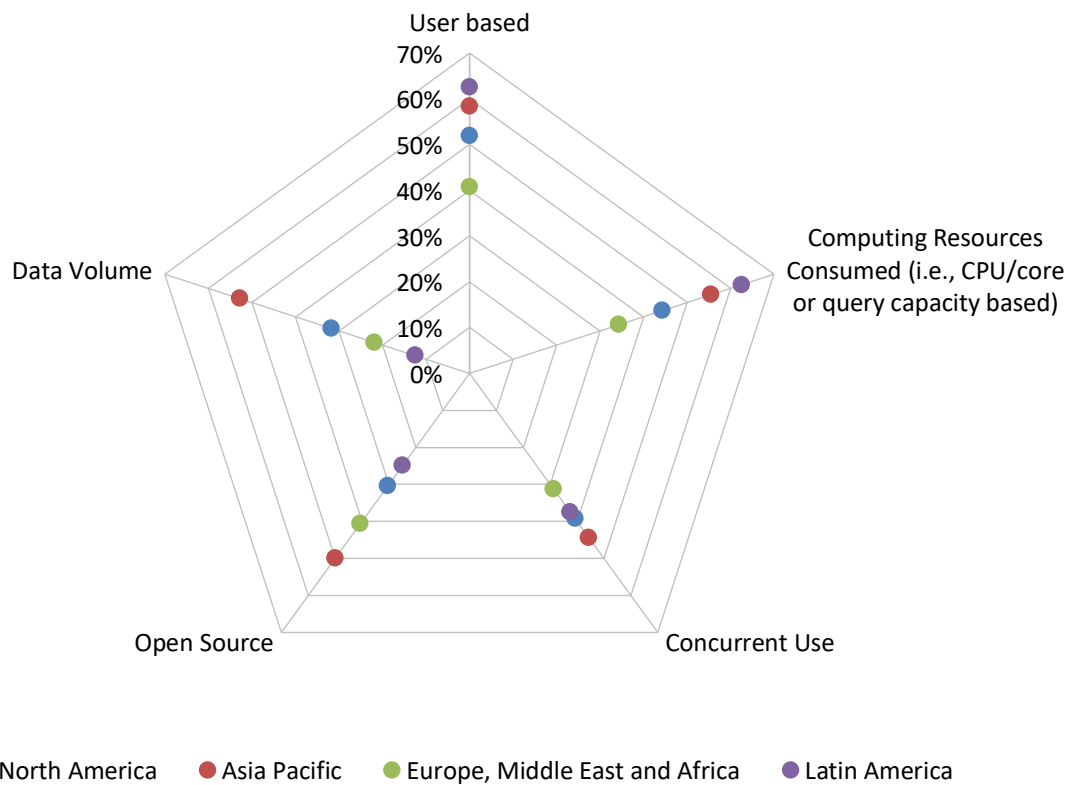


Figure 48 - Data Engineering Licensing Options by Geography

Licensing preferences for data engineering vary significantly across organizations of different sizes. Very large organizations with more than 10,000 employees prefer user-based licensing with 41% frequency, far lower than organizations of other smaller size (fig. 49). These very large organizations are more likely to prefer resource-based approaches (50% frequency) compared to smaller organizations. Medium (101-1,000 employees) and large (1,001-10,000 employees) organizations prefer a mix of licensing models at more moderate levels of frequency, although they rate user-based and resource-based as most preferable. Small organizations (1-100 employees) prefer user-based approaches at a 55% frequency, well above their preferences for all other model types.

Data Engineering Licensing Preferences by Organization Size

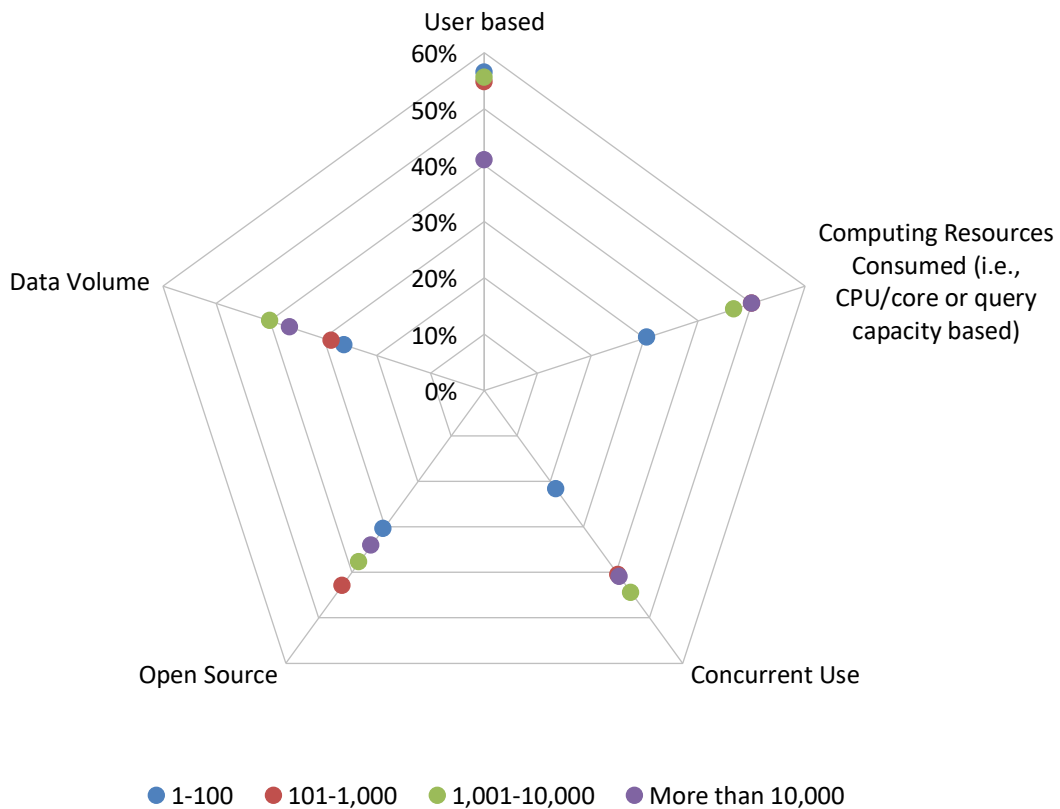


Figure 49 - Data Engineering Licensing Preferences by Organization Size

Continuing the theme of diversity, our survey data shows a clear link between diverse data engineering licensing preferences and organizations' success with BI. Those indicating their BI efforts are extremely successful also more often prefer all licensing approaches than do their peers experiencing lower levels of BI success (fig. 50). More than 50% of those extremely successful organizations indicate a preference for all licensing models other than concurrent use. In contrast, those indicating they are very successful or only moderately successful with BI tend to emphasize a single licensing approach, as well as show less frequent preference across all options.

