

From Data Infrastructure to Customer Intelligence: Analytics Value Delivered for a Global Retail GCC

75%

Cloud Cost
Reduction

40%

Pipeline Run
time Saved

3

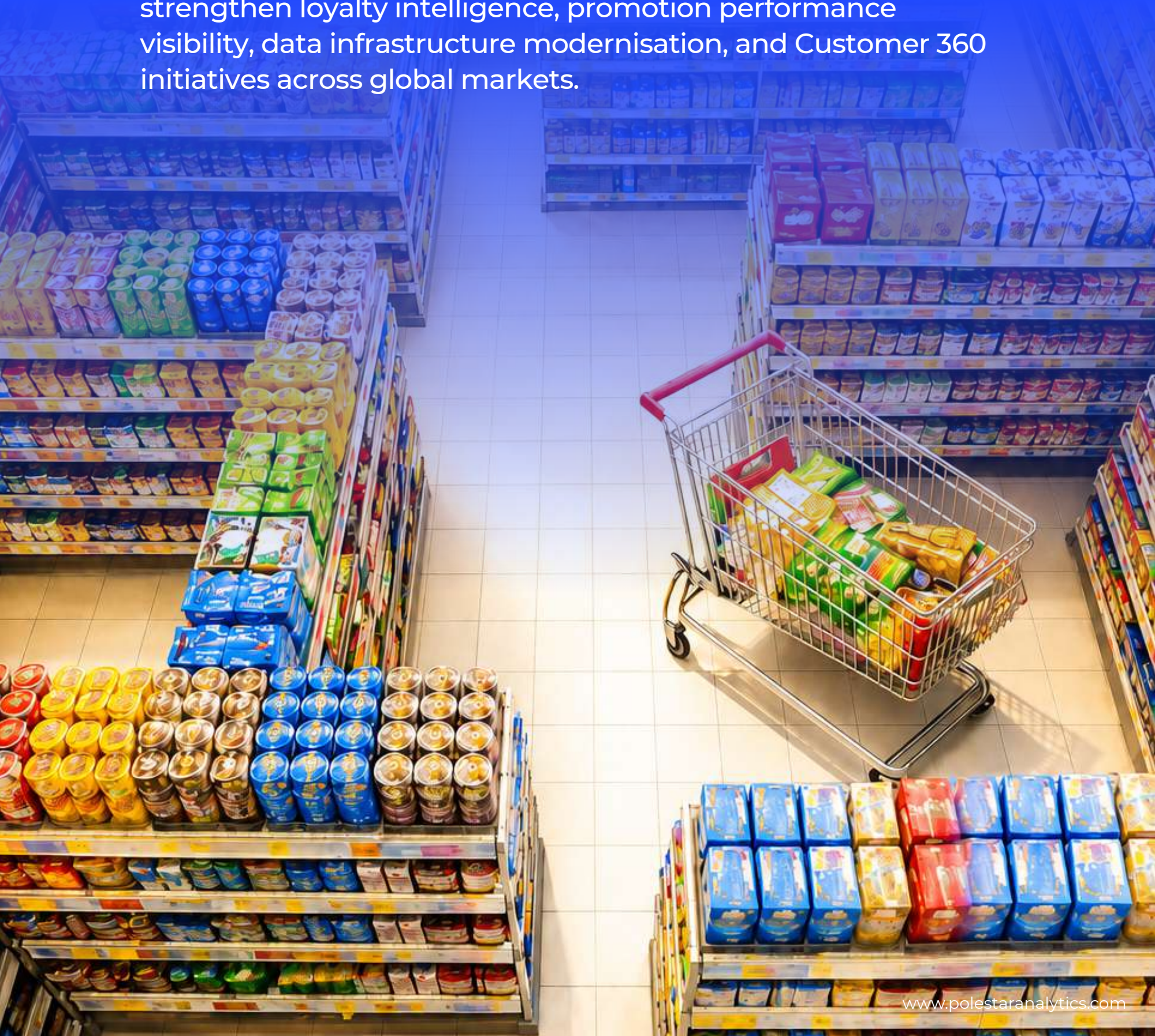
Practice Areas
Delivered



Client Overview

The client is a **global retail enterprise** with a GCC focused on data, analytics, and ML operations for global markets. **The GCC supports three core practice areas:** Customer Loyalty & Management, Promotion Analytics, and MLOps & Data Engineering.

Polestar Analytics works as an embedded analytics partner with the client's customer analytics function, helping strengthen loyalty intelligence, promotion performance visibility, data infrastructure modernisation, and Customer 360 initiatives across global markets.



The Engagement



GCC Location

Gurgaon, India

Focused on data, analytics & ML operations for global markets



Platform

Databricks (primary)

Transitioning to Snowflake

Migrating from legacy Hive to Unity Catalog



Practice Areas

Customer Loyalty & Management

Promotion Analytics

MLOps & Data Engineering



Key Stakeholder

Director, Customer Analytics

Polestar Analytics team: 4 embedded analysts + MLOps

Contribution #1

A2K: Polestar Analytics Self-Led Initiative Turning Anonymous Shoppers into Known, High-Value Targets

1 Identify Non-Loyal Shoppers

Separated loyal transactions from total. Remaining = non-loyal pool. Used payment method linkage (Payment Account Reference Identifier(PAR IDs) in Europe, raw card-based token identification in North America) to construct unique customer fingerprints from anonymous transactions.



