



Working checklist

Shopify KPI Checklist

Reliable data, formulas and operational decisions for reading a Shopify store without mixing up the numbers.

Francesco Guiducci

Edition October 2026 · English

ifgdigital.com

For personal or internal business use; unauthorized distribution is prohibited.

Contents

How to use this checklist	3
Open a review	3
K01 · Sources, access and definitions (P0)	3
K02 · Test events before you interpret them (P0)	4
K03 · Conversion: the right denominator (P1)	5
K04 · Cart and checkout abandonment (P1)	5
K05 · AOV and order composition (P1)	5
K06 · Traffic, engagement and sources (P1)	6
K07 · CAC, CPA, ROAS and margin (P0)	6
K08 · Repeat purchases, cohorts and customer value (P1)	7
K09 · Returns, refunds and feedback (P1)	8
K10 · Interpret changes and choose experiments (P1/P2)	8
K11 · Routine and decision log	9
K12 · Close the review	9

How to use this checklist

When a metric changes, the first question is: what am I comparing? Two values called "conversion" can have different denominators, and rising revenue can sit alongside falling margin. Use this checklist to verify the data before you decide, and to turn every analysis into an action you can review.

Use it for the weekly or monthly review, after any tracking change and before you raise spend. The formulas and examples let you work without a custom report. The examples are illustrative: they are not industry benchmarks or forecasts.

Open a review

Field	To complete
Store, channels and markets included	...
Period analysed and comparison	...
Currency, time zone and extraction date	...
Owner of the analysis and of the decision	...
Business question to resolve	...
Changes to campaigns, prices, stock or site	...

Give each of K01–K12 one of these outcomes: **OK**, **NEEDS FIXING**, **NOT VERIFIED**, **NOT APPLICABLE**. OK means a definition, evidence and a consistent result all exist. PO marks errors that make financial decisions unreliable. P1 marks interpretation or measurement problems to resolve before you act. P2 marks later analysis and improvements.

Choose periods that match in length and weekday mix. For seasonal businesses, add the same period of the previous year when you have it. Note sales, holidays, market launches, stockouts and consent changes. Do not compare a full month with a few days that are still updating.

K01 - Sources, access and definitions (P0)

Open the actual report and read the field definition. For orders and sales, use a declared transactional source; for on-site behaviour, pick the tool that fits the question. Shopify and GA4 can differ on sessions, attribution, consent and processing time: do not "fix" one number just because it differs from the other.

The main Shopify Analytics features are available on every Shopify plan. Check the permissions and features you actually have; do not assume you need Advanced or Plus. In large reports, make sure the export includes every row you need, not only the visible ones.

Fill in a small dictionary before you calculate:

Metric	Definition and formula	Source and filters	Owner
Online store conversion	...	...	...
AOV	...	...	...
CAC	...	...	...
Contribution after acquisition	...	...	...

- ☐ Orders, customers, users, sessions and events are kept distinct.
- ☐ Tax, shipping, discounts, refunds and test orders have a stated treatment.
- ☐ Currency and time zone are the same across periods.
- ☐ Saved reports keep their filters and date range.

Done K01. Another person can rebuild the value from the same data. If a definition changes, start a new series or annotate the break.

K02 - Test events before you interpret them (P0)

For GA4, check the current implementation: the Google & YouTube channel lets you connect a property. Old instructions based on Universal Analytics, or on code pasted into legacy thank-you-page settings, are not a procedure to reuse.

Inventory your tags: which integration sends each event, on which surfaces and under which consent conditions. Look for overlaps between channels, custom pixels, apps and theme code. Name one owner for sending each event; any double sending needs deduplication that has been designed and tested.

Walk through a journey with test data: product page, add to cart, checkout start, purchase and refund where supported. In the debug tools, look at the parameters actually sent. A purchase needs a consistent transaction ID, value, currency and items, following the schema you have adopted. The GA4 ecommerce documentation describes the events.

