SAMPLE FICTIONAL EXERCISE

This case study is a fictitious scenario constructed solely to demonstrate the author's strategic and technical approach to solving complex organizational marketing analytics problems. All organizations, individuals, metrics, and events described herein are entirely fictional. No real companies, institutions, or individuals are represented.

CASE STUDY | PRISM RETAIL GROUP

Optimizing Multi-Channel Marketing Performance with Unified Analytics

Retail / E-Commerce | Customer Intelligence & Marketing Analytics

ORGANIZATION **Prism Retail Group**

SOLUTION ARCHITECT **Manuel Munoz Jr.**

DOMAIN **Retail / E-Commerce — Marketing Analytics**

VERSION **1.0**

This document presents a complete project lifecycle narrative — from initial business case and commission through data architecture, customer segmentation modeling, and measurable marketing outcomes. It is designed to provide full transparency into how fragmented channel data was unified into an actionable intelligence platform.

01 EXECUTIVE SUMMARY

From Disconnected Channels to a Unified Customer Intelligence Platform

Prism Retail Group operated a growing e-commerce and omnichannel retail business across four distinct brands, spending approximately $14.2M annually across paid search, social media, email, and organic channels. Despite significant marketing investment, leadership had no unified view of what was working, who their most valuable customers were, or how to allocate budget across channels with confidence.

This engagement was commissioned to replace disconnected, channel-silo reporting with a unified analytics platform — one that could model customer lifetime value, segment the customer base for targeted activation, and give marketing leadership a single, trusted view of performance across all channels.

34%	2.1x	$2.8M	4 Brands
ROAS IMPROVEMENT	**CLV UPLIFT**	**EST. REVENUE IMPACT**	**UNIFIED**
Budget reallocated to high-CLV segments	Top segment vs. prior unmanaged baseline	Year-1 incremental attributed revenue	Single platform across all business lines

Core Outcome

Leadership shifted from reporting on what happened last month to making forward-looking decisions about which customers to invest in, which channels to scale, and which campaigns to retire — based on lifetime value, not last-click attribution.

Engagement Snapshot

INDUSTRY **Retail / E-Commerce**	SCOPE **4 Brands, Enterprise Marketing Analytics**	STACK **Microsoft Fabric | Power BI | OneLake | PySpark**	DURATION **~7 Months**

02 BUSINESS CASE

Why This Project Was Commissioned

In the lead-up to Q4 planning, Prism Retail Group's CMO and CFO jointly presented an analysis to the Executive Committee revealing a critical gap: the company was spending $14.2M annually on marketing with no reliable way to measure true return, no customer-level performance data, and no shared definition of a profitable acquisition.

The Executive Committee authorized a cross-functional analytics initiative with the following stated mandate:

> **Project Charter — Stated Mandate**
> *To design and implement a unified marketing analytics platform that consolidates customer and campaign data across all brands and channels, establishes a governed CLV and segmentation model, and enables marketing leadership to optimize spend allocation based on lifetime value intelligence rather than last-click metrics.*

Commissioning Drivers

R	**ROAS Opacity** Marketing teams across the four brands were each reporting their own ROAS using different attribution windows, different channel definitions, and different conversion events. The CFO had no way to compare investment efficiency across brands or make cross-brand budget decisions.
C	**Customer Blind Spot** The business had no customer lifetime value model. Acquisition cost decisions were made at the campaign level without any understanding of whether the acquired customers were high-value repeat buyers or one-time purchasers who would never return.
S	**Siloed Channel Data** Each channel team — paid search, social, email, and organic — maintained its own performance reporting in separate platforms. No mechanism existed to attribute revenue to a customer journey that crossed multiple channels, which virtually all journeys did.
G	**Growth Stage Inflection** Prism was entering a Series C capital raise. Institutional investors required evidence of marketing efficiency and scalable customer acquisition economics. The existing reporting infrastructure could not produce those metrics in any defensible form.

