A Modern Data Architecture

How to upgrade your stack to make faster, smarter business decisions

Table of Contents

What is a modern data stack?

Assembling the perfect data stack is a task that's never quite done. The most valuable stacks allow companies to take advantage of rapidly evolving technology, adding best-of-breed solutions for collecting, transforming, and analyzing data without requiring a large commitment of time and money. But too many companies see their plans for improvement impeded by the pitfalls of monolithic, legacy applications, and outdated architecture.

A modern stack is composed of dozens of tools and applications that work harmoniously to help teams get to insights faster. Each component should be both flexible enough to be swapped out independently of each other and robust enough to get answers to complex business questions, making continuous improvement more affordable and rewarding.

During a company's early stages, the data team can get by on **quickly assembling** a stack from open source and freemium tools. But as the business grows, analysts and stakeholders need to solve more nuanced problems, relying on increasingly large volumes of data in more varied formats to make decisions. To stay competitive, data teams must learn to modernize their stacks without overspending on costly implementations, falling victim to vendor lock-in, or diverting engineering resources from core initiatives.

Use this guide to get there

Understand the key characteristics of modern data architecture, decide when to make improvements, and future-proof your approach to analytics as you grow. The right mix of technology will be unique to your business, but the underlying concepts are the same: build a stack that empowers everyone to get answers faster, helps you build better products, and moves your data strategy to a new level of maturity.

Why modular
stacks scale better

Today's most talked-about data solutions aren't the monolithic solutions of previous decades. The proliferation of microservices and REST APIs changed the way companies approach data architecture, paving the way for a more modular stack that allows them to choose the best tools to solve discrete problems without disrupting the system as a whole.

Modular means scalable. As a company's growth goals become more ambitious, the data stack should evolve to meet them. This may mean that data solutions need more customization, more ability to connect to a broad set of business applications, and greater ease of adoption across departments and teams with very different initiatives.

Modular also means flexible. Rather than selecting a single vendor to cover the functions of data collection, transformation, analysis, visualization, and sharing, data teams can experiment with the best technology for each use case, testing new solutions without worrying about vendor lock-in or having to retool the entire stack.

Kitu Super Coffee worked with Sisu to streamline their stack by combining disparate data sources into a smooth, end-to-end data pipeline. As orders are placed through Amazon, Shopify, and Walmart.com, data from their fulfillment channel, Whitebox, is piped into Snowflake using Matillion, an ETL tool. Sisu continuously monitors new order data as it arrives in the Snowflake warehouses, providing weekly diagnoses on Kitu's key business questions, and helps their small analytics team carve out time for on-demand requests.

Read the case study
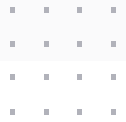
Stack components easily connect to each other

Data teams move faster when their stack components are interchangeable. Many popular tools today are built with this in mind, with dozens of pre-built connectors to common business apps and a REST API for custom integrations. Some are created for the sole purpose of piping data from one place to another, like Fivetran or Stitch, that let teams create single, automated pipelines with a standardized schema. This allows analysis to quickly explore data in new ways without spending months building connectors to each new tool they try.

Rapidly model and explore data to get to insights

To get insights faster, data teams must have the flexibility to dissect data in whatever way best addresses the question at hand. The best tool for the job will vary based on a number of factors: what language an analyst is most comfortable in, what question they're trying to answer, and what type of stakeholder is asking for insights.

A modern data stack doesn't limit analysts in their approach. It includes tools for working in several programming languages, visualizing data in many ways, and customizing dashboards and reports to focus on each department's most important metrics. Unlike monolithic solutions that often prescribe analytics workflows to their users, modular stacks made of flexible components let users dictate the best workflows for their goals. Analysts can freely slice and dice data to find new patterns and trends, while business stakeholders see the output in easy-to-use, no-code environments.

At Shopify, analysts are constantly busy with requests from everyone from project managers to salespeople. These questions often relate to conversion, retention, and events within the product that can help them make decisions on what to do next. To make reporting less cumbersome, Shopify analysts use Mode to quickly load reports for each department in the format they need without replacing their BI tool or overhauling their data infrastructure.

Read the case study
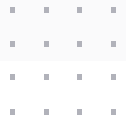
Enable global governance

Monolithic solutions typically include governance features that standardize data formats, check for duplicate or inaccurate data, manage compliance and security requirements, and provide a framework for the company's data policy. This poses a problem when it's time to change vendors—if governance lives inside a single business intelligence (BI) and analytics tool, you'll have to start from square one when implementing a new platform.

