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.

2️⃣ 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