

# AI Agent for Google BigQuery

## Revefi's Unified FinOps, DataOps, and Data Observability

By combining cost control, operational resilience, and data trust, Revefi transforms BigQuery into a strategic asset rather than an unpredictable expense line item.

- Eliminate wasted BigQuery spend, identify inefficient queries and anomalies, and deliver AI-driven optimization recommendations that cut cloud data costs
- Detect schema drift, stale tables, and performance issues early to get instant root-cause analysis (RCA) and quality alerts
- Correlate cost, performance, and data health in one view to enforce proactive, automated governance without slowing innovation

## How Can Organizations Leverage Revefi?

Revefi retrieves metadata from Google BigQuery in minutes, requiring no code changes or complex integrations. It continuously monitors query patterns, slot usage, table access, and cost drivers at a granular level, delivering real-time visibility across projects, datasets, and teams.

### 1. Automated Cost Management, and Optimization

Revefi delivers comprehensive slot and cost visibility across BigQuery workloads by analyzing query activity, bytes processed, and reservation slot usage. The platform identifies expensive query patterns, inefficient pipelines, and unused resources while providing optimization insights.

#### Key capabilities:

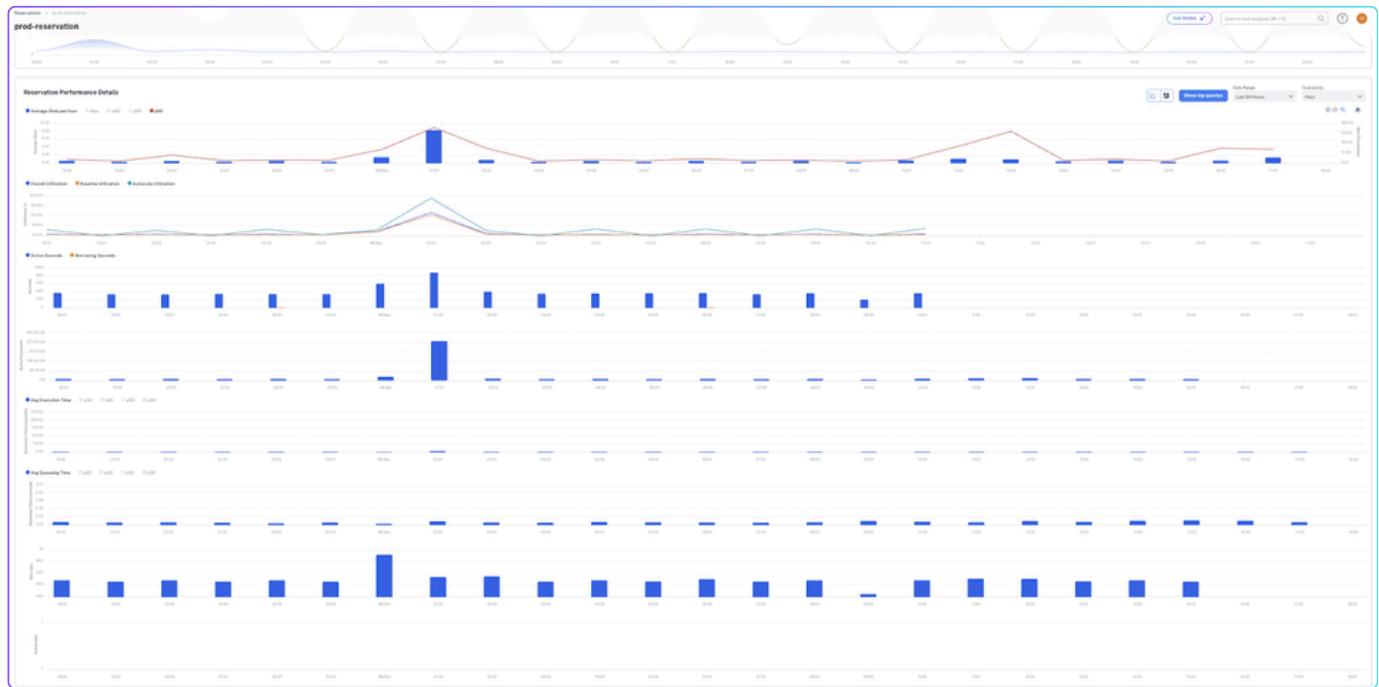
- Unified visibility into BigQuery slot consumption patterns and cost drivers
- Identification of high-cost queries by slot-ms usage, bytes processed, and runtime
- Automatic comparison of utilization across reservations (dev, prod, QA)



## 2. Automated Performance and Efficiency Insights

Using metadata from and job execution logs, Revefi surfaces trends in execution time, queue time, job concurrency, throughput and more. This allows data teams to diagnose bottlenecks such as:

- Overloaded reservations
- Queries with unusually high slot consumption
- Long-running operations

