

Modernizing Your Data: Taking a Dip Into a Data Lakehouse

We all know data helps organizations manage current business challenges and unlock future opportunities. And, over the years, we've learned that a successful data strategy hinges on what happens early in the data value chain: collection, storage, and other preliminary activities.

Modernizing your data processes upstream of business intelligence enables your organization to focus on the art of the possible, and ask the question, "What are we able to do now, that we couldn't do before?"

Today, the foundation of any modern data and analytics architecture should be a data lakehouse.



What is a data lakehouse?

Simply put, a data lakehouse combines the reliability and performance of a data warehouse with the openness, flexibility, and machine learning support of a data lake, enabling business intelligence and advanced analytics to coexist in a single platform. A data lakehouse provides a means to support all data (even voice and video), as well as all analytics and AI use cases.

A data lakehouse enables you to more effectively use data to accelerate your business. As a result, your organization can operate more efficiently, innovate faster, and support the breadth of use cases that today's hyper-competitive and constantly changing business environment demands.

How can I tell if a platform a true data lakehouse?

As word spreads about the benefits of data lakehouses, and the success many organizations are already realizing from them, some technology providers are latching onto the term, with varying degrees of accuracy. Let's look at the critical capabilities of a genuine data lakehouse.

A true data lakehouse:

- **Eliminates data silos:** Because all data, and the transformations performed against that data, live in one place, every team in an organization, (data, analytics, business, and others) can leverage data for an expansive range of use cases.
- **Has efficient, centralized, and transparent data governance:** A true lakehouse can provide centralized and transparent governance across data assets including files, tables, machine learning models, and dashboards on any cloud, while future-proofing your data and AI governance with the flexibility to leverage existing data catalogs and governance solutions.
- **Is endlessly flexible:** A real data lakehouse supports all types of data, and enables use cases at any stage in the analytics maturity curve. It's also able to produce outputs to support integrations with other technologies in the data, analytics, and AI ecosystems.

- **Accelerates innovation:** By combining insight into what's happening at any given moment with advanced analytics that predict what will happen next, and then suggesting which actions to take, a data lakehouse speeds use cases to implementation and delivers measurable results, quickly. A real data lakehouse delivers more use cases with the ability to support both batch and real-time streaming use cases.
- **Boosts efficiencies:** Engineers don't need to spend valuable time copying data from the lake to the warehouse, or dealing with complex management and tracking processes. Faster model and pipeline deployment, shorter production lifecycles, and shortened time to productivity combine to significantly reduce lost opportunity costs.
- **Makes the most of your time and money:** A data lakehouse offers a lower total cost of ownership by reducing labor costs for usage and maintenance; while delivering efficiencies in compute costs for ETL transformations; simplifying governing and allocation costs; and enabling predictable cost forecasting.

Which data lakehouse is the best choice?

Lovelytics has worked with companies of all sizes, each with unique needs for their data environment. Across these engagements, we've consistently found the Databricks solution to be the best fit for the vast majority of companies. Databricks excels in multiple areas, including cost, features, and future-proofing. Here are a few specific reasons why we've come to this conclusion:

A platform that delivers value across the spectrum.

Databricks generates value across different types of workloads, such as data processing for data engineering and machine learning. Databricks customers can easily scale their analytics business as the volume of data increases, orchestrate workflows, build predictive models, and productionize pipelines, all with minimal infrastructure management. That enables organizations to focus resources where they matter most, and encourage innovation by letting software manage the automatable pieces.

Integration with the broader ecosystem.

Because Databricks recognizes that other best-in-class technologies may deserve a place in a data and analytics stack, they've made it easy to integrate other solutions you're already using, such as Tableau, Fivetran, or Alteryx. That ease of integration gives customers more ways to easily make the most of their data.

The lowest total cost of ownership.

Databricks has consistently proven to be the best option to strategically align with an organization's time and money objectives, and drive maximum value over the long term. Databricks:

- Requires the least labor to use and maintain
- Is the most cost-effective platform for pipeline transformations
- Reduces the cost and technical and organizational debt a multi-tool architecture can incur
- Delivers faster model and pipeline deployment, shorter production lifecycle, and faster time to productivity
- Mitigates vendor lock-in risk and maximizes investment reusability

A team with a true vision and practical guidance.

From the founders to the sales team to the solution architects, the Databricks team is made up of smart, insightful people who think big about what data can do, and provide practical guidance on how to advance along the maturity curve. They're committed to accelerating time-to-value, starting with initial implementation and migration.

Conclusion

Today, any forward-thinking organization seeking to make more of their data should consider making a data lakehouse the foundation of their data and analytics architecture. The flexibility, efficiency, and expansive capabilities of a data lakehouse equip organizations to drive successful outcomes and take advantage of opportunity. With its innovative vision, vast flexibility, and exceptional long-term value, Databricks is the clear leader in the data lakehouse marketplace.

About Lovelytics

Lovelytics is a data and analytics consulting and services company. We help many of the world's leading organizations accelerate, optimize, and modernize their data ecosystems. The solutions we build in partnership with our clients are designed to drive business outcomes and create sustainable value through self-sufficiency and hands-on enablement. Successful organizations love Lovelytics – and you will, too.