

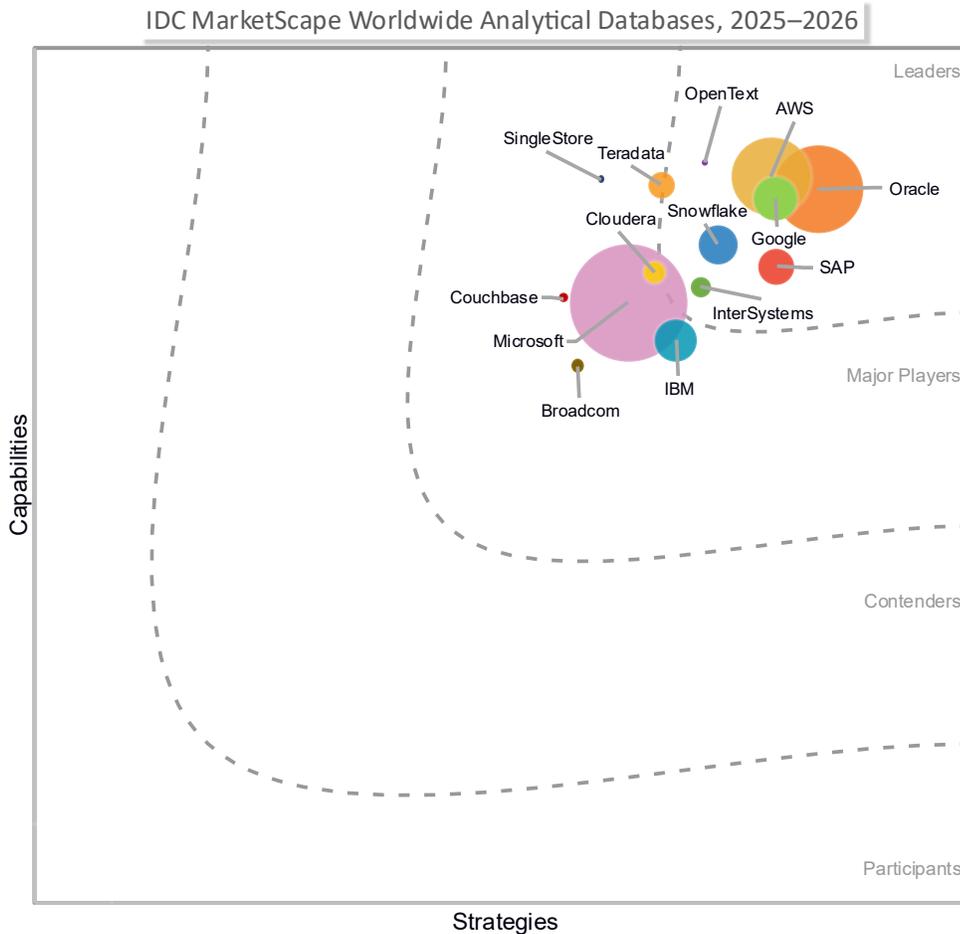
IDC MarketScape: Worldwide Analytical Databases 2025–2026 Vendor Assessment

Devin Pratt

IDC MARKETSCAPE FIGURE

FIGURE 1

IDC MarketScape Worldwide Analytical Databases Vendor Assessment



Source: IDC, 2025

See the Appendix for detailed methodology, market definition, and scoring criteria.

IDC OPINION

The analytical database market is entering a new phase of innovation as advances in storage and compute architectures redefine how enterprises extract value from data. For technology buyers, this research provides a strategic framework to evaluate analytical databases based on flexibility, interoperability, and AI readiness, the new drivers of competitive differentiation in enterprise analytics.

The decoupling of storage and compute has fundamentally redefined scalability, elasticity, and cost efficiency, extending the market beyond traditional columnar systems to include lakehouse, object store, and distributed query architectures that deliver analytics wherever data resides. This evolution shifts the focus from raw performance to architectural flexibility, operational efficiency, and long-term cost optimization. As enterprises modernize their data estates, analytical databases are becoming the performance and intelligence layer linking data storage to analytics, the foundation of modern data strategy.

In hybrid and multicloud environments, storage efficiency now determines scalability and cost predictability. Modern architectures that combine columnar compression, vectorized compute, and tiered or object-based storage help enterprises maintain high-speed access to active data while managing colder data cost effectively.

Competitive differentiation now centers on openness, elasticity, and intelligent optimization. Support for open table formats such as Apache Iceberg and Delta enhances durability and interoperability, while embedded machine learning and AI-driven tuning reduce administrative effort and accelerate performance. As organizations advance toward AI and agentic AI-enabled analytics, analytical databases are emerging as the high-performance bridge between data storage and intelligent decisioning.

For technology buyers, the takeaway is clear: success in choosing the right analytical database platform comes from selecting solutions that unite performance with intelligence. Organizations that adopt platforms combining optimized storage, elastic compute, strong governance, and open interoperability gain the agility, scalability, and insight needed to lead in the data-driven era. The vendors evaluated in this IDC MarketScape demonstrate these capabilities, providing the trusted foundation for intelligent, data-driven enterprises.

IDC MARKETSCOPE VENDOR INCLUSION CRITERIA

A critical point in this research effort is to meet the following criteria:

- **Core offering:** The vendor must offer an analytical database solution as a core product. This solution should be purpose built for analytics rather than transactional workloads.
- **High-performance querying:** The database must support fast, complex analytical queries optimized for read-heavy operations.
- **Scalability:** The solution must demonstrate the ability to scale horizontally (e.g., by adding nodes) or vertically (e.g., increasing resources on a single node) without performance degradation.
- **Data storage optimization:** The database must employ advanced storage models optimized for analytics, such as columnar storage or hybrid approaches, to enhance query efficiency and compression.
- **Concurrency:** The database must handle high levels of concurrent queries and users effectively.
- **Global presence:** Vendors must operate in North America and generate at least 10% of their revenue from the Americas, EMEA, and APAC regions.
- **Revenue threshold:** Vendors must have at least \$25 million in company revenue for 2024.
- **Vendor IP ownership:** The core components of the analytical database (e.g., query engine and storage optimization) must be the intellectual property of the vendor. Partnered or OEM-provided IP is acceptable for supplementary features such as data transformation or governance.
- **Strategic importance:** The analytical database being evaluated must represent a strategic offering within the vendor's portfolio.

ADVICE FOR TECHNOLOGY BUYERS

Based on IDC's research and observed market dynamics, organizations evaluating analytical database platforms should consider the following guidance:

- **Prioritize scalability and workload adaptability.** Analytical databases must deliver predictable performance across diverse workloads, including batch, interactive, and real-time analytics. Buyers should select solutions that automatically scale resources to meet performance objectives as workloads grow or fluctuate. Evaluate whether platforms scale horizontally and vertically without performance loss. Architectures that separate compute from storage or use massively parallel processing (MPP) provide the elasticity needed to meet

dynamic business demand. Organizations should ensure that scaling strategies align with governance, reliability, and cost management goals to sustain performance, predictability, and efficiency. These capabilities form the foundation for both current analytical workloads and emerging AI and agentic AI initiatives.

- **Evaluate openness, interoperability, and deployment flexibility.** Hybrid and multi-cloud architectures are now the standard for enterprise analytics. Analytical databases should support open formats such as Apache Iceberg, Delta, or Parquet and integrate with catalogs, orchestration, and governance tools. Confirm that platforms support SaaS, bring your own cloud, or containerized deployment models to meet compliance, sovereignty, and data residency needs. Assess how easily a platform integrates with existing data lakes, pipelines, and BI tools to reduce implementation effort and avoid vendor lock-in. Solutions that deliver consistent performance and governance across environments simplify modernization, strengthen interoperability, and establish the open data foundation needed for both current analytics and future AI ecosystems.
- **Align analytical depth with AI and emerging workload readiness.** The analytical database market is rapidly evolving to support AI and machine learning directly within the database engine. Buyers should evaluate how effectively vendors embed capabilities such as in database machine learning, vector search, graph analytics, and integration with generative AI frameworks. Determine whether these capabilities align with the organization's data science maturity and operational goals. AI features should be usable and governable within existing analytics workflows. Platforms that unify analytics and AI simplify architectures, improve performance, and accelerate innovation. Readiness for retrieval augmented generation and vector-enabled querying is becoming a key differentiator as enterprises expand AI-driven and agentic workloads.
- **Consider converged workloads built on hybrid transactional and analytical processing (HTAP) architectures for real-time decisioning.** Modern analytical database platforms increasingly support converged workloads that unify transactional and analytical processing within a single system. HTAP enables organizations to analyze live operational data without replication or delay, providing the foundation for real-time decisioning. Buyers should assess whether these capabilities meet requirements such as fraud detection, personalization, or supply chain visibility while maintaining transactional integrity and analytical performance on the same data. Platforms that support both workloads improve responsiveness, reduce data movement, and strengthen governance. HTAP architectures are also becoming essential for agentic AI, where real-time, context-rich data enables autonomous, data-driven decisioning across the enterprise.

VENDOR SUMMARY PROFILES

This section briefly explains IDC's key observations resulting in a vendor's position in the IDC MarketScape. While every vendor is evaluated against each of the criteria outlined in the Appendix, the description here provides a summary of each vendor's strengths and challenges.

AWS

After a thorough evaluation of AWS strategies and capabilities, IDC has positioned AWS in the Leaders category in this 2025 IDC MarketScape for worldwide analytical databases.

Amazon Web Services (AWS) is a global provider of cloud infrastructure, and Amazon Redshift is positioned as its flagship analytical database platform. Delivered as a fully managed, cloud-native service, Amazon Redshift supports enterprise-scale workloads across structured and semistructured data. Its architecture separates compute from storage, giving organizations the choice between provisioned clusters for predictable operations and Amazon Redshift Serverless for elastic, consumption-based analytics.

Amazon Redshift consolidates analytical workloads into a single environment, supporting SQL-based querying, semistructured formats, and direct access to data in Amazon S3. Features such as AI-driven scaling and optimization, workload management, materialized views, and data sharing across clusters improve operational efficiency while sustaining performance consistency. Integration with AWS data services — including relational and non-relational databases like Amazon Aurora and Amazon DynamoDB via zero-ETL, Amazon Bedrock for GenAI apps, Amazon Kinesis and Amazon Managed Streaming for Apache Kafka for near-real-time analytics, AWS Glue Data Catalog for cataloging, AWS Lake Formation for governance, and Amazon SageMaker AI for advanced analytics — extends Redshift's role beyond the warehouse into a unified, end-to-end data platform. Amazon Redshift also integrates with the next generation of Amazon SageMaker, supporting high-performance data warehouse workloads on unified data across the Amazon SageMaker lakehouse architecture. These capabilities help enterprises balance elasticity with governance, scalability, and cost optimization.

As part of the broader AWS ecosystem, Amazon Redshift provides a unified foundation for analytics modernization within AWS, combining scalability, security, and reliability for mission-critical workloads.

