Customer Lifecycle & Campaign Analysis

RFM Segmentation · Churn Prediction · Attribution · Customer Lifetime Value

Amrit Sharma | MSc Business Analytics | Durham University | 2025

Agenda

The Business Problem

The UK E-Commerce Reality

1. ~1 million transactions across 2 years of trading data
2. No segmentation — every customer gets the same email
3. No early warning system for customers about to leave
4. No channel attribution — unknown which marketing drives revenue
5. No understanding of which customers are worth fighting for

01 EDA & Cleaning
Understand and trust the data

02 RFM Segments
Know who each customer is

03 Churn Model
Predict who is about to leave

04 Attribution
Know what marketing works

05 CLV Model
Quantify forward revenue value

Data & Cleaning — From Messy to Trusted

1,067,371
Raw rows loaded

287,946
Rows removed (27%)

779,425
Clean rows kept

5,878
Unique customers

Step 1
Duplicates removed
34,335 exact duplicate rows — system export error

Step 2
No Customer ID
235,151 anonymous rows — cannot be linked to a customer

Step 3
Cancellations (C...)
18,390 reversed invoices — no money changed hands

Step 4
Invalid qty/price
70 rows with zero or negative values

Key EDA Findings: Peak trading Thursday 11am–1pm · Q4 seasonal revenue spike · Top intl markets: Netherlands, EIRE · Heavy Pareto revenue skew → justifies RFM segmentation

RFM Segmentation — 7 Behavioural Segments

R Recency

F Frequency

M Monetary

Each scored 1–5 using pd.qcut (quantile bins) so scores are always relative to the full customer base.

Segment	Count	RFM Score	Campaign Strategy
Champions	1,482	R4–5, F4–5	VIP — bought recently, buy often, spend most
Loyal Customers	1,221	R3+, F3+	Regular buyers — solid second tier
Need Attention	551	R2, F3+	Declining — starting to drift away
At Risk	89	R1, F4+	Used to buy often — have not returned
Lost	1,523	R1–2, F1–2	Effectively gone — suppress from paid ads
Potential Loyal	828	R3+, F1–2	Recent first-time buyers — nurture to repeat
New Customers	184	varies, F1	Single purchase only — welcome series

Churn Prediction Model

Churn Definition: no purchase in last 90 days. A business decision, not a statistical one. 90 days suits a seasonal gift retailer.

Logistic Regression

SELECTED

AUC 0.777

Interpretable — can explain each prediction to a non-technical stakeholder. Coefficients show exact feature weights.

Random Forest

AUC 0.79

Higher accuracy, captures non-linear patterns. Used as comparison benchmark.

Key Predictive Features

- Recency score
- Avg days between purchases
- Frequency score
- Revenue trend (recent vs earlier)
- Basket size
- Product categories purchased

Campaign Attribution — 3 Model Comparison

First-Touch

Credit:

100% first channel

Best for:

Awareness campaigns

Limitation:

Ignores closing channels

Last-Touch

Credit:

100% final channel

Best for:

Conversion campaigns

Limitation:

Ignores awareness spend

Linear

Credit:

Equal across all

Best for:

Balanced view

Limitation:

Treats all touchpoints equal

Google Ads Measurement Certification connection: these three models are exactly what GA4 attribution covers natively. Replicating them in Python demonstrates understanding of the underlying mechanics, not just dashboard usage.

Note: Channel touchpoints simulated using RFM proxies (methodology demonstration). Production version uses GA4 API event-level data.

Notebook 05: Customer Lifetime Value Model

Survival-Weighted Discounted Cash Flow · 5,878 Customers · 12M & 24M Horizons

"The churn model told us who is leaving. The CLV model tells us what it actually costs — in forward revenue."

CLV Methodology — Survival-Weighted DCF

$$CLV = \Sigma \; [\; AOV \times purchase\_rate \times P(alive\ at\ month\ m) \;] \;/\; (1 + r)^m$$

Step 1: AOV

Monetary ÷ Frequency
Mean spend per transaction

Step 2: Purchase Rate

Frequency ÷ active_days × 30
Monthly purchase frequency

Step 3: Survival Decay

P(alive) × (1 − monthly_churn)
Decays each month
compounding

Step 4: Discounting

÷ (1 + 0.83%)^m
10% annual rate, standard
convention

Churn conversion: 90-day model output → monthly rate: P(churn_monthly) = 1 − (1 − Churn_Probability)^(1/3) | Prevents 3× overstatement of monthly risk

Production note: BG/NBD + Gamma-Gamma (lifetimes library) would be used on raw transaction data. This survival-DCF demonstrates understanding of the underlying model.

