

GB English Version: Customer Segmentation Analysis (RM Report)

1. Methodology & Research Process

The analysis was performed on the Olist e-commerce dataset. The process involved pivoting from theoretical assumptions to a fact-based diagnosis:

- **Initial Statistics:** Descriptive data suggested an average of **1.04** orders per customer (Frequency).
- **Technical Discovery:** During table joining (JOIN), a one-to-many relationship was identified between orders and payments. It was discovered that installment payments were incorrectly counted as separate orders.
- **Data Correction:** Implementing COUNT(DISTINCT order_id) revealed that actual customer retention is near **0%**.
- **Model Pivot:** The classic RFM model was pivoted to an **RM (Recency-Monetary) model**. Since almost all customers purchased only once, Frequency (F) lost its statistical utility as a segmentation criterion.

2. Key Business Insights

- **Lack of Loyalty:** The business model relies entirely on continuous customer acquisition. The company loses nearly every customer after their first transaction.
- **High-Value Segments:** Identified 468 customers in the **"At Risk High-Value"** group (spent >1000 but haven't returned for a long time). This represents the most significant capital leak.
- **Revenue Structure:** Most revenue is generated by the **"New / Active"** and **"Occasional"** segments, confirming a high dependency on ongoing paid marketing campaigns (Ads).

3. Recommendations & Solutions

- **CRM Implementation:** Immediate launch of marketing automation (post-purchase emails with discount codes for the next order).
- **Win-back Campaign:** Targeting a dedicated offer to the **"At Risk High-Value"** segment to recover the highest-spending customers before they churn completely.
- **Loyalty Program:** Establishing a points-based mechanism to incentivize a second purchase, thereby increasing Customer Lifetime Value (LTV).