Baseline Behavioural Revenue Modelling

📅 28 February 2026 | 11:00 AM

1️⃣ Executive Context

The company operates a multi-category retail platform with growing user volume but volatile net revenue.

Management has observed:

- Uneven purchase frequency across users
- Significant basket size variability
- Concentrated revenue among a small user segment
- Return-prone behavior eroding margins

The Board has commissioned a structured behavioral revenue diagnostic.

You are appointed as the **Customer Revenue Analytics Advisory Team**.

This stage focuses on:

Understanding user behavior, modeling revenue generation, and identifying structural fragility.

No optimization decisions are required at this stage.

2 Data Provided

📁 **customer_transactions.xlsx**

Historical transactional dataset (training window).

Core Variables (Representative Structure)

Variable Name	Data Type	Level	Description
EventID	Nominal	Basket	Unique identifier assigned to each user interaction (basket/invoice). Multiple rows may share the same EventID.
ProductID	Nominal	Line Item	Unique identifier assigned to each product.
ProductName	Nominal (Text)	Line Item	Text description of the product.
Quantity	Numeric	Line Item	Number of units involved in the interaction. May represent purchase or return activity depending on dataset encoding.
EventDateTime	Timestamp	Basket / Line Item	Date and time at which the interaction occurred. Must be treated with strict temporal integrity.
UnitPrice	Numeric (INR)	Line Item	Price per unit of the product at time of interaction.
UserID	Nominal	User	Unique identifier assigned to each customer.

3 Stage 1 Analytical Mandate

You must build a structured behavioral model addressing:

A. Purchase Likelihood Modeling

- Estimate probability of repeat purchase

- Model transaction frequency per user

- Identify high-engagement segments

- Detect dormant vs active clusters

B. Basket Size Modeling

- Estimate expected basket value per transaction

- Analyze SKU/category contribution

- Detect heavy-tail revenue effects

- Identify product mix drivers

C. Net Revenue Estimation

Define: (Expected Net Revenue) = E(Gross Value) - E(Return Value)

Compute: Expected\Revenue Per User (ERPU)]

This becomes your Stage 1 baseline.

D. Behavioral Segmentation

You must quantify:

- Revenue concentration (Top 10% share)

- Return-prone user clusters

- SKU dependency exposure

- Category contribution risk

4 Required Outputs (Stage 1)

Your submission must include:

- Behavioral model structure
- ERPU baseline estimate
- Purchase probability framework
- Basket value model
- Revenue concentration profile
- Return risk segmentation
- Clear documentation of assumptions

5 Analytical Discipline Expectations

Teams must:

- Clearly separate probability modeling from value modeling
- Avoid single-metric summaries
- Quantify uncertainty where appropriate
- Ensure reproducibility

Black-box outputs without explanation will be penalized.

6 What This Stage Tests

Dimension	Evaluation Focus
Modeling Depth	Behavioral drivers identified
Structure	Logical revenue decomposition
Statistical Discipline	Probability vs value separation
Fragility Awareness	Concentration & return risk quantified

N. L. Dalmia®
Institute of Management
Studies and Research
(A School of Excellence of N. L. Dalmia Educational Society)

Stage 2: In-Space Validation & Model Robustness Check

📅 28 February 2026 | 7:00 PM

1️⃣ **Data Update**

A validation dataset is now released.

📁 **customer_validation_set.xlsx**

This dataset is an **in-space holdout sample** drawn from the same operating regime as the Stage 1 training data.

You should assume:

- Same policy and market conditions as training
- Similar behavioral regime (no structural shock implied)
- Differences may arise due to sampling, user mix, SKU mix, or noise

This stage evaluates **model quality, stability, and overfitting control** — not drift adaptation.

2️⃣ **Purpose of Stage 2**

Stage 2 is a robustness checkpoint.