CLV Results — Portfolio Overview

£6.7M
Total 12-Month CLV
5,878 customers

£10.7M
Total 24-Month CLV
60% uplift from 12M

£4.8M
Champions CLV 12M
71.4% of total portfolio

£445K
At-Risk CLV 12M
Revised from £5.5M historical

Critical Upgrade: Notebook 03 flagged £5.5M "at risk" — that was historical spend. The CLV model replaces it with £445K projected forward value. That is the actual business case number.

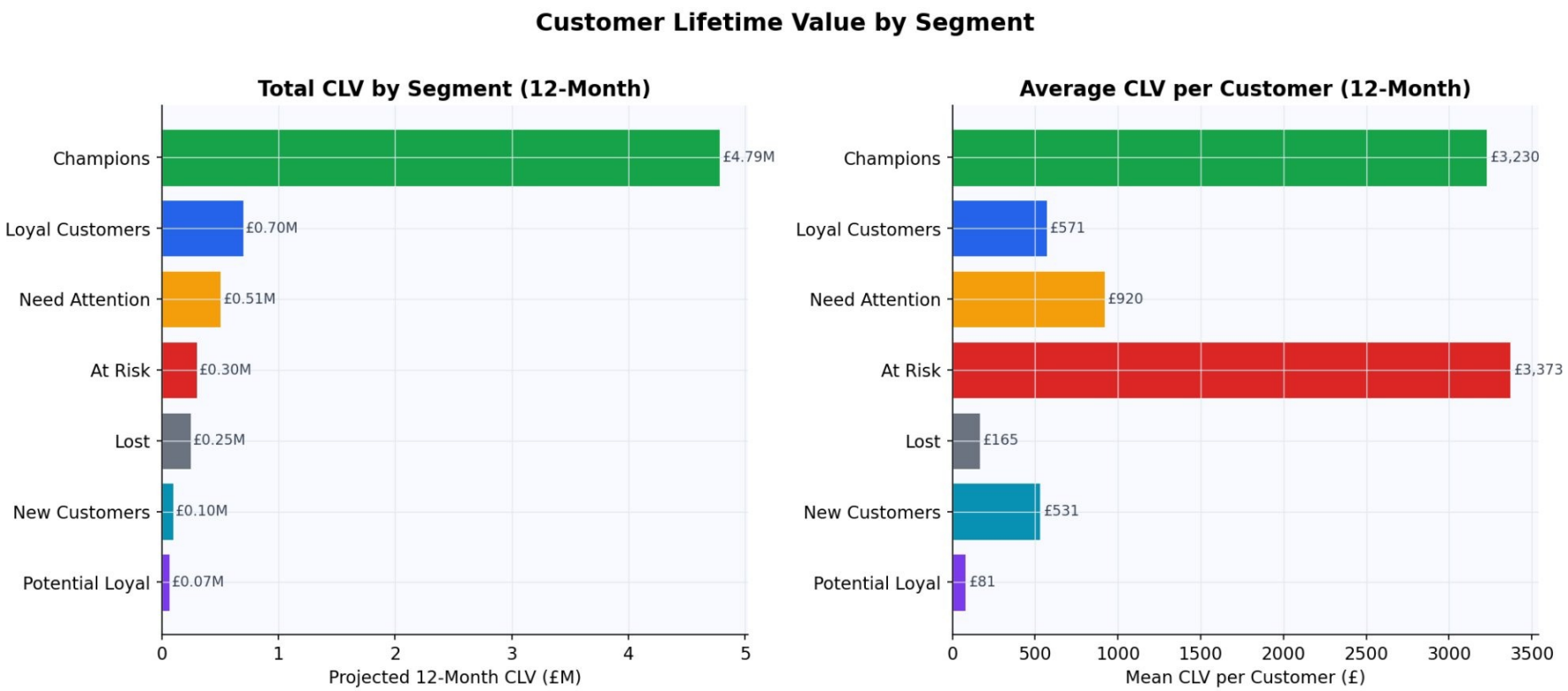
Segment	n	CLV 12M	Mean CLV	Portfolio %
Champions	1,482	£4,786,184	£3,230	71.4%
Loyal Customers	1,221	£697,630	£571	10.4%
Need Attention	551	£507,002	£920	7.6%
At Risk	89	£300,206	£3,373	4.5%
Lost	1,523	£250,593	£165	3.7%

