GREYFIELD DATA

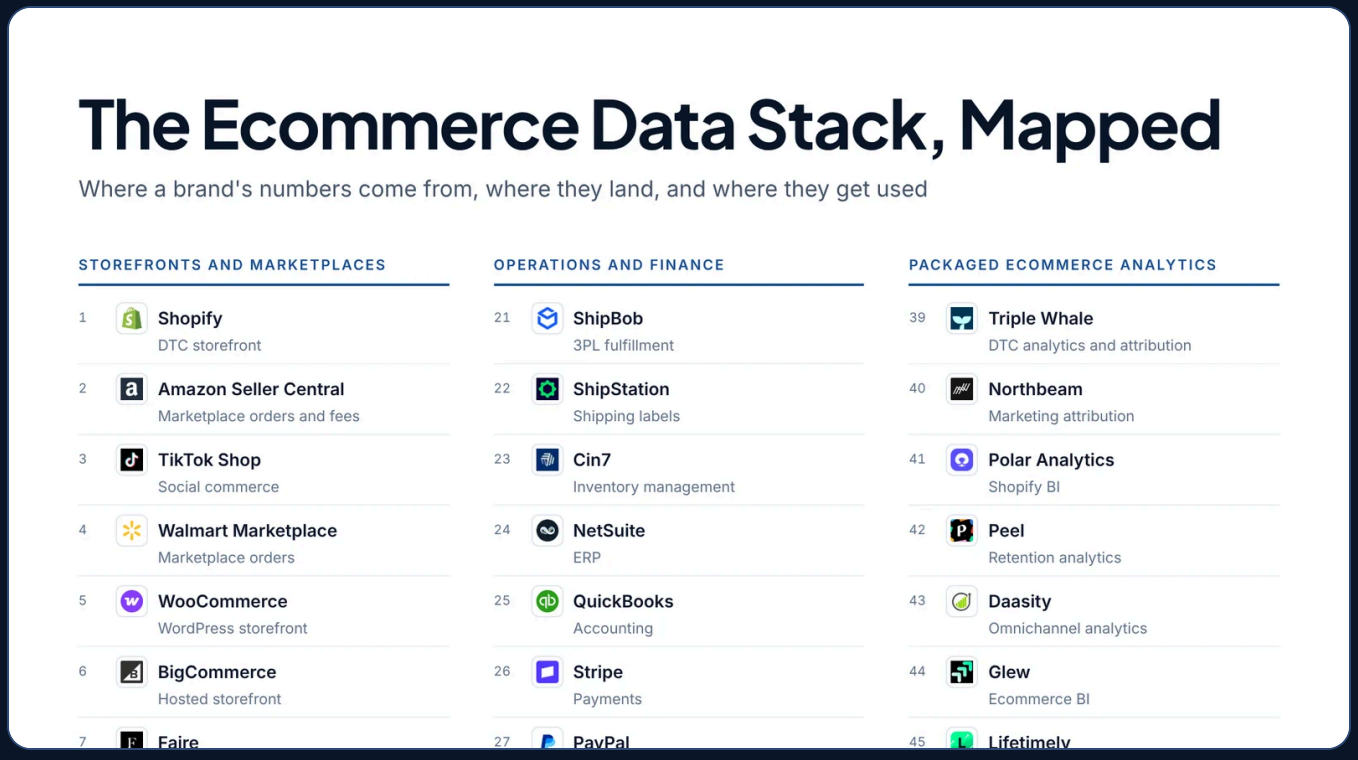
WHITE PAPER

FOR DTC AND OMNICHANNEL CONSUMER BRANDS

The Ecommerce and Consumer Brand Data Stack

Tool selection by stage, and a phased implementation plan.

Sources | Ingestion | Warehouse | Modeling | Semantic layer | BI | Activation



Linus Tse, Greyfield Data
November 2026

greyfielddata.com

01

Executive summary

A consumer brand's numbers start in a dozen systems: storefronts, marketplaces, ad platforms, subscriptions, returns, fulfillment, and accounting. A data stack moves those numbers into one place, defines them once, and puts them in front of the people who decide. This paper helps you choose those tools for the stage your brand is at, and build them in an order that pays off at each step.

Choose by stage, not by logo

A single-channel Shopify brand, a brand on Amazon and TikTok Shop, and a brand with wholesale and retail need different stacks. Section 04 gives one for each.

