January 28, 2021

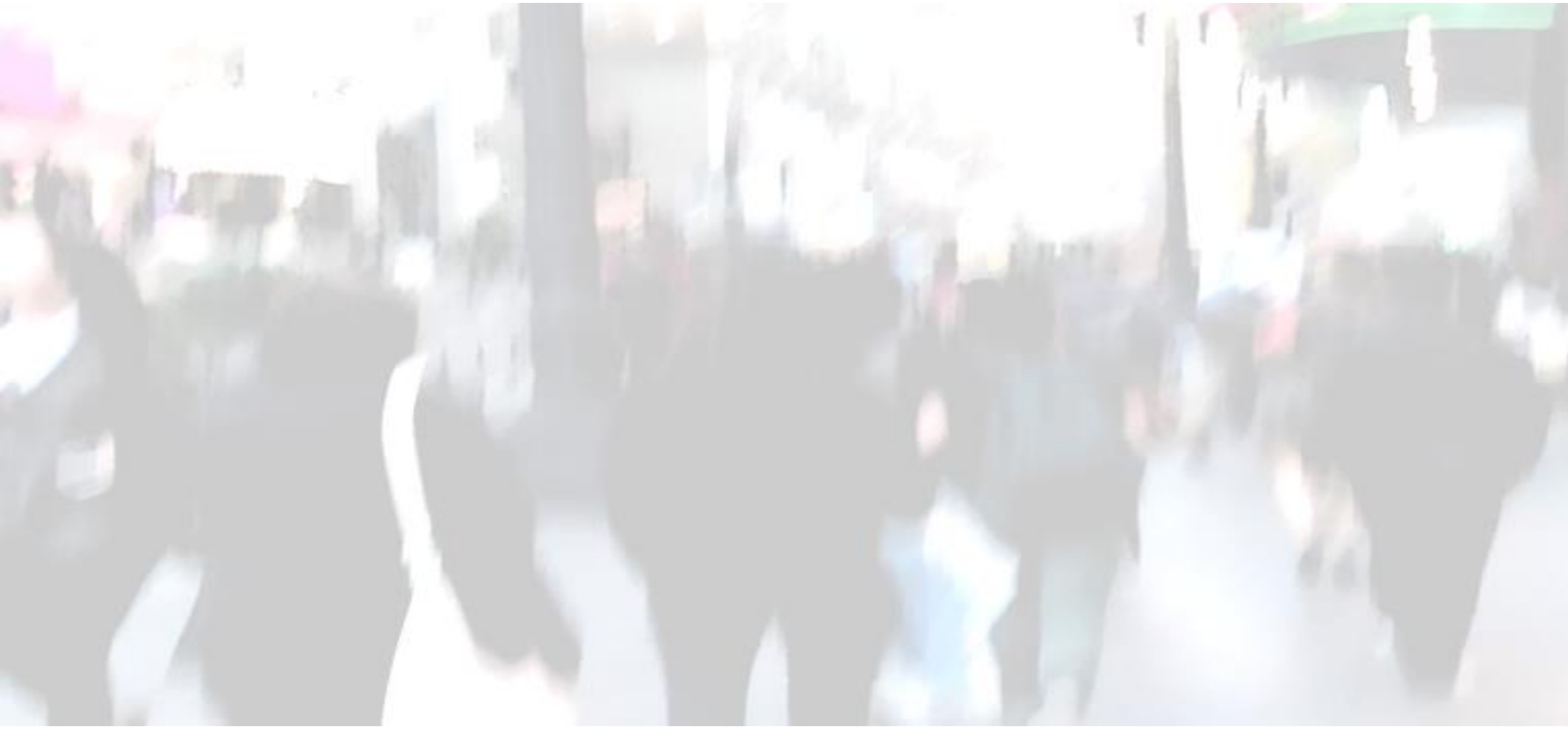
Dresner Advisory Services, LLC

2021 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.

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 2021, we mark the 14th anniversary of Dresner Advisory Services and the fifth edition of this report. Our thanks to all of you for your continued support and ongoing encouragement. Since our founding in 2007, we worked hard to set the "bar" high—challenging ourselves to innovate and lead the market—offering ever greater value with each successive year.

At the time of publication of this report, the COVID-19 pandemic continues to affect millions worldwide and impacts businesses and how they leverage data and business intelligence. As our data collection took place during Q3 and Q4 of 2020, the data and resulting analyses reflect the pandemic's impact.

Through this period, we separately conducted specific COVID-19 research, which is not reflected in this report but is available on our blog at no cost. Additionally, we will continue to collect this data and will continue to publish research through the duration of the pandemic.

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 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.

We trust you will benefit from this ADI report. It was a pleasure to develop this report, and our entire team looks forward to serving you in the future. We thank our clients, colleagues, and community members for their support, which helps us to develop this important research. We look forward to hearing from you after you explore the study findings in this report.

Howard Dresner
Chief Research Officer
Dresner Advisory Services

Contents

Definitions 3
Business Intelligence Defined 3
Analytical Data Infrastructure Defined 3
Introduction 4
Contents 5
About Brian Wood 7
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
ADI Analytical Features 82
Vendor Ratings 89
Customer Experience Model 90
Vendor Credibility Model 92
Detailed Vendor Ratings 94
Amazon Detailed Score 95
AtScale Detailed Score 96
Cloudera Detailed Score 97
Couchbase Detailed Score 98
Exasol Detailed Score 99
Google Detailed Score 100
IBM Detailed Score 101
Incorta Detailed Score 102
Microsoft Detailed Score 103
MongoDB Detailed Score 104
Oracle Detailed Score 105
SAP Detailed Score 106
SAS Detailed Score 107
SingleStore Detailed Score 108
Snowflake Detailed Score 109
Teradata Detailed Score 110
Other Dresner Advisory Services Research Reports 111
Appendix: Analytical Data Infrastructure Survey Instrument 112

About Brian Wood

Brian Wood is a Research Director with Dresner Advisory Services. Brian has over 30 years in the consulting, systems integration, and software business, working with multinational clients in industries that ranged from Financial Services, Consumer Packaged Goods, Life Sciences, and Manufacturing.

Brian was formerly a Research Director and Analyst at Gartner covering CRM, Architecture and Integration, "Governance, Risk, and Compliance" (GRC), and Corporate Performance Management (CPM).

Brian holds a BS in Finance and an MBA in International Business from the University of Rhode Island (URI).

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.

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

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

- Wisdom of Crowds® Flagship BI Market Study
- Cloud Computing and Business Intelligence
- Data Catalog
- Data Preparation
- Data Science and Machine Learning
- Embedded Business Intelligence
- Self-Service BI

Howard (www.twitter.com/howarddresner) conducts a bi-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 2021 report during Q3 and Q4 of 2020.

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

Analytical Data Infrastructures provide the platform and technology for integrating, preparing, modeling, managing, storing, and providing analytic capabilities. Previously, the market called analytic data infrastructures by names such as “data warehouses” or “analytic databases.” But today’s diversity and scale of BI and analytic use cases require new technology, architectures, and data/analytic workflows at high scale and security and with new licensing models, hence, the emergence of Analytic Data Infrastructures.

Various analytic and BI use cases consume the analytic content from the analytic data infrastructure. In our 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 different combinations of data and analytical workflows to an ADI platform. The goal of this report is to provide better understanding of the priority of use cases and preferences for Analytical Data Infrastructure features/capabilities. Understanding these capabilities, uses, and adoption helps prioritization planning, developing, and execution of a BI and analytics strategy for any size organization.

- When asked to prioritize the use cases for their ADI platforms, the use case most respondents identified as a top priority for ADI platforms is business user reporting and dashboards (fig 5). Accordingly, we believe that support for this use case is really table stakes in the ADI market and not a differentiator. The data science use case is a high priority for over 50 percent of respondents—down slightly from 51 percent last year. 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 42 percent of the respondents.
- Corporate standards have been 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. Lack of corporate standards also will lead to data governance challenges. We also see the importance of cost increasing somewhat this year

largely as a result of the need to do more with less (or what an organization already has).

- The majority of business function respondents prefer an ADI platform accessed/licensed via a cloud deployment ("as-a-cloud service" versus on-premises software). However, respondents from Marketing and Sales, and Business Intelligence Competency Center (BICC) organizations do not have this preference, with Sales and Marketing preferring a hybrid model and BICC rating cloud and on-premises deployment equally.
- Across multiple use cases, the "as-a-cloud service" option is the top deployment priority for respondents (fig. 20). Few prefer the hybrid model, which is the reality for many organizations today; they have ADI use cases that prefer cloud services and others that prefer on-premises deployment. Multiple buying centers, and the lack of a formal governance process in most organizations can result in the creation of a "hybrid" deployment model for ADI platforms.
- Cross data center integration and management capabilities to support hybrid deployments (cloud, multi-cloud, and on premises) is a growing ADI platform priority. We expect cross data center integration and management tools for hybrid ADI platform deployments will continue to 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.
- According to survey responses, multi-tenancy support is often an ADI capability used to promote cloud offerings when compared to other ADI development and deployment priorities. Multi-tenancy priority is comparatively low.

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 52.7 percent of respondents (fig. 1). EMEA accounts for 33.2 percent of respondents. Asia Pacific accounts for 8.7 percent, and Latin America represents 5.4 percent of the respondents.

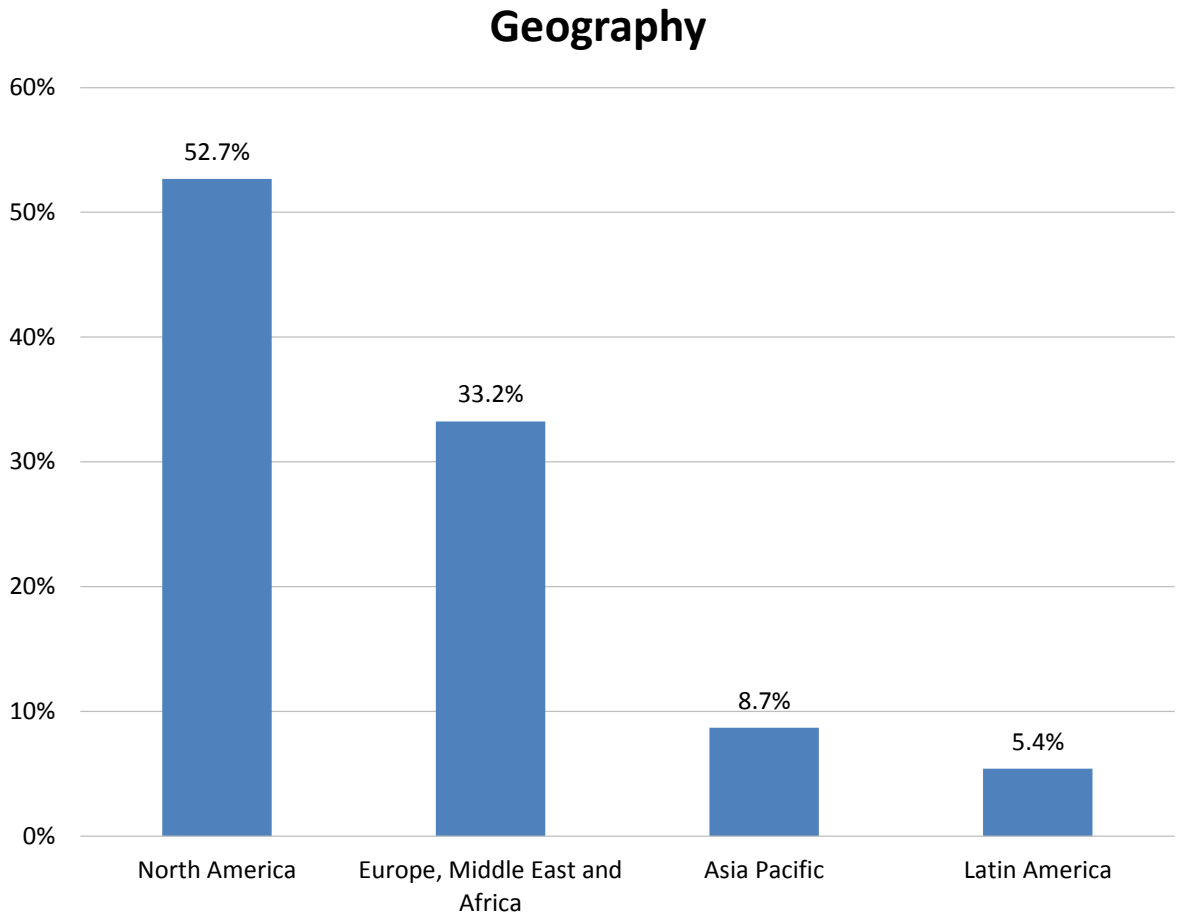


Figure 1 – Geography

Functions

IT (37.1 percent) and BICC (21.9 percent) are the largest groups represented in our ADI sample (fig. 2). Executive Management comprises 13.2 percent. The remaining respondents come from Finance, R&D, Operations, Strategic Planning, and Marketing functions, with 6 percent from other functions.

Examining ADI priorities and preferences by function provides insight to 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.

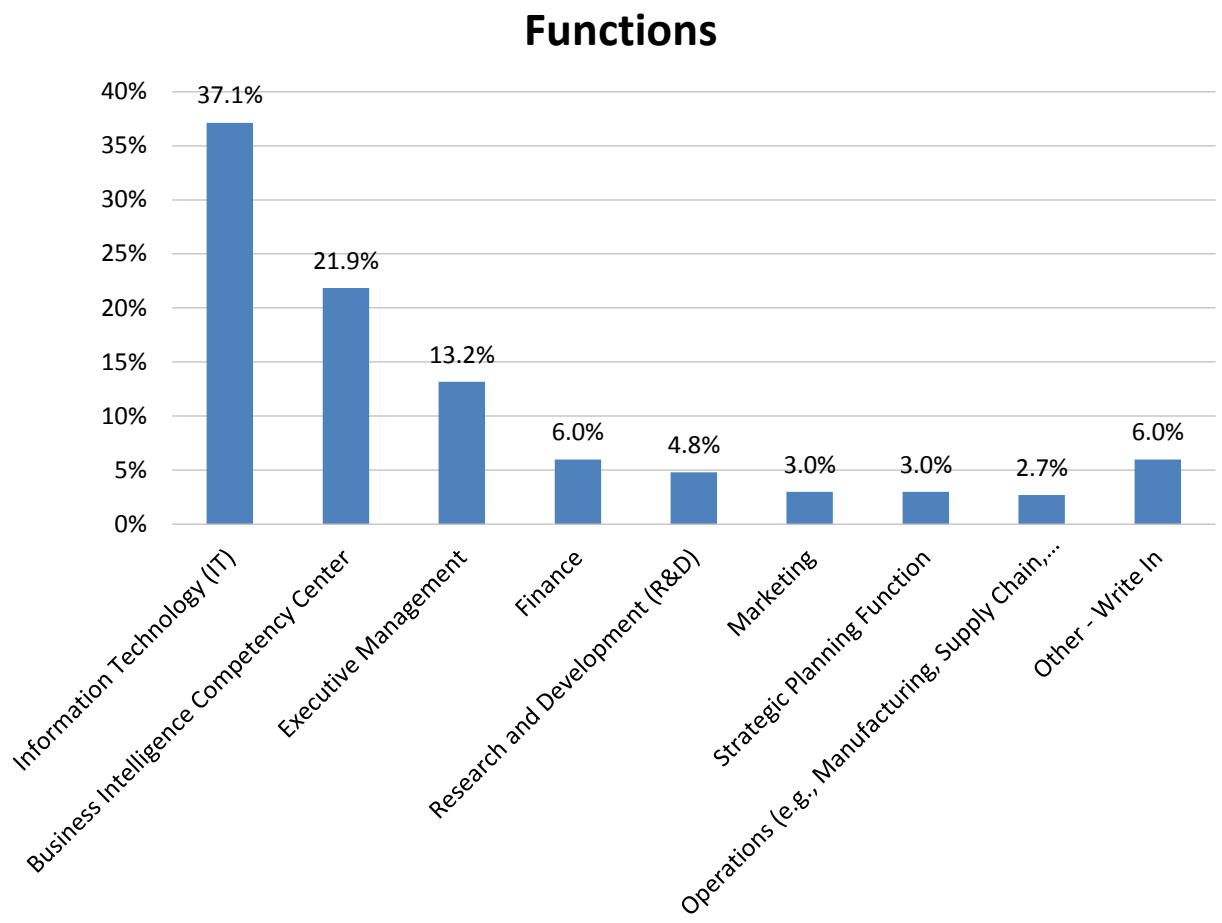


Figure 2 – Functions

Vertical Industries

Respondent organizations represent a mix of vertical industries. Technology accounts for 14 percent, Financial Services 10 percent, and Healthcare 9 percent. Other industries include Manufacturing at 7 percent, Retail/Wholesale and Consulting at 6 percent, Insurance at 5 percent, and Telecommunications and Higher Education at 4 percent. The balance of respondents come from Not for Profit, Automotive, Transportation, Real Estate, Advertising, Food Beverage, & Tobacco, Pharmaceuticals, State and Local Government, Utilities, Chemical, Entertainment and Leisure, Hospitality, Distribution & Logistics, and Federal Government, each with 3 percent or less, and an additional 4 percent in the "Other" category.

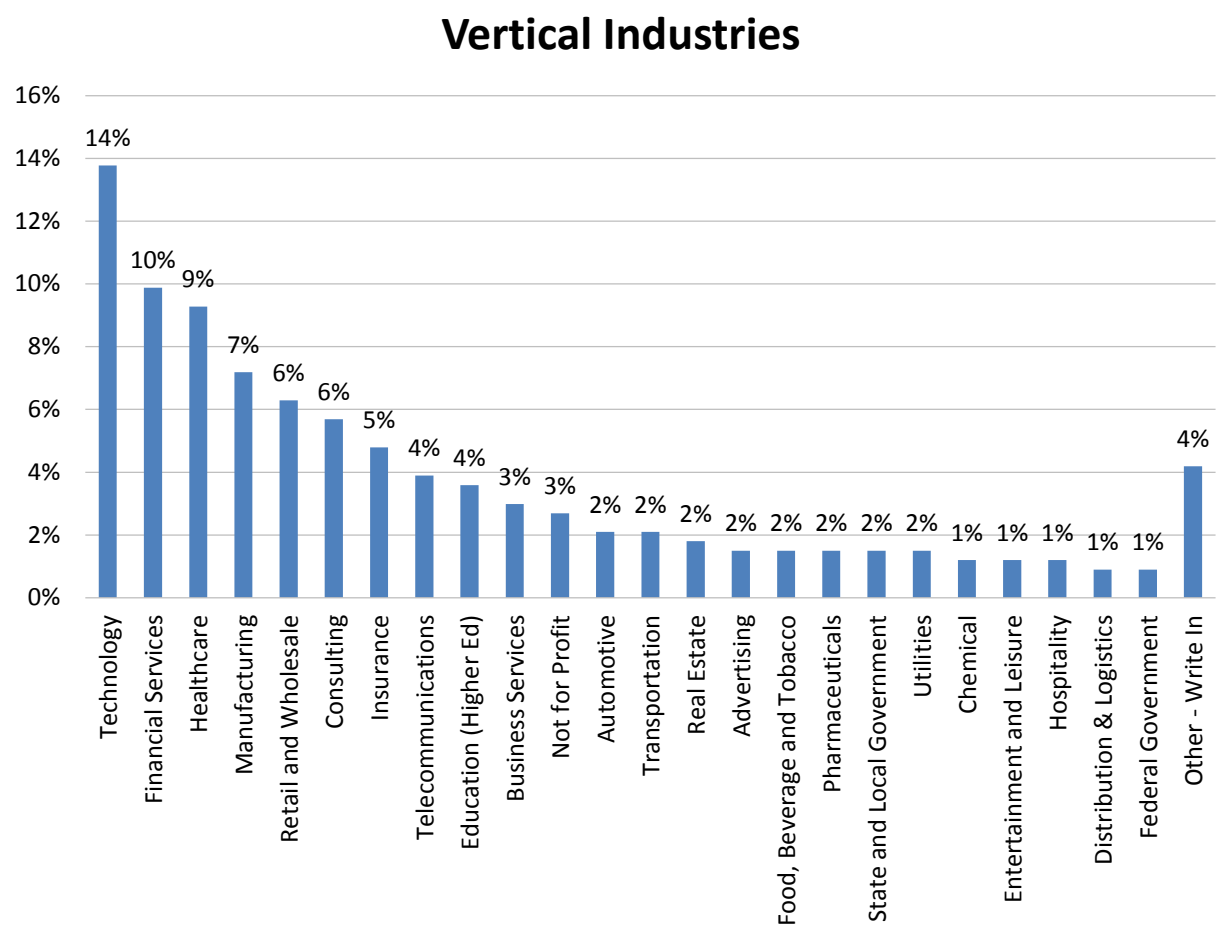


Figure 3 – Vertical industries

Organization Size

Analyzing our market survey 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 important just to large enterprises; 22.8 percent of respondents are from small organizations with fewer than 100 employees, and mid-sized organizations with 101-1,000 employees account for another 27.8 percent.

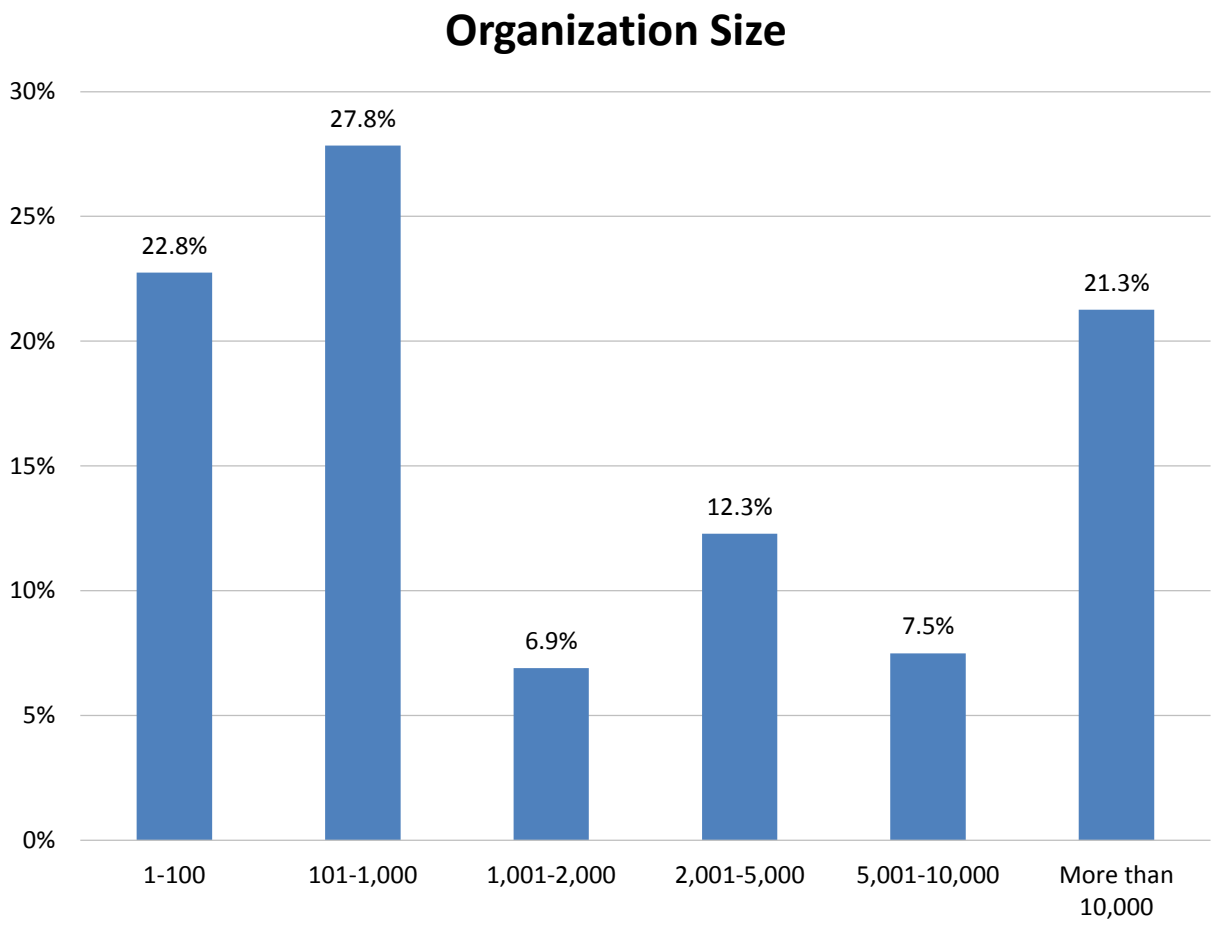


Figure 4 – Organization size

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 (a scale of one to five with five being the 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

ADI use cases are diverse. Different use cases often access different data sets for analysis and use different workflows and data processing pipelines for moving the data from the sources, storage, doing the analysis, and delivery to the analytic consumer. Therefore, we see a diversity of buying requirements and priority. Buying requirements for ADI platforms often represent priorities for more than one use case.

We asked respondents to rank the importance of four types of ADI use cases. The largest proportion (78 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 65 percent of respondents rank as "critical" or "very important." Forty-nine percent rank "data science (advanced and predictive analytics or data mining)" use cases as their highest priority, and 42 percent rank "embedded analytics with business applications" as their highest priority use cases.

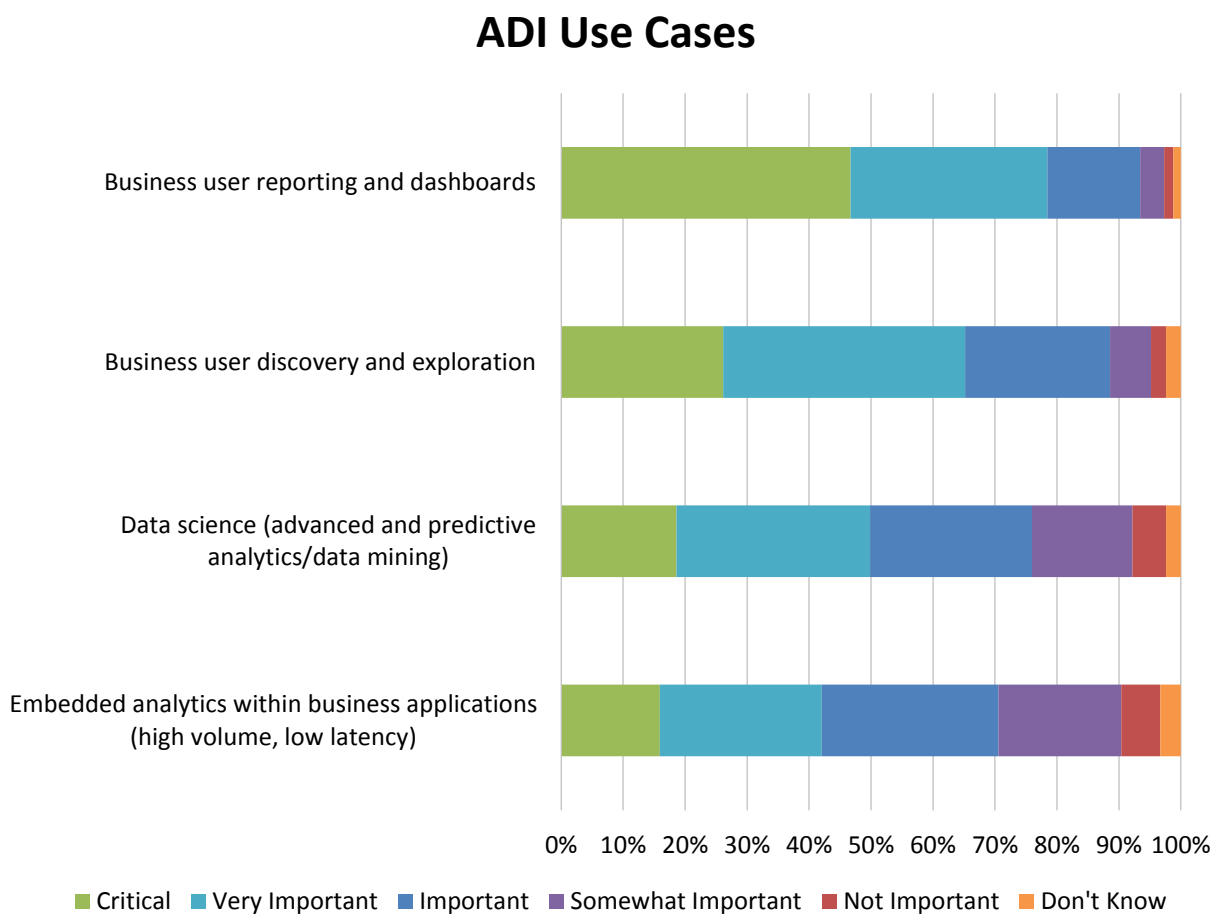


Figure 5 – ADI use cases

Year over year, “business user reporting and dashboards” and “business user discovery and exploration” are the top use cases; however, they both decrease very slightly in priority (fig. 6) year over year. The changes in ADI priorities for the data science use case and embedded analytics use cases also decreases slightly. 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.

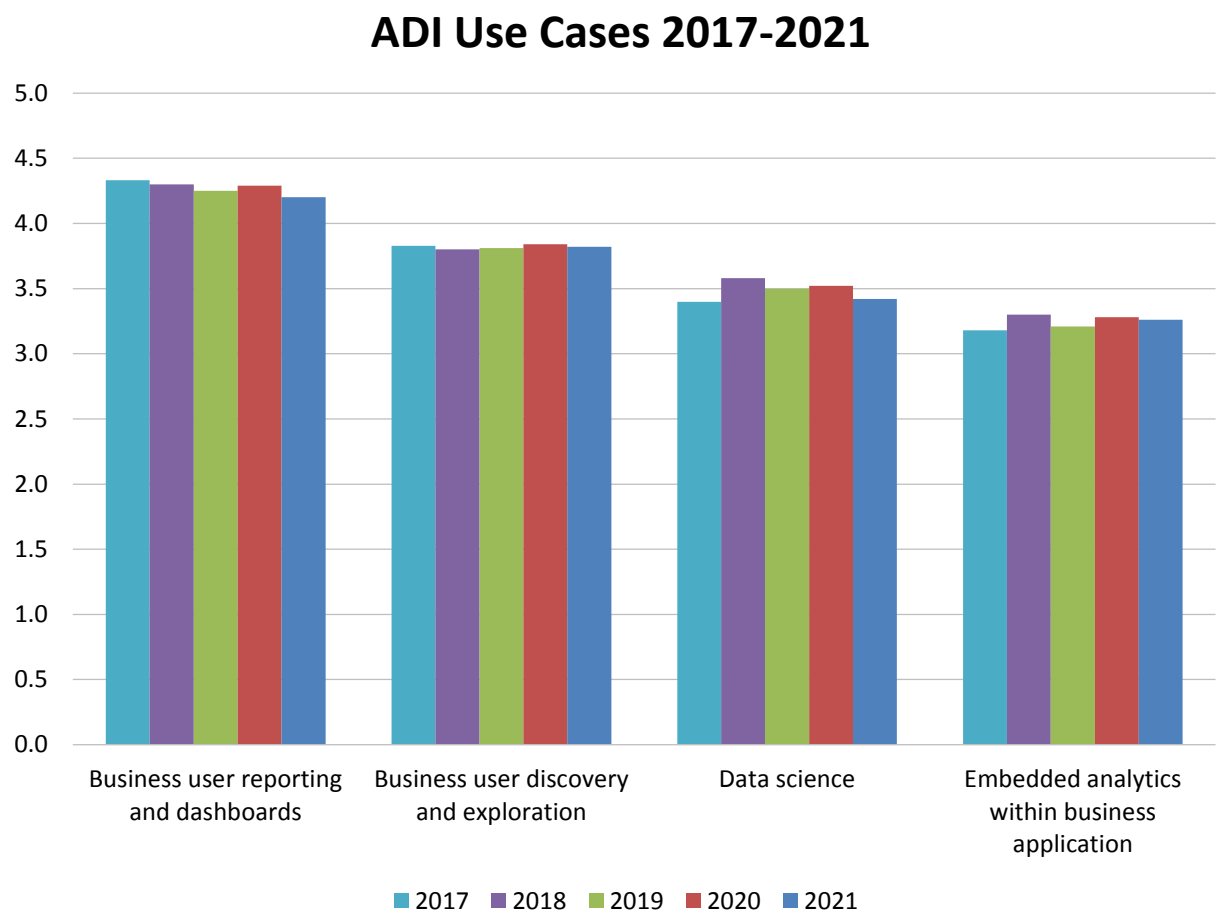


Figure 6 – ADI use cases 2017-2021

In 2021, “business user reporting and dashboards” and “business user discovery and exploration” are top use-case priorities for ADI platforms and investments (fig. 7) for all geographies. While this is obviously an important use case, we believe this is essentially table stakes and not likely to be a differentiator.

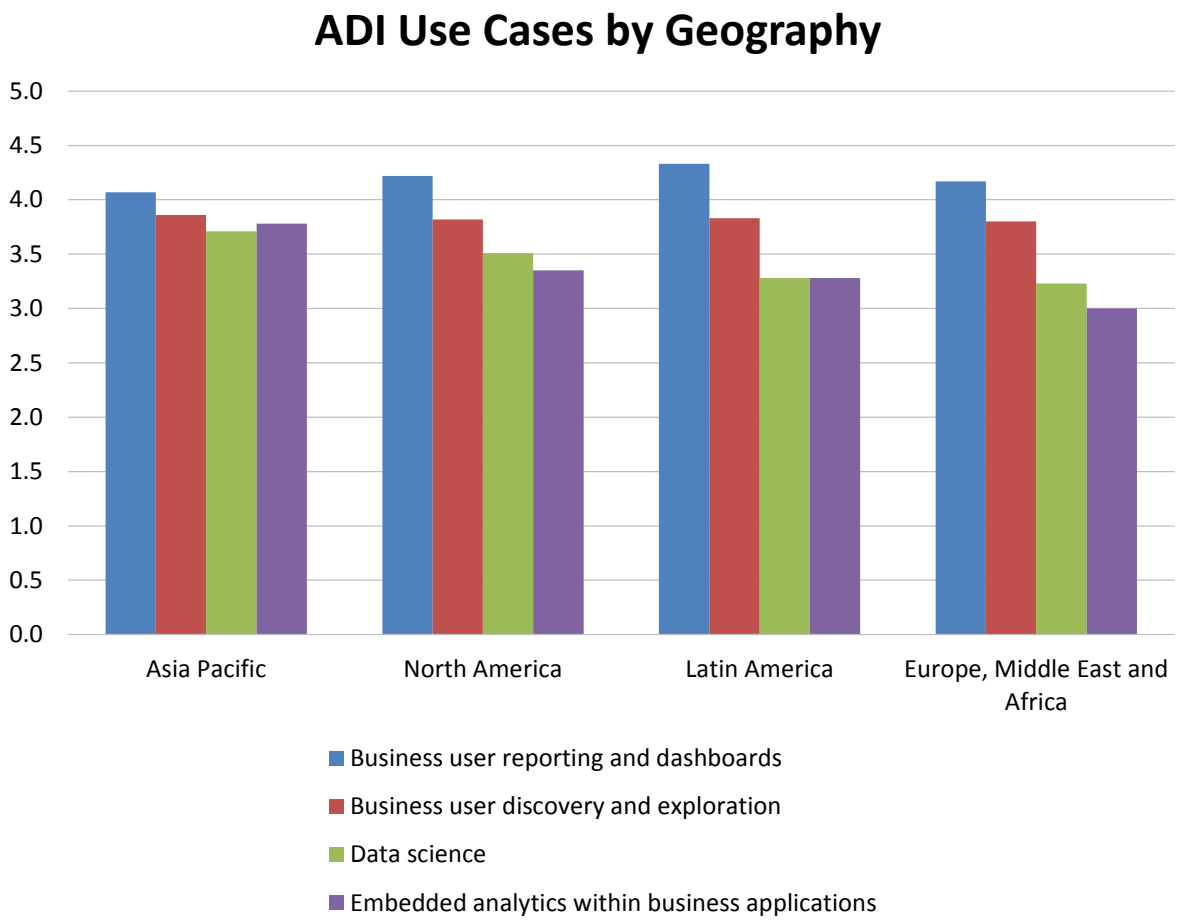


Figure 7 – ADI use cases by geography

Survey respondents from different organizational functions indicate differing use-case priorities for their ADI requirements, making the selection of ADI platforms more complicated (fig. 8). Most organizational functions have similar profiles in use-case priorities, apart from Sales/Marketing respondents. Sales/Marketing functions place a higher-than-average priority on "business user discovery and exploration."

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

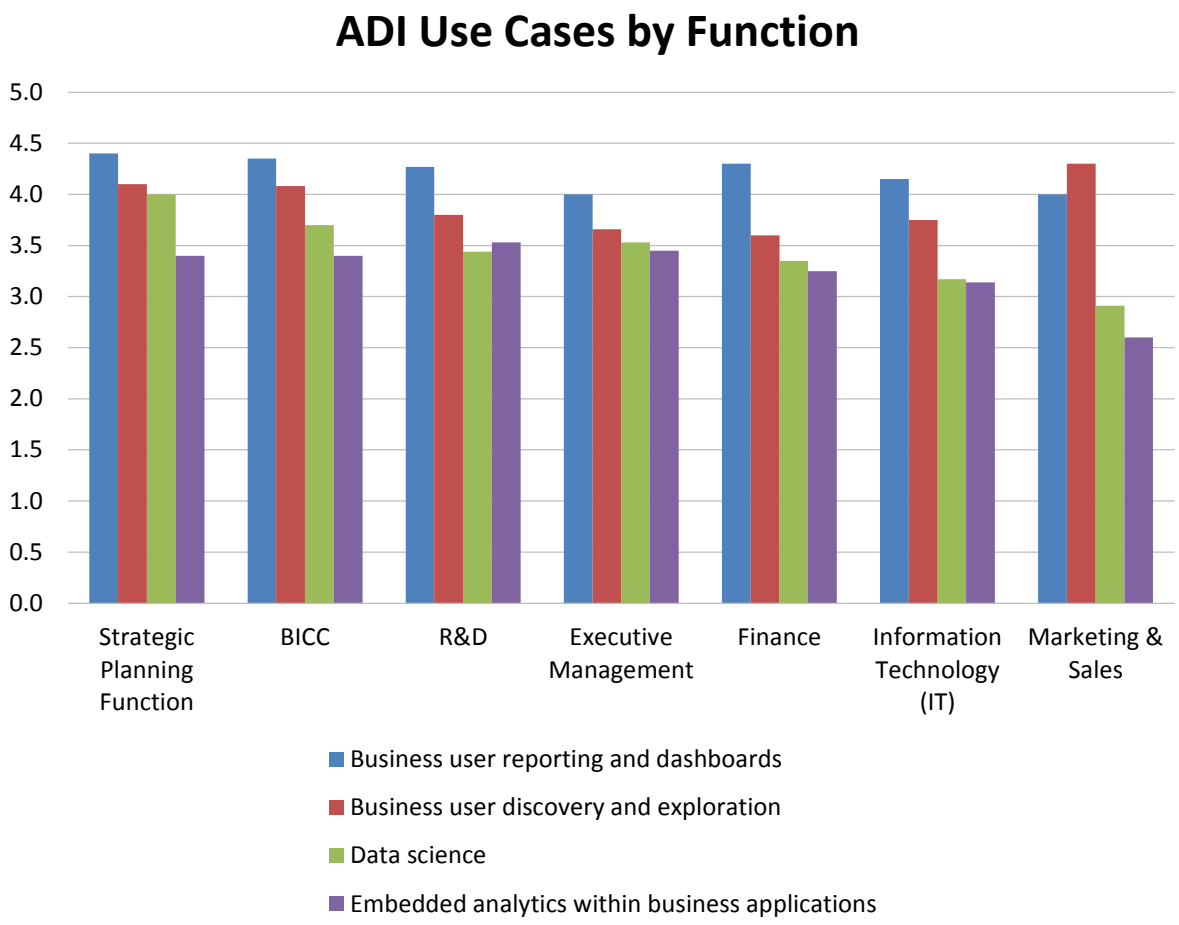


Figure 8 – ADI use cases by function

Organizations from different industries have similar priorities for ADI platforms (fig. 9). "Business user reporting and dashboards" and "business user discovery and exploration" are the top use cases across all industries. Respondents from Technology and Business Services place a higher relative priority on "embedded analytics," and Financial Services places more emphasis on "data science" use cases.

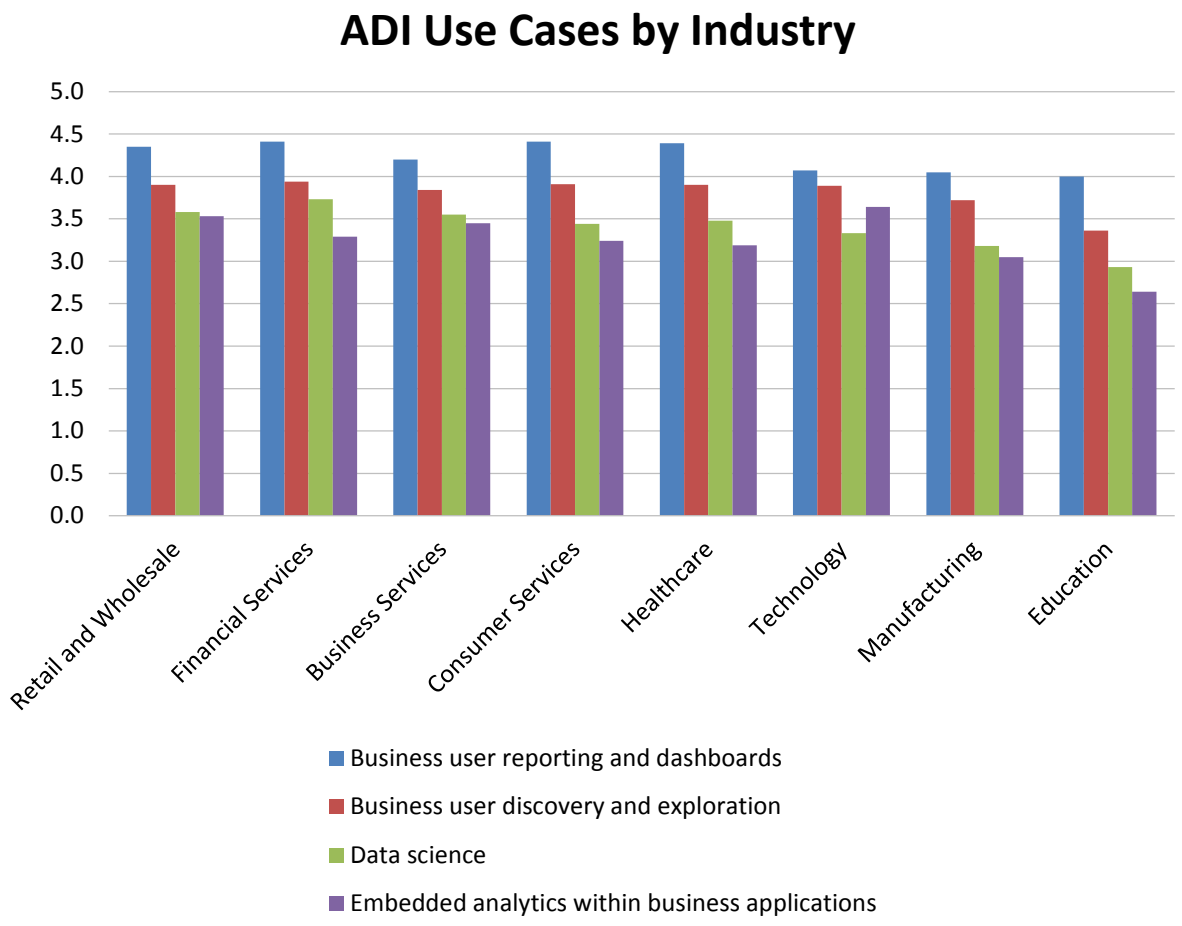


Figure 9 – ADI use cases by industry

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. While the priorities do not seem to change, larger organizations tend to rate all use cases slightly higher than smaller ones, with the exception of organizations of less than 100 employees, who rate “embedded analytics within business applications” higher.