Strengths

- **Elastic architecture and scalability:** Amazon Redshift separates compute from storage, offering provisioned clusters for predictable operations and Amazon

Redshift Serverless for elastic, consumption-based workloads. This design enables organizations to manage both steady reporting and highly variable workloads without overprovisioning infrastructure. AI-driven scaling and optimization, workload isolation, and automated resource management allow enterprises to adapt capacity dynamically as demand shifts. By aligning infrastructure directly with business needs, Redshift helps organizations balance scalability and cost efficiency while maintaining reliability for mission-critical analytics environments.

- **Performance optimization at scale:** Amazon Redshift uses columnar storage, massively parallel processing, result caching, materialized views, and advanced workload management to deliver predictable performance under heavy concurrency. These optimizations allow enterprises to sustain throughput across a wide range of queries, from interactive dashboards to complex joins on petabyte-scale data sets. Performance consistency ensures that teams can run diverse workloads on a shared platform without degradation. By embedding optimization features directly into the platform, Redshift reduces administrative overhead and helps enterprises maintain efficiency as analytical demands grow.
- **Deep integration with the AWS Ecosystem:** Amazon Redshift integrates seamlessly with AWS storage, analytics, streaming, AI, and database, including the next generation of Amazon SageMaker, creating an end-to-end environment for data ingestion, transformation, analysis, and machine learning. This integration enables enterprises to govern data consistently across the AWS platform while reducing reliance on third-party tools. IAM integration, data sharing across accounts, and compatibility with open formats further simplify operations. By leveraging existing AWS skills and services, organizations can accelerate modernization initiatives, shorten time to value, and create unified pipelines within their existing cloud environment.

Challenges

- **Cross-cloud portability and ecosystem strategy:** As organizations adopt multicloud and hybrid analytics models, ensuring consistent governance, cost alignment, and interoperability across environments remains a strategic challenge. Amazon Redshift's deep integration within the AWS ecosystem offers efficiency advantages but requires thoughtful planning to optimize participation in broader, cross-cloud data ecosystems.
- **Cost governance for variable workloads:** Amazon Redshift provides flexibility through provisioned clusters, serverless usage, and concurrency scaling, but these options require strong governance to maintain predictable economics. Enterprises with highly variable workloads should balance elasticity with cost efficiency through deliberate workload planning and monitoring.

- **Tuning for mixed workloads:** Amazon Redshift is designed to run diverse workloads, but balancing interactive queries, batch jobs, and data sharing on shared resources requires active tuning. Administrators must configure workload management rules, optimize sort and distribution keys, and adjust resource allocation to sustain consistent performance. As environments grow in complexity, operational discipline becomes increasingly important. These capabilities allow Amazon Redshift to serve as a consolidated analytics platform but also require specialized expertise and ongoing oversight, which increases administrative burden for organizations without mature data operations.

Consider AWS When

Consider Amazon Redshift when modernizing analytics within AWS is a priority. The platform is well suited for organizations requiring elastic scalability, SQL familiarity, and strong governance for regulated industries. Amazon Redshift is also relevant for enterprises analyzing structured and semistructured data, leveraging Amazon S3 directly, or managing dynamic workloads with a mix of provisioned and serverless deployments. Organizations seeking to balance modernization, governance, and performance will find Amazon Redshift a practical option.

Broadcom

After a thorough evaluation of Broadcom's strategies and capabilities, IDC has positioned Broadcom in the Major Players category in this 2025 IDC MarketScape for worldwide analytical databases.

Broadcom is a global technology provider with a diversified portfolio spanning semiconductors, networking, infrastructure, and enterprise software. Within its enterprise software division, VMware Tanzu Data Intelligence is positioned as one of the prominent data lakehouse-based solution for private cloud. Tanzu Data Intelligence includes an analytical database offering, combining a PostgreSQL foundation with the scalability of a massively parallel processing architecture. This design enables enterprises to manage high-volume analytical workloads while leveraging the familiarity and flexibility of SQL.

Tanzu's analytical database capability, Tanzu Greenplum, supports deployment across on-premises and public cloud environments, with options for containerized and Kubernetes-based operations that align with hybrid and multicloud strategies. Its MPP architecture distributes queries across nodes to deliver predictable performance for complex workloads. Support for advanced SQL, geospatial, and time-series analytics, along with compatibility with open source extensions, allows organizations to address diverse analytical requirements without managing multiple siloed systems.

With proven adoption across industries such as financial services, government, and healthcare, Tanzu Greenplum is recognized for its maturity and reliability in enterprise-scale environments. Integration with Broadcom's broader data, infrastructure, and security portfolio extends Tanzu Greenplum's role in modernization initiatives that demand scalable performance, hybrid deployment flexibility, and enterprise-grade support.

Strengths

- **MPP scalability and predictable performance:** Tanzu Greenplum uses a shared-nothing, massively parallel processing architecture that distributes data and queries across nodes. This design delivers predictable throughput for complex SQL workloads at enterprise scale. Enterprises can increase capacity linearly by adding nodes, allowing the platform to handle terabyte- to petabyte-scale data sets while preserving performance. By maintaining familiar SQL development patterns, Tanzu Greenplum enables organizations to consolidate workloads and sustain performance consistency across diverse analytical requirements.
- **Deployment flexibility and control:** Tanzu Greenplum supports on-premises, public cloud, and Kubernetes deployments, aligning with private, hybrid, and multicloud strategies. Enterprises retain control over data locality and infrastructure choices, meeting governance and compliance requirements in regulated industries. This flexibility enables organizations to modernize incrementally without being forced into a full cloud migration. By supporting multiple deployment models, Tanzu Greenplum provides continuity and control while extending analytical capabilities.
- **SQL extensibility and ecosystem familiarity:** Tanzu Greenplum builds on the PostgreSQL ecosystem, offering advanced SQL functionality, extensions, and in-database analytics. This heritage provides compatibility with a large set of open source tools and a wide skills base, lowering adoption friction. Enterprises can leverage existing SQL expertise while extending analytics to cover geospatial, time-series, and machine learning workloads. This extensibility helps organizations reduce reliance on multiple specialized data marts by consolidating them into a unified analytical environment.

Challenges

- **Interoperability and open format alignment:** Tanzu Greenplum's alignment with open table formats and cross-engine interoperability is narrower than alternatives designed for shared lake storage. Organizations adopting open data architectures may require additional integration to ensure smooth participation in heterogeneous environments. This limitation can increase complexity for

enterprises standardizing on open formats for long-term durability. Developing a clear integration strategy for catalogs and formats is often necessary to reduce friction in hybrid or multi-engine estates.

- **Cost efficiency and capacity planning at scale:** While Tanzu Greenplum operates under a fixed-cost model, achieving efficient storage and compute utilization at scale still requires careful planning and governance. Enterprises must implement workload partitioning, life-cycle management, and optimization strategies to ensure consistent performance and predictable economics. Tanzu Greenplum provides tools for resource management, resource utilization, and user chargebacks to aid in this process. Without proactive management, resource contention and inefficiencies may erode performance, even if total cost remains fixed.
- **Integrated AI and data science experience:** Tanzu Greenplum is optimized for SQL-based analytics, while advanced AI or machine learning workloads typically depend on external frameworks such as TensorFlow, Spark, or Python libraries. This reliance adds orchestration complexity for enterprises prioritizing deeply integrated AI. Although Tanzu Greenplum supports in-database functions, its native AI and ML capabilities remain narrower than those of cloud-native providers. Enterprises pursuing AI-driven analytics should plan for complementary tools and pipelines, ensuring integration and governance frameworks are in place to align these workloads with modernization objectives.

Consider Broadcom When

Consider Broadcom Tanzu Greenplum as part of Tanzu Data Intelligence when seeking a proven analytical database with a long history of enterprise adoption, scalable MPP performance, and support for diverse data formats. It is particularly relevant for organizations consolidating large-scale data warehouses, running complex queries on petabyte-scale data sets, or adopting hybrid deployment models in regulated industries such as healthcare. Enterprises prioritizing open source compatibility and hybrid flexibility with enterprise-grade support will find Tanzu Greenplum a practical option for balancing performance, control, and modernization objectives.

Cloudera

After a thorough evaluation of Cloudera's strategies and capabilities, IDC has positioned Cloudera in the Major Players category in this 2025 IDC MarketScape for worldwide analytical databases.

Cloudera is a global enterprise data company whose analytical capabilities are delivered through the Cloudera Platform. Within this platform, Cloudera Data Warehouse is positioned as the primary analytical database service, combining massively parallel processing with open table formats to support diverse workloads.

Cloudera Data Warehouse enables organizations to run consistent analytics on premises and in the cloud, providing the flexibility to manage sensitive data locally while extending scalability and performance through cloud infrastructure.

The platform integrates open standards such as Apache Iceberg to support schema evolution, governance, and cross-engine compatibility, reducing vendor lock-in. Cloudera Data Warehouse unifies interactive BI, batch processing, and advanced analytics on shared open tables, helping enterprises simplify data landscapes and minimize system duplication. Integration with Cloudera's Data Engineering and AI services extends Cloudera Data Warehouse's reach across the full analytics life cycle, enabling seamless governance and interoperability.

Adopted across all major industries, particularly regulated industries such as financial services, healthcare, and government, Cloudera positions its analytical database as both a warehouse and lakehouse foundation. Its combination of hybrid deployment flexibility, enterprise-grade governance, and scalability makes it a practical choice for organizations modernizing analytics while maintaining compliance, performance consistency, and control. As part of Cloudera's broader hybrid data ecosystem, Cloudera Data Warehouse underscores the company's role in enabling open, governed analytics across complex enterprise environments.

Strengths

- **Hybrid deployment flexibility:** Cloudera Data Warehouse operates consistently across on-premises and cloud environments, giving IT leaders control over data locality while scaling to meet enterprise demands. This flexibility allows modernization at the organization's own pace, integrating with existing systems and maintaining consistent performance across hybrid architectures. For regulated industries with strict compliance or residency requirements, the ability to place workloads and data where they ideally helps reduce migration risk and ensures long-term operational resilience.
- **Open standards foundation:** Cloudera's alignment with open standards such as Apache Iceberg provides a consistent framework for analytics and governance across diverse data environments. This approach reduces vendor lock-in, supports interoperability among data engines, and ensures continuity as technology evolves. For organizations seeking to future-proof analytics investments, Cloudera's open foundation offers flexibility in cloud strategy while maintaining visibility, governance, and auditability across ecosystems. Recent enhancements, such as the Iceberg REST Catalog, extend this openness by enabling third-party engines to access governed data directly without duplication, reinforcing Cloudera's position as a neutral, interoperable platform within multi-engine environments.

- **Enterprise-scale performance and analytics:** Cloudera combines mature enterprise governance with the scalability to handle high-volume, multiteam analytics. The platform supports complex workloads through distributed query execution, caching, and workload isolation, helping organizations maintain responsiveness even under heavy concurrency. For large enterprises operating regulated or mission-critical analytics environments, Cloudera provides both performance consistency and the governance maturity required to align with corporate and regulatory standards.

