



March 3, 2025

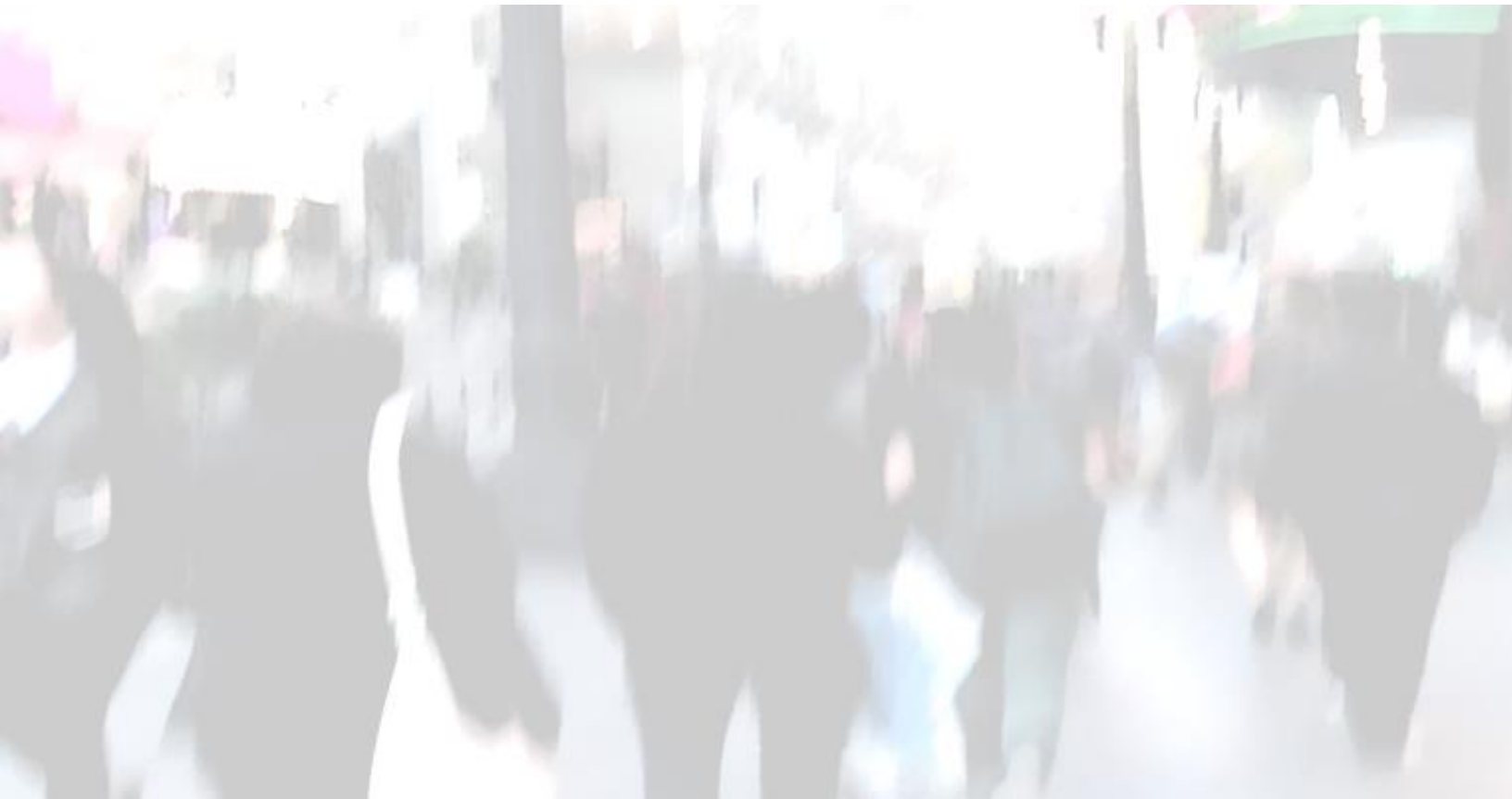
Dresner Advisory Services, LLC

2025 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 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 BI/analytic use cases. Deployments may be use-case specific, i.e., focused on a small number of users doing data science projects with data sources/transformations that are required or may be required to be fault-tolerant, highly secure, and scalable to span large data volumes, multiple analytic steps, and multiple analytic models, and to support multiple BI use cases and tools.

Introduction

As Dresner Advisory Services approaches its 18th anniversary in 2025, we are pleased to present the 11th edition of this report. We sincerely appreciate our clients' and subscribers' ongoing support and collaboration. Since our founding in 2007, we have remained committed to setting high standards, driving innovation, and increasing the value we deliver each year.

This report continues to explore market requirements and priorities for data orchestration, integration, and transformations including advanced analytics in the data engineering pipeline workflow.

We hope the insights in this report support your operational and strategic goals. Developing this study has been a valuable endeavor, and we look forward to continuing to serve you.

Thank you to our clients, colleagues, and the broader community—your support is integral to our research. We welcome the opportunity to engage further as you explore the study's findings.

Best,



Howard Dresner
Chief Research Officer
Dresner Advisory Services

Contents

Definitions	3
Business Intelligence Defined	3
Data Engineering 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
About Bill Hostmann	9
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	10
Data Quality	10
Executive Summary	12
Study Demographics	15
Geography	15
Functions	16
Vertical Industries	17
Organization Size	18
Analysis and Trends	20
Importance of Data Engineering	20
Use and Plans for Data Engineering	26
Use Cases for Data Engineering	32
Data Engineering Features for Data Processing and Transformation	37

Frequency of Data Engineering	43
Data Engineering Usability	49
Frequency of Enrichment with Third-Party Data	55
Data Sources and Targets for Data Engineering	62
Effectiveness with Current Approach to Data Engineering	65
Deployment of Data Engineering Capabilities	70
Industry and Vendor Analysis.....	76
Industry Support for Data Engineering Deployment Options	76
Data Engineering Pricing Models.....	77
Industry Support for Development and Usability Features	78
Industry Support for Data Processing and Transformation Features	79
Vendor Ratings	81
Other Dresner Advisory Services Research Reports	82
Appendix: 2025 Data Engineering Survey Instrument.....	83

Benefits of the Study

The 2025 Dresner Advisory Services Data Engineering Market Study provides a wealth of information and analysis, offering value to both consumers and producers of business intelligence 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 collaborative BI and related technologies.

From our unique vendor performance measurement system, users glean key insights into BI 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 business intelligence 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 two 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®
- AI, Data Science and Machine Learning
- Analytical Platforms
- Cloud Computing and Business Intelligence
- Collective Insights®
- Data Engineering
- Data Governance
- Guided Analytics®
- Self-Service BI
- Semantic Layer

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

About Bill Hostmann



Bill Hostmann is a VP and Research Fellow with Dresner Advisory. His area of focus includes trends in analytic data infrastructures (ADI)—integrating and managing the information and information models used by BI, advanced analytics, and CPM/PM applications.

Bill has more than 20 years of experience at the intersection of business intelligence/analytics and data analytics infrastructure, including positions in product and general management at Gemstone Systems, Informix, and Informatica.

He spent 14 years as a research analyst at Gartner, including several years as a VP and Distinguished Analyst for BI/Analytics. Bill's education includes a BS in electrical engineering, an MS in computer science and engineering, and an MBA.

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 Sherry Fairchok

Sherry 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.

We include our own research community of over 8,000 organizations as well as crowdsourcing and vendors' customer communities.

Data Quality

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

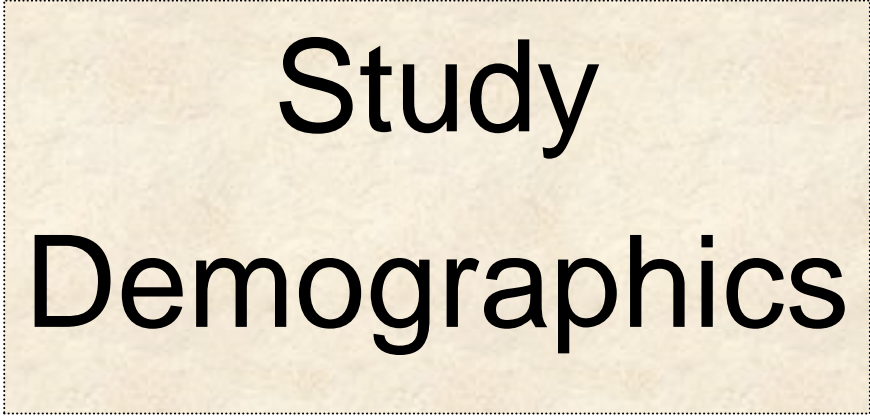
Executive Summary

Executive Summary

- Eighty-one percent of our user survey respondents indicate data engineering is critical (fig. 5) or very important, up from 77 percent in our 2024 report. The increased importance of data engineering reflects the rising demand for access to analytical data as well as the ongoing fragmentation of data sources and targets. Less than 1 percent of user survey respondents indicate that data engineering is not important.
- Seventy-five percent of user respondents say their organizations use data engineering capabilities today and 20 percent of those respondents indicate plans to expand the use. Thirteen percent of respondents indicate plans to use data engineering tools within the next 12 months (fig. 10). Only 7 percent have no plans to use data engineering capabilities.
- There appears to be room for improvement in current approaches to data engineering. Thirty-seven percent of our user survey respondents rate their current approach as highly effective (fig. 53). Forty-eight percent of our survey respondents rate their current approach as somewhat effective.
- Data engineering products are often purchased and used for multiple BI use cases (fig 16). Thirty-two percent of our user respondents say they employ the majority (over 60 percent) of their data engineering capabilities for the use case of data integration, cleansing, and transformation workflows for a data warehouse supporting dashboards and reporting, and 22 percent indicated that most of their use of data engineering supports data flows between operation systems. When comparing year-over-year trends of use cases, we see a new data engineering trend may be emerging from its increased use in data science and augmented analytics use cases and replatforming/replication of an existing data warehouse (fig. 17).
- The top features for data engineering workflows include the ability to aggregate and group data and ETL and ELT workflows, as well as the management of engineering workflows, such as alerting/job monitoring and execution plan-, time-, and event-based schedulers for jobs. Least important is support for Kafka and Apache big data services (fig. 21).
- Private and public cloud (SaaS) lead in the deployment of data engineering capabilities (fig. 61). Continuing an upward trend over the past three years, we see on-premises deployment increasing in priority in 2025 (fig. 61).
- We asked our survey respondents about the use and frequency of data enrichment with third-party data as part of data engineering workflows (fig. 37). Twenty-three percent of the user responses said they rarely use third-party data for enrichment of data engineering processes. Twenty-six percent occasionally enrich their data engineering workflows with third-party data. Eighteen percent

constantly enrich with third-party data and 28 percent frequently enrich their data sets with third-party data as part of their data engineering workflows. Only 5 percent of organizations never use third-party data for enrichment.

- Overall, respondents indicate that relational databases, file systems and applications are the top data sources and targets for data engineering workflows. Analytical databases are also a primary target of data engineering pipelines, but according to this year's survey data, their use appears to have declined. Of note is object stores (e.g., Amazon S3, Google Cloud Storage, and Microsoft Azure blob storage), which are an important target for data engineering workflows but also highly utilized as a source. This suggests their role is expanding as offline storage and staging repository for data engineering workflows (fig. 42).



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 54 percent, followed by EMEA (24 percent). Asia Pacific and Latin America account for the balance (22 percent) of respondents (fig.1)

Geographies Represented

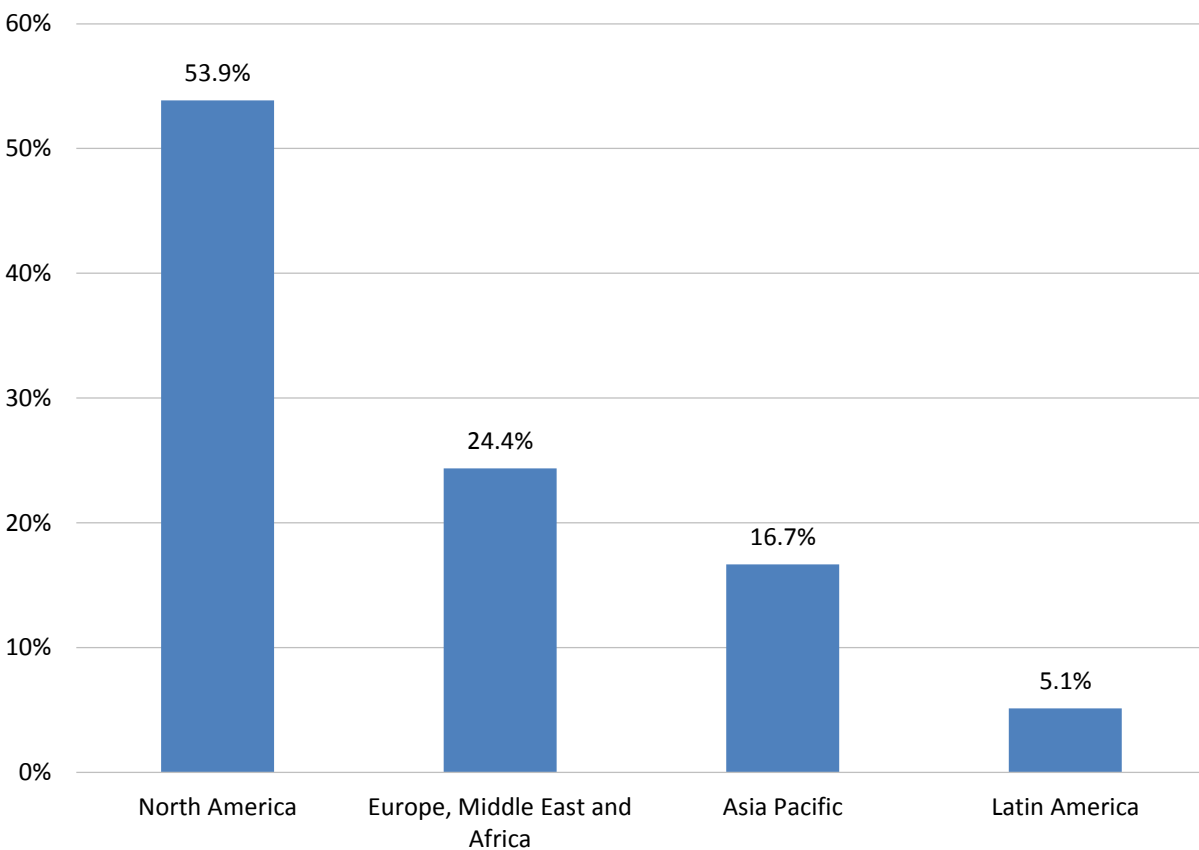


Figure 1 – Geographies represented

Functions

In 2025, respondents from IT functions account for 40 percent of our sample, followed by the Business Intelligence Competency Center (BICC) at 18 percent of the respondents (fig. 2). Participants from several business functions responded, including data science functions, R&D, finance, operations, and executive management.

Functions Represented

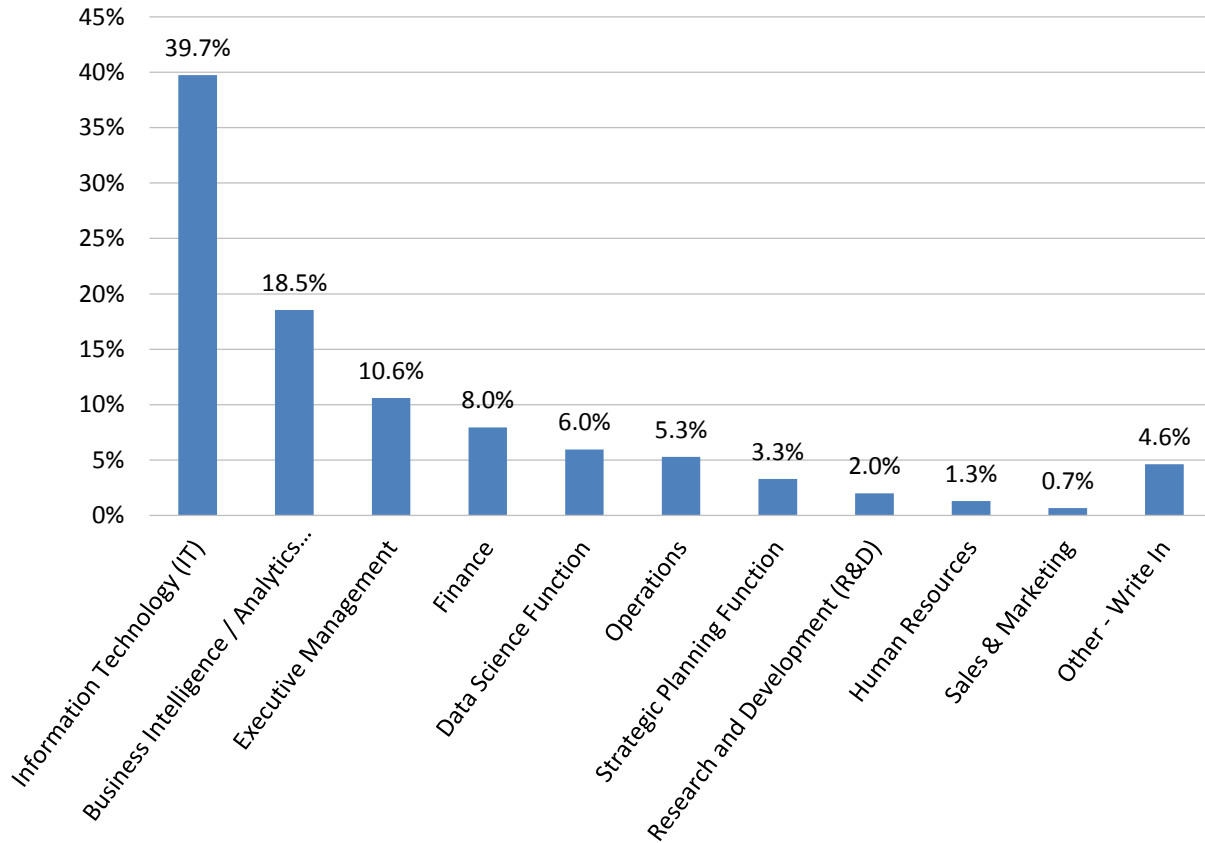


Figure 2 – Functions represented

Vertical Industries

In 2025, business services (which includes consulting, telecommunications, and transportation) led the vertical industry distribution of respondents (fig. 3), followed by technology (20 percent), manufacturing (19 percent), healthcare (9 percent), and financial services (8 percent) as the industries most represented in the sample.

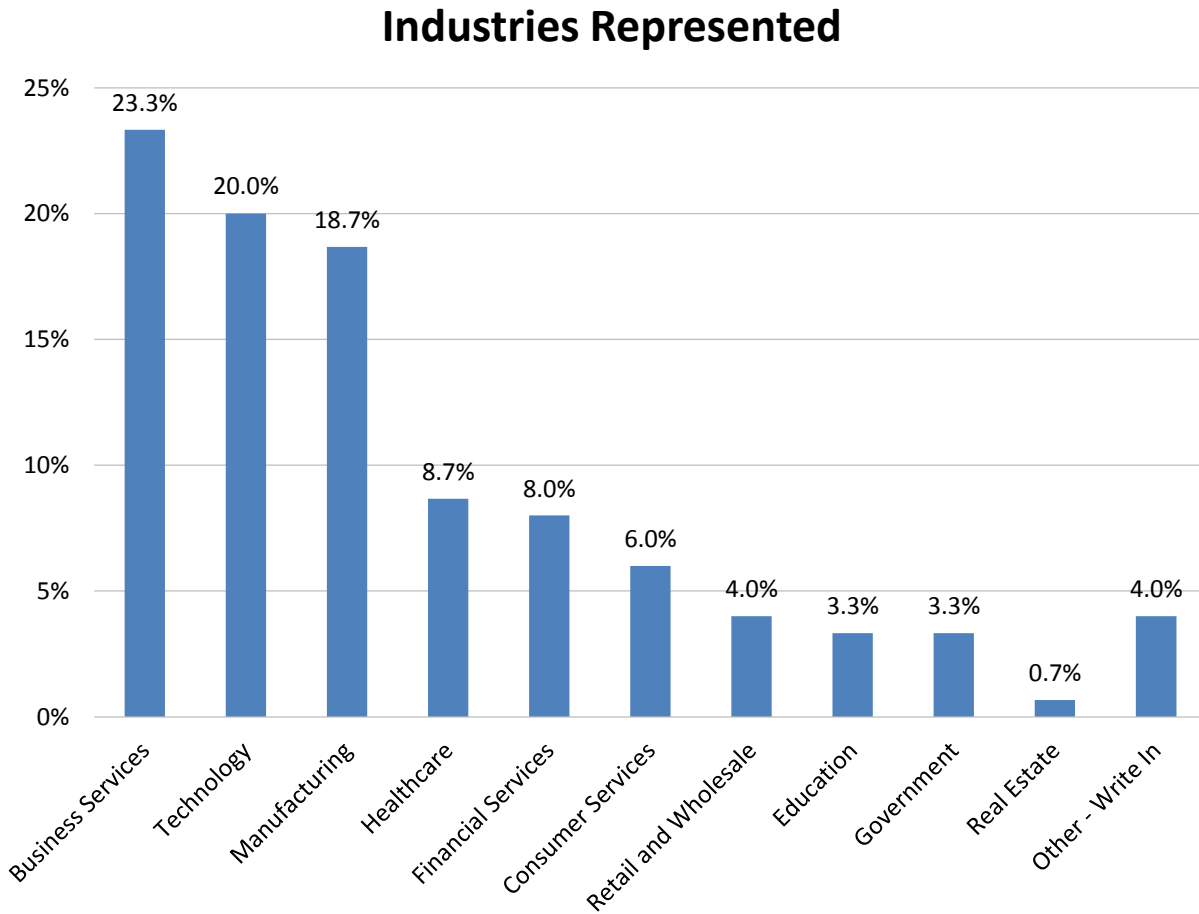


Figure 3 –Industries represented

Organization Size

In 2025, our survey includes small organizations (1-100 employees), midsize organizations (101-1,000 employees), and large organizations (more than 1,000 employees) (fig. 4). This year, small organizations account for 27 percent of our sample, midsize organizations account for 25 percent, and organizations larger than 1,000 employees account for 48 percent. Twenty-two percent of the respondents are from organizations larger than 10,000 employees.

Organization Sizes Represented

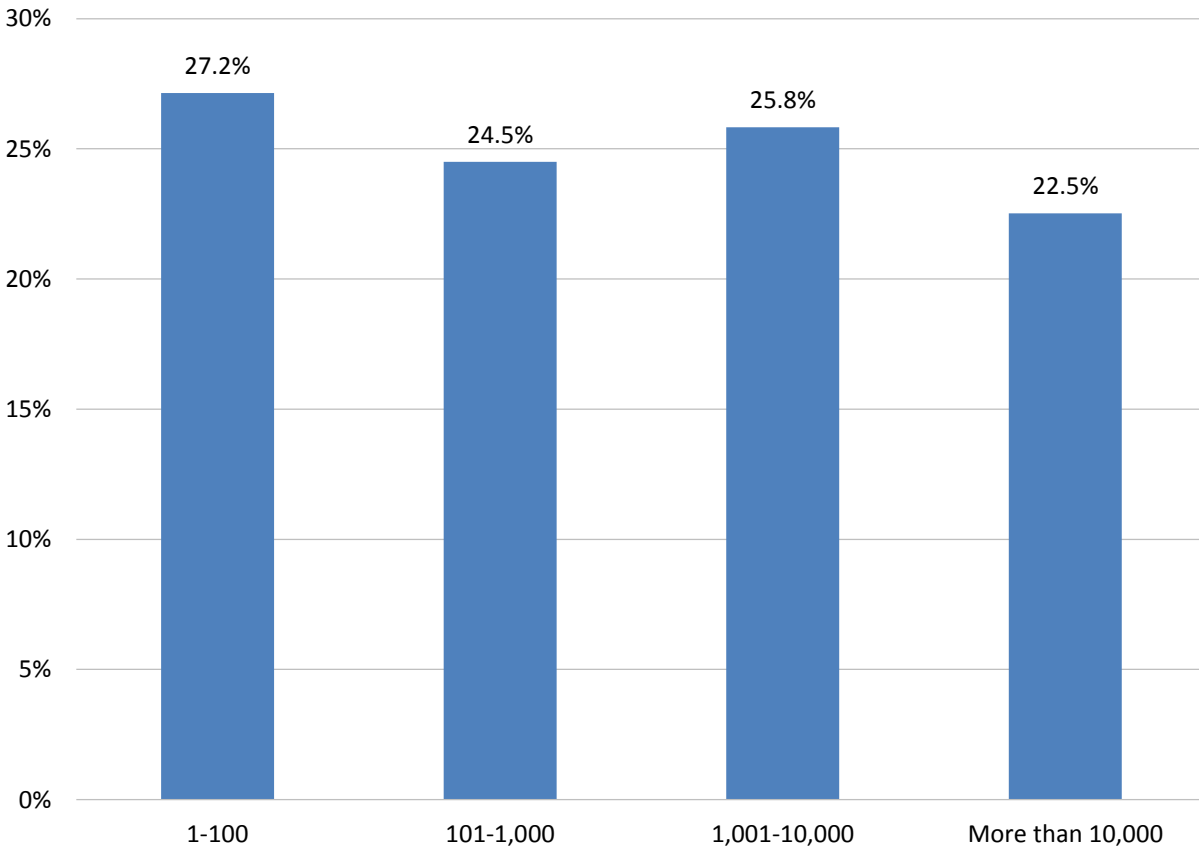


Figure 4 – Organization sizes represented



Analysis & Trends

Analysis and Trends

Importance of Data Engineering

We asked respondents about the importance of data engineering. In our 2025 survey, 44 percent of respondents indicate data engineering is of critical importance (fig. 5) while 37 percent indicate it is very important. Less than two percent indicate that data engineering is not important.

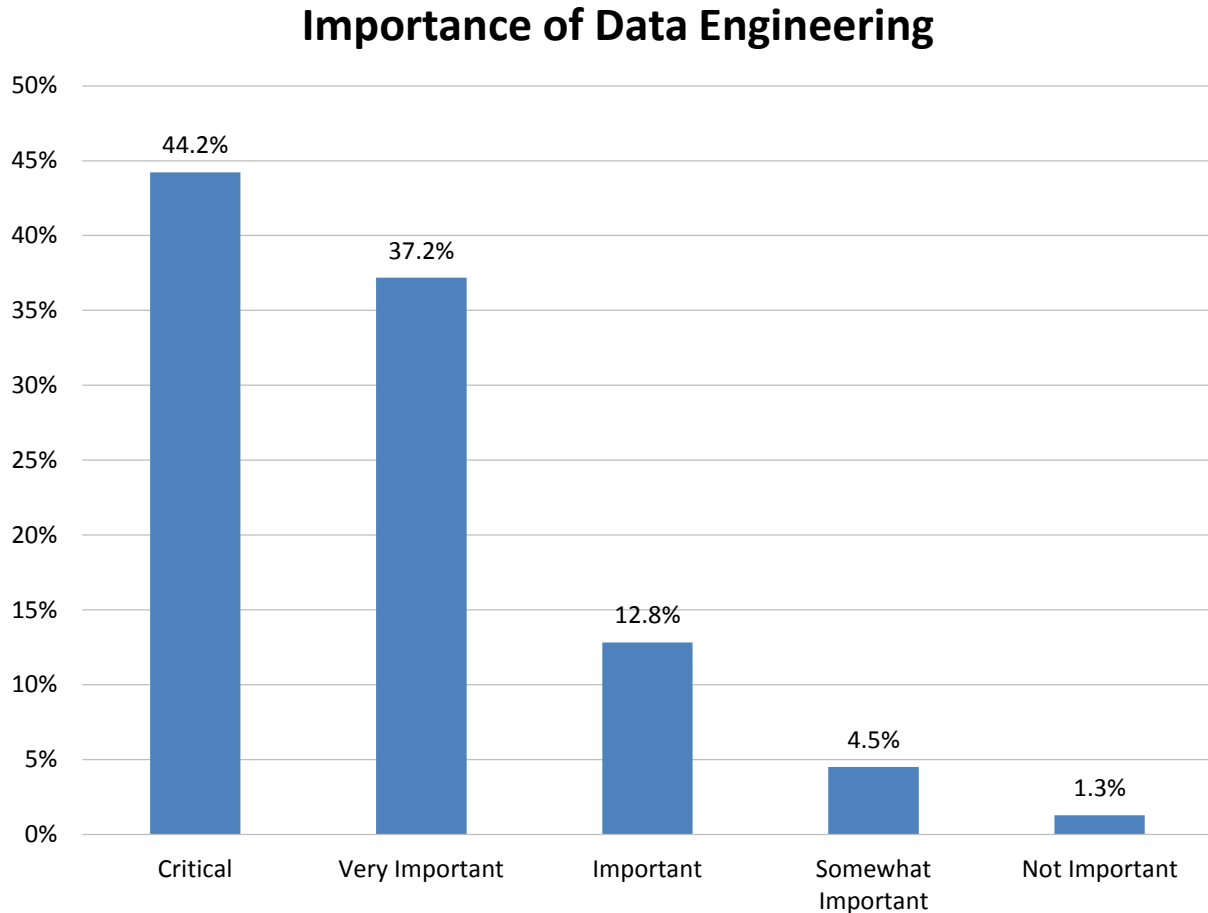


Figure 5– Importance of data engineering

Data engineering continues its year-over-year trend of escalating in importance, driven by an increase in critical importance responses (fig. 6). We expect it will remain highly important for some time to come as organizations continue to move to hybrid and multi-cloud deployments which require more complex data-engineering workflows.