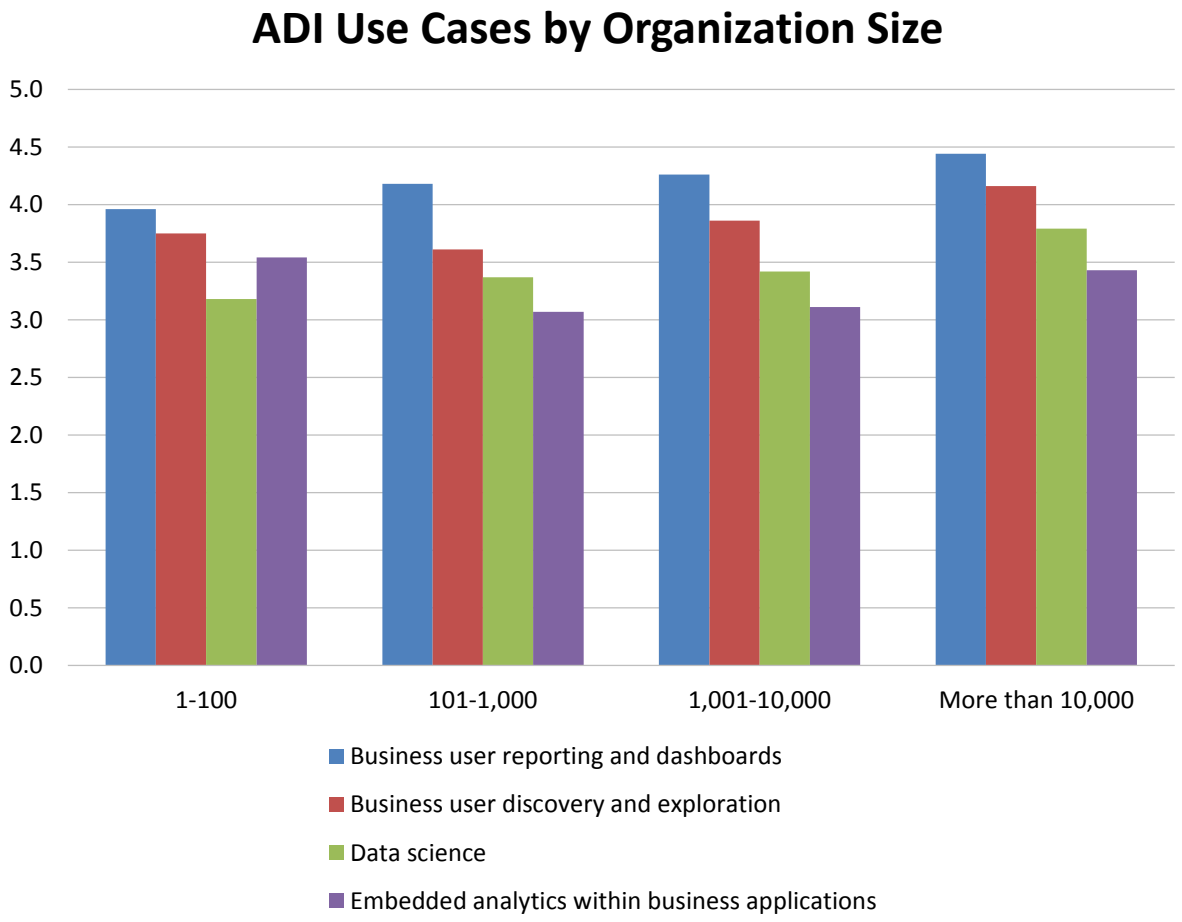


Figure 10 – ADI use cases by organization size

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 remain the highest selection priorities for ADI platforms in 2021. The ever-growing volume of data, algorithms, and number of users arising from combinations of use cases drives this priority. 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. 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" is a relatively low priority, demonstrating the diversity of ADI platforms/skills/workflows and that, for most organizations, no one ADI platform can be a "corporate standard." Multi-cloud support and single-vendor solution appear to be the lowest priorities.

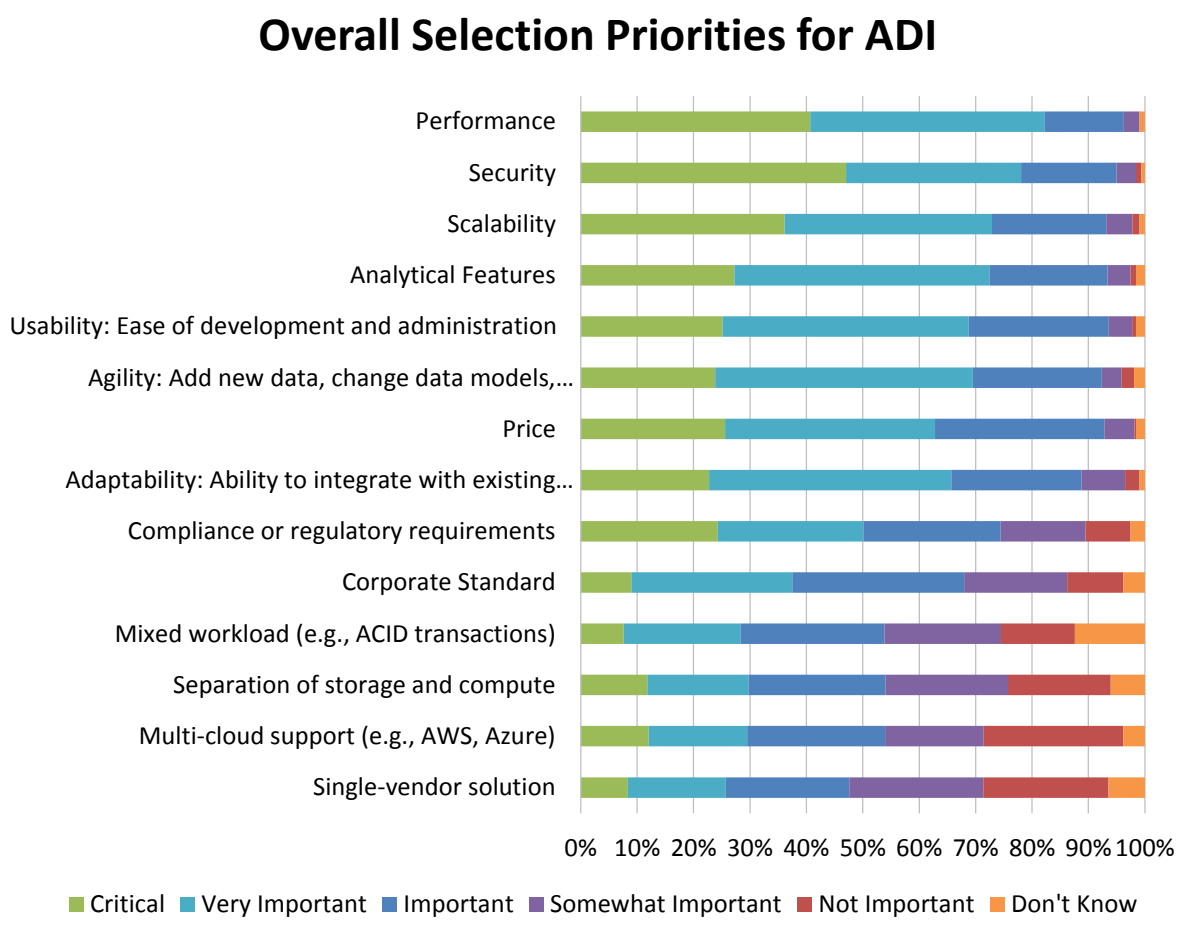


Figure 11 – Overall selection priorities for ADI

Performance leads the selection priorities for ADI platforms again in 2021. Security deviates from the rising trend upward in priority over the past four years with the priority slightly lower than last year (fig. 12). Usability remains a high priority year over year. And, like last year, respondents rank corporate standards as their lowest selection priority, ranking the priority of compliance only slightly higher. The priority of price increases year over year, which is consistent with the overall trend to do more with less (as identified in our COVID research).

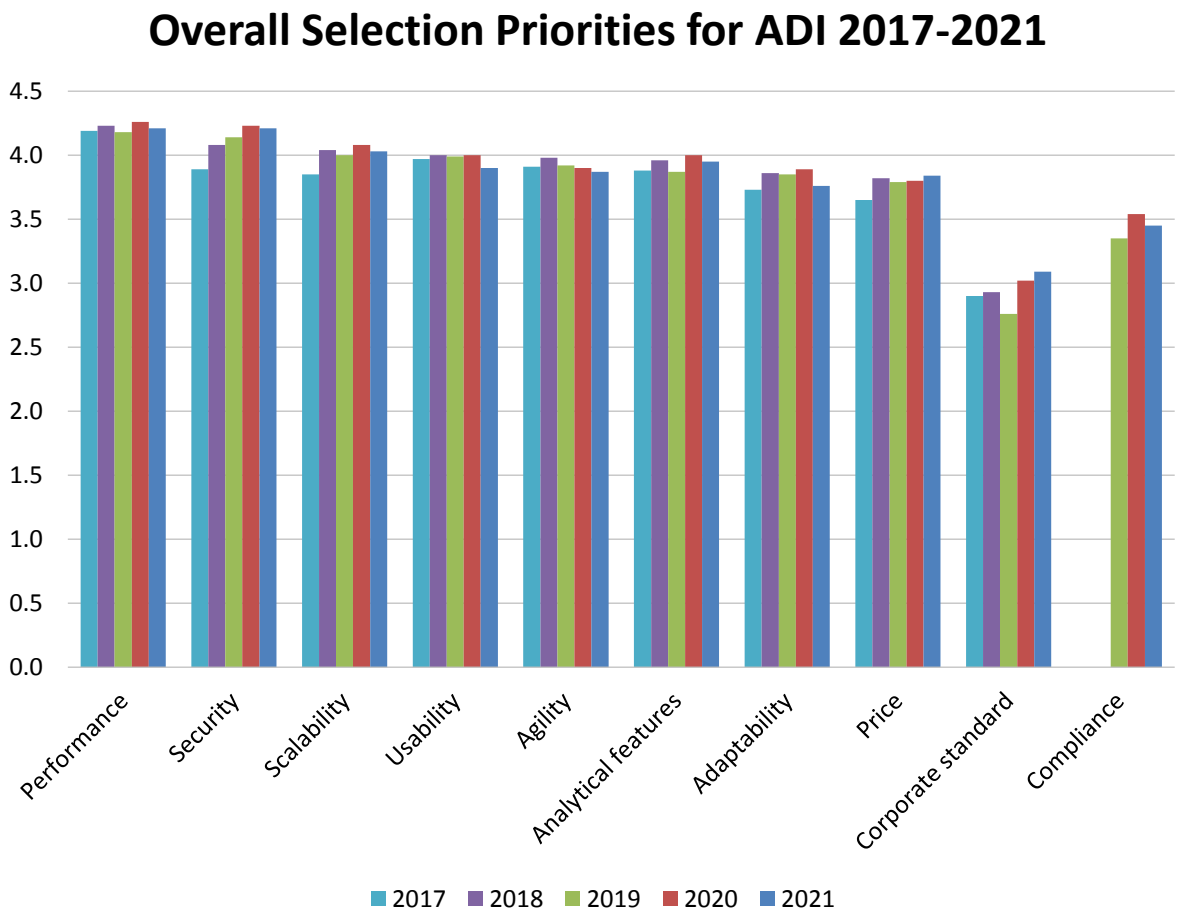


Figure 12 – Overall selection priorities for ADI 2017-2021

Selection priorities for ADI platform technologies vary only slightly, by use-case workload and workflow (fig. 13). Performance and security are the top priorities across all use cases. In the past, the market appeared to not be price sensitive, placing specific ADI capabilities ahead of the price; however, recent survey data indicate an increasing priority on price.

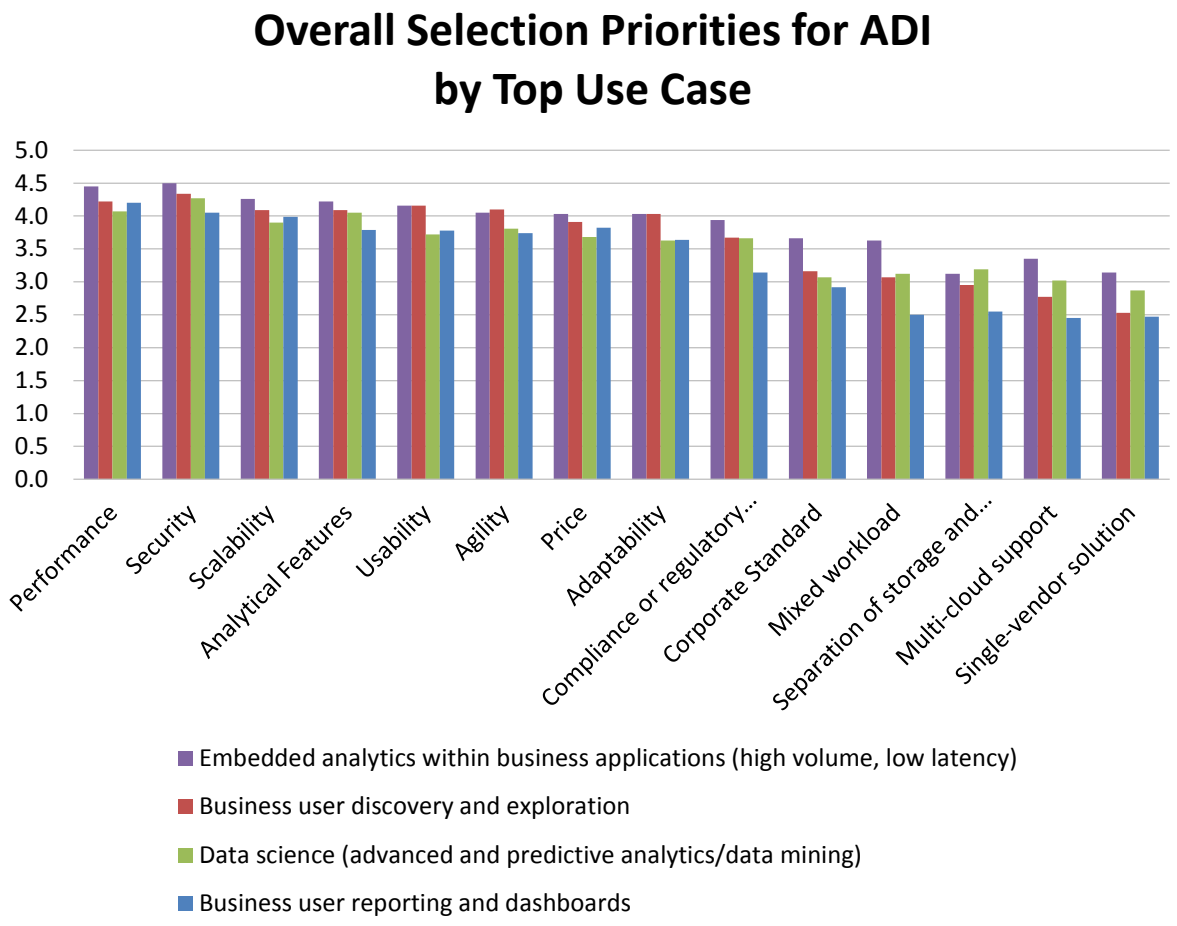


Figure 13 – Overall selection priorities for ADI by top use case

There are specific differences in ADI priorities across the geographic market dimension. All regions consider a single-vendor solution as the lowest priority for selecting ADI (fig. 14). However, in contrast, Asia Pacific ranks several selection priorities higher including mixed workload, corporate standard, and single-vendor solution.

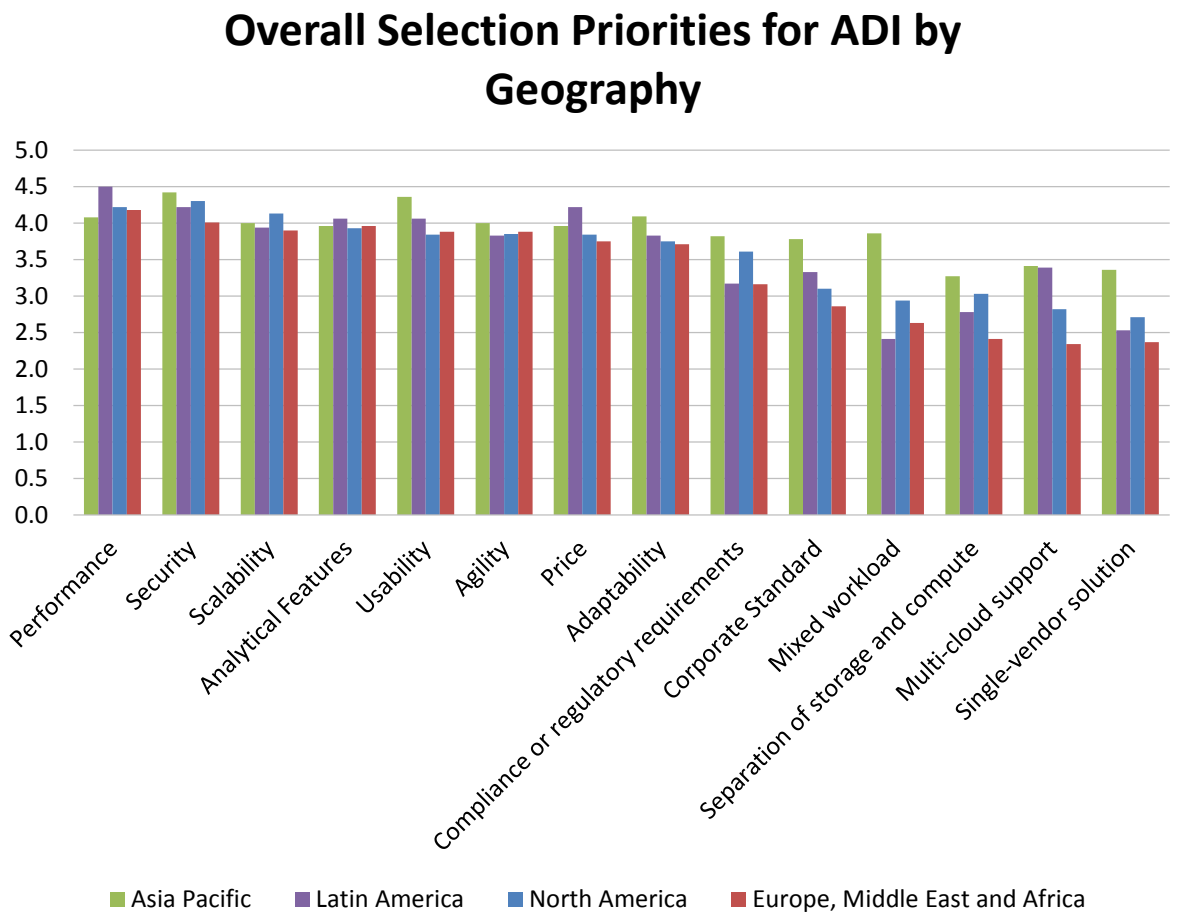


Figure 14 – Overall selection priorities for ADI by geography

As shown below (fig. 15), the relative priorities for ADI platforms vary by organization functions. Performance is the top selection priority across all organization functions, followed by security. All organization functions rate single-vendor solution as their lowest selection priority. Some organizations are more price sensitive than others regarding their selection priorities (e.g., BICC is the most price sensitive, assuming they meet or exceed all their other priorities).

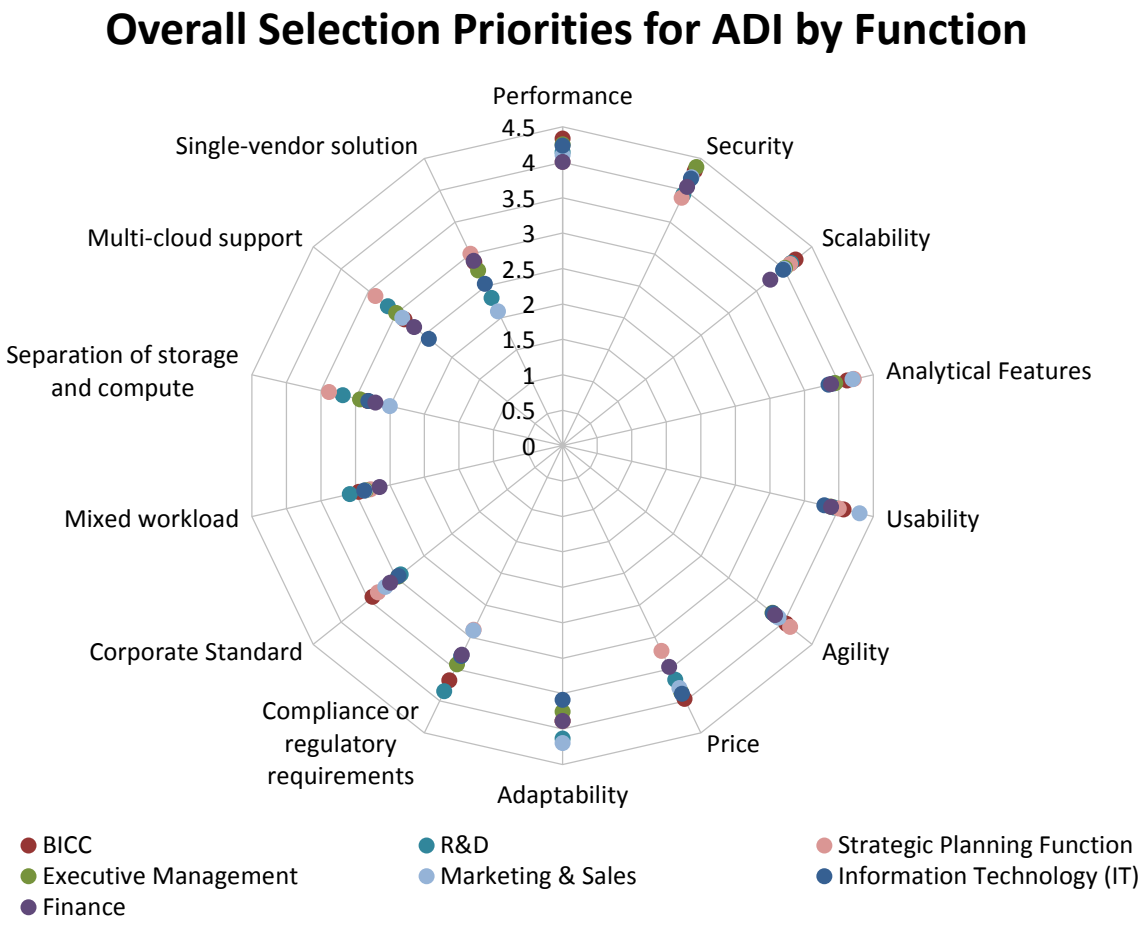


Figure 15 – Overall selection priorities for ADI by function

ADI platform priorities vary across industries (fig 16). Some variation of priorities across industries is notable, and some ADI platforms are better suited for different industries and their users/use cases. Security, performance, and scalability are generally the highest priorities. All other sectors place performance as the highest ADI platform priority. Government and Healthcare sector respondents place compliance significantly higher than other industry sectors. Government respondents also appear to be the least price sensitive and perhaps are justified, given the number and level of critical priorities (security and compliance) they place on their ADI platform decisions. Financial Services indicates a higher priority across the board, and particularly with respect to security, performance, scalability, and analytical features.

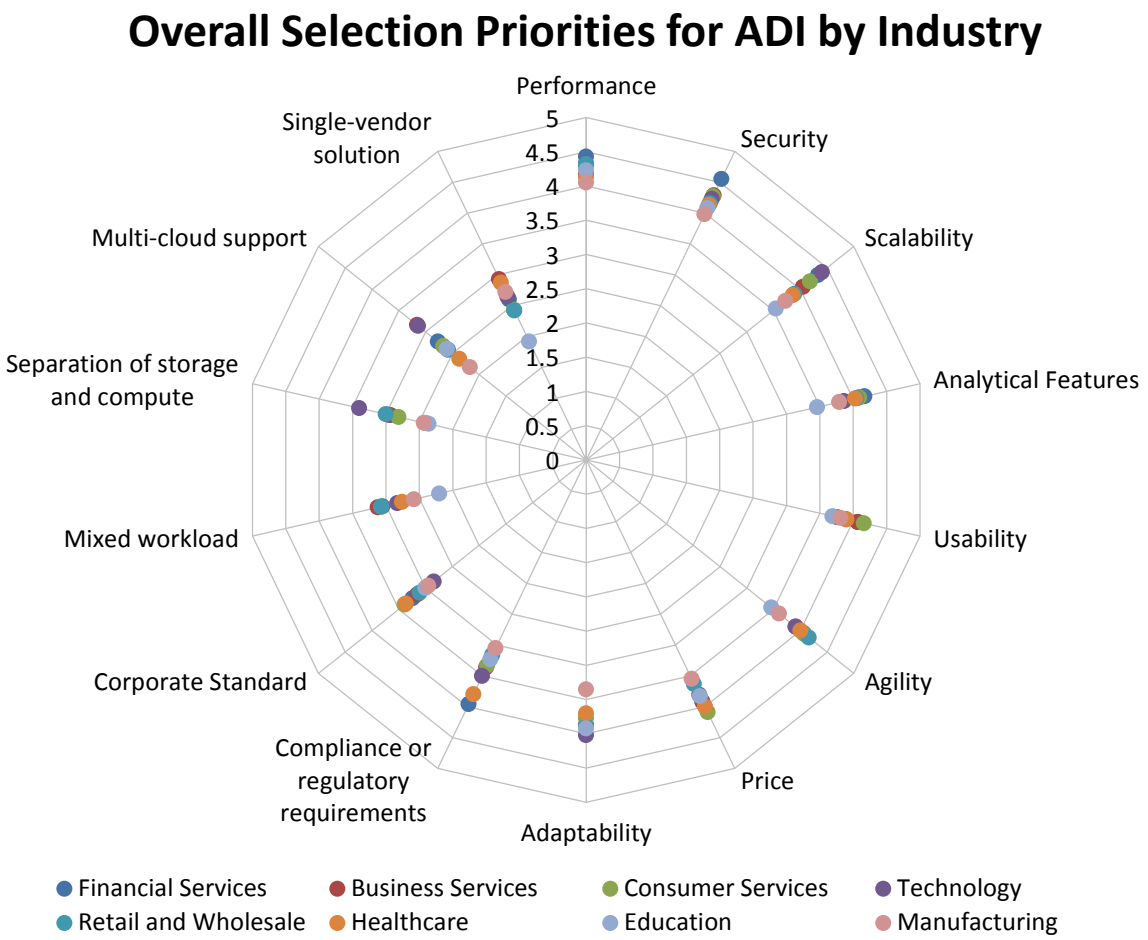


Figure 16 – Overall selection priorities for ADI by industry

ADI platform priorities also vary by size of organization. The largest organizations place higher priority on all selection criteria. Except for mid-size organizations (101-1,000 employees), security is the highest priority. Mid-sized organizations indicate performance as their top selection priority, followed by security as their second-highest priority (fig. 17). Compliance is a low priority for small and mid-sized organizations (1–1,000 employees). The smallest organizations (1-100 employees) and largest (more than 10,000 employees) are the most price sensitive.

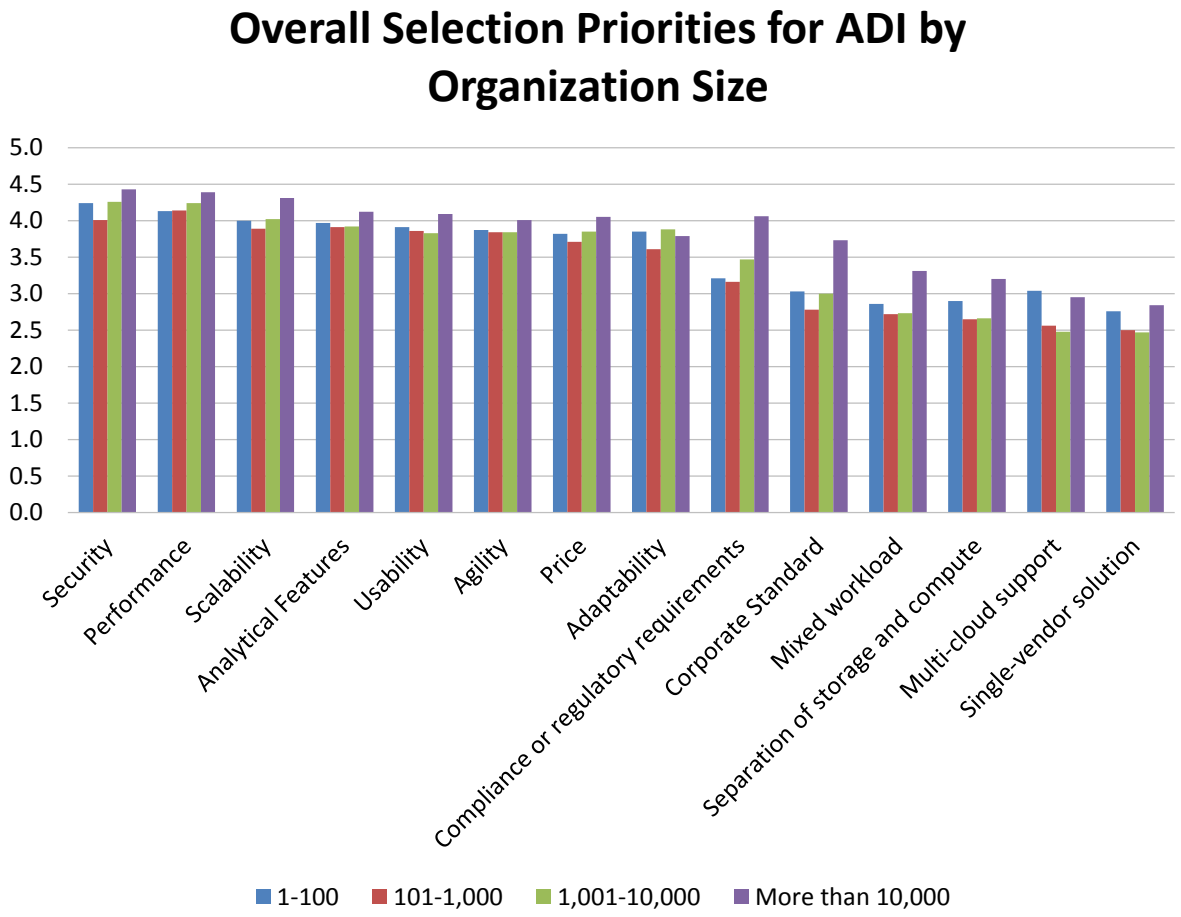


Figure 17 – Overall selection priorities for ADI by organization size

Analytical Data Infrastructure Deployment and Licensing

ADI Deployment Priorities

"As-a-cloud service" deployment of ADI platforms is the highest priority ("critical" or "very important") in deployment options for the majority of our survey respondents. "Cross data center integration and management capabilities" and "hybrid deployment (cloud and on premises)" are top ADI priorities for 26-32 percent of respondents (fig. 18). Notably, 40 percent of respondents rate on-premises software as a top priority.

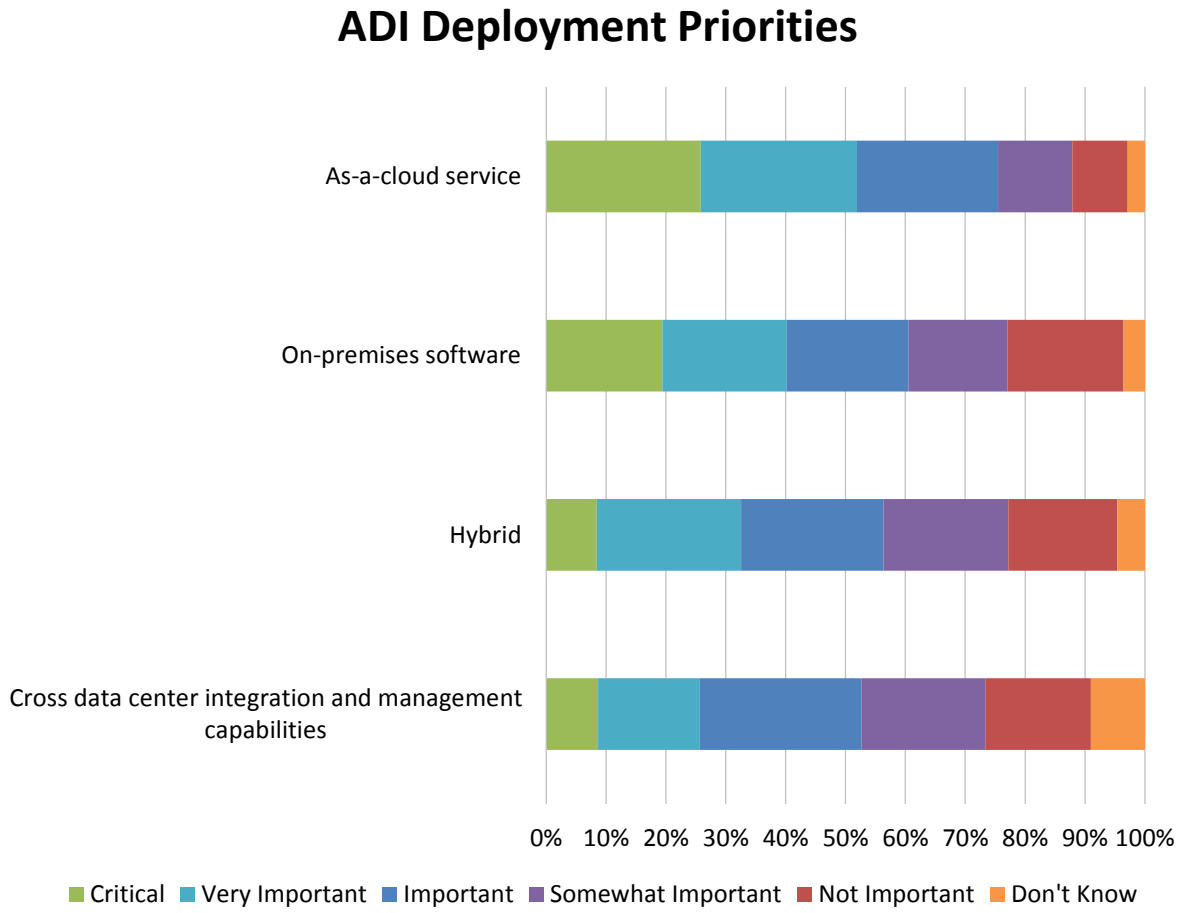


Figure 18 – ADI deployment priorities

A trend of note regarding deployment priorities emerges when looking at the year-over-year data sets (fig. 19). Choosing an ADI offered as an "as-a-cloud service" deployment picks up momentum as a preference, whereas the on-premises ADI preference softens but still is a priority for over 40 percent of organizations. Correspondingly, the priority for hybrid deployment trends upward; hybrid is a waypoint on the journey to cloud. Cross data center integration and management capabilities for ADI platforms also trends upward.

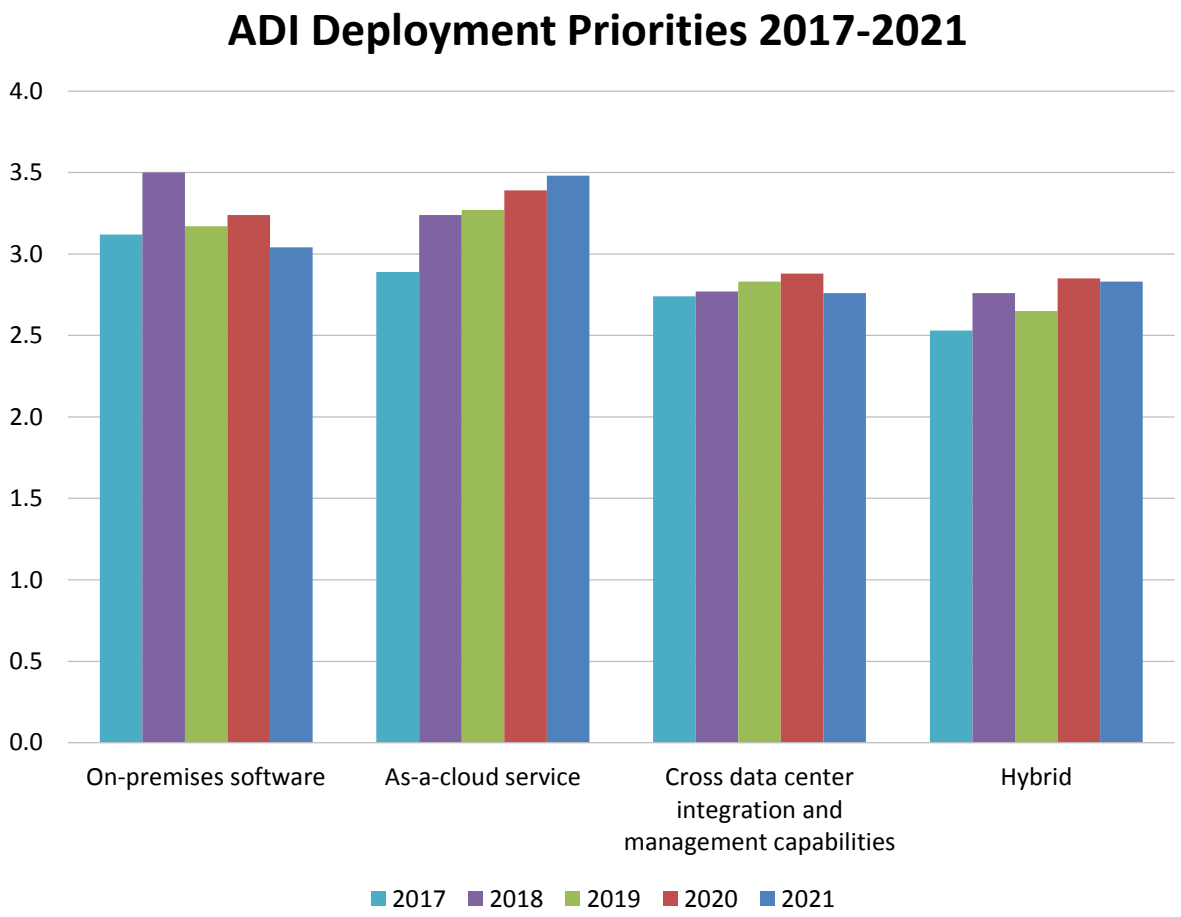


Figure 19 – ADI deployment priorities 2017-2021

Across all use cases, the “as-a-cloud service” option is the top deployment priority for respondents (fig. 20). Fewer prefer the hybrid model. Multiple buying centers and use cases and ill-defined/managed governance processes in most organizations end up creating a “hybrid” deployment model for ADI platforms by default. The associated analytic content contained in hybrid ADI deployments creates a data workflow challenge; users must somehow integrate and operationalize the workflow to serve potentially multiple use cases.

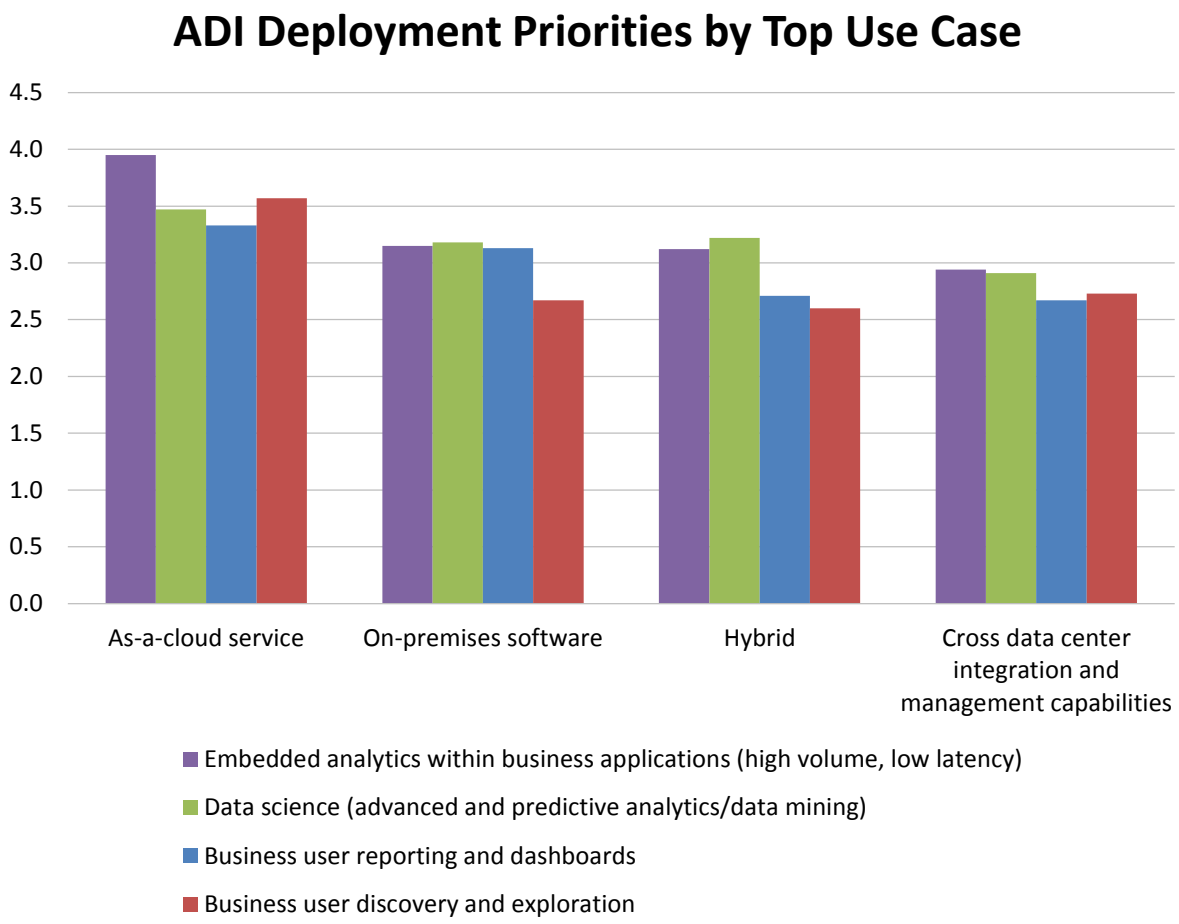


Figure 20 – ADI deployment priorities by top use case

Deployment priorities vary by geography. All geographies, with the exception of EMEA, show a preference for “as-a-cloud service” for their ADI platforms. EMEA shows a slight preference for on-premises options. Hybrid deployment options and “cross data center integration and management capabilities” rate higher in organizations in Asia-Pacific and Latin American geographies compared to other regions.