Project Authorization & Scope Boundaries

- Authorized budget: $620,000 (internal labor, tooling, Microsoft Fabric licensing)

- Timeline: 7 months from kickoff to full production deployment
- In scope: All four Prism brands — Prism Home, Prism Apparel, Prism Beauty, Prism Kids; all digital marketing channels
- Out of scope: Offline / in-store attribution; changes to CRM platform; media buying strategy
- Success criteria: Unified cross-brand dashboard live; CLV model deployed with segment activation capability; ROAS variance across brands resolved to single governed definition

03 STAKEHOLDER REGISTER

Project Team & Accountable Parties

The following individuals represented the core project team and primary stakeholders across the engagement. Roles reflect both formal organizational titles and project-specific responsibilities.

NAME	TITLE	DEPARTMENT	PROJECT ROLE
Manuel Munoz Jr.	Solution Architect & Developer	External Engagement	Led end-to-end solution design including Fabric Lakehouse architecture, CLV modeling, RFM segmentation framework, semantic model design, and Power BI reporting layer.
Daniela Voss	Chief Marketing Officer	Marketing	Executive sponsor. Defined campaign performance requirements, approved CLV segment definitions, and served as primary business decision authority.
Raymond Chu	Chief Financial Officer	Finance	Co-sponsor. Drove ROAS governance requirements, validated attribution model logic, and defined CFO-level reporting needs for investor communication.
Sofia Larsen	VP of Growth Marketing	Marketing	Functional lead for channel performance requirements. Led business-side validation of attribution outputs and defined campaign optimization thresholds.
Dev Anand	Director of Data Engineering	Technology	Owned Fabric platform provisioning, OneLake access governance, and pipeline operationalization. Managed ongoing data engineering team post-deployment.
Camille Okonkwo	Senior Marketing Analyst	Marketing Analytics	Subject matter expert for channel-level reporting logic. Primary tester for dashboard accuracy, segment validation, and CLV model calibration.
Jae Park	E-Commerce Director	Prism Apparel	Brand-level stakeholder representing the highest-revenue business unit. Provided transaction data documentation and validated customer journey assumptions.
Lena Hoffmann	CRM & Loyalty Manager	Customer Experience	Provided customer identity and loyalty data as the primary source for CLV historical transactions. Validated segment-to-channel mapping assumptions.
Marcus Webb	Project Manager	Enterprise PMO	Managed timeline, stakeholder communication, and budget tracking. Liaison between technical team and executive steering committee.

04 THE REALITY OF THE PROBLEM

Four Brands. Four Realities. Zero Shared Truth.

Prism Retail Group's marketing problem was not a lack of data. It was a surplus of incompatible data — each brand team working from its own channel dashboards, its own conversion definitions, and its own interpretation of what a successful campaign looked like.

At the enterprise level, leadership was asked to make $14.2M in annual budget allocation decisions using a collection of screenshots from four different platforms, assembled into a PowerPoint by a junior analyst the night before every monthly review.

> **The Real Problem**
> *The business was flying blind at the moment it needed clarity most — during peak campaign windows, when budget reallocation decisions had to be made in days, not weeks. No one could answer a simple question: which customers, acquired through which channels, were worth acquiring again?*

What Was Happening

Each of the four brand teams used different attribution windows — some last-click, some linear, one using a custom model built in a spreadsheet. Paid search reported a 4.8x ROAS. Social reported 3.2x. Neither number was comparable. Both were contested in every budget meeting.

Where It Broke Down

Customer data lived in three separate systems — the e-commerce platform, the loyalty CRM, and a third-party email tool — with no common customer identifier across them. A customer who purchased through paid search, returned via email, and converted a second time through direct was counted as three separate customers.

05 WHY THIS WAS HARD

Identity, Attribution, and Organizational Resistance

This engagement presented complexity across four distinct dimensions. Each required a different intervention — and solving one without the others would have produced the wrong answer with high confidence.

I	**Customer Identity Fragmentation** No unified customer identifier existed across the e-commerce platform, loyalty CRM, and email marketing system. A single customer could appear under three different IDs depending on which touchpoint they last used. Any CLV or segmentation model built on top of this data would compound the fragmentation.
A	**Attribution Model Conflict** Each brand used a different attribution methodology with different lookback windows. Unifying these into a single model required stakeholder consensus on a new shared definition — before any technical work could begin. Channel teams were resistant to any model that might reduce their reported contribution.
D	**Data Volume & Velocity** Prism processed approximately 2.4M customer events per day across all brands and channels. The existing reporting infrastructure — built on Excel and siloed channel platforms — had no mechanism for handling this volume at the speed required for campaign-cycle decision-making.
O	**Organizational Silos** Each brand operated with near-complete marketing autonomy. A unified analytics platform was perceived by some brand teams as a threat to that autonomy. Stakeholder alignment required framing the platform as an enabler of brand-level decision-making, not as a corporate control mechanism.