Importance of Data Engineering 2015-2025

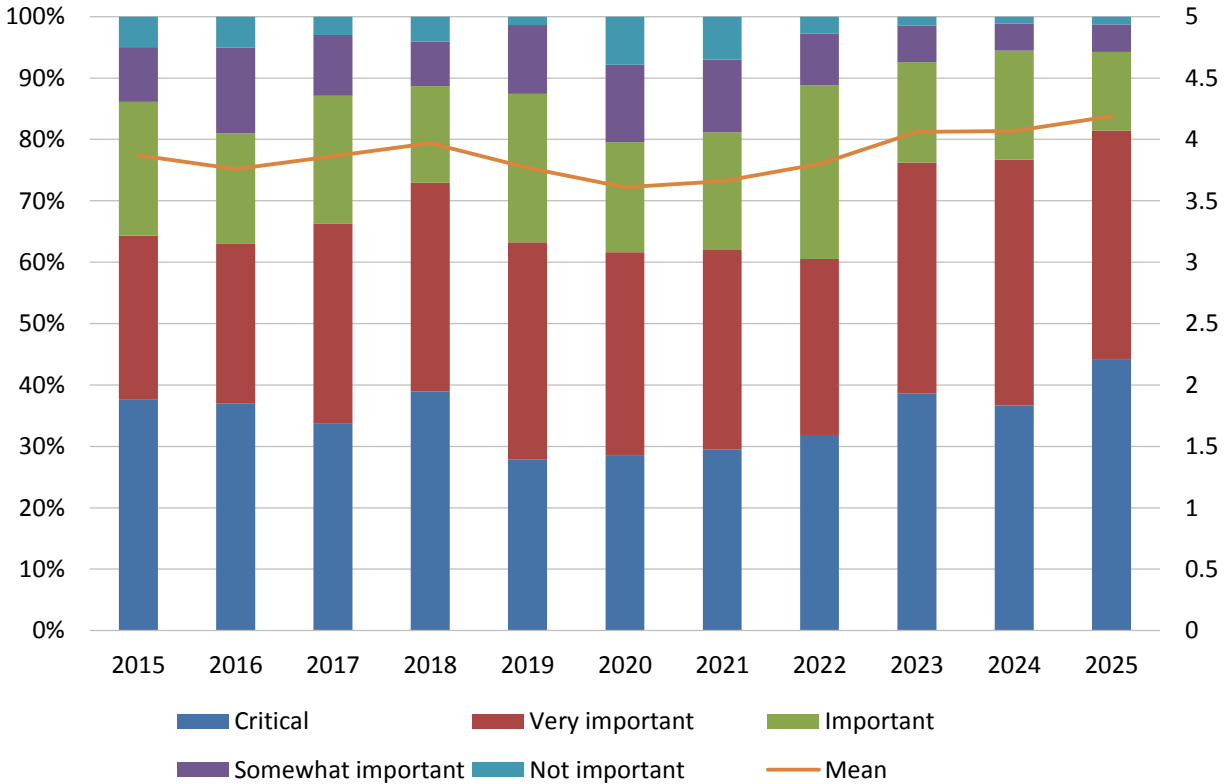


Figure 6 – Importance of data engineering 2015-2025

This year, respondents from data science functions emphatically place the highest overall importance on data engineering. Over 54 percent of survey respondents from the BICC functions indicate data engineering is of critical importance (fig. 7). Respondents from the BICC and executive management functions were more likely than other functions to place a higher level of overall importance on data engineering capabilities.

Importance of Data Engineering by Function

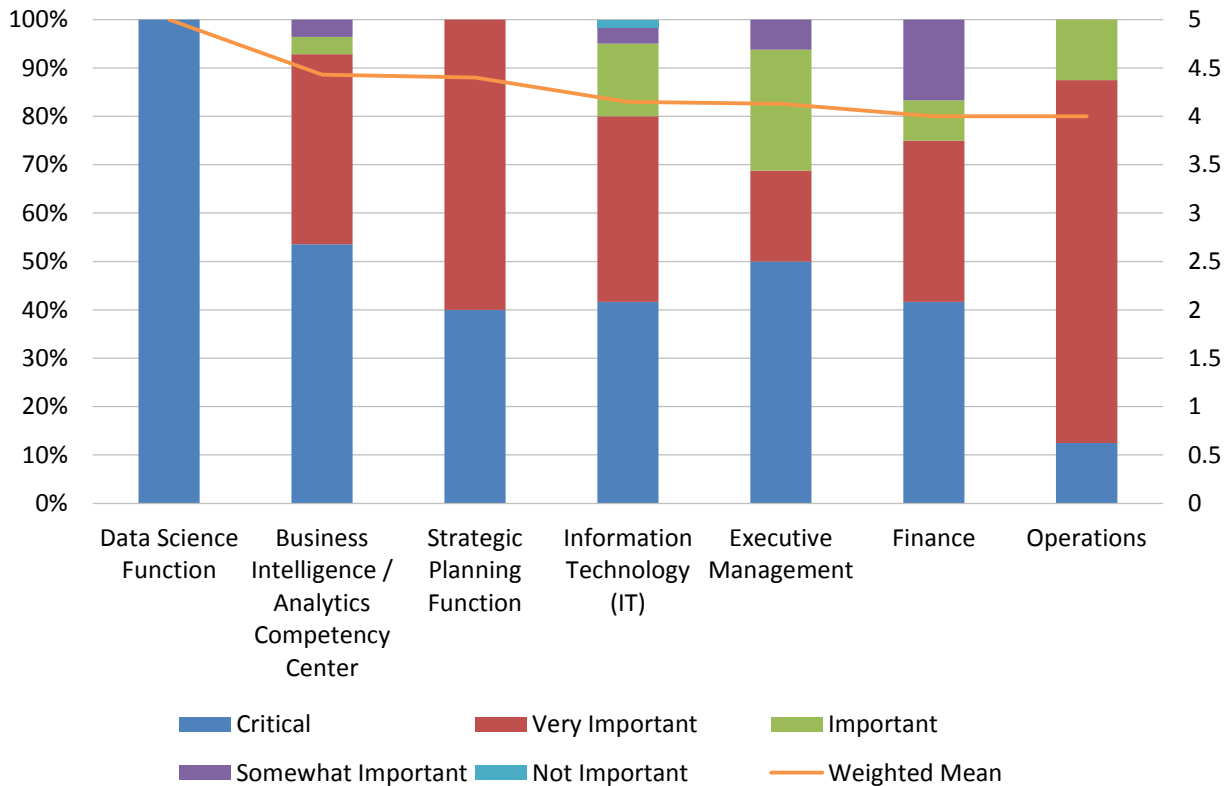


Figure 7 – Importance of data engineering by function

Data engineering is an increasing priority as organization size expands beyond 1,000 employees, likely due to the greater complexity of information infrastructure and scale of digital operations/analysis (fig. 8). Smaller organizations (1-100 employees) often use simple query interfaces (e.g., SQL via ODBC/JDBC) and spreadsheets to extract their data for analysis—perhaps into an analyst’s personal data prep and integration tool. Even so, smaller organizations’ data-integration requirements can be very complex, as they may need to analyze data originating from as many data sources as larger organizations. Unlike their smaller counterparts, however, larger organizations often have already made investments into data-engineering platforms and associated best practices and approaches.

Importance of Data Engineering by Organization Size

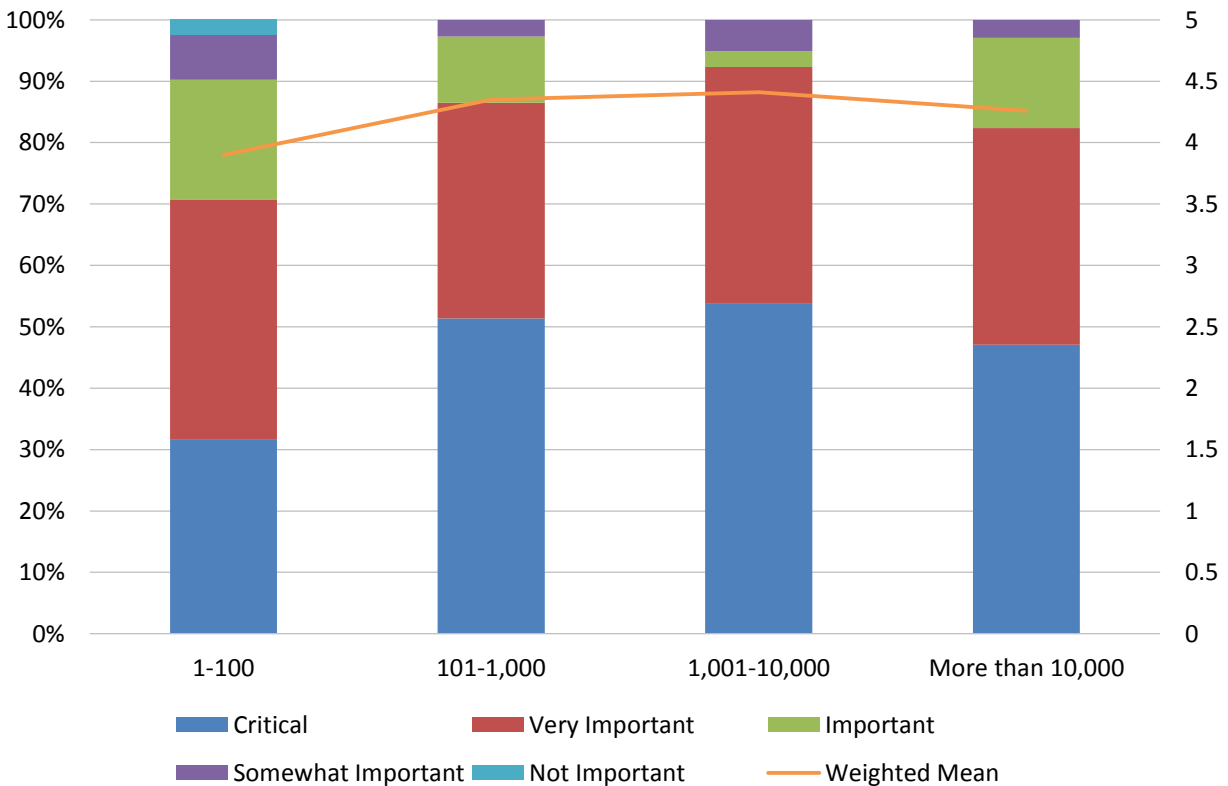


Figure 8 – Importance of data engineering by organization size

Although all industries place a high priority on data engineering, respondents from data-intensive industries like government, manufacturing, and healthcare give it the highest relative priority, followed by financial services and technology (fig. 9). Consumer services and retail and wholesale respondents gave data engineering lower importance scores.

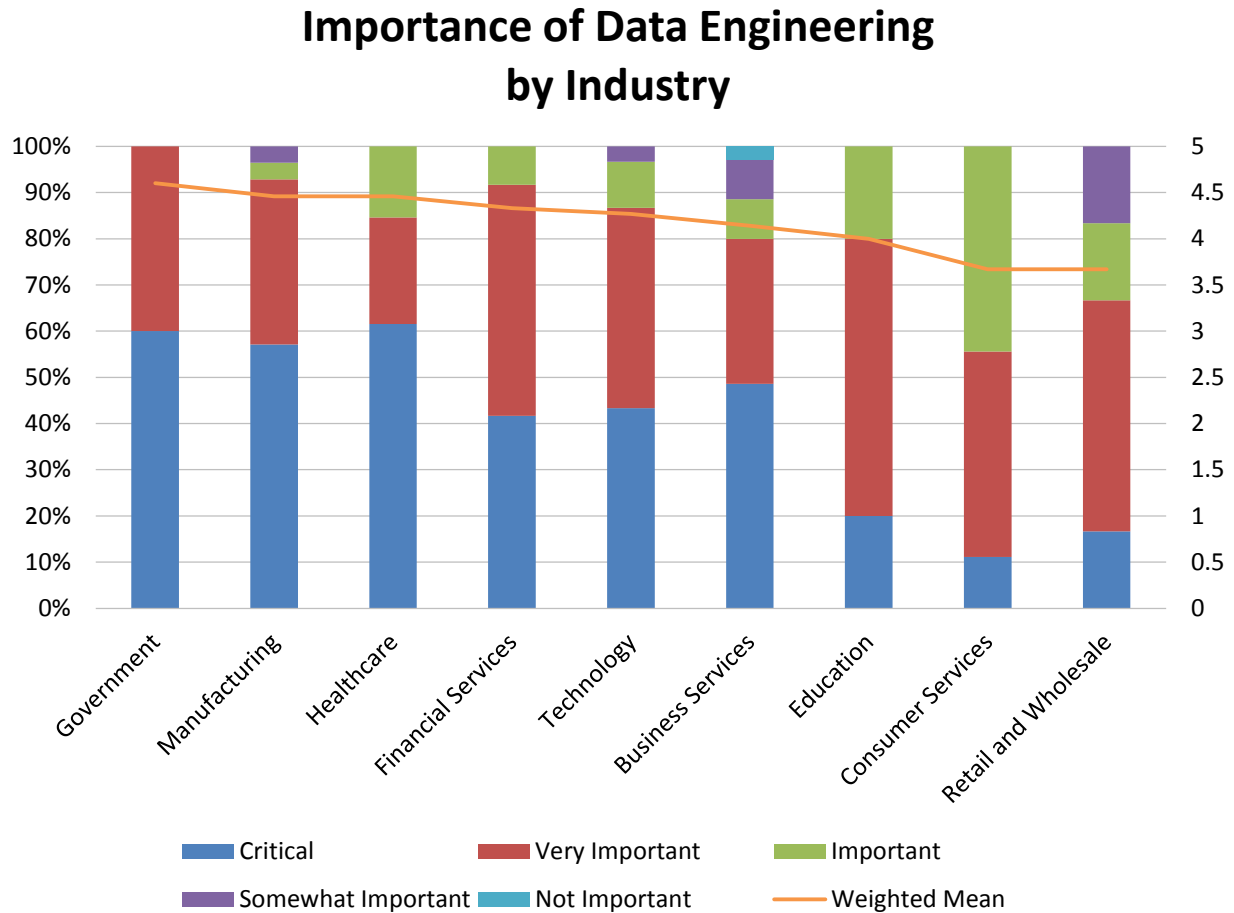


Figure 9 – Importance of data engineering by industry

Taking a data-engineering approach to accessing, combining, and preparing data speaks to a degree of maturity and sophistication in the leverage of data as a strategic asset (fig. 10). It isn't surprising, then, that respondents that rate their BI initiatives as successful place a much higher relative priority on data engineering than do those organizations which believe their BI initiatives are less successful.

Importance of Data Engineering by Success with BI

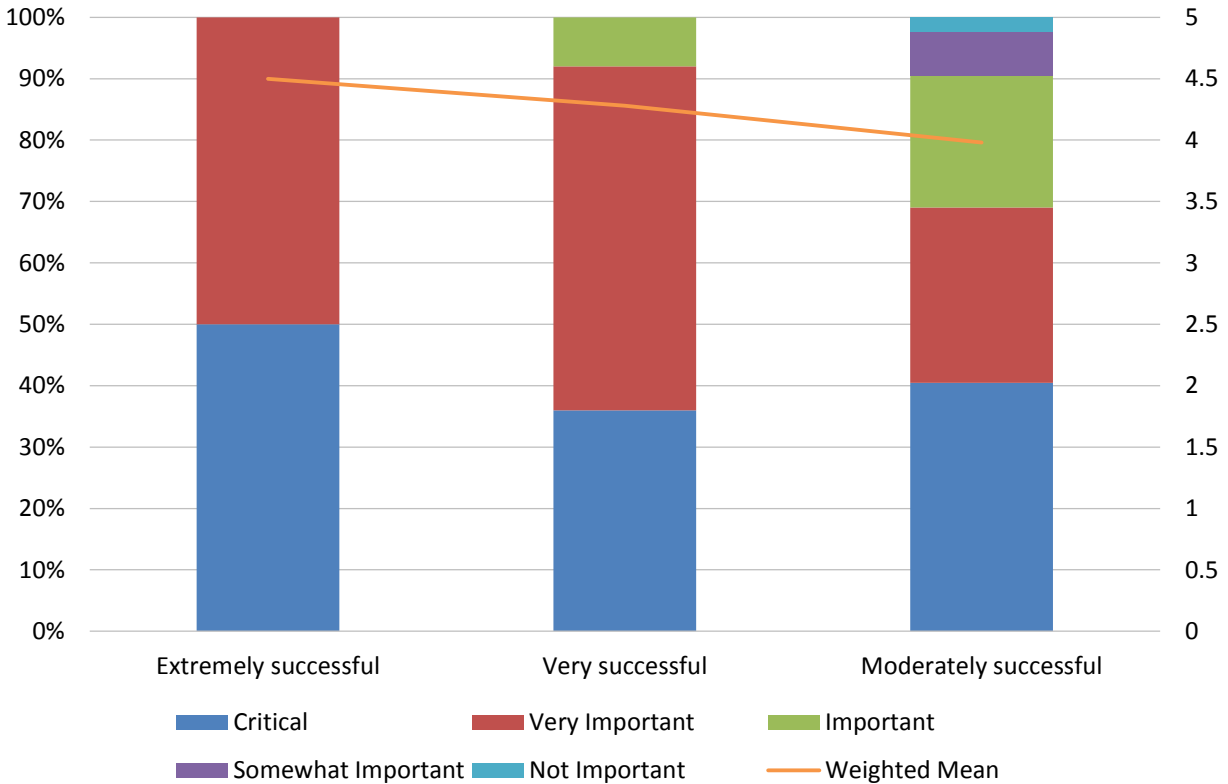


Figure 10 – Importance of data engineering by success with BI

Use and Plans for Data Engineering

When we asked respondents about their current use and plans for data engineering, 75 percent said their organizations use data-engineering capabilities today and 20 percent indicate plans to expand their use. Thirteen percent of respondents indicate plans to use data-engineering tools within the next 12 months (fig. 11). Only 7 percent have no plans at all to use data engineering capabilities.

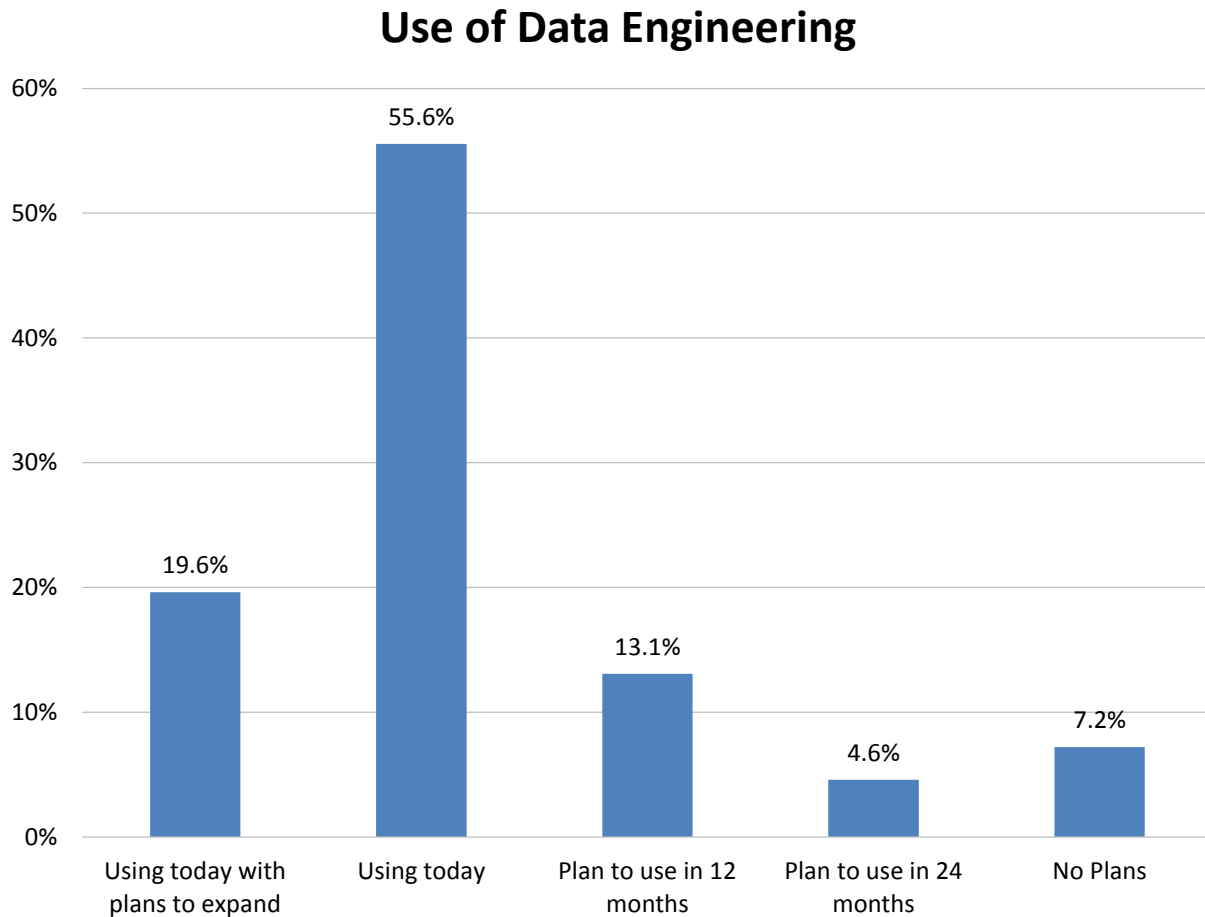


Figure 11 – Use of data engineering

The use of data engineering has risen over the past few years, however, plans to expand its use within the installed base (i.e., by those using it today) have declined slightly year over year (fig.12). Additionally, the number of those who do not use data engineering today but plan to do so in future increased slightly year over year.

Use of Data Engineering 2022-2025

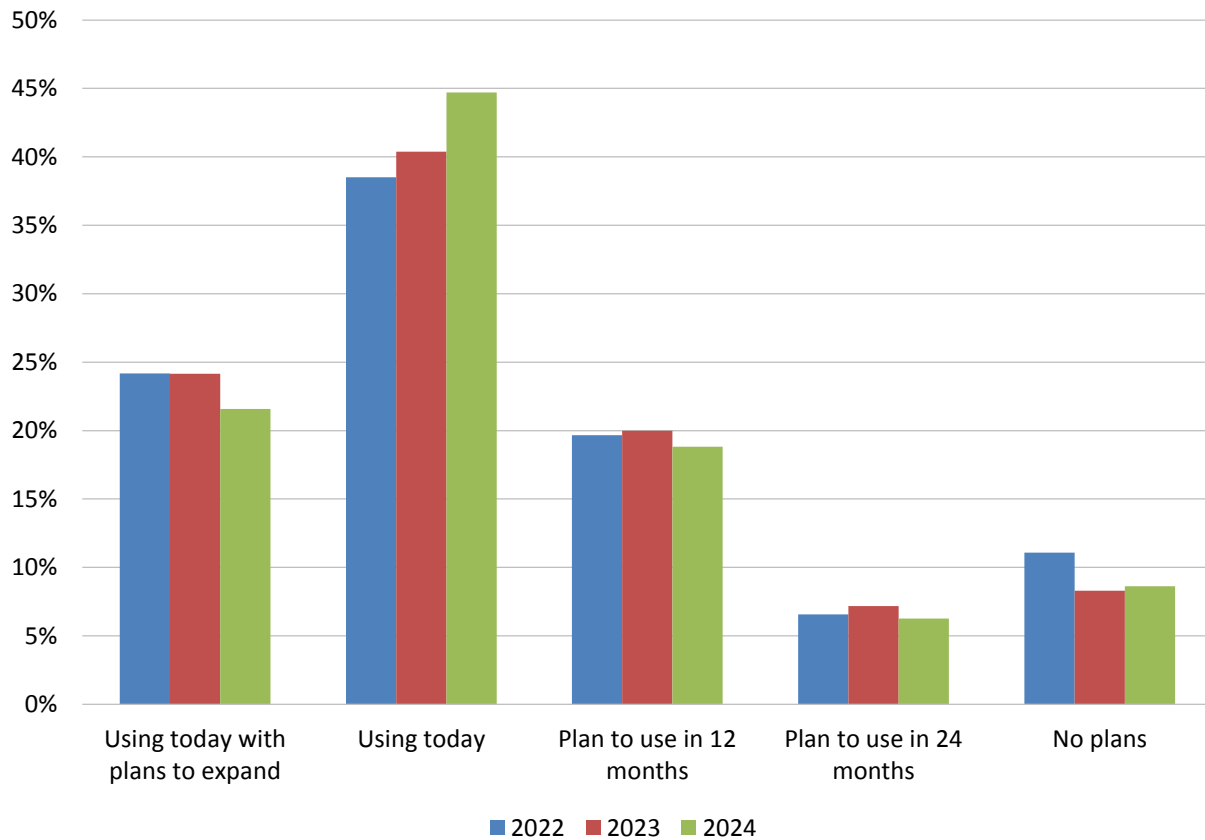


Figure 12 – Use of data engineering 2022-2025

Across geographies, Asia Pacific respondents indicated the highest levels of use today along with plans to expand their data-engineering capabilities (fig. 13). EMEA respondents show the lowest percentage who are using data engineering today and plan to expand its use. Latin American respondents indicated the highest percentage (25 percent) among those who plan to begin using data engineering within 12 months.

Use of Data Engineering by Geography

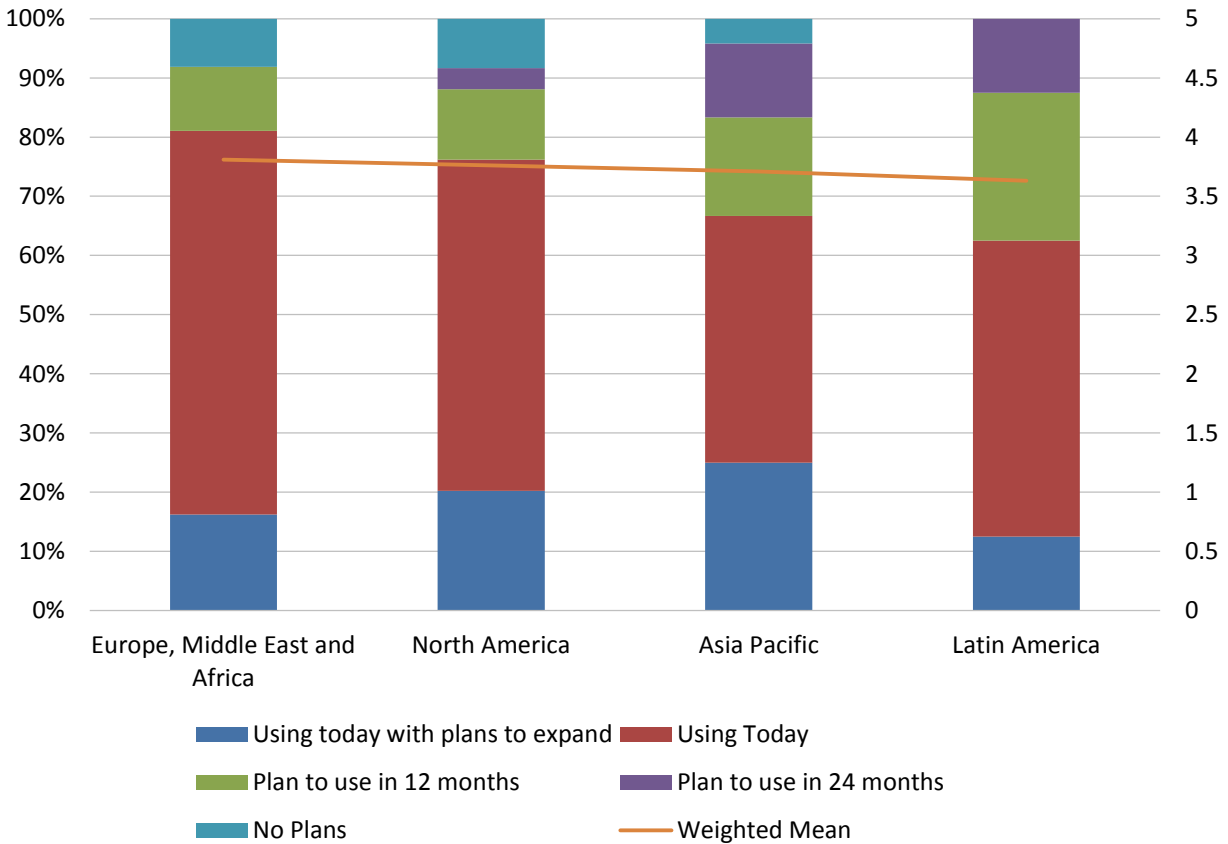


Figure 13 – Use of data engineering by geography

Survey respondents from government are among the highest users of data engineering today (fig. 14), but also have no plans to expand their usage. Consumer services and technology industry respondents have the highest percentage indicating plans to use of data engineering. Respondents from the education segment comprise the largest percentage with no plans for the use of data engineering.