Top 50 Customers

£2.5M

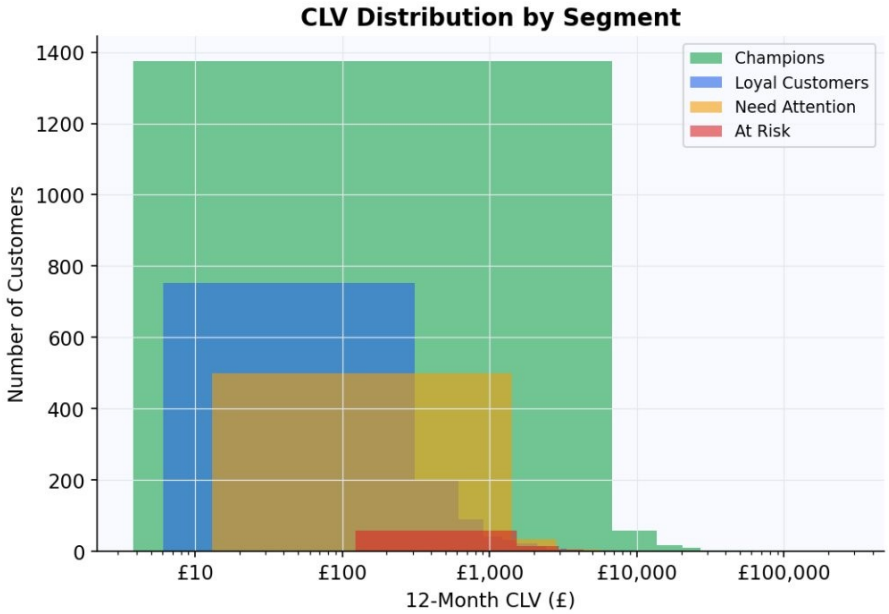
37.9% of portfolio CLV

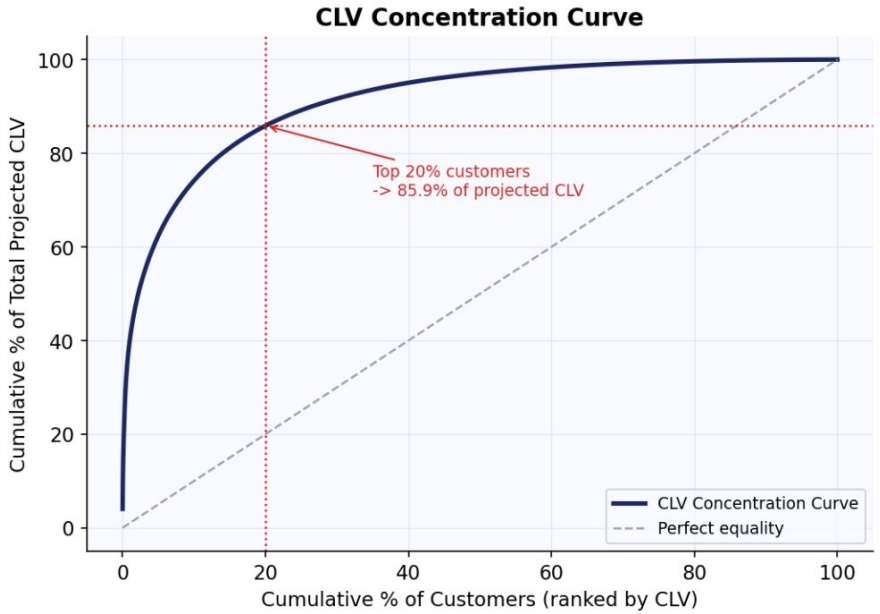
CLV by Segment — Where Portfolio Value Lives



At Risk mean CLV = £3,373 — nearly matches Champions despite only 89 customers. Each justifies bespoke Klaviyo outreach. | Champions: 71.4% of portfolio CLV from just 25% of the customer base.

CLV Concentration — Top 20% → 85.9% of Projected Portfolio Revenue





Extreme Pareto

Top 20% of customers by CLV generate 85.9% of total projected 12M revenue — more concentrated than typical e-commerce benchmarks.

