Business Rules & Data Assumptions

1. Data Scope & Authority

This analytical system operates strictly within the boundaries of the provided Sales_enriched.csv dataset. No external datasets, benchmarks, market assumptions, inflation adjustments, or inferred customer behavior are permitted. All insights, summaries, and conclusions must be fully traceable to this dataset. This rule exists to prevent hallucinated business context and ensure analytical reproducibility.

2. Revenue Computation Rules

Revenue is computed at the row level using the formula Revenue = Price × Quantity. The revenue column present in the dataset is treated as authoritative and pre-computed. No discounts, refunds, taxes, promotions, or cost adjustments are assumed unless explicitly present. Negative revenue values are treated as data-quality violations.

3. Order & Transaction Definition

Each unique sale_id represents exactly one order. Orders cannot be assumed to represent unique customers. No repeat-purchase or customer behavior assumptions are allowed.

4. Time & Period Rules

Temporal analysis is conducted at the monthly level using sale_year_month derived from sale_date. Partial months may exist and must not be normalized or extrapolated.

5. Product Validity & Lifecycle

Products are identified by product_id. Launch_Date indicates product introduction but does not invalidate observed sales. All sales are assumed valid unless explicitly flagged.

6. Store & Organizational Rules

Stores are identified by store_id with no assumed hierarchy or geography. Store churn refers only to activity presence, not closure or relocation.

7. Data Quality & Integrity

Missing values, duplicates, or anomalies must be surfaced. No silent corrections, imputations, or synthetic data generation are allowed.

8. Analytical Boundaries

Customer churn, CLV, profitability, and marketing attribution are explicitly prohibited due to missing inputs.