



Unlocking Customer Insights: A Transaction Data Deep Dive

This project delves into customer purchasing habits using transaction data from nearly 4,000 orders. Our objective is to uncover critical insights into spending trends, customer segmentation, product preferences, and subscription engagement to inform strategic business decisions and drive growth.



Project Overview: Data-Driven Strategies

By analyzing a comprehensive dataset, we aim to provide actionable intelligence across key business areas:

Spending Trends

Identify patterns in customer expenditure across various product categories.

Customer Groups

Segment customers based on demographics and purchasing behavior for targeted strategies.

Product Preferences

Understand which products resonate most with our audience and why.

Subscription Activity

Evaluate the impact of subscription models on customer loyalty and revenue.

Dataset at a Glance: The Foundation of Our Analysis

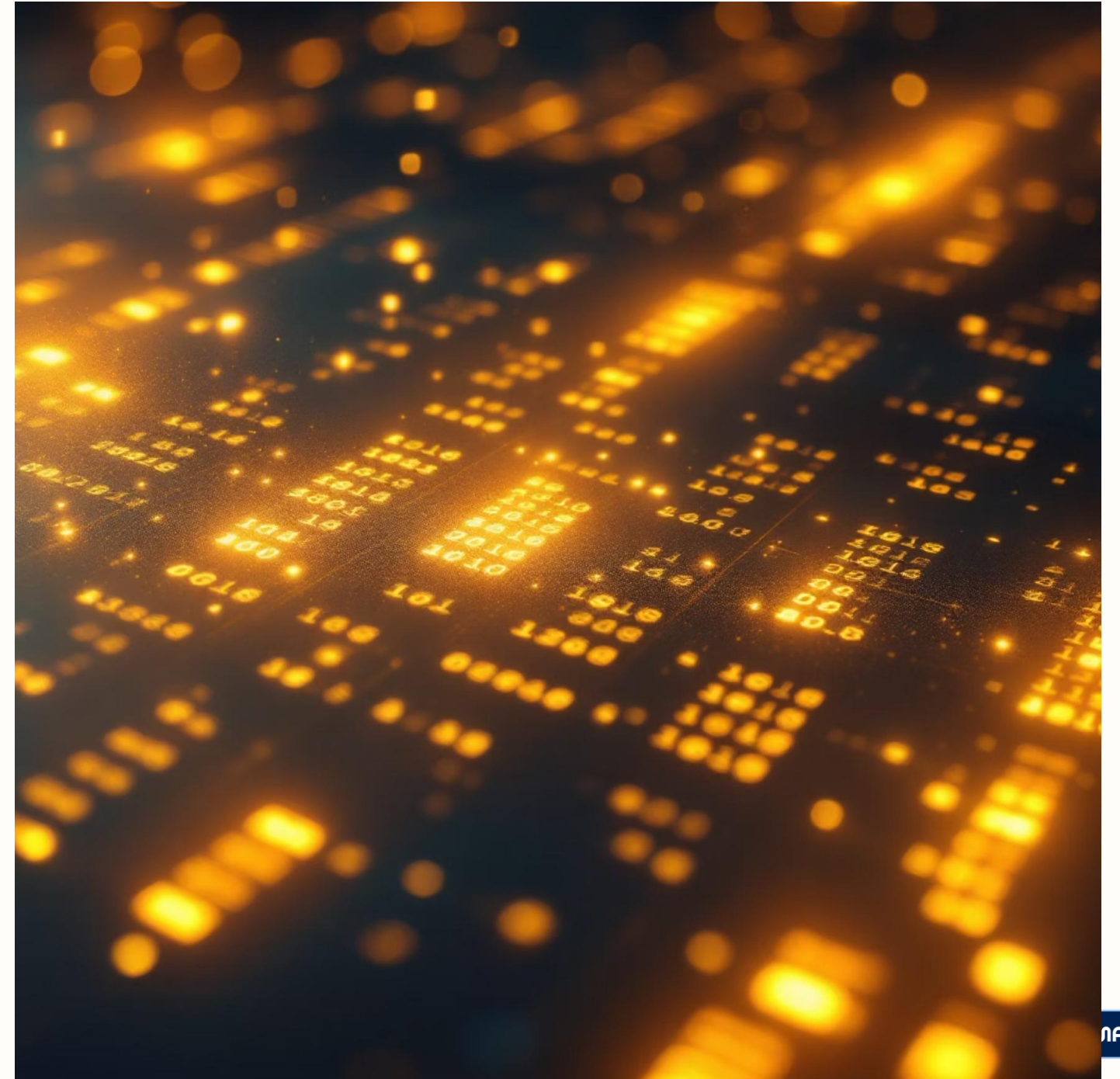
Our analysis is built upon a robust dataset containing detailed information on customer transactions.

