



January 16, 2019
Dresner Advisory Services, LLC

2019 Edition

Analytical Data Infrastructure Market Study

Wisdom of Crowds® Series

Licensed to Snowflake



Disclaimer

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

The information contained in all Wisdom of Crowds® Market Study Reports reflects the opinions expressed in the online responses of individuals who chose to respond to our online questionnaire and does not represent a scientific sampling of any kind. Dresner Advisory Services, LLC shall not be liable for the content of reports, study results, or for any damages incurred or alleged to be incurred by any of the companies included in the reports as a result of 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, OLAP (online analytical processing), data mining, and advanced analytics, end-user tools for ad hoc query and analysis, and dashboards for performance monitoring.

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

Analytical Data Infrastructure Defined

Analytical data infrastructure (ADI) defines a set of technology components for integrating, modeling, managing, storing, and accessing the data sets that serve as sources for analytic/BI consumers, e.g., analytic/business applications, tools, and users.

Introduction

In 2019, we look forward to our 12th anniversary at Dresner Advisory Services and our 3rd annual [Real Business Intelligence Conference](#), May 14-15, 2019 on the campus of MIT. We thank our clients, colleagues and community members for their support, which helped us to reach this important milestone.

The choice of which ADI platform to use is a significant decision for business intelligence and analytics users, data administrators/managers, and application developers. In this, our third annual ADI market study, we examine the market preferences and priorities for ADI platforms including deployment and licensing priorities, data types, data model / management of data associated with ADI, data preparation and loading priorities, preferences in ADI development and deployment features, as well as ADI interfaces and analytical features.

Like our other flagship research reports, this ADI market study explores user perceptions and intentions as well as vendor rankings—driven by the voice of the customer, making it a valuable tool for anyone considering investing in ADI products and services.

We hope you enjoy and benefit from this ADI report. It has been a pleasure to develop this report, and our entire team looks forward to serving you in the future. We look forward to hearing from you after you explore the study findings in this report.



Bill Hostmann
Research Fellow
Dresner Advisory Services

Contents

Definitions	3
Business Intelligence Defined.....	3
Analytical Data Infrastructure Defined	3
Introduction	4
Contents.....	5
About Bill Hostmann.....	7
About Howard Dresner and Dresner Advisory Services.....	8
Benefits of the Study	9
A Consumer Guide	9
A Supplier Tool	9
External Awareness	9
Internal Planning	9
Survey Method and Data Collection	10
Data Quality	10
Executive Summary.....	12
Study Demographics	14
Geography.....	14
Functions	15
Vertical Industries	16
Organization Size	17
Analysis and Trends.....	19
Analytical Data Infrastructure Use Cases	20
Selection Priorities for Analytical Data Infrastructure	26
Analytical Data Infrastructure Deployment and Licensing.....	33
ADI Deployment Priorities	33
ADI Licensing Preferences.....	40
Analytical Data Infrastructure Development and Deployment Features.....	47
Analytical Data Infrastructure Data Types	54
Analytical Data Infrastructure Data Preparation and Loading	61
Analytical Data Infrastructure Data Model / Management of Data	68
Analytical Data Infrastructure Interfaces	75

2019 Analytical Data Infrastructure Market Study

ADI Analytical Features	82
Vendor Ratings	89
Customer Experience Model	90
Vendor Credibility Model.....	92
Detailed Vendor Ratings	94
Amazon Detailed Score	95
Cloudera Detailed Score.....	96
Exasol Detailed Score	97
Google Detailed Score.....	98
Hortonworks Detailed Score	99
IBM Detailed Score.....	100
Microsoft Detailed Score.....	101
Oracle Detailed Score.....	102
Postgres Detailed Score	103
SAP Detailed Score	104
SAS Detailed Score	105
Snowflake Detailed Score.....	106
Teradata Detailed Score	107
Other Dresner Advisory Services Research Reports	108
Appendix: Analytical Data Infrastructure Survey Instrument	109

About Bill Hostmann

Bill Hostmann is a 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 product management 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 served as conference chair of the Gartner BI/Analytics Conference for many years, growing the number of conference attendees from hundreds to several thousand attendees.

About Howard Dresner and Dresner Advisory Services

DAS market studies are 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



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” Business Intelligence Market study](#)
- [Advanced and Predictive Analytics](#)
- [Cloud Computing and Business Intelligence](#)
- [End User Data Preparation](#)
- [Enterprise Planning](#)
- [Location Intelligence](#)
- [IoT Intelligence®](#)

Howard (www.twitter.com/howarddresner) conducts a weekly Twitter “tweetchat” on Fridays at 1:00 p.m. ET. The hashtag is #BIWisdom. During these live events, the #BIWisdom community discusses a wide range of business intelligence topics.

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

Benefits of the Study

The Dresner Advisory Services' Analytical Data Infrastructure Market Study provides a wealth of information and analysis, offering value to consumers and producers of BI/analytics technology and services.

A Consumer Guide

As an objective source of industry research, consumers may use this report to understand how their peers use and invest in business intelligence and related technologies. Using our trademark vendor performance measurement system, users can glean key insights into BI software supplier performance, enabling:

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

A Supplier Tool

Vendor licensees may use the report in several important ways:

External Awareness

- Build awareness for the business intelligence market and supplier brand, citing the Analytical Data Infrastructure Market Study trends and vendor performance
- Create lead and demand generation for supplier offerings through association with the report findings, webinars, etc.

Internal Planning

- Refine internal product plans and align with market priorities and realities as identified in the Analytical Data Infrastructure Market Study
- Better understand customer priorities, concerns, and issues
- Identify competitive pressures and opportunities

Survey Method and Data Collection

As with all our Wisdom of Crowds® Business Intelligence Market Studies, we constructed the study from a survey instrument to collect data and used social media and crowdsourcing techniques as well as our own research community of over 5,000 organizations to recruit study participants. We conducted the study and collected the data for this 2019 report during Q3 and Q4 2018.

Data Quality

We scrutinized and verified all respondent entries to ensure that only qualified participants are included in the study. Among this study's "qualified participants" are analytic product users, developers, and data (integration, management, etc.) professionals involved with ADI product uses and implementation decisions.

Executive Summary

Executive Summary

In our 2019 Analytical Data Infrastructure Market Study, we used the following use cases to analyze the market survey responses:

- Business user reporting and dashboards
- Business user discovery and exploration
- Data science (e.g., advanced and predictive workloads and workflows)
- Embedded analytics (e.g., analytic functions and data embedded within business applications for higher volume/low latency applications).

Each use case presents a different combination of data workloads and analytical workflows to an ADI platform. The goal of this report is to better understand the priority of use cases and preferences for Analytical Data Infrastructure features/capabilities such as performance versus cost, data-integration priorities, and development and deployment preferences. Understanding these capabilities, uses, and adoption will help with the prioritization planning, developing, and execution of a BI and analytics strategy for any size organization.

- The top use case most identified as a top priority for ADI platforms is business user reporting and dashboards. For most organizations (83 percent of respondents), this is the most important use case when considering ADI platforms. The data science use case is a priority for 43 percent of respondents. Embedded analytics, i.e., a use case within business operational applications work and data flows and requiring low latency analytics and typically high data volumes is a priority for evaluating ADI platforms for 35 percent of the market.
- Cost and corporate standards are a low relative priority for ADI platforms compared to performance and security priorities. We think this will lead to further ADI platform fragmentation and associated data and analytics fragmentation across organizations.
- The preference (60 percent of responses) is for a single ADI platform that can support multiple use cases and workloads/workflows (e.g., it must provide capabilities for business user reporting and dashboards as well as business discovery and exploration or data science analytic use cases).
- The majority of respondents prefer an ADI platform accessed/licensed via a cloud deployment (“as a cloud service” versus “on-premises software”).
- Data integration and management capabilities to support hybrid deployments (cloud and on premises) is a top ADI platform priority for 30 percent of the market. We expect cross data center integration and management tools for hybrid ADI platform deployments will increase in priority and the range of capabilities and options will become extensive as the technology/market develops.

- The range of innovation and variety of ADI platform capabilities and diversity of use cases in the market today and the lack of priority on corporate standards and governance, makes developing a business and technical strategy (for using data and analytics to drive business change) for larger scale, cross-functional, multi-use case BI and analytics projects, more difficult than ever for business and technical leaders.

Study Demographics

The respondents in this survey provide a cross-section of geographies, functions, organization sizes, and vertical industries. We believe this is a representative sample and more useful indicator of true market dynamics. We constructed cross-tab analyses using these demographics to identify and illustrate important industry trends.

Geography

North America, which includes the U.S., Canada, and Puerto Rico, represents 65 percent of respondents (fig. 1). EMEA accounts for 27 percent of respondents. Asia Pacific and Latin America each represent 4 percent of the respondents.

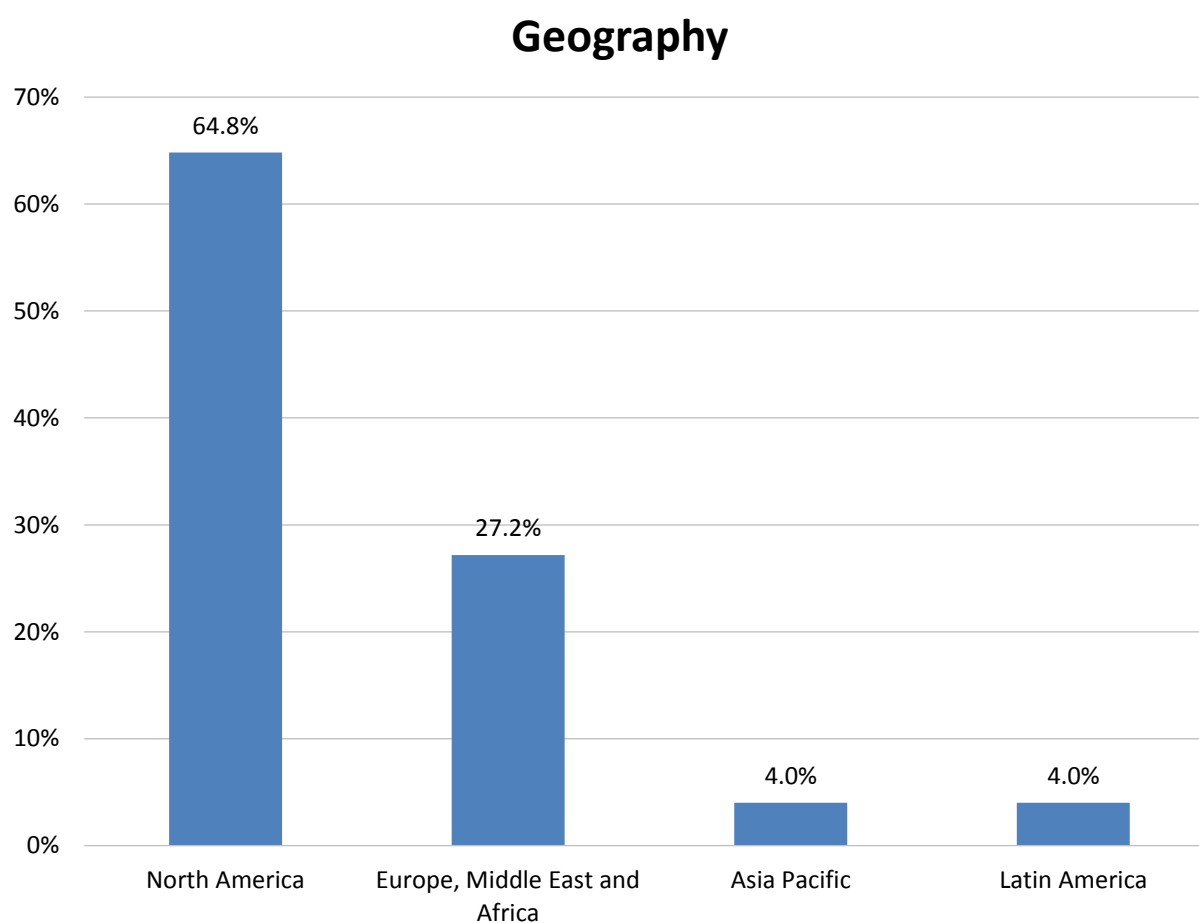


Figure 1 – Geography

Functions

IT (41 percent) and the Business Intelligence Competency Center (BICC) (22 percent) are the largest groups represented in our ADI sample (fig. 2). Executive Management comprises 11 percent. The remaining respondents include R&D, Finance, and Sales/Marketing functions.

Examining preferences by function helps us compare and contrast different priorities for different buying centers. And, as you will see in the report, the differences in priorities are of note and have an impact on best practices, licensing, skills, technologies, and assumptions about economies of scale from ADI investments.

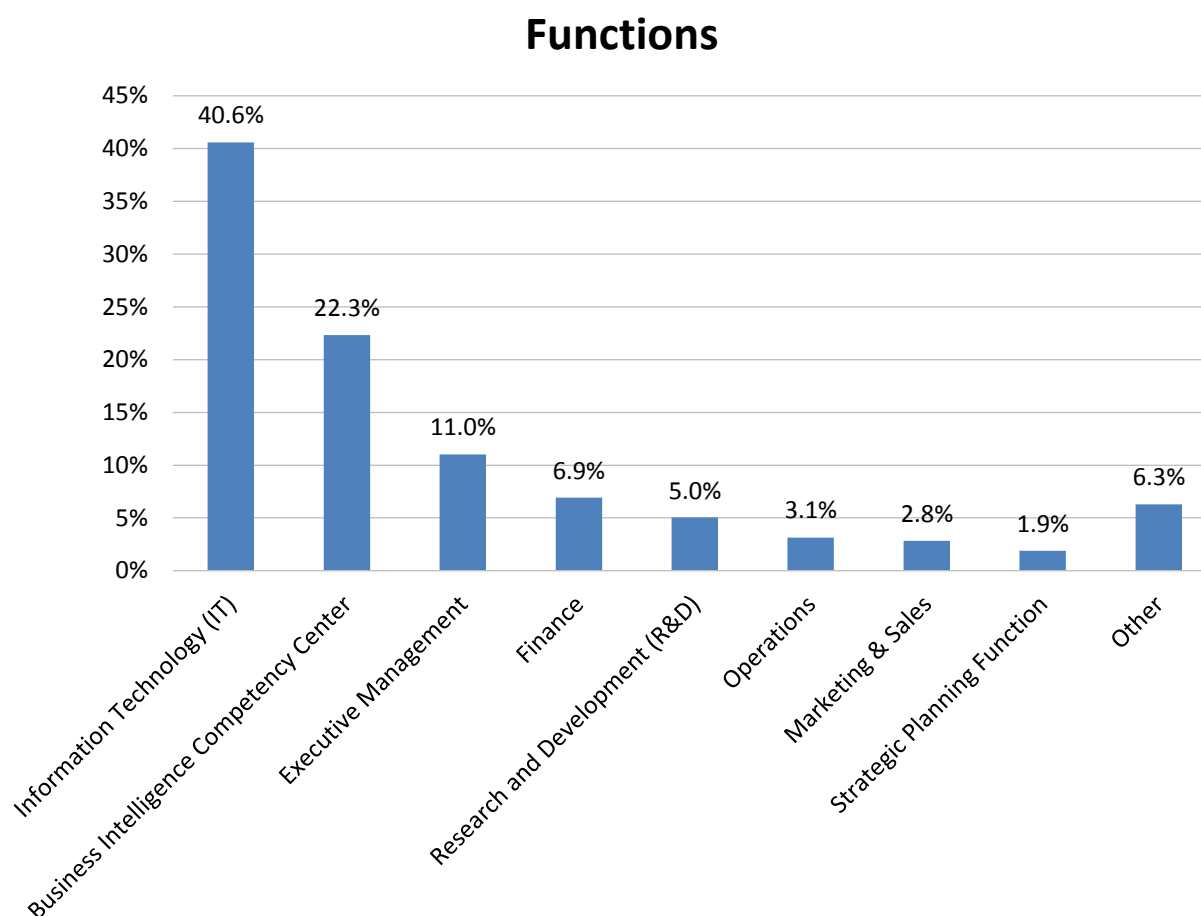


Figure 2 – Functions

Vertical Industries

Respondent organizations represent a mix of vertical industries. Technology (12 percent), Healthcare (10 percent), Financial Services (9 percent), Retail/Wholesale (6 percent), Education (5 percent), Manufacturing (5 percent), Consulting (5 percent) and Telecommunications (5 percent) represent the majority of survey respondents. We include responses from consultants—who often have high a degree of interaction with ADI-related initiatives and deeper industry knowledge. This also yields insight into the partner ecosystem for ADI vendors (fig. 3).

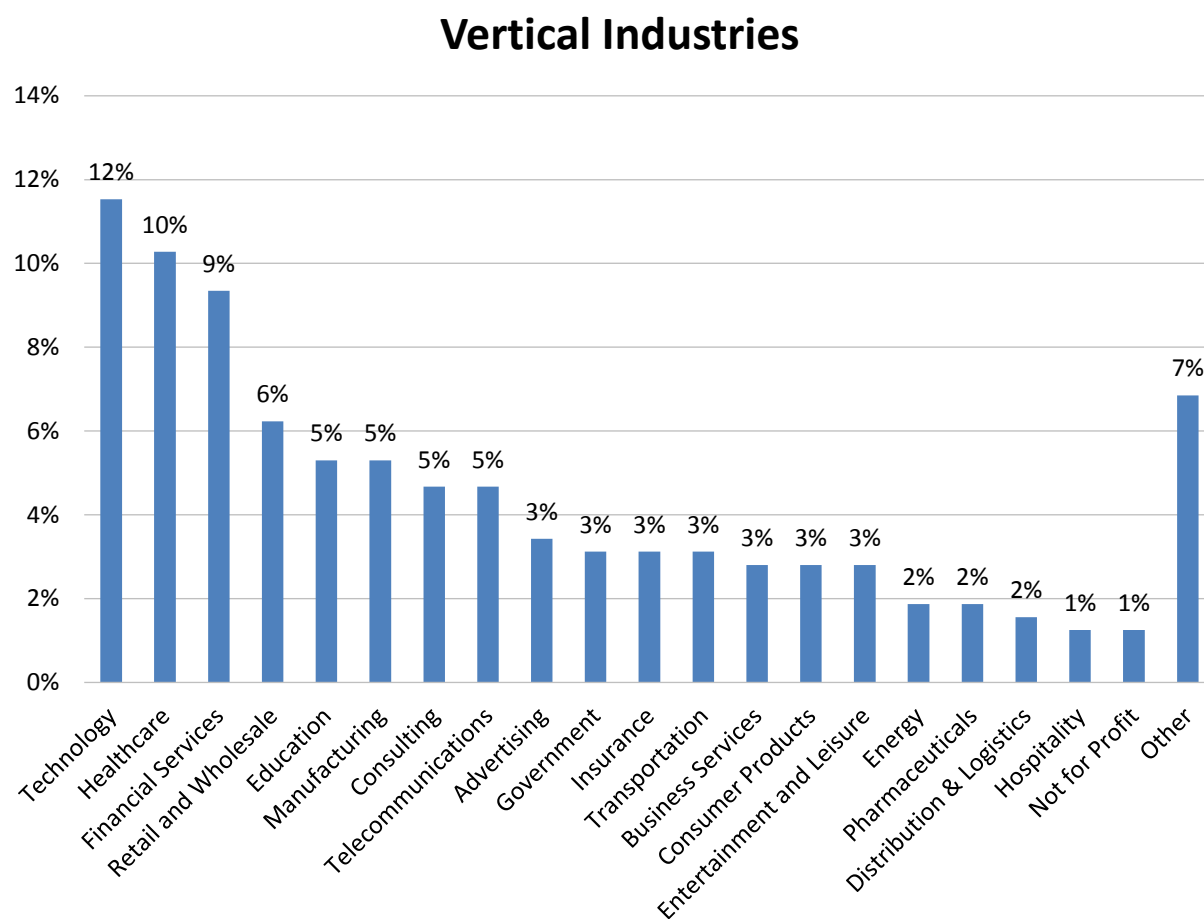


Figure 3 – Vertical industries

Organization Size

Tabulating results by organization size reveals important differences in practices and product priorities. Respondents reflect a mix of organizational sizes and structures (fig. 4). ADI investments and technologies are not just important to very large enterprises; 21 percent of respondents are from small organizations with fewer than 100 employees, and mid-sized organizations with 100-1,000 employees account for another 31 percent.

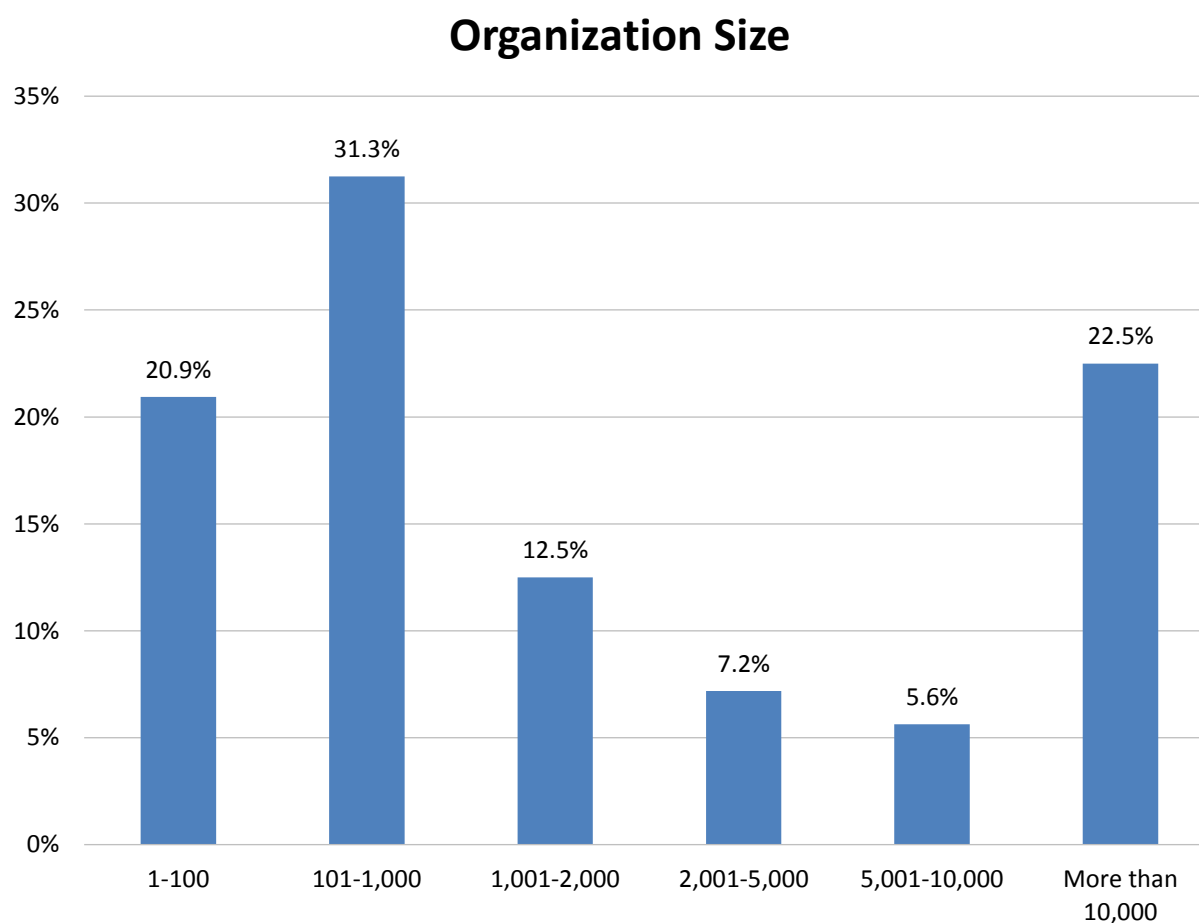


Figure 4 – Organization size

Analysis and Trends

Analysis and Trends

This report describes the current priorities and trends for the analytical data infrastructure market for business intelligence. We sampled organizations' experience with use cases, selection criteria, and ADI priorities (scale of one to five with five being highest priority) including development features, ADI deployment priorities (cloud, on premises, etc.), licensing preferences, data types, data modeling, interfaces, analytical features, and more.

We also asked respondents to rate the performance of ADI vendors on a scale ranging from "very poor" to "excellent" across sales/acquisition experience, value for price paid, quality and usefulness of product, quality of technical support, quality and value of consulting services, whether the vendor is recommended, and integrity.

Analytical Data Infrastructure Use Cases

There is a diversity of ADI use cases. Different use cases have different data and analysis workloads/workflows and buying requirements. Buying requirements for ADI platforms can reflect combinations of use-case priorities. We asked respondents to rank the importance of four types of ADI use cases. The majority (83 percent) of respondents indicate “business user reporting and dashboards” as their highest use case driving ADI requirements and priorities, ranking it as “critical” or “very important” (fig. 5). The second-highest use case is “business user discovery and exploration,” which 67 percent of respondents rank as “critical” or “very important.” Fifty-three percent of respondents rank “data science (advanced and predictive analytics or data mining)” use cases as their highest priority, and 45 percent rank “embedded analytics with business applications” as their highest priority use cases. Sixty percent of respondents identify multiple use cases as top priorities for their ADI platforms.

ADI Use Cases

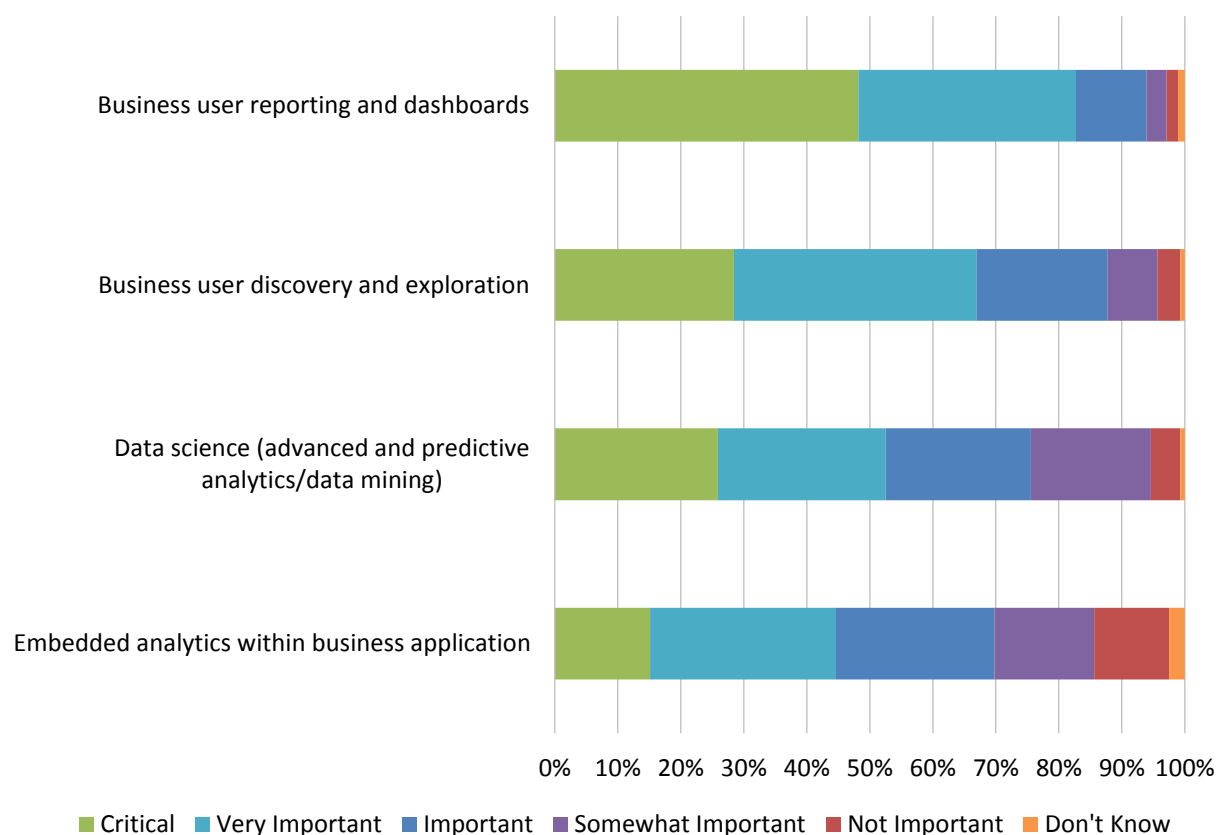


Figure 5 – ADI use cases

An early, albeit slight, trend is emerging in changing priorities for ADI use cases. Year over year, “business user reporting and dashboards” and “business user discovery and exploration” are the top use cases; however, they decline very slightly in priority (fig. 6) year over year. The year-over-year changes in ADI priorities for the data science and embedded analytics use cases continues to increase slightly, much as they did in our 2018 report. We believe that as organizations become more mature in their experience with BI, they expand to more advanced functionality and workflows found in data science and embedded analytics use cases.

ADI Use Cases 2017-2019

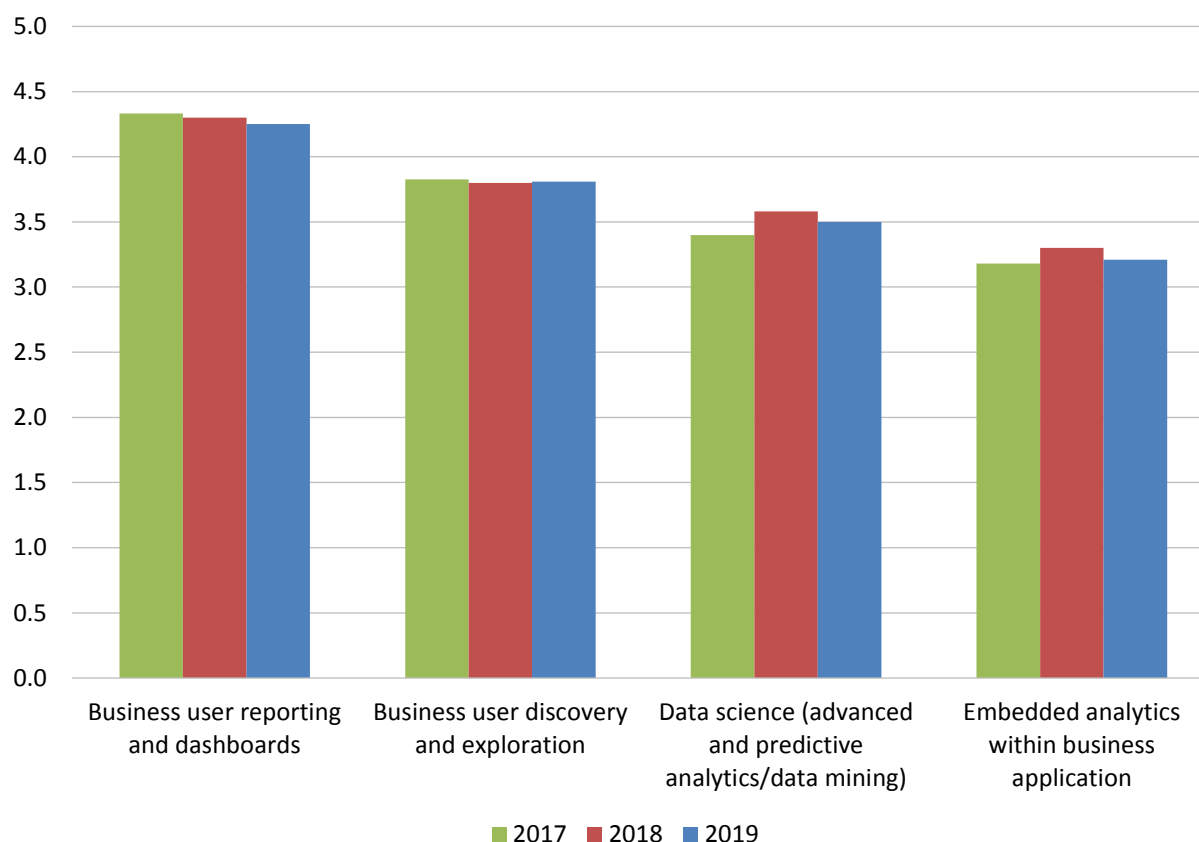


Figure 6 – ADI use cases 2017-2019

In 2018, across all geographies, “Business user reporting and dashboards” and “business user discovery and exploration” are top use-case priorities for ADI platforms and investments (fig. 7). Latin American respondents indicate a lower priority for “business user discovery and exploration.” In Asia-Pacific and Latin America, “embedded analytics” use cases are higher priority than “data science” use cases. “Embedded analytics in business applications” is the lowest, albeit rising, use-case priority driving ADI investments across all geographies (fig. 7).

ADI Use Cases by Geography

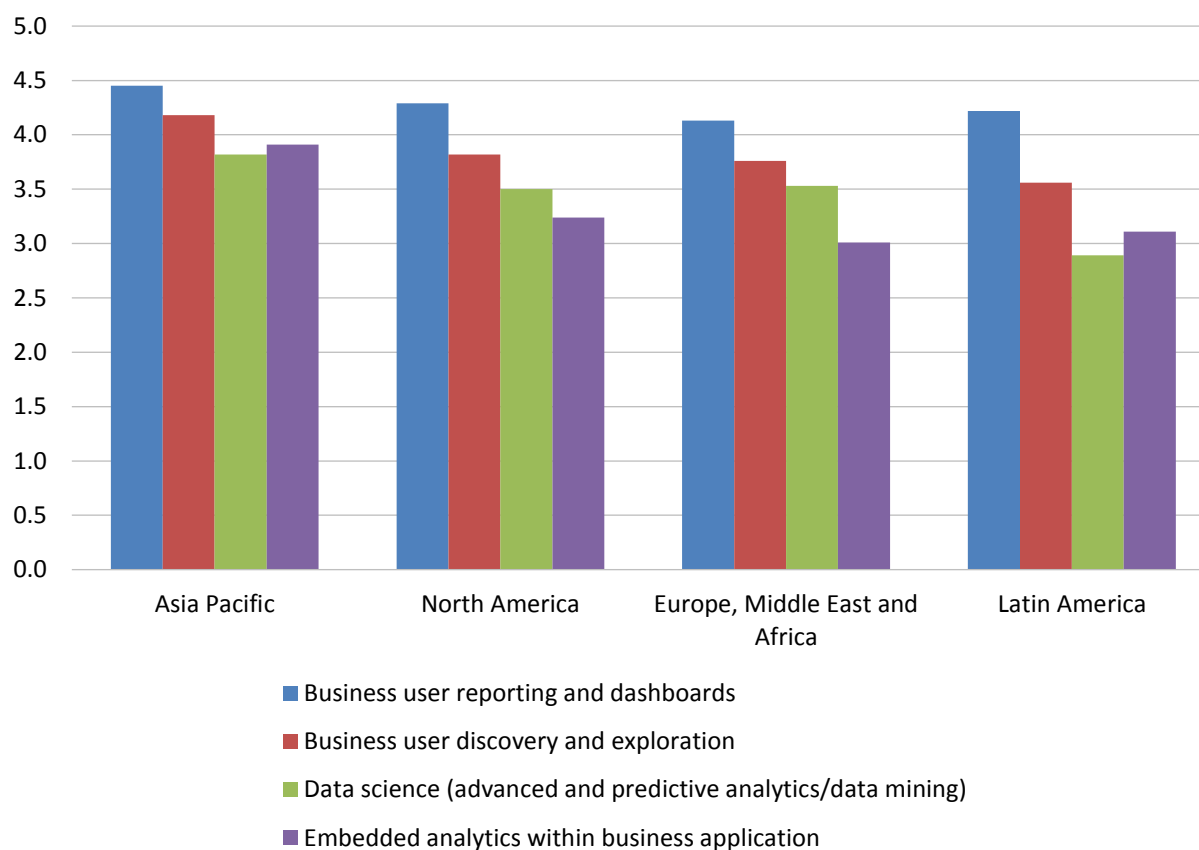


Figure 7 – ADI use cases by geography

Survey respondents from different organizational functions indicate differing use-case priorities for their ADI requirements. Across all functions, “business user reporting and dashboards” is the highest-priority use case for ADI, followed by “business user discovery and exploration” (fig. 8). Most organizational functions have similar profiles in use-case priorities, apart from Sales/Marketing, Operations, and R&D respondents. R&D and Marketing/Sales functions place a higher-than-average priority on “embedded analytics within business applications.”

Recognizing the different use case priorities by function is important to factor into the evaluation of ADI vendors and products. One ADI platform may or may not serve all the combination of functional use case priorities a company may have.

ADI Use Cases by Function

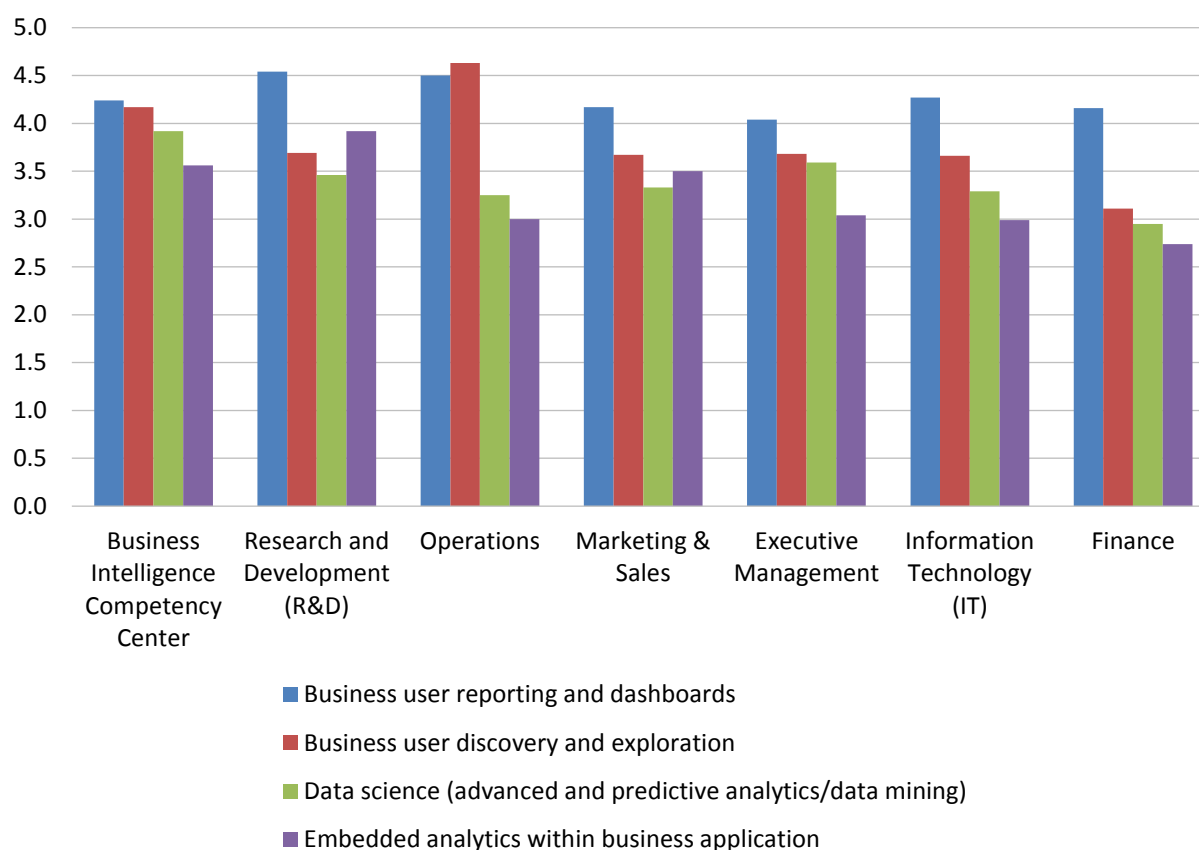


Figure 8 – ADI use cases by function

Organizations from different industries have different priorities for ADI platforms. This is shown by the combination of use-case priorities (fig. 9). The respondents from Telecommunications and Financial Services industries place the highest priorities on multiple use cases for ADI platforms. Respondents from the Advertising industry place the highest priority on the “data science” use case; in fact, this use case is a higher priority than even “business user reporting and dashboards.” Respondents from Government organizations place equal priority on “business user and dashboards” and “business user discovery and exploration.”

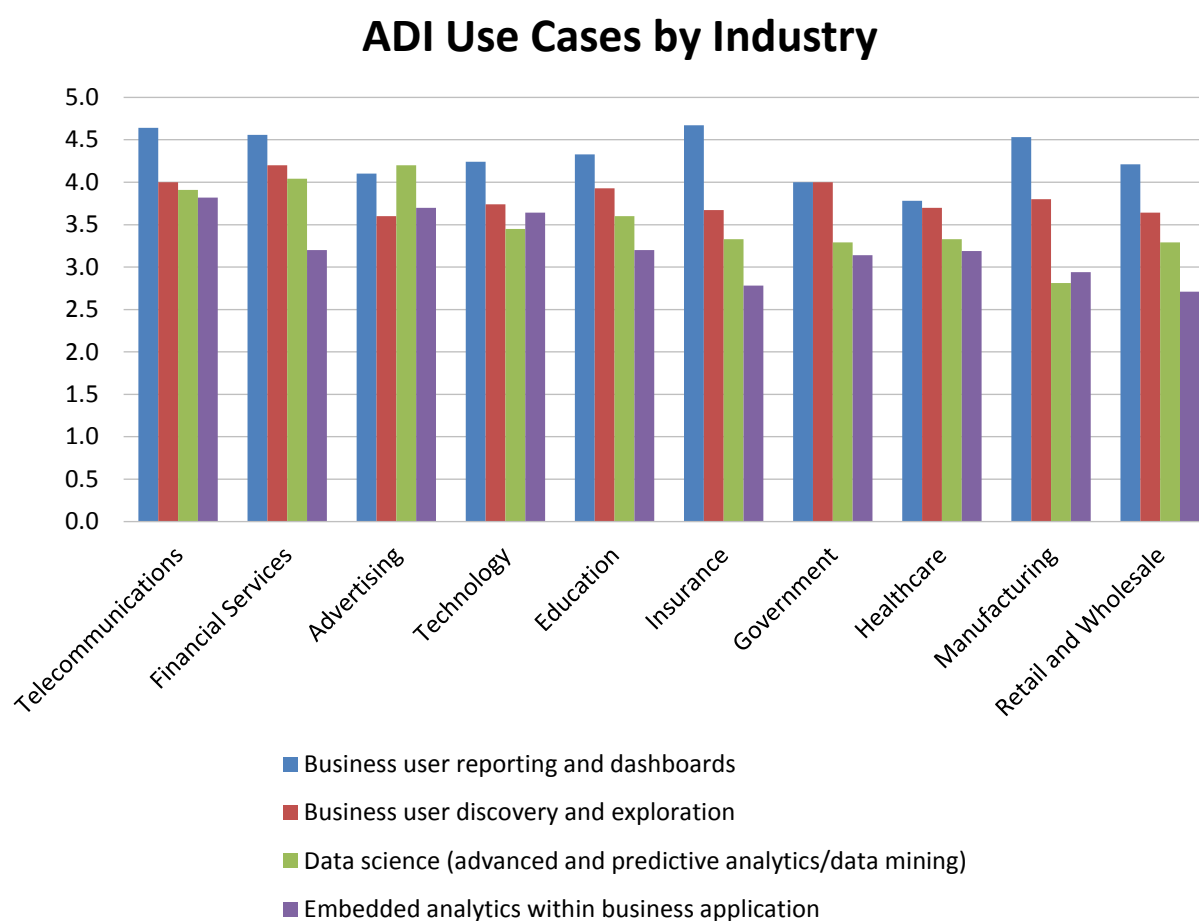


Figure 9 – ADI use cases by industry

Like other market dimensions, organizations, regardless of size, place “business users reporting and dashboards” as their highest (“critical” and “very important”) ADI priority (fig 10). The overall ADI use-case priority profiles are similar across the organization-size market dimension; that is, ADI use-case priorities are somewhat independent of organization size. Respondents from organizations with more than 5,000 employees place a higher priority on “business user discovery and exploration,” “data science,” and “embedded analytics” than the rest of the market. This is perhaps due to a longer-term experience base with BI and analytical data infrastructure investments.

