Retail Transaction Audit

Specialty gift & homeware retailer, UK · figures in USD · 541,909 transaction lines · 25,900 invoices · 1 Dec 2010 – 9 Dec 2011

$2,815,990

of a year's sales cannot be tied to any customer. 16.5% of gross sales, spread across 3,710 invoices, carries no customer identifier — so a sixth of the business cannot be segmented, retained, or valued.

Measured, not estimated. Gross sales $17,112,562 · net of returns $15,638,312.

GROSS SALES	NET SALES	IDENTIFIED CUSTOMERS	REPORTED RETURN RATE	REVENUE IN TOP 5%
$17.11M	**$15.64M**	**4,372**	**8.6%**	**48.6%**
530,104 positive lines	after all reversals	38 countries	actual rate is 2.2%	219 customers

TOP 5 FINDINGS

1 $2,815,990 of sales has no customer attached to it

135,080 transaction lines across **3,710 invoices** carry no customer identifier — **16.5% of gross sales**. Every downstream customer metric in this business is computed on the remaining 83.5% and then described as if it covered the whole. Lifetime value, repeat rate, and segment mix are all understated by an unknown amount, because nobody knows whether those 3,710 invoices belong to new buyers or to existing ones ordering again.

> **ACTION** Capture an identifier on every order before anything else on this list. Until then, treat every customer figure in the business as covering five-sixths of it.

2 The return rate is reported at 8.6%. It is actually 2.2%

$1,474,250 of negative lines sits in the ledger. Only **$373,916** of it is a customer returning a product. **$706,233** is non-product adjustment — marketplace fees, manual corrections, bad debt, bank charges — and **$394,101** is two same-day order-and-cancel pairs that never shipped. Anyone reading the ledger at face value is managing a returns problem that is roughly a quarter the size they think it is, and missing a $706,233 adjustment problem entirely.

> **ACTION** Move fees, manual corrections and bad debt out of the returns line into their own adjustment ledger. Reconciliation on page 2.

3 44 customers carry 30% of the revenue — and 63 of the top accounts have gone quiet

Of $13,315,796 in identified net revenue, the top 1% of customers (**44 accounts**) produce 29.9% and the top 5% (219 accounts) produce 48.6%. Within the top 20% by value, **63 customers have not ordered in 90 days or more** — median 138 days silent — representing **$375,061** of trailing-year revenue with no one assigned to it.

> **ACTION** Pull those 63 accounts as a named list this week. At this concentration, losing four of them costs more than a bad quarter of new-customer acquisition.

4 A third of customers buy once and never come back

1,494 customers (34.2%) placed exactly one order all year. A customer who orders a second time is worth **7.8× more** over the year — $4,392 against $560. The gap is not gradual; it opens at order two. Getting **10% of the one-time group to place a single additional order** at their own average basket is worth **$83,377**, using no assumption beyond what those customers already spent once.

> **ACTION** A second-order campaign against 1,494 named buyers is the cheapest revenue on this page. Measure it on second-order rate, not on opens.

5 $335,137 of the catalogue's own attach pattern is going unsold

Across **66 product families** that customers demonstrably buy as sets — colour and scent variants of the same item — baskets routinely arrive holding every member but one. Where the pattern is strongest, the pull is near-mechanical: a basket with the green and roses teacup contains the pink one **82%** of the time (lift 21.7×). Summing the missing member across all 66 families at its own average basket value gives **$335,137** of demand the catalogue is already generating and not capturing.

> **ACTION** Start with the eight largest families — $174,562 of the $335,137. A shelf-adjacency and basket-prompt fix, not a discount.