Key Metrics

- **Total Records:** 3,900 customer orders
- **Number of Columns:** 18 distinct attributes

Critical Attributes Explored

- **Customer Demographics:** Age, Gender, Location, Subscription Status
- **Purchase Details:** Product, Category, Amount Spent, Season, Size, Color
- **Shopping Behavior:** Discount Usage, Prior Purchases, Buying Frequency, Review Ratings, Shipping Type



Python-Powered Data Preparation: Cleansing & Structuring

Rigorous data preparation in Python ensured the reliability of our analysis. Using pandas, we transformed raw data into a clean, actionable format.

01

Data Import & Initial Review

Loaded the dataset and performed initial checks using `df.info()` for structure and `.describe()` for a statistical overview.



02

Missing Data Handling

Null values in `review_rating` were imputed using the median rating specific to each product category.

03

Column Formatting & Consistency

Converted all column names to `snake_case` for uniform accessibility and readability.

04

Feature Engineering

Created an `age_group` variable by binning ages and derived `purchase_frequency_days` from transaction timestamps.



05

Redundancy Check

Identified and removed the redundant `promo_code_used` column as it overlapped with `discount_applied`.

Seamless Integration: Python to MySQL

Once cleaned and structured in Python, the dataset was integrated into MySQL for advanced, query-based analysis. This step allowed us to leverage SQL's powerful capabilities for specific business questions.

Why MySQL?

- Efficient querying of large datasets.
- Structured environment for complex aggregations.
- Facilitates clear, repeatable analytical operations.



SQL-Powered Insights: Revenue & Discounts

Our SQL-based analysis revealed compelling patterns in revenue generation and discount utilization.

Revenue by Gender

Comparative analysis of total sales generated by male versus female customers, highlighting key differences in purchasing power.



\$

High-Spending Discount Users

Identified 839 customers who used discounts but still spent above the average purchase value, indicating strategic discount engagement.