2 Score & Segment

Polestar Analytics devised a new methodology and logic integration to correct the major data leakage. Applied an Advanced, Composite Engagement Scoring Model. Built behavioral engagement score:

Active Weeks (40%) + Visits Per Week(30%) + Spend PerWeek (20%) + Consistency (10%).

Segmented the pure B2C base into statistical tiers using dynamic percentiles:

High(top 25%),Medium (mid 50%), and Low (bottom 25%)

3 Surface High-Value Targets

Delivered a ranked list of B2C non-loyal customers with high spend propensity — ready for loyalty programme outreach. A2K evolved from minor exploratory project to a major, ongoing GCC priority.



Delivered for both North America and Europe — with region-specific data models due to different payment infrastructure and regulatory frameworks.

Contribution #2

Promotions Analytics: Closing the Planning–Execution Gap

The Bottleneck

The client's retail store GMs plan promotions via separate dashboards. Once promotions ran, there was no single view to compare planned vs. actual outcomes.

Leadership had to toggle between multiple dashboards —with no unified way to learn and optimise future promotions.



What We Built

Unified Pre vs. Post promotion dashboard connecting planned parameters (expected sales, product mix, store targets) with actual results - curated backend tables feeding Power BI.



Better Decisions

Leadership can now compare planned vs. actual revenue at any store, any promotion



Future Optimisation

Insights surface what changed, what worked, and what to do differently next cycle

Contribution #3

MLOps & Data Infrastructure Modernisation



Unity Catalog Migration

Migrated legacy Hive-based data to Databricks Unity Catalog enabling table lineage tracking, access control, and governance at scale. Previous model: ad-hoc, teams deleting/mismanaging tables.



ML Model Versioning

Introduced MLflow for clustering and other models previously saved to local Hive meta store. Enabled version rollback and audit trail critical since model scores surface directly in the Extra 2.0 loyalty app.



CI/CD & Service Principals

Established GitHub + CI/CD pipeline for production deployments. Setup cloud-based service principals to remove individual engineer dependencies from production runs critical for operational resilience.



Pipeline Cost & Performance Optimisation

Migrated CLM pipelines from expensive serverless compute to job clusters. CLV historical backfill: cost reduced from **\$2,000 → \$200 per run (90% reduction)**.
Runtime: 17+ hours → 6 hours.

Contribution #4

Customer 360: Integrating External Intelligence

Enriching client's store-level and customer-level view with third-party data to power smarter decisions.

1

ChainXY - Competitive Intelligence

Integrated competitor-level data for each store. Enables leadership to understand the competitive landscape at a granular, store-by-store level across geographies.

2

PopStats - Demographic Context

Added demographic profiles around each store: household density, education levels, family composition. Helps tailor promotions and loyalty offers to local market characteristics.

3

Traffic Data - Footfall Potential

Outcome: A richer Customer 360 schema that connects store footfall, competitor pressure, and demographic reality — giving loyalty and marketing teams a multi-dimensional view unavailable before.

Quantified Business Impact

FLAGSHIP RESULT - Offer Model Pipeline Optimisation (Poland BU)

Applied right-sized dataset, data caching strategy & job cluster configuration to the optimised offer model pipeline.

\$45 → \$10
per pipeline run

1h 34m → 58m
runtime



75%

Lower
Cloud Cost

\$45 → \$10 per run
Offer Model
Pipeline (Poland
BU)



40%

Faster
Runtime

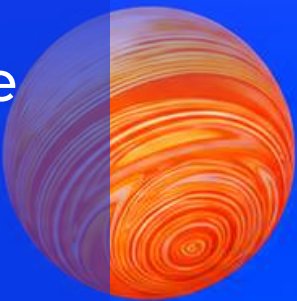
1h 34m → 58 min
Same Pipeline
Optimisation



90%

CLV Pipeline
Savings

\$2,000 → \$200
per run
Historical CLV
Backfill



What Polestar Analytics Delivered Across the GCC Analytics Journey



Embedded, Not Outsourced

Our team operates inside client's GCC as embedded partners - attending standups, working in their Databricks workspace, and co-owning delivery with their leadership.



Full-Stack Analytics Capability

From raw data pipelines and Unity Catalog governance to ML model versioning, customer segmentation, and executive dashboards - one team, end-to-end.



Domain Depth in CPG & Retail

Promotions analytics, loyalty programme intelligence, and customer360 are core competencies - not adjacent skills. We speak the business language your stakeholders speak.



Proven ROI Delivery

We don't just build - we optimise. 75% cloud cost reduction on offer model pipelines, 90% CLV pipeline savings, 40% faster runtimes. Impact that shows up in your HQ conversations.

About Polestar Analytics

Polestar Analytics is a leader in Data, Analytics, AI, and Enterprise Planning helping organizations to unlock intelligent outcomes through our proprietary products like 1Platform, accelerators, and services. Our expertise spans data engineering, data science, agentic and generative AI, and advanced planning for CPG/Retail, Pharmaceuticals, Manufacturing, IT/ITeS, and Financial Services.

Reach out to us today!