Data Engineering Licensing Preferences by BI Success

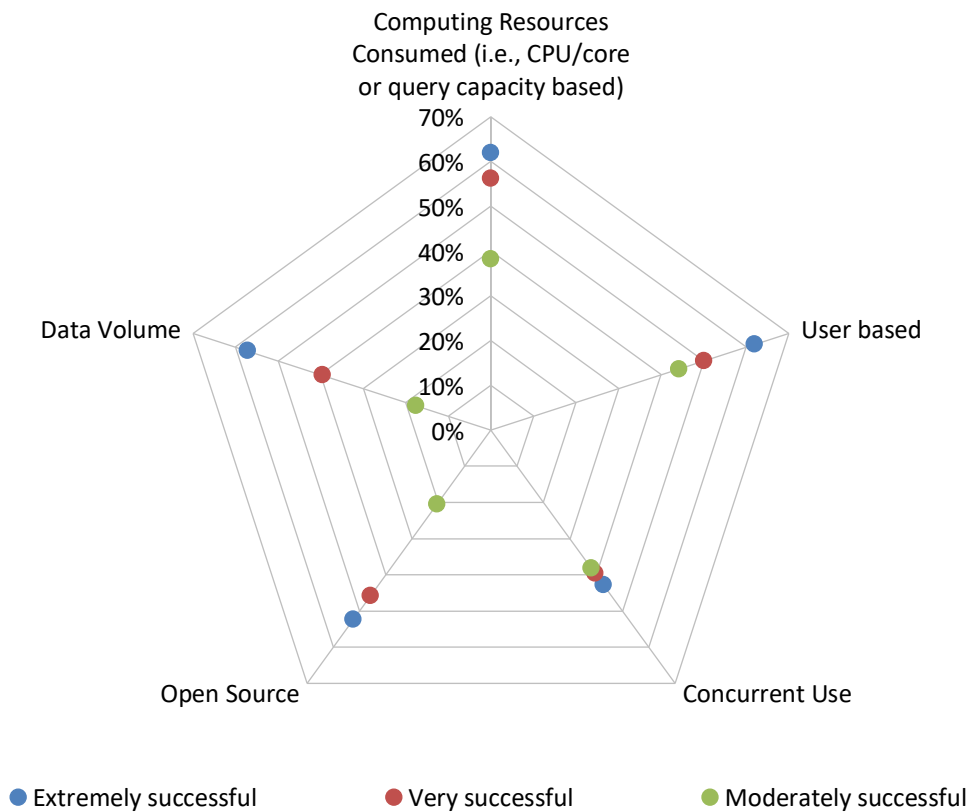


Figure 50 - Data Engineering Licensing Preferences by BI Success

Sourcing of Data Engineering Capabilities

As a component of ADI, data engineering capabilities are increasingly bundled with other ADI components by vendors with large product portfolios. At the same time, numerous providers of stand-alone data engineering technology continue to enter the market. We observe buyers adopting a range of sourcing strategies for data engineering (fig. 51). The current most popular approach, adopted by 68% of data engineering respondents, is choosing stand-alone best-of-breed vendors. Forty-six percent choose vendors that partner and integrate with the buyer’s preferred ADI platform vendor. Thirty-three percent of organizations get data engineering technology as a component of their primary ADI vendor solution. And selecting a vendor that offers data engineering from the primary ADI vendor’s marketplace is the least popular approach, at 31%.

Data Engineering Buying Approaches

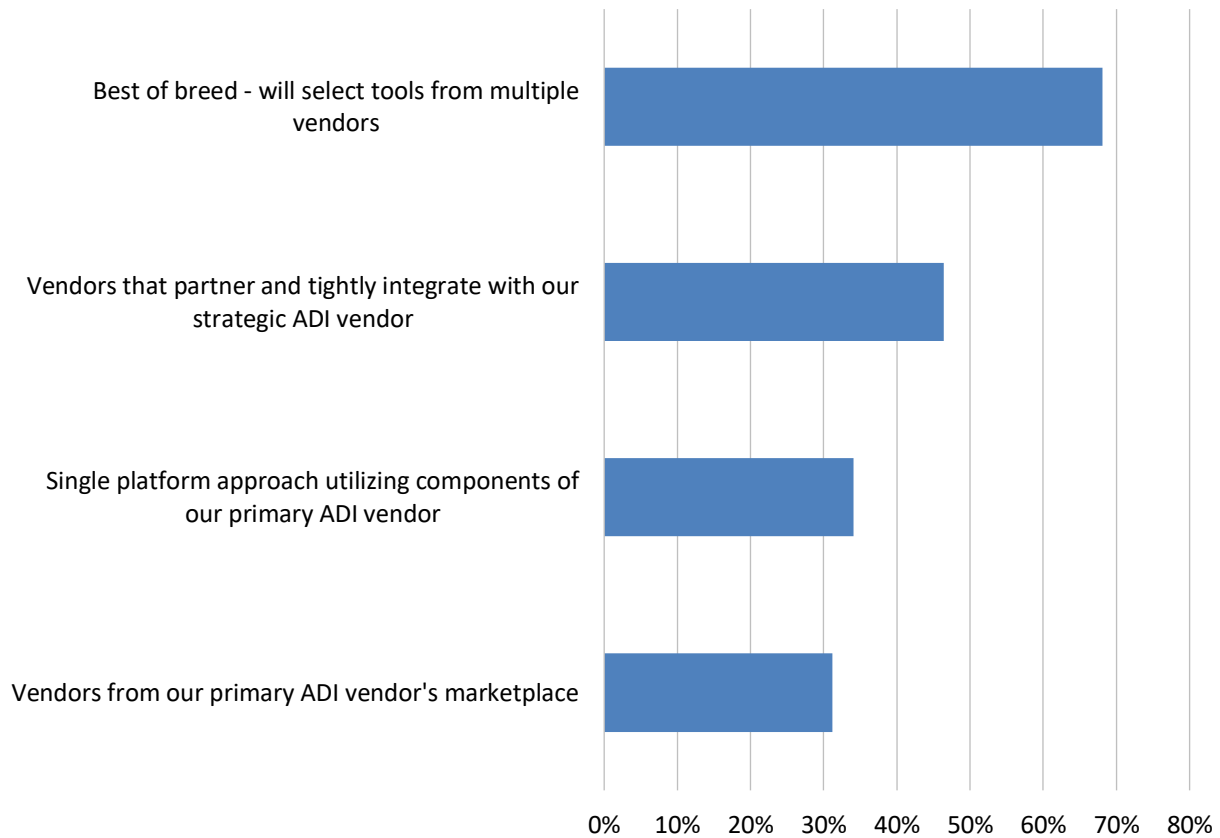


Figure 51 - Data Engineering Buying Approaches

Preferences for data engineering sourcing vary across regions. Best-of-breed providers are most often the preference of North American and EMEA organizations, with 70% and 80% frequency, respectively (fig. 52). In North America, that preference far exceeds all other options, whereas EMEA reflects more diversity of preference. Asia Pacific organizations prefer best of breed and vendors that partner with their strategic ADI vendor at nearly equal frequencies of 56% and 59%. Seventy-five percent of Latin American organizations indicate a preference for sourcing data engineering from partners that are integrated with their strategic ADI vendor, far beyond other options.