Challenges

- **Interoperability across engines:** Cloudera's support for open formats such as Apache Iceberg provides strong cross-engine compatibility, but integration across diverse analytics frameworks can still require planning and governance. For organizations pursuing multi-engine or multicloud strategies, achieving seamless interoperability may involve additional effort to align catalogs, security, and data life-cycle management. Establishing consistent integration practices helps ensure interoperability while preserving compliance and performance across environments.
- **Operational overhead in scaling:** While Cloudera provides strong control over resources, its deployments typically require more configuration and tuning than fully serverless systems. Recent additions such as Lakehouse Optimizer and fast autoscaling for Operational Database (in preview) reduce manual effort and improve elasticity, but their availability and impact vary by deployment model. As a result, organizations running traditional or hybrid environments may still encounter higher operational complexity compared with cloud-native platforms.
- **Perceptions of legacy positioning:** Cloudera's long association with Hadoop-era big data technologies can still influence how the platform is perceived, even as it has evolved into a modern hybrid SQL and lakehouse foundation. Organizations evaluating cloud analytics options may not fully recognize the platform's advancements in openness and hybrid deployment. Enterprises should assess Cloudera based on its current platform capabilities and hybrid architecture maturity rather than historical associations.

Consider Cloudera When

Consider Cloudera when delivering consistent analytics across hybrid and multicloud environments is a strategic priority. The platform is particularly relevant for enterprises that require strong governance, compliance alignment, and control over data locality while modernizing analytical operations. With adoption across industries such as financial services, healthcare, and government, Cloudera offers a mature, production-tested foundation for organizations seeking to unify data warehousing and lakehouse architectures under a governed, open framework. Enterprises standardizing on open

table formats such as Apache Iceberg or operating in regulated, multisite environments will find Cloudera a practical choice for balancing modernization, interoperability, governance, and performance predictability.

Couchbase

After a thorough evaluation of Couchbase's strategies and capabilities, IDC has positioned Couchbase in the Major Players category in this 2025 IDC MarketScape for worldwide analytical databases.

Couchbase is a global enterprise database provider offering both self-managed (Couchbase Enterprise) and fully managed database-as-a-service options through Couchbase Capella, available on AWS, Azure, and Google Cloud. This dual model provides deployment flexibility for enterprises operating across cloud, edge, on premises, and hybrid environments, aligning modernization goals with governance and compliance needs.

Capella integrates Couchbase's document database engine with its analytics database engine, allowing organizations to query operational and historical data in one environment without extensive data movement. The platform combines JSON document storage with the SQL-compatible query language SQL++, supporting high-performance querying and real-time analytics on structured, semistructured, and evolving data sets. This integration of operational and analytical capabilities helps enterprises simplify architectures and accelerate insight generation.

Recognized for operational reliability and hybrid deployment flexibility, Couchbase positions its platform as a unified environment for transactional, mobile, and analytical workloads. By reducing data duplication, improving governance, and supporting decision-making directly on live data, the platform serves as a practical option for organizations modernizing data estates while balancing compliance, scalability, and real-time performance.

Strengths

- **Unified operational and analytical workloads:** Couchbase enables enterprises to run both transactional and analytical queries within a single environment, eliminating the need for separate systems and ETL processes. This design supports real-time, zero-ETL insight directly on operational data, allowing organizations to accelerate decision-making and automate actions for use cases such as customer engagement, fraud detection, and IoT monitoring. By consolidating operational and analytical workloads, Couchbase helps reduce data duplication, simplify architectures, and align governance across environments, enabling enterprises to modernize while maintaining efficiency.

- **Strong security and governance:** The platform delivers enterprise-grade protections, including encryption, audit logging, and privacy controls, all of which are critical for organizations operating in regulated industries such as financial services and healthcare. These capabilities ensure compliance and data protection while extending analytics into mission-critical workloads. By embedding strong governance mechanisms into the platform, Couchbase provides organizations with the assurance needed to trust their analytical environments while still enabling real-time decision-making. Enterprises in highly regulated sectors such as healthcare and banking benefit from Couchbase's ability to maintain strict compliance standards while enabling agile data-driven operations.
- **High-performance querying:** Couchbase provides strong query performance across operational and analytical workloads, supporting low-latency access to semistructured and evolving data sets. Its SQL-compatible query language (SQL++) allows teams to work with JSON documents using familiar syntax, improving adoption for SQL-trained users. Combined with its ability to support real-time analytics, uniquely columnarizing JSON data, Couchbase offers enterprises a platform that delivers speed and accessibility, making it well suited for time-sensitive use cases that demand responsiveness.

Challenges

- **Maturity of analytic services:** Couchbase's columnar analytic services are newer compared with those of long-established analytical platforms. While suitable for many scenarios, they lack the same depth of optimizations, ecosystem integrations, and advanced features as more mature offerings. Organizations with complex analytical needs or large-scale workloads require additional tools or complementary platforms to meet specialized requirements. Enterprises adopting Couchbase for analytics should evaluate its road map and consider pilots to ensure the platform aligns with long-term analytical requirements.
- **Performance at extreme scale:** Although Couchbase is effective for operational workloads, its performance at very large analytical scales, such as multi-petabyte deployments or high-concurrency environments, remains less proven. Enterprises consolidating significant analytical estates should validate performance against representative workloads before broad rollout. Capacity planning, stress testing, and proof-of-concepts are important steps to confirm that Couchbase meets enterprise performance expectations when deployed at extreme scale.
- **Licensing flexibility:** Couchbase's licensing model is less adaptable than usage-based consumption models offered by some cloud-native competitors. This can

create challenges for organizations with highly variable or unpredictable workloads, where aligning costs with actual consumption is critical. Enterprises planning to deploy Couchbase at scale should evaluate its licensing terms carefully and establish governance practices to ensure predictable economics and cost alignment with business demand.

Consider Couchbase When

Consider Couchbase when consolidating operational and analytical workloads on a single platform is a priority. The platform is particularly relevant for organizations that require real-time analytics on semistructured data, zero-ETL architecture, SQL accessibility alongside JSON flexibility, and operational reliability across hybrid or multicloud environments. Enterprises seeking to simplify data estates while maintaining strong governance and consistent performance will find Couchbase a practical choice for modernization and real-time analytics.

Google

After a thorough evaluation of Google's strategies and capabilities, IDC has positioned Google in the Leaders category in this 2025 IDC MarketScape for worldwide analytical databases.

Google is a global technology company whose cloud division positions BigQuery as its flagship analytical database and a central part of the Google Cloud data platform. Delivered as a fully managed, serverless service, BigQuery separates compute from storage, enabling independent scaling for interactive queries, batch workloads, and real-time analytics. The platform supports structured, semistructured, and increasingly unstructured data, providing elasticity for both predictable and highly variable demand.

BigQuery embeds machine learning and generative AI directly into the database and connects seamlessly with services such as Vertex AI and Looker. These capabilities allow organizations to run predictive modeling, large-scale inference, and advanced analytics directly on governed data sets without moving data. Integration with the broader Google Cloud data and AI ecosystem, including Dataplex, Dataflow, AlloyDB, and BigLake, extends BigQuery's role in unified analytics and governance across multicloud environments.

With interoperability across Google Cloud services and multicloud reach through BigQuery Omni, and integration with the expanding Data Cloud marketplace, the platform is positioned as an AI-enabled environment for enterprises modernizing analytics. For organizations balancing scalability, governance, and cross-cloud flexibility, BigQuery provides a unified foundation to combine data and AI at enterprise scale.

Strengths

- **Scalability and elastic performance:** Google BigQuery's serverless architecture separates compute from storage, allowing each to scale independently for interactive queries, batch processing, and real-time analytics. This flexibility reduces operational overhead and supports petabyte-scale analysis with consistent performance. Enterprises benefit from automatic resource management that eliminates manual provisioning, making it easier to adapt to both predictable reporting and highly variable workloads. By providing elasticity without the need for infrastructure tuning, BigQuery enables organizations to focus on business outcomes rather than platform management.
- **Integrated AI/ML and advanced analytics:** BigQuery embeds machine learning and generative AI capabilities directly in the database and integrates with services such as Vertex AI and Looker, along with Gemini incorporated across the platform. These features allow enterprises to build predictive models, run real-time inference, and operationalize AI-driven workloads directly on governed data sets. By reducing data movement, BigQuery lowers complexity while accelerating adoption of advanced analytics. This integration positions the platform as a strong option for organizations seeking to embed AI capabilities into everyday analytical workflows.
- **Interoperability and ecosystem integration:** BigQuery seamlessly integrates with Google Cloud services such as Dataflow, Dataproc, and Pub/Sub while also supporting governance through Dataplex Universal Catalog and multicloud access via BigQuery Omni. This ecosystem alignment allows enterprises to extend analytics into diverse environments while maintaining governance and consistency. Compatibility with open standards provides flexibility and reduces lock-in risk for organizations adopting hybrid or multicloud strategies. By unifying data engineering, BI, and AI workflows, BigQuery helps enterprises streamline modernization initiatives and expand use cases.

Challenges

- **Architectural alignment:** While BigQuery has significantly expanded flexibility through native Apache Iceberg support and Object Tables for unstructured data, organizations pursuing a fully open, multi-engine lakehouse strategy should assess how Google's integrated, managed approach aligns with their interoperability and data-portability goals. BigQuery supports open table and file formats within Google Cloud, allowing compatibility with open ecosystems, while layering additional value through its integrated, serverless architecture.
- **Cost predictability controls:** BigQuery now offers extensive flexibility across on-demand, capacity-based, and autoscaling pricing models, enabling customers to optimize for performance and cost. However, as usage scales across distributed

teams and projects, maintaining predictable economics still depends on strong FinOps practices, proactive budget controls, and governance visibility to prevent fragmented spend.

- **Operational readiness for advanced capabilities:** BigQuery's architecture is designed to simplify, not complicate, advanced analytics by integrating capabilities such as machine learning, streaming, and geospatial analysis within a unified, serverless platform. This consolidation reduces infrastructure overhead and minimizes the need for specialized tools. However, realizing the full value of these capabilities still requires organizational maturity, particularly in areas like SQL proficiency, model governance, and data operations. The introduction of autonomous agentic capabilities is fundamentally changing this by democratizing access to these features, without needing extensive up-front investment in training or specialized in-house expertise.

Consider Google When

Consider Google BigQuery when modernizing enterprise analytics with a serverless database is a priority. The platform is particularly relevant for organizations that need elastic performance at petabyte scale, want to embed AI and machine learning directly into analytical workflows, and require strong interoperability across cloud environments. Enterprises adopting multicloud strategies where governance, interoperability, and AI-driven analytics are strategic priorities will find BigQuery a practical choice for balancing modernization and scalability.

IBM

After a thorough evaluation of IBM's strategies and capabilities, IDC has positioned IBM in the Major Players category in this 2025 IDC MarketScape for worldwide analytical databases.