Packaged tools are often right

Triple Whale, Polar, Peel, Daasity, and others are faster and cheaper than a build when their definitions fit your business. Section 03 shows when they stop fitting.

Definitions come before tools

Most reporting problems are definition problems. A written metric dictionary is the cheapest layer in the stack and the one everything else reads.

Build in phases with checks

Five phases, each with deliverables, owners, acceptance checks, and risks. You move on when the numbers tie out, not when a date arrives.

CONTENTS

Who this paper is for	3	Modeling, semantic layer, BI, activation	11
The seven layers of the stack	4	The implementation plan	12
The ecommerce stack, mapped	5	Phases 0 and 1	13
Packaged analytics or a warehouse	6	Phases 2 and 3	14
Stage 1: single-channel Shopify	7	Phase 4, and what Phase 3 produces	15
Stage 2: multi-channel	8	Where to start: readiness checklist	16
Stage 3: omnichannel	9	Common failure modes	17
Ingestion	10	About Greyfield Data	18

Take the quiz. The free ecommerce data stack quiz applies this framework to ten answers about your brand and returns one pick per layer and the phase to start with.

02

Who this paper is for

Founders, finance leads, growth leads, and the first data hire at direct-to-consumer and omnichannel consumer brands. It assumes you sell physical products online, and that someone has asked a question your current reports cannot answer, or answers three different ways.

THREE STAGES

The tools change with how many places you sell and how many teams rely on the numbers. Revenue size matters less than channel count and who needs the figures to tie to the books.

1 **Single-channel Shopify brand**

One storefront, maybe a few Shopify stores. Questions are about ads, profit per order, and repeat purchase. Often no data person. Shopify's own reports or a packaged tool usually come first.

2 **Multi-channel: Shopify plus Amazon, TikTok Shop, or Walmart**

The same customer and the same product now live in several systems, each with its own fees, dates, and SKU codes. Channel totals stop adding up to the bank. This is where a warehouse usually starts to earn its cost.

3 **Omnichannel: online plus wholesale or retail**

An ERP such as NetSuite, wholesale through Faire or direct accounts, and retailer or POS data sit next to online orders. Finance, operations, and planning need one model. A packaged omnichannel product or a warehouse build.

WHAT THIS PAPER DOES NOT DO

It does not quote durations or prices. Both depend on your sources, history, data quality, and how quickly owners agree on definitions, and the paper says what drives each one instead (page 12). Vendor facts come from Greyfield's sourced comparison pages at greyfielddata.com/compare. Check current plans with each vendor.

03

The seven layers of the stack

Every stack, packaged or built, does the same seven jobs. A packaged tool does several of them inside one product with its own definitions. A build picks a tool for each.

Layer	Job	Common tools	What decides the pick
Sources	Create the data: orders, refunds, fees, spend, subscriptions, costs	Shopify, Amazon Seller Central, TikTok Shop, Faire, Meta Ads, Google Ads, Klaviyo, Recharge, Loop, ShipBob, NetSuite, QuickBooks	Your channels and operations; you rarely choose these for reporting
Ingestion	Copy source data into the warehouse on a schedule	Fivetran, Airbyte, Daton, Funnel, Windsor.ai, Supermetrics, custom code	Which sources and tables each connector covers, who runs it, how it is priced
Warehouse	Store everything in one place you can query	BigQuery, Snowflake, Databricks	The cloud you already use, and what your connectors and tools support
Modeling	Turn raw tables into one order, customer, and margin model	dbt (managed platform or Core), prebuilt vendor models	Who maintains the SQL, and how specific your rules are
Semantic layer	Define each metric once: grain, date, sources, owner	A metric dictionary, dbt marts, Snowflake semantic views, a BI semantic model	How many tools read the numbers, including AI assistants
BI	Show the numbers to the people who decide	Looker Studio, Sigma, Power BI, Tableau, Metabase, Shopify Analytics	Who builds reports, your cloud, and whether people enter plans
Activation	Send audiences and scores back to marketing tools	Hightouch, Census, native Shopify and Klaviyo syncs	Whether the customer view is trusted yet

Packaged analytics sits across layers. Triple Whale, Polar, Peel, Daasity, Lifetimely, and similar tools bundle ingestion, a model, metrics, and dashboards. Several now include a warehouse you can query, which makes them a reasonable first step toward a build rather than a dead end.

