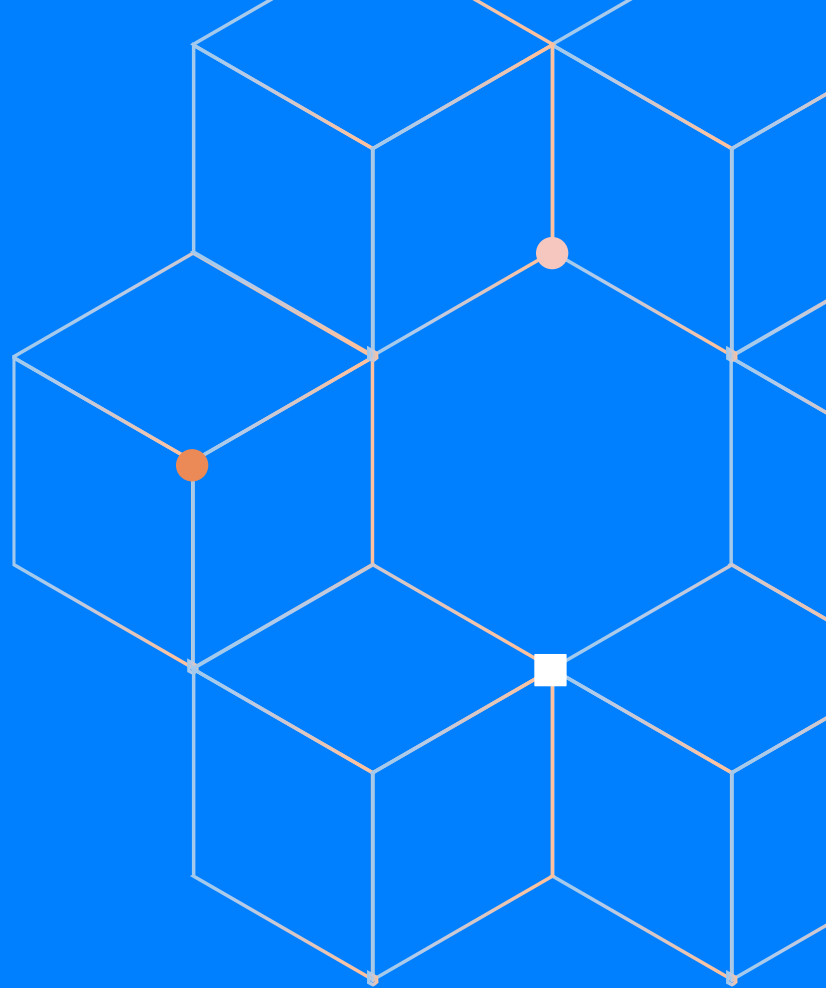




Scalable data infrastructure in under an hour

Affordable and easy to launch, a modern data stack can take you from startup to enterprise

Table of contents

01	Introduction: Address current & future data needs without building infrastructure	3
-----------	--	----------

02	The modern data stack: Origin, benefits, capabilities	5
	• What is a modern data stack?	5
	• How startup data teams can drive growth with an MDS	6
	• Real-world analytics projects powered by an MDS	7
	• Finance & marketing analytics	7
	• Product analytics	7
	• Support analytics	7
	• Sales analytics	7
	• Operational analytics	7

03	How to launch a modern data stack in under an hour	8
	• Setup process	8
	• Demo video	9

04	How to choose modern data stack components	10
	• Data integration criteria	10
	• Data warehouse criteria	11
	• Business intelligence criteria	12

05	Looking ahead: How a modern data stack enables data maturity	13
	• Defining data maturity	13
	• The benefits of being data-mature	14
	• Data maturity & the modern data stack	15

01

Introduction: Address current & future data needs without building infrastructure

Data leaders at startups and smaller companies face a set of somewhat paradoxical challenges. They need to:

- Generate and deliver reliable metrics that support rapid growth
- Build the foundation of a modern data infrastructure that can scale to meet future needs
- Minimize costs and engineering headcount, partly by limiting engineering hours dedicated to basic ETL tasks

Satisfying analytics stakeholders while honoring budget constraints can be a tricky balancing act. These days, however, startup data teams can set up scalable data infrastructure in under an hour and pay as little as a few hundred dollars a month for it.