Since data must be accessible to a range of solutions, governed data should be consumable by all of them and free from disruption when one tool is swapped for another. You can address this by transforming data before it loads into any particular solution, creating scripts with a tool like dbt to ensure that no matter what you add or remove from your stack, you'll be able to maintain data quality.

Protect against vendor lock-in

Switching vendors is painful when a single solution accounts for most of your data infrastructure. It requires migrating large volumes of data without knowing whether the new solution is going to meet your needs, discouraging change by making it risky. Monolithic solutions often require more of an upfront investment of time and budget, creating a sunk cost mentality among teams that might otherwise choose a more fitting option.

Modular data stacks also make the buying process much easier. Each solution requires a smaller budget and fewer calls with sales reps, making the purchase less likely to go through IT procurement and endure a long approval process.

Learn more about future-proofing your stack with modular technology in this **webinar**.

Making moves: when to update your stack

Even the most mature data-driven organizations should be continually exploring new technologies, replacing stack components with better-suited options for the company. But in some cases, the need to change is more acute—here are some telltale signs that it's time to improve your data stack.

Your infrastructure can no longer support your needs

Growing companies need infrastructure that supports them as they mature. When you start to notice performance issues in your data workflows, it's a clear sign to re-evaluate your stack. While legacy solutions may offer pricing tiers for increased server capacity, sometimes the issue lies elsewhere in the process of piping, transforming, and querying data. To avoid overburdening engineering with a backlog of fixes, start looking for tools specifically designed to handle heavier data workloads, optimize pipelines, and shorten query times, like **Snowflake** or **BigQuery**.

In 2019, Instacart wrote about their experience moving from their Redshift cloud data storage warehouse to Snowflake. They had grown tired of spending engineering resources dealing with backlog, waiting for weekly VACUUM cleanup jobs to complete, and pausing important processes for data issues. With their modular data stack, they were able to migrate to Snowflake without affecting dependencies or disrupting daily operations, testing the new system for a month before rolling over.

Read the blog post

Your dashboards aren't telling you enough

Many companies invest heavily in a centralized BI tool and expect its dashboards to guide important business decisions. But having a warehouse full of data doesn't mean your data team can respond to each team's constant requests for information with actionable, timely insights. Rather than scramble to produce more dashboards for stakeholders, your team should have access to tools that allow them to dig deeper. This will help you answer every question from stakeholders, uncover new patterns, and quickly get to the bottom of why each teams' metrics are changing.

To get more proactive insights faster, add technology to your stack that challenges the traditional limitations of BI tools. In Mode, for example, **shared notebooks** let analysts examine data in many ways in SQL, Python, or R in the same environment. In Sisu, collaborative query libraries help analytics teams share and use queries efficiently across every application.

What happens when basic dashboards aren't enough to get to the bottom of the problem? See how data teams go deeper with complex exploratory analysis.

Read the paper

## You want to deploy analytics in your product

SaaS companies often use embedded analytics to create additional revenue streams. Consumer applications can benefit from this too, showing users' progress and rewarding them for certain engagement milestones. You can do this by investing in solutions like Mode that let you add custom dashboards to your product offering without disrupting existing data workflows. Before you launch new analytics features to your customer base, be sure to address the performance issues mentioned above so they don't impact your users, revenue, or retention rates.

Molecule uses Mode in its stack to embed analytics that look and feel just the Molecule app. Its clients get stunning, custom reports and dashboards while Molecule keeps its small development team focused on building core software.

Read the case study
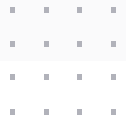
Mounting requests leave no time for proactive analysis

Many organizations never leave the reactive phase of data analysis. Even if you have well-governed data, a team devoted to analytics, and the ability to do in-depth analysis, you may still spend too much time responding to questions from each department. As you dig through the data to find the answers, it feels like more and more questions are piling up.

Data teams often spend 80% of their time preparing their data, and only 20% actually analyzing it. To get ahead, teams need to spend fewer hours each week manually slicing and dicing data or building new reports and dashboards for each question they're asked, and instead make time for proactive analysis. Look for tools like Sisu that automate the most tedious analytics tasks—like data prep or continuous analysis—and augment your analysis by comprehensively testing every factor in your data and surface new and actionable insights. These augmented intelligence solutions use advanced machine learning technology to diagnose changing metrics as they happen, investigate every dimension of your data, and give answers to questions no one knew they should ask.