ADI Use Cases by Organization Size

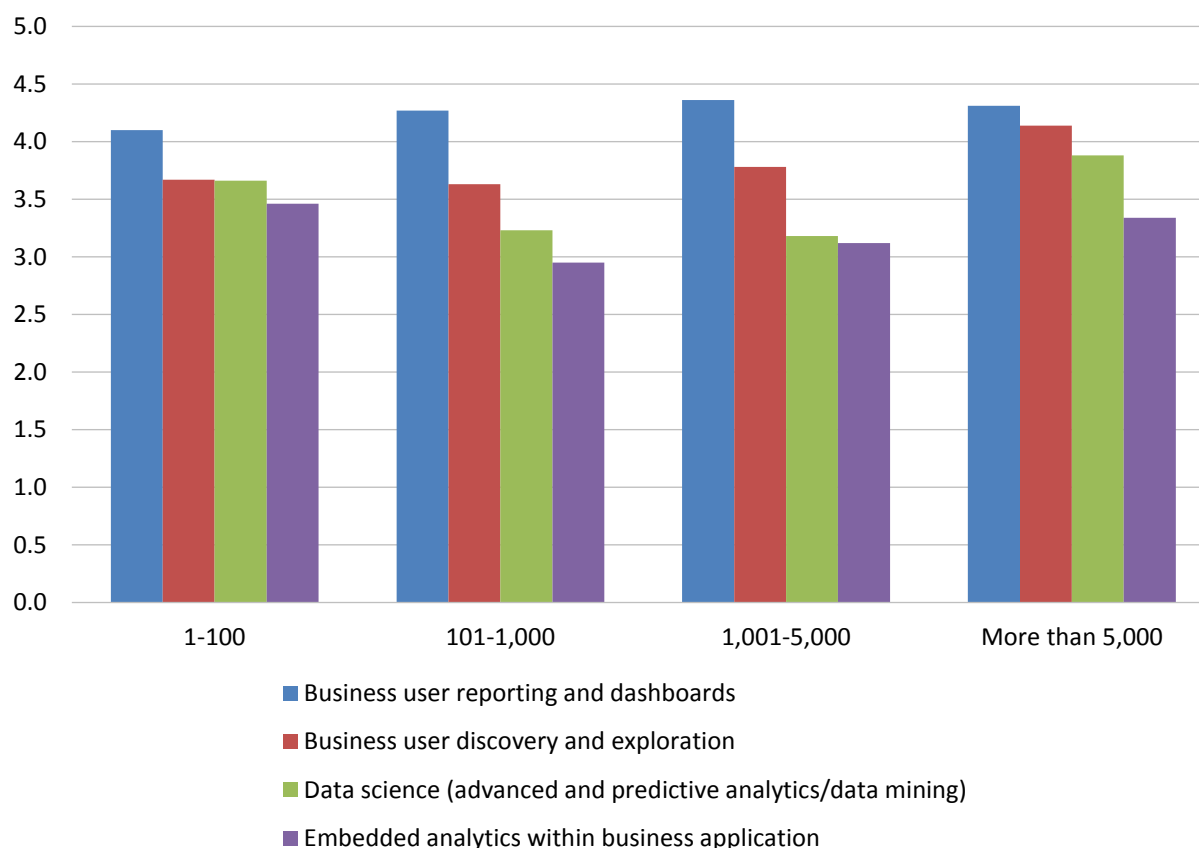


Figure 10 – ADI use cases by organization size

Selection Priorities for Analytical Data Infrastructure

In this section, we discuss respondents' priorities covering price/performance product-related qualities such as scalability, usability, etc. Performance and security are the highest selection priorities for ADI platforms in 2019. This is driven by the ever-growing volume of data, algorithms, and number of users arising from combinations of use cases. An ADI platform must support the volume and combinations of data, workloads, integrations, etc. required from the combinations of use cases and their workloads and workflows. We added a new question to our survey this year, which shows respondents place a relatively lower priority on ADI platforms' ability to meet compliance or regulatory requirements, ranking these features lower than most other ADI requirements (fig. 11). "Corporate standards" and "price" are relatively low priorities, demonstrating the diversity of ADI platforms/skills/workflows and that, for most organizations, no one ADI platform can be a "corporate standard."

Overall Selection Priorities for ADI

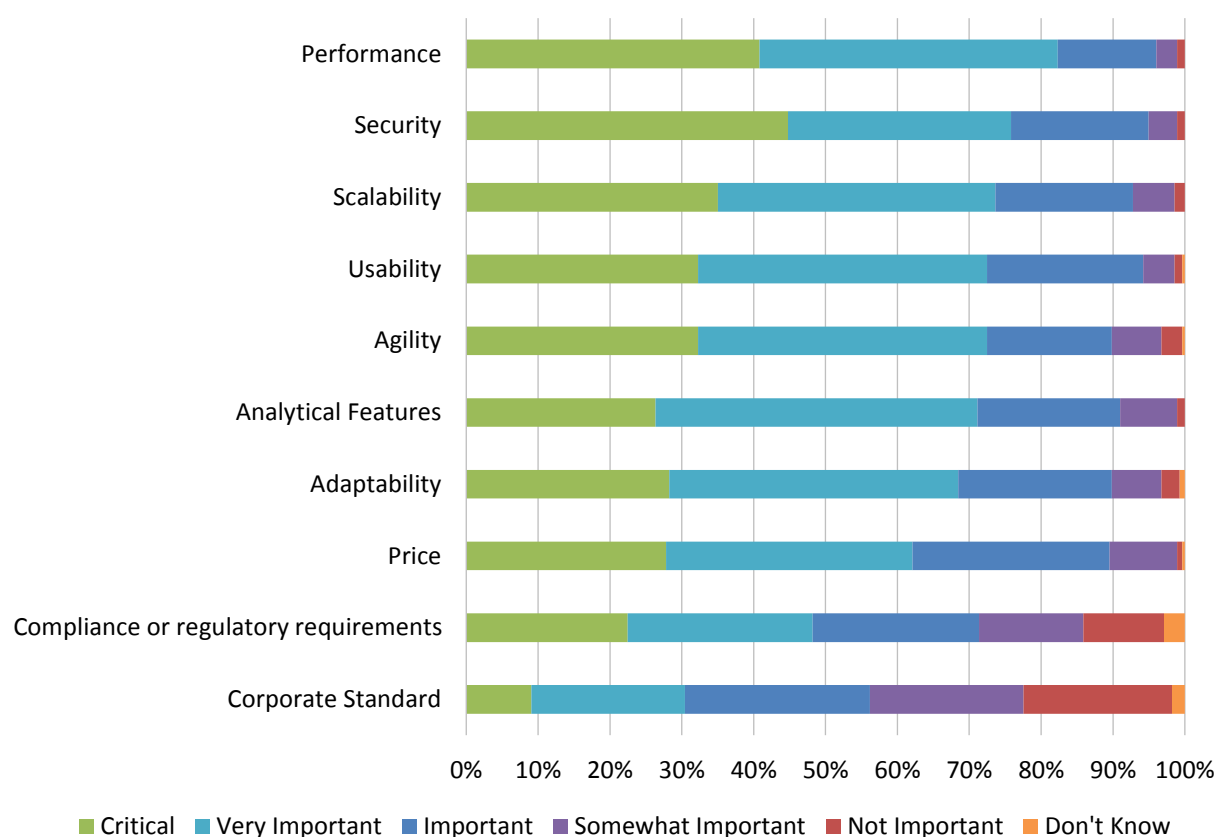


Figure 11 – Overall selection priorities for ADI

Performance leads the selection priorities for ADI platforms in 2019. Security shows a rising trend upward in priority over the past three years (fig. 12). Usability remains a high priority year over year. And, like last year, respondents rank price and corporate standards as their lowest selection priorities, with the priority of corporate standards declining slightly year over year.

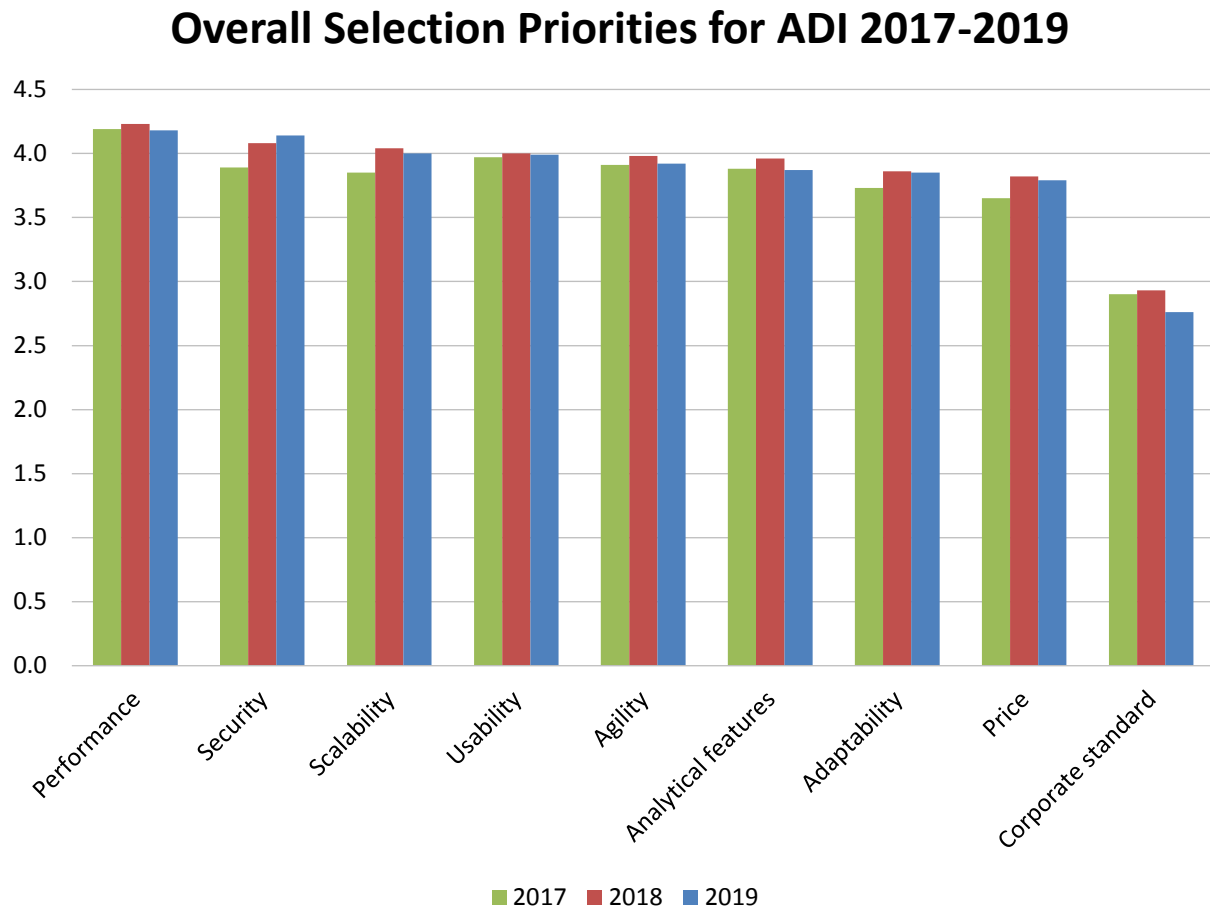


Figure 12 – Overall selection priorities for ADI 2017-2019

Selection priorities for ADI platforms varies by use-case workload and workflow (fig. 13). Price is becoming comparatively a higher priority for embedded analytics and business user discovery and exploration. “Business discovery and exploration” and data science use cases place a higher priority on analytic functions within the ADI platform compared to other use cases.

Overall Selection Priorities for ADI by Top Use Case

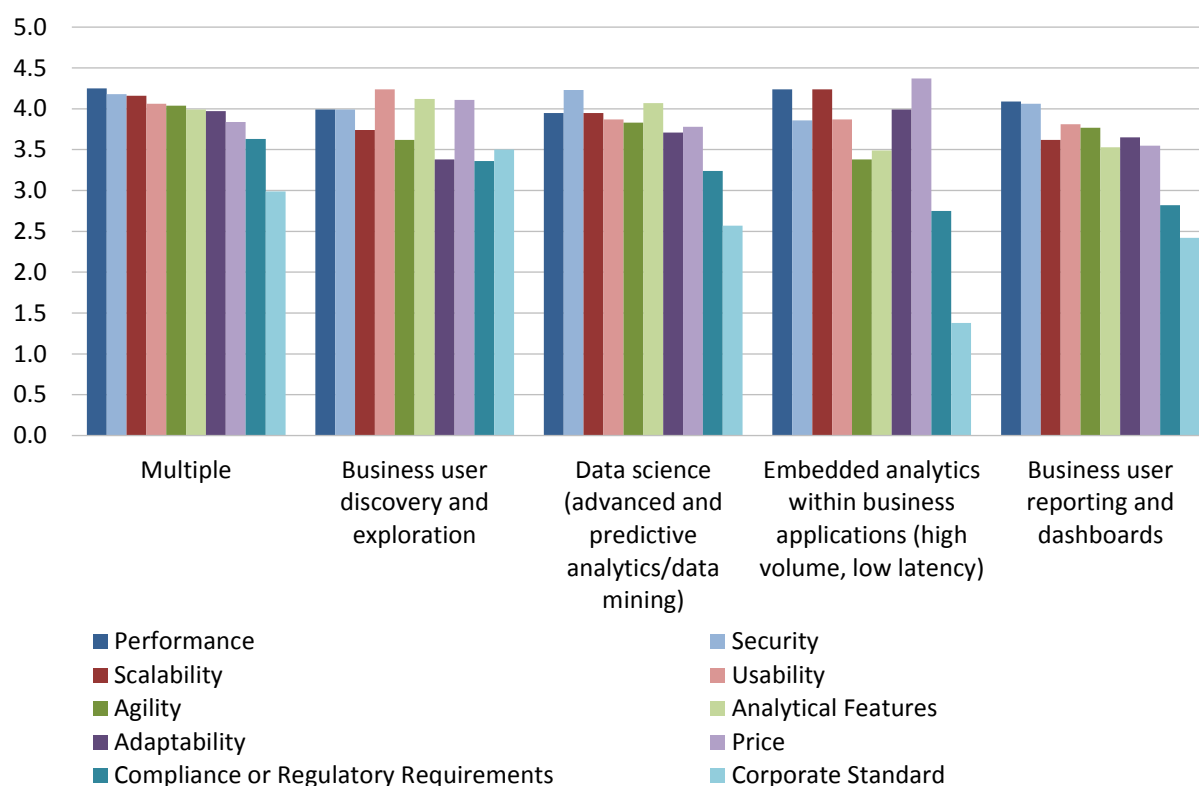


Figure 13 – Overall selection priorities for ADI by top use case

There are specific differences in ADI priorities across the geographic market dimension. All regions consider a “corporate standard” as the lowest priority for selecting ADI (fig. 14). Performance and security are the top selection priorities across all geographies; however Asia-Pacific respondents rate “security” as a higher priority than “performance.” There are distinct differences in respondents’ view of “compliance/regulatory requirements,” with Asia Pacific ranking it a high priority and Latin American respondents ranking it as a low priority.

Overall Selection Priorities for ADI by Geography

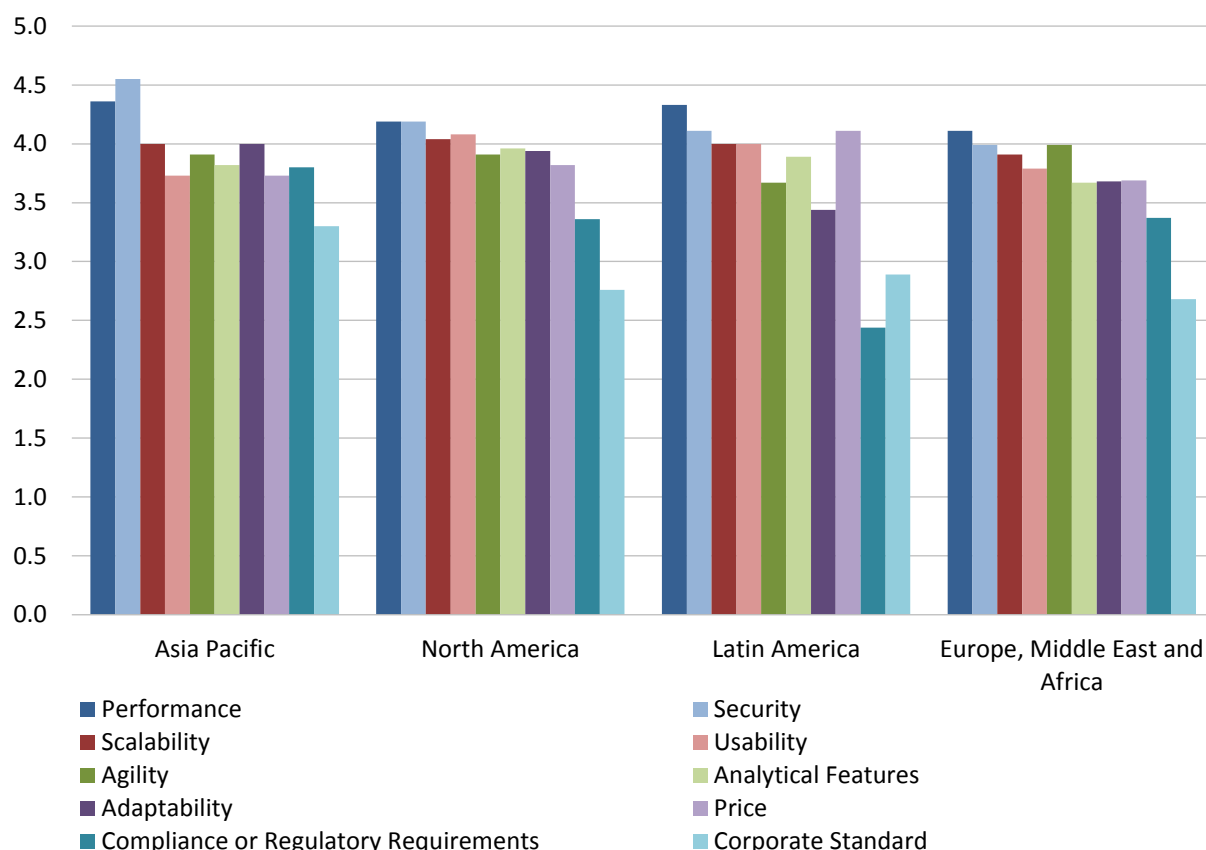


Figure 14 – Overall selection priorities for ADI by geography

As shown below, the relative priorities for ADI platforms vary by organization functions. “Performance” is a top selection priority among BICC, R&D, and Operations functions (fig. 15). Operations, however, places “Security” as its highest priority. Adaptability and security are the top two priorities for Marketing/Sales. Security is the top priority for both the Finance and Executive Management functions. Analytic features rank highest for respondents from the Operations function. Interestingly, respondents from the Finance functions rank “compliance or regulatory requirements” and “corporate standards” as their lowest ADI platform priorities. All organization functions rate “corporate standard” as their lowest selection priority and are fairly price insensitive, assuming all their other priorities are met/exceeded.

Overall Selection Priorities for ADI by Function

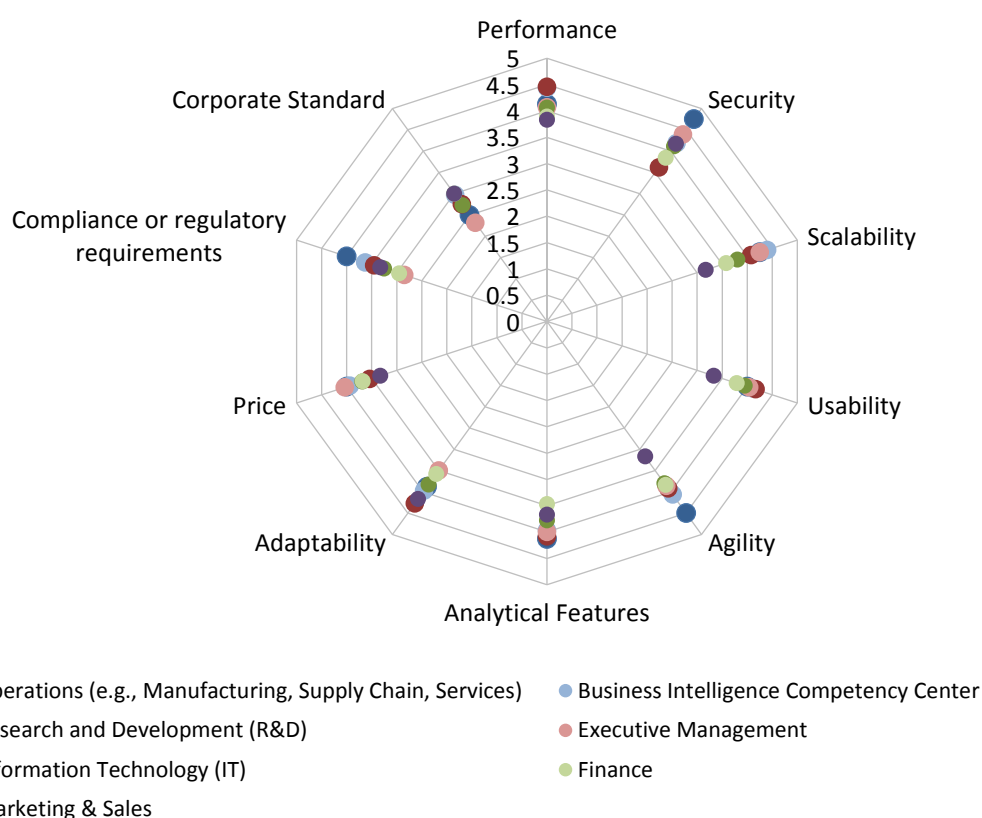


Figure 15 – Overall selection priorities for ADI by function

ADI platform priorities vary across Industries (fig 16). Some variation of priorities across industries is notable and, as a result, there are some ADI platforms that are better suited for different industries and their users/use cases. The priority for a “corporate standard” is low but varies across industries. Security is the highest priority for the Financial, Healthcare, Education, and Retail/Wholesale sector users / use cases. Advertising, Retail/Wholesale, and Manufacturing rank performance as their top ADI platform priority. Usability of the ADI platform is the top priority for our Insurance industry survey participants. The Telecommunication respondents rank “agility” as their top priority.

Overall Selection Priorities for ADI by Industry

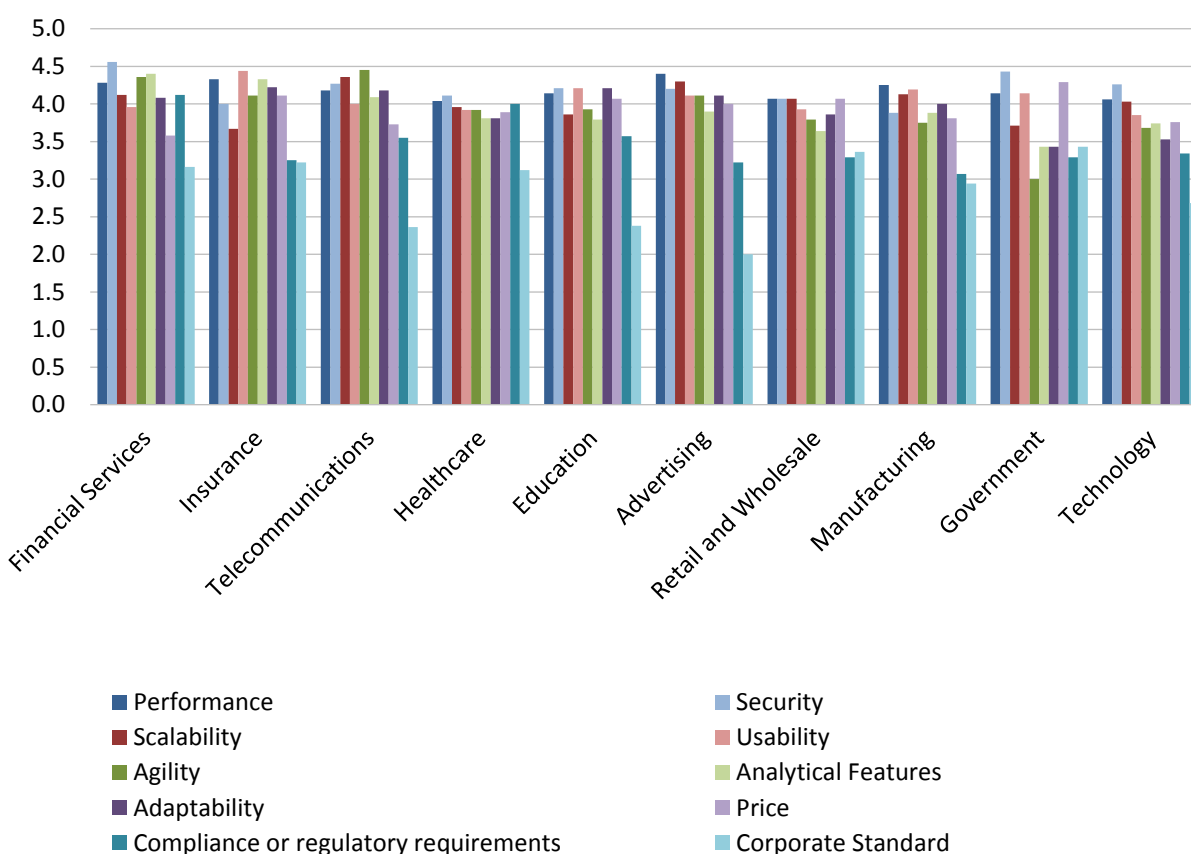


Figure 16 – Overall selection priorities for ADI by industry

ADI platform priorities also vary by size of organization. Except for mid-size organizations (1,000-5,000 employees), performance is the highest priority. Mid-sized organizations indicate security as their top selection priority, followed by performance as their second-highest priority (fig. 17). Not surprisingly, larger organizations (more than 5,000 employees) prioritize performance, security, and scalability as their top needs. “Compliance or regulatory requirements” and “corporate standard” are the lowest selection priorities across organizations of all sizes for their ADI platforms.

Overall Selection Priorities for ADI by Organization Size

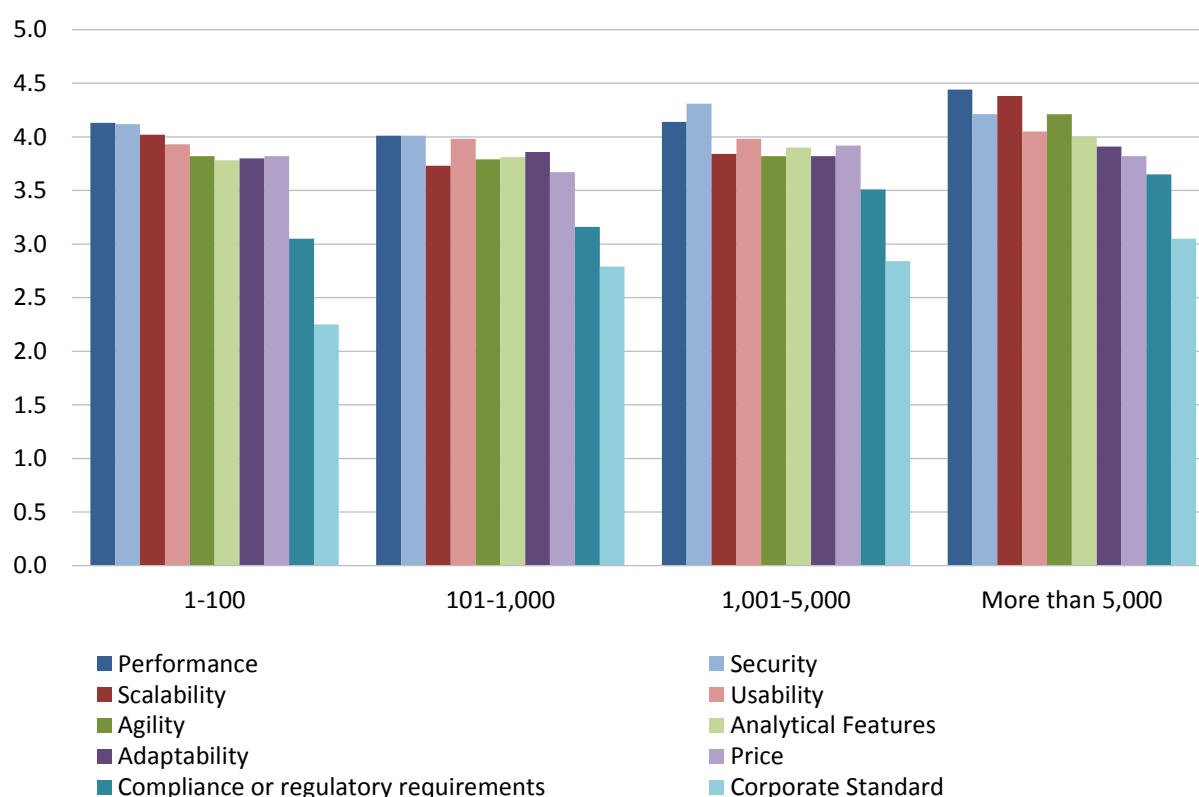


Figure 17 – Overall selection priorities for ADI by organization size

Analytical Data Infrastructure Deployment and Licensing

ADI Deployment Priorities

A trend of note—and for the first time—cloud deployment of ADI platforms is the highest priority (“critical” or “very important”) deployment option for the majority of our survey respondents. A significant number of organizations also need a distributed/hybrid deployment model with cross data center integration and management capabilities. “Cross data center integration and management capabilities” and “hybrid deployment (cloud and on premises)” are a top ADI priority for more than 30 percent of respondents and also shows an increase from last year’s priorities (fig. 18).

ADI Deployment Priorities

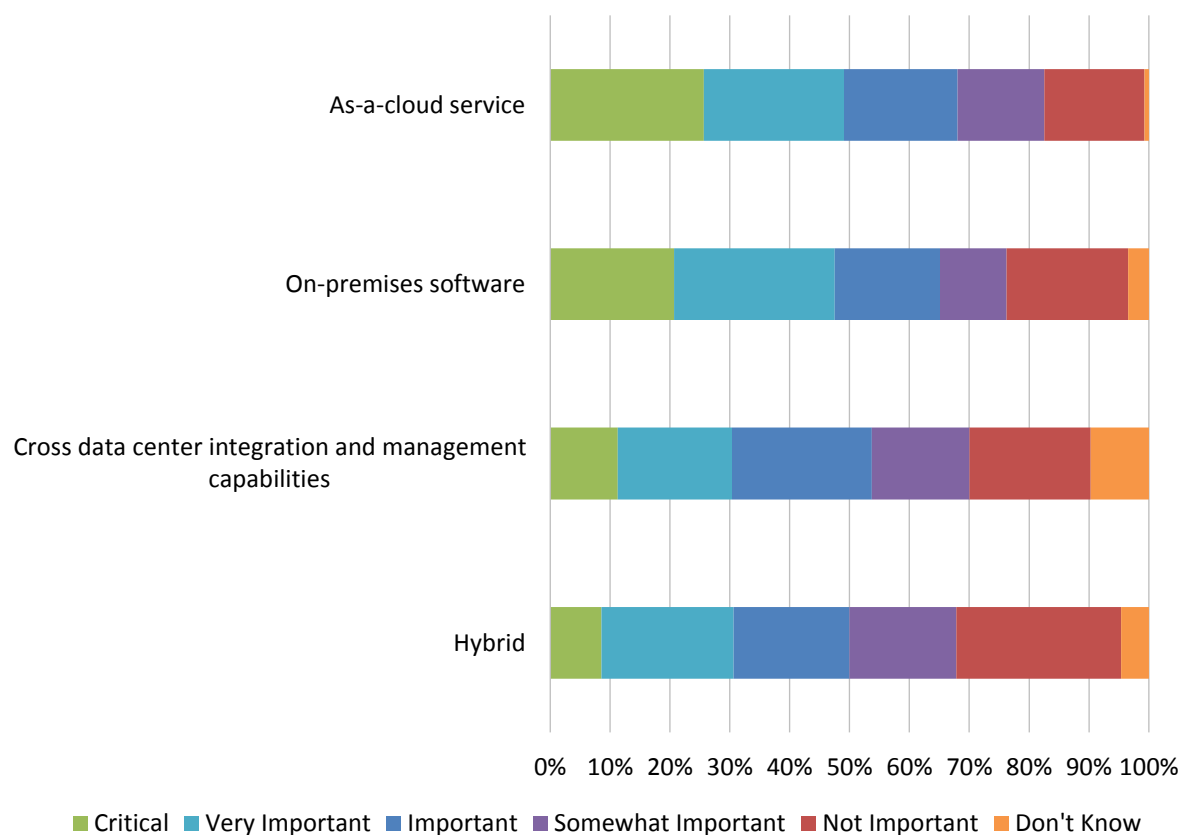


Figure 18 – ADI deployment priorities

Some response trends of note regarding deployment priorities emerge when looking at the year-over-year data sets (fig. 19). Choosing an ADI offered as a cloud service picks up momentum as a preference whereas the on-premises ADI preference softens. Cross data center integration and management capabilities for ADI platforms trend upward year over year as well.

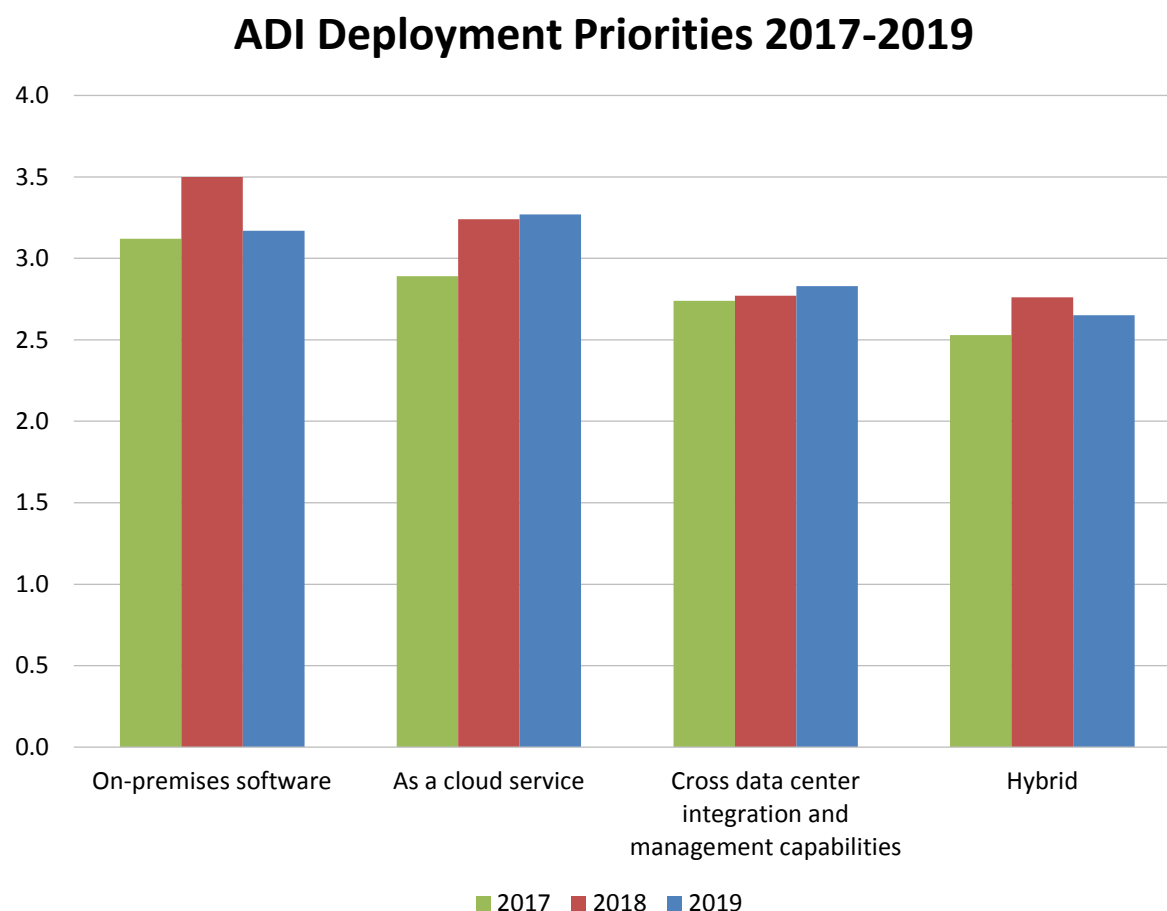


Figure 19 – ADI deployment priorities 2017-2019

When the ADI platform needs to serve multiple use cases, the “as a cloud service” option is the deployment priority for respondents (fig. 20). On premises is the deployment preference for “business user discovery and exploration” and “business user reporting and dashboards.” The deployment priority for embedded use cases is “as a cloud service.”

ADI Deployment Priorities by Top Use Case

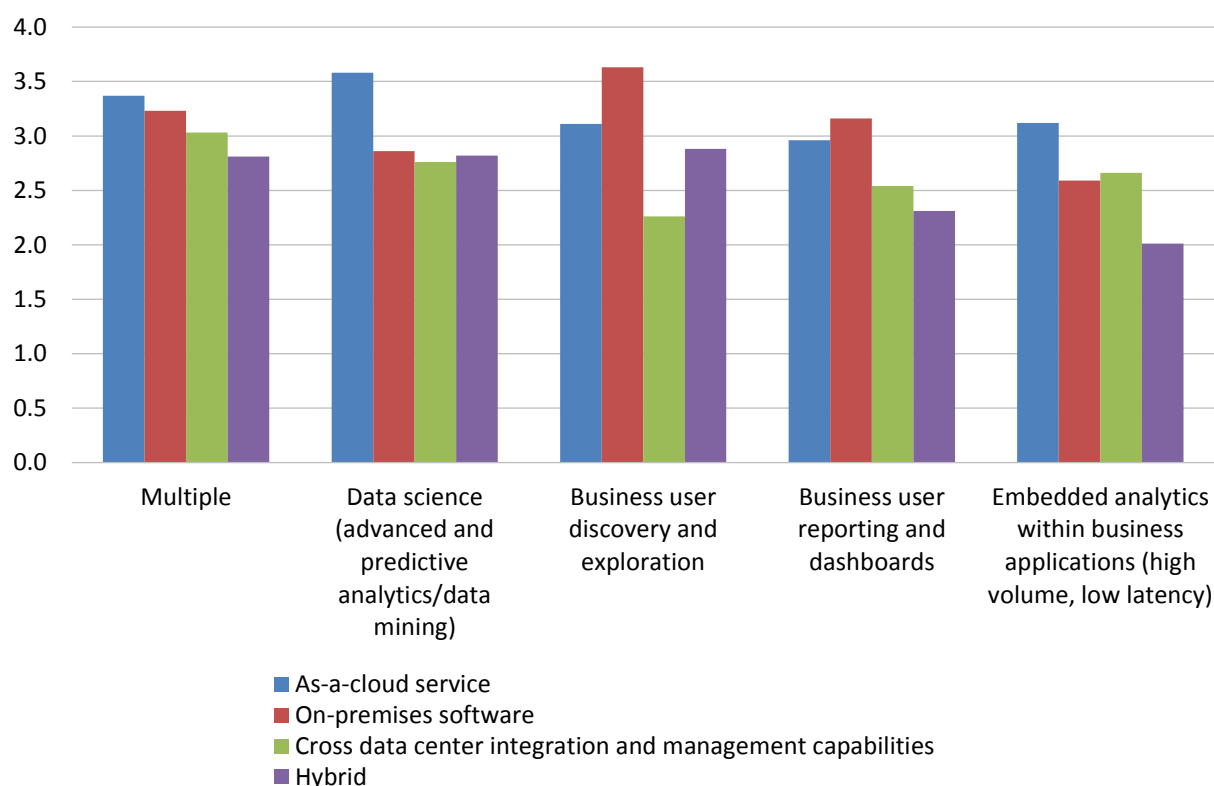


Figure 20 – ADI deployment priorities by top use case

Deployment priorities vary by geography. Asia Pacific and North America show a preference for “as a cloud service” for their ADI platforms. Europe, Middle East, Africa and Latin America respondents indicate “on-premises software” as their top deployment option for ADI platforms (fig. 21). Hybrid deployment options and “cross data center integration and management capabilities” rate higher in organizations in Asia-Pacific geographies compared to other regions.

