



Manage Production Analytics with Databricks SQL

Module 07



Module Agenda

Manage Production Analytics with Databricks SQL

Introduction to Databricks SQL

DE 7.1 – Navigating Databricks SQL and attaching to SQL Warehouse

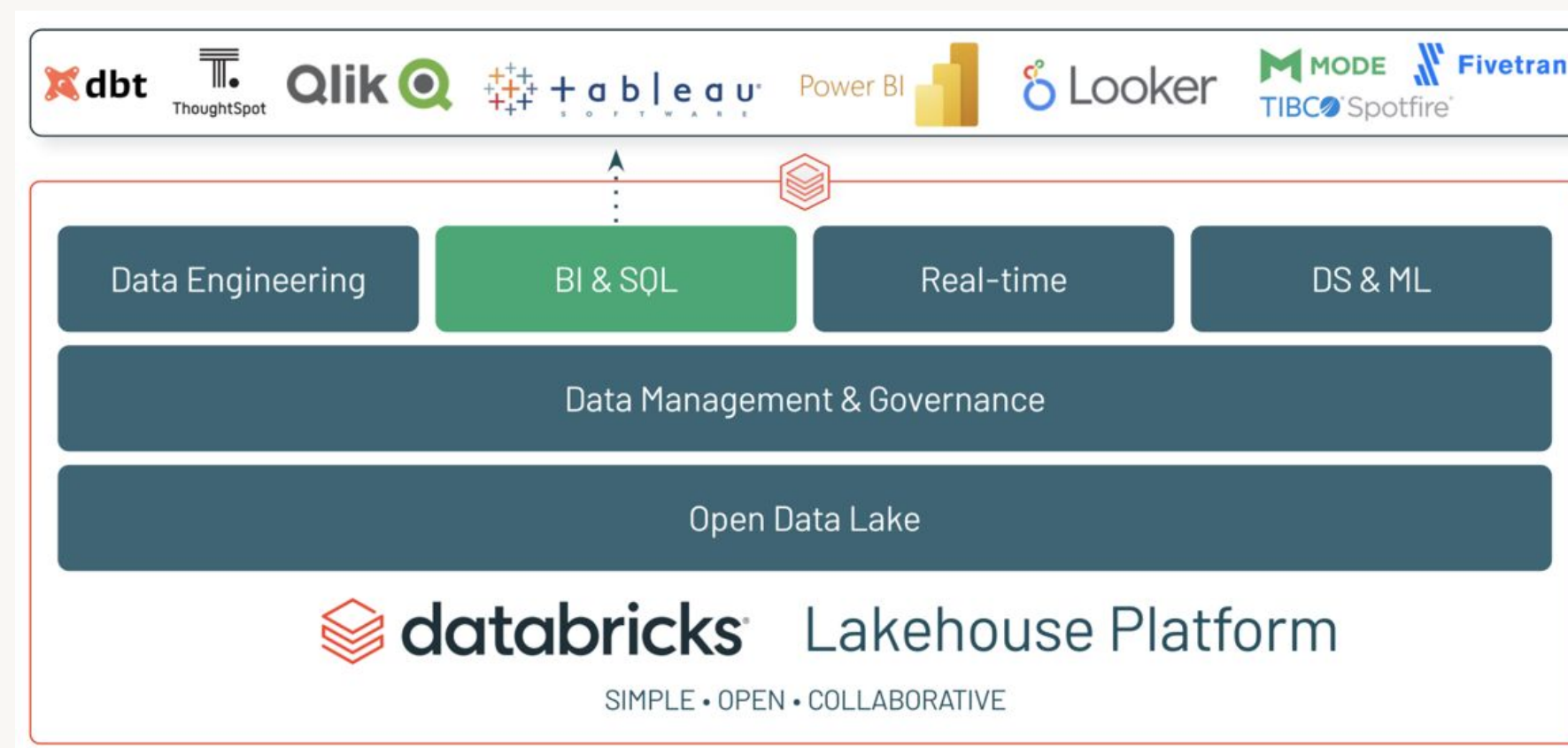
DE 7.2 – Last Mile ETL with Databricks SQL