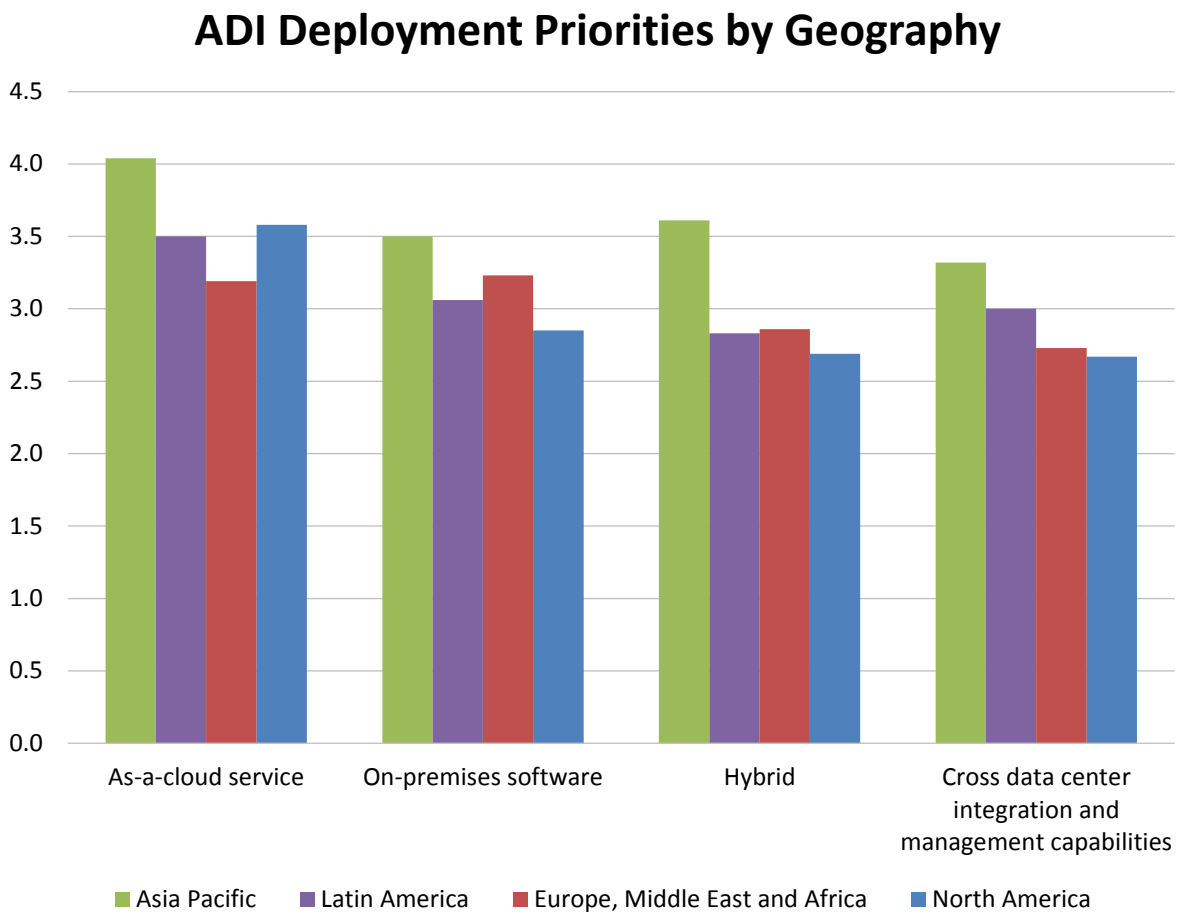


Figure 21 – ADI deployment priorities by geography

ADI deployment priorities can vary significantly, depending on the organization function indicated by the respondent (fig 22). This can make it difficult to align ADI platform priorities, especially if an application spans multiple business functions. Almost every business function shows a preference for “as-a-cloud service” for their ADI platform deployment, with the exception of BICC, where the preference for on-premises deployment is rated the same as "as-a-cloud service.”

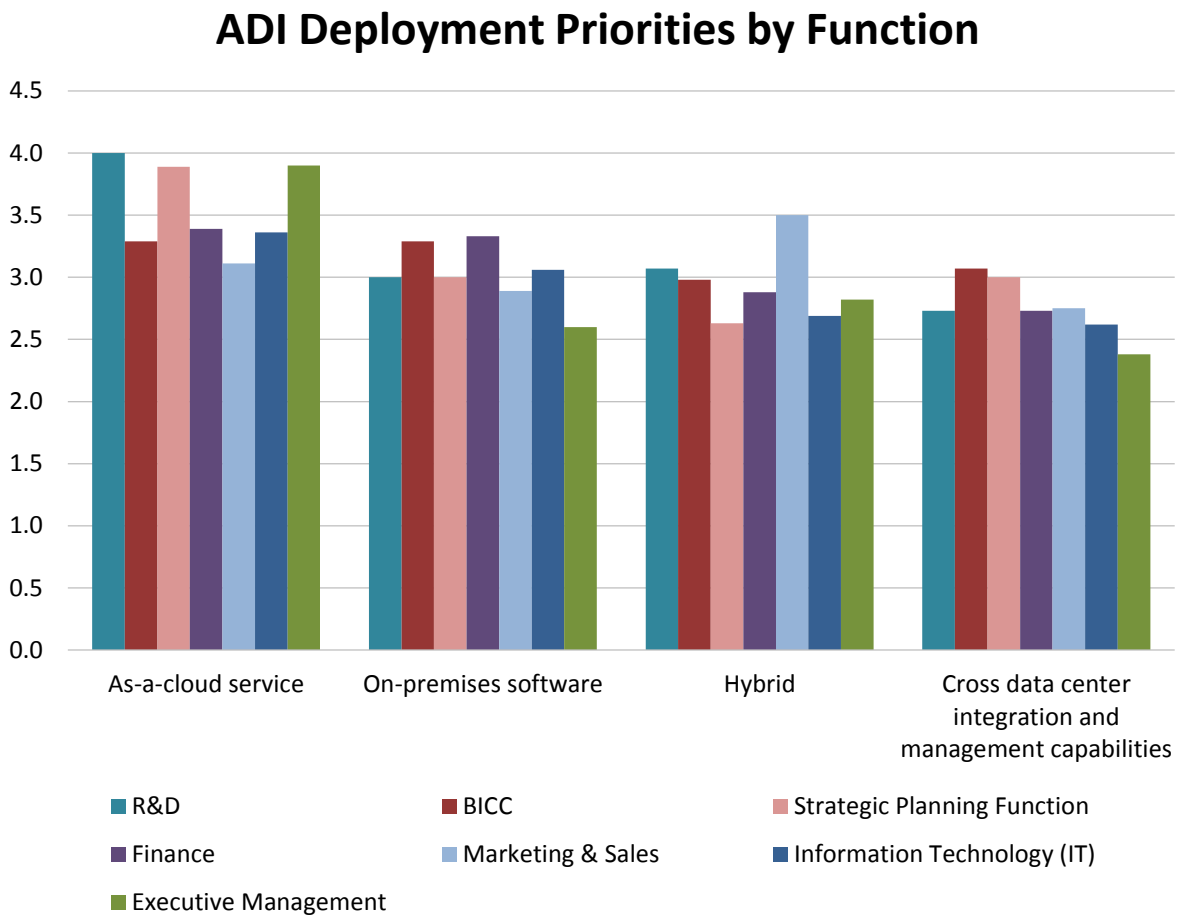


Figure 22 – ADI deployment priorities by function

ADI deployment priorities vary quite a bit by industry. Again this year, respondents from the Financial Services industry rank on-premises software as their highest deployment priority (fig. 23). Technology organizations indicate the highest preference for cloud deployments. Healthcare and Business Services respondents indicate a higher-than-average priority for cross data center integration and management capabilities for their ADI deployments.

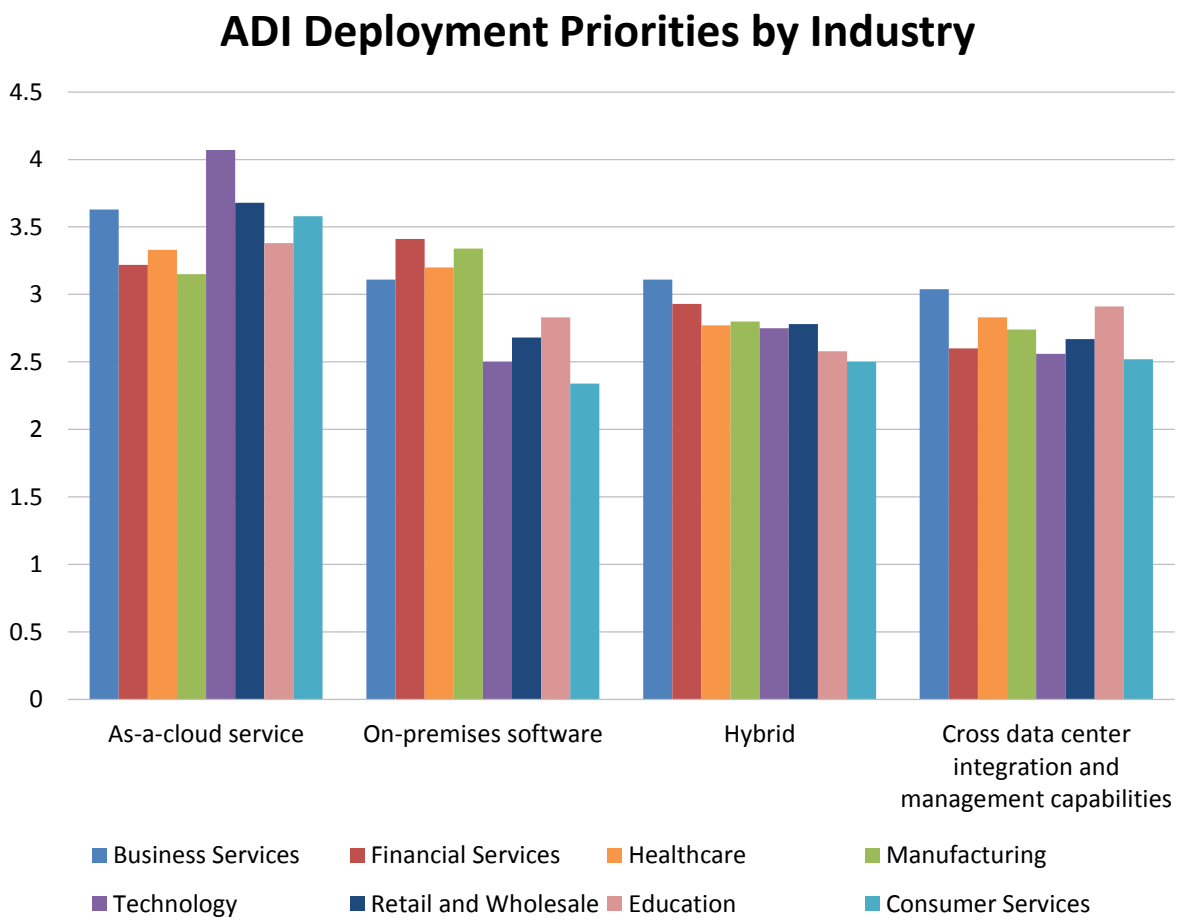


Figure 23 – ADI deployment priorities by industry

Deployment priorities vary by organization size. There are often multiple ADI platforms and deployments in any organization over the size of 1,000 employees (fig. 24). One key trend emerges: the priority of hybrid deployments and cross data integration increases with organization size. It is not uncommon for smaller organizations to lack the staff to support on-premises ADI platform deployments; hence, they turn to cloud services. It is interesting to note that the smallest and largest organizations rate cloud deployments highest.

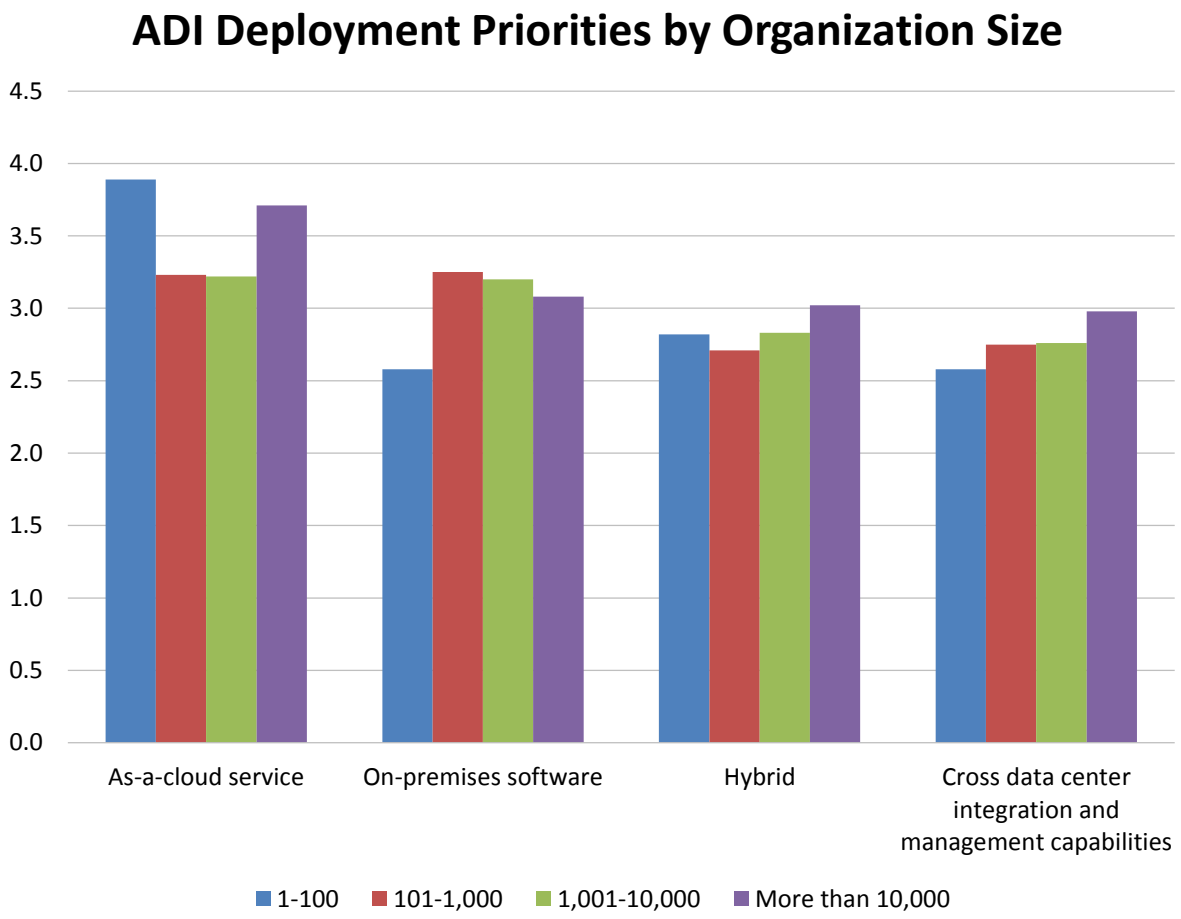


Figure 24 – ADI deployment priorities by organization size

ADI Licensing Preferences

This year, respondents show a slight preference for “concurrent use” licensing and very little preference for “open source community and commercial” licensing models. Respondents rated concurrent use almost 44 percent and all other models between 39-41 percent preference, with the exception of the “open source” model, which received 30 percent (fig. 25).

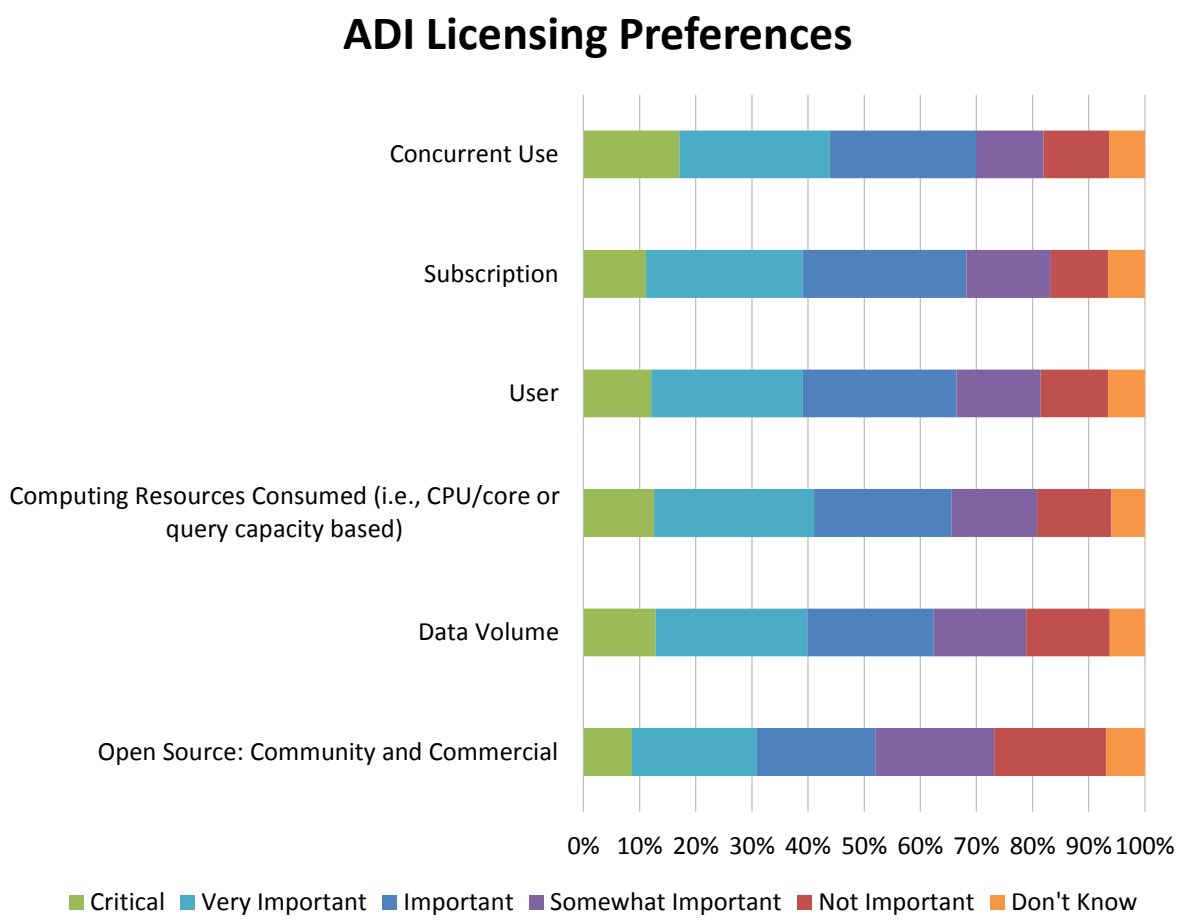


Figure 25 – ADI licensing preferences

While the “concurrent use” licensing model continues to be the top choice, the “data volume” licensing preference continues to increase slightly year over year (fig. 26) and now approaches “subscription” licensing in priority. The priority of “computing resources consumed” increases as a preference and is now equal to the “user” model. These trends in licensing preferences reflect the trend in the preferences for cloud services (see fig 21).

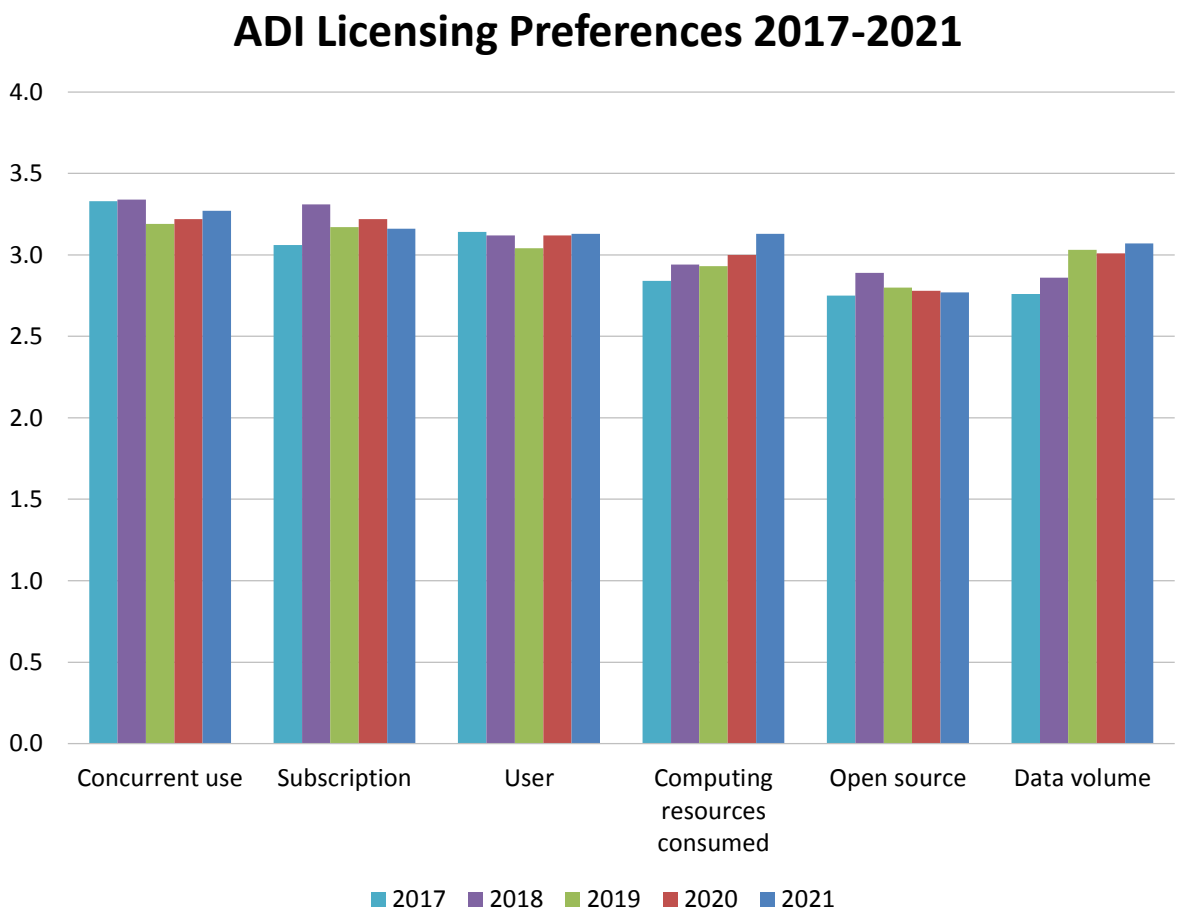


Figure 26 – ADI licensing preferences 2017-2021

When we look at the ADI platform responses on licensing priorities by top use case, we see limited variability in licensing model preferences, with the exception of "embedded analytics" use cases, with a higher preference for "data volume," "computing resources consumed," and "user" licensing models than all other use cases (fig. 27). For organizations preferring an open source licensing model, the top preferences are the data science and embedded analytics use cases and less so for "reporting and dashboards."

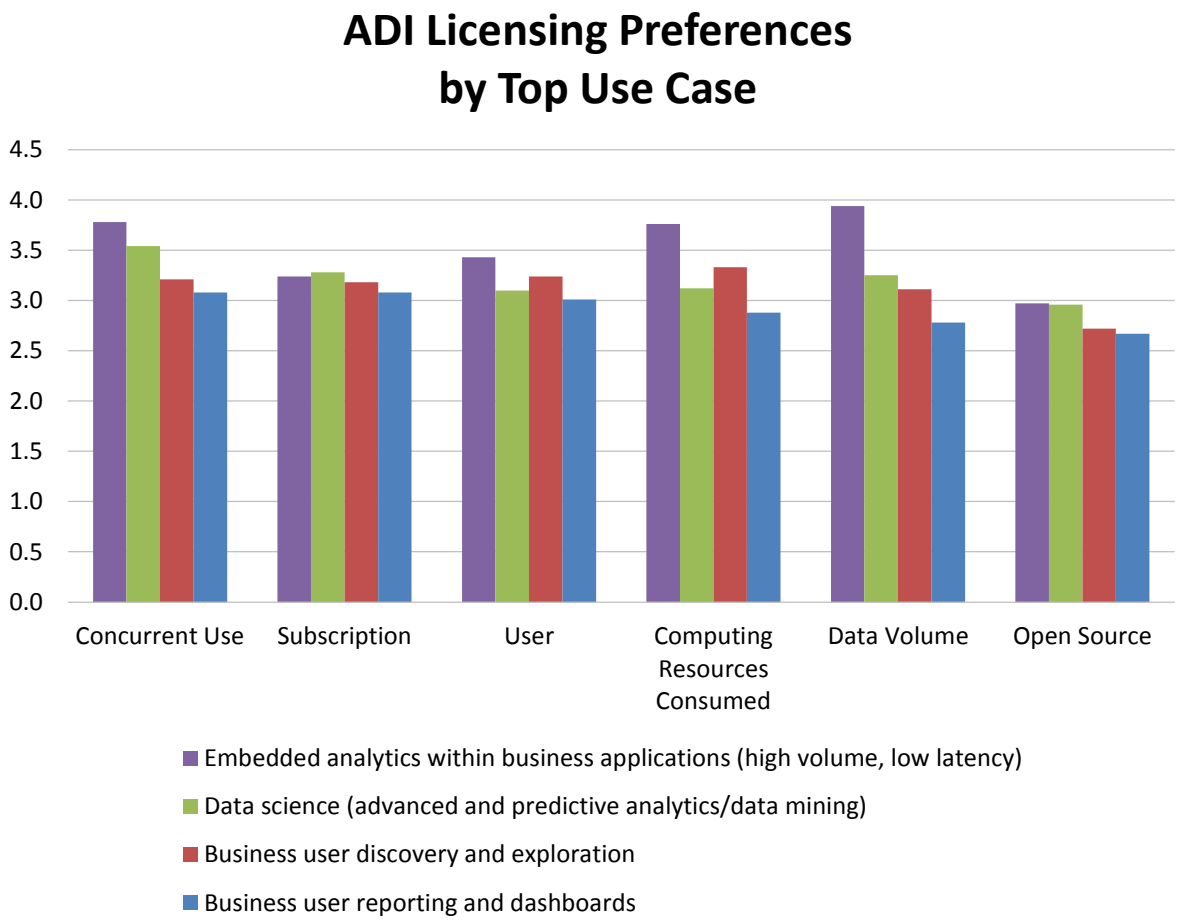


Figure 27 – ADI licensing priorities by top use case

Licensing preferences vary by geography (fig. 28). Asia-Pacific respondents tend to have a higher preference for consumption models and show higher preference for open source licensing models compared to other geographies. North America and Latin America favor “concurrent use” models, and EMEA favors a user-based model.

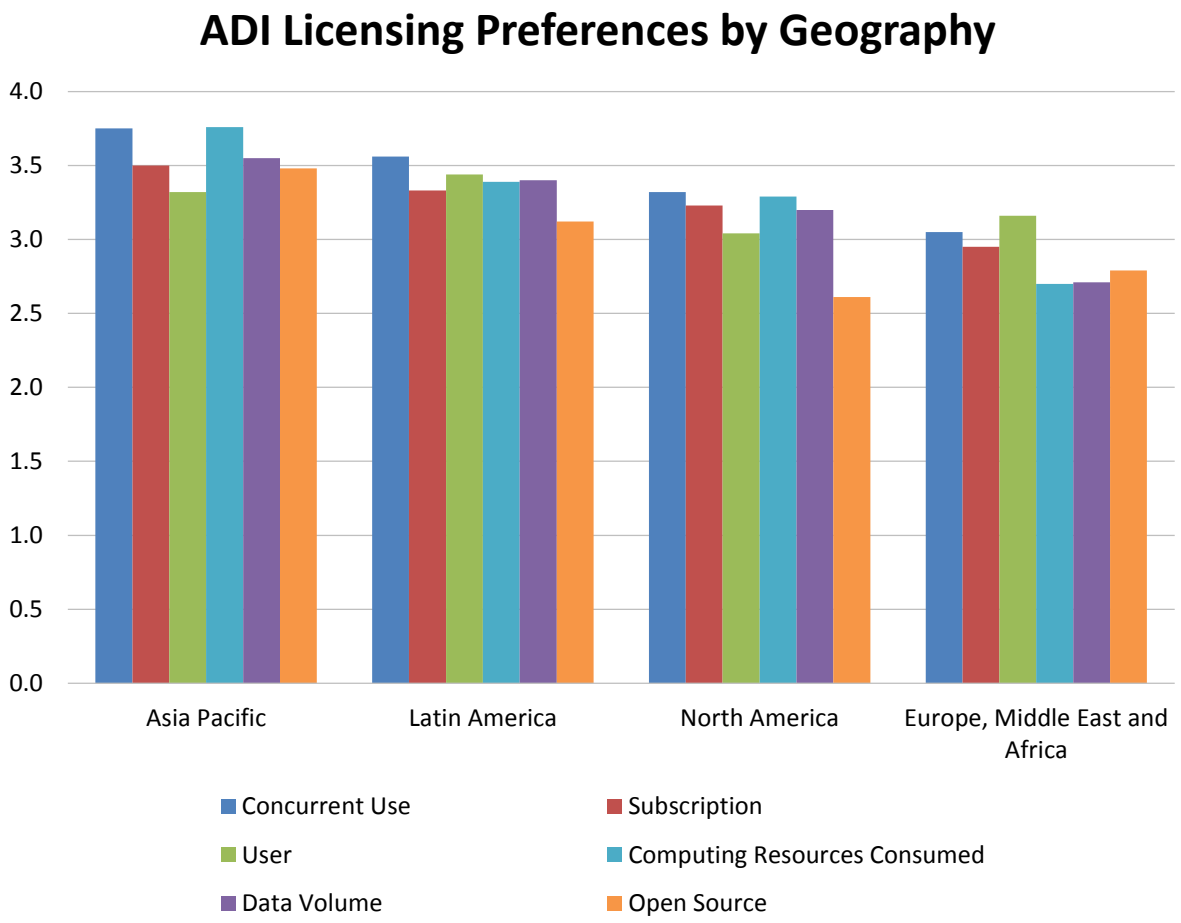


Figure 28 – ADI licensing preferences by geography

Licensing preferences vary by organizational functions; with the IT function having the least variation in responses regarding their preferences. Respondents from the Operations function have the highest preference for “subscription” and “user” licensing models (fig. 29). Surprisingly, open source is not the top licensing model for R&D functions, given their desire to modify source code of community-developed software to achieve a higher level of required customization and integration. Executive Management respondents indicate a preference for “concurrent use” and subscription licensing models. Respondents from the Finance organization prefer “user” and “concurrent use” licensing models.

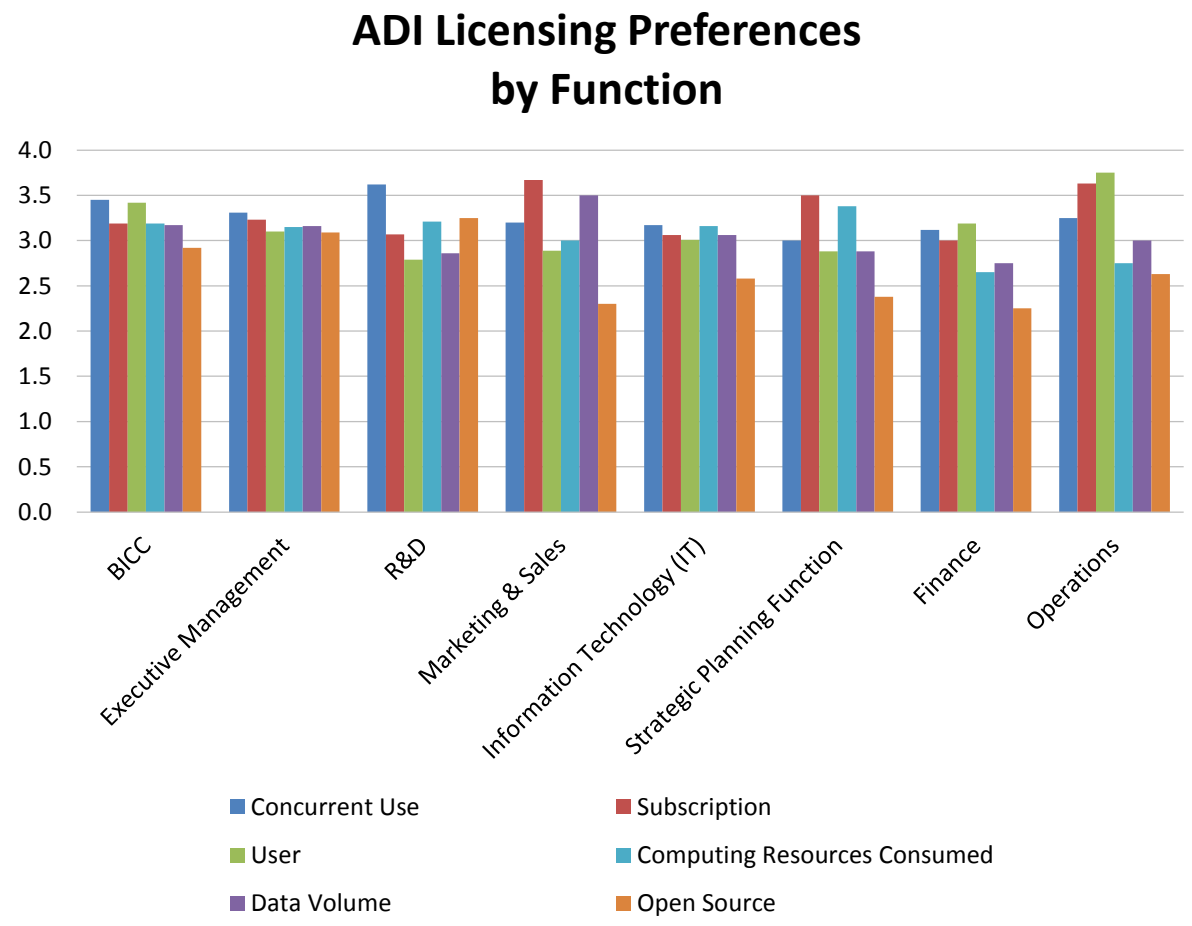


Figure 29 – ADI licensing preferences by function

ADI licensing preferences vary by industry. Concurrent-use and subscription-based licensing models are a relatively high preference across all industries, with the exception of Government, which has a stronger preference for open source than other segments (fig. 30).

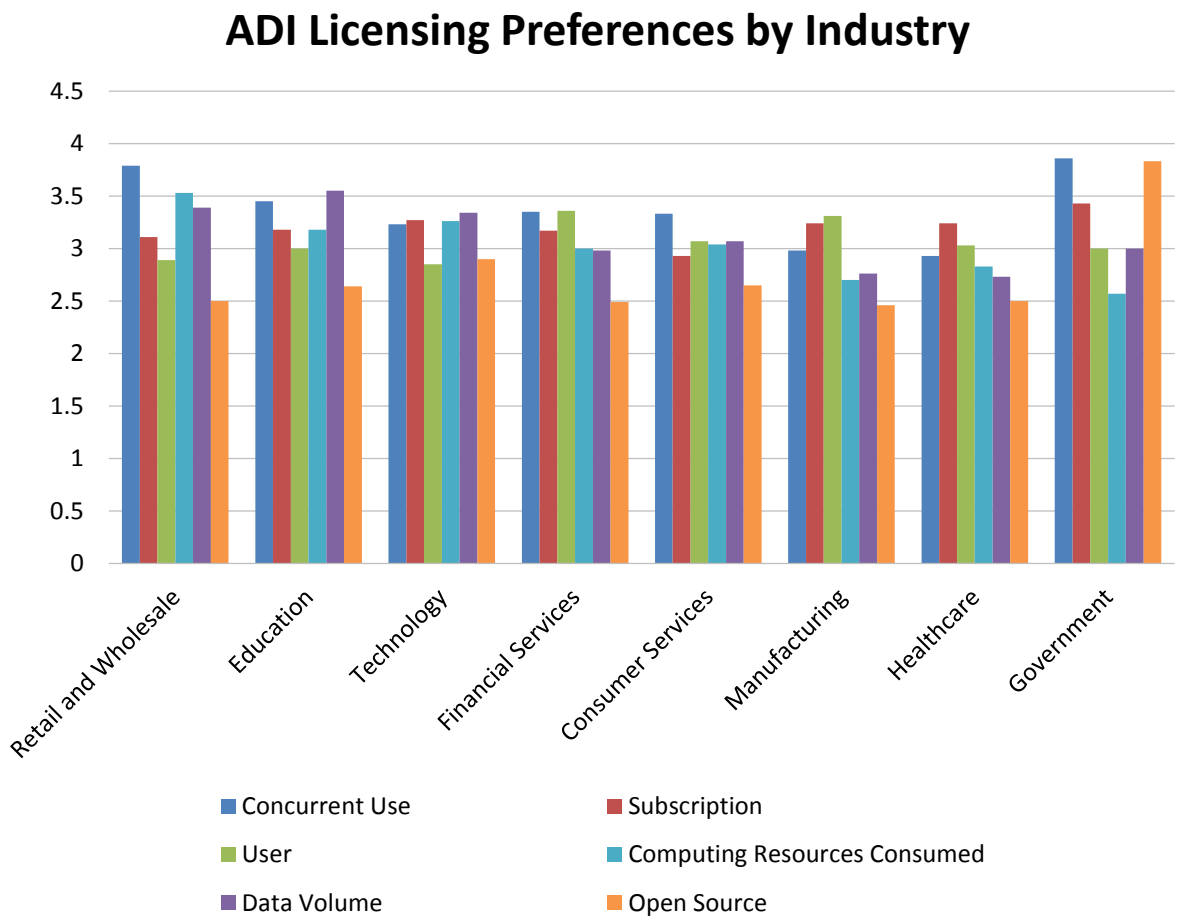


Figure 30 – ADI licensing preferences by industry

Not surprisingly, smaller organizations that show a preference for cloud deployments of their ADI platform (see fig 24) also rank subscription-based models as their licensing preference for ADI offerings (fig. 31). Larger organizations prefer every licensing model over open source licensing for their ADI platforms. Concurrent-use, data-volume, and subscription licensing give larger organizations more flexibility in administering their deployment since they do not have to identify and manage/administer large numbers of users as they do with a user-based licensing model.

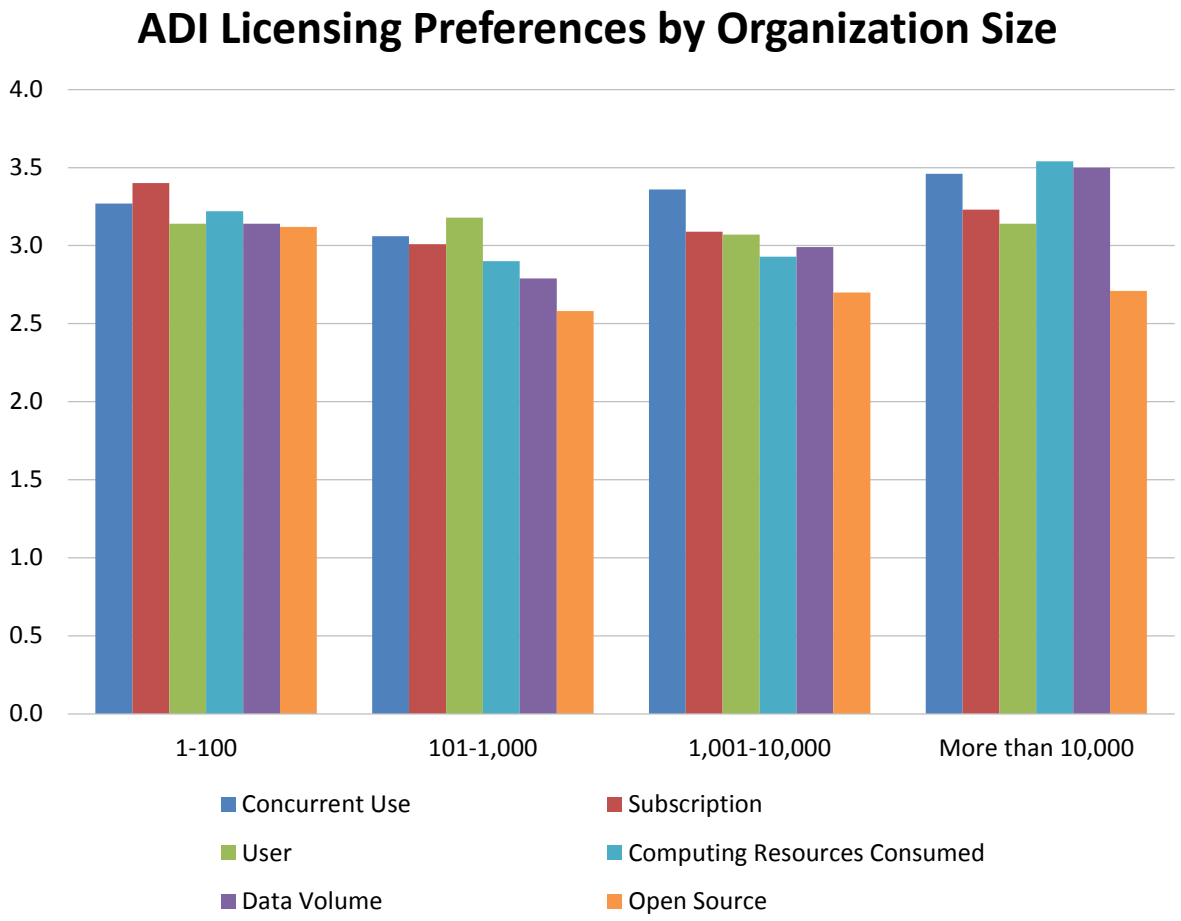


Figure 31 – ADI licensing preferences by organization sizes

Analytical Data Infrastructure Development and Deployment Features

We asked respondents to prioritize nine ADI development and deployment features. "Scale up and scale out" capabilities lead as the highest priority for ADI development features, followed by "data life cycle management" capabilities. Like last year, "pre-built data models" is the lowest priority development capability. Pre-built data models have a spotty history in analytic applications, given the difficulty of modifying/adapting the models 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. "Containerization and orchestration" is a new option for 2021.