ADI Deployment Priorities by Geography

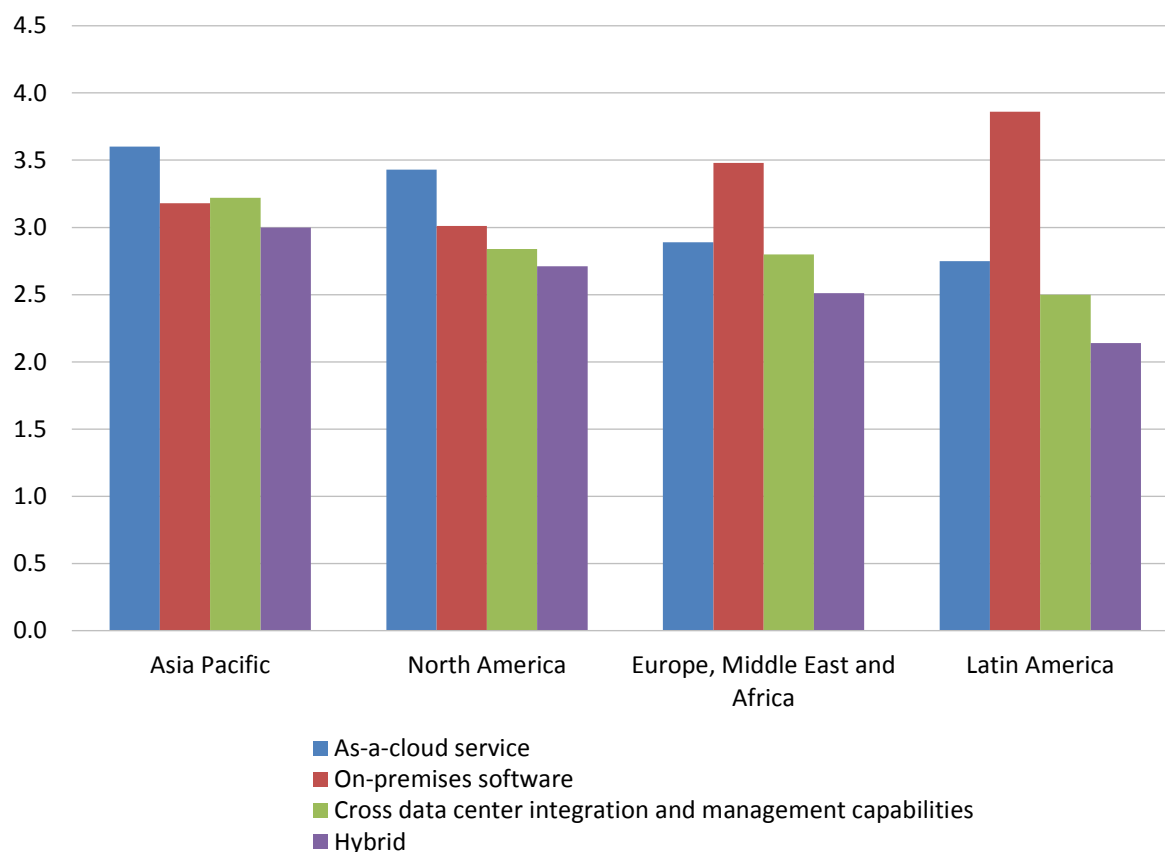


Figure 21 – ADI deployment priorities by geography

ADI deployment priorities vary significantly, depending on the organization function (fig 22). This makes it difficult to align ADI platform priorities. IT, Marketing/Sales, Operations, and Finance respondents rank on-premises deployment as their highest priority, whereas the BICC, Executive Management, and R&D place “as a cloud service” as their deployment priority.

ADI Deployment Priorities by Function

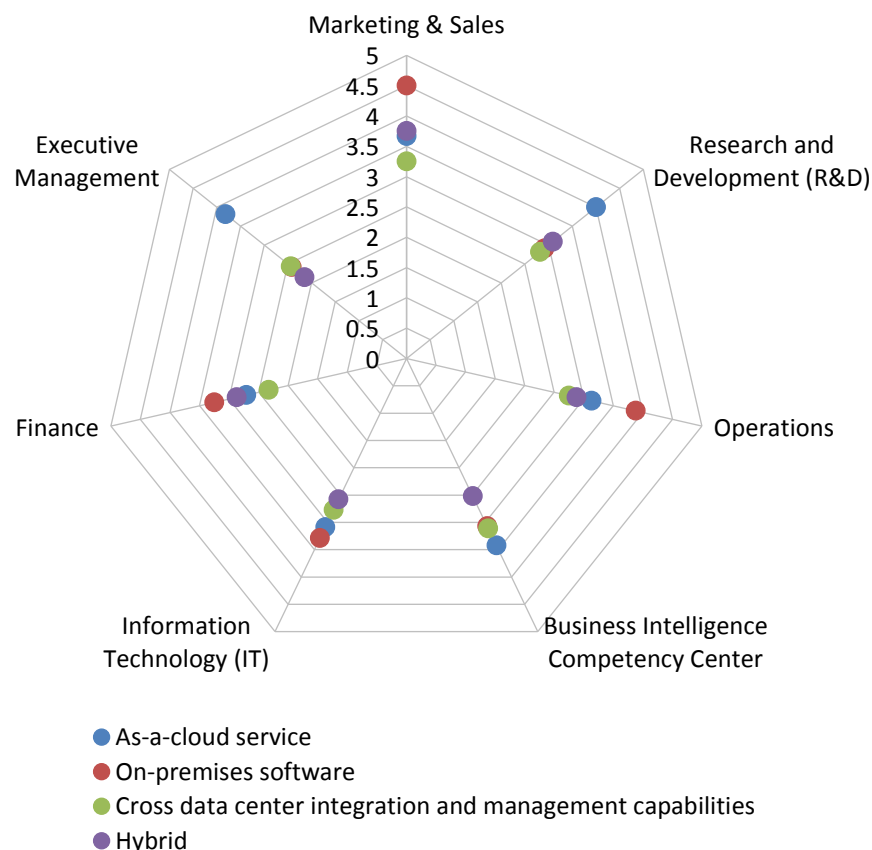


Figure 22 – ADI deployment priorities by function

ADI deployment priorities vary by industry. This year's respondents from the Financial Services, Insurance, Manufacturing, Telecommunications, Healthcare, and Manufacturing rank "on-premises software" as their highest deployment priority (fig. 23). Technology organizations indicate a higher preference for cloud deployments. Education industry participants indicate a slight preference for cloud deployments over on-premises deployments.

ADI Deployment Priorities by Industry

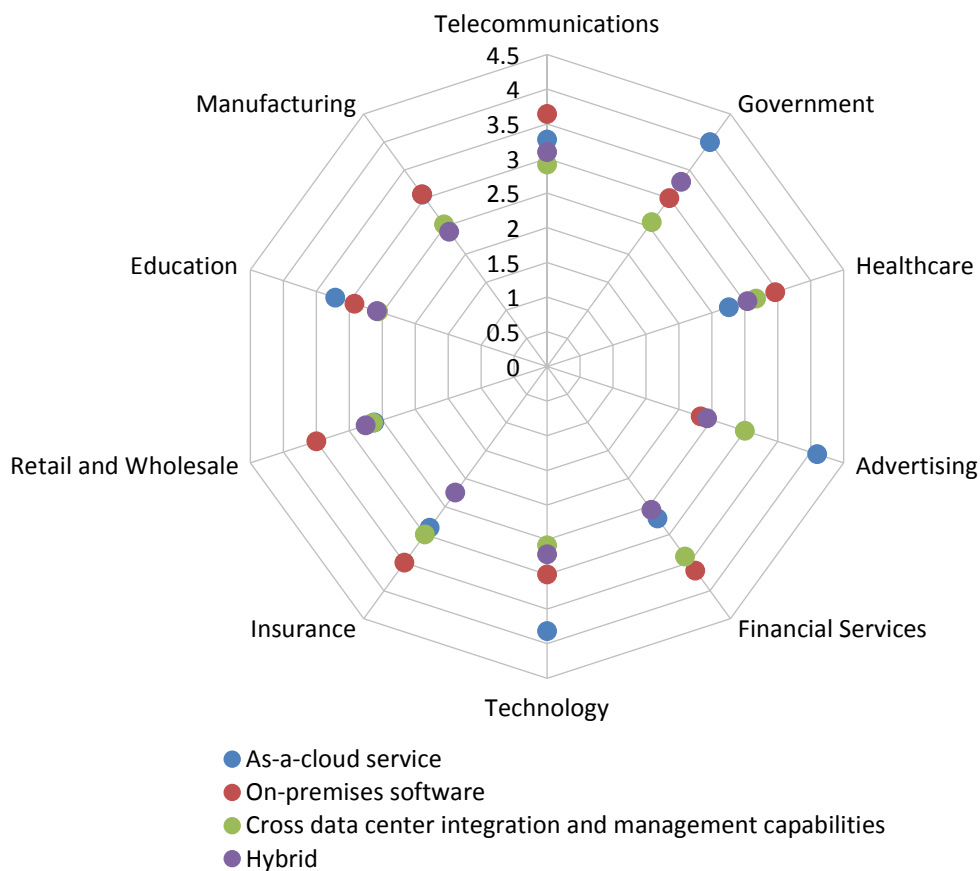


Figure 23 – ADI deployment priorities by industry

Deployment priorities vary by organization size. Organizations with less than 1,000 employees place a priority on cloud services deployment, whereas respondents from larger organizations with more than 1,000 employees place their deployment priority on on-premises deployments. Organizations with more than 5,000 employees prioritize “cross data center integration and management” (fig. 24) higher than “as a cloud service” deployments. It is likely that smaller organizations lack the staff to support ADI platform deployments and, hence, turn to cloud services. “Cross data center integration and management capabilities” increases relative to company size—the larger the company, the higher the priority for cross data center integration. Hybrid deployment priorities are similar across organizations smaller than 5,000 employees; however, its priority increases as the organization size grows beyond 5,000 employees.

ADI Deployment Priorities by Organization Size

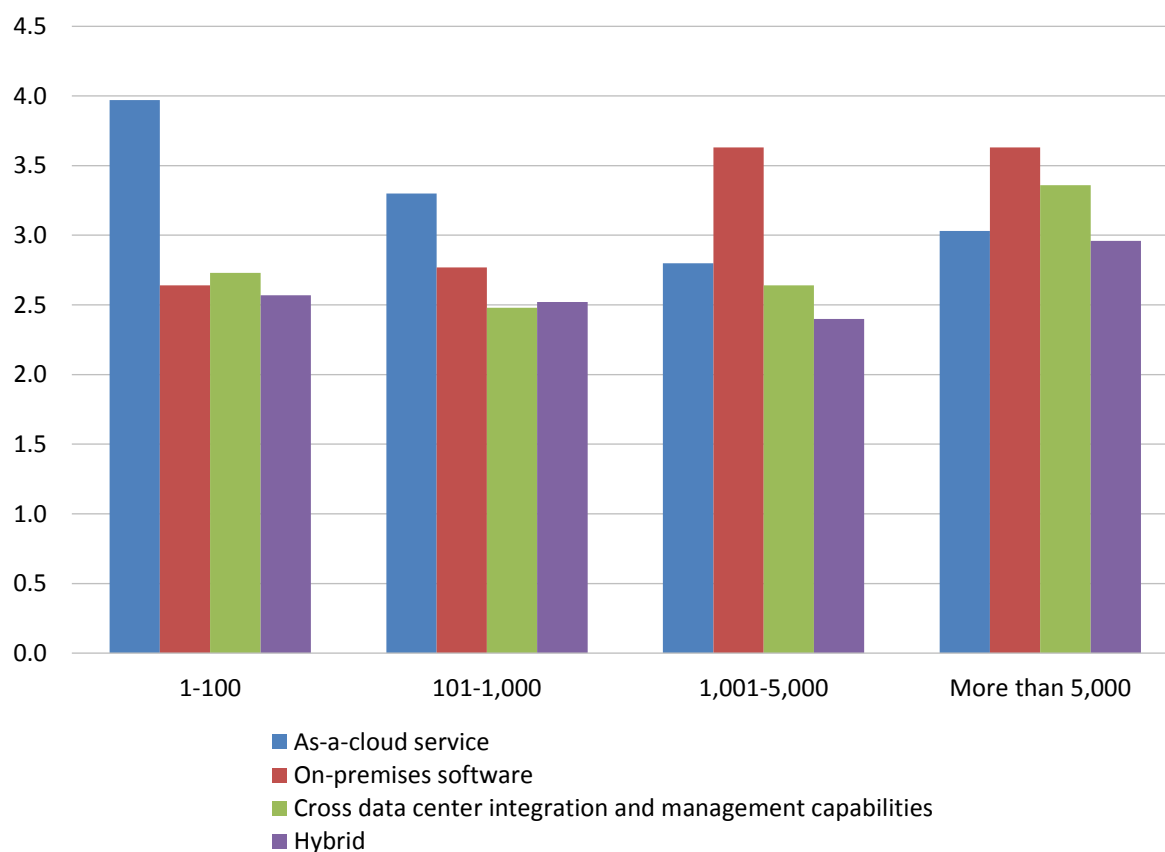


Figure 24 – ADI deployment priorities by organization size

ADI Licensing Preferences

“Concurrent use” and “subscription-based” models are the top licensing preferences among respondents, with 43 percent and 40 percent preference respectively (fig. 25). “Data volume” surpasses “user” licensing models as a preference this year. Respondents indicate “open source” as their lowest preference for ADI licensing. Interestingly there are many cloud services that offer/use open source ADI platforms.

ADI Licensing Preferences

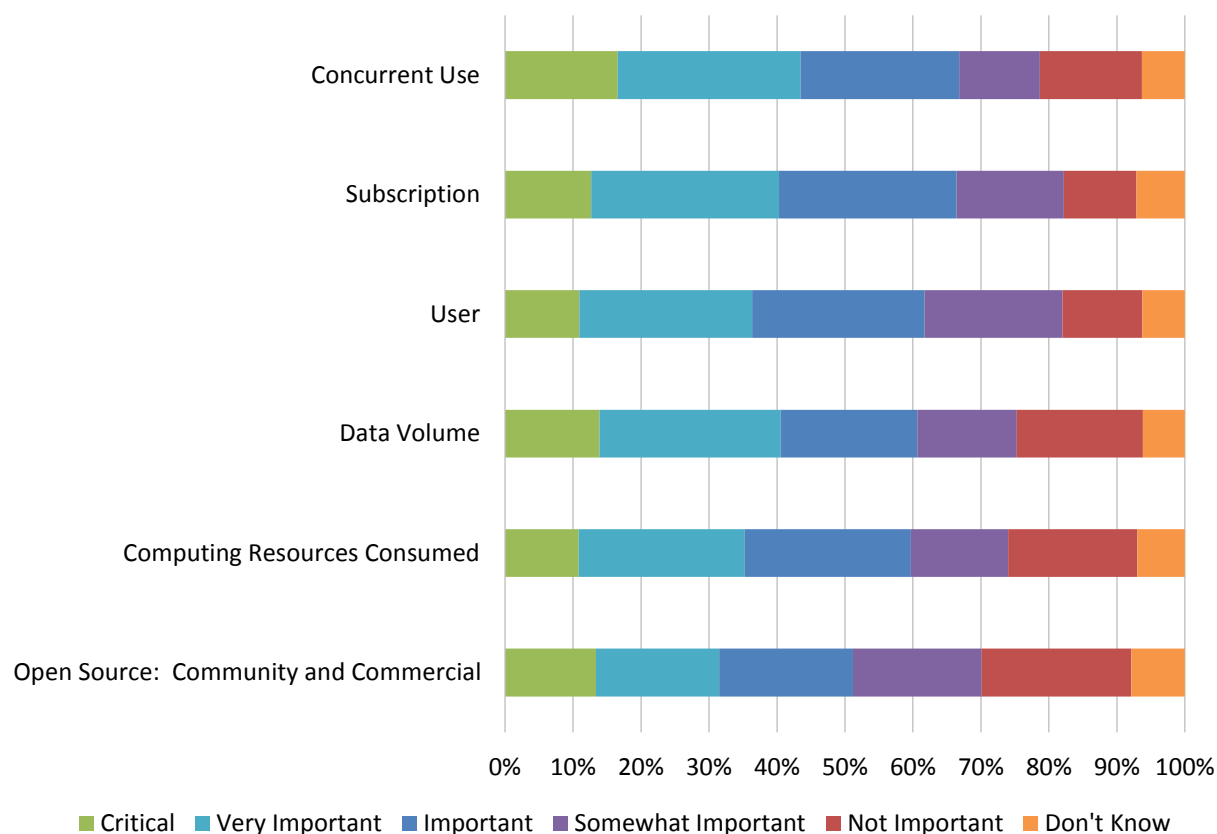


Figure 25 – ADI licensing preferences

ADI platform licensing preferences are evolving, shifting from user based to subscription and volume/resources consumed. The user licensing preference continues to decline slightly year over year (fig. 26). Respondent preferences increase for subscription, computing resources consumed, and data volume licensing options. This matches the increased preferences for cloud deployments (fig 18), which are typically licensed on a subscription or capacity model.

ADI Licensing Preferences 2017-2019

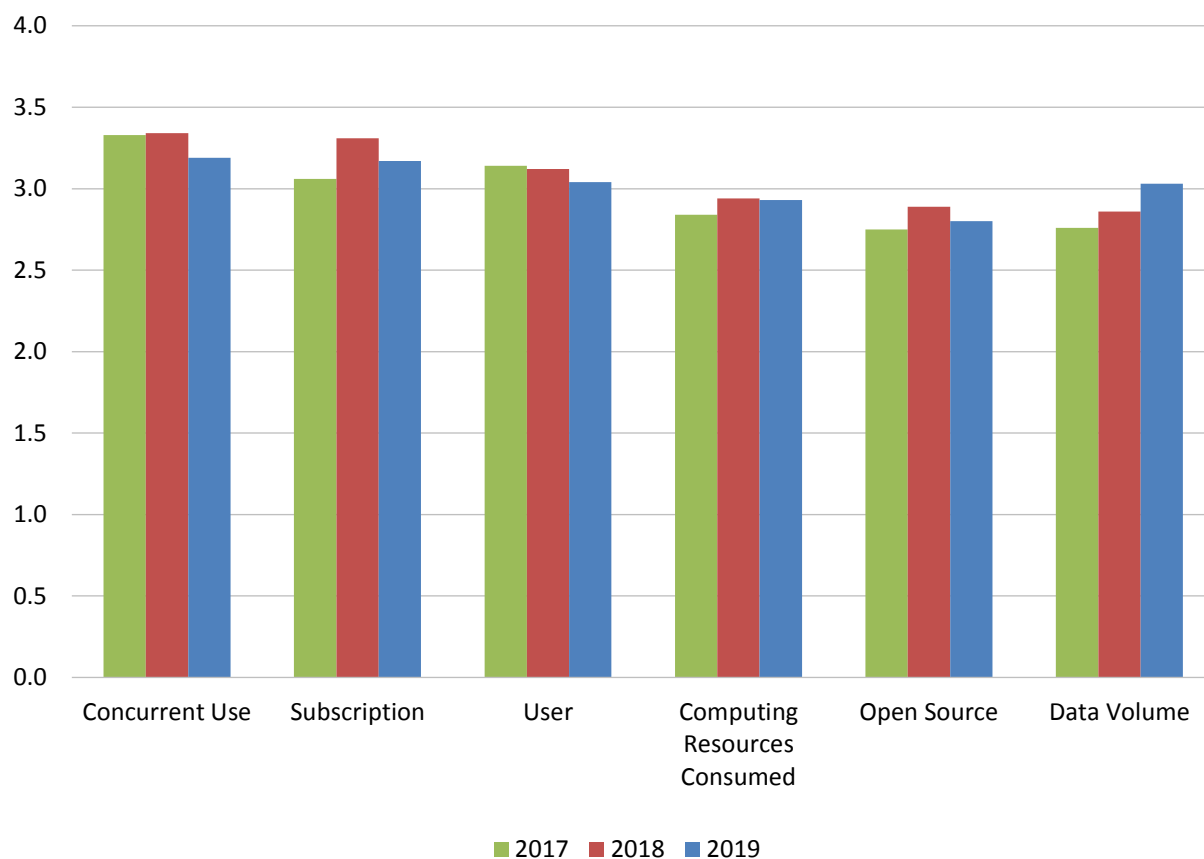


Figure 26 – ADI licensing preferences 2017-2019

When we look at the ADI platform responses on licensing priorities by top use case, we can see a relatively high variability in licensing model preferences (fig. 27). When data science and “business user discovery and exploration” use cases are top priorities, the preference is for a computing resources consumed license model which would provide the users with flexibility of scale. Open source licensing models are a top preference for both the data science and embedded analytics use cases.

ADI Licensing Priorities by Top Use Case

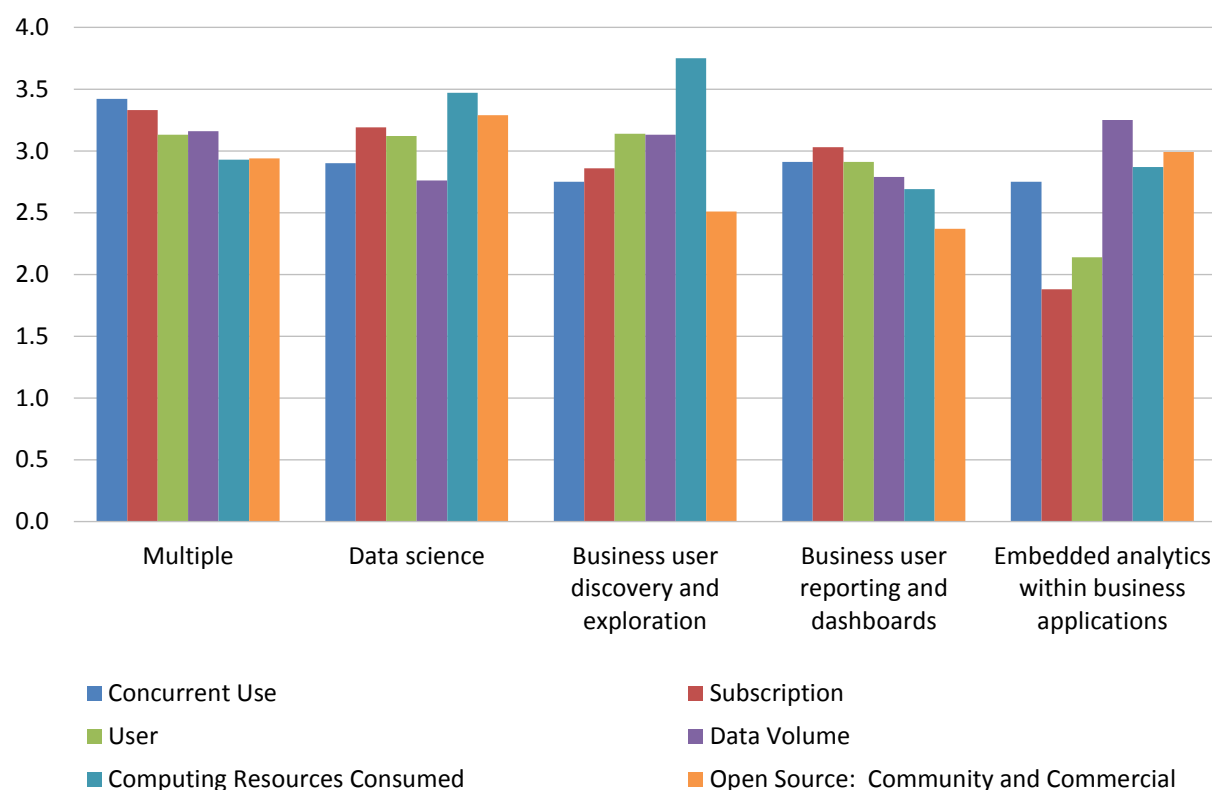


Figure 27 – ADI licensing priorities by top use case

Licensing preferences vary by geographies (fig. 28). Asia-Pacific respondents have a higher preference for volume/consumption models and show higher preference for open source licensing models compared to other geographies. Latin America shows a highest preference for user-based licensing, which matches their higher preference for on-premises deployments (fig 21).

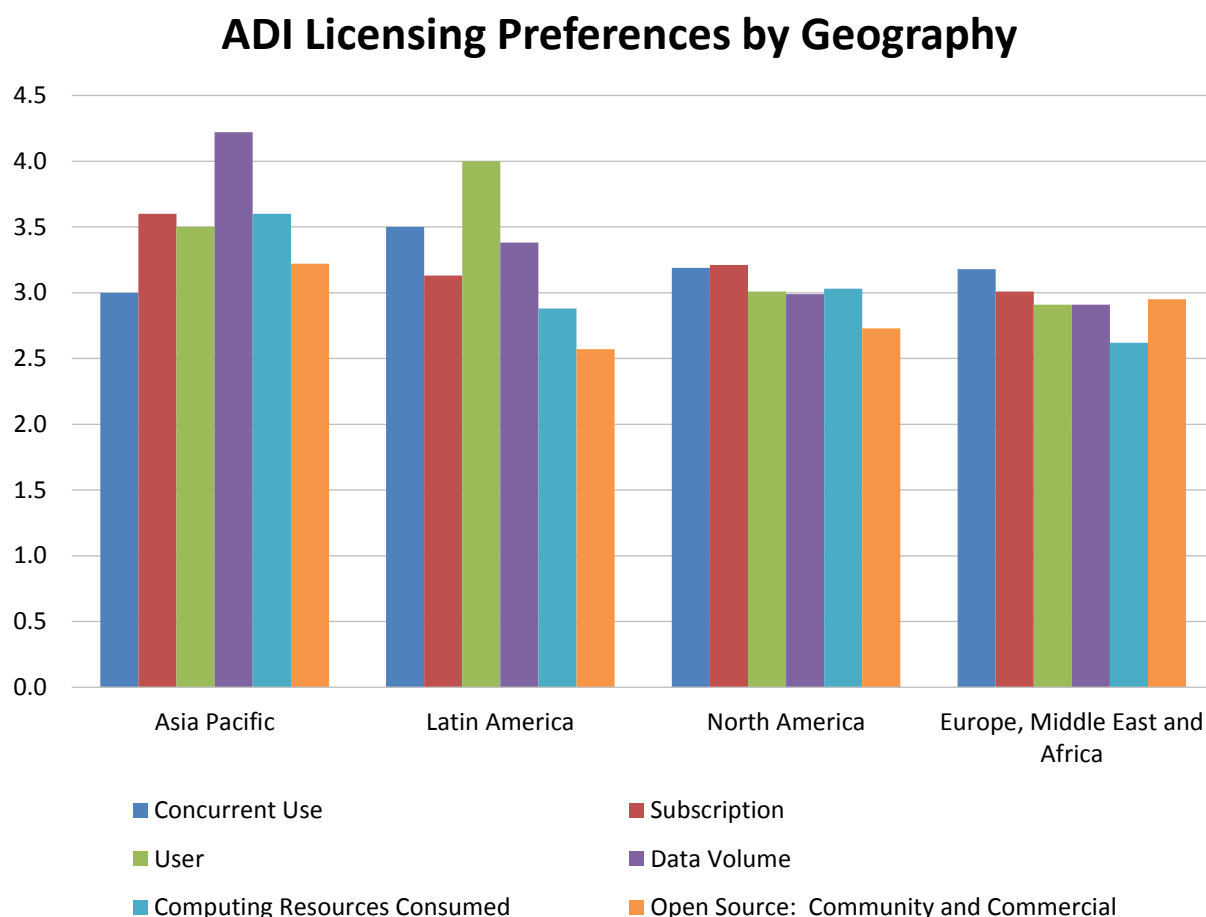


Figure 28 – ADI licensing preferences by geography

Licensing preferences vary by organizational functions. Respondents from the Finance, Marketing/Sales, and Executive Management functions give highest preference for “subscription” licensing models (fig. 29). Not surprisingly, open source is the preferred licensing model for R&D functions, given their need to often modify the source code of community-developed software and the higher level of development, customization, and integration with which they might be involved. Respondents from the Operations functions prefer “concurrent use” and user-based licenses for their ADI platforms.

ADI Licensing Preferences by Function

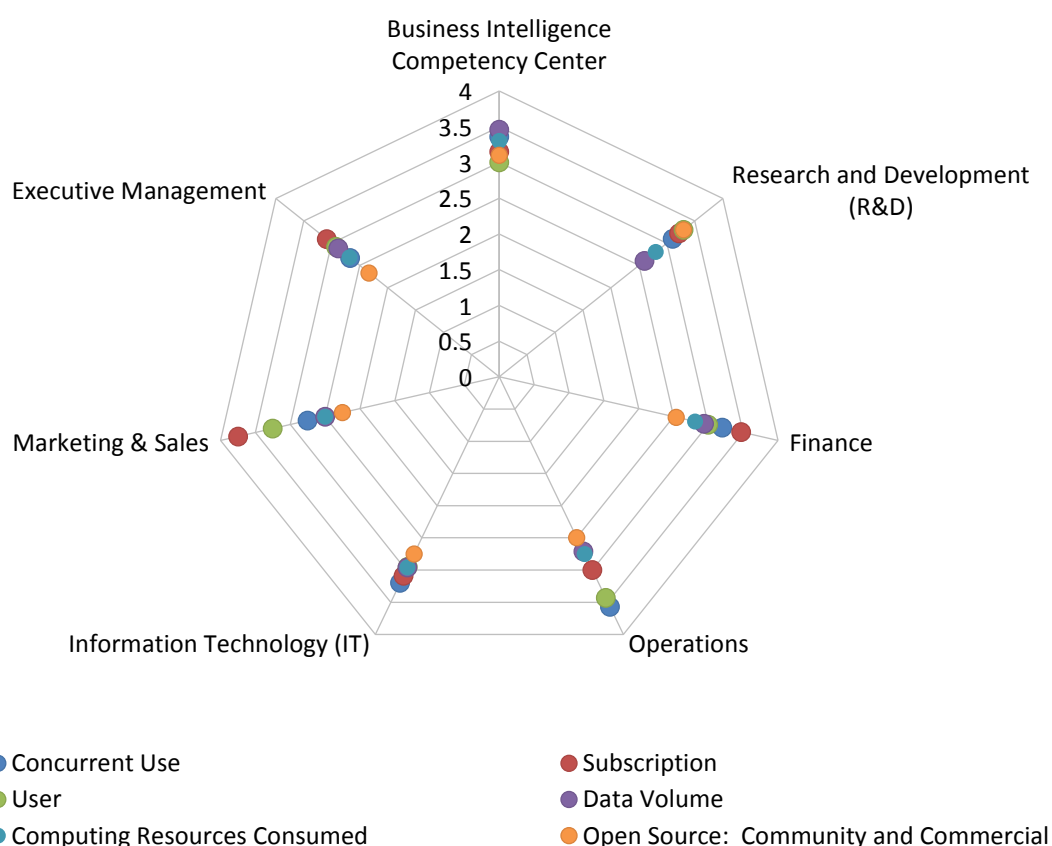


Figure 29 – ADI licensing preferences by function

ADI licensing preferences vary by industry. Concurrent-use and subscription-based licensing models are a relatively high preference across all industries (fig. 30). Telecommunication respondents have the highest preference for concurrent-use licensing. Data volume is the highest preference licensing model for Healthcare. Open source licensing models are the top licensing preference for Education respondents.

ADI Licensing Preferences by Industry

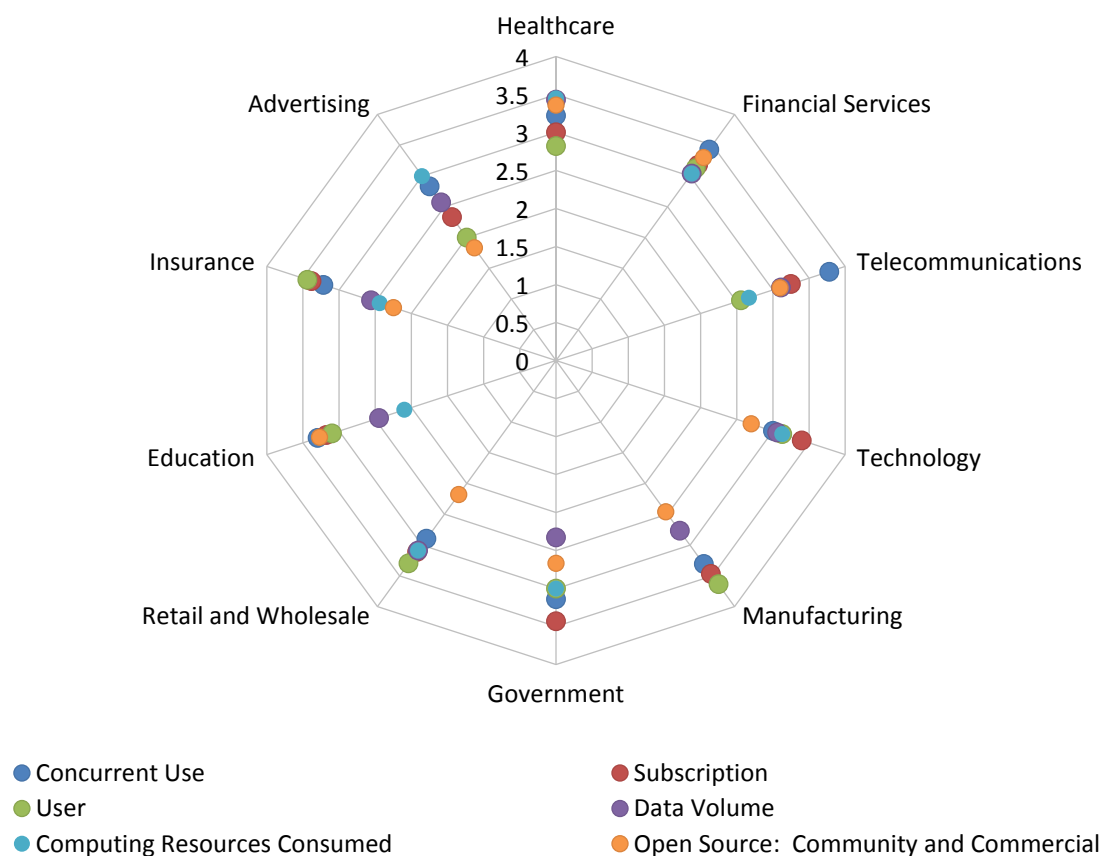


Figure 30 – ADI licensing preferences by industry

Licensing option preferences segmented by company size do not vary as much as other factors (i.e., geography, organizational function, industry). Not surprisingly, smaller organizations that show a preference for cloud deployment rank their licensing top priority in subscription-based models (fig. 31). Larger organizations prefer almost every licensing model over user-based licensing for their ADI platforms. Concurrent-use, data-volume, and subscription licensing gives larger organizations more flexibility in administering their deployment since they don't have to identify and manage/administer large numbers of users as they do with a user-based licensing model. Of interest is the rising preference for open source licensing for ADI platforms in the largest of organizations.

ADI Licensing Preferences by Organization Size

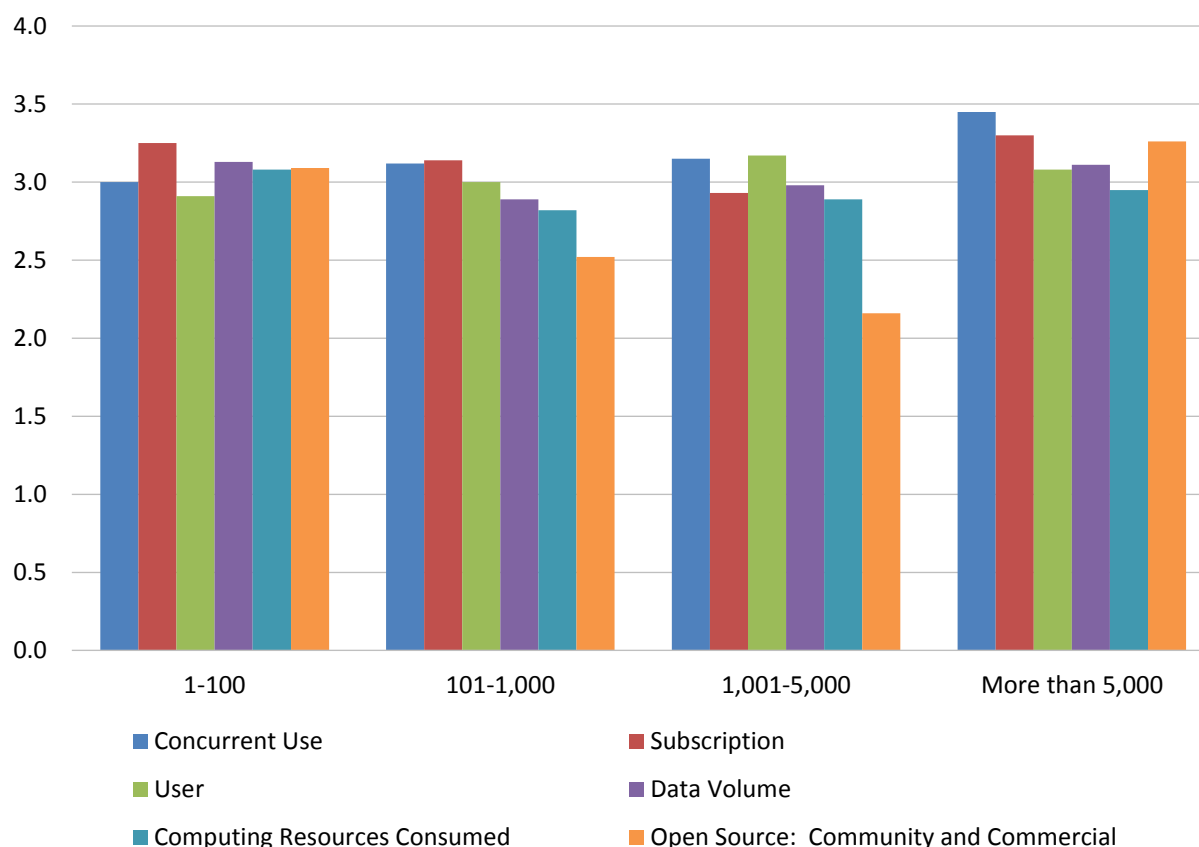


Figure 31 – ADI licensing preferences by organization size

Analytical Data Infrastructure Development and Deployment Features

Respondents indicate this year that performance and scalability are top priorities for their ADI platform requirements (fig. 32). We track eight ADI development and deployment features. Not surprisingly, “scale up and scale out” capabilities lead as the highest priority for ADI development features. Security, also a top ADI overall selection priority is reflected in the ADI development priority of data life cycle management capabilities. ADI platforms require investments in data integration, cleansing, and analytic applications development; hence, support for programming languages and application development tools rank increasingly in priority. Like last year, “pre-built data models” is the lowest priority development capability/priority. Pre-built data models have a spotted history in analytic applications ,given the difficulty of modifying and adapting them to new/different requirements than what they were originally designed for. While multi-tenancy support is often an ADI capability used to promote cloud offerings, its priority is comparatively low.

ADI Development and Deployment Features

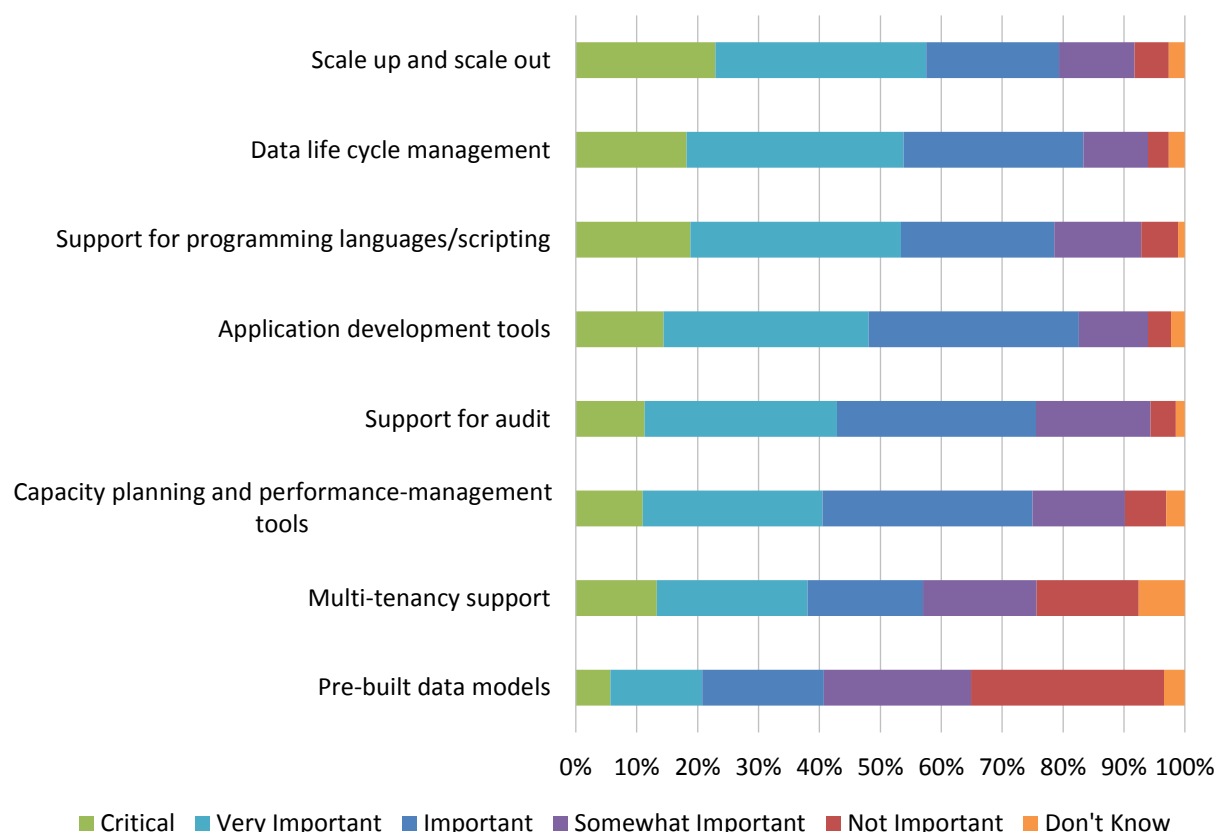


Figure 32 – ADI development and deployment features

Year-over-year responses show little change in the priority for ADI development and deployment features. Last year, “scale up and scale out” was the top development and deployment feature priority (fig. 33). Pre-built data models are the lowest priority for all years.

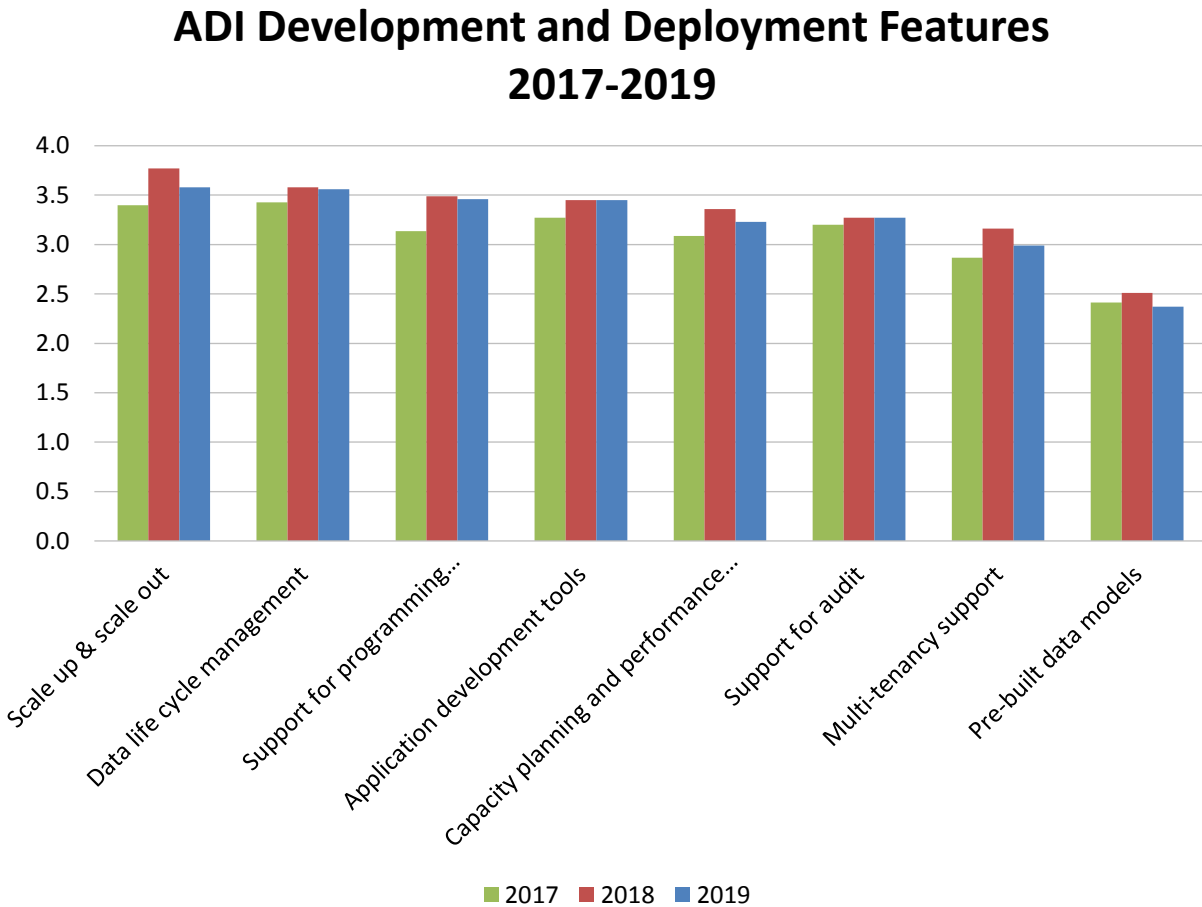


Figure 33 – ADI development and deployment features 2017-2019

“Scale up and scale out” is the top priority for ADI platforms regardless of use case. When we look at the results by specific use cases, we see a few trends of note (fig. 34). Embedded analytics can be taxing and challenging from a scale perspective, depending on the workload of the embedding business application. The ability to scale up and scale out the embedded analytics is, therefore, a top priority for those use cases. Not surprising is the embedded use-case priority placed on support for programming languages and run times supported by the ADI platform.