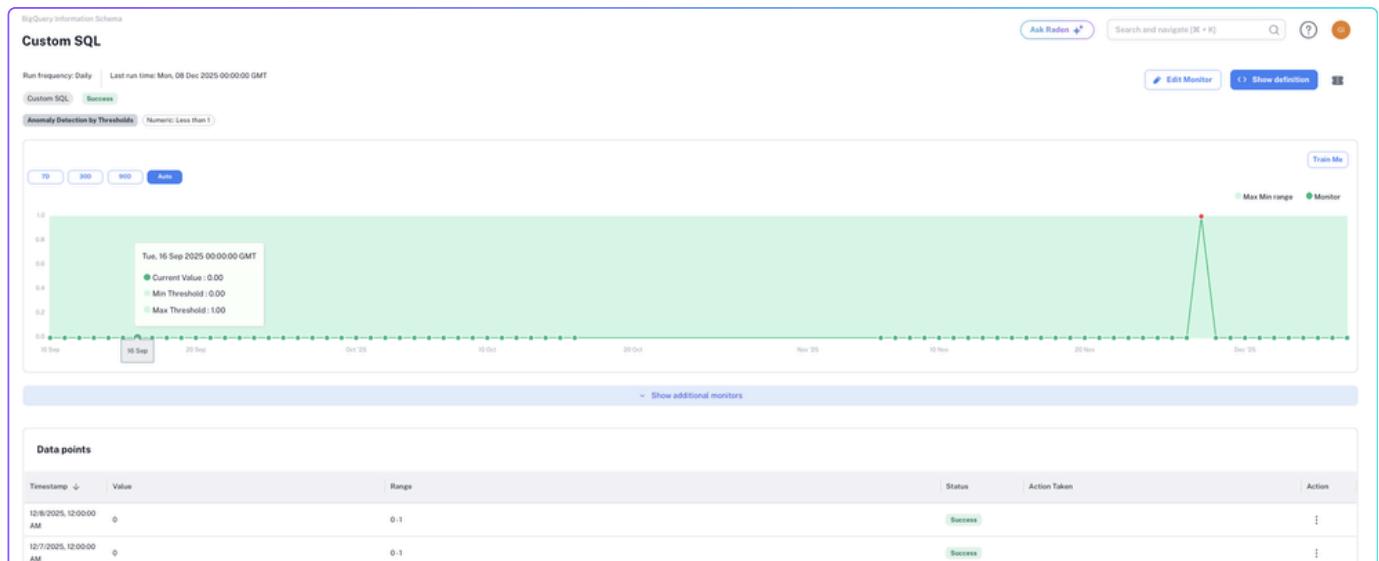


## 3. Automated Data Quality, and Observability

Revefi automatically generates and maintains data quality monitors across every table in an organization's BigQuery environment. These monitors track schema drift, data freshness, load success/failure, row count anomalies, and threshold-based custom SQL rules.

### Our platform provides:

- Daily checks (see image below)
- Detailed trend lines for each monitor
- Custom SQL monitors to enable high-sensitivity anomaly detection



## 4. Granular Data Usage, and Access Intelligence

Revefi provides granular insights into BigQuery compute costs, broken down to individual users, tables, and workloads. This makes it possible to understand how data assets are being accessed, which workloads are responsible for cost spikes, and which teams are under or over-utilizing compute resources.

User-level, and job-level breakdowns include:

- Slots consumed
- Bytes processed
- Runtime and queue time
- Behavior analysis

The screenshot shows the Revefi Usage dashboard with the following sections:

- Total Slots Cost (USD):** Last 24 hrs: \$1.63, Last 7 days: \$11.18, Last 30 days: \$50.97.
- Total Queries:** Last 24 hrs: 19 k, Last 7 days: 150 k, Last 30 days: 648 k.
- Total Bytes Processed:** Last 24 hrs: 516.4 GB, Last 7 days: 4.7 TB, Last 30 days: 22.7 TB.
- Total Storage:** Data Size: 2.34 TB, Annual cost (approx): \$291.03.
- Annual Saving Potential:** Storage Saving Potential: \$64.30, Compute Saving Potential: \$660.07.
- Query Time Distribution:** in last 30 days. Total Query time: 249.37 hours. Breakdown: Total Queued Time (28h 24m 91%), Total Query Execution Time (9d 5h 89%), Others (0.00 hours 0%).
- Overview by Project:** A table showing project details and costs for four projects: shasta-403921, shasta-qa-452921, shasta-dev-638554, and bigquery-sample-data-3... The table includes columns for Last 24h Query Count, Last 24h Cost (USD), Last 7d Query Count, Last 7d Cost (USD), Last 30d Query Count, Last 30d Cost (USD), and Action.

