

Drive *profitable* retail growth with dynamic data pipelines



If you're a data leader who has worked in the retail industry for a while—particularly over the last few years—you know that change may be one of the only things that hasn't changed. Customer expectations have changed. Economic conditions have changed. The way we measure profits and make business decisions have all changed. With all of this change, particularly in direct-to-consumer businesses, it's neither sufficient nor accurate to assume that topline revenue or even more customers equates with profitability.

To promote positive business outcomes and profitable growth, more and more organizations are turning to data. In a [survey of over 1000 IT leaders and line-of-business executives](#), respondents reported that in 2023, data and business analytics will be a primary area of focus and investment priority, second only to cybersecurity and risk management.

But even the savviest, data-driven decision makers must be discerning in the quality of the data they collect. For example, revenue-based models that measure customer lifetime value (CLV) but disregard variable costs—including the rising costs of goods, delivery, discounting, and marketing, plus customer expectations around shipping, cancellations, and returns—may not only paint a misleading profitability picture but also a costly one. [Gartner analysts estimate](#) that on average, low-quality data can cost organizations \$12.9 million USD and lead to subpar decision making. Long gone are the days when retail organizations can rely solely on hunches, balance sheets, or bad data to make sound business decisions. Now more than ever, success requires decision making based on granular, high-fidelity data.



Data- and insight-driven organizations are more likely to report above-market business growth, plus before-tax earnings increases ranging from 15–25%. With intelligent real-time data flows, key stakeholders can access data when they need it and gain both comprehensive visibility and nuanced business insights to make better business decisions.

The downside of data

Across industries, there's no shortage of available data. In 2022, enterprise data was estimated at 2.02 petabytes, more than 42% higher than the preceding two years. **But what good is an abundance of data that your company's key stakeholders can't access when and how they need it? What good are data pipelines that monopolize your engineering team's time—not on delivering strategic value, but rather on troubleshooting and maintenance? What good is data that adds little value to business outcomes?** As a retail data leader, you must understand your business needs and translate them into strategic solutions. You're accountable to make the best use of your resources, personnel, and data. But as you know—ingesting, integrating, and leveraging data to drive profitable growth isn't without its challenges.



Data architecture challenges

Exact, transform, load (**ETL**) and extract, load, transform (**ELT**) pipelines are two common data-ingestion processes that move raw data from various sources to destination databases such as **data warehouses** or **data lakes**. ETL pipelines will produce cleaner data, but they're time intensive, inflexible in real time, and can be prone to breakage. Alternatively, ELT processes are faster, more conducive for real-time data access, and require lower maintenance, but they can be messy. In the absence of precise metadata, raw and unorganized data stored in data lakes can quickly turn from a business asset to a festering liability—becoming unusable, exploitable, and irretrievable. It can transform a data lake—that is, a cost-effective solution for storing copious amounts of raw data—into a **data swamp**. But the additional modeling required to make raw data useful, consumable, auditable, compliant, and accessible in real time, often comes with additional challenges and costs.

Retail-specific data pipeline challenges

The retail industry is anything but swamp-like. It's dynamic, changing, and in perpetual motion. Retail inputs and data sources are too. Variable costs are steep and can fluctuate daily, affecting both earnings and income. Promotions, for example, [can diminish grocer sales and margins](#) by as much as 10–15%, but when executed optimally, they can increase sales 4–8% and net income by as much as 3%. Inflexible data pipelines with point-to-point connections aren't built for variability, and as such, they're more likely to break.

Additionally, most end-to-end platforms designed to connect, model, transform, store, visualize, and activate data aren't built specific to retail. To produce industry-relevant and useful data in real time, they often require additional work, time, and money.





Challenges in data enablement

Once data pipelines are flowing optimally, data must still be translated and accessible to be useful. But despite advances in analytics and infrastructure, many retailers' data is siloed and dispersed across multiple platforms, applications, software systems, and technology—each with diverse conventions. Consequently, data is often underused or lost. [According to one Forrester analyst](#), 60–73% of enterprise data goes unused for analytics. Further, unifying and standardizing fragmented data can bottleneck decision making, contributing to a common drawback of many data-infrastructure projects—failure to link them to business decisions and outcomes. [Gartner analysts estimate](#) that through 2022, only 20% of analytic insights will deliver business value.