ADI Development and Deployment Features by Top Use Case

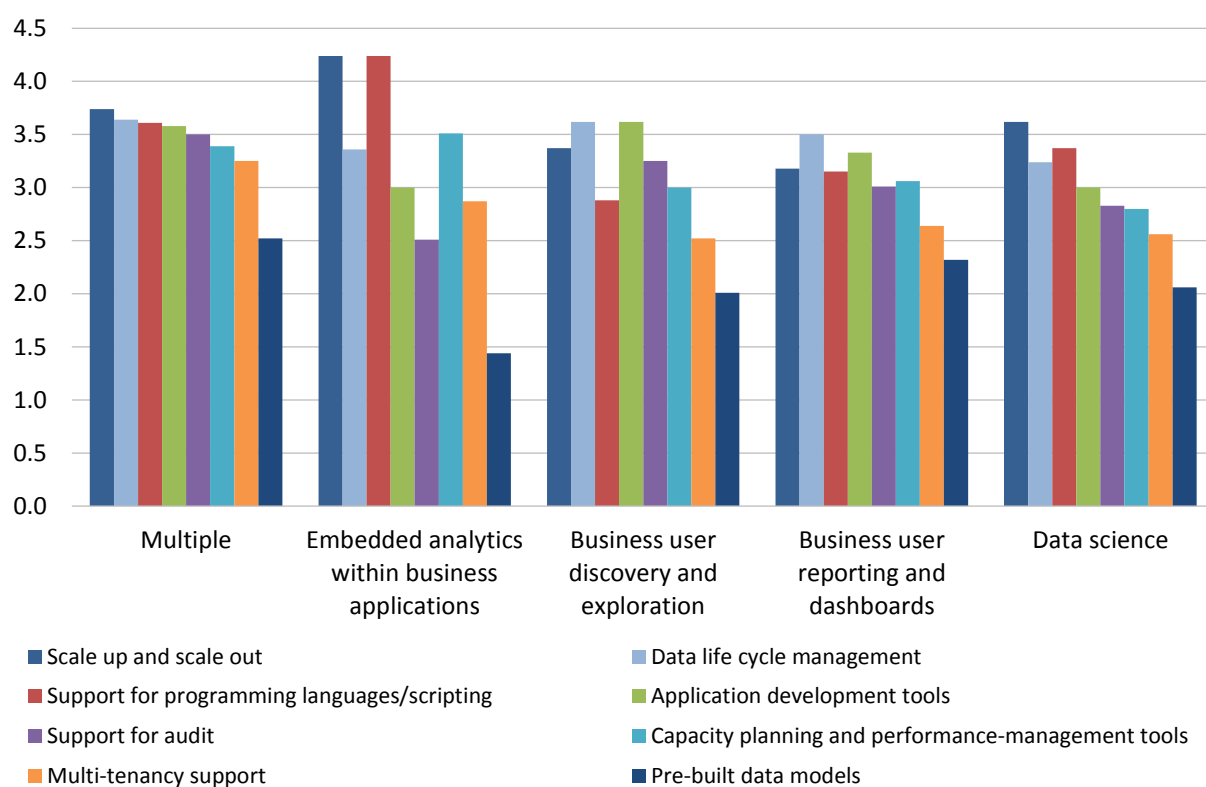


Figure 34 – ADI development and deployment features by top use case

The ADI Development and Deployment features vary quite a bit when we look at the responses by geography (fig. 35). The top priority for Asia Pacific and Latin America is “capacity planning and performance-management tools,” however, that is the lowest priority for EMEA respondents. Pre-built data models have higher appeal/preference in Asia Pacific than other geographies. For North America, “scale up and scale out” and “data life cycle management” lead respondents’ feature preferences.

ADI Development and Deployment Features by Geography

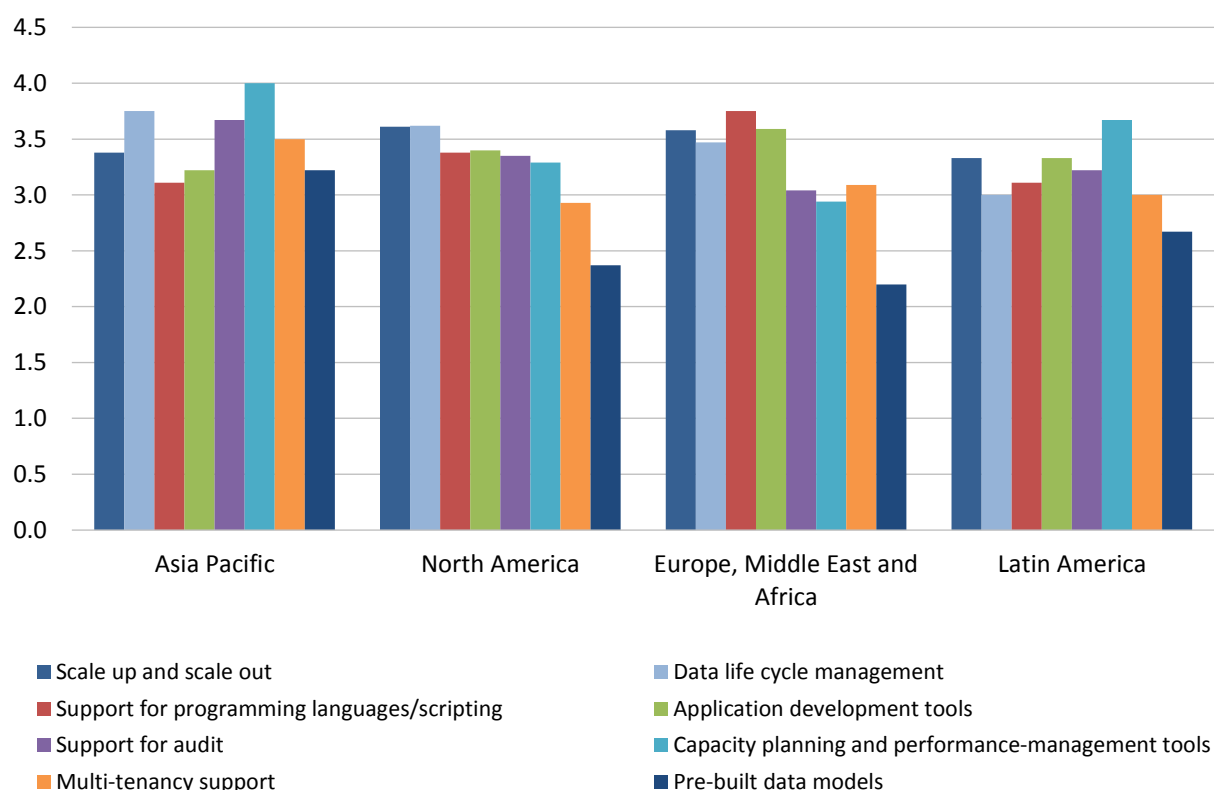


Figure 35 – ADI development and deployment features by geography

ADI development and deployment features are typically consistent across organizational functions (fig. 36). The BICC, R&D, and Operations place the highest priority on “scale up and scale out” features. The priority of “data life cycle management” tools and “support for audit” is fairly high across organizational functions, indicating a possible increase in understanding for the need for these features, especially when it comes to security controls. Interestingly, Finance respondents show a high interest in “application development tools” this year.

ADI Development and Deployment Features by Function

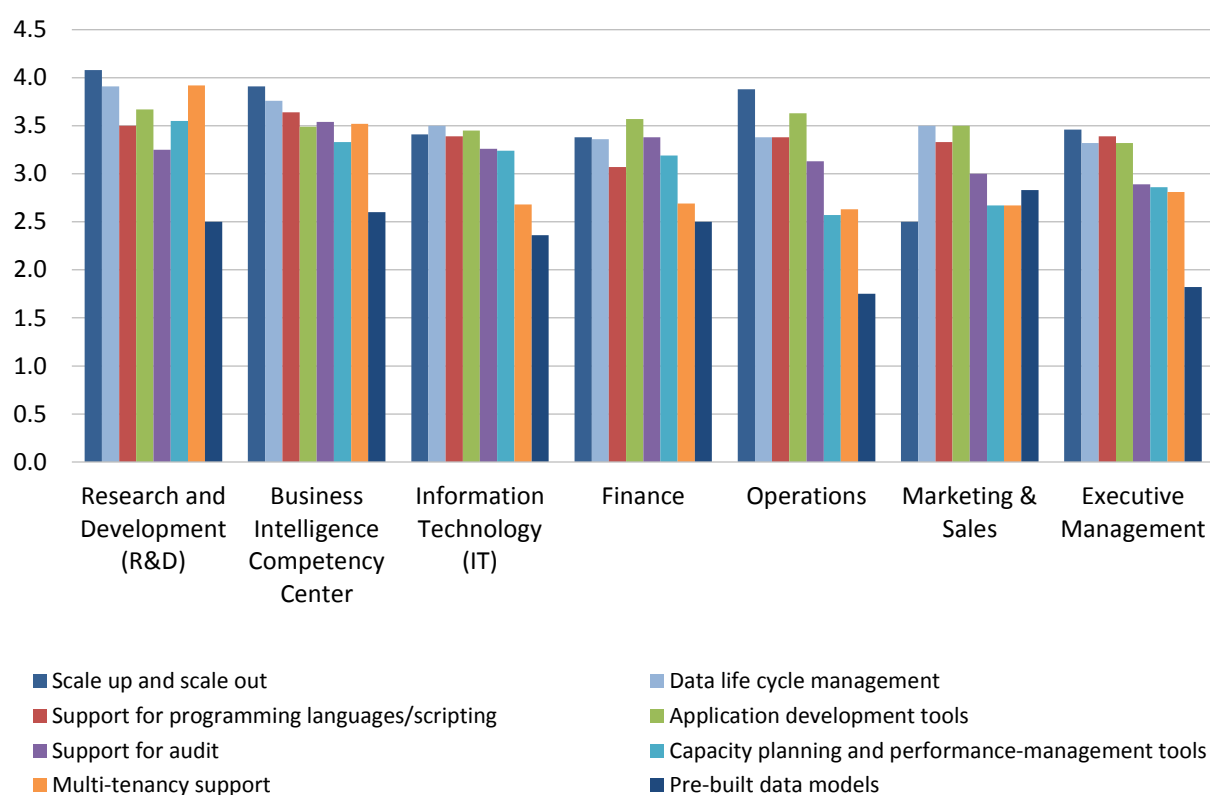


Figure 36 – ADI development and deployment features by function

ADI priorities for development and deployment features vary along the industry dimension. In general, “scale up and scale out” capability is a very high or the highest ADI priority across all industries (fig. 37). “Data life cycle management” is a high priority for Government, Telecommunication, and Education. Telecommunications respondents place their highest priority on multi-tenancy support (which is consistent with last year’s results). Advertising respondents place their highest priorities on support for programming languages. Across several years, there is consistently low priority for pre-built data models in some industries, showing above-average interest (e.g., Financial services, Government, Retail/Wholesale, and Insurance).

ADI Development and Deployment Features by Industry

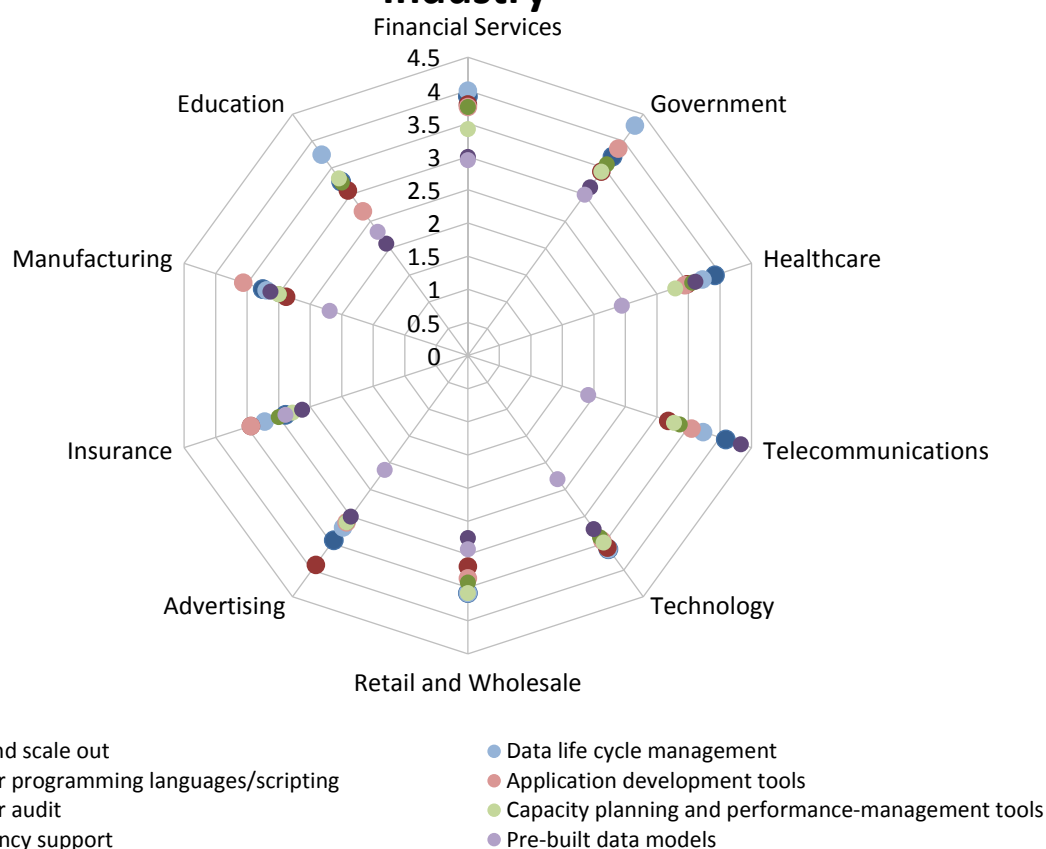


Figure 37 – ADI development and deployment features by industry

The priorities for ADI development and deployment features segmented by organization size shows relatively low variability. In general, larger organizations place a higher priority on development and deployment features compared to smaller organizations. Not surprisingly, larger organizations with 1,001 or more employees place high priority on “data life cycle management” and “scale up and scale out” capabilities. Across most companies, regardless of size, “scale up and scale out” is the highest priority and “pre-built data models” are the lowest priority (fig. 38).

ADI Development and Deployment Features by Organization Size

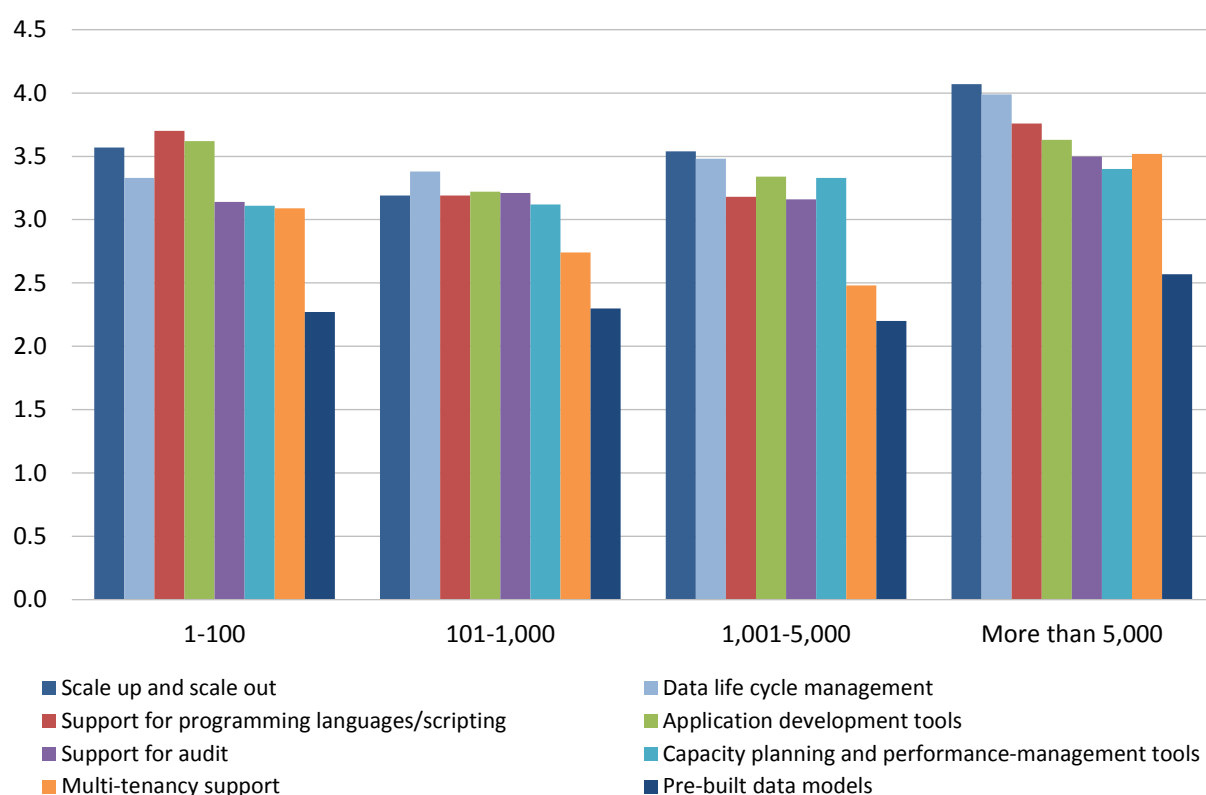


Figure 38 – ADI development and deployment features by organization size

Analytical Data Infrastructure Data Types

More than 80 percent of respondents have analytic workloads and workflows based on “transactional data” sources (i.e., from business applications). The next highest data type priority for ADI platforms and associated analytic use case workflows are metadata and Excel/CSV data (fig. 39). Given the role of metadata in data lineage and impact analysis, the high priority for metadata echoes respondents’ priority for “data life cycle management” capabilities in the ADI development and deployment sections (see fig. 32, p. 47). In general use, images and video data type support has, historically, been a lower priority for ADI applications/use cases, and this set of data types responses shows a similar priority for ADI platforms today. It is worth noting that “machine and events / log data” are important data types for a large number of analytic use-case workflows/workloads.

Data Type Priorities for ADI

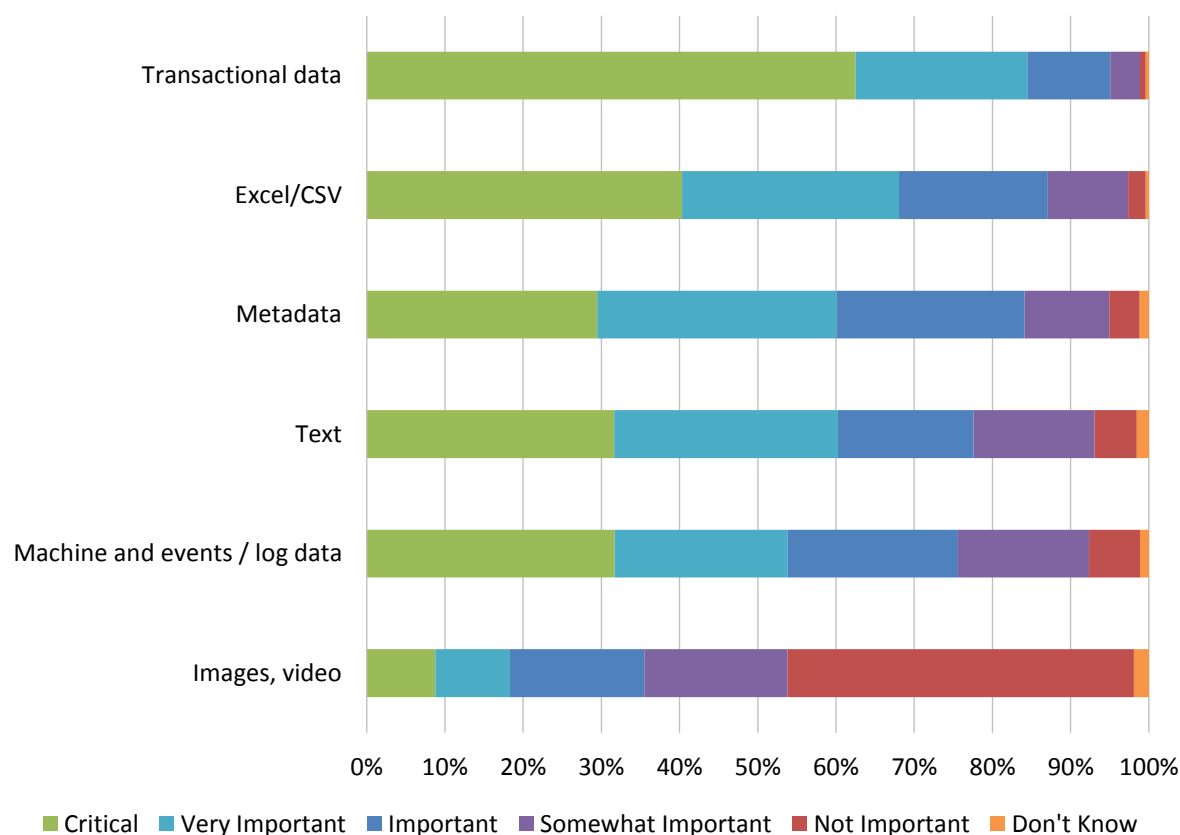


Figure 39 – Data type priorities for ADI

The relative priority of data types supported by an ADI platform changes little between 2018 and 2019. It is interesting to note the importance of metadata is equal to the importance of an ADI's ability to support Excel data types (e.g., import / export / manage Excel data) (fig. 40). Metadata is needed as organizations scale up/out their implementation and focus more on data life cycle management. Text and “machine and events / log data” also rise in importance compared to our 2018 report.

Data Type Priorities for ADI 2017-2019

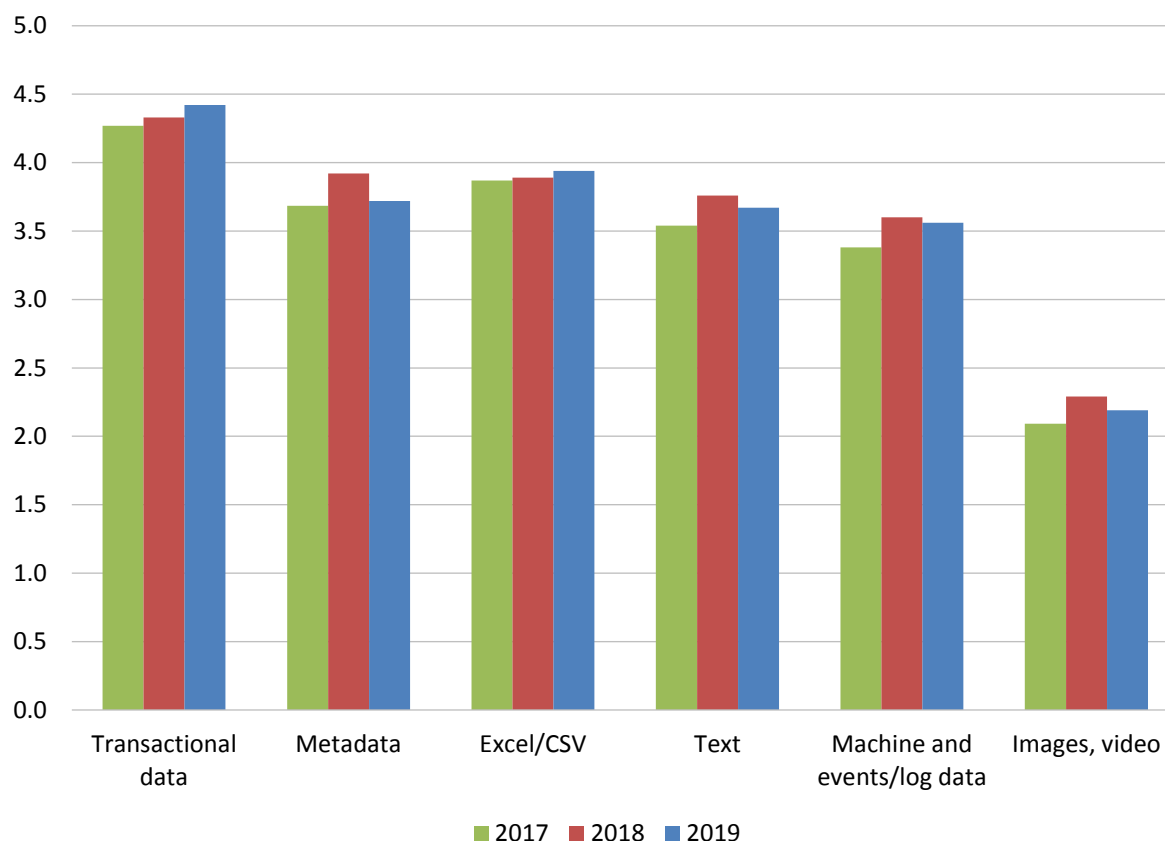


Figure 40 – Data type priorities for ADI 2017-2019

The top data type priority for all ADI use cases is transaction data. The lowest priority data type across all ADI use cases is “images and video” (fig. 41). Embedded analytics use cases place a higher priority on text data types compared to other use cases.

Data Type Priorities for ADI by Top Use Case

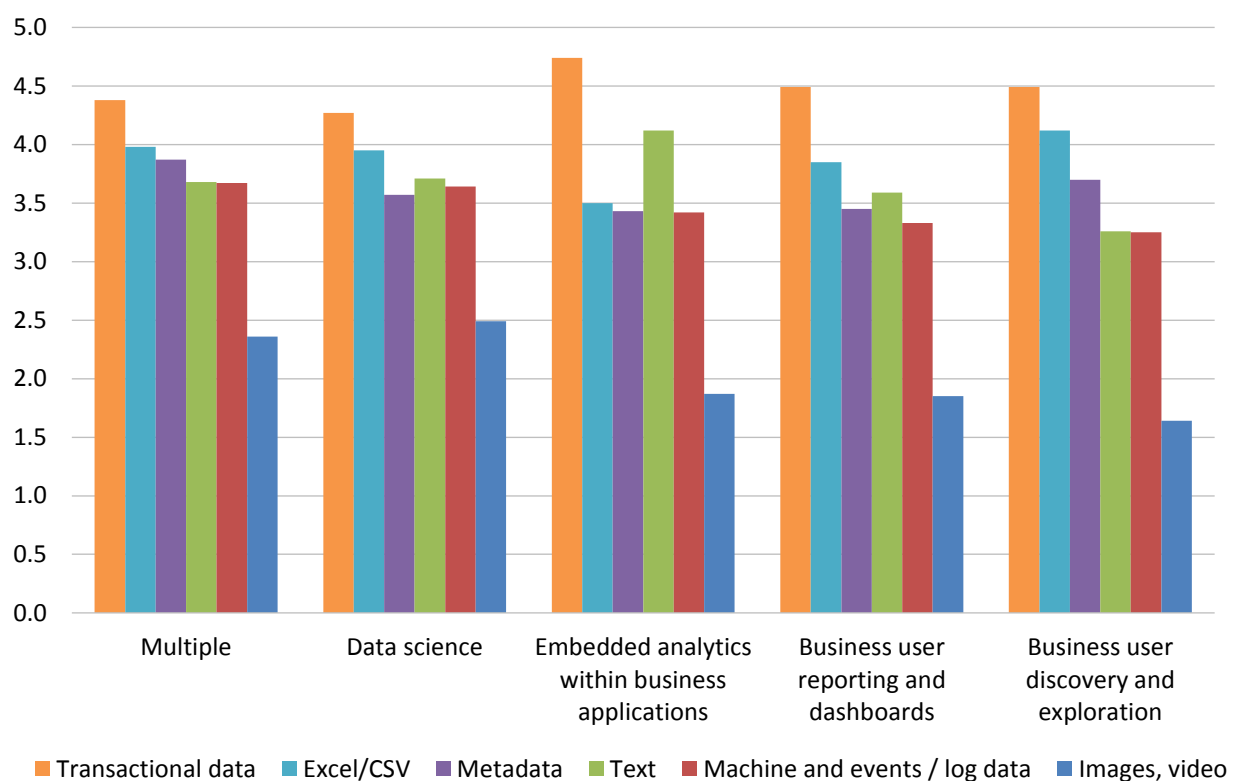


Figure 41 – Data type priorities for ADI by top use case

The relative priority of data types is similar across all regions. Transactional data is the highest priority data type across all geographies (fig. 42). Metadata support is slightly higher in priority for North America than other regions. All geographies rank “images and video” as a low priority data type to access / manage / analyze as a part of the ADI platform functionality.

Data Type Priorities for ADI by Geography

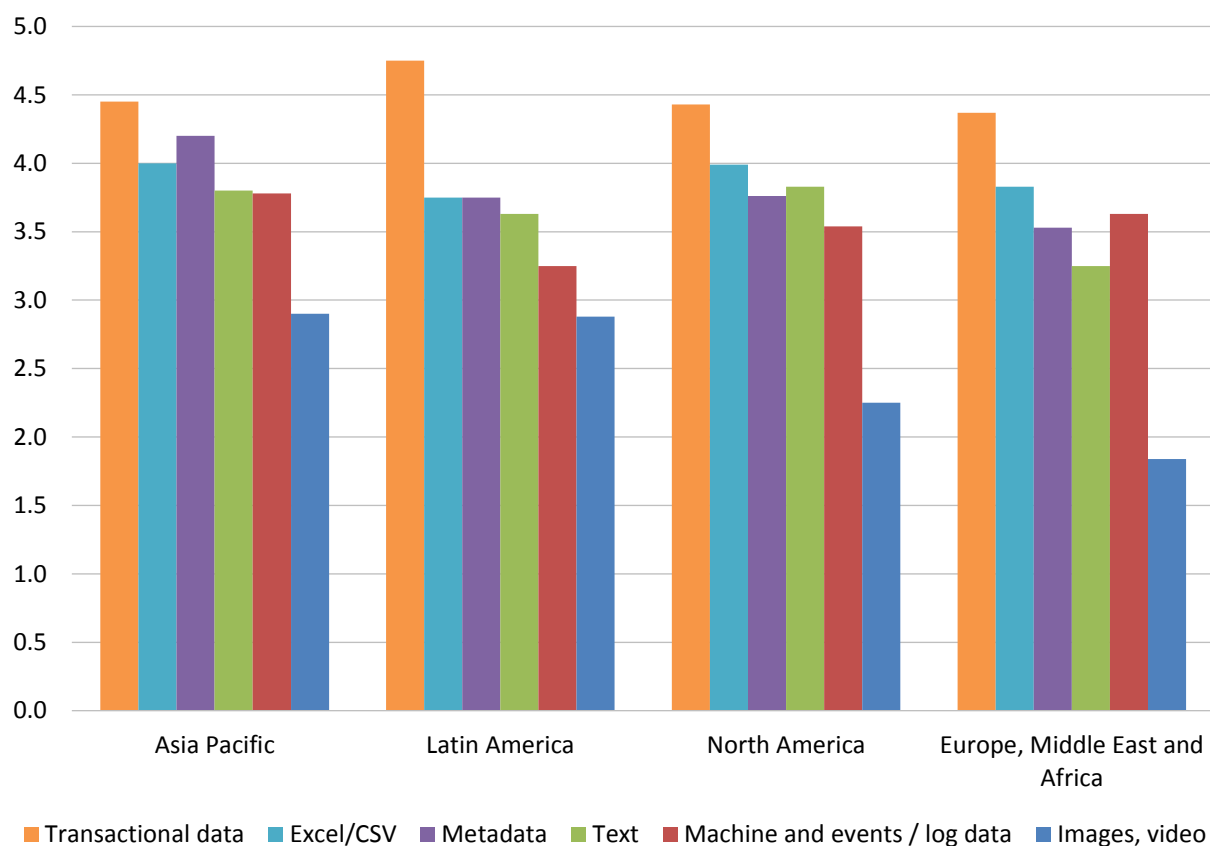


Figure 42 – Data type priorities for ADI by geography

The priority of ADI data type capabilities varies by respondents' function in their organization. Marketing/Sales respondents rate Excel/CSV data type capabilities as a higher priority than other data types including transactional data (fig. 43). The BICC shows the highest relative priority for metadata. This is probably due to the data life cycle management techniques that involve the BICC. Compare this to Marketing and Sales respondents, who place metadata at the lowest level of priority. It appears that there are still a lot of opportunities for educating different business functions on the role and importance of different data types especially meta-data and its role in data life cycle management and security/governance requirements.

Data Type Priorities for ADI by Function

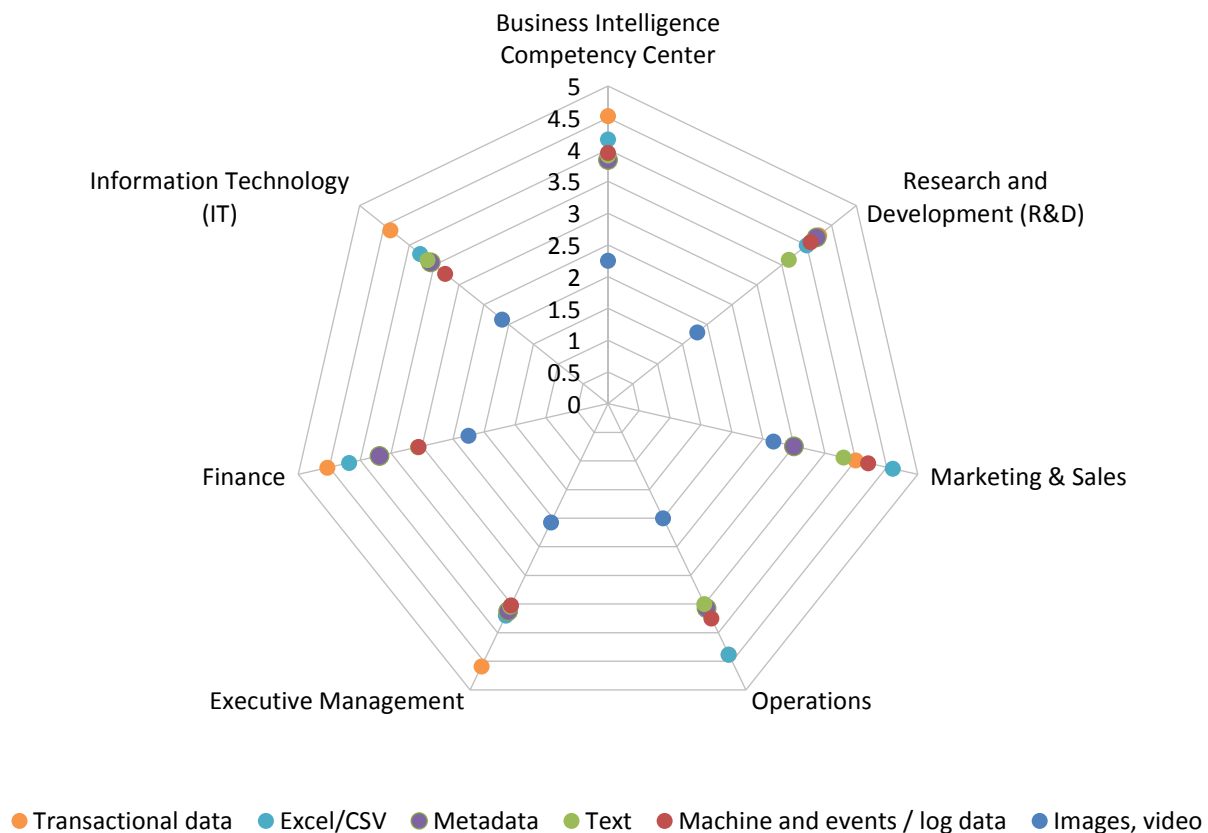


Figure 43 – Data type priorities for ADI by function

In general, the relative priorities for ADI data type capabilities by industry are like other dimensions we study (i.e., function, geographies, organization size), noting that transaction data is the highest priority and images/video are among the lowest priorities (fig. 44). On the other hand, we see a large variety in priority for metadata, images/video, “machine and events / log data,” and text across industries. The variety in data types by industry reflects the variation of different analytical workflows and workloads by industry.

Data Type Priorities for ADI by Industry

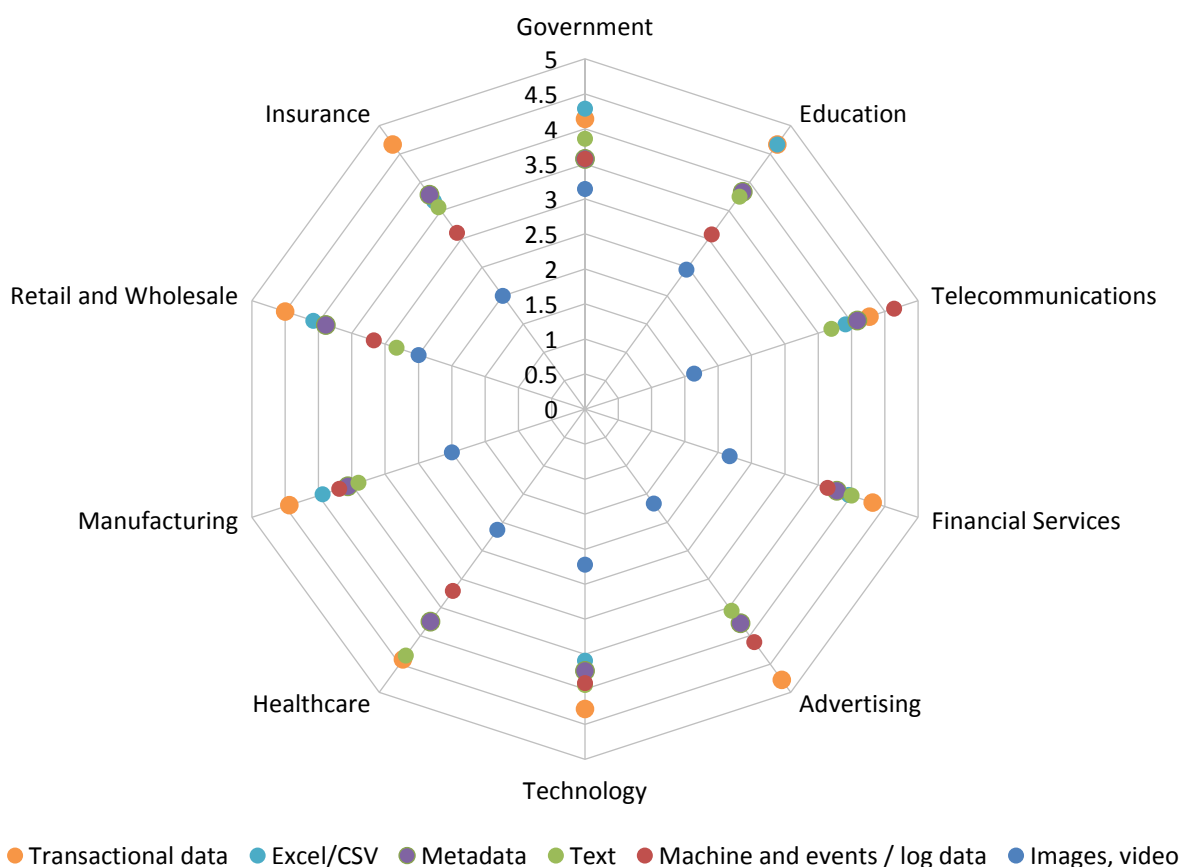


Figure 44 – Data type priorities for ADI by industry

Organizations of all sizes indicate a priority for use of transactional data types, followed by Excel/CSV and metadata (fig. 45) as requirements for their ADI platform.

Images/video is a relatively higher data type priority for smaller organizations and is a lower priority for larger organizations. Larger organizations place a higher priority on metadata. This mirrors their higher priority for data life cycle management capabilities (among development and deployment feature priorities) which require the use of metadata. Somewhat surprisingly, smaller organizations match larger organizations in their priority for the ability to import/analyze text and “machine and events / log data.”

