



# Yellowbrick

## Stacking up versus Presto





## Yellowbrick Data Warehouse Overview

The Yellowbrick Data Warehouse delivers a modern data warehouse experience in your own cloud account as well as on-premises. We support all the benefits of truly cloud-native data warehousing – separate storage and compute and built-in elasticity, managed through SQL – without your data ever needing to leave your VPC or data center.

Yellowbrick customers are the world's largest global enterprises, running ad-hoc analytics, and operational data warehouses supporting business-critical workloads with high concurrency. Wherever you run Yellowbrick, you'll find rich workload management, real-time data ingest, the ability to load and query data together on the same instance, high availability, and replication across clouds and on-premises for disaster recovery. You can say “goodbye” to the instability of Hadoop and “hello” to a stable, reliable, and trustworthy data warehouse.

Pricing with Yellowbrick is open, simple, and predictable. We've saved many customers millions of dollars per year in legacy and cloud spend. You use and pay for your storage and compute, making use of your cloud credits without paying

smaller companies to mark up your infrastructure costs. You can be comfortable meeting your security, regulatory, and compliance requirements.

If you're considering a new migration to the cloud, Yellowbrick avoids lock-in to any particular cloud vendor or database. We use open-source PostgreSQL as our SQL syntax of choice and are backward compatible with on-premises ETL tools like Informatica PowerCenter and CDC tools, Oracle GoldenGate or Qlik (Attunity) Replicate. We don't do lock-in. We don't want to be the next Oracle (nor did our founders come from there) and we use open standards to ensure the portability of your data and workload.

We're proud to have the best performance in the industry, at the lowest possible cost. No one runs data warehousing workloads faster than Yellowbrick.

Keep reading to see how we stack up versus our competitors in the market.



# Presto

The Yellowbrick Data Warehouse fits easily into your Data Lake architecture, without needing to replace it. Most importantly, Yellowbrick is an enterprise-quality database. We have enterprise-class stability; you don't need to manage storage or file layouts or partitions; we have built-in multi-site replication, transactional consistency, and the ability to plan complex SQL queries, run stored procedures, role-based access control; and all the other things you'd expect from a "database."

Yellowbrick supports high concurrency and low-latency interactive queries. Badly written run-away queries can't disrupt other users' queries, and you can have thousands of users querying your data

at the same time. Yellowbrick guarantees correct results, isn't written in Java – a platform choice that leads to fundamentally unpredictable performance – and doesn't periodically crash. Furthermore, it interoperates with all of your existing ETL. At one of the world's largest logistics companies, a 15-node Yellowbrick cluster ran ad-hoc, interactive queries 30x faster than a 120-node Presto cluster.

If you have a need for operational data warehousing, high concurrency, ad-hoc queries, predictable response times, or simply software that doesn't crash all the time, Presto – like other SQL-on-Hadoop products – will do nothing for you. Follow in the footsteps of the world's largest insurance, logistics, credit card, and risk companies and use Yellowbrick to enable business users to get value out of your Data Lake.

**At one of the world's largest logistics companies, a 15-node Yellowbrick cluster ran ad-hoc, interactive queries 30x faster than a 120-node Presto cluster.**

	Yellowbrick	Presto
Is it a database?	✓	✗
Enterprise-class stability	✓	✗
Built-in multi-site replication	✓	✗
Fast inserts, updates, deletes	✓	✗
Autonomous without indexing/tuning	✓	✗
Hardware-specific, native code performance optimizations	✓	✗
Cloud data warehouse in your VPC	✓	✓
SaaS cloud data warehouse (data elsewhere)	✓	✓
Elastic with separate storage/compute	✓	✓
Kubernetes cloud-native architecture	✓	✓
High concurrency, predictable latency for operational workloads	✓	✗
Advanced Workload management	✓	✗
Real-time streaming data ingest	✓	✗