Use of Data Engineering by Industry

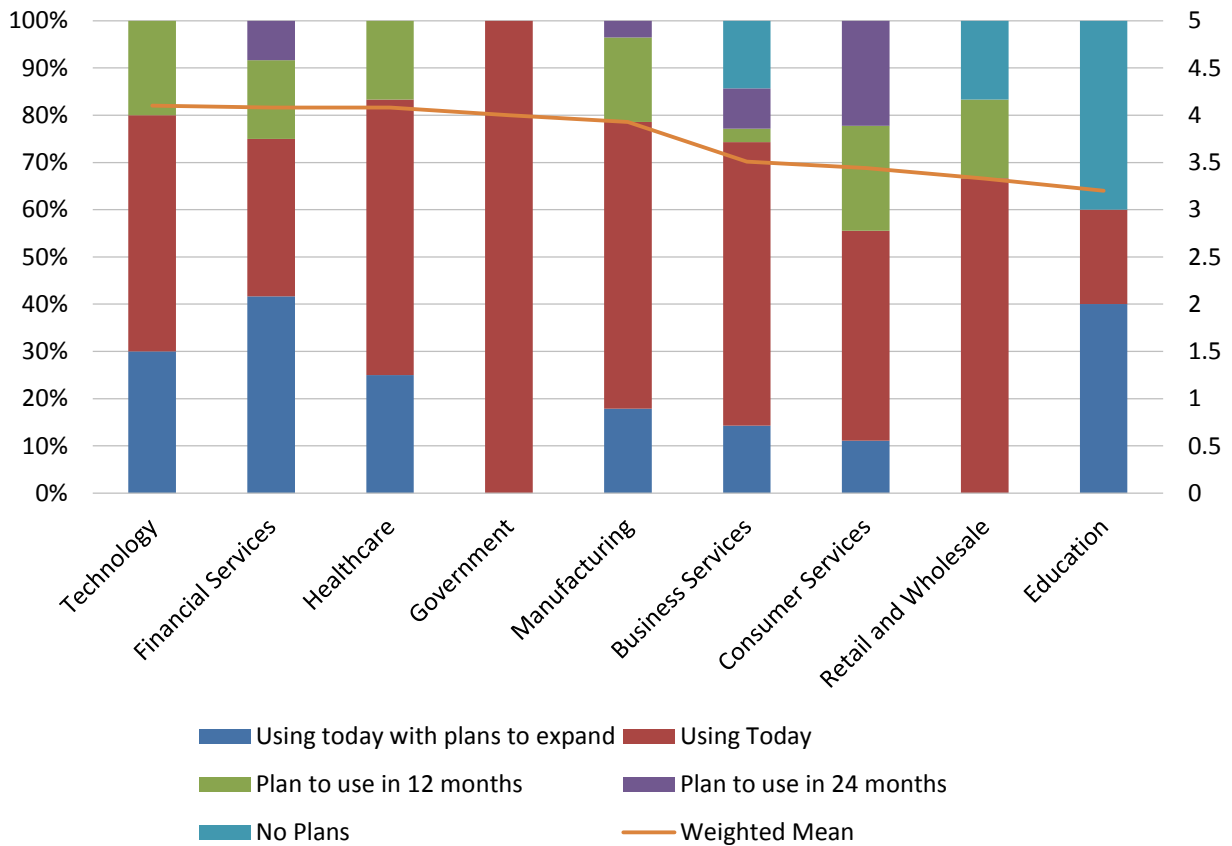


Figure 14 – Use of data engineering by industry

This year, respondents from strategic planning functions indicated a high use of data engineering today with plans to expand use (fig. 15). Only a small percentage of respondents from executive management are using data engineering today, but they also include the highest percentage (27%) of those who plan to expand its use.

Use of Data Engineering by Function

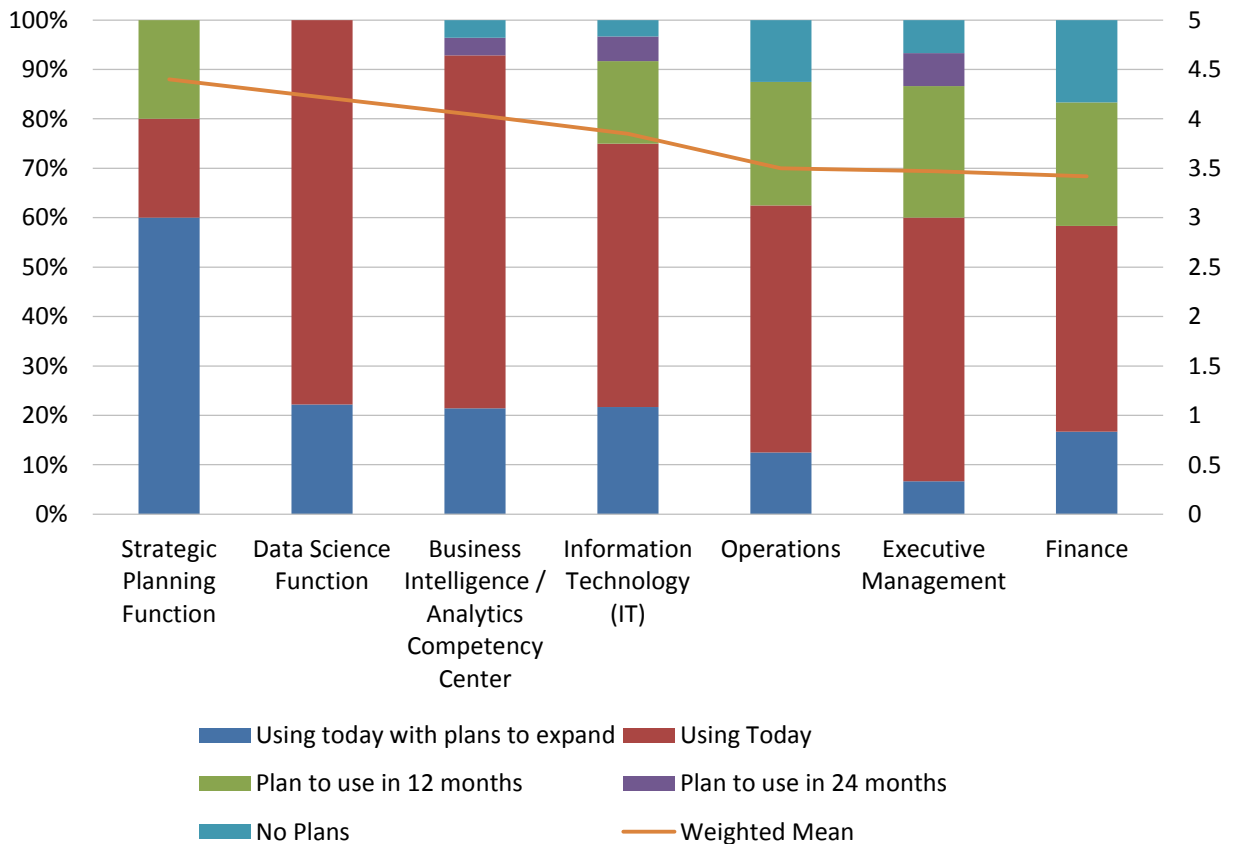


Figure 15 – Use of data engineering by function

Usage of and plans for data engineering vary by organization size (fig. 16). Smaller organizations have the lowest usage and plans for data-engineering capabilities today, but their current usage and plans scale upward along with the size of the organization. Larger organizations often require greater investment in data engineering due to the complexity and multitude of their data sources. In contrast, smaller organizations with less complex businesses may use spreadsheets to manage data and may not yet see a need for data-engineering capabilities.

Use of Data Engineering by Organization Size

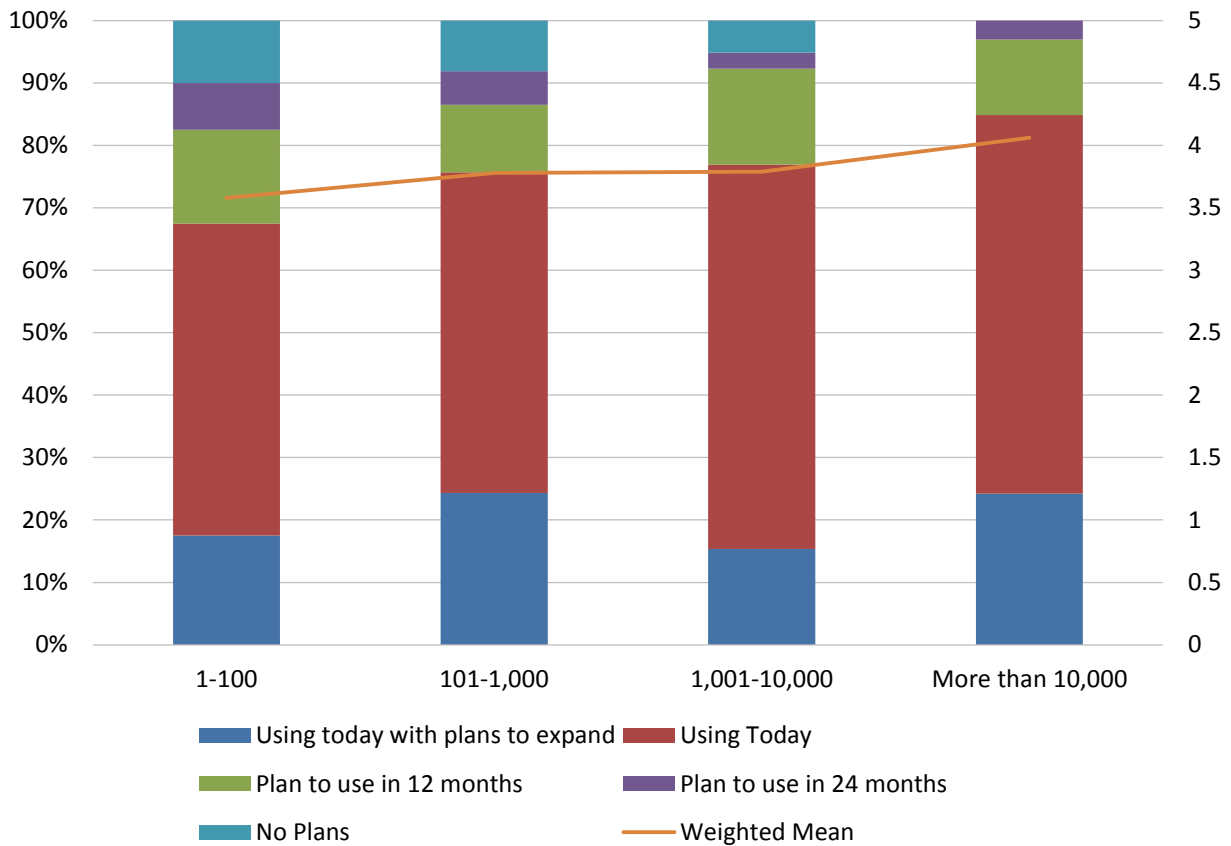


Figure 16 – Use of data engineering by organization size

Use Cases for Data Engineering

Data-engineering products are often purchased and used for multiple BI use cases (fig 17). We asked our survey respondents to indicate the estimated percentage of BI use cases for which they purchase and use data engineering products. Thirty-two percent say they employ 60 percent of their data-engineering capabilities for data integration, cleansing, and transformation workflows for a data warehouse that supports dashboards and reporting. These percentages reflect most organizations' related use of the technology within various business intelligence, analytics, and data science activities.

Data Engineering Use Cases

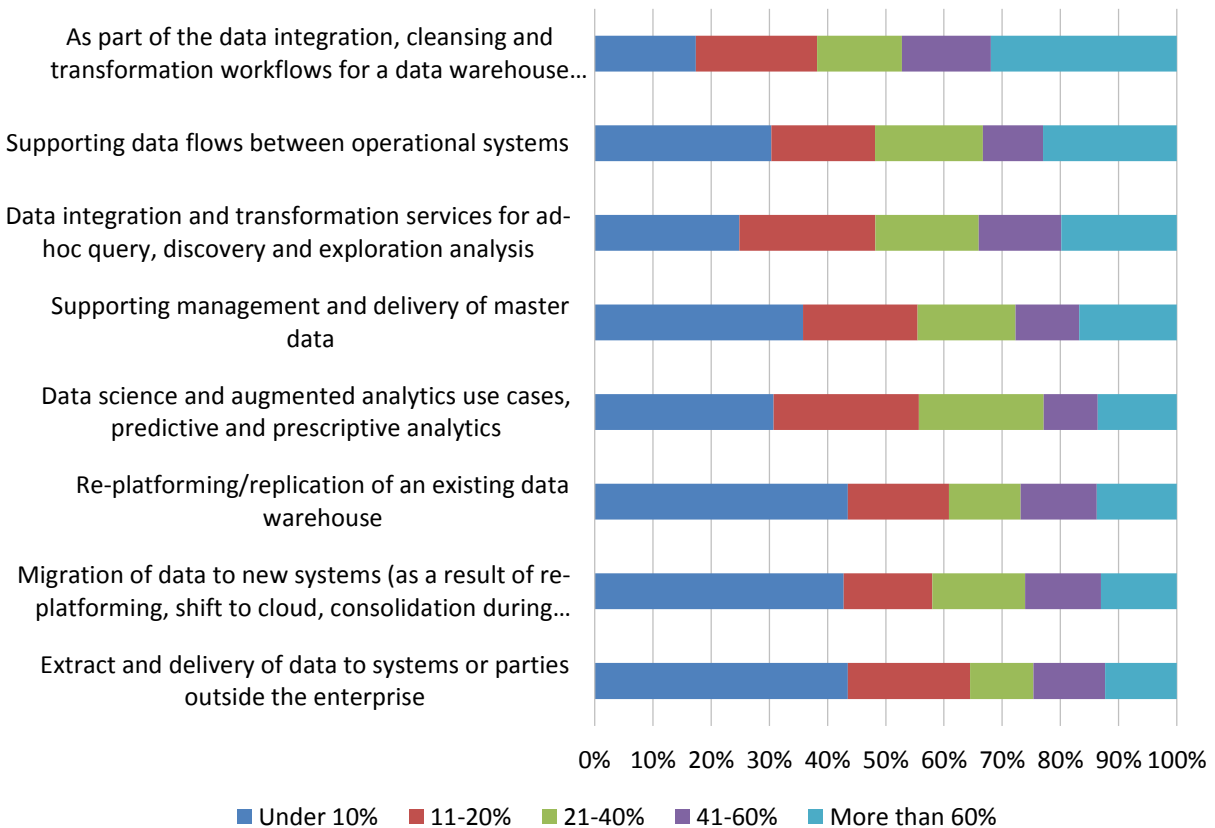


Figure 17 – Data-engineering use cases

When comparing year-over-year trends in use cases, we see a new data-engineering trend may be emerging, as evidenced by increasing use in data science and augmented analytics and re-platforming/replication of an existing data warehouse (fig. 18). The predominant use case driving data-engineering investments remains "as part of the data integration, cleansing and transformation workflows for a data warehouse supporting dashboards and reporting," but a secondary and also important use case is "data integration and transformation services for ad hoc query, discovery and exploration analysis."

Data Engineering Use Cases 2022-2025

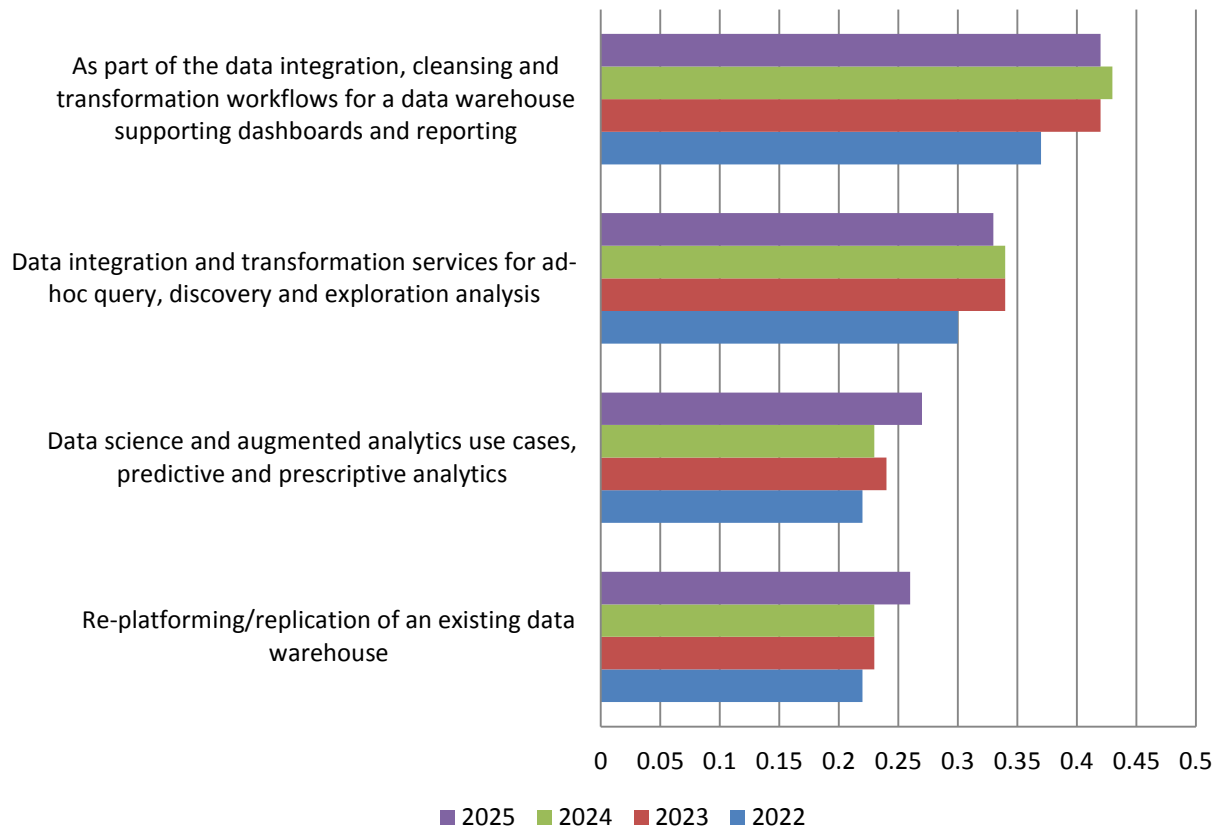


Figure 18 - Data engineering use cases 2022-2025

We filtered data-engineering usage according to the highest percentage of use case responses and geographic regions (fig. 19). The percentage of use cases varies by geography. North America and EMEA show similar percentages in data-engineering investments for supporting dashboards and reporting use cases. Latin American respondents indicate a low use of data-engineering capabilities employed for supporting data flows between operational systems.

Data Engineering Use Cases by Geography

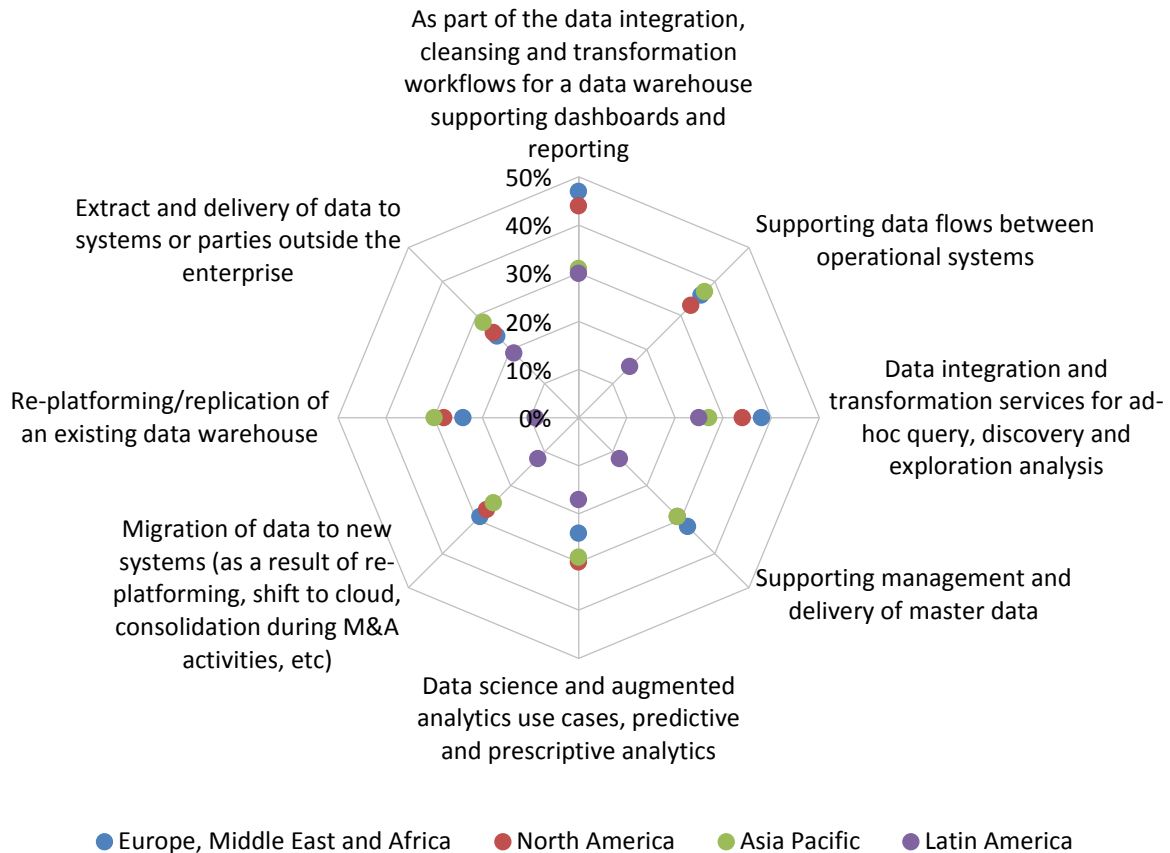


Figure 19 – Data engineering use cases by geography

Data-engineering use cases vary by industry (fig 20). The highest use case across all industries is for “as part of the data integration, cleansing and transformation workflows for a data warehouse supporting dashboards and reporting,” especially with government respondents. “Data science and augmented analytics use cases, predictive and prescriptive analytics” is the highest-ranked use case for technology industry respondents. “Supporting management and delivery of master data” and “re-platforming/replication of an existing data warehouse” are the most important use cases for the healthcare industry.

Data Engineering Use Cases by Industry

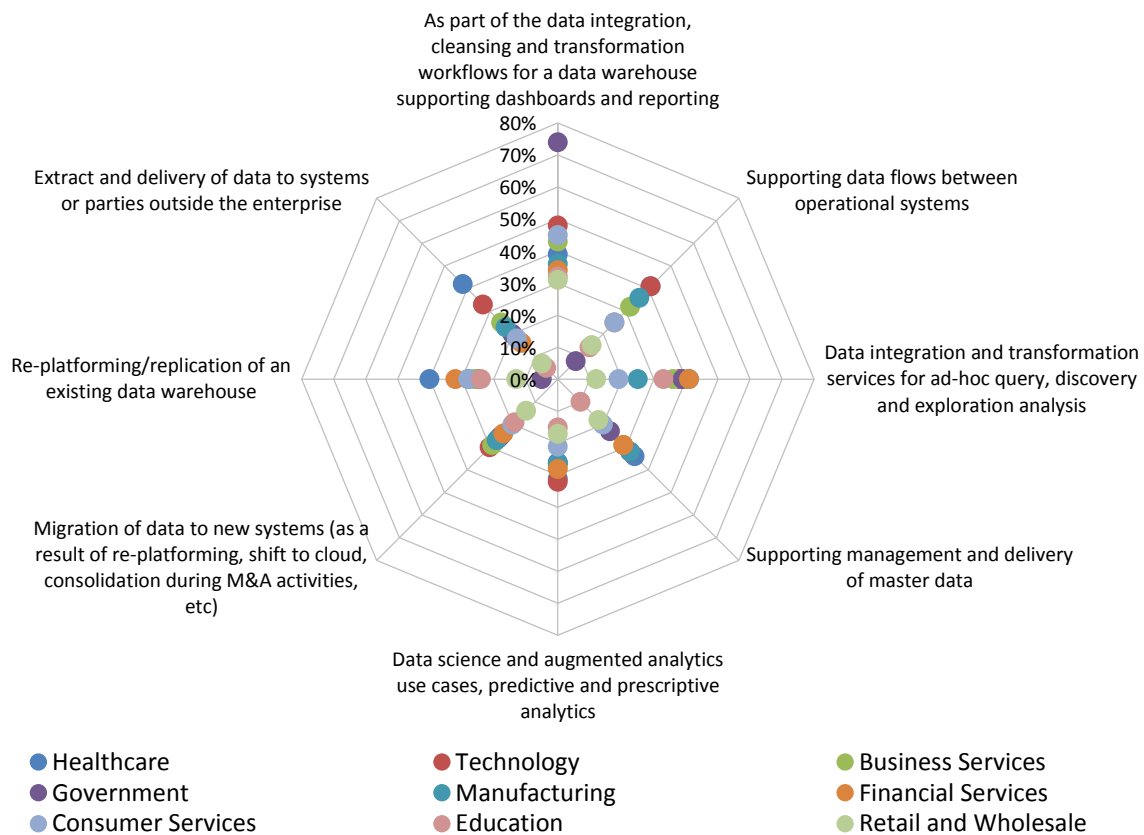


Figure 20 – Data engineering use cases by industry

Using data engineering for data integration, cleansing, and transformation workflows for a data warehouse to support dashboards and reporting is a high use case for respondents in all functions. The BICC function rates this use case the highest overall. Data science and R&D function respondents rated the use of data engineering for data science and augmented analytics higher than did other functions (fig. 21).

Data Engineering Use Cases by Function

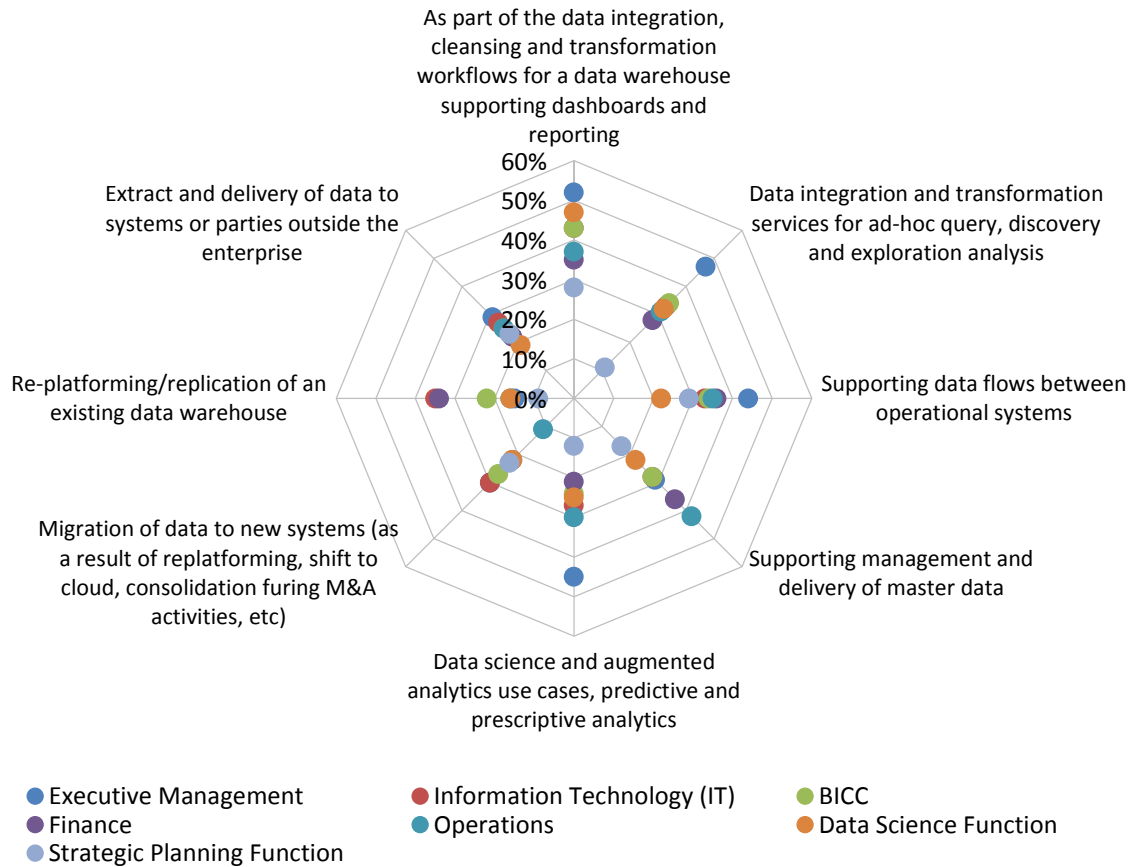


Figure 21 – Data engineering use cases by function

Data Engineering Features for Data Processing and Transformation

We asked our survey respondents to indicate the importance of the data- engineering features for processing and transforming data listed in fig. 22. We've sorted the 30 top results by the weighted mean of the importance of the features (such as Critical, Very Important, etc.). The top features for data-engineering workflows include the ability to aggregate & group data and ETL and ELT workflows, as well as the management of engineering workflows, such as alerting/job monitoring and execution plan, time-, and event-based schedulers for jobs. Least important is support for Kafka and Apache big data services.

Features for Data Processing and Transformation

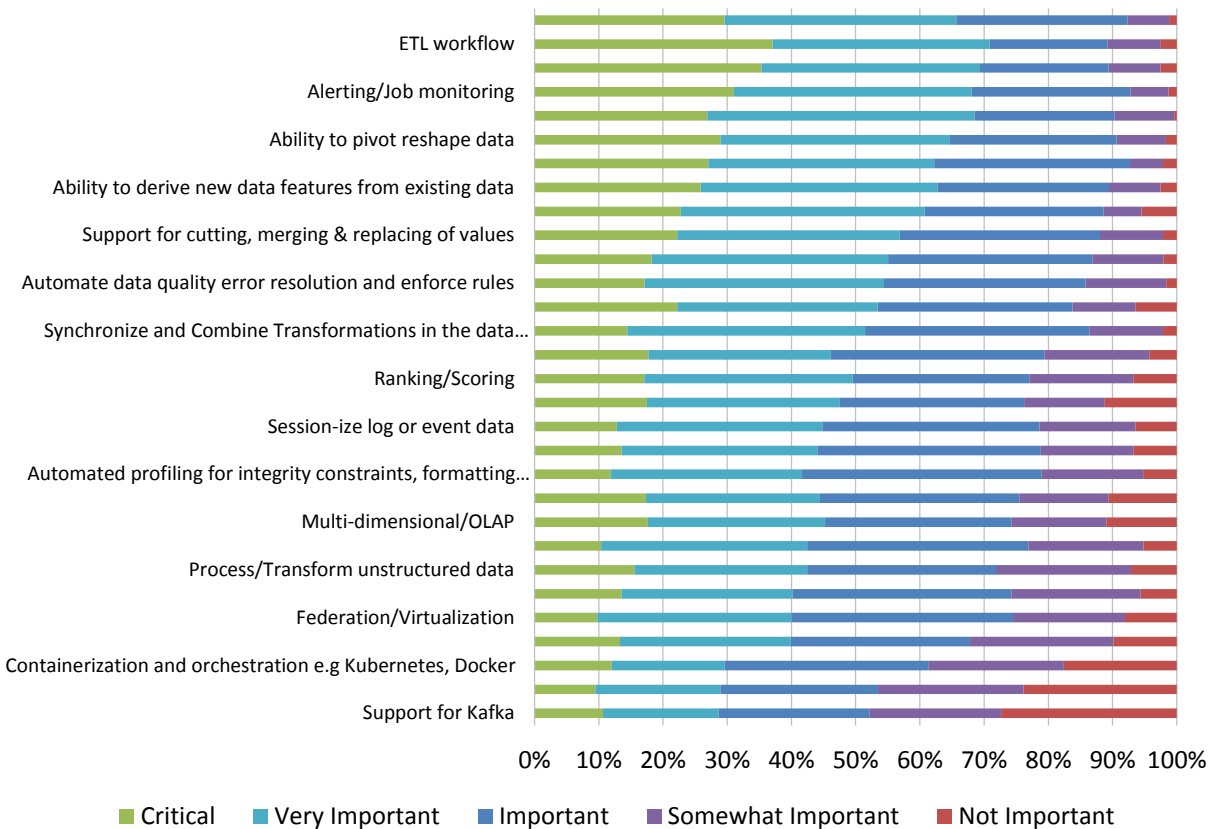


Figure 22 – Features for data processing and transformation

Overall, our survey responses show a trend of data-driven organizations investing strategically in modern data-engineering capabilities, with a clear emphasis on real-time data, automation, data quality, and cloud-native technologies (fig. 23).