Data Type Priorities for ADI by Organization Size

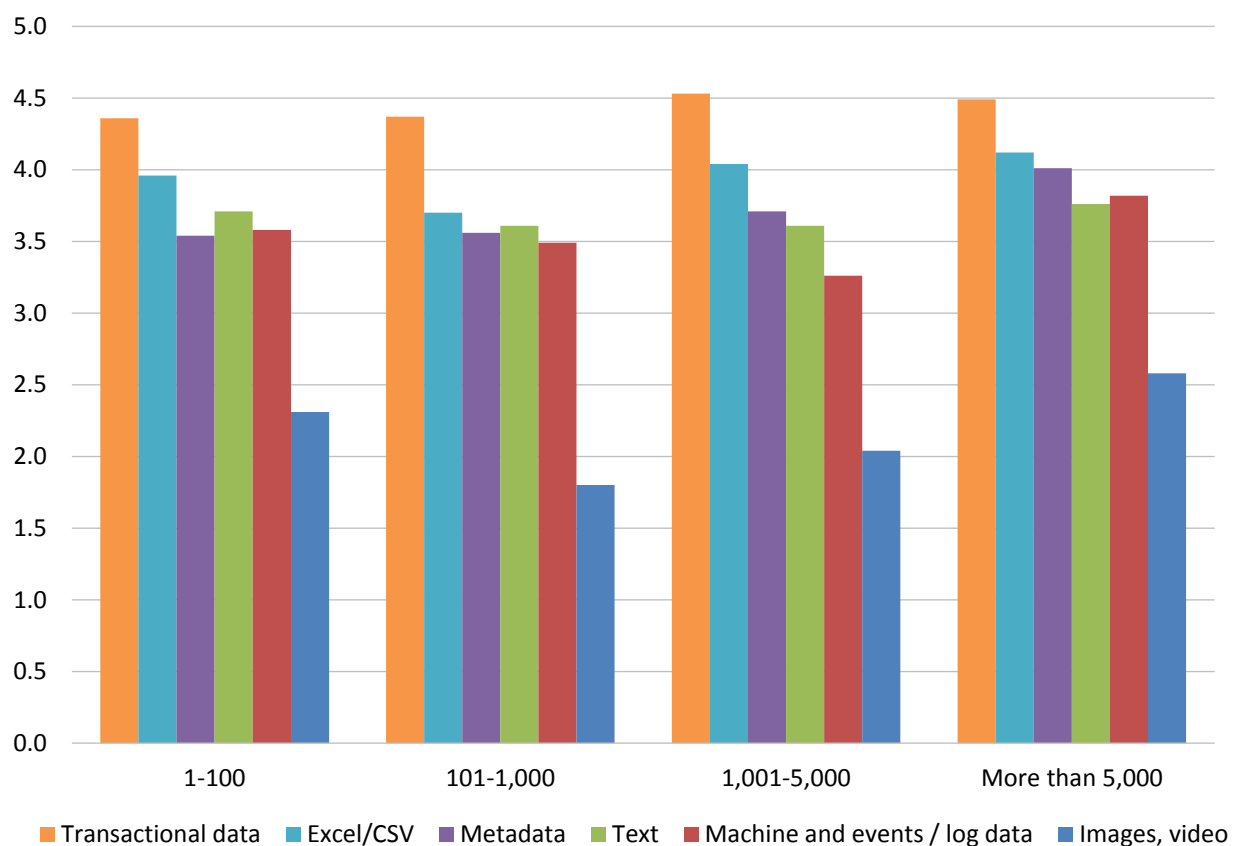


Figure 45 – Data type priorities for ADI by organization size

Analytical Data Infrastructure Data Preparation and Loading

Earlier in this report, we note increasing growth in preferences for ADI platforms to support or provide “cross data center integration and management” capabilities. Traditional bulk load data preparation/transformation and loading capabilities such as ELT / ETL (Extract, Transform, Load) remain the highest priority associated with data preparation and loading for ADI platforms (fig. 46). Inserts/updates/upserts, as a means for quickly updating data that changes often, is seeing increased use, followed by standards-based capabilities for loading data (i.e., JDBC, ODBC, Web/RESTful services), followed by end-user data prep. Interestingly, metadata import and management is a lower priority in the context of data prep and loading compared to respondents’ use of metadata in other ADI preferences and priorities. Slightly more than 46 percent indicate “real-time / streaming, trickle, increments / change capture” is a priority for data prep and loading of ADI. Less than 30 percent of respondents consider support for Apache big data services “critical” or “very important.”

ADI Data Preparation and Loading

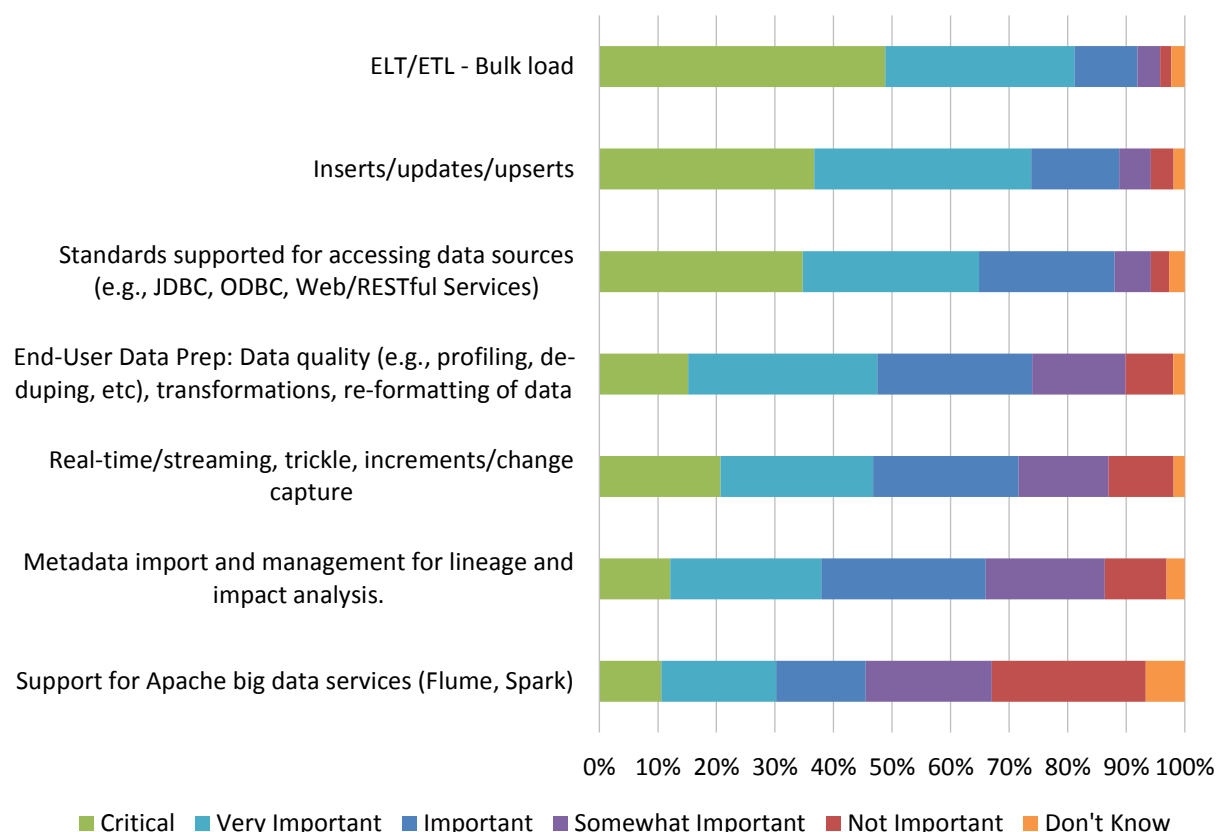


Figure 46 – ADI data preparation and loading

With the exception of the trend towards using inserts/updates/upserts style for updating data, the relative priority of data preparation and loading features does not change appreciably year over year (fig. 47). While respondents indicate metadata is a high priority for development and deployment, the importance of metadata for data loading and prep is not as high as one might expect, making it a problem in the future for governance and data management projects/practices.

ADI Data Preparation and Loading 2017-2019

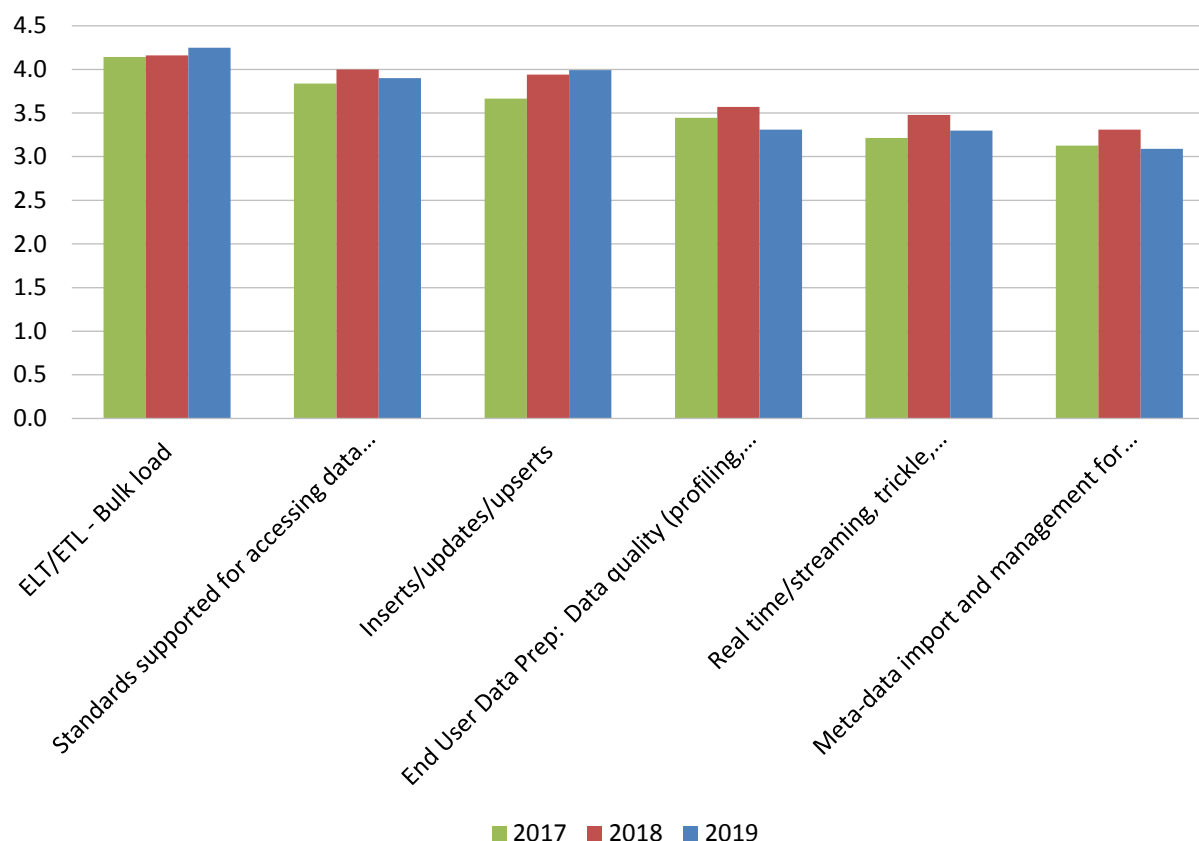


Figure 47 – ADI data preparation and loading 2017-2019

There are small variations across relative priorities and use cases. ET/ETL – Bulk loading is the top data prep and loading functionality across all use cases, and big data services are the lowest data prep and loading functionality across all use cases (fig. 48).

ADI Data Preparation and Loading by Top Use Case

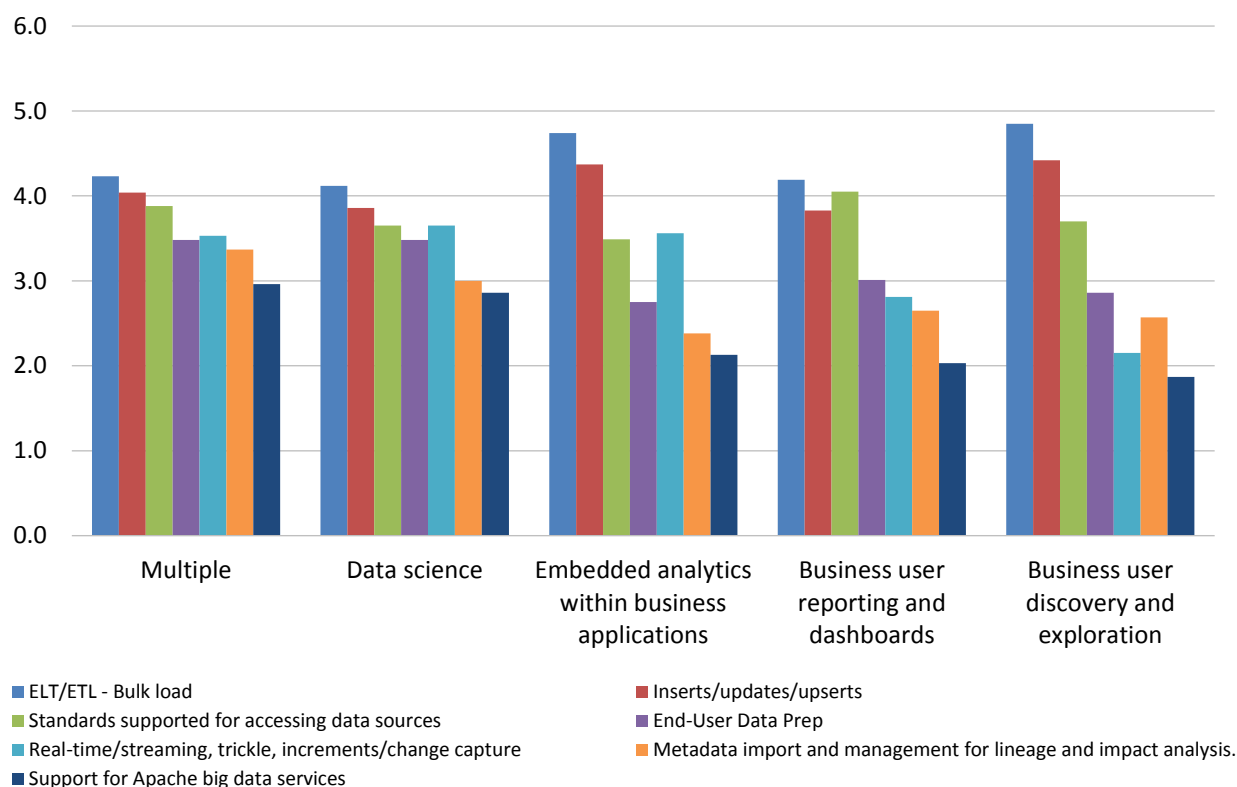


Figure 48 – ADI data preparation and loading by top use case

Asia-Pacific respondents' data preparation and loading priorities differ slightly from other regions. Asia-Pacific respondents indicate ETL / ELT and bulk load capabilities are a lower priority than support for standard interfaces (JDBC, ODBC, etc.) (fig. 49). The use of inserts/updates/upserts is more important in North America than other geographies. Support for Apache big data services ranks lowest in priority across all geography respondents.

ADI Data Preparation and Loading by Geography

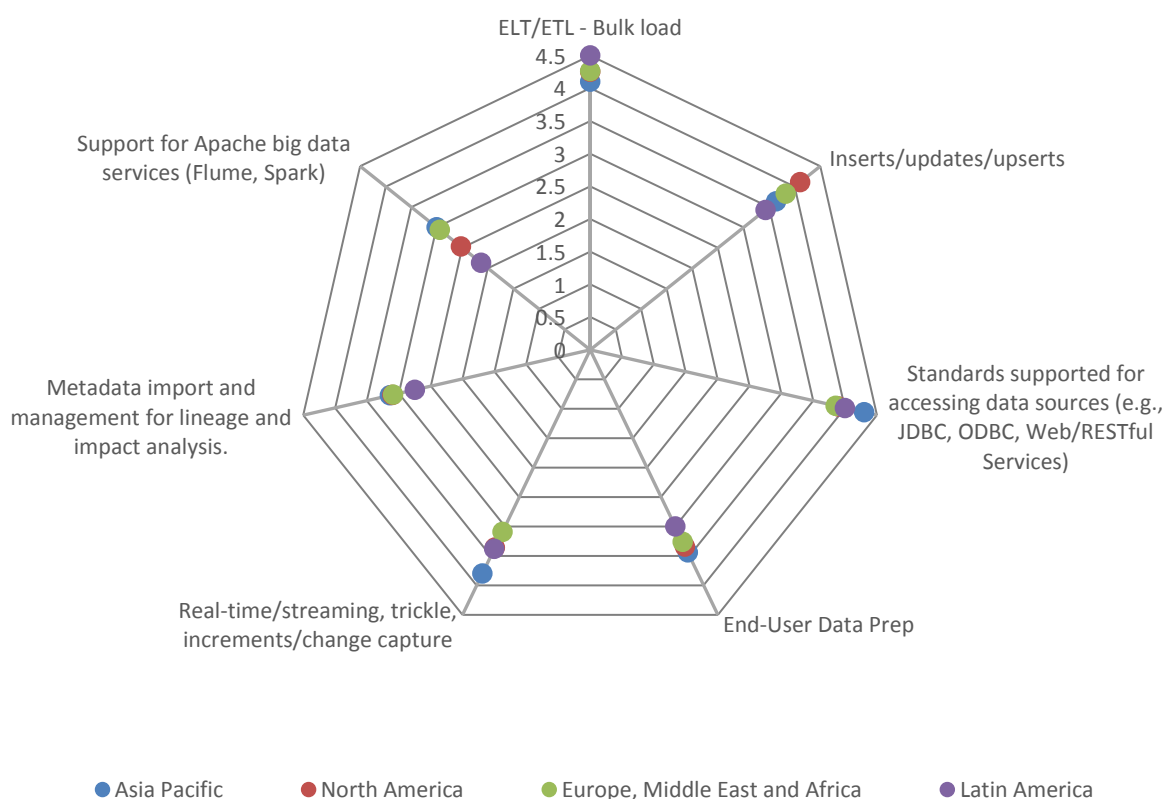


Figure 49 – ADI data preparation and loading by geography

With a few exceptions, most functions have similar relative priorities on data preparation and loading functions (fig. 50). R&D functions place a higher priority on big data integration services.

ADI Data Preparation and Loading by Function

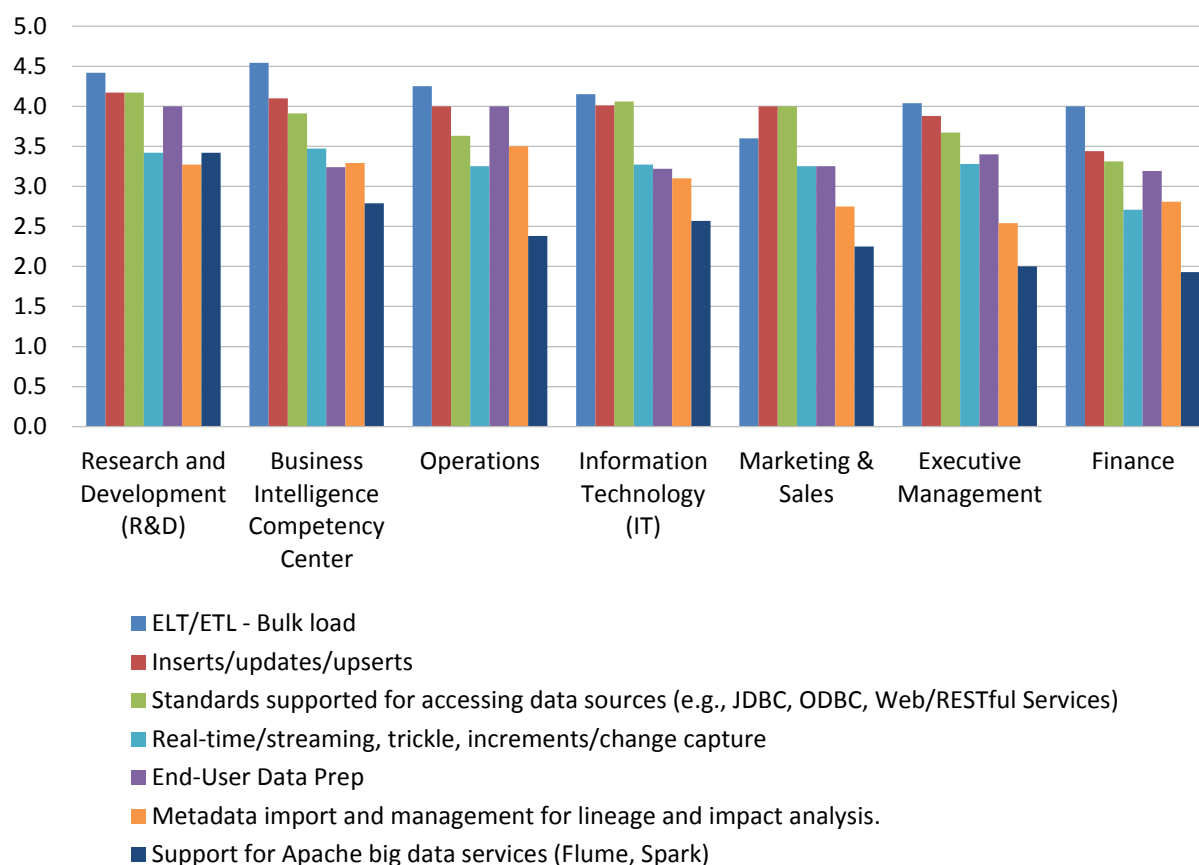


Figure 50 – ADI data preparation and loading by function

The relative priority of ADI data preparation and loading capabilities is similar across industry respondents; for example, ELT / ETL bulk loading is at the top of priorities, and support for Apache big data services is the lowest priority (fig. 51).

ADI Data Preparation and Loading by Industry

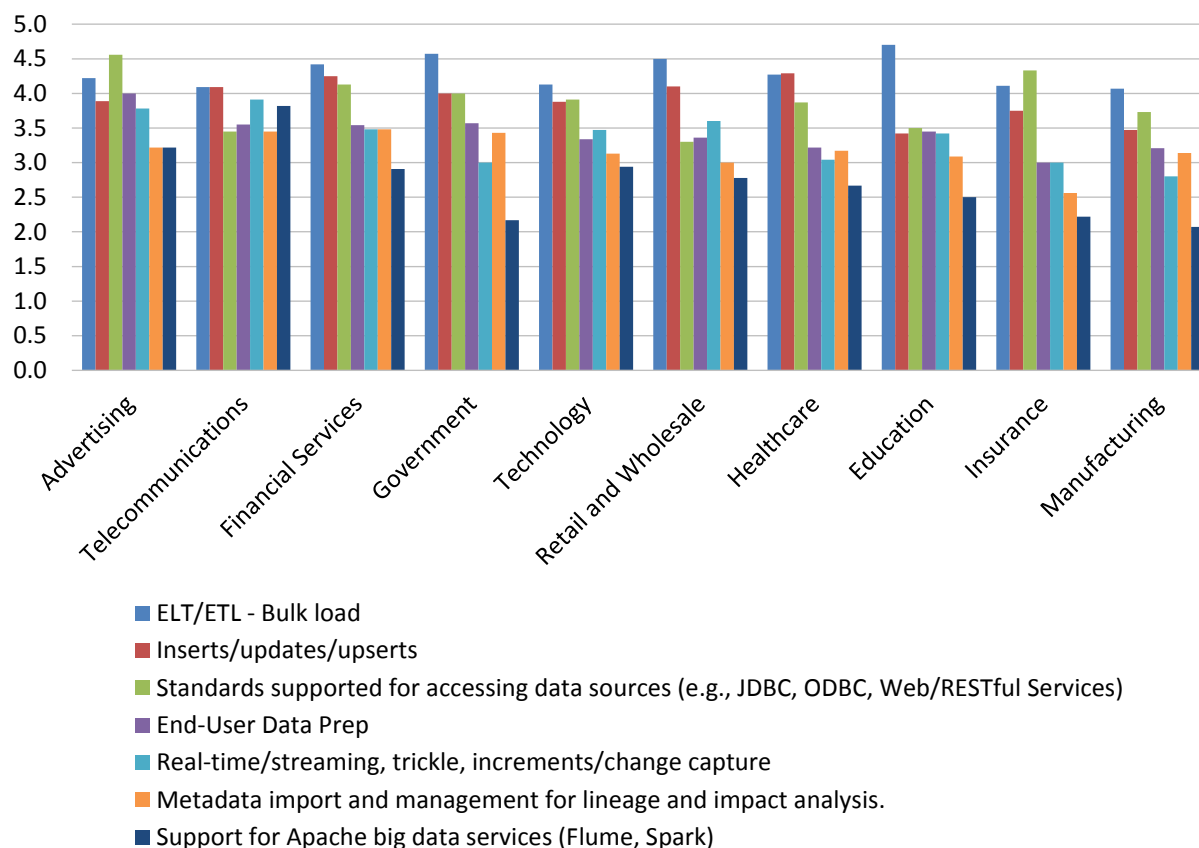


Figure 51 – ADI data preparation and loading by industry

We note only small differences in data prep and loading priorities by organization size (fig. 52). Larger organizations place a higher priority on ELT/ETL bulk load and metadata import and management capabilities. This is probably due to their higher priority in data life cycle management as a priority among development and deployment preferences (see fig. 32, p. 47).

ADI Data Preparation and Loading by Organization Size

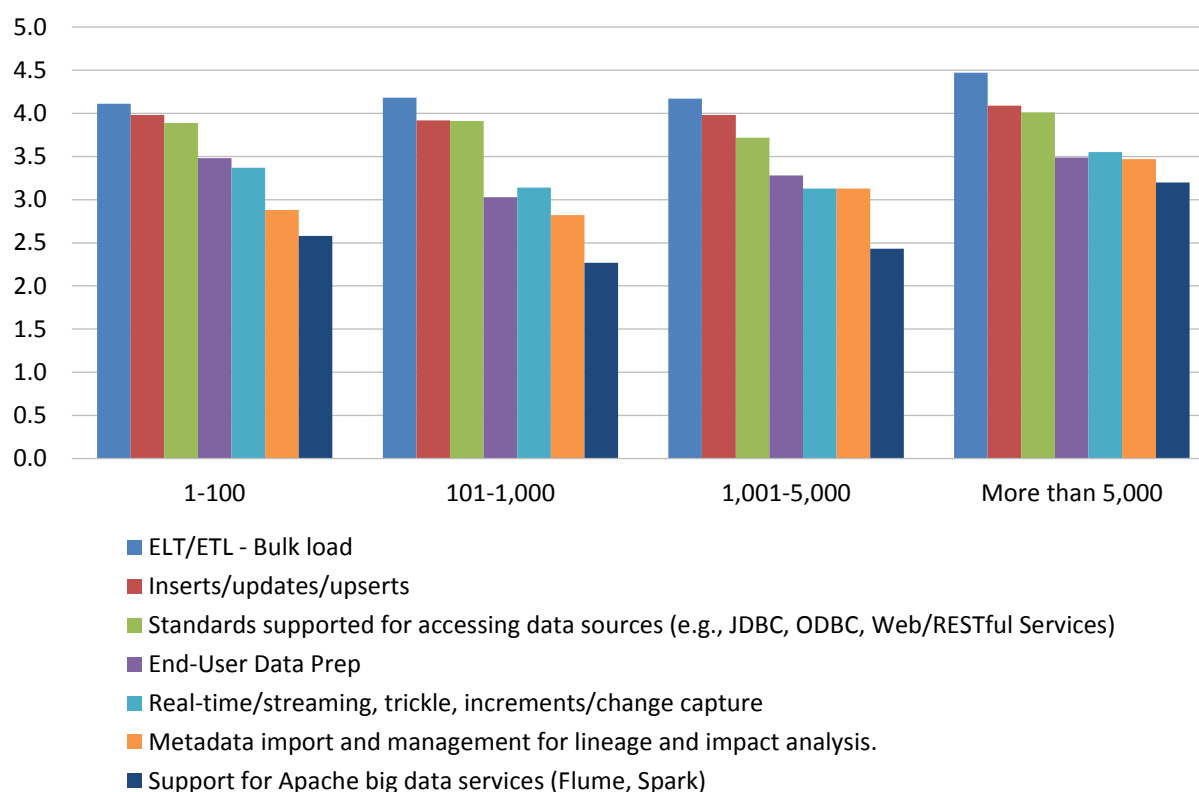


Figure 52 – ADI data preparation and loading by organization size

Analytical Data Infrastructure Data Model / Management of Data

SQL data is the top data model / management capability (fig. 53). This matches the top priority respondents place on transactional data (see fig. 39, p. 54).

Given the amount of market attention, in-memory data capabilities is a high priority for ADI platforms for more than 50 percent of respondents.

Non-SQL and hierarchical files (e.g., Hadoop HDFS) are high priorities for less than 40 percent of respondents.

ADI Data Model / Management of Data

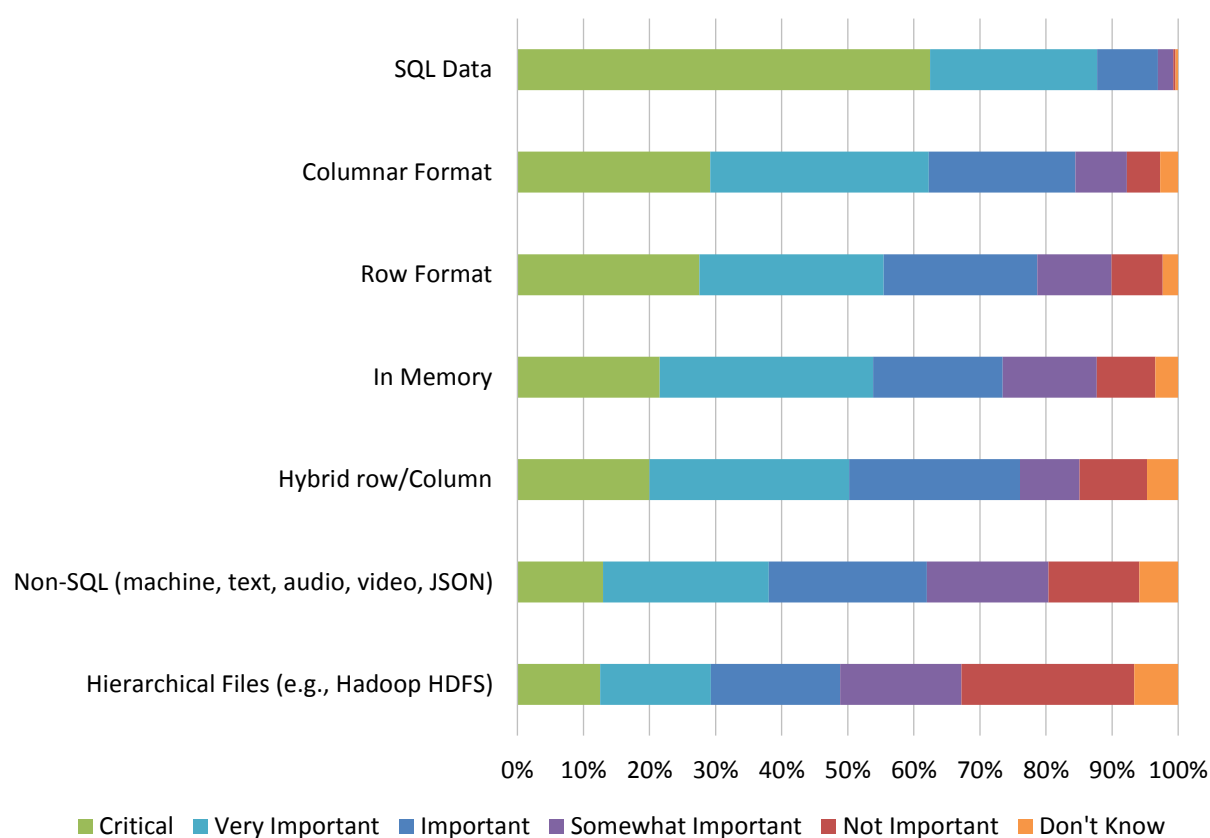


Figure 53 – ADI data model / management of data

Non-SQL and hierarchical files/HDFS support remain the lowest relative priority for ADI data models/data and show variability in importance year over year (fig. 54). Columnar and hybrid row/column models increase in preference year over year.

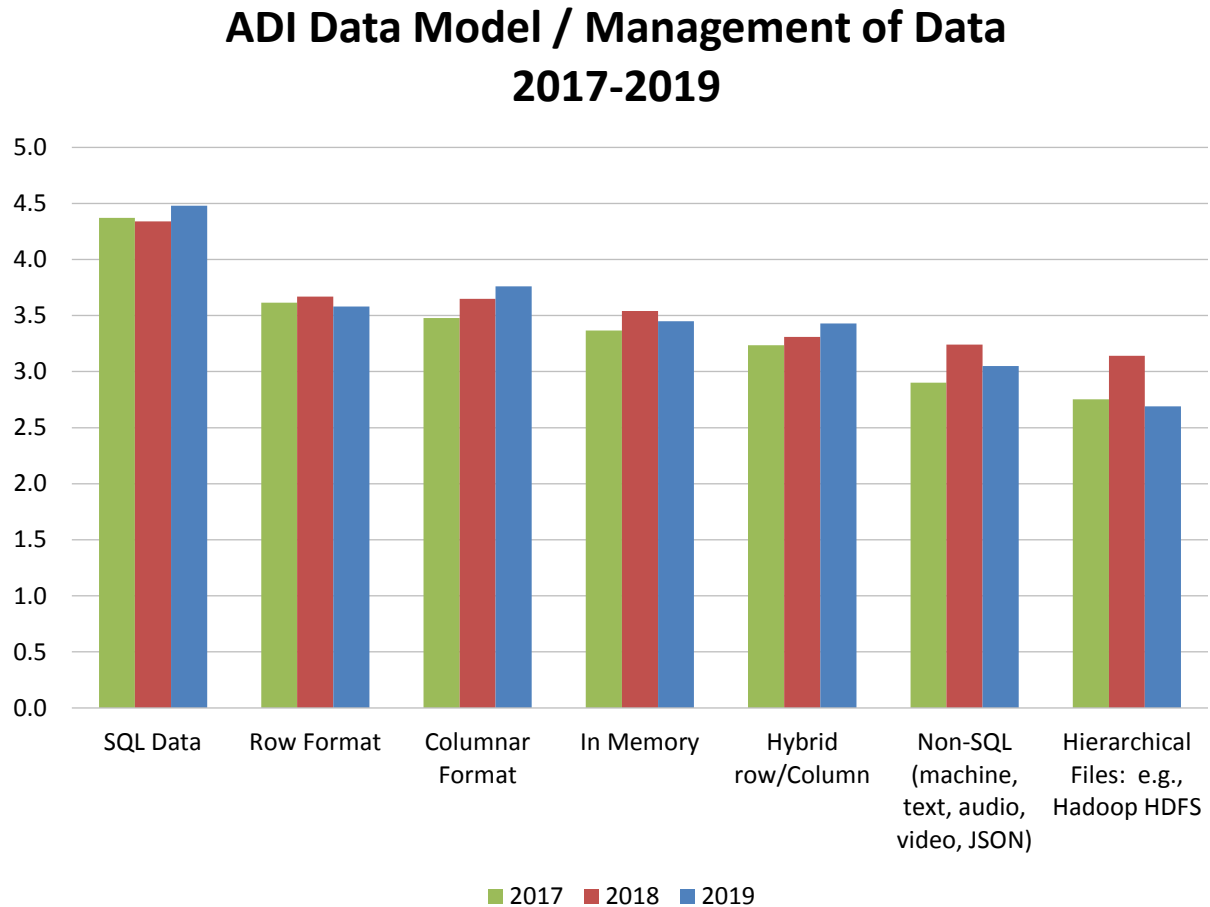


Figure 54 – ADI data model / management of data 2017-2019

SQL data models are the primary data management models for most analytic applications and use cases (fig. 55). Both “business user reporting and dashboards” and “business user discovery and exploration” place a higher priority on SQL data than other use cases. Hierarchical files is the lowest priority across all use cases.

ADI Data Model / Management of Data by Top Use Case

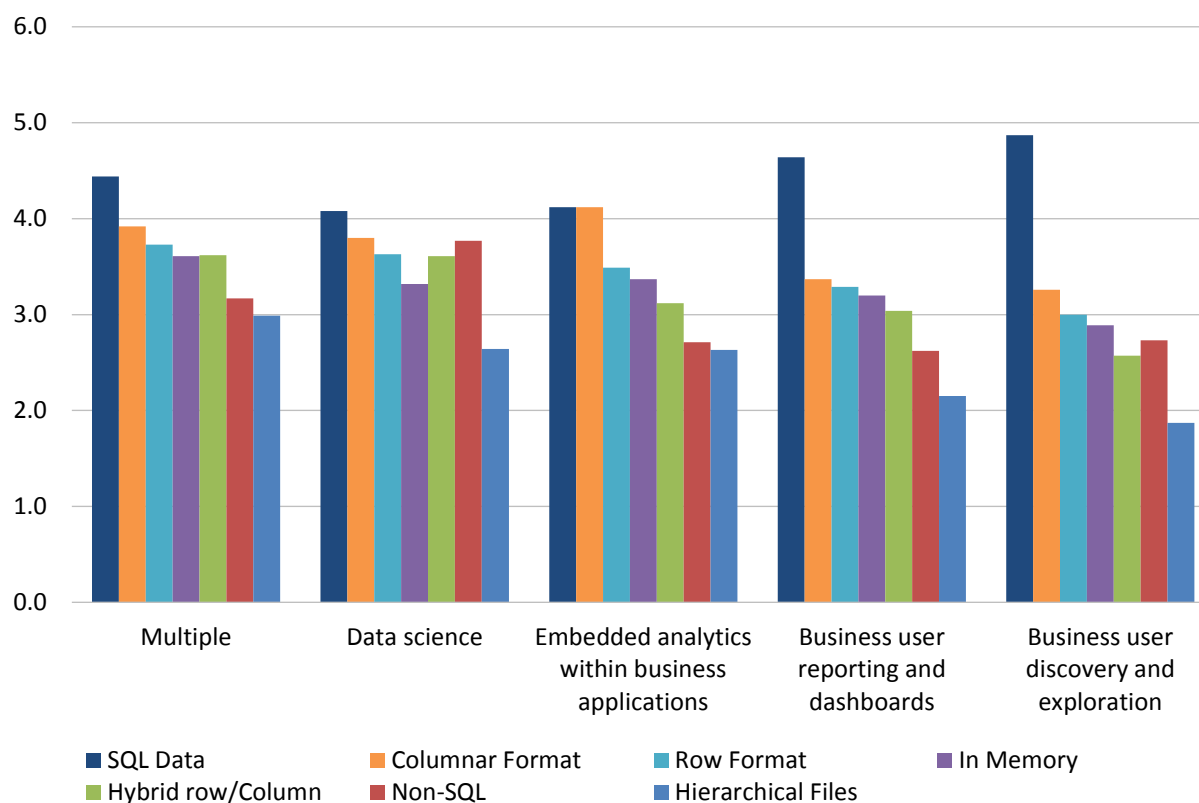


Figure 55 – ADI data model / management of data by top use case

Respondents in all geographies place their highest priority on support for SQL data types; however, we see some regional variations for the data model/management choices for ADI platforms (fig. 56).

The largest variation in priorities for data model / management capabilities for ADI platforms across geographies is the priority for the use of hierarchical files as a part of the ADI platform. The lowest priority is in Latin America, and the highest is in Asia Pacific.

ADI Data Model / Management of Data by Geography

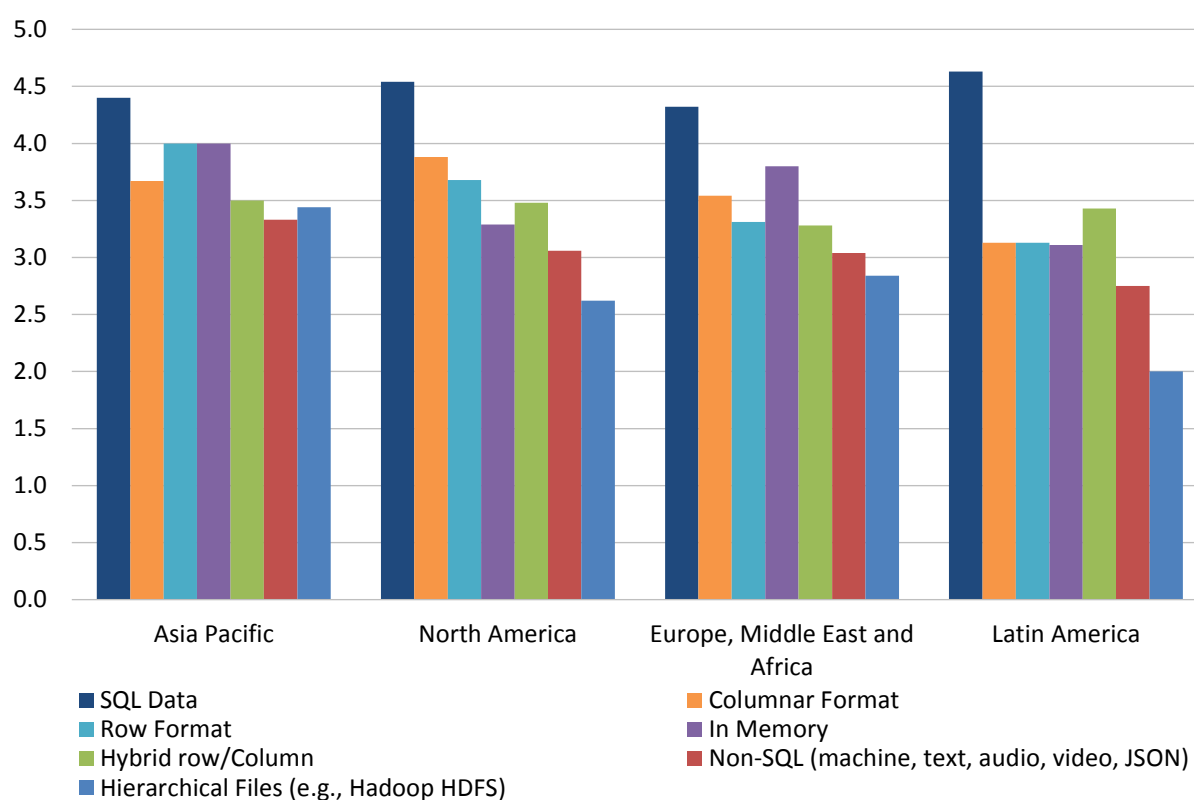


Figure 56 – ADI data model / management of data by geography

ADI data model / management priorities vary by respondents' organizational function (fig. 57). Variation in priorities by functions can cause challenges to leaders of analytic initiatives if they need involvement and alignment of priorities across technical and economic buyers from different functions within an organization.

Marketing/Sales respondents place a high relative priority on hybrid row and column capabilities. The BICC respondents place a higher priority on in-memory capabilities compared to other organization functions.

Only R&D respondents indicate a high priority for hierarchical files, which perhaps indicates use cases still developing for this data management capability.