Architect's Note

The most technically difficult problem in this engagement was customer identity resolution — not the CLV model itself. A mathematically precise lifetime value calculation built on fragmented identities is precisely wrong. We solved the identity layer first, or we would have built the rest on a false foundation.

06 MY APPROACH

Unify the Customer First. Everything Else Follows.

The instinct in marketing analytics engagements is to start with the dashboard — stakeholders want to see something quickly, and channel performance is the most visible output. The correct instinct is to start with the customer identity layer, because every downstream analysis depends on knowing whether two records represent one person or two.

Discovery Objectives

- Audit all customer data sources across brands and map the existence and consistency of customer identifiers
- Establish a probabilistic identity resolution approach for customers with no common key across systems
- Define a shared attribution model — including lookback window, channel weighting, and conversion event definitions — with explicit sign-off from Finance and Marketing
- Determine the CLV calculation methodology and the minimum historical transaction depth required to produce reliable scores
- Map how campaign decisions were currently made — and identify the exact moments where data quality was causing those decisions to be wrong

Core Design Principles

Identity Before Intelligence
No segmentation, CLV score, or attribution report is trustworthy until the customer records underneath it are resolved to a single, unified identity.

CLV as the Unit of Analysis
Every campaign performance question would ultimately be answered in terms of customer lifetime value impact — not click-through rate, impressions, or even single-transaction ROAS.

One Attribution Contract
All four brands would share a single, governing attribution definition. Brand-level customization was permitted only within that contract — not in violation of it

Activation-Ready Architecture
The segmentation model was designed not just to describe customers, but to produce outputs that could be directly consumed by channel activation tools — email, paid media, and CRM.

07 CUSTOMER SEGMENTATION & CLV MODEL

The Central Differentiator: Knowing Who Is Worth What

The customer intelligence layer was the primary differentiator of this engagement — the capability that transformed the platform from a reporting tool into a decision engine.

The CLV model was built using a three-year transaction history across all four brands, incorporating recency, frequency, and monetary value (RFM) as the foundation, with probabilistic lifetime extension modeling applied on top.

RFM Segmentation Framework

Customers were scored across three dimensions — Recency (days since last purchase), Frequency (number of purchases in the trailing 12 months), and Monetary value (total spend) — and assigned to one of six actionable segments:

SEGMENT	RFM PROFILE	% OF CUSTOMERS	AVG. CLV	PRIMARY CHANNEL	CAMPAIGN STRATEGY
Champions	High R, High F, High M	8%	$1,840	Email + Loyalty	Exclusive early access, VIP retention programs, referral activation
Loyalists	High R, High F, Med M	14%	$920	Email + Retargeting	Cross-brand expansion offers, subscription upsell, frequency rewards
Promising	High R, Low F, Med M	22%	$410	Paid Social + Email	Second-purchase incentives, product education, brand affinity building
At Risk	Low R, High F, High M	11%	$680	Email Win-back	Personalized win-back sequences, loyalty re-engagement, exclusive offers
Hibernating	Low R, Low F, Low M	27%	$90	Paid Search	Low-cost awareness only; exclude from high-CPM social inventory
New Customers	High R, Low F, Any M	18%	TBD	All Channels	Onboarding sequence, first-30-day engagement scoring, early CLV signal capture

CLV Calculation — Certified Measure (Power BI DAX)

The CLV measure was implemented as a certified calculation in the Power BI semantic model to ensure consistency across all reports and dashboards. All teams — regardless of brand — referenced this single definition.

```
CLV_Score =
  VAR _AvgOrderValue =
    CALCULATE(
      AVERAGEX( Customer_Transactions, Customer_Transactions[order_value] ),
      Customer_Transactions[transaction_date] >= TODAY() - 365
    )
  VAR _PurchaseFrequency =
    CALCULATE(
      COUNTROWS( Customer_Transactions ),
      Customer_Transactions[transaction_date] >= TODAY() - 365
    )
  VAR _AvgCustomerLifespan = 3.2  -- trailing cohort average in years
  VAR _GrossMargin = 0.42        -- blended margin across all brands
RETURN
  IF(
    _PurchaseFrequency = 0, BLANK(),
    _AvgOrderValue * _PurchaseFrequency * _AvgCustomerLifespan * _GrossMargin
  )
```

Design note: _AvgCustomerLifespan and _GrossMargin are parameterized constants stored in a governance table — not hardcoded — enabling Finance to update assumptions without model redeployment.

