1. Methodology & Research Process

The analysis was performed on the Olist e-commerce dataset. The process involved a deep-dive audit to move from "noisy" raw data to a reliable business diagnosis:

- Initial Statistics: Preliminary data suggested a Frequency (F) of 1.04. However, initial queries showed "0" repeat customers due to the use of transactional customer_id.

- Technical Discovery: 1. Row Inflation: I identified a one-to-many relationship between orders and payments. Installment payments (up to 29 rows per order) were incorrectly inflating transaction counts. 2. ID Mapping: I discovered that customer_id is unique to each session. To find returning humans, I had to join the customers table to access the customer_unique_id.

- Data Correction: By implementing COUNT(DISTINCT order_id) grouped by customer_unique_id, I successfully identified the real loyal customer base (e.g., a top customer with 15 unique orders).

- Model Strategy: I applied a full RFM model, but focused heavily on Recency and Monetary due to the low overall retention rate (most customers are "Single" buyers).

2. Key Business Insights

- Low but Existing Loyalty: Contrary to the first impression, loyalty exists but is very rare. We identified over 2,400 repeat customers across the "New/Active" and "Occasional" segments.

- The "Leaky Bucket" Challenge: 97% of the database consists of one-time buyers. This confirms that the company is highly dependent on expensive, continuous customer acquisition (Ads).

- Recent Champions vs. At-Risk: The "Recent Champions" group (225 people) has a 10% retention rate, which is much higher than the company average. Meanwhile, the "At Risk High-Value" segment (468 people) represents the most significant revenue loss.

3. Strategic Recommendations

- Focus on the "Champions": Analyze the 24 repeat customers in the "Recent Champions" segment to understand why they returned (e.g., specific product categories or fast delivery).

- Win-back Campaign: Target the 468 "At Risk High-Value" customers with a personalized "We Miss You" offer. These customers have already proven they are willing to spend >1000 units.

- CRM & Loyalty Loop: Implement automated post-purchase communication. The data shows that if a customer doesn't return within 90 days (Recency score 5), the chance of them returning drops significantly.