ADI Data Model / Management of Data by Function

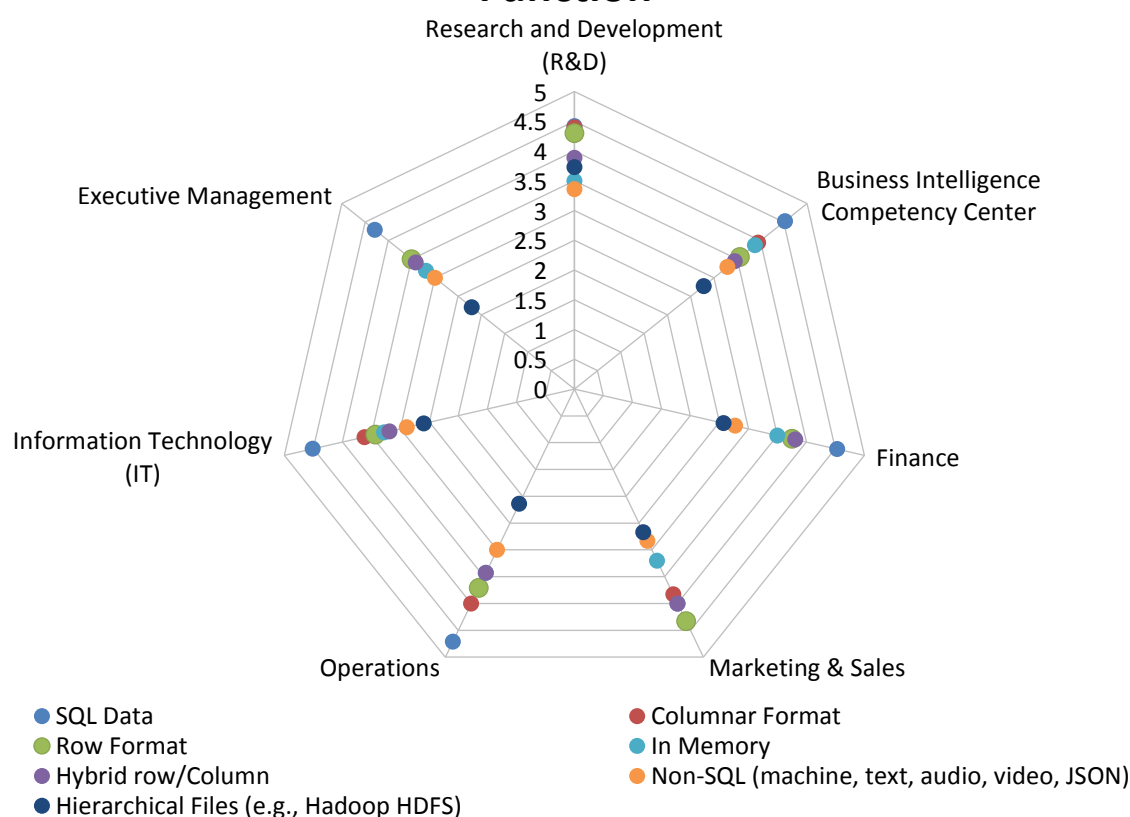


Figure 57 – ADI data model / management of data by function

The relative priority for ADI data model / management capabilities varies quite a bit by industry of the respondent. The largest variability of priorities is for hierarchical files, e.g., Hadoop HDFS, followed by in memory (fig. 58). Telecommunications and Financial services respondents indicate priority for a variety of data model / management types compared to other industries. Manufacturing has the lowest variability of data type priorities, with columnar format ranked the second highest priority (behind SQL data).

ADI Data Model / Management of Data by Industry

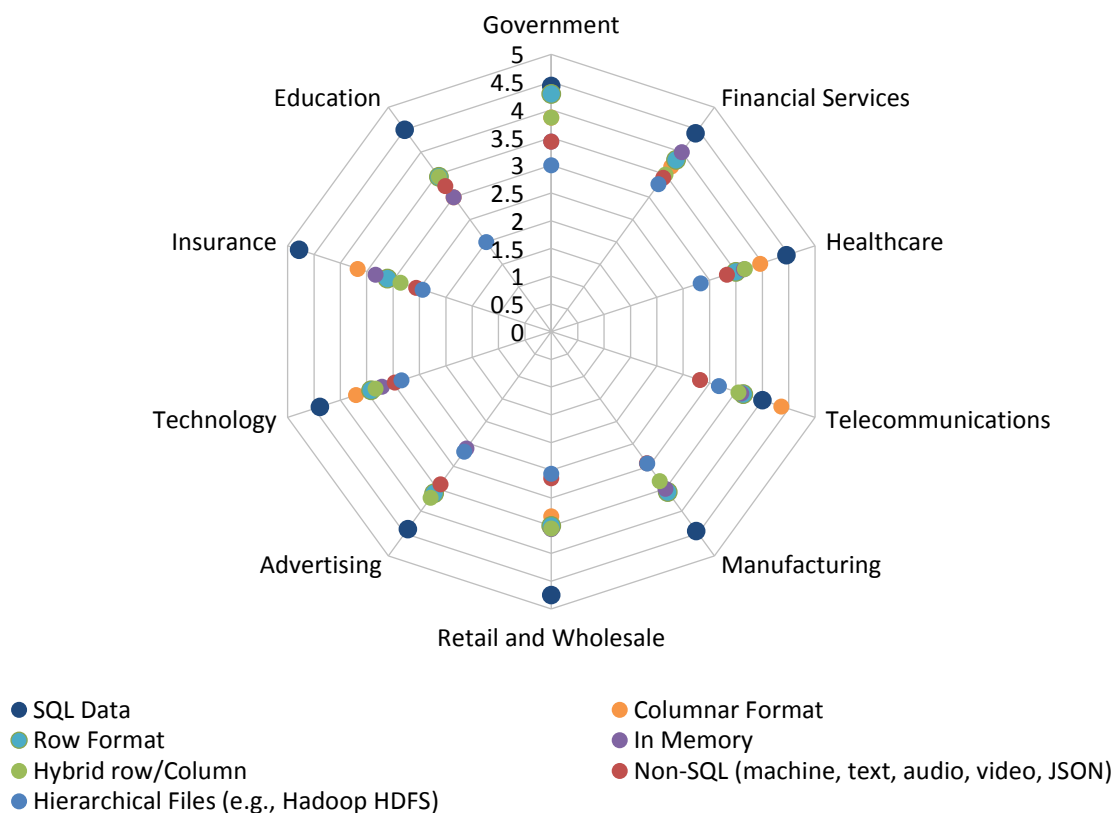


Figure 58 – ADI data model / management of data by industry

Larger organizations place higher priority on data models other than SQL data. For example, the priority for hierarchical files, in-memory, and row and columnar data models is higher for larger organizations with more than 5,000 employees) compared to the priorities of smaller organizations.

The relative priorities (i.e., SQL data as the top priority and hierarchical files as the lowest priority) of ADI data models / management of data are similar to the results of other dimensions (i.e., industry, geography, and function) (fig. 59).

ADI Data Model / Management of Data by Organization Size

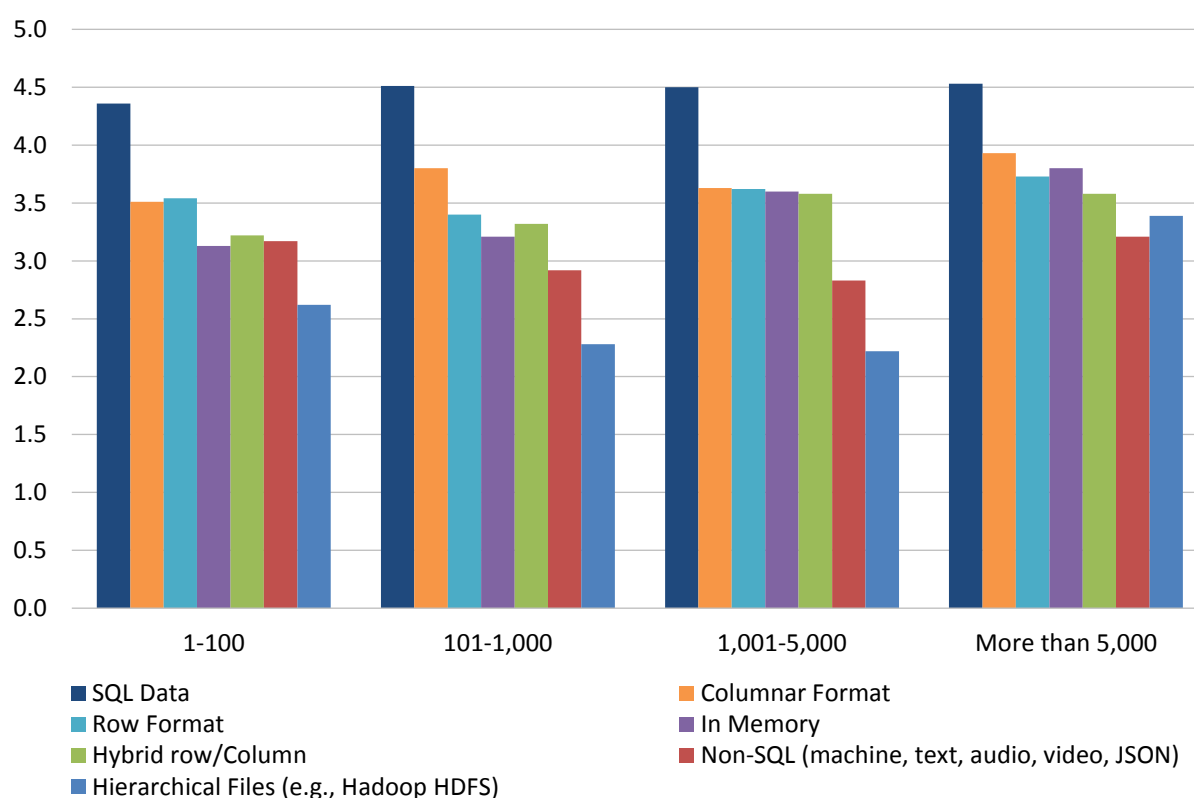


Figure 59 – ADI data model / management of data by organization size

Analytical Data Infrastructure Interfaces

ADI platforms need to support interfaces for Excel/CSV and ODBC and lead the way for ADI interface priorities (fig. 60). The popularity of in-memory capabilities is likely the cause for the low priority/use of multi-dimensional models (MDX). Many respondents indicate “don’t know” or “not important” regarding several Apache-based services such as Impala, Avro, Parquet, indicating the lack of use/maturity of the technology.

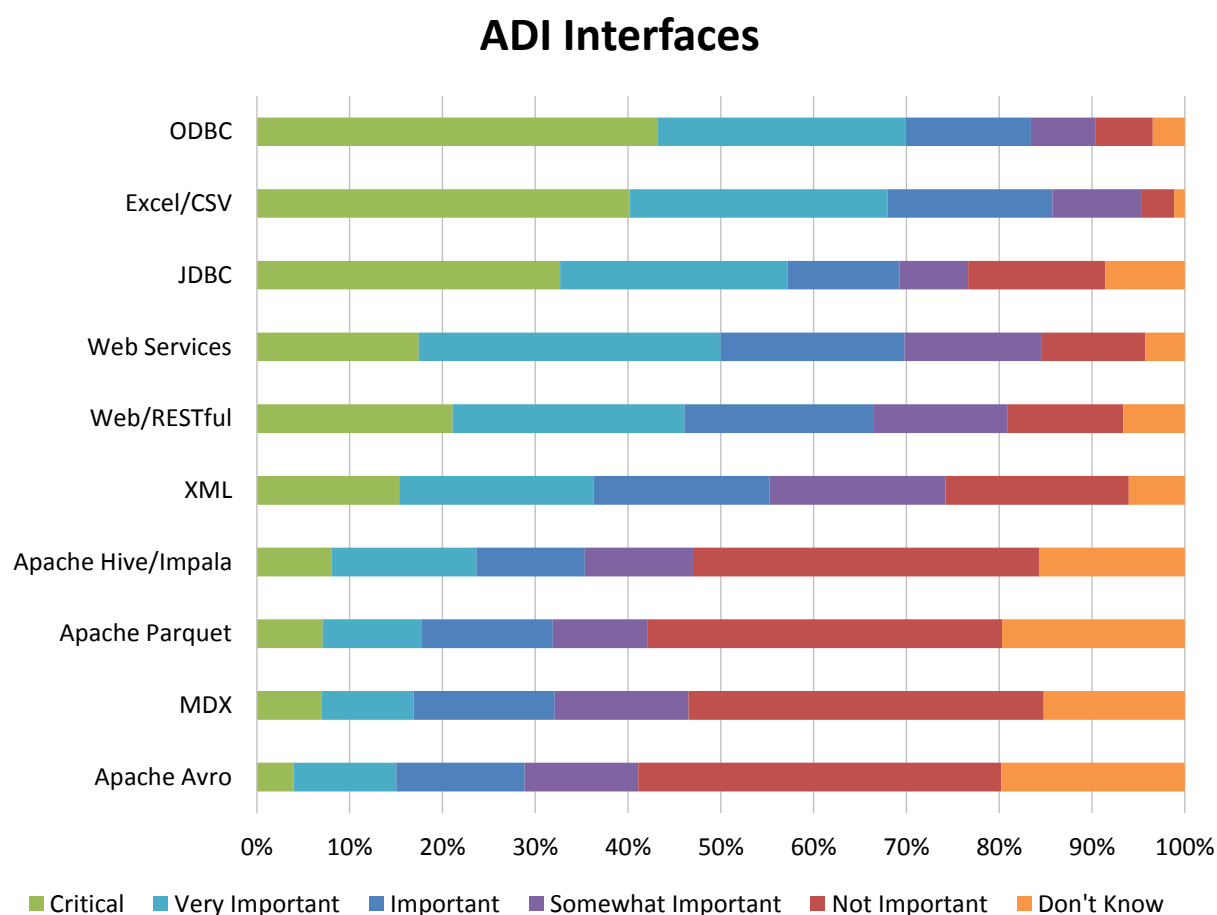


Figure 60 – ADI interfaces

Not much changes in terms of ADI platform interface priorities over the past three years. While the Apache interfaces (i.e., Hive/Impala, Parquet, Avro) are low in terms of relative priority, we see year-over-year growth in importance (fig. 61). XML interfaces shows a slight trend in decline of importance.

ADI Interfaces 2017-2019

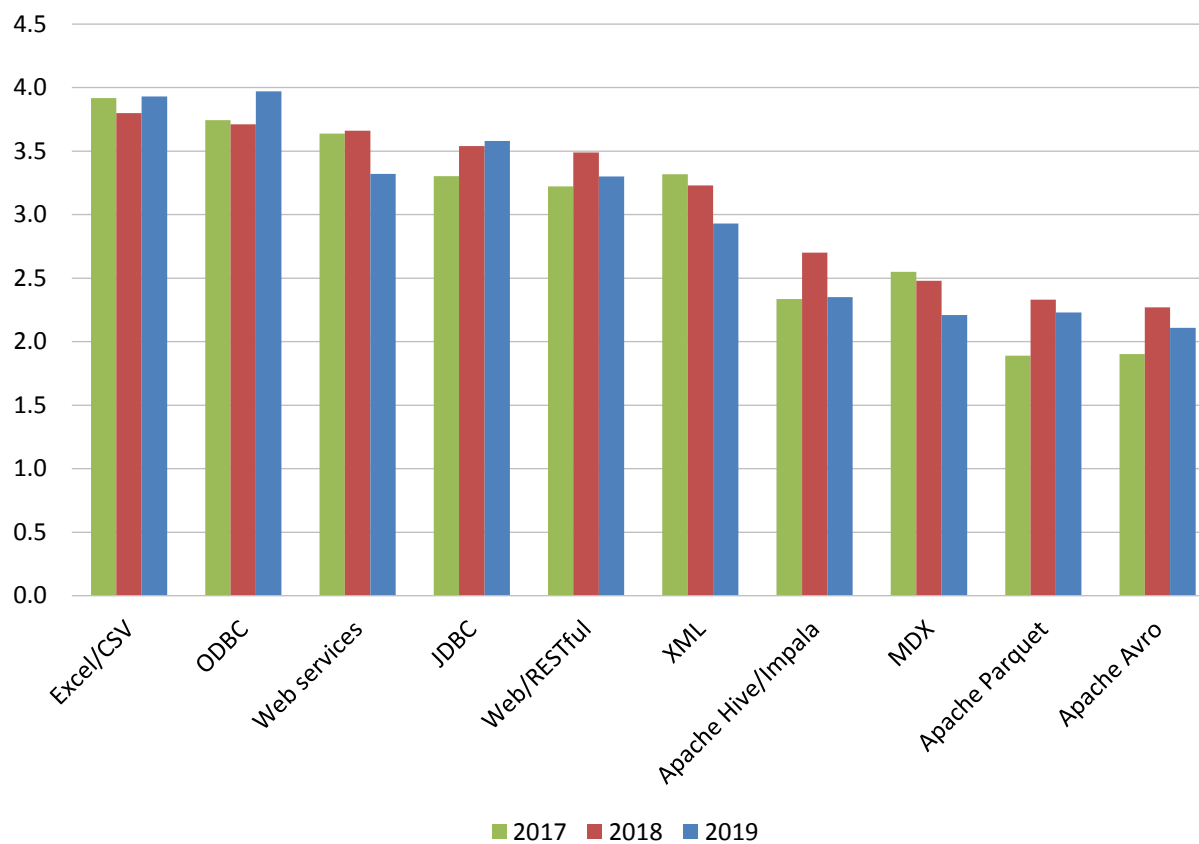


Figure 61 – ADI interfaces 2017-2019

ADI interface relative priorities do not vary a lot by use case, but there are some interesting variations of note (fig. 62). Respondents with a “business user discovery and exploration” use case priority rank interfaces for Excel/CSV data as important as ODBC interfaces. Respondents with data science use cases rank Excel/CSV interfaces and Web services at the top of their interface priorities. Embedded analytic use cases place a top priority on JDBC and Web/RESTful APIs. Business user reporting and dashboards place a top priority on JDBC and Web/RESTful APIs. Business user discovery and exploration place a top priority on JDBC and Web/RESTful APIs.

ADI Interfaces by Top Use Case

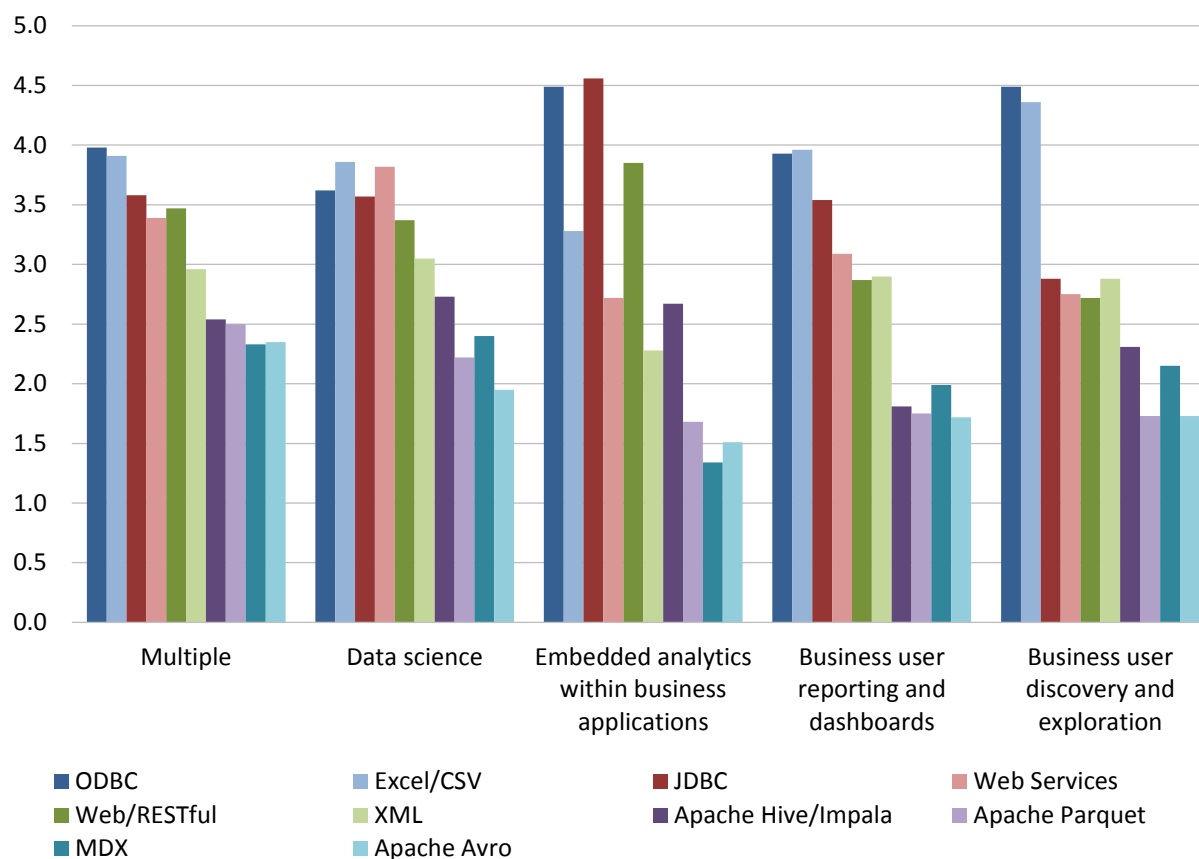


Figure 62 – ADI interfaces by top use case

As fig. 63 shows, there is not a lot of variation in respondents' priorities for ADI interfaces by geography. Asia-Pacific respondents place a comparatively higher priority for MDX and Apache services.

ADI Interfaces by Geography

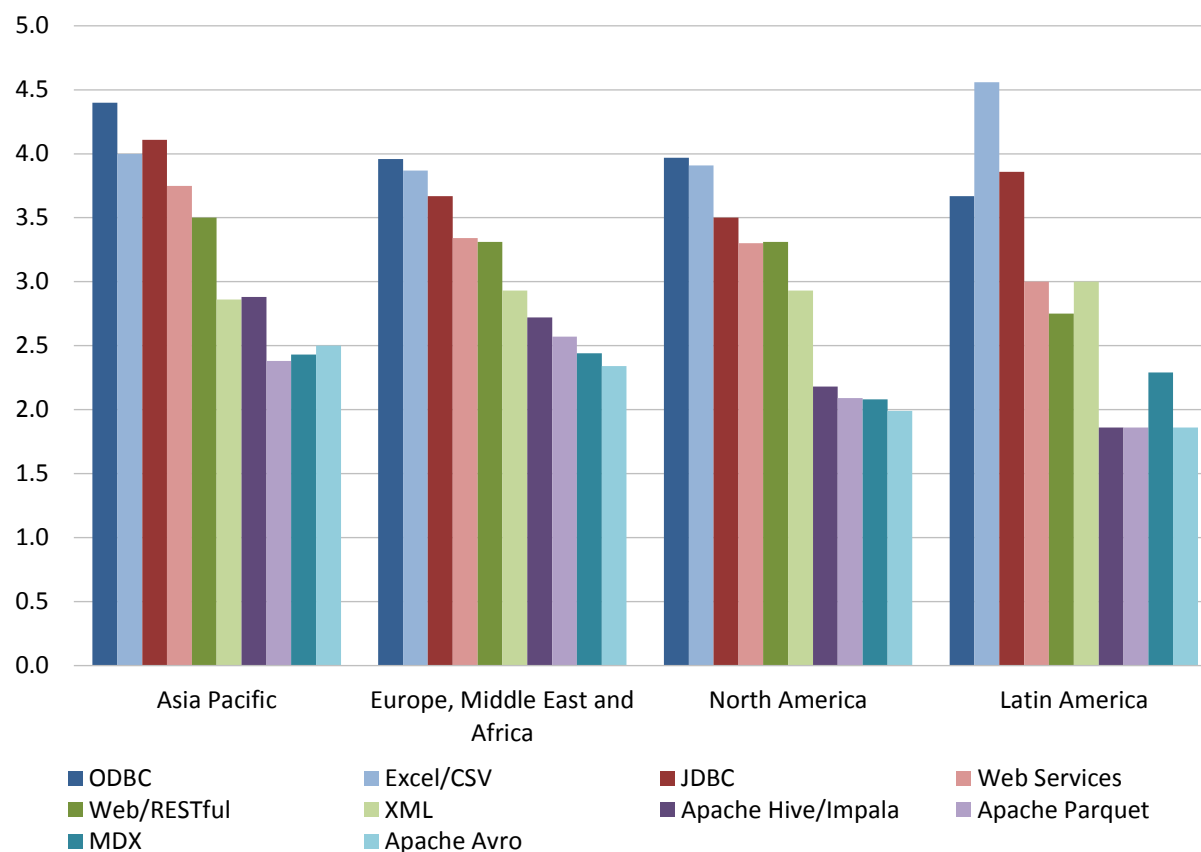


Figure 63 – ADI interfaces by geography

There is quite a bit of variability in priorities for ADI interfaces by organizational function. This may cause alignment problems when organizations rank priorities for a vendor's ADI proposal that spans multiple functions. Finance and Operations respondents place their highest priority on Excel/CSV interfaces (fig. 64) and their lowest priority on Apache services. This is not surprising, given the amount of work done in spreadsheets by Finance and Operations organizations. The Finance function also places a high priority on MDX interfaces. This is not surprising, given the use case of budgeting analysis modeled as multi-dimensional models.

R&D functions place a higher priority on Apache interfaces in keeping with the emerging big data use cases developed by R&D and data types, which require these interfaces.

ADI Interfaces by Function

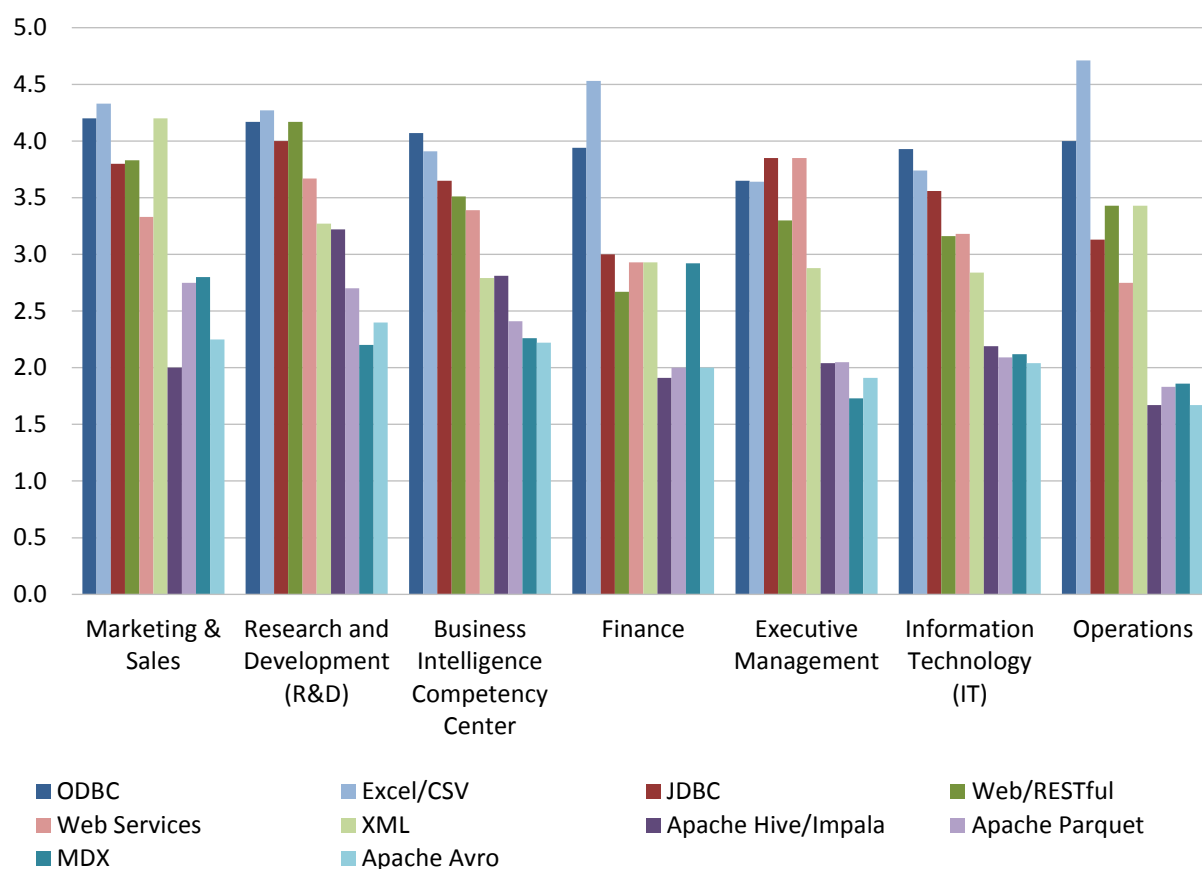


Figure 64 – ADI interfaces by function

This year, we see quite a bit of variability in priorities across industries for their ADI interface preferences/priorities (fig. 65).

ODBC and Excel/CSV interfaces rank as the top preferences for accessing data within an ADI platform. MDX, once a popular multi-dimensional interface, still sees some preferences from respondents within Manufacturing and Retail/Wholesale industry segments. Hive/Impala and Parquet interfaces show up as preferences with respondents from the Telecommunications and Advertising industries.

ADI Interfaces by Industry

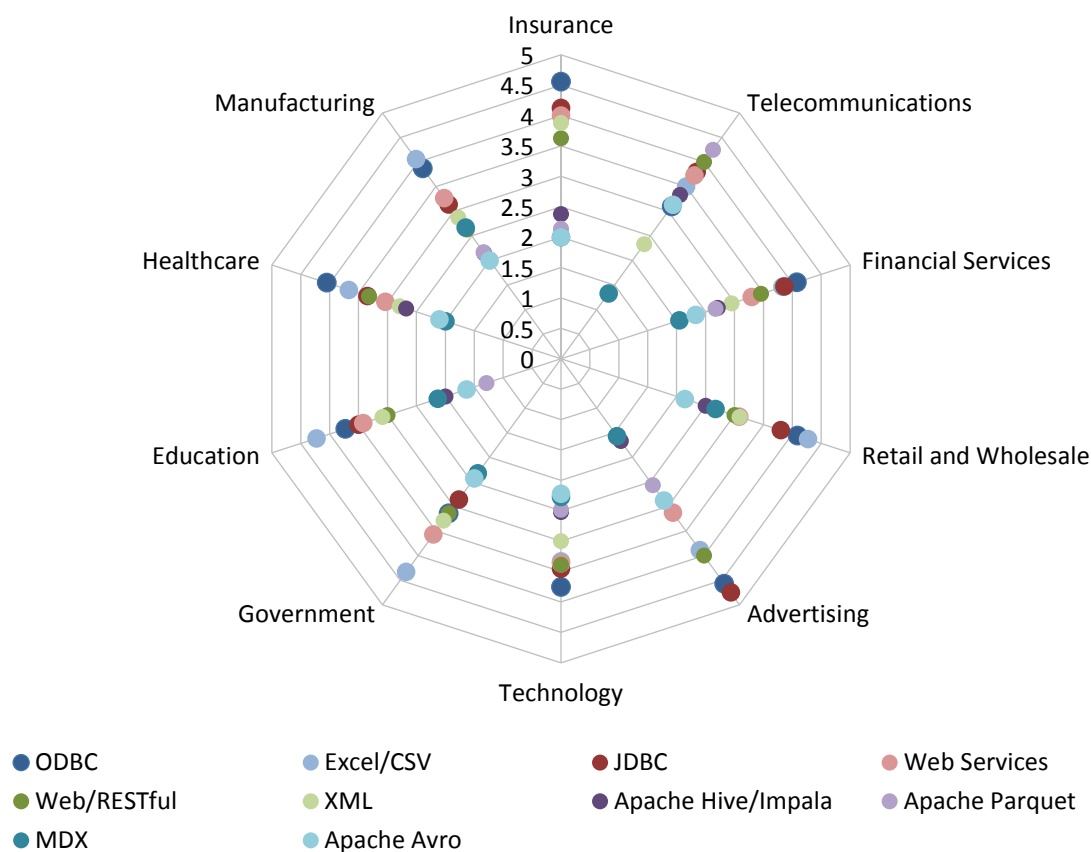


Figure 65 – ADI interfaces by industry

ODBC, Excel/CSV, and JDBC interfaces are the primary interfaces for accessing ADI platform data. Larger-sized organizations place higher relative priority on Apache interfaces (fig. 66). Smaller organizations with 1-100 employees place higher priority on Web/Restful interfaces. Smaller organizations' preference for Web/Restful interfaces is consistent with their indicated preference for cloud deployments (fig. 24 p. 39).

ADI Interfaces by Organization Size

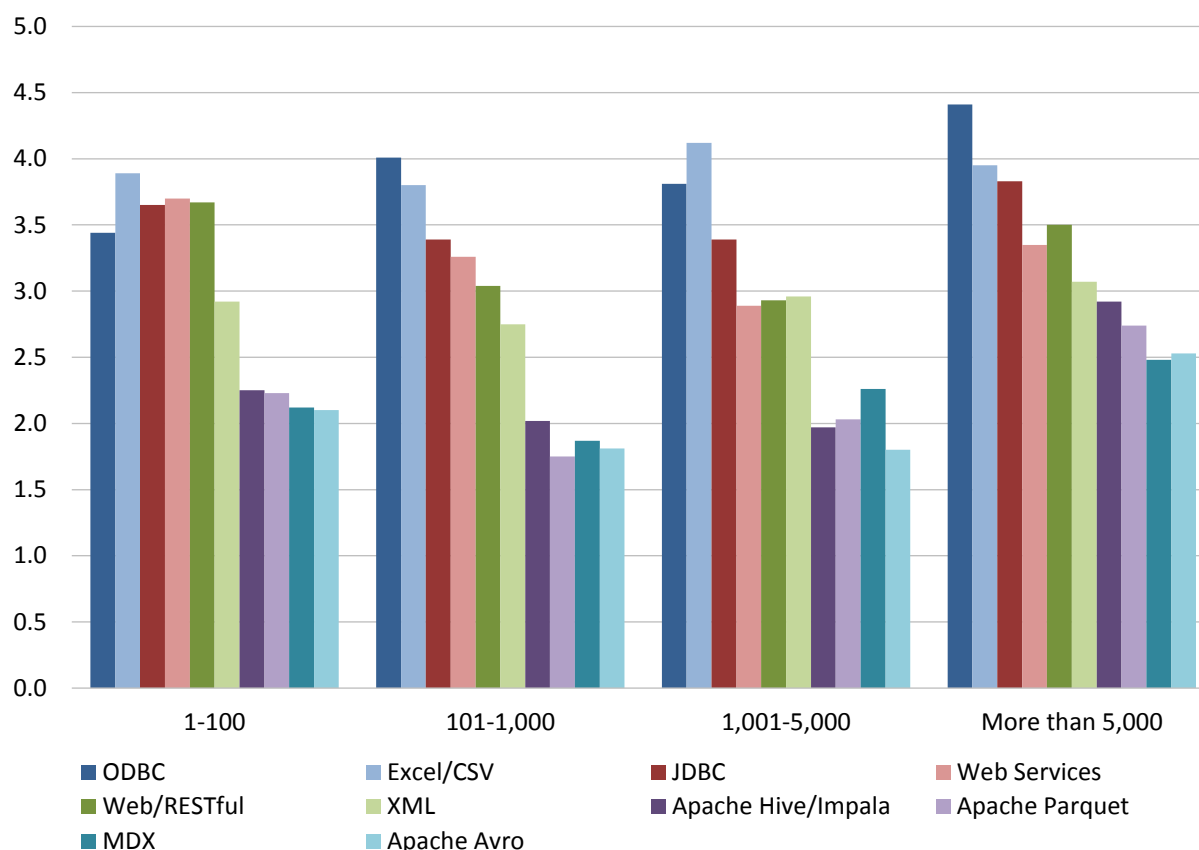


Figure 66 – ADI interfaces by organization size

ADI Analytical Features

The diversity and priority for analytical features within an ADI platform vary widely. Aggregations, statistical analysis/R, and multi-dimensional/OLAP features rank highest, while sentiment analysis and path/link analysis rank lowest (fig. 67). Many respondents indicate they “don’t know” or rank as “not important” such ADI analytical features as Spark, Map Reduce, sentiment analysis, and path/link analysis. Statistical analysis and support for R ranks as a high priority, demonstrating continuing interest and growth of R. While multi-dimensional/OLAP analytical features rank as high priority, it is interesting to contrast this finding with the low priority for MDX as an interface (see fig. 60, p. 75). While 60 percent of respondents indicate a high priority for multi-dimensional/OLAP analysis, only a small percent of respondents say MDX interfaces are a high-priority capability.

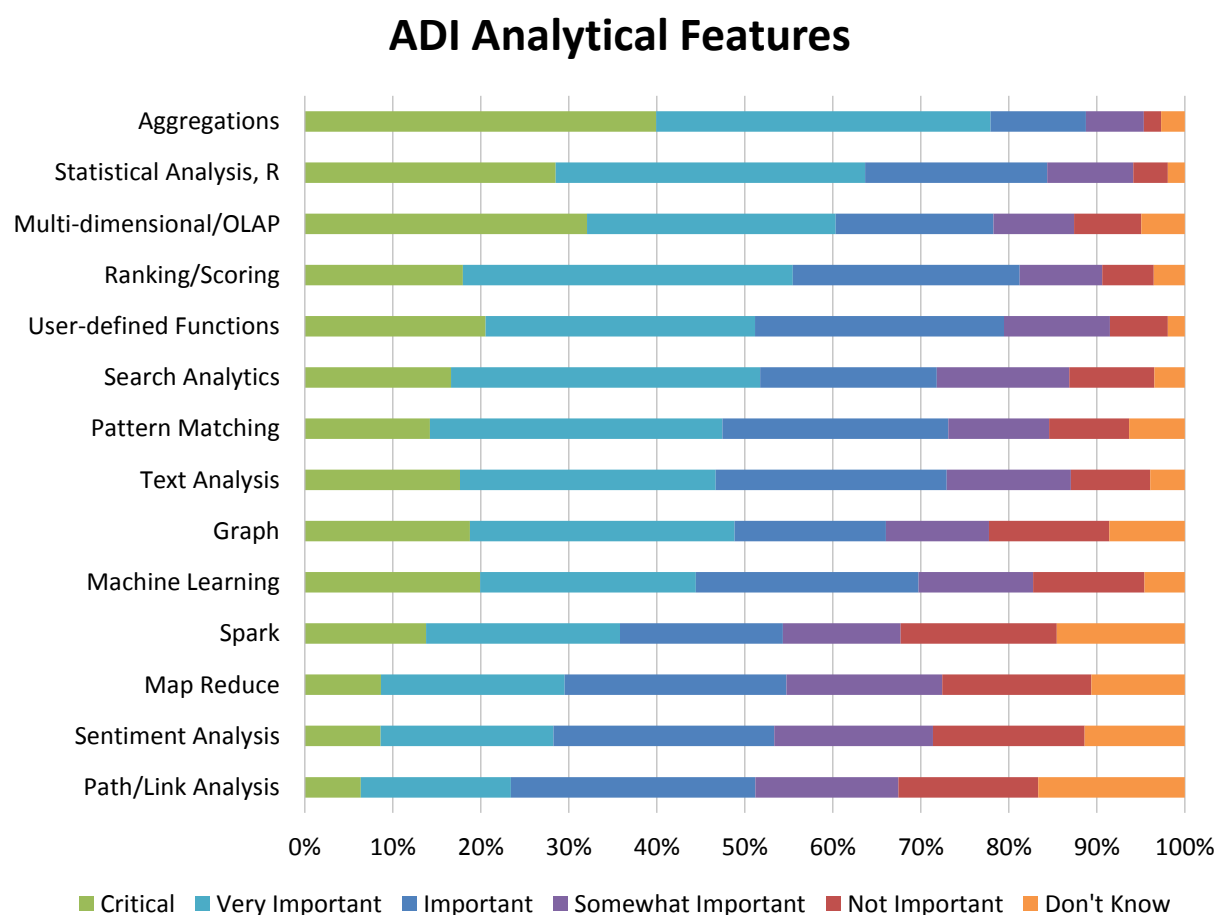


Figure 67 – ADI analytical features

There are some modest year-over-year changes to ADI analytical feature requirements. Statistical analysis/R features increase in priority year over year. More advanced analytic features, such as machine learning, MapReduce, and Spark, also increase in priority for ADI platforms (fig. 68). Spark analytic features for ADI platforms increase the most year over year with an increase in priority of approximately 10 percent.

ADI Analytical Features 2017-2019

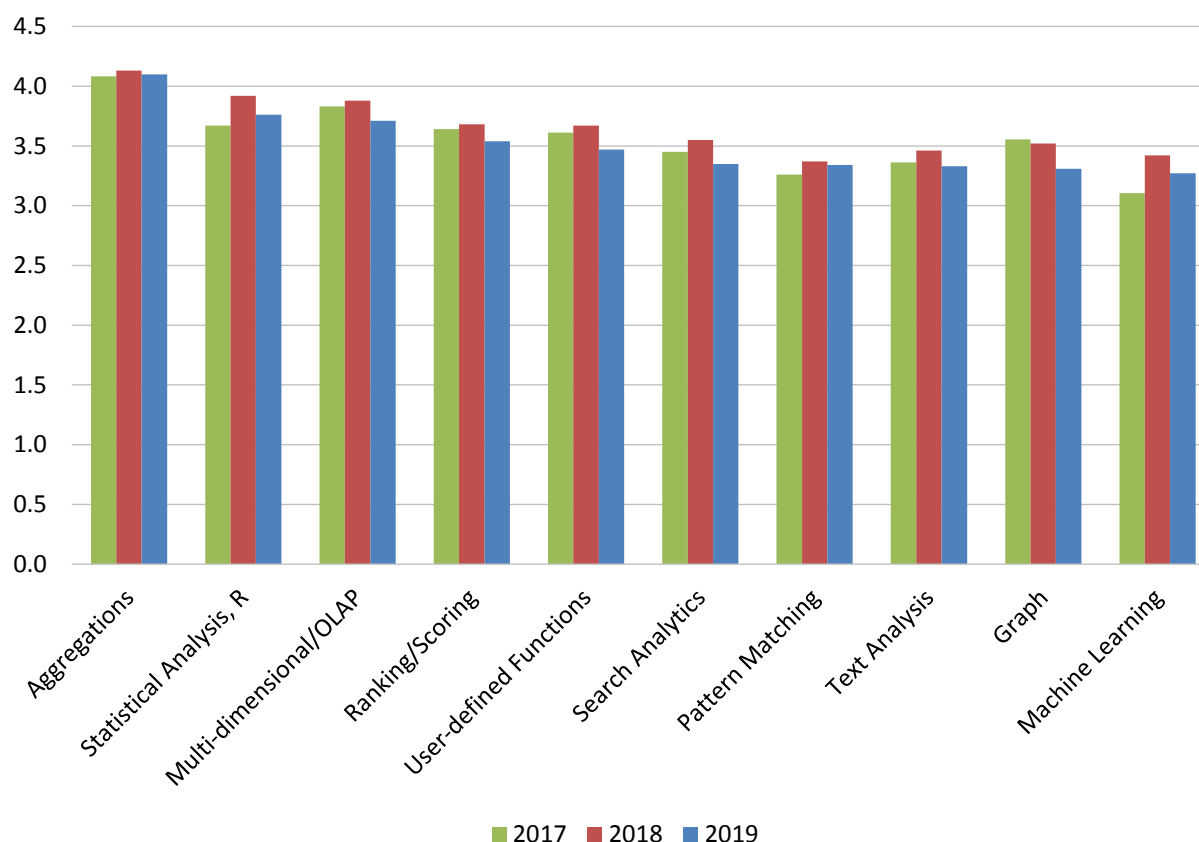


Figure 68 – ADI analytical features 2017-2019

While aggregations are the top analytical feature across all use cases, other features vary more in priority by use case (fig. 69). Respondents with data science as their top use case place their highest priority on statistical analysis and support for languages such as R. Machine learning is also a top priority for data science use cases. Graph analysis is the lowest priority for embedded use-case priorities.

