Yellowbrick

USE CASE DETAILS

U.S. Health Insurance Company

Headquartered on the East Coast of the U.S. is the country's second largest integrated health care delivery and financing network based on revenue. This company and its subsidiaries and affiliates provide health insurance to nearly 5 million members in mid-Atlantic states, as well as dental insurance, vision care and related health products through a national network of other businesses.

The Company's Analytics Environment

The insurance company has a 7-node Hortonworks data lake comprising hundreds of terabytes of data that the company wants to use for self-service analytics. In addition, the company has a Teradata environment against which it runs SAS, and the company recently purchased three Exadatas from Oracle. The insurance company uses its analytics environment for the following:

- **Ad hoc queries:** Ad hoc queries are used for examining policy-writing and claims trends. Actuaries use this information daily to make decisions on writing policies and addressing claims.
- **Reporting:** The business uses SAS for critical reporting against the Teradata system; over 1200 users use the tool and often write complex SQL queries.
- **Month-end close:** The company uses its systems to measure and understand overall risk and examine historical loss and exposure trends. Month-end analysis is vital to making sound decisions that affect company exposure.

Use Case Examples and POC Results

The company wanted to provide self-service analytics against their existing Hortonworks data lake but could not get a high enough level of performance. The company had also considered shifting some Teradata-based queries to the data lake, but its users balked at the loss of performance. The insurance company tested Yellowbrick's ability to support a high-concurrency set of users and streaming data updates. The proof-of-concept (POC)'s results included:

- **Ad hoc queries:** With no tuning, the Yellowbrick system ran the company's SQL queries between 2X and 80X faster than Teradata using the company's existing tool set with full compatibility
- **Concurrency:** Yellowbrick easily handled over 1000 concurrent users running mixed workloads, at the same time automatically sequencing fast and complex queries for optimal performance
- **Simple extracts:** Yellowbrick accepted high-speed extracts from Teradata and data streams from the Hortonworks data lake