Introduction to Databricks SQL

Databricks SQL

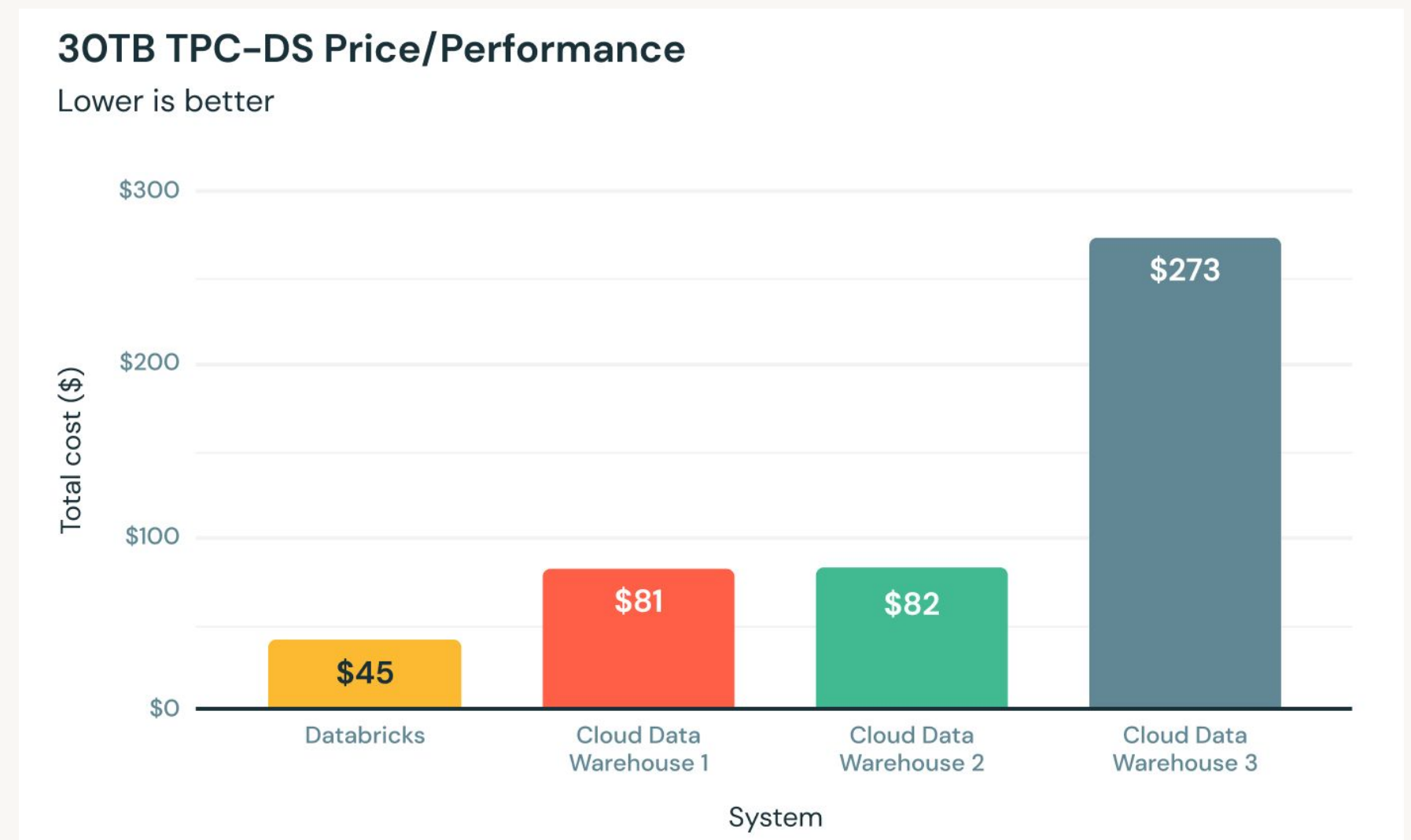
Delivering analytics on the freshest data with data warehouse performance and data lake economics

- Better price/performance than other cloud data warehouses
- Simplify discovery and sharing of new insights
- Connect to familiar BI tools, like Tableau or Power BI
- Simplified administration and governance



Better price / performance

Run SQL queries on your lakehouse and analyze your freshest data with **up to 6x better price/performance** than traditional cloud data warehouses.