- ☐ Reloading the confirmation page does not create a duplicate purchase.
- ☐ Currency and value match the chosen definition.
- ☐ A failed payment is not counted as a purchase.
- ☐ Consent acceptance, refusal and changes have been tested.
- ☐ Email, phone number and other unintended personal data do not end up in parameters.
- ☐ Refunds are observed separately from the purchase event.

Keep the test ID, the expected and observed events, the limits of the method and the date. Technical tests do not prove that aggregate reports are complete: after processing, check a sample in the relevant reports too, keeping in mind that some simulated orders may be excluded.

K03 - Conversion: the right denominator (P1)

For Shopify online store conversion, use sessions that completed checkout $\div$ sessions $\times$ 100. Do not divide all orders, including POS or manually created ones, by web sessions. One person can open several sessions, and an event count is not a headcount.

Example: 10,000 sessions and 200 sessions with a completed checkout give 2%. The next period, 12,000 sessions and 300 completions give 2.5%: up 0.5 percentage points, or 25% in relative terms. Purchases also rise because sessions rose, so do not credit everything to the new button.

If the rate falls, work in this order: tracking quality, errors in the journey, availability and prices, traffic mix. Then segment by device, market, source and landing page. More informational visits can pull the average down without any worsening among people looking for a specific product.

Check. Numerator ... / denominator ... / rate ... / segment ... / main difference in mix Do not average the daily rates: add up consistent numerators and denominators, then recalculate the ratio.

K04 - Cart and checkout abandonment (P1)

Define the start and end points. "Abandonment" is not a single metric: it can describe sessions with a cart and no purchase, checkouts started and not completed, or a group of users followed over time.

For a sequential, session-based funnel: checkout abandonment = $1 - \text{completing sessions in the group} \div \text{checkout-starting sessions in the group}$. Use the same group and the same window. Do not use raw add_to_cart events as the denominator: one session can add five products.

Example: 500 sessions enter checkout and 350 of them complete within the defined journey. Completion is 70%, abandonment 30%. If the report uses an open funnel, or accelerated checkouts skip steps, the counts may not nest: read the definition before you subtract.

- ☐ The basis, sessions or users, is stated and consistent.
- ☐ Open or closed funnel, and the window, are explicit.
- ☐ Shipping, final total, errors and payment methods have been checked.
- ☐ A recovery attributed to email is not presented as an incremental sale.

Record the problem stage and reproduce at least one case. If you cannot reproduce it, keep it as a hypothesis and gather more evidence: a blanket discount is not the first response.

K05 - AOV and order composition (P1)

Shopify's AOV is (gross sales – discounts) $\div$ orders, excluding adjustments made after the order. It is not the total collected including tax and shipping. For details, use the definition in Shopify's sales report.

Example: €12,000 gross sales, €2,000 discounts and 200 orders give an AOV of €50. Tax and shipping are not part of the formula. To measure economic value after returns and costs, build a separate metric with an explicit name and formula.

Before you introduce bundles or thresholds, look at order distribution, units per order, product mix, discounts and margin. The average can rise because many small orders disappear. A threshold far above the usual cart can push people to give up rather than add an item.

To compare. AOV ... / orders ... / margin per order ... / conversion ... / returns Close the check only if the interpretation covers order volume and whether the offer is sustainable.

K06 - Traffic, engagement and sources (P1)

In GA4, bounce rate is the share of sessions that were not engaged. An engaged session meets at least one of these conditions: it lasts longer than the configured time threshold, it records a key event, or it has at least two page or screen views. The default threshold is 10 seconds. Do not apply the old Universal Analytics definition, which was based on a single pageview.

An unexpected value calls for an event check: duplicate page_view events distort engagement. To judge traffic quality, cross landing page, source, device and useful action. An article can answer a question in one visit; no immediate sale does not make it useless.