IBM is a global enterprise technology provider with a long history in data management and analytics. Db2 is positioned as IBM's core relational database platform supporting both transactional (OLTP) and analytical (OLAP) applications. Within the analytical segment, Db2 is available as Db2 Warehouse (self-managed) and Db2 Warehouse on Cloud (SaaS), each optimized for high-performance analytics through IBM's BLU Acceleration technology. BLU Acceleration combines columnar storage, in-memory processing, compression, and vectorized execution to enhance performance for scan-heavy SQL queries at scale. For enterprises with large, complex environments, Db2 Warehouse extends this capability through a massively parallel processing architecture designed for high concurrency and complex joins. In-database analytics for SQL, Python, and R reduce data movement and enable advanced statistical workloads to run close to the data.

Deployment flexibility remains a key differentiator for Db2. Organizations can run it self-managed on premises or in public clouds, consume it as a managed service on IBM Cloud or AWS, or extend managed services into their own virtual private cloud through a bring-your-own-cloud (BYOC) model, now also supported on Microsoft Azure to broaden hybrid and multicloud options. Integration with IBM Cloud Pak for Data and watsonx reinforces Db2's role within IBM's broader analytics and AI ecosystem, linking data management with machine learning and governance initiatives. This range of deployment choices allows enterprises the flexibility to run Db2 Warehouse wherever it is an ideal match, on premises or in the cloud, while aligning data platforms with governance, security, and residency requirements while maintaining portability across environments.

Db2 Warehouse offers predictable analytical performance, hybrid deployment flexibility, and enterprise-grade reliability. Its combination of scale-out architecture, columnar acceleration, and governance alignment makes it well suited for regulated industries and organizations modernizing legacy data warehouse environments.

Strengths

- **Cloud-native architecture and cost-efficient performance:** The latest generation of IBM Db2 Warehouse introduces cloud-native support for Db2 column-organized tables stored directly in cloud object storage such as Amazon S3. This enhancement provides a scalable, available, and durable storage tier suited to large data volumes. Combined with new caching and query-optimization features, Db2 Warehouse improves performance efficiency and scalability for analytical workloads. Built on IBM's long-standing database engine innovation, these updates position Db2 Warehouse as a flexible platform for managing and analyzing large data sets across hybrid and multicloud environments. Fully managed service options also enable elastic scaling to support enterprise-class workloads.
- **Open data format integration and unified analytics:** The latest release extends Db2's data management capabilities with support for open table and file formats such as Iceberg, Parquet, ORC, and CSV, enhancing interoperability across diverse analytics ecosystems. Db2 Warehouse integrates with watsonx.data, IBM's open lakehouse platform, helping organizations unify data across warehouses and lakes for consolidated analytics and AI development. Features such as schema-level backup and automatic storage expansion simplify administration and business-continuity planning, supporting modernization efforts while maintaining governance and compliance standards.
- **Hybrid deployment and BYOC flexibility:** Db2 Warehouse supports deployment across on-premises environments, IBM Cloud, AWS, and Azure. In addition to managed services on IBM Cloud and AWS, a bring-your-own-cloud

option on Azure enables organizations to deploy managed Db2 environments within their own virtual private clouds. This model provides additional control over data governance and security while maintaining consistency with IBM-managed environments. The ability to align deployments with regulatory, security, and data-residency requirements supports modernization across hybrid and multicloud strategies.

Challenges

- **Operational model compared with serverless platforms:** While Db2's compute sizing and scaling require up-front configuration, its workload and storage management are largely automatic and adaptive, minimizing ongoing administrative effort. Manual tuning is generally reserved for advanced scenarios where organizations want explicit control over resource allocation. This model differs from fully serverless platforms that automate all scaling, but it offers more predictable performance and cost, offering flexibility for enterprises that prefer greater control.
- **Workload sensitivity and optimization flexibility:** Db2's BLU Acceleration technology is optimized for analytical workloads and supports both scan-intensive and highly selective queries through optional primary and secondary indexes on column-organized tables. The design allows enterprises to run diverse workloads within a single schema design. However, tuning decisions around table organization, indexing, and concurrency remain important to ensure consistent performance in mixed or unpredictable environments. Db2's support for both columnar and row-based storage structures helps balance performance and efficiency across operational and analytical workloads.
- **Comparative breadth of advanced analytics:** Db2 provides built-in analytics for SQL, Python, and R, enabling predictive and statistical functions to execute close to the data with minimal latency. For advanced AI and machine learning workflows, such as automated feature engineering or model life-cycle management, Db2 integrates with IBM watsonx and open source frameworks. This modular approach offers flexibility to build scalable AI pipelines using their preferred tools, though it can introduce additional integration considerations compared with platforms offering fully embedded machine learning capabilities.

Consider IBM When

Consider IBM Db2 when your organization requires consistent, high-performance analytics and seeks flexible deployment across on-premises, hybrid, and public cloud environments. Db2 is well suited for enterprises modernizing legacy analytical platforms, particularly in highly regulated industries that prioritize governance, data residency control, and business continuity. Its support for open data formats, in-database analytics, and native vector operations makes it a strong fit for organizations

advancing toward AI-enabled analytical workloads. Organizations that already leverage IBM technologies, such as watsonx, Cloud Pak for Data, or IBM Cloud, can benefit from Db2's ecosystem integration, shared governance capabilities, and alignment with IBM's broader data and AI strategy.

InterSystems

After a thorough evaluation of InterSystems' strategies and capabilities, IDC has positioned InterSystems in the Leaders category in this 2025 IDC MarketScape for worldwide analytical databases.

InterSystems is a global data technology provider with a focus on industries where performance, reliability, and compliance are critical. Its flagship database platform, InterSystems IRIS, combines transactional and analytical capabilities in a unified multimodel engine. The platform supports relational, document, key value, object, vector, and time-series data models, enabling hybrid transactional and analytical processing within a single environment.

InterSystems IRIS is designed for both vertical and horizontal scalability, delivering high throughput and low latency for mission-critical workloads. Its architecture allows organizations to run real-time analytics directly on operational data, reducing duplication and accelerating decision-making. Its applications that are built on InterSystems IRIS, including InterSystems HealthShare and TrakCare, reinforce InterSystems IRIS's position in regulated, data-intensive industries, particularly healthcare and life sciences. With portability across on-premises, private cloud, and major public cloud providers, InterSystems IRIS fits seamlessly into hybrid and multicloud strategies.

By consolidating operational and analytical workloads on one platform, InterSystems IRIS simplifies governance and improves efficiency for enterprises operating in regulated sectors. Its strengths in healthcare, financial services, and supply chain highlight its suitability for organizations that require compliance, governance, and performance at scale.

Strengths

- **Integrated multimodel and multiworkload architecture:** InterSystems IRIS unifies relational, document, key value, object, vector, and time-series models in a single engine. This multimodel approach allows enterprises to support operational, analytical, and emerging AI-driven workloads without relying on multiple systems. By enabling hybrid transactional and analytical processing directly on live data, InterSystems IRIS reduces duplication and accelerates

insight. This integration supports diverse applications across industries while simplifying governance and improving consistency in enterprise architectures.

- **Performance and scalability for real-time analytics:** InterSystems IRIS is designed for high throughput and low latency, enabling real-time analytics at enterprise scale. The platform supports both vertical scaling for resource-intensive workloads and horizontal scaling across distributed deployments, providing flexibility for organizations of all sizes. This dual approach ensures performance predictability in demanding scenarios such as financial trading, healthcare systems, and logistics optimization. Organizations can rely on InterSystems IRIS to maintain responsiveness while managing mixed workloads that combine operational and analytical demands.
- **Hybrid and multicloud deployment flexibility:** InterSystems IRIS supports deployment across on premises, private cloud, and major hyperscalers including AWS, Azure, and Google Cloud. Its portability and support for containerized, Kubernetes-managed environments make it well suited for hybrid and multicloud strategies. Bring-your-own-cloud options allow organizations to extend managed services into their own environments while maintaining governance and compliance. This deployment flexibility is particularly valuable for enterprises in regulated industries that must balance modernization with strict requirements for data locality and oversight.

Challenges

- **Accessibility for nontechnical users:** While InterSystems IRIS delivers strong technical performance, it is oriented toward developers and administrators rather than business users. Compared with cloud-native self-service platforms, InterSystems IRIS requires more specialized expertise to configure and operate. This can limit adoption across nontechnical teams and increase reliance on skilled staff. Enterprises prioritizing broad accessibility and intuitive self-service may need to invest in enablement, additional tooling, or partner solutions to expand adoption across the organization. InterSystems is addressing this challenge with its InterSystems Data Studio offering.
- **Market recognition beyond healthcare:** InterSystems has strong recognition in healthcare and clinical data platforms but is less visible in broader enterprise analytics markets. Although InterSystems IRIS offers competitive performance, hybrid flexibility, and governance features, alternatives from hyperscale vendors are often more top of mind for buyers in financial services, retail, or manufacturing. This relative lack of visibility can hinder competitive evaluations.
- **Maturing self-service and AI features:** InterSystems IRIS provides embedded support for SQL, Python, and R, as well as integration points for AI and ML workloads. These capabilities enable developers to build and deploy models

close to operational data, but more advanced automation features, such as feature engineering or full model life-cycle management, still depend on external frameworks or complementary tools. This reliance can increase integration effort and operational complexity. Organizations with ambitious AI or predictive analytics road maps should evaluate how InterSystems IRIS fits within their broader data and AI ecosystem to ensure scalability and long-term alignment.

Consider InterSystems When

Consider InterSystems IRIS when the priority is to deliver real-time analytical insight directly on operational data while supporting hybrid transactional and analytical processing. The platform is particularly relevant for enterprises in regulated industries such as healthcare, financial services, and supply chain that require high concurrent operational and analytical performance, hybrid deployment flexibility, and governance. Organizations seeking to consolidate multiple workloads on a single, multimodel engine will find InterSystems IRIS a practical option for balancing modernization, governance, and performance.

Microsoft

After a thorough evaluation of Microsoft's strategies and capabilities, IDC has positioned Microsoft in the Major Players category in this 2025 IDC MarketScape for worldwide analytical databases.

Microsoft is a global enterprise technology provider with a broad portfolio that spans cloud infrastructure, productivity platforms, and analytics services. Within this ecosystem, Azure Synapse Analytics is positioned as Microsoft's SaaS analytical database service offering a cloud-native platform designed for enterprise-scale data warehousing and analytics. Built on a massively parallel processing foundation, Synapse separates compute from storage and provides both provisioned and serverless SQL options. This dual model gives organizations flexibility to optimize for either predictable, high-volume reporting or variable, ad hoc workloads.

Synapse integrates directly with Azure Data Lake Storage, enabling organizations to query structured and semistructured formats such as Parquet and JSON without significant preprocessing or duplication. It also interoperates with Spark, extending the platform's reach into data engineering and advanced transformations within the same environment. Tight integration with Power BI, Azure Machine Learning, and Purview further connects analytics, AI, and governance, allowing enterprises to unify data discovery, modeling, and policy management under a single framework. Embedded in the broader Azure ecosystem, Synapse inherits enterprise-grade security, identity management, and compliance frameworks, making it particularly relevant for industries operating under stringent regulatory requirements.