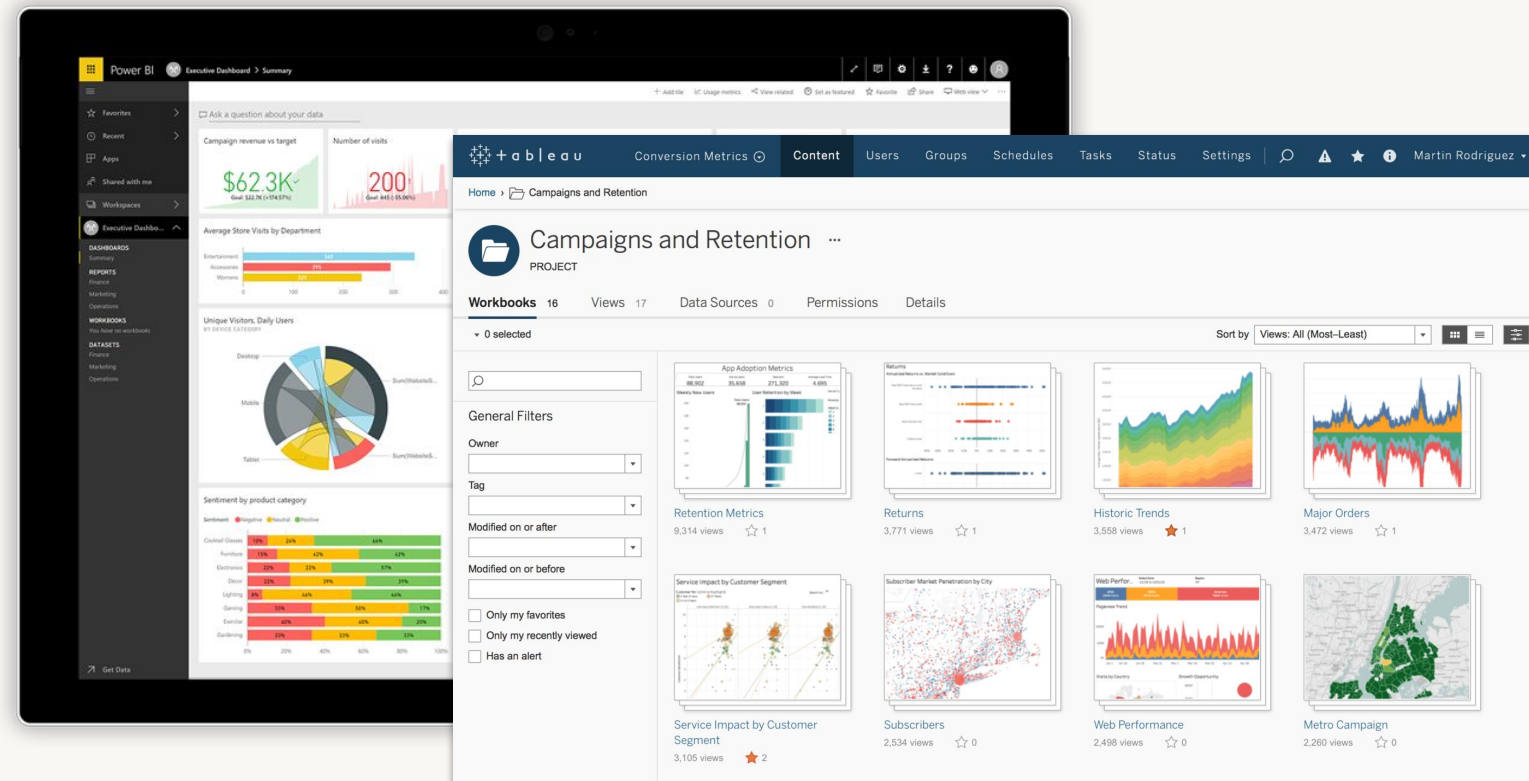
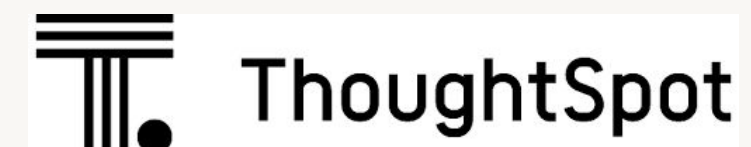
Source: Performance Benchmark with Barcelona Supercomputing Center

Broad integration with BI tools

Connect your preferred BI tools with optimized connectors that provide fast performance, low latency, and high user concurrency to your data lake for your existing BI tools.