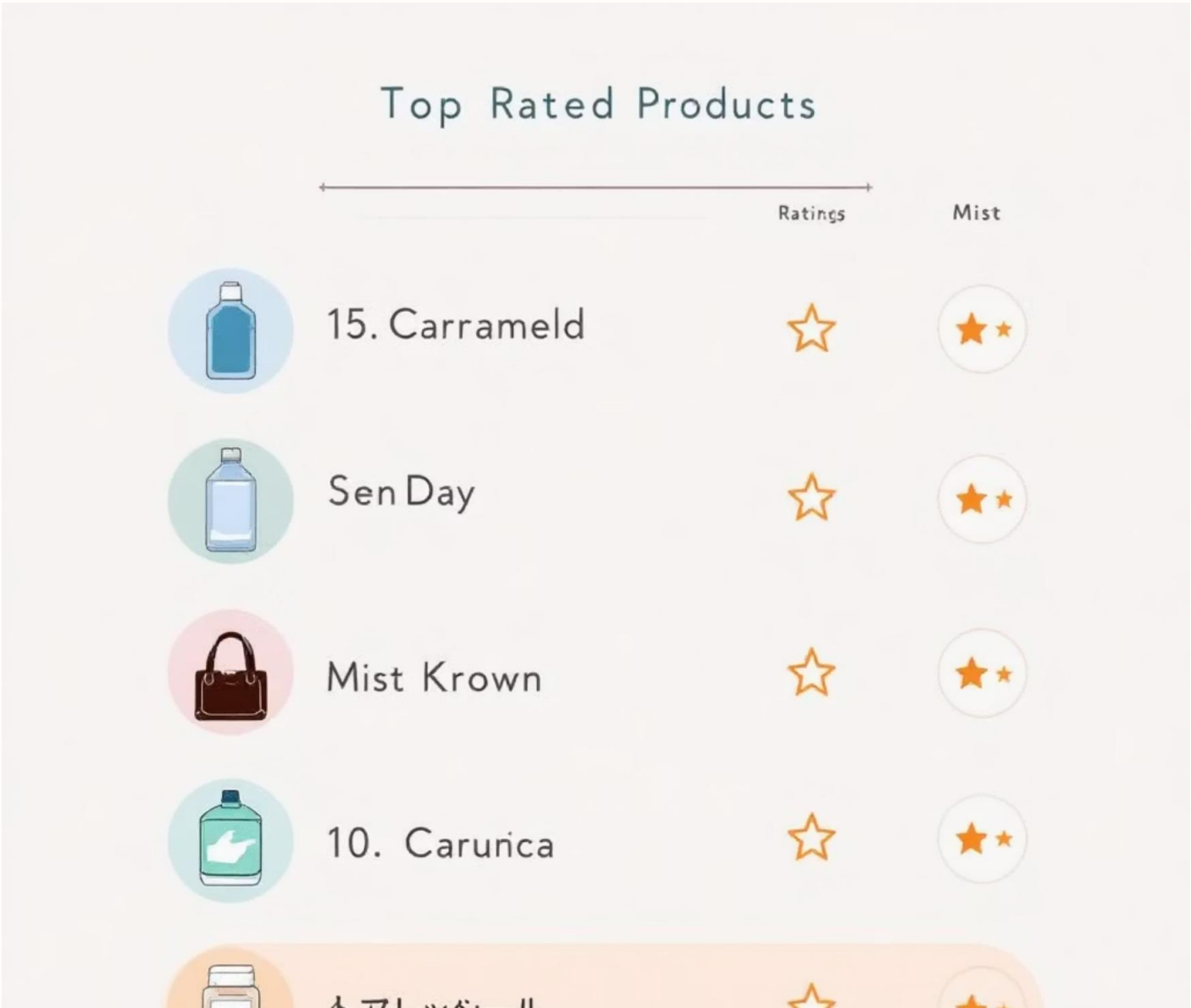


Product Performance & Shipping Choices

Delving deeper, we examined product ratings and the financial impact of different shipping preferences.

Top 5 Products by Rating

- Revealed products consistently achieving the highest mean review scores.
- Insights into customer satisfaction and product quality drivers.



Shipping Type Comparison

- Analyzed average spending between Standard and Express delivery options.
- Provided data on customer willingness to pay for expedited shipping.



Subscribers vs. Non-Subscribers: A Key Distinction

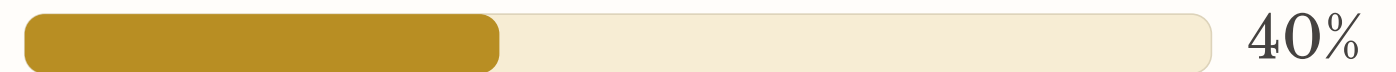
Understanding the value of our subscription program is vital. This analysis compares the average spend and overall revenue contributions of subscribers against non-subscribers.



Higher Average Spend

Subscribers consistently show a higher average purchase value.

These findings underscore the importance of nurturing our subscriber base and exploring strategies to convert non-subscribers.



Increased Overall Revenue

Subscription holders contribute significantly more to total revenue.