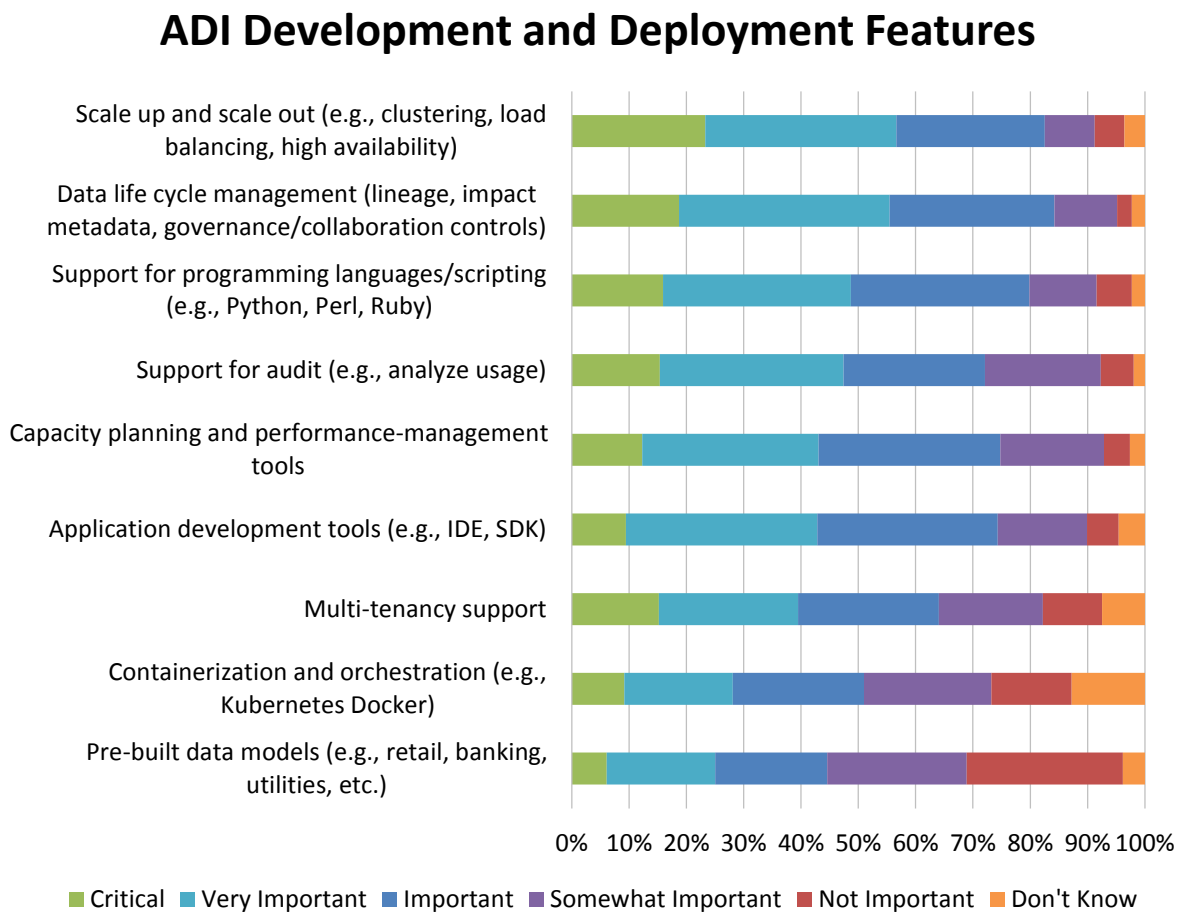


Figure 32 – ADI development and deployment features

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

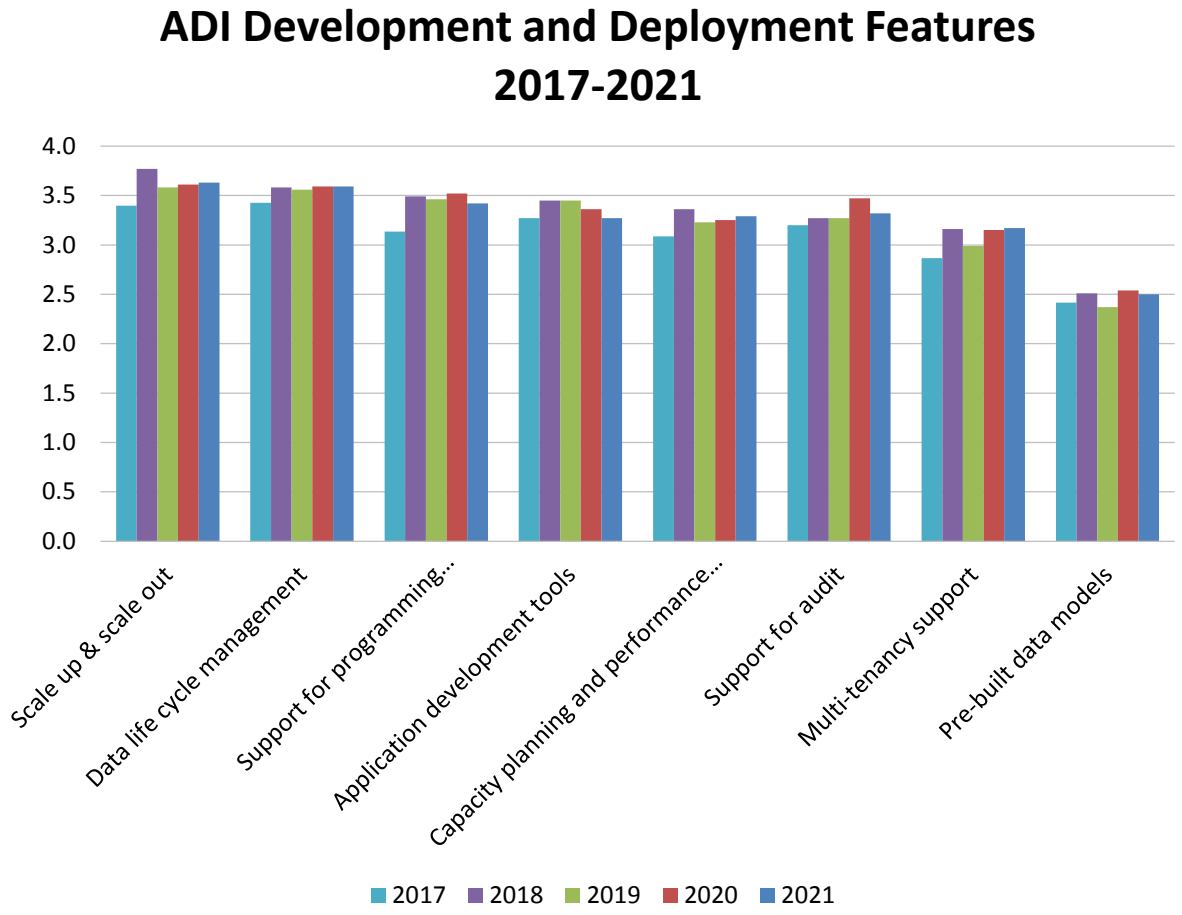


Figure 33 – ADI development and deployment features 2017-2021

"Scale up and scale out" capabilities are the top development and deployment priority for ADI platforms regardless of use case, with the exception of "business user reporting," where respondents rate data life cycle management slightly higher. When we look at the results by specific use cases, we see a few trends of note (fig. 34). Of note, the embedded analytics use case makes multi-tenancy support a priority while data science places a higher-than-average priority on support for programming languages.

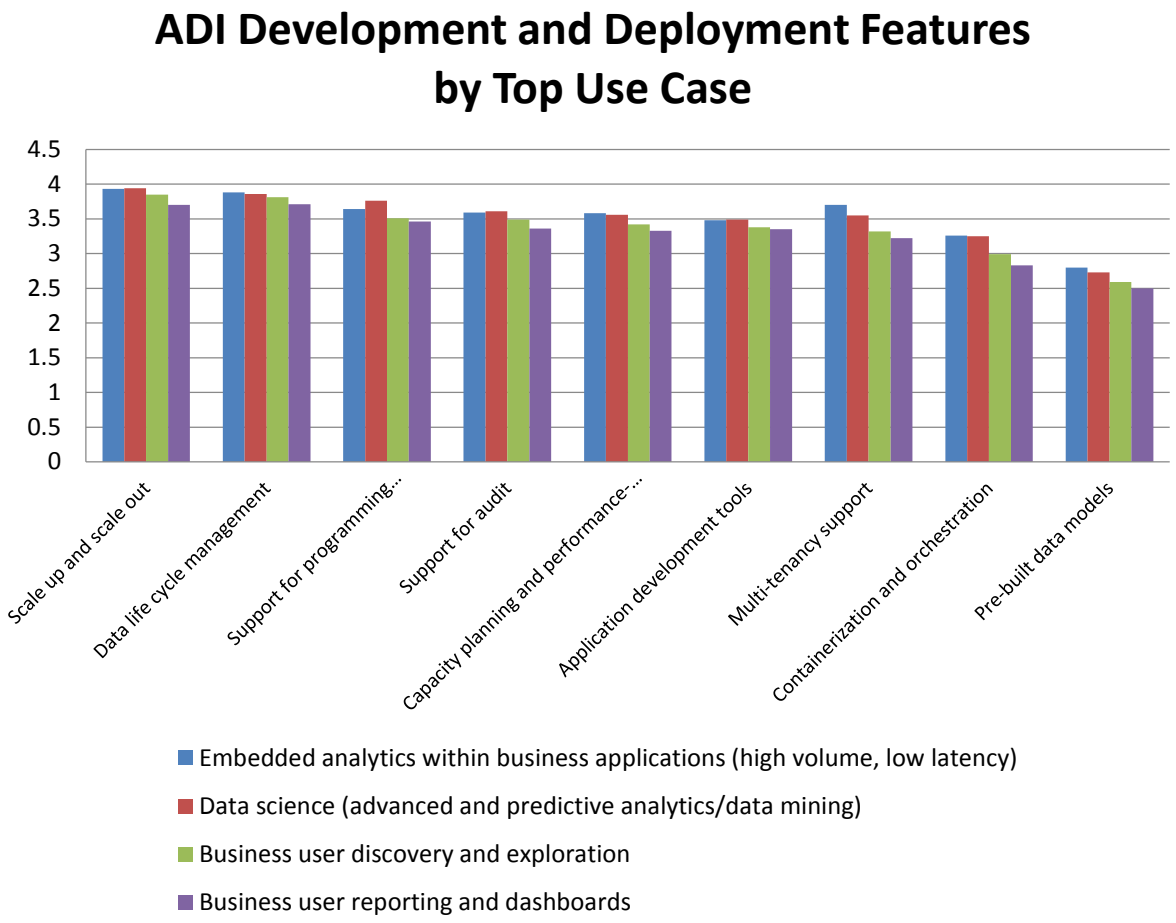


Figure 34 – ADI development and deployment features by top use case

The priority of ADI development and deployment features can vary quite a bit when we look at the responses by geography (fig. 35). The top priority for Latin America, and a close second for North America respondents, is “data life cycle management.” 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.

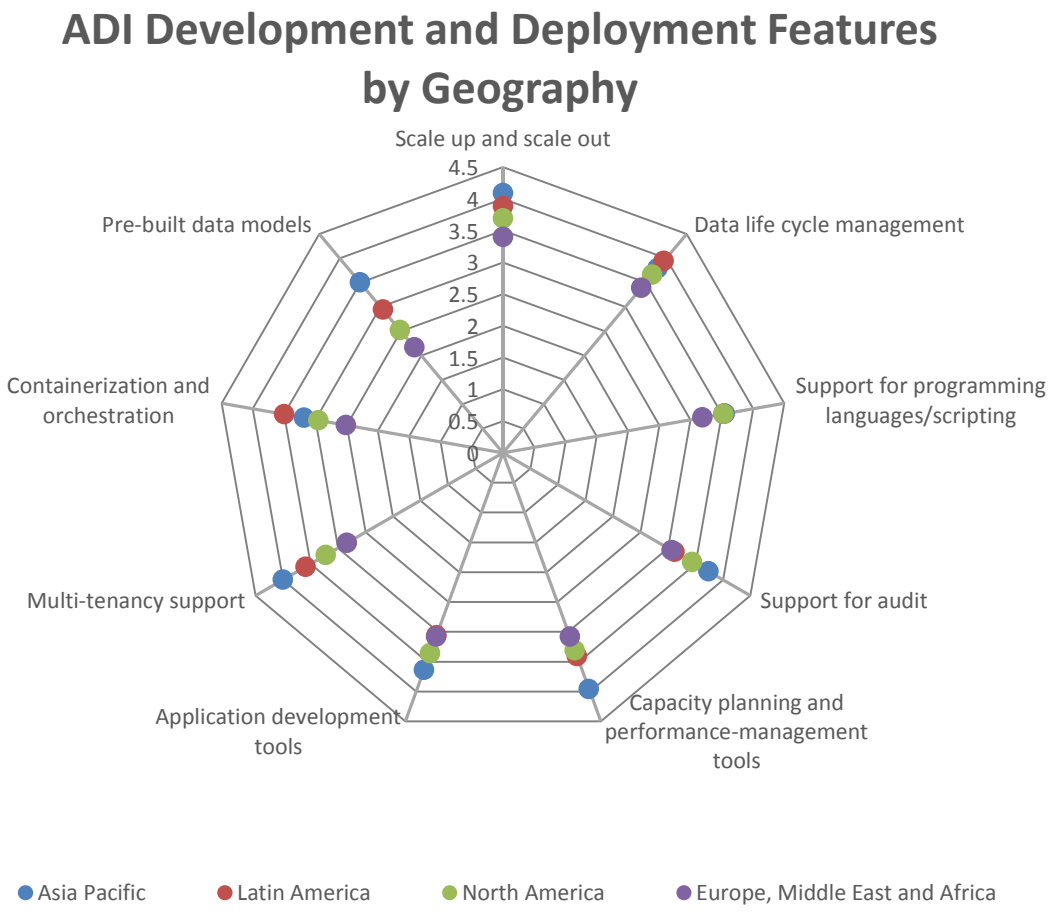


Figure 35 – ADI development and deployment features by geography

ADI development and deployment features vary depending on organizational function of the respondent (fig. 36). Most functions, with the exception of Marketing, place the highest priority on “scale up and scale out” features. The priority of “data life cycle management” tools is fairly high across all organizational functions, indicating a possible increase in understanding for the need for these features, especially when it comes to support for multiple data sources, multiple users and use cases, and security controls. Strategic Planning, BICC, and IT place a greater priority on application development tools,” while Finance respondents show a high interest in “capacity planning and performance management tools.”

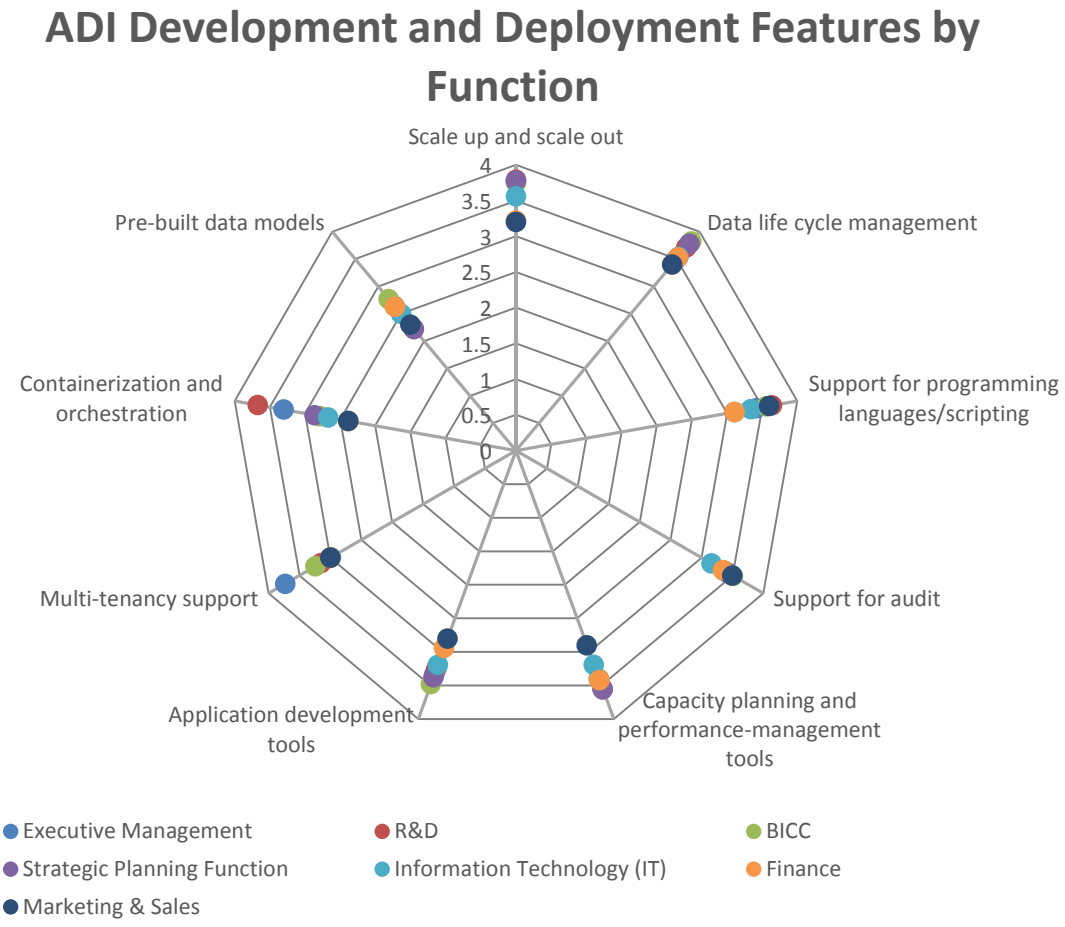


Figure 36 – ADI development and deployment features by function

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 (fig. 37), followed by “data life cycle management.” Financial Services respondents place a higher-than-average priority on “support for audit” development and deployment features, while Government favors application development tools. Across several years and across all industries, there is consistently low priority for pre-built data models.

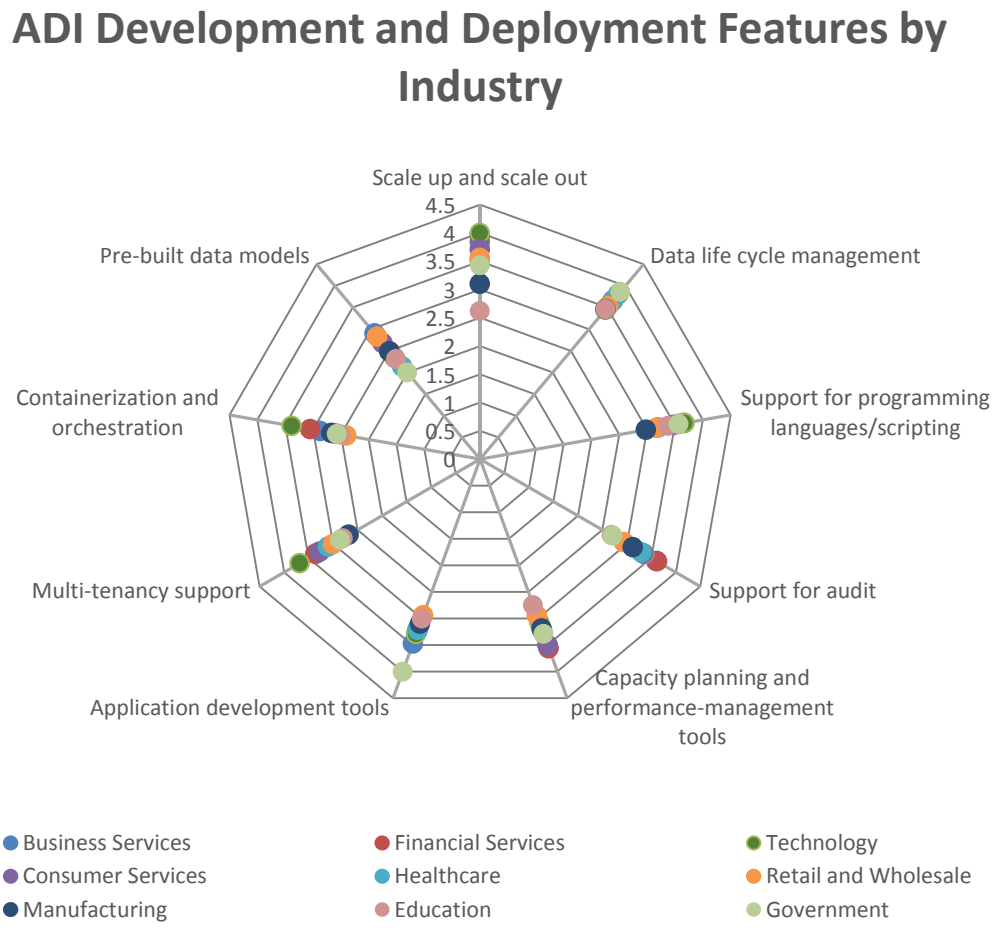


Figure 37 – ADI development and deployment features by industry

When segmented by organization size, the priorities for ADI development and deployment features show relatively low variability. In general, larger organizations place a higher priority on development and deployment features compared to smaller organizations. Larger organizations with 1,001 or more employees place high priority on "data life cycle management" and "scale up and scale out" capabilities. Organizations with more than 10,000 employees place significant value on "containerization and orchestration." Across all companies, regardless of size, "pre-built data models" are the lowest priority (fig. 38).

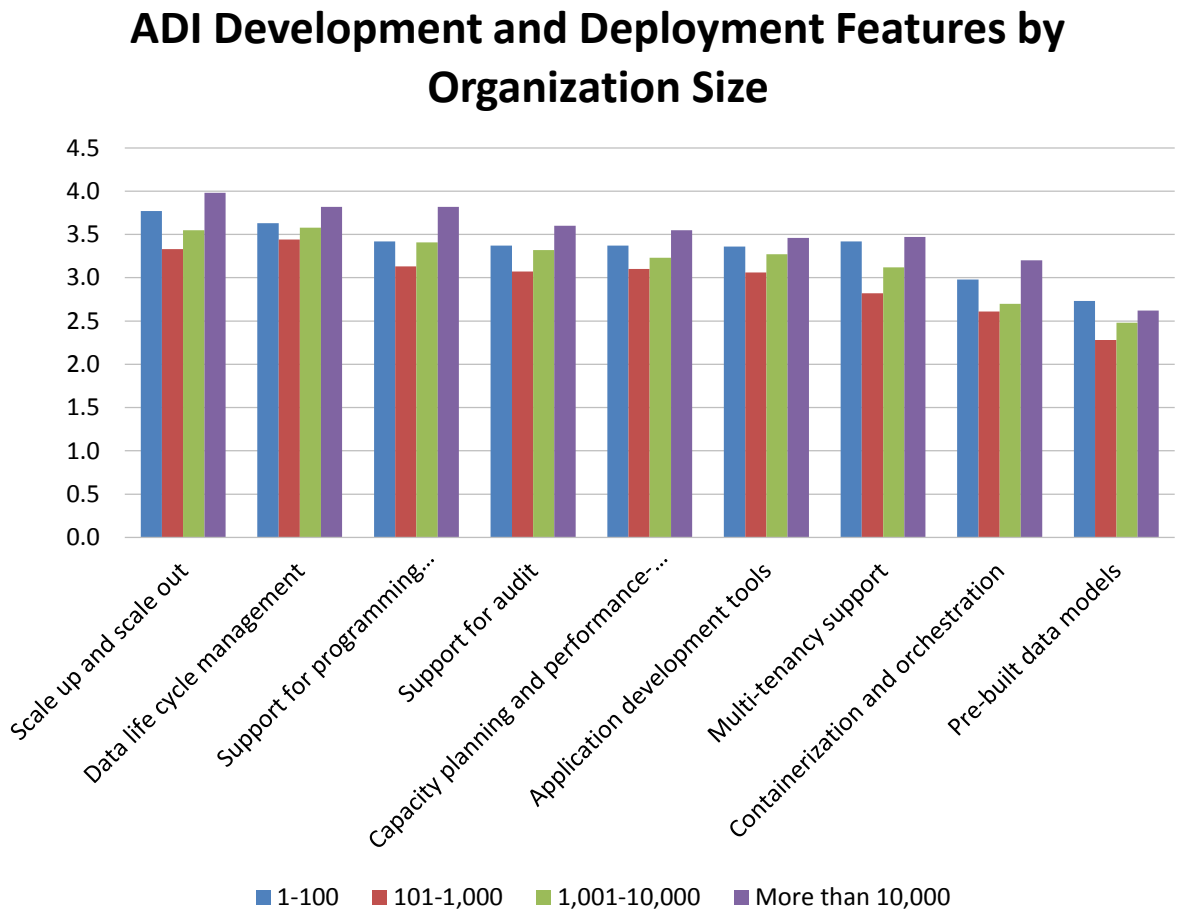


Figure 38 – ADI development and deployment features by organization size

Analytical Data Infrastructure Data Types

More than 84 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 Excel/CSV data (fig. 39) and metadata. 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 general use, images and video data type support was, historically a lower priority for ADI applications / use cases; but this year's responses show this as a higher priority for ADI platforms. It is worth noting that "machine and events / log data" are an important priority for a large number of analytic use-case workflows/workloads, with over 50 percent rating this as either critical or very important.

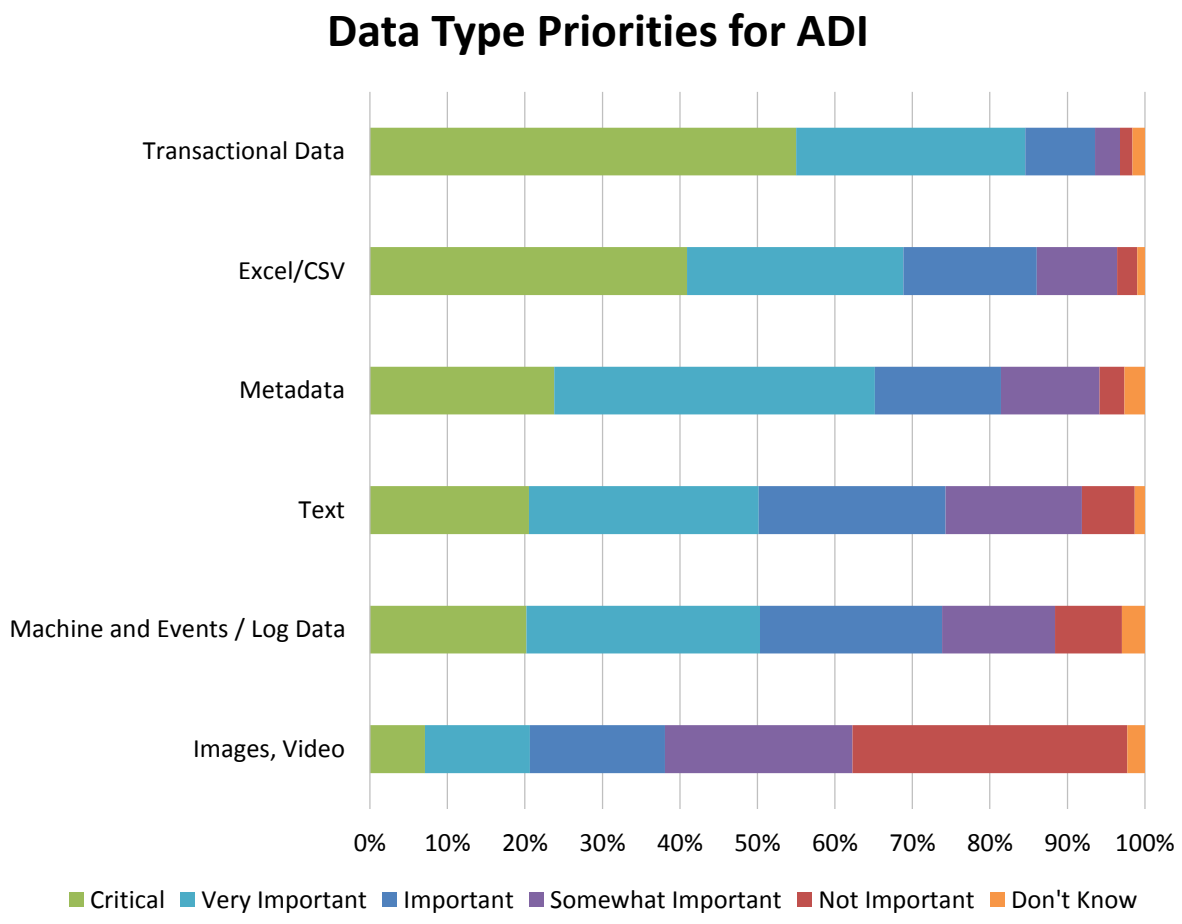


Figure 39 – Data type priorities for ADI

The relative priority of data types supported by an ADI platform changes little between 2020 and 2021 survey results. It is interesting to note the importance of the metadata capabilities of an ADI is equal to the importance of an ADI's ability to support Excel data types (e.g., import / export / manage Excel data that is used for discovery and exploration tools) (fig. 40). Metadata is needed as organizations scale up/out their implementation and focus more on data life cycle management. Images and video data types remains the lowest data type priority, however, over 20 percent rate this as critical or very important.

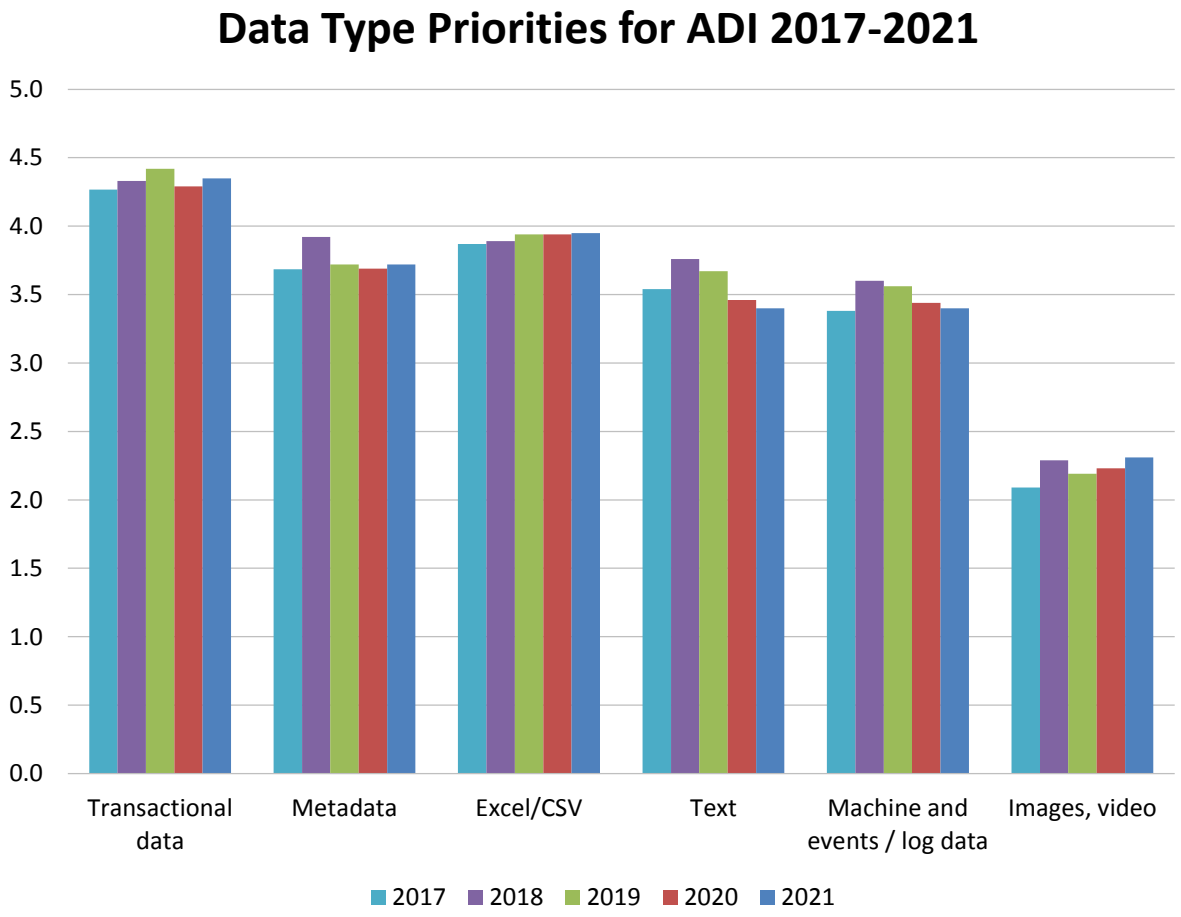


Figure 40 – Data type priorities for ADI 2017-2021

The top data type priority for all ADI use cases is still transactional data. The lowest priority data type across all ADI use cases is "images and video," although over 20 percent of respondents rate this capability as either critical or very important (fig. 41). Embedded analytics use cases place a higher priority on "text" data types compared to other use cases, while data science and machine learning place a higher priority upon Excel/CSV and metadata.

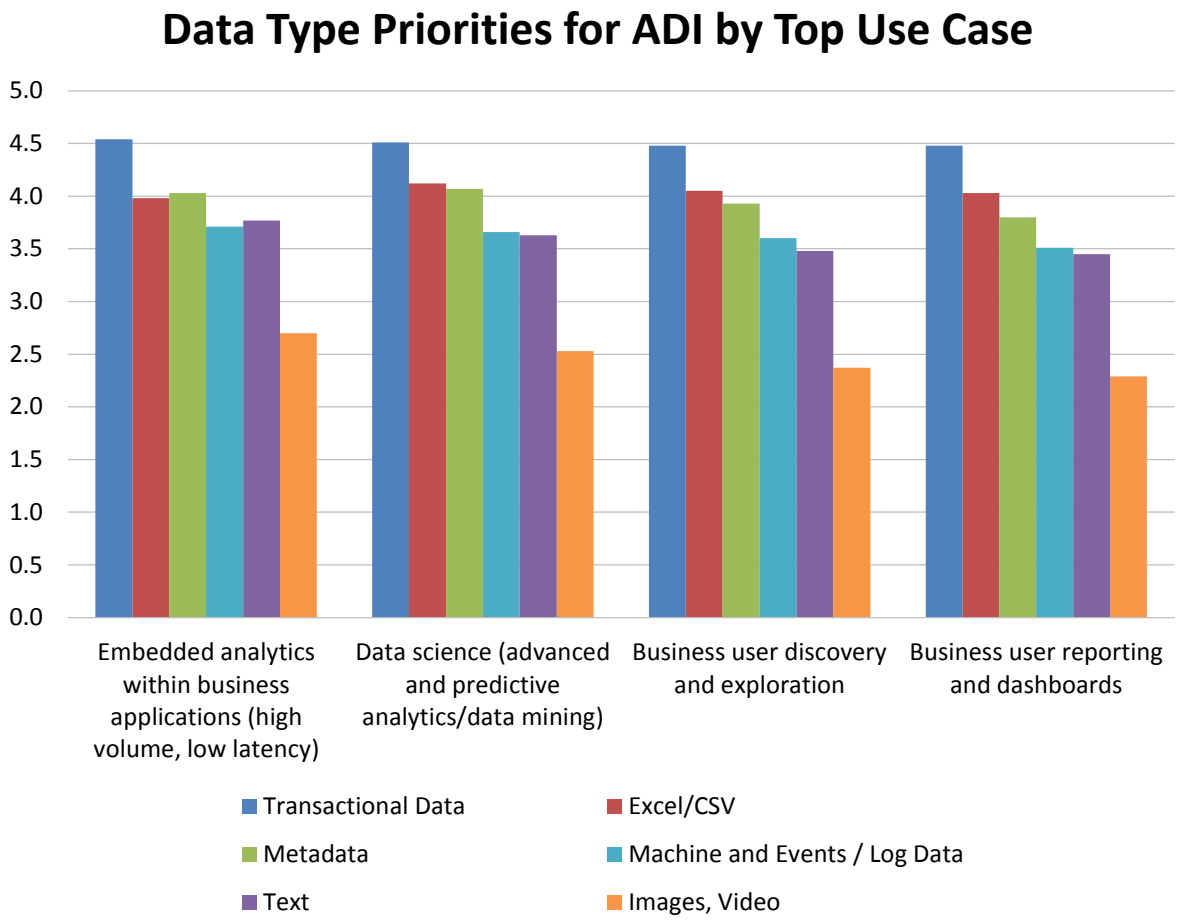


Figure 41 – Data type priorities for ADI by top use case

The relative priority of data types is similar across geographies. Transactional data is the highest priority data type across geographies (fig. 42), with the exception of Asia Pacific, where respondents rate Excel/CSV highest. All geographies rank “images and video” as a low priority data type to access / manage / analyze as a part of the ADI platform functionality.

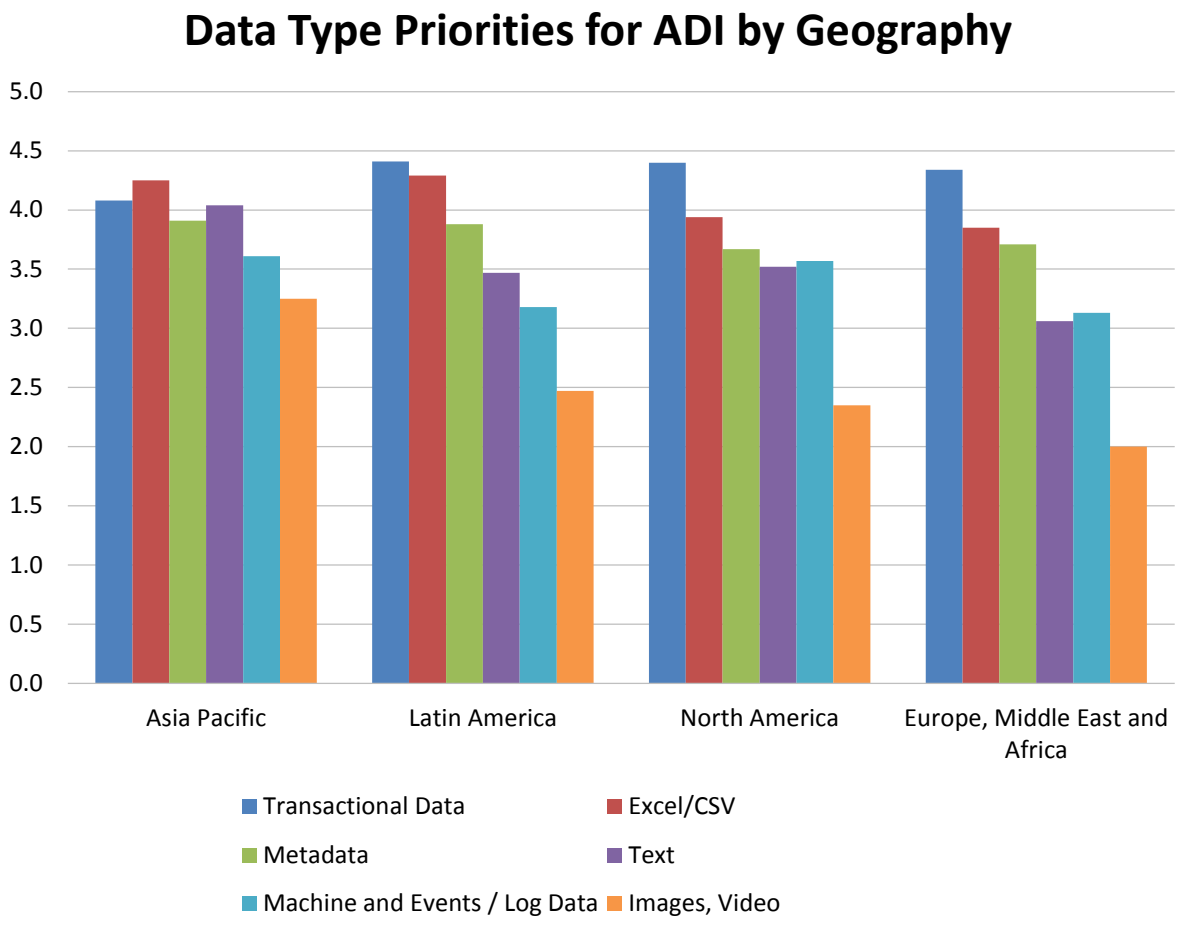


Figure 42 – Data type priorities for ADI by geography

The priority of ADI data type capabilities varies by respondents' function in their organization. For a majority of functions, respondents rate Excel/CSV data type capabilities as a priority, following transactional data (fig. 43). For Strategic Planning, R&D, BICC, and IT, metadata is the third-highest priority. Across all functions, image data holds the lowest 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, including metadata and its role in data life cycle management and security/governance requirements.

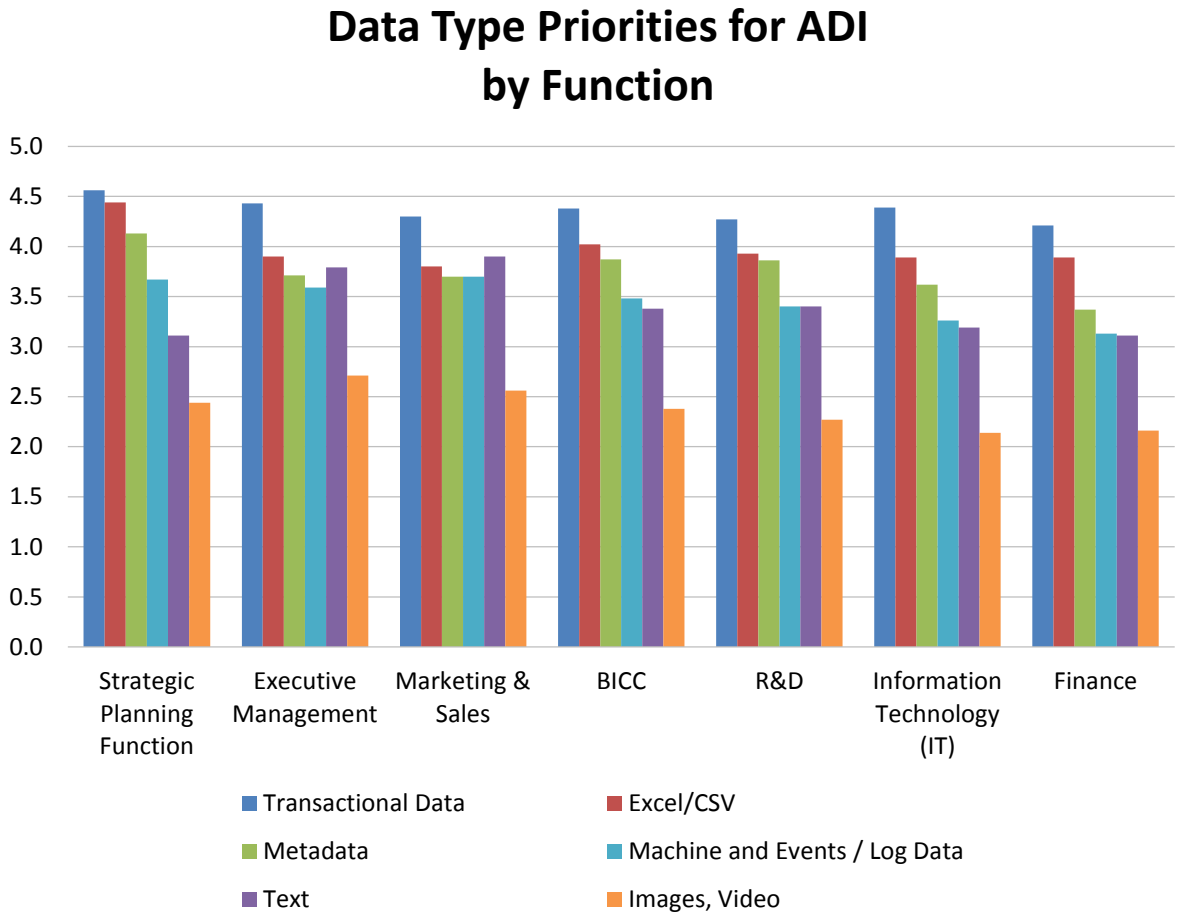


Figure 43 – Data type priorities for ADI by function

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 "transactional data" and "Excel/CSV" are the highest priority and images/video is among the lowest priorities (fig. 44). The variety in data types by industry reflects the variation of different analytical workflows and workloads by industry.

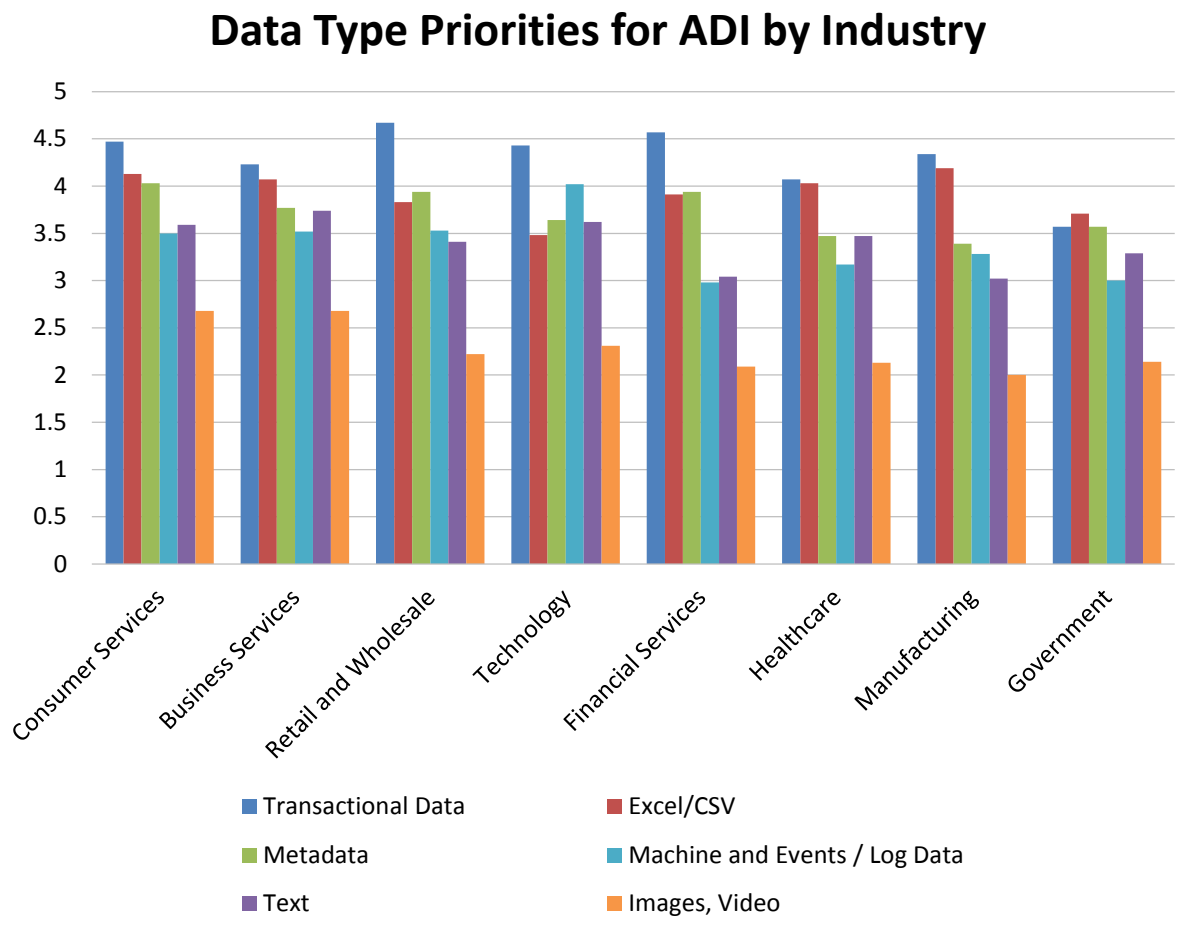


Figure 44 – Data type priorities for ADI by industry

Organizations of all sizes indicate their top priority is support for transactional data types, followed by Excel/CSV and metadata (fig. 45) when comparing requirements for their ADI platform. Support for images/video is a relatively lower data type priority for all organization sizes. 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) that require the use of metadata.