Segment-to-Channel Activation Mapping

A key design requirement was that segmentation outputs could be consumed directly by downstream activation tools. Fabric pipelines wrote segment assignments back to the CRM and email platform on a nightly basis, enabling channel teams to build audiences without requiring manual data exports.

★	~	?	!
Champions + Loyalists	**At Risk Customers**	**Promising Segment**	**Hibernating Segment**
Exported nightly to Klaviyo VIP lists and Meta Custom Audiences for suppression from prospecting campaigns	Triggered automated win-back email sequence within 24hrs of RFM score crossing threshold	Fed into paid social lookalike seed audience — modeling expansion against high-propensity second-purchase profile	Excluded from high-CPM Meta and TikTok inventory; retained only in low-cost Google search retargeting

08 ARCHITECTURE & DESIGN DECISIONS

Microsoft Fabric as the Unification Layer

The platform was built on Microsoft Fabric as the end-to-end data and analytics foundation, with Power BI as the governed reporting and activation layer. The choice of Fabric was driven by a specific architectural requirement: all four brands needed to share a single Lakehouse while maintaining logical isolation of brand-level data — a capability that Fabric's OneLake architecture supports natively.

End-to-End Architecture

SOURCE CHANNELS Paid Search (Google Ads) | Social Media (Meta, TikTok) | Email (Klaviyo) | Organic / SEO | Direct / Referral

↓

INGESTION LAYER Microsoft Fabric Data Factory Pipelines — Unified event schema, channel normalization, deduplication

↓

LAKEHOUSE (FABRIC) OneLake — Raw / Silver / Gold layer architecture | Delta format | Medallion pattern

↓

TRANSFORMATION LAYER Fabric Notebooks (PySpark) — CLV modeling, RFM scoring, attribution logic, cohort construction

↓

SEMANTIC MODEL Power BI Dataset | DAX measures | Row-Level Security by brand / region | Certified KPIs

↓

REPORTING & ACTIVATION Power BI — CMO Dashboard | Channel Performance | Segment Explorer | Automated Campaign Alerts

Component Rationale

Microsoft Fabric / OneLake	Selected as the unified storage and compute layer. OneLake's multi-workspace shortcut architecture enabled all four brands to contribute data to a shared Lakehouse without losing brand-level access boundaries — eliminating the need for separate storage accounts per brand.
Medallion Architecture (Bronze/Silver/Gold)	Raw channel data landed in Bronze with full fidelity. Silver handled identity resolution, deduplication, and schema normalization. Gold produced the CLV-scored, segment-assigned customer master and the channel performance tables consumed by Power BI.

Fabric Notebooks (PySpark)	Customer identity resolution and CLV modeling required iterative computation across hundreds of millions of event records — beyond the practical reach of SQL-only transformation. PySpark enabled the probabilistic matching and cohort modeling logic at the required scale.
Power BI Semantic Model	Served as the governed reporting contract between the data layer and all consumers. All KPI definitions — ROAS, CLV, CAC, CLTV-to-CAC ratio — were implemented as certified measures here, ensuring consistency regardless of which brand or report referenced them.
Row-Level Security (RLS)	Implemented at the semantic model level to ensure brand teams could see their own performance data and enterprise rollups, but not raw data from other brands. Single shared dataset, filtered by identity — same pattern as the healthcare engagement, different domain.

09 INSIDE THE BUILD

Selected Artifacts & Design Details

The following artifacts represent key decisions made during implementation. They are included to make the reasoning behind technical choices transparent — not as a showcase of tools.