The general version of this map, beyond ecommerce, is at greyfielddata.com/blog/modern-data-stack-tools-mapped.

03 CONTINUED

The ecommerce stack, mapped

57 common tools in nine categories, from the systems that create data to the tools that report on it and send it back out.

The first four columns are where data is created. Connectors move it, the warehouse and models define it, and the last groups report on it or activate it.

Numbers are for reference, not a ranking, and inclusion is not an endorsement.

Full-size map and a short note on every tool: greyfielddata.com/blog/ecommerce-data-stack-tools-mapped

The Ecommerce Data Stack, Mapped

Where a brand's numbers come from, where they land, and where they get used

STOREFRONTS AND MARKETPLACES

1 **Shopify** — DTC storefront
2 **Amazon Seller Central** — Marketplace orders and fees
3 **TikTok Shop** — Social commerce
4 **Walmart Marketplace** — Marketplace orders
5 **WooCommerce** — WordPress storefront
6 **BigCommerce** — Hosted storefront
7 **Faire** — Wholesale marketplace

SUBSCRIPTIONS, RETURNS, AND SUPPORT

8 **Recharge** — Subscriptions
9 **Skio** — Subscriptions
10 **Loop Returns** — Returns and exchanges
11 **Redo** — Returns coverage
12 **Gorgias** — Ecommerce helpdesk
13 **Yotpo** — Reviews and loyalty

MARKETING AND ADS

14 **Meta Ads** — Facebook and Instagram ads
15 **Google Ads** — Search and Shopping ads
16 **TikTok Ads** — Short-video ads
17 **Amazon Ads** — Sponsored products
18 **Klaviyo** — Email and SMS
19 **Attentive** — SMS marketing
20 **GA4** — Site analytics

OPERATIONS AND FINANCE

21 **ShipBob** — 3PL fulfillment
22 **ShipStation** — Shipping labels
23 **Cin7** — Inventory management
24 **NetSuite** — ERP
25 **QuickBooks** — Accounting
26 **Stripe** — Payments
27 **PayPal** — Payments

ECOMMERCE CONNECTORS

28 **Daton** — Ecommerce ELT
29 **Fivetran** — Managed connectors
30 **Airbyte** — Open-source connectors
31 **Funnel** — Marketing data hub
32 **Windsor.ai** — Marketing connectors
33 **Supermetrics** — Ad data to sheets and warehouses
34 **Improvado** — Marketing data pipeline

WAREHOUSE AND MODELS

35 **Snowflake** — Cloud data warehouse
36 **BigQuery** — Serverless warehouse
37 **Databricks** — Lakehouse platform
38 **dbt** — Order and margin models

PACKAGED ECOMMERCE ANALYTICS

39 **Triple Whale** — DTC analytics and attribution
40 **Northbeam** — Marketing attribution
41 **Polar Analytics** — Shopify BI
42 **Peel** — Retention analytics
43 **Daasity** — Omnichannel analytics
44 **Glew** — Ecommerce BI
45 **Lifetimely** — LTV and profit
46 **Klar** — Profit analytics
47 **Saras Analytics** — Ecommerce data platform
48 **BeProfit** — Profit tracking
49 **TrueProfit** — Profit tracking
50 **Sellerboard** — Amazon profit

BI AND REPORTING

51 **Sigma** — Spreadsheet-style BI
52 **Power BI** — Microsoft BI
53 **Tableau** — Visual analytics
54 **Looker Studio** — Free Google dashboards
55 **Shopify Analytics** — Built-in store reports

ACTIVATION

56 **Hightouch** — Warehouse audiences
57 **Census** — Warehouse to app sync

57 tools in 9 categories · greyfielddata.com · Linus Tse

Logos belong to their owners. No ranking or endorsement implied.

03 CONTINUED

Packaged analytics or a warehouse?

Choose packaged analytics when its definitions match how you run the business, and a warehouse when they don't. Many brands run both: a packaged tool for daily marketing views and a warehouse for finance, margin, and anything that has to tie out.