Data Engineering Buying Approaches by Geography

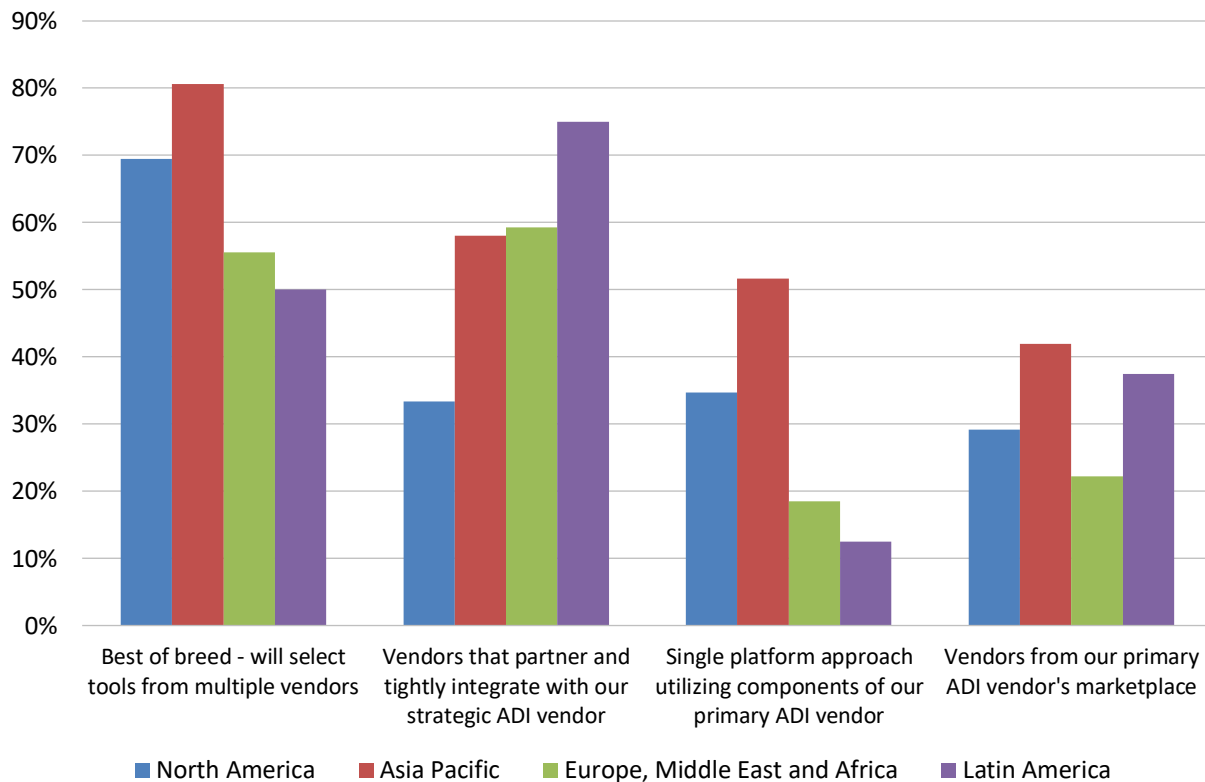


Figure 52 - Data Engineering Buying Approaches by Geography

Across industry sectors, data engineering buying preferences reflect the overall preference toward best of breed (fig. 53). A majority of 55%-75% of organizations in virtually all industries indicate they prefer this approach. Retail and wholesale is an exception, with 67% of respondents in that industry preferring vendors that partner/integrate with their strategic ADI vendor. Most industries show a clear preference for one approach far above all others. However, manufacturing and education show a different pattern—while both sectors prefer best of breed with highest frequency, they also indicate a balanced level of preference for the remaining styles at a moderately lower degree.

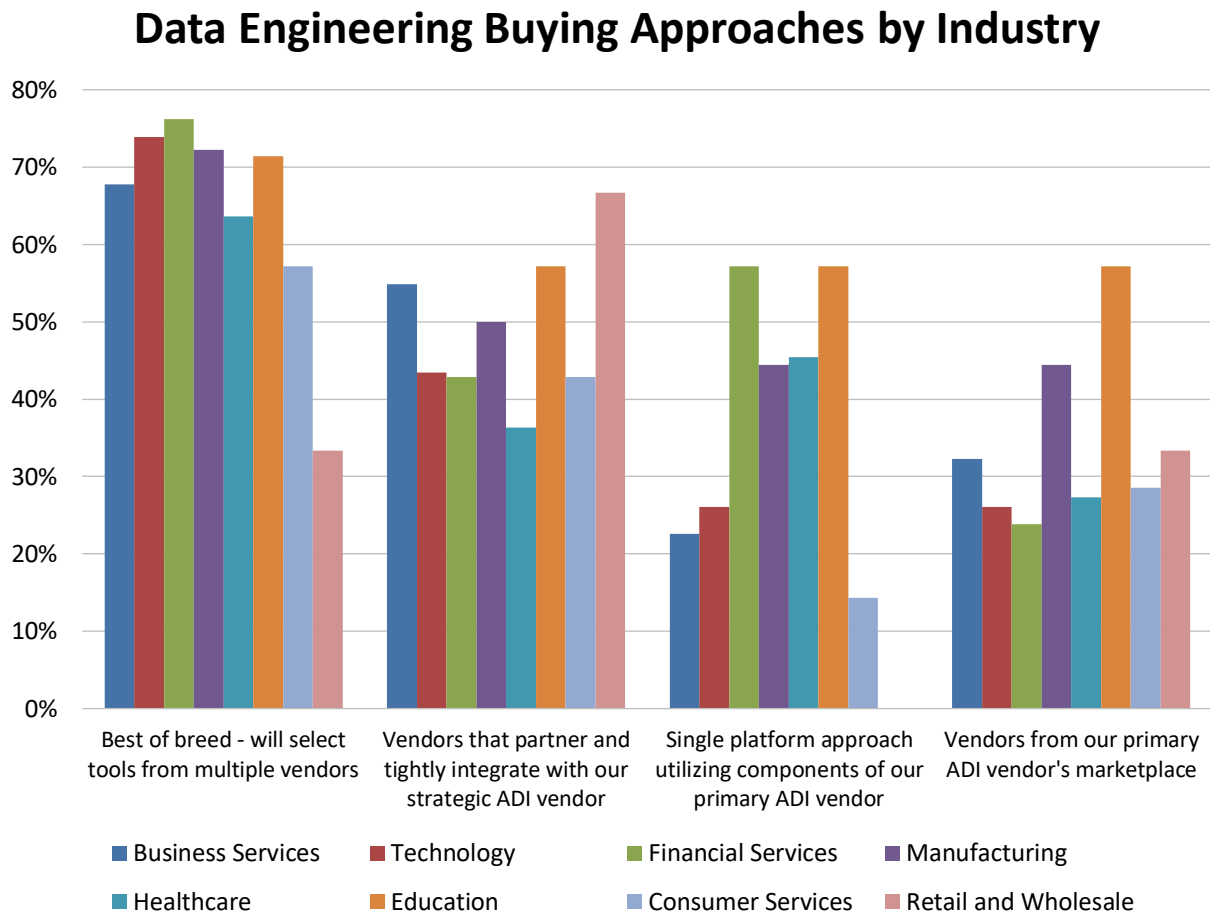


Figure 53 - Data Engineering Buying Approaches by Industry

It is useful to consider any relationship between how buyers source data engineering technology and their degree of success in BI efforts, even though it is not the most significant driver of that success. Seventy-four percent of organizations reporting they are extremely successful with BI indicate a preference for best-of-breed vendors for data engineering capabilities, compared to less than 60% for all other options (fig. 54). A similar pattern holds for those organizations indicating they are very successful with BI. Moderately successful organizations also prioritize best of breed, with 60% indicating it is their sourcing method of choice for data engineering.