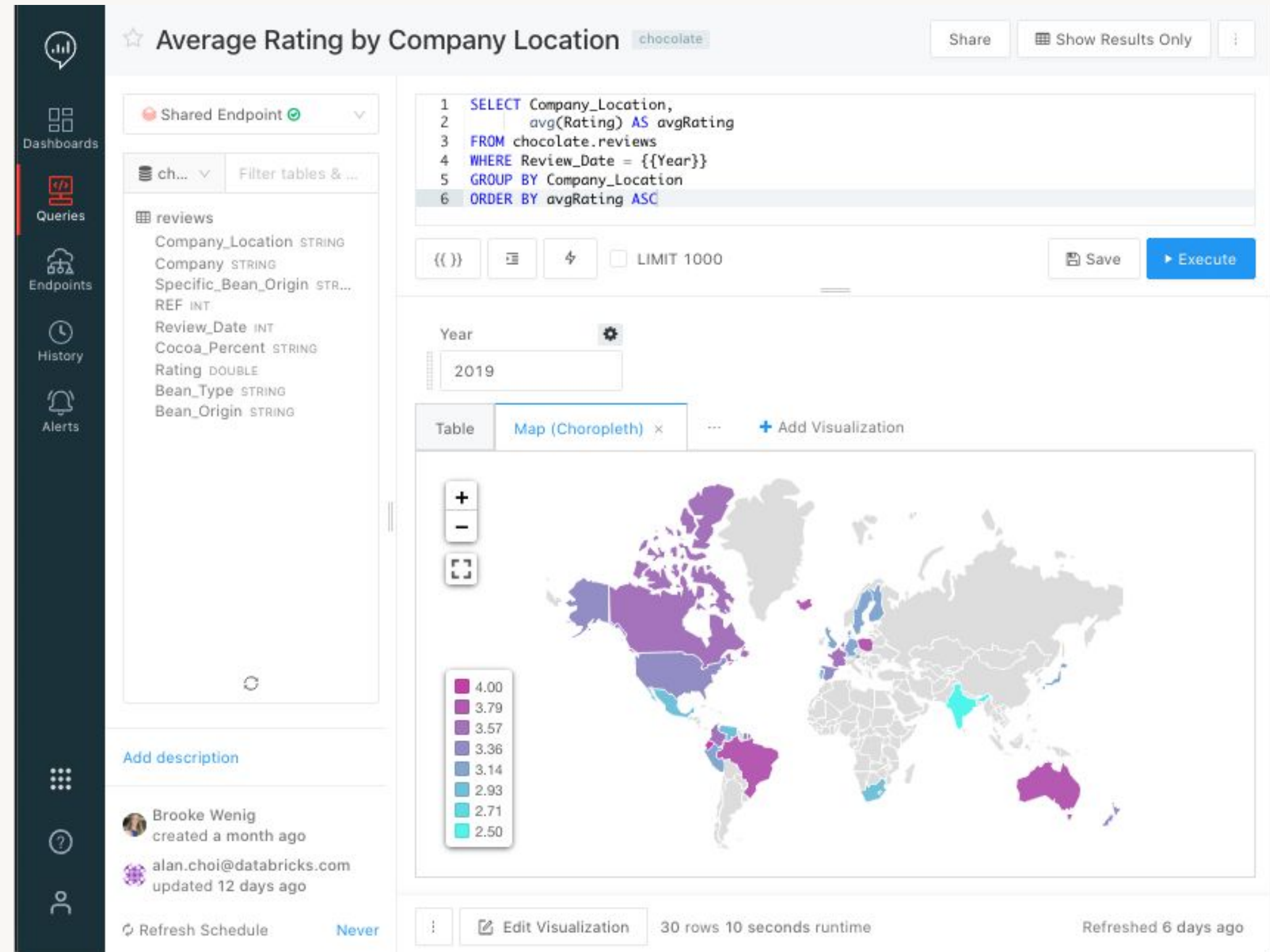


Coming soon:



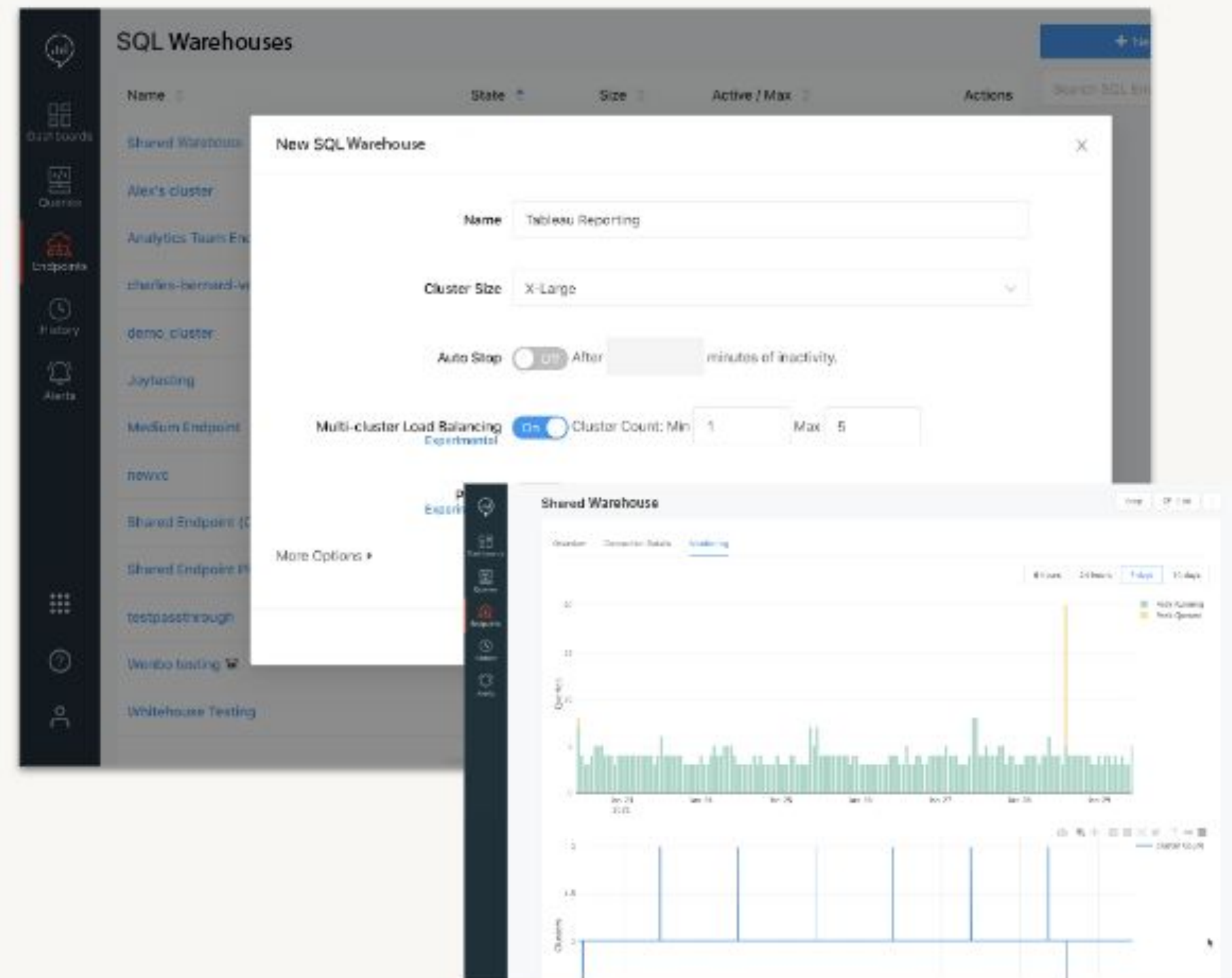
A new home for data analysts

Enable data analysts to quickly **perform ad-hoc and exploratory data analysis**, with a new SQL query editor, visualizations and dashboards. Automatic alerts can be triggered for critical changes, allowing to respond to business needs faster.



