Belle Curve Insights

PUBLIC DATA

Transaction Analysis

POINT OF SALE · WHAT THE BASKET SHOWS

A UK specialty gift retailer · 541,909 transaction lines, 1 Dec 2010 – 9 Dec 2011 · figures in USD · prepared by Sandra Stokke

$2,815,990

of a year's sales **cannot be tied to any customer** — 16.5% of gross, spread across 3,710 invoices.

$17.11M	$15.64M	4,372	8.6%
GROSS SALES · 530,104 POSITIVE LINES	NET SALES · AFTER ALL REVERSALS	IDENTIFIED CUSTOMERS · 38 COUNTRIES	REPORTED RETURN RATE · ACTUAL IS 2.2%

A method sample, not a client result. This runs on a real retailer's complete transaction log, published under Creative Commons Attribution 4.0, so every number here can be checked against the source rather than taken on trust. It also shows the limit of point-of-sale data honestly: the file carries no cost field, so every figure is revenue and none of it is profit.

FIVE FINDINGS — AND WHAT TO DO ABOUT EACH

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 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 invoices belong to new buyers or to existing ones ordering again.

→ 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 roughly a quarter the size they think it is, and missing a $706,233 adjustment problem entirely.

→ Move fees, manual corrections and bad debt out of the returns line into their own adjustment ledger. Both numbers become usable the moment they are separated.

3 44 accounts 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** — a median 138 days silent — representing **$375,061** of trailing-year revenue with no assignee.

→ 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 is worth $83,377 at their own average basket — assuming nothing beyond what those customers already spent once.

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 **82%** of the time (lift 21.7×).

→ 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 one.**
2. **Split adjustments out of the returns line.** $706,233 of fees and corrections is currently indistinguishable from customers sending goods back. Both numbers become usable the moment they are separated. **Week one.**
3. **Work the 63 lapsed top-tier accounts.** $375,061 of trailing-year revenue, a named list, no new data required. The highest return per hour on this page. **Week two.**
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 three.**
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 two.**

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

Component	Lines	Value	Share
Non-product adjustments — marketplace fees, manual, bad debt, other	586	$706,233	47.9%
Two same-day order/cancel pairs that never shipped	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 = 2.2% — $373,916 against $16,718,461 of clean gross sales. The 8.6% headline divides every negative line by unadjusted gross, counting marketplace fees and a cancelled order that never shipped as customer returns.

WHERE THE REVENUE SITS · IDENTIFIED CUSTOMERS ONLY

Customer tier	Accounts	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 entirely. **One order versus more than one:** 1,494 customers ordered once, averaging $560; 2,845 ordered twice or more, averaging $4,392 across six orders — a 7.8× value multiple. The $83,377 assumes 149 of the 1,494 place one further order at their own average, not the repeat-buyer average.

LARGEST ATTACH GAPS · TOP 3 OF 66 VARIANT FAMILIES

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, multiplied by that item's average value when present.

WHAT THIS ANALYSIS CANNOT SUPPORT

— Any statement of profit or margin — the file holds no cost field, so the $335,137 attach gap is gross sales, not contribution.
— Customer conclusions covering the whole business — they cover the 83.5% of sales that carries an identifier.
— Basket affinity below 40 baskets: nothing is ranked under that floor. Affinity uses the 16,782 invoices holding 2–60 items.

Everything here 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. Eleven non-product stock codes are held out of product-level analysis and reported separately; the two order/cancel pairs above 10,000 units are held out of gross sales and returns alike, because leaving them in overstates both by $394,101 and moves the return rate 6.4 points.

Your transaction file will answer the same five questions.

One export from your point-of-sale system is usually enough to start. sandra@bellecurveinsights.com · 239-469-4021

Source: Online Retail, UCI Machine Learning Repository dataset 352 — a UK specialty gift retailer's complete transaction log, 1 December 2010 – 9 December 2011, under Creative Commons Attribution 4.0. No client data appears here. Sterling converted at $1.6043 per £1 (Federal Reserve G.5A, 3 January 2012). Excluded and reported separately: eleven non-product stock codes — fees, manual, bad debt, postage, carriage, commission, bank charges, discount, samples, vouchers, pads. Figures are unaudited. A method sample, not a client result. Not tax, legal or accounting advice.