Data Engineering Buying Approaches by BI Success

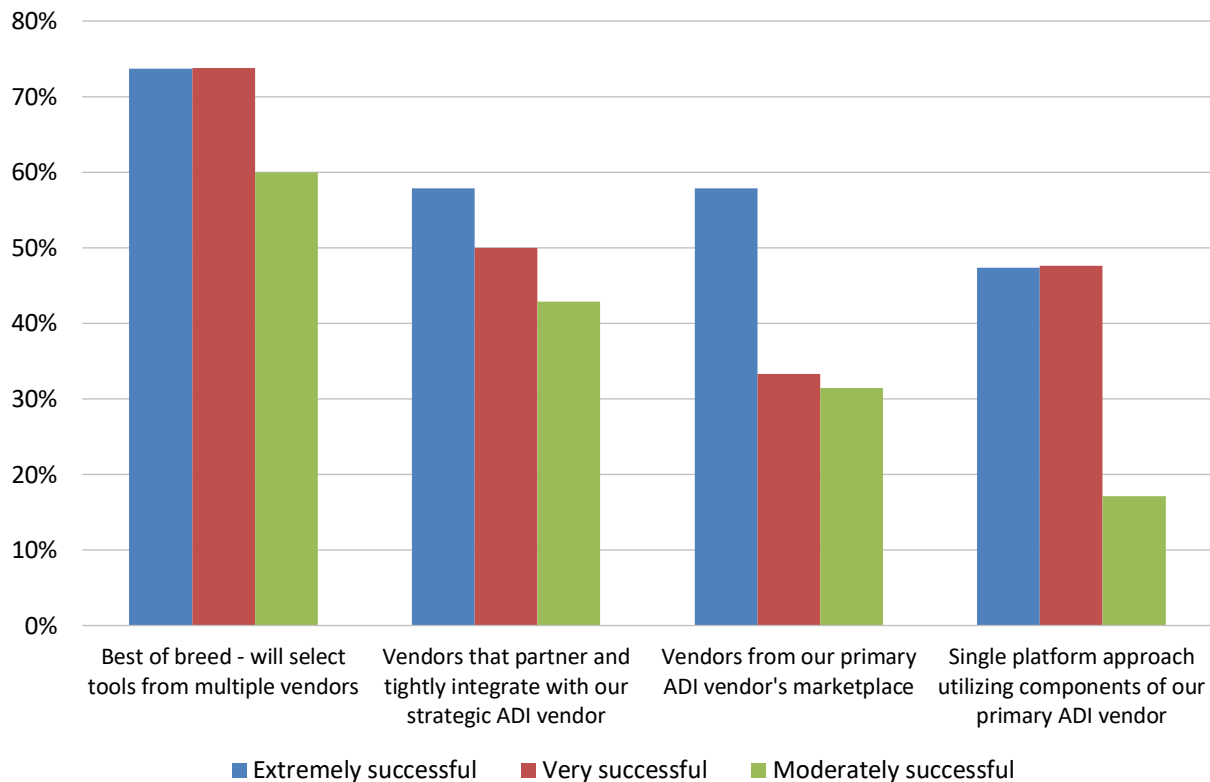



Figure 54 - Data Engineering Buying Approaches by BI Success



Industry and Vendor Analysis

Industry and Vendor Analysis

ADI is a market ecosystem in constant evolution, where an increasing number of vendors provide an ever-widening range of ADI-relevant technology components as part of their product portfolio. Data engineering functionality is growing more prevalent in the market as large providers of very broad capabilities continue to bundle it with other ADI components. And specialist vendors with a specific emphasis on enabling technology that supports data engineering roles keep entering the market.

Industry Support for Data Engineering Capabilities Within ADI

In 2026, 63% of responding vendors indicated data engineering is core to their technology platform (fig. 55). With another 23% stating they provide such capabilities via integration with other vendor’s technology, an overwhelming majority can supply data engineering functionality to customers. Customers’ familiarity with data engineering concepts helps vendors market solutions, but also means they face strong competition.

Industry Support of ADI Components

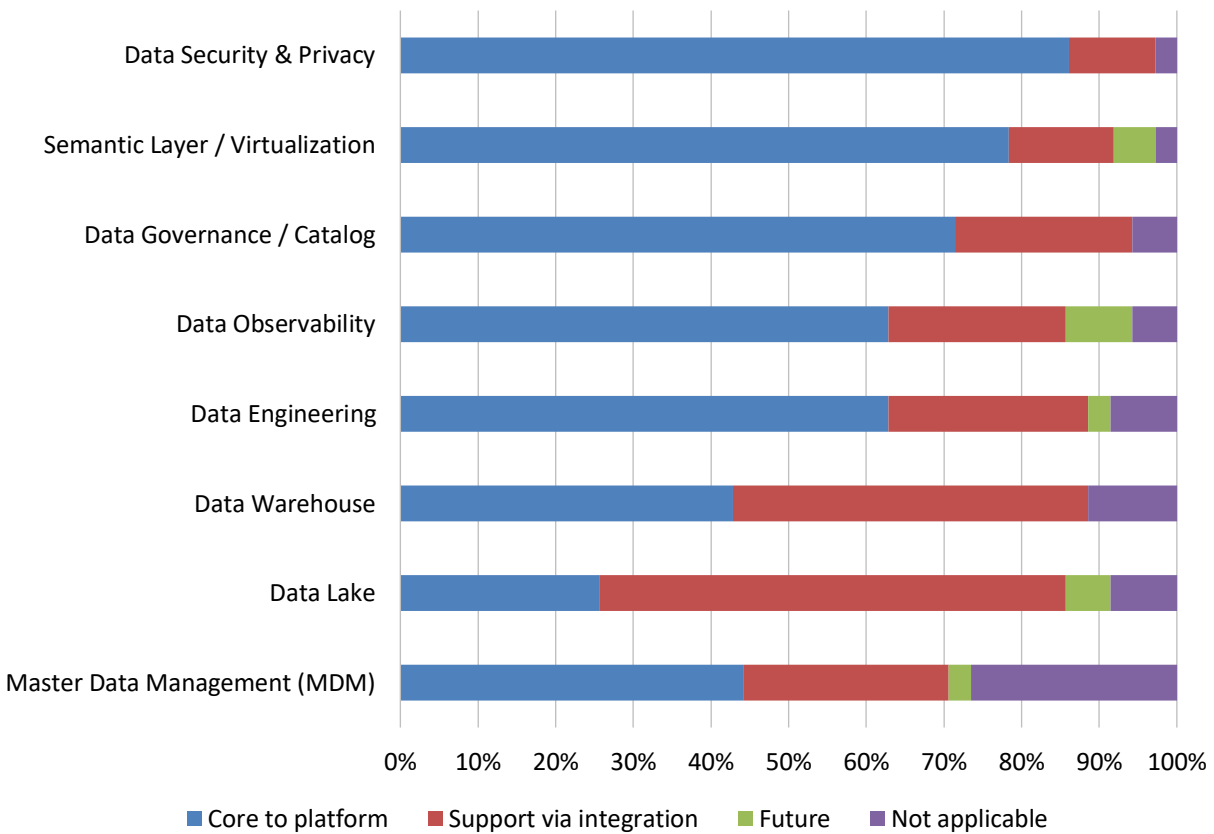


Figure 55 – Industry Support of ADI Components

Industry Support of Sources and Targets

Organizations require that their data engineering technology support a growing range of data sources and targets. Vendors are responding to these needs by steadily adding support for more data storage platforms. All vendors claim to support the full range of required data source types (fig. 56). Their ability to target that full range is a work in progress, with 94% supporting analytical databases, 87% supporting relational databases, and 80% supporting various file structures. Support for other targets is variable, with less than 80% of vendors supporting object stores, graph databases, business applications, noSQL, Hadoop, and speciality data platforms.

