

Customer Shopping Behavior Analysis

Analysis of 3,900 transactions to reveal spending patterns, segments, product preferences, and subscription behavior to inform strategy.



Dataset Summary

Core Facts

3,900 rows × 18 columns. Key features: demographics, purchase details, shopping behavior, and subscription status. 37 missing values in Review Rating.

- Demographics: Age, Gender, Location, Subscription
- Purchases: Item, Category, Amount, Season, Size, Color
- Behavior: Discount, Promo, Frequency, Rating, Shipping





Data Prep & EDA (Python)

Load & Inspect

Imported with pandas; used df.info() and .describe() to profile structure and summary stats.

Clean & Impute

Imputed missing Review Rating values using median per product category; standardized column names to snake_case.

Feature Engineering

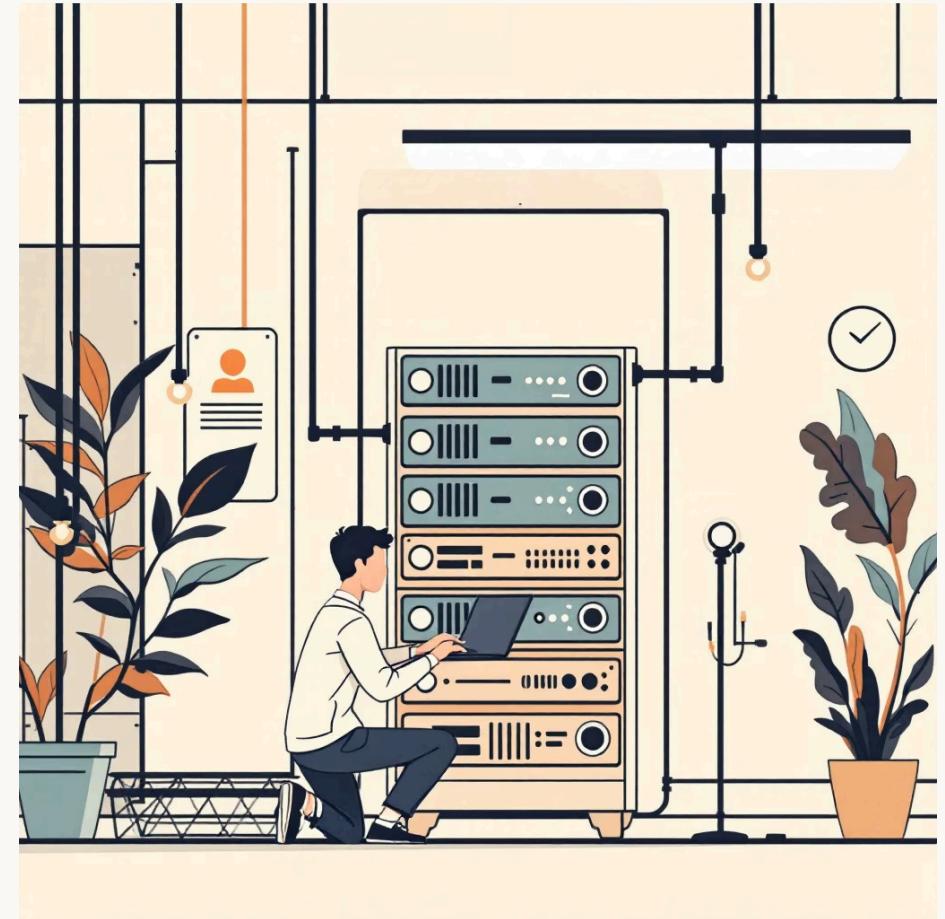
Created age_group and purchase_frequency_days; evaluated redundancy and dropped promo_code_used.

Database Integration

PostgreSQL Load

Cleaned DataFrame uploaded to PostgreSQL for scalable SQL analysis across business questions.

- Ensures reproducibility
- Supports complex aggregation queries



Revenue & Customer Segments (SQL)



Revenue by Gender

Male: 157,890 USD • Female: 75,191 USD — clear skew toward male-driven revenue.



Customer Segments

Loyal: 3,116 • Returning: 701 • New: 83 — majority categorized as Loyal.



Repeat Buyers & Subscriptions

Repeat buyers more likely non-subscribers (2,518 No vs. 958 Yes among repeat buyers).