ADI Analytical Features by Top Use Case

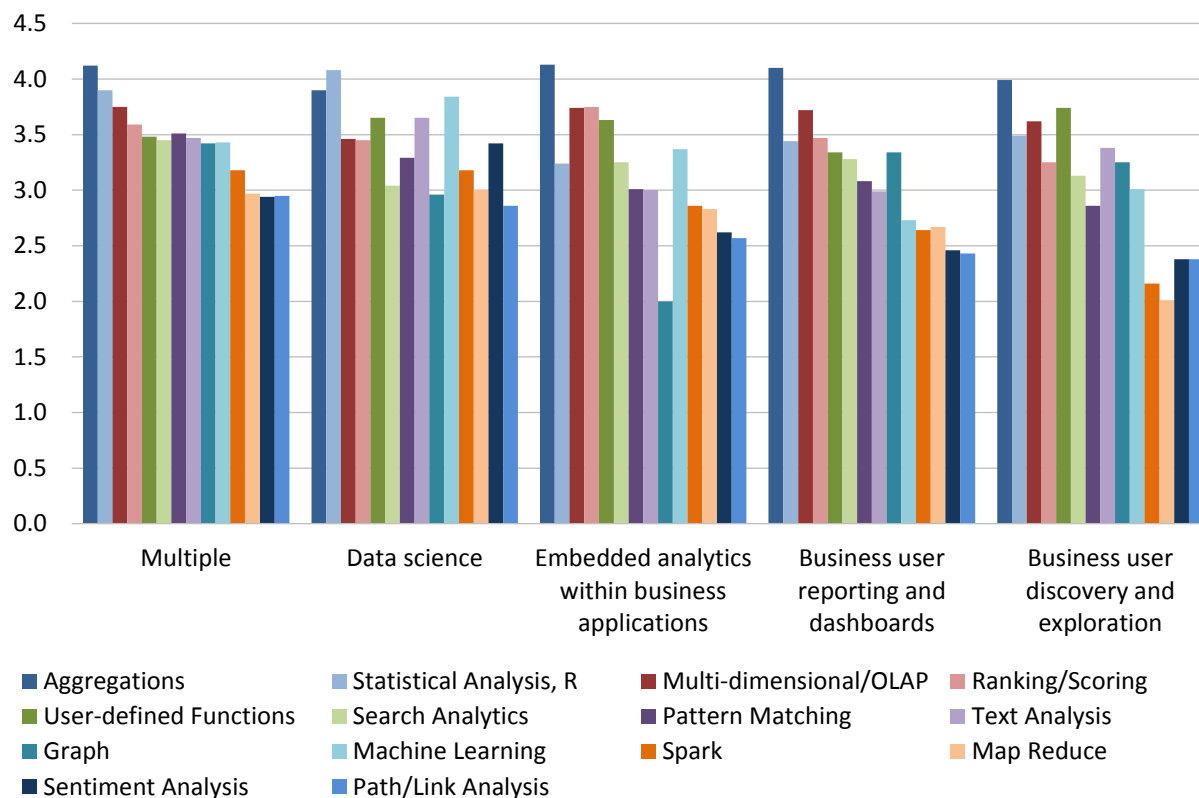


Figure 69 – ADI analytical features by top use case

Asia-Pacific respondents prioritize path/link analysis as a top priority for the ADI platform; other regions rank path/link analysis as their lowest priority (fig. 70). Latin America has a higher-than-average preference for multi-dimensional/OLAP and sentiment analytical features. Spark features are lower priority for most organizations, with the exception of EMEA respondents, who rate it of high importance.

ADI Analytical Features by Geography

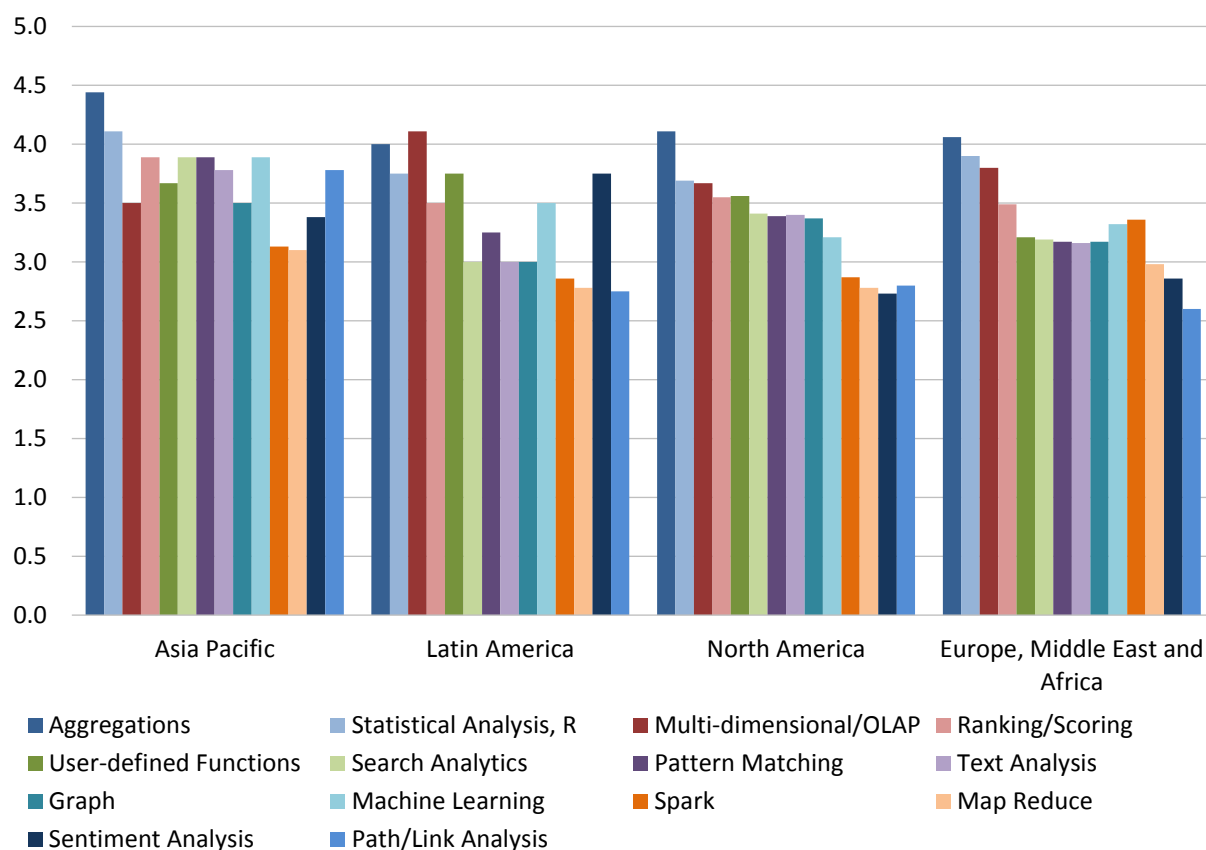


Figure 70 – ADI analytical features by geography

The priority of ADI analytic features by organization functions varies quite a bit. This is like the results for other sections of this report where priorities vary by organization function. Marketing/Sales functions, followed by the BICC, have the highest priority for almost all analytical features (fig. 71) including more advanced analysis such as sentiment, path/link, and Spark analytic features. In contrast, IT respondents place lower priority on more advanced analytical features such as machine learning, Map-Reduce, path/link analysis, and sentiment analysis. This may create challenges when trying to gain agreement and align IT and non-IT priorities for ADI analytic functions.

ADI Analytical Features by Function

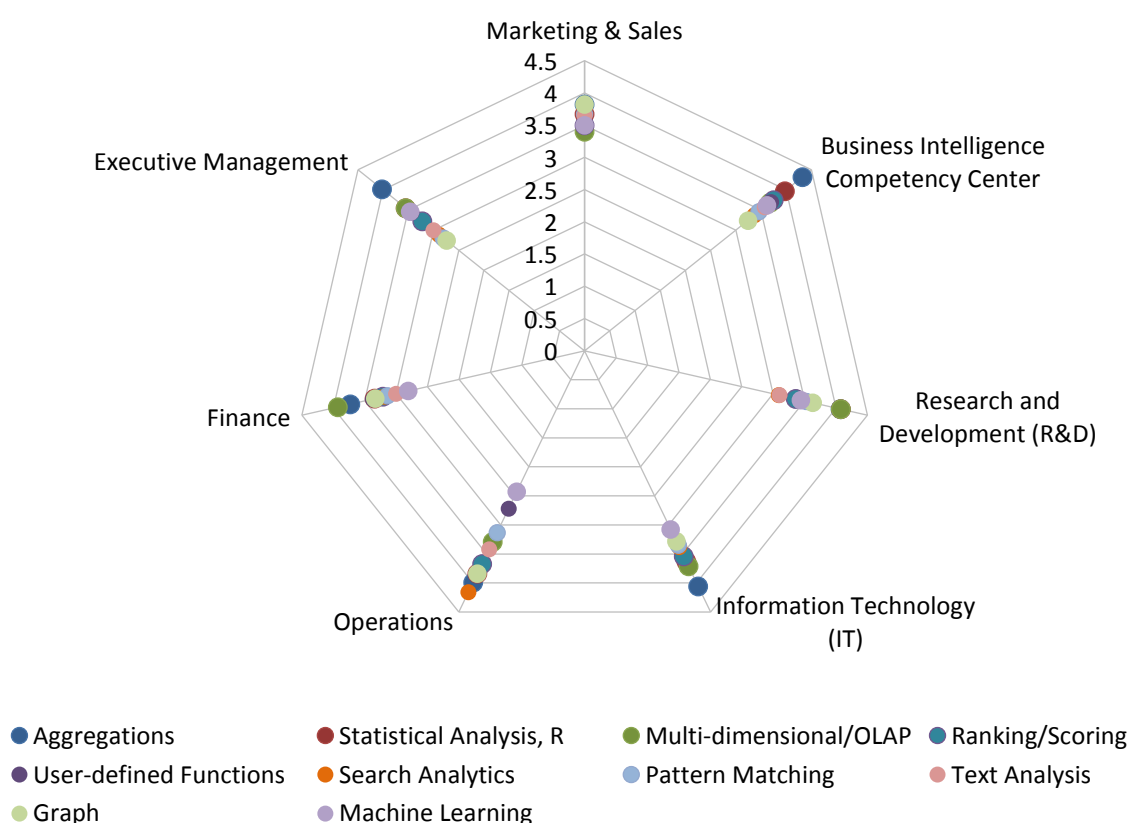


Figure 71 – ADI analytical features by function

When we analyze ADI analytical features by industry, we see a fair amount of variation in priorities. The Financial, Technology, and Telecommunications sectors place high priority on a wide range of analytical features, whereas the Manufacturing sector shows the lowest priority across the range of analytical features (fig. 72). The largest variability of priority across industries is in text and Spark analytic features.

ADI Analytical Features by Industry

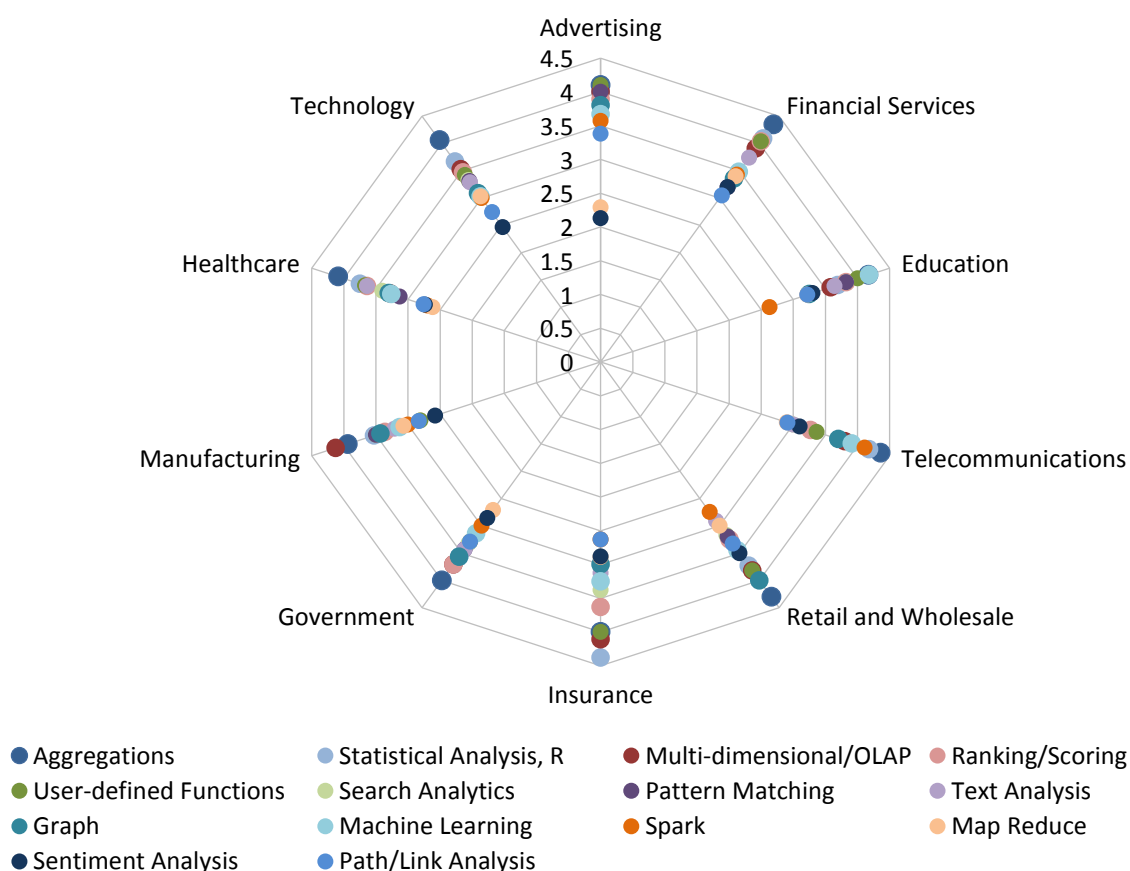


Figure 72 – ADI analytical features by industry

The market has smaller variations when considering ADI analytical feature priorities by organization size. Larger organizations place higher relative priority on more advanced analytical features being a part of the ADI platform (i.e., sentiment analysis, path/link analysis, custom R packages, Spark, MapReduce, pattern matching, machine learning, and search analytics) (fig. 73). Smaller organizations place a lower priority on these advanced analytical features.

ADI Analytical Features by Organization Size

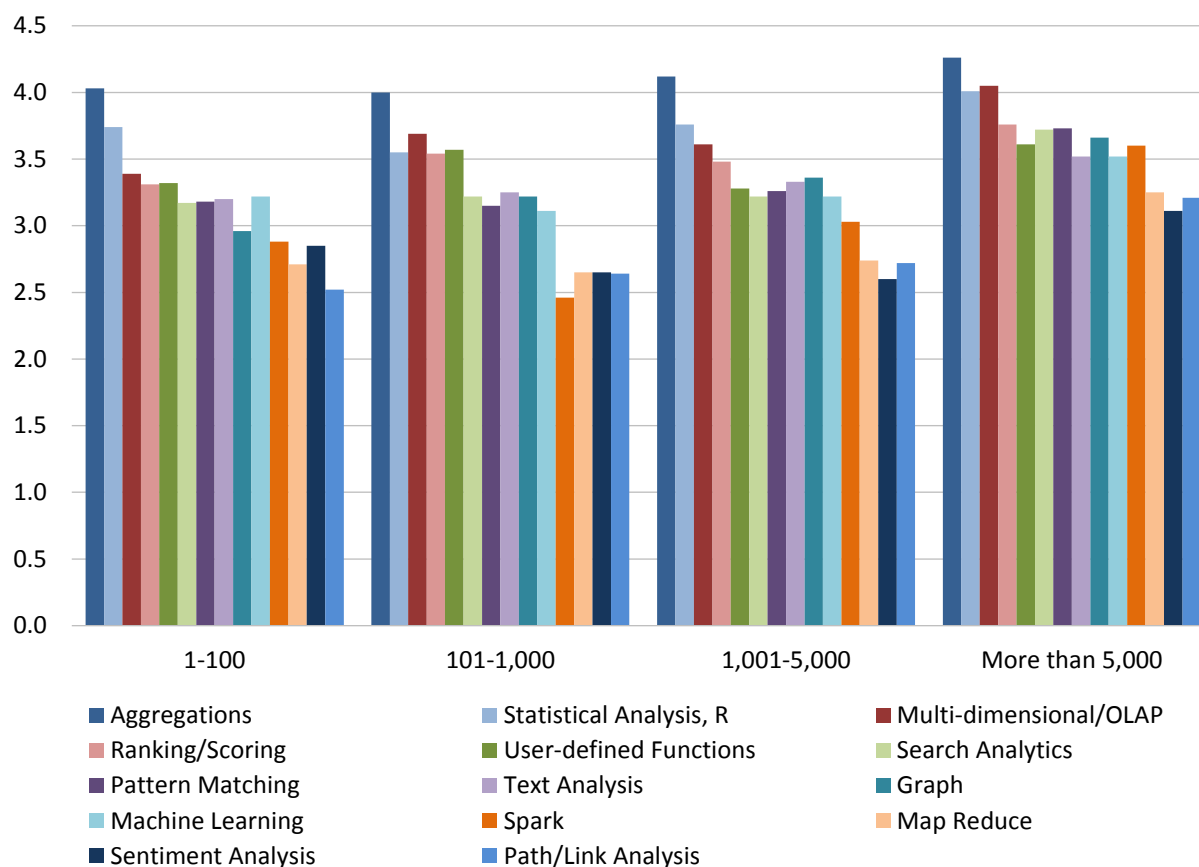


Figure 73 – ADI analytical features by organization size

Vendor Ratings

In this section, we offer ratings of analytical data infrastructure vendors. We rate vendors using 33 different criteria, on a five-point scale for each. Criteria covers sales/acquisition experience (eight criteria), value for price paid (1), quality and usefulness of product (12), quality of technical support (5), quality and value of consulting services (5), whether the vendor is recommended (1), and integrity (1).

As we explore vendor performance in more detail, it is important to understand the scale we use in scoring the industry and vendors:

- 5.0 = Excellent
- 4.0 = Very good
- 3.0 = Adequate
- 2.0 = Poor
- 1.0 = Very poor

Based on our scoring methodology, all vendors perform at a level that is considered more than “adequate” for all criteria categories.

Please note that “average score” is the mathematical mean of all items included in vendor ratings. Each column in the chart represents a scale consisting of varying numbers of items (for example, “sales” is a scale consisting of eight items, while “value for price paid” is one item). As such, each column is weighted differently (based upon the number of items represented and the number of respondents rating those items) in calculating the overall average rating. The average score cannot be calculated by simply averaging across the subscale scores.

Analytical Data Infrastructure Market Models

Starting in 2015, we began using two new models for examining and understanding the analytical data infrastructure market. Using quadrants, we plot aggregated user sentiment into x and y axes.

Customer Experience Model

The Customer Experience Model considers the real-world experience of customers working with ADI products daily (fig. 74). For the x axis, we combine all vendor touchpoints—including the sales and acquisition process (eight measures), technical support (five measures), and consulting services (five measures)—into a single “sales and service” dimension. On the y axis, we plot customer sentiment surrounding product, derived from the 12 product and technology measures used to rank vendors. On the resulting four quadrants, we plot vendors based on these measures.

The upper-right quadrant contains the highest-scoring vendors and is named “overall experience leaders.” Technology leaders (upper-left quadrant) identifies vendors with strong product offerings but relatively lower services scores. Contenders (lower-left quadrant) would benefit from varying degrees of improvement to product, services, or both.

User sentiment surrounding outliers (outside of the four quadrants) suggests that significant improvements are required to product and services.

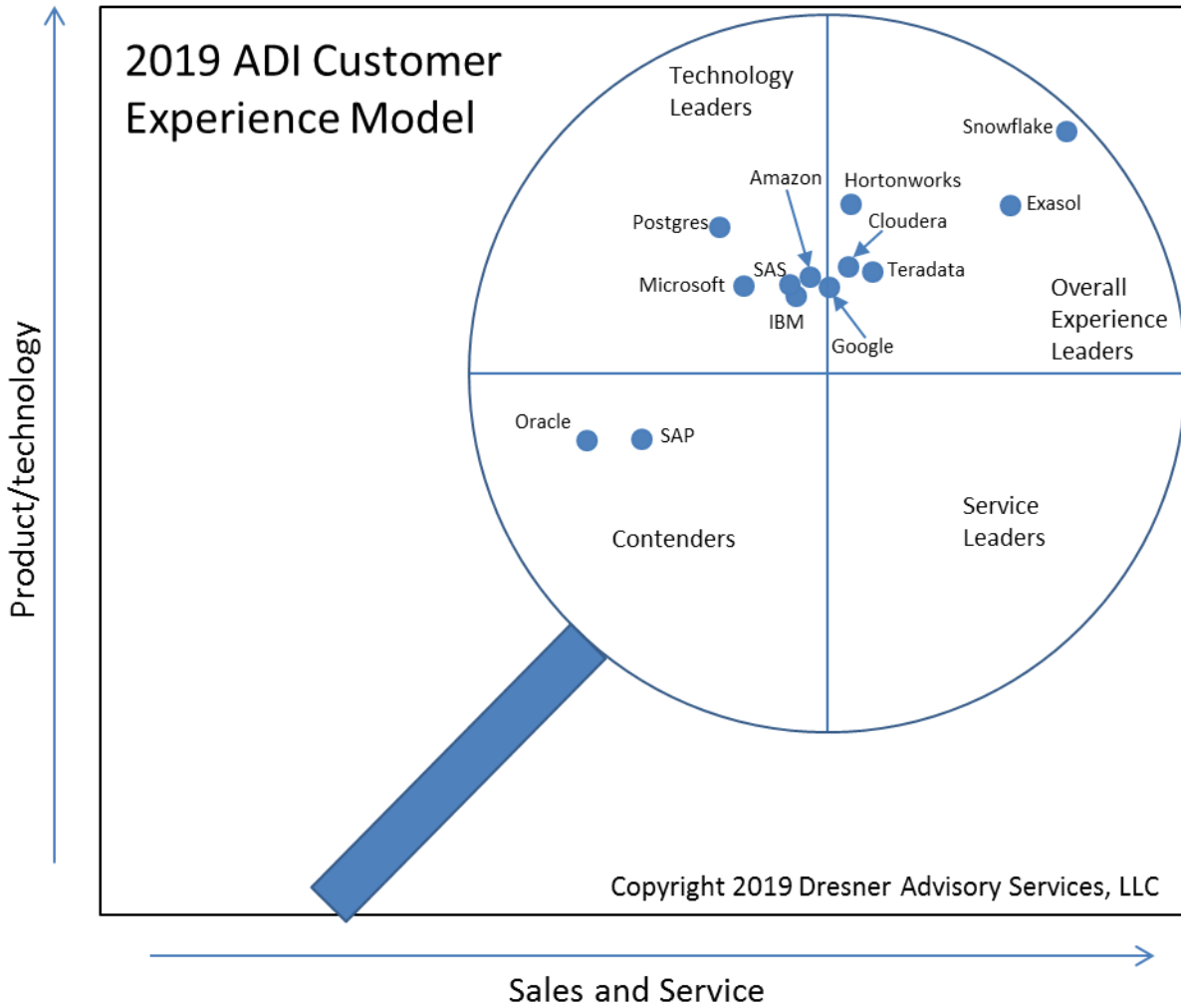


Figure 74 – Customer experience model

Vendor Credibility Model

The Vendor Credibility Model considers how customers “feel” about their vendor (fig. 75). The x axis plots perceived value for the price paid. The y axis combines the integrity and recommend measures, creating a “confidence” dimension. The resulting four quadrants position vendors based on these dimensions.

The upper-right quadrant contains the highest-scoring vendors and is named “credibility leaders.” Trust leaders (upper-left quadrant) identifies vendors with solid perceived confidence but relatively lower value scores. Contenders (lower-left quadrant) would benefit by working to improve customer value, confidence, or both.

User sentiment surrounding outliers (outside of the four quadrants) suggests that significant improvements are required to improve perceived value and confidence.

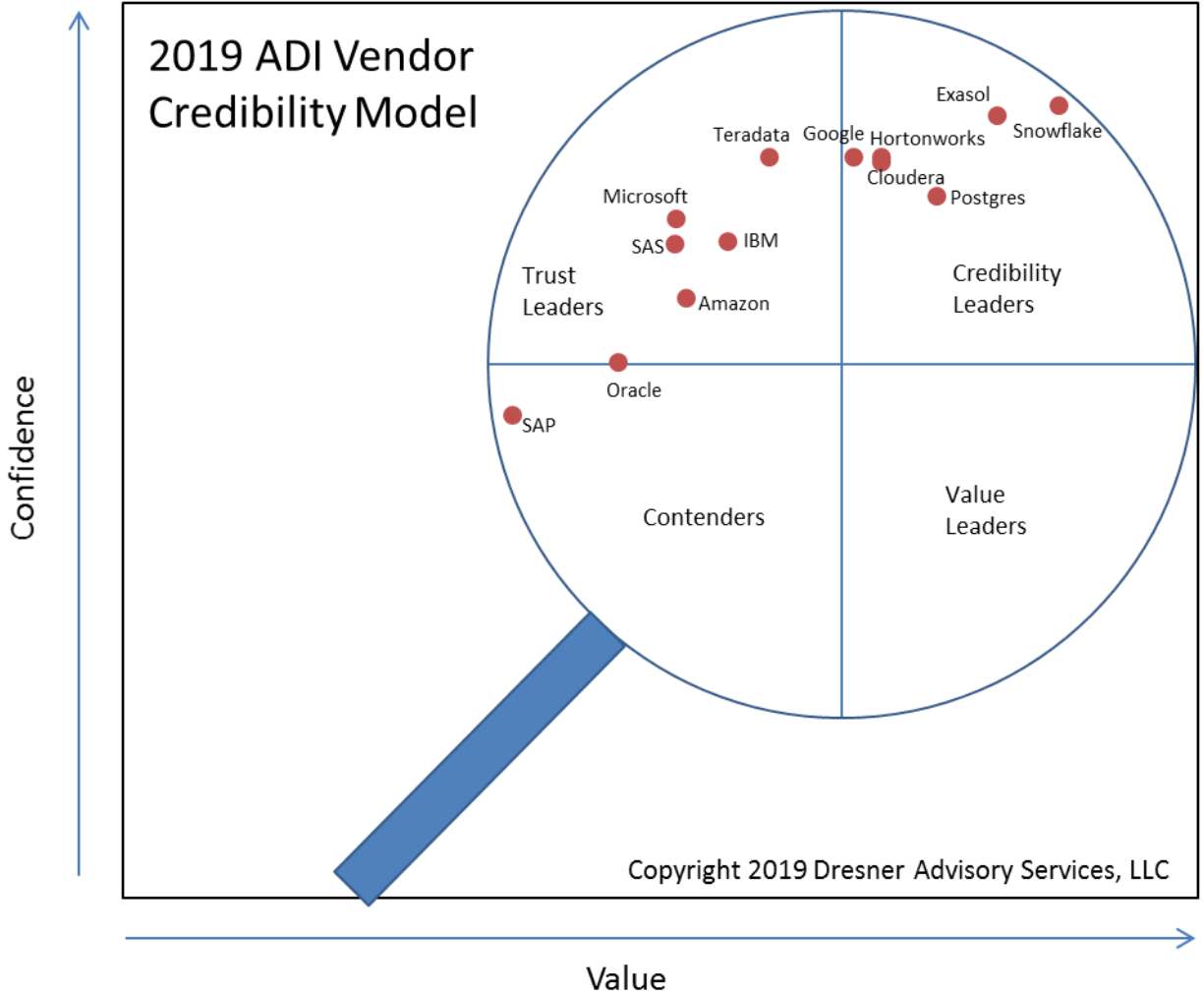


Figure 75 – Vendor credibility model

Detailed Vendor Ratings

In this section, we offer detailed vendor scores. Using our 33-criteria evaluation model (table 1), we compare each vendor's performance to their previous year's performance and to the average for all vendors (all records in the study population).

The detailed criteria are below. We added "clock" position information to assist in locating specific scores.

Table 1- Detailed vendor rating criteria

- Sales/acquisition experience (12 - 2 o'clock) - o Professionalism - o Product knowledge - o Understanding our business/needs - o Responsiveness - o Flexibility/accommodation - o Business practices - o Contractual terms and conditions - o Follow-up after the sale - Value for price (3 o'clock) - Quality and usefulness of product (3 - 7 o'clock) - o Robustness/sophistication of technology - o Completeness of functionality - o Reliability of technology - o Scalability - o Integration of components within product - o Integration with third-party technologies - o Overall usability - o Ease of installation - o Ease of administration	- Quality and usefulness of product (continued) - o Customization and extensibility - o Ease of upgrade/migration to new versions - o Online forums and documentation - Quality of technical support (8 - 9 o'clock) - o Professionalism - o Product knowledge - o Responsiveness - o Continuity of personnel - o Time to resolve problems - Quality and value of consulting services (9 - 10 o'clock) - o Professionalism - o Product knowledge - o Experience - o Continuity - o Value - Integrity (11 o'clock) - Whether vendor is recommended (12 o'clock)
---	--

Amazon Detailed Score

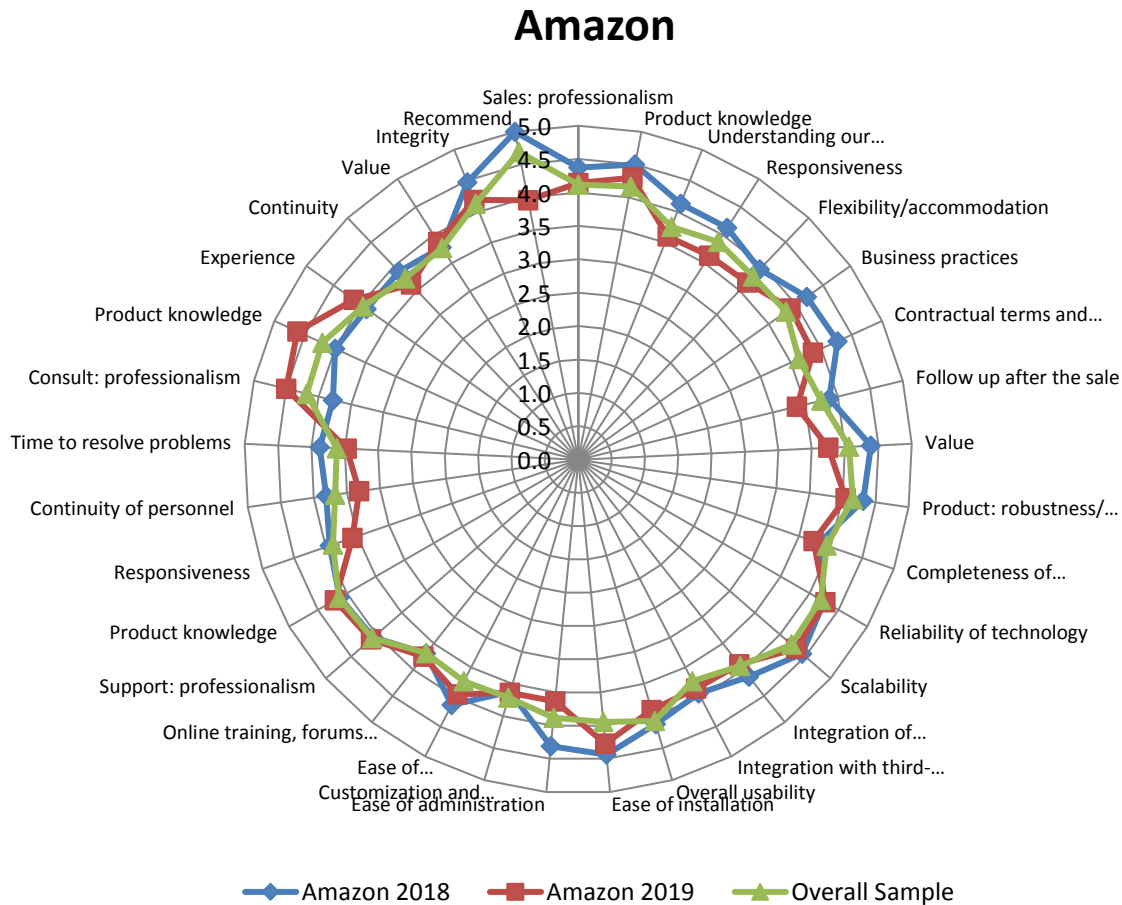


Figure 76 – Amazon detailed score

With scores generally in line with or below the overall sample for all categories, Amazon is a Technology Leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model. While its overall performance dipped versus 2018, Amazon saw an improvement in its consulting score.

Cloudera Detailed Score

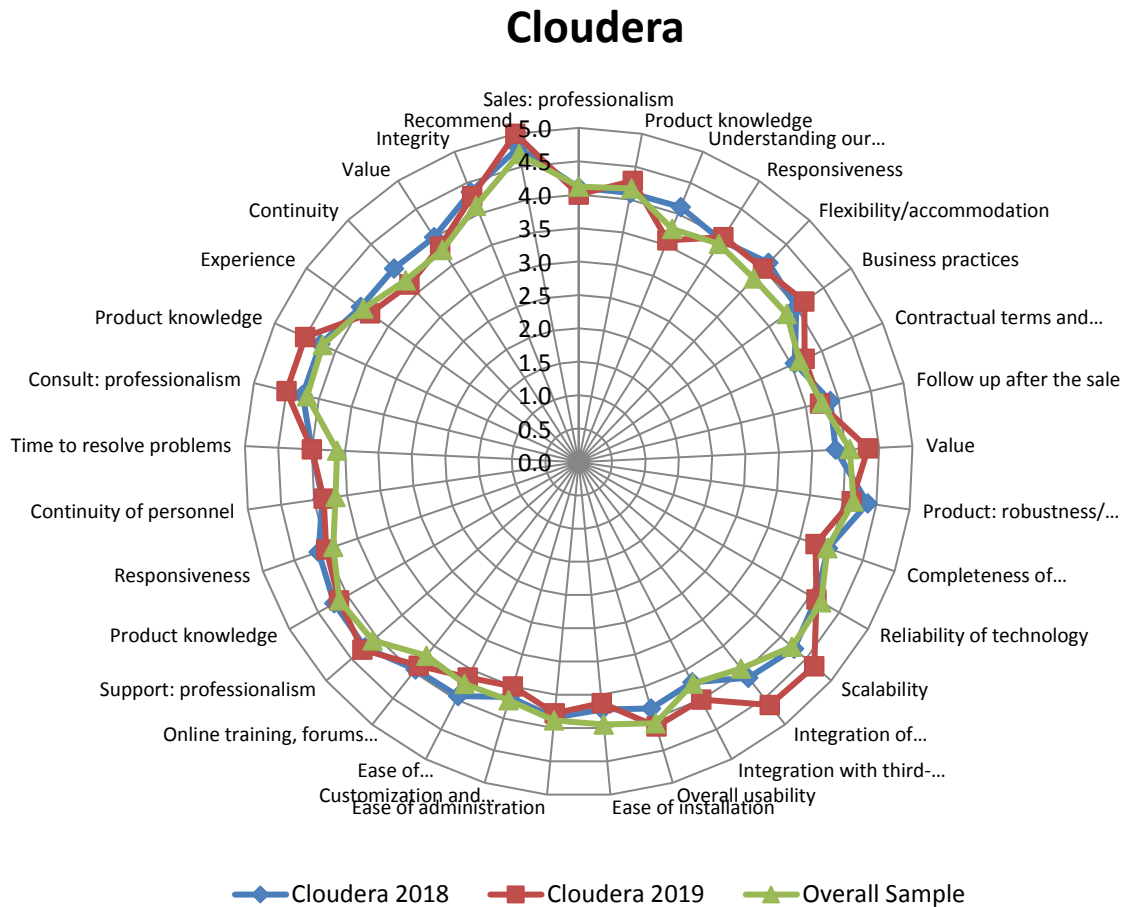


Figure 77 – Cloudera detailed score

With scores generally above or in line with the overall sample, Cloudera is an overall leader in both the Customer Experience and Vendor Credibility models. It saw key improvements versus 2018 and is best in class for product Integration of components within product and has a perfect recommend score.

Exasol Detailed Score

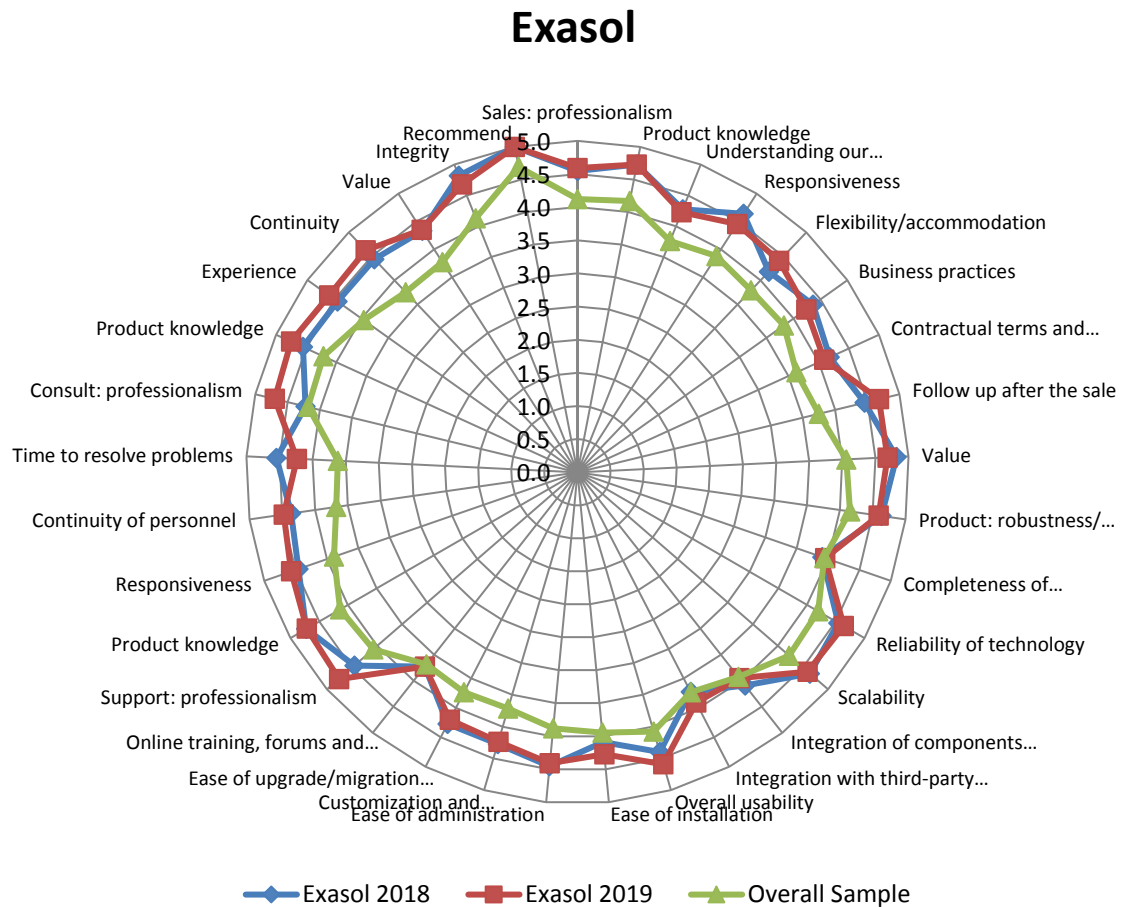


Figure 78 – Exasol detailed score

With scores significantly above the overall sample, Exasol is an Overall Leader in both Customer Experience and Vendor Credibility models. It is best in class for product reliability of technology and customization and extensibility. It has a perfect recommend score.

Google Detailed Score

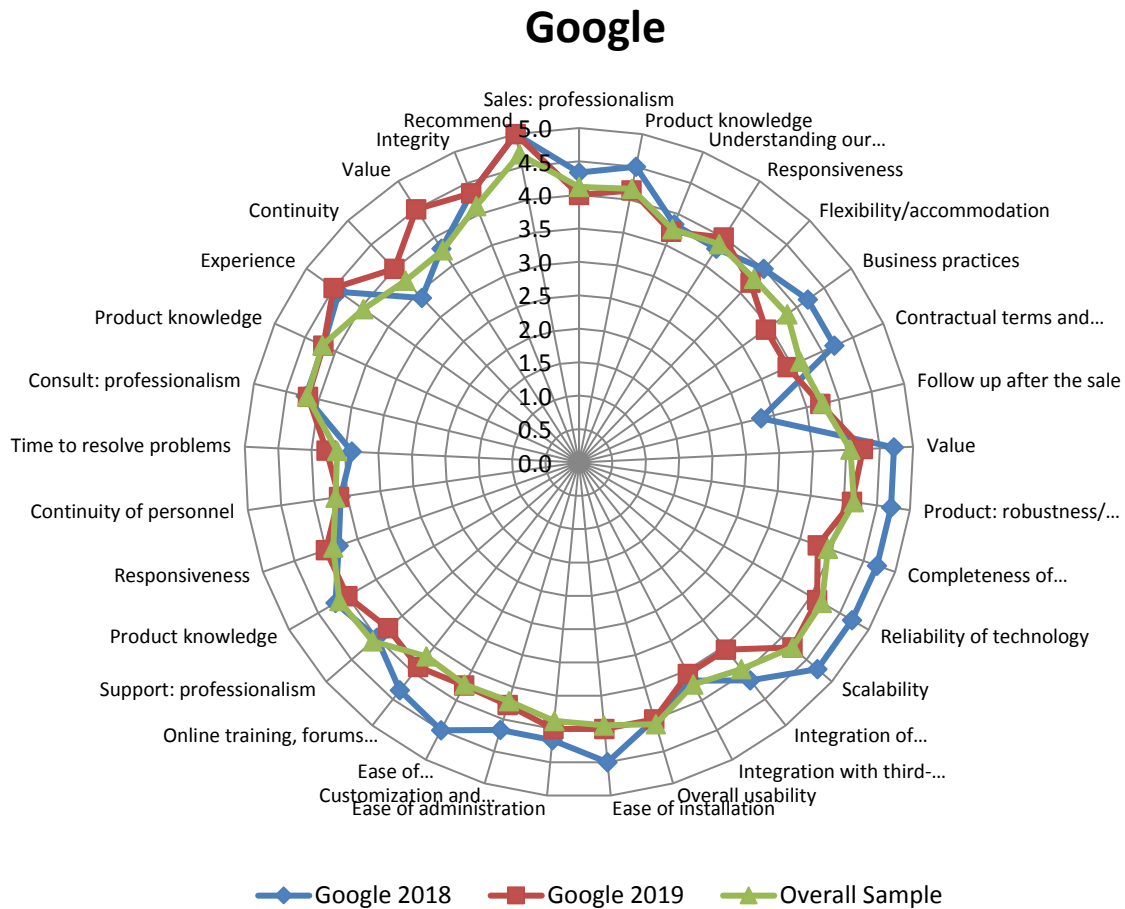


Figure 79 – Google detailed score

With scores generally above or in line with the overall sample, Google is an Overall Leader in both Customer Experience and Vendor Credibility models. It saw key improvements in consulting and technical support scores versus 2018 and maintains a perfect recommend score.