Simple administration and governance

Quickly setup SQL / BI optimized compute with SQL warehouses. Databricks automatically determines instance types and configuration for the best price/performance. Then, easily manage usage, perform quick auditing, and troubleshooting with query history.

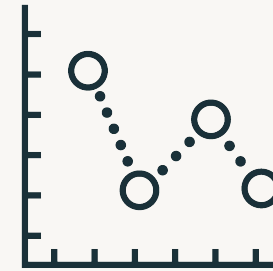


Use Cases



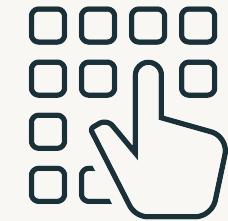
Connect existing BI tools to one source of truth for all your data

Maximize existing investments by connecting your preferred BI tools to your data lake with Databricks SQL Warehouses. Re-engineered and optimized connectors ensure fast performance, low latency, and high user concurrency to your data lake. Now analysts can use the best tool for the job on one single source of truth for your data while minimizing more ETL and data silos.



Collaboratively explore the latest and freshest data

Respond to business needs faster with a self-served experience designed for every analysts in your organization. Databricks SQL Analytics provides a simple and secure access to data, ability to create or reuse SQL queries to analyze the data that sits directly on your data lake, and quickly mock-up and iterate on visualizations and dashboards that fit best the business.