What makes this possible is a set of cloud-based technologies collectively known as the modern data stack. A modern data stack (MDS) features prebuilt, automated data integration that eliminates the need to build and maintain data pipelines and delivers normalized, analysis-ready data into a destination, usually a cloud data warehouse or data lake.

An MDS enables any data team to deliver reliable, up-to-date business metrics derived from multiple data sources, without investing in infrastructure, adding data engineering talent or diverting existing data engineering hours.

"These days, startup data teams can set up scalable data infrastructure in under an hour and pay as little as a few hundred dollars a month for it."

Because MDS technologies are cloud-native, they scale infinitely, so you won't need to continually engineer your analytics stack as you grow — an MDS can be the first step on your journey to data maturity.

Startups tend to focus on the foundational stage of data maturity (descriptive analytics), but an MDS can easily address your data needs as you progress to the proactive stage (diagnostic analytics) and finally to the innovation stage, where you're leveraging prescriptive analytics, AI/ML and more.

02

The modern data stack: Origin, benefits, capabilities

What is a modern data stack?

The most important difference between a modern data stack and a legacy data stack is cloud technology. An MDS is hosted in the cloud and requires little technical configuration or maintenance. Legacy stacks — and specifically the extract-transform-load data integration sequence — evolved at a time when computing power, storage and bandwidth were scarce and expensive. Data pipelines were often built for specific analytics projects, and data transformation occurred before data was loaded into a destination. Upstream schema changes, as well as downstream changes to business requirements or data models, could break the pipelines.

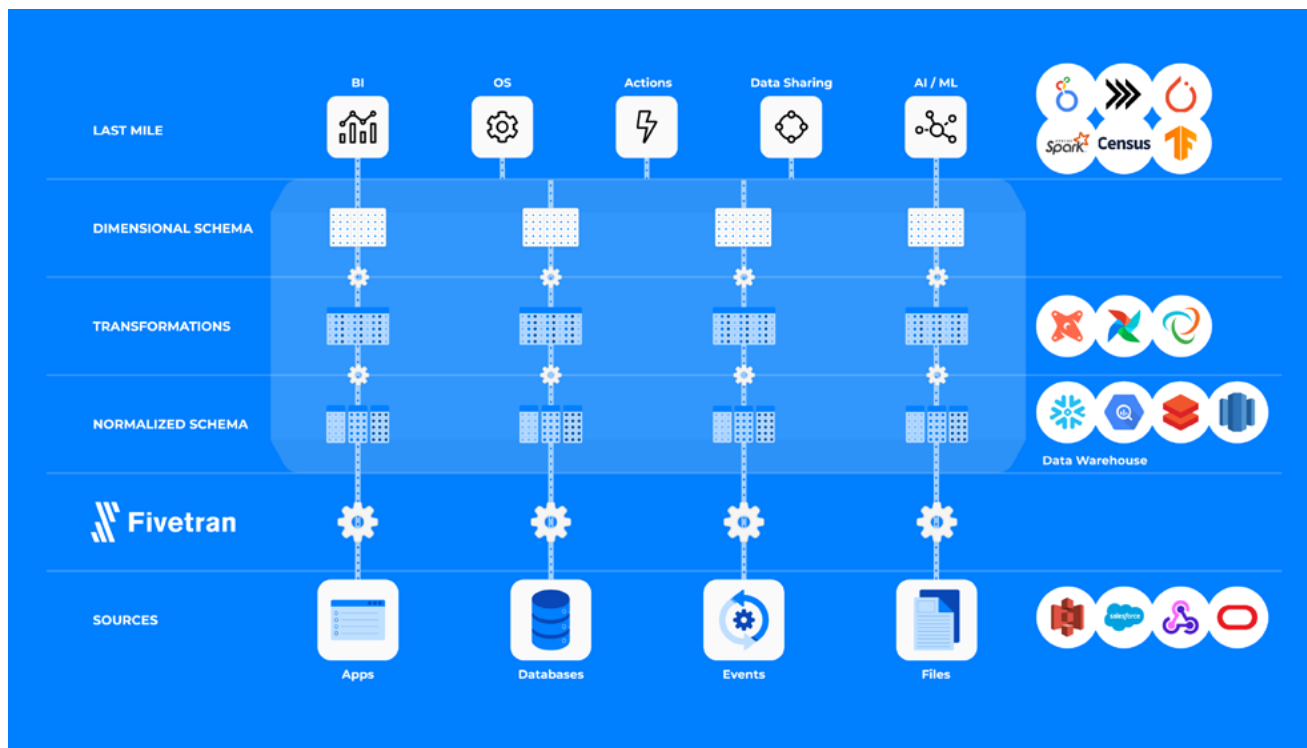
The off-the-shelf components of a modern data stack are far more automated, robust and scalable, and they're built with analysts and business users in mind. An MDS makes data analysts far less dependent on data engineers — analysts can add new data sources in minutes and maintenance is handled by vendors. Data engineers are free to focus on complex strategic projects.