Hortonworks Detailed Score

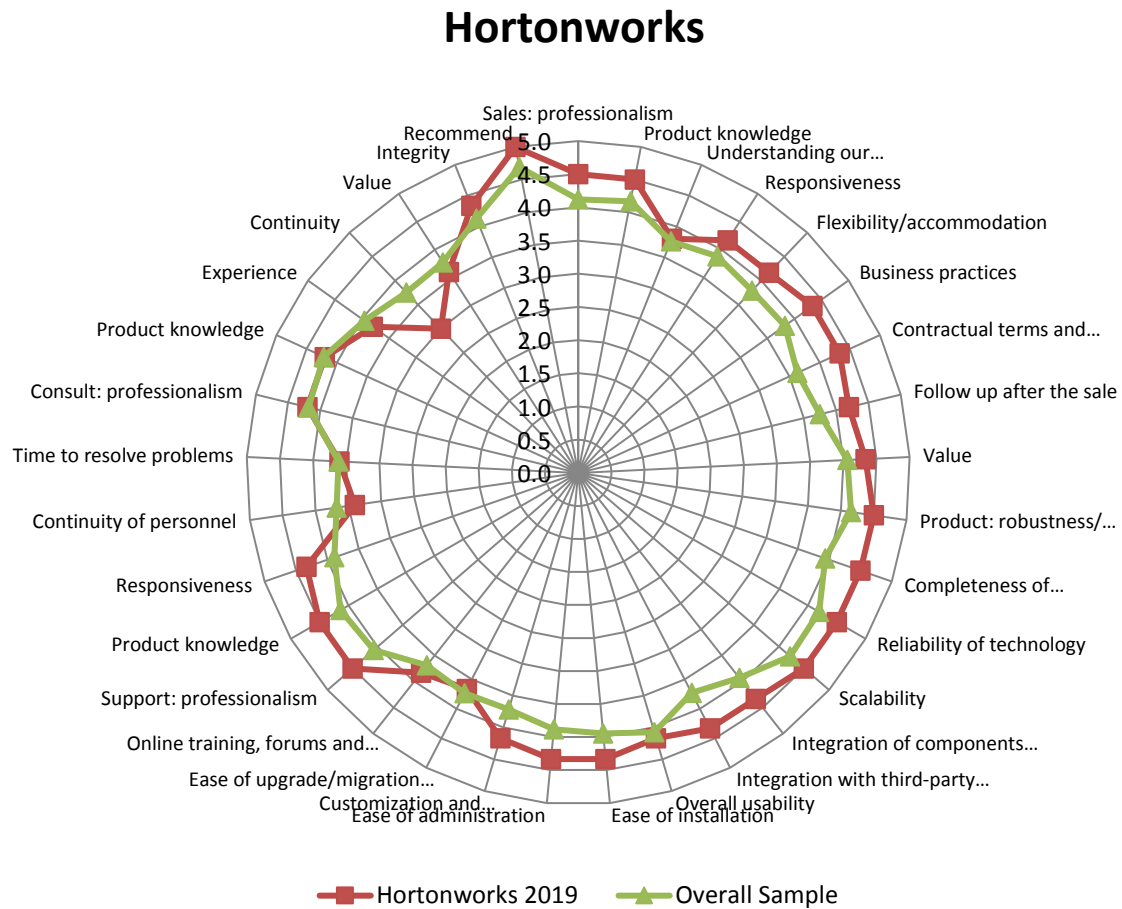


Figure 80 – Hortonworks detailed score

In its first year of inclusion, Hortonworks (now Cloudera) has scores that are generally above the overall sample and is an Overall Leader in both Customer Experience and Vendor Credibility models. It is best in class for product completeness of functionality, and integration with third-party technologies. It has a perfect recommend score.

IBM Detailed Score

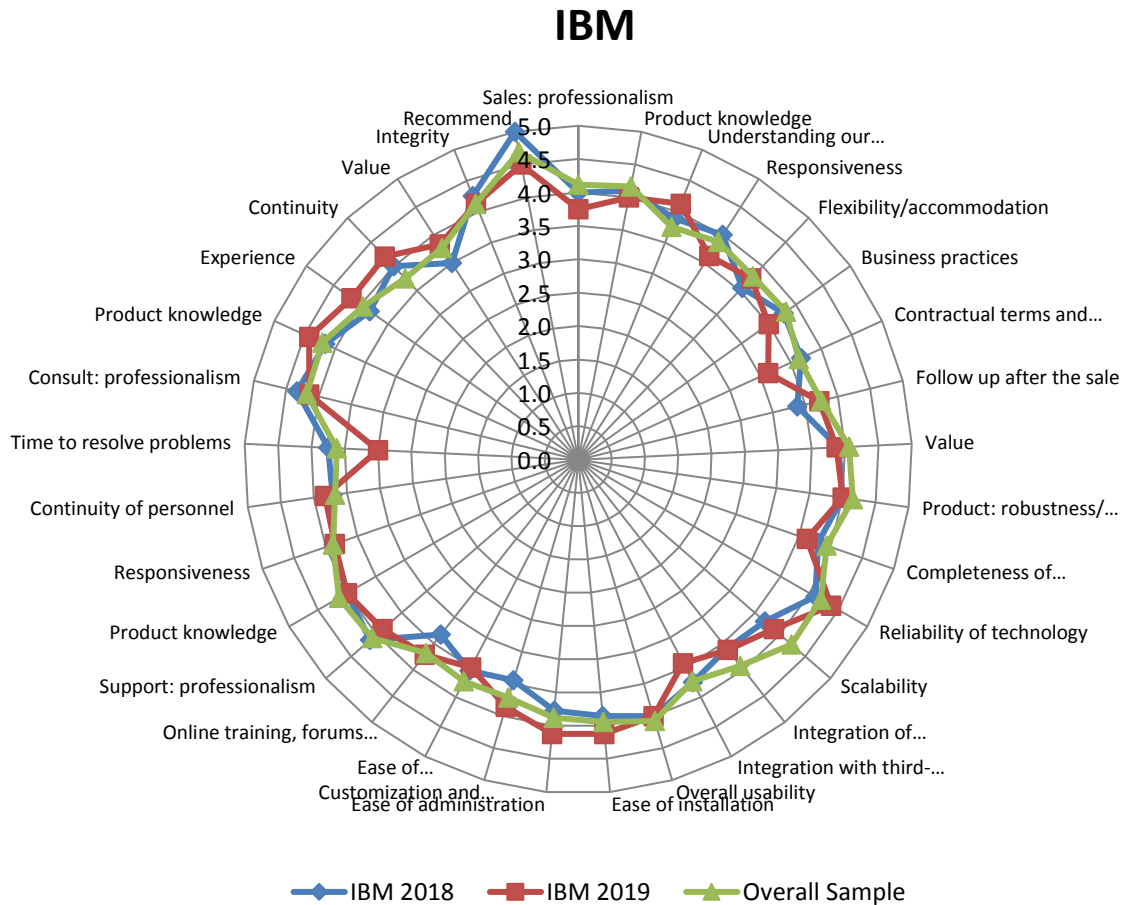


Figure 81 – IBM detailed score

With somewhat improved scores for product and consulting, in 2019, IBM is generally in line with our overall sample. It is a Technology Leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model.

Microsoft Detailed Score

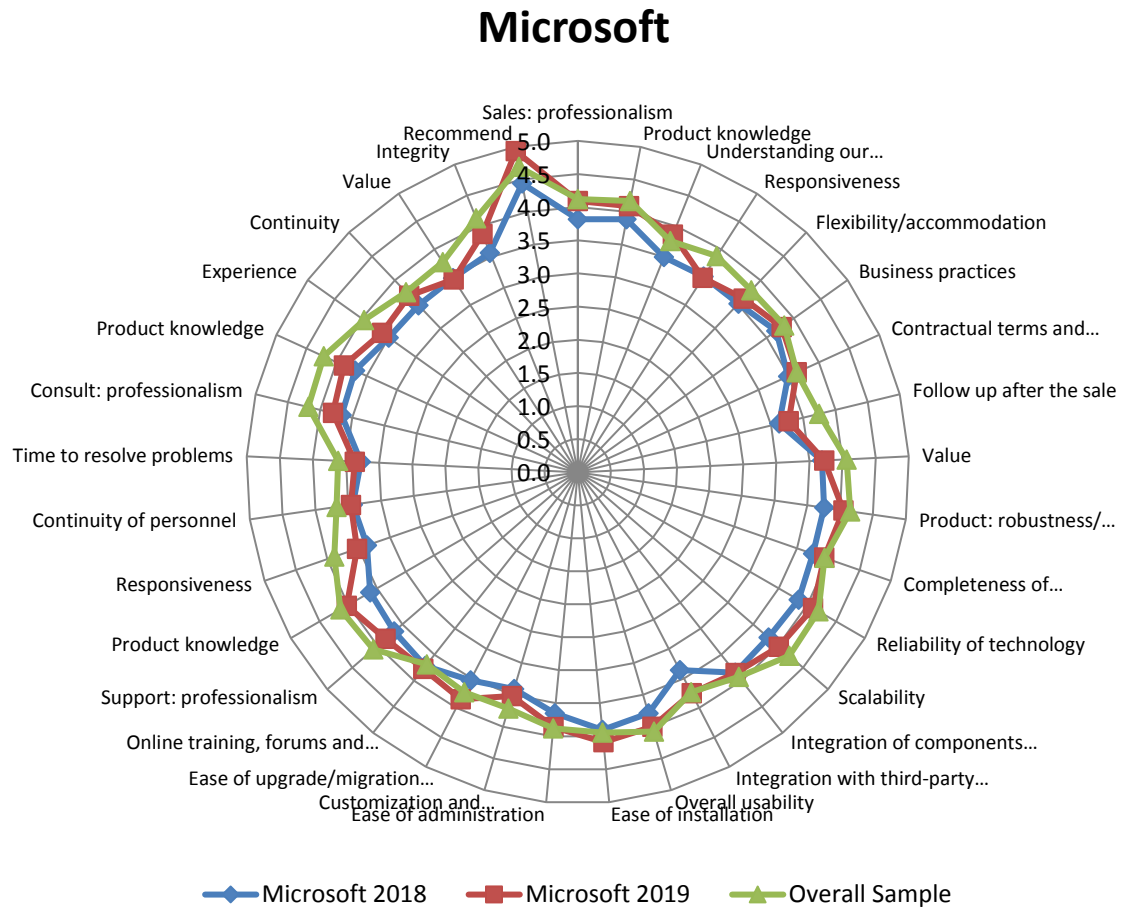


Figure 82 – Microsoft detailed score

In 2019, Microsoft has key improvements in all categories of measurement versus 2018 and is a Technology Leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility model.

Oracle Detailed Score

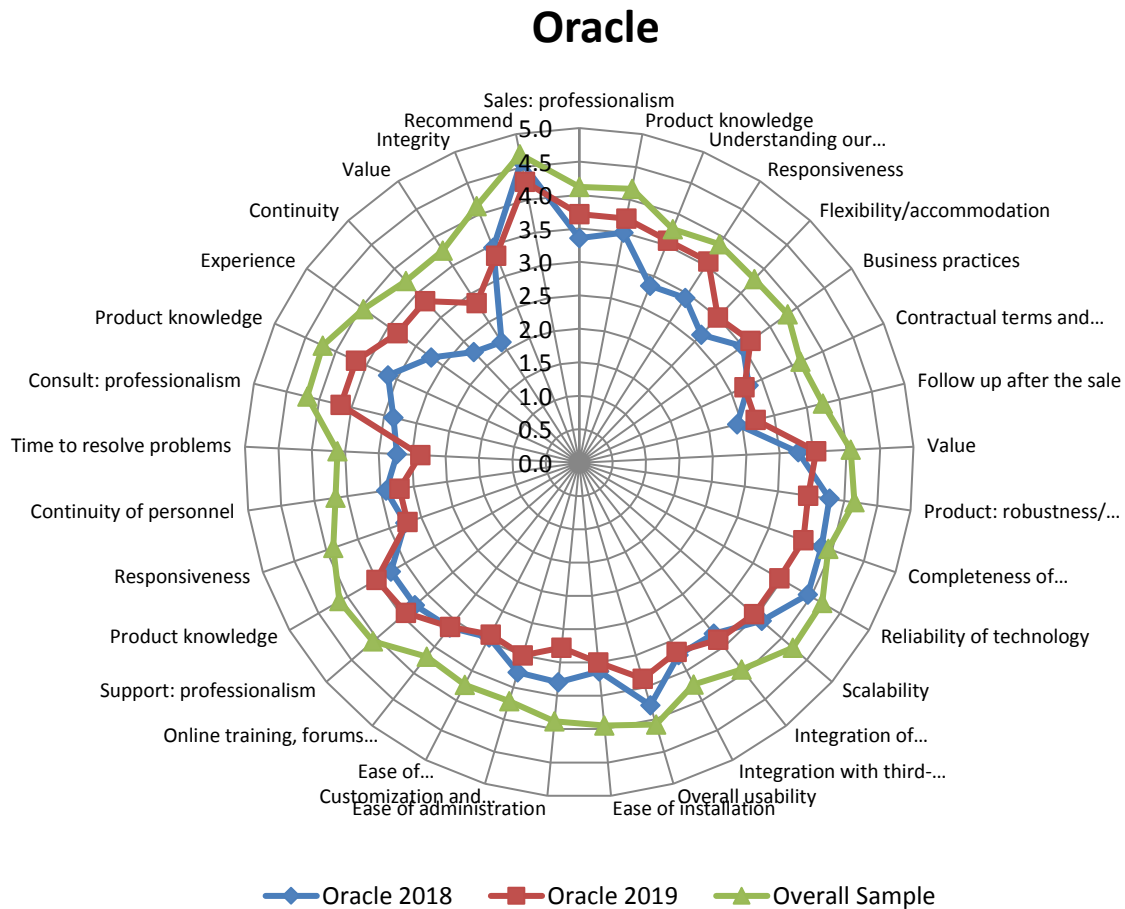


Figure 83 – Oracle detailed score

Although Oracle's scores for 2019 remain below the overall sample, it has key improvements in 2019 for sales, value, and consulting, making it a Contender in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model.

Postgres Detailed Score

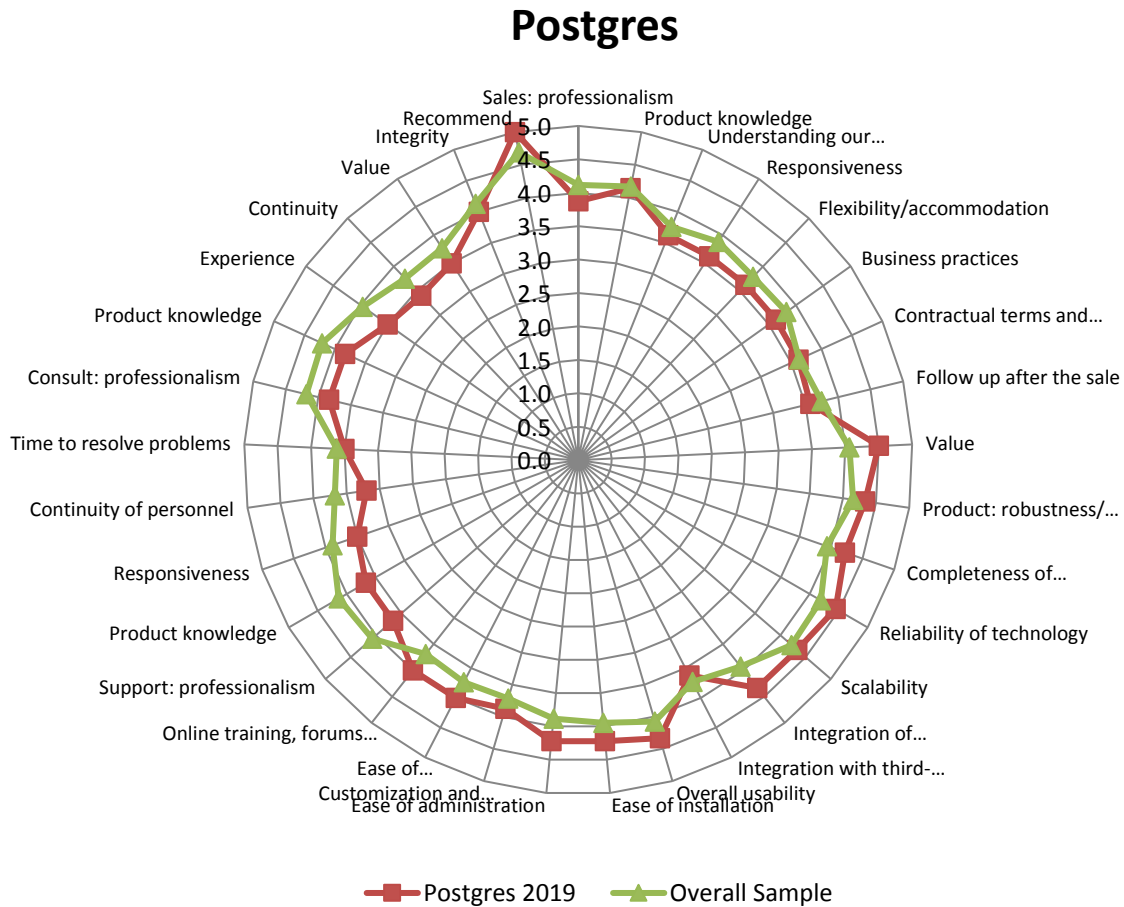


Figure 84 – Postgres detailed score

For 2019, Postgres' scores are generally in line with our overall sample, making it a Technology Leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model. It has a perfect recommend score.

SAP Detailed Score

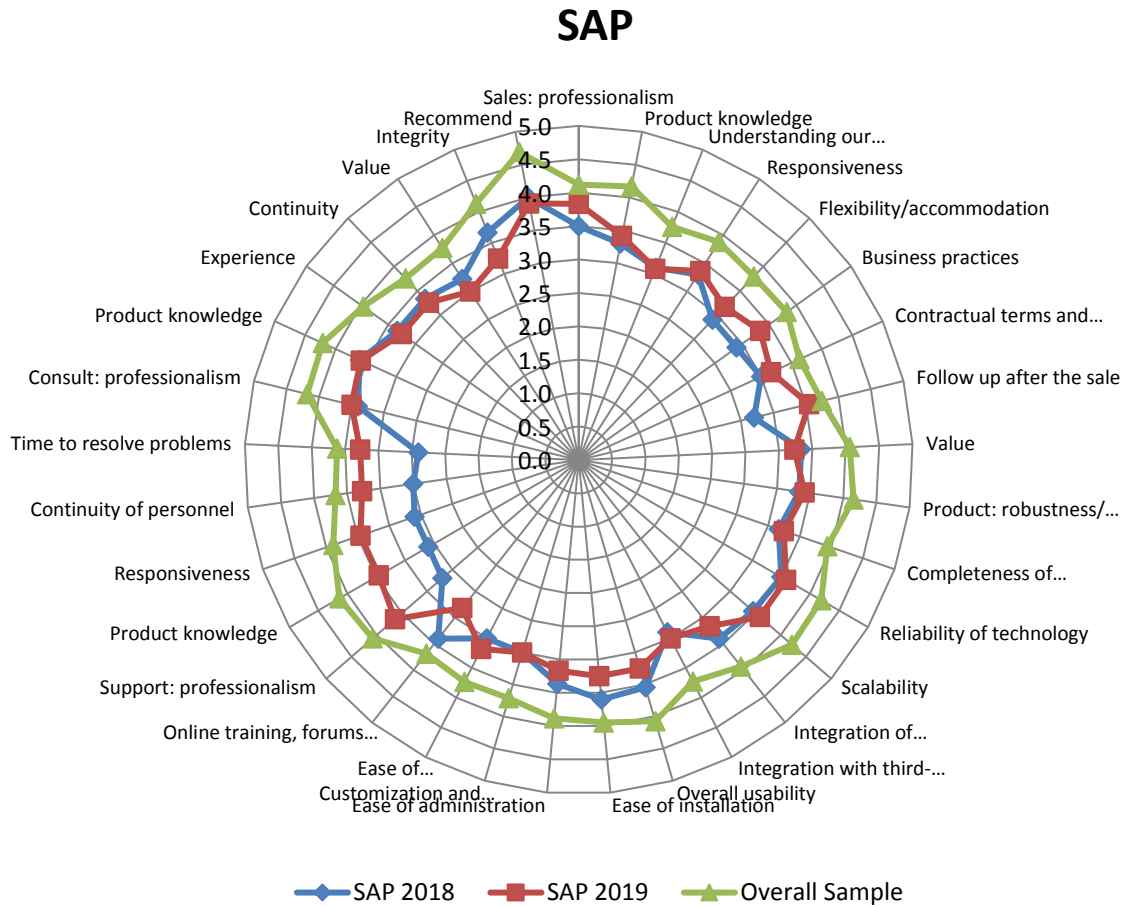


Figure 85 – SAP detailed score

Although still below our overall sample, SAP saw important improvements in its sales, technical support, and overall scores. It is a Contender in both Customer Experience and Vendor Credibility models.

SAS Detailed Score

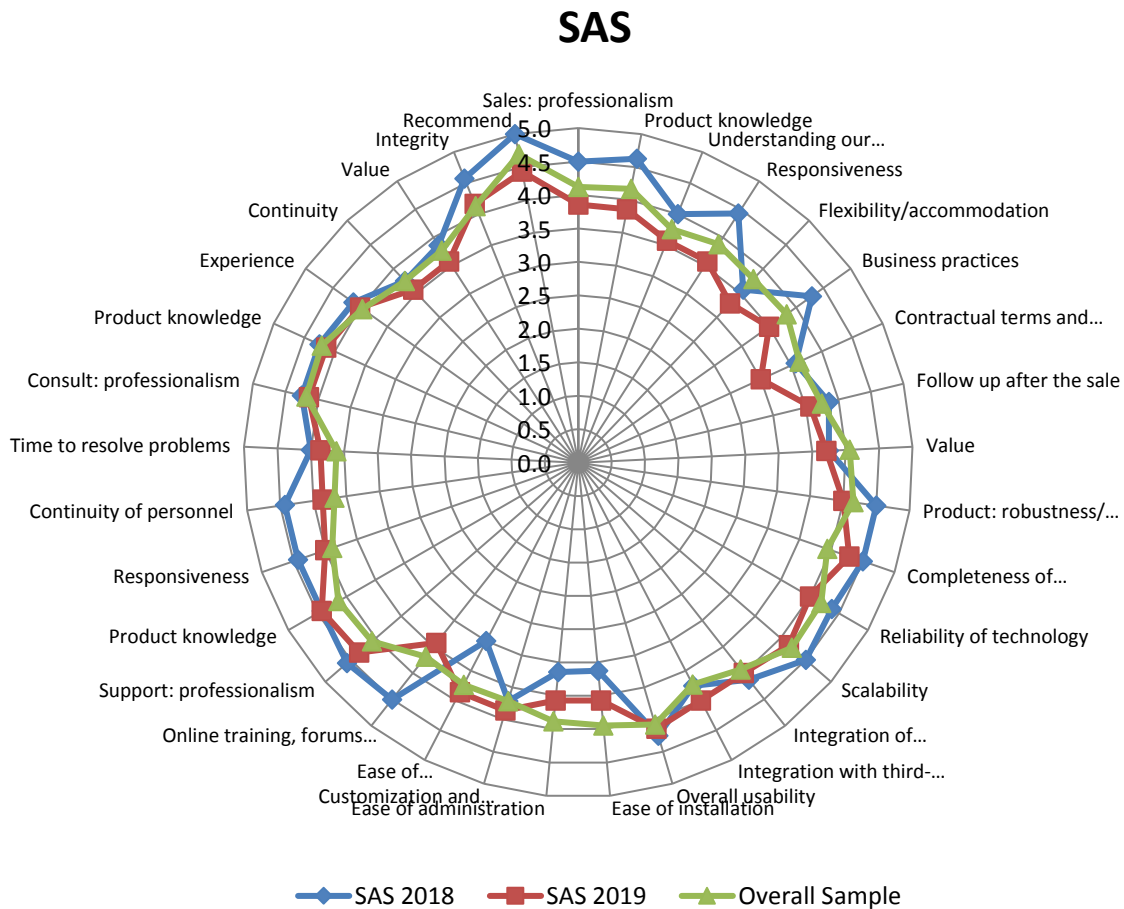


Figure 86 – SAS detailed score

With scores generally in line with or below the overall sample, SAS is a Technology Leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model.

Snowflake Detailed Score

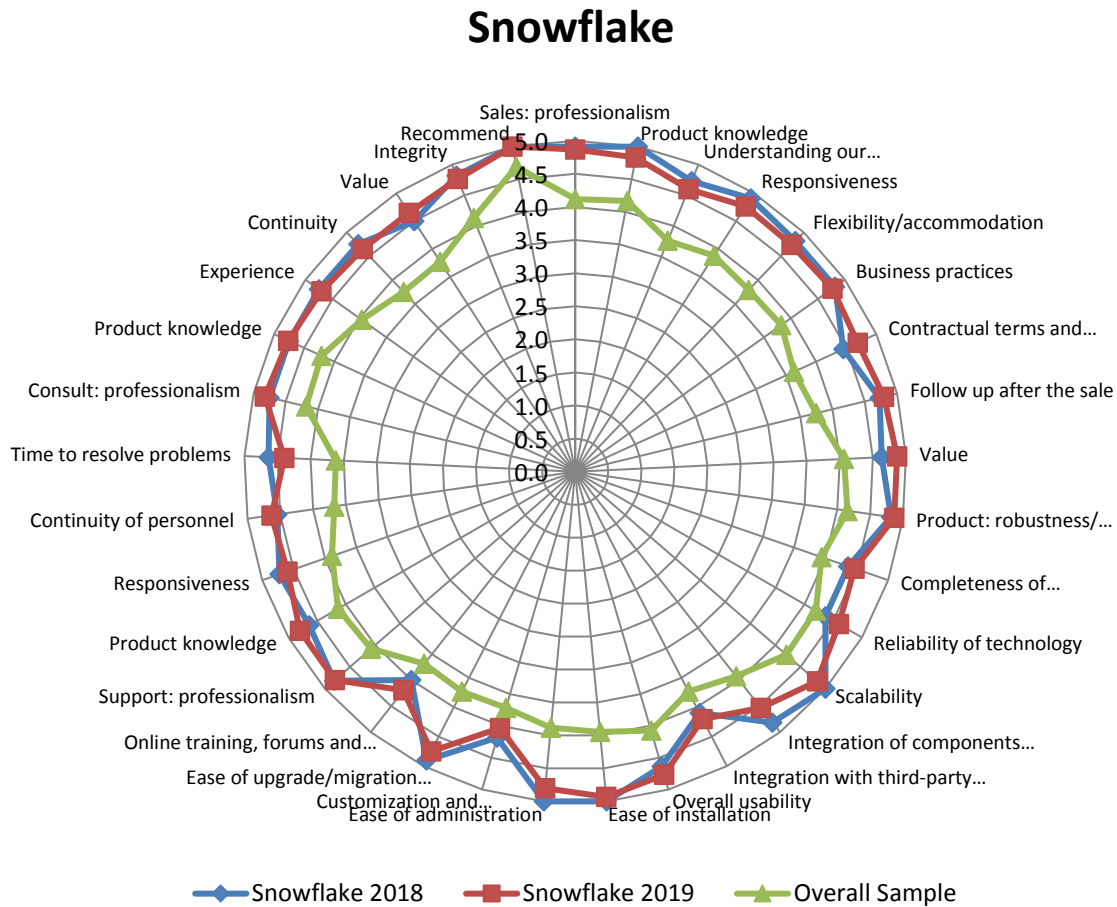


Figure 87 – Snowflake detailed score

In its second year of inclusion, Snowflake maintains exceptionally high scores. With scores well above our overall sample, Snowflake is an Overall Leader in both the Customer Experience and Vendor Credibility models. It is best in class for virtually all measures across sales, value, product/technology, technical support, and consulting. It is also best in class for integrity and has a perfect recommend score.

Teradata Detailed Score

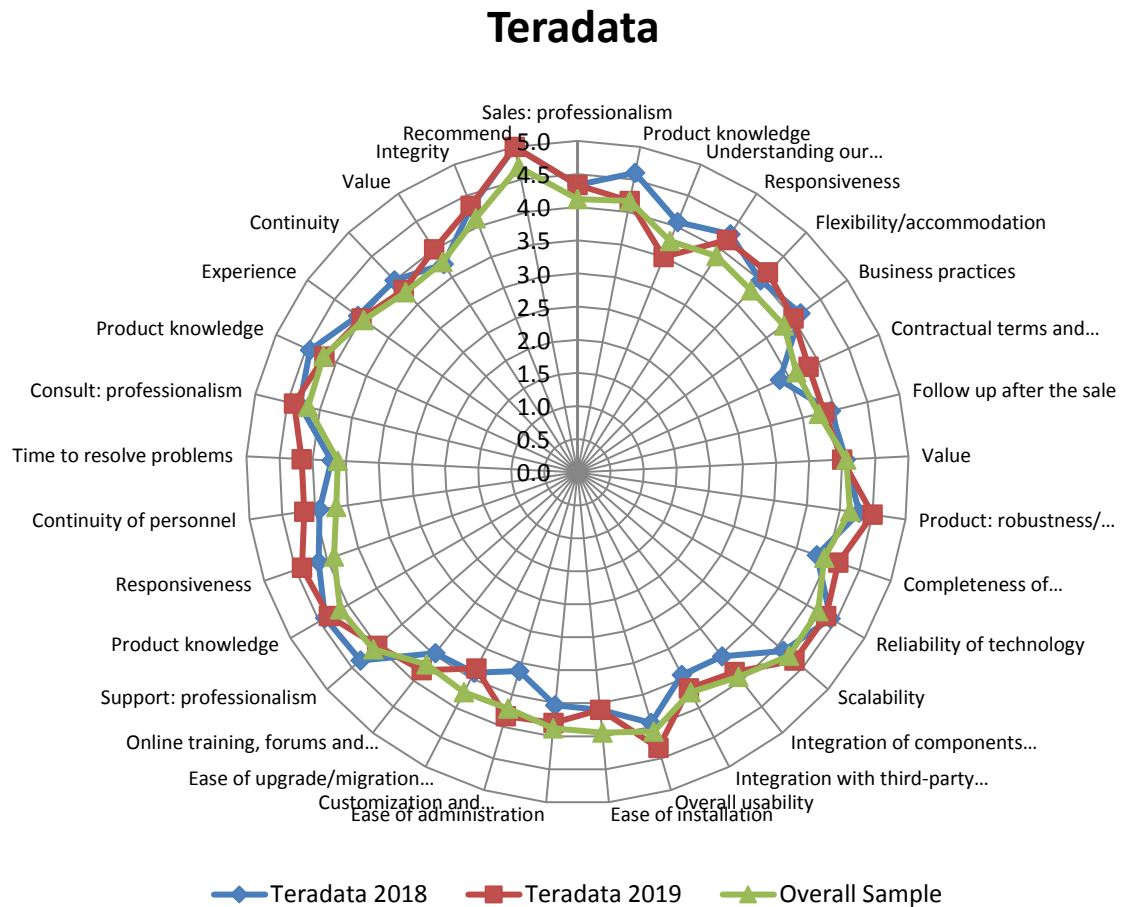


Figure 88 – Teradata detailed score

With scores generally in line with our overall sample, Teradata is an Overall Leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model. It maintains a perfect recommend score.

Other Dresner Advisory Services Research Reports

- [Wisdom of Crowds “Flagship” Business Intelligence Market study](#)
- [Wisdom of Crowds “Flagship” Enterprise Planning](#)
- [Advanced and Predictive Analytics](#)
- [Big Data Analytics](#)
- [Business Intelligence Competency Center](#)
- [Cloud Computing and Business Intelligence](#)
- [Collective Insights®](#)
- [Data Catalog](#)
- [Embedded Business Intelligence](#)
- [End User Data Preparation](#)
- [IoT Intelligence®](#)
- [IT Analytics](#)
- [Location Intelligence](#)
- [Sales Planning](#)
- [Small and Mid-Sized Enterprise Business Intelligence](#)
- [Small and Mid-Sized Planning](#)

Appendix: Analytical Data Infrastructure 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: _____

4) 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?

- ☐ 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 & 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

Please prioritize the following use cases for ADI*

	Critical	Very Important	Important	Some-what Important	Not Important	Don't Know
Business user reporting and dashboards	()	()	()	()	()	()
Business user discovery and exploration	()	()	()	()	()	()
Data science (advanced and predictive analytics/data mining)	()	()	()	()	()	()
Embedded analytics within business applications (high volume, low latency)	()	()	()	()	()	()

What are your overall priorities for using/selecting an analytical data infrastructure product?*

	Cri- ti- cal	Very Im- portant	Im- portant	Some- what Im- portant	Not Im- portant	Don' t Kno w
Performance	()	()	()	()	()	()
Scalability	()	()	()	()	()	()
Price	()	()	()	()	()	()
Security	()	()	()	()	()	()
Analytical Features	()	()	()	()	()	()
Corporate Stand- ard	()	()	()	()	()	()
Agility: Add new data, change data models, change analytics (mod- els/algorithms)	()	()	()	()	()	()
Usability: Ease of development and administration	()	()	()	()	()	()
Adaptability: Ability to integrate with existing business applications, data infrastructure, and processes	()	()	()	()	()	()
Compliance or regulatory require- ments	()	()	()	()	()	()

Deployment and licensing priorities

Please prioritize your deployment preferences for data infrastructure centers

	Critical	Very Important	Important	Some-what Important	Not Important	Don't Know
As-a-cloud service	()	()	()	()	()	()
On-premises software	()	()	()	()	()	()
Hybrid	()	()	()	()	()	()
Cross data center integration and management capabilities	()	()	()	()	()	()

Please prioritize your preferences for ADI licensing

	Critical	Very Important	Important	Some-what Important	Not Important	Don't Know
User	()	()	()	()	()	()
Subscription	()	()	()	()	()	()
Concurrent use	()	()	()	()	()	()

2019 Analytical Data Infrastructure Market Study

Open source: community and commercial	()	()	()	()	()	()
Data volume	()	()	()	()	()	()
Computing resources consumed (i.e., CPU/core or query capacity based).	()	()	()	()	()	()

Please prioritize the following ADI development and deployment features

	Critical	Very Important	Important	Some-what Important	Not Important	Don't Know
Data life cycle management (lineage, impact metadata, governance/collaboration controls)	()	()	()	()	()	()
Application development tools (e.g., IDE, SDK)	()	()	()	()	()	()

2019 Analytical Data Infrastructure Market Study

Support for programming languages/scripting (e.g., Python, Perl, Ruby)	()	()	()	()	()	()
Pre-built data models (e.g., retail, banking, utilities, etc.)	()	()	()	()	()	()
Support for audit (e.g., analyze usage)	()	()	()	()	()	()
Capacity planning and performance-management tools	()	()	()	()	()	()
Scale up and scale out (e.g., clustering, load balancing, high availability)	()	()	()	()	()	()
Multi-tenancy support	()	()	()	()	()	()

Data Source Access, Preparation and Load

Please prioritize your preferences for the types of data that you will load into your

 analytical data infrastructure (ADI)

	Critical	Very Important	Important	Some-what Important	Not Important	Don't Know
Images, video	()	()	()	()	()	()
Text	()	()	()	()	()	()
Machine and events / log data	()	()	()	()	()	()
Transactional data	()	()	()	()	()	()
Metadata	()	()	()	()	()	()
Excel/CSV	()	()	()	()	()	()

Other sources?

Please prioritize your preferences for loading and preparing data

	Critical	Very Important	Important	Some-what Important	Not Important	Don't Know
ELT/ETL - Bulk load	()	()	()	()	()	()
In-	()	()	()	()	()	()

2019 Analytical Data Infrastructure Market Study

serts/updates/updates						
Real-time/streaming, trickle, increments/change capture	()	()	()	()	()	()
Support for Apache big data services (Flume, Spark)	()	()	()	()	()	()
End-User Data Prep: Data quality (e.g., profiling, deduping, etc), transformations, re-formatting of data	()	()	()	()	()	()
Metadata import and management for lineage and impact analysis.	()	()	()	()	()	()
Standards supported for accessing data sources (e.g., JDBC, ODBC, Web/RESTful Services)	()	()	()	()	()	()

Please prioritize your preferences for data model, and management and organization of data

	Critical	Very Important	Important	Some-what Important	Not Important	Don't Know
Hierarchical files (e.g., Hadoop HDFS)	()	()	()	()	()	()
SQL data	()	()	()	()	()	()
Row format	()	()	()	()	()	()
Columnar format	()	()	()	()	()	()
Hybrid row/column	()	()	()	()	()	()
In memory	()	()	()	()	()	()
Non-SQL (machine, text, audio, video, JSON)	()	()	()	()	()	()

What interface(s) will your end user BI/analytics tool use to access the data stored in your analytical data infrastructure?

	Critical	Very Important	Important	Some-what Important	Not Important	Don't Know
JDBC	()	()	()	()	()	()
ODBC	()	()	()	()	()	()
MDX	()	()	()	()	()	()
Apache Hive/Impala	()	()	()	()	()	()
Web/RESTful	()	()	()	()	()	()
XML	()	()	()	()	()	()
Web Services	()	()	()	()	()	()
Excel/CSV	()	()	()	()	()	()
Apache Avro	()	()	()	()	()	()
Apache Parquet	()	()	()	()	()	()

Which analytical features should an analytical data infrastructure (ADI) product support?

	Cri ti- cal	Very Im- portant	Im- portant	Some- what Im- portant	Not Im- portant	Don' t Kno w
Multi-dimensional/OLAP	()	()	()	()	()	()
Map Reduce	()	()	()	()	()	()
Path/link analysis	()	()	()	()	()	()
Pattern matching	()	()	()	()	()	()
Aggregations	()	()	()	()	()	()
Statistical analysis, R	()	()	()	()	()	()
Spark	()	()	()	()	()	()
Graph	()	()	()	()	()	()
Search analytics	()	()	()	()	()	()
Text analysis	()	()	()	()	()	()
Ranking/scoring	()	()	()	()	()	()
User-defined functions	()	()	()	()	()	()
Machine learning	()	()	()	()	()	()
Sentiment analysis	()	()	()	()	()	()

Analytical Data Infrastructure Vendor Ratings

Please select an ADI vendor to rate

- ☐ 1010data
- ☐ Actian
- ☐ Amazon (including Redshift)
- ☐ AtScale
- ☐ Attivio
- ☐ Cloudera
- ☐ Couchbase
- ☐ DataStax
- ☐ EMC Greenplum
- ☐ Exasol
- ☐ Firebird
- ☐ Google (including BigQuery)
- ☐ Hortonworks
- ☐ HP (including Vertica)
- ☐ IBM
- ☐ Infobright
- ☐ JethroData
- ☐ Kognitio
- ☐ MapR
- ☐ MariaDB
- ☐ MarkLogic
- ☐ MemSQL
- ☐ Microsoft (including SQL Server)
- ☐ MongoDB
- ☐ Neo4j
- ☐ Oracle
- ☐ Pivotal
- ☐ PostgreSQL
- ☐ SAP (including HANA)
- ☐ SAS
- ☐ SnappyData
- ☐ Snowflake Computing
- ☐ Splunk
- ☐ Teradata (including Aster)
- ☐ Other - Write In: _____

Please specify the product name and version for the selected vendor

How long has this product been in use?

- ☐ Less than 1 year
- ☐ 1 - 2 years
- ☐ 3 - 5 years
- ☐ 6 - 10 years
- ☐ More than 10 years

How many users currently use this product?

- ☐ 1-100
- ☐ 101-500
- ☐ 501-1,000
- ☐ 1,001-10,000
- ☐ More than 10,000

How would you characterize the sales/acquisition experience with this vendor?

	Excel- lent	Very Goo d	Ade- quate	Poo r	Ver y Poo r	Don't Kno w
Professionalism	☐	☐	☐	☐	☐	☐
Product Knowledge	☐	☐	☐	☐	☐	☐
Understanding Our Business/Needs	☐	☐	☐	☐	☐	☐
Responsiveness	☐	☐	☐	☐	☐	☐
Flexibil- ity/Accommodation	☐	☐	☐	☐	☐	☐
Business Practices	☐	☐	☐	☐	☐	☐
Contractual Terms and Conditions	☐	☐	☐	☐	☐	☐
Follow-up After the Sale	☐	☐	☐	☐	☐	☐

How would you characterize the value for the price paid?

- ☐ Great Value (well exceeded expectations)
- ☐ Good Value (somewhat exceeded expectations)
- ☐ Average Value (met expectations)
- ☐ Poor Value (fell short of expectations)
- ☐ Very Poor Value (fell far short of expectations)

How would you characterize the quality and usefulness of the product?

	Ex- cel- lent	Very Goo d	Ade- quat e	Poor	Ver y Poo r	Don't Kno w
Robustness/Sophistication of Technology	☐	☐	☐	☐	☐	☐
Completeness of Functionality	☐	☐	☐	☐	☐	☐
Reliability of Technology	☐	☐	☐	☐	☐	☐
Scalability	☐	☐	☐	☐	☐	☐
Integration of Components within Product	☐	☐	☐	☐	☐	☐
Integration with Third-Party Technologies	☐	☐	☐	☐	☐	☐
Overall Usability	☐	☐	☐	☐	☐	☐
Ease of Installation	☐	☐	☐	☐	☐	☐
Ease of Administration	☐	☐	☐	☐	☐	☐
Customization and Extensibility	☐	☐	☐	☐	☐	☐
Ease of Upgrade/Migration to New Versions	☐	☐	☐	☐	☐	☐
Online Training, Forums and Documentation	☐	☐	☐	☐	☐	☐

How would you characterize the vendor's technical support?

	Excellent	Very Good	Adequate	Poor	Very Poor	Don't Know
Professionalism	()	()	()	()	()	()
Product Knowledge	()	()	()	()	()	()
Responsiveness	()	()	()	()	()	()
Continuity of Personnel	()	()	()	()	()	()
Time to Resolve Problems	()	()	()	()	()	()

How would you characterize the vendor's consulting services?

	Excellent	Very Good	Adequate	Poor	Very Poor	Don't Know
Professionalism	()	()	()	()	()	()
Product Knowledge	()	()	()	()	()	()
Responsiveness	()	()	()	()	()	()
Continuity of Personnel	()	()	()	()	()	()
Time to Resolve Problems	()	()	()	()	()	()

How would you rate the "integrity" (i.e., truthfulness, honesty) of this vendor?

- ☐ Excellent
- ☐ Very Good
- ☐ Adequate
- ☐ Poor
- ☐ Very Poor
- ☐ Don't Know

Did this vendor's overall performance improve, remains the same or decline from last year?

- ☐ Improved
- ☐ Stayed the Same
- ☐ Declined

Would you recommend this vendor/product?

- ☐ Yes, I would recommend this vendor/product
- ☐ No, I would NOT recommend this vendor/product

Please enter any additional comments regarding this vendor and/or its products