Artifact 1 — Customer Identity Resolution (PySpark Logic)

The identity resolution process was the most foundational step in the entire engagement. Customers were matched across the three source systems using a probabilistic scoring approach combining email address hash, device fingerprint, and behavioral sequence similarity.

1	**Deterministic Match**	Exact email hash match across e-commerce, CRM, and email platform. Records sharing a verified email are assigned a shared master_customer_id immediately.
2	**Probabilistic Match**	For records with no email match: score similarity across device_id, session_ip_range, and purchase_sequence_hash. Records scoring above 0.82 confidence are soft-matched and flagged for review.
3	**Golden Record Creation**	For each resolved identity, the Silver layer produces a single golden record selecting the most recent, most complete attribute set. All source IDs are retained as foreign keys for lineage.
4	**Validation Pass**	Post-resolution, a PySpark validation job checks match rate by brand and flags anomalies. Any brand falling below 78% match rate triggers a data quality alert before proceeding to Gold.

[Visual Placeholder: Fabric Notebook — identity resolution PySpark logic, probabilistic scoring function]

Artifact 2 — Attribution Model Contract

Before any attribution reporting was built, a formal Attribution Contract was produced and signed off by Marketing and Finance. This document defined the single governing attribution model for all four brands — ending the ROAS disagreement at the source, not in the dashboard.

PARAMETER	DEFINITION	RATIONALE
Attribution Model	Data-driven (position-based with ML weighting)	Replaces last-click for all brands; weighting updated monthly via Fabric Notebook job
Lookback Window	30-day click / 1-day view	Standardized across all channels; eliminates window discrepancy between paid search and social reports
Conversion Event	Confirmed order (post-return window)	Excludes returns from ROAS numerator; Finance requirement for P&L alignment

Cross-Brand Revenue	Attributed to acquisition brand only	Revenue from a customer acquired by Prism Apparel who later buys from Prism Home is credited to Apparel
ROAS Definition	Revenue attributed / Media spend (inclusive of agency fees)	CFO sign-off; ensures comparability across brands and with investor reporting

[Visual Placeholder: Power BI semantic model view — ROAS certified measure, CLV_Score measure, attribution lineage]

10 FROM DATA TO DECISION

The Reporting Experience

The reporting layer was designed around three distinct decision-making contexts — each requiring a different depth of information and a different pace of interaction.

CMO Dashboard

- Cross-brand ROAS, CAC, and CLV-to-CAC ratio — single-screen view
- Segment health summary — size, trend, and migration velocity
- Budget efficiency heatmap: spend vs. CLV-weighted return by channel

[Visual Placeholder: CMO dashboard — cross-brand KPI overview with CLV trend]

Campaign Performance View

- Channel-level ROAS under the unified attribution contract
- Segment acquisition mix by campaign — which segments did this campaign win?
- CLV projection for newly acquired cohort vs. historical benchmark

[Visual Placeholder: Campaign drillthrough — segment acquisition mix by channel]

Segment Explorer

A dedicated segment exploration view gave marketing analysts the ability to interrogate any of the six customer segments across dimensions including acquisition channel, brand, product category, cohort period, and geographic market. This view was intentionally designed to support hypothesis generation — not just reporting — enabling the team to ask questions the dashboard had not anticipated.

[Visual Placeholder: Segment Explorer — RFM scatter plot with CLV overlay, segment migration Sankey]

> **The Shift**
> *From: Monthly PowerPoint assembled from channel screenshots | To: Daily automated dashboard with segment-level CLV visibility and same-day campaign performance reporting under a single attribution model.*

11 BUSINESS IMPACT

Outcomes That Changed How the Business Grows

The following metrics represent outcomes measured at 6 months post-deployment, based on internal marketing benchmarks, Finance-validated attribution reporting, and cohort analysis comparing pre- and post-platform campaign performance.

