



A GUIDE FROM STONEFISH CONSULTING

The SMB Guide to Data Platform Readiness

How to know you're ready — and which platform actually fits

Why This Guide Exists

If you're a \$10M–\$50M company, you've probably already been pitched Snowflake, Databricks, or Microsoft Fabric by someone with a reseller margin riding on your answer. Here's the uncomfortable truth most vendors won't lead with: **the platform is rarely the hard part**. The hard part is knowing whether your business is actually ready to use one well — and if it is, which one fits your specific situation rather than the salesperson's quota.

This guide walks through both: a straightforward readiness self-assessment, and an honest, no-commission comparison of the five platforms dominating this market in 2026.

Part 1 — Are You Actually Ready for a Data Platform?

A data platform doesn't fix bad data, unclear ownership, or a team that isn't ready to use it. Work through this checklist honestly before you evaluate any vendor.

Business Alignment

- ☐ We have a specific business outcome this would drive (not just “better data”)
- ☐ Someone senior enough to fund and champion this has said so explicitly
- ☐ We know roughly what budget range we're working with

Data Foundation

- ☐ We know where our core data actually lives today (systems, spreadsheets, apps)
- ☐ We could name our 3 biggest “we don't trust this number” problems right now
- ☐ Our data volume and growth rate are roughly known, even if imprecise

Team & Process

- ☐ At least one person could realistically own this platform operationally
- ☐ We have SQL skills somewhere in the business, even outside IT
- ☐ We've thought about who would actually use this day-to-day, not just at launch

Governance

- ☐ We have at least an informal sense of who owns which data
- ☐ We know what compliance requirements apply to us (if any)



Scoring Yourself

Score	What It Means
9–11 checked	You're genuinely ready to evaluate platforms. Move to Part 2.
5–8 checked	Real gaps worth closing first — usually data foundation or ownership issues — but a phased approach (start smaller, prove value, then scale) can work.
0–4 checked	A platform purchase right now would very likely become shelfware. The right next step is a readiness assessment, not a vendor demo.

The single biggest predictor of failure isn't the platform — it's buying one before the business has decided who owns it and what specific problem it solves.



Part 2 — The Five Platforms, Explained Honestly

There is no universally “best” platform — only the one that fits your workload, your existing cloud footprint, your team's skills, and your budget shape. Here's an honest breakdown of each, without a partner margin attached to the answer.

Snowflake

A fully managed, near-zero-maintenance cloud data warehouse — the safest, lowest-operational-overhead default for SQL-first analytics teams. It runs natively across AWS, Azure, and GCP, and added AI/ML inference capability (Cortex AI) in 2025, narrowing its gap with Databricks on AI use cases.

Best fit if: Your team is SQL-first, you want the lowest operational burden, and you don't already have a strong ML/data-engineering team.

Databricks

A “lakehouse” platform combining data lake flexibility with data warehouse structure — the strongest option for teams doing serious machine learning or complex data engineering. It added an OLTP engine (Lakebase) in 2025 to further round out its capabilities, but it's built for teams with real Spark expertise and a meaningful data engineering capability, not a lightweight starting point.

Best fit if: You have (or plan to build) a real data engineering/ML capability, and AI use cases are a core priority, not an experiment.

Amazon Redshift

A cloud data warehouse tightly integrated into the AWS ecosystem. If your infrastructure already runs on AWS, Redshift avoids coordinating multiple vendors and remains dominant for AWS-native teams willing to invest time tuning workload management for meaningful cost savings versus serverless alternatives.

Best fit if: You're already deep in AWS and want one vendor relationship rather than several.

Google BigQuery

A fully serverless data warehouse with a pay-per-query pricing model. It's the simplest path if you're on Google Cloud and want minimal management overhead, though per-query scanning costs can escalate unpredictably on tables that aren't well partitioned — cost discipline matters more here than the sticker price suggests.

Best fit if: You're Google Cloud-native, your workloads are query-heavy but not constant, and you're disciplined about table structure.

Microsoft Fabric

A unified analytics platform built around Power BI and the Microsoft ecosystem. It's strongest when reporting and stakeholder workflows already live in Microsoft tools, providing a unified semantic layer across BI, notebooks, and data science under a single capacity model.

Best fit if: Your business already runs on Microsoft 365/Azure and your reporting culture is built around Power BI.



Part 3 — A Simple Decision Framework

Answer these three questions before anything else — they eliminate more wrong choices than any feature comparison.

1. Where does your data and infrastructure already live?

Environment	Likely Fit
Mostly AWS	Redshift or Snowflake
Mostly Google Cloud	BigQuery or Snowflake
Mostly Microsoft/Azure with Power BI in use	Fabric or Snowflake
Multi-cloud or undecided	Snowflake or Databricks (only two native across all three clouds)

2. What's the primary workload — Reporting/Bi, or Machine Learning/AI?

Primarily reporting and dashboards → Snowflake, BigQuery, Redshift, or Fabric.

Primarily ML, AI use cases, or complex data engineering → Databricks (or Snowflake, which has narrowed this gap).

3. What does your team actually look like today?

SQL-first analysts, no dedicated data engineering team → Snowflake, BigQuery, or Fabric (lower operational burden).

Real data engineering capability, or willingness to build one → Databricks or Redshift.

One honest caveat that applies to all five: each platform prices itself in a different currency — Snowflake credits, Databricks DBUs, BigQuery slots — which makes direct cost comparison difficult without normalizing for your actual workload. Any vendor quote you receive should be stress-tested against your real usage pattern, not their example calculator.



Part 4 — Common SMB Mistakes to Avoid

- **Buying based on a demo, not your data.** Every platform looks impressive on the vendor's sample dataset. Insist on testing against a real, messy slice of your own data before committing.
- **Underestimating the “who owns this” question.** The platform choice matters less than whether a real person is accountable for using and maintaining it after the contract is signed.
- **Following brand recognition instead of workload fit.** “Everyone's using Snowflake” isn't a reason if your workload is genuinely ML-heavy and your team has real engineering depth — Databricks may serve you better.
- **Ignoring your existing cloud footprint.** Multi-cloud sounds flexible but adds real operational cost; matching your platform to where your data already lives usually wins for a company at this size.
- **Skipping the readiness check entirely.** The platform is not the hard part. The business alignment, data foundation, and ownership questions in Part 1 determine whether any platform succeeds.

Next Step

If you scored well on the Part 1 readiness checklist and have a directional answer from the Part 3 framework, you're in a strong position to start a focused vendor conversation — and to walk in with enough knowledge that you won't be sold something that doesn't fit.

If there are real gaps, or you want an outside, non-commissioned second opinion before signing anything, that's exactly what a **Data & AI Readiness Assessment** is built to resolve — an unbiased look at your specific environment, workload, and team, with a clear recommendation and no reseller margin riding on the answer.

Stonefish Consulting | stonefishconsulting.com

Objective data and AI advisory. No platform to sell. No hype to buy.