Tool	Fits when	A warehouse fits better when
Shopify Analytics	Nearly all revenue goes through the store; included in every plan, with cohorts and profit from unit costs	Revenue happens outside the store, or profit must include fees and ad spend
Lifetimely	A Shopify store wants daily profit and LTV without a project; free plan, priced by orders	Other channels matter, or finance needs reconciled numbers
Triple Whale	Shopify is most of revenue and paid ads are the main lever; pixel and attribution	Amazon, TikTok Shop, wholesale, or retail are a large share; warehouse sync is a paid add-on
Northbeam	Heavy paid media spend needing attribution, incrementality, and MMM	Finance needs margin, inventory, and channel P&L
Polar Analytics	Shopify-led brand wants a warehouse without a data team; dedicated Snowflake on every plan	Revenue logic, sources, or brands do not fit a predefined model
Peel	Retention and subscription LTV on Shopify and Amazon; Snowflake access on every plan	TikTok Shop, wholesale, or ERP data matter, or full contribution margin is needed
Daasity	DTC plus retail with standard omnichannel needs; prebuilt models, retail and syndicated data	Your models need rules specific to you, or you want to pick the BI tool
Sellerboard	Amazon is most of revenue; fee detail, PPC automation, restock planning	Shopify, TikTok Shop, or wholesale need one profit number across channels

SIGNALS YOU HAVE OUTGROWN A PACKAGED TOOL

- Finance rebuilds the tool's numbers in a spreadsheet before every month-end or board meeting.
- A channel, fee, or cost the tool does not model is now a large share of revenue or cost.
- You run several brands or stores and need the same definitions on each.
- You want the data in a warehouse you own, so any BI tool or AI assistant can query it later.

Each row summarizes a sourced comparison at greyfielddata.com/compare. Greyfield Data builds warehouses, so the comparisons also say where the packaged tool is the better choice.

04 TOOL SELECTION BY STAGE

01 Single-channel Shopify brand

One store holds every order. The questions are about ads, profit per order, and who comes back.

Signals you are here

Shopify is nearly all revenue. A founder checks sales and spend daily. Nobody owns data models, and the main report is a spreadsheet of exports.

Start packaged. Shopify Analytics first. Then one focused tool: a profit app such as Lifetimely, BeProfit, or TrueProfit; Triple Whale for attribution; Peel for retention; or Polar if you want a warehouse without a data team.

THE STACK TO BUILD WHEN YOU OUTGROW IT

Layer	Recommended	What decides it
Ingestion	**Fivetran or Airbyte**	Shopify, Klaviyo, Meta, and Google Ads are covered by most connectors. Pay for managed if nobody will run infrastructure.
Warehouse	**BigQuery or Snowflake**	BigQuery if you already use Google Cloud and Looker Studio; Snowflake if your tools or a packaged product already write to it.
Modeling	**dbt**	One order table with refunds and discounts, a dated product cost table, a customer table.
Semantic layer	**A metric dictionary**	Net sales, contribution margin, CAC, and repeat rate written down once.
BI	**Looker Studio or Sigma**	Looker Studio for a free founder scorecard; Sigma when finance wants a spreadsheet on live data.
Activation	**Native Shopify and Klaviyo syncs**	Basic audiences already flow between these tools.

WHEN A BUILD EARNS ITS PLACE AT THIS STAGE

- Finance needs numbers that tie to the books, and the packaged tool's figures do not.
- You run several Shopify stores or brands and want one set of definitions across them.
- Your margin rules (bundles, gift cards, shipping income) differ from the tool's model.

04 TOOL SELECTION BY STAGE

02 Multi-channel: Shopify, Amazon, TikTok Shop

The same customer and the same product now live in several systems, each with its own fees and dates.

Signals you are here

Channel totals do not add up to the bank. Amazon profit moves after the sale as fees settle. One product has a different SKU code on each channel. Someone reconciles by hand every month.

Packaged tools that can fit. Polar or Peel if Shopify still leads and their definitions fit; Peel reads Shopify and Amazon. Sellerboard if Amazon is most of revenue. Triple Whale's TikTok Shop integration is US-only and leaves out platform and affiliate fees for now.

RECOMMENDED STACK

Layer	Recommended	What decides it
Ingestion	**Daton or Fivetran**	Whether the connector loads the Amazon settlement and fee reports you need, not only orders. Daton targets marketplace sources at low cost.
Warehouse	**BigQuery or Snowflake**	The cloud you already use; both are supported by the connectors and packaged tools named here.
Modeling	**dbt**	Order, refund, and fee models per channel; a SKU map across channels; deduplication so no order counts twice.
Semantic layer	**A metric dictionary**	Which date a refund uses, and estimated versus settled fees, written down.
BI	**Sigma, Power BI, or Looker Studio**	Who builds reports: finance in spreadsheets (Sigma), a Microsoft shop (Power BI), or a free scorecard (Looker Studio).
Packaged	**Optional, alongside**	Keep an attribution or retention tool marketing relies on, and load its data into the warehouse.