Industry Support of Data Engineering Sources and Targets

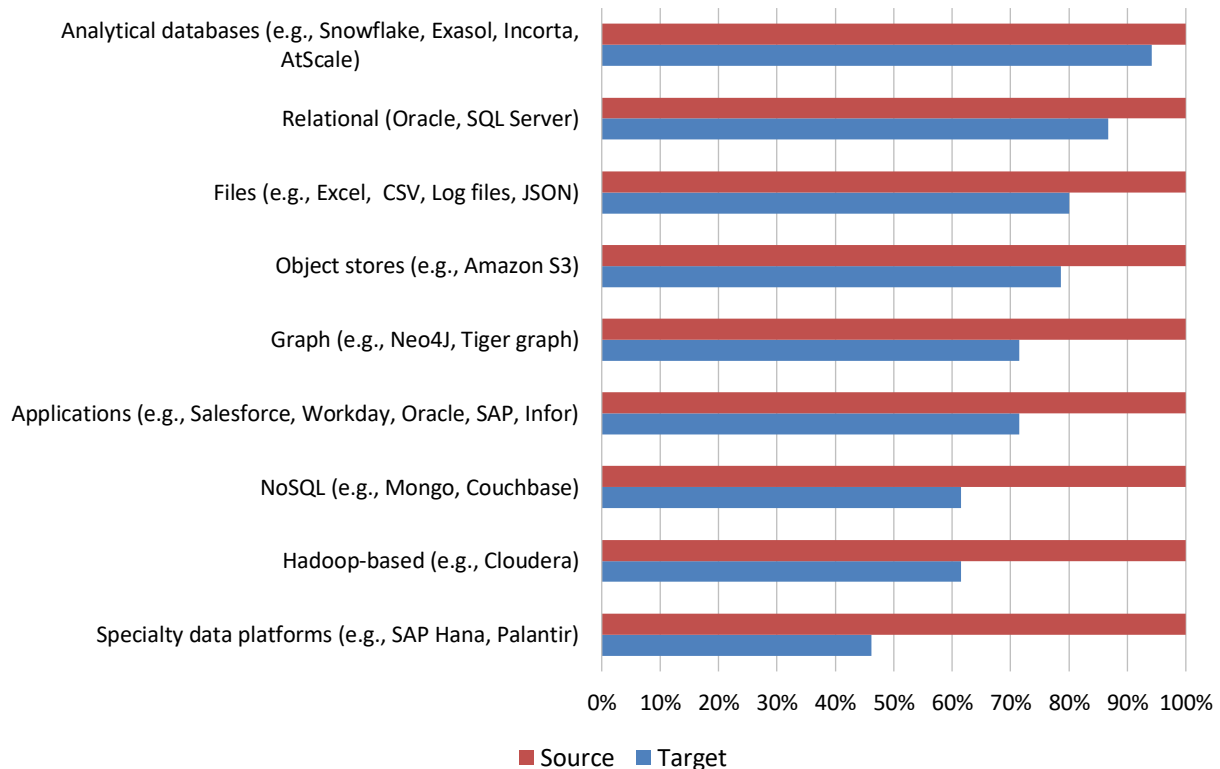


Figure 56 – Industry Support of Data Engineering Sources and Targets

Industry Support of Interface Types

Meeting customer demand for data source and target support is more nuanced than just promoting how many data platforms the data engineering tools can interact with. Equally important is HOW that interaction happens—what type of interfaces are used to allow data engineering tools to talk to those data platforms. Most vendors employ a combination of native interfaces and third-party, standards-based interfaces (fig. 57). Most vendors support native interfaces for a wide range of sources and targets, presumably enabling simpler integration and better performance. Some also employ standards-based interfaces for certain data types. More than 50% of vendors provide native interfaces for all critical data sources and targets, while third-party interfaces are offered by less than 80% of vendors for all types.

Industry Support of Data Engineering Interface Types

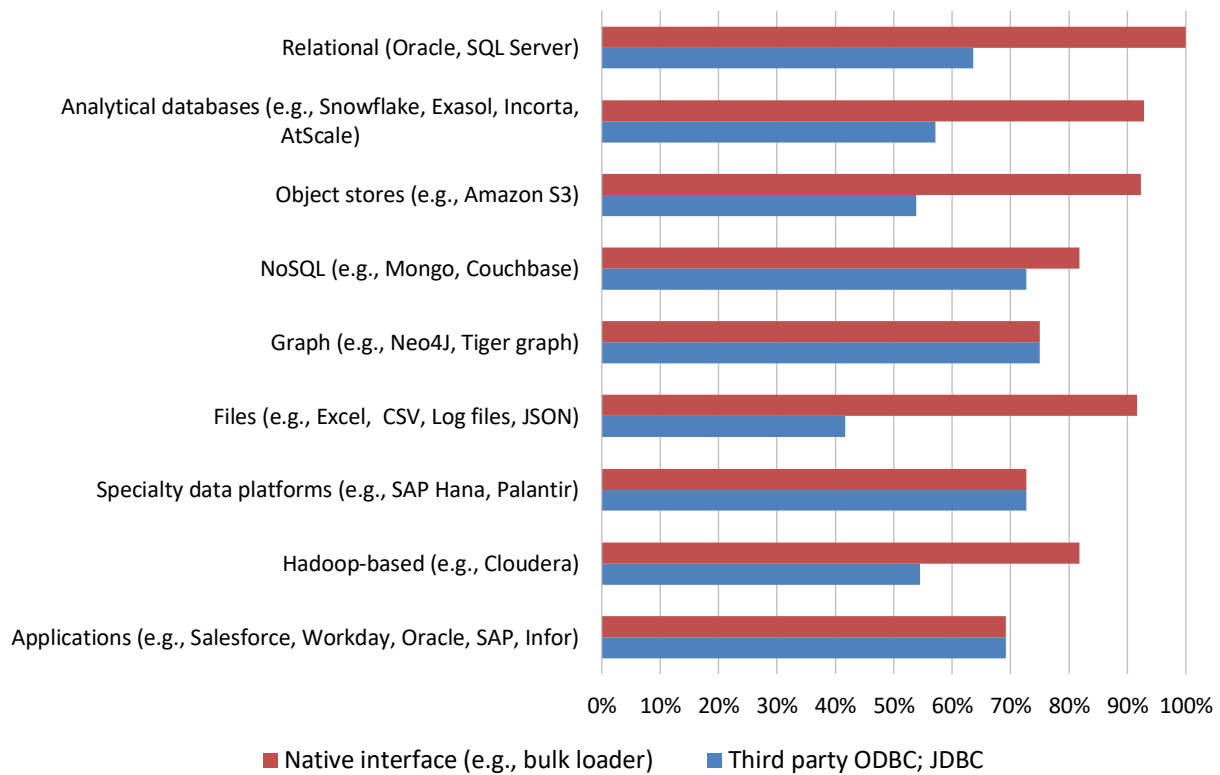


Figure 57 – Industry Support of Data Engineering Interface Types

Industry Support for Development and Usability Features

Customers view strong capabilities for development of data pipelines and overall usability for data engineers as critical to their success and the perceived value of the tools. Vendors serving data engineering demand appear to offer strong support for the main features customers seek in this regard. All surveyed vendors currently support features such as simple development of workflows, pre-built components, governance capabilities, and creation of integrated views of data (fig. 58). Some features garner more sporadic vendor support, with less than 70% of vendors currently supporting ML-based and other automated recommendations and automated data quality rules. However, vendor plans indicate that all development and usability features for data engineering will be supported by over 90% of vendors within 24 months.