Product & Rating Insights

Top Products by Rating

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

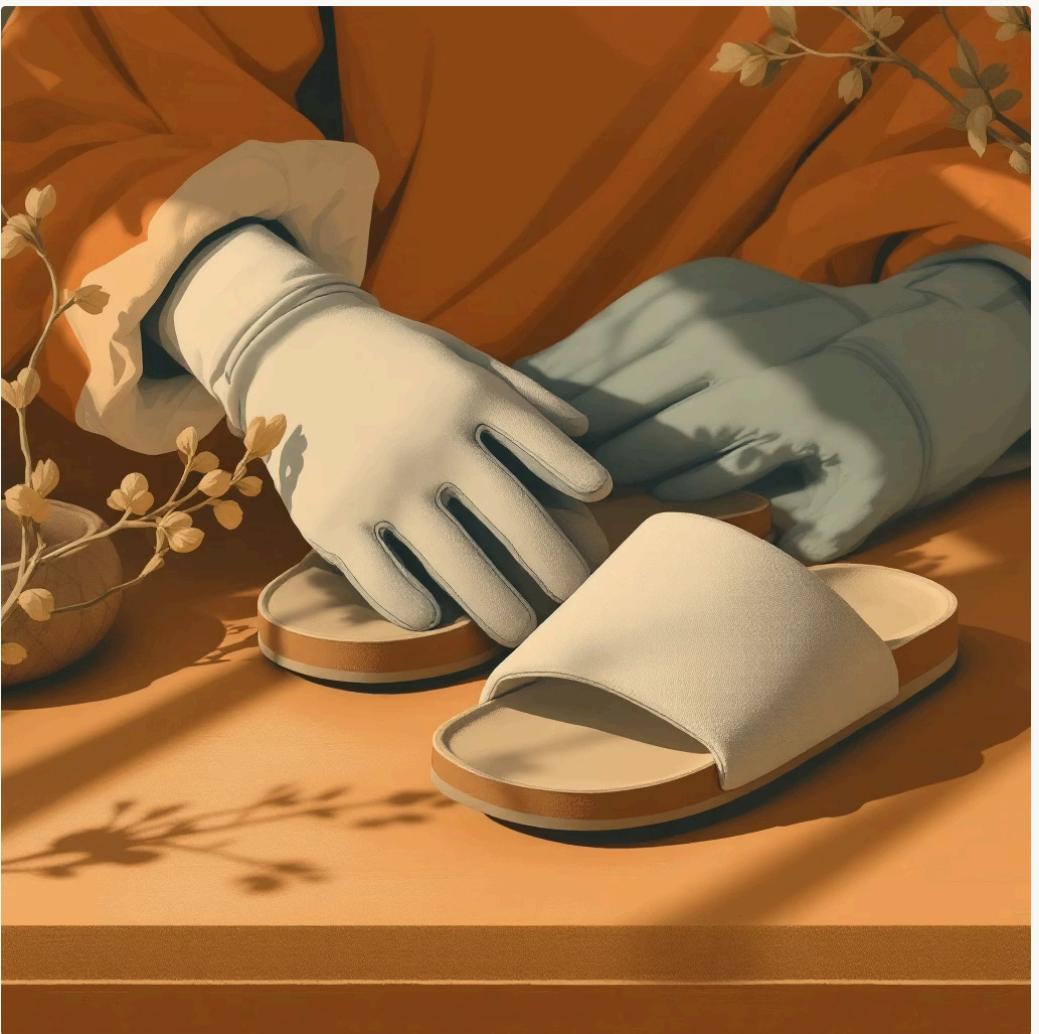
Top Items by Category

Accessories: Jewelry, Sunglasses, Belt.

Clothing: Blouse, Pants, Shirt.

Footwear: Sandals, Shoes, Sneakers.

Outerwear: Jacket, Coat.



Discounts & Price Behavior

839

High-Spending Discount
Users

839 rows identified where
customers used discounts yet
spent above average purchase
amount.

Top 5

Discount-Dependent
Products

Hat 50.00%, Sneakers 49.66%,
Coat 49.07%, Sweater 48.17%,
Pants 47.37% — these rely heavily
on discounts.