"Direct" does not mean loyal customers typing the domain: it can include visits with no attribution available. Check campaign parameters and cross-domain hops. For organic queries, use Search Console: GA4 acquisition reports do not return every search term.

- ☐ Campaigns use consistent UTM naming.
- ☐ Unknown sources are analysed without inventing an origin.
- ☐ Volume and quality are read together.
- ☐ Heatmaps have a specific purpose, masking and appropriate data handling.

Write down an observable question, for example: "can mobile users find the size selector?" A session recording can explain an obstacle, but on its own it does not tell you how widespread it is.

K07 - CAC, CPA, ROAS and margin (PO)

Keep three ratios apart:

Metric	Formula	What to clarify
CAC	Defined acquisition costs ÷ new customers	Costs included, and customers from the same scope
Advertising CPA	Spend ÷ acquisitions attributed by the platform	Which event and which window
ROAS	Attributed revenue ÷ ad spend	Revenue definition and attribution model

An attributed purchase can come from a returning customer, so CPA and CAC are not interchangeable. For blended CAC, include the acquisition costs you have chosen, such as media, production and attributable labour, without counting any twice. For channel CAC, you need new customers attributed consistently to that channel: do not use all of the store's new customers for every platform.

Example: €1,200 of defined costs and 40 new customers give a CAC of €30. If the platform attributes 60 purchases to the same spend, its CPA is €20. That does not show you are acquiring a new customer for €20.

Check sustainability

With €5,000 of attributed revenue and €1,000 of advertising, ROAS is 5. If the margin after variable costs, before advertising, is 50%, €2,500 remains; subtract the ads and €1,500 remains. That is residual contribution, not net profit, because fixed costs, taxes and other items are missing.

Break-even ROAS = $1 \div \text{pre-ad margin percentage}$ holds only if that margin includes all relevant variable costs and revenue is on the same basis. At a 40% margin the theoretical ad break-even is 2.5. It is not a universal target and does not cover fixed costs on its own.

- ☐ Attributed revenue is not summed across platforms that may claim the same order.
- ☐ Windows, models, currency and refund treatment are stated.
- ☐ The proposed budget is judged on margin, not only on ROAS.
- ☐ Costs that are still estimates and delayed returns are visible.

Before you raise spend, write down. Expected margin ... / CAC limit ... / payback period ... / stop condition A high ROAS on campaigns aimed at existing customers does not show that new-customer acquisition works.

K08 - Repeat purchases, cohorts and customer value (P1)

Separate the share of returning customers in a period from a cohort's repeat-purchase rate. A cohort groups the customers whose first purchase fell in the same interval: you can follow them after one, three or six months, using equal windows.

Example: 100 new customers in January, 25 of them with at least a second purchase within 90 days: a 90-day repeat rate of 25%. Compare it with another cohort that has had the full 90 days. A recent cohort is not worse just because it has not yet had time to buy again.

Revenue per customer observed over 12 months is not automatically lifetime value. CLV estimates need assumptions about lifespan, frequency, margin, churn and the timing of cash. Label observed data and forecasts separately.

- ☐ Cohort, window and the rule for a new customer are stated.
- ☐ Orders and customers are deduplicated consistently.
- ☐ The CAC versus customer value comparison uses compatible margins and timing.
- ☐ Repeat-purchase emails and incentives are judged on cost, returns and unsubscribes too.

A useful decision: what reorder interval shows up in the data, and what message could be relevant? Do not send every customer the same sequence without considering the product bought and its usage cycle.

K09 - Returns, refunds and feedback (P1)

Define the return rate by units or by orders. By units: units returned ÷ units sold in the same cohort. By orders: orders with at least one return ÷ orders in the cohort. A monetary refund does not always mean a physical return, and several partial refunds are not several orders.

Wait long enough for returns to arrive and record the lag. Segment by product, variant, reason and market. Before you cut returns with a policy change, check sizing, descriptions, damage, picking errors and delivery promises.