- While ETL and ELT workflow already were relatively mature in 2023, these core data transformation processes continue to increase in importance, reflecting ongoing optimization efforts.
- Support for Kafka increased in importance significantly from 2023 to 2025. This highlights a strong emphasis on real-time data streaming and processing, which is crucial for modern data architectures.
- Likewise, growth in containerization and orchestration reflects the industrywide shift towards cloud-native technologies and the need for scalable and manageable data pipelines.
- Also of note, embed advanced analytics models in data transformation workflow continues to rise in importance year over year.

Data-engineering features vary in importance in minor ways across geographies. Here are the top 15 data engineering features according to the geographic location of our survey respondents. The chart suggests a global trend toward modern data-engineering practices, but with regional variations in implementation priorities and maturity levels. North America's responses indicate a potentially more advanced data-engineering landscape, while other regions are catching up, focusing on specific areas aligned with their needs.

Features for Data Processing and Transformation by Geography

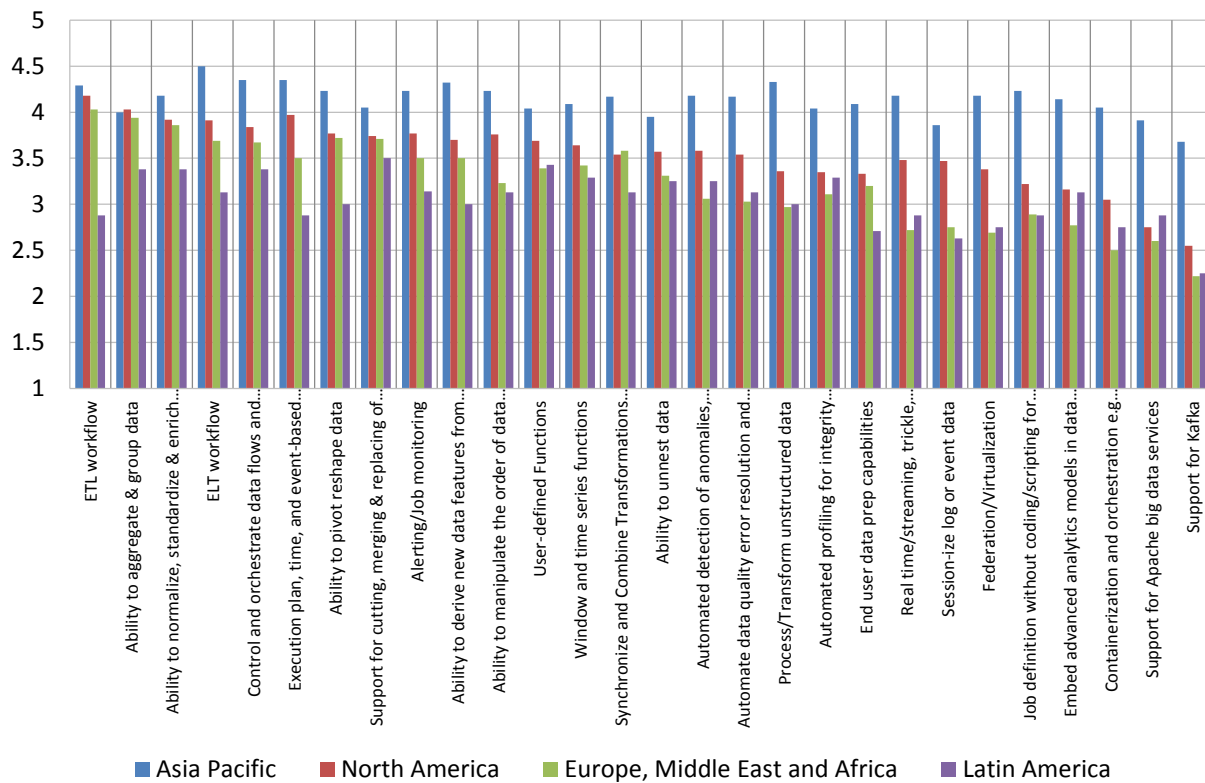


Figure 23 - Features for data processing and transformation by geography

Here we rank the top 15 data engineering features according to their importance by our survey respondents' industry (fig. 24). Overall, the chart highlights a significant disparity in data-engineering feature priorities across industries. Technology- and data-driven sectors like technology and financial services lead in terms of ranking importance, while others like government and education appear to be in the earlier stages of adoption, possibly due to varying levels of investment or different data needs.

Features for Data Processing and Transformation by Industry

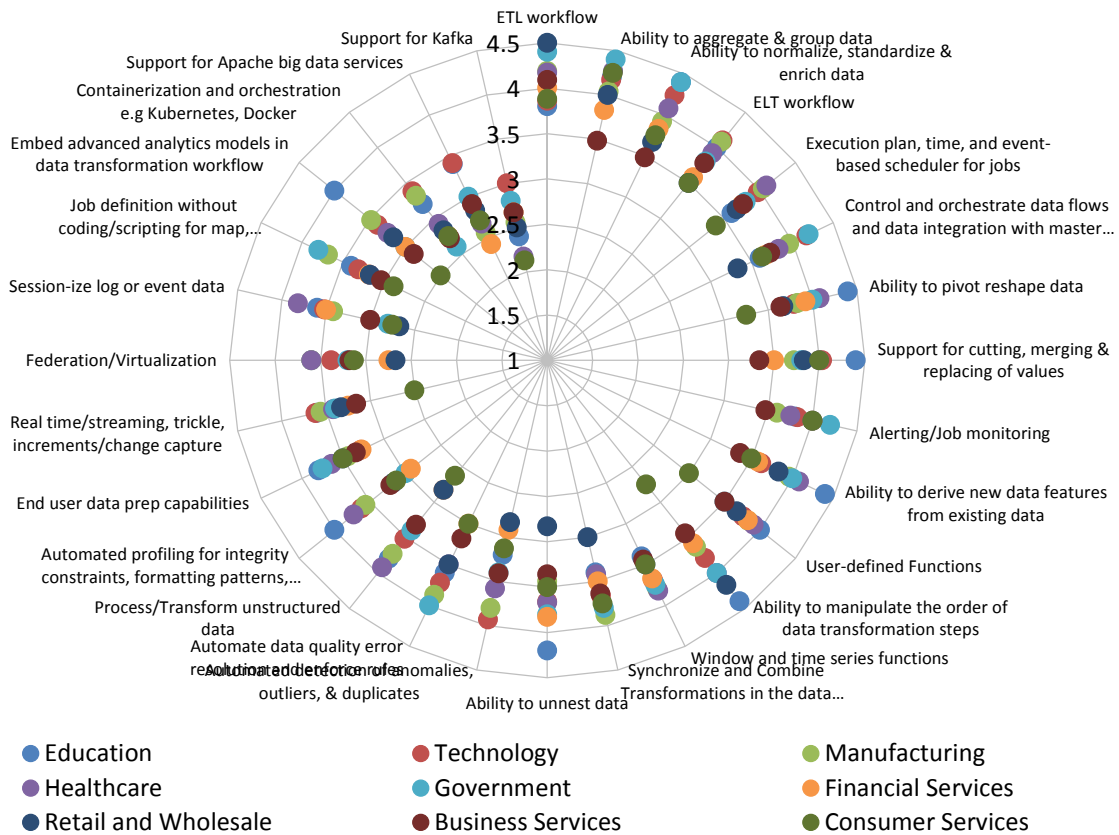


Figure 24 – Features for data processing and transformation by industry

The top-rated data-engineering features vary slightly across functions. Some functions, such as the Business Intelligence/Analytics Competency Center and information technology (IT), have a higher need than others for more features because they're very data-driven and rely heavily on data-engineering capabilities. On the other hand, functions like executive management and finance seem to place less importance on certain features, perhaps indicating a lower level of understanding of the importance of data engineering needs (fig. 25).

Features for Data Processing and Transformation by Function

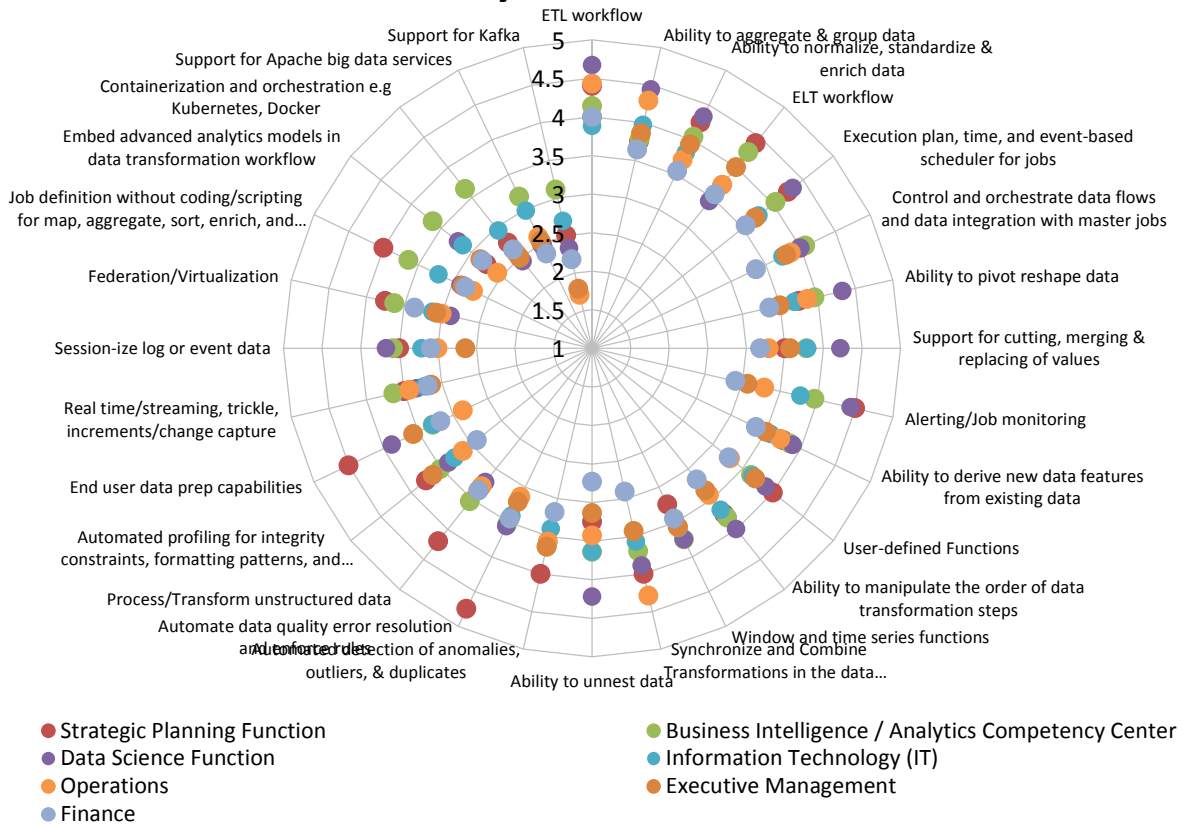


Figure 25 – Features for data processing and transformation by function

The survey responses in fig. 26 illustrate the priority assigned to data-engineering features across organizations of varying sizes. Larger organizations (more than 10,000 employees) consistently give the highest priority to nearly all features, notably in support for Kafka, containerization and orchestration, and automated profiling. Organizations with 1,001-10,000 employees generally rank second in priorities, closely trailing the largest organizations. Small organizations (1-100 employees) generally give the lowest priority to capabilities and score most features lower, possibly because they have limited resources or a different focus.

The data suggests a correlation between organization size and data-engineering maturity. Larger organizations likely have greater resources and more complex data needs, and they tend to invest more heavily in advanced data-engineering capabilities. Smaller organizations may prioritize differently or face resource constraints, resulting in a less-mature adoption of these features.

Features for Data Processing and Transformation by Organization Size



Figure 26 – Features for data processing and transformation by organization size

Frequency of Data Engineering

Our survey asked respondents, “How frequently do people in your organization engage in data engineering prior to analysis?” Forty-seven percent of respondents engage in data engineering constantly or frequently (fig. 27). Twenty-two percent use data engineering occasionally, and less than 11 percent use data engineering rarely or never.

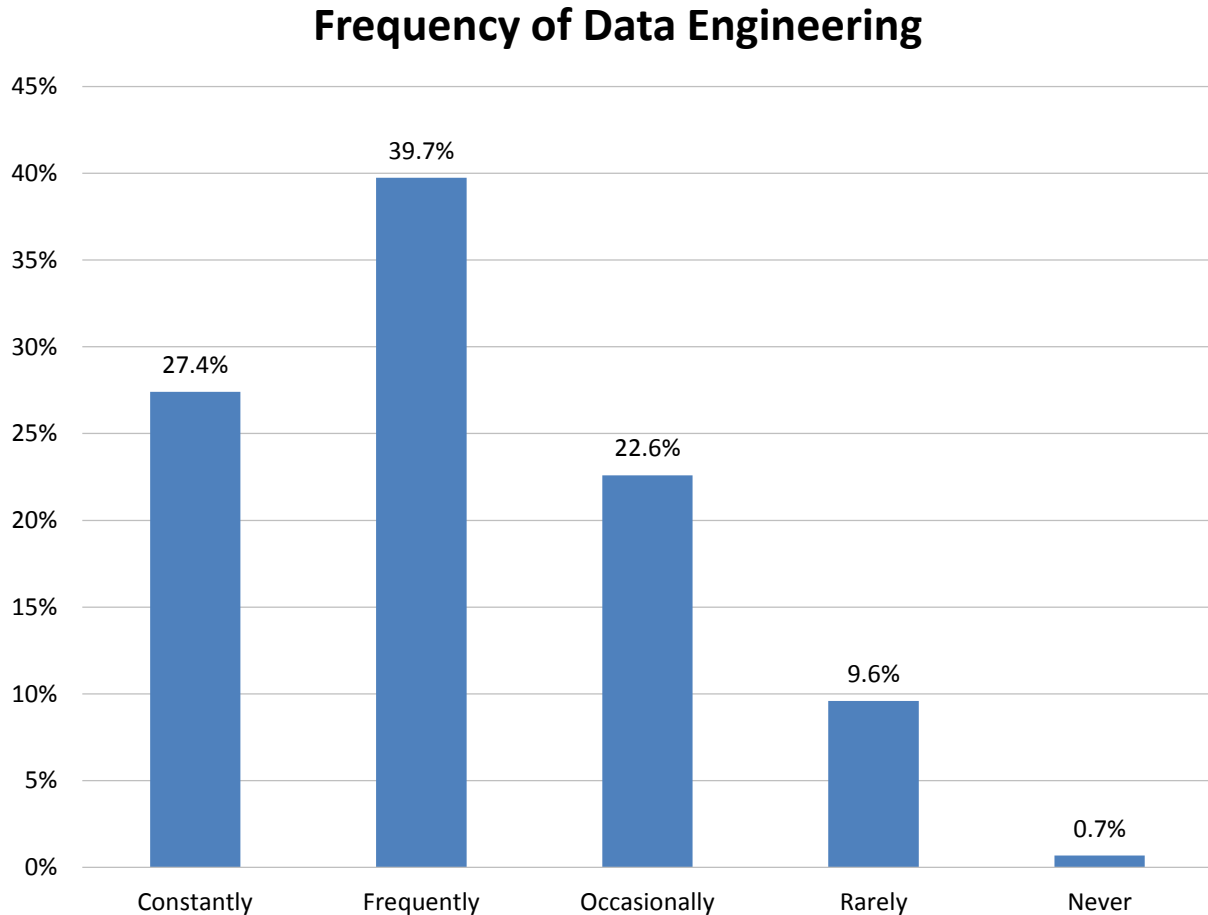


Figure 27 – Frequency of data engineering

The stacked bar chart in fig. 28 depicts the frequency of data-engineering practices from 2015 to 2025, revealing a clear trend: The constant and frequent use of data engineering has remained relatively constant over time. Conversely, rare and never responses have dwindled over the past four years, indicating wider adoption and integration of data engineering across organizations. The shift toward more frequent and constant practices highlights the evolving data landscape and the need for robust data infrastructure and processes to support business operations and decision making.

Frequency of Data Engineering 2015-2025

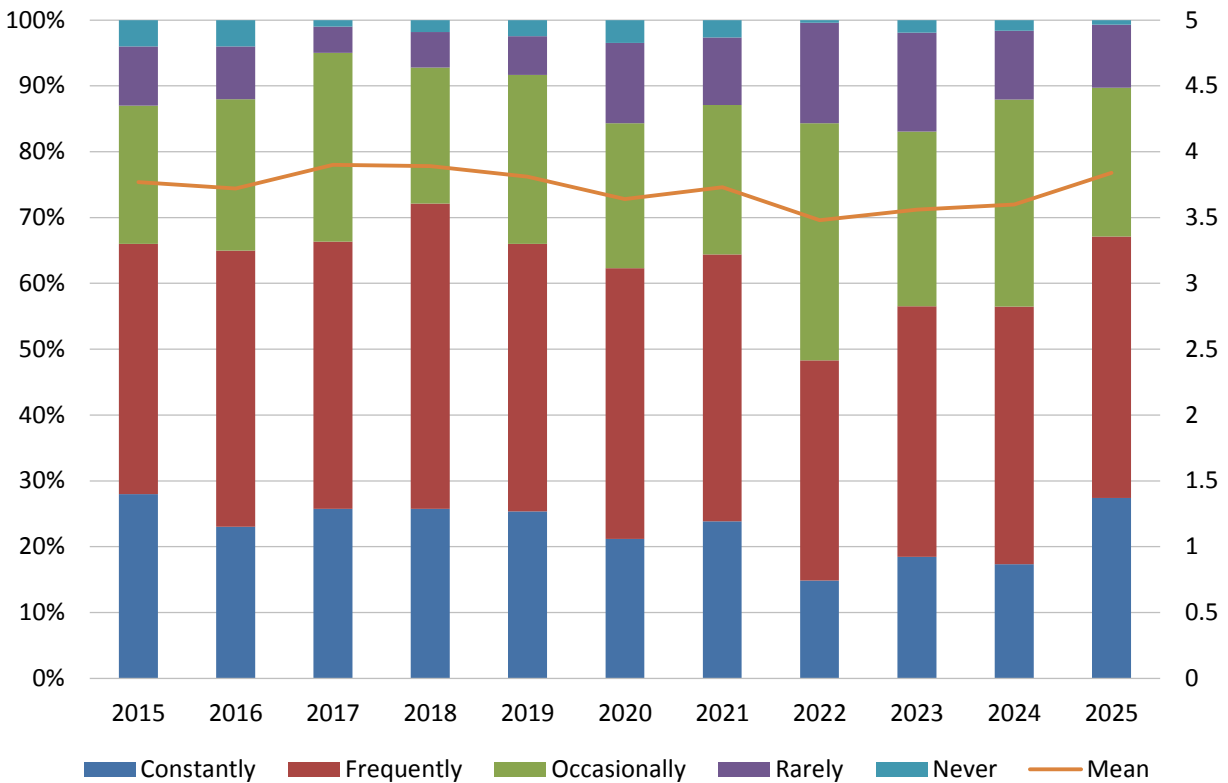


Figure 28 - Frequency of Data Engineering 2015-2025

Looking at survey data broken out by respondents' industry, those in healthcare show the highest use of constant and frequent data engineering (fig. 29). Technology industry responses show a more balanced distribution across frequency levels. Manufacturing and business services indicated occasional usage. Financial services and retail and wholesale exhibit a more scattered frequency of use for data engineering.

Frequency of Data Engineering by Industry

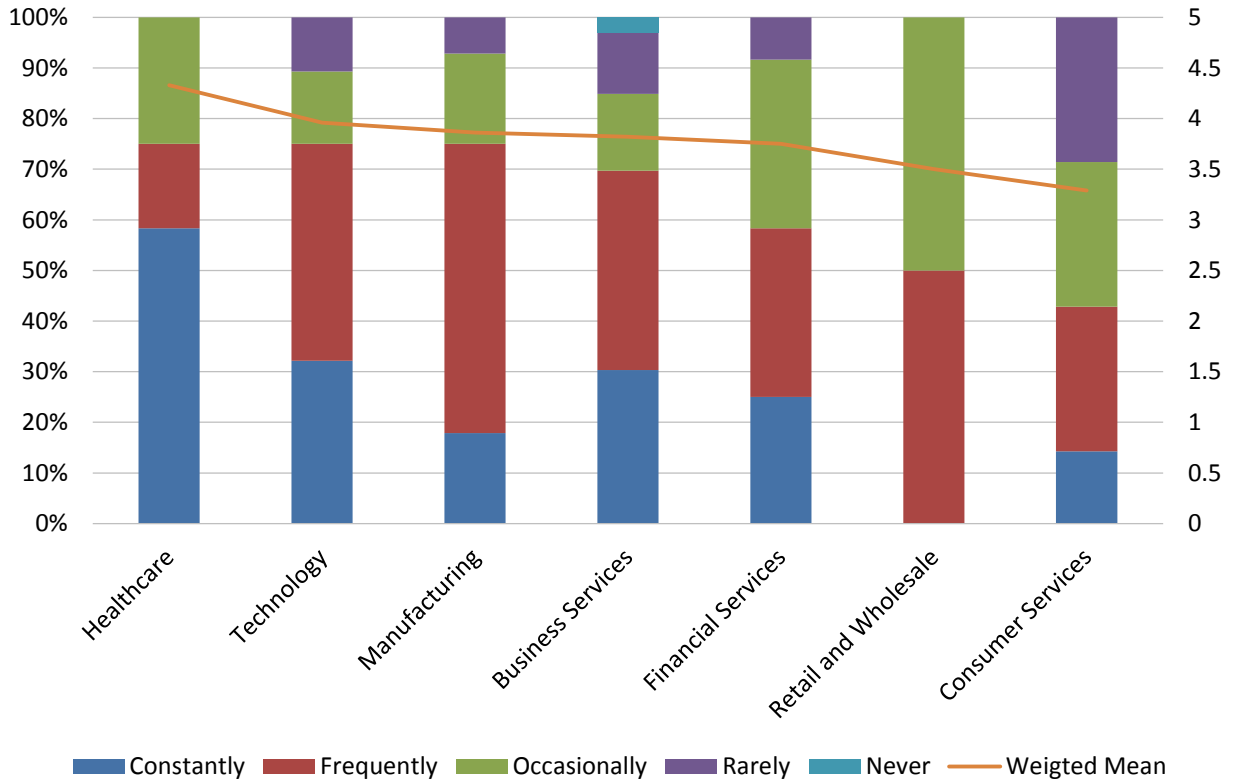


Figure 29 – Frequency of data engineering by industry

Slicing the responses on frequency of data engineering by function, we find some interesting responses (fig. 30). The strategic planning, functions, the BICC, and data science functions indicate the most use of data engineering constantly.

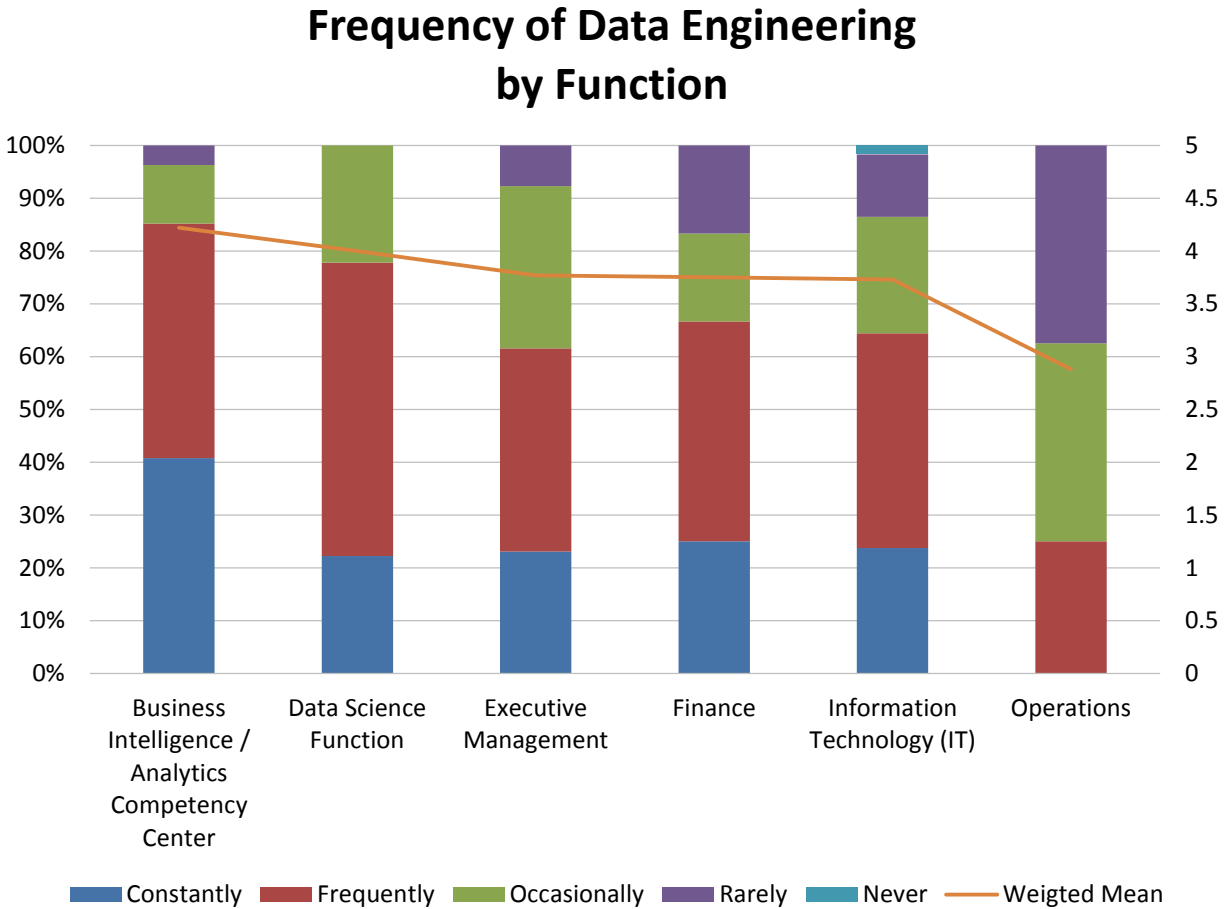


Figure 30 – Frequency of data engineering by function

Because larger organizations have more complex operational and analytical data infrastructures, they generally employ data engineering more frequently than smaller organizations (fig. 31). Of particular interest is our finding that many organizations report that they rarely or never use data engineering, perhaps due to their extensive reliance on spreadsheets to run and analyze the business.

Frequency of Data Engineering by Organization Size

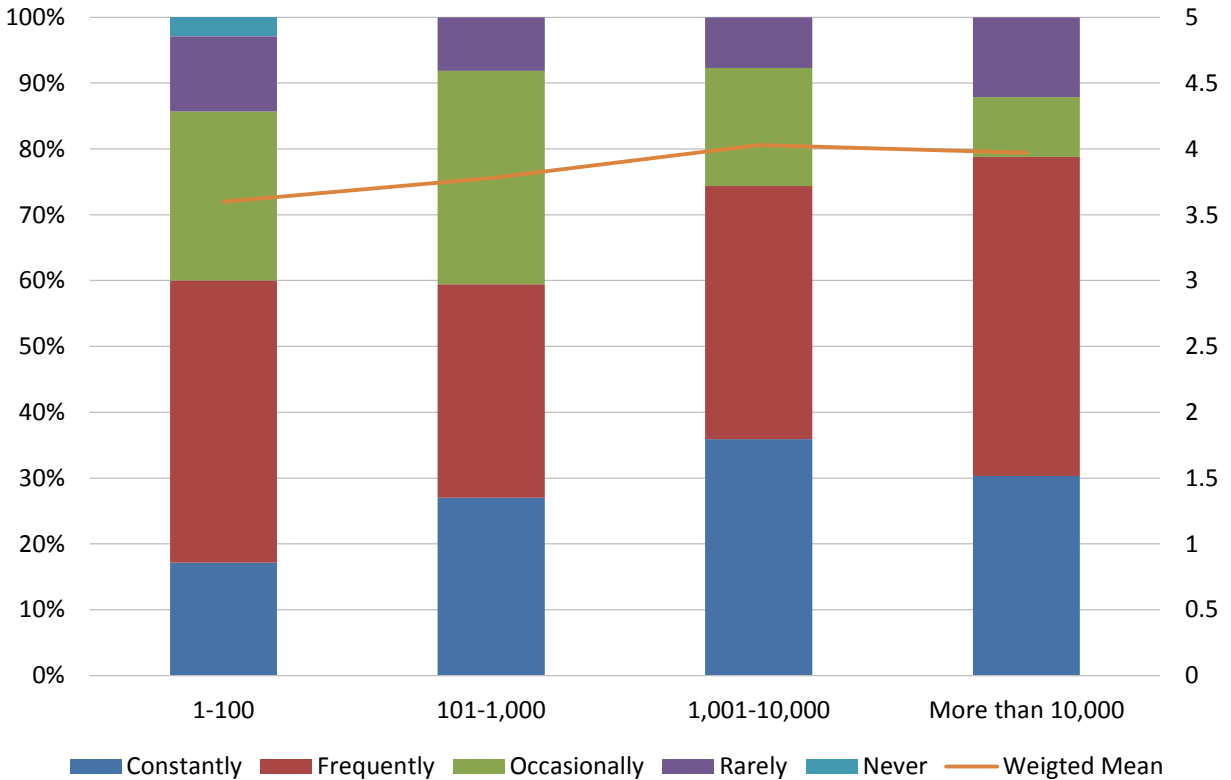


Figure 31 – Frequency of data engineering by organization size

The data in fig. 32 suggests that the more frequently data engineering is employed as a part of data analysis, the more successful an organization is likely to be with BI. While such success isn't guaranteed, organizations that use data engineering more frequently can better meet BI users' demand for more timely and relevant data.