"An MDS makes it easy for analysts to generate cross-source metrics that point the way to faster product improvement, better customer retention and more."

Here are the components of an MDS, in order of how data flows through the stack:

- Fully managed [data pipelines](#) that automatically adjust to API and schema changes
- Cloud-based columnar warehouse or data lake as a destination
- Data transformation tool
- Business intelligence or data visualization platform

The modern data stack



How startup data teams can drive growth with an MDS

To survive, startups need to grow, and a modern data stack can play a major role in driving growth. By centralizing and normalizing an organization's data, an MDS makes it easy for analysts to generate cross-source metrics that can point the way to more rapid product improvement, better customer retention and more.

By building out near real-time growth dashboards, for example, data teams can ensure that smaller marketing departments with limited budgets are maximizing every dollar spent to acquire new customers.

Go-to-market teams are another good example. By using an MDS to join marketing, sales and support data, they can better understand where prospects and customers sit in the buying cycle, allowing them to create a more tailored customer journey and better purchasing experience.

Real-world analytics projects powered by an MDS

An MDS allows your team to tackle any kind of analytics project — and often enables stakeholders to answer ad hoc questions as well, because it makes data exploration accessible to nontechnical users. Here are a few examples of the kinds of projects data teams tend to take on after adopting an MDS.

Finance & marketing analytics

Finance teams sit on a wealth of data, but it's often siloed. Finance and marketing analysts use an MDS to centralize revenue, fulfillment and inventory data from sources like NetSuite and Quickbooks. They can then join it with marketing data to build metrics like customer lifetime value and conversion rate. Insights into the customer lifecycle help both marketing and sales teams improve audience targeting and cross-selling.

Product analytics

By aggregating product data, growing businesses can learn how customers are using their product. A startup gaming company, for example, could analyze data on in-game behavior and use it to further develop the features customers love.

Support analytics

Customer success platforms like Zendesk collect data on how many support tickets are being handled per day and how quickly each is resolved. Centralizing that data allows analysts to aggregate metrics and create an individual customer health score, an invaluable tool for reducing churn.

Sales analytics

Data from Salesforce and other CRMs plays a critical role in the customer lifecycle. Joining sales and support data from multiple sources provides key customer data prior to an engagement. For example, analysts can help sales teams understand how to tailor their approach depending on which marketing materials a prospective customer has engaged with.

Operational analytics

Centralizing and analyzing operational data from ERP systems like SAP and Oracle can help businesses boost efficiency and expand margins. Teams can measure the progress of a product build, do job costing to understand margins, and tackle other operational analytics projects.

03

How to launch a modern data stack in under an hour

Setup process

We recently set up a modern data stack in under an hour, and we have the video to prove it.
The components of our stack:



