

Snowflake vs. Databricks

Unvarnished Reviews · Independent Report · unvarnishedreviews.com · Free · No subscription

Unvarnished Reviews Research

This report synthesizes data from 1,400+ verified user reviews and practitioner community posts collected from G2 (Databricks 707 reviews at 4.6/5, Snowflake 692 reviews at 4.6/5), Gartner Peer Insights, TrustRadius, Capterra, Reddit r/dataengineering and r/snowflake, Stack Overflow data engineering communities, and independent pricing analyses. Pricing data reflects vendor pricing pages, CostBench independent pricing analysis, CheckThat.ai verified pricing data, and independent TCO modeling current as of June 2026. Full research methodology at unvarnishedreviews.com/methodology. Research Notes available on request at editorial@unvarnishedreviews.com.

The Verdict Up Front

Snowflake is the dominant enterprise cloud data warehouse, 18.33% market share, approximately \$50 billion market cap as of June 2026, and the platform that defined the separation of compute and storage as a commercial data architecture. Its virtual warehouse model, T-shirt sized compute clusters (XS to 6XL) billed per second, delivers real elasticity for structured data workloads. It is also the platform where a 6XL virtual warehouse burns 512 credits per hour, 512 times the cost of an XS warehouse, where inactive warehouses silently consume credits if auto-suspend is not configured, and where a single Medium warehouse running 8 hours per day for 20 business days costs \$1,920/month in compute alone, before storage or serverless add-ons. Effective Snowflake cost optimization typically delivers 40%-60% savings, a finding that implies most organizations are overpaying by 40%-60% before they engage a FinOps practice.

Databricks is the AI/ML-first data platform, 8.67% market share growing at 57% year-over-year, a \$62 billion private valuation (2023), a Gartner Magic Quadrant Leader for Data Science and Machine Learning, and the platform that unified data engineering, ML, and generative AI on a single lakehouse architecture. Its May 2025 acquisition of Neon and subsequent launch of Lakebase, a serverless PostgreSQL offering, marks a historic architectural expansion: for the first time, Databricks can serve transactional OLTP workloads natively alongside analytics and ML on the same platform. This positions Databricks as a full-stack data platform, directly challenging the traditional separation between operational databases and analytical platforms. Its bill shock mechanism is different from Snowflake but equally real: Databricks charges DBUs plus a separate cloud provider bill for underlying VMs and storage, two meters running simultaneously, with All-Purpose Compute at \$0.40-\$0.55/DBU often more expensive than running the same workload as an automated job.

The straight 2026 framing: These platforms are converging. Snowflake has added Python execution (Snowpark), ML model inference (Cortex AI), streaming ingestion, and native Iceberg support. Databricks has added SQL Pro warehouses optimized for BI, Unity Catalog for enterprise governance, serverless compute, and Lakebase for transactional workloads. The architectural gap is narrowing, and the most sophisticated enterprise data teams are using both platforms simultaneously

in a "Factory and Showroom" pattern, not choosing one.

Recommendations: For SQL-first analytics teams, business intelligence workloads, and organizations that prioritize ease of use over ML depth: Snowflake. For data engineering and ML/AI-first organizations building agentic systems, LLM pipelines, and complex training workloads: Databricks. For large enterprises with both requirements: model the Factory/Showroom hybrid architecture explicitly before assuming one platform can do everything.

The Agentic AI Context: Why This Comparison Matters More in 2026 Than It Did in 2025

The Silicon Angle headline that prompted this report, "Snowflake, Databricks and the model makers: The battle for the agentic client and AI back end", captures the defining competitive dynamic of 2026.

Enterprise AI is moving from model inference (asking a single model a question) to agentic workflows (AI systems that reason, retrieve data, take actions, and orchestrate multi-step tasks across enterprise systems). Both Snowflake and Databricks are repositioning from "data platform" to "AI infrastructure platform", and the battle for where enterprise AI agents live and run is the strategic context for every data platform evaluation.

Snowflake's agentic play: Cortex AI adds LLM-powered functions directly into SQL. Snowflake Horizon Catalog (GA February 2026) enables cross-platform data governance. The pitch: agents query governed enterprise data through familiar SQL interfaces without leaving the Snowflake environment.

Databricks' agentic play: Mosaic AI includes tools for building advanced agent-based systems that reason, retrieve data, and take actions across multiple steps. Unity AI Gateway governs AI agents through the same Unity Catalog framework that governs data and ML models. Lakebase positions Databricks as the platform agents are built and governed on, not just queried from. The IDC MarketScape awarded Databricks the highest Strategies score of all vendors for unified AI governance in December 2025.