Further reading: avoiding double-counted orders, mapping SKUs across channels, and why Amazon profit changes after the sale.

04 TOOL SELECTION BY STAGE

03 Omnichannel: online plus wholesale and retail

Finance, operations, and planning need one model that covers every place the product is sold.

Signals you are here

NetSuite or another ERP holds wholesale orders and costs. Retailers send sell-through data as portal exports or files. Several brands or regions share a team. Planning happens in spreadsheets copied from reports.

Packaged option: Daasity. Prebuilt omnichannel models in Snowflake or BigQuery, with retail and syndicated data connectors. Brands with standard omnichannel needs should look hard at it before a build.

RECOMMENDED STACK

Layer	Recommended	What decides it
Ingestion	**Fivetran, plus custom code**	ERP and back-office sources favor a broad catalog. Retailer portals, EDI files, and wholesale spreadsheets often have no connector.
Warehouse	**Snowflake or BigQuery**	The cloud you use, and whether a packaged product (Daasity, Polar, Peel) already writes to it.
Modeling	**dbt, or Daasity's models extended**	How far your rules differ from standard omnichannel models.
Semantic layer	**Dictionary, then semantic views or a BI model**	Snowflake semantic views if BI and AI tools both read metrics; a shared Power BI model in a Microsoft shop.
BI	**Sigma or Power BI**	Sigma input tables let planners type forecasts next to actuals; Power BI fits Microsoft 365.
Activation	**Hightouch or Census**	Warehouse audiences once the customer view is trusted (Phase 3).

THE EXTRA WORK AT THIS STAGE

Retailer data rarely shares your product IDs or your customer IDs, and it arrives on the retailer's calendar. Budget modeling time for a product master that maps every retailer's item code, and keep sell-in (your shipments) and sell-through (retailer sales) as separate metrics.

05

Ingestion: getting each source into the warehouse

Connecting a source is the quick part. Most brands end up with two or three ingestion answers across their stack, chosen source by source.

Option	What it is	Best when
Fivetran	Managed connectors with nothing to host	Nobody will run infrastructure, or sources go beyond ecommerce apps
Airbyte	Open-source connectors, self-hosted or on Airbyte Cloud	Engineers can run servers, upgrades, and monitoring to lower the bill
Daton	Ecommerce-focused ELT from Saras Analytics into your own warehouse	Marketplaces and ad platforms dominate and cost matters
Funnel	Marketing data hub that harmonizes many ad and analytics platforms	The source is mostly ad spend, and a plan includes the destination you need (Snowflake export is on Enterprise)
Custom code	Your own scripts, or a library such as dlt, with cursors and checkpoints	No connector covers the tables you report on, or a connector's pricing is a problem at your volume

FOUR QUESTIONS PER SOURCE

- [ ] **Does a managed connector cover the tables you report on?**
  Check the exact objects: settlement and fee reports for Amazon, refund lines for Shopify, not only "orders".
- [ ] **Can someone on your team run servers, upgrades, and monitoring?**
  If not, self-hosted options cost more in time than they save.
- [ ] **Is pricing by rows or credits a problem at your volume?**
  High-volume event or ad sources are where connector bills grow. Estimate with each vendor.
- [ ] **Is the source mostly ad spend and marketing performance?**
  Marketing hubs and cheaper marketing pipes (Supermetrics, Windsor.ai) can be enough for those.

For every answer: alert on freshness per source, check row counts against the source, and review cost per connector every month. A longer comparison is at greyfielddata.com/blog/ecommerce-data-ingestion-tools.

06

Modeling, the metric dictionary, BI, and activation

MODELING IN THREE LAYERS

Bronze keeps data as it arrived, silver cleans and matches it (one row per order, duplicates removed, refunds matched, one ID per person), and gold shapes it for business questions. dbt is the common tool: on dbt Labs' managed platform if you want scheduling and docs hosted, or dbt Core on your own scheduler. A worked example with one order, a duplicate, and a late refund is at greyfielddata.com/blog/bronze-silver-gold-ecommerce-order-example.