It tests whether your Stage 1 model:

- Generalizes to unseen users/events within the same regime
- Maintains calibration and ranking quality
- Separates signal vs noise
- Provides stable ERPU estimates without overfitting

3️⃣ **Stage 2 Analytical Mandate**

You must:

A. Score Validation Users

Using your Stage 1 model, compute for each user:

- Purchase likelihood score
- Expected basket value
- Expected return risk (if modeled)
- Expected Revenue per User (ERPU)

NLD Synapse 2026 | N. L. Dalmia Institute of Management Studies & Research.

B. Evaluate Generalization Performance

Quantify:

- Calibration quality (probability reliability)
- Ranking quality (top-decile lift / gain)
- Basket value error
- Return-risk separation (if included)
- ERPU prediction error

You must explicitly compare:

Training vs Validation on the same metric definitions.

C. Diagnose Overfitting Risk

Demonstrate whether:

- Performance drops materially
- High scores are unstable
- Results depend on narrow features
- Concentration risk is an artifact of model bias

D. Recalibration (Only If Justified)

You may recalibrate:

- Probability calibration (Platt / isotonic / scaling)
- Segment-level adjustments
- Regularization or feature pruning

You must not claim "regime shift" here.

All modifications must be justified as **generalization fixes**.

■ **Interim Submission – 11:00 PM**

Submit:

- Stage 1 baseline model summary
- Train vs validation performance dashboard
- Overfitting diagnosis
- Any calibration adjustments + justification
- Updated ERPU and targeting logic direction (no final selection)

N. L. Dalmia®
Institute of Management
Studies and Research
(A School of Excellence of N. L. Dalmia Educational Society)

Revenue Targeting Optimization Under Structural Constraint

📅 01 March 2026 | 1:00 AM

1️⃣ **Final Dataset Release**

📁 **customer_test.csv**

This dataset represents an **out-of-space behavioral window**.

Unlike validation:

- User mix may differ
- Category exposure may shift
- Purchase intensity may change
- Return patterns may evolve

You must assume this window represents a new operational phase.

No clarification regarding shift magnitude will be provided.

2️⃣ **Board Directive**

The Board now requires:

Maximize expected net revenue through intelligent user targeting while preserving structural stability.

This is no longer a modeling stage.
It is a **constrained optimization decision stage.**

3️⃣ **Decision Constraints**

You must satisfy **all** of the following:

1. **Targeting Budget Constraint**

   You may target **at most 30% of users**
   (based on your ranking logic).

   Mass targeting is not permitted.

   You must justify your selection rule.

2. **Return Risk Control**

   Users in the **top decile of predicted return risk** must not exceed **25% of the targeted group**.

3. **Concentration Stability Constraint**

   Revenue share of the **top 10% predicted users**
   must not increase by more than **15% relative to Stage 1 baseline concentration**.

⚡ Stage 3 Analytical Mandate

You must:

A. Score Test Users

For each user:

- Purchase probability
- Expected basket value
- Expected return loss
- Expected net revenue

B. Rank and Select

Construct an optimization rule that:

- Maximizes total expected net revenue
- Respects targeting cap
- Controls return risk exposure
- Controls concentration amplification

This may require:

- Threshold optimization
- Segment-based targeting
- Risk-adjusted ranking
- Multi-objective scoring

C. Quantify Trade-Offs

You must explicitly show:

- Revenue vs risk trade-off
- Revenue vs concentration trade-off
- Marginal gain of expanding target set
- Fragility under small parameter shifts

D. Stress Sensitivity

Demonstrate:

- What happens if targeting cap = 25% or 35%
- What happens if return risk threshold tightens
- Which assumption most destabilizes outcome

5 Required Outputs (Final Submission – 9:00 AM)

Your final submission must integrate:

- Stage 1 baseline model
- Stage 2 validation calibration
- Stage 3 targeting decision

Must include:

- Total expected net revenue
- Targeted user count
- Concentration metrics
- Return exposure metrics
- Trade-off mapping
- Structural fragility assessment

Single coherent strategy required.