The architectural difference that matters for agentic AI: Databricks was built for data science, ML, and generative AI from inception. Snowflake assembled AI capabilities over time, largely through acquisition. The difference between native architecture and assembled capability becomes most visible under agentic workloads where latency, model governance, and retrieval-augmented generation performance are operational requirements, not benchmarks.

Platform Ratings at a Glance

Platform	G2	Market Share	Growth	Private/Public
Snowflake	4.6 / 5 (692 reviews)	18.33%	Steady	Public (SNOW)
Databricks	4.6 / 5 (707 reviews)	8.67%	57% YoY	Private (\$62B valuation)

Both platforms rate identically at 4.6/5 on G2, the distinction is not in aggregate satisfaction but in use case fit, cost management complexity, and AI/ML depth.

Architecture: The Core Differences

Snowflake: Compute/Storage Separation, SQL-First

Snowflake's architectural breakthrough was separating compute and storage, enabling organizations to scale each independently, pay only for what they use, and run multiple compute clusters against the same data simultaneously. Virtual warehouses are the compute layer: XS to 6XL T-shirt sized clusters that start, stop, and scale on demand.

The storage layer is Snowflake's proprietary columnar format, fast and highly optimized for SQL analytics, but historically creating lock-in because native format tables could only be queried by Snowflake engines. Apache Iceberg support reaching GA in 2025 addresses this, though the openness is narrower than Databricks' Delta Lake model where data lives in open files on your own object storage.

Databricks: Lakehouse, Delta Lake, ML-Native

Databricks built the Lakehouse architecture, combining the low-cost storage of a data lake with the reliability and performance of a data warehouse on a single platform. Delta Lake, Apache Spark, and the Photon native vectorized execution engine form the core. Data lives in open Delta Lake files on cloud object storage, fully portable, queryable by any engine, not locked to Databricks' storage layer.

The ML and AI layer is native, not assembled: MLflow for experiment tracking, Feature Store for ML features, Mosaic AI for model serving and agent building, Unity Catalog for governance across data, ML, and AI. Lakebase (2026) adds serverless PostgreSQL for transactional workloads, completing the full-stack architecture.

The Bill Shock Mechanisms: Both Platforms, Different Traps

Snowflake: The Credit Doubling Problem

Snowflake virtual warehouses consume credits at exponential rates as size increases. The math is documented and important:

Warehouse Size	Credits/Hour	Monthly Cost (8hr/day, 20 days, \$3/credit)
XS	1	\$480
S	2	\$960
M	4	\$1,920
L	8	\$3,840

Warehouse Size	Credits/Hour	Monthly Cost (8hr/day, 20 days, \$3/credit)
XL	16	\$7,680
2XL	32	\$15,360
6XL	512	\$245,760

The silent credit burn: Virtual warehouses bill per second with a 60-second minimum per start or resize. Warehouses that are not configured with appropriate auto-suspend settings continue billing when idle. This is documented as one of the most common sources of unexpected Snowflake costs.

The optimization opportunity: Third-party case studies document \$21,000/month in savings from warehouse rightsizing alone. Effective optimization typically delivers 40%-60% cost reduction, implying most organizations are significantly overpaying before engaging FinOps practices. Compute represents 70%-85% of typical Snowflake bills.

The spot instance gap: Databricks supports up to 90% spot instance ratio on cluster nodes, dropping sustained ETL costs by 50%-65%. Snowflake does not offer equivalent spot pricing flexibility, a significant cost disadvantage for high-volume data engineering workloads.

Databricks: The Double Meter Problem

Databricks charges on two separate meters simultaneously:

1. **DBUs** (Databricks Units), proprietary processing units billed per second
2. **Cloud provider compute**, AWS, Azure, or GCP charges for the underlying VMs and storage

All-Purpose Compute (interactive notebooks) runs \$0.40-\$0.55/DBU, often more expensive than running the same workload as an automated job. Organizations that run production workflows in interactive compute clusters rather than automated jobs are routinely overpaying.

The DBU rate varies by compute type, pricing tier, and cloud provider, making pre-deployment cost modeling more complex than Snowflake's credit-based model.

The savings lever: Running non-critical streaming pipelines on spot instances can significantly reduce costs. Automated jobs run at lower DBU rates than interactive compute. Organizations that profile their workloads and migrate appropriate jobs to automated runs consistently find material cost reductions.

What Practitioners Report

Snowflake: What Works

G2 reviewers consistently praise three areas: SQL analytics performance, ease of use for data analysts, and the Data Marketplace for sharing and acquiring datasets.

