Customer Shopping Behavior Analysis

This project analyzes 3,900 customer purchases to uncover spending patterns, segments, product preferences, and subscription behavior. Our goal is to guide strategic business decisions.

Dataset Summary: Key Features

Customer Demographics

Age, Gender, Location, Subscription Status

Purchase Details

Item, Category, Amount, Season, Size, Color

Shopping Behavior

Discount, Promo Code, Previous Purchases, Frequency, Review Rating, Shipping Type

The dataset comprises 3,900 rows and 18 columns, with 37 missing values in the Review Rating column.

Exploratory Data Analysis (Python)

Data Preparation & Cleaning

- Data Loading: Imported dataset using pandas.

- Initial Exploration: Checked structure and summary statistics.

- Missing Data Handling: Imputed missing Review Ratings using median.

Feature Engineering & Database Integration

Age Group Column

Binned customer ages for segmentation.

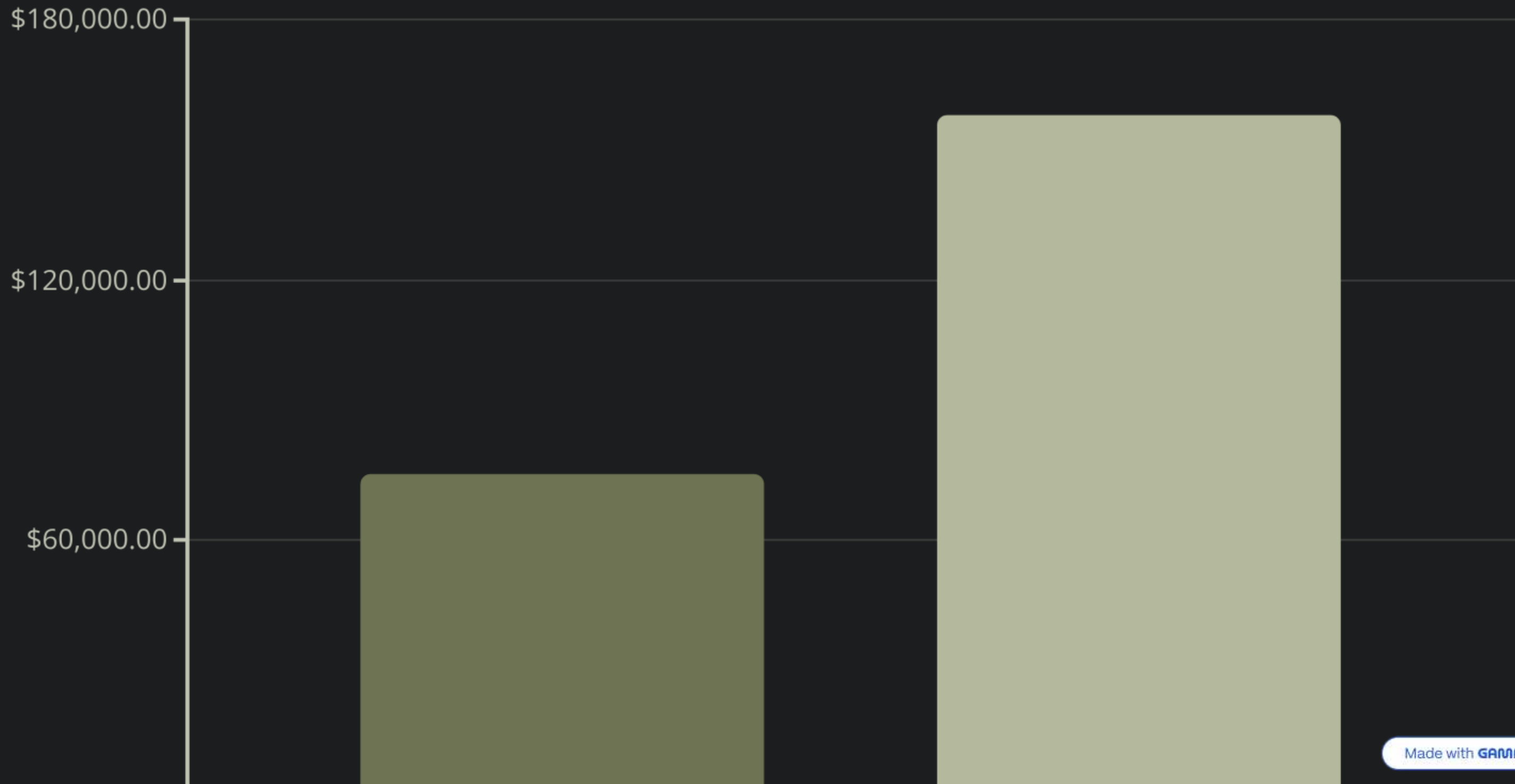### Purchase Frequency

Derived from purchase data.