Azure Synapse Analytics provides a central analytical database service within the Microsoft cloud, positioned to unify data warehouse and data lake capabilities. Its scalability, security alignment, and deep integration across Azure services make it a practical choice for enterprises standardizing on Azure and seeking to balance flexibility, governance, and analytical scale.

As Microsoft evolves its analytics portfolio, Synapse continues to serve enterprise-scale analytical workloads while Microsoft introduces Fabric as its emerging unified analytics platform. Fabric brings together data engineering, analytics, and AI under a single framework, providing a potential modernization path for organizations as Microsoft's analytics road map shifts toward Fabric.

Strengths

- **Dual compute modes for flexibility:** Azure Synapse Analytics offers both provisioned and serverless SQL compute in a single service, giving enterprises flexibility to match resources with workload needs. Dedicated pools provide predictable performance for steady-state analytical workloads, while serverless options allow organizations to query data on demand without pre-provisioning infrastructure. This dual model allows enterprises to balance performance and cost efficiency depending on demand patterns. The flexibility to use both compute models within the same platform helps organizations modernize incrementally while avoiding overprovisioning or underutilization.
- **Support for structured and semistructured data:** Synapse natively integrates with Azure Data Lake Storage, enabling direct querying of Parquet, CSV, JSON, and other semistructured formats. This capability allows organizations to unify warehouse and lake workloads, eliminating the need for extensive data movement or transformation. By supporting both relational and semistructured data formats, Synapse expands the range of use cases it can address while simplifying architectures. This direct interoperability with Azure Data Lake makes Synapse especially useful for enterprises seeking to consolidate analytics and data lake operations.
- **Enterprise security and compliance alignment:** Azure Synapse Analytics benefits from Microsoft's enterprise-grade security and compliance framework, which includes encryption at rest and in transit, role-based access controls, and a broad set of certifications. These capabilities make Synapse suitable for regulated industries where compliance and data protection are critical. Integration with Azure Active Directory simplifies identity and access management, while audit logging and governance features provide visibility and control. By aligning with Microsoft's unified security and compliance ecosystem, Synapse gives enterprises confidence to deploy mission-critical analytics at scale.

Challenges

- **Concurrency and workload management:** Although Synapse supports high concurrency, achieving consistent performance across hundreds or thousands of concurrent queries requires careful configuration. Workload management features, such as resource classes and workload isolation, need to be tuned to avoid contention and uneven performance. The service provides burst capacity for simultaneous queries, but overall throughput depends on workload design and configuration. This reliance on manual optimization increases administrative overhead compared with platforms that scale concurrency automatically. Microsoft has indicated that Fabric Warehouse introduces more automated workload management and elasticity to address this concurrency and workload management challenge.
- **Performance optimization and tuning requirements:** Peak performance in Synapse often depends on careful data distribution, partitioning, and indexing strategies. While the platform can deliver strong throughput, suboptimal schema design or workload configuration can reduce efficiency. This places more responsibility on database administrators and architects to apply best practices for performance optimization. Compared with cloud-native systems that emphasize automation, Synapse may require more manual intervention to achieve consistent efficiency at scale, particularly for enterprises managing highly variable or mixed workloads.
- **Cost predictability in dedicated pools:** While Synapse's dedicated pools provide strong and predictable performance, they require up-front provisioning and capacity planning. This model can complicate cost predictability for enterprises with highly variable or bursty workloads. Organizations must monitor utilization closely to avoid overprovisioning, which increases costs, or underprovisioning, which impacts performance. Although serverless options provide flexibility, enterprises relying heavily on dedicated pools should implement governance practices to align costs with business priorities.

Consider Microsoft When

Consider Microsoft Azure Synapse Analytics when consolidating data warehouse and data lake workloads on a single cloud-native service within the Azure ecosystem is a priority. The platform is particularly relevant for enterprises that require both provisioned and serverless compute options, need to query structured and semistructured data directly in Azure Data Lake, and must align analytics with Microsoft's enterprise-grade security and compliance standards. Organizations standardizing on Azure for modernization, governance, and analytics integration will find Azure Synapse Analytics a practical option for balancing flexibility, control, and scale.

OpenText

After a thorough evaluation of OpenText's strategies and capabilities, IDC has positioned OpenText in the Leaders category in this 2025 IDC MarketScape for worldwide analytical databases.

OpenText is a global enterprise technology provider with a focus on information management and analytics solutions. Its flagship analytical platform, the OpenText Analytics Database (OTAD), is based on a distributed, columnar, massively parallel processing architecture. OTAD supports structured, semistructured, and time-series workloads, providing enterprises with predictable performance for complex analytical requirements.

The platform incorporates advanced encoding and compression techniques to optimize storage efficiency and query performance. By supporting separation of compute and storage, as well as in-database machine learning, OTAD helps enterprises reduce duplication and improve cost efficiency. Integration with the wider OpenText portfolio, spanning data management, governance, and analytics applications, extends OTAD's role from a standalone analytical database to a foundation for enterprise information modernization. Its compatibility with major public clouds and support for containerized deployments provide flexibility for hybrid and multicloud strategies.

As part of OpenText's broader information management portfolio, OTAD is positioned as a mature and technically robust option for large-scale analytics. It is particularly well suited for organizations in regulated industries that require high performance, efficient storage economics, and deployment flexibility across on-premises and cloud environments.

Strengths

- **Performance and scalability at enterprise scale:** OTAD builds on a massively parallel processing columnar architecture that delivers predictable performance for large-scale analytical workloads. The platform supports petabyte-scale deployments with high concurrency and complex queries across structured, semistructured, and time-series data. Enterprises can rely on its ability to sustain performance across diverse analytical requirements without degradation. By combining mature query optimization with distributed scalability, OTAD provides a foundation for organizations seeking to run mission-critical analytics at enterprise scale.
- **Advanced compression and storage efficiency:** OTAD applies advanced encoding and compression techniques that reduce storage requirements while accelerating scan performance. Columnar storage, combined with optimized encoding, allows enterprises to minimize total cost of ownership (TCO) while

sustaining throughput for analytical workloads. These features make it especially well suited for organizations with large, complex data sets where efficiency and speed are critical. By improving storage economics without compromising performance, OTAD helps enterprises scale analytics while maintaining cost efficiency.

- **Deployment flexibility across hybrid environments:** OpenText OTAD supports deployment across on-premises, private cloud, and major public cloud platforms, giving enterprises choice and control in aligning analytics with governance and compliance requirements. Its architecture allows portability across environments, including containerized deployments managed with Kubernetes. This hybrid flexibility enables organizations to modernize incrementally, maintaining existing infrastructure investments while adopting cloud services. For regulated industries, OTAD offers a balanced approach to modernization by supporting hybrid strategies with consistency and control.

Challenges

- **Complexity in workload optimization:** Although OTAD provides strong workload management features, achieving peak performance often requires specialized expertise in schema design, partitioning, and tuning. Compared with cloud-native platforms that emphasize automation, OTAD demands more administrative oversight to configure workloads effectively. This complexity can increase reliance on experienced administrators and limit accessibility for nontechnical teams. Enterprises considering OTAD should plan for the necessary skills and governance processes to sustain consistent performance as workloads scale.
- **Cost considerations at lower scale:** While OTAD's storage efficiency and compression features reduce total data volumes, overall costs are optimized for scale so they can still be higher compared with consumption-based cloud-native services for smaller organizations or those with less demanding data requirements. When an enterprise has highly variable demand, the licensing models and dedicated infrastructure requirements of OTAD can present challenges. Therefore, organizations deploying OTAD at scale should carefully evaluate cost structures, capacity planning, and governance models to ensure predictable economics. Balancing efficiency gains with long-term total cost of ownership requires deliberate planning to align the platform with business priorities.
- **Concurrency management under peak loads:** OTAD supports high concurrency, but sustaining predictable performance during peak demand often requires careful workload management and resource configuration. Elastic scaling is less automated than in serverless platforms, which can lead to

contention under extreme concurrency. Enterprises with diverse workloads must invest in tuning strategies and governance to maintain service levels during spikes. This challenge highlights the need for planning and oversight when using OTAD in environments with large numbers of concurrent users.

Consider OpenText When

Consider the OpenText Analytics Database when the priority is to manage large-scale, complex analytical workloads with strong performance, advanced compression, and hybrid deployment flexibility. The platform is particularly relevant for enterprises seeking predictable scalability, efficient storage economics, and governance across on-premises and cloud environments. Organizations operating in regulated industries or managing large data sets will find OTAD a practical choice for balancing modernization, performance, and compliance.

Oracle

After a thorough evaluation of Oracle's strategies and capabilities, IDC has positioned Oracle in the Leaders category in this 2025 IDC MarketScape for worldwide analytical databases.

Oracle is a global enterprise technology provider with a long history in databases and enterprise data management. Its flagship cloud service, Oracle Autonomous Database, automates provisioning, tuning, patching, and scaling to reduce administrative overhead and improve reliability. In October 2025, Oracle introduced Oracle Autonomous AI Lakehouse, expanding upon the capabilities of Oracle Autonomous Database to unify data warehousing, data lake, and AI-driven analytics within a single managed service. Delivered on Oracle Exadata infrastructure and Oracle Cloud, the platform is designed for mission-critical workloads requiring performance, governance, and security at scale.

The database supports structured, semistructured, and JSON data, as well as graph, spatial, and vector workloads, consolidating diverse requirements into a single platform. It also includes in-database machine learning capabilities, allowing organizations to run advanced models without moving data. With workload optimization, compression, and automated scaling features, Autonomous Database provides consistent performance for enterprises consolidating complex estates.

By embedding automation into core operations, Oracle positions Autonomous Database as a strategic choice for enterprises modernizing mission-critical environments. It is particularly relevant for regulated industries, global enterprises, and organizations with large analytical estates that require predictable performance, governance assurance, and automation-driven efficiency.

Strengths

- **Autonomous operations and performance management:** Oracle Autonomous Database automates traditionally manual tasks such as provisioning, tuning, patching, and workload optimization. By embedding automation into the core platform, it reduces reliance on manual DBA intervention while ensuring consistent performance. Autonomous features help organizations minimize downtime and administrative overhead, enabling IT teams to focus on higher-value activities. This automation provides enterprises with predictable performance and operational resilience, particularly in mission-critical environments where consistency and availability are essential.
- **Scale and performance on Exadata Infrastructure:** Autonomous Database runs on Oracle Exadata infrastructure, leveraging optimized hardware and software integration to deliver exceptional throughput and scalability. This combination enables petabyte-scale deployments, high concurrency, and low-latency analytics across structured, semistructured, and JSON data. Exadata's compression, caching, and workload management features further enhance efficiency. By providing enterprises with a proven platform for performance at scale, Oracle positions Autonomous Database as a strong choice for organizations requiring continuous availability and reliability for analytical workloads.
- **Broad in-database analytics capabilities:** Oracle Autonomous Database includes extensive in-database analytics, supporting machine learning, graph, spatial, JSON, and vector workloads natively. These capabilities allow enterprises to build advanced models and perform complex queries without moving data into external systems. By consolidating diverse analytical capabilities within the database, Oracle reduces data movement and governance complexity. This broad analytics portfolio helps enterprises accelerate innovation while ensuring compliance, making the platform suitable for organizations with diverse and demanding use cases.