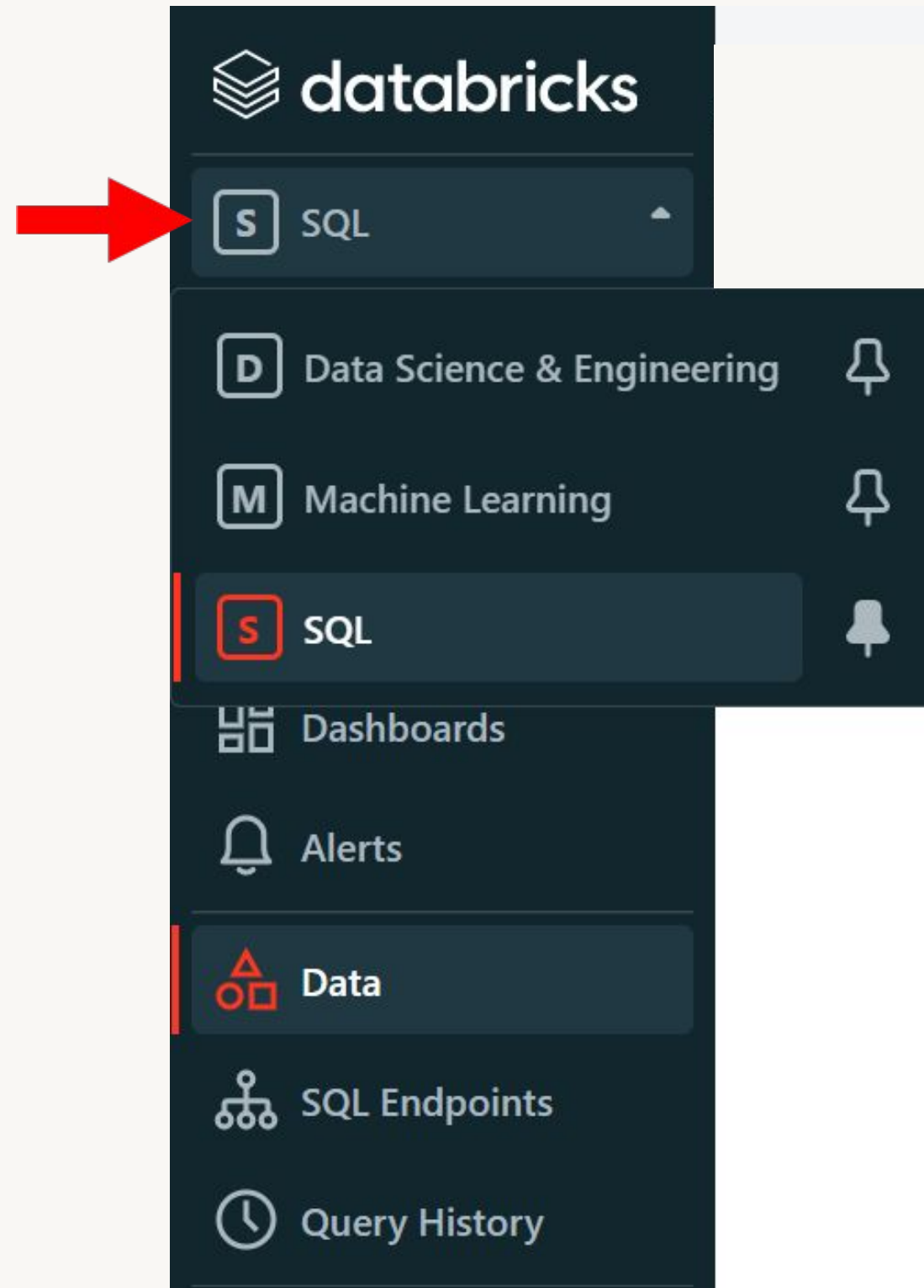


Build data-enhanced applications

Build rich and custom data enhanced applications for your own organization or your customers. Benefit from the ease of connectivity, management, and better price / performance of Databricks SQL Analytics to simplify development of data-enhanced applications at scale, all served from your data lake.



Navigating to Databricks SQL Persona



SQL Warehouse

An SQL warehouse is a computation resource that lets you run SQL commands on objects within Databricks SQL

To reduce latency, increase cluster size

Scale to handle more concurrent users

Serverless endpoints starts instantly

Tags for tracking in cloud provider

New features and release

New SQL Warehouse ⓘ

Name

mott whse

Cluster size ⓘ

Medium

24 DBU / h ▾

Auto stop



After

10

minutes of inactivity.

Scaling ⓘ

Min.

1

Max.

1

clusters (24 DBU)

Type



Serverless ⓘ



Pro ⓘ



Classic



Serverless SQL warehouses contain all advanced features and are Databricks' fastest warehouse type.

Prices are reduced (up to 40%) until Apr 30, 2023. Try a Serverless SQL warehouse today! [Learn more](#)