MySQL Integration

Loaded cleaned data for SQL analysis.

We also verified and dropped redundant columns like 'promo_code_used' for data consistency.

Revenue by Gender

$180,000.00

$120,000.00

$60,000.00

Top Products & Shipping Insights

Top 5 Products by Rating

- Gloves: 3.86
- Sandals: 3.84
- Boots: 3.82
- Hat: 3.80
- Skirt: 3.78

Shipping Type Comparison

Average purchase amounts:

- Standard Shipping: $58.46
- Express Shipping: $60.48

Subscription & Discount Analysis

Subscribers vs. Non-Subscribers

Subscribers (1053) have a slightly lower average spend ($59.49) than non-subscribers (2847) ($59.87), but contribute significantly to total revenue.
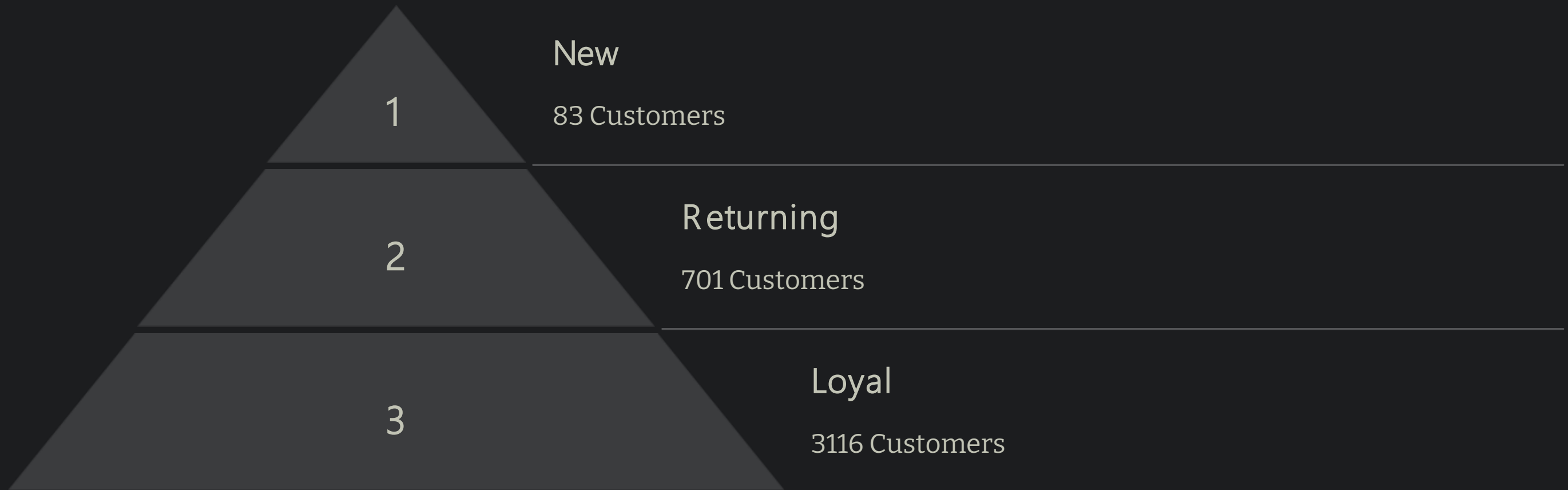
Discount-Dependent Products