34% **ROAS IMPROVEMENT** Budget shifted to high-CLV channel mix	**2.1x** **CLV UPLIFT** Champions vs. pre-platform unmanaged avg	**$2.8M** **INCREMENTAL REVENUE** Year-1, Finance-validated attribution	**68%** **CAC REDUCTION** Hibernating segment excluded from high-CPM

BEFORE	AFTER
ROAS calculated four different ways across four brand teams; monthly budget reviews consumed by definition disputes	Single governed ROAS definition under unified attribution contract — budget meetings focused on strategy, not reconciliation
No customer lifetime value model; acquisition decisions made on single-transaction ROAS with no view of long-term value	Six-segment CLV model deployed; Champions and Loyalists segments identified, generating 2.1x average CLV vs. prior unmanaged baseline
Hibernating segment (27% of customer base) receiving same media investment as high-CLV segments	Hibernating segment excluded from $3.1M of annual high-CPM Meta/TikTok spend; 68% CAC reduction on that cohort
Campaign reporting available 5-7 business days after campaign close via manual analyst assembly	Campaign performance visible within 24 hours of launch; segment acquisition mix available same day
Acquisition lender flagged inability to produce customer economics data as a risk to Series C terms	CLV model and cohort retention data provided to lender; financing terms improved, closing on schedule

Most Important Outcome

The business stopped optimizing for the cheapest click and started optimizing for the most valuable customer. That single shift in analytical frame — from transaction-level ROAS to customer lifetime value — was the source of all measurable financial improvement.

12 TRADE-OFFS & LESSONS

What Was Optimized, What Would Be Refined

Every architectural decision involves trade-offs. Naming them explicitly — rather than presenting the solution as optimal — is more honest and more instructive.

DECISION POINT	CHOSEN APPROACH	RATIONALE
Identity resolution method	Probabilistic (0.82 threshold)	Chose coverage over precision. A higher threshold would have produced cleaner matches but left 18% of customers unresolved, undermining CLV model population. Threshold calibrated against downstream false-positive rate.
Attribution model complexity	Data-driven position-based	More sophisticated than last-click, less complex than full Shapley value. Chosen to balance accuracy with stakeholder comprehension — a model no one trusts is less useful than a slightly imperfect one everyone uses.
CLV model granularity	Brand-level margin assumption	A product-level margin model would have been more precise, but required product catalog data that was incomplete across two brands. Brand-level was the most granular defensible calculation available at the time of deployment.
Segment count	Six segments	Tested four, six, and eight segment configurations. Six optimized the balance between behavioral distinctiveness and channel team operability — fewer than six collapsed actionable differences; more than six exceeded what channel teams could act on.

What I Would Refine in a Next Version

- Implement a full Shapley value attribution model using Fabric ML capabilities — the current position-based model is defensible but not theoretically optimal for multi-touch journeys with long consideration cycles

- Build a real-time segment scoring pipeline that updates CLV scores intra-day rather than nightly — enabling same-session personalization at the edge for the e-commerce platform
- Add a predictive churn probability score to the At Risk segment definition, replacing the reactive RFM threshold with a forward-looking propensity model
- Integrate segment activation directly into the paid media platforms via API rather than nightly file export — reducing the activation latency from 24 hours to near real-time

Key Principle Reinforced

In marketing analytics, the most dangerous number is the one everyone agrees is correct but nobody has defined the same way. Before building any model, establish the contract. The attribution contract saved more time than any technical optimization in this engagement.

13 KEY TAKEAWAY

The Customer Is the Unit of Analysis

This engagement was ultimately about reframing how a growth-stage retail business understood its own economics. Not: which campaign generated the most clicks last week. But: which customers are worth acquiring, retaining, and expanding — and which channels and campaigns produce them reliably.

The technology — Fabric, Power BI, PySpark — was in service of that reframing. The architecture was in service of the segmentation model. The segmentation model was in service of better decisions.

> *"Optimizing marketing spend without a customer lifetime value model*
>
> *is just finding a faster way to make the wrong decision."*
>
> — Manuel Munoz Jr.

This Case Study Demonstrates

- ✓ Customer-first analytical design — identity resolution was solved before any model or dashboard was built
- ✓ Governance through contracts — the Attribution Contract eliminated a structural conflict that no dashboard could have resolved
- ✓ Segmentation as a decision engine — CLV segments were designed for activation, not just classification
- ✓ Microsoft Fabric at enterprise scale — OneLake Medallion architecture handling 2.4M daily events across four brands in a single governed platform
- ✓ Trade-off transparency — what was optimized, what was deferred, and what the next architectural iteration would address

Manuel Munoz Jr. | Solution Architect & Developer | Prism Retail Group Engagement

MUNOZDATAWORKS.COM