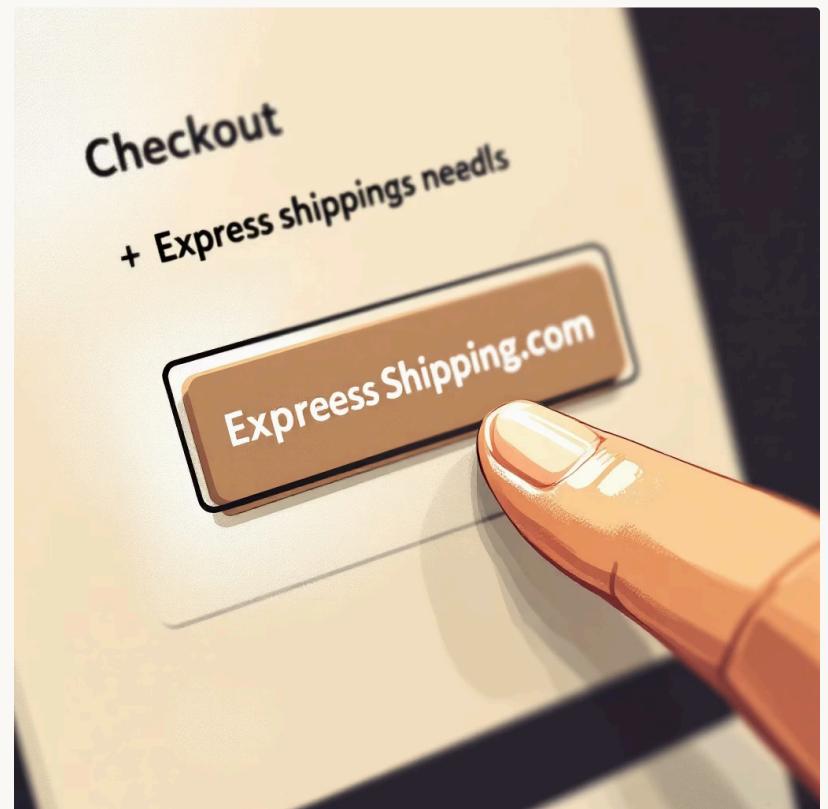
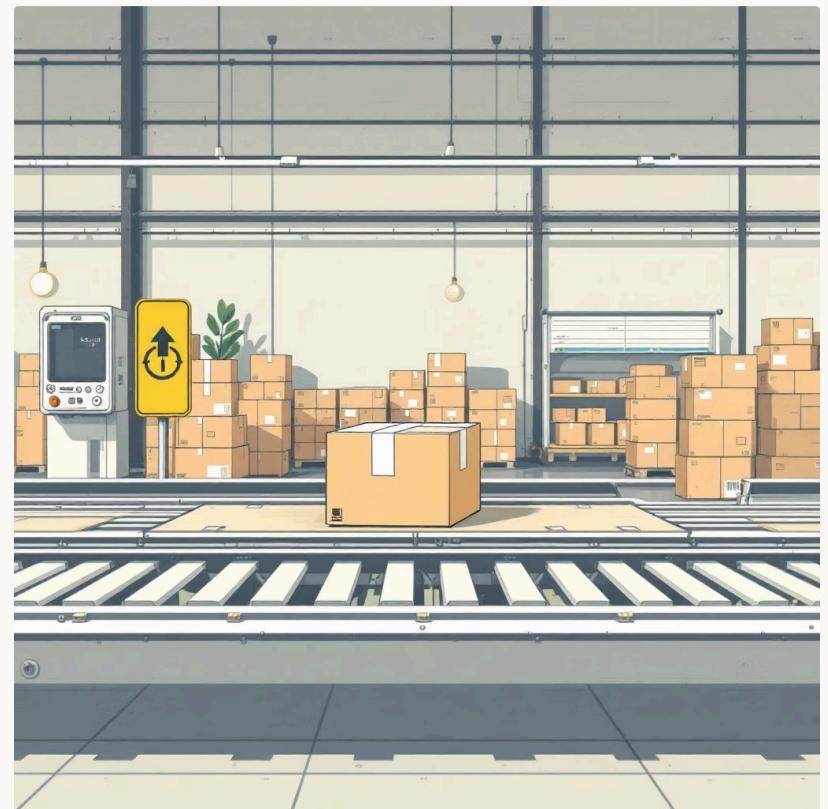
Behavioral & Shipping Patterns

Shipping Comparison

Average purchase: Standard = \$58.46 • Express = \$60.48 — express buyers spend slightly more.

Payment & Frequency

Popular payment method: PayPal. Frequent cadence: Every 3 Months for a notable group.



Revenue by Age & Subscription

Young...

Total Revenue: \$62,143

Middle-...

Total Revenue: \$59,197

Adult

Total Revenue: \$55,978

Senior

Total Revenue: \$55,763

Subscribers: 1,053 customers (avg spend \$59.49, total revenue \$62,645). Non-subscribers: 2,847 (avg \$59.87, total \$170,436).

Business Recommendations



Boost Subscriptions

Promote exclusive benefits to increase subscriber base and lifetime value.



Loyalty Programs

Reward repeat buyers to convert Returning customers into Loyal segment.



Review Discount Policy

Balance promotional lift with margin control; target discount-dependent products carefully.



Product Positioning & Targeting

Highlight top-rated and best-selling items; focus marketing on high-revenue age groups and express-shipping users.