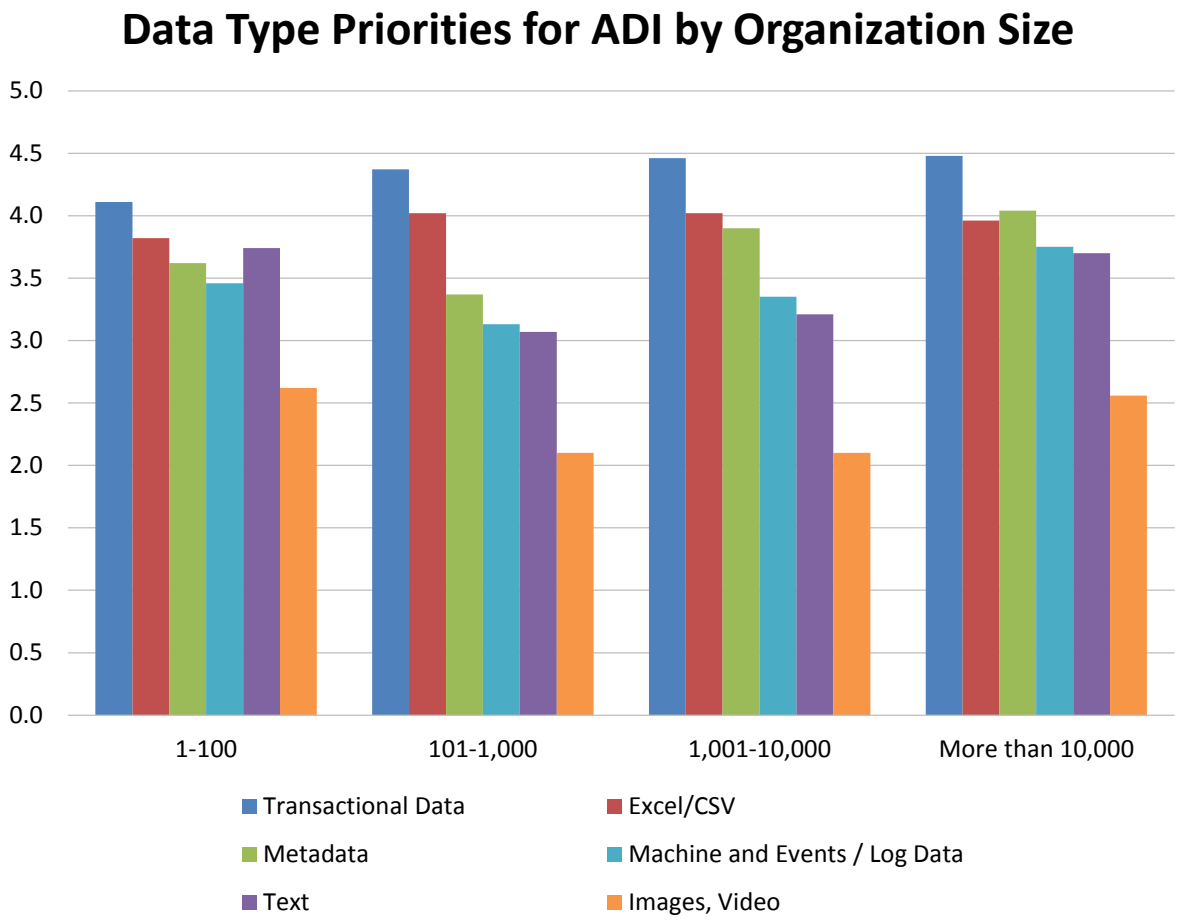


Figure 45 – Data type priorities for ADI by organization size

Analytical Data Infrastructure Data Preparation and Loading

Earlier in this report, we note increasing growth in preferences for "cross data center integration and management" capabilities. This also reflects the relative priority of data integration as the second-highest strategic technology across BI use cases. 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 to support (fig. 46). Slightly more than 49 percent indicate "real-time / streaming, trickle, increments / change capture" is a priority for data prep and loading of ADI. Only 24 percent of respondents consider support for Apache big data services critical or very important, a drop from 2020 (27%).

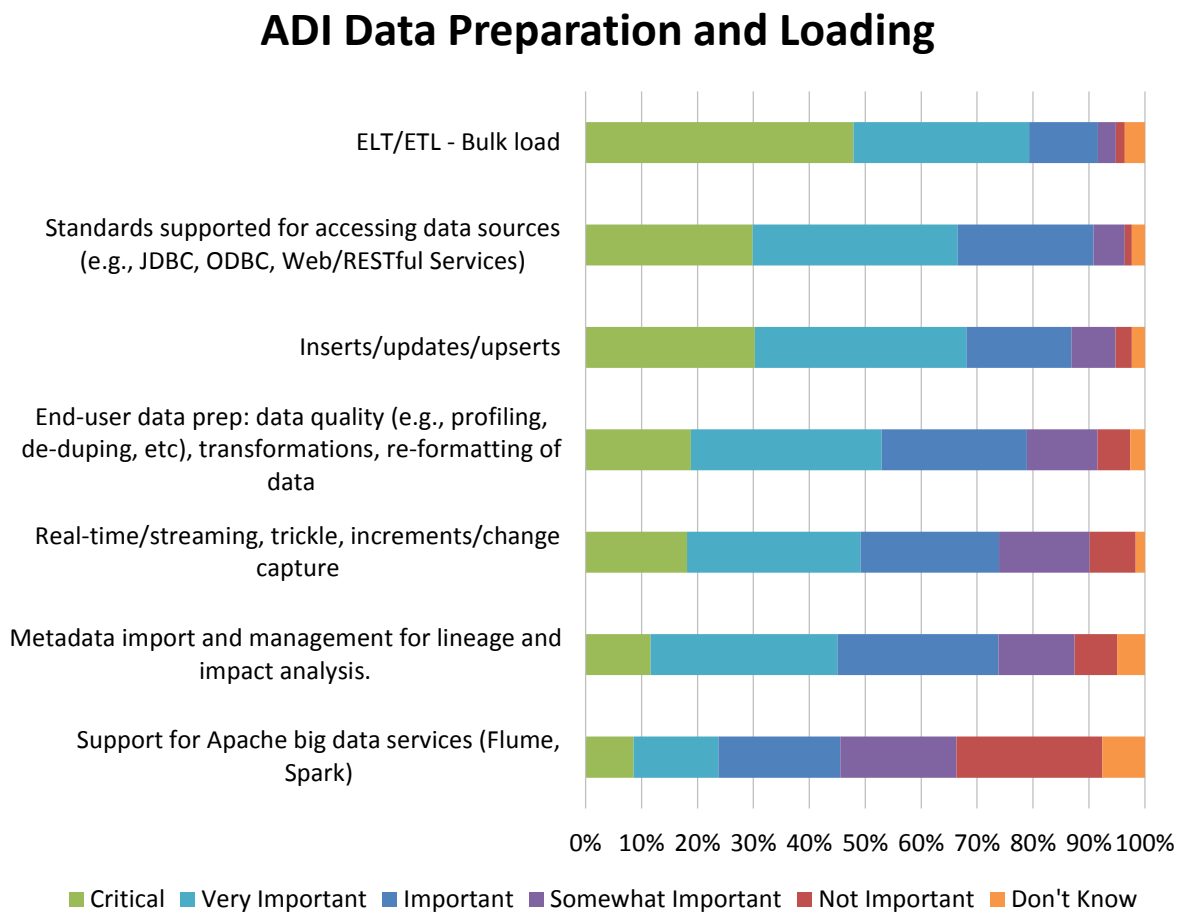


Figure 46 – ADI data preparation and loading

The relative priority of data preparation and loading features did 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, which could be a problem in the future for governance and data management projects/practices. However, it increased slightly since last year, as did "end-user data prep: data quality (e.g., profiling, de-duping, etc.), transformations, re-formatting of data."

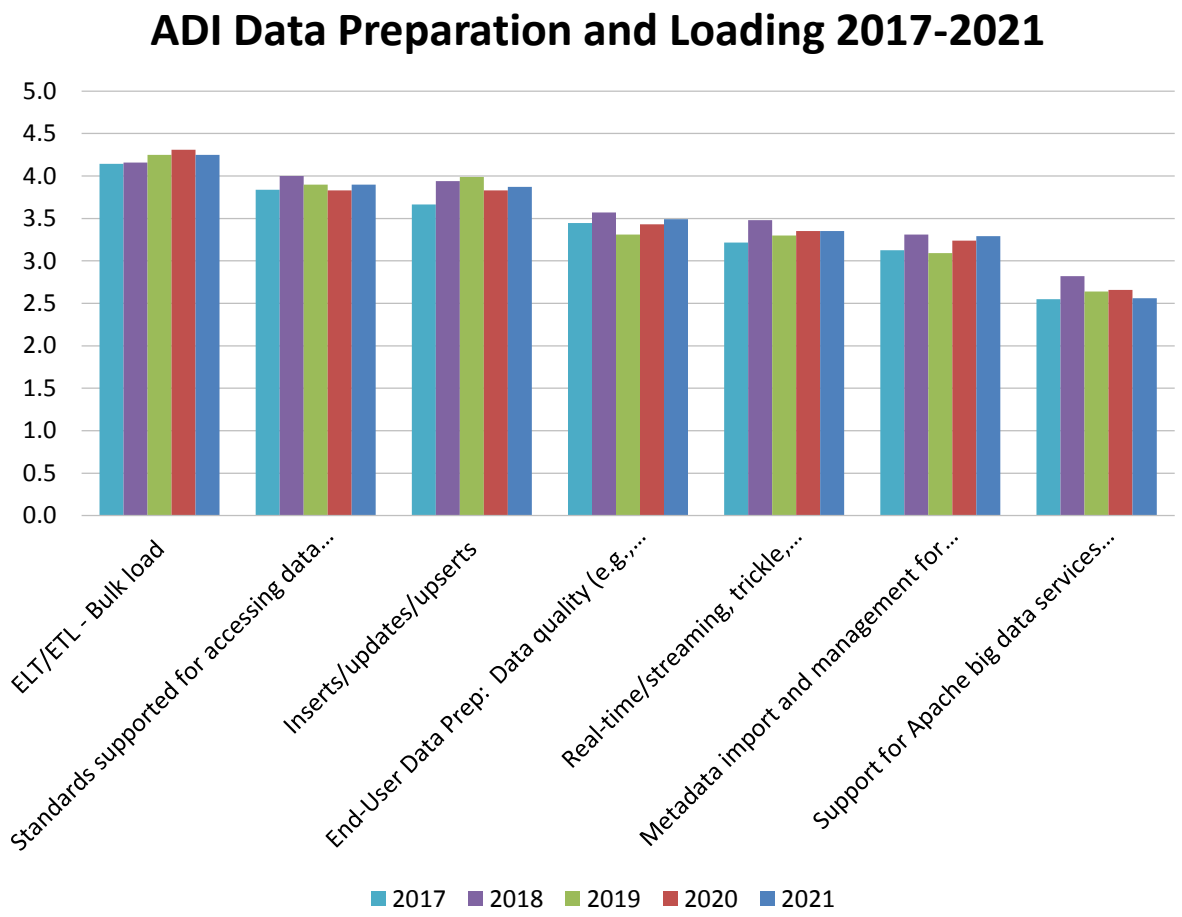


Figure 47 – ADI data preparation and loading 2017-2021

All regions indicate ELT/ETL and bulk load capabilities as their highest priority for their ADI platform to support; their second priority is the support for standard interfaces (JDBC, ODBC, etc.) (fig. 48). The use of inserts / updates / upserts is more important in Asia Pacific than in other geographies. Support for Apache big data services ranks lowest in priority across all geography respondents.

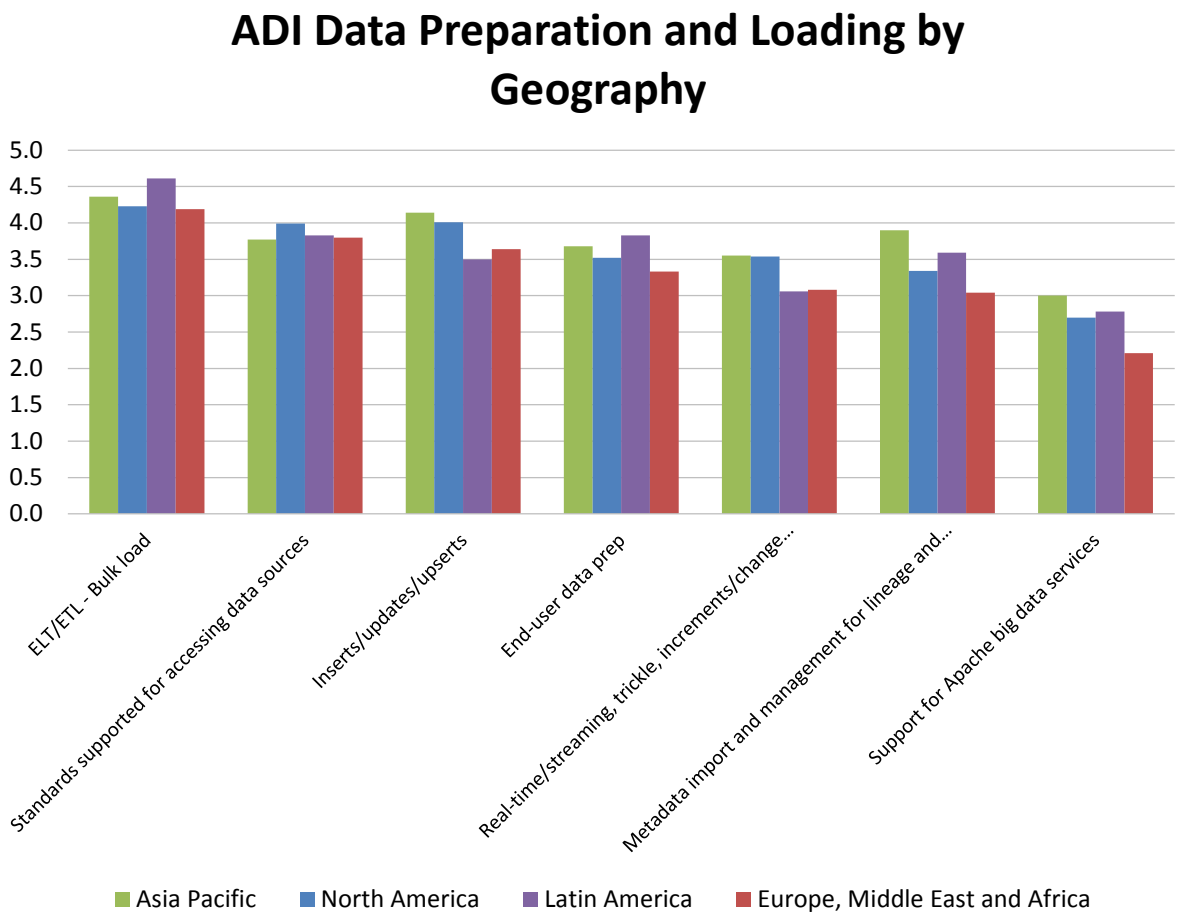


Figure 48 – ADI data preparation and loading by geography

For data preparation and loading, respondents from all use cases prioritize bulk ELT/ETL capabilities over all others. Embedded use cases, which often includes analytical features within operational applications, places a high priority upon "inserts/updates/upserts." Data loading from Apache big data services is the lowest priority for "reporting and dashboards," even for "data science" use cases.

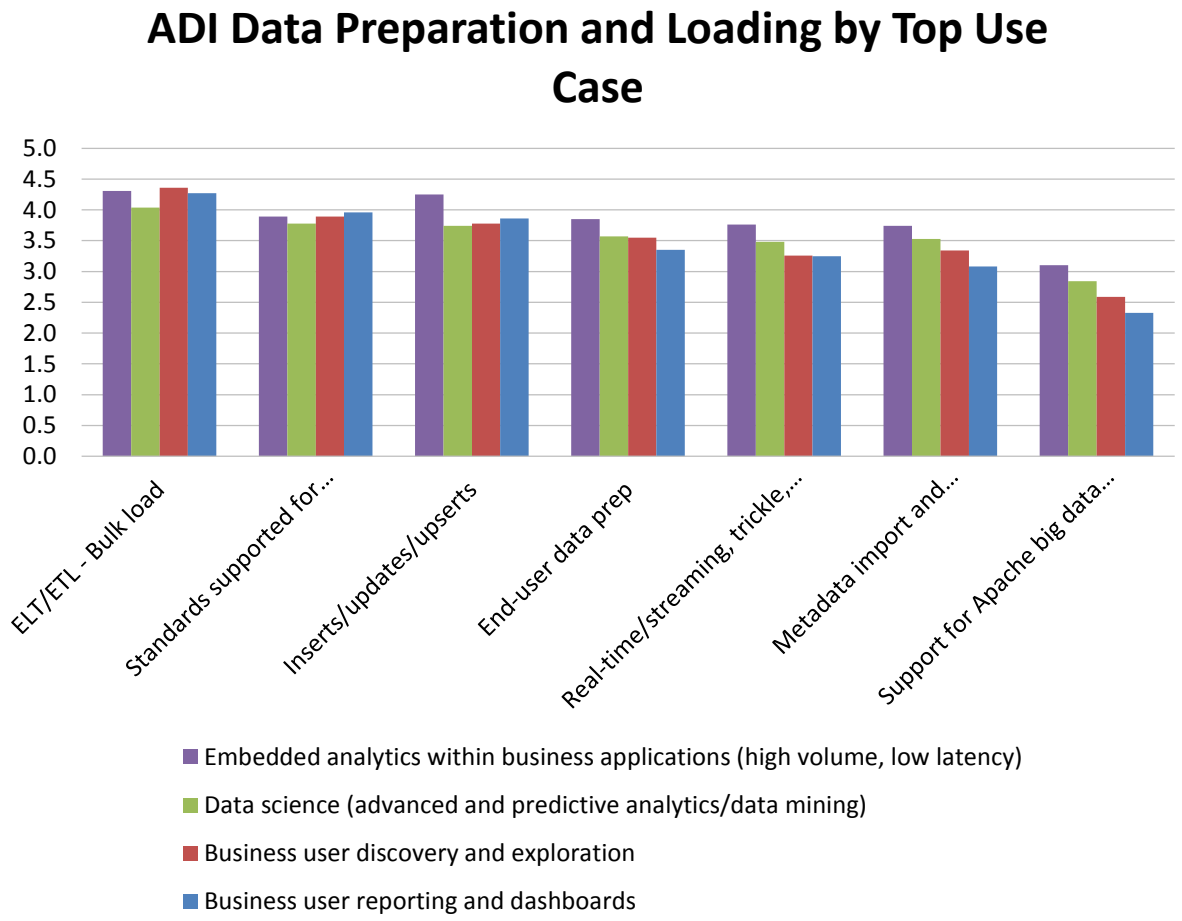


Figure 49 – ADI data preparation and loading by top use case

With a few exceptions, most functions have similar relative priorities on data preparation and loading functions, favoring ELT/ETL bulk load (fig. 50). R&D places a higher priority on real-time streaming content and big data integration services, while the Strategic Planning function favors metadata import and standards support.

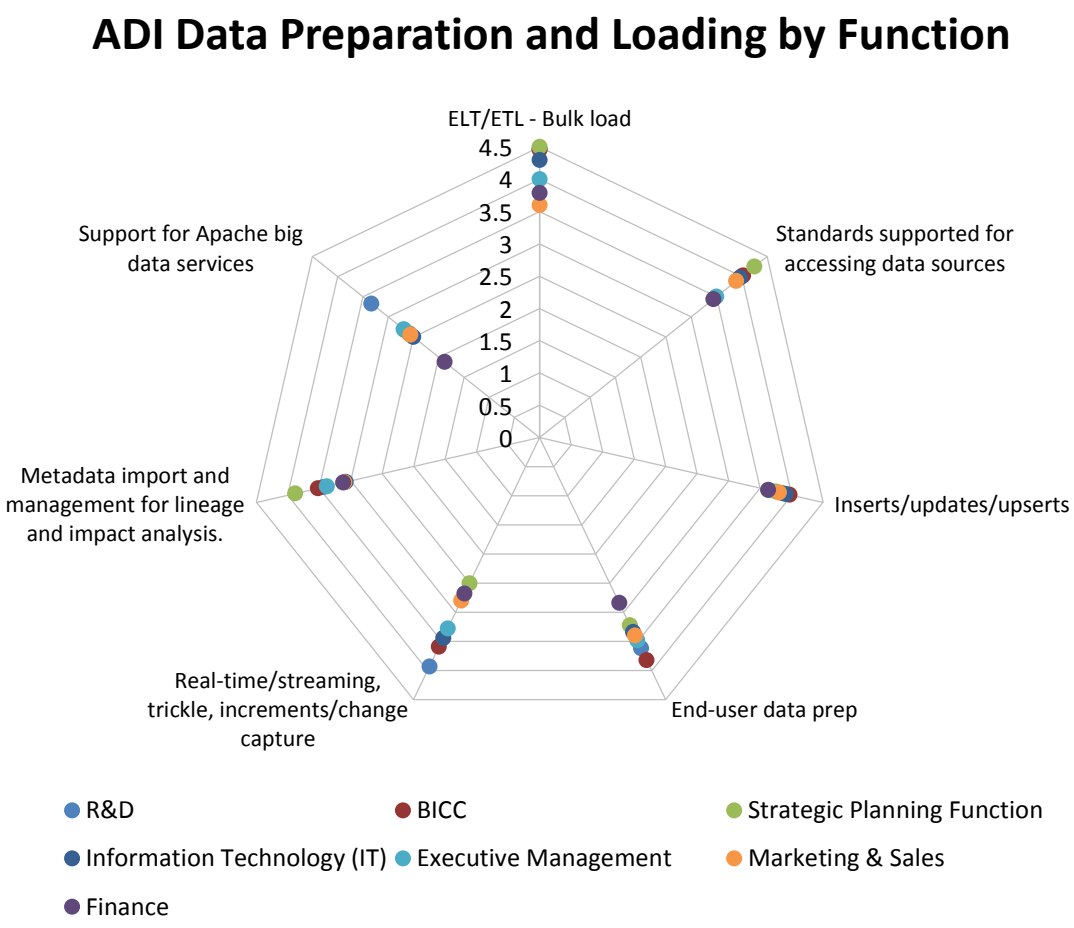


Figure 50 – ADI data preparation and loading by function

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, with the exception of Technology, which rates "real-time/streaming, trickle, increments/change capture" higher (fig. 51).

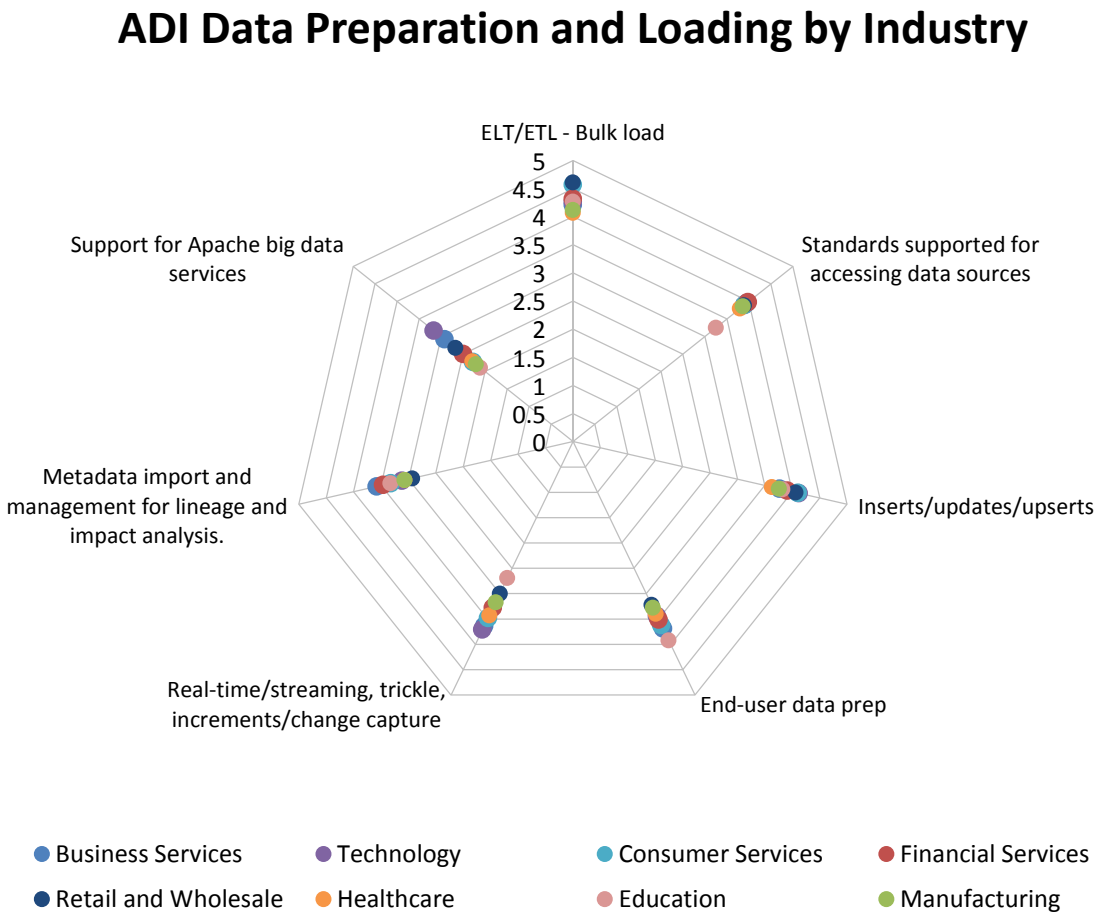


Figure 51 – ADI data preparation and loading by industry

We note only small differences in data prep and loading priorities by organization size (fig. 52). Larger organizations place a higher priority on all data preparation and loading capabilities. This is likely due to their higher priority in data life cycle management as a priority among development and deployment preferences (see fig. 32). “Support for Apache big data services” is a somewhat higher priortiy for the largest organizations.

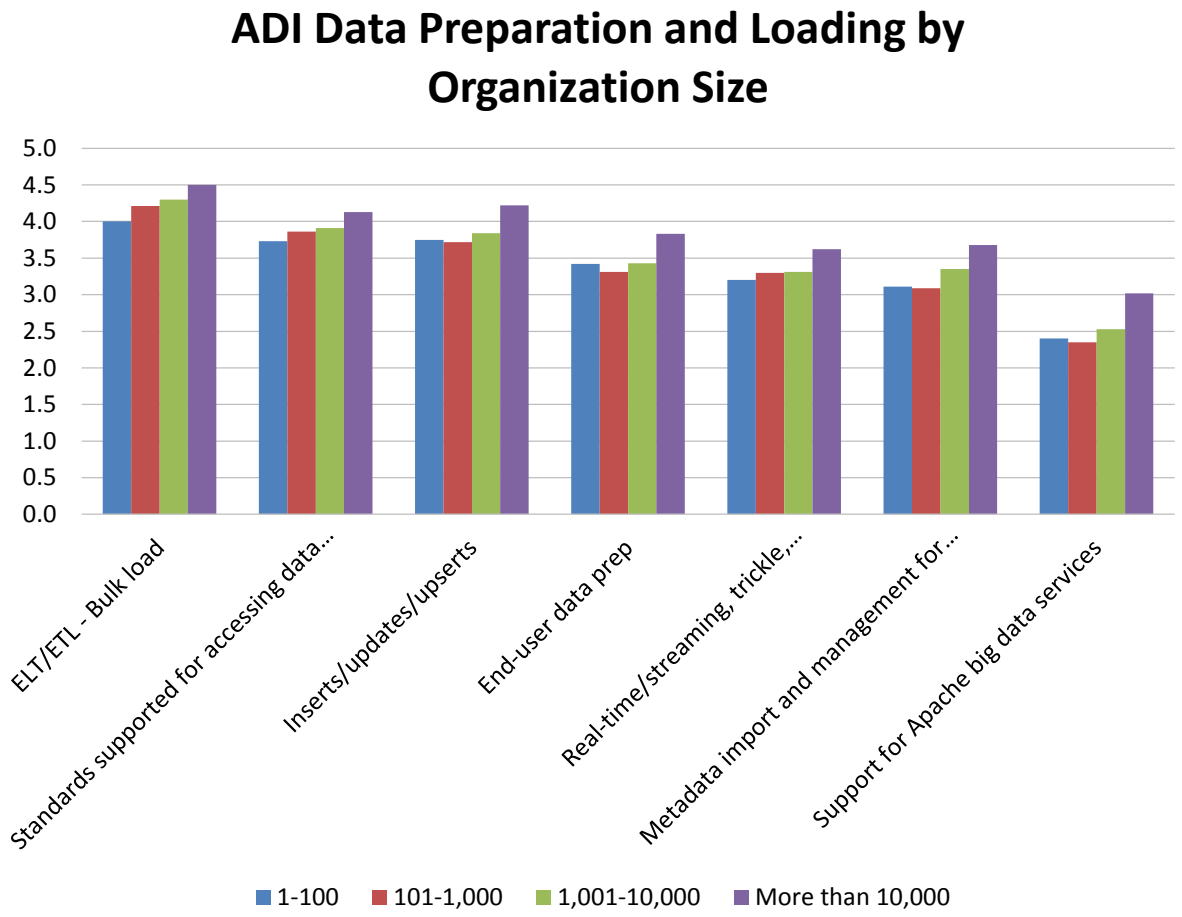


Figure 52 – ADI data preparation and loading by organization size

Analytical Data Infrastructure Data Model / Management of Data

SQL data capabilities are, by far, the top data model / management priority for an ADI platform (fig. 53). This matches the top priority respondents place on transactional data (see fig. 39).

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 35 percent of respondents.

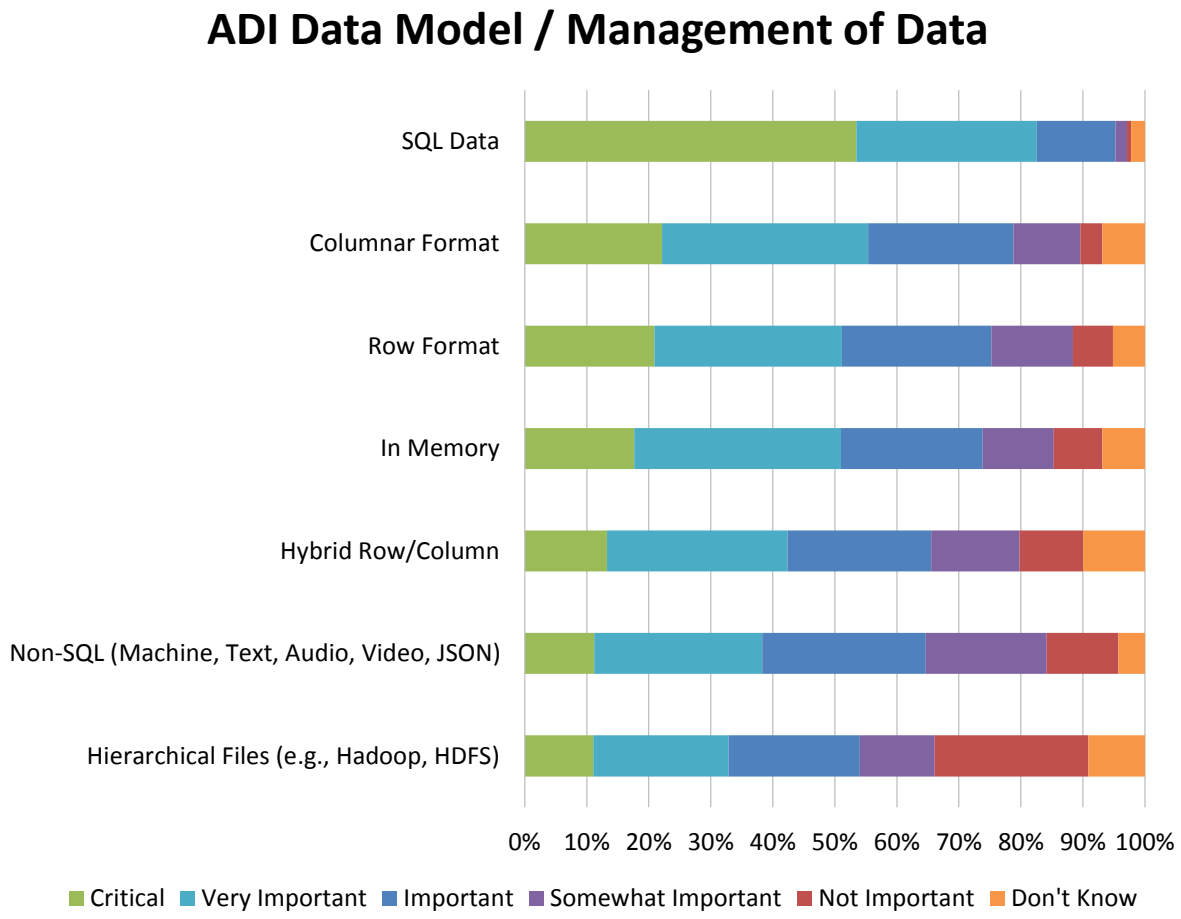


Figure 53 – ADI data model / management of data

Hierarchical files / HDFS capabilities remains the lowest relative priority for ADI data models/data and, although showing an increase in interest this year, declines in importance (fig. 54). Columnar data model preference remains relatively constant and surpasses row-format-based models for organizing analytic data. The priority of in-memory capabilities remains steady as a third top priority following row and columnar.

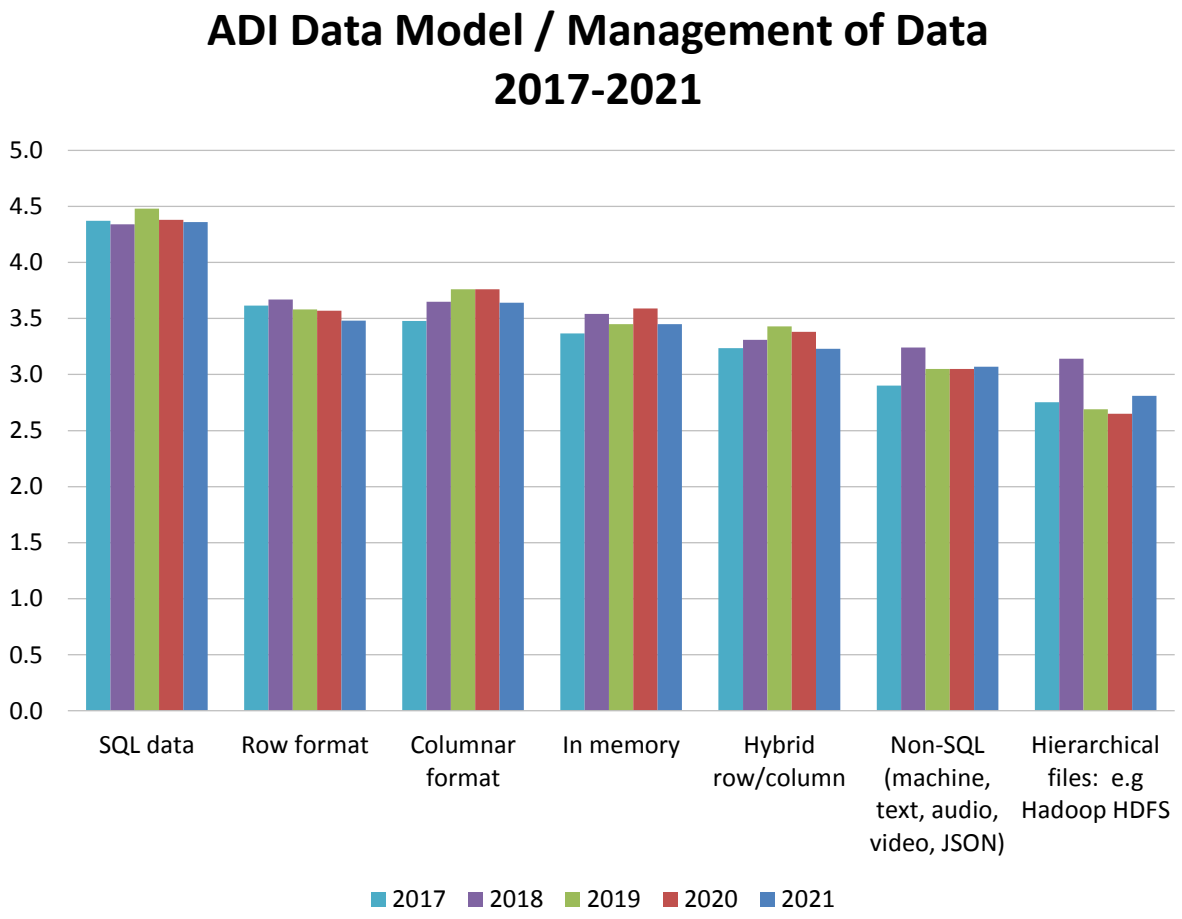


Figure 54 – ADI data model / management of data 2017-2021

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.

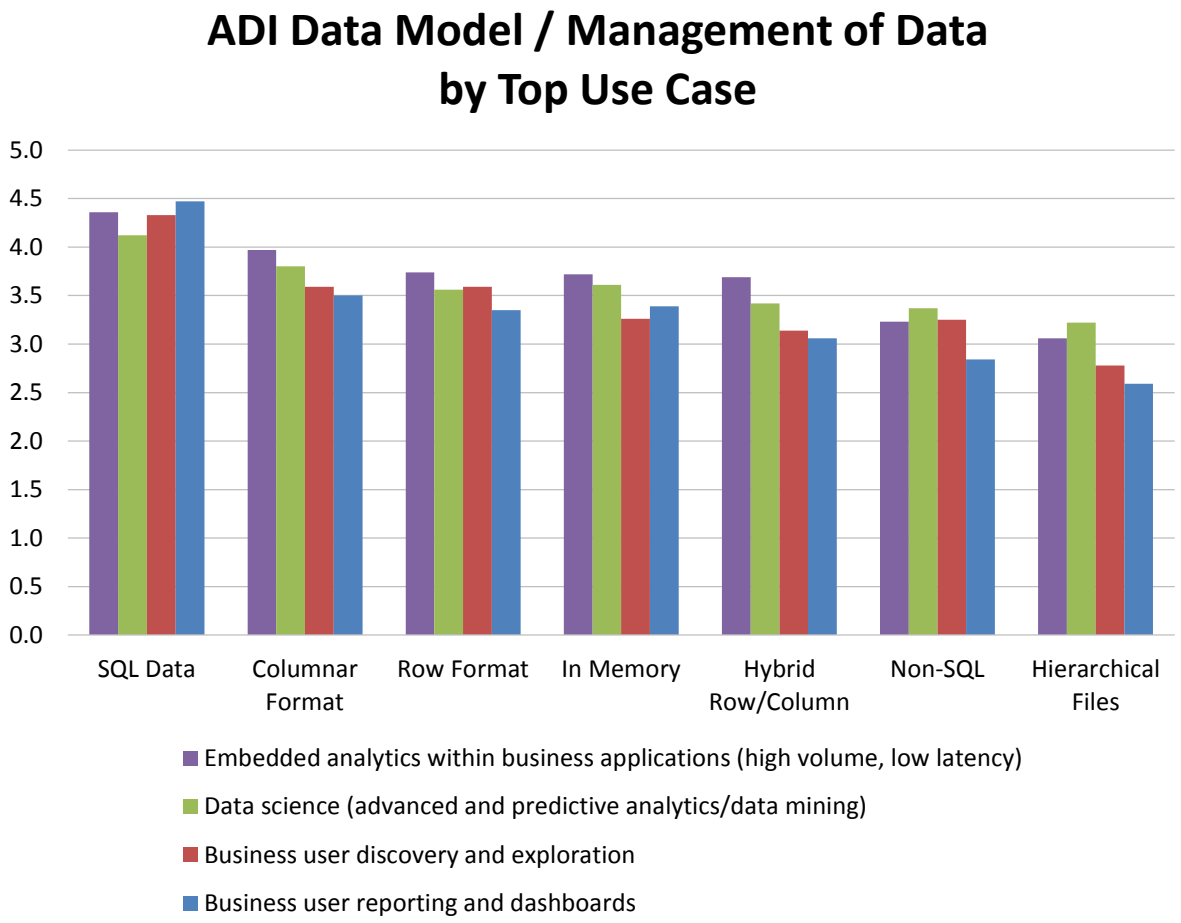


Figure 55 – ADI data model / management of data by top use case

Respondents in all geographies represented 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 EMEA, and the highest is in Asia Pacific.

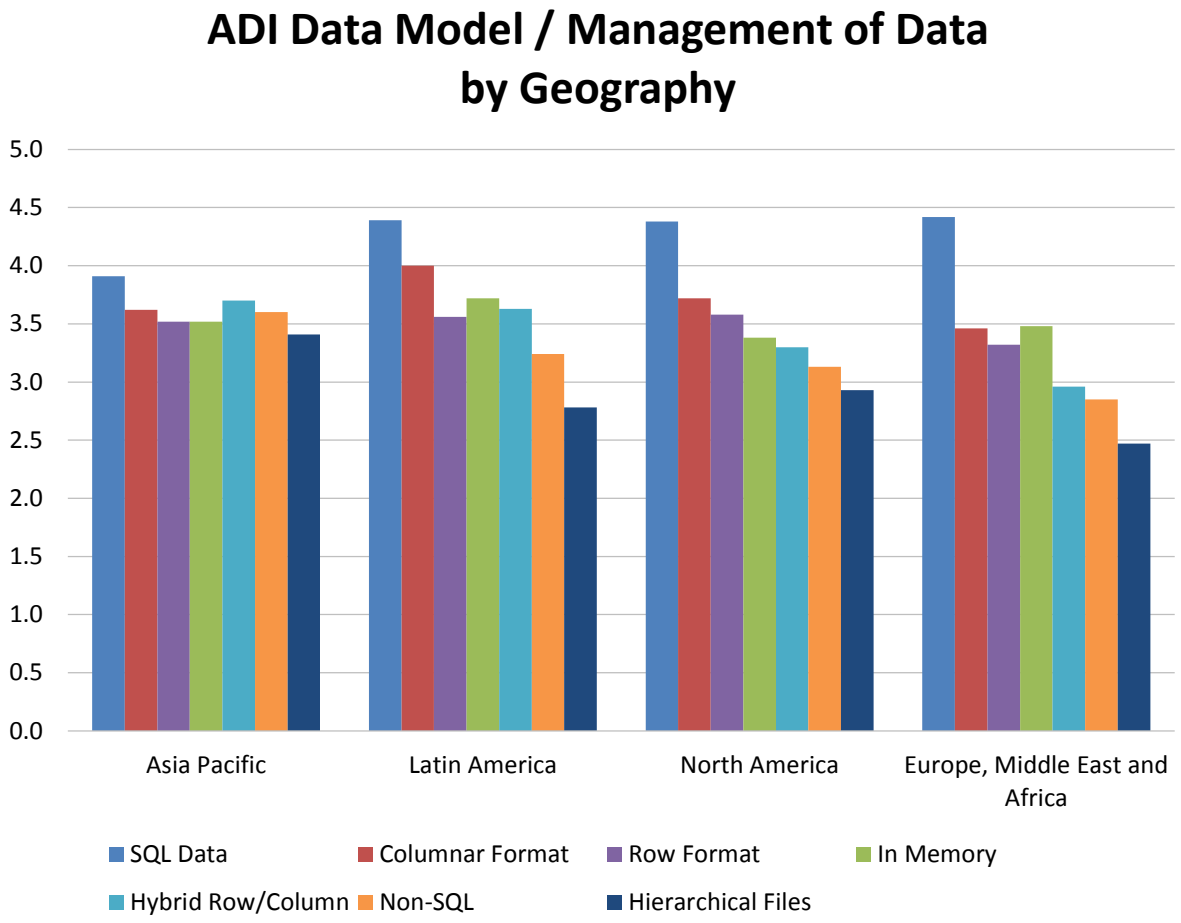


Figure 56 – ADI data model / management of data by geography

ADI data model / management priorities vary by respondents' organizational function (fig. 57). Variation in priorities across organizational 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.