WHERE TO START — IN ORDER

1	**Capture a customer identifier on every order** Nothing else here can be measured properly while 16.5% of sales is anonymous. A checkout change, not an analytics project.	WEEK 1 · FIX
2	**Split adjustments out of the returns line** $706,233 of fees and corrections is indistinguishable from customers sending goods back. Both numbers become usable the moment they are separated.	WEEK 1 · FIX
3	**Work the 63 lapsed top-tier accounts** $375,061 of trailing-year revenue, a named list, no new data required. Highest return per hour on this page.	WEEK 2 · REVENUE
4	**Run a second-order campaign at the 1,494 one-time buyers** $83,377 at a conservative 10% conversion — and a second order is where the 7.8× value gap begins.	WEEK 3 · REVENUE
5	**Fix attach on the eight largest variant families** $174,562 of the $335,137 sits in eight families. Merchandising first; test before extending to the other 58.	MONTH 2 · REVENUE

EVIDENCE

RECONCILIATION — WHAT THE $1,474,250 OF NEGATIVE LINES ACTUALLY IS

COMPONENT	LINES	VALUE	SHARE
Non-product adjustments (marketplace fee $377,462 · manual $235,486 · bad debt $35,494 · other $57,791)	586	$706,233	47.9%
Two same-day order/cancel pairs (80,995 units cancelled 12 min after ordering; 74,215 units cancelled 16 min after)	2	$394,101	26.7%
Genuine product returns	8,702	$373,916	25.4%
Total negative lines	**9,290**	**$1,474,250**	**100%**

True return rate = $373,916 ÷ $16,718,461 clean gross sales = **2.2%**. The 8.6% headline divides all negative lines by unadjusted gross, counting fees and a cancelled order that never shipped as customer returns.

REVENUE CONCENTRATION

CUSTOMER TIER	ACCTS	NET REVENUE	SHARE
Top 1%	44	$3,984,751	29.9%
Top 5%	219	$6,470,254	48.6%
Top 10%	437	$7,990,401	60.0%
Top 20%	874	$9,830,112	73.8%
All identified	**4,372**	**$13,315,796**	**100%**

Excludes the $2,815,990 of anonymous sales.

ONE ORDER VS. MORE THAN ONE

GROUP	CUSTS	AVG NET	AVG ORDERS
Ordered once	1,494	$560	1.0
Ordered 2+ times	2,845	$4,392	6.0
Value multiple	**—**	**7.8×**	**—**

$83,377 assumes 149 of the 1,494 place one further order at $560 — their own average, not the repeat-buyer average.

LARGEST ATTACH GAPS — VARIANT FAMILIES BOUGHT AS SETS (TOP 3 OF 66)

PRODUCT FAMILY	VARIANTS	BASKETS WITH 2+	ATTACH GAP
Picnic basket, wicker	2	122	$43,956
Heart card holder	2	81	$35,329
Harmonica in box	2	214	$20,586
All 66 qualifying families	**—**	**—**	**$335,137**

Attach gap = baskets holding every family member but one, × that item's average value when present. Strongest pair: green + roses Regency teacup predicts the pink at 82% (lift 21.7×).

METHOD, CAVEATS & SOURCE

Source. Online Retail, UCI Machine Learning Repository (dataset 352) — a real UK specialty gift retailer's complete transaction log, 1 Dec 2010 – 9 Dec 2011, under Creative Commons Attribution 4.0. No client data appears here. Every figure reconciles: positive lines $17,112,562 less negative lines $1,474,250 equals net $15,638,312, matching the sum of all 541,909 rows exactly.

Currency. The source records sterling; figures here are converted at **$1.6043 per £1**, the 2011 annual average published by the Federal Reserve (G.5A, 3 Jan 2012). Every percentage, ratio and rank is unaffected by the conversion.

Excluded, and why. Eleven non-product stock codes (fees, manual, bad debt, postage, carriage, commission, bank charges, discount, samples, vouchers, pads) are held out of product-level analysis and reported separately. Two order/cancel pairs above 10,000 units are held out of gross sales and returns alike; left in, they overstate both by $394,101 and move the return rate 6.4 points.

What this cannot tell you. No cost or margin field exists, so every figure is revenue, never profit — the $335,137 attach gap is gross sales, not contribution. Customer findings cover only the 83.5% of sales carrying an identifier. Basket affinity uses 16,782 invoices holding 2–60 items; nothing is ranked on under 40 baskets.