Challenges

- **Migration complexity:** Moving databases and applications to Oracle Autonomous Database can require planning and validation to ensure compatibility, adjust dataflows, and redesign workflows. These activities may extend implementation timelines and increase initial effort. Oracle provides tools such as Zero Downtime Migration, free heterogeneous migration utilities, and SQL Developer Migration Workbench to simplify the process of moving both Oracle and non-Oracle workloads. The SQL Translation Framework further supports migrations by translating SQL statements from other database platforms into Oracle syntax.

- **Vendor lock-in considerations:** Dependence on a single cloud provider can increase the effort required to migrate or integrate workloads across different environments due to variations in APIs, networking, and identity services. Oracle offers multicloud deployment options, including Oracle Database@Hyperscaler on Microsoft Azure, Google Cloud, AWS, and OCI, to help support flexibility and integration across platforms.
- **Operational adaptation:** The move to an autonomous, AI-driven database can require adjustments to processes and roles as routine administrative tasks become automated. Organizations often reallocate staff time toward higher-value activities such as data modeling, analytics, and innovation following deployment.

Consider Oracle When

Consider Oracle Autonomous Database when the priority is to modernize mission-critical analytics with automation, Exadata-class performance, and enterprise-grade governance. The platform is particularly relevant for organizations that require predictable performance, continuous availability, and embedded analytics such as machine learning, graph, and spatial directly in the database. Enterprises seeking to reduce administrative burden, improve compliance, and ensure operational resilience at scale will find Oracle Autonomous Database a practical option for both hybrid and cloud-native deployments.

SAP

After a thorough evaluation of SAP's strategies and capabilities, IDC has positioned SAP in the Leaders category in this 2025 IDC MarketScape for worldwide analytical databases.

SAP is a global enterprise technology provider with a long-standing focus on business applications, data management, and analytics. Its flagship cloud database platform, SAP HANA Cloud, is a multicloud, in-memory, columnar database designed to deliver high-speed performance and real-time insight. The platform supports hybrid transactional and analytical processing, enabling organizations to combine operational and analytical workloads within a single environment.

HANA Cloud integrates advanced capabilities including predictive analytics, spatial, graph, text, and time-series processing, reducing reliance on external systems for specialized workloads. Its elastic, cloud-native design provides scalability for complex analytical use cases while simplifying governance and data management. As part of SAP's broader Business Technology Platform (BTP), HANA Cloud connects analytics, AI, and application data to create a unified foundation for real-time decision-making and enterprise modernization.

HANA Cloud is positioned for enterprises to modernize and optimize on-premises database environments. It is particularly relevant for organizations that require real-time analytics, convergence of transactional and analytical data, and advanced functionality to support strategic decision-making at scale.

Strengths

- **In-memory performance for real-time analytics:** SAP HANA Cloud uses an in-memory, columnar architecture that enables high-speed query performance and real-time analytics. This design reduces latency and accelerates insight for time-sensitive workloads, making it well suited for operational reporting and advanced analytics. The platform allows enterprises to run complex queries and aggregations and joins at scale without significant degradation in responsiveness. By optimizing workloads directly in memory, HANA Cloud provides predictable performance and responsiveness for mission-critical use cases such as finance, supply chain, and customer engagement.
- **Converged analytical and transactional processing (HTAP):** HANA Cloud combines analytical and transactional processing within a single platform, allowing organizations to run hybrid transactional and analytical processing on live data. This eliminates the need for separate systems, data movement, or duplication, reducing complexity and improving governance. The ability to run OLTP and OLAP workloads together enables enterprises to deliver real-time operational insights and accelerate decision-making. By converging these workloads, HANA Cloud supports modernization initiatives where agility, speed, and consolidation are business priorities.
- **Advanced analytical functions:** SAP HANA Cloud provides integrated support for advanced analytics including predictive modeling, spatial, property graph, text, and time-series processing as well as native vector and knowledge graph capabilities supporting GenAI use cases. These capabilities allow organizations to expand use cases beyond traditional BI and reporting, enabling advanced applications such as geospatial optimization, network analysis, and forecasting. By embedding these functions natively in the database, enterprises reduce reliance on external systems and simplify data governance. This breadth of advanced analytics helps organizations innovate faster while maintaining control and compliance across their data estate.

Challenges

- **Perceived cost efficiency at scale:** While HANA Cloud's in-memory architecture delivers strong performance, it can be more expensive than disk-based or cloud-native providers at very large scale. Memory-resident data storage increases infrastructure costs as data sets grow into the multi-terabyte or petabyte range.

Although compression and optimization features reduce this impact, organizations with unpredictable or rapidly expanding workloads may face challenges in maintaining predictable economics. HANA Cloud's elastic compute capabilities as well as its native multitier data processing to seamlessly leverage disk, petabyte-scale relational data lake, and object store persistency options, can significantly improve TCO and flexibility. Enterprises should assess long-term TCO carefully and apply governance frameworks to align performance requirements with budget constraints.

- **Adoption outside the SAP Ecosystem:** HANA Cloud is strongly aligned with SAP's application ecosystem, making it an obvious choice for existing SAP customers. However, adoption outside the SAP base is more limited, as some organizations perceive the platform as SAP centric. This can reduce visibility in competitive evaluations against providers with broader cross-industry positioning. Enterprises without significant SAP investments may need to weigh the platform's technical advantages against the need for wider ecosystem interoperability.
- **Competitive pressure from cloud-native warehouses:** SAP HANA Cloud competes against cloud-native warehouses and lakehouse platforms that emphasize elasticity, openness, and broad adoption. While HANA Cloud offers advanced features and integration within the SAP stack, some buyers may view cloud-native alternatives as more flexible for heterogeneous environments. This dynamic can affect perception and adoption, particularly in enterprises standardizing on multicloud data strategies.

Consider SAP When

Consider SAP HANA Cloud when consolidating transactional and analytical processing with real-time, in-memory performance is a priority. The platform is particularly relevant for organizations with significant SAP investments that want to unify analytics and applications on a single foundation. Enterprises that require advanced capabilities such as predictive, spatial, or graph analysis alongside HTAP will find SAP HANA Cloud a practical choice for regulated industries and SAP-centric modernization strategies.

SingleStore

After a thorough evaluation of SingleStore's strategies and capabilities, IDC has positioned SingleStore in the Major Players category in this 2025 IDC MarketScape for worldwide analytical databases.

SingleStore is a global enterprise database provider focused on delivering high-performance analytics through a distributed, cloud-native platform. Its flagship offering, SingleStoreDB, combines in-memory rowstore with disk-based columnstore to support hybrid transactional and analytical processing. This design enables enterprises

to run real-time analytics and operational workloads within a single relational environment.

The platform supports horizontal scale-out across distributed clusters, providing predictable performance for high-concurrency, high-ingest scenarios. By combining relational SQL compatibility with modern distributed architecture, SingleStoreDB allows organizations to manage streaming data, transactional workloads, and analytical queries without moving data between systems. Integration with prominent BI, AI, and data-engineering tools through marketplace partnerships continues to broaden SingleStore's enterprise reach and ecosystem alignment.

SingleStore offers deployment flexibility through its fully managed cloud service, SingleStore Helios, as well as self-managed options for on-premises or private cloud environments. This flexibility allows enterprises to balance modernization with governance and compliance needs. The platform is positioned for organizations seeking workload consolidation, real-time decision-making, and simplified architectures across hybrid and multicloud strategies.

Strengths

- **Distributed performance for real-time analytics:** SingleStoreDB delivers high-performance analytics through a distributed relational engine that combines in-memory rowstore with disk-based columnstore. This architecture supports sub-second response times for complex analytical queries while handling high ingest rates of streaming and transactional data. By scaling horizontally across distributed clusters, SingleStore maintains predictable performance under growing concurrency and workload diversity. These capabilities make it particularly effective for operational decision-making scenarios where speed and responsiveness are business critical, such as fraud detection, IoT analytics, and personalized customer engagement.
- **Converged transactional and analytical processing (HTAP):** SingleStore supports hybrid transactional and analytical processing, allowing organizations to run OLTP and OLAP workloads directly within one unified platform. This eliminates the need for ETL processes or separate systems, reducing latency and simplifying data governance. By processing live operational data alongside analytical queries, the platform enables real-time operational intelligence and faster insights. This HTAP capability helps enterprises streamline architectures, reduce duplication, and improve agility, making it a strong fit for dynamic use cases that require continuous analysis of transactional data.
- **Deployment flexibility across managed and self-managed models:** SingleStore provides deployment flexibility through its fully managed cloud service, SingleStore Helios, and self-managed options for on-premises or private

cloud environments. This choice allows organizations to align database operations with governance, compliance, and cost requirements. Containerized deployments with Kubernetes support enable portability and ease of scaling across hybrid or multicloud environments. This flexibility makes SingleStore relevant for enterprises that want to modernize incrementally while maintaining control of sensitive data and ensuring operational resilience across varied infrastructures.

Challenges

- **Cost efficiency at extreme scale:** Although SingleStore provides strong performance for real-time analytics, cost efficiency can be a challenge at multi-petabyte scale compared with hyperscale cloud-native warehouses. The platform's architecture, while effective for sub-second performance, may require additional infrastructure and governance to maintain predictable economics. Organizations with rapidly growing data estates should evaluate long-term total cost of ownership and capacity management strategies. Workload governance, monitoring, and clear planning help enterprises maximize SingleStore's performance benefits while maintaining cost control at scale.
- **Ecosystem maturity and integration breadth:** Compared with hyperscale vendors, SingleStore's partner ecosystem and breadth of integrations remain less extensive. While it offers strong compatibility with SQL and popular programming languages, enterprises require additional effort to connect SingleStore with third-party BI, AI, and data management tools. The company continues to expand its marketplace partnerships and integrations, but organizations should evaluate how its evolving ecosystem aligns with broader data strategies and interoperability requirements.
- **Competitive positioning against hyperscale vendors:** SingleStore competes in a market dominated by hyperscale cloud providers and independent vendors. While SingleStore differentiates through its HTAP capabilities and performance, its overall market visibility and adoption levels are lower. This can create challenges in competitive evaluations where larger vendors carry stronger brand recognition and enterprise presence. Organizations should weigh SingleStore's unique strengths against the broader reach and perceived stability of hyperscale competitors when making platform decisions.

Consider SingleStore When

Consider SingleStore when real-time analytics, operational decision-making, or HTAP capabilities are key requirements. The platform is particularly relevant for organizations that need to ingest streaming data, run high-volume transactions, and deliver sub-second analytical responses in a single environment. Enterprises seeking deployment flexibility between managed cloud services and self-managed models, where workload

consolidation and real-time analytics are strategic priorities, will find SingleStore a practical option for balancing modernization, governance, and performance.