Industry Support for Data Engineering Development and Usability Features

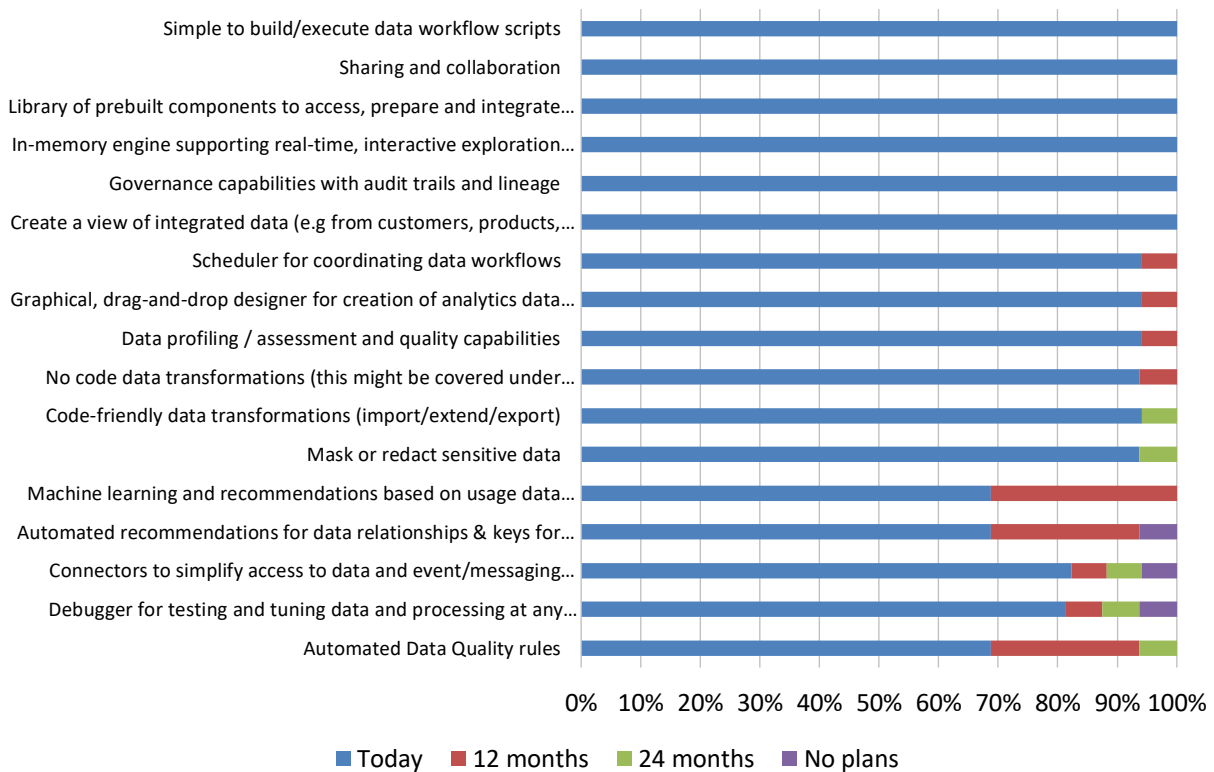


Figure 58 – Industry Support for Data Engineering Development and Usability Features

Industry Support for Data Processing and Transformation Features

Customers likewise view core functionality for manipulating data as a high priority in their selection and use of tools for data engineering. Topping the list as well-supported are features like ETL, the ability to derive new data features, and transformations such as grouping, aggregating, cutting and merging values and so forth. All surveyed vendors state they offer these features today (fig. 59). Over 80% of vendors currently support most other critical features in this category, underscoring the wide availability of solutions to serve demand. Support is more variable for specialized features, such as support for unstructured data, big data-specific platforms such as Kafka and Apache, and automation of data quality rules, with less than 75% of vendors providing these features today.

Industry Support of Data Engineering Data Processing and Transformation Features

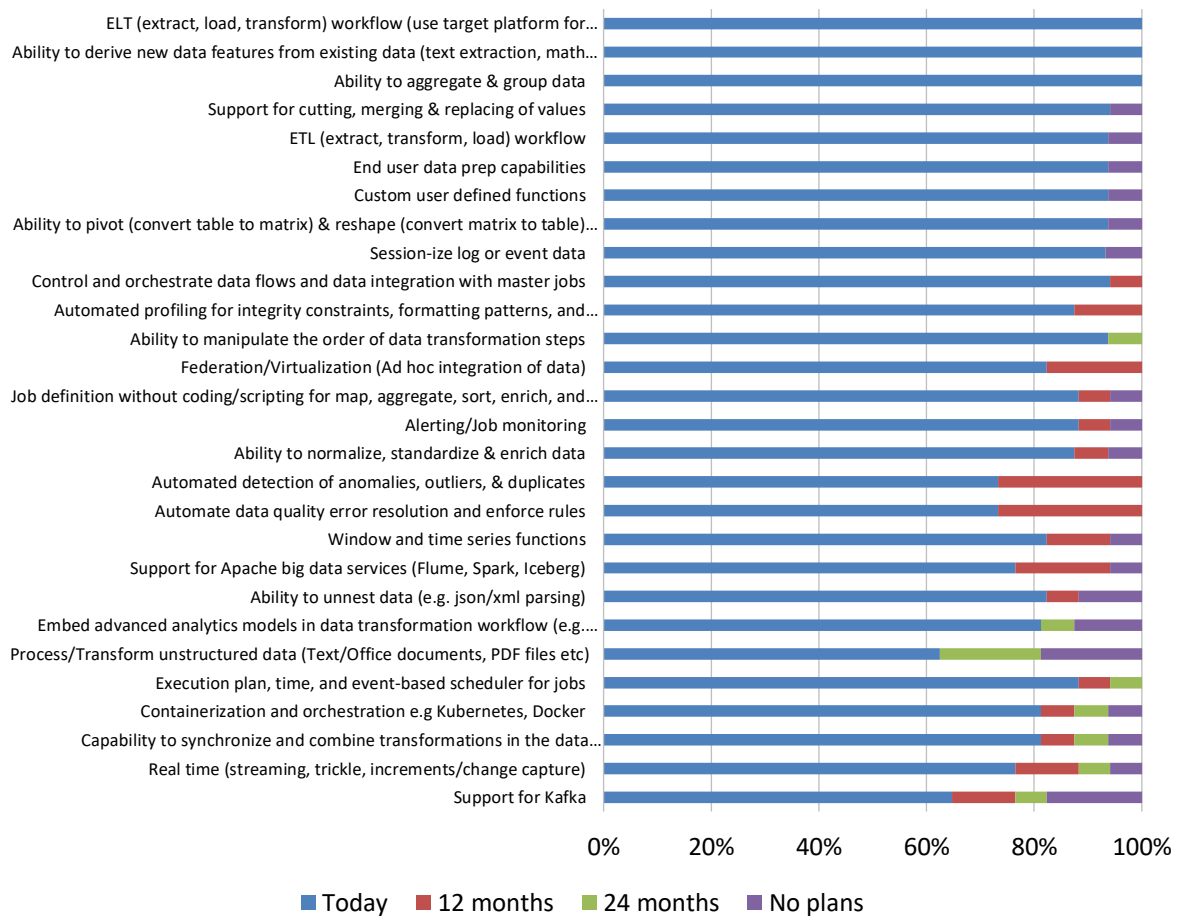


Figure 59 – Industry Support of Data Engineering Data Processing and Transformation Features

Industry Support for Single Product Data Engineering Functionality

Customers tend to prioritize vendor offerings that provide all their desired functionality within a single product, rather than limited functionality that requires them to purchase, deploy and possibly integrate multiple products. A very strong majority of 88% of vendors working in this space claims to provide all their data engineering functionality as a single product (fig. 60). The remaining 12% of vendors support customers' data engineering needs through some combination of multiple products.

