

Migration Assessment Report

BigQuery → Databricks Lakehouse · Prepared for: Meridian Retail Group
June 2026 · Engagement MRG-2026-014 · Confidential

△ SAMPLE DOCUMENT — "Meridian Retail Group" is fictional and all figures are illustrative. This sample shows the exact structure and depth of evidence every Replatform client receives.

1. Executive Summary

Meridian Retail Group runs a 14-person data team on Google BigQuery supporting analytics for 2,300 employees. Rising costs (~\$62k/month, growing 4% monthly) and a stalled ML roadmap prompted this assessment of a move to Databricks on the existing GCP footprint.

Finding	Number
Assets in scope (tables, views, scheduled queries, UDFs)	3,847 (31% unused 90+ days — candidates to retire, not migrate)
Estimated 3-year TCO saving vs. status quo	\$610k – \$890k (32–41%)
Recommended timeline to full cutover	16 weeks, 4 waves, pilot in week 3

Bottom line: this migration is economically justified primarily by workload right-sizing and the retirement of unused assets, secondarily by compute pricing. We recommend proceeding, starting with the Wave-0 pilot scoped in Section 5. Fixed-fee execution price: Section 8.

2. Scope & Method

Inputs analyzed (read-only): INFORMATION_SCHEMA across 6 projects, 90 days of query logs (412,089 jobs), scheduled-query definitions, 214 UDF/procedure bodies, current GCP billing export.

Method: the Replatform Engine built the full inventory and dependency graph from metadata and query logs; complexity scoring per object; translation feasibility tested by converting a 5% stratified sample (192 objects) and validating results; TCO modeled from actual usage, not list prices.

No row-level customer data was accessed at any point.

3. Inventory & Usage

Object type	Count	Unused 90d+	Note
Tables	2,210	744 (34%)	1.8 PB total; 612 TB attributable to unused tables
Views	1,184	381 (32%)	Heavy nesting in finance marts (max depth 9)

Scheduled queries	239	41 (17%)	Map to Databricks Workflows / DLT
UDFs & procedures	214	37 (17%)	Highest manual-effort category

Cold-asset finding: retiring rather than migrating the unused 31% reduces migration scope by ≈5 weeks and saves ≈\$9.4k/month in storage/maintenance immediately — value that exists whether or not you proceed with the migration.

4. Complexity & Automation Forecast

Tier	Share	What it means
LOW — auto-convertible	61%	Vanilla DDL/SQL; pipeline converts and validates with engineer spot-checks
MEDIUM — assisted	31%	Dialect-specific functions (SAFE_CAST, FORMAT_DATE, UNNEST patterns); pipeline converts, engineer reviews every diff
HIGH — engineer-led	8%	Legacy SQL, complex UDFs, _PARTITIONTIME dependencies, ARRAY_AGG ordering semantics; engineers translate with pipeline assistance

Sample-conversion results on the 192-object test set: 89% of LOW and 71% of MEDIUM objects passed deterministic validation on first pass. These rates set the effort model behind the fixed fee — we price from measured rates, not optimism.

5. Recommended Wave Plan

Wave	Content	Duration	Exit criterion
0 — Pilot	Marketing mart: 118 tables, 64 views, 9 scheduled queries	3 weeks	100% reconciliation evidence; client signoff
1	Product & ops marts + shared dims	4 weeks	Dual-run parity for 2 weeks
2	Finance (highest complexity, audited)	5 weeks	Checksum + query-diff parity; finance signoff
3	Long tail + BI repoint + decommission	4 weeks	Source read-traffic = 0; cutover approved

6. 3-Year TCO Comparison

Cost line	Status quo (BQ)	Target (Databricks)	Delta
Compute (measured workload)	\$1.61M	\$1.04M – \$1.18M	–27% to –35%
Storage (after cold-asset retirement)	\$402k	\$214k	–47%

Platform/licence & support	\$0	\$108k	+\$108k
Migration cost (Section 8) amortized	—	\$184k	one-time
3-year total	\$2.01M	\$1.12M – \$1.40M	–\$610k to –\$890k

Ranges reflect workload-tuning uncertainty; assumptions in Appendix A. We model from your actual 90-day usage, escalated 4%/month for growth.

7. Validation: How You'll Know Nothing Broke

- Row counts and aggregate checksums (SUM/MIN/MAX/COUNT DISTINCT on keys and measures) for every migrated table — delivered as a signed evidence report per wave.
- Query-result diffing: your top 25 production queries by usage, executed on both platforms, results compared with float-tolerant, ordering-normalized diffs.
- Two-week dual-run before each cutover; finance wave gets four.

All validation is independent, deterministic code — conversion and verification are strictly separated. Evidence artifacts are yours to keep and audit.

8. Risks (top 5 of 14 — full register in workbook)

Risk	Severity	Mitigation
ARRAY_AGG ordering semantics differ in 23 finance views	High	Engineer-led translation + targeted query-diff tests (already scoped)
_PARTITIONTIME dependencies in 61 tables	High	Explicit partition columns added at ingestion; design doc in week 1
Looker dashboards (148) need repointing	Medium	Wave-3 workstream; connection swap tested in pilot
Concurrent BQ usage during dual-run doubles cost briefly	Medium	Budgeted in TCO; capped at 6 weeks total
Key-person dependency: 2 client engineers own finance logic	Medium	Their review time contractually reserved (4 hrs/week)

9. Fixed-Fee Proposal

Item	Fee	Terms
Assessment (this document)	\$12,000	Paid; 100% credited against execution if signed within 60 days
Execution — Waves 0–3	\$172,000	Milestone-billed 30/30/30/10 at wave exits

30-day hypercare	Included	—
Optional: FinOps retainer post-migration	\$6,500/month	Cancel anytime after month 3

Price protection: this fee is fixed against the scope in Section 5. Scope changes are priced from the same measured-rate model before any work proceeds — never invoiced retroactively.

Appendix A — Assumptions

- Databricks pricing: premium tier on GCP, 1-year commit discount; compute sized from measured BQ slot-utilization equivalents.
- Growth: 4%/month query-volume growth applied to both scenarios equally.
- Client engineering availability: 2 named engineers, 4 hrs/week each, for domain review.
- Cold-asset retirement list requires client approval; figures assume 90% of identified candidates are retired.
- Exchange/list prices as of June 2026; provider funding (Databricks migration co-funding) not yet applied — it can reduce the execution fee and will be pursued on signature.