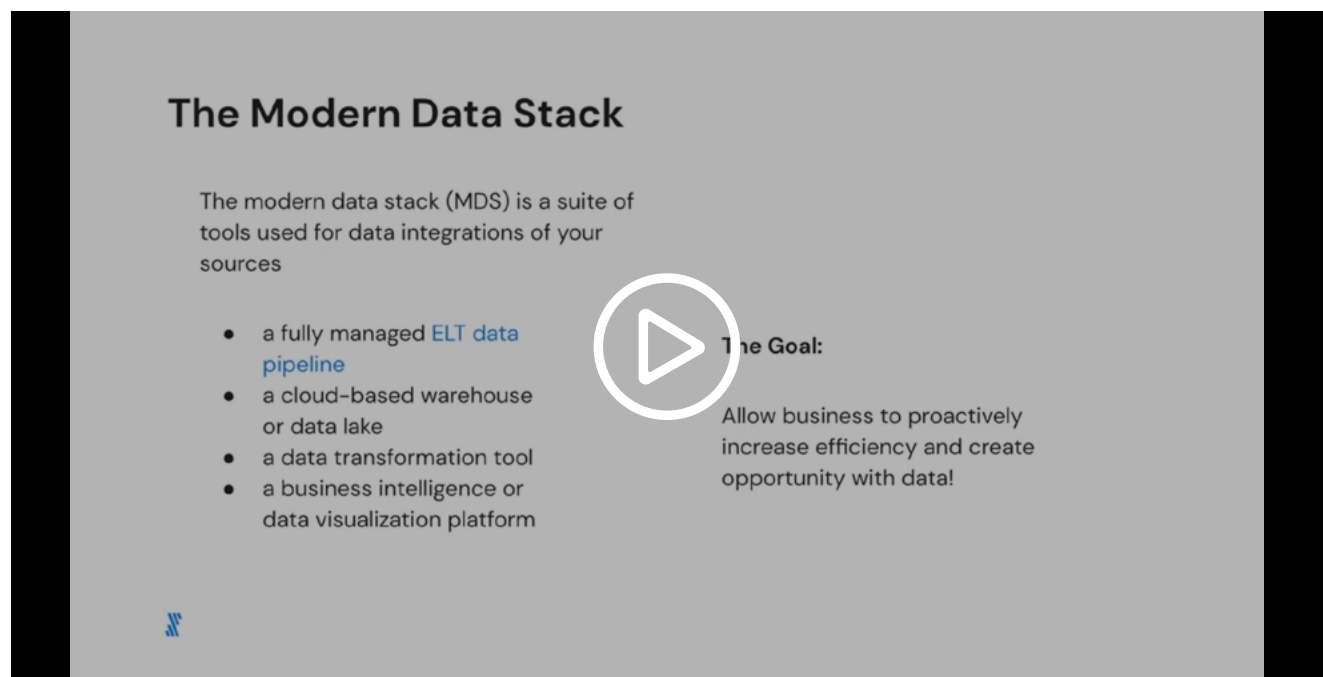
That sub-hour elapsed time includes loading Salesforce data into Snowflake and running an in-warehouse transformation — a simple JOIN of **Opportunity** and **Accounts** to see **Closed Won** deals sorted by dollar amount. We also had time to make a couple of real-time calculations and visualizations using our BI tool, Sigma.

"In minutes, we were able to start a new data warehouse, ingest Salesforce data, orchestrate a transformation and make real-time calculations."

In a matter of minutes, then, we were able to start a new data warehouse, ingest Salesforce data, orchestrate a transformation on the warehouse and make real-time calculations with our landed data. You'll see all that in the video below.

But that's about as minimalistic as a modern data stack gets. What could we do next? If we wanted to understand how marketing spend impacted sales, we could effortlessly add connectors for Google Analytics, Marketo and one or two other marketing sources, and create that metric. Or we could load our production application's statistics into Snowflake and infer how interactions with our customer success team are accelerating customer application usage.

Our modern data stack can accommodate all manner of analytics projects — check out how we set it up below.



The Modern Data Stack

The modern data stack (MDS) is a suite of tools used for data integrations of your sources

- a fully managed [ELT data pipeline](#)
- a cloud-based warehouse or data lake
- a data transformation tool
- a business intelligence or data visualization platform

The Goal:

Allow business to proactively increase efficiency and create opportunity with data!

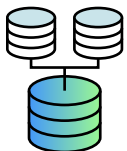
04

How to choose modern data stack components

That's how easy it can be to set up a modern data stack. Choosing the individual components, however, may require a bit more thought, depending on how much experience you have with the tools.

There are many capable, mutually compatible MDS technologies in the market, including some built by tech giants like Amazon, Google and Microsoft. Most offer affordable pricing plans for startups and smaller businesses.

It does make sense to consider your current and future use cases, however, and choose tools that can accommodate them. Here are the key criteria to consider when choosing each tool.



Data integration criteria

The technical approaches and feature sets of data integration tools vary quite a bit. Here are the factors to keep in mind:

1. **Data connector quality.** (a) *Proprietary or open-source?* The former tend to be of higher quality and integrate more seamlessly with other tools. (b) *Standardized schemas and normalization?* Normalization fosters data integrity; standardization enables automation and support for the same solution across customers. (c) *Incremental or full updates?* Log-based incremental updates can be more frequent, without impacting operations.