The separation of compute and storage, enabling queries to run against current data without performance degradation from concurrent users, is specifically cited as the platform's most distinctive operational advantage for BI and analytics workloads. Business analysts and data analysts

with SQL skills can be productive immediately without deep data engineering knowledge.

The Data Marketplace, where organizations can share, discover, and acquire external datasets without data movement, is a clearly differentiated capability that Databricks does not replicate. For organizations that consume external data (financial data, weather, demographics, market data), this is a meaningful operational differentiator.

One G2 reviewer: "Snowflake simplifies advanced data warehousing, making it a go-to choice for those looking for a fast and scalable solution without the burden of heavy infrastructure."

Snowflake: What Doesn't Work

Cost management complexity is the dominant practitioner complaint. The credit system, described as "confusing initially" in independent reviews, combined with the exponential cost scaling of virtual warehouse sizes creates a FinOps requirement that organizations frequently underestimate. One Capterra reviewer: "Costly if data is too big, you can't own the data, you have to depend on the Data Engineering team."

Proprietary storage lock-in (partially addressed by Iceberg support but not fully resolved) is a documented concern for organizations that want data portability across query engines.

ML and AI capabilities assembled through acquisition rather than built natively, Snowpark, Cortex AI, are less mature and less integrated than Databricks' native ML architecture. For organizations building production ML systems, Snowflake's AI capabilities require more workarounds.

Databricks: What Works

G2 reviewers consistently praise three areas: scalability for large-scale ML and AI workloads, the unified platform from data ingestion to model serving, and autoscaling for cost management.

The unified data + ML + AI architecture, where data engineering, model training, experiment tracking, feature serving, and model deployment share the same governance framework (Unity Catalog), eliminates the integration overhead that organizations face when assembling a multi-tool ML stack.

Mosaic AI's agent building capabilities, enabling multi-step reasoning, retrieval-augmented generation, and action execution across enterprise systems, are specifically praised as the platform's most distinctive 2026 capability. For organizations building agentic AI systems, Databricks' architectural alignment is a genuine operational advantage.

The 90% spot instance support for non-critical workloads is specifically cited by practitioners as the most impactful cost lever available.

Databricks: What Doesn't Work

The double billing complexity, DBUs plus cloud provider compute, is the most consistent practitioner complaint. Understanding the true cost of a Databricks workload requires modeling two meters simultaneously, and the interaction between them is not always intuitive.

Steeper learning curve for SQL-first teams. Data analysts accustomed to pure SQL workflows find Databricks' notebook-centric, Python-heavy interface less accessible than Snowflake. Databricks SQL addresses this but is newer and less feature-complete than Snowflake's core SQL experience.

Lakebase is very new (launched 2026). The serverless PostgreSQL offering that completes Databricks' full-stack positioning is early-stage. Organizations planning to use Lakebase for production OLTP workloads should evaluate it in pilot before committing at scale.

Pricing Reality (June 2026)

Snowflake

Component	Rate	Notes
Credits (on-demand)	\$2-\$4/credit	Cloud and tier dependent
Credits (annual commitment)	\$1.50-\$2.50/credit	25%-37% discount
Storage	~\$23/TB/month	Compressed
Standard Edition	\$2/credit	Basic features
Enterprise Edition	\$3/credit	Multi-cluster, materialized views
Business Critical	\$4/credit	HIPAA, PCI compliance

The real cost driver: Compute is 70%-85% of typical Snowflake bills. Warehouse sizing decisions and auto-suspend configuration are the most impactful cost levers. Benchmark: \$2,870-\$3,328/month for 10TB data volumes at modest utilization.

Databricks

Compute Type	DBU Rate	Notes
All-Purpose (interactive)	\$0.40-\$0.55/DBU	Most expensive; use for development only
Jobs Compute (automated)	\$0.15-\$0.22/DBU	Lowest rate; use for production
SQL Pro Warehouse	\$0.22/DBU	BI-optimized; competitive with Snowflake for SQL
Photon acceleration	+\$0.04-\$0.06/DBU	Performance multiplier

Plus cloud provider bill: AWS, Azure, or GCP charges for underlying VMs and storage separately.

Comparative TCO Benchmark (1TB analytical workload, moderate usage)

Platform	Estimated Monthly	Key Variable
Snowflake Standard	\$500-\$3,000+	Warehouse size and utilization
Databricks (jobs compute)	\$400-\$2,500+	DBU type and cloud VM costs
Snowflake vs. Databricks gap	Varies significantly	ML-heavy: Databricks cheaper. SQL-heavy: Snowflake simpler

At 1TB scale for pure analytics, costs are comparable. At enterprise scale with ML training, Databricks' spot instance support makes it materially cheaper for ML workloads. At enterprise scale

for high-concurrency SQL analytics, Snowflake's multi-cluster architecture often produces better price/performance.