THE METRIC DICTIONARY IS YOUR SEMANTIC LAYER

"Revenue" is at least four different metrics in an ecommerce business. Give each its own entry, and decide which one answers when someone says the word.

Field	What to write	Example for net sales
Definition	One sentence a non-analyst understands	Item sales after discounts and returns
Grain and date	What one row means; the date that decides the month	Order line; order date for sales, one rule for returns
Sources and owner	Tables it reads; who approves a change	Order and refund lines; finance
Tie-out check	The total it must match	Shopify net sales, same dates

The full dictionary, from gross sales to LTV, with a CSV template: greyfielddata.com/blog/ecommerce-semantic-layer. Once it exists, it can be encoded as dbt marts, Snowflake semantic views, or a shared BI model.

BI: pick by who builds

Looker Studio for free scorecards; Power BI in a Microsoft 365 company; Sigma when finance and operations want a spreadsheet on live warehouse data and input tables for plans; Metabase for a SQL team on a budget. The free BI tool quiz at greyfielddata.com/tools/bi-tool-quiz goes deeper.

Activation: last, not first

Hightouch and Census push warehouse audiences and scores to Klaviyo, Meta, and other tools, so targeting uses the same definitions as reporting. Until the customer view is trusted, native Shopify and Klaviyo syncs are enough.

07

The implementation plan

Five phases, each ending with acceptance checks. Each phase produces reports people use, so the work pays off before the whole stack exists. You move on when the checks pass.

Phase 0
Audit and definitions

Phase 1
Orders and margin

Phase 2
Marketing and CAC

Phase 3
Retention and customer view

Phase 4
Forecasting and activation

Phase	Question it answers	Done when
0	What do our numbers mean, and where do they come from?	Every core metric has a written entry and an owner
1	What did we sell, and what did we make, by channel?	Net sales and margin tie to the books for a closed month
2	What does a new customer cost, by channel?	Spend ties to platform billing; CAC uses the same new-customer count as Phase 1
3	Who comes back, and what are they worth?	One ID per person; cohorts only count eligible customers
4	What will we sell, and who should we target?	Forecasts backtested; synced audiences match warehouse counts

WHAT DRIVES TIMELINE AND COST

This paper gives no durations or prices because they vary too much between brands. These are the things that move them:

Timeline

- Number of sources, and whether connectors exist for each
- History depth needed, and API limits on backfills
- Data quality: duplicates, SKU mismatches, missing costs
- How fast metric owners decide, and are available for questions

Cost

- Connector pricing at your row or source volume
- Warehouse compute, driven by refresh frequency and query load
- BI seats or platform fees, and packaged tool plans
- Build effort, and who maintains the models after launch

Phase 0 Audit and definitions | **Phase 1** Orders and margin | **Phase 2** Marketing and CAC | **Phase 3** Retention and customer view | **Phase 4** Forecasting and activation

0 Audit and definitions

Agree what the numbers mean, and what they must tie to, before anything is built.

DELIVERABLES

- Source inventory: system, owner, access, history available
- Metric dictionary draft for the core metrics
- A tie-out month with finance's target totals
- A pick per layer, and the first three questions to answer

OWNERS

- Finance: revenue and margin definitions
- Growth: CAC and retention definitions
- Data lead (in house or partner): inventory and stack
- An executive sponsor breaks ties

ACCEPTANCE CHECKS

- Every metric on today's main dashboards maps to one entry with an owner
- Tie-out month and totals written down
- Access confirmed for every source in Phase 1

RISKS

- Definitions debated without end: the owner decides, with a date
- Sources found late: ask finance and ops, not only marketing
- Scope grows to every report: start with three questions

1 Orders and margin

One order table across channels, with refunds, fees, and costs, that ties to the books.

DELIVERABLES

- Connectors for storefronts, marketplaces, payments, and fulfillment
- Order, refund, and fee models; a SKU map; a dated cost table
- Daily KPIs by channel and contribution margin per order

OWNERS

- Data lead builds the models
- Finance signs off the tie-out
- Operations owns the SKU map and product costs

ACCEPTANCE CHECKS

- Net sales by channel match source reports for the tie-out month, within a tolerance finance sets
- No order counted twice; refunds follow the dictionary's date rule
- Margin reconciles to the ledger for the closed month

RISKS

- Marketplace fees settle later: label estimated and settled
- No cost history: use a dated table and flag estimates
- Time zones differ by system: convert to one in the warehouse

2 Marketing and CAC

Spend next to new customers, with CAC defined once and margin after marketing.