Although all functions place the highest priority upon SQL data, the Strategic Planning function places a high priority upon columnar and row format data while the BICC, Marketing, and Finance place a relatively high priority upon in-memory.

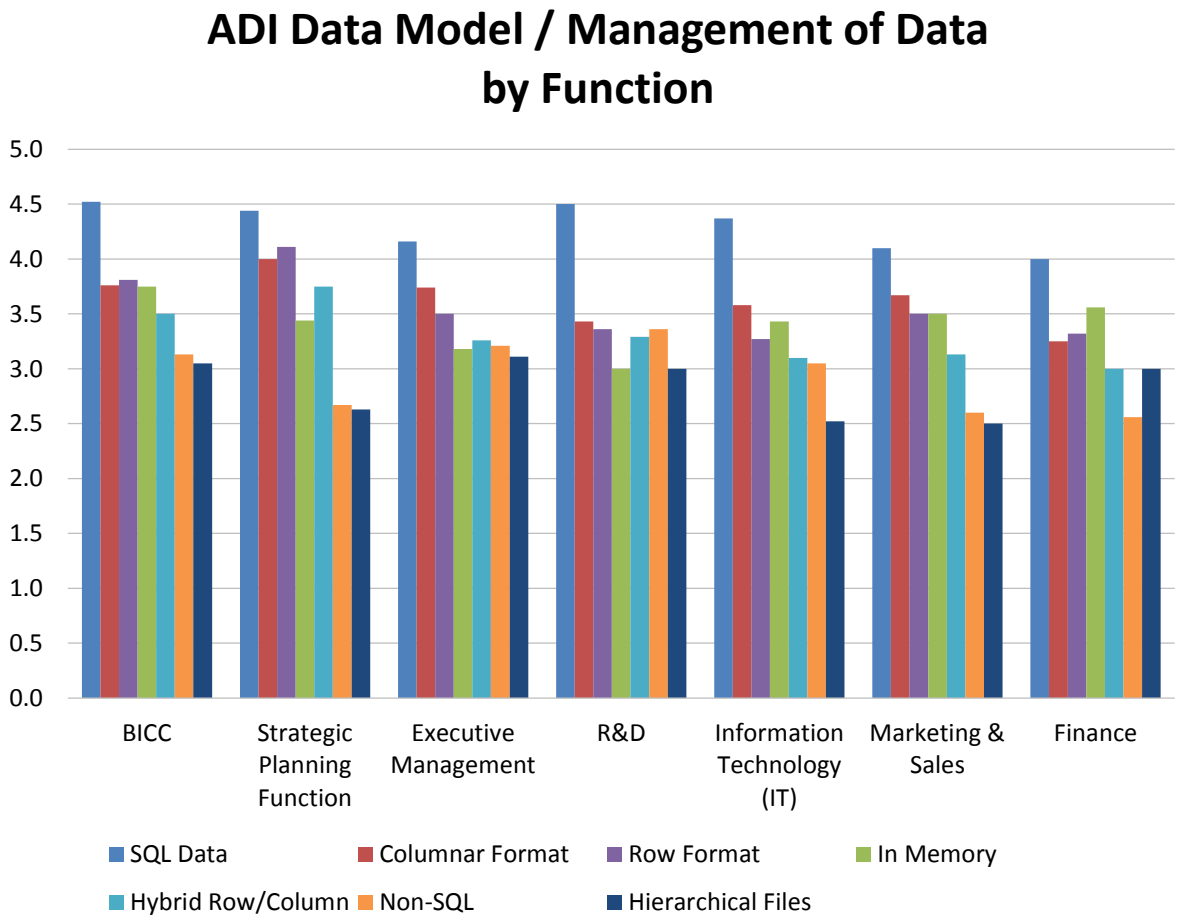


Figure 57 – ADI data model / management of data by function

The relative priority for ADI data model / management capabilities varies quite a bit by industry of the respondent. SQL data models/management and columnar formats are the most common priorities across industries. One of the larger variabilities of priorities is for hierarchical files, e.g., Hadoop and HDFS (fig. 58).

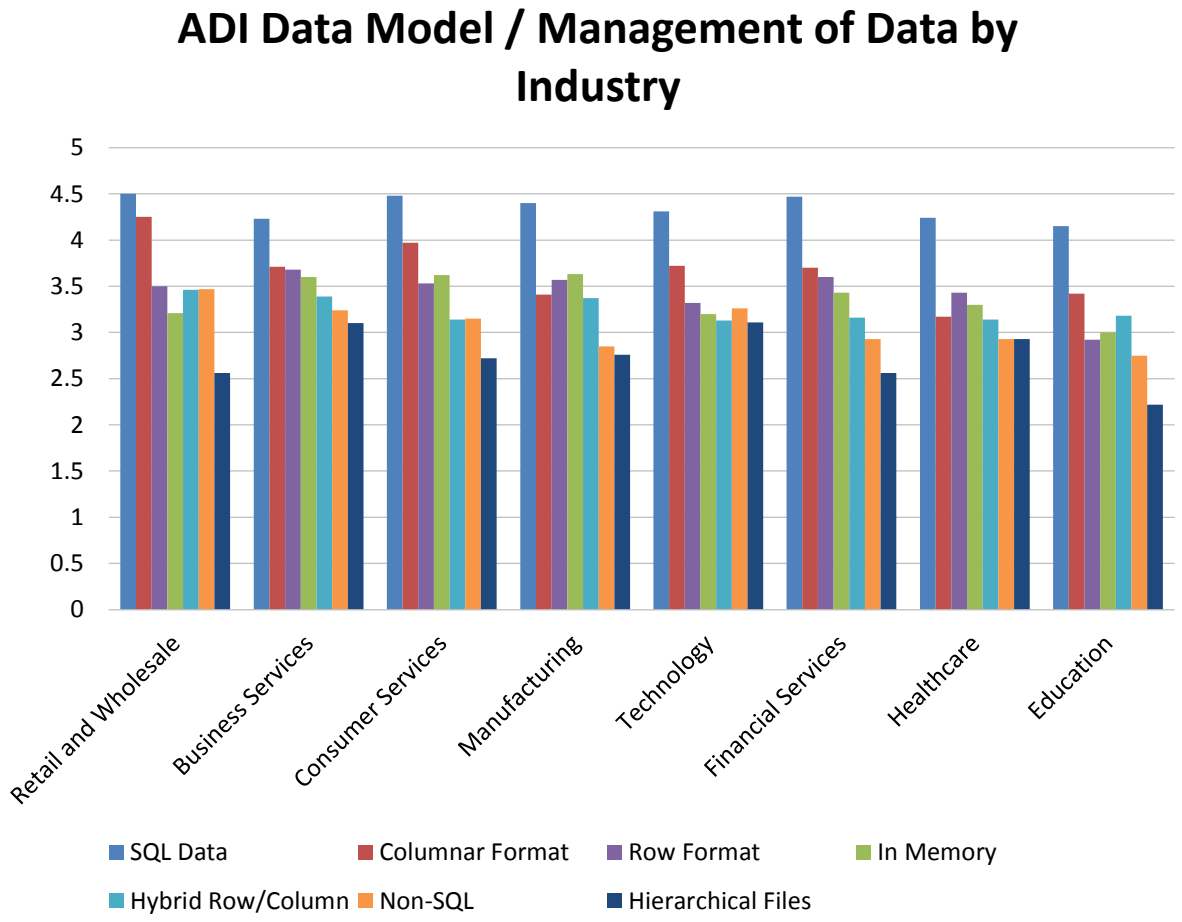


Figure 58 – ADI data model / management of data by industry

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 like the results of other dimensions (i.e., industry, geography, and function) (fig. 59).

Large organizations place a slightly 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.

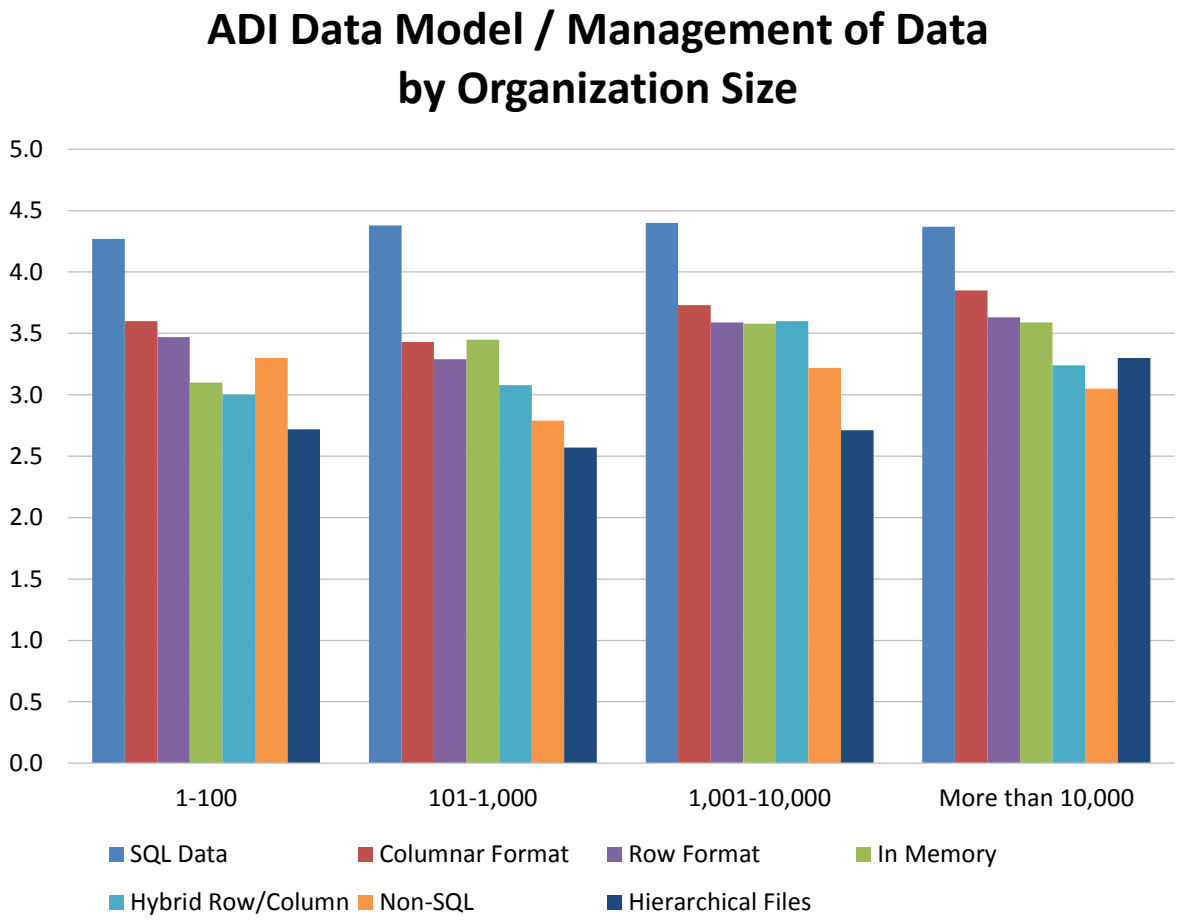


Figure 59 – ADI data model / management of data by organization size

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.

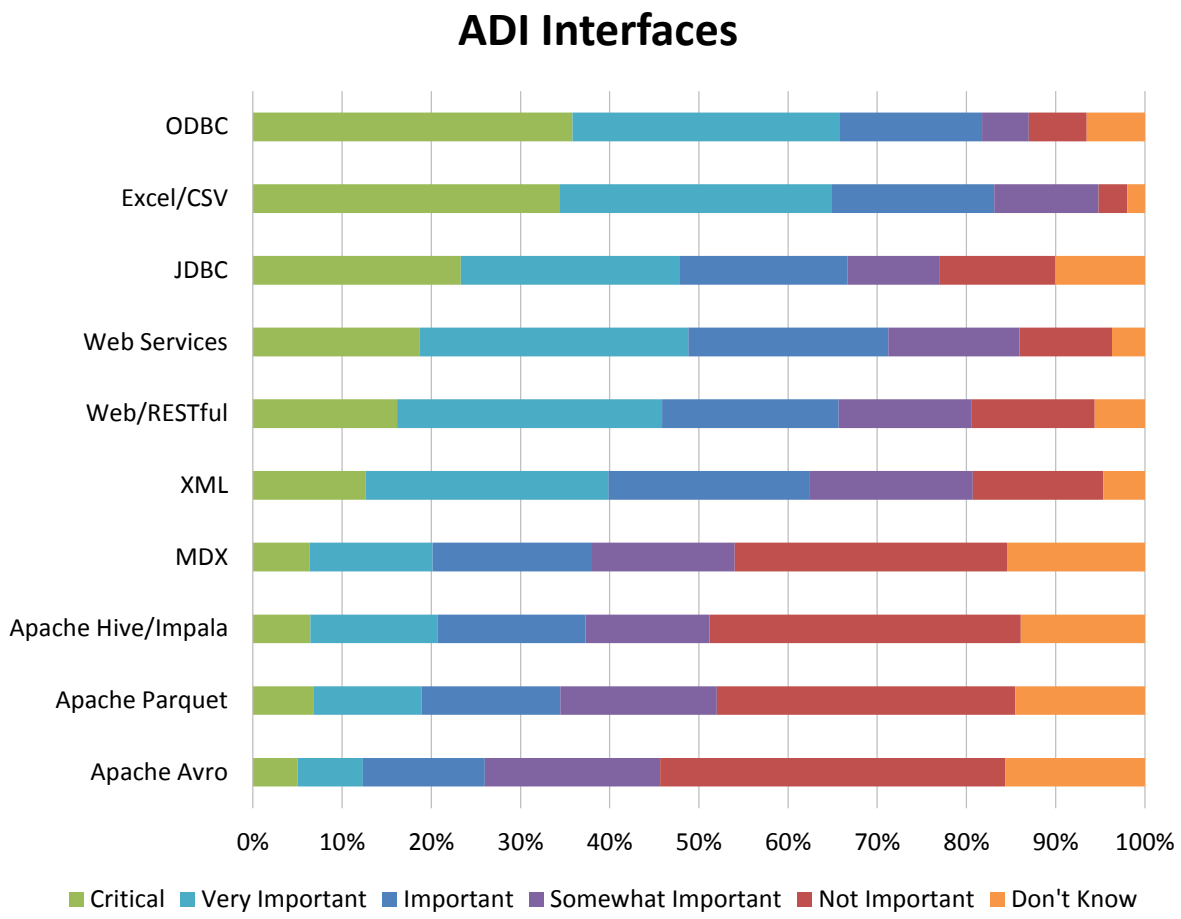


Figure 60 – ADI interfaces

Not much changes in terms of ADI platform interface priorities over the past three years. While the Apache interfaces (i.e., Hive/Impala, Parquet) are low in terms of relative priority, we see year-over-year growth in importance (fig. 61). XML interfaces increase in importance for the past two years, as does MDX.

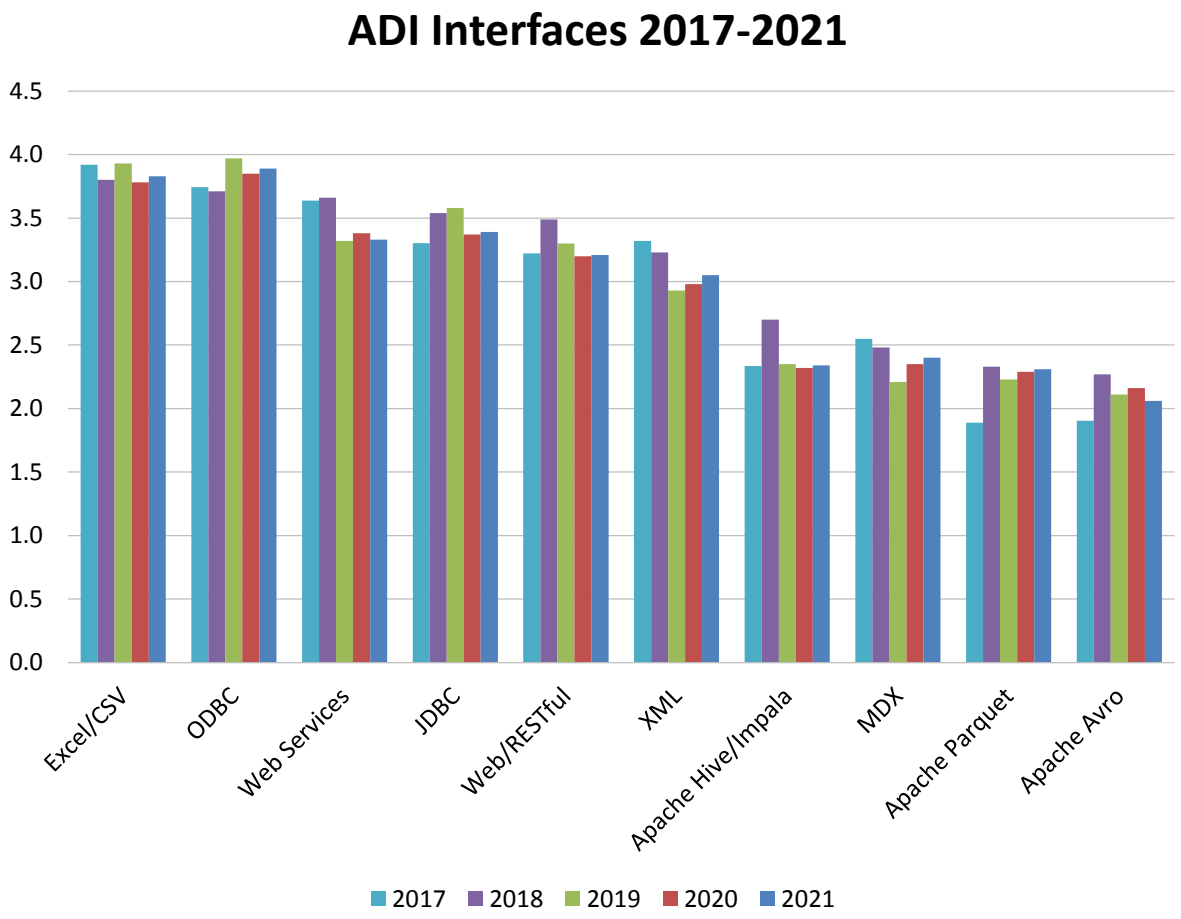


Figure 61 – ADI interfaces 2017-2021

ADI interface relative priorities do not vary a lot by use case, but there are some interesting variations of note (fig. 62). Both embedded and "business user discovery and exploration" use-cases prioritize ODBC interfaces. Respondents with data science use cases rank Excel/CSV as their top interface priority.

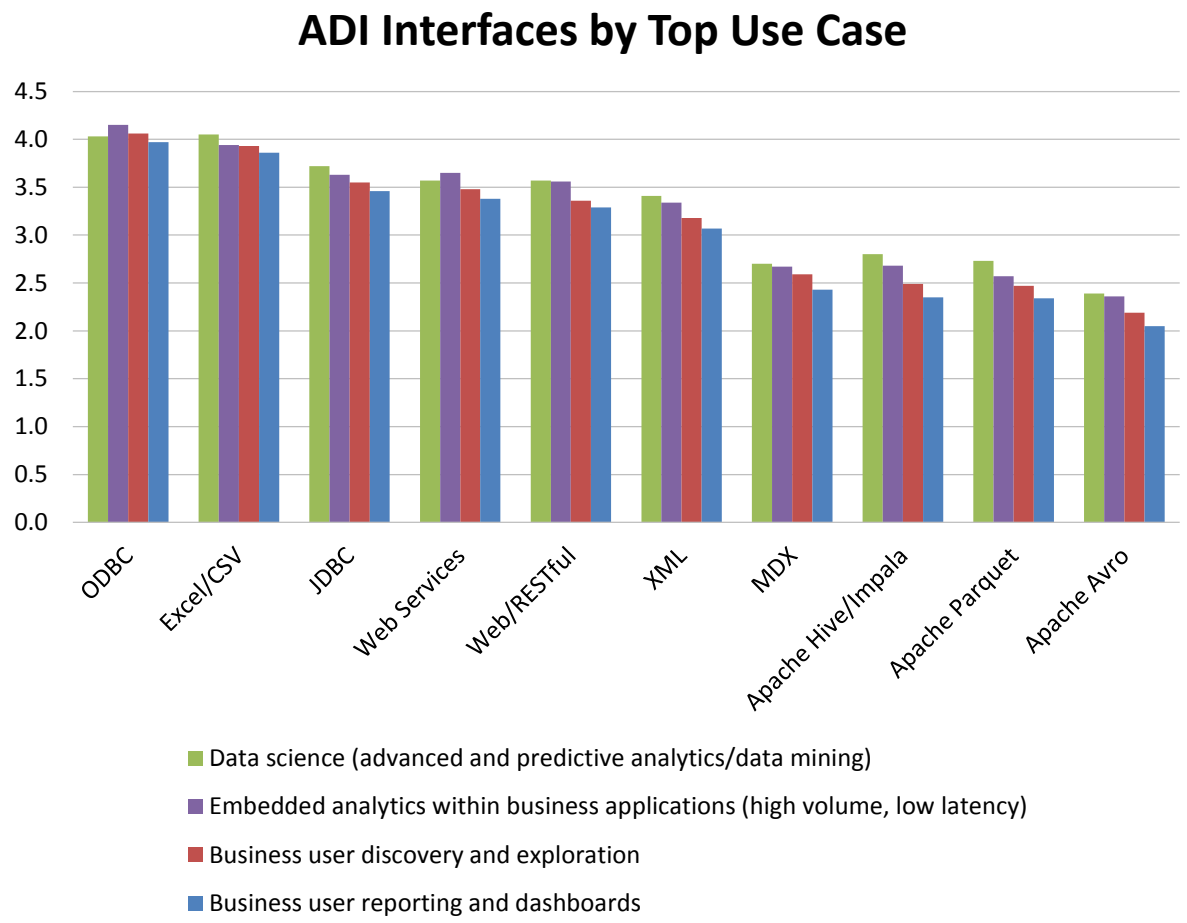


Figure 62 – ADI interfaces by top use case

As fig. 63 shows, there is not a lot of variation in respondents' relative priorities for ADI interfaces by geography. Asia-Pacific respondents place a comparatively higher priority for Apache-based interface services, and Latin America places a disproportionate priority upon Excel/CSV.

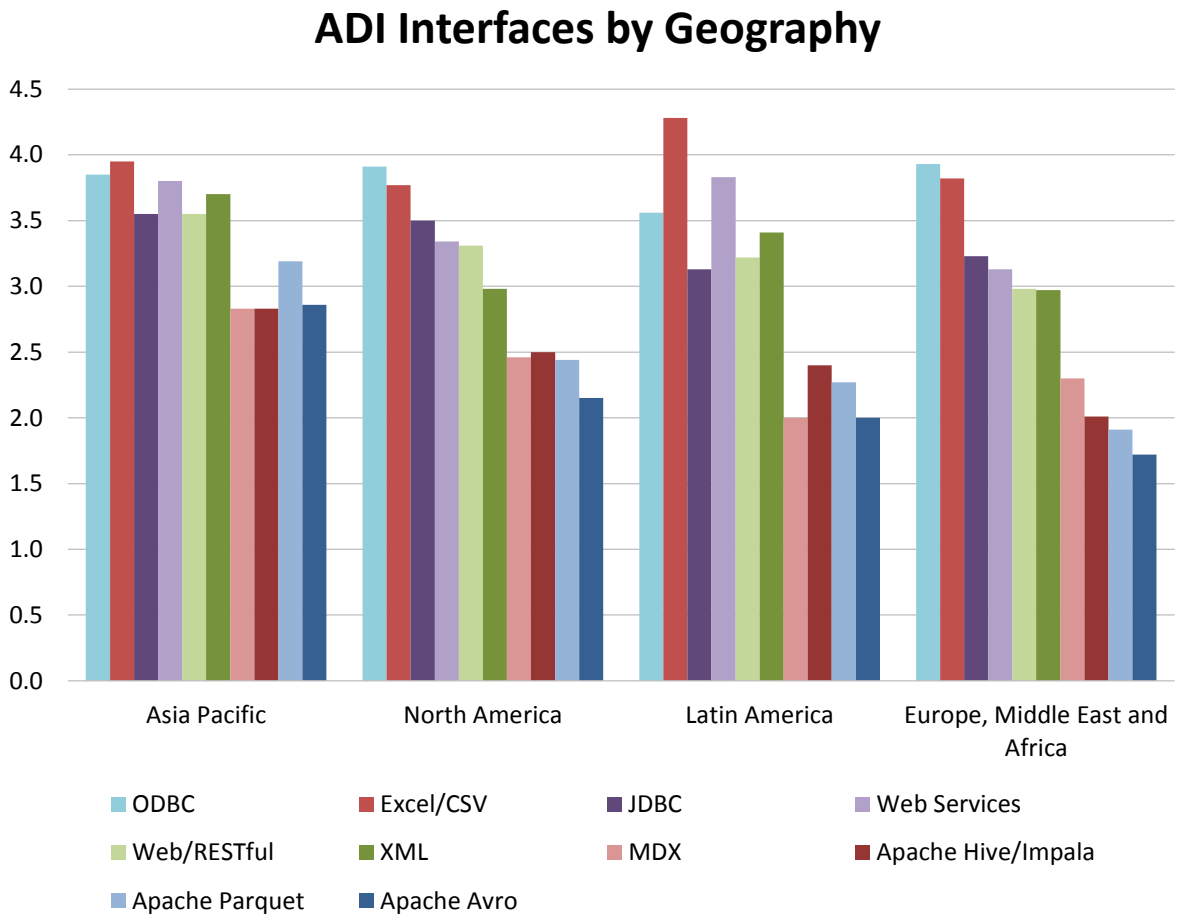


Figure 63 – ADI interfaces by geography

There can be quite a bit of variability in priorities for ADI interfaces by organizational function. This may cause alignment problems when organizations try to rank priorities for a vendor's ADI proposal that spans multiple functions. Many respondents place their highest priority on Excel/CSV interfaces (fig. 64), even over ODBC-based interfaces, and their lowest priority on Apache services. This is not surprising, given the amount of work done in spreadsheets.

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

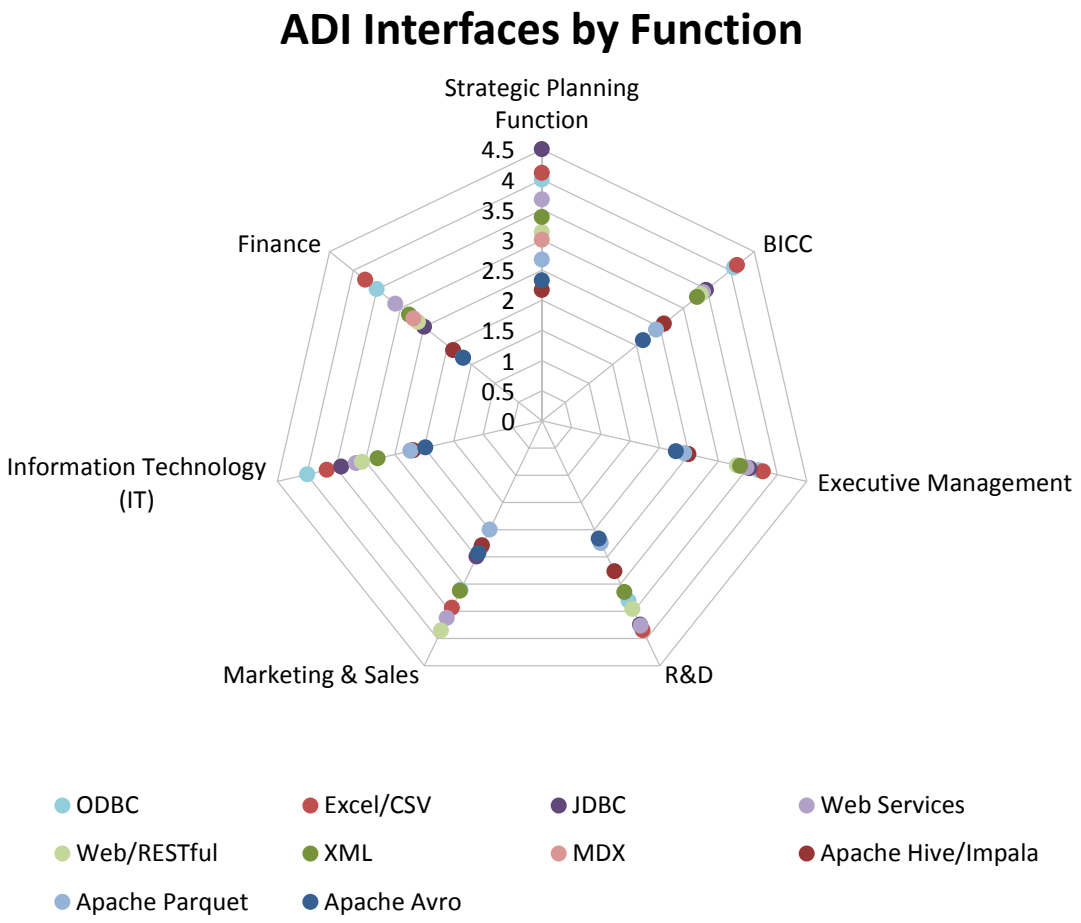


Figure 64 – ADI interfaces by function

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. This is followed by JDBC, with strongest support from Technology and Retail/ Wholesale. MDX, once a popular multi-dimensional interface, still sees some limited support from respondents within Business Services and Healthcare industry segments. Although a relatively low priority, Hive/Impala and Parquet interfaces garner limited support with respondents from the Technology and Business Services segments.

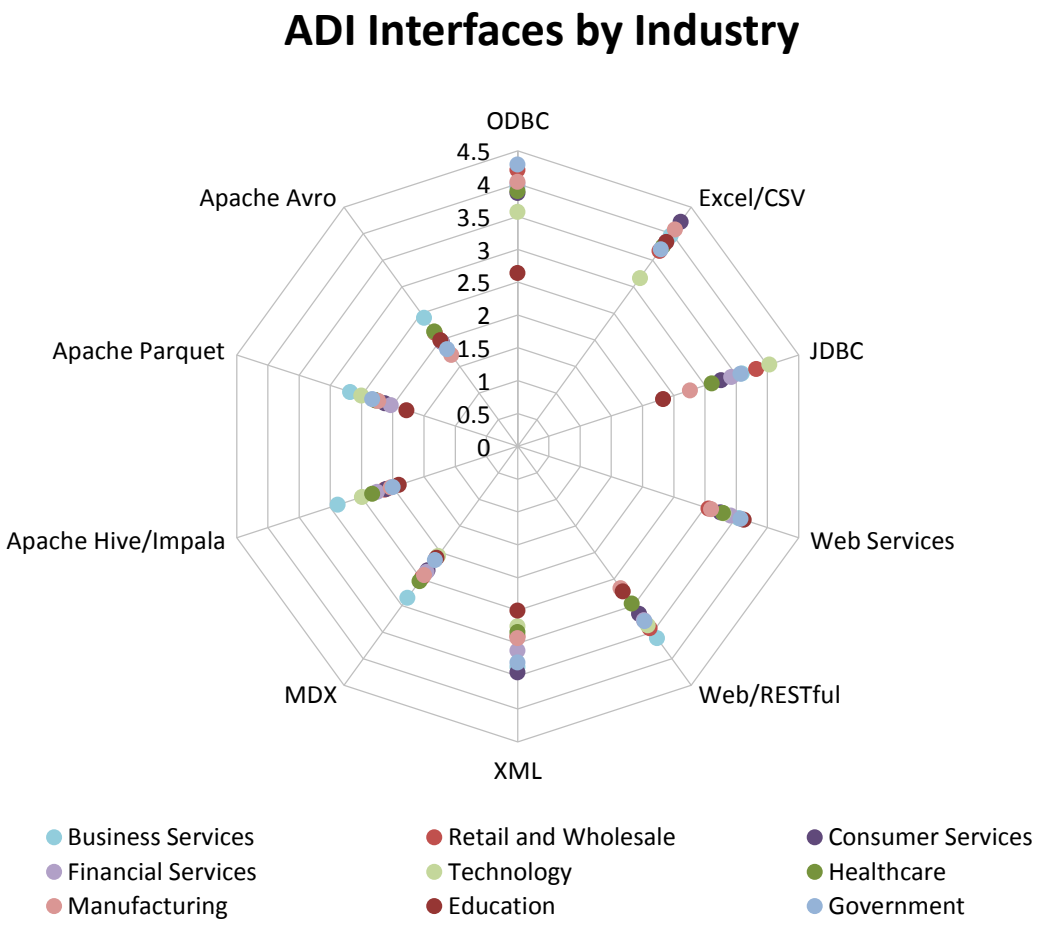


Figure 65 – ADI interfaces by industry

Examining interface preferences by organization size, ODBC, Excel/CSV, Web Services, and JDBC interfaces are the primary interfaces for accessing ADI platform data. Both larger and smaller-sized organizations place higher relative priority on Web Services and Web/Restful interfaces (fig 66).

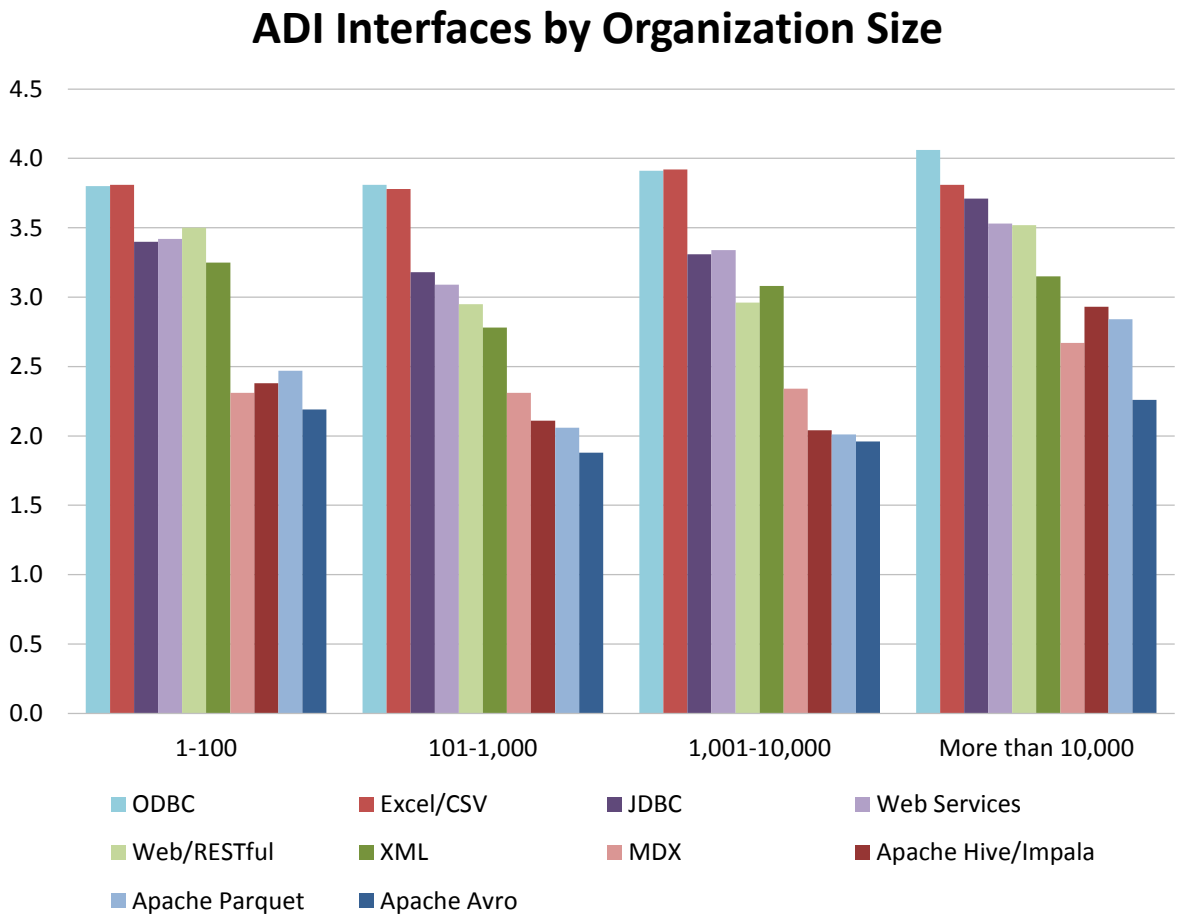


Figure 66 – ADI interfaces by organization size

ADI Analytical Features

The diversity and priority of analytical features within an ADI platform can vary widely. Aggregations, statistical analysis/R, and multi-dimensional/OLAP features rank highest year over year, 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.

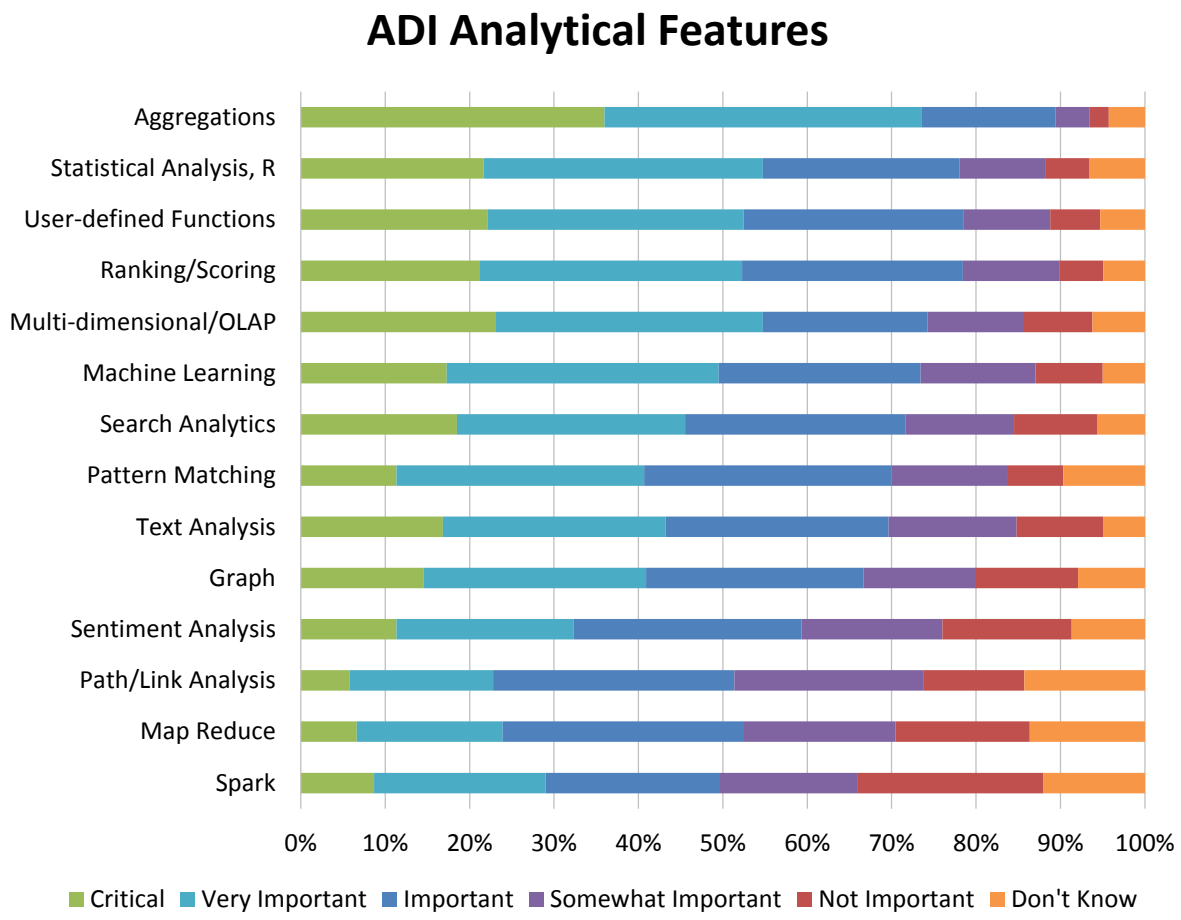


Figure 67 – ADI analytical features

There are some modest year-over-year changes to ADI analytical feature requirements. Statistical analysis/R features shows a slight decline in priority year over year but remains the number-two priority. More advanced analytic features, such as machine learning, show an increasing trend in priority. Other features including multi-dimensional OLAP, Graph, and Spark trend downwards (fig. 68).

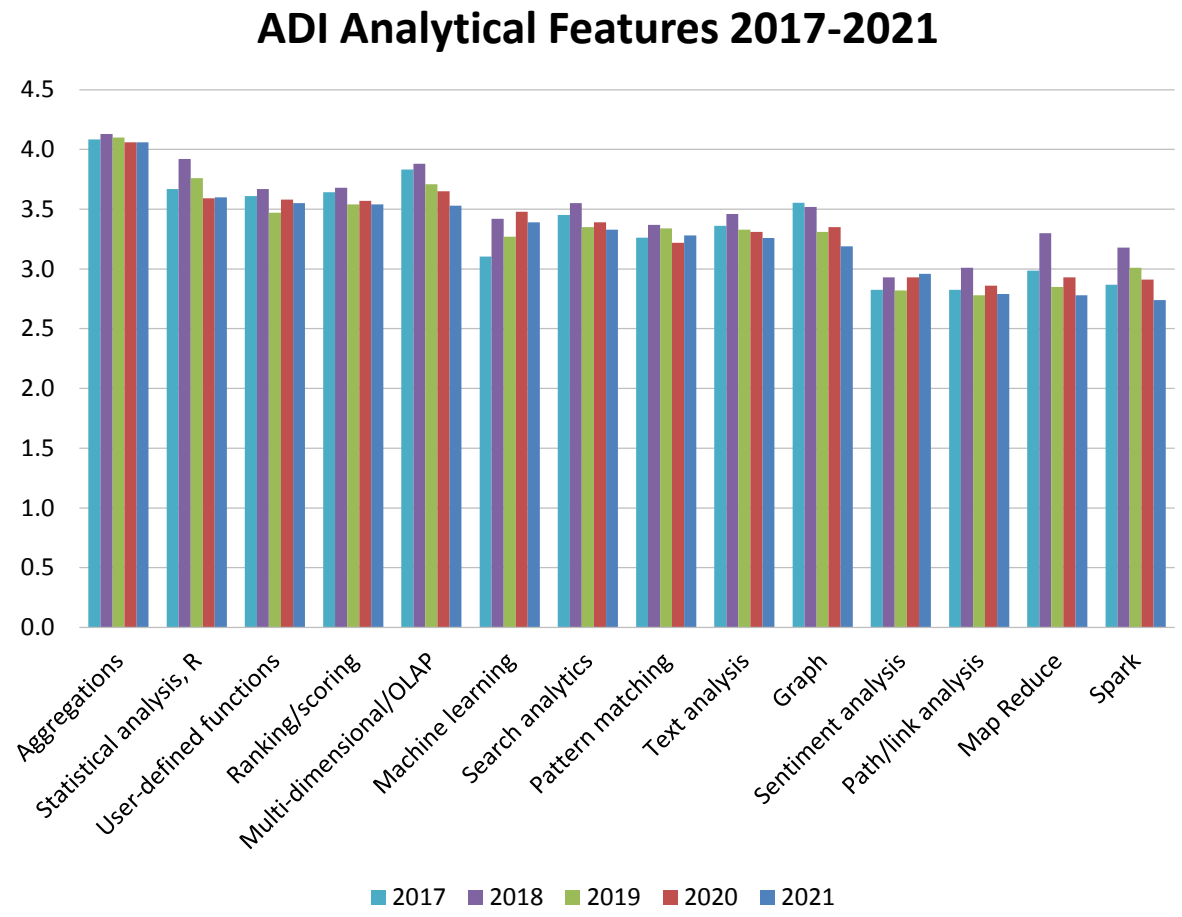


Figure 68 – ADI analytical features 2017-2021

In general, the data science use cases place higher priorities on “analytic features” in the ADI than other use cases. 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 “machine learning.” Graph analysis is the lowest priority for embedded use-case priorities.