Housecall Pro, a platform that enables billions of jobs and transactions a year, turned to augmented intelligence to quickly diagnose and act on changes across their dynamic platform. Their analytics team is tasked with answering hundreds of questions for the business, but on average it would take their team three days to answer each request. At best, they could answer 10 out of every 100 queries. After adding Sisu's augmented intelligence platform to their modern data stack, they can now search through billions of factors to test millions of hypotheses in seconds, finding their analysts can get to the facts 96% faster than using traditional, descriptive analytics tools alone.

Read the case study

# The final step: future-proofing your data strategy

There's more to choosing the perfect data stack than technology itself. Here are several ways to strengthen your data strategy and help the team easily transition the next time you need to make a change.

Empower stakeholders to act on your analyses

When business stakeholders across every department are involved in the analytics process, companies see a higher return on their investment in data. From choosing data solutions that are accessible to non-analysts to providing analysis that's easy to interpret and act on, your data stack should help each department understand what needs to be done in order to improve their KPIs and explore deeper.

Practically speaking, this means meeting the consumers of your data (both internal and external) where they are. Analysis is only as good as people's ability to consume it, and you can't expect stakeholders to dramatically change the way they work. Involve them in decisions about your data stack early so the systems you choose connect to the products they use daily, like Salesforce or Zendesk, and decide on the best way to share the results of your analysis.

When stakeholders can find insights in ways that make sense to them, they can act more quickly. Instead of delivering just charts and graphs, your data team should be able to tell the whole story, from in-depth analytics to clean visualizations to straightforward recommendations in tools people already know how to use.
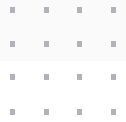
Stay agile and open to feedback

It's easy to overplan for analytics. Many companies that want to go from simply having data to using it in a way that makes them competitive spend a great deal of time meeting and planning for when they launch their perfect data strategy. But the most successful companies are constantly delivering incremental progress instead of perfection.

Rather than trying to decide exactly what events to track, which dashboards will be ideal for each stakeholder, and which metrics will be the most powerful a year out, start with something that can provide value quickly. Remember to choose modular, easily interchangeable technology that doesn't require a big financial outlay or put a burden on engineering resources, then let the team test it to learn what works and what doesn't.

Build a culture of data innovation

Access to data isn't enough. Even the best solutions and workflows can't guarantee that the various stakeholders at your company will know how to make the most of their available data. The more you encourage people to follow their natural curiosity, the more answers they'll find on their own using the self-serve tools you've provided.

To build a strong culture around data and exploration, constantly put data in front of people. Hold meetings where the data team shares insights and inspires other departments to ask more questions. Teach stakeholders different ways to dig into data without requesting reports from engineering. Show them how your data strategy helps them succeed, feeding them new facts and uncovering new patterns on a weekly basis that relate directly to each department's key metrics.

A modern, modular data stack paired with a strong, company-wide vision for how to use it can be the catalyst for entering the next phase of data maturity —a proactive, fast-moving team that helps the entire company make important decisions every day.
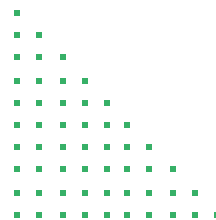
MODE

Mode is an advanced analytics solution providing the ideal environment for an analyst workflow—writing SQL, analysis in Python and R notebooks, visualizing results, and sharing insights—all in a single, seamless experience.

Get in touch

Try Mode with your data
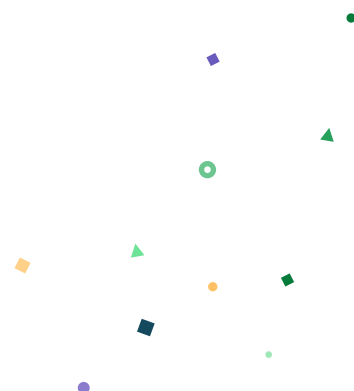
Sisu accelerates data exploration for analytics and product teams. As the first augmented intelligence solution built to quickly & completely analyze millions of dimensions and trends in cloud-scale data, Sisu helps teams understand, collaborate, and act using all of their data, faster.

Get in touch

Try Sisu with your data