



MARKET REPORT

UPDATED
10/28/2025

Databricks at \$4B ARR growing 50% YoY

TEAM

Jan-Erik Asplund
Co-Founder
jan@sacra.com

DISCLAIMERS

This report is for information purposes only and is not to be used or considered as an offer or the solicitation of an offer to sell or to buy or subscribe for securities or other financial instruments. Nothing in this report constitutes investment, legal, accounting or tax advice or a representation that any investment or strategy is suitable or appropriate to your individual circumstances or otherwise constitutes a personal trade recommendation to you.

This research report has been prepared solely by Sacra and should not be considered a product of any person or entity that makes such report available, if any.

Information and opinions presented in the sections of the report were obtained or derived from sources Sacra believes are reliable, but Sacra makes no representation as to their accuracy or completeness. Past performance should not be taken as an indication or guarantee of future performance, and no representation or warranty, express or implied, is made regarding future performance. Information, opinions and estimates contained in this report reflect a determination at its original date of publication by Sacra and are subject to change without notice.

Sacra accepts no liability for loss arising from the use of the material presented in this report, except that this exclusion of liability does not apply to the extent that liability arises under specific statutes or regulations applicable to Sacra. Sacra may have issued, and may in the future issue, other reports that are inconsistent with, and reach different conclusions from, the information presented in this report. Those reports reflect different assumptions, views and analytical methods of the analysts who prepared them and Sacra is under no obligation to ensure that such other reports are brought to the attention of any recipient of this report.

All rights reserved. All material presented in this report, unless specifically indicated otherwise is under copyright to Sacra. Sacra reserves any and all intellectual property rights in the report. All trademarks, service marks and logos used in this report are trademarks or service marks or registered trademarks or service marks of Sacra. Any modification, copying, displaying, distributing, transmitting, publishing, licensing, creating derivative works from, or selling any report is strictly prohibited. None of the material, nor its content, nor any copy of it, may be altered in any way, transmitted to, copied or distributed to any other party, without the prior express written permission of Sacra. Any unauthorized duplication, redistribution or disclosure of this report will result in prosecution.

Published on Oct 28th, 2025

Databricks at \$4B ARR growing 50% YoY

By Jan-Erik Asplund

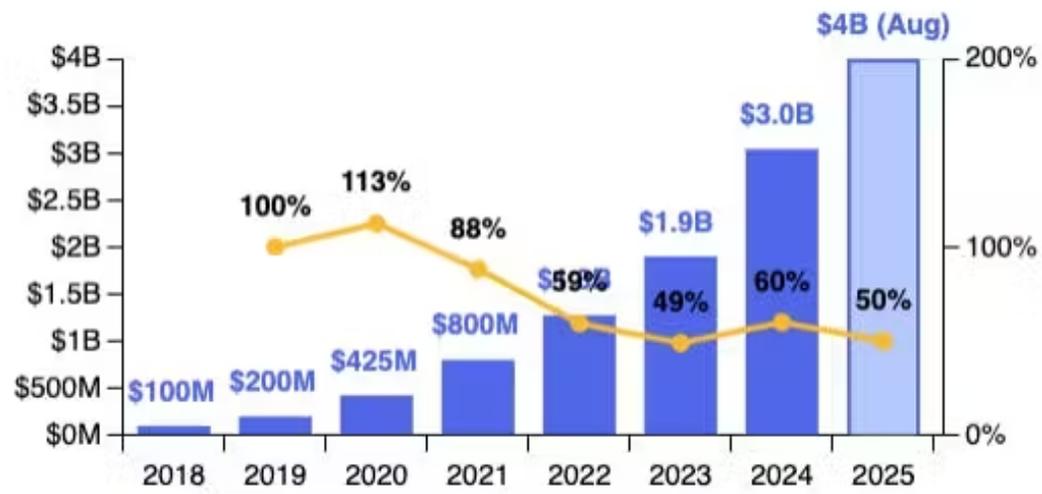
TL;DR: After becoming cash-flow positive with 85% gross margins and 650+ customers spending over \$1M annually, Databricks is positioned for what could be the largest software IPO since Snowflake in 2020—riding three secular waves of big data, cloud migration, and now enterprise AI where its model training products hit \$1B run-rate within 18 months. Sacra estimates Databricks hit \$4B in annualized revenue in August 2025, growing 50% YoY. For more, check out our [full report](#) and [dataset](#).