Top 5 products with highest percentage of discounted purchases: Hat (50%), Sneakers (49.66%), Coat (49.07%), Sweater (48.17%), Pants (47.37%).
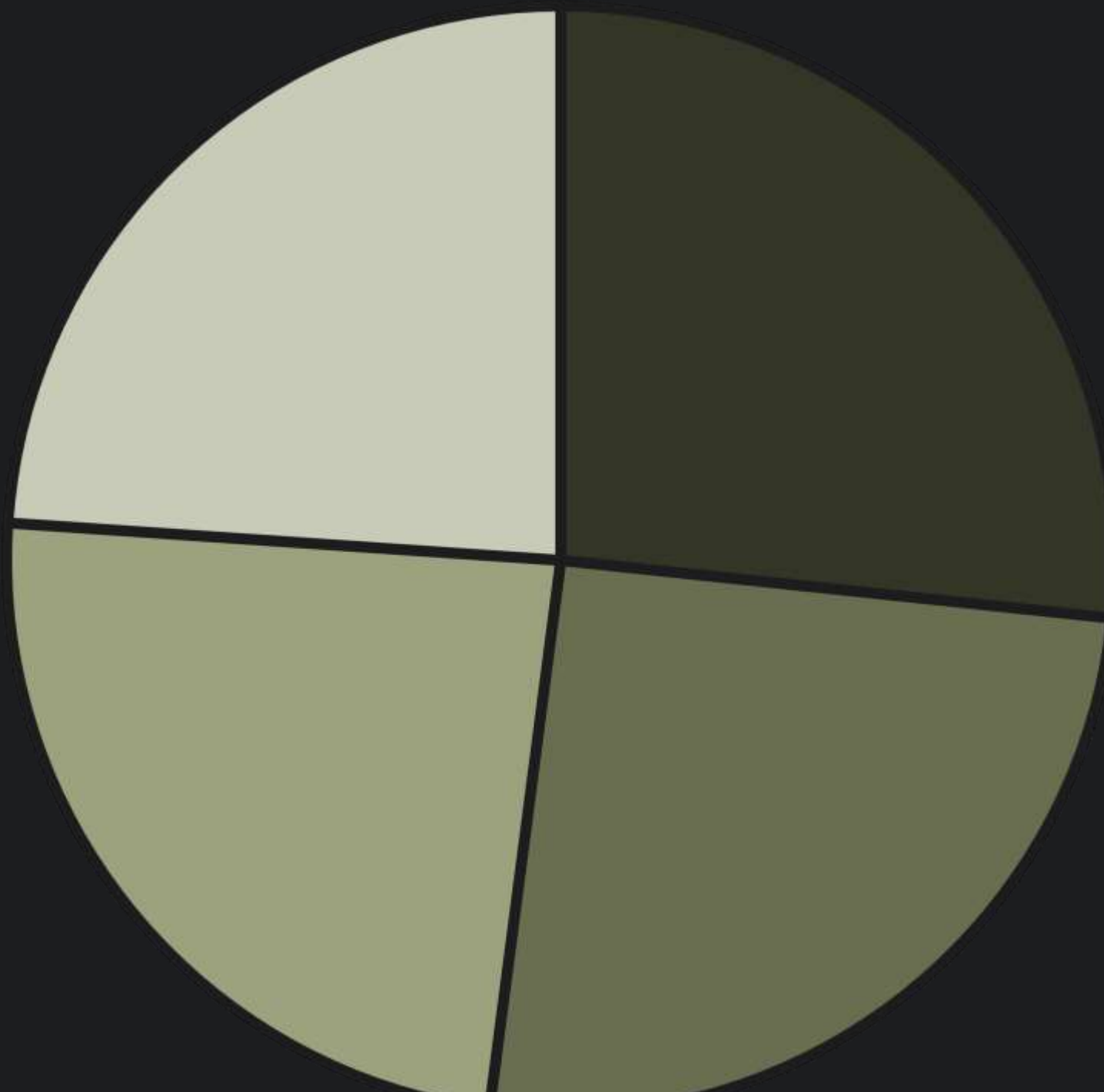
This highlights the impact of discounts on specific product sales and subscriber engagement.

Customer Segmentation & Loyalty

New

1

83 Customers

Returning

2

701 Customers

Loyal

3

3116 Customers

The majority of our customer base falls into the 'Loyal' segment, indicating strong retention. Repeat buyers (those with >5 purchases) are more likely to subscribe (958 subscribers vs. 2518 non-subscribers).

Revenue by Age Group

Business Recommendations

Boost Subscriptions

Promote exclusive benefits to increase subscriber base.

Customer Loyalty Programs

Reward repeat buyers to foster loyalty and retention.

Targeted Marketing

Focus efforts on high-revenue age groups and express-shipping users.

Product Positioning

Highlight top-rated and best-selling products in campaigns.