Snowflake

After a thorough evaluation of Snowflake's strategies and capabilities, IDC has positioned Snowflake in the Leaders category in this 2025 IDC MarketScape for worldwide analytical databases.

Snowflake is a global enterprise data platform provider that delivers the Snowflake AI Data Cloud, a cloud-native analytical database and ecosystem designed to support elasticity, concurrency, and secure collaboration across multiple public clouds. Its architecture separates compute from storage, allowing independent scaling and efficient workload isolation. Snowflake provides a multitenant, distributed environment that enables high concurrency and flexible consumption-based economics.

The platform supports structured, semistructured, and increasingly unstructured data formats, including JSON, Avro, and Parquet, while integrating with AI and ML pipelines through Snowpark and partner ecosystems. Its expanding AI Data Cloud marketplace connects enterprises, data providers, and developers in a unified ecosystem, supporting collaboration and accelerating data-driven innovation. Multicloud compute allows organizations to sustain performance for thousands of concurrent users, while native data-sharing features simplify collaboration across teams, regions, and partners. This combination enables enterprises to consolidate data silos and accelerate adoption of advanced analytical workloads.

Deployed natively on AWS, Microsoft Azure, and Google Cloud, Snowflake offers enterprises multicloud flexibility with consistent performance and governance across environments. Its positioning as a fully cloud-native service makes it a strong option for organizations seeking scalability, secure data collaboration, and modernized analytics within public cloud strategies.

Strengths

- **Multicloud compute for high concurrency:** Snowflake's cloud-native architecture supports multicloud compute, allowing enterprises to handle thousands of concurrent users without performance degradation. This design ensures that BI dashboards, batch analytics, and ad hoc queries can run simultaneously without impacting throughput or response times. By dynamically provisioning additional clusters during demand spikes, Snowflake maintains predictable service levels even under heavy workloads. This makes the platform especially relevant for organizations with large user bases or diverse workloads that require consistent performance at scale.

- **Elastic multicloud deployment:** Snowflake is natively available across AWS, Microsoft Azure, and Google Cloud, giving enterprises the ability to run analytics consistently across multiple providers. Its architecture separates compute from storage, allowing independent scaling and seamless workload portability across regions and clouds. This multicloud capability provides resilience, flexibility, and vendor choice while simplifying governance across global deployments. For organizations adopting multicloud strategies or operating across geographies, Snowflake offers a proven way to unify analytics with consistent performance.
- **Support for diverse data and workloads:** Snowflake supports structured, semistructured, and increasingly unstructured data types, enabling enterprises to consolidate data silos into a single analytical environment. The platform also integrates with AI and ML pipelines, expanding its applicability to advanced workloads such as predictive modeling and real-time inference. By providing native support for JSON, Avro, Parquet, and more, Snowflake simplifies data engineering while broadening analytical use cases. This breadth of support helps enterprises modernize their data ecosystems and extend analytics to emerging workloads without relying on multiple platforms. As adoption of AI and ML accelerates across industries, Snowflake's integrated support for these workloads allows organizations to innovate faster and operationalize insights at enterprise scale.

Challenges

- **Cost predictability at scale:** Snowflake's consumption-based pricing aligns costs with actual usage, but forecasting spend at scale is challenging. Highly variable workloads, concurrency spikes, or experimental AI and ML projects lead to unpredictable billing. Snowflake continues to enhance cost visibility and governance through resource monitoring, usage dashboards, and workload optimization features, but effective FinOps practices remain essential to sustain predictability. For enterprises with large analytical estates or rapidly shifting demand, proactive planning and governance frameworks are critical to balancing flexibility with financial control.
- **Dependence on public cloud infrastructure:** Snowflake runs exclusively on AWS, Microsoft Azure, and Google Cloud, making it dependent on these providers for infrastructure. This model delivers resilience and reach but limits options for organizations requiring hybrid or on-premises deployments. Enterprises in regulated industries with strict data residency requirements may need to evaluate whether public cloud-only availability aligns with governance mandates. Organizations seeking hybrid strategies should assess integration with complementary platforms to address gaps in on-premises capabilities.

- **Egress and data movement costs:** Snowflake enables secure data sharing and multiregion replication, but cross-region and multicloud data movement introduces additional egress fees. Enterprises with globally distributed teams or workloads spanning multiple regions may face rising costs as replication grows. While the platform delivers flexibility in data access, careful planning is needed to balance convenience with budgetary impact. Organizations should evaluate replication patterns, optimize region usage, and align data-sharing strategies with financial governance policies to manage long-term costs.

Consider Snowflake When

Consider Snowflake AI Data Cloud when the priority is to deliver a cloud-native analytical database with elasticity, concurrency, and simplified operations. The platform is particularly relevant for organizations requiring support for thousands of concurrent users, structured and semistructured data at scale, and deployments across multiple public clouds. Enterprises seeking secure data sharing, cost governance, and integration with advanced workloads such as AI and ML will find Snowflake a practical choice for balancing flexibility, scalability, and modernization.

Teradata

After a thorough evaluation of Teradata's strategies and capabilities, IDC has positioned Teradata in the Major Players category in this 2025 IDC MarketScape for worldwide analytical databases.

Teradata is a global enterprise technology provider with a focus on data management and large-scale analytics. Its flagship platform, Teradata Vantage, serves as the foundation of Teradata's Autonomous AI and Knowledge Platform, designed to unify data, analytics, and AI for intelligence and autonomous decisioning across the enterprise. Teradata VantageCloud is a cloud-native analytical platform designed for mission-critical workloads that require high concurrency, scalability, and strict governance.

VantageCloud supports deployment across AWS, Azure, and Google Cloud, as well as on premises, providing flexibility for hybrid and multicloud strategies. With QueryGrid, organizations can federate queries across environments, enabling seamless access to distributed data without complex movement. Integration with ClearScape Analytics and Teradata AI Factory strengthens Teradata's role as a unified platform for data management, AI, and machine learning execution, helping enterprises build, orchestrate, and scale agentic and generative AI applications with governance and consistent performance across cloud and on-premises environments. The platform also supports advanced analytical functions including time-series, geospatial, and path analysis to extend its value across varied workloads.

As part of Teradata's evolution from traditional data warehousing to an Autonomous AI and Knowledge Platform, the company continues to expand its role from managing data to operationalizing AI-driven intelligence and decisioning at enterprise scale. It is particularly relevant for organizations in industries such as financial services, telecommunications, and retail that require reliable workload optimization, integrated analytics, and predictable performance across hybrid and multicloud environments.

Strengths

- **Advanced workload optimization:** Teradata Vantage provides sophisticated workload management capabilities, enabling enterprises to optimize performance across thousands of concurrent queries. Features such as workload isolation, priority-based scheduling, and intelligent resource management allow organizations to maintain consistent service levels in complex environments. These capabilities are particularly valuable in mission-critical industries where predictable performance under heavy concurrency is essential. By reducing contention and ensuring fairness across workloads, Vantage helps enterprises balance diverse user demands with operational efficiency.
- **Scalability across hybrid and multicloud environments:** VantageCloud supports deployment across AWS, Azure, and Google Cloud, as well as on premises, providing flexibility for hybrid and multicloud strategies. With QueryGrid, organizations can federate queries across environments, enabling seamless access to distributed data without complex data movement. This capability supports modernization initiatives by allowing enterprises to integrate cloud resources with existing investments. The ability to scale elastically across multiple clouds ensures consistency, governance, and resilience for large-scale analytical workloads.
- **Broad and integrated analytics capabilities:** Teradata integrates advanced analytical functions directly within Vantage through ClearScape Analytics, supporting machine learning, time-series, geospatial, and path analysis. The platform's compatibility with Python, R, and open source ecosystems enables enterprises to embed advanced analytics within operational workflows. By unifying data management with in-database analytics, Teradata reduces data movement and governance complexity. This breadth of integrated capabilities makes Vantage a strategic platform for enterprises seeking to combine traditional BI with advanced AI/ML workloads.

Challenges

- **Cost competitiveness outside enterprise scale:** Teradata Vantage provides strong value for large-scale enterprise deployments, particularly where performance, concurrency, and governance are paramount. For smaller or

midmarket workloads, some of VantageCloud's advanced capabilities may exceed immediate requirements. Enterprises considering Vantage should evaluate total cost of ownership relative to alternatives, particularly if workloads do not require high concurrency or advanced optimization. For organizations operating at very large scale, however, the cost profile aligns well with the performance and governance benefits provided. This makes Vantage ideal for organizations operating at global enterprise scale, where workload optimization and concurrency management are priorities.

- **Evolving market perception:** As Teradata continues to reposition itself as an Autonomous AI and Knowledge Platform, some organizations are still updating their perceptions of the company's heritage in data warehousing. Enterprises should evaluate Teradata based on the current maturity and innovation of its platform rather than historical positioning.
- **Ease of use:** VantageCloud's powerful workload optimization and scalability features come with a level of product sophistication that may exceed the requirements of some enterprises. Teradata often requires greater expertise in tuning and workload management. This can increase administrative effort, particularly for organizations without mature data engineering practices. Enterprises adopting VantageCloud should plan for enablement and governance frameworks to fully leverage its advanced capabilities while maintaining efficient operations.

Consider Teradata When

Consider Teradata VantageCloud when the enterprises need to operationalize AI and analytics at scale, with high concurrency and strict governance. The platform is well suited for enterprises pursuing hybrid or multicloud strategies that demand seamless data federation and consistent performance across environments. Organizations prioritizing advanced workload optimization, integrated AI and ML capabilities, and enterprise-grade reliability will find Teradata VantageCloud a practical option for bridging traditional analytics with the next generation of trusted, autonomous intelligence.

Vendors to Watch

The vendors mentioned in the sections that follow are included as vendors to watch in this year's assessment based on recent developments that expand their capabilities and strengthen alignment with IDC's analytical database inclusion criteria. Each continues to advance its offerings and may be considered for full evaluation in future editions.

Databricks

Databricks provides an analytical database platform built on an open lakehouse architecture that integrates data warehousing and data lake capabilities within a single environment. The platform supports high-performance SQL analytics and data engineering workloads at scale through its optimized query engine and Delta Lake storage layer, which delivers ACID transactions, schema enforcement, and time travel. Databricks SQL enables interactive and batch query processing with high concurrency and low latency. Unified governance through Unity Catalog and support for Vector Search extend analytical workloads into AI via AI Functions and retrieval augmented generation. Databricks' open, cloud-native architecture positions it as a vendor to watch in the analytical database market.

MariaDB

MariaDB offers a unified cloud database platform that brings together transactional, analytical, and AI (vector) database engines into a single, high-performance platform. The MariaDB Enterprise Platform delivers high-performance analytics with MariaDB Exa, a massively parallel processing and in-memory engine designed for multi-terabyte, complex analytical workloads on operational data, delivering high-throughput query performance and near-real-time insight generation. MariaDB Exa is the result of a strategic partnership between MariaDB and Exasol, allowing organizations to extract real-time insights from large operational data sets and gain value from transactional information without requiring data movement to a separate analytics system. With managed service delivery through MariaDB Cloud and flexible on-premises deployment, MariaDB provides an open and scalable architecture that positions it as a vendor to watch in the analytical database market.

