Do Discounts Improve Customer Retention?

Customer-level retention and value analysis using SQL, Python, and Tableau

Problem Statement

Discounting is commonly used to encourage repeat purchases. This analysis evaluates whether customers who receive deeper discounts actually return more often and contribute higher long-term value.

Initial Assumption (Data Quality and Distribution)

An initial exploration of customer-level behaviour revealed a highly skewed distribution of purchase frequency, with most customers placing only one or two orders and a small number contributing disproportionately large values.

To ensure that retention patterns reflected meaningful customer behaviour rather than artifacts of skewed data, the analysis assumed that:
A customer-level summary with corrected repeat indicators and balanced distribution was necessary for reliable retention analysis.

Based on this assumption:
- Repeat purchase flags were validated and corrected
- Order counts were explicitly modeled at the customer level
- Customers were segmented using discount exposure buckets rather than raw discount values

This preparation step ensured that subsequent retention insights were not driven by data imbalance alone.

Data, Setup and Tools

- Simulated but behaviorally realistic e-commerce dataset
- ~500 customers and ~3,000 orders
- Transactional data aggregated to the customer level using SQL

Customer-level features engineered in SQL:
- Order count per customer
- Total revenue per customer
- Repeat purchase flag
- Average discount exposure (bucketed as Low, Medium, High)

Python was used for analysis and validation, and results were visualized in Tableau.

Project artifacts: SQL schema & queries, Python analysis, Tableau dashboard.

Analytical Approach

- Aggregated transaction-level data into a customer summary table
- Validated retention indicators to mitigate skew effects
- Segmented customers by discount exposure bucket
- Compared segments using:
  - Repeat purchase rate
  - Average orders per customer
  - Average revenue per customer

This approach isolates long-term customer behaviour while controlling for distributional imbalance.

Key Insight

- Customers receiving medium discounts showed the highest repeat purchase rates
- These customers also demonstrated stronger engagement, placing more orders per customer
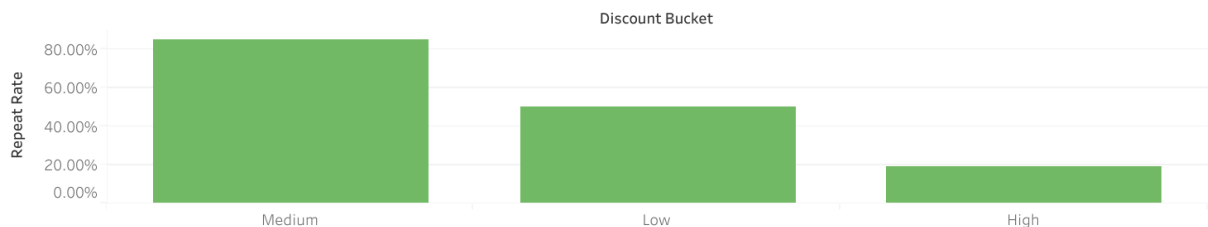- High discount customers performed worst across all metrics, including retention and customer value

The retention patterns remained consistent after correcting for skewed customer behaviour.

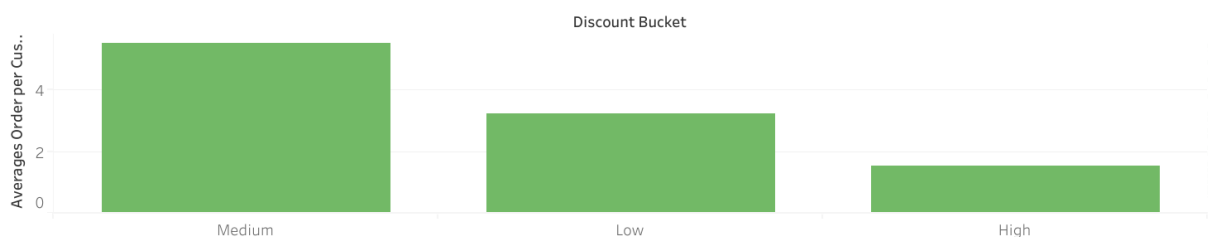**Customer Retention Dashboard**

Do Discounts Improve Customer Retention?

Repeat behaviour and customer value by discount exposure

Repeat Rate by Discount Bucket

Avg Orders per Customer by Discount Bucket

Business Implication

While discounts may increase short-term purchasing activity, this analysis suggests that deeper discounts do not necessarily lead to stronger customer retention.

Moderate incentives performed best across retention, engagement, and value metrics in this dataset, indicating that discount depth alone may not be an effective long-term loyalty strategy.

Project artifacts: SQL schema & queries, Python analysis, Tableau dashboard.