Advanced options ▾

Tags ⓘ

Key

Value

Channel ⓘ



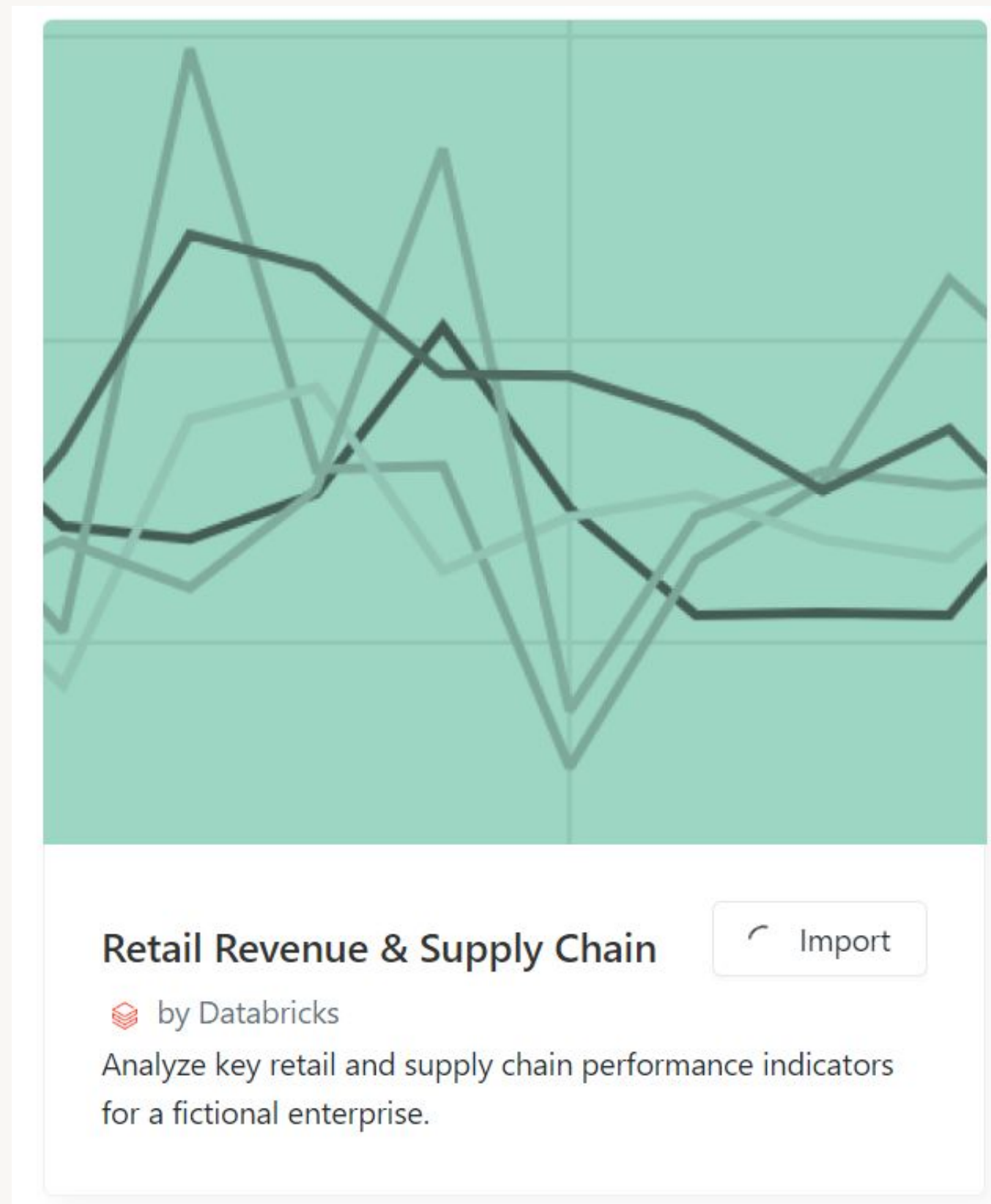
Current



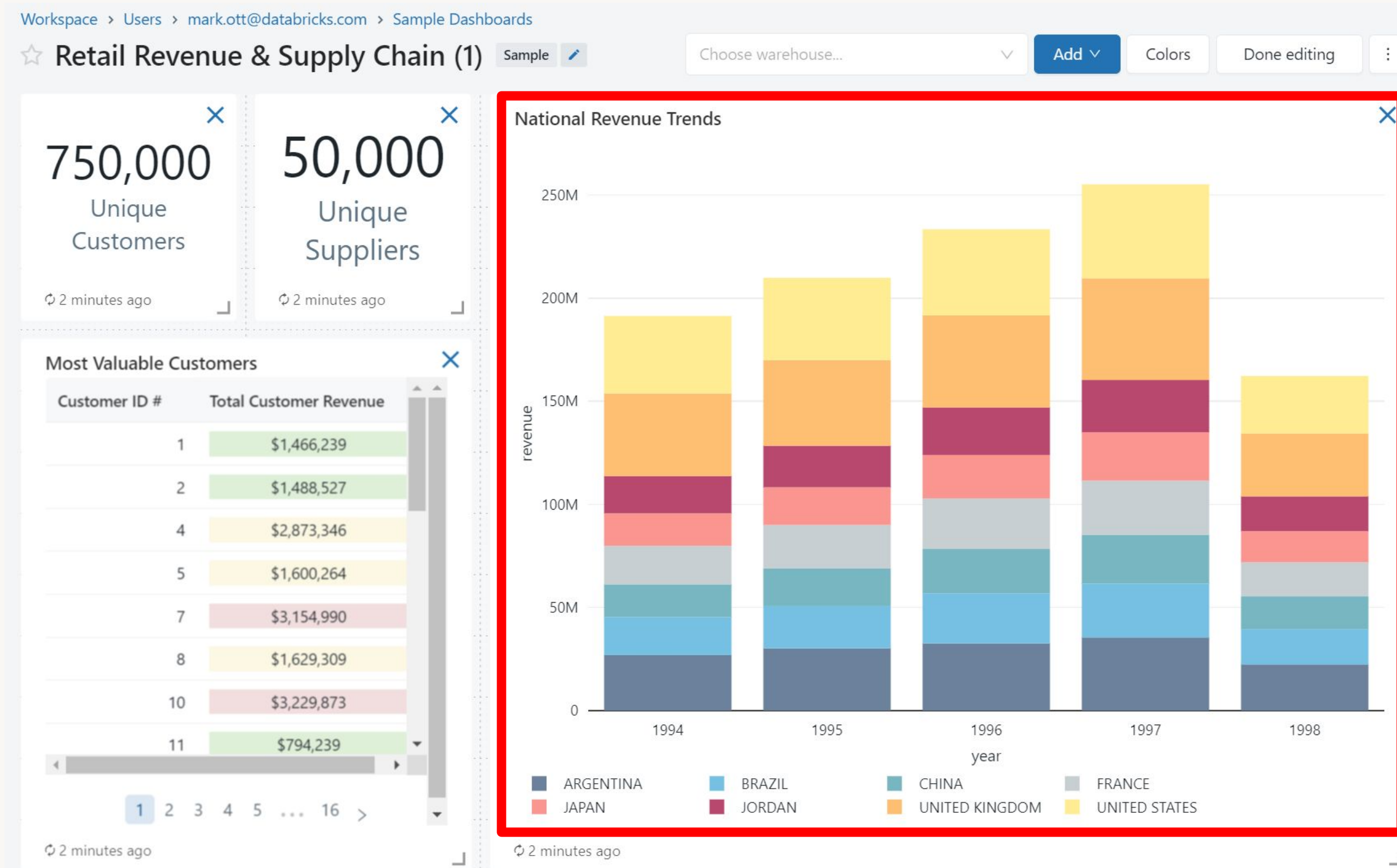
Preview



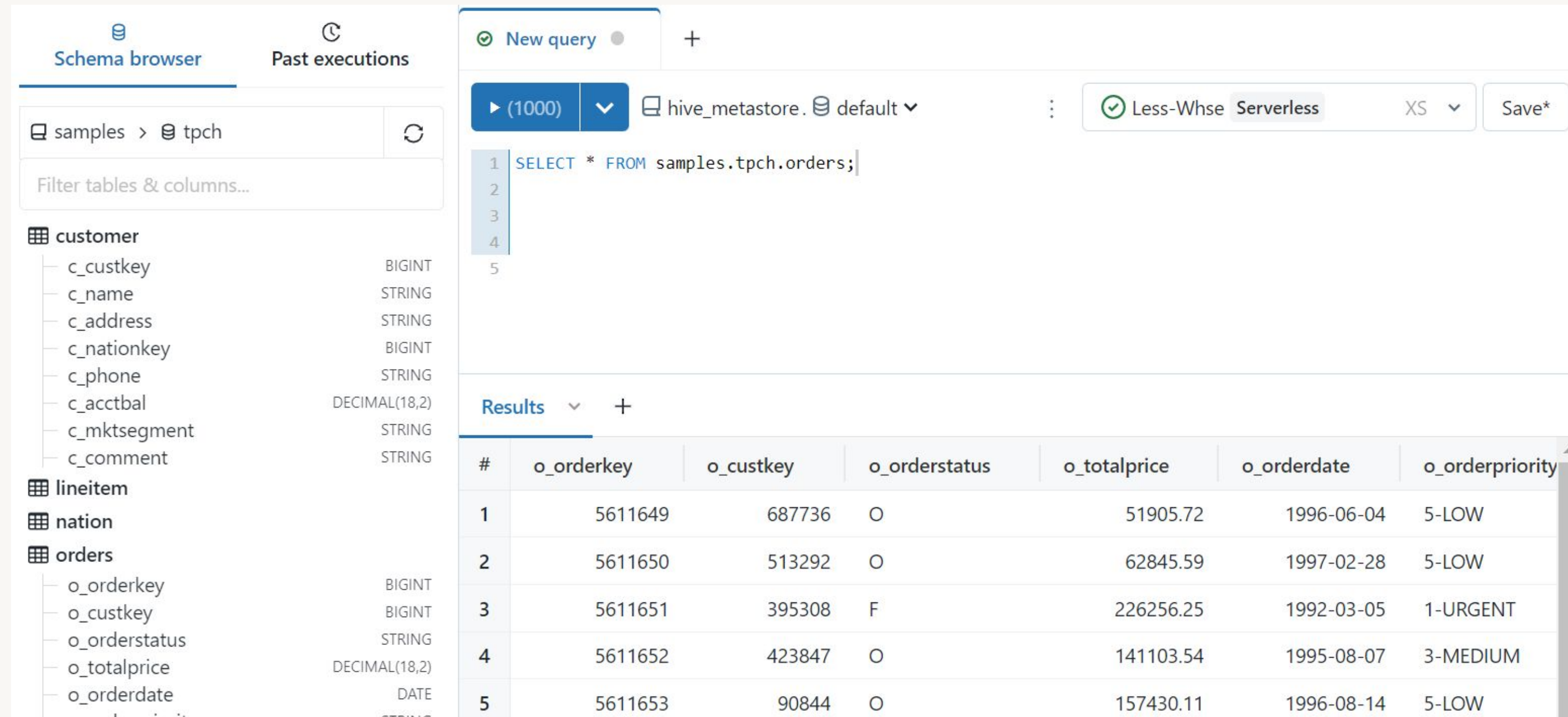
Import Existing Dashboard



Add Visualization to Dashboard



Run a SQL query



The screenshot displays the Databricks SQL interface. On the left, the 'Schema browser' shows a hierarchy of tables: 'customer', 'lineitem', 'nation', and 'orders'. The 'orders' table is selected, showing its columns: 'o_orderkey' (BIGINT), 'o_custkey' (BIGINT), 'o_orderstatus' (STRING), 'o_totalprice' (DECIMAL(18,2)), 'o_orderdate' (DATE), and 'o_orderpriority' (STRING). On the right, the 'New query' tab is active, showing a SQL query: 'SELECT * FROM samples.tpch.orders;'. Below the query editor, the 'Results' tab is active, displaying a table with 5 rows of order data.

#	o_orderkey	o_custkey	o_orderstatus	o_totalprice	o_orderdate	o_orderpriority
1	5611649	687736	O	51905.72	1996-06-04	5-LOW
2	5611650	513292	O	62845.59	1997-02-28	5-LOW
3	5611651	395308	F	226256.25	1992-03-05	1-URGENT
4	5611652	423847	O	141103.54	1995-08-07	3-MEDIUM
5	5611653	90844	O	157430.11	1996-08-14	5-LOW

Demo: Navigating Databricks SQL and Attaching to Warehouses

- Navigate to Databricks SQL
- Create SQL Warehouse
- Load a sample dashboard
- Attach to a SQL Warehouse

Lab: Run a Databricks SQL Query

- Create a sample query
- Execute sample query
- Add a visualization
- Modify query and rerun
- Create a new query

Lab: Last Mile ETL with Databricks SQL

- Create Queries
- Set a Query Refresh Schedule
- Create a Bar Graph Visualization
- Create a New Dashboard
- Add a Line Plot Visualization to your Dashboard
- Create a Query to Report Summary Statistics
- Add the Summary Table to your Dashboard
- Review and Refresh your Dashboard
- Share your Dashboard
- Set Up an Alert and Review Alert Destination Options