Progress Software

Progress MarkLogic is an analytical data management platform that integrates document, graph, and relational capabilities with a high-performance search and query engine. It enables complex analytical workloads by unifying structured and unstructured data within a single architecture. Delivered through Progress Data Cloud, the platform provides horizontal scalability, high availability, and managed operations across major public clouds. MarkLogic supports semantic modeling, vector search, and retrieval augmented generation for AI-driven analytics. Its ability to combine multimodal data processing, semantic enrichment, and advanced search positions Progress as a vendor to watch in the analytical database market.

Reading an IDC MarketScape Graph

For the purposes of this analysis, IDC divided potential key measures for success into two primary categories: capabilities and strategies.

Positioning on the y-axis reflects the vendor's current capabilities and menu of services and how well aligned the vendor is to customer needs. The capabilities category focuses on the capabilities of the company and product today, here and now. Under this category, IDC analysts will look at how well a vendor is building/delivering capabilities that enable it to execute its chosen strategy in the market.

Positioning on the x-axis, or strategies axis, indicates how well the vendor's future strategy aligns with what customers will require in three to five years. The strategies category focuses on high-level decisions and underlying assumptions about offerings, customer segments, and business and go-to-market plans for the next three to five years.

The size of the individual vendor markers in the IDC MarketScape represents the market share of each individual vendor within the specific market segment being assessed. For this IDC MarketScape, IDC determined the vendor market size using IDC's Software Tracker, 2024, which represents an estimate of each vendor's 2024 software revenue.

IDC MarketScape Methodology

IDC MarketScape criteria selection, weightings, and vendor scores represent well-researched IDC judgment about the market and specific vendors. IDC analysts tailor the range of standard characteristics by which vendors are measured through structured discussions, surveys, and interviews with market leaders, participants, and end users. Market weightings are based on user interviews, buyer surveys, and the input of IDC experts in each market. IDC analysts base individual vendor scores, and ultimately vendor positions on the IDC MarketScape, on detailed surveys and interviews with the vendors, publicly available information, and end-user experiences in an effort to provide an accurate and consistent assessment of each vendor's characteristics, behavior and capability.

Market Definition

An analytical database is a specialized data management system designed to store, process, and query large volumes of data for analytical workloads rather than transactional operations. Modern analytical databases support a wide range of use cases, from historical reporting and business intelligence (BI) to real-time analytics,

machine learning (ML), and AI-assisted decisioning. They are optimized for high-performance querying and scalable data processing that enable enterprises to derive actionable insights from both structured and semistructured data.

Analytical databases now span on-premises, cloud native, and hybrid architectures, including lakehouse and distributed query models that enable analytics wherever data resides. The separation of storage and compute is a core design principle, enabling elastic scaling, workload isolation, and cost-efficient performance across hybrid and multicloud environments. Many platforms now incorporate automation, vectorized execution, and AI-driven optimization to improve efficiency, reduce administrative overhead, and accelerate time to insight. Together, these innovations define the next generation of analytical databases powering enterprise intelligence.

Strategies and Capabilities Criteria

IDC divides the assessment criteria into two primary categories of strategies and capabilities, as shown in Tables 1 and 2, for the success of the analytical database market assessment. IDC analysts look at how well a vendor is building and delivering capabilities that enable it to execute its chosen strategy in the market.

In the strategies category, IDC assesses whether a vendor's plans align with customer requirements and spending priorities over a defined future time frame, including business and go-to-market objectives. The most influential criteria within the strategies category are financial and funding strength, functionality and offering strategy, and overall growth momentum. In the capabilities category, the most influential criteria within this category are data storage capabilities and the breadth of services offered.

TABLE 1**Key Strategy Measures for Success: Worldwide Analytical Databases**

Strategies Criteria	Definition	Weight (%)
Financial/funding	<ul style="list-style-type: none"> ▪ Age of company, number of employees, and revenue to employee ratio ▪ Range of regions that the company operates in ▪ Full company revenue in 2023 and 2024 	21.0
Partner ecosystem	<ul style="list-style-type: none"> ▪ Number of channel, technology, and implementation partners ▪ Range of developer community initiatives and programming languages supported ▪ Range of deployment options supported ▪ Range of public cloud availability 	17.0
Innovation	<ul style="list-style-type: none"> ▪ Approach to research and development ▪ Approach to infusing AI into the database 	12.0
Growth	<ul style="list-style-type: none"> ▪ Range of new partners planned for 2025, planned improvements for partner strategy ▪ Number of additional human languages and public cloud regions supported in 2025 ▪ Range and depth of significant milestones in the past year 	16.0
Functionality or offering strategy	<ul style="list-style-type: none"> ▪ Road map update frequency, availability to customers, and execution of delivery ▪ Range of security and privacy certifications ▪ Range of user experience and API enhancements planned through 2026 ▪ Range of architectural and cloud-specific enhancements planned through 2026 ▪ Range of industry-specific enhancements planned through 2026 ▪ Range of general database enhancements, frequency of major product releases 	19.0
Customer engagement	<ul style="list-style-type: none"> ▪ Approach to customer onboarding and engagement; range of support options for all customers ▪ Range of customer training options, community initiatives, and road map plans in the next 12 months ▪ Range of community initiatives 	9.0

TABLE 1**Key Strategy Measures for Success: Worldwide Analytical Databases**

Strategies Criteria	Definition	Weight (%)
Customer delivery strategy	<ul style="list-style-type: none"> ▪ How would you rate the ability of the vendor to solve your customer service issues and keep you happy as a customer? ▪ Is the product road map relevant to you? Does the vendor do a good job informing you what is coming? ▪ Is the vendor providing you with the tools to succeed and adapt in the foreseeable future with a new innovation, advanced modules, or ease of customization? ▪ How would you rate the vendor's AI strategy? 	6.0
Total		100.0

Source: IDC, 2025

TABLE 2**Key Capability Measures for Success: Worldwide Analytical Databases**

Capabilities Criteria	Definition	Weight (%)
Product information	<ul style="list-style-type: none"> ▪ Time that product has been generally available ▪ Number of customers and product revenue for 2023 and 2024 ▪ 2024 net retention rate 	8.0
Data storage	<ul style="list-style-type: none"> ▪ Range of storage architectures natively offered; range of automated storage tiering offered ▪ Range of open table formats supported ▪ Ability to optimize storage costs while maintaining high performance ▪ Range of scaling capabilities, including automatic, vertical, and horizontal scaling ▪ Range of disaster recovery and monitoring capabilities 	24.0
Data organization	<ul style="list-style-type: none"> ▪ Range of data formats supported for processing ▪ Range of integration approaches supported ▪ Ability to monitor, ensure, validate, and schedule complex data integration pipelines ▪ Ability to do real-time data processing 	14.0
Security and compliance	<ul style="list-style-type: none"> ▪ Range of security certifications and security controls supported ▪ Range of data protection capabilities ▪ Audit logging capabilities 	12.0

TABLE 2**Key Capability Measures for Success: Worldwide Analytical Databases**

Capabilities Criteria	Definition	Weight (%)
Functionality or offering	<ul style="list-style-type: none"> Maximum number of concurrent queries, storage model impact on performance, and query optimization features 	5.0
Range of services	<ul style="list-style-type: none"> Range of advanced analytics supported natively Range of AI/ML frameworks supported, ability for model versioning, monitoring, training, and governance Ability to support all user personas, cater to nontechnical users, and track user engagement with data Range of agentic AI capabilities and AI features in analytics and collaboration 	16.0
Pricing model or structure of product/offering	<ul style="list-style-type: none"> Range of licensing options and subscription terms Range of base units and capacity options Range of cost optimization features supported Flexibility for exceeding paid capacity 	13.0
Customer satisfaction	<ul style="list-style-type: none"> Relative to the price you paid for the software, did the value of the system meet your expectations? How easy is it for users to understand and utilize the software? How well does the product perform and is the level of availability provided sufficient for users? Focusing on the software, not partners who implemented the software, was the software easily configurable and deployed in a timely manner? 	8.0
Total		100.0

Source: IDC, 2025

LEARN MORE**Related Research**

- *IDC FutureScape: Worldwide Data and Analytics 2026 Predictions* (IDC #US53860225, October 2025)
- *IDC MarketScape: Worldwide Data Platform Software 2025 Vendor Assessment* (IDC #US52974125, July 2025)
- *IDC Quick Poll, 2Q25: Agentic AI and Data* (IDC #US53638725, June 2025)

- *Worldwide Database Management Systems Software Market Shares, 2024: Powering AI Workloads and Enabling Autonomous DBMS Operations* (IDC #US53032425, June 2025)
- *Worldwide Database Management Systems Software Forecast, 2025–2029* (IDC #US53032525, June 2025)

Synopsis

The IDC study evaluates worldwide analytical database software platforms for 2025, highlighting how advances in storage and compute architectures are transforming enterprise analytics. Vendors evaluated include AWS, Broadcom, Cloudera, Couchbase, Google, IBM, InterSystems, Microsoft, OpenText, Oracle, SAP, SingleStore, Snowflake, and Teradata. The market is shifting beyond traditional columnar systems toward cloud-native, lakehouse, and distributed query architectures that separate storage and compute to improve elasticity, performance, and cost control.

As organizations modernize data infrastructure and accelerate AI adoption, analytical databases are emerging as the intelligent, high-performance foundation for real-time analytics and data-driven decisioning. Vendors are embedding AI-driven optimization, vector processing, and automation to enhance scalability, accelerate insights, and prepare for agentic AI workloads. The market's growth signals a new phase where analytics, AI, and storage innovation converge to enable faster and more intelligent enterprise decisions.

"Analytics is the engine of enterprise intelligence, and analytical databases are evolving to make it faster, smarter, and more adaptive," said Devin Pratt, research director, Data Management Software at IDC. "As organizations advance their analytics strategies and incorporate AI and agentic AI, leadership will belong to databases that seamlessly unite storage, compute, and automation to turn data into trusted, real-time decisions."

ABOUT IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology, IT benchmarking and sourcing, and industry opportunities and trends in over 110 countries. IDC's analysis and insight helps IT professionals, business executives, and the investment community to make fact-based technology decisions and to achieve their key business objectives. Founded in 1964, IDC is a wholly owned subsidiary of International Data Group (IDG, Inc.).

Global Headquarters

140 Kendrick Street
Building B
Needham, MA 02494
USA
508.872.8200
Twitter: @IDC
blogs.idc.com
www.idc.com

Copyright and Trademark Notice

This IDC research document was published as part of an IDC continuous intelligence service, providing written research, analyst interactions, and web conference and conference event proceedings. Visit www.idc.com to learn more about IDC subscription and consulting services. To view a list of IDC offices worldwide, visit www.idc.com/about/worldwideoffices. Please contact IDC at customerservice@idc.com for information on additional copies, web rights, or applying the price of this document toward the purchase of an IDC service.

Copyright 2025 IDC. Reproduction is forbidden unless authorized. All rights reserved.