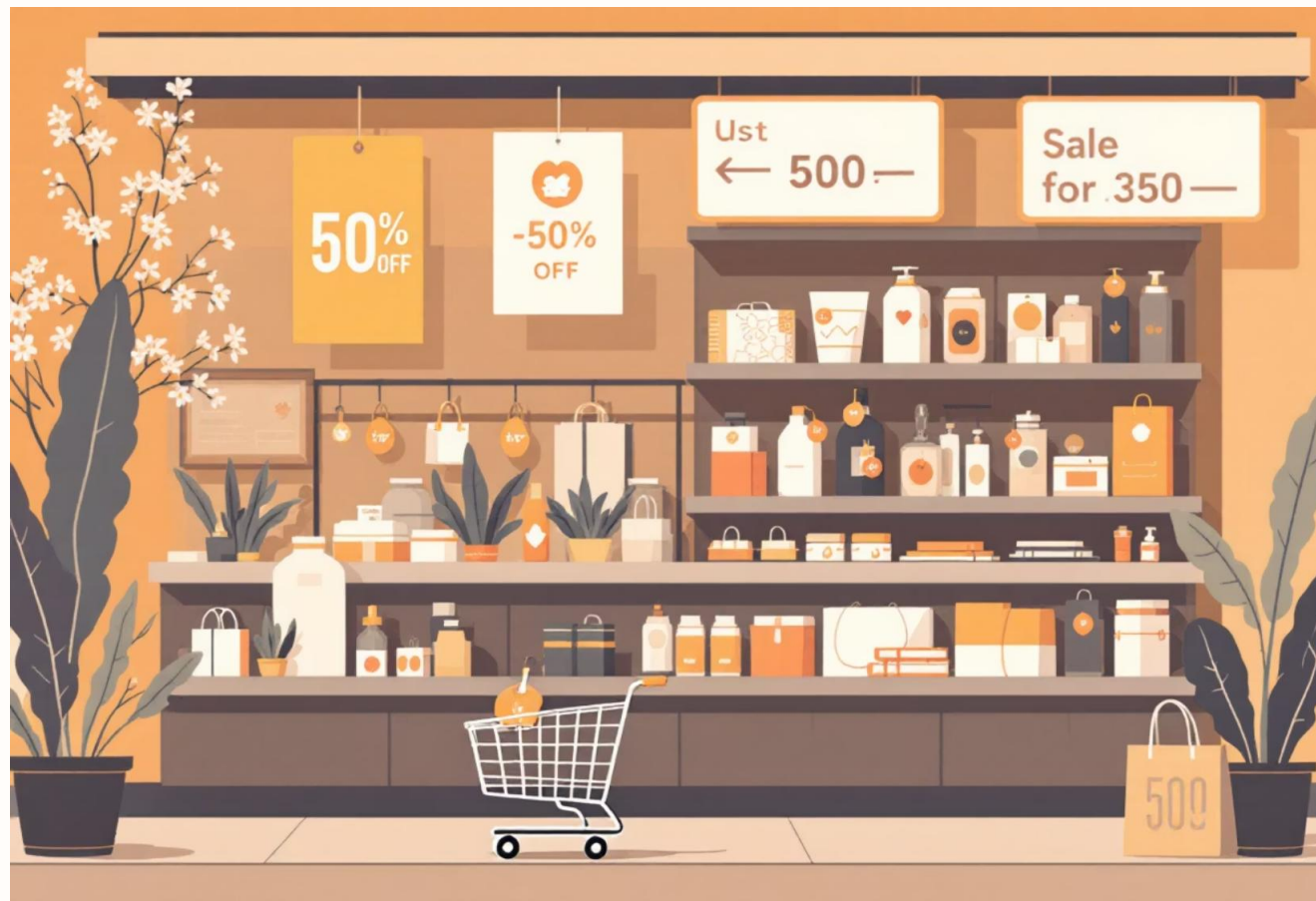
Discount Dependency & Customer Segmentation

Our analysis identified products heavily reliant on discounts and segmented our customer base for tailored marketing efforts.

1

Discount-Dependent Products

Highlighted five products most frequently purchased with discounts, suggesting potential price elasticity.



2

Customer Segmentation

Categorized customers into **New**, **Returning**, and **Loyal** groups based on their purchase history for personalized engagement strategies.



Customer Segmentation: Unveiling Distinct Behaviors

Understanding our customer base through detailed segmentation allows us to tailor marketing efforts, product development, and retention strategies effectively. We've identified three core segments with unique purchasing patterns and loyalty levels.