Industry Support of Single Product Data Engineering Functionality

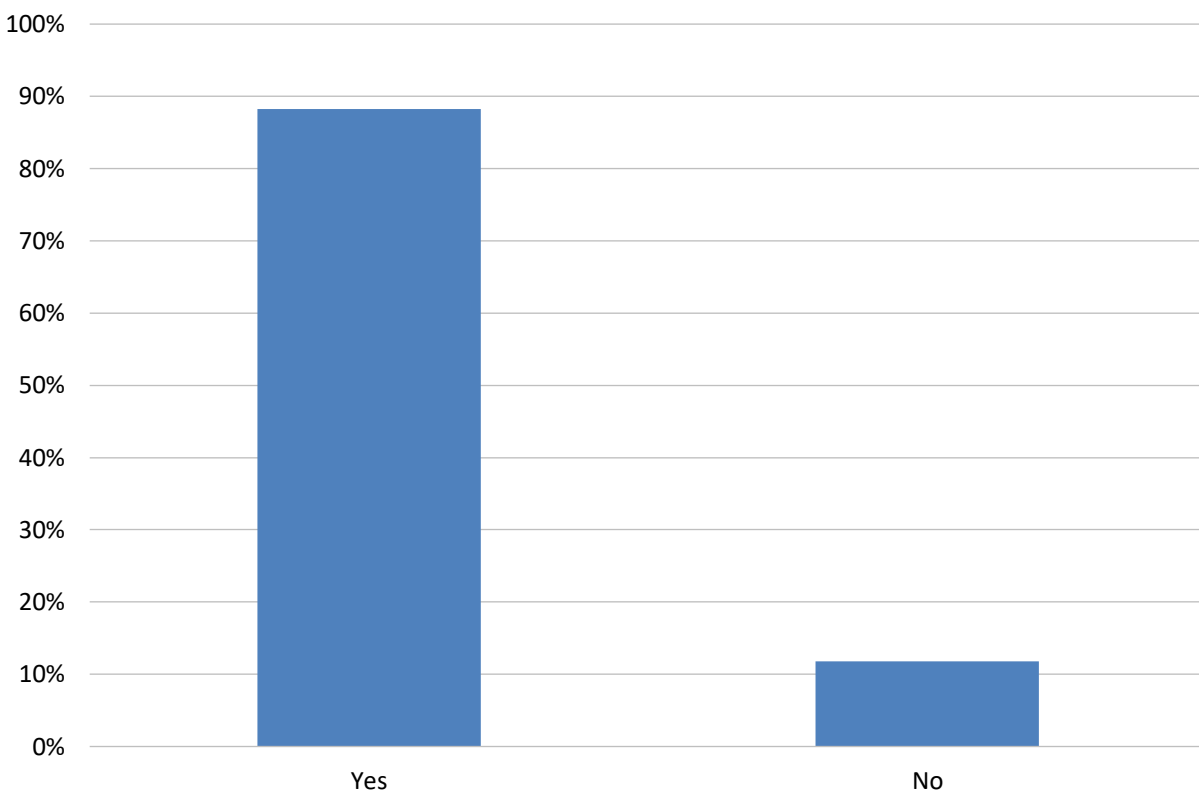


Figure 60 – Industry Support of Single Product Data Engineering Functionality



Vendor Ratings

Vendor Ratings

We include 32 vendors in our 2026 data-engineering ratings (fig. 61). For each vendor, we consider the following features: development and usability, process and transformation of data, and data sources and targets.

Scores for the top half of vendors ranked are relatively close, with the top and bottom of that group varying by only 15%, with multiple ties.

Data Engineering Vendor Ratings

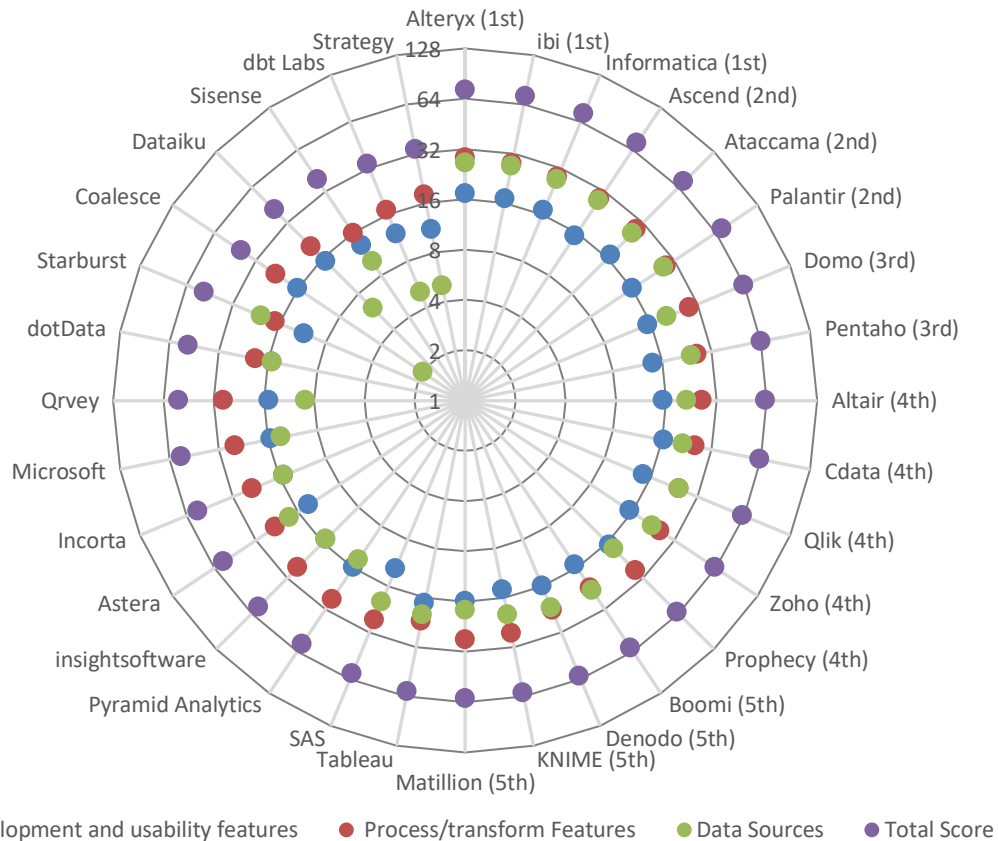


Figure 61 – Data engineering vendor ratings

Other Dresner Advisory Services Research Reports

- Wisdom of Crowds® “Flagship” Business Intelligence Market Study
- AI Development Platforms
- Agentic AI Assisted Analytics
- Agentic AI Automation Platforms
- Analytical Data Infrastructure
- Analytical Data Products
- Carbon Planning and Analysis
- Cloud Computing and Business Intelligence
- Data Catalog
- AI, Data and Analytics Governance
- Embedded Business Intelligence
- Enterprise Performance Management
- Enterprise Resource Planning (ERP)
- ESG Reporting
- Financial Consolidation, Close Management and Reporting
- Guided Analytics
- ModelOps
- Self-Service Business Intelligence
- Semantic Layer and Data Virtualization
- Supply Chain Planning and Analysis
- Workforce Planning and Analysis