Frequency of Data Engineering by Success with BI

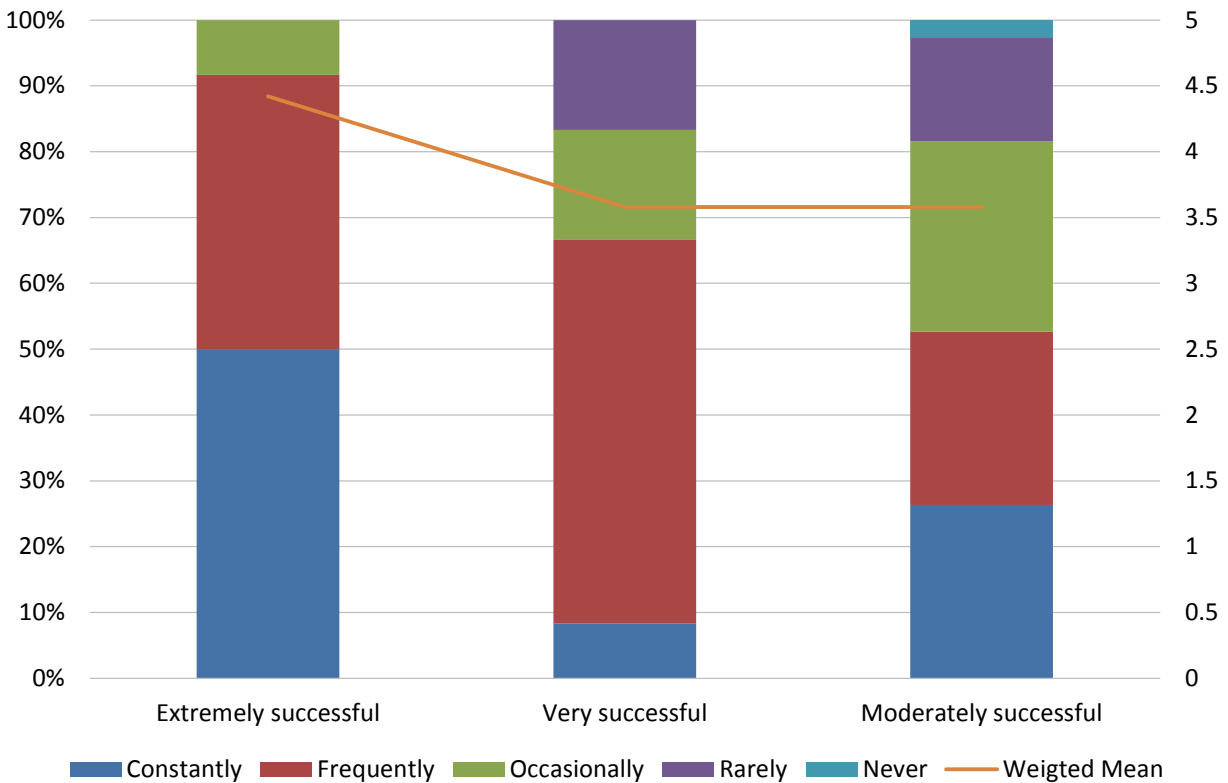


Figure 32 – Frequency of data engineering by success with BI

Data Engineering Usability

Simple to build data workflows, scheduler for coordinating data workflows, and mask or redact sensitive data are the top three usability features our survey respondents considered critical, very important, or important (fig. 33). The features most rated as not important are no code data transformations and machine learning and recommendations based on usage data gathered across users, groups, or organizations. We also note the low importance of graphical drag-and-drop capabilities.

Data Engineering Usability Features

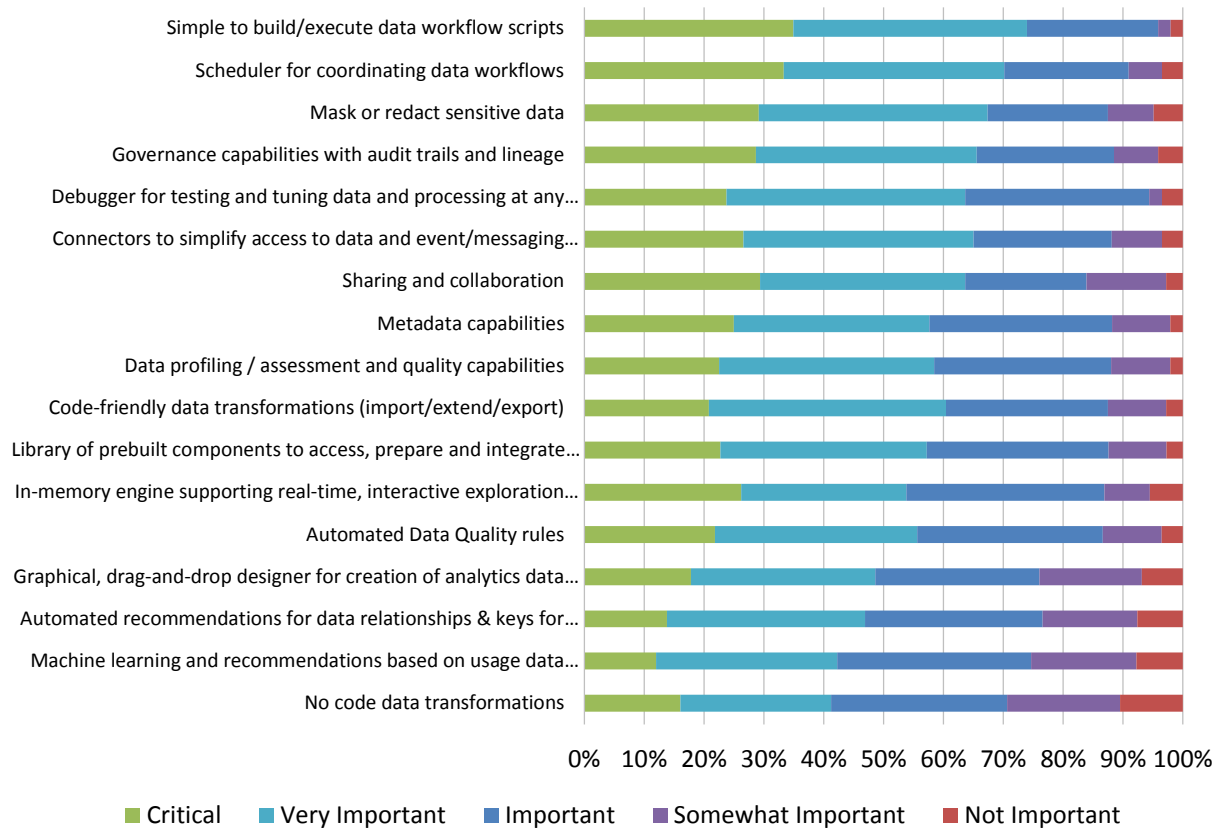


Figure 33– Data engineering usability features

Several trends over 2022-2025 are noticeable in fig. 34. Simple to build/execute data workflow scripts consistently ranks as the most important feature, increasing its score steadily over the years. Features related to automation, such as automated recommendations for data relationships and automated data quality rules, also are becoming increasingly important. Data governance and security features like mask or redact sensitive data and governance capabilities maintain high scores and show growth. Conversely, some features like graphical, drag-and-drop designer have declined slightly in perceived importance. Overall, there is a trend toward greater emphasis on automation, governance, and ease of use in data-engineering tools.

Data Engineering Usability Features 2022-2025

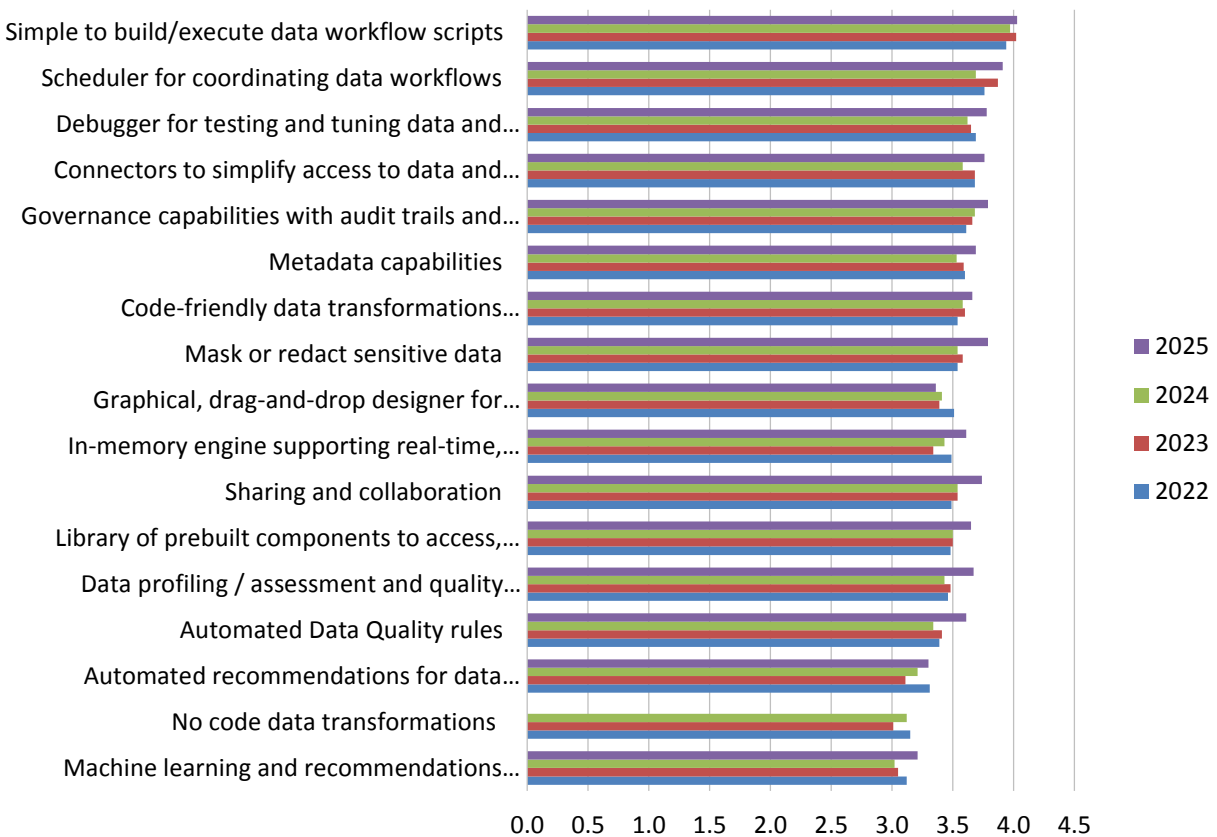


Figure 34 - Data engineering usability features, 2022-2025

All geographies give the highest rating to simple to build/execute data workflow scripts, with average importance scores of more than 4.2 (out of 5; fig. 35). Asia Pacific respondents rate it as the most important. All geographies rated no code data transformation and machine learning and recommendations as the lowest in importance; however, machine learning and recommendations has increased year over year.

Data Engineering Usability Features by Geography

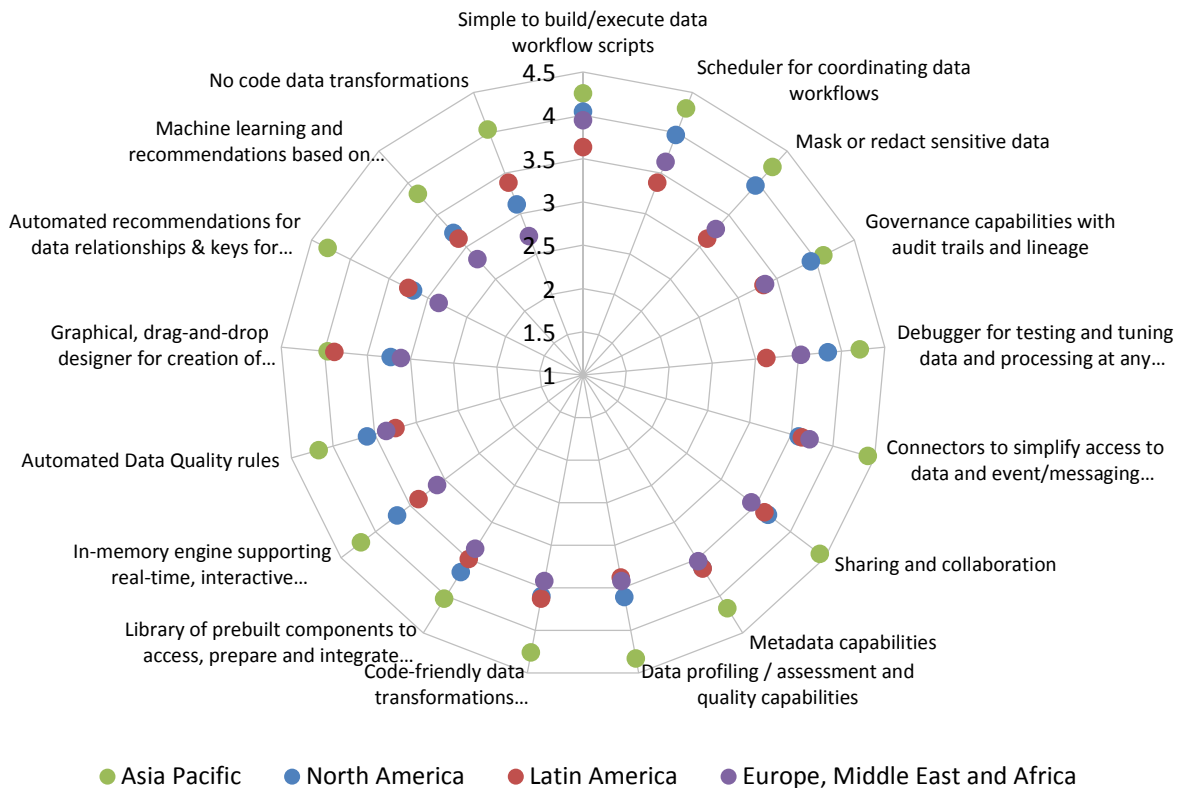


Figure 35 – Data Engineering Usability Features by Geography

Across of our respondents' industries, simple to build/execute data workflow scripts is assigned the highest importance among data engineering usability features (fig. 36). Also all industries agree that the two lowest-importance usability features are those related to machine learning recommendations and no-code data transformations.

Data Engineering Usability Features by Industry

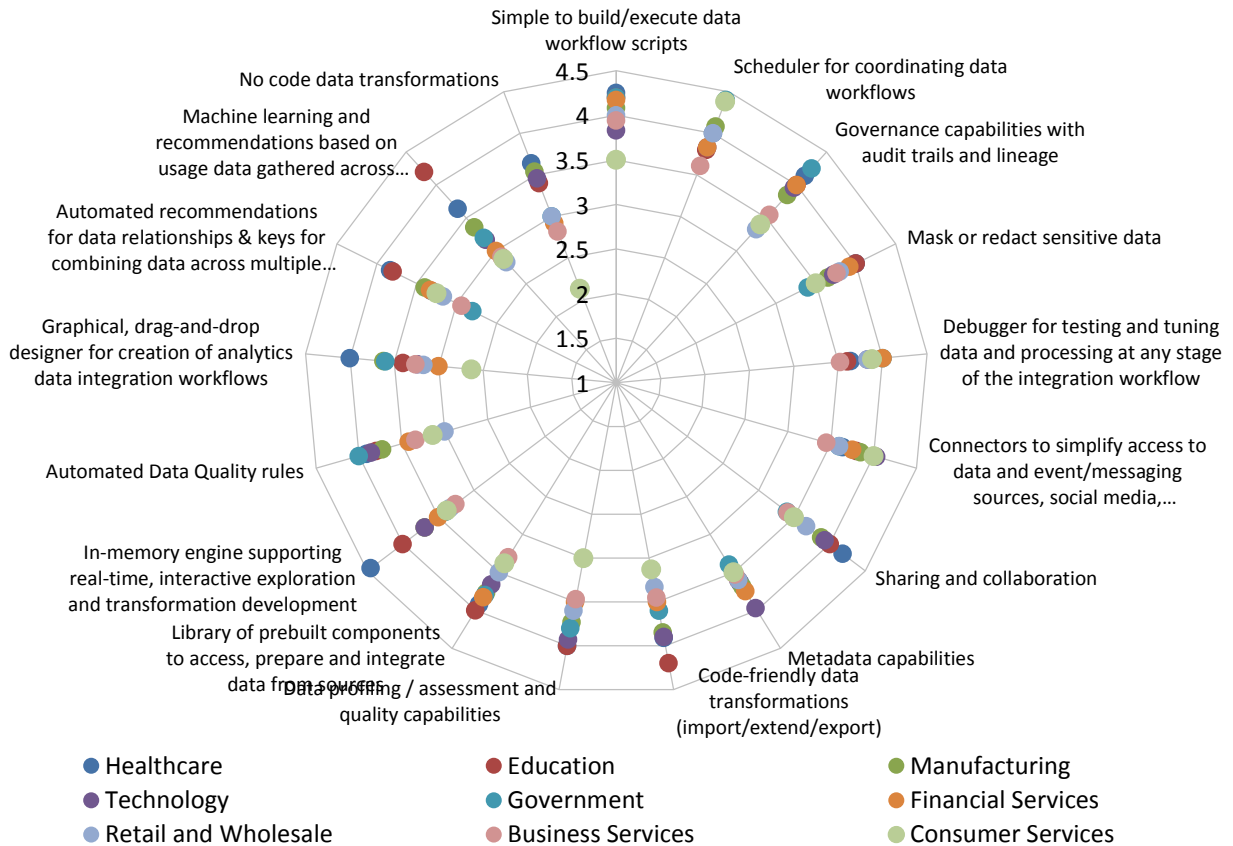


Figure 36 – Data engineering usability features by industry

Similarly, while respondents in each function rate simple to build/execute data workflow scripts as having the highest importance for usability features (fig. 37), the lowest importance across all functions is no-code data transformations.

Data Engineering Usability Features by Function

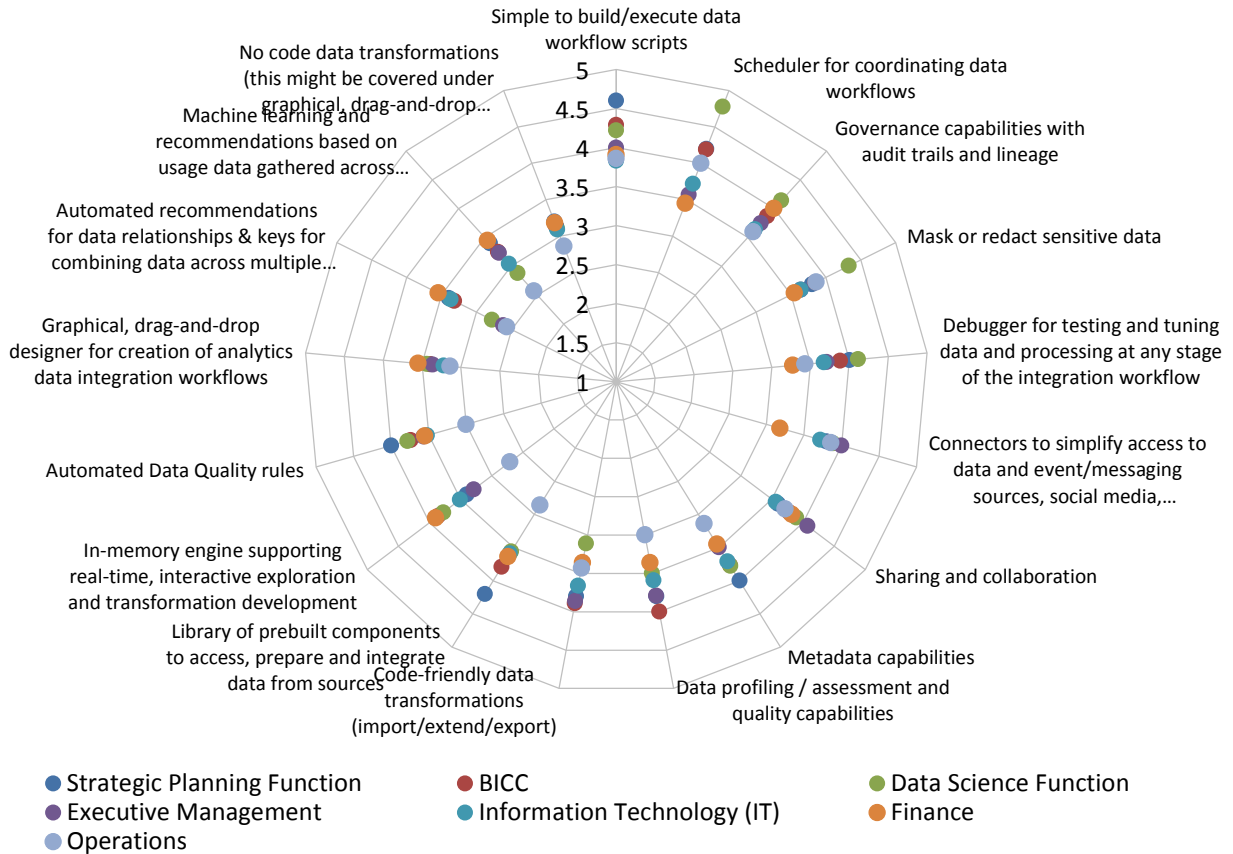


Figure 37 – Data engineering usability features by function

Our respondents' scores show small variations in the importance assigned to usability features across organization sizes. Organizations of all sizes rate simple to build/execute data workflow scripts as their top two priorities for usability features (fig. 38). Very large organizations' (more than 10,000 employees) ratings for data profiling/assessment and quality capabilities are tied as their second-highest-rated usability features.

Data Engineering Usability Features by Organization Size

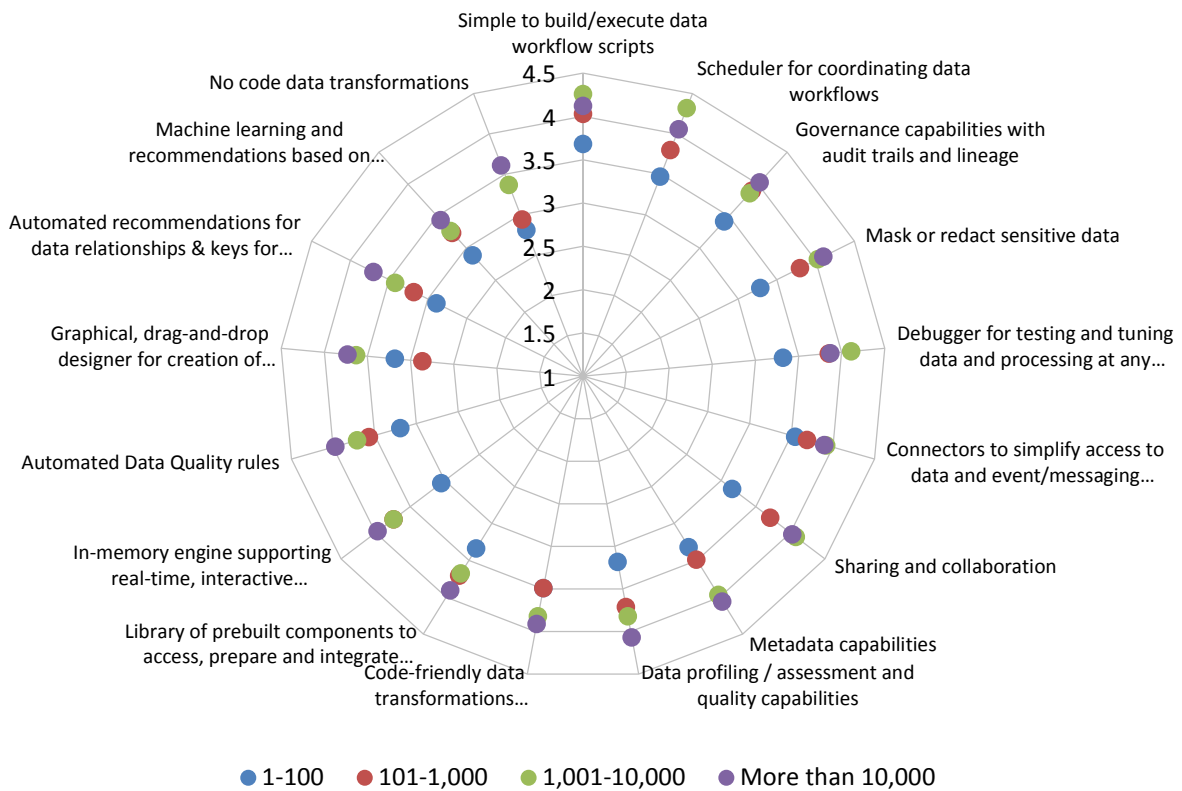


Figure 38 – Data engineering usability features by organization size

Frequency of Enrichment with Third-Party Data

We asked our survey respondents about the use and frequency of data enrichment with third-party data as part of data-engineering workflows (fig. 39). Twenty-three percent of responses say they rarely use third-party data for enrichment of data-engineering processes. Twenty-six percent occasionally enrich their data-engineering workflows with third-party data. Eighteen percent constantly enrich with third-party data and 28 percent frequently enrich with third-party data. Five percent of organizations never use third-party data for enrichment.

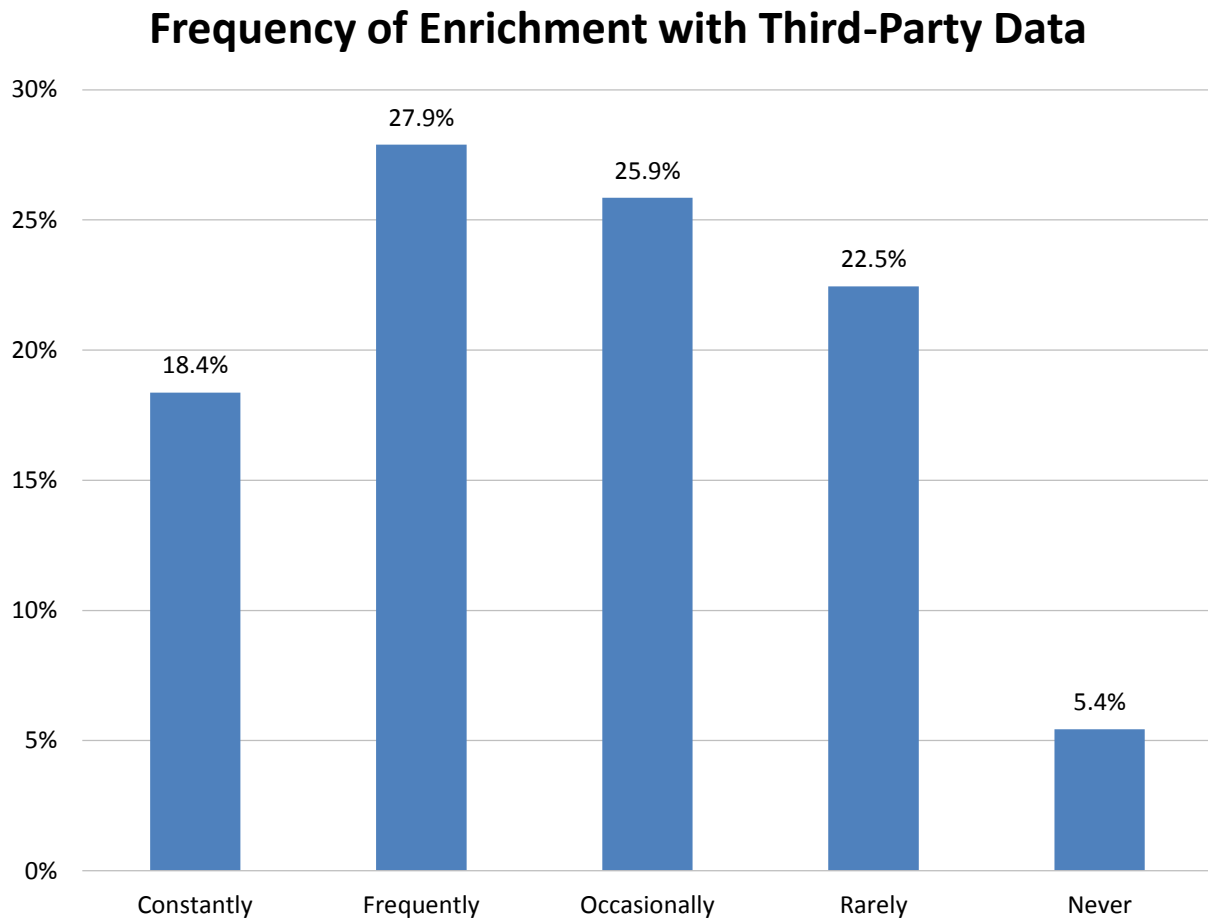


Figure 39– Frequency of enrichment with third-party data

The past seven years have seen a trend toward increasing enrichment of data engineering pipelines with third-party data. In 2025, an increasing percentage of respondents indicate they constantly and frequently use third-party data enrichment (fig. 40). As organizations use more of this data, their governance programs need to expand to address related legal and regulatory compliance and data quality issues and new risks.