DELIVERABLES

- Daily ad spend by platform, campaign, and channel
- Blended and attributed CAC, labeled differently
- Contribution after marketing by channel
- Packaged attribution data loaded, if marketing uses a tool

OWNERS

- Growth owns CAC definitions
- Finance decides which costs count as acquisition
- Data lead builds and maintains the models

ACCEPTANCE CHECKS

- Spend per platform matches platform billing for the month
- New customers in CAC equal new customers in the order model
- Attributed credits are labeled; they can sum above the total

RISKS

- Platform attribution read as revenue
- Agency fees and creative costs missing from spend
- Currency and time zones differ by ad account

3 Retention and customer view

One ID per person, cohorts that are fair to new customers, and one customer timeline.

DELIVERABLES

- Identity table joining storefront, email, subscription, and support IDs
- Cohorts, 90-day repeat rate, frequency, and reactivation
- LTV on contribution margin by acquisition channel

OWNERS

- Growth owns retention definitions
- Finance co-owns LTV
- Data lead owns identity rules

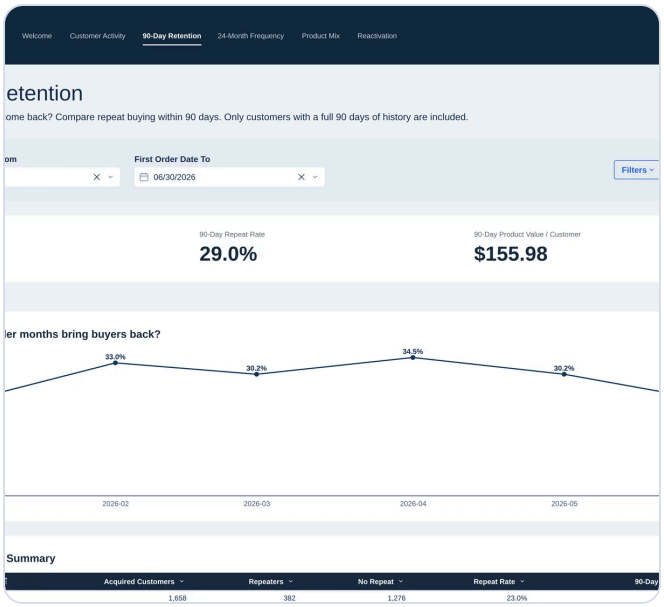
ACCEPTANCE CHECKS

- Duplicate customers checked on a sample of records
- Repeat rates count only customers who had the full window
- Channel LTV sums back to cohort contribution

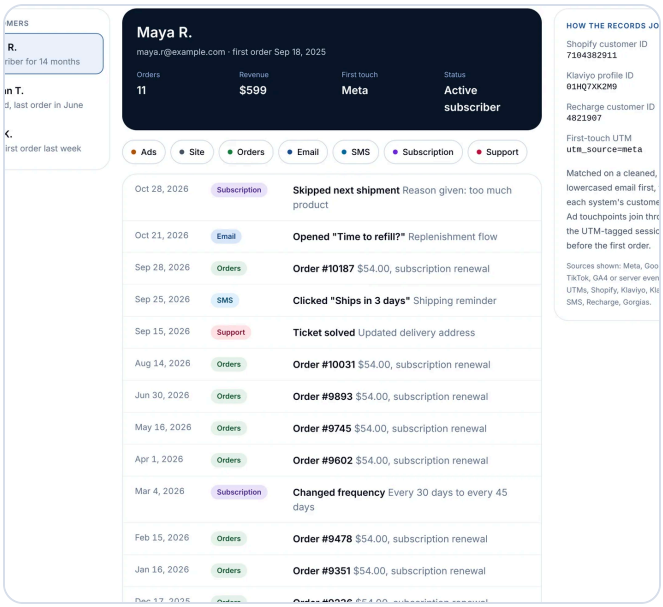
RISKS

- Marketplaces share little buyer identity: report them separately
- Subscription renewals inflate voluntary repeat rate
- Guest checkouts split one person into two

WHAT PHASE 3 PRODUCES



A 90-day retention report in Sigma that only counts eligible customers. Synthetic data for Willow, a made-up brand.