"There are many capable MDS technologies in the market with affordable pricing plans for startups."

2. **Support for sources and destinations.** Does the provider support your sources and destinations? Do they often add new sources and solicit new connector requests?
3. **Configuration vs. zero-touch.** Zero-touch, fully managed tools are very accessible, with standardized, stress-tested, maintenance-free connectors.
4. **Automation.** A tool should remove as much manual intervention and effort as possible with features like automatic schema migration and adjustment to API changes.
5. **Transforming within vs. before the data warehouse.** Under ELT (extract-load-transform), transformations occur in an elastic cloud warehouse, allowing limitless resource scaling.
6. **Recovery from failure.** To avoid permanent data loss, find out whether your prospective tools are [idempotent](#) and perform net-additive integration.
7. **Security and compliance.** Learn how prospective tools deal with regulatory compliance, data retention, role-based access, and column blocking and hashing.

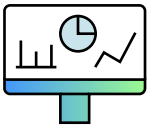


Data warehouse criteria

A data warehouse is the repository of record for your structured data, and different warehouses offer different feature sets and tradeoffs. Here's what you should focus on:

1. **Centralized vs. decentralized data storage.** Is data stored on one machine or distributed across multiple machines, trading redundancy for performance?
2. **Elasticity.** Can the data warehouse scale compute and storage resources up and down quickly? Are compute and storage independent or coupled together?
3. **Concurrency.** How well does the data warehouse accommodate multiple simultaneous queries?
4. **Load and query performance.** How quickly can you complete typical loads and queries?
5. **Data governance and metadata management.** How does the data warehouse handle permissions and regulatory compliance?

6. **SQL dialect.** Which dialect of SQL does the warehouse use? Will your analysts have to adjust the syntax they use?
7. **Backup and recovery support.** If your data warehouse somehow gets corrupted or breaks, can you easily revert to a previous state?
8. **Resilience and availability.** What about preventing database failures in the first place?
9. **Security.** Does the data warehouse follow current security best practices?



Business intelligence criteria

Business intelligence tools enable you to easily build reports and dashboards, but different tools have different strengths and weaknesses. Here are the key factors to consider:

1. **Seamless integration with cloud data warehouses.** Is it easy to connect this BI tool to your cloud data warehouse of choice?
2. **Ease of use and drag-and-drop interfaces.** Ease of use is especially important to popularizing data-driven decisions across your organization.
3. **Automated reporting and notifications.** Writing reports by hand gets tedious. Does the BI tool allow you to schedule reports to publish automatically? What about alerting users when the data changes?
4. **Ability to conduct ad hoc calculations and reports by ingesting and exporting data files.** Your analysts and data scientists might want to explore data without the overhead of having to go through a data warehouse first.
5. **Speed, performance and responsiveness.** Basic quality-of-life considerations are important, like dashboards and visualizations loading in a timely manner.
6. **Modeling layer with version control and development mode.** Does the BI tool allow your analysts to work collaboratively by sharing data models and code?
7. **Library of visualizations.** Pie charts, column charts, trendlines and other basic visualizations can only take you so far. Does the BI tool feature more specialized visualizations, like heat maps or radar charts? Does it allow you to build your own?