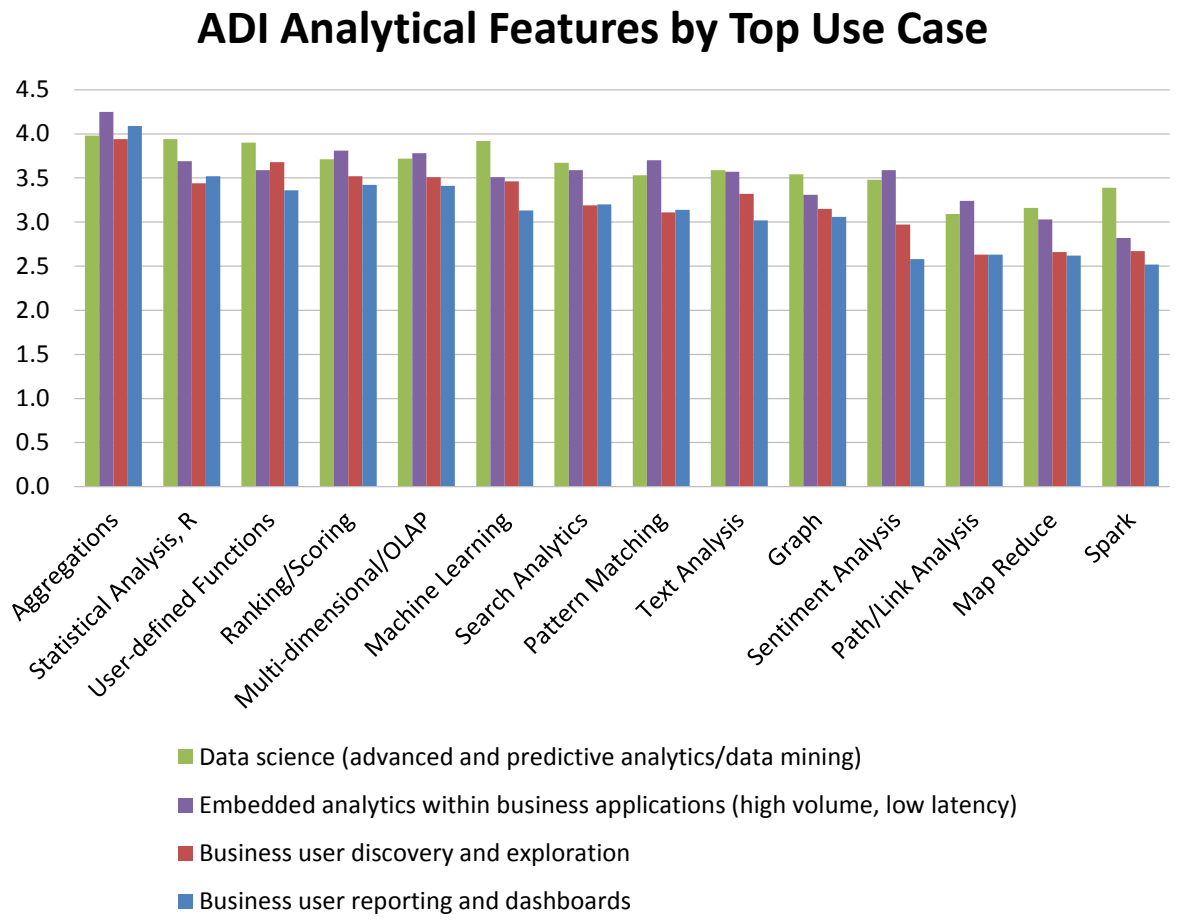


Figure 69 – ADI analytical features by top use case

Although aggregations and R are top choices across all geographies, Asia-Pacific respondents prioritize the need for more diverse analytical features than other regions (fig. 70). Except for a few minor variations, North American and EMEA respondents have similar relative priority profiles.

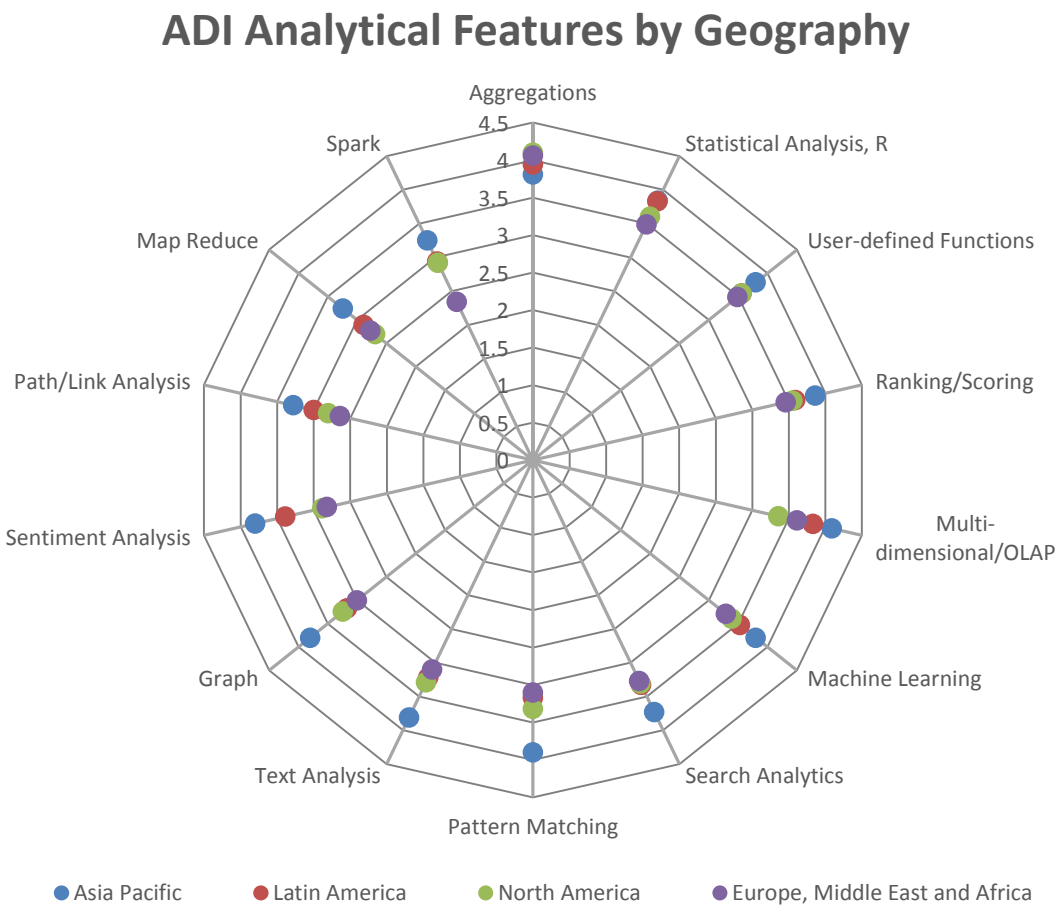


Figure 70 – ADI analytical features by geography

The priority of ADI analytic features varies quite a bit by organization function (fig. 71) including capabilities for more advanced analysis such as sentiment, path/link, and Spark analytic features. This may create challenges when trying to gain agreement and align IT and non-IT priorities for ADI analytic functions.

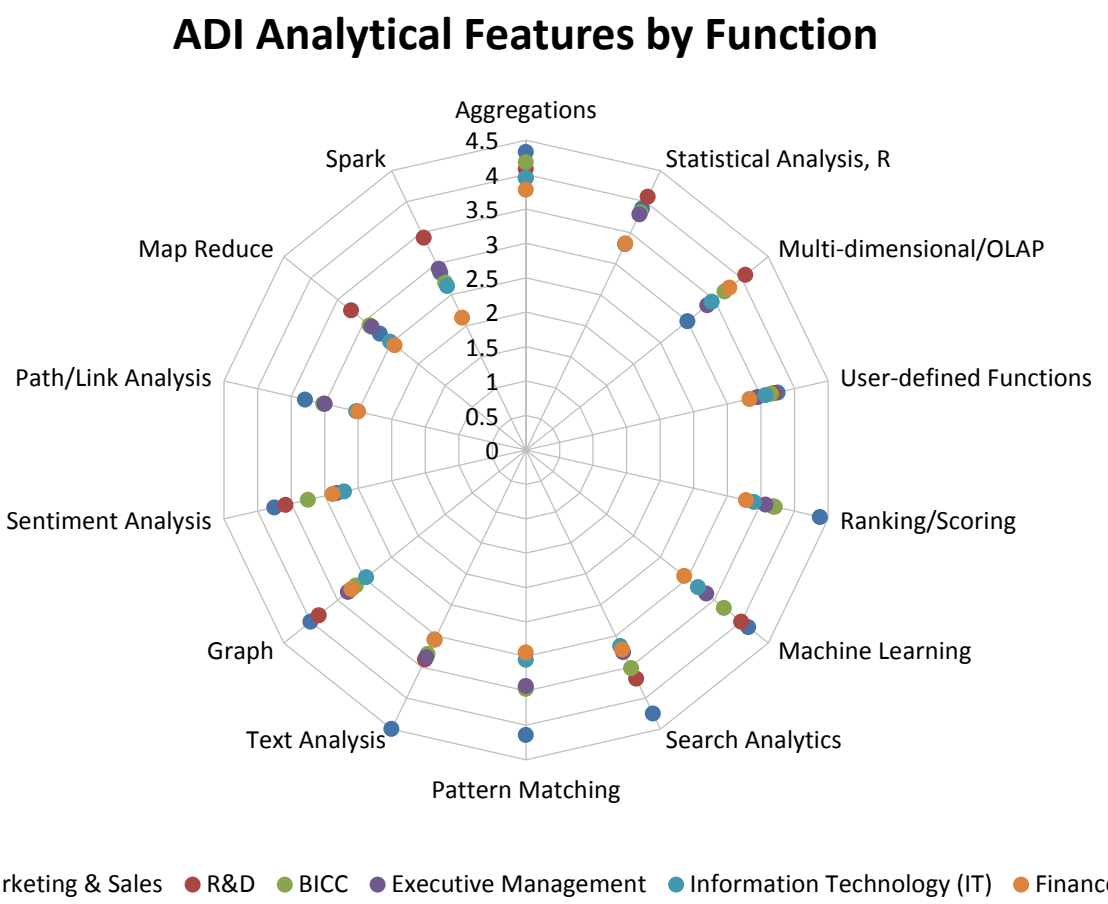


Figure 71 – ADI analytical features by function

When we analyze ADI analytical features by industry, we see a fair amount of variation in priorities (fig. 72). By way of example, Higher Education places a high priority upon multi-dimensionality/OLAP and machine learning while Financial Services favors statistical analysis, R, and ranking/scoring. In contrast, Healthcare places a relatively high priority upon text analysis and search.

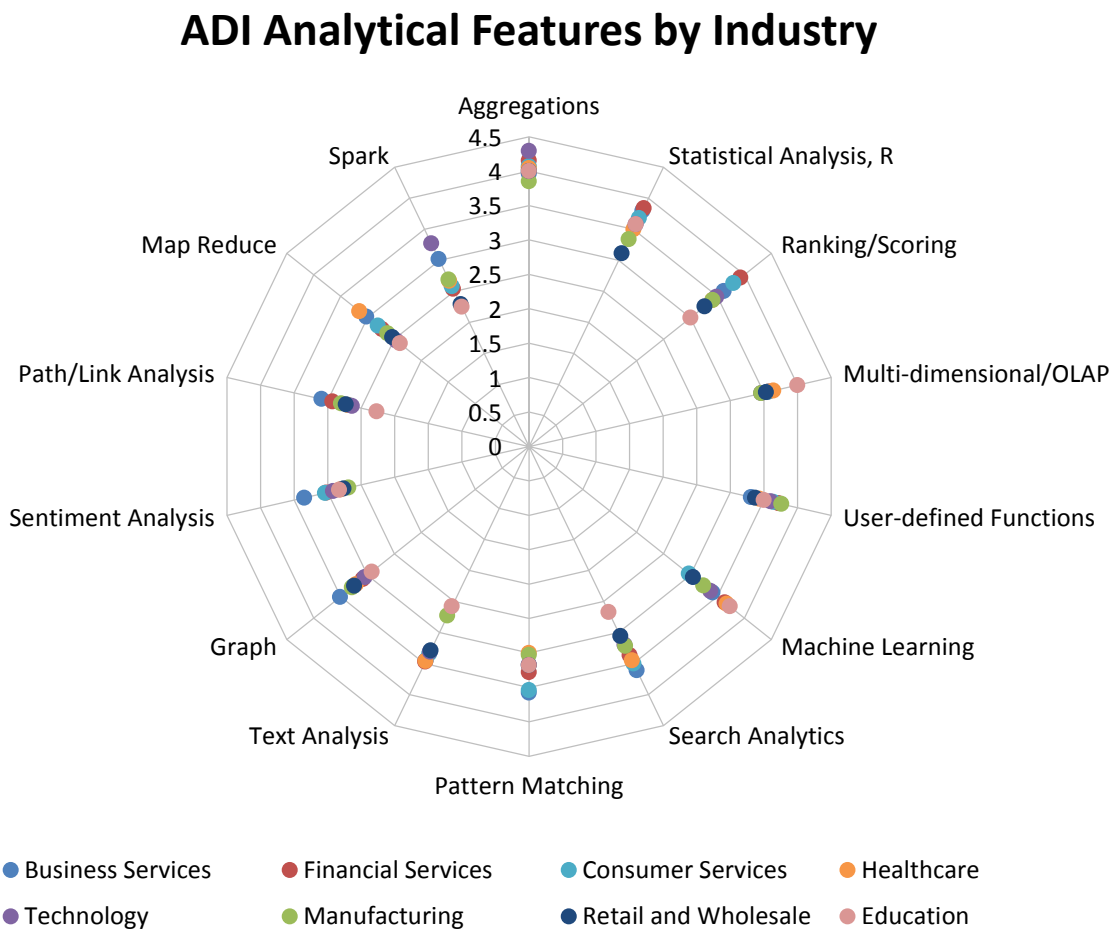


Figure 72 – ADI analytical features by industry

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 slightly lower priority on these advanced analytical features.

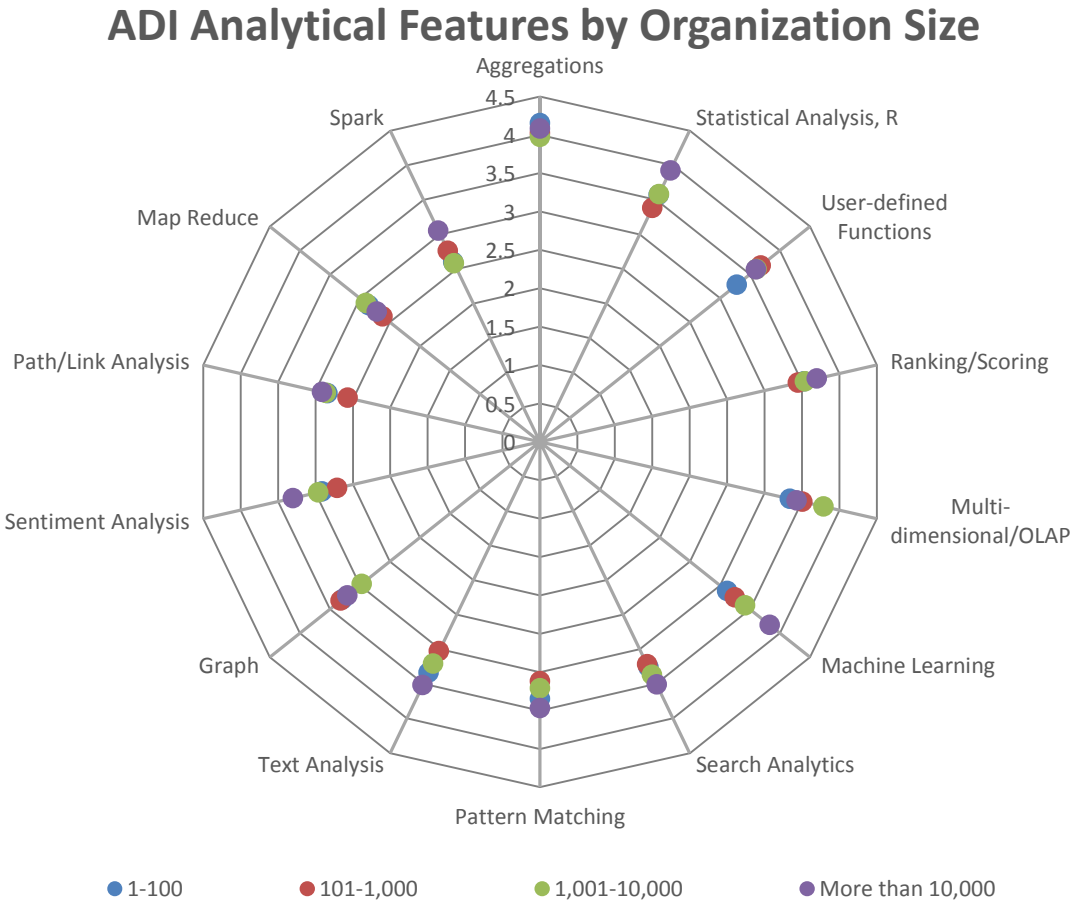


Figure 73 – ADI analytical features by organization size

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 cover 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 touch points—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.

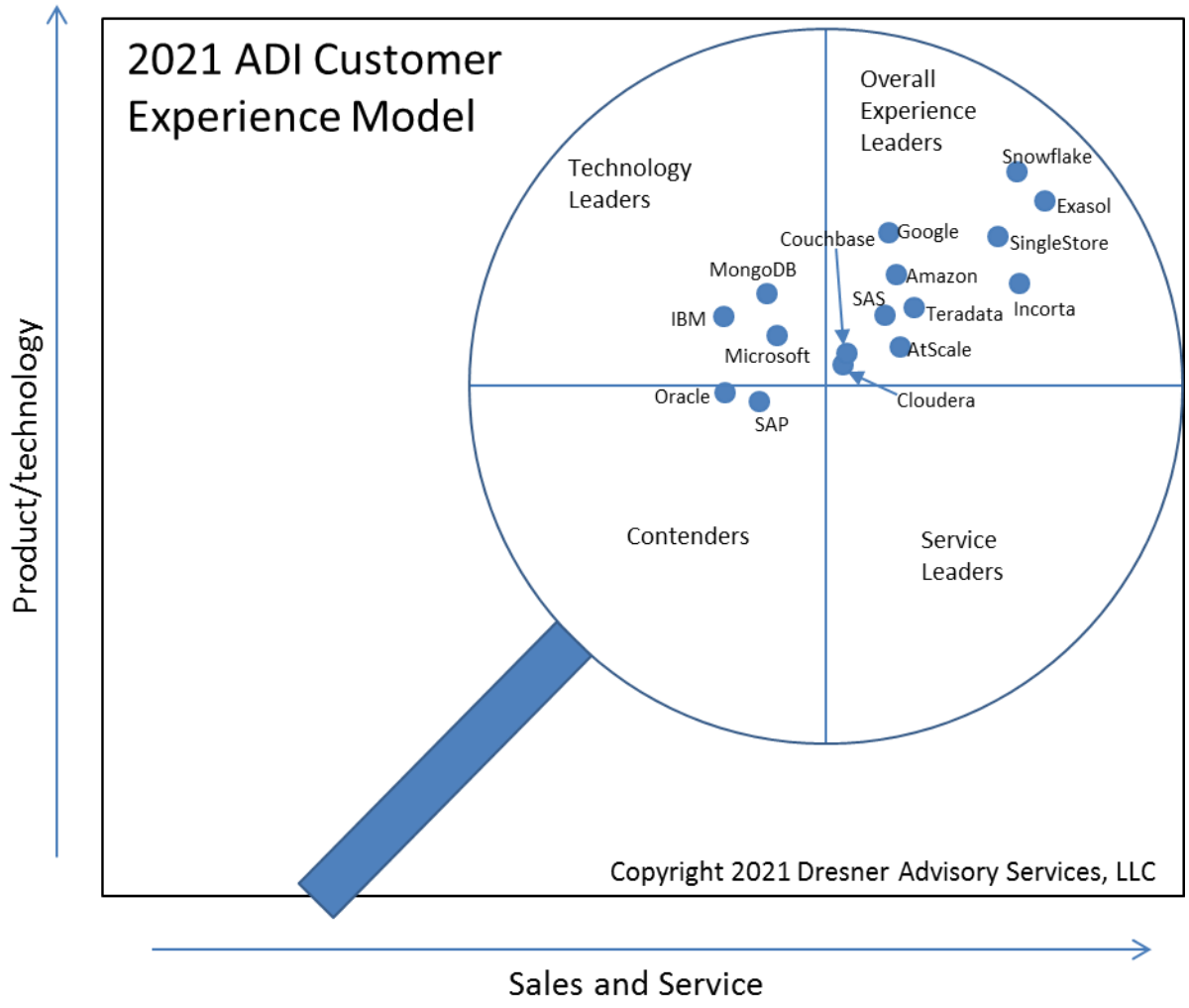


Figure 74 – Customer experience model

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.

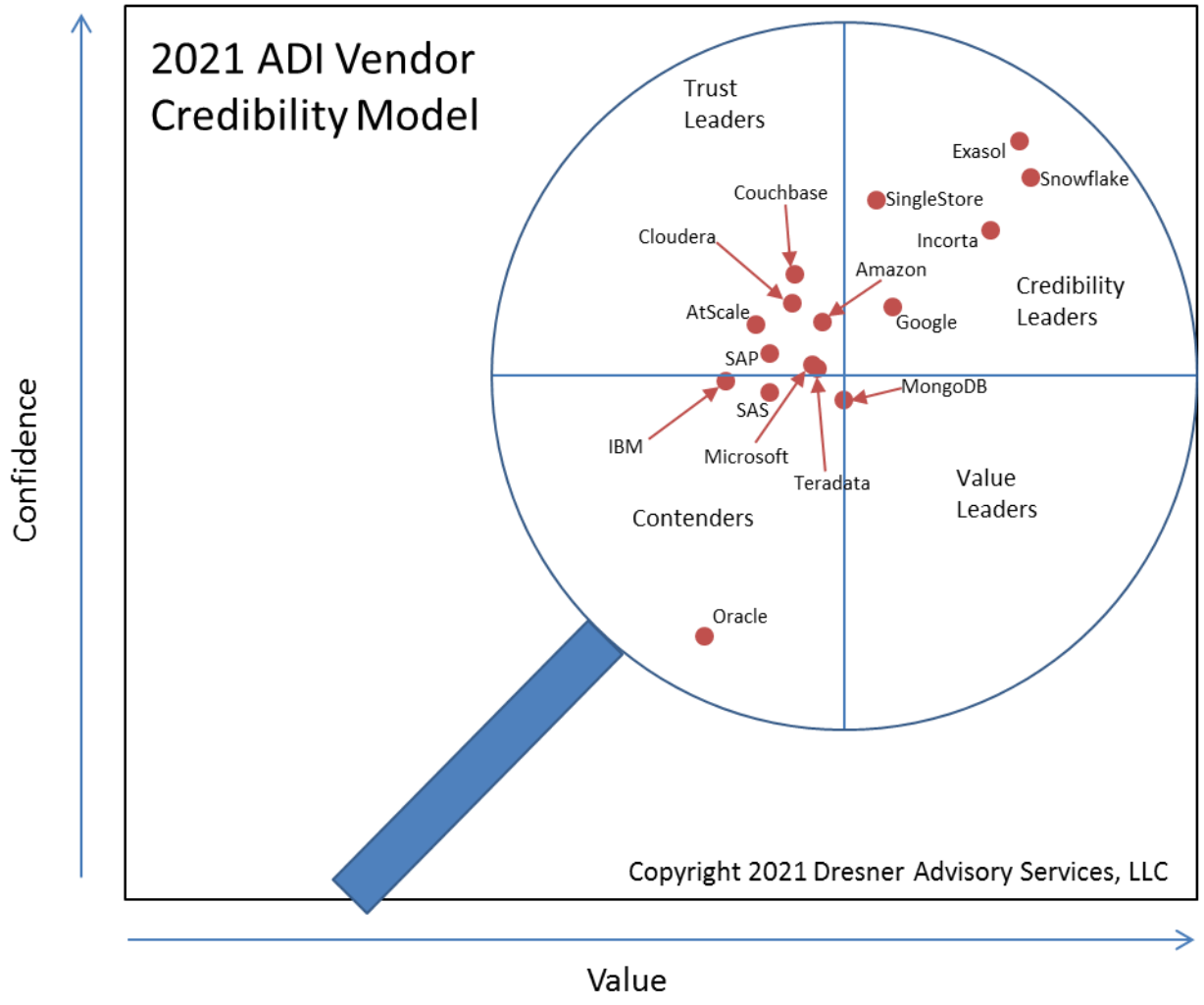


Figure 75 – Vendor credibility model

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)*
  ○ Professionalism
  ○ Product knowledge
  ○ Understanding our business/needs
  ○ Responsiveness
  ○ Flexibility/accommodation
  ○ Business practices
  ○ Contractual terms and conditions
  ○ Follow-up after the sale
 	
- **Value for price** *(3 o'clock)*
 	
- **Quality and usefulness of product** *(3 - 7 o'clock)*
  ○ 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
	- **Quality and usefulness of product (continued)**
	  ○ Customization and extensibility
	  ○ Ease of upgrade/migration to new versions
	  ○ Online forums and documentation
	-
	- **Quality of technical support** *(8 - 9 o'clock)*
	  ○ Professionalism
	  ○ Product knowledge
	  ○ Responsiveness
	  ○ Continuity of personnel
	  ○ Time to resolve problems
	
	- **Quality and value of consulting services** *(9 - 10 o'clock)*
	  ○ Professionalism
	  ○ Product knowledge
	  ○ Experience
	  ○ Continuity
	  ○ Value
	
	- **Integrity** *(11 o'clock)*
	
	- **Whether vendor is recommended** *(12 o'clock)*

Amazon Detailed Score

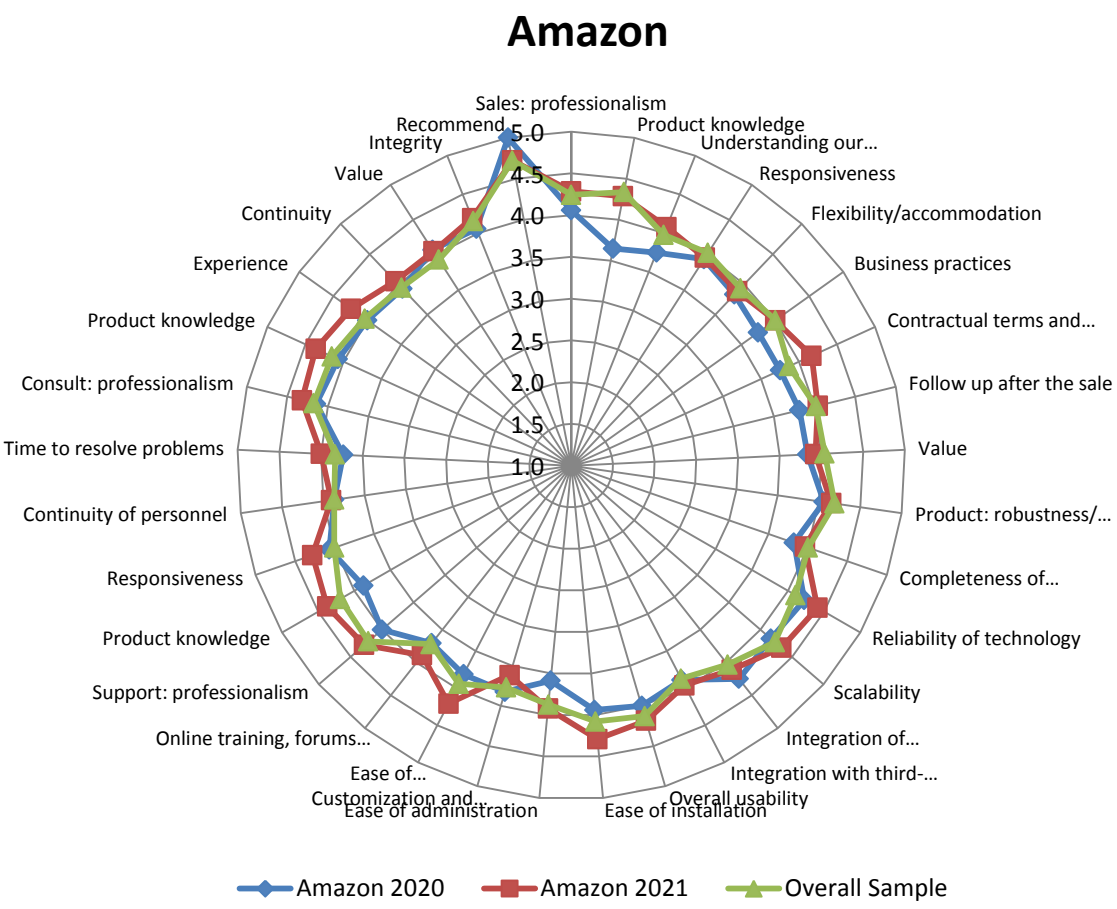


Figure 76 – Amazon detailed score

In 2021, Amazon's performance improves across most categories of measurement, including sales, value, product, technical support, and integrity. Its scores are generally above the overall sample, and it is now considered an Overall Leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model.

AtScale Detailed Score

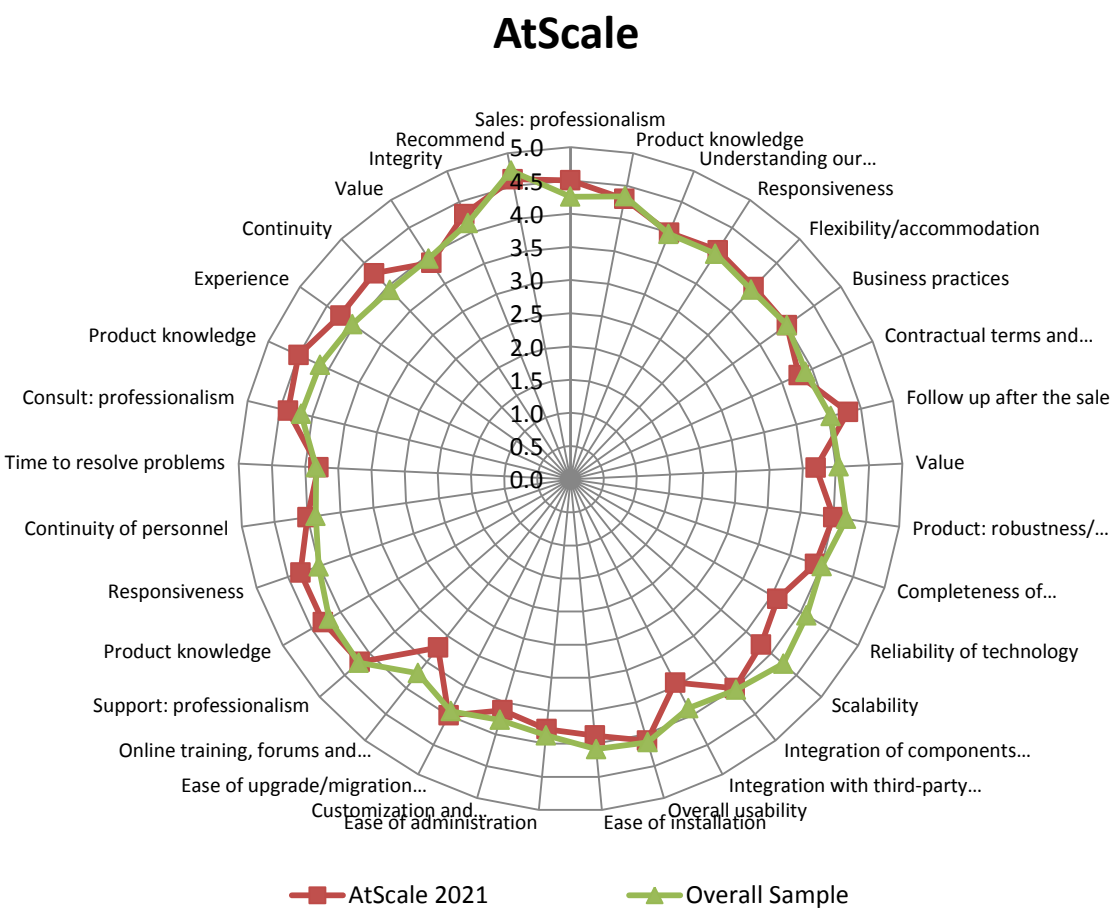


Figure 77 – AtScale detailed score

In its first year of coverage, AtScale's scores are generally in line with the overall sample, with the exception of consulting, which is generally above the sample. It is considered an Overall Leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model.

Cloudera Detailed Score

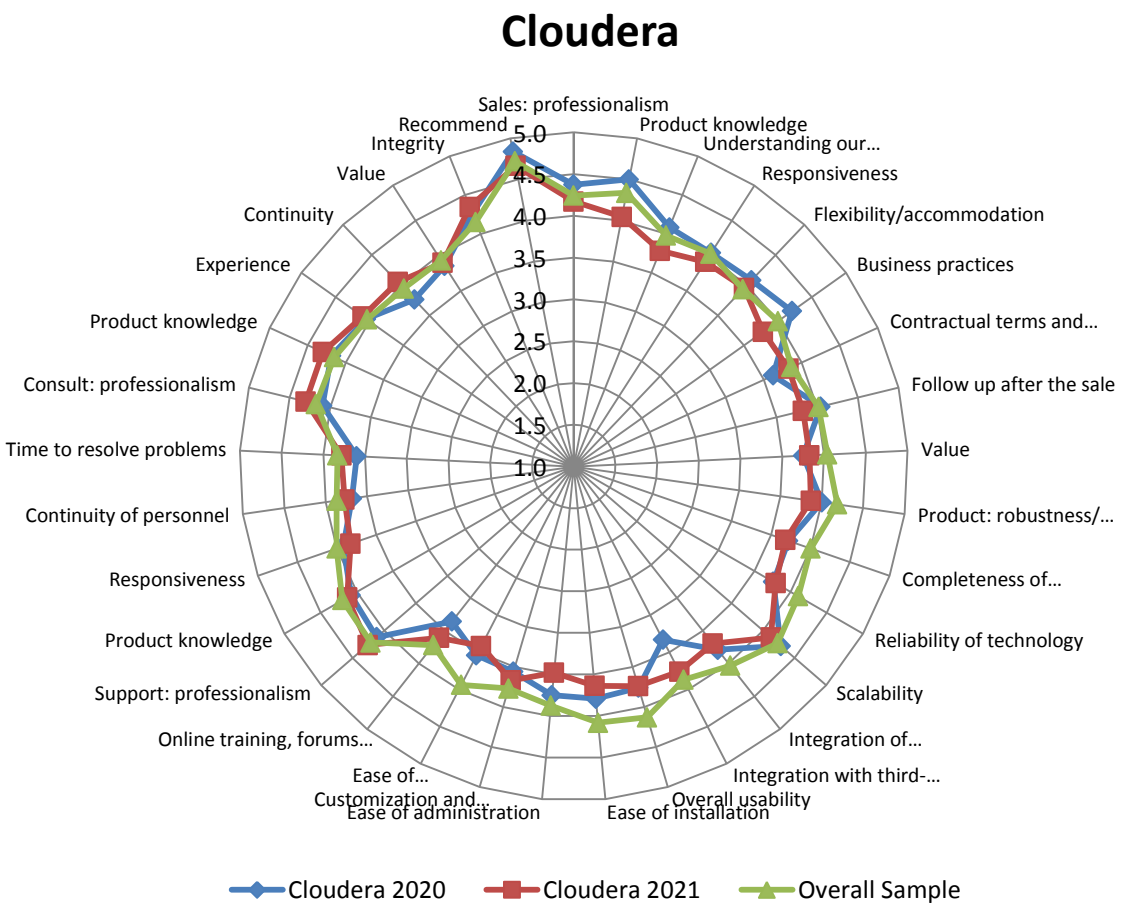


Figure 78 – Cloudera detailed score

With scores generally in line with the overall sample, Cloudera is an overall leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model. For 2021, it has improvements in several areas including value, product reliability, third-party integration, overall integrity, and most consulting and technical support measures.

Couchbase Detailed Score

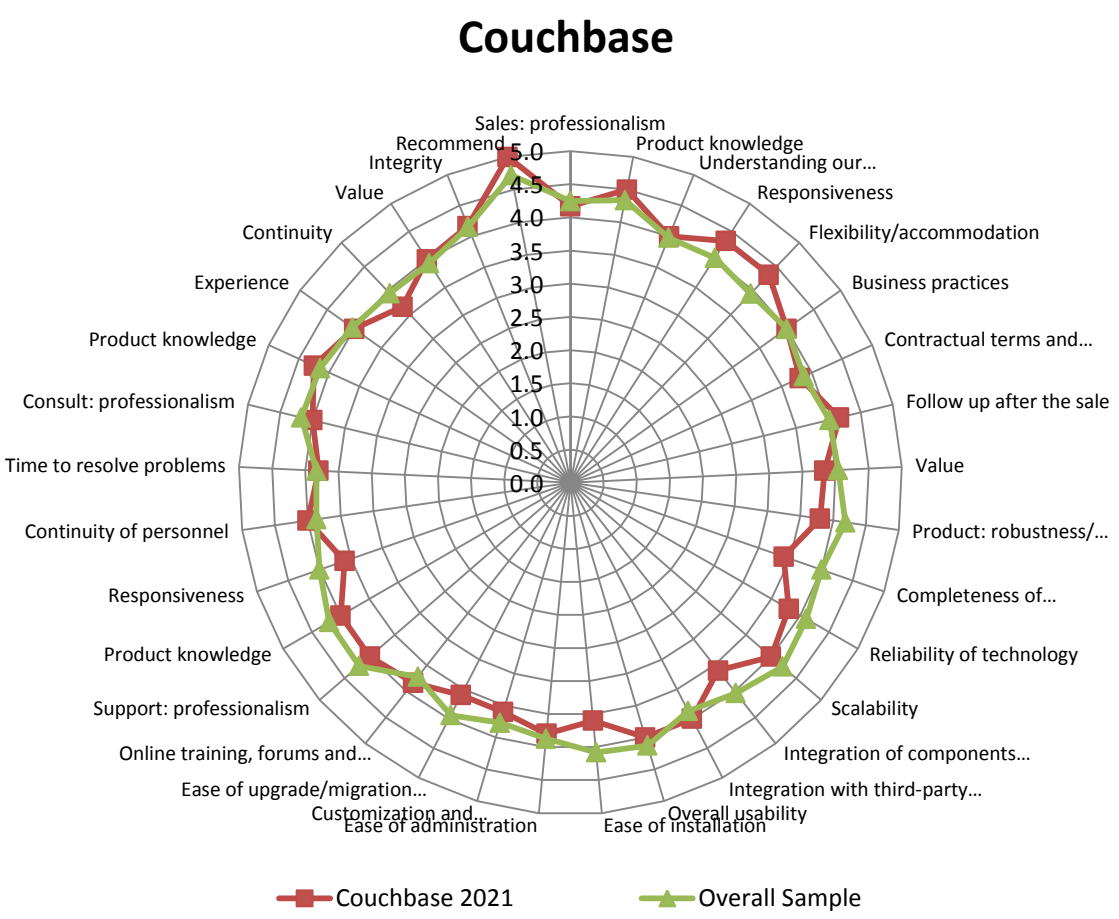


Figure 79 - Couchbase detailed score

In its first year of inclusion, Couchbase's performance is generally in line with the overall sample. It is considered an overall leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model and has a perfect customer recommend score.

Exasol Detailed Score

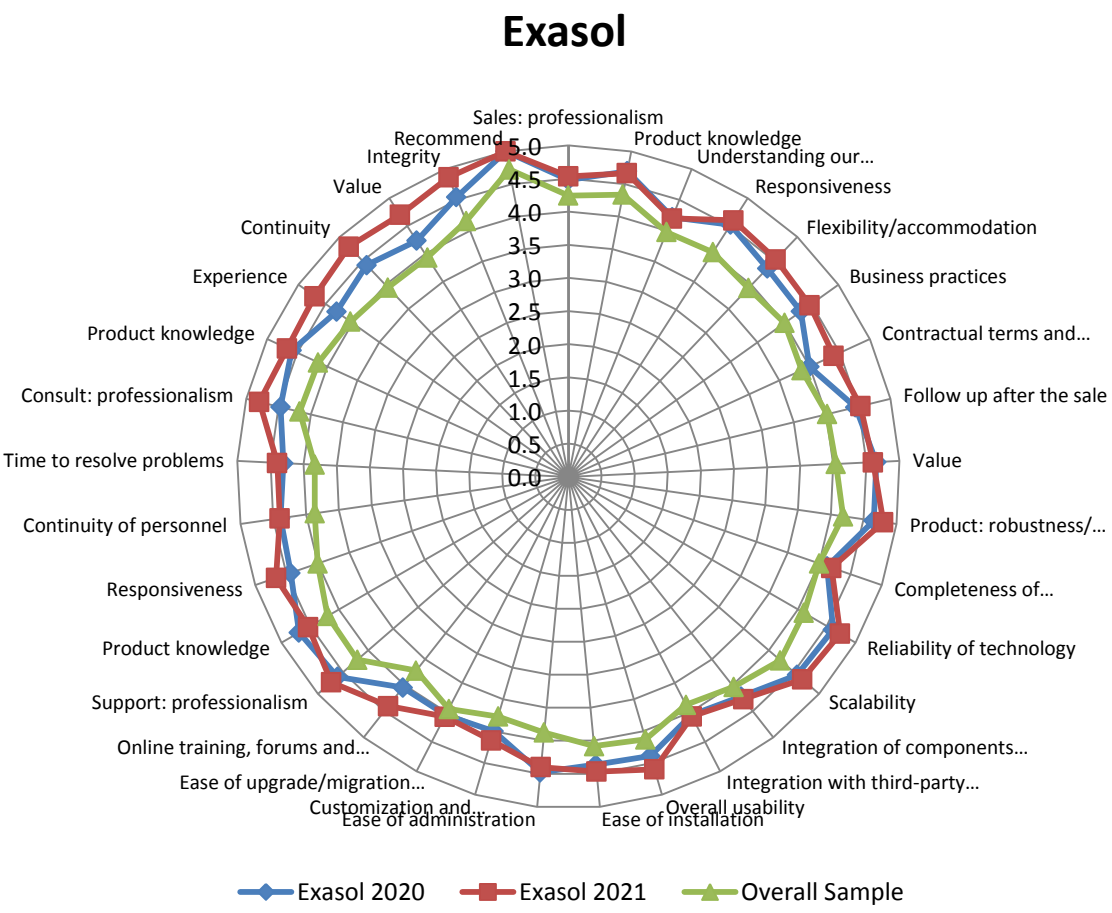


Figure 80 – Exasol detailed score

With scores significantly above the overall sample, Exasol remains an Overall Leader in both Customer Experience and Vendor Credibility models for 2021. It is best in class for product robustness/sophistication of technology, reliability of technology, online training, forums and documentation, technical support time to resolve problems, all consulting measures, and integrity. It maintains a perfect recommend score.