Overworked yet underutilized engineering teams

To maximize investments, many organizations will double down on efforts and put their engineering teams to work. But even with best-in-class tooling, in-house data pipeline projects are often burdensome and complex. They pull data engineers away from strategic projects and tasks designed to derive business value from data insights. [In one study of over 23,000 data professionals](#), respondents reported spending 40% of their time gathering and cleaning data, plus an additional 20% building and selecting models. As a data and technology leader in your organization and an ever-evolving industry, your time and team are invaluable. **Frankly, if your priority is profitable business growth—building, troubleshooting, and maintaining data flows isn't a great use of your time and resources.**

Unlock the value of retail data with dynamic data pipelines

To drive profitable growth, you must equip your organization to manage variable costs in real time. This better data awareness and enablement often requires employing tools to help you collect, match, and organize your data, as well as to standardize key business metrics. To begin, you must adopt modern cloud architecture.

Modernize your data architecture

Again, retail data fluctuates and is ever-changing. As such, it must not only be accessible in real time, but it must also be auditable. Big-data analytics, AI, reporting, BI, and more require data at rest, however monolithic legacy architecture alone is insufficient to meet the variability that retail data requires. Your organization must be able to ingest and store both defined and undefined data. Further, in-the-moment data access requires that your architecture can process data in motion. With architecture bolstered by technology such as Apache Kafka, you can process and quickly operationalize variable data, as well as add both flexibility and real-time access to your data architecture. Additionally, you can simplify storage requirements, complexity, and costs, while also maintaining data fidelity.

Make it retail specific

To drive profitable growth and optimize lifetime value from your customers (CLV), you must have the right resources and tools in place not only to make real-time decisions, but to make strategic business decisions informed and driven by retail-specific insights and data models. When considering CLV, for example, a revenue-only lens is insufficient. You must also consider profitability over a lifetime of transactions. Your data should help you to identify your most impactful programs, to predict your customers' buying behaviors, to set retention-cost thresholds, to evaluate and determine optimal customer acquisition costs (CAC), again, across a lifetime of transactions. To leverage data to drive profitable growth in this industry, your organization should have real-time visibility into marketing, operations, discounting, variable revenue, costs, and margin inputs.



Make it useful

Once your data is transformed into a retail-specific context, leaders in your organization can access current data whenever they need it. They can use it to make informed decisions and take immediate action. The best tools will store all your data in one place, standardize disparate data sources, apply metadata to appropriately catalog your data, then structure and organize your data so it's not only readily accessible, but also meaningful. They will visualize data in such a way that users across your organization can infer key metrics and insights in the moment.

Make better use of your time

You and your team have better things to do than preparing data and endlessly building and maintaining data pipelines. There is a better way! By mapping and cataloging data earlier, you can reduce all those manual and time-consuming ad hoc tasks. Rather than building from scratch, you can leverage pre-built connectors, data catalogs, and semantic structures—all custom built for retail. You can outsource connection management and minimize valuable time spent maintaining data ingestion. You can minimize costs and data loss by storing event logs in a data lake. And you can stop toggling between disparate tools by consolidating your data into a single platform. With the right solution, you can increase the transparency and auditability of your data, operationalize and activate quality data in real time, and leverage your data and data pipelines to directly impact business outcomes.



Discover profitable growth with SoundCommerce

You needn't manage your data pipelines alone, however. With SoundCommerce—the dynamic data-ingestion platform custom-built for retail organizations—you can quickly and easily build and maintain complex data pipelines and transformations. Built on Google Cloud, SoundCommerce integrates, standardizes, and structures data as it streams into your data warehouse for increased flexibility, plus advanced retail analytics and data activation. Through pre-built connections, data models, and tools, SoundCommerce connects and models marketing, operations, and merchandising data. The platform supports transparent and auditable data preparation in a low-code no-code (LCNC) user experience. It integrates with existing data infrastructure, reducing the time required for configurations. And since data is managed through an auditable, governable, self-service experience, you can take back more of your valuable time and resources.

With SoundCommerce you can:

- Reduce data latency.
- Improve data quality and availability.
- Promote an organizational culture of visibility and trust.
- Improve operational efficiency.
- Minimize wasted time.
- Drive profitable growth with every data-driven decision.

On average, retail organizations that have partnered with SoundCommerce have experienced a **19% increase** in CLV, a **30% increase** in average order value (AOV), and a **50% year-over-year decrease** in CAC. With SoundCommerce, you can too.

[Visit SoundCommerce.com](https://www.soundcommerce.com) to learn more about how SoundCommerce can help you drive profitable growth.