\$4.0B

↑ 50.0% YoY

All 3Y 5Y



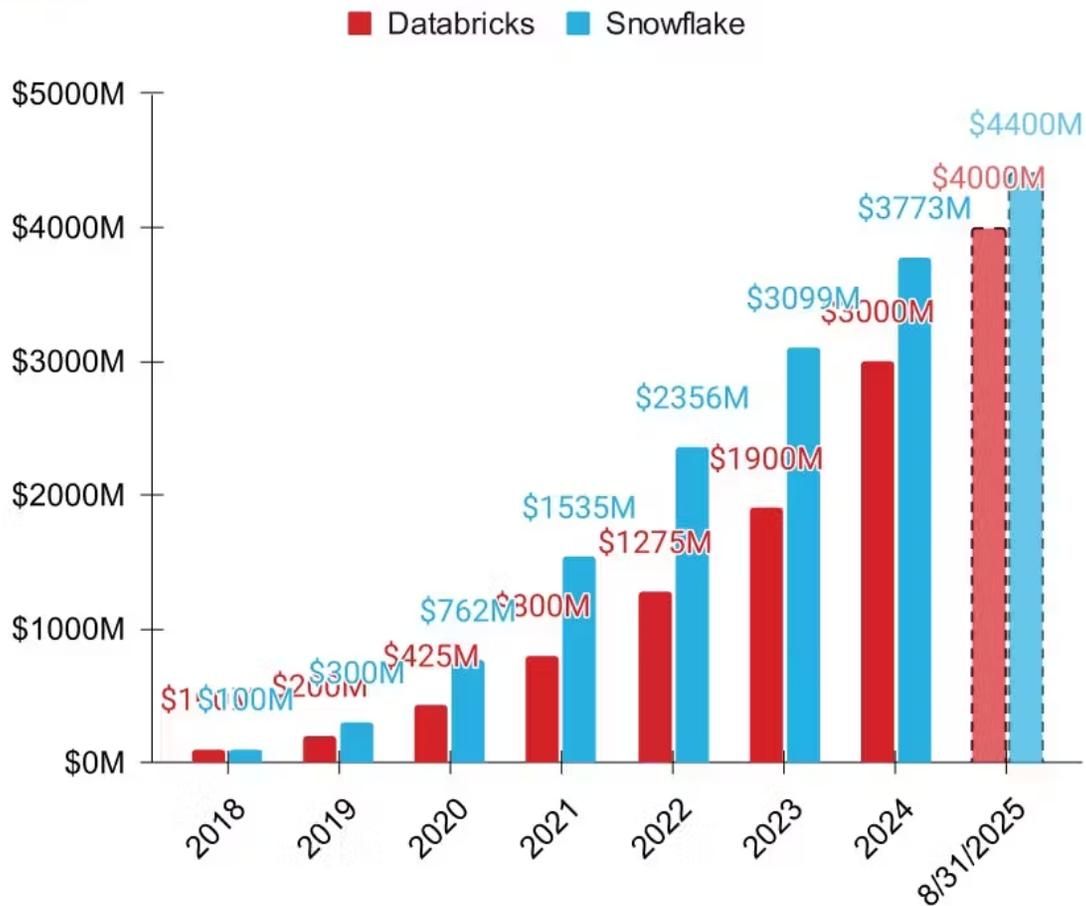
Source: Sacra

Updated: Oct 28, 2025

ARR (\$M) Growth



databricks vs. **snowflake**
Annual recurring revenue (ARR)



Key points via Sacra AI:

