dremio

QUICK GUIDE TO THE
APACHE ICEBERG LAKEHOUSE

Data Lakehouses eliminate the need to move data between a lake and a warehouse by turning your data lake into your data warehouse reducing the risk of inconsistencies, delays and operational bottlenecks from multiple data copies and excessive data movement.

At the heart of a data lakehouse are several composable components:

- **STORAGE SYSTEM**
- **FILE FORMAT**
- **TABLE FORMAT**
- **CATALOG**
- **QUERY ENGINE**

WHY APACHE ICEBERG FOR THE LAKEHOUSE?

Apache Iceberg stands out for its powerful and unique features that go beyond what other table formats.

FEATURES

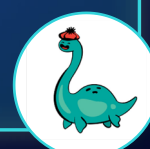**Partition Evolution** allows you to change partitioning schemes over time without rewriting the table

Catalog integration, offering enterprise-grade features via tools like Apache Polaris (incubating)

Table Level Branching, which enables version control for your datasets

With **Project Nessie**, Iceberg supports catalog-level versioning

ANOTHER MAJOR REASON TO CHOOSE APACHE ICEBERG IS ITS RICH AND EXPANDING ECOSYSTEM.

- Flink
- Apache Spark
- dremio
- presto
- trino
- Apache Airflow

As an Apache Project, Iceberg is built with a strong commitment to transparent, community-driven development.