05

Looking ahead: How a modern data stack enables data maturity

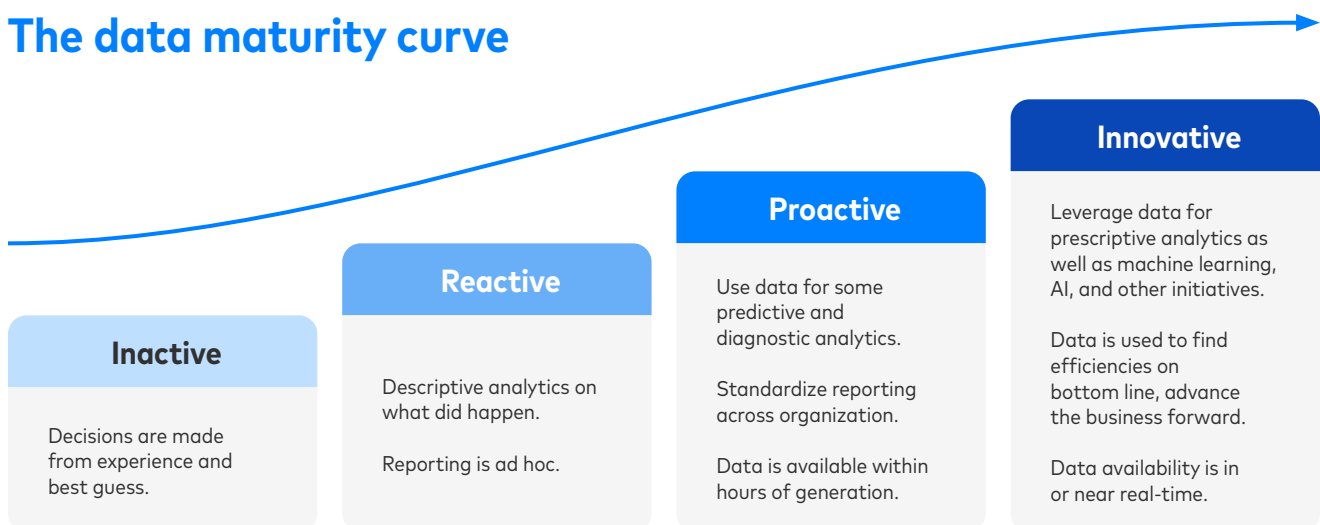
Defining data maturity

The concept of data maturity is a tool for assessing how effectively a business utilizes its data, the ultimate goal being to ensure that data-derived insights inform strategic decision-making across the organization.

"Data-mature" businesses typically wield forward-looking analytics — they don't just understand what happened in the past, and why; they can respond to market signals in real time and anticipate the future effects of a range of strategic decisions.

Progress towards data maturity is often measured in stages along a curve. Here's one of the models we use at Fivetran:

The data maturity curve



Data-mature businesses tend to have an organization-wide single source of truth, with standardized reporting, and analysts and decision-makers work with real-time or near real-time data. These businesses have moved from descriptive analytics to diagnostic and predictive analytics and finally to prescriptive analytics, often in tandem with AI/ML initiatives.

The benefits of being data-mature

The concept of data maturity is broader than the concept of data-drivenness, but a data-mature business is by definition data-driven. And we have a lot of evidence that data-driven organizations outperform their competition. Take these findings from Forrester's "State of Insights-Driven Business Maturity":

- Data-driven companies are 178% more likely to grow revenue
- They are 240% more likely to have a competitive edge
- Between 2019 and 2021, they took \$1.8 trillion from their competitors

Data-mature organizations also look a lot like what Thomas Davenport, in the Harvard Business Review and his [recent book](#), has called "[analytics competitors](#)": businesses that leverage analytics to consistently outcompete their peers.

Analytics competitors go well beyond basic analytics. As Davenport observes, they use predictive modeling to understand which customers are most profitable, how much they will buy over a lifetime, what prices they will pay, and what will trigger them to buy more. They set prices in real time; anticipate supply chain bottlenecks and simulate solutions; and model the impact of operational costs on profit margins.

To achieve this, they need to rely on modern data infrastructure — a modern data stack.

"A modern data stack supports data maturity by automating data integration and standardizing distribution processes and channels."

Data maturity & the modern data stack

A modern data stack supports data maturity and advanced analytics in a number of ways. Because data integration, storage and analysis are largely automated and infinitely scalable, organizations can easily access and leverage all of their data — there are no incomplete data sets or missing data sources, and plenty of fuel for AI/ML.

A modern data stack also standardizes all data distribution processes and channels; everyone is operating from the same source of truth and reports all emerge from the same tool set. Data access is universal and data sharing is seamless, so nearly all decisions are influenced by relevant metrics visible to everyone.

This is how analytics competitors operate. They eliminate siloed data *and* siloed analytics. Analytics initiatives are centrally managed and rely on the same technology, tools and data.

So a modern data stack isn't just easy to set up and affordable for startups and smaller companies — it's a flexible, scalable infrastructure solution that will support growth and a maturing data program for years to come.

Start building your modern data stack for free



Fivetran is the global leader in modern data integration. Our mission is to make access to data as simple and reliable as electricity. Built for the cloud, Fivetran enables data teams to effortlessly centralize and transform data from hundreds of SaaS and on-prem data sources into high-performance cloud destinations. Fast-moving startups to the world's largest companies use Fivetran to accelerate modern analytics and operational efficiency, fueling data-driven business growth. For more info, visit [Fivetran.com](https://fivetran.com).