Dremio & dbt

Perform easy SQL transformations in your open data lakehouse

Analytics Engineering Meets Open Data Lakehouse

Cloud data lakes represent the first and primary destination for the newest and fastest-growing sources of customer and operational data for businesses. The Dremio open data lakehouse provides the data management, data governance, and analytic capabilities of a data warehouse directly on data lake storage.

dbt gives data teams a development environment to transform data in their data lake into tables or views using simple SQL statements. Together, Dremio and dbt deliver an easy and open data lakehouse solution that gets insights into the hands of data consumers quickly.

Dremio

The Need for an Open Data Lakehouse

Cloud data lakes give organizations a massively scalable and flexible storage solution for large volumes and a variety of data. To make that data available to data consumers, data teams need to build Extract, Transform & Load (ETL) pipelines to move that data into data warehouse platforms. These pipelines are often manual and ad hoc. Business users often require custom data copies in the form of BI cubes and extracts to meet query performance SLAs. The result is a massive backlog of data access requests that can take weeks and even months for data teams to fulfill.

The Dremio open data lakehouse combines the scalability and flexibility of the data lake with the data management, data governance, and data analytics capabilities of the data warehouse. Dremio's no-copy architecture enables data consumers to analyze data in place, either in the data lake or in other cloud data lakes, relational and NoSQL databases, and on-premises HDFS and object storage solutions.

Dremio simplifies data architectures, so data access requests now take just hours or days instead of weeks or months, and enables self-service analytics for a wide range of technical and non-technical data consumers.

dbt

Analytics Engineering

As businesses move towards self-service analytics, data teams need to transform raw data into production-ready datasets that anyone can explore using their BI tool of choice to get consistent and reliable answers.

dbt Core is a development framework that gives anyone who knows SQL the ability to transform data in the data lake using software development principles like version control, modularity, and documentation.

dbt replaces the need to write DDL/DML with simple SQL SELECT statements to build tables and views, transforming raw data into trusted data products. dbt enables data teams to develop, test, and deploy analytics code rapidly.

Dremio and dbt give data teams the ability to load data into the lakehouse once, transform it into tables and views, and make it instantly accessible to data consumers across the organization for high-performance SQL queries and interactive dashboards, without moving or copying the data to another repository.

Dremio & dbt, Better Together

Open

Dremio and dbt are based on open source projects with large community support. Data remains in the data lake in open formats, so organizations can use best-in-class processing engines and reduce the risk of vendor lock-in.

Sub-Second Performance

Dremio provides customers with sub-second SQL queries and enables high concurrency so customers have access to enterprise-grade BI directly on the data lake.
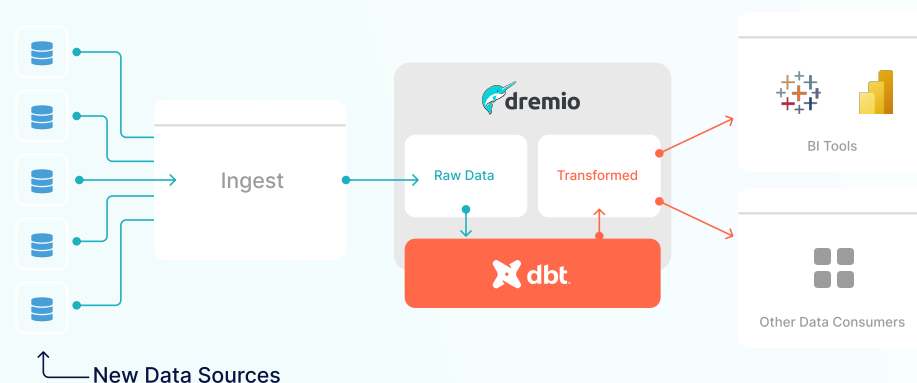
Self-Service SQL Analytics

Dremio and dbt use SQL to transform, join, and query data using the BI tools of their choice. Technical and non-technical data consumers have easy and efficient access to consistent and reliable datasets.

Collaborative

With Dremio and dbt, data analysts can now collaborate with data engineers and data scientists to build, test, and deploy data transformation pipelines in a shared workspace.

Analyze in Place

Dremio and dbt enable data teams to move data into the data lake and analyze it where it resides, instead of moving or copying the data to a proprietary data platform, or building and maintaining BI cubes and extracts. Dremio gives data consumers the ability to access, join, and query data in disparate repositories, including multiple clouds, relational and NoSQL databases, HDFS, and more.

Get Started Today

Get started with Dremio Cloud for free today. All you pay for is AWS resources. Sign up at

Dremio.com/get-started

Dremio on dbt setup guide. Details at

https://docs.getdbt.com/reference/warehouse-setups/dremio-setup

Try dbt Core for free today. Get started at

https://docs.getdbt.com/docs/get-started/getting-started-dbt-core