- **The 2000s explosion of web and mobile data—clickstreams, sensors, logs—combined with plummeting storage costs made Netflix-style personalization algorithms economically viable, but running experiments on terabytes of data remained prohibitively slow, inspiring a team of UC Berkeley researchers to build Apache Spark (2009), an open-source engine that let data teams **prototype ML models in hours vs. days**. At the same time that cloud data warehouses like BigQuery (2010) and Snowflake (2012) emerged to help analysts run SQL queries over structured data for dashboards & reports, Spark's creators founded Databricks (2013) as a managed platform for training models on semi-structured data (logs, clickstreams, JSON) with a set of defaults to eliminate the configuration nightmares that had bottlenecked Spark adoption.**
- **Where earlier frameworks like Hadoop MapReduce required teams of Java engineers to write low-level code & manually manage clusters—feasible only for Yahoo, Facebook, and other tech giants with dedicated infra teams—Databricks by 2014 found product-market fit**



wrapping Spark with collaborative notebooks where data scientists could write familiar Python or Scala & point-and-click cluster management so companies could use it without a DevOps team. Databricks monetized through consumption-based pricing (\$0.07-0.40 per compute hour, typically \$2K-10K/month) that undercut both Oracle/Teradata licenses (\$50K-200K/year) and the cost of hiring DevOps engineers (\$150K-250K fully loaded) to maintain on-premise Hadoop clusters.

- **Going multi-product with MLflow for ML lifecycle management (2018), Delta Lake (2019) for data storage, Databricks SQL (2021) to take on Snowflake in analytics workloads, and AI infrastructure via their \$1.3B acquisition of MosaicML acquisition, Sacra estimates that Databricks grew from \$275M in annual recurring revenue (ARR) in 2020 to \$4B as of August 2025, growing 50% YoY at a \$100B valuation for a 25x forward revenue multiple. Compare to cloud data warehouse & now competitive all-in-one Snowflake (NYSE: SNOW) at \$4.4B in ARR, growing 27% YoY, valued at \$90B for an 20.5x multiple, which is larger than Databricks by revenue scale but growing half as fast—at their present growth rates, Databricks will pass Snowflake in 2026.**
- **As AI became a C-suite priority across enterprises, Databricks has pushed to turn its vertically-integrated infrastructure for data storage, processing & ML into the core platform for enterprise AI, adding model training capabilities (\$1B revenue within 18 months) and acquiring Neon for \$1B (May 2025) to capture operational databases as 80% of new Postgres databases are now being created by AI agents rather than humans. The Neon acquisition enables Databricks to monetize the full stack of AI agent development where agents require both analytical processing for training and inference (Databricks' historical strength) and transactional databases for production applications (Neon's Postgres), positioning Databricks to capture value each time an AI agent spins up infrastructure.**
- **After becoming cash-flow positive with 85% gross margins (vs. Snowflake at ~70%) and maintaining 50%+ growth at \$4B+ revenue scale, Databricks is positioned for a late-2025 or early-2026 IPO that could be the largest software listing since Snowflake's \$33B debut in 2020, even as the product's technical complexity and the company's engineer-first, academic DNA potentially creates tensions with the enterprise dealmaking that will be key to**



sustaining growth in the public markets. With 650+ customers spending over \$1M annually (up from near-zero in 2019) and AI products hitting \$1B run-rate within 18 months of launch, Databricks has demonstrated it can ride multiple secular waves—first big data, then cloud migration, now AI.

For more, check out this other research from our platform:

- [Salesforce, Amplitude, and the fat data layer in B2B SaaS](#)
- [Fivetran: the \\$200M/yr Zapier of ETL](#)
- [Conor McCarter, co-founder of Prequel, on Fivetran's existential risk](#)
- [Sean Lynch, co-founder of Census, on reverse ETL's role in the modern data stack](#)
- [dbt Labs vs Databricks vs Snowflake](#)
- [Cribl at \\$200M ARR](#)
- [Databricks at \\$2.4B ARR growing 60%](#)
- [Grafana at \\$270M/year growing 69%](#)
- [Salesforce vs. Snowflake](#)
- [Julia Schottenstein, Product Manager at dbt Labs, on the business model of open source](#)
- [Earl Lee, co-founder and CEO of HeadsUp, on the modern data stack value chain](#)
- [Tristan Handy, CEO of dbt Labs, on dbt's multi-cloud tailwinds](#)