For NPS, use one consistent 0–10 question: percentage of promoters (9–10) – percentage of detractors (0–6); 7–8 are passives. The result runs from –100 to +100 and is not the same as an average rating. Bain describes the definition in the Net Promoter System.

Example: 50 responses, 30 promoters, 10 passives and 10 detractors give NPS 60% – 20% = 40. Also record invitations, response rate and when the survey was sent. A sample made up only of very satisfied customers does not represent the whole customer base.

Evidence required: denominators, recurring reasons and at least one action tied to a concrete problem. Where the questions allow, keep product quality, delivery and support separate.

K10 - Interpret changes and choose experiments (P1/P2)

Use a short sequence: observation → hypothesis → evidence → decision. "Mobile conversion fell" is an observation; "the new size selector creates friction" is a hypothesis; reproducing the error or running an experiment provides the evidence.

Before you change anything, write down the primary metric, the minimum useful improvement, the metrics that must not get worse, the audience and the duration. With low traffic, favour fixing reproducible defects and qualitative research; do not call a variant the winner after a few purchases.

An A/B test needs consistent assignment, comparable groups, a suitable tool and an analysis plan. Do not stop it the moment a favourable result appears. Google Optimize was discontinued on 30 September 2023: do not use it as a current operating instruction.

Three diagnostic examples:

- **Falling mobile conversion.** Check errors and traffic mix; test variants and checkout on the affected device before you propose a new theme.
- **AOV up, margin down.** Separate discounts, products sold, shipping costs and returns: a threshold may have rewarded orders that were coming anyway.
- **Rising CAC.** Check costs, new customers, attribution windows and stock; then assess creative, audience and landing page.

K11 - Routine and decision log

On each working day, check the anomalies that need a fast response: payments, orders, unexpected revenue and event outages. The weekly review covers funnel, campaigns, stock and reported issues. The monthly review brings together margin, costs, mature cohorts and the results of changes. Adjust the frequency to your volume, keeping one named owner for each check.

Observation with data	Hypothesis or evidence	Action and owner	Review date and criterion
...	...	...	...
...	...	...	...

Save enough extracts and definitions to reproduce the analysis, with appropriate access and retention. Avoid customer-level copies in shared spreadsheets when totals or test identifiers are enough. In email reports, remember the limits of open rates and read clicks, purchases and unsubscribes too, using the platform's definitions.

K12 - Close the review

Area	Outcome and evidence	Open point	Owner and deadline
K01-K02: definitions and collection	...	...	...
K03-K06: funnel and traffic	...	...	...
K07-K09: economics and customers	...	...	...
K10-K11: experiments and routine	...	...	...

If the checks give different outcomes, split the rows.

Closing the review. No unresolved PO in the data used to decide; reproducible formulas; attribution and sample limits stated; a decision tied to the initial question; an owner and a review date. Not every KPI has to improve in the same period.

Decision made ... Data supporting it ... Remaining material uncertainty ... What would change the decision ...

Sources

- Shopify Help Center, Shopify reports: <https://help.shopify.com/en/manual/reports-and-analytics/shopify-reports>
- Shopify Help Center, sales report: <https://help.shopify.com/en/manual/reports-and-analytics/shopify-reports/report-types/default-reports/sales-report>
- Shopify Help Center, Google & YouTube channel: <https://help.shopify.com/en/manual/online-sales-channels/marketplaces/google/getting-setup/connect>
- Google Analytics, GA4 ecommerce events: <https://developers.google.com/analytics/devguides/collection/ga4/ecommerce>
- Google Analytics, engagement and bounce rate: <https://support.google.com/analytics/answer/12195621?hl=en>
- Google Analytics, Google Optimize sunset: <https://support.google.com/analytics/answer/12979939?hl=en>
- Bain, Net Promoter System: <https://www.netpromotersystem.com/about/measuring-your-net-promoter-score/>