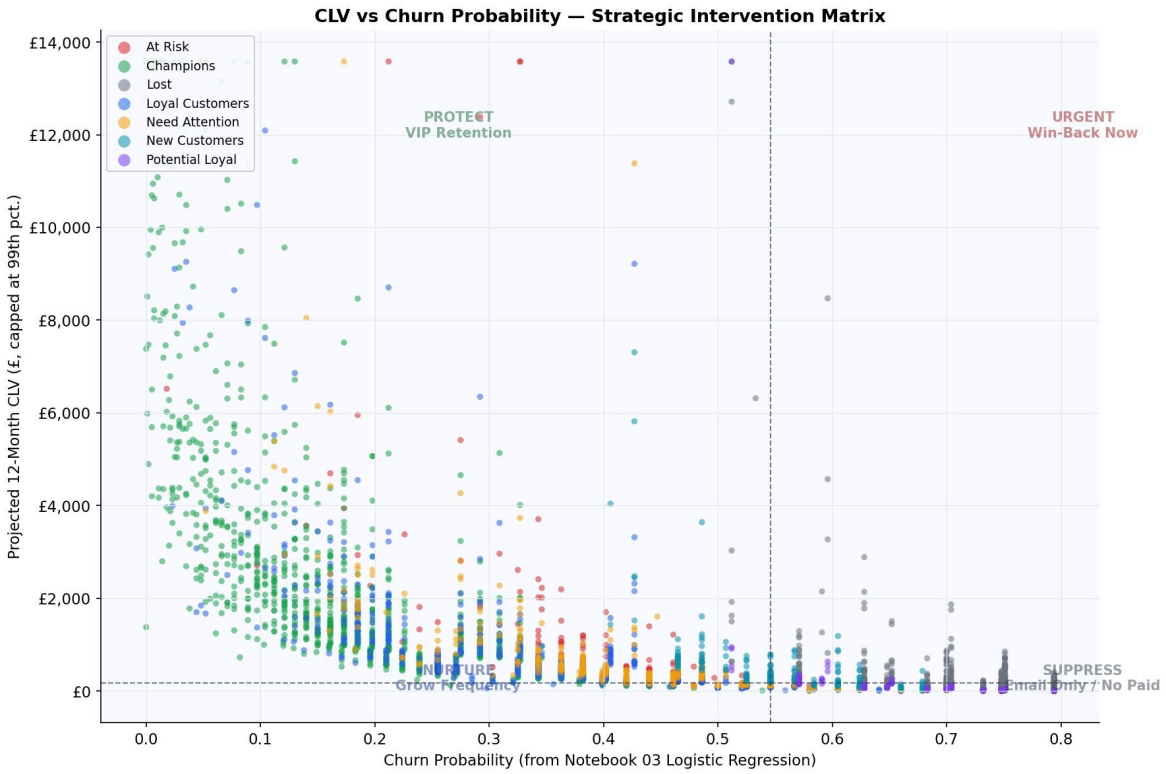
Log-scale distribution

Champions cluster at £1K–£10K. Lost customers are tightly compressed below £200, with negligible spread.

Portfolio implication

Retaining the top quintile protects 86p of every £1 in projected portfolio value. Campaign priority is clear.

Strategic 2×2 — CLV vs Churn Risk: Every Customer Has an Action Zone



PROTECT

High CLV · Low churn
Champions — VIP treatment.
Do not over-market.

URGENT

High CLV · High churn
At Risk — intervene now.
Bespoke Klaviyo, no paid ads.

NURTURE

Low CLV · Low churn
New/Potential Loyal.
Lifecycle programmes.

SUPPRESS

Low CLV · High churn
Lost — exclude all paid.
One final email, then stop.

Google Ads Campaign — £12,500/Month Plan

01 Champions Retention

Customer Match +30% bid · Target ROAS 600%

£3,500/mo

Segment: 512 customers

02 Loyal RLSA Boost

RLSA +20% · Target CPA £18

£4,000/mo

Segment: 1,221 customers

03 Win-Back Display

3-phase creative sequence · Target CPA £18

£3,500/mo

Segment: 551 customers

04 Lost Suppression

Negative audiences · One final email then suppress

£1,500/mo

Segment: 1,523 customers

At Risk (89 customers) — EXCLUDED from all paid campaigns. Bespoke Klaviyo outreach only. | Thursday 09:00–13:00: +25–30% bid boost (EDA peak trading window)

Summary: The Full Business Case

£6.7M 12M Portfolio CLV	**£445K** Forward Revenue at Risk	**AUC 0.777** Churn Model Accuracy
£12,500 Monthly Campaign Budget	**5,878** Customers Analysed	**4 Campaigns** Google Ads Activation

CV Bullet Point

Built end-to-end customer lifecycle analysis in Python — RFM segmentation, churn prediction, CLV modelling, and multi-touch attribution — translating findings into Klaviyo lifecycle flows and a £12,500/month Google Ads campaign strategy validated against survival-weighted DCF revenue projections.

Amrit Sharma · MSc Business Analytics · Durham University · 2025 · github.com/macash916/Customer-Lifecycle-Campaign-Analysis