## 5. Pipeline and Orchestration Observability

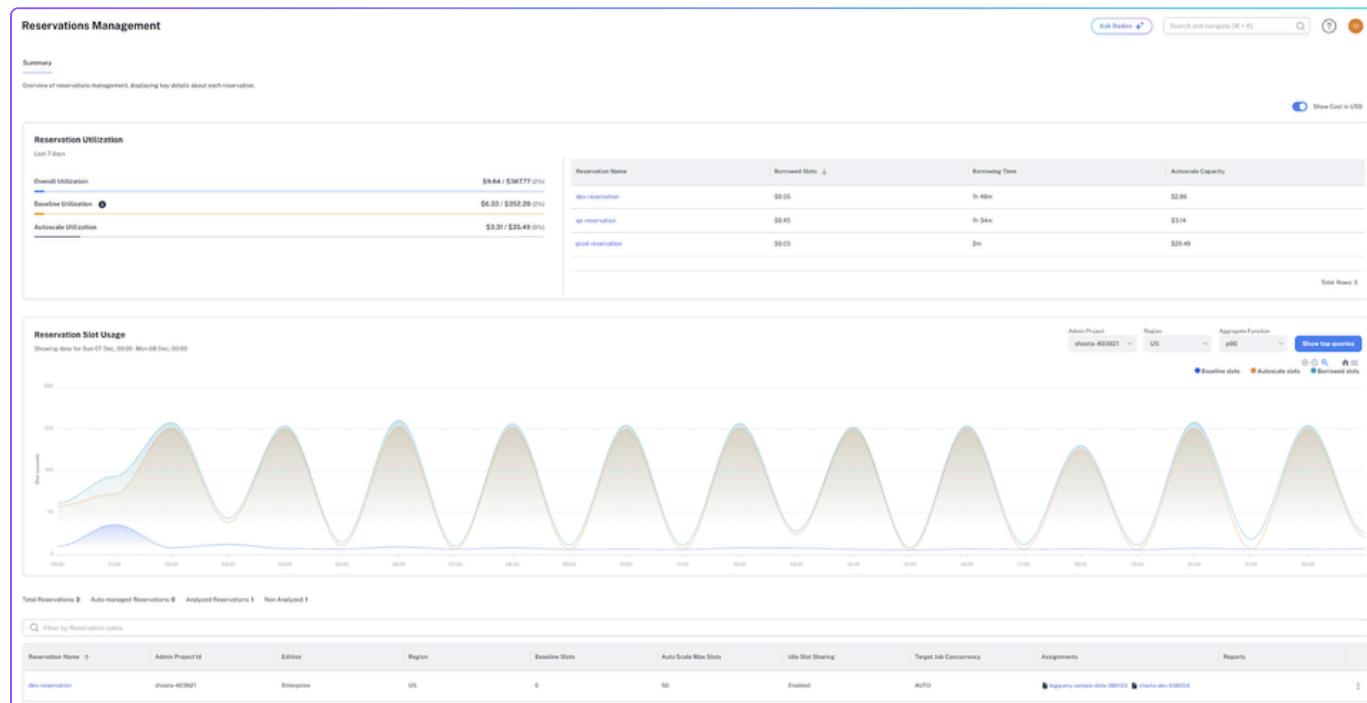
Revefi integrates with orchestration systems such as Cloud Composer (Airflow). It monitors DAG state, execution duration, queueing patterns, and success/failure trends. Engineers gain a unified view of operational pipeline health across BigQuery and workflow systems.

The screenshot shows the Revefi DAG monitoring dashboard for the DAG Name: airflow\_monitoring. It includes the following sections:

- DAG Details:** Monitors, Get Source Code, Open Airflow UI.
- DAG Metadata:** Name: airflow\_monitoring, Description:iveness monitoring dag, Status: Unknown, Owner: airflow.
- Execution Information:** A table showing execution counts and durations for Last 1 day, Last 7 days, and Last 30 days.
- Daily Summary - Last 7 days:** A line chart showing the distribution of execution durations over the last 7 days. The Y-axis represents the number of runs (0 to 800) and the X-axis represents dates from 08 Dec to 15 Dec. The chart shows a significant peak on 09 Dec.
- DAG Runs - Last 7 days:** A table listing the last 7 DAG runs with columns for Run ID, Status, Started On, Completed On, Execution Duration, Queued Duration, Run Type, and External Trigger.

## 6. BigQuery Table Health, Lineage, and Metadata Intelligence

Table-level dashboards show daily incremental row counts, load durations, total load bytes, and schema metadata. Revefi's lineage visualization reconstructs upstream and downstream dependencies using BigQuery job history, enabling impact analysis during failures or schema changes.



## Put Google BigQuery on Autopilot with Revefi's Autonomous AI

With Revefi, data teams now eliminate cost surprises, are able to debug slow dashboards, reverse-engineer why data pipelines silently broke, continuously for every query, slot, table, and job, and schema change across all projects.

## Integrations



## Organizations That Trust Revefi