Google Detailed Score

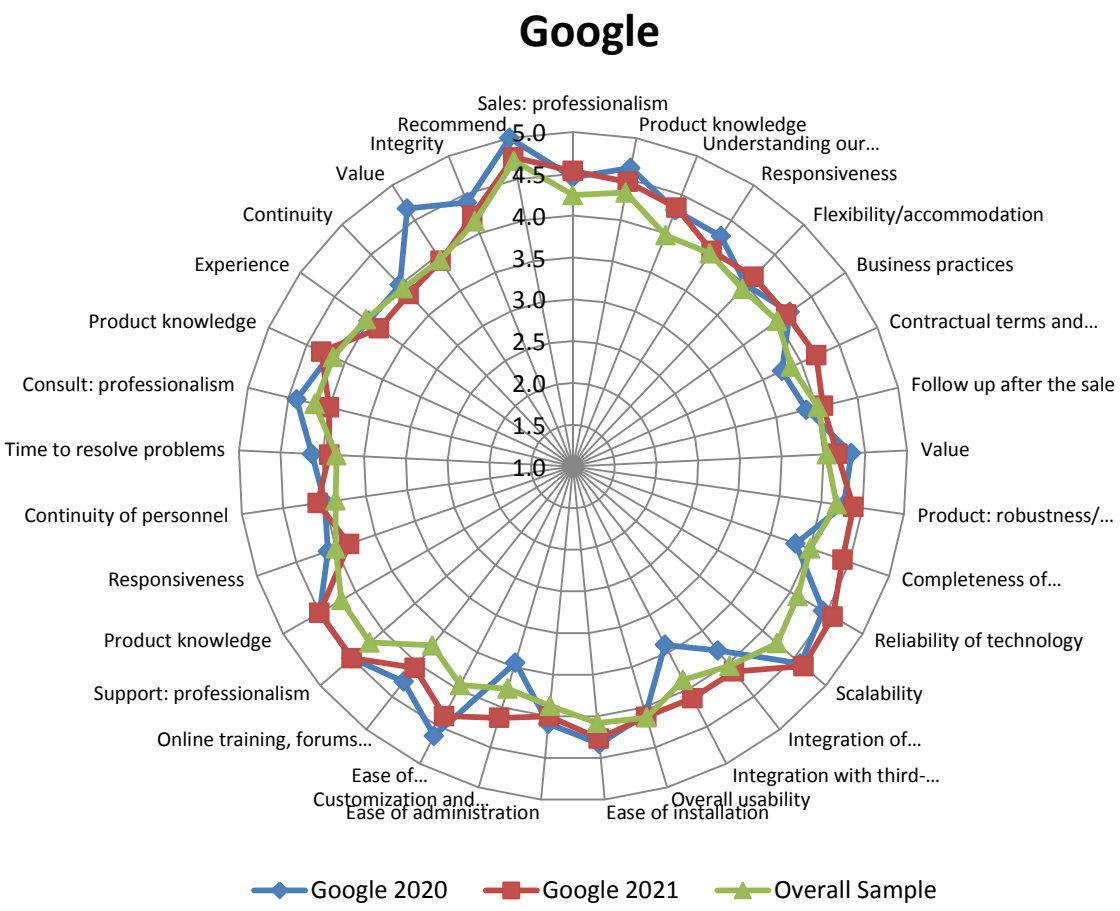


Figure 81 – Google detailed score

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 has improvements in several categories, including sales and product in 2021 and is best in class for product completeness of functionality.

IBM Detailed Score

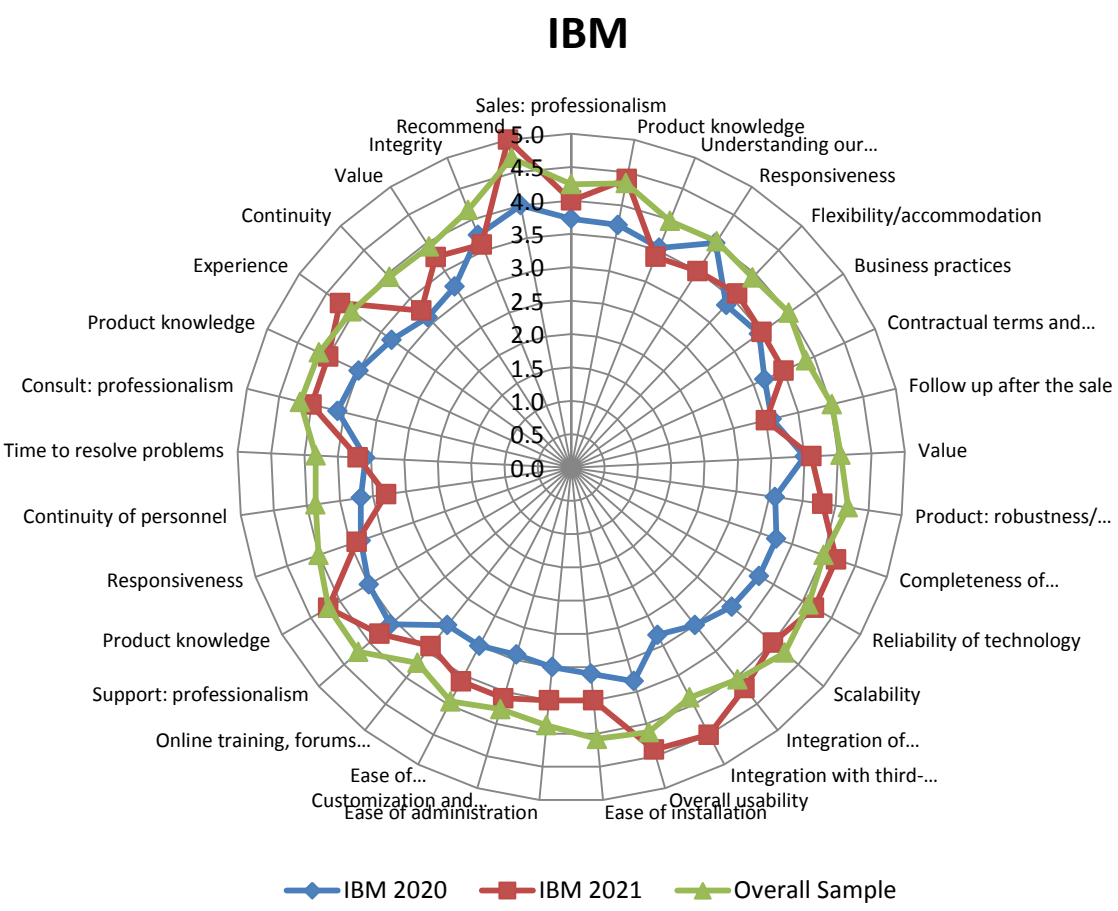


Figure 82 – IBM detailed score

For 2021, IBM's scores recover versus 2020, especially in the product category; but it still remains below the overall sample. It is considered a Technology Leader in the Customer Experience Model and a contender in the Vendor Credibility Model.

Incorta Detailed Score

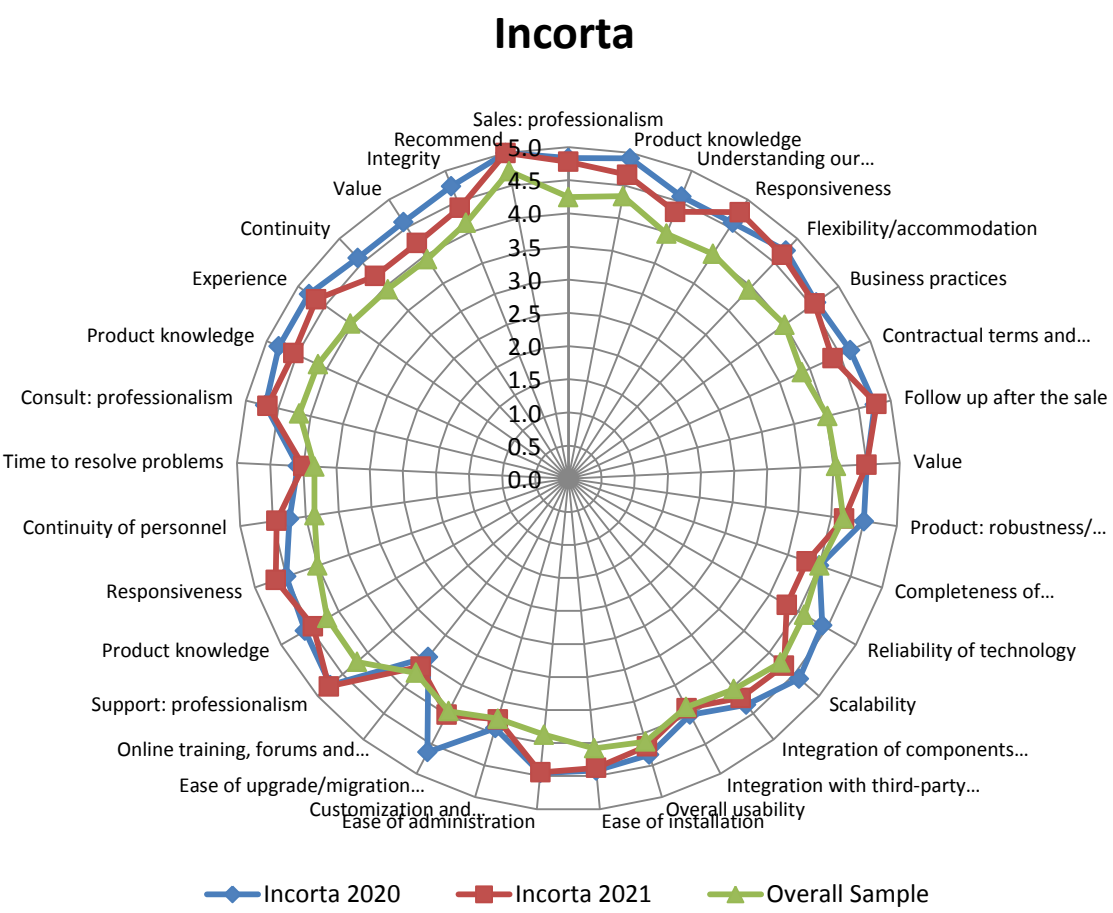


Figure 83 – Incorta detailed score

In 2021 Incorta is generally above the overall sample for most measures and is considered an Overall Leader in both the Customer Experience and Vendor Credibility models. It is best in class for a majority of sales measures, technical support professionalism, and continuity of personnel. It maintains a perfect recommend score.

Microsoft Detailed Score

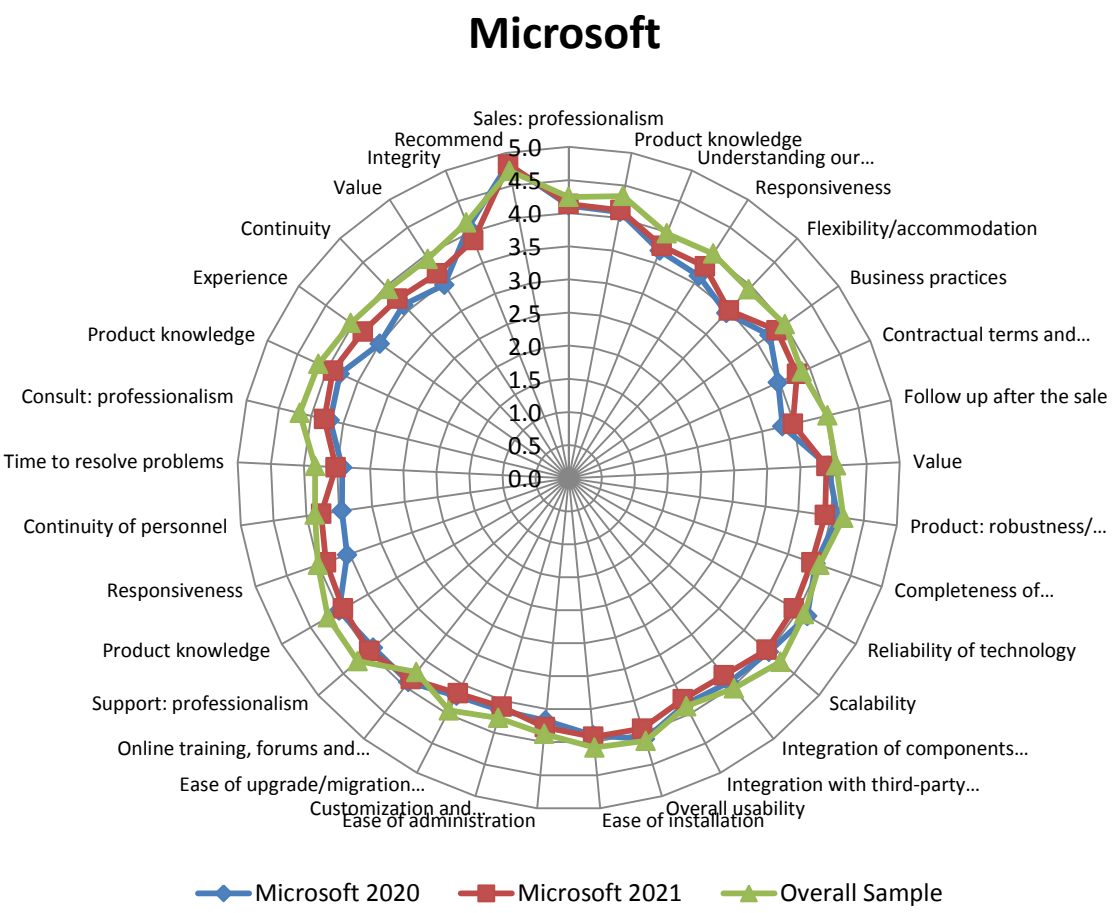


Figure 84 – Microsoft detailed score

Although generally below the overall sample, in 2021 Microsoft's scores improve in key areas of sales, consulting, and technical support. It remains a Technology Leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility model.

MongoDB Detailed Score

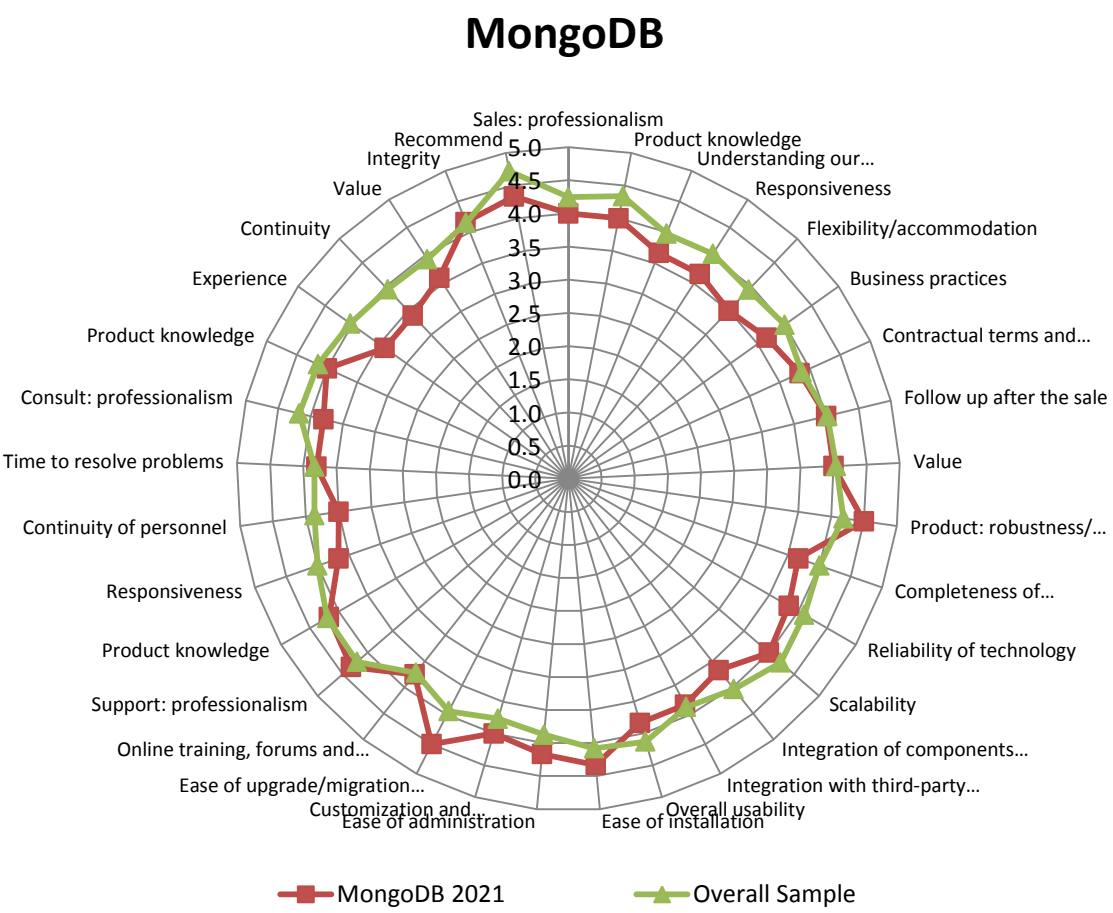


Figure 85 - MongoDB detailed score

In its first year of coverage, MongoDB's scores are generally below the overall sample. It is considered a Technology Leader in the Customer Experience Model and a Contender in the Vendor Credibility Model.

Oracle Detailed Score

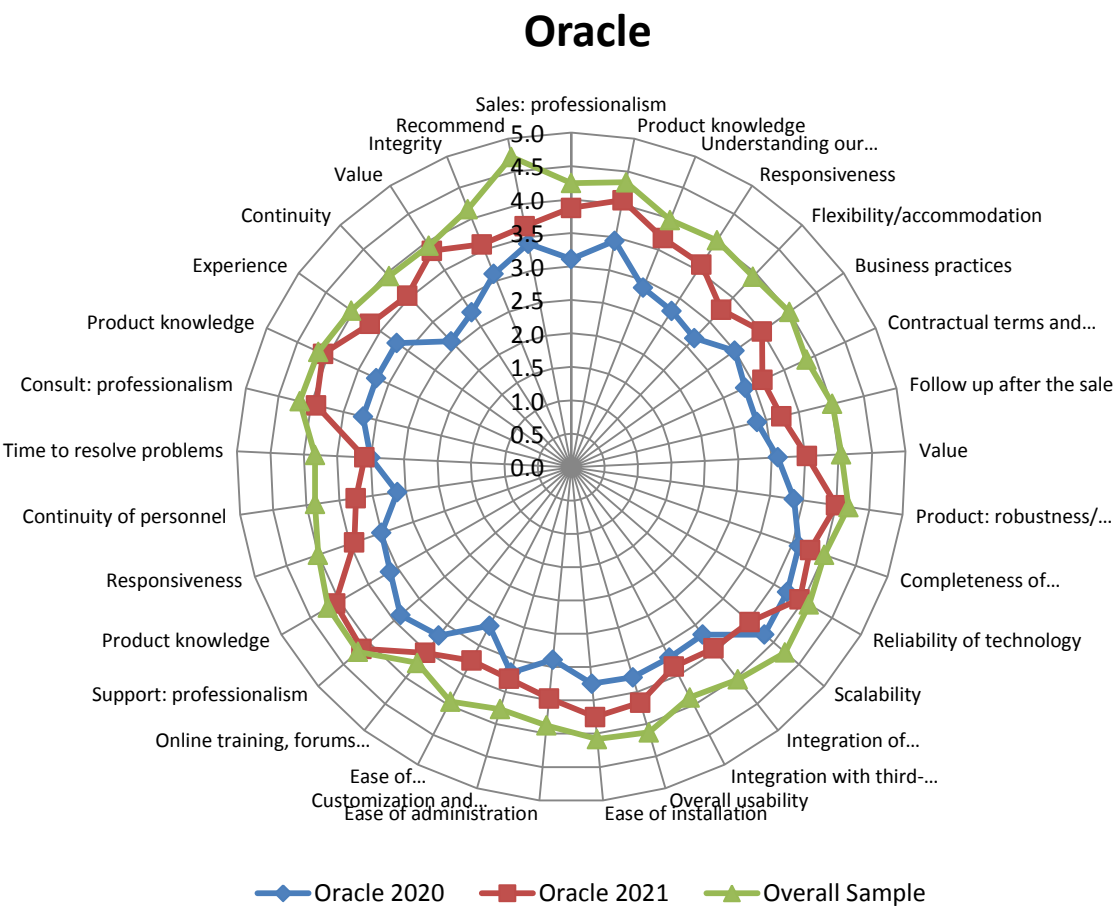


Figure 86 – Oracle detailed score

Although still below the overall sample, in 2021, Oracle's scores have improvements for most categories of measurement. It is a Contender in both the Customer Experience and Vendor Credibility models.

SAP Detailed Score

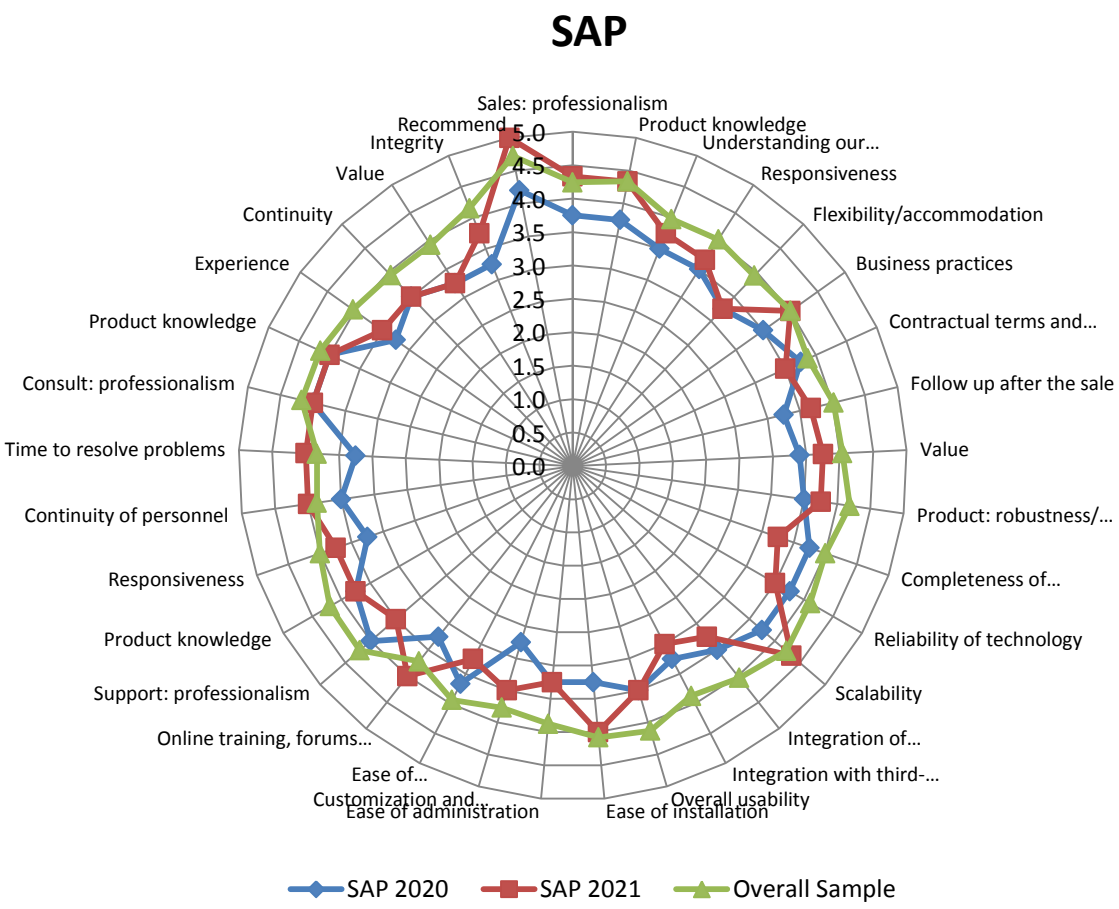


Figure 87 - SAP detailed score

Although still generally below the overall sample, in 2021, SAP continues to make improvements in most categories of measurement, as well as its overall score. It is a Contender in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model.

SAS Detailed Score

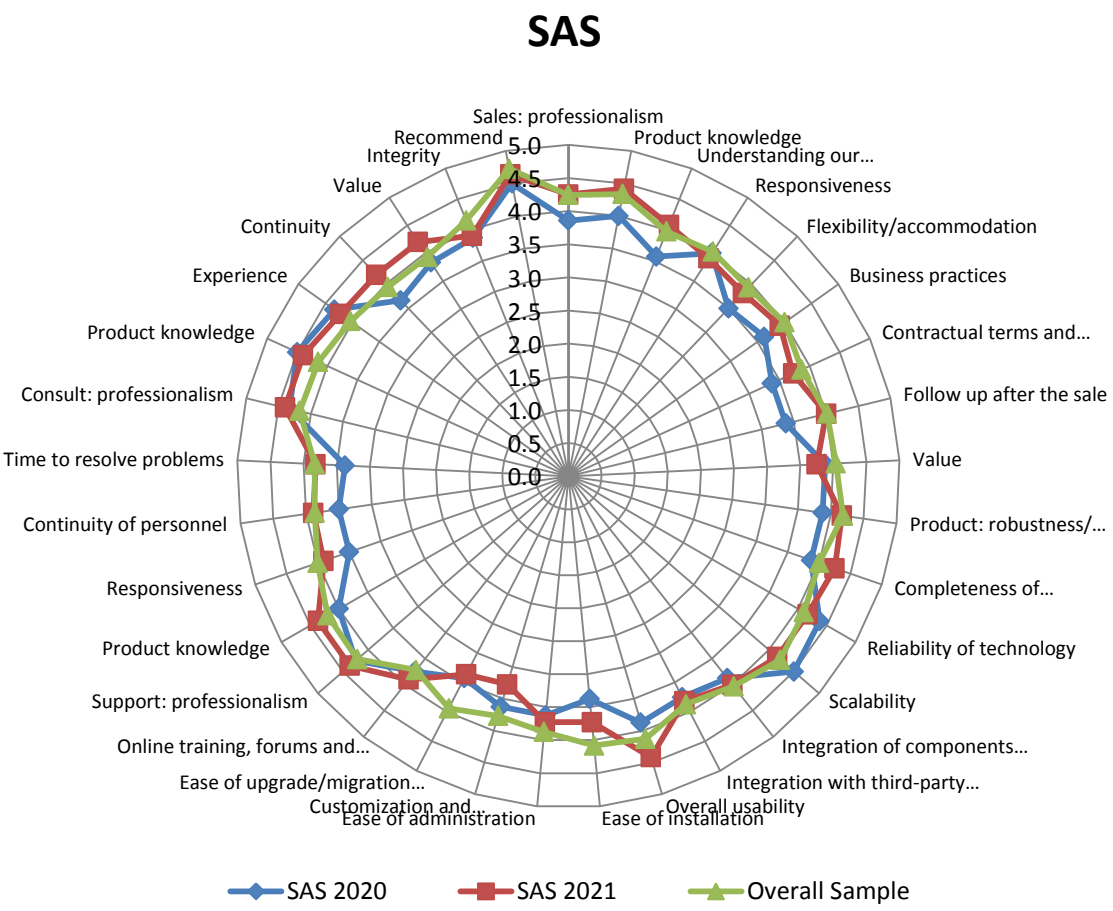


Figure 88 – SAS detailed score

In 2021, SAS has improvements in most categories of measurement and is generally in line with the overall sample. It is considered an Overall Leader in the Customer Experience Model and a Contender in the Vendor Credibility Model.

SingleStore Detailed Score

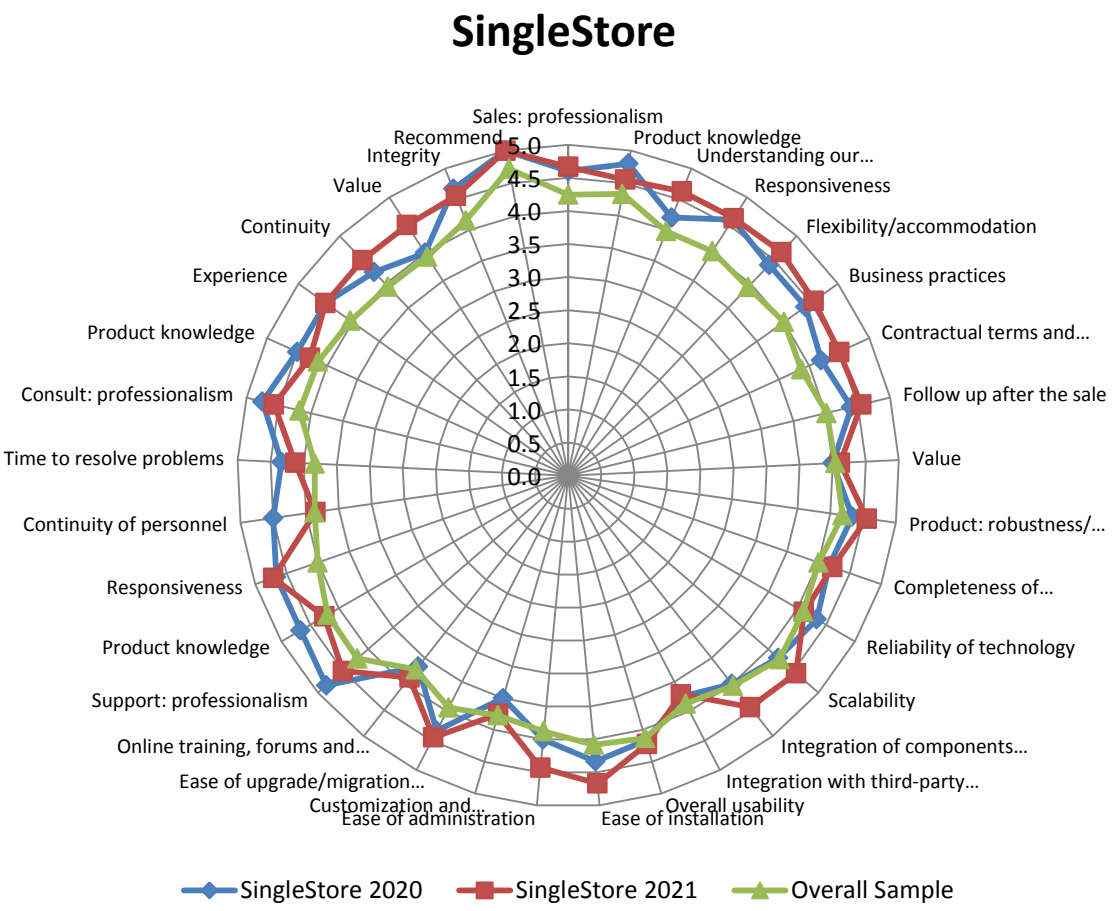


Figure 89 SingleStore detailed score

For 2021, SingleStore (fka MemSQL) has increases for a number of measures in sales, product, and consulting. In general, it is above the overall sample for all measures and is considered an Overall Leader in the both the Customer Experience and Vendor Credibility models. It is best in class for a number of sales measures including sales understanding of business needs, flexibility/accommodation, business practices, and contractual terms and conditions. It is also best in class for integration of components within product, and technical support responsiveness. It maintains a perfect recommend score.

Snowflake Detailed Score

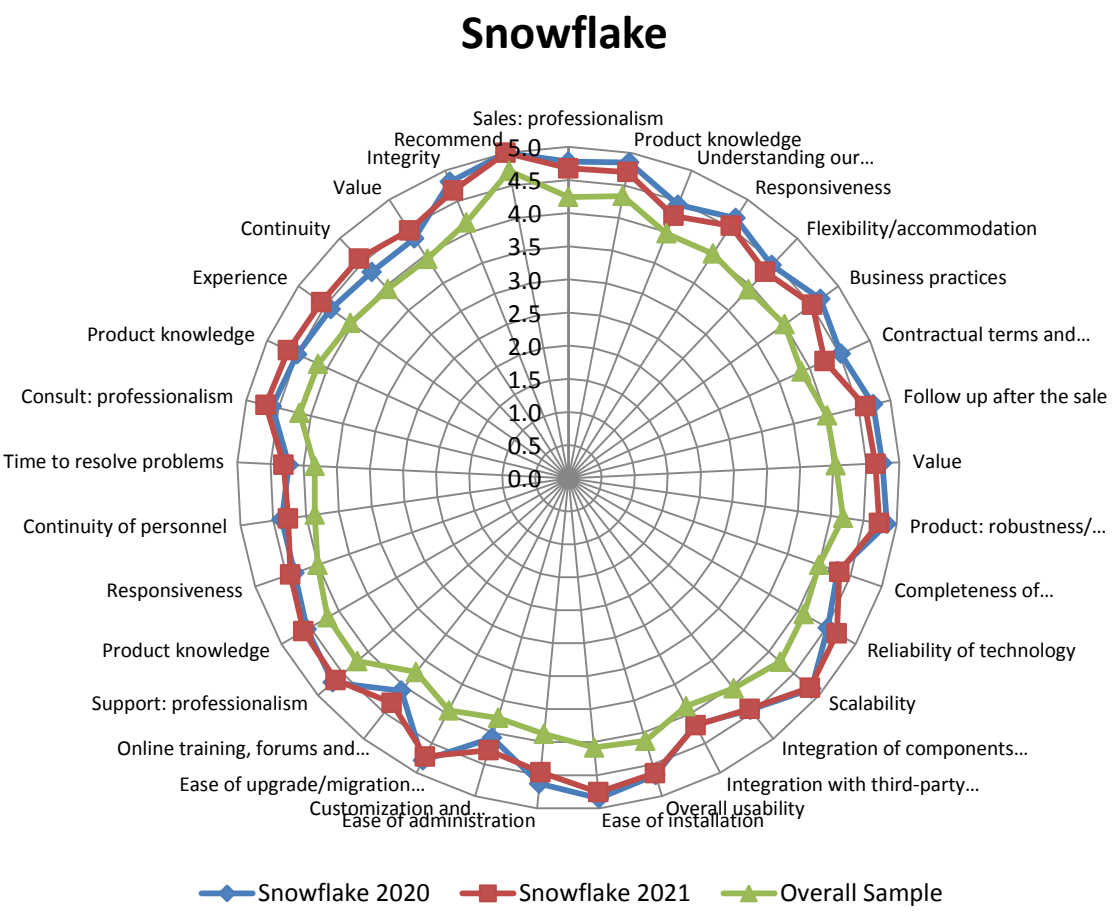


Figure 90 – Snowflake detailed score

For 2021, Snowflake continues to maintain very high scores. With scores well above the overall sample, Snowflake is an Overall Leader in both the Customer Experience and Vendor Credibility models. It is best in class for a number of product measures including scalability, usability, ease of installation, ease of administration, customization and extensibility, and ease of upgrade/migration to new versions. It is also best in class for sales product knowledge, support product knowledge, and overall value. It maintains a perfect recommend score.

Teradata Detailed Score

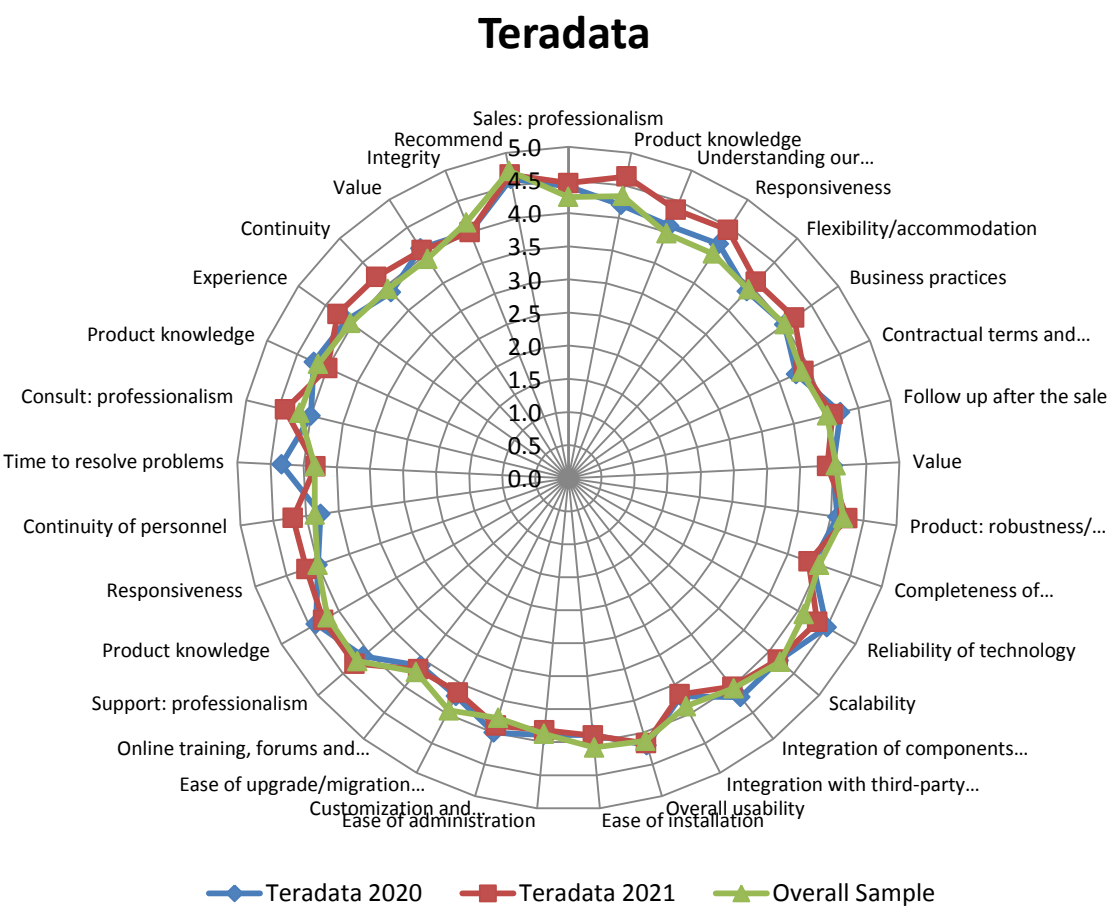


Figure 91 – Teradata detailed score

With some key improvements in sales, technical support, and consulting, for 2021, Teradata's scores are generally in line with, or slightly above, the overall sample. It is an Overall Leader in the Customer Experience Model and a Trust Leader in the Vendor Credibility Model.

Other Dresner Advisory Services Research Reports

- Wisdom of Crowds® "Flagship" Business Intelligence Market Study
- BI Competency Center
- Big Data Analytics
- Cloud Computing and Business Intelligence
- Data Catalog
- Data Pipelines and Integration
- Data Preparation
- Data Science and Machine Learning
- Embedded Business Intelligence
- Enterprise Performance Management
- Sales Performance Management
- Self-Service BI
- Small and Mid-Sized Enterprise Business Intelligence
- Small and Mid-Sized Enterprise Performance Management

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*

	Criti-cal	**Very Im-porta nt**	**Im-portan t**	**Some-what Im-portant**	**Not Im-portant**	**Don't Know**
Business user reporting and dashboards	( )	( )	( )	( )	( )	( )
Business user discovery and exploration	( )	( )	( )	( )	( )	( )
Data science (advanced and predictive analyt-ics/data mining)	( )	( )	( )	( )	( )	( )
Embedded ana-lytics within busi-ness applications (high volume, low latency)	( )	( )	( )	( )	( )	( )

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

product?*

	Critical	Very Important	Important	Somewhat Important	Not Important	Don't Know
Performance	( )	( )	( )	( )	( )	( )
Scalability	( )	( )	( )	( )	( )	( )
Price	( )	( )	( )	( )	( )	( )
Security	( )	( )	( )	( )	( )	( )
Analytical Features	( )	( )	( )	( )	( )	( )
Corporate Standard	( )	( )	( )	( )	( )	( )
Agility: Add new data, change data models, change analytics (models/algorithms)	( )	( )	( )	( )	( )	( )
Usability: Ease of development and administration	( )	( )	( )	( )	( )	( )
Adaptability: Ability to integrate with existing business applications, data infrastructure, and processes	( )	( )	( )	( )	( )	( )
Compliance or regulatory requirements	( )	( )	( )	( )	( )	( )

Deployment and licensing priorities

Please prioritize your deployment preferences for data infrastructure centers

	Criti-cal	Very Im-portant	Im-portant	Some-what Im-portant	Not Im-portant	Don't Know
As-a-cloud service	( )	( )	( )	( )	( )	( )
On-premises software	( )	( )	( )	( )	( )	( )
Hybrid	( )	( )	( )	( )	( )	( )
Cross data center inte-gration and manage-ment capa-bilities	( )	( )	( )	( )	( )	( )

Please prioritize your preferences for ADI licensing

	Criti-cal	Very Im-portant	Im-portant	Some-what Im-portant	Not Im-portant	Don't Know
User	( )	( )	( )	( )	( )	( )
Subscrip-tion	( )	( )	( )	( )	( )	( )
Concurrent use	( )	( )	( )	( )	( )	( )
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**	**Somewhat Important**	**Not Important**	**Don't Know**
Data life cycle management (lineage, impact metadata, governance/collaboration controls)	( )	( )	( )	( )	( )	( )
Application development tools (e.g., IDE, SDK)	( )	( )	( )	( )	( )	( )

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)

	Criti-cal	Very Im-Im-portant	Im-portant	Some-what Im-portant	Not Im-portant	Don't Know
Images, video	( )	( )	( )	( )	( )	( )
Text	( )	( )	( )	( )	( )	( )
Machine and events / log data	( )	( )	( )	( )	( )	( )
Transac-tional data	( )	( )	( )	( )	( )	( )
Metadata	( )	( )	( )	( )	( )	( )
Excel/CSV	( )	( )	( )	( )	( )	( )

Other sources?

_____
_____
_____
_____

Please prioritize your preferences for loading and preparing data

	Criti-cal	Very Im-portant	Im-portant	Some-what Im-portant	Not Im-Im-portant	Don't Know
ELT/ETL - Bulk load	( )	( )	( )	( )	( )	( )
In-	( )	( )	( )	( )	( )	( )

serts/updates/upserts						
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, reformatting 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

	Criti-cal	**Very Im-portant**	**Im-portant**	**Some-what Im-portant**	**Not Im-portant**	**Don't Kno w**
Hierar-chical files (e.g., Ha-doop HDFS)	( )	( )	( )	( )	( )	( )
SQL data	( )	( )	( )	( )	( )	( )
Row for-mat	( )	( )	( )	( )	( )	( )
Columnar format	( )	( )	( )	( )	( )	( )
Hybrid row/colum n	( )	( )	( )	( )	( )	( )
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?

	Criti-cal	**Very Im-Im-portant**	**Im-portant**	**Some-what Im-portant**	**Not Im-portant**	**Don't Know**
JDBC	( )	( )	( )	( )	( )	( )
ODBC	( )	( )	( )	( )	( )	( )
MDX	( )	( )	( )	( )	( )	( )
Apache Hive/Impala	( )	( )	( )	( )	( )	( )
Web/RESTful	( )	( )	( )	( )	( )	( )
XML	( )	( )	( )	( )	( )	( )
Web Ser-vices	( )	( )	( )	( )	( )	( )
Excel/CSV	( )	( )	( )	( )	( )	( )
Apache Avro	( )	( )	( )	( )	( )	( )
Apache Parquet	( )	( )	( )	( )	( )	( )

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

	Critical	**Very Important**	**Important**	**Somewhat Important**	**Not Important**	**Don't Know**
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
( ) PostgresSQL
( ) 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 Good**	**Ade-quate**	**Poor**	**Very Poor**	**Don't Know**
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 Good**	**Ade-quate**	**Poor**	**Very Poor**	**Don't Know**
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

_____
_____
_____
_____