New Customers

- **Average Spend:** \$50
- **Purchase Frequency:** 1.2 orders/month
- **Discount Usage:** 70%
- **Customer Lifetime Value:** \$150
- **Retention Rate:** 20% (3-month)
- **Avg. Review Rating:** 3.8/5

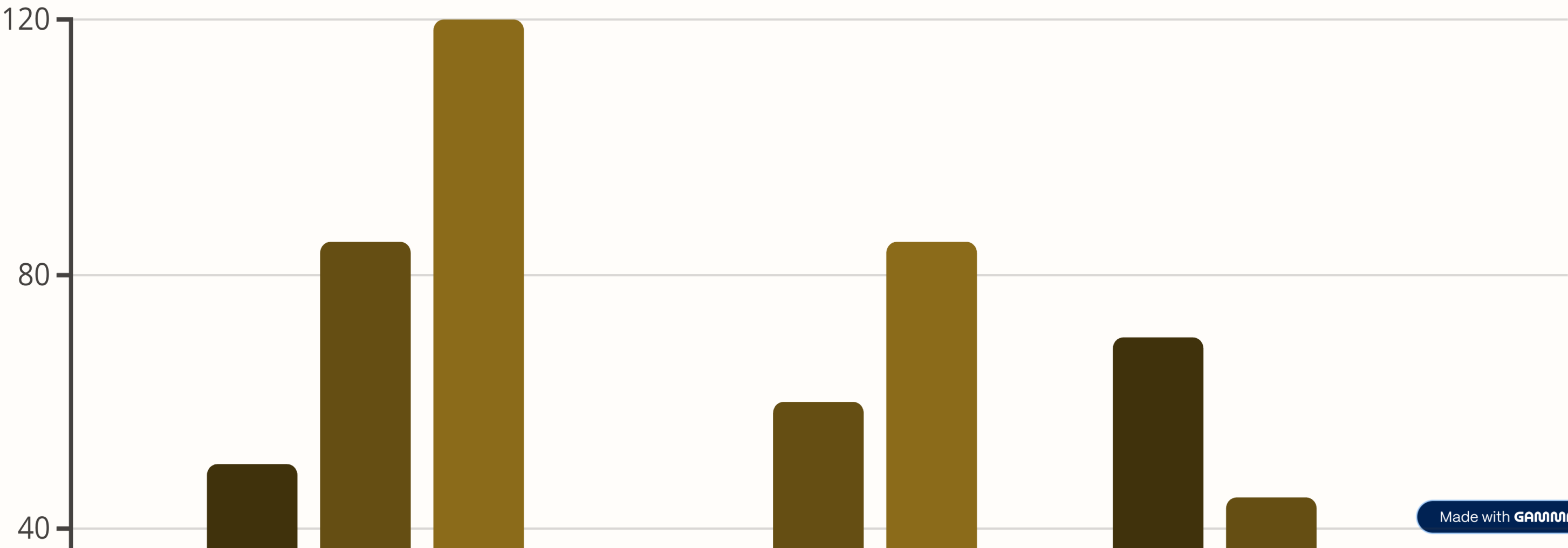
Returning Customers

- **Average Spend:** \$85
- **Purchase Frequency:** 1.8 orders/month
- **Discount Usage:** 45%
- **Customer Lifetime Value:** \$400
- **Retention Rate:** 60% (3-month)
- **Avg. Review Rating:** 4.2/5

Loyal Customers

- **Average Spend:** \$120
- **Purchase Frequency:** 2.5 orders/month
- **Discount Usage:** 20%
- **Customer Lifetime Value:** \$900+
- **Retention Rate:** 85% (3-month)
- **Avg. Review Rating:** 4.7/5

This comparative overview highlights the varying needs and values of each segment, informing where we can best allocate resources for engagement and growth.





Actionable Insights for Strategic Growth

The data-driven insights from this project empower us to refine business strategies and optimize customer engagement.



Targeted Marketing

Develop campaigns tailored to specific customer segments and product preferences.



Optimize Discount Strategies

Strategically apply discounts to maximize revenue without compromising perceived value.



Enhance Subscription Value

Focus on increasing subscriber retention and conversion through enhanced benefits.



Product Development

Leverage top-rated product insights for future product innovation and improvement.