One customer timeline across Shopify, Klaviyo, Recharge, and support. Synthetic data for Brightleaf, a made-up brand.

4 Forecasting and activation

Plan demand on trusted history, and send warehouse audiences back to the tools that act.

DELIVERABLES

- Demand forecast by SKU and channel, with an inventory and reorder view
- Plans entered next to actuals, for example in Sigma input tables
- Audiences and scores synced to Klaviyo and ad platforms

OWNERS

- Operations or planning owns the forecast
- Growth owns audiences and their use
- Data lead maintains syncs and monitoring

ACCEPTANCE CHECKS

- Forecast backtested against a held-out period
- Plan versus actual uses the same dictionary metrics
- Synced audience counts match warehouse counts

RISKS

- Stockouts in history read as low demand
- Audiences synced before identity is trusted
- Consent and suppression fields not carried through

08

Where to start, and a readiness checklist

If this describes you today	Start with
Finance, marketing, and Shopify do not agree on last month's revenue, or nobody has checked	Phase 0
Platform reports or a packaged tool, and the revenue gaps are not written down	Phase 0
Platform reports or a packaged tool, and teams already agree on revenue	Phase 1
Raw data already loads into a warehouse, but little is modeled	Phase 1
Orders and margin are modeled and trusted	Phase 2, or 3 if retention is the main question

READY TO START WHEN

- [ ] **A sponsor is named**
  Someone senior can break ties between finance and marketing.
- [ ] **Each core metric has an owner**
  Net sales, contribution margin, CAC, repeat rate, and LTV each have one person who approves changes.
- [ ] **A tie-out month is chosen**
  A closed month, with finance's totals for it, that every new number must match.
- [ ] **Sources are listed with access**
  Every system, who owns the login, and how much history it holds.
- [ ] **Product costs have history**
  Costs by SKU and date, not only today's cost.
- [ ] **The first three questions are written**
  Specific decisions the first reports will support.
- [ ] **Someone will maintain it after launch**
  An in-house owner or a partner for models, connectors, and monitoring.

09

Common failure modes

Most stalled data projects fail in one of these ways. Each has a fix that costs less than the problem.

1. **Dashboards before definitions**
   Reports get built, then each team disputes them. Fix: Phase 0 first, and a dictionary entry behind every number on a dashboard.

2. **Connecting everything at once**
   Twenty connectors run, nothing is modeled, and the bill grows. Fix: load what Phase 1 needs, then add sources per phase.

3. **Platform attribution read as revenue**
   Ad platforms credit the same order more than once. Fix: label attributed numbers, and report revenue from the order model.

4. **Orders counted twice**
   An order reaches the warehouse from two systems or two connectors. Fix: deduplicate in silver and test for it.

5. **Today's cost on last year's orders**
   A single cost field rewrites history every time a supplier price changes. Fix: a dated cost table.

6. **Fees taken at the estimate**
   Marketplace fees settle after the sale, so early margin is wrong. Fix: restate when fees settle, and label estimated periods.

7. **Two sets of definitions, unlabeled**
   A packaged tool and the warehouse both report "revenue" differently. Fix: name each version in the dictionary.

8. **Activation before identity**
   Audiences sync before one person is one customer, so targeting misfires. Fix: Phase 3 acceptance checks before Phase 4.

9. **No owner after launch**
   Connectors break quietly and numbers drift. Fix: freshness alerts, row count checks, and a named maintainer.

Related: why reliable ecommerce reporting takes time.

GREYFIELD DATA

Want help choosing or building your stack?

Greyfield Data is Linus Tse's analytics consultancy. I build reporting for consumer brands on modern data warehouses: the connectors, the order, margin, and customer models underneath, and the dashboards people open every week, often in Sigma.

If your channel numbers do not agree, a short call is a good place to start. Bring the reports you use today and the question they cannot answer.

Book a call with Linus | **greyfielddata.com**

Get your stack
greyfielddata.com/tools/ecommerce-data-stack-quiz

See the map
The ecommerce data stack, mapped: 57 tools

Compare tools
Packaged analytics and BI tools, compared with sources

Screenshots use synthetic data for made-up brands. Vendor details summarize Greyfield's comparison pages as of November 2026; plans and features change, so confirm with each vendor. Product names belong to their owners. Greyfield Data builds warehouses and sells Sigma consulting.