Frequency of Enrichment with Third-Party Data 2015-2025

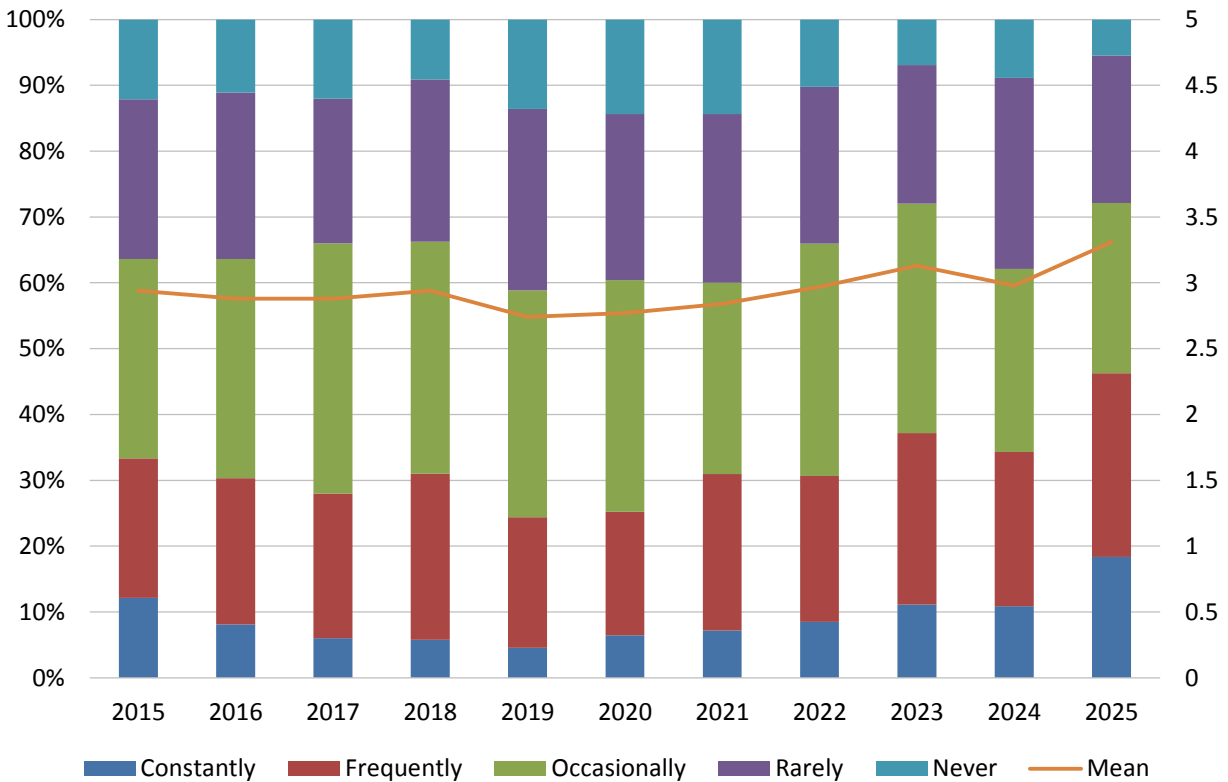


Figure 40 - Frequency of enrichment with third-party data 2021-2025

Asia Pacific respondents enrich their data sets with third-party data more often than respondents in other geographies (fig. 41). EMEA respondents are the least likely to enrich their data-engineering workflows with third-party data, with 41 percent of the respondents indicating that they rarely enrich their data sets in this way. This may be a consequence of the greater focus on privacy within the EMEA market.

Frequency of Enrichment with Third-Party Data by Geography

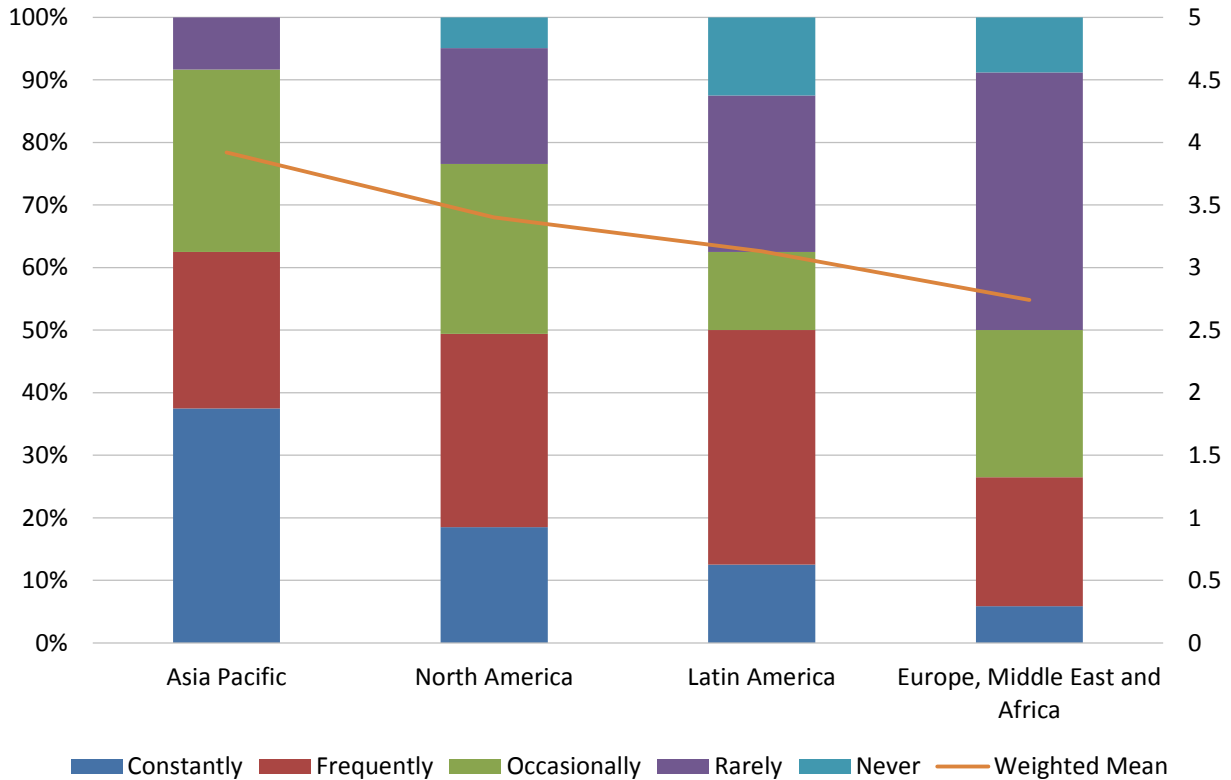


Figure 41- Frequency of enrichment with third-party data by geography

Our survey’s healthcare, technology, and financial services respondents use third party data to enrich their data-engineering workflows the most when compared with other industries (fig. 42). Education segment respondents are the most likely to never use third-party data enrichment.

Frequency of Enrichment with Third-Party Data by Industry

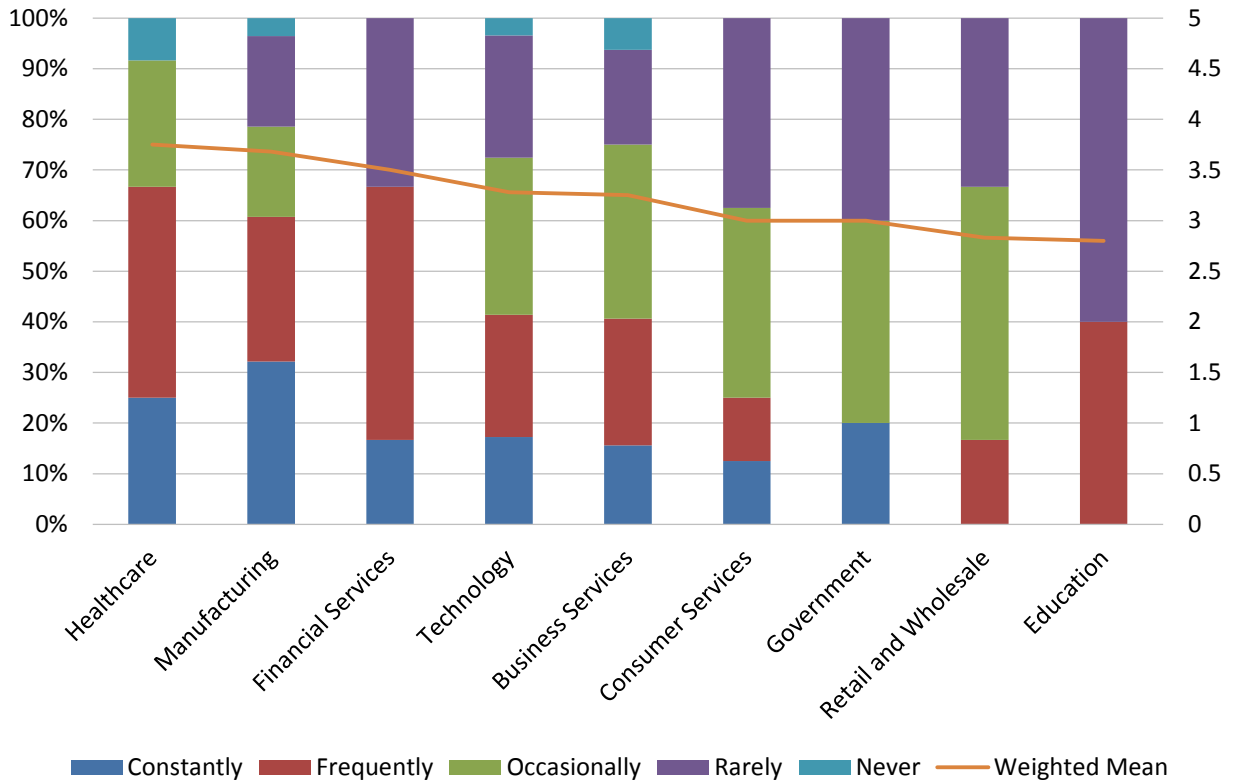


Figure 42- Frequency of enrichment with third-party data by industry

Reviewing our survey data by function, BICC functions had the highest use of enrichment of their data sets with third-party data via their data-engineering workflows (fig. 43). Operations functions, which mostly rely on internal operational system data sets, showed the lowest frequency and placed the lowest importance on data enrichment from third-party sources.

Frequency of Enrichment with Third-Party Data by Function

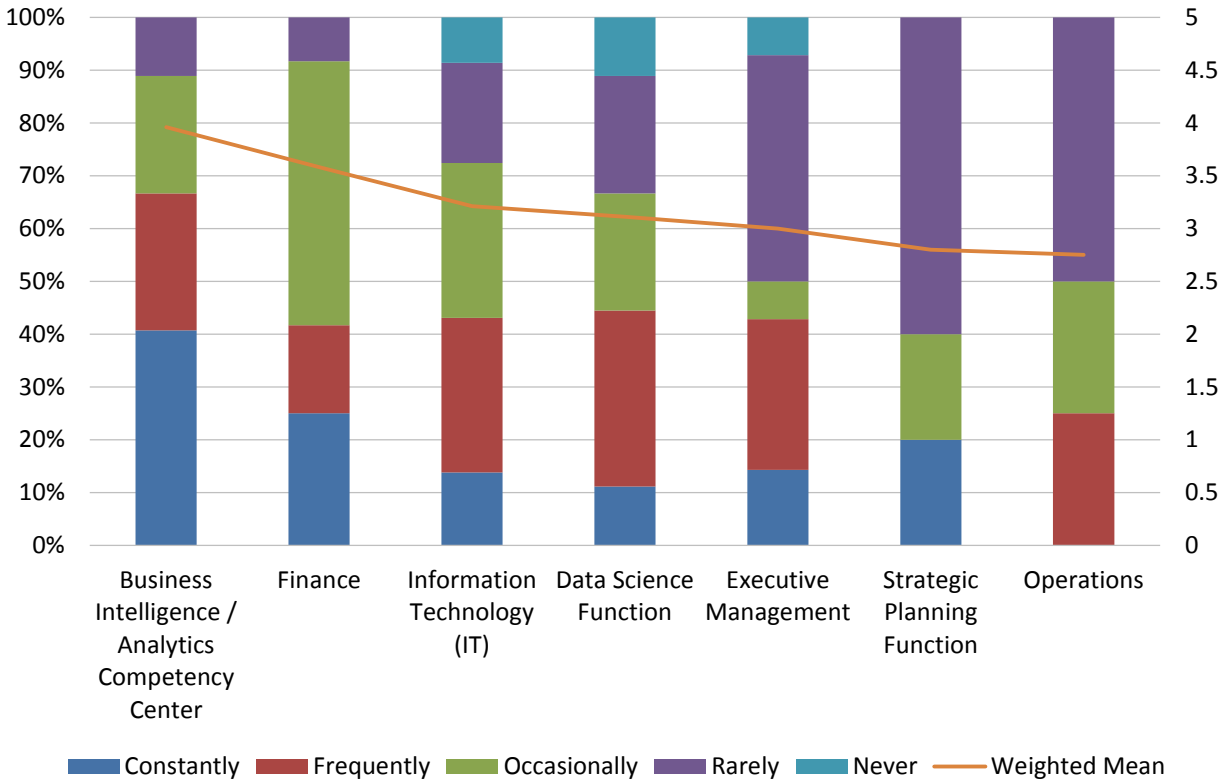


Figure 43 - Frequency of enrichment with third-party data by function

Enrichment of data engineering pipeline workflows with third-party data varies by organization size (fig. 44). The smallest organizations (1-100 employees) indicated the lowest frequency of use of third-party data and the largest organizations (more than 10,000 employees) indicate higher frequency of enrichment of their data sets with third-party data.

Frequency of Enrichment with Third-Party Data by Organization Size

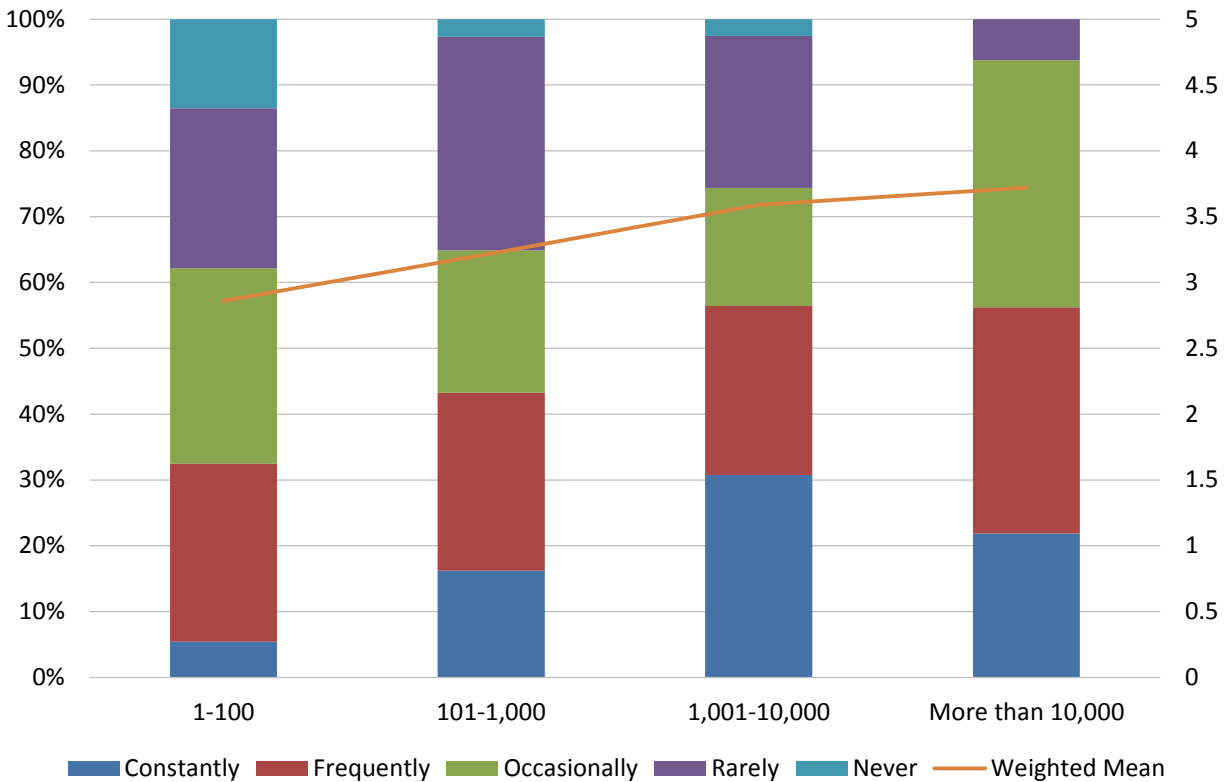


Figure 44 - Frequency of enrichment with third-party data by organization size

There appears to be a reasonable correlation between BI success and the use of third-party data set enrichment with internal data sets (fig. 45). Organizations with successful BI initiatives report far higher use of third-party data than those with less successful BI initiatives.

Frequency of Enrichment with Third-Party Data by Success with BI

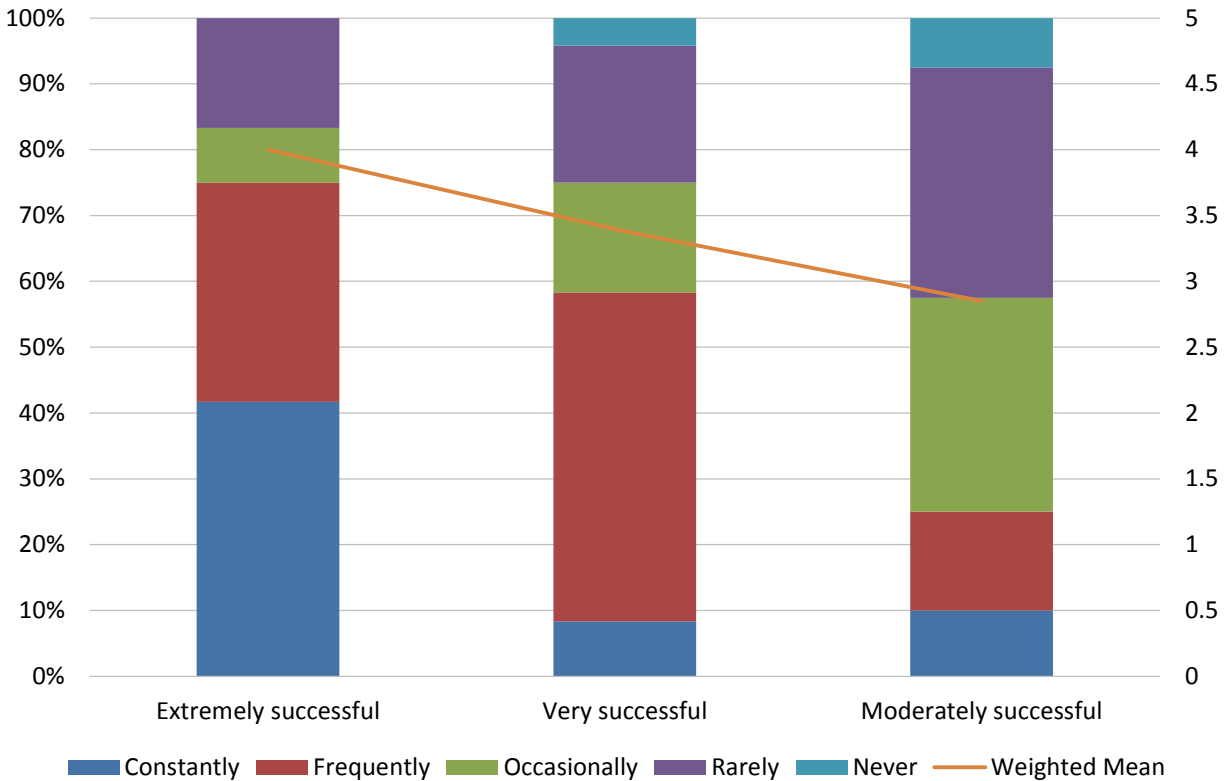


Figure 45 - Frequency of enrichment with third-party data by success with BI

Data Sources and Targets for Data Engineering

Overall, respondents indicate that relational databases, file systems, and applications are the top data sources and targets for data-engineering workflows (fig. 46). Analytical databases are also a primary target of data-engineering pipelines. Of note is the high score for object stores (e.g., Amazon S3, Google Cloud Storage, Microsoft Azure blob storage), which are an important target for data-engineering workflows but are also frequently used as a source. This is due to their role as an offline storage and staging repository for those workflows.

Data Engineering Sources and Targets

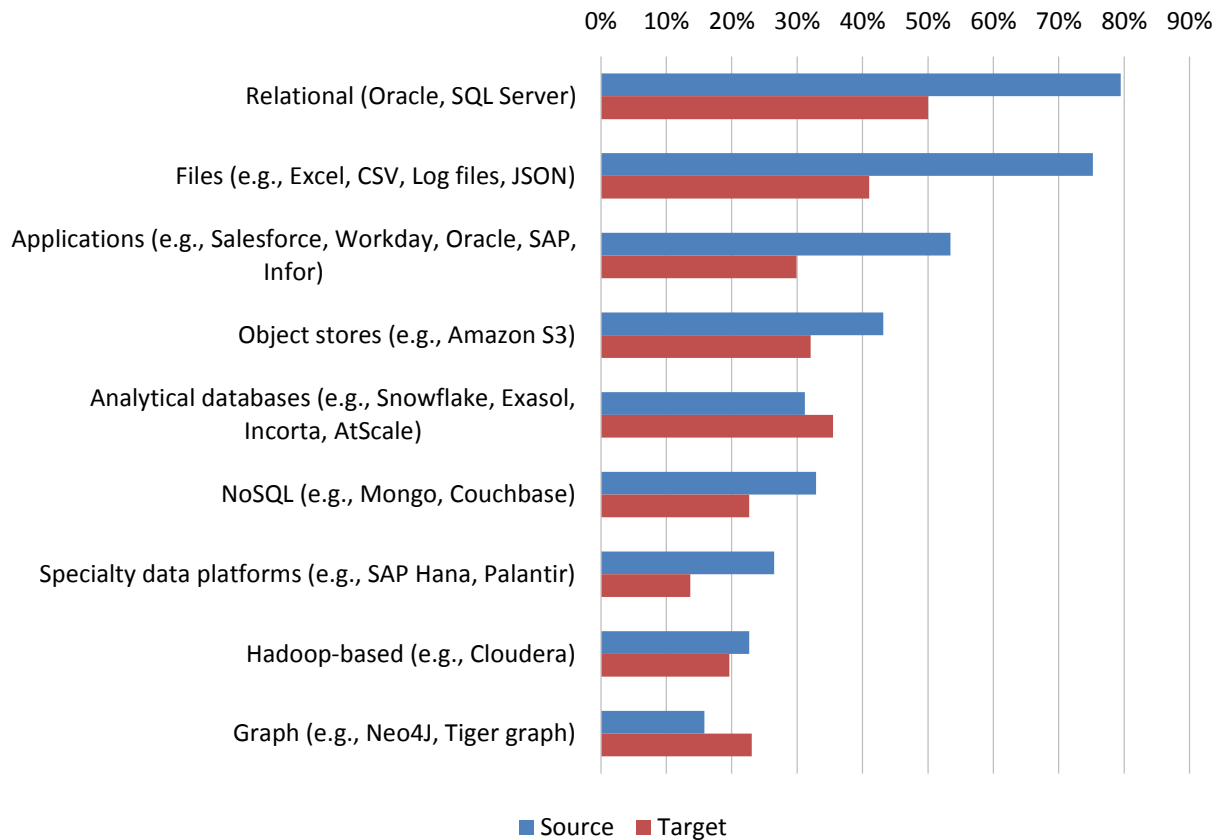


Figure 46 – Data engineering sources and targets

The use of file storage (e.g., Excel, CSV, Log files, JSON) and object storage as sources for data engineering workflows is increasing year over year (fig. 47). In fact, this year usage of files actually eclipsed the use of relational sources in data engineering workflows! There is also a trend toward greater use of NoSQL databases as sources in data-engineering workflows.

Data Engineering Sources 2022-2025

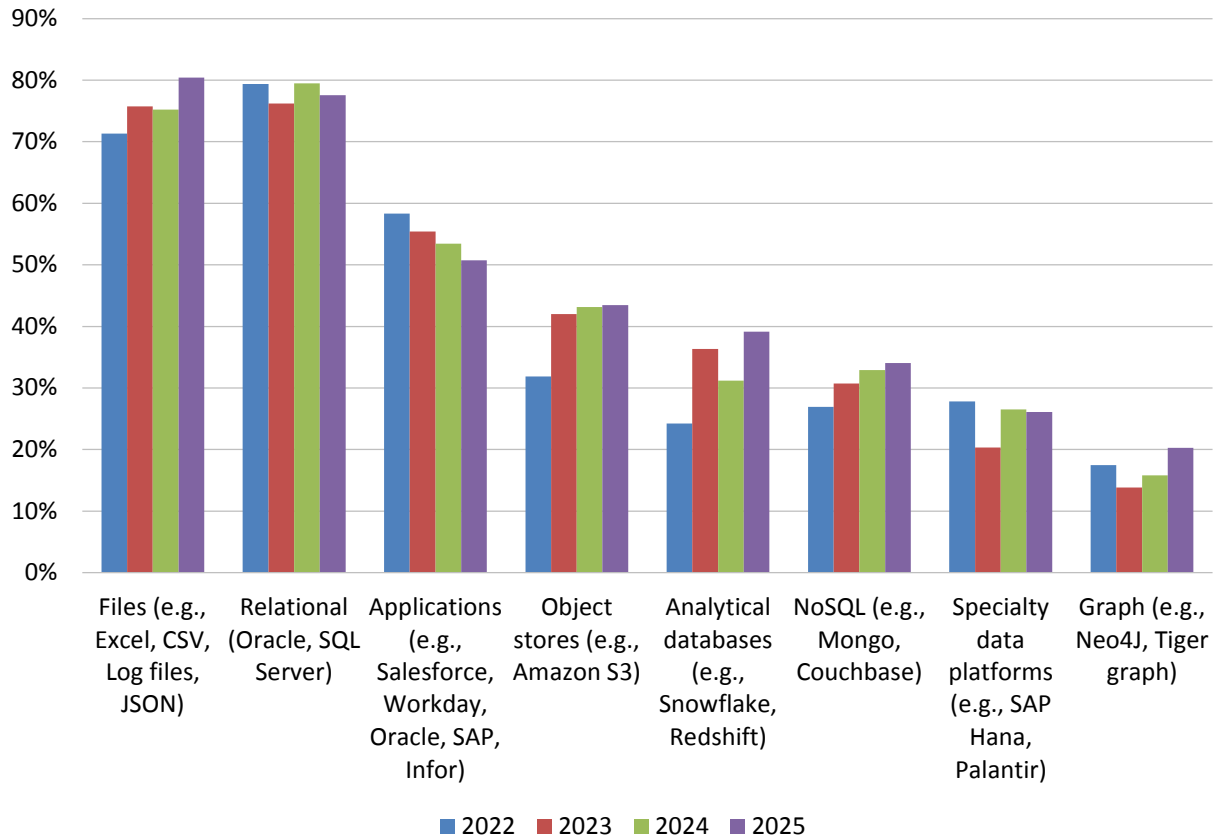


Figure 47 - Data Engineering Sources 2022-2025

With one exception, the targets for data-engineering workflows remain the same year over year (fig. 48). Whereas last year, analytical databases showed the largest year-over-year increase as targets, this year they show the largest decline as targets. Relational databases remain consistent as a top choice among targets.