The Factory and Showroom Pattern

The most important finding for organizations evaluating Snowflake vs. Databricks as a binary choice: the most sophisticated enterprise data teams don't choose one. They use both in a documented "Factory and Showroom" architecture:

Databricks as the Factory: Heavy data engineering, ML model training, feature engineering, and AI agent development. Open Delta Lake format, Spark-based processing, GPU clusters for model training.

Snowflake as the Showroom: Serving governed, curated data to business analysts, BI tools, and external partners through high-concurrency SQL warehouses. Data Marketplace for external data consumption.

The enabling technology: Open formats (Apache Iceberg, Delta Lake) and lakehouse federation (Databricks' Lakehouse Federation queries Snowflake directly without data copying) make the hybrid architecture increasingly practical. Organizations don't have to move data between platforms, they query it in place.

For organizations being pressured to choose one platform, the real question is: does your data infrastructure actually have one dominant workload profile, or are you trying to optimize two genuinely different workload types on one platform?

The Decision Framework

Choose Snowflake if:

- SQL analytics and business intelligence for non-technical business users is the primary workload
- Data sharing and marketplace access to external datasets is operationally important
- Your team is SQL-first and Python/ML expertise is limited
- Ease of administration and managed infrastructure are priorities
- You have implemented, or plan to implement, warehouse rightsizing, auto-suspend, and FinOps practices to manage credit consumption
- You have modeled your virtual warehouse sizes against actual query requirements, not theoretical maximums

Choose Databricks if:

- ML model training, feature engineering, and production ML serving are primary workloads
- You are building agentic AI systems that require unified data + model + governance infrastructure
- Your team has data engineering and Python expertise
- Open data format (Delta Lake, Iceberg) and avoiding storage lock-in are priorities
- High-volume ETL and data transformation workloads benefit from spot instance cost optimization

- Lakebase's serverless PostgreSQL capability for transactional workloads fits your architecture

Consider both (Factory and Showroom) if:

- Your organization has both ML/AI workloads and high-concurrency SQL analytics at scale
- Open formats make bi-directional data access practical without data movement
- The cost of running non-optimal workloads on a single platform exceeds the operational overhead of managing two

The pre-deployment cost checklist for both platforms:

Snowflake:

1. Define each workload's query patterns and concurrency requirements before sizing virtual warehouses
2. Configure auto-suspend aggressively, the default is often too generous
3. Start with XS or S warehouses and scale up based on actual performance data, never start at L or XL
4. Enable resource monitors with spending alerts before production workloads run
5. Audit existing warehouses for idle credit consumption before any cost optimization

Databricks:

1. Profile workloads by compute type, interactive, jobs, streaming, ML training
2. Migrate production workloads from All-Purpose to Jobs Compute immediately
3. Enable spot instances for non-critical streaming and ETL jobs
4. Monitor the cloud provider bill alongside DBU costs, both meters matter
5. Use Photon acceleration only for compute-intensive SQL workloads where the DBU premium pays off in speed

The Bottom Line

Snowflake and Databricks are both excellent platforms, rated identically at 4.6/5 on G2, serving different primary use cases that are converging toward the same agentic AI infrastructure battle.

Snowflake is the most appropriate choice for SQL-first analytics teams, high-concurrency BI workloads, and organizations that value managed simplicity over ML depth. Its credit-based billing model creates bill shock for organizations that don't implement FinOps discipline, the 40%-60% documented savings from optimization implies most organizations are significantly overpaying without it.

Databricks is the most appropriate choice for ML/AI-first data organizations, agentic AI infrastructure, and workloads where open data formats and spot instance cost optimization matter. Its double-meter billing (DBUs + cloud provider) creates cost modeling complexity that requires explicit management.

The most explosive finding in this comparison: A Snowflake 6XL virtual warehouse burns 512 credits per hour. At Enterprise Edition pricing (\$3/credit), that's \$1,536/hour, \$36,864 for a 24-hour period. Organizations that provision large warehouses for development or testing without auto-suspend, a documented common pattern, generate this cost with no business value. The optimization opportunity is proportional: \$21,000/month in documented monthly savings from warehouse rightsizing alone.

The question that determines platform selection before any technical evaluation: Is your primary workload SQL analytics for business users, or ML/AI training and agentic infrastructure for data engineers? The answer to that question narrows the field more definitively than any feature comparison.