Appendix: 2026 Data Engineering Survey Instrument

Please provide your contact information below:

First Name*: _____

Last Name*: _____

Company: _____

Email Address*: _____

Major Geography*

- Asia Pacific
- Europe, Middle East and Africa
- Latin America
- North America

Please specify your city and country

City: _____

Country: _____

Please provide your contact information below:

Address 1: _____

Address 2: _____

City: _____

State: _____

Zip: _____

Country: _____

Phone Number: _____

What is your current title?

What function are you a part of?

- Data Science Function
- Executive Management
- Finance
- Business Intelligence / Analytics Competency Center
- Human Resources
- Information Technology (IT)
- Marketing
- Operations (e.g., Manufacturing, Supply Chain, Services)
- Research and Development (R&D)
- Sales
- Strategic Planning Function
- Other - Write In: _____

Please select an 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 and Wholesale
- Sports
- State and Local Government
- Technology
- Telecommunications
- Transportation
- Utilities
- Other - Write In: _____

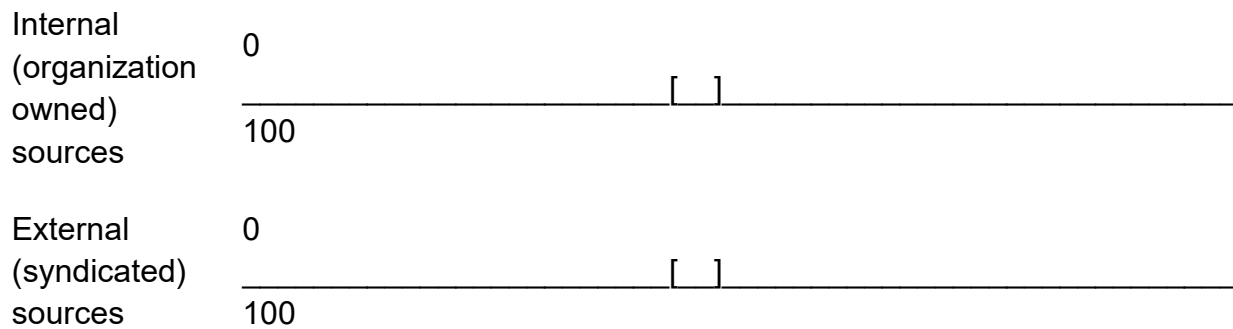
How many employees does your company employ worldwide?

- 1-100
- 101-1,000
- 1,001-2,000
- 2,001-5,000
- 5,001-10,000
- More than 10,000

How old is your company?

- () Less than 5 years old
- () 5-10 years old
- () 11-16 years old
- () 16 or more years old

On average, how many different sources of data (source systems) are being included for data engineering purposes?



Please indicate the estimated percent of your use cases for which your data engineering products are purchased and being used;

	Under 10%	11-20%	21-40%	41-60%	More than 60%
As part of the data integration, cleansing and transformation workflows for a data warehouse supporting dashboards and reporting	()	()	()	()	()
Data integration and transformation services for ad-hoc	()	()	()	()	()

query, discovery and exploration analysis					
Data science and augmented analytics use cases, predictive and prescriptive analytics	()	()	()	()	()
Re-platforming/replication of an existing data warehouse	()	()	()	()	()
Migration of data to new systems (as a result of replatforming, shift to cloud, consolidation during M&A activities, etc)	()	()	()	()	()
Extract and delivery of data to systems or parties outside the enterprise	()	()	()	()	()
Supporting management and delivery of master data	()	()	()	()	()
Supporting data flows between operational systems	()	()	()	()	()

Please indicate the importance of the following data engineering features for processing and transforming data

	Critical	Very Important	Important	Somewhat Important	Not Important
ETL (extract, transform, load) workflow	()	()	()	()	()
ELT (extract, load, transform) workflow (use target platform for transformation processing)	()	()	()	()	()
Process/Transform unstructured data (Text/Office documents, PDF files etc)	()	()	()	()	()
Federation/Virtualization (Ad hoc integration of data; i.e. to support Ad hoc analytics in the case of discovery and exploration and data sciences/modeling using machine learning and advanced analytic techniques such as prescriptive and predictive analytics)	()	()	()	()	()
Capability to synchronize and combine transformations in the data integration workflow, including	()	()	()	()	()

2026 Data Engineering Market Study

notifications and alerts (e.g. Business Activity Monitoring)					
Control and orchestrate data flows and data integration with master jobs	()	()	()	()	()
Execution plan, time, and event-based scheduler for jobs	()	()	()	()	()
Automate data quality error resolution and enforce rules	()	()	()	()	()
Automated detection of anomalies, outliers, & duplicates	()	()	()	()	()
Real time/streaming, trickle, increments/change capture	()	()	()	()	()
Job definition without coding/scripting for map, aggregate, sort, enrich, and merge data operations	()	()	()	()	()
Containerization and orchestration e.g Kubernetes, Docker	()	()	()	()	()
Support for Kafka	()	()	()	()	()

2026 Data Engineering Market Study

Support for Apache big data services (Flume, Spark)	()	()	()	()	()
Embed advanced analytics models in data transformation workflow (e.g. integrate models from R, Python, Scala and Weka to operationalize analytic models within data workflow.	()	()	()	()	()
Alerting/Job monitoring	()	()	()	()	()
Ability to unnest data (e.g. json/xml parsing)	()	()	()	()	()
Ability to normalize, standardize & enrich data	()	()	()	()	()
Support for cutting, merging & replacing of values	()	()	()	()	()
Ability to aggregate & group data	()	()	()	()	()
Ability to pivot (convert table to matrix) & reshape (convert matrix to table) data	()	()	()	()	()
Ability to derive new data features from existing data (text extraction, math)	()	()	()	()	()

2026 Data Engineering Market Study

expressions, date expressions, etc.)					
Ability to manipulate the order of data transformation steps	()	()	()	()	()
Session-ize log or event data	()	()	()	()	()
Window and time series functions	()	()	()	()	()
Custom user defined functions	()	()	()	()	()
End user data prep capabilities	()	()	()	()	()
Automated profiling for integrity constraints, formatting patterns, and dependencies between columns and datasets.	()	()	()	()	()
Use of AI agents to automate creation or revision of data models (e.g., based on source schema, lineage, or usage patterns)	()	()	()	()	()
Use of AI agents for automated ingestion and transformation of new data sources (e.g., parsing files, mapping fields, applying formatting rules)	()	()	()	()	()

Use of AI agents to detect and recommend fixes for data quality issues (e.g., missing values, inconsistent types, anomaly detection)i.e. modeling, movement, transformations.	()	()	()	()	()
---	-----	-----	-----	-----	-----