Data Engineering Targets 2022-2025

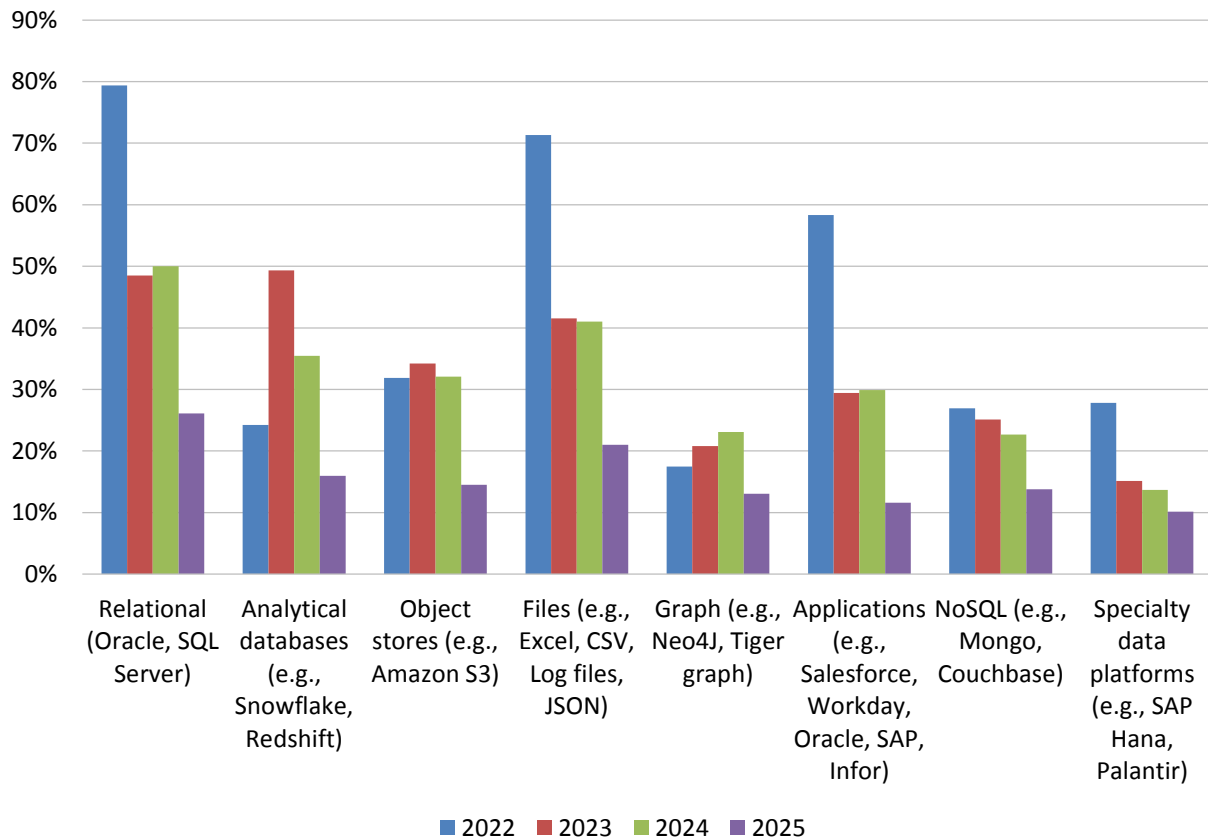


Figure 48-Data Engineering Targets 2022-2025

Effectiveness with Current Approach to Data Engineering

There appears to be room for improvement in current approaches to data engineering. Thirty-seven percent of our survey respondents rate their current approach to data engineering as highly effective (fig. 49). Forty-eight percent of our respondents rate their current approach to data engineering as somewhat effective. Fifteen percent of respondents rate their current approach as somewhat or totally ineffective. Data leaders who champion investments in best practices initiatives can help improve the effectiveness of data=engineering investments.

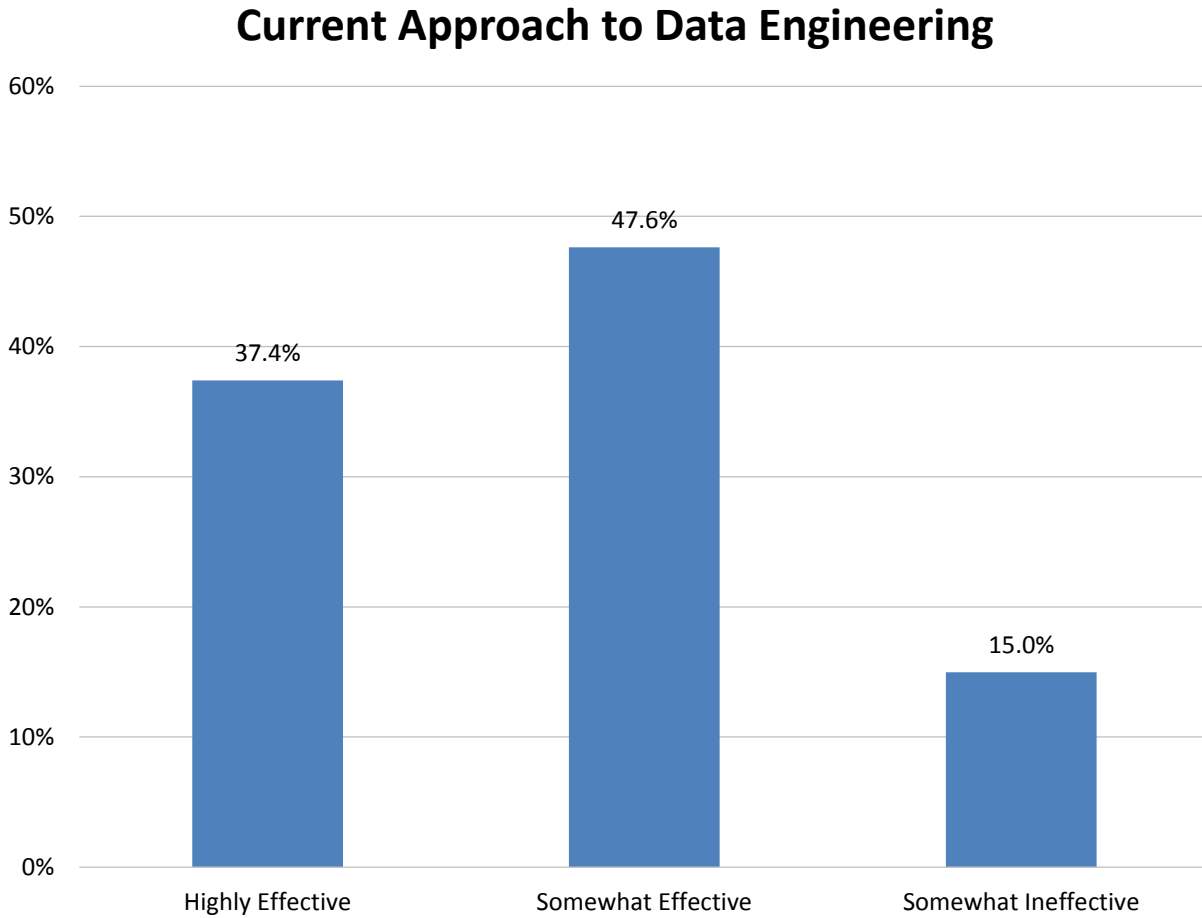


Figure 49 – Current approach to data engineering

Overall, our survey data suggests organizational approaches to data engineering are becoming more effective (fig. 50). The percentage of organizations that are highly effective is trending upward while the percentage of organizations that are somewhat ineffective or totally ineffective declines.

Current Approach to Data Engineering 2015-2025

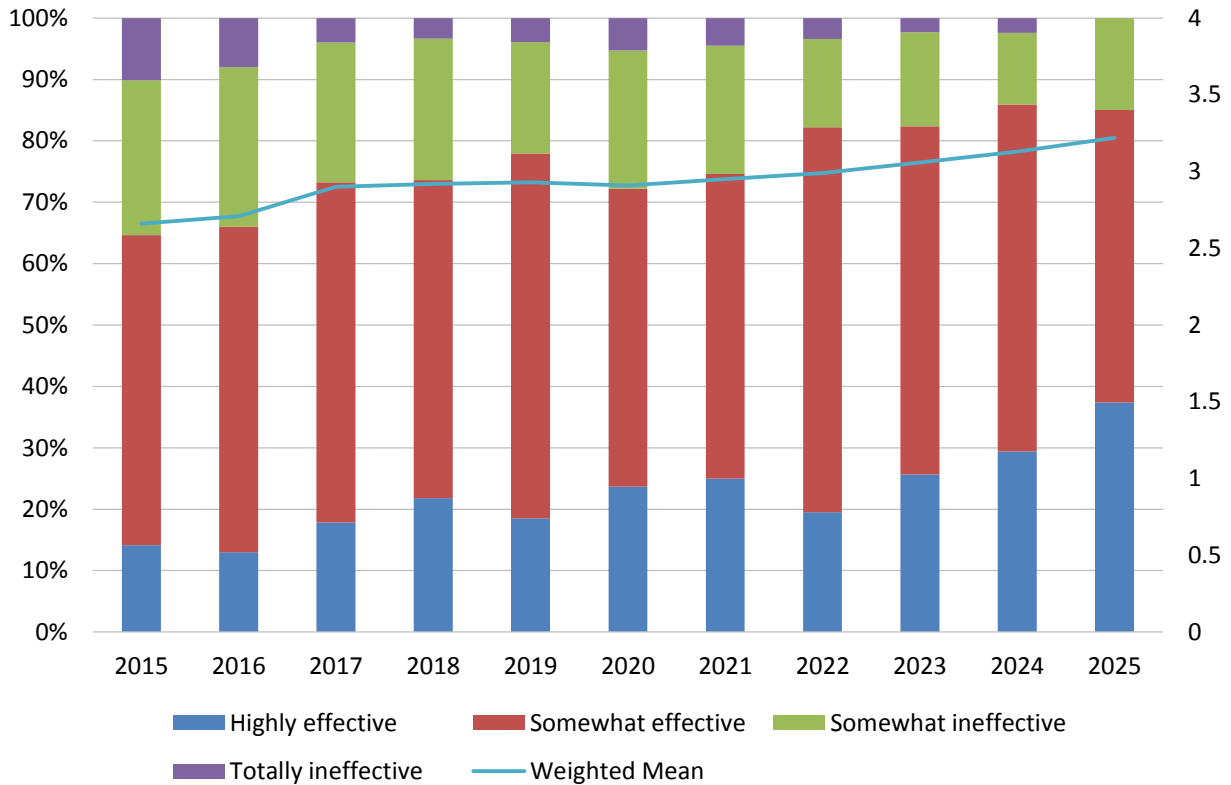


Figure 50- Current approach to data engineering 2015-2025

Reviewing the data by region, Asia Pacific and North American respondents are the most satisfied with their current approach to data engineering. EMEA and Latin American respondents are less likely to rate their data-engineering approaches as effective (fig. 51).

Current Approach to Data Engineering by Geography

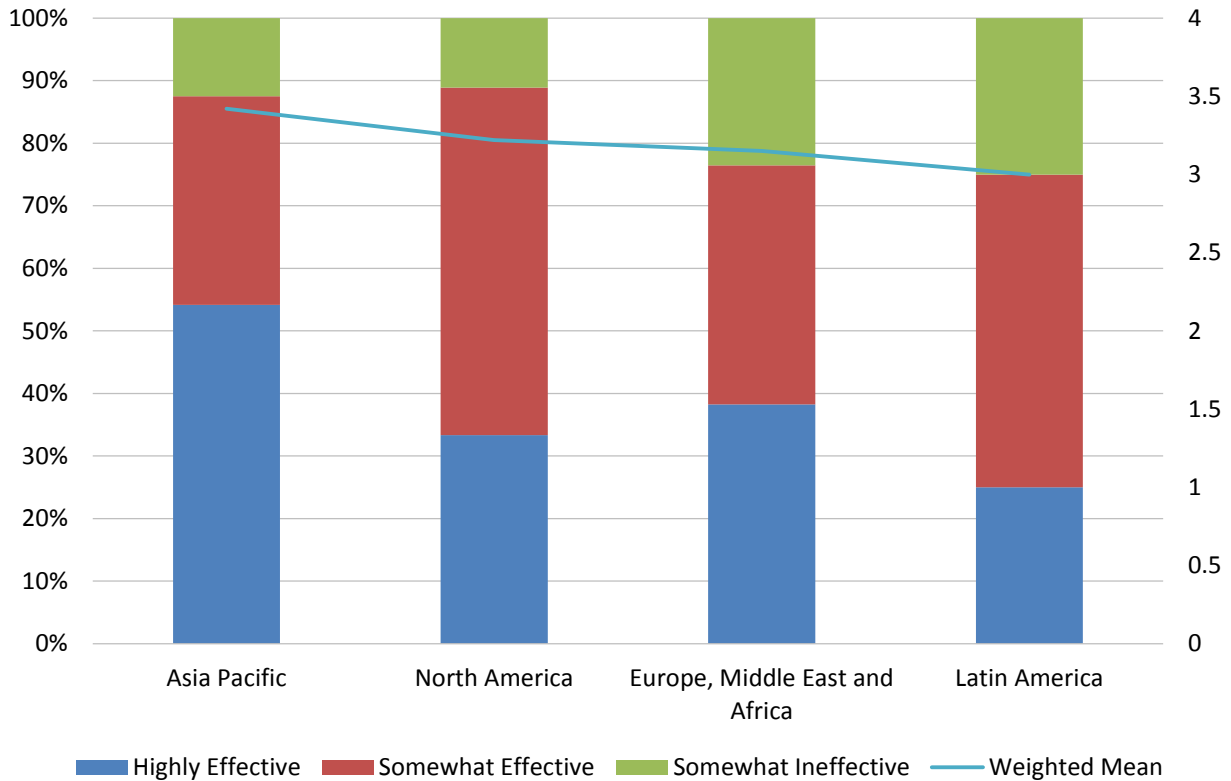


Figure 51 — Current approach to data engineering by geography

In general, the effectiveness of data engineering increases along with organizational size (fig. 52). However, as it approaches and exceeds 10,000 employees, effectiveness declines. Smaller organizations most likely have smaller and lower-skilled staff available to support data-engineering efforts. The smallest organizations probably are using scripts, spreadsheets, and self-service data prep tools for data engineering. As organizations attain the very largest size, their effectiveness declines due to the complexity and multiple locations of data sources and targets (e.g., cloud, hybrid, and on-premises deployments) and the large number and diversity of their use cases.

Current Approach to Data Engineering by Organization Size

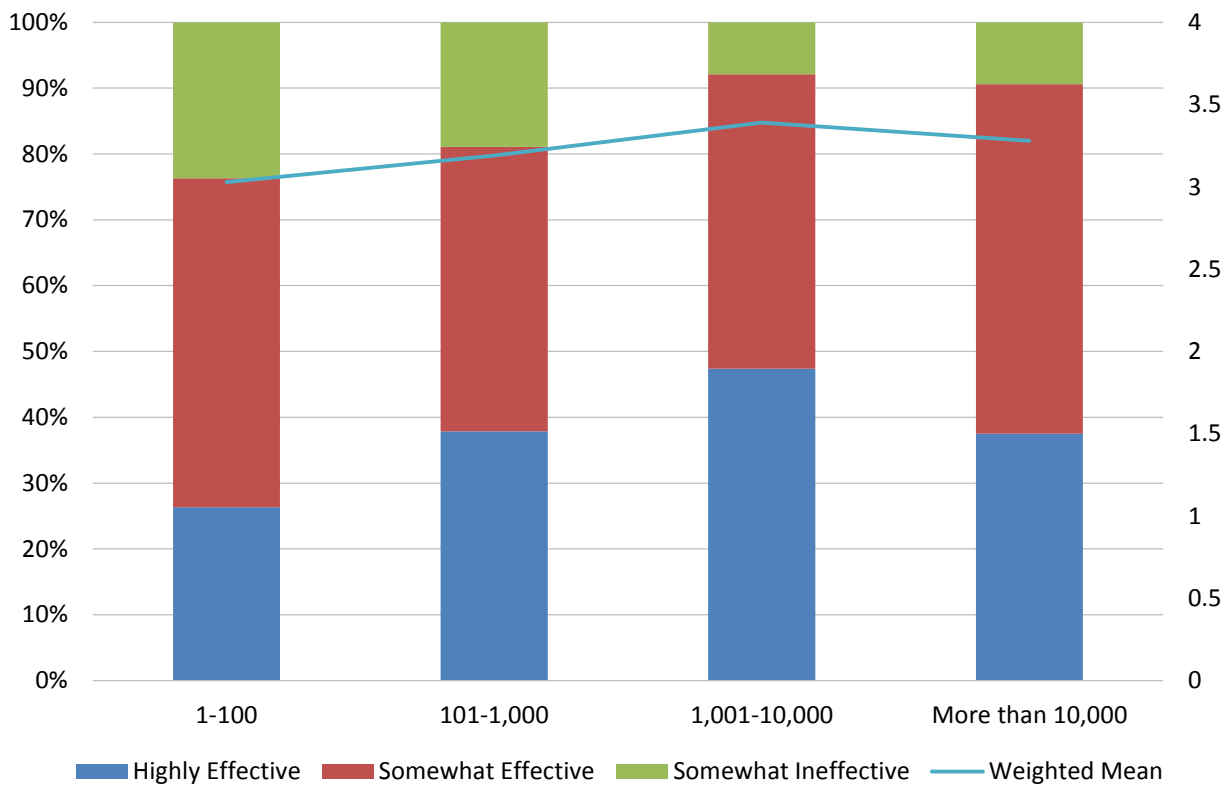


Figure 52 – Current approach to data engineering by organization size

Fig. 53 suggests that the effectiveness of data-engineering approaches can materially impact the overall success of a BI initiative. Organizations shouldn't overlook investments in data-engineering approaches and best practices and skills when evaluating how to increase the success of their BI initiatives.

Current Approach to Data Engineering by Success with BI

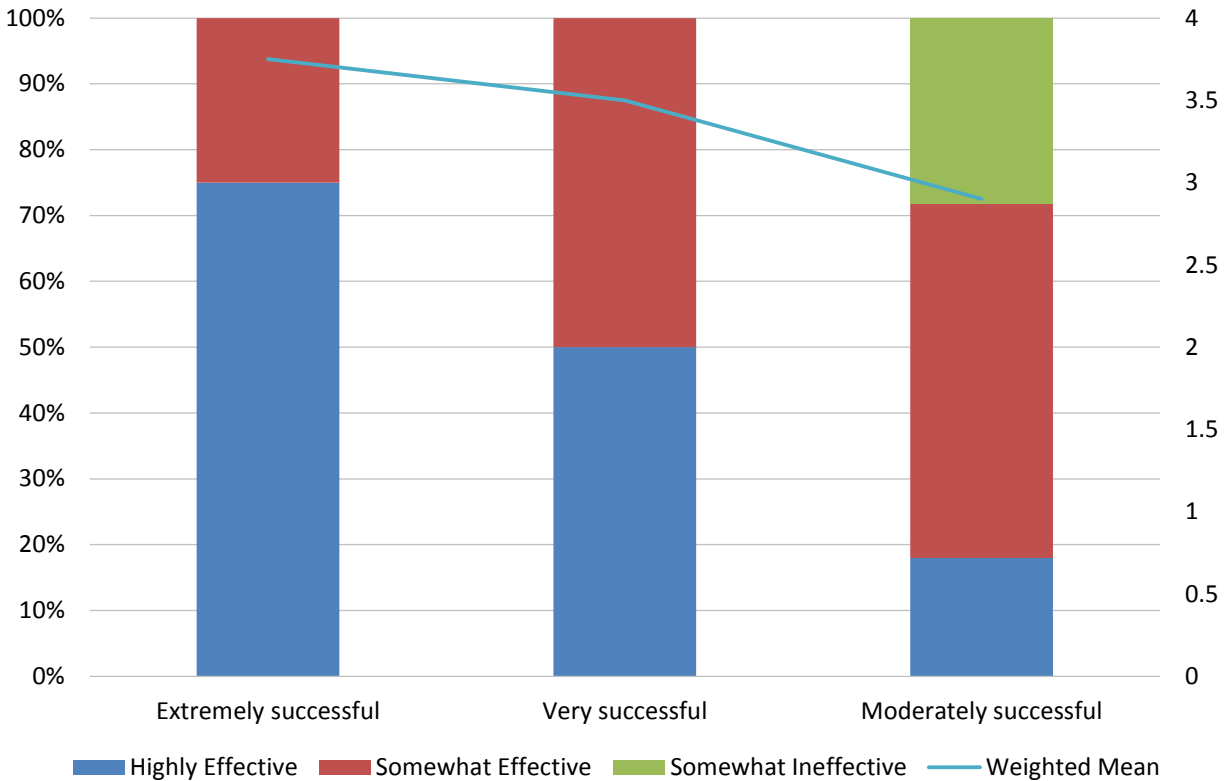


Figure 53 – Current approach to data engineering by success with BI

Deployment of Data Engineering Capabilities

Fig. 54 shows the distribution of survey responses regarding the importance of data-engineering functionality across different deployment options. The responses are categorized by perceived importance (critical, very important, important, somewhat important, not important) and the deployment options (on premises, public cloud [SaaS], hybrid [on premises and cloud], private cloud). The percentages indicate the proportion of respondents who assigned each level of importance to data-engineering functionality within each deployment option. Private cloud and public cloud are the preferred deployments for data-engineering capabilities.

Deployment of Data Engineering Capabilities

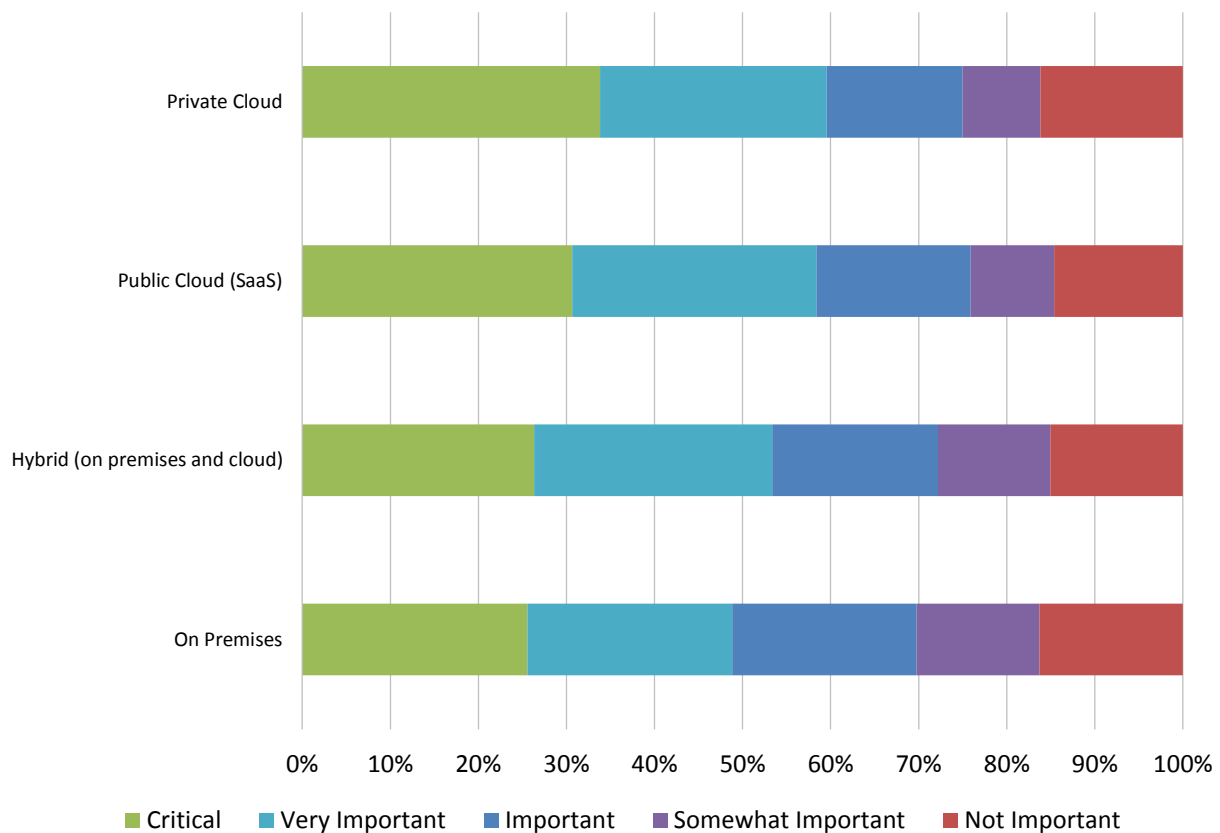


Figure 54- Deployment of data engineering capabilities

Fig. 55 represents the evolution of deployment preferences for data- engineering capabilities across different platforms from 2017-2025. The numerical values represent the perceived importance or adoption level for each deployment option during the respective year. Public cloud (SaaS) and private cloud deployment options are the top priority in 2025.

Deployment of Data Engineering Capabilities 2017-2025

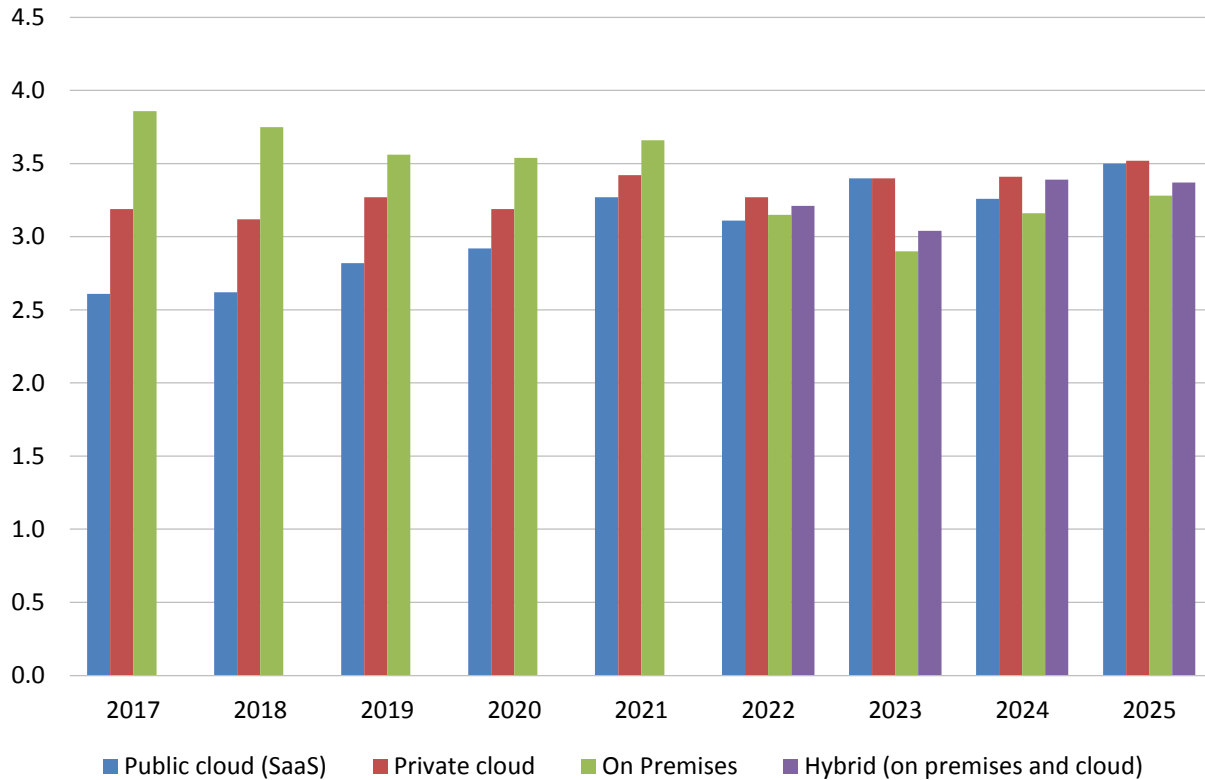


Figure 55- Deployment of data engineering capabilities 2017-2025

Fig. 56 presents numeric responses indicating the priority for the location of data-engineering functionality based on geographic regions. The top deployment priority for Asia Pacific and Latin America is public cloud (SaaS). In North America, the top deployment priority is slightly in favor of public clouds over private deployments (second in priority). On-premises deployment is the top priority for EMEA respondents.

Deployment of Data Engineering Capabilities by Geography

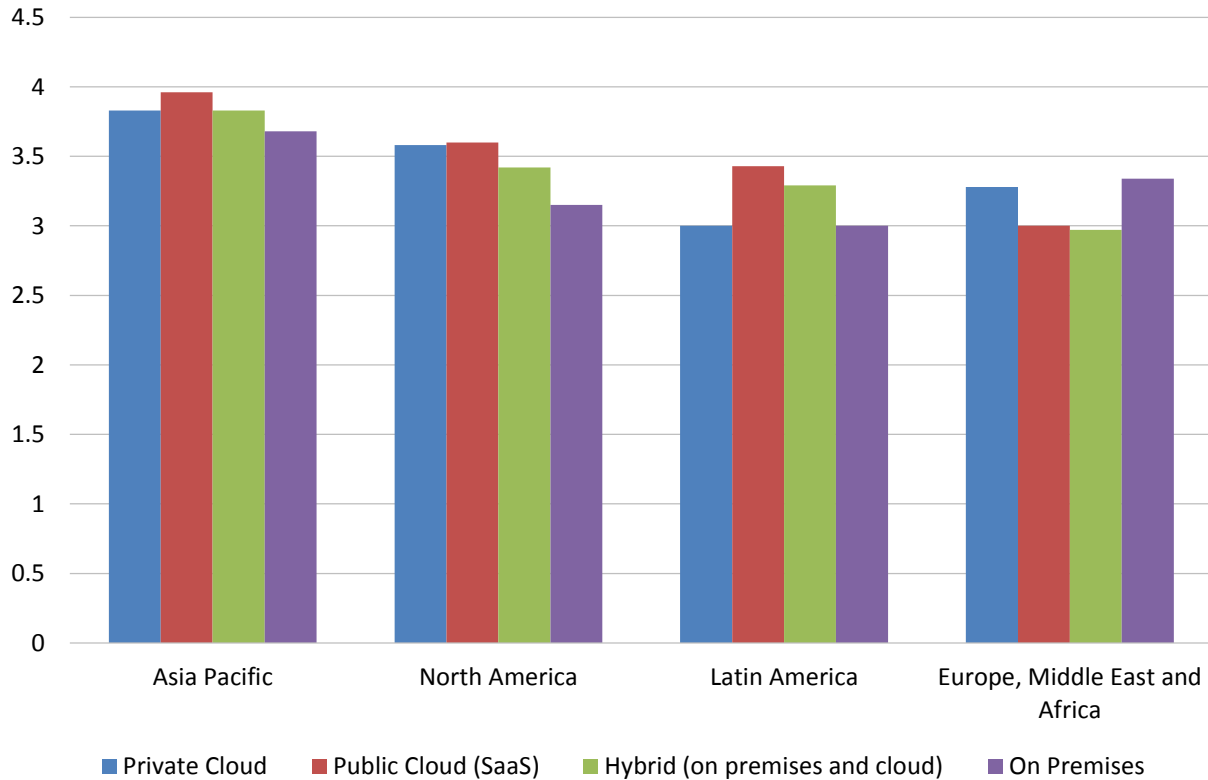


Figure 56- Deployment of data engineering capabilities by geography

Overall, industry preferences vary, with public and private cloud as prominent choices in different sectors. Fig. 57 indicates industry preferences for the location of data engineering functionality. The retail and wholesale sector favors public deployment while education prefers deployment on premises. Financial services firms prefer public cloud; consumer services prefer private cloud; consumer services prefer private cloud; consumer services prefer private cloud; consumer services prefer private cloud.

Deployment of Data Engineering Capabilities by Industry

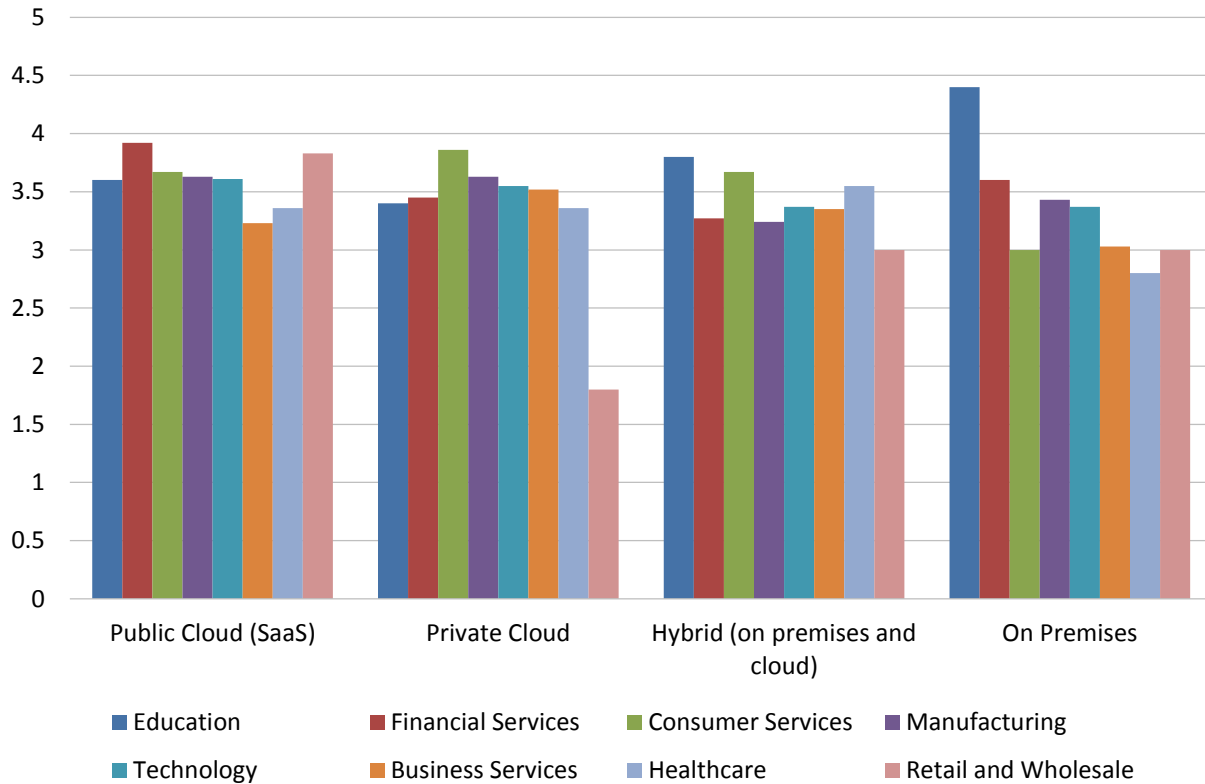


Figure 57 - Deployment of data engineering capabilities by industry

The data also suggests preferences vary according to organizational size, with the largest organizations generally favoring hybrid and private cloud deployments (fig. 58). Although the smallest of organizations place the highest priority on public cloud (SaaS), it's a popular option across all sizes of organizations.

Deployment of Data Engineering Capabilities by Organization Size

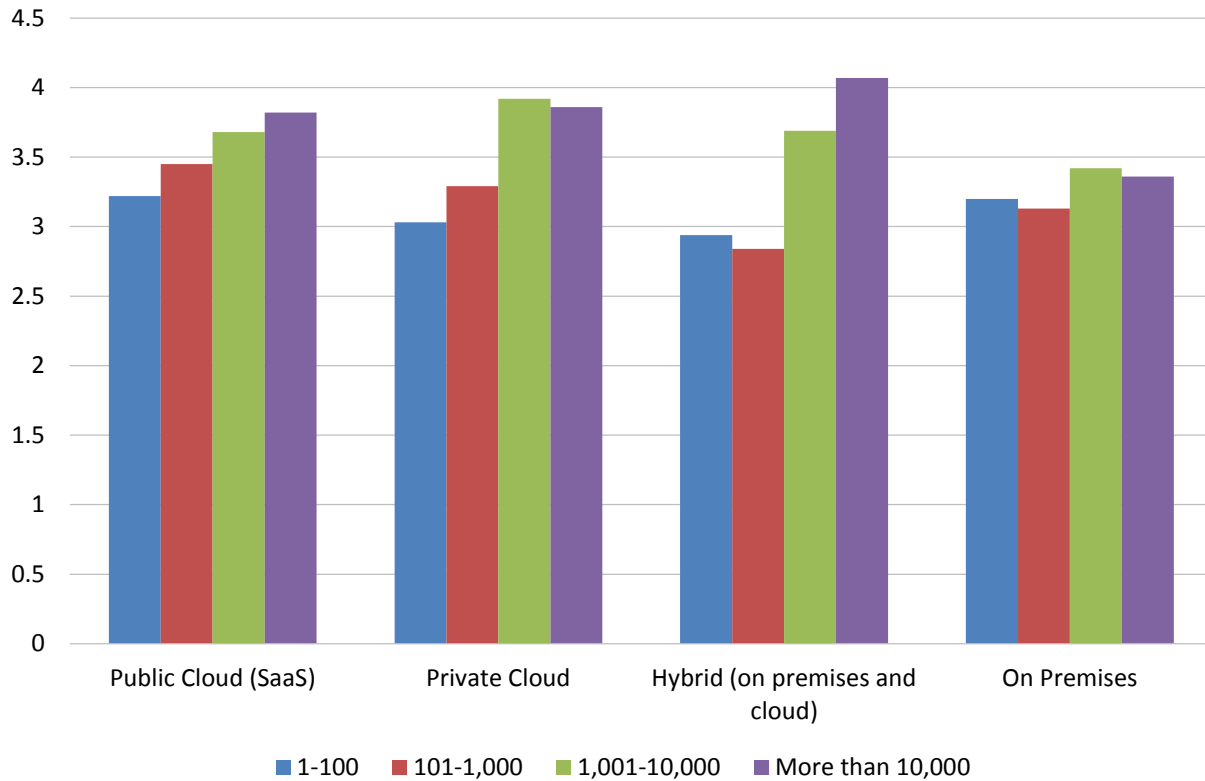


Figure 58 - Deployment of data engineering capabilities by organization size

Industry and Vendor Analysis

Industry and Vendor Analysis

Industry Support for Data Engineering Deployment Options

Alongside their users, vendors have made the shift to cloud-based deployment, with SaaS/cloud support at 92 percent (fig. 59). For cloud and database independence, 88 percent of vendors responded stating that they are independent of both. Additionally, 12 percent of the vendors said they have no plans for offering an on-premises deployment option.

Industry Support of Deployment Options

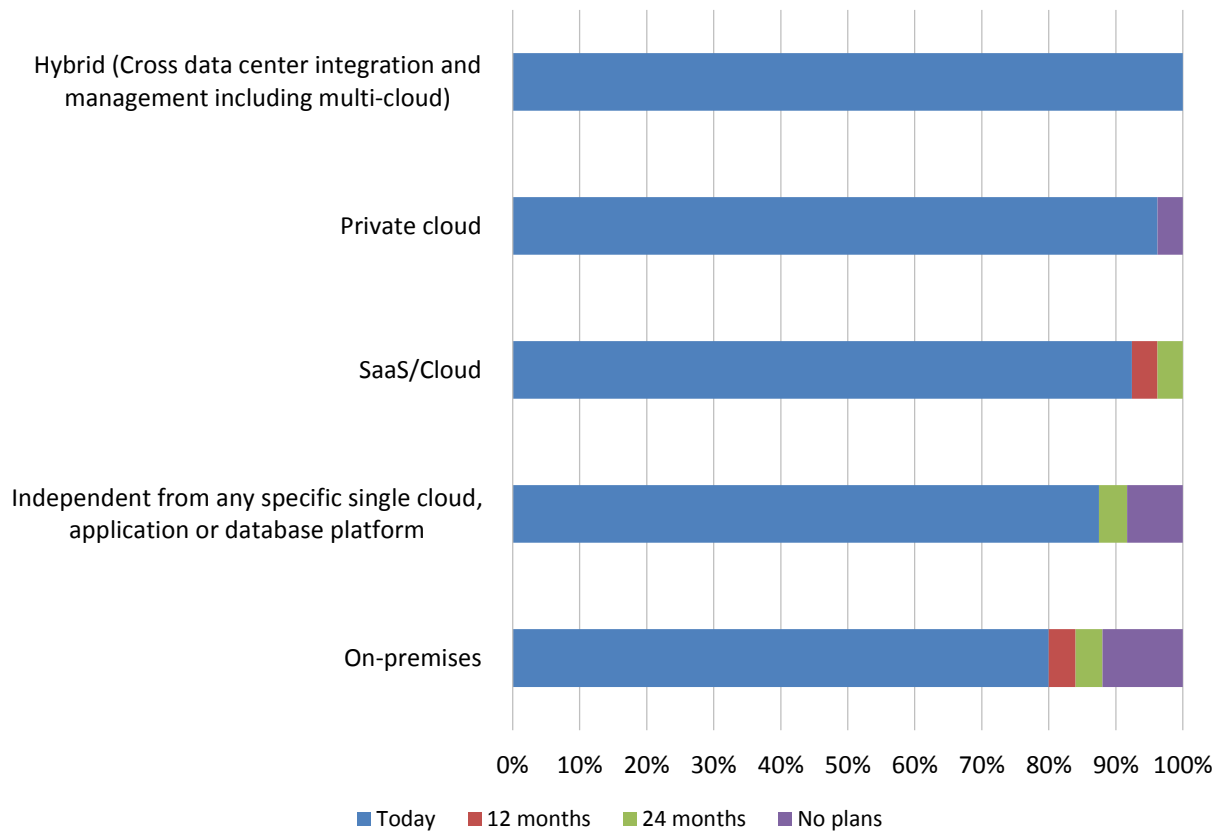


Figure 59 – Industry support for data engineering deployment options

Data Engineering Pricing Models

In 2025, our survey data suggests that vendor organizations strive for a balance across multiple licensing models (fig. 60). The top three models are enterprise licensing, data volume/capacity-based pricing, and named user pricing.

Industry Support of Data Engineering Pricing Models

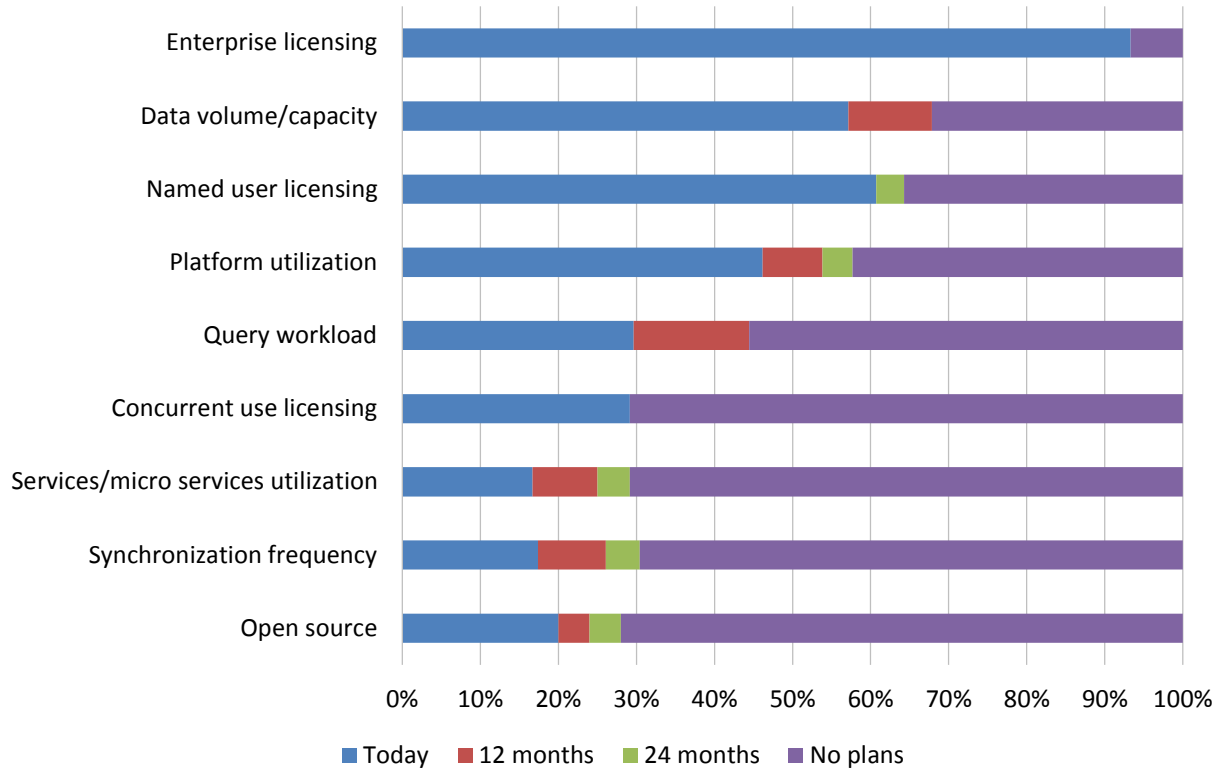


Figure 60 – Industry support of data engineering pricing models

Industry Support for Development and Usability Features

In large part, users can expect most vendors to meet their usability requirements for data-engineering solutions. That said, user and vendor priorities do not always match (fig. 62). For example, the ability to mask or redact sensitive data is the third most-important item for respondents asked the same question in the user survey, while vendors ranked this feature ranks at No. 12.

Industry Support of Data Engineering Product Development and Usability Features

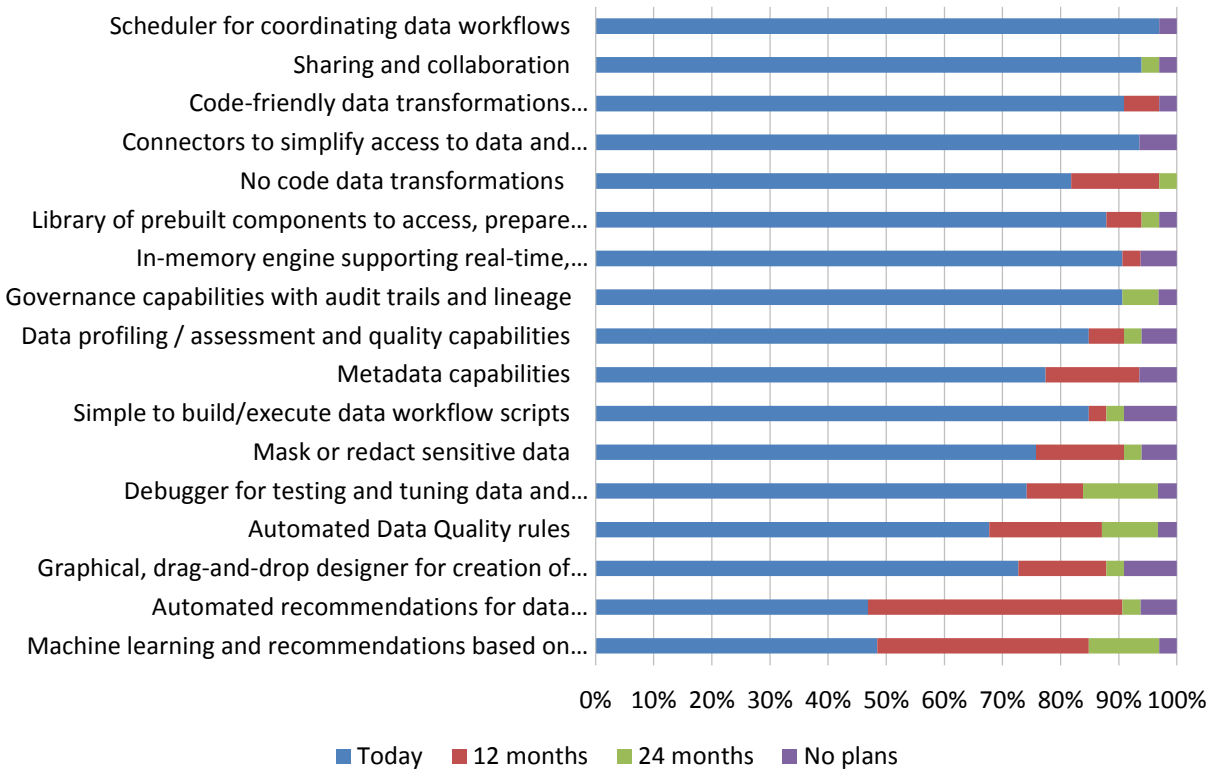


Figure 61 – Industry support of data engineering product development and usability features

Industry Support for Data Processing and Transformation Features

As with usability, users can expect a majority of vendors to meet their transformation requirements for data-engineering solutions (fig. 63). However, there are differences worth examining. For example, process/transform unstructured data is No. 19 in importance in the user survey, but 33 percent of the vendors don't offer that capability today.

Industry Support of Data Engineering Features for Processing and Transforming Data

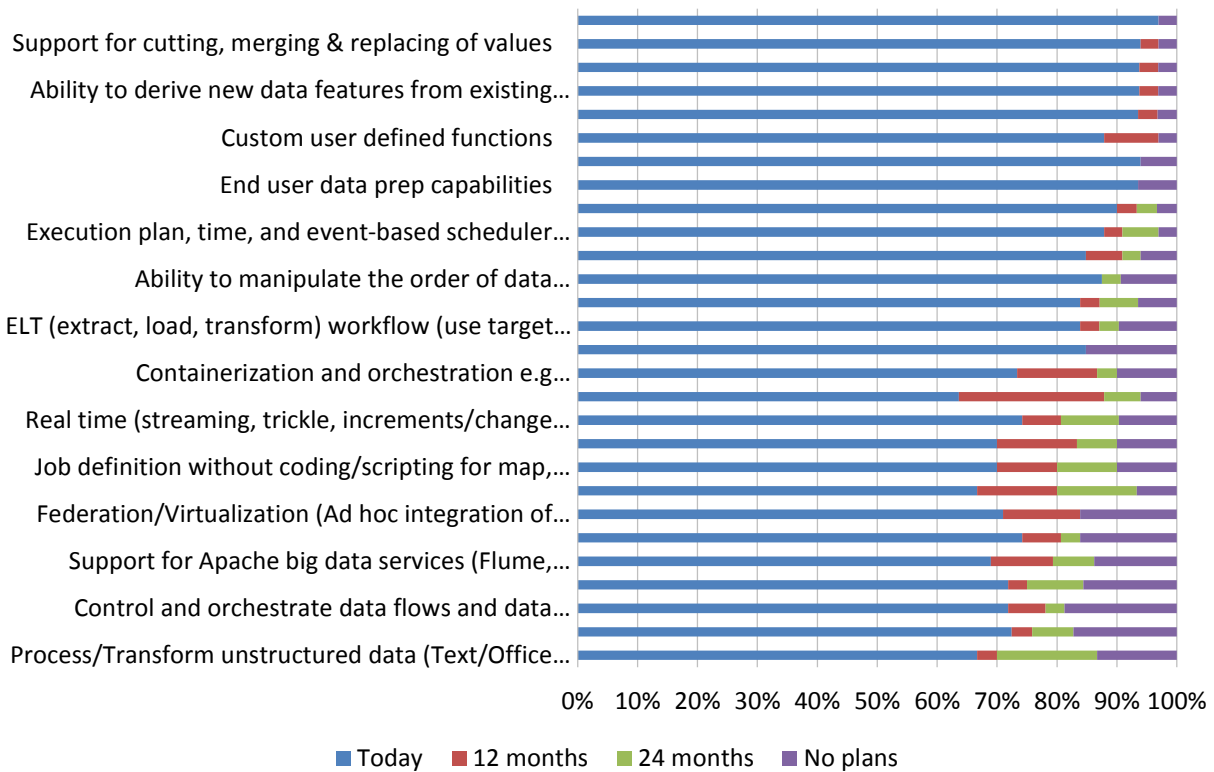


Figure 62 – Industry support of data engineering features for processing and transforming data



Vendor Ratings

Vendor Ratings

We include 31 vendors in our 2025 data-engineering ratings (fig. 63). For each vendor, we consider the following features: development and usability, process and transformation of data, and data sources and targets. In this report, we include vendors that scored 50 percent or greater.

The top ten vendors had relatively close scores, with minor differences among them. The variations within the top 20 were somewhat greater, but still relatively modest.

Data Engineering Vendor Ratings

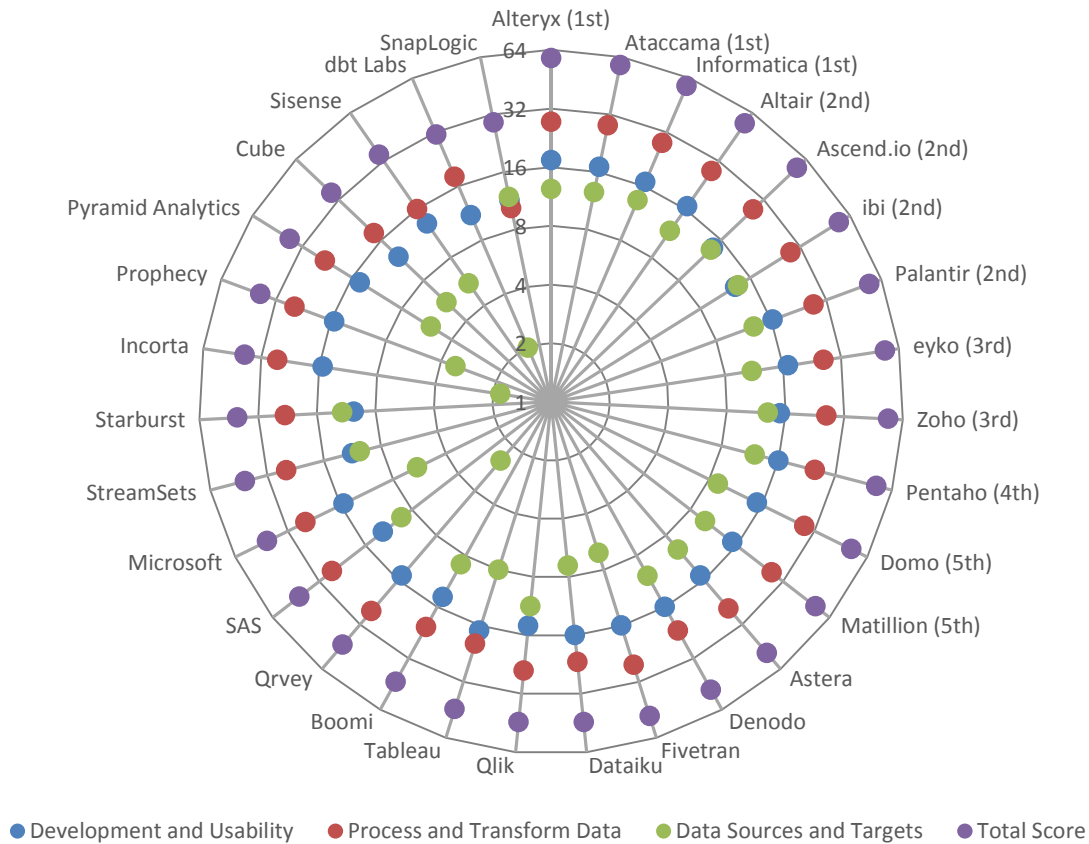


Figure 63 – Data engineering vendor ratings

Other Dresner Advisory Services Research Reports

- Wisdom of Crowds® “Flagship” Business Intelligence Market Study
- Active Data Architecture®
- AI, Data Science, and Machine Learning
- Analytical Data Infrastructure
- Analytical Platforms
- Cloud Computing and Business Intelligence
- Collective Insights®
- Data Catalog
- Data Governance
- Embedded Business Intelligence
- Enterprise Performance Management
- ESG Reporting
- Financial Consolidation, Close Management and Reporting
- Guided Analytics®
- Master Data Management
- ModelOps
- Sales Performance Management
- Self-Service Business Intelligence
- Small and Midsize Enterprise Business Intelligence
- Small and Midsize Enterprise Performance Management
- Supply Chain Planning and Analysis
- Workforce Planning and Analysis

Appendix: 2025 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: _____

What is your current title?

What function are you a part of?

- Business Intelligence Competency Center
- Executive Management
- Finance
- 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 and Accessories
- Automotive
- Aviation
- Biotechnology
- Broadcasting
- Business Services
- Chemical
- Construction
- Consulting
- Consumer Products
- Defense

- Distribution and 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 important is data engineering to your organization?*

- Critical
- Very Important
- Important
- Somewhat Important
- Not Important

What are your plans for employing data engineering?

- Using today
- Using today with plans to expand
- Plan to use in 12 months
- Plan to use in 24 months
- No plans

44) Vendors and Products

Which vendor/product are you using for data engineering?

How satisfied are you with your vendor and product for data engineering?

- Extremely satisfied
- Mostly satisfied
- Somewhat satisfied

- Somewhat unsatisfied
- Unsatisfied

How effective is the current approach to data engineering today?

- Highly Effective
- Somewhat Effective
- Somewhat Ineffective
- Totally Ineffective

How frequently do people in your organization engage in data engineering prior to analysis?

- Constantly
- Frequently
- Occasionally
- Rarely
- Never

How often do users enrich internal data with third-party data (e.g., Google, Facebook, Dun & Bradstreet, U.S. Census)?

- Constantly
- Frequently
- Occasionally
- Rarely
- Never

How do you define, generate, and extract features for machine learning models?

	Primary	Secondary	Tertiary	N/A
Custom SQL, Python, Java Script	()	()	()	()
End user data prep tool	()	()	()	()
ETL tool	()	()	()	()
Data prep capabilities of analytic data infrastructure platform	()	()	()	()

Please indicate the estimated percentage of your BI 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	()	()	()	()	()

Should data engineering be a standalone capability or part of another tool?

- () Standalone
- () Part of business intelligence tools
- () Other - Write In: _____

Please indicate the importance of the following development and usability features for data engineering software:

	Critical	Very Important	Important	Somewhat Important	Not Important
Simple to build/execute data workflow scripts	()	()	()	()	()
Graphical, drag-and-drop designer for creation of analytics data integration workflows	()	()	()	()	()

In-memory engine supporting real-time, interactive exploration and transformation development	()	()	()	()	()
Automated recommendations for data relationships & keys for combining data across multiple data sets and sources	()	()	()	()	()
Machine learning and recommendations based on usage data gathered across users, groups, or organizations	()	()	()	()	()
Library of prebuilt components to access, prepare, and integrate data from sources	()	()	()	()	()
Scheduler for coordinating data workflows	()	()	()	()	()
Debugger for testing and tuning data and processing at any stage of the integration workflow	()	()	()	()	()
Metadata capabilities: Create a view of integrated	()	()	()	()	()

2025 Data Engineering Market Study

data (e.g., from customers, products, suppliers) and securely manage and deliver meta data for analytical use cases/tools					
Connectors to simplify access to data and event/messaging sources, social media, enterprise cloud applications, security services	()	()	()	()	()
Mask or redact sensitive data	()	()	()	()	()
Data profiling/assessment and quality capabilities	()	()	()	()	()
Automated data quality rules	()	()	()	()	()
Sharing and collaboration	()	()	()	()	()
No-code data transformations (this might be covered under graphical, drag-and-drop capabilities)	()	()	()	()	()
Code-friendly data transformations (import/extend/export)	()	()	()	()	()

2025 Data Engineering Market Study

Governance capabilities with audit trails and lineage	()	()	()	()	()
---	-----	-----	-----	-----	-----

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 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.,	()	()	()	()	()

2025 Data Engineering Market Study

Kubernetes, Docker)					
Support for Kafka	()	()	()	()	()
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 values	()	()	()	()	()
Ability to aggregate & group data	()	()	()	()	()
Ability to pivot (convert table to matrix) & reshape (convert matrix to table) data	()	()	()	()	()

2025 Data Engineering Market Study

Ability to derive new data features from existing data (text extraction, math 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	()	()	()	()	()

Which of the following advanced transformations are important to you?

	Critical	Very Important	Important	Somewhat Important	Not Important	Don't Know
Multi-dimensional/OLAP	()	()	()	()	()	()
Path/Link Analysis	()	()	()	()	()	()
Pattern Matching	()	()	()	()	()	()
Aggregations	()	()	()	()	()	()
Statistical Analysis, R	()	()	()	()	()	()
Spark	()	()	()	()	()	()
Graph	()	()	()	()	()	()
Search Analytics	()	()	()	()	()	()
Text Analysis	()	()	()	()	()	()
Ranking/Scoring	()	()	()	()	()	()
User-defined Functions	()	()	()	()	()	()
Machine Learning	()	()	()	()	()	()
Sentiment Analysis	()	()	()	()	()	()

What are the top data sources, targets, and interfaces for use as a part of data engineering?

	Source	Target	Native interface (e.g., bulk loader)	Third party ODBC; JDBC
Relational (Oracle, SQL Server)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Graph (e.g., Neo4J, Tiger graph)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Object stores (e.g., Amazon S3)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Files (e.g., Excel, CSV, Log files, JSON)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Analytical databases (e.g., Snowflake, Exasol, Incorta, AtScale)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
NoSQL (e.g., Mongo, Couchbase)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Hadoop-based (e.g., Cloudera)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Specialty data platforms (e.g., SAP Hana, Palantir)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Applications (e.g., Salesforce, Workday, Oracle, SAP, Infor)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Where should data engineering functionality reside?

	Critical	Very Important	Important	Somewhat Important	Not Important
On Premises	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Private Cloud	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Public Cloud